MSB 101/NSB 111/BPT 111

Cell Biology, Genetics & General Embryology

Lecturer Dr. Kiptoo

Outline

- Genetic Code
- Protein Biosynthesis

Genetic Code

- In protein synthesis the Ribosomes synthesize proteins by reading the nucleotide sequence of mRNAs and polymerizing amino acids in an N–→C direction.
- The nucleotide sequence of an mRNA molecule is translated into the **amino acid** sequence of a protein molecule.
- ***To understand this process, the other questions need to be resolved***
  1) What is the **genetic code** that allows the information specified in a sequence of bases to be translated into the amino acid sequence of a polypeptide?
  2) How is the **4-letter language of nucleic acids** translated into the 20-letter language of proteins?
- Francis Crick reasoned that **adapter molecules** must bridge this information gap.
- These adapter molecules must interact specifically with both nucleic acids (mRNAs) and amino acids.
- At least 20 different adapter molecules would be needed, at least one for each amino acid.

- The **various adapter molecules** would be able to read the genetic code in an mRNA template and align the amino acids according to the template's directions so that they could be polymerized into a unique polypeptide.
- Transfer RNAs (tRNAs) are the adapter molecules. Amino acids are attached to the 3-OH at the 3-CCA end of tRNAs as aminoacyl esters.
- The formation of these aminoacyl-tRNAs, so-called **charged tRNAs,** is catalyzed by specific **aminoacyl-tRNA synthetases.**
- There is one of these enzymes for each of the 20 amino acids and each aminoacyl-tRNA synthetase loads its amino acid only onto tRNAs designed to carry it.
- In turn, these tRNAs specifically recognize unique sequences of bases in the mRNA through complementary base pairing.

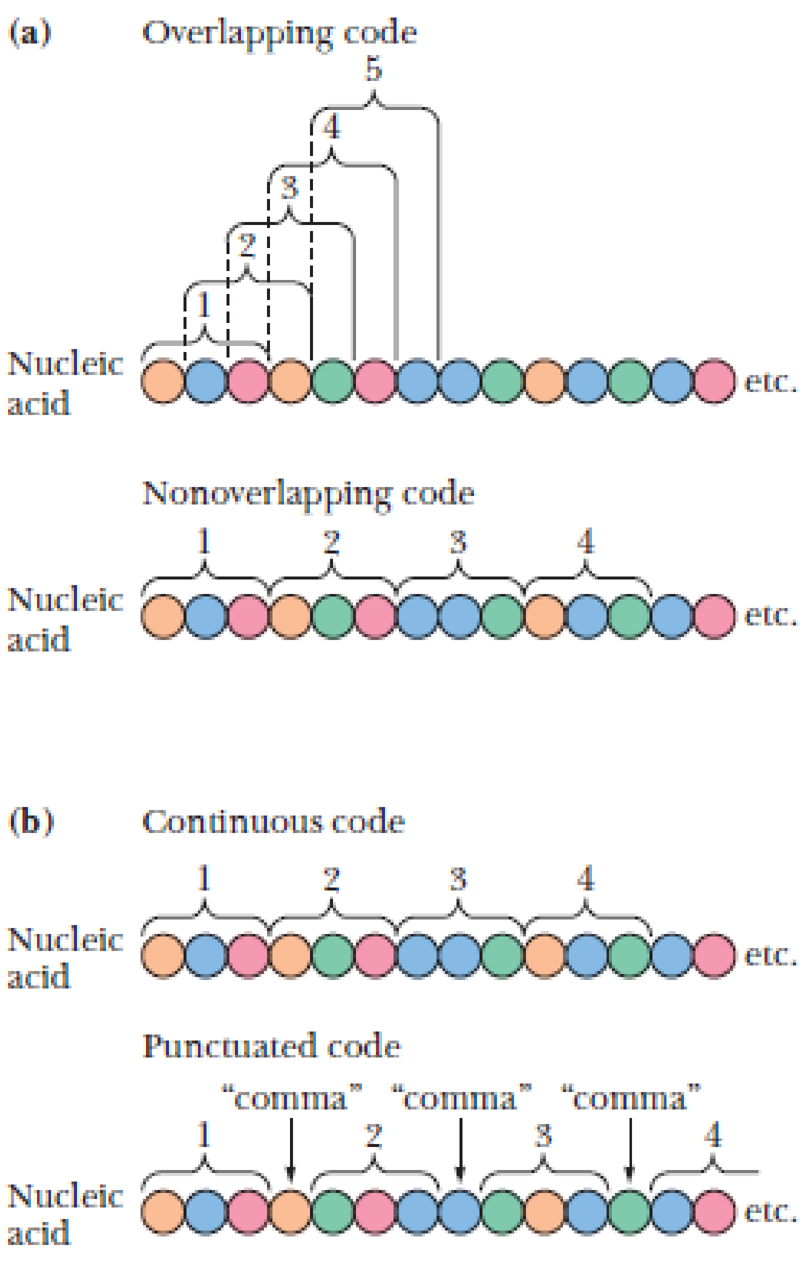


FIGURE 30.2 **(a)** An overlapping versus a nonoverlapping code. **(b)** A continuous versus a punctuated code.

What is Genetic Code

- Once it was realized that the **sequence of bases in a gene** specified the sequence of amino acids in a protein, various possibilities for such a genetic code were considered.
- How many bases are necessary to specify each amino acid?
  - a) Is the code overlapping or nonoverlapping?
  - b) Is the code punctuated or continuous?
- Mathematical considerations favored a **triplet of bases** as the minimal code word, or **codon,** for each amino acid:
- **A doublet code** based on pairs of the four possible bases, A, C, G, and U, has $4^2 = 16$ unique arrangements, an insufficient number to encode the 20 amino acids.
- **A triplet code** of four bases has $4^3 = 64$ possible code words, more than enough for the task.

The Genetic Code is a Triplet Code

- The genetic **code is a triplet code** read **continuously** from a fixed starting point in each mRNA.
- Specifically, it is defined by the following:
  1) A group of three bases codes for one amino acid.
  2) The code is not overlapping.
  3) The base sequence is read **from a fixed starting point without punctuation**. That is, the mRNA sequences contain no "commas" signifying appropriate groupings of triplets. If the reading frame is displaced by one base, it remains shifted throughout the subsequent message; no **"commas"** are present to restore the "correct" frame.
  4) The code is **degenerate,** meaning that, in most cases, each amino acid can be **coded by any of several triplets**. Recall that a triplet code yields 64 codons for 20 amino acids. Most codons (61 of 64) code for some amino acid.

Codons Specify Amino Acids

- Codons, like other nucleotide sequences, are read 5→3. Codons represent triplets of bases in mRNA or, replacing U with T, triplets along the nontranscribed (nontemplate) strand of DNA.

Several noteworthy features characterize the genetic code:

1. *All the codons have meaning.*

- Of the 64 codons, 61 specify particular amino acids. The remaining 3—UAA, UAG, and UGA—specify no amino acid and thus they are **nonsense codons.**
- Nonsense codons serve as **termination codons;** they are "stop" signals indicating that the end of the protein has been reached.

2. *The genetic code is unambiguous*

- Each of the 61 "sense" codons **encodes only one amino acid**.

3. *The genetic code is degenerate*

- With the exception of **Met** and **Trp,** every amino acid is coded by more than one codon. Several—Arg, Leu, and Ser—are represented by **six different codons**. Codons coding for the same amino acid are called **synonymous codons.**

4. *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, so, for example, all four codons in the GGX family specify Gly, and the UCX family specifies Ser. This feature is known as **third-base degeneracy.**
- Note also that codons with a **pyrimidine** as second base likely encode amino acids with **hydrophobic side chains**, and codons with a **purine** in the second-base position typically specify **polar** or **charged amino acids.**
- The two **negatively charged amino acids**, Asp and Glu, are encoded by GAX codons; GA pyrimidine gives Asp and GA–purine specifies Glu.
- The consequence **of these** similarities is that mutations are **less likely to be harmful** because single base changes in a codon will result either in no change or in a substitution with an amino acid similar to the original amino acid.
- The **degeneracy of the code** is evolution's buffer against mutational disruption.

5. *The genetic code is "universal."*

- Although certain minor exceptions in codon usage occur, the more striking feature of the code is its **universality**: Codon assignments are virtually the same **throughout all organisms**—archaea, bacteria, and eukaryotes.
- **This conformity** means that all extant organisms use the same genetic code, providing strong evidence that they all evolved from a common primordial ancestor.

The complete translation of the Genetic code

TABLE 30.1 The Genetic Code

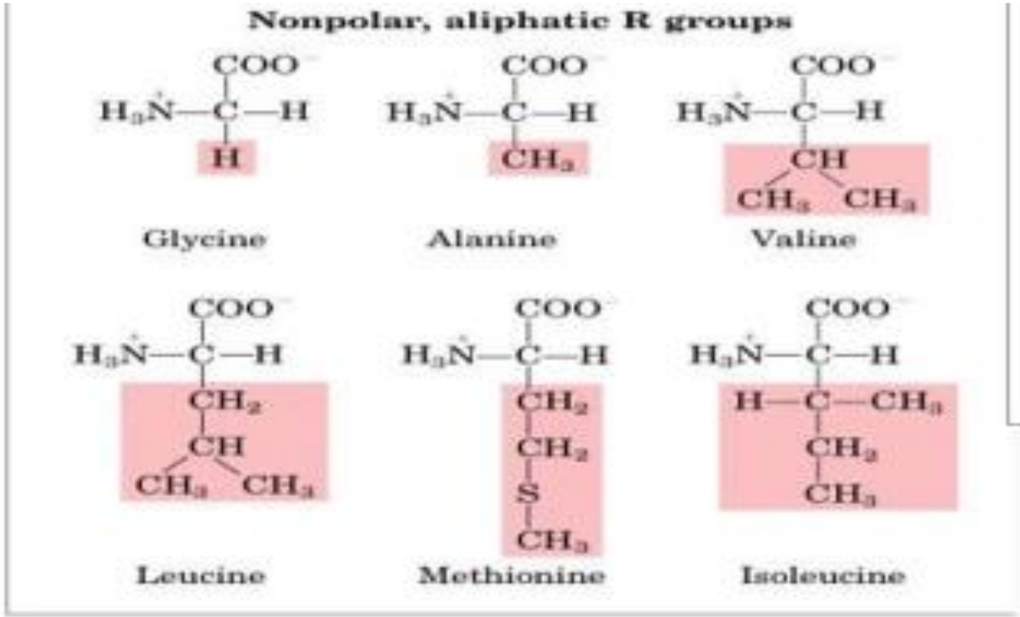
First Position (5′-end)	Second Position U	C	A	G	Third Position (3′-end)
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA Stop	UGA Stop	A
	UUG Leu	UCG Ser	UAG Stop	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	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

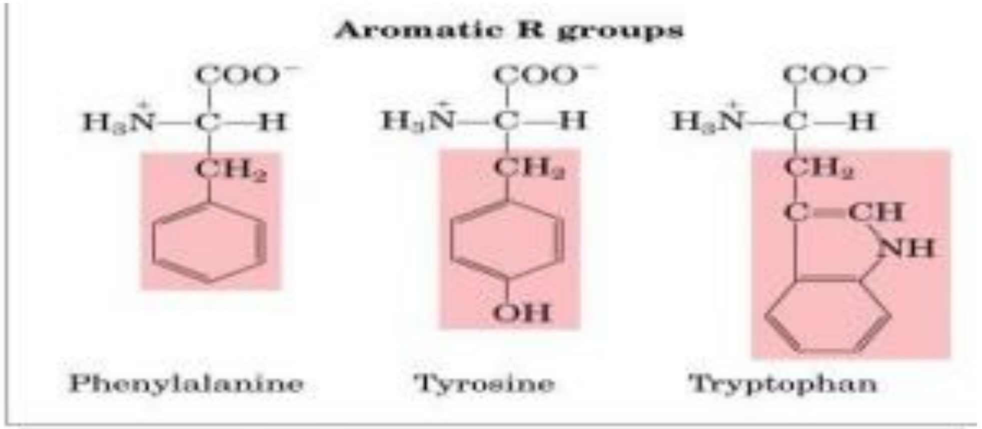
Third-Base Degeneracy Is Color-Coded

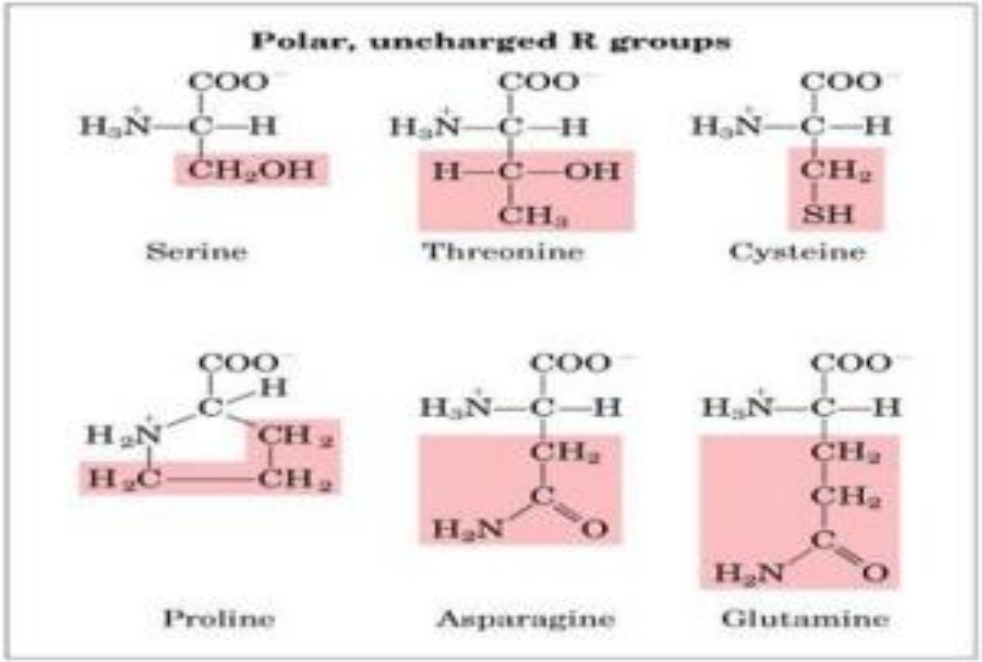
Third-Base Relationship	Third Bases with Same Meaning	Number of Codons
Third base irrelevant	U, C, A, G	32 (8 families)
Purines	A or G	12 (6 pairs)
Pyrimidines	U or C	14 (7 pairs)
Three out of four	U, C, A	3 (AUX = Ile)
Unique definitions	G only	2 (AUG = Met) (UGG = Trp)
Unique definition	A only	1 (UGA = Stop)

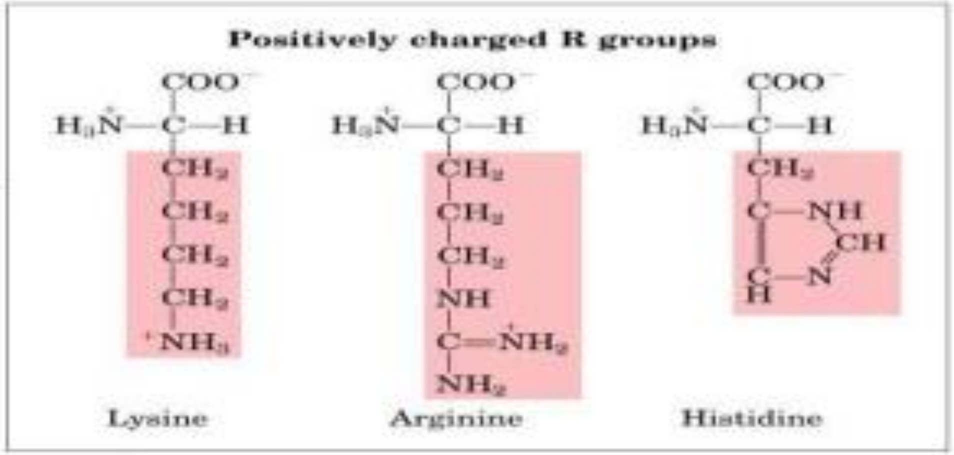
*AUG signals translation initiation as well as coding for Met residues.

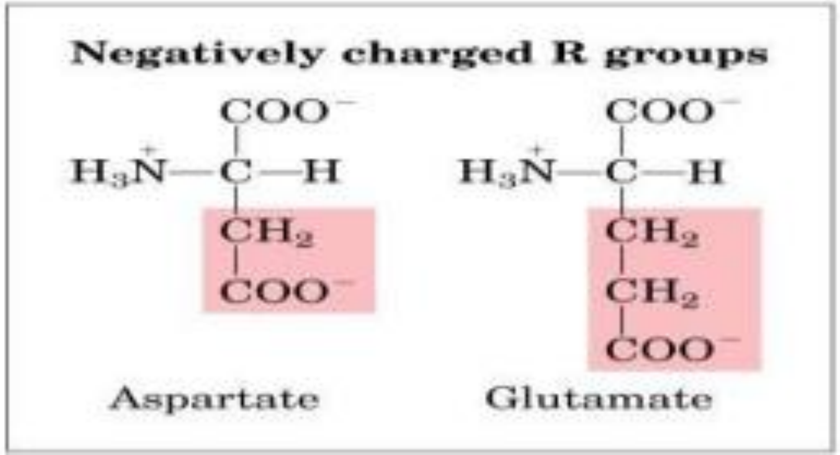
Amino Acids Classification











Protein Biosynthesis

- Protein synthesis can be described in three phases: **Initiation, Elongation and Termination**

Initiation

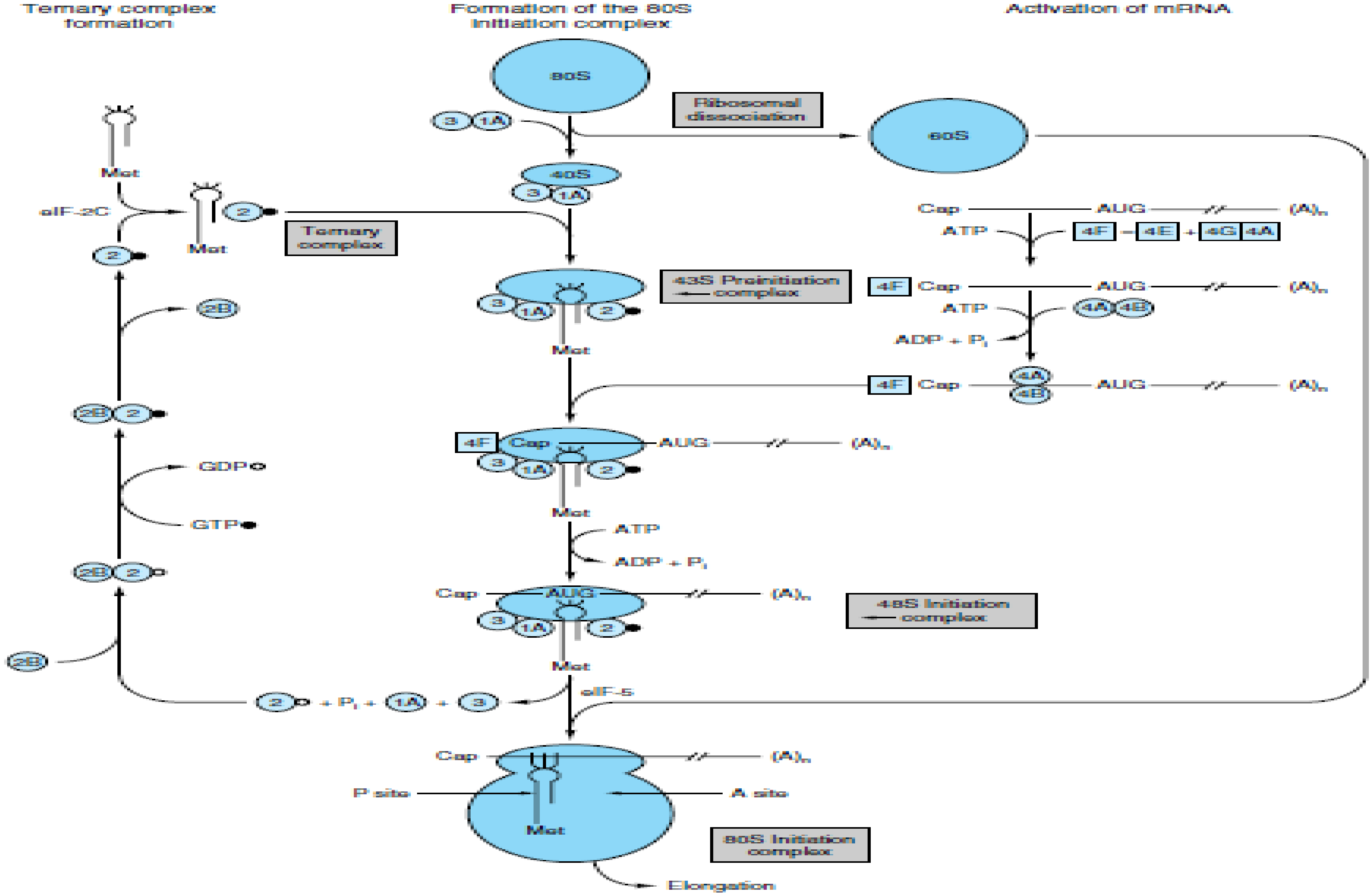
- The translation of the mRNA commences near its 5′ terminal with the formation of the corresponding amino terminal of the protein molecule.
- The message is read from 5′ to 3′, concluding with the formation of the carboxyl terminal of the protein.
- Initiation Involves Several Protein-RNA Complexes
- Initiation of protein synthesis requires that an mRNA molecule be selected for translation by a ribosome.
- Once the mRNA binds to the ribosome, the latter finds the correct reading frame on the mRNA, and translation begins.
- This process involves tRNA, rRNA, mRNA, and at **least ten eukaryotic initiation factors (eIFs),** some of which have multiple (three to eight) subunits.
- Also involved are GTP, ATP, and amino acids.

➢ ***Initiation can be divided into <u>four steps</u>***

A. Dissociation of the ribosome into its 40S and 60S subunits.
B. Binding of a ternary complex consisting of met-tRNAi, GTP, and eIF-2 to the 40S ribosome to form a preinitiation complex.
C. Binding of mRNA to the 43S preinitiation complex to form a 48S initiation complex.
D. Combination of the 48S initiation complex with the 60S ribosomal subunit to form the 80S initiation complex.

A. RIBOSOMAL DISSOCIATION

- Two initiation factors, **eIF-3 and eIF-1A**, bind to the newly dissociated 40S ribosomal subunit.
- This delays its reassociation with the 60S subunit and allows other translation initiation factors to associate with the 40S subunit.
- The first step in this process involves the binding of GTP by eIF-2. This **binary complex** then binds to met-$tRNA^{i}$, a tRNA specifically involved in binding to the initiation codon AUG.
- (There are two tRNAs for methionine. **One** specifies methionine for the initiator codon, the **other** for internal methionines. Each has a unique nucleotide sequence.)
- This **ternary complex** binds to the 40S ribosomal subunit to form the **43S preinitiation complex**, which is stabilized by association with **eIF-3** and **eIF-1A.**


Ternary complex formation
Formation of the 80S initiation complex
Activation of mRNA
80S
3
1A
Ribosomal dissociation
60S
40S
Met
eIF-2C
2
Ternary complex
Cap
AUG
(A)n
ATP
4F
4E
4G
4A
43S Preinitiation complex
2B
4B
ADP + Pi
GDP
GTP
48S Initiation complex
eIF-5
2 + Pi + 1A + 3
P site
A site
80S Initiation complex
Elongation

B. FORMATION OF THE 43S INITIATION COMPLEX

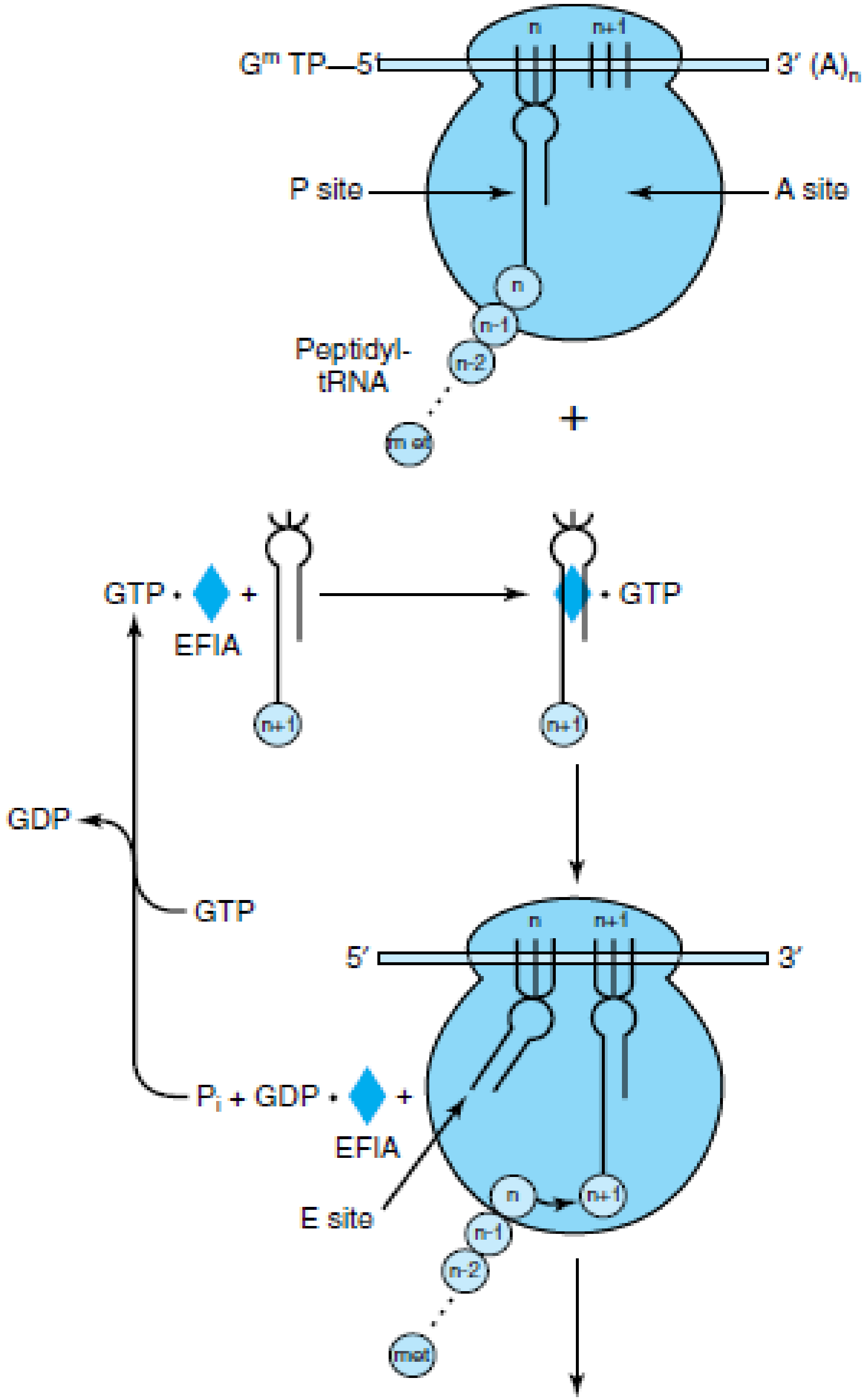
- The 5′ terminals of most mRNA molecules in eukaryotic cells are "**capped,**" with **methyl-guanosyl triphosphate cap** that facilitates the binding of mRNA to the 43S preinitiation complex.
- A **cap binding protein complex**, eIF-4F (4F), which consists of eIF-4E and the eIF-4G (4G)-eIF4A (4A) complex, binds to the cap through the 4E protein.
- Then eIF-4A (4A) and eIF-4B (4B) bind and **reduce the complex secondary structure of the 5′ end of the mRNA through ATPase and ATP-dependent helicase activities.**
- The association of mRNA with the 43S preinitiation complex to form the **48S initiation complex requires ATP hydrolysis.**
- Following **association of the 43S preinitiation complex with the mRNA cap** and reduction ("melting") of the secondary structure near the 5′ end of the mRNA, the complex scans the mRNA for a suitable initiation codon.
- Generally this is the 5′-most AUG, but the precise initiation codon is determined by so-called **Kozak consensus sequences** that surround the AUG: GCCA / GCCAUGG
- Most preferred is the **presence of a purine** at positions −3 and +4 relative to the AUG.

C. ROLE OF THE POLY(A) TAIL IN INITIATION

- The 3′ poly(A) tail and its binding protein, **Pab1p**, are required for efficient initiation of protein synthesis.
- Further, it is required to stimulates recruitment of the 40S ribosomal subunit to the mRNA through a complex set of interactions.
- **Pab1p**, bound to the poly(A) tail, interacts with eIF-4G, which in turn binds to eIF-4E that is bound to the cap structure.
- It is possible that a circular structure is formed and that this helps direct the 40S ribosomal subunit to the 5′ end of the mRNA.
- This helps explain how the cap and poly(A) tail structures have a synergistic effect on protein synthesis.

D. FORMATION OF THE 80S INITIATION COMPLEX

- The binding of the **60S ribosomal subunit** to the **48S initiation complex** involves hydrolysis of the GTP bound to eIF-2 by eIF-5.
- This reaction results in **release of the initiation factors** bound to the 48S initiation complex (these factors then are recycled) and the **rapid association of the 40S and 60S subunits to form the 80S ribosome**.
- At this point, the met-tRNA^{i} is on the **P site** of the ribosome, ready for the elongation cycle to commence.

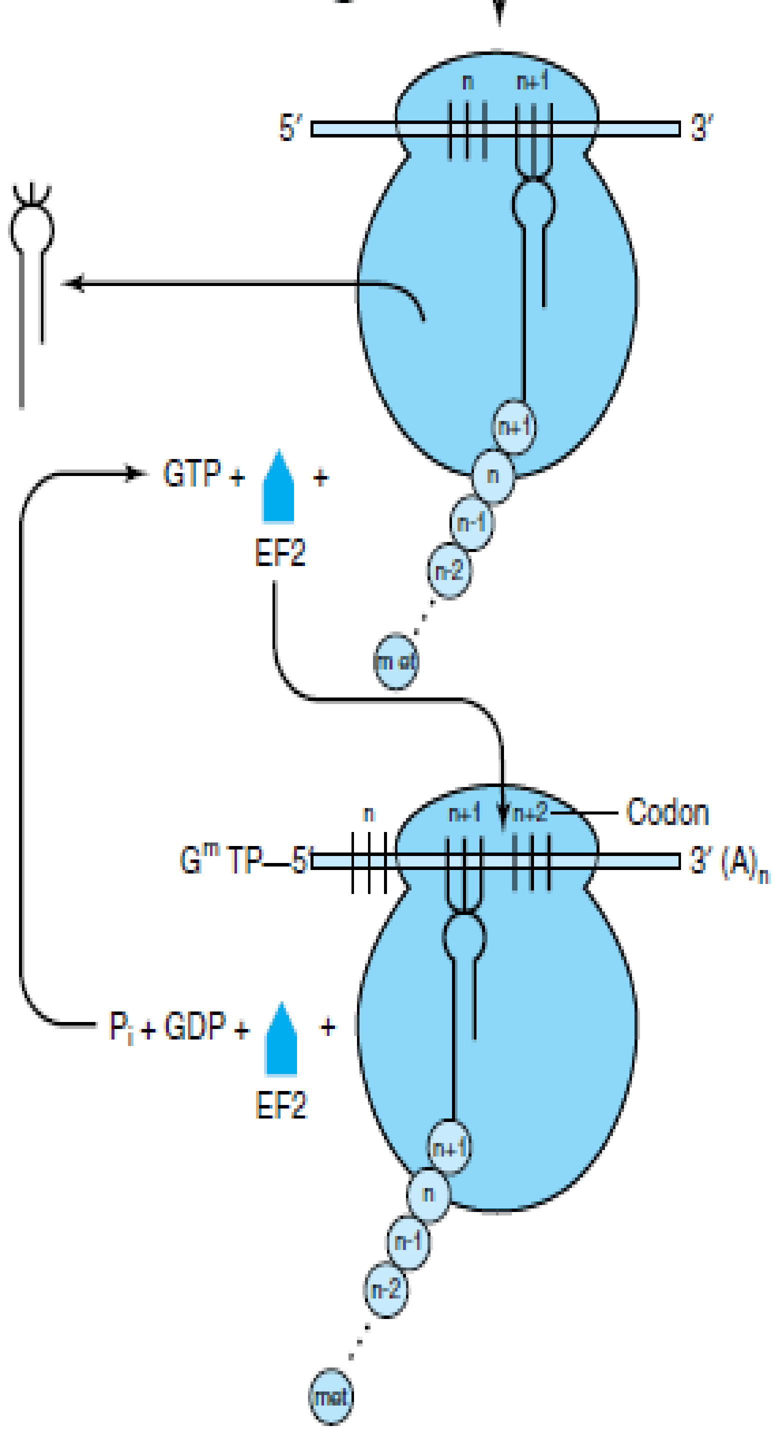


Elongation

- *It is also Is a Multistep Process*
- Elongation is a cyclic process on the ribosome in which one amino acid at a time is added to the nascent peptide chain.
- The **peptide sequence is determined by the order of the codons in the mRNA**.
- *Elongation involves several steps catalyzed by proteins called elongation factors (EFs).*
- *These steps are:*
    - A. Binding of aminoacyl-tRNA to the A site,
    - B. Peptide bond formation,
    - C. Translocation.

A. BINDING OF AMINOACYL-TRNA TO THE A SITE

- In the complete 80S ribosome formed during the **process of initiation**, the A site (aminoacyl or acceptor site) is free.
- The binding of the **proper aminoacyltRNA in the A site** requires proper codon recognition.
- **Elongation factor EF1A** forms a ternary complex with GTP and the entering aminoacyl-tRNA.
- **This complex** then allows the aminoacyl-tRNA to enter the A site with the release of EF1A•GDP and phosphate.
- GTP hydrolysis is catalyzed by an active site on the ribosome. The EF1A-GDP is then recycles to EF1A-GTP with the aid of other soluble protein factors and GTP.

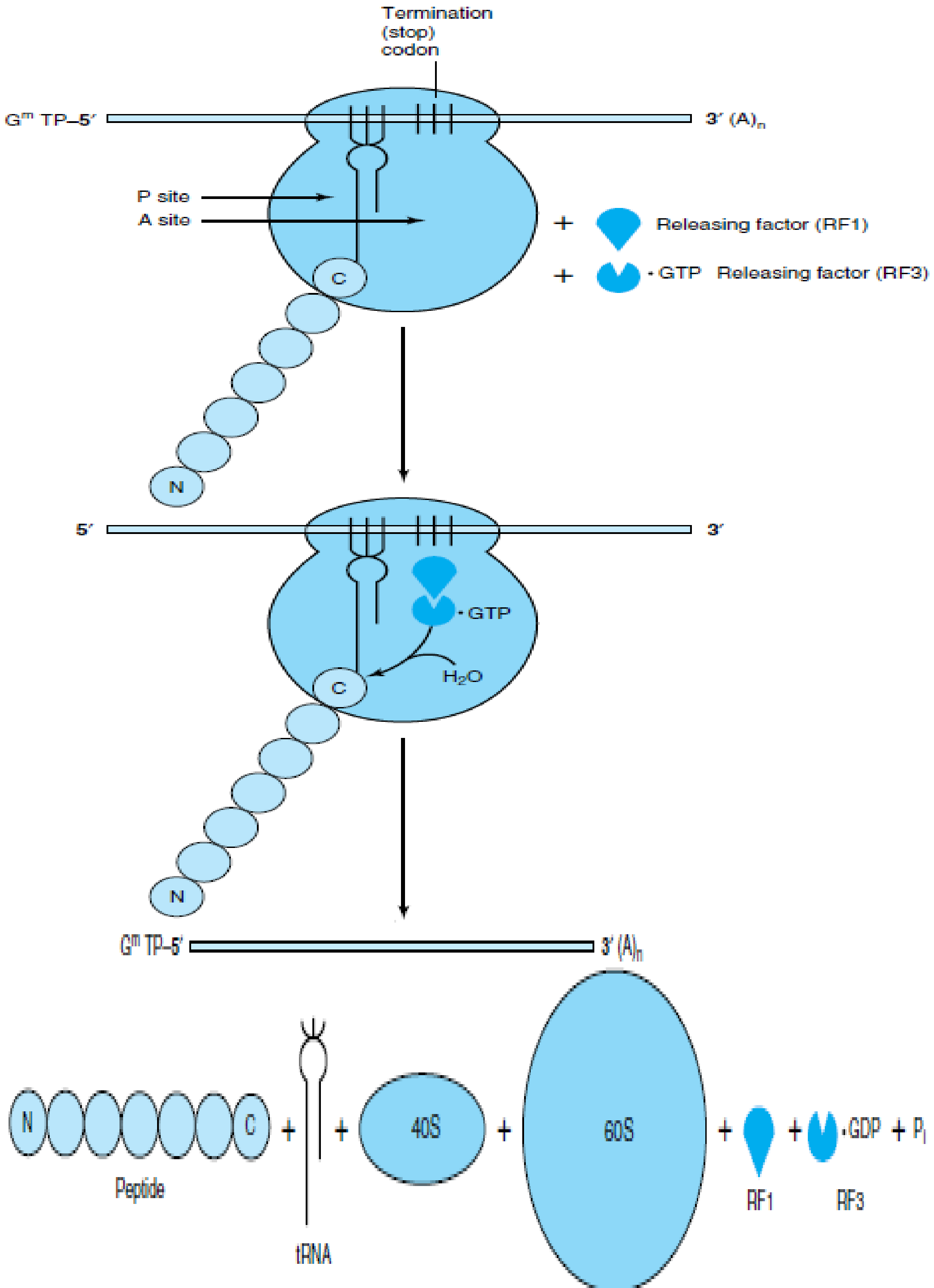


B. PEPTIDE BOND FORMATION

- The **α-amino group** of the **new aminoacyl-tRNA** in the A site carries out a nucleophilic attack on the **esterified carboxyl group of the peptidyl-tRNA occupying the P site** (peptidyl or polypeptide site).
- **At initiation**, this site is occupied by aminoacyl-tRNA meti.
- This reaction is catalyzed by a **peptidyltransferase,** a component of the 28S RNA of the 60S ribosomal subunit.
- And because the amino acid on the aminoacyl-tRNA is already "activated," no further energy source is required for this reaction.
- The reaction results in attachment of the growing peptide chain to the tRNA in the A site.

C. TRANSLOCATION

- The deacylated tRNA is attached by its anticodon to the P site at one end and by the open CCA tail to an **exit (E) site** on the large ribosomal subunit.
- At this point, **elongation factor 2 (EF2)** binds to and displaces the peptidyl tRNA from the A site to the P site.
- In turn, the deacylated tRNA is on the E site, from which it leaves the ribosome.
- The **EF2-GTP complex** is hydrolyzed to EF2-GDP, effectively moving the mRNA forward by one codon and leaving the A site open for occupancy by another **ternary complex of amino acid tRNA-EF1A-GTP** and another cycle of elongation.



Termination

- It occurs when a stop codon is recognized
- After multiple cycles of elongation culminating in polymerization of the specific amino acids into a protein molecule, the stop or terminating codon of mRNA (UAA, UAG, UGA) appears in the A site.
- Normally, there is no tRNA with an anticodon capable of recognizing such a termination signal.
- **Releasing factor RF1** recognizes that a stop codon resides in the A site.
- RF1 is bound by a **complex** consisting of **releasing factor RF3** with bound GTP.
- This complex, with the **peptidyl transferase**, promotes hydrolysis of the bond **between the peptide and the tRNA** occupying the **P site**.
- Thus, a **water molecule** rather than an amino acid is added.
- This hydrolysis releases the protein and the tRNA from the P site.
- Upon hydrolysis and release, the **80S ribosome dissociates** into its 40S and 60S subunits, which are then recycled.
- Therefore, the releasing factors are proteins **that hydrolyze the peptidyl-tRNA bond when a stop codon occupies the A site.**
- The **mRNA is then released** from the ribosome, which **dissociates into its component 40S and 60S subunits,** and another cycle can be repeated.

END