3

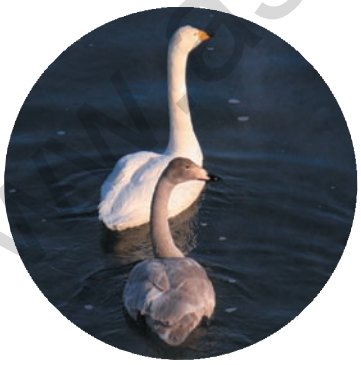
Water and Life

KEY CONCEPTS

3.1 Polar covalent bonds in water molecules result in hydrogen bonding

3.2 Four emergent properties of water contribute to Earth's suitability for life

3.3 Acidic and basic conditions affect living organisms



▲ A young whooper swan paddles after its parent.

▲ **Figure 3.1** How does the habitat of a whooper swan depend on the chemistry of water?

The Molecule That Supports All of Life

Life on Earth began in water and evolved there for 3 billion years before spreading onto land. Water is the substance that makes possible life as we know it here on Earth. All organisms familiar to us are made mostly of water and live in an environment dominated by water. Water is the biological medium here on Earth, and possibly on other planets as well.

Three-quarters of Earth's surface is covered by water. Although most of this water is in liquid form, water is also present on Earth as a solid (ice) and a gas (water vapor). Water is the only common substance to exist in the natural environment in all three physical states of matter. Furthermore, the solid state of water floats on the liquid, a rare property emerging from the chemistry of the water molecule. All three states of water can be seen in **Figure 3.1**, which shows water vapor rising from hot springs that feed into a partially frozen lake in Hokkaido, Japan. The lake is a migratory stop for the elegant whooper swan (*Cygnus cygnus*). The growing young require a watery habitat because their legs can't support their body weight on land for long periods of time.

In this chapter, you will learn how the structure of a water molecule allows it to interact with other molecules, including other water molecules. This ability leads to water's unique emergent properties that help make Earth suitable for life.

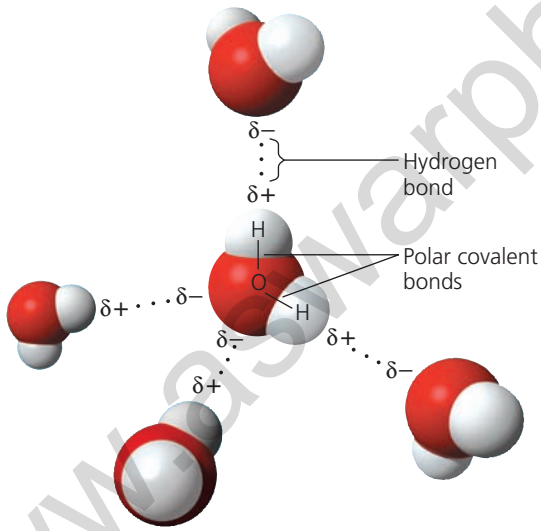
CONCEPT 3.1

Polar covalent bonds in water molecules result in hydrogen bonding

Water is so familiar to us that it is easy to overlook its many extraordinary qualities. Following the theme of emergent properties, we can trace water's unique behavior to the structure and interactions of its molecules.

Studied on its own, the water molecule is deceptively simple. It is shaped like a wide V, with its two hydrogen atoms joined to the oxygen atom by single covalent bonds. Oxygen is more electronegative than hydrogen, so the electrons of the covalent bonds spend more time closer to oxygen than to hydrogen; these are **polar covalent bonds** (see Figure 2.11). This unequal sharing of electrons and water's V-like shape make it a **polar molecule**, meaning that its overall charge is unevenly distributed. In water, the oxygen region of the molecule has a partial negative charge (δ–), and each hydrogen has a partial positive charge (δ+).

The properties of water arise from attractions between oppositely charged atoms of different water molecules: The slightly positive hydrogen of one molecule is attracted to the slightly negative oxygen of a nearby molecule. The two molecules are thus held together by a hydrogen bond **(Figure 3.2)**. When water is in its liquid form, its hydrogen bonds are very fragile, each only about 1/20 as strong as a covalent bond. The hydrogen bonds form, break, and re-form with great frequency. Each lasts only a few trillionths of a second, but the molecules are constantly forming new hydrogen bonds with a succession of partners. Therefore, at any instant, most of the water molecules are hydrogen-bonded to their neighbors. The extraordinary properties of water emerge from this hydrogen bonding , which organizes water molecules into a higher level of structural order.

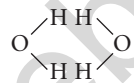


▲ **Figure 3.2 Hydrogen bonds between water molecules.** The charged regions in a water molecule are due to its polar covalent bonds. Oppositely charged regions of neighboring water molecules are attracted to each other, forming hydrogen bonds. Each molecule can hydrogen-bond to multiple partners, and these associations are constantly changing.

DRAW IT *Draw partial charges on the water molecule at the far left, and draw two more water molecules hydrogen-bonded to it.*

CONCEPT CHECK 3.1

1. **MAKE CONNECTIONS** What is electronegativity, and how does it affect interactions between water molecules? (Review Figure 2.11.)
2. Why is it unlikely that two neighboring water molecules would be arranged like this?



3. **WHAT IF?** What would be the effect on the properties of the water molecule if oxygen and hydrogen had equal electronegativity?

For suggested answers, see Appendix A.

CONCEPT 3.2

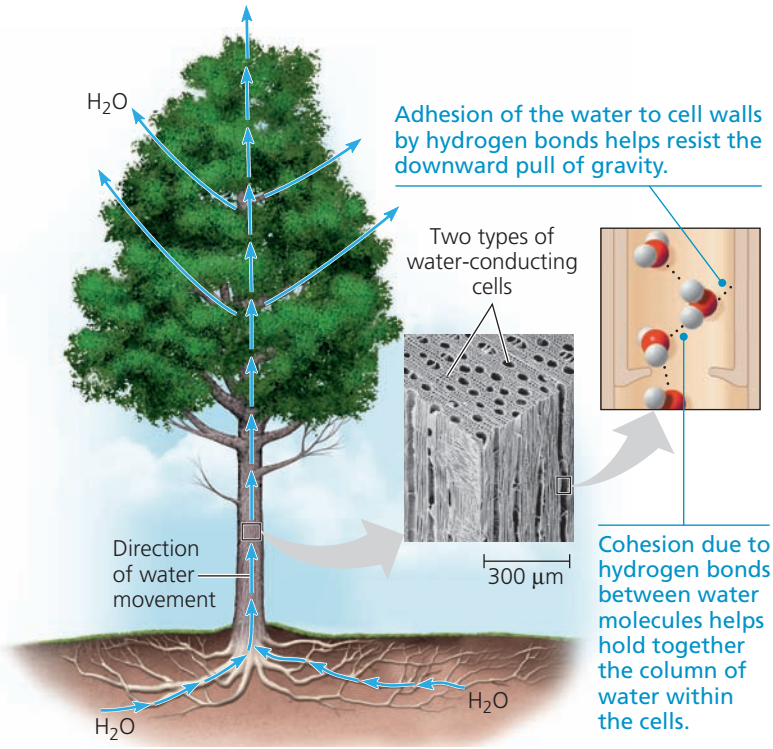
Four emergent properties of water contribute to Earth's suitability for life

We will examine four emergent properties of water that contribute to Earth's suitability as an environment for life: cohesive behavior, ability to moderate temperature, expansion upon freezing, and versatility as a solvent.

Cohesion of Water Molecules

Water molecules stay close to each other as a result of hydrogen bonding. Although the arrangement of molecules in a sample of liquid water is constantly changing, at any given moment many of the molecules are linked by multiple hydrogen bonds. These linkages make water more structured than most other liquids. Collectively, the hydrogen bonds hold the substance together, a phenomenon called **cohesion**.

Cohesion due to hydrogen bonding contributes to the transport of water and dissolved nutrients against gravity in plants. Water from the roots reaches the leaves through a network of water-conducting cells **(Figure 3.3)**. As water evaporates from a leaf, hydrogen bonds cause water molecules leaving the veins to tug on molecules farther down, and the upward pull is transmitted through the water-conducting cells all the way to the roots. **Adhesion**, the clinging of one substance to another, also plays a role. Adhesion of



▲ Figure 3.3 Water transport in plants. Evaporation from leaves pulls water upward from the roots through water-conducting cells. Because of the properties of cohesion and adhesion, the tallest trees can transport water more than 100 m upward—approximately one-quarter the height of the Empire State Building in New York City.

BioFlix Visit the Study Area in **MasteringBiology** for the BioFlix ® 3-D Animation on Water Transport in Plants.

water by hydrogen bonds to the molecules of cell walls helps counter the downward pull of gravity (see Figure 3.3).

Related to cohesion is **surface tension**, a measure of how difficult it is to stretch or break the surface of a liquid. At the interface between water and air is an ordered arrangement of water molecules, hydrogen-bonded to one another and to the water below. This gives water an unusually high surface tension, making it behave as though it were coated with an invisible film. You can observe the surface tension of water by slightly overfilling a drinking glass; the water will stand above the rim. The spider in **Figure 3.4** takes advantage of the surface tension of water to walk across a pond without breaking the surface.

▼ Figure 3.4 Walking on water. The high surface tension of water, resulting from the collective strength of its hydrogen bonds, allows this raft spider to walk on the surface of a pond.

Moderation of Temperature by Water

Water moderates air temperature by absorbing heat from air that is warmer and releasing the stored heat to air that is cooler. Water is effective as a heat bank because it can absorb or release a relatively large amount of heat with only a slight change in its own temperature. To understand this capability of water, let's first look at temperature and heat.

Temperature and Heat

Anything that moves has **kinetic energy**, the energy of motion. Atoms and molecules have kinetic energy because they are always moving, although not necessarily in any particular direction. The faster a molecule moves, the greater its kinetic energy. The kinetic energy associated with the random movement of atoms or molecules is called **thermal energy**. Thermal energy is related to temperature, but they are not the same thing. **Temperature** is a measure of energy that represents the *average* kinetic energy of the molecules in a body of matter, regardless of volume, whereas the *total* thermal energy depends in part on the matter's volume. When water is heated in a coffeemaker, the average speed of the molecules increases, and the thermometer records this as a rise in temperature of the liquid. The total amount of thermal energy also increases in this case. Note, however, that although the pot of coffee has a much higher temperature than, say, the water in a swimming pool, the swimming pool contains more thermal energy because of its much greater volume.

Whenever two objects of different temperature are brought together, thermal energy passes from the warmer to the cooler object until the two are the same temperature. Molecules in the cooler object speed up at the expense of the thermal energy of the warmer object. An ice cube cools a drink not by adding coldness to the liquid, but by absorbing thermal energy from the liquid as the ice itself melts. Thermal energy in transfer from one body of matter to another is defined as **heat**.

One convenient unit of heat used in this book is the **calorie (cal)**. A calorie is the amount of heat it takes to raise the temperature of 1 g of water by 1°C. Conversely, a calorie is also the amount of heat that 1 g of water releases when it cools by 1°C. A **kilocalorie (kcal)**, 1,000 cal, is the quantity of heat required to raise the temperature of 1 kilogram (kg) of water by 1°C. (The "calories" on food packages are actually kilocalories.) Another energy unit used in this book is the **joule (J)**. One joule equals 0.239 cal; one calorie equals 4.184 J.

Water's High Specific Heat

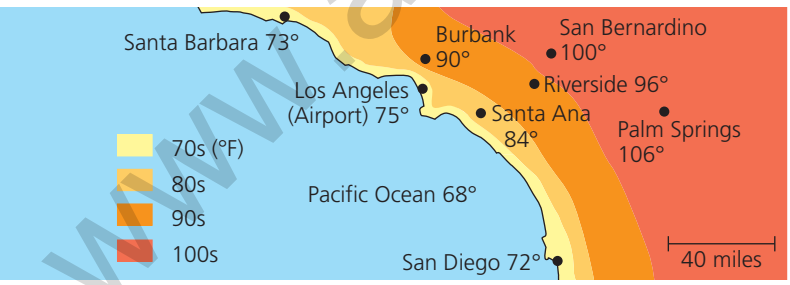
The ability of water to stabilize temperature stems from its relatively high specific heat. The **specific heat** of a substance is defined as the amount of heat that must be absorbed or lost for 1 g of that substance to change its temperature by 1°C. We

already know water's specific heat because we have defined a calorie as the amount of heat that causes 1 g of water to change its temperature by 1°C. Therefore, the specific heat of water is 1 calorie per gram and per degree Celsius, abbreviated as 1 cal/g · °C. Compared with most other substances, water has an unusually high specific heat. For example, ethyl alcohol, the type of alcohol in alcoholic beverages, has a specific heat of 0.6 cal/g · °C; that is, only 0.6 cal is required to raise the temperature of 1 g of ethyl alcohol by 1°C.

Because of the high specific heat of water relative to other materials, water will change its temperature less than other liquids when it absorbs or loses a given amount of heat. The reason you can burn your fingers by touching the side of an iron pot on the stove when the water in the pot is still lukewarm is that the specific heat of water is ten times greater than that of iron. In other words, the same amount of heat will raise the temperature of 1 g of the iron much faster than it will raise the temperature of 1 g of the water. Specific heat can be thought of as a measure of how well a substance resists changing its temperature when it absorbs or releases heat. Water resists changing its temperature; when it does change its temperature, it absorbs or loses a relatively large quantity of heat for each degree of change.

We can trace water's high specific heat, like many of its other properties, to hydrogen bonding. Heat must be absorbed in order to break hydrogen bonds; by the same token, heat is released when hydrogen bonds form. A calorie of heat causes a relatively small change in the temperature of water because much of the heat is used to disrupt hydrogen bonds before the water molecules can begin moving faster. And when the temperature of water drops slightly, many additional hydrogen bonds form, releasing a considerable amount of energy in the form of heat.

What is the relevance of water's high specific heat to life on Earth? A large body of water can absorb and store a huge amount of heat from the sun in the daytime and during summer while warming up only a few degrees. At night and during winter, the gradually cooling water can warm the air. This capability of water serves to moderate air temperatures in coastal areas **(Figure 3.5)**. The high specific heat of water also tends to stabilize ocean temperatures, creating a favorable environment for marine life. Thus, because of its high specific heat, the water that covers most of Earth keeps temperature fluctuations on land and in water within limits that permit life. Also, because organisms are made primarily of water, they are better able to resist changes in their own temperature than if they were made of a liquid with a lower specific heat.



▲ **Figure 3.5** **Temperatures for the Pacific Ocean and Southern California on an August day.**

INTERPRET THE DATA *Explain the pattern of temperatures shown in this diagram.*

Evaporative Cooling

Molecules of any liquid stay close together because they are attracted to one another. Molecules moving fast enough to overcome these attractions can depart the liquid and enter the air as a gas (vapor). This transformation from a liquid to a gas is called vaporization, or evaporation. Recall that the speed of molecular movement varies and that temperature is the *average* kinetic energy of molecules. Even at low temperatures, the speediest molecules can escape into the air. Some evaporation occurs at any temperature; a glass of water at room temperature, for example, will eventually evaporate completely. If a liquid is heated, the average kinetic energy of molecules increases and the liquid evaporates more rapidly.

Heat of vaporization is the quantity of heat a liquid must absorb for 1 g of it to be converted from the liquid to the gaseous state. For the same reason that water has a high specific heat, it also has a high heat of vaporization relative to most other liquids. To evaporate 1 g of water at 25°C, about 580 cal of heat is needed—nearly double the amount needed to vaporize a gram of alcohol or ammonia. Water's high heat of vaporization is another emergent property resulting from the strength of its hydrogen bonds, which must be broken before the molecules can exit from the liquid in the form of water vapor (see Figure 3.1).

The high amount of energy required to vaporize water has a wide range of effects. On a global scale, for example, it helps moderate Earth's climate. A considerable amount of solar heat absorbed by tropical seas is consumed during the evaporation of surface water. Then, as moist tropical air circulates poleward, it releases heat as it condenses and forms rain. On an organismal level, water's high heat of vaporization accounts for the severity of steam burns. These burns are caused by the heat energy released when steam condenses into liquid on the skin.

As a liquid evaporates, the surface of the liquid that remains behind cools down (its temperature decreases). This **evaporative cooling** occurs because the "hottest" molecules, those with the greatest kinetic energy, are the most likely to leave as gas. It is as if the hundred fastest runners at a college transferred to another school; the average speed of the remaining students would decline.

Evaporative cooling of water contributes to the stability of temperature in lakes and ponds and also provides a mechanism that prevents terrestrial organisms from overheating. For example, evaporation of water from the leaves of a plant

helps keep the tissues in the leaves from becoming too warm in the sunlight. Evaporation of sweat from human skin dissipates body heat and helps prevent overheating on a hot day or when excess heat is generated by strenuous activity. High humidity on a hot day increases discomfort because the high concentration of water vapor in the air inhibits the evaporation of sweat from the body.

Floating of Ice on Liquid Water

Water is one of the few substances that are less dense as a solid than as a liquid. In other words, ice floats on liquid water. While other materials contract and become denser when they solidify, water expands. The cause of this exotic behavior is, once again, hydrogen bonding. At temperatures above 4°C, water behaves like other liquids, expanding as it warms and contracting as it cools. As the temperature falls from 4°C to 0°C, water begins to freeze because more and more of its molecules are moving too slowly to break hydrogen bonds. At 0°C, the molecules become locked into a crystalline lattice, each water molecule hydrogen-bonded to four partners **(Figure 3.6)**. The hydrogen bonds keep the molecules at "arm's length," far enough apart to make ice about 10% less dense (10% fewer molecules for the same volume) than liquid water at 4°C. When ice absorbs enough heat for its temperature to rise above 0°C, hydrogen bonds between molecules are disrupted. As the crystal collapses, the ice melts, and molecules are free to slip closer together. Water reaches its greatest density at 4°C and then begins to expand as the molecules move faster. Even in liquid water, many of the molecules are connected by hydrogen bonds, though only transiently: The hydrogen bonds are constantly breaking and re-forming.

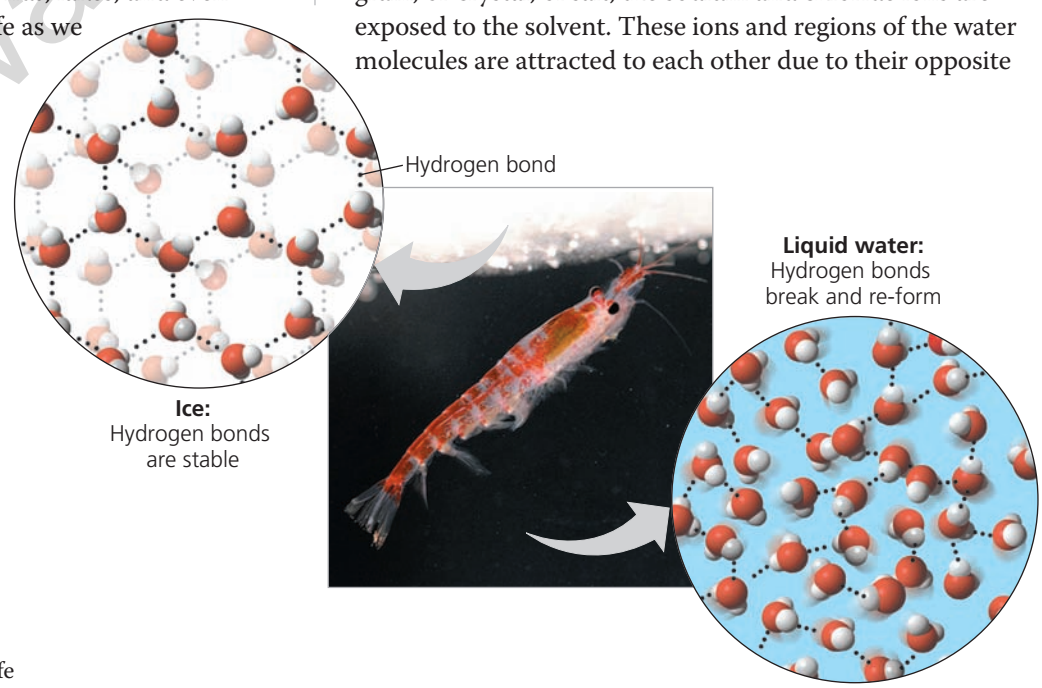
The ability of ice to float due to its lower density is an important factor in the suitability of the environment for life. If ice sank, then eventually all ponds, lakes, and even oceans would freeze solid, making life as we know it impossible on Earth. During summer, only the upper few inches of the ocean would thaw. Instead, when a deep body of water cools, the floating ice insulates the liquid water below, preventing it from freezing and allowing life to exist under the frozen surface, as shown in the photo in Figure 3.6. Besides insulating the water below, ice also provides a solid habitat for some animals, such as polar bears and seals.

Many scientists are worried that these bodies of ice are at risk of disappearing. Global warming, which is caused by carbon dioxide and other "greenhouse" gases in the atmosphere, is having a profound effect on icy environments around the globe. In the Arctic, the average air temperature has risen 1.4°C just since 1961. This temperature increase has affected the seasonal balance between Arctic sea ice and liquid water, causing ice to form later in the year, to melt earlier, and to cover a smaller area. The rate at which glaciers and Arctic sea ice are disappearing is posing an extreme challenge to animals that depend on ice for their survival.

Water: The Solvent of Life

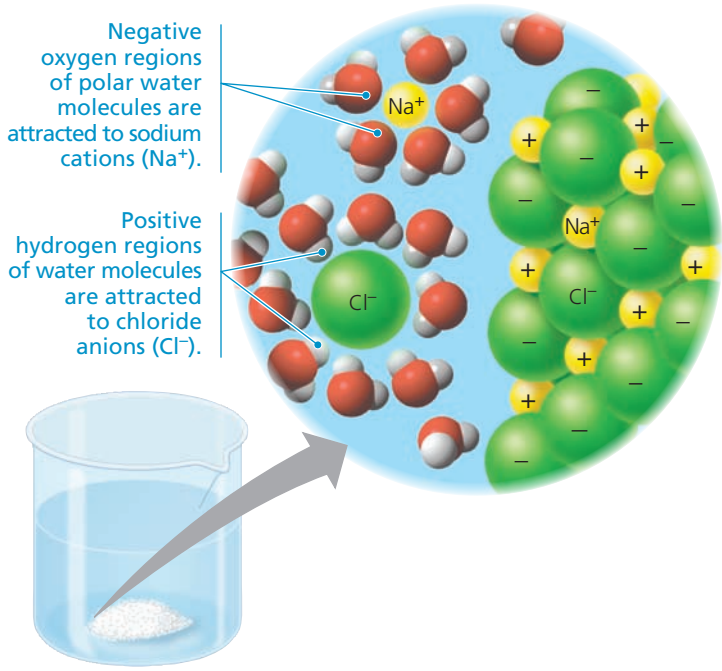
A sugar cube placed in a glass of water will dissolve with a little stirring. The glass will then contain a uniform mixture of sugar and water; the concentration of dissolved sugar will be the same everywhere in the mixture. A liquid that is a completely homogeneous mixture of two or more substances is called a **solution**. The dissolving agent of a solution is the **solvent**, and the substance that is dissolved is the **solute**. In this case, water is the solvent and sugar is the solute. An **aqueous solution** is one in which the solute is dissolved in water; water is the solvent.

Water is a very versatile solvent, a quality we can trace to the polarity of the water molecule. Suppose, for example, that a spoonful of table salt, the ionic compound sodium chloride (NaCl), is placed in water **(Figure 3.7)**. At the surface of each grain, or crystal, of salt, the sodium and chloride ions are exposed to the solvent. These ions and regions of the water molecules are attracted to each other due to their opposite



▶ Figure 3.6 Ice: crystalline structure and floating barrier. In ice, each molecule is hydrogen-bonded to four neighbors in a three-dimensional crystal. Because the crystal is spacious, ice has fewer molecules than an equal volume of liquid water. In other words, ice is less dense than liquid water. Floating ice becomes a barrier that insulates the liquid water below from the colder air. The marine organism shown here is a type of shrimp called krill; it was photographed beneath floating ice in the Southern Ocean near Antarctica.

WHAT IF? *If water did not form hydrogen bonds, what would happen to the shrimp's habitat, shown here?*



▲ **Figure 3.7 Table salt dissolving in water.** A sphere of water molecules, called a hydration shell, surrounds each solute ion.

WHAT IF? *What would happen if you heated this solution for a long time?*

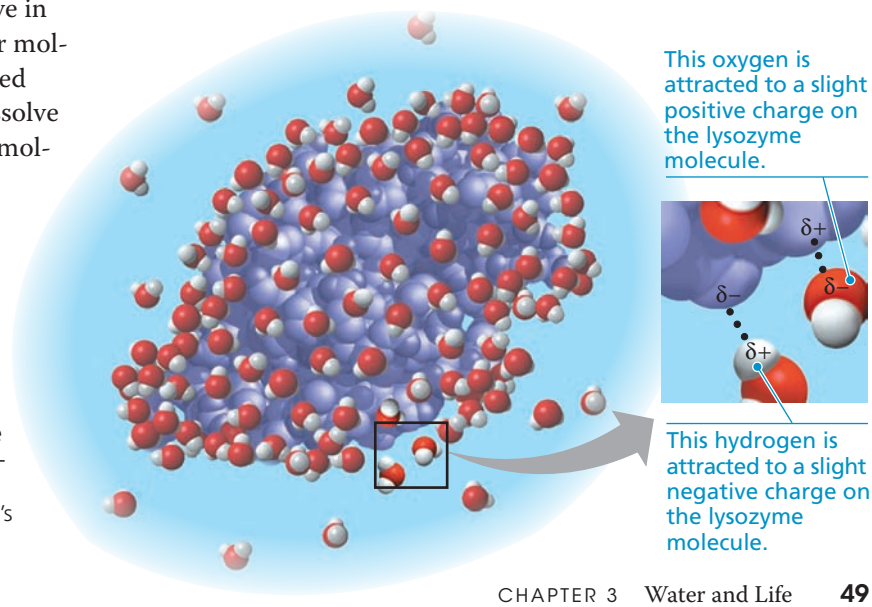
charges. The oxygen regions of the water molecules are negatively charged and are attracted to sodium cations. The hydrogen regions are positively charged and are attracted to chloride anions. As a result, water molecules surround the individual sodium and chloride ions, separating and shielding them from one another. The sphere of water molecules around each dissolved ion is called a **hydration shell**. Working inward from the surface of each salt crystal, water eventually dissolves all the ions. The result is a solution of two solutes, sodium cations and chloride anions, homogeneously mixed with water, the solvent. Other ionic compounds also dissolve in water. Seawater, for instance, contains a great variety of dissolved ions, as do living cells.

A compound does not need to be ionic to dissolve in water; many compounds made up of nonionic polar molecules, such as the sugar in the sugar cube mentioned earlier, are also water-soluble. Such compounds dissolve when water molecules surround each of the solute molecules, forming hydrogen bonds with them. Even molecules as large as proteins can dissolve in water if they have ionic and polar regions on their surface **(Figure 3.8)**. Many different kinds of polar compounds are dissolved (along with ions) in the water of such biological fluids as blood, the sap of plants, and the liquid within all cells. Water is the solvent of life.

Hydrophilic and Hydrophobic Substances

Any substance that has an affinity for water is said to be **hydrophilic** (from the Greek *hydro*, water, and *philos*, loving). In some cases, substances can be hydrophilic without actually dissolving. For example, some molecules in cells are so large that they do not dissolve. Another example of a hydrophilic substance that does not dissolve is cotton, a plant product. Cotton consists of giant molecules of cellulose, a compound with numerous regions of partial positive and partial negative charges that can form hydrogen bonds with water. Water adheres to the cellulose fibers. Thus, a cotton towel does a great job of drying the body, yet it does not dissolve in the washing machine. Cellulose is also present in the walls of water-conducting cells in a plant; you read earlier how the adhesion of water to these hydrophilic walls helps water move up the plant against gravity.

There are, of course, substances that do not have an affinity for water. Substances that are nonionic and nonpolar (or otherwise cannot form hydrogen bonds) actually seem to repel water; these substances are said to be **hydrophobic** (from the Greek *phobos*, fearing). An example from the kitchen is vegetable oil, which, as you know, does not mix stably with water-based substances such as vinegar. The hydrophobic behavior of the oil molecules results from a prevalence of relatively nonpolar covalent bonds, in this case bonds between carbon and hydrogen, which share electrons almost equally. Hydrophobic molecules related to oils are major ingredients of cell membranes. (Imagine what would happen to a cell if its membrane dissolved!)



▶ **Figure 3.8 A water-soluble protein.** Human lysozyme is a protein found in tears and saliva that has antibacterial action. This model shows the lysozyme molecule (purple) in an aqueous environment. Ionic and polar regions on the protein's surface attract water molecules.

Solute Concentration in Aqueous Solutions

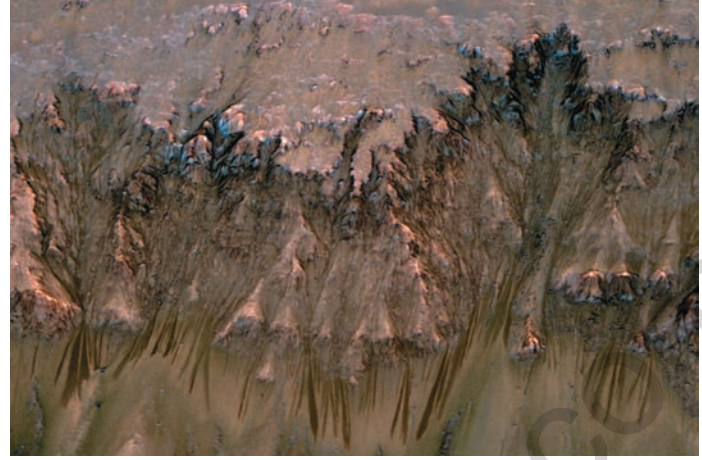
Most of the chemical reactions in organisms involve solutes dissolved in water. To understand such reactions, we must know how many atoms and molecules are involved and calculate the concentration of solutes in an aqueous solution (the number of solute molecules in a volume of solution).

When carrying out experiments, we use mass to calculate the number of molecules. We must first calculate the **molecular mass**, which is the sum of the masses of all the atoms in a molecule. As an example, let's calculate the molecular mass of table sugar (sucrose), $C_{12}H_{22}O_{11}$. In round numbers of daltons, the mass of a carbon atom is 12, the mass of a hydrogen atom is 1, and the mass of an oxygen atom is 16. Thus, sucrose has a molecular mass of $(12 \times 12) + (22 \times 1) + (11 \times 16) = 342$ daltons. Because we can't weigh out small numbers of molecules, we usually measure substances in units called moles. Just as a dozen always means 12 objects, a **mole (mol)** represents an exact number of objects: 6.02×10^{23}, which is called Avogadro's number. Because of the way in which Avogadro's number and the unit *dalton* were originally defined, there are 6.02×10^{23} daltons in 1 g. Once we determine the molecular mass of a molecule such as sucrose, we can use the same number (342), but with the unit *gram*, to represent the mass of 6.02×10^{23} molecules of sucrose, or 1 mol of sucrose (this is sometimes called the *molar mass*). To obtain 1 mol of sucrose in the lab, therefore, we weigh out 342 g.

The practical advantage of measuring a quantity of chemicals in moles is that a mole of one substance has exactly the same number of molecules as a mole of any other substance. If the molecular mass of substance A is 342 daltons and that of substance B is 10 daltons, then 342 g of A will have the same number of molecules as 10 g of B. A mole of ethyl alcohol (C_2H_6O) also contains 6.02×10^{23} molecules, but its mass is only 46 g because the mass of a molecule of ethyl alcohol is less than that of a molecule of sucrose. Measuring in moles makes it convenient for scientists working in the laboratory to combine substances in fixed ratios of molecules.

How would we make a liter (L) of solution consisting of 1 mol of sucrose dissolved in water? We would measure out 342 g of sucrose and then gradually add water, while stirring, until the sugar was completely dissolved. We would then add enough water to bring the total volume of the solution up to 1 L. At that point, we would have a 1-molar (1 *M*) solution of sucrose. **Molarity**—the number of moles of solute per liter of solution—is the unit of concentration most often used by biologists for aqueous solutions.

Water's capacity as a versatile solvent complements the other properties discussed in this chapter. Since these remarkable properties allow water to support life on Earth so well, scientists who seek life elsewhere in the universe look for water as a sign that a planet might sustain life.



▲ Figure 3.9 Evidence for subsurface liquid water on Mars. The dark streaks running down the lower portion of the photo are proposed to be streams of subsurface flowing water because they appear only during the warm season. The gullies in the middle of the photo could have been formed by flowing water.

Possible Evolution of Life on Other Planets

EVOLUTION Biologists who look for life elsewhere in the universe (known as *astrobiologists*) have concentrated their search on planets that might have water. More than 800 planets have been found outside our solar system, and there is evidence for the presence of water vapor on a few of them. In our own solar system, Mars has been a focus of study. Like Earth, Mars has an ice cap at both poles. Images from spacecraft sent to Mars show that ice is present just under the surface of Mars and enough water vapor exists in its atmosphere for frost to form. **Figure 3.9** shows streaks that form along steep slopes during the Mars spring and summer, features that vanish during the winter. Some scientists have proposed that these are seasonal streams of flowing water occurring when subsurface ice melts during the warm season, while others think they are the result of CO_2 rather than water. Drilling below the surface may be the next step in the search for signs of life on Mars. If any life-forms or fossils are found, their study will shed light on the process of evolution from an entirely new perspective.

CONCEPT CHECK 3.2

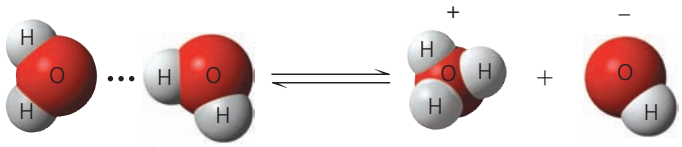
1. Describe how properties of water contribute to the upward movement of water in a tree.
2. Explain the saying "It's not the heat; it's the humidity."
3. How can the freezing of water crack boulders?
4. **WHAT IF?** A water strider (which can walk on water) has legs that are coated with a hydrophobic substance. What might be the benefit? What would happen if the substance were hydrophilic?
5. **INTERPRET THE DATA** The concentration of the appetite-regulating hormone ghrelin is about 1.3×10^{-10} *M* in the blood of a fasting person. How many molecules of ghrelin are in 1 L of blood?

For suggested answers, see Appendix A.

CONCEPT 3.3

Acidic and basic conditions affect living organisms

Occasionally, a hydrogen atom participating in a hydrogen bond between two water molecules shifts from one molecule to the other. When this happens, the hydrogen atom leaves its electron behind, and what is actually transferred is a **hydrogen ion** (H^+), a single proton with a charge of 1+. The water molecule that lost a proton is now a **hydroxide ion** (OH^-), which has a charge of 1−. The proton binds to the other water molecule, making that molecule a **hydronium ion** (H_3O^+). We can picture the chemical reaction as follows:



By convention, H^+ (the hydrogen ion) is used to represent H_3O^+ (the hydronium ion), and we follow that practice in this book. Keep in mind, though, that H^+ does not exist on its own in an aqueous solution. It is always associated with a water molecule in the form of H_3O^+.

As indicated by the double arrows, this is a reversible reaction that reaches a state of dynamic equilibrium when water molecules dissociate at the same rate that they are being reformed from H^+ and OH^-. At this equilibrium point, the concentration of water molecules greatly exceeds the concentrations of H^+ and OH^-. In pure water, only one water molecule in every 554 million is dissociated; the concentration of each ion in pure water is 10^{-7} *M* (at 25°C). This means there is only one ten-millionth of a mole of hydrogen ions per liter of pure water and an equal number of hydroxide ions. (Even so, this is a huge number—over 60,000 *trillion*—of each ion.)

Although the dissociation of water is reversible and statistically rare, it is exceedingly important in the chemistry of life. H^+ and OH^- are very reactive. Changes in their concentrations can drastically affect a cell's proteins and other complex molecules. As we have seen, the concentrations of H^+ and OH^- are equal in pure water, but adding certain kinds of solutes, called acids and bases, disrupts this balance. Biologists use something called the pH scale to describe how acidic or basic (the opposite of acidic) a solution is. In the remainder of this chapter, you will learn about acids, bases, and pH and why changes in pH can adversely affect organisms.

Acids and Bases

What would cause an aqueous solution to have an imbalance in H^+ and OH^- concentrations? When acids dissolve in water, they donate additional H^+ to the solution. An **acid** is a substance that increases the hydrogen ion concentration of a solution. For example, when hydrochloric acid (HCl) is added to water, hydrogen ions dissociate from chloride ions:

$$HCl \rightarrow H^+ + Cl^-$$

This source of H^+ (dissociation of water is the other source) results in an acidic solution—one having more H^+ than OH^-.

A substance that reduces the hydrogen ion concentration of a solution is called a **base**. Some bases reduce the H^+ concentration directly by accepting hydrogen ions. Ammonia (NH_3), for instance, acts as a base when the unshared electron pair in nitrogen's valence shell attracts a hydrogen ion from the solution, resulting in an ammonium ion (NH_4^+):

$$NH_3 + H^+ \rightleftharpoons NH_4^+$$

Other bases reduce the H^+ concentration indirectly by dissociating to form hydroxide ions, which combine with hydrogen ions and form water. One such base is sodium hydroxide (NaOH), which in water dissociates into its ions:

$$NaOH \rightarrow Na^+ + OH^-$$

In either case, the base reduces the H^+ concentration. Solutions with a higher concentration of OH^- than H^+ are known as basic solutions. A solution in which the H^+ and OH^- concentrations are equal is said to be neutral.

Notice that single arrows were used in the reactions for HCl and NaOH. These compounds dissociate completely when mixed with water, so hydrochloric acid is called a strong acid and sodium hydroxide a strong base. In contrast, ammonia is a weak base. The double arrows in the reaction for ammonia indicate that the binding and release of hydrogen ions are reversible reactions, although at equilibrium there will be a fixed ratio of NH_4^+ to NH_3.

Weak acids are acids that reversibly release and accept back hydrogen ions. An example is carbonic acid:

$$\underset{\text{Carbonic acid}}{H_2CO_3} \rightleftharpoons \underset{\text{Bicarbonate ion}}{HCO_3^-} + \underset{\text{Hydrogen ion}}{H^+}$$

Here the equilibrium so favors the reaction in the left direction that when carbonic acid is added to pure water, only 1% of the molecules are dissociated at any particular time. Still, that is enough to shift the balance of H^+ and OH^- from neutrality.

The pH Scale

In any aqueous solution at 25°C, the *product* of the H^+ and OH^- concentrations is constant at 10^{-14}. This can be written

$$[H^+][OH^-] = 10^{-14}$$

In such an equation, brackets indicate molar concentration. In a neutral solution at 25°C (close to room temperature), $[H^+] = 10^{-7}$ and $[OH^-] = 10^{-7}$, so in this case, 10^{-14} is the

product of 10^{-7} and 10^{-7}. If enough acid is added to a solution to increase $[H^+]$ to 10^{-5} *M*, then $[OH^-]$ will decline by an equivalent factor to 10^{-9} *M* (note that $10^{-5} \times 10^{-9} = 10^{-14}$). This constant relationship expresses the behavior of acids and bases in an aqueous solution. An acid not only adds hydrogen ions to a solution, but also removes hydroxide ions because of the tendency for H^+ to combine with OH^-, forming water. A base has the opposite effect, increasing OH^- concentration but also reducing H^+ concentration by the formation of water. If enough of a base is added to raise the OH^- concentration to 10^{-4} *M*, it will cause the H^+ concentration to drop to 10^{-10} *M*. Whenever we know the concentration of either H^+ or OH^- in an aqueous solution, we can deduce the concentration of the other ion.

Because the H^+ and OH^- concentrations of solutions can vary by a factor of 100 trillion or more, scientists have developed a way to express this variation more conveniently than in moles per liter. The pH scale (**Figure 3.10**) compresses the range of H^+ and OH^- concentrations by employing logarithms. The **pH** of a solution is defined as the negative logarithm (base 10) of the hydrogen ion concentration:

$$pH = -\log [H^+]$$

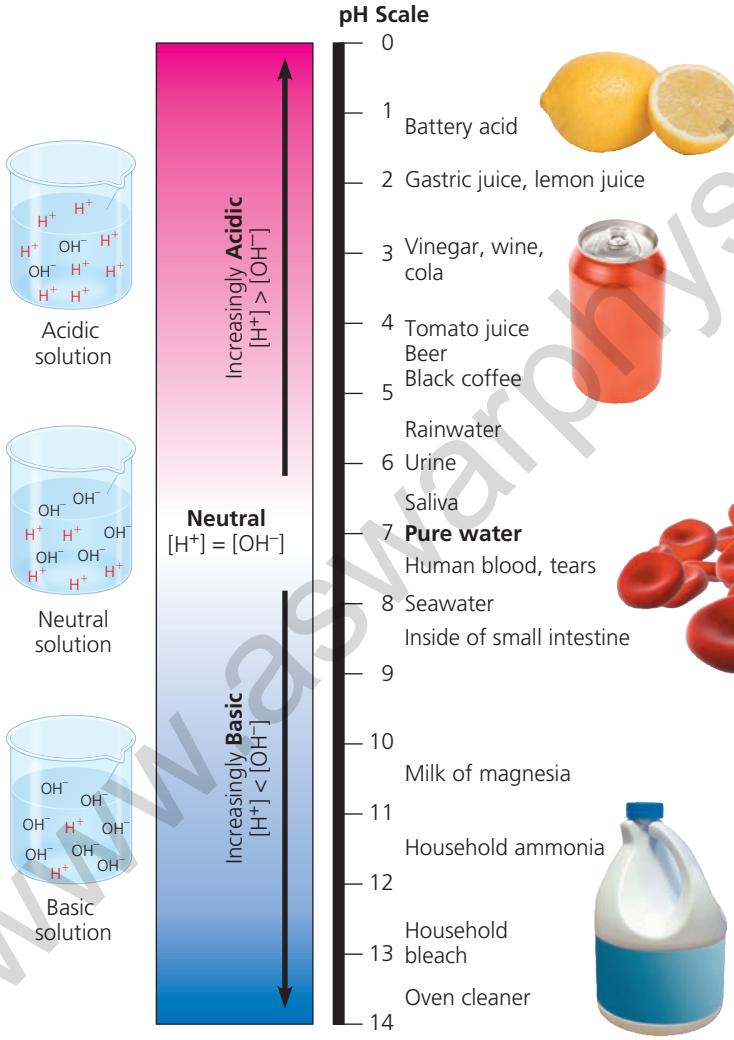
For a neutral aqueous solution, $[H^+]$ is 10^{-7} *M*, giving us

$$-\log 10^{-7} = -(-7) = 7$$

Notice that pH *declines* as H^+ concentration *increases*. Notice, too, that although the pH scale is based on H^+ concentration, it also implies OH^- concentration. A solution of pH 10 has a hydrogen ion concentration of 10^{-10} *M* and a hydroxide ion concentration of 10^{-4} *M*.

The pH of a neutral aqueous solution at 25°C is 7, the midpoint of the pH scale. A pH value less than 7 denotes an acidic solution; the lower the number, the more acidic the solution. The pH for basic solutions is above 7. Most biological fluids, such as blood and saliva, are within the range of pH 6–8. There are a few exceptions, however, including the strongly acidic digestive juice of the human stomach, which has a pH of about 2.

Remember that each pH unit represents a tenfold difference in H^+ and OH^- concentrations. It is this mathematical feature that makes the pH scale so compact. A solution of pH 3 is not twice as acidic as a solution of pH 6, but a thousand times ($10 \times 10 \times 10$) more acidic. When the pH of a solution changes slightly, the actual concentrations of H^+ and OH^- in the solution change substantially.



▲ Figure 3.10 The pH scale and pH values of some aqueous solutions.

Buffers

The internal pH of most living cells is close to 7. Even a slight change in pH can be harmful, because the chemical processes of the cell are very sensitive to the concentrations of hydrogen and hydroxide ions. The pH of human blood is very close to 7.4, which is slightly basic. A person cannot survive for more than a few minutes if the blood pH drops to 7 or rises to 7.8, and a chemical system exists in the blood that maintains a stable pH. If 0.01 mol of a strong acid is added to a liter of pure water, the pH drops from 7.0 to 2.0. If the same amount of acid is added to a liter of blood, however, the pH decrease is only from 7.4 to 7.3. Why does the addition of acid have so much less of an effect on the pH of blood than it does on the pH of water?

The presence of substances called buffers allows biological fluids to maintain a relatively constant pH despite the addition of acids or bases. A **buffer** is a substance that minimizes changes in the concentrations of H^+ and OH^- in a solution. It does so by accepting hydrogen ions from the solution when they are in excess and donating hydrogen ions to the solution when they have been depleted. Most buffer solutions contain a weak acid and its corresponding base, which combine reversibly with hydrogen ions.

Several buffers contribute to pH stability in human blood and many other biological solutions. One of these is

carbonic acid (H_2CO_3), which is formed when CO_2 reacts with water in blood plasma. As mentioned earlier, carbonic acid dissociates to yield a bicarbonate ion (HCO_3^-) and a hydrogen ion (H^+):

$$\underset{\substack{H^+ \text{ donor}\\(\text{acid})}}{H_2CO_3} \underset{\substack{\text{Response to}\\\text{a drop in pH}}}{\overset{\substack{\text{Response to}\\\text{a rise in pH}}}{\rightleftharpoons}} \underset{\substack{H^+ \text{ acceptor}\\(\text{base})}}{HCO_3^-} + \underset{\substack{\text{Hydrogen}\\\text{ion}}}{H^+}$$

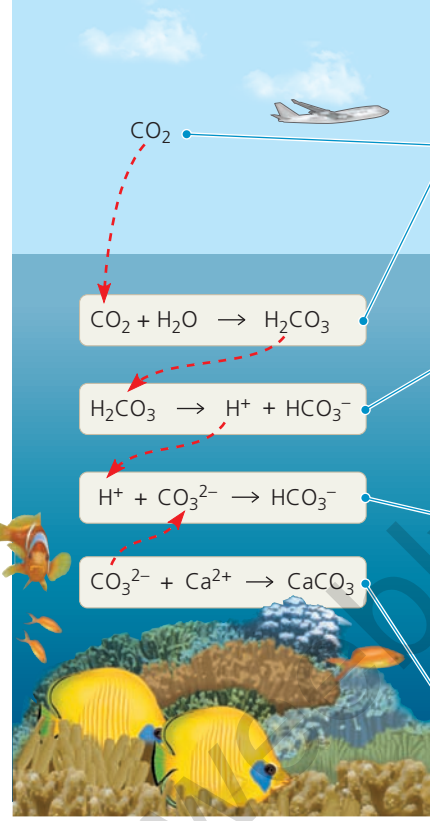
The chemical equilibrium between carbonic acid and bicarbonate acts as a pH regulator, the reaction shifting left or right as other processes in the solution add or remove hydrogen ions. If the H^+ concentration in blood begins to fall (that is, if pH rises), the reaction proceeds to the right and more carbonic acid dissociates, replenishing hydrogen ions. But when the H^+ concentration in blood begins to rise (when pH drops), the reaction proceeds to the left, with HCO_3^- (the base) removing the hydrogen ions from the solution and forming H_2CO_3. Thus, the carbonic acid–bicarbonate buffering system consists of an acid and a base in equilibrium with each other. Most other buffers are also acid-base pairs.

Acidification: A Threat to Water Quality

Among the many threats to water quality posed by human activities is the burning of fossil fuels, which releases gaseous compounds into the atmosphere. When certain of these compounds react with water, the water becomes more acidic, altering the delicate balance of conditions for life on Earth. Carbon dioxide is the main product of fossil fuel combustion. About 25% of human-generated CO_2 is absorbed by the oceans. In spite of the huge volume of water in the oceans, scientists worry that the absorption of so much CO_2 will harm marine ecosystems.

Recent data have shown that such fears are well founded. When CO_2 dissolves in seawater, it reacts with water to form carbonic acid, which lowers ocean pH, a process known as **ocean acidification**. Based on measurements of CO_2 levels in air bubbles trapped in ice over thousands of years, scientists calculate that the pH of the oceans is 0.1 pH unit lower now than at any time in the past 420,000 years. Recent studies predict that it will drop another 0.3–0.5 pH unit by the end of this century.

As seawater acidifies, the extra hydrogen ions combine with carbonate ions (CO_3^{2-}) to form bicarbonate ions (HCO_3^-), thereby reducing the carbonate ion concentration **(Figure 3.11)**. Scientists predict that ocean acidification will cause the carbonate ion concentration to decrease by 40% by the year 2100. This is of great concern because carbonate ions are required for calcification, the production of calcium carbonate ($CaCO_3$) by many marine organisms, including reef-building corals and animals that build shells. The **Scientific Skills Exercise** allows you to work with data from an experiment examining the effect of carbonate ion concentration on coral reefs. Coral reefs are sensitive ecosystems that act as havens for a great diversity of marine life. The disappearance of coral reef ecosystems would be a tragic loss of biological diversity.



▲ Figure 3.11 Atmospheric CO_2 from human activities and its fate in the ocean.

If there is any reason for optimism about the future quality of water resources on our planet, it is that we have made progress in learning about the delicate chemical balances in oceans, lakes, and rivers. Continued progress can come only from the actions of informed individuals, like yourselves, who are concerned about environmental quality. This requires understanding the crucial role that water plays in the suitability of the environment for continued life on Earth.

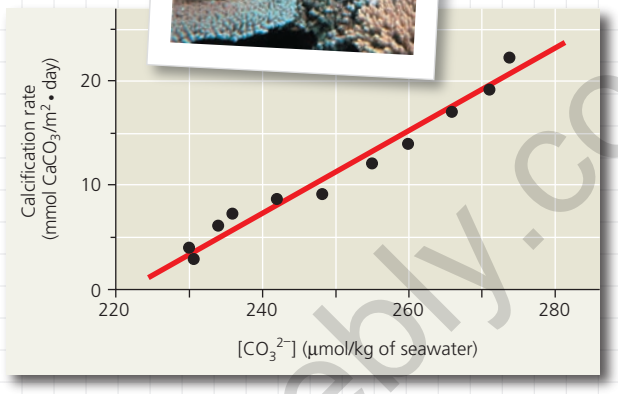
CONCEPT CHECK 3.3

1. Compared with a basic solution at pH 9, the same volume of an acidic solution at pH 4 has _____ times as many hydrogen ions (H^+).
2. HCl is a strong acid that dissociates in water: $HCl \rightarrow H^+ + Cl^-$. What is the pH of 0.01 *M* HCl?
3. Acetic acid (CH_3COOH) can be a buffer, similar to carbonic acid. Write the dissociation reaction, identifying the acid, base, H^+ acceptor, and H^+ donor.
4. **WHAT IF?** Given a liter of pure water and a liter solution of acetic acid, what would happen to the pH if you added 0.01 mol of a strong acid to each? Use the reaction equation from question 3 to explain the result.

For suggested answers, see Appendix A.

SCIENTIFIC SKILLS EXERCISE

Interpreting a Scatter Plot with a Regression Line



How Does the Carbonate Ion Concentration of Seawater Affect the Calcification Rate of a Coral Reef? Scientists predict that acidification of the ocean due to higher levels of atmospheric CO_2 will lower the concentration of dissolved carbonate ions, which living corals use to build calcium carbonate reef structures. In this exercise, you will analyze data from a controlled experiment that examined the effect of carbonate ion concentration ($[CO_3^{2-}]$) on calcium carbonate deposition, a process called calcification.

How the Experiment Was Done The Biosphere 2 aquarium in Arizona contains a large coral reef system that behaves like a natural reef. For several years, a group of researchers measured the rate of calcification by the reef organisms and examined how the calcification rate changed with differing amounts of dissolved carbonate ions in the seawater.

Data from the Experiment The black data points in the graph form a scatter plot. The red line, known as a linear regression line, is the best-fitting straight line for these points.

Interpret the Data

1. When presented with a graph of experimental data, the first step in analysis is to determine what each axis represents. (a) In words, explain what is being shown on the *x*-axis. Be sure to include the units. (b) What is being shown on the *y*-axis (including units)? (c) Which variable is the independent variable—the variable that was *manipulated* by the researchers? (d) Which variable is the dependent variable—the variable that responded to or depended on the treatment, which was *measured* by the researchers? (For additional information about graphs, see the Scientific Skills Review in Appendix F and in the Study Area in MasteringBiology.)
2. Based on the data shown in the graph, describe in words the relationship between carbonate ion concentration and calcification rate.
3. (a) If the seawater carbonate ion concentration is 270 μmol/kg, what is the approximate rate of calcification, and approximately how many days would it take 1 square meter of reef to accumulate 30 mmol of calcium carbonate ($CaCO_3$)? (b) If the seawater carbonate ion concentration is 250 μmol/kg, what is the approximate rate of calcification, and approximately how many days would it take 1 square meter of reef to accumulate 30 mmol of calcium carbonate? (c) If carbonate ion concentration decreases, how does the calcification rate change, and how does that affect the time it takes coral to grow?
4. (a) Referring to the equations in Figure 3.11, determine which step of the process is measured in this experiment. (b) Are the results of this experiment consistent with the hypothesis that increased atmospheric $[CO_2]$ will slow the growth of coral reefs? Why or why not?

MB A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

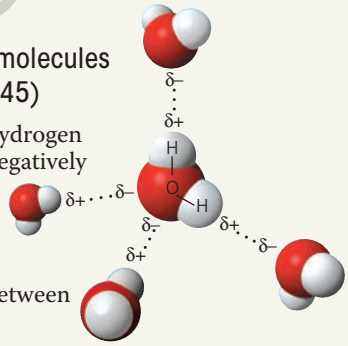
Data from C. Langdon et al., Effect of calcium carbonate saturation state on the calcification rate of an experimental coral reef, *Global Biogeochemical Cycles* 14:639–654 (2000).

3 Chapter Review

SUMMARY OF KEY CONCEPTS

CONCEPT 3.1

Polar covalent bonds in water molecules result in hydrogen bonding (p. 45)



- Water is a **polar molecule**. A hydrogen bond forms when the slightly negatively charged oxygen of one water molecule is attracted to the slightly positively charged hydrogen of a nearby water molecule. Hydrogen bonding between water molecules is the basis for water's properties.

DRAW IT *Label a hydrogen bond and a polar covalent bond in this figure. Is a hydrogen bond a covalent bond? Explain.*

CONCEPT 3.2

Four emergent properties of water contribute to Earth's suitability for life (pp. 45–50)

- Hydrogen bonding keeps water molecules close to each other, and this **cohesion** helps pull water upward in the microscopic water-conducting cells of plants. Hydrogen bonding is also responsible for water's **surface tension**.
- Water has a high **specific heat**: Heat is absorbed when hydrogen bonds break and is released when hydrogen bonds form. This helps keep temperatures relatively steady, within limits that permit life. **Evaporative cooling** is based on water's high **heat of vaporization**. The evaporative loss of the most energetic water molecules cools a surface.
- Ice floats because it is less dense than liquid water. This property allows life to exist under the frozen surfaces of lakes and polar seas.
- Water is an unusually versatile **solvent** because its polar molecules are attracted to ions and polar substances that can form

hydrogen bonds. **Hydrophilic** substances have an affinity for water; **hydrophobic** substances do not. **Molarity**, the number of moles of **solute** per liter of **solution**, is used as a measure of solute concentration in solutions. A **mole** is a certain number of molecules of a substance. The mass of a mole of a substance in grams is the same as the **molecular mass** in daltons.

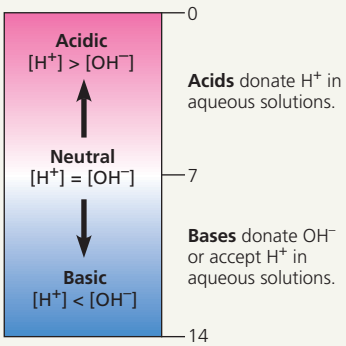
- The emergent properties of water support life on Earth and may contribute to the potential for life to have evolved on other planets.

? *Describe how different types of solutes dissolve in water. Explain what a solution is.*

CONCEPT 3.3

Acidic and basic conditions affect living organisms (pp. 51–54)

- A water molecule can transfer an H^+ to another water molecule to form H_3O^+ (represented simply by H^+) and OH^-.
- The concentration of H^+ is expressed as **pH**; pH = −log $[H^+]$. A **buffer** consists of an acid-base pair that combines reversibly with hydrogen ions, allowing it to resist pH changes.
- The burning of fossil fuels increases the amount of CO_2 in the atmosphere. Some CO_2 dissolves in the oceans, causing **ocean acidification**, which has potentially grave consequences for coral reefs.



? *Explain how increasing amounts of CO_2 dissolving in the ocean leads to ocean acidification. How does this change in pH affect carbonate ion concentration and the rate of calcification?*

TEST YOUR UNDERSTANDING

LEVEL 1: KNOWLEDGE/COMPREHENSION

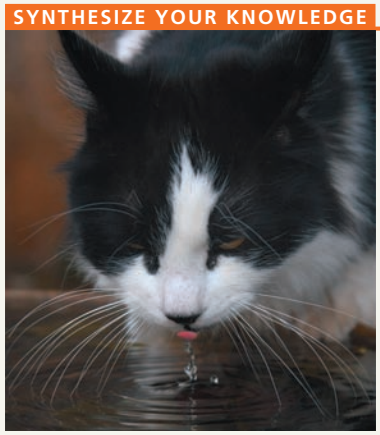
1. Which of the following is a hydrophobic material?
   a. paper
   b. table salt
   c. wax
   d. sugar
2. We can be sure that a mole of table sugar and a mole of vitamin C are equal in their
   a. mass.
   b. volume.
   c. number of atoms.
   d. number of molecules.
3. Measurements show that the pH of a particular lake is 4.0. What is the hydrogen ion concentration of the lake?
   a. 4.0 M
   b. 10^{-10} M
   c. 10^{-4} M
   d. 10^{4} M
4. What is the *hydroxide* ion concentration of the lake described in question 3?
   a. 10^{-10} M
   b. 10^{-4} M
   c. 10^{-7} M
   d. 10.0 M

LEVEL 2: APPLICATION/ANALYSIS

5. A slice of pizza has 500 kcal. If we could burn the pizza and use all the heat to warm a 50-L container of cold water, what would be the approximate increase in the temperature of the water? (*Note*: A liter of cold water weighs about 1 kg.)
   a. 50°C
   b. 5°C
   c. 100°C
   d. 10°C
6. **DRAW IT** Draw the hydration shells that form around a potassium ion and a chloride ion when potassium chloride (KCl) dissolves in water. Label the positive, negative, and partial charges on the atoms.

LEVEL 3: SYNTHESIS/EVALUATION

7. In agricultural areas, farmers pay close attention to the weather forecast. Right before a predicted overnight freeze, farmers spray water on crops to protect the plants. Use the properties of water to explain how this method works. Be sure to mention why hydrogen bonds are responsible for this phenomenon.
8. **EVOLUTION CONNECTION**
   This chapter explains how the emergent properties of water contribute to the suitability of the environment for life. Until fairly recently, scientists assumed that other physical requirements for life included a moderate range of temperature, pH, atmospheric pressure, and salinity, as well as low levels of toxic chemicals. That view has changed with the discovery of organisms known as extremophiles, which have been found flourishing in hot, acidic sulfur springs, around hydrothermal vents deep in the ocean, and in soils with high levels of toxic metals. Why would astrobiologists be interested in studying extremophiles? What does the existence of life in such extreme environments say about the possibility of life on other planets?
9. **SCIENTIFIC INQUIRY**
   Design a controlled experiment to test the hypothesis that water acidification caused by acidic rain would inhibit the growth of *Elodea*, a freshwater plant (see Figure 2.17).
10. **WRITE ABOUT A THEME: ORGANIZATION**
    Several emergent properties of water contribute to the suitability of the environment for life. In a short essay (100–150 words), describe how the ability of water to function as a versatile solvent arises from the structure of water molecules.
11. **SYNTHESIZE YOUR KNOWLEDGE**



How do cats drink? While dogs form their tongues into spoons and scoop water into their mouths, scientists using high-speed video have shown that cats use a different technique to drink aqueous substances like water and milk. Four times a second, the cat touches the tip of its tongue to the water and draws a column of water up into its mouth (as you can see in the photo), which then shuts before gravity can pull the water back down. Describe how the properties of water allow cats to drink in this fashion, including how water's molecular structure contributes to the process.

For selected answers, see Appendix A.

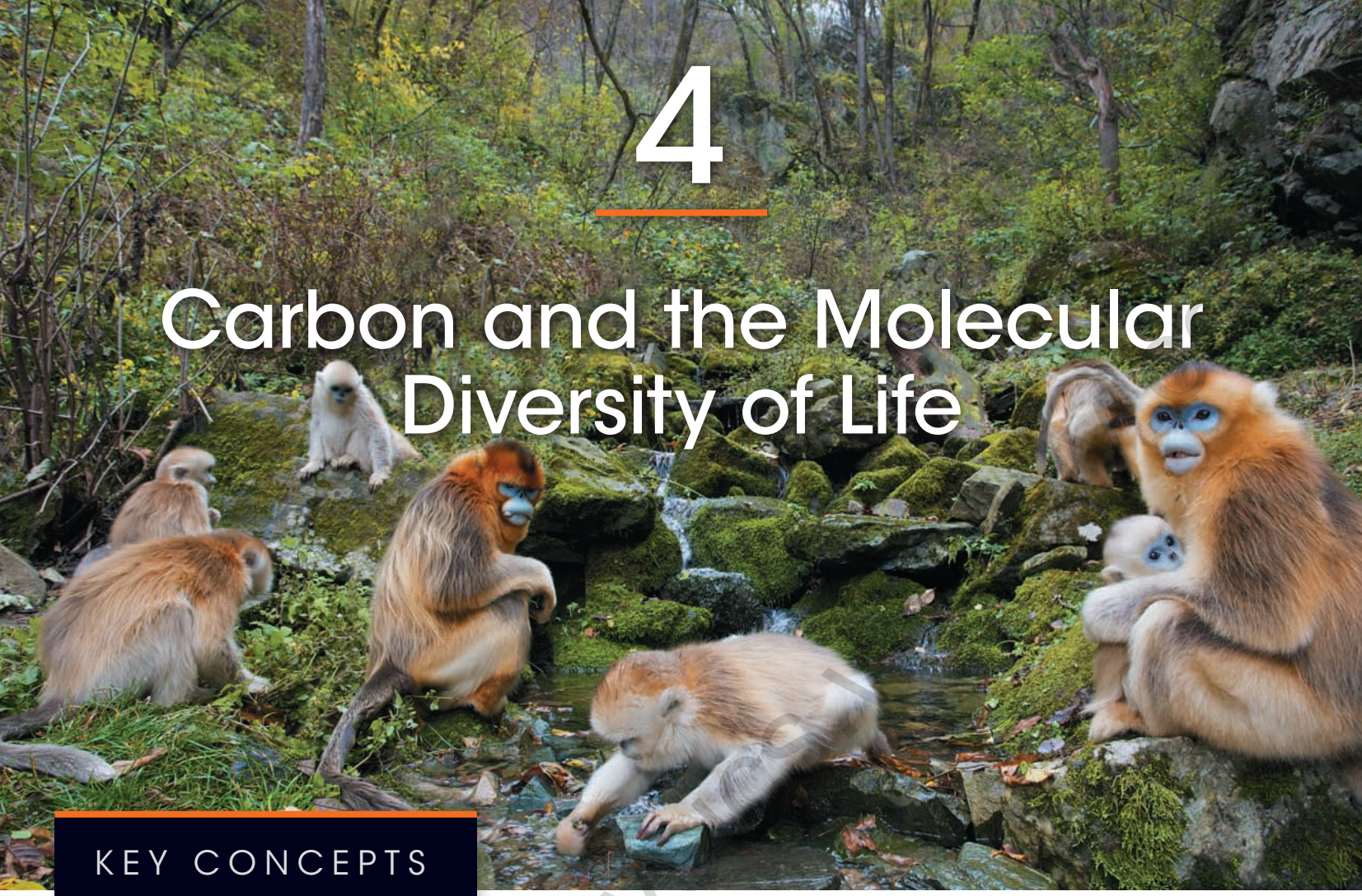
MasteringBiology®

Students Go to **MasteringBiology** for assignments, the eText, and the Study Area with practice tests, animations, and activities.

Instructors Go to **MasteringBiology** for automatically graded tutorials and questions that you can assign to your students, plus Instructor Resources.

4

Carbon and the Molecular Diversity of Life



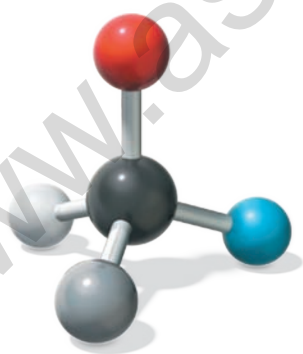
▲ **Figure 4.1** **What properties make carbon the basis of all life?**

KEY CONCEPTS

4.1 Organic chemistry is the study of carbon compounds

4.2 Carbon atoms can form diverse molecules by bonding to four other atoms

4.3 A few chemical groups are key to molecular function



▲ Carbon can bond to four other atoms or groups of atoms, making a large variety of molecules possible.

Carbon: The Backbone of Life

Living organisms, such as the plants and the Qinling golden snub-nosed monkeys shown in **Figure 4.1**, are made up of chemicals based mostly on the element carbon. Carbon enters the biosphere through the action of plants and other photosynthetic organisms. Plants use solar energy to transform atmospheric CO_2 into the molecules of life, which are then taken in by plant-eating animals.

Of all the chemical elements, carbon is unparalleled in its ability to form molecules that are large, complex, and varied, making possible the diversity of organisms that have evolved on Earth. Proteins, DNA, carbohydrates, and other molecules that distinguish living matter from inanimate material are all composed of carbon atoms bonded to one another and to atoms of other elements. Hydrogen (H), oxygen (O), nitrogen (N), sulfur (S), and phosphorus (P) are other common ingredients of these compounds, but it is the element carbon (C) that accounts for the enormous variety of biological molecules.

Large biological molecules, such as proteins, are the main focus of Chapter 5. In this chapter, we investigate the properties of smaller molecules. We will use these small molecules to illustrate concepts of molecular architecture that will help explain why carbon is so important to life, at the same time highlighting the theme that emergent properties arise from the organization of matter in living organisms.

CONCEPT 4.1

Organic chemistry is the study of carbon compounds

For historical reasons, compounds containing carbon are said to be organic, and their study is called **organic chemistry**. By the early 1800s, chemists had learned to make simple compounds in the laboratory by combining elements under the right conditions. Artificial synthesis of the complex molecules extracted from living matter seemed impossible, however. Organic compounds were thought to arise only in living organisms, which were believed to contain a life force beyond the jurisdiction of physical and chemical laws.

Chemists began to chip away at this notion when they learned to synthesize organic compounds in the laboratory. In 1828, Friedrich Wöhler, a German chemist, tried to make an "inorganic" salt, ammonium cyanate, by mixing solutions of ammonium ions (NH_4^+) and cyanate ions (CNO^-). Wöhler was astonished to find that instead he had made urea, an organic compound present in the urine of animals.

The next few decades saw laboratory synthesis of increasingly complex organic compounds, supporting the view that physical and chemical laws govern the processes of life. Organic chemistry was redefined as the study of carbon compounds, regardless of origin. Organic compounds range from simple molecules, such as methane (CH_4), to colossal ones, such as proteins, with thousands of atoms.

Organic Molecules and the Origin of Life on Earth

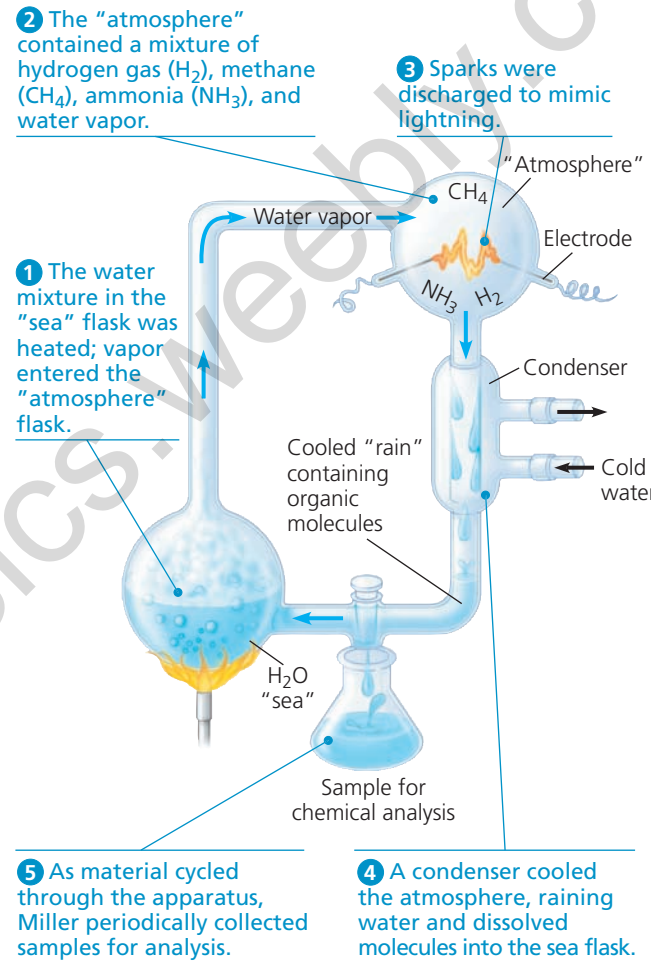
EVOLUTION In 1953, Stanley Miller, a graduate student of Harold Urey's at the University of Chicago, helped bring the abiotic (nonliving) synthesis of organic compounds into the context of evolution. Study **Figure 4.2** to learn about his classic experiment. From his results, Miller concluded that complex organic molecules could arise spontaneously under conditions thought at that time to have existed on the early Earth. You can work with the data from a related experiment in the **Scientific Skills Exercise**. These experiments support the idea that abiotic synthesis of organic compounds, perhaps near volcanoes, could have been an early stage in the origin of life (see Chapter 25).

The overall percentages of the major elements of life—C, H, O, N, S, and P—are quite uniform from one organism to another, reflecting the common evolutionary origin of all life. Because of carbon's ability to form four bonds, however, this limited assortment of atomic building blocks can be used to build an inexhaustible variety of organic molecules. Different species of organisms, and different individuals within a species, are distinguished by variations in the types

▼ Figure 4.2 Inquiry

Can organic molecules form under conditions estimated to simulate those on the early Earth?

Experiment In 1953, Stanley Miller set up a closed system to mimic conditions thought at that time to have existed on the early Earth. A flask of water simulated the primeval sea. The water was heated so that some vaporized and moved into a second, higher flask containing the "atmosphere"—a mixture of gases. Sparks were discharged in the synthetic atmosphere to mimic lightning.



Results Miller identified a variety of organic molecules that are common in organisms. These included simple compounds, such as formaldehyde (CH_2O) and hydrogen cyanide (HCN), and more complex molecules, such as amino acids and long chains of carbon and hydrogen known as hydrocarbons.

Conclusion Organic molecules, a first step in the origin of life, may have been synthesized abiotically on the early Earth. Although new evidence indicates that the early Earth's atmosphere was different from the "atmosphere" used by Miller in this experiment, recent experiments using the revised list of chemicals also produced organic molecules. (We will explore this hypothesis in more detail in Chapter 25.)

Source: S. L. Miller, A production of amino acids under possible primitive Earth conditions, *Science* 117:528–529 (1953).

WHAT IF? *If Miller had increased the concentration of NH_3 in his experiment, how might the relative amounts of the products HCN and CH_2O have differed?*

SCIENTIFIC SKILLS EXERCISE

Working with Moles and Molar Ratios

Could the First Biological Molecules Have Formed Near Volcanoes on Early Earth? In 2007, Jeffrey Bada, a former graduate student of Stanley Miller's, discovered some vials of samples that had never been analyzed from an experiment performed by Miller in 1958. In this experiment, Miller used hydrogen sulfide gas (H_2S) as one of the gases in the reactant mixture. Since H_2S is released by volcanoes, the H_2S experiment was designed to mimic conditions near volcanoes on early Earth. In 2011, Bada and colleagues published the results of their analysis of these "lost" samples. In this exercise, you will make calculations using the molar ratios of reactants and products from the H_2S experiment.

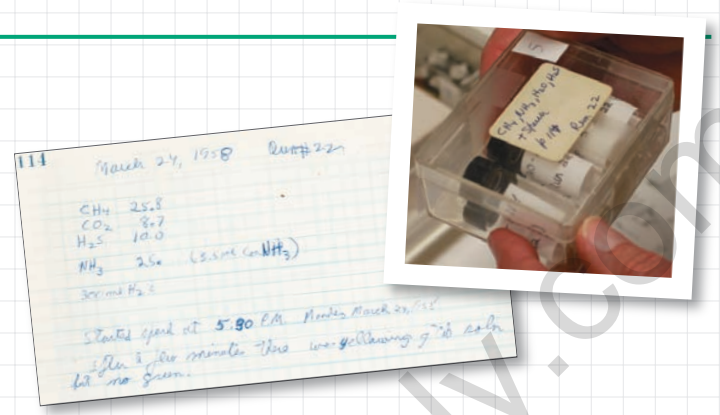
How the Experiment Was Done According to his laboratory notebook, Miller used the same apparatus as in his original experiment (see Figure 4.2), but the mixture of gaseous reactants included methane (CH_4), carbon dioxide (CO_2), hydrogen sulfide (H_2S), and ammonia (NH_3). After three days of simulated volcanic activity, he collected samples of the liquid, partially purified the chemicals, and sealed the samples in sterile vials. In 2011, Bada's research team used modern analytical methods to analyze the products in the vials for the presence of amino acids, the building blocks of proteins.

Data from the Experiment The table below shows 4 of the 23 amino acids detected in the samples from Miller's 1958 H_2S experiment.

Product Compound	Molecular Formula	Molar Ratio (Relative to Glycine)
Glycine	$C_2H_5NO_2$	1.0
Serine	$C_3H_7NO_3$	3.0×10^{-2}
Methionine	$C_5H_{11}NO_2S$	1.8×10^{-3}
Alanine	$C_3H_7NO_2$	1.1

Interpret the Data

1. A *mole* is the number of grams of a substance that equals its molecular (or atomic) mass in daltons. There are 6.02×10^{23} molecules (or atoms) in 1.0 mole (Avogadro's number; see Concept 3.2). The data table shows the "molar ratios" of some of the products from the Miller H_2S experiment. In a molar ratio, each unitless value is expressed relative to a standard for that experiment. Here, the standard is the number of moles of the amino acid glycine, which is set to a value of 1.0. For instance, serine has a molar ratio of 3.0×10^{-2}, meaning that for every mole of glycine, there is 3.0×10^{-2} mole of serine. (a) Give the molar ratio of methionine to glycine and explain what it means. (b) How many molecules of glycine are present in 1.0 mole? (c) For every 1.0 mole of glycine in the sample, how many molecules of methionine are present? (Recall that to multiply two numbers with exponents, you add their exponents; to divide them, you subtract the exponent in the denominator from that in the numerator.)
2. (a) Which amino acid is present in higher amounts than glycine? (b) How many more molecules of that amino acid are present than the number of molecules in 1.0 mole of glycine?
3. The synthesis of products is limited by the amount of reactants. (a) If one mole each of CH_4, NH_3, H_2S, and CO_2 is added to 1 liter of water (= 55.5 moles of H_2O) in a flask, how many moles of hydrogen, carbon, oxygen, nitrogen, and sulfur are in the flask? (b) Looking at the molecular formula in the table, how many moles of each element would be needed to make 1.0 mole of glycine? (c) What is the maximum number of moles of glycine that could be made in that flask, with the specified ingredients, if no other molecules were made? Explain. (d) If serine or methionine were made individually, which element(s) would be used up first for each? How much of each product could be made?
4. The earlier published experiment carried out by Miller did not include H_2S in the reactants (see Figure 4.2). Which of the compounds shown in the data table can be made in the H_2S experiment but could not be made in the earlier experiment?



▲ Some of Stanley Miller's notes from his 1958 hydrogen sulfide (H_2S) experiment along with his original vials.

MB A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

Data from E. T. Parker et al., Primordial synthesis of amines and amino acids in a 1958 Miller H_2S-rich spark discharge experiment, *Proceedings of the National Academy of Sciences USA* 108:5526-5531 (2011). www.pnas.org/cgi/doi/10.1073/pnas.1019191108.

of organic molecules they make. In a sense, the great diversity of living organisms we see on the planet (and in fossil remains) is made possible by the unique chemical versatility of the element carbon.

CONCEPT CHECK 4.1

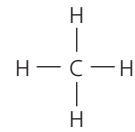
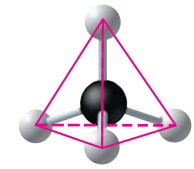
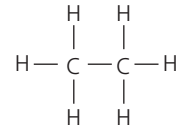
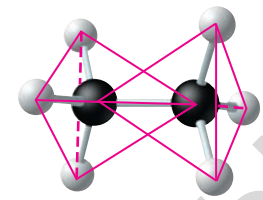
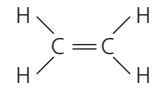
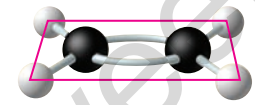
1. Why was Wöhler astonished to find he had made urea?
2. **WHAT IF?** Miller carried out a control experiment without the electrical discharge and found no organic compounds. What might explain this result?

For suggested answers, see Appendix A.

CONCEPT 4.2

Carbon atoms can form diverse molecules by bonding to four other atoms

The key to an atom's chemical characteristics is its electron configuration. This configuration determines the kinds and number of bonds an atom will form with other atoms. Recall that it is the valence electrons, those in the outermost shell, that are available to form bonds with other atoms.

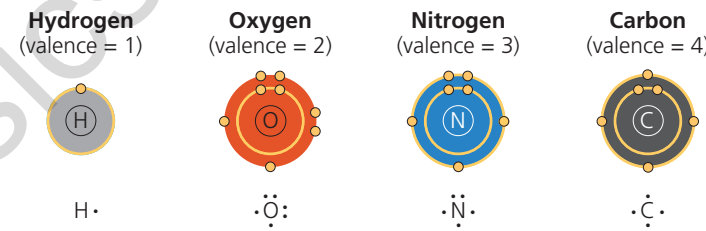
Molecule and Molecular Shape	Molecular Formula	Structural Formula	Ball-and-Stick Model (molecular shape in pink)	Space-Filling Model
(a) Methane. When a carbon atom has four single bonds to other atoms, the molecule is tetrahedral.	CH_4			
(b) Ethane. A molecule may have more than one tetrahedral group of single-bonded atoms. (Ethane consists of two such groups.)	C_2H_6			
(c) Ethene (ethylene). When two carbon atoms are joined by a double bond, all atoms attached to those carbons are in the same plane, and the molecule is flat.	C_2H_4			

▲ **Figure 4.3** **The shapes of three simple organic molecules.**

The Formation of Bonds with Carbon

Carbon has 6 electrons, with 2 in the first electron shell and 4 in the second shell; thus, it has 4 valence electrons in a shell that can hold up to 8 electrons. A carbon atom usually completes its valence shell by sharing its 4 electrons with other atoms so that 8 electrons are present. Each pair of shared electrons constitutes a covalent bond (see Figure 2.10d). In organic molecules, carbon usually forms single or double covalent bonds. Each carbon atom acts as an intersection point from which a molecule can branch off in as many as four directions. This enables carbon to form large, complex molecules.

When a carbon atom forms four single covalent bonds, the arrangement of its four hybrid orbitals causes the bonds to angle toward the corners of an imaginary tetrahedron. The bond angles in methane (CH_4) are 109.5° **(Figure 4.3a)**, and they are roughly the same in any group of atoms where carbon has four single bonds. For example, ethane (C_2H_6) is shaped like two overlapping tetrahedrons **(Figure 4.3b)**. In molecules with more carbons, every grouping of a carbon bonded to four other atoms has a tetrahedral shape. But when two carbon atoms are joined by a double bond, as in ethene (C_2H_4), the bonds from both carbons are all in the same plane, so the atoms joined to those carbons are in the same plane as well **(Figure 4.3c)**. We find it convenient to write molecules as structural formulas, as if the molecules being represented are two-dimensional, but keep in mind that molecules are three-dimensional and that the shape of a molecule is central to its function.



▲ **Figure 4.4** **Valences of the major elements of organic molecules.** Valence is the number of covalent bonds an atom can form. It is generally equal to the number of electrons required to complete the valence (outermost) shell (see Figure 2.7). All the electrons are shown for each atom in the electron distribution diagrams (top). Only the valence shell electrons are shown in the Lewis dot structures (bottom). Note that carbon can form four bonds.

MAKE CONNECTIONS *Draw the Lewis dot structures for sodium, phosphorus, sulfur, and chlorine. (Refer to Figure 2.7.)*

The electron configuration of carbon gives it covalent compatibility with many different elements. **Figure 4.4** shows the valences of carbon and its most frequent bonding partners—hydrogen, oxygen, and nitrogen. These are the four major atomic components of organic molecules. These valences are the basis for the rules of covalent bonding in organic chemistry—the building code for the architecture of organic molecules.

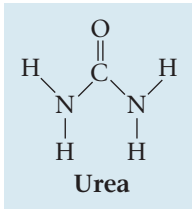
How do the rules of covalent bonding apply to carbon atoms with partners other than hydrogen? We'll look at two examples, the simple molecules carbon dioxide and urea.

In the carbon dioxide molecule (CO_2), a single carbon atom is joined to two atoms of oxygen by double covalent bonds. The structural formula for CO_2 is shown here:

$$O{=}C{=}O$$

Each line in a structural formula represents a pair of shared electrons. Thus, the two double bonds in CO_2 have the same number of shared electrons as four single bonds. The arrangement completes the valence shells of all atoms in the molecule. Because CO_2 is a very simple molecule and lacks hydrogen, it is often considered inorganic, even though it contains carbon. Whether we call CO_2 organic or inorganic, however, it is clearly important to the living world as the source of carbon for all organic molecules in organisms.

Urea, $CO(NH_2)_2$, is the organic compound found in urine that Wöhler synthesized in the early 1800s. Again, each atom has the required number of covalent bonds. In this case, one carbon atom participates in both single and double bonds.



Urea

Urea and carbon dioxide are molecules with only one carbon atom. But as Figure 4.3 shows, a carbon atom can also use one or more valence electrons to form covalent bonds to other carbon atoms, each of which can also form four bonds. Thus, the atoms can be linked into chains of seemingly infinite variety.

Molecular Diversity Arising from Variation in Carbon Skeletons

Carbon chains form the skeletons of most organic molecules. The skeletons vary in length and may be straight, branched, or arranged in closed rings **(Figure 4.5)**. Some carbon skeletons have double bonds, which vary in number and location. Such variation in carbon skeletons is one important source of the molecular complexity and diversity that characterize living matter. In addition, atoms of other elements can be bonded to the skeletons at available sites.

Hydrocarbons

All of the molecules that are shown in Figures 4.3 and 4.5 are **hydrocarbons**, organic molecules consisting of only carbon and hydrogen. Atoms of hydrogen are attached to the carbon skeleton wherever electrons are available for covalent bonding. Hydrocarbons are the major components of petroleum, which is called a fossil fuel because it consists of the partially decomposed remains of organisms that lived millions of years ago.

Although hydrocarbons are not prevalent in most living organisms, many of a cell's organic molecules have regions consisting of only carbon and hydrogen. For example, the molecules known as fats have long hydrocarbon tails attached to a nonhydrocarbon component **(Figure 4.6)**. Neither petroleum nor fat dissolves in water; both are hydrophobic compounds because the great majority of their bonds are relatively nonpolar carbon-to-hydrogen linkages. Another characteristic of hydrocarbons is that they can undergo reactions that release a relatively large amount of energy. The gasoline that fuels a car consists of hydrocarbons, and the hydrocarbon tails of fats serve as stored fuel for plant embryos (seeds) and animals.

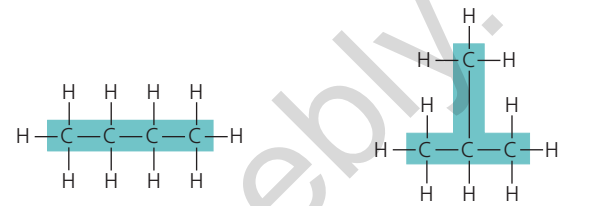
▼ Figure 4.5 Four ways that carbon skeletons can vary.

(a) Length


Ethane Propane

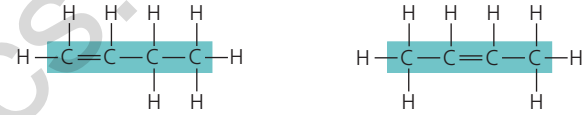
Carbon skeletons vary in length.

(b) Branching


Butane 2-Methylpropane (commonly called isobutane)

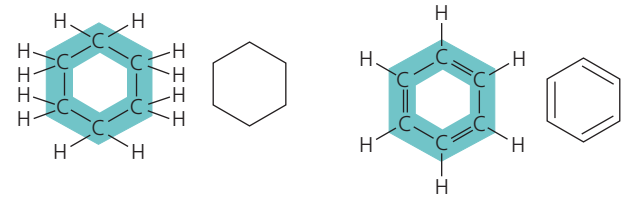
Skeletons may be unbranched or branched.

(c) Double bond position

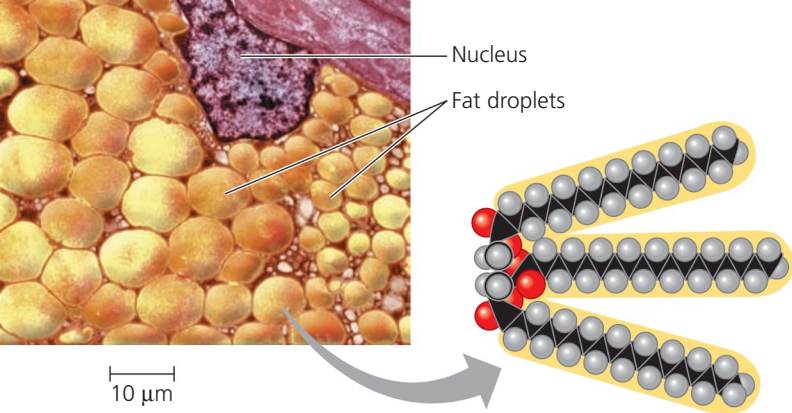

1-Butene 2-Butene

The skeleton may have double bonds, which can vary in location.

(d) Presence of rings


Cyclohexane Benzene

Some carbon skeletons are arranged in rings. In the abbreviated structural formula for each compound (at the right), each corner represents a carbon and its attached hydrogens.



(a) Part of a human adipose cell (b) A fat molecule

▲ **Figure 4.6 The role of hydrocarbons in fats. (a)** Mammalian adipose cells stockpile fat molecules as a fuel reserve. This colorized micrograph shows part of a human adipose cell with many fat droplets, each containing a large number of fat molecules. **(b)** A fat molecule consists of a small, nonhydrocarbon component joined to three hydrocarbon tails that account for the hydrophobic behavior of fats. The tails can be broken down to provide energy. (Black = carbon; gray = hydrogen; red = oxygen.)

MAKE CONNECTIONS *How do the tails account for the hydrophobic nature of fats? (See Concept 3.2.)*

Isomers

Variation in the architecture of organic molecules can be seen in **isomers**, compounds that have the same numbers of atoms of the same elements but different structures and hence different properties. We will examine three types of isomers: structural isomers, *cis-trans* isomers, and enantiomers.

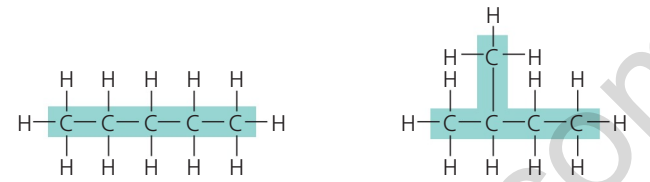
Structural isomers differ in the covalent arrangements of their atoms. Compare, for example, the two five-carbon compounds in **Figure 4.7a**. Both have the molecular formula C_5H_{12}, but they differ in the covalent arrangement of their carbon skeletons. The skeleton is straight in one compound but branched in the other. The number of possible isomers increases tremendously as carbon skeletons increase in size. There are only three forms of C_5H_{12} (two of which are shown in Figure 4.7a), but there are 18 variations of C_8H_{18} and 366,319 possible structural isomers of $C_{20}H_{42}$. Structural isomers may also differ in the location of double bonds.

In ***cis-trans* isomers** (formerly called *geometric isomers*), carbons have covalent bonds to the same atoms, but these atoms differ in their spatial arrangements due to the inflexibility of double bonds. Single bonds allow the atoms they join to rotate freely about the bond axis without changing the compound. In contrast, double bonds do not permit such rotation. If a double bond joins two carbon atoms, and each C also has two different atoms (or groups of atoms) attached to it, then two distinct *cis-trans* isomers are possible. Consider a simple molecule with two double-bonded carbons, each of which has an H and an X attached to it **(Figure 4.7b)**. The arrangement with both Xs on the same side of the double bond is called a *cis isomer*, and that with the Xs on opposite sides is called a *trans isomer*. The subtle difference in shape between such isomers can dramatically affect the biological activities of organic molecules. For example, the biochemistry of vision involves a light-induced change of retinal, a chemical compound in the eye, from the *cis* isomer to the *trans* isomer (see Figure 50.17). Another example involves *trans* fats, which are discussed in Chapter 5.

Enantiomers are isomers that are mirror images of each other and that differ in shape due to the presence of an *asymmetric carbon*, one that is attached to four different atoms or groups of atoms. (See the middle carbon in

▼ **Figure 4.7 Three types of isomers, compounds with the same molecular formula but different structures.**

(a) Structural isomers



Structural isomers differ in covalent partners, as shown in this example of two isomers of C_5H_{12}.

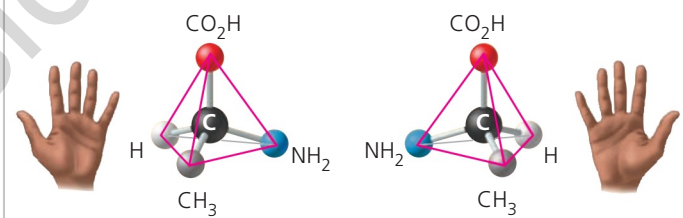
(b) *Cis-trans* isomers

cis isomer: The two Xs are on the same side.

trans isomer: The two Xs are on opposite sides.

Cis-trans isomers differ in arrangement about a double bond. In these diagrams, X represents an atom or group of atoms attached to a double-bonded carbon.

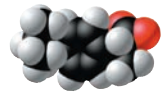
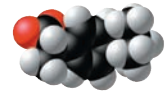
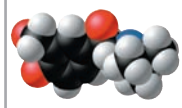
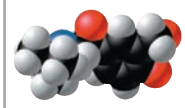
(c) Enantiomers



L isomer D isomer

Enantiomers differ in spatial arrangement around an asymmetric carbon, resulting in molecules that are mirror images, like left and right hands. The two isomers here are designated the L and D isomers from the Latin for "left" and "right" (*levo* and *dextro*). Enantiomers cannot be superimposed on each other.

DRAW IT *There are three structural isomers of C_5H_{12}; draw the one not shown in (a).*

Drug	Effects	Effective Enantiomer	Ineffective Enantiomer
Ibuprofen	Reduces inflammation and pain	 *S*-Ibuprofen	 *R*-Ibuprofen
Albuterol	Relaxes bronchial (airway) muscles, improving airflow in asthma patients	 *R*-Albuterol	 *S*-Albuterol

▲ Figure 4.8 The pharmacological importance of enantiomers. Ibuprofen and albuterol are drugs whose enantiomers have different effects. (*S* and *R* are used here to distinguish between enantiomers.) Ibuprofen is commonly sold as a mixture of the two enantiomers; the *S* enantiomer is 100 times more effective than the *R* form. Albuterol is synthesized and sold only as the *R* form of the drug; the *S* form counteracts the active *R* form.

the ball-and-stick models shown in **Figure 4.7c**.) The four groups can be arranged in space around the asymmetric carbon in two different ways that are mirror images. Enantiomers are, in a way, left-handed and right-handed versions of the molecule. Just as your right hand won't fit into a left-handed glove, a "right-handed" molecule won't fit into the same space as the "left-handed" version. Usually, only one isomer is biologically active because only that form can bind to specific molecules in an organism.

The concept of enantiomers is important in the pharmaceutical industry because the two enantiomers of a drug may not be equally effective, as is the case for both ibuprofen and the asthma medication albuterol **(Figure 4.8)**. Methamphetamine also occurs in two enantiomers that have very different effects. One enantiomer is the highly addictive stimulant drug known as "crank," sold illegally in the street drug trade. The other has a much weaker effect and is the active ingredient in an over-the-counter vapor inhaler for treatment of nasal congestion. The differing effects of enantiomers in the body demonstrate that organisms are sensitive to even the most subtle variations in molecular architecture. Once again, we see that molecules have emergent properties that depend on the specific arrangement of their atoms.

CONCEPT CHECK 4.2

1. **DRAW IT** (a) Draw a structural formula for C_2H_4. (b) Draw the *trans* isomer of $C_2H_2Cl_2$.
2. Which molecules in Figure 4.5 are isomers? For each pair, identify the type of isomer.
3. How are gasoline and fat chemically similar?
4. Can propane (C_3H_8) form isomers? Explain.

For suggested answers, see Appendix A.

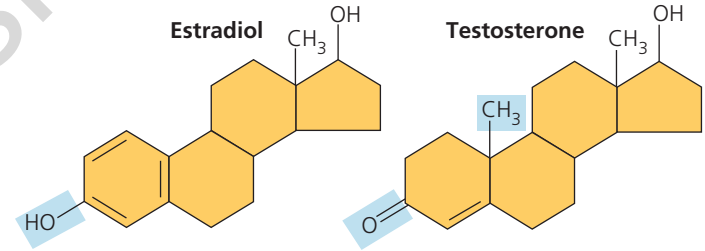
CONCEPT 4.3

A few chemical groups are key to molecular function

The properties of an organic molecule depend not only on the arrangement of its carbon skeleton but also on the chemical groups attached to that skeleton. We can think of hydrocarbons, the simplest organic molecules, as the underlying framework for more complex organic molecules. A number of chemical groups can replace one or more hydrogens of the hydrocarbon. These groups may participate in chemical reactions or may contribute to function indirectly by their effects on molecular shape; they help give each molecule its unique properties.

The Chemical Groups Most Important in the Processes of Life

Consider the differences between estradiol (a type of estrogen) and testosterone. These compounds are female and male sex hormones, respectively, in humans and other vertebrates. Both are steroids, organic molecules with a common carbon skeleton in the form of four fused rings. They differ only in the chemical groups attached to the rings (shown here in abbreviated form); the distinctions in molecular architecture are shaded in blue:

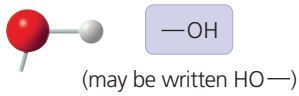
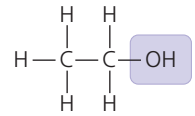
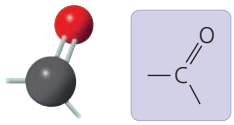
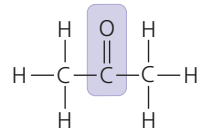
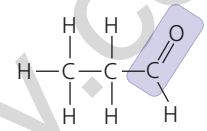
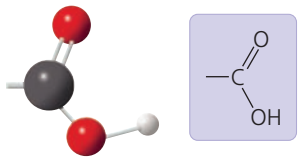
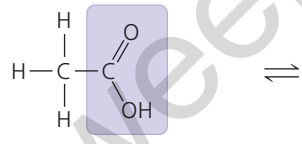
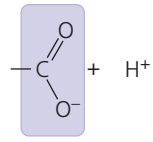
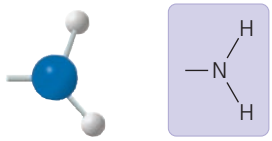
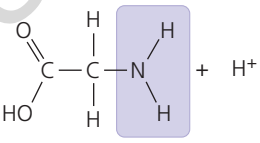
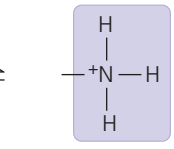
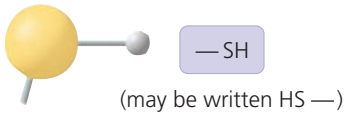
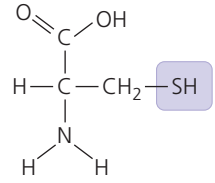
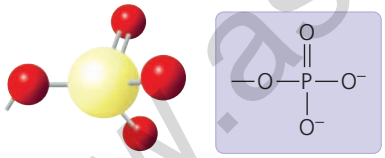
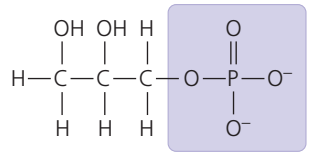
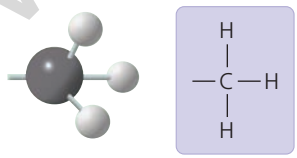
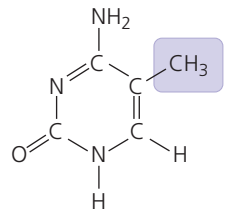


The different actions of these two molecules on many targets throughout the body are the basis of gender, producing the contrasting features of male and female vertebrates. In this case, the chemical groups are important because they affect molecular shape, contributing to function.

In other cases, chemical groups are directly involved in chemical reactions; such groups are known as **functional groups**. Each has certain properties, such as shape and charge, that cause it to participate in chemical reactions in a characteristic way.

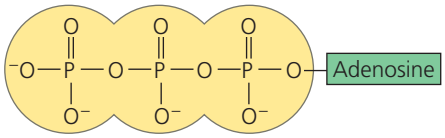
The seven chemical groups most important in biological processes are the hydroxyl, carbonyl, carboxyl, amino, sulfhydryl, phosphate, and methyl groups. The first six groups can be chemically reactive; of these, all except the sulfhydryl group are also hydrophilic and thus increase the solubility of organic compounds in water. The methyl group is not reactive, but instead often serves as a recognizable tag on biological molecules. Study **Figure 4.9** to become familiar with these biologically important chemical groups.

▼ Figure 4.9 Some biologically important chemical groups.

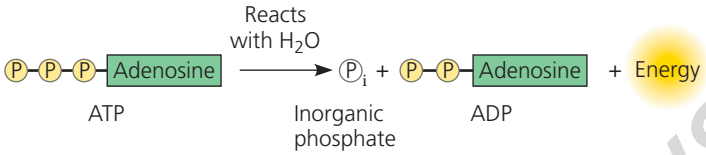
Chemical Group	Group Properties and Compound Name	Examples
Hydroxyl group (—OH)  (may be written HO—)	Is polar due to electronegative oxygen. Forms hydrogen bonds with water, helping dissolve compounds such as sugars. Compound name: Alcohol (specific name usually ends in *-ol*)	 **Ethanol**, the alcohol present in alcoholic beverages
Carbonyl group ($>C=O$) 	Sugars with ketone groups are called ketoses; those with aldehydes are called aldoses. Compound name: Ketone (carbonyl group is within a carbon skeleton) or aldehyde (carbonyl group is at the end of a carbon skeleton)	 **Acetone**, the simplest ketone  **Propanal**, an aldehyde
Carboxyl group (—COOH) 	Acts as an acid (can donate H^+) because the covalent bond between oxygen and hydrogen is so polar. Compound name: Carboxylic acid, or organic acid	 **Acetic acid**, which gives vinegar its sour taste  Ionized form of —COOH (carboxylate ion), found in cells
Amino group ($—NH_2$) 	Acts as a base; can pick up an H^+ from the surrounding solution (water, in living organisms). Compound name: Amine	 **Glycine**, an amino acid (note its carboxyl group)  Ionized form of $—NH_2$, found in cells
Sulfhydryl group (—SH)  (may be written HS —)	Two —SH groups can react, forming a "cross-link" that helps stabilize protein structure. Hair protein cross-links maintain the straightness or curliness of hair; in hair salons, permanent treatments break cross-links, then re-form them while the hair is in the desired shape. Compound name: Thiol	 **Cysteine**, a sulfur-containing amino acid
Phosphate group ($—OPO_3^{2-}$) 	Contributes negative charge (1– when positioned inside a chain of phosphates; 2– when at the end). When attached, confers on a molecule the ability to react with water, releasing energy. Compound name: Organic phosphate	 **Glycerol phosphate**, which takes part in many important chemical reactions in cells
Methyl group ($—CH_3$) 	Affects the expression of genes when on DNA or on proteins bound to DNA. Affects the shape and function of male and female sex hormones. Compound name: Methylated compound	 **5-Methyl cytosine**, a component of DNA that has been modified by addition of a methyl group

ATP: An Important Source of Energy for Cellular Processes

The "Phosphate group" row in Figure 4.9 shows a simple example of an organic phosphate molecule. A more complicated organic phosphate, **adenosine triphosphate**, or **ATP**, is worth mentioning here because its function in the cell is so important. ATP consists of an organic molecule called adenosine attached to a string of three phosphate groups:



Where three phosphates are present in series, as in ATP, one phosphate may be split off as a result of a reaction with water. This inorganic phosphate ion, $HOPO_3^{2-}$, is often abbreviated $\textcircled{P}_i$ in this book, and a phosphate group in an organic molecule is often written as $\textcircled{P}$. Having lost one phosphate, ATP becomes adenosine *di*phosphate, or ADP. Although ATP is sometimes said to store energy, it is more accurate to think of it as storing the potential to react with water. This reaction releases energy that can be used by the cell. You will learn about this in more detail in Chapter 8.



CONCEPT CHECK 4.3

1. What does the term *amino acid* signify about the structure of such a molecule?
2. What chemical change occurs to ATP when it reacts with water and releases energy?
3. **WHAT IF?** Suppose you had an organic molecule such as cysteine (see Figure 4.9, sulfhydryl group example), and you chemically removed the $—NH_2$ group and replaced it with —COOH. Draw the structural formula for this molecule and speculate about its chemical properties. Is the central carbon asymmetric before the change? After?

For suggested answers, see Appendix A.

The Chemical Elements of Life: *A Review*

Living matter, as you have learned, consists mainly of carbon, oxygen, hydrogen, and nitrogen, with smaller amounts of sulfur and phosphorus. These elements all form strong covalent bonds, an essential characteristic in the architecture of complex organic molecules. Of all these elements, carbon is the virtuoso of the covalent bond. The versatility of carbon makes possible the great diversity of organic molecules, each with particular properties that emerge from the unique arrangement of its carbon skeleton and the chemical groups appended to that skeleton. This variation at the molecular level provides the foundation for the rich biological diversity found on our planet.

4 Chapter Review

SUMMARY OF KEY CONCEPTS

CONCEPT 4.1

Organic chemistry is the study of carbon compounds (pp. 57–58)

- Organic compounds, once thought to arise only within living organisms, were finally synthesized in the laboratory.
- Living matter is made mostly of carbon, oxygen, hydrogen, and nitrogen. Biological diversity results from carbon's ability to form a huge number of molecules with particular shapes and properties.

? *How did Stanley Miller's experiments support the idea that, even at life's origins, physical and chemical laws govern the processes of life?*

CONCEPT 4.2

Carbon atoms can form diverse molecules by bonding to four other atoms (pp. 58–62)

- Carbon, with a valence of 4, can bond to various other atoms, including O, H, and N. Carbon can also bond to other carbon atoms, forming the carbon skeletons of organic compounds. These skeletons vary in length and shape and have bonding sites for atoms of other elements.
- **Hydrocarbons** consist of carbon and hydrogen.
- **Isomers** are compounds that have the same molecular formula but different structures and therefore different properties. Three types of isomers are **structural isomers**, ***cis-trans* isomers**, and **enantiomers**.

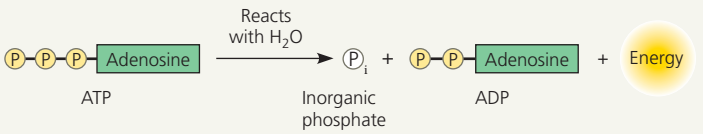
? *Refer back to Figure 4.9. What type of isomers are acetone and propanal? How many asymmetric carbons are present in acetic acid, glycine, and glycerol phosphate? Can these three molecules exist as forms that are enantiomers?*

CONCEPT 4.3

A few chemical groups are key to molecular function (pp. 62–64)

- Chemical groups attached to the carbon skeletons of organic molecules participate in chemical reactions (**functional groups**) or contribute to function by affecting molecular shape (see Figure 4.9).
- **ATP (adenosine triphosphate)** consists of adenosine attached to three phosphate groups. ATP can react with water, forming

inorganic phosphate and ADP (adenosine diphosphate). This reaction releases energy that can be used by the cell.

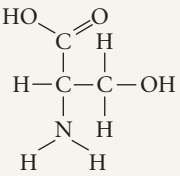


? *In what ways does a methyl group differ chemically from the other six important chemical groups shown in Figure 4.9?*

TEST YOUR UNDERSTANDING

LEVEL 1: KNOWLEDGE/COMPREHENSION

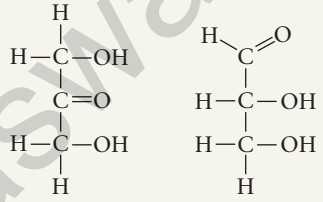
1. Organic chemistry is currently defined as
   a. the study of compounds made only by living cells.
   b. the study of carbon compounds.
   c. the study of natural (as opposed to synthetic) compounds.
   d. the study of hydrocarbons.
2. Which functional group is *not* present in this molecule?
   a. carboxyl
   b. sulfhydryl
   c. hydroxyl
   d. amino



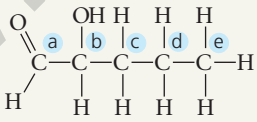
3. **MAKE CONNECTIONS** Which chemical group is most likely to be responsible for an organic molecule behaving as a base (see Concept 3.3)?
   a. hydroxyl
   b. carbonyl
   c. amino
   d. phosphate

LEVEL 2: APPLICATION/ANALYSIS

4. Which of the following hydrocarbons has a double bond in its carbon skeleton?
   a. C_3H_8
   b. C_2H_6
   c. C_2H_4
   d. C_2H_2
5. Choose the term that correctly describes the relationship between these two sugar molecules:
   a. structural isomers
   b. *cis-trans* isomers
   c. enantiomers
   d. isotopes



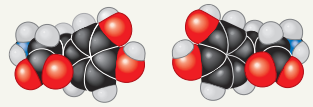
6. Identify the asymmetric carbon in this molecule:



7. Which action could produce a carbonyl group?
   a. the replacement of the —OH of a carboxyl group with hydrogen
   b. the addition of a thiol to a hydroxyl
   c. the addition of a hydroxyl to a phosphate
   d. the replacement of the nitrogen of an amine with oxygen
8. Which of the molecules shown in question 5 has an asymmetric carbon? Which carbon is asymmetric?

LEVEL 3: SYNTHESIS/EVALUATION

9. **EVOLUTION CONNECTION**
   DRAW IT Some scientists think that life elsewhere in the universe might be based on the element silicon, rather than on carbon, as on Earth. Look at the electron distribution diagram for silicon in Figure 2.7 and draw the Lewis dot structure for silicon. What properties does silicon share with carbon that would make silicon-based life more likely than, say, neon-based life or aluminum-based life?
10. **SCIENTIFIC INQUIRY**
    50 years ago, pregnant women who were prescribed thalidomide for morning sickness gave birth to children with birth defects. Thalidomide is a mixture of two enantiomers; one reduces morning sickness, but the other causes severe birth defects. Today, the FDA has approved this drug for non-pregnant individuals with Hansen's disease (leprosy) or newly diagnosed multiple myeloma, a blood and bone marrow cancer. The beneficial enantiomer can be synthesized and given to patients, but over time, both the beneficial *and* the harmful enantiomer can be detected in the body. Propose a possible explanation for the presence of the harmful enantiomer.
11. **WRITE ABOUT A THEME: ORGANIZATION**
    In 1918, an epidemic of sleeping sickness caused an unusual rigid paralysis in some survivors, similar to symptoms of advanced Parkinson's disease. Years later, L-dopa (below, left), a chemical used to treat Parkinson's disease, was given to some of these patients. L-dopa was remarkably effective at eliminating the paralysis, at least temporarily. However, its enantiomer, D-dopa (right), was subsequently shown to have no effect at all, as is the case for Parkinson's disease. In a short essay (100–150 words), discuss how the effectiveness of one enantiomer and not the other illustrates the theme of structure and function.



12. **SYNTHESIZE YOUR KNOWLEDGE**

Explain how the chemical structure of the carbon atom accounts for the differences between the male and female lions seen in the photo.

For selected answers, see Appendix A.

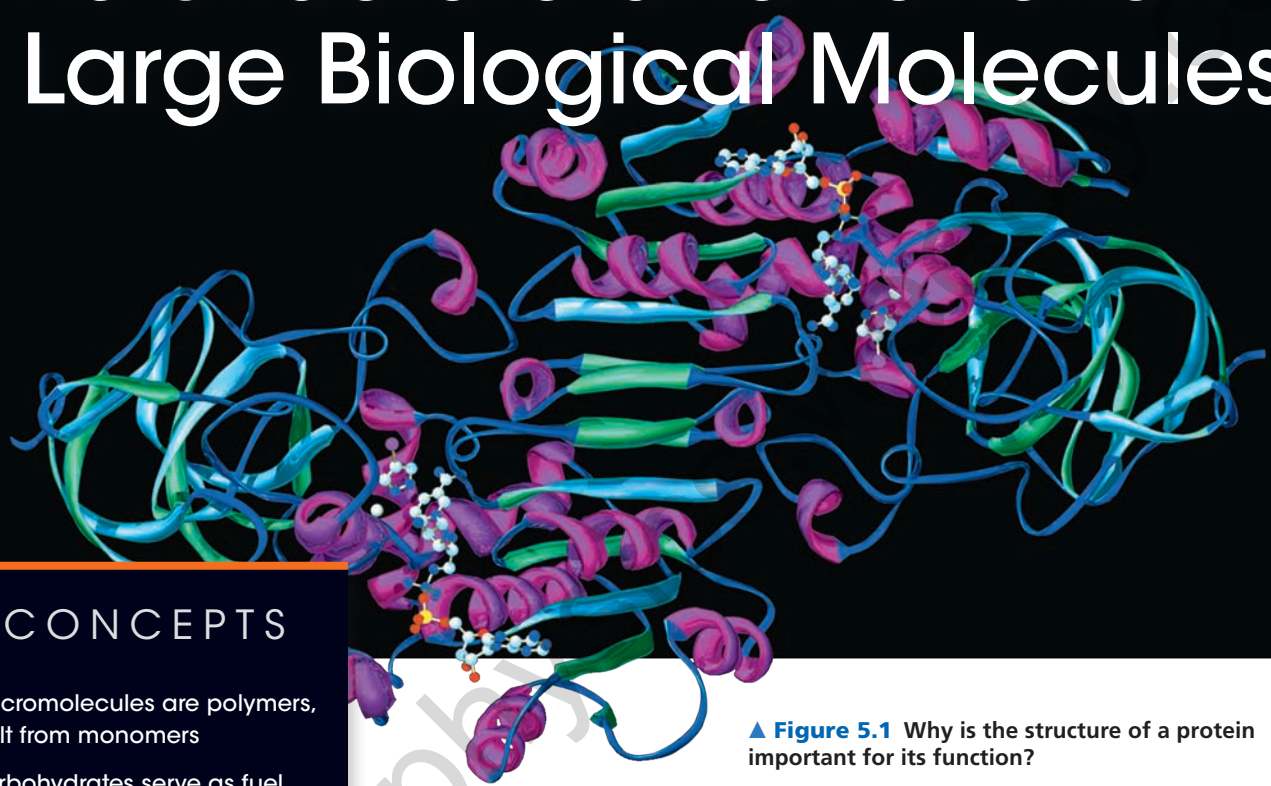
MasteringBiology®

Students Go to **MasteringBiology** for assignments, the eText, and the Study Area with practice tests, animations, and activities.

Instructors Go to **MasteringBiology** for automatically graded tutorials and questions that you can assign to your students, plus Instructor Resources.

5

The Structure and Function of Large Biological Molecules



▲ **Figure 5.1** Why is the structure of a protein important for its function?

KEY CONCEPTS

5.1 Macromolecules are polymers, built from monomers

5.2 Carbohydrates serve as fuel and building material

5.3 Lipids are a diverse group of hydrophobic molecules

5.4 Proteins include a diversity of structures, resulting in a wide range of functions

5.5 Nucleic acids store, transmit, and help express hereditary information

5.6 Genomics and proteomics have transformed biological inquiry and applications

The Molecules of Life

Given the rich complexity of life on Earth, it might surprise you that the most important large molecules found in all living things—from bacteria to elephants—can be sorted into just four main classes: carbohydrates, lipids, proteins, and nucleic acids. On the molecular scale, members of three of these classes—carbohydrates, proteins, and nucleic acids—are huge and are therefore called **macromolecules**. For example, a protein may consist of thousands of atoms that form a molecular colossus with a mass well over 100,000 daltons. Considering the size and complexity of macromolecules, it is noteworthy that biochemists have determined the detailed structure of so many of them. The image in **Figure 5.1** is a molecular model of a protein called alcohol dehydrogenase, which breaks down alcohol in the body.

The architecture of a large biological molecule plays an essential role in its function. Like water and simple organic molecules, large biological molecules exhibit unique emergent properties arising from the orderly arrangement of their atoms. In this chapter, we'll first consider how macromolecules are built. Then we'll examine the structure and function of all four classes of large biological molecules: carbohydrates, lipids, proteins, and nucleic acids.

CONCEPT 5.1

Macromolecules are polymers, built from monomers

The macromolecules in three of the four classes of life's organic compounds—carbohydrates, proteins, and nucleic acids, all except lipids—are chain-like molecules called polymers (from the Greek *polys*, many, and *meros*, part). A **polymer** is a long molecule consisting of many similar or identical building blocks linked by covalent bonds, much as a train consists of a chain of cars. The repeating units that serve as the building blocks of a polymer are smaller molecules called **monomers** (from the Greek *monos*, single). Some monomers also have other functions of their own.

The Synthesis and Breakdown of Polymers

Although each class of polymer is made up of a different type of monomer, the chemical mechanisms by which cells make and break down polymers are basically the same in all cases. In cells, these processes are facilitated by **enzymes**, specialized macromolecules that speed up chemical reactions. Monomers are connected by a reaction in which two molecules are covalently bonded to each other, with the loss of a water molecule; this is known as a **dehydration reaction** **(Figure 5.2a)**. When a bond forms between two monomers, each monomer contributes part of the water molecule that is released during the reaction: One monomer provides a hydroxyl group (—OH), while the other provides a hydrogen (—H). This reaction is repeated as monomers are added to the chain one by one, making a polymer.

Polymers are disassembled to monomers by **hydrolysis**, a process that is essentially the reverse of the dehydration reaction **(Figure 5.2b)**. Hydrolysis means water breakage (from the Greek *hydro*, water, and *lysis*, break). The bond between monomers is broken by the addition of a water molecule, with a hydrogen from water attaching to one monomer and the hydroxyl group attaching to the other. An example of hydrolysis within our bodies is the process of digestion. The bulk of the organic material in our food is in the form of polymers that are much too large to enter our cells. Within the digestive tract, various enzymes attack the polymers, speeding up hydrolysis. Released monomers are then absorbed into the bloodstream for distribution to all body cells. Those cells can then use dehydration reactions to assemble the monomers into new, different polymers that can perform specific functions required by the cell.

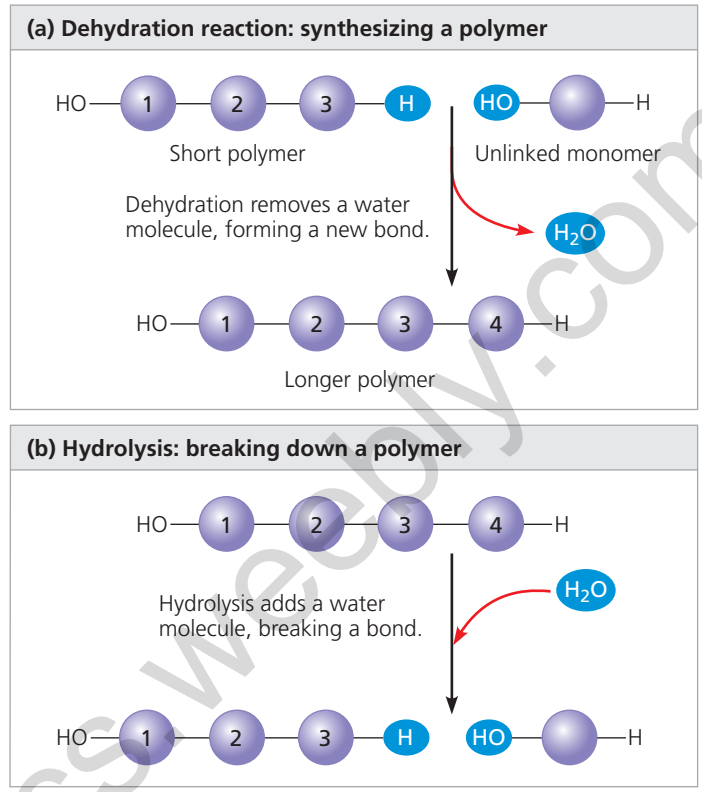
The Diversity of Polymers

A cell has thousands of different macromolecules; the collection varies from one type of cell to another. The inherited differences between close relatives such as human siblings reflect small variations in polymers, particularly DNA and proteins. Molecular differences between unrelated individuals are more extensive, and those between species greater still. The diversity of macromolecules in the living world is vast, and the possible variety is effectively limitless.

What is the basis for such diversity in life's polymers? These molecules are constructed from only 40 to 50 common monomers and some others that occur rarely. Building a huge variety of polymers from such a limited number of monomers is analogous to constructing hundreds of thousands of words from only 26 letters of the alphabet. The key is arrangement—the particular linear sequence that the units follow. However, this analogy falls far short of describing the great diversity of macromolecules because most biological polymers have many more monomers than the number of letters a word, even the longest ones. Proteins, for example, are built from 20 kinds of amino acids arranged in chains that are typically hundreds of amino acids long. The molecular logic of life is simple but elegant: Small molecules common to all organisms are ordered into unique macromolecules.

Despite this immense diversity, molecular structure and function can still be grouped roughly by class. Let's examine each of the four major classes of large biological molecules. For each class, the large molecules have emergent properties not found in their individual building blocks.

▼ Figure 5.2 The synthesis and breakdown of polymers.



CONCEPT CHECK 5.1

1. What are the four main classes of large biological molecules? Which class does not consist of polymers?
2. How many molecules of water are needed to completely hydrolyze a polymer that is ten monomers long?
3. **WHAT IF?** If you eat a piece of fish, what reactions must occur for the amino acid monomers in the protein of the fish to be converted to new proteins in your body?

For suggested answers, see Appendix A.

CONCEPT 5.2

Carbohydrates serve as fuel and building material

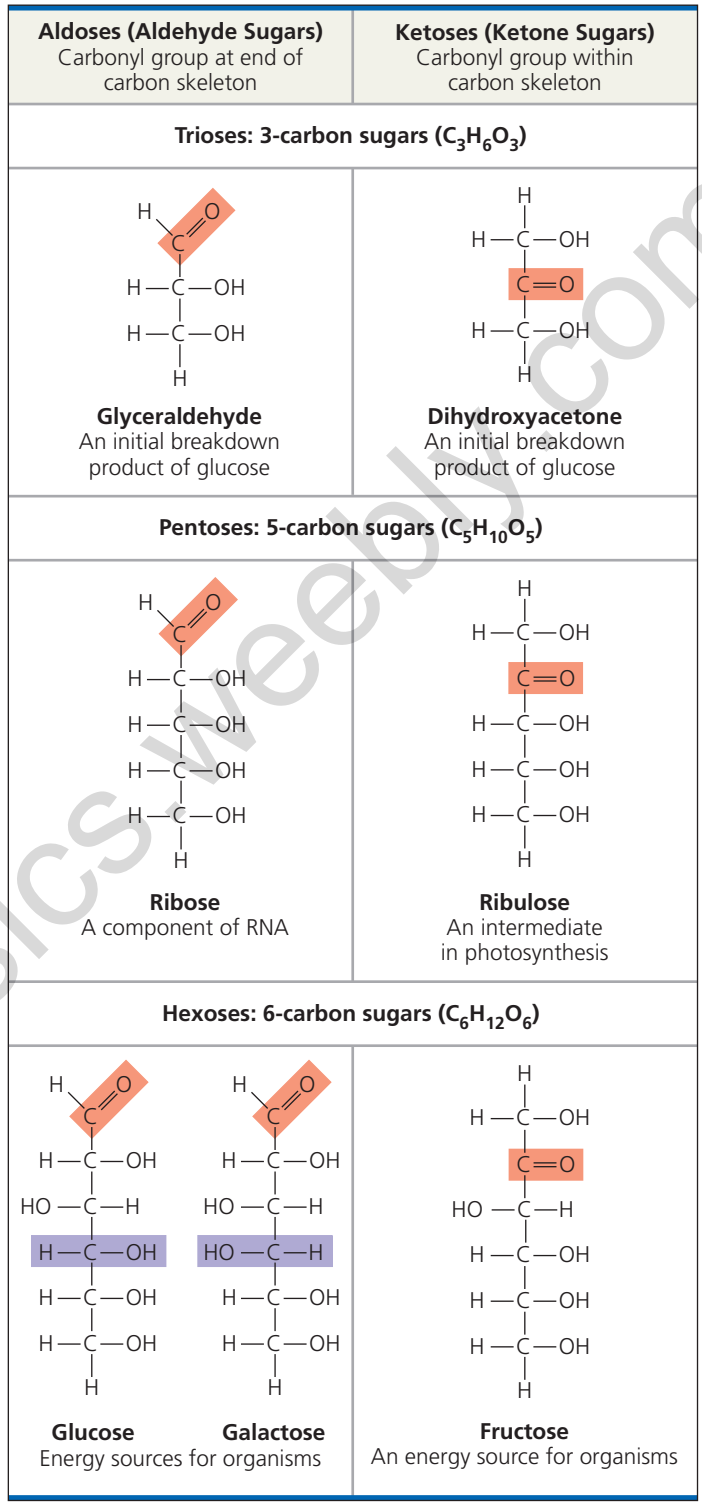
Carbohydrates include sugars and polymers of sugars. The simplest carbohydrates are the monosaccharides, or simple sugars; these are the monomers from which more complex carbohydrates are built. Disaccharides are double sugars, consisting of two monosaccharides joined by a covalent bond. Carbohydrate macromolecules are polymers called polysaccharides, composed of many sugar building blocks.

Sugars

Monosaccharides (from the Greek *monos*, single, and *saccharon*, sugar) generally have molecular formulas that are some multiple of the unit CH_2O. Glucose ($C_6H_{12}O_6$), the most common monosaccharide, is of central importance in the chemistry of life. In the structure of glucose, we can see the trademarks of a sugar: The molecule has a carbonyl group (CO) and multiple hydroxyl groups (—OH) **(Figure 5.3)**. Depending on the location of the carbonyl group, a sugar is either an aldose (aldehyde sugar) or a ketose (ketone sugar). Glucose, for example, is an aldose; fructose, an isomer of glucose, is a ketose. (Most names for sugars end in *-ose*.) Another criterion for classifying sugars is the size of the carbon skeleton, which ranges from three to seven carbons long. Glucose, fructose, and other sugars that have six carbons are called hexoses. Trioses (three-carbon sugars) and pentoses (five-carbon sugars) are also common.

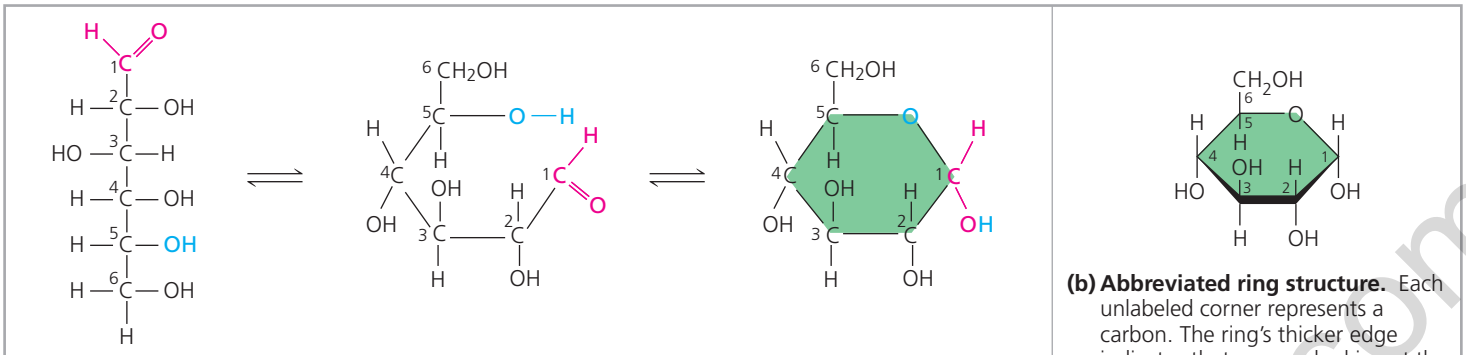
Still another source of diversity for simple sugars is in the spatial arrangement of their parts around asymmetric carbons. (Recall that an asymmetric carbon is a carbon attached to four different atoms or groups of atoms.) Glucose and galactose, for example, differ only in the placement of parts around one asymmetric carbon (see the purple boxes in Figure 5.3). What seems like a small difference is significant enough to give the two sugars distinctive shapes and binding activities, thus different behaviors.

Although it is convenient to draw glucose with a linear carbon skeleton, this representation is not completely accurate.



▲ Figure 5.3 The structure and classification of some monosaccharides. Sugars vary in the location of their carbonyl groups (orange), the length of their carbon skeletons, and the spatial arrangement around asymmetric carbons (compare, for example, the purple portions of glucose and galactose).

MAKE CONNECTIONS *In the 1970s, a process was developed that converts the glucose in corn syrup to its sweeter-tasting isomer, fructose. High-fructose corn syrup, a common ingredient in soft drinks and processed food, is a mixture of glucose and fructose. What type of isomers are glucose and fructose? (See Figure 4.7.)*



(a) Linear and ring forms. Chemical equilibrium between the linear and ring structures greatly favors the formation of rings. The carbons of the sugar are numbered 1 to 6, as shown. To form the glucose ring, carbon 1 (magenta) bonds to the oxygen (blue) attached to carbon 5.

(b) Abbreviated ring structure. Each unlabeled corner represents a carbon. The ring's thicker edge indicates that you are looking at the ring edge-on; the components attached to the ring lie above or below the plane of the ring.

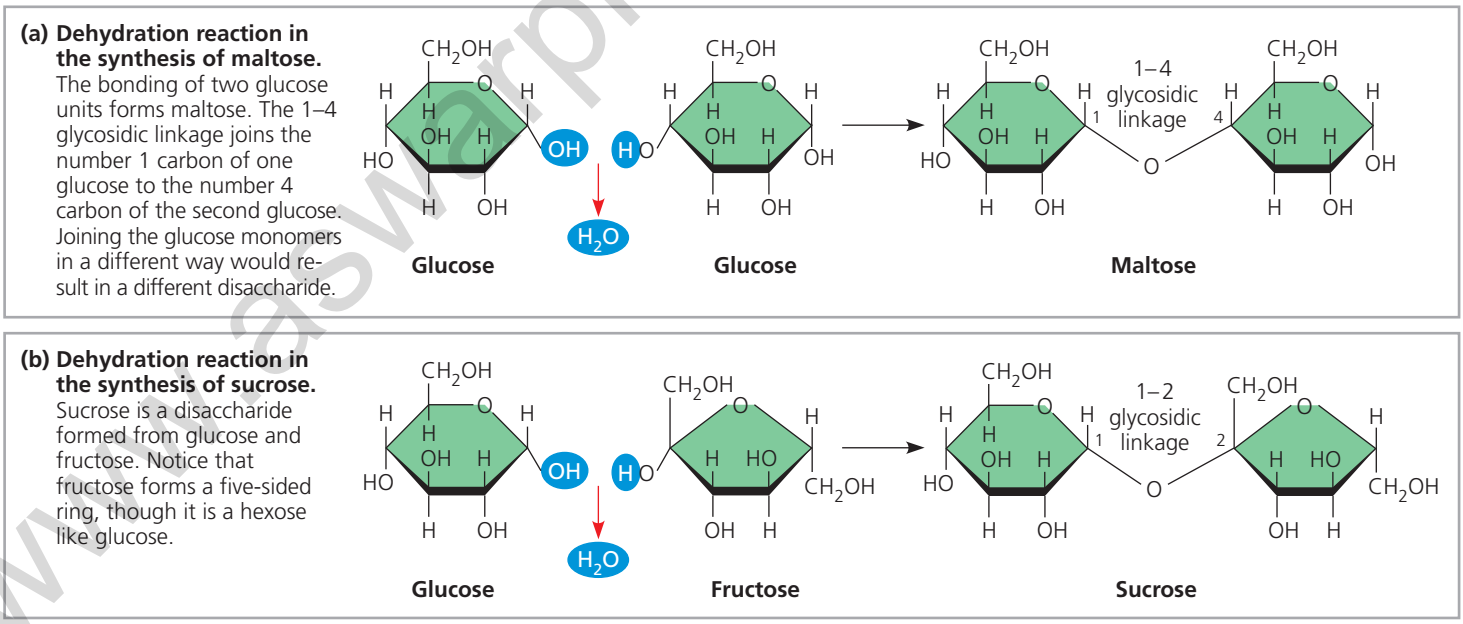
▲ Figure 5.4 Linear and ring forms of glucose.

DRAW IT *Start with the linear form of fructose (see Figure 5.3) and draw the formation of the fructose ring in two steps. First, number the carbons starting at the top of the linear structure. Then draw the molecule in the same orientation as the glucose in the middle of (a) above, attaching carbon 5 via its oxygen to carbon 2. Compare the number of carbons in the fructose and glucose rings.*

In aqueous solutions, glucose molecules, as well as most other five- and six-carbon sugars, form rings **(Figure 5.4)**.

Monosaccharides, particularly glucose, are major nutrients for cells. In the process known as cellular respiration, cells extract energy from glucose molecules by breaking them down in a series of reactions. Not only are simple-sugar molecules a major fuel for cellular work, but their carbon skeletons also serve as raw material for the synthesis of other types of small organic molecules, such as amino acids and fatty acids. Sugar molecules that are not immediately used in these ways are generally incorporated as monomers into disaccharides or polysaccharides.

A **disaccharide** consists of two monosaccharides joined by a **glycosidic linkage**, a covalent bond formed between two monosaccharides by a dehydration reaction. For example, maltose is a disaccharide formed by the linking of two molecules of glucose **(Figure 5.5a)**. Also known as malt sugar, maltose is an ingredient used in brewing beer. The most prevalent disaccharide is sucrose, which is table sugar. Its two monomers are glucose and fructose **(Figure 5.5b)**. Plants generally transport carbohydrates from leaves to roots and other nonphotosynthetic organs in the form of sucrose. Lactose, the sugar present in milk, is another disaccharide, in this case a glucose molecule joined to a galactose molecule.



(a) Dehydration reaction in the synthesis of maltose. The bonding of two glucose units forms maltose. The 1–4 glycosidic linkage joins the number 1 carbon of one glucose to the number 4 carbon of the second glucose. Joining the glucose monomers in a different way would result in a different disaccharide.

(b) Dehydration reaction in the synthesis of sucrose. Sucrose is a disaccharide formed from glucose and fructose. Notice that fructose forms a five-sided ring, though it is a hexose like glucose.

▲ Figure 5.5 Examples of disaccharide synthesis.

DRAW IT *Referring to Figures 5.3 and 5.4, number the carbons in each sugar in this figure. Insert arrows linking the carbons to show how the numbering is consistent with the name of each glycosidic linkage.*

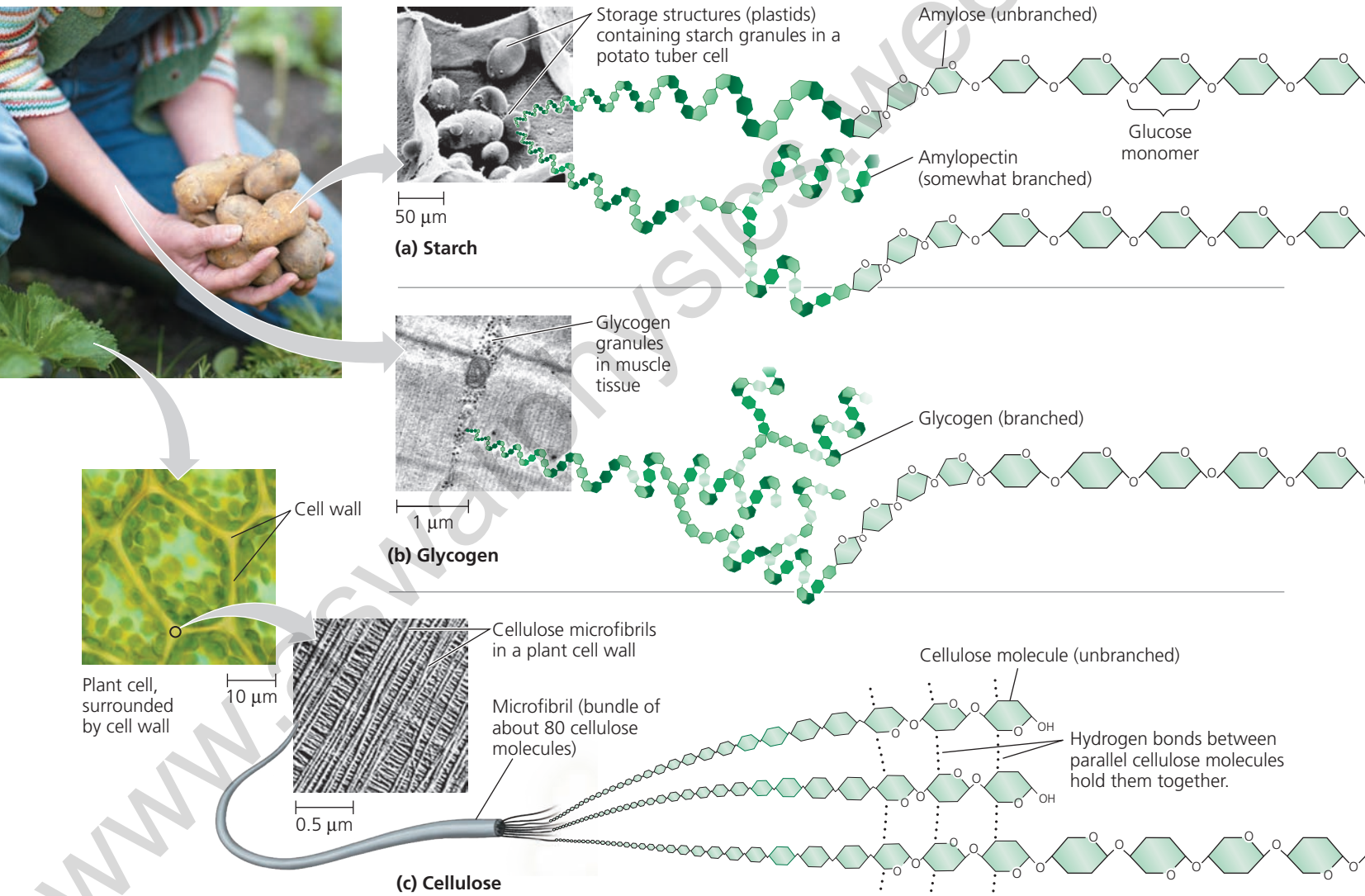
Polysaccharides

Polysaccharides are macromolecules, polymers with a few hundred to a few thousand monosaccharides joined by glycosidic linkages. Some polysaccharides serve as storage material, hydrolyzed as needed to provide sugar for cells. Other polysaccharides serve as building material for structures that protect the cell or the whole organism. The architecture and function of a polysaccharide are determined by its sugar monomers and by the positions of its glycosidic linkages.

Storage Polysaccharides

Both plants and animals store sugars for later use in the form of storage polysaccharides **(Figure 5.6)**. Plants store **starch**, a polymer of glucose monomers, as granules within cellular structures known as plastids, which include chloroplasts. Synthesizing starch enables the plant to stockpile surplus glucose. Because glucose is a major cellular fuel, starch represents stored energy. The sugar can later be withdrawn from this carbohydrate "bank" by hydrolysis, which breaks the bonds between the glucose monomers. Most animals, including humans, also have enzymes that can hydrolyze plant starch, making glucose available as a nutrient for cells. Potato tubers and grains are the major sources of starch in the human diet.

Most of the glucose monomers in starch are joined by 1–4 linkages (number 1 carbon to number 4 carbon), like the glucose units in maltose (see Figure 5.5a). The simplest form of starch, amylose, is unbranched. Amylopectin, a more complex starch, is a branched polymer with 1–6 linkages at the branch points. Both of these starches are shown in **Figure 5.6a**.



▲ Figure 5.6 Polysaccharides of plants and animals. (a) Starch stored in plant cells, **(b)** glycogen stored in muscle cells, and **(c)** structural cellulose fibers in plant cell walls are all polysaccharides composed entirely of glucose monomers (green hexagons). In starch and glycogen, the polymer chains tend to form helices in unbranched regions because of the angle of the linkages between glucose molecules. There are two kinds of starch: amylose and amylopectin. Cellulose, with a different kind of glucose linkage, is always unbranched.

Animals store a polysaccharide called **glycogen**, a polymer of glucose that is like amylopectin but more extensively branched **(Figure 5.6b)**. Vertebrates store glycogen mainly in liver and muscle cells. Hydrolysis of glycogen in these cells releases glucose when the demand for sugar increases. This stored fuel cannot sustain an animal for long, however. In humans, for example, glycogen stores are depleted in about a day unless they are replenished by consumption of food. This is an issue of concern in low-carbohydrate diets, which can result in weakness and fatigue.

Structural Polysaccharides

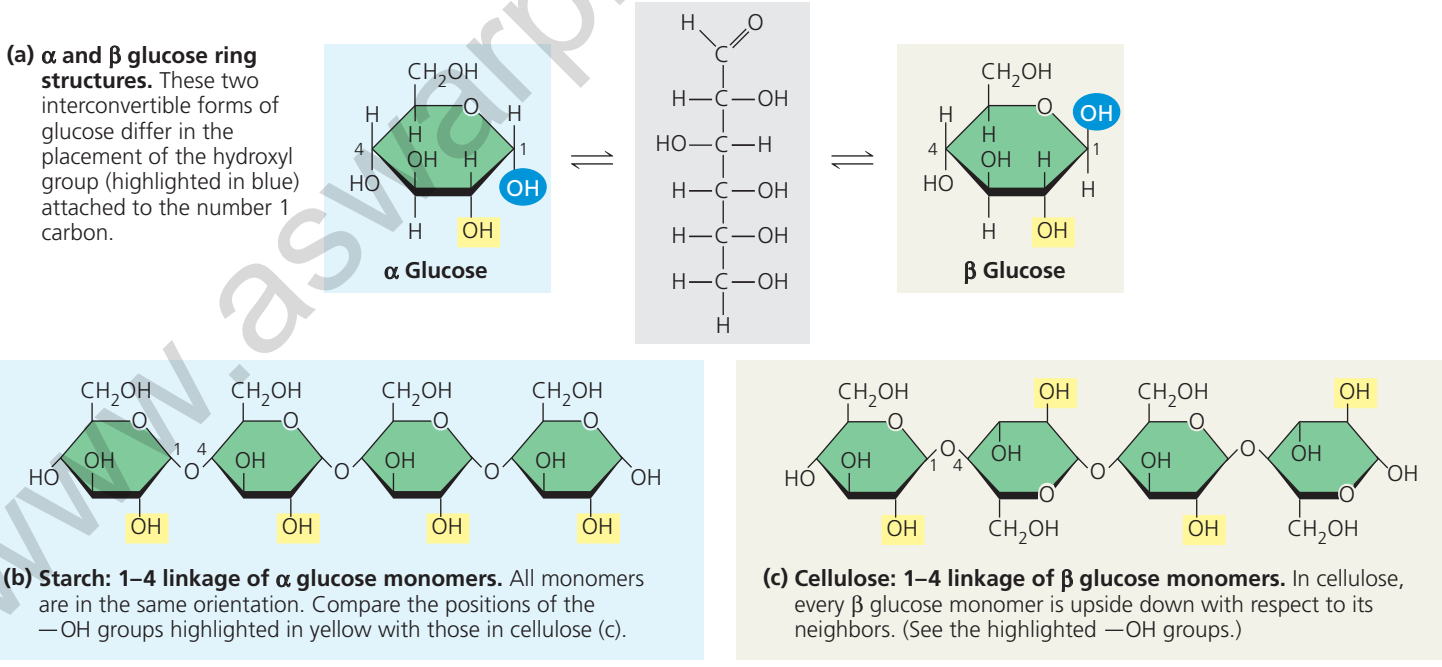
Organisms build strong materials from structural polysaccharides. For example, the polysaccharide called **cellulose** is a major component of the tough walls that enclose plant cells **(Figure 5.6c)**. On a global scale, plants produce almost 10^{14} kg (100 billion tons) of cellulose per year; it is the most abundant organic compound on Earth.

Like starch, cellulose is a polymer of glucose, but the glycosidic linkages in these two polymers differ. The difference is based on the fact that there are actually two slightly different ring structures for glucose **(Figure 5.7a)**. When glucose forms a ring, the hydroxyl group attached to the number 1 carbon is positioned either below or above the plane of the ring. These two ring forms for glucose are called alpha (α) and beta (β), respectively. (Greek letters are often used as a "numbering" system for different versions of biological structures, much as we use the letters a, b, c, and so on for the parts of a question or a figure.) In starch, all the glucose monomers are in the α configuration **(Figure 5.7b)**, the arrangement we saw in Figures 5.4 and 5.5. In contrast, the glucose monomers of cellulose are all in the β configuration, making every glucose monomer "upside down" with respect to its neighbors **(Figure 5.7c;** see also Figure 5.6c).

The differing glycosidic linkages in starch and cellulose give the two molecules distinct three-dimensional shapes. Whereas certain starch molecules are largely helical, a cellulose molecule is straight. Cellulose is never branched, and some hydroxyl groups on its glucose monomers are free to hydrogen-bond with the hydroxyls of other cellulose molecules lying parallel to it. In plant cell walls, parallel cellulose molecules held together in this way are grouped into units called microfibrils (see Figure 5.6c). These cable-like microfibrils are a strong building material for plants and an important substance for humans because cellulose is the major constituent of paper and the only component of cotton.

Enzymes that digest starch by hydrolyzing its α linkages are unable to hydrolyze the β linkages of cellulose due to the different shapes of these two molecules. In fact, few organisms possess enzymes that can digest cellulose. Almost all animals, including humans, do not; the cellulose in our food passes through the digestive tract and is eliminated with the feces. Along the way, the cellulose abrades the wall of the digestive tract and stimulates the lining to secrete mucus, which aids in the smooth passage of food through the tract. Thus, although cellulose is not a nutrient for humans, it is an important part of a healthful diet. Most fruits, vegetables, and whole grains are rich in cellulose. On food packages, "insoluble fiber" refers mainly to cellulose.

Some microorganisms can digest cellulose, breaking it down into glucose monomers. A cow harbors

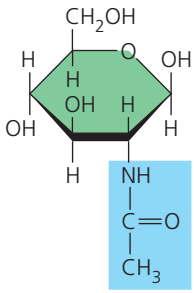


(a) α and β glucose ring structures. These two interconvertible forms of glucose differ in the placement of the hydroxyl group (highlighted in blue) attached to the number 1 carbon.

(b) Starch: 1–4 linkage of α glucose monomers. All monomers are in the same orientation. Compare the positions of the —OH groups highlighted in yellow with those in cellulose (c).

(c) Cellulose: 1–4 linkage of β glucose monomers. In cellulose, every β glucose monomer is upside down with respect to its neighbors. (See the highlighted —OH groups.)

▲ **Figure 5.7** **Starch and cellulose structures.**



◀ The structure of the chitin monomer

◀ Chitin, embedded in proteins, forms the exoskeleton of arthropods. This cicada is molting—shedding its old exoskeleton and emerging in adult form.

▶ Chitin is used to make a strong and flexible surgical thread that decomposes after the wound or incision heals.

▲ Figure 5.8 Chitin, a structural polysaccharide.

cellulose-digesting prokaryotes and protists in its gut. These microbes hydrolyze the cellulose of hay and grass and convert the glucose to other compounds that nourish the cow. Similarly, a termite, which is unable to digest cellulose by itself, has prokaryotes or protists living in its gut that can make a meal of wood. Some fungi can also digest cellulose in soil and elsewhere, thereby helping recycle chemical elements within Earth's ecosystems.

Another important structural polysaccharide is **chitin**, the carbohydrate used by arthropods (insects, spiders, crustaceans, and related animals) to build their exoskeletons **(Figure 5.8)**. An exoskeleton is a hard case that surrounds the soft parts of an animal. Made up of chitin embedded in a layer of proteins, the case is leathery and flexible at first, but becomes hardened when the proteins are chemically linked to each other (as in insects) or encrusted with calcium carbonate (as in crabs). Chitin is also found in fungi, which use this polysaccharide rather than cellulose as the building material for their cell walls. Chitin is similar to cellulose, with β linkages, except that the glucose monomer of chitin has a nitrogen-containing appendage (see Figure 5.8, top right).

CONCEPT CHECK 5.2

1. Write the formula for a monosaccharide that has three carbons.
2. A dehydration reaction joins two glucose molecules to form maltose. The formula for glucose is $C_6H_{12}O_6$. What is the formula for maltose?
3. **WHAT IF?** After a cow is given antibiotics to treat an infection, a vet gives the animal a drink of "gut culture" containing various prokaryotes. Why is this necessary?

For suggested answers, see Appendix A.

CONCEPT 5.3

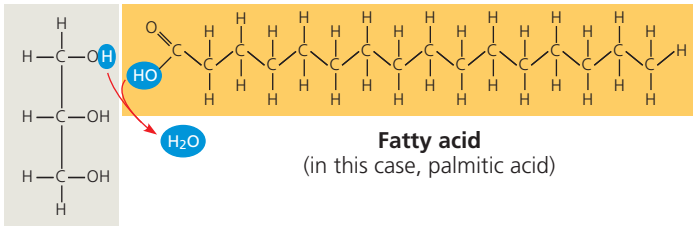
Lipids are a diverse group of hydrophobic molecules

Lipids are the one class of large biological molecules that does not include true polymers, and they are generally not big enough to be considered macromolecules. The compounds called **lipids** are grouped with each other because they share one important trait: They mix poorly, if at all, with water. The hydrophobic behavior of lipids is based on their molecular structure. Although they may have some polar bonds associated with oxygen, lipids consist mostly of hydrocarbon regions. Lipids are varied in form and function. They include waxes and certain pigments, but we will focus on the types of lipids that are most biologically important: fats, phospholipids, and steroids.

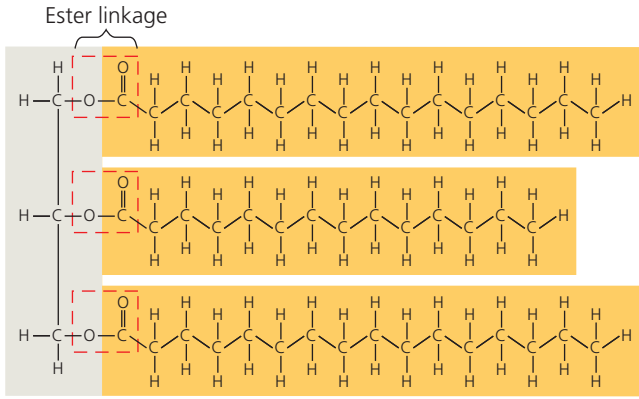
Fats

Although fats are not polymers, they are large molecules assembled from smaller molecules by dehydration reactions. A **fat** is constructed from two kinds of smaller molecules: glycerol and fatty acids **(Figure 5.9a)**. Glycerol is an alcohol; each of its three carbons bears a hydroxyl group. A **fatty acid** has a long carbon skeleton, usually 16 or 18 carbon atoms in length. The carbon at one end of the skeleton is part of a carboxyl group, the functional group that gives these molecules the name fatty *acid*. The rest of the skeleton consists of a hydrocarbon chain. The relatively nonpolar C—H bonds in the hydrocarbon chains of fatty acids are the reason fats are hydrophobic. Fats separate from water because the water molecules hydrogen-bond to one another and exclude the fats. This is the reason that vegetable oil (a liquid fat) separates from the aqueous vinegar solution in a bottle of salad dressing.

In making a fat, three fatty acid molecules are each joined to glycerol by an ester linkage, a bond formed by a dehydration reaction between a hydroxyl group and a carboxyl group. The resulting fat, also called a **triacylglycerol**, thus consists of three fatty acids linked to one glycerol molecule.



(a) One of three dehydration reactions in the synthesis of a fat



(b) Fat molecule (triacylglycerol)

▲ Figure 5.9 The synthesis and structure of a fat, or triacylglycerol. The molecular building blocks of a fat are one molecule of glycerol and three molecules of fatty acids. **(a)** One water molecule is removed for each fatty acid joined to the glycerol. **(b)** A fat molecule with three fatty acid units, two of them identical. The carbons of the fatty acids are arranged zigzag to suggest the actual orientations of the four single bonds extending from each carbon (see Figure 4.3a).

(Still another name for a fat is *triglyceride*, a word often found in the list of ingredients on packaged foods.) The fatty acids in a fat can all be the same, or they can be of two or three different kinds, as in **Figure 5.9b**.

The terms *saturated* fats and *unsaturated* fats are commonly used in the context of nutrition **(Figure 5.10)**. These terms refer to the structure of the hydrocarbon chains of the fatty acids. If there are no double bonds between carbon atoms composing a chain, then as many hydrogen atoms as possible are bonded to the carbon skeleton. Such a structure is said to be *saturated* with hydrogen, and the resulting fatty acid is therefore called a **saturated fatty acid (Figure 5.10a)**. An **unsaturated fatty acid** has one or more double bonds, with one fewer hydrogen atom on each double-bonded carbon. Nearly all double bonds in naturally occurring fatty acids are *cis* double bonds, which cause a kink in the hydrocarbon chain wherever they occur **(Figure 5.10b)**. (See Figure 4.7b to remind yourself about *cis* and *trans* double bonds.)

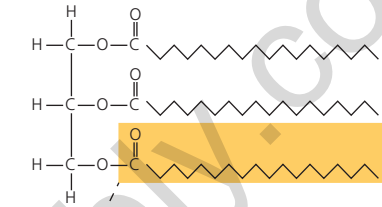
A fat made from saturated fatty acids is called a saturated fat. Most animal fats are saturated: The hydrocarbon chains of their fatty acids—the "tails" of the fat molecules—lack double bonds, and their flexibility allows the fat molecules to pack together tightly. Saturated animal fats—such as lard and butter—are solid at room temperature. In contrast, the fats of plants and fishes are generally unsaturated, meaning that they are built of one or more types of unsaturated fatty acids. Usually liquid at room temperature, plant and fish fats are referred to as oils—olive oil and cod liver oil are examples. The kinks where the *cis* double bonds are located prevent the molecules from packing together closely enough

▼ Figure 5.10 Saturated and unsaturated fats and fatty acids.

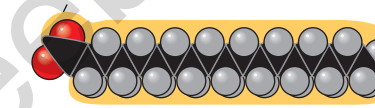
(a) Saturated fat

At room temperature, the molecules of a saturated fat, such as the fat in butter, are packed closely together, forming a solid.

Structural formula of a saturated fat molecule (Each hydrocarbon chain is represented as a zigzag line, where each bend represents a carbon atom and hydrogens are not shown.)



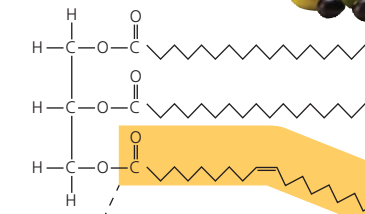
Space-filling model of stearic acid, a saturated fatty acid (red = oxygen, black = carbon, gray = hydrogen)



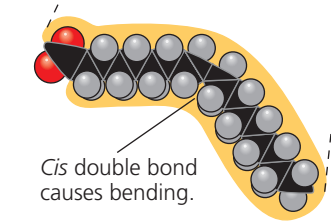
(b) Unsaturated fat

At room temperature, the molecules of an unsaturated fat such as olive oil cannot pack together closely enough to solidify because of the kinks in some of their fatty acid hydrocarbon chains.

Structural formula of an unsaturated fat molecule



Space-filling model of oleic acid, an unsaturated fatty acid



to solidify at room temperature. The phrase "hydrogenated vegetable oils" on food labels means that unsaturated fats have been synthetically converted to saturated fats by adding hydrogen. Peanut butter, margarine, and many other products are hydrogenated to prevent lipids from separating out in liquid (oil) form.

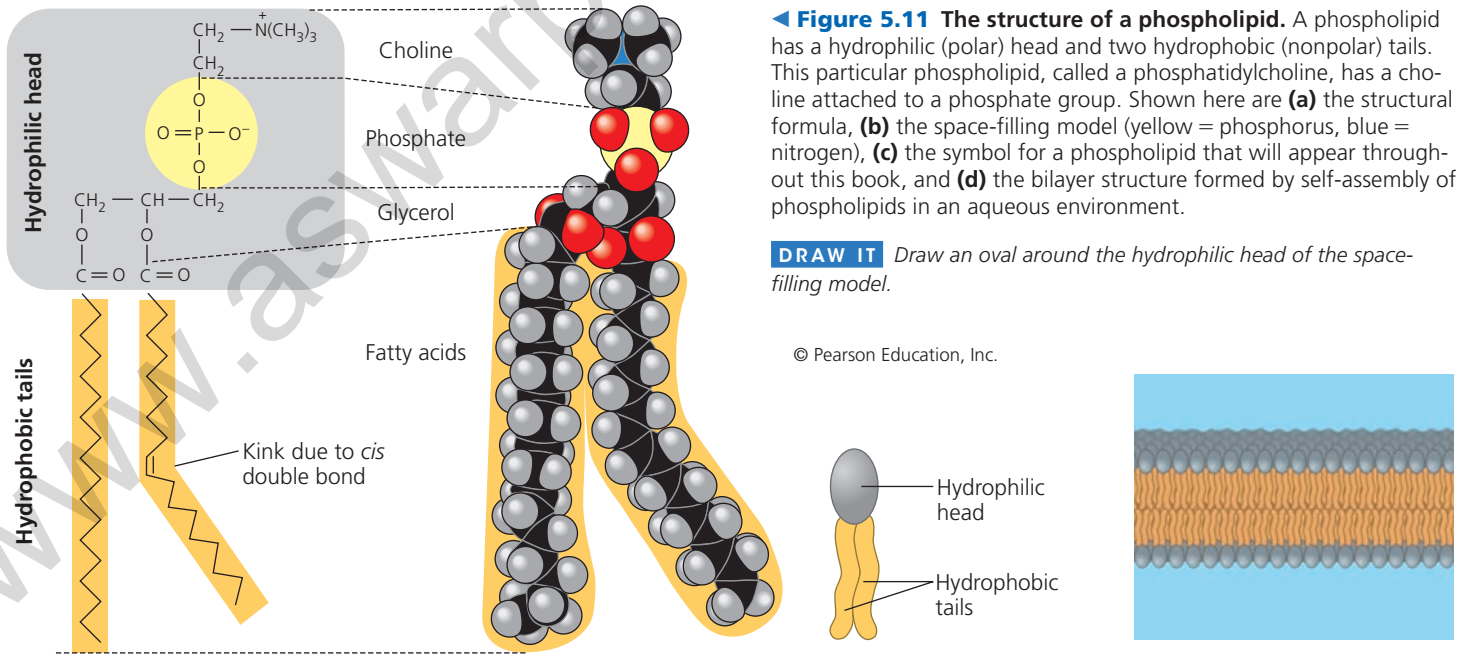
A diet rich in saturated fats is one of several factors that may contribute to the cardiovascular disease known as atherosclerosis. In this condition, deposits called plaques develop within the walls of blood vessels, causing inward bulges that impede blood flow and reduce the resilience of the vessels. Recent studies have shown that the process of hydrogenating vegetable oils produces not only saturated fats but also unsaturated fats with *trans* double bonds. These **trans fats** may contribute more than saturated fats to atherosclerosis (see Chapter 42) and other problems. Because trans fats are especially common in baked goods and processed foods, the U.S. Department of Agriculture requires nutritional labels to include information on trans fat content. Some U.S. cities and at least two countries—Denmark and Switzerland—have even banned the use of trans fats in restaurants.

The major function of fats is energy storage. The hydrocarbon chains of fats are similar to gasoline molecules and just as rich in energy. A gram of fat stores more than twice as much energy as a gram of a polysaccharide, such as starch. Because plants are relatively immobile, they can function with bulky energy storage in the form of starch. (Vegetable oils are generally obtained from seeds, where more compact storage is an asset to the plant.) Animals, however, must carry their energy stores with them, so there is an advantage to having a more compact reservoir of fuel—fat. Humans and other mammals stock their long-term food reserves in adipose cells (see Figure 4.6a), which swell and shrink as fat is deposited and withdrawn from storage. In addition to storing energy, adipose tissue also cushions such vital organs as the kidneys, and a layer of fat beneath the skin insulates the body. This subcutaneous layer is especially thick in whales, seals, and most other marine mammals, protecting them from cold ocean water.

Phospholipids

Cells as we know them could not exist without another type of lipid—phospholipids. Phospholipids are essential for cells because they are major constituents of cell membranes. Their structure provides a classic example of how form fits function at the molecular level. As shown in **Figure 5.11**, a **phospholipid** is similar to a fat molecule but has only two fatty acids attached to glycerol rather than three. The third hydroxyl group of glycerol is joined to a phosphate group, which has a negative electrical charge in the cell. Typically, an additional small charged or polar molecule is also linked to the phosphate group. Choline is one such molecule (see Figure 5.11), but there are many others as well, allowing formation of a variety of phospholipids that differ from each other.

The two ends of phospholipids show different behavior toward water. The hydrocarbon tails are hydrophobic and are excluded from water. However, the phosphate group and its attachments form a hydrophilic head that has an affinity for water. When phospholipids are added to water, they self-assemble into double-layered structures called



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

◀ Figure 5.11 The structure of a phospholipid. A phospholipid has a hydrophilic (polar) head and two hydrophobic (nonpolar) tails. This particular phospholipid, called a phosphatidylcholine, has a choline attached to a phosphate group. Shown here are **(a)** the structural formula, **(b)** the space-filling model (yellow = phosphorus, blue = nitrogen), **(c)** the symbol for a phospholipid that will appear throughout this book, and **(d)** the bilayer structure formed by self-assembly of phospholipids in an aqueous environment.

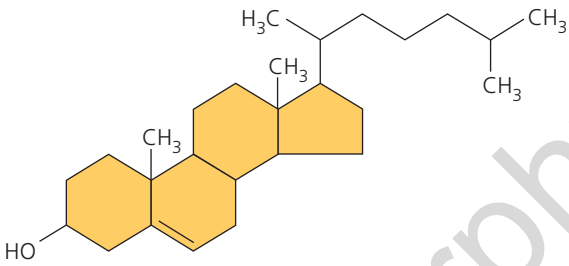
DRAW IT *Draw an oval around the hydrophilic head of the space-filling model.*

"bilayers," shielding their hydrophobic portions from water (Figure 5.11d).

At the surface of a cell, phospholipids are arranged in a similar bilayer. The hydrophilic heads of the molecules are on the outside of the bilayer, in contact with the aqueous solutions inside and outside of the cell. The hydrophobic tails point toward the interior of the bilayer, away from the water. The phospholipid bilayer forms a boundary between the cell and its external environment; in fact, the existence of cells depends on the properties of phospholipids.

Steroids

Steroids are lipids characterized by a carbon skeleton consisting of four fused rings. Different steroids are distinguished by the particular chemical groups attached to this ensemble of rings. **Cholesterol**, a type of steroid, is a crucial molecule in animals **(Figure 5.12)**. It is a common component of animal cell membranes and is also the precursor from which other steroids, such as the vertebrate sex hormones, are synthesized. In vertebrates, cholesterol is synthesized in the liver and is also obtained from the diet. A high level of cholesterol in the blood may contribute to atherosclerosis. In fact, both saturated fats and trans fats exert their negative impact on health by affecting cholesterol levels.



▲ **Figure 5.12 Cholesterol, a steroid.** Cholesterol is the molecule from which other steroids, including the sex hormones, are synthesized. Steroids vary in the chemical groups attached to their four interconnected rings (shown in gold).

MAKE CONNECTIONS *Compare cholesterol with the sex hormones shown in the figure at the beginning of Concept 4.3. Circle the chemical groups that cholesterol has in common with estradiol; put a square around the chemical groups that cholesterol has in common with testosterone.*

CONCEPT CHECK 5.3

1. Compare the structure of a fat (triglyceride) with that of a phospholipid.
2. Why are human sex hormones considered lipids?
3. **WHAT IF?** Suppose a membrane surrounded an oil droplet, as it does in the cells of plant seeds and in some animal cells. Describe and explain the form it might take.

For suggested answers, see Appendix A.

CONCEPT 5.4

Proteins include a diversity of structures, resulting in a wide range of functions

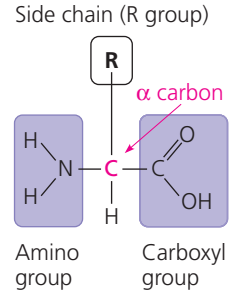
Nearly every dynamic function of a living being depends on proteins. In fact, the importance of proteins is underscored by their name, which comes from the Greek word *proteios*, meaning "first," or "primary." Proteins account for more than 50% of the dry mass of most cells, and they are instrumental in almost everything organisms do. Some proteins speed up chemical reactions, while others play a role in defense, storage, transport, cellular communication, movement, or structural support. **Figure 5.13** shows examples of proteins with these functions, which you'll learn more about in later chapters.

Life would not be possible without enzymes, most of which are proteins. Enzymatic proteins regulate metabolism by acting as **catalysts**, chemical agents that selectively speed up chemical reactions without being consumed by the reaction. Because an enzyme can perform its function over and over again, these molecules can be thought of as workhorses that keep cells running by carrying out the processes of life.

A human has tens of thousands of different proteins, each with a specific structure and function; proteins, in fact, are the most structurally sophisticated molecules known. Consistent with their diverse functions, they vary extensively in structure, each type of protein having a unique three-dimensional shape.

Diverse as proteins are, they are all constructed from the same set of 20 amino acids, linked in unbranched polymers. The bond between amino acids is called a peptide bond, so a polymer of amino acids is called a **polypeptide**. A **protein** is a biologically functional molecule made up of one or more polypeptides, each folded and coiled into a specific three-dimensional structure.

Amino Acid Monomers



All amino acids share a common structure. An **amino acid** is an organic molecule with both an amino group and a carboxyl group (see Figure 4.9). The figure at the right shows the general formula for an amino acid. At the center of the amino acid is an asymmetric carbon atom called the *alpha* (α) *carbon*. Its four different partners are an amino group, a carboxyl group, a hydrogen atom, and a variable group symbolized by R. The R group, also called the side chain, differs with each amino acid.

▼ Figure 5.13 An overview of protein functions.

Enzymatic proteins

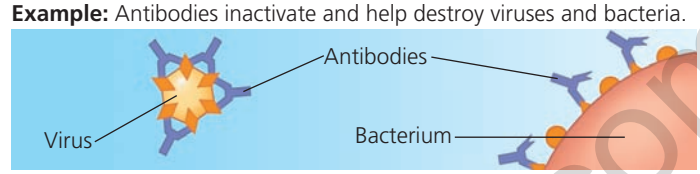
Function: Selective acceleration of chemical reactions

Example: Digestive enzymes catalyze the hydrolysis of bonds in food molecules.

Defensive proteins

Function: Protection against disease

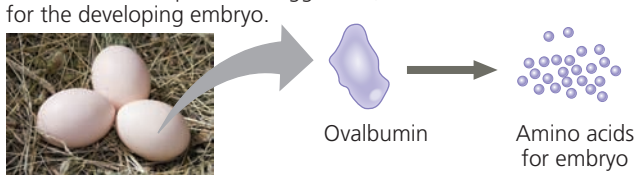
Example: Antibodies inactivate and help destroy viruses and bacteria.



Storage proteins

Function: Storage of amino acids

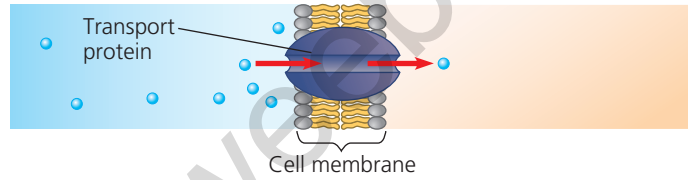
Examples: Casein, the protein of milk, is the major source of amino acids for baby mammals. Plants have storage proteins in their seeds. Ovalbumin is the protein of egg white, used as an amino acid source for the developing embryo.



Transport proteins

Function: Transport of substances

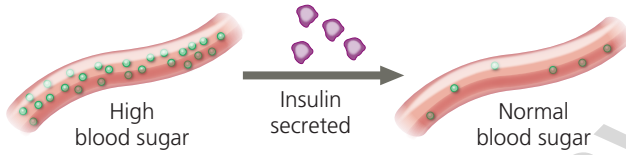
Examples: Hemoglobin, the iron-containing protein of vertebrate blood, transports oxygen from the lungs to other parts of the body. Other proteins transport molecules across membranes, as shown here.



Hormonal proteins

Function: Coordination of an organism's activities

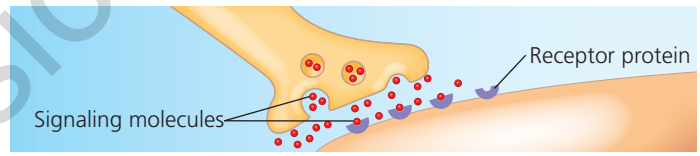
Example: Insulin, a hormone secreted by the pancreas, causes other tissues to take up glucose, thus regulating blood sugar concentration.



Receptor proteins

Function: Response of cell to chemical stimuli

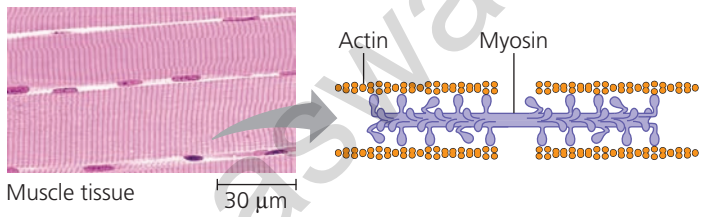
Example: Receptors built into the membrane of a nerve cell detect signaling molecules released by other nerve cells.



Contractile and motor proteins

Function: Movement

Examples: Motor proteins are responsible for the undulations of cilia and flagella. Actin and myosin proteins are responsible for the contraction of muscles.



Structural proteins

Function: Support

Examples: Keratin is the protein of hair, horns, feathers, and other skin appendages. Insects and spiders use silk fibers to make their cocoons and webs, respectively. Collagen and elastin proteins provide a fibrous framework in animal connective tissues.

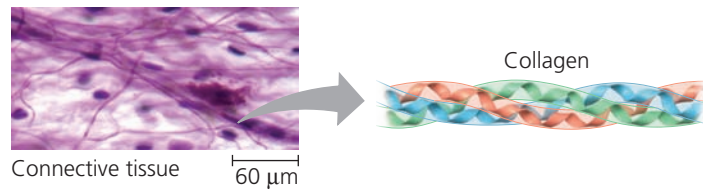
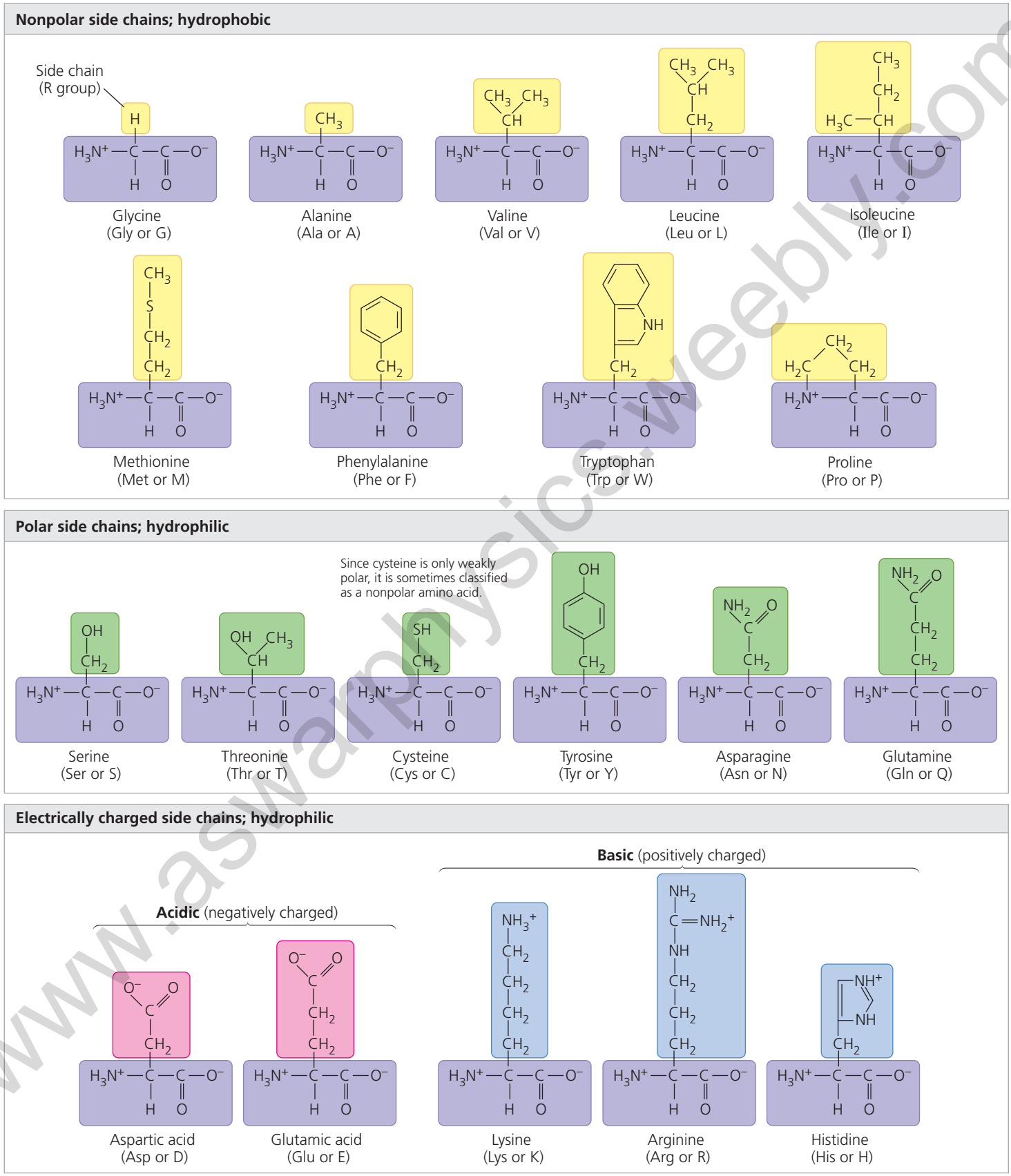


Figure 5.14 shows the 20 amino acids that cells use to build their thousands of proteins. Here the amino groups and carboxyl groups are all depicted in ionized form, the way they usually exist at the pH found in a cell.

The physical and chemical properties of the side chain determine the unique characteristics of a particular amino acid, thus affecting its functional role in a polypeptide. In Figure 5.14, the amino acids are grouped according to the properties of their side chains. One group consists of amino acids with nonpolar side chains, which are hydrophobic.

Another group consists of amino acids with polar side chains, which are hydrophilic. Acidic amino acids are those with side chains that are generally negative in charge due to the presence of a carboxyl group, which is usually dissociated (ionized) at cellular pH. Basic amino acids have amino groups in their side chains that are generally positive in charge. (Notice that *all* amino acids have carboxyl groups and amino groups; the terms *acidic* and *basic* in this context refer only to groups in the side chains.) Because they are charged, acidic and basic side chains are also hydrophilic.

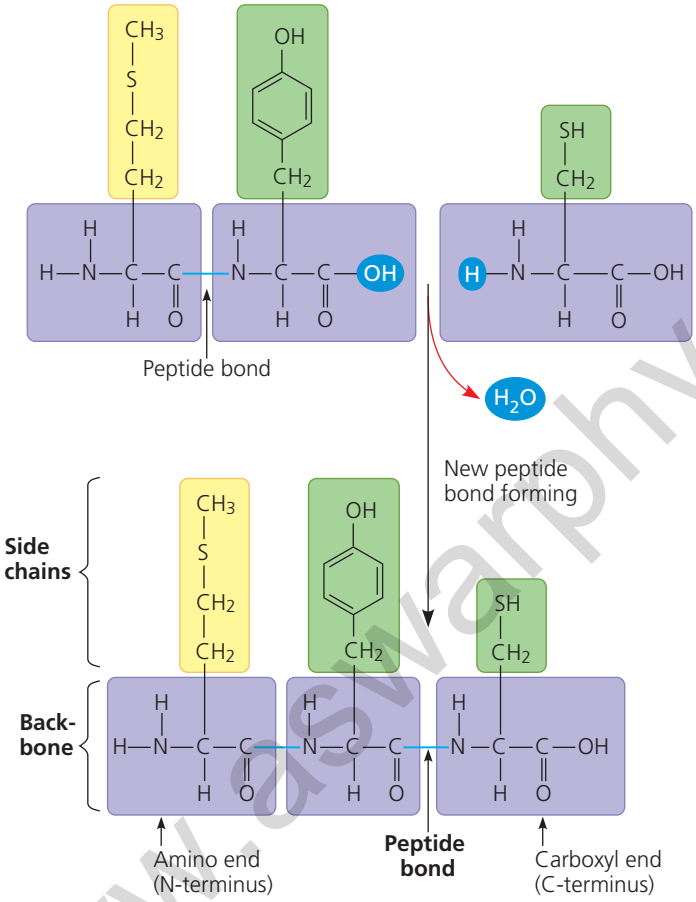
▼ Figure 5.14 The 20 amino acids of proteins. The amino acids are grouped here according to the properties of their side chains (R groups) and shown in their prevailing ionic forms at pH 7.2, the pH within a cell. The three-letter and one-letter abbreviations for the amino acids are in parentheses. All of the amino acids used in proteins are L enantiomers (see Figure 4.7c).



Polypeptides (Amino Acid Polymers)

Now that we have examined amino acids, let's see how they are linked to form polymers **(Figure 5.15)**. When two amino acids are positioned so that the carboxyl group of one is adjacent to the amino group of the other, they can become joined by a dehydration reaction, with the removal of a water molecule. The resulting covalent bond is called a **peptide bond**. Repeated over and over, this process yields a polypeptide, a polymer of many amino acids linked by peptide bonds.

The repeating sequence of atoms highlighted in purple in Figure 5.15 is called the *polypeptide backbone*. Extending from this backbone are the different side chains (R groups) of the amino acids. Polypeptides range in length from a few amino acids to a thousand or more. Each specific polypeptide has a unique linear sequence of amino acids. Note that one end of the polypeptide chain has a free amino group, while the opposite end has a free carboxyl group. Thus, a polypeptide of any length has a single amino end (N-terminus) and a single carboxyl end (C-terminus). In a polypeptide of any significant size, the side chains far outnumber the terminal groups, so the chemical nature of the molecule as a whole is determined by the kind and sequence of the side chains. The immense variety of polypeptides in nature illustrates an important concept introduced earlier—that cells can make many different polymers by linking a limited set of monomers into diverse sequences.



▲ Figure 5.15 Making a polypeptide chain. Peptide bonds are formed by dehydration reactions, which link the carboxyl group of one amino acid to the amino group of the next. The peptide bonds are formed one at a time, starting with the amino acid at the amino end (N-terminus). The polypeptide has a repetitive backbone (purple) to which the amino acid side chains (yellow and green) are attached.

DRAW IT *Label the three amino acids in the upper part of the figure using three-letter and one-letter codes. Circle and label the carboxyl and amino groups that will form the new peptide bond.*

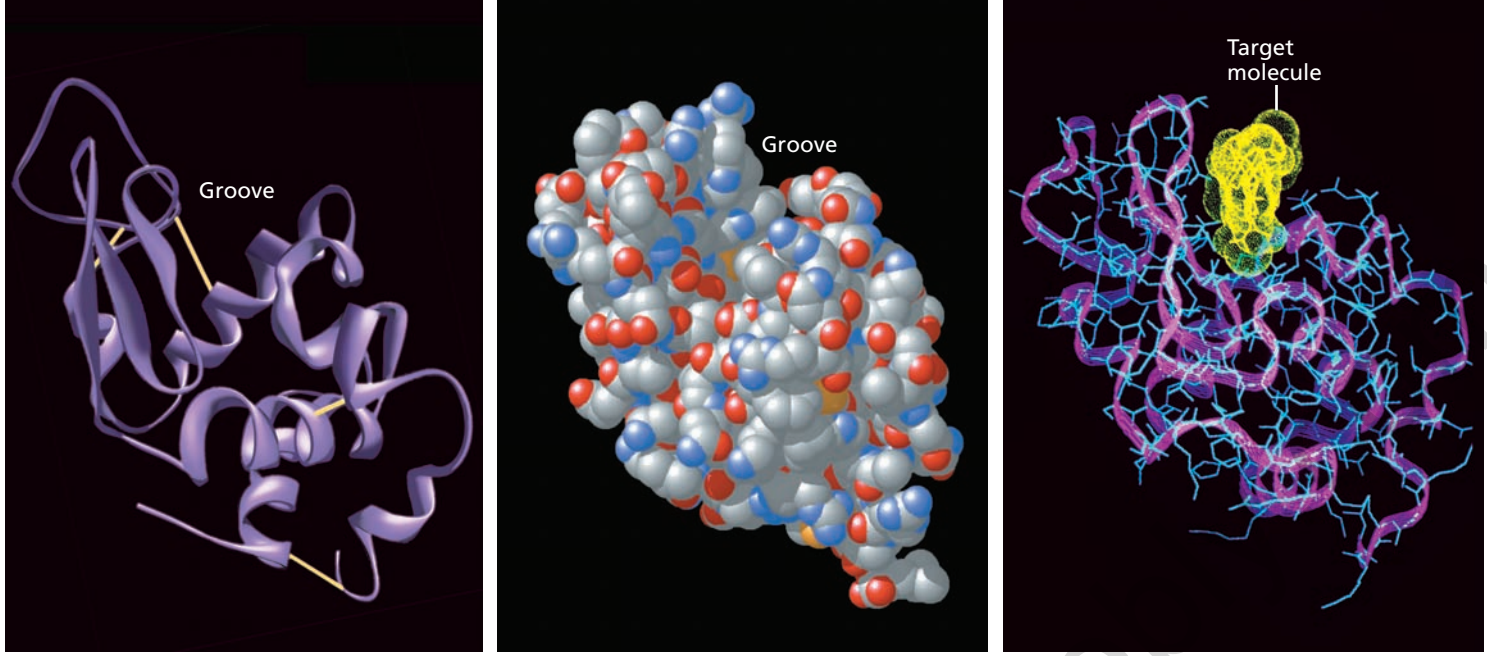
Protein Structure and Function

The specific activities of proteins result from their intricate three-dimensional architecture, the simplest level of which is the sequence of their amino acids. The pioneer in determining the amino acid sequence of proteins was Frederick Sanger, who, with his colleagues at Cambridge University in England, worked on the hormone insulin in the late 1940s and early 1950s. He used agents that break polypeptides at specific places, followed by chemical methods to determine the amino acid sequence in these small fragments. Sanger and his co-workers were able, after years of effort, to reconstruct the complete amino acid sequence of insulin. Since then, the steps involved in sequencing a polypeptide have been automated.

Once we have learned the amino acid sequence of a polypeptide, what can it tell us about the three-dimensional structure (commonly referred to simply as the "structure") of the protein and its function? The term *polypeptide* is not synonymous with the term *protein*. Even for a protein consisting of a single polypeptide, the relationship is somewhat analogous to that between a long strand of yarn and a sweater of particular size and shape that can be knit from the yarn. A functional protein is not *just* a polypeptide chain, but one or more polypeptides precisely twisted, folded, and coiled into a molecule of unique shape, which can be shown in several different types of models **(Figure 5.16)**. And it is the amino acid sequence of each polypeptide that determines what three-dimensional structure the protein will have under normal cellular conditions.

When a cell synthesizes a polypeptide, the chain may fold spontaneously, assuming the functional structure for that protein. This folding is driven and reinforced by the formation of various bonds between parts of the chain, which in turn depends on the sequence of amino acids. Many proteins are roughly spherical (*globular proteins*), while others are shaped like long fibers (*fibrous proteins*). Even within these broad categories, countless variations exist.

A protein's specific structure determines how it works. In almost every case, the function of a protein depends on its ability to recognize and bind to some other molecule. In an especially striking example of the marriage of form and



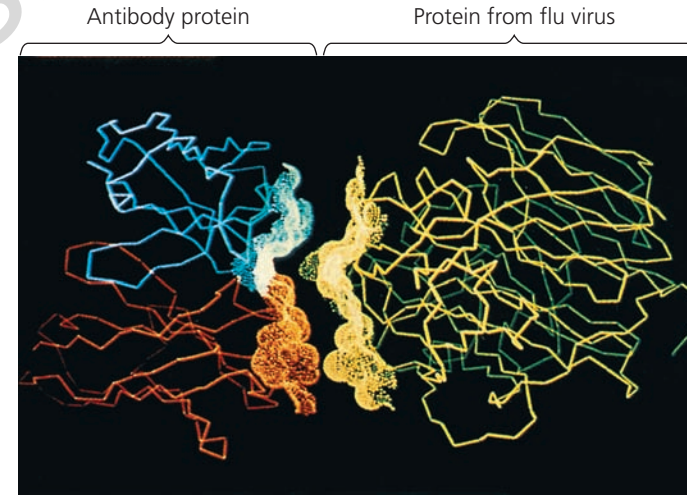
(a) A **ribbon model** shows how the single polypeptide chain folds and coils to form the functional protein. (The yellow lines represent disulfide bridges that stabilize the protein's shape.)

(b) A **space-filling model** shows more clearly the globular shape seen in many proteins, as well as the specific three-dimensional structure unique to lysozyme.

(c) In this view, a ribbon model is superimposed on a **wireframe model**, which shows the backbone with the side chains extending from it. The yellow structure is the target molecule.

▲ Figure 5.16 Structure of a protein, the enzyme lysozyme. Present in our sweat, tears, and saliva, lysozyme is an enzyme that helps prevent infection by binding to and catalyzing the destruction of specific molecules on the surface of many kinds of bacteria. The groove is the part of the protein that recognizes and binds to the target molecules on bacterial walls.

function, **Figure 5.17** shows the exact match of shape between an antibody (a protein in the body) and the particular foreign substance on a flu virus that the antibody binds to and marks for destruction. In Chapter 43, you'll learn more about how the immune system generates antibodies that match the shapes of specific foreign molecules so well. Also, you may recall from Chapter 2 that natural signaling molecules called endorphins bind to specific receptor proteins on the surface of brain cells in humans, producing euphoria and relieving pain. Morphine, heroin, and other opiate drugs are able to mimic endorphins because they all share a similar shape with endorphins and can thus fit into and bind to endorphin receptors in the brain. This fit is very specific, something like a lock and key (see Figure 2.16). Thus, the function of a protein—for instance, the ability of a receptor protein to bind to a particular pain-relieving signaling molecule—is an emergent property resulting from exquisite molecular order.



▲ Figure 5.17 An antibody binding to a protein from a flu virus. A technique called X-ray crystallography was used to generate a computer model of an antibody protein (blue and orange, left) bound to a flu virus protein (green and yellow, right). Computer software was then used to back the images away from each other, revealing the exact complementarity of shape between the two protein surfaces.

Four Levels of Protein Structure

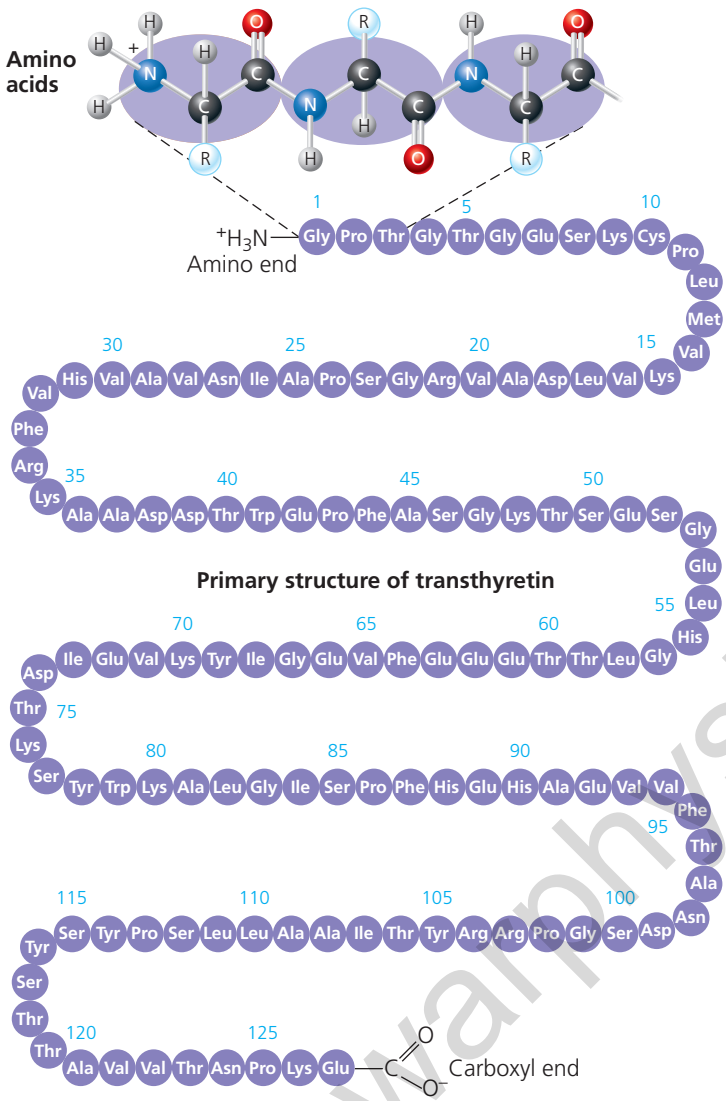
With the goal of understanding the function of a protein, learning about its structure is often productive. In spite of their great diversity, all proteins share three superimposed levels of structure, known as primary, secondary, and tertiary structure. A fourth level, quaternary structure, arises when a protein consists of two or more polypeptide chains. **Figure 5.18** describes these four levels of protein structure. Be sure to study this figure thoroughly before going on to the next section.

▼ Figure 5.18

Exploring Levels of Protein Structure

Primary Structure

Linear chain of amino acids

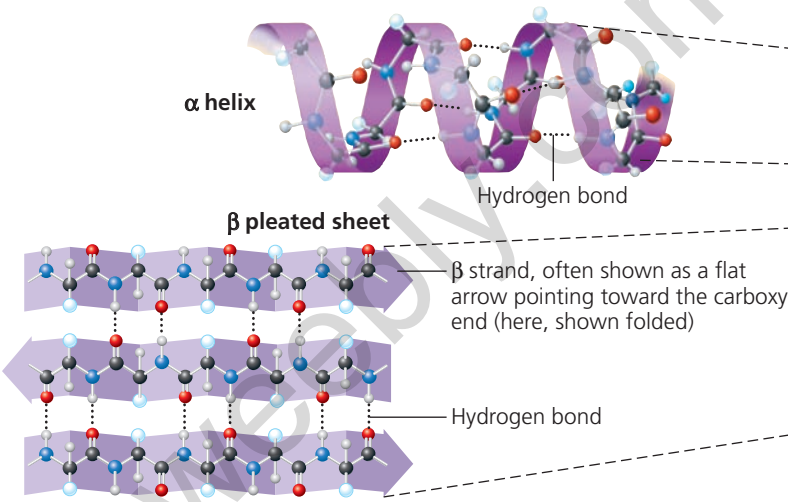


The **primary structure** of a protein is its sequence of amino acids. As an example, let's consider transthyretin, a globular blood protein that transports vitamin A and one of the thyroid hormones throughout the body. Transthyretin is made up of four identical polypeptide chains, each composed of 127 amino acids. Shown here is one of these chains unraveled for a closer look at its primary structure. Each of the 127 positions along the chain is occupied by one of the 20 amino acids, indicated here by its three-letter abbreviation.

The primary structure is like the order of letters in a very long word. If left to chance, there would be 20^{127} different ways of making a polypeptide chain 127 amino acids long. However, the precise primary structure of a protein is determined not by the random linking of amino acids, but by inherited genetic information. The primary structure in turn dictates secondary and tertiary structure, due to the chemical nature of the backbone and the side chains (R groups) of the amino acids along the polypeptide.

Secondary Structure

Regions stabilized by hydrogen bonds between atoms of the polypeptide backbone



Most proteins have segments of their polypeptide chains repeatedly coiled or folded in patterns that contribute to the protein's overall shape. These coils and folds, collectively referred to as **secondary structure**, are the result of hydrogen bonds between the repeating constituents of the polypeptide backbone (not the amino acid side chains). Within the backbone, the oxygen atoms have a partial negative charge, and the hydrogen atoms attached to the nitrogens have a partial positive charge (see Figure 2.14); therefore, hydrogen bonds can form between these atoms. Individually, these hydrogen bonds are weak, but because there are so many of them over a relatively long region of the polypeptide chain, they can support a particular shape for that part of the protein.

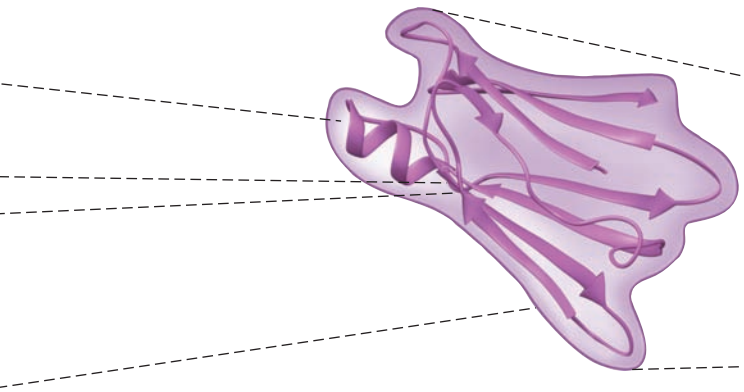
One such secondary structure is the **α helix**, a delicate coil held together by hydrogen bonding between every fourth amino acid, as shown above. Although each transthyretin polypeptide has only one α helix region (see tertiary structure), other globular proteins have multiple stretches of α helix separated by nonhelical regions (see hemoglobin on the next page). Some fibrous proteins, such as α-keratin, the structural protein of hair, have the α helix structure over most of their length.

The other main type of secondary structure is the **β pleated sheet**. As shown above, in this structure two or more segments of the polypeptide chain lying side by side (called β strands) are connected by hydrogen bonds between parts of the two parallel segments of the polypeptide backbone. β pleated sheets make up the core of many globular proteins, as is the case for transthyretin (see tertiary structure), and dominate some fibrous proteins, including the silk protein of a spider's web. The teamwork of so many hydrogen bonds makes each spider silk fiber stronger than a steel strand of the same weight.

▼ Spiders secrete silk fibers made of a structural protein containing β pleated sheets, which allow the spider web to stretch and recoil.

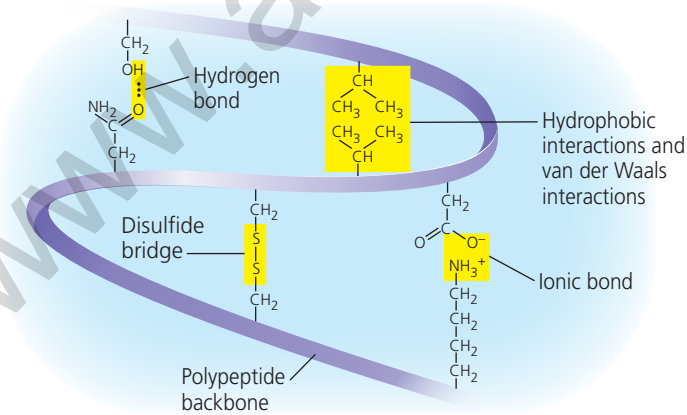
Tertiary Structure

Three-dimensional shape stabilized by interactions between side chains



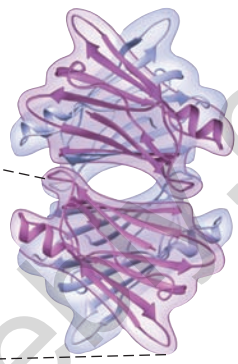
Superimposed on the patterns of secondary structure is a protein's tertiary structure, shown above in a ribbon model of the transthyretin polypeptide. While secondary structure involves interactions between backbone constituents, **tertiary structure** is the overall shape of a polypeptide resulting from interactions between the side chains (R groups) of the various amino acids. One type of interaction that contributes to tertiary structure is called—somewhat misleadingly—a **hydrophobic interaction**. As a polypeptide folds into its functional shape, amino acids with hydrophobic (nonpolar) side chains usually end up in clusters at the core of the protein, out of contact with water. Thus, a "hydrophobic interaction" is actually caused by the exclusion of nonpolar substances by water molecules. Once nonpolar amino acid side chains are close together, van der Waals interactions help hold them together. Meanwhile, hydrogen bonds between polar side chains and ionic bonds between positively and negatively charged side chains also help stabilize tertiary structure. These are all weak interactions in the aqueous cellular environment, but their cumulative effect helps give the protein a unique shape.

Covalent bonds called **disulfide bridges** may further reinforce the shape of a protein. Disulfide bridges form where two cysteine monomers, which have sulfhydryl groups (—SH) on their side chains (see Figure 4.9), are brought close together by the folding of the protein. The sulfur of one cysteine bonds to the sulfur of the second, and the disulfide bridge (—S—S—) rivets parts of the protein together (see the yellow lines in Figure 5.16a). All of these different kinds of interactions can contribute to the tertiary structure of a protein, as shown here in a small part of a hypothetical protein:



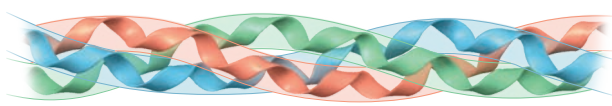
Quaternary Structure

Association of two or more polypeptides (some proteins only)

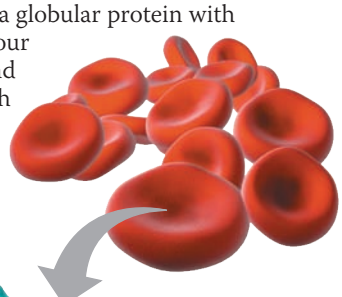


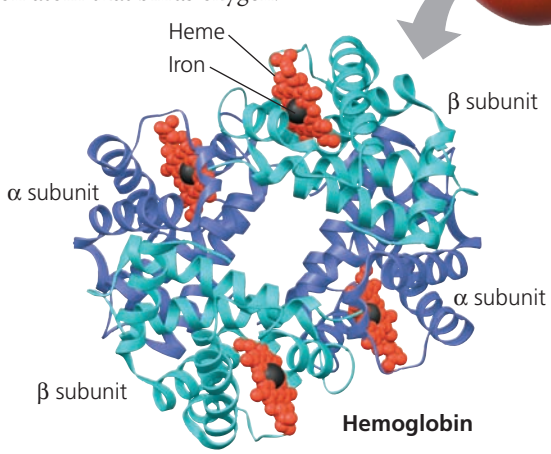
Some proteins consist of two or more polypeptide chains aggregated into one functional macromolecule. **Quaternary structure** is the overall protein structure that results from the aggregation of these polypeptide subunits. For example, shown above is the complete globular transthyretin protein, made up of its four polypeptides.

Another example is collagen, shown below, which is a fibrous protein that has three identical helical polypeptides intertwined into a larger triple helix, giving the long fibers great strength. This suits collagen fibers to their function as the girders of connective tissue in skin, bone, tendons, ligaments, and other body parts. (Collagen accounts for 40% of the protein in a human body.)

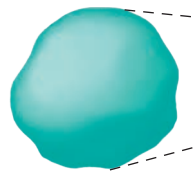
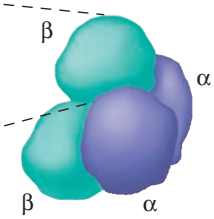
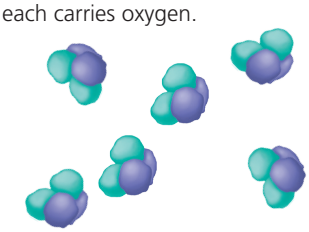
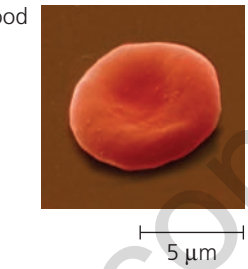
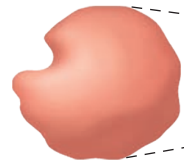
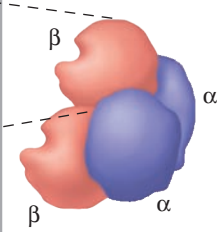
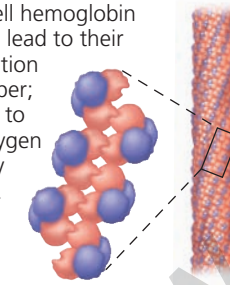
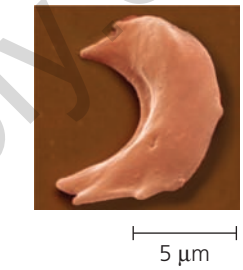


Hemoglobin, the oxygen-binding protein of red blood cells shown below, is another example of a globular protein with quaternary structure. It consists of four polypeptide subunits, two of one kind (α) and two of another kind (β). Both α and β subunits consist primarily of α-helical secondary structure. Each subunit has a nonpolypeptide component, called heme, with an iron atom that binds oxygen.





Hemoglobin

	Primary Structure	Secondary and Tertiary Structures	Quaternary Structure	Function	Red Blood Cell Shape
Normal hemoglobin	1 Val 2 His 3 Leu 4 Thr 5 Pro 6 Glu 7 Glu	Normal β subunit 	Normal hemoglobin 	Normal hemoglobin proteins do not associate with one another; each carries oxygen. 	Normal red blood cells are full of individual hemoglobin proteins. 
Sickle-cell hemoglobin	1 Val 2 His 3 Leu 4 Thr 5 Pro 6 Val 7 Glu	Sickle-cell β subunit 	Sickle-cell hemoglobin 	Hydrophobic interactions between sickle-cell hemoglobin proteins lead to their aggregation into a fiber; capacity to carry oxygen is greatly reduced. 	Fibers of abnormal hemoglobin deform red blood cell into sickle shape. 

▲ Figure 5.19 A single amino acid substitution in a protein causes sickle-cell disease.

MAKE CONNECTIONS *Considering the chemical characteristics of the amino acids valine and glutamic acid (see Figure 5.14), propose a possible explanation for the dramatic effect on protein function that occurs when valine is substituted for glutamic acid.*

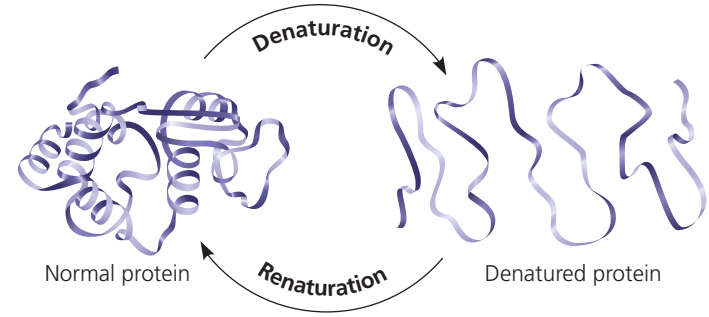
Sickle-Cell Disease: A Change in Primary Structure

Even a slight change in primary structure can affect a protein's shape and ability to function. For instance, **sickle-cell disease**, an inherited blood disorder, is caused by the substitution of one amino acid (valine) for the normal one (glutamic acid) at a particular position in the primary structure of hemoglobin, the protein that carries oxygen in red blood cells. Normal red blood cells are disk-shaped, but in sickle-cell disease, the abnormal hemoglobin molecules tend to aggregate into chains, deforming some of the cells into a sickle shape **(Figure 5.19)**. A person with the disease has periodic "sickle-cell crises" when the angular cells clog tiny blood vessels, impeding blood flow. The toll taken on such patients is a dramatic example of how a simple change in protein structure can have devastating effects on protein function.

What Determines Protein Structure?

You've learned that a unique shape endows each protein with a specific function. But what are the key factors determining protein structure? You already know most of the answer: A polypeptide chain of a given amino acid sequence can be arranged into a three-dimensional shape determined by the interactions responsible for secondary and tertiary structure. This folding normally occurs as the protein is being synthesized in the crowded environment within a cell, aided by other proteins. However, protein structure also depends on the physical and chemical conditions of the protein's environment. If the pH, salt concentration, temperature, or other aspects of its environment are altered, the weak chemical bonds and interactions within a protein may be destroyed, causing the protein to unravel and lose its native shape, a change called **denaturation (Figure 5.20)**. Because it is misshapen, the denatured protein is biologically inactive.

Most proteins become denatured if they are transferred from an aqueous environment to a nonpolar solvent, such as ether or chloroform; the polypeptide chain refolds so that its hydrophobic regions face outward toward the solvent. Other denaturation agents include chemicals that disrupt



▲ Figure 5.20 Denaturation and renaturation of a protein. High temperatures or various chemical treatments will denature a protein, causing it to lose its shape and hence its ability to function. If the denatured protein remains dissolved, it may renature when the chemical and physical aspects of its environment are restored to normal.

the hydrogen bonds, ionic bonds, and disulfide bridges that maintain a protein's shape. Denaturation can also result from excessive heat, which agitates the polypeptide chain enough to overpower the weak interactions that stabilize the structure. The white of an egg becomes opaque during cooking because the denatured proteins are insoluble and solidify. This also explains why excessively high fevers can be fatal: Proteins in the blood tend to denature at very high body temperatures.

When a protein in a test-tube solution has been denatured by heat or chemicals, it can sometimes return to its functional shape when the denaturing agent is removed. (Sometimes this is not possible: For example, a fried egg will not become liquefied when placed back into the refrigerator!) We can conclude that the information for building specific shape is intrinsic to the protein's primary structure. The sequence of amino acids determines the protein's shape—where an α helix can form, where β pleated sheets can exist, where disulfide bridges are located, where ionic bonds can form, and so on. But how does protein folding occur in the cell?

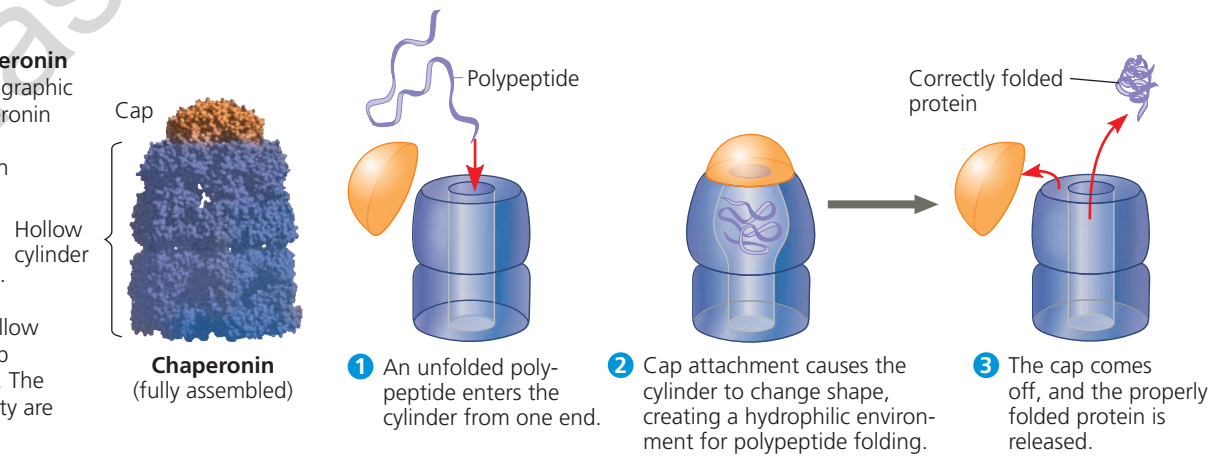
Protein Folding in the Cell

Biochemists now know the amino acid sequence for more than 24 million proteins, with about 1 million added each month, and the three-dimensional shape for more than 25,000. Researchers have tried to correlate the primary structure of many proteins with their three-dimensional structure to discover the rules of protein folding. Unfortunately, however, the protein-folding process is not that simple. Most proteins probably go through several intermediate structures on their way to a stable shape, and looking at the mature structure does not reveal the stages of folding required to achieve that form. However, biochemists have developed methods for tracking a protein through such stages.

Crucial to the folding process are **chaperonins** (also called chaperone proteins), protein molecules that assist in the proper folding of other proteins **(Figure 5.21)**. Chaperonins do not specify the final structure of a polypeptide. Instead, they keep the new polypeptide segregated from disruptive chemical conditions in the cytoplasmic environment while it folds spontaneously. The chaperonin shown in Figure 5.21, from the bacterium *E. coli*, is a giant multiprotein complex shaped like a hollow cylinder. The cavity provides a shelter for folding polypeptides, and recent research suggests that minute amounts of water are present, ensuring a hydrophilic environment that aids the folding process. Molecular systems have been identified that interact with chaperonins and check whether proper folding has occurred. Such systems either refold the misfolded proteins correctly or mark them for destruction.

Misfolding of polypeptides is a serious problem in cells that has come under increasing scrutiny by medical researchers. Many diseases—such as cystic fibrosis, Alzheimer's, Parkinson's, and mad cow disease—are associated with an accumulation of misfolded proteins. In fact, misfolded versions of the transthyretin protein featured in Figure 5.18 have been implicated in several diseases, including one form of senile dementia.

Even when scientists have a correctly folded protein in hand, determining its exact three-dimensional structure is not simple, for a single protein molecule has thousands of atoms. The first 3-D structures were worked out in the late 1950s for hemoglobin and a related protein called myoglobin. The method that made these feats possible was **X-ray crystallography**, which has since been used to determine the 3-D structure of many other proteins. In a recent example, Roger Kornberg and his colleagues at Stanford University used this method to elucidate the structure of RNA polymerase, an enzyme that plays a crucial role in the expression of genes **(Figure 5.22)**. Another method for analyzing protein structure is nuclear magnetic resonance (NMR) spectroscopy, which does not require protein crystallization. A still newer approach employs bioinformatics (see Concept 5.6) to predict the 3-D structure of polypeptides from their amino acid sequence. X-ray crystallography, NMR spectroscopy, and bioinformatics are complementary approaches to understanding protein structure and function.

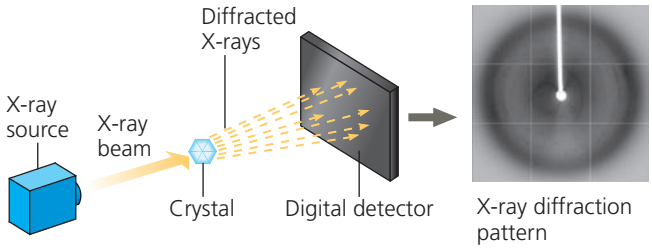


▶ Figure 5.21 A chaperonin in action. The computer graphic (left) shows a large chaperonin protein complex from the bacterium *E. coli*. It has an interior space that provides a shelter for the proper folding of newly made polypeptides. The complex consists of two proteins: One is a hollow cylinder; the other is a cap that can fit on either end. The steps of chaperonin activity are shown at the right.

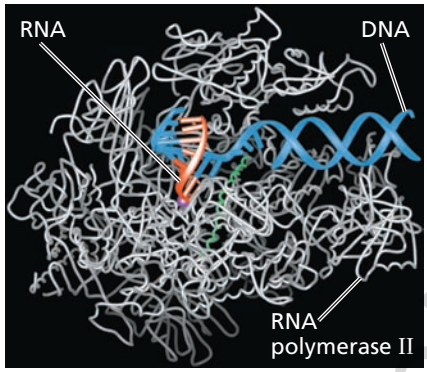
▼ Figure 5.22 **Inquiry**

What can the 3-D shape of the enzyme RNA polymerase II tell us about its function?

Experiment In 2006, Roger Kornberg was awarded the Nobel Prize in Chemistry for using X-ray crystallography to determine the 3-D shape of RNA polymerase II, which binds to the DNA double helix and synthesizes RNA. After crystallizing a complex of all three components, Kornberg and his colleagues aimed an X-ray beam through the crystal. The atoms of the crystal diffracted (bent) the X-rays into an orderly array that a digital detector recorded as a pattern of spots called an X-ray diffraction pattern.



Results Using data from X-ray diffraction patterns, as well as the amino acid sequence determined by chemical methods, the researchers built a 3-D model of the complex with the help of computer software.



Conclusion Analysis of the model led to a hypothesis about the functions of different regions of RNA polymerase II. For example, the region above the DNA may act as a clamp that holds the nucleic acids in place. (You'll learn more about this enzyme in Chapter 17.)

Source: A. L. Gnatt et al., Structural basis of transcription: an RNA polymerase II elongation complex at 3.3Å, *Science* 292:1876–1882 (2001). Computer graphic copyright © 2001 by AAAS. Reprinted with permission.

WHAT IF? *Looking at the model, can you identify any elements of secondary structure?*

CONCEPT CHECK 5.4

1. What parts of a polypeptide participate in the bonds that hold together secondary structure? Tertiary structure?
2. Thus far in the chapter, the Greek letters α and β have been used to specify at least three different pairs of structures. Name and briefly describe them.
3. **WHAT IF?** Where would you expect a polypeptide region rich in the amino acids valine, leucine, and isoleucine to be located in a folded polypeptide? Explain.

For suggested answers, see Appendix A.

CONCEPT 5.5

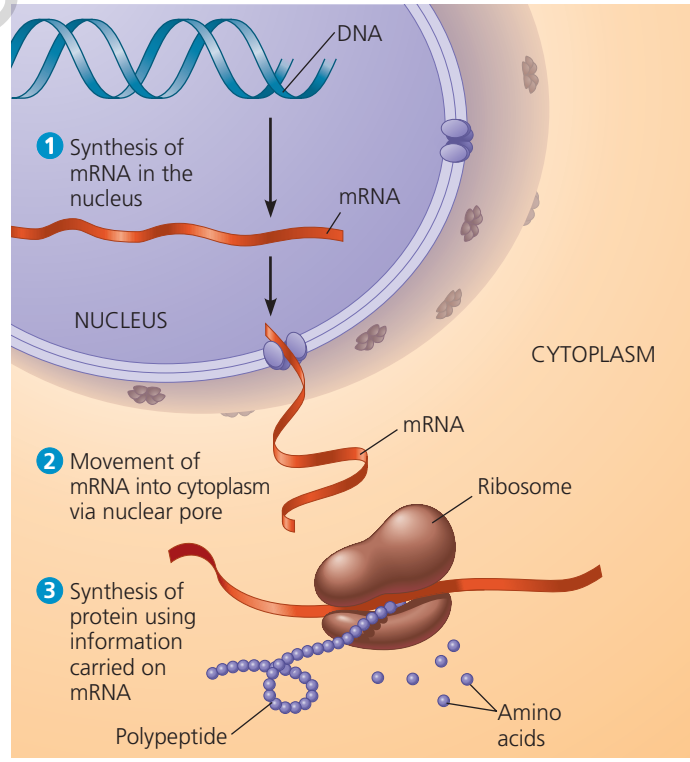
Nucleic acids store, transmit, and help express hereditary information

If the primary structure of polypeptides determines a protein's shape, what determines primary structure? The amino acid sequence of a polypeptide is programmed by a discrete unit of inheritance known as a **gene**. Genes consist of DNA, which belongs to the class of compounds called nucleic acids. **Nucleic acids** are polymers made of monomers called nucleotides.

The Roles of Nucleic Acids

The two types of nucleic acids, **deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**, enable living organisms to reproduce their complex components from one generation to the next. Unique among molecules, DNA provides directions for its own replication. DNA also directs RNA synthesis and, through RNA, controls protein synthesis; this entire process is called **gene expression** (Figure 5.23).

DNA is the genetic material that organisms inherit from their parents. Each chromosome contains one long DNA molecule, usually carrying several hundred or more genes. When a cell reproduces itself by dividing, its DNA molecules are copied and passed along from one generation of



▲ Figure 5.23 **Gene expression: DNA → RNA → protein.** In a eukaryotic cell, DNA in the nucleus programs protein production in the cytoplasm by dictating synthesis of messenger RNA (mRNA).

cells to the next. Encoded in the structure of DNA is the information that programs all the cell's activities. The DNA, however, is not directly involved in running the operations of the cell, any more than computer software by itself can read the bar code on a box of cereal. Just as a scanner is needed to read a bar code, proteins are required to implement genetic programs. The molecular hardware of the cell—the tools for biological functions—consists mostly of proteins. For example, the oxygen carrier in red blood cells is the protein hemoglobin that you saw earlier (see Figure 5.17), not the DNA that specifies its structure.

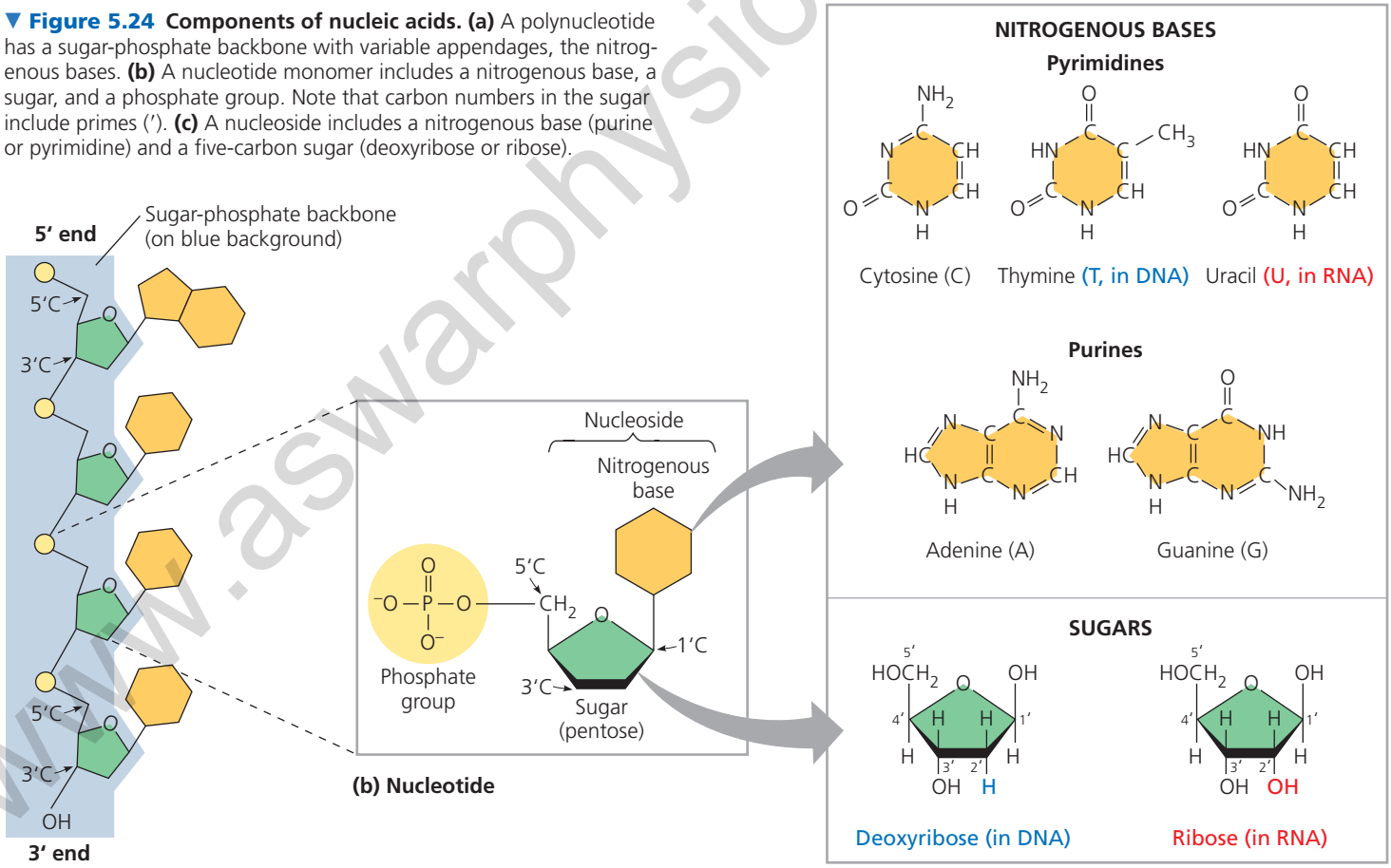
How does RNA, the other type of nucleic acid, fit into gene expression, the flow of genetic information from DNA to proteins? Each gene along a DNA molecule directs synthesis of a type of RNA called *messenger RNA* (*mRNA*). The mRNA molecule interacts with the cell's protein-synthesizing machinery to direct production of a polypeptide, which folds into all or part of a protein. We can summarize the flow of genetic information as DNA → RNA → protein (see Figure 5.23). The sites of protein synthesis are cellular structures called ribosomes. (In the Unit 1 interview before Chapter 2, Venki Ramakrishnan describes how the structure of ribosomes was determined by X-ray crystallography.) In a eukaryotic cell, ribosomes are in the region between the nucleus and the plasma membrane (the cytoplasm), but DNA resides in the nucleus. Messenger RNA conveys genetic instructions for building proteins from the nucleus to the cytoplasm. Prokaryotic cells lack nuclei but still use mRNA to convey a message from the DNA to ribosomes and other cellular equipment that translate the coded information into amino acid sequences. In Chapter 18, you'll read about other functions of some recently discovered RNA molecules.

The Components of Nucleic Acids

Nucleic acids are macromolecules that exist as polymers called **polynucleotides** **(Figure 5.24a)**. As indicated by the name, each polynucleotide consists of monomers called **nucleotides**. A nucleotide, in general, is composed of three parts: a five-carbon sugar (a pentose), a nitrogen-containing (nitrogenous) base, and one or more phosphate groups **(Figure 5.24b)**. In a polynucleotide, each monomer has only one phosphate group. The portion of a nucleotide without any phosphate groups is called a *nucleoside*.

To build a nucleotide, let's first consider the nitrogenous bases **(Figure 5.24c)**. Each nitrogenous base has one or two rings that include nitrogen atoms. (They are called nitrogenous *bases* because the nitrogen atoms tend to take up

▼ Figure 5.24 Components of nucleic acids. (a) A polynucleotide has a sugar-phosphate backbone with variable appendages, the nitrogenous bases. **(b)** A nucleotide monomer includes a nitrogenous base, a sugar, and a phosphate group. Note that carbon numbers in the sugar include primes (′). **(c)** A nucleoside includes a nitrogenous base (purine or pyrimidine) and a five-carbon sugar (deoxyribose or ribose).



H^+ from solution, thus acting as bases.) There are two families of nitrogenous bases: pyrimidines and purines. A **pyrimidine** has one six-membered ring of carbon and nitrogen atoms. The members of the pyrimidine family are cytosine (C), thymine (T), and uracil (U). **Purines** are larger, with a six-membered ring fused to a five-membered ring. The purines are adenine (A) and guanine (G). The specific pyrimidines and purines differ in the chemical groups attached to the rings. Adenine, guanine, and cytosine are found in both DNA and RNA; thymine is found only in DNA and uracil only in RNA.

Now let's add the sugar to which the nitrogenous base is attached. In DNA the sugar is **deoxyribose**; in RNA it is **ribose** (see Figure 5.24c). The only difference between these two sugars is that deoxyribose lacks an oxygen atom on the second carbon in the ring; hence the name *deoxy*ribose.

So far, we have built a nucleoside (nitrogenous base plus sugar). To complete the construction of a nucleotide, we attach a phosphate group to the 5′ carbon of the sugar (see Figure 5.24b). The molecule is now a nucleoside monophosphate, more often called a nucleotide.

Nucleotide Polymers

The linkage of nucleotides into a polynucleotide involves a dehydration reaction. (You will learn the details in Chapter 16). In the polynucleotide, adjacent nucleotides are joined by a phosphodiester linkage, which consists of a phosphate group that links the sugars of two nucleotides. This bonding results in a repeating pattern of sugar-phosphate units called the *sugar-phosphate backbone* (see Figure 5.24a). (Note that the nitrogenous bases are not part of the backbone.) The two free ends of the polymer are distinctly different from each other. One end has a phosphate attached to a 5′ carbon, and the other end has a hydroxyl group on a 3′ carbon; we refer to these as the *5′ end* and the *3′ end*, respectively. We can say that a polynucleotide has a built-in directionality along its sugar-phosphate backbone, from 5′ to 3′, somewhat like a one-way street. All along this sugar-phosphate backbone are appendages consisting of the nitrogenous bases.

The sequence of bases along a DNA (or mRNA) polymer is unique for each gene and provides very specific information to the cell. Because genes are hundreds to thousands of nucleotides long, the number of possible base sequences is effectively limitless. A gene's meaning to the cell is encoded in its specific sequence of the four DNA bases. For example, the sequence 5′-AGGTAACTT-3′ means one thing, whereas the sequence 5′-CGCTTTAAC-3′ has a different meaning. (Entire genes, of course, are much longer.) The linear order of bases in a gene specifies the amino acid sequence—the primary structure—of a protein, which in turn specifies that protein's three-dimensional structure and its function in the cell.

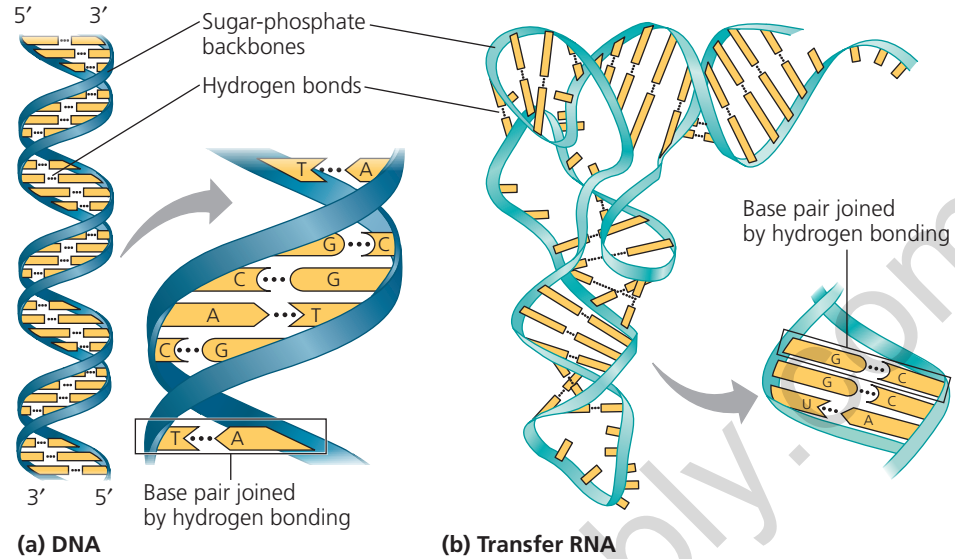
The Structures of DNA and RNA Molecules

DNA molecules have two polynucleotides, or "strands," that wind around an imaginary axis, forming a **double helix** **(Figure 5.25a)**. The two sugar-phosphate backbones run in opposite 5′ → 3′ directions from each other; this arrangement is referred to as **antiparallel**, somewhat like a divided highway. The sugar-phosphate backbones are on the outside of the helix, and the nitrogenous bases are paired in the interior of the helix. The two strands are held together by hydrogen bonds between the paired bases (see Figure 5.25a). Most DNA molecules are very long, with thousands or even millions of base pairs. For example, the one long DNA double helix in a eukaryotic chromosome includes many genes, each one a particular segment of the molecule.

In base pairing, only certain bases in the double helix are compatible with each other. Adenine (A) in one strand always pairs with thymine (T) in the other, and guanine (G) always pairs with cytosine (C). Reading the sequence of bases along one strand of the double helix would tell us the sequence of bases along the other strand. If a stretch of one strand has the base sequence 5′-AGGTCCG-3′, then the base-pairing rules tell us that the same stretch of the other strand must have the sequence 3′-TCCAGGC-5′. The two strands of the double helix are *complementary*, each the predictable counterpart of the other. It is this feature of DNA that makes it possible to generate two identical copies of each DNA molecule in a cell that is preparing to divide. When the cell divides, the copies are distributed to the daughter cells, making them genetically identical to the parent cell. Thus, the structure of DNA accounts for its function of transmitting genetic information whenever a cell reproduces.

RNA molecules, by contrast, exist as single strands. Complementary base pairing can occur, however, between regions of two RNA molecules or even between two stretches of nucleotides in the *same* RNA molecule. In fact, base pairing within an RNA molecule allows it to take on the particular three-dimensional shape necessary for its function. Consider, for example, the type of RNA called *transfer RNA* (*tRNA*), which brings amino acids to the ribosome during the synthesis of a polypeptide. A tRNA molecule is about 80 nucleotides in length. Its functional shape results from base pairing between nucleotides where complementary stretches of the molecule can run antiparallel to each other **(Figure 5.25b)**.

Note that in RNA, adenine (A) pairs with uracil (U); thymine (T) is not present in RNA. Another difference between RNA and DNA is that DNA almost always exists as a double helix, whereas RNA molecules are more variable in shape. RNAs are very versatile molecules, and many biologists believe RNA may have preceded DNA as the carrier of genetic information in early forms of life (see Concept 25.1).



▶ **Figure 5.25 The structures of DNA and tRNA molecules. (a)** The DNA molecule is usually a double helix, with the sugar-phosphate backbones of the antiparallel polynucleotide strands (symbolized here by blue ribbons) on the outside of the helix. Hydrogen bonds between pairs of nitrogenous bases hold the two strands together. As illustrated here with symbolic shapes for the bases, adenine (A) can pair only with thymine (T), and guanine (G) can pair only with cytosine (C). Each DNA strand in this figure is the structural equivalent of the polynucleotide diagrammed in Figure 5.24a. **(b)** A tRNA molecule has a roughly L-shaped structure, with complementary base pairing of antiparallel stretches of RNA. In RNA, A pairs with U.

CONCEPT CHECK **5.5**

1. **DRAW IT** Go to Figure 5.24a and, for the top three nucleotides, number all the carbons in the sugars, circle the nitrogenous bases, and star the phosphates.
2. **DRAW IT** In a DNA double helix, a region along one DNA strand has this sequence of nitrogenous bases: 5′-TAGGCCT-3′. Copy this sequence, and write down its complementary strand, clearly indicating the 5′ and 3′ ends of the complementary strand.

For suggested answers, see Appendix A.

CONCEPT 5.6

Genomics and proteomics have transformed biological inquiry and applications

Experimental work in the first half of the 20th century established the role of DNA as the bearer of genetic information, passed from generation to generation, that specified the functioning of living cells and organisms. Once the structure of the DNA molecule was described in 1953, and the linear sequence of nucleotide bases was understood to specify the amino acid sequence of proteins, biologists sought to "decode" genes by learning their base sequences.

The first chemical techniques for *DNA sequencing*, or determining the sequence of nucleotides along a DNA strand, one by one, were developed in the 1970s. Researchers began to study gene sequences, gene by gene, and the more they learned, the more questions they had: How was expression of genes regulated? Genes and their protein products clearly interacted with each other, but how? What was the function, if any, of the DNA that is not part of genes? To fully understand the genetic functioning of a living organism, the entire sequence of the full complement of DNA, the organism's *genome*, would be most enlightening. In spite of the apparent impracticality of this idea, in the late 1980s several prominent biologists put forth an audacious proposal to launch a project that would sequence the entire human genome—all 3 billion bases of it! This endeavor began in 1990 and was effectively completed in the early 2000s.

An unplanned but profound side benefit of this project—the Human Genome Project—was the rapid development of faster and less expensive methods of sequencing. This trend has continued apace: The cost for sequencing 1 million bases in 2001, well over $5,000, has decreased to less than $0.10 in 2012. And a human genome, the first of which took over 10 years to sequence, could be completed at today's pace in just a few days. The number of genomes that have been fully sequenced has burgeoned, generating reams of data and prompting development of *bioinformatics*, the use of computer software and other computational tools that can handle and analyze these large data sets.

The reverberations of these developments have transformed the study of biology and related fields. Biologists often look at problems by analyzing large sets of genes or even comparing whole genomes of different species, an approach called **genomics**. A similar analysis of large sets of proteins, including their sequences, is called **proteomics**. (Protein sequences can be determined either by using biochemical techniques or by translating the DNA sequences that code for them.) These approaches permeate all fields of biology, some examples of which are shown in **Figure 5.26**.

Perhaps the most significant impact of genomics and proteomics on the field of biology as a whole has been their contributions to our understanding of evolution. In addition to confirming evidence for evolution from the study of fossils and characteristics of currently existing species, genomics has helped us tease out relationships among different groups of organisms that had not been resolved by previous types of evidence, and thus infer evolutionary history.

▼ Figure 5.26

MAKE CONNECTIONS

Contributions of Genomics and Proteomics to Biology

Nucleotide sequencing and the analysis of large sets of genes and proteins can be done rapidly and inexpensively due to advances in technology and information processing. Taken together, genomics and proteomics have advanced our understanding of biology across many different fields.

Paleontology

New DNA sequencing techniques have allowed decoding of minute quantities of DNA found in ancient tissues from our extinct relatives, the Neanderthals (*Homo neanderthalensis*). Sequencing the Neanderthal genome has informed our understanding of their physical appearance as well as their relationship with modern humans. *See Figure 34.49.*

Medical Science

Identifying the genetic basis for human diseases like cancer helps researchers focus their search for potential future treatments. Currently, sequencing the sets of genes expressed in an individual's tumor can allow a more targeted approach to treating the cancer, a type of "personalized medicine." *See Figures 12.20 and 18.27.*

Evolution

A major aim of evolutionary biology is to understand the relationships among species, both living and extinct. For example, genome sequence comparisons have identified the hippopotamus as the land mammal sharing the most recent common ancestor with whales. *See Figure 22.20.*

Hippopotamus

Short-finned pilot whale

Conservation Biology

The tools of molecular genetics and genomics are increasingly used by ecologists to identify which species of animals and plants are killed illegally. In one case, genomic sequences of DNA from illegal shipments of elephant tusks were used to track down poachers and pinpoint the territory where they were operating. *See Figure 56.9.*

Species Interactions

Over 90% of all plant species exist in a mutually beneficial partnership with fungi that are associated with the plants' roots. Genome sequencing and analysis of gene expression in several plant-fungal pairs promise major advances in our understanding of such interactions and may have implications for agricultural practices. *(See the Scientific Skills Exercise in Chapter 31.)*

MAKE CONNECTIONS *Considering the examples provided here, describe how the approaches of genomics and proteomics help us to address a variety of biological questions.*

DNA and Proteins as Tape Measures of Evolution

EVOLUTION We are accustomed to thinking of shared traits, such as hair and milk production in mammals, as evidence of shared ancestry. Because DNA carries heritable information in the form of genes, sequences of genes and their protein products document the hereditary background of an organism. The linear sequences of nucleotides in DNA molecules are passed from parents to offspring; these sequences determine the amino acid sequences of proteins. As a result, siblings have greater similarity in their DNA and proteins than do unrelated individuals of the same species.

Given our evolutionary view of life, we can extend this concept of "molecular genealogy" to relationships between species: We would expect two species that appear to be closely related based on anatomical evidence (and possibly fossil evidence) to also share a greater proportion of their DNA and protein sequences than do less closely related species. In fact, that is the case. An example is the comparison of the β polypeptide chain of human hemoglobin with the corresponding hemoglobin polypeptide in other vertebrates. In this chain of 146 amino acids, humans and gorillas differ in just 1 amino acid, while humans and frogs, more distantly related, differ in 67 amino acids. In the **Scientific Skills Exercise**, you can apply this sort of reasoning to additional species. And this conclusion holds true as well when comparing whole genomes: The human genome is 95–98% identical to that of the chimpanzee, but only roughly 85% identical to that of the mouse, a more distant evolutionary relative. Molecular biology has added a new tape measure to the toolkit biologists use to assess evolutionary kinship.

CONCEPT CHECK 5.6

1. How would sequencing the entire genome of an organism help scientists to understand how that organism functioned?
2. Given the function of DNA, why would you expect two species with very similar traits to also have very similar genomes?

For suggested answers, see Appendix A.

SCIENTIFIC SKILLS EXERCISE

Analyzing Polypeptide Sequence Data

▶ Human  ▶ Rhesus monkey  ▶ Gibbon 

Are Rhesus Monkeys or Gibbons More Closely Related to Humans? DNA and polypeptide sequences from closely related species are more similar to each other than are sequences from more distantly related species. In this exercise, you will look at amino acid sequence data for the β polypeptide chain of hemoglobin, often called β-globin. You will then interpret the data to hypothesize whether the monkey or the gibbon is more closely related to humans.

How Such Experiments Are Done Researchers can isolate the polypeptide of interest from an organism and then determine the amino acid sequence. More frequently, the DNA of the relevant gene is sequenced, and the amino acid sequence of the polypeptide is deduced from the DNA sequence of its gene.

Data from the Experiments In the data below, the letters give the sequence of the 146 amino acids in β-globin from humans, rhesus monkeys, and gibbons. Because a complete sequence would not fit on one line here, the sequences are broken into three segments. The sequences for the three different species are aligned so that you can compare them easily. For example, you can see that for all three species, the first amino acid is V (valine) and the 146th amino acid is H (histidine).

Species	Alignment of Amino Acid Sequences of β-globin
Human	1 VHLTPEEKSA VTALWGKVNV DEVGGEALGR LLVVYPWTQR FFESFGDLST
Monkey	1 VHLTPEEKNA VTTLWGKVNV DEVGGEALGR LLLVYPWTQR FFESFGDLSS
Gibbon	1 VHLTPEEKSA VTALWGKVNV DEVGGEALGR LLVVYPWTQR FFESFGDLST
Human	51 PDAVMGNPKV KAHGKKVLGA FSDGLAHLDN LKGTFATLSE LHCDKLHVDP
Monkey	51 PDAVMGNPKV KAHGKKVLGA FSDGLNHLDN LKGTFAQLSE LHCDKLHVDP
Gibbon	51 PDAVMGNPKV KAHGKKVLGA FSDGLAHLDN LKGTFAQLSE LHCDKLHVDP
Human	101 ENFRLLGNVL VCVLAHHFGK EFTPPVQAAY QKVVAGVANA LAHKYH
Monkey	101 ENFKLLGNVL VCVLAHHFGK EFTPQVQAAY QKVVAGVANA LAHKYH
Gibbon	101 ENFRLLGNVL VCVLAHHFGK EFTPQVQAAY QKVVAGVANA LAHKYH

Interpret the Data

1. Scan the monkey and gibbon sequences, letter by letter, circling any amino acids that do not match the human sequence. (a) How many amino acids differ between the monkey and the human sequences? (b) Between the gibbon and human?
2. For each nonhuman species, what percent of its amino acids are identical to the human sequence of β-globin?
3. Based on these data alone, state a hypothesis for which of these two species is more closely related to humans. What is your reasoning?
4. What other evidence could you use to support your hypothesis?

MB A version of this Scientific Skills Exercise can be assigned in MasteringBiology.

Data from Human: http://www.ncbi.nlm.nih.gov/protein/AAA21113.1; rhesus monkey: http://www.ncbi.nlm.nih.gov/protein/122634; gibbon: http://www.ncbi.nlm.nih.gov/protein/122616

5 Chapter Review

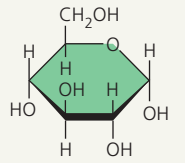
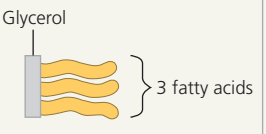
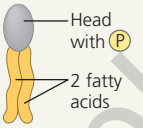
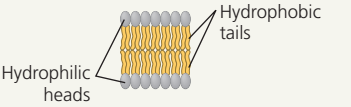
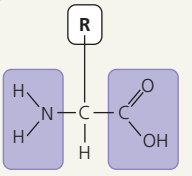
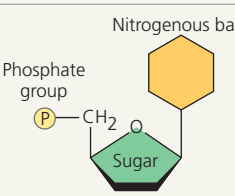
SUMMARY OF KEY CONCEPTS

CONCEPT 5.1

Macromolecules are polymers, built from monomers (pp. 67–68)

- Large carbohydrates (polysaccharides), proteins, and nucleic acids are **polymers**, which are chains of **monomers**. The components of lipids vary. Monomers form larger molecules by **dehydration reactions**, in which water molecules are released. Polymers can disassemble by the reverse process, **hydrolysis**. An immense variety of polymers can be built from a small set of monomers.

? *What is the fundamental basis for the differences between large carbohydrates, proteins, and nucleic acids?*

Large Biological Molecules	Components	Examples	Functions
CONCEPT 5.2 **Carbohydrates serve as fuel and building material (pp. 68–72)** ? *Compare the composition, structure, and function of starch and cellulose. What role do starch and cellulose play in the human body?*	 Monosaccharide monomer	**Monosaccharides:** glucose, fructose	Fuel; carbon sources that can be converted to other molecules or combined into polymers
		Disaccharides: lactose, sucrose	
		Polysaccharides: • Cellulose (plants) • Starch (plants) • Glycogen (animals) • Chitin (animals and fungi)	• Strengthens plant cell walls • Stores glucose for energy • Stores glucose for energy • Strengthens exoskeletons and fungal cell walls
CONCEPT 5.3 **Lipids are a diverse group of hydrophobic molecules (pp. 72–75)** ? *Why are lipids not considered to be polymers or macromolecules?*		**Triacylglycerols** (fats or oils): glycerol + 3 fatty acids	Important energy source  
		**Phospholipids:** glycerol + phosphate group + 2 fatty acids	Lipid bilayers of membranes 
	 Steroid backbone	**Steroids:** four fused rings with attached chemical groups	• Component of cell membranes (cholesterol) • Signaling molecules that travel through the body (hormones)
CONCEPT 5.4 **Proteins include a diversity of structures, resulting in a wide range of functions (pp. 75–84)** ? *Explain the basis for the great diversity of proteins.*	 Amino acid monomer (20 types)	• Enzymes • Structural proteins • Hormones • Receptor proteins • Motor proteins • Defensive proteins	• Catalyze chemical reactions • Provide structural support • Coordinate organismal responses • Receive signals from outside cell • Function in cell movement • Protect against disease
CONCEPT 5.5 **Nucleic acids store, transmit, and help express hereditary information (pp. 84–87)** ? *What role does complementary base pairing play in the functions of nucleic acids?*	 Nucleotide monomer	**DNA:**  • Sugar = deoxyribose • Nitrogenous bases = C, G, A, T • Usually double-stranded	Stores hereditary information
		RNA:  • Sugar = ribose • Nitrogenous bases = C, G, A, U • Usually single-stranded	Various functions in gene expression, including carrying instructions from DNA to ribosomes

CONCEPT 5.6

Genomics and proteomics have transformed biological inquiry and applications (pp. 87–89)

- Recent technological advances in DNA sequencing have given rise to **genomics**, an approach that analyzes large sets of genes or whole genomes, and **proteomics**, a similar approach for large sets of proteins. Bioinformatics is the use of computational tools and computer software to analyze these large data sets.
- The more closely two species are related evolutionarily, the more similar their DNA sequences are. DNA sequence data confirms models of evolution based on fossils and anatomical evidence.

? *Given the sequences of a particular gene in fruit flies, fish, mice, and humans, predict the relative similarity of the human sequence to that of each of the other species.*

TEST YOUR UNDERSTANDING

LEVEL 1: KNOWLEDGE/COMPREHENSION

1. Which of the following categories includes all others in the list?
   a. monosaccharide
   b. polysaccharide
   c. starch
   d. carbohydrate
2. The enzyme amylase can break glycosidic linkages between glucose monomers only if the monomers are in the α form. Which of the following could amylase break down?
   a. glycogen, starch, and amylopectin
   b. glycogen and cellulose
   c. cellulose and chitin
   d. starch, chitin, and cellulose
3. Which of the following is true of *unsaturated* fats?
   a. They are more common in animals than in plants.
   b. They have double bonds in the carbon chains of their fatty acids.
   c. They generally solidify at room temperature.
   d. They contain more hydrogen than do saturated fats having the same number of carbon atoms.
4. The structural level of a protein *least* affected by a disruption in hydrogen bonding is the
   a. primary level.
   b. secondary level.
   c. tertiary level.
   d. quaternary level.
5. Enzymes that break down DNA catalyze the hydrolysis of the covalent bonds that join nucleotides together. What would happen to DNA molecules treated with these enzymes?
   a. The two strands of the double helix would separate.
   b. The phosphodiester linkages of the polynucleotide backbone would be broken.
   c. The pyrimidines would be separated from the deoxyribose sugars.
   d. All bases would be separated from the deoxyribose sugars.

LEVEL 2: APPLICATION/ANALYSIS

6. The molecular formula for glucose is $C_6H_{12}O_6$. What would be the molecular formula for a polymer made by linking ten glucose molecules together by dehydration reactions?
   a. $C_{60}H_{120}O_{60}$
   b. $C_{60}H_{102}O_{51}$
   c. $C_{60}H_{100}O_{50}$
   d. $C_{60}H_{111}O_{51}$
7. Which of the following pairs of base sequences could form a short stretch of a normal double helix of DNA?
   a. 5′-AGCT-3′ with 5′-TCGA-3′
   b. 5′-GCGC-3′ with 5′-TATA-3′
   c. 5′-ATGC-3′ with 5′-GCAT-3′
   d. All of these pairs are correct.
8. Construct a table that organizes the following terms, and label the columns and rows.

Monosaccharides	Polypeptides	Phosphodiester linkages
Fatty acids	Triacylglycerols	Peptide bonds
Amino acids	Polynucleotides	Glycosidic linkages
Nucleotides	Polysaccharides	Ester linkages

9. **DRAW IT** Copy the polynucleotide strand in Figure 5.24a and label the bases G, T, C, and T, starting from the 5′ end. Assuming this is a DNA polynucleotide, now draw the complementary strand, using the same symbols for phosphates (circles), sugars (pentagons), and bases. Label the bases. Draw arrows showing the 5′ → 3′ direction of each strand. Use the arrows to make sure the second strand is antiparallel to the first. *Hint*: After you draw the first strand vertically, turn the paper upside down; it is easier to draw the second strand from the 5′ toward the 3′ direction as you go from top to bottom.

LEVEL 3: SYNTHESIS/EVALUATION

10. **EVOLUTION CONNECTION**
    Comparisons of amino acid sequences can shed light on the evolutionary divergence of related species. If you were comparing two living species, would you expect all proteins to show the same degree of divergence? Why or why not?
11. **SCIENTIFIC INQUIRY**
    Suppose you are a research assistant in a lab studying DNA-binding proteins. You have been given the amino acid sequences of all the proteins encoded by the genome of a certain species and have been asked to find candidate proteins that could bind DNA. What type of amino acids would you expect to see in the DNA-binding regions of such proteins? Why?
12. **WRITE ABOUT A THEME: ORGANIZATION**
    Proteins, which have diverse functions in a cell, are all polymers of the same kinds of monomers—amino acids. Write a short essay (100–150 words) that discusses how the structure of amino acids allows this one type of polymer to perform so many functions.
13. **SYNTHESIZE YOUR KNOWLEDGE**

    Given that the function of egg yolk is to nourish and support the developing chick, explain why egg yolks are so high in fat, protein, and cholesterol.

For selected answers, see Appendix A.

MasteringBiology®

Students Go to **MasteringBiology** for assignments, the eText, and the Study Area with practice tests, animations, and activities.

Instructors Go to **MasteringBiology** for automatically graded tutorials and questions that you can assign to your students, plus Instructor Resources.

UNIT 2

THE CELL

AN INTERVIEW WITH

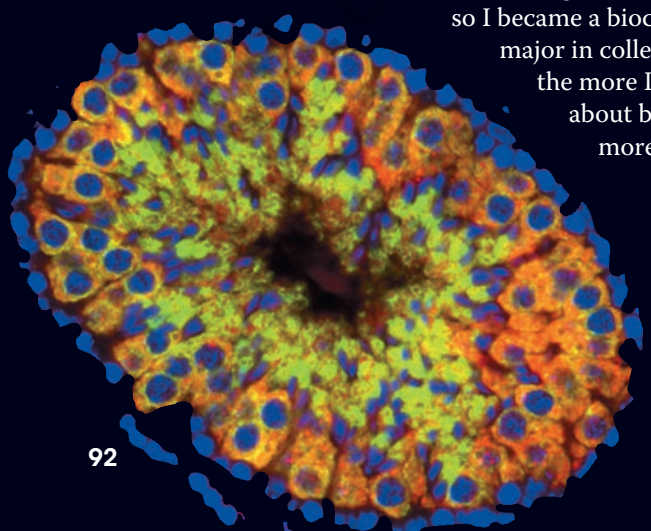
Haifan Lin

Born in China, Haifan Lin majored in biochemistry at Fudan University in Shanghai. He then earned a Ph.D. in genetics and development from Cornell University and was a postdoctoral fellow at the Carnegie Institution of Washington (now the Carnegie Institution for Science). There, he started using the fruit fly (*Drosophila melanogaster*) as a model to explore fundamental questions in stem cells. Dr. Lin then spent 12 years as a faculty member at Duke University, broadening his study of stem cells by working on mammalian models and clinical applications. He is one of the discoverers of Piwi-interacting RNAs, a finding that was heralded by *Science* magazine as a Discovery of the Year in 2006. That same year, Dr. Lin moved to Yale University, where he founded and now directs the Yale Stem Cell Center.

How did you get interested in science?

As a child I liked to build things, so I imagined myself a ship builder or an architect, something like that. I didn't get attracted to biology until high school. Genetic engineering had become a very fashionable term in China, and I thought, "That's cool. That's the engineering of life." I was more attracted by the word "engineering" than "genetics." However, people told me it was important to have a solid biochemistry foundation in order to become a genetic engineer, so I became a biochemistry major in college. And the more I learned about biology, the more I loved it.

"If we hadn't started by working on basic cell biology in *Drosophila*, I don't think we could have found this connection to cancer so quickly."



◀ Cross section of a tubule in the testis of a mouse, showing the Piwi protein (red-orange in this fluorescence micrograph).

What did you study in graduate school and as a postdoc?

At Cornell, I thought about the very first cell division of the embryo. To me, it's literally the first step of life—the division of a fertilized egg. Working on a cell division process with developmental significance was really intellectually rewarding. For my postdoc, I felt that I should continue to study cell division with developmental consequences but expand to a different cell type, so I turned to stem cells.

What is a stem cell?

Stem cells are really the mother of all cells. Embryonic stem cells lead to the development of all tissues—the entire adult body. Tissue stem cells are responsible for the generation and/or maintenance of a specific tissue. All stem cells share a unique property—they can self-renew (reproduce) as well as give rise to more specialized cells. In theory, stem cells are immortal; they are like a fountain of youth that goes on and on.

How do you study stem cells?

To study stem cells, you have to identify the cell unambiguously, so cell biology is the first step. Cell biology defines a problem, describes the phenomenon, and provides the biological context for further mechanistic studies. It's crucially important. Then we move on to genetics, and, in my style of research, biochemistry usually comes as a third component.

What is the most interesting thing you have discovered about stem cells?

Using the genetic approach, we found a fruit fly (***Drosophila***) gene that encodes a protein called Piwi. The Piwi protein is also required in mammalian stem cells that make the testis (see micrograph). Piwi proteins bind to a kind of small RNA we and others independently discovered and called Piwi-interacting RNAs (or piRNAs). One of the wonderful things about working with fruit flies is that as soon as you identify new genes in flies and confirm that they function in stem cells, you can immediately look in humans to see whether these same genes become overactivated in cancer. It turns out the human Piwi gene is expressed at least sixfold more in a common kind of testicular cancer. We published the Piwi gene family in 1998, and amazingly, in 2002, we already had the results on this human cancer. If we hadn't started by working on basic cell biology in ***Drosophila***, I don't think we could have found this connection to cancer so quickly.

MB For an extended interview and video clip, go to the Study Area in MasteringBiology.

6

A Tour of the Cell

KEY CONCEPTS

6.1 Biologists use microscopes and the tools of biochemistry to study cells

6.2 Eukaryotic cells have internal membranes that compartmentalize their functions

6.3 The eukaryotic cell's genetic instructions are housed in the nucleus and carried out by the ribosomes

6.4 The endomembrane system regulates protein traffic and performs metabolic functions in the cell

6.5 Mitochondria and chloroplasts change energy from one form to another

6.6 The cytoskeleton is a network of fibers that organizes structures and activities in the cell

6.7 Extracellular components and connections between cells help coordinate cellular activities

▲ **Figure 6.1** How do your cells help you learn about biology?

The Fundamental Units of Life

Cells are as fundamental to the living systems of biology as the atom is to chemistry. Many different types of cells are working for you right now. The contraction of muscle cells moves your eyes as you read this sentence. **Figure 6.1** shows extensions from a nerve cell (orange) making contact with muscle cells (red). The words on the page are translated into signals that nerve cells carry to your brain, where they are passed on to other nerve cells. As you study, you are making cell connections like these that solidify memories and permit learning to occur.

All organisms are made of cells. In the hierarchy of biological organization, the cell is the simplest collection of matter that can be alive. Indeed, many forms of life exist as single-celled organisms. Larger, more complex organisms, including plants and animals, are multicellular; their bodies are cooperatives of many kinds of specialized cells that could not survive for long on their own. Even when cells are arranged into higher levels of organization, such as tissues and organs, the cell remains the organism's basic unit of structure and function.

All cells are related by their descent from earlier cells. During the long evolutionary history of life on Earth, cells have been modified in many different ways. But although cells can differ substantially from one another, they share common features. In this chapter, we'll first examine the tools and techniques that allow us to understand cells, then tour the cell and become acquainted with its components.