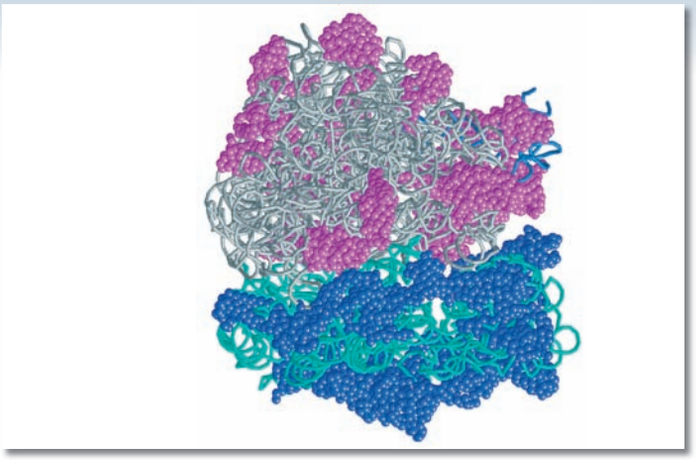
## CHAPTER OUTLINE





***A molecular model for the structure of a ribosome.*** *This is a model of ribosome structure based on X-ray crystallography. Ribosomes are needed to synthesize polypeptides, using mRNA as a template. A detailed description of this model is discussed later in Figure 13.15.*

# 13 TRANSLATION OF mRNA

The synthesis of cellular proteins occurs via the translation of the sequence of codons within mRNA into a sequence of amino acids that constitute a polypeptide. The general steps that occur in this process were already outlined in Chapter 1. In this chapter, we will explore the current state of knowledge regarding translation, with an eye toward the specific molecular interactions responsible for this process. During the past few decades, the concerted efforts of geneticists, cell biologists, and biochemists have profoundly advanced our understanding of translation. Even so, many questions remain unanswered, and this topic continues to be an exciting area of investigation.

We will begin by considering classic experiments that revealed the purpose of some genes is to encode proteins that function as enzymes. Next, we examine how the genetic code is used to decipher the information within mRNA to produce a polypeptide with a specific amino acid sequence. The rest of this chapter is devoted to an understanding of translation at the molecular level as it occurs in living cells. This will involve an examination of the cellular components—including many different proteins, RNAs, and small molecules—needed for the translation process. We will consider the structure and function of tRNA molecules, which act as the translators of the genetic information within mRNA, and then examine the composition of ribosomes. Finally, we will explore the differences between translation in bacterial cells and eukaryotic cells.

## 13.1 THE GENETIC BASIS FOR PROTEIN SYNTHESIS

Proteins are critically important as active participants in cell structure and function. The primary role of DNA is to store the information needed for the synthesis of all the proteins that an organism makes. As we discussed in Chapter 12, genes that encode an amino acid sequence are known as **structural genes.** The RNA transcribed from structural genes is called **messenger RNA (mRNA).** The main function of the genetic material is to encode the production of cellular proteins in the correct cell, at the proper time, and in suitable amounts. This is an extremely complicated task because living cells make thousands of different proteins. Genetic analyses have shown that a typical bacterium can make a few thousand different proteins, and estimates for eukaryotes range from several thousand in simple eukaryotic

organisms, such as yeast, to tens of thousands in plants and animals.

In this section, we will begin by considering early experiments that showed the role of genes is to encode proteins. We then examine the general features of the genetic code—the sequence of bases in a codon that specifies an amino acid—and explore the experiments through which the code was deciphered, or "cracked." Finally, we will look at the biochemistry of polypeptide synthesis to see how this determines the structure and function of proteins, which are ultimately responsible for the characteristics of living cells and an organism's traits.

## Archibald Garrod Proposed That Some Genes Code for the Production of a Single Enzyme

The idea that a relationship exists between genes and the production of proteins was first suggested at the beginning of the twentieth century by Archibald Garrod, a British physician. Prior to Garrod's studies, biochemists had studied many metabolic pathways within living cells. These pathways consist of a series of metabolic conversions of one molecule to another, each step catalyzed by a specific enzyme. Each enzyme is a distinctly different protein that catalyzes a particular chemical reaction. **Figure 13.1** illustrates part of the metabolic pathway for the degradation of phenylalanine, an amino acid commonly found in human diets. The enzyme phenylalanine hydroxylase catalyzes the conversion of phenylalanine to tyrosine, and a different enzyme, tyrosine aminotransferase, converts tyrosine into *p*-hydroxyphenylpyruvic acid, and so on. In all of the steps shown in Figure 13.1, a specific enzyme catalyzes a single type of chemical reaction.

Garrod studied patients who had defects in their ability to metabolize certain compounds. He was particularly interested in the inherited disease known as **alkaptonuria.** In this disorder, the patient's body accumulates abnormal levels of homogentisic acid (also called alkapton), which is excreted in the urine, causing it to appear black on exposure to air. In addition, the disease is characterized by bluish black discoloration of cartilage and skin (ochronosis). Garrod proposed that the accumulation of homogentisic acid in these patients is due to a missing enzyme, namely, homogentisic acid oxidase (see Figure 13.1).

How did Garrod realize that certain genes encode enzymes? He already knew that alkaptonuria is an inherited trait that follows an autosomal recessive pattern of inheritance. Therefore, an individual with alkaptonuria must have inherited the mutant (defective) gene that causes this disorder from both parents. From these observations, Garrod proposed that a relationship exists between the inheritance of the trait and the inheritance of a defective enzyme. Namely, if an individual inherited the mutant gene (which causes a loss of enzyme function), she or he would not produce any normal enzyme and would be unable to metabolize homogentisic acid. Garrod described alkaptonuria as an **inborn error of metabolism.** This hypothesis was the first suggestion that a connection exists between the function of genes and the production of enzymes. At the turn of the century, this idea was particularly insightful, because the structure and function of the genetic material were completely unknown.

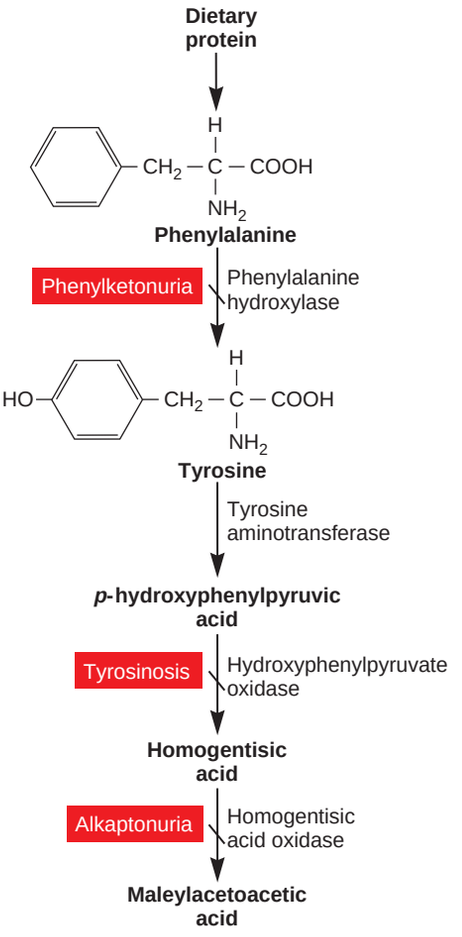



**FIGURE 13.1 The metabolic pathway of phenylalanine breakdown.** This diagram shows part of the pathway of phenylalanine metabolism, which consists of enzymes that successively convert one molecule to another. Certain human genetic diseases (shown in red boxes) are caused when enzymes in this pathway are missing or defective.

**Genes → Traits** When a person inherits two defective copies of the gene that encodes homogentisic acid oxidase, he or she cannot convert homogentisic acid into maleylacetoacetic acid. Such a person accumulates large amounts of homogentisic acid in the urine and has other symptoms of the disease known as alkaptonuria. Similarly, if a person has two defective copies of the gene encoding phenylalanine hydroxylase, he or she is unable to synthesize the enzyme phenylalanine hydroxylase and has the disease called phenylketonuria (PKU).

## Beadle and Tatum's Experiments with *Neurospora* Led Them to Propose the One-Gene/One-Enzyme Hypothesis

In the early 1940s, George Beadle and Edward Tatum were also interested in the relationship among genes, enzymes, and traits. They developed an experimental system for investigating the connection between genes and the production of particular enzymes. Consistent with the ideas of Garrod, the underlying assumption behind their approach was that a relationship exists between genes and the production of enzymes. However, the quantitative nature of this relationship was unclear. In particular, they asked the question, Does one gene control the production

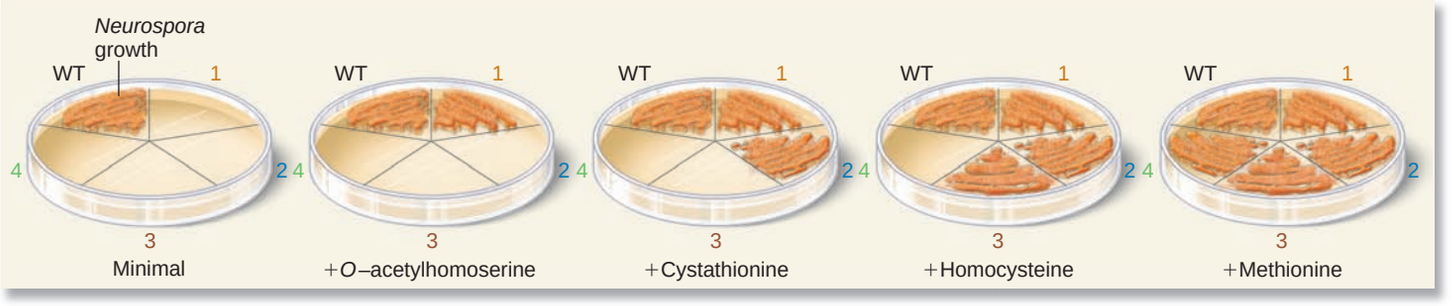
of one enzyme, or does one gene control the synthesis of many enzymes involved in a complex biochemical pathway?

At the time of their studies, many geneticists were trying to understand the nature of the gene by studying morphological traits. However, Beadle and Tatum realized that morphological traits are likely to be based on systems of biochemical reactions so complex as to make analysis exceedingly difficult. Therefore, they turned their genetic studies to the analysis of simple nutritional requirements in *Neurospora crassa*, a common bread mold. *Neurospora* can be easily grown in the laboratory and has few nutritional requirements: a carbon source (sugar), inorganic salts, and the vitamin biotin. Normal *Neurospora* cells produce many different enzymes that can synthesize the organic molecules, such as amino acids and other vitamins, which are essential for growth.

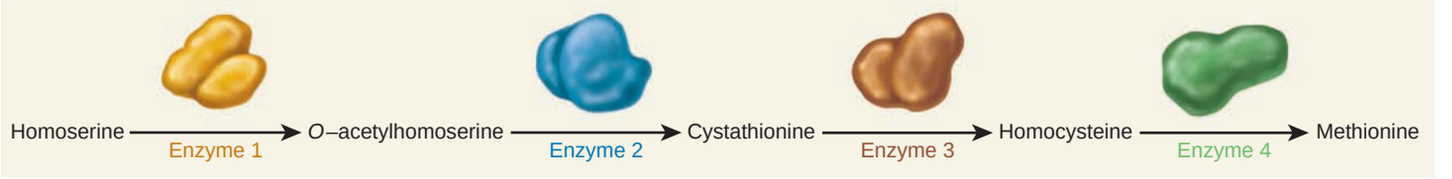
Beadle and Tatum wanted to understand how enzymes are controlled by genes. They reasoned that a mutation in a gene, causing a defect in an enzyme needed for the cellular synthesis of an essential molecule, would prevent that mutant strain from growing on minimal medium, which contains only a carbon source, inorganic salts, and biotin. In their original study of 1941, Beadle and Tatum exposed *Neurospora* cells to X-rays, which caused mutations to occur, and studied the growth of the resulting cells. By plating the cells on media with or without vitamins, they were able to identify mutant strains that required vitamins for growth. In each case, a single mutation resulted in the requirement for a single type of vitamin in the growth media.

This early study by Beadle and Tatum led to additional research to study enzymes involved with the synthesis of other substances, including the amino acid methionine. They first isolated several different mutant strains that required methionine for growth. They hypothesized that each mutant strain might be blocked at only a single step in the consecutive series of reactions that lead to methionine synthesis. To test this hypothesis, the mutant strains were examined for their ability to grow in the presence of *O*-acetylhomoserine, cystathionine, homocysteine, or methionine. *O*-Acetylhomoserine, cystathionine, and homocysteine are intermediates in the synthesis of methionine from homoserine. A simplified depiction of the results is shown in **Figure 13.2a**. The wild-type strain could grow on minimal growth media that contained the minimum set of nutrients that is required for growth. The minimal media did not contain *O*-acetylhomoserine, cystathionine, homocysteine, or methionine. Based on their growth properties, the mutant strains that had been originally identified as requiring methionine for growth could be placed into four groups designated strains 1, 2, 3, and 4 in this figure. A strain 1 mutant was missing enzyme 1, needed for the conversion of homoserine into *O*-acetylhomoserine. The cells could grow only if *O*-acetylhomoserine, cystathionine, homocysteine, or methionine was added to the growth medium. A strain 2 mutant was missing the second enzyme in this pathway that is needed for the conversion of *O*-acetylhomoserine into cystathionine and a strain 3 mutant was unable to convert cystathionine into homocysteine. Finally, a strain 4 mutant could not make methionine from homocysteine. Based on these results, the researchers could order the enzymes into a biochemical pathway as depicted in **Figure 13.2b**. Taken together, the analysis of these mutants allowed Beadle and Tatum to conclude that a single gene controlled the synthesis of a single enzyme. This was referred to as the **one-gene/one-enzyme hypothesis.**

In later decades, this hypothesis had to be modified in four ways. First, enzymes are only one category of cellular proteins. All proteins are encoded by genes, and many of them do not function as enzymes. Second, some proteins are composed of two or more different polypeptides. Therefore, it is more accurate




**(a) Growth of strains on minimal and supplemented growth media**




**(b) Simplified pathway for methionine biosynthesis**

**FIGURE 13.2 An example of an experiment that supported Beadle and Tatum's one-gene/one-enzyme hypothesis.** (a) Growth of wild-type (WT) and mutant strains on minimal media or in the presence of *O*-acetylhomoserine, cystathionine, homocysteine, or methionine. (b) A simplified pathway for methionine biosynthesis. Note: Homoserine is made by *Neurospora* via enzymes and precursor molecules not discussed in this experiment.

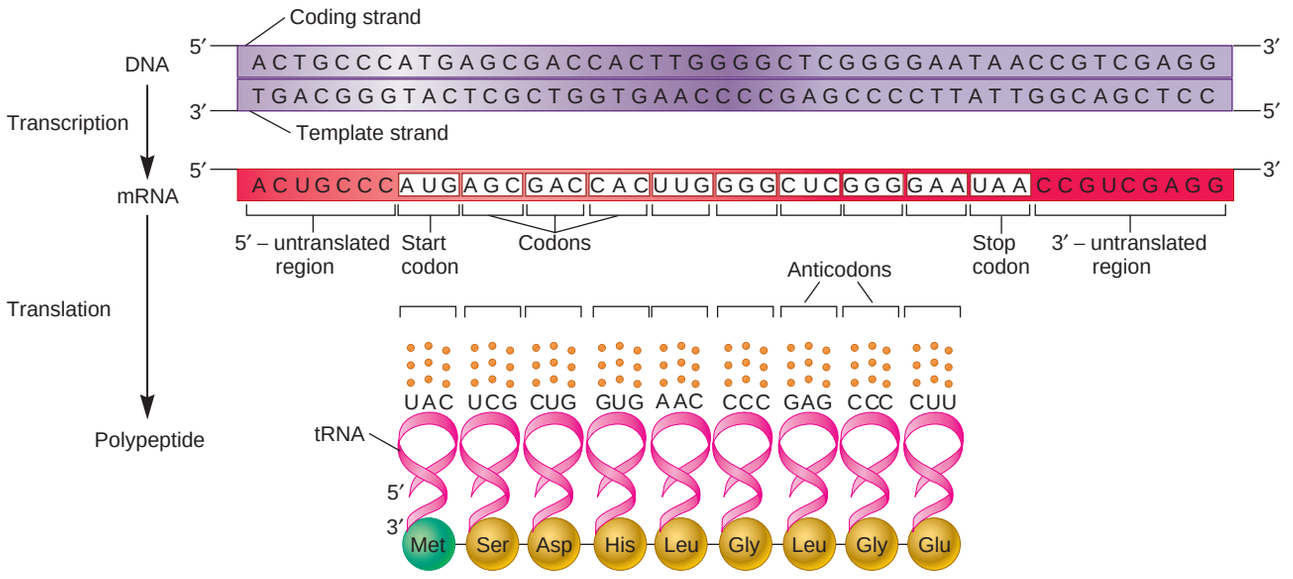



**FIGURE 13.3 The relationships among the DNA coding sequence, mRNA codons, tRNA anticodons, and amino acids in a polypeptide.** The sequence of nucleotides within DNA is transcribed to make a complementary sequence of nucleotides within mRNA. This sequence of nucleotides in mRNA is translated into a sequence of amino acids of a polypeptide. tRNA molecules act as intermediates in this translation process.

to say that a structural gene encodes a polypeptide. The term **polypeptide** refers to a structure; it is a linear sequence of amino acids. A structural gene encodes a polypeptide. By comparison, the term **protein** denotes function. Some proteins are composed of one polypeptide. In such cases, a single gene does encode a single protein. In other cases, however, a functional protein is composed of two or more different polypeptides. An example is hemoglobin, which is composed of two α-globin and two β-globin polypeptides. In this case, the expression of two genes—the α-globin and β-globin genes—is needed to create one functional protein. A third reason why the one-gene/one-polypeptide hypothesis needed revision is that we now know that many genes do not encode polypeptides. As discussed in Chapter 12, several types of genes specify functional RNA molecules that are not used to encode polypeptides (refer back to Table 12.1). Finally, as discussed in Chapter 15, one gene can encode multiple polypeptides due to alternative splicing and RNA editing.

## During Translation, the Genetic Code Within mRNA Is Used to Make a Polypeptide with a Specific Amino Acid Sequence

Let's now turn to a general description of translation. Why have researchers named this process translation? At the molecular level, **translation** involves an interpretation of one language—the language of mRNA, a nucleotide sequence—into the language of proteins—an amino acid sequence. The ability of mRNA to be translated into a specific sequence of amino acids relies on the **genetic code.** The sequence of bases within an mRNA molecule provides coded information that is read in groups of three nucleotides known as codons (**Figure 13.3**). The sequence of three bases in most codons specifies a particular amino acid. These codons are termed **sense codons.** For example, the codon AGC specifies the amino acid serine, whereas the codon GGG encodes the amino acid glycine. The codon AUG, which specifies methionine, is used as a **start codon;** it is usually the first codon that begins a polypeptide sequence. The AUG codon can also be used to specify additional methionines within the coding sequence. Finally, three codons are used to end the process of translation. These are UAA, UAG, and UGA, which are known as **stop codons.** They are also known as **termination** or **nonsense codons.**

The codons in mRNA are recognized by the anticodons in transfer RNA (tRNA) molecules (see Figure 13.3). **Anticodons** are three-nucleotide sequences that are complementary to codons in mRNA. The tRNA molecules carry the amino acids that correspond to the codons in the mRNA. In this way, the order of codons in mRNA dictates the order of amino acids within a polypeptide.

The details of the genetic code are shown in **Table 13.1.** Because polypeptides are composed of 20 different kinds of amino acids, a minimum of 20 codons is needed in order to specify each type. With four types of bases in mRNA (A, U, G, and C), a genetic code containing two bases in a codon would not be sufficient because it would only have $4^2$, or 16, possible types. By comparison, a three-base codon system can specify $4^3$, or 64, different codons. Because the number of possible codons exceeds 20—which is the number of different types of amino acids—the genetic code is termed **degenerate.** This means that more than one codon can specify the same amino acid. For example, the codons GGU, GGC, GGA, and GGG all specify the amino acid glycine. Such codons are termed **synonymous codons.** In most instances, the third base in the codon is the base that varies. The third base is sometimes referred to as the **wobble base.** This term is derived from the idea that the complementary base in the tRNA can "wobble" a bit during the recognition of the third base of the codon in mRNA. The significance of the wobble base will be discussed later in this chapter.

The start codon (AUG) defines the **reading frame** of an mRNA—a sequence of codons determined by reading bases in

**TABLE 13.1**

**The Genetic Code**

| First base | Second base: U | Second base: C | Second base: A | Second base: G | Third base |
|---|---|---|---|---|---|
| U | UUU Phenylalanine (Phe); UUC Phenylalanine (Phe); UUA Leucine (Leu); UUG Leucine (Leu) | UCU Serine (Ser); UCC Serine (Ser); UCA Serine (Ser); UCG Serine (Ser) | UAU Tyrosine (Tyr); UAC Tyrosine (Tyr); UAA Stop codon; UAG Stop codon | UGU Cysteine (Cys); UGC Cysteine (Cys); UGA Stop codon; UGG Tryptophan (Trp) | U C A G |
| C | CUU Leucine (Leu); CUC Leucine (Leu); CUA Leucine (Leu); CUG Leucine (Leu) | CCU Proline (Pro); CCC Proline (Pro); CCA Proline (Pro); CCG Proline (Pro) | CAU Histidine (His); CAC Histidine (His); CAA Glutamine (Gln); CAG Glutamine (Gln) | CGU Arginine (Arg); CGC Arginine (Arg); CGA Arginine (Arg); CGG Arginine (Arg) | U C A G |
| A | AUU Isoleucine (Ile); AUC Isoleucine (Ile); AUA Isoleucine (Ile); AUG Methionine (Met); start codon | ACU Threonine (Thr); ACC Threonine (Thr); ACA Threonine (Thr); ACG Threonine (Thr) | AAU Asparagine (Asn); AAC Asparagine (Asn); AAA Lysine (Lys); AAG Lysine (Lys) | AGU Serine (Ser); AGC Serine (Ser); AGA Arginine (Arg); AGG Arginine (Arg) | U C A G |
| G | GUU Valine (Val); GUC Valine (Val); GUA Valine (Val); GUG Valine (Val) | GCU Alanine (Ala); GCC Alanine (Ala); GCA Alanine (Ala); GCG Alanine (Ala) | GAU Aspartic acid (Asp); GAC Aspartic acid (Asp); GAA Glutamic acid (Glu); GAG Glutamic acid (Glu) | GGU Glycine (Gly); GGC Glycine (Gly); GGA Glycine (Gly); GGG Glycine (Gly) | U C A G |

groups of three, beginning with the start codon. This concept is best understood with a few examples. The mRNA sequence shown below encodes a short polypeptide with 7 amino acids:

5′–AUGCCCGGAGGCACCGUCCAAU–3′

Met–Pro–Gly–Gly–Thr–Val–Gln

If we remove one base (C) adjacent to the start codon, this changes the reading frame to produce a different polypeptide sequence:

5′–AUGCCGGAGGCACCGUCCAAU–3′

Met–Pro–Glu–Ala–Pro–Ser–Asn

Alternatively, if we remove three bases (CCC) next to the start codon, the resulting polypeptide has the same reading frame as the first polypeptide, though one amino acid (Pro, proline) has been deleted:

5′–AUGGGAGGCACCGUCCAAU–3′

Met–Gly–Gly–Thr–Val–Gln

How did researchers discover that the genetic code is read in triplets? The first evidence came from studies of Francis Crick and his colleagues in 1961. These experiments involved the isolation of mutants in a bacteriophage called T4. As described in Chapter 7, mutations in T4 genes that affect plaque morphology are easily identified (see Figure 7.16). In particular, loss-of-function mutations within certain T4 genes, designated *rII,* resulted in plaques that were larger and had a clear boundary. In comparison, wild-type phages, designated $r^+$, produced smaller plaques with a fuzzy boundary. Crick and colleagues exposed T4 phages to a chemical called proflavin, which causes single-nucleotide additions or deletions in gene sequences. The mutagenized phages were plated to identify large (*rII*) plaques. Though proflavin can cause either single-nucleotide additions or deletions, the first mutant strain that the researchers identified was arbitrarily called a (+) mutation. Many years later, when methods of DNA sequencing became available, it was determined that the (+) mutation is a single-nucleotide addition. **Table 13.2** shows a hypothetical wild-type sequence in a phage gene (first line) and considers how nucleotide additions and/or deletions could affect the resulting amino acid sequence. A single-nucleotide addition (+) would alter the reading frame beyond the point of insertion, thereby abolishing the proper function of the encoded protein. This is called a **frameshift mutation,** because it has changed the reading frame. This mutation resulted in a loss of function for the protein encoded by this viral gene and thereby produced an *rII* plaque phenotype.

The (+) mutant strain was then subjected to a second round of mutagenesis via proflavin. Several plaques were identified that had reverted to a wild-type ($r^+$) phenotype. By analyzing these strains using methods described in Chapter 7, it was determined that each one contained a second mutation that was close to the original (+) mutation. These second mutations were designated (–) mutations. Three different (–) mutations, designated a, b, and c, were identified. Each of these (–) mutations was a single-nucleotide deletion that was close to the original (+) mutation. Therefore, it restored the reading frame and produced a protein with a nearly normal amino acid sequence.

The critical experiment that suggested the genetic code is read in triplets came by combining different (–) mutations together. Mutations in different phages can be brought together into the same phage via crossing over, as described in Chapter 7. Using such an approach, the researchers constructed strains containing one, two, or three (–) mutations. The results showed that

**TABLE 13.2**

**Evidence That the Genetic Code Is Read in Triplets***

| Strain | Plaque Phenotype | DNA Coding Sequence/Polypeptide Sequence§ | Downstream Sequence‡ |
|---|---|---|---|
| Wild type | $r^+$ | ATG GGG CCC GTC CAT CCG TAC GCC GGA ATT ATA<br>Met Gly Pro Val His Pro Tyr Ala Gly Ile Ile---------- | In frame |
| (+) | *rII* | ↓A<br>ATG GGG ACC CGT CCA TCC GTA CGC CGG AAT TAT A<br>Met Gly Thr Arg Pro Ser Val Arg Arg Asn Tyr--------- | Out of frame |
| $(+)(-)_a$ | $r^+$ | ↓A ↑C<br>ATG GGG ACC GTC CAT CCG TAC GCC GGA ATT ATA<br>Met Gly Thr Val His Pro Tyr Ala Gly Ile Ile---------- | In frame |
| $(+)(-)_b$ | $r^+$ | ↓A ↑T<br>ATG GGG ACC CGC CAT CCG TAC GCC GGA ATT ATA<br>Met Gly Thr Arg His Pro Tyr Ala Gly Ile Ile---------- | In frame |
| $(+)(-)_c$ | $r^+$ | ↓A ↑G<br>ATG GGG ACC CTC CAT CCG TAC GCC GGA ATT ATA<br>Met Gly Thr Leu His Pro Tyr Ala Gly Ile Ile---------- | In frame |
| $(-)_a$ | *rII* | ↑C<br>ATG GGG CCG TCC ATC CGT ACG CCG GAA TTA TA<br>Met Gly Pro Ser Ile Arg Thr Pro Glu Leu---------- | Out of frame |
| $(-)_a(-)_b$ | *rII* | ↑C ↑T<br>ATG GGG CCG CCA TCC GTA CGC CGG AAT TAT A<br>Met Gly Pro Pro Ser Val Arg Arg Asn Tyr---------- | Out of frame |
| $(-)_a(-)_b(-)_c$ | $r^+$ | ↑↑↑GTC<br>ATG GGG CCC CAT CCG TAC GCC GGA ATT ATA<br>Met Gly Pro His Pro Tyr Ala Gly Ile Ile---------- | In frame<br>[Only Val is missing] |

*This table shows only a small portion of a hypothetical coding sequence.
§A down arrow (↓) indicates the location of a single nucleotide addition; an up arrow (↑) indicates the location of a single nucleotide deletion.
‡The term downstream sequence refers to the remaining part of the sequence that is not shown in this figure. It could include hundreds of codons. An "in-frame" sequence is wild type, whereas an "out-of-frame" sequence (caused by the addition or deletion of one or two base pairs) is not.

a wild-type plaque morphology was obtained only when three (–) mutations were combined in the same phage (see Table 13.2). The three (–) mutations in the same phage restored the normal reading frame. These results were consistent with the hypothesis that the genetic code is read in multiples of three nucleotides.

## Exceptions to the Genetic Code Are Known to Occur, Which Includes the Incorporation of Selenocysteine and Pyrrolysine into Polypeptides

From the analysis of many different species, including bacteria, protists, fungi, plants, and animals, researchers have found that the genetic code is nearly universal. However, a few exceptions to the genetic code have been noted (**Table 13.3**). The eukaryotic organelles known as mitochondria have their own DNA, which encodes a few structural genes. In mammals, the mitochondrial genetic code contains differences such as AUA = methionine and UGA = tryptophan. Also, in mitochondria and certain ciliated protists, AGA and AGG specify stop codons instead of arginine.

**Selenocysteine** (Sec) and **pyrrolysine** (Pyl) are sometimes called the 21st and 22nd amino acids in polypeptides. Their structures are shown later in Figure 13.7f. Selenocysteine is found in several enzymes involved in oxidation-reduction reactions in bacteria, archaea, and eukaryotes. Pyrrolysine is found in a few enzymes of methane-producing archaea. Selenocysteine and pyrrolysine are encoded by UGA and UAG codons, respectively, which normally function as stop codons. Like the standard 20 amino acids, selenocysteine and pyrrolysine are bound to tRNAs that specifically carry them to the ribosome for their incorporation into polypeptides. The anticodon of the tRNA that carries selenocysteine is complementary to a UGA codon, and the tRNA that carries pyrrolysine has an anticodon that is complementary to UAG.

**TABLE 13.3**
**Examples of Exceptions to the Genetic Code***

| Codon | Universal Meaning | Exception |
|---|---|---|
| AUA | Isoleucine | Methionine in yeast and mammalian mitochondria |
| UGA | Stop | Tryptophan in mammalian mitochondria |
| CUU, CUA, CUC, CUG | Leucine | Threonine in yeast mitochondria |
| AGA, AGG | Arginine | Stop codon in ciliated protozoa and in yeast and mammalian mitochondria |
| UAA, UAG | Stop | Glutamine in ciliated protozoa |
| UGA | Stop | Encodes selenocysteine in certain genes found in bacteria, archaea, and eukaryotes |
| UAC | Stop | Encodes pyrrolysine in certain genes found in methane-producing archaea |

*Several other exceptions, sporadically found among various species, are also known.

How are UGA and UAG codons "recoded" to specify the incorporation of selenocysteine or pyrrolysine, respectively? In the case of selenocysteine, a UGA codon is followed by a sequence called the selenocysteine insertion sequence (SECIS), which forms a stem-loop structure. In bacteria, a SECIS may be located immediately following the UGA codon, whereas a SECIS may be further downstream in the 3′-untranslated region of the mRNA in archaea and eukaryotes. The SECIS is recognized by proteins that favor the binding of a UGA codon to a tRNA carrying selenocysteine instead of the binding of release factors that are needed for polypeptide termination. Similarly, pyrrolysine incorporation may involve sequences downstream from a UAG codon that form a stem-loop structure.

## EXPERIMENT 13A

### Synthetic RNA Helped to Decipher the Genetic Code

Having determined that the genetic code is read in triplets, how did scientists determine the functions of the 64 codons of the genetic code? During the early 1960s, three research groups headed by Marshall Nirenberg, Severo Ochoa, and H. Gobind Khorana set out to decipher the genetic code. Though they used different methods, all of these groups used synthetic mRNA in their experimental approaches to "crack the code." We first consider the work of Nirenberg and his colleagues. Prior to their studies, several laboratories had already determined that extracts from bacterial cells, containing a mixture of components including ribosomes, tRNAs, and other factors required for translation, are able to synthesize polypeptides if mRNA and amino acids are added. This mixture is termed an in vitro, or **cell-free translation system.** If radiolabeled amino acids are added to a cell-free translation system, the synthesized polypeptides are radiolabeled and easy to detect.

To decipher the genetic code, Nirenberg and colleagues needed to gather information regarding the relationship between mRNA composition and polypeptide composition. To accomplish this goal, they made mRNA molecules of a known base composition, added them to a cell-free translation system, and then analyzed the amino acid composition of the resultant polypeptides. For example, if an mRNA molecule consisted of a string of adenine-containing nucleotides (e.g., 5′–AAAAAAAAAAAAAAA–3′), researchers could add this polyA mRNA to a cell-free translation system and ask the question, Which amino acid is specified by a codon that contains only adenine nucleotides? (As Table 13.1 shows, it is lysine.)

Before discussing the details of this type of experiment, let's consider how the synthetic mRNA molecules were made. To synthesize mRNA, an enzyme known as polynucleotide phosphorylase was used. In the presence of excess ribonucleoside diphosphates, also called nucleoside diphosphates (NDPs), this enzyme catalyzes the covalent linkage of nucleotides to make a polymer of RNA. Because it does not use a template, the order of the nucleotides is random. For example, if only uracil-containing diphosphates (UDPs) are added, then a polyU mRNA (5′–UUUUUUUUUUUUUUU–3′) is made. If nucleotides containing two different bases, such as uracil and guanine, are added, then the phosphorylase makes a random polymer containing both nucleotides (5′–GGGUGUGUGGUGGGUG–3′). An experimenter can control the amounts of the nucleotides that are added. For example, if 70% G and 30% U are mixed together with polynucleotide phosphorylase, the predicted amounts of the codons within the random polymer are as follows:

| *Codon Possibilities* | *Percentage in the Random Polymer* |
|---|---|
| GGG | $0.7 \times 0.7 \times 0.7 = 0.34 =$ 34% |
| GGU | $0.7 \times 0.7 \times 0.3 = 0.15 =$ 15% |
| GUU | $0.7 \times 0.3 \times 0.3 = 0.06 =$ 6% |
| UUU | $0.3 \times 0.3 \times 0.3 = 0.03 =$ 3% |
| UUG | $0.3 \times 0.3 \times 0.7 = 0.06 =$ 6% |
| UGG | $0.3 \times 0.7 \times 0.7 = 0.15 =$ 15% |
| UGU | $0.3 \times 0.7 \times 0.3 = 0.06 =$ 6% |
| GUG | $0.7 \times 0.3 \times 0.7 = 0.15 =$ 15% |
| | 100% |

By controlling the amounts of the NDPs in the phosphorylase reaction, the relative amounts of the possible codons can be predicted.

The first experiment that demonstrated the ability to synthesize polypeptides from synthetic mRNA was performed by Marshall Nirenberg and J. Heinrich Matthaei in 1961. As shown in **Figure 13.4**, a cell-free translation system was added to 20

different tubes. An mRNA template made via polynucleotide phosphorylase was then added to each tube. In this example, the mRNA was made using 70% G and 30% U. Next, the 20 amino acids were added to each tube, but each tube differed with regard to the type of radiolabeled amino acid. For example, radiolabeled glycine would be found in only 1 of the 20 tubes. The tubes were incubated a sufficient length of time to allow translation to occur. The newly made polypeptides were then precipitated onto a filter by treatment with trichloroacetic acid. This step precipitates polypeptides but not amino acids. A washing step caused amino acids that had not been incorporated into polypeptides to pass through the filter. Finally, the amount of radioactivity trapped on the filter was determined by liquid scintillation counting.

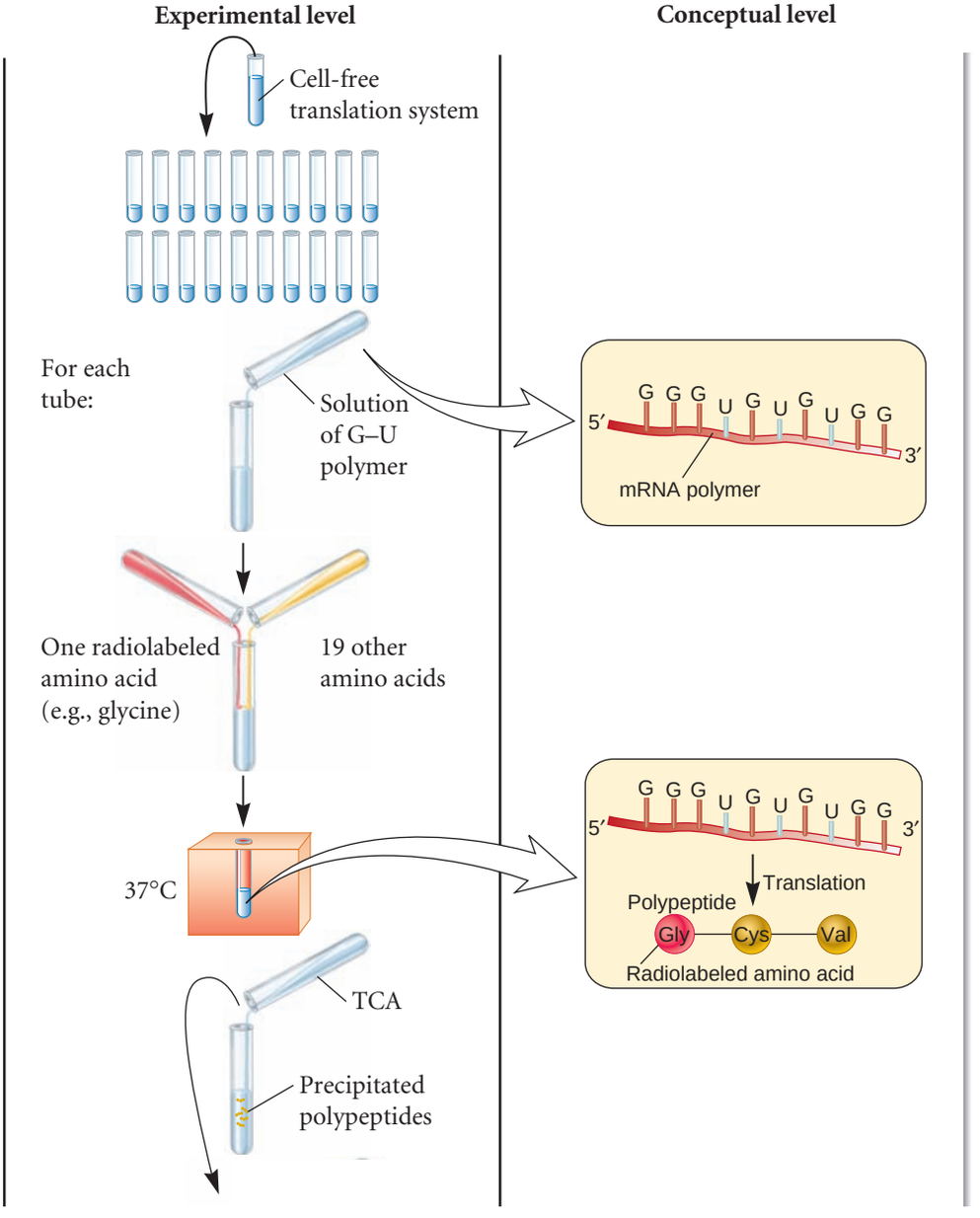
## THE GOAL

The researchers assumed that the sequence of bases in mRNA determines the incorporation of specific amino acids into a polypeptide. The purpose of this experiment was to provide information that would help to decipher the relationship between base composition and particular amino acids.

## ACHIEVING THE GOAL — FIGURE 13.4 Elucidation of the genetic code.

**Starting material:** A cell-free translation system that can synthesize polypeptides if mRNA and amino acids are added.

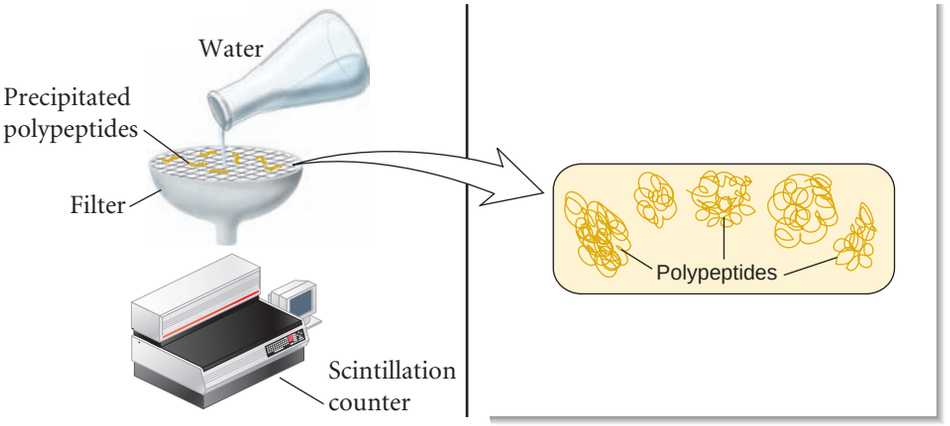
1. Add the cell-free translation system to each of 20 tubes.
2. To each tube, add random mRNA polymers of G and U made via polynucleotide phosphorylase using 70% G and 30% U.
3. Add a different radiolabeled amino acid to each tube, and add the other 19 non-radiolabeled amino acids. The translation system contained enzymes (discussed later) that attach amino acids to the appropriate tRNAs.
4. Incubate for 60 minutes to allow translation to occur.
5. Add 15% trichloroacetic acid (TCA), which precipitates polypeptides but not amino acids.



6. Place the precipitate onto a filter and wash to remove unused amino acids.

7. Count the radioactivity on the filter in a scintillation counter (see the Appendix for a description).

8. Calculate the amount of radiolabeled amino acids in the precipitated polypeptides.




## THE DATA

| *Radiolabeled Amino Acid Added* | *Relative Amount of Radiolabeled Amino Acid Incorporated into Translated Polypeptides (% of total)* |
|---|---|
| Alanine | 0 |
| Arginine | 0 |
| Asparagine | 0 |
| Aspartic acid | 0 |
| Cysteine | 6 |
| Glutamic acid | 0 |
| Glutamine | 0 |
| Glycine | 49 |
| Histidine | 0 |
| Isoleucine | 0 |
| Leucine | 6 |
| Lysine | 0 |
| Methionine | 0 |
| Phenylalanine | 3 |
| Proline | 0 |
| Serine | 0 |
| Threonine | 0 |
| Tryptophan | 15 |
| Tyrosine | 0 |
| Valine | 21 |

Adapted from Nirenberg, Marshall W., and Matthaei, J.H. (1961) The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc Natl Acad Sci USA 47*, 1588–1602.

## INTERPRETING THE DATA

According to the calculation previously described, codons should occur in the following percentages: 34% GGG, 15% GGU, 6% GUU, 3% UUU, 6% UUG, 15% UGG, 6% UGU, and 15% GUG. In the data shown in Figure 13.4, the value of 49% for glycine is due to two codons: GGG (34%) and GGU (15%). The 6% cysteine is due to UGU, and so on. It is important to realize that the genetic code was not deciphered in a single experiment such as the one described here. Furthermore, this kind of experiment yields information regarding only the nucleotide content of codons, not the specific order of bases within a single codon. For example, this experiment indicates that a cysteine codon contains two U's and one G. However, it does not tell us that a cysteine codon is UGU. Based on these data alone, a cysteine codon could be UUG, GUU, or UGU. However, by comparing many different RNA polymers, the laboratories of Nirenberg and Ochoa established patterns between the specific base sequences of codons and the amino acids they encode. In their first experiments, Nirenberg and Matthaei showed that a random polymer containing only uracil produced a polypeptide containing only phenylalanine. From this result, they inferred that UUU specifies phenylalanine. This idea is consistent with the results shown in the data table. In the random 70% G and 30% U polymer, 3% of the codons will be UUU. Likewise, 3% of the amino acids within the polypeptides were found to be phenylalanine.

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

### The Use of RNA Copolymers and the Triplet-Binding Assay Also Helped to Crack the Genetic Code

In the 1960s, H. Gobind Khorana and colleagues developed a novel method to synthesize RNA. They first created short RNA molecules, two to four nucleotides in length, that had a defined sequence. For example, RNA molecules with the sequence 5′–AUC–3′ were synthesized chemically. These short RNAs were then linked together enzymatically, in a 5′ to 3′ manner, to create long copolymers with the sequence

5′–AUCAUCAUCAUCAUCAUCAUCAUCAUC–3′

This is called a copolymer, because it is made from the linkage of several smaller molecules. Depending on the reading frame, such a copolymer would contain three different codons: AUC,

**TABLE 13.4**

**Examples of Copolymers That Were Analyzed by Khorana and Colleagues**

| Synthetic RNA* | Codon Possibilities | Amino Acids Incorporated into Polypeptides |
|---|---|---|
| UC | UCU, CUC | Serine, leucine |
| AG | AGA, GAG | Arginine, glutamic acid |
| UG | UGU, GUG | Cysteine, valine |
| AC | ACA, CAC | Threonine, histidine |
| UUC | UUC, UCU, CUU | Phenylalanine, serine, leucine |
| AAG | AAG, AGA, GAA | Lysine, arginine, glutamic acid |
| UUG | UUG, UGU, GUU | Leucine, cysteine, valine |
| CAA | CAA, AAC, ACA | Glutamine, asparagine, threonine |
| UAUC | UAU, AUC, UCU, CUA | Tyrosine, isoleucine, serine, leucine |
| UUAC | UUA, UAC, ACU, CUU | Leucine, tyrosine, threonine |

*The synthetic RNAs were linked together to make copolymers.

UCA, and CAU. Using a cell-free translation system like the one described in Figure 13.4, such a copolymer produced polypeptides containing isoleucine, serine, and histidine. **Table 13.4** summarizes some of the copolymers that were made using this approach and the amino acids that were incorporated into polypeptides.

Finally, another method that helped to decipher the genetic code also involved the chemical synthesis of short RNA molecules. In 1964, Marshall Nirenberg and Philip Leder discovered that RNA molecules containing three nucleotides—a triplet—could stimulate ribosomes to bind a tRNA. In other words, the RNA triplet acted like a codon. Ribosomes were able to bind RNA triplets, and then a tRNA with the appropriate anticodon could subsequently bind to the ribosome. To establish the relationship between triplet sequences and specific amino acids, samples containing ribosomes and a particular triplet were exposed to tRNAs with different radiolabeled amino acids.

As an example, in one experiment the researchers began with a sample of ribosomes that were mixed with 5′–CCC–3′ triplets. Portions of this sample were then added to 20 different tubes that had tRNAs with different radiolabeled amino acids. For example, one tube contained radiolabeled histidine, a second tube had radiolabeled proline, a third tube contained radiolabeled glycine, and so on. Only one radiolabeled amino acid was added to each tube. After allowing sufficient time for tRNAs to bind to the ribosomes, the samples were filtered; only the large ribosomes and anything bound to them were trapped on the filter (**Figure 13.5**). Unbound tRNAs passed through the filter. Next, the researchers determined the amount of radioactivity trapped on each filter. If the filter contained a large amount of radioactivity, the results indicated that the added triplet encoded the amino acid that was radiolabeled.

Using the triplet-binding assay, Nirenberg and Leder were able to establish relationships between particular triplet sequences and the binding of tRNAs carrying specific (radiolabeled) amino acids. In the case of the 5′–CCC–3′ triplet, they determined that tRNAs carrying radiolabeled proline were bound to the ribosomes. Unfortunately, in some cases, a triplet could not promote sufficient tRNA binding to yield unambiguous results. Nevertheless, the triplet-binding assay was an important tool in the identification of the majority of codons.

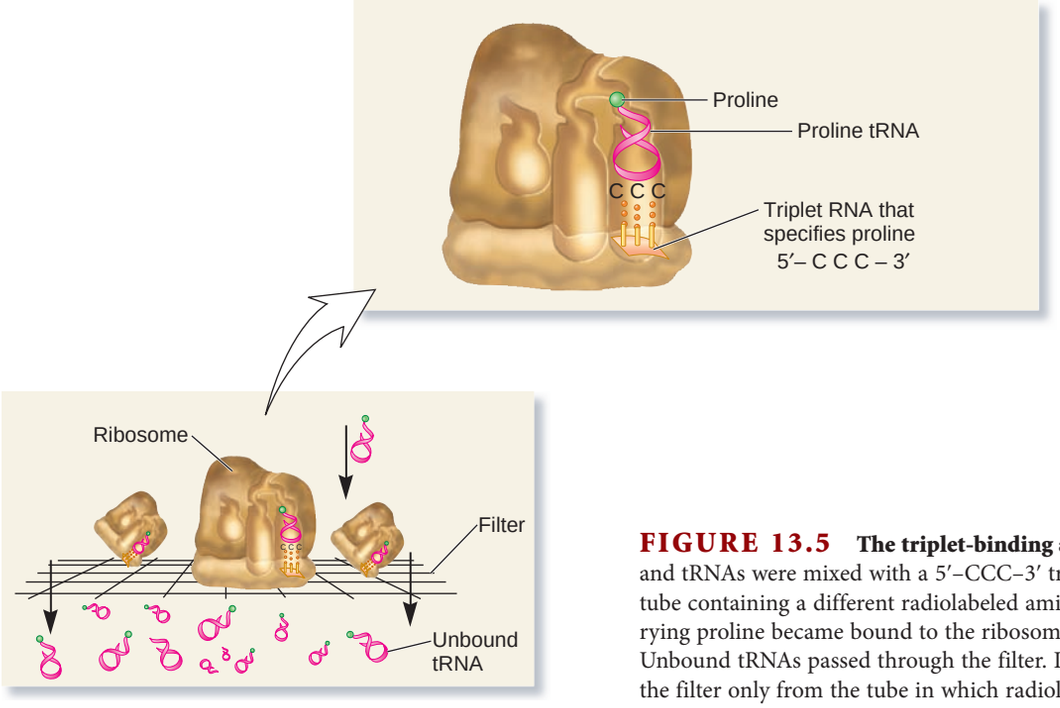
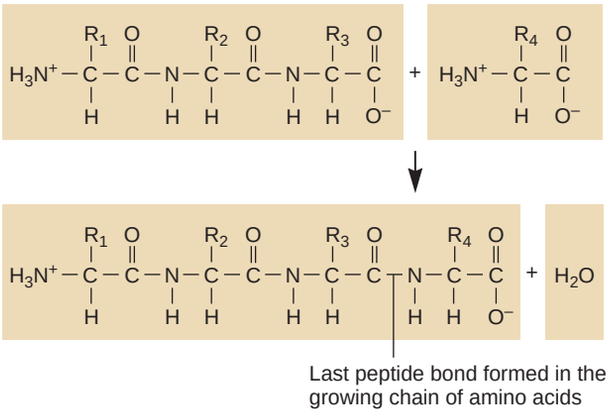
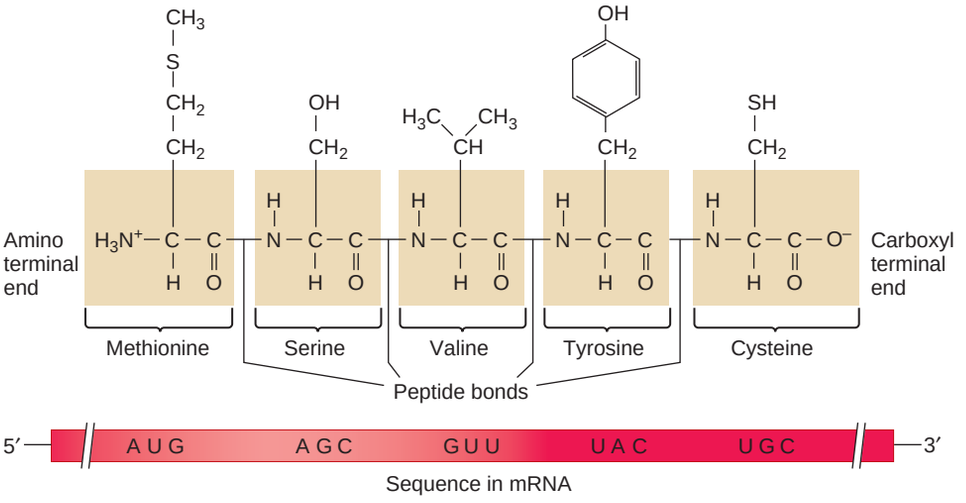



**FIGURE 13.5** **The triplet-binding assay.** In this experiment, ribosomes and tRNAs were mixed with a 5′–CCC–3′ triplet in 20 separate tubes, with each tube containing a different radiolabeled amino acid (not shown). Only tRNAs carrying proline became bound to the ribosome, which became trapped on the filter. Unbound tRNAs passed through the filter. In this case, radioactivity was trapped on the filter only from the tube in which radiolabeled proline was added.




**(a) Attachment of an amino acid to a peptide chain**




**(b) Directionality in a polypeptide and mRNA**

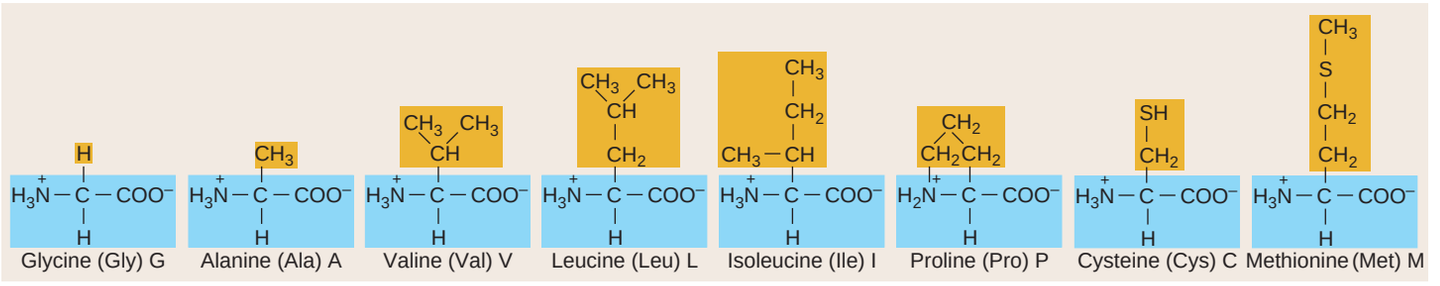
**FIGURE 13.6 The directionality of polypeptide synthesis.** (**a**) An amino acid is connected to a polypeptide chain via a condensation reaction that releases a water molecule. The letter R is a general designation for an amino acid side chain. **(b)** The first amino acid in a polypeptide chain (usually methionine) is located at the amino-terminal end, and the last amino acid is at the carboxyl-terminal end. Thus, the directionality of amino acids in a polypeptide chain is from the amino terminal-end to the carboxyl-terminal end, which corresponds to the 5′ to 3′ orientation of codons in mRNA.

## A Polypeptide Chain Has Directionality from Its Amino-Terminal to Its Carboxyl-Terminal End

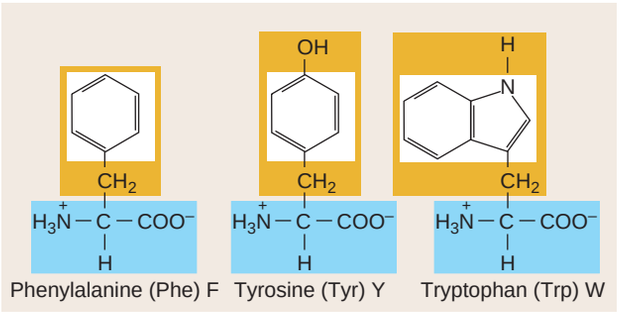
Let's now turn our attention to polypeptide biochemistry. Polypeptide synthesis has a directionality that parallels the order of codons in the mRNA. As a polypeptide is made, a **peptide bond** is formed between the carboxyl group in the last amino acid of the polypeptide chain and the amino group in the amino acid being added. As shown in **Figure 13.6a**, this occurs via a condensation reaction that releases a water molecule. The newest amino acid added to a growing polypeptide always has a free carboxyl group. **Figure 13.6b** compares the sequence of a very short polypeptide with the mRNA that encodes it. The first amino acid is said to be at the **N-terminus,** or **amino-terminal end,** of the polypeptide. An amino group ($NH_3^+$) is found at this site. The term N-terminus refers to the presence of a nitrogen atom (N) at this end. The first amino acid is specified by a codon that is near the 5′ end of the mRNA. By comparison, the last amino acid in a completed polypeptide is located at the **C-terminus,** or **carboxyl-terminal end.** A carboxyl group (COO–) is always found at this site in the polypeptide chain. This last amino acid is specified by a codon that is closer to the 3′ end of the mRNA.

## The Amino Acid Sequences of Polypeptides Determine the Structure and Function of Proteins

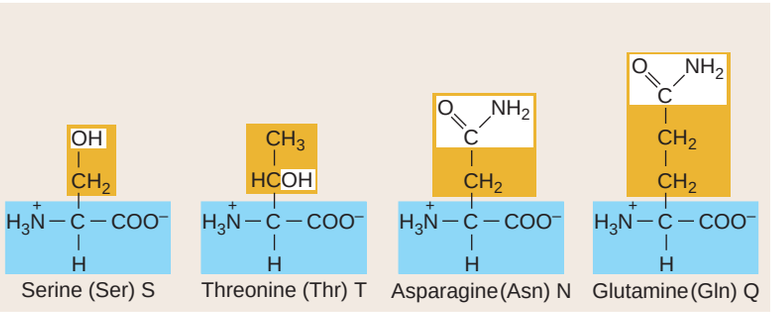
Now that we understand how mRNAs encode polypeptides, let's consider the structure and function of the gene product, namely, polypeptides. **Figure 13.7** shows the 20 different amino acids that may be found within polypeptides. Each amino acid contains a unique **side chain,** or **R group,** that has its own particular chemical properties. For example, aliphatic and aromatic amino acids are relatively nonpolar, which means they are less likely to associate with water. These hydrophobic (meaning water-fearing) amino acids are often buried within the interior of a folded protein. In contrast, the polar amino acids are hydrophilic (water-loving) and are more likely to be on the surface of a protein, where they can favorably interact with the surrounding water. The chemical properties of the amino acids and their sequences



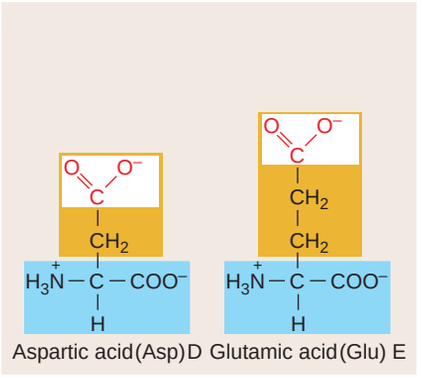

(a) Nonpolar, aliphatic amino acids



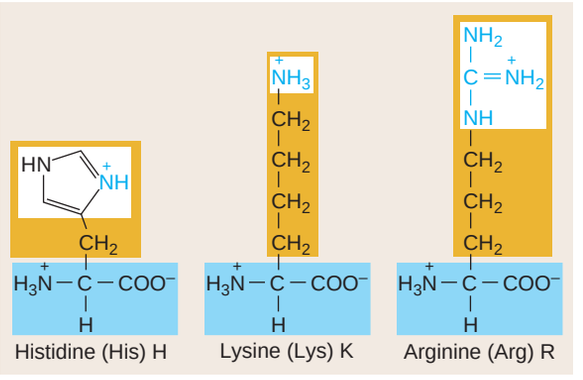

(b) Aromatic amino acids



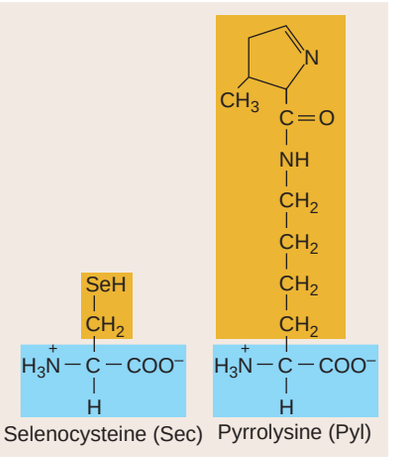

(c) Polar, neutral amino acids




(d) Polar, acidic amino acids




(e) Polar, basic amino acids




(f) Nonstandard amino acids

**FIGURE 13.7** **The amino acids that are incorporated into polypeptides during translation.** Parts **(a)** through **(e)** show the 20 standard amino acids, and part **(f)** shows two amino acids that are occasionally incorporated into polypeptides by the use of stop codons (see Table 13.3). The structures of amino acid side chains can also be covalently modified after a polypeptide is made, a phenomenon called post-translational modification.

in a polypeptide are critical factors that determine the unique structure of that polypeptide.

Following gene transcription and mRNA translation, the end result is a polypeptide with a defined amino acid sequence. This sequence is the **primary structure** of a polypeptide. **Figure 13.8** shows the primary structure of an enzyme called lysozyme, a relatively small protein containing 129 amino acids. The primary structure of a typical polypeptide may be a few hundred or even a couple of thousand amino acids in length. Within a living cell, a newly made polypeptide is not usually found in a long linear state for a significant length of time. Rather, to become a functional unit, most polypeptides quickly adopt a compact three-dimensional structure. The folding process begins while the polypeptide is still being translated. The progression from the primary structure of a polypeptide to the three-dimensional structure of a protein is dictated by the amino acid sequence within the polypeptide. In particular, the chemical properties of the amino acid side chains play a central role in determining the folding pattern of a protein. In addition, the folding of some polypeptides is aided by **chaperones**—proteins that bind to polypeptides and facilitate their proper folding.

This folding process of polypeptides is governed by the primary structure and occurs in multiple stages (**Figure 13.9**). The first stage involves the formation of a regular, repeating shape known as a **secondary structure.** The two types of secondary structures are the **α helix** and the **β sheet** (Figure 13.9b). A single polypeptide may have some regions that fold into an α helix and other regions that fold into a β sheet. Because of the geometry of

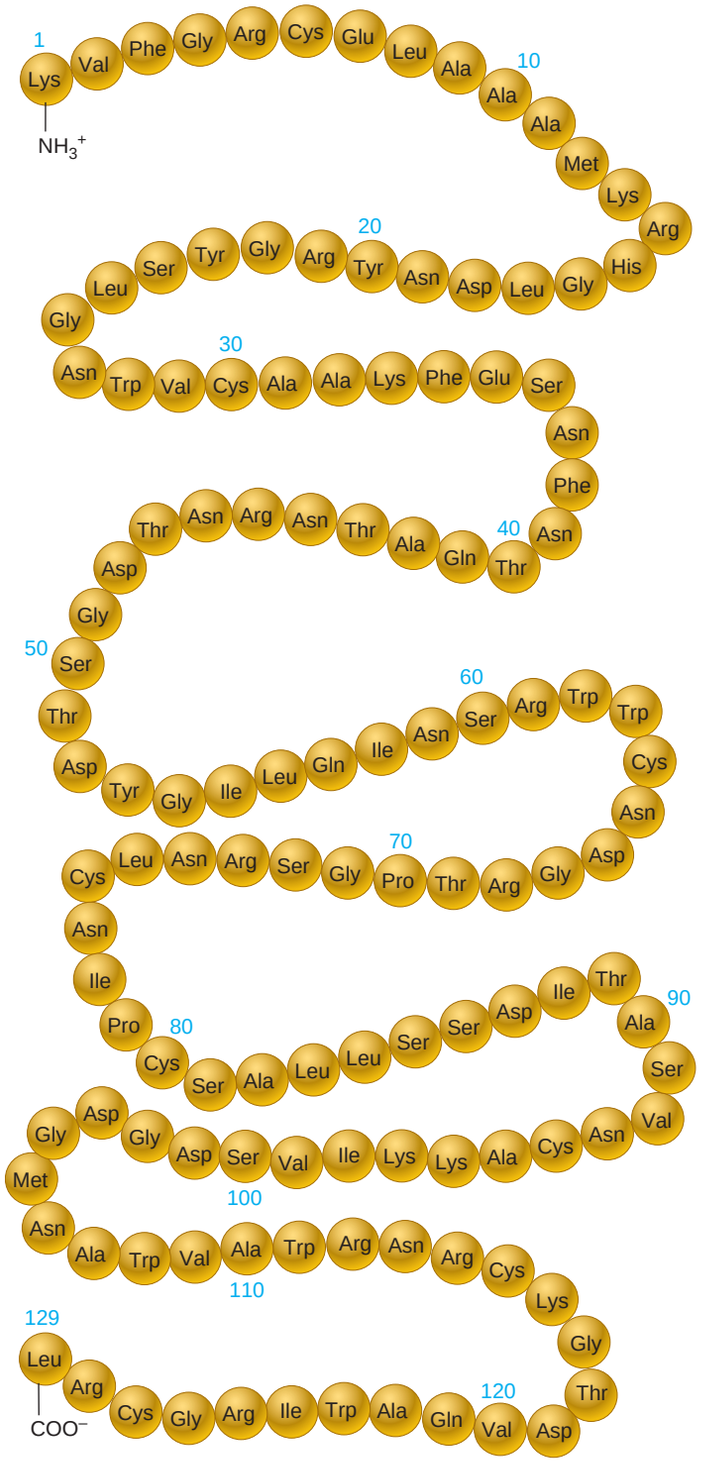



**FIGURE 13.8** **An example of a protein's primary structure.** This is the amino acid sequence of the enzyme lysozyme, which contains 129 amino acids in its primary structure. As you may have noticed, the first amino acid is not methionine; instead, it is lysine. The first methionine residue in this polypeptide sequence is removed after or during translation. The removal of the first methionine occurs in many (but not all) proteins.

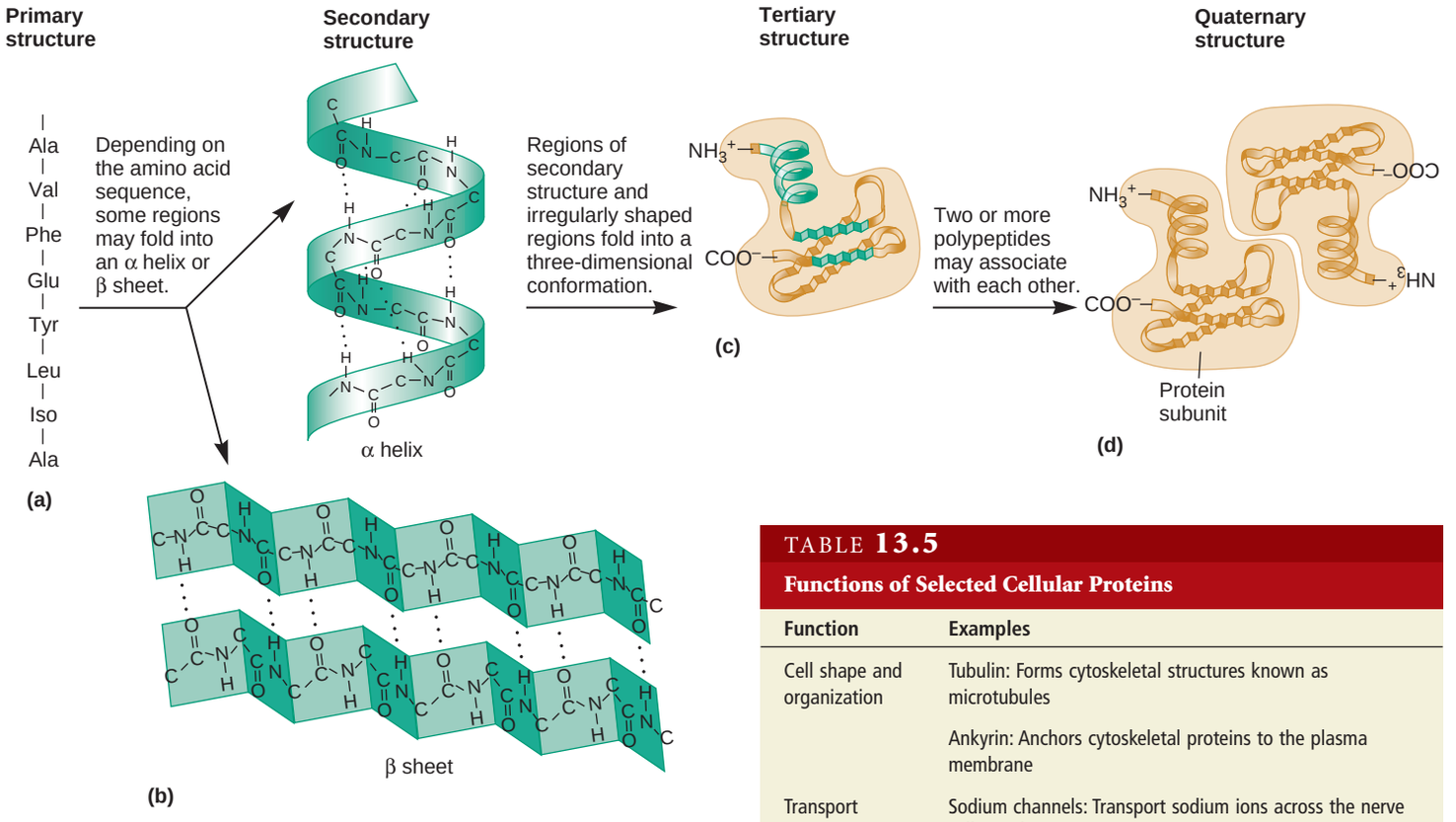
secondary structures, certain amino acids, such as glutamic acid, alanine, and methionine, are good candidates to form an α helix. Other amino acids, such as valine, isoleucine, and tyrosine, are more likely to be found in a β-sheet conformation. Secondary structures within polypeptides are primarily stabilized by the formation of hydrogen bonds between atoms that are located in the polypeptide backbone. In addition, some regions do not form a repeating secondary structure. Such regions have shapes that look very irregular in their structure because they do not follow a repeating folding pattern.

The short regions of secondary structure within a polypeptide are folded relative to each other to make the **tertiary structure** of a polypeptide. As shown in Figure 13.9c, α-helical regions and β-sheet regions are connected by irregularly shaped segments to determine the tertiary structure of the polypeptide. The folding of a polypeptide into its secondary and then tertiary conformation can usually occur spontaneously because it is a thermodynamically favorable process. The structure is determined by various interactions, including the tendency of hydrophobic amino acids to avoid water, ionic interactions among charged amino acids, hydrogen bonding among amino acids in the folded polypeptide, and weak bonding known as van der Waals interactions.

A protein is a functional unit that can be composed of one or more polypeptides. Some proteins are composed of a single polypeptide. Many proteins, however, are composed of two or more polypeptides that associate with each other to make a functional protein with a **quaternary structure** (Figure 13.9d). The individual polypeptides are called **subunits** of the protein, each of which has its own tertiary structure. The association of multiple subunits is the quaternary structure of a protein.

## Cellular Proteins Are Primarily Responsible for the Characteristics of Living Cells and an Organism's Traits

Why is the genetic material largely devoted to storing the information to make proteins? To a great extent, the characteristics of a cell depend on the types of proteins that it makes. In turn, the traits of multicellular organisms are determined by the properties of their cells. Proteins perform a variety of functions critical to the life of cells and to the morphology and function of organisms (**Table 13.5**). Some proteins are important in determining the shape and structure of a given cell. For example, the protein tubulin assembles into large cytoskeletal structures known as microtubules, which provide eukaryotic cells with internal structure and organization. Some proteins are inserted into the cell membrane and aid in the transport of ions and small molecules across the membrane. An example is a sodium channel that transports sodium ions into nerve cells. Another interesting category of proteins are those that function as biological motors, such as myosin, which is involved in the contractile properties of muscle cells. Within multicellular organisms, certain proteins function in cell signaling and cell surface recognition. For example, proteins, such as the hormone insulin, are secreted by endocrine cells and




ONLINE ANIMATION

**FIGURE 13.9 Levels of structures formed in proteins.** **(a)** The primary structure of a polypeptide within a protein is its amino acid sequence. **(b)** Certain regions of a primary structure will fold into a secondary structure; the two types of secondary structures are called α helices and β sheets. **(c)** Both of these secondary structures can be found within the tertiary structure of a polypeptide. **(d)** Some polypeptides associate with each other to form a protein with a quaternary structure.

bind to the insulin receptor proteins found within the plasma membrane of target cells.

Many proteins are **enzymes,** which function to accelerate chemical reactions within the cell. Some enzymes assist in the breakdown of molecules or macromolecules into smaller units. These are known as catabolic enzymes and are important in utilizing cellular energy. In contrast, anabolic enzymes function in the synthesis of molecules and macromolecules. Several anabolic enzymes are listed in Table 13.5, including DNA polymerase, which is required for the synthesis of DNA from nucleotide building blocks. Throughout the cell, the synthesis of molecules and macromolecules relies on enzymes and accessory proteins. Ultimately, then, the construction of a cell greatly depends on its anabolic enzymes because these are required to synthesize all cellular macromolecules.

**TABLE 13.5**
**Functions of Selected Cellular Proteins**

| Function | Examples |
|---|---|
| Cell shape and organization | Tubulin: Forms cytoskeletal structures known as microtubules |
| | Ankyrin: Anchors cytoskeletal proteins to the plasma membrane |
| Transport | Sodium channels: Transport sodium ions across the nerve cell membrane |
| | Lactose permease: Transports lactose across the bacterial cell membrane |
| | Hemoglobin: Transports oxygen in red blood cells |
| Movement | Myosin: Involved in muscle cell contraction |
| | Kinesin: Involved in the movement of chromosomes during cell division |
| Cell signaling | Insulin: A hormone that influences target cell metabolism and growth |
| | Epidermal growth factor: A growth factor that promotes cell division |
| | Insulin receptor: Recognizes insulin and initiates a cell response |
| Cell surface recognition | Integrins: Bind to large extracellular proteins |
| Enzymes | Hexokinase: Phosphorylates glucose during the first step in glycolysis |
| | β-Galactosidase: Cleaves lactose into glucose and galactose |
| | Glycogen synthetase: Uses glucose molecules as building blocks to synthesize a large carbohydrate known as glycogen |
| | Acyl transferase: Links together fatty acids and glycerol phosphate during the synthesis of phospholipids |
| | RNA polymerase: Uses ribonucleotides as building blocks to synthesize RNA |
| | DNA polymerase: Uses deoxyribonucleotides as building blocks to synthesize DNA |

## 13.2 STRUCTURE AND FUNCTION OF tRNA

Thus far, we have considered the general features of translation and surveyed the structure and functional significance of cellular proteins. The rest of this chapter is devoted to a molecular understanding of translation as it occurs in living cells. Biochemical studies of protein synthesis and tRNA molecules began in the 1950s. As work progressed toward an understanding of translation, research revealed that different kinds of RNA molecules are involved in the incorporation of amino acids into growing polypeptides. Francis Crick proposed the **adaptor hypothesis.** According to this idea, the position of an amino acid within a polypeptide chain is determined by the binding between the mRNA and an adaptor molecule carrying a specific amino acid. Later, work by Paul Zamecnik and Mahlon Hoagland suggested that the adaptor molecule is tRNA. During translation, a tRNA has two functions: (1) It recognizes a three-base codon sequence in mRNA, and (2) it carries an amino acid specific for that codon. In this section, we will examine the general function of tRNA molecules. We begin by considering an experiment that was critical in supporting the adaptor hypothesis and then explore some of the important structural features that underlie tRNA function.

### The Function of a tRNA Depends on the Specificity Between the Amino Acid It Carries and Its Anticodon

The adaptor hypothesis proposes that tRNA molecules recognize the codons within mRNA and carry the correct amino acids to the site of polypeptide synthesis. During mRNA-tRNA recognition, the anticodon in a tRNA molecule binds to a codon in mRNA due to their complementary sequences (**Figure 13.10**). Importantly, the anticodon in the tRNA corresponds to the amino acid that it carries. For example, if the anticodon in the tRNA is 3′–AAG–5′, it is complementary to a 5′–UUC–3′ codon. According to the genetic code, described earlier in this chapter, the UUC codon specifies phenylalanine. Therefore, the tRNA with a 3′–AAG–5′ anticodon must carry a phenylalanine. As another example, if the tRNA has a 3′–GGC–5′ anticodon, it is complementary to a 5′–CCG–3′ codon that specifies proline. This tRNA must carry proline.

Recall that the genetic code has 64 codons. Of these, 61 are sense codons that specify the 20 amino acids. Therefore, to synthesize proteins, a cell must produce many different tRNA molecules having specific anticodon sequences. To do so, the chromosomal DNA contains many distinct tRNA genes that encode tRNA molecules with different sequences. According to the adaptor hypothesis, the anticodon in a tRNA specifies the type of amino acid that it carries. Due to this specificity, tRNA molecules are named according to the type of amino acid they carry. For example, a tRNA that attaches to phenylalanine is described as $tRNA^{Phe}$, whereas a tRNA that carries proline is $tRNA^{Pro}$.

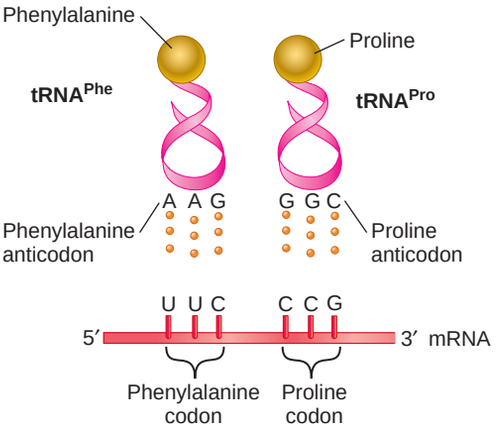



**FIGURE 13.10** **Recognition between tRNAs and mRNA.** The anticodon in the tRNA binds to a complementary sequence in the mRNA. At its other end, the tRNA carries the amino acid that corresponds to the codon in the mRNA via the genetic code.

## EXPERIMENT 13B

### tRNA Functions as the Adaptor Molecule Involved in Codon Recognition

In 1962, François Chapeville and his colleagues conducted experiments aimed at testing the adaptor hypothesis. Their technical strategy was similar to that of the Nirenberg experiments that helped to decipher the genetic code (see Experiment 13A). In this approach, a cell-free translation system was made from cell extracts that contained the components necessary for translation. These components include ribosomes, tRNAs, and other translation factors. A cell-free translation system can synthesize polypeptides in vitro if mRNA and amino acids are added. Such a translation system can be used to investigate the role of specific factors by adding a particular mRNA template and varying individual components required for translation.

According to the adaptor hypothesis, the amino acid attached to a tRNA is not directly involved in codon recognition. Chapeville reasoned that if this were true, the alteration of an amino acid already attached to a tRNA should cause that altered amino acid to be incorporated into the polypeptide instead of the normal amino acid. For example, consider a $tRNA^{Cys}$ that carries the amino acid cysteine. If the attached cysteine were changed to an alanine, this $tRNA^{Cys}$ should insert an alanine into a polypeptide where it would normally put a cysteine. Fortunately, Chapeville could carry out this strategy because he had a reagent, known as Raney nickel, that can chemically convert cysteine to alanine.

A key aspect of the experimental design was the choice of the mRNA template. Chapeville and his colleagues synthesized an mRNA template that contained only U and G. Therefore, this

template contained only the following codons (refer back to the genetic code in Table 13.1):

| | |
|---|---|
| UUU = phenylalanine | GGU = glycine |
| UUG = leucine | GUU = valine |
| UGG = tryptophan | GUG = valine |
| GGG = glycine | UGU = cysteine |

Among the eight possible codons, one cysteine codon occurs, but no alanine codons can be formed from a polyUG template.

As shown in the experiment of **Figure 13.11**, Chapeville began with a cell-free translation system that contained tRNA molecules. Amino acids, which would become attached to tRNAs, were added to this mixture. Of the 20 amino acids, only cysteine was radiolabeled. After allowing sufficient time for the amino acids to become attached to the correct tRNAs, the sample was divided into two tubes. One tube was treated with Raney nickel, whereas the control tube was not. As mentioned, Raney nickel converts cysteine into alanine by removing the –SH (sulfhydryl) group. However, it did not remove the radiolabel, which was a $^{14}$C-label within the cysteine amino acid. Next, the polyUG mRNA was added as a template, and the samples were incubated to allow the translation of the mRNA into a polypeptide. In the control tube, we would expect the polypeptide to contain phenylalanine, leucine, tryptophan, glycine, valine, and cysteine, because these are the codons that contain only U and G. However, in the Raney nickel-treated sample, if the $tRNA^{Cys}$ was using its anticodon region to recognize the mRNA, we would expect to see alanine instead of cysteine.

Following translation, the polypeptides were isolated and hydrolyzed via a strong acid treatment, and then the individual amino acids were separated by column chromatography. The column separated cysteine from alanine; alanine eluted in a later fraction. The amount of radioactivity in each fraction was determined by liquid scintillation counting.

## THE HYPOTHESIS

Codon recognition is dictated only by the tRNA anticodon; the chemical structure of the amino acid attached to the tRNA does not play a role.

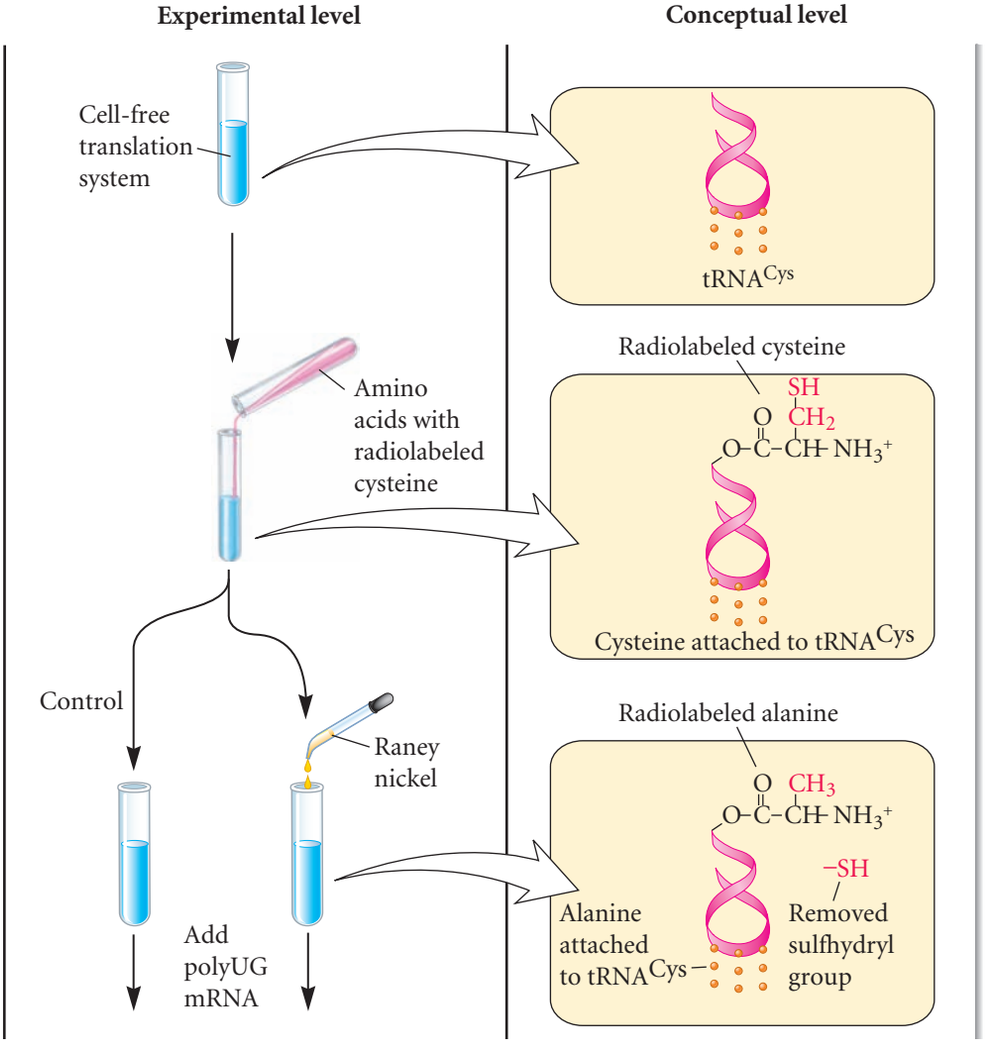
## TESTING THE HYPOTHESIS — FIGURE 13.11 Evidence that tRNA uses its anticodon sequence to recognize mRNA.

**Starting material:** A cell-free translation system that can synthesize polypeptides if mRNA and amino acids are added.

1. Place cell-free translation system into a tube. Note: This drawing emphasizes only $tRNA^{Cys}$, even though the cell-free translation system contains all types of tRNAs, and other components, such as ribosomes. In the translation system, a substantial proportion of the tRNAs do not have an attached amino acid. The translation system also contains enzymes that attach amino acids to tRNAs. (These enzymes will be described later in the chapter.)

2. Add amino acids, including radiolabeled cysteine. An enzyme within the translation system will specifically attach the radiolabeled cysteine to $tRNA^{Cys}$. The other tRNAs will have unlabeled amino acids attached to them. Incubate and divide into two tubes.

3. In one tube, treat the tRNAs with Raney nickel. This removes the –SH group from cysteine, converting it to alanine. In the control tube, do not add Raney nickel.



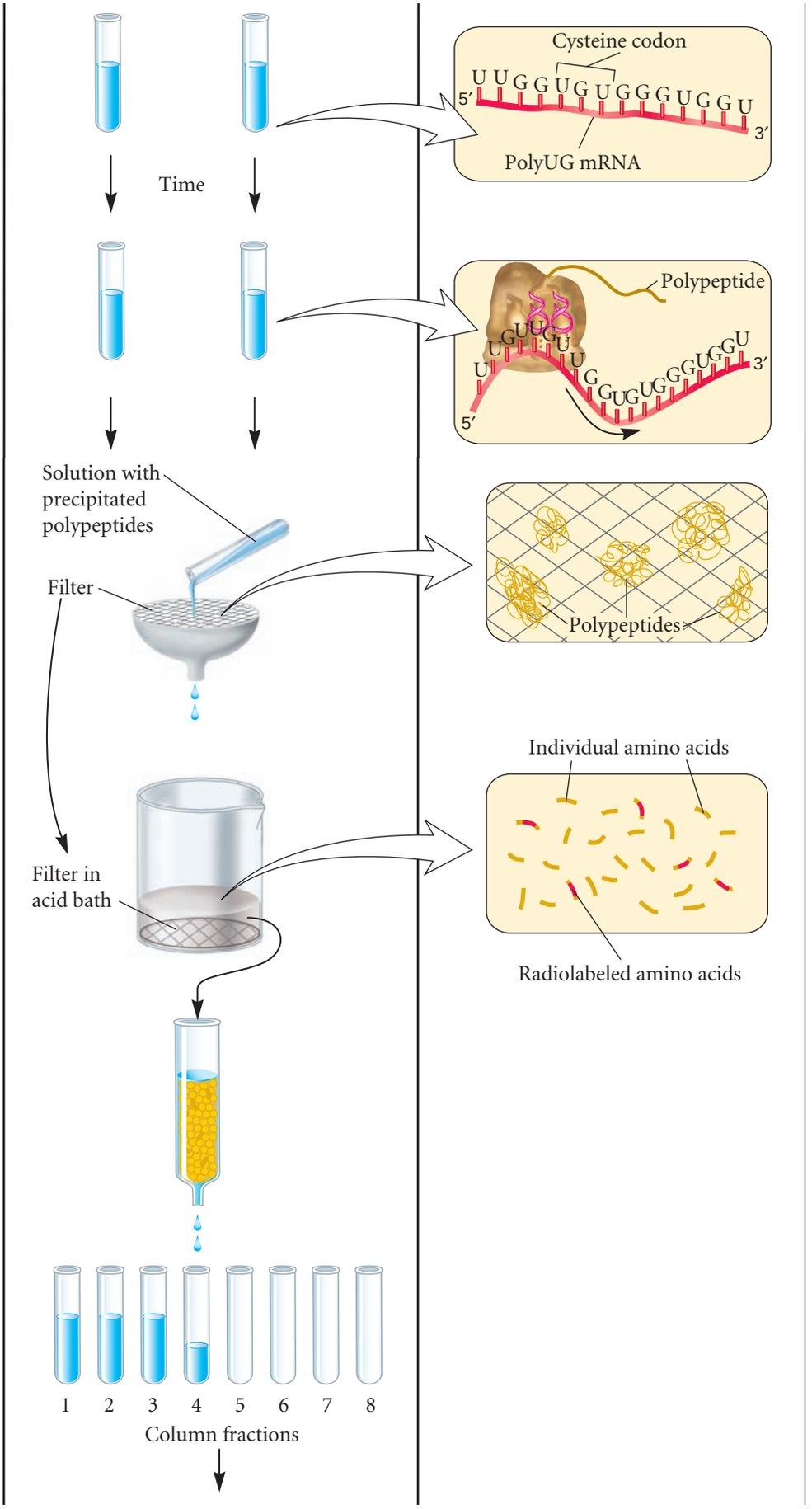
4. Add polyUG mRNA made via polynucleotide phosphorylase as a template. A polyUG mRNA contains cysteine codons but no alanine codons.

5. Allow translation to proceed.

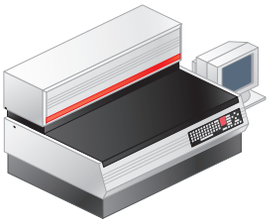
6. Precipitate the newly made polypeptides with trichloroacetic acid and then isolate the polypeptides on a filter.

7. Hydrolyze the polypeptides to their individual amino acids by treatment with a solution containing concentrated hydrochloric acid.

8. Run the sample over a column that separates cysteine and alanine. (See the Appendix for a description of column chromatography.) Separate into fractions. Note: Cysteine runs through the column more quickly and comes out in fraction 3. Alanine comes out later, in fraction 7.



9. Determine the amount of radioactivity in the fractions that contain alanine and cysteine.



## THE DATA

| Conditions | Amount of Radiolabeled Amino Acids Incorporated into Polypeptide (cpm)* | | |
|---|---|---|---|
| | *Cysteine* | *Alanine* | *Total* |
| Control, untreated tRNA | 2835 | 83 | 2918 |
| Raney nickel-treated tRNA | 990 | 2020 | 3010 |

*cpm is the counts per minute of radioactivity in the sample.
Adapted from Chapeville F., Lipmann F., von Ehrenstein G., et al. (1962) On the role of soluble ribonucleic acid in coding of amino acids. *Proc Natl Acad Sci USA 48*, 1086–1092.

## INTERPRETING THE DATA

In the control sample, nearly all of the radioactivity was found in the fraction containing cysteine. This result was expected because the only radiolabeled amino acid attached to tRNAs was cysteine. The low radioactivity in the alanine fraction (83 counts per minute [cpm]) probably represents contamination of this fraction by a small amount of cysteine. By comparison, when the tRNAs were treated with Raney nickel, a substantial amount of radiolabeled alanine became incorporated into polypeptides. This occurred even though the mRNA template did not contain any alanine codons. How do we explain these results? They are consistent with the explanation that a $tRNA^{Cys}$, which carried alanine instead of cysteine, incorporated alanine into the synthesized polypeptide. These observations indicate that the codons in mRNA are identified directly by the tRNA rather than the attached amino acid.

As seen in the data of Figure 13.11, the Raney nickel-treated sample still had 990 cpm of cysteine incorporated into polypeptides. This is about one-third of the total amount of radioactivity (namely, 990/3010). In other experiments conducted in this study, the researchers showed that the Raney nickel did not react with about one-third of the $tRNA^{Cys}$. Therefore, this proportion of the Raney nickel-treated $tRNA^{Cys}$ would still carry cysteine. This observation was consistent with the data shown here. Overall, the results of this experiment supported the adaptor hypothesis, indicating that tRNAs act as adaptors to carry the correct amino acid to the ribosome based on their anticodon sequence.

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

## Common Structural Features Are Shared by All tRNAs

To understand how tRNAs act as carriers of the correct amino acids during translation, researchers have examined the structural characteristics of these molecules in great detail. Though a cell makes many different tRNAs, all tRNAs share common structural features. As originally proposed by Robert W. Holley in 1965, the secondary structure of tRNAs exhibits a cloverleaf pattern. A tRNA has three stem-loop structures, a few variable sites, and an acceptor stem with a 3′ single-stranded region (**Figure 13.12**). The acceptor stem is where an amino acid becomes attached to a tRNA (see inset). A conventional numbering system for the nucleotides within a tRNA molecule begins at the 5′ end and proceeds toward the 3′ end. Among different types of tRNA molecules, the variable sites (shown in blue) can differ in the number of nucleotides they contain. The anticodon is located in the second loop region.

The actual three-dimensional, or tertiary, structure of tRNA molecules involves additional folding of the secondary structure. In the tertiary structure of tRNA, the stem-loop regions are folded into a much more compact molecule. The ability of RNA molecules to form stem-loop structures and the tertiary folding of tRNA molecules are described in Chapter 9 (see Figure 9.23). Interestingly, in addition to the normal A, U, G, and C nucleotides, tRNA molecules commonly contain modified nucleotides within their primary structures. For example, Figure 13.12 illustrates a tRNA that contains several modified bases. Among many different species, researchers have found that more than 80 different nucleotide modifications can occur in tRNA molecules. We will explore the significance of modified bases in codon recognition later in this chapter.

## Aminoacyl-tRNA Synthetases Charge tRNAs by Attaching the Appropriate Amino Acid

To function correctly, each type of tRNA must have the appropriate amino acid attached to its 3′ end. How does an amino acid get attached to a tRNA with the correct anticodon? Enzymes

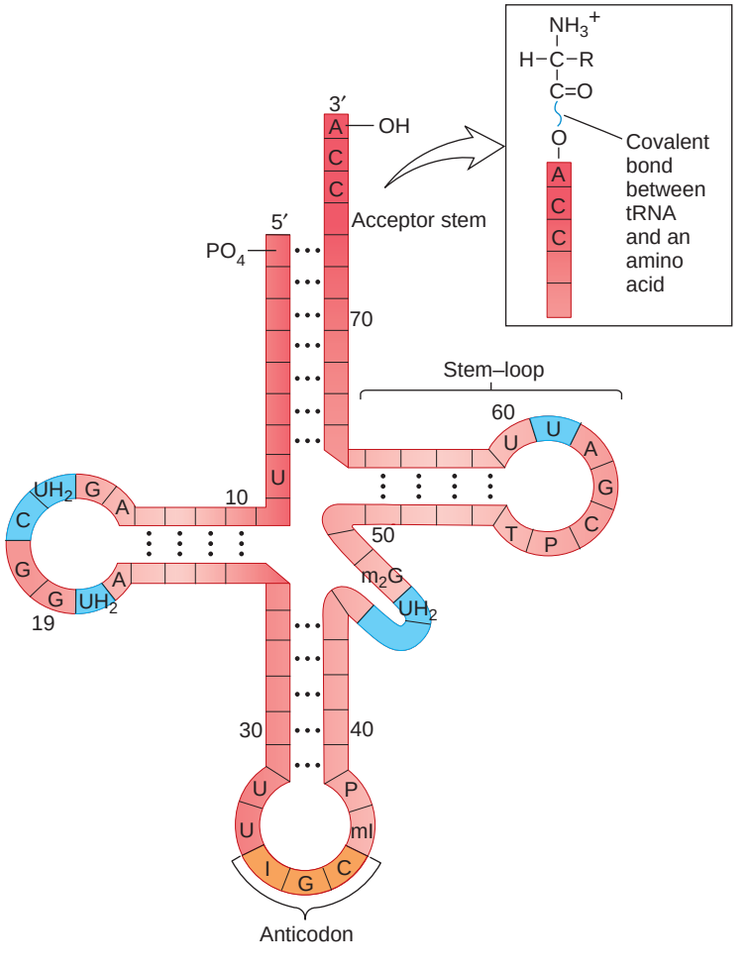
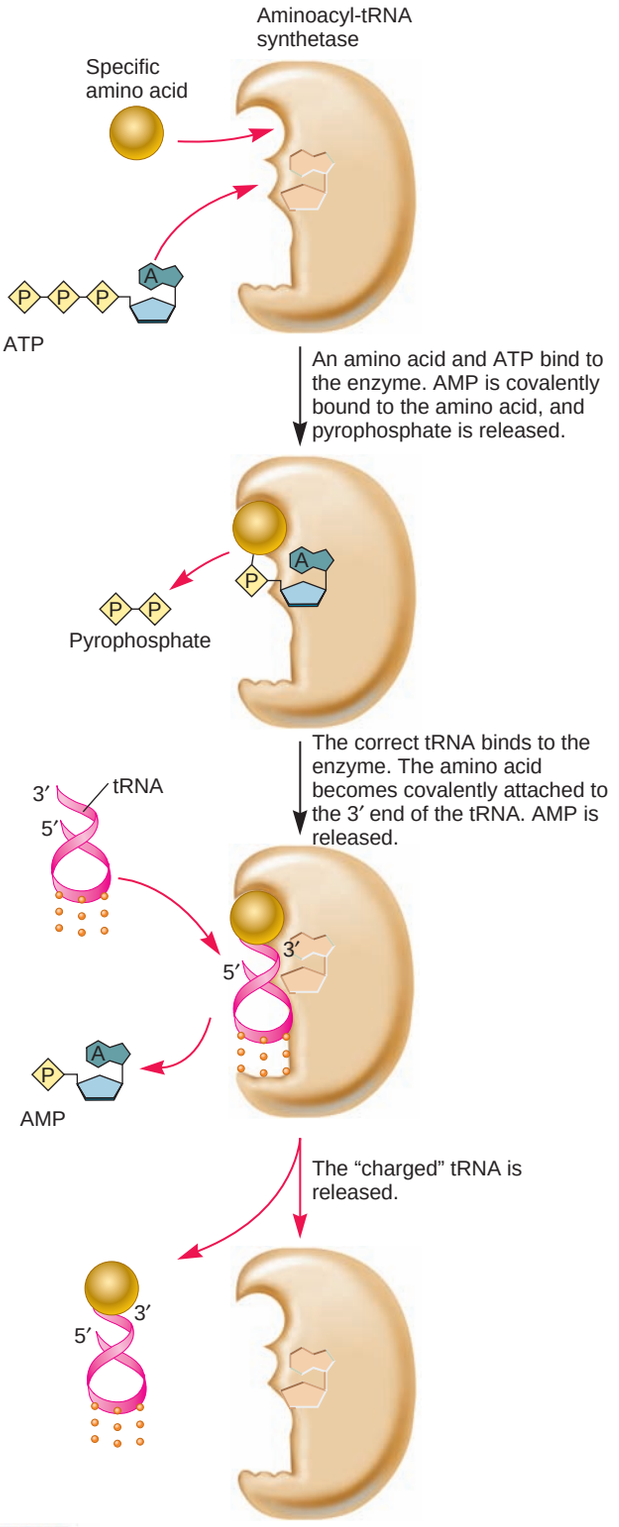



**FIGURE 13.12** **Secondary structure of tRNA.** The conventional numbering of nucleotides begins at the 5′ end and proceeds toward the 3′ end. In all tRNAs, the nucleotides at the 3′ end contain the sequence CCA. Certain locations can have additional nucleotides not found in all tRNA molecules. These variable sites are shown in blue. The figure also shows the locations of a few modified bases specifically found in a yeast tRNA that carries alanine. The modified bases are as follows: I = inosine, mI = methylinosine, T = ribothymidine, $UH_2$ = dihydrouridine, $m_2G$ = dimethylguanosine, and P = pseudouridine. The inset shows an amino acid covalently attached to the 3′ end of a tRNA.

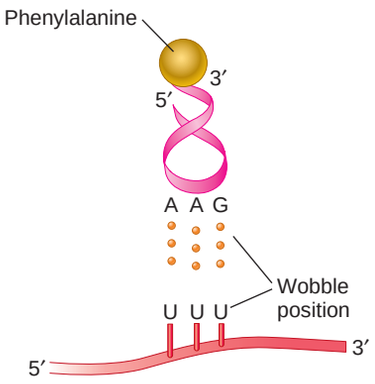
in the cell known as **aminoacyl-tRNA synthetases** catalyze the attachment of amino acids to tRNA molecules. Cells produce 20 different aminoacyl-tRNA synthetase enzymes, 1 for each of the 20 distinct amino acids. Each aminoacyl-tRNA synthetase is named for the specific amino acid it attaches to tRNA. For example, alanyl-tRNA synthetase recognizes a tRNA with an alanine anticodon—$tRNA^{Ala}$—and attaches an alanine to it.

Aminoacyl-tRNA synthetases catalyze a chemical reaction involving three different molecules: an amino acid, a tRNA molecule, and ATP. In the first step of the reaction, a synthetase recognizes a specific amino acid and also ATP (**Figure 13.13**). The ATP is hydrolyzed, and AMP becomes attached to the amino acid; pyrophosphate is released. During the second step, the correct tRNA binds to the synthetase. The amino acid becomes covalently attached to the 3′ end of the tRNA molecule at the




ONLINE ANIMATION

**FIGURE 13.13** **Catalytic function of aminoacyl-tRNA synthetase.** Aminoacyl-tRNA synthetase has binding sites for a specific amino acid, ATP, and a particular tRNA. In the first step, the enzyme catalyzes the covalent attachment of AMP to an amino acid, yielding an activated amino acid. In the second step, the activated amino acid is attached to the appropriate tRNA.




(a) Location of wobble position

| Nucleotide of tRNA anticodon | Third nucleotide of mRNA codon |
|---|---|
| G | C, U |
| C | G |
| A | U, C, G, (A) |
| U | A, U, G, (C) |
| I | U, C, A |
| $xm^5s^2U$<br>$xm^5Um$<br>Um<br>$xm^5U$ | A, (G) |
| $xo^5U$ | U, A, G |
| $k^2C$ | A |

(b) Revised wobble rules

**FIGURE 13.14 Wobble position and base-pairing rules. (a)** The wobble position occurs between the first base (meaning the first base in the 5′ to 3′ direction) in the anticodon and the third base in the mRNA codon. **(b)** The revised wobble rules are slightly different from those originally proposed by Crick. The standard bases found in RNA are G, C, A, and U. In addition, the structures of bases in tRNAs may be modified. Some modified bases that may occur in the wobble position in tRNA are I = inosine; $xm^5s^2U$ = 5-methyl-2-thiouridine; $xm^5Um$ = 5-methyl-2′-*O*-methyluridine; Um = 2′-*O*-methyluridine; $xm^5U$ = 5-methyluridine; $xo^5U$ = 5-hydroxyuridine; $k^2C$ = lysidine (a cytosine derivative). The mRNA bases in parentheses are recognized very poorly by the tRNA.

acceptor stem, and AMP is released. Finally, the tRNA with its attached amino acid is released from the enzyme. At this stage, the tRNA is called a **charged tRNA** or an **aminoacyl-tRNA.** In a charged tRNA molecule, the amino acid is attached to the 3′ end of the tRNA by a covalent bond (see Figure 13.12 inset).

The ability of the aminoacyl-tRNA synthetases to recognize tRNAs has sometimes been called the "second genetic code." This recognition process is necessary to maintain the fidelity of genetic information. The frequency of error for aminoacyl-tRNA synthetases is less than $10^{-5}$. In other words, the wrong amino acid is attached to a tRNA less than once in 100,000 times! As you might expect, the anticodon region of the tRNA is usually important for precise recognition by the correct aminoacyl-tRNA synthetase. In studies of *Escherichia coli* synthetases, 17 of the 20 types of aminoacyl-tRNA synthetases recognize the anticodon region of the tRNA. However, other regions of the tRNA are also important recognition sites. These include the acceptor stem and bases in the stem-loop regions.

As mentioned previously, tRNA molecules frequently contain bases within their structure that have been chemically modified. These modified bases can have important effects on tRNA function. For example, modified bases within tRNA molecules affect the rate of translation and the recognition of tRNAs by aminoacyl-tRNA synthetases. Positions 34 and 37 contain the largest variety of modified nucleotides; position 34 is the first base in the anticodon that matches the third base in the codon of mRNA. As discussed next, a modified base at position 34 can have important effects on codon-anticodon recognition.

### Mismatches That Follow the Wobble Rule Can Occur at the Third Position in Codon-Anticodon Pairing

After considering the structure and function of tRNA molecules, let's reexamine some subtle features of the genetic code. As discussed earlier, the genetic code is degenerate, which means that more than one codon can specify the same amino acid. Degeneracy usually occurs at the third position in the codon. For example, valine is specified by GUU, GUC, GUA, and GUG. In all four cases, the first two bases are G and U. The third base, however, can be U, C, A, or G. To explain this pattern of degeneracy, Francis Crick proposed in 1966 that it is due to "wobble" at the third position in the codon-anticodon recognition process. According to the **wobble rules,** the first two positions pair strictly according to the AU/GC rule. However, the third position can tolerate certain types of mismatches (**Figure 13.14**). This proposal suggested that the base at the third position in the codon does not have to hydrogen bond as precisely with the corresponding base in the anticodon.

Because of the wobble rules, some flexibility is observed in the recognition between a codon and anticodon during the process of translation. When two or more tRNAs that differ at the wobble base are able to recognize the same codon, these are termed **isoacceptor tRNAs.** As an example, tRNAs with an anticodon of 3′–CCA–5′ or 3′–CCG–5′ can recognize a codon with the sequence of 5′–GGU–3′. In addition, the wobble rules enable a single type of tRNA to recognize more than one codon. For example, a tRNA with an anticodon sequence of 3′–AAG–5′ can recognize a 5′–UUC–3′ and a 5′–UUU–3′ codon. The 5′–UUC–3′ codon is a perfect match with this tRNA. The 5′–UUU–3′ codon is mismatched according to the standard RNA-RNA hybridization rules (namely, G in the anticodon is mismatched to U in the codon), but the two can fit according to the wobble rules described in Figure 13.14. Likewise, the modification of the wobble base to an inosine allows a tRNA to recognize three different codons. At the cellular level, the ability of a single tRNA to recognize more than one codon makes it unnecessary for a cell to make 61 different tRNA molecules with anticodons that are complementary to the 61 possible sense codons. *E. coli* cells, for example, make a population of tRNA molecules that have just 40 different anticodon sequences.

## 13.3 RIBOSOME STRUCTURE AND ASSEMBLY

In Section 13.2, we examined how the structure and function of tRNA molecules are important in translation. According to the adaptor hypothesis, tRNAs bind to mRNA due to complementarity between the anticodons and codons. Concurrently, the tRNA molecules have the correct amino acid attached to their 3′ ends.

To synthesize a polypeptide, additional events must occur. In particular, the bond between the 3′ end of the tRNA and the amino acid must be broken, and a peptide bond must be formed between the adjacent amino acids. To facilitate these events, translation occurs on the surface of a macromolecular complex known as the **ribosome.** The ribosome can be thought of as the macromolecular arena where translation takes place.

In this section, we will begin by outlining the biochemical compositions of ribosomes in bacterial and eukaryotic cells. We will then examine the key functional sites on ribosomes for the translation process.

## Bacterial and Eukaryotic Ribosomes Are Assembled from rRNA and Proteins

Bacterial cells have one type of ribosome that is found within the cytoplasm. Eukaryotic cells contain biochemically distinct ribosomes in different cellular locations. The most abundant type of ribosome functions in the cytosol, which is the region of the eukaryotic cell that is inside the plasma membrane but outside the membrane-bound organelles. Besides the cytosolic ribosomes, all eukaryotic cells have ribosomes within the mitochondria. In addition, plant cells and algae have ribosomes in their chloroplasts. The compositions of mitochondrial and chloroplast ribosomes are quite different from that of the cytosolic ribosomes. Unless other-wise noted, the term eukaryotic ribosome refers to ribosomes in the cytosol, not to those found within organelles. Likewise, the description of eukaryotic translation refers to translation via cytosolic ribosomes.

Each ribosome is composed of structures called the large and small subunits. This term is perhaps misleading because each ribosomal subunit itself is formed from the assembly of many different proteins and RNA molecules called ribosomal RNA or rRNA. In bacterial ribosomes, the 30S subunit is formed from the assembly of 21 different ribosomal proteins and a 16S rRNA molecule; the 50S subunit contains 34 different proteins and 5S and 23S rRNA molecules (**Table 13.6**). The designations 30S and 50S refer to the rate that these subunits sediment when subjected to a centrifugal force. This rate is described as a sedimentation coefficient in Svedberg units (S), in honor of Theodor Svedberg, who invented the ultracentrifuge. Together, the 30S and 50S subunits form a 70S ribosome. (Note: Svedberg units do not add up linearly.) In bacteria, the ribosomal proteins and rRNA molecules are synthesized in the cytoplasm, and the ribosomal subunits are assembled there.

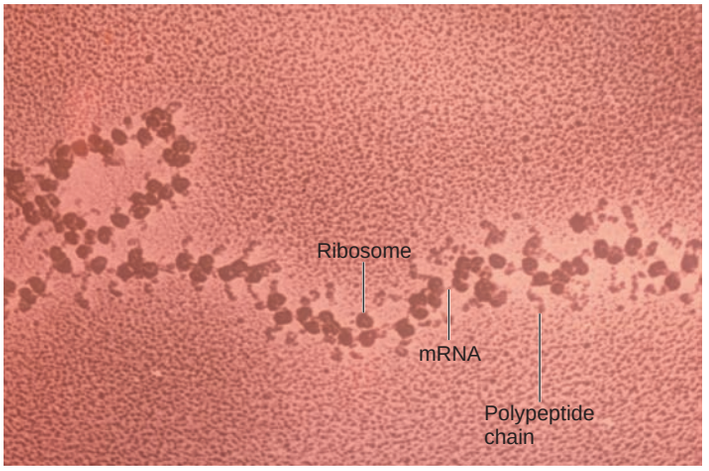
The synthesis of eukaryotic rRNA occurs within the nucleus, and the ribosomal proteins are made in the cytosol, where translation takes place. The 40S subunit is composed of 33 proteins and an 18S rRNA; the 60S subunit is made of 49 proteins and 5S, 5.8S, and 28S rRNAs (see Table 13.6). The assembly of the rRNAs and ribosomal proteins to make the 40S and 60S subunits occurs within the **nucleolus,** a region of the nucleus specialized for this purpose. The 40S and 60S subunits are then exported into the cytosol, where they associate to form an 80S ribosome during translation.

## Components of Ribosomal Subunits Form Functional Sites for Translation

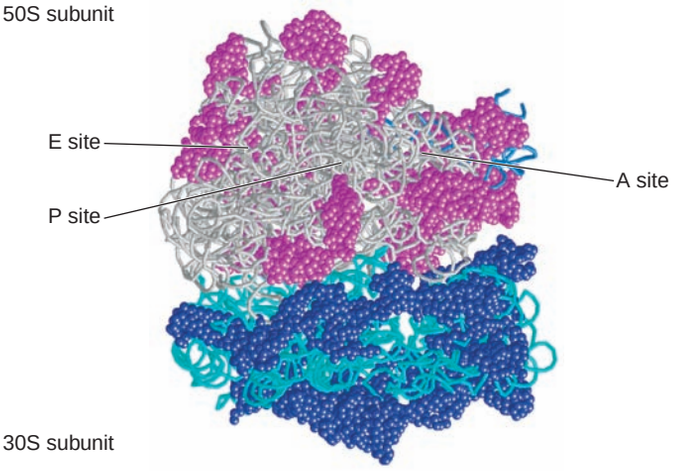
To understand the structure and function of the ribosome at the molecular level, researchers must determine the locations and functional roles of the individual ribosomal proteins and rRNAs. In recent years, many advances have been made toward a molecular understanding of ribosomes. Microscopic and biophysical

**TABLE 13.6**
**Composition of Bacterial and Eukaryotic Ribosomes**

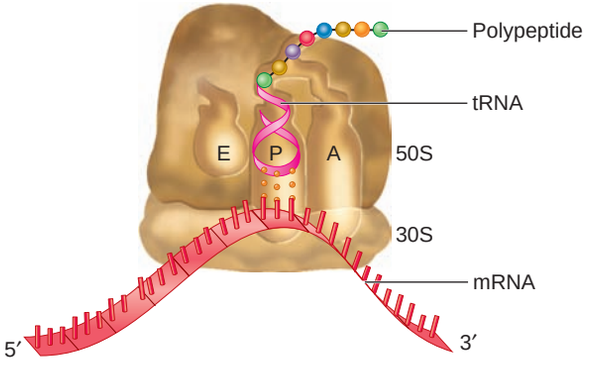
| | Small subunit | Large subunit | Assembled ribosome |
|---|---|---|---|
| **Bacterial** | | | |
| Sedimentation coefficient | 30S | 50S | 70S |
| Number of proteins | 21 | 34 | 55 |
| rRNA | 16S rRNA | 5S rRNA, 23S rRNA | 16S rRNA, 5S rRNA, 23S rRNA |
| **Eukaryotic** | | | |
| Sedimentation coefficient | 40S | 60S | 80S |
| Number of proteins | 33 | 49 | 82 |
| rRNA | 18S rRNA | 5S rRNA, 5.8S rRNA, 28S rRNA | 18S rRNA, 5S rRNA, 5.8S rRNA, 28S rRNA |




(a) Ribosomes as seen with electron microscope




(b) Bacterial ribosome model based on X–ray diffraction studies




(c) Model for ribosome structure

**FIGURE 13.15** **Ribosomal structure.** **(a)** Electron micrograph of ribosomes attached to a bacterial mRNA molecule. **(b)** Crystal structure of the 50S and 30S subunits in bacteria. This model shows the interface between the two subunits. The rRNA is shown in gray strands (50S subunit) and turquoise strands (30S subunit), and proteins are shown in violet (50S subunit) and navy blue (30S subunit). **(c)** A model depicting the sites where tRNA and mRNA bind to an intact ribosome. The mRNA lies on the surface of the 30S subunit. The E, P, and A sites are formed at the interface between the large and small subunits. The growing polypeptide chain exits through a hole in the 50S subunit.

methods have been used to study ribosome structure. An electron micrograph of bacterial ribosomes is shown in **Figure 13.15a.** More recently, a few research groups have succeeded in crystallizing ribosomal subunits, and even intact ribosomes. This is an amazing technical feat, because it is difficult to find the right conditions under which large macromolecules will form highly ordered crystals. **Figure 13.15b** shows the crystal structure of bacterial ribosomal subunits. The overall shape of each subunit is largely determined by the structure of the rRNAs, which constitute most of the mass of the ribosome. The interface between the 30S and 50S subunits is primarily composed of rRNA. Ribosomal proteins cluster on the outer surface of the ribosome and on the periphery of the interface.

During bacterial translation, the mRNA lies on the surface of the 30S subunit within a space between the 30S and 50S subunits. As the polypeptide is being synthesized, it exits through a channel within the 50S subunit (**Figure 13.15c**). Ribosomes contain discrete sites where tRNAs bind and the polypeptide is synthesized. In 1964, James Watson was the first to propose a two-site model for tRNA binding to the ribosome. These sites are known as the **peptidyl site (P site)** and **aminoacyl site (A site).** In 1981, Knud Nierhaus, Hans Sternbach, and Hans-Jörg Rheinberger proposed a three-site model. This model incorporated the observation that uncharged tRNA molecules can bind to a site on the ribosome that is distinct from the P and A sites. This third site is now known as the **exit site (E site).** The locations of the E, P, and A sites are shown in Figure 13.15c. Next, we will examine the roles of these sites during the three stages of translation.

## 13.4 STAGES OF TRANSLATION

Like transcription, the process of translation can be viewed as occurring in three stages: initiation, elongation, and termination. **Figure 13.16** presents an overview of these stages. During **initiation,** the ribosomal subunits, mRNA, and the first tRNA assemble to form a complex. After the initiation complex is formed, the ribosome slides along the mRNA in the 5′ to 3′ direction, moving over the codons. This is the **elongation** stage of translation. As the ribosome moves, tRNA molecules sequentially bind to the mRNA at the A site in the ribosome, bringing with them the appropriate amino acids. Therefore, amino acids are linked in

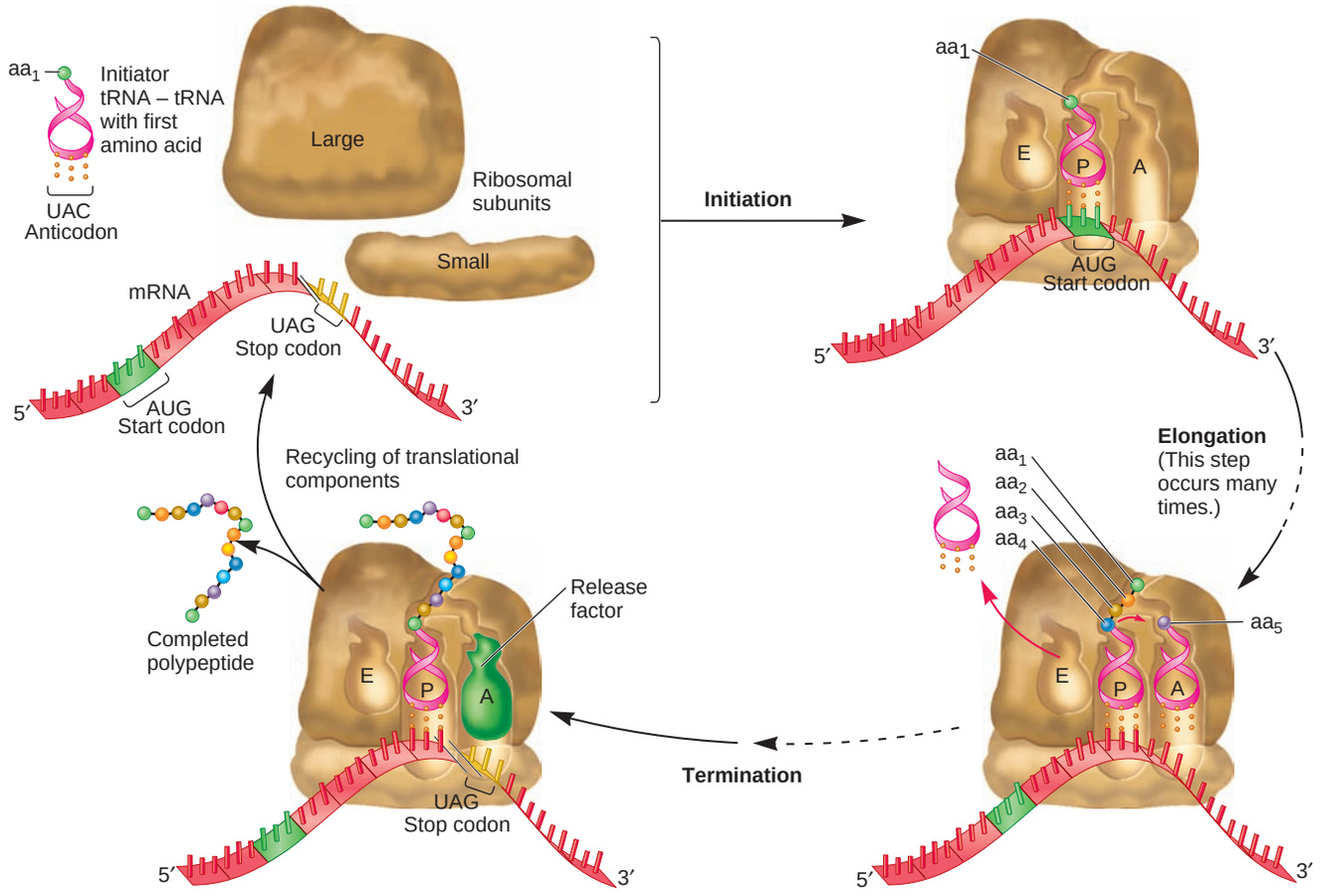



**FIGURE 13.16** **Overview of the stages of translation.** The initiation stage involves the assembly of the ribosomal subunits, mRNA, and the initiator tRNA carrying the first amino acid. During elongation, the ribosome slides along the mRNA and synthesizes a polypeptide chain. Translation ends when a stop codon is reached and the polypeptide is released from the ribosome. (Note: In this and succeeding figures in this chapter, the ribosomes are drawn schematically to emphasize different aspects of the translation process. The structures of ribosomes are described in Figure 13.15.)

**Genes → Traits** The ability of genes to produce an organism's traits relies on the molecular process of gene expression. During translation, the codon sequence within mRNA (which is derived from a gene sequence during transcription) is translated into a polypeptide sequence. After polypeptides are made within a living cell, they function as proteins to govern an organism's traits. For example, once the β-globin polypeptide is made, it functions within the hemoglobin protein and provides red blood cells with the ability to carry oxygen, a vital trait for survival. Translation allows functional proteins to be made within living cells.

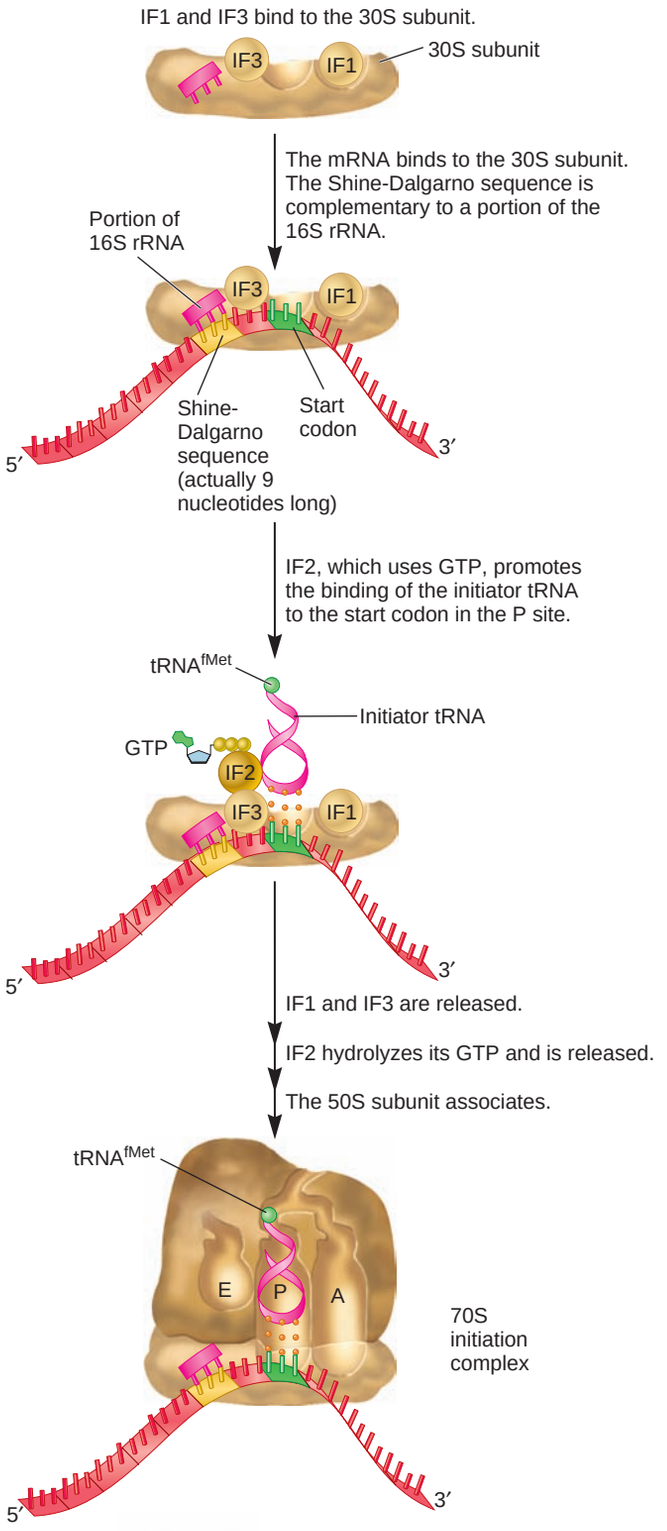
the order dictated by the codon sequence in the mRNA. Finally, a stop codon is reached, signaling the **termination** of translation. At this point, disassembly occurs, and the newly made polypeptide is released. In this section, we will examine the components required for the translation process and consider their functional roles during the three stages of translation.

## The Initiation Stage Involves the Binding of mRNA and the Initiator tRNA to the Ribosomal Subunits

During initiation, an mRNA and the first tRNA bind to the ribosomal subunits. A specific tRNA functions as the **initiator tRNA,** which recognizes the start codon in the mRNA. In bacteria, the initiator tRNA, which is also designated $tRNA^{fMet}$, carries a methionine that has been covalently modified to *N*-formylmethionine. In this modification, a formyl group (—CHO) is attached to the nitrogen atom in methionine after the methionine has been attached to the tRNA.

Figure 13.17 describes the initiation stage of translation in bacteria during which the mRNA, $tRNA^{fMet}$, and ribosomal subunits associate with each other to form an initiation complex. The formation of this complex requires the participation of three initiation factors: IF1, IF2, and IF3. First, IF1 and IF3 bind to the 30S subunit. IF1 and IF3 prevent the association of the 50S subunit. Next, the mRNA binds to the 30S subunit. This binding is facilitated by a nine-nucleotide sequence within the bacterial mRNA called the **Shine-Dalgarno sequence.** The location of this sequence is shown in Figure 13.17 and in more detail in Figure 13.18. How does the Shine-Dalgarno sequence facilitate the binding of mRNA to the ribosome? The Shine-Dalgarno sequence is complementary to a short sequence within the 16S rRNA, which promotes the hydrogen bonding of the mRNA to the 30S subunit.

Next, $tRNA^{fMet}$ binds to the mRNA that is already attached to the 30S subunit (see Figure 13.17). This step requires the function of IF2, which uses GTP. The $tRNA^{fMet}$ binds to the start




ONLINE ANIMATION INTERACTIVE EXERCISE **FIGURE 13.17 The initiation stage of translation in bacteria.**

codon, which is typically a few nucleotides downstream from the Shine-Dalgarno sequence. The start codon is usually AUG, but in some cases it can be GUG or UUG. Even when the start codon is GUG (which normally encodes valine) or UUG (which normally encodes leucine), the first amino acid in the polypeptide is still a formylmethionine because only a $tRNA^{fMet}$ can initiate translation. During or after translation of the entire polypeptide, the formyl group or the entire formylmethionine may be removed. Therefore, some polypeptides may not have formylmethionine or methionine as their first amino acid. As noted in Figure 13.17, the $tRNA^{fMet}$ binds to the P site on the ribosome. IF1 is thought to occupy a portion of the A site, thereby preventing the binding of $tRNA^{fMet}$ to the A site during initiation. By comparison, during the elongation stage that is discussed later, all of the other tRNAs initially bind to the A site.

After the mRNA and $tRNA^{fMet}$ have become bound to the 30S subunit, IF1 and IF3 are released, and then IF2 hydrolyzes its GTP and is also released. This allows the 50S ribosomal subunit to associate with the 30S subunit. Much later, after translation is completed, IF1 binding is necessary to dissociate the 50S and 30S ribosomal subunits so that the 30S subunit can reinitiate with another mRNA molecule.

In eukaryotes, the assembly of the initiation complex bears similarities to that in bacteria. However, as described in **Table 13.7**, additional factors are required for the initiation process. Note that the initiation factors are designated eIF (for eukaryotic Initiation Factor) to distinguish them from bacterial initiation factors. The initiator tRNA in eukaryotes carries methionine rather than formylmethionine, as in bacteria. A eukaryotic initiation factor, eIF2, binds directly to $tRNA^{Met}$ to recruit it to the 40S subunit. Eukaryotic mRNAs do not have a Shine-Dalgarno sequence. How then are eukaryotic mRNAs recognized by the ribosome? The mRNA is recognized by eIF4, which is a multiprotein complex that recognizes the 7-methylguanosine cap and facilitates the binding of the mRNA to the 40S subunit.

The identification of the correct AUG start codon in eukaryotes differs greatly from that in bacteria. After the initial binding of mRNA to the ribosome, the next step is to locate an AUG start codon that is somewhere downstream from the 5′ cap structure. In 1986, Marilyn Kozak proposed that the ribosome begins at the 5′ end and then scans along the mRNA in the 3′ direction in search of an AUG start codon. In many, but not all, cases, the ribosome uses the first AUG codon that it encounters as a start codon. When a start codon is identified, the 60S subunit assembles onto the 40S subunit with the aid of eIF5.

By analyzing the sequences of many eukaryotic mRNAs, researchers have found that not all AUG codons near the 5′ end of mRNA can function as start codons. In some cases, the scanning ribosome passes over the first AUG codon and chooses an AUG farther down the mRNA. The sequence of bases around the AUG codon plays an important role in determining whether or not it is selected as the start codon by a scanning ribosome. The consensus sequence for optimal start codon recognition in more complex eukaryotes, such as vertebrates and vascular plants, is shown here.

| | | | | | | *Start Codon* | | | |
|---|---|---|---|---|---|---|---|---|---|
| G | C | C | (A/G) | C | C | A | U | G | G |
| −6 | −5 | −4 | −3 | −2 | −1 | +1 | +2 | +3 | +4 |

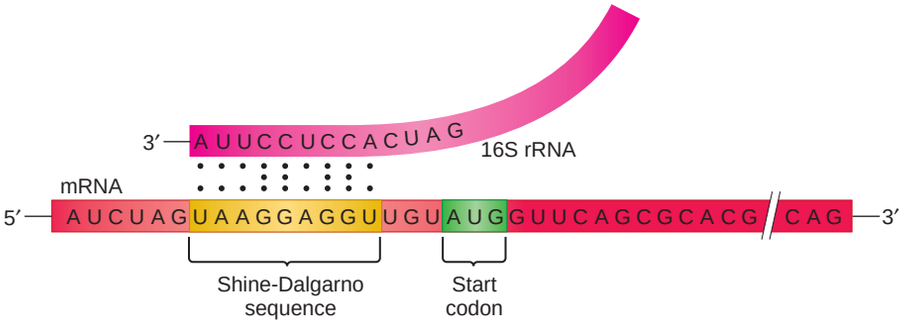



**FIGURE 13.18 The locations of the Shine-Dalgarno sequence and the start codon in bacterial mRNA.** The Shine-Dalgarno sequence is complementary to a sequence in the 16S rRNA. It hydrogen bonds with the 16S rRNA to promote initiation. The start codon is typically a few nucleotides downstream from the Shine-Dalgarno sequence.

Aside from an AUG codon itself, a guanine at the +4 position and a purine, preferably an adenine, at the –3 position are the most important sites for start codon selection. These rules for optimal translation initiation are called **Kozak's rules.**

TABLE **13.7**

**A Simplified Comparison of Translational Protein Factors in Bacteria and Eukaryotes**

| Bacterial Factors | Eukaryotic Factors* | Function |
|---|---|---|
| **Initiation Factors** | | |
| | eIF4 | Involved with the recognition of the 7-methylguanosine cap and the binding of the mRNA to the small ribosomal subunit |
| IF1, IF3 | eIF1, eIF3, eIF6 | Prevent the association between the small and large ribosomal subunits and favor their disassociation |
| IF2 | eIF2 | Promotes the binding of the initiator tRNA to the small ribosomal subunit |
| | eIF5 | Helps to dissociate the other elongation factors, which allows the large ribosomal subunit to bind |
| **Elongation Factors** | | |
| EF-Tu | eEF1α | Involved in the binding of tRNAs to the A site |
| EF-Ts | eEF1βγ | Nucleotide exchange factors required for the functioning of EF-Tu and eEF1α, respectively |
| EF-G | eEF2 | Required for translocation |
| **Release Factors** | | |
| RF1, RF2 | eRF1 | Recognize a stop codon and trigger the cleavage of the polypeptide from the tRNA |
| RF3 | eRF3 | GTPases that are also involved in termination |

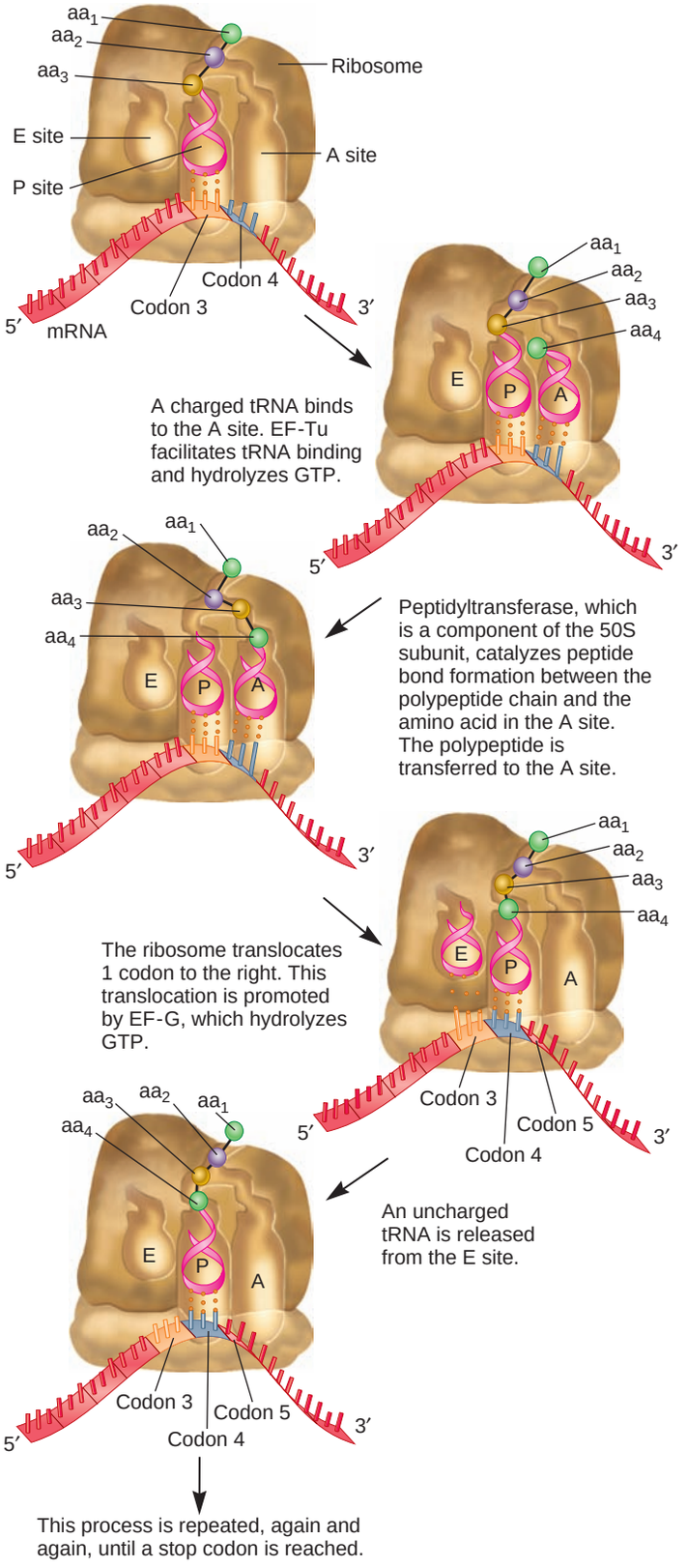
*Eukaryotic translation factors are typically composed of multiple proteins.

## Polypeptide Synthesis Occurs During the Elongation Stage

During the elongation stage of translation, amino acids are added, one at a time, to the polypeptide chain (**Figure 13.19**). Even though this process is rather complex, it occurs at a remarkable rate. Under normal cellular conditions, a polypeptide chain can elongate at a rate of 15 to 20 amino acids per second in bacteria and 2 to 6 amino acids per second in eukaryotes!

To begin elongation, a charged tRNA brings a new amino acid to the ribosome so it can be attached to the end of the growing polypeptide chain. At the top of Figure 13.19, which describes bacterial translation, a short polypeptide is attached to the tRNA located at the P site of the ribosome. A charged tRNA carrying a single amino acid binds to the A site. This binding occurs because the anticodon in the tRNA is complementary to the codon in the mRNA. The hydrolysis of GTP by the elongation factor, EF-Tu, provides energy for the binding of a tRNA to the A site. In addition, the 16S rRNA, which is a component of the small 30S ribosomal subunit, plays a key role that ensures the proper recognition between the mRNA and correct tRNA. The 16S rRNA can detect when an incorrect tRNA is bound at the A site and will prevent elongation until the mispaired tRNA is released from the A site. This phenomenon, termed the **decoding function** of the ribosome, is important in maintaining high fidelity of mRNA translation. An incorrect amino acid is incorporated into a growing polypeptide at a rate of approximately one mistake per 10,000 amino acids, or $10^{-4}$.

The next step of elongation is the **peptidyl transfer** reaction—the polypeptide is removed from the tRNA in the P site and transferred to the amino acid at the A site. This transfer is accompanied by the formation of a peptide bond between the amino acid at the A site and the polypeptide chain, lengthening the chain by one amino acid. The peptidyl transfer reaction is catalyzed by a component of the 50S subunit known as **peptidyl transferase,** which is composed of several proteins and rRNA. Interestingly, based on the crystal structure of the 50S subunit,




ONLINE ANIMATION INTERACTIVE EXERCISE

**FIGURE 13.19 The elongation stage of translation in bacteria.** This process begins with the binding of an incoming tRNA. The hydrolysis of GTP by EF-Tu provides the energy for the binding of the tRNA to the A site. A peptide bond is then formed between the incoming amino acid and the last amino acid in the growing polypeptide chain. This moves the polypeptide chain to the A site. The ribosome then translocates in the 3′ direction so that the two tRNAs are moved to the E and P sites. The tRNA carrying the polypeptide is now back in the P site. This translocation requires the hydrolysis of GTP via EF-G. The uncharged tRNA in the E site is released from the ribosome. Now the process is ready to begin again. Each cycle of elongation causes the polypeptide chain to grow by one amino acid.

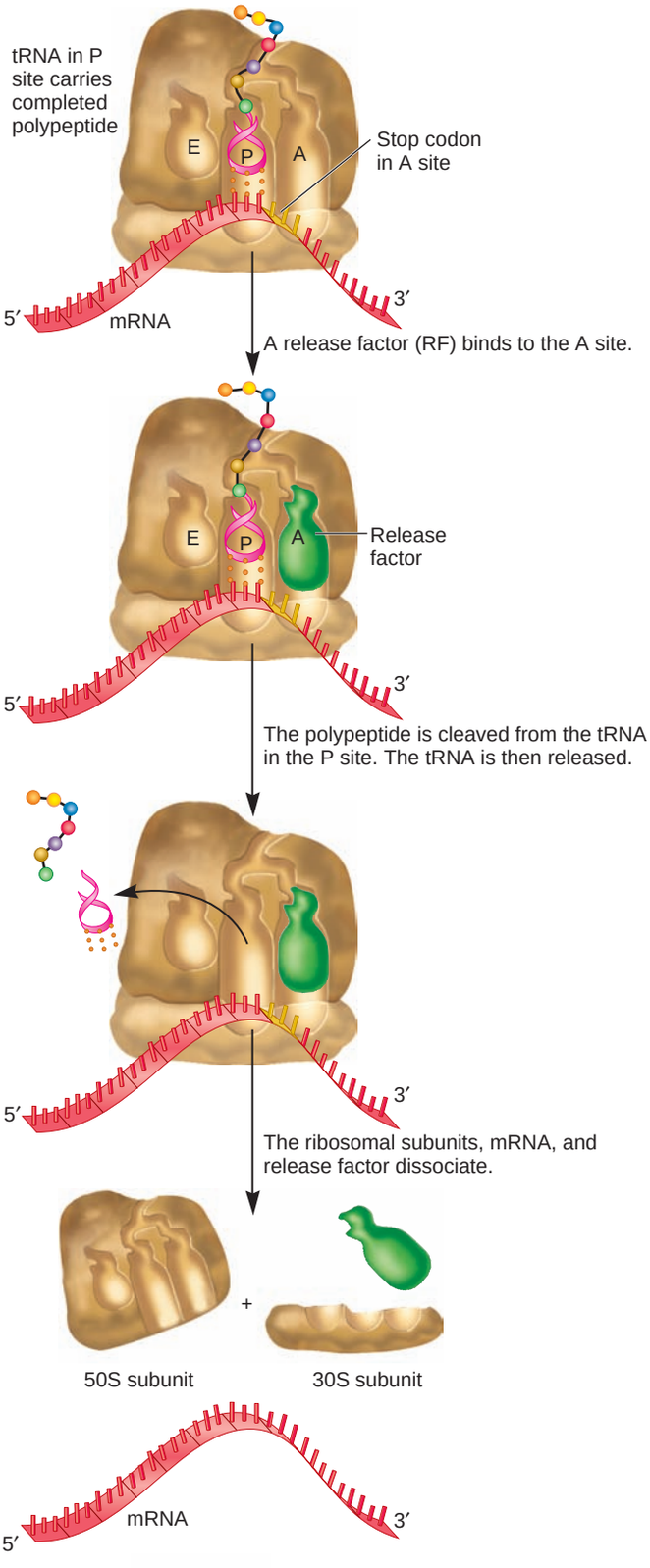
Thomas Steitz, Peter Moore, and their colleagues concluded that the 23S rRNA—not the ribosomal protein—catalyzes bond formation between adjacent amino acids. In other words, the ribosome is a ribozyme!

After the peptidyl transfer reaction is complete, the ribosome moves, or translocates, to the next codon in the mRNA. This moves the tRNAs at the P and A sites to the E and P sites, respectively. Finally, the uncharged tRNA exits the E site. You should notice that the next codon in the mRNA is now exposed in the unoccupied A site. At this point, a charged tRNA can enter the empty A site, and the same series of steps can add the next amino acid to the polypeptide chain. As you may have realized, the A, P, and E sites are named for the role of the tRNA that is usually found there. The A site binds an aminoacyl-tRNA (also called a charged tRNA), the P site usually contains the peptidyl-tRNA (a tRNA with an attached peptide), and the E site is where the uncharged tRNA exits.

## Termination Occurs When a Stop Codon Is Reached in the mRNA

The final stage of translation, known as termination, occurs when a stop codon is reached in the mRNA. In most species, the three stop codons are UAA, UAG, and UGA. The stop codons are not recognized by a tRNA with a complementary sequence. Instead, they are recognized by proteins known as **release factors** (see Table 13.7). Interestingly, the three-dimensional structures of release factor proteins are "molecular mimics" that resemble the structure of tRNAs. Such proteins can specifically bind to a stop codon sequence. In bacteria, RF1 recognizes UAA and UAG, and RF2 recognizes UGA and UAA. A third release factor, RF3, is also required. In eukaryotes, a single release factor, eRF, recognizes all three stop codons and eRF3 is also required for termination.

**Figure 13.20** illustrates the termination stage of translation in bacteria. At the top of this figure, the completed polypeptide chain is attached to a tRNA in the P site. A stop codon is located at the A site. In the first step, RF1 or RF2 binds to the stop codon at the A site and RF3 (not shown) binds at a different location on the ribosome. After RF1 (or RF2) and RF3 have bound, the bond between the polypeptide and the tRNA is hydrolyzed. The polypeptide and tRNA are then released from the ribosome.




ONLINE ANIMATION INTERACTIVE EXERCISE

**FIGURE 13.20 The termination stage of translation in bacteria.** When a stop codon is reached, RF1 or RF2 binds to the A site. (RF3 binds elsewhere and uses GTP to facilitate the termination process.) The polypeptide is cleaved from the tRNA in the P site and released. The tRNA is released, and the rest of the components disassemble.

The final step in translational termination is the disassembly of ribosomal subunits, mRNA, and the release factors.

## Bacterial Translation Can Begin Before Transcription Is Completed

Although most of our knowledge concerning transcription and translation has come from genetic and biochemical studies, electron microscopy (EM) has also been an important tool in elucidating the mechanisms of transcription and translation. As described earlier in this chapter, EM has been a critical technique in facilitating our understanding of ribosome structure. In addition, it has been employed to visualize genetic processes such as translation.

The first success in the EM observation of gene expression was achieved by Oscar Miller, Jr., and his colleagues in 1967. **Figure 13.21** shows an electron micrograph of a bacterial gene in the act of gene expression. Prior to this experiment, biochemical and genetic studies had suggested that the translation of a bacterial structural gene begins before the mRNA transcript is completed. In other words, as soon as an mRNA strand is long enough, a ribosome attaches to the 5′ end and begins translation, even before RNA polymerase has reached the transcriptional termination site within the gene. This phenomenon is termed the coupling between transcription and translation in bacterial cells. Note that coupling of these processes does not usually occur in eukaryotes, because transcription takes place in the nucleus while translation occurs in the cytosol.

As shown in Figure 13.21, several RNA polymerase enzymes have recognized a gene and begun to transcribe it. Because the transcripts on the right side are longer than those on the left, Miller concluded that transcription was proceeding from left to right in the micrograph. This EM image also shows the process of translation. Relatively small mRNA transcripts, near the left side of the figure, have a few ribosomes attached to them. As the transcripts become longer, additional ribosomes are attached to them. The term **polyribosome,** or **polysome,** is used to describe an mRNA transcript that has many bound ribosomes in the act of translation. In this electron micrograph, the nascent polypeptide chains were too small for researchers to observe. In later studies, as EM techniques became more refined, the polypeptide chains emerging from the ribosome were also visible (see Figure 13.15a).

## Bacterial and Eukaryotic Translation Show Similarities and Differences

Throughout this chapter, we have compared translation in both bacteria and eukaryotic organisms. The general steps of translation are similar in all forms of life, but we have also seen some striking differences between bacteria and eukaryotes. **Table 13.8** compares translation between these groups.

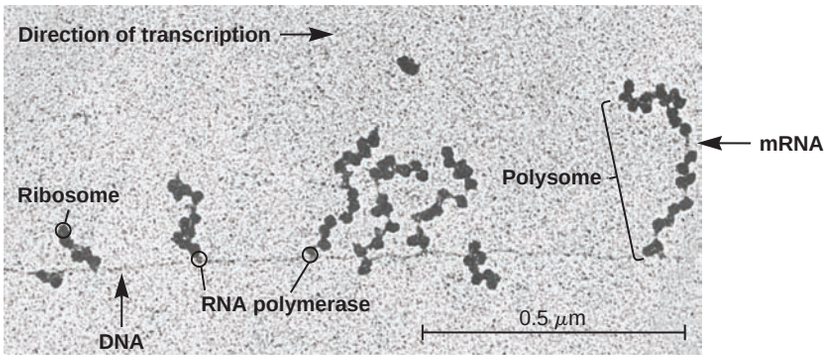



**FIGURE 13.21 Coupling between transcription and translation in bacteria.** An electron micrograph showing the simultaneous transcription and translation processes. The DNA is transcribed by many RNA polymerases that move along the DNA from left to right. Note that the RNA transcripts are getting longer as you go from left to right. Ribosomes attach to the mRNA, even before transcription is completed. The complex of many ribosomes bound to the same mRNA is called a polyribosome or a polysome. Several polyribosomes are seen here.

TABLE **13.8**

**A Comparison of Bacterial and Eukaryotic Translation**

| | Bacterial | Eukaryotic |
|---|---|---|
| Ribosome composition: | 70S ribosomes:<br>30S subunit—<br>21 proteins + 1 rRNA<br>50S subunit—<br>34 proteins + 2 rRNAs | 80S ribosomes:<br>40S subunit—<br>33 proteins + 1 rRNA<br>60S subunit—<br>49 proteins + 3 rRNAs |
| Initiator tRNA: | $tRNA^{fmet}$ | $tRNA^{Met}$ |
| Formation of the initiation complex: | Requires IF1, IF2, and IF3 | Requires more initiation factors compared to bacterial initiation |
| Initial binding of mRNA to the ribosome: | Requires a Shine-Dalgarno sequence | Requires a 7-methylguanosine cap |
| Selection of a start codon: | AUG, GUG, or UUG located just downstream from the Shine-Dalgarno sequence | According to Kozak's rules |
| Elongation rate | Typically 15 to 20 amino acids per second | Typically 2 to 6 amino acids per second |
| Termination: | Requires RF1, RF2, and RF3 | Requires eRF1 and eRF3 |
| Coupled to transcription: | Yes | No |

## KEY TERMS

**Page 326.** structural genes, messenger RNA (mRNA)
**Page 327.** alkaptonuria, inborn error of metabolism
**Page 328.** one-gene/one-enzyme hypothesis
**Page 329.** polypeptide, protein, translation, genetic code, sense codons, start codon, stop codons, termination codons, nonsense codons, anticodons, degenerate, synonymous codons, wobble base, reading frame
**Page 330.** frameshift mutation
**Page 331.** selenocysteine, pyrrolysine
**Page 332.** cell-free translation system
**Page 336.** peptide bond, N-terminus, amino-terminal end, C-terminus, carboxyl-terminal end, side chain, R group
**Page 337.** primary structure, chaperones, secondary structure, α helix, β sheet
**Page 338.** tertiary structure, quaternary structure, subunits
**Page 339.** enzymes
**Page 340.** adaptor hypothesis
**Page 344.** aminoacyl-tRNA synthetases
**Page 345.** charged tRNA, aminoacyl-tRNA, wobble rules, isoacceptor tRNAs
**Page 346.** ribosome, nucleolus
**Page 347.** peptidyl site (P site), aminoacyl site (A site), exit site (E site), initiation, elongation
**Page 348.** termination, initiator tRNA, Shine-Dalgarno sequence
**Page 350.** Kozak's rules, decoding function, peptidyl transfer, peptidyl transferase
**Page 351.** release factors
**Page 352.** polyribosome, polysome

## CHAPTER SUMMARY

- Cellular proteins are made via the translation of mRNA.

### 13.1 The Genetic Basis for Protein Synthesis

- Garrod studied the disease called alkaptonuria and suggested that some genes may encode enzymes (see Figure 13.1).
- Beadle and Tatum studied *Neurospora* mutants that were altered in their nutritional requirements and hypothesized that one gene encodes one enzyme. This hypothesis was later modified because: (1) some proteins are not enzymes; (2) some proteins are composed of two or more different polypeptides; and (3) some genes encode RNAs that are not translated into polypeptides (see Figure 13.2).
- During translation, the sequence of codons in mRNA is used via tRNA molecules to make a polypeptide with a specific amino acid sequence (see Figure 13.3).
- The genetic code refers to the relationship between three-base codons in the mRNA and the amino acids that are incorporated into a polypeptide. One codon (AUG) is a start codon, which determines the reading frame of the mRNA. Three codons (UAA, UAG, and UGA) can function as stop codons (see Table 13.1).
- Crick studied mutations in T4 phage and determined that the genetic code is read in triplets (see Table 13.2).
- The genetic code is largely universal but some exceptions are known to occur (see Table 13.3).
- Nirenberg and colleagues used synthetic RNA and a cell-free translation system to decipher the genetic code (see Figure 13.4).
- Other methods to decipher the genetic code included the synthesis of copolymers by Khorana and the triplet-binding assay of Nirenberg and Leder (see Table 13.4, Figure 13.5).
- A polypeptide is made by the formation of peptide bonds between adjacent amino acids. Each polypeptide has a directionality from its amino terminus to its carboxyl terminus that parallels the arrangement of codons in mRNA in the 5′ to 3′ direction (see Figure 13.6).
- Amino acids differ in their side chain structure (see Figure 13.7).
- Cellular proteins carry out a variety of functions. The structures and functions of proteins are largely responsible for an organism's traits (see Table 13.5).
- Protein structure can be viewed at different levels, which include primary structure (sequence of amino acids), secondary structure (repeating folding patterns such as the α helix and the β sheet), tertiary structure (additional folding), and quaternary structure (the binding of multiple subunits to each other) (see Figures 13.8, 13.9).

### 13.2 Structure and Function of tRNA

- The anticodon in a tRNA is complementary to a codon in mRNA. The tRNA carries a specific amino acid that corresponds to the codon in the mRNA according to the genetic code (see Figure 13.10).
- Chapeville used Raney nickel to determine that the recognition between tRNA and mRNA is not due to the chemistry of the amino acid that is attached to the tRNA (see Figure 13.11).
- The secondary structure of tRNA resembles a cloverleaf. The anticodon is in the second loop and the amino acid is attached to the 3′ end (see Figure 13.12).
- Aminoacyl-tRNA synthetases are a group of enzymes that attach the correct amino acid to a tRNA. The resulting tRNA is called a charged tRNA or an aminoacyl-tRNA (see Figure 13.13).
- Mismatches are allowed between the pairing of tRNAs and mRNA according to the wobble rules (see Figure 13.14).

### 13.3 Ribosome Structure and Assembly

- Ribosomes are the site of polypeptide synthesis. The small and large subunit of ribosomes are composed of rRNAs and multiple proteins (see Table 13.6).
- A ribosome contains an A (aminoacyl), P (peptidyl) and E (exit) site, which are occupied by tRNA molecules (see Figure 13.15).

### 13.4 Stages of Translation

- The three stages of translation are initiation, elongation, and termination (see Figure 13.16).
- During the initiation stage of translation, the mRNA, initiator tRNA, and ribosomal subunits assemble. Initiation factors are involved in the process. In bacteria, the Shine-Dalgarno sequence promotes the binding of the mRNA to the small ribosomal subunit (see Figures 13.17, 13.18, Table 13.7).
- During elongation, tRNAs bring amino acids to the A site and a series of peptidyl transferase reactions creates a polypeptide. At each step, the polypeptide is transferred from the P site to the A site. The tRNAs are released from the E site. Elongation factors are involved in this process (see Figure 13.19).
- Start codon selection in complex eukaryotes follows Kozak's rules.
- During termination, a release factor binds to a stop codon in the A site. This promotes the cleavage of the polypeptide from the tRNA and the subsequent disassembly of the tRNA, mRNA, and ribosomal subunits (see Figure 13.20).
- Bacterial translation can begin before transcription is completed (see Figure 13.21).
- Bacterial and eukaryotic translation show many similarities and differences (see Table 13.8).

## PROBLEM SETS & INSIGHTS

### Solved Problems

S1. The first amino acid in a certain bacterial polypeptide is methionine. The start codon in the mRNA is GUG, which codes for valine. Why isn't the first amino acid formylmethionine or valine?

***Answer:*** The first amino acid in a polypeptide is carried by the initiator tRNA, which always carries formylmethionine. This occurs even when the start codon is GUG (valine) or UUG (leucine). The formyl group can be later removed to yield methionine as the first amino acid.

S2. A tRNA has the anticodon sequence 3′–CAG–5′. What amino acid does it carry?

***Answer:*** Because the anticodon is 3′–CAG–5′, it would be complementary to a codon with the sequence 5′–GUC–3′. According to the genetic code, this codon specifies the amino acid valine. Therefore, this tRNA must carry valine at its acceptor stem.

S3. In eukaryotic cells, the assembly of ribosomal subunits occurs in the nucleolus. As discussed in Chapter 12 (see Figure 12.16), a single 45S rRNA transcript is cleaved to produce the three rRNA fragments—18S, 5.8S, and 28S rRNA—that play a key role in creating the structure of the ribosome. The genes that encode the 45S precursor are found in multiple copies (i.e., they are moderately repetitive). The segments of chromosomes that contain the 45S rRNA genes align themselves at the center of the nucleolus. This site is called the nucleolar-organizing center. In this region, active transcription of the 45S gene takes place. Briefly explain how the assembly of the ribosomal subunits occurs.

***Answer:*** In the nucleolar-organizing center, the 45S RNA is cleaved to the 18S, 5.8S, and 28S rRNAs. The other components of the ribosomal subunits, 5S rRNA and ribosomal proteins, must also be imported into the nucleolar region. Because proteins are made in the cytosol, they must enter the nucleus through the nuclear pores. When all the components are present, they assemble into 40S and 60S ribosomal subunits. Following assembly, the ribosomal subunits exit the nucleus through the nuclear pores and enter the cytosol.

S4. An antibiotic is a drug that kills or inhibits the growth of microorganisms. The use of antibiotics has been of great importance in the battle against many infectious diseases caused by microorganisms. For many antibiotics, their mode of action is to inhibit the translation process within bacterial cells. Certain antibiotics selectively bind to bacterial (70S) ribosomes but do not inhibit eukaryotic (80S) ribosomes. Their ability to inhibit translation can occur at different steps in the translation process. For example, tetracycline prevents the attachment of tRNA to the ribosome, whereas erythromycin inhibits the translocation of the ribosome along the mRNA. Why would an antibiotic bind to a bacterial ribosome but not to a eukaryotic ribosome? How does inhibition of translation by antibiotics such as tetracycline prevent bacterial growth?

***Answer:*** Because bacterial ribosomes have a different protein and rRNA composition than eukaryotic ribosomes, certain drugs can recognize these different components, bind specifically to bacterial ribosomes, thereby interferimg with the process of translation. In other words, the surface of a bacterial ribosome must be somewhat different from the surface of a eukaryotic ribosome so that the drugs bind to the surface of only bacterial ribosomes. If a bacterial cell is exposed to tetracycline or other antibiotics, it cannot synthesize new polypeptides. Because polypeptides form functional proteins needed for processes such as cell division, the bacterium is unable to grow and proliferate.

### Conceptual Questions

C1. An mRNA has the following sequence:

5′–GGCGAUGGGCAAUAAACCGGGCCAGUAAGC–3′

Identify the start codon and determine the complete amino acid sequence that would be translated from this mRNA.

C2. What does it mean when we say that the genetic code is degenerate? Discuss the universality of the genetic code.

C3. According to the adaptor hypothesis, are the following statements true or false?

A. The sequence of anticodons in tRNA directly recognizes codon sequences in mRNA, with some room for wobble.

B. The amino acid attached to the tRNA directly recognizes codon sequences in mRNA.

C. The amino acid attached to the tRNA affects the binding of the tRNA to a codon sequence in mRNA.

C4. In bacteria, researchers have isolated strains that carry mutations within tRNA genes. These mutations can change the sequence of the anticodon. For example, a normal $tRNA^{Trp}$ gene would encode a tRNA with the anticodon 3′–ACC–5′. A mutation could change this sequence to 3′–CCC–5′. When this mutation occurs, the tRNA still carries a tryptophan at its 3′ acceptor stem, even though the anticodon sequence has been altered.

A. How would this mutation affect the synthesis of polypeptides within the bacterium?

B. What does this mutation tell you about the recognition between tryptophanyl-tRNA synthetase and $tRNA^{Trp}$? Does the enzyme primarily recognize the anticodon or not?

C5. The covalent attachment of an amino acid to a tRNA is an endergonic reaction. In other words, it requires an input of energy for the reaction to proceed. Where does the energy come from to attach amino acids to tRNA molecules?

C6. The wobble rules for tRNA-mRNA pairing are shown in Figure 13.14. If we assume that the tRNAs do not contain modified bases, what is the minimum number of tRNAs needed to efficiently recognize the codons for the following types of amino acids?

A. Leucine

B. Methionine

C. Serine

C7. How many different sequences of mRNA could encode a peptide with the sequence proline-glycine-methionine-serine?

C8. If a tRNA molecule carries a glutamic acid, what are the two possible anticodon sequences that it could contain? Be specific about the 5′ and 3′ ends.

C9. A tRNA has an anticodon sequence 3′–GGU–5′. What amino acid does it carry?

C10. If a tRNA has an anticodon sequence 3′–CCI–5′, what codon(s) can it recognize?

C11. Describe the anticodon of a single tRNA that could recognize the codons 5′–AAC–3′ and 5′–AAU–3′. How would this tRNA need to be modified for it to also recognize 5′–AAA–3′?

C12. Describe the structural features that all tRNA molecules have in common.

C13. In the tertiary structure of tRNA, where is the anticodon region relative to the attachment site for the amino acid? Are they adjacent to each other?

C14. What is the role of aminoacyl-tRNA synthetase? The ability of the aminoacyl-tRNA synthetases to recognize tRNAs has sometimes been called the "second genetic code." Why has the function of this type of enzyme been described this way?

C15. What is an activated amino acid?

C16. Discuss the significance of modified bases within tRNA molecules.

C17. How and when does formylmethionine become attached to the initiator tRNA in bacteria?

C18. Is it necessary for a cell to make 61 different tRNA molecules, corresponding to the 61 codons for amino acids? Explain your answer.

C19. List the components required for translation. Describe the relative sizes of these different components. In other words, which components are small molecules, macromolecules, or assemblies of macromolecules?

C20. Describe the components of eukaryotic ribosomal subunits and where the assembly of the subunits occurs within living cells.

C21. The term subunit can be used in a variety of ways. Compare the use of the term subunit in proteins versus ribosomal subunit.

C22. Do the following events during bacterial translation occur primarily within the 30S subunit, within the 50S subunit, or at the interface between these two ribosomal subunits?

A. mRNA-tRNA recognition

B. Peptidyl transfer reaction

C. Exit of the polypeptide chain from the ribosome

D. Binding of initiation factors IF1, IF2, and IF3

C23. What are the three stages of translation? Discuss the main events that occur during these three stages.

C24. Describe the sequence in bacterial mRNA that promotes recognition by the 30S subunit.

C25. For each of the following initiation factors, how would eukaryotic initiation of translation be affected if it were missing?

A. eIF2

B. eIF4

C. eIF5

C26. How does a eukaryotic ribosome select its start codon? Describe the sequences in eukaryotic mRNA that provide an optimal context for a start codon.

C27. For each of the following sequences, rank them in order (from best to worst) as sequences that could be used to initiate translation according to Kozak's rules.

GACGCCAUGG

GCCUCCAUGC

GCCAUCAAGG

GCCACCAUGG

C28. Explain the functional roles of the A, P, and E sites during translation.

C29. An mRNA has the following sequence: 5′–AUG UAC UAU GGG GCG UAA–3′

Describe the amino acid sequence of the polypeptide that would be encoded by this mRNA. Be specific about the amino and carboxyl terminal ends.

C30. Which steps during the translation of bacterial mRNA involve an interaction between complementary strands of RNA?

C31. What is the function of the nucleolus?

C32. In which of the ribosomal sites, the A site, P site, and/or E site, could the following be found?

A. A tRNA without an amino acid attached

B. A tRNA with a polypeptide attached

C. A tRNA with a single amino acid attached

C33. What is a polysome?

C34. According to Figure 13.19, explain why the ribosome translocates along the mRNA in a 5′ to 3′ direction rather than a 3′ to 5′ direction.

C35. The lactose permease of *E. coli* is a protein composed of a single polypeptide that is 417 amino acids in length. By convention, the amino acids within a polypeptide are numbered from the amino-terminal end to the carboxyl-terminal end. Are the following questions about the lactose permease true or false?

A. Because the 64th amino acid is glycine and the 68th amino acid is aspartic acid, the codon for glycine, 64, is closer to the 3′ end of the mRNA than the codon for aspartic acid, 68.

B. The mRNA that encodes the lactose permease must be greater than 1241 nucleotides in length.

C36. An mRNA encodes a polypeptide that is 312 amino acids in length. The 53rd codon in this polypeptide is a tryptophan codon. A mutation in the gene that encodes this polypeptide changes this tryptophan codon into a stop codon. How many amino acids would be in the resulting polypeptide: 52, 53, 259, or 260?

C37. Explain what is meant by the coupling of transcription and translation in bacteria. Does coupling occur in bacterial and/or eukaryotic cells? Explain.

## Experimental Questions

E1. In the experiment of Figure 13.4, what would be the predicted amounts of amino acids incorporated into polypeptides if the RNA was a random polymer containing 50% C and 50% G?

E2. With regard to the experiment described in Figure 13.11, answer the following questions:

A. Why was a polyUG mRNA template used?

B. Would you radiolabel the cysteine with the isotope $^{14}C$ or $^{35}S$? Explain your choice.

C. What would be the expected results if the experiment was followed in the same way except that a polyGC template was used? Note: A polyGC template could contain two different alanine codons (GCC and GCG), but it could not contain any cysteine codons.

E3. An experimenter has a chemical reagent that modifies threonine to another amino acid. Following the protocol described in Figure 13.11, an mRNA is made composed of 50% C and 50% A. The amino acid composition of the resultant polypeptides is 12.5% lysine, 12.5% asparagine, 25% serine, 12.5% glutamine, 12.5% histidine, and 25% proline. One of the amino acids present in this polypeptide is due to the modification of threonine. Which amino acid is it? Based on the structure of the amino acid side chains, explain how the structure of threonine has been modified.

E4. Polypeptides can be translated in vitro. Would a bacterial mRNA be translated in vitro by eukaryotic ribosomes? Would a eukaryotic mRNA be translated in vitro by bacterial ribosomes? Why or why not?

E5. Discuss how the elucidation of the structure of the ribosome can help us to understand its function.

E6. Figure 13.21 shows an electron micrograph of a bacterial gene as it is being transcribed and translated. In this figure, label the 5′ and 3′ ends of the DNA and RNA strands. Place an arrow where you think the start codons are found in the mRNA transcripts.

E7. Chapter 18 describes a blotting method known as Western blotting that can be used to detect the production of a polypeptide that is translated from a particular mRNA. In this method, a protein is detected with an antibody that specifically recognizes and binds to its amino acid sequence. The antibody acts as a probe to detect the presence of the protein. In a Western blotting experiment, a mixture of cellular proteins is separated using gel electrophoresis according to their molecular masses. After the antibody has bound to the protein of interest within a blot of a gel, the protein is visualized as a dark band. For example, an antibody that recognizes the β-globin polypeptide could be used to specifically detect the β-globin polypeptide in a blot. As shown here, the method of Western blotting can be used to determine the amount and relative size of a particular protein that is produced in a given cell type.


Lane 1 is a sample of proteins isolated from normal red blood cells.

Lane 2 is a sample of proteins isolated from kidney cells. Kidney cells do not produce β globin.

Lane 3 is a sample of proteins isolated from red blood cells from a patient with β-thalassemia. This patient is homozygous for a mutation that results in the shortening of the β-globin polypeptide.

Now here is the question. A protein called troponin contains 334 amino acids. Because each amino acid weighs 120 daltons (Da) (on average), the molecular mass of this protein is about 40,000 Da, or 40 kDa. Troponin functions in muscle cells, and it is not expressed in nerve cells. Draw the expected results of a Western blot for the following samples:

Lane 1: Proteins isolated from muscle cells

Lane 2: Proteins isolated from nerve cells

Lane 3: Proteins isolated from the muscle cells of an individual who is homozygous for a mutation that introduces a stop codon at codon 177

E8. The technique of Western blotting can be used to detect proteins that are translated from a particular mRNA. This method is described in Chapter 18 and also in experimental question E7. Let's suppose a researcher was interested in the effects of mutations on the expression of a structural gene that encodes a protein we will call protein X. This protein is expressed in skin cells and contains 572 amino acids. Its molecular mass is approximately 68,600 Da, or 68.6 kDa. Make a drawing that shows the expected results of a Western blot using proteins isolated from the skin cells obtained from the following individuals:

Lane 1: A normal individual

Lane 2: An individual who is homozygous for a deletion, which removes the promoter for this gene

Lane 3: An individual who is heterozygous in which one gene is normal and the other gene has a mutation that introduces an early stop codon at codon 421

Lane 4: An individual who is homozygous for a mutation that introduces an early stop codon at codon 421

Lane 5: An individual who is homozygous for a mutation that changes codon 198 from a valine codon into a leucine codon

E9. The protein known as tyrosinase is needed to make certain types of pigments. Tyrosinase is composed of a single polypeptide with 511 amino acids. Because each amino acid weighs 120 Da (on average), the molecular mass of this protein is approximately 61,300 Da, or 61.3 kDa. People who carry two defective copies of the tyrosinase gene have the condition known as albinism. They are unable to make pigment in the skin, eyes, and hair. Western blotting can be used to detect proteins that are translated from a particular mRNA. This method is described in Chapter 18 and also in experimental question E7. Skin samples were collected from a pigmented individual (lane 1) and from three unrelated albino individuals (lanes 2, 3, and 4) and subjected to a Western blot analysis using an antibody that recognizes tyrosinase. Explain the possible cause of albinism in the three albino individuals.


E10. Although 61 codons specify the 20 amino acids, most species display a codon bias. This means that certain codons are used much more frequently than other codons. For example, UUA, UUG, CUU, CUC, CUA, and CUG all specify leucine. In yeast, however, the UUG codon is used to specify leucine approximately 80% of the time.

A. The experiment of Figure 13.4 shows the use of an in vitro or cell-free translation system. In this experiment, the RNA, which was used for translation, was chemically synthesized. Instead of using a chemically synthesized RNA, researchers can isolate mRNA from living cells and then add the mRNA to the cell-free translation system. If a researcher isolated mRNA from kangaroo cells and then added it to a cell-free translation system that came from yeast cells, how might the phenomenon of codon bias affect the production of proteins?

B. Discuss potential advantages and disadvantages of codon bias for translation.

## Questions for Student Discussion/Collaboration

1. Discuss why you think the ribosome needs to contain so many proteins and rRNA molecules. Does it seem like a waste of cellular energy to make such a large structure so that translation can occur?
2. Discuss and make a list of the similarities and differences in the events that occur during the initiation, elongation, and termination stages of transcription (see Chapter 12) and translation (Chapter 13).
3. Which events during translation involve molecular recognition of a nucleotide base sequence within RNA? Which events involve recognition between different protein molecules?

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