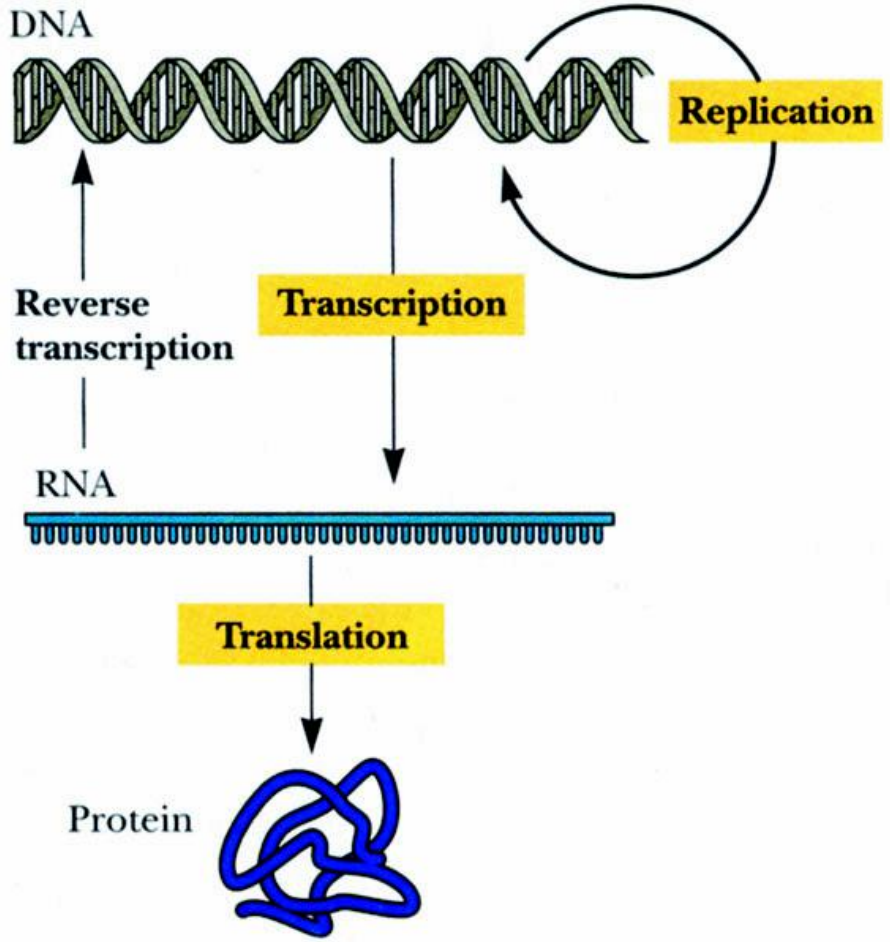
Protein metabolism

Translation:

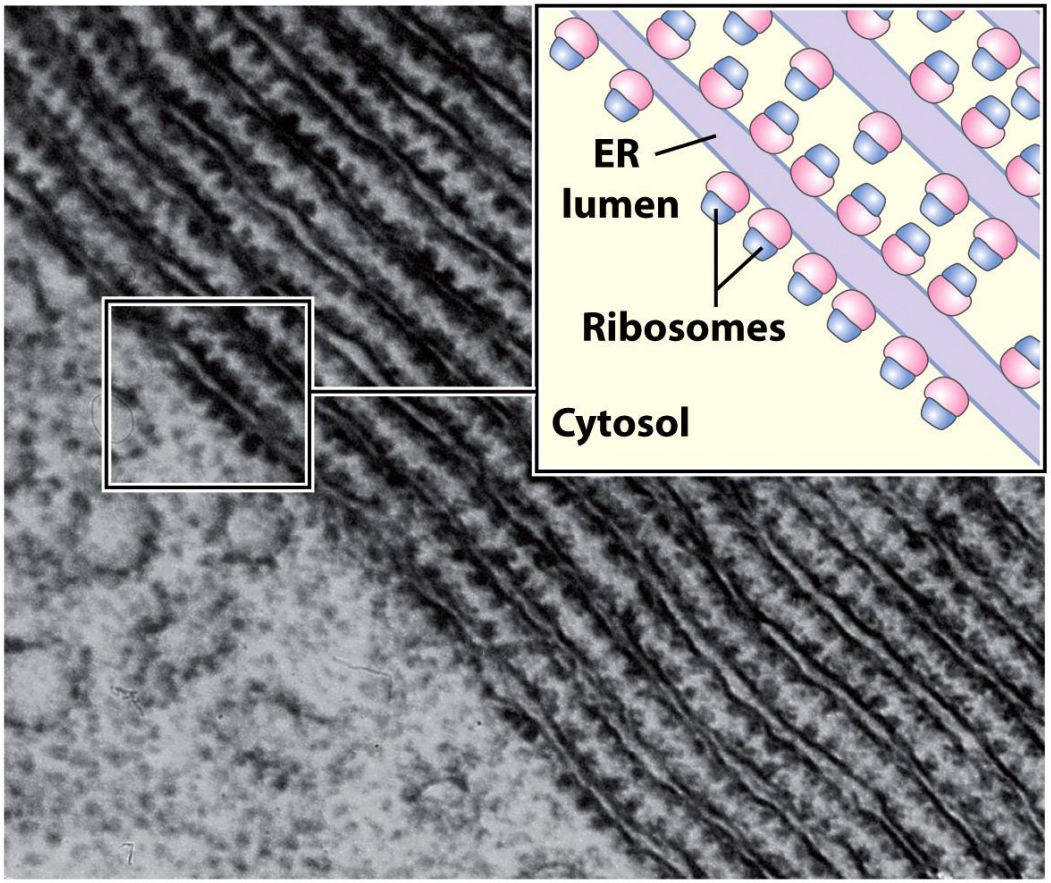
The process in which the genetic information present in an mRNA molecule specifies the sequence of amino acids during protein synthesis.



Central dogma of molecular biology

Protein synthesis can account for up to 90% of the chemical energy used by a cell for all biosynthetic reactions.

Location of protein synthesis:
ribosomes and endoplasmic reticulum




Paul Zamecnik

1950s, by injecting radioactive amino acids into rats

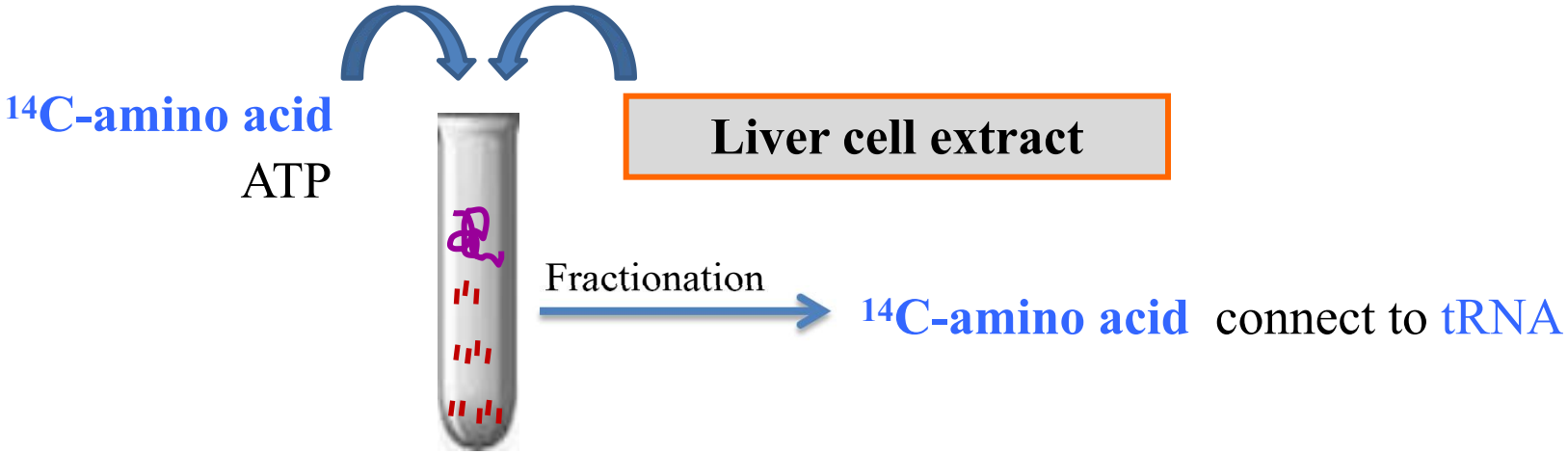
Ribosomes are the site of protein synthesis from amino acids

Discovery of aminoacyl-tRNA


Paul Zamecnik

In 1956, Mahlon Hoagland & Zamecnik using ^{14}C-amino acids in cell-free system found amino acids were attached to a heat-stable soluble RNA, later named transfer RNA (tRNA).

Amino acid + tRNA + ATP → Aminoacyl-tRNA + AMP + 2Pi



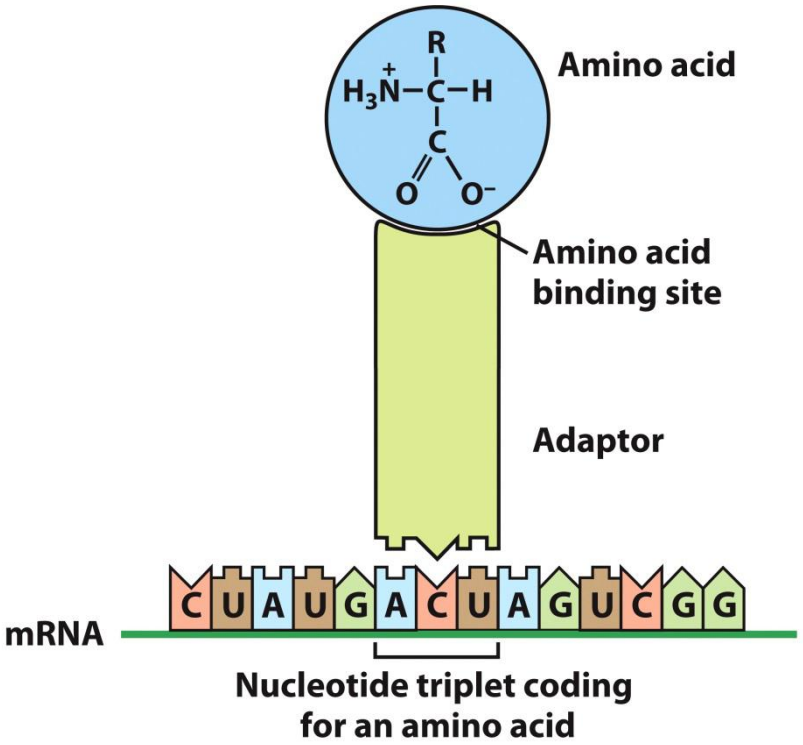
Question:

How is the 4-letter language of nucleic acid translated into the 20-letter language of proteins? What are the adapter molecules?

Crick's adaptor hypothesis

The adapter molecules must bridge the information gap from mRNAs to proteins:

1. must interact specifically with both nucleic acids and proteins (amino acids);
2. be able to read the genetic code in mRNA template and align the amino acids according to the template's direction;
3. at least 20 different adapter molecules are needed.



The genetic code

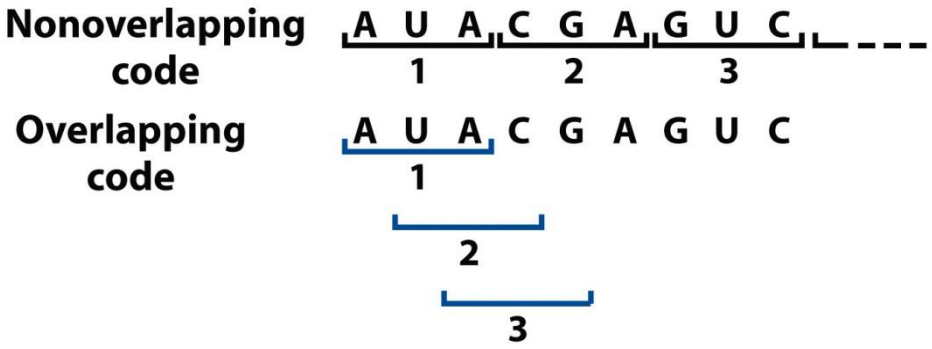
Codon (密码子): the set of triplet nucleotides in mRNA (or DNA) coding for the amino acids of proteins.

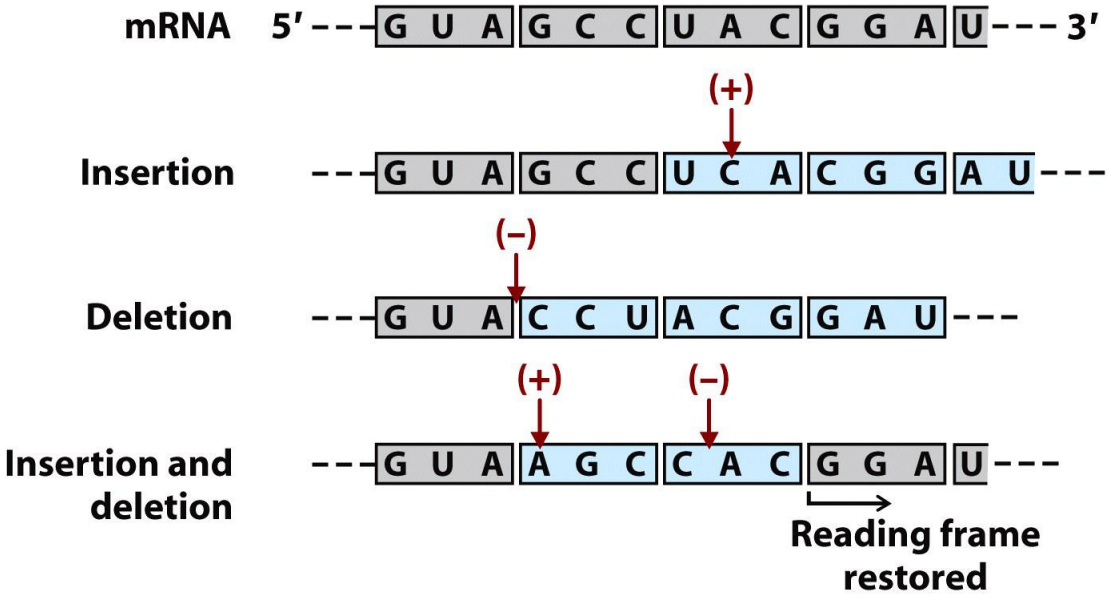
The four code letters of DNA (A, T, G, and C) in groups of two can yield only $4^2=16$ different combinations, insufficient to encode 20 amino acids. Groups of three, however, yield $4^3=64$ different combinations.

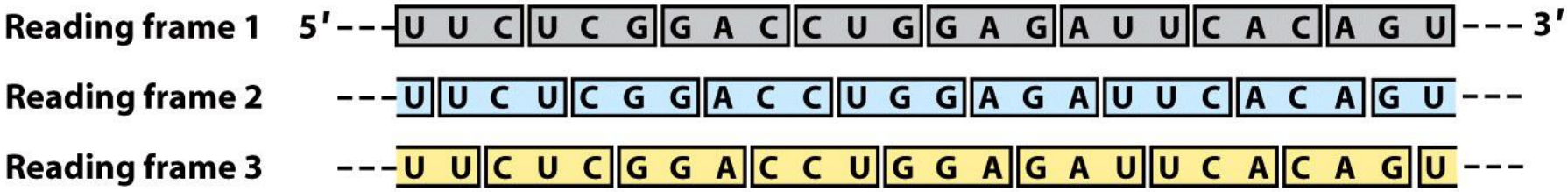
Reading frames in the genetic code

1. A group of three bases codes for one amino acid.

2. The code is not overlapping.





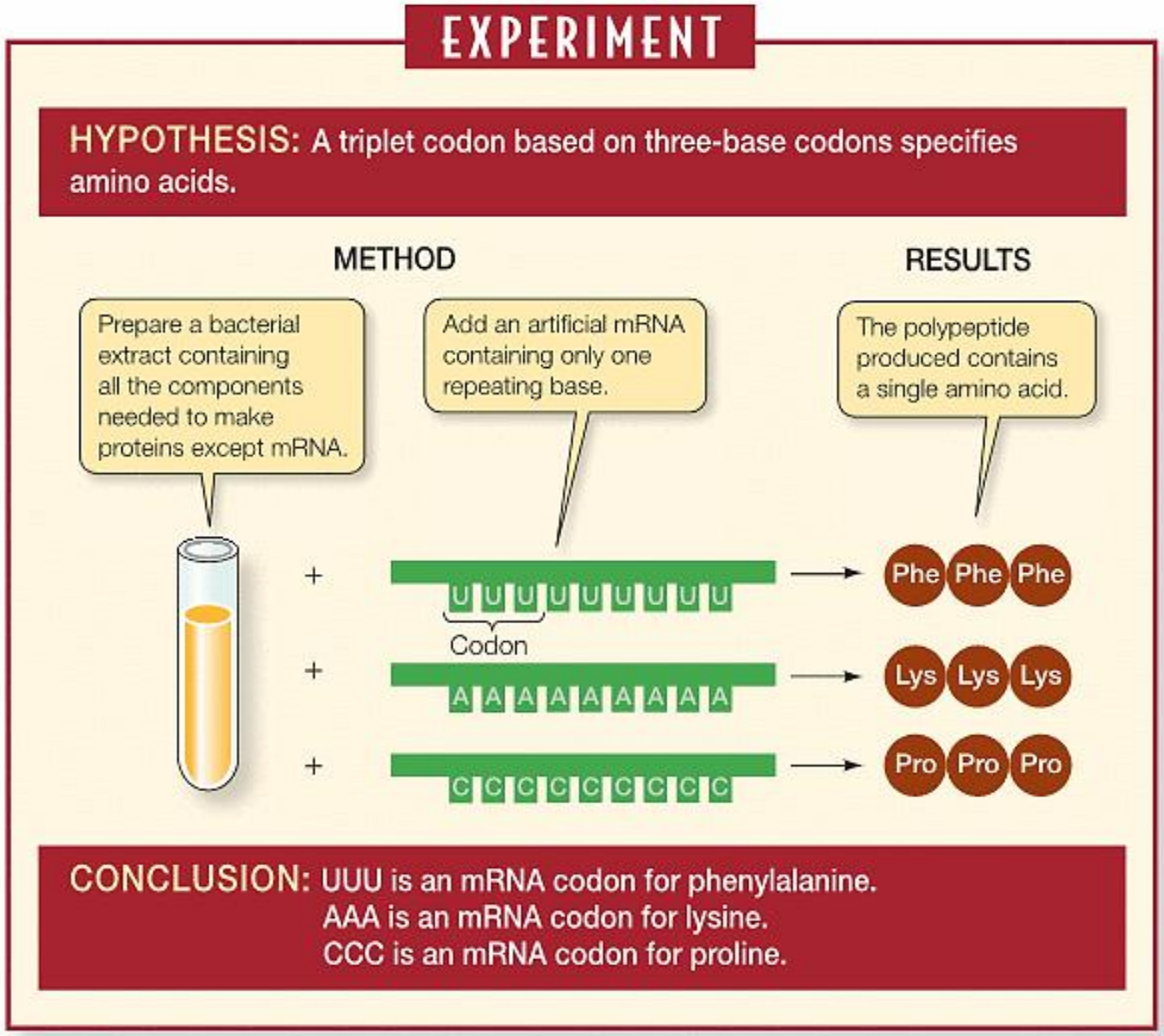


In principle, any given single-stranded DNA or mRNA sequence has three possible reading frames. Each reading frame gives a different sequence of codons, but only one is likely to encode a given protein.

Question:

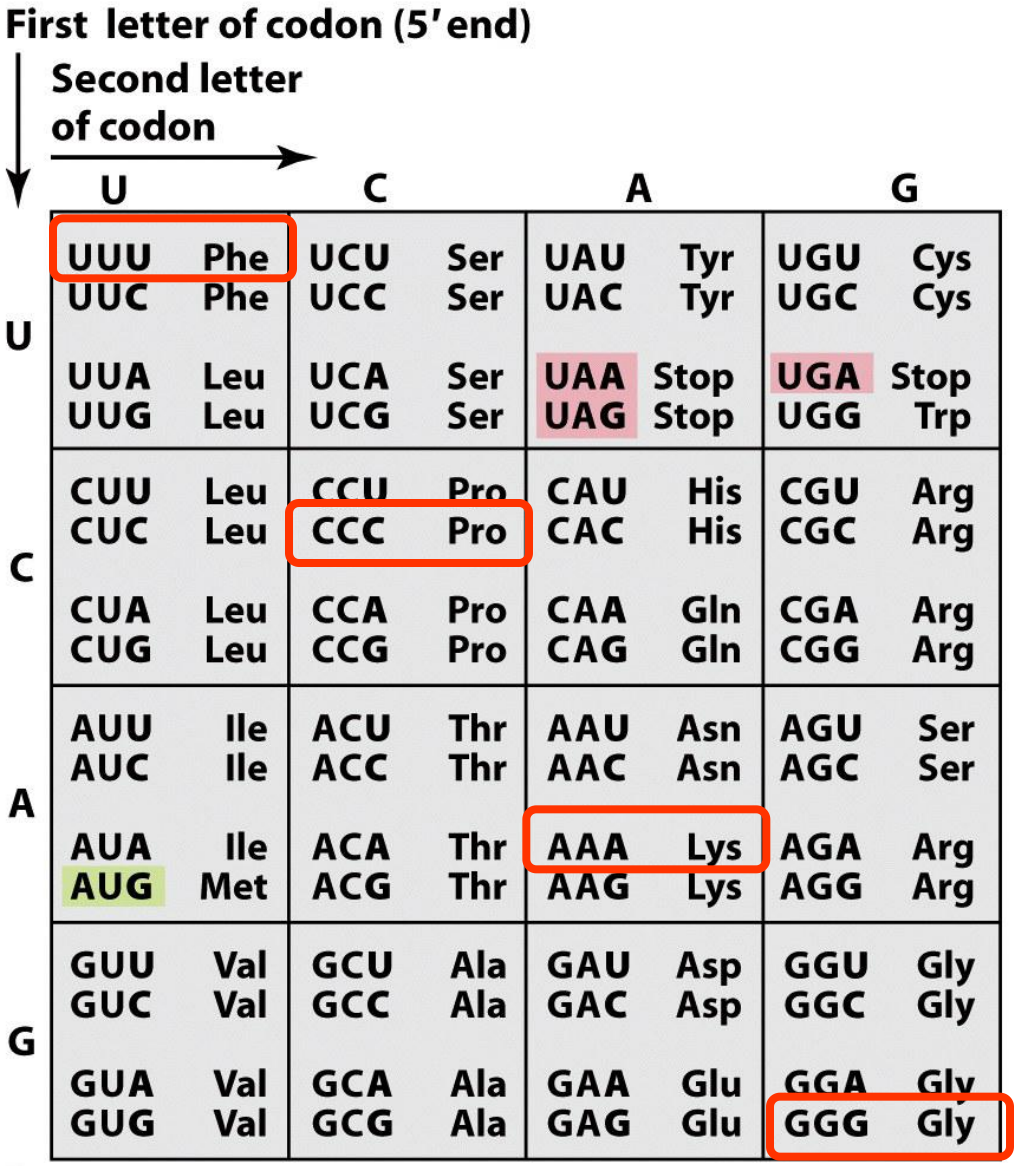
What are the three-letter code words for each amino acid?

Elucidating the genetic code



Marshall Nirenberg

Crack the genetic code



Additional templates

- Nirenberg and his colleagues got synthetic polynucleotides by randomly joining nucleotides using polynucleotide phosphorylase.

$$(\text{NMP})_n + \text{NDP} \rightleftharpoons (\text{NMP})_{n+1} + \text{P}_i$$

Lengthened polynucleotide

- For example, using ATP & CTP (5:1) to get AAA, AAC, ACC, ACA, CAA, CCA, CAC, and CCC.

TABLE 27–1 **Incorporation of Amino Acids into Polypeptides in Response to Random Polymers of RNA**

Amino acid	Observed frequency of incorporation (Lys = 100)	Tentative assignment for nucleotide composition of corresponding codon*	Expected frequency of incorporation based on assignment (Lys = 100)
Asparagine	24	A_2C	20
Glutamine	24	A_2C	20
Histidine	6	AC_2	4
Lysine	100	AAA	100
Proline	7	AC_2, CCC	4.8
Threonine	26	A_2C, AC_2	24

Synthetic template


H. Gobind Khorana

- 1965, H. Gobind Khorana synthesized repeating sequences of 2 to 4 bases
- CA+CA+CA+……→ CACACACACACACACA

histidine & threonine

Already known that codon for histidine contains 2 C, therefore must be CAC for histidine and ACA for threonine.

"Dictionary" of amino acid code words in mRNAs

First letter of codon (5′ end)

Second letter of codon

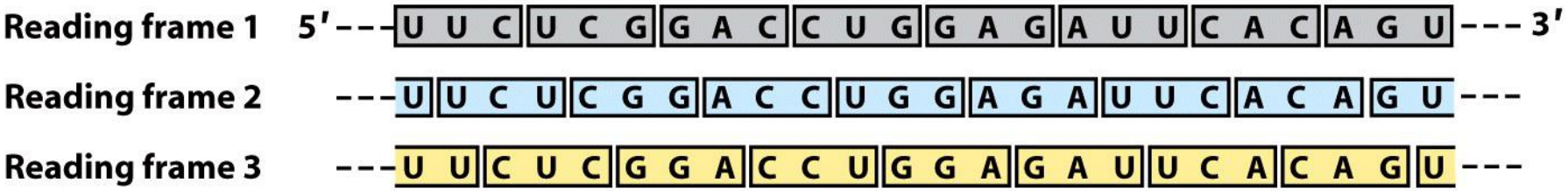
	U		C		A		G	
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
	UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop
	UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
A	AUU	Ile	ACU	Thr	AAU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

Meanings for all the triplet codons were established by 1966.

The cracking of the genetic code is regarded as one of the most important scientific discoveries of the twentieth century.

Start codon: AUG
Stop codons: UAA, UAG, UGA

Reading frame



- **Reading frame**: the way of breaking a sequence of nucleotides in DNA or RNA into three letter codons which can be translated into amino acid sequence
- **Frameshift**: reading from a different frame
- How to ensure correct frame? **Start codon**
- **Open reading frame**: DNA sequence that does not contain a **stop codon** in a given *reading frame*

The nature of the genetic code

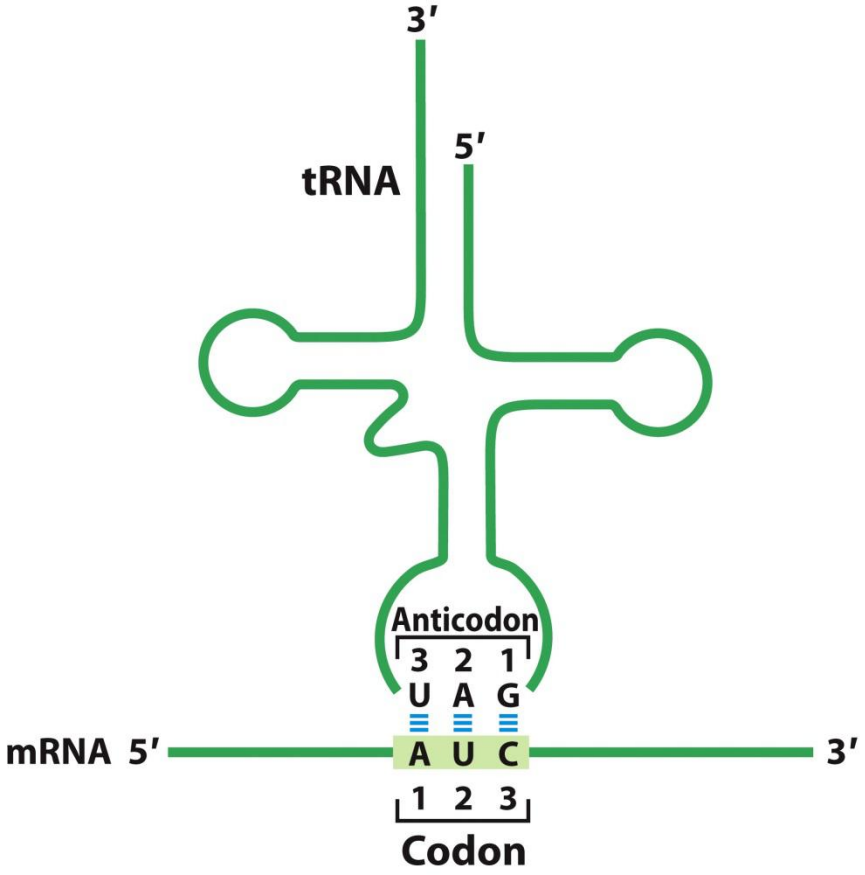
1. Codons are read 5'→3', they represent triplets of bases in mRNA.
2. 61 of the 64 codons specify particular amino acids, and the remaining 3 are termination codons (nonsense codons: **UAG**, **UAA**, & **UGA**).
3. Each of the 61 "sense" codons encodes only one amino acid.
4. The genetic code is degenerate. With the exception of Met and Trp, every amino acid is coded by more than one codon. Codons coding for the same amino acid are called synonymous codons.
5. Codons representing the same amino acid or chemically similar amino acids tend to be similar in sequence. Often the third base in a codon is irrelevant, this feature is known as third-base degeneracy.
6. The genetic code is universal. Codon assignments are virtually the same throughout all organism.

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

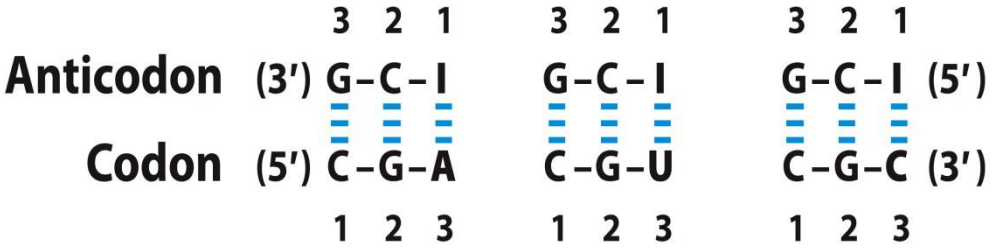
Degeneracy of the genetic code is to reduce the deleterious effect of mutations on organisms, because single base changes in a codon might result either in no change or in a substitution with an amino acid similar to the original amino acid.

Pairing relationship of codon and anticodon



Anticodon: three-base sequence on the tRNA for paring with mRNA.

Codon-anticodon pairs in antiparallel fashion.



Three different codon pairing relationships are possible when the tRNA anticodon contains inosinate (次黄嘌呤核苷酸）

Wobble hypothesis

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

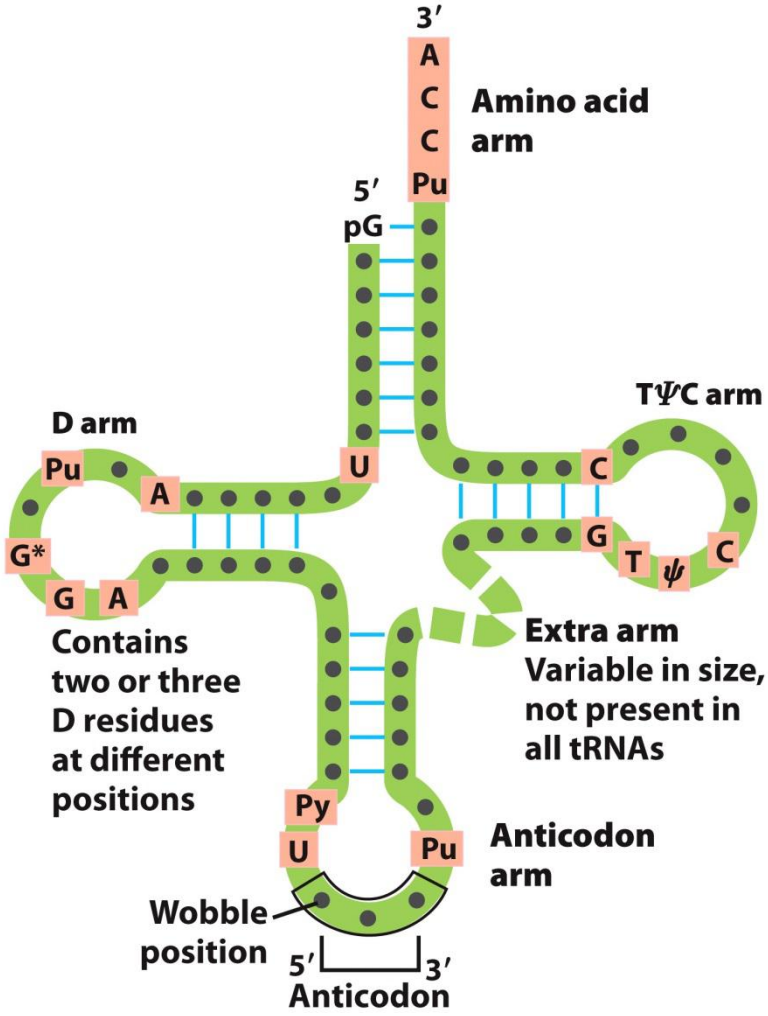
1. One codon recognized:		
Anticodon	(3′) X–Y–**C** (5′)	(3′) X–Y–**A** (5′)
Codon	(5′) X′–Y′–**G** (3′)	(5′) X′–Y′–**U** (3′)
2. Two codons recognized:		
Anticodon	(3′) X–Y–**U** (5′)	(3′) X–Y–**G** (5′)
Codon	(5′) X′–Y′–A/G (3′)	(5′) X′–Y′–C/U (3′)
3. Three codons recognized:		
Anticodon	(3′) X–Y–**I** (5′)	
Codon	(5′) X′–Y′–A/U/C (3′)	

The purpose of wobble:

Codon-anticodon interactions balance the requirements for accuracy and speed.

A minimum of 32 tRNAs are required to translate all 61 codons (31 to encode the amino acids and 1 for initiation)

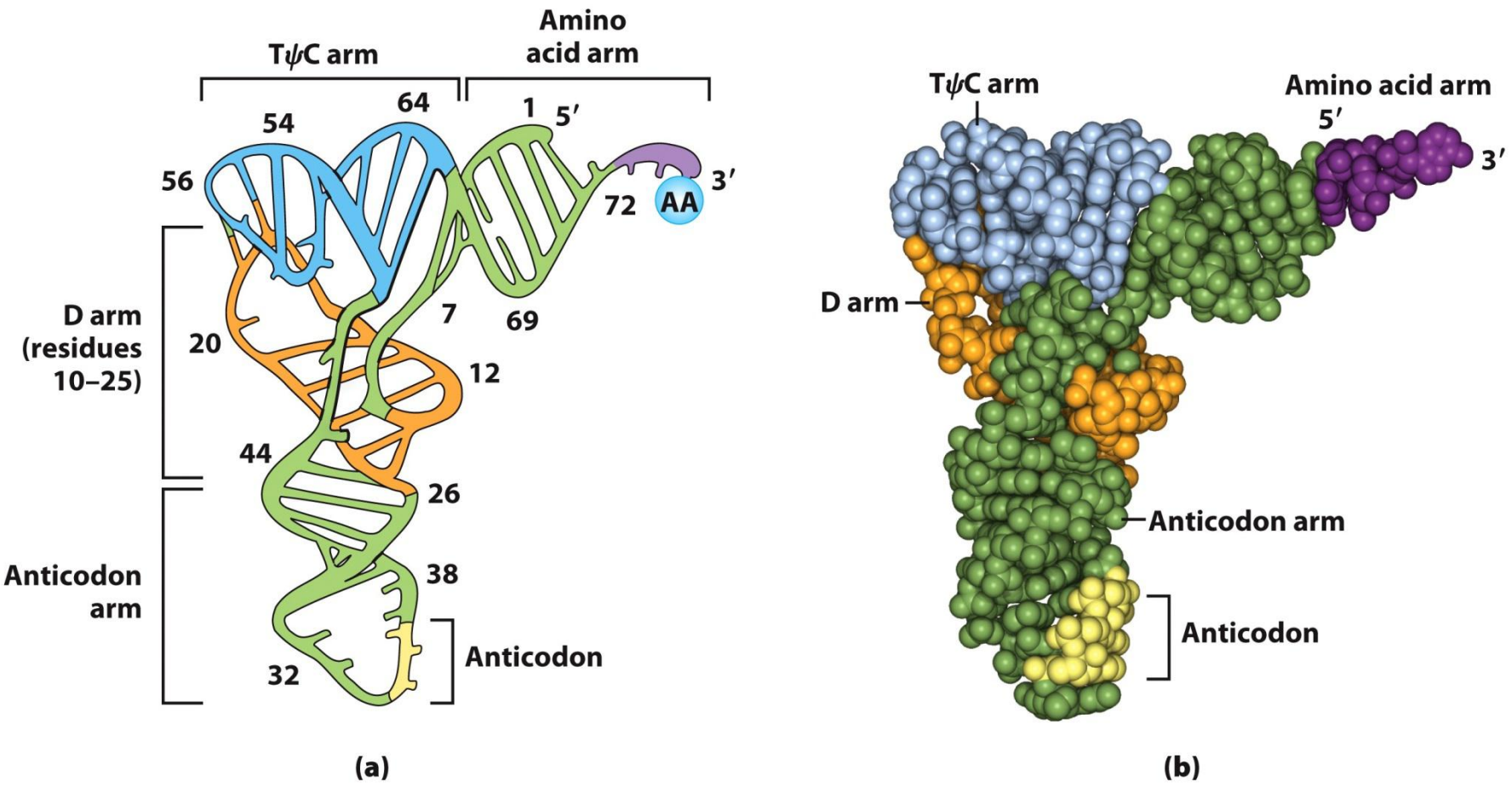
General cloverleaf secondary structure of tRNAs



Ψ, pseudouridine; T, ribothymidine; D, 5,6-dihydrouridine; Pu, purine nucleotide; Py, pyrimidine nucleotide; G*, guanylate *or* 2-*O*-methylguanylate

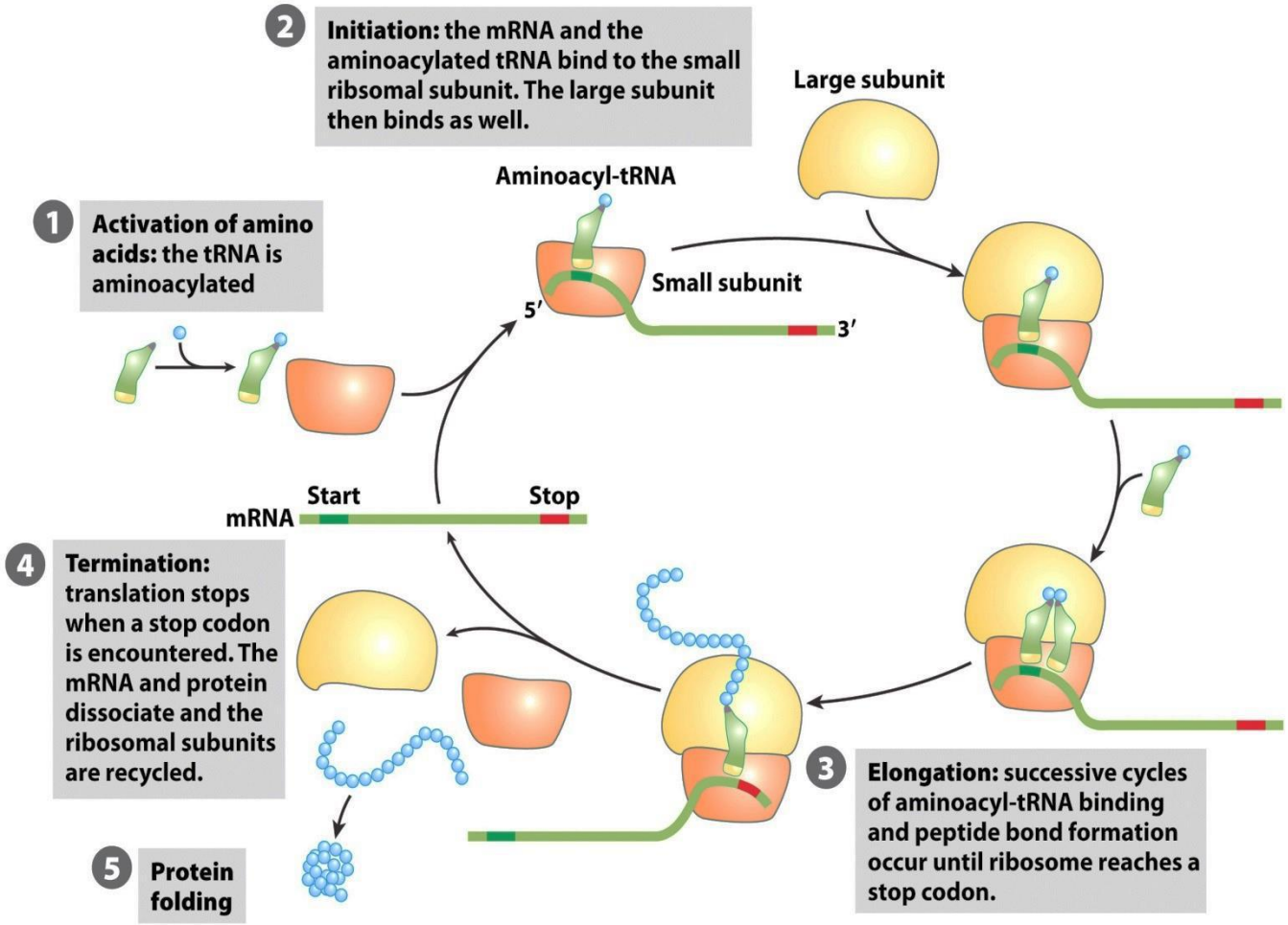
- 73 - 93 nucleotide residues;
- 5’ guanylate (pG) residue and 3’;
- The amino acid arm can carry a specific amino acid;
- The anticodon arm contains the anticodon;
- The the D arm contains the unusual nucleotide dihydrouridine (D);
- The TΨC arm contains ribothymidine (T) and pseudouridine (Ψ). The D and TΨC arms contribute important interactions for the overall folding of tRNA molecules, and
- The TΨC arm interacts with the large-subunit rRNA.

Tertiary structure of tRNA



Five stages of protein biosynthesis

1. Activation of amino acids
2. Initiation
3. Elongation
4. Termination
5. Folding & posttranslational modifications



Stage 1: Activation of amino acids

$$\text{Amino acid} + \text{tRNA} + \text{ATP} \xrightarrow{Mg^{2+}} \text{aminoacyl-tRNA} + \text{AMP} + 2P_i$$

$$\Delta G'^{\circ} \approx -29 \text{ kJ/mol}$$

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

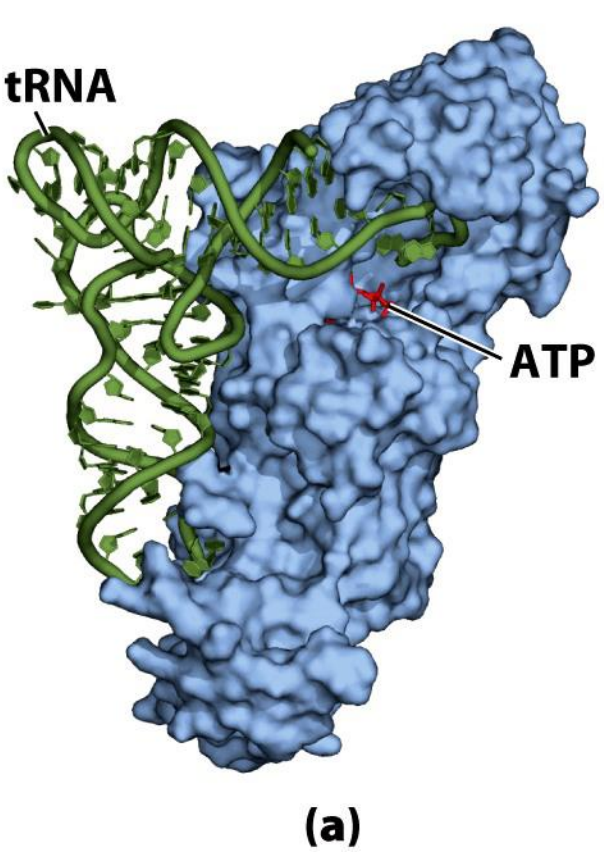
Class I		Class II	
Arg	Leu	Ala	Lys
Cys	Met	Asn	Phe
Gln	Trp	Asp	Pro
Glu	Tyr	Gly	Ser
Ile	Val	His	Thr

The reaction is catalyzed by an aminoacyl-tRNA synthetase. This reaction occurs in two steps in the enzyme's active site.

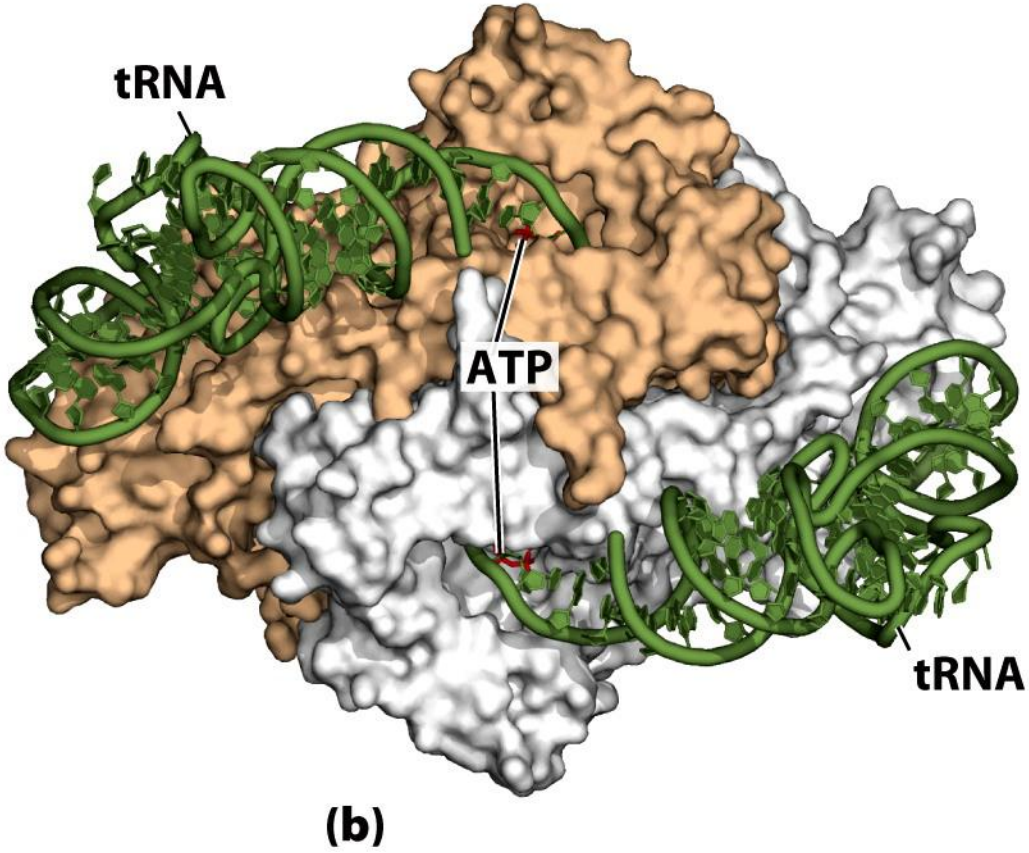
There are two classes aminoacyl-tRNA synthetase based on substantial differences in primary and tertiary structure and in reaction mechanism.

- Each enzyme is specific for one amino acid and one or more corresponding tRNAs.
- Most organisms have one aminoacyl-tRNA synthetase for each amino acid.
- For amino acids with two or more corresponding tRNAs, the same enzyme usually aminoacylates all of them.

Structure of aminoacyl-tRNA synthetases

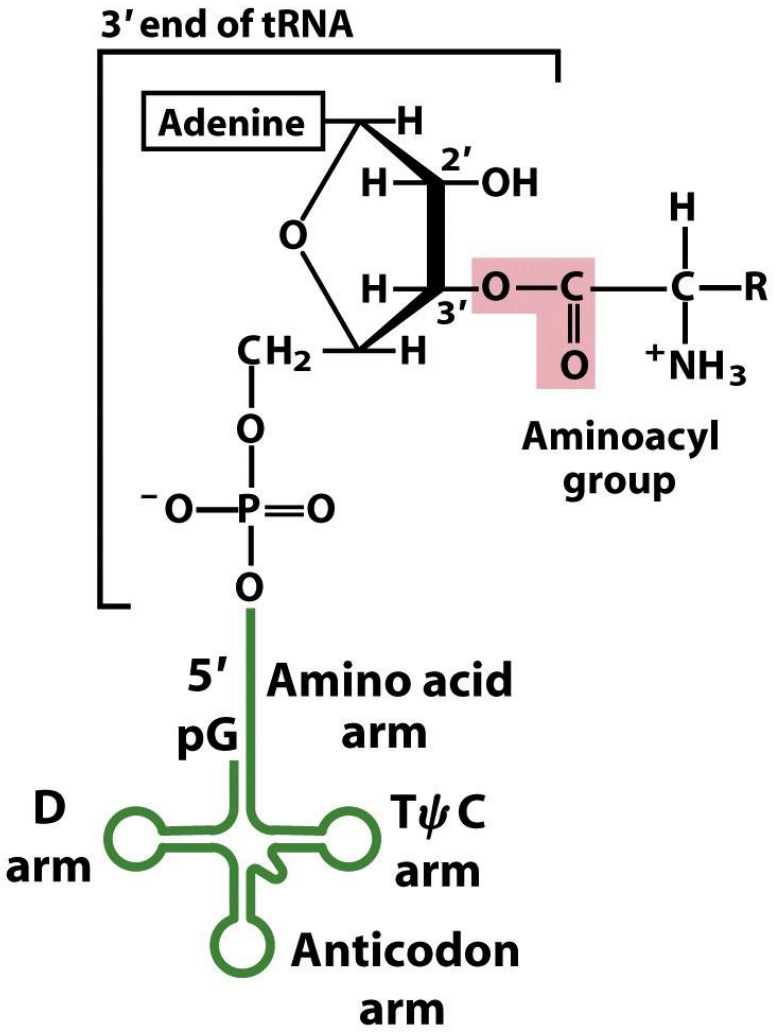


Type I



Type II

General structure of aminoacyl-tRNAs



The aminoacylation of tRNA accomplishes two ends:

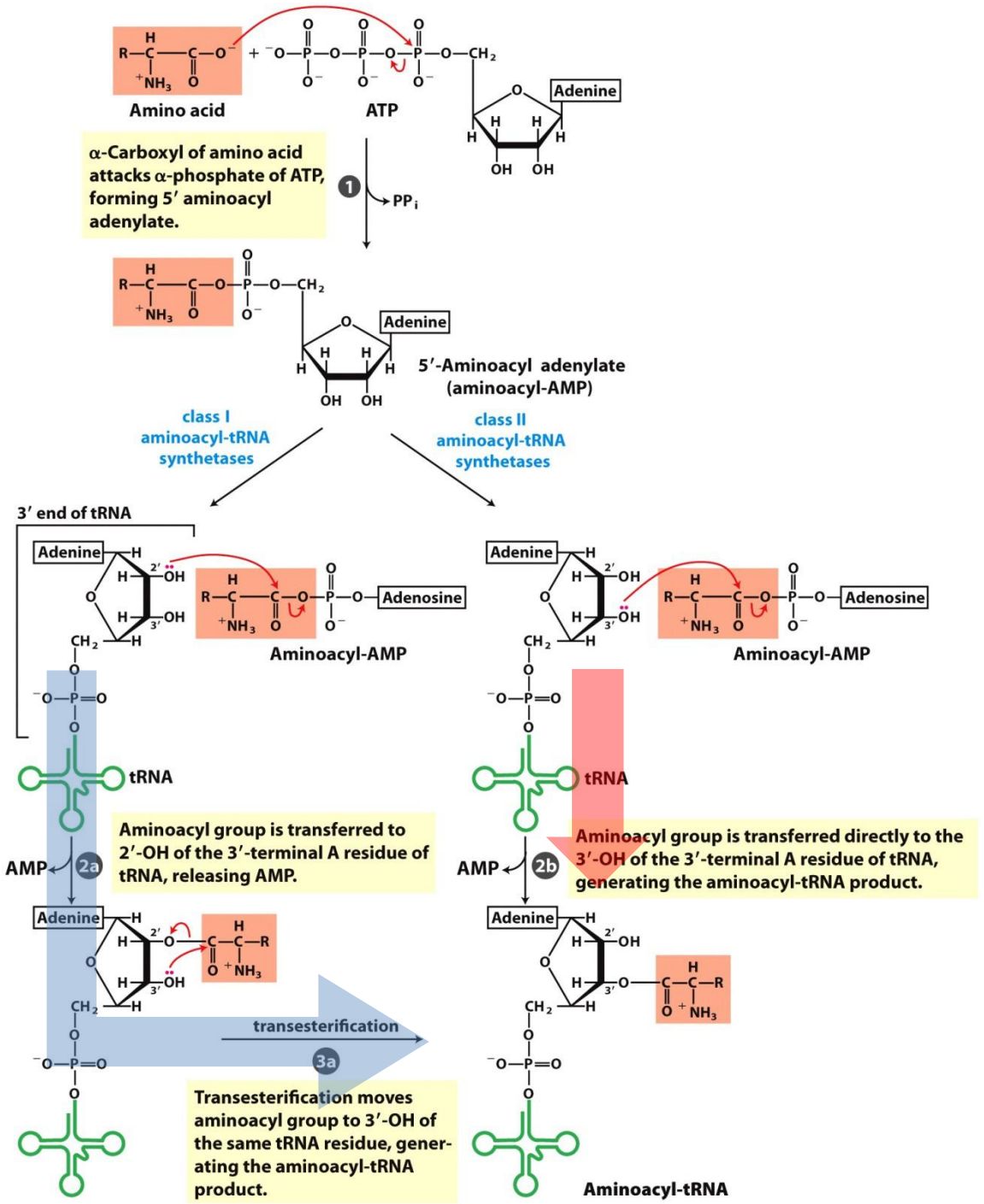
((1))activation of an amino acid for peptide bond formation

(2) attachment of the correct amino acid to an correct adaptor tRNA that ensures appropriate placement of the amino acid in a growing polypeptide.

Aminoacylation of tRNA by aminoacyl-tRNA synthetases

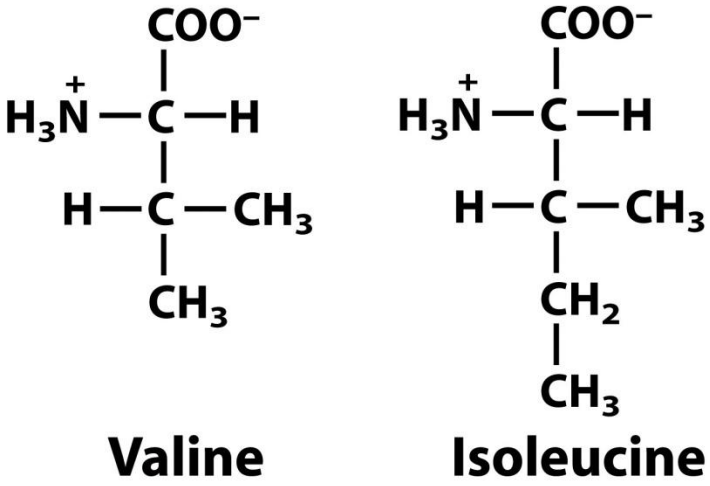
Step 1:
Forming aminoacyl-AMP

Step 2:
Transferring aminoacyl group to its corresponding specific tRNA



Proofreading by aminoacyl-tRNA synthetases

The identity of the amino acid attached to a tRNA is not checked on the ribosome, so attachment of the correct amino acid to the tRNA is essential to the fidelity of protein synthesis.

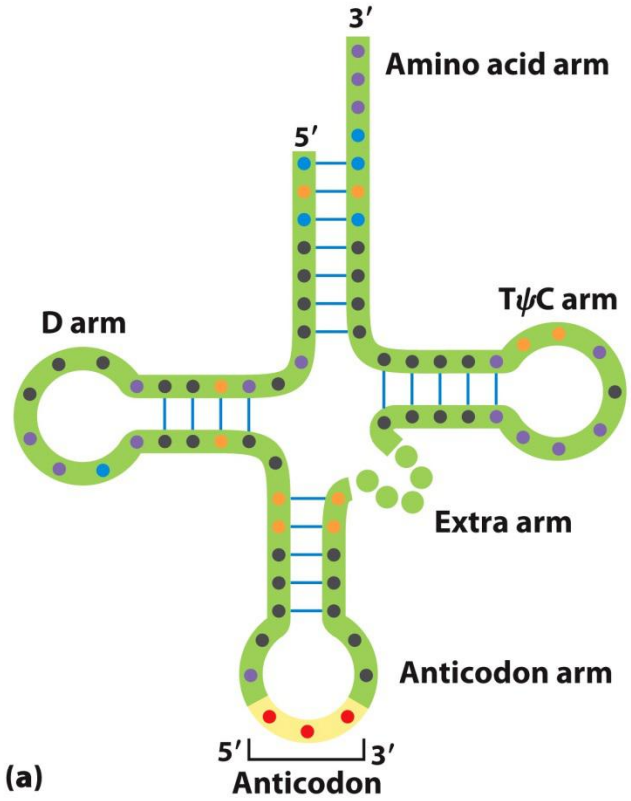


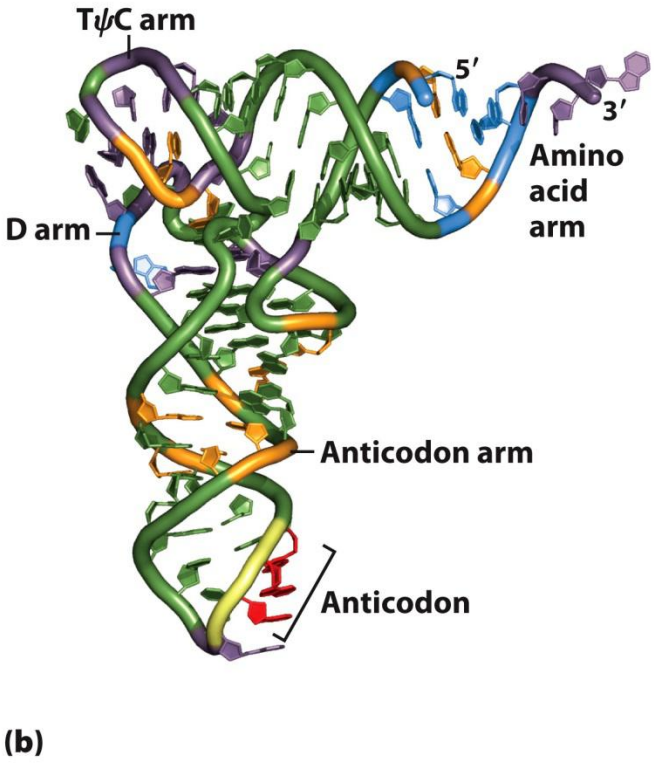
Incorrect aminoacyl-AMP products binding to a aminoacyl-tRNA synthetase will be hydrolyzed by a separate active site on the enzyme.

Specificity of tRNAs

Purple dots: same in all tRNAs

Other colors: recognition points

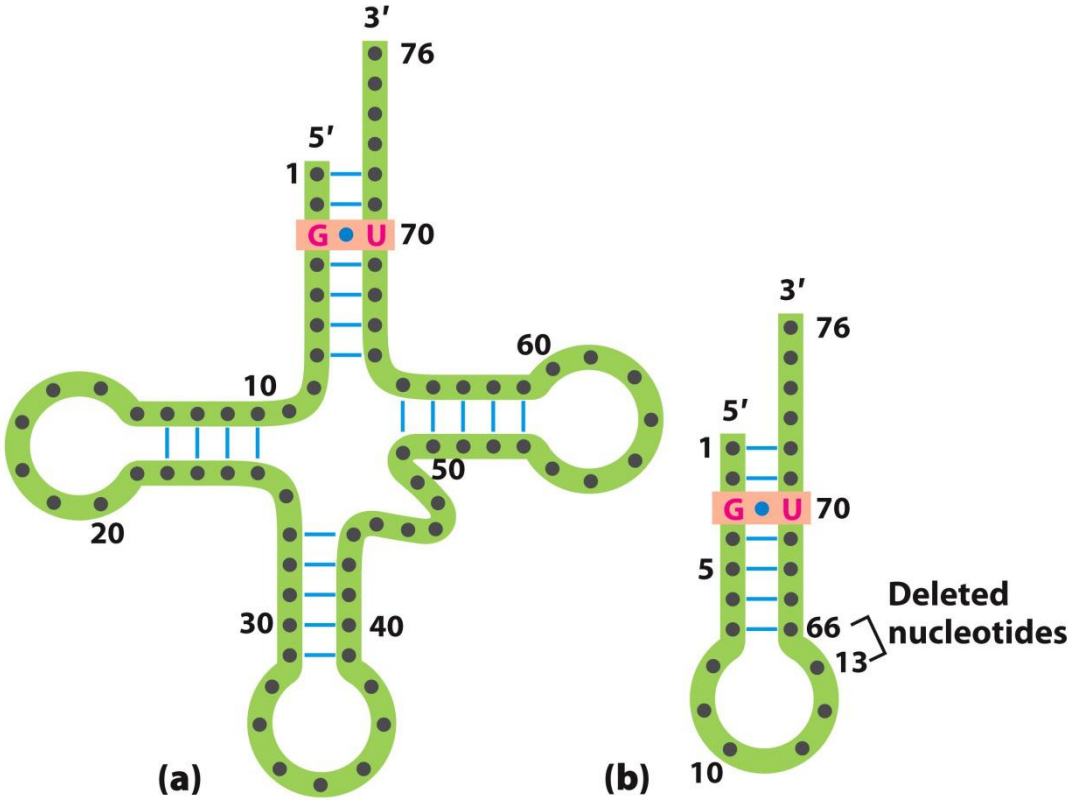




"Second genetic code": tRNA recognition by its specific aminoacyl-tRNA synthetase

Recognition mechanism could be simple

Specific nucleotides for tRNA recognition by its specific aminoacyl-tRNA synthetase.



tRNA^{Ala}

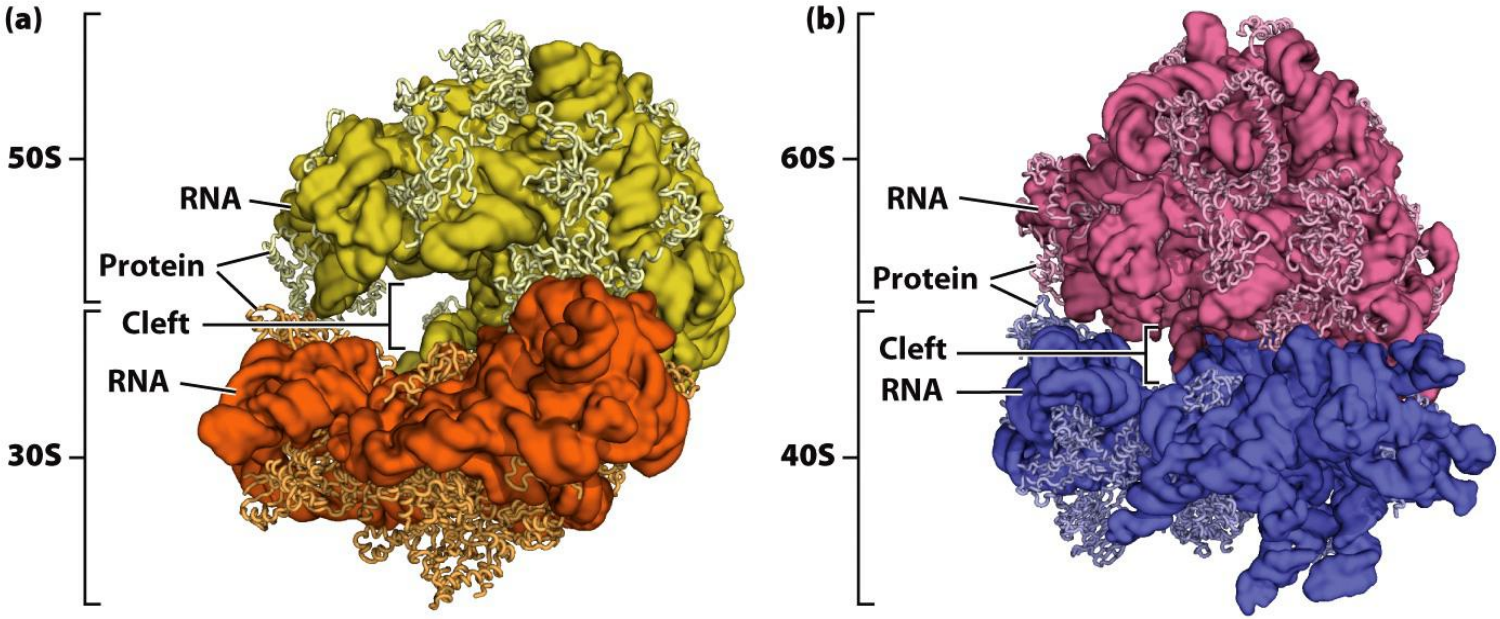
A short RNA with as few as 7 bp arranged in a simple hairpin minihelix is efficiently aminoacylated by the Ala-tRNA synthetase, as long as the RNA contains the critical G=U.

Five stages of protein biosynthesis

1. Activation of amino acids
2. Initiation
3. Elongation
4. Termination
5. Folding & posttranslational modifications

Machinery for translation: ribosome, a complex supramolecular machine

Ribosome is a ribozyme, in which rRNAs form the structural core



Bacterial ribosome

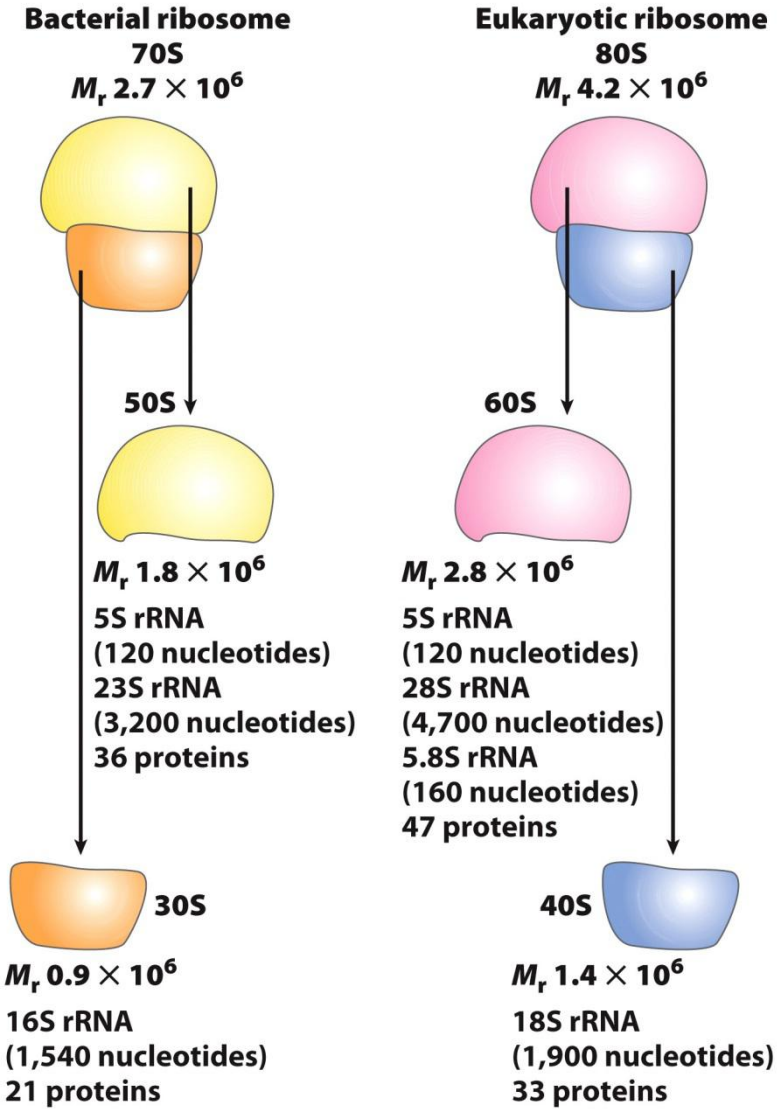
Yeast ribosome

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*	2 (5S and 23S rRNAs)

Bacterial Ribosome : 65% rRNA and 35% protein

Composition and mass of ribosomes in prokaryotes and eukaryotes

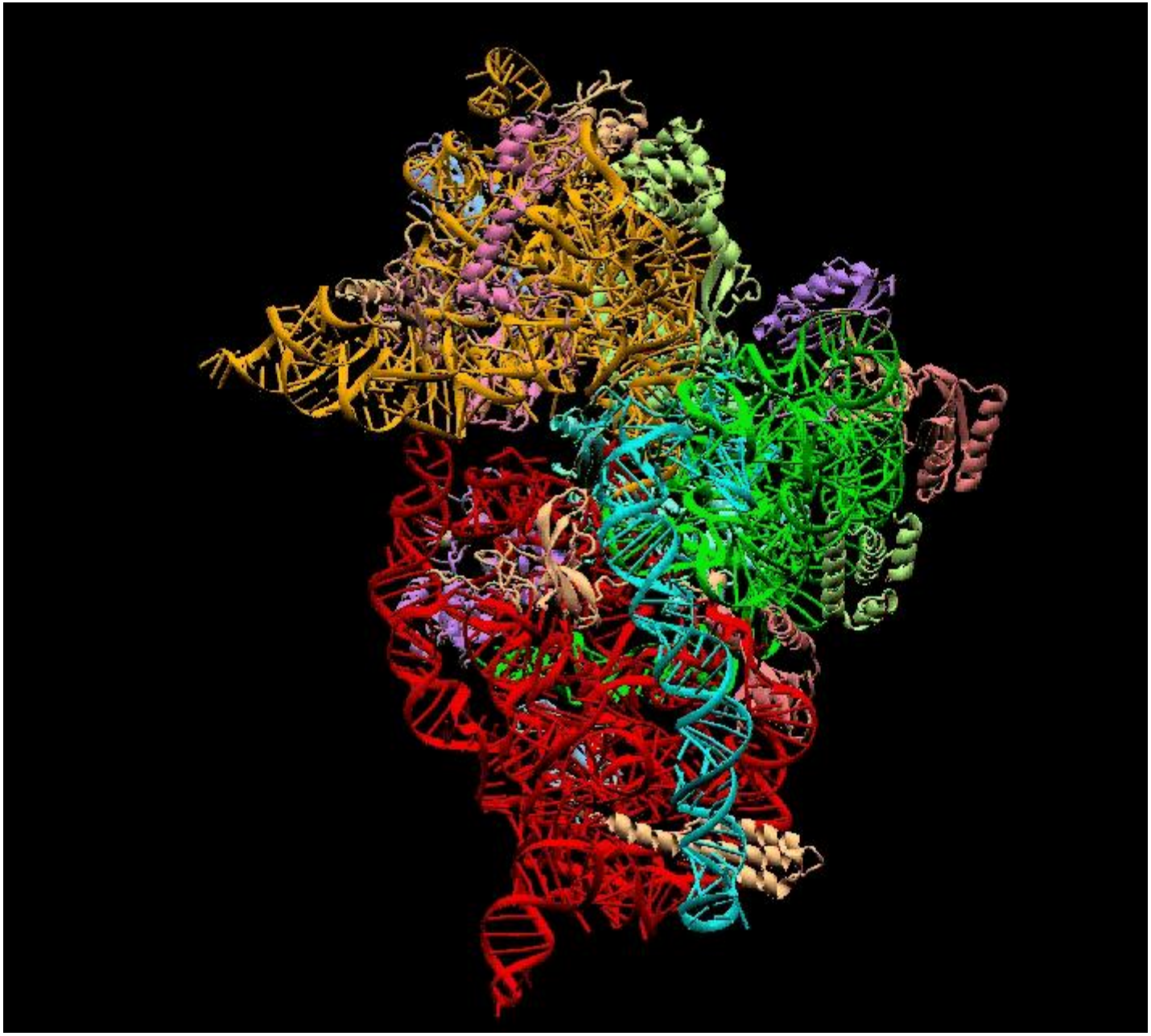


Prokaryotic rRNA

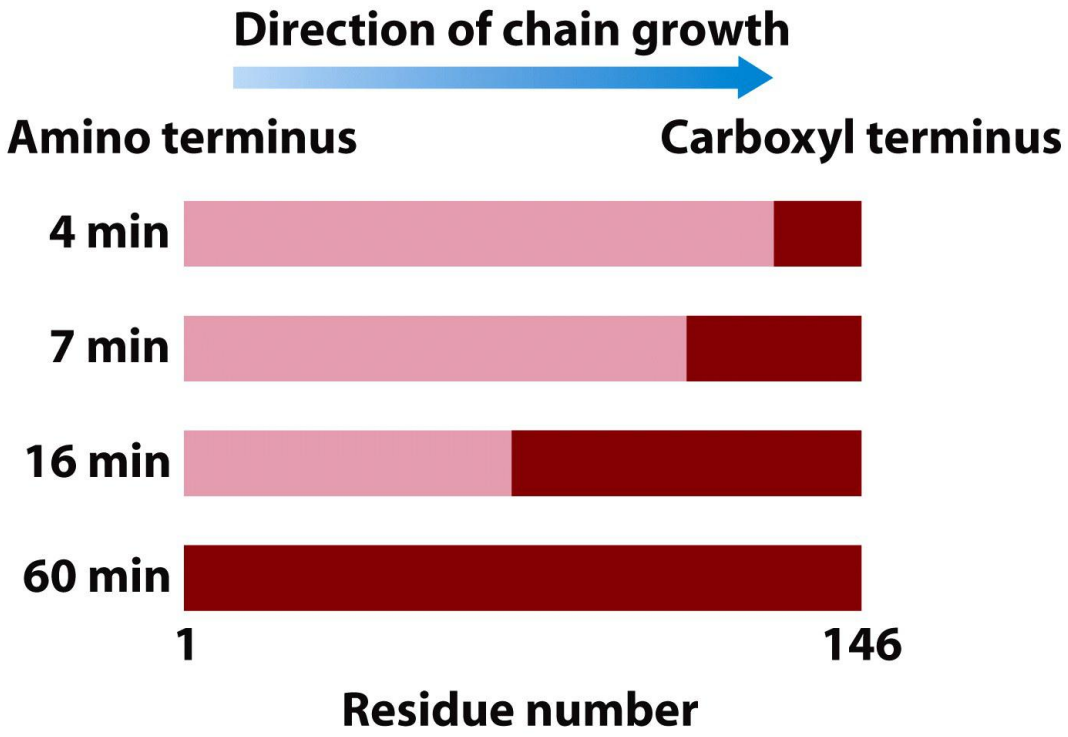
- 70S
  - 50S
    - 5S rRNA，
    - 23S rRNA
    - 36 proteins
  - 30S
    - 16S rRNA
    - 21 proteins

Eukaryotic rRNA

- 80S
  - 60S
    - 5S rRNA，5.8S rRNA，
    - 28S rRNA
    - 49 proteins
  - 40S
    - 18S rRNA
    - 33 proteins



Direction of protein synthesis



Protein synthesis begins at the amino-terminal end and proceeds by the stepwise addition of amino acids to the carboxyl-terminal end of the growing polypeptide.

Completed polypeptides were isolated at various times after incubated with radioactive amino acids.

AUG: the start codon

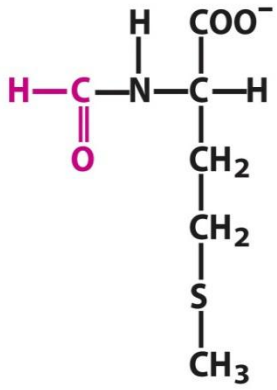
The **AUG** codon is for methionine (Met).

Two tRNAs for Met: initial Met & internal Met.

In bacteria:

Use *N*-formylmethionine (fMet) for initiation.

Two types of tRNA for Met: **tRNA^{Met}** and **tRNA^{fMet}**



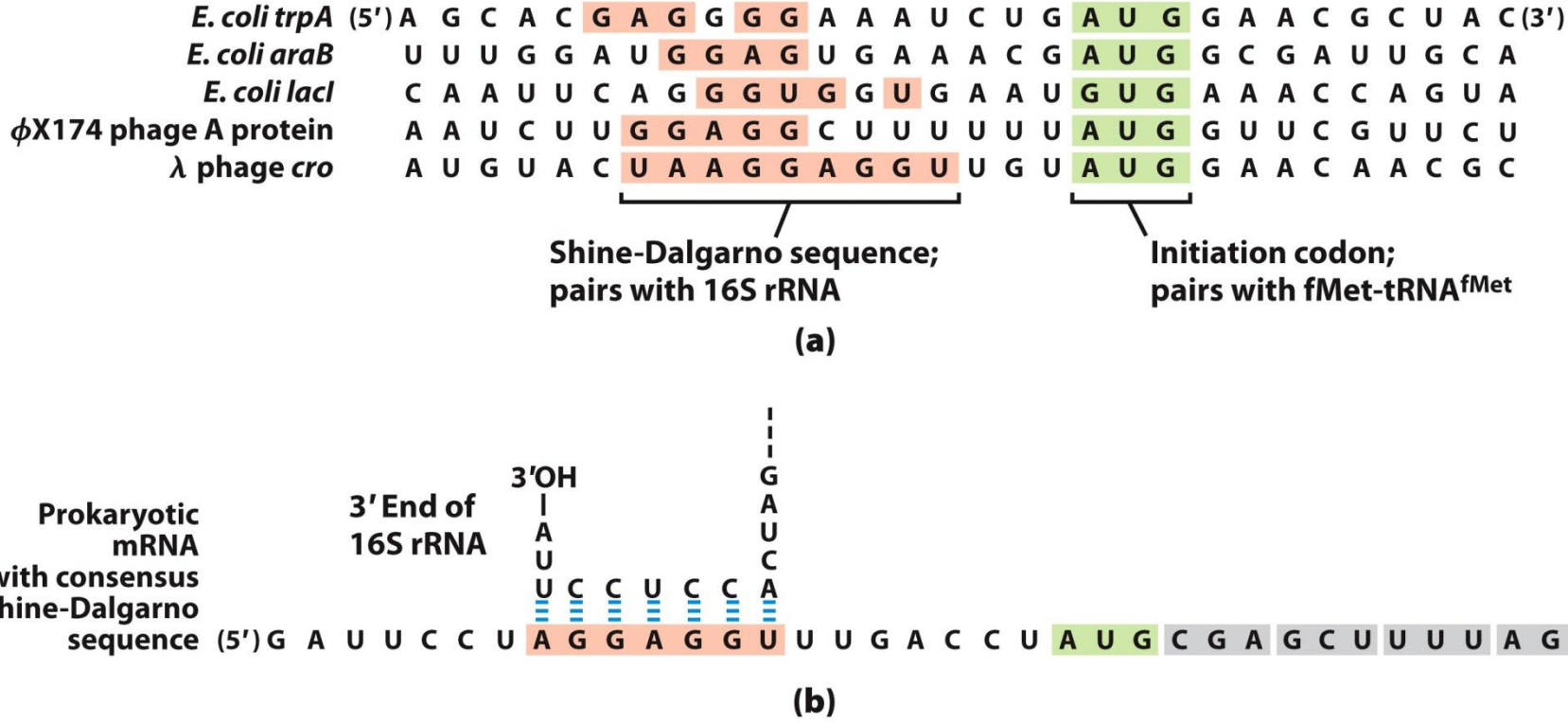
***N*-Formylmethionine**

In eukaryotic cells:

No fMet, but the cell uses a special tRNA for initial Met.

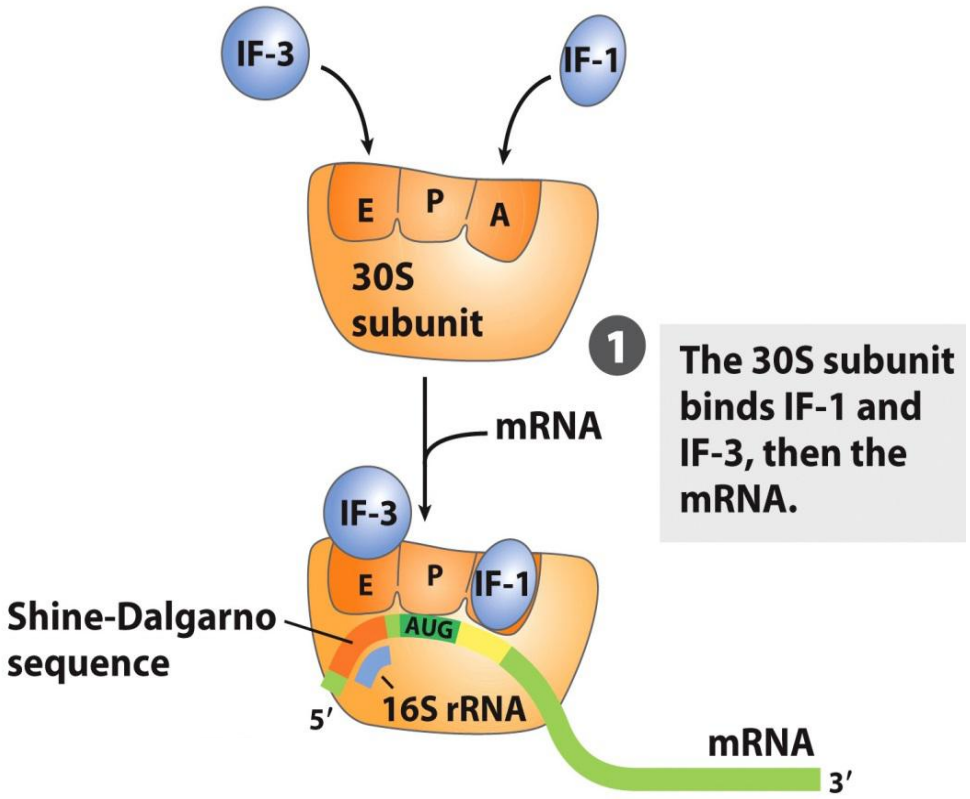
Where to start?

Determine reading frame: Shine-Dalgarno sequences



Messenger RNA sequence itself contains signal for initiation of protein synthesis in bacteria.

Formation of the initiation complex



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

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

Step 1:

30S subunit binds to mRNA

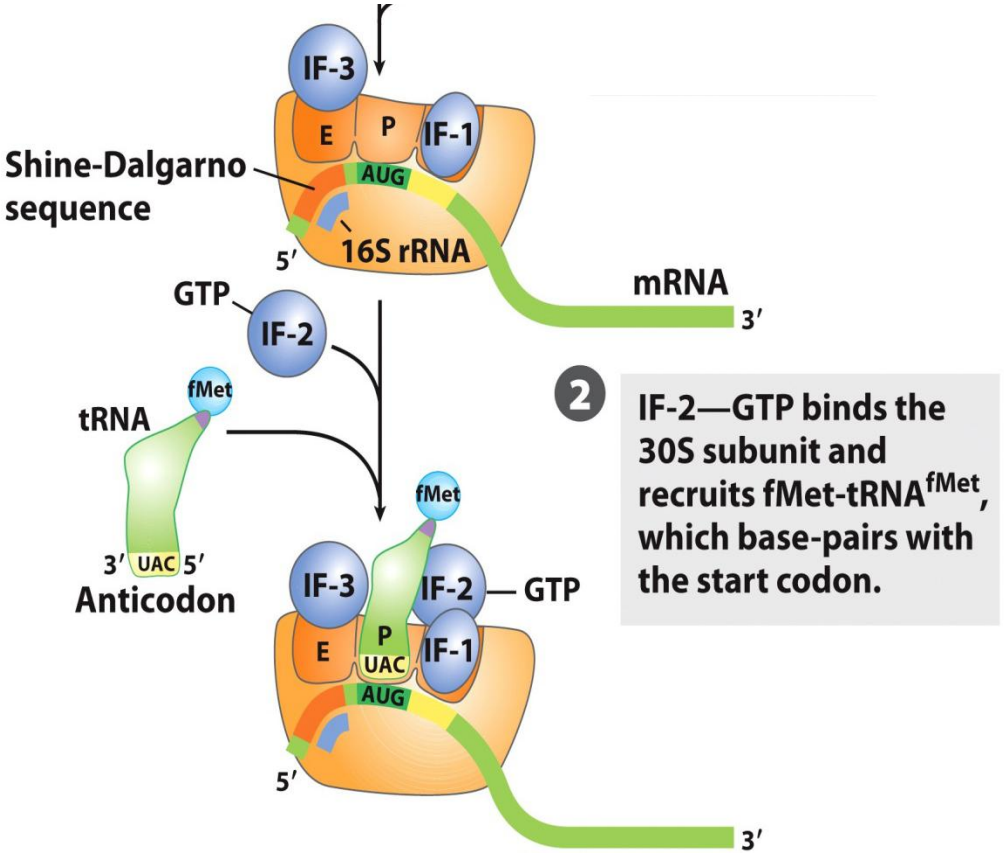
1) initiation factors IF-1 and IF-3;

2) then binds to mRNA.

A site (aminoacyl site):
IF-1 now binds here
prevent tRNA-aa binding during initiation.

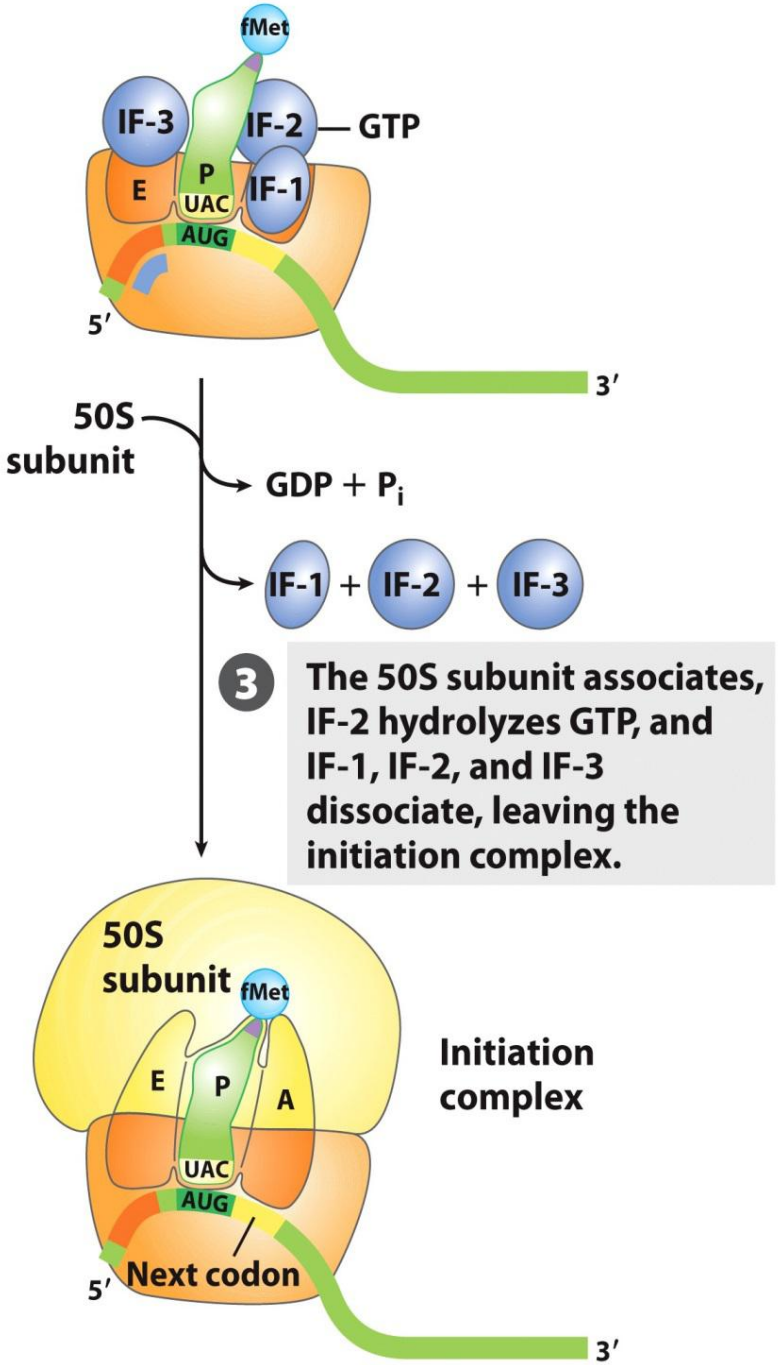
P site (peptidyl site): at AUG codon
A special Met: f Met-tRNA^{fMet}
Structure similar to a peptide

E (exit site): for "uncharged" tRNAs leave during elongation, on 50S subunit.



Step 2:

GTP-bound IF-2 delivers the initiating fMet-tRNA^{fMet}.



Step 3:

GTP on IF-2 is hydrolyzed to GDP and Pi.

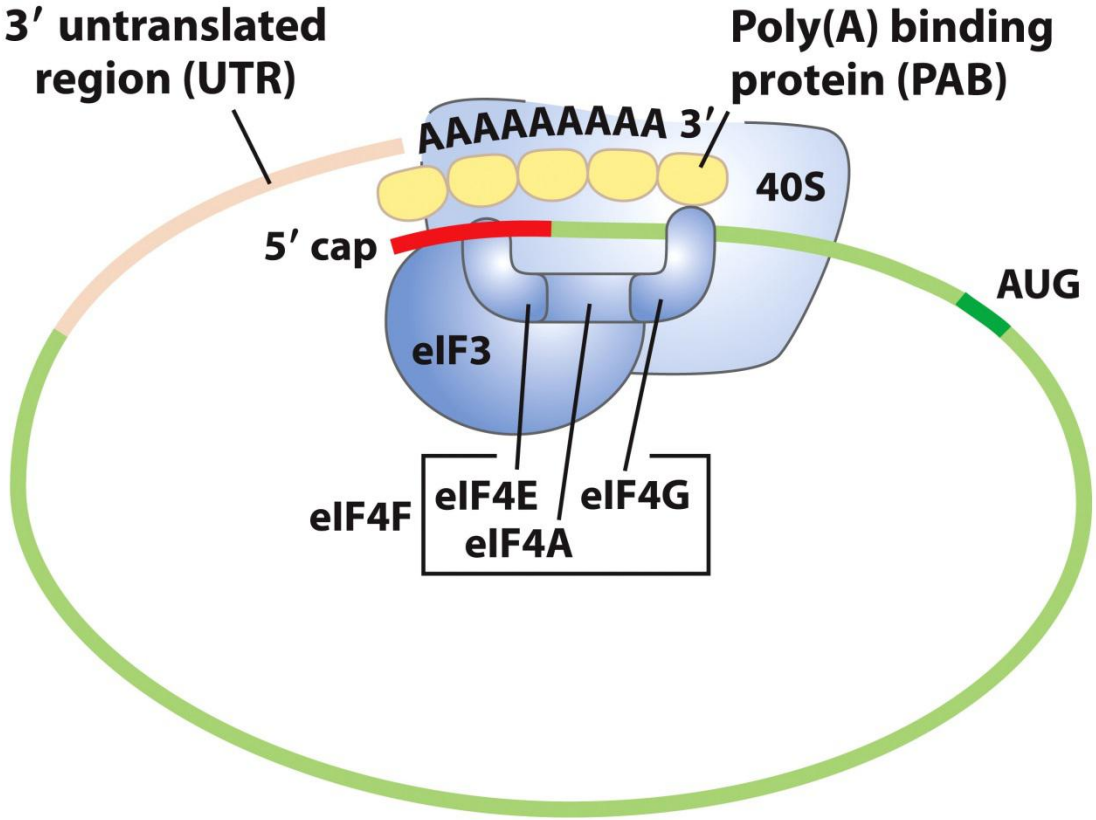
IF-1, 2, 3 all release from 30S subunit.

50S subunit associates with 30S subunit.

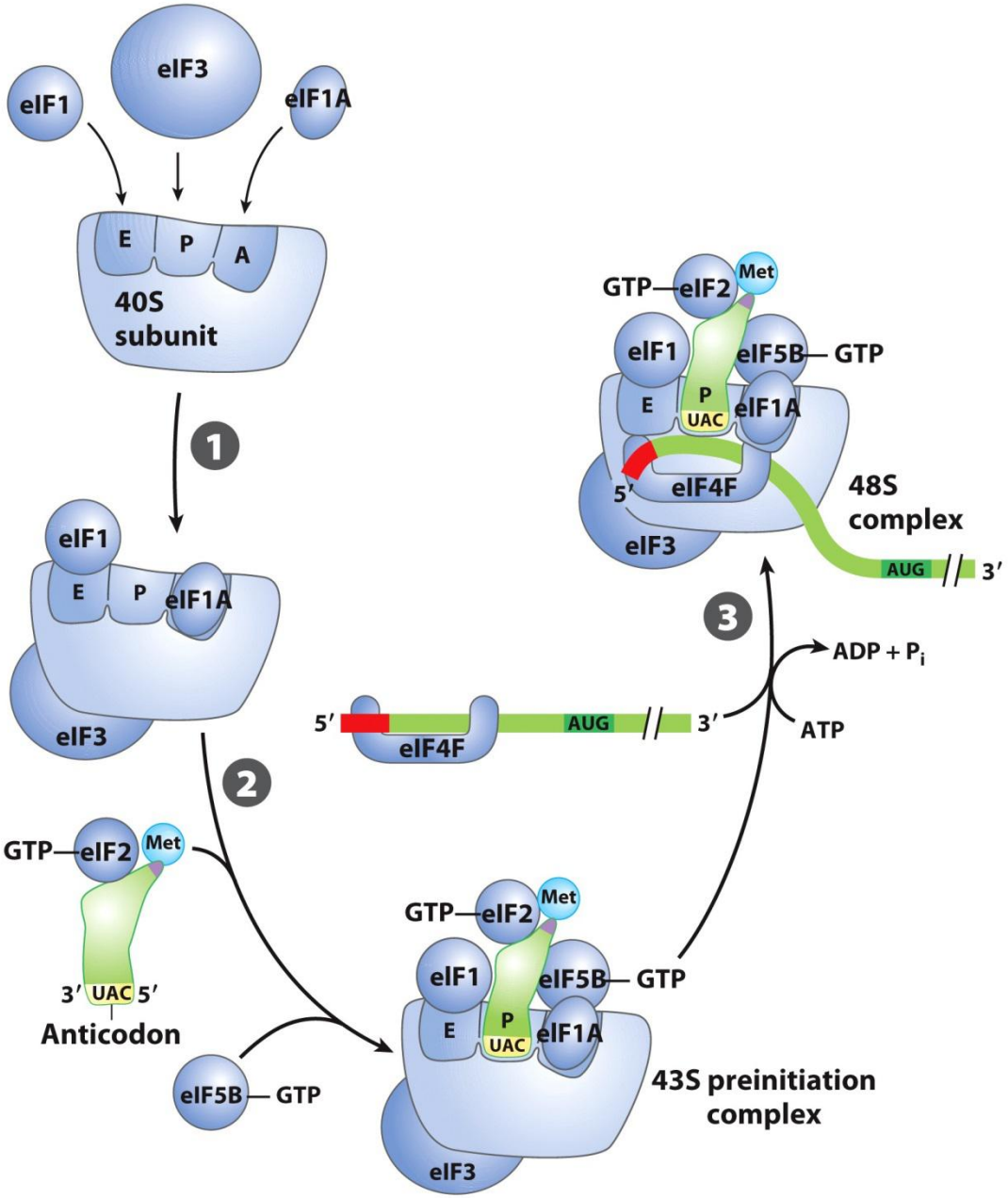
Initiation complex { 70S ribosome
mRNA
fMet-tRNA^{fMet}.

Initiation in eukaryotic cells

- No Shine-Dalgarno-like sequence
- Scan mRNA from the 5’ end until the first AUG



eIF4F complex contains eIF4E (binding to the 5’ cap), eIF4A (an ATPase and RNA helicase), and eIF4G (binding to eIF3 and the poly(A) binding protein)



Step 1:

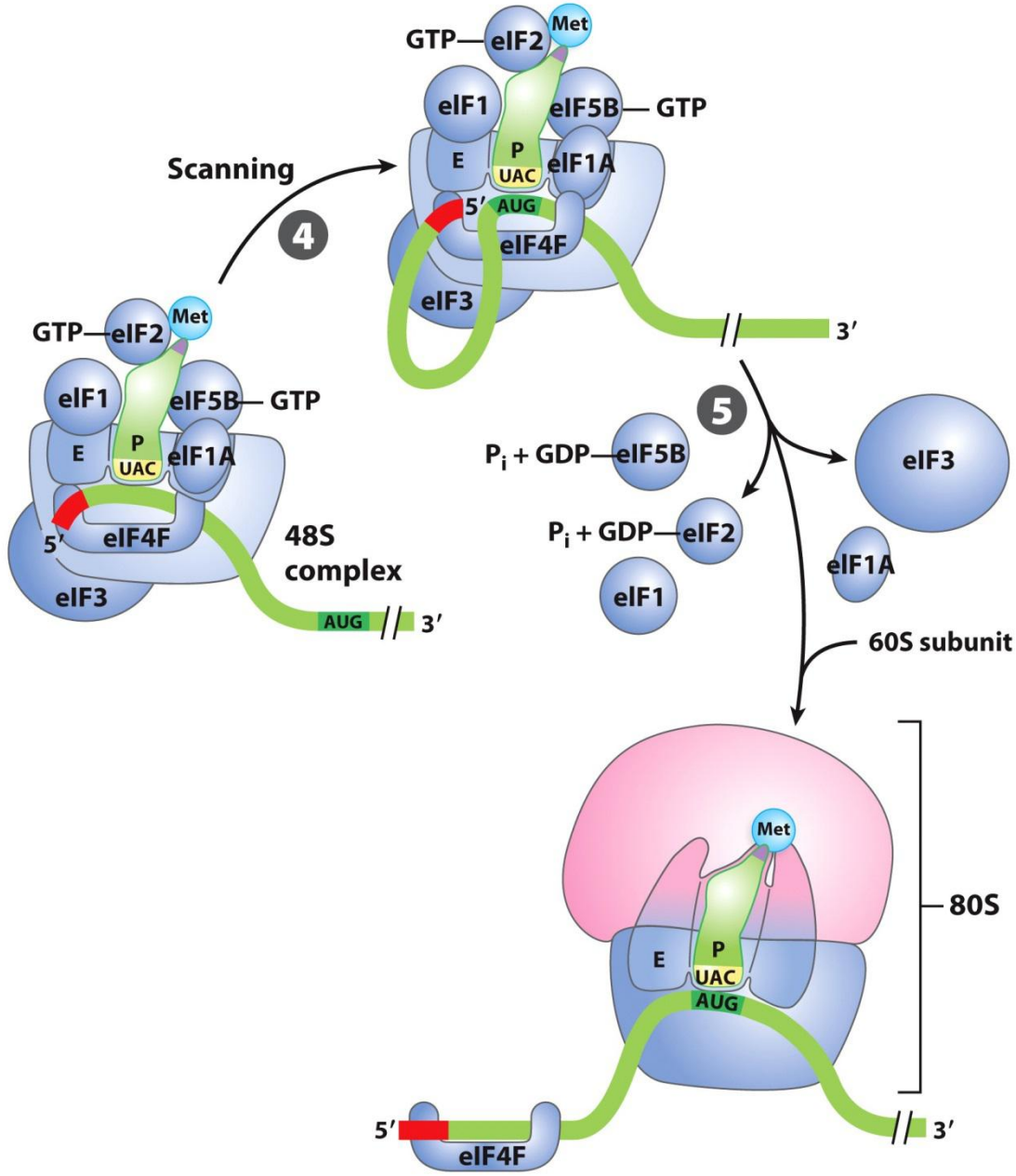
Initiation factors eIF1, eIF1A, and eIF3 bind to 40S subunit.

Step 2:

GTP-bound eIF2 delivers the charged initiator tRNA to the complex to form the 43S preinitiation complex together with eIF5 and eIF5B.

Step 3:

The 43S preinitiation complex, eIF4F complex, and mRNA together create a 48S complex.



Step 4:

Scan for the first AUG.

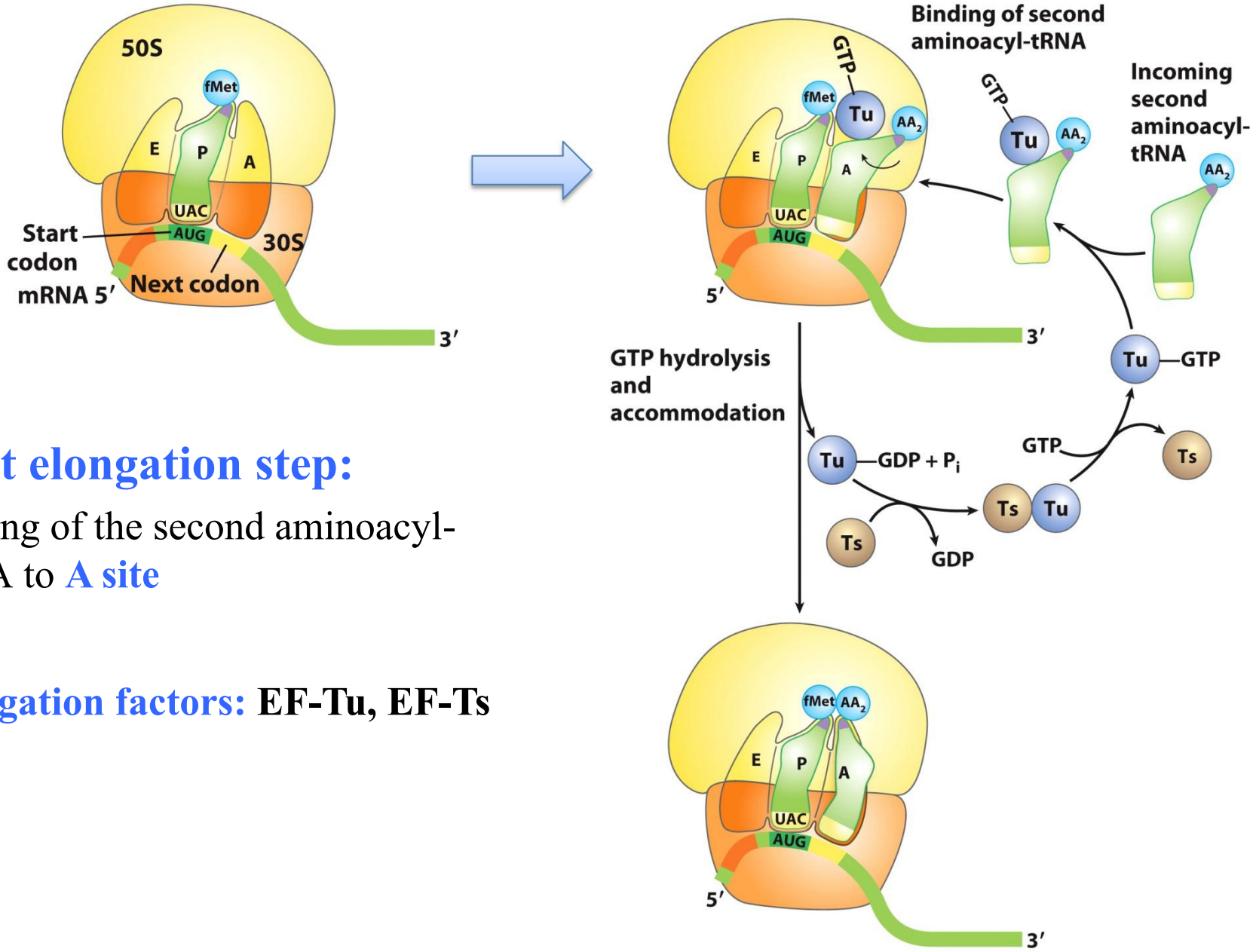
Step 5:

Release the initiation factors, associate with 60S subunit to form the 80S initiation complex.

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	
eIF1	Binds to the E site of the 40S subunit; facilitates interaction between eIF2-tRNA-GTP ternary complex and the 40S subunit
eIF1A	Homolog of bacterial IF-1; prevents premature binding of tRNAs to A site
eIF2	GTPase; facilitates binding of initiating Met-tRNA^{Met} to 40S ribosomal subunit
eIF2B*, eIF3	First factors to bind 40S subunit; facilitate subsequent steps
eIF4F	Complex consisting of eIF4E, eIF4A, and eIF4G
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 (PABP); 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
eIF5b	GTPase homologous to bacterial IF-2; promotes dissociation of initiation factors prior to final ribosome assembly

Stage 3: Elongation

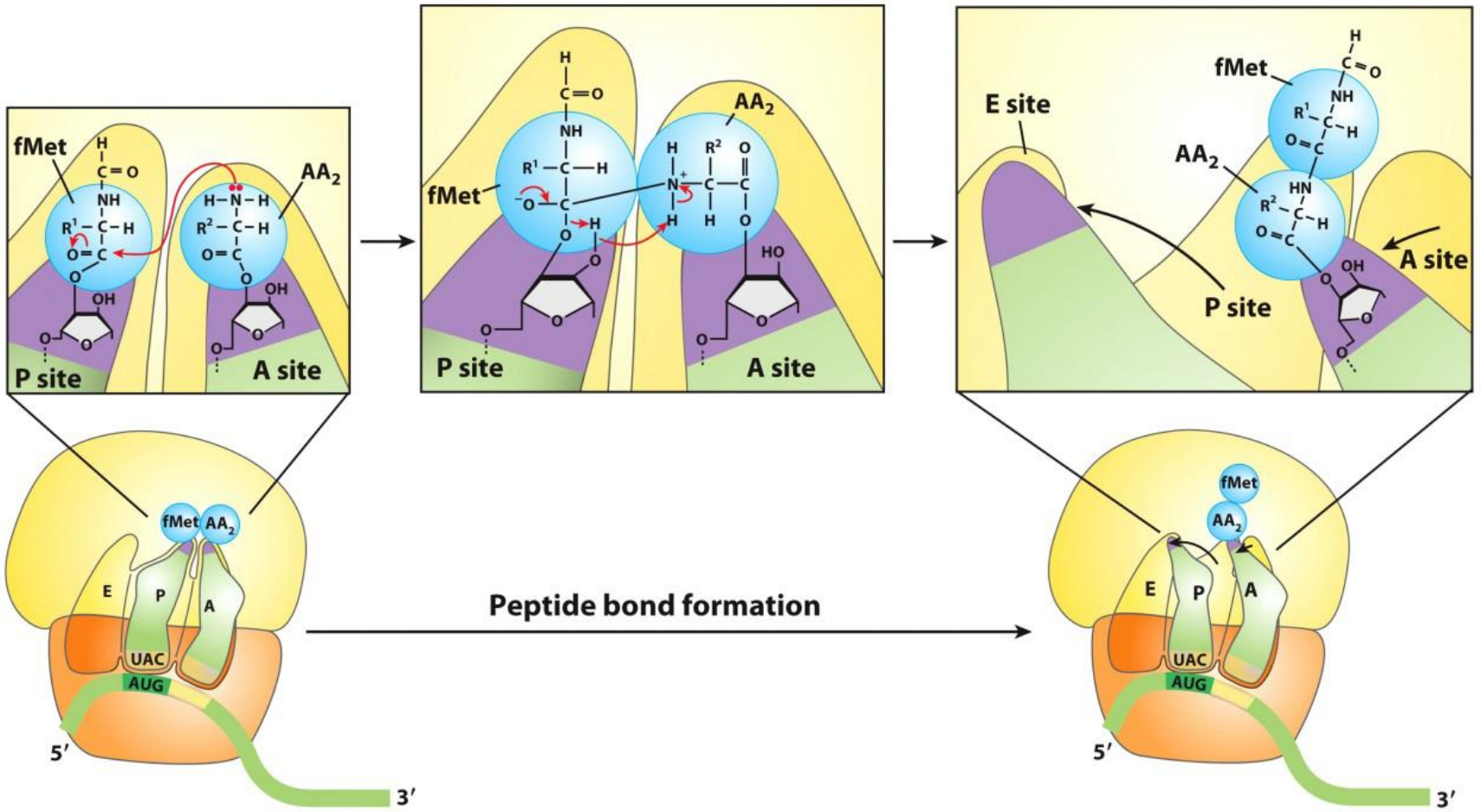


First elongation step:

binding of the second aminoacyl-tRNA to **A site**

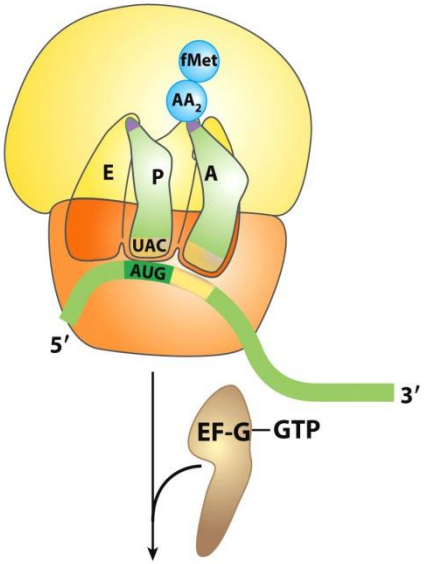
Elongation factors: **EF-Tu, EF-Ts**

Second elongation step: formation of the first peptide bond

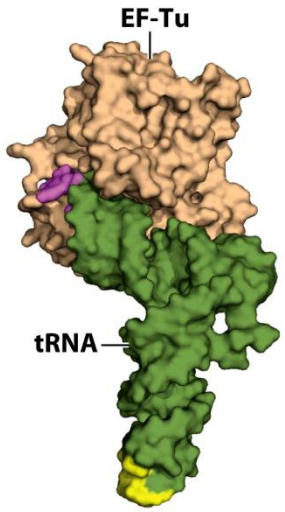


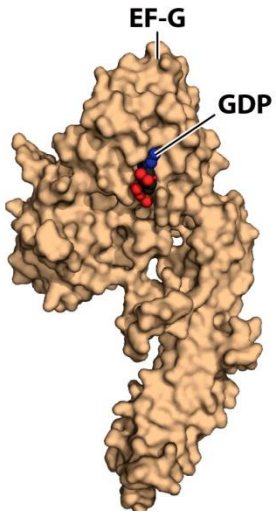
The α-amino group of the amino acid in the A site acts as a nucleophile, displacing the tRNA in the P site to form the peptide bond. It therefore produces a dipeptidyl-tRNA in the A site, and the now "uncharged" $tRNA^{fMet}$ remains bound to the P site. This reaction is catalyzed by the 23S rRNA.

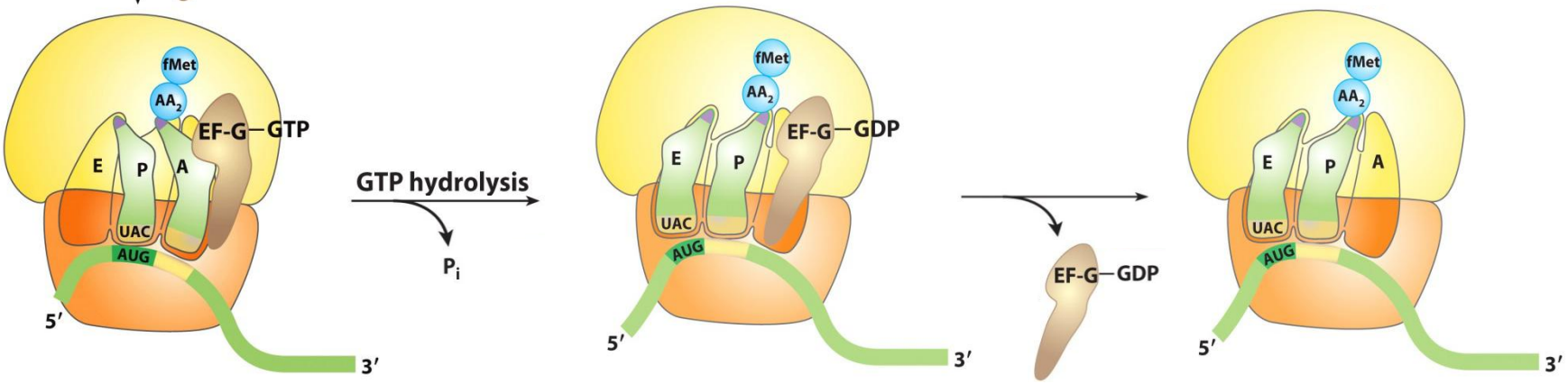
Third elongation step: translocation



EF-G mimics the structure of the EF-Tu–tRNA complex, EF-G can bind the A site and presumably displace the peptidyl-tRNA.

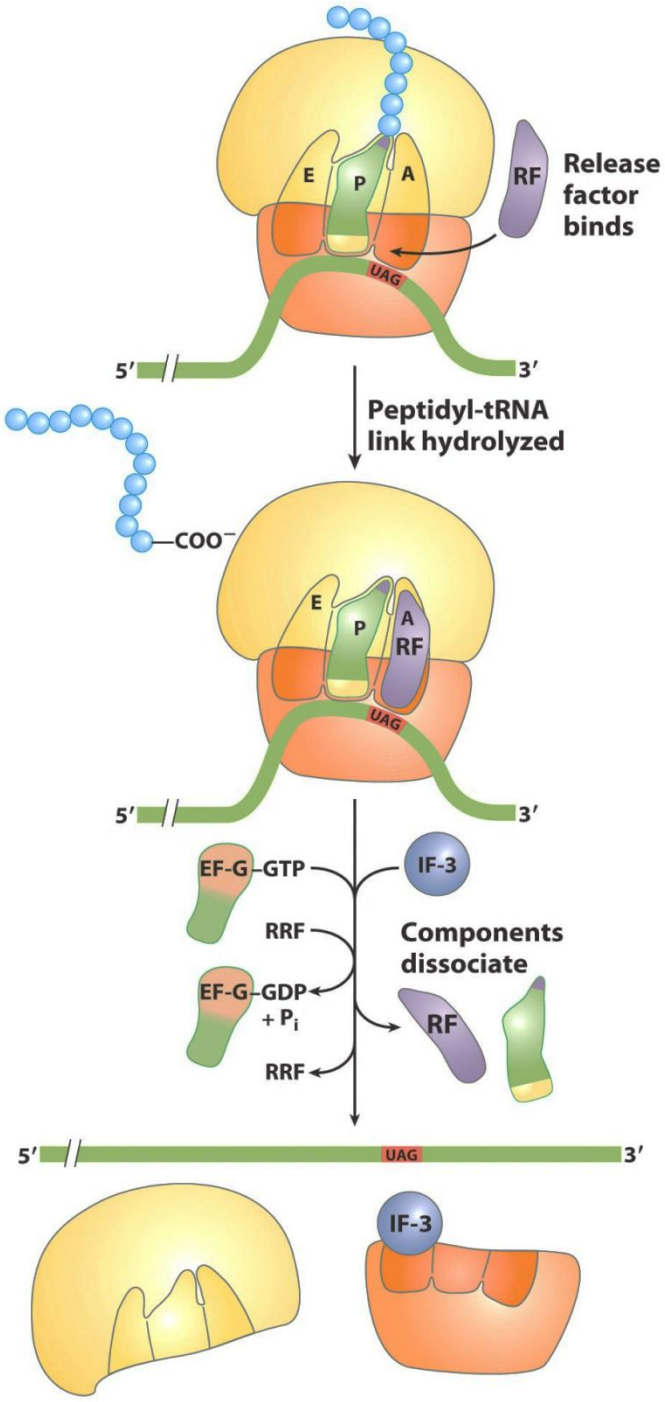






The ribosome moves one codon toward the 3’ end of the mRNA, which requires **EF-G (translocase)** and the energy provided by hydrolysis of a molecule of GTP. A change in the three-dimensional conformation of the entire ribosome results in its movement along the mRNA.

Stage 4: Termination



Termination codons: UAA, UAG, UGA

Termination factors:

- In bacteria:

  RF-1: recognizes UAA, UAG

  RF-2: recognizes UAA, UGA

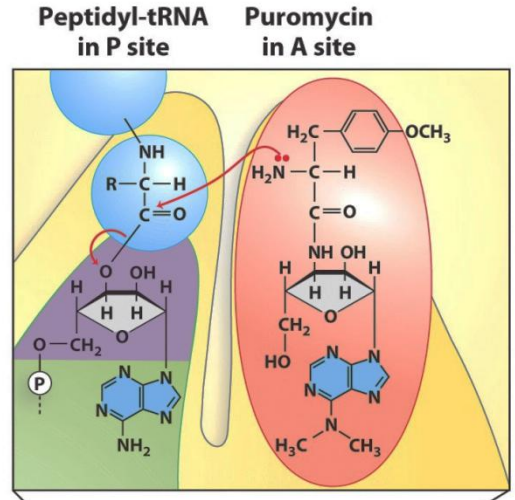
  RF-3: releases the ribosomal subunit

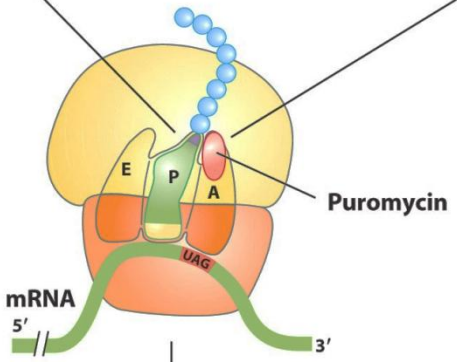
- In eukaryotes:

  eRF, a single release factor, recognizes all three termination codons.

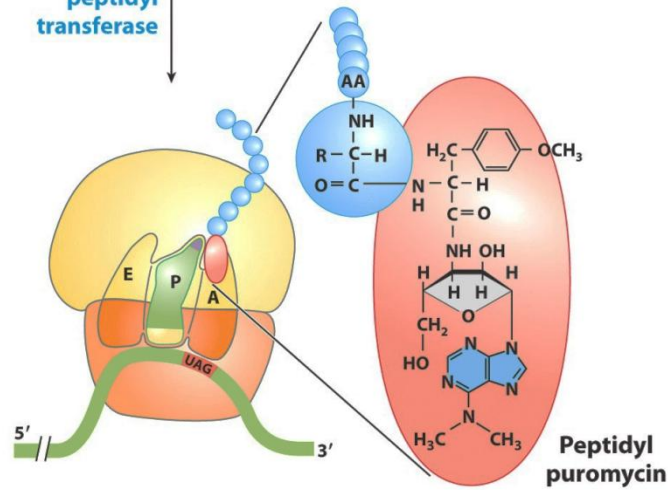
1. Hydrolysis of the terminal peptidyl-tRNA bond;
2. Release of the free polypeptide and the last tRNA, now uncharged, from the P site;
3. Dissociation of the 70S ribosome into its 30S and 50S subunits, ready to start a new cycle of polypeptide synthesis.

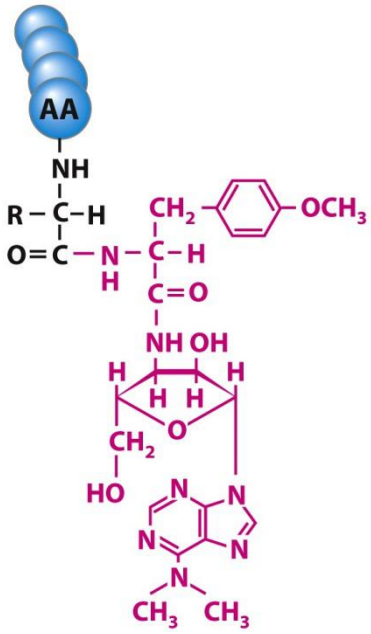
Inhibition of protein synthesis by many antibiotics and Toxins



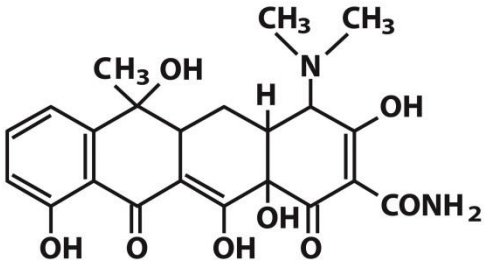


peptidyl transferase

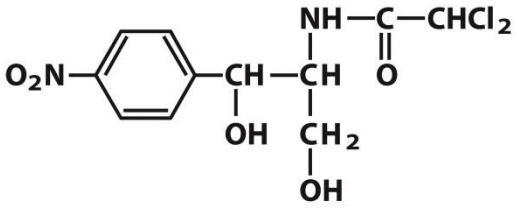




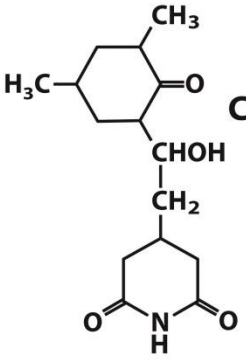
Puromycin



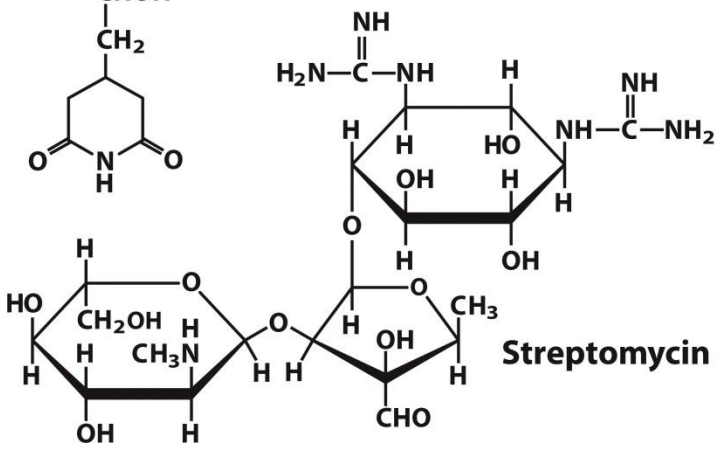
Tetracycline



Chloramphenicol



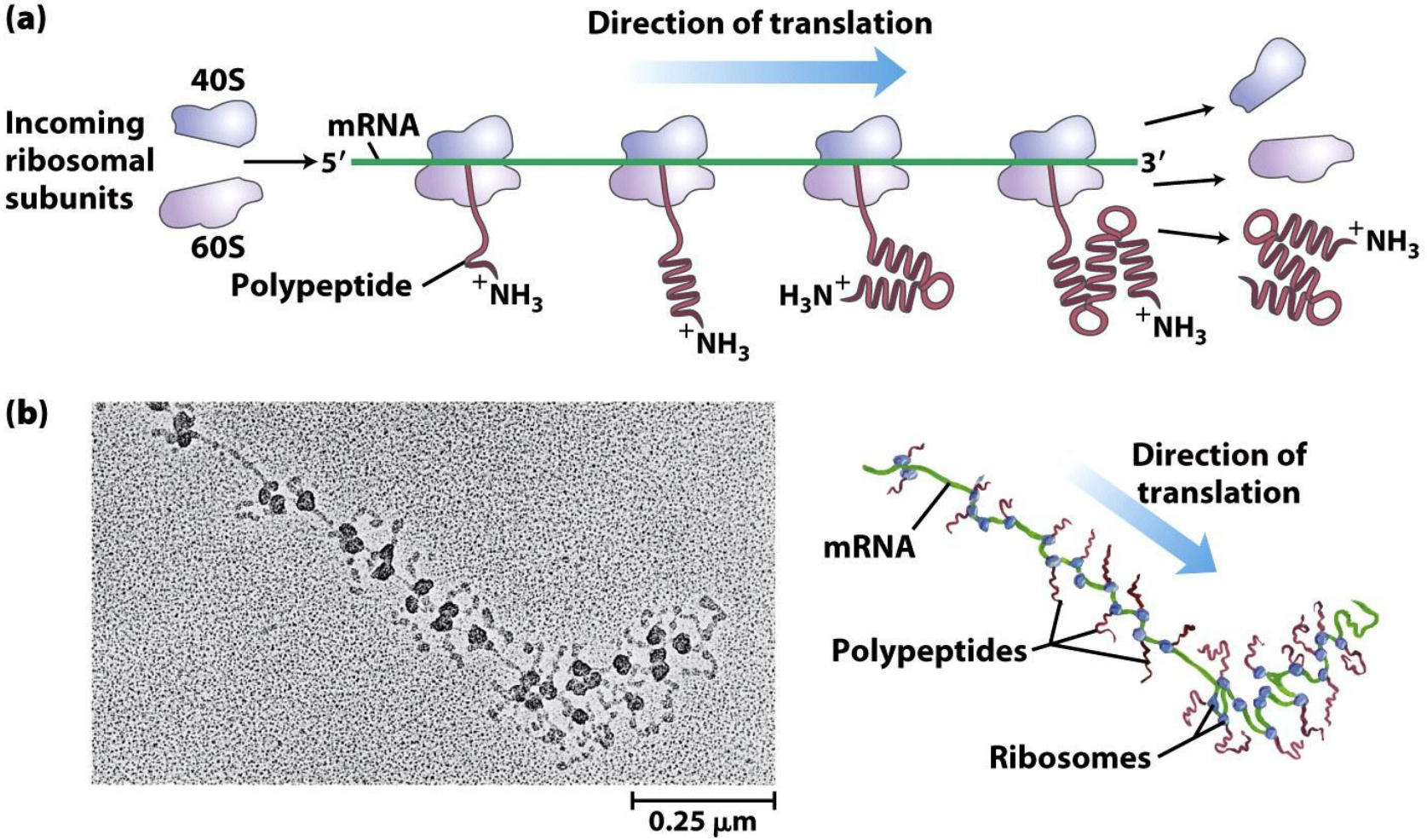
Cycloheximide



Streptomycin

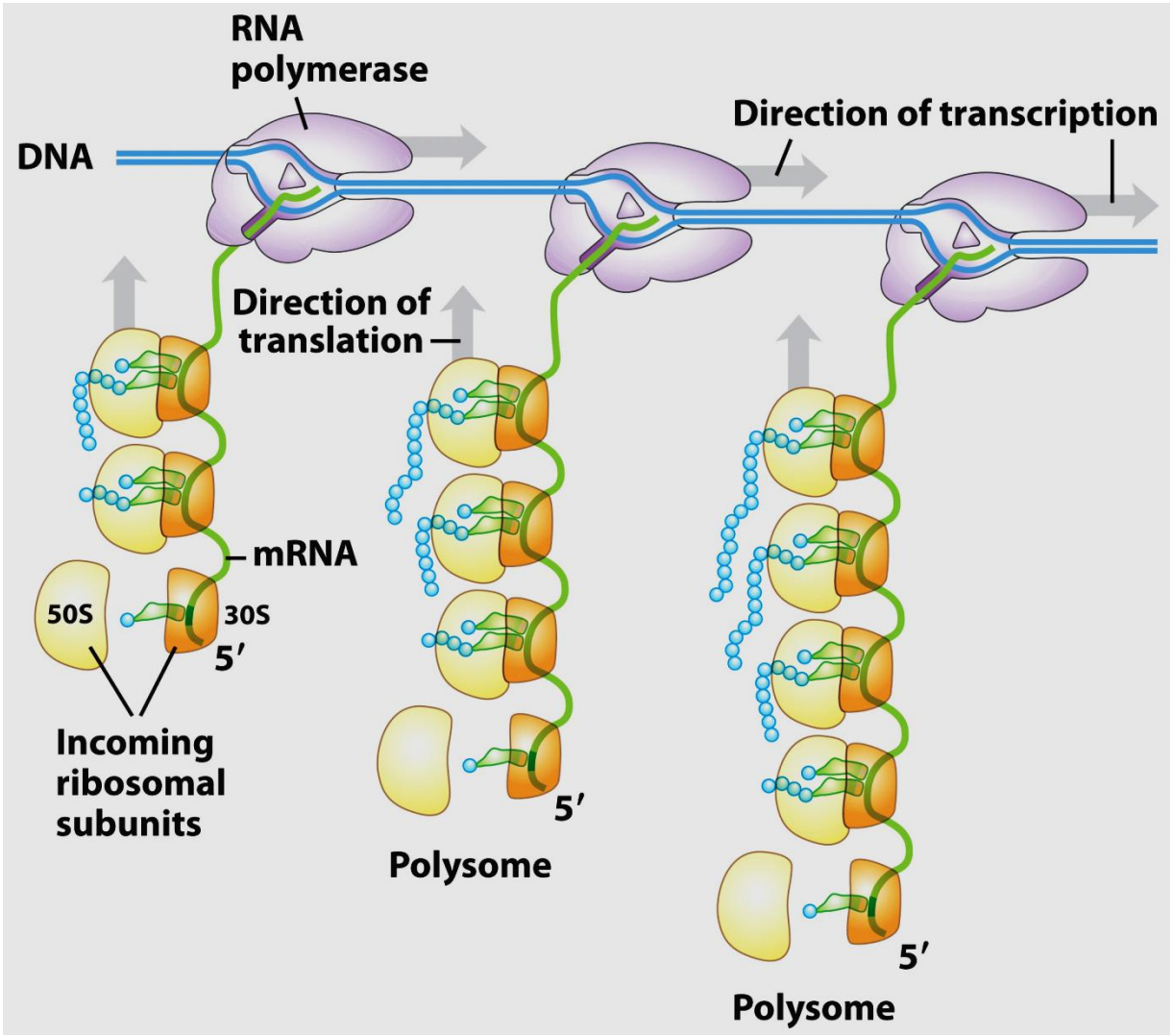

HHMI

Rapid translation of a single message by polysomes



Polysome: a cluster 10 to 100 of ribosomes on a single molecule of mRNA simultaneously synthesize proteins that allows the highly efficient use of the mRNA in both eukaryotic and bacterial cells.

Coupling of transcription and translation in bacteria



The mRNA is translated by ribosomes while it is still being transcribed from DNA by RNA polymerase.

Not in eukaryotic cells.

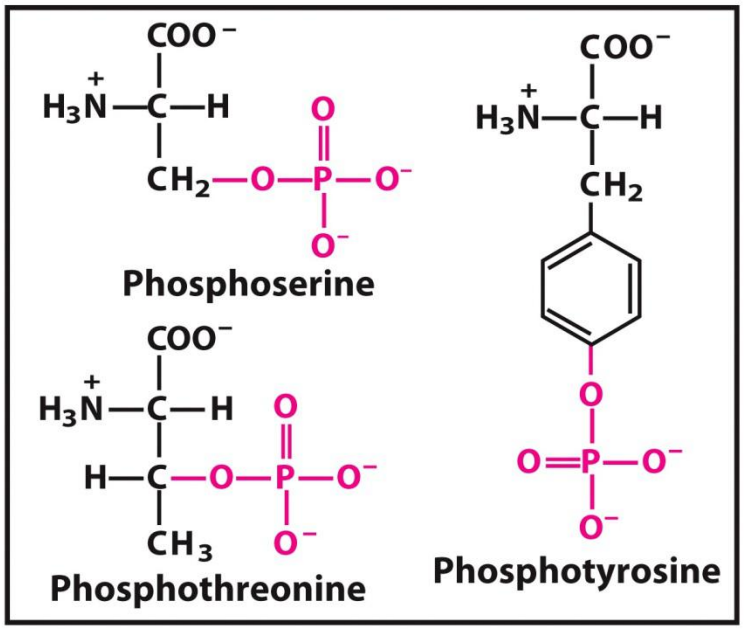
Stage 5: Protein folding & posttranslational modifications

- Amino-terminal and carboxyl-terminal modifications
- Loss of signal sequences
- Modification of individual amino acids: phosphorylation, methylation, acetylation, etc.
- Glycosylation: attachment of carbohydrate side chains
- Addition of isoprenyl groups: addition of groups derived from isoprene (isoprenyl groups).
- Addition of prosthetic Groups: biotin molecule of acetyl-CoA carboxylase and the heme group of hemoglobin or cytochrome *c*.
- Proteolytic processing
- Formation of disulfide cross-Links

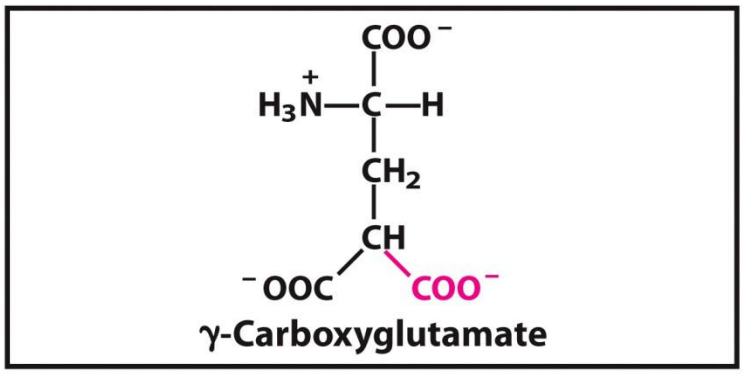
Why are proteins modified?

- Regulation of activity
  - modification may turn activity on
  - modification may turn activity off
  - modification may generate a different function
- Protein-protein interaction
  - modification site may be a binding interface
- Subcellular localization
  - modification site may be a targeting signal
  - modification may be a membrane anchor
- Aging
  - modification may identify the protein for degradation
  - modification may target a protein to be scavenged

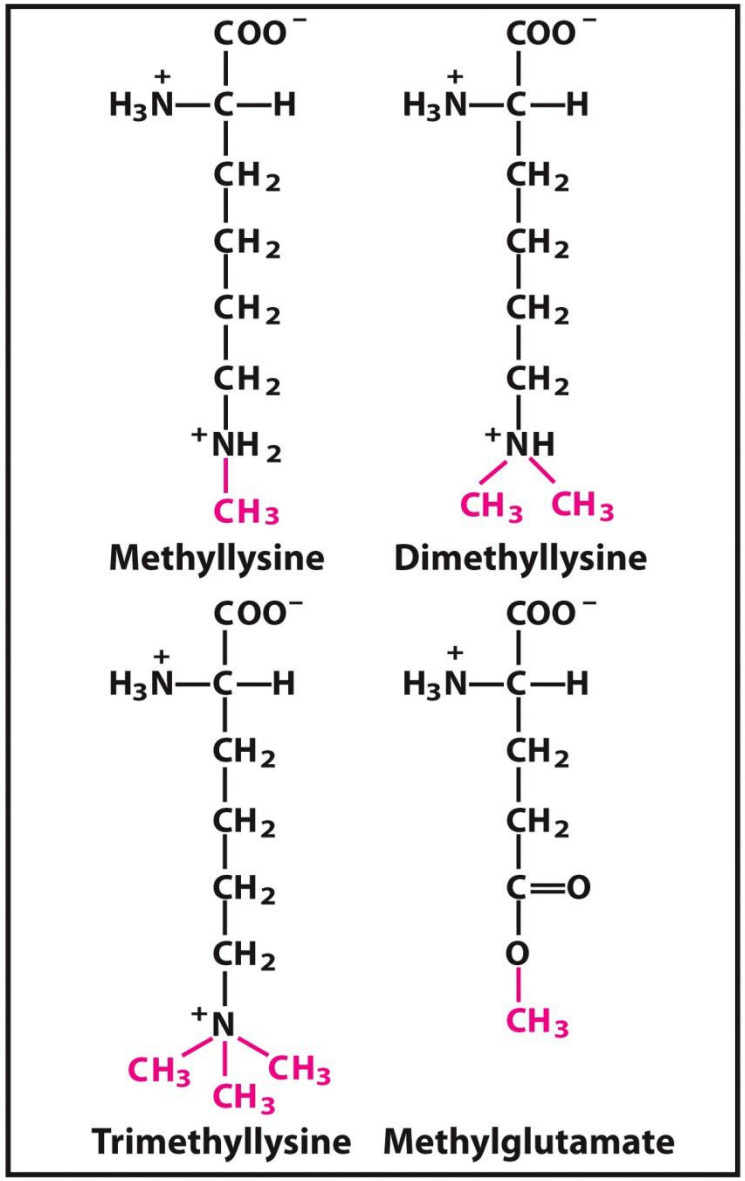
Some modified amino acid residues



(a)

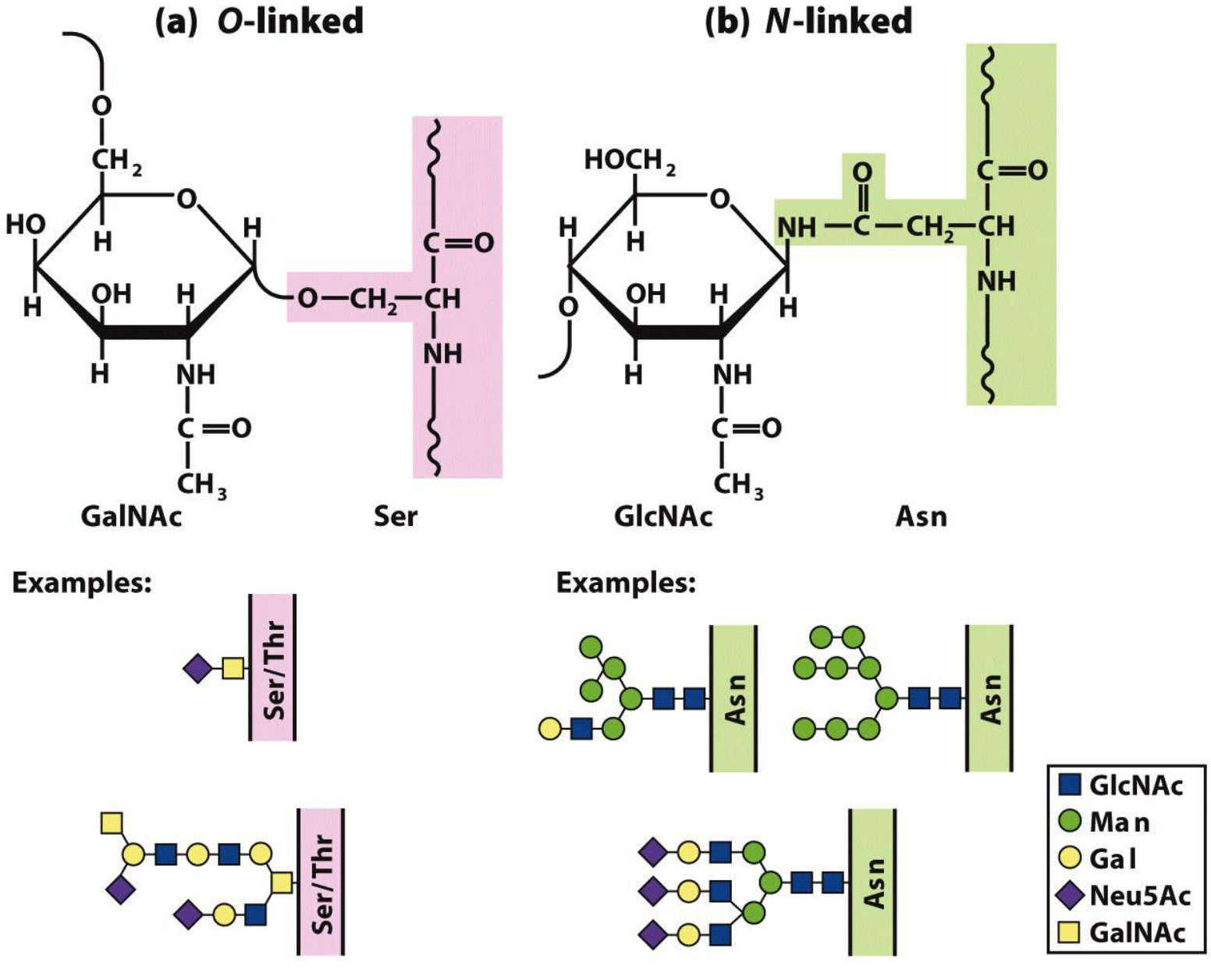


(b)

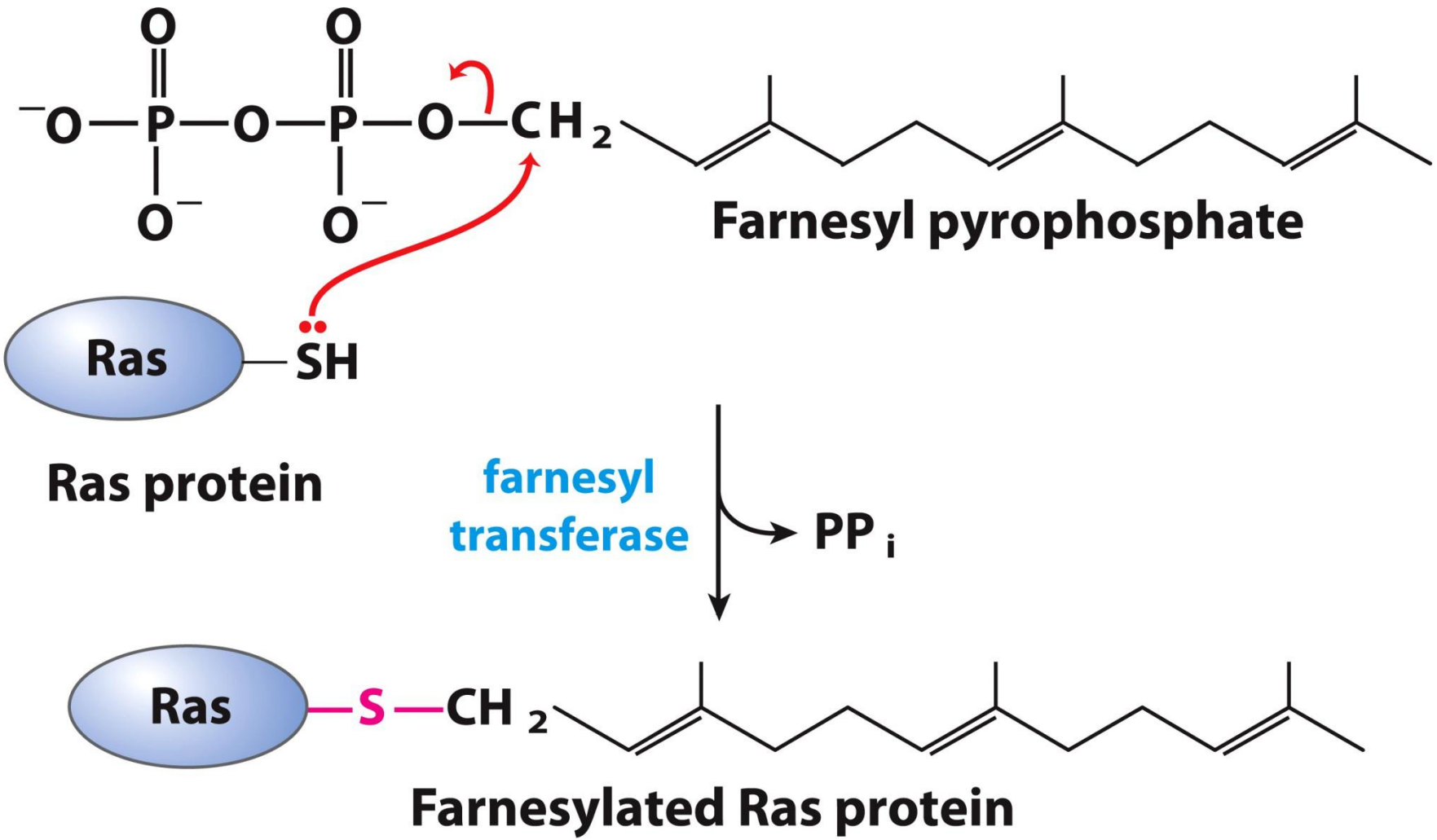


(c)

Oligosaccharide linkages in glycoproteins



Farnesylation of a Cys residue



Protein targeting

Signal sequence: a short sequence of amino acids in a protein that directs the protein to the final cellular destination

Targeting places:

nucleus,

endoplasmic reticulum (ER),

mitochondria,

chloroplasts,

secret out to cell matrix,

integration in the plasma membrane,

inclusion in lysosomes,

etc.

Targeting to ER

Posttranslational modification of many eukaryotic proteins begins in the ER (endoplasmic reticulum)

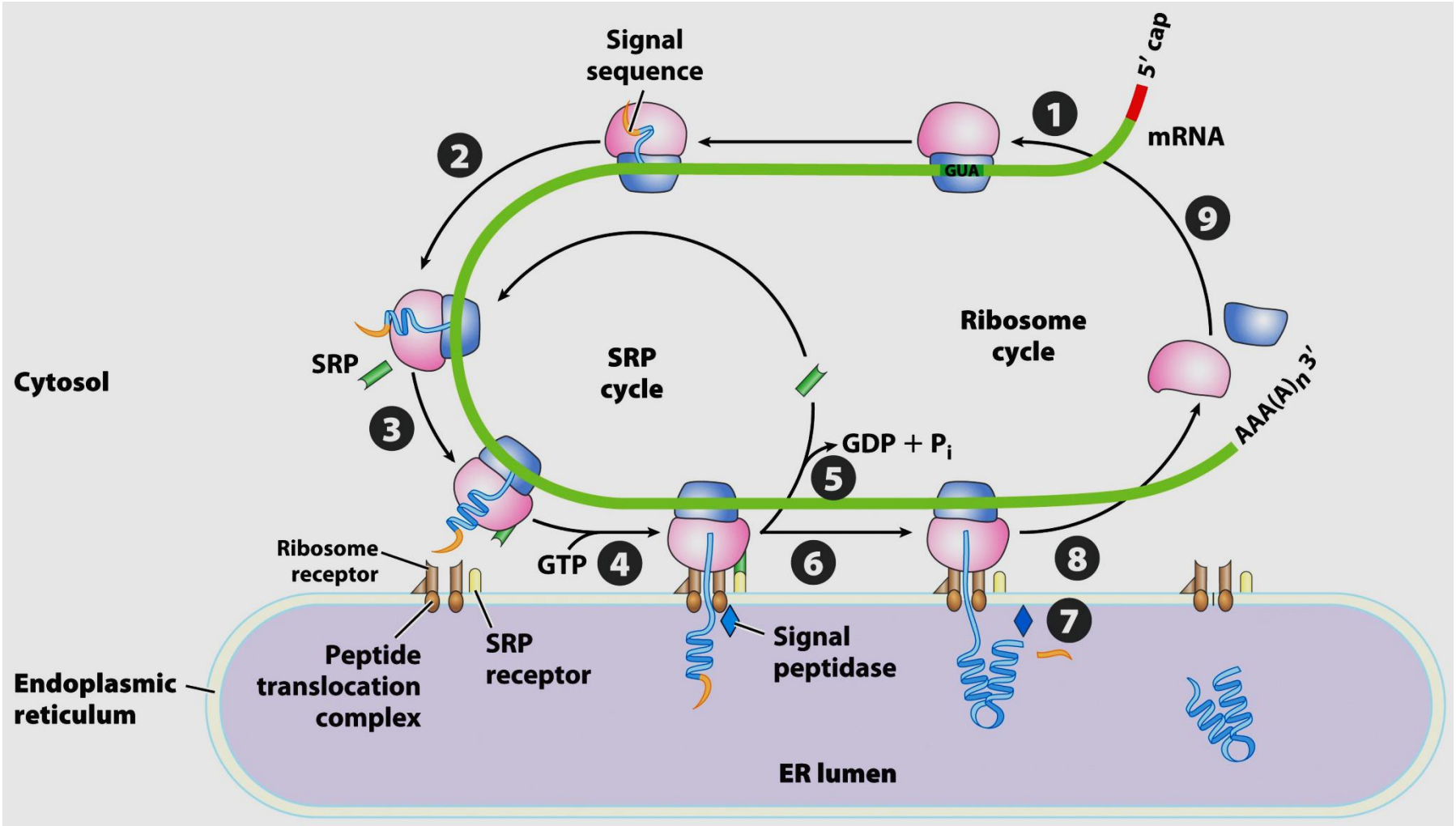
ER signal sequences (13 to 36 amino acids) have the following features:

(1) about 10 to 15 hydrophobic amino acid residues;

(2) one or more positively charged residues, usually near the amino terminus, preceding the hydrophobic sequence;

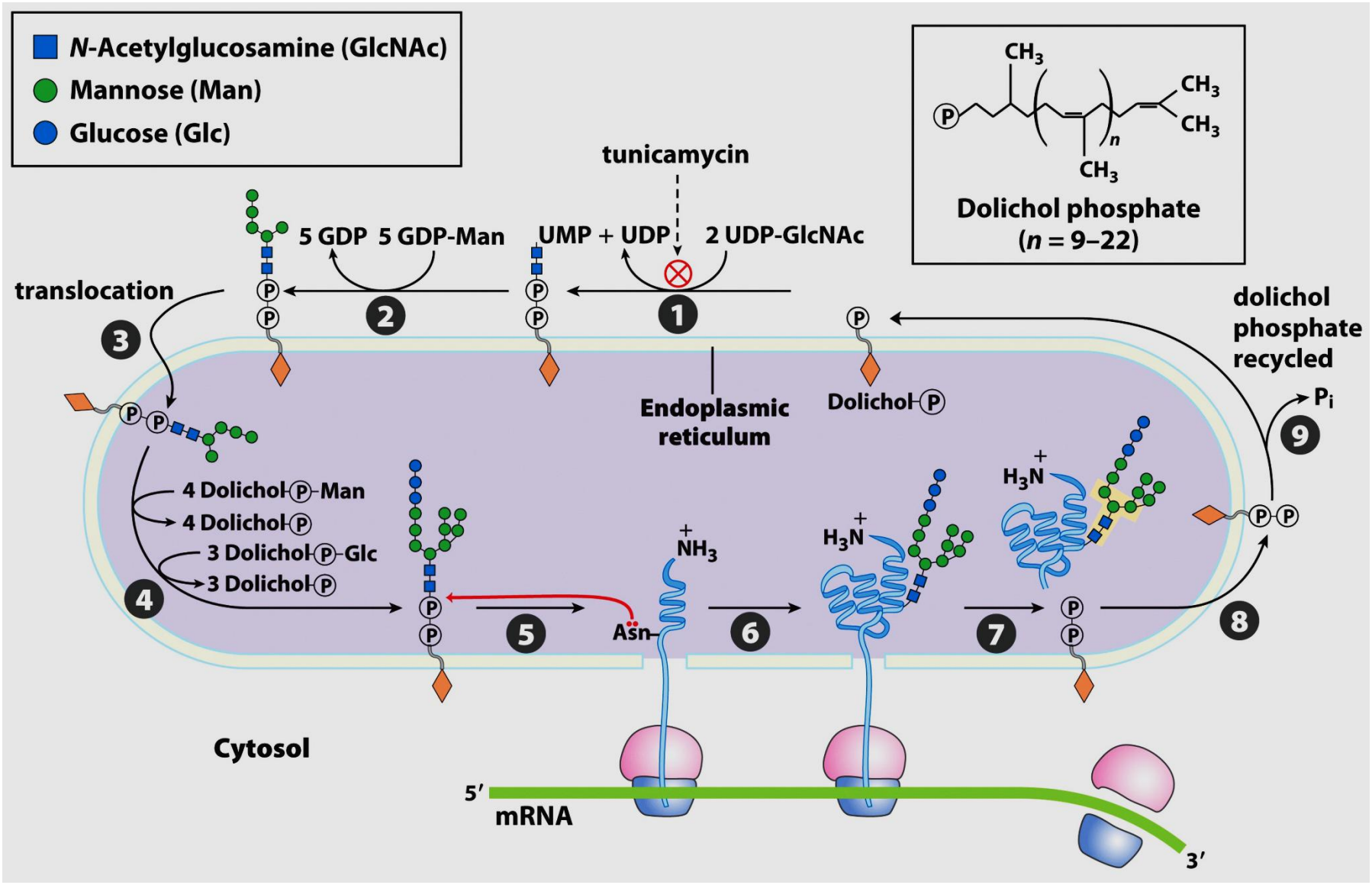
(3) a short sequence at the carboxyl terminus (near the cleavage site) that is relatively polar, typically having amino acid residues with short side chains (especially Ala) at the positions closest to the cleavage site.

Directing eukaryotic proteins to the endoplasmic reticulum

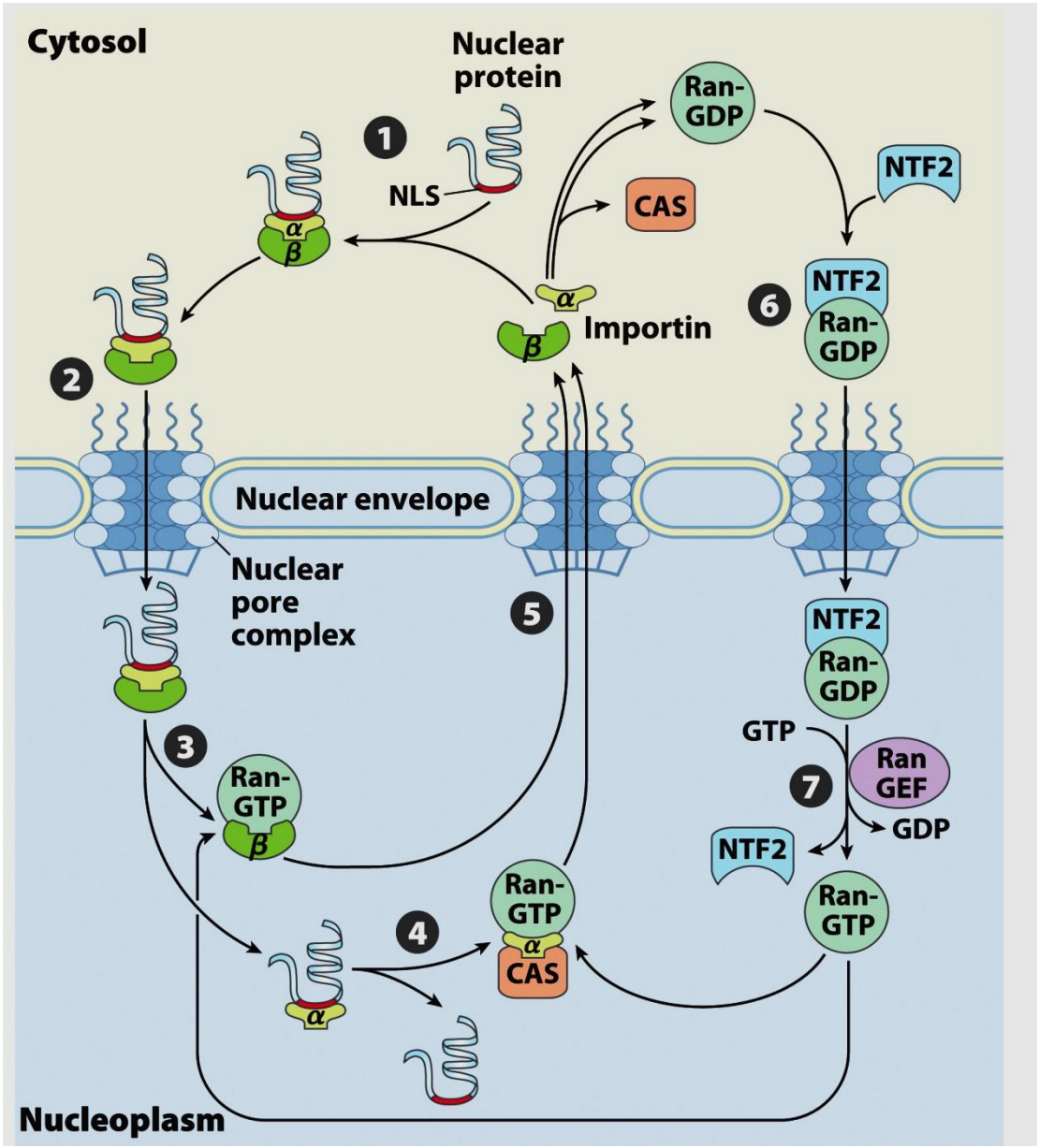


SRP: **Signal recognition particle**

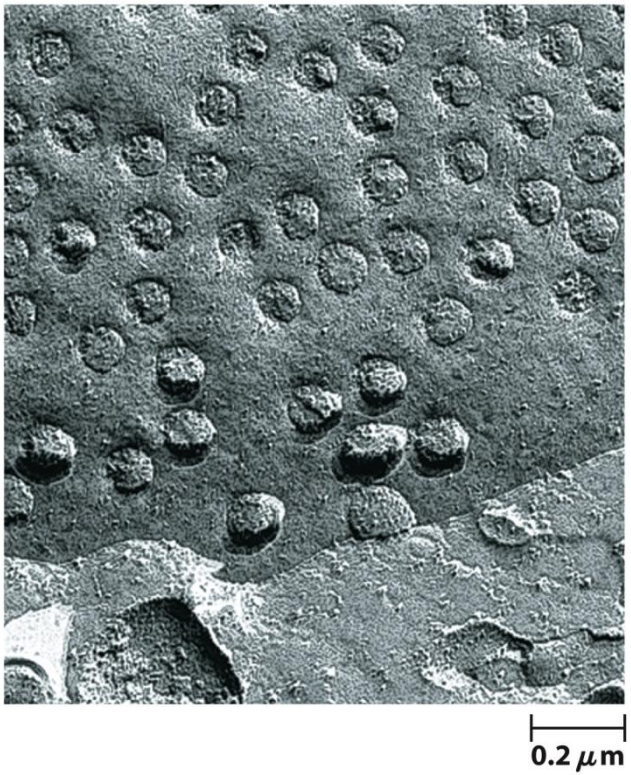
Synthesis of the core oligosaccharide of glycoproteins in ER



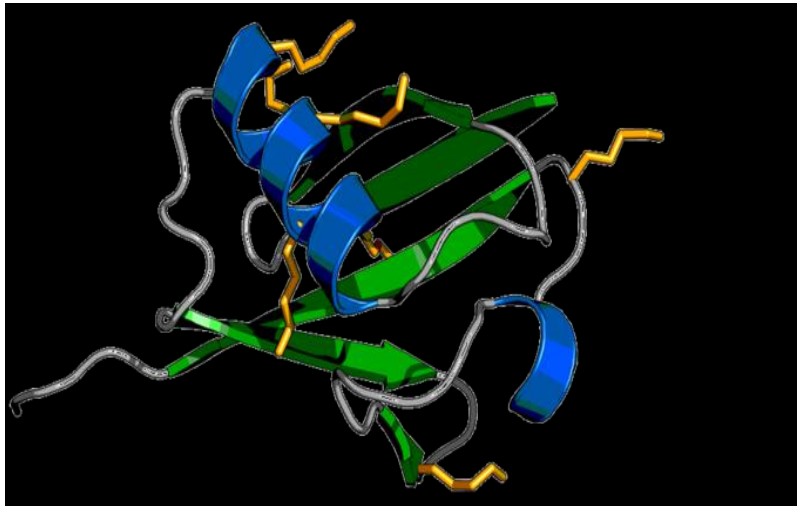
Targeting of nuclear proteins

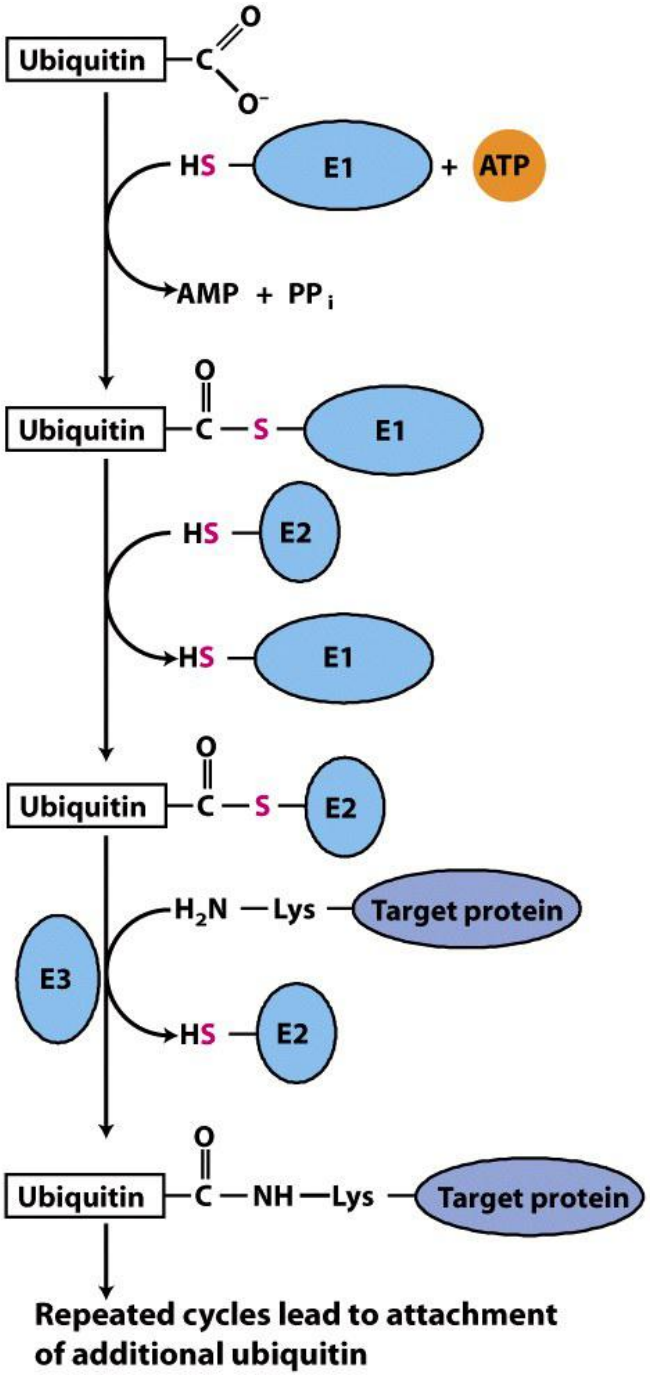


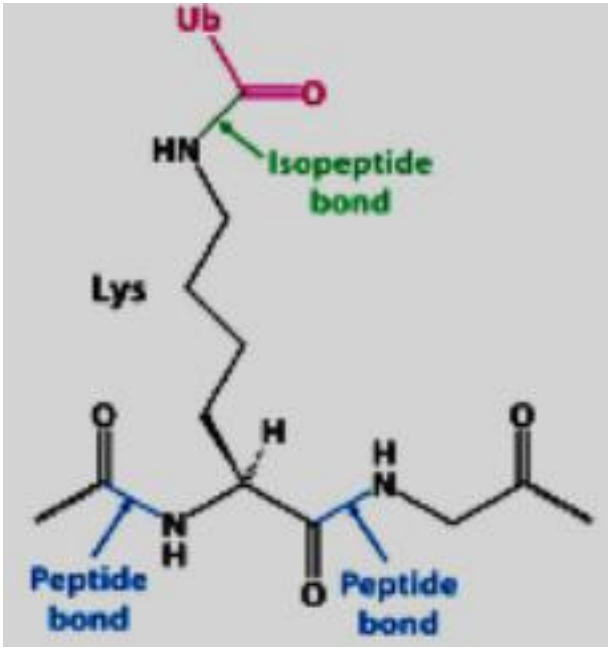
NLS: nuclear localization sequence



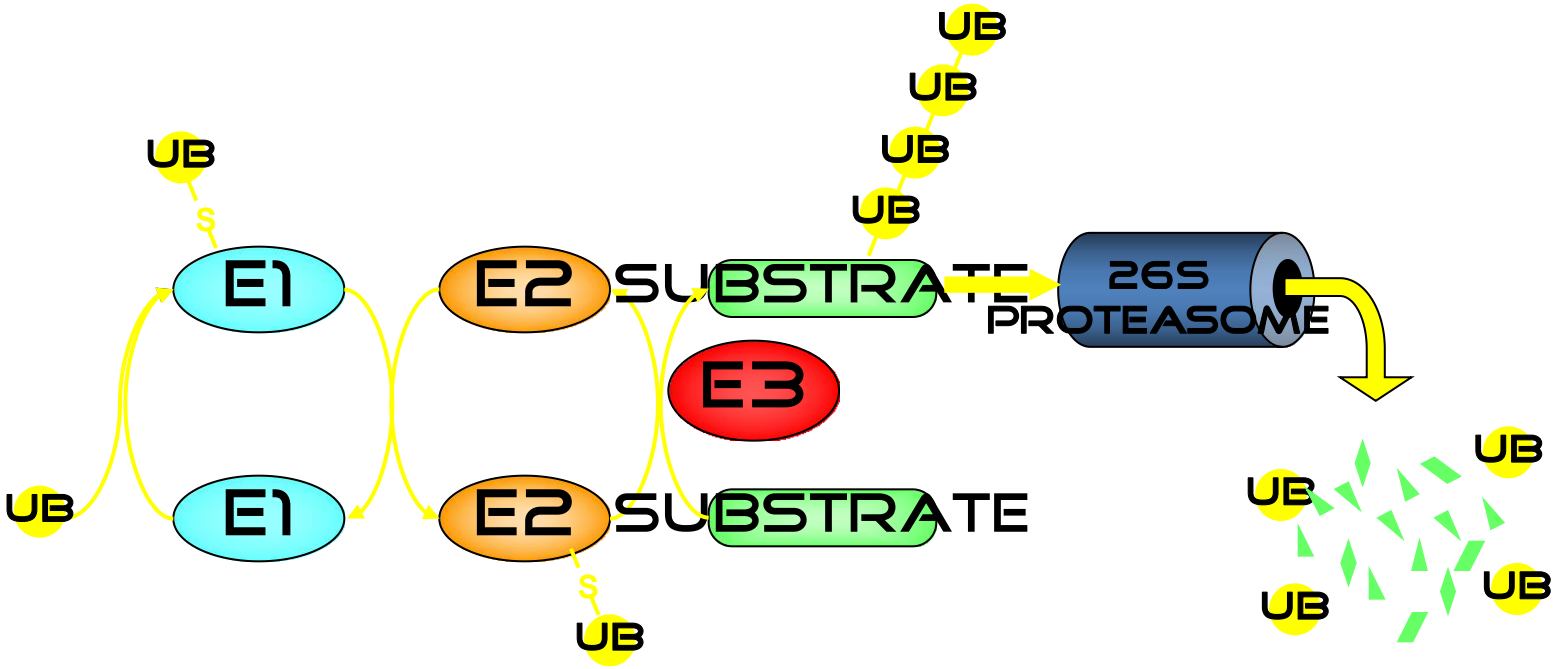
Ubiquitination







UBIQUITIN-PROTEASOME PA



Tightly controlled
protein degradation