

# 1

## LIFE BEGINS WITH CELLS



A single ~200 micrometer ( $\mu\text{m}$ ) cell, the human egg, with sperm, which are also single cells. From the union of an egg and sperm will arise the 10 trillion cells of a human body.

[Photo Researchers, Inc.]

Like ourselves, the individual cells that form our bodies can grow, reproduce, process information, respond to stimuli, and carry out an amazing array of chemical reactions. These abilities define life. We and other multicellular organisms contain billions or trillions of cells organized into complex structures, but many organisms consist of a single cell. Even simple unicellular organisms exhibit all the hallmark properties of life, indicating that the cell is the fundamental unit of life. As the twenty-first century opens, we face an explosion of new data about the components of cells, what structures they contain, how they touch and influence each other. Still, an immense amount remains to be learned, particularly about how information flows through cells and how they decide on the most appropriate ways to respond.

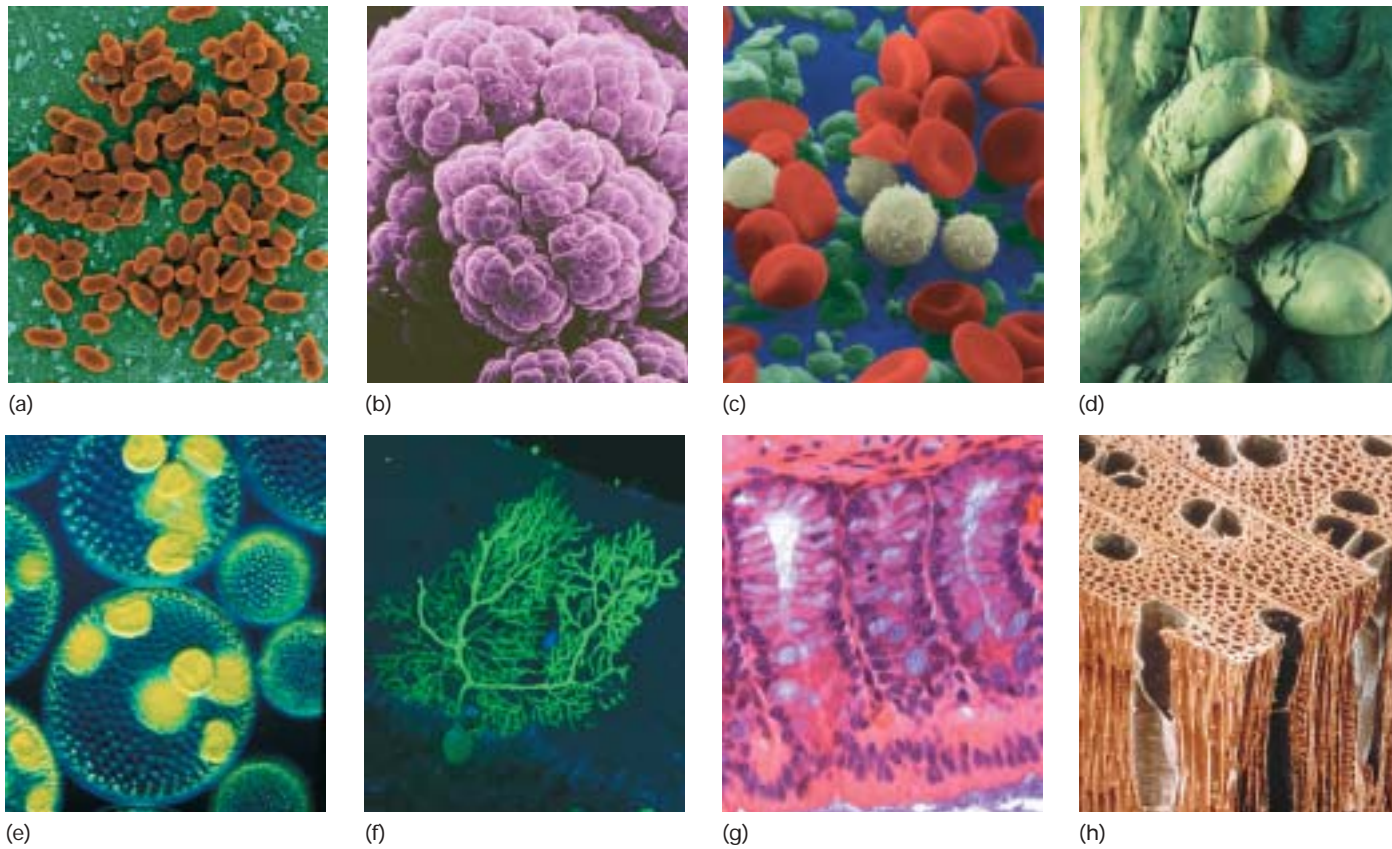
Molecular cell biology is a rich, integrative science that brings together biochemistry, biophysics, molecular biology, microscopy, genetics, physiology, computer science, and developmental biology. Each of these fields has its own emphasis and style of experimentation. In the following chapters, we will describe insights and experimental approaches drawn from all of these fields, gradually weaving the multifaceted story of the birth, life, and death of cells. We start in this prologue chapter by introducing the diversity of cells, their basic constituents and critical functions, and what we can learn from the various ways of studying cells.

### 1.1 The Diversity and Commonality of Cells

Cells come in an amazing variety of sizes and shapes (Figure 1-1). Some move rapidly and have fast-changing structures, as we can see in movies of amoebae and rotifers. Others are largely stationary and structurally stable. Oxygen kills some cells but is an absolute requirement for others. Most cells in multicellular organisms are intimately involved with other cells. Although some unicellular organisms live in isolation, others form colonies or live in close association with other types of organisms, such as the bacteria that help plants to extract nitrogen from the air or the bacteria that live in our intestines and help us digest food. Despite these and numerous

#### OUTLINE

- 1.1 The Diversity and Commonality of Cells
- 1.2 The Molecules of a Cell
- 1.3 The Work of Cells
- 1.4 Investigating Cells and Their Parts
- 1.5 A Genome Perspective on Evolution



▲ **FIGURE 1-1 Cells come in an astounding assortment of shapes and sizes.** Some of the morphological variety of cells is illustrated in these photographs. In addition to morphology, cells differ in their ability to move, internal organization (prokaryotic versus eukaryotic cells), and metabolic activities. (a) Eubacteria; note dividing cells. These are *Lactococcus lactis*, which are used to produce cheese such as Roquefort, Brie, and Camembert. (b) A mass of archaeobacteria (*Methanosarcina*) that produce their energy by converting carbon dioxide and hydrogen gas to methane. Some species that live in the rumen of cattle give rise to >150 liters of methane gas/day. (c) Blood cells, shown in false color. The red blood cells are oxygen-bearing erythrocytes, the white blood cells (leukocytes) are part of the immune system and fight infection, and the green cells are platelets that provide substances to make blood clot at a wound. (d) Large single cells: fossilized dinosaur eggs. (e) A colonial single-celled green alga, *Volvox aureus*. The large spheres are made up of many individual cells, visible as blue or green dots. The yellow masses inside are daughter colonies, each made up of many cells. (f) A single

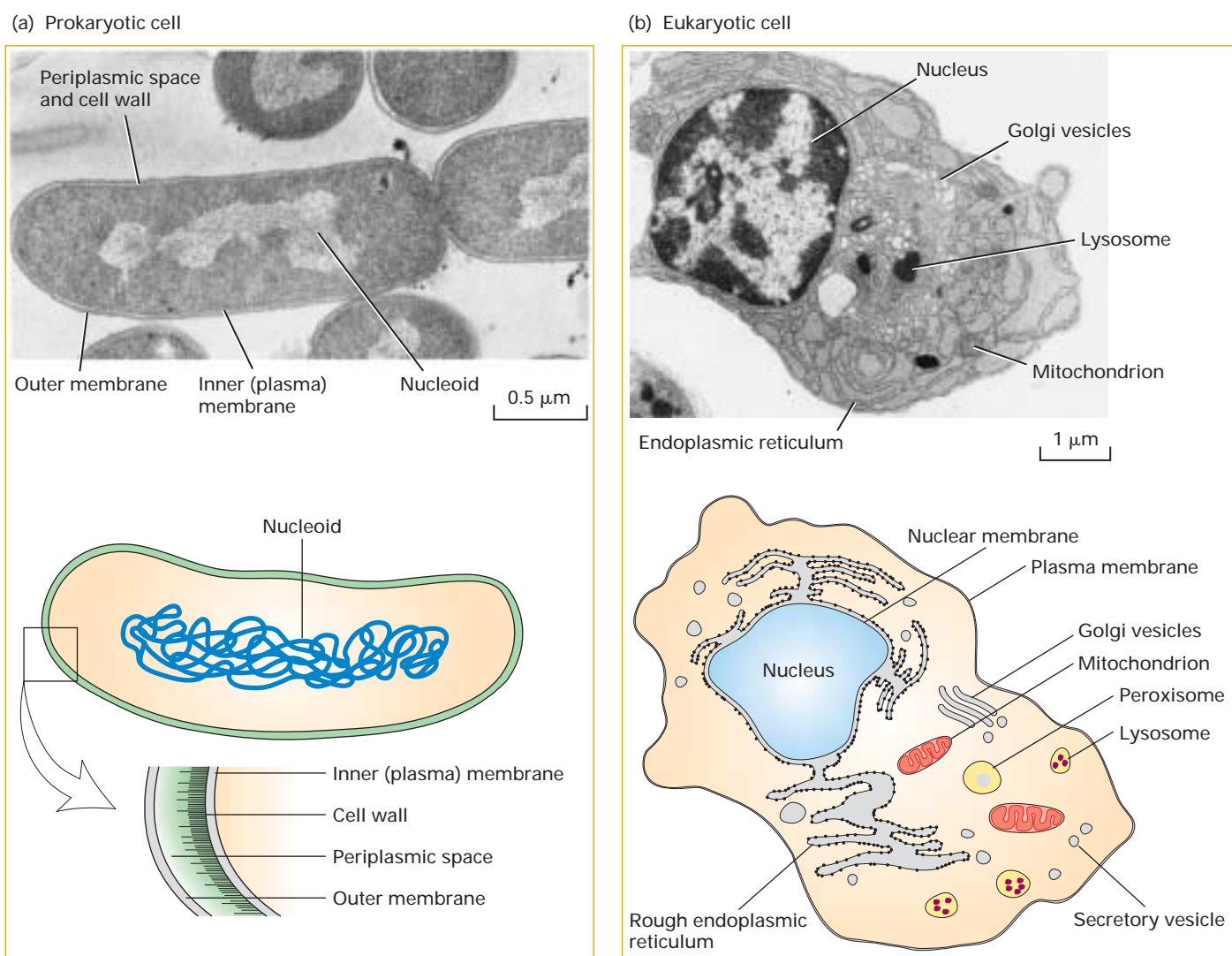
Purkinje neuron of the cerebellum, which can form more than a hundred thousand connections with other cells through the branched network of dendrites. The cell was made visible by introduction of a fluorescent protein; the cell body is the bulb at the bottom. (g) Cells can form an epithelial sheet, as in the slice through intestine shown here. Each finger-like tower of cells, a villus, contains many cells in a continuous sheet. Nutrients are transferred from digested food through the epithelial sheet to the blood for transport to other parts of the body. New cells form continuously near the bases of the villi, and old cells are shed from the top. (h) Plant cells are fixed firmly in place in vascular plants, supported by a rigid cellulose skeleton. Spaces between the cells are joined into tubes for transport of water and food. [Part (a) Gary Gaugler/ Photo Researchers, Inc. Part (b) Ralph Robinson/ Visuals Unlimited, Inc. Part (c) NIH/Photo Researchers, Inc. Part (d) John D. Cunningham/Visuals Unlimited, Inc. Part (e) Carolina Biological/Visuals Unlimited, Inc. Part (f) Helen M. Blau, Stanford University. Part (g) Jeff Gordon, Washington University School of Medicine. Part (h) Richard Kessel and C. Shih/Visuals Unlimited, Inc.]

other differences, all cells share certain structural features and carry out many complicated processes in basically the same way. As the story of cells unfolds throughout this book, we will focus on the molecular basis of both the differences and similarities in the structure and function of various cells.

### All Cells Are Prokaryotic or Eukaryotic

The biological universe consists of two types of cells—prokaryotic and eukaryotic. Prokaryotic cells consist of a sin-

gle closed compartment that is surrounded by the **plasma membrane**, lacks a defined **nucleus**, and has a relatively simple internal organization (Figure 1-2a). All **prokaryotes** have cells of this type. Bacteria, the most numerous prokaryotes, are single-celled organisms; the cyanobacteria, or blue-green algae, can be unicellular or filamentous chains of cells. Although bacterial cells do not have membrane-bounded compartments, many proteins are precisely localized in their aqueous interior, or **cytosol**, indicating the presence of internal organization. A single *Escherichia coli* bacterium has a dry weight of about



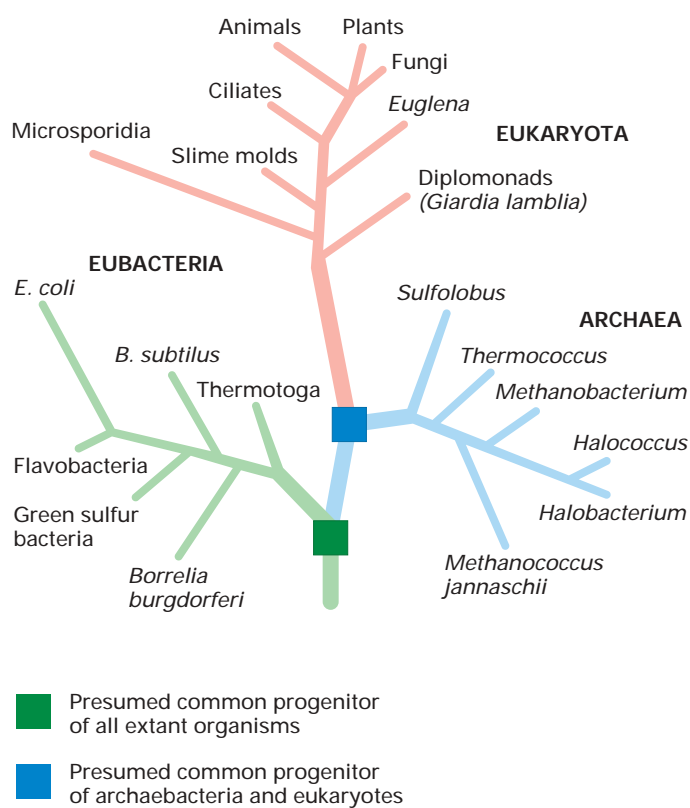
▲ **FIGURE 1-2 Prokaryotic cells have a simpler internal organization than eukaryotic cells.** (a) Electron micrograph of a thin section of *Escherichia coli*, a common intestinal bacterium. The nucleoid, consisting of the bacterial DNA, is not enclosed within a membrane. *E. coli* and some other bacteria are surrounded by two membranes separated by the periplasmic space. The thin cell wall is adjacent to the inner membrane. (b) Electron micrograph of a plasma cell, a type of white blood cell that secretes antibodies. Only a single membrane (the plasma membrane) surrounds the cell, but the interior contains many membrane-limited compartments, or organelles. The defining

characteristic of eukaryotic cells is segregation of the cellular DNA within a defined nucleus, which is bounded by a double membrane. The outer nuclear membrane is continuous with the rough endoplasmic reticulum, a factory for assembling proteins. Golgi vesicles process and modify proteins, mitochondria generate energy, lysosomes digest cell materials to recycle them, peroxisomes process molecules using oxygen, and secretory vesicles carry cell materials to the surface to release them. [Part (a) courtesy of I. D. J. Burdett and R. G. E. Murray. Part (b) from P. C. Cross and K. L. Mercer, 1993, *Cell and Tissue Ultrastructure: A Functional Perspective*, W. H. Freeman and Company.]

$25 \times 10^{-14}$  g. Bacteria account for an estimated 1–1.5 kg of the average human's weight. The estimated number of bacteria on earth is  $5 \times 10^{30}$ , weighing a total of about  $10^{12}$  kg. Prokaryotic cells have been found 7 miles deep in the ocean and 40 miles up in the atmosphere; they are quite adaptable! The carbon stored in bacteria is nearly as much as the carbon stored in plants.

Eukaryotic cells, unlike prokaryotic cells, contain a defined membrane-bound nucleus and extensive internal mem-

branes that enclose other compartments, the **organelles** (Figure 1-2b). The region of the cell lying between the plasma membrane and the nucleus is the **cytoplasm**, comprising the cytosol (aqueous phase) and the organelles. **Eukaryotes** comprise all members of the plant and animal kingdoms, including the fungi, which exist in both multicellular forms (molds) and unicellular forms (yeasts), and the protozoans (*proto*, primitive; *zoan*, animal), which are exclusively unicellular. Eukaryotic cells are commonly about 10–100  $\mu\text{m}$  across,



▲ **FIGURE 1-3 All organisms from simple bacteria to complex mammals probably evolved from a common, single-celled progenitor.** This family tree depicts the evolutionary relations among the three major lineages of organisms. The structure of the tree was initially ascertained from morphological criteria: Creatures that look alike were put close together. More recently the sequences of DNA and proteins have been examined as a more information-rich criterion for assigning relationships. The greater the similarities in these macromolecular sequences, the more closely related organisms are thought to be. The trees based on morphological comparisons and the fossil record generally agree well with those those based on molecular data. Although all organisms in the eubacterial and archaean lineages are prokaryotes, archaea are more similar to eukaryotes than to eubacteria ("true" bacteria) in some respects. For instance, archaean and eukaryotic genomes encode homologous histone proteins, which associate with DNA; in contrast, bacteria lack histones. Likewise, the RNA and protein components of archaean ribosomes are more like those in eukaryotes than those in bacteria.

generally much larger than bacteria. A typical human fibroblast, a connective tissue cell, might be about 15  $\mu\text{m}$  across with a volume and dry weight some thousands of times those of an *E. coli* bacterial cell. An amoeba, a single-celled protozoan, can be more than 0.5 mm long. An ostrich egg begins as a single cell that is even larger and easily visible to the naked eye.

All cells are thought to have evolved from a common progenitor because the structures and molecules in all cells have

so many similarities. In recent years, detailed analysis of the DNA sequences from a variety of prokaryotic organisms has revealed two distinct types: the so-called "true" bacteria, or **eubacteria**, and **archaea** (also called archaeobacteria or archaeans). Working on the assumption that organisms with more similar genes evolved from a common progenitor more recently than those with more dissimilar genes, researchers have developed the evolutionary lineage tree shown in Figure 1-3. According to this tree, the archaea and the eukaryotes diverged from the true bacteria before they diverged from each other.

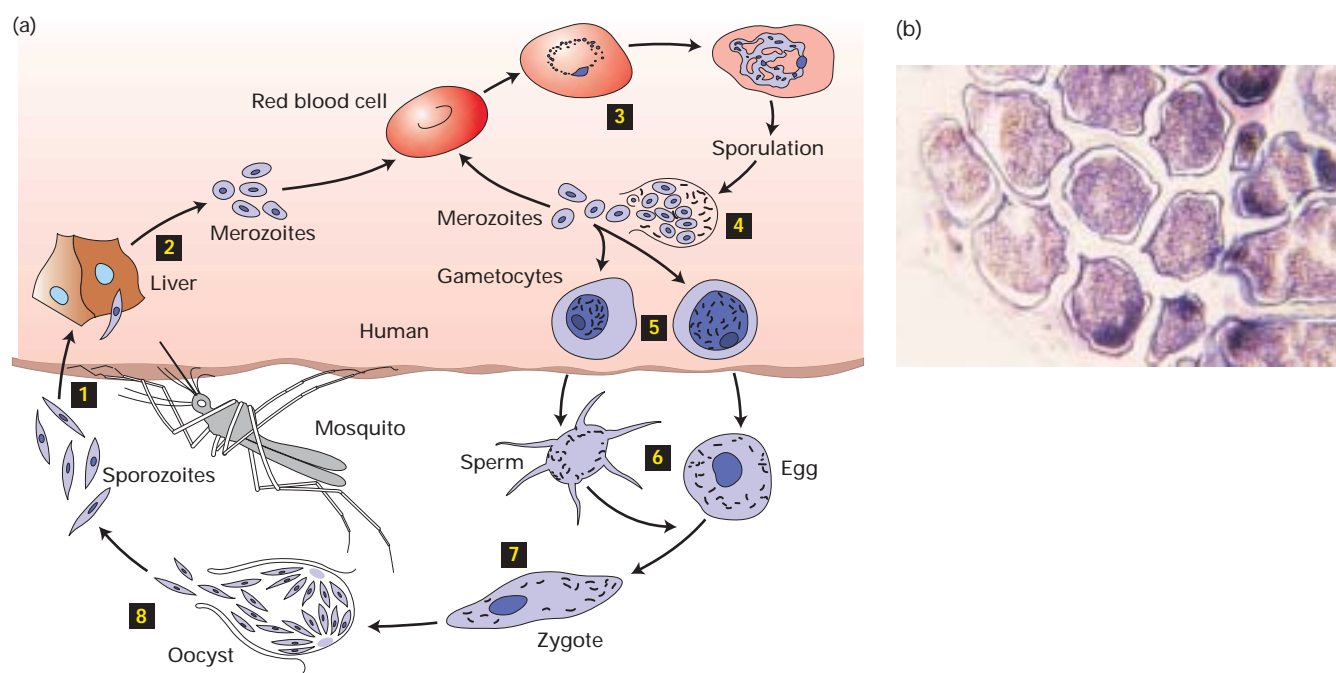
Many archaeans grow in unusual, often extreme, environments that may resemble ancient conditions when life first appeared on earth. For instance, halophiles ("salt loving") require high concentrations of salt to survive, and thermoacidophiles ("heat and acid loving") grow in hot (80° C) sulfur springs, where a pH of less than 2 is common. Still other archaeans live in oxygen-free milieus and generate methane ( $\text{CH}_4$ ) by combining water with carbon dioxide.

### Unicellular Organisms Help and Hurt Us

Bacteria and archaeobacteria, the most abundant single-celled organisms, are commonly 1–2  $\mu\text{m}$  in size. Despite their small size and simple architecture, they are remarkable biochemical factories, converting simple chemicals into complex biological molecules. Bacteria are critical to the earth's ecology, but some cause major diseases: bubonic plague (Black Death) from *Yersinia pestis*, strep throat from *Streptomyces*, tuberculosis from *Mycobacterium tuberculosis*, anthrax from *Bacillus anthracis*, cholera from *Vibrio cholerae*, food poisoning from certain types of *E. coli* and *Salmonella*.

Humans are walking repositories of bacteria, as are all plants and animals. We provide food and shelter for a staggering number of "bugs," with the greatest concentration in our intestines. Bacteria help us digest our food and in turn are able to reproduce. A common gut bacterium, *E. coli* is also a favorite experimental organism. In response to signals from bacteria such as *E. coli*, the intestinal cells form appropriate shapes to provide a niche where bacteria can live, thus facilitating proper digestion by the combined efforts of the bacterial and the intestinal cells. Conversely, exposure to intestinal cells changes the properties of the bacteria so that they participate more effectively in digestion. Such communication and response is a common feature of cells.

The normal, peaceful mutualism of humans and bacteria is sometimes violated by one or both parties. When bacteria begin to grow where they are dangerous to us (e.g., in the bloodstream or in a wound), the cells of our immune system fight back, neutralizing or devouring the intruders. Powerful antibiotic medicines, which selectively poison prokaryotic cells, provide rapid assistance to our relatively slow-developing immune response. Understanding the molecular biology of bacterial cells leads to an understanding of how bacteria are normally poisoned by antibiotics, how they become resistant to antibiotics, and what processes or structures present in bacterial but not human cells might be usefully targeted by new drugs.



▲ **FIGURE 1-4** *Plasmodium* organisms, the parasites that cause malaria, are single-celled protozoans with a remarkable life cycle. Many *Plasmodium* species are known, and they can infect a variety of animals, cycling between insect and vertebrate hosts. The four species that cause malaria in humans undergo several dramatic transformations within their human and mosquito hosts. (a) Diagram of the life cycle. Sporozoites enter a human host when an infected *Anopheles* mosquito bites a person **1**. They migrate to the liver where they develop into merozoites, which are released into the blood **2**. Merozoites differ substantially from sporozoites, so this transformation is a metamorphosis (Greek, “to transform” or “many shapes”). Circulating merozoites invade red blood cells (RBCs) and reproduce within them **3**. Proteins produced by some *Plasmodium* species move to the surface of infected RBCs, causing the cells to adhere to the walls of blood vessels. This prevents infected RBCs cells from circulating to the spleen where cells of the immune system would destroy the RBCs and the *Plasmodium* organisms they harbor. After growing and reproducing in RBCs for a period of time characteristic of each *Plasmodium* species, the merozoites suddenly burst forth in synchrony from large numbers of infected cells **4**. It is this

event that brings on the fevers and shaking chills that are the well-known symptoms of malaria. Some of the released merozoites infect additional RBCs, creating a cycle of production and infection. Eventually, some merozoites develop into male and female gametocytes **5**, another metamorphosis. These cells, which contain half the usual number of chromosomes, cannot survive for long unless they are transferred in blood to an *Anopheles* mosquito. In the mosquito's stomach, the gametocytes are transformed into sperm or eggs (gametes), yet another metamorphosis marked by development of long hairlike flagella on the sperm **6**. Fusion of sperm and eggs generates zygotes **7**, which implant into the cells of the stomach wall and grow into oocysts, essentially factories for producing sporozoites. Rupture of an oocyst releases thousands of sporozoites **8**; these migrate to the salivary glands, setting the stage for infection of another human host. (b) Scanning electron micrograph of mature oocysts and emerging sporozoites. Oocysts abut the external surface of stomach wall cells and are encased within a membrane that protects them from the host immune system. [Part (b) courtesy of R. E. Sinden.]

Like bacteria, protozoa are usually beneficial members of the food chain. They play key roles in the fertility of soil, controlling bacterial populations and excreting nitrogenous and phosphate compounds, and are key players in waste treatment systems—both natural and man-made. These unicellular eukaryotes are also critical parts of marine ecosystems, consuming large quantities of phytoplankton and harboring photosynthetic algae, which use sunlight to produce biologically useful energy forms and small fuel molecules.

However, some protozoa do give us grief: *Entamoeba histolytica* causes dysentery; *Trichomonas vaginalis*, vagini-

tis; and *Trypanosoma brucei*, sleeping sickness. Each year the worst of the protozoa, *Plasmodium falciparum* and related species, is the cause of more than 300 million new cases of malaria, a disease that kills 1.5 to 3 million people annually. These protozoans inhabit mammals and mosquitoes alternately, changing their morphology and behavior in response to signals in each of these environments. They also recognize receptors on the surfaces of the cells they infect. The complex life cycle of *Plasmodium* dramatically illustrates how a single cell can adapt to each new challenge it encounters (Figure 1-4). All of the transformations in cell type that

occur during the *Plasmodium* life cycle are governed by instructions encoded in the genetic material of this parasite and triggered by environmental inputs.

The other group of single-celled eukaryotes, the yeasts, also have their good and bad points, as do their multicellular cousins, the molds. Yeasts and molds, which collectively constitute the fungi, have an important ecological role in breaking down plant and animal remains for reuse. They also

make numerous antibiotics and are used in the manufacture of bread, beer, wine, and cheese. Not so pleasant are fungal diseases, which range from relatively innocuous skin infections, such as jock itch and athlete's foot, to life-threatening *Pneumocystis carinii* pneumonia, a common cause of death among AIDS patients.

### Even Single Cells Can Have Sex

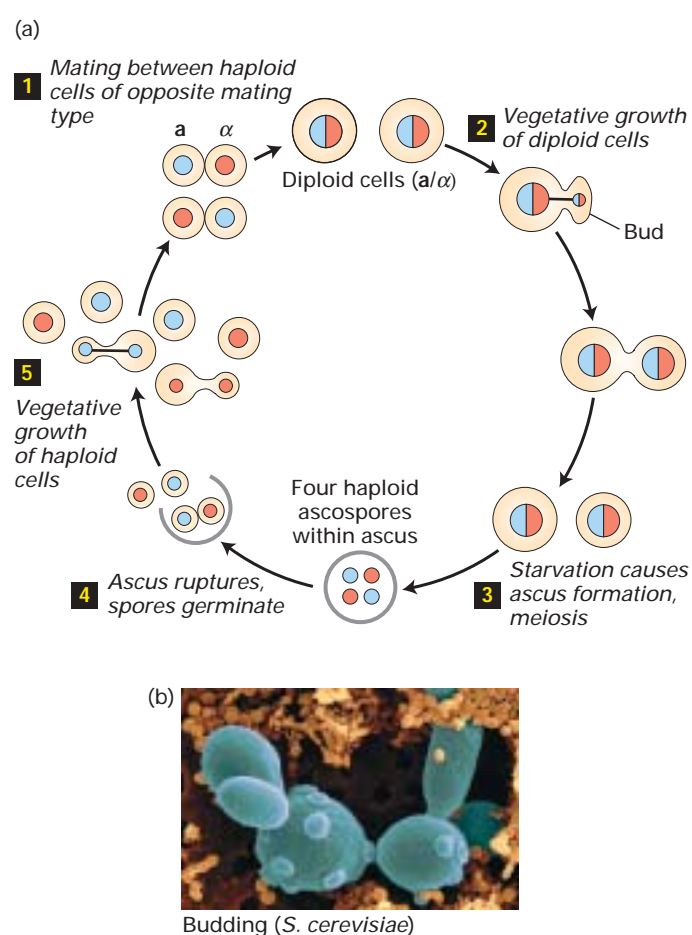
The common yeast used to make bread and beer, *Saccharomyces cerevisiae*, appears fairly frequently in this book because it has proven to be a great experimental organism. Like many other unicellular organisms, yeasts have two mating types that are conceptually like the male and female gametes (eggs and sperm) of higher organisms. Two yeast cells of opposite mating type can fuse, or mate, to produce a third cell type containing the genetic material from each cell (Figure 1-5). Such sexual life cycles allow more rapid changes in genetic inheritance than would be possible without sex, resulting in valuable adaptations while quickly eliminating detrimental mutations. That, and not just Hollywood, is probably why sex is so ubiquitous.

### Viruses Are the Ultimate Parasites

Virus-caused diseases are numerous and all too familiar: chicken pox, influenza, some types of pneumonia, polio, measles, rabies, hepatitis, the common cold, and many others. Smallpox, once a worldwide scourge, was eradicated by a decade-long global immunization effort beginning in the mid-1960s. Viral infections in plants (e.g., dwarf mosaic virus in corn) have a major economic impact on crop production. Planting of virus-resistant varieties, developed by traditional breeding methods and more recently by genetic engineering techniques, can reduce crop losses significantly. Most viruses have a rather limited host range, infecting certain bacteria, plants, or animals (Figure 1-6).

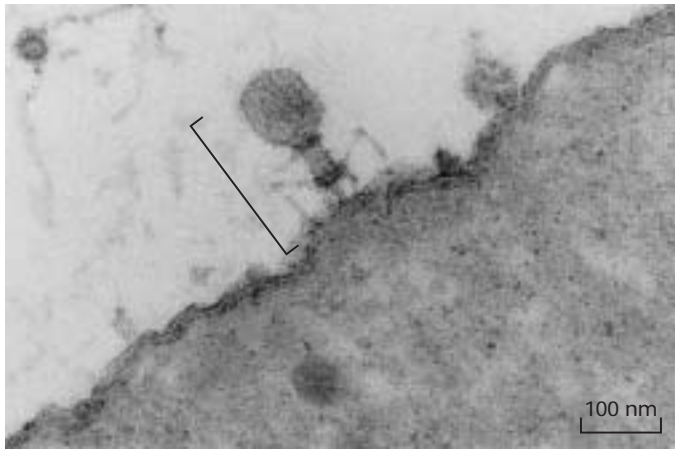
Because viruses cannot grow or reproduce on their own, they are not considered to be alive. To survive, a virus must infect a host cell and take over its internal machinery to synthesize viral proteins and in some cases to replicate the viral genetic material. When newly made viruses are released, the cycle starts anew. Viruses are much smaller than cells, on the order of 100 nanometer (nm) in diameter; in comparison, bacterial cells are usually  $>1000$  nm ( $1\text{ nm} = 10^{-9}$  meters). A virus is typically composed of a protein coat that encloses a core containing the genetic material, which carries the information for producing more viruses (Chapter 4). The coat protects a virus from the environment and allows it to stick to, or enter, specific host cells. In some viruses, the protein coat is surrounded by an outer membrane-like envelope.

The ability of viruses to transport genetic material into cells and tissues represents a medical menace and a medical opportunity. Viral infections can be devastatingly destructive, causing cells to break open and tissues to fall apart. However, many methods for manipulating cells depend upon using



▲ **FIGURE 1-5** The yeast *Saccharomyces cerevisiae* reproduces sexually and asexually. (a) Two cells that differ in mating type, called **a** and **α**, can mate to form an **a/α** cell **1**. The **a** and **α** cells are haploid, meaning they contain a single copy of each yeast chromosome, half the usual number. Mating yields a diploid **a/α** cell containing two copies of each chromosome. During vegetative growth, diploid cells multiply by mitotic budding, an asexual process **2**. Under starvation conditions, diploid cells undergo meiosis, a special type of cell division, to form haploid ascospores **3**. Rupture of an ascus releases four haploid spores, which can germinate into haploid cells **4**. These also can multiply asexually **5**. (b) Scanning electron micrograph of budding yeast cells. After each bud breaks free, a scar is left at the budding site so the number of previous buds can be counted. The orange cells are bacteria. [Part (b) M. Abbey/Visuals Unlimited, Inc.]

(a) T4 bacteriophage

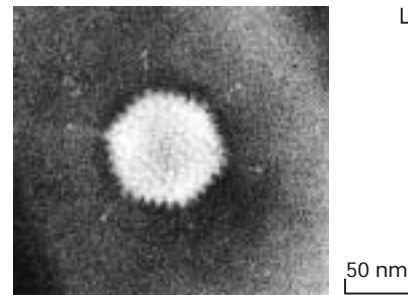


▲ **FIGURE 1-6 Viruses must infect a host cell to grow and reproduce.** These electron micrographs illustrate some of the structural variety exhibited by viruses. (a) T4 bacteriophage (bracket) attaches to a bacterial cell via a tail structure. Viruses that infect bacteria are called bacteriophages, or simply phages. (b) Tobacco mosaic virus causes a mottling of the leaves of

(b) Tobacco mosaic virus



(c) Adenovirus



infected tobacco plants and stunts their growth. (c) Adenovirus causes eye and respiratory tract infections in humans. This virus has an outer membranous envelope from which long glycoprotein spikes protrude. [Part (a) from A. Levine, 1991, *Viruses*, Scientific American Library, p. 20. Part (b) courtesy of R. C. Valentine. Part (c) courtesy of Robley C. Williams, University of California.]

viruses to convey genetic material into cells. To do this, the portion of the viral genetic material that is potentially harmful is replaced with other genetic material, including human genes. The altered viruses, or **vectors**, still can enter cells totting the introduced genes with them (Chapter 9). One day, diseases caused by defective genes may be treated by using viral vectors to introduce a normal copy of a defective gene into patients. Current research is dedicated to overcoming the considerable obstacles to this approach, such as getting the introduced genes to work at the right places and times.

### We Develop from a Single Cell

In 1827, German physician Karl von Baer discovered that mammals grow from eggs that come from the mother's ovary. Fertilization of an egg by a sperm cell yields a **zygote**, a visually unimpressive cell 200  $\mu\text{m}$  in diameter. Every human being begins as a zygote, which houses all the necessary instructions for building the human body containing about 100 trillion ( $10^{14}$ ) cells, an amazing feat. **Development** begins with the fertilized egg cell dividing into two, four, then eight cells, forming the very early embryo (Figure 1-7). Continued cell proliferation and then **differentiation** into distinct cell types gives rise to every tissue in the body. One initial cell, the fertilized egg (zygote), generates hundreds of different kinds of cells that differ in contents, shape, size, color, mobility, and surface composition. We will see how genes and signals control cell diversification in Chapters 15 and 22.

Making different kinds of cells—muscle, skin, bone, neuron, blood cells—is not enough to produce the human body. The cells must be properly arranged and organized into tis-

sues, organs, and appendages. Our two hands have the same kinds of cells, yet their different arrangements—in a mirror image—are critical for function. In addition, many cells exhibit distinct functional and/or structural asymmetries, a property often called **polarity**. From such polarized cells arise



(a)



(b)



(c)

◀ **FIGURE 1-7 The first few cell divisions of a fertilized egg set the stage for all subsequent development.** A developing mouse embryo is shown at (a) the two-cell, (b) four-cell, and (c) eight-cell stages. The embryo is surrounded by supporting membranes. The corresponding steps in human development occur during the first few days after fertilization. [Claude Edelman/Photo Researchers, Inc.]

asymmetric, polarized tissues such as the lining of the intestines and structures like hands and hearts. The features that make some cells polarized, and how they arise, also are covered in later chapters.

### Stem Cells, Cloning, and Related Techniques Offer Exciting Possibilities but Raise Some Concerns

Identical twins occur naturally when the mass of cells composing an early embryo divides into two parts, each of which develops and grows into an individual animal. Each cell in an eight-cell-stage mouse embryo has the potential to give rise to any part of the entire animal. Cells with this capability are referred to as *embryonic stem (ES) cells*. As we learn in Chapter 22, ES cells can be grown in the laboratory (cultured) and will develop into various types of differentiated cells under appropriate conditions.

The ability to make and manipulate mammalian embryos in the laboratory has led to new medical opportunities as well as various social and ethical concerns. *In vitro* fertilization, for instance, has allowed many otherwise infertile couples to have children. A new technique involves extraction of nuclei from defective sperm incapable of normally fertilizing an egg, injection of the nuclei into eggs, and implantation of the resulting fertilized eggs into the mother.

In recent years, nuclei taken from cells of adult animals have been used to produce new animals. In this procedure, the nucleus is removed from a body cell (e.g., skin or blood cell) of a donor animal and introduced into an unfertilized mammalian egg that has been deprived of its own nucleus. This manipulated egg, which is equivalent to a fertilized egg, is then implanted into a foster mother. The ability of such a donor nucleus to direct the development of an entire animal suggests that all the information required for life is retained in the nuclei of some adult cells. Since all the cells in an animal produced in this way have the genes of the single original donor cell, the new animal is a **clone** of the donor (Figure 1-8). Repeating the process can give rise to many clones. So far, however, the majority of embryos produced by this technique of *nuclear-transfer cloning* do not survive due to birth defects. Even those animals that are born live have shown abnormalities, including accelerated aging. The “rooting” of plants, in contrast, is a type of cloning that is readily accomplished by gardeners, farmers, and laboratory technicians.

The technical difficulties and possible hazards of nuclear-transfer cloning have not deterred some individuals from pursuing the goal of human cloning. However, cloning of humans per se has very limited scientific interest and is opposed by most scientists because of its high risk. Of greater scientific and medical interest is the ability to generate specific cell types starting from embryonic or adult stem cells. The scientific interest comes from learning the signals that can unleash the potential of the genes to form a certain cell type. The medical interest comes from the possibility of treating the nu-



▲ **FIGURE 1-8 Five genetically identical cloned sheep.** An early sheep embryo was divided into five groups of cells and each was separately implanted into a surrogate mother, much like the natural process of twinning. At an early stage the cells are able to adjust and form an entire animal; later in development the cells become progressively restricted and can no longer do so. An alternative way to clone animals is to replace the nuclei of multiple single-celled embryos with donor nuclei from cells of an adult sheep. Each embryo will be genetically identical to the adult from which the nucleus was obtained. Low percentages of embryos survive these procedures to give healthy animals, and the full impact of the techniques on the animals is not yet known. [Geoff Tompkinson/Science Photo Library/Photo Researchers, Inc.]

merous diseases in which particular cell types are damaged or missing, and of repairing wounds more completely.

## 1.2 The Molecules of a Cell

Molecular cell biologists explore how all the remarkable properties of the cell arise from underlying molecular events: the assembly of large molecules, binding of large molecules to each other, catalytic effects that promote particular chemical reactions, and the deployment of information carried by giant molecules. Here we review the most important kinds of molecules that form the chemical foundations of cell structure and function.

### Small Molecules Carry Energy, Transmit Signals, and Are Linked into Macromolecules

Much of the cell's contents is a watery soup flavored with small molecules (e.g., simple sugars, amino acids, vitamins) and ions (e.g., sodium, chloride, calcium ions). The locations and concentrations of small molecules and ions within the cell are controlled by numerous proteins inserted in cellular membranes. These pumps, transporters, and ion channels move nearly all small molecules and ions into or out of the cell and its organelles (Chapter 7).



One of the best-known small molecules is **adenosine triphosphate (ATP)**, which stores readily available chemical energy in two of its chemical bonds (see Figure 2-24). When cells split apart these energy-rich bonds in ATP, the released energy can be harnessed to power an energy-requiring process like muscle contraction or protein biosynthesis. To obtain energy for making ATP, cells break down food molecules. For instance, when sugar is degraded to carbon dioxide and water, the energy stored in the original chemical bonds is released and much of it can be “captured” in ATP (Chapter 8). Bacterial, plant, and animal cells can all make ATP by this process. In addition, plants and a few other organisms can harvest energy from sunlight to form ATP in **photosynthesis**.

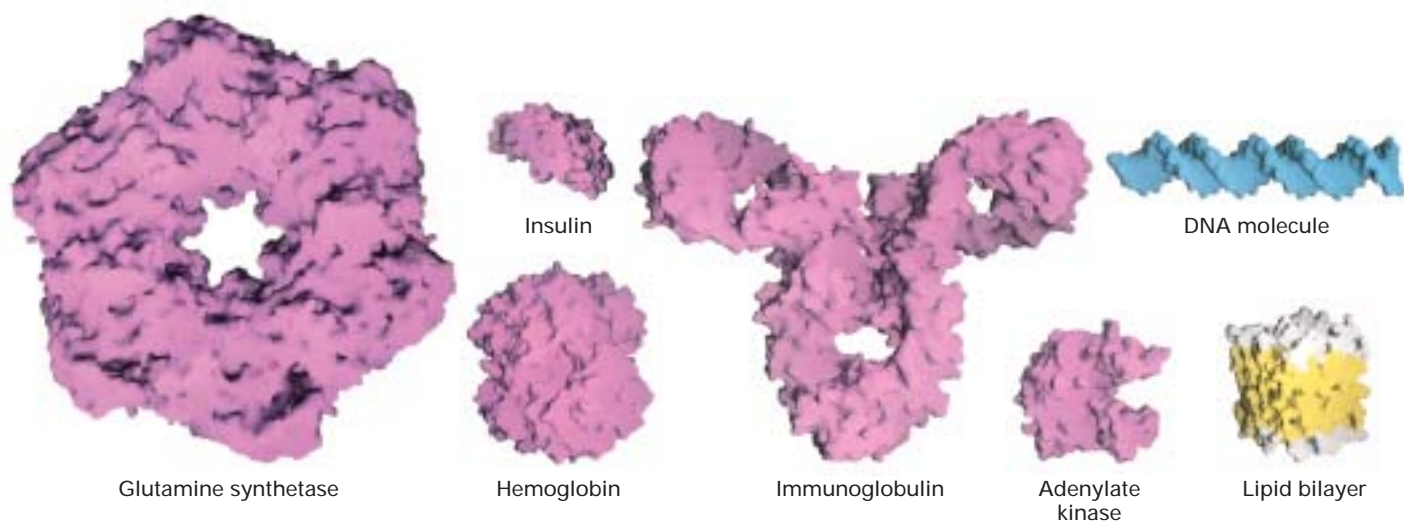
Other small molecules act as signals both within and between cells; such signals direct numerous cellular activities (Chapters 13–15). The powerful effect on our bodies of a frightening event comes from the instantaneous flooding of the body with epinephrine, a small-molecule **hormone** that mobilizes the “fight or flight” response. The movements needed to fight or flee are triggered by nerve impulses that flow from the brain to our muscles with the aid of **neurotransmitters**, another type of small-molecule signal that we discuss in Chapter 7.

Certain small molecules (**monomers**) in the cellular soup can be joined to form **polymers** through repetition of a single type of chemical-linkage reaction (see Figure 2-11). Cells produce three types of large polymers, commonly called **macromolecules**: polysaccharides, proteins, and nucleic acids. Sugars, for example, are the monomers used to form

**polysaccharides**. These macromolecules are critical structural components of plant cell walls and insect skeletons. A typical polysaccharide is a linear or branched chain of repeating identical sugar units. Such a chain carries information: the number of units. However if the units are *not* identical, then the order and type of units carry additional information. As we see in Chapter 6, some polysaccharides exhibit the greater informational complexity associated with a linear code made up of different units assembled in a particular order. This property, however, is most typical of the two other types of biological macromolecules—**proteins** and **nucleic acids**.

### Proteins Give Cells Structure and Perform Most Cellular Tasks

The varied, intricate structures of proteins enable them to carry out numerous functions. Cells string together 20 different **amino acids** in a linear chain to form a protein (see Figure 2-13). Proteins commonly range in length from 100 to 1000 amino acids, but some are much shorter and others longer. We obtain amino acids either by synthesizing them from other molecules or by breaking down proteins that we eat. The “essential” amino acids, from a dietary standpoint, are the eight that we cannot synthesize and must obtain from food. Beans and corn together have all eight, making their combination particularly nutritious. Once a chain of amino acids is formed, it folds into a complex shape, conferring a distinctive three-dimensional structure and function on each protein (Figure 1-9).



▲ **FIGURE 1-9 Proteins vary greatly in size, shape, and function.** These models of the water-accessible surface of some representative proteins are drawn to a common scale and reveal the numerous projections and crevices on the surface. Each protein has a defined three-dimensional shape (conformation) that is stabilized by numerous chemical interactions discussed in Chapters 2 and 3. The illustrated proteins include enzymes

(glutamine synthetase and adenylate kinase), an antibody (immunoglobulin), a hormone (insulin), and the blood’s oxygen carrier (hemoglobin). Models of a segment of the nucleic acid DNA and a small region of the lipid bilayer that forms cellular membranes (see Section 1.3) demonstrate the relative width of these structures compared with typical proteins. [Courtesy of Gareth White.]

Some proteins are similar to one another and therefore can be considered members of a **protein family**. A few hundred such families have been identified. Most proteins are designed to work in particular places within a cell or to be released into the extracellular (*extra*, “outside”) space. Elaborate cellular pathways ensure that proteins are transported to their proper intracellular (*intra*, within) locations or secreted (Chapters 16 and 17).

Proteins can serve as structural components of a cell, for example, by forming an internal skeleton (Chapters 5, 19, and 20). They can be sensors that change shape as temperature, ion concentrations, or other properties of the cell change. They can import and export substances across the plasma membrane (Chapter 7). They can be **enzymes**, causing chemical reactions to occur much more rapidly than they would without the aid of these protein **catalysts** (Chapter 3). They can bind to a specific gene, turning it on or off (Chapter 11). They can be extracellular signals, released from one cell to communicate with other cells, or intracellular signals, carrying information within the cell (Chapters 13–15). They can be motors that move other molecules around, burning chemical energy (ATP) to do so (Chapters 19 and 20).

How can 20 amino acids form all the different proteins needed to perform these varied tasks? Seems impossible at first glance. But if a “typical” protein is about 400 amino acids long, there are  $20^{400}$  possible different protein sequences. Even assuming that many of these would be functionally equivalent, unstable, or otherwise discountable, the number of possible proteins is well along toward infinity.

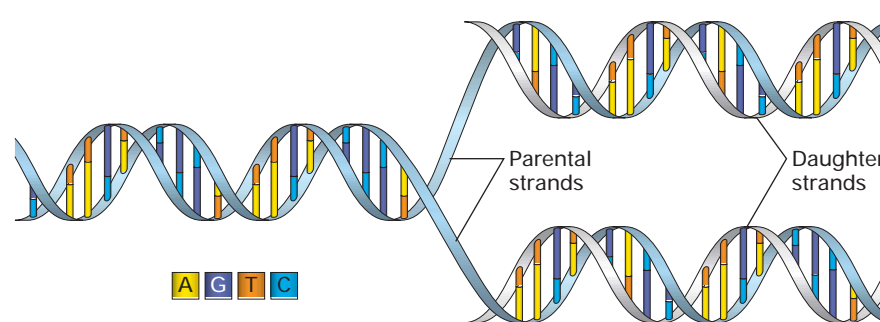
Next we might ask how many protein molecules a cell needs to operate and maintain itself. To estimate this number, let’s take a typical eukaryotic cell, such as a hepatocyte (liver cell). This cell, roughly a cube  $15\ \mu\text{m}$  ( $0.0015\ \text{cm}$ ) on a side, has a volume of  $3.4 \times 10^{-9}\ \text{cm}^3$  (or milliliters). Assuming a cell density of  $1.03\ \text{g/ml}$ , the cell would weigh  $3.5 \times 10^{-9}\ \text{g}$ . Since protein accounts for approximately 20 percent of a cell’s weight, the total weight of cellular protein is  $7 \times 10^{-10}\ \text{g}$ . The average yeast protein has a mo-

lecular weight of  $52,700\ (\text{g/mol})$ . Assuming this value is typical of eukaryotic proteins, we can calculate the total number of protein molecules per liver cell as about  $7.9 \times 10^9$  from the total protein weight and Avogadro’s number, the number of molecules per mole of any chemical compound ( $6.02 \times 10^{23}$ ). To carry this calculation one step further, consider that a liver cell contains about 10,000 different proteins; thus, a cell contains close to a million molecules of each type of protein on average. In actuality the abundance of different proteins varies widely, from the quite rare insulin-binding receptor protein (20,000 molecules) to the abundant structural protein actin ( $5 \times 10^8$  molecules).

### Nucleic Acids Carry Coded Information for Making Proteins at the Right Time and Place

The information about how, when, and where to produce each kind of protein is carried in the genetic material, a polymer called **deoxyribonucleic acid (DNA)**. The three-dimensional structure of DNA consists of two long helical strands that are coiled around a common axis, forming a **double helix**. DNA strands are composed of monomers called **nucleotides**; these often are referred to as *bases* because their structures contain cyclic organic bases (Chapter 4).

Four different nucleotides, abbreviated A, T, C, and G, are joined end to end in a DNA strand, with the base parts projecting out from the helical backbone of the strand. Each DNA double helix has a simple construction: wherever there is an A in one strand there is a T in the other, and each C is matched with a G (Figure 1-10). This **complementary** matching of the two strands is so strong that if complementary strands are separated, they will spontaneously zip back together in the right salt and temperature conditions. Such **hybridization** is extremely useful for detecting one strand using the other. For example, if one strand is purified and attached to a piece of paper, soaking the paper in a solution containing the other complementary strand will lead to zippering,



▲ **FIGURE 1-10 DNA consists of two complementary strands wound around each other to form a double helix.** (Left) The double helix is stabilized by weak hydrogen bonds between the A and T bases and between the C and G bases. (Right) During replication, the two strands are unwound and used

as templates to produce complementary strands. The outcome is two copies of the original double helix, each containing one of the original strands and one new daughter (complementary) strand.

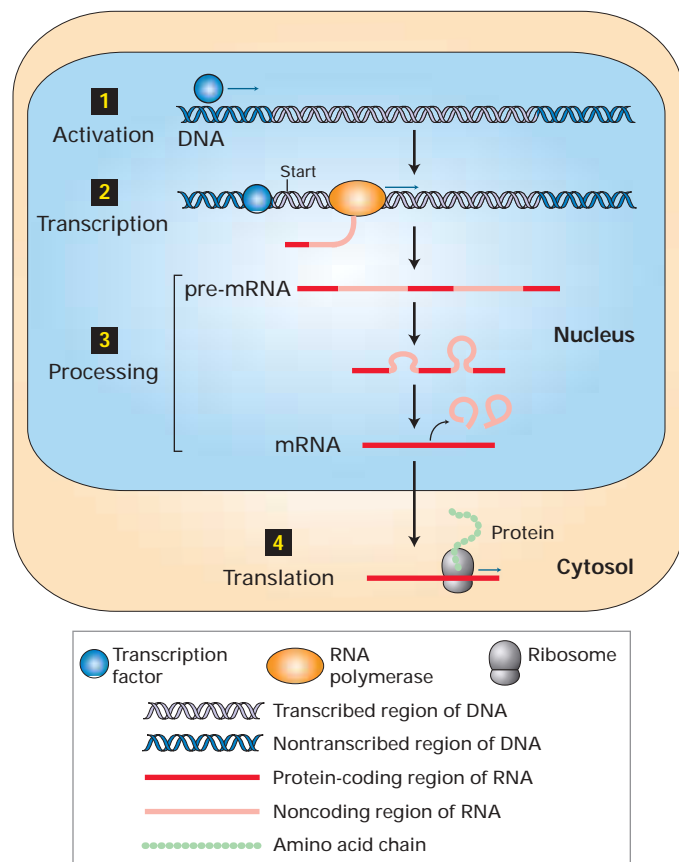
even if the solution also contains many other DNA strands that do not match.

The genetic information carried by DNA resides in its sequence, the linear order of nucleotides along a strand. The information-bearing portion of DNA is divided into discrete functional units, the **genes**, which typically are 5000 to 100,000 nucleotides long. Most bacteria have a few thousand genes; humans, about 40,000. The genes that carry instructions for making proteins commonly contain two parts: a *coding region* that specifies the amino acid sequence of a protein and a *regulatory region* that controls when and in which cells the protein is made.

Cells use two processes in series to convert the coded information in DNA into proteins (Figure 1-11). In the first, called **transcription**, the coding region of a gene is copied into a single-stranded **ribonucleic acid (RNA)** version of the double-stranded DNA. A large enzyme, **RNA polymerase**, catalyzes the linkage of nucleotides into a RNA chain using DNA as a template. In eukaryotic cells, the initial RNA product is processed into a smaller **messenger RNA (mRNA)** molecule, which moves to the cytoplasm. Here the **ribosome**, an enormously complex molecular machine composed of both RNA and protein, carries out the second process, called **translation**. During translation, the ribosome assembles and links together amino acids in the precise order dictated by the mRNA sequence according to the nearly universal **genetic code**. We examine the cell components that carry out transcription and translation in detail in Chapter 4.

All organisms have ways to control when and where their genes can be transcribed. For instance, nearly all the cells in our bodies contain the full set of human genes, but in each cell type only some of these genes are active, or turned on, and used to make proteins. That's why liver cells produce some proteins that are not produced by kidney cells, and vice versa. Moreover, many cells can respond to external signals or changes in external conditions by turning specific genes on or off, thereby adapting their repertoire of proteins to meet current needs. Such control of gene activity depends on DNA-binding proteins called **transcription factors**, which bind to DNA and act as switches, either activating or repressing transcription of particular genes (Chapter 11).

Transcription factors are shaped so precisely that they are able to bind preferentially to the regulatory regions of just a few genes out of the thousands present in a cell's DNA. Typically a DNA-binding protein will recognize short DNA sequences about 6–12 base pairs long. A segment of DNA containing 10 base pairs can have  $4^{10}$  possible sequences (1,048,576) since each position can be any of four nucleotides. Only a few copies of each such sequence will occur in the DNA of a cell, assuring the specificity of gene activation and repression. Multiple copies of one type of transcription factor can coordinately regulate a set of genes if binding sites for that factor exist near each gene in the set. Transcription factors often work as multiprotein complexes, with more than one protein contributing its own DNA-binding specificity to selecting the regulated genes. In complex organisms,



▲ **FIGURE 1-11 The coded information in DNA is converted into the amino acid sequences of proteins by a multistep process.** Step **1**: Transcription factors bind to the regulatory regions of the specific genes they control and activate them. Step **2**: Following assembly of a multiprotein initiation complex bound to the DNA, RNA polymerase begins transcription of an activated gene at a specific location, the start site. The polymerase moves along the DNA linking nucleotides into a single-stranded pre-mRNA transcript using one of the DNA strands as a template. Step **3**: The transcript is processed to remove noncoding sequences. Step **4**: In a eukaryotic cell, the mature messenger RNA (mRNA) moves to the cytoplasm, where it is bound by ribosomes that read its sequence and assemble a protein by chemically linking amino acids into a linear chain.

hundreds of different transcription factors are employed to form an exquisite control system that activates the right genes in the right cells at the right times.

### The Genome Is Packaged into Chromosomes and Replicated During Cell Division

Most of the DNA in eukaryotic cells is located in the nucleus, extensively folded into the familiar structures we know as **chromosomes** (Chapter 10). Each chromosome contains a single linear DNA molecule associated with certain proteins. In prokaryotic cells, most or all of the genetic information resides

in a single circular DNA molecule about a millimeter in length; this molecule lies, folded back on itself many times, in the central region of the cell (see Figure 1-2a). The **genome** of an organism comprises its entire complement of DNA. With the exception of eggs and sperm, every normal human cell has 46 chromosomes (Figure 1-12). Half of these, and thus half of the genes, can be traced back to Mom; the other half, to Dad.

Every time a cell divides, a large multiprotein replication machine, the replisome, separates the two strands of double-helical DNA in the chromosomes and uses each strand as a template to assemble nucleotides into a new complementary strand (see Figure 1-10). The outcome is a pair of double helices, each identical to the original. **DNA polymerase**, which is responsible for linking nucleotides into a DNA strand, and the many other components of the replisome are described in Chapter 4. The molecular design of DNA and the remarkable properties of the replisome assure rapid, highly accurate copying. Many DNA polymerase molecules work in concert, each one copying part of a chromosome. The entire genome of fruit flies, about  $1.2 \times 10^8$  nucleotides long, can be copied in three minutes! Because of the accuracy of DNA replication, nearly all the cells in our bodies carry the same genetic instructions, and we can inherit Mom's brown hair and Dad's blue eyes.

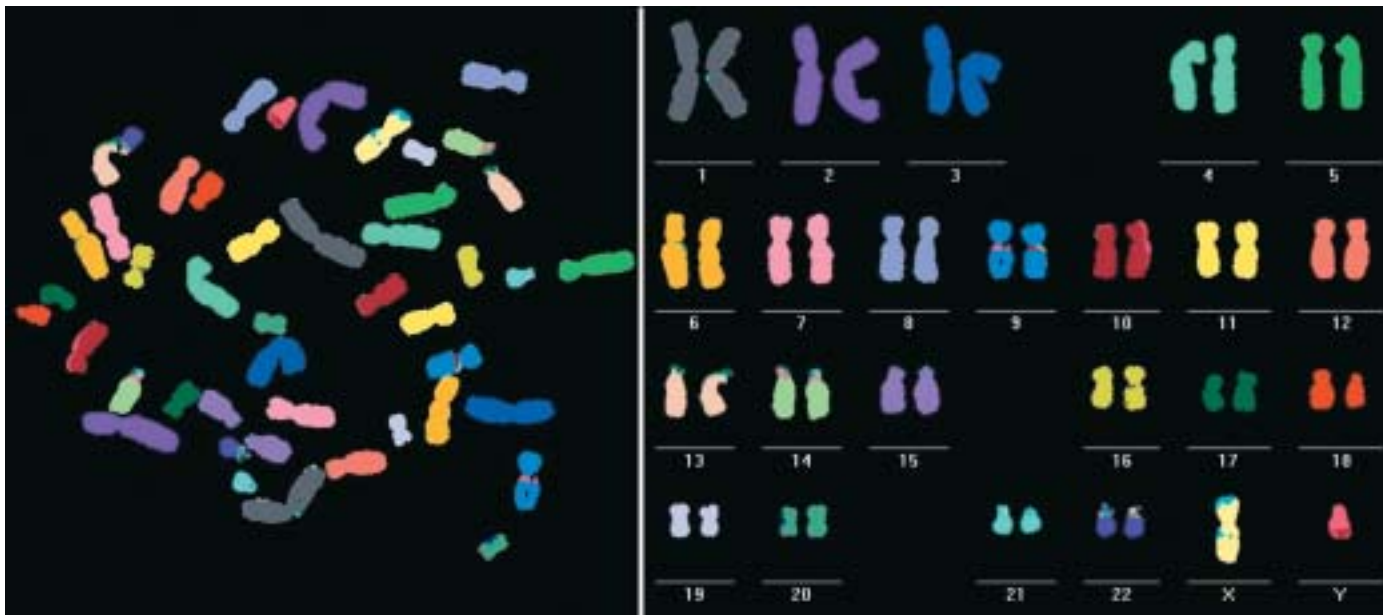
A rather dramatic example of gene control involves inactivation of an entire chromosome in human females. Women have two X chromosomes, whereas men have one

X chromosome and one Y chromosome, which has different genes than the X chromosome. Yet the genes on the X chromosome must, for the most part, be equally active in female cells (XX) and male cells (XY). To achieve this balance, one of the X chromosomes in female cells is chemically modified and condensed into a very small mass called a Barr body, which is inactive and never transcribed.

Surprisingly, we inherit a small amount of genetic material entirely and uniquely from our mothers. This is the circular DNA present in mitochondria, the organelles in eukaryotic cells that synthesize ATP using the energy released by the breakdown of nutrients. Mitochondria contain multiple copies of their own DNA genomes, which code for some of the mitochondrial proteins (Chapter 10). Because each human inherits mitochondrial DNA only from his or her mother (it comes with the egg but not the sperm), the distinctive features of a particular mitochondrial DNA can be used to trace maternal history. Chloroplasts, the organelles that carry out photosynthesis in plants, also have their own circular genomes.

### Mutations May Be Good, Bad, or Indifferent

Mistakes occasionally do occur spontaneously during DNA replication, causing changes in the sequence of nucleotides. Such changes, or **mutations**, also can arise from radiation



▲ **FIGURE 1-12 Chromosomes can be “painted” for easy identification.** A normal human has 23 pairs of morphologically distinct chromosomes; one member of each pair is inherited from the mother and the other member from the father. *(Left)* A chromosome spread from a human body cell midway through mitosis, when the chromosomes are fully condensed. This preparation was treated with fluorescent-labeled staining reagents that allow each of the 22 pairs and the X and Y

chromosomes to appear in a different color when viewed in a fluorescence microscope. This technique of multiplex fluorescence in situ hybridization (M-FISH) sometimes is called chromosome painting (Chapter 10). *(Right)* Chromosomes from the preparation on the left arranged in pairs in descending order of size, an array called a karyotype. The presence of X and Y chromosomes identifies the sex of the individual as male. [Courtesy of M. R. Speicher.]

that causes damage to the nucleotide chain or from chemical poisons, such as those in cigarette smoke, that lead to errors during the DNA-copying process (Chapter 23). Mutations come in various forms: a simple swap of one nucleotide for another; the deletion, insertion, or inversion of one to millions of nucleotides in the DNA of one chromosome; and translocation of a stretch of DNA from one chromosome to another.

In sexually reproducing animals like ourselves, mutations can be inherited only if they are present in cells that potentially contribute to the formation of offspring. Such **germ-line cells** include eggs, sperm, and their precursor cells. Body cells that do not contribute to offspring are called **somatic cells**. Mutations that occur in these cells never are inherited, although they may contribute to the onset of cancer. Plants have a less distinct division between somatic and germ-line cells, since many plant cells can function in both capacities.

Mutated genes that encode altered proteins or that cannot be controlled properly cause numerous inherited diseases. For example, sickle cell disease is attributable to a single nucleotide substitution in the hemoglobin gene, which encodes the protein that carries oxygen in red blood cells. The single amino acid change caused by the sickle cell mutation reduces the ability of red blood cells to carry oxygen from the lungs to the tissues. Recent advances in detecting disease-causing mutations and in understanding how they affect cell functions offer exciting possibilities for reducing their often devastating effects.

Sequencing of the human genome has shown that a very large proportion of our DNA does not code for any RNA or have any discernible regulatory function, a quite unexpected finding. Mutations in these regions usually produce no immediate effects—good or bad. However, such “indifferent” mutations in nonfunctional DNA may have been a major player in evolution, leading to creation of new genes or new regulatory sequences for controlling already existing genes. For instance, since binding sites for transcription factors typically are only 10–12 nucleotides long, a few single-nucleotide mutations might convert a nonfunctional bit of DNA into a functional protein-binding regulatory site.

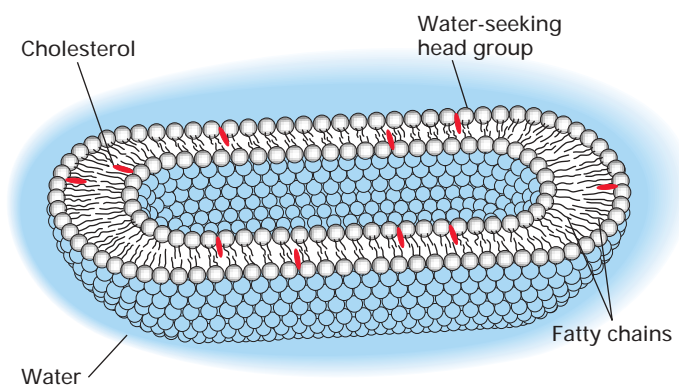
Much of the nonessential DNA in both eukaryotes and prokaryotes consists of highly repeated sequences that can move from one place in the genome to another. These **mobile DNA elements** can jump (transpose) into genes, most commonly damaging but sometimes activating them. Jumping generally occurs rarely enough to avoid endangering the host organism. Mobile elements, which were discovered first in plants, are responsible for leaf color variegation and the diverse beautiful color patterns of Indian corn kernels. By jumping in and out of genes that control pigmentation as plant development progresses, the mobile elements give rise to elaborate colored patterns. Mobile elements were later found in bacteria in which they often carry and, unfortunately, disseminate genes for antibiotic resistance.

Now we understand that mobile elements have multiplied and slowly accumulated in genomes over evolutionary

time, becoming a universal property of genomes in present-day organisms. They account for an astounding 45 percent of the human genome. Some of our own mobile DNA elements are copies—often highly mutated and damaged—of genomes from viruses that spend part of their life cycle as DNA segments inserted into host-cell DNA. Thus we carry in our chromosomes the genetic residues of infections acquired by our ancestors. Once viewed only as molecular parasites, mobile DNA elements are now thought to have contributed significantly to the evolution of higher organisms (Chapter 10).

### 1.3 The Work of Cells

In essence, any cell is simply a compartment with a watery interior that is separated from the external environment by a surface membrane (the plasma membrane) that prevents the free flow of molecules in and out of cells. In addition, as we’ve noted, eukaryotic cells have extensive internal membranes that further subdivide the cell into various compartments, the organelles. The plasma membrane and other cellular membranes are composed primarily of two layers of **phospholipid** molecules. These bipartite molecules have a “water-loving” (**hydrophilic**) end and a “water-hating” (**hydrophobic**) end. The two phospholipid layers of a membrane are oriented with all the hydrophilic ends directed toward the inner and outer surfaces and the hydrophobic ends buried within the interior (Figure 1-13). Smaller amounts of



▲ FIGURE 1-13 The watery interior of cells is surrounded by the plasma membrane, a two-layered shell of phospholipids. The phospholipid molecules are oriented with their fatty acyl chains (black squiggly lines) facing inward and their water-seeking head groups (white spheres) facing outward. Thus both sides of the membrane are lined by head groups, mainly charged phosphates, adjacent to the watery spaces inside and outside the cell. All biological membranes have the same basic phospholipid bilayer structure. Cholesterol (red) and various proteins (not shown) are embedded in the bilayer. In actuality, the interior space is much larger relative to the volume of the plasma membrane depicted here.

other lipids, such as cholesterol, and many kinds of proteins are inserted into the phospholipid framework. The lipid molecules and some proteins can float sidewise in the plane of the membrane, giving membranes a fluid character. This fluidity allows cells to change shape and even move. However, the attachment of some membrane proteins to other molecules inside or outside the cell restricts their lateral movement. We learn more about membranes and how molecules cross them in Chapters 5 and 7.

The cytosol and the internal spaces of organelles differ from each other and from the cell exterior in terms of acidity, ionic composition, and protein contents. For example, the composition of salts inside the cell is often drastically different from what is outside. Because of these different “microclimates,” each cell compartment has its own assigned tasks in the overall work of the cell (Chapter 5). The unique functions and micro-climates of the various cell compartments are due largely to the proteins that reside in their membranes or interior.

We can think of the entire cell compartment as a factory dedicated to sustaining the well-being of the cell. Much cellular work is performed by molecular machines, some housed in the cytosol and some in various organelles. Here we quickly review the major tasks that cells carry out in their pursuit of the good life.

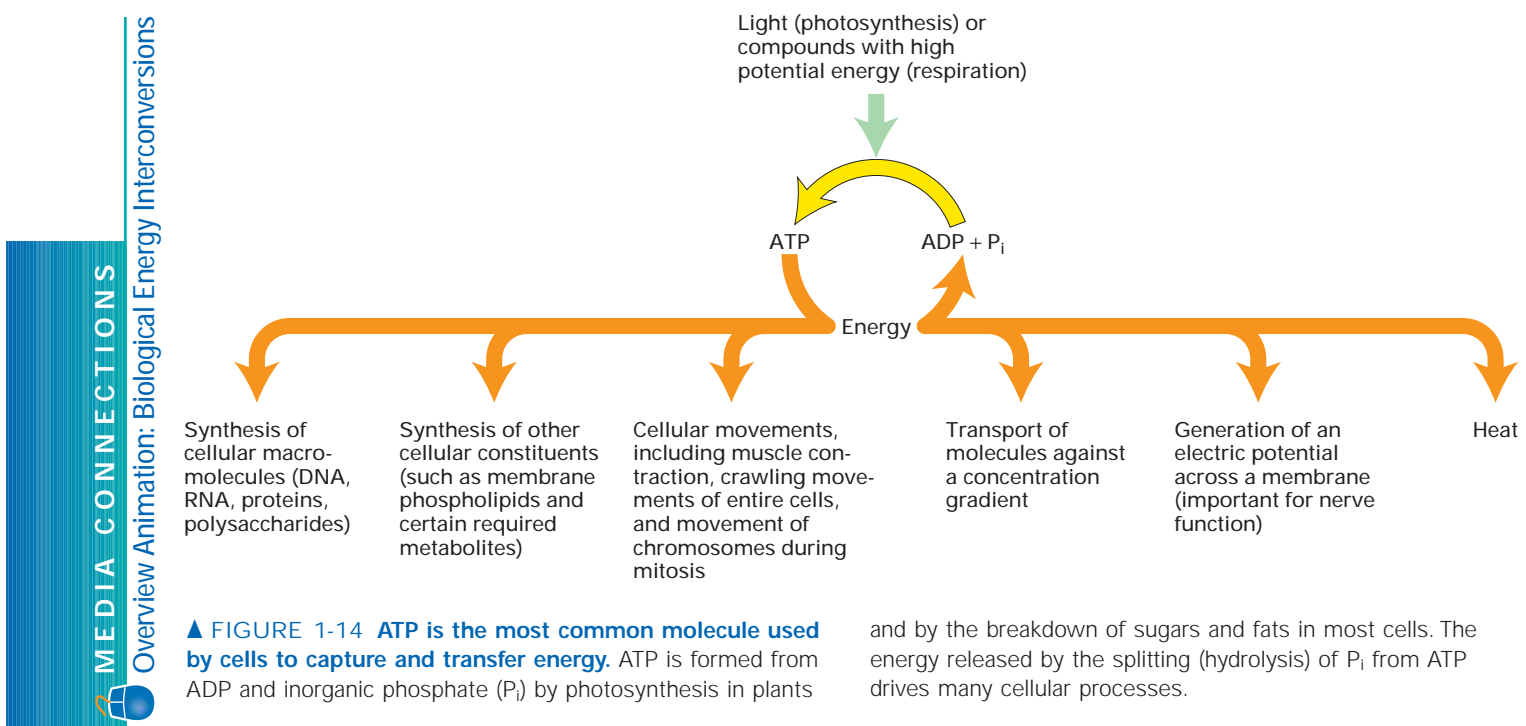
### Cells Build and Degrade Numerous Molecules and Structures

As chemical factories, cells produce an enormous number of complex molecules from simple chemical building blocks. All

of this synthetic work is powered by chemical energy extracted primarily from sugars and fats or sunlight, in the case of plant cells, and stored primarily in ATP, the universal “currency” of chemical energy (Figure 1-14). In animal and plant cells, most ATP is produced by large molecular machines located in two organelles, **mitochondria** and **chloroplasts**. Similar machines for generating ATP are located in the plasma membrane of bacterial cells. Both mitochondria and chloroplasts are thought to have originated as bacteria that took up residence inside eukaryotic cells and then became welcome collaborators (Chapter 8). Directly or indirectly, all of our food is created by plant cells using sunlight to build complex macromolecules during photosynthesis. Even underground oil supplies are derived from the decay of plant material.

Cells need to break down worn-out or obsolete parts into small molecules that can be discarded or recycled. This housekeeping task is assigned largely to **lysosomes**, organelles crammed with degradative enzymes. The interior of lysosomes has a pH of about 5.0, roughly 100 times more acidic than that of the surrounding cytosol. This aids in the breakdown of materials by lysosomal enzymes, which are specially designed to function at such a low pH. To create the low pH environment, proteins located in the lysosomal membrane pump hydrogen ions into the lysosome using energy supplied from ATP (Chapter 7). Lysosomes are assisted in the cell’s cleanup work by **peroxisomes**. These small organelles are specialized for breaking down the lipid components of membranes and rendering various toxins harmless.

Most of the structural and functional properties of cells depend on proteins. Thus for cells to work properly, the nu-



merous proteins composing the various working compartments must be transported from where they are made to their proper locations (Chapters 16 and 17). Some proteins are made on ribosomes that are free in the cytosol. Proteins secreted from the cell and most membrane proteins, however, are made on ribosomes associated with the **endoplasmic reticulum (ER)**. This organelle produces, processes, and ships out both proteins and lipids. Protein chains produced on the ER move to the **Golgi apparatus**, where they are further modified before being forwarded to their final destinations. Proteins that travel in this way contain short sequences of amino acids or attached sugar chains (oligosaccharides) that serve as addresses for directing them to their correct destinations. These addresses work because they are recognized and bound by other proteins that do the sorting and shipping in various cell compartments.

### Animal Cells Produce Their Own External Environment and Glues

The simplest multicellular animals are single cells embedded in a jelly of proteins and polysaccharides called the **extracellular matrix**. Cells themselves produce and secrete these materials, thus creating their own immediate environment (Chapter 6). **Collagen**, the single most abundant protein in the animal kingdom, is a major component of the extracellular matrix in most tissues. In animals, the extracellular matrix cushions and lubricates cells. A specialized, especially tough matrix, the **basal lamina**, forms a supporting layer underlying sheetlike cell layers and helps prevent the cells from ripping apart.

The cells in animal tissues are “glued” together by **cell-adhesion molecules (CAMs)** embedded in their surface membranes. Some CAMs bind cells to one another; other types bind cells to the extracellular matrix, forming a cohesive unit. The cells of higher plants contain relatively few such molecules; instead, plants cells are rigidly tied together by extensive interlocking of the cell walls of neighboring

cells. The cytosols of adjacent animal or plant cells often are connected by functionally similar but structurally different “bridges” called **gap junctions** in animals and **plasmodesmata** in plants. These structures allow cells to exchange small molecules including nutrients and signals, facilitating coordinated functioning of the cells in a tissue.

### Cells Change Shape and Move

Although cells sometimes are spherical, they more commonly have more elaborate shapes due to their internal skeletons and external attachments. Three types of protein filaments, organized into networks and bundles, form the **cytoskeleton** within animal cells (Figure 1-15). The cytoskeleton prevents the plasma membrane of animal cells from relaxing into a sphere (Chapter 5); it also functions in cell locomotion and the intracellular transport of vesicles, chromosomes, and macromolecules (Chapters 19 and 20). The cytoskeleton can be linked through the cell surface to the extracellular matrix or to the cytoskeleton of other cells, thus helping to form tissues (Chapter 6).

All cytoskeletal filaments are long polymers of protein subunits. Elaborate systems regulate the assembly and disassembly of the cytoskeleton, thereby controlling cell shape. In some cells the cytoskeleton is relatively stable, but in others it changes shape continuously. Shrinkage of the cytoskeleton in some parts of the cell and its growth in other parts can produce coordinated changes in shape that result in cell locomotion. For instance, a cell can send out an extension that attaches to a surface or to other cells and then retract the cell body from the other end. As this process continues due to coordinated changes in the cytoskeleton, the cell moves forward. Cells can move at rates on the order of 20  $\mu\text{m}/\text{second}$ . Cell locomotion is used during embryonic development of multicellular animals to shape tissues and during adulthood to defend against infection, to transport nutrients, and to heal wounds. This process does not play a role in the growth and development of multicellular plants because new plant cells



▲ **FIGURE 1-15 The three types of cytoskeletal filaments have characteristic distributions within cells.** Three views of the same cell. A cultured fibroblast was treated with three different antibody preparations. Each antibody binds specifically to the protein monomers forming one type of filament and is chemically linked to a differently colored fluorescent dye (green,

blue, or red). Visualization of the stained cell in a fluorescence microscope reveals the location of filaments bound to a particular dye-antibody preparation. In this case, intermediate filaments are stained green; microtubules, blue; and microfilaments, red. All three fiber systems contribute to the shape and movements of cells. [Courtesy of V. Small.]

are generated by the division of existing cells that share cell walls. As a result, plant development involves cell enlargement but not movement of cells from one position to another.

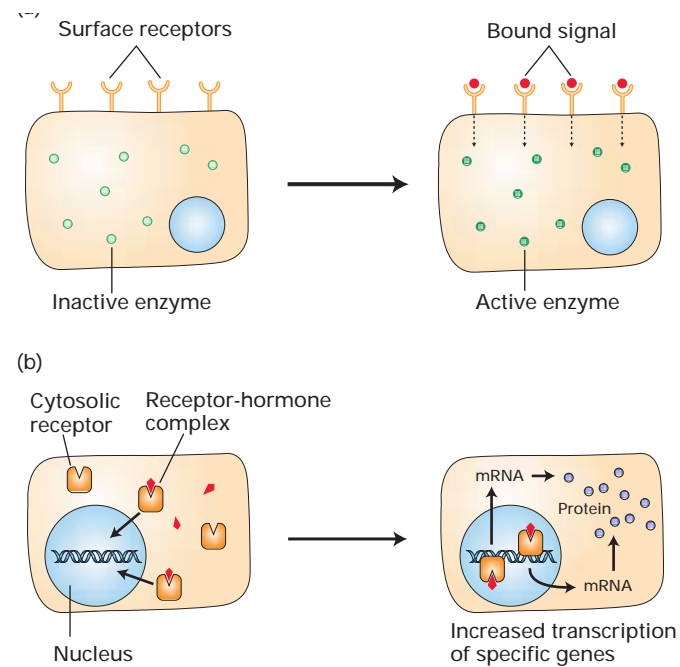
### Cells Sense and Send Information

A living cell continuously monitors its surroundings and adjusts its own activities and composition accordingly. Cells also communicate by deliberately sending signals that can be received and interpreted by other cells. Such signals are common not only within an individual organism, but also between organisms. For instance, the odor of a pear detected by us and other animals signals a food source; consumption of the pear by an animal aids in distributing the pear's seeds. Everyone benefits! The signals employed by cells include simple small chemicals, gases, proteins, light, and mechanical movements. Cells possess numerous receptor proteins for detecting signals and elaborate pathways for transmitting them within the cell to evoke a response. At any time, a cell may be able to sense only some of the signals around it, and how a cell responds to a signal may change with time. In some cases, receiving one signal primes a cell to respond to a subsequent different signal in a particular way.

Both changes in the environment (e.g., an increase or decrease in a particular nutrient or the light level) and signals received from other cells represent external information that cells must process. The most rapid responses to such signals generally involve changes in the location or activity of preexisting proteins. For instance, soon after you eat a carbohydrate-rich meal, glucose pours into your bloodstream. The rise in blood glucose is sensed by  $\beta$  cells in the pancreas, which respond by releasing their stored supply of the protein hormone **insulin**. The circulating insulin signal causes glucose transporters in the cytoplasm of fat and muscle cells to move to the cell surface, where they begin importing glucose. Meanwhile, liver cells also are furiously taking in glucose via a different glucose transporter. In both liver and muscle cells, an intracellular signaling pathway triggered by binding of insulin to cell-surface receptors activates a key enzyme needed to make **glycogen**, a large glucose polymer (Figure 1-16a). The net result of these cell responses is that your blood glucose level falls and extra glucose is stored as glycogen, which your cells can use as a glucose source when you skip a meal to cram for a test.

The ability of cells to send and respond to signals is crucial to development. Many developmentally important signals are secreted proteins produced by specific cells at specific times and places in a developing organism. Often a receiving cell integrates multiple signals in deciding how to behave, for example, to differentiate into a particular tissue type, to extend a process, to die, to send back a confirming signal (yes, I'm here!), or to migrate.

The functions of about half the proteins in humans, roundworms, yeast, and several other eukaryotic organisms have been predicted based on analyses of genomic sequences (Chapter 9). Such analyses have revealed that at least 10–15 percent of the proteins in eukaryotes function as secreted ex-



▲ FIGURE 1-16 External signals commonly cause a change in the activity of preexisting proteins or in the amounts and types of proteins that cells produce. (a) Binding of a hormone or other signaling molecule to its specific receptors can trigger an intracellular pathway that increases or decreases the activity of a preexisting protein. For example, binding of insulin to receptors in the plasma membrane of liver and muscle cells leads to activation of glycogen synthase, a key enzyme in the synthesis of glycogen from glucose. (b) The receptors for steroid hormones are located within cells, not on the cell surface. The hormone-receptor complexes activate transcription of specific target genes, leading to increased production of the encoded proteins. Many signals that bind to receptors on the cell surface also act, by more complex pathways, to modulate gene expression.

tracellular signals, signal receptors, or intracellular **signal-transduction** proteins, which pass along a signal through a series of steps culminating in a particular cellular response (e.g., increased glycogen synthesis). Clearly, signaling and signal transduction are major activities of cells.

### Cells Regulate Their Gene Expression to Meet Changing Needs

In addition to modulating the activities of existing proteins, cells often respond to changing circumstances and to signals from other cells by altering the amount or types of proteins they contain. **Gene expression**, the overall process of selectively reading and using genetic information, is commonly controlled at the level of transcription, the first step in the production of proteins. In this way cells can produce a particular mRNA only when the encoded protein is needed, thus minimizing wasted energy. Producing a mRNA is, however, only the first in a chain of regulated events that together determine whether an active protein product is produced from a particular gene.



Transcriptional control of gene expression was first decisively demonstrated in the response of the gut bacterium *E. coli* to different sugar sources. *E. coli* cells prefer glucose as a sugar source, but they can survive on lactose in a pinch. These bacteria use both a DNA-binding *repressor* protein and a DNA-binding *activator* protein to change the rate of transcription of three genes needed to metabolize lactose depending on the relative amounts of glucose and lactose present (Chapter 4). Such dual positive/negative control of gene expression fine tunes the bacterial cell's enzymatic equipment for the job at hand.

Like bacterial cells, unicellular eukaryotes may be subjected to widely varying environmental conditions that require extensive changes in cellular structures and function. For instance, in starvation conditions yeast cells stop growing and form dormant spores (see Figure 1-4). In multicellular organisms, however, the environment around most cells is relatively constant. The major purpose of gene control in us and in other complex organisms is to tailor the properties of various cell types to the benefit of the entire animal or plant.

Control of gene activity in eukaryotic cells usually involves a balance between the actions of transcriptional activators and repressors. Binding of activators to specific DNA regulatory sequences called **enhancers** turns on transcription, and binding of repressors to other regulatory sequences called **silencers** turns off transcription. In Chapters 11 and 12, we take a close look at transcriptional activators and repressors and how they operate, as well as other mechanisms for controlling gene expression. In an extreme case, expression of a particular gene could occur only in part of the brain, only during evening hours, only during a certain stage of development, only after a large meal, and so forth.

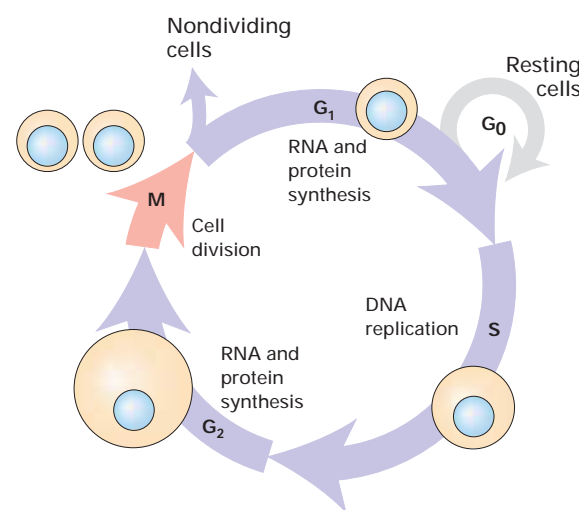
Many external signals modify the activity of transcriptional activators and repressors that control specific genes. For example, lipid-soluble steroid hormones, such as estrogen and testosterone, can diffuse across the plasma membrane and bind to their specific receptors located in the cytoplasm or nucleus (Figure 1-16b). Hormone binding changes the shape of the receptor so that it can bind to specific enhancer sequences in the DNA, thus turning the receptor into a transcriptional activator. By this rather simple signal-transduction pathway, steroid hormones cause cells to change which genes they transcribe (Chapter 11). Since steroid hormones can circulate in the bloodstream, they can affect the properties of many or all cells in a temporally coordinated manner. Binding of many other hormones and of growth factors to receptors on the cell surface triggers different signal-transduction pathways that also lead to changes in the transcription of specific genes (Chapters 13–15). Although these pathways involve multiple components and are more complicated than those transducing steroid hormone signals, the general idea is the same.

### Cells Grow and Divide

The most remarkable feature of cells and entire organisms is their ability to reproduce. Biological reproduction, combined

with continuing evolutionary selection for a highly functional body plan, is why today's horseshoe crabs look much as they did 300 million years ago, a time span during which entire mountain ranges have risen or fallen. The Teton Mountains in Wyoming, now about 14,000 feet high and still growing, did not exist a mere 10 million years ago. Yet horseshoe crabs, with a life span of about 19 years, have faithfully reproduced their ancient selves more than half a million times during that period. The common impression that biological structure is transient and geological structure is stable is the exact opposite of the truth. Despite the limited duration of our individual lives, reproduction gives us a potential for immortality that a mountain or a rock does not have.

The simplest type of reproduction entails the division of a “parent” cell into two “daughter” cells. This occurs as part of the **cell cycle**, a series of events that prepares a cell to divide followed by the actual division process, called **mitosis**. The eukaryotic cell cycle commonly is represented as four stages (Figure 1-17). The chromosomes and the DNA they carry are copied during the **S (synthesis) phase**. The replicated chromosomes separate during the **M (mitotic) phase**, with each daughter cell getting a copy of each chromosome during cell division. The M and S phases are separated by two gap stages, the **G<sub>1</sub> phase** and **G<sub>2</sub> phase**, during which mRNAs and proteins are made. In single-celled organisms, both daughter cells



▲ **FIGURE 1-17** During growth, eukaryotic cells continually progress through the four stages of the cell cycle, generating new daughter cells. In most proliferating cells, the four phases of the cell cycle proceed successively, taking from 10–20 hours depending on cell type and developmental state. During interphase, which consists of the G<sub>1</sub>, S, and G<sub>2</sub> phases, the cell roughly doubles its mass. Replication of DNA during S leaves the cell with four copies of each type of chromosome. In the mitotic (M) phase, the chromosomes are evenly partitioned to two daughter cells, and the cytoplasm divides roughly in half in most cases. Under certain conditions such as starvation or when a tissue has reached its final size, cells will stop cycling and remain in a waiting state called G<sub>0</sub>. Most cells in G<sub>0</sub> can reenter the cycle if conditions change.

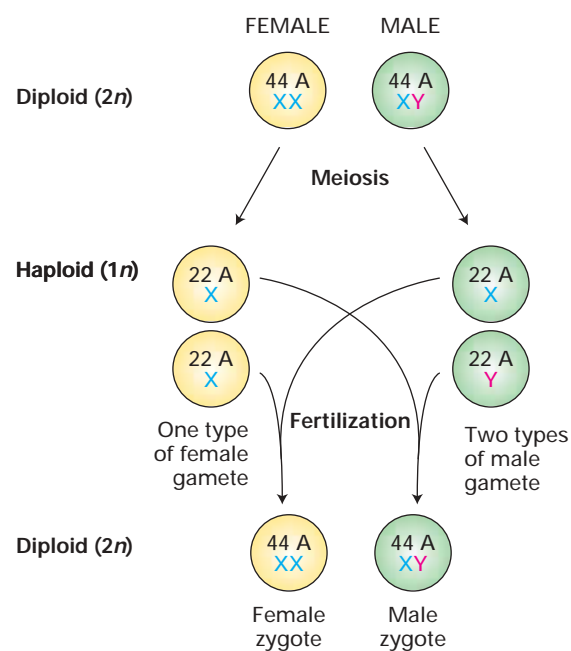
often (though not always) resemble the parent cell. In multicellular organisms, **stem cells** can give rise to two different cells, one that resembles the parent cell and one that does not. Such asymmetric cell division is critical to the generation of different cell types in the body (Chapter 22).

During growth the cell cycle operates continuously, with newly formed daughter cells immediately embarking on their own path to mitosis. Under optimal conditions bacteria can divide to form two daughter cells once every 30 minutes. At this rate, in an hour one cell becomes four; in a day one cell becomes more than  $10^{14}$ , which if dried would weigh about 25 grams. Under normal circumstances, however, growth cannot continue at this rate because the food supply becomes limiting.

Most eukaryotic cells take considerably longer than bacterial cells to grow and divide. Moreover, the cell cycle in adult plants and animals normally is highly regulated (Chapter 21). This tight control prevents imbalanced, excessive growth of tissues while assuring that worn-out or damaged cells are replaced and that additional cells are formed in response to new circumstances or developmental needs. For instance, the proliferation of red blood cells increases substantially when a person ascends to a higher altitude and needs more capacity to capture oxygen. Some highly specialized cells in adult animals, such as nerve cells and striated muscle cells, rarely divide, if at all. The fundamental defect in cancer is loss of the ability to control the growth and division of cells. In Chapter 23, we examine the molecular and cellular events that lead to inappropriate, uncontrolled proliferation of cells.

Mitosis is an *asexual* process since the daughter cells carry the exact same genetic information as the parental cell. In *sexual* reproduction, fusion of two cells produces a third cell that contains genetic information from each parental cell. Since such fusions would cause an ever-increasing number of chromosomes, sexual reproductive cycles employ a special type of cell division, called **meiosis**, that reduces the number of chromosomes in preparation for fusion (see Figure 9-3). Cells with a full set of chromosomes are called **diploid** cells. During meiosis, a diploid cell replicates its chromosomes as usual for mitosis but then divides twice without copying the chromosomes in-between. Each of the resulting four daughter cells, which has only half the full number of chromosomes, is said to be **haploid**.

Sexual reproduction occurs in animals and plants, and even in unicellular organisms such as yeasts (see Figure 1-5). Animals spend considerable time and energy generating eggs and sperm, the haploid cells, called **gametes**, that are used for sexual reproduction. A human female will produce about half a million eggs in a lifetime, all these cells form before she is born; a young human male, about 100 million sperm each day. Gametes are formed from diploid precursor germ-line cells, which in humans contain 46 chromosomes. In humans the X and Y chromosomes are called sex chromosomes because they determine whether an individual is male or female. In human diploid cells, the 44 remaining chromosomes, called **autosomes**, occur as pairs of 22 different kinds. Through meiosis, a man produces sperm that have 22 chromosomes plus either an X or a Y, and a woman produces ova (unfertilized eggs) with



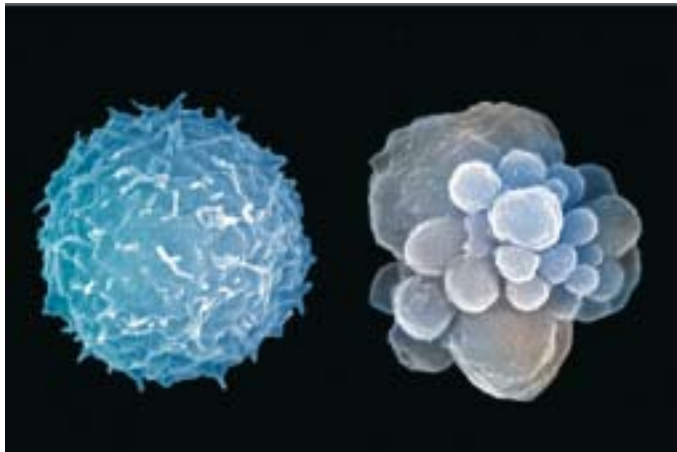
▲ **FIGURE 1-18 Dad made you a boy or girl.** In animals, meiosis of diploid precursor cells forms eggs and sperm (gametes). The male parent produces two types of sperm and determines the sex of the zygote. In humans, as shown here, X and Y are the sex chromosomes; the zygote must receive a Y chromosome from the male parent to develop into a male. A=autosomes (non-sex chromosomes).

22 chromosomes plus an X. Fusion of an egg and sperm (fertilization) yields a fertilized egg, the zygote, with 46 chromosomes, one pair of each of the 22 kinds and a pair of X's in females or an X and a Y in males (Figure 1-18). Errors during meiosis can lead to disorders resulting from an abnormal number of chromosomes. These include Down's syndrome, caused by an extra chromosome 21, and Klinefelter's syndrome, caused by an extra X chromosome.

### Cells Die from Aggravated Assault or an Internal Program

When cells in multicellular organisms are badly damaged or infected with a virus, they die. Cell death resulting from such a traumatic event is messy and often releases potentially toxic cell constituents that can damage surrounding cells. Cells also may die when they fail to receive a life-maintaining signal or when they receive a death signal. In this type of programmed cell death, called **apoptosis**, a dying cell actually produces proteins necessary for self-destruction. Death by apoptosis avoids the release of potentially toxic cell constituents (Figure 1-19).

Programmed cell death is critical to the proper development and functioning of our bodies (Chapter 22). During fetal life, for instance, our hands initially develop with "webbing" between the fingers; the cells in the webbing subsequently die in an orderly and precise pattern that leaves the



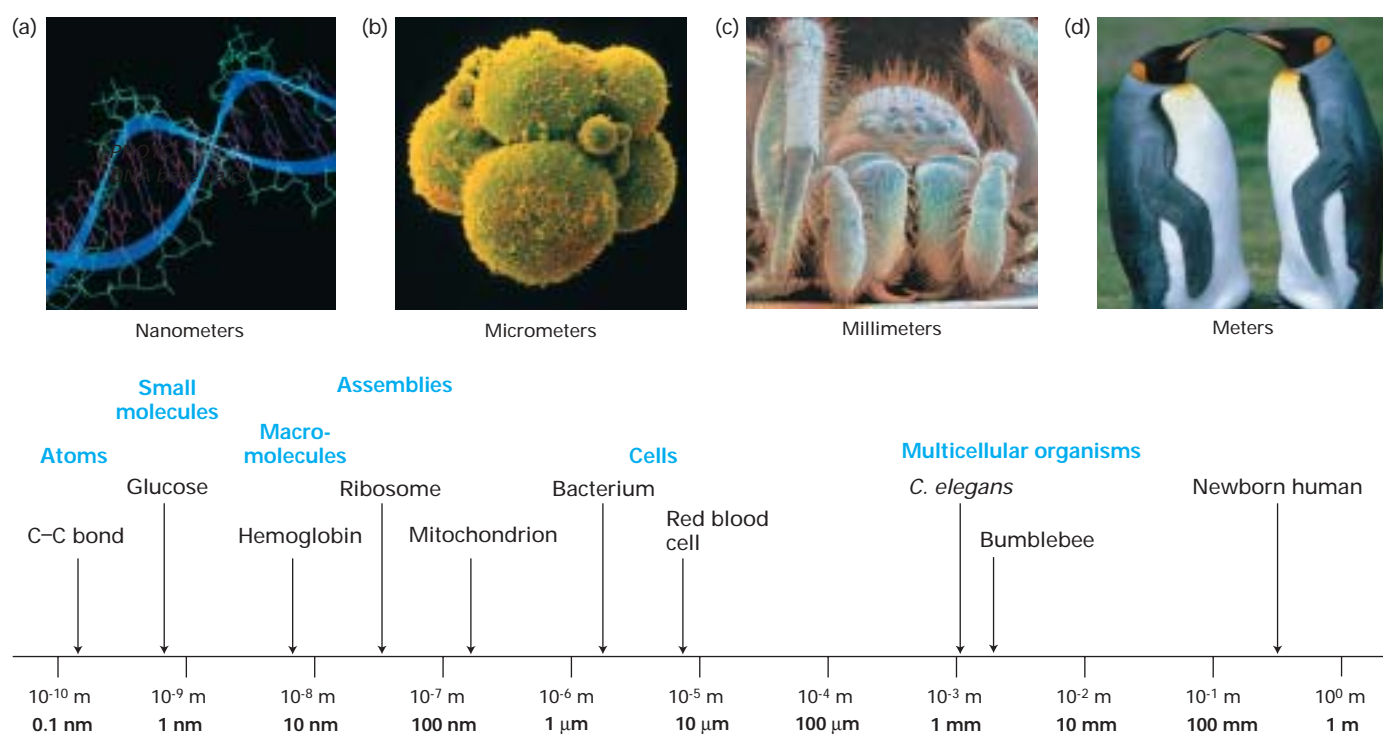
▲ **FIGURE 1-19 Apoptotic cells break apart without spewing forth cell constituents that might harm neighboring cells.** White blood cells normally look like the cell on the left. Cells undergoing programmed cell death (apoptosis), like the cell on the right, form numerous surface blebs that eventually are released. The cell is dying because it lacks certain growth signals. Apoptosis is important to eliminate virus-infected cells, remove cells where they are not needed (like the webbing that disappears as fingers develop), and to destroy immune system cells that would react with our own bodies. [Gopal Murti/Visuals Unlimited, Inc.]

fingers and thumb free to play the piano. Nerve cells in the brain soon die if they do not make proper or useful electrical connections with other cells. Some developing **lymphocytes**, the immune-system cells intended to recognize foreign proteins and polysaccharides, have the ability to react against our own tissues. Such self-reactive lymphocytes become programmed to die before they fully mature. If these cells are not weeded out before reaching maturity, they can cause autoimmune diseases in which our immune system destroys the very tissues it is meant to protect.

## 1.4 Investigating Cells and Their Parts

To build an integrated understanding of how the various molecular components that underlie cellular functions work together in a living cell, we must draw on various perspectives. Here, we look at how five disciplines—cell biology, biochemistry, genetics, genomics, and developmental biology—can contribute to our knowledge of cell structure and function. The experimental approaches of each field probe the cell's inner workings in different ways, allowing us to ask different types of questions about cells and what they do. Cell division provides a good example to illustrate the role of different perspectives in analyzing a complex cellular process.

The realm of biology ranges in scale more than a billion-fold (Figure 1-20). Beyond that, it's ecology and earth science



▲ **FIGURE 1-20 Biologists are interested in objects ranging in size from small molecules to the tallest trees.** A sampling of biological objects aligned on a logarithmic scale. (a) The DNA double helix has a diameter of about 2 nm. (b) Eight-cell-stage human embryo three days after fertilization, about 200 μm

across. (c) A wolf spider, about 15 mm across. (d) Emperor penguins are about 1 m tall. [Part (a) Will and Deni McIntyre. Part (b) Yorgas Nikas/Photo Researchers, Inc. Part (c) Gary Gaugler/Visuals Unlimited, Inc. Part (d) Hugh S. Rose/Visuals Unlimited, Inc.]

at the “macro” end, chemistry and physics at the “micro” end. The visible plants and animals that surround us are measured in meters ( $10^0$ – $10^2$  m). By looking closely, we can see a biological world of millimeters ( $1\text{ mm} = 10^{-3}$  m) and even tenths of millimeters ( $10^{-4}$  m). Setting aside oddities like chicken eggs, most cells are 1–100 micrometers ( $1\text{ }\mu\text{m} = 10^{-6}$  m) long and thus clearly visible only when magnified. To see the structures within cells, we must go farther down the size scale to 10–100 nanometers ( $1\text{ nm} = 10^{-9}$  m).

### Cell Biology Reveals the Size, Shape, and Location of Cell Components

Actual observation of cells awaited development of the first, crude microscopes in the early 1600s. A compound microscope, the most useful type of light microscope, has two lenses. The total magnifying power is the product of the magnification by each lens. As better lenses were invented, the magnifying power and the ability to distinguish closely spaced objects, the **resolution**, increased greatly. Modern compound microscopes magnify the view about a thousandfold, so that a bacterium 1 micrometer ( $1\text{ }\mu\text{m}$ ) long looks like it’s a millimeter long. Objects about  $0.2\text{ }\mu\text{m}$  apart can be discerned in these instruments.

Microscopy is most powerful when particular components of the cell are stained or labeled specifically, enabling them to be easily seen and located within the cell. A simple example is staining with dyes that bind specifically to DNA to visualize the chromosomes. Specific proteins can be detected by harnessing the binding specificity of **antibodies**, the proteins whose normal task is to help defend animals against infection and foreign substances. In general, each type of antibody binds to one protein or large polysaccharide and no other (Chapter 3). Purified antibodies can be chemically linked to a fluorescent molecule, which permits their detection in a special fluorescence microscope (Chapter 5). If a cell or tissue is treated with a detergent that partially dissolves cell membranes, fluorescent antibodies can drift in and bind to the specific protein they recognize. When the sample is viewed in the microscope, the bound fluorescent antibodies identify the location of the target protein (see Figure 1-15).

Better still is pinpointing proteins in living cells with intact membranes. One way of doing this is to introduce an engineered gene that codes for a hybrid protein: part of the hybrid protein is the cellular protein of interest; the other part is a protein that fluoresces when struck by ultraviolet light. A common fluorescent protein used for this purpose is *green fluorescent protein (GFP)*, a natural protein that makes some jellyfish colorful and fluorescent. GFP “tagging” could reveal, for instance, that a particular protein is first made on the endoplasmic reticulum and then is moved by the cell into the lysosomes. In this case, first the endoplasmic reticulum and later the lysosomes would glow in the dark.



▲ **FIGURE 1-21** During the later stages of mitosis, microtubules (red) pull the replicated chromosomes (black) toward the ends of a dividing cell. This plant cell is stained with a DNA-binding dye (ethidium) to reveal chromosomes and with fluorescent-tagged antibodies specific for tubulin to reveal microtubules. At this stage in mitosis, the two copies of each replicated chromosome (called chromatids) have separated and are moving away from each other. [Courtesy of Andrew Bajer.]

Chromosomes are visible in the light microscope only during mitosis, when they become highly condensed. The extraordinary behavior of chromosomes during mitosis first was discovered using the improved compound microscopes of the late 1800s. About halfway through mitosis, the replicated chromosomes begin to move apart. Microtubules, one of the three types of cytoskeletal filaments, participate in this movement of chromosomes during mitosis. Fluorescent tagging of tubulin, the protein subunit that polymerizes to form microtubules, reveals structural details of cell division that otherwise could not be seen and allows observation of chromosome movement (Figure 1-21).

Electron microscopes use a focused beam of electrons instead of a beam of light. In transmission electron microscopy, specimens are cut into very thin sections and placed under a high vacuum, precluding examination of living cells. The resolution of transmission electron microscopes, about 0.1 nm, permits fine structural details to be distinguished, and their powerful magnification would make a  $1\text{-}\mu\text{m}$ -long bacterial cell look like a soccer ball. Most of the organelles in eukaryotic cells and the double-layered structure of the plasma membrane were first observed with electron microscopes (Chapter 5). With new specialized electron microscopy techniques, three-dimensional models of organelles and large protein complexes can be constructed from multiple images. But to obtain a more detailed look at the individual macromolecules within cells, we must turn to techniques within the purview of biochemistry.

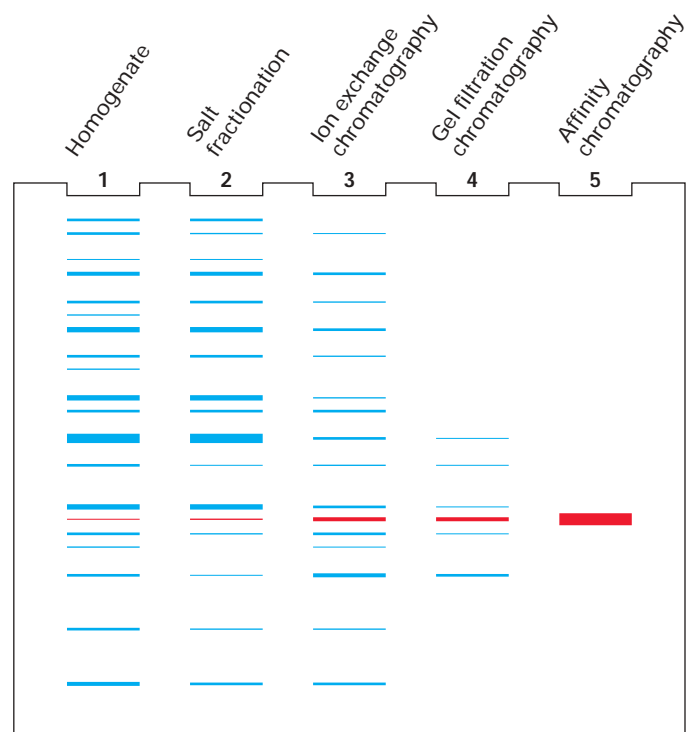
### Biochemistry Reveals the Molecular Structure and Chemistry of Purified Cell Constituents

Biochemists extract the contents of cells and separate the constituents based on differences in their chemical or physical properties, a process called *fractionation*. Of particular interest are proteins, the workhorses of many cellular processes. A typical fractionation scheme involves use of various separation techniques in a sequential fashion. These separation techniques commonly are based on differences in the size of molecules or the electrical charge on their surface (Chapter 3). To purify a particular protein of interest, a purification scheme is designed so that each step yields a preparation with fewer and fewer contaminating proteins, until finally only the protein of interest remains (Figure 1-22).

The initial purification of a protein of interest from a cell extract often is a tedious, time-consuming task. Once a small amount of purified protein is obtained, antibodies to it can be produced by methods discussed in Chapter 6. For a biochemist, antibodies are near-perfect tools for isolating larger amounts of a protein of interest for further analysis. In effect, antibodies can “pluck out” the protein they specifically recognize and bind from a semipure sample containing numerous different proteins. An increasingly common alternative is to engineer a gene that encodes a protein of interest with a small attached protein “tag,” which can be used to pull out the protein from whole cell extracts.

Purification of a protein is a necessary prelude to studies on how it catalyzes a chemical reaction or carries out other functions and how its activity is regulated. Some enzymes are made of multiple protein chains (subunits) with one chain catalyzing a chemical reaction and other chains regulating when and where that reaction occurs. The molecular machines that perform many critical cell processes constitute even larger assemblies of proteins. By separating the individual proteins composing such assemblies, their individual catalytic or other activities can be assessed. For example, purification and study of the activity of the individual proteins composing the DNA replication machine provided clues about how they work together to replicate DNA during cell division (Chapter 4).

The folded, three-dimensional structure, or **conformation**, of a protein is vital to its function. To understand the relation between the function of a protein and its form, we need to know both what it does and its detailed structure. The most widely used method for determining the complex structures of proteins, DNA, and RNA is **x-ray crystallography**. Computer-assisted analysis of the data often permits the location of every atom in a large, complex molecule to be determined. The double-helix structure of DNA, which is key to its role in heredity, was first proposed based on x-ray crystallographic studies. Throughout this book you will encounter numerous examples of protein structures as we zero in on how proteins work.



▲ FIGURE 1-22 Biochemical purification of a protein from a cell extract often requires several separation techniques.

The purification can be followed by gel electrophoresis of the starting protein mixture and the fractions obtained from each purification step. In this procedure, a sample is applied to wells in the top of a gelatin-like slab and an electric field is applied. In the presence of appropriate salt and detergent concentrations, the proteins move through the fibers of the gel toward the anode, with larger proteins moving more slowly through the gel than smaller ones (see Figure 3-32). When the gel is stained, separated proteins are visible as distinct bands whose intensities are roughly proportional to the protein concentration. Shown here are schematic depictions of gels for the starting mixture of proteins (lane 1) and samples taken after each of several purification steps. In the first step, salt fractionation, proteins that precipitated with a certain amount of salt were re-dissolved; electrophoresis of this sample (lane 2) shows that it contains fewer proteins than the original mixture. The sample then was subjected in succession to three types of column chromatography that separate proteins by electrical charge, size, or binding affinity for a particular small molecule (see Figure 3-34). The final preparation is quite pure, as can be seen from the appearance of just one protein band in lane 5. [After J. Berg et al., 2002, *Biochemistry*, W. H. Freeman and Company, p. 87]

### Genetics Reveals the Consequences of Damaged Genes

Biochemical and crystallographic studies can tell us much about an individual protein, but they cannot prove that it is required for cell division or any other cell process. The importance of a protein is demonstrated most firmly if a mu-

tation that prevents its synthesis or makes it nonfunctional adversely affects the process under study.

We define the **genotype** of an organism as its composition of genes; the term also is commonly used in reference to different versions of a single gene or a small number of genes of interest in an individual organism. A diploid organism generally carries two versions (**alleles**) of each gene, one derived from each parent. There are important exceptions, such as the genes on the X and Y chromosomes in males of some species including our own. The **phenotype** is the visible outcome of a gene's action, like blue eyes versus brown eyes or the shapes of peas. In the early days of genetics, the location and chemical identity of genes were unknown; all that could be followed were the observable characteristics, the phenotypes. The concept that genes are like "beads" on a long "string," the chromosome, was proposed early in the 1900s based on genetic work with the fruit fly *Drosophila*.

In the classical genetics approach, mutants are isolated that lack the ability to do something a normal organism can do. Often large genetic "screens" are done, looking for many different mutant individuals (e.g., fruit flies, yeast cells) that are unable to complete a certain process, such as cell division or muscle formation. In experimental organisms or cultured cells, mutations usually are produced by treatment with a **mutagen**, a chemical or physical agent that promotes mutations in a largely random fashion. But how can we isolate and maintain mutant organisms or cells that are defective in some process, such as cell division, that is necessary for survival? One way is to look for **temperature-sensitive mutants**. These mutants are able to grow at one temperature, the *permissive* temperature, but not at another, usually higher temperature, the *nonpermissive* temperature. Normal cells can grow at either temperature. In most cases, a temperature-sensitive mutant produces an altered protein that works at the permissive temperature but unfolds and is nonfunctional at the nonpermissive temperature. Temperature-sensitive screens are readily done with viruses, bacteria, yeast, roundworms, and fruit flies.

By analyzing the effects of numerous different temperature-sensitive mutations that altered cell division, geneticists discovered all the genes necessary for cell division without knowing anything, initially, about which proteins they encode or how these proteins participate in the process. The great power of genetics is to reveal the existence and relevance of proteins without prior knowledge of their biochemical identity or molecular function. Eventually these "mutation-defined" genes were isolated and replicated (cloned) with **recombinant DNA** techniques discussed in Chapter 9. With the isolated genes in hand, the encoded proteins could be produced in the test tube or in engineered bacteria or cultured cells. Then the biochemists could investigate whether the proteins associate with other proteins or DNA or catalyze particular chemical reactions during cell division (Chapter 21).

The analysis of genome sequences from various organisms during the past decade has identified many previously unknown DNA regions that are likely to encode proteins

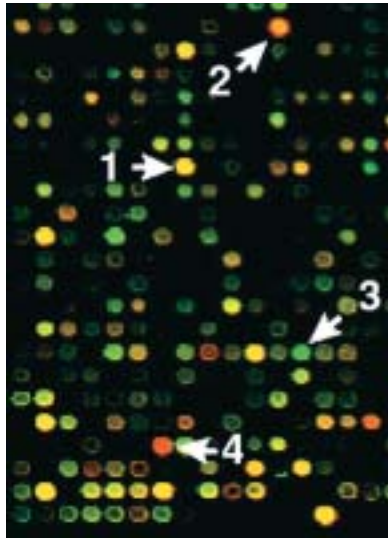
(i.e., protein-coding genes). The general function of the protein encoded by a sequence-identified gene may be deduced by analogy with known proteins of similar sequence. Rather than randomly isolating mutations in novel genes, several techniques are now available for inactivating specific genes by engineering mutations into them (Chapter 9). The effects of such deliberate gene-specific mutations provide information about the role of the encoded proteins in living organisms. This application of genetic techniques starts with a gene/protein sequence and ends up with a mutant phenotype; traditional genetics starts with a mutant phenotype and ends up with a gene/protein sequence.

### Genomics Reveals Differences in the Structure and Expression of Entire Genomes

Biochemistry and genetics generally focus on one gene and its encoded protein at a time. While powerful, these traditional approaches do not give a comprehensive view of the structure and activity of an organism's genome, its entire set of genes. The field of **genomics** does just that, encompassing the molecular characterization of whole genomes and the determination of global patterns of gene expression. The recent completion of the genome sequences for more than 80 species of bacteria and several eukaryotes now permits comparisons of entire genomes from different species. The results provide overwhelming evidence of the molecular unity of life and the evolutionary processes that made us what we are (see Section 1.5). Genomics-based methods for comparing thousands of pieces of DNA from different individuals all at the same time are proving useful in tracing the history and migrations of plants and animals and in following the inheritance of diseases in human families.

New methods using **DNA microarrays** can simultaneously detect all the mRNAs present in a cell, thereby indicating which genes are being transcribed. Such global patterns of gene expression clearly show that liver cells transcribe a quite different set of genes than do white blood cells or skin cells. Changes in gene expression also can be monitored during a disease process, in response to drugs or other external signals, and during development. For instance, the recent identification of all the mRNAs present in cultured fibroblasts before, during, and after they divide has given us an overall view of transcriptional changes that occur during cell division (Figure 1-23). Cancer diagnosis is being transformed because previously indistinguishable cancer cells have distinct gene expression patterns and prognoses (Chapter 23). Similar studies with different organisms and cell types are revealing what is universal about the genes involved in cell division and what is specific to particular organisms.

The entire complement of proteins in a cell, its **proteome**, is controlled in part by changes in gene transcription. The regulated synthesis, processing, localization, and degradation of specific proteins also play roles in determining the proteome of a particular cell, and the association of certain proteins with one another is critical to the functional abilities



▲ **FIGURE 1-23 DNA microarray analysis gives a global view of changes in transcription following addition of serum to cultured human cells.** Serum contains growth factors that stimulate nondividing cells to begin growing and dividing. DNA microarray analysis can detect the relative transcription of genes in two different cell populations (see Figure 9-35). The microarray consists of tiny spots of DNA attached to a microscope slide. Each spot contains many copies of a DNA sequence from a single human gene. One preparation of RNA, containing all the different types of RNA being made in nongrowing cells cultured without serum, is labeled with green fluorescent molecules. Another RNA population from growing, serum-treated, cells is labeled with red. The two are mixed and hybridized to the slide, where they "zipper up" with their corresponding genes. Green spots (e.g., spot 3) therefore indicate genes that are transcribed in nondividing (serum-deprived) cells; red spots (e.g., spot 4) indicate genes that are transcribed in dividing cells, and yellow spots (e.g., spots 1 and 2) indicate genes that are transcribed equally in dividing and nondividing cells. [From V. R. Iyer et al., 1999, *Science* 283:83.]

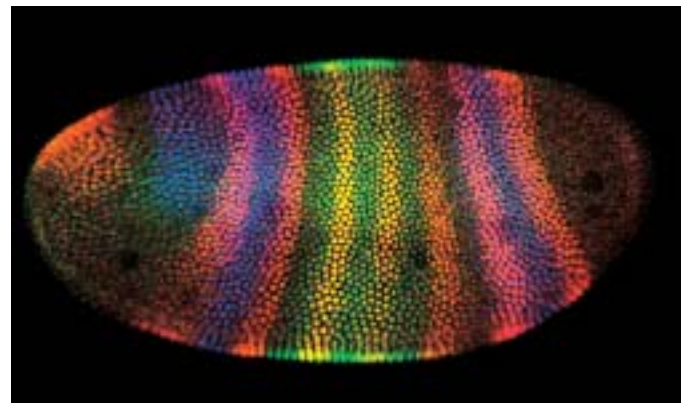
of cells. New techniques for monitoring the presence and interactions of numerous proteins simultaneously, called **proteomics**, are one way of assembling a comprehensive view of the proteins and molecular machines important for cell functioning. The field of proteomics will advance dramatically once high-throughput x-ray crystallography, currently under development, permits researchers to rapidly determine the structures of hundreds or thousands of proteins.

### Developmental Biology Reveals Changes in the Properties of Cells as They Specialize

Another approach to viewing cells comes from studying how they change during development of a complex organism. Bacteria, algae, and unicellular eukaryotes (protozoans, yeasts) often, but by no means always, can work solo. The concerted actions of the trillions of cells that compose our

bodies require an enormous amount of communication and division of labor. During the development of multicellular organisms, differentiation processes form hundreds of cell types, each specialized for a particular task: transmission of electrical signals by neurons, transport of oxygen by red blood cells, destruction of infecting bacteria by macrophages, contraction by muscle cells, chemical processing by liver cells.

Many of the differences among differentiated cells are due to production of specific sets of proteins needed to carry out the unique functions of each cell type. That is, only a subset of an organism's genes is transcribed at any given time or in any given cell. Such **differential gene expression** at different times or in different cell types occurs in bacteria, fungi, plants, animals, and even viruses. Differential gene expression is readily apparent in an early fly embryo in which all the cells look alike until they are stained to detect the proteins encoded by particular genes (Figure 1-24). Transcription can change within one cell type in response to an external signal or in accordance with a biological clock; some genes, for instance, undergo a daily cycle between low and high transcription rates.



▲ **FIGURE 1-24 Differential gene expression can be detected in early fly embryos before cells are morphologically different.** An early *Drosophila* embryo has about 6000 cells covering its surface, most of which are indistinguishable by simple light microscopy. If the embryo is made permeable to antibodies with a detergent that partially dissolves membranes, the antibodies can find and bind to the proteins they recognize. In this embryo we see antibodies tagged with a fluorescent label bound to proteins that are in the nuclei; each small sphere corresponds to one nucleus. Three different antibodies were used, each specific for a different protein and each giving a distinct color (yellow, green, or blue) in a fluorescence microscope. The red color is added to highlight overlaps between the yellow and blue stains. The locations of the different proteins show that the cells are in fact different at this early stage, with particular genes turned on in specific stripes of cells. These genes control the subdivision of the body into repeating segments, like the black and yellow stripes of a hornet. [Courtesy of Sean Carroll, University of Wisconsin.]

Producing different kinds of cells is not enough to make an organism, any more than collecting all the parts of a truck in one pile gives you a truck. The various cell types must be organized and assembled into all the tissues and organs. Even more remarkable, these body parts must work almost immediately after their formation and continue working during the growth process. For instance, the human heart begins to beat when it is less than 3 mm long, when we are mere 23-day-old embryos, and continues beating as it grows into a fist-size muscle. From a few hundred cells to billions, and still ticking.

In the developing organism, cells grow and divide at some times and not others, they assemble and communicate, they prevent or repair errors in the developmental process, and they coordinate each tissue with others. In the adult organism, cell division largely stops in most organs. If part of an organ such as the liver is damaged or removed, cell division resumes until the organ is regenerated. The legend goes that Zeus punished Prometheus for giving humans fire by chaining him to a rock and having an eagle eat his liver. The punishment was eternal because, as the Greeks evidently knew, the liver regenerates.

Developmental studies involve watching where, when, and how different kinds of cells form, discovering which signals trigger and coordinate developmental events, and understanding the differential gene action that underlies differentiation (Chapters 15 and 22). During development we can see cells change in their normal context of other cells. Cell biology, biochemistry, cell biology, genetics, and genomics approaches are all employed in studying cells during development.

### Choosing the Right Experimental Organism for the Job

Our current understanding of the molecular functioning of cells rests on studies with viruses, bacteria, yeast, protozoa, slime molds, plants, frogs, sea urchins, worms, insects, fish, chickens, mice, and humans. For various reasons, some organisms are more appropriate than others for answering particular questions. Because of the evolutionary conservation of genes, proteins, organelles, cell types, and so forth, discoveries about biological structures and functions obtained with one experimental organism often apply to others. Thus researchers generally conduct studies with the organism that is most suitable for rapidly and completely answering the question being posed, knowing that the results obtained in one organism are likely to be broadly applicable. Figure 1-25 summarizes the typical experimental uses of various organisms whose genomes have been sequenced completely or nearly so. The availability of the genome sequences for these organisms makes them particularly useful for genetics and genomics studies.

Bacteria have several advantages as experimental organisms: They grow rapidly, possess elegant mechanisms for controlling gene activity, and have powerful genetics. This

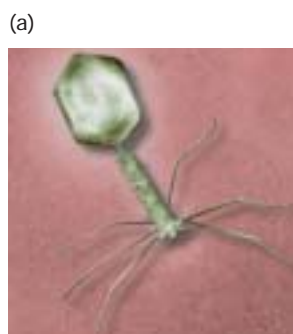
### ► FIGURE 1-25 Each experimental organism used in cell biology has advantages for certain types of studies.

Viruses and bacteria have small genomes amenable to genetic dissection. Many insights into gene control initially came from studies with these organisms. The yeast *Saccharomyces cerevisiae* has the cellular organization of a eukaryote but is a relatively simple single-celled organism that is easy to grow and to manipulate genetically. In the nematode worm *Caenorhabditis elegans*, which has a small number of cells arranged in a nearly identical way in every worm, the formation of each individual cell can be traced. The fruit fly *Drosophila melanogaster*, first used to discover the properties of chromosomes, has been especially valuable in identifying genes that control embryonic development. Many of these genes are evolutionarily conserved in humans. The zebrafish *Danio rerio* is used for rapid genetic screens to identify genes that control development and organogenesis. Of the experimental animal systems, mice (*Mus musculus*) are evolutionarily the closest to humans and have provided models for studying numerous human genetic and infectious diseases. The mustard-family weed *Arabidopsis thaliana*, sometimes described as the *Drosophila* of the plant kingdom, has been used for genetic screens to identify genes involved in nearly every aspect of plant life. Genome sequencing is completed for many viruses and bacterial species, the yeast *Saccharomyces cerevisiae*, the roundworm *C. elegans*, the fruit fly *D. melanogaster*, humans, and the plant *Arabidopsis thaliana*. It is mostly completed for mice and in progress for zebrafish. Other organisms, particularly frogs, sea urchins, chickens, and slime molds, continue to be immensely valuable for cell biology research. Increasingly, a wide variety of other species are used, especially for studies of evolution of cells and mechanisms. [Part (a) Visuals Unlimited, Inc. Part (b) Kari Lountmaa/Science Photo Library/Photo Researchers, Inc. Part (c) Scimat/Photo Researchers, Inc. Part (d) Photo Researchers, Inc. Part (e) Darwin Dale/Photo Researchers, Inc. Part (f) Inge Spence/Visuals Unlimited, Inc. Part (g) J. M. Labat/Jancana/Visuals Unlimited, Inc. Part (h) Darwin Dale/Photo Researchers, Inc.]

latter property relates to the small size of bacterial genomes, the ease of obtaining mutants, the availability of techniques for transferring genes into bacteria, an enormous wealth of knowledge about bacterial gene control and protein functions, and the relative simplicity of mapping genes relative to one another in the genome. Single-celled yeasts not only have some of the same advantages as bacteria, but also possess the cell organization, marked by the presence of a nucleus and organelles, that is characteristic of all eukaryotes.

Studies of cells in specialized tissues make use of animal and plant “models,” that is, experimental organisms with attributes typical of many others. Nerve cells and muscle cells, for instance, traditionally were studied in mammals or in creatures with especially large or accessible cells, such as the giant neural cells of the squid and sea hare or the flight muscles of birds. More recently, muscle and nerve development have been extensively studied in fruit flies (*Drosophila melanogaster*), roundworms (*Caenorhabditis elegans*), and zebrafish in which mutants can be readily isolated. Organisms with large-celled embryos that develop outside the

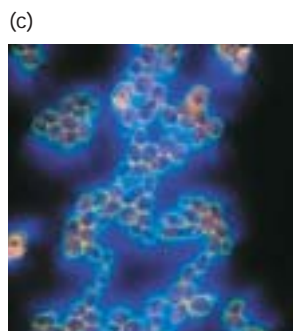




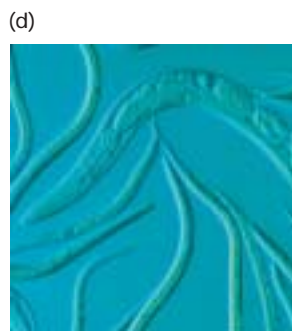
**Viruses**  
 Proteins involved in DNA, RNA, protein synthesis  
 Gene regulation  
 Cancer and control of cell proliferation  
 Transport of proteins and organelles inside cells  
 Infection and immunity  
 Possible gene therapy approaches



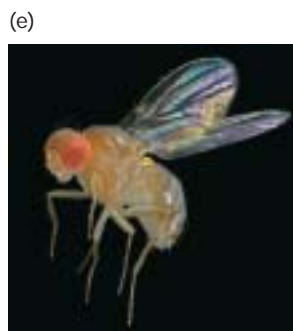
**Bacteria**  
 Proteins involved in DNA, RNA, protein synthesis, metabolism  
 Gene regulation  
 Targets for new antibiotics  
 Cell cycle  
 Signaling



**Yeast (*Saccharomyces cerevisiae*)**  
 Control of cell cycle and cell division  
 Protein secretion and membrane biogenesis  
 Function of the cytoskeleton  
 Cell differentiation  
 Aging  
 Gene regulation and chromosome structure



**Roundworm (*Caenorhabditis elegans*)**  
 Development of the body plan  
 Cell lineage  
 Formation and function of the nervous system  
 Control of programmed cell death  
 Cell proliferation and cancer genes  
 Aging  
 Behavior  
 Gene regulation and chromosome structure



**Fruit fly (*Drosophila melanogaster*)**  
 Development of the body plan  
 Generation of differentiated cell lineages  
 Formation of the nervous system, heart, and musculature  
 Programmed cell death  
 Genetic control of behavior  
 Cancer genes and control of cell proliferation  
 Control of cell polarization  
 Effects of drugs, alcohol, pesticides



**Zebrafish**  
 Development of vertebrate body tissues  
 Formation and function of brain and nervous system  
 Birth defects  
 Cancer



**Mice, including cultured cells**  
 Development of body tissues  
 Function of mammalian immune system  
 Formation and function of brain and nervous system  
 Models of cancers and other human diseases  
 Gene regulation and inheritance  
 Infectious disease



**Plant (*Arabidopsis thaliana*)**  
 Development and patterning of tissues  
 Genetics of cell biology  
 Agricultural applications  
 Physiology  
 Gene regulation  
 Immunity  
 Infectious disease

mother (e.g., frogs, sea urchins, fish, and chickens) are extremely useful for tracing the fates of cells as they form different tissues and for making extracts for biochemical studies. For

instance, a key protein in regulating mitosis was first identified in studies with frog and sea urchin embryos and subsequently purified from extracts (Chapter 21).

Using recombinant DNA techniques researchers can engineer specific genes to contain mutations that inactivate or increase production of their encoded proteins. Such genes can be introduced into the embryos of worms, flies, frogs, sea urchins, chickens, mice, a variety of plants, and other organisms, permitting the effects of activating a gene abnormally or inhibiting a normal gene function to be assessed. This approach is being used extensively to produce mouse versions of human genetic diseases. New techniques specifically for inactivating particular genes by injecting short pieces of RNA are making quick tests of gene functions possible in many organisms.

Mice have one enormous advantage over other experimental organisms: they are the closest to humans of any animal for which powerful genetic approaches are feasible. Engineered mouse genes carrying mutations similar to those associated with a particular inherited disease in humans can be introduced into mouse embryonic stem (ES) cells. These cells can be injected into an early embryo, which is then implanted into a pseudopregnant female mouse (Chapter 9). If the mice that develop from the injected ES cells exhibit diseases similar to the human disease, then the link between the disease and mutations in a particular gene or genes is supported. Once mouse models of a human disease are available, further studies on the molecular defects causing the disease can be done and new treatments can be tested, thereby minimizing human exposure to untested treatments.

A continuous unplanned genetic screen has been performed on human populations for millennia. Thousands of inherited traits have been identified and, more recently, mapped to locations on the chromosomes. Some of these traits are inherited propensities to get a disease; others are eye color or other minor characteristics. Genetic variations in virtually every aspect of cell biology can be found in human populations, allowing studies of normal and disease states and of variant cells in culture.

Less-common experimental organisms offer possibilities for exploring unique or exotic properties of cells and for studying standard properties of cells that are exaggerated in a useful fashion in a particular animal. For example, the ends of chromosomes, the **telomeres**, are extremely dilute in most cells. Human cells typically contain 92 telomeres (46 chromosomes  $\times$  2 ends per chromosome). In contrast, some protozoa with unusual “fragmented” chromosomes contain millions of telomeres per cell. Recent discoveries about telomere structure have benefited greatly from using this natural variation for experimental advantage.

## 1.5 A Genome Perspective on Evolution

Comprehensive studies of genes and proteins from many organisms are giving us an extraordinary documentation of the history of life. We share with other eukaryotes thousands of

individual proteins, hundreds of macromolecular machines, and most of our organelles, all as a result of our shared evolutionary history. New insights into molecular cell biology arising from genomics are leading to a fuller appreciation of the elegant molecular machines that arose during billions of years of genetic tinkering and evolutionary selection for the most efficient, precise designs. Despite all that we currently know about cells, many new proteins, new macromolecular assemblies, and new activities of known ones remain to be discovered. Once a more complete description of cells is in hand, we will be ready to fully investigate the rippling, flowing dynamics of living systems.

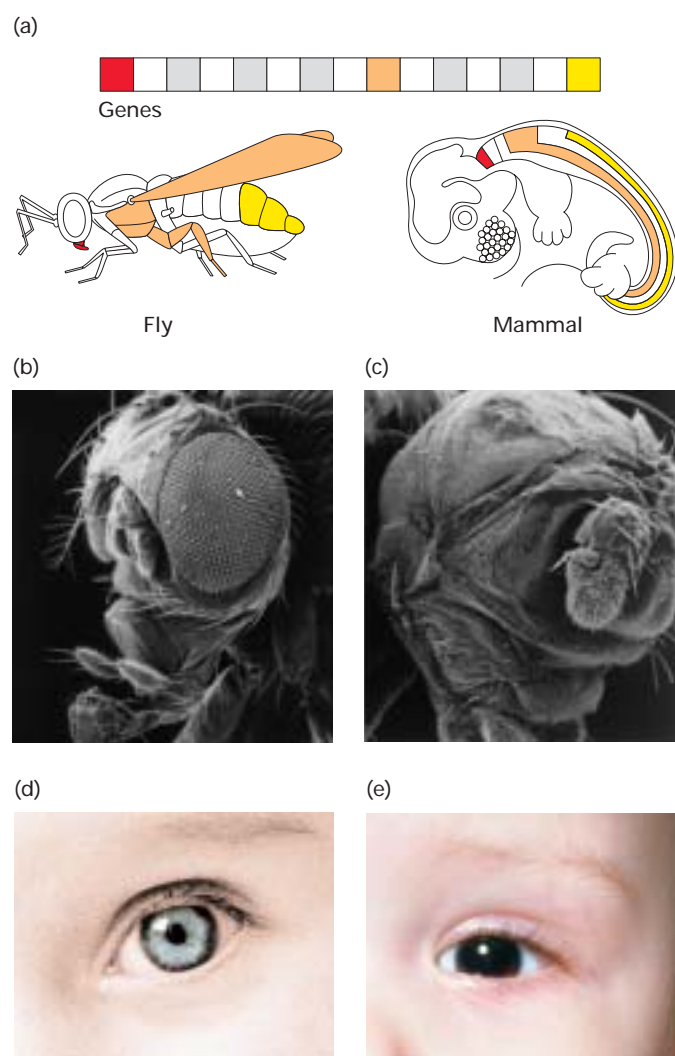
### Metabolic Proteins, the Genetic Code, and Organelle Structures Are Nearly Universal

Even organisms that look incredibly different share many biochemical properties. For instance, the enzymes that catalyze degradation of sugars and many other simple chemical reactions in cells have similar structures and mechanisms in most living things. The genetic code whereby the nucleotide sequences of mRNA specifies the amino acid sequences of proteins can be read equally well by a bacterial cell and a human cell. Because of the universal nature of the genetic code, bacterial “factories” can be designed to manufacture growth factors, insulin, clotting factors, and other human proteins with therapeutic uses. The biochemical similarities among organisms also extend to the organelles found in eukaryotic cells. The basic structures and functions of these subcellular components are largely conserved in all eukaryotes.

Computer analysis of DNA sequence data, now available for numerous bacterial species and several eukaryotes, can locate protein-coding genes within genomes. With the aid of the genetic code, the amino acid sequences of proteins can be deduced from the corresponding gene sequences. Although simple conceptually, “finding” genes and deducing the amino acid sequences of their encoded proteins is complicated in practice because of the many noncoding regions in eukaryotic DNA (Chapter 9). Despite the difficulties and occasional ambiguities in analyzing DNA sequences, comparisons of the genomes from a wide range of organisms provide stunning, compelling evidence for the conservation of the molecular mechanisms that build and change organisms and for the common evolutionary history of all species.

### Many Genes Controlling Development Are Remarkably Similar in Humans and Other Animals

As humans, we probably have a biased and somewhat exaggerated view of our status in the animal kingdom. Pride in our swollen forebrain and its associated mental capabilities may blind us to the remarkably sophisticated abilities of other species: navigation by birds, the sonar system of bats, homing by salmon, or the flight of a fly.



◀ **FIGURE 1-26 Similar genes, conserved during evolution, regulate many developmental processes in diverse animals.**

Insects and mammals are estimated to have had a common ancestor about half a billion years ago. They share genes that control similar processes, such as growth of heart and eyes and organization of the body plan, indicating conservation of function from ancient times. (a) *Hox* genes are found in clusters on the chromosomes of most or all animals. *Hox* genes encode related proteins that control the activities of other genes. *Hox* genes direct the development of different segments along the head-to-tail axis of many animals as indicated by corresponding colors. Each gene is activated (transcriptionally) in a specific region along the head-to-tail axis and controls the growth of tissues there. For example, in mice the *Hox* genes are responsible for the distinctive shapes of vertebrae. Mutations affecting *Hox* genes in flies cause body parts to form in the wrong locations, such as legs in lieu of antennae on the head. These genes provide a head-to-tail address and serve to direct formation of the right structures in the right places. (b) Development of the large compound eyes in fruit flies requires a gene called *eyeless* (named for the mutant phenotype). (c) Flies with inactivated *eyeless* genes lack eyes. (d) Normal human eyes require the human gene, called *Pax6*, that corresponds to *eyeless*. (e) People lacking adequate *Pax6* function have the genetic disease *aniridia*, a lack of irises in the eyes. *Pax6* and *eyeless* encode highly related proteins that regulate the activities of other genes, and are descended from the same ancestral gene. [Parts (a) and (b) Andreas Hefti, Interdepartmental Electron Microscopy (IEM) Biocenter, University of Basel. Part (d) © Simon Fraser/Photo Researchers, Inc.]

This is not to say that all genes or proteins are evolutionarily conserved. Many striking examples exist of proteins that, as far as we can tell, are utterly absent from certain lineages of animals. Plants, not surprisingly, exhibit many such differences from animals after a billion-year separation in their evolution. Yet certain DNA-binding proteins differ between peas and cows at only two amino acids out of 102!

### Darwin's Ideas About the Evolution of Whole Animals Are Relevant to Genes

Darwin did not know that genes exist or how they change, but we do: the DNA replication machine makes an error, or a mutagen causes replacement of one nucleotide with another or breakage of a chromosome. Some changes in the genome are innocuous, some mildly harmful, some deadly; a very few are beneficial. Mutations can change the sequence of a gene in a way that modifies the activity of the encoded protein or alters when, where, and in what amounts the protein is produced in the body.

Gene-sequence changes that are harmful will be lost from a population of organisms because the affected individuals cannot survive as well as their relatives. This selection process is exactly what Darwin described without knowing the underlying mechanisms that cause organisms to vary. Thus the selection of whole organisms for survival is really a selection of genes, or more accurately sets of genes. A population of organisms often contains many variants that are

Despite all the evidence for evolutionary unity at the cellular and physiological levels, everyone expected that genes regulating animal development would differ greatly from one phylum to the next. After all, insects and sea urchins and mammals look *so* different. We must have many unique proteins to create a brain like ours . . . or must we? The fruits of research in developmental genetics during the past two decades reveal that insects and mammals, which have a common ancestor about half a billion years ago, possess many similar development-regulating genes (Figure 1-26). Indeed, a large number of these genes appear to be conserved in many and perhaps all animals. Remarkably, the developmental functions of the proteins encoded by these genes are also often preserved. For instance, certain proteins involved in eye development in insects are related to protein regulators of eye development in mammals. Same for development of the heart, gut, lungs, and capillaries and for placement of body parts along the head-to-tail and back-to-front body axes (Chapter 15).

all roughly equally well-suited to the prevailing conditions. When conditions change—a fire, a flood, loss of preferred food supply, climate shift—variants that are better able to adapt will survive, and those less suited to the new conditions will begin to die out. In this way, the genetic composition of a population of organisms can change over time.

### Human Medicine Is Informed by Research on Other Organisms

Mutations that occur in certain genes during the course of our lives contribute to formation of various human cancers. The normal, wild-type forms of such “cancer-causing” genes generally encode proteins that help regulate cell proliferation or death (Chapter 23). We also can inherit from our parents mutant copies of genes that cause all manner of genetic diseases, such as cystic fibrosis, muscular dystrophy, sickle cell anemia, and Huntington’s disease. Happily we can also inherit genes that make us robustly resist disease. A remarkable number of genes associated with cancer and other human

diseases are present in evolutionarily distant animals. For example, a recent study shows that more than three-quarters of the known human disease genes are related to genes found in the fruit fly *Drosophila*.

With the identification of human disease genes in other organisms, experimental studies in experimentally tractable organisms should lead to rapid progress in understanding the normal functions of the disease-related genes and what occurs when things go awry. Conversely, the disease states themselves constitute a genetic analysis with well-studied phenotypes. All the genes that can be altered to cause a certain disease may encode a group of functionally related proteins. Thus clues about the normal cellular functions of proteins come from human diseases and can be used to guide initial research into mechanism. For instance, genes initially identified because of their link to cancer in humans can be studied in the context of normal development in various model organisms, providing further insight about the functions of their protein products.