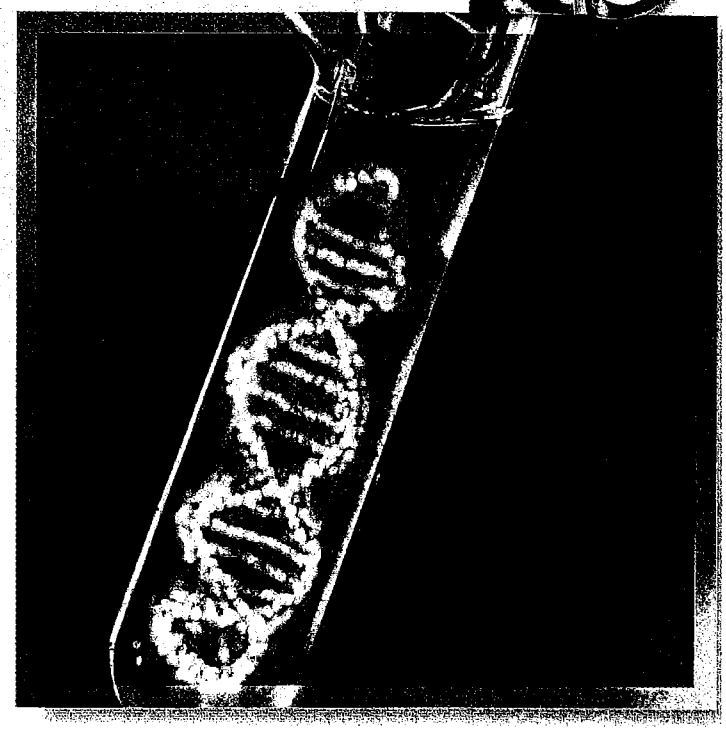
CHAPTER 25

Molecular Basis of Inheritance

Chapter Concepts



Test tube containing DNA whose structure resembles a spiral staircase (digital composite).

One of the most exciting periods of scientific activity in history occurred during the thirty short years between the 1930s and 1960s. Geneticists knew that chromosomes contain protein and DNA (deoxyribonucleic acid). Of these two organic molecules, proteins are seemingly more complicated; they consist of countless sequences of 20 amino acids, which can coil and fold into complex shapes. DNA, on the other hand, contains only four different nucleotides. Surely, the diversity of life forms on earth must be the result of the endless varieties of proteins.

Due to several elegantly executed experiments, by the mid-1950s researchers realized that DNA, not protein, is the genetic material. But this finding only led to another fundamental question—what exactly is the structure of DNA? *The biological community at the time knew that whoever determined the structure of DNA would get a Nobel Prize, and would go down in history. Consequently, researchers were racing against time and each other. The story of the discovery of DNA resembles a mystery, with each clue adding to the total picture until the breathtaking design of DNA—a double helix—was finally unraveled.*

25.1 DNA Structure and Replication

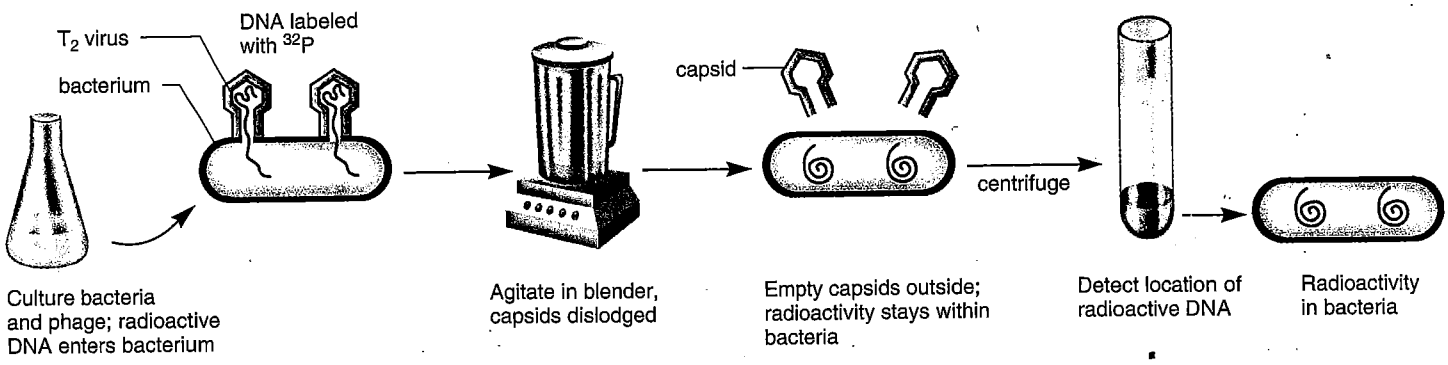
In the mid-1900s, scientists knew that the chromosomes contained genetic information. But because the chromosomes were composed of both **DNA (deoxyribonucleic acid)** and proteins, they were uncertain which one was the genetic material. Certain investigators turned to experiments with viruses to resolve this question since they knew that viruses are tiny particles having just two parts: an inner nucleic acid core and an outer protein coat, called a capsid. A virus called the T_2 virus (the T_2 simply means *type* 2) enters bacteria and reproduces. Researchers thought that if they could determine which part of the virus—DNA or protein—enters a bacterium and produces more viruses, they would be able to determine whether genes were made up of DNA or protein.

Two experiments were done (Fig. 25.1). In the first experiment, phage DNA was labeled with radioactive ^{32}P. The phages were allowed to attach to and inject their genetic material into *E. coli* cells. Then the culture was agitated in a kitchen blender to remove whatever remained of the phages on the outside of the bacterial cells. Finally, the culture was centrifuged (spun at high speed) so that the bacterial cells collected as a pellet at the bottom of the centrifuge tube. In this experiment, as you would predict, investigators found most of the ^{32}P-labeled DNA in the cells and not in the liquid medium. Why? Because the DNA had entered the cells.

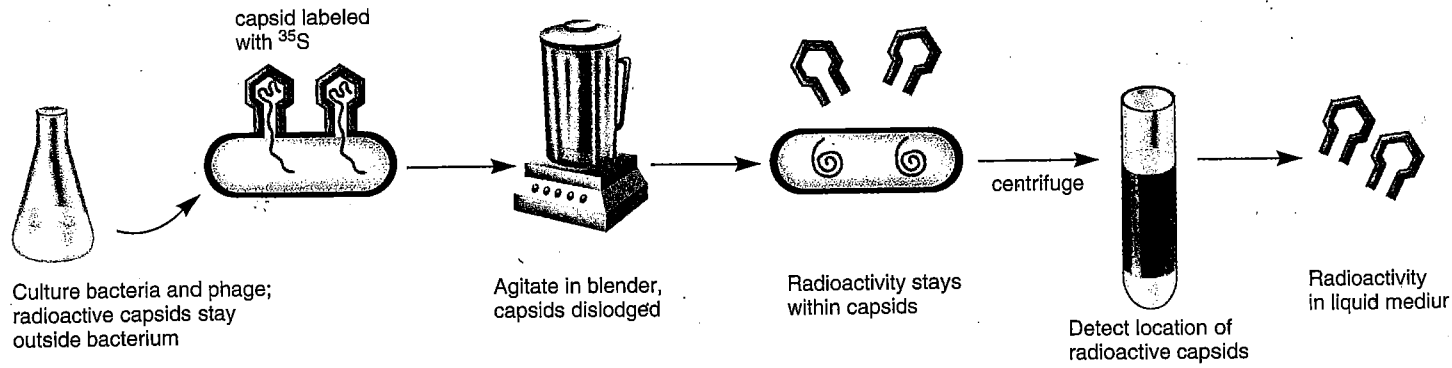
In the second experiment, phage protein in capsids was labeled with radioactive ^{35}S. The phages were allowed to attach to and inject their genetic material into *E. coli* bacterial cells. Then the culture was agitated in a kitchen blender to remove whatever remained of the phages on the outside of the bacterial cells. Finally, the culture was centrifuged so that the bacterial cells collected as a pellet at the bottom of the centrifuge tube. In this experiment, as you would predict, scientists found ^{35}S-labeled protein in the liquid medium and not in the cells. Why? Because the radioactive capsids remained on the outside of the cells and were removed by the blender.

These results indicated that the DNA of a virus, (not a protein), enters the host, where viral reproduction takes place. Therefore, DNA is the genetic material. It transmits all the genetic information needed to produce new viruses.

Viral experiments showed that DNA, not protein, is the genetic material.

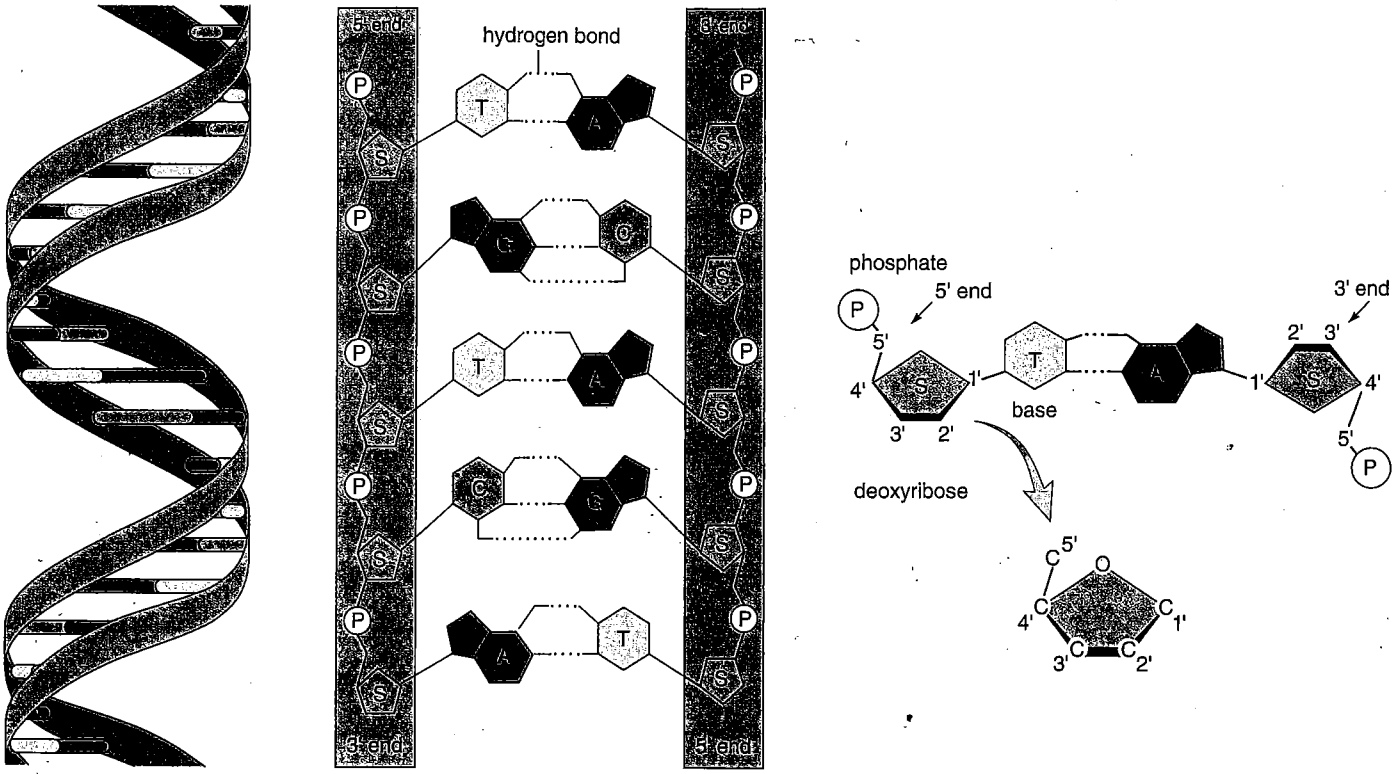


a. Viral DNA is labeled (red).



b. Viral capsid is labeled (red).

Figure 25.1 Discovering that DNA is the genetic material.



a. Double helix b. Ladder structure c. One pair of bases

Figure 25.2 Overview of DNA structure.

a. DNA double helix. **b.** When the helix is unwound, a ladder configuration shows that the uprights are composed of sugar and phosphate molecules and the rungs are complementary bases. Notice that the bases in DNA pair in such a way that the phosphate-sugar groups are oriented in different directions. This means that the strands of DNA end up running antiparallel to one another, with the 3′ end of one strand opposite the 5′ end of the other strand. **c.** Notice that 3′ and 5′ refer to a numbering system for the carbon atoms that make up the sugar.

Structure of DNA

The structure of DNA was determined by James Watson and Francis Crick in the early 1950s. The data they used and how they used the data to deduce DNA's structure are reviewed in the Science Focus on the previous page.

DNA is a polynucleotide; each nucleotide is a complex of three subunits—phosphoric acid (phosphate), a pentose sugar (deoxyribose), and a nitrogen-containing base. There are four possible bases: two are **purines** with a double ring, and two are **pyrimidines** with a single ring. The names of the bases are as follows:

Purines	Pyrimidines
Adenine (A)	Thymine (T)
Guanine (G)	Cytosine (C)

A polynucleotide *strand* has a backbone made up of alternating phosphate and sugar molecules. The bases are attached to the sugar but project to one side. DNA has two such strands, and the two strands twist about one another in the form of a **double helix** (Fig. 25.2*a*). The strands are held together by hydrogen bonding between the bases: A pairs with T by forming two hydrogen bonds, and G pairs with C by forming three hydrogen bonds, or vice versa. This is called **complementary base pairing.**

When the DNA helix unwinds, it resembles a ladder (Fig. 25.2*b*). The sides of the ladder are the phosphate-sugar backbones, and the rungs of the ladder are the complementary paired bases. Notice that a purine is always bonded to a pyrimidine.

The two DNA strands are antiparallel—that is, they run in opposite directions, which you can verify by noticing that the sugar molecules are oriented differently. The carbon atoms in a sugar molecule are numbered, and the fifth carbon atom (5′) is uppermost in the strand on the left, while the third carbon atom (3′)—attached to a phosphate—is uppermost in the strand on the right (Fig. 25.2*c*).

DNA is a double helix with phosphate-sugar backbones on the outside and paired bases on the inside. Complementary base pairing occurs: adenine (A) pairs with thymine (T), and guanine (G) pairs with cytosine (C).

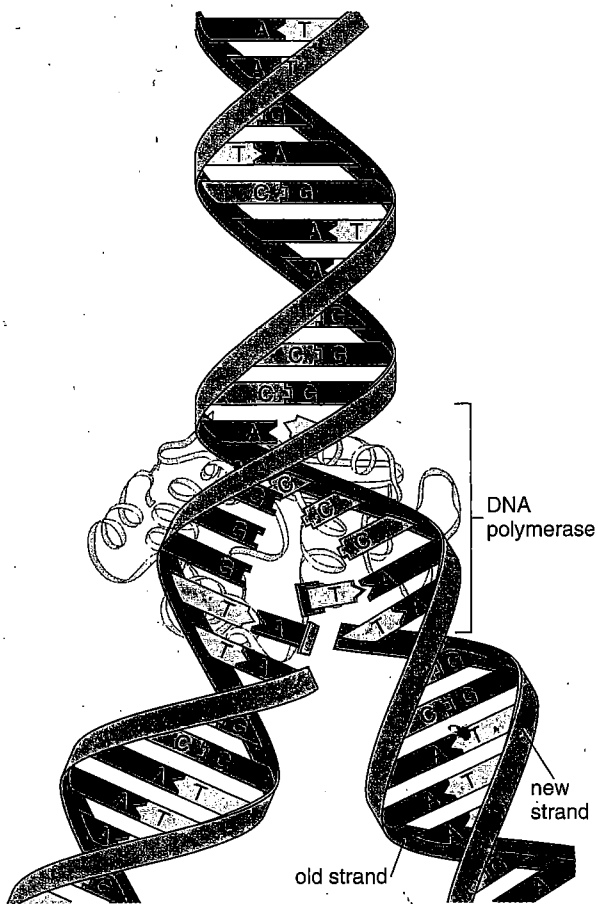


Figure 25.3 Overview of DNA replication.
Replication is called semiconservative because each new double helix is composed of an old (parental) strand and a new (daughter) strand.

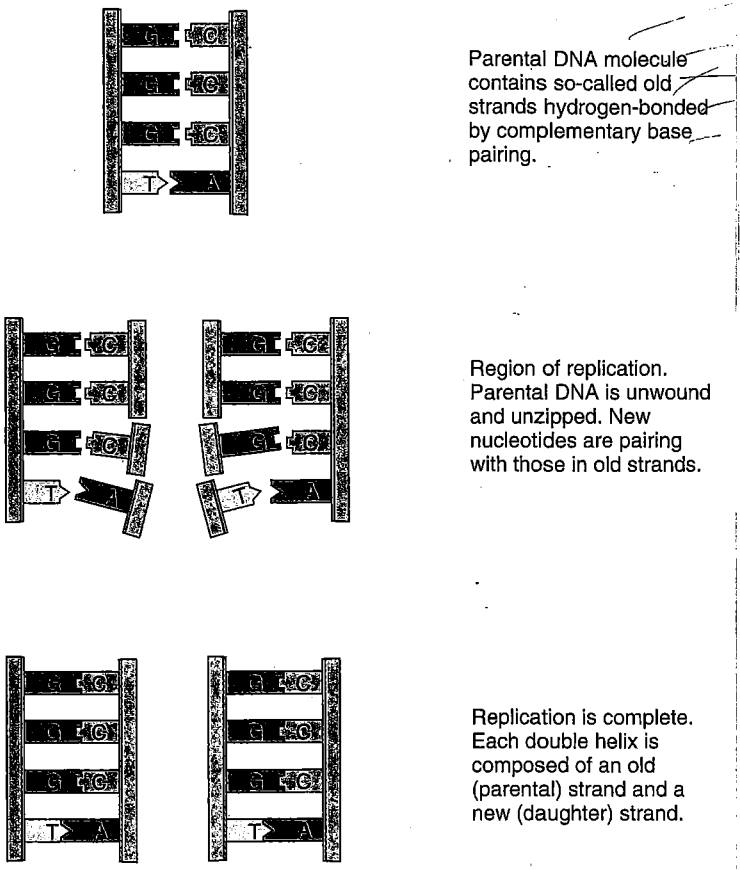


Figure 25.4 Ladder configuration and DNA replication.
Use of the ladder configuration better illustrates how complementary nucleotides available in the cell pair with those of each old strand before they are joined together to form a daughter strand.

Replication of DNA

During **replication,** an exact copy of the DNA is produced with the aid of an enzyme called **DNA polymerase.** The double-stranded structure of DNA aids replication because each strand can serve as a template for the formation of a complementary strand. A **template** is most often a mold used to produce a shape opposite to itself. In this case, each old (parental) strand is a template for each new (daughter) strand.

DNA replication is termed *semiconservative* because each new double helix has one old strand and one new strand. In other words, one of the parental strands is conserved, or present, in each new double helix. Because each old strand has produced a new strand through complementary base pairing, there are now two DNA helices identical to each other and to the original molecule (Fig. 25.3).

Figure 25.4 shows how replication is carried out.

1. Before replication begins, the two strands that make up parental DNA are hydrogen-bonded to one another.
2. During replication, DNA unwinds and "unzips" (i.e., the weak hydrogen bonds between the paired bases break). New complementary nucleotides, always present in the nucleus, fit into place by the process of complementary base pairing, and then they are joined to one another
3. When replication is complete, two double helix molecules are identical.

Cancer, which is characterized by rapidly dividing cells, is treated with chemotherapeutic drugs that stop replication and therefore cell division. Some chemotherapeutic drugs are analogs that have a similar, but not identical, structure to the four nucleotides in DNA. When these are mistakenly used by the cancer cells to synthesize DNA, replication stops and the cells die off.

During DNA replication, DNA unwinds and unzips, and new strands that are complementary to the original strands form.

25.2 Gene Expression

As mentioned in Chapter 24, an individual with a genetic disorder has a missing or malfunctioning enzymatic or structural protein in the cell. Persons with Duchenne muscular dystrophy, for example, are missing the protein called dystrophin, which is normally present in muscle cells. In the genetic disorder Huntington disease, the protein huntingtin is altered and unable to carry out its usual functions. People with cystic fibrosis have a malfunctioning Cl^- channel protein in the plasma membrane. By studying various metabolic disorders, geneticists have confirmed many times over that proteins are the link between genotype and phenotype.

Table 25.1 DNA Structure Compared to RNA Structure

	DNA	RNA
Sugar	Deoxyribose	Ribose
Bases	Adenine, guanine, thymine, cytosine	Adenine, guanine, uracil, cytosine
Strands	Double stranded with base pairing	Single stranded
Helix	Yes	No

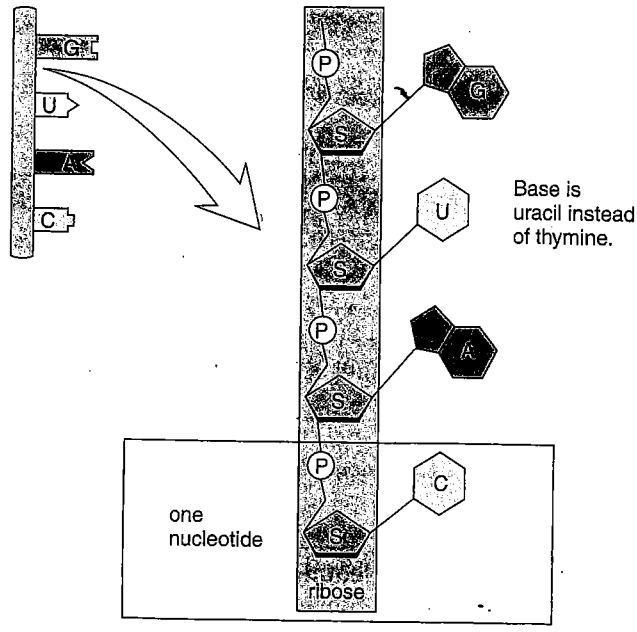


Figure 25.5 Structure of RNA.
Like DNA, RNA is a polymer of nucleotides. In an RNA nucleotide, the sugar ribose is attached to a phosphate molecule and to a base, either G, U, A, or C. Notice that in RNA, the base uracil replaces thymine as one of the pyrimidine bases. RNA is single stranded, whereas DNA is double stranded.

Granted that there is a connection between genes and proteins, what exactly is it that genes do? A **gene** is a segment of DNA that specifies the amino acid sequence of a protein. This is the information that DNA stores and the reason DNA activity brings about the development of the unique structures that make up a particular organism. In this chapter, you will learn how a difference in base sequence causes a difference in protein structure and determines, for example, whether you have blue, brown, or hazel eye pigments.

A gene does not directly control protein synthesis; instead, it passes its genetic information on to RNA, which is more directly involved in protein synthesis.

RNA

Like DNA, **RNA (ribonucleic acid)** is a polynucleotide (Fig. 25.5). However, the nucleotides in RNA contain the sugar ribose, not deoxyribose. Also, the bases in RNA are adenine (A), cytosine (C), guanine (G), and uracil (U). In other words, the base uracil replaces the thymine found in DNA. Finally, RNA is single stranded and does not form a double helix in the same manner as DNA (Fig. 25.5). Table 25.1 summarizes the similarities and differences between DNA and RNA.

There are three major classes of RNA, each with specific functions in protein synthesis:

Messenger RNA (mRNA) takes a message from DNA to the ribosomes.

Ribosomal RNA (rRNA), along with proteins, makes up the ribosomes, where proteins are synthesized.

Transfer RNA (tRNA) transfers amino acids to the ribosomes.

Gene expression requires two steps, known as transcription and translation. First, information is transferred from DNA to RNA during **transcription.** Consider that transcription means making a close copy of a document. During transcription, one type of polynucleotide (DNA) becomes another type of polynucleotide (RNA). Second, during **translation,** an RNA transcript directs the sequence of amino acids in a polypeptide. Consider that a translator needs to understand two languages. Similarly, the cell understands two different languages: nucleotide sequences and amino acid sequences. During translation, a nucleotide sequence directs an amino acid sequence (see Fig. 25.7).

With the help of RNA, a gene (a segment of DNA) specifies the sequence of amino acids in a polypeptide. In this way, genes control the structure and the metabolism of cells.

The Genetic Code

DNA has a particular sequence of bases, and a polypeptide has a particular sequence of amino acids. This suggests that DNA contains coded information. Can four bases provide enough combinations to code for 20 amino acids? If the code were a doublet (any two bases stand for one amino acid), it would not be possible to code for 20 amino acids, but if the code were a triplet, then the four bases could supply 64 different triplets, far more than needed to code for 20 different amino acids. It should come as no surprise, then, to learn that the code is a **triplet code.**

To crack the code, a cell-free experiment was done: artificial RNA was added to a medium containing bacterial ribosomes and a mixture of amino acids. Comparison of the bases in the RNA with the resulting polypeptide allowed investigators to decipher the code. Each three-letter unit of an mRNA molecule is called a **codon.** All 64 mRNA codons have been determined (Fig. 25.6). Sixty-one triplets correspond to a particular amino acid; the remaining three are stop codons, which signal polypeptide termination. The one codon that stands for the amino acid methionine is also a start codon signaling polypeptide initiation. Notice too that most amino acids have more than one codon; leucine, serine, and arginine have six different codons, for example. This offers some protection against possibly harmful mutations that change the sequence of the bases.

The genetic code is just about universal in living things. This suggests that the code dates back to the very first organisms on earth and that all living things are related.

First Base	Second Base				Third Base
	U	C	A	G	
U	UUU phenylalanine	UCU serine	UAU tyrosine	UGU cysteine	U
	UUC phenylalanine	UCC serine	UAC tyrosine	UGC cysteine	C
	UUA leucine	UCA serine	UAA *stop*	UGA *stop*	A
	UUG leucine	UCG serine	UAG *stop*	UGG tryptophan	G
C	CUU leucine	CCU proline	CAU histidine	CGU arginine	U
	CUC leucine	CCC proline	CAC histidine	CGC arginine	C
	CUA leucine	CCA proline	CAA glutamine	CGA arginine	A
	CUG leucine	CCG proline	CAG glutamine	CGG arginine	G
A	AUU isoleucine	ACU threonine	AAU asparagine	AGU serine	U
	AUC isoleucine	ACC threonine	AAC asparagine	AGC serine	C
	AUA isoleucine	ACA threonine	AAA lysine	AGA arginine	A
	AUG (*start*) methionine	ACG threonine	AAG lysine	AGG arginine	G
G	GUU valine	GCU alanine	GAU aspartate	GGU glycine	U
	GUC valine	GCC alanine	GAC aspartate	GGC glycine	C
	GUA valine	GCA alanine	GAA glutamate	GGA glycine	A
	GUG valine	GCG alanine	GAG glutamate	GGG glycine	G

Figure 25.6 Messenger RNA codons.
Notice that in this chart, each of the codons (white rectangles) is composed of three letters representing the first base, second base, and third base. For example, find the rectangle where C for the first base and A for the second base intersect. You will see that U, C, A, or G can be the third base. CAU and CAC are codons for histidine; CAA and CAG are codons for glutamine.

Central Concept

The central concept of genetics can be summarized as in Figure 25.7. DNA has a sequence of bases that is transcribed into a sequence of bases in mRNA. Every three bases is a codon that stands for a particular amino acid. In this way, DNA specifies the sequence of amino acids in a protein when translation occurs at the ribosomes.

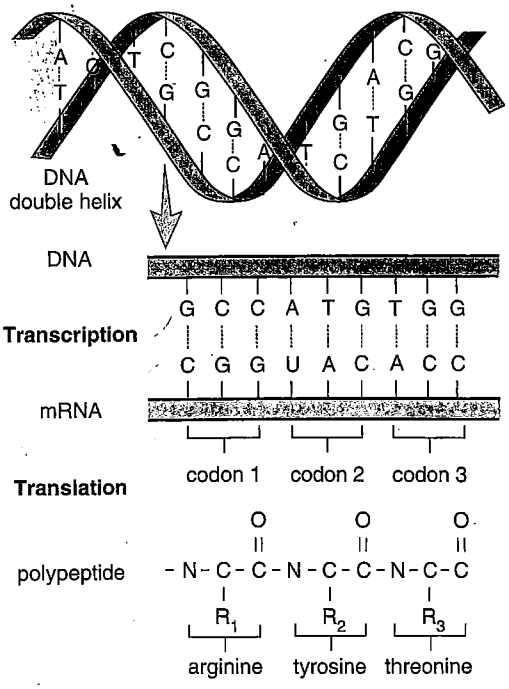


Figure 25.7 Overview of gene expression.
Transcription occurs when DNA acts as a template for mRNA synthesis. Translation occurs when the sequence of the mRNA codons determines the sequence of amino acids in a protein.

Transcription

During transcription, a segment of the DNA helix unwinds and unzips, and complementary RNA nucleotides from an RNA nucleotide pool in the nucleus pair with the DNA nucleotides of one strand. The RNA nucleotides are joined by an enzyme called **RNA polymerase,** and an RNA molecule results. Although all three classes of RNA are formed by transcription, we will focus on transcription to form mRNA (Fig. 25.8). When mRNA forms, it has a sequence of bases complementary to DNA; wherever A, T, G, or C is present in the DNA template, U, A, C, or G is incorporated into the mRNA molecule. In this way, the code is transcribed, or copied. Now mRNA has a sequence of codons that are complementary to the DNA triplet code.

Following transcription, mRNA has a sequence of bases complementary to one of the DNA strands. Now mRNA contains codons that are complementary to the DNA triplet code.

Processing of mRNA

Most genes in humans are interrupted by segments of DNA that are not part of the gene. These portions are called *introns* because they are *intra*gene segments. The other portions of the gene are called *exons* because they are ultimately *ex*pressed. Only exons result in a protein product.

When DNA is transcribed, the mRNA contains bases that are complementary to both exons and introns, but before the mRNA exits the nucleus, it is *processed*. During processing, the introns are removed, and the exons are joined to form an mRNA molecule consisting of continuous exons. This splicing of mRNA is done by ribozymes. **Ribozymes** are organic catalysts composed of RNA, not protein.

In eukaryotic cells, processing occurs in the nucleus. The newly formed mRNA is called the *primary* mRNA molecule, and the processed mRNA is called the *mature* mRNA molecule. The mature mRNA molecule passes from the cell nucleus into the cytoplasm. There it becomes associated with ribosomes.

Ordinarily, processing brings together all the exons of a gene. In some instances, however, mRNA splicing, particularly during development, brings together just some of exons and not others. This so-called alternative splicing produces different mRNA molecules (Fig. 25.9). Then the cell will go on to produce different but related proteins.

In eukaryotic cells, the primary mRNA molecule is processed; introns are removed, so that the mature mRNA molecule contains only exons. Mature RNA leaves the nucleus and becomes associated with ribosomes.

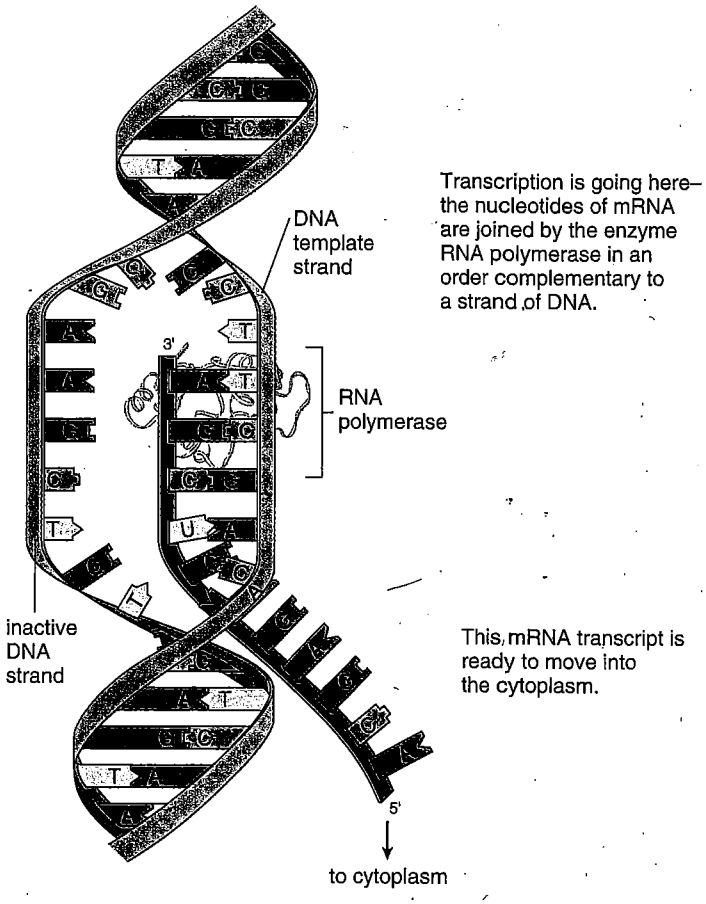


Figure 25.8 Transcription to form mRNA.
During transcription, complementary RNA is made from a DNA template. A portion of DNA unwinds and unzips at the point of attachment of RNA polymerase. A strand of mRNA is produced when complementary bases join in the order dictated by the sequence of bases in DNA. Transcription occurs in the nucleus, and the mRNA passes out of the nucleus to enter the cytoplasm.

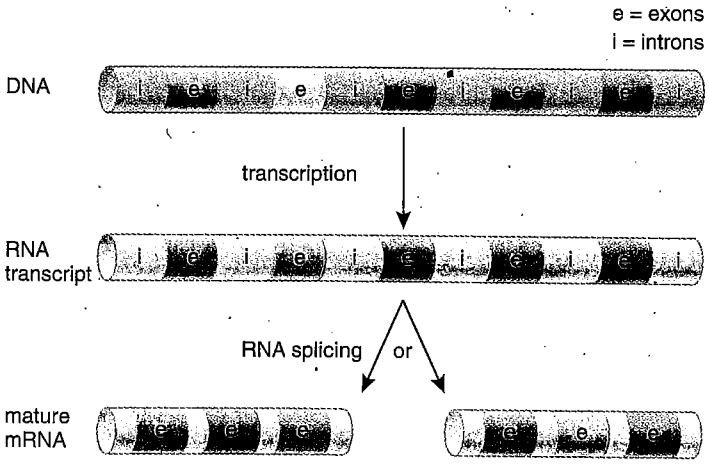


Figure 25.9 Function of introns.
Introns allow alternative splicing and therefore the production of different versions of mature mRNA from the same gene.