3.1 | Synthesis of Biological Macromolecules

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

- Understand macromolecule synthesis
- Explain dehydration (or condensation) and hydrolysis reactions

As you've learned, **biological macromolecules** are large molecules, necessary for life, that are built from smaller organic molecules. There are four major biological macromolecule classes (carbohydrates, lipids, proteins, and nucleic acids). Each is an important cell component and performs a wide array of functions. Combined, these molecules make up the majority of a cell's dry mass (recall that water makes up the majority of its complete mass). Biological macromolecules are organic, meaning they contain carbon. In addition, they may contain hydrogen, oxygen, nitrogen, and additional minor elements.

Dehydration Synthesis

Most macromolecules are made from single subunits, or building blocks, called **monomers**. The monomers combine with each other using covalent bonds to form larger molecules known as **polymers**. In doing so, monomers release water molecules as byproducts. This type of reaction is **dehydration synthesis**, which means "to put together while losing water."
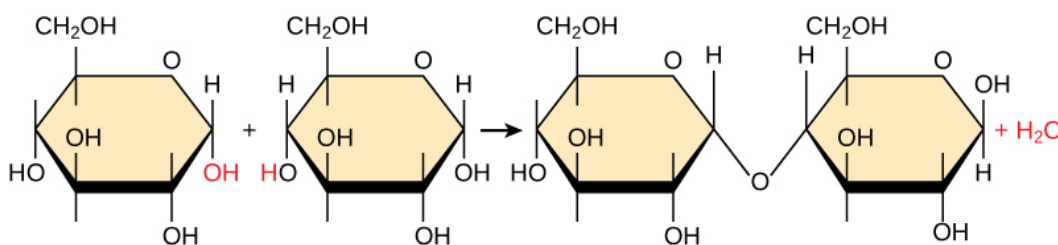

Figure 3.2 In the dehydration synthesis reaction above, two glucose molecules link to form the disaccharide maltose. In the process, it forms a water molecule.

In a dehydration synthesis reaction (**Figure 3.2**), the hydrogen of one monomer combines with the hydroxyl group of another monomer, releasing a water molecule. At the same time, the monomers share electrons and form covalent bonds. As additional monomers join, this chain of repeating monomers forms a polymer. Different monomer types can combine in many configurations, giving rise to a diverse group of macromolecules. Even one kind of monomer can combine in a variety of ways to form several different polymers. For example, glucose monomers are the constituents of starch, glycogen, and cellulose.

Hydrolysis

Polymers break down into monomers during hydrolysis. A chemical reaction occurs when inserting a water molecule across the bond. Breaking a covalent bond with this water molecule in the compound achieves this (**Figure 3.3**). During these reactions, the polymer breaks into two components: one part gains a hydrogen atom (H+) and the other gains a hydroxyl molecule (OH–) from a split water molecule.
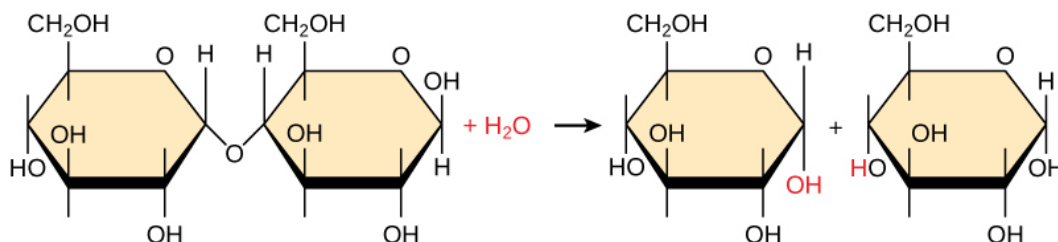

Figure 3.3 In the hydrolysis reaction here, the disaccharide maltose breaks down to form two glucose monomers by adding a water molecule. Note that this reaction is the reverse of the synthesis reaction in **Figure 3.2**.

Dehydration and **hydrolysis reactions** are catalyzed, or "sped up," by specific enzymes; dehydration reactions

involve the formation of new bonds, requiring energy, while hydrolysis reactions break bonds and release energy. These reactions are similar for most macromolecules, but each monomer and polymer reaction is specific for its class. For example, catalytic enzymes in the digestive system hydrolyze or break down the food we ingest into smaller molecules. This allows cells in our body to easily absorb nutrients in the intestine. A specific enzyme breaks down each macromolecule. For instance, amylase, sucrase, lactase, or maltase break down carbohydrates. Enzymes called proteases, such as pepsin and peptidase, and hydrochloric acid break down proteins. Lipases break down lipids. These broken down macromolecules provide energy for cellular activities.

LINK TO LEARNING

Visit **this site (http://openstaxcollege.org/l/hydrolysis)** to see visual representations of dehydration synthesis and hydrolysis.

3.2 | Carbohydrates

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

- Discuss the role of carbohydrates in cells and in the extracellular materials of animals and plants

- Explain carbohydrate classifications

- List common monosaccharides, disaccharides, and polysaccharides

Most people are familiar with carbohydrates, one type of macromolecule, especially when it comes to what we eat. To lose weight, some individuals adhere to "low-carb" diets. Athletes, in contrast, often "carb-load" before important competitions to ensure that they have enough energy to compete at a high level. Carbohydrates are, in fact, an essential part of our diet. Grains, fruits, and vegetables are all natural carbohydrate sources that provide energy to the body, particularly through glucose, a simple sugar that is a component of **starch** and an ingredient in many staple foods. Carbohydrates also have other important functions in humans, animals, and plants.

Molecular Structures

The stoichiometric formula $(CH_2O)_n$, where n is the number of carbons in the molecule represents **carbohydrates**. In other words, the ratio of carbon to hydrogen to oxygen is 1:2:1 in carbohydrate molecules. This formula also explains the origin of the term "carbohydrate": the components are carbon ("carbo") and the components of water (hence, "hydrate"). Scientists classify carbohydrates into three subtypes: monosaccharides, disaccharides, and polysaccharides.

Monosaccharides

Monosaccharides (mono- = "one"; sacchar- = "sweet") are simple sugars, the most common of which is glucose. In monosaccharides, the number of carbons usually ranges from three to seven. Most monosaccharide names end with the suffix -ose. If the sugar has an aldehyde group (the functional group with the structure R-CHO), it is an aldose, and if it has a ketone group (the functional group with the structure RC(=O)R'), it is a ketose. Depending on the number of carbons in the sugar, they can be trioses (three carbons), pentoses (five carbons), and/or hexoses (six carbons). **Figure 3.4** illustrates monosaccharides.
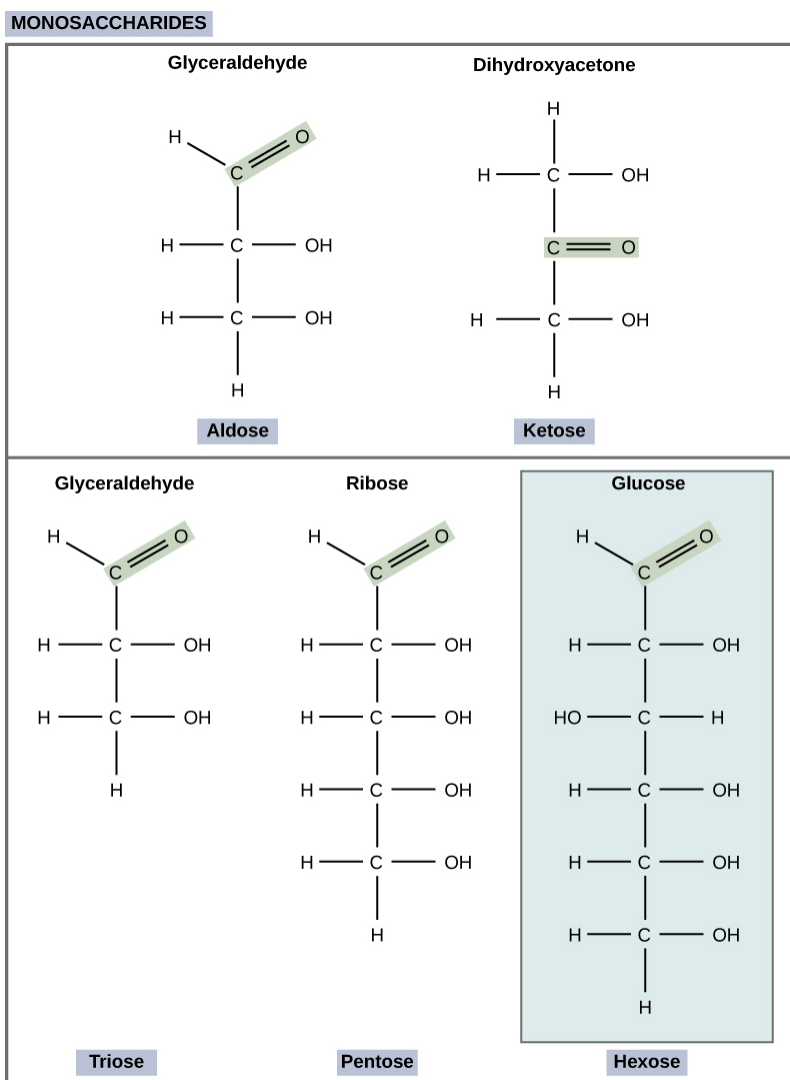
Figure 3.4 Scientists classify monosaccharides based on the position of their carbonyl group and the number of carbons in the backbone. Aldoses have a carbonyl group (indicated in green) at the end of the carbon chain, and ketoses have a carbonyl group in the middle of the carbon chain. Trioses, pentoses, and hexoses have three-, five-, and six- carbon backbones, respectively.

The chemical formula for glucose is $C_6H_{12}O_6$. In humans, glucose is an important source of energy. During cellular respiration, energy releases from glucose, and that energy helps make adenosine triphosphate (ATP). Plants synthesize glucose using carbon dioxide and water, and glucose in turn provides energy requirements for the plant. Humans and other animals that feed on plants often store excess glucose as catabolized (cell breakdown of larger molecules) starch.

Galactose (part of lactose, or milk sugar) and fructose (found in sucrose, in fruit) are other common monosaccharides. Although glucose, galactose, and fructose all have the same chemical formula ($C_6H_{12}O_6$), they differ structurally and chemically (and are isomers) because of the different arrangement of functional groups around the asymmetric carbon. All these monosaccharides have more than one asymmetric carbon (**Figure 3.5**).

visual CONNECTION

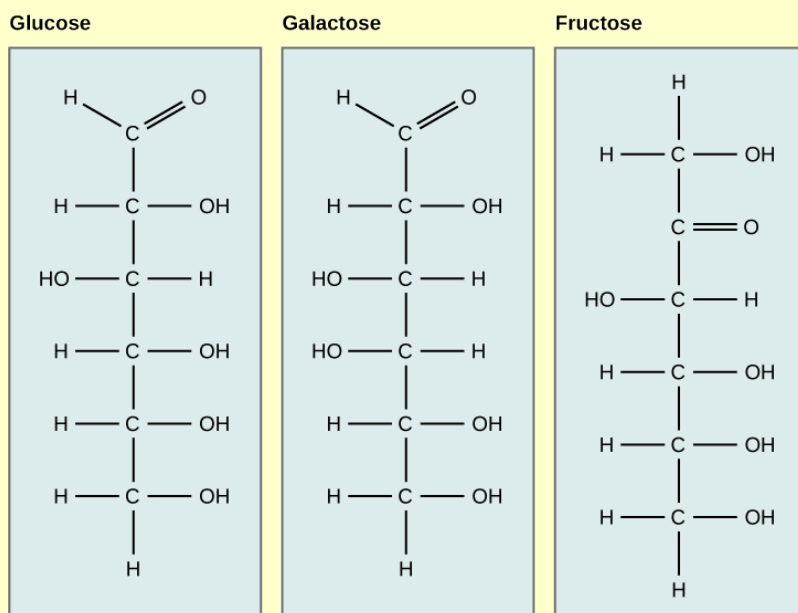**Glucose**   **Galactose**   **Fructose**

Figure 3.5 Glucose, galactose, and fructose are all hexoses. They are structural isomers, meaning they have the same chemical formula ($C_6H_{12}O_6$) but a different atom arrangement.

What kind of sugars are these, aldose or ketose?

Glucose, galactose, and fructose are isomeric monosaccharides (hexoses), meaning they have the same chemical formula but have slightly different structures. Glucose and galactose are aldoses, and fructose is a ketose.

Monosaccharides can exist as a linear chain or as ring-shaped molecules. In aqueous solutions they are usually in ring forms (**Figure 3.6**). Glucose in a ring form can have two different hydroxyl group arrangements (OH) around the anomeric carbon (carbon 1 that becomes asymmetric in the ring formation process). If the hydroxyl group is below carbon number 1 in the sugar, it is in the alpha (α) position, and if it is above the plane, it is in the beta (β) position.
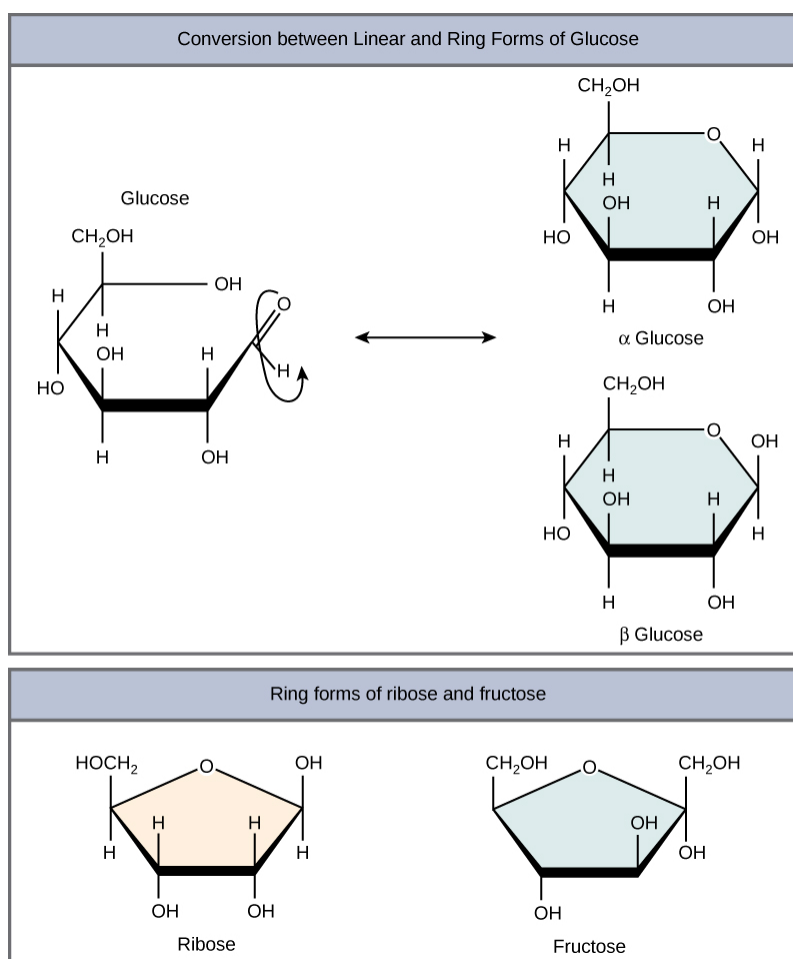
Figure 3.6 Five and six carbon monosaccharides exist in equilibrium between linear and ring forms. When the ring forms, the side chain it closes on locks into an *α* or *β* position. Fructose and ribose also form rings, although they form five-membered rings as opposed to the six-membered ring of glucose.

Disaccharides

Disaccharides (di- = "two") form when two monosaccharides undergo a dehydration reaction (or a condensation reaction or dehydration synthesis). During this process, one monosaccharide's hydroxyl group combines with another monosaccharide's hydrogen, releasing a water molecule and forming a covalent bond. A covalent bond forms between a carbohydrate molecule and another molecule (in this case, between two monosaccharides). Scientists call this a **glycosidic bond** (**Figure 3.7**). Glycosidic bonds (or glycosidic linkages) can be an alpha or beta type. An alpha bond is formed when the OH group on the carbon-1 of the first glucose is below the ring plane, and a beta bond is formed when the OH group on the carbon-1 is above the ring plane.
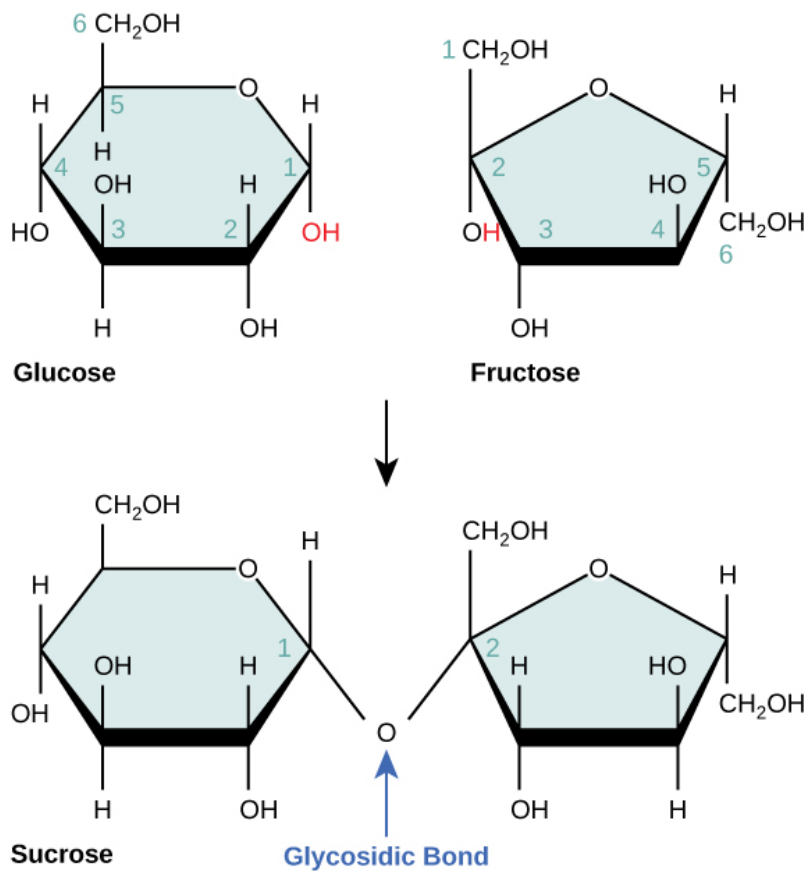
Figure 3.7 Sucrose forms when a glucose monomer and a fructose monomer join in a dehydration reaction to form a glycosidic bond. In the process, a water molecule is lost. By convention, the carbon atoms in a monosaccharide are numbered from the terminal carbon closest to the carbonyl group. In sucrose, a glycosidic linkage forms between carbon 1 in glucose and carbon 2 in fructose.

Common disaccharides include lactose, maltose, and sucrose (**Figure 3.8**). Lactose is a disaccharide consisting of the monomers glucose and galactose. It is naturally in milk. Maltose, or malt sugar, is a disaccharide formed by a dehydration reaction between two glucose molecules. The most common disaccharide is sucrose, or table sugar, which is comprised of glucose and fructose monomers.
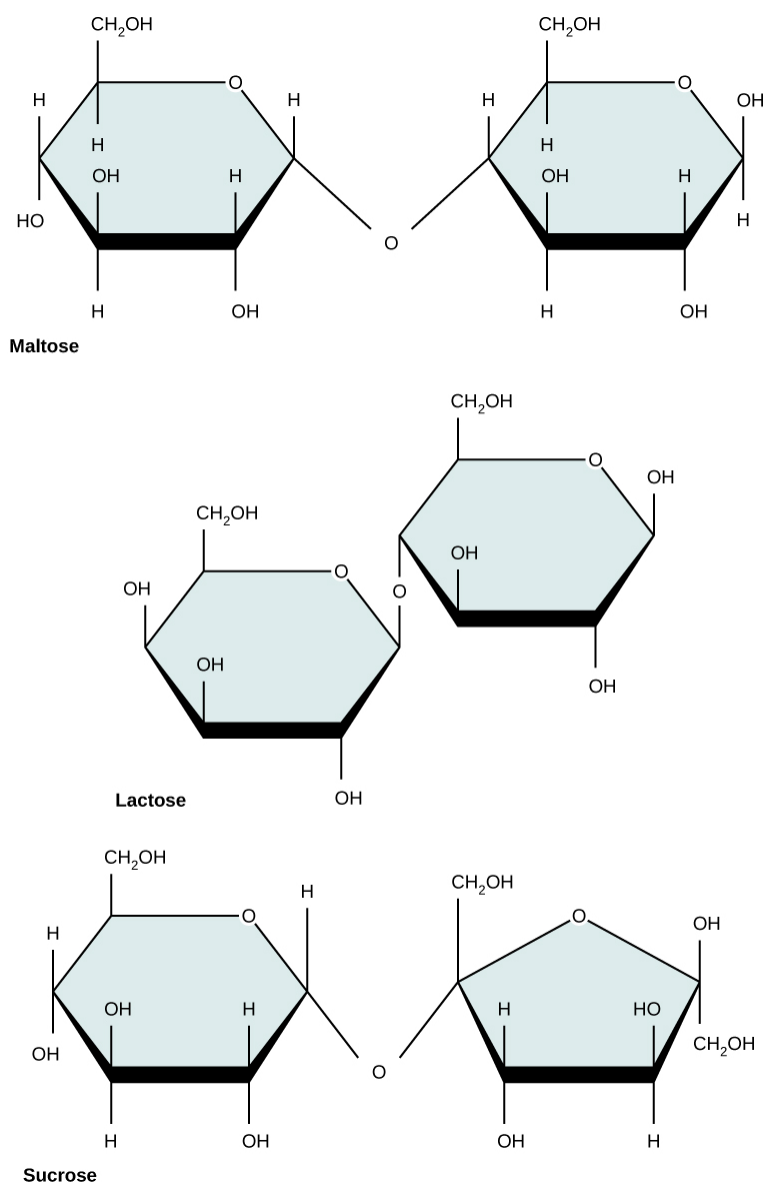
Figure 3.8 Common disaccharides include maltose (grain sugar), lactose (milk sugar), and sucrose (table sugar).

Polysaccharides

A long chain of monosaccharides linked by glycosidic bonds is a **polysaccharide** (poly- = "many"). The chain may be branched or unbranched, and it may contain different types of monosaccharides. The molecular weight may be 100,000 daltons or more depending on the number of joined monomers. Starch, glycogen, cellulose, and chitin are primary examples of polysaccharides.

Plants store starch in the form of sugars. In plants, an amylose and amylopectic mixture (both glucose polymers) comprise these sugars. Plants are able to synthesize glucose, and they store the excess glucose, beyond the their immediate energy needs, as starch in different plant parts, including roots and seeds. The starch in the seeds provides food for the embryo as it germinates and can also act as a food source for humans and animals. Enzymes break down the starch that humans consume. For example, an amylase present in saliva catalyzes, or breaks down this starch into smaller molecules, such as maltose and glucose. The cells can then absorb the glucose.

Glucose starch comprises monomers that are joined by α 1-4 or α 1-6 glycosidic bonds. The numbers 1-4 and 1-6 refer to the carbon number of the two residues that have joined to form the bond. As **Figure 3.9** illustrates, unbranched glucose monomer chains (only α 1-4 linkages) form the starch; whereas, amylopectin is a branched polysaccharide (α 1-6 linkages at the branch points).
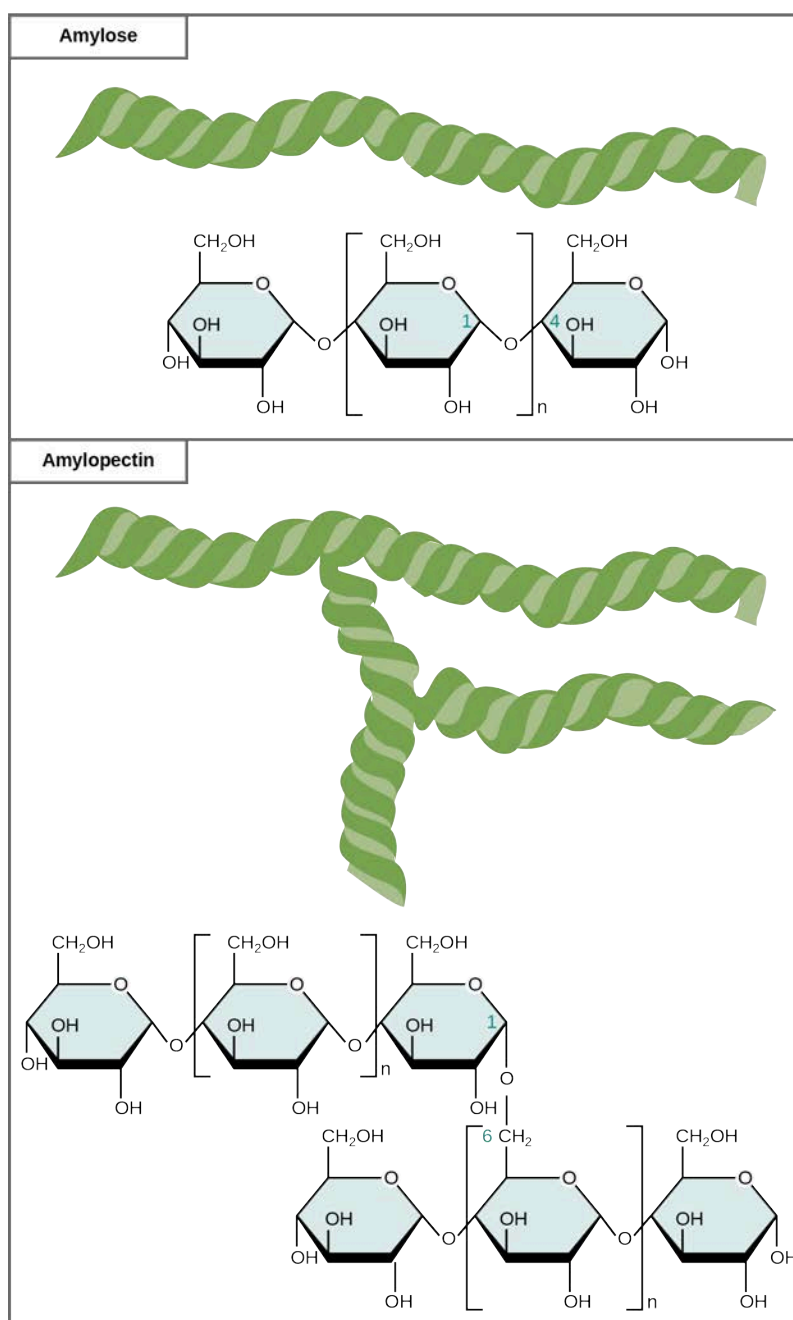
Figure 3.9 Amylose and amylopectin are two different starch forms. Unbranched glucose monomer chains comprise amylose by α 1-4 glycosidic linkages. Unbranched glucose monomer chains comprise amylopectin by α 1-4 and α 1-6 glycosidic linkages. Because of the way the subunits are joined, the glucose chains have a helical structure. Glycogen (not shown) is similar in structure to amylopectin but more highly branched.

Glycogen is the storage form of glucose in humans and other vertebrates and is comprised of monomers of glucose. Glycogen is the animal equivalent of starch and is a highly branched molecule usually stored in liver and muscle cells. Whenever blood glucose levels decrease, glycogen breaks down to release glucose in a process scientists call glycogenolysis.

Cellulose is the most abundant natural biopolymer. Cellulose mostly comprises a plant's cell wall. This provides the cell structural support. Wood and paper are mostly cellulosic in nature. Glucose monomers comprise cellulose that β 1-4 glycosidic bonds link (**Figure 3.10**).
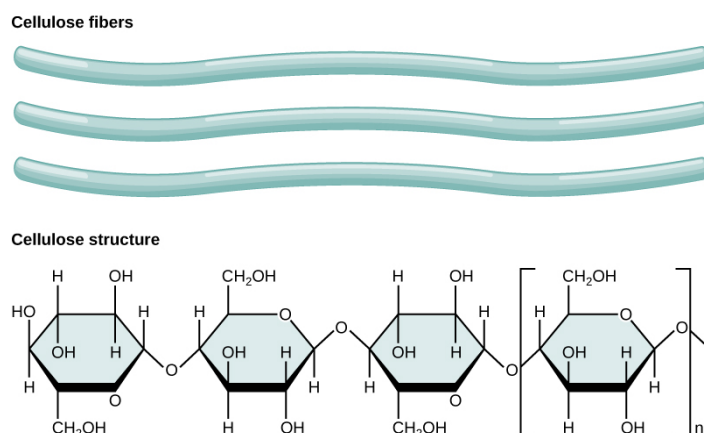
Cellulose fibers

Cellulose structure

Figure 3.10 In cellulose, glucose monomers are linked in unbranched chains by β 1-4 glycosidic linkages. Because of the way the glucose subunits are joined, every glucose monomer is flipped relative to the next one resulting in a linear, fibrous structure.

As **Figure 3.10** shows, every other glucose monomer in cellulose is flipped over, and the monomers are packed tightly as extended long chains. This gives cellulose its rigidity and high tensile strength—which is so important to plant cells. While human digestive enzymes cannot break down the β 1-4 linkage, herbivores such as cows, koalas, and buffalos are able, with the help of the specialized flora in their stomach, to digest plant material that is rich in cellulose and use it as a food source. In some of these animals, certain species of bacteria and protists reside in the rumen (part of the herbivore's digestive system) and secrete the enzyme cellulase. The appendix of grazing animals also contains bacteria that digest cellulose, giving it an important role in ruminants' digestive systems. Cellulases can break down cellulose into glucose monomers that animals use as an energy source. Termites are also able to break down cellulose because of the presence of other organisms in their bodies that secrete cellulases.

Carbohydrates serve various functions in different animals. Arthropods (insects, crustaceans, and others) have an outer skeleton, the exoskeleton, which protects their internal body parts (as we see in the bee in **Figure 3.11**). This exoskeleton is made of the biological macromolecule **chitin**, which is a polysaccharide-containing nitrogen. It is made of repeating N-acetyl-β-d-glucosamine units, which are a modified sugar. Chitin is also a major component of fungal cell walls. Fungi are neither animals nor plants and form a kingdom of their own in the domain Eukarya.

Figure 3.11 Insects have a hard outer exoskeleton made of chitin, a type of polysaccharide. (credit: Louise Docker)

CAREER CONNECTION

Registered Dietitian

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

To become a registered dietitian, one needs to earn at least a bachelor's degree in dietetics, nutrition, food technology, or a related field. In addition, registered dietitians must complete a supervised internship program and pass a national exam. Those who pursue careers in dietetics take courses in nutrition, chemistry, biochemistry, biology, microbiology, and human physiology. Dietitians must become experts in the chemistry and physiology (biological functions) of food (proteins, carbohydrates, and fats).

Benefits of Carbohydrates

Are carbohydrates good for you? Some often tell people who wish to lose weight that carbohydrates are bad and they should avoid them. Some diets completely forbid carbohydrate consumption, claiming that a low-carbohydrate diet helps people to lose weight faster. However, carbohydrates have been an important part of the human diet for thousands of years. Artifacts from ancient civilizations show the presence of wheat, rice, and corn in our ancestors' storage areas.

As part of a well balanced diet, we should supplement carbohydrates with proteins, vitamins, and fats. Calorie-wise, a gram of carbohydrate provides 4.3 Kcal. For comparison, fats provide 9 Kcal/g, a less desirable ratio. Carbohydrates contain soluble and insoluble elements. The insoluble part, fiber, is mostly cellulose. Fiber has many uses. It promotes regular bowel movement by adding bulk, and it regulates the blood glucose consumption rate. Fiber also helps to remove excess cholesterol from the body. Fiber binds to the cholesterol in the small intestine, then attaches to the cholesterol and prevents the cholesterol particles from entering the bloodstream. Cholesterol then exits the body via the feces. Fiber-rich diets also have a protective role in reducing the occurrence of colon cancer. In addition, a meal containing whole grains and vegetables gives a feeling of fullness. As an immediate source of energy, glucose breaks down during the cellular respiration process, which produces ATP, the cell's energy currency. Without consuming carbohydrates, we reduce the availability of "instant energy". Eliminating carbohydrates from the diet is not the best way to lose weight. A low-calorie diet that is rich in whole grains, fruits, vegetables, and lean meat, together with plenty of exercise and plenty of water, is the more sensible way to lose weight.

LINK TO LEARNING

For an additional perspective on carbohydrates, explore "Biomolecules: the Carbohydrates" through this **interactive animation (http://openstaxcollege.org/l/carbohydrates)** .

3.3 | Lipids

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

- Describe the four major types of lipids
- Explain the role of fats in storing energy
- Differentiate between saturated and unsaturated fatty acids
- Describe phospholipids and their role in cells
- Define the basic structure of a steroid and some steroid functions
- Explain how cholesterol helps maintain the plasma membrane's fluid nature

Lipids include a diverse group of compounds that are largely nonpolar in nature. This is because they are hydrocarbons that include mostly nonpolar carbon–carbon or carbon–hydrogen bonds. Non-polar molecules are hydrophobic ("water fearing"), or insoluble in water. Lipids perform many different functions in a cell. Cells store energy for long-term use in the form of fats. Lipids also provide insulation from the environment for plants and animals (**Figure 3.12**). For example, they help keep aquatic birds and mammals dry when forming a protective layer over fur or feathers because of their water-repellant hydrophobic nature. Lipids are also the building blocks of many hormones and are an important constituent of all cellular membranes. Lipids include fats, oils, waxes, phospholipids, and steroids.

Figure 3.12 Hydrophobic lipids in aquatic mammals' fur, such as this river otter, protect them from the elements. (credit: Ken Bosma)

Fats and Oils

A fat molecule consists of two main components—glycerol and fatty acids. Glycerol is an organic compound (alcohol) with three carbons, five hydrogens, and three hydroxyl (OH) groups. Fatty acids have a long chain of hydrocarbons to which a carboxyl group is attached, hence the name "fatty acid." The number of carbons in the fatty acid may range from 4 to 36. The most common are those containing 12–18 carbons. In a fat molecule, the fatty acids attach to each of the glycerol molecule's three carbons with an ester bond through an oxygen atom (**Figure 3.13**).
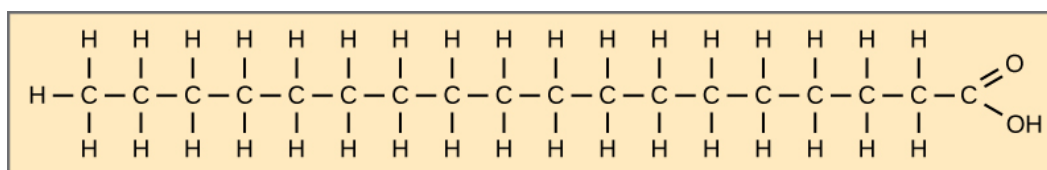
Glycerol

+

Fatty Acid

Triacylglycerol

Figure 3.13 Joining three fatty acids to a glycerol backbone in a dehydration reaction forms triacylglycerol. Three water molecules release in the process.

During this ester bond formation, three water molecules are released. The three fatty acids in the triacylglycerol may be similar or dissimilar. We also call fats **triacylglycerols** or **triglycerides** because of their chemical structure. Some fatty acids have common names that specify their origin. For example, palmitic acid, a **saturated fatty acid**, is derived from the palm tree. Arachidic acid is derived from *Arachis hypogea,* the scientific name for groundnuts or peanuts.

Fatty acids may be saturated or unsaturated. In a fatty acid chain, if there are only single bonds between neighboring carbons in the hydrocarbon chain, the fatty acid is saturated. Saturated fatty acids are saturated with hydrogen. In other words, the number of hydrogen atoms attached to the carbon skeleton is maximized. Stearic acid is an example of a saturated fatty acid (**Figure 3.14**).

Figure 3.14 Stearic acid is a common saturated fatty acid.

When the hydrocarbon chain contains a double bond, the fatty acid is **unsaturated**. Oleic acid is an example of an unsaturated fatty acid (**Figure 3.15**).
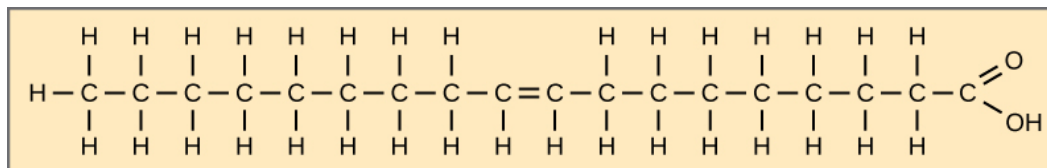

Figure 3.15 Oleic acid is a common unsaturated fatty acid.

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

When a fatty acid has no double bonds, it is a saturated fatty acid because it is not possible to add more hydrogen to the chain's carbon atoms. A fat may contain similar or different fatty acids attached to glycerol. Long straight fatty acids with single bonds generally pack tightly and are solid at room temperature. Animal fats with stearic acid and palmitic acid (common in meat) and the fat with butyric acid (common in butter) are examples of saturated fats. Mammals store fats in specialized cells, or adipocytes, where fat globules occupy most of the cell's volume. Plants store fat or oil in many seeds and use them as a source of energy during seedling development. Unsaturated fats or oils are usually of plant origin and contain *cis* unsaturated fatty acids. *Cis* and *trans* indicate the configuration of the molecule around the double bond. If hydrogens are present in the same plane, it is a cis fat. If the hydrogen atoms are on two different planes, it is a **trans fat**. The *cis* double bond causes a bend or a "kink" that prevents the fatty acids from packing tightly, keeping them liquid at room temperature (**Figure 3.16**). Olive oil, corn oil, canola oil, and cod liver oil are examples of unsaturated fats. Unsaturated fats help to lower blood cholesterol levels; whereas, saturated fats contribute to plaque formation in the arteries.
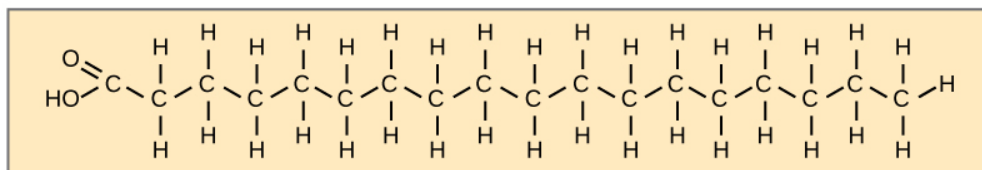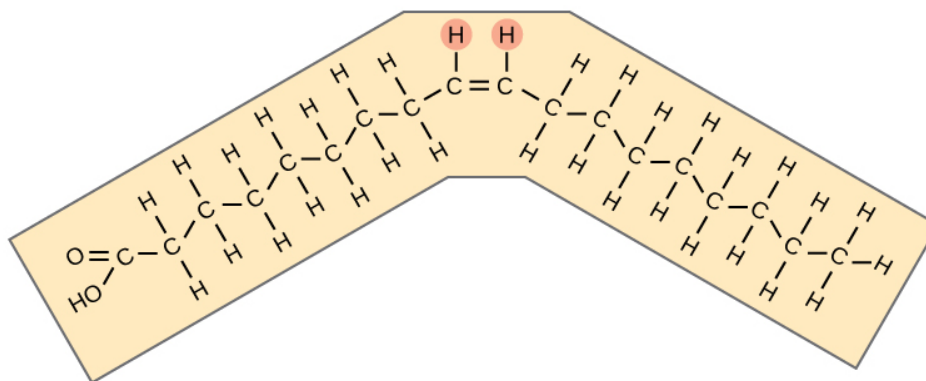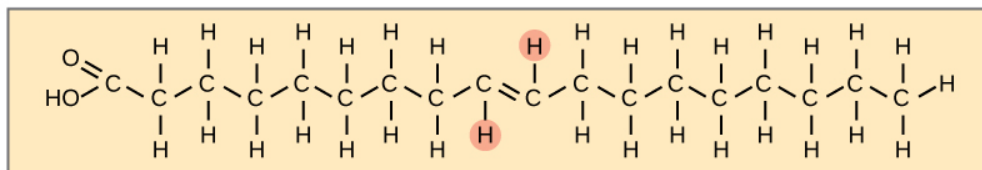
Saturated fatty acid

Stearic acid

Unsaturated fatty acids

Cis oleic acid

Trans oleic acid

Figure 3.16 Saturated fatty acids have hydrocarbon chains connected by single bonds only. Unsaturated fatty acids have one or more double bonds. Each double bond may be in a *cis* or *trans* configuration. In the *cis* configuration, both hydrogens are on the same side of the hydrocarbon chain. In the *trans* configuration, the hydrogens are on opposite sides. A *cis* double bond causes a kink in the chain.

Trans Fats

The food industry artificially hydrogenates oils to make them semi-solid and of a consistency desirable for many processed food products. Simply speaking, hydrogen gas is bubbled through oils to solidify them. During this hydrogenation process, double bonds of the *cis-* conformation in the hydrocarbon chain may convert to double bonds in the *trans-* conformation.

Margarine, some types of peanut butter, and shortening are examples of artificially hydrogenated trans fats. Recent studies have shown that an increase in trans fats in the human diet may lead to higher levels of low-density lipoproteins (LDL), or "bad" cholesterol, which in turn may lead to plaque deposition in the arteries, resulting in heart disease. Many fast food restaurants have recently banned using trans fats, and food labels are required to display the trans fat content.

Omega Fatty Acids

Essential fatty acids are those that the human body requires but does not synthesize. Consequently, they have to be supplemented through ingestion via the diet. **Omega**-3 fatty acids (like those in **Figure 3.17**) fall into this category and are one of only two known for humans (the other is omega-6 fatty acid). These are polyunsaturated fatty acids and are omega-3 because a double bond connects the third carbon from the hydrocarbon chain's end to its neighboring carbon.
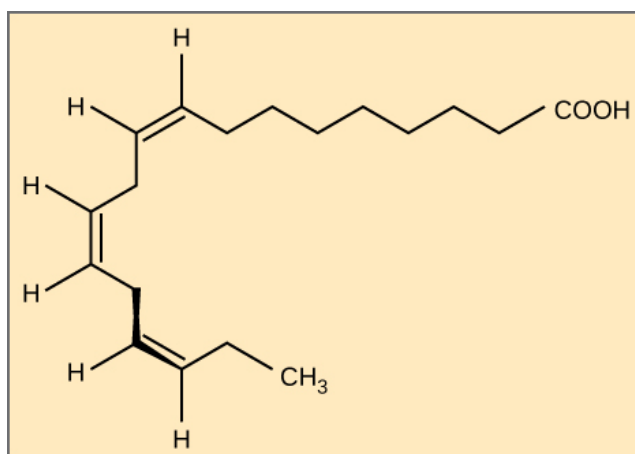
Figure 3.17 Alpha-linolenic acid is an example of an omega-3 fatty acid. It has three *cis* double bonds and, as a result, a curved shape. For clarity, the diagram does not show the carbons. Each singly bonded carbon has two hydrogens associated with it, which the diagram also does not show.

The farthest carbon away from the carboxyl group is numbered as the omega (ω) carbon, and if the double bond is between the third and fourth carbon from that end, it is an omega-3 fatty acid. Nutritionally important because the body does not make them, omega-3 fatty acids include alpha-linoleic acid (ALA), eicosapentaenoic acid (EPA), and docosahexaenoic acid (DHA), all of which are polyunsaturated. Salmon, trout, and tuna are good sources of omega-3 fatty acids. Research indicates that omega-3 fatty acids reduce the risk of sudden death from heart attacks, lower triglycerides in the blood, decrease blood pressure, and prevent thrombosis by inhibiting blood clotting. They also reduce inflammation, and may help lower the risk of some cancers in animals.

Like carbohydrates, fats have received considerable bad publicity. It is true that eating an excess of fried foods and other "fatty" foods leads to weight gain. However, fats do have important functions. Many vitamins are fat soluble, and fats serve as a long-term storage form of fatty acids: a source of energy. They also provide insulation for the body. Therefore, we should consume "healthy" fats in moderate amounts on a regular basis.

Waxes

Wax covers some aquatic birds' feathers and some plants' leaf surfaces. Because of waxes' hydrophobic nature, they prevent water from sticking on the surface (**Figure 3.18**). Long fatty acid chains esterified to long-chain alcohols comprise waxes.

Figure 3.18 Lipids comprise waxy coverings on some leaves. (credit: Roger Griffith)

Phospholipids

Phospholipids are major plasma membrane constituents that comprise cells' outermost layer. Like fats, they are comprised of fatty acid chains attached to a glycerol or sphingosine backbone. However, instead of three fatty acids attached as in triglycerides, there are two fatty acids forming diacylglycerol, and a modified phosphate

group occupies the glycerol backbone's third carbon (**Figure 3.19**). A phosphate group alone attached to a diaglycerol does not qualify as a phospholipid. It is phosphatidate (diacylglycerol 3-phosphate), the precursor of phospholipids. An alcohol modifies the phosphate group. Phosphatidylcholine and phosphatidylserine are two important phospholipids that are in plasma membranes.
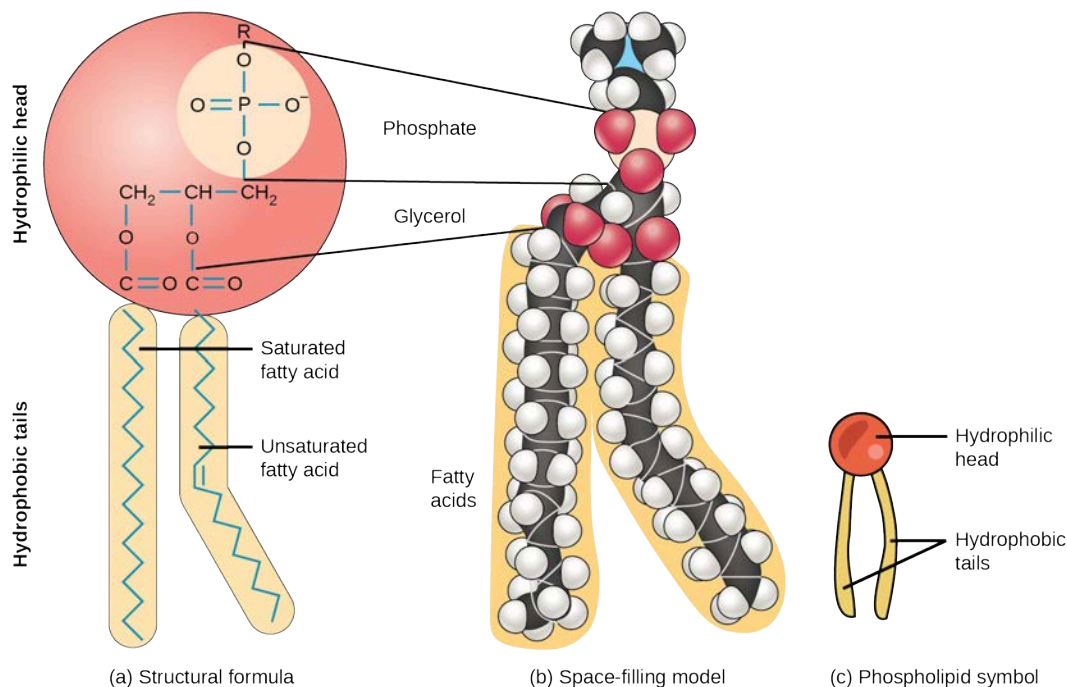

(a) Structural formula   (b) Space-filling model   (c) Phospholipid symbol

Figure 3.19 A phospholipid is a molecule with two fatty acids and a modified phosphate group attached to a glycerol backbone. Adding a charged or polar chemical group may modify the phosphate.

A phospholipid is an amphipathic molecule, meaning it has a hydrophobic and a hydrophilic part. The fatty acid chains are hydrophobic and cannot interact with water; whereas, the phosphate-containing group is hydrophilic and interacts with water (**Figure 3.20**).
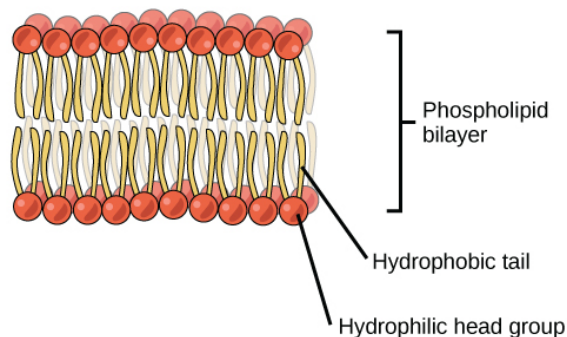

Figure 3.20 The phospholipid bilayer is the major component of all cellular membranes. The hydrophilic head groups of the phospholipids face the aqueous solution. The hydrophobic tails are sequestered in the middle of the bilayer.

The head is the hydrophilic part, and the tail contains the hydrophobic fatty acids. In a membrane, a bilayer of phospholipids forms the structure's matrix, phospholipids' fatty acid tails face inside, away from water; whereas, the phosphate group faces the outside, aqueous side (**Figure 3.20**).

Phospholipids are responsible for the plasma membrane's dynamic nature. If a drop of phospholipids is placed in water, it spontaneously forms a structure that scientists call a micelle, where the hydrophilic phosphate heads face the outside and the fatty acids face the structure's interior.
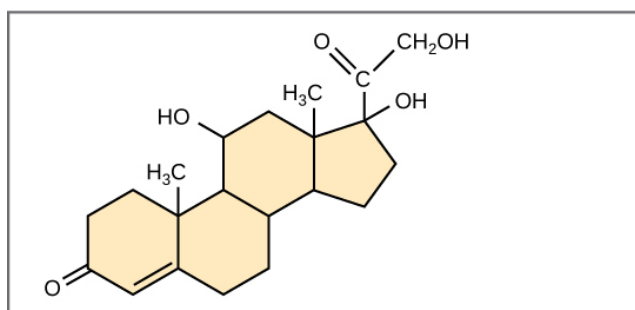
Steroids

Unlike the phospholipids and fats that we discussed earlier, **steroids** have a fused ring structure. Although they do not resemble the other lipids, scientists group them with them because they are also hydrophobic and

insoluble in water. All steroids have four linked carbon rings and several of them, like cholesterol, have a short tail (**Figure 3.21**). Many steroids also have the –OH functional group, which puts them in the alcohol classification (sterols).
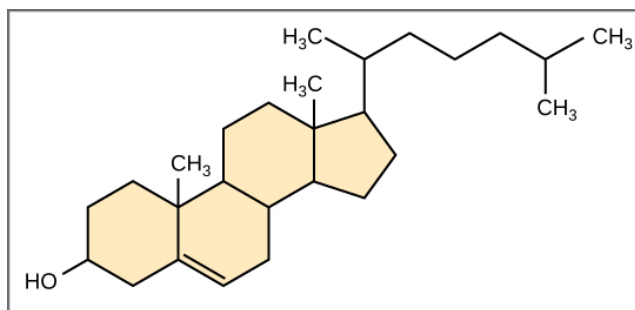
Cholesterol

Cortisol

Figure 3.21 Four fused hydrocarbon rings comprise steroids such as cholesterol and cortisol.

Cholesterol is the most common steroid. The liver synthesizes cholesterol and is the precursor to many steroid hormones such as testosterone and estradiol, which gonads and endocrine glands secrete. It is also the precursor to Vitamin D. Cholesterol is also the precursor of bile salts, which help emulsifying fats and their subsequent absorption by cells. Although lay people often speak negatively about cholesterol, it is necessary for the body's proper functioning. Sterols (cholesterol in animal cells, phytosterol in plants) are components of the plasma membrane of cells and are found within the phospholipid bilayer.

LINK TO LEARNING

For an additional perspective on lipids, explore the interactive animation **"Biomolecules: The Lipids" (http://openstaxcollege.org/l/lipids)** .

3.4 | Proteins

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

- Describe the functions proteins perform in the cell and in tissues
- Discuss the relationship between amino acids and proteins
- Explain the four levels of protein organization
- Describe the ways in which protein shape and function are linked

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

Types and Functions of Proteins

Enzymes, which living cells produce, are catalysts in biochemical reactions (like digestion) and are usually complex or conjugated proteins. Each enzyme is specific for the substrate (a reactant that binds to an enzyme) upon which it acts. The enzyme may help in breakdown, rearrangement, or synthesis reactions. We call enzymes that break down their substrates catabolic enzymes. Those that build more complex molecules from their substrates are anabolic enzymes, and enzymes that affect the rate of reaction are catalytic enzymes. Note that all enzymes increase the reaction rate and, therefore, are organic catalysts. An example of an enzyme is salivary amylase, which hydrolyzes its substrate amylose, a component of starch.

Hormones are chemical-signaling molecules, usually small proteins or steroids, secreted by endocrine cells that act to control or regulate specific physiological processes, including growth, development, metabolism, and reproduction. For example, insulin is a protein hormone that helps regulate the blood glucose level. **Table 3.1** lists the primary types and functions of proteins.

Protein Types and Functions

Type	Examples	Functions
Digestive Enzymes	Amylase, lipase, pepsin, trypsin	Help in food by catabolizing nutrients into monomeric units
Transport	Hemoglobin, albumin	Carry substances in the blood or lymph throughout the body
Structural	Actin, tubulin, keratin	Construct different structures, like the cytoskeleton
Hormones	Insulin, thyroxine	Coordinate different body systems' activity
Defense	Immunoglobulins	Protect the body from foreign pathogens
Contractile	Actin, myosin	Effect muscle contraction
Storage	Legume storage proteins, egg white (albumin)	Provide nourishment in early embryo development and the seedling

Table 3.1

Proteins have different shapes and molecular weights. Some proteins are globular in shape; whereas, others are fibrous in nature. For example, hemoglobin is a globular protein, but collagen, located in our skin, is a fibrous protein. Protein shape is critical to its function, and many different types of chemical bonds maintain this shape. Changes in temperature, pH, and exposure to chemicals may lead to permanent changes in the protein's shape, leading to loss of function, or **denaturation**. Different arrangements of the same 20 types of amino acids comprise all proteins. Two rare new amino acids were discovered recently (selenocystein and pirrolysine), and additional new discoveries may be added to the list.

Amino Acids

Amino acids are the monomers that comprise proteins. Each amino acid has the same fundamental structure, which consists of a central carbon atom, or the alpha (α) carbon, bonded to an amino group (NH_2), a carboxyl group (COOH), and to a hydrogen atom. Every amino acid also has another atom or group of atoms bonded to the central atom known as the R group (**Figure 3.22**).
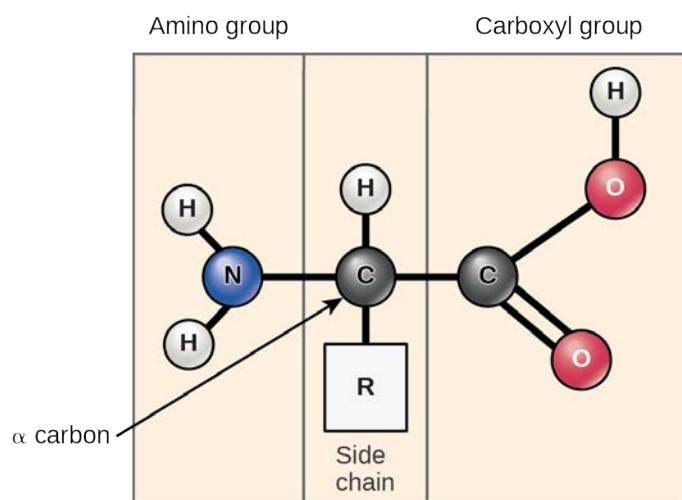
Figure 3.22 Amino acids have a central asymmetric carbon to which an amino group, a carboxyl group, a hydrogen atom, and a side chain (R group) are attached.

Scientists use the name "amino acid" because these acids contain both amino group and carboxyl-acid-group in their basic structure. As we mentioned, there are 20 common amino acids present in proteins. Nine of these are essential amino acids in humans because the human body cannot produce them and we obtain them from our diet. For each amino acid, the R group (or side chain) is different (**Figure 3.23**).
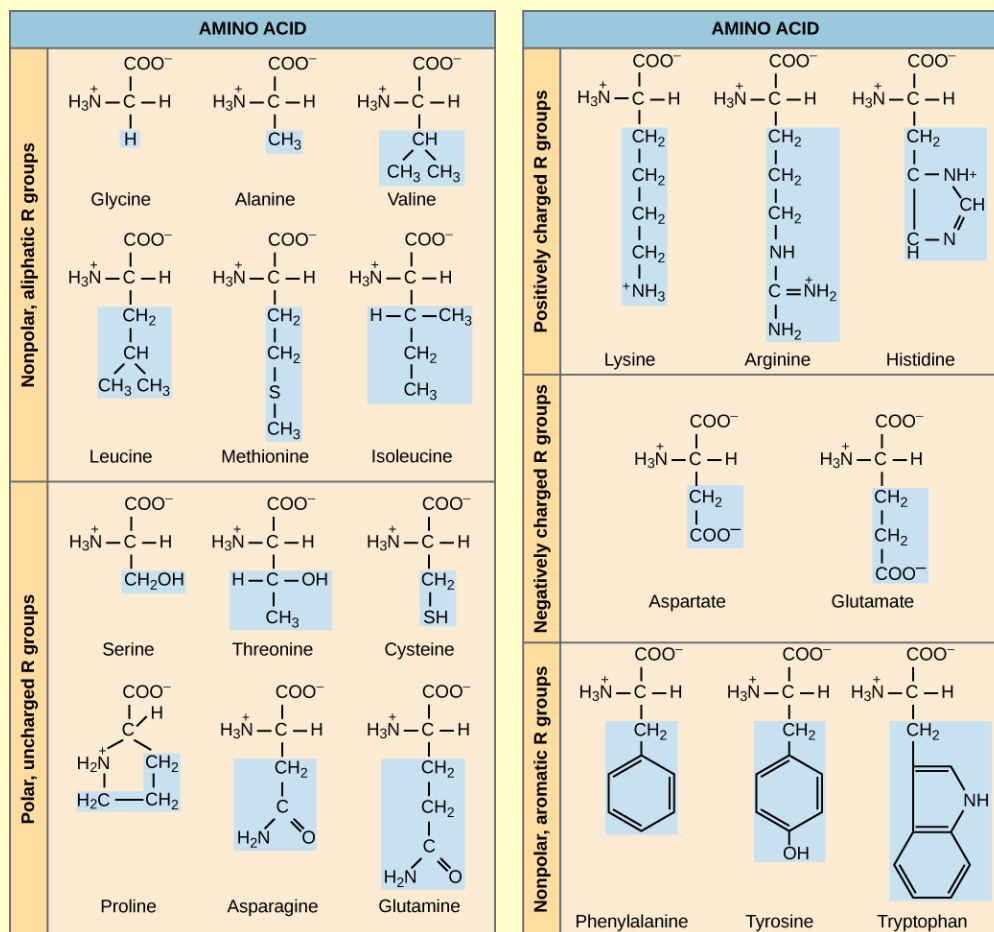
Figure 3.23 There are 20 common amino acids commonly found in proteins, each with a different R group (variant group) that determines its chemical nature.

Which categories of amino acid would you expect to find on a soluble protein's surface and which would you expect to find in the interior? What distribution of amino acids would you expect to find in a protein embedded in a lipid bilayer?

The chemical nature of the side chain determines the amino acid's nature (that is, whether it is acidic, basic, polar, or nonpolar). For example, the amino acid glycine has a hydrogen atom as the R group. Amino acids such as valine, methionine, and alanine are nonpolar or hydrophobic in nature, while amino acids such as serine, threonine, and cysteine are polar and have hydrophilic side chains. The side chains of lysine and arginine are positively charged, and therefore these amino acids are also basic amino acids. Proline has an R group that is linked to the amino group, forming a ring-like structure. Proline is an exception to the amino acid's standard structure since its amino group is not separate from the side chain (**Figure 3.23**).

A single upper case letter or a three-letter abbreviation represents amino acids. For example, the letter V or the three-letter symbol val represent valine. Just as some fatty acids are essential to a diet, some amino acids also are necessary. These essential amino acids in humans include isoleucine, leucine, and cysteine. Essential amino acids refer to those necessary to build proteins in the body, but not those that the body produces. Which amino acids are essential varies from organism to organism.

The sequence and the number of amino acids ultimately determine the protein's shape, size, and function. A covalent bond, or **peptide bond**, attaches to each amino acid, which a dehydration reaction forms. One amino acid's carboxyl group and the incoming amino acid's amino group combine, releasing a water molecule. The resulting bond is the peptide bond (**Figure 3.24**).
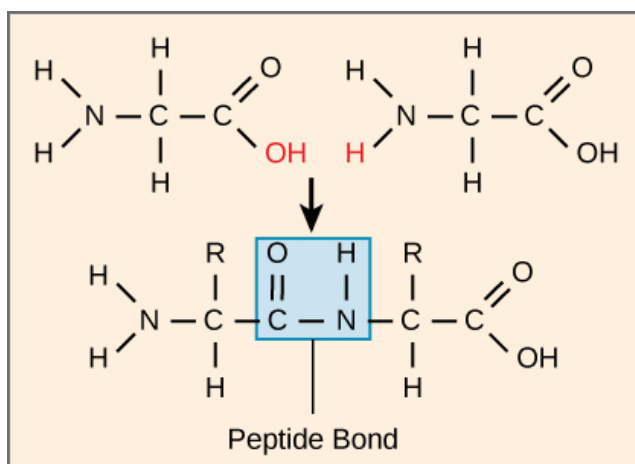
Figure 3.24 Peptide bond formation is a dehydration synthesis reaction. The carboxyl group of one amino acid is linked to the incoming amino acid's amino group. In the process, it releases a water molecule.

The products that such linkages form are peptides. As more amino acids join to this growing chain, the resulting chain is a polypeptide. Each polypeptide has a free amino group at one end. This end the N terminal, or the amino terminal, and the other end has a free carboxyl group, also the C or carboxyl terminal. While the terms polypeptide and protein are sometimes used interchangeably, a polypeptide is technically a polymer of amino acids, whereas the term protein is used for a polypeptide or polypeptides that have combined together, often have bound non-peptide prosthetic groups, have a distinct shape, and have a unique function. After protein synthesis (translation), most proteins are modified. These are known as post-translational modifications. They may undergo cleavage, phosphorylation, or may require adding other chemical groups. Only after these modifications is the protein completely functional.

LINK TO LEARNING

Click through the steps of protein synthesis in this **interactive tutorial (http://openstaxcollege.org/l/ protein_synth)** .

evolution CONNECTION

The Evolutionary Significance of Cytochrome c

Cytochrome c is an important component of the electron transport chain, a part of cellular respiration, and it is normally located in the cellular organelle, the mitochondrion. This protein has a heme prosthetic group, and the heme's central ion alternately reduces and oxidizes during electron transfer. Because this essential protein's role in producing cellular energy is crucial, it has changed very little over millions of years. Protein sequencing has shown that there is a considerable amount of cytochrome c amino acid sequence homology among different species. In other words, we can assess evolutionary kinship by measuring the similarities or differences among various species' DNA or protein sequences.

Scientists have determined that human cytochrome c contains 104 amino acids. For each cytochrome c molecule from different organisms that scientists have sequenced to date, 37 of these amino acids appear in the same position in all cytochrome c samples. This indicates that there may have been a common ancestor. On comparing the human and chimpanzee protein sequences, scientists did not find a sequence difference. When researchers compared human and rhesus monkey sequences, the single difference was in one amino acid. In another comparison, human to yeast sequencing shows a difference in the 44th position.

Protein Structure

As we discussed earlier, a protein's shape is critical to its function. For example, an enzyme can bind to a specific substrate at an active site. If this active site is altered because of local changes or changes in overall protein structure, the enzyme may be unable to bind to the substrate. To understand how the protein gets its final shape or conformation, we need to understand the four levels of protein structure: primary, secondary, tertiary, and quaternary.

Primary Structure

Amino acids' unique sequence in a polypeptide chain is its **primary structure**. For example, the pancreatic hormone insulin has two polypeptide chains, A and B, and they are linked together by disulfide bonds. The N terminal amino acid of the A chain is glycine; whereas, the C terminal amino acid is asparagine (**Figure 3.25**). The amino acid sequences in the A and B chains are unique to insulin.
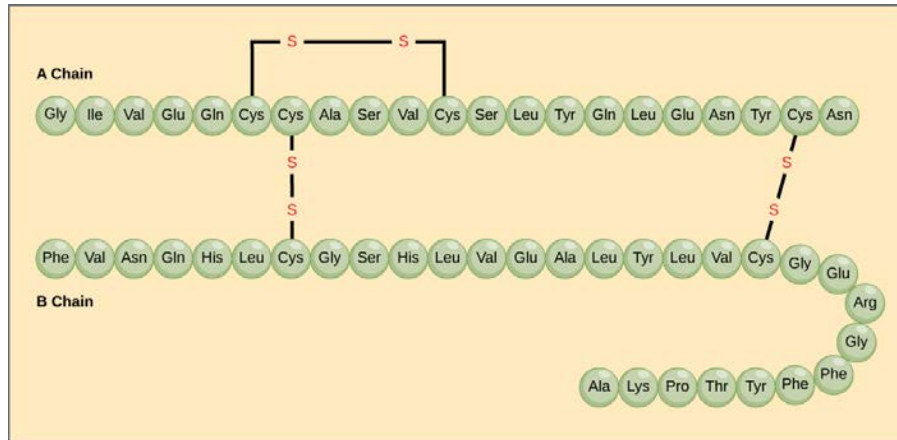

Figure 3.25 Bovine serum insulin is a protein hormone comprised of two peptide chains, A (21 amino acids long) and B (30 amino acids long). In each chain, three-letter abbreviations that represent the amino acids' names in the order they are present indicate primary structure. The amino acid cysteine (cys) has a sulfhydryl (SH) group as a side chain. Two sulfhydryl groups can react in the presence of oxygen to form a disulfide (S-S) bond. Two disulfide bonds connect the A and B chains together, and a third helps the A chain fold into the correct shape. Note that all disulfide bonds are the same length, but we have drawn them different sizes for clarity.

The gene encoding the protein ultimately determines the unique sequence for every protein. A change in nucleotide sequence of the gene's coding region may lead to adding a different amino acid to the growing polypeptide chain, causing a change in protein structure and function. In sickle cell anemia, the hemoglobin β chain (a small portion of which we show in **Figure 3.26**) has a single amino acid substitution, causing a change in protein structure and function. Specifically, valine in the β chain substitutes the amino acid glutamic. What is most remarkable to consider is that a hemoglobin molecule is comprised of two alpha and two beta chains that each consist of about 150 amino acids. The molecule, therefore, has about 600 amino acids. The structural difference between a normal hemoglobin molecule and a sickle cell molecule—which dramatically decreases life expectancy—is a single amino acid of the 600. What is even more remarkable is that three nucleotides each encode those 600 amino acids, and a single base change (point mutation), 1 in 1800 bases causes the mutation.
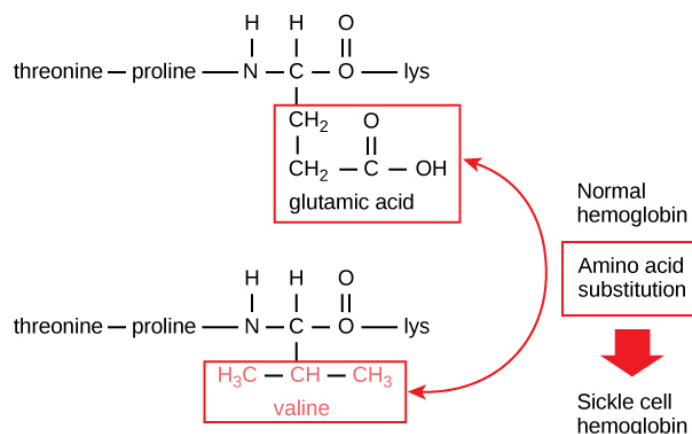
Figure 3.26 The beta chain of hemoglobin is 147 residues in length, yet a single amino acid substitution leads to sickle cell anemia. In normal hemoglobin, the amino acid at position seven is glutamate. In sickle cell hemoglobin, a valine replaces glutamate.

Because of this change of one amino acid in the chain, hemoglobin molecules form long fibers that distort the biconcave, or disc-shaped, red blood cells and causes them to assume a crescent or "sickle" shape, which clogs blood vessels (**Figure 3.27**). This can lead to myriad serious health problems such as breathlessness, dizziness, headaches, and abdominal pain for those affected by this disease.
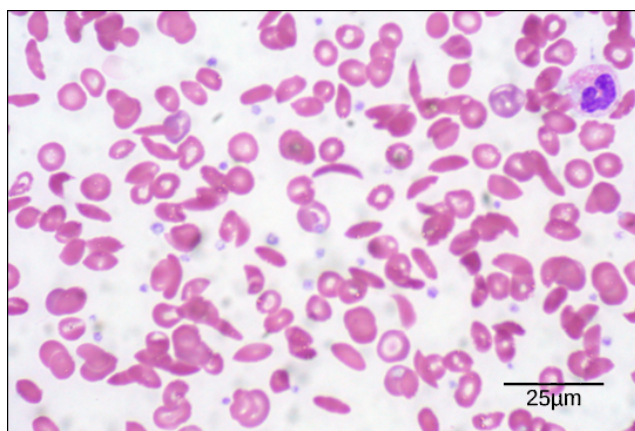

Figure 3.27 In this blood smear, visualized at 535x magnification using bright field microscopy, sickle cells are crescent shaped, while normal cells are disc-shaped. (credit: modification of work by Ed Uthman; scale-bar data from Matt Russell)

Secondary Structure

The local folding of the polypeptide in some regions gives rise to the **secondary structure** of the protein. The most common are the **α-helix** and **β-pleated sheet** structures (**Figure 3.28**). Both structures are held in shape by hydrogen bonds. The hydrogen bonds form between the oxygen atom in the carbonyl group in one amino acid and another amino acid that is four amino acids farther along the chain.
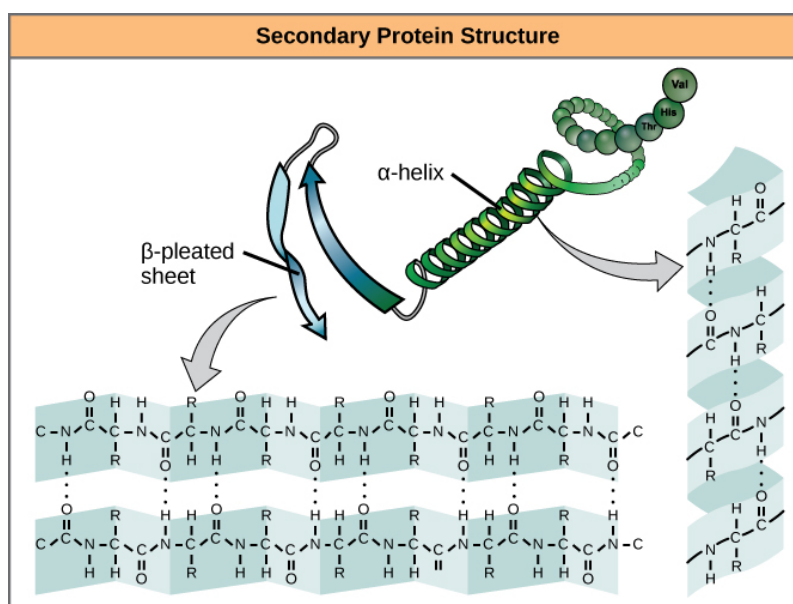
Figure 3.28 The α-helix and β-pleated sheet are secondary structures of proteins that form because of hydrogen bonding between carbonyl and amino groups in the peptide backbone. Certain amino acids have a propensity to form an α-helix, while others have a propensity to form a β-pleated sheet.

Every helical turn in an alpha helix has 3.6 amino acid residues. The polypeptide's R groups (the variant groups) protrude out from the α-helix chain. In the β-pleated sheet, hydrogen bonding between atoms on the polypeptide chain's backbone form the "pleats". The R groups are attached to the carbons and extend above and below the pleat's folds. The pleated segments align parallel or antiparallel to each other, and hydrogen bonds form between the partially positive nitrogen atom in the amino group and the partially negative oxygen atom in the peptide backbone's carbonyl group. The α-helix and β-pleated sheet structures are in most globular and fibrous proteins and they play an important structural role.

Tertiary Structure

The polypeptide's unique three-dimensional structure is its **tertiary structure** (**Figure 3.29**). This structure is in part due to chemical interactions at work on the polypeptide chain. Primarily, the interactions among R groups create the protein's complex three-dimensional tertiary structure. The nature of the R groups in the amino acids involved can counteract forming the hydrogen bonds we described for standard secondary structures. For example, R groups with like charges repel each other and those with unlike charges are attracted to each other (ionic bonds). When protein folding takes place, the nonpolar amino acids' hydrophobic R groups lie in the protein's interior; whereas, the hydrophilic R groups lie on the outside. Scientists also call the former interaction types hydrophobic interactions. Interaction between cysteine side chains forms disulfide linkages in the presence of oxygen, the only covalent bond that forms during protein folding.
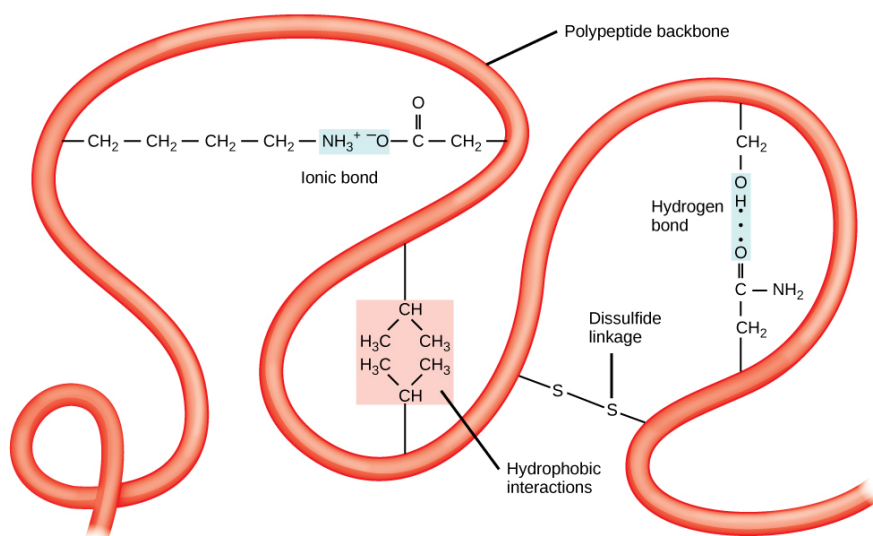
Figure 3.29 A variety of chemical interactions determine the proteins' tertiary structure. These include hydrophobic interactions, ionic bonding, hydrogen bonding, and disulfide linkages.

All of these interactions, weak and strong, determine the protein's final three-dimensional shape. When a protein loses its three-dimensional shape, it may no longer be functional.

Quaternary Structure

In nature, some proteins form from several polypeptides, or subunits, and the interaction of these subunits forms the **quaternary structure**. Weak interactions between the subunits help to stabilize the overall structure. For example, insulin (a globular protein) has a combination of hydrogen and disulfide bonds that cause it to mostly clump into a ball shape. Insulin starts out as a single polypeptide and loses some internal sequences in the presence of post-translational modification after forming the disulfide linkages that hold the remaining chains together. Silk (a fibrous protein), however, has a β-pleated sheet structure that is the result of hydrogen bonding between different chains.

Figure 3.30 illustrates the four levels of protein structure (primary, secondary, tertiary, and quaternary).
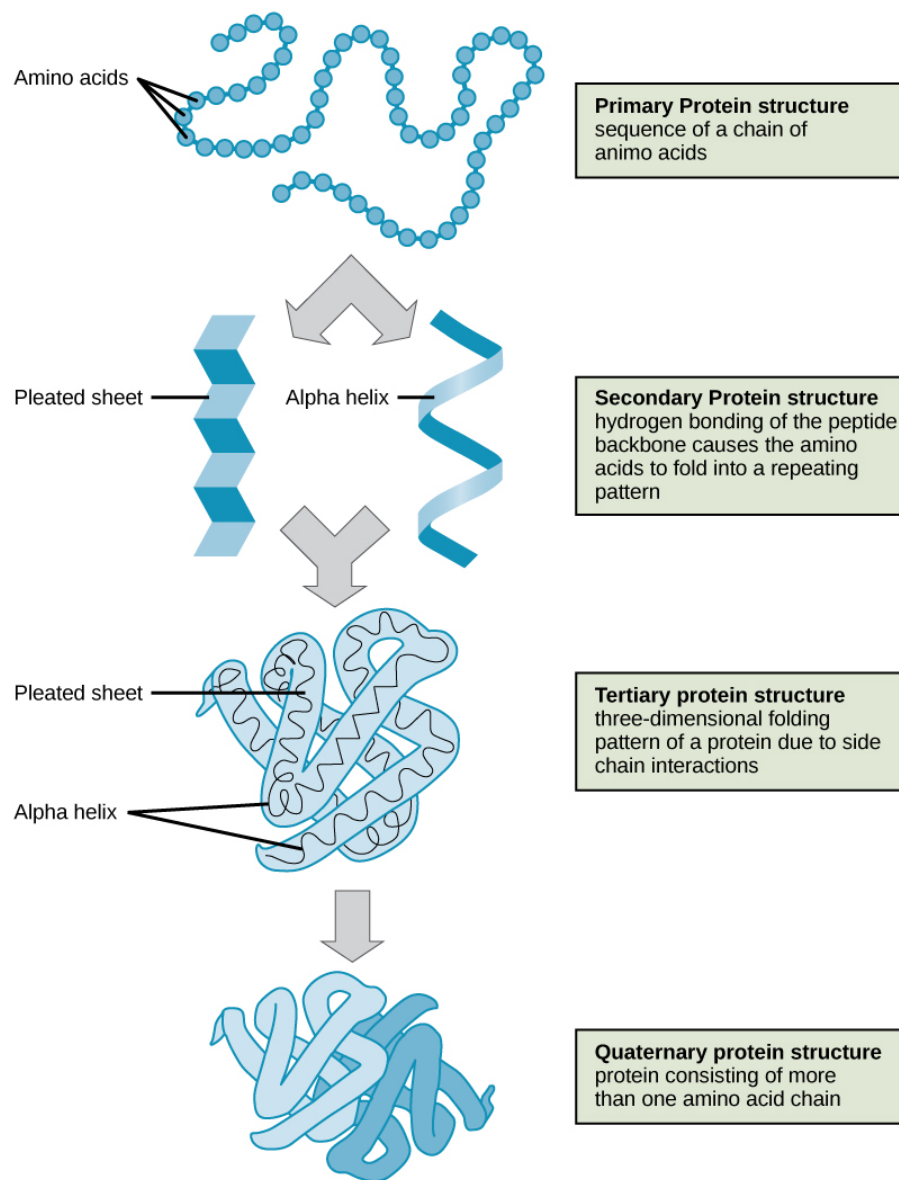
Figure 3.30 Observe the four levels of protein structure in these illustrations. (credit: modification of work by National Human Genome Research Institute)

Denaturation and Protein Folding

Each protein has its own unique sequence and shape that chemical interactions hold together. If the protein is subject to changes in temperature, pH, or exposure to chemicals, the protein structure may change, losing its shape without losing its primary sequence in what scientists call denaturation. Denaturation is often reversible because the polypeptide's primary structure is conserved in the process if the denaturing agent is removed, allowing the protein to resume its function. Sometimes denaturation is irreversible, leading to loss of function. One example of irreversible protein denaturation is frying an egg. The albumin protein in the liquid egg white denatures when placed in a hot pan. Not all proteins denature at high temperatures. For instance, bacteria that survive in hot springs have proteins that function at temperatures close to boiling. The stomach is also very acidic, has a low pH, and denatures proteins as part of the digestion process; however, the stomach's digestive enzymes retain their activity under these conditions.

Protein folding is critical to its function. Scientists originally thought that the proteins themselves were responsible for the folding process. Only recently researchers discovered that often they receive assistance in the folding process from protein helpers, or **chaperones** (or chaperonins) that associate with the target protein during the folding process. They act by preventing polypeptide aggregation that comprise the complete protein

structure, and they disassociate from the protein once the target protein is folded.

LINK TO LEARNING

For an additional perspective on proteins, view **this animation (http://openstaxcollege.org/l/proteins)** called "Biomolecules: The Proteins."

3.5 | Nucleic Acids

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

- Describe nucleic acids' structure and define the two types of nucleic acids
- Explain DNA's structure and role
- Explain RNA's structure and roles

Nucleic acids are the most important macromolecules for the continuity of life. They carry the cell's genetic blueprint and carry instructions for its functioning.

DNA and RNA

The two main types of nucleic acids are **deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**. DNA is the genetic material in all living organisms, ranging from single-celled bacteria to multicellular mammals. It is in the nucleus of eukaryotes and in the organelles, chloroplasts, and mitochondria. In prokaryotes, the DNA is not enclosed in a membranous envelope.

The cell's entire genetic content is its genome, and the study of genomes is genomics. In eukaryotic cells but not in prokaryotes, DNA forms a complex with histone proteins to form chromatin, the substance of eukaryotic chromosomes. A chromosome may contain tens of thousands of genes. Many genes contain the information to make protein products. Other genes code for RNA products. DNA controls all of the cellular activities by turning the genes "on" or "off."

The other type of nucleic acid, RNA, is mostly involved in protein synthesis. The DNA molecules never leave the nucleus but instead use an intermediary to communicate with the rest of the cell. This intermediary is the **messenger RNA (mRNA)**. Other types of RNA—like rRNA, tRNA, and microRNA—are involved in protein synthesis and its regulation.

DNA and RNA are comprised of monomers that scientists call **nucleotides**. The nucleotides combine with each other to form a **polynucleotide**, DNA or RNA. Three components comprise each nucleotide: a nitrogenous base, a pentose (five-carbon) sugar, and a phosphate group (**Figure 3.31**). Each nitrogenous base in a nucleotide is attached to a sugar molecule, which is attached to one or more phosphate groups.
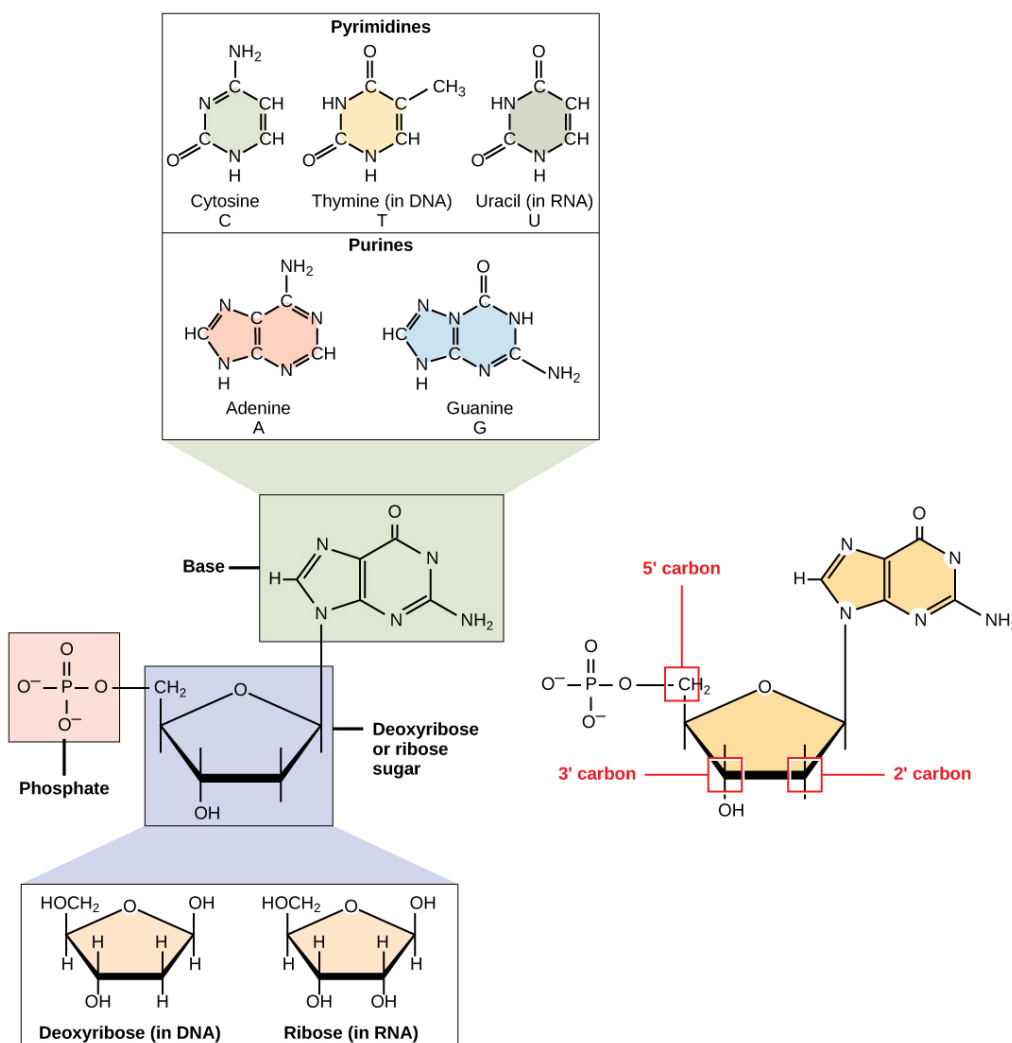
Figure 3.31 Three components comprise a nucleotide: a nitrogenous base, a pentose sugar, and one or more phosphate groups. Carbon residues in the pentose are numbered 1′ through 5′ (the prime distinguishes these residues from those in the base, which are numbered without using a prime notation). The base is attached to the ribose's 1′ position, and the phosphate is attached to the 5′ position. When a polynucleotide forms, the incoming nucleotide's 5′ phosphate attaches to the 3′ hydroxyl group at the end of the growing chain. Two types of pentose are in nucleotides, deoxyribose (found in DNA) and ribose (found in RNA). Deoxyribose is similar in structure to ribose, but it has an H instead of an OH at the 2′ position. We can divide bases into two categories: purines and pyrimidines. Purines have a double ring structure, and pyrimidines have a single ring.

The nitrogenous bases, important components of nucleotides, are organic molecules and are so named because they contain carbon and nitrogen. They are bases because they contain an amino group that has the potential of binding an extra hydrogen, and thus decreasing the hydrogen ion concentration in its environment, making it more basic. Each nucleotide in DNA contains one of four possible nitrogenous bases: adenine (A), guanine (G) cytosine (C), and thymine (T).

Scientists classify adenine and guanine as **purines**. The purine's primary structure is two carbon-nitrogen rings. Scientists classify cytosine, thymine, and uracil as **pyrimidines** which have a single carbon-nitrogen ring as their primary structure (**Figure 3.31**). Each of these basic carbon-nitrogen rings has different functional groups attached to it. In molecular biology shorthand, we know the nitrogenous bases by their symbols A, T, G, C, and U. DNA contains A, T, G, and C; whereas, RNA contains A, U, G, and C.

The pentose sugar in DNA is deoxyribose, and in RNA, the sugar is ribose (**Figure 3.31**). The difference between the sugars is the presence of the hydroxyl group on the ribose's second carbon and hydrogen on the deoxyribose's second carbon. The carbon atoms of the sugar molecule are numbered as 1′, 2′, 3′, 4′, and 5′ (1′ is read as "one prime"). The phosphate residue attaches to the hydroxyl group of the 5′ carbon of one sugar and the hydroxyl group of the 3′ carbon of the sugar of the next nucleotide, which forms a 5′–3′ **phosphodiester**

linkage. A simple dehydration reaction like the other linkages connecting monomers in macromolecules does not form the phosphodiester linkage. Its formation involves removing two phosphate groups. A polynucleotide may have thousands of such phosphodiester linkages.

DNA Double-Helix Structure

DNA has a double-helix structure (**Figure 3.32**). The sugar and phosphate lie on the outside of the helix, forming the DNA's backbone. The nitrogenous bases are stacked in the interior, like a pair of staircase steps. Hydrogen bonds bind the pairs to each other. Every base pair in the double helix is separated from the next base pair by 0.34 nm. The helix's two strands run in opposite directions, meaning that the 5′ carbon end of one strand will face the 3′ carbon end of its matching strand. (Scientists call this an antiparallel orientation and is important to DNA replication and in many nucleic acid interactions.)
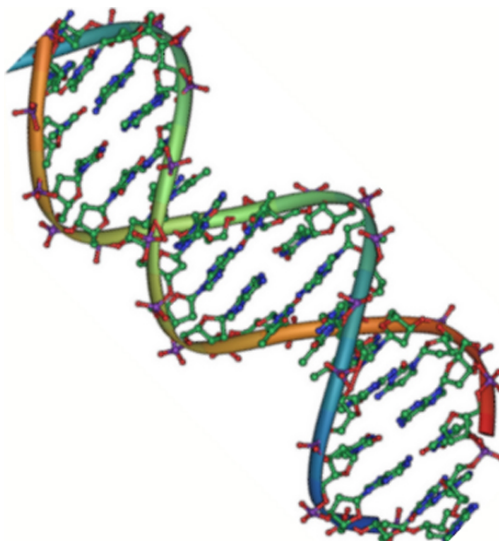
Figure 3.32 Native DNA is an antiparallel double helix. The phosphate backbone (indicated by the curvy lines) is on the outside, and the bases are on the inside. Each base from one strand interacts via hydrogen bonding with a base from the opposing strand. (credit: Jerome Walker/Dennis Myts)

Only certain types of base pairing are allowed. For example, a certain purine can only pair with a certain pyrimidine. This means A can pair with T, and G can pair with C, as **Figure 3.33** shows. This is the base complementary rule. In other words, the DNA strands are complementary to each other. If the sequence of one strand is AATTGGCC, the complementary strand would have the sequence TTAACCGG. During DNA replication, each strand copies itself, resulting in a daughter DNA double helix containing one parental DNA strand and a newly synthesized strand.
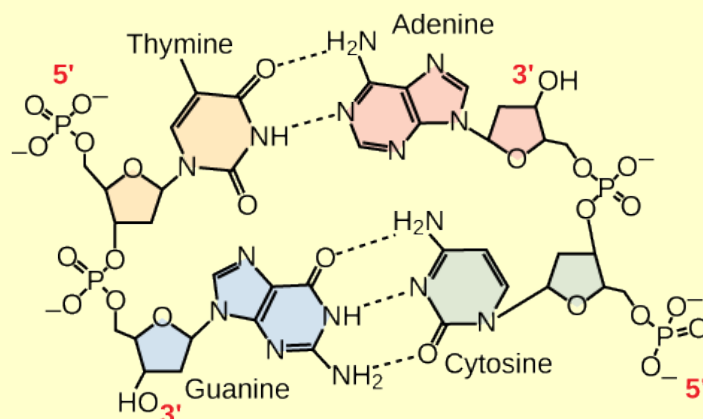
visual CONNECTION

Figure 3.33 In a double stranded DNA molecule, the two strands run antiparallel to one another so that one strand runs 5′ to 3′ and the other 3′ to 5′. The phosphate backbone is located on the outside, and the bases are in the middle. Adenine forms hydrogen bonds (or base pairs) with thymine, and guanine base pairs with cytosine.

A mutation occurs, and adenine replaces cytosine. What impact do you think this will have on the DNA structure?

RNA

Ribonucleic acid, or RNA, is mainly involved in the process of protein synthesis under the direction of DNA. RNA is usually single-stranded and is comprised of ribonucleotides that are linked by phosphodiester bonds. A ribonucleotide in the RNA chain contains ribose (the pentose sugar), one of the four nitrogenous bases (A, U, G, and C), and the phosphate group.

There are four major types of RNA: messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), and microRNA (miRNA). The first, mRNA, carries the message from DNA, which controls all of the cellular activities in a cell. If a cell requires synthesizing a certain protein, the gene for this product turns "on" and the messenger RNA synthesizes in the nucleus. The RNA base sequence is complementary to the DNA's coding sequence from which it has been copied. However, in RNA, the base T is absent and U is present instead. If the DNA strand has a sequence AATTGCGC, the sequence of the complementary RNA is UUAACGCG. In the cytoplasm, the mRNA interacts with ribosomes and other cellular machinery (**Figure 3.34**).
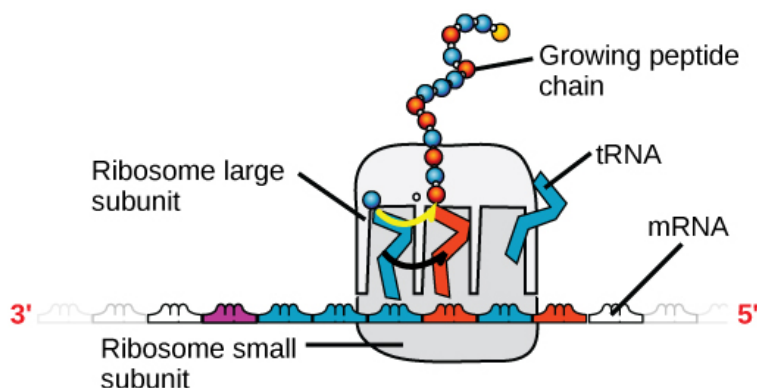

Figure 3.34 A ribosome has two parts: a large subunit and a small subunit. The mRNA sits in between the two subunits. A tRNA molecule recognizes a codon on the mRNA, binds to it by complementary base pairing, and adds the correct amino acid to the growing peptide chain.

The mRNA is read in sets of three bases known as codons. Each codon codes for a single amino acid. In this way, the mRNA is read and the protein product is made. **Ribosomal RNA (rRNA)** is a major constituent of

ribosomes on which the mRNA binds. The rRNA ensures the proper alignment of the mRNA and the Ribosomes. The ribosome's rRNA also has an enzymatic activity (peptidyl transferase) and catalyzes peptide bond formation between two aligned amino acids. **Transfer RNA (tRNA)** is one of the smallest of the four types of RNA, usually 70–90 nucleotides long. It carries the correct amino acid to the protein synthesis site. It is the base pairing between the tRNA and mRNA that allows for the correct amino acid to insert itself in the polypeptide chain. MicroRNAs are the smallest RNA molecules and their role involves regulating gene expression by interfering with the expression of certain mRNA messages. **Table 3.2** summarizes DNA and RNA features.

DNA and RNA Features

	DNA	RNA
Function	Carries genetic information	Involved in protein synthesis
Location	Remains in the nucleus	Leaves the nucleus
Structure	Double helix	Usually single-stranded
Sugar	Deoxyribose	Ribose
Pyrimidines	Cytosine, thymine	Cytosine, uracil
Purines	Adenine, guanine	Adenine, guanine

Table 3.2

Even though the RNA is single stranded, most RNA types show extensive intramolecular base pairing between complementary sequences, creating a predictable three-dimensional structure essential for their function.

As you have learned, information flow in an organism takes place from DNA to RNA to protein. DNA dictates the structure of mRNA in a process scientists call **transcription**, and RNA dictates the protein's structure in a process scientists call **translation**. This is the Central Dogma of Life, which holds true for all organisms; however, exceptions to the rule occur in connection with viral infections.

LINK TO LEARNING

To learn more about DNA, explore the **Howard Hughes Medical Institute BioInteractive animations (http://openstaxcollege.org/l/DNA)** on the topic of DNA.