Genetic code & Translation

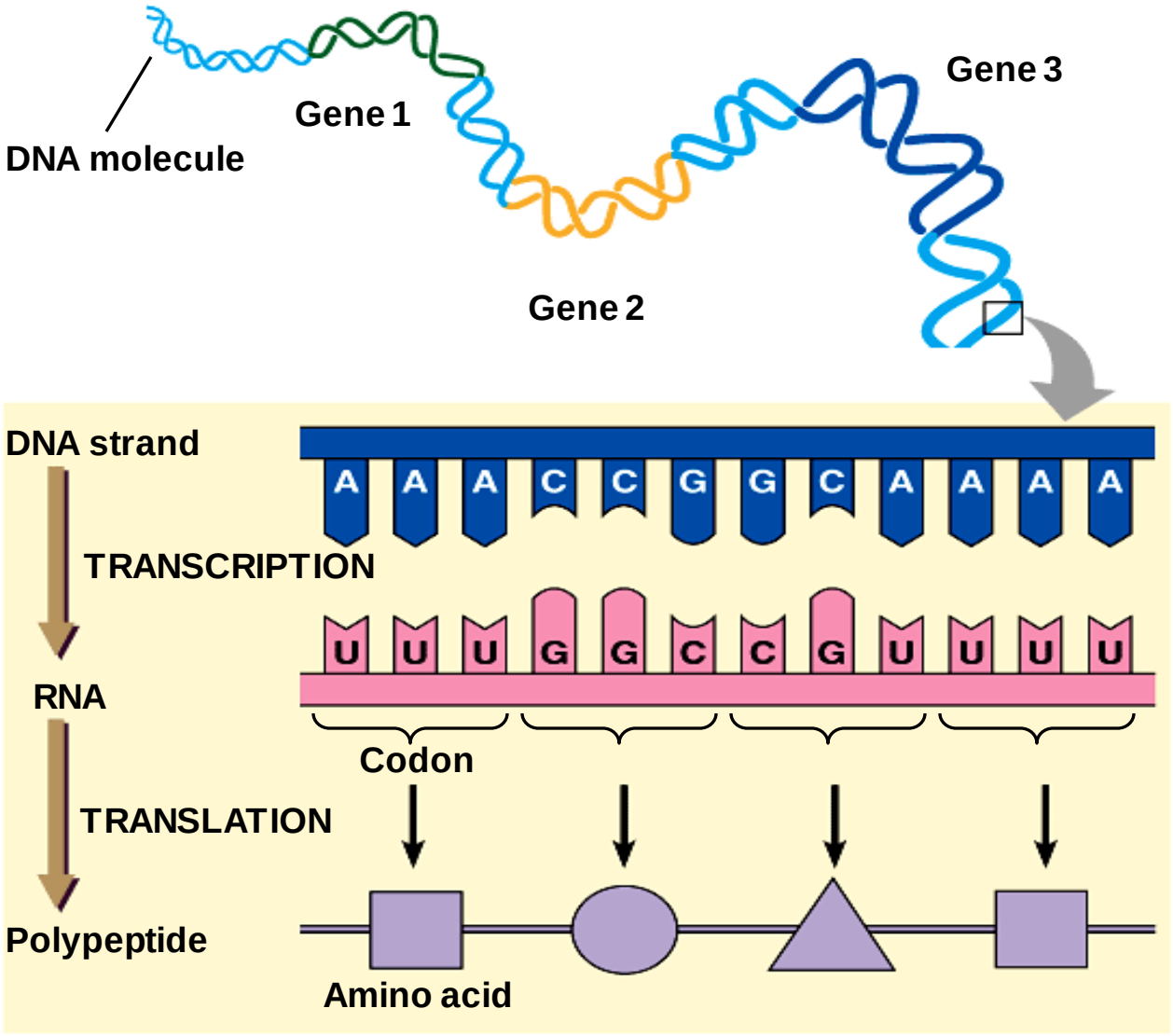


Figure 10.7

The Genetic Code

*The **genetic code** is the set of rules by which information encoded in genetic material (DNA or mRNA sequences) is translated into proteins (amino acid sequences) by living cells.*

It is a set of 3 nucleotides on the mRNA known as CODON.

(Since there are 4 bases, there are $4^3 = 64$ possible codons, which must code for 20 different amino acids.)

The Genetic Code

- is almost universal, used in both prokaryotes and eukaryotes.
- The code is written in the 5' → 3' direction.
- AUG is used as the start codon. All proteins are initially translated with methionine in the first position, although it is often removed after translation. There are also internal methionine in most proteins, coded by the same AUG codon.

- There are 3 stop codons (UAA, UGA, UAG), also called "nonsense" codons. Proteins end in a stop codon, which codes for no amino acid.

- The first codon establishes the **Reading frame.**

- Known to be nonoverlapping.

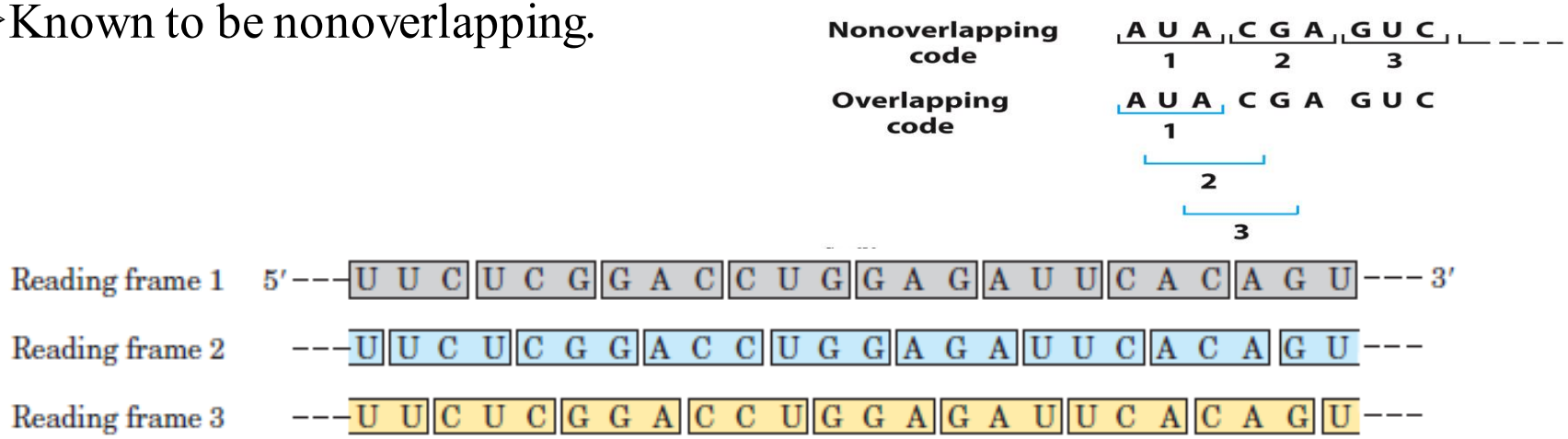


FIGURE 27–5 Reading frames in the genetic code. In a triplet, nonoverlapping code, all mRNAs have three potential reading frames, shaded here in different colors. The triplets, and hence the amino acids specified, are different in each reading frame.

Breaking the Genetic Code

- One of the greatest scientific developments of the twentieth century
- First clue from Nirenberg and Matthaei:
  - poly(U) and 20 radiolabeled amino acids fed to *E.coli* →only *Phe produced
    - Hence, UUU codes for Phe.
- Another step forward from Khorana:
  - used defined mRNAs in planned patterns
    - that is, $(AC)_n$ (alternating ACA and CAC codons → His and Thr)

Virtually all organisms share the same genetic code

"unity of life"

First Base	Second Base: U	Second Base: C	Second Base: A	Second Base: G	Third Base
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 (start)	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

There are 20 common, genetically encoded amino acids, by 61/64 codons.

Most Amino Acids Have More Than One Codon

- Some codons are less subject to causing a mutation in an amino acid sequence because of degeneracy or because of the abundance of such tRNAs.
- There are 20 amino acids with 61 possible codons.
- Only Met and Trp have a **single** codon.

Codon Correlation with Amino Acids

TABLE 27-3	Degeneracy of the Genetic Code		
Amino acid	**Number of codons**	**Amino acid**	**Number of codons**
Met	1	Tyr	2
Trp	1	Ile	3
Asn	2	Ala	4
Asp	2	Gly	4
Cys	2	Pro	4
Gln	2	Thr	4
Glu	2	Val	4
His	2	Arg	6
Lys	2	Leu	6
Phe	2	Ser	6

The Genetic Code Is Resistant to Mutations

- Degenerate code allows certain mutations to still code for the same amino acid.
  - "silent" mutations—different nucleotide in DNA but same amino acid in protein
- Mutation in the first base of a codon usually produces a conservative substitution.
  - Example: GUU → Val but AUU →Leu

The Genetic Code Is Universal, With a Few Exceptions

- It is used by prokaryotes and eukaryotes, across species.
- Mitochondria contain DNA and use a slightly different code.
  - UGA encodes Trp in vertebrate mtDNA (instead of STOP).
  - AGA/AGG encodes STOP in vertebrate mtNDA (instead of Arg).
- Mitochondria encode their own tRNAs, using 22 instead of 32.

Molecular Recognition of Codons in mRNA by tRNA

- The codon sequence is complementary with the anticodon sequence.
- The codon in mRNA base pairs with the anticodon in mRNA via hydrogen bonding.
- The alignment of two RNA segments is antiparallel.
- A reading frame without a termination codon among 50 or more codons is referred to as an **open reading frame (ORF).**

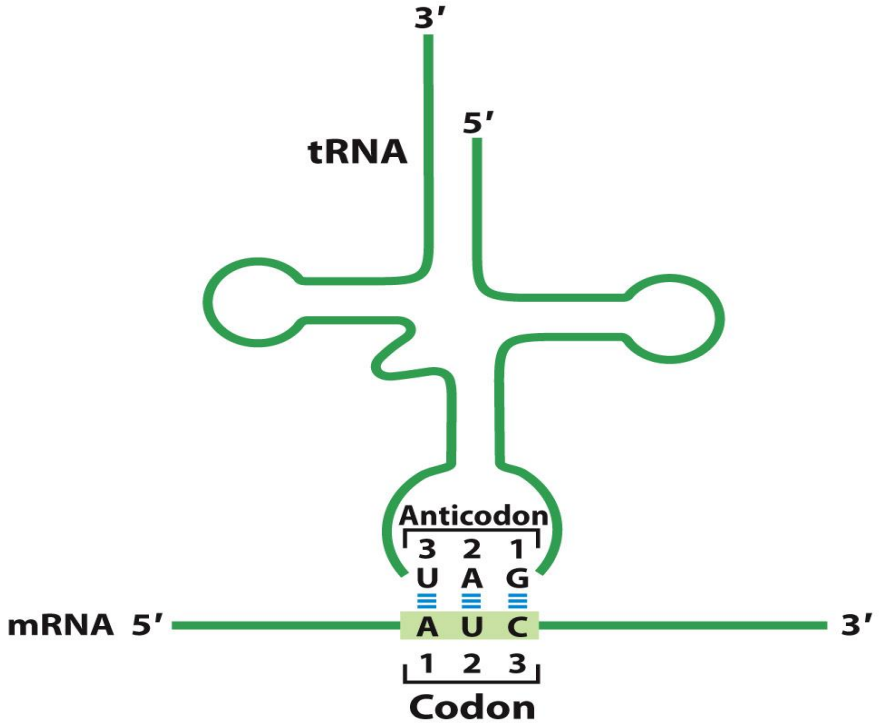


Figure 27-8a
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

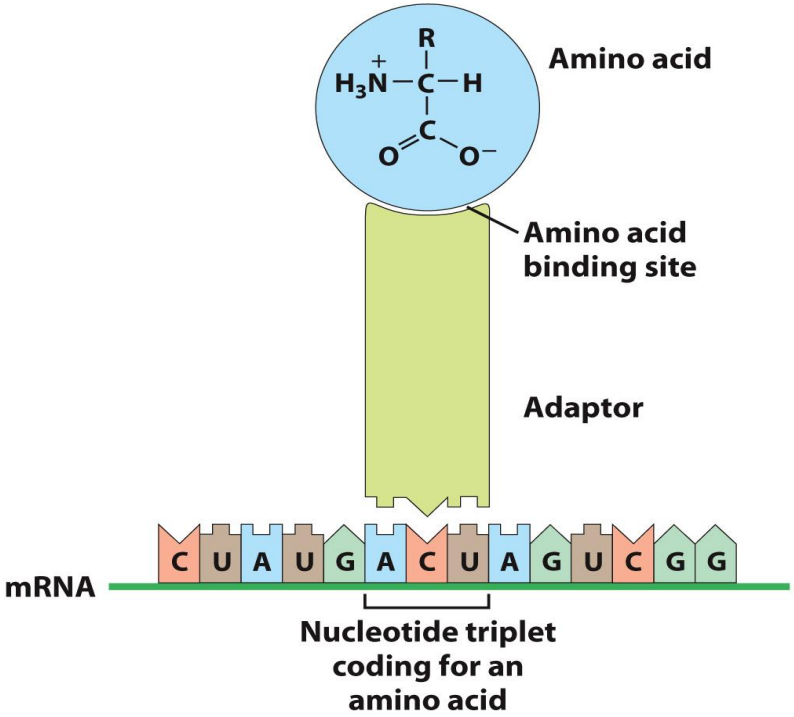


Figure 27-2
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

"Wobble" Pairings in tRNA with mRNA Can Occur in the Third Base

- The third base of a codon can form noncanonical base pairs with its complement (**anticodon**) in tRNA.
- Some tRNAs contain Inosinate (I), which can H-bond with U,C, and A.
  - These H-bonds are weaker and were named by Crick as "wobble" base pairs.
  - Example: In yeast, CGA, GCU, and CGC all bind to $tRNA^{Arg}$, which has the anticodon $^{3'}$-GCI-$^{5'}$.
    - Although sequences are usually written 5'→3', the anticodon here is written 3'→5' to illustrate its bonding to the mRNA codons.

Wobble Hypothesis

1. The first two bases of codon always form strong base pairs with the corresponding bases anticodon and confer the coding specificity.
2. When an amino acid is specifled by several different codons, the codons that differ in either of the first two bases require different tRNAs.
3. A minimum of 32 tRNAs are required to translate all 61 codons (31 to encode the amino acids and 1 for initiation).
4. The first base of the anticodon (reading in the 5'-3' direction; this pairs with the third base of the codon) determines the number of codons recognized by the tRNA.

a. When the first base of the anticodon is C or A, base pairing is specific and only one codon is recognized.
b. When the first base is U or G, binding is less specific and two different codons may be read.
c. When inosine (I) is the first (wobble) nucleotide of an anticodon, three different codons can be recognized-the maximum number for any tRNA.

TABLE 27-4 How the Wobble Base of the Anticodon Determines the Number of Codons a tRNA Can Recognize

1. One codon recognized:

Anticodon	(3′) X–Y– **C** (5′)	(3′) X–Y– **A** (5′)
Codon	(5′) X′–Y′– **G** (3′)	(5′) X′–Y′– **U** (3′)

2. Two codons recognized:

Anticodon	(3′) X–Y– **U** (5′)	(3′) X–Y– **G** (5′)
Codon	(5′) X′–Y′– A/G (3′)	(5′) X′–Y′– C/U (3′)

3. Three codons recognized:

Anticodon	(3′) X–Y– **I** (5′)
Codon	(5′) X′–Y′– A/U/C (3′)

Note: X and Y denote bases complementary to and capable of strong Watson-Crick base pairing with X′ and Y′, respectively. Wobble bases—in the 3′ position of codons and 5′ position of anticodons—are shaded in white.

Some mRNAs Are Edited Before Protein Synthesis

- Editing involves alteration, addition, or deletion of nucleotides in mRNA.
- Editing uses guide RNAs (gRNAs) that temporarily hybridize with the mRNA and act as templates for editing.

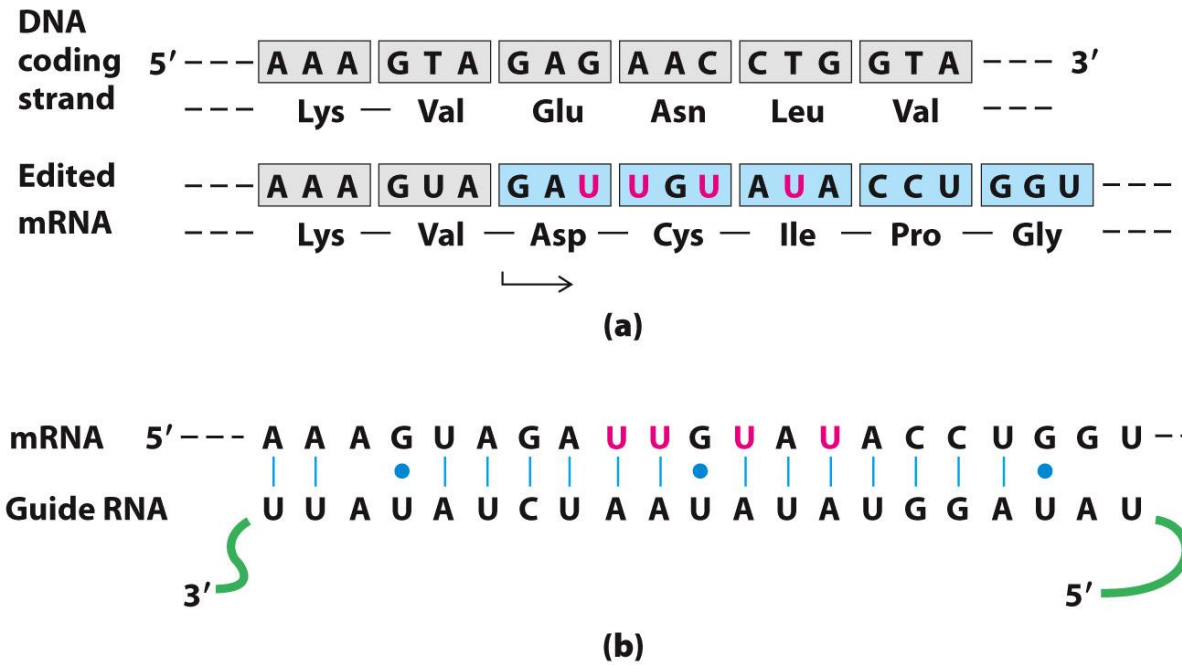


Figure 27-10
Lehninger Principles of Biochemistry, Seventh Edition

Deamination Reactions Yield

- RNA editing by alteration of nucleotides most commonly involves the enzymatic deamination of adenosine or cytidine residues.

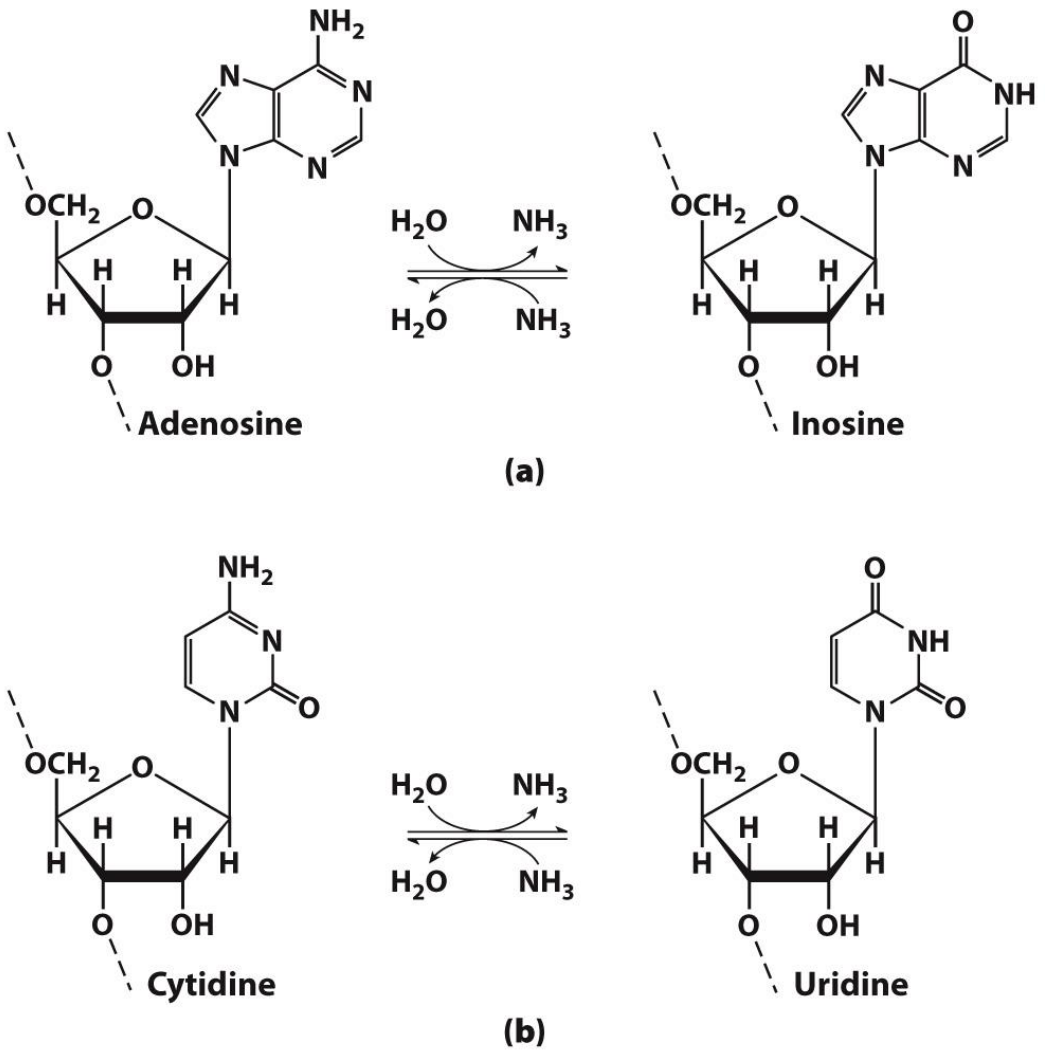


Figure 27-11
Lehninger Principles of Biochemistry, Seventh Edition

Five Stages of Protein Synthesis

1. Activation of amino acids
   - tRNA aminoacylated
2. Initiation of translation
   - mRNA and aminoacylated tRNA bind to ribosome
3. Elongation
   - cycles of aminoacyl-tRNA binding and peptide bond formation…until a STOP codon is reached
4. Termination and ribosome recycling
   - mRNA and protein dissociate, ribosome recycled
5. Folding and posttranslational processing
   - catalyzed by a variety of enzymes

Overview of Protein Synthesis

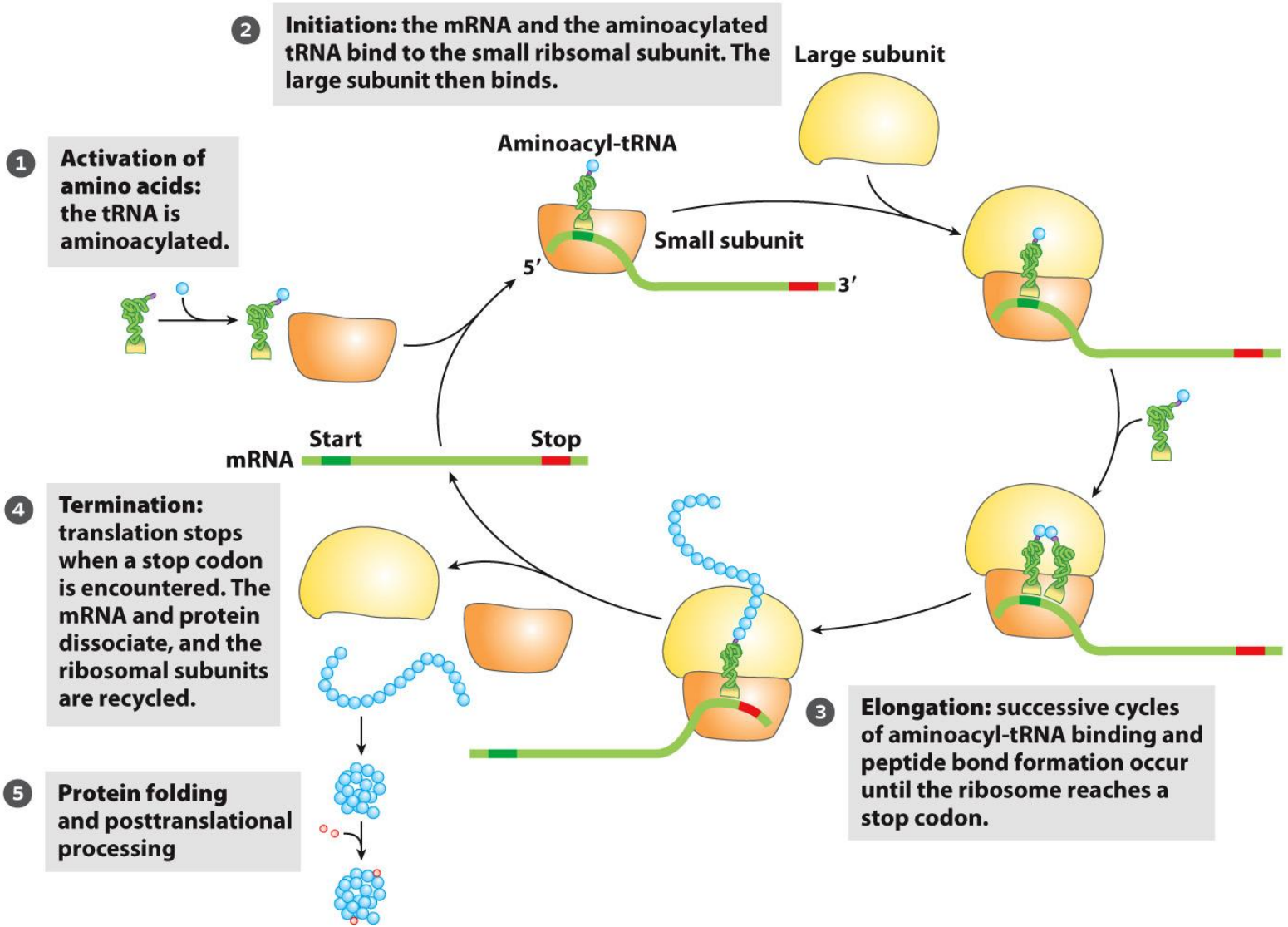


Figure 27-13
Lehninger Principles of Biochemistry, Seventh Edition

TABLE 27-5 Components Required for the Five Major Stages of Protein Synthesis in *E. Coli*

Stage	Essential components
1. Activation of amino acids	20 amino acids 20 aminoacyl-tRNA syntheses 32 or more tRNAs ATP Mg^{2+}
2. Initiation	mRNA N-Formylmethionyl-tRNAfMet Initiation codon in mRNA (AUG) 30S ribosomal subunit 50S ribosomal subunit Initiation factors (IF1, IF2, IF3) GTP Mg^{2+}
3. Elongation	Functional 70S ribosomes (initiation complex) Aminoacyl-tRNAs specified by codons Elongation factors (EF-Tu, EF-Ts, EF-G) GTP Mg^{2+}
4. Termination and ribosome recycling	Termination codon in mRNA Release factors (RF1, RF2, RF3, RRF) EF-G IF3
5. Folding and posttranslational processing	Chaperones and folding enzymes (PPI, PDI); specific enzymes, cofactors, and other components for removal of initiating residues and signal sequences, additional proteolytic processing, modification of terminal residues, and attachment of acetyl, phosporyl, methyl, carboxyl, carbohydrate, or prosthetic groups

The Ribosome Is a Key Player in Protein Synthesis

- Make up ~25% of dry weight of bacteria
- ~65% rRNA, 35% protein
  - rRNA forms the core.
  - RNA does the catalysis of peptide bond formation.
- Made of two subunits bound together (30S and 50S) in bacteria, with mRNA running through them

TABLE 27-6 RNA and Protein Components of the *E. Coli* Ribosome

Subunit	Number of different proteins	Total number of proteins	Protein designations	Number and type of rRNAs
30S	21	21	S1–S21	1 (16S rRNA)
50S	33	36	L1–L36[a]	2 (5S and 23S rRNAs)

[a]The L1 to L36 protein designations do not correspond to 36 different proteins. The protein originally designated L7 is a modified form of L12, and L8 is a complex of three other proteins. Also, L26 proved to be the same protein as S20 (and not part of the 50S subunit). This gives 33 different proteins in the large subunit. There are four copies of the L7/L12 protein, with the three extra copies bringing the total protein count to 36.

Ribosomes in Bacteria and Eukaryotes

- Overall, very similar
- Still two subunits with mRNA running between
- In eukaryotes, larger (80S), more complex, contain > 80 proteins
- Chloroplasts and mitochondria have ribosomes simpler than those in bacteria.

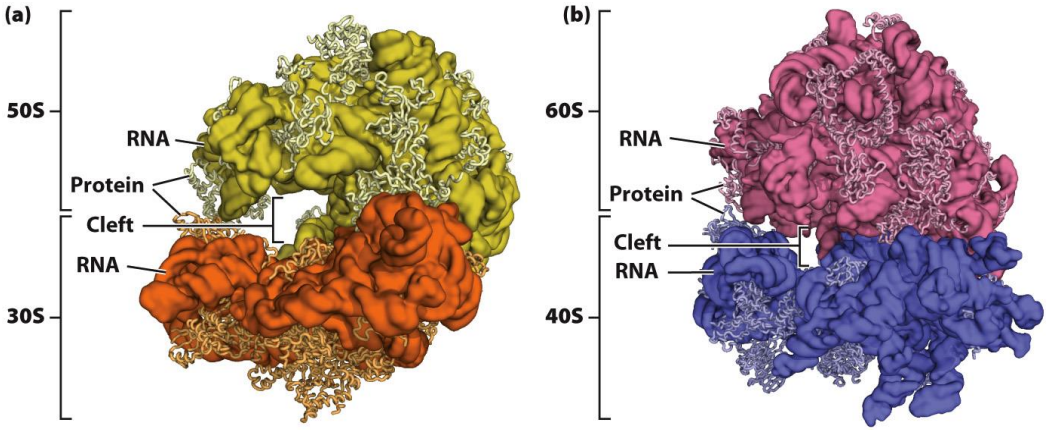


Figure 27-14
Lehninger Principles of Biochemistry, Seventh Edition

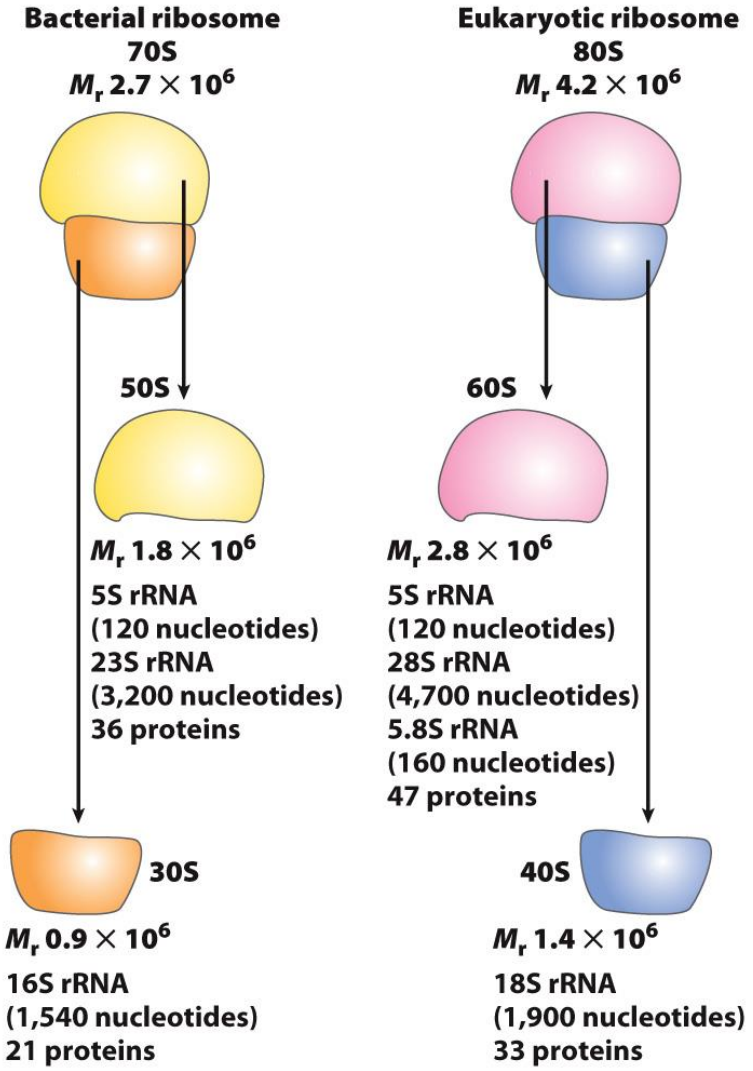


Figure 27-16
Lehninger Principles of Biochemistry, Seventh Edition

rRNAs Have Complex Secondary Structures

- The three ss rRNAs having specific 3-D structure with extensive intrachain base pairs.
- Shape of rRNAs are highly conserved.
- *The red, yellow, and purple indicate areas where the structures of the rRNAs from bacteria, archaea, and eukaryotes have diverged.*
- *Conserved regions are shown in green.*

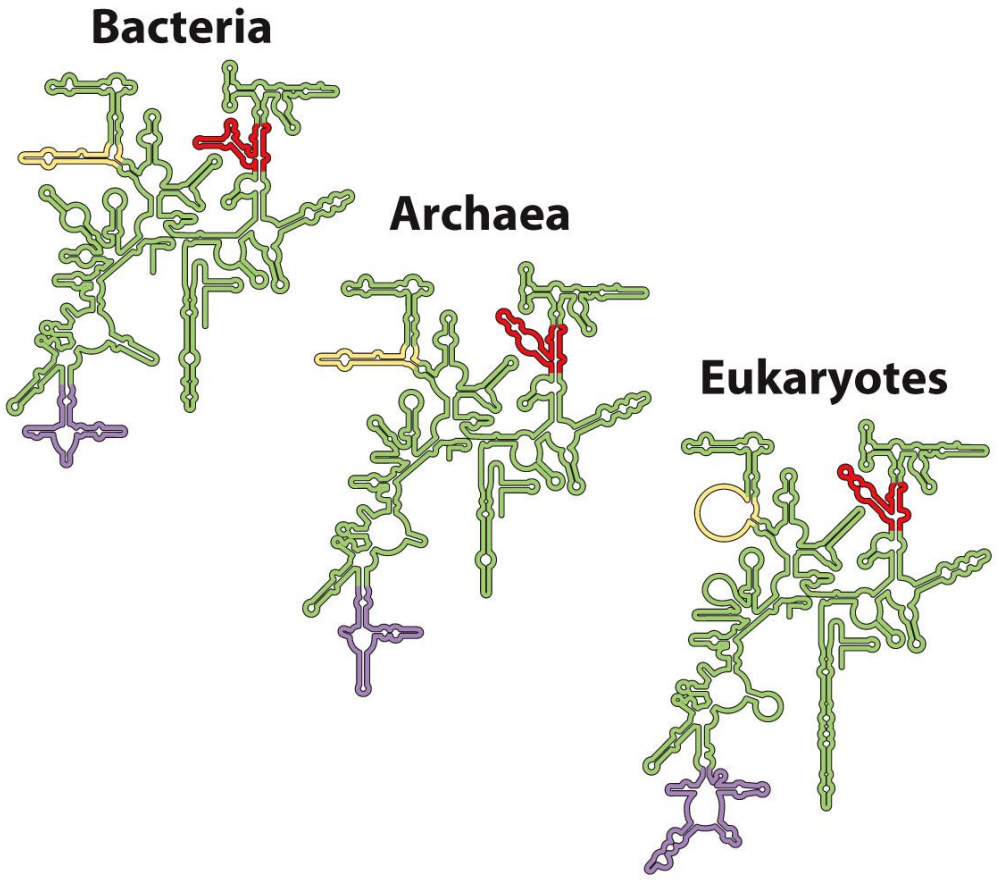


Figure 27-15
Lehninger Principles of Biochemistry, Seventh Edition

Characteristics of tRNAs

- ssRNA of 73–93 nucleotides in both bacteria and eukaryotes
- Cloverleaf structure in 2-D, "Twisted L" shape in 3-D
- Serve as interpreters during translation
- Most have G at 5'-end; all have CAA at 3'-end
- Have modified bases
  - methylated bases, and so on

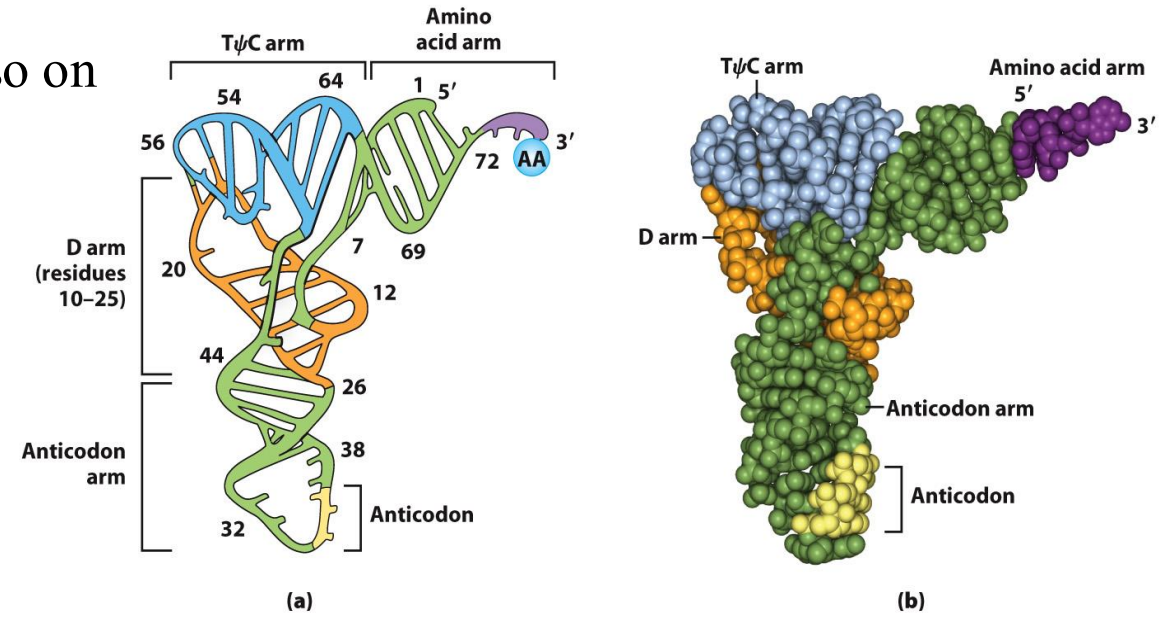


Figure 27-18
Lehninger Principles of Biochemistry, Seventh Edition

3-D Structure of Yeast tRNA^{Phe}

Characteristics of tRNAs

- **Amino acid arm**
  - has amino acid esterified via carboxyl group to the 2'-OH or 3'-OH of the A of the terminal CAA codon
- **Anticodon arm**
- **D arm**
  - contains dihydrouridine (D)
  - contributes to folding
- **TψC arm**
  - contains pseudouridine (ψ)—has bonding between base and ribose
  - helps in folding

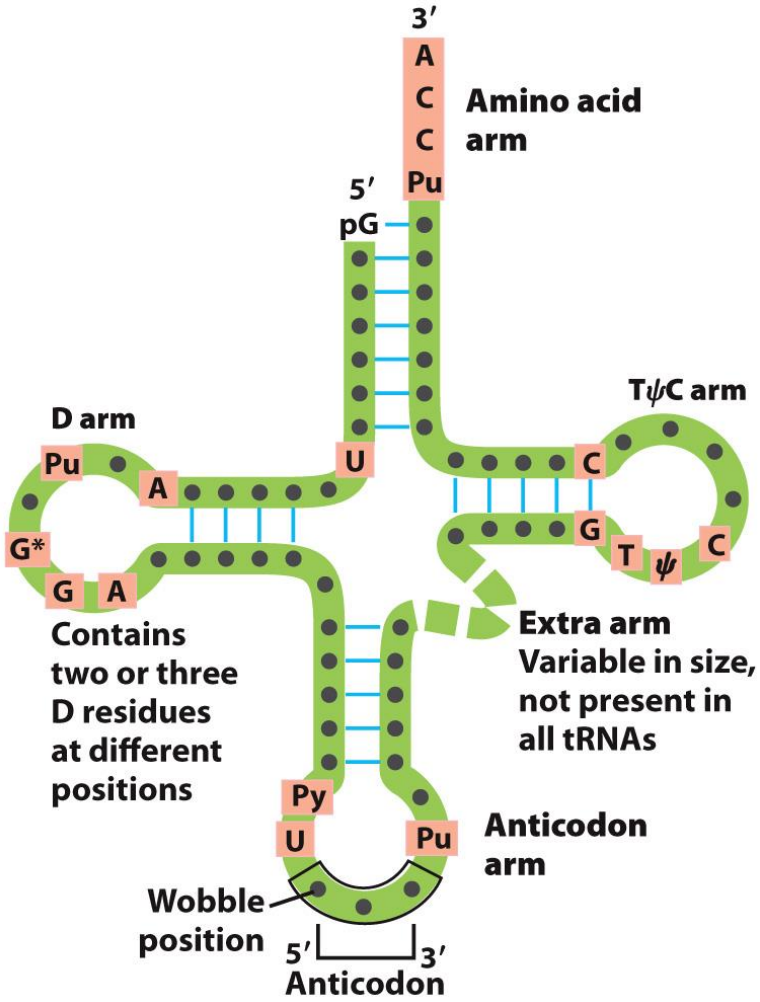
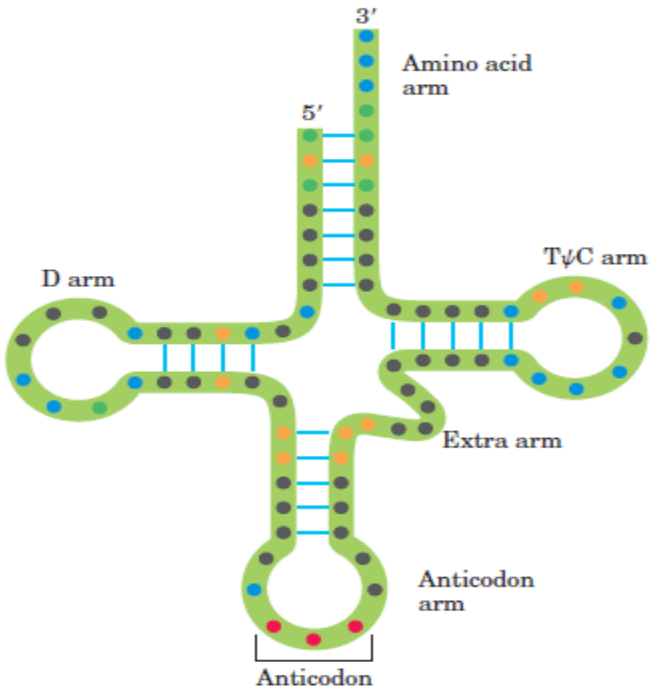
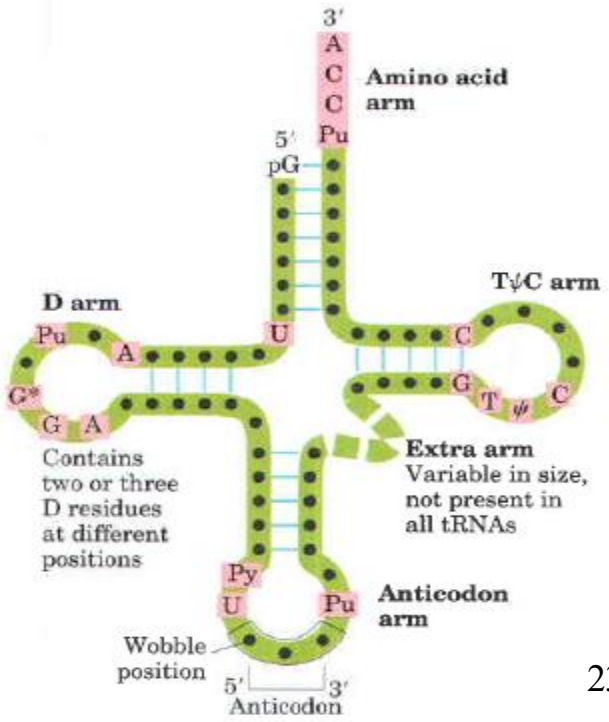


Figure 27-17
Lehninger Principles of Biochemistry, Seventh Edition

Recognition points of aminoacyl-tRNA synthetases

- Each tRNA molecule has a triplet anticodon on one end and an amino acid attachment site on the other
- **Black dots** are same in all tRNAs
- **Green & Blue Dots are the Recognition points** (orange for one) or (green for more) aminoacyl-tRNA synthetases.





23

Stage 1: Activation of Amino Acids

Aminoacyl-tRNA synthetases esterify the 20 amino acids to their corresponding tRNAs. Each enzyme is specific for one amino acid attachment to one or more corresponding tRNAs.

•*Two classes i.e. Class I and Class II: Based on substantial differences in reaction mechanism*

Amino acid + tRNA + ATP → aminoacyl-tRNA + AMP + Ppi

(1) *Activation*: it activates an amino acid for peptide bond formation

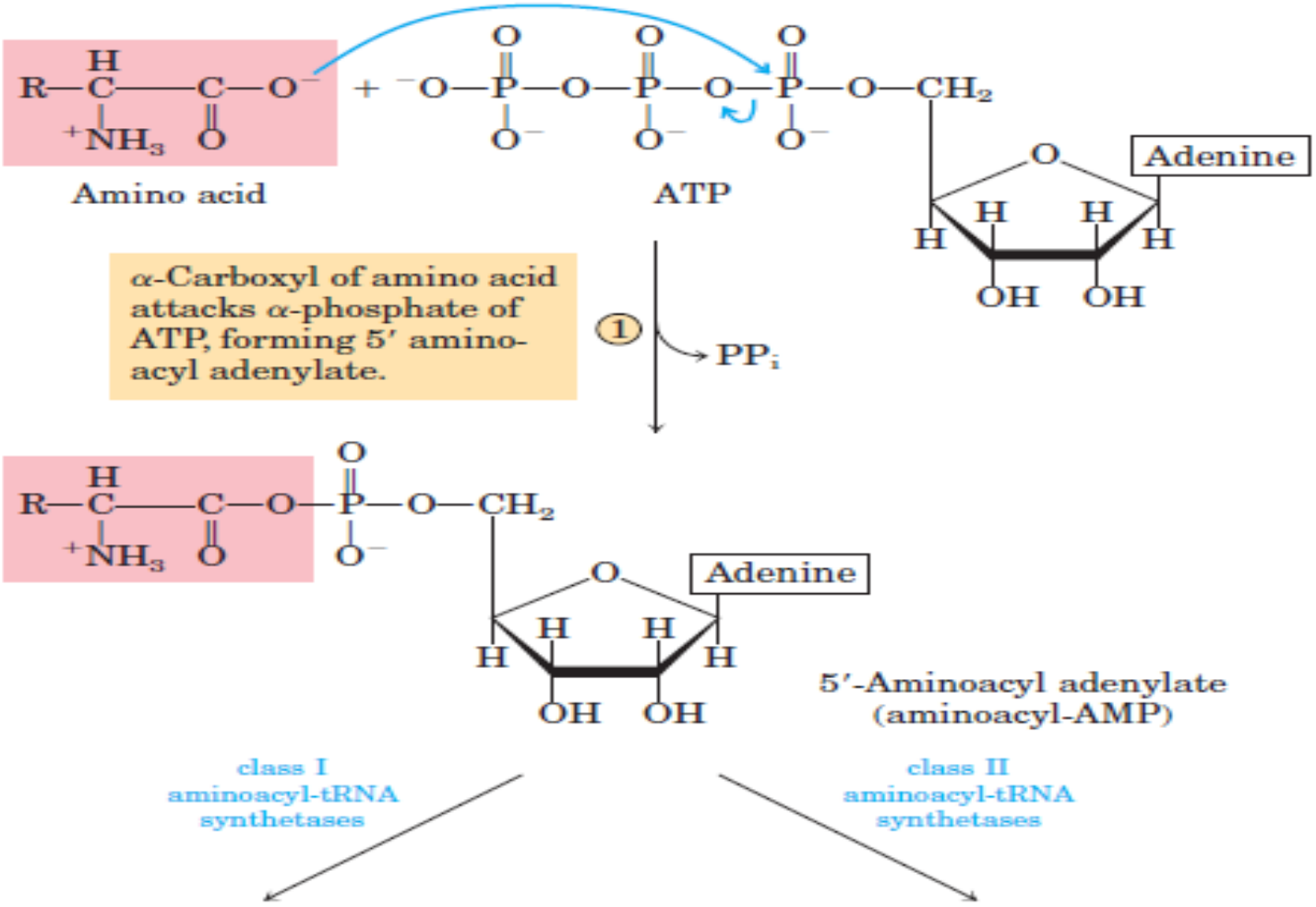
(2) *Proof reading*: it ensures appropriate placement of the amino acid in a growing polypeptide.

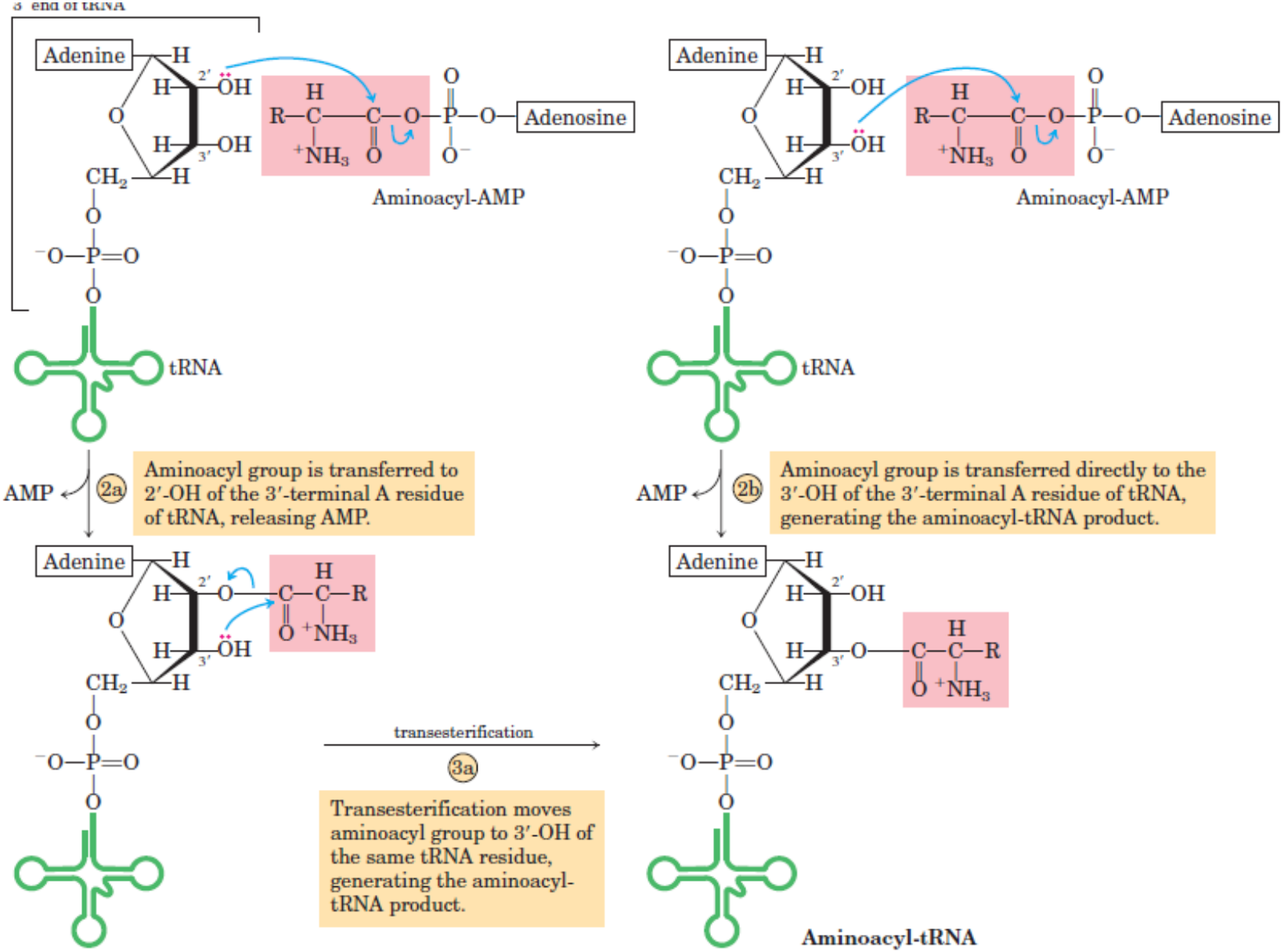
TABLE 27–7 The Two Classes of Aminoacyl-tRNA Synthetases

Class I		Class II	
Arg	Leu	Ala	Lys
Cys	Met	Asn	Phe
Gln	Trp	Asp	Pro
Glu	Tyr	Gly	Ser
Ile	Val	His	Thr

Note: Here, Arg represents arginyl-tRNA synthetase, and so forth. The classification applies to all organisms for which tRNA synthetases have been analyzed and is based on protein structural distinctions and on the mechanistic distinction outlined in Figure 27–19.

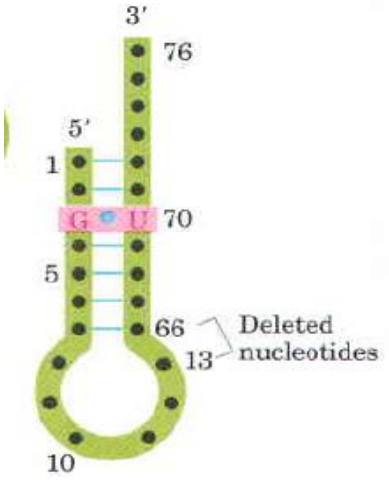
Activation of Amino Acids




3′ end of tRNA
Adenine
Adenosine
Aminoacyl-AMP
tRNA
AMP
2a
Aminoacyl group is transferred to 2′-OH of the 3′-terminal A residue of tRNA, releasing AMP.
2b
Aminoacyl group is transferred directly to the 3′-OH of the 3′-terminal A residue of tRNA, generating the aminoacyl-tRNA product.
transesterification
3a
Transesterification moves aminoacyl group to 3′-OH of the same tRNA residue, generating the aminoacyl-tRNA product.
Aminoacyl-tRNA

The Second Genetic Code

- Aminoacyl-tRNA synthetases *must be specific for both amino acid and tRNA.*
  - Matching each amino acid with correct tRNA can be viewed as the "second genetic code."
  - The "code" is in molecular recognition of a specific tRNA molecule by a specific synthetase.
- Only a few nucleotides in tRNA confer the binding specificity.
  - anticodon region
  - The primary determinant in Ala-tRNA is a single G=U in the amino acid arm.



Shine-Dalgarno sequence:

Consensus Sequence of 4-9 purine residue provide the initiation signals and binds to 16S rRNA

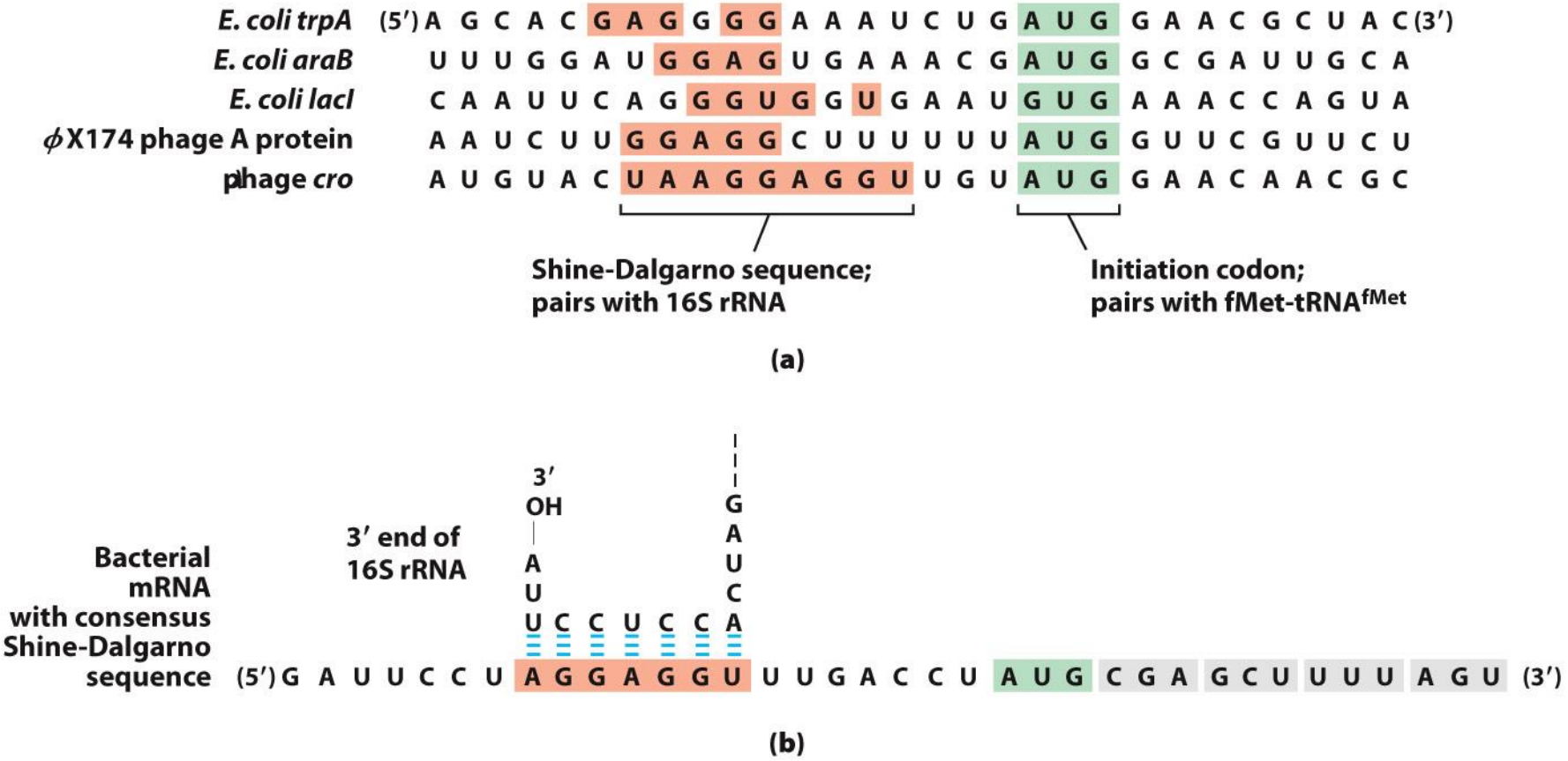


Figure 27-26
Lehninger Principles of Biochemistry, Seventh Edition

(named for Australian researchers John Shine and Lyrrn Dalgarno, who identified it)

Stage 2: Initiation

Bacterial Ribosomes have three sites that bind tRNAs,

Aminoacyl (A) site, the peptidyl (P) site, and exit (E) site

Step 1: The 30S ribosomal subunit binds two initiation factors, IF-1 and IF-3.

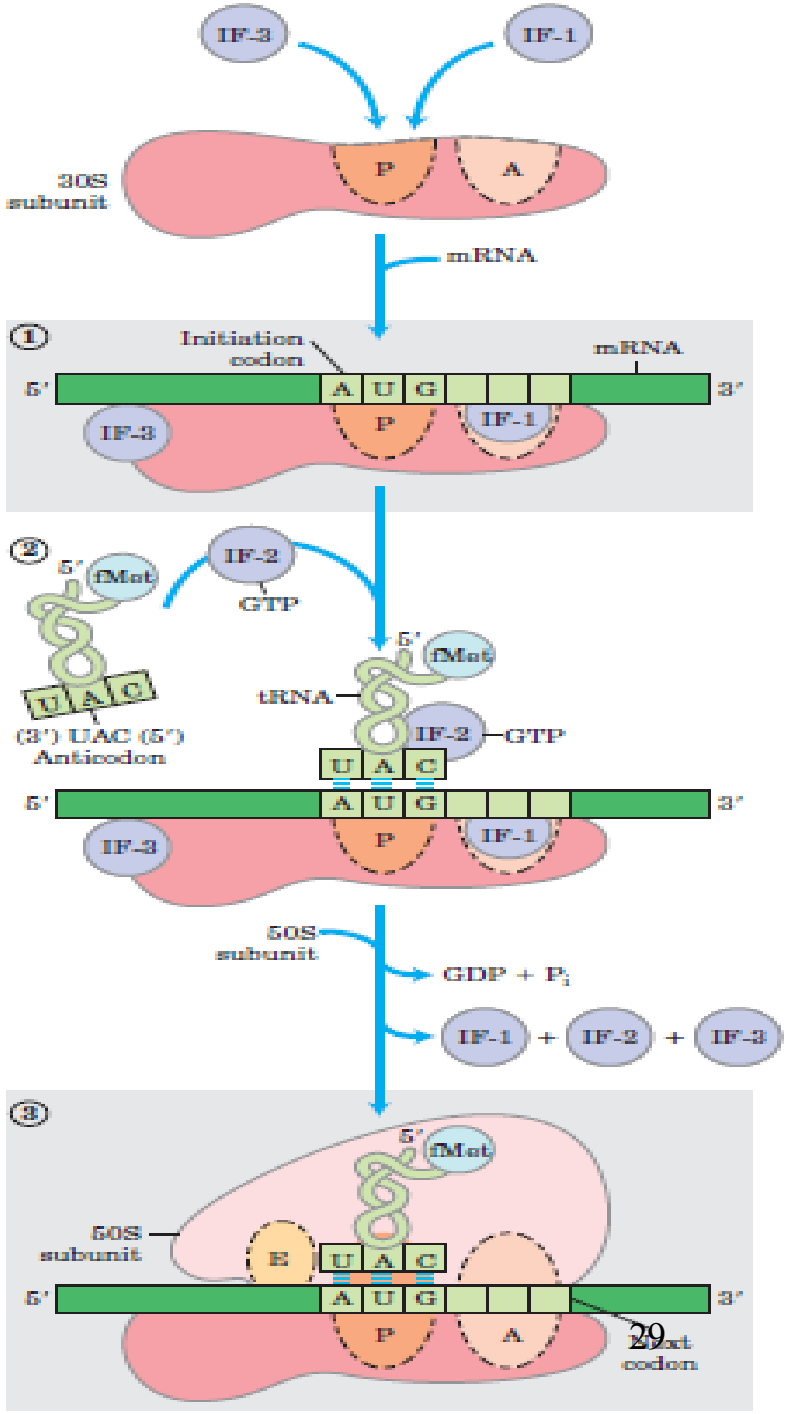
Factor IF-3 prevents the 30S and 50S subunits from combining prematurely.

Factor IF-1 binds at the A site and prevents tRNA binding at this site during initiation.

The mRNA then binds to the 30S subunit. The initiating (5)AUG is guided to its correct position by the ShineDalgarno sequence.

Step 2: The complex consisting of the 30S ribosomal subunit, IF-3, and mRNA is joined by both GTP-bound IF-2 and the initiating fMet-tRNAfMet. The anticodon of this tRNA now pairs correctly with the mRNA's initiation codon.

Step 3: This large complex combines with the 50S ribosomal subunit; simultaneously, the GTP bound to IF-2 is hydrolyzed to GDP and Pi , which are released from the complex. All three initiation factors depart from the ribosome at this point.



Initiation in Eukaryotes

Eukaryotic mRNAs are bound to the ribosome as a complex with a number of specific binding proteins.

1. At the 3 end, the mRNA is bound by the poly(A) binding Proteins (PAB).
2. The protein **eIF4A** has an RNA helicase activity removes secondary structure within the mRNA seq.
3. **eIF4F** complex (eIF4A, E, & G) and **eIF4B** scan mRNA from the 5 end until the first AUG is encountered.

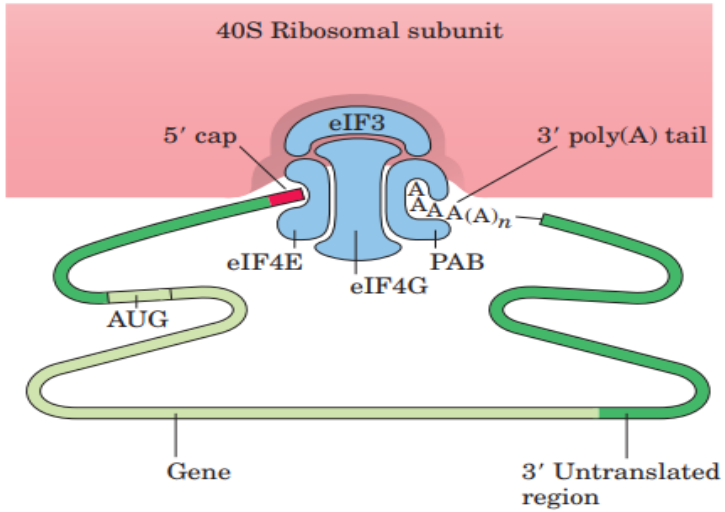


FIGURE 27–22 Protein complexes in the formation of a eukaryotic initiation complex. The 3′ and 5′ ends of eukaryotic mRNAs are linked by a complex of proteins that includes several initiation factors and the poly(A) binding protein (PAB). The factors eIF4E and eIF4G are part of a larger complex called eIF4F. This complex binds to the 40S ribosomal subunit.

TABLE 27–8 Protein Factors Required for Initiation of Translation in Bacterial and Eukaryotic Cells

Factor	*Function*
Bacterial	
IF-1	Prevents premature binding of tRNAs to A site
IF-2	Facilitates binding of fMet-tRNA^{fMet} to 30S ribosomal subunit
IF-3	Binds to 30S subunit; prevents premature association of 50S subunit; enhances specificity of P site for fMet-tRNA^{fMet}
Eukaryotic*	
eIF2	Facilitates binding of initiating Met-tRNA^{Met} to 40S ribosomal subunit
eIF2B, eIF3	First factors to bind 40S subunit; facilitate subsequent steps
eIF4A	RNA helicase activity removes secondary structure in the mRNA to permit binding to 40S subunit; part of the eIF4F complex
eIF4B	Binds to mRNA; facilitates scanning of mRNA to locate the first AUG
eIF4E	Binds to the 5′ cap of mRNA; part of the eIF4F complex
eIF4G	Binds to eIF4E and to poly(A) binding protein (PAB); part of the eIF4F complex
eIF5	Promotes dissociation of several other initiation factors from 40S subunit as a prelude to association of 60S subunit to form 80S initiation complex
eIF6	Facilitates dissociation of inactive 80S ribosome into 40S and 60S subunits

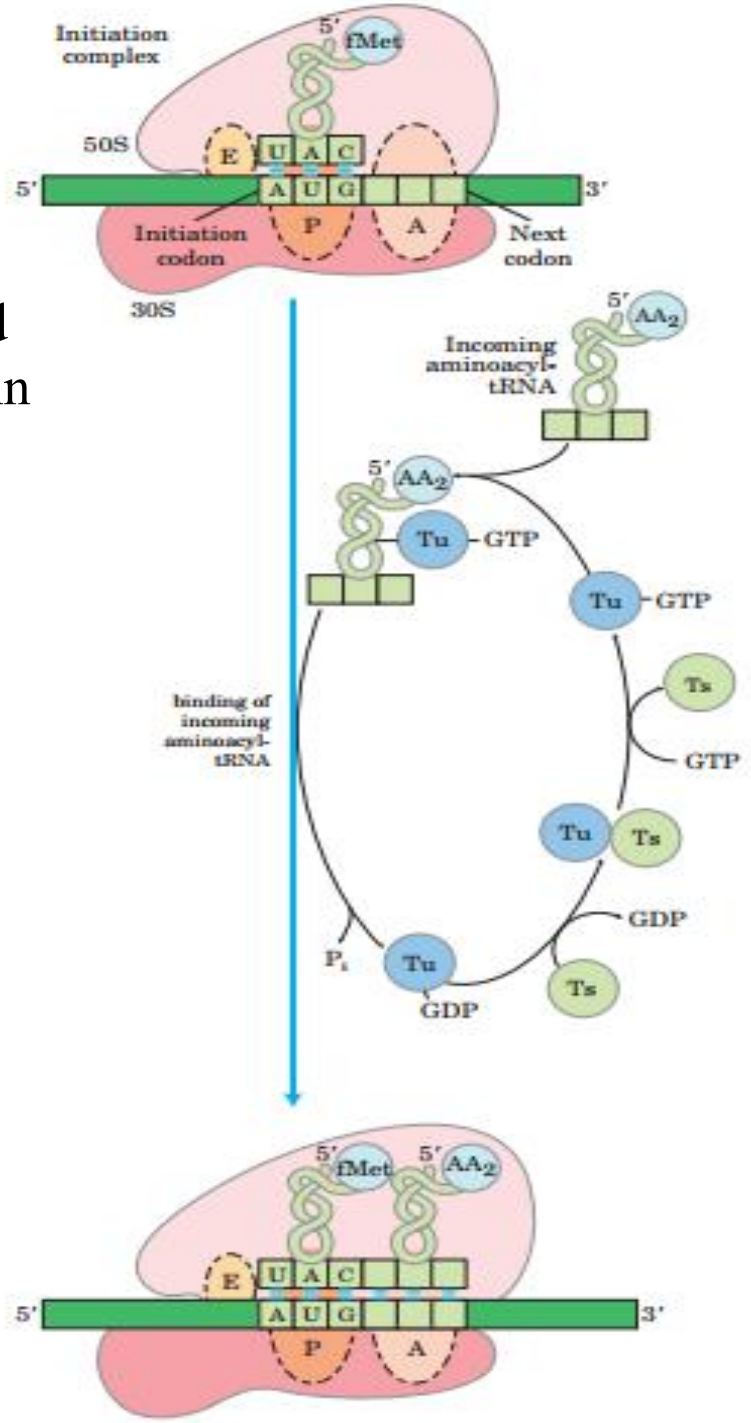
Stage 3: Elongation

Elongation requires

1. the initiation complex described above,
2. aminoacyl-tRNAs,
3. a set of three soluble cytosolic proteins called elongation factors (EF-Tu, EF-Ts, and EF-G in bacteria).
4. GTP.

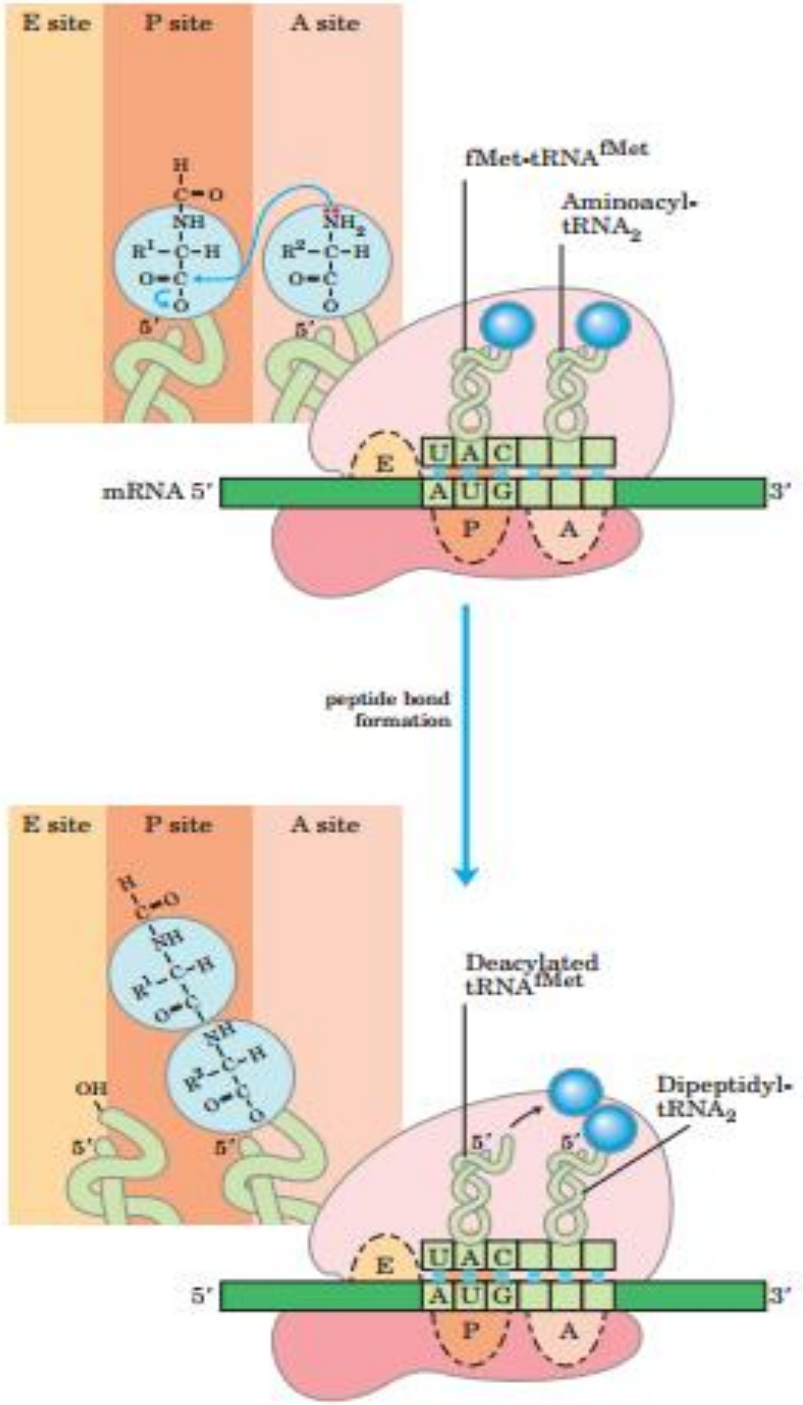
Elongation Step 1: Binding of an Incoming Aminoacyl-tRNA

a) The appropriate incoming aminoacyl-tRNA binds to a complex of GTP-bound EF-Tu.
b) The resulting aminoacyltRNA–EF-Tu–GTP complex binds to the A site of the 70S initiation complex.
c) The GTP is hydrolyzed and an EF-Tu–GDP complex is released from the 70S ribosome.
d) The EF-Tu–GTP complex is regenerated in a process involving EF-Ts and GTP.



Elongation Step 2: Peptide Bond Formation

a) A peptide bond is formed by the transfer of the initiating N-formylmethionyl group from its tRNA to the amino group of the second amino acid, now in the A site.
b) The amino group of the amino acid in the A site acts as a nucleophile, displacing the tRNA in the P site to form the peptide bond.
c) This reaction produces a dipeptidyl tRNA in the A site, and the now "uncharged" (deacylated) tRNAfMet remains bound to the P site.
d) The enzymatic activity that catalyzes peptide bond formation has historically been referred to as ***peptidyl transferase (***assumed to be intrinsic to one or more of the proteins in the large ribosomal subunit). We now know that this reaction is catalyzed by the 23S rRNA adding to the known catalytic property of ribozymes.



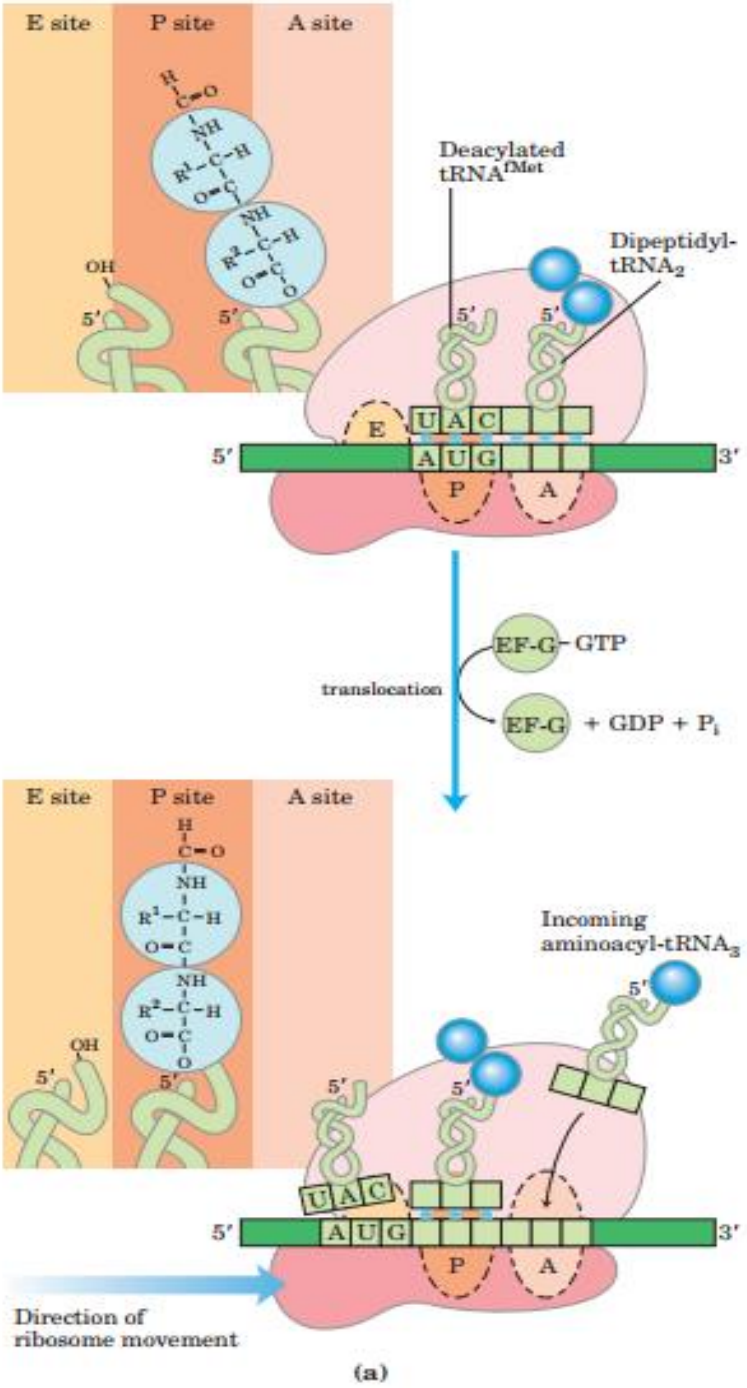
Elongation Step 3: Translocation

a) The ribosome moves one codon toward the 3 end of the mRNA. This movement shifts the anticodon of the dipeptidyl tRNA, which is still attached to the second codon of the mRNA, from the A site to the P site, and shifts the deacylated tRNA from the P site to the E site, from where the tRNA is released into the cytosol.

b) The third codon of the mRNA now lies in the A site and the second codon in the P site.

c) Movement of the ribosome along the mRNA requires ***EF-G (also known as translocase)*** and the energy provided by hydrolysis of another molecule of GTP.

d) A change in the three-dimensional conformation of the entire ribosome results in its movement along the mRNA. Because the structure of EF-G (bind to A site and displace the peptidyl-tRNA) mimics the structure of the EF-Tu–tRNA complex.

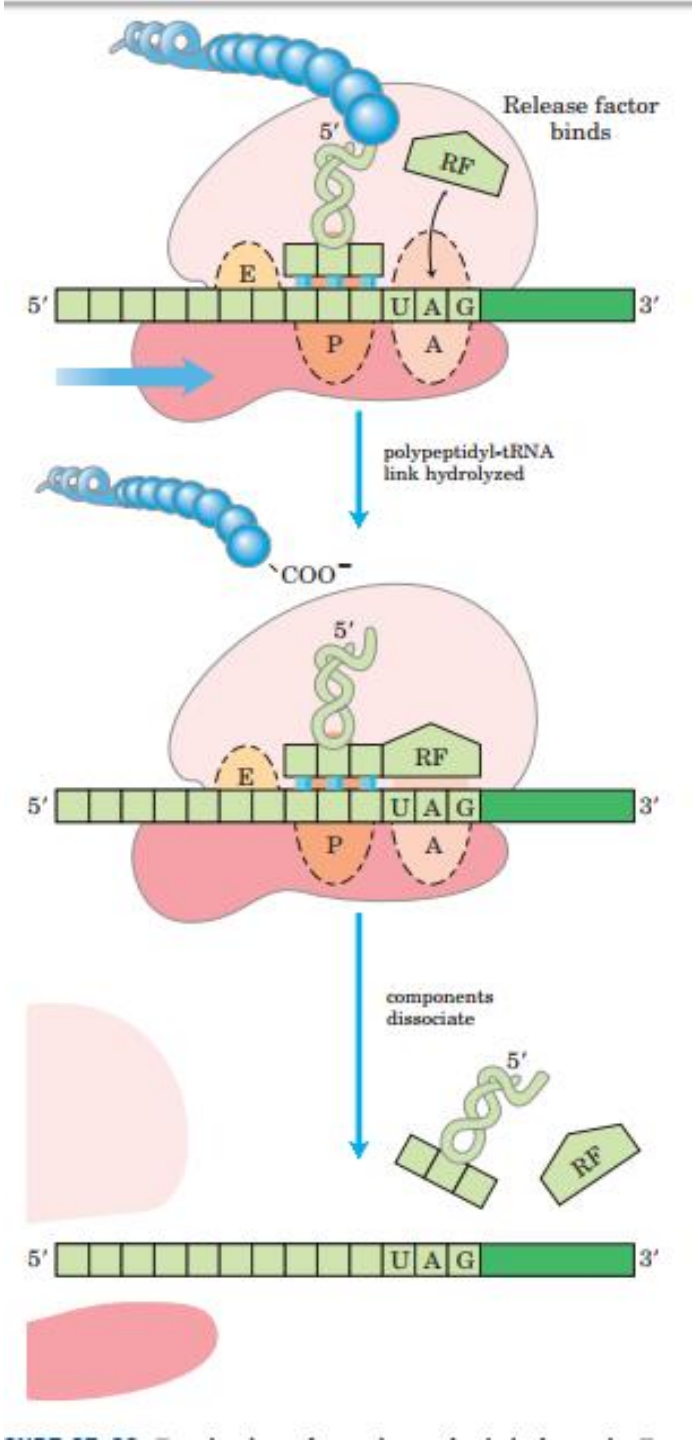


Stage 3: Elongation (Eukaryotic)

- Steps are similar to bacteria
- Elongation factors – eEF1α (EF-Tu), eEF1$\beta\gamma$ (EF-Ts), eEF2 (EF-G)
- Difference: Eukaryotic ribosomes do NOT have an E site; the uncharged tRNAs are released from the P site.

Stage 4:Termination

- Termination occurs in response To a termination codon In the A site.
- First, a release factor, (RF-1or RF-2,depending on which termination Codon is present) binds to the A site.
- This leads to hydrolysis of the ester linkage between the nascent polypeptide and the tRNA in the P site and release of the completed polypeptide.
- Finally the mRNA, deacylated tRNA, and release factor leave the ribosome which dissociates in to its 30S and 50S subunits aided by ribosome recycling factor (RRF). IF-3, and energy. provided by EF-G-mediated GTP hydrolysis.
- The 30S subunit complex with IF-3 is ready to begin another cycle of translation.



Features of Protein Synthesis

- Large energy cost (122 kJ/mol of phosphodiester bond energy, to generate a peptide bond, which has a standard free energy of hydrolysis of about -21 kJ/mol)
- Can be rapid when accomplished on clusters of ribosomes called a polysome
- In bacteria, tightly coupled to transcription
  - Translation can begin before transcription is finished.

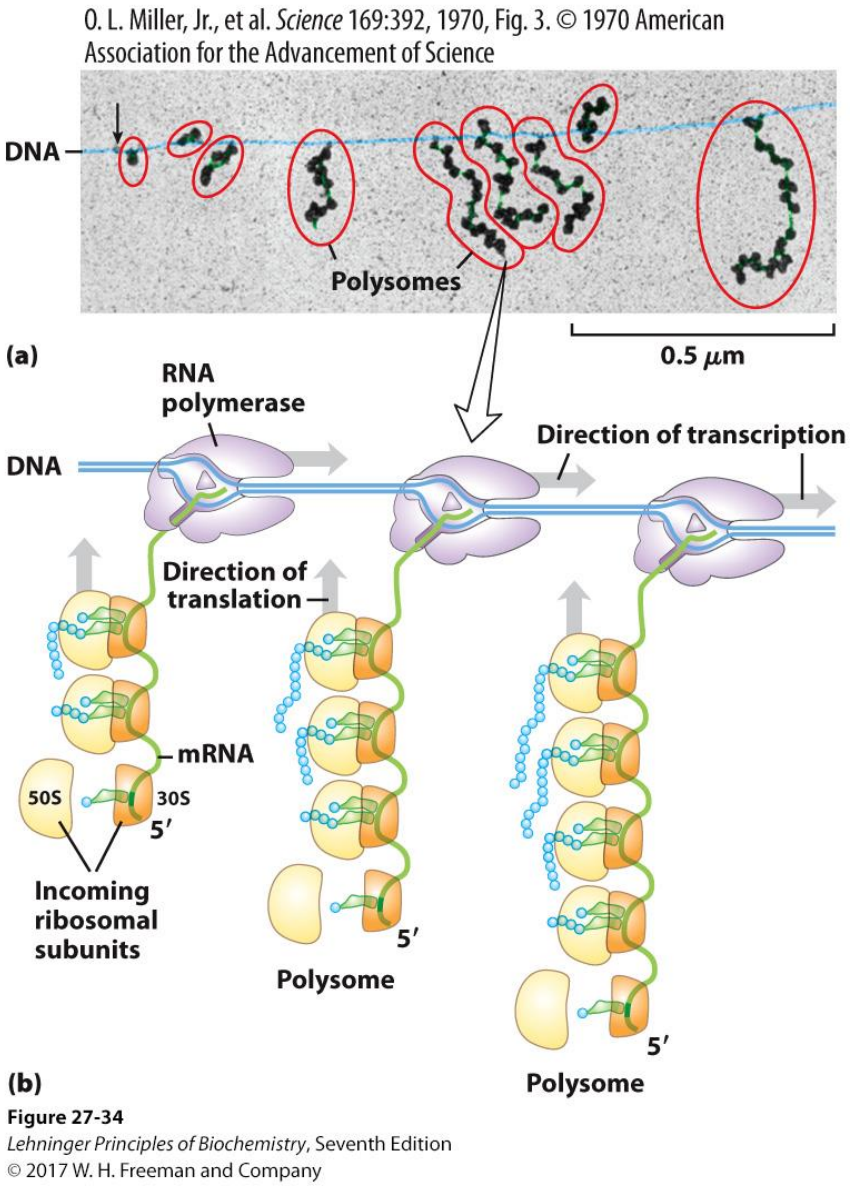


Figure 27-34
Lehninger Principles of Biochemistry, Seventh Edition

Posttranslational Modifications

Newly Synthesize polypeptide chains Undergo Folding and Processing

1. Amino-Terminal and Carboxy-Terminal Modifications
   - Removal of Formyl-Methionine and Methionine
   - Removal of Additional Amine Terminal
   - Removal of Additional Carboxy Terminal
   - Acetylation at N-terminal (50% of eukaryotic proteins)
2. Removal of Signal Sequence (Direct the movement of proteins in ER)
3. Phosphorylation: ex: Casien
4. Attachment of Carbohydrate Side Chains
5. Addition of Isoprenyl Groups
6. Addition of Prosthetic Groups *(Biotin in carboxylases, heme in hemoglobin and cytochromes)*
7. Proteolytic Processing
8. Formation of Disulfide Cross-Links

Posttranslational Modifications

- Modifying amino acids with additional phosphates, carboxylic acid groups, and so on

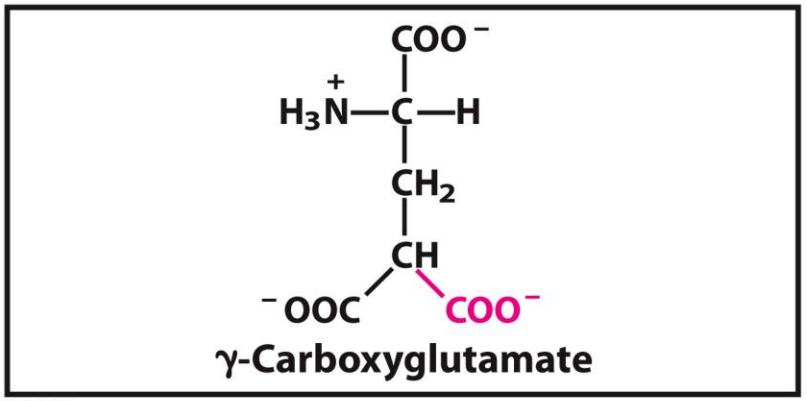


Figure 27-35b
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

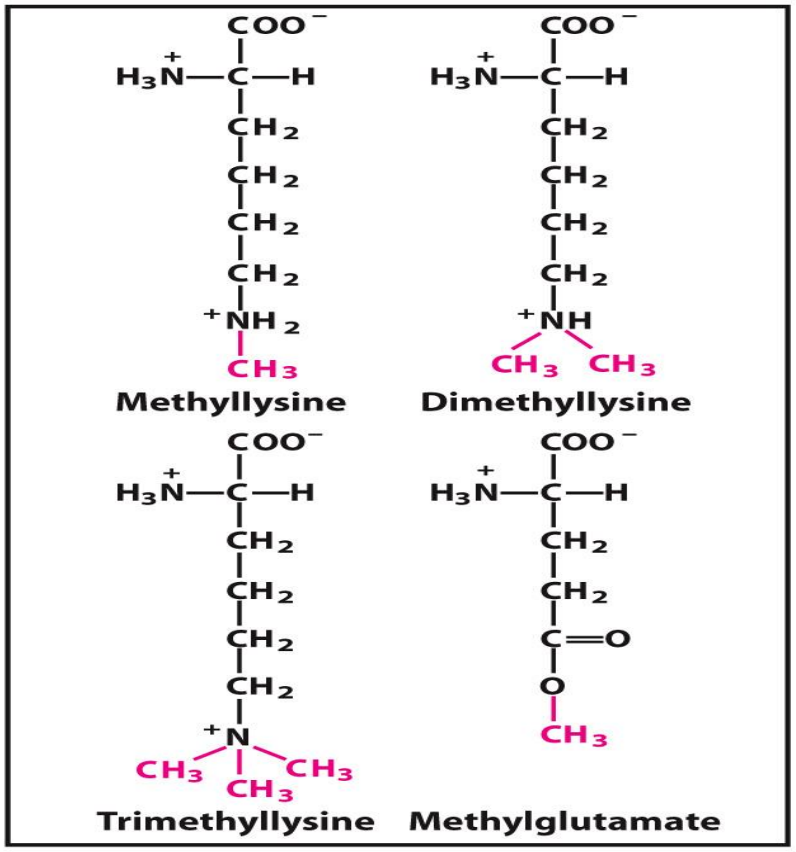


Figure 27-35c
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

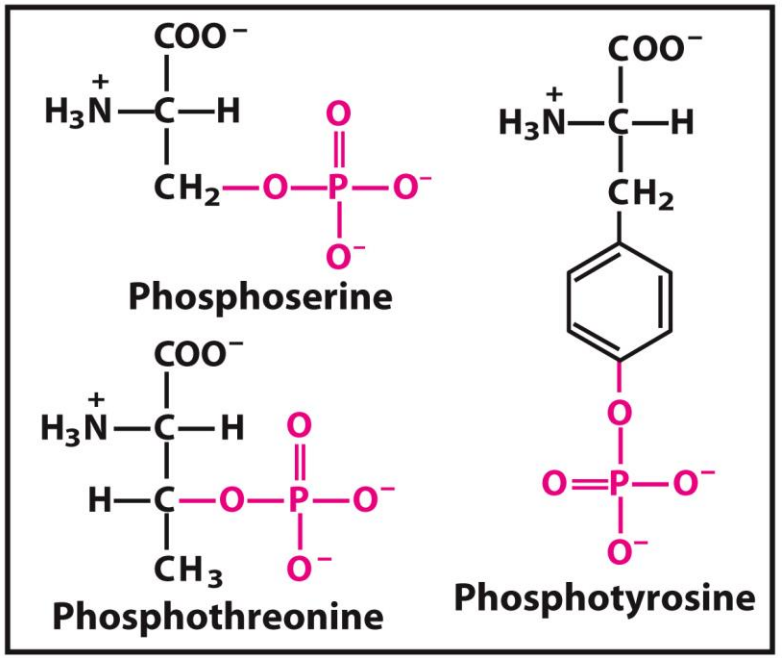


Figure 27-35a
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

More Modified Amino Acids

Addition of isoprenyl groups (such as farnesyl pyrophosphate) from intermediates of cholesterol synthesis pathway

Isoprene group helps anchor proteins in membranes for carcinogenic activity.

Product of *ras* oncogene

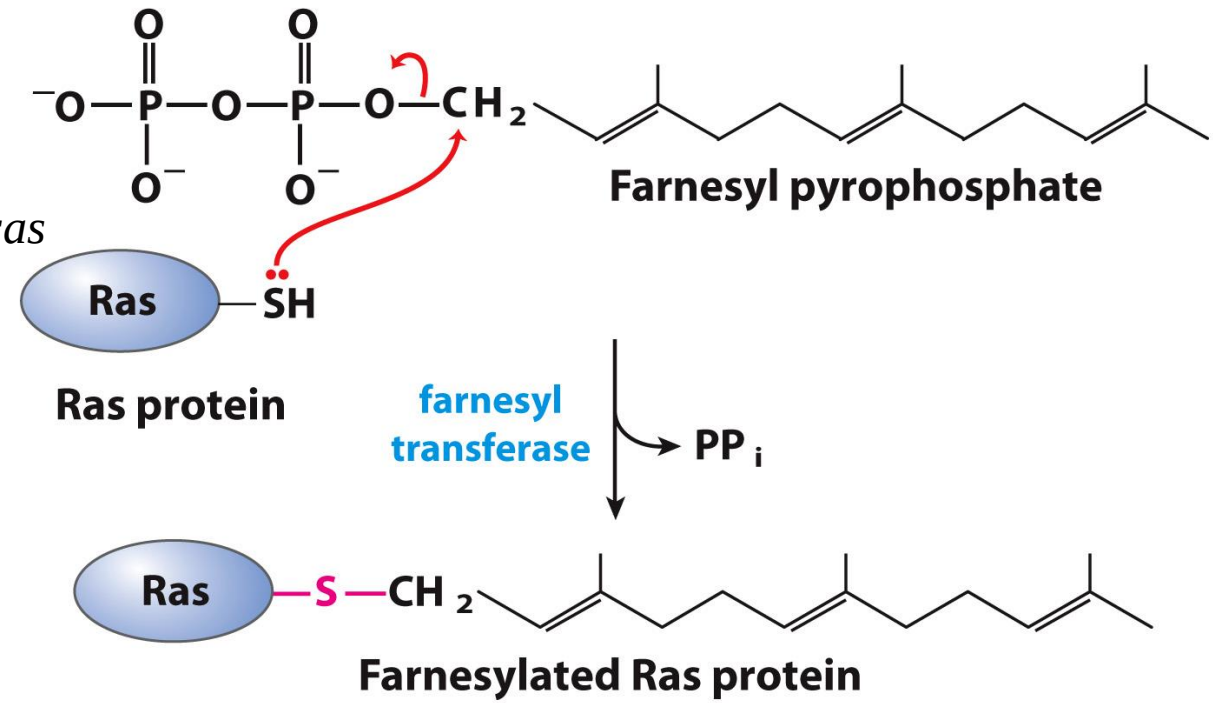
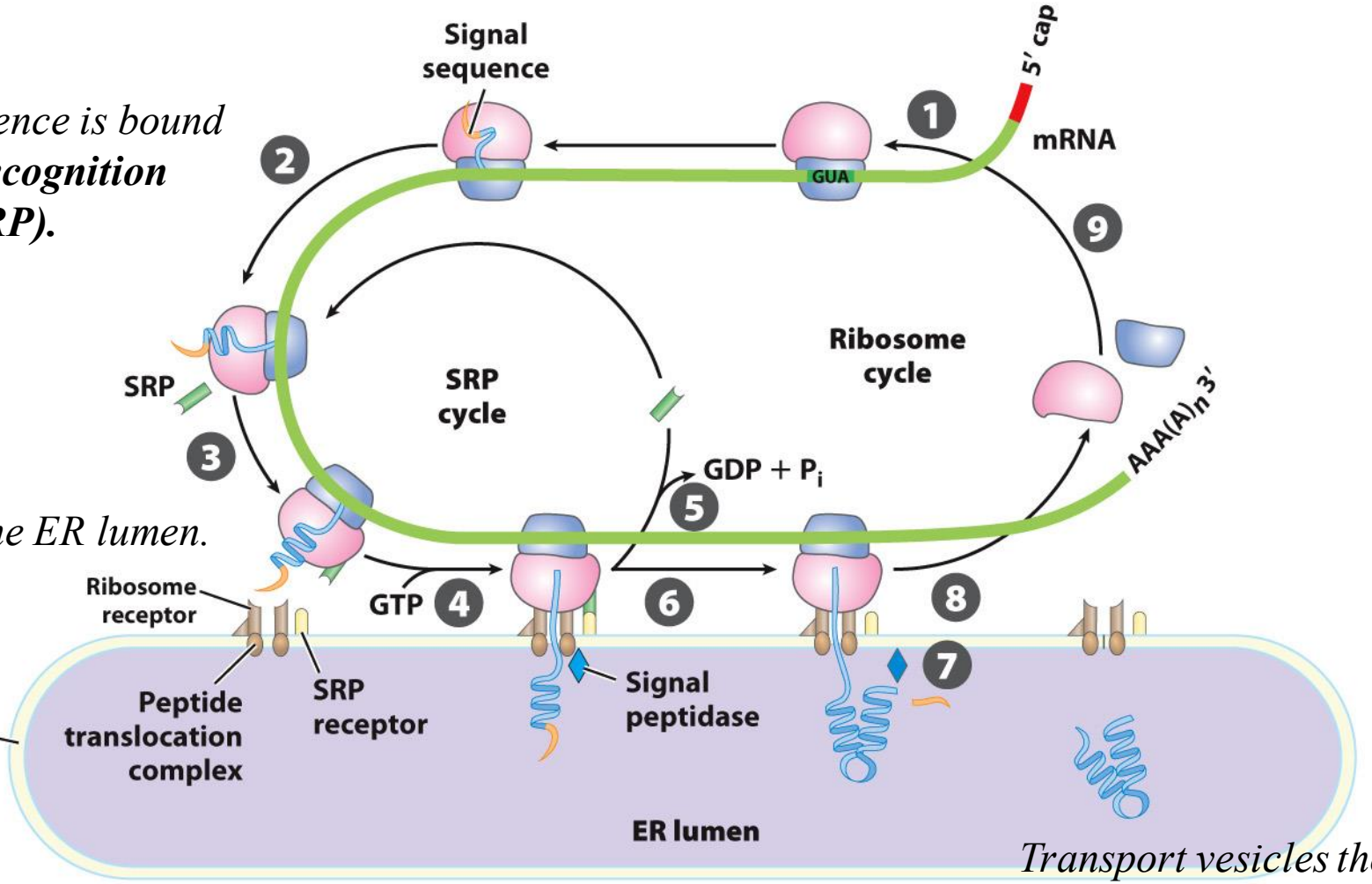


Figure 27-36
Lehninger Principles of Biochemistry, Seventh Edition

Directing Peptides to the ER and Beyond by signal sequence

*signal sequence is bound by **signal recognition particle (SRP).***



delivered to the ER lumen.

modification takes place in ER (i.e. glycosylation)

Transport vesicles then take proteins to Golgi apparatus for final dispatch.

Figure 27-40
Lehninger Principles of Biochemistry, Seventh Edition

Glycosylation of Proteins

- Glycoproteins form by linking an oligosaccharide to a side group on the peptide.
- N-linked (nitrogen of Asn or Arg)
  - begins in the RER

- O-linked (oxygen of Ser, Thr, Tyr)
  typically occurs in the golgi or cytosol for proteins not entering the ER
- Proteins are sorted in the Golgi and sent on to their final destinations.

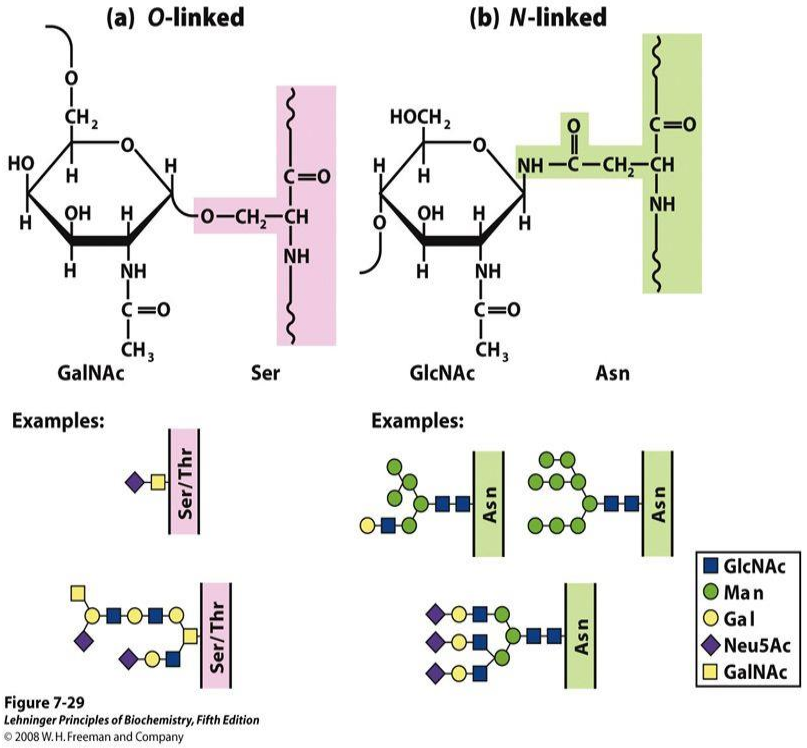


Figure 7-29
Lehninger Principles of Biochemistry, Fifth Edition

Glycosylation of Proteins

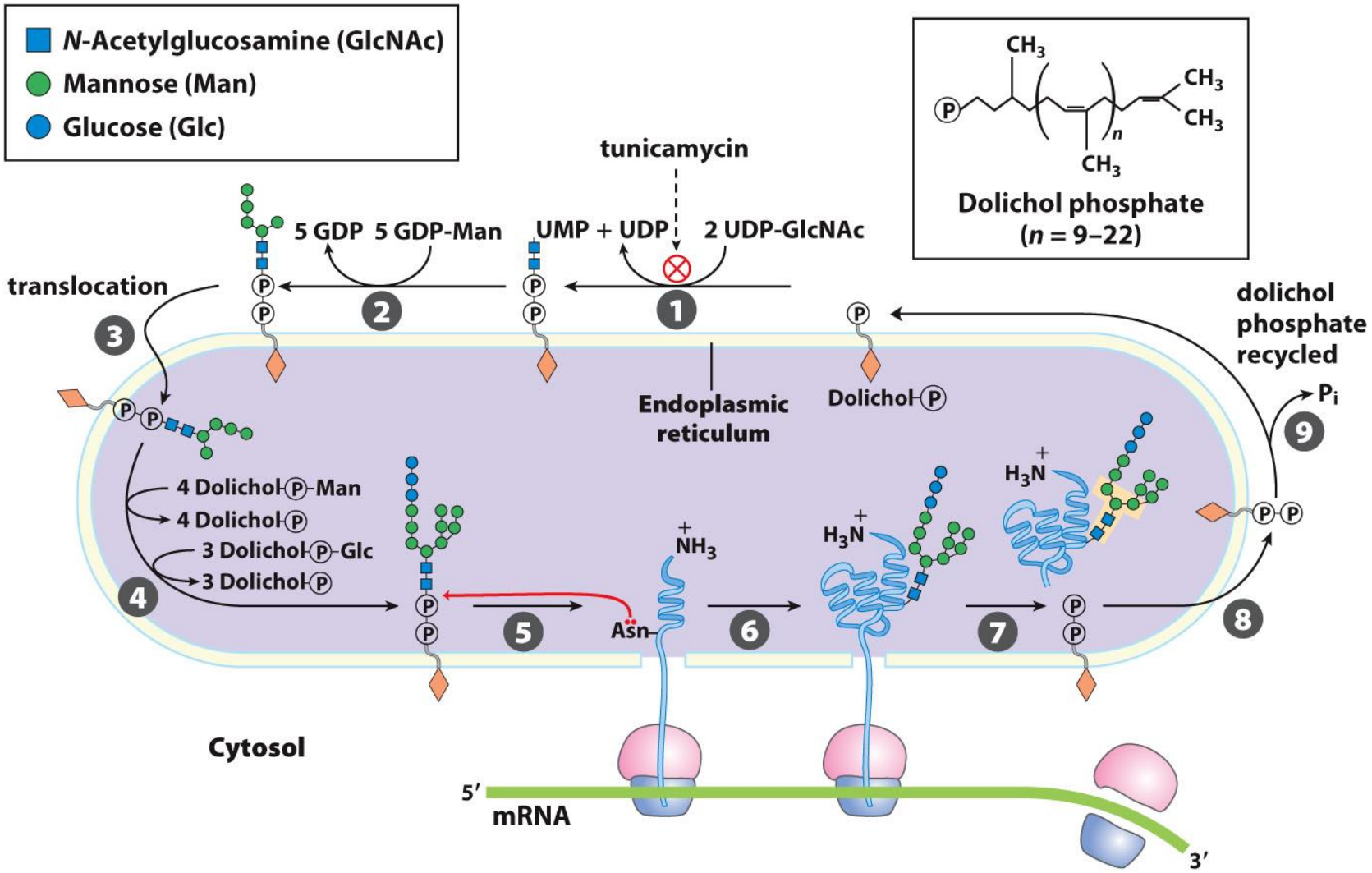


Figure 27-41
Lehninger Principles of Biochemistry, Seventh Edition