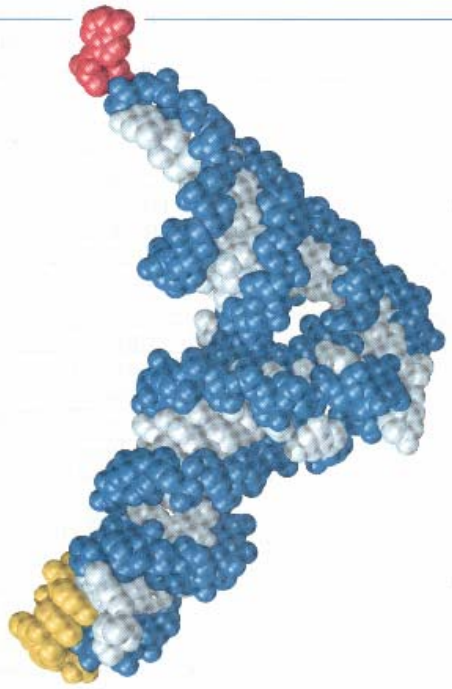
CHAPTER

34



Protein Synthesis

In Chapter 5, we considered the flow of genetic information from DNA to RNA to protein and the nature of the genetic code. We turn now to the mechanism of protein synthesis, a process called *translation* because the four-letter alphabet of nucleic acids is translated into the entirely different alphabet of proteins. As might be expected, translation is a more complex process than either replication or transcription, which take place within the framework of a common base-pairing language. In fact, translation necessitates the coordinated interplay of more than a hundred macromolecules. Transfer RNA molecules, mRNA, and many proteins are required in addition to ribosomes. The focus of this chapter will be on protein synthesis in *Escherichia coli* because it illustrates many general principles and is well understood. Some distinctive features of protein synthesis in eukaryotes will also be presented.

Let us take an overview of protein synthesis before examining it in some detail. A protein is synthesized in the amino-to-carboxyl direction by the sequential addition of amino acids to the carboxyl end of the growing peptide chain. The activated precursors are aminoacyl-tRNAs, in which the carboxyl group of an amino acid is joined to the 3′-OH of a transfer RNA (tRNA). The linking of an amino acid to its corresponding

Opening Image: Transfer RNAs, the central molecules of protein synthesis, serve as adaptors between the four-base language of nucleic acids and the twenty-amino acid language of proteins. The anticodon is shown in yellow and the attachment site for the amino acid in red. The RNA backbone is shown in dark blue and the bases in light blue. [Drawn from 6tna.pdb. J.L. Sussman, S.R. Holbrook, R.W. Warrant, G.M. Church, and S.-H. Kim. J. Mol. Biol. 123(1978):607.]

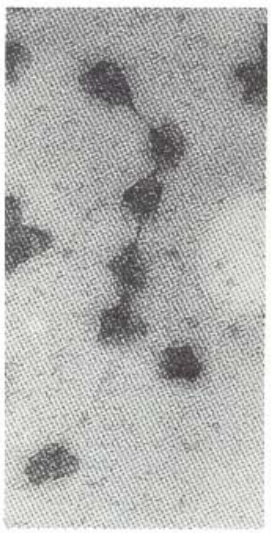


Figure 34-1
Electron micrograph of ribosomes on an mRNA molecule. Ribosomes are template-directed catalysts of peptide-bond formation. [Courtesy of Dr. Alex Rich.]

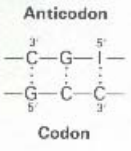
tRNA is catalyzed by an aminoacyl-tRNA synthetase. This activation reaction, which is analogous to the activation of fatty acids, is driven by ATP. For each amino acid, there is at least one kind of tRNA and activating enzyme.

Protein synthesis takes place in three stages: initiation, elongation, and termination. *Initiation* results in the binding of the initiator tRNA to the start signal of mRNA. The initiator tRNA occupies the P (peptidyl) site on a ribosome. *Elongation* starts with the binding of an aminoacyl-tRNA to the A (aminoacyl) site, a distinct tRNA-binding site on the ribosome. A peptide bond then forms between the amino group of the incoming aminoacyl-tRNA and the carboxyl group of the formylmethionine carried by the initiator tRNA. The resulting dipeptidyl-tRNA then moves from the A site to the P site, and the initiator tRNA molecule moves to the E (exit) site before leaving the ribosome. The binding of aminoacyl-tRNA, the movement of peptidyl-tRNA, and the associated movement of the ribosome to the next codon are powered by the hydrolysis of GTP. An aminoacyl-tRNA then binds to the vacant A site to start another round of elongation, which proceeds as described above. *Termination* occurs when a stop signal on the mRNA is read by a protein release factor, which leads to the release of the completed polypeptide chain from the ribosome. *Ribosomes, in essence, are enzymes that catalyze mRNA-directed formation of peptide bonds.*

Ribosomes are large ribonucleoprotein assemblies consisting of a small (30S) and a large (50S) subunit. They contain 55 different proteins and 3 RNA molecules. The name *ribosome* is apt because two-thirds of the mass is RNA. Indeed, ribosomal RNA molecules are information-rich molecules that play a central role in translation. Their action in contemporary ribosomes provides clues as to how primitive ribosomes consisting only of RNA may have carried out protein synthesis.

TRANSFER RNA (tRNA) MOLECULES HAVE A COMMON DESIGN

We begin our consideration of protein synthesis with tRNA because it serves as the adaptor molecule that recognizes both the enzyme that attaches the correct amino acid and the anticodon on mRNA. The base sequence of a tRNA molecule was first determined by Robert Holley in 1965, the culmination of seven years of effort. Indeed, his study of yeast alanine tRNA provided the first complete sequence of any nucleic acid and suggested how tRNA functions. This adaptor molecule is a single chain of 76 ribonucleotides (Figure 34-2). The 5′ terminus is phosphorylated (pG), whereas the 3′ terminus has a free hydroxyl group. A striking feature of this RNA is its high content of bases other than A, U, G, and C. Many unusual nucleosides are present: inosine (I, p. 887), pseudouridine (ψ, p. 850), dihydrouridine (UH_2, p. 878), ribothymidine (T, p. 850), and methylated derivatives of guanosine and inosine. The *amino acid attachment site* is the 3′-hydroxyl group of the adenosine residue at the 3′ terminus of the molecule. The sequence IGC in the middle of the molecule is the *anticodon*. It is complementary to GCC, one of the codons for alanine.



The sequences of several other tRNA molecules were determined a short time later. More than 100 sequences are now known. The striking finding is that all of them can be written in a cloverleaf pattern in which about half the residues are base-paired. Hence, *tRNA molecules have many common structural features.* This finding is not unexpected, because all

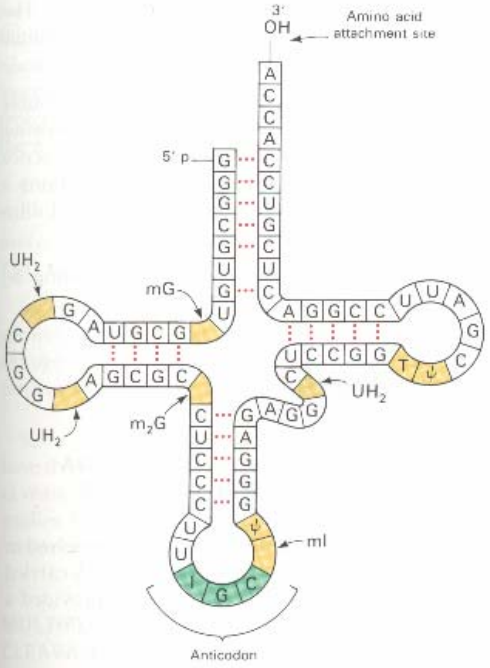


Figure 34-2
Base sequence of yeast alanine tRNA. Modified nucleosides (shown in yellow) are abbreviated as follows: methylinosine (mI), dihydrouridine (UH_2), ribothymidine (T), pseudouridine (ψ), methylguanosine (mG), and dimethylguanosine (m_2G). Inosine (I), another modified nucleoside, is part of the anticodon (green).

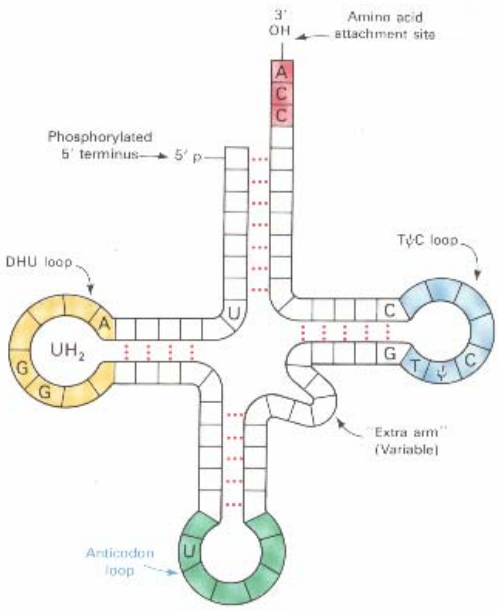


Figure 34-3
Common features of tRNA molecules.

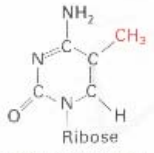
tRNA molecules must be able to interact in nearly the same way with ribosomes, mRNAs, and elongation factors. For example, they must fit into the A, P, and E sites on the ribosome and interact with the enzymatic site that catalyzes peptide-bond formation.

All known transfer RNA molecules share the following features (Figure 34-3):

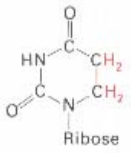
1. They are single chains containing between *73 and 93 ribonucleotides* (~25 kd).

2. They contain *many unusual bases*, typically between 7 and 15 per molecule. Some are methylated or dimethylated derivatives of A, U, C, and G that are formed by enzymatic modification of a precursor tRNA (p. 850). Methylation prevents the formation of certain base pairs, thereby rendering some of the bases accessible for other interactions. Also, methylation imparts a hydrophobic character to some regions of tRNAs, which may be important for their interaction with synthetases and ribosomal proteins and for their folding. Other modifications alter codon recognition, as will be discussed shortly.

3. The 5′ end of tRNAs is phosphorylated. The 5′ terminal residue is usually pG.



5-Methylcytidine

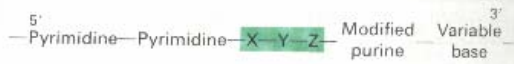


Dihydrouridine

4. The base sequence at the 3′ end of mature tRNAs is CCA. The activated amino acid is attached to the 3′-hydroxyl group of the terminal adenosine.

5. About half the nucleotides in tRNAs are base-paired to form double helices. Five groups of bases are not base paired: the 3′ *CCA terminal region;* the *TψC loop,* which acquired its name from the sequence ribothymine-pseudouracil-cytosine; the "*extra arm,*" which contains a variable number of residues; the *DHU loop,* which contains several dihydrouracil residues; and the *anticodon loop.*

6. The *anticodon loop* consists of seven bases, with the following sequence:



THE ACTIVATED AMINO ACID AND ANTICODON OF tRNA ARE AT OPPOSITE ENDS OF THE L-SHAPED MOLECULE

The three-dimensional structure of a tRNA molecule was first solved in 1974. X-ray crystallographic studies of yeast phenylalanine tRNA carried out in the laboratories of Alexander Rich and Aaron Klug provided a wealth of structural information:

1. The molecule is *L-shaped* (Figure 34-4).

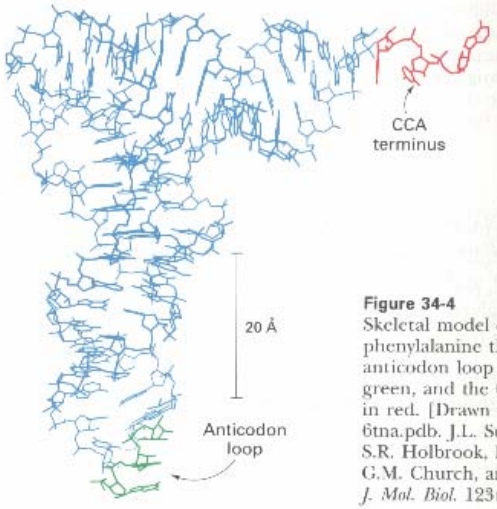


Figure 34-4
Skeletal model of yeast phenylalanine tRNA. The anticodon loop is shown in green, and the CCA terminus in red. [Drawn from 6tna.pdb. J.L. Sussman, S.R. Holbrook, R.W. Warrant, G.M. Church, and S.-H. Kim. *J. Mol. Biol.* 123(1978):607.]

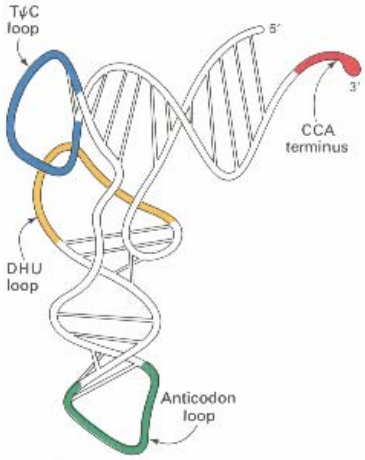


Figure 34-5
Schematic diagram of the three-dimensional structure of yeast phenylalanine tRNA. [After a drawing kindly provided by Dr. Sung-Hou Kim.]

2. There are *two segments of double helix.* They are like A-DNA, as expected for an RNA duplex (p. 790). Each of these helices contains about 10 base pairs, which correspond to one turn of helix. The helical segments are perpendicular to each other, giving the molecule its L shape (Figure 34-5). The base pairing in the cloverleaf model (p. 877), postulated on the basis of sequence studies, is correct.

3. Most of the bases in the nonhelical regions participate in unusual hydrogen-bonding interactions. These *tertiary interactions* are between bases that are not usually complementary (e.g., GG, AA, and AC). More-

over, the ribose-phosphate backbone interacts with some bases and even with another region of the backbone itself. The 2′–OH groups of the ribose units act as hydrogen bond donors or acceptors in many of these interactions. In addition, most bases are stacked. These hydrophobic interactions between adjacent aromatic rings play a major role in stabilizing the architecture of the molecule.

4. The CCA terminus containing the *amino acid attachment site* is at one end of the L. The other end of the L is the *anticodon loop*. Thus, *the amino acid in aminoacyl-tRNA is far from the anticodon* (about 80 Å). The DHU and TψC loops form the corner of the L.

5. The CCA terminus and the adjacent helical region do not interact strongly with the rest of the molecule. This part of the molecule may change its conformation during amino acid activation and also during protein synthesis on the ribosome.

Subsequent x-ray analyses of other prokaryotic and eukaryotic tRNAs have shown that their molecular architecture follows the same plan as that of yeast phenylalanine tRNA. As will be discussed shortly, crystallographic studies have also revealed how tRNAs interact with synthetases when amino acids become attached to their 3′ CCA terminus.

MULTIPLE TRANSFER RNA MOLECULES ARISE FROM CLEAVAGE OF A LARGE PRECURSOR BY RIBONUCLEASE P

The 60 genes for transfer RNA molecules in *E. coli* are clustered in 25 units that are transcribed into multimeric precursors. Some units encode ribosomal RNAs as well, whereas others contain only tRNAs. One of these transcripts is the precursor of seven tRNAs: one specific for leucine, two for internal methionines, and two each for two kinds of glutamine codons (Figure 34-6). The primary transcript of 950 nucleotides is cleaved by *ribonuclease P* (RNase P) on the 5′ side of the first nucleotide of each mature tRNA-to-be. *Ribonuclease D* (RNase D) then trims the exposed 3′ end of each precursor until it reaches the CCA sequence, which becomes the 3′ terminus of the mature molecule.

Ribonuclease P is a ribonucleoprotein enzyme comprising a 377-nucleotide M1 RNA molecule and a 20-kd protein. In 1983, Sidney Altman discovered that the RNA component alone possesses enzymatic activity. At higher than physiologic concentrations of Mg^{2+} (~60 mM), *M1 RNA by itself recognizes target sites in primary transcripts and cleaves them at an appreciable rate. This experiment showed that M1 RNA interacts specifically with the substrate and possesses the catalytic groups.* The role of the protein is subsidiary—it increases the hydrolytic rate and enables the reaction to occur at a much lower concentration of Mg^{2+}. We see here another example of *catalytically active RNA* and catch a glimpse of the early RNA world.

AMINO ACIDS ARE ACTIVATED AND LINKED TO PARTICULAR TRANSFER RNAs BY SPECIFIC SYNTHETASES

The formation of a peptide bond between the amino group of one amino acid and the carboxyl group of another is thermodynamically unfavorable. This barrier is overcome by activating the carboxyl group of the precursor amino acids. *The activated intermediates in protein synthesis are amino acid esters,* in which the carboxyl group of an amino acid is linked to

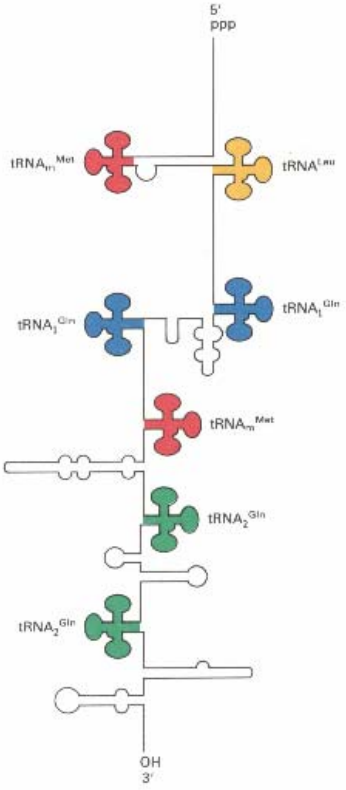


Figure 34-6
Seven tRNA molecules are formed by cleavage of this 950-nucleotide primary transcript. [After N. Nakajima, H. Ozeki, and Y. Shimura. *Cell* 23(1981):245.]

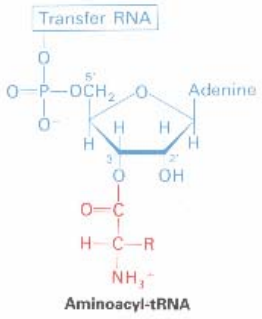
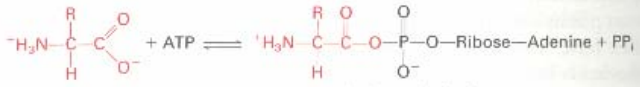


Figure 34-7
An amino acid is esterified to the 2′- or the 3′-hydroxyl group of the terminal adenosine in an aminoacyl-tRNA.

either the 2′- or the 3′-hydroxyl group of the ribose unit at the 3′ end of tRNA. In some tRNAs, the activated amino acid migrates very rapidly between the 2′- and 3′-hydroxyl groups. An amino acid ester of tRNA is called an *aminoacyl-tRNA* (Figure 34-7); it is sometimes called a *charged tRNA*.

The attachment of an amino acid to a tRNA is important not only because it activates the carboxyl group, but also because amino acids by themselves cannot recognize the codons on mRNA. Rather, amino acids must be carried to the ribosomes by specific tRNAs, which do recognize codons on mRNA and thereby act as adaptor molecules.

In 1957, Paul Zamecnik and Mahlon Hoagland discovered that the activation of amino acids and their subsequent linkage to tRNAs are catalyzed by specific *aminoacyl-tRNA synthetases*, which are also called *activating enzymes*. The first step is the formation of an *aminoacyl-adenylate* from an amino acid and ATP. This activated species is a mixed anhydride in which the carboxyl group of the amino acid is linked to the phosphoryl group of AMP; hence, it is also known as *aminoacyl-AMP*.



Aminoacyl adenylate (Aminoacyl-AMP)

The next step is the transfer of the aminoacyl group of aminoacyl-AMP to a tRNA molecule to form *aminoacyl-tRNA*.

$$\text{Aminoacyl-AMP} + \text{tRNA} \rightleftharpoons \text{aminoacyl-tRNA} + \text{AMP}$$

The sum of these activation and transfer steps is

$$\text{Amino acid} + \text{ATP} + \text{tRNA} \rightleftharpoons \text{aminoacyl-tRNA} + \text{AMP} + \text{PP}_i$$

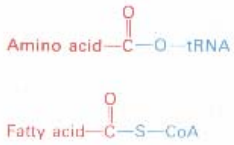
The $\Delta G^{\circ\prime}$ of this reaction is close to 0, because the free energy of hydrolysis of the ester bond of aminoacyl-tRNA is similar to that of the terminal phosphoryl group of ATP. What then drives the synthesis of aminoacyl-tRNA? As expected, the reaction is driven by the hydrolysis of pyrophosphate. The sum of these three reactions is highly exergonic:

$$\text{Amino acid} + \text{ATP} + \text{tRNA} + \text{H}_2\text{O} \longrightarrow \text{aminoacyl-tRNA} + \text{AMP} + 2\ \text{P}_i$$

Thus, *two ~P are consumed in the synthesis of an aminoacyl-tRNA*. One of them is consumed in forming the ester linkage of aminoacyl-tRNA, whereas the other is consumed in driving the reaction forward.

The activation and transfer steps for a particular amino acid are catalyzed by the same aminoacyl-tRNA synthetase. In fact, *the aminoacyl-AMP intermediate does not dissociate from the synthetase*. Rather, it is tightly bound to the active site of the enzyme by noncovalent interactions. The aminoacyl-AMP is normally a transient intermediate in the synthesis of aminoacyl-tRNA, but it is quite stable and readily isolated if tRNA is absent from the reaction mixture.

We have already encountered an acyl adenylate intermediate in fatty acid activation (p. 607). In fact, Paul Berg first discovered this intermediate in fatty acid activation, and then recognized that it is also formed in amino acid activation. The major difference between these reactions is that the acceptor of the acyl group is CoA in the former and tRNA in the latter. The energetics of these biosyntheses are very similar: both are made irreversible by the hydrolysis of pyrophosphate.



AMINOACYL-tRNA SYNTHETASES BELONG TO TWO STRUCTURAL CLASSES

At least one aminoacyl-tRNA synthetase exists for each amino acid. The diverse sizes, subunit composition, and sequences of these enzymes were bewildering for many years. However, recent studies have shown that synthetases can be grouped into two classes, termed *class I* and *class II*, according to the presence of short signature sequences. The synthetases for 10 of the basic set of 20 amino acids belong to class I enzymes, and those for the other 10 to class II (Table 34-1). It is interesting to note that the smaller amino acids are generally activated by class II synthetases, whereas the larger amino acids and also the more hydrophobic ones are activated by class I enzymes. The class II enzymes may be the more ancient ones. Another difference is that class I enzymes acylate the 2′-hydroxyl group of the terminal adenosine of tRNA, whereas class II enzymes (except the one for Phe) acylate the 3′-hydroxyl. The amino acid activating domain of the two classes is also different. Class I enzymes have a parallel β domain (the classical dinucleotide binding fold, p. 498), whereas class II enzymes have an antiparallel β domain. They also differ in how they recognize tRNA (p. 886).

Table 34-1
Classification and subunit structure of aminoacyl-tRNA synthetases in *E. coli*

Class I	*Class II*
Arg (α)	Ala (α_4)
Cys (α)	Asn (α_2)
Gln (α)	Asp (α_2)
Glu (α)	Gly ($\alpha_2\beta_2$)
Ile (α)	His (α_2)
Leu (α)	Lys (α_2)
Met (α)	Phe ($\alpha_2\beta_2$)
Trp (α_2)	Ser (α_2)
Tyr (α_2)	Pro (α_2)
Val (α)	Thr (α_2)

TYROSYL-AMP FORMATION IS GREATLY ACCELERATED BY THE BINDING OF γ PHOSPHATE IN THE TRANSITION STATE

X-ray crystallographic and protein engineering studies have provided insight into the catalytic action of *E. coli* tyrosyl-tRNA synthetase, a class I dimer of 47-kd subunits. The amino-terminal 320 residues are needed for the activation reaction, whereas the carboxyl-terminal 99 residues participate in the binding of tRNA and the formation of tyrosyl-tRNA. The crystal structure of the synthetase containing bound tyrosyl-AMP has been solved at high resolution. This activated intermediate is stable in the absence of the matching tRNA and is bound to the enzyme by some 12 hydrogen bonds (Figure 34-8).

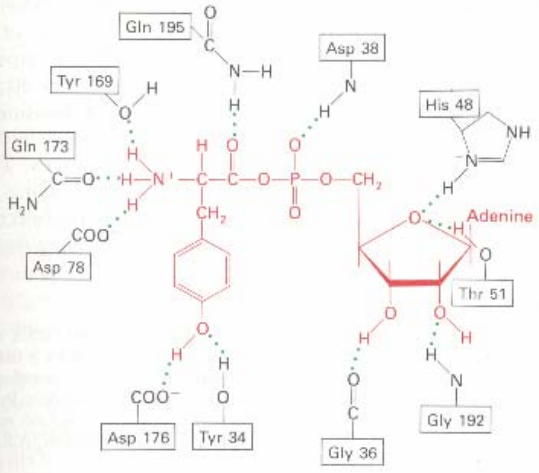
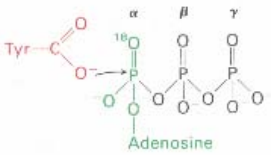


Figure 34-8
Tyrosyl-AMP (shown in red) is bound to its synthetase by multiple hydrogen bonds. [Courtesy of Dr. David Blow.]



The formation of tyrosyl-AMP from tyrosine and ATP stereospecifically labeled with ^{18}O at the α phosphoryl group leads to an *inversion* of configuration. Hence, the reaction probably proceeds by an *in-line displacement* in which the tyrosyl carboxylate is the attacking nucleophile and pyrophosphate is the leaving group. The α phosphorus atom in the transition state is pentacovalent and has the geometry of a trigonal bipyramid, as in the hydrolytic reaction catalyzed by ribonuclease (p. 218) and the self-splicing of a ribozyme (p. 868).

A plausible structure of the transition state was deduced by model building (Figure 34-9). A key feature is the hydrogen bonding of the γ phosphate group to the side chains of threonine 40 and histidine 45.

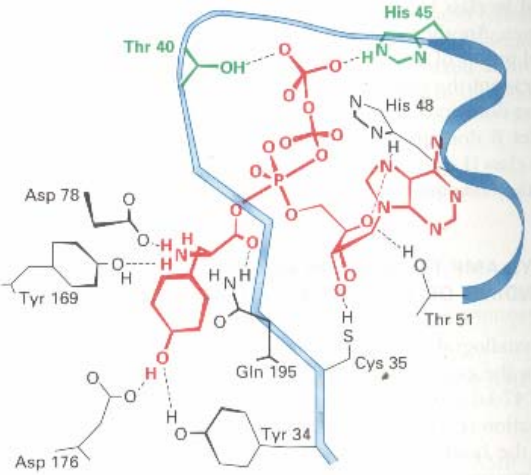
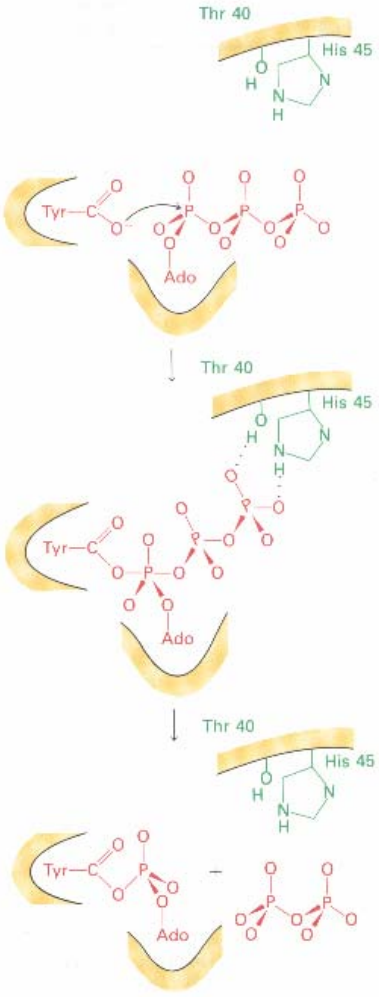


Figure 34-9
Proposed structure of the transition state in the formation of tyrosyl-AMP from tyrosine and ATP (shown in red). The side chains of threonine 40 and histidine 45 (shown in green) play a critical role in catalysis by selectively binding the γ phosphate in the transition state. [After R.J. Leatherbarrow, A.R. Fersht, and G. Winter. *Proc. Nat. Acad. Sci.* 82(1985):7841.]



Their importance was assessed by mutating both to alanine, to eliminate the possibility of hydrogen bonding. Indeed, the catalytic activity of the double mutant was less than that of the wild type by a factor of 3×10^6, whereas the binding affinities of the enzyme for ATP and tyrosine were altered relatively little. This experiment shows that threonine 40 and histidine 45 are essential for catalysis but not for substrate binding (Figure 34-10). They probably interact strongly with the γ phosphate in the transition state but not in the initial enzyme-substrate complex. The large change in position of the pyrophosphate unit accompanying the shift from a tetrahedral to a bipyramidal geometry triggers this selective binding. Recall that *the essence of catalysis is selective stabilization of the transition state* (p. 188).

Figure 34-10
Proposed mechanism for the formation of tyrosyl-AMP by tyrosyl-tRNA synthetase. The α phosphorus atom of ATP is nucleophilically attacked by a carboxylate oxygen of tyrosine. The transition state is stabilized by hydrogen bonding of the γ phosphate to the side chains of a threonine and a histidine residue. Pyrophosphate bound to Mg^{2+} (not shown) leaves the pentacovalent transition state to give tyrosyl-AMP. Ado denotes adenosine. [After R.J. Leatherbarrow, A.R. Fersht, and G. Winter. *Proc. Nat. Acad. Sci.* 82(1985):7841.]

Which groups on the enzyme directly participate in making and breaking bonds? Probably none. The carboxylate group of tyrosine is an intrinsically effective nucleophile, ATP is already activated, and Mg^{2+}–pyrophosphate is a good leaving group. The enzyme accelerates catalysis by a factor of about 4×10^4 simply by bringing tyrosine and ATP together, and it gains another factor of 3×10^5 mainly by binding γ phosphate in the transition state. It is noteworthy that the new interactions in the transition state are at some distance from the α phosphorus atom, the reaction center. Thus, *catalysis can be delocalized*—what counts is selective binding of the transition state, however achieved.

PROOFREADING BY AMINOACYL-tRNA SYNTHETASES INCREASES THE FIDELITY OF PROTEIN SYNTHESIS

Aminoacyl-tRNA synthetases are highly selective in their recognition of both the amino acid to be activated and the prospective tRNA acceptor. As will be discussed shortly, tRNA molecules that accept different amino acids have different base sequences, and so they can be readily distinguished by their synthetases. A much more demanding task for these enzymes is to discriminate between similar amino acids. For example, the only difference between isoleucine and valine is that isoleucine contains a methylene group not present in valine (Figure 34-11). The additional binding energy (~−3 kcal/mol) contributed by this extra $-CH_2-$ group favors the activation of isoleucine over valine by isoleucyl-tRNA synthetase by a factor of about 200. Even so, the concentration of valine in vivo is about five times that of isoleucine, and so valine would be mistakenly incorporated into proteins in place of isoleucine 1 in 40 times. However, the observed error frequency in vivo is only 1 in 3000, indicating that there must be a subsequent editing step to enhance fidelity.

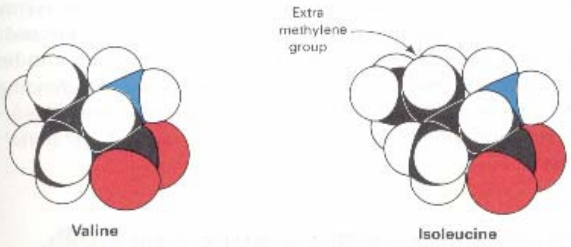


Figure 34-11
Molecular models of valine and isoleucine. The extra methylene group in isoleucine is marked. The synthetases specific for these amino acids are highly discerning.

In fact, *the synthetase corrects its own errors.* Mistakenly activated valine is not transferred to tRNA specific for isoleucine. Instead, *this tRNA promotes the hydrolysis of valine–AMP and thereby prevents its erroneous incorporation into proteins* (Figure 34-12). Furthermore, hydrolysis frees the synthetase to activate and transfer isoleucine, the correct amino acid. How does the synthetase avoid hydrolyzing isoleucine–AMP, the correct intermediate? Most likely, the hydrolytic site is just large enough to accommodate valine–AMP but too small to allow the entry of isoleucine–AMP.

How do synthetases distinguish between valine and threonine, a pair of amino acids that are very nearly identical in size? Choosing between these

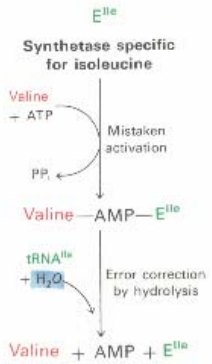


Figure 34-12
Proofreading by hydrolysis of an incorrect aminoacyl-AMP. The entry of tRNA specific for isoleucine induces hydrolysis of valyl-AMP.

amino acids is demanding because threonine differs from valine only in having an $-OH$ in place of a $-CH_3$ group. The aminoacyl-tRNA synthetase for valine contains two adjacent catalytic sites, one for the acylation of tRNA and the other for hydrolysis of incorrectly acylated tRNA (Figure 34-13). Valine is preferred over threonine in the acylation reaction because the acylation site is hydrophobic. By contrast, threonyl-tRNA is hydrolyzed much more rapidly than valyl-tRNA because the hydrolytic site is hydrophilic. The synthetase for valine does most of its editing at the level of aminoacyl-tRNA, whereas the one for isoleucine does so at the level of aminoacyl-AMP.

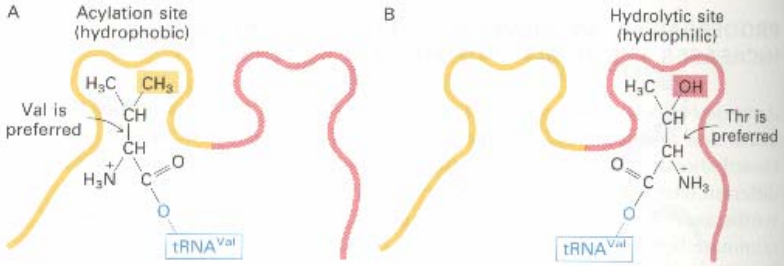


Figure 34-13
Valyl-tRNA synthetase rejects threonine at two stages. (A) The hydrophobic acylation site for the transfer of activated amino acid to tRNA prefers valine over threonine because valine is more hydrophobic. (B) The adjacent hydrolytic site, by contrast, is hydrophilic. Hence, threonine erroneously linked to the tRNA is preferentially hydrolyzed because the binding energy of the hydroxyl group of threonine is used to stabilize the transition state. [After A. Fersht. *Enzyme Structure and Mechanism,* 2nd ed. (W.H. Freeman, 1984), p. 356.]

Most aminoacyl-tRNA synthetases contain hydrolytic sites in addition to acylation sites. Complementary pairs of sites function as a *double sieve* to assure very high fidelity. The acylation site rejects amino acids that are *larger* than the correct one because there is insufficient room for them, whereas the hydrolytic site destroys activated intermediates that are *smaller* than the correct species. Hydrolytic proofreading is central to the fidelity of many aminoacyl-tRNA synthetases, as it is to DNA polymerases (p. 800). However, a few synthetases achieve high accuracy without editing their covalently attached intermediates. For example, tyrosyl-tRNA synthetase has no difficulty discriminating between tyrosine and phenylalanine; the hydroxyl group on the tyrosine ring enables it to be bound to the enzyme 10^4 times as strongly as phenylalanine. *Proofreading is costly in energy and time and hence is selected in the course of evolution only when fidelity must be enhanced beyond what can be obtained through an initial binding interaction.*

SYNTHETASES RECOGNIZE THE ANTICODON LOOP AND ACCEPTOR STEM OF TRANSFER RNA MOLECULES

How do synthetases choose their tRNA partners? Precise recognition of tRNAs is equally important for high-fidelity protein synthesis as is accurate selection of amino acids. A priori, the anticodon of tRNA would seem to be a good choice for establishing its identity because each tRNA has a different one. Indeed, *some synthetases recognize their tRNA partner primarily on the basis of its anticodon.* The most direct evidence comes from *identity swap* experiments. The CCA anticodon of tryptophanyl-tRNA (abbreviated $tRNA^{Trp}$) and the UAC anticodon of $tRNA^{Val}$ were replaced with CAU, the anticodon for methionine (Figure 34-14A). The genes for these altered tRNAs were transcribed in vitro, and the kinetics of acylation by the synthetase specific for methionine were measured. Changing the anticodon of $tRNA^{Trp}$ and $tRNA^{Val}$ to CAU increases the rate of aminoacylation with methionine more than 50,000-fold. In fact, $tRNA^{Val}$

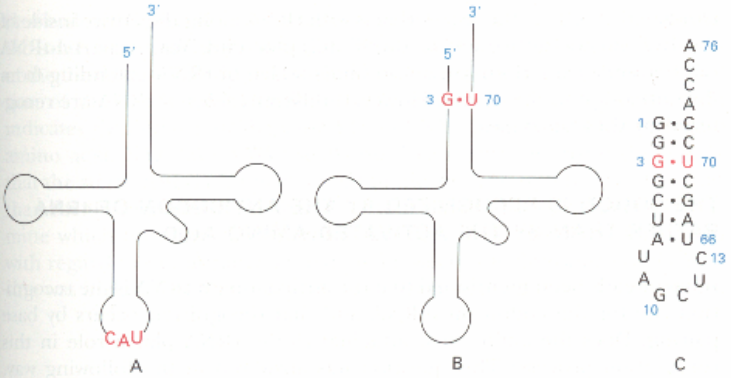


Figure 34-14
Determinants of the identity of tRNA molecules in recognition by synthetases. (A) The identity of $tRNA^{Val}$ is changed from Val to Met by mutating its anticodon from UAC to CAU. (B) The identity of $tRNA^{Cys}$ is changed from Cys to Ala by changing its 3:70 base pair from C · G to G · U. (C) A "microhelix" containing only 24 of the 76 nucleotides of $tRNA^{Ala}$ is recognized by the cognate synthetase.

bearing the CAU anticodon is aminoacylated by methionyl-tRNA synthetase at nearly the same rate as is $tRNA^{Met}$.

In contrast, the anticodon of $tRNA^{Ala}$ can be changed with no effect on the aminoacylation specificity. Rather, *the crucial determinant is the presence of a G · U base pair at the 3:70 position in the 3' acceptor stem of this 76-nucleotide molecule* (Figure 34-14B). Cysteine tRNA differs from $tRNA^{Ala}$ at 40 positions and contains a C · G base pair at the 3:70 position. When this C · G base pair is changed to G · U, $tRNA^{Cys}$ is recognized by alanyl-tRNA synthetase as though it were $tRNA^{Ala}$. This finding raised the question as to whether a fragment of tRNA suffices for aminoacylation by alanine-tRNA synthetase. Indeed, *a "microhelix" containing just 24 of the 76 nucleotides of the native tRNA is specifically aminoacylated by the synthetase.* This microhelix contains only the 3' acceptor stem and a hairpin loop (Figure 34-14C).

The identities of other tRNAs are conferred by multiple determinants. For example, *both the anticodon and acceptor stems of $tRNA^{Gln}$ are recognized by glutaminyl-tRNA synthetase.* X-ray crystallographic studies of a complex of this synthetase, $tRNA^{Gln}$, and ATP have provided a revealing view of how this enzyme recognizes its tRNA partner (Figure 34-15). This highly

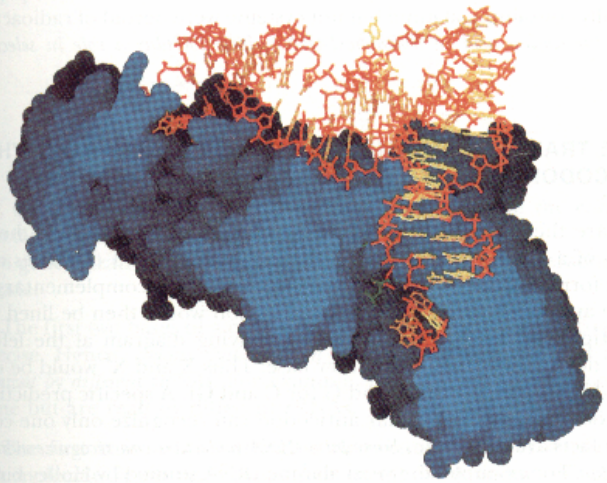
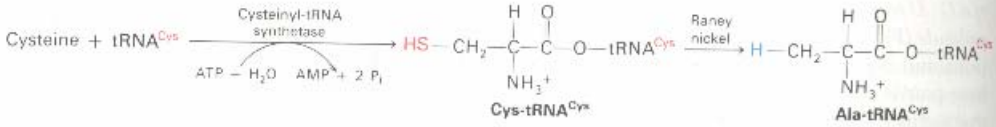


Figure 34-15
Structure of a tRNA bound to its cognate synthetase. The sugar-phosphate backbone of $tRNA^{Gln}$ is shown in red and the bases in yellow. The solvent-accessible surface of glutaminyl-tRNA synthetase is depicted in blue. ATP (green) is also bound to the synthetase. [After M.A. Rould, J.J. Perona, D. Söll, and T.A. Steitz. *Science* 246(1989):1135.]

elongated class I synthetase interacts with tRNA along the entire inside of the L, from the anticodon to the 3′ acceptor end. Yeast aspartyl-tRNA synthetase, a class II enzyme, also binds a face of tRNA extending from the anticodon to the 3′ end. However, different sides of tRNA are recognized by these enzymes.

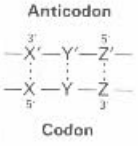
THE CODON IS RECOGNIZED BY THE ANTICODON OF tRNA RATHER THAN BY THE ACTIVATED AMINO ACID

It has already been mentioned that the anticodon on tRNA is the recognition site for the codon on mRNA and that recognition occurs by base pairing. Does the amino acid attached to the tRNA play a role in this recognition process? This question was answered in the following way. First, cysteine was attached to its cognate tRNA. The attached cysteine unit was then converted into alanine by reacting Cys-$tRNA^{Cys}$ with Raney nickel, which removed the sulfur atom from the activated cysteine residue without affecting its linkage to tRNA. Thus, a *hybrid aminoacyl-tRNA* was produced in which alanine was covalently attached to a tRNA specific for cysteine.



Does this hybrid tRNA recognize the codon for alanine or for cysteine? The answer came on adding the tRNA to a cell-free protein-synthesizing system. The template was a random copolymer of U and G in the ratio of 5:1, which normally leads to the incorporation of cysteine (encoded by UGU) but not of alanine (encoded by GCX). However, alanine was incorporated into a polypeptide when Ala-$tRNA^{Cys}$ was added to the incubation mixture, because it was attached to the tRNA specific for cysteine. The same result was obtained when mRNA for hemoglobin served as the template and ^{14}C alanyl-$tRNA^{Cys}$ was used as the hybrid aminoacyl-tRNA. The only radioactive tryptic peptide produced was one that normally contained cysteine but not alanine. On the other hand, peptides normally containing alanine but not cysteine were devoid of radioactivity. Thus, *the amino acid in aminoacyl-tRNA does not play a role in selecting a codon.*

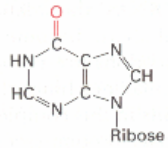
SOME TRANSFER RNA MOLECULES RECOGNIZE MORE THAN ONE CODON BECAUSE OF WOBBLE IN BASE PAIRING



What are the rules that govern the recognition of a codon by the anticodon of a tRNA? A simple hypothesis is that each of the bases of the codon forms a Watson-Crick type of base pair with a complementary base on the anticodon. The codon and anticodon would then be lined up in an antiparallel fashion. In the accompanying diagram at the left, the prime denotes the complementary base. Thus X and X′ would be either A and U (or U and A) or G and C (or C and G). A specific prediction of this model is that a particular anticodon can recognize only one codon.

The facts are otherwise. *Some pure tRNA molecules can recognize more than one codon.* For example, the yeast alanine tRNA studied by Holley binds to

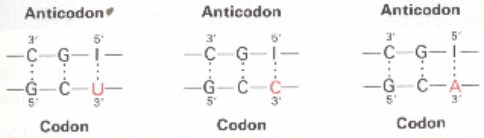
three codons: GCU, GCC, and GCA. The first two bases of these codons are the same, whereas the third is different. Could it be that the recognition of the third base of a codon is sometimes less discriminating than recognition of the other two? The pattern of degeneracy of the genetic code indicates that this might be so. XYU and XYC always code for the same amino acid; XYA and XYG usually do. Crick surmised from these data that the steric criteria for pairing of the third base might be less stringent than for the other two. Models of various base pairs were built to determine which ones are similar to the standard A · U and G · C base pairs with regard to the distance and angle between the glycosidic bonds. Inosine was included in this study because it appeared in several anticodons.



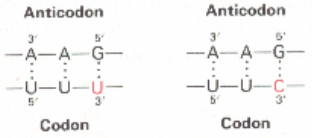
Inosine

Assuming some steric freedom (*"wobble"*) in the pairing of the third base of the codon, the combinations shown in Table 34-2 seemed plausible.

The wobble hypothesis is now firmly established. The anticodons of tRNAs of known sequence bind to the codons predicted by this hypothesis. For example, the anticodon of yeast alanine tRNA is IGC. This tRNA recognizes the codons GCU, GCC, and GCA. Recall that, by convention, nucleotide sequences are written in the 5′ → 3′ direction unless otherwise noted. Hence, I (the 5′ base of this anticodon) pairs with U, C, or A (the 3′ base of the codon), as predicted.



Phenylalanine tRNA, which has the anticodon GAA, recognizes the codons UUU and UUC but not UUA and UUG.



Thus, G pairs with either U or C in the third position of the codon, as predicted by the wobble hypothesis.

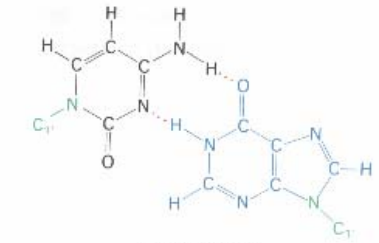
Two generalizations concerning the codon-anticodon interaction can be made:

1. The first two bases of a codon pair in the standard way. Recognition is precise. Hence, *codons that differ in either of their first two bases must be recognized by different tRNAs.* For example, both UUA and CUA code for leucine but are read by different tRNAs.

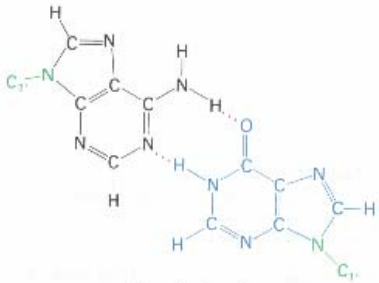
2. The first base of an anticodon determines whether a particular tRNA molecule reads one, two, or three kinds of codons: C or A (1

Table 34-2
Allowed pairings at the third base of the codon according to the wobble hypothesis

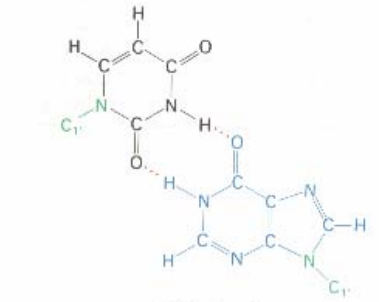
First base of anticodon	*Third base of codon*
C	G
A	U
U	A or G
G	U or C
I	U, C, or A



Cytidine-inosine base pair



Adenosine-inosine base pair



Uridine-inosine base pair

Figure 34-16
Inosine can base-pair with cytosine, adenine, or uracil because of wobble.

codon), U or G (2 codons), or I (3 codons). Thus, *part of the degeneracy of the genetic code arises from imprecision (wobble) in the pairing of the third base of the codon with the first base of the anticodon.* We see here a strong reason for the frequent appearance of inosine, one of the unusual nucleosides, in anticodons. *Inosine maximizes the number of codons that can be read by a particular tRNA molecule* (Figure 34-16). The inosines in tRNA are formed by deamination of adenosine following synthesis of the primary transcript.

RIBOSOMES ARE RIBONUCLEOPROTEIN PARTICLES (70S) MADE OF A SMALL (30S) AND A LARGE (50S) SUBUNIT

We have seen how transfer RNAs, the adaptor molecules between two fundamentally different alphabets, become specifically aminoacylated and, in turn, how their anticodons are recognized by codons. We turn now to ribosomes, the molecular machines that coordinate the interplay of tRNAs, mRNA, and proteins in this complex process and catalyze peptide-bond formation. An *E. coli* ribosome is a ribonucleoprotein assembly with a mass of about 2700 kd, a diameter of approximately 200 Å, and a

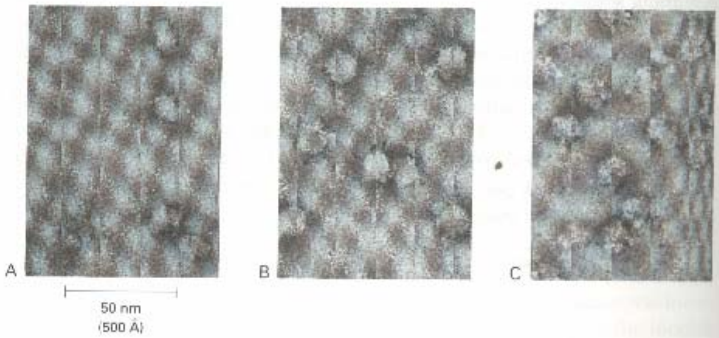


Figure 34-17
Electron micrographs of (A) 30S subunits, (B) 50S subunits, and (C) 70S ribosomes. [Courtesy of Dr. James Lake.]

sedimentation coefficient of 70S (Figure 34-17). The 20,000 ribosomes in a bacterial cell constitute nearly a fourth of its mass. A ribosome can be dissociated into a *large subunit* (*50S*) and a *small subunit* (*30S*) (Figure 34-18). These subunits can be further split into their constituent proteins

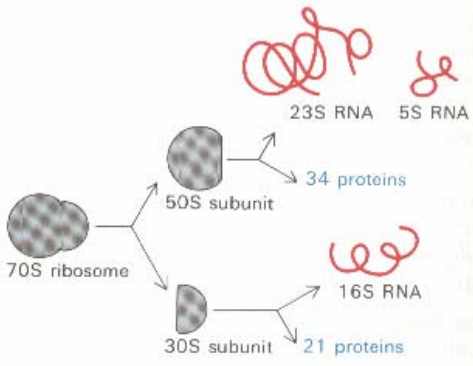


Figure 34-18
Ribosomes can be dissociated into 55 kinds of proteins and 3 RNA molecules.

and RNAs. The 30S subunit contains 21 different proteins (labeled S1 to S21) and a 16S RNA molecule. The 50S subunit contains 34 different proteins (labeled L1 to L34) and two RNA molecules, a 23S and a 5S species. A ribosome contains one copy of each RNA molecule, two of the L7 and L12 proteins, and one of each other protein. L7 is identical with L12 except that its amino terminus is acetylated. Both the 30S and 50S subunits can be reconstituted in vitro from their constituent proteins, as was first achieved by Masayasu Nomura in 1968.

RIBOSOMAL RNAs (5S, 16S, AND 23S rRNA) PLAY A CENTRAL ROLE IN PROTEIN SYNTHESIS

The prefix *ribo* in the name *ribosome* is apt, for RNA constitutes nearly two-thirds of the mass of these large molecular assemblies. The three RNAs present—5S, 16S, and 23S—are critical for ribosomal architecture and function. They are formed by cleavage of primary 30S transcripts and further processing. The base-pairing patterns of these molecules have been deduced by carrying out chemical modification and digestion experiments and by comparing nucleotide sequences of many species to detect conserved features (Figure 34-19). The striking finding is that *ribosomal RNAs (rRNAs) are folded into defined structures with many short duplex regions.*

For many years, it was presumed that ribosomal proteins orchestrate protein synthesis and that ribosomal RNAs serve primarily as a structural scaffold. The current view is almost the reverse. The discovery of catalytic RNA made us receptive to the possibility of a much more active role for RNA in ribosomal function. Indeed, several lines of evidence now suggest that *ribosomal RNAs have directive roles in protein synthesis and may indeed be dominant:*

1. Cleavage of a single bond in 16S rRNA by colicin E3, a nuclease secreted by some bacteria, abolishes protein synthesis.

2. The omission of single proteins in the in vitro reconstitution of 30S subunits leads to decreased ribosomal activity rather than total loss of function.

3. A sequence in 16S rRNA selects the start site in mRNA.

4. The wobble base of tRNA occupying the P site in a ribosome is paired with a base located in a 14-nucleotide sequence in 16S rRNA that is identical in more than a thousand molecules sequenced thus far. The identity of this sequence in archaebacteria, eubacteria, and eukaryotes indicates that it plays a critical role, one that has been conserved over several billion years of evolution.

5. The 3′ acceptor end of tRNA interacts with a conserved region of 23S rRNA.

6. Ribosomes virtually depleted of protein can still catalyze the formation of peptide bonds. 23S rRNA, by contrast, is essential for peptidyl transferase activity.

7. Most antibiotics that interfere with protein synthesis do so by interacting specifically with a ribosomal RNA rather than a ribosomal protein.

RIBOSOMAL ARCHITECTURE, INTERACTIONS, AND DYNAMICS ARE BEING MAPPED BY MULTIPLE TECHNIQUES

Elucidating the relations between ribosome structure, dynamics, and function is a formidable challenge because of the large size (megadaltons) and complexity of these assemblies. Nevertheless, investigators have succeeded in delineating the overall shape of the ribosome, its surface topography, and the location of its protein and RNA constituents. This impressive progress comes from the application of a wide range of techniques by many laboratories. The shape of the ribosome and its 30S and 50S subunits has been reconstructed from a large number of electron-microscopic images. Immunoelectron microscopy using antibodies specific for particular proteins has revealed the identity of many surface features (Figure 34-20). The mRNA binding site and the 3' end of 16S RNA are situated on a platform located between the upper and lower parts of the 30S subunit. In a cleft formed by this platform and the upper third of the subunit are the two tRNA binding sites. The 50S subunit contains three protuberances. The peptidyl transferase site that catalyzes peptide-bond formation is located in the valley between two of them; a fingerlike projection formed by a tetramer of L7 and L12 proteins contains the GTPase site that powers the movements of tRNAs and mRNA. The growing polypeptide chain emerges from the ribosome on the opposite side of the 50S subunit.

The locations of all 21 proteins of the 30S subunit have been determined by neutron diffraction analyses of concentrated solutions. Neutrons rather than x-rays were used because neutrons are scattered very differently by deuterium and hydrogen. The 30S subunit was reconstituted with two of its proteins deuterated—the subunits were obtained from bacteria grown in D_2O. Neutron scattering then revealed the distance between the centers of mass of the two deuterated proteins in the reconstituted particle. Many measurements of this kind on subunits containing different pairs of deuterated proteins led to an unequivocal map of their positions (Figure 34-21).

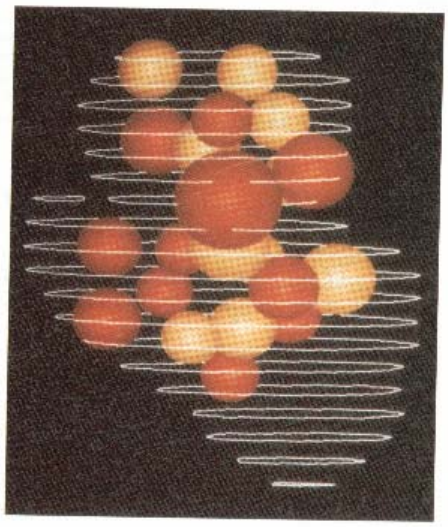


Figure 34-21
The location of all 21 proteins in the 30S ribosomal subunit has been mapped by neutron diffraction analyses of reconstituted ribosomes containing pairs of deuterated proteins. The white arcs depict the surface of the 30S particle, and the colored spheres show the positions of the proteins (the different colors serve only to help distinguish individual spheres). RNA occupies the unmarked volume of the particle. [Courtesy of Drs. M.S. Capel, D.M. Engelman, B.R. Freeborn, M. Kjeldgaard, J.A. Langer, V. Ramakrishnan, D.G. Schindler, D.K. Schneider, B.P. Schoenborn, I.-Y. Sillers, S. Yabuki, and P.B. Moore.]

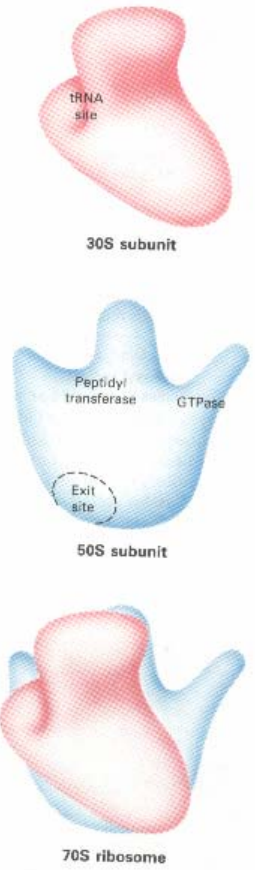
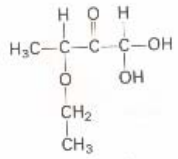
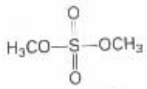


Figure 34-20
Surface topography and functional sites of the 30S subunit, 50S subunit, and intact 70S ribosome. [Courtesy of Dr. James Lake.]



Kethoxal
(Modifies G)



Dimethylsulfate
(Modifies A and C)

Hydroxyl radical
(Modifies ribose units)

Chemical modification studies of rRNAs have also been very informative. Naked RNA is vulnerable to attack by reagents such as *kethoxal, dimethylsulfate,* and *hydroxyl radical.* Nucleotides interacting with proteins and other RNA molecules are usually shielded from attack by these small chemical probes. Regions of 16S rRNA that bind each of the proteins in the 30S subunit have been mapped in this way. These *footprinting experiments* by Harry Noller have led to a model for the three-dimensional folding of 16S rRNA (Figure 34-22). Changes in footprints recorded at

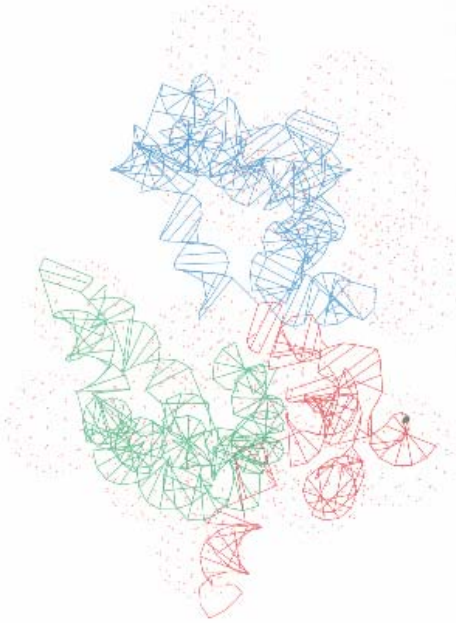


Figure 34-22
Model of the folding of 16S ribosomal RNA in the 30S subunit. The 5′, central, and 3′ domains of the RNA are shown in red, green, and blue, respectively. The dotted spheres mark the positions of individual proteins, as determined by neutron diffraction (see Figure 34-21). The proximity of particular regions of RNA to these proteins was determined by footprinting experiments. [After S. Stern, T. Powers, L.-M. Changchien, and H.F. Noller. *Science* 244(1989):783.]

different steps in peptide elongation provide revealing glimpses of the dynamics of ribosome function. *Photocross-linking experiments* too have been valuable in identifying groups that are in close proximity. For example, a pyrimidine dimer is formed between the wobble base of the anticodon of tRNA occupying the P site and the cytosine at nucleotide 1400 in 16S rRNA.

PROTEINS ARE SYNTHESIZED IN THE AMINO-TO-CARBOXYL DIRECTION

One of the first questions asked about the mechanism of protein synthesis was whether proteins are synthesized in the amino-to-carboxyl direction or the reverse direction. Pulse-labeling studies by Howard Dintzis pro-

vided a clear-cut answer. Reticulocytes (young red blood cells) that were actively synthesizing hemoglobin were exposed to ^{3}H-leucine. Completed hemoglobin was sampled frequently during a period shorter than required to synthesize a complete chain. Each sample was separated into α and β chains and then treated with trypsin to fingerprint them (p. 171). In the earliest samples, only peptides from the carboxyl ends were labeled. Later samples yielded labeled peptides closer and closer to the amino ends. Over all the samples, *a gradient of radioactivity increasing from the amino to the carboxyl end of each chain was found* (Figure 34-23). This would be expected if the amino part of the sampled chains was already synthesized prior to the addition of radioactive leucine. If the carboxyl end was synthesized last, radioactive label would appear there first, in chains that were almost complete when label was added to the medium. This experiment demonstrated that *the direction of chain growth is from the amino to the carboxyl end.*

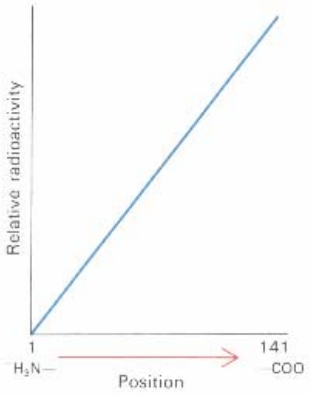
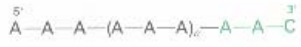


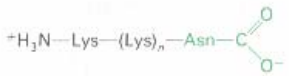
Figure 34-23
Distribution of ^{3}H-leucine in α chains of hemoglobin synthesized after exposure of reticulocytes to tritiated leucine. The higher radioactivity of the carboxyl ends relative to the amino ends indicates that the carboxyl end of each chain was synthesized last.

MESSENGER RNA IS TRANSLATED IN THE 5′ → 3′ DIRECTION

The direction of reading of mRNA was determined by using the synthetic polynucleotide



as the template in a cell-free protein-synthesizing system. AAA codes for lysine, whereas AAC codes for asparagine. The polypeptide product was



Because asparagine was the carboxyl-terminal residue, the codon AAC was the last to be read. Hence, *the direction of translation is 5′ → 3′.*

If the direction of translation were opposite to that of transcription, only fully synthesized mRNA could be translated. In contrast, if the directions were the same, mRNA could be translated while it is being synthesized. In fact, mRNA too is synthesized in the 5′ → 3′ direction (p. 845). In *E. coli,* almost no time is lost between transcription and translation. The 5′ end of mRNA interacts with ribosomes very soon after it is made, much before the 3′ end of mRNA is finished (Figure 34-24). *An important feature of prokaryotic gene expression is that translation and transcription are closely coupled in space and time.*

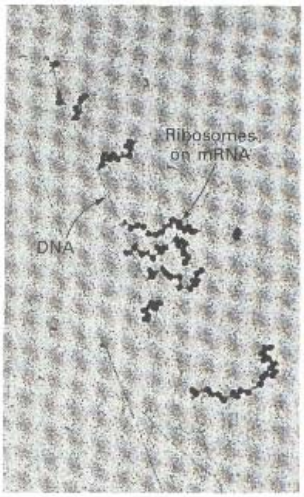


Figure 34-24
Transcription of a section of the DNA of *E. coli* and translation of the nascent mRNA. Only part of the chromosome is being transcribed. [From O.L. Miller, Jr., Barbara A. Hamkalo, and C.A. Thomas, Jr. Visualization of bacterial genes in action. *Science* 169(1970):392.]

SEVERAL RIBOSOMES SIMULTANEOUSLY TRANSLATE A MESSENGER RNA MOLECULE

Many ribosomes can simultaneously translate an mRNA molecule. This parallel synthesis markedly increases the efficiency of utilization of the mRNA. The group of ribosomes bound to an mRNA molecule is called a *polyribosome* or a *polysome.* The ribosomes in this unit operate independently, each synthesizing a complete polypeptide chain. The maximum density of ribosomes on mRNA is about 1 ribosome per 80 nucleotides. Polyribosomes synthesizing hemoglobin (which contains about 145 amino acids per chain, or 500 nucleotides per mRNA) typically consist of

five ribosomes bound to an mRNA molecule. Ribosomes closest to the 5′ end of the messenger have the shortest polypeptide chains, whereas those nearest the 3′ end have almost finished chains (Figure 34-25). Ribosomes dissociate into 30S and 50S subunits soon after the polypeptide product is released.

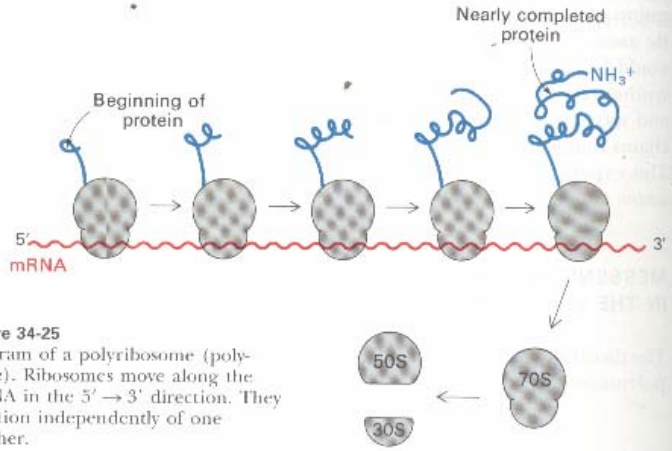


Figure 34-25
Diagram of a polyribosome (polysome). Ribosomes move along the mRNA in the 5′ → 3′ direction. They function independently of one another.

PROTEIN SYNTHESIS IN BACTERIA IS INITIATED BY FORMYLMETHIONYL TRANSFER RNA

How does protein synthesis start? The simplest possibility a priori is that the first three nucleotides of each mRNA serve as the first codon; no special start signal would then be needed. However, the experimental fact is that translation does not begin immediately at the 5′ terminus of mRNA. Indeed, the first translated codon is nearly always more than 25 nucleotides away from the 5′ end. Furthermore, many mRNA molecules in prokaryotes are *polycistronic*—that is, they code for two or more polypeptide chains. For example, a single mRNA molecule about 7000 nucleotides long specifies five enzymes in the biosynthetic pathway for tryptophan in *E. coli*. Each of these five proteins has its own start and stop signals on the mRNA. In fact, *all known mRNA molecules contain signals that define the beginning and end of each encoded polypeptide chain.*

A clue to the mechanism of initiation was the finding that nearly half the amino-terminal residues of proteins in *E. coli* are methionine, yet this amino acid is uncommon at other positions of the polypeptide chain. Furthermore, the amino terminus of nascent proteins is usually modified, which suggests that a derivative of methionine participates in initiation. In fact, *protein synthesis in bacteria starts with N-formylmethionine (fMet).* A special tRNA brings formylmethionine to the ribosome to initiate protein synthesis. This *initiator tRNA* (abbreviated as $tRNA_f$) is different from the one that inserts methionine in internal positions (abbreviated as $tRNA_m$). The subscript *f* indicates that methionine attached to the initiator tRNA can be formylated, whereas it cannot be formylated when attached to $tRNA_m$.

Methionine is linked to these two kinds of tRNAs by the same aminoacyl-tRNA synthetase (Figure 34-26). A specific enzyme then formylates

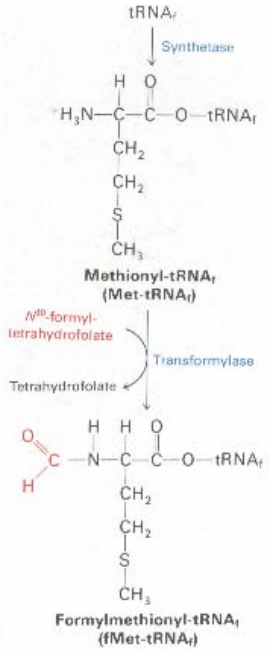


Figure 34-26
Formation of formylmethionyl-$tRNA_f$.

the amino group of methionine that is attached to $tRNA_f$. The activated formyl donor in this reaction is N^{10}-formyltetrahydrofolate (p. 720). It is significant that free methionine and methionyl-$tRNA_m$ are not substrates for this transformylase.

THE START SIGNAL IS AUG (OR GUG) PRECEDED BY SEVERAL BASES THAT PAIR WITH 16S rRNA

The initiating codon in mRNA is AUG (methionine) or, much less frequently, GUG (valine). How is an initiating AUG or GUG distinguished from one that encodes an internal residue of a protein? The first step toward answering this question was the isolation of initiator regions from a number of mRNAs. This was accomplished by digesting mRNA-ribosome complexes (formed under conditions of chain initiation but not elongation) with pancreatic ribonuclease. In each case, a sequence of about 30 nucleotides was protected from digestion. As expected, each of these initiator regions displays an AUG (or GUG) codon (Figure 34-27). In addition, each initiator region contains a purine-rich sequence centered about 10 nucleotides on the 5′ side of the initiator codon.

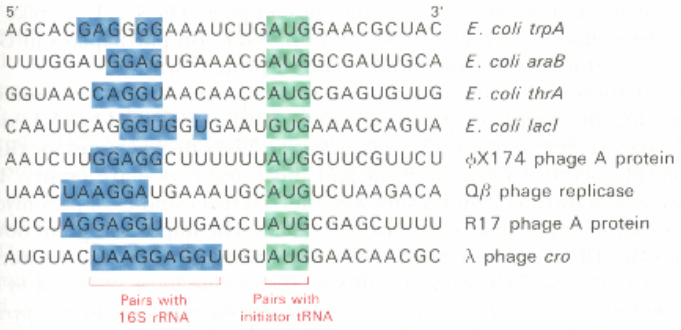


Figure 34-27
Sequences of initiation sites for protein synthesis in some bacterial and viral mRNA molecules.

The role of this purine-rich region (called the *Shine-Dalgarno sequence*) became evident when the sequence of 16S rRNA was elucidated. The 3′ end of this RNA component of the 30S subunit contains a sequence of several bases that is complementary to the purine-rich region in the initiator sites of mRNA. In fact, a complex of the 3′ end of 16S rRNA and the initiator region of mRNA has been isolated from an enzymatic digest of the initiation complex (Figure 34-28). The sequences of more than 70

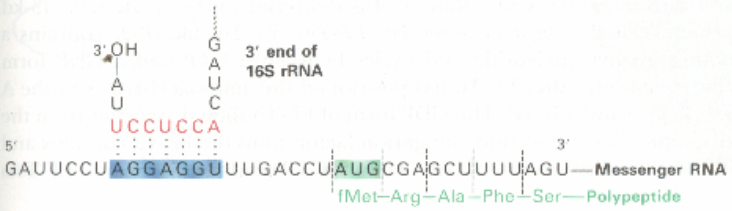


Figure 34-28
Base pairing between the purine-rich region (blue) in the initiator region of an mRNA and the 3′ end of 16S rRNA (red). The AUG codon (green) defines the start of the polypeptide chain. The mRNA shown here codes for the A protein of R17 phage. The purine-rich region is known as a Shine-Dalgarno sequence.

known initiator sites show that the number of base pairs between mRNA and 16S rRNA ranges from three to nine. Mutagenesis of the CCUCC sequence near the 3′ end of 16S rRNA to ACACA markedly interferes with the recognition of start sites in mRNA. Thus, *two kinds of interactions determine where protein synthesis starts: the pairing of mRNA bases with the 3′ end of 16S rRNA, and the pairing of the initiator codon on mRNA with the anticodon of fMet initiator tRNA.*

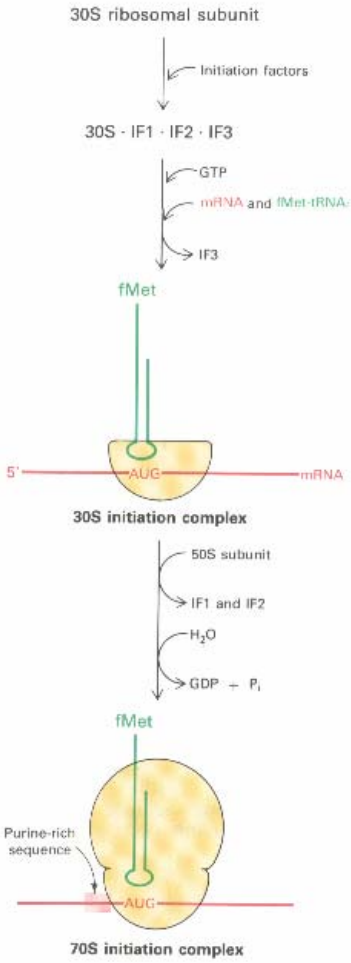


Figure 34-29
Initiation phase of protein synthesis: formation of a 30S initiation complex, followed by a 70S initiation complex.

FORMYLMETHIONYL-tRNA$_f$ IS PLACED IN THE P SITE OF THE RIBOSOME BY THE 70S INITIATION COMPLEX

How are mRNA and formylmethionyl-tRNA$_f$ brought to the ribosome to initiate protein synthesis? Three protein *initiation factors* (IF1, IF2, and IF3) are essential. The 30S ribosomal subunit first forms a complex with these three factors (Figure 34-29). The binding of GTP to IF2 enables mRNA and the initiator tRNA to join the complex as IF3 is released. Formylmethionyl-tRNA$_f$ is specifically recognized by IF2, and the release of IF3 allows the 50S subunit to join the complex. Hydrolysis of GTP bound to IF2 on entry of the 50S subunit leads to the release of IF1 and IF2. The result is a *70S initiation complex.* The L7 and nearly identical L12 proteins of the 50S subunit participate in GTP hydrolysis, which is essential for forming a productive initiation complex. The L7/L12 spike on the large subunit (see Figure 34-20 on p. 891) also participates in GTP hydrolysis during the elongation phase of protein synthesis.

Formation of the 70S initiation complex signifies that the ribosome is ready for the elongation phase of protein synthesis. *The fMet-tRNA$_f$ molecule occupies the P (peptidyl) site on the ribosome.* The other two sites for tRNA molecules—the *A (aminoacyl) site* and the *E (exit) site*—are empty. The existence of distinct P and A sites was inferred from studies of puromycin, an antibiotic to be discussed shortly (p. 902). The important point now is that fMet-tRNA$_f$ is positioned so that its anticodon pairs with the initiating AUG (or GUG) codon on mRNA. *The reading frame is defined by this interaction and by the pairing of the adjoining purine-rich sequence to a pyrimidine-rich sequence in 16S RNA* (see Figures 34-27 and 34-28). How were poly(U) and other synthetic polypeptides without start signals translated in the studies leading to the elucidation of the genetic code (p. 105)? Fortunately, nonspecific initiation occurred because the concentration of Mg^{2+} in the reaction mixture was much higher than it is in vivo.

THE GTP FORM OF ELONGATION FACTOR Tu DELIVERS AMINOACYL-tRNA TO THE A SITE OF THE RIBOSOME

The elongation cycle in protein synthesis begins with the insertion of an aminoacyl-tRNA into the empty A site on the ribosome. The particular species inserted depends on the mRNA codon that is positioned in the A site. The cognate aminoacyl-tRNA does not simply leave the synthetase and diffuse to the A site. Rather, it is delivered to the A site by a 43-kd protein called *elongation factor Tu (EF-Tu).* EF-Tu, like IF2, contains a bound guanyl nucleotide and cycles between a GTP and a GDP form (Figure 34-30). After EF-Tu has positioned the aminoacyl-tRNA in the A site, GTP is hydrolyzed. The GDP form of EF-Tu then dissociates from the ribosome. *EF-Ts,* a second elongation factor, joins the EF-Tu complex and induces the dissociation of GDP. Finally, GTP binds to EF-Tu, and EF-Ts is

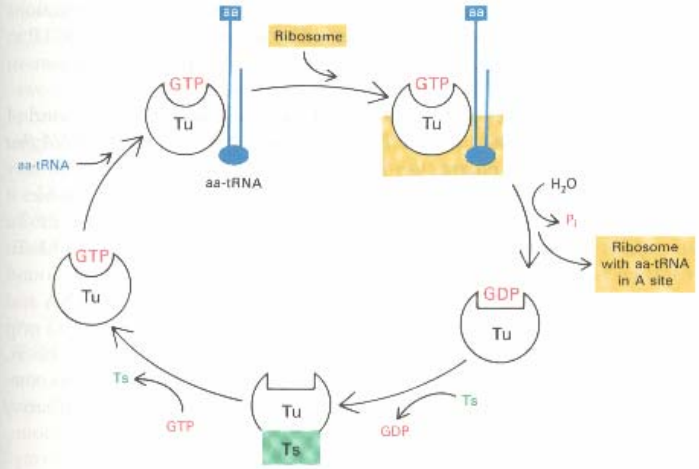


Figure 34-30
Reaction cycle of elongation factor Tu (EF-Tu).

concomitantly released. EF-Tu containing bound GTP is now ready to pick up another aminoacyl-tRNA and deliver it to the A site of the ribosome.

It is noteworthy that *EF-Tu does not interact with fMet-tRNA$_f$*. Hence, this initiator tRNA is not delivered to the A site. In contrast, Met-tRNA$_m$, like all other aminoacyl-tRNAs, does bind to EF-Tu. These findings account for the fact that *internal AUG codons are not read by the initiator tRNA*. Conversely, initiation factor 2 recognizes fMet-tRNA$_f$ but no other tRNA.

This GTP-GDP cycle of EF-Tu is reminiscent of that of transducin in vision (p. 336), the stimulatory G protein in hormone action (p. 342), and the ras protein in growth control (p. 355). Indeed, the amino-terminal domain (the G-domain) of EF-Tu is structurally similar to the GTP-binding subunit of these signal-transducing proteins. The other two do-

Figure 34-31
Structure of the GTP form of EF-Tu, a tripartite protein. The GTPase domain is shown in red, and the other two domains in blue and green. The bound hydrolysis-resistant analog of GTP is shown in yellow. The location of the binding site for aminoacyl-tRNA was inferred from labeling and mutagenesis studies. (A) Schematic diagram. (B) Space-filling model. [Drawn from left.pdb. M. Kjeldgaard, P. Nissen, S. Thirup, and J. Nyborg. *Structure* 1(1993):35.]