

Protein Synthesis

Chapter 13

Approximately two-thirds of the organic matter of a typical cell consists of protein. Proteins are the final products of gene expression in most cases, and the term **proteome** refers to the complete set of proteins encoded by a genome, or the total number of proteins found in an organism at one point in time. They are made by translating **messenger RNA (mRNA)** using a molecular decoding machine known as a ribosome. **Translation** refers to the conversion of genetic information in nucleic acid “language” to protein “language.” Proteins are polymers made of amino acids, and each of the three bases of DNA or RNA corresponds to a single amino acid in the protein chain. Not only is the genetic information carried to the ribosome by mRNA, but other RNA molecules take part in the process of translation itself. **Ribosomal RNA (rRNA)** catalyses the formation of the peptide bonds between the individual amino acids and **transfer RNA (tRNA)** acts as an adaptor between the nucleic acid and protein “languages.” Thus, RNA lies at the very heart of gene expression to give proteins.

messenger RNA (mRNA) The type of RNA molecule that carries genetic information from the genes to the rest of the cell.

proteome The total set of proteins encoded by a genome or the total protein complement of an organism.

ribosomal RNA (rRNA) Class of RNA molecule that makes up part of the structure of a ribosome.

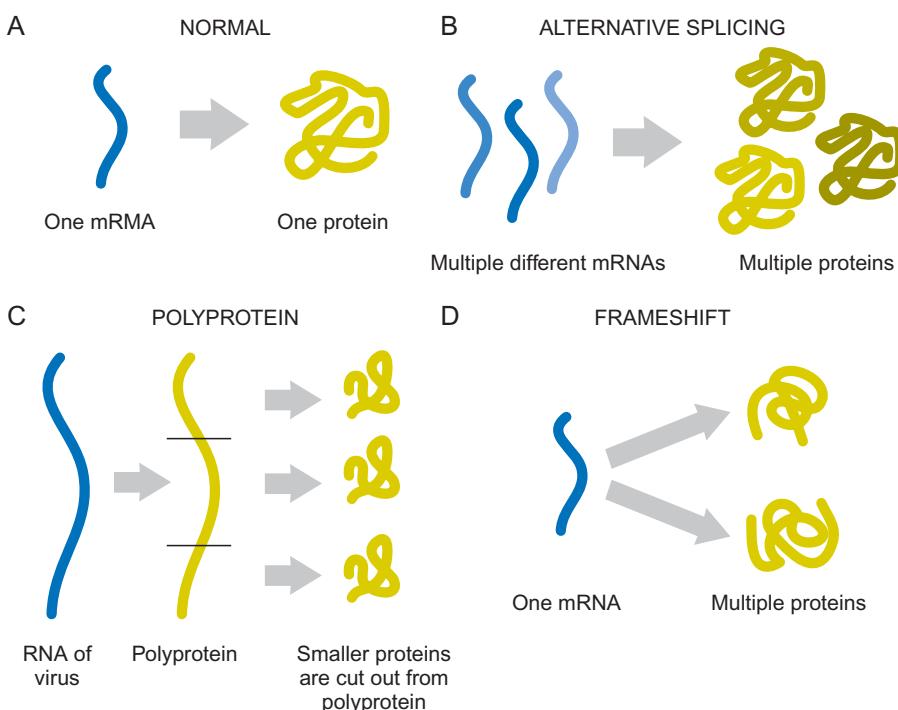
translation Making a protein using the information provided by mRNA.

transfer RNA (tRNA) RNA molecules that carry amino acids to the ribosome.

1. Overview of Protein Synthesis.....398
2. Proteins Are Chains of Amino Acids.....399
3. Decoding the Genetic Information403
4. The Ribosome: The Cell’s Decoding Machine409
5. Three Possible Reading Frames Exist.....412
6. The tRNA Occupies Three Sites During Elongation of the Polypeptide.....418
7. Bacterial mRNA Can Code for Several Proteins421
8. Some Ribosomes Become Stalled and Are Rescued ..424
9. Differences Between Eukaryotic and Prokaryotic Protein Synthesis.....425
10. Protein Synthesis Is Halted When Resources Are Scarce429
11. A Signal Sequence Marks a Protein for Export From the Cell430
12. Protein Synthesis Occurs in Mitochondria and Chloroplasts432
13. Mistranslation Usually Results in Mistakes in Protein Synthesis435
14. Many Antibiotics Work by Inhibiting Protein Synthesis.....435
15. Post-Translational Modifications of Proteins..436
16. Selenocysteine and Pyrrolysine: Rare Amino Acids436
17. Degradation of Proteins439
- Review Questions442
- Further Reading443

FIGURE 13.01
How Many Proteins Per Gene?

(A) Normally each gene is transcribed giving one mRNA and this is translated into a single protein. Variations in the normal theme are (B) alternative splicing, (C) polyproteins, and (D) multiple proteins due to the use of different reading frames.



1. Overview of Protein Synthesis

Each protein in every organism is made using the genetic information stored in the genome. The genetic information is transmitted in two stages. First, the information in the DNA is transcribed into mRNA. Next, the information carried by the mRNA is read to give the sequence of amino acids making up a polypeptide chain. This overall flow of information in biological cells from DNA to RNA to protein is known as the central dogma of molecular biology (see Chapter 3: Nucleic Acids and Proteins, Fig. 3.20) and was first formulated by Sir Francis Crick.

The decoding of mRNA is carried out by the **ribosome**, which binds the mRNA and translates it with tRNA. The ribosome moves along the mRNA reading the message, recruiting the proper tRNA at each step, and synthesizing a new polypeptide chain.

The correspondence between genes and mRNA dictates the way in which proteins are made. An early rule of molecular biology was Beadle and Tatum's dictum: "One gene—one enzyme" (see Chapter 2: Basic Genetics). This rule was later broadened to include other proteins in addition to enzymes. Proteins are therefore often referred to as "**gene products**." However, it must be remembered that some RNA molecules (such as tRNA, rRNA, and small nuclear RNA) are never translated into protein and are therefore also gene products.

Furthermore, instances are now known where one gene may encode multiple proteins (Fig. 13.01). Two relatively widespread cases of this are known—alternative splicing and polyproteins. In eukaryotic cells, the coding sequences of genes are often interrupted by noncoding regions, the introns. These introns are removed by splicing at the level of mRNA. Alternative splicing schemes may generate multiple mRNA molecules and therefore multiple proteins from the same gene. This is especially frequent in higher eukaryotes, in particular vertebrates (see Chapter 12: Processing of RNA). A set of proteins generated in this manner shares much of their sequence and structure.

Ribosomes use the information carried by mRNA to make proteins.

Gene products include proteins as well as noncoding RNA.

gene product End product of gene expression; usually a protein but includes various untranslated RNAs such as rRNA, tRNA, and snRNA.
ribosome The cell's machinery for making proteins.

In eukaryotic cells, most mRNAs only carry information from a single gene and therefore can only be translated into a single protein. This causes problems for certain viruses that infect eukaryotic cells and that have RNA genomes (see Chapter 24: Viruses, Viroids, and Prions). To circumvent the problem, these viruses make a huge “polyprotein” from an extremely long coding sequence in their RNA. This polyprotein is then cut up into several smaller proteins.

Finally, there are occasional oddities, such as the generation of two proteins from the same gene due to frameshifting (discussed later). Despite these exceptions, it is still generally true that most genes give rise to a single protein.

To help understand the very complex process of translation, this chapter is divided into different sections. First, the key components (the genetic code, amino acids, tRNA, and ribosomes) are described in detail. Then the process of translation is broken into four steps: Initiation, elongation, termination, and ribosome recycling. The simpler bacterial process of translation is first explained, and then the more complex eukaryotic version is presented. Finally, protein synthesis in mitochondria and chloroplasts is discussed. The final part of the chapter deals with getting proteins to the correct location in the cell and some of the post-translational modifications and amino acid substitutions that affect a protein’s ultimate function.

Although there are exceptions, most genes give rise to a single protein.

2. Proteins Are Chains of Amino Acids

Proteins consist of linear chains of monomers, known as amino acids, and are folded into a variety of complex 3D shapes. A chain of amino acids is called a polypeptide chain. What is the difference between a polypeptide chain and a protein? Firstly, some proteins consist of more than one polypeptide chain and secondly, many proteins contain additional components such as metal ions or small organic molecules known as co-factors, in addition to their polypeptide portion (see Chapter 14: Protein Structure and Function).

Proteins are linear polymers made from amino acids. Most proteins fold into complex 3D structures.

Twenty different **amino acids** are incorporated when polypeptide chains are synthesized. (Strictly speaking, there are 22, but the other two are inserted only under special conditions—discussed later.) They all have a central carbon atom, the **alpha carbon**, surrounded by an amino group, a carboxyl group, a hydrogen atom, and a side chain or **R-group**, as shown in Fig. 13.02A (proline is an exception—discussed later). The simplest amino acid is **glycine** (Fig. 13.02B) in which the R-group is just a single hydrogen atom. In solution, under physiological conditions, the amino group and the carboxyl group of amino acids are both ionized to give a **zwitterion** or **dipolar ion** with one positive and one negative charge (Fig. 13.02C).

Amino acids are joined together by **peptide bonds** (Fig. 13.03) to give a **polypeptide chain**. The first amino acid in the chain retains its free amino ($-\text{NH}_2$) group, and this end is therefore called the **amino- or N-terminus** of the polypeptide chain. The last amino acid to be added is left with a free carboxyl ($-\text{COOH}$) group, so this end is the **carboxy- or C-terminus**. When synthesized, the polypeptide is elongated from the amino terminus toward the carboxy terminus.

Apart from glycine, amino acids have four different chemical groups surrounding the alpha carbon atom. This is called a **chiral** or **asymmetric center** (Fig. 13.04). Consequently, such amino acids exist as two alternative mirror-image isomers with

alpha (α) carbon Central carbon atom of an amino acid that carries both the amino group and the carboxyl group.

amino acid Monomer from which polypeptide chains are built.

amino- or N-terminus The end of a poly peptide chain that is made first and has a free amino group.

asymmetric center Carbon atom with four different groups attached. This results in optical isomerism.

carboxy- or C-terminus The end of a poly peptide chain that is made last and has a free carboxyl-group.

chiral center Same as asymmetric center.

dipolar ion Same as zwitterion; a molecule with both a positive and a negative charge.

glycine The simplest amino acid.

peptide bond Type of chemical linkage holding amino acids together in a polypeptide chain.

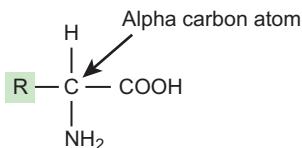
polypeptide chain A polymer that consists of amino acids.

R-group Any unspecified chemical group; in particular, the side chain of an amino acid.

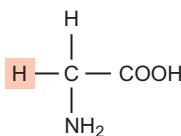
zwitterion Same as dipolar ion; a molecule with both a positive and a negative charge.

FIGURE 13.02
General Structure of Amino Acids

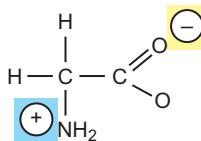
(A) A generalized amino acid contains an alpha carbon atom, an R-group, an $-\text{NH}_2$ group, and a $-\text{COOH}$ group. (B) Glycine is the simplest amino acid with an H atom as the R-group. (C) When glycine is placed in solution at neutral pH it ionizes to form a zwitterion.



A GENERAL AMINO ACID STRUCTURE



B GLYCINE



C GLYCINE IN SOLUTION
(ZWITTERION)

Amino acids occur as pairs of optical isomers. Natural proteins are made from the L-isomers.

The D-isomers of amino acids are found in bacterial cell walls and some antibiotics.

The 20 amino acids that comprise proteins vary greatly in their chemical and physical properties.

different “chirality” or “handedness.” A pair of mirror-image isomers is known as **enantiomers** or **optical isomers**. They are referred to as the **L- and D-forms** because solutions of chiral molecules rotate the plane of polarization of light in either a left-handed (L- = levorotatory) or right-handed (D- = dextrorotatory) direction. The amino acids found in proteins are all of the L-form. Although “L-amino acid” are sometimes referred to as the “natural” isomers, D-amino acids do exist in nature. The **peptidoglycan** that is found in bacterial cell walls contain several different D-amino acids. Several peptide antibiotics, also made by prokaryotes (e.g., bacitracin, polymixin B, and actinomycin D), also contain D-amino acids.

2.1. Twenty Amino Acids Form Biological Polypeptides

The 20 amino acids found in proteins possess a variety of different chemical groups (Fig. 13.05). This wide choice of possible monomers makes proteins very versatile, with a wide range of properties and capabilities, including a great variety of possible 3D structures. The amino acids may be classified into groups depending on their physical and chemical characteristics. The major division is between those with **hydrophilic** (water-loving) and those with **hydrophobic** (water-hating) R-groups. Glycine has only a single hydrogen atom as its side chain, so it does not really fit into either group.

The hydrophilic amino acids may be subdivided into basic, acidic, and neutral. Basic amino acids contribute a positive charge to the protein, whereas acidic residues provide a negative charge. Strictly, this refers to the situation in solution within

enantiomers A pair of mirror-image optical isomers (i.e., D- and L-isomers).

hydrophilic Water-loving; readily dissolves in water.

hydrophobic Water-hating; repelled by water and dissolves in water only with great difficulty.

L- and D-forms The two isomeric forms of an optically active substance; also called L- and D-isomers.

optical isomers Isomers where the molecules differ only in their 3D arrangement and consequently affect the rotation of polarized light.

peptidoglycan Polymer that makes up eubacterial cell walls; consists of long chains of sugar derivatives, cross-linked at intervals with short chains of amino acids.

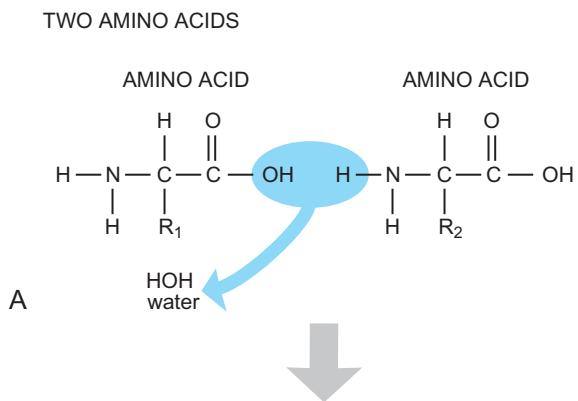
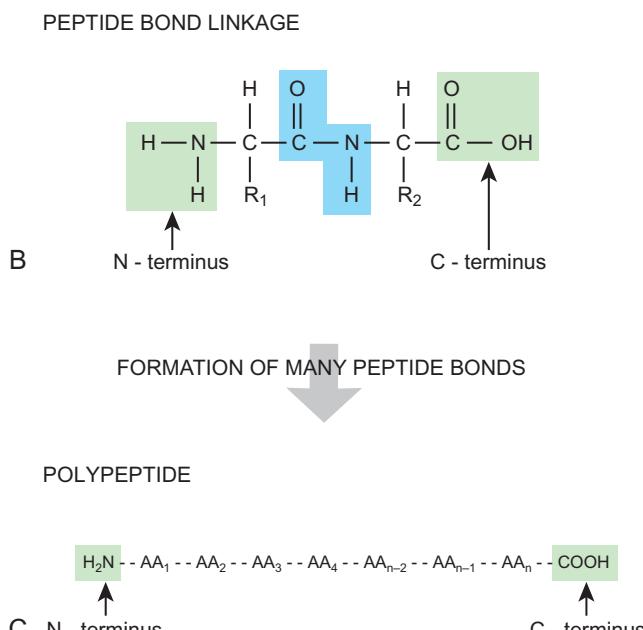


FIGURE 13.03
Polypeptide Chain Is Made of Amino Acids

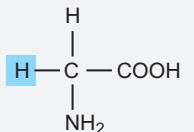
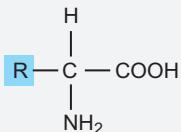
(A) Two generic amino acids are shown. The R-groups, R_1 and R_2 , represent the side chains of any of the 20 different amino acids that make up proteins. Each amino acid has an amino ($-NH_2$) and carboxyl ($-COOH$) group. (B) A peptide bond is formed between one $-NH_2$ and one $-COOH$ group, with water eliminated in the process (dehydration synthesis). (C) Successive amino acids are joined in a similar manner by peptide bonds to form a polypeptide. The polypeptide contains an amino (or N-) terminus and a carboxy (or C-) terminus.



■ **FIGURE 13.04**
The L- and D-Forms of an Amino Acid

The four groups are arranged around the alpha carbon differently in the L-form and the D-form of an amino acid. Although they share the same molecular formula, one is the mirror image of the other.

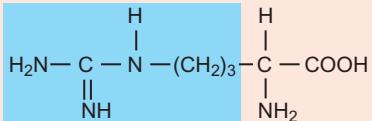
GENERAL STRUCTURE OF AMINO ACID



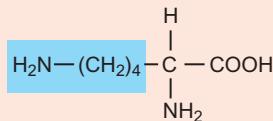
GLYCINE - A SIMPLE AMINO ACID
(HYDROPHOBIC)

HYDROPHILIC AMINO ACIDS

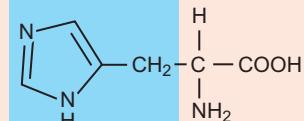
BASIC AMINO ACIDS



ARGININE

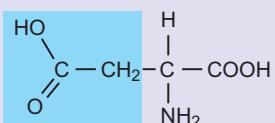


LYSINE

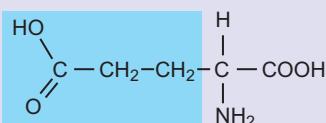


HISTIDINE

ACIDIC AMINO ACIDS

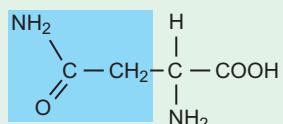


ASPARTIC ACID

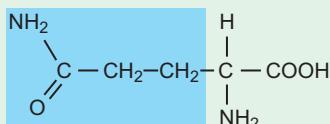


GLUTAMIC ACID

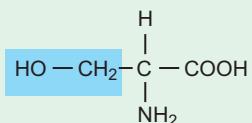
NEUTRAL POLAR AMINO ACIDS



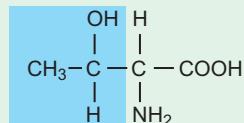
ASPARAGINE



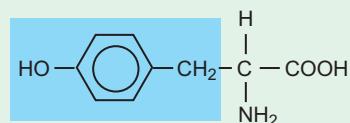
GLUTAMINE



SERINE

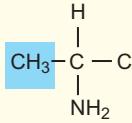


THREONINE

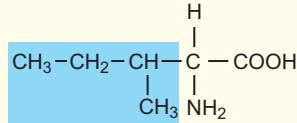


TYROSINE

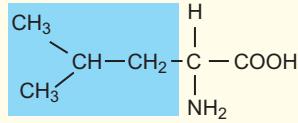
HYDROPHOBIC AMINO ACIDS



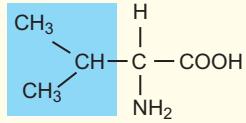
ALANINE



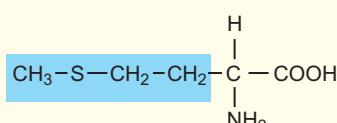
ISOLEUCINE



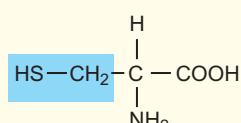
LEUCINE



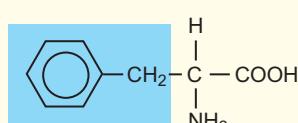
VALINE



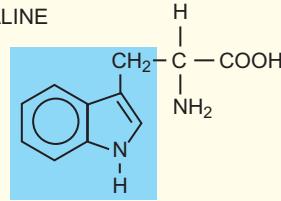
METHIONINE



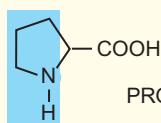
CYSTEINE



PHENYLALANINE



TRYPTOPHAN



PROLINE

FIGURE 13.05

The 20 Amino Acids Found in Proteins

Amino acids can be grouped by their physical and chemical properties. The R-group for each amino acid is highlighted.

the physiological pH range. Neutral polar residues have side chains that are capable of forming hydrogen bonds. The side chains of the hydrophilic amino acids carry chemical groups that can take part in reactions. The active sites of enzymes (discussed later) often contain serine, histidine, and basic or acidic amino acids.

The hydrophobic amino acids may be subdivided into those that are aliphatic (Ala, Leu, Ile, Val, and Met) and those containing aromatic rings (Phe, Trp, and Tyr). These amino acid residues are largely structural in function, except for tyrosine, which has a hydroxyl group attached to its aromatic ring and can therefore take part in a variety of reactions. The classification of tyrosine is ambiguous as its hydroxyl group is polar in nature. Proline has a nonaromatic ring and contains a secondary amino group (—NH—) rather than a primary amino group (NH_2) like the other imino acids. Such structures are sometimes referred to as imino acids (rather than amino acids).

Two of the hydrophobic amino acids, Met and Cys, contain sulfur. When a polypeptide chain is first synthesized, methionine is always the first amino acid, although it may be trimmed off later. Cysteine is important for 3D structure as it forms disulfide bonds. The free **sulphydryl group** of cysteine is highly reactive and is often used in enzyme active sites or to attach various chemical groups to proteins (see Chapter 14: Protein Structure and Function).

The 20 different amino acids normally found in proteins may be represented by both three-letter and one-letter abbreviations (Table 13.01). These mostly correspond to the first letter(s) of the name, but since several amino acids start with the same letter of the alphabet, the others need a little imagination. They are used especially when writing out protein sequences. The amides, asparagine and glutamine, are relatively unstable and break down easily into the corresponding acids, aspartate and glutamate. Consequently, some analyses do not distinguish the acids from their amides. The abbreviations Asx and Glx were invented for these ambiguous pairs.

3. Decoding the Genetic Information

There are 20 amino acids in proteins but only four different bases in the mRNA, so one base of a nucleic acid cannot code for a single amino acid when making a protein. During translation, the bases of mRNA are read off in groups of three, which are known as **codons**. Each codon represents a particular amino acid. Four different bases give 64 possible groups of three bases; that is, there are 64 different codons in the **genetic code**. Because there are only 20 different amino acids, some are encoded by more than one codon. In addition, three of the codons are used for punctuation. Those are the **stop codons** that signal the end of a polypeptide chain. Fig. 13.06 shows nature's genetic code.

Each amino acid in a protein is encoded by three bases in the DNA or RNA sequence.

To read the codons, tRNAs recognize the codon on the mRNA. At one end, each tRNA has an **anticodon** consisting of three bases that are complementary to the three bases of the corresponding codon on the mRNA. The codon and anticodon recognize each other by base pairing and are held together by hydrogen bonds (Fig. 13.07). At its other end, each tRNA carries the amino acid encoded by the codon it recognizes. This amino acid is sometimes known as the tRNA's "cognate" amino acid.

The anticodon of tRNA recognizes the codon on mRNA by base pairing.

The genetic code is not quite universal. Despite this, the term "**universal genetic code**" is used to refer to the codon table shown in Fig. 13.06, since it applies to almost all organisms. Rarely, exceptions to the code are found in some protozoans and mycoplasmas and in the mitochondrial genome of animals and fungi (Table 13.02). Mycoplasmas are parasitic bacteria with unusually small genomes. *Paramecium* and *Euplotes* are ciliated protozoans and *Candida* is a type of yeast.

Each tRNA carries one particular amino acid.

anticodon Group of three complementary bases on tRNA that recognize and bind to a codon on the mRNA.

codon Group of three RNA or DNA bases that encodes a single amino acid.

genetic code System for encoding amino acids as groups of three bases (codons) of DNA or RNA.

stop codon Codon that signals the end of a protein.

sulphydryl group —SH ; Chemical group of sulfur and hydrogen.

universal genetic code Version of the genetic code used by almost all organisms.

TABLE 13.01 Amino Acids and Their Properties

Amino Acid	3-Letter Code	1-Letter Code	Physical Properties
Alanine	Ala	A	hydrophobic
Arginine	Arg	R	basic
Asparagine	Asn	N	neutral polar
Aspartic acid	Asp	D	acidic
Cysteine	Cys	C	hydrophobic
Glutamic acid	Glu	E	acidic
Glutamine	Gln	Q	neutral polar
Glycine	Gly	G	—
Histidine	His	H	basic
Isoleucine	Ile	I	hydrophobic
Leucine	Leu	L	hydrophobic
Lysine	Lys	K	basic
Methionine	Met	M	hydrophobic
Phenylalanine	Phe	F	hydrophobic
Proline	Pro	P	hydrophobic
Serine	Ser	S	neutral polar
Threonine	Thr	T	neutral polar
Tryptophan	Trp	W	hydrophobic
Tyrosine	Tyr	Y	neutral polar/hydrophobic
Valine	Val	V	hydrophobic
Aspartic acid or Asparagine	Asx	B	
Glutamic acid or Glutamine	Glx	Z	
Unspecified amino acid		X	

FIGURE 13.06 The Genetic Code

The 64 codons as found in mRNA are shown with their corresponding amino acids. As usual, bases are read from 5' to 3' so that the first base is at the 5' end of the codon. Three codons (UAA, UAG, and UGA) have no cognate amino acid but signal stop. AUG (encoding methionine) and, less often, GUG (encoding valine) act as start codons. To locate a codon, find the first base in the vertical column on the left, the second base in the horizontal row at the top, and the third base in the vertical column on the right.

2nd (middle) base					
1st base	U	C	A	G	3rd base
U	UUU Phe UUC Phe UUA Leu UUG Leu	UCU Ser UCC Ser UCA Ser UCG Ser	UAU Tyr UAC Tyr UAA stop UAG stop	UGU Cys UGC Cys UGA stop UGG Trp	U C A G
	CUU Leu CUC Leu CUA Leu CUG Leu	CCU Pro CCC Pro CCA Pro CCG Pro	CAU His CAC His CAA Gln CAG Gln	CGU Arg CGC Arg CGA Arg CGG Arg	U C A G
A	AUU Ile AUC Ile AUA Ile AUG Met	ACU Thr ACC Thr ACA Thr ACG Thr	AAU Asn AAC Asn AAA Lys AAG Lys	AGU Ser AGC Ser AGA Arg AGG Arg	U C A G
	GUU Val GUC Val GUA Val GUG Val	GCU Ala GCC Ala GCA Ala GCG Ala	GAU Asp GAC Asp GAA Glu GAG Glu	GGU Gly GGC Gly GGA Gly GGG Gly	U C A G

Minor variations in the genetic code are found in mitochondria and certain microorganisms.

Note that there is no general mitochondrial genetic code. Although fungal and animal mitochondria share similarities (e.g., UGA = Trp), there are also differences (e.g., CUA = Thr in fungi but Leu in animals). Mammalian mitochondria also have a reduced set of tRNAs that results in an increase in the number of tRNAs that read multiple codons. However, plant mitochondria and chloroplasts use the universal genetic code.

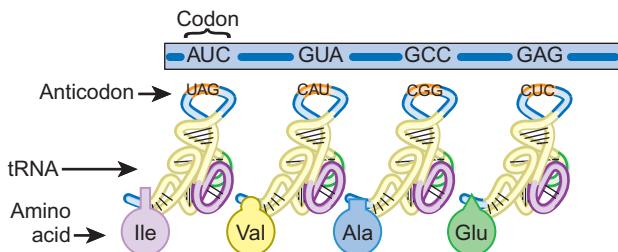


FIGURE 13.07
Transfer RNA Recognizes Codons

Several tRNAs are seen bound to mRNA codons by their anticodons. Each tRNA carries a different amino acid at the end of the adaptor stem. This diagram is intended to show the principle of mRNA decoding. It does NOT illustrate the actual mechanism of protein synthesis. In real life, the codons are contiguous and there are no spacers in between and only two tRNAs are bound at any given time.

TABLE 13.02 Exceptions to the Universal Genetic Code

Exceptions in the Chromosomal Genome					
Codon	Universal	Mycoplasma	Paramecium	Euplotes	Candida
UGA	Stop	Trp	Stop	Cys	Stop
UAA/UAG	Stop	Stop	Gln	Stop	Stop
CUG	Leu	Leu	Leu	Leu	Ser

Exceptions in the Mitochondrial Genome					
Codon	Universal	Fungi	Protozoa	Mammals	Flatworm
UGA	Stop	Trp	Trp	Trp	Trp
UAA	Stop	Stop	Stop	Stop	Tyr
AUA	Ile	Met	Met	Met	Ile
AGA/AGG	Arg	Arg	Arg	Stop	Ser
AAA	Lys	Lys	Lys	Lys	Asn
CUA	Leu	Thr	Leu	Leu	Leu

3.1. Transfer RNA Forms a Folded "L" Shape With Modified Bases

Transfer RNA molecules are about 80 nucleotides in length. About half the bases are paired to form double helical segments. A typical tRNA has four short base-paired stems and three loops (Fig. 13.08). This is shown best in the **cloverleaf structure**, intended to reveal details of base pairing, which shows the tRNA spread out flat in only two dimensions. (Such a diagram is sometimes called a secondary structure map). The tRNA is in reality folded up further to give an L-shaped 3D structure, in which the T_ψC-loop (or T-loop) and the D-loop are pushed together. The anticodon and attached amino acid are located at the two ends of the L-structure. Different tRNA molecules vary considerably in sequence, but they all conform to this same overall structure. Variations in length (from 73–93 nucleotides) occur, due mostly to the variable loop.

The **acceptor stem** is made by pairing of the 5' end, which almost always ends in G and is phosphorylated, and the 3' end, which ends in CCA-OH (Box 13.01). The amino acid is bound to the 3'-hydroxyl group of the adenosine at the free 3' end of the acceptor stem. The anticodon is about halfway round the sequence, in the

acceptor stem Base-paired stem of tRNA to which the amino acid is attached.

anticodon loop Loop of tRNA molecule that contains the anticodon.

cloverleaf structure 2D structure showing base pairing in a tRNA molecule.

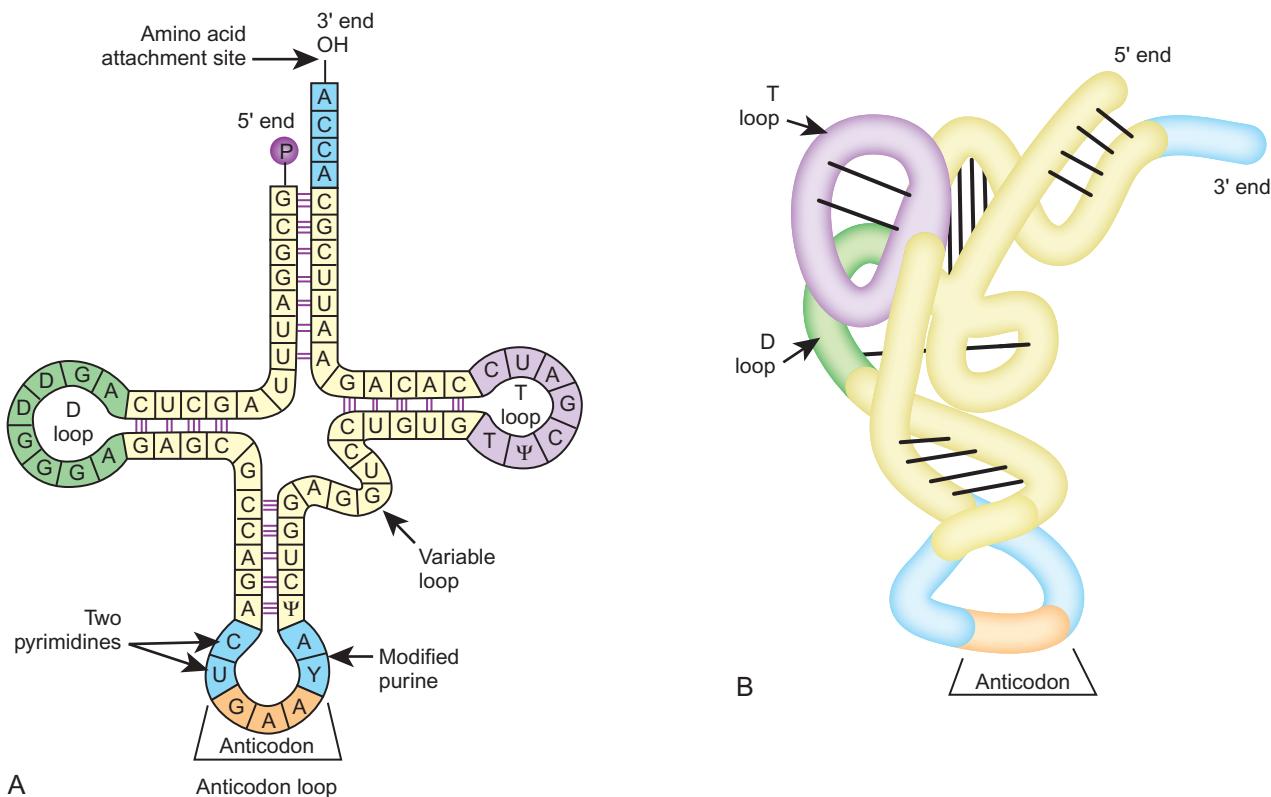


FIGURE 13.08
Structure of tRNA

(A) A planar view (secondary structure) of a tRNA shows its cloverleaf structure comprised of the 3' and 5' acceptor stem, the T- (or T_ψC) and D-loops, and the anticodon loop. A variable loop, which varies in length in different tRNA molecules, is also found. (B) The folded (tertiary) configuration resembles an "L."

anticodon loop. This consists of seven bases with the three anticodon bases in the middle. The anticodon is always preceded, on the 5' side, by two pyrimidines and followed by a modified purine (Fig. 13.08).

The other two loops of tRNA are named after **modified bases**. These bases are modified after the RNA has been transcribed and are especially common in tRNA, although they are also found in other classes of RNA. The T_ψC-loop contains “ψ” (spelled “psi” but pronounced “sigh”), which stands for pseudouracil; and the D-loop or DHU-loop has “D” for dihydrouracil (Fig. 13.09). In pseudouridine, the uracil itself is not altered, but is attached to ribose by carbon-5 instead of nitrogen-1, as in normal uridine. Note that thymine (=5-methyl uracil), which is normally only found in DNA, is also found in the T_ψC-loop of tRNA, where it is attached to ribose and is made by methylation of uracil after transcription. In addition to uracil, guanine, adenine, and cytosine may also be modified by methylation. Other modifications include **inosine**, which is a nucleotide with a base called hypoxanthine. However, it is written as “I” in sequences and sometimes called I-base to avoid confusion. Similarly, the bases of queuosine and wynosine are referred to as Q-base and Y-base. Methylation of bases prevents pairing of certain bases and also aids binding of ribosomal proteins. The modified bases are required for proper folding and operation of the tRNA. For example, the T_ψC-loop and the D-loop are needed for binding to the ribosome and other protein factors involved in translation.

Some bases are chemically modified after RNA has been transcribed. Modified bases are especially common in tRNA.

modified base Nucleic acid base that is chemically altered after the nucleic acid has been synthesized.
inosine An unusual modified nucleoside derived from guanosine.

Box 13.01 CCA-OH Tails of tRNA

Not every tRNA gene encodes the CCA tail. In some higher organisms, a special enzyme called the CCA-adding enzyme adds 5' –CCA 3' after transcription is complete. CCA-adding enzyme is found in bacteria and functions to repair any tRNA that has been degraded by nucleases, even though most eubacterial tRNA genes encode the CCA tail. The crystal structure of CCA-adding enzyme shows that it is quite a remarkable enzyme. Most enzymes catalyze the same exact reaction using the same exact mechanism. This enzyme is capable of recognizing all of the different tRNA molecules, meaning the binding pocket for tRNA is flexible. This enzyme also adds each nucleotide individually, which means that the enzyme has to also recognize whether the tRNA has a partial or complete CCA tail. If the CCA tail is partially added, CCA-adding enzyme then discerns whether it needs to add the second C or third A. Even more amazing, the enzyme does not have a nucleotide template, determining which nucleotide to add and when it is accomplished with the protein structure alone. The CCA-adding enzyme can use either ATP or CTP to add the tRNA tail but is able to exclude UTP and GTP. Additionally, some CCA-adding enzymes tag dysfunctional tRNAs, which are unstable and contain acceptor stem mismatches, for degradation by addition of a 3' CCACCA.

3.2. Some tRNA Molecules Read More Than One Codon

Each tRNA carries only a single amino acid, so at least 20 different tRNAs are needed for the 20 different amino acids. On the other hand (excluding the stop codons), there are 61 codons to be recognized, since some amino acids have more than one codon. Indeed, some tRNAs can read more than one codon, though, of course, these must all code for the same amino acid. The minimum set of different tRNA molecules needed to read all 61 codons is 31. The actual number found is usually somewhat higher and varies from species to species.

Since only complementary bases can pair, how does a tRNA with one anticodon read more than one codon? Remember that the standard base pairing rules apply to bases that form part of a DNA double helix. Since the codon and anticodon do not form a standard double helix, slightly different rules for base pairing apply. The last two bases of the tRNA anticodon, which pair with the first two bases of the mRNA codon, pair strictly according to normal rules. However, the first base of the tRNA anticodon (which pairs with the third base of the mRNA codon) can wobble around a little because it is not squeezed between other bases as in a helix structure. Consequently, the codon/anticodon base pairing rules are known as the **wobble rules** ([Table 13.03](#)).

If the first anticodon base is G it can pair with C, as usual, or, in wobble mode, with U. For example, tRNA for histidine, with GUG as anticodon, can recognize both the CAC and CAU codons. Similarly, if the first anticodon base is U, it can pair with A or G. Whenever an amino acid is encoded by a pair of codons, the third codon bases are U and C (e.g., histidine, and tyrosine) or A and G (e.g., lysine, and glutamic acid), but never other combinations. Similarly, those privileged amino acids with four or six codons may be regarded as having two or three such pairs. Due to wobble pairing, only a single tRNA is needed to read each such pair of codons. It is possible for a single tRNA to read three codons by making use of inosine. The I-base is occasionally used as the first anticodon base because it can pair with any of U, C, or A.

Wobbling between the anticodon and codon allows some tRNA molecules to read more than one codon.

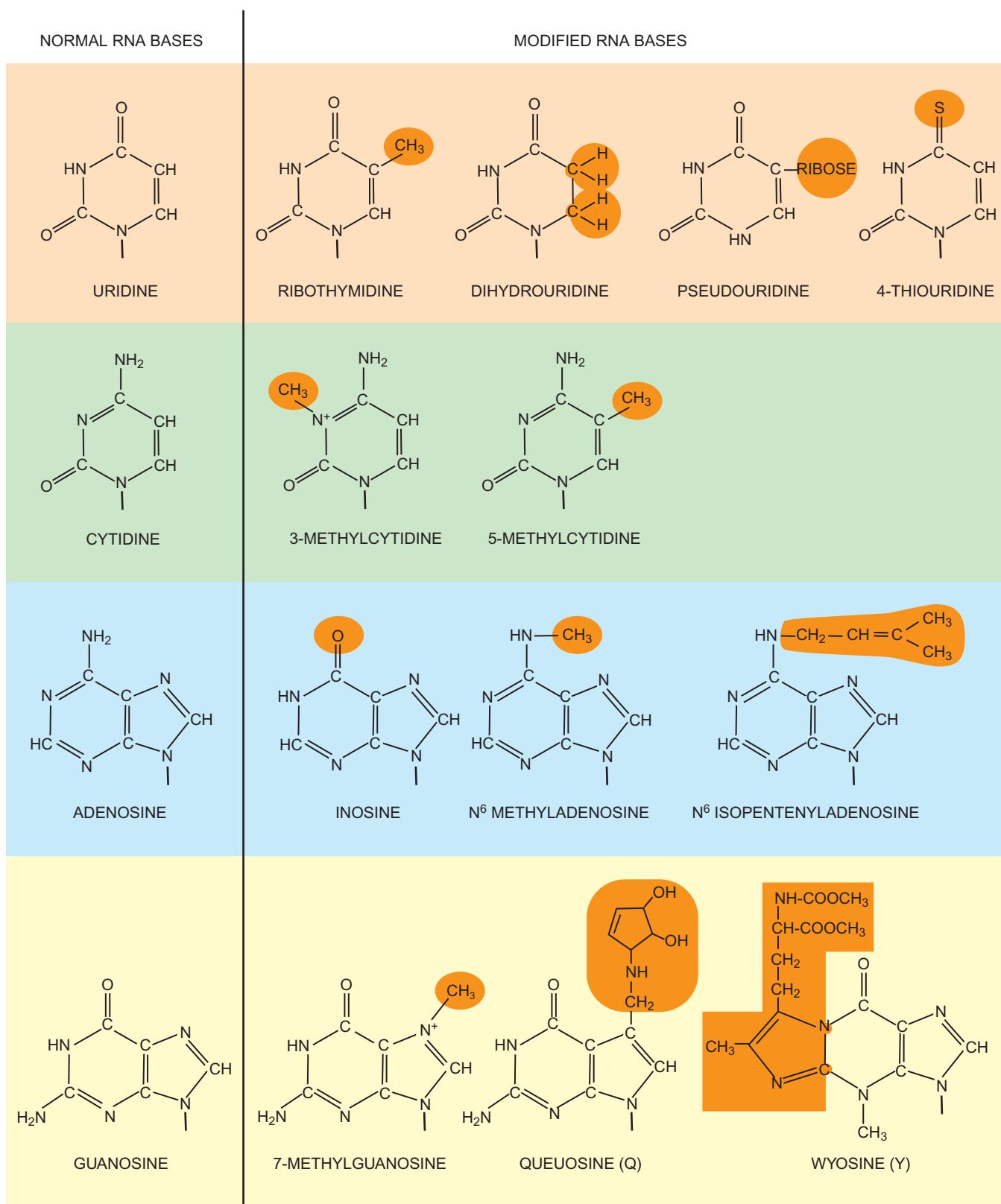


FIGURE 13.09
Modified Bases in tRNA

All four bases normally found in RNA have modified derivatives that may be found in tRNA. The names given are those of the corresponding nucleosides (i.e., base plus ribose).

TABLE 13.03 Wobble Rules for Codon/Anticodon Pairing

Pairs with Third Codon Base		
First Anticodon Base	Normal	by wobble
G	C	U
U	A	G
I	—	C or U or A
C	G	no wobble
A	U	no wobble

3.3. Charging the tRNA with the Amino Acid

For each tRNA there is a specific enzyme that recognizes both the tRNA and the corresponding amino acid. These enzymes, known as **aminoacyl tRNA synthetases**, attach the cognate amino acid to the tRNA. This is called “charging the tRNA.” Empty tRNA is known as **uncharged tRNA** while tRNA with its amino acid is **charged tRNA**.

Charging occurs in two steps (Fig. 13.10). First the amino acid reacts with ATP to form aminoacyl-AMP (also known as aminoacyl-adenylate). Next, the aminoacyl-group is transferred to the 3' end of the tRNA.

- (1) amino acid + ATP → aminoacyl-AMP + PPi
- (2) aminoacyl-AMP + tRNA → aminoacyl-tRNA + AMP

The aminoacyl tRNA synthetases are highly specific for both the correct amino acid and the correct tRNA. In some cases, they recognize the correct tRNA by its anticodon and in others by the sequence of the acceptor stem. Some aminoacyl tRNA synthetases recognize both regions of the tRNA. Fig. 13.11 shows an aminoacyl tRNA synthetase bound to its cognate tRNA.

A specific enzyme, called aminoacyl tRNA synthetase, attaches the correct amino acid to the correct tRNA.

4. The Ribosome: The Cell's Decoding Machine

The decoding process is carried out by the ribosome, a submicroscopic machine that binds mRNA and charged tRNA molecules. The mRNA is translated into protein starting at the 5' end. After binding to the mRNA, the ribosome moves along it, adding a new amino acid to the growing polypeptide chain each time it reads a codon. Each codon on the mRNA is actually read by an anticodon on the corresponding tRNA and so the information in the mRNA is used to synthesize a polypeptide chain from the amino acids carried by the tRNAs.

Ribosomes consist of RNA plus proteins and their role is to synthesize new proteins.

The ribosome and its components were originally analyzed by ultracentrifugation. Consequently, sizes are referred to in Svedberg units (S-value), which measure sedimentation velocity. Although higher S-values indicate larger particles, the S-value is not directly proportional to molecular weight. The **bacterial (70S) ribosome** consists of two subunits, the **50S or large subunit** and the **30S or small subunit** (Fig. 13.12). **Eukaryotic (80S) ribosomes** are somewhat larger, consisting of **60S** and **40S** subunits (discussed later).

The bacterial ribosome contains three rRNA molecules that make up about two-thirds of its weight and about 50 smallish proteins that make up the remaining third.

aminoacyl tRNA synthetase Enzyme that attaches an amino acid to tRNA.

bacterial (70S) ribosome Type of ribosome found in bacterial cells.

charged tRNA tRNA with an amino acid attached.

eukaryotic (80S) ribosome Type of ribosome found in cytoplasm of eukaryotic cell and encoded by genes in the nucleus.

30S subunit Small subunit of a 70S ribosome.

40S subunit Small subunit of an 80S ribosome.

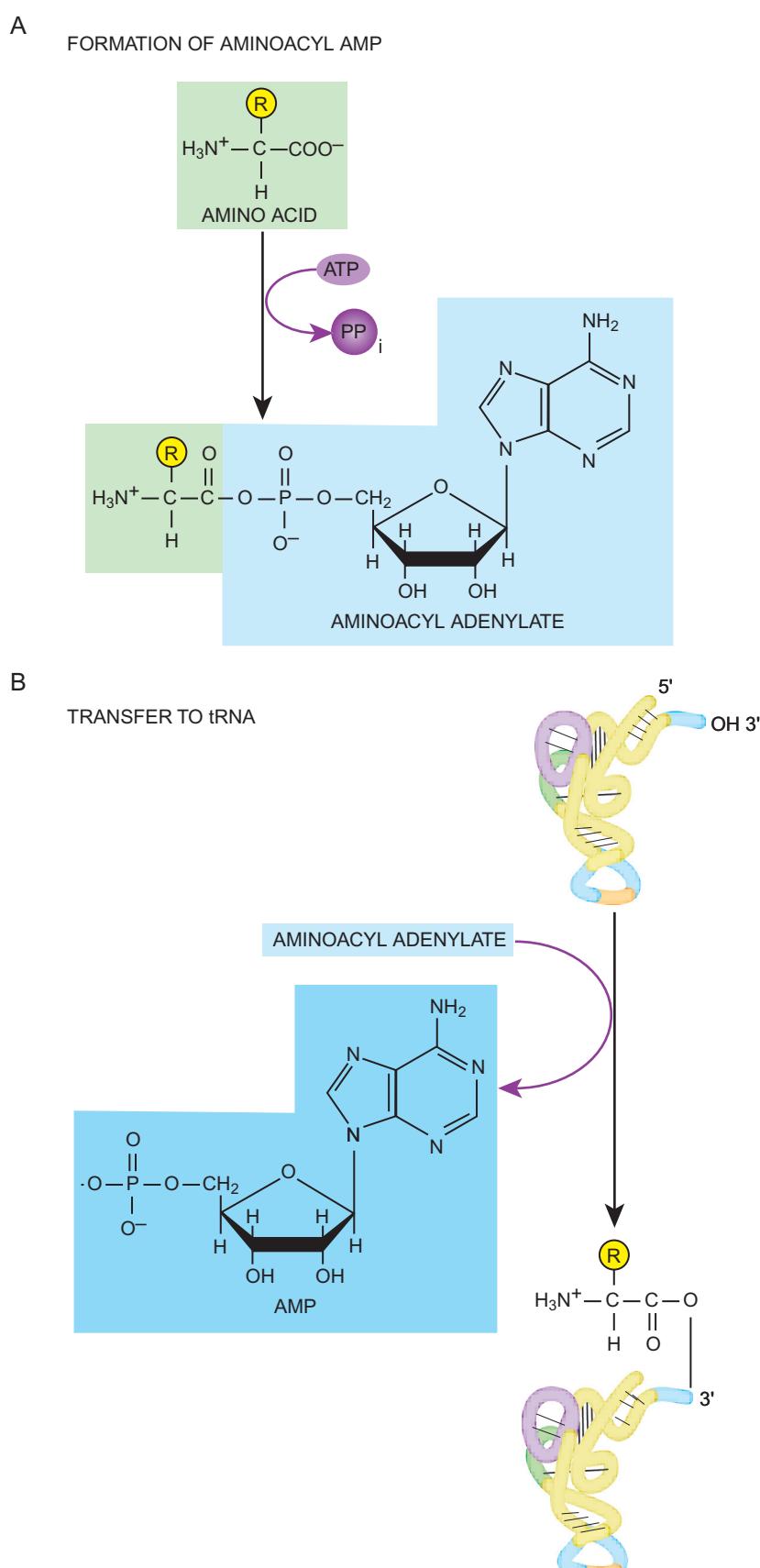
50S subunit Large subunit of a 70S ribosome.

60S subunit Large subunit of an 80S ribosome.

uncharged tRNA tRNA without an amino acid attached.

FIGURE 13.10
Charging tRNA With the Amino Acid

This two-step procedure begins (A) by attachment of the amino acid to adenosine monophosphate (AMP) to give aminoacyl-AMP or aminoacyl-adenylate. This involves splitting ATP and the release of inorganic pyrophosphate. Then, in the second step (B), the amino acid is transferred to the hydroxyl group of the ribose at the 3' end of the tRNA, yielding AMP as a byproduct.



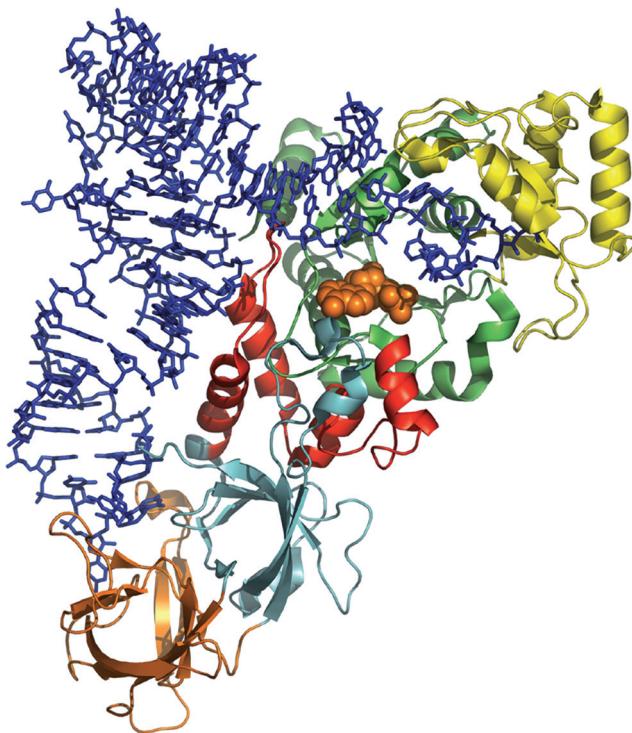


FIGURE 13.11
Glutamine tRNA Bound to its Aminoacyl tRNA Synthetase

Structure of glutaminyl-tRNA synthetase bound to tRNA(Gln) and a glutamyl adenylate analog. The analog is in orange and is shown in a space-filling representation. The tRNA is depicted in dark blue. Domains of the enzyme are color-coded as follows: Active-site Rossman fold, green; acceptor-end binding domain, yellow; connecting helical subdomain, red; proximal beta-barrel, light blue; distal beta-barrel, orange. The image was made in PyMol by John Perona, Department of Chemistry and Biochemistry, University of California at Santa Barbara.

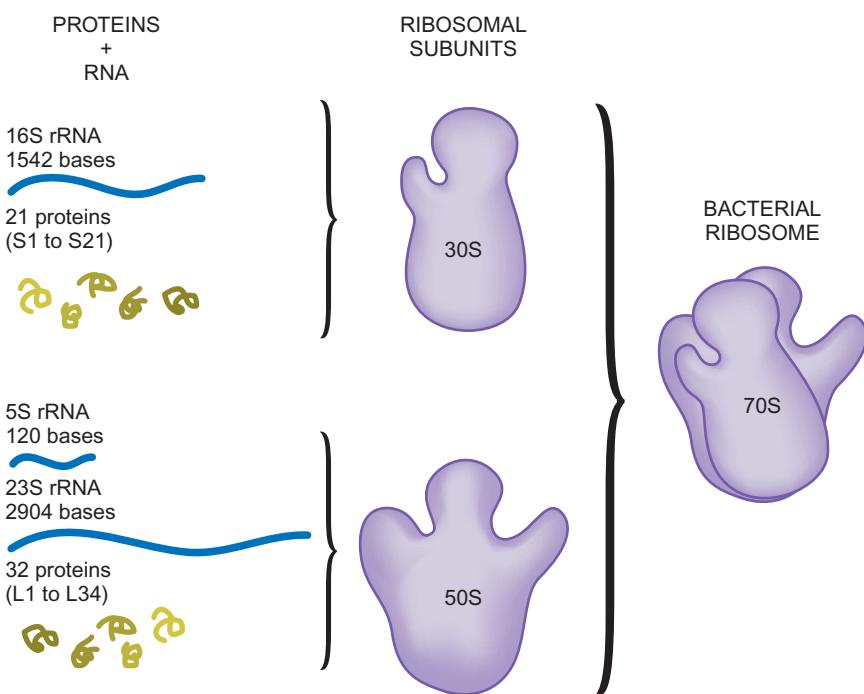


FIGURE 13.12
Components of a Bacterial Ribosome

The ribosome is composed of 30S and 50S subunits. The 30S subunit consists of one 16S rRNA together with 21 proteins, and the 50S subunit has a 5S and 23S rRNA plus 34 proteins.

The 30S subunit contains the 16S rRNA and the 50S subunit contains the 5S and 23S rRNA (Fig. 13.12). The 3D structure of a 70S ribosome is shown in Figs. 13.13 and 13.14. The structure of the eukaryotic ribosome has also been solved (Fig. 13.15).

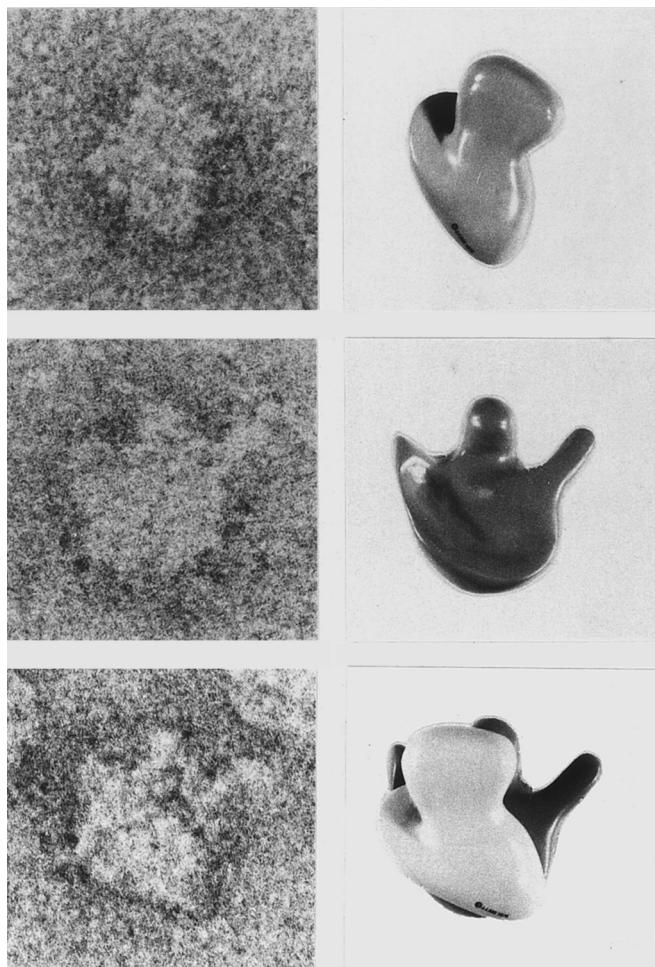
The rRNA molecules have highly defined secondary structures with many stems and loops (Fig. 13.16). Although it was originally believed to have a largely structural role, recent work indicates that the rRNA is responsible for most of the critical reactions of protein synthesis. In particular, the 23S rRNA of the large subunit is a

The peptide bond linking amino acids in the growing protein is made by the largest rRNA, which acts as a ribozyme.

ribozyme RNA molecule that acts as an enzyme.

FIGURE 13.13
3D Structure of a Ribosome by EM

This structure was deduced from negatively stained electron microscope images of a bacterial 70S ribosome.



ribozyme that catalyzes the synthesis of the peptide bonds between the amino acids; that is, it is the **peptidyl transferase**. Indeed, X-ray crystallography of the 50S subunit has shown that no ribosomal proteins are close enough to the catalytic center to take part in the reaction. Alteration by mutation of the catalytic residues in typical ribozymes either abolishes activity completely or reduces it by many-fold. However, the peptidyl transferase center of 23S rRNA behaves in an atypical manner. Alteration of A2451 or G2447 (*E. coli* numbering) did not greatly reduce catalytic activity, although these residues are present in the catalytic center. These results suggest that the ribosome does not operate via direct chemical catalysis. Rather, the ribosome acts by correctly positioning the two substrates. The activated aminoacyl-tRNA then reacts spontaneously with the end of the growing polypeptide chain.

5. Three Possible Reading Frames Exist

Before mRNA is translated into protein, the issue of **reading frame** must be dealt with. The bases of mRNA are read off in groups of three, with each codon corresponding to one amino acid. But how is the sequence divided into codons? For any given nucleotide sequence there are three alternatives, depending on where the start is considered to be. Consider the following sequence:

GAAAUGUAUGCAUGCCAAAGGAGGCAUCUAAGG

peptidyl transferase Enzyme activity on the ribosome that makes peptide bonds; actually 23S rRNA (bacterial) or 28S rRNA (eukaryotic).
reading frame One of three alternative ways of dividing up a sequence of bases in DNA or RNA into codons.

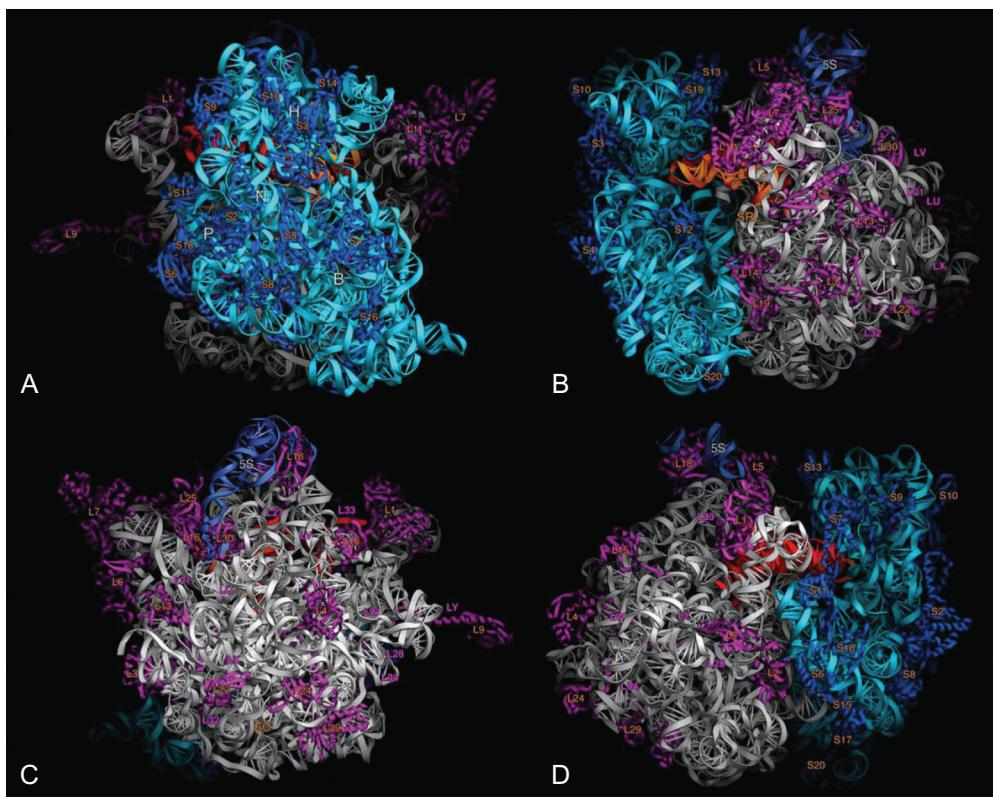


FIGURE 13.14
3D Structure of a Ribosome by X-Ray

Views of the structure of the *Thermus thermophilus* 70S ribosome. A, B, C, and D are successive 90° rotations about the vertical axis. (A) view from the back of the 30S subunit. H, head; P, platform; N, neck; B, body. (B) view from the right-hand side, showing the subunit interface cavity, with the 30S subunit on the left and the 50S on the right. The anticodon arm of the A tRNA (gold) is visible in the interface cavity. (C) View from the back of the 50S subunit. EC, the end of the polypeptide exit channel. (D) View from the left-hand side, with the 50S subunit on the left and the 30S on the right. The anticodon arm of the E-tRNA (red) is partly visible. The different molecular components are colored for identification: Cyan, 16S rRNA; grey, 23S rRNA; light blue, 5S rRNA; dark blue, 30S proteins; magenta, 50S proteins. (Credit: Yusupov et al., *Crystal Structure of the Ribosome at 5.5 Å Resolution*. *Science* 292 (2001) 883–96.)

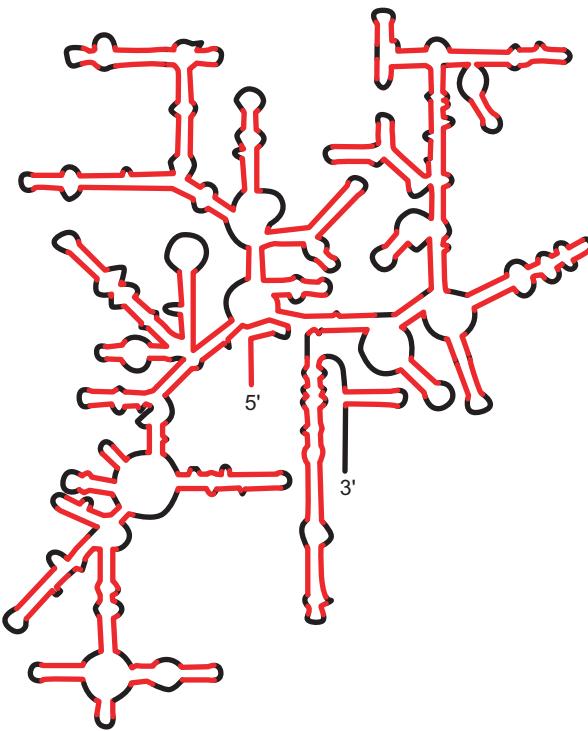


FIGURE 13.15
3D Structure of a Eukaryotic Ribosome

Crystal structure of the eukaryotic ribosome. Image from the RCSB PDB (www.pdb.org) of PDB ID 3O30 (Ben-Shem, A, et al. (2010) *Science* 330: 1203–1209).

FIGURE 13.16
Secondary Structure of an rRNA

The 16S rRNA from the small ribosomal subunit of *E. coli* is complex with extensive secondary structure, forming loops and stems. Red indicates regions of base pairing.



If we start at base #1 we get the following codons:

GAA | AUG | UAU | GCA | UGC | CAA | AGG | AGG | CAU | CUA | AGG

If translated this would give the following amino acid sequence:

Glu | Met | Tyr | Ala | Cys | Gln | Arg | Arg | His | Leu | Arg

If we start at base #2 we get the following codons:

G | AAA | UGU | AUG | CAU | GCC | AAA | GGA | GGC | AUC | UAA | GG

If translated this would give the following amino acid sequence:

- | Lys | Cys | Met | His | Ala | Lys | Gly | Gly | Ile | Stop | -

And if we start at base #3 we get the following codons:

GA | AAU | GUA | UGC | AUG | CCA | AAG | GAG | GCA | UCU | AAG | G

If translated this would give the following amino acid sequence:

- | Asn | Val | Cys | Met | Pro | Lys | Glu | Ala | Ser | Lys | -

Since the genetic code is read in groups of three bases, any nucleic acid sequence contains three possible reading frames.

Between the very front of the mRNA and the coding sequence is a short untranslated region called the 5'-UTR.

Each set of codons gives a translation completely out of step with each of the others. These three possibilities are known as reading frames. As there are three bases in a codon, there are only three possible reading frames. Changing the reading frame by three (or a multiple of three) provides the same sequence as the first example above.

Any sequence of DNA or RNA, beginning with a start codon, and which can, at least theoretically, be translated into a protein, is known as an **open reading frame**, often abbreviated to (and pronounced!) **ORF**. Since ORFs are derived by examining nucleic acid sequences, deciding whether an ORF is a genuine protein coding sequence requires further information. Any mRNA will possess several possible ORFs. The correct one is what matters. Note that the message on an mRNA molecule does not start exactly at the 5' end. Between the 5' end and the coding sequence is a short region that is not translated—the **5'-untranslated region or 5'-UTR**.

open reading frame (ORF) Sequence of mRNA or corresponding region of DNA that can be translated to give a protein.
5'-untranslated region (5'-UTR) Short sequence at the 5' end of mRNA that is not translated into protein.

(sometimes 5'-nontranslated region or 5'-NTR). Hence, the reading frame cannot be defined simply by starting at the front end of the mRNA.

One way to define the reading frame is by choosing the **start codon**. The first codon is almost always AUG, encoding methionine. This will define both the start of translation and the reading frame. In the example considered above, there are three possible start codons (underlined), each of which starts at a slightly different point and gives a different reading frame:

GAAAAUGCAUGCCAAAGGAGGAUCUAAGGA

The start codon begins the coding sequence and is read by a special tRNA that carries methionine.

5.1. The Start Codon Is Chosen

A special tRNA, the **initiator tRNA**, is charged with methionine and binds to the AUG start codon (Fig. 13.17). In prokaryotes, chemically tagged methionine, **N-formyl-methionine (fMet)**, is attached to the initiator tRNA, whereas in eukaryotes unmodified methionine is used. Consequently, all polypeptide chains begin with methionine, at least when first made. Sometimes the initial methionine (in eukaryotes), or N-formyl-methionine (in prokaryotes), is snipped off later, so mature proteins do not always begin with methionine. In bacteria, even when the fMet is

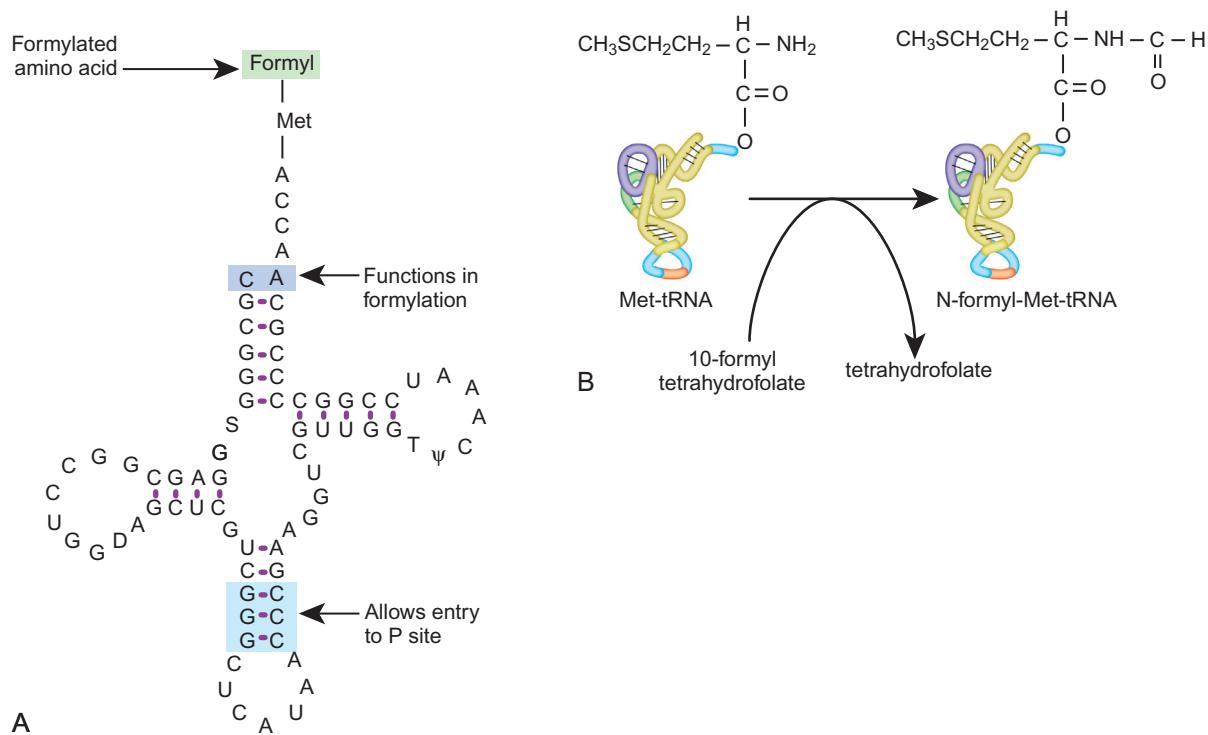


FIGURE 13.17
Initiator tRNA Carries N-Formyl-Methionine

(A) The structure of the initiator tRNA, fMet-tRNA, is unique. A CA base pair at the top of the acceptor stem is needed to allow formylation (violet). The initiator tRNA must enter the P-site directly (discussed later), which requires the three GC base pairs in the anticodon stem (blue).

(B) The initiator tRNA is first charged with unmodified methionine. Then a formyl group carried by the tetrahydrofolate cofactor is added to the methionine.

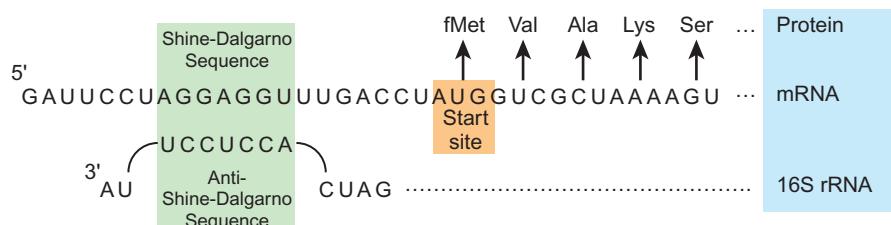
initiator tRNA The tRNA that brings the first amino acid to the ribosome when starting a new polypeptide chain.

N-formyl-methionine or fMet Modified methionine used as the first amino acid during protein synthesis in bacteria.

start codon The special AUG codon that signals the start of a protein.

FIGURE 13.18
Shine-Dalgarno Sequence of mRNA Binds to 16S rRNA

The Shine-Dalgarno sequence on the mRNA is recognized by base pairing with the anti-Shine-Dalgarno sequence on the 16S rRNA. The first AUG downstream of the S-D/anti-S-D site serves as the start codon.



To choose the correct start codon, the mRNA binds to 16S rRNA at a specific sequence.

not removed as a whole, the N-terminal formyl group is often removed leaving unmodified methionine at the N-terminus of the polypeptide chain.

AUG codons also occur in the middle of messages and result in the incorporation of methionines in the middle of polypeptide chains. So how does the ribosome know which AUG codon to start with? Near the front (the 5' end) of the mRNA of prokaryotes is a special sequence, the **ribosome-binding site (RBS)**, often called the **Shine-Dalgarno or S-D sequence**, after its two discoverers (Fig. 13.18). The sequence complementary to this, the **anti-Shine-Dalgarno sequence**, is found close to the 3' end of the 16S rRNA. Consequently, the mRNA and the 16S rRNA of the ribosome bind together by base pairing between these two sequences. The start codon is the next AUG codon after the ribosome-binding site. Typically, there are about seven bases between the S-D sequence and the start codon. In some cases, the S-D sequence exactly matches the anti-S-D sequence and the mRNAs are translated efficiently. In other cases, the match is poorer and translation is less efficient. (Note that eukaryotes do not use an S-D sequence to locate the start of translation; instead, they scan the mRNA starting from the 5'-cap—discussed later.)

Occasionally, coding sequences even start with GUG (normally encoding valine) instead of AUG. This leads to inefficient initiation and is mostly found for proteins required only in very low amounts, such as regulatory proteins, for example, LacI, the repressor of the *lac* operon (see Chapter 16: Regulation of Transcription in Prokaryotes). Note that when GUG acts as the start codon, the same initiator fMet-tRNA is used as when AUG is the start codon. Consequently, fMet is the first amino acid, even for proteins that start with a GUG codon. This is apparently due to the involvement of the initiation factors, especially IF3 (discussed later).

5.2. The Initiation Complex Assembles

Before protein synthesis starts, the two subunits of the ribosome are floating around separately. Because the 16S rRNA, with the anti-Shine-Dalgarno sequence, is in the small subunit of the ribosome, the mRNA binds to a free small subunit. Next the initiator tRNA, carrying fMet, recognizes the AUG start codon. Assembly of this **30S initiation complex** needs three proteins (IF1, IF2, and IF3), known as **initiation factors**, which help arrange all the components correctly. IF2 physically contacts the acceptor stem of the fMet-tRNA, and this interaction is essential to stabilize the initiation complex.

IF3 recognizes the start codon and the matching anticodon end of the initiator tRNA. IF3 prevents the 50S subunit from binding prematurely to the small subunit

anti-Shine-Dalgarno sequence Sequence on 16S rRNA that is complementary to the Shine-Dalgarno sequence of mRNA.

initiation factors Proteins that are required for the initiation of a new polypeptide chain.

Shine-Dalgarno (S-D) sequence Same as RBS; sequence close to the front of mRNA that is recognized by the ribosome; only found in prokaryotic cells.

ribosome-binding site (RBS) Same as Shine-Dalgarno sequence; sequence close to the front of mRNA that is recognized by the ribosome; only found in prokaryotic cells.

30S initiation complex Initiation complex for translation that contains only the small subunit of the bacterial ribosome.

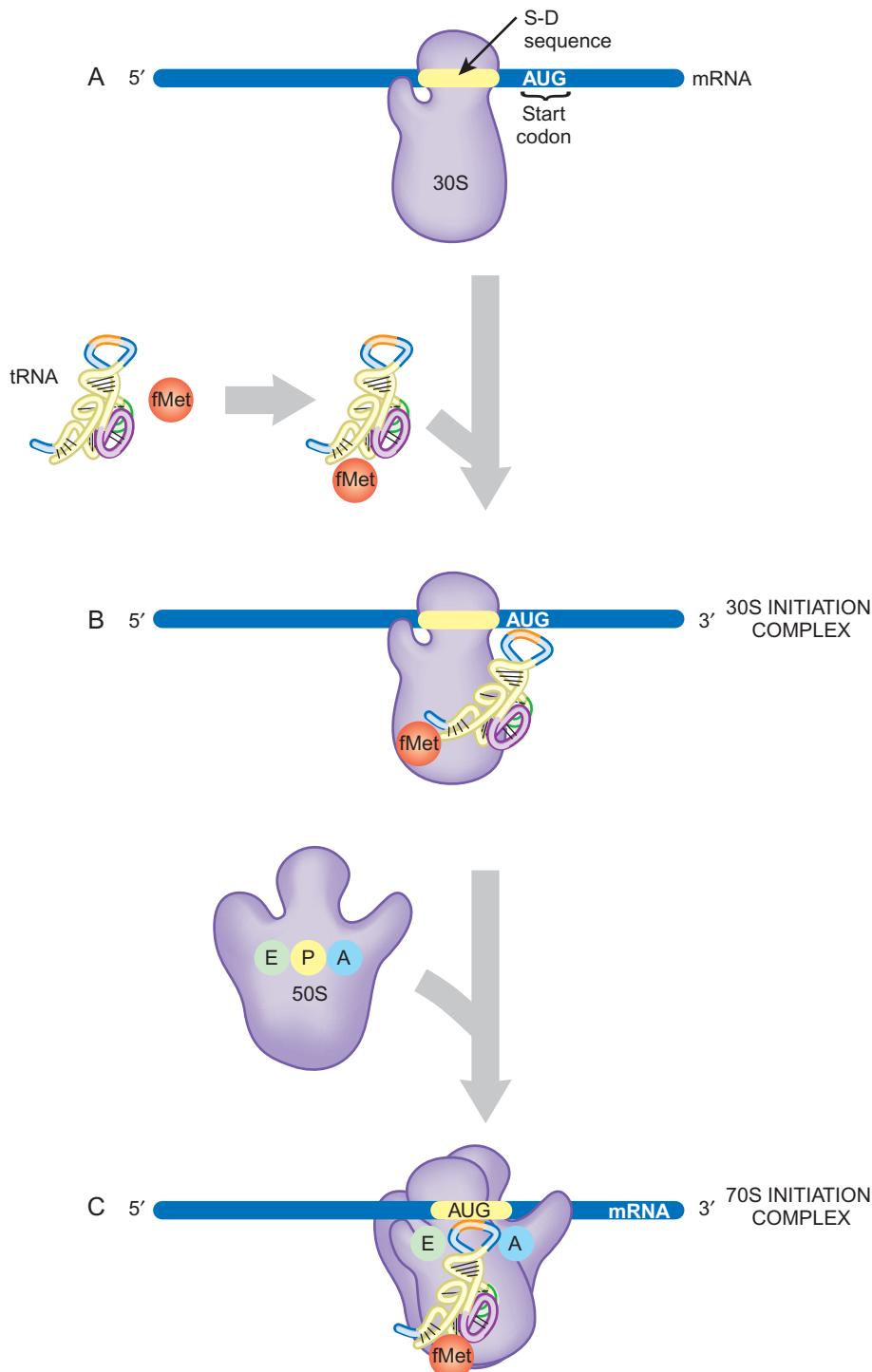


FIGURE 13.19
Formation of 30S and 70S Initiation Complexes

(A) The small subunit and the mRNA bind to each other at the Shine-Dalgarno sequence. The start codon, AUG, is just downstream of this site. (B) The initiator tRNA becomes tagged with fMet and binds to the AUG codon on the mRNA. (C) The large ribosomal subunit joins the small subunit and accommodates the tRNA at the P-site.

before the correct initiator tRNA is present. Once the 30S initiation complex has been assembled, IF3 departs and the 50S subunit binds. IF1 and IF2 are now released, resulting in the **70S initiation complex** (Fig. 13.19). This process consumes energy in the form of GTP, which is split by IF2.

Proteins known as initiation factors help the ribosomal subunits, mRNA and tRNA assemble correctly.

70S initiation complex Initiation complex for translation that contains both subunits of the bacterial ribosome.

6. The tRNA Occupies Three Sites During Elongation of the Polypeptide

Only two tRNA molecules can occupy the ribosome at any instant.

After peptide bond formation the tRNA carrying the growing polypeptide chain moves sideways between sites on the ribosome.

After the large subunit of the ribosome has arrived, the polypeptide can be made. Amino acids are linked together by the peptidyl transferase reaction, which is catalyzed by the 23S rRNA of the large subunit. The amino acids are carried to the ribosome attached to tRNA. The ribosome has three sites for tRNA: The **A (acceptor) site**, the **P (peptide) site**, and the **E (exit) site**. However, only two charged tRNA molecules can be accommodated on the ribosome at any given instant (Fig. 13.20).

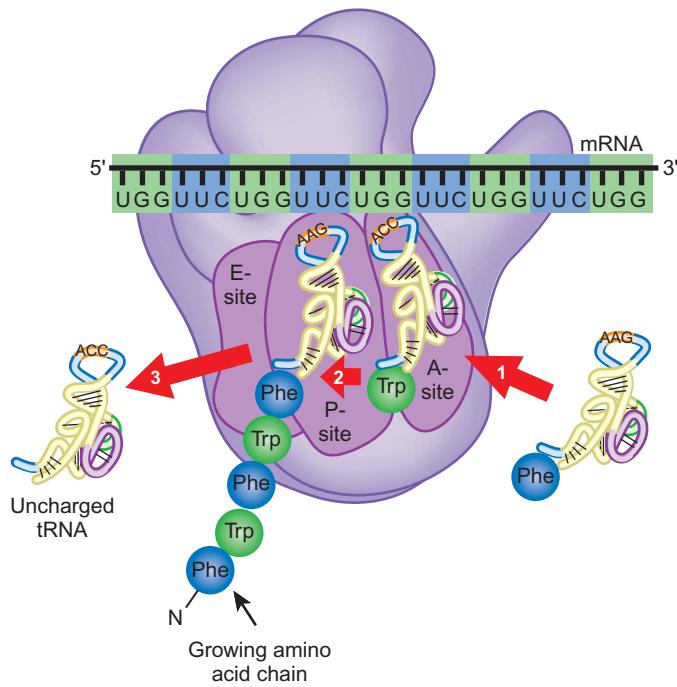
The fMet initiator tRNA starts out in the P-site. Another tRNA, carrying the next amino acid, arrives and enters the A-site. Next, fMet is cut loose from its tRNA and bonded to amino acid #2. So tRNA #2 now carries two linked amino acids, the beginnings of a growing polypeptide chain. The enzyme activity that joins two amino acids together is referred to as the peptidyl transferase activity, as the growing peptide chain is transferred from the tRNA carrying it at each step. The ribosome is necessary for this enzyme activity because it provides an environment that shields the active site from water and provides the proper charged residues to orient the incoming substrates.

After peptide bond formation, the two tRNAs are tilted relative to the A- and P-sites, a process called ratcheting (Fig. 13.21). The tRNA carrying the growing polypeptide chain now occupies part of the A-site on the 30S subunit but part of the P-site on the 50S subunit. This movement is facilitated by the **elongation factor** EF-G, which uses energy from GTP hydrolysis to maintain the ratcheted state and prevent the ribosome from returning to the original conformation.

The next step is **translocation**, in which the mRNA moves one codon sideways relative to the ribosome (Fig. 13.21). This moves the two tRNAs into the P- and

FIGURE 13.20
Overview of the
Elongation Cycle on the
Ribosome

(1) The incoming charged tRNA first occupies the A-site. (2) The peptide bond is formed between the amino acid at the A-site and the growing polypeptide chain in the P-site. (3) The uncharged tRNA exits the ribosome.



A- (acceptor) site Binding site on the ribosome for the tRNA that brings in the next amino acid.

E- (exit) site Site on the ribosome that a tRNA occupies just before leaving the ribosome.

elongation factors Proteins that are required for the elongation of a growing polypeptide chain.

P- (peptide) site Binding site on the ribosome for the tRNA that is holding the growing polypeptide chain.

translocation a) Transport of a newly made protein across a membrane by means of a translocase; b) sideways movement of the ribosome on mRNA during translation; and c) removal of a segment of DNA from a chromosome and its reinsertion in a different place.

ACCEPTANCE OF A NEW tRNA

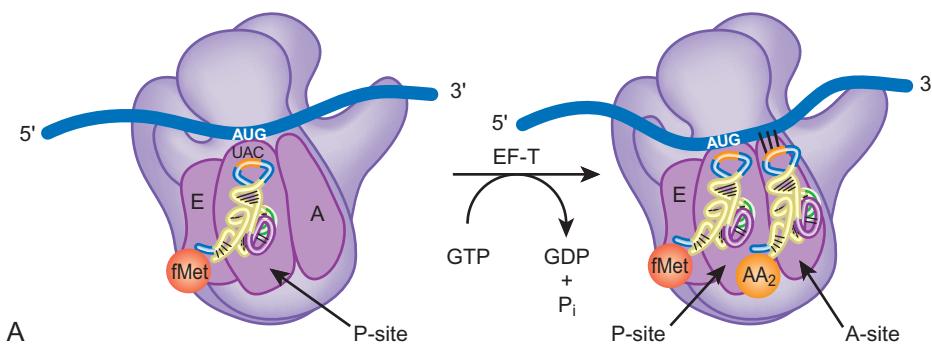
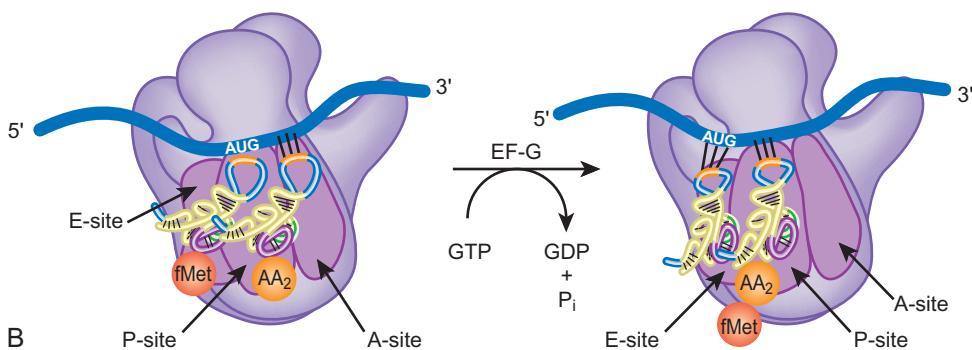


FIGURE 13.21
Elongation Factors and Site Occupancy

(A) The EF-T factor helps charged tRNA to occupy the A-site. (B) The EF-G factor helps translocate the tRNAs from the A- and P-sites to the P- and E-sites, respectively. Note that during translocation the tRNA temporarily binds "diagonally" across two sites.

TRANSLOCATION OF tRNA



E-sites, leaving the A-site empty. Once the two tRNAs are in the E- and P-sites, EF-G exits the complex. This dislodges the tRNA in the E-site and relocks the ribosome with tRNA #2 now in the P-site. The A- and E-sites cannot be simultaneously occupied. Once the previous tRNA has exited, another charged tRNA can enter the A-site. As the peptide chain continues to grow, it is constantly cut off from the tRNA holding it and joined instead to the newest amino acid brought by its tRNA into the A-site; hence, the name “acceptor” site. This process is repeated for each codon until the stop codon.

Bacterial elongation requires another elongation factor, EF-T, which uses energy in the form of GTP. EF-T actually consists of a pair of proteins, EF-Tu and EF-Ts. Incoming charged tRNA is delivered to the ribosome and installed into the A-site by elongation factor EF-Tu. This requires energy from the hydrolysis of GTP. EF-Ts is responsible for exchanging the GDP left bound to EF-Tu for a fresh GTP (Fig. 13.21).

The empty tRNA leaves the ribosome at the E-site and a new, charged tRNA enters at the A-site.

6.1. Termination of Translation and Ribosome Recycling

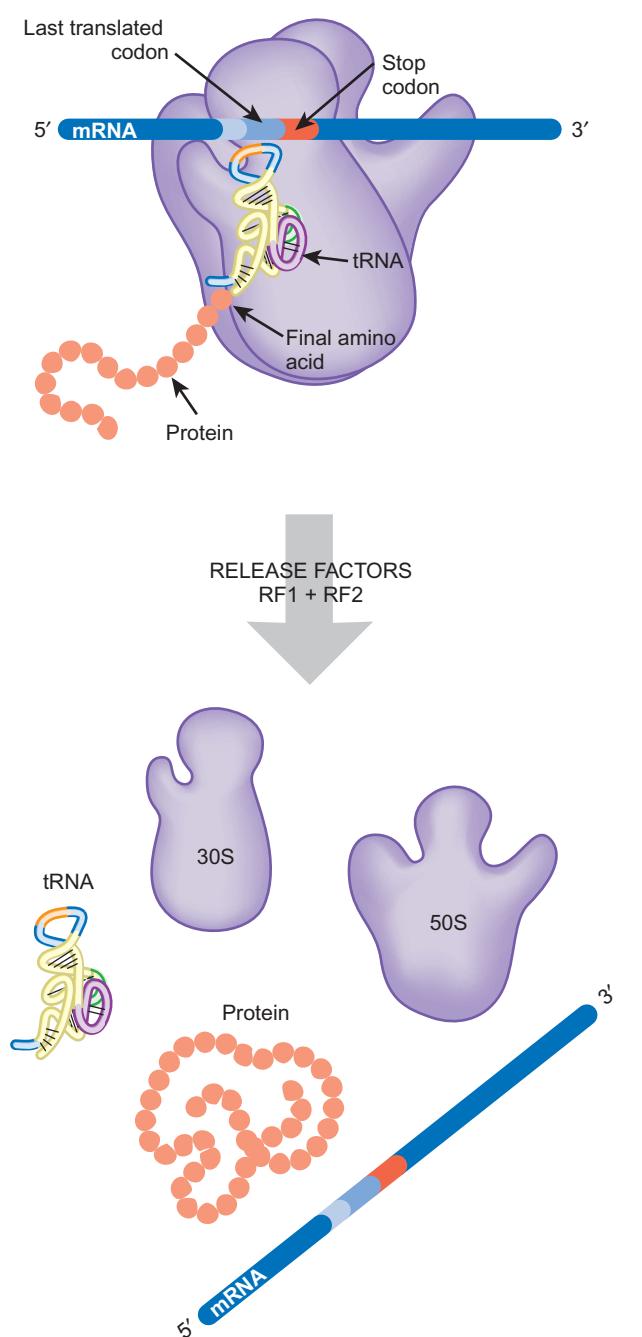
Eventually the ribosome reaches the end of the message. This is marked by one of three possible stop codons, UGA, UAG, and UAA. As no tRNA exists to read these three codons, the polypeptide chain can no longer grow. Instead, proteins known as **release factors (RF)** read the stop signals (Fig. 13.22). RF1 recognizes UAA or UAG and RF2 recognizes UAA or UGA. The completed polypeptide chain is now released from the last tRNA. This is actually done by the peptidyl transferase. Binding of the release factor activates the peptidyl transferase that hydrolyzes the bond between the finished polypeptide chain and the tRNA in the P-site. Then, RF3 releases RF1 and RF2 from the ribosome using GTP as an energy source.

The stop codon is read by a protein, the release factor, not by a tRNA.

release factor Protein that recognizes a stop codon and brings about the release of a finished polypeptide chain from the ribosome.

FIGURE 13.22
Termination and Release of Finished Polypeptide

In prokaryotes, after the ribosome has added the final amino acid, release factors (RF1 and RF2) recognize the stop codon and cause the ribosome complex to dissociate.



After polypeptide release, the ribosome complex is dissociated and recycled to translate a new mRNA. Two factors aid in dissociation: **Ribosome recycling factor (RRF)** and EF-G, which remove the large 50S subunit. Next, IF3 dissociates the last tRNA and mRNA from the small subunit. All the components are now free to be used again.

6.2. Several Ribosomes Usually Read the Same Message at Once

Once the first ribosome has begun to move, another can associate with the same mRNA and travel along behind. In practice, several ribosomes will move along the

ribosome recycling factor (RRF) Protein that dissociates the ribosomal subunits after a polypeptide chain has been finished and released.

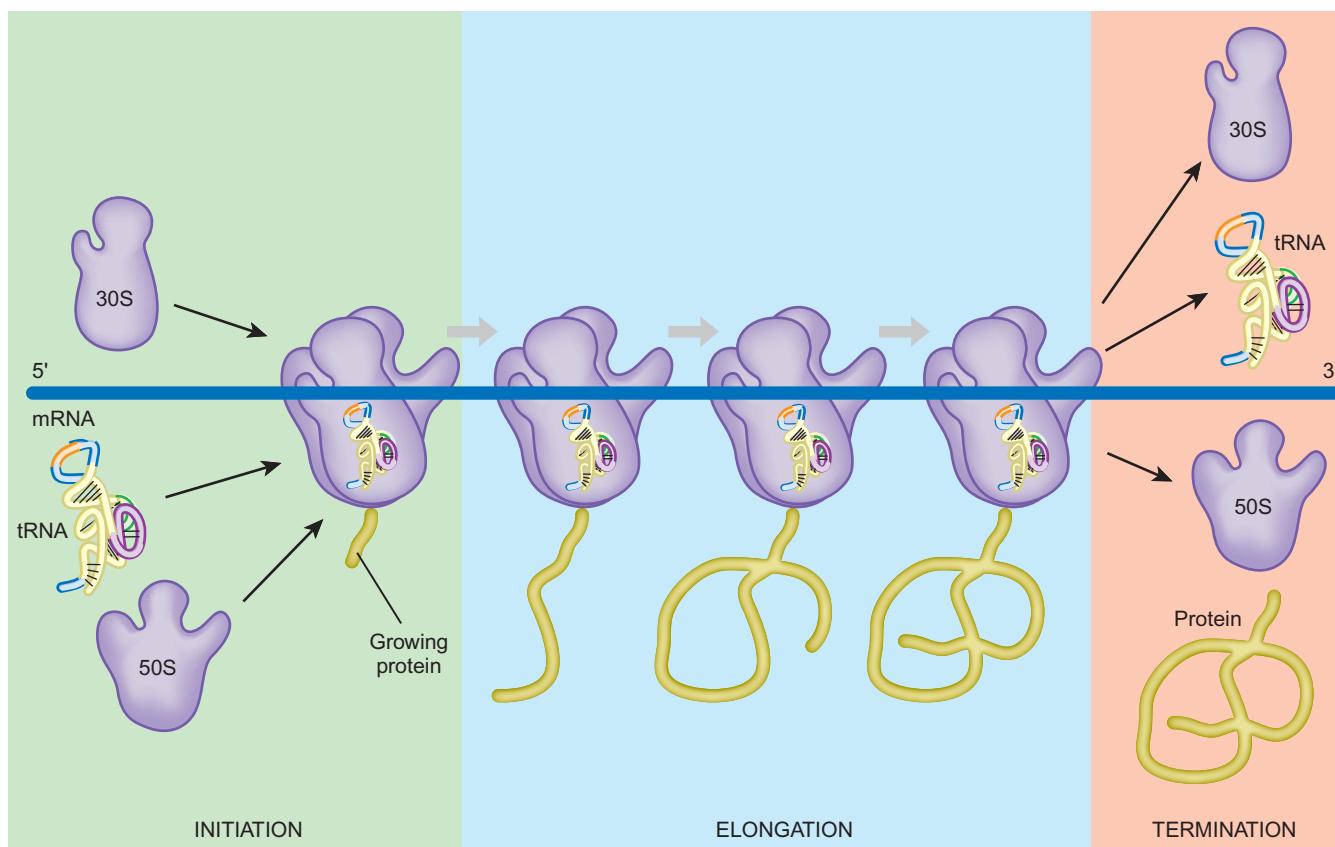


FIGURE 13.23
Polysome

A single mRNA molecule is associated with several ribosomes. At initiation the two subunits assemble; during chain elongation several ribosomes are at different stages in reading the same mRNA message; at termination the ribosome complex disassembles.

same mRNA about a hundred bases apart (Fig. 13.23). An mRNA with several attached ribosomes is called a **polysome** (short for polyribosome).

Electron microscope observations have suggested that the polysomes of eukaryotic cells are circular (Fig. 13.24). Apparently, the 3' end of the mRNA is attached to the 5' end by protein-protein contact between the poly(A) binding protein (attached to the 3'-poly(A) tail) and the eukaryotic initiation factor, eIF4 (attached to the cap at the 5' end). In prokaryotic cells, such circularization cannot occur as the 3' end of the mRNA is still being elongated by RNA polymerase while ribosomes have begun translating from the 5' end.

Messenger RNA is long enough for several ribosomes to translate it simultaneously.

7. Bacterial mRNA Can Code for Several Proteins

In bacteria, several genes may be transcribed to give a single mRNA. The term **operon** refers to clusters of genes that are co-transcribed. The result is that several proteins may be encoded by the same mRNA. As long as each ORF has its own Shine-Dalgarno sequence in front of it, the ribosome will bind and start translating. ORFs that are translated into proteins are sometimes known as cistrons; consequently, mRNA which carries several of these is called **polycistronic mRNA** (Fig. 13.25).

Messenger RNA in bacteria often carries several coding sequences.

operon A cluster of prokaryotic genes that are transcribed together to give a single mRNA (i.e., polycistronic mRNA).

polycistronic mRNA mRNA carrying multiple coding sequences that may be translated to give several different protein molecules; only found in prokaryotic (bacterial) cells.

polysome Group of ribosomes bound to and translating the same mRNA.

FIGURE 13.24
False Color TEM of Polysome

This false color transmission electron micrograph (TEM) shows a polysome from a human brain cell. Polysomes consist of several individual ribosomes connected by slender strands of mRNA. Magnification: $\times 240,000$. (Credit: CNRI/Science Photo Library.)

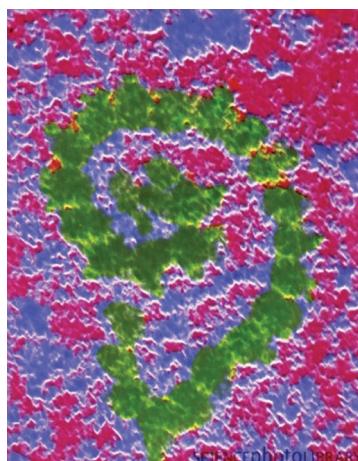
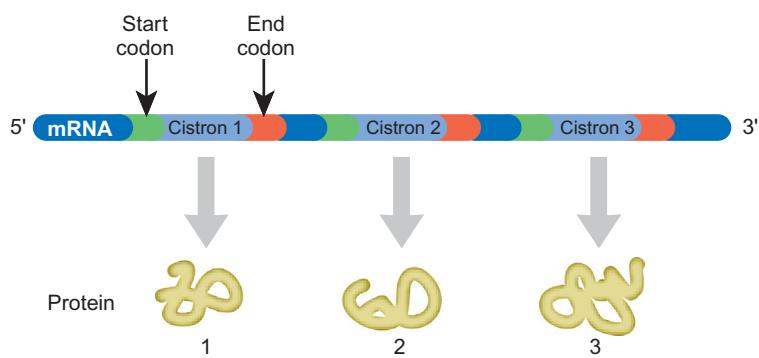


FIGURE 13.25
Polycistronic mRNA in Bacteria

The mRNA contains several cistrons or ORFs, each of which codes for a protein.



Eukaryotic mRNA molecules each code for a single protein.

In higher organisms operons are rare and neighboring genes are usually not co-transcribed. Each individual gene is transcribed separately to give an individual molecule of RNA. Apart from a few exceptional cases, each molecule of eukaryotic mRNA only carries a single protein coding sequence. Furthermore, eukaryotic mRNA does not make use of the Shine-Dalgarno sequence. Instead, the front (5' end) of the mRNA molecule is recognized by its cap structure. Consequently, in eukaryotes only the first ORF would normally be translated, even if multiple ORFs were present (discussed later).

7.1. Transcription and Translation Are Coupled in Bacteria

When mRNA is transcribed from the original DNA template, its synthesis starts at the 5' end. The mRNA is also read by the ribosome starting at the 5' end. In prokaryotic cells, the chromosome and ribosomes are all in the same single cellular compartment. Therefore, ribosomes can start translating the message before synthesis of the mRNA molecule has actually been finished. The result is that partly finished mRNA, still attached to the bacterial chromosome via RNA polymerase, may have several ribosomes already moving along it making polypeptide chains. This is known as **coupled transcription-translation** (Fig. 13.26). This is impossible in higher eukaryotic cells because the DNA is inside the nucleus and the ribosomes are outside, in the cytoplasm.

How are the ribosomes and the RNA polymerase kept in sync? It now appears that a complex of two proteins, NusE and NusG, binds the small subunit of the ribosome directly to the RNA polymerase (Fig. 13.27). When a ribosome binds mRNA

In prokaryotes, the ribosomes can begin to translate a message before the RNA polymerase has finished transcribing it.

coupled transcription-translation When ribosomes of bacteria start translating an mRNA molecule that is still being transcribed from the DNA.

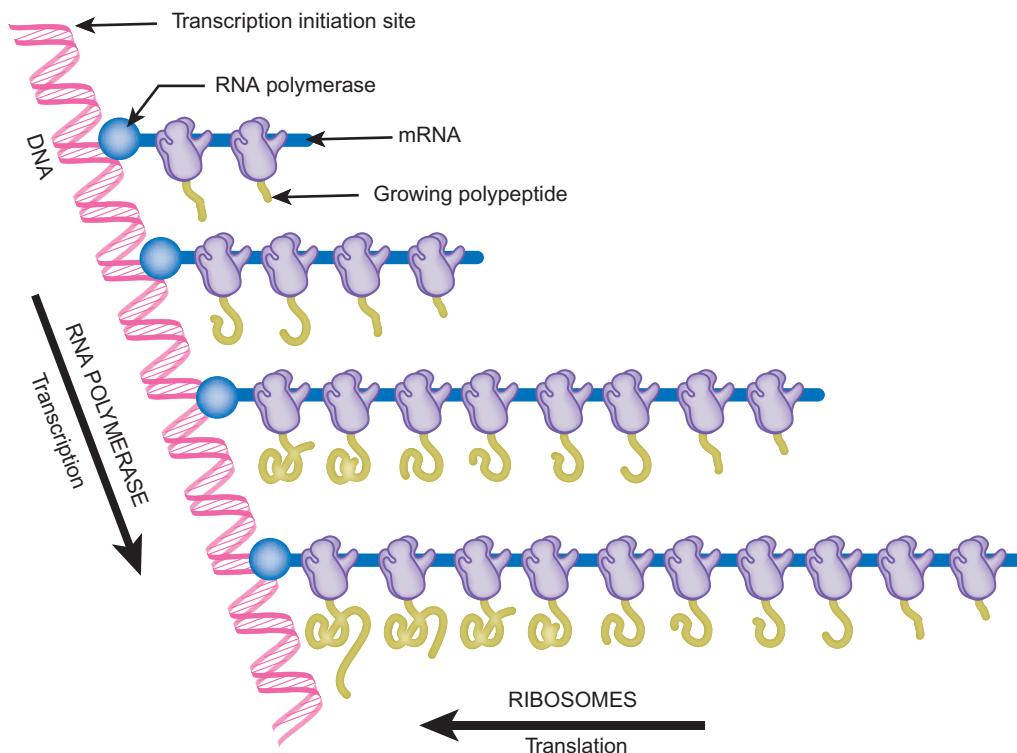


FIGURE 13.26
Coupled Transcription-Translation in Bacteria

Even as the DNA is being transcribed to give mRNA, ribosomes sequentially attach to the growing mRNA and initiate protein synthesis.

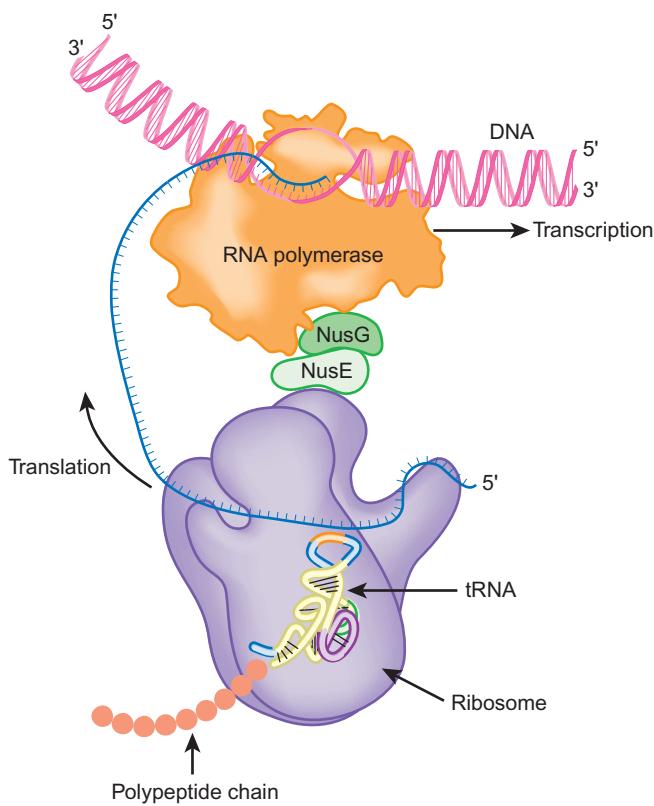


FIGURE 13.27
NusEG Couples Transcription With Translation

Ribosomes are directly attached to RNA polymerase via the NusEG complex.

and begins to make a polypeptide chain, this stimulates the RNA polymerase to increase the speed of RNA synthesis in order to keep up. Conversely, if bacteria are treated with an antibiotic that halts protein synthesis, RNA polymerase will slow down.

8. Some Ribosomes Become Stalled and Are Rescued

Cellular metabolism is not perfect and cells must allow for errors. One problem ribosomes sometimes run across is defective mRNA that lacks a stop codon. Whether synthesis of the mRNA was never completely finished or whether it was mistakenly snipped short by a ribonuclease, problems ensue. In the normal course of events, a ribosome that is translating a message into protein will, sooner or later, come across a stop codon. Even if an mRNA molecule comes to an abrupt end, ribosomes may be released only by release factor and this in turn needs a stop codon. If the mRNA is defective and there is no stop codon, a ribosome that reaches the end could just sit there forever and the ribosomes behind it will all be stalled, too.

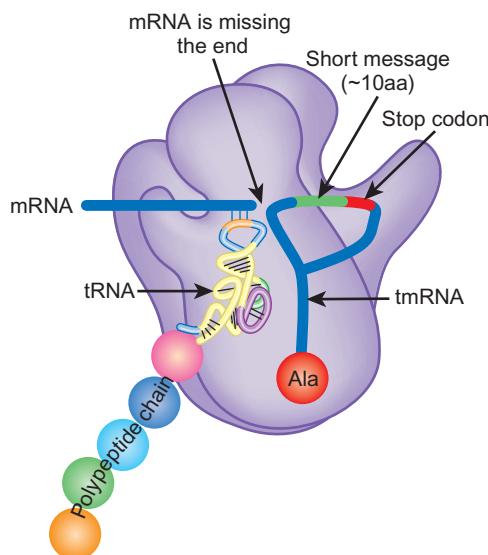
Ribosomes that have stalled due to defective mRNA can be rescued by a special RNA—tmRNA.

Bacterial cells contain a small RNA molecule that rescues stalled ribosomes. This is named **tmRNA** because it acts partly like tRNA and partly like mRNA. Like a tRNA, the tmRNA carries alanine, an amino acid. When it finds a stalled ribosome, it binds beside the defective mRNA (Fig. 13.28). Protein synthesis now continues, first using the alanine carried by tmRNA, and then continuing on to translate the short stretch of message that is also part of the tmRNA. Finally, the tmRNA provides a proper stop codon so that release factor can disassemble the ribosome and free it for continued protein synthesis. The tRNA domain of tmRNA lacks an anticodon loop and a D-loop. A protein known as SmpB (not shown in Fig. 13.28 for clarity) binds to the tRNA domain and makes contacts to the ribosome that would normally be made by the missing D-loop.

Clearly, the protein that has just been made is defective and should be degraded. As might be supposed, the tmRNA has signaled that the protein that was made is defective. The short stretch of 11 amino acids specified by the message part of tmRNA and added to the end of the defective protein acts as a signal, known as the **ssrA tag**. (SsrA stands for small stable RNA A, a name used for tmRNA before its function was elucidated.) The ssrA tag is recognized by several proteases (originally referred to as “**tail specific proteases**”), which degrade all proteins carrying this signal. These include the Clp proteases and the HflB protease involved in the heat shock response (see Chapter 16: Regulation of Transcription in Prokaryotes). Eukaryotic cells lack tmRNA, but do have a process called nonsense-mediated decay that degrades defective mRNA (see Chapter 12: Processing of RNA).

FIGURE 13.28
Stalled Ribosome Liberated by tmRNA

Binding of a tmRNA carrying alanine allows the translation of a damaged message to continue. First, alanine is added, then a short sequence of about 10 amino acids encoded by the tmRNA. Finally, the stop codon of the tmRNA allows proper termination of the polypeptide chain.



tail specific protease Enzyme that destroys defective proteins by degrading them tail-first; that is, from the carboxy terminal end.
tmRNA Specialized RNA used to terminate protein synthesis when a ribosome is stalled by damaged mRNA.

9. Differences Between Eukaryotic and Prokaryotic Protein Synthesis

The overall scheme of protein synthesis is similar in all living cells. However, there are significant differences between bacteria and eukaryotes. These are summarized in **Table 13.04** and discussed in the following sections. Note that eukaryotic cells contain mitochondria and chloroplasts, which have their own DNA and their own ribosomes. The ribosomes of these organelles operate similarly to those of bacteria and will be considered separately below. In eukaryotic protein synthesis, it is usually the cytoplasmic ribosomes that translate nuclear genes. Several aspects of eukaryotic protein synthesis are more complex. The ribosomes of eukaryotic cells are larger and contain more rRNA and protein molecules than those of prokaryotes. In addition, eukaryotes have more initiation factors and a more complex initiation procedure.

A few aspects of protein synthesis are actually less complex in eukaryotes. In prokaryotes, mRNA is polycistronic and may carry several genes that are translated to give several proteins. In eukaryotes, each mRNA is monocistronic and carries only a single gene, which is translated into a single protein. In prokaryotes, the genome and the ribosomes are both in the cytoplasm, whereas in eukaryotes the genome is in the nucleus. Consequently, coupled transcription and translation is not possible for eukaryotes (except for their organelles; discussed later).

Eukaryotic ribosomes are larger and more complex than those of prokaryotes.

Both prokaryotes and eukaryotes have a special initiator tRNA that recognizes the start codon and inserts methionine as the first amino acid. In prokaryotes, this first methionine has a formyl group on its amino group (i.e., it is N-formyl-methionine), but in eukaryotes unmodified methionine is used.

9.1. Initiation, Elongation, and Termination of Protein Synthesis in Eukaryotes

Initiation of protein synthesis differs significantly between prokaryotes and eukaryotes. Eukaryotic mRNA has no ribosome-binding site (RBS). Instead recognition and binding to the ribosome rely on a component that is lacking in prokaryotes: The cap structure at the 5' end, which is added to eukaryotic mRNA before it leaves the

TABLE 13.04 Comparison of Protein Synthesis

Prokaryotes	Eukaryotes (Cytoplasm)
Polycistronic mRNA	Monocistronic mRNA
Coupled transcription and translation	No coupled transcription and translation for nuclear genes
Linear polyribosomes	Circular polyribosomes
No cap on mRNA	5' end of mRNA is recognized by cap
Start codon is next AUG after ribosome-binding site	No ribosome-binding site so first AUG in mRNA is used
First amino acid is formyl-Met	First Met is unmodified
70S ribosomes made of: 30S and 50S subunits	80S ribosomes made of: 40S and 60S subunits
Small 30S subunit: 16S rRNA and 21 proteins	Small 40S subunit: 18S rRNA and 33 proteins
Large 50S subunit: 23S and 5S rRNA plus 31 proteins	Large 60S subunit: 28S, 5.8S and 5S rRNA plus 49 proteins
Elongation factors: EF-T (2 subunits) and EF-G	Elongation factors: eEF1 (3 subunits) and eEF2
Three initiation factors: IF1, IF2 and IF3	Multiple initiation factors: eIF2 (3 subunits), eIF3, eIF4 (4 subunits), eIF5
Shut-off by dimerization of ribosomes in nongrowing cells	Control via eIF sequestration

TABLE 13.05 Translation Factors: Prokaryotes vs Eukaryotes

	Prokaryotes	Eukaryotes
Initiation	IF1 IF2 IF3	eIF1A eIF5B (GTPase) eIF1 eIF2 (α , β , γ) (GTPase) eIF2B (α , β , γ , δ , ϵ) eIF3 (13 subunits) eIF4A (RNA helicase) eIF4B (activates eIF4A) eIF4E (cap-binding protein) eIF4G (eIF4 complex scaffold) eIF4H eIF5 eIF6 PABP (Poly(A)-binding protein)
Elongation	EF-Tu EF-Ts EF-G	eEF1A eEF1B (2–3 subunits) SBP2 eEF2
Termination	RF1 RF2 RF3	eRF1 eRF3
Recycling	RRF EF-G	eIF3 eIF3j eIF1A eIF1

Functionally homologous factors are in the same row.

Adapted from Table 1 of Rodnina MV and Wintermeyer W. (2009) Recent mechanistic insights into eukaryotic ribosomes. *Curr. Op. Cell Biol.* 21: 435–443.

nucleus (see Chapter 12: Processing of RNA). Cap-binding protein (one of the subunits of eIF4) binds to the cap of the mRNA.

Eukaryotes also have more initiation factors than prokaryotes and the order of assembly of the initiation complex is different (see Table 13.05). Two different complexes assemble before binding to mRNA. The first is the 43S pre-initiation complex. This is an assembly of the small 40S subunit of the ribosome attached to several eukaryotic initiation factors (eIFs). These include eIF1, eIF1A, eIF3, and eIF5. This binds the charged initiator tRNA, Met-tRNA_i^{Met}, plus eIF2. The second complex, the cap-binding complex, contains cap-binding protein (eIF4E), eIF4G, eIF4A, eIF4B, and poly(A)-binding protein (PABP).

During eukaryotic initiation, cap-binding complex first attaches to the mRNA via its cap. Next, the poly(A) tail is bound by PABP so that the mRNA forms a ring. This structure can now bind the 43S assembly. In order to align the Met-tRNA_i^{Met} with the correct AUG codon, the two structures work together to scan each codon from the 5' end. This scanning process uses energy from ATP (Fig. 13.29). Normally, the first AUG is used as the start codon (see Box 13.02 for exceptions), although the sequence surrounding the AUG is important. The consensus is GCCRCCAUGG (R = A or G). If its surrounding sequence is too far from consensus an AUG may be

Eukaryotic mRNA is recognized by its cap structure (not by base pairing to rRNA).

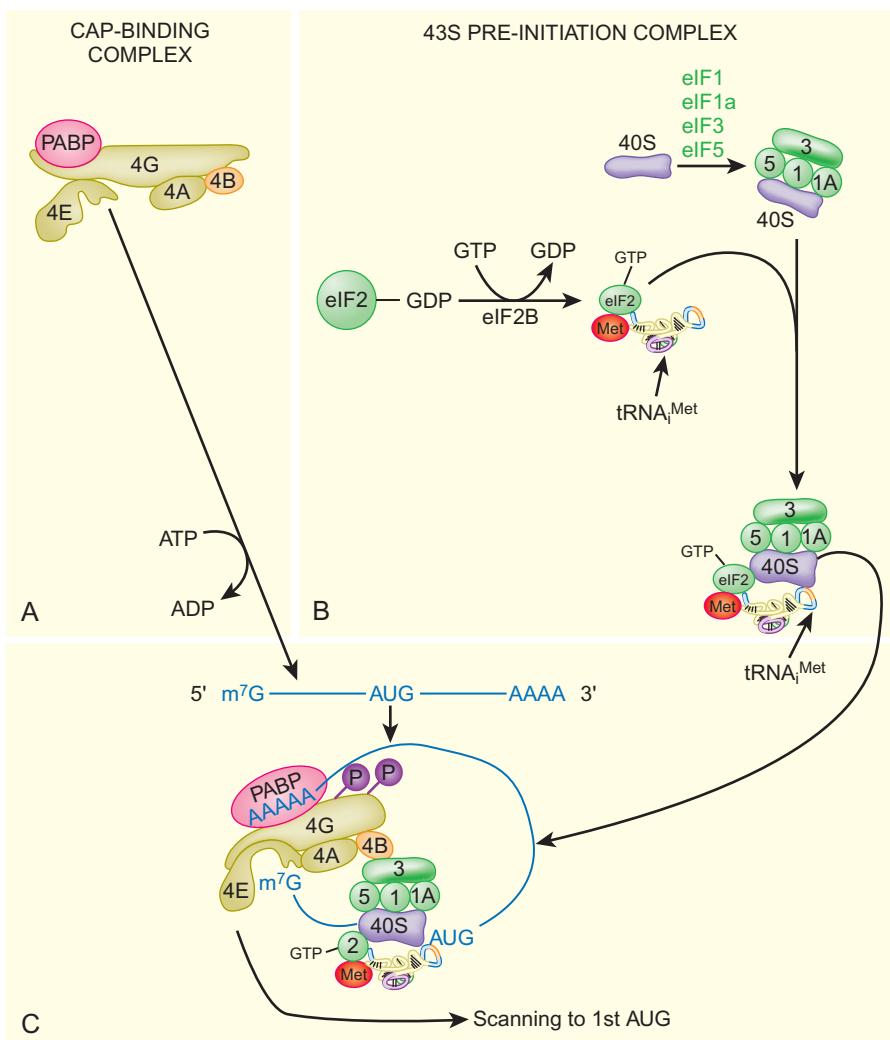


FIGURE 13.29
Assembly of the Eukaryotic Initiation Complex

(A) The cap-binding complex includes poly(A)-binding protein (PABP), eIF4A, eIF4B, eIF4E, and eIF4G, which is in an unphosphorylated state when unbound to mRNA. ATP transfers phosphates to the complex to make it competent for binding the mRNA. (B) The 43S initiation complex forms bringing the small ribosomal subunit together with the tRNA_i^{met}. This complex uses GTP to attach the tRNA to the 40S subunit via eIF2. In addition, initiation factors eIF1, eIF1A, eIF3, and eIF5, and eIF2B guide and make the complex competent to bind to the 5'-UTR of mRNA. (C) The mRNA is recognized by the cap-binding complex via the connections between eIF4E and PABP which bind the 5' and 3' ends of the mRNA, respectively. These two connections cause the rest of the mRNA to loop out. When this is established, then the 43S pre-initiation complex can attach and start scanning for the first AUG. After pausing at the first AUG, then the 50S subunit of the ribosome can bind and initiate translation.

skipped. Once a suitable AUG has been located, eIF5 joins the complex, which in turn allows the 60S subunit to join and the cap-binding protein, eIF2, eIF1, eIF3, and maybe eIF5 to depart. eIF5 uses energy from GTP to accomplish this remodeling of the ribosome.

The next stage is elongation (Fig. 13.30). Of all the stages of translation, elongation in bacteria and eukaryotes is the most similar. As in bacteria, elongation factors work to decode the mRNA and bind the tRNA into the A-site of the ribosome. Rather than EF-Tu and EF-Ts, eukaryotes use eEF1A to deliver the tRNA using GTP hydrolysis for energy and eEF1B to replace the depleted GDP with fresh GTP. The only difference is that eukaryotic elongation factors include more subunits. The remaining steps are the same. The peptidyl transferase activity of the 28S rRNA of the large subunit links the incoming amino acid to the polypeptide chain. Then elongation factor eEF2 (direct counterpart to bacterial EF-G) uses GTP to drive the conformational changes in the ribosome and ratchet the tRNAs from the P- and A-sites into the E- and P-sites. Elongation continues until a stop codon enters the A-site.

Eukaryotic termination differs from prokaryotic termination in two ways. First, rather than having two different release factors (RF1 and RF2) to recognize different stop codons, eukaryotes have a single release factor (eRF1) that recognizes all three stop codons. eRF1 binds the stop codon, but this does not affect peptide bond formation. Instead, eRF3 carrying a GTP molecule binds to eRF1. GTP hydrolysis then rearranges the factors and the final amino acid attaches to the polypeptide. Therefore, eukaryotes require GTP for polypeptide completion, whereas in bacteria, RF1 or RF2 is sufficient.

Box 13.02 Internal Ribosome Entry Sites

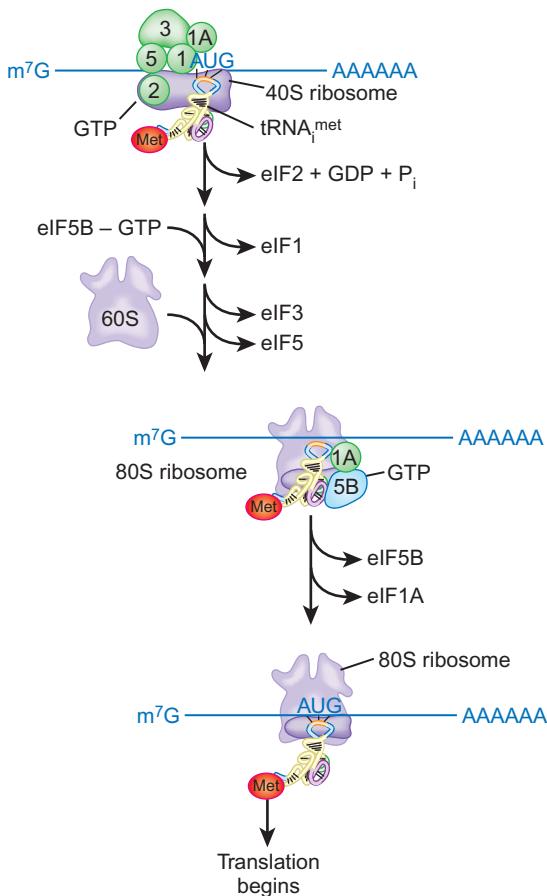
Although most eukaryotic mRNA is scanned by the 40S subunit to find the first AUG, exceptions do occur. Sequences known as internal ribosome entry sites (IRES) are found in a few mRNA molecules. As the name indicates, these allow ribosomes to initiate translation internally, rather than at the 5' end of the mRNA. IRES sequences were first found in certain viruses that have polycistronic mRNA despite infecting eukaryotic cells. In this case, the presence of IRES sequences in front of each coding sequence allows a single mRNA to be translated to give multiple proteins. The best known examples are members of the Picornavirus family, which includes poliovirus (causative agent of polio) and rhinovirus (one of the agents of common cold).

More recently, it has been found that a few special mRNA molecules encoded by eukaryotic cells themselves also possess IRES sequences. During major stress situations, such as heat shock or energy deficit, synthesis of the majority of proteins is greatly decreased. Much of this regulation occurs at the initiation stage of translation (discussed later). However, a few proteins are exempted from this down-regulation as they are needed under stress conditions. The mRNAs encoding these proteins often contain an IRES sequence. In these cases, the mRNA carries only a single coding sequence and the IRES is located in the 5'-UTR, between the 5' end of the mRNA and the start of the coding sequence. This allows translation to be initiated at the IRES even in the absence of the standard initiation/scanning procedure.

Finally, as in bacteria, eukaryotic ribosomes are recycled. eIF3 triggers the release of the 60S subunit, and then eIF1 releases the final tRNA. An additional factor, eIF3j, then removes the mRNA. The components are then recycled.

FIGURE 13.30
Beginning Eukaryotic Translation Elongation

Once the eukaryotic 40S subunit complex finds the first AUG, then the remaining 60S subunit and associated factors combine to form the final 80S ribosome.



10. Protein Synthesis Is Halted When Resources Are Scarce

Proteins make up about two-thirds of the organic matter in a cell and their synthesis consumes a major part of the cell's energy and raw materials. Clearly, when cells run low on nutrients or energy they cannot continue to synthesize proteins at the normal rate. In bacteria, ribosomes are taken out of service during stationary phase or periods of slow growth. A small basic protein, **ribosome modulation factor (RMF)**, binds to ribosomes and inactivates them (Fig. 13.31). The inactive ribosomes exist as dimers. When favorable conditions return, the inactive dimers are disassembled and the ribosomes are reactivated. In *E. coli*, starvation induces the **stringent response**, where the cell only transcribes genes that are essential for survival and virulence. Most other genes are turned off at the transcriptional level. This response is triggered, in part, when an uncharged tRNA loads into the A-site of the ribosome. This will only occur if charged tRNA is in short supply—due to lack of energy or amino acids. When this happens, RelA, a protein associated with the ribosome during translation, starts making pppGpp, which binds to RNA polymerase to modulate what genes are transcribed. This response is also important for bacteria such as *Mycobacterium tuberculosis*, which causes tuberculosis. These bacteria are highly resistant to antibiotic treatment and the human immune system and can live for years inside immune cells. They are able to persist because of their stringent response. Recently, abscesses caused by *Staphylococcus aureus* and *Pseudomonas aeruginosa* (both use the stringent response as part of their pathogenicity) have been diminished using a peptide that targets ppGpp in the stringent response. Targeting the stringent response might provide useful tools against antibiotic-resistant strains of bacterial pathogens.

Higher organisms also stop protein synthesis when nutrients or energy run low. However, they do so by inactivating the initiation factors rather than the ribosomes (Fig. 13.32). Initiation factor eIF2 uses energy by hydrolyzing GTP to GDP. After initiation is over, it is released from the ribosome with the GDP still bound. It then binds to eIF2B, which exchanges GDP for GTP, so recycling the eIF2. In times of stress, a kinase phosphorylates eIF2 and prevents the removal of GDP. The GDP bound form of eIF2 cannot initiate translation and protein synthesis is halted. Some viruses have taken advantage of this mechanism and use phosphorylation of eIF2 as a way to shut down host protein synthesis.

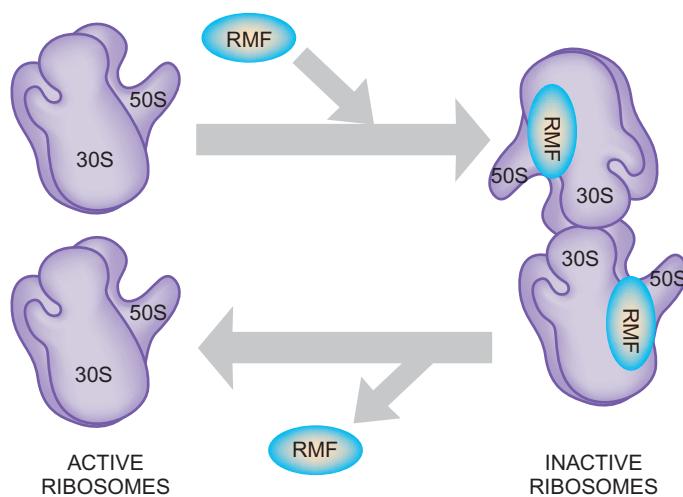


FIGURE 13.31
Bacterial Ribosomes on Standby During Bad Conditions

Active bacterial ribosomes can become inactive when the RMF protein binds to them. The ribosomes form dimers with the 30S subunits attached to one another. When conditions are favorable, dissociation occurs.

ribosome modulation factor (RMF) Protein that inactivates surplus ribosomes during slow growth or stationary phase in bacteria.
stringent response Decreasing transcription of nonessential genes when nutrients are in limited supply.

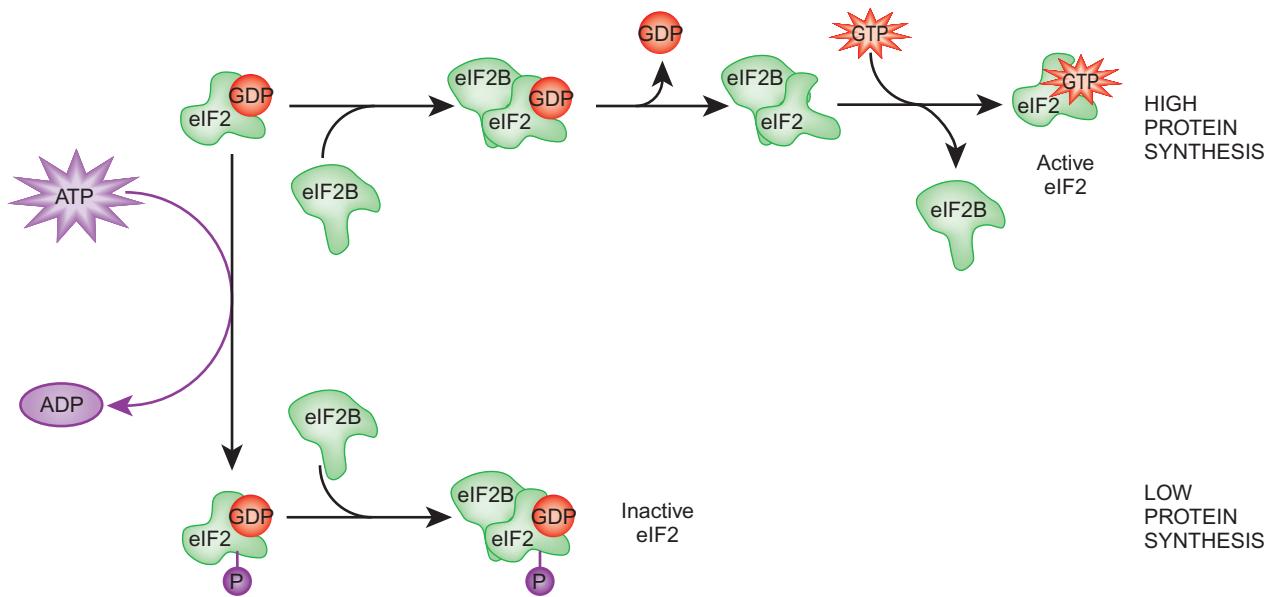


FIGURE 13.32
Recycling of Initiation Factor eIF2 Is Controlled

When eukaryotes down-regulate the level of protein synthesis, a protein kinase phosphorylates eIF2/GDP. This prevents eIF2B from removing the GDP and eIF2/GDP stays locked in an inactive complex with eIF2B. Absence of active eIF2 decreases the rate of initiation.

As usual, eukaryotes have multiple pathways of control. Another way the cell controls protein synthesis is to regulate the cap-binding complex. When the cell is stressed or starved, mTOR (target of the antibiotic rapamycin) is prevented from phosphorylating eIFs 4E, 4G, and 4B. Without these phosphates, the eIF complex cannot bind to mRNA, and translation is halted.

11. A Signal Sequence Marks a Protein for Export From the Cell

Once a protein has been made, it must find its correct location within the cell. Although cytoplasmic proteins are made in the cell compartment where they belong, other proteins, which do not reside in the cytoplasm, must be transported. Proteins destined to be exported to the exterior of the cell must be exported through the cell membrane. Similar systems exist in bacterial and eukaryotic cells. Proteins destined for export are tagged at the N-terminus with a **signal sequence**. This is cut off after export by proteases attached to the outside of the membrane and is therefore not present in the mature protein. The signal sequence consists of approximately 20 amino acids that form an α -helix. There is little specific sequence homology between signal sequences from different exported proteins. A positively charged, basic N-terminus of two to eight amino acids is followed by a long stretch of hydrophobic amino acids. The amino acid just before the cleavage site has a short side chain (Fig. 13.33).

Exported proteins have a signal sequence at the front.

A polypeptide destined for export is recognized by its signal sequence. In bacteria, the signal recognition protein (SecA) binds the signal sequence and guides it to the **translocase** complex in the cell membrane. The rest of the protein being exported is synthesized and follows the signal sequence into and through the membrane via the translocase. This is known as **cotranslational export**, since the protein

cotranslational export Export of a protein across a membrane while it is still being synthesized by a ribosome.

signal sequence Short, largely hydrophobic sequence of amino acids at the front of a protein that label it for export.

translocase Enzyme complex that transports proteins across membranes.

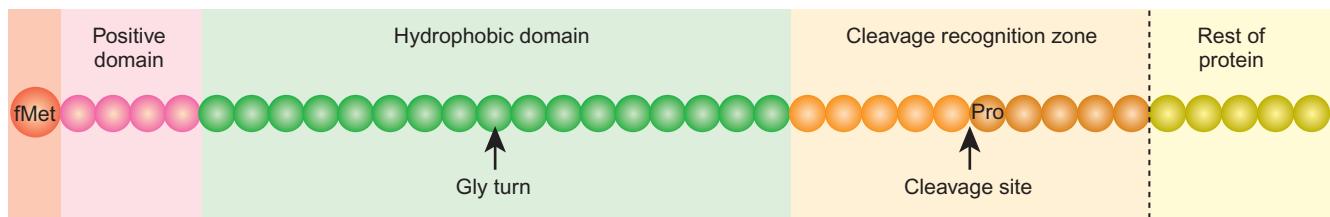


FIGURE 13.33
Standard Signal Sequence for Exported Proteins

The signal sequence contains a positively charged domain (containing lysine and/or arginine), a α -helical hydrophobic domain (rich in alanine, leucine and valine), and a cleavage site preceded by a glycine or serine and followed by a proline. A reverse turn due to glycine is found approximately halfway through the hydrophobic domain.

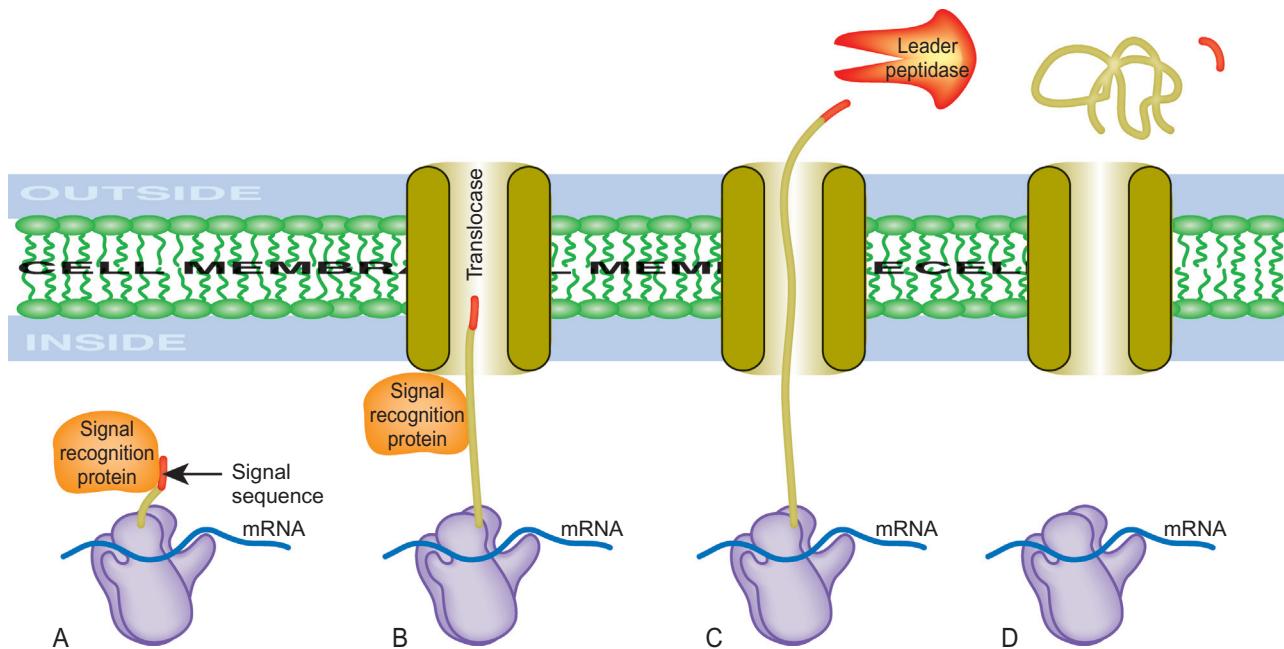


FIGURE 13.34
Cotranslational Export of Proteins

(A) The ribosome making the polypeptide chain approaches the cell membrane. The polypeptide with its signal sequence binds to the signal recognition protein. (B) The signal recognition protein recognizes the translocase and binds to it, allowing the polypeptide chain to begin its journey through the membrane. (C) After the signal sequence exits the translocase, leader peptidase cuts the polypeptide chain, liberating the signal peptide. (D) Final folding of the protein occurs outside the cell.

is exported as it is made. The signal sequence is cut off by the **leader peptidase** (or signal peptidase) after translocation (Fig. 13.34).

There are approximately 500 translocases per *E. coli* cell. Each cell exports about 1×10^6 proteins from the cytoplasm prior to dividing. In a cell that doubles in 20 minutes, 100 proteins are exported per minute per translocase. Protein export is 10-fold faster than protein synthesis. So the demand for a growing protein chain will allow the translocase to be ready for a new chain as fast as the ribosome can make it. Note that in gram-negative bacteria such as *E. coli*, most of these exported proteins are structural components of the outer membrane that are being made constantly, rather than enzymes being excreted outside the cell for digestive purposes.

After export of a protein across the cell membrane via the translocase, the signal sequence is cut off.

leader peptidase Enzyme that removes the leader sequence after protein export.

In eukaryotes, cotranslational export occurs across the membranes of the endoplasmic reticulum. In multicellular eukaryotes proteins involved in digestion, such as amylases and proteases, must be exported. So must proteins located in blood and other body fluids, such as antibodies, albumins, and circulating peptide hormones. When the animal genes for preproinsulin or ovalbumin are put into *E. coli*, correct export across the cell membrane occurs and cleavage of the signal sequence by the *E. coli* leader peptidase happens at the correct position. Conversely, yeast cells correctly process and excrete bacterial β -lactamase. Thus, the export machinery is highly conserved between diverse organisms.

11.1. Molecular Chaperones Oversee Protein Folding

Chaperonins are proteins that promote the correct folding of other proteins.

Molecular **chaperones**, or **chaperonins**, are proteins that oversee the correct folding of other proteins. Many chaperonins belong to the family of **heat shock proteins (HSPs)**, as their levels increase at high temperature (see Chapter 16: Regulation of Transcription in Prokaryotes). Chaperonins may be divided into two main classes: “holders” and “folders”—respectively, those that prevent premature folding and those that attempt to rectify misfolding. Obviously, chaperonins cannot “know” the correct 3D structure for several other proteins. Mechanistically, they act to prevent incorrect folding, rather than actively creating a correct structure.

During bacterial protein export, the secretory chaperonin SecB keeps the polypeptide chain from folding up prematurely. Secreted proteins must travel through a narrow translocase channel and so must remain unfolded until they reach the other side of the membrane. The Hsp70 set of chaperonins tends to bind to newly synthesized or highly uncoiled proteins (Fig. 13.35).

The more complex GroE (=Hsp60/Hsp10) chaperonin machine attempts to refold damaged or misfolded proteins. When polypeptide chains unfold, they expose hydrophobic regions that are normally clustered in the center of the folded protein. Left to themselves, many proteins could refold. However, inside a cell, there is a high concentration of protein. Consequently, exposed hydrophobic regions from multiple proteins bind to each other and the proteins aggregate together. The GroE chaperonin machine forms a cavity in which a single polypeptide can refold on its own, protected from interactions with other polypeptide chains.

Newly made proteins face a folding problem. The N-terminus of the growing polypeptide has already left the ribosome, while the C-terminal region is still being made. Consequently, the N-terminal region does not yet have access to any folding information that resides in later regions of the protein. To prevent misfolding at this stage, the emerging protein is sheltered by a chaperonin known as trigger factor (Fig. 13.36). This binds to the large subunit of the ribosome close to the polypeptide exit tunnel.

12. Protein Synthesis Occurs in Mitochondria and Chloroplasts

Mitochondria and chloroplasts are thought to be of prokaryotic origin. The symbiotic hypothesis of organelle origins argues that symbiotic prokaryotes evolved into organelles by specializing in energy production and progressively losing their genetic independence (see Chapter 4: Genes, Genomes, and DNA for further details). Both mitochondria and chloroplasts contain circular DNA that encodes some of their own genes and they divide by binary fission. They contain their own ribosomes and make some of their own proteins. Organelle ribosomes resemble the ribosomes of bacteria rather than the ribosomes of the eukaryotic cytoplasm (Box 13.03). The initiation and elongation factors of organelles are also bacterial in nature. Nonetheless, there

chaperone Sometimes “molecular chaperone”; same as chaperonin.

chaperonin Protein that oversees the correct folding of other proteins.

heat shock protein (HSP) Protein induced in response to high temperature. Many heat shock proteins are chaperonins.

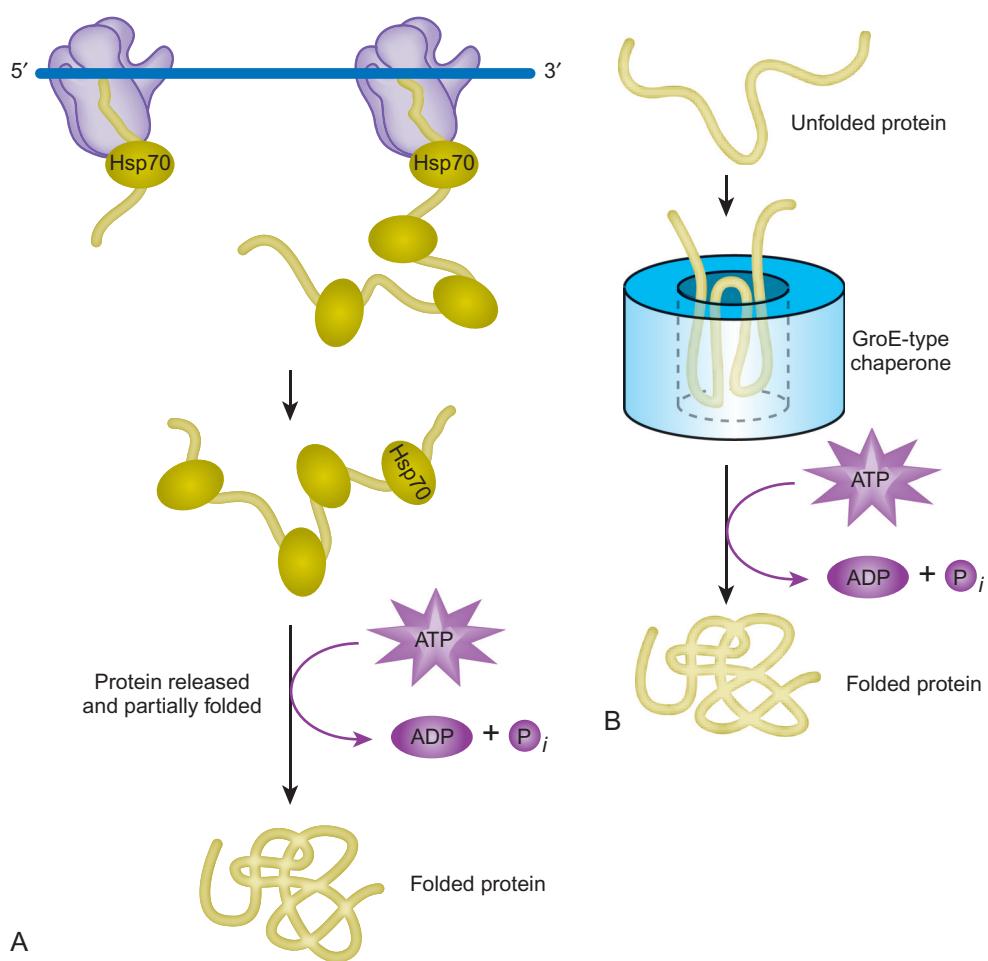


FIGURE 13.35
Chaperonins Act by Two General Mechanisms

(A) Chaperonins of the Hsp70 type act during protein formation by binding to hydrophobic patches of the protein. Once chaperonins are released, the protein automatically folds. (B) Large chaperonins, such as GroE, act after translation by sequestering misfolded protein in a central cavity. Freed from the influences of other molecules in the cytoplasm, the protein will fold correctly.

are differences in composition between organelle and bacterial ribosomes, as shown in [Table 13.06](#).

12.1. Proteins Are Imported Into Mitochondria and Chloroplasts by Translocases

The size of organelle genomes varies considerably from organism to organism. Generally, the more advanced eukaryotes have smaller organelle genomes. The mitochondria of mammals make only around 10 proteins and in higher plants the chloroplasts make approximately 50 proteins. The other organelle proteins are encoded by nuclear genes and made on the cytoplasmic ribosomes. They are then transported into the organelles.

Proteins for import into mitochondria have a leader sequence at the N-terminus. This consists of 20 or more amino acids with a positively charged lysine or arginine every three or four residues and no negatively charged residues. The leader forms an α -helix with a positively charged face and a hydrophobic face. This is recognized by a receptor on the mitochondrial surface. The protein is imported successively through two translocase complexes known as TOM (translocase, outer mitochondrial) and TIM (translocase, inner mitochondrial) that lie in the outer and inner membranes of the mitochondria, respectively. After importing the protein, its leader sequence is trimmed off.

Plant cells are more complex than animal cells as they possess not only mitochondria but also chloroplasts. The principle of protein import is similar. The chloroplast contains two translocases equivalent to TIM and TOM, which are known as TIC and TOC (C for chloroplast). The leader sequences for chloroplast proteins

Protein synthesis in mitochondria and chloroplasts resembles that of bacteria in many respects.

Many mitochondrial and chloroplast proteins are made in the eukaryotic cytoplasm and enter the organelle after synthesis.

FIGURE 13.36
Trigger Factor and the Ribosome

Trigger factor shelters newly made polypeptides as they emerge from the ribosome. Trigger factor protects about 70% of newly made proteins. The other 30% need the DnaJ/K or GroEL chaperonins. (Credit: Hoffmann A, Bukau B, Kramer G. (2010) Structure and function of the molecular chaperone Trigger Factor. *Biochim Biophys Acta* 1803:650–61.)

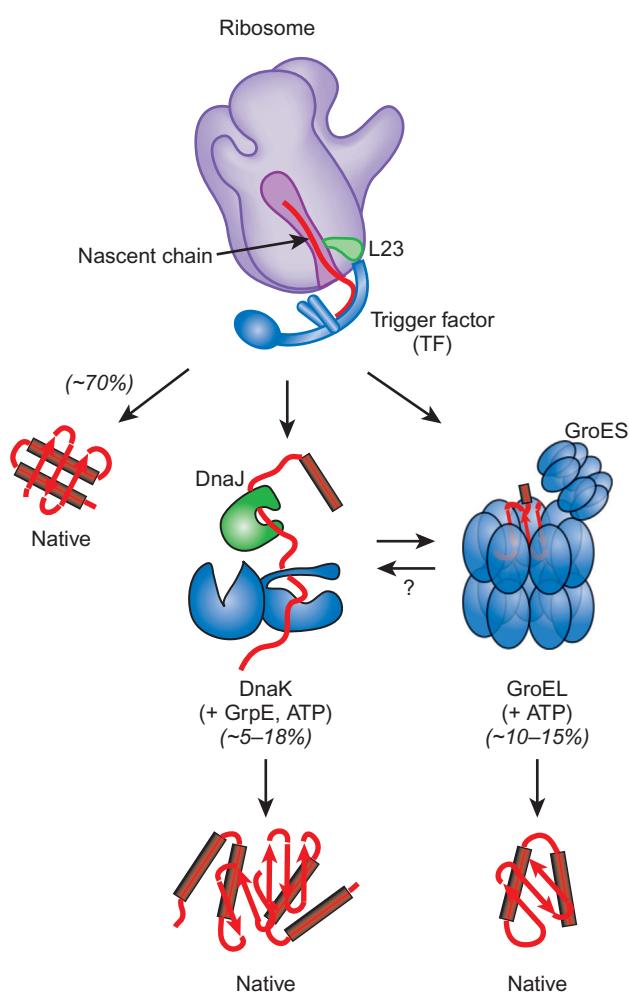


TABLE 13.06 Components of Cytoplasmic, Organelle, and Bacterial Ribosomes

Location	Subunits	Ribosomal RNA	Proteins
Animal Cytoplasm	40S	18S	33
	60S	28S, 5.8S, 5S	49
Animal Mitochondria	28S	12S	31
	39S	16S	48
Plant Cytoplasm	40S	18S	~35
	60S	28S, 5.8S, 5S	~50
Plant Chloroplast	30S	16S	22–31
	50S	23S, 5S, 4.5S	32–36
Plant Mitochondria	30S	18S	>25
	50S	26S, 5S	>30
Bacterial	30S	16S	21
	50S	23S, 5S	31
Archaeal	30S	16S	26–27
	50S	23S, 5S	30–31

resemble those for mitochondria, and in fact only plant cells can tell them apart. Thus, the mitochondria of fungi will import chloroplast proteins if genes encoding these are artificially introduced into the fungal cell. It is still unclear how plants

Box 13.03 Hybrid Ribosomes

Although they are larger, the cytoplasmic/nucleus-encoded ribosomes of eukaryotes resemble those from Archaea in the way they operate. A similar relation holds for the associated initiation and elongation factors. In fact, it is possible to make hybrid ribosomes containing one subunit from yeast and one from *Sulfolobus* (an archaeon). These still make protein, albeit less efficiently than native ribosomes. In contrast, hybrid ribosomes made by mixing subunits from yeast and *E. coli* are totally functionless.

decide between chloroplast and mitochondrial leader sequences; however, it seems that the leaders for the two kinds of organelle form different secondary structures.

Protein import by organelles also needs chaperonins on both sides of the membrane. An imported protein must travel through the narrow translocase channel in an uncoiled conformation. To avoid premature folding, newly synthesized organelle proteins are kept in a loosely folded conformation by chaperonins. Later, when the imported protein emerges from the translocase into the inside of the organelle, it is bound by another set of chaperonins. In particular, an Hsp70-type chaperonin is responsible for hauling in the incoming protein. The Hsp70 acts as a ratchet, binding to successive segments of unfolded polypeptide chain. Each binding and release of Hsp70 consumes energy in the form of ATP.

13. Mistranslation Usually Results in Mistakes in Protein Synthesis

Ribosomes are not perfect and make occasional mistakes. Perhaps 1 in 10,000 codons is misread and results in the wrong amino acid being incorporated. Two amino acids whose codons differ by only one base are most likely to be confused. Other possible errors are shifts in the reading frame (“**frameshift**”) or reading through stop codons. These assorted errors are collectively known as **mistranslation**.

Although such events are rare, a few weird genes actually require such errors for proper expression. For example, the *pol* gene of retroviruses (see Chapter 24: Viruses, Viroids, and Prions) is only translated if the ribosome frameshifts or reads through a stop codon while translating the preceding *gag* gene. In Chapter 10, Cell Division and DNA Replication, it was noted that the *dnaX* gene of *E. coli* gives rise to two proteins, tau and gamma, both subunits of DNA polymerase. The gamma protein is made only as a result of frameshifting. Release factor 2 (RF2) of *E. coli* also requires a frameshift for its synthesis.

14. Many Antibiotics Work by Inhibiting Protein Synthesis

Many well-known antibiotics work by inhibiting protein synthesis. Most of these are specific for prokaryotic ribosomes. However, very high concentrations of these agents will inhibit the ribosomes of mitochondria and chloroplasts, which are of prokaryotic ancestry.

Aminoglycoside antibiotics bind to the 30S subunit. **Streptomycin** binds to the 16S rRNA near where the two ribosomal subunits touch. The presence of streptomycin distorts the A-site and hinders binding of incoming charged tRNA. In particular, binding of initiator tRNA-Met is inhibited and so initiation of translation is prevented. Streptomycin-resistant mutants have alterations in nucleotide 523 of 16S rRNA or in ribosomal protein S12 (RpsL), which assists antibiotic binding. Many of

Streptomycin and related antibiotics bind to rRNA in the small subunit of the bacterial ribosome.

aminoglycosides Class of antibiotics that inhibits protein synthesis; includes streptomycin, neomycin, kanamycin, amikacin, and gentamycin.

frameshift Alteration in the reading frame during polypeptide synthesis.

mistranslation Errors made during translation.

streptomycin An antibiotic of the aminoglycoside family that inhibits protein synthesis.

Tetracycline binds rRNA in the small subunit of both prokaryotic and eukaryotic ribosomes.

Chloramphenicol binds to 23S rRNA and prevents peptide bond formation.

the other aminoglycosides, such as gentamycin and kanamycin, bind to multiple sites on the 30S subunit and mainly inhibit the translocation step of protein synthesis. Streptomycin and other aminoglycosides also cause misreading of the mRNA.

Tetracyclines inhibit both bacterial and eukaryotic ribosomes. They bind to the 16S (or 18S) rRNA of the small subunit and block the attachment of charged tRNA. Despite inhibiting both types of ribosome, tetracyclines inhibit bacteria preferentially due to the fact that bacteria actively take them up, whereas eukaryotic cells actively export them.

Chloramphenicol binds to the 50S subunit to the loop of 23S rRNA that interacts with the acceptor stem of the tRNA and inhibits the peptidyl transferase. **Cycloheximide** binds to the 60S subunit of eukaryotic ribosomes and inhibits the peptidyl transferase. **Erythromycin** and related macrolide antibiotics bind to the 23S rRNA of bacterial ribosomes and inhibit the translocation step.

Fusidic acid is a steroid derivative that binds to prokaryotic elongation factor EF-G. In the presence of fusidic acid, EF-G, with its bound GDP, is frozen in place on the ribosome. Fusidic acid also inhibits the corresponding eukaryotic elongation factor EF-2; however, in practice, animal cells are unaffected as they do not take up the antibiotic.

15. Post-Translational Modifications of Proteins

Although the genetic code has codons for only 20 amino acids, many other amino acids are occasionally found in proteins. Apart from selenocysteine and pyrrolysine (discussed later), these extra amino acids are made by modifying genetically encoded amino acids after the polypeptide chain has been assembled. This is known as **post-translational modification**. See Fig. 13.37 for a summary of some common post-translational modifications.

An example of medical importance is **diphthamide**, which is derived from histidine by post-translational modification (Fig. 13.38). It is found only in elongation factor eEF2 of eukaryotes and Archaea, in a region of the amino acid sequence that is highly conserved. The corresponding bacterial factor, EF-G, does not contain diphthamide.

Diphthamide was named after diphtheria, an infectious disease caused by the bacterium *Corynebacterium diphtheriae*. Diphtheria toxin attaches an ADP-ribose fragment to elongation factor eEF2 via diphthamide and this inhibits protein synthesis and kills the target cells. eEF2 normally splits GTP and uses the energy released to move the peptidyl-tRNA from the A-site to the P-site. ADP-ribosylated eEF2 still binds GTP but cannot hydrolyze it or translocate the peptidyl-tRNA.

16. Selenocysteine and Pyrrolysine: Rare Amino Acids

Selenocysteine (Sec) is not one of the standard 20 amino acids and yet it is incorporated into a few rare proteins during translation of the mRNA by the ribosome. This occurs both in bacteria and in eukaryotes, including humans. Sequencing of the genes and proteins involved has shown that selenocysteine is encoded by UGA. However, UGA is one of the stop codons. Apparently, UGA is normally read as “stop” but is occasionally translated to give selenocysteine, which therefore has the honor of being the 21st genetically encoded amino acid. The choice between “stop” and selenocysteine depends on a special recognition sequence in the following part

chloramphenicol An antibiotic that inhibits bacterial protein synthesis.

cycloheximide An antibiotic that inhibits eukaryotic protein synthesis.

diphthamide Modified amino acid found only in eukaryotic elongation factor eEF2 that is the target for diphtheria toxin.

erythromycin An antibiotic that inhibits bacterial protein synthesis.

fusidic acid An antibiotic that inhibits protein synthesis.

post-translational modification Modification of a protein or its constituent amino acids after translation is finished.

selenocysteine (Sec) Amino acid resembling cysteine but containing selenium instead of sulfur.

tetracyclines Family of antibiotics that inhibit protein synthesis.

GROUP NAME	GROUP ADDED
METHYLATION	protein — CH ₃
HYDROXYLATION	protein — OH
ACETYLATION	$\text{protein} - \overset{\text{O}}{\underset{\text{H}_3\text{C}}{\underset{\parallel}{\text{C}}} - \text{O}^-$
PHOSPHORYLATION	$\text{protein} - \overset{\text{O}}{\underset{\text{OH}}{\underset{\parallel}{\text{P}}} - \text{OH}$
GLYCOSYLATION (N-LINKED)	<p>N-acetylglucosamine</p>
GLYCOSYLATION (O-LINKED)	<p>N-acetylgalactosamine</p>
ADENYLYATION	<p>Adenine</p>

■ **FIGURE 13.37**
Common Post-Translational Modifications

The structure of some of the more common protein modification groups are shown.

of the gene—the **selenocysteine insertion sequence (SECIS element)**. Selenocysteine has its own tRNA and a special protein initiation factor to escort charged tRNA-Sec to the ribosome. In fact, selenocysteine-tRNA is initially charged with serine. Then the attached serine is enzymatically modified to form selenocysteine.

Very rarely, the stop codon UGA is read as the unusual amino acid selenocysteine.

selenocysteine insertion sequence (SECIS element) Recognition sequence that signals for insertion of selenocysteine at a UGA stop codon.

FIGURE 13.38
Histidine and Diphthamide

After the polypeptide chain of eEF2 has been made on the ribosome one specific histidine residue is converted to diphthamide. Diphtheria toxin stops protein synthesis by adding an ADP-ribose group to the side chain of diphthamide. This inhibits the action of eEF2.

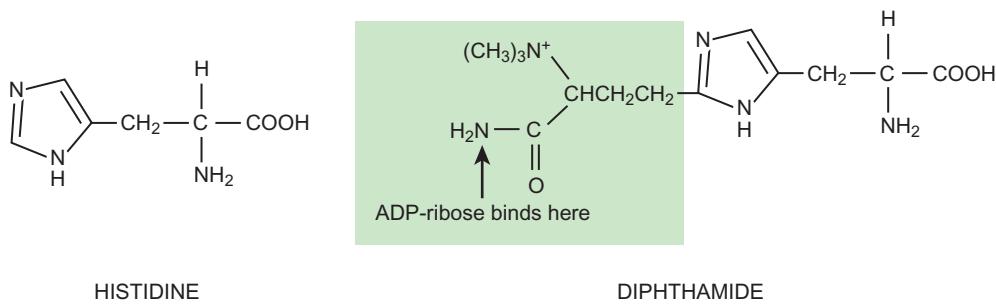
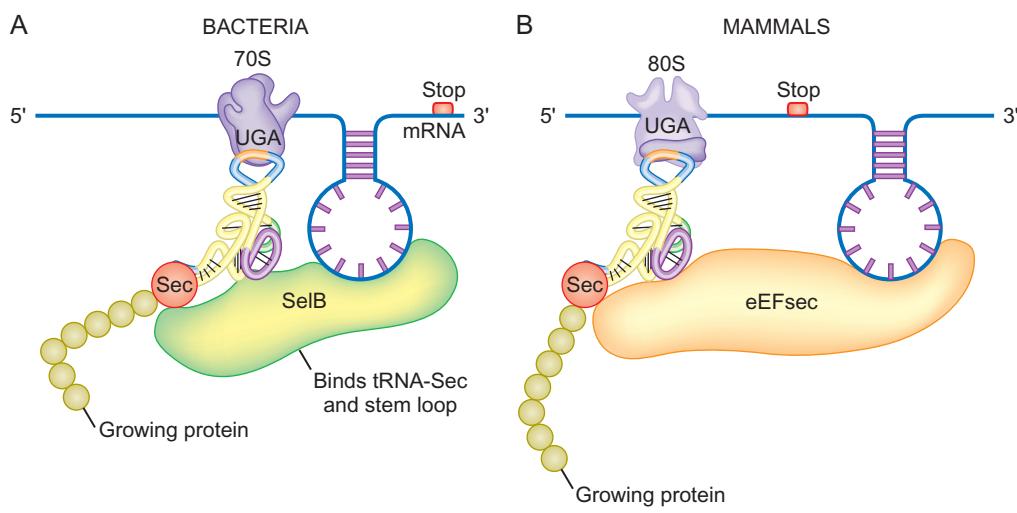


FIGURE 13.39
Delivery of tRNA With Selenocysteine to an Internal UGA Stop Codon

(A) In bacteria, the tRNA carrying selenocysteine (Sec) first binds to SelB and the complex then binds to a stem and loop in the mRNA. This aligns the tRNA-Sec with a UGA codon within the coding sequence on the mRNA. Selenocysteine is then inserted as part of the growing polypeptide. Only the fully bound complex is shown. (B) In mammals, the protein that binds the stem and loop and the tRNA-Sec is called eEFsec. In addition, the stem and loop are more distant, being found after the stop codon.



When bacteria use selenocysteine, the selenocysteine insertion sequence forms a stem and loop structure in the mRNA molecule just after the UGA. SelB protein recognizes both charged tRNA-Sec and this stem and loop. Thus, selenocysteine bound to tRNA is delivered to the right place (Fig. 13.39A). In bacteria, the stem and loop form temporarily from part of the coding sequence and this section of the mRNA is therefore translated after insertion of the selenocysteine.

In mammals, the stem and loop structure is found beyond the end of the coding sequence, in the 3'-untranslated region—not next to the critical UGA codon! SECIS binding protein 2 and the selenocysteine specific elongation factor, eEFsec, are needed for binding the tRNA-Sec and recognizing the SECIS stem and loop. The stem of this structure has a conserved sequence that forms a K-turn via nonstandard base pairing. This allows specific recognition. The tRNA-Sec is then delivered to the correct position for insertion (Fig. 13.39B). Humans have about 25 genes that encode selenoproteins. These are divided into two groups for regulation. Some are expressed under all conditions and others only when needed to combat oxidative stress.

Selenocysteine is an analog of cysteine, but has selenium instead of sulfur (Fig. 13.40). Selenium is more susceptible to oxidation than sulfur and so proteins that contain it must be protected from oxygen. Examples are the formate dehydrogenases found in many bacteria. These contain selenocysteine in their active sites and function in anaerobic metabolism. They are inactivated by oxygen and are normally made only in the absence of air. It has been suggested that the occurrence of selenoproteins in different groups of organisms is related to their oxygen sensitivity. Higher plants, which make oxygen, completely lack proteins that contain selenocysteine. Fungal genomes also totally lack selenoproteins. Conversely, fish, which live in the sea where oxygen levels are lower than on land, have more selenoproteins.

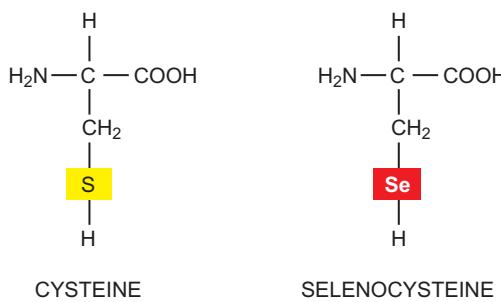


FIGURE 13.40
Selenocysteine and Cysteine

Selenocysteine is identical to cysteine except for the replacement of sulfur by selenium.

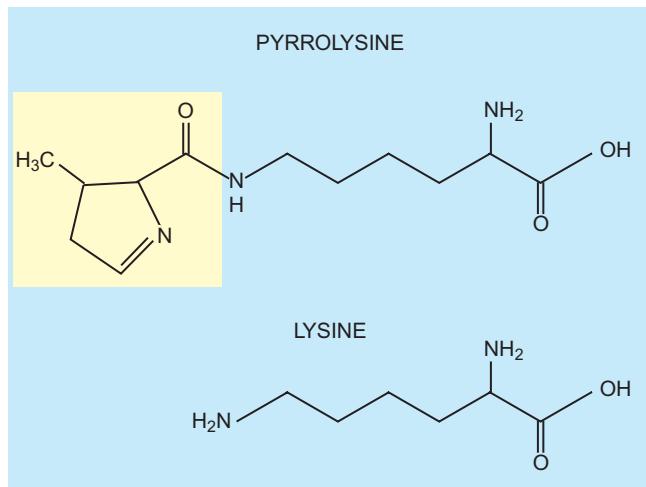


FIGURE 13.41
Pyrrolysine and Lysine

Pyrrolysine is (5R,5R)-4-methyl-pyrroline carboxylate.

than typical mammals. Indeed, zebrafish selenoprotein P contains 17 Sec residues, the largest number in any known protein.

In 2002, a 22nd genetically encoded amino acid was discovered—**pyrrolysine (Pyl)**, a derivative of lysine with an attached pyrrolidine ring (Fig. 13.41). This is found in a few Archaea where it is encoded by the stop codon UAG in occasional proteins. Pyrrolysine was first discovered in the active site of methylamine methyl-transferases found in methane-producing Archaea of the genus *Methanosaerina*. An unusual aminoacyl-tRNA synthase, a special tRNA, and genes for three accessory proteins are also found in organisms with pyrrolysine.

By analogy with selenocysteine, it was thought that pyrrolysine-tRNA was first charged with lysine, which was then modified to form pyrrolysine. However, this proved to be wrong. Pyrrolysine is made first as a free amino acid and then attached to tRNA-Pyl. There is no pyrrolysine specific elongation factor. Moreover, the sequence determinants that specify which UAG codons should be used for pyrrolysine insertion are unclear. Genome sequencing has found genes homologous to those for the pyrrolysine system in occasional Eubacteria suggesting that pyrrolysine may be present. However, pyrrolysine itself has not yet been identified directly in these organisms. The fact that the same five genes in the same order are found in both Archaea and Eubacteria suggests that horizontal gene transfer has occurred.

The stop codon UAG is very rarely translated as the rare amino acid, pyrrolysine.

17. Degradation of Proteins

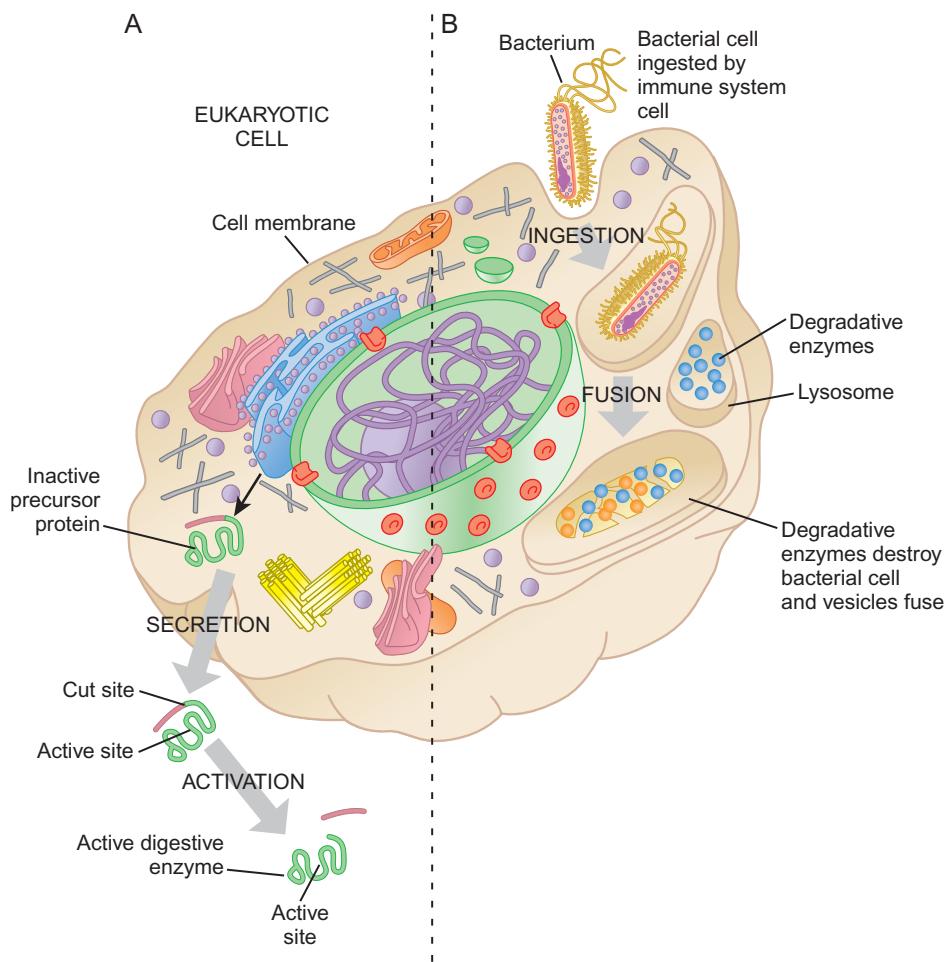
Living cells not only synthesize proteins, they also degrade them. Although protein degradation is nowhere near as complex as synthesis, it is nonetheless carefully regulated and often highly specific. **Proteases** (or proteinases) are enzymes that degrade

protease Same as proteinase; an enzyme that degrades proteins.

pyrrolysine (Pyl) 22nd Genetically encoded amino acid, derived from lysine.

FIGURE 13.42
Digestive Enzymes Are Activated on Location

(A) Proteases destined for export are made as precursors and are cleaved to form the active protease once safely outside the cell. (B) Proteases in the membrane-bound lysosome degrade ingested material.



proteins. They are potentially dangerous to the organism that makes them and must be carefully controlled. Proteases are often located in separate compartments where they can act without endangering other components of the organism. Alternatively, proteases may be designed so that they only accept specifically tagged proteins for degradation.

Proteases are found in three main locations: Extracellular, inside special compartments, and free in the cytoplasm. Animals secrete proteases into their digestive tracts. These enzymes are usually synthesized as inactive precursors and only activated once they are safely outside the cells of the animal that made them (Fig. 13.42A). Examples are trypsin (and its precursor trypsinogen) and pepsin (and its precursor pepsinogen). Plants that catch insects, fungi that trap nematodes, and bacteria that live in rotting animal or plant tissue also secrete proteases. As with animals, these proteases are generally secreted as inactive precursors and only activated once outside the cells of the producer organism.

Lysosomes are membrane-bound organelles found in eukaryotic cells. They contain a variety of digestive enzymes, including proteases, and function in self-defense. When cells of the immune system have engulfed bacteria or virus particles, the vesicles containing the invader are merged with lysosomes and the infectious agent is, hopefully, digested (Fig. 13.42B). Bacteria do not always cooperate—for example, many pathogenic strains of *Salmonella* can survive the toxins and digestive enzymes inside lysosomes.

Proteases located in the cytoplasm itself must be very carefully controlled. Nonetheless, the cell needs some internal proteases to degrade damaged or

Enzymes that degrade proteins are dangerous. They are frequently kept in separate compartments and often made as inactive precursors.

lysosome Membrane-bound organelle of eukaryotic cells that contains degradative enzymes.

misfolded proteins. Many proteases found inside bacterial cells form hollow cylinders, with the dangerous active site on the inside of the cylinder. Proteins slated for destruction are ferried to the protease cylinder and pushed into its center by accessory proteins. The number of misfolded proteins and consequently the level of protein degradation increases greatly under certain conditions, in particular when cells are exposed to uncomfortably high temperatures that tend to disrupt protein structure. This induces the heat shock response described in more detail in Chapter 16, Regulation of Transcription in Prokaryotes.

Eukaryotes have more sophisticated structures known as **proteasomes**. These are cylindrical, with the protease active sites inside. The top and bottom of the cylinder are covered by protein complexes that recognize and bind damaged or unwanted proteins. Proteins destined for degradation are identified by tagging with **ubiquitin**. This is a small protein that is fixed to damaged or misfolded proteins and also to certain proteins that are needed only for a brief period (Fig. 13.43). Ubiquitin tagged proteins are unfolded and then fed into the barrel of the proteasome where they are degraded into short peptides. The ubiquitin tags themselves are cleaved off and recycled.

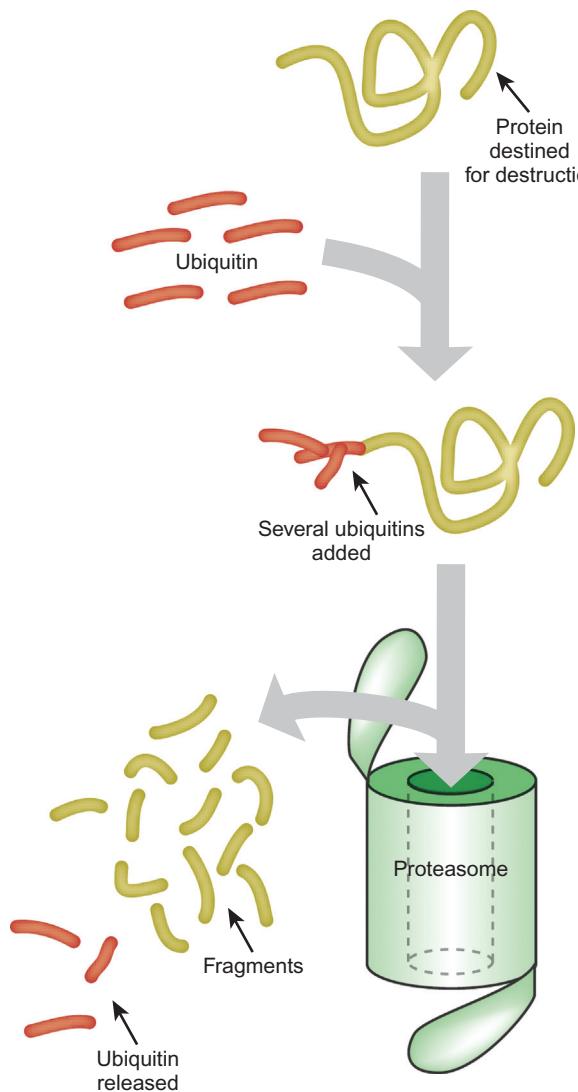


FIGURE 13.43
Operation of Proteasome

Ubiquitin tags damaged proteins and is recognized by the cylindrical proteasome. After degradation, the polypeptide fragments and ubiquitin are extruded.

proteasome Protein assembly found in eukaryotic cells that degrades proteins.

ubiquitin Small protein attached to other proteins as a signal that they should be degraded; used by eukaryotic cells, not bacteria.

Key Concepts

- Polypeptide chains are linear polymers of amino acids. There are 20 common genetically encoded amino acids plus two rare ones.
- Each amino acid is encoded by a codon consisting of three bases. Each codon is recognized by the corresponding anticodon on a transfer RNA (tRNA) to which the amino acid is attached.
- Translation is the synthesis of proteins using information from RNA and is carried out by the ribosome.
- Codons can be read in three possible reading frames, thus it is important to locate the correct start codon by using recognition sequences.
- During elongation of the polypeptide chain, the tRNA moves between three separate sites on the ribosome.
- Translation is terminated when the stop codon is read by a protein known as release factor. The ribosome subunits are then recycled.
- Several ribosomes usually read the same mRNA simultaneously.
- In bacteria, a single mRNA may encode several proteins. In addition, since there is no nucleus, transcription and translation are coupled.
- There are differences between protein synthesis in eukaryotic and prokaryotic cells.
- Proteins to be exported from the cell have signal sequences that direct them to the export machinery.
- Molecular chaperones are specialized proteins that oversee the correct folding of other proteins.
- Some proteins are made inside mitochondria and chloroplasts, but most organelle proteins are made on cytoplasmic ribosomes and transported into the organelles.
- Many antibiotics inhibit protein synthesis. These include tetracycline, aminoglycosides (such as streptomycin), and chloramphenicol.
- The rare amino acids selenocysteine and pyrrolysine are inserted at special stop codons.
- Proteins are broken down by proteases. In eukaryotes proteins tagged with ubiquitin are degraded in cylindrical assemblies called proteasomes.

Review Questions

1. What is a gene product? Are all gene products translated?
2. Describe three gene products that are not translated.
3. How common is the generalization “one gene—one protein” true? Briefly give three examples of how one gene can yield multiple proteins.
4. Would you expect to find the same number of genes in the genome as there are proteins in the proteome? Why or why not?
5. Translate the following mRNA sequence into protein:
5'-AUGCUAGCUCCUGAUUUCUA-3'
6. Name two modified bases commonly found in the T ψ C-loop and D-loops of tRNA molecules and explain why they are necessary.
7. If there are 61 codons (excluding the stop codons), then why are only 31 different tRNA molecules needed to carry out translation?
8. How is a tRNA molecule “charged” with an amino acid?
9. Compare and contrast ribosomes and their subunits in prokaryotes and eukaryotes.

10. What role does 23S rRNA play in synthesis of protein by a ribosome?
11. Why is the correct reading frame important in translating an mRNA into protein?
12. Which mutation is more detrimental for an ORF: Insertion of one extra nucleotide or insertion of three extra nucleotides? Why?
13. Summarize the initiation of translation in prokaryotes. Include the following terms: Initiator tRNA, N-formyl-methionine, Shine-Dalgarno sequence, 16S rRNA, AUG start codon, and initiation factors.
14. What other codon is sometimes used as a start codon instead of AUG? What types of proteins tend to use this codon?
15. Summarize the elongation step of translation in prokaryotes. Include the following terms: E-site, P-site, A-site, peptidyl transferase, translocation, elongation factors, and GTP.
16. What are the three stop codons? What proteins recognize them?
17. What is a polysome?
18. What is polycistronic mRNA? Is it found in prokaryotes, eukaryotes, or both?
19. How does a prokaryote recognize an mRNA that needs to be translated? How about a eukaryote?
20. What is meant by “coupled transcription and translation”? Does this occur in prokaryotes, eukaryotes, or both? Explain.
21. What is tmRNA? What role does it play during translation?
22. How does tmRNA tag proteins for degradation?
23. In what ways is translation in eukaryotes simpler than translation in prokaryotes?
24. Summarize initiation of translation in eukaryotes. In what ways is it different from prokaryotes?
25. How does the halting of protein synthesis due to nutrient deprivation differ between prokaryotes and eukaryotes?
26. What is a signal sequence and what purpose does it serve?
27. Summarize the process of cotranslational export.
28. What purposes do chaperone proteins serve?
29. Compare and contrast bacterial protein export with mitochondrial and chloroplast protein import.
30. Provide an example of how some organisms utilize “mistranslation” to create a functional protein.
31. Why is the genetic code not completely “universal”?
32. What is post-translational modification?
33. Why are selenocysteine and pyrrolysine considered the 21st and 22nd amino acids, whereas diphthamide is not considered the 23rd amino acid?
34. How does a ribosome know to incorporate selenocysteine when it encounters a UGA stop codon?
35. Which stop codon occasionally encodes pyrrolysine?
36. Describe the mechanism of action of the aminoglycoside antibiotics.
37. Do you think spontaneous streptomycin resistance is likely to occur in bacteria? Why or why not?
38. How does a cell protect itself from proteases that it normally exports after synthesis?
39. What are lysosomes?
40. How do eukaryotes dispose of unwanted proteins?

Further Reading

- Agris, P.F., et al., 2017. The importance of being modified: the role of RNA modifications in translational fidelity. *The Enzyme* 41:, 1–50. In Press.
- Asnani, M., Kumar, P., Hellen, C.U.T., 2015. Widespread distribution and structural diversity of type IV IRESSs in members of *Picornaviridae*. *Virology* 478, 61–74.
- Duechler, M., et al., 2016. Nucleoside modifications in the regulation of gene expression: focus on tRNA. *Cell Mol. Life Sci.* 73 (16), 3075–3095.
- Kuhn, C.-D., et al., 2015. On-enzyme refolding permits small RNA and tRNA surveillance by the CCA-adding enzyme. *Cell* 160, 644–658.
- Mansour, S.C., et al., 2016. Bacterial abscess formation is controlled by the stringent stress response and can be targeted therapeutically. *EBioMedicine* 12, 219–226.

- Wall, M.L., Lewis, S.M., 2017. Methylarginines within the RGG-motif region of hnRNP A1 affect its IRES *trans*-acting factor activity and are required for hnRNP A1 stress granule localization and formation. *J. Mol. Biol.* 429 (2), 295–307.
- Werner, F., 2012. A nexus for gene expression—molecular mechanisms of Spt5 and NusG in the three domains of life. *J. Mol. Biol.* 417, 13–27.
- Wilusz, J.E., et al., 2011. tRNAs marked with CCACCA are targeted for degradation. *Science* 334, 817–821.
- Yamashita, S., Tomita, K., 2016. Mechanism of 3'-matured tRNA discrimination from 3'-immature tRNA by Class-II CCA-adding enzyme. *Structure* 24 (6), 918–925.