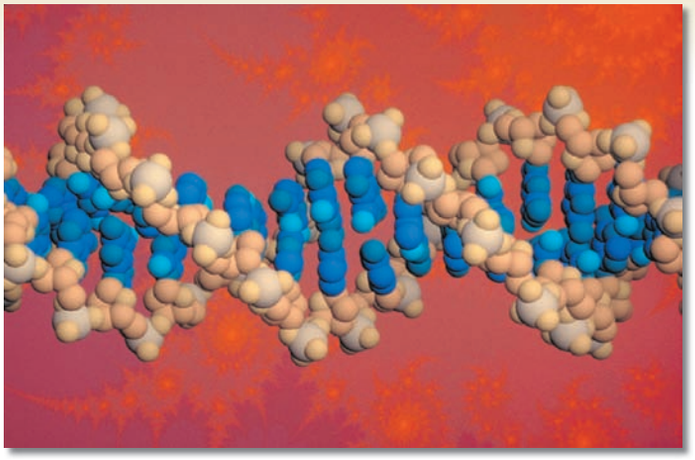
CHAPTER OUTLINE



A molecular model showing the structure of the DNA double helix.

9 MOLECULAR STRUCTURE OF DNA AND RNA

In Chapters 2 through 8, we focused on the relationship between the inheritance of genes and chromosomes and the outcome of an organism's traits. In Chapter 9, we will shift our attention to **molecular genetics**—the study of DNA structure and function at the molecular level. An exciting goal of molecular genetics is to use our knowledge of DNA structure to understand how DNA functions as the genetic material. Using molecular techniques, researchers have determined the organization of many genes. This information, in turn, has helped us understand how the expression of such genes governs the outcome of an individual's inherited traits.

The past several decades have seen dramatic advances in techniques and approaches to investigate and even to alter the genetic material. These advances have greatly expanded our understanding of molecular genetics and also have provided key insights into the mechanisms underlying transmission and population genetics. Molecular genetic technology is also widely used in supporting disciplines such as biochemistry, cell biology, and microbiology.

To a large extent, our understanding of genetics comes from our knowledge of the molecular structure of **DNA (deoxyribonucleic acid)** and **RNA (ribonucleic acid).** In this chapter, we will begin by considering classic experiments which showed that DNA is the genetic material. We will then survey the molecular features of DNA and RNA that underlie their function.

9.1 IDENTIFICATION OF DNA AS THE GENETIC MATERIAL

To fulfill its role, the genetic material must meet four criteria.

1. **Information:** The genetic material must contain the information necessary to construct an entire organism. In other words, it must provide the blueprint to determine the inherited traits of an organism.
2. **Transmission:** During reproduction, the genetic material must be passed from parents to offspring.
3. **Replication:** Because the genetic material is passed from parents to offspring, and from mother cell to daughter cells during cell division, it must be copied.
4. **Variation:** Within any species, a significant amount of phenotypic variability occurs. For example, Mendel studied several characteristics in pea plants that varied among different plants. These included height (tall versus dwarf) and seed color (yellow versus green). Therefore, the genetic material must also vary in ways that can account for the known phenotypic differences within each species.

Along with Mendel's work, the data of many other geneticists in the early 1900s were consistent with these four properties: information, transmission, replication, and variation. However, the experimental study of genetic crosses cannot, by itself, identify the chemical nature of the genetic material.

In the 1880s, August Weismann and Carl Nägeli championed the idea that a chemical substance within living cells is responsible for the transmission of traits from parents to offspring. The chromosome theory of inheritance was developed, and experimentation demonstrated that the chromosomes are the carriers of the genetic material (see Chapter 3). Nevertheless, the story was not complete because chromosomes contain both DNA and proteins. Also, RNA is found in the vicinity of chromosomes. Therefore, further research was needed to precisely identify the genetic material. In this section, we will examine the first experimental approaches to achieve this goal.

Experiments with Pneumococcus Suggested That DNA Is the Genetic Material

Some early work in microbiology was important in developing an experimental strategy to identify the genetic material. Frederick Griffith studied a type of bacterium known then as pneumococci and now classified as *Streptococcus pneumoniae*. Certain strains of *S. pneumoniae* secrete a polysaccharide capsule, whereas other strains do not. When streaked onto petri plates containing a solid growth medium, capsule-secreting strains have a smooth colony morphology, whereas those strains unable to secrete a capsule have a rough appearance.

The different forms of *S. pneumoniae* also affect their virulence, or ability to cause disease. When smooth strains of *S. pneumoniae* infect a mouse, the capsule allows the bacteria to escape attack by the mouse's immune system. As a result, the bacteria can grow and eventually kill the mouse. In contrast, the nonencapsulated (rough) bacteria are destroyed by the animal's immune system.

In 1928, Griffith conducted experiments that involved the injection of live and/or heat-killed bacteria into mice. He then observed whether or not the bacteria caused a lethal infection. Griffith was working with two strains of *S. pneumoniae*, a type S (S for smooth) and a type R (R for rough). When injected into a live mouse, the type S bacteria proliferated within the mouse's bloodstream and ultimately killed the mouse (**Figure 9.1a**). Following the death of the mouse, Griffith found many type S bacteria within the mouse's blood. In contrast, when type R bacteria were injected into a mouse, the mouse lived (**Figure 9.1b**). To verify that the proliferation of the smooth bacteria was causing the death of the mouse, Griffith killed the smooth bacteria with heat treatment before injecting them into the mouse. In this case, the mouse also survived (**Figure 9.1c**).

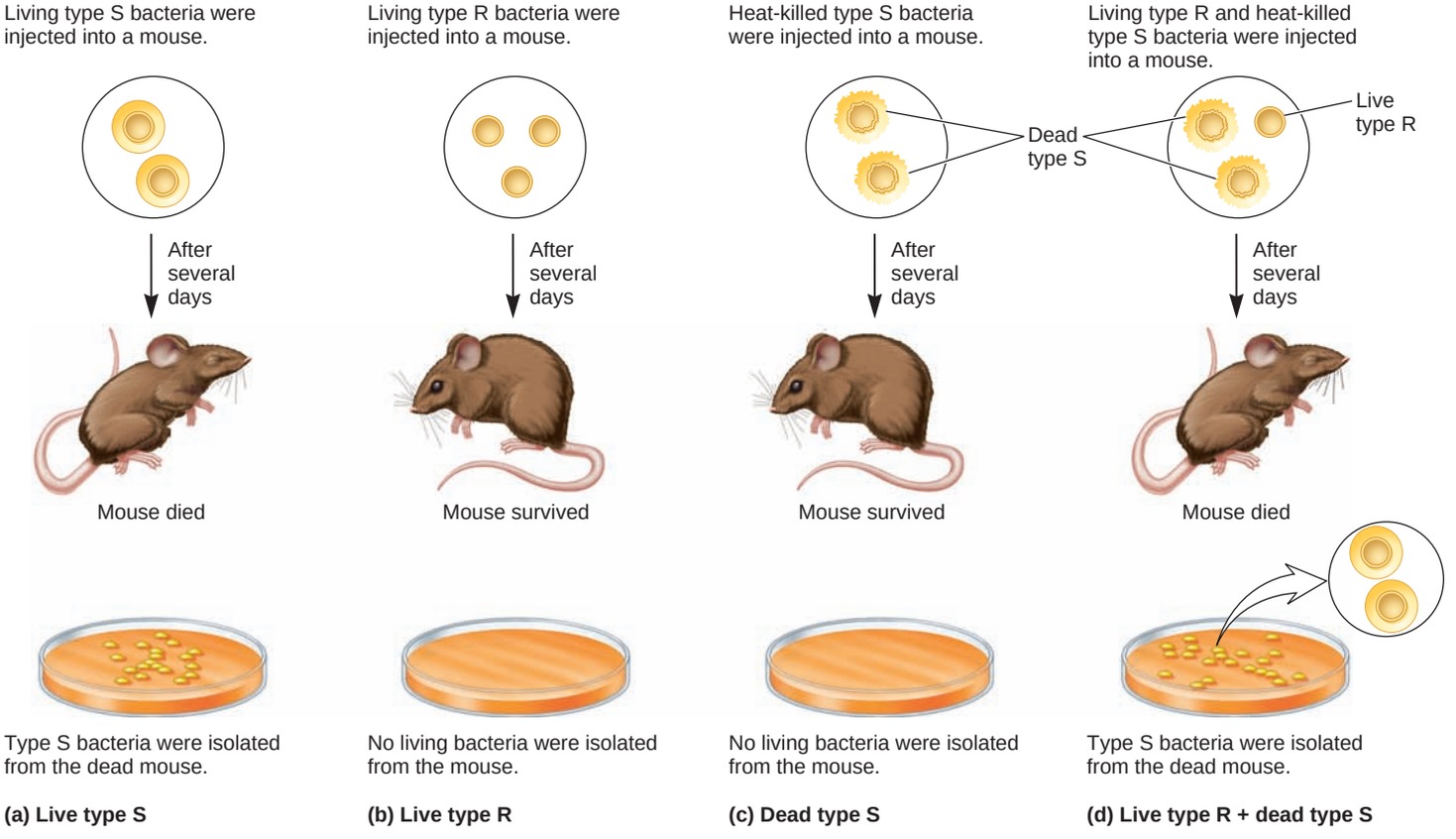


FIGURE 9.1 **Griffith's experiments on genetic transformation in pneumococcus.**

The critical and unexpected result was obtained in the experiment outlined in **Figure 9.1d**. In this experiment, live type R bacteria were mixed with heat-killed type S bacteria. As shown here, the mouse died. Furthermore, extracts from tissues of the dead mouse were found to contain living type S bacteria! What can account for these results? Because living type R bacteria alone could not proliferate and kill the mouse (Figure 9.1b), the interpretation of the data in Figure 9.1d is that something from the dead type S bacteria was transforming the type R bacteria into type S. Griffith called this process **transformation,** and the unidentified substance causing this to occur was termed the transformation principle. The steps of bacterial transformation are described in Chapter 7 (see Figure 7.13).

At this point, let's look at what Griffith's observations mean in genetic terms. The transformed bacteria acquired the *information* to make a capsule. Among different strains, *variation* exists in the ability to create a capsule and to cause mortality in mice. The genetic material that is necessary to create a capsule must be *replicated* so that it can be *transmitted* from mother to daughter cells during cell division. Taken together, these observations are consistent with the idea that the formation of a capsule is governed by the bacteria's genetic material, meeting the four criteria described previously. Griffith's experiments showed that some genetic material from the dead bacteria had been transferred to the living bacteria and provided them with a new trait. However, Griffith did not know what the transforming substance was.

Important scientific discoveries often take place when researchers recognize that someone else's experimental observations can be used to address a particular scientific question. Oswald Avery, Colin MacLeod, and Maclyn McCarty realized that Griffith's observations could be used as part of an experimental strategy to identify the genetic material. They asked the question, What substance is being transferred from the dead type S bacteria to the live type R? To answer this question, they incorporated additional biochemical techniques into their experimental methods.

At the time of these experiments in the 1940s, researchers already knew that DNA, RNA, proteins, and carbohydrates are major constituents of living cells. To separate these components and to determine if any of them was the genetic material, Avery, MacLeod, and McCarty used established biochemical purification procedures and prepared bacterial extracts from type S strains containing each type of these molecules. After many repeated attempts with different types of extracts, they discovered that only one of the extracts, namely, the one that contained purified DNA, was able to convert the type R bacteria into type S. As shown in **Figure 9.2**, when this extract was mixed with

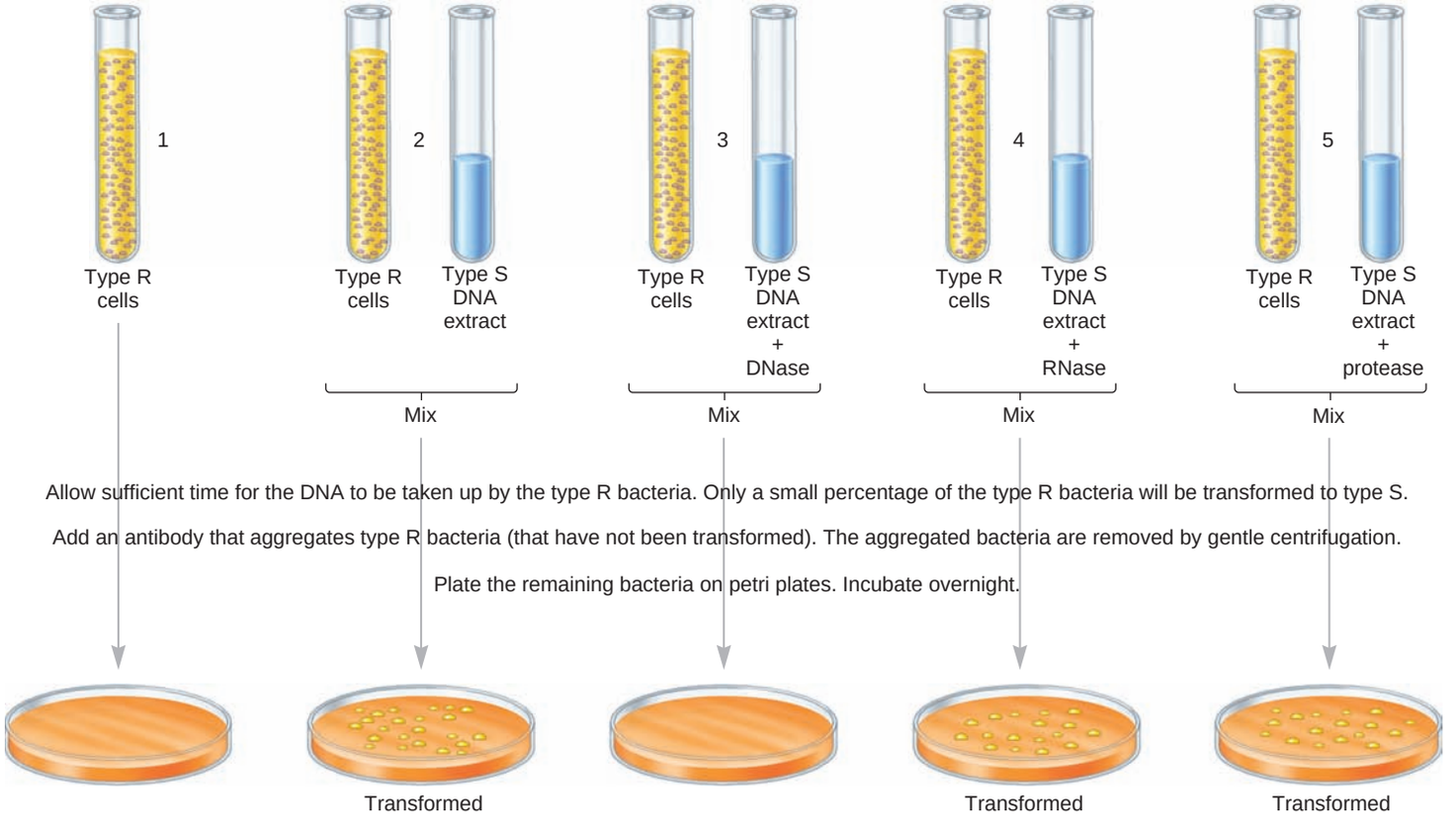


FIGURE 9.2 Experimental protocol used by Avery, MacLeod, and McCarty to identify the transforming principle. Samples of *Streptococcus pneumoniae* cells were either not exposed to a type S DNA extract (tube 1) or exposed to a type S DNA extract (tubes 2–5). Tubes 3, 4, and 5 also contained DNase, RNase, or protease, respectively. After incubation, the cells were exposed to antibodies, which are molecules that can specifically recognize the molecular structure of macromolecules. In this experiment, the antibodies recognized the cell surface of type R bacteria and caused them to clump together. The clumped bacteria were removed by a gentle centrifugation step. Only the bacteria that were not recognized by the antibody (namely, the type S bacteria) remained in the supernatant. The cells in the supernatant were plated on solid growth media. After overnight incubation, visible colonies may be observed.

type R bacteria, some of the bacteria were converted to type S. However, if no DNA extract was added, no type S bacterial colonies were observed on the petri plates.

A biochemist might point out that a DNA extract may not be 100% pure. In fact, any purified extract might contain small traces of some other substances. Therefore, one can argue that a small amount of contaminating material in the DNA extract might actually be the genetic material. The most likely contaminating substances in this case would be RNA or protein. To further verify that the DNA in the extract was indeed responsible for the transformation, Avery, MacLeod, and McCarty treated samples of the DNA extract with enzymes that digest DNA (called **DNase**), RNA **(RNase),** or protein **(protease)** (see Figure 9.2). When the DNA extracts were treated with RNase or protease, they still converted type R bacteria into type S. These results indicated that any remaining RNA or protein in the extract was not acting as the genetic material. However, when the extract was treated with DNase, it lost its ability to convert type R into type S bacteria. These results indicated that the degradation of the DNA in the extract by DNase prevented conversion of type R to type S. This interpretation is consistent with the hypothesis that DNA is the genetic material. A more elegant way of saying this is that the transforming principle is DNA.

EXPERIMENT 9A

Hershey and Chase Provided Evidence That DNA Is the Genetic Material of T2 Phage

A second experimental approach indicating that DNA is the genetic material came from the studies of Alfred Hershey and Martha Chase in 1952. Their research centered on the study of a virus known as T2. This virus infects *Escherichia coli* bacterial cells and is therefore known as a **bacteriophage,** or simply a **phage.** As shown in **Figure 9.3**, the external structure of the T2 phage, known as the capsid or phage coat, consists of a head, sheath, tail fibers, and base plate. Biochemically, the phage coat is composed entirely of protein, which includes several different polypeptides. DNA is found inside the head of the T2 capsid. From a molecular point of view, this virus is rather simple, because it is composed of only two types of macromolecules: DNA and proteins.

Although the viral genetic material contains the blueprint to make new viruses, a virus itself cannot synthesize new viruses. Instead, a virus must introduce its genetic material into the cytoplasm of a living cell. In the case of T2, this first involves the attachment of its tail fibers to the bacterial cell wall and the subsequent injection of its genetic material into the cytoplasm of the cell (**Figure 9.4**). The phage coat remains attached on the outside of the bacterium and does not enter the cell. After the entry of the

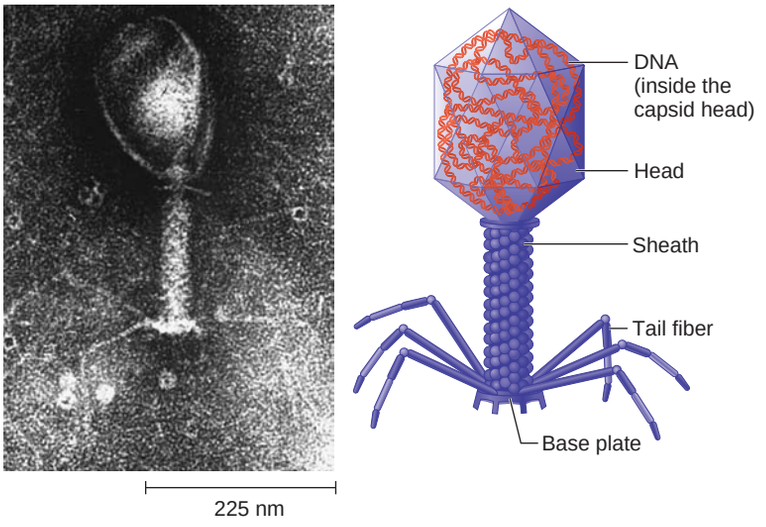
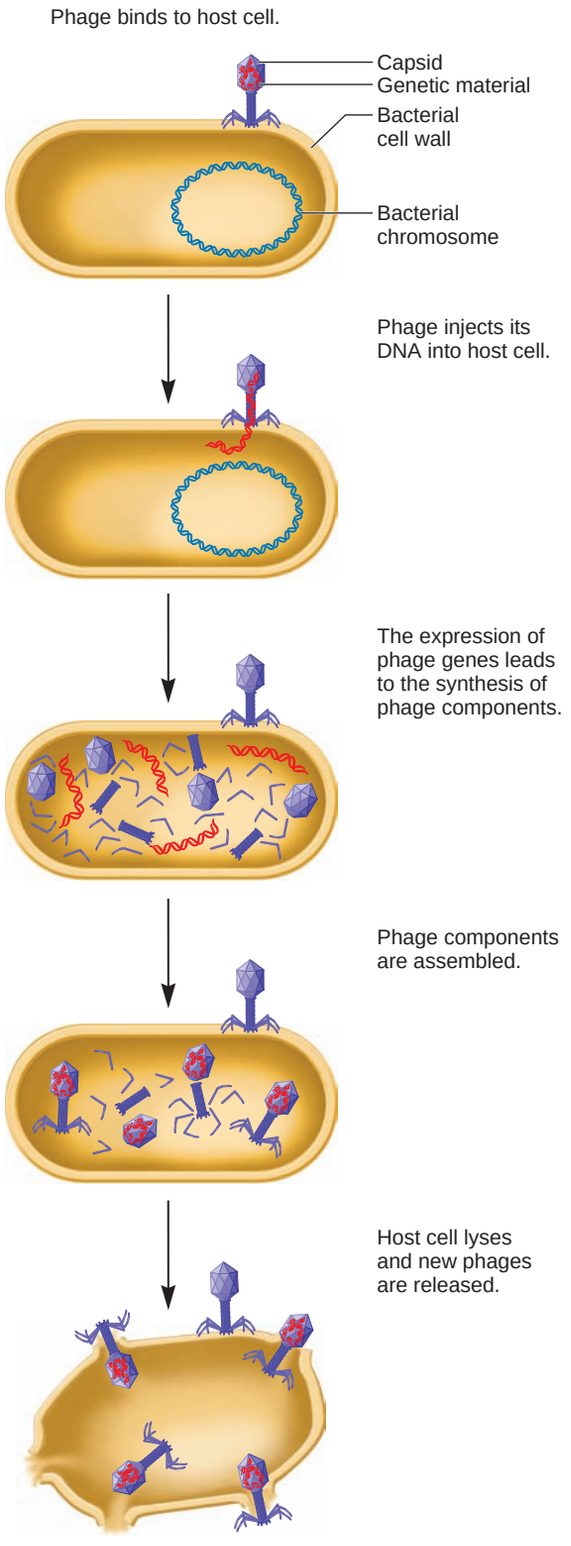


FIGURE 9.3 **Structure of the T2 bacteriophage.** The T2 bacteriophage is composed of a phage coat, or capsid, with genetic material inside the head of the capsid. The capsid is divided into regions called the head, sheath, tail fibers, and base plate. These components are composed of proteins. The genetic material is composed of DNA.

Genes → Traits The genetic material of a bacteriophage contains many genes, which provide the blueprint for making new viruses. When the bacteriophage injects its genetic material into a bacterium, these genes are activated and direct the host cell to make new bacteriophages, as described in Figure 9.4.



ONLINE ANIMATION

FIGURE 9.4 **Life cycle of the T2 bacteriophage.**

viral genetic material, the bacterial cytoplasm provides all of the machinery necessary to make viral proteins and DNA. The viral proteins and DNA assemble to make new viruses that are subsequently released from the cell by **lysis** (i.e., cell breakage).

To verify that DNA is the genetic material of T2, Hershey and Chase devised a method to separate the phage coat, which is attached to the outside of the bacterium, from the genetic material, which is injected into the cytoplasm. They were aware of microscopy experiments by Thomas Anderson showing that the T2 phage attaches itself to the outside of a bacterium by its tail fibers. Hershey and Chase reasoned that this is a fairly precarious attachment that could be disrupted by subjecting the bacteria to high shear forces, such as those produced in a kitchen blender. Their method was to expose bacteria to T2 phage, allowing sufficient time for the viruses to attach to bacteria and inject their genetic material. They then sheared the phage coats from the surface of the bacteria by a blender treatment. In this way, the phages' genetic material, which had been injected into the cytoplasm of the bacterial cells, could be separated from the phage coats that were sheared away.

Hershey and Chase used radioisotopes to distinguish proteins from DNA. Sulfur atoms are found in proteins but not in DNA, whereas phosphorus atoms are found in DNA but not in phage proteins. Therefore, ^{35}S (a radioisotope of sulfur) and ^{32}P (a radioisotope of phosphorus) were used to specifically label phage proteins and DNA, respectively. The researchers grew *E. coli* cells in media that contained ^{35}S or ^{32}P and then infected the *E. coli* cells with T2 phages. When new phages were produced, they were labeled with ^{35}S or ^{32}P. In the experiment described in **Figure 9.5**, they began with *E. coli* cells and two preparations of T2 phage that were obtained in this manner. One preparation was labeled with ^{35}S to label the phage proteins, and the other preparation was labeled with ^{32}P to label the phage DNA. In separate flasks, each type of phage was mixed with a new sample of *E. coli* cells. The phages were given sufficient time to inject their genetic material into the bacterial cells, and then the sample was subjected to shearing force using a blender. This treatment was expected to remove the phage coat from the surface of the bacterial cell. The sample was then subjected to centrifugation at a speed that would cause the heavier bacterial cells to form a pellet at the bottom of the tube, whereas the light phage coats would remain in the supernatant, the liquid found above the pellet. The amount of radioactivity in the supernatant (emitted from either ^{35}S or ^{32}P) was determined using a Geiger counter.

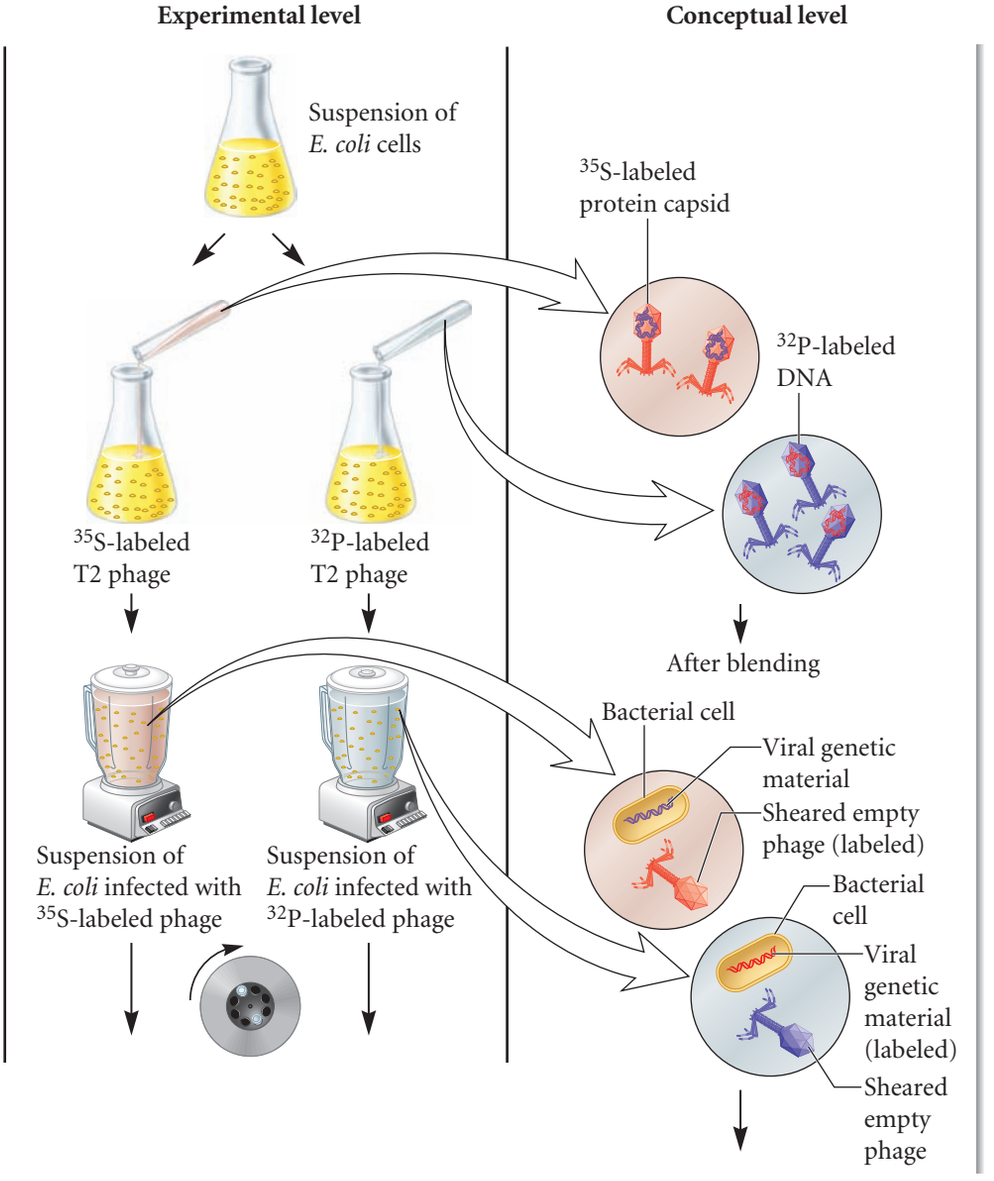
THE HYPOTHESIS

Only the genetic material of the phage is injected into the bacterium. Isotope labeling will reveal if it is DNA or protein.

TESTING THE HYPOTHESIS — FIGURE 9.5 — Evidence that DNA is the genetic material of T2 bacteriophage.

Starting materials: The starting materials were *E. coli* cells and two preparations of T2 phage. One phage preparation had phage proteins labeled with ^{35}S, and the other preparation had phage DNA labeled with ^{32}P.

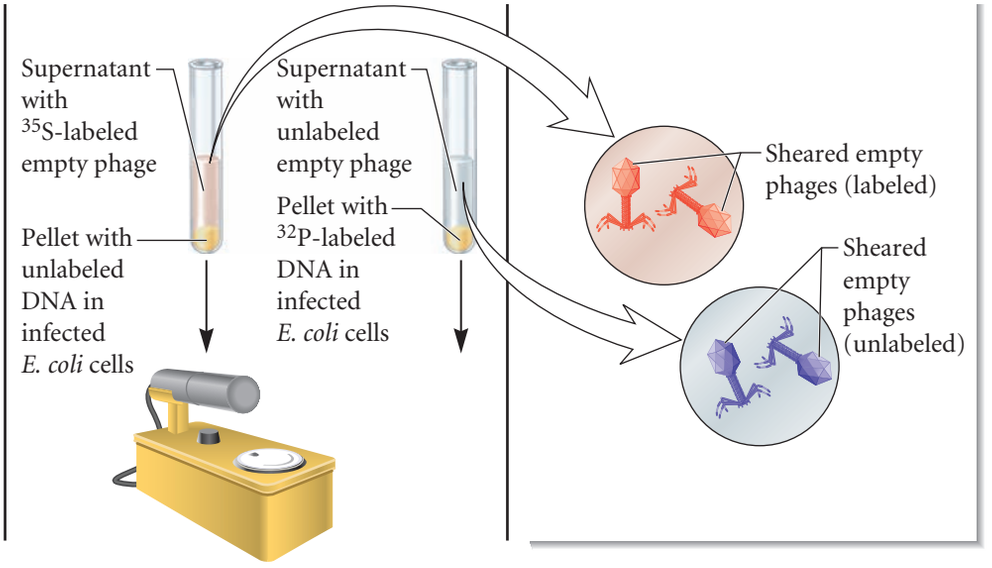
1. Grow bacterial cells. Divide into two flasks.
2. Into one flask, add ^{35}S-labeled phage; in the second flask, add ^{32}P-labeled phage.
3. Allow infection to occur.
4. Agitate solutions in blenders for different lengths of time to shear the empty phages off the bacterial cells.
5. Centrifuge at 10,000 rpm.



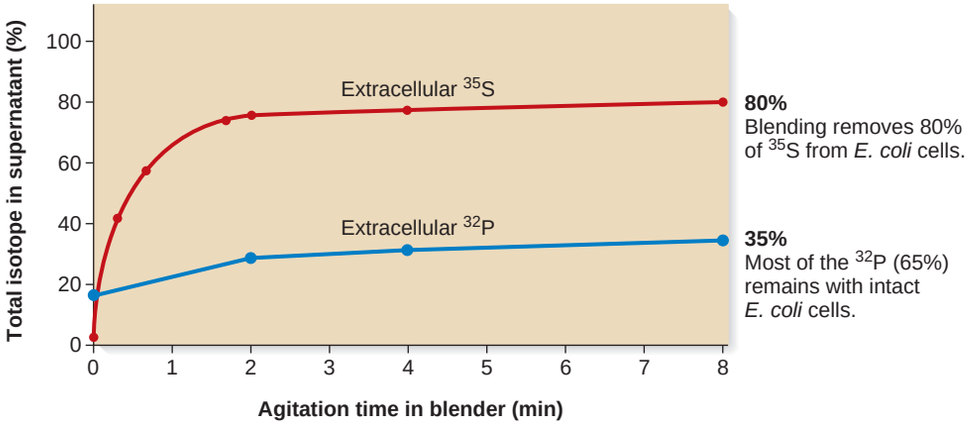
(continued)

6. The heavy bacterial cells sediment to the pellet, while the lighter phages remain in the supernatant. (See Appendix for explanation of centrifugation.)

7. Count the amount of radioisotope in the supernatant with a Geiger counter. Compare it with the starting amount.



THE DATA



Data from A. D. Hershey and Martha Chase (1952) Independent Functions of Viral Protein and Nucleic Acid in Growth of Bacteriophage. *Journal of General Physiology 36*, 39–56.

INTERPRETING THE DATA

As seen in the data, most of the ^{35}S isotope was found in the supernatant. Because the shearing force was expected to remove the phage coat, this result indicates that the empty phages contain primarily protein. By comparison, only about 35% of the ^{32}P was found in the supernatant following shearing. Therefore, most of the DNA was located within the bacterial cells in the pellet. These results are consistent with the idea that the DNA is injected into the bacterial cytoplasm during infection, which would be the expected result if DNA is the genetic material.

By themselves, the results described in Figure 9.5 were not conclusive evidence that DNA is the genetic material. For example, you may have noticed that less than 100% of the phage protein was found in the supernatant. Therefore, some of the phage protein could have been introduced into the bacterial cells (and could function as the genetic material). Nevertheless, the results of Hershey and Chase were consistent with the conclusion that the genetic material is DNA rather than protein. Overall, their studies of the T2 phage were quite influential in convincing the scientific community that DNA is the genetic material.

A self-help quiz involving this experiment can be found at www.mhhe.com/brookergenetics4e.

RNA Functions as the Genetic Material in Some Viruses

We now know that bacteria, archaea, protists, fungi, plants, and animals all use DNA as their genetic material. As mentioned, viruses also have their own genetic material. Hershey and Chase concluded from their experiments that this genetic material is DNA. In the case of T2 bacteriophage, that is the correct conclusion. However, many viruses use RNA, rather than DNA, as their genetic material. In 1956, Alfred Gierer and Gerhard Schramm

TABLE 9.1
Examples of DNA- and RNA-Containing Viruses

Virus	Host	Nucleic Acid
Tomato bushy stunt virus	Tomato	RNA
Tobacco mosaic virus	Tobacco	RNA
Influenza virus	Humans	RNA
HIV	Humans	RNA
f2	*E. coli*	RNA
Qβ	*E. coli*	RNA
Cauliflower mosaic virus	Cauliflower	DNA
Herpesvirus	Humans	DNA
SV40	Primates	DNA
Epstein-Barr virus	Humans	DNA
T2	*E. coli*	DNA
M13	*E. coli*	DNA

isolated RNA from the tobacco mosaic virus (TMV), which infects plant cells. When this purified RNA was applied to plant tissue, the plants developed the same types of lesions that occurred when they were exposed to intact tobacco mosaic viruses. Gierer and Schramm correctly concluded that the viral genome of TMV is composed of RNA. Since that time, many other viruses have been found to contain RNA as their genetic material. **Table 9.1** compares the genetic compositions of several different types of viruses.

9.2 NUCLEIC ACID STRUCTURE

DNA and its molecular cousin, RNA, are known as **nucleic acids.** This term is derived from the discovery of DNA by Friedrich Miescher in 1869. He identified a novel phosphorus-containing substance from the nuclei of white blood cells found in waste surgical bandages. He named this substance nuclein. As the structure of DNA and RNA became better understood, it was found they are acidic molecules, which means they release hydrogen ions (H^+) in solution and have a net negative charge at neutral pH. Thus, the name nucleic acid was coined.

Geneticists, biochemists, and biophysicists have been interested in the molecular structure of nucleic acids for several decades. Both DNA and RNA are macromolecules composed of smaller building blocks. To fully appreciate their structures, we need to consider four levels of complexity (**Figure 9.6**):

1. **Nucleotides** form the repeating structural unit of nucleic acids.
2. Nucleotides are linked together in a linear manner to form a **strand** of DNA or RNA.
3. Two strands of DNA (and sometimes RNA) interact with each other to form a **double helix.**
4. The three-dimensional structure of DNA results from the folding and bending of the double helix. Within living cells, DNA is associated with a wide variety of proteins that influence its structure. Chapter 10 examines the roles of these proteins in creating the three-dimensional structure of DNA found within chromosomes.

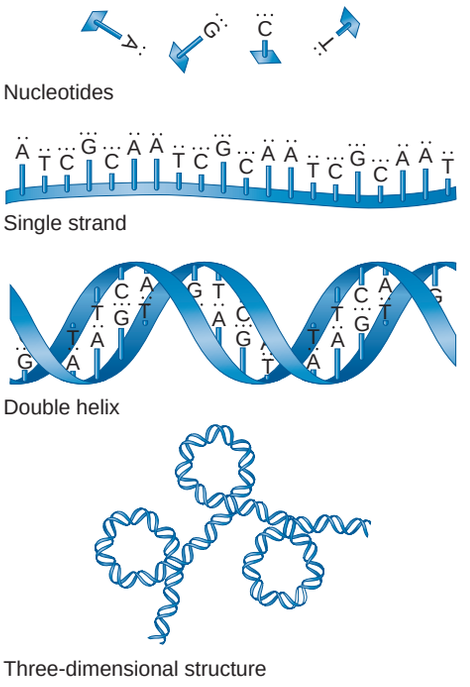
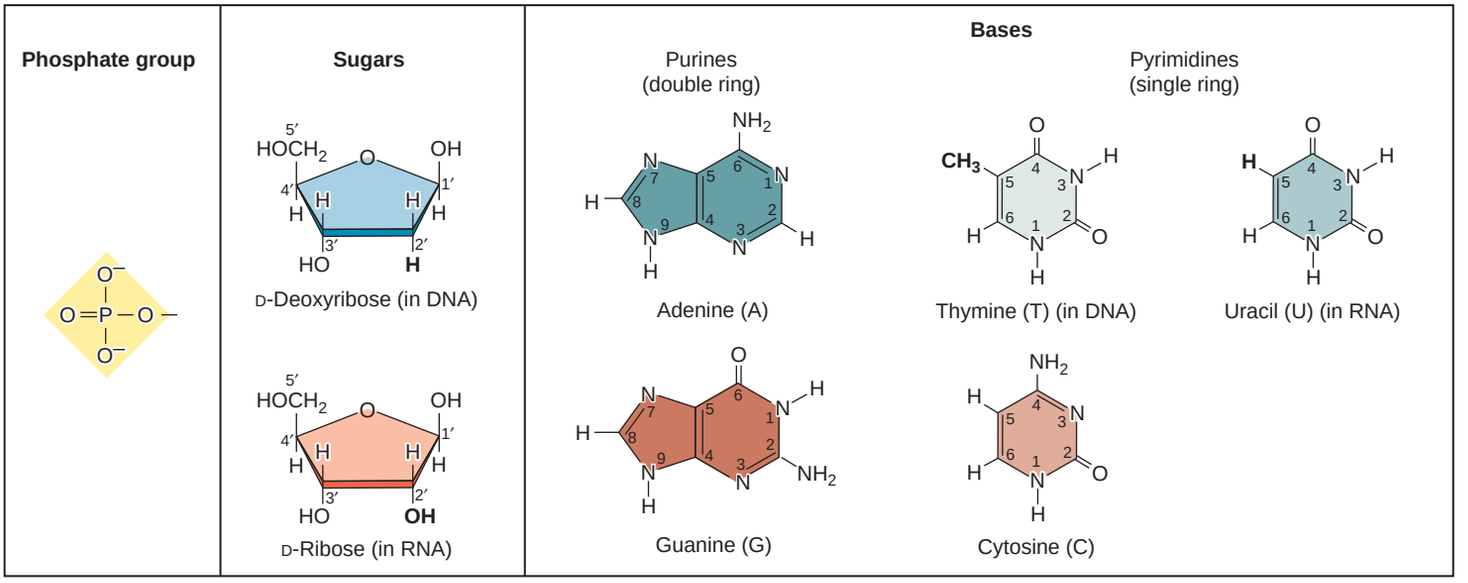


FIGURE 9.6 **Levels of nucleic acid structure.**

In this section, we will first examine the structure of individual nucleotides and then progress through the structural features of DNA and RNA. Along the way, we will also review some of the pivotal experiments that led to the discovery of the double helix.

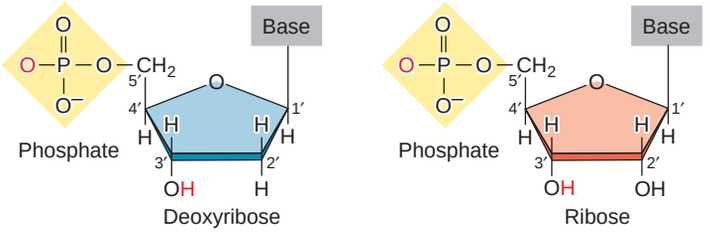
Nucleotides Are the Building Blocks of Nucleic Acids

The nucleotide is the repeating structural unit of DNA and RNA. A nucleotide has three components: at least one phosphate group, a pentose sugar, and a nitrogenous base. As shown in **Figure 9.7**, nucleotides can vary with regard to the sugar and the nitrogenous base. The two types of sugars are **deoxyribose** and **ribose,** which are found in DNA and RNA, respectively. The five different bases are subdivided into two categories: the **purines** and the **pyrimidines.** The purine bases, **adenine (A)** and **guanine (G),** contain a double-ring structure; the pyrimidine bases, **thymine (T), cytosine (C),** and **uracil (U),** contain a single-ring structure. The sugar in DNA is always deoxyribose. In RNA, the sugar is ribose. Also, the base thymine is not found in RNA. Rather, uracil is found in RNA instead of thymine. Adenine, guanine, and cytosine occur in both DNA and RNA. As noted in Figure 9.7, the bases and sugars have a standard numbering system. The nitrogen and carbon atoms found in the ring structure of the bases are given numbers 1 through 9 for the purines and 1 through 6 for the pyrimidines. In comparison, the five carbons found in the sugars are designated with primes, such as 1′, to distinguish them from the numbers found in the bases.



INTERACTIVE EXERCISE

FIGURE 9.7 The components of nucleotides. The three building blocks of a nucleotide are one or more phosphate groups, a sugar, and a base. The bases are categorized as purines (adenine and guanine) and pyrimidines (thymine, cytosine, and uracil).



(a) Repeating unit of deoxyribonucleic acid (DNA)

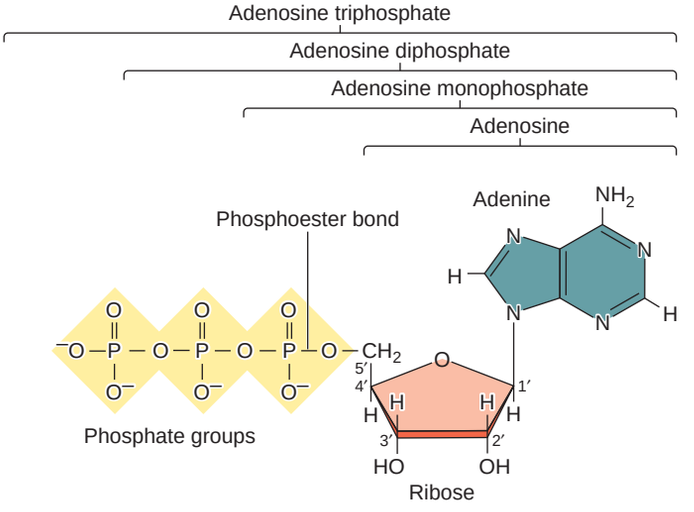
(b) Repeating unit of ribonucleic acid (RNA)

INTERACTIVE EXERCISE

FIGURE 9.8 The structure of nucleotides found in (a) DNA and (b) RNA. DNA contains deoxyribose as its sugar and the bases A, T, G, and C. RNA contains ribose as its sugar and the bases A, U, G, and C. In a DNA or RNA strand, the oxygen on the 3′ carbon is linked to the phosphorus atom of phosphate in the adjacent nucleotide. The two atoms (O and H) shown in red would be found within individual nucleotides but not when nucleotides are joined together to make strands of DNA and RNA.

Figure 9.8 shows the repeating unit of nucleotides found in DNA and RNA. The locations of the attachment sites of the base and phosphate to the sugar molecule are important to the nucleotide's function. In the sugar ring, carbon atoms are numbered in a clockwise direction, beginning with a carbon atom adjacent to the ring oxygen atom. The fifth carbon is outside the ring structure. In a single nucleotide, the base is always attached to the 1′ carbon atom, and one or more phosphate groups are attached at the 5′ position. As discussed later, the –OH group attached to the 3′ carbon is important in allowing nucleotides to form covalent linkages with each other.

The terminology used to describe nucleic acid units is based on three structural features: the type of base, the type of sugar, and the number of phosphate groups. When a base is attached to only a sugar, we call this pair a **nucleoside.** If adenine is attached to ribose, this nucleoside is called adenosine (**Figure 9.9**). Nucleosides containing guanine, thymine, cytosine, or uracil are called guanosine, thymidine, cytidine, and uridine, respectively. When only the bases are attached to deoxyribose, they are called deoxyadenosine,



INTERACTIVE EXERCISE

FIGURE 9.9 A comparison between the structures of an adenine-containing nucleoside and nucleotides.

deoxyguanosine, deoxythymidine, and deoxycytidine. The covalent attachment of one or more phosphate molecules to a nucleoside creates a nucleotide. If a nucleotide contains adenine, ribose, and one phosphate, it is adenosine monophosphate, abbreviated AMP. If a nucleotide contains adenine, ribose, and three phosphate groups, it is called adenosine triphosphate, or ATP. If it contains guanine, ribose, and three phosphate groups, it is guanosine triphosphate, or GTP. A nucleotide can also be composed of adenine, deoxyribose, and three phosphate groups. This nucleotide is deoxyadenosine triphosphate (dATP).

Nucleotides Are Linked Together to Form a Strand

A strand of DNA or RNA has nucleotides that are covalently attached to each other in a linear fashion. **Figure 9.10** depicts a short strand of DNA with four nucleotides. A few structural features are worth noting. First, the linkage involves an ester bond between a phosphate group on one nucleotide and the sugar molecule on the adjacent nucleotide. Another way of viewing this linkage is to notice that a phosphate group connects two sugar molecules. For this reason, the linkage in DNA or RNA strands is called a **phosphodiester linkage.** The phosphates and sugar molecules form the **backbone** of a DNA or RNA strand. The bases project from the backbone. The backbone is negatively charged due to a negative charge on each phosphate.

A second important structural feature is the orientation of the nucleotides. As mentioned, the carbon atoms in a sugar molecule are numbered in a particular way. A phosphodiester linkage involves a phosphate attachment to the 5′ carbon in one nucleotide and to the 3′ carbon in the other. In a strand, all sugar molecules are oriented in the same direction. As shown in Figure 9.10, the 5′ carbons in every sugar molecule are above the 3' carbons. Therefore, a strand has a **directionality** based on the orientation of the sugar molecules within that strand. In Figure 9.10, the direction of the strand is 5′ to 3′ when going from top to bottom.

A critical aspect regarding DNA and RNA structure is that a strand contains a specific sequence of bases. In Figure 9.10, the sequence of bases is thymine–adenine–cytosine–guanine, abbreviated TACG. Furthermore, to show the directionality, the strand should be abbreviated 5′–TACG–3′. The nucleotides within a strand are covalently attached to each other, so the sequence of bases cannot shuffle around and become rearranged. Therefore, the sequence of bases in a DNA strand remains the same over time, except in rare cases when mutations occur. As we will see throughout this textbook, the sequence of bases within DNA and RNA is the defining feature that allows them to carry information.

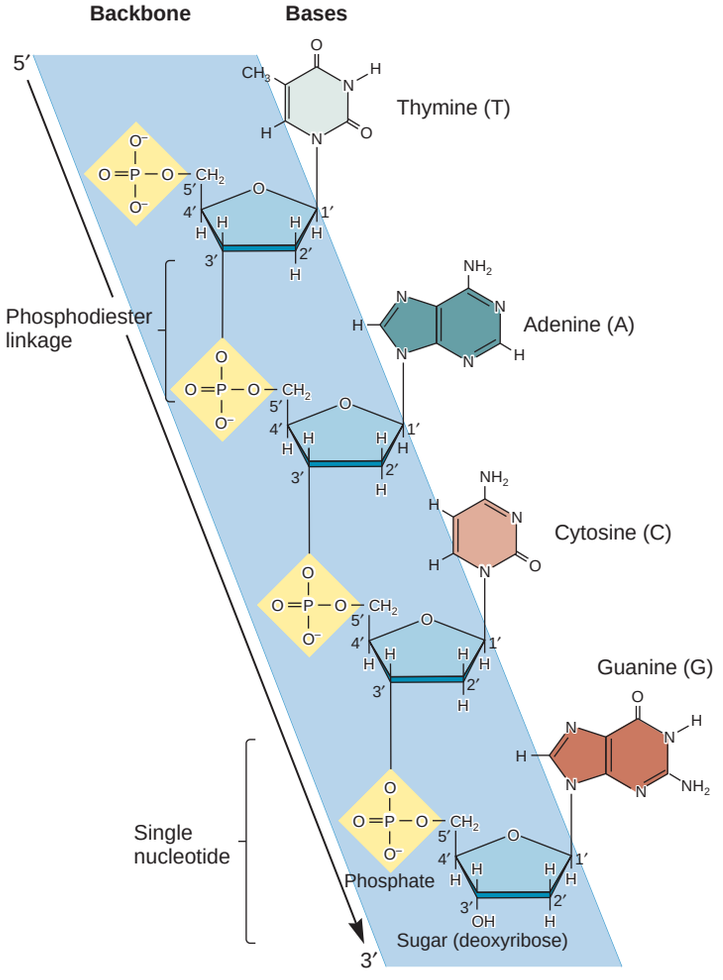
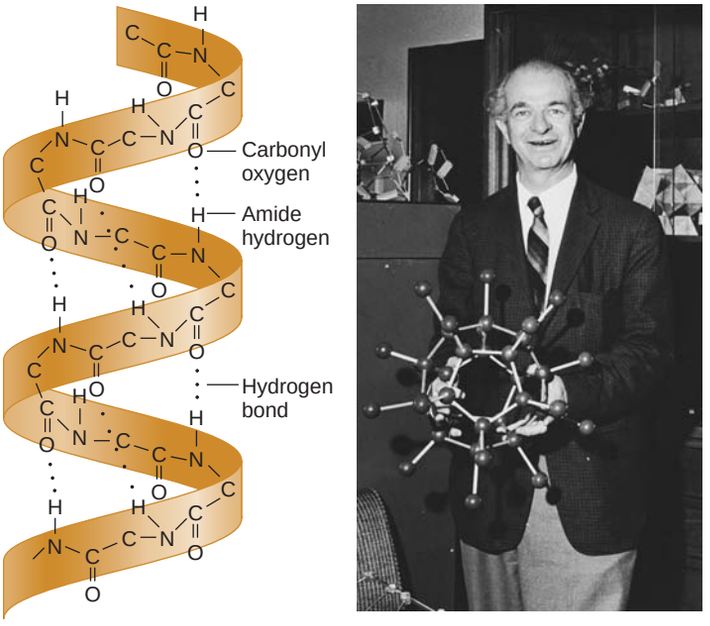


FIGURE 9.10 A short strand of DNA containing four nucleotides. Nucleotides are covalently linked together to form a strand of DNA.

A Few Key Events Led to the Discovery of the Double-Helix Structure

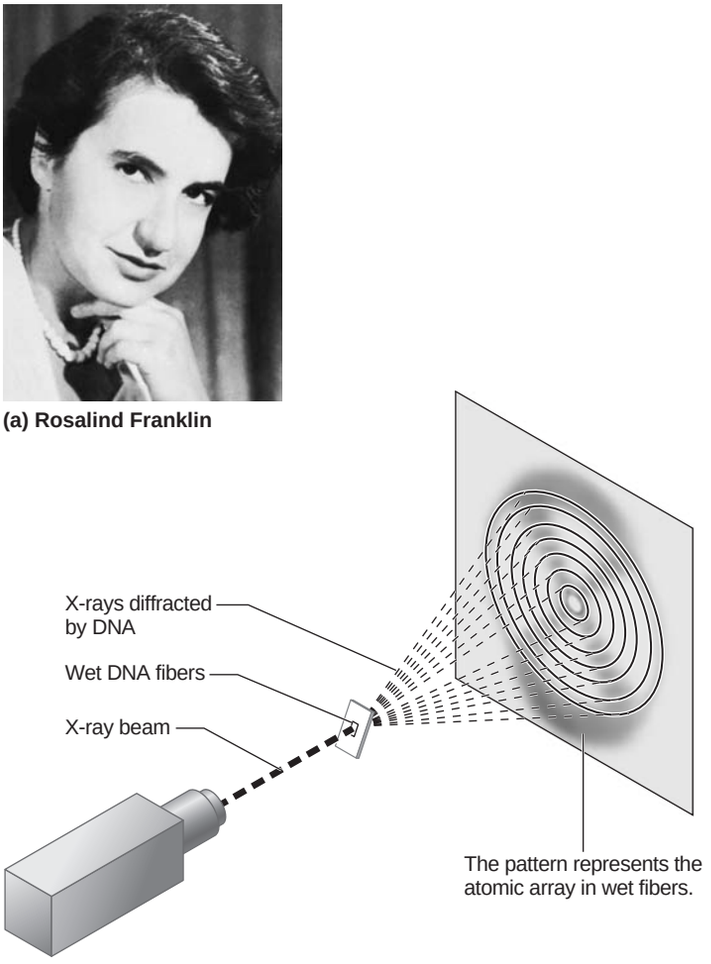
A major discovery in molecular genetics was made in 1953 by James Watson and Francis Crick. At that time, DNA was already known to be composed of nucleotides. However, it was not understood how the nucleotides are bonded together to form the structure of DNA. Watson and Crick committed themselves to determine the structure of DNA because they felt this knowledge was needed to understand the functioning of genes. Other researchers, such as Rosalind Franklin and Maurice Wilkins, shared this view. Before we examine the characteristics of the double helix, let's consider the events that provided the scientific framework for Watson and Crick's breakthrough.

In the early 1950s, Linus Pauling proposed that regions of proteins can fold into a secondary structure known as an α helix (**Figure 9.11a**). To elucidate this structure, Pauling built large models by linking together simple ball-and-stick units (**Figure 9.11b**). By carefully scaling the objects in his models, he could visualize if atoms fit together properly in a complicated three-dimensional structure. Is this approach still used today? The answer is "Yes," except that today researchers construct their three-dimensional models on computers. As we will see, Watson and Crick also used a ball-and-stick approach to solve the structure of the DNA double helix. Interestingly, they were well aware that Pauling might figure out the structure of DNA before they did. This provided a stimulating rivalry between the researchers.



(a) An α helix in a protein (b) Linus Pauling

FIGURE 9.11 Linus Pauling and the α-helix protein structure. **(a)** An α-helix is a secondary structure found in proteins. This structure emphasizes the polypeptide backbone (shown as a tan ribbon), which is composed of amino acids linked together in a linear fashion. Hydrogen bonding between hydrogen and oxygen atoms stabilizes the helical conformation. **(b)** Linus Pauling with a ball-and-stick model.



(a) Rosalind Franklin

(b) X-ray diffraction of wet DNA fibers

FIGURE 9.12 X-ray diffraction of DNA.

A second important development that led to the elucidation of the double helix was X-ray diffraction data. When a purified substance, such as DNA, is subjected to X-rays, it produces a well-defined diffraction pattern if the molecule is organized into a regular structural pattern. An interpretation of the diffraction pattern (using mathematical theory) can ultimately provide information concerning the structure of the molecule. Rosalind Franklin (**Figure 9.12a**), working in the same laboratory as Maurice Wilkins, used X-ray diffraction to study wet DNA fibers. Franklin made marked advances in X-ray diffraction techniques while working with DNA. She adjusted her equipment to produce an extremely fine beam of X-rays. She extracted finer DNA fibers than ever before and arranged them in parallel bundles. Franklin also studied the fibers' reactions to humid conditions. The diffraction pattern of Franklin's DNA fibers is shown in **Figure 9.12b**. This pattern suggested several structural features of DNA. First, it was consistent with a helical structure. Second, the diameter of the helical structure was too wide to be only a single-stranded helix. Finally, the diffraction pattern indicated that the helix contains about 10 base pairs (bp) per complete turn. These observations were instrumental in solving the structure of DNA.

EXPERIMENT 9B

Chargaff Found That DNA Has a Biochemical Composition in Which the Amount of A Equals T and the Amount of G Equals C

Another piece of information that led to the discovery of the double-helix structure came from the studies of Erwin Chargaff. In the 1940s and 1950s, he pioneered many of the biochemical techniques for the isolation, purification, and measurement of nucleic acids from living cells. This was not a trivial undertaking, because the biochemical composition of living cells is complex. At the time of Chargaff's work, researchers already knew that the building blocks of DNA are nucleotides containing the bases adenine, thymine, guanine, or cytosine. Chargaff analyzed the base composition of DNA, which was isolated from many different species. He expected that the results might provide important clues concerning the structure of DNA.

The experimental protocol of Chargaff is described in **Figure 9.13**. He began with various types of cells as starting material. The chromosomes were extracted from cells and then treated with protease to separate the DNA from chromosomal proteins. The DNA was then subjected to a strong acid treatment that cleaved the bonds between the sugars and bases. Therefore, the strong acid treatment released the individual bases from the DNA strands. This mixture of bases was then subjected to paper chromatography to separate the four types. The amounts of the four bases were determined spectroscopically.

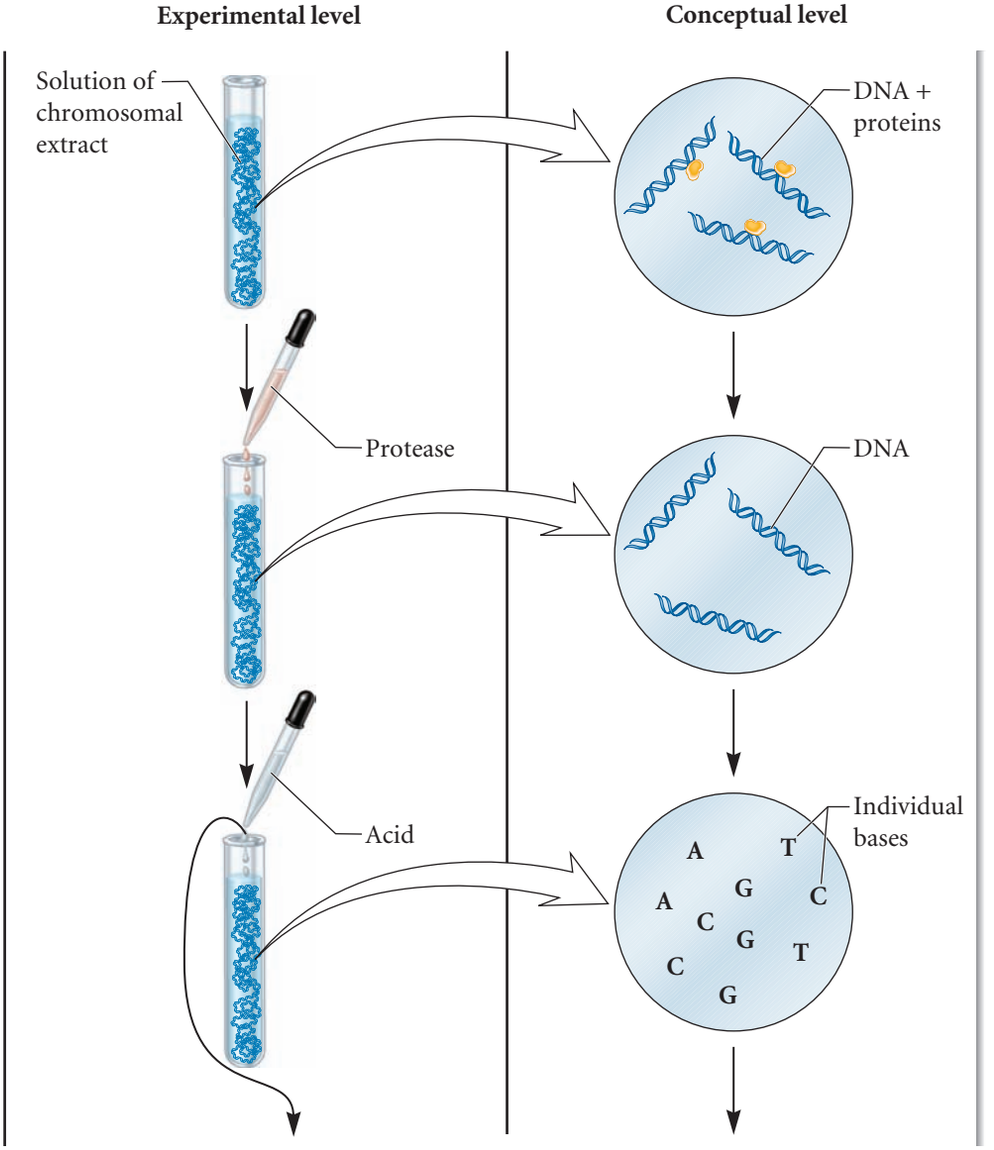
THE GOAL

An analysis of the base composition of DNA in different organisms may reveal important features about the structure of DNA.

ACHIEVING THE GOAL — FIGURE 9.13 An analysis of base composition among different DNA samples.

Starting material: The following types of cells were obtained: *Escherichia coli*, *Streptococcus pneumoniae*, yeast, turtle red blood cells, salmon sperm cells, chicken red blood cells, and human liver cells.

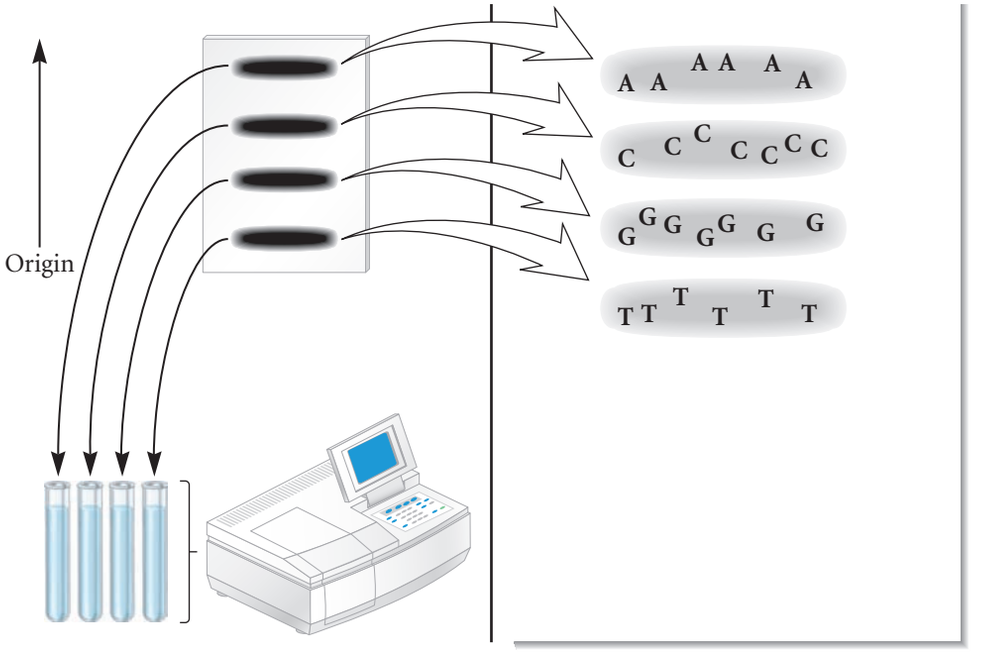
1. For each type of cell, extract the chromosomal material. This can be done in a variety of ways, including the use of high salt, detergent, or mild alkali treatment. Note: The chromosomes contain both DNA and protein.
2. Remove the protein. This can be done in several ways, including treament with protease.
3. Hydrolyze the DNA to release the bases from the DNA strands. A common way to do this is by strong acid treatment.



4. Separate the bases by chromatography. Paper chromatography provides an easy way to separate the four types of bases. (The technique of chromatography is described in the Appendix.)

5. Extract bands from paper into solutions and determine the amounts of each base by spectroscopy. Each base will absorb light at a particular wavelength. By examining the absorption profile of a sample of base, it is then possible to calculate the amount of the base. (Spectroscopy is described in the Appendix.)

6. Compare the base content in the DNA from different organisms.



THE DATA

Base Content in the DNA from a Variety of Organisms*

	Percentage of Bases (based on molarity)			
Organism	*Adenine*	*Thymine*	*Guanine*	*Cytosine*
Escherichia coli	26.0	23.9	24.9	25.2
Streptococcus pneumoniae	29.8	31.6	20.5	18.0
Yeast	31.7	32.6	18.3	17.4
Turtle red blood cells	28.7	27.9	22.0	21.3
Salmon sperm	29.7	29.1	20.8	20.4
Chicken red blood cells	28.0	28.4	22.0	21.6
Human liver cells	30.3	30.3	19.5	19.9

*When the base compositions from different tissues within the same species were measured, similar results were obtained. These data were compiled from several sources. See E. Chargaff and J. Davidson, Eds. (1995) *The Nucleic Acids*. Academic Press, New York.

INTERPRETING THE DATA

The data shown in Figure 9.13 are only a small sampling of Chargaff's results. During the late 1940s and early 1950s, Chargaff published many papers concerned with the chemical composition of DNA from biological sources. Hundreds of measurements were made. The compelling observation was that the amount of adenine was similar to that of thymine, and the amount of guanine was similar to cytosine. The idea that the amount of A in DNA equals the amount of T, and the amount of G equals C, is known as **Chargaff's rule.**

These results were not sufficient to propose a model for the structure of DNA. However, they provided the important clue that DNA is structured so that each molecule of adenine interacts with thymine, and each molecule of guanine interacts with cytosine. A DNA structure in which A binds to T, and G to C, would explain the equal amounts of A and T, and G and C observed in Chargaff's experiments. As we will see, this observation became crucial evidence that Watson and Crick used to elucidate the structure of the double helix.

A self-help quiz involving this experiment can be found at www.mhhe.com/brookergenetics4e.

Watson and Crick Deduced the Double-Helical Structure of DNA

Thus far, we have examined key pieces of information used to determine the structure of DNA. In particular, the X-ray diffraction work of Franklin suggested a helical structure composed of two or more strands with 10 bases per turn. In addition, the work of Chargaff indicated that the amount of A equals T, and the amount of G equals C. Furthermore, Watson and Crick were familiar with Pauling's success in using ball-and-stick models to deduce the secondary structure of proteins. With these key observations, they set out to solve the structure of DNA.

Watson and Crick assumed DNA is composed of nucleotides that are linked together in a linear fashion. They also assumed the chemical linkage between two nucleotides is always the same. With these ideas in mind, they tried to build ball-and-

stick models that incorporated the known experimental observations. Because the diffraction pattern suggested the helix must have two (or more) strands, a critical question was, How could two strands interact? As discussed in his book, *The Double Helix*, James Watson noted that in an early attempt at model building, they considered the possibility that the negatively charged phosphate groups, together with magnesium ions, were promoting an interaction between the backbones of DNA strands (**Figure 9.14**).

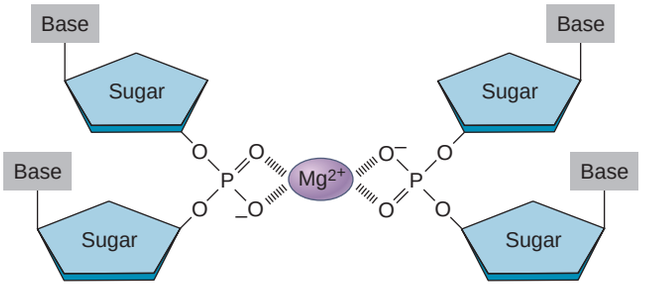
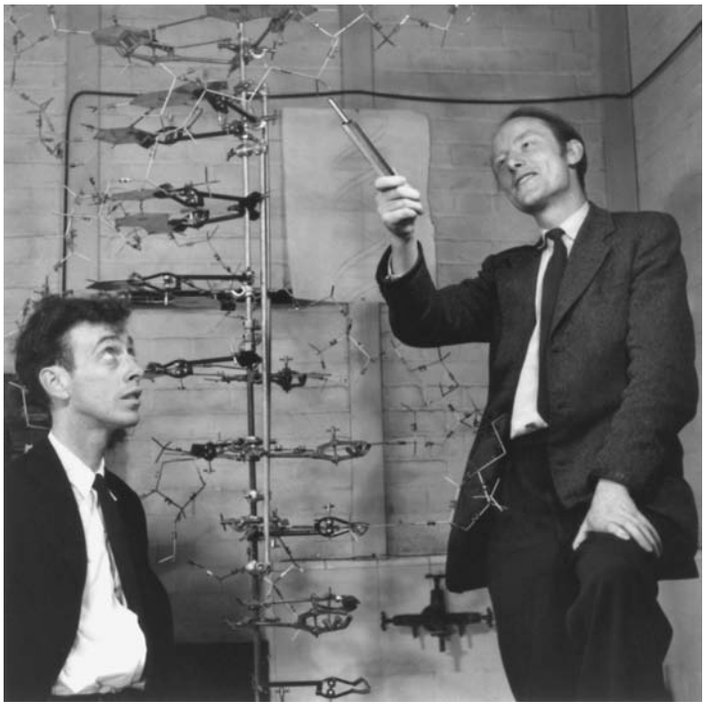


FIGURE 9.14 An incorrect hypothesis for the structure of the DNA double helix. This illustration shows an early hypothesis of Watson and Crick's, suggesting that two DNA strands interact by a cross-link between the negatively charged phosphate groups in the backbone and divalent Mg^{2+} cations.

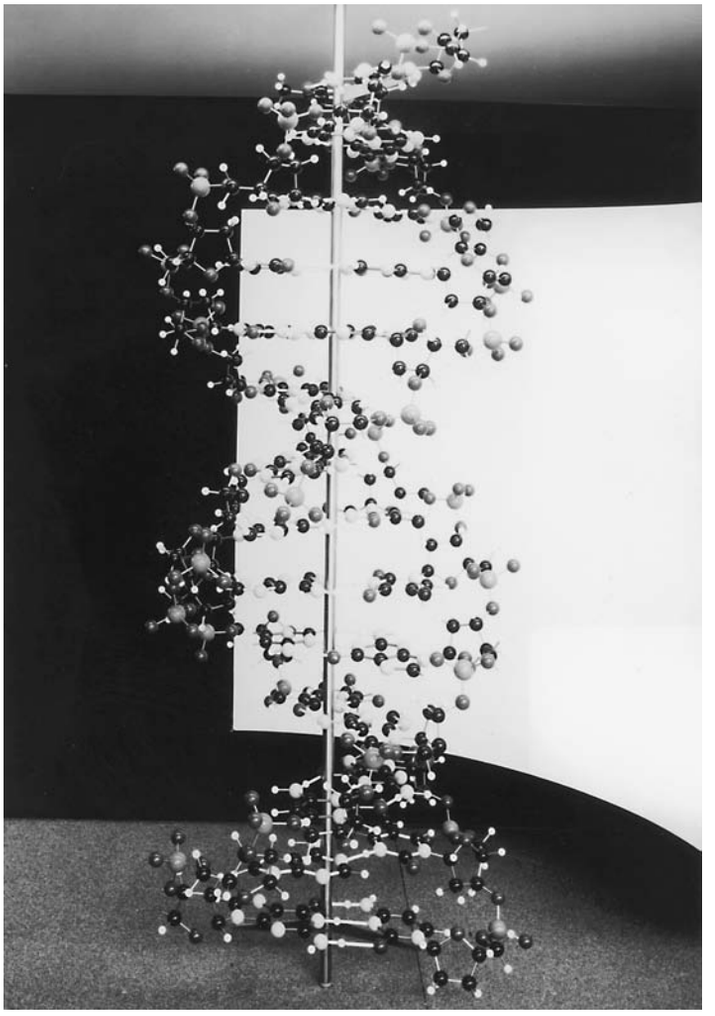


(a) Watson and Crick

FIGURE 9.15 Watson and Crick and their model of the DNA double helix. **(a)** James Watson is shown here on the left and Francis Crick on the right. **(b)** The molecular model they originally proposed for the double helix. Each strand contains a sugar-phosphate backbone. In opposite strands, A hydrogen bonds to T, and G hydrogen bonds with C.

However, more detailed diffraction data were not consistent with this model.

Because the magnesium hypothesis for DNA structure appeared to be incorrect, it was back to the drawing board (or back to the ball-and-stick units) for Watson and Crick. During this time, Rosalind Franklin had produced even clearer X-ray diffraction patterns, which provided greater detail concerning the relative locations of the bases and backbone of DNA. This major breakthrough suggested a two-strand interaction that was helical. In their model building, Watson and Crick's emphasis shifted to models containing the two backbones on the outside of the model, with the bases projecting toward each other. At first, a structure was considered in which the bases form hydrogen bonds with the identical base in the opposite strand (A to A, T to T, G to G, and C to C). However, the model building revealed that the bases could not fit together this way. The final hurdle was overcome when it was realized that the hydrogen bonding of adenine to thymine was structurally similar to that of guanine to cytosine. With an interaction between A and T and between G and C, the ball-and-stick models showed that the two strands would fit together properly. This ball-and-stick model, shown in **Figure 9.15**, was consistent with all of the known data regarding DNA structure.



(b) Original model of the DNA double helix

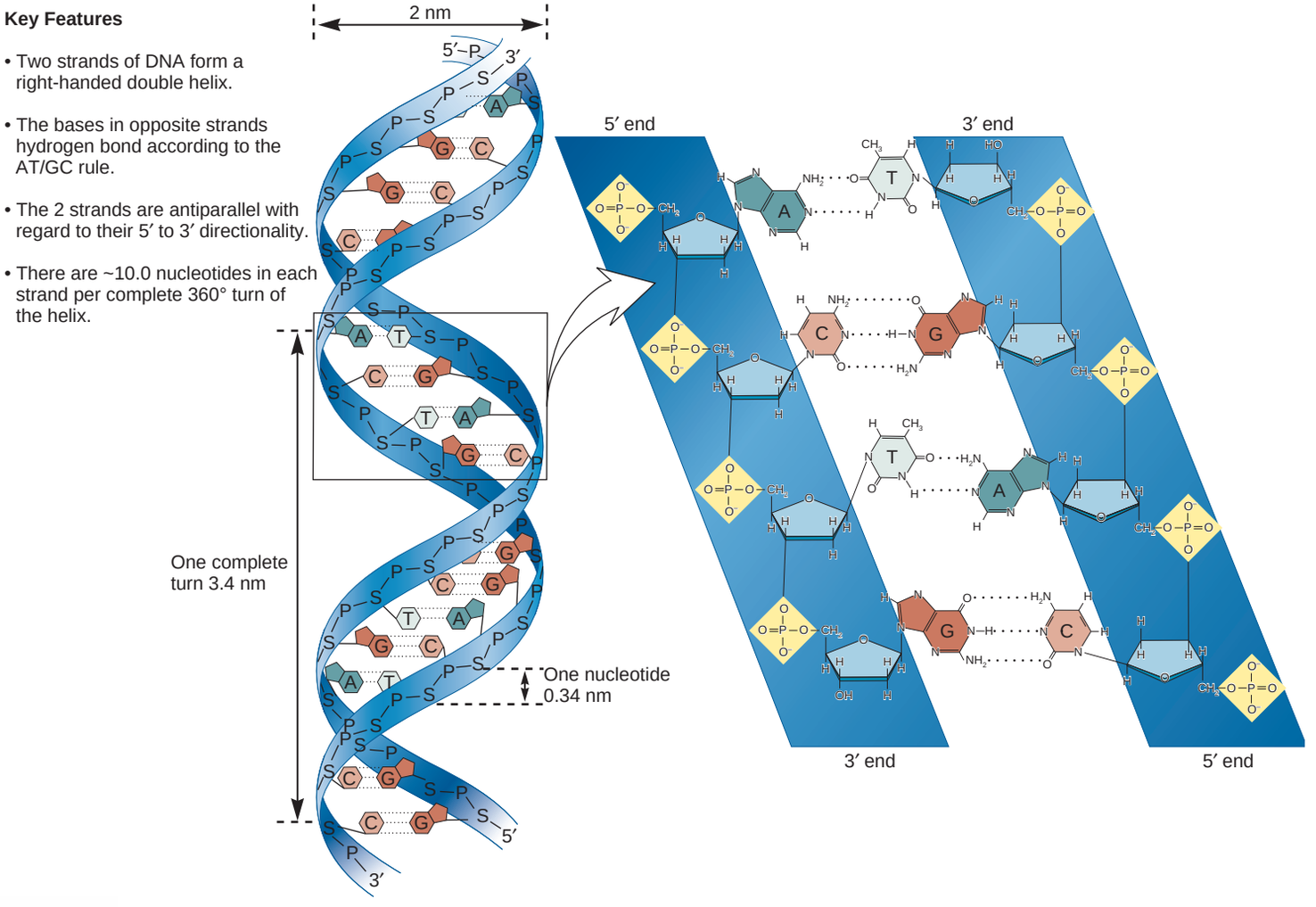
For their work, Watson, Crick, and Maurice Wilkins were awarded the 1962 Nobel Prize in physiology or medicine. The contribution of Rosalind Franklin to the discovery of the double helix was also critical and has been acknowledged in several books and articles. Franklin was independently trying to solve the structure of DNA. However, Wilkins, who worked in the same laboratory, shared Franklin's X-ray data with Watson and Crick, presumably without her knowledge. This provided important information that helped them solve the structure of DNA, which was published in the journal *Nature* in April 1953. Though she was not given credit in the original publication of the double-helix structure, Franklin's key contribution became known in later years. Unfortunately, however, Rosalind Franklin died in 1958, and the Nobel Prize is not awarded posthumously.

The Molecular Structure of the DNA Double Helix Has Several Key Features

The general structural features of the double helix are shown in **Figure 9.16**. In a DNA double helix, two DNA strands are twisted together around a common axis to form a structure that resembles a spiral staircase. This double-stranded structure is stabilized by **base pairs (bp)**—pairs of bases in opposite strands that are hydrogen bonded to each other. Counting the bases, if you move past 10 bp, you have gone 360° around the backbone. The linear distance of a complete turn is 3.4 nm; each base pair traverses 0.34 nm.

A distinguishing feature of the hydrogen bonding between base pairs is its specificity. An adenine base in one strand hydrogen bonds with a thymine base in the opposite strand, or a guanine base hydrogen bonds with a cytosine. This **AT/GC rule** explained the earlier data of Chargaff showing that the DNA from many organisms contains equal amounts of A and T, and equal amounts of G and C (see Figure 9.13). The AT/GC rule indicates that purines (A and G) always bond with pyrimidines (T and C). This keeps the width of the double helix relatively constant. As noted in Figure 9.16, three hydrogen bonds occur between G and C but only two between A and T. For this reason, DNA sequences that have a high proportion of G and C tend to form more stable double-stranded structures.

The AT/GC rule implies that we can predict the sequence in one DNA strand if the sequence in the opposite strand is known.



ONLINE ANIMATION

FIGURE 9.16 **Key features of the structure of the double helix.** Note: In the inset, the planes of the bases and sugars are shown parallel to each other in order to depict the hydrogen bonding between the bases. In an actual DNA molecule, the bases would be rotated about 90° so that the planes of the bases would be facing each other.

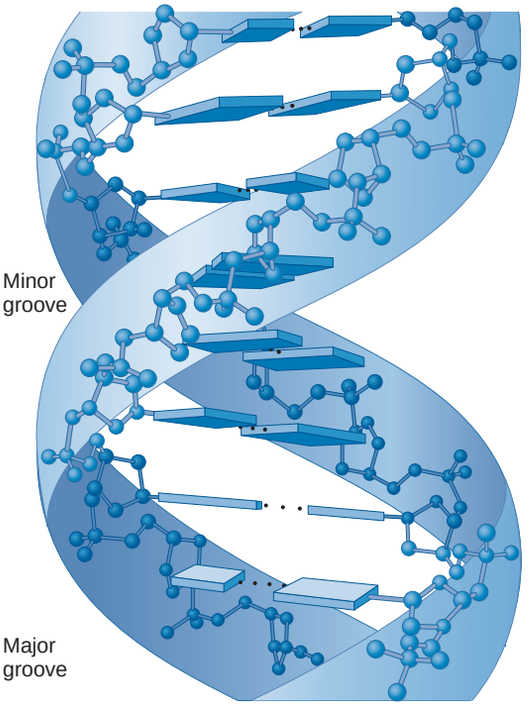
For example, let's consider a DNA strand with the sequence of 5′–ATGGCGGATTT–3′. The opposite strand would have to be 3′–TACCGCCTAAA–5′. In genetic terms, we would say that these two sequences are **complementary** to each other or that the two sequences exhibit complementarity. In addition, you may have noticed that the sequences are labeled with 5′ and 3′ ends. These numbers designate the direction of the DNA backbones. The direction of DNA strands is depicted in the inset to Figure 9.16. When going from the top of this figure to the bottom, one strand is running in the 5′ to 3′ direction, and the other strand is 3′ to 5′. This opposite orientation of the two DNA strands is referred to as an **antiparallel** arrangement. An antiparallel structure was initially proposed in the models of Watson and Crick.

Figure 9.17a is a schematic model that emphasizes certain molecular features of DNA structure. The bases in this model are depicted as flat rectangular structures that hydrogen bond in pairs. (The hydrogen bonds are the dotted lines.) Although the bases are not actually rectangular, they do form flattened planar structures. Within DNA, the bases are oriented so that the flattened regions are facing each other, an arrangement referred to as base stacking. In other words, if you think of the bases as flat plates, these plates are stacked on top of each other in the double-stranded DNA structure. Along with hydrogen bonding, base stacking is a structural feature that stabilizes the double helix by excluding water molecules. The helical structure of the DNA backbone depends on the hydrogen bonding between base pairs and also on base stacking.

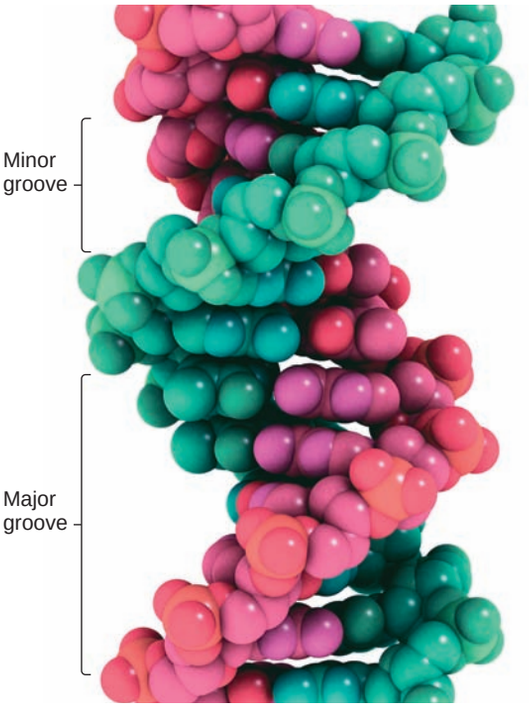
By convention, the direction of the DNA double helix shown in Figure 9.17a spirals in a direction that is called "right-handed." To understand this terminology, imagine that a double helix is laid on your desk; one end of the helix is close to you, and the other end is at the opposite side of the desk. As it spirals away from you, a right-handed helix turns in a clockwise direction. By comparison, a left-handed helix would spiral in a counterclockwise manner. Both strands in Figure 9.17a spiral in a right-handed direction.

Figure 9.17b is a space-filling model for DNA in which the atoms are represented by spheres. This model emphasizes the surface features of DNA. Note that the backbone—composed of sugar and phosphate groups—is on the outermost surface. In a living cell, the backbone has the most direct contact with water. In contrast, the bases are more internally located within the double-stranded structure. Biochemists use the term **grooves** to describe the indentations where the atoms of the bases are in contact with the surrounding water. As you travel around the DNA helix, the structure of DNA has two grooves: the **major groove** and the **minor groove.**

As discussed in later chapters, proteins can bind to DNA and affect its conformation and function. For example, some proteins can hydrogen bond to the bases within the major groove. This hydrogen bonding can be very precise so that a protein interacts with a particular sequence of bases. In this way, a protein can recognize a specific gene and affect its ability to be transcribed. We will consider such proteins in Chapters 12, 14, and 15. Alternatively, other proteins bind to the DNA backbone. For example, histone proteins, which are discussed in Chapter 10, form ionic interactions with the negatively charged phosphates in the DNA backbone. The histones are important for the proper compaction of DNA in eukaryotic cells and also play a role in gene transcription.



(a) Ball-and-stick model of DNA



(b) Space-filling model of DNA

ONLINE ANIMATION

FIGURE 9.17 Two models of the double helix. **(a)** Ball-and-stick model of the double helix. The deoxyribose-phosphate backbone is shown in detail, whereas the bases are depicted as flattened rectangles. **(b)** Space-filling model of the double helix.

DNA Can Form Alternative Types of Double Helices

The DNA double helix can form different types of structures. **Figure 9.18** compares the structures of **A DNA, B DNA,** and **Z DNA.** The highly detailed structures shown here were deduced by X-ray crystallography on short segments of DNA. B DNA is the predominant form of DNA found in living cells. However, under certain in vitro conditions, the two strands of DNA can twist into A DNA or Z DNA, which differ significantly from B DNA. A and B DNA are right-handed helices; Z DNA has a left-handed conformation. In addition, the helical backbone in Z DNA appears to zigzag slightly as it winds itself around the double-helical structure. The numbers of base pairs per 360° turn

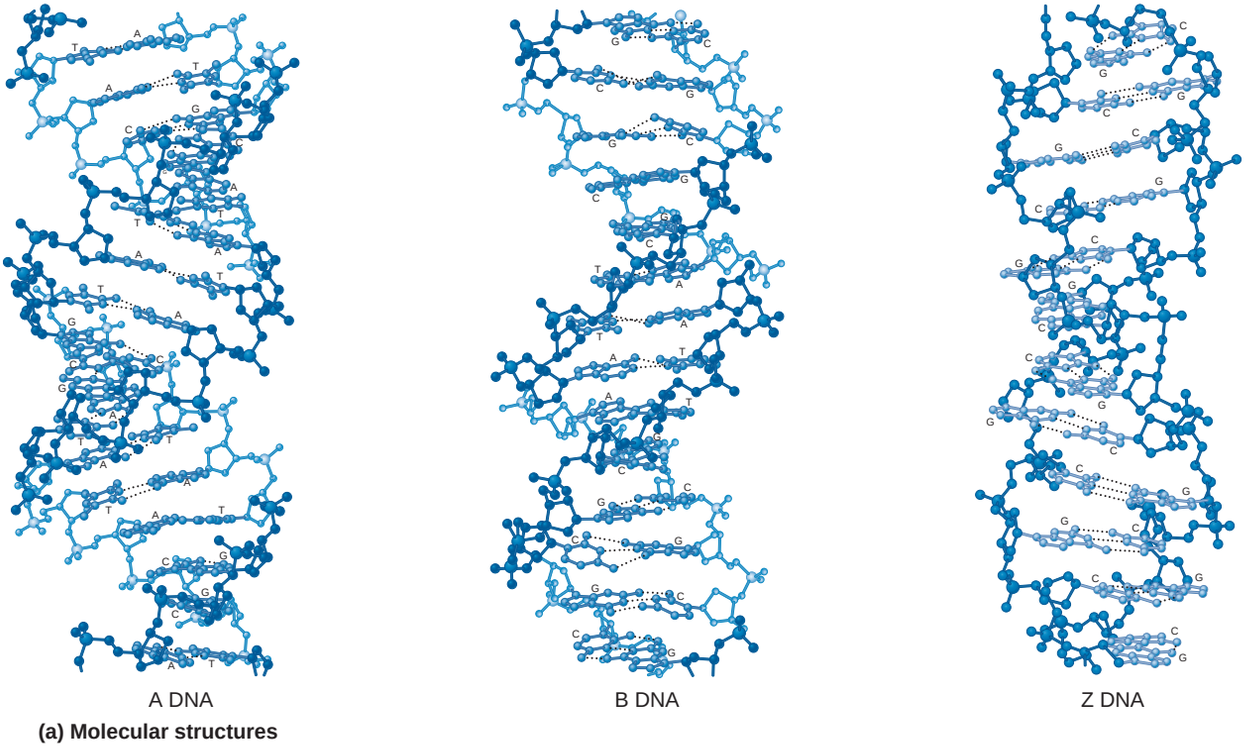


FIGURE 9.18 Comparison of the structures of A DNA, B DNA, and Z DNA. **(a)** The highly detailed structures shown here were deduced by X-ray crystallography performed on short segments of DNA. In contrast to the less detailed structures obtained from DNA wet fibers, the diffraction pattern obtained from the crystallization of short segments of DNA provides much greater detail concerning the exact placement of atoms within a double-helical structure. Alexander Rich, Richard Dickerson, and their colleagues were the first researchers to crystallize a short piece of DNA. **(b)** Space-filling models of the B-DNA and Z-DNA structures. In the case of Z DNA, the black lines connect the phosphate groups in the DNA backbone. As seen here, they travel along the backbone in a zigzag pattern.

are 11.0, 10.0, and 12.0 in A, B, and Z DNA, respectively. In B DNA, the bases tend to be centrally located, and the hydrogen bonds between base pairs occur relatively perpendicular to the central axis. In contrast, the bases in A DNA and Z DNA are substantially tilted relative to the central axis.

The ability of the predominant B DNA to adopt A-DNA and Z-DNA conformations depends on certain conditions. In X-ray diffraction studies, A DNA occurs under conditions of low humidity. The ability of a double helix to adopt a Z-DNA conformation depends on various factors. At high ionic strength (i.e., high salt concentration), formation of a Z-DNA conformation is favored by a sequence of bases that alternates between purines and pyrimidines. One such sequence is

5'–GCGCGCGCG–3'
3'–CGCGCGCGC–5'

At lower ionic strength, the methylation of cytosine bases can favor Z-DNA formation. Cytosine **methylation** occurs when a cellular enzyme attaches a methyl group ($-CH_3$) to the cytosine base. In addition, negative supercoiling (a topic discussed in Chapter 10) favors the Z-DNA conformation.

What is the biological significance of A and Z DNA? Research has not found any biological role for A DNA. However, accumulating evidence suggests a possible biological role for Z DNA in the process of transcription. Recent research has identified cellular proteins that specifically recognize Z DNA. In 2005, Alexander Rich and colleagues reported that the Z-DNA-binding region of one such protein played a role in regulating the transcription of particular genes. In addition, other research has suggested that Z DNA may play a role in chromosome structure by affecting the level of compaction.

DNA Can Form a Triple Helix, Called Triplex DNA

A surprising discovery made in 1957 by Alexander Rich, David Davies, and Gary Felsenfeld was that DNA can form a triple-helical structure called **triplex DNA.** This triplex was formed in vitro using pieces of DNA that were made synthetically. Although this result was interesting, it seemed to have little, if any, biological relevance.

About 30 years later, interest in triplex DNA was renewed by the observation that triplex DNA can form in vitro by mixing natural double-stranded DNA and a third short strand that is synthetically made. The synthetic strand binds into the major groove of the naturally occurring double-stranded DNA (**Figure 9.19**). As shown here, an interesting feature of triplex DNA formation is that it is sequence-specific. In other words, the synthetic third strand incorporates itself into a triple helix due to specific interactions between the synthetic DNA and the biological DNA. The pairing rules are that a thymine in the synthetic DNA hydrogen bonds at an AT pair in the biological DNA and that a cytosine in the synthetic DNA hydrogen bonds at a GC pair.

The formation of triplex DNA has been implicated in several cellular processes, including recombination, which is described in Chapter 17. In addition, researchers are interested in triplex DNA due to its potential as a tool to specifically inhibit particular genes. As shown in Figure 9.19, the synthetic DNA

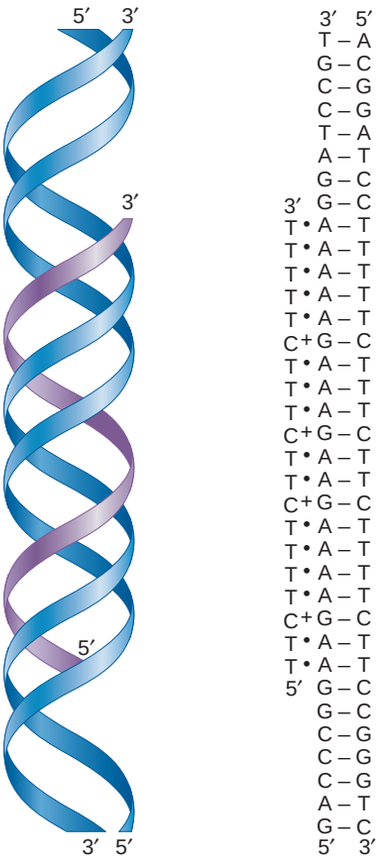


FIGURE 9.19 **The structure of triplex DNA.** **(a)** As seen in the ribbon model, the third, synthetic strand binds within the major groove of the double-stranded structure. **(b)** Within triplex DNA, the third strand hydrogen bonds according to the rule T to AT, and C to GC. The cytosine bases in the third strand are protonated (i.e., positively charged).

strand binds into the major groove according to specific base-pairing rules. Therefore, researchers can design a synthetic DNA to recognize the base sequence found in a particular gene. When the synthetic DNA binds to a gene, it inhibits transcription. In addition, the synthetic DNA can cause mutations in a gene that inactivate its function. Researchers are excited about the possibility of using such synthetic DNA to silence the expression of particular genes. For example, this approach could be used to silence genes that become overactive in cancer cells. However, further research is needed to develop effective ways to promote the uptake of synthetic DNAs into the appropriate target cells.

The Three-Dimensional Structure of DNA Within Chromosomes Requires Additional Folding and the Association with Proteins

To fit within a living cell, the long double-helical structure of chromosomal DNA must be extensively compacted into a three-dimensional conformation. With the aid of DNA-binding proteins, such as histone proteins, the double helix becomes greatly twisted and folded. **Figure 9.20** depicts the relationship between the DNA

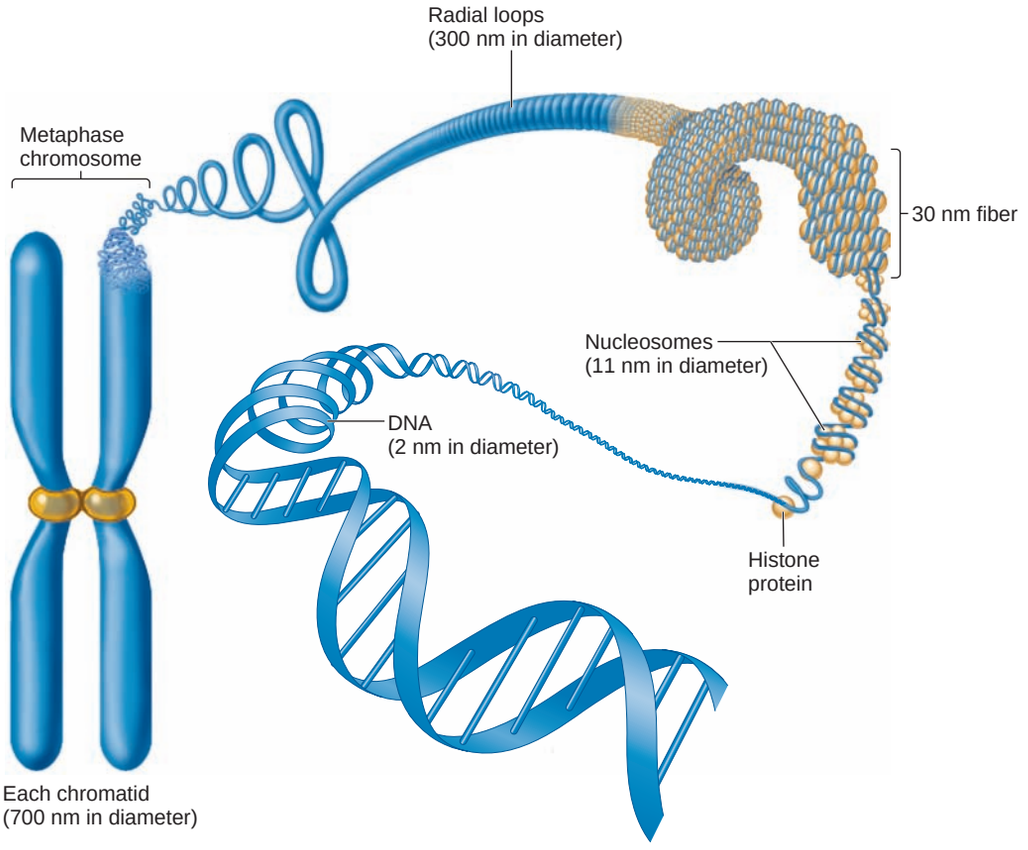


FIGURE 9.20 **The steps in eukaryotic chromosomal compaction leading to the metaphase chromosome.** The DNA double helix is wound around histone proteins and then is further compacted to form a highly condensed metaphase chromosome. The levels of DNA compaction will be described in greater detail in Chapter 10.

double helix and the compaction that occurs within a eukaryotic chromosome. Chapter 10 is devoted to the topic of chromosome organization and the molecular mechanisms responsible for the packaging of genetic material in cells.

RNA Molecules Are Composed of Strands That Fold into Specific Structures

Let's now turn our attention to RNA structure, which bears many similarities to DNA structure. The structure of an RNA strand is much like a DNA strand (**Figure 9.21**). Strands of RNA are typically several hundred or several thousand nucleotides in length, which is much shorter than chromosomal DNA. When RNA is made during transcription, the DNA is used as a template to make a copy of single-stranded RNA. In most cases, only one of the two DNA strands is used as a template for RNA synthesis. Therefore, only one complementary strand of RNA is usually made. Nevertheless, relatively short sequences within one RNA molecule or between two separate RNA molecules can form double-stranded regions.

The helical structure of RNA molecules is due to the ability of complementary regions to form base pairs between A and U and between G and C. This base pairing allows short segments to form a double-stranded region. As shown in **Figure 9.22**, different types of structural patterns are possible. These include bulge loops, internal loops, multibranched junctions, and stem-loops (also called hairpins). These structures contain regions of complementarity punctuated by regions of noncomplementarity. As shown in Figure 9.22, the complementary regions are held together by connecting hydrogen bonds, whereas the noncomplementary regions have their bases projecting away from the double-stranded region.

Many factors contribute to the structure of RNA molecules. These include the base-paired double-stranded helices, stacking between bases, and hydrogen bonding between bases and backbone regions. In addition, interactions with ions, small molecules, and large proteins may influence RNA structure. **Figure 9.23** depicts the structure of a transfer RNA molecule known as $tRNA^{Phe}$, which is a tRNA molecule that carries the amino acid phenylalanine. It was the first naturally occurring RNA to have its structure elucidated. This RNA molecule has several double-stranded and single-stranded regions. RNA double helices are antiparallel and right-handed, with 11 to 12 bp per turn. In a living cell, the various regions of an RNA molecule fold and interact with each other to produce the three-dimensional structure.

The folding of RNA into a three-dimensional structure is important for its function. For example, as discussed in Chapter 13, a tRNA molecule has two key functional sites—an anticodon and a 3′ acceptor site—that play important roles in translation. In a folded tRNA molecule, these sites are exposed on the surface of the molecule and can perform their roles (see

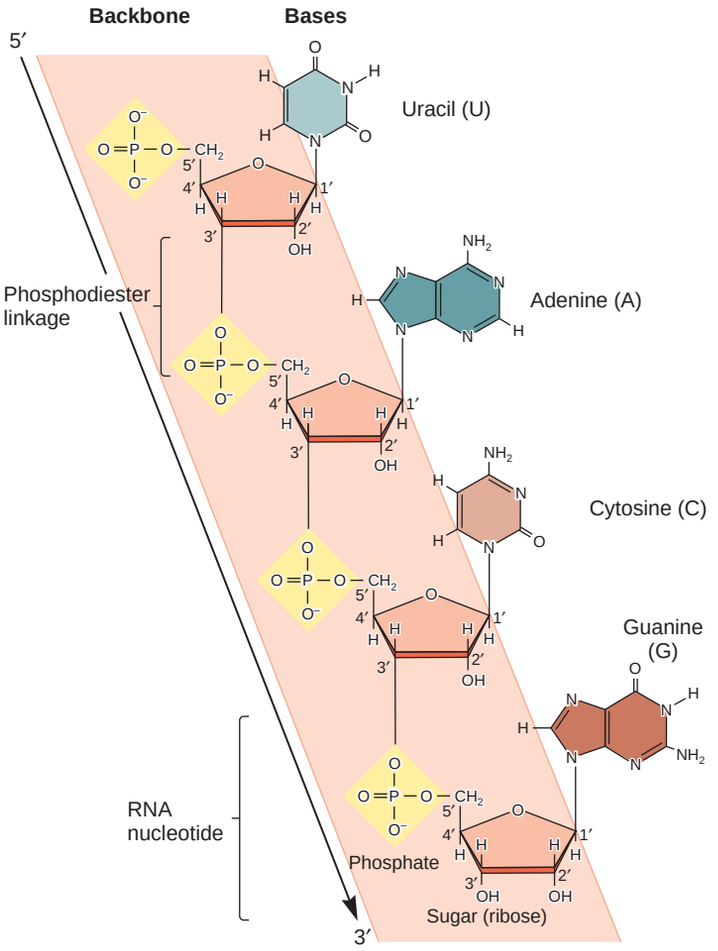
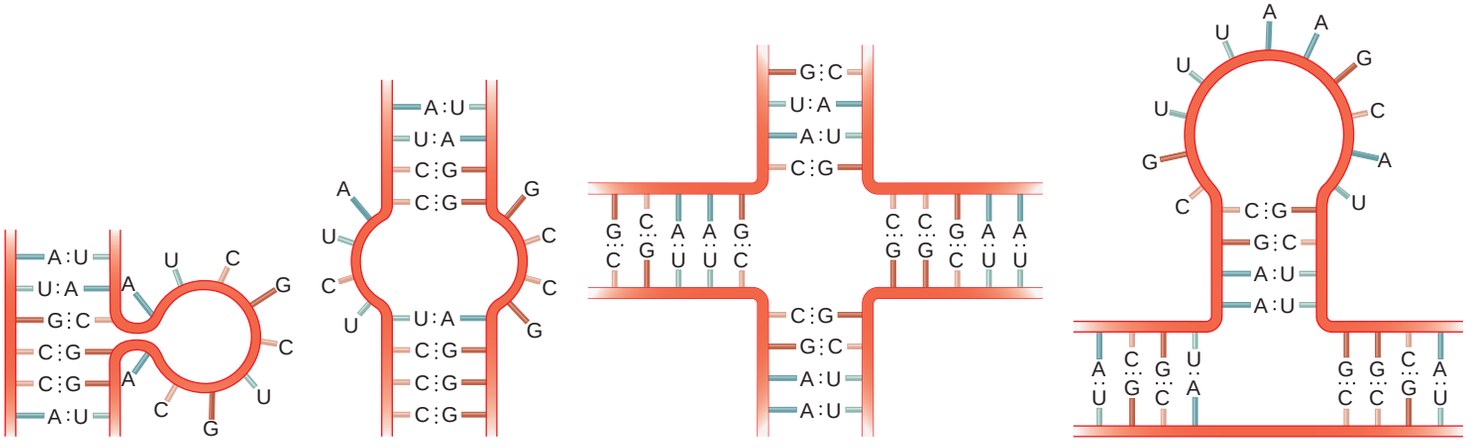
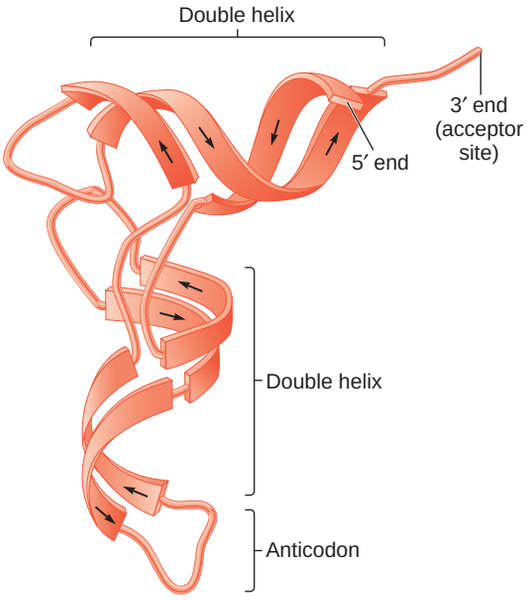


FIGURE 9.21 **A strand of RNA.** This structure is very similar to a DNA strand (see Figure 9.10), except that the sugar is ribose instead of deoxyribose, and uracil is substituted for thymine.



(a) Bulge loop (b) Internal loop (c) Multibranched junction (d) Stem-loop

FIGURE 9.22 **Possible structures of RNA molecules.** The double-stranded regions are depicted by connecting hydrogen bonds. Loops are noncomplementary regions that are not hydrogen bonded with complementary bases. Double-stranded RNA structures can form within a single RNA molecule or between two separate RNA molecules.



(a) Ribbon model

(b) Space-filling model

FIGURE 9.23 The structure of $tRNA^{Phe}$, the transfer RNA molecule that carries phenylalanine. **(a)** The double-stranded regions of the molecule are shown as antiparallel ribbons. **(b)** A space-filling model of $tRNA^{Phe}$.

Figure 9.23a). Many other examples are known in which RNA folding is key to its structure and function. These include the folding of ribosomal RNAs (rRNAs), which are important components of ribosomes, and ribozymes, which are RNA molecules with catalytic function.

KEY TERMS

Page 222. molecular genetics, DNA (deoxyribonucleic acid), RNA (ribonucleic acid)
Page 224. transformation
Page 225. DNase, RNase, protease, bacteriophage, phage
Page 226. lysis
Page 229. nucleic acids, nucleotides, strand, double helix, deoxyribose, ribose, purines, pyrimidines, adenine (A), guanine (G), thymine (T), cytosine (C), uracil (U)
Page 230. nucleoside
Page 231. phosphodiester linkage, backbone, directionality
Page 234. Chargaff's rule
Page 236. base pairs (bp), AT/GC rule
Page 237. complementary, antiparallel, grooves, major groove, minor groove
Page 238. A DNA, B DNA, Z DNA
Page 239. methylation, triplex DNA

CHAPTER SUMMARY

- Molecular genetics is the study of DNA structure and function at the molecular level.

9.1 Identification of DNA as the Genetic Material

- To fulfill its role, a genetic material must meet four criteria: information, transmission, replication, and variation.
- Griffith showed that the genetic material from type S bacteria could transform type R bacteria into type S (see Figure 9.1).
- Avery, MacLeod, and McCarty discovered that the transforming substance is DNA (see Figure 9.2).
- Hershey and Chase determined that the genetic material of T2 phage is DNA (see Figures 9.3–9.5).
- Viruses may use DNA or RNA as their genetic material (see Table 9.1).

9.2 Nucleic Acid Structure

- DNA and RNA are types of nucleic acids.
- In DNA, nucleotides are linked together to form strands, which then form a double helix that is found within chromosomes (see Figure 9.6).
- A nucleotide is composed of one or more phosphates, a sugar, and a base. The purine bases are adenine and guanine, whereas the pyrimidine bases are thymine (DNA only), cytosine, and uracil (RNA only) (see Figures 9.7–9.9).
- In a DNA strand, nucleotides are covalently attached to one another via phosphodiester linkages (see Figure 9.10).
- Pauling used ball-and-stick models to deduce the structure of an α helix in a protein (see Figure 9.11).
- Franklin performed X-ray diffraction studies that helped to determine the structure of DNA (see Figure 9.12).

- Chargaff determined that, in DNA, the amount of A equals T and the amount of G equals C (see Figure 9.13).
- Watson and Crick deduced the structure of DNA, though they proposed incorrect models along the way (see Figures 9.14, 9.15).
- DNA is a right-handed double helix in which A hydrogen bonds to T and G hydrogen bonds to C. The two strands are antiparallel and contain about 10 bp per turn (see Figure 9.16).
- The spiral structure of DNA has a major groove and a minor groove (see Figure 9.17).
- B DNA is major form of DNA found in living cells. A DNA and Z DNA are alternative conformations for DNA (see Figure 9.18).
- Under certain conditions, DNA can form a triple helix structure that obeys specific base-pairing rules (see Figure 9.19).
- In a chromosome, DNA interacts with proteins and is highly compacted (see Figure 9.20)
- RNA is also a strand of nucleotides (see Figure 9.21).
- RNA can form double-stranded helical regions, and it folds into a three-dimensional structure (see Figures 9.22, 9.23).

PROBLEM SETS & INSIGHTS

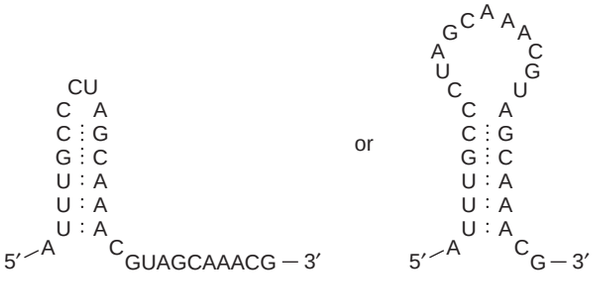
Solved Problems

S1. A hypothetical sequence at the beginning of an mRNA molecule is

5′–AUUUGCCCUAGCAAACGUAGCAAACG . . . rest of the coding sequence

Using two out of the three underlined sequences, draw two possible models for potential stem-loop structures at the 5' end of this mRNA.

Answer:



S2. Describe the previous experimental evidence that led Watson and Crick to the discovery of the DNA double helix.

Answer:

1. The chemical structure of single nucleotides was understood by the 1950s.
2. Watson and Crick assumed DNA is composed of nucleotides linked together in a linear fashion to form a strand. They also assumed the chemical linkage between two nucleotides is always the same.
3. Franklin's diffraction patterns suggested several structural features. First, it was consistent with a helical structure. Second, the diameter of the helical structure was too wide to be only a single-stranded helix. Finally, the line spacing on the diffraction pattern indicated the helix contains about 10 bp per complete turn.
4. In the chemical analysis of the DNA from different species, the work of Chargaff indicated that the amount of adenine equaled the amount of thymine, and the amount of cytosine equaled that of guanine.
5. In the early 1950s, Linus Pauling proposed that regions of proteins can fold into a secondary structure known as an α helix. To discover this, Pauling built large models by linking together simple ball-and-stick units. In this way, he could determine if atoms fit together properly in a complicated three-dimensional structure. A similar approach was used by Watson and Crick to solve the structure of the DNA double helix.

S3. Within living cells, many different proteins play important functional roles by binding to DNA and RNA. As described throughout your textbook, the dynamic interactions between nucleic acids and proteins lie at the heart of molecular genetics. Some proteins bind to DNA (or RNA) but not in a sequence-specific manner. For example, histones are proteins important in the formation of chromosome structure. In this case, the positively charged histone proteins actually bind to the negatively charged phosphate groups in DNA. In addition, several other proteins interact with DNA but do not require a specific nucleotide sequence to carry out their function. For example, DNA polymerase, which catalyzes the synthesis of new DNA strands, does not bind to DNA in a sequence-dependent manner. By comparison, many other proteins do interact with nucleic acids in a sequence-dependent fashion. This means that a specific sequence of bases can provide a structure that is recognized by a particular protein. Throughout the textbook, the functions of many of these proteins will be described. Some examples include transcription factors that affect the rate of transcription, proteins that bind to centromeres, and proteins that bind to origins of replication. With regard to the three-dimensional structure of DNA, where would you expect DNA-binding proteins to bind if they recognize a specific base sequence? What about DNA-binding proteins that do not recognize a base sequence?

Answer: DNA-binding proteins that recognize a base sequence must bind into a major or minor groove of the DNA, which is where the bases are accessible to a DNA-binding protein. Most DNA-binding proteins, which recognize a base sequence, fit into the major groove. By comparison, other DNA-binding proteins, such as histones, which do not recognize a base sequence, bind to the DNA backbone.

S4. The formation of a double-stranded structure must obey the rule that adenine hydrogen bonds to thymine (or uracil) and cytosine hydrogen bonds to guanine. Based on your previous understanding of genetics (from this course or a general biology course), discuss reasons why complementarity is an important feature of DNA and RNA structure and function.

Answer: Note: Many of the topics described next are discussed in Chapters 10 through 13. One way that complementarity underlies function is that it provides the basis for the synthesis of new strands of DNA and RNA. During replication, the synthesis of the new DNA strands occurs in such a way that adenine hydrogen bonds to thymine, and cytosine hydrogen bonds to guanine. In other words, the molecular feature of a complementary double-stranded structure makes it possible to produce

exact copies of DNA. Likewise, the ability to transcribe DNA into RNA is based on complementarity. During transcription, one strand of DNA is used as a template to make a complementary strand of RNA.

In addition to the synthesis of new strands of DNA and RNA, complementarity is important in other ways. As mentioned in this chapter, the folding of RNA into a particular structure is driven by the hydrogen bonding of complementary regions. This event is necessary to produce functionally active tRNA molecules. Likewise, stem-loop structures also occur in other types of RNA. For example, the rapid formation of stem-loop structures is known to occur as RNA is being transcribed and to affect the termination of transcription.

A third way that complementarity can be functionally important is that it can promote the interaction of two separate RNA molecules. During translation, codons in mRNA bind to the anticodons in tRNA (see Chapter 13). This binding is due to complementarity. For example, if a codon is 5′–AGG–3′, the anticodon is 3′–UCC–5′. This type of specific interaction between codons and anticodons is an important step that enables the nucleotide sequence in mRNA to code for an amino acid sequence within a protein. In addition, many other examples of RNA-RNA interactions are known and will be described throughout this textbook.

S5. An important feature of triplex DNA formation is that it is sequence-specific. The synthetic third strand incorporates itself into a triple helix, so a thymine in the synthetic DNA binds near an AT pair in the biological DNA, and a cytosine in the synthetic DNA binds near a GC pair. From a practical point of view, this opens the possibility of synthesizing a short strand of DNA that forms a triple helix at a particular target site. For example, if the sequence of a particular gene is known, researchers can make a synthetic piece of DNA that forms a triple helix somewhere within that gene according to the T to AT, and C to GC rule. Triplex DNA formation is known to inhibit gene transcription. In other words, when the synthetic DNA binds within the DNA of a gene, the formation of triplex DNA prevents that gene from being transcribed into RNA. Discuss how this observation might be used to combat diseases.

Answer: Triplex DNA formation opens the exciting possibility of designing synthetic pieces of DNA to inhibit the expression of particular genes. Theoretically, such a tool could be used to combat viral diseases or to inhibit the growth of cancer cells. To combat a viral disease, a synthetic DNA could be made that specifically binds to an essential viral gene, thereby preventing viral proliferation. To inhibit cancer, a synthetic DNA could be made to bind to an oncogene. (Note: As described in Chapter 22, an oncogene is a gene that promotes cancerous growth.) Inhibition of an oncogene could prevent cancer. At this point, a primary obstacle in applying this approach is devising a method of getting the synthetic DNA into living cells.

Conceptual Questions

C1. What is the meaning of the term genetic material?

C2. After the DNA from type S bacteria is exposed to type R bacteria, list all of the steps that you think must occur for the bacteria to start making a capsule.

C3. Look up the meaning of the word transformation in a dictionary and explain whether it is an appropriate word to describe the transfer of genetic material from one organism to another.

C4. What are the building blocks of a nucleotide? With regard to the 5′ and 3′ positions on a sugar molecule, how are nucleotides linked together to form a strand of DNA?

C5. Draw the structure of guanine, guanosine, and deoxyguanosine triphosphate.

C6. Draw the structure of a phosphodiester linkage.

C7. Describe how bases interact with each other in the double helix. This discussion should address the issues of complementarity, hydrogen bonding, and base stacking.

C8. If one DNA strand is 5′–GGCATTACACTAGGCCT–3′, what is the sequence of the complementary strand?

C9. What is meant by the term DNA sequence?

C10. Make a side-by-side drawing of two DNA helices, one with 10 bp per 360° turn and the other with 15 bp per 360° turn.

C11. Discuss the differences in the structural features of A DNA, B DNA, and Z DNA.

C12. What parts of a nucleotide (namely, phosphate, sugar, and/or bases) occupy the major and minor grooves of double-stranded DNA, and what parts are found in the DNA backbone? If a DNA-binding protein does not recognize a specific nucleotide sequence, do you expect that it recognizes the major groove, the minor groove, or the DNA backbone? Explain.

C13. List the structural differences between DNA and RNA.

C14. Draw the structure of deoxyribose and number the carbon atoms. Describe the numbering of the carbon atoms in deoxyribose with regard to the directionality of a DNA strand. In a DNA double helix, what does the term antiparallel mean?

C15. Write out a sequence of an RNA molecule that could form a stem-loop with 24 nucleotides in the stem and 16 nucleotides in the loop.

C16. Compare the structural features of a double-stranded RNA structure with those of a DNA double helix.

C17. Which of the following DNA double helices would be more difficult to separate into single-stranded molecules by treatment with heat, which breaks hydrogen bonds?

A. GGCGTACCAGCGCAT
CCGCATGGTCGCGTA

B. ATACGATTTACGAGA
TATGCTAAATGCTCT

Explain your choice.

C18. What structural feature allows DNA to store information?

C19. Discuss the structural significance of complementarity in DNA and in RNA.

C20. An organism has a G + C content of 64% in its DNA. What are the percentages of A, T, G, and C?

C21. Let's suppose you have recently identified an organism that was scraped from an asteroid that hit the earth. (Fortunately, no one was injured.) When you analyze this organism, you discover that its DNA is a triple helix, composed of six different nucleotides: A, T, G, C, X, and Y. You measure the chemical composition of the bases and find the following amounts of these six bases: A = 24%,

T = 23%, G = 11%, C = 12%, X = 21%, Y = 9%. What rules would you propose that govern triplex DNA formation in this organism? Note: There is more than one possibility.

C22. On further analysis of the DNA described in conceptual question C21, you discover that the triplex DNA in this alien organism is composed of a double helix, with the third helix wound within the major groove (just like the DNA in Figure 9.19). How would you propose that this DNA is able to replicate itself? In your answer, be specific about the base pairing rules within the double helix and which part of the triplex DNA would be replicated first.

C23. A DNA-binding protein recognizes the following double-stranded sequence:

5′–GCCCGGGC–3′
3′–CGGGCCCG–5′

This type of double-stranded structure could also occur within the stem region of an RNA stem-loop molecule. Discuss the structural differences between RNA and DNA that might prevent this DNA-binding protein from recognizing a double-stranded RNA molecule.

C24. Within a protein, certain amino acids are positively charged (e.g., lysine and arginine), some are negatively charged (e.g., glutamate and aspartate), some are polar but uncharged, and some are nonpolar. If you knew that a DNA-binding protein was recognizing the DNA backbone rather than base sequences, which amino acids in the protein would be good candidates for interacting with the DNA?

C25. In what ways are the structures of an α helix in proteins and the DNA double helix similar, and in what ways are they different?

C26. A double-stranded DNA molecule contains 560 nucleotides. How many complete turns would be found in this double helix?

C27. As the minor and major grooves of the DNA wind around a DNA double helix, do they ever intersect each other, or do they always run parallel to each other?

C28. What chemical group (phosphate group, hydroxyl group, or a nitrogenous base) is found at the 3′ end of a DNA strand? What group is found at the 5′ end?

C29. The base composition of an RNA virus was analyzed and found to be 14.1% A, 14.0% U, 36.2% G, and 35.7% C. Would you conclude that the viral genetic material is single-stranded RNA or double-stranded RNA?

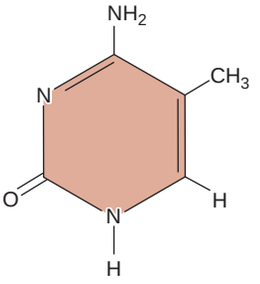
C30. The genetic material found within some viruses is single-stranded DNA. Would this genetic material contain equal amounts of A and T and equal amounts of G and C?

C31. A medium-sized human chromosome contains about 100 million bp. If the DNA were stretched out in a linear manner, how long would it be?

C32. A double-stranded DNA molecule is 1 cm long, and the percentage of adenine is 15%. How many cytosines would be found in this DNA molecule?

C33. Could single-stranded DNA form a stem-loop structure? Why or why not?

C34. As described in Chapter 15, the methylation of cytosine bases can have an important effect on gene expression. For example, the methylation of cytosines may inhibit the transcription of genes. A methylated cytosine base has the following structure:



Would you expect the methylation of cytosine to affect the hydrogen bonding between cytosine and guanine in a DNA double helix? Why or why not? (Hint: See Figure 9.16 for help.) Take a look at solved problem S3 and speculate as to how methylation could affect gene expression.

C35. An RNA molecule has the following sequence:

Region 1 Region 2 Region 3

5′-CAUCCAUCCAUUCCCCAUCCGAUAAGGGGAAUGGAUCCGAAUGGAUAAC-3′

Parts of region 1 can form a stem-loop with region 2 and with region 3. Can region 1 form a stem-loop with region 2 and region 3 at the same time? Why or why not? Which stem-loop would you predict to be more stable: a region 1/region 2 interaction or a region 1/region 3 interaction? Explain your choice.

Experimental Questions

E1. Genetic material acts as a blueprint for an organism's traits. Explain how the experiments of Griffith indicated that genetic material was being transferred to the type R bacteria.

E2. With regard to the experiment described in Figure 9.2, answer the following:

A. List several possible reasons why only a small percentage of the type R bacteria was converted to type S.

B. Explain why an antibody must be used to remove the bacteria that are not transformed. What would the results look like, in all five cases, if the antibody/centrifugation step had not been included in the experimental procedure?

C. The DNA extract was treated with DNase, RNase, or protease. Why was this done? (In other words, what were the researchers trying to demonstrate?)

E3. An interesting trait that some bacteria exhibit is resistance to killing by antibiotics. For example, certain strains of bacteria are resistant to tetracycline, whereas other strains are sensitive to tetracycline. Describe an experiment you would carry out to demonstrate that tetracycline resistance is an inherited trait encoded by the DNA of the resistant strain.

E4. With regard to the experiment of Figure 9.5, answer the following:

A. Provide possible explanations why some of the DNA is in the supernatant.

B. Plot the results if the radioactivity in the pellet, rather than in the supernatant, had been measured.

C. Why were ^{32}P and ^{35}S chosen as radioisotopes to label the phages?

D. List possible reasons why less than 100% of the phage protein was removed from the bacterial cells during the shearing process.

E5. Does the experiment of Figure 9.5 rule out the possibility that RNA is the genetic material of T2 phage? Explain your answer. If it does not, could you modify the approach of Hershey and Chase to show that it is DNA and not RNA that is the genetic material of T2 bacteriophage? Note: It is possible to specifically label DNA or RNA by providing bacteria with radiolabeled thymine or uracil, respectively.

E6. In this chapter, we considered two experiments—one by Avery, MacLeod, and McCarty and the second by Hershey and Chase—that indicated DNA is the genetic material. Discuss the strengths and weaknesses of the two approaches. Which experimental approach did you find the most convincing? Why?

E7. The type of model building used by Pauling and Watson and Crick involved the use of ball-and-stick units. Now we can do model building on a computer screen. Even though you may not be familiar with this approach, discuss potential advantages of using computers in molecular model building.

E8. With regard to Chargaff's experiment described in Figure 9.13, answer the following:

A. What is the purpose of paper chromatography?

B. Explain why it is necessary to remove the bases in order to determine the base composition of DNA.

C. Would Chargaff's experiments have been convincing if they had been done on only one species? Discuss.

E9. Gierer and Schramm exposed plant tissue to purified RNA from tobacco mosaic virus, and the plants developed the same types of lesions as if they were exposed to the virus itself. What would be the results if the RNA was treated with DNase, RNase, or protease prior to its exposure to the plant tissue?

Questions for Student Discussion/Collaboration

1. Try to propose structures for a genetic material that are substantially different from the double helix. Remember that the genetic material must have a way to store information and a way to be faithfully replicated.
2. How might you provide evidence that DNA is the genetic material in mice?

Note: All answers appear at the website for this textbook; the answers to even-numbered questions are in the back of the textbook.

www.mhhe.com/brookergenetics4e

Visit the website for practice tests, answer keys, and other learning aids for this chapter. Enhance your understanding of genetics with our interactive exercises, quizzes, animations, and much more.