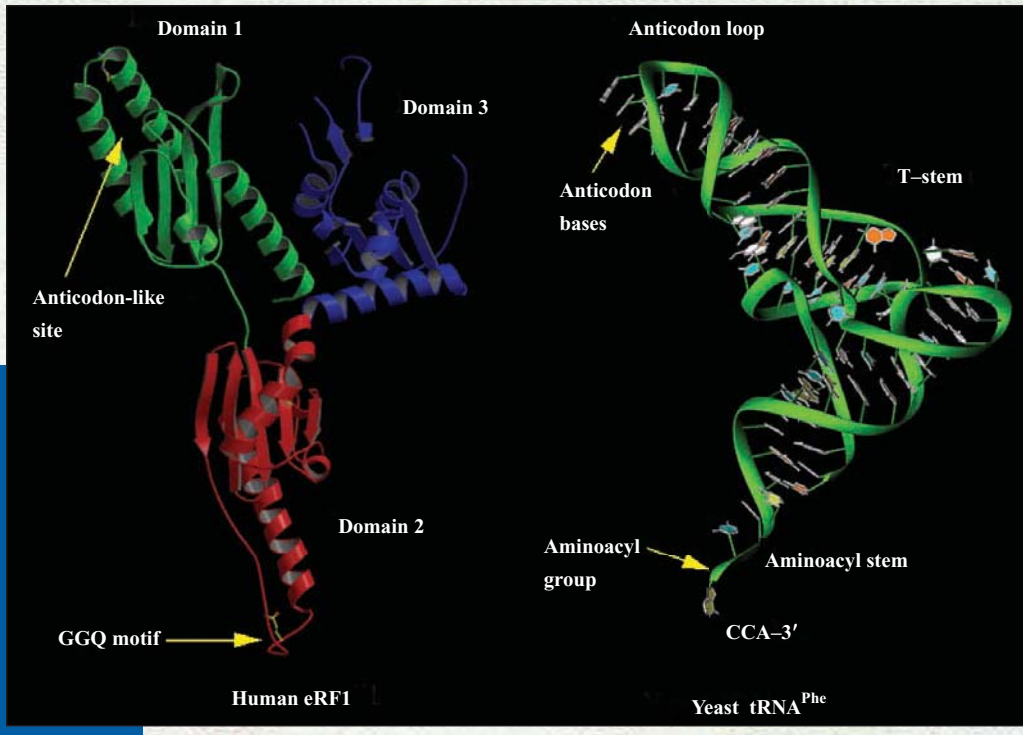
CHAPTER 18

The Mechanism of Translation II: Elongation and Termination



Comparison of crystal structures of human eRF1 and yeast $tRNA^{Phe}$.

The elongation processes in bacteria and eukaryotes are very similar. Accordingly, we will consider the processes together, discussing the bacterial system first, then noting some differences in the eukaryotic system.

As we learned in Chapter 17, the initiation process in bacteria creates a ribosome primed with an mRNA and the initiating aminoacyl-tRNA, fMet-$tRNA_f^{Met}$, ready to begin elongating a polypeptide chain. Before we look at the steps involved in this elongation process, let us consider some fundamental questions about the nature of elongation: (1) In what direction is a polypeptide synthesized? (2) In what direction does the ribosome read the mRNA? (3) What is the nature of the genetic code that dictates which amino acids will be incorporated in response to the mRNA?

18.1 The Direction of Polypeptide Synthesis and of mRNA Translation

Proteins are made one amino acid at a time, but where does synthesis begin? Do protein chains grow in the amino-to-carboxyl direction, or the reverse? In other words, which amino acid is inserted first into a growing polypeptide—the amino-terminal amino acid, or the carboxyl-terminal one? Howard Dintzis provided definitive proof of the amino → carboxyl direction in 1961 with a study of α- and β-globin synthesis in isolated rabbit reticulocytes (immature red blood cells). He labeled the growing globin chains for various short lengths of time with [^{3}H]leucine, and for a long time with [^{14}C]leucine. Then he separated the α- and β-globins, cut them into peptides with trypsin, and separated the peptides. He then plotted the relative amounts of [^{3}H]leucine incorporated into the peptides versus the positions of the peptides, from N-terminus to C-terminus, in the proteins. The long labeling with [^{14}C]leucine should have labeled all peptides equally, so it could be used as an internal control for losses of certain peptides during purification, and for differences in leucine content from one peptide to another.

Figure 18.1 shows how this procedure can tell us the direction of translation. It is important to notice that the protein chains are in all stages of completion when the ^{3}H- labeled amino acid is added. Thus, some are just starting, some are partly finished, and some are almost finished. This means that label will be incorporated into the first peptide only in those proteins whose synthesis had just begun when the label was added. The others will be labeled in downstream peptides, but not in the first one. By contrast, the end of the protein where protein synthesis ends will be

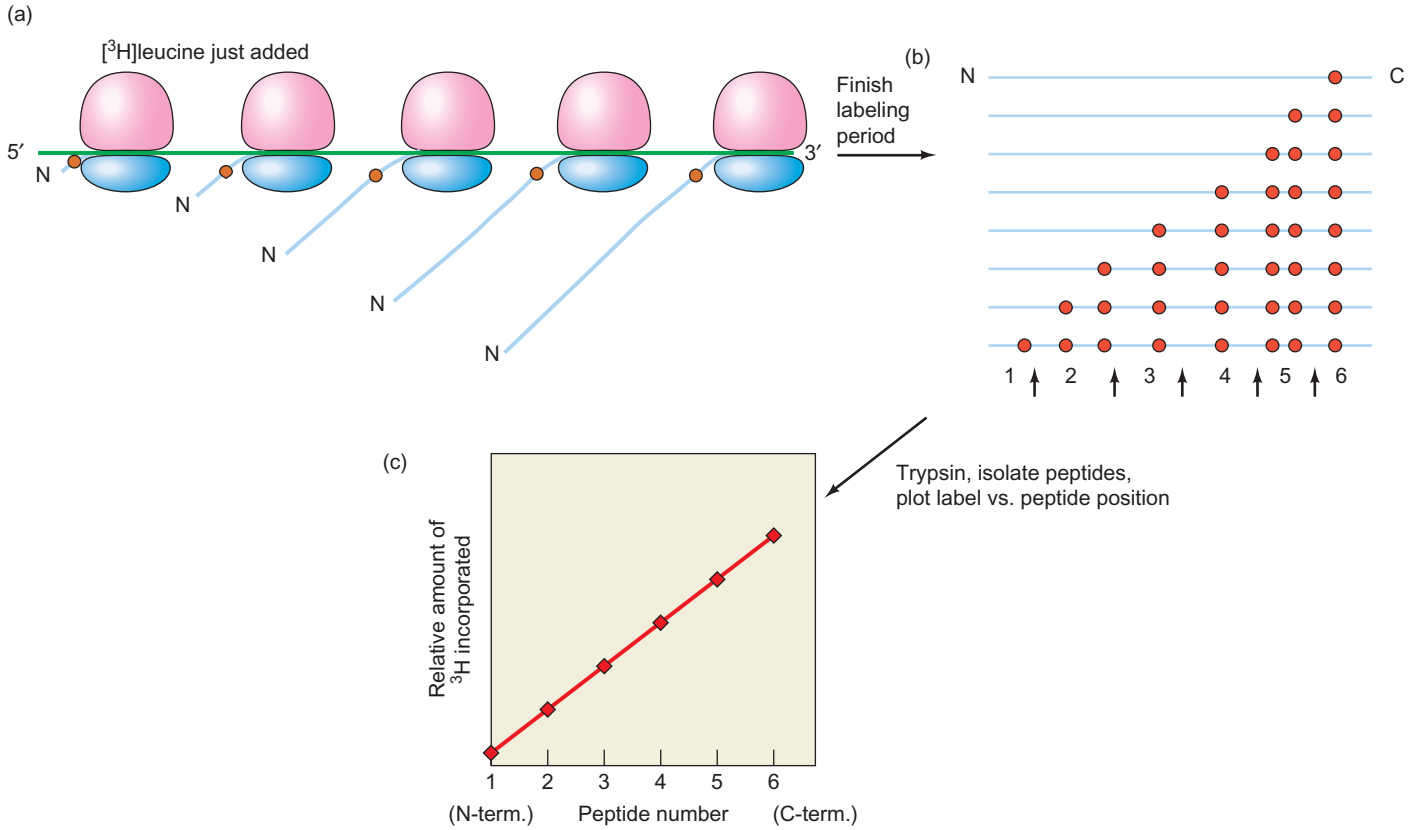


Figure 18.1 Experimental strategy to determine the direction of translation. (a) Labeling the protein. Consider an mRNA (green) being translated by several ribosomes (pink and blue), assuming that the mRNA is translated in the 5′→3′ direction and the proteins are made in the amino (N) to carboxyl (C) direction. A labeled amino acid ([^{3}H]leucine) has just been added to the system, so it has begun to be incorporated into the growing protein chains (blue), as indicated by the red dots. It is incorporated near the N-terminus in the polypeptides on the left, where protein synthesis has just begun, but only near the C-terminus in the polypeptides on the right, which are almost completed. **(b)** Distribution of label in completed proteins after a moderate labeling period. The proteins near the top, with label only near the C-terminus correspond to the nearly completed proteins near the right in panel (a). Those near the bottom, with label distributed toward the N-terminus, correspond to the growing proteins near the left in panel (a). These have had time to incorporate label throughout a greater length of the protein. Cutting sites for trypsin within the protein are indicated by arrows at bottom, and the resulting peptides are numbered 1–6 according to their positions in the protein. **(c)** Model experimental results. One plots the relative amount of ^{3}H labeling in each of the peptides, 1–6, and finds that the C-terminal peptides are the most highly labeled. This is what we expect if translation started at the N-terminus. If it had started at the C-terminus (opposite to the picture in panel [a]), then the N-terminal peptides would be the most highly labeled.

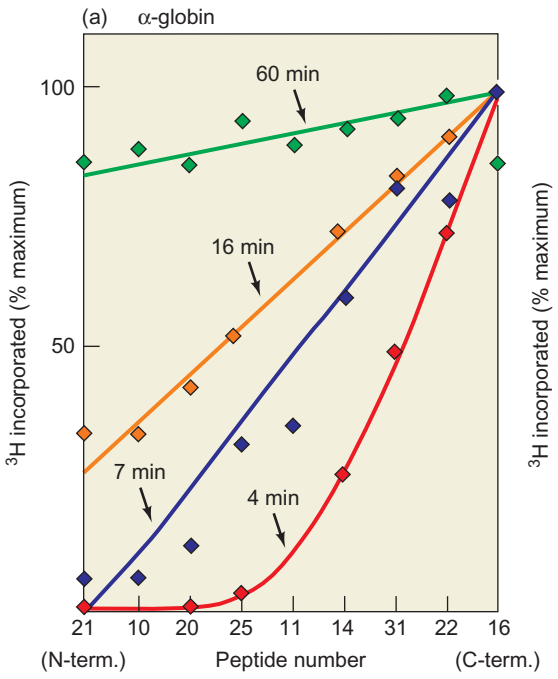


Figure 18.2 Determining the direction of translation. Dintzis carried out the experimental plan outlined in Figure 18.1 with rabbit reticulocytes, which make almost nothing but α- and β-globins. He labeled the reticulocytes with [^{3}H]leucine for various lengths of time, then separated the α- and β-globins, cut each protein into peptides with trypsin, and determined the label in each peptide. He plotted the relative amount of ^{3}H label against the peptide number, with the N-terminal peptide on the left, and the C-terminal peptide on the right. The curves for α- and β-globin showed the most label in the C-terminal peptides, especially after short labeling times. (Only the α-globin results are shown here.) This is what we expect if translation starts at the N-terminus of a protein. Note that the peptide numbers are not related to their position in the protein, as they are in the example in Figure 18.1. (*Source:* Adapted from Dintzis, H.M., Assembly of the peptide chains of hemoglobin. *Proceedings of the National Academy of Sciences USA* 47:255, 1961.)

relatively rich in label after a short labeling time. Intermediate peptides will show intermediate levels of labeling. Thus, if translation starts at the amino terminus, labeling will be strongest in carboxyl-terminal peptides. Figure 18.2 shows the results. Labeling of the peptides of both α- and β-globins increased from the amino terminus to the carboxyl terminus, and this disparity was especially noticeable with short labeling times. Therefore, protein synthesis starts at the amino terminus of the protein.

Is the mRNA read in the 5′→3′ direction or the reverse? Knowing that proteins grow in the amino→carboxyl direction, it is easy to show that mRNAs are read in the 5′→3′ direction. When molecular biologists first started using synthetic mRNAs as templates for protein synthesis in the 1960s, some of these messages held the answer to our question. For example, when Ochoa and his colleagues translated the mRNA: 5′-AUGUUU$_n$-3′, they obtained fMet-Phe$_n$, where the fMet was at the amino terminus. We know that AUG codes for fMet and UUU codes for phenylalanine (Phe). We see that fMet is incorporated into the amino terminal position of the protein, which means it was added first, before any of the phenylalanines. Therefore the mRNA must have been read from the 5′-end, because that is where the fMet codon is.

SUMMARY Messenger RNAs are read in the 5′→3′ direction, the same direction in which they are synthesized. Proteins are made in the amino→carboxyl direction, which means that the amino terminal amino acid is added first.

18.2 The Genetic Code

The term **genetic code** refers to the set of three-base code words (**codons**) in mRNAs that stand for the 20 amino acids in proteins. Like any code, this one had to be broken before we knew what the codons stood for. Indeed, before 1960, other more basic questions about the code were still unanswered. These included: Do the codons overlap? Are there gaps, or "commas," in the code? How many bases make up a codon? These questions were answered in the 1960s by a series of imaginative experiments, which we will examine here.

Nonoverlapping Codons

In a nonoverlapping code, each base is part of at most one codon. In an overlapping code, one base may be part of two or even three codons. Consider the following micromessage:

AUGUUC

Assuming that the code is triplet (three bases per codon) and this message is read from the beginning, the codons will be AUG and UUC if the code is nonoverlapping. On the other hand, an overlapping code might yield four codons: AUG, UGU, GUU, and UUC. As early as 1957, Sydney Brenner concluded on theoretical grounds that a fully overlapping triplet code like this would be impossible.

However, given the data available in 1957, a *partially* overlapping code remained possible, but A. Tsugita and H. Frankel-Conrat laid it to rest with the following line of reasoning: If the code is nonoverlapping, a change of one base in an mRNA (a missense mutation) would change no more than one amino acid in the resulting protein. For example, consider another micromessage:

AUGCUA

Assuming that the code is triplet (three bases per codon) and this message is read from the beginning, the codons

will be AUG and CUA if the code is nonoverlapping. A change in the fourth base (C) would change only one codon (CUA) and therefore at most only one amino acid. On the other hand, if the code were overlapping, base C could be part of three adjacent codons (UGC, GCU, and CUA). Therefore, if the C were changed, up to three adjacent amino acids could be changed in the resulting protein. But when the investigators introduced one-base alterations into mRNA from tobacco mosaic virus (TMV), they found that these never caused changes in more than one amino acid. Hence, the code must be nonoverlapping.

No Gaps in the Code

If the code contained untranslated gaps, or "commas," mutations that add or subtract a base from a message might change a few codons, but we would expect the ribosome to get back on track after the next comma. In other words, these mutations might frequently be lethal, but in many cases the mutation should occur just before a comma in the message and therefore have little, if any, effect. If no commas were present to get the ribosome back on track, these mutations would be lethal except when they occur right at the end of a message.

Such mutations do occur, and they are called **frameshift mutations;** they work as follows. Consider another tiny message:

AUGCAGCCAACG

If translation starts at the beginning, the codons will be AUG, CAG, CCA, and ACG. If we insert an extra base (X) right after base U, we get:

AUXGCAGCCAACG

Now this would be translated from the beginning as AUX, GCA, GCC, AAC. Notice that the extra base changes not only the codon (AUX) in which it appears, but every codon from that point on. The **reading frame** has shifted one base to the left; whereas C was originally the first base of the second codon, G is now in that position.

On the other hand, a code with commas would be one in which each codon is flanked by one or more untranslated bases, represented by Z's in the following message. The commas would serve to set off each codon so the ribosome could recognize it:

AUGZCAGZCCAZACGZ

Deletion or insertion of a base anywhere in this message would change only a single codon. The comma (Z) at the end of the damaged codon would then put the ribosome back on the right track. Thus, addition of an extra base (X) to the first codon would give the message:

AUXGZCAGZCCAZACGZ

The first codon (AUXG) is now wrong, but all the others, still neatly set off by Z's, would be translated normally.

When Francis Crick and his colleagues treated bacteria with acridine dyes that usually cause single-base insertions or deletions, they found that such mutations were very severe; the mutant genes gave no functional product. This is what we would expect of a "comma-less" code with no gaps; base insertions or deletions cause a shift in the reading frame of the message that persists until the end of the message.

Moreover, Crick found that adding a base could cancel the effect of deleting a base, and vice versa. This phenomenon is illustrated in Figure 18.3, where we start with an artificial gene composed of the same codon, CAT, repeated over and over. When we add a base, G, in the third position, we change the reading frame so that all codons thereafter read TCA. When we start with the wild-type gene and delete the fifth base, A, we change the reading frame in the other direction, so that all subsequent codons read ATC. Crossing these two mutants sometimes gives a recombined "pseudo-wild-type" gene like the one on line 4 of the figure. Its first two codons, CAG and TCT, are wrong, but thereafter the insertion and deletion cancel, and the original reading frame is restored. All codons from that point on read CAT.

The Triplet Code

Francis Crick and Leslie Barnett discovered that a presumed set of three insertions or deletions could produce a

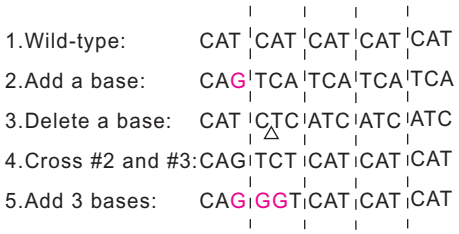


Figure 18.3 Frameshift mutations. Line 1: An imaginary gene has the same codon, CAT, repeated over and over. The vertical dashed lines show the reading frame, starting from the beginning. Line 2: Adding a base, G (pink), in the third position changes the first codon to CAG and shifts the reading frame one base to the left so that every subsequent codon reads TCA. Line 3: Deleting the fifth base, A (marked by the triangle), from the wild-type gene changes the second codon to CTC and shifts the reading frame one base to the right so that every subsequent codon reads ATC. Line 4: Crossing the mutants in lines 2 and 3 occasionally gives a recombined "pseudo-wild-type" revertant with an insertion and a deletion close together. The end result is a DNA with its first two codons altered, but all the other ones put back into the correct reading frame. Line 5: Adding three bases, GGG (pink), after the first two bases disrupts the first two codons, but leaves the reading frame unchanged. The same would be true of deleting three bases.

pseudo-wild-type gene (Figure 18.3, line 5). This of course demands that a codon consist of three bases. As Crick remarked to Barnett when he saw the experimental result, "We're the only two [who] know it's a triplet code!" Actually, Crick and Bartlett were inferring that their pseudo-wild-type genes contained three insertions or deletions. They had no way of sequencing the genes to make sure, so more experiments were needed.

In 1961, Marshall Nirenberg and Johann Heinrich Matthaei performed a groundbreaking experiment that laid the foundation for confirming the triplet nature of the code and for breaking the genetic code itself. The experiment was deceptively simple; it showed that synthetic RNA could be translated in vitro. In particular, when Nirenberg and Matthaei translated poly(U), a synthetic RNA composed only of U's, they made polyphenylalanine. Of course, that told them that a codon for phenylalanine contains only U's. This finding by itself was important, but the long-range implication was that one could design synthetic mRNAs of defined sequence and analyze the protein products to shed light on the nature of the code. Gobind Khorana and his colleagues were the chief practitioners of this strategy.

Here is how Khorana's synthetic messenger experiments confirmed that the codons contain three bases: First, if the codons contain an odd number of bases, then a repeating dinucleotide poly(UC) or UCUCUCUC . . . should contain two alternating codons (UCU and CUC, in this case), no matter where translation starts. The resulting protein would be a repeating dipeptide—two amino acids alternating with each other. If codons have an even number of bases, only one codon (UCUC, for example) should be repeated over and over. Of course, if translation started at the second base, the single repeated codon would be different (CUCU). In either case, the resulting protein would be a homopolypeptide, containing only one amino acid repeated over and over. Khorana found that poly(UC) translated to a repeating dipeptide, poly(serine-leucine) (Figure 18.4a), proving that the codons contained an odd number of bases.

Repeating triplets were translated to homopolypeptides, as had been expected if the number of bases in a codon was three or a multiple of three. For example, poly(UUC) translated to polyphenylalanine plus polyserine plus polyleucine (Figure 18.4b). The reason for three different products is that translation can start at any point in the synthetic message. Therefore, poly(UUC) can be read as UUC, UUC, and so on, UCU, UCU, and so on, or CUU, CUU, and so on, depending on where translation starts. In all cases, once translation begins, only one codon is encountered, as long as the number of bases in a codon is divisible by 3.

Repeating tetranucleotides were translated to repeating tetrapeptides. For example, poly(UAUC) yielded poly(tyrosine-leucine-serine-isoleucine) (Figure 18.4c). As an exercise, you can write out the sequence of such a message

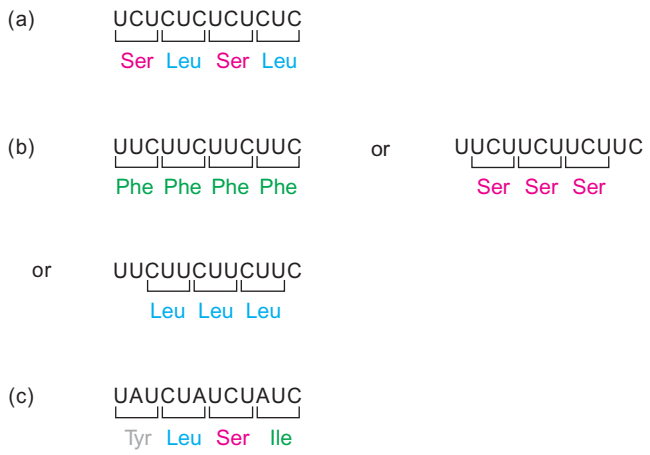


Figure 18.4 Coding properties of several synthetic mRNAs. **(a)** Poly(UC) contains two alternating codons, UCU and CUC, which code for serine (Ser) and leucine (Leu), respectively. Thus, the product is poly(Ser-Leu). **(b)** Poly(UUC) contains three codons, UUC, UCU, and CUU, which code for phenylalanine (Phe), serine (Ser), and leucine (Leu), respectively. The product is therefore poly(Phe), or poly(Ser), or poly(Leu), depending on which of the three reading frames the ribosome uses. **(c)** Poly(UAUC) contains four codons in a repeating sequence: UAU, CUA, UCU, and AUC, which code for tyrosine (Tyr), leucine (Leu), serine (Ser), and isoleucine (Ile), respectively. The product is therefore poly(Tyr-Leu-Ser-Ile).

and satisfy yourself that it is compatible with codons having three bases, or nine, or even more, but not six. (We already know six cannot be right because it is not an odd number.) Because codons are not likely to be as cumbersome as nine bases long, three is the best choice. Look at the problem another way: Three is the lowest number that gives enough different codons to specify all 20 amino acids. (The number of permutations of four different bases taken 3 at a time is 4^3, or 64.) There would be only 16 two-base codons ($4^2 = 16$), not quite enough. But there would be over 200,000 ($4^9 = 262{,}144$) nine-base codons. Nature is usually more economical than that.

SUMMARY The genetic code is a set of three-base code words, or codons, in mRNA that instruct the ribosome to incorporate specific amino acids into a polypeptide. The code is nonoverlapping: that is, each base is part of only one codon. It is also devoid of gaps, or commas; that is, each base in the coding region of an mRNA is part of a codon.

Breaking the Code

Obviously, Khorana's synthetic mRNAs gave strong hints about some of the codons. For example, because poly(UC) yields poly(serine-leucine), we know that one of the codons (UCU or CUC) codes for serine and the other codes

for leucine. The question remains: Which is which? Nirenberg developed a powerful assay to answer this question. He found that a trinucleotide was usually enough like an mRNA to cause a specific aminoacyl-tRNA to bind to ribosomes. For example, the triplet UUU will cause phenylalanyl-tRNA to bind, but not lysyl-tRNA or any other aminoacyl-tRNA. Therefore, UUU is a codon for phenylalanine. This method was not perfect; some codons did not cause any aminoacyl-tRNA to bind, even though they were authentic codons for amino acids. But it provided a nice complement to Khorana's method, which by itself would not have given all the answers either, at least not easily.

Here is an example of how the two methods could be used together: Translation of the polynucleotide poly(AAG) yielded polylysine plus polyglutamate plus polyarginine. There are three different codons in that synthetic message: AAG, AGA, and GAA. Which one codes for lysine? All three were tested by Nirenberg's assay, yielding the results shown in Figure 18.5. Clearly, AGA and GAA caused no binding of [^{14}C]lysyl-tRNA to ribosomes, but AAG did. Therefore, AAG is the lysine codon in poly(AAG). Something else to notice about this experiment is that the triplet AAA also caused lysyl-tRNA to bind. Therefore, AAA is another lysine codon. This illustrates a general feature of the code: In most cases, more than one triplet codes for a given amino acid. In other words, the code is **degenerate.**

Figure 18.6 shows the entire genetic code. As predicted, there are 64 different codons and only 20 different amino acids, yet all of the codons are used. Three are "stop" codons found at the ends of messages, but all the others specify amino acids, which means that the code is highly degenerate. Leucine, serine, and arginine have six different codons; several others, including proline, threonine, and alanine, have four; isoleucine has three; and many others have two. Just two amino acids, methionine and tryptophan, have only one codon.

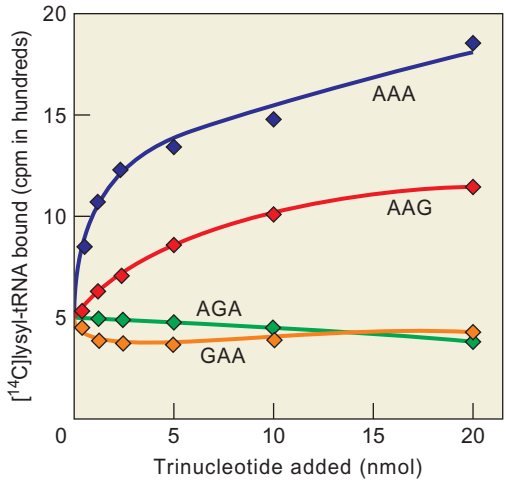


Figure 18.5 Binding of lysyl-tRNA to ribosomes in response to various codons. Lysyl-tRNA was labeled with radioactive carbon (^{14}C) and mixed with *E. coli* ribosomes in the presence of the following trinucleotides: AAA, AAG, AGA, and GAA. Lysyl-tRNA-ribosome complex formation was measured by binding to nitrocellulose filters. (Unbound lysyl-tRNA does not stick to these filters, but a lysyl-tRNA–ribosome complex does.) AAA was a known lysine codon, so binding was expected with this trinucleotide. (*Source:* Adapted from Khorana, H.G., Synthesis in the study of nucleic acids, *Biochemical Journal* 109:715, 1968.)

SUMMARY The genetic code was broken by using either synthetic messengers or synthetic trinucleotides and observing the polypeptides synthesized or aminoacyl-tRNAs bound to ribosomes, respectively. There are 64 codons in all. Three are stop signals, and the rest code for amino acids. This means that the code is highly degenerate.

Second position

First position (5′-end)	U	C	A	G	Third position (3′-end)
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA STOP	UGA STOP	A
	UUG Leu	UCG Ser	UAG STOP	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
	GUG Val	GCG Ala	GAG Glu	GGG Gly	G

Figure 18.6 The genetic code. All 64 codons are listed, along with the amino acid for which each codes. To find a given codon—ACU, for example—we start with the wide horizontal row labeled with the name of the first base of the codon (A) on the left border. Then we move across to the vertical column corresponding to the second base (C). This brings us to a box containing all four codons beginning with AC. It is now a simple matter to find the one among these four we are seeking, ACU. We see that this triplet codes for threonine (Thr), as do all the other codons in the box: ACC, ACA, and ACG. This is an example of the degeneracy of the code. Notice that three codons (pink) do not code for amino acids; instead, they are stop signals.

Unusual Base Pairs Between Codon and Anticodon

How does an organism cope with multiple codons for the same amino acid? One way would be to have multiple tRNAs (**isoaccepting species**) for the same amino acid, each one specific for a different codon. This is part of the answer, and indeed a given organism contains about 60 different tRNAs. But, in principle, we can get along with considerably fewer tRNAs than that simple hypothesis would predict. Again Francis Crick anticipated experimental results with insightful theory. In this case, Crick hypothesized that the first two bases of a codon must pair correctly with the anticodon according to Watson–Crick base-pairing rules (Figure 18.7a), but the last base of the codon can "wobble" from its normal position to form unusual base pairs with the anticodon. This proposal was called the **wobble hypothesis.** In particular, Crick proposed that a G in an anticodon can pair not only with a C in the third position of a codon (the **wobble position**), but also with a U. This would give the **wobble base pair** shown in Figure 18.7b. Notice how the U has moved, or wobbled from its normal position to form this base pair.

Furthermore, Crick noted that one of the unusual nucleosides found in tRNA is **inosine (I),** which has a structure similar to that of guanosine. This nucleoside can ordinarily pair like G, so we would expect it to pair with C (Watson–Crick base pair) or U (wobble base pair) in the third position (the wobble position) of a codon. But Crick proposed that inosine could form still another kind of wobble pair, this time with A in the third position of a codon (Figure 18.7c). That means an anticodon with I in the first position can potentially pair with three different codons ending with C, U, or A.

The wobble phenomenon reduces the number of tRNAs required to translate the genetic code. For example, consider the two codons for phenylalanine, UUU and UUC, listed at the top left of Figure 18.6. According to the wobble hypothesis, they can both be recognized by an anticodon that reads 3′-AAG-5′ (Figure 18.8a). The G in the 5′-position of the anticodon could form a Watson–Crick G–C base pair with the C in the UUC, or a G–U wobble base pair with the U in UUU. Similarly, the two leucine codons in the same box, UUA and UUG, can both be recognized by the anticodon 3′-AAU-5′ (Figure 18.8b). The U can form a Watson–Crick pair with the A in UUA, or a wobble pair with the G in UUG.

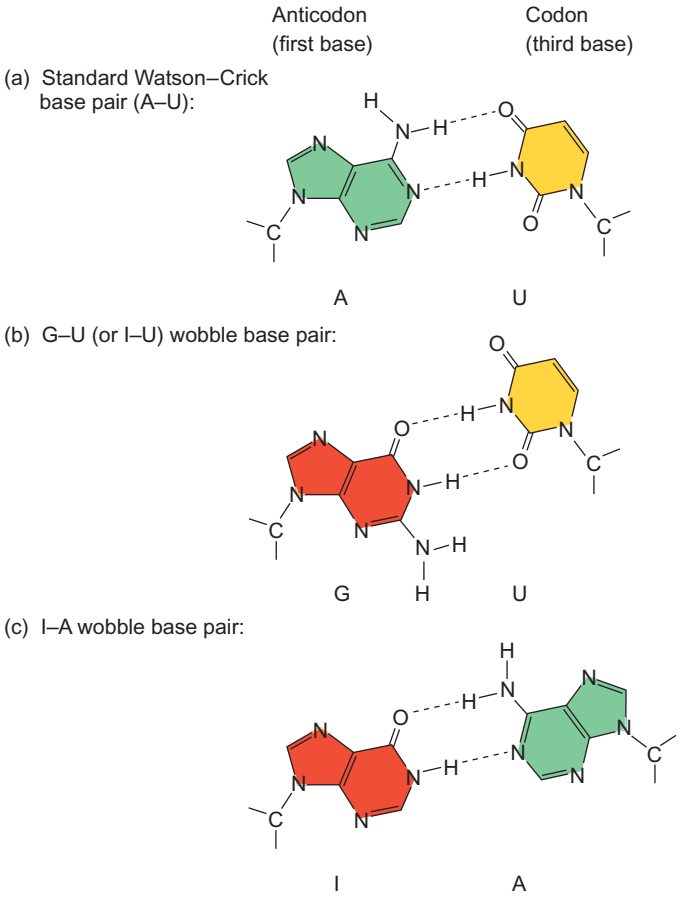


Figure 18.7 Wobble base pairs. (a) Relative positions of bases in a standard (A–U) base pair. The base on the left here and in the wobble base pairs (b) and (c) is the first base in the anticodon. The base on the right is the third base in the codon. **(b)** Relative positions of bases in a G–U (or I–U) wobble base pair. Notice that U has to "wobble" upward to pair with the G (or I). **(c)** Relative positions of bases in an I–A wobble base pair. The A has to "wobble" upward in order to form this pair.

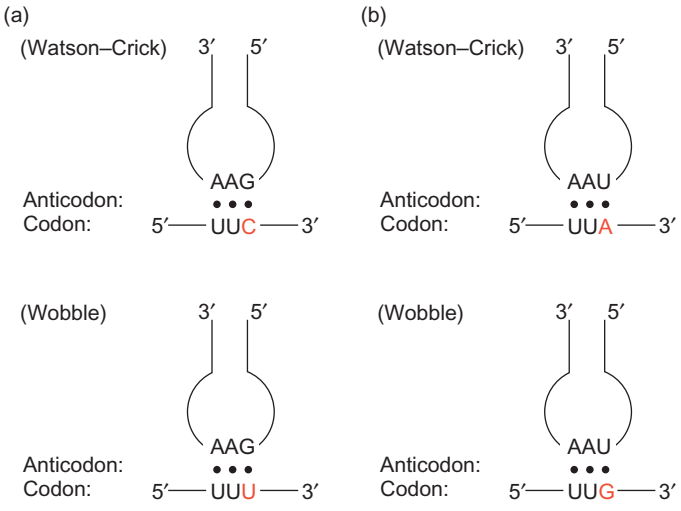


Figure 18.8 The wobble position. (a) An abbreviated tRNA with anticodon 3′-AAG-5′ is shown base-pairing with two different codons for phenylalanine: UUC and UUU. The wobble position (the third base of the codon) is highlighted in red. The base-pairing with the UUC codon (top) uses only Watson–Crick pairs; the base-pairing with the UUU codon (bottom) uses two Watson–Crick pairs in the first two positions of the codon, but requires a wobble pair (G–U) in the wobble position. **(b)** A similar situation, in which a tRNA with anticodon AAU base-pairs with two different codons for leucine: UUA and UUG. Pairing with the UUG codon requires a G–U wobble pair in the wobble position.

According to the wobble hypothesis, a cell should be able to get by with only 31 tRNAs to read all 64 codons, assuming no tRNA is needed to read the UAA and UAG stop codons. But human mitochondria and plant plastids contain fewer than 31 tRNAs, so something besides wobble appears to be in play. This has led to the **superwobble** hypothesis, which holds that a single tRNA with a U in its wobble position (the first base in its anticodon) can, at least in certain circumstances, recognize codons ending in any of the four bases.

Ralph Bock and colleagues put the superwobble hypothesis to the test in 2008 when they knocked out both $tRNA^{Gly}$ genes in tobacco plastids, then added back only $tRNA^{Gly}$(UCC), which, using superwobble, should be able to translate all four glycine codons. The resulting tobacco cells were indeed viable, though translation efficiency was reduced. Thus, superwobble appears to work, but not perfectly, which probably explains why it has not evolved very often.

SUMMARY Part of the degeneracy of the genetic code is accommodated by isoaccepting species of tRNA that bind the same amino acid but recognize different codons. The rest is handled by wobble, in which the third base of a codon is allowed to move slightly from its normal position to form a non-Watson–Crick base pair with the anticodon. This allows the same aminoacyl-tRNA to pair with more than one codon. The wobble pairs are G–U (or I–U) and I–A. Some organelles have evolved with fewer tRNAs than are required to translate all the sense codons. In these cases, codons with U in the wobble position can apparently translate codons with all four bases in the last position by superwobble.

The (Almost) Universal Code

In the years after the genetic code was broken, all organisms examined, from bacteria to humans, were shown to share the same code. Therefore it was generally assumed (incorrectly, as we will see) that the code was universal, with no deviations whatsoever. This apparent universality led in turn to the notion of a single origin of present life on earth.

The reasoning for this idea goes like this: Nothing is inherently advantageous about each specific codon assignment we see. There is no obvious reason, for example, why UUC should make a good codon for phenylalanine, whereas AAG is a good one for lysine. Rather, the genetic code may be an "accident"; it just happened to evolve that way. However, once these codons were established, there was a very good reason why they did not change: A change that fundamental would almost certainly be lethal.

Consider, for instance, a tRNA for the amino acid cysteine and the codon it recognizes, UGU. For that relationship to change, the anticodon of the cysteinyl-tRNA would have to change so it can recognize a different codon, say UCU, which is a serine codon. At the same time, all the UCU codons in that organism's genome that code for important serines would have to change to alternate serine codons so they would not be recognized as cysteine codons. The chances of all these things happening together, even over vast evolutionary time, are negligible. That is why the genetic code is sometimes called a "frozen accident"; once it was established, for whatever reasons, it had to stay that way. So a universal code would be powerful evidence for a single origin of life. After all, if life started independently in two places, we would hardly expect the two lines to evolve the same genetic code by accident!

In light of all this, it is remarkable that the genetic code is not absolutely universal; there are some exceptions to the rule. The first of these to be discovered were in the genomes of mitochondria. In mitochondria of the fruit fly *D. melanogaster,* UGA is a codon for tryptophan rather than for "stop." Even more remarkably, AGA in these mitochondria codes for serine, whereas it is an arginine codon in the standard code. Mammalian mitochondria show some deviations, too. Both AGA and AGG, though they are arginine codons in the standard code, have a different meaning in human and bovine mitochondria; there they code for "stop." Furthermore, AUA, ordinarily an isoleucine codon, codes for methionine in these mitochondria.

These aberrations might be dismissed as relatively unimportant, occurring as they do in mitochondria, which have very small genomes coding for only a few proteins and therefore more latitude to change than nuclear genomes. But exceptional codons also occur in nuclear genomes and bacterial genomes. In at least three ciliated protozoa, including *Paramecium,* UAA and UAG, which are normally stop codons, code for glutamine. In the prokaryote *Mycoplasma capricolum,* UGA, normally a stop codon, codes for tryptophan. In the pathogenic yeast, *Candida albicans,* CTG, usually a leucine codon, codes for serine. Deviations from the standard genetic code are summarized in Table 18.1.

Clearly, the so-called universal code is not really universal. Does this mean that the evidence now favors more than one origin of present life on earth? If the deviant codes were radically different from the standard one, this might be an attractive possibility, but they are not. In many cases, the novel codons are stop codons that have been recruited to code for an amino acid: glutamine or tryptophan. There is a well-established mechanism for this sort of occurrence, as we will see later in this chapter. The vast majority of known examples of codons that have switched their meaning from one amino acid to another occur in mitochondria. Again, mitochondrial genomes, because they code for far fewer proteins than nuclear genomes or even bacterial genomes, might be expected to change a codon safely every now and then. In summary, even if the code is not universal, a standard code does exist from which the deviant ones almost certainly evolved. Therefore, the evidence still strongly favors a single origin of life.

Table 18.1 Deviations from the "Universal" Genetic Code

Source	Codon	Usual meaning	New meaning
Fruit fly mitochondria	UGA	Stop	Tryptophan
	AGA & AGG	Arginine	Serine
	AUA	Isoleucine	Methionine
Mammalian mitochondria	AGA & AGG	Arginine	Stop
	AUA	Isoleucine	Methionine
	UGA	Stop	Tryptophan
Yeast mitochondria	CUN*	Leucine	Threonine
	AUA	Isoleucine	Methionine
	UGA	Stop	Tryptophan
Higher plant mitochondria	UGA	Stop	Tryptophan
	CGG	Arginine	Tryptophan
Candida albicans nuclei	CTG	Leucine	Serine
Protozoa nuclei	UAA & UAG	Stop	Glutamine
Mycoplasma	UGA	Stop	Tryptophan

*N = Any base.

What about the argument that the code is random: that the existing codons have no inherent advantage? Actually, when we consider the code's effectiveness in dealing with mutations, we find that it is an excellent code indeed. First, consider the fact that single-base changes in the code are likely to result in a shift to a chemically similar amino acid. For example, leucine, isoleucine, and valine all have very similar hydrophobic side chains. And their codons are also very similar, differing only in the first base. So, to pick a particularly advantageous example, a mutation in the first base of the isoleucine codon AUA, could yield UUA, CUA, or GUA. The first two are leucine codons, and the last is a valine codon. Thus, none of these mutations would cause much change in the corresponding amino acid, which minimizes the chance of causing serious damage to the protein product of the mutated gene.

When we consider two other factors, the code looks even better: First, **transitions** (the change of one purine to another, or one pyrimidine to another), are much more common mutations than **transversions,** the change of a purine to a pyrimidine, or vice versa. Second, the ribosome is much more likely to misread the first and third bases in a codon than the second. Considering these things, we can calculate the probability that a single base change will result in no change or just a modest change in the encoded amino acid, for all the possible three-base codes. Then we can see how our natural code stacks up against the others. Figure 18.9 presents a result of this mathematical analysis, which shows that our code is literally one in a million. Only one in a million other possible codes would work better than ours in minimizing the effects of mutations. Given those odds, it seems less likely that our code is just an accident, and not the result of honing by evolution.

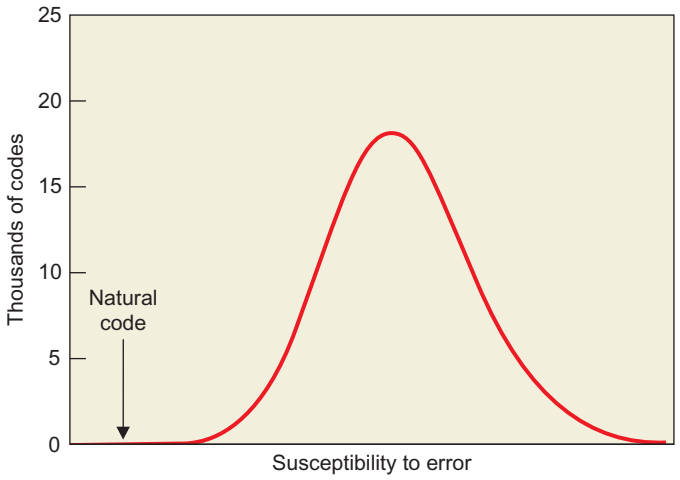


Figure 18.9 Susceptibility of genetic codes to error. The susceptibility to error of all possible triplet genetic codes with four bases is plotted against the number of codes (in thousands) having each susceptibility value. Our own natural code lies far outside the normal distribution, with a very low susceptibility to error. In fact, only one code in a million has a lower susceptibility. (*Source*: Adapted from Vogel, G. Tracking the history of the genetic code. *Science* 281 (17 Jul 1998) 329–331.)

SUMMARY The genetic code is not strictly universal. In certain eukaryotic nuclei and mitochondria and in at least one bacterium, codons that cause termination in the standard genetic code can code for amino acids such as tryptophan and glutamine. In several mitochondrial genomes, and in the nuclei of at least one yeast, the sense of a codon is changed from one amino acid to another. These deviant codes are still

closely related to the standard one from which they probably evolved. It is not clear whether the genetic code is a frozen accident or the product of evolution, but its ability to cope with mutations suggests that it has been subject to evolution.

18.3 The Elongation Cycle

Elongation of a polypeptide chain occurs in a three-step cycle (the **elongation cycle**) that is repeated over and over. We will survey these steps first, then come back and fill in the details, along with experimental evidence.

Overview of Elongation

Figure 18.10 schematically depicts the elongation cycle through two rounds (adding two amino acids to a growing polypeptide chain) in *E. coli*. We start with mRNA and fMet-tRNA$_f^{Met}$ bound to a ribosome. There are three binding sites for aminoacyl-tRNAs on the ribosome. Two of these are called the **P (peptidyl) site** and the **A (aminoacyl) site.** In our schematic diagram, the P site is on the left and the A site is on the right. The fMet-tRNA$_f^{Met}$ is in the P site. A binding site for deacylated tRNA called the **E (exit) site** is empty because the translation process has just begun. Detailed below are the elongation events as shown in Figure 18.10:

a. To begin elongation, we need another amino acid to join with the first. This second amino acid arrives

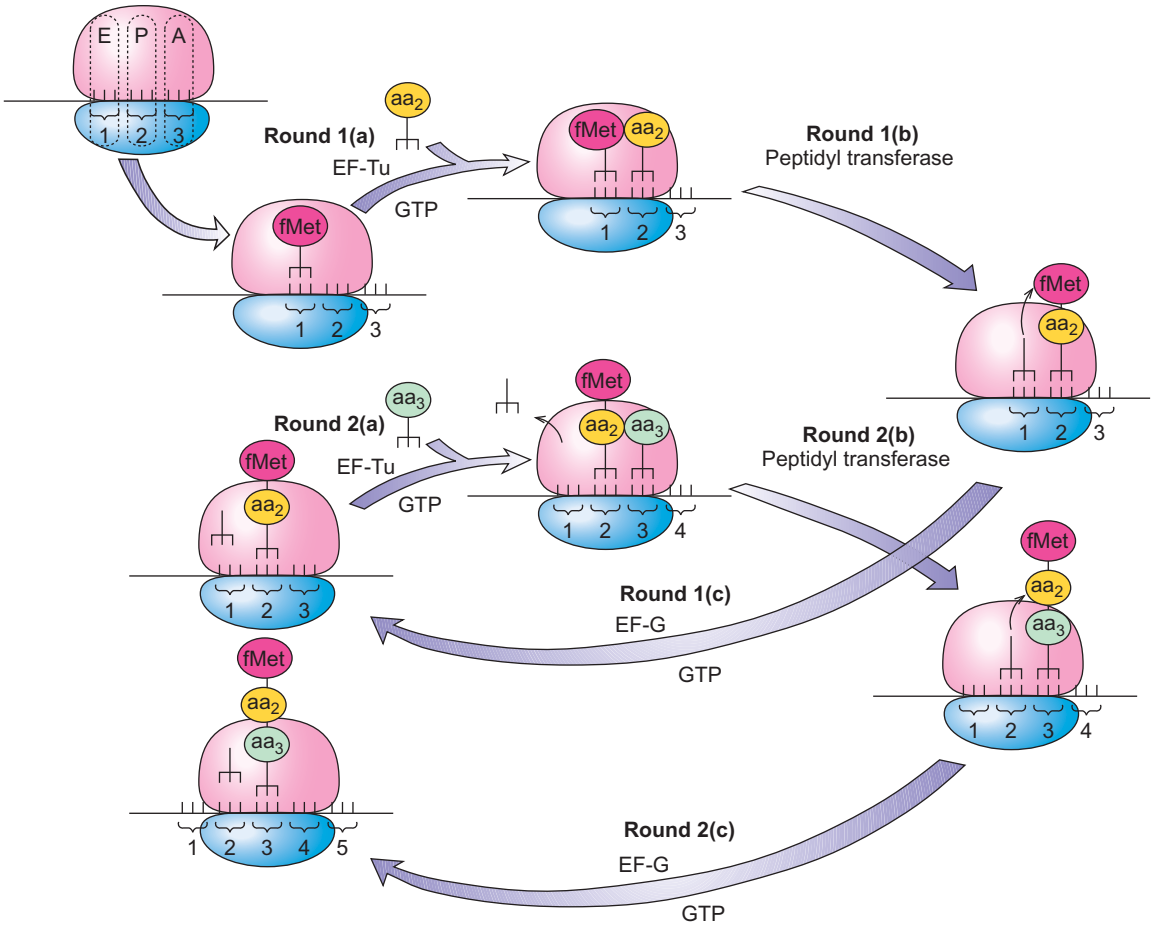


Figure 18.10 Elongation in translation. Note first of all that this is a highly schematic view of protein synthesis. For example, tRNAs are represented by fork-like structures that merely suggest the two business ends of the molecule. Upper left: A ribosome with an mRNA attached is shown to illustrate three sites, E, P and A, indicated with dotted lines. Round I: **(a)** EF-Tu brings in the second aminoacyl-tRNA (yellow) to the A site on the ribosome. The P site is already occupied by fMet-tRNA (magenta). **(b)** Peptidyl transferase forms a peptide bond between fMet and the second aminoacyl-tRNA. **(c)** In the translocation step, EF-G shifts the message and the tRNAs one codon's width to the left. This moves the dipeptidyl-tRNA into the P site, moves the deacylated tRNA in the P site into the E site, and opens up the A site for a new aminoacyl-tRNA. In round 2, these steps are repeated to add one more amino acid (green) to the growing polypeptide. This time, there is a deacylated tRNA in the E site. When EF-Tu brings in the third aminoacyl-tRNA, hydrolysis of the bound GTP allows release of the tRNA from the E site. This opens up the E site for the next translocation step.

bound to a tRNA, and the nature of this aminoacyl-tRNA is dictated by the second codon in the message. The second codon is in the A site, which is otherwise empty, so our second aminoacyl-tRNA will bind to this site. Such binding requires a protein **elongation factor** known as **EF-Tu** (where EF stands for elongation factor) and GTP.

b. Next, the first peptide bond forms. An enzyme called **peptidyl transferase**—an integral part of the large ribosomal subunit—transfers the fMet from its tRNA in the P site to the aminoacyl-tRNA in the A site. This forms a two-amino acid unit called a dipeptide linked to the tRNA in the A site. This whole assembly in the A site is a dipeptidyl-tRNA. What remains in the P site is a deacylated tRNA—a tRNA without its amino acid.

   The formation of the first peptide bond in bacteria is aided by an essential factor known as **EF-P.** Its role appears to be to position the fMet-$tRNA_f^{Met}$ properly for peptide bond formation. A eukaryotic homolog called **eIF5A** probably plays the same role in eukaryotic cells.

c. In the next step, called **translocation,** the mRNA with its peptidyl-tRNA attached in the A site moves one codon's length to the left. This has the following results: (1) The deacylated tRNA in the P site (the one that lost its amino acid during the peptidyl transferase step when the peptide bond formed) moves to the E site. (2) The dipeptidyl-tRNA in the A site, along with its corresponding codon, moves into the P site. (3) The codon that was "waiting in the wings" to the right moves into the A site, ready to interact with an aminoacyl-tRNA. Translocation requires an elongation factor called **EF-G** plus GTP.

The process then repeats itself to add another amino acid: (a) EF-Tu, in conjunction with GTP, brings the appropriate aminoacyl-tRNA to match the new codon in the A site. Upon hydrolysis of GTP by EF-Tu, the deacylated tRNA is ejected from the E site, which makes room for another deacylated tRNA at the end of the second round of elongation. (b) Peptidyl transferase brings the dipeptide from the P site and joins it to the aminoacyl-tRNA in the A site, forming a tripeptidyl-tRNA. (c) EF-G translocates the tripeptidyl-tRNA, together with its mRNA codon, to the P site. At the same time, the deacylated tRNA in the P site moves to the E site.

We have now completed two rounds of peptide chain elongation. We started with an aminoacyl-tRNA (fMet-$tRNA_f^{Met}$) in the P site, and we have lengthened the chain by two amino acids to a tripeptidyl-tRNA. This process continues over and over until the ribosome reaches the last codon in the message. The polypeptide is now complete; it is time for chain termination. The elongation process has been greatly simplified in this brief presentation. It will be fleshed out later in this chapter, and even more in Chapter 19.

SUMMARY Elongation takes place in three steps: (1) EF-Tu, with GTP, binds an aminoacyl-tRNA to the ribosomal A site. (2) Peptidyl transferase forms a peptide bond between the peptide in the P site and the newly arrived aminoacyl-tRNA in the A site. This lengthens the peptide by one amino acid and shifts it to the A site. (3) EF-G, with GTP, translocates the growing peptidyl-tRNA, with its mRNA codon, to the P site, and moves the deacylated tRNA in the P site to the E site.

A Three-Site Model of the Ribosome

The previous section introduced the concept of the three-site ribosome. But what is the evidence for these three sites? We will begin our discussion with the evidence for the A and P sites, and then examine the evidence for the E site. The existence of the A and P sites was originally based on experiments with the antibiotic **puromycin** (Figure 18.11). This drug is an amino acid coupled to an adenosine analog. Thus, it resembles the aminoacyl adenosine at the end of an aminoacyl-tRNA. In fact, it looks enough like an aminoacyl-tRNA that it binds to the A site of a ribosome. Then it can form a peptide bond with the peptide in the P site, yielding a peptidyl puromycin. At this point the ruse is over. The peptidyl puromycin is not tightly bound to the ribosome and so is soon released, aborting translation prematurely. This is why puromycin kills bacteria and other cells.

The link between puromycin and the two-site model is this: Before translocation, because the A site is occupied by a peptidyl-tRNA, puromycin cannot bind and release the peptide; after translocation, the peptidyl-tRNA has moved to the P site, and the A site is open. At this point puromycin can bind and release the peptide. We therefore see two states the ribosome can assume: puromycin reactive and puromycin unreactive. Those two states require at least two binding sites on the ribosome for the peptidyl-tRNA.

Puromycin can be used to show whether an aminoacyl-tRNA is in the A or the P site. If it is in the P site, it can form a peptide bond with puromycin and be released. However, if it is in the A site, it prevents puromycin from binding to the ribosome and is not released.

This same procedure can be used to show that fMet-tRNA goes to the P site in the 70s initiation complex. In our discussion of initiation in Chapter 17, we assumed that the fMet-$tRNA_f^{Met}$ goes to the P site. This certainly makes sense, because it would leave the A site open for the second aminoacyl-tRNA. Using the puromycin assay, M.S. Bretscher and Marcker showed in 1966 that it does indeed go to the P site. They mixed [^{35}S]fMet-$tRNA_f^{Met}$ with ribosomes, the trinucleotide AUG, and puromycin. If AUG attracted fMet-$tRNA_f^{Met}$ to the P site, then the labeled fMet should have been able to react with puromycin, releasing

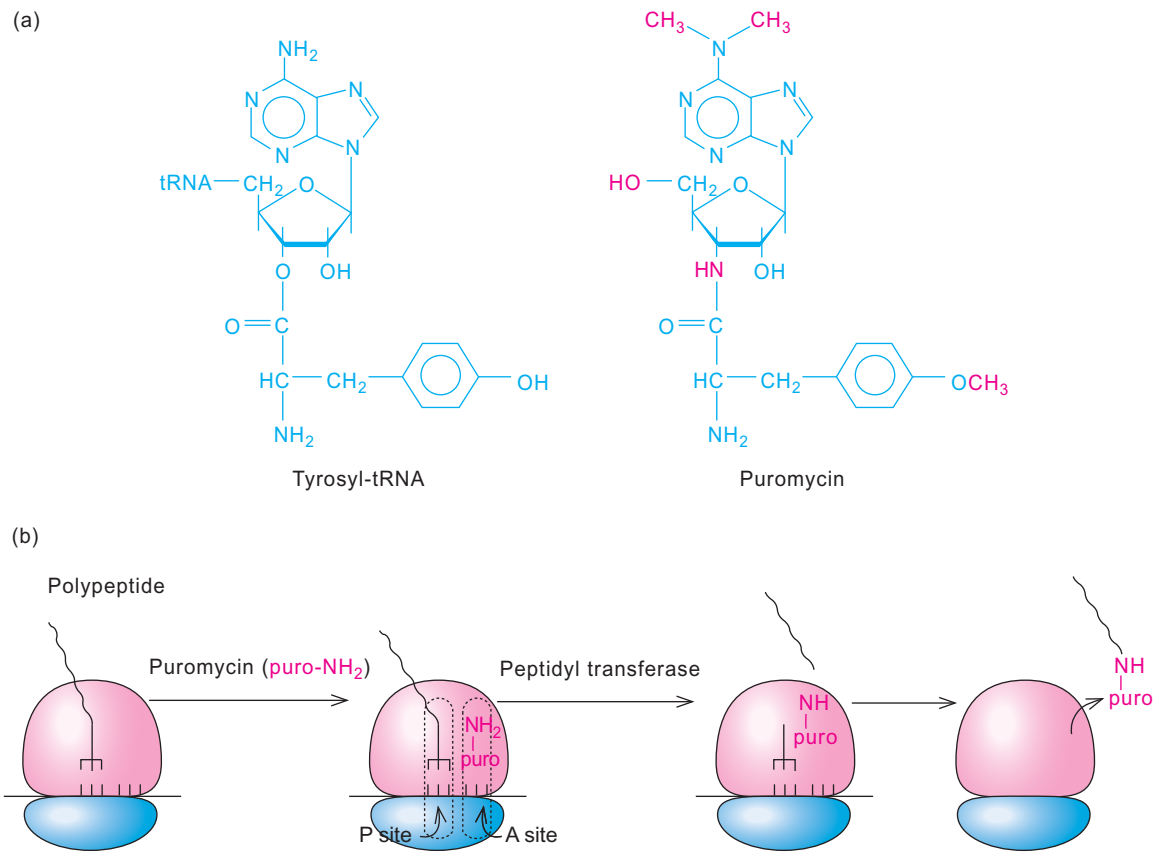


Figure 18.11 Puromycin structure and activity. (a) Comparison of structures of tyrosyl-tRNA and puromycin. Note the rest of the tRNA attached to the 5′-carbon in the aminoacyl-tRNA, where there is only a hydroxyl group in puromycin. The differences between puromycin and tyrosyl-tRNA are highlighted in magenta. **(b)** Mode of action of puromycin. First, puromycin (puro-NH_2) binds to the open A site on the ribosome. (The A site must be open for puromycin to bind.) Next, peptidyl transferase joins the peptide in the P site to the amino group of puromycin in the A site. Finally, the peptidyl-puromycin dissociates from the ribosome, terminating translation prematurely.

labeled fMet-puromycin. On the other hand, if the fMet-$tRNA_f^{Met}$ went to the A site, puromycin should not have been able to bind, so no release of labeled amino acid should have occurred. Figure 18.12 shows that the fMet attached to $tRNA_f^{Met}$ was indeed released by puromycin, whereas the methionine attached to $tRNA_m^{Met}$ was not. Thus, fMet-$tRNA_f^{Met}$ goes to the P site, but methionyl-$tRNA_m^{Met}$ goes to the A site. One could argue that it was the fMet, not the $tRNA_f^{Met}$ that made the difference in this experiment. To eliminate that possibility, Bretscher and Marcker performed the same experiment with Met-$tRNA_f^{Met}$ and found that its methionine was also released by puromycin (Figure 18.12c). Thus, the tRNA, not the formyl group on the methionine, is what targets the aminoacyl-tRNA to the P site.

Actually, x-ray crystallography studies in 2009 showed that fMet-$tRNA_f^{Met}$ does not automatically go to the P site. Instead, on its own, it goes first into a hybrid state called the **P/I state** in which the anticodon of the tRNA is in the P site of the 30S subunit, but the fMet and acceptor stem of the tRNA are not in the P site of the 50S subunit, which encompasses the peptidyl transferase center. Instead, the fMet and acceptor stem are in an "initiator" site to the left of the P site (toward the E site) as the ribosome is conventionally depicted (recall Figure 18.10). It is the job of a protein factor called EF-P to bind to the left of fMet-$tRNA_f^{Met}$ and nudge the fMet and acceptor stem to the right into the peptidyl transferase center. That action puts the fMet-$tRNA_f^{Met}$ fully in the P site.

In 1981, Knud Nierhaus and coworkers presented evidence for a third ribosomal site called the E site. Their experimental strategy was to bind radioactive deacylated $tRNA^{Phe}$ ($tRNA^{Phe}$ lacking phenylalanine), or Phe-$tRNA^{Phe}$, or acetyl-Phe-$tRNA^{Phe}$ to *E. coli* ribosomes and to measure the number of molecules bound per 70S ribosome. Table 18.2 shows the results of binding experiments carried out in the presence or absence of poly(U) mRNA. Only one molecule of acetyl-Phe-$tRNA^{Phe}$ could bind at a time to a ribosome, and the binding site could be either the A site or P site. On the other hand, two molecules of Phe-$tRNA^{Phe}$ could bind, one to the A site, and the other to the P site. Finally, three molecules of deacylated $tRNA^{Phe}$ could bind. We can explain these results most easily by postulating a third site that presumably binds deacylated tRNA on its way out of the ribosome. Hence the E, for exit. In the absence of mRNA, only one tRNA can bind. This can be either deacylated $tRNA^{Phe}$

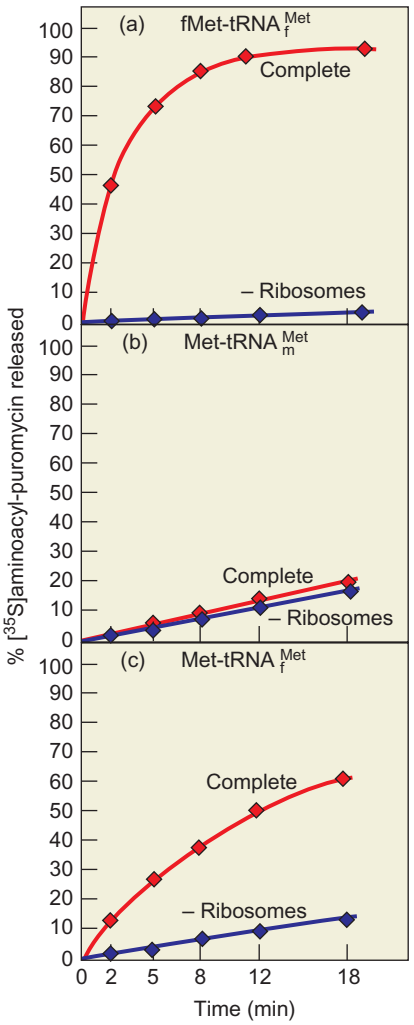


Figure 18.12 fMet-tRNA$_f^{Met}$ occupies the ribosomal P site. Bretscher and Marcker used a puromycin-release assay to determine the location of fMet-tRNA$_f^{Met}$ on the ribosome. They mixed ^{35}S-labeled fMet-tRNA$_f^{Met}$ **(a),** Met-tRNA$_m^{Met}$ **(b),** or Met-tRNA$_f^{Met}$ **(c)** with ribosomes, AUG, and puromycin, and tested for release of labeled fMet- or Met-puromycin by precipitating tRNA and protein with perchloric acid. Aminoacyl-puromycin released from the ribosome is acid-soluble, whereas aminoacyl-tRNA bound to the ribosome is acid-insoluble. The complete reactions contained all ingredients; control reactions lacked one ingredient, as indicated beside each curve. Met or fMet attached to tRNA$_f^{Met}$ went to the P site and was released. Met attached to tRNA$_m^{Met}$ stayed in the A site and was not released by puromycin. (*Source*: Adapted from Bretscher, M.S. and K.A. Marcker, Peptidyl-sRibonucleic acid and amino-acyl-sRibonucleic acid binding sites on ribosomes. *Nature* 211:382–83, 1966.)

or acetyl-Phe-tRNA^{Phe}. Nierhaus and colleagues speculated that the binding site was the P site, and subsequent work has confirmed this suspicion.

We will discuss the E site in greater detail in Chapter 19, but we should note at this point that the E site is not just a way station for deacylated tRNA on its way out of the ribosome. It plays a critical role in maintaining the reading frame of an mRNA. Ordinarily, reading frame shifts occur only about once in 30,000 codons, which is a good thing because such shifts generally give rise to meaningless proteins. But proper translation of some mRNAs actually depends on frameshifting.

Table 18.2 Binding of tRNAs and Aminoacyl-tRNAs to *E. coli* Ribosomes

	tRNA	Binding sites	
mRNA	**Species**	**No.**	**Location**
Poly(U)	Acetyl-Phe-tRNA^{Phe}	1	P or A
Poly(U)	Phe-tRNA^{Phe}	2	P and A
Poly(U)	tRNA^{Phe}	3	P, E, and A
None	tRNA^{Phe}	1	P
None	Phe-tRNA^{Phe}	0	—
None	Acetyl-Phe-tRNA^{Phe}	1	P

Source: Rheinberger, H.-J., H. Sternbach, and K.H. Nierhaus, Three tRNA binding sites on *Escherichia coli* ribosomes, *Proceedings of the National Academy of Sciences USA* 78(9):5310–14, September 1981. Reprinted with permission.

An example is the *E. coli prfB* gene, which encodes RF2, a release factor we will study later in this chapter. In order for the *prfB* mRNA to be translated correctly, a frameshift to the +1 reading frame must occur within the mRNA. Thus, the sequence CUUUGAC would normally be read: CUU UGA (Leu, Stop). But, with the +1 frameshift, it is read CUU*U*GAC (Leu, Asp). The italicized U is skipped, and the next codon is the underlined GAC, which encodes aspartate.

In 2004, Knud Nierhaus and colleagues examined translation of the *prfB* mRNA in vitro and found that the presence of a deacylated tRNA in the E site prevented this frameshift. When they removed the deacylated tRNA from the E site, the frameshift occurred with high frequency. Thus, they concluded that deacylated tRNA in the E site is normally required for the vital purpose of maintaining the proper reading frame. When frameshifting is required for proper translation of a particular mRNA, the cell must remove the deacylated tRNA from the E site.

SUMMARY Puromycin resembles an aminoacyl-tRNA and so can bind to the A site, couple with the peptide in the P site, and release it as peptidyl puromycin. On the other hand, if the peptidyl-tRNA is in the A site, puromycin will not bind to the ribosome, and the peptide will not be released. This defines two sites on the ribosome: a puromycin-reactive site (P), and a puromycin unreactive site (A). fMet-tRNA$_f^{Met}$ is puromycin reactive in the 70S initiation complex, so it is in the P site. Other studies have identified a third binding site (the E site) for deacylated tRNA. Such tRNAs presumably bind to the E site as they exit the ribosome, and this binding helps maintain the reading frame of the mRNA.

Elongation Step 1: Binding an Aminoacyl-tRNA to the A Site of the Ribosome

Our detailed understanding of the elongation process began in 1965 when Yasutomi Nishizuka and Fritz Lipmann used anion exchange chromatography to separate two protein factors required for peptide bond formation in *E. coli*. They named one factor T, for transfer, because it transfers aminoacyl-tRNAs to the ribosome. The second factor they called G because of its GTPase activity. (T also has GTPase activity, as we will see.) Then Jean Lucas-Lenard and Lipmann showed that T is actually composed of two different proteins, which they called Tu (where the **u** stands for unstable) and Ts (where the **s** stands for stable). These three factors, which we now call EF-Tu (or EF1A), **EF-Ts** (or EF1B), and EF-G (or EF2), participate in the first and third steps in elongation. (In eukaryotes, the roles of EF-Tu and EF-Ts are played by a three-subunit protein known as **EF1.** The EF1 α subunit performs the EF-Tu role, and the β and γ subunits perform the EF-Ts role. The EF-G role in eukaryotes is played by **EF2.**) Let us consider first the activities of EF-Tu and -Ts because they are involved in the first elongation step.

Joanne Ravel showed in 1967 that unfractionated EF-T (Tu plus Ts) had GTPase activity, and that EF-T required GTP to bind an aminoacyl-tRNA to the ribosome. To demonstrate this phenomenon, she made [^{14}C]Phe-tRNA^{Phe} and added it to washed ribosomes along with EF-T and an increasing concentration of GTP. Then she filtered the ribosomes through nitrocellulose. Labeled Phe-tRNA^{Phe} that bound to ribosomes stuck to the filter, but unbound Phe-tRNA^{Phe} washed through. Figure 18.13a depicts the results. Background nonenzymatic binding of the Phe-tRNA^{Phe} to the ribosomes was rather high in the absence of EF-T and GTP, but this was not physiologically significant. Ignoring that background, we can see that GTP was necessary for EF-T-dependent binding of Phe-tRNA^{Phe} to the ribosomes.

When Ravel added both EF-T and EF-G to washed ribosomes in the presence of poly(U) and labeled Phe-tRNA^{Phe} she found that the ribosomes made labeled polyphenylalanine. And this polymerization of amino acids required an even higher concentration of GTP than the aminoacyl-tRNA-binding reaction did.

When we examined initiation of translation, we learned that IF-2-mediated binding of fMet-tRNA$_f^{Met}$ to ribosomes also required GTP, but that GTP hydrolysis was not required. Could the same be true of EF-T and binding of ordinary aminoacyl-tRNAs to ribosomes? Anne-Lise Haenni and Lucas-Lenard showed in 1968 that this is indeed the case. They labeled *N*-acetyl-Phe-tRNA with ^{14}C and Phe-tRNA with ^{3}H. Then they mixed these labeled aminoacyl-tRNAs with EF-T and either GTP or the unhydrolyzable analog, GDPCP. Under the non-physiological conditions of this experiment, the *N*-acetyl-Phe-tRNA^{Phe} went to the P site. These workers measured binding of aminoacyl-tRNAs to ribosomes by filter binding, as described in Figure 18.13. They also measured

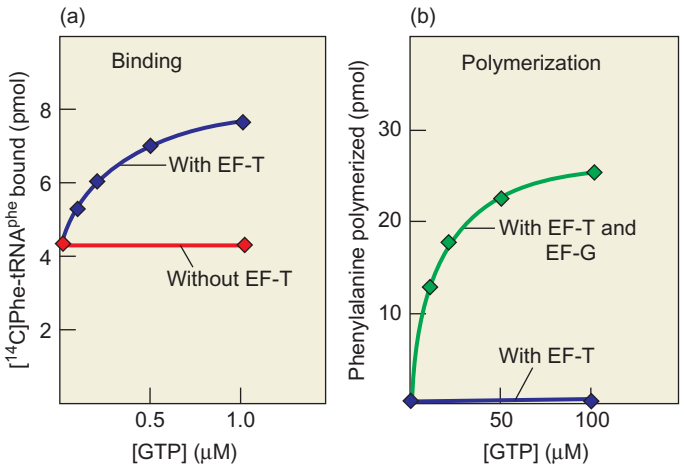


Figure 18.13 Effects of EF-T and GTP on Phe-tRNA^{Phe} binding to ribosomes and on poly-Phe synthesis. (a) Binding Phe-tRNA^{Phe} to ribosomes. Ravel mixed ^{14}C-Phe-tRNA^{Phe} with washed ribosomes and various concentrations of GTP in the presence or absence of EF-T. She measured Phe-tRNA^{Phe}–ribosome binding by filtering the mixture and determining the labeled Phe bound to the ribosomes on the filter. Considerable nonenzymatic binding occurred in the absence of EF-T and GTP, but the EF-T-dependent binding required GTP. **(b)** Polymerization of phenylalanine. Ravel mixed labeled Phe-tRNA^{Phe} with ribosomes, EF-T, and various concentrations of GTP in the presence and absence of EF-G. She measured polymerization of Phe by acid precipitation as follows: She precipitated the poly(Phe) with trichloroacetic acid (TCA), heated the precipitate in the presence of TCA to hydrolyze any phe-tRNA^{Phe}, and trapped the precipitated poly(Phe) on filters. Polymerization required both EF-T and EF-G and a high concentration of GTP. (*Source*: Adapted from Ravel, J.M., Demonstration of a guanosine triphosphate-dependent enzymatic binding of aminoacyl-ribonucleic acid to *Escherichia coli* ribosomes. *Proceedings of the National Academy of Sciences USA* 57:1815, 1967.)

peptide bond formation between the *N*-acetyl-Phe in the P site and the Phe-tRNA^{Phe} in the A site by extracting the dipeptide product and identifying it by paper electrophoresis. Table 18.3 shows that *N*-acetyl-Phe-tRNA^{Phe} could bind to the P site, and that Phe-tRNA^{Phe} could bind to the A site, with the help of EF-T and either GTP or GDPCP. (In fact, *N*-acetyl-Phe-tRNA^{Phe} did not even need EF-T to bind to the P site.) Thus, GTP hydrolysis is not needed for EF-T to promote aminoacyl-tRNA binding to the ribosomal A site. In marked contrast, formation of the peptide bond between *N*-acetyl-Phe and Phe-tRNA^{Phe} required GTP hydrolysis. This is analogous to the situation in initiation, where IF-2 can bind fMet-tRNA$_f^{Met}$ to the P site without GTP hydrolysis, but subsequent events are blocked until GTP is hydrolyzed.

These same scientists also demonstrated that *both* EF-Tu and EF-Ts are required for Phe-tRNA^{Phe} binding to the ribosome. The assay was the same as in Table 18.3, except that no GDPCP was used and that EF-Tu and EF-Ts were separated from each other (except for some residual contamination of the EF-Tu fraction with EF-Ts), and added separately. Table 18.4 shows that both EF-Tu and -Ts are required for Phe-tRNA^{Phe}-ribosome binding. The small amount of binding seen with EF-Tu alone resulted from contamination of the factor by EF-Ts.

Table 18.3 Effect of GTP and GDPCP on Aminoacyl-tRNA Binding to Ribosomes and on Binding Plus Peptide Bond Formation

Additions	*N*-acetyl-Phe-tRNA^{Phe} bound (^{14}C) (pmol)	*N*-acetyl diPhe-tRNA formed (^{14}C or ^{3}H) (pmol)	Phe-tRNA bound (^{3}H) (pmol)
None	7.6	0.4	0.1
EF-T + GTP	3.0	4.5	2.8
EF-T + GDPCP	7.0	0.5	4.8

Source: Haenni, A.L. and J. Lucas-Lenard, Stepwise synthesis of a tripeptide, *Proceedings of the National Academy of Sciences, USA* 61:1365, 1968. Reprinted by permission.

Table 18.4 Requirement for Both EF-Ts and EF-Tu to Bind [^{3}H]Phe-tRNA to Ribosomes Carrying Prebound *N*-acetyl-[^{14}C]Phe-tRNA

Additions	[^{3}H]Phe-tRNA bound (pmol)
None	2.8
EF-Ts + GTP	2.8
EF-Tu + GTP	5.2
EF-Ts + EF-Tu + GTP	11.6

Source: Naenni, A.L., and J. Lucas-Lenard, Stepwise synthesis of a tripeptide, *Proceedings of the National Academy of Sciences USA* 61:1365, 1968. Reprinted with permission.

Figure 18.14 presents a model for the detailed mechanism by which EF-Tu and EF-Ts cooperate to cause transfer of aminoacyl-tRNAs to the ribosome. First, EF-Tu and GTP form a binary (two-part) complex. Then aminoacyl-tRNA joins the complex, forming a ternary (three-part) complex composed of EF-Tu, GTP, and aminoacyl-tRNA. This ternary complex then delivers its aminoacyl-tRNA to the ribosome's A site. EF-Tu and GTP remain bound to the ribosome. Next, GTP is hydrolyzed and an EF-Tu–GDP complex dissociates from the ribosome. Finally, EF-Ts exchanges GTP for GDP on the complex, yielding an EF-Tu–GTP complex.

What is the evidence for this scheme? Herbert Weissbach and colleagues found in 1967 that an EF-T preparation and GTP could form a complex that was retained by a nitrocellulose filter. They labeled GTP, mixed it with EF-T, and found that the labeled nucleotide bound to the filter. This meant that GTP had bound to a protein in the EF-T preparation,

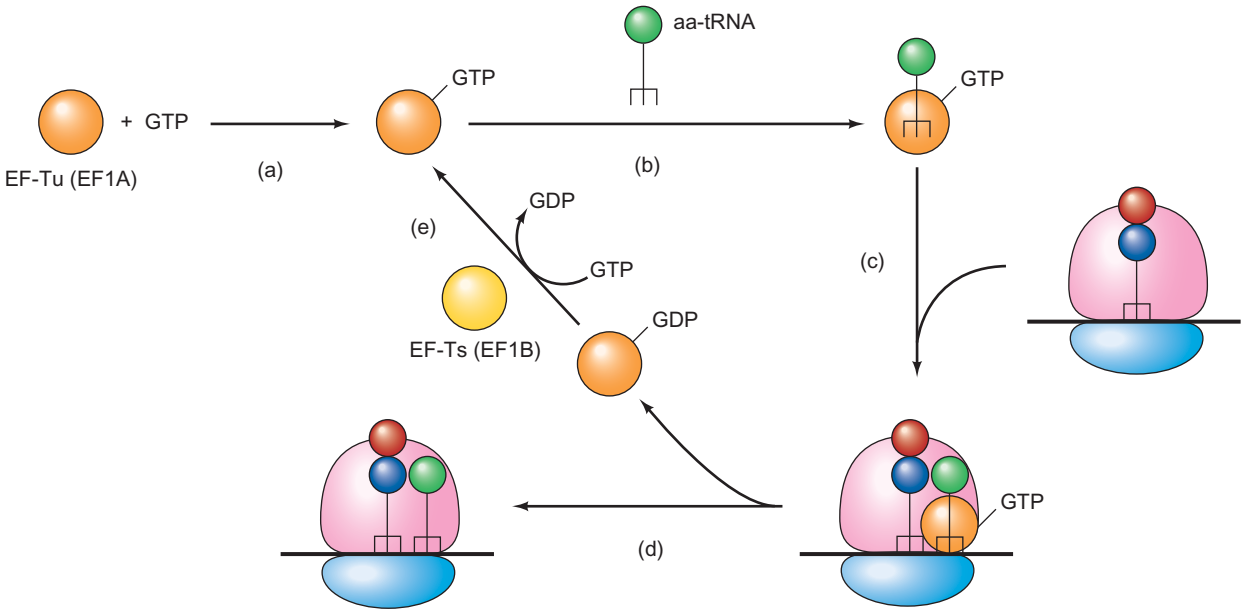


Figure 18.14 Model of binding aminoacyl-tRNAs to the ribosome A site. (a) EF-Tu couples with GTP to form a binary complex. **(b)** This complex associates with an aminoacyl-tRNA to form a ternary complex. **(c)** The ternary complex binds to a ribosome with a peptidyl-tRNA in its P site and an empty A site. **(d)** GTP is hydrolyzed and the resulting EF-Tu–GDP complex dissociates from the ribosome, leaving the new aminoacyl-tRNA in the A site. **(e)** EF-Ts exchanges GTP for GDP on EF-Tu, regenerating the EF-Tu–GTP complex.

presumably EF-T itself, to form a complex. Julian Gordon then discovered that adding an aminoacyl-tRNA to the EF-Tu–GTP complex caused the complex to be released from the filter. One interpretation of this behavior is that the aminoacyl-tRNA joined the EF-Tu–GTP complex to form a ternary complex that could no longer bind to the filter.

Ravel and her collaborators gave us additional evidence for the formation of the ternary complex with the following experiment. They labeled GTP with ^{3}H and ^{32}P, and Phe-tRNA^{Phe} with ^{14}C, and mixed them with EF-T, then subjected the mixture to gel filtration on Sephadex G100 (Chapter 5). This gel filtration resin excludes relatively large proteins, such as EF-T, so they flow through rapidly in a fraction called the void volume. By contrast, relatively small substances like GTP, and even Phe-tRNA^{Phe}, enter the pores in the resin and are thereby retarded; they emerge later from the column, after the void volume. In fact, the smaller the molecule, the longer it takes to elute from the column. Figure 18.15 shows the results of this gel filtration experiment. A fraction of both labeled substances, GTP and Phe-tRNA^{Phe}, emerged relatively late, in their usual positions. These fractions represented free GTP and Phe-tRNA^{Phe}, although very little free GTP was observed. However, significant fractions of both substances eluted much earlier, around fraction 20, demonstrating that they must be complexed to something larger. The predominant larger substance in this experiment was EF-T, and the experiments we have already discussed implicate EF-T in this complex, so we infer that a ternary complex, involving Phe-tRNA^{Phe}, GTP, and EF-T has formed.

So far, we have not distinguished between EF-Ts and EF-Tu in these experiments. Herbert Weissbach and his collaborators did this by separating the two proteins and testing them separately. They found that EF-Tu is the factor

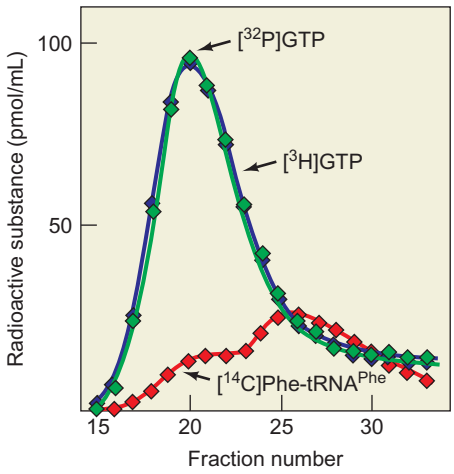


Figure 18.15 Formation of a ternary complex among EF-T, aminoacyl-tRNA, and GTP. Ravel and colleagues mixed [^{14}C]Phe-tRNA^{Phe} with GTP (labeled in the guanine part with ^{3}H and in the γ-phosphate with ^{32}P), and EF-T. Then they passed the mixture through a Sephadex G100 gel filtration column to separate large molecules, such as EF-T, from relatively small molecules such as GTP and Phe-tRNA^{Phe}. They assayed each fraction for the three radioisotopes to detect GTP and Phe-tRNA^{Phe}. Both of these substances were found at least partly in a large-molecule fraction (around fraction 20), so they were bound to the EF-T in a complex. *(Source:* Adapted from Ravel, J.M., R.L. Shorey, and W. Shive, The composition of the active intermediate in the transfer of aminoacyl-RNA to ribosomes. *Biochemical and Biophysical Research Communications* 32:12, 1968.)

that binds GTP in the binary complex. What then is the role of EF-Ts? These investigators demonstrated that this factor is essential for conversion of the EF-Tu–GDP complex to the EF-Tu–GTP complex. However, EF-Ts has little, if any, effect when it is presented with the pre-formed EF-Tu–GTP complex or with EF-Tu itself (Figure 18.16).

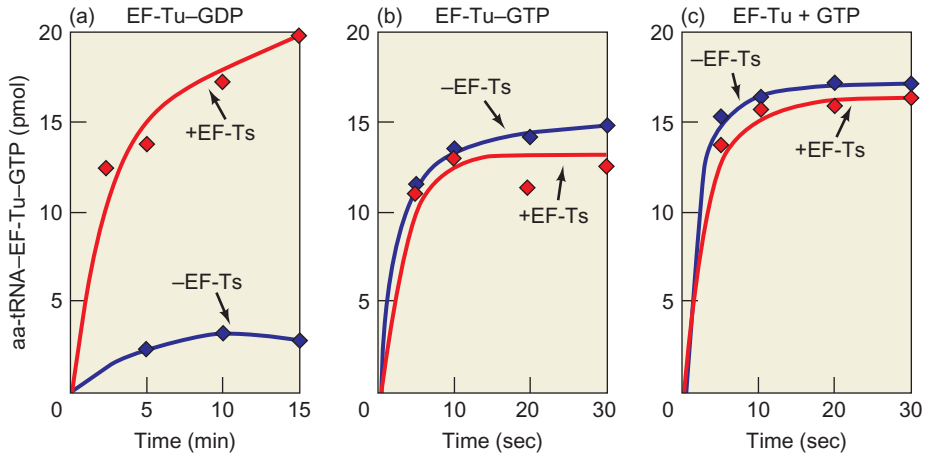


Figure 18.16 Effect of EF-Ts on ternary complex formation. Weissbach and colleagues attempted to form the ternary complex with [^{14}C]Phe-tRNA, [^{3}H]GTP, and the EF-Tu preparations listed at top, with (red) and without (blue) EF-Ts. They measured ternary complex formation by loss of radioactivity trapped by nitrocellulose filtration. EF-Ts stimulated complex formation only when EF-Tu–GDP was the substrate (panel **a**). EF-Tu–GTP (panel **b**) or EF-Tu+GTP (panel **c**) could form the complex spontaneously, with no help from EF-Ts. (aa-tRNA = aminoacyl-tRNA). (*Source:* Adapted from Weissbach, H., D.L. Miller, and J. Hachmann, Studies on the role of factor Ts in polypeptide synthesis. *Archives of Biochemistry and Biophysics,* 137:267, 1970.)

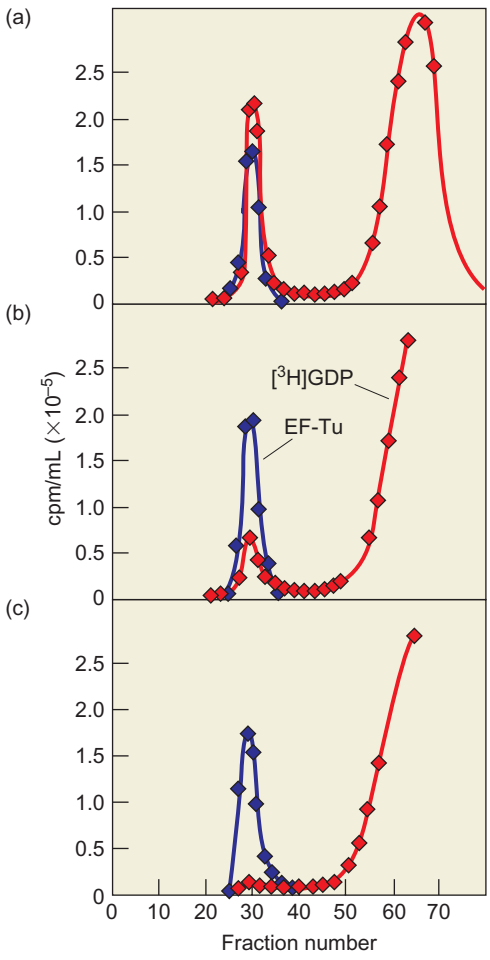


Figure 18.17 Displacement of GDP from an EF-Tu–GDP complex by EF-Ts. Miller and Weissbach mixed an EF-Tu–[^{3}H]GDP complex with three different amounts of EF-Ts, then detected the amount of GDP remaining in the complex by gel filtration through Sephadex G-25. The three panels contained the following amounts of EF-Ts: **(a),** 500 units; **(b),** 14,000 units; **(c),** 25,000 units. Red, [^{3}H]GDP; blue, EF-Tu. (*Source*: Adapted from Miller, D.L. and H. Weissbach, Interactions between the elongation factors: The displacement of GDP from the Tu-GDP complex by factor Ts. *Biochemical and Biophysical Research Communications* 38:1019, 1970.)

Thus, it seems that EF-Ts does not form a complex directly from EF-Tu and GTP. Instead, it converts EF-Tu–GDP to EF-Tu–GTP by exchanging the guanine nucleotide.

How does EF-Ts perform its exchange duty? David Miller and Weissbach showed that EF-Ts can displace GDP from EF-Tu–GDP (Figure 18.17) by forming an EF-Ts–EF-Tu complex. How does this displacement work? X-ray crystallography studies on EF-Tu–EF-Ts complexes by Reuben Leberman and colleagues have shown that one of the main consequences of EF-Ts binding to EF-Tu–GDP is disruption of the Mg^{2+}-binding center of EF-Tu. The weakened binding between EF-Tu and Mg^{2+} leads to dissociation of GDP, which opens the way for binding of GTP to EF-Tu.

Why is EF-Tu needed to escort aminoacyl-tRNAs to the ribosome? The ester bond joining the amino acid to its cognate tRNA is easily broken, and sequestering the aminoacyl-tRNA within the EF-Tu protein protects this labile compound from hydrolysis. But the concentration of aminoacyl-tRNAs in the cell is quite high. Is there enough EF-Tu to go around? Yes, because EF-Tu is one of the most abundant proteins in the cell. For example, EF-Tu constitutes 5% of the total protein in *E. coli* cells, and the reason for this abundance appears to be the important protective role that EF-Tu plays.

SUMMARY A ternary complex formed from EF-Tu, aminoacyl-tRNA, and GTP delivers an aminoacyl-tRNA to the ribosome's A site, without hydrolysis of the GTP. In the next step, EF-Tu hydrolyzes GTP with its ribosome-dependent GTPase activity, and an EF-Tu–GDP complex dissociates from the ribosome. EF-Ts regenerates an EF-Tu–GTP complex by exchanging GTP for GDP attached to EF-Tu. Addition of aminoacyl-tRNA then reconstitutes the ternary complex for another round of translation elongation.

Proofreading As we will see in Chapter 19, part of the accuracy of protein synthesis comes from charging of tRNAs with the correct amino acids. But part also comes in elongation step 1: The ribosome usually binds the aminoacyl-tRNA called for by the codon in the A site. However, if it makes a mistake in this initial recognition step, it still has a chance to correct it by rejecting an incorrect aminoacyl-tRNA before it can donate its amino acid to the growing polypeptide. This process is called **proofreading.**

Proofreading can occur at two steps within step 1 of elongation: First, the ternary complex can dissociate from the ribosome after binding, and this happens more readily if a ternary complex with the wrong aminoacyl-tRNA has bound. Second, the aminoacyl-tRNA (derived from the ternary complex) can dissociate from the ribosome. Again, this happens at a much higher rate if the aminoacyl-tRNA is incorrect than it does when it is correct, because of the weakness of the imperfect codon–anticodon base pairing. This is generally fast enough that an incorrect aminoacyl-tRNA dissociates from the ribosome before its amino acid has a chance to be incorporated into the nascent polypeptide.

A general principle that emerges from the analysis of accuracy in translation is that a high degree of accuracy and a high rate of translation are incompatible. In fact, accuracy and speed are inversely related: The faster translation goes, the less accurate it becomes. This is because the ribosome must allow enough time for incorrect ternary complexes and aminoacyl-tRNAs to leave before the incorrect amino acid is irreversibly incorporated into the growing polypeptide. If translation goes faster, more incorrect amino acids will be incorporated. Conversely, if translation goes more

slowly, accuracy will be higher, but then proteins may not be made fast enough to sustain life. So there is a delicate balance between speed and accuracy of translation.

One of the most important factors in this balance is the rate of hydrolysis of GTP by EF-Tu. If the rate were higher, less time would be available for the first proofreading step: EF-Tu would hydrolyze GTP to GDP quickly without giving sufficient time for ternary complexes bearing improper aminoacyl-tRNAs to dissociate from the ribosome. On the other hand, if the rate were lower, there would be ample time for proofreading, but translation would be too slow. What is the proper rate? In *E. coli,* the average time between binding of the ternary complex and hydrolysis of GTP is several milliseconds. Then it takes several milliseconds more for EF-Tu–GDP to dissociate from the ribosome. Proofreading takes place during both of these pauses, and shortening one or both could be devastating to accuracy of translation.

How large an error rate in translation can a cell tolerate? What if it were 1%, for example? Ninety-nine percent accuracy sounds pretty good until you consider that the lengths of most polypeptides are much more than 100 amino acids. They average about 300 amino acids long, and some are more than 1000 amino acids long. The probability p of producing an error-free polypeptide, given an error rate per amino acid (ε) and a polypeptide length (n) is given by the following expression:

$$p = (1 - \varepsilon)^n$$

For example, with an error rate of 1%, an average-size polypeptide would be produced error-free only about 5% of the time, and a 1000-amino-acid polypeptide would almost never be error-free. With a 10-fold better error rate, 0.1%, an average-size polypeptide would be produced error-free about 74% of the time, but a 1000-amino-acid polypeptide would be made error-free only about 37% of the time. This would still pose a problem for large polypeptides. What if the error rate were only 0.01%? At that rate, about 97% of average-size polypeptides, and about 91% of 1000-amino-acid polypeptides would be produced error-free. That seems like an acceptable rate of production of defective proteins, and the observed error rate per amino acid added, at least in *E. coli,* is in fact close to 0.01%.

An important antibiotic known as **streptomycin** interferes with proofreading so the ribosome makes more mistakes. For example, normal ribosomes incorporate phenylalanine almost exclusively in response to the synthetic message poly(U). But streptomycin greatly stimulates the incorporation of isoleucine and, to a lesser extent, serine and leucine in response to poly(U).

Certain natural conditions allow us to see what happens when the rate of translation is either faster or slower than normal. For example, mutants in ribosomal proteins, such as *ram,* or in EF-Tu, such as *tufAr,* double the rate of peptide bond formation. In these mutants, accuracy of translation suffers because not enough time is available for incorrect aminoacyl-tRNAs to dissociate from the ribosome.

By contrast, in streptomycin-resistant mutants such as *strA,* the rate of peptide bond formation is only half the normal value. This allows extra time for incorrect aminoacyl-tRNAs to leave the ribosome, so translation is extra accurate.

SUMMARY The protein-synthesizing machinery achieves accuracy during elongation in a two-step process. First, it gets rid of ternary complexes bearing the wrong aminoacyl-tRNA before GTP hydrolysis occurs. If this screen fails, it can still eliminate the incorrect aminoacyl-tRNA in the proofreading step before the wrong amino acid can be incorporated into the growing protein chain. Presumably, both these screens rely on the weakness of incorrect codon–anticodon base pairing to ensure that dissociation will occur more rapidly than either GTP hydrolysis or peptide bond formation. The balance between speed and accuracy of translation is delicate. If peptide bond formation goes too fast, incorrect aminoacyl-tRNAs do not have enough time to leave the ribosome, so their amino acids are incorporated into protein. But if translation goes too slowly, proteins are not made fast enough for the organism to grow successfully. The actual error rate, about 0.01% per amino acid added, strikes a good balance between speed and accuracy.

Elongation Step 2: Peptide Bond Formation

After the initiation factors and EF-Tu have done their jobs, the ribosome has fMet-tRNA$_f^{Met}$ in the P site and an aminoacyl-tRNA in the A site. Now it is time to form the first peptide bond. You might be expecting a new group of elongation factors to participate in this event, but there are none. Instead, the ribosome itself contains the enzymatic activity, called **peptidyl transferase,** that forms peptide bonds. No soluble factors are needed.

The peptidyl transferase step in prokaryotes is inhibited by an important antibiotic called **chloramphenicol.** This drug has no effect on most eukaryotic ribosomes, which makes it selective for bacterial invaders in higher organisms. However, the mitochondria of eukaryotes have their own ribosomes, and chloramphenicol does inhibit their peptidyl transferase. Thus, chloramphenicol's selectivity for bacteria is not absolute.

The classic assay for peptidyl transferase was invented by Robert Traut and Robert Monro and uses a labeled aminoacyl-tRNA or peptidyl-tRNA bound to the ribosomal P site, and puromycin. The release of labeled aminoacyl- or

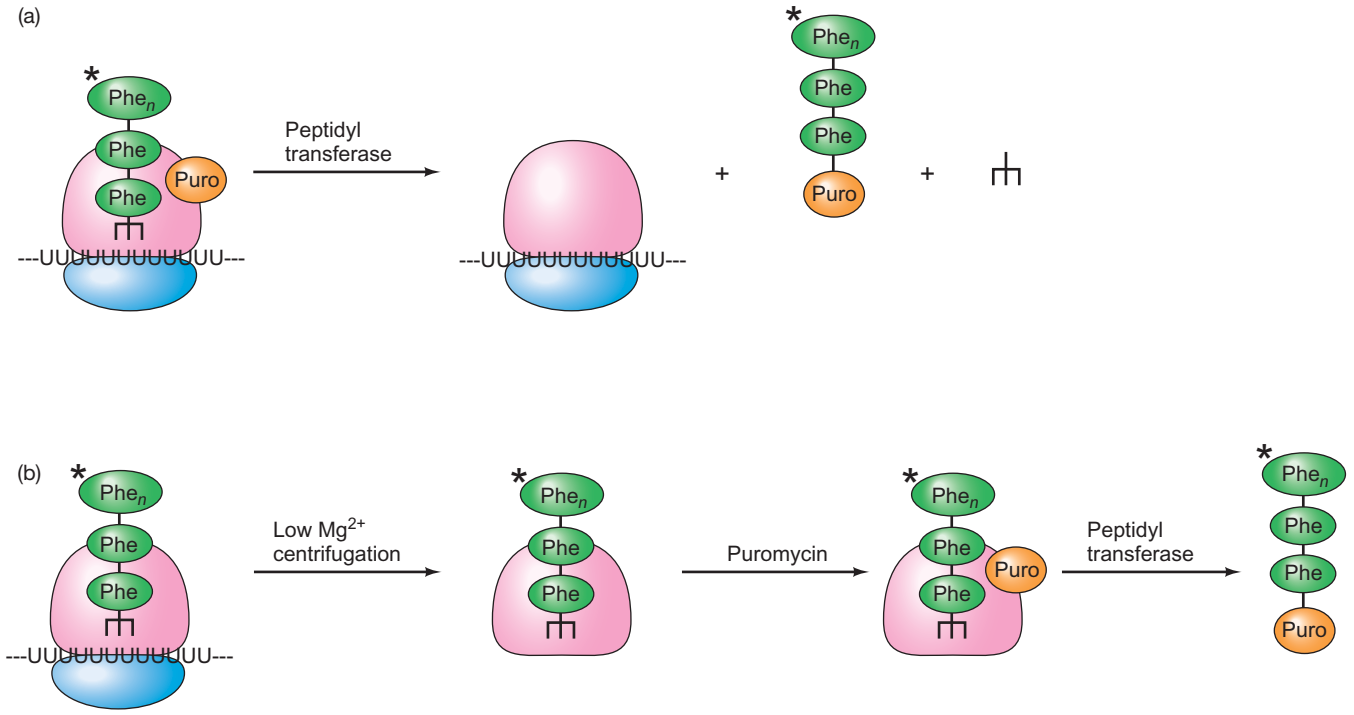


Figure 18.18 The puromycin reaction as an assay for peptidyl transferase. (a) Standard puromycin reaction. Add labeled poly(Phe)–tRNA into the P site by running a translation reaction with poly(U) as messenger. Then add puromycin. When a peptide bond forms between the labeled poly(Phe) and puromycin, the labeled peptidyl-puromycin is released. **(b)** Reaction with 50S subunits only. Again, add labeled poly(Phe)-tRNA into the ribosome's P site, incubate in a low Mg^{2+} buffer, and then centrifuge to separate the 50S-poly(Phe)–tRNA complex from the 30S subunit and the mRNA. Then add puromycin and detect peptidyl transferase by the release of labeled peptidyl-puromycin. The by-products of the reaction (50S subunits and tRNA) are not pictured. The asterisks denote the label in the poly(Phe).

peptidyl-puromycin depends on forming a peptide bond between the amino acid or peptide in the P site, and puromycin in the A site, as depicted in Figure 18.18a. Traut and Monro also discovered that this system could be modified somewhat to show that the 50S ribosomal subunit, without any help from the 30S subunit or soluble factors, could carry out the peptidyl transferase reaction (Figure 18.18b). First, they allowed ribosomes to carry out poly(Phe) synthesis, using poly(U) as mRNA. This placed labeled poly(Phe)-tRNA in the P site. Then they removed the 30S subunits by incubation with buffer having a low Mg^{2+} concentration, followed by ultracentrifugation. Then they washed away any remaining initiation or elongation factors with salt solutions, leaving the 50S subunits bound to poly(Phe)-tRNA. Ordinarily, such primed 50S subunits would be unreactive with puromycin, but these workers found that they could elicit puromycin reactivity with 33% methanol (ethanol also worked). In both assays, one must distinguish the released peptidyl-tRNA from the peptidyl-tRNA still bound to ribosomes. Traut and Monro originally accomplished this by sucrose gradient centrifugation, as shown in Figure 18.19. Later, a more convenient filter-binding assay was developed. Figure 18.19a is a negative control with no puromycin, and the poly(Phe) remained bound to the 50S subunit, as expected. Figure 18.19b is a positive control in which the poly(Phe) was released by destroying the ribosomes with urea and RNase. Figure 18.19c and 18.19d show the experimental results with puromycin plus and minus GTP, respectively. Peptidyl transferase appeared to be working, since puromycin could release the poly(Phe). This reaction occurred even in the absence of GTP, as the peptidyl transferase reaction should.

The puromycin reaction with 50S subunits seems to demonstrate that the 50S subunit contains the peptidyl transferase activity, but could the rather unphysiological conditions (33% methanol, puromycin) be distorting the picture? One encouraging sign is that the reaction of a peptide with puromycin seems to follow the same mechanism as normal peptide synthesis. Also, M.A. Gottesmann substituted poly(A) for poly(U), and therefore poly(Lys) for poly(Phe), and also substituted lysyl-tRNA for puromycin, and found the same kind of reaction, demonstrating that the puromycin reaction is a valid model for peptide bond formation. Furthermore, these reactions are all blocked by chloramphenicol and other antibiotics that inhibit the normal peptidyl transferase reaction, suggesting that the model reactions use the same pathway as the normal one.

For decades, no one knew what part of the 50S subunit had the peptidyl transferase activity. However, as soon as Thomas Cech and coworkers demonstrated in the early

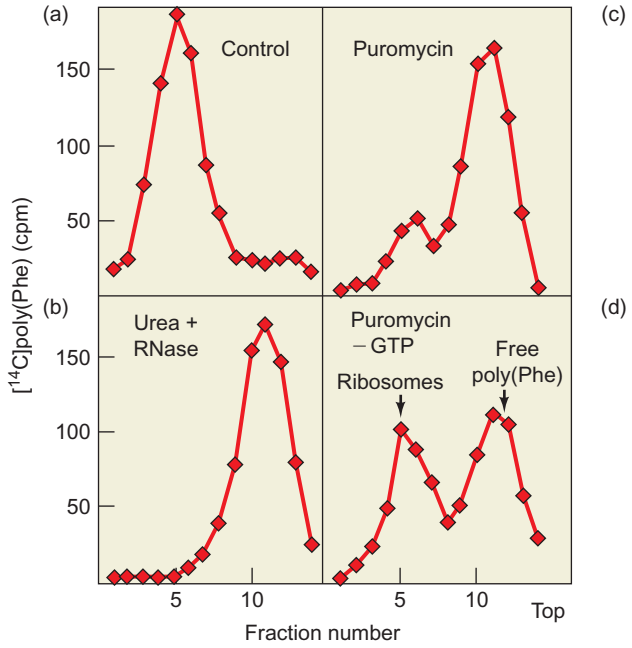


Figure 18.19 Puromycin assay for peptide bond formation. Traut and Monro loaded ribosomes with [^{14}C]polyphenylalanine, incubated them with or without puromycin, then subjected the products to sucrose gradient centrifugation to separate ribosome-bound poly(Phe) from free poly(Phe) that had been released from the ribosomes. The poly(Phe)-loaded ribosomes were treated as follows: **(a)** no treatment; **(b)** treated with urea and RNase; **(c)** treated with puromycin; **(d)** treated with puromycin in the absence of GTP. The positions of ribosomes and free poly(Phe) are indicated in (d). (*Source*: Adapted from Traut, R.R. and R.E. Monro, The puromycin reaction and its relation to protein synthesis. *Journal of Molecular Biology,* 10:63–72, 1964.)

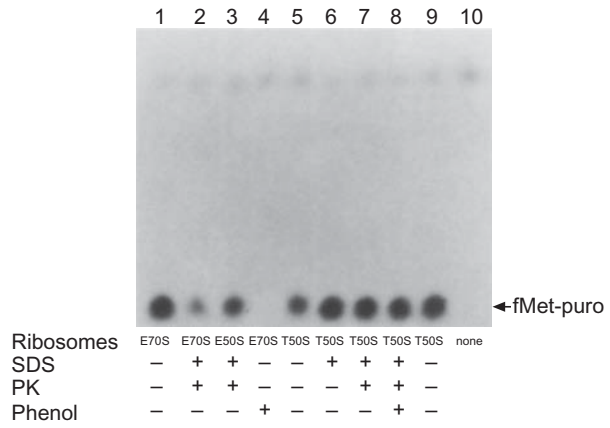


Figure 18.20 Effects of protein-removing reagents on peptidyl transferase activities of *E. coli* and *Thermus aquaticus* ribosomes. Noller and collaborators treated ribosomes with SDS, proteinase K (PK), or phenol, or combinations of these treatments, as indicated at bottom. Then they tested the treated ribosomes for peptidyl transferase by the fragment reaction using CAACCA-f[^{35}S]Met. They isolated f[^{35}S]Met-puromycin by high-voltage paper electrophoresis and detected it by autoradiography. The ribosome source is listed at bottom: E70S and E50S are 70S ribosomes and 50S ribosomal subunits, respectively, from *E. coli;* T50S refers to 50S ribosomal subunits from *Thermus aquaticus.* The position of fMet-puromycin is indicated at right. (*Source:* Adapted from Noller, H.F., V. Hoffarth, and L. Zimniak, Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* 256 (1992) p. 1417, f. 2.)

1980s that some RNAs have catalytic activity, some molecular biologists began to suspect that the 23S rRNA might actually catalyze the peptidyl transferase reaction. In 1992, Harry Noller and his coworkers presented evidence that this is so. As their assay for peptidyl transferase, they used a modification of the puromycin reaction called the **fragment reaction.** This procedure, pioneered by Monro in the 1960s, uses a fragment of labeled fMet-tRNA$_f^{Met}$ in the P site and puromycin in the A site. The fragment can be CCA-fMet, or CAACCA-fMet. Either one resembles the whole fMet-tRNA$_f^{Met}$ enough that it can bind to the P site. Then the labeled fMet can react with puromycin to release labeled fMet-puromycin.

The task facing Noller and collaborators was to show that they could remove all the protein from 50S particles, leaving only the rRNA, and that this rRNA could catalyze the fragment reaction. To remove the protein from the rRNA, these workers treated 50S subunits with three harsh agents known for their ability to denature or degrade protein: phenol, SDS, and proteinase K (PK). Figure 18.20, lanes 1–4, shows that the peptidyl transferase activity of *E. coli* 50S subunits survived SDS and proteinase K treatment, but not extraction with phenol. The ability to withstand SDS and PK was impressive, but it leaves us wondering why phenol extraction would disrupt the peptidyl transferase any more than the other two agents.

Noller and colleagues reasoned that phenol might be disrupting some higher-order RNA structure that is essential for peptidyl transferase activity. If so, they postulated that the rRNA from a thermophilic bacterium might be more sturdy and therefore might keep its native structure even after phenol extraction. To test this hypothesis, they tried the same experiment with 50S subunits from a thermophilic bacterium, *Thermus aquaticus*, that inhabits scalding hot springs. Lanes 5–9 of Figure 18.20 demonstrate that the peptidyl transferase activity of *T. aquaticus* 50S subunits survives treatment with all three of these agents.

If the fragment activity really represents peptidyl transferase, it should be blocked by peptidyl transferase inhibitors like chloramphenicol and carbomycin. Furthermore, if rRNA is a key factor in peptidyl transferase, then the fragment reaction should be inhibited by RNase. Noller and colleagues verified both of these predictions. The fragment reactions carried out by either intact or treated *T. aquaticus* 50S subunits are inhibited by carbomycin, chloramphenicol, and RNase, just as they should be.

Do these experiments show that ribosomal RNA is the only component of peptidyl transferase? Noller and coworkers stopped short of that conclusion, in part because they could not eliminate all protein from their preparations, even after vigorous treatment with protein-destroying agents. In fact, their subsequent work in collaboration with

Alexander Mankin demonstrated that eight ribosomal proteins remained associated with rRNA even after such vigorous treatment.

Mankin, Noller, and colleagues subjected *T. aquaticus* 50S ribosomal particles to the same protein-destroying agents used in Noller's original experiments. Then they performed sucrose gradient ultracentrifugation on the remaining material, and found that the material retaining peptidyl transferase activity sedimented as 50S and 80S particles, which they called *KSP50* and *KSP80* particles. The K, S, and P stand for proteinase K, SDS, and Phenol, respectively. Next, they examined intact 50S particles, as well as KSP50 and KSP80 particles, to see which RNAs and proteins they contained. They identified 23S and 5S rRNAs by gel electrophoresis. To separate and identify the protein in these particles, they used two-dimensional electrophoresis (Chapter 5). Amazingly, eight proteins remained more-or-less intact, and four of them (L2, L3, L13, and L22) were present in near-stoichiometric quantities. The other four (L15, L17, L18, and L21) were reduced in quantity. Mankin, Noller, and colleagues double-checked the identities of these eight proteins by sequencing N-terminal peptides derived from each one. Because identical proteins and RNAs appeared in both particles, it is likely that the KSP80 particles are simply dimers of KSP50 particles.

Earlier studies on reconstitution of peptidyl transferase from purified components had shown that peptidyl transferase activity could be reconstituted from just 23S rRNA and proteins L2, L3, and L4. Of these, only L4 was missing from the KSP particles. Thus, considering the reconstitution data together with the KSP particle data, Mankin, Noller, and colleagues concluded that the minimum components necessary for peptidyl transferase activity are 23S rRNA and proteins L2 and L3.

What role does 23S rRNA play in the peptidyl transferase activity? It is tempting to speculate that it has a catalytic role, but we cannot reach that conclusion based on the data presented so far. However, in 2000 Thomas Steitz and his colleagues performed x-ray crystallography studies on 50S ribosomal particles and they found no proteins—only 23S rRNA—near the peptidyl transferase active center. So it appears that 23S rRNA really does have the peptidyl transferase catalytic activity. We will examine this subject in detail in Chapter 19.

SUMMARY Peptide bonds are formed by a ribosomal enzyme called peptidyl transferase, which resides on the 50S ribosomal particle. The minimum components necessary for peptidyl transferase activity in vitro are 23S rRNA and proteins L2 and L3. X-ray crystallography studies show that 23S rRNA is at the catalytic center of peptidyl transferase and therefore appears to have peptidyl transferase activity in vivo.

Elongation Step 3: Translocation

Once the peptidyl transferase has done its job, the ribosome has a peptidyl-tRNA in the A site and a deacylated tRNA in the P site. The next step, translocation, moves the mRNA and peptidyl-tRNA one codon's length through the ribosome. This places the peptidyl-tRNA in the P site and moves the deacylated tRNA to the E site. The translocation process requires the elongation factor EF-G, which hydrolyzes GTP after translocation is complete. In this section we will examine the translocation process in more detail.

Three-Nucleotide Movement of mRNA During Translocation First of all, it certainly makes sense that translocation should move the mRNA exactly 3 nt (one codon's length) through the ribosome; any other length of movement would tend to shift the ribosome into a different reading frame, yielding aberrant protein products. But what is the evidence? Peter Lengyel and colleagues provided data in support of the 3-nt hypothesis in 1971. They created a pretranslocation complex with a phage mRNA, ribosomes, and aminoacyl-tRNAs, but left out EF-G and GTP to prevent translocation. Then they made a posttranslocation complex by adding EF-G and GTP. They treated each of these complexes with pancreatic ribonuclease to digest any mRNA not protected by the ribosome, then released the protected RNA fragment and sequenced it. They found that the sequence of the 3′-end of the fragment was UUU in the pretranslocation complex, and UUUACU in the posttranslocation complex. This indicated that translocation moved the mRNA 3 nt to the left, so three additional nucleotides (ACU) entered the ribosome and became protected. As an added check on the 3′-terminal sequences of the protected RNAs, these workers finished translating them before they released them for sequencing. They found that the protected mRNA fragment in the pretranslocation complex produced a peptide ending in phenylalanine, encoded by UUU, but the protected mRNA fragment in the posttranslocation complex produced a peptide ending in threonine, encoded by ACU. Thus, translocation had moved the mRNA exactly 3 nt, one codon's worth, through the ribosome.

SUMMARY Each translocation event moves the mRNA one codon's length, 3 nt, through the ribosome.

Role of GTP and EF-G Translocation in *E. coli* depends on GTP and a GTP-binding protein called EF-G, as we learned earlier in this chapter. In eukaryotes, a homologous protein known as **EF-2** carries out the same process. Yoshito Kaziro and colleagues demonstrated this dependence on GTP and EF-G in 1970. Then, in 1974, they amplified their findings by showing *when* during the translocation process GTP is required. First, they created the translocation substrate

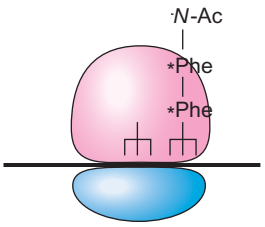


Figure 18.21 Translocation substrate used to measure dependence of translocation on EF-G and GTP. Kaziro and colleagues created a translocation substrate by loading ribosomes with *N*-acetyl-di-Phe-tRNA in the A site and a deacylated tRNA in the P site as follows: First, they mixed ribosomes and poly(U) RNA with *N*-acetyl-Phe-tRNA, which went to the P site. Then they added ordinary Phe-tRNA, which went to the A site. Peptidyl transferase then formed a peptide bond, yielding *N*-acetyl-diPhe-tRNA in the A site, and a deacylated tRNA in the P site.

pictured in Figure 18.21, with ^{14}C-labeled *N*-acetyl-diPhe-tRNA in the A site and a deacylated tRNA in the P site. This substrate is poised to undergo translocation, which can be measured in two ways: The first assay was the release of the deacylated tRNA from the ribosome. This is a non-physiological reaction. In vivo, the deacylated tRNA would simply go to the E site. The second assay for translocation was puromycin reactivity. As soon as translocation occurs, the labeled dipeptide in the P site can join with puromycin and be released. Table 18.5 shows that neither GTP nor EF-G alone caused significant translocation, but that both together did promote translocation, measured by the release of deacylated tRNA.

Table 18.5 Roles of EF-G and GTP in Translocation

Additions	tRNA released (pmol)	Δ
Experiment 1		
None	0.8	
GTP	1.8	1.0
EF-G	2.4	1.6
EF-G, GTP	12.6	11.8
EF-G, GDPCP	7.5	6.7
Experiment 2		
None	1.6	
EF-G	1.5	0
EF-G, GTP	5.1	3.5
EF-G, GTP, fusidic acid	6.7	5.1
EF-G, GDPCP	4.3	2.7
EF-G, GDPCP, fusidic acid	4.7	3.1

Source: Inove-Yokosawa, N., C. Ishikawa, and Y. Kaziro, The role of guanosine triphosphate in translocation reaction catalyzed by elongation factor G. *Journal of Biological Chemistry* 249:4322, 1974. Copyright © 1974. The American Society for Biochemistry & Molecular Biology, Bethesda, MD. Reprinted by permission.

At what point in this process is GTP hydrolyzed? There are two main possibilities: Model I calls for GTP hydrolysis pretranslocation. Model II allows for GTP hydrolysis after translocation has occurred. As unlikely as it may sound, model II was once the preferred hypothesis, based on the following experiments.

Kaziro and colleagues performed experiments with an unhydrolyzable analog of GTP, GDPCP. If GTP is not needed until after translocation, then this GTP analog should promote translocation, as natural GTP does. Table 18.5 shows that GDPCP does yield a significant amount of translocation, though not quite as much as GTP does. However, when they used GDPCP, the investigators found that they had to add stoichiometric quantities of EF-G (equimolar with the ribosomes). Ordinarily, translation requires only catalytic amounts of EF-G, because EF-G can be recycled over and over. But when GTP hydrolysis is not possible, as with GDPCP, recycling cannot occur. This suggested a function for GTP hydrolysis: release of EF-G from the ribosome, so both EF-G and ribosome can participate in another round of elongation.

Experiment 2, reported in the bottom part of Table 18.5, includes data on the effect of the antibiotic **fusidic acid.** This substance blocks the release of EF-G from the ribosome after GTP hydrolysis. This would normally greatly inhibit translation because it would halt the process after only one round of translocation. In this experiment, however, one round of translocation was all that could occur in any event, so fusidic acid had no effect. Kaziro and colleagues repeated these same experiments, using puromycin reactivity as their assay for translocation, and obtained essentially the same results. They also tried GDP in place of GTP and found that it could not support translocation.

Kaziro and colleagues concluded that GTP hydrolysis is not absolutely required for translocation (although it did help). Therefore, they reasoned, GTP hydrolysis must follow translocation. But their assays took several minutes, much longer than the millisecond time scale at which translation reactions take place. So they could not measure GTP hydrolysis and translocation and really tell which one happened first. To answer the question rigorously, we need a **kinetic experiment** that can measure events from one millisecond to the next. In 1997, Wolfgang Wintermeyer and colleagues performed such kinetic experiments and showed conclusively that GTP hydrolysis is very rapid and occurs before translocation.

Part of these workers' experimental plan was to load pretranslocation ribosomes in vitro with a fluorescent peptidyl-tRNA in the A site and a deacylated tRNA in the P site. Then they added EF-G–GTP and instantly began measuring the fluorescence of the complex. Such kinetic experiments on millisecond (ms) time scales are possible using a **stopped-flow apparatus** in which two or more

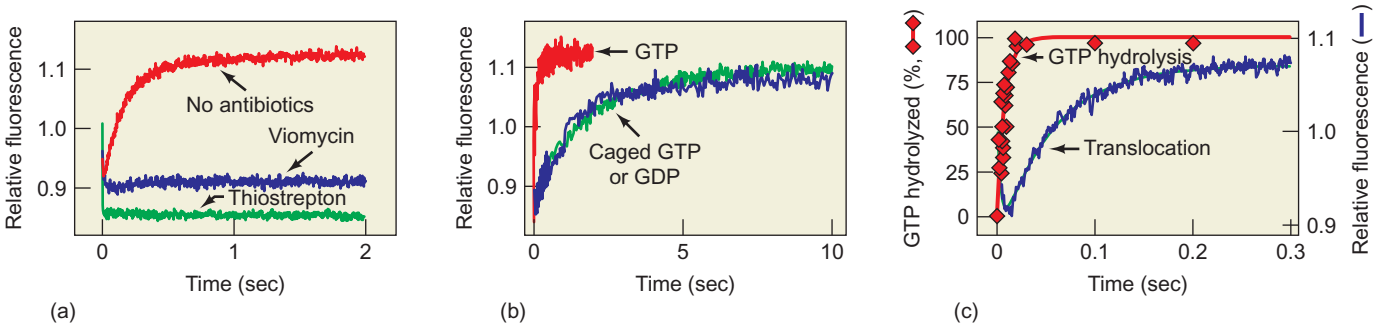


Figure 18.22 Kinetics of translocation. Wintermeyer and colleagues used stopped-flow kinetic experiments to measure translocation. They plotted the relative fluorescence of a fluorescent derivative of fMet-Phe-tRNA^{Phe} bound to the A site as a function of time in seconds. The rise in fluorescence was taken as a measure of translocation. **(a)** Effect of antitranslocation antibiotics as follows: red, no antibiotics; blue, viomycin; green, thiostrepton. **(b)** Effect of GTP analogs. The following GTP analogs were added to the translocation reaction: red, GTP; blue, an unhydrolyzable GTP analog (caged GTP); green, GDP. **(c)** Timing of GTP hydrolysis and translocation. Wintermeyer and colleagues measured translocation by stopped-flow kinetics as in panels (a) and (b) and GTP hydrolysis by release of $^{32}P_j$ from [^{32}P]GTP in a stopped-flow device. GTP hydrolysis occurs first, and about five times faster than translocation. (*Source:* Adapted from (a) Rodnina, M.V., A. Savelsbergh, V.I. Katunin, and W. Wintermeyer, Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature* 385 (2 Jan 1997) f. 1, p. 37. (b) f. 1, p. 37. (c) f. 2, p. 38.)

solutions are forced simultaneously into a mixing chamber, and then immediately into another chamber for analysis. The mixing time in these experiments is of the order of only 2 ms. After an initial drop, the fluorescence increased significantly, as shown in Figure 18.22a, red trace. This increase in fluorescence appears to be related to translocation, because it is prevented by two antibiotics that block translocation, viomycin and thiostrepton (Figure 18.22a, blue and green traces, respectively). Translocation worked much better with GTP (Figure 18.22b, red trace) than with an unhydrolyzable GTP analog (a "caged" GTP, Figure 18.22b, blue trace) or with GDP (Figure 18.22b, green trace).

Next, Wintermeyer and colleagues compared the timing and speed of GTP hydrolysis and translocation. They measured GTP hydrolysis with [γ-^{32}P]GTP, again with a stopped-flow device. This time, they rapidly mixed the radioactive GTP with the other components and then, after only milliseconds, forced the mixture into another chamber where the reaction was stopped with perchlorate solution. They measured $^{32}P_i$ released by liquid scintillation counting. Again, they assayed translocation by fluorescence increase. Figure 18.22c shows that GTP hydrolysis occurred first, and about five times faster than translocation. Thus, Wintermeyer and colleagues concluded that GTP hydrolysis precedes and drives translocation.

It is clear that EF-G, using energy from GTP, catalyzes the translocation process. Does that mean that no translocation can occur in the absence of EF-G? Actually, certain in vitro conditions have been found to allow some translocation even in the absence of EF-G. In 2003, Kurt Fredrick and Harry Noller performed the most convincing study to date on this topic, demonstrating that the antibiotic sparsomycin can catalyze translocation in the absence of EF-G and GTP. This finding suggests that the ribosome itself has the ability to perform translocation even without help from EF-G, and that the energy required for translocation is stored in the complex of ribosome, tRNAs, and mRNA after each peptide bond forms.

SUMMARY GTP and EF-G are necessary for translocation, although translocation activity appears to be inherent in the ribosome and can be expressed without EF-G and GTP in vitro. GTP hydrolysis precedes translocation and significantly accelerates it. For a new round of elongation to occur, EF-G must be released from the ribosome, and that release depends on GTP hydrolysis.

G Proteins and Translation

We have now seen two examples of proteins that use hydrolysis of GTP to drive important steps in the elongation phase of translation: EF-Tu and EF-G. Recall from Chapter 17 that IF2 plays a similar role in the initiation phase. Finally, at the end of this chapter we will discover that another factor (RF3) plays the same role in translation termination.

What do all of these processes have in common? All use energy from GTP to drive molecular movements essential for translation. IF2 and EF-Tu both bring aminoacyl-tRNAs to the ribosome (IF2 transports the initiating aminoacyl-tRNA (fMet-tRNA$_f^{Met}$) to the P site of the ribosome, while EF-Tu transports the elongating aminoacyl-tRNAs to the A site of the ribosome). EF-G sponsors translocation, in which the mRNA and the peptidyl-tRNA move from the A site to the P site and the deacylated tRNA moves from the P site to the E site of the ribosome. And RF3 helps catalyze termination, in which the bond linking the finished polypeptide to the tRNA is broken and the polypeptide exits the ribosome.

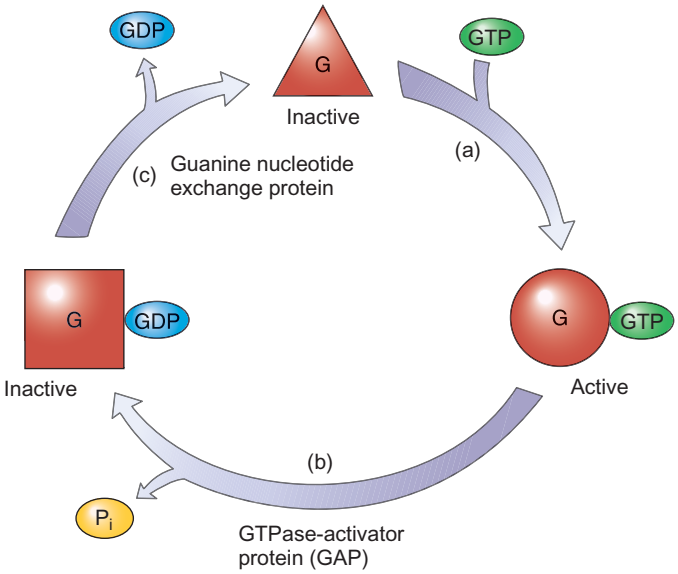


Figure 18.23 Generalized G protein cycle. The G protein at top (red triangle) is in the unbound state with neither GDP nor GTP bound. This state is normally short-lived. **(a)** GTP binds to the unbound G protein, changing its conformation (represented by the change from triangular to circular shape), and thereby activating it. **(b)** A GTPase-activator protein (GAP) stimulates the intrinsic GTPase activity of the G protein, causing it to hydrolyze its GTP to GDP. This results in another conformational change, represented by the change to square shape, which inactivates the G protein. **(c)** A guanine nucleotide exchange protein removes the GDP from the G protein, changing it back to the original unbound state, which is ready to accept another GTP.

All of these factors belong to a large class of proteins known as **G proteins** that perform a wide variety of cellular functions. Most of the G proteins share the following features, illustrated in Figure 18.23:

1. They are GDP- and GTP-binding proteins. In fact the "G" in "G protein" comes from "guanine nucleotide."
2. They cycle among three conformational states, depending on whether they are bound to GDP, GTP, or neither nucleotide, and these conformational states determine their activities.
3. When they are bound to GTP they are activated to carry out their functions.
4. They have intrinsic GTPase activity.
5. Their GTPase activity is stimulated by another agent called a **GTPase activator protein (GAP).**
6. When a GAP stimulates their GTPase activity, they cleave their bound GTP to GDP, inactivating themselves.
7. They are reactivated by another protein called a **guanine nucleotide exchange protein.** This factor removes GDP from the inactive G protein and allows another molecule of GTP to bind. One guanine nucleotide exchange protein comes immediately to mind: EF-Ts. We have seen that EF-Ts is essential for replacing GDP with GTP on EF-Tu.

The GTPases of all the G proteins involved in translation are stimulated by the ribosome. Thus, we might predict that the GAP for all these G proteins would be a protein or proteins at some site(s) on the ribosome. In fact, a set of ribosomal proteins and parts of a ribosomal RNA, collectively known as the **GTPase-associated site** or **GTPase center** has been discovered on the ribosome. It consists of the ribosomal protein L11, a complex of the ribosomal proteins L10 and L12, and the 23S rRNA. Note that the GTPase center merely stimulates the GTPase activity of the associated G protein; it does not have GTPase activity of its own.

The GTPase center is located on a stalk of the 50S subunit conventionally shown on the right of the ribosome, and called either the **L7/L12 stalk**, or the **L10–L12 stalk**. L7 and L12 are 50S ribosomal proteins that have identical amino acid sequences, but L7 is acetylated on its N-terminal amino group. One molecule each of L7 and L12 form a dimer that binds to the rest of the 50S particle via protein L10. *E. coli* ribosomes have two dimers of L7/L12. *Thermus thermophilus* and some other thermophilic bacterial ribosomes have three dimers of L12. The L12 molecules in these bacteria are not acetylated, but some of them are phosphorylated.

SUMMARY Several translation factors harness the energy of GTP to catalyze molecular motions. These factors belong to a large class of G proteins that are activated by GTP, have intrinsic GTPase activity that is activated by an external factor (GAP), are inactivated when they cleave their own GTP to GDP, and are reactivated by another external factor (a guanine nucleotide exchange protein) that replaces GDP with GTP.

The Structures of EF-Tu and EF-G

If EF-Tu and EF-G really bind to the same ribosomal GTPase center, then the two factors should have similar structures, just as two keys that fit the same lock must have similar shapes. X-ray crystallography studies on the two proteins have shown that this is true, with one qualification: It is actually the EF-Tu–tRNA–GTP ternary complex that has a shape very similar to that of the EF-G–GTP binary complex. This makes sense because EF-Tu binds to the ribosome as a ternary complex with tRNA and GTP, whereas EF-G binds as a binary complex with GTP only. To avoid GTP hydrolysis, the experimenters used unhydrolyzable GTP analogs, GDP in the case of EF-G, and GDPNP in the case of EF-Tu–tRNA.

Figure 18.24 depicts the three-dimensional structures of the two complexes. We can see that the lower part of the

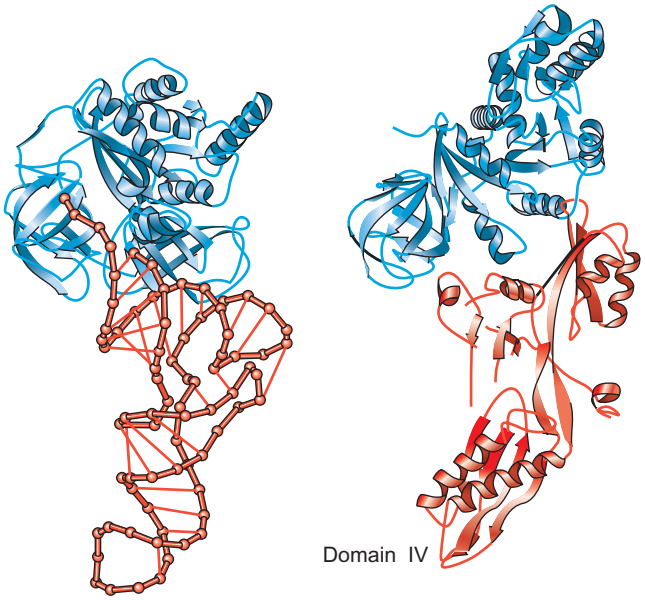


Figure 18.24 Comparison of the three-dimensional shapes of the EF-Tu–tRNA–GDPNP ternary complex (left) and the EF-G–GDP binary complex (right). The tRNA part of the ternary complex and the corresponding protein part of the binary complex are highlighted in red.

EF-G protein (**domain IV**) mimics the shape of the anticodon stem loop portion of the tRNA (red, left) in the EF-Tu ternary complex. This presumably allows both complexes to bind at or close to the same site on the ribosome.

Two other translation factors also have ribosome-dependent GTPase activities: the prokaryotic initiation factor IF2 (Chapter 17) and the termination factor RF3 (see later in this chapter). Because they also seem to rely on the same GTPase-activating center on the ribosome, it is reasonable to predict that they are structurally similar to at least parts of the two complexes depicted in Figure 18.24. Later in this chapter, we will learn that the structure of *E. coli* RF3-GDP is indeed very similar to that of EF-Tu–GTP.

Furthermore, if EF-G and IF2 bind to the same GTPase center of the ribosome, we would expect the two to compete for binding there. In fact, Albert Dahlberg and colleagues demonstrated in 2002 that IF2 does indeed compete with EF-G for ribosome binding. Moreover, they showed that two antibiotics, thiostrepton and micrococcin, that were known to bind to the GTPase center, also interfere with binding of both EF-G and IF2 at that site. Thus, IF2, EF-G, EF-Tu, and, quite probably, RF3 all bind to at least overlapping GTPase centers on the ribosome.

SUMMARY The three-dimensional shapes of the EF-Tu–tRNA–GDPNP ternary complex and the EF-G–GDP binary complex have been determined by x-ray crystallography. As predicted, they are very similar.

18.4 Termination

The elongation cycle repeats over and over, adding amino acids one at a time to the growing polypeptide product. Finally, the ribosome encounters a stop codon, signaling that it is time for the last step in translation: termination.

Termination Codons

The first termination codon (the **amber codon**) was discovered by Seymour Benzer and Sewell Champe in 1962 as a conditional mutation in a T4 phage. The amber mutation was conditional in that the mutant phage was unable to replicate in wild-type *E. coli* cells, but could replicate in a mutant, **suppressor** strain. Certain mutations in the *E. coli* alkaline phosphatase gene were also suppressed by the same suppressor strain, so it appeared that they were also **amber mutations.** We now know that amber mutations create termination codons that cause translation to stop prematurely in the middle of an mRNA, and therefore give rise to incomplete proteins. What was the evidence for this conclusion?

First of all, amber mutations have severe effects. Ordinary missense mutations change at most one amino acid in a protein, which may or may not affect the function of the protein, but even if the protein is inactive, it can usually be detected with an antibody. By contrast, *E. coli* strains with amber mutations in the alkaline phosphatase gene produce no detectable alkaline phosphatase activity or protein. This fits the hypothesis that the amber mutations caused premature termination of the alkaline phosphatase, so no full-size protein could be found.

A genetic experiment by Benzer and Champe further strengthened this hypothesis. They introduced a deletion into the adjacent *rIIA* and *B* genes of phage T4 that fused the two genes together, as shown in Figure 18.25. The fused gene gave a fusion protein with B activity, but no A activity. Then they introduced an amber mutation into the *rIIA* part of the fused gene. This mutation blocked *rIIB* activity, and this block was removed by an amber suppressor. How could a mutation in the *A* cistron block the expression of the *B* cistron, which lies downstream? Translation termination at the amber mutation is an obvious explanation. If translation stops at the amber codon, it would never reach the *B* cistron. Moreover, according to this logic, the amber suppressor overrides the translation termination at the amber codon and allows translation to continue on into the *B* cistron.

More direct evidence for the amber mutation as a translation terminator came from studies by Brenner and colleagues on the head protein gene of phage T4. When this phage infects *E. coli* B, head protein accounts for more than 50% of the protein made late in infection, which makes it easy to purify. When these investigators introduced amber mutations into the head protein gene, they were unable to isolate intact head protein from infected cells, but they could isolate fragments of head protein. And tryptic digestion of these fragments yielded peptides that could be

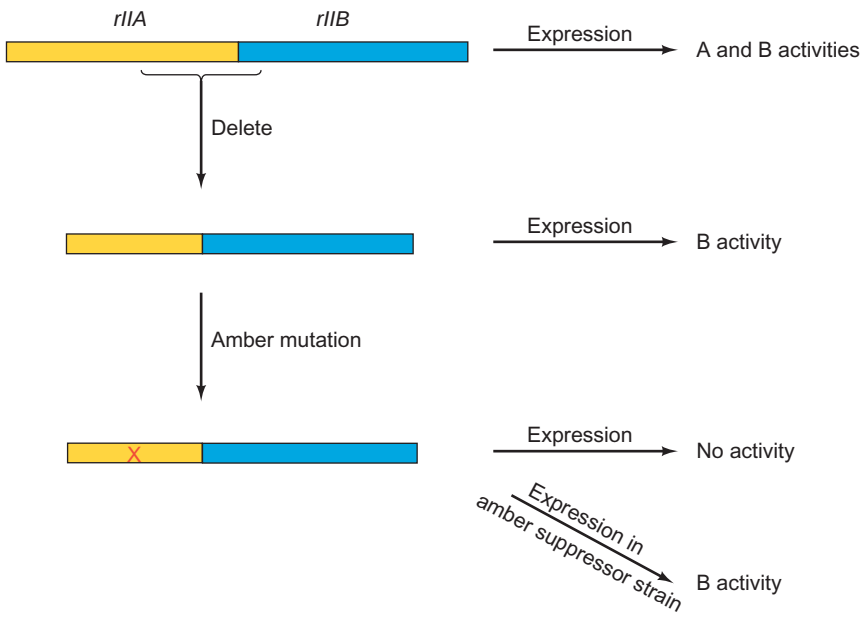


Figure 18.25 Effects of an amber mutation in a fused gene. Benzer and Champe deleted the DNA shown by the bracket, fusing the *rIIA* and *B* cistrons together. Expression of this fused gene yielded *B* activity, but no *A* activity. An amber mutation in the *A* cistron inactivated *B* activity, which could be restored by transferring the gene to an amber suppressor strain (*E. coli* CR63). The amber mutation caused premature translation termination in the *A* cistron, and the amber suppressor prevented this termination, allowing production of the *B* part of the fusion protein.

identified as amino-terminal peptides. Thus, the products of head protein genes with amber mutations were all amino-terminal protein fragments. Because translation starts at a protein's amino terminus, this experiment demonstrated that the amber mutations caused termination of translation before it had a chance to reach the carboxyl terminus.

The amber mutation defined one translation stop codon, but the two others have similarly colorful names, **ochre** and **opal.** Ochre mutations were originally distinguished by the fact that they were not suppressed by **amber suppressors.** Instead, they have their own class of **ochre suppressors.** Similarly, opal mutations are suppressed by **opal suppressors.**

How did the amber mutation get its name? In was named in honor of the mother of a graduate student named Harris Bernstein to settle a bet he made with two fellow students about the mutant they were making. He accurately predicted the properties of the mutant, so it now bears his mother's (and his) name—translated into English (German: *bernstein* = amber). Mutants that create the other two stop codons were named in the same colorful style.

Since amber mutations are caused by mutagens that give rise to missense mutations, we suspect that these mutations come from the conversion of an ordinary codon to a stop codon by a one-base change. We know that only three unassigned "nonsense" codons occur in the genetic code: UAG, UAA, and UGA. We assume these are stop codons, so the simplest explanation for the results we have seen so far is that one of these is the **amber codon,** one is the **ochre codon,** and one is the **opal codon.** But which is which?

Martin Weigert and Alan Garen answered this question in 1965, not by sequencing DNA or RNA, but by sequencing protein. They studied an amber mutation at one position in the alkaline phosphatase gene of *E. coli.* The amino acid at this position in wild-type cells was tryptophan, whose sole codon is UGG. Because the amber mutation originated with a one-base change, we already know that the amber codon is related to UGG by a one-base change. To find out what that change was, Weigert and Garen determined the amino acids inserted in this position by several different revertants. The revertants presumably arose by one-base changes from the amber codon. Some of these had tryptophan in the key position, but most had other amino acids: serine, tyrosine, leucine, glutamate, glutamine, and lysine. These other amino acids could substitute for tryptophan well enough to give at least some alkaline phosphatase activity. The puzzle is to deduce the one codon that is related by one-base changes to at least one codon for each of these amino acids, including tryptophan. Figure 18.26 demonstrates

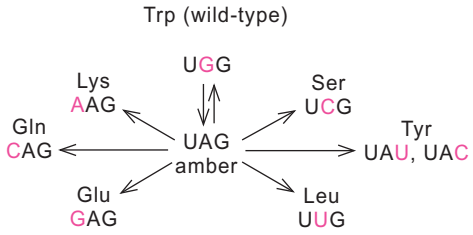


Figure 18.26 The amber codon is UAG. The amber codon (middle) came via a one-base change from the tryptophan codon (UGG), and the gene reverts to a functional condition in which one of the following amino acids replaces tryptophan: serine, tyrosine, leucine, glutamate, glutamine, or lysine. The pink color represents the single base that is changed in all these revertants, including the wild-type revertant that codes for tryptophan.

that UAG is the solution to this puzzle and therefore must be the amber codon.

By the same logic, including the fact that amber mutants can mutate by single-base changes to ochre mutants, Sydney Brenner and collaborators reasoned that the ochre codon must be UAA. Severo Ochoa and colleagues verified that UAA is a stop signal when they showed that the synthetic message $AUGUUUUAAA_n$ directed the synthesis and release of the dipeptide fMet-Phe. (AUG codes for fMet; UUU codes for Phe; and UAA codes for stop.) With UAG and UAA assigned to the amber and ochre codons, respectively, UGA must be the opal codon, by elimination. Now that we have the base sequences of thousands of genes, it is abundantly clear that these three codons really do serve as stop signals. Sometimes we even find two stop codons in a row (e.g., UAAUAG), which provides a fail-safe stop signal even if termination at one codon is suppressed.

SUMMARY Amber, ochre, and opal mutations create termination codons (UAG, UAA, and UGA, respectively) within an mRNA and thereby cause premature termination of translation. These three codons are also the natural stop signals at the ends of coding regions in mRNAs.

Stop Codon Suppression

How do suppressors overcome the lethal effects of premature termination signals? Mario Capecchi and Gary Gussin showed in 1965 that tRNA from a suppressor strain of *E. coli* could suppress an amber mutation in the coat cistron of phage R17 mRNA. This identified tRNA as the suppressor molecule, but how does it work? Brenner and collaborators found the answer when they sequenced a suppressor tRNA. They placed the gene for an amber suppressor tRNA on a ϕ80 phage and used this recombinant phage to infect *E. coli* cells bearing an amber mutation in the *lacZ* gene. Because of this suppressor tRNA, infected cells were able to suppress the amber mutation by inserting a tyrosine instead of terminating. When Brenner and colleagues sequenced this suppressor tRNA they found only one difference from the sequence of the wild-type $tRNA^{Tyr}$: a change from C to G in the first base of the anticodon, as shown in Figure 18.27.

Figure 18.28 illustrates how this altered tRNA can suppress an amber codon. We start with a codon, CAG, which encodes glutamine (Gln). It pairs with the anticodon 3′-GUC-5′ on a $tRNA^{Gln}$. Assume that the CAG codon is mutated to UAG. Now it can no longer pair with the $tRNA^{Gln}$; instead, it attracts the termination machinery to stop translation. Now a second mutation occurs in the anticodon of a $tRNA^{Tyr}$, changing it from AUG to AUC (again reading 3′→5′). This new tRNA is a suppressor tRNA because it has an anticodon complementary to the amber codon

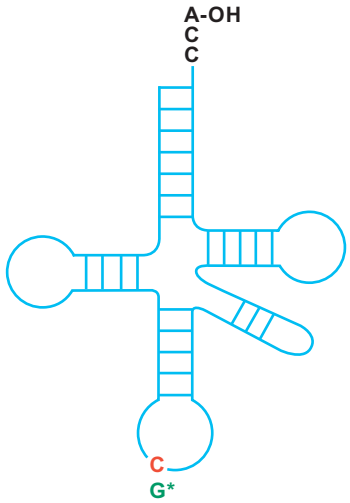


Figure 18.27 Comparison of sequence of wild-type *E. coli* $tRNA^{Tyr}$ and *E. coli* amber suppressor tRNA. The G* (green) present in the wild-type $tRNA^{Tyr}$ is replaced by a C (red) in the suppressor tRNA. (*Source:* Adapted from Goodman, H.M., J. Abelson, A. Landy, S. Brenner, and J.D. Smith, Amber suppression: A nucleotide change in the anticodon of a tyrosine transfer RNA. *Nature* 217:1021, 1968.)

UAG. Thus, it can pair with the UAG stop codon and insert tyrosine into the growing polypeptide, allowing the ribosome to get past the stop codon without terminating translation.

SUMMARY Most suppressor tRNAs have altered anticodons that can recognize stop codons and prevent termination by inserting an amino acid and allowing the ribosome to move on to the next codon.

Release Factors

Because the stop codons are triplets, just like ordinary codons, one might expect that these stop codons would be decoded by tRNAs, just as other codons are. However, work begun by Capecchi in 1967 proved that tRNAs do not ordinarily recognize stop codons. Instead, proteins called **release factors** (**RFs**) do. Capecchi devised the following scheme to identify the release factors: He began with *E. coli* ribosomes plus an R17 phage mRNA that was mutated in the seventh codon of the coat cistron to UAG (amber). The codon preceding this amber codon was ACC, which codes for threonine. He incubated the ribosomes with this mRNA in the absence of threonine so they would make a pentapeptide and then stall at the threonine codon. Then he isolated the ribosomes with the pentapeptide attached and placed them in a system containing only EF-Tu, EF-G (attached to the ribosomes) and [^{14}C]threonyl-tRNA. The ribosomes incorporated the labeled threonine into the peptide, producing a labeled hexapeptide in the P site, poised on the brink of release. To find the release factor, Capecchi added

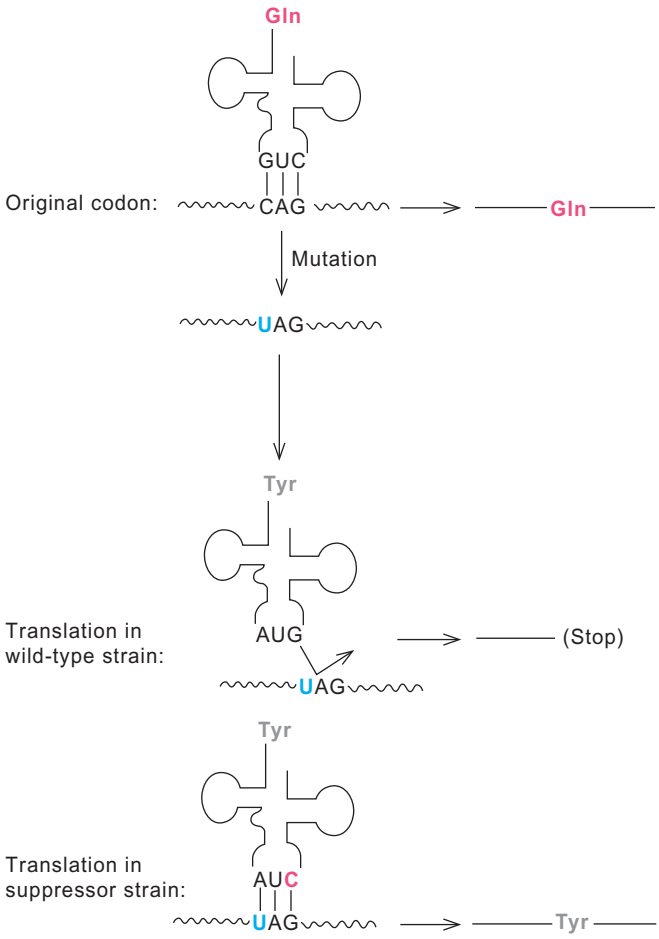


Figure 18.28 Mechanism of suppression. Top: The original codon in the wild-type *E. coli* gene was CAG, which was recognized by a glutamine tRNA. Middle: This codon mutated to UAG, which was translated as a stop codon by a wild-type strain of *E. coli.* Notice the tyrosine tRNA, whose anticodon (AUG) cannot translate the amber codon. Bottom: A suppressor strain contains a mutant tyrosine tRNA with the anticodon AUC instead of AUG. This altered anticodon recognizes the amber codon and causes the insertion of tyrosine (gray) instead of allowing termination.

ribosomal supernatant fractions until one released the labeled peptide. He discovered that this factor, which he called release factor (RF), was not a tRNA, but a protein.

Nirenberg and colleagues devised a simpler technique (Figure 18.29), which was a takeoff on their assay for identifying codons, examined earlier in this chapter. They formed a ternary complex with ribosomes, the triplet AUG, and [^{3}H]fMet-tRNA$_f^{Met}$. The initiation codon and aminoacyl-tRNA went to the P site in the complex, and the labeled amino acid was therefore eligible for release. Incubation of this complex with a crude release factor preparation and any of the three termination codons (UAG, UAA, or UGA) caused release of the labeled fMet. In this assay, the termination trinucleotide went to the A site and dictated release if the appropriate release factor was present. Table 18.6 shows that one factor (**RF1**)

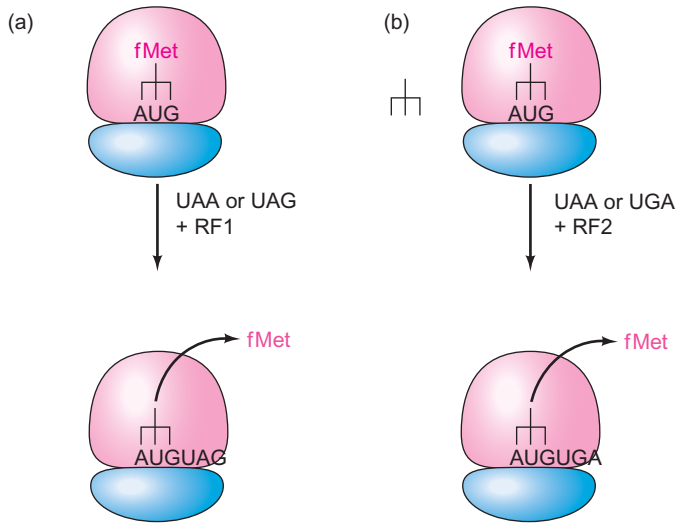


Figure 18.29 Nirenberg's assay for release factors. Nirenberg loaded the P site of ribosomes with the initiation codon AUG and [^{3}H] fMet-tRNA$_f^{Met}$. Then he added one of the termination codons plus a release factor, which released the labeled fMet. **(a)** RF1 is active with UAA or UAG. **(b)** RF2 is active with UAA or UGA.

Table 18.6 Response of RF1 and RF2 to Stop Codons

Additions		pmol [^{3}H]fMet released in presence of:	
Release factor	**Stop codon**	**0.012 M Mg^{2+}**	**0.030 M Mg^{2+}**
RF1	None	0.12	0.15
RF1	UAA	0.47	0.86
RF1	UAG	0.53	1.20
RF1	UGA	0.08	0.10
RF2	None	0.02	0.14
RF2	UAA	0.22	0.77
RF2	UAG	0.02	0.14
RF2	UGA	0.33	1.08

Source: From "Release Factors Differing in Specificity for Terminator codons," by W. Scolnick, R. Tompkins, T. Caskey, and M. Nirenberg, *Proceedings of the National Academy of Sciences, USA,* 61:772, 1968. Reprinted with permission of the authors.

cooperated with the stop codons UAA and UAG to cause release of the fMet, while another factor (**RF2**) cooperated with UAA and UGA. Subsequent studies showed that UAA or UAG could direct the binding of purified RF1 to the ribosome, while UAA or UGA could direct RF2 binding. This reinforced the idea that the RFs could recognize specific translation stop signals. A third release factor, (**RF3**), a ribosome-dependent GTPase, binds GTP, then binds to the ribosome and induces a large conformational change in the ribosome that apparently facilitates the

release of RF1 or RF2 after they have done their jobs. Based on EF-G's mimicry of the shape of EF-Tu bound to a tRNA, it was predicted that RF3 would have a structure resembling the protein part of the EF-Tu–tRNA–GTP ternary complex. In fact, the crystal structure of *E. coli* RF3-GDP is very similar to that of EF-Tu–GTP. It was further predicted that RF1 and RF2 mimic the structure of tRNA. The facts that RF1 and RF2 compete with tRNA for binding to the ribosome, recognize codons as tRNAs do, and are about the same size as tRNAs are consistent with this hypothesis. Indeed, in 2008 Harry Noller and colleagues determined the crystal structure of a complex including the 70S ribosome, RF1, and tRNA (Chapter 19). They showed that parts of RF1 really do occupy essentially the same position in the A site that an aminoacyl-tRNA normally would.

What about eukaryotic release factors? The first such factor (eRF) was discovered by a technique similar to Nirenberg's in 1971. Then, in 1994, a collaborative group led by Lev Kisselev finally purified eRF, still using an assay based on Nirenberg's, and succeeded in cloning and sequencing the eRF gene. Their approach to cloning and sequencing the gene was a widely used one: Using an fMet release assay similar to Nirenberg's to detect eRF, they purified the eRF activity until it gave one major band on SDS-PAGE, then subjected this protein to two-dimensional gel electrophoresis to purify it away from all other proteins. They cut out the eRF spot from this electrophoresis step, cleaved the protein with trypsin, and subjected four of the tryptic peptides to microsequencing. The sequences strongly resembled those of proteins from humans, *Xenopus laevis,* yeast, and the small flowering plant *Arabidopsis thaliana*. Thus, they were able to use the *Xenopus* gene (C11), which had already been cloned, as a probe to find the corresponding human gene in a human cDNA library. To verify that the products of the cloned *Xenopus* and human genes (C11 and TB3-1, respectively) had eRF activity, Kisselev and colleagues expressed these genes in bacteria or yeast, respectively, and tested them in the fMet release assay with tetranucleotides, some of which contained stop codons. Both proteins released fMet from loaded ribosomes, but only in the presence of a stop codon. The *Xenopus* protein was expressed with an oligohistidine (His) tag, so Kisselev and colleagues included unrelated His-tagged proteins as negative controls. They also showed that an antibody against C11 blocked its release factor activity, but an irrelevant antibody (anti-Eg5) did not.

Furthermore, eRF can recognize all three stop codons, unlike either of the two prokaryotic release factors, which can recognize only two. Does eRF collaborate with a G protein as prokaryotic RF1 and RF2 do? Michel Philippe and colleagues found that the answer is yes when they discovered a protein factor, now called **eRF3,** in *X. laevis* cells in 1995. Another member of the eRF3 family, a yeast protein known as Sup35, has a guanine nucleotide-binding domain and is essential for yeast growth. With the discovery of eRF3, eRF has been renamed **eRF1.** Interestingly, the function of eRF3 is much different from that of bacterial RF3. It collaborates with eRF1 both in recognizing the three stop codons, and in releasing the finished polypeptide from the ribosome.

SUMMARY Prokaryotic translation termination is mediated by three factors: RF1, RF2, and RF3. RF1 recognizes the termination codons UAA and UAG; RF2 recognizes UAA and UGA. RF3 is a GTP-binding protein that facilitates release of RF1 and RF2 from the ribosome. Eukaryotes have two release factors: eRF1, which recognizes all three termination codons, and eRF3, a ribosome-dependent GTPase that helps eRF1 recognize stop codons and release the finished polypeptide.

Dealing with Aberrant Termination

Two kinds of aberrant mRNAs can lead to aberrant termination. First, as we have seen, "nonsense" mutations can occur that cause premature termination. Second, some mRNAs (**non-stop mRNAs**) lack termination codons, sometimes because the synthesis of the mRNA was aborted upstream of the termination codon. Ribosomes translate through these non-stop mRNAs and then stall. Both of these events cause problems for the cell. Either premature termination or a stalled ribosome yields incomplete proteins that might have adverse effects on the cell. Stalled ribosomes present a cell with the additional problem that the stalled ribosome is out of action and unable to participate in any further protein synthesis. Let us first examine the ways that cells deal with non-stop mRNAs, then we will look at mechanisms for degrading the products of premature termination.

Non-Stop mRNAs To deal with non-stop mRNAs, cells need to degrade the aberrant protein product and release the ribosomal subunits so they can participate in productive translation instead of remaining stalled forever. The mechanisms of this process differ between bacteria and eukaryotes. Bacteria use so called **transfer-messenger RNAs (tmRNAs)** to rescue stalled ribosomes and tag the non-stop mRNAs for destruction (**tmRNA-mediated ribosome rescue**). The tmRNAs are about 300 nt long, and their 5′- and 3′-ends come together to form a **tRNA-like domain (TLD)** that resembles a tRNA (Figure 18.30). In fact, the resemblance is so strong that a tmRNA can be charged with alanine. Once charged, the alanyl-tmRNA can bind to the ribosome's A site and, via the ribosome's peptidyl transferase, can donate its alanine to the stalled polypeptide.

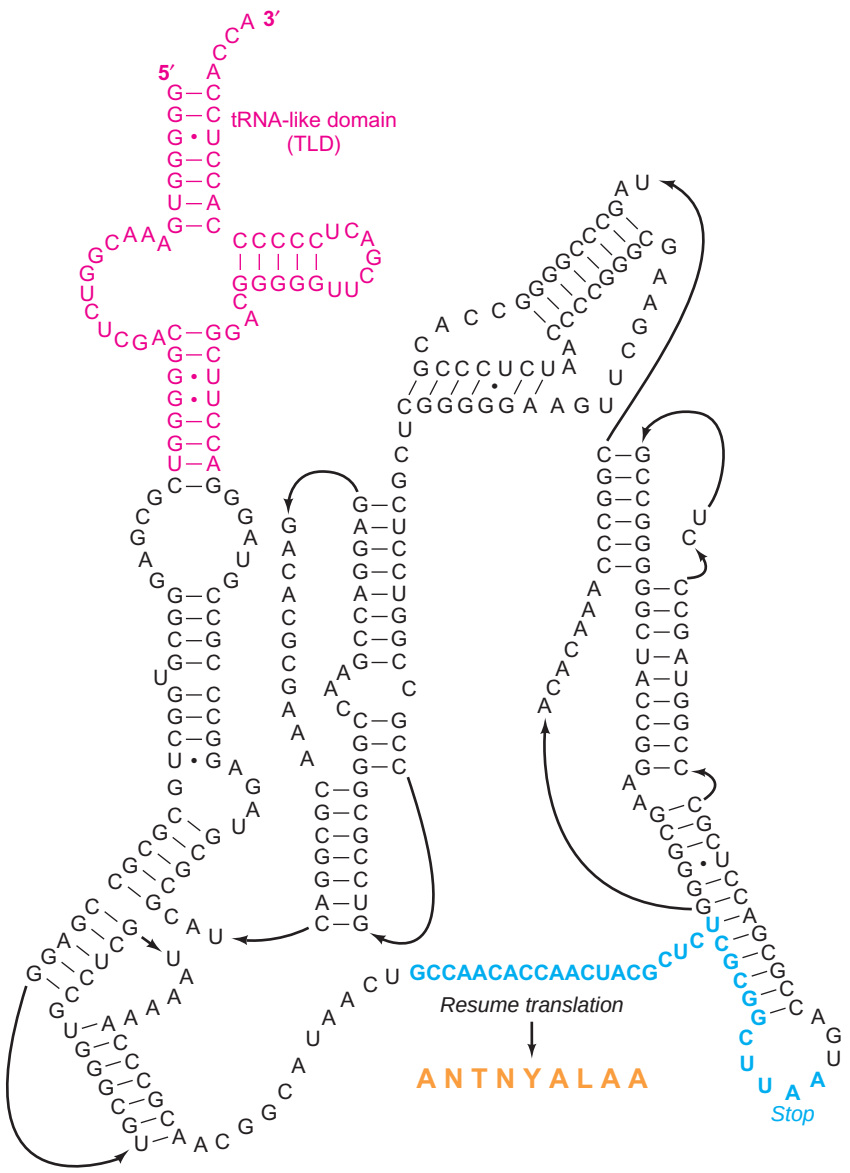


Figure 18.30 Structure of the *Thermus thermophilus* tmRNA. The TLD is at upper left in pink, and the ORF is at bottom in blue. The peptide encoded by the ORF is in orange. (*Source*: Adapted from Valle et al., Visualizing tmRNA entry into a stalled ribosome, *Science* 300:128, fig.1, 2003.)

After this peptidyl transferase reaction, the central part of the tmRNA comes into play (Figure 18.31). This part of the tmRNA contains a short open reading frame (ORF) that is positioned in the A site such that the ribosome switches from translating the non-stop mRNA to translating the tmRNA, a process called ***trans*-translation.** The ORF of the tmRNA encodes a short, hydrophobic peptide that is added to the carboxyl terminus of the stalled polypeptide. This peptide targets the whole polypeptide for destruction, minimizing its ability to harm the cell.

Obviously, a tmRNA is not just like a tRNA. For one thing, it lacks an anticodon, so there can be no codon–anticodon pairing. And, as we have seen, codon–anticodon pairing is essential to avoid dissociation of an aminoacyl-tRNA during proofreading. A second difference between a tmRNA and a real tRNA is that the tmRNA does not have a standard D loop. But the tmRNA systems gets around these problems using a protein known as **SmpB.** In 2003, Joachim Frank and V. Ramakrishnan obtained cryo-electron microscopy images of a complex of EF-Tu, tmRNA, and SmpB bound to ribosomes from *Thermus thermophilus*. This study showed that SmpB binds to tmRNA and EF-Tu and makes contacts with the ribosome that would normally come from the D loop of an RNA. Thus, SmpB helps to hold the tmRNA to the ribosome even though the tmRNA lacks some of the elements it needs to bind tightly by itself.

What happens to the non-stop mRNA once the ribosome has been released by tmRNA? We do not know the answer for sure, but tmRNAs do copurify with a 3′→5′ exonuclease known as **RNase R.** It is an attractive hypothesis

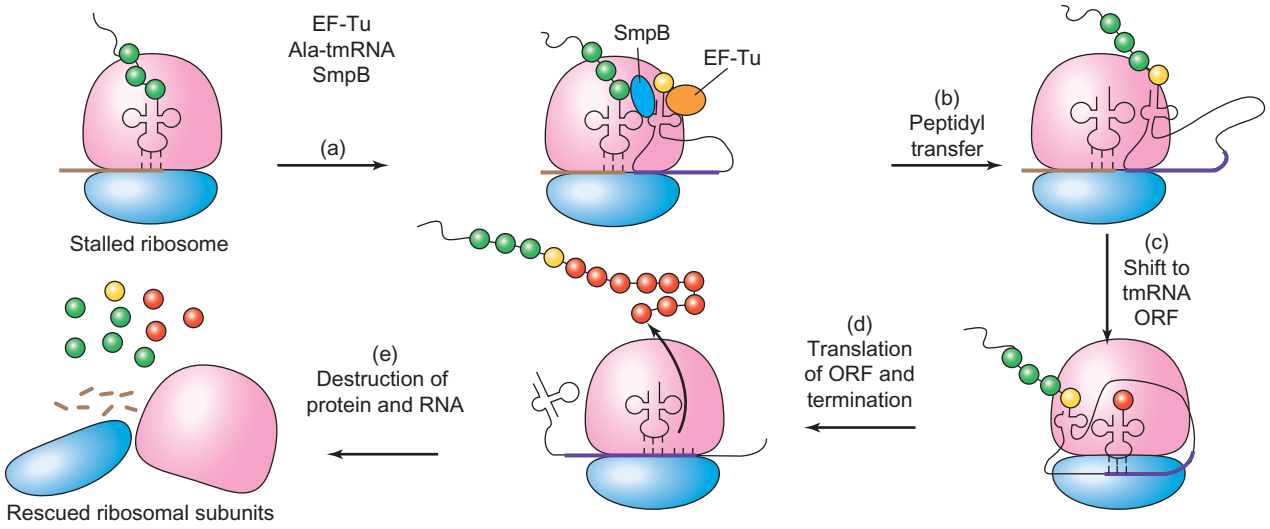


Figure 18.31 Mechanism of tmRNA-mediated release of non-stop mRNA and polypeptide. **(a)** EF-Tu, alanyl-tmRNA and SmpB (turquoise) bind to the A site of the ribosome stalled on a non-stop mRNA (brown). SmpB helps the tRNA-like domain of the tmRNA bind to the ribosome. **(b)** The ribosome's peptidyl transferase transfers the alanine (yellow) from the tmRNA to the stalled polypeptide (green). **(c)** The ribosome shifts to reading the ORF (purple) of the tmRNA. **(d)** The ribosome completes translating the ORF of the tmRNA, adding nine more amino acids (red) to the end of the stalled polypeptide and releasing it. **(e)** Together, these extra amino acids target the whole polypeptide for destruction. At the same time, the non-stop mRNA is destroyed, perhaps by RNase R, which associates with tmRNA. (*Source:* Adapted from Moore, S.D., K.E. McGinness, and R.T. Sauer, A glimpse into tmRNA-mediated ribosome rescue. *Science* 300 [2003] p. 73, f. 1.)

that RNase R degrades the non-stop mRNA before it can complex with a new ribosome.

Eukaryotes do not have tmRNAs, so how do they deal with non-stop mRNAs? Figure 18.32 illustrates the current hypothesis. The A site of a ribosome stalled at the end of a nonstop mRNA will contain zero to three nucleotides of the terminal poly(A). This state is recognized by the carboxyl-terminal domain of a protein called **Ski7p.** This protein domain resembles the GTPase domains of the elongation and termination factors EF1A and eRF3, respectively. These domains normally associate with the ribosomal A site, and so does Ski7p. In addition, Ski7p associates tightly with the cytoplasmic **exosome,** a complex of 9-11 proteins, including a 3′→5′ exonuclease that degrades RNA. The Ski7p–exosome complex then recruits the **Ski complex** to the ribosomal A site, adajacent to the end of the non-stop mRNA. Finally, the exosome degrades the non-stop mRNA. (in a process known as **non-stop decay** [**NSD**])

SUMMARY Prokaryotes deal with non-stop mRNAs by tmRNA-mediated ribosome rescue. An alanyl-tmRNA, which resembles an alanyl-tRNA, binds to the vacant A site of a ribosome stalled on a non-stop mRNA, and donates its alanine to the stalled polypeptide. Then the ribosome shifts to translating an ORF on the tmRNA, adding another nine amino acids to the polypeptide before terminating. These extra amino acids target the polypeptide for destruction, and a nuclease destroys the non-stop mRNA. Eukaryotic ribosomes at the end of the poly(A) tail of a non-stop mRNA recruit the Ski7p–exosome complex to the vacant A site. Next, the Ski complex is recruited to the A site, and the exosome, positioned just at the end of the non-stop mRNA, degrades that RNA. The aberrant polypeptide is presumably also destroyed.

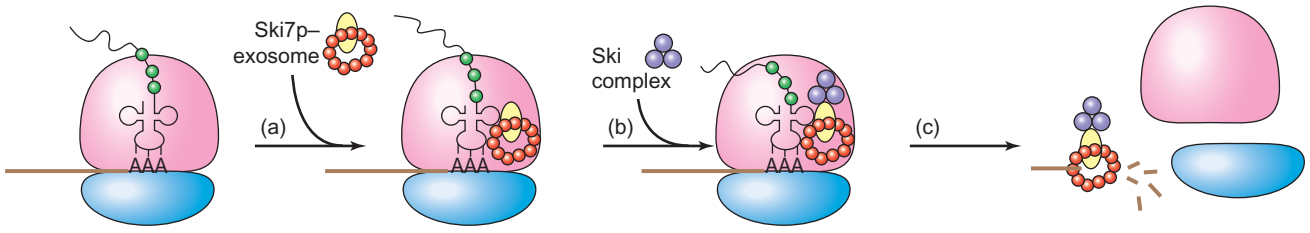


Figure 18.32 Model for exosome-mediated degradation of eukaryotic non-stop mRNA. **(a)** The A site of a ribosome stalled at the end of a non-stop mRNA (brown) contains zero to three nucleotides of the mRNA's poly(A) tail. Here, no A's are in the A site. This state of the ribosome is attractive to the Ski7p–exosome complex (yellow and red), which binds to the vacant A site. **(b)** Next, the Ski complex (purple) binds to the A site, and **(c)** this triggers degradation of the non-stop mRNA and release of the ribosomal subunits.

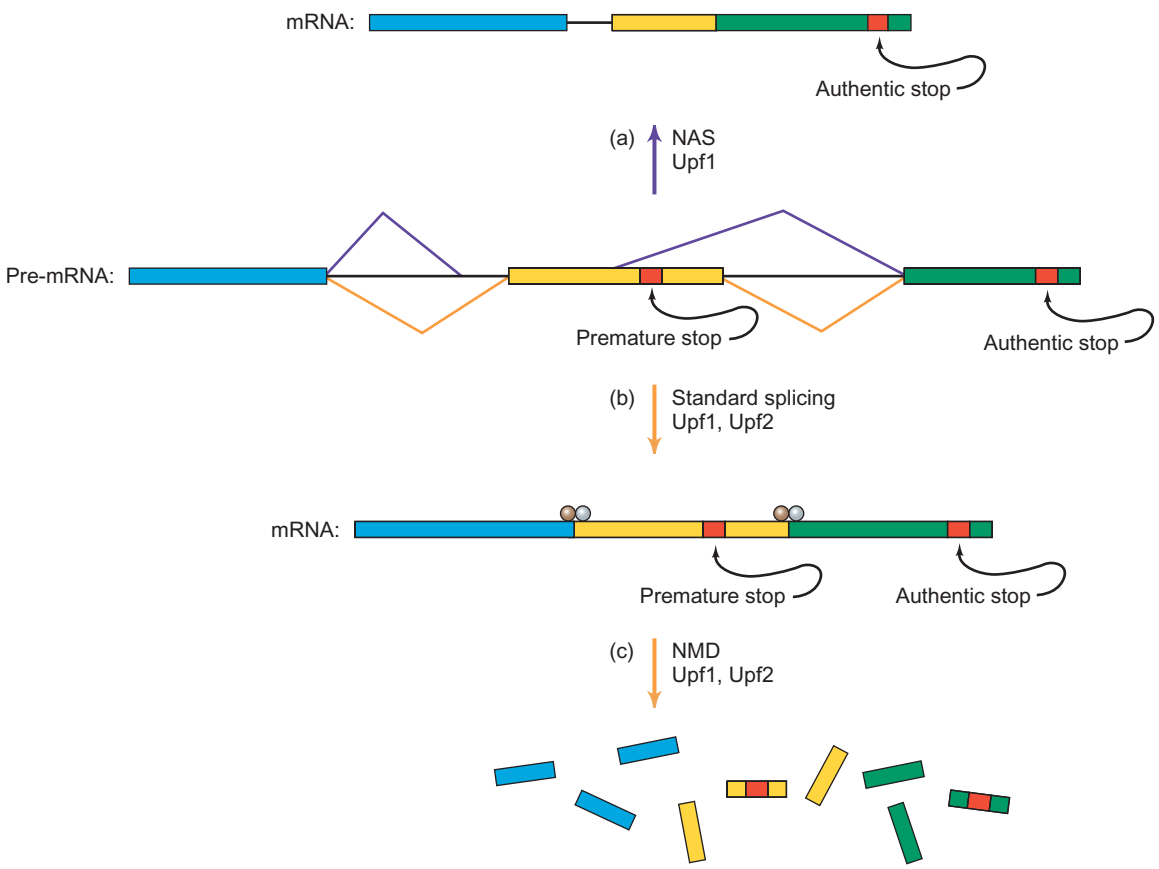


Figure 18.33 Models for NAS and NMD. (a) NAS. Upf1, perhaps in conjunction with other proteins, senses a premature stop codon in the reading frame of the future mRNA and induces an alternative splicing pattern (purple) to produce the mature mRNA at top, which lacks the premature stop codon. **(b)** Standard splicing (orange) produces a mature mRNA with a premature stop codon, and Upf1 and Upf2 bound at the exon/exon boundaries. **(c)** NMD. Upf1 and Upf2 (brown and gray), perhaps in conjunction with other proteins, sense the in-frame premature stop codon too close to the second exon/exon boundary and induce destruction of the mRNA.

Premature Termination Messenger RNAs with premature termination codons (nonsense codons) also give rise to aberrant, truncated protein products that are potentially harmful to the cell. Eukaryotic cells have evolved two ways of dealing with this problem (Figure 18.33): **nonsense-mediated mRNA decay (NMD) and nonsense-associated altered splicing (NAS).**

NMD depends on indentifying a stop codon as premature (a **premature termination codon** [**PTC**]). Obviously, there is an authentic stop codon at the end of every mRNA, and the cell must somehow discriminate between authentic and premature stop codons. Mammalian cells do this by measuring the distance between the stop codon and the **exon junction complex** (**EJC**) during the pioneer round of translation. (The EJC is a collection of proteins deposited about 20 to 25 nt upstream of exon-exon junctions at the time of splicing. If the distance between the stop codon and the EJC is short (less than about 55 nt), the stop codon is likely to be authentic, but if it is longer than about 55 nt, the stop codon is likely to be premature.

Two of the EJC proteins that are active in mammalian T cells are **Upf1** and **Upf2.** If either of these proteins is removed from a cell by RNAi (Chapter 16), NMD is inhibited. When these proteins are bound to an mRNA at a sufficiently long distance downstream of a stop codon, they recognize the stop codon as premature and activate the NMD process. On the other hand, if these proteins are relatively close to the stop codon, they are simply removed by the ribosome translating the mRNA in the pioneer round.

Lynne Macquat and colleagues presented data in 2008 that further illuminated the role of Upf1 in human NMD. They found that when translation terminates prematurely at a PTC, Upf1 binds to the downstream EJC and becomes phosphorylated. Phospho-Upf1 then binds to eIF3 and prevents the eIF3-dependent conversion of the 48S initiation complex to the 80S initiation complex that is competent to begin translation. Thus, translation is repressed, and the PTC-bearing mRNA is degraded, probably in P bodies (Chapter 16). If this model, which critically involves eIF3,

is correct, then eIF3-independent translation should not exhibit NMD. Indeed, Macquat and colleagues found that eIF3-independent translation of cricket paralysis virus (CrPV) mRNA is not subject to NMD.

In contrast to the model just described, Elisa Izaurralde and colleagues reported in 2003 that the components of the EJC are not required for NMD in *Drosophila* cells, raising the possibility that the mechanism of NMD varies from one class of organisms to another. Then, in 2004, Allan Jacobson and colleagues reported on an investigation of NMD in yeast, showing that the mechanism of premature termination is itself aberrant.

In particular, Jacobson and colleagues used a toeprinting assay (Chapter 17) to show that ribosomes, once they had terminated prematurely, did not dissociate from the mRNA, but moved upstream to a start codon (AUG). This behavior could be blocked by removing the yeast Upf1 protein, or by placing a normal 3′-UTR near the premature stop codon. Furthermore, an mRNA containing a premature stop codon could be stabilized by tethering a poly(A)-binding protein (Pab1p) to the mRNA. All these findings support a model in which the ribosome recognizes a normal stop codon by its context near a 3′-UTR, or near a poly(A), and terminates normally. By contrast, the ribosome recognizes a premature stop codon as aberrant by its remoteness from these normal cues, and terminates abnormally by going back to an upstream AUG. In principle, any eukaryotic cell should be able to recognize this unusual termination and degrade the associated mRNA, but it is not yet clear how uniform the NMD mechanism is in eukaryotes.

NAS is more mysterious than NMD. When the NAS machinery detects an in-frame (but not an out-of-frame) premature stop codon, it causes the splicing apparatus to splice the pre-mRNA in an alternative way that eliminates the premature stop codon from the mature mRNA. But that scheme raises a very intriguing question: How does the NAS machinery detect the future reading frame before the pre-mRNA is even spliced?

So far, we have no answer to that question, but we do know that one of the essential players in NAS is also one of the key agents in NMD: Upf1. Harry Dietz and colleagues used RNAi to show that Upf1, but not Upf2, is required for NAS. Then they refined their technique to ask whether the same parts of Upf1 are required for both NMD and NAS. To do this, they used **allele-specific RNAi** as follows: They made an altered *Upf1* gene that was not subject to RNAi caused by the double-stranded RNA that blocks expression of the endogenous gene. Then they introduced this altered gene, on a plasmid, into cells experiencing RNAi directed at the endogenous *Upf1* gene. The altered gene could rescue both NAS and NMD, which would otherwise have been blocked due to loss of *Upf1* expression.

Next, Dietz and colleagues made mutations to conserved regions of the altered *Upf1* gene. One of these mutations knocked out the ability of the altered gene to rescue NMD, but had no effect on the ability to rescue NAS. Thus, although NMD and NAS both depend on Upf1, they apparently rely on different functions of the protein.

SUMMARY Eukaryotes deal with premature termination codons by two different mechanisms: NMD and NAS. NMD in mammalian cells relies on the ribosome during the pioneer round to measure the distance between the stop codon and the EJC. If it is too long, the mRNA is destroyed. In yeast, the cell appears to recognize a premature stop codon by the absence of a normal 3′-UTR or poly(A) nearby. When a ribosome stops at a premature stop codon, it moves to an upstream AUG, and this may mark the mRNA for destruction. The NAS machinery senses a stop codon in the middle of a reading frame and changes the splicing pattern such that the premature stop codon is spliced out of the mature mRNA. Like NMD, this process also requires Upf1.

No-go Decay In 2006, Meenakshi Doma and Roy Parker identified another kind of mRNA decay, which they dubbed "**no-go decay (NGD).**" They artificially induced a ribosome stall by creating an mRNA with a very stable stem-loop that the ribosome was incapable of traversing. Yeast cells degraded this mRNA faster than they did the wild-type mRNA lacking the stem-loop.

Doma and Parker found that this accelerated decay occurred in cells that were deficient in either decapping or 3′→5′ exonucleases, which are key elements of the usual 5′→3′ and 3′→5′ decay, respectively, in yeast. And they found that decay is also accelerated in cells defective in NMD because of a mutation in *Upf1*.

If decay is not happening by the usual pathways, how is it accomplished? Doma and Parker showed that the no-go mRNA was cleaved by an endonuclease at a site near the stable stem-loop that had stalled the ribosome. This cut within the mRNA created new 3′- and 5′-ends that are substrates for degradation by the usual 3′- and 5′-endonucleases.

Natural mRNAs are not likely to contain stable stem-loops that arrest ribosomes, so no-go decay probably acts on ribosomes that are stalled because of natural causes such as defective mRNAs or ribosomes. It also provides another potential means of post-transcriptional control by selective degradation of mRNAs.

SUMMARY Stalled ribosomes can trigger no-go decay of mRNA, which begins with an endonucleolytic cleavage near the stalled ribosome.

Use of Stop Codons to Insert Unusual Amino Acids

Most proteins contain only the 20 amino acids pictured in Figure 3.2. However, a few proteins require unusual amino acids. The first unusual amino acids to be discovered, such as hydroxyproline, were found to arise through posttranslational modification of proteins made from the standard 20 amino acids. More recently, other unusual amino acids, such as selenocysteine and pyrrolysine, have been shown to be incorporated directly into growing polypeptides. In these cases, mechanisms have evolved to take advantage of stop codons in the middle of coding regions. Cells interpret these stop codons, not as termination signals, but as codons for unusual amino acids.

The first unusual amino acid discovered in proteins (the "21st amino acid") was **selenocysteine,** which looks just like cysteine except that it has a selenium atom in place of the sulfur atom. Some enzymes, such as glutathione peroxidase and formate dehydrogenase, do not work without selenocysteine. Each requires a single selenocysteine residue as part of its active site. But how can this unusual amino acid be incorporated into proteins? The genes that encode these enzymes produce mRNAs with UGA stop codons in the positions where selenocysteine is needed. Furthermore, in the absence of selenium, translation stops prematurely at these stop codons. These findings suggest that the cell somehow interprets these UGA codons as selenocysteine codons. But how?

A special tRNA with an anticodon that recognizes the UGA stop codon can be charged with serine by a normal seryl-tRNA synthetase. Then, the serine in this special seryl-tRNA is converted to selenocysteine. A special EF-Tu can then deliver this altered aminoacyl-tRNA to the ribosome in response to the UGA codon in the middle of the mRNA—but not to UGA codons at the ends of coding regions. If the latter were the case, selenocysteines would be incorporated in response to authentic stop codons, hindering termination.

Thus, the UGA codons within an mRNA are only part of the signal that recruits the selenocysteinyl-tRNA. Other parts of the mRNA must also play a role. In the case of the formate dehydrogenase mRNA, this is a region about 40 nt downstream of the internal UGA, and in another mRNA, it is a region about 1000 nt downstream, in the 3′-untranslated region of the mRNA. Such an mRNA region, which dictates that a UGA codon should be recognized as a selenocysteine (Sec) codon, is called a **Sec insertion sequence,** or **SECIS**. A SECIS is a stem-loop in the mRNA with three short conserved motifs. These conserved sequences are clearly important, because mutations within them prevent selenocysteine incorporation.

The "22nd amino acid" is **pyrrolysine,** which has the structure shown in Figure 18.34. Unlike selenocysteine, which is widespread, pyrrolysine has so far been found only in certain methanogenic (methane-producing) archaea. Also

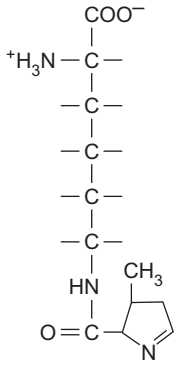


Figure 18.34 Pyrrolysine.

unlike selenocysteine, which is built from a normal amino acid (serine) on seryl-tRNA, pyrrolysine is first synthesized and then added to a special tRNA by a special pyrrolysyl-tRNA synthetase. This is the 21st aminoacyl-tRNA synthetase ever found—the only one aside from the 20 that charge normal tRNAs with the 20 normal amino acids.

E. coli cells cannot normally incorporate pyrrolysine into their proteins. But Joseph Krzycki and colleagues showed in 2004 that they could endow *E. coli* cells with the ability to do this incorporation if they added three things to the cells: a gene for the special tRNA, a gene for the special pyrrolysyl-tRNA synthetase, and pyrrolysine itself. Furthermore, they showed that the tRNA can accept preformed pyrrolysine in vitro, strongly suggesting that this is the way it works in vivo.

As is the case with selenocysteine, pyrrolysine is incorporated into growing polypeptides in response to a stop codon, but it is the UAG codon instead of UGA. This implies that the anticodon of the special tRNA is 5′-CUA-3′, and that is indeed the case.

SUMMARY The unusual amino acids selenocysteine and pyrrolysine are incorporated into growing polypeptides in response to the termination codons UGA and UAG, respectively, as follows: (1) Selenocysteine: A special tRNA (with an anticodon that recognizes the UGA codon) is charged with serine, which is then converted to selenocysteine, and the selenocysteyl-tRNA is escorted to the ribosome by a special EF-Tu. (2) Pyrrolysine: A special pyrrolysyl-tRNA synthetase joins preformed pyrrolysine with a special tRNA that has an anticodon that recognizes the codon UAG.

18.5 Posttranslation

The story of translation does not end with termination. Proteins must fold properly and ribosomes need to be released from the mRNA so they can engage in further

rounds of translation. Strictly speaking, the first of these processes does not occur after translation; rather, it is a cotranslational event that occurs as the nascent polypeptide is being made. However it is convenient to deal with it separately, as it has no direct relationship to the initiation, elongation, and termination events we have been discussing. Let us consider the folding problem first, then the ribosomal release problem.

Folding Nascent Proteins

Native proteins are folded so that any hydrophobic (Greek: "water-fearing") regions are buried in the interiors of the proteins, away from the aqueous environment in the cell. But most proteins do not fold into their proper shapes by themselves. They need help from molecular chaperones, just as proteins that have been unfolded by heat shock do (Chapter 8). The problem is that any exposed hydrophobic sections of a nascent polypeptide would try to interact with any other exposed hydrophobic regions they could find, to hide from the water surrounding them. But the nearest hydrophobic region is likely to be the wrong partner, so that interaction would lead to a misfolded and therefore inactive protein. In fact, some misfolded proteins, such as the one involved in bovine spongiform encephalopathy (BSE, or "mad cow disease") can be deadly toxic to a cell.

Here is another example of the importance of proper protein folding: Silent mutations occur when a codon for an amino acid is changed into another codon for that same amino acid. Ordinarily, such mutations have no effect, which is why we call them silent. Occasionally, however, "silent" mutations can actually cause problems. This has been documented to occur in several ways: The change of one codon for an amino acid to another codon for the same amino acid sounds harmless, but if the new codon is much rarer for that organism (a phenomenon known as **codon bias**), the corresponding tRNA is probably also rare, so the ribosome slows down at that codon waiting for the rare aminoacyl-tRNA to appear. Some proteins fold differently depending on their rate of synthesis, so slowing down translation while waiting for a rare aminoacyl-tRNA can cause misfolding, and perhaps inactivation, of the protein product. Michael Gottesman and colleagues demonstrated in 2007 that a mutation in the human *multidrug resistant* 1 (*MDR*1) gene, though it is a "silent" mutation, creates a rare codon and yields a product with altered, and less effective, activity, presumably because of misfolding.

On the other hand, ribosomal pausing between domains (independently folded parts) of a protein can be beneficial because it allows these domains to fold without interference from irrelevant other parts of the protein. Thus, it was intriguing that Joseph Watts, Kevin Weeks, and their colleagues showed in 2009 that the HIV (human immunodeficiency virus) RNA, which serves as both genome and mRNA, has its highest levels of secondary structure in the regions of the mRNAs that encode loops between protein domains. These regions of secondary structure (intramolecular base-pairing) would presumably impede the progress of the ribosome and allow the recently completed protein domain to fold before beginning the synthesis of the next domain.

To probe the secondary structure of the HIV RNA, Watts and Weeks and their colleagues used a technique known as **selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE).** This method relies on the fact that certain reagents, such as 1-methyl-7-nitroisatoic anhydride (1M7), selectively acylate the 2′-hydroxyl groups of RNA nucleotides that are conformationally flexible. Nucleotides that are base-paired are rigid and relatively protected from acylation. After reacting the RNA with 1M7, the investigators subjected it to primer extension (Chapter 5) with reverse transcriptase and fluorescent primers. Then they analyzed the lengths of the extended primers to locate regions of base-pairing, where the primer extension tends to stop.

Combining this direct analysis of secondary structure with computational analysis of likely secondary structure allowed Watts, Weeks and colleagues to build a low-resolution model of secondary structure encompassing the entire RNA. The HIV RNA encodes 15 mature proteins. Three of its nine open reading frames encode polyproteins that must be cleaved by a protease to yield the mature proteins. For example, the Gag-Pol polyprotein contains the protease, the reverse transcriptase, and the integrase. In Chapter 23 we will discuss HIV and other retroviruses in more detail. The secondary structure model showed a striking correspondence between likely secondary structure and the coding regions for the loops between protein domains, and between mature protein sequences in the polyproteins. Thus, the RNA appears to have a regulatory code written into its sequence that would cause ribosomes to encounter RNA secondary structure and pause between coding regions for protein domains. And this pausing should help with protein folding during translation.

Joshua Plotkin and colleagues enriched this discussion in 2009 when they created a library of 154 genes encoding green fluorescent protein (GFP), all containing "silent" mutations that did not change the coding of the gene. But, when these genes were expressed in *E. coli*, they yielded protein levels that differed by a factor of 250. Codon bias played little or no role in this variation; instead, the stability of mRNA folding, particularly around the Shine-Dalgarno sequence, was the most important factor.

To minimize misfolding, the cell needs a mechanism to hide hydrophobic sections of a nascent polypeptide until the right partner is made. Ordinary molecular chaperones do this by enveloping exposed hydrophobic protein regions in a hydrophobic pocket of their own, and preventing inappropriate associations with other exposed hydrophobic regions. But *E. coli* has a special chaperone called **trigger factor** that associates with the large ribosomal subunit and

catches newly-synthesized hydrophobic regions in a hydrophobic basket to protect them from water.

To see how trigger factor does its job, it would be ideal to have the crystal structure of the chaperone bound to its ribosomal docking site. But that presents a problem: The only large ribosomal subunit that has been crystallized is from the archaeon *Haloarcula marismortui* (Chapter 19), but archaea do not have trigger factor. So Nenad Ban and colleagues crystallized the whole *E. coli* trigger factor to see its shape, and then crystallized the ribosome-binding part of *E. coli* trigger factor together with the archaeal large ribosomal subunit, in hopes that the ribosomal binding site was conserved well enough between archaea and bacteria that such a cross-kingdom complex would form.

And the strategy worked! The binding site for trigger factor (on ribosomal protein L23) is highly conserved between bacteria and archaea, so the ribosomal subunit for an archaeon can bind to a bacterial trigger factor. The crystal structure of trigger factor alone suggested to Ban and colleagues a "crouching dragon" with a head, back, arms, and tail, as illustrated in Figure 18.35. Based on the cocrystal structure of the 50S ribosomal subunit with the tail domain of trigger factor, Ban and colleagues positioned trigger factor as shown in Figure 18.35, with the "dragon crouching" upside down. This places the hydrophobic surface of the tail and arm domains in perfect position to catch the nascent polypeptide as it exits through the ribosomal exit tunnel. This would effectively sequester any exposed hydrophobic regions of the nascent polypeptide until they can associate with the appropriate partner hydrophobic regions.

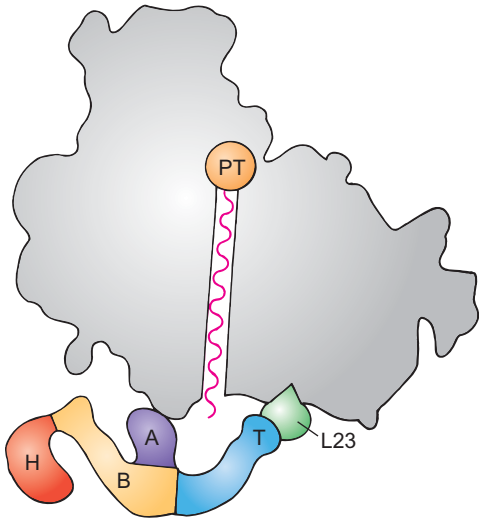


Figure 18.35 A model for trigger factor bound to a ribosome. The chaperone protein, trigger factor, is bound like an upside-down crouching dragon to the bottom of the ribosome, covering the exit tunnel. In this position, the hydrophobic domains of trigger factor (arm [A] and tail [T], purple and blue, respectively) can catch hydrophobic regions of a nascent polypeptide as they emerge from the exit tunnel, and keep them in a hydrophobic environment until they can pair with other hydrophobic regions of the nascent polypeptide, promoting proper folding. The other domains of trigger factor are the head (H, red), and the back (B, yellow). L23 (green) is one of the proteins of the large ribosomal subunit, and is the site of major contacts with trigger factor. PT (orange) is the peptidyl transferase site at the beginning of the exit tunnel. (*Source:* Adapted from Ferbitz, L., T. Maier, H. Patzelt, B. Bukau, E. Deverling, and N. Ban, Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins, *Nature* 431:593, 2004.)

Trigger factor is not essential for *E. coli* life, because bacteria have a backup system: a chaperone called DnaK. It is freestanding protein, rather than a ribosome-associated protein like trigger factor. Instead of a basket to catch nascent proteins, DnaK has a hydrophobic arch that protects exposed hydrophobic regions of nascent proteins until they can fold properly. Archaea and eukaryotes lack trigger-factor-like proteins entirely, so they rely exclusively on freestanding chaperones for proper folding of nascent proteins.

SUMMARY Most newly-made polypeptides do not fold properly by themselves, but require help from molecular chaperones. *E. coli* cells have a protein called trigger factor that associates with the ribosome in such a way as to catch the nascent polypeptide as it emerges from the ribosome's exit tunnel. Thus, hydrophobic regions of the nascent polypeptide are protected from inappropriate associations until the appropriate partner is available. Archaea and eukaryotes lack trigger factor, so they must use freestanding chaperones, which are also present in bacteria. "Silent" mutations can affect translation rates, even though they do not change the sequence of the protein product.

Release of Ribosomes from mRNA

Early studies on termination used model systems, including just AUG and UAG as mRNA analogs, and these studies did not detect a need for ribosome release, in part because some of the model mRNAs dissociated from ribosomes spontaneously.

Then A. Kaji and colleagues discovered a protein factor that could release ribosomes from natural mRNAs in **post-termination complexes** (**post-TCs**). They named it **ribosomal recycling factor** (**RRF**). Then in 1994, Kaji and colleagues demonstrated that RRF is essential for bacterial life. In temperature-sensitive mutants in the gene for RRF, shift to the nonpermissive temperature killed bacteria in lag phase and arrested the growth of bacteria in log phase. Thus, release of ribosomes from mRNAs after termination of translation is essential.

Kaji and colleagues purified RRF from the bacterium *Thermotoga maritima* using the following assay to detect RRF: They treated bacterial polysomes with puromycin to

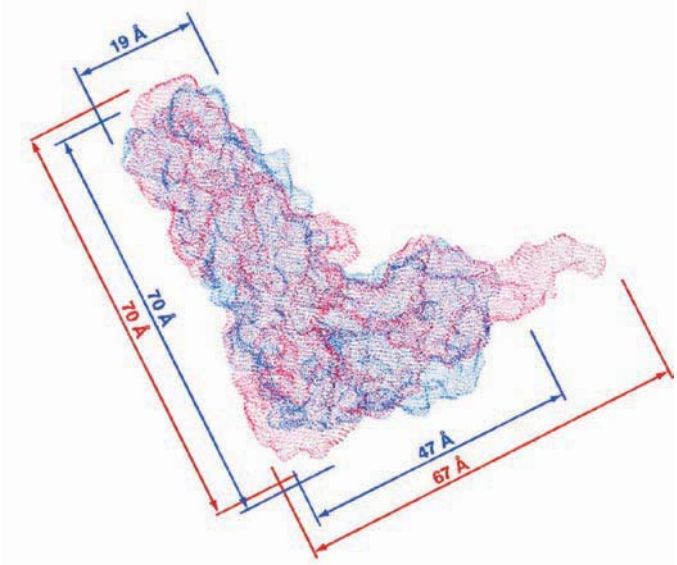


Figure 18.36 Superimposition of the structures of RRF and a tRNA. The surfaces of the *Thermotoga maritima* RRF (blue) and yeast $tRNA^{Phe}$ (red) are superimposed to show their great similarity. (*Source:* From Selmer M., Al-Karadaghi S., Hirokawa G., Kaji A., and Liljas A. 1999. Crystal structure of Thermotoga maritima ribosome recycling factor: A tRNA mimic. Science 286:2349. © 1999 AAAS.)

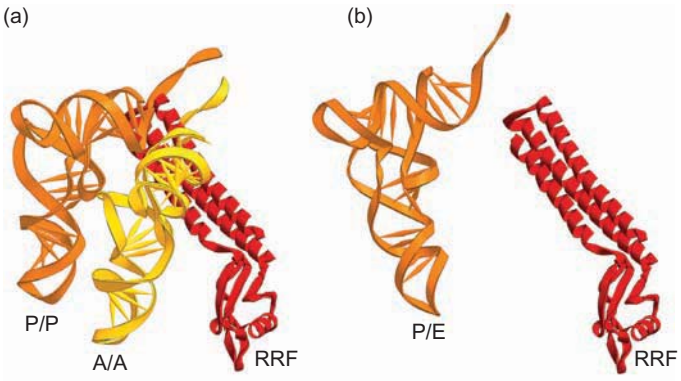


Figure 18.37 Model for the position of RRF in the ribosome. **(a)** Position of RRF (red) relative to tRNAs bound in the pure A site (A/A, yellow) and the pure P site (P/P, orange). **(b)** Position of RRF (red) relative to a tRNA in the hybrid P/E site (orange). (*Source:* Reprinted from Cell v. III, Lancaster et al., p. 444 © 2002, with permission from Elsevier Science.)

release the nascent polypeptide. This left each of the ribosomes with two deacylated tRNAs, one in the P site and one in the E site. Thus, each of the ribosomes in these polysomes resembled a ribosome that had just experienced termination, except that there was no termination codon in the A site. To these puromycin-treated polysomes, these workers added RRF, which converted the polysomes to monosomes. Once it was purified, Kaji and colleagues, in collaboration with Anders Liljas and colleagues, determined the crystal structure of RRF.

The crystal structure was striking—an almost perfect mimic of a tRNA. Figure 18.36 shows the structure of the *T. maritima* RRF superimposed on the structure of $tRNA^{Phe}$. The fit is nearly perfect; the only things missing from RRF are amino acids to fill in the space normally occupied by the terminal CCA of the tRNA, and a small piece of the anticodon. Based on this structure and other information, Kaji and colleagues proposed that RRF binds to the A site, just like an aminoacyl-tRNA would, thereby allowing translocation to occur in the presence of EF-G, and then somehow releases the ribosome from the mRNA.

Then in 2002, Kaji and colleagues, in collaboration with Noller and colleagues, performed structural studies on RRF–ribosome complexes using **hydroxyl radical probing.** They employed this method as follows: First, they used site-directed mutagenesis to replace the single cysteine in the RRF molecule with serine. Then they mutagenized this cysteine-free RRF, which still retained activity, to place cysteine at each of 10 different locations throughout the RRF molecule. Each of these RRF molecules with a single cysteine could be coupled to a molecule bearing Fe^{2+}, and then the RRF-Fe^{2+} could be bound to ribosomes. The Fe^{2+} creates hydroxyl radicals that break nearby segments of rRNA, and these breaks can be detected by primer extension (Chapter 5). Because we know exactly where each part of the 16S and 23S rRNAs are located in the ribosome (Chapter 19), different parts of RRF could be mapped to specific locations on the ribosome.

This experiment demonstrated that, despite its near-perfect structural resemblance to tRNA, RRF does not behave just like a tRNA in binding to the ribosome. It binds to the A site of the ribosome in an orientation very different from that of a tRNA in the A site (Figure 18.37a). This result called into question the simple model of Kaji and colleagues. In fact, it even raised the question of how RRF could bind to the ribosome in the way it does because the end of RRF would overlap with the acceptor stem of a deacylated tRNA bound in the P site. But Kaji, Noller, and colleagues noted that a tRNA deacylated by puromycin, or presumably by RF1 or RF2, does not exist in the pure P site-bound state. Instead, as Noller and colleagues have shown, it is in a hybrid P/E state, with its acceptor end in the E site and its anticodon in the P site. In this position, it would not interfere with RRF's binding, as illustrated in Figure 18.37b.

What happens after RRF binds to the A site? That is still poorly understood, though we know it acts with EF-G to release the ribosome from the mRNA. Some of the time, it could release just the 50S subunit, leaving the 30S subunit to be released by another mechanism, perhaps by binding to IF3.

Eukaryotes do not encode an RRF, so how do they dissociate post-TCs? Tatyana Pestova and colleages showed in 2007 that eIF3 is the most important factor in eukaryotic ribosome release, and it gets help from eIF1, eIF1A, and eIF3j, which is a loosely bound subunit of eIF3.

SUMMARY Ribosomes do not release from the mRNA spontaneously after termination. Bacterial ribosomes need help from ribosome recycling factor (RRF) and EF-G. RRF strongly resembles a tRNA and can bind to the ribosome's A site, but in a position not normally taken by a tRNA. Then it collaborates with EF-G in releasing either the 50S ribosomal subunit, or the whole ribosome. Eukaryotic ribosomes are released from post-TCs by eIF3, aided by eIF1, eIF1A, and eIF3j.

SUMMARY

Messenger RNAs are read in the 5′→3′ direction, the same direction in which they are synthesized. Proteins are made in the amino to carboxyl direction, which means that the amino-terminal amino acid is added first.

The genetic code is a set of three-base code words, or codons, in mRNA that instruct the ribosome to incorporate specific amino acids into a polypeptide. The code is nonoverlapping: that is, each base is part of only one codon. It is also devoid of gaps, or commas; that is, each base in the coding region of an mRNA is part of a codon. There are 64 codons in all. Three are stop signals, and the rest code for amino acids. This means that the code is highly degenerate.

Part of the degeneracy of the genetic code is accommodated by isoaccepting species of tRNA that bind the same amino acid but recognize different codons. The rest is handled by wobble, in which the third base of a codon is allowed to move slightly from its normal position to form a non-Watson–Crick base pair with the anticodon. This allows the same aminoacyl-tRNA to pair with more than one codon. The wobble pairs are G–U (or I–U) and I–A.

The genetic code is not strictly universal. In certain eukaryotic nuclei and mitochondria and in at least one bacterium, codons that cause termination in the standard genetic code can code for amino acids such as tryptophan and glutamine. In several mitochondrial genomes, the sense of a codon is changed from one amino acid to another. These deviant codes are still closely related to the standard one from which they probably evolved.

Elongation takes place in three steps: (1) EF-Tu, with GTP, binds an aminoacyl-tRNA to the ribosomal A site. (2) Peptidyl transferase forms a peptide bond between the peptide in the P site and the newly arrived aminoacyl-tRNA in the A site. This lengthens the peptide by one amino acid and shifts it to the A site. (3) EF-G, with GTP, translocates the growing peptidyl-tRNA, with its mRNA codon, to the P site, and moves the deacylated tRNA in the P site to the E site.

Puromycin resembles an aminoacyl-tRNA, and so can bind to the A site, couple with the peptide in the P site, and release it as peptidyl puromycin. On the other hand, if the peptidyl-tRNA is in the A site, puromycin will not bind to the ribosome, and the peptide will not be released. This defines two sites on the ribosome: the P site, in which the peptide in a peptidyl-tRNA is puromycin reactive, and the A site, in which the peptide in a peptidyl-tRNA is puromycin unreactive. fMet-tRNA$_f^{Met}$ is puromycin reactive in the 70S initiation complex, so it is in the P site. Binding and structural studies have identified a third binding site (the E site) for deacylated tRNA. Such tRNAs bind to the E site as they exit the ribosome, and this binding helps maintain the reading frame of the mRNA.

A ternary complex formed from EF-Tu, aminoacyl-tRNA, and GTP delivers an aminoacyl-tRNA to the ribosome's A site, without hydrolysis of the GTP. In the next step, GTP is hydrolyzed by a ribosome-dependent GTPase activity of EF-Tu, and an EF-Tu–GDP complex dissociates from the ribosome. EF-Ts regenerates an EF-Tu–GTP complex by exchanging GTP for GDP attached to EF-Tu. Addition of aminoacyl-tRNA then reconstitutes the ternary complex for another round of translation elongation.

The protein-synthesizing machinery achieves accuracy during elongation in a two-step process. First, it gets rid of ternary complexes bearing the wrong aminoacyl-tRNA before GTP hydrolysis occurs. If this screen fails, it can still eliminate the incorrect aminoacyl-tRNA in the proofreading step before the wrong amino acid can be incorporated into the growing protein chain. Both these screens may rely on the weakness of incorrect codon–anticodon base pairing to ensure that dissociation will occur more rapidly than either GTP hydrolysis or peptide bond formation. The balance between speed and accuracy of translation is delicate. If peptide bond formation goes too fast, incorrect aminoacyl-tRNAs do not have enough time to leave the ribosome, so their amino acids are incorporated into protein. But if translation goes too slowly, proteins are not made fast enough for the organism to grow successfully.

Peptide bonds are formed by a ribosomal enzyme called peptidyl transferase. This activity resides on the 50S subunit. The 23S rRNA contains the catalytic center of the peptidyl transferase.

Each translocation event moves the mRNA one codon's length, 3 nt, through the ribosome. GTP and EF-G are necessary for translocation, although translocation activity can be expressed without EF-G and GTP in vitro. For a new round of elongation to occur, GTP hydrolysis releases EF-G from the ribosome. The three-dimensional shapes of the EF-Tu–tRNA–GDPNP ternary complex and the EF-G–GDP binary complex have been determined by x-ray crystallography. As predicted, they are very similar.

Amber, ochre, and opal mutations create termination codons (UAG, UAA, and UGA, respectively) in the middle of a message and thereby cause premature termination of translation. These three codons are also the natural stop signals at the ends of coding regions in mRNAs. Most suppressor tRNAs have altered anticodons that can recognize stop codons and prevent termination by inserting an amino acid and allowing the ribosome to move on to the next codon.

Prokaryotic translation termination is mediated by three factors: RF1, RF2, and RF3. RF1 recognizes the termination codons UAA and UAG; RF2 recognizes UAA and UGA. RF3 is a GTP-binding protein that facilitates release of RF1 and RF2 from the ribosome. Eukaryotes have two release factors: eRF1, which recognizes all three termination codons, and eRF3, a ribosome-dependent GTPase that helps eRF1 recognize stop codons and release the finished polypeptide.

Prokaryotes deal with non-stop mRNAs by tmRNA-mediated ribosome rescue. An alanyl-tmRNA, which resembles an alanyl-tRNA, binds to the vacant A site of a ribosome stalled on a non-stop mRNA and donates its alanine to the stalled polypeptide. Then the ribosome shifts to translating an ORF on the tmRNA, adding another nine amino acids to the polypeptide before terminating. These extra amino acids target the polypeptide for destruction, and a nuclease destroys the non-stop mRNA. Eukaryotic ribosomes at the end of the poly(A) tail of a non-stop mRNA recruit the Ski7p–exosome complex to the vacant A site. Next, the Ski complex is recruited to the A site, and the exosome, positioned just at the end of the non-stop mRNA, degrades that RNA. The aberrant polypeptide is presumably also destroyed.

Eukaryotes deal with premature termination codons by two different mechanisms: NMD and NAS. NMD in mammalian cells involves a downstream destabilizing element, including Upf1 and Upf2 bound to an mRNA at exon–exon junctions that measures the distance to a stop codon. If the codon is far enough upstream, it looks like a premature stop codon and activates the downstream destabilizing element to degrade the mRNA. In yeast, the absence of a normal 3′-UTR or poly(A) near a stop codon may identify it as abnormal. The NAS machinery senses a stop codon in the middle of a reading frame and changes the splicing pattern such that the premature stop codon is spliced out of the mature mRNA. Like NMD, this process also requires Upf1.

The unusual amino acids selenocysteine and pyrrolysine are incorporated into growing polypeptides in response to the termination codons UGA and UAG, respectively, as follows: (1) Selenocysteine: A special tRNA (with an anticodon that recognizes the UGA codon) is charged with serine, which is then converted to selenocysteine, and the selenocysteyl-tRNA is escorted to the ribosome by a special EF-Tu. (2) Pyrrolysine: A special pyrrolysyl-tRNA synthetase joins preformed pyrrolysine with a special tRNA that has an anticodon that recognizes the codon UAG.

Most newly-made polypeptides do not fold properly by themselves, but require help from molecular chaperones. *E. coli* cells have a protein called trigger factor that associates with the ribosome in such a way as to catch the nascent polypeptide as it emerges from the ribosome's exit tunnel. Thus, hydrophobic regions of the nascent polypeptide are protected from inappropriate associations until the appropriate partner is available. Archaea and eukaryotes lack trigger factor, so they must use freestanding chaperones, which are also present in bacteria.

Ribosomes do not release from the mRNA spontaneously after termination; they need help from ribosome recycling factor (RRF) and EFG. RRF strongly resembles a tRNA and can bind to the ribosome's A site, but in a position not normally taken by a tRNA. Then it collaborates with EFG in releasing either the 50S ribosomal subunit, or the whole ribosome, by an unknown mechanism.

REVIEW QUESTIONS

1. Describe and give the results of an experiment that shows that translation starts at the amino terminus of a protein.
2. How do we know that mRNAs are read in the 5′→3′ direction?
3. How do we know that the genetic code is: (a) nonoverlapping; (b) commaless; (c) triplet; (d) degenerate?
4. Describe and give the results of an experiment that reveals two of the codons for an amino acid.
5. Diagram a wobble base pair. You do not have to show the positions of all the atoms, just the shape of the base pair. Contrast this with the shape of a Watson–Crick base pair. What is the importance of wobble in translation?
6. Diagram the translation elongation process in prokaryotes.
7. Diagram the mode of action of puromycin.
8. Describe and give the results of an experiment that shows that fMet-tRNA$_f^{Met}$ occupies the P site of the ribosome.
9. Describe and give the results of an experiment that shows that EF-Ts releases GDP from EF-Tu.
10. What step in translation does chloramphenicol block?
11. Diagram the roles of EF-Tu and EF-Ts in translation.
12. Present evidence for the formation of a ternary complex among EF-Tu, GTP, and aminoacyl-tRNA.
13. Describe and give the results of an experiment that shows that ribosomal RNA is likely to be the catalytic agent in peptidyl transferase.
14. What are the initial recognition and proofreading steps in protein synthesis?

15. Describe and give the results of an experiment that shows that the mRNA moves in 3-nt units in the translocation step.
16. Describe and give the results of an experiment that shows that EF-G and GTP are both required for translocation. What are the effects of (a) substituting GDPCP for GTP, and (b) adding fusidic acid in this single-translocation event assay?
17. Describe an experiment that shows that GTP hydrolysis precedes translocation.
18. Present direct evidence that the amber codon is a translation terminator.
19. Present evidence that the amber codon is UAG.
20. Explain how an amber suppressor works.
21. Present evidence that the amber suppressor is a tRNA.
22. Describe an assay for a release factor.
23. What are the roles of RF1, RF2, and RF3?
24. How do we know which termination codons RF1 and RF2 recognize?
25. What are the roles of eRF1 and eRF3?
26. Diagram the mechanism by which prokaryotes deal with non-stop mRNAs.
27. What differences between tmRNAs and tRNAs limit the ability of tmRNAs to bind tightly to the ribosome? How does the cell deal with these deficiencies?
28. Diagram the mechanism by which mammalian cells deal with non-stop mRNAs.
29. Diagram two mechanisms by which eukaryotic cells deal with premature termination codons.
30. Describe the mechanisms by which selenocysteine and pyrrolysine are incorporated into proteins.
31. How does trigger factor's cellular location help it in its chaperone function?

ANALYTICAL QUESTIONS

1. What would be the effect on a G protein's activity if:
   a. its GAP were inhibited?
   b. its guanine nucleotide exchange protein were inhibited?
2. You have isolated an *E. coli* mutant with an aminoacyl-tRNA synthetase that causes a tRNA with the anticodon 3′-UUC-5′ to be charged with asparagine at the elevated temperature of 42°C. What effect would you expect this to have on protein synthesis in these cells at 42°C, and why? You then isolate another mutant that suppresses the first mutation, and you trace the second mutation to a tRNA gene. What tRNA would you expect to be altered in the second mutant, and where? Predict the nature of this alteration.
3. Consider this short mRNA: 5′-AUGGCAGUGCCA-3′. Answer the following questions, assuming first that the code is fully overlapping and then that it is nonoverlapping.
   a. How many codons would be represented in this oligonucleotide?
   b. If the second G were changed to a C, how many codons would be changed?
4. What would be the effect on reading frame and gene function if
   a. two bases were inserted into the middle of an mRNA?
   b. three bases were inserted into the middle of an mRNA?
   c. one base were inserted into one codon and one subtracted from the next?
5. If codons were six bases long, what kind of product would you expect from a repeating tetranucleotide such as poly (UUCG)?
6. How many codons would exist in a genetic code that had codons that were four bases long?
7. A certain ochre suppressor inserts glutamine in response to the ochre codon. What is the likeliest change in the anticodon of a $tRNA^{Gln}$ that created this suppressor strain?
8. Describe the evolutionary changes that had to occur to give an organism the ability to incorporate pyrrolysine into its proteins. In what order do you think these changes occurred? Why? *Hint:* See Wang, L. (2003). Expanding the genetic code. *Science* 302:584–85.
9. Each of the 20 amino acids can be found in natural proteins adjacent to each of the other amino acids. How does this prove that the genetic code is nonoverlapping?

SUGGESTED READINGS

General References and Reviews

Horwich, A. 2004. Sight at the end of the tunnel. *Nature* 431:520–22.

Kaji, A., M.C. Kiel, G. Hirokawa, A.R. Muto, Y. Inokuchi, and H. Kaji. 2001. The fourth step of protein synthesis: Disassembly of the posttermination complex is catalyzed by elongation factor G and ribosome recycling factor, a near-perfect mimic of tRNA. *Cold Spring Harbor Symposia on Quantitative Biology* 66:515–29.

Kaziro, Y. 1978. The role of guanosine 5′-triphosphate in polypeptide chain elongation. *Biochimica et Biophysica Acta* 505:95–127.

Khorana, H.G. 1968. Synthesis in the study of nucleic acids. *Biochemistry Journal* 109:709–25.

Maquat, L.E. 2002. Skiing toward nonstop mRNA decay. *Science* 295:2221–22.

Moore, M.J. 2002. No end to nonsense. *Science* 298:370–71.

Moore, S.D., K.E. McGinness, and R.T. Sauer. 2003. A glimpse into tmRNA-mediated ribosome rescue, *Science* 300:72–73.

Nakamura, Y., K. Ito, and M. Ehrenberg. 2000. Mimicry grasps reality in translation termination. *Cell* 101:349–52.

Nakamura, Y., K. Ito, and L.A. Isaksson. 1996. Emerging understanding of translation termination. *Cell* 87:147–50.

Nierhaus, K.H. 1996. An elongation factor turn-on. *Nature* 379:491–92.

Ramakrishnan, V. 2002. Ribosome structure and the mechanism of translation. *Cell* 108:557–72.

Schimmel, P. and K. Beebe. 2004. Genetic code seizes pyrrolysine. *Nature* 431:257–58.

Schmeing, T.M. and V. Ramakrishnan. 2009. What recent ribosome structures have revealed about the mechanism of translation. *Nature* 461:1234–42.
Thompson, R.C. 1988. EFTu provides an internal kinetic standard for translational accuracy. *Trends in Biochemical Sciences* 13:91–93.
Tuite, M.F. and I. Stansfield. 1994. Knowing when to stop. *Nature* 372:614–15.

Research Articles

Benzer, S. and S.P. Champe. 1962. A change from nonsense to sense in the genetic code. *Proceedings of the National Academy of Sciences USA* 48:1114–21.
Brenner, S., A.O.W. Stretton, and S. Kaplan. 1965. Genetic code: The "nonsense" triplets for chain termination and their suppression. *Nature* 206:994–98.
Bretscher, M.S., and K.A. Marcker. 1966. Peptidyl-sRibonucleic acid and amino-acyl-sRibonucleic acid binding sites on ribosomes. *Nature* 211:380–84.
Crick, F.H.C., L. Barnett, S. Brenner, and R.J. Watts-Tobin. 1961. General nature of the genetic code for proteins. *Nature* 192:1227–32.
Dintzis, H.M. 1961. Assembly of the peptide chains of hemoglobin. *Proceedings of the National Academy of Sciences USA* 47:247–61.
Ferbitz, L., T. Maier, H. Patzelt, B. Bukau, E. Deuerling, and N. Ban. 2004. Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* 431:590–96.
Fredrick, K. and H.F. Noller. 2003. Catalysis of ribosomal translocation by sparsomycin. *Science* 300:1159–62.
Goodman, H.M., J. Abelson, A. Landy, S. Brenner, and J.D. Smith. 1968. Amber suppression: A nucleotide change in the anticodon of a tyrosine transfer RNA. *Nature* 217:1019–24.
Haenni, A.-L. and J. Lucas-Lenard. 1968. Stepwise synthesis of a tripeptide. *Proceedings of the National Academy of Sciences USA* 61:1363–69.
Inoue-Yokosawa, N., C. Ishikawa, and Y. Kaziro. 1974. The role of guanosine triphosphate in translocation reaction catalyzed by elongation factor G. *Journal of Biological Chemistry* 249:4321–23.
Ito, K., M. Uno, and Y. Nakamura. 2000. A tripeptide "anticodon" deciphers stop codons in messenger RNA. *Nature* 403:680–84.
Khaitovich, P., A.S. Mankin, R. Green, L. Lancaster, and H.F. Noller. 1999. Characterization of functionally active subribosomal particles from *Thermus aquaticus*. *Proceedings of the National Academy of Sciences USA* 96:85–90.
Lancaster, L., M.C. Kiel, A. Kaji, and H.F. Noller. 2002. Orientation of ribosome recycling factor in the ribosome from directed hydroxyl radical probing. *Cell* 111:129–40.
Last, J.A., W.M. Stanley, Jr., M. Salas, M.B. Hille, A.J. Wahba, and S. Ochoa. 1967. Translation of the genetic message, IV. UAA as a chain termination codon. *Proceedings of the National Academy of Sciences USA* 57:1062–67.
Miller, D.L. and H. Weissbach. 1970. Interactions between the elongation factors: The displacement of GDP from the Tu-GDP complex by factor Ts. *Biochemical and Biophysical Research Communications* 38:1016–22.
Nirenberg, M. and P. Leder. 1964. RNA codewords and protein synthesis: The effect of trinucleotides upon binding of sRNA to ribosomes. *Science* 145:1399–1407.
Nissen, P., M. Kjeldgaard, S. Thirup, G. Polekhina, L. Reshetnikova, B.F.C. Clark, and J. Nyborg. 1995. Crystal structure of the ternary complex of Phe-tRNA^{Phe}, EF-Tu, and a GTP analog. *Science* 270:1464–71.
Noller, H.F., V. Hoffarth, and L. Zimniak. 1992. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* 256:1416–19.
Ravel, J.M. 1967. Demonstration of a guanine triphosphate-dependent enzymatic binding of aminoacyl-ribonucleic acid to *Escherichia coli* ribosomes. *Proceedings of the National Academy of Sciences USA* 57:1811–16.
Ravel, J.M., R.L. Shorey, and W. Shire. 1968. The composition of the active intermediate in the transfer of aminoacyl-RNA to ribosomes. *Biochemical and Biophysical Research Communications* 32:9–14.
Rheinberger, H.-J., H. Sternbach, and K.H. Nierhaus. 1981. Three tRNA binding sites on *Escherichia coli* ribosomes. *Proceedings of the National Academy of Sciences USA* 78:5310–14.
Rodnina, M.V., A. Savelsbergh, V.I. Katunin, and W. Wintermeyer. 1997. Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature* 385:37–41.
Sarabhai, A.S., A.O.W. Stretton, S. Brenner, and A. Bolle. 1964. Co-linearity of the gene with the polypeptide chain. *Nature* 201:13–17.
Scolnick, E., R. Tompkins, T. Caskey, and M. Nirenberg. 1968. Release factors differing in specificity for terminator codons. *Proceedings of the National Academy of Sciences USA* 61:768–74.
Thach, S.S. and R.E. Thach. 1971. Translocation of messenger RNA and "accommodation" of fMet-tRNA. *Proceedings of the National Academy of Sciences USA* 68:1791–95.
Traut, R.R. and R.E. Monro. 1964. The puromycin reaction and its relation to protein synthesis. *Journal of Molecular Biology* 10:63–72.
Valle, M., R. Gillet, S. Kaur, A. Henne, V. Ramakrishnan, and J. Frank. 2003 Visualizing tmRNA entry into a stalled ribosome. *Science* 300:127–30.
Weigert, M.G. and A. Garen. 1965. Base composition of nonsense codons in *E. coli*. *Nature* 206:992–94.
Weissbach, H., D.L. Miller, and J. Hachmann. 1970. Studies on the role of factor Ts in polypeptide synthesis. *Archives of Biochemistry and Biophysics* 137:262–69.
Zhouravleva, G., L. Frolova, X. Le Goff, R. Le Guellec, S. Inge-Vechtomov, L. Kisselev, and M. Philippe. 1995. Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. *EMBO Journal* 14:4065–72.