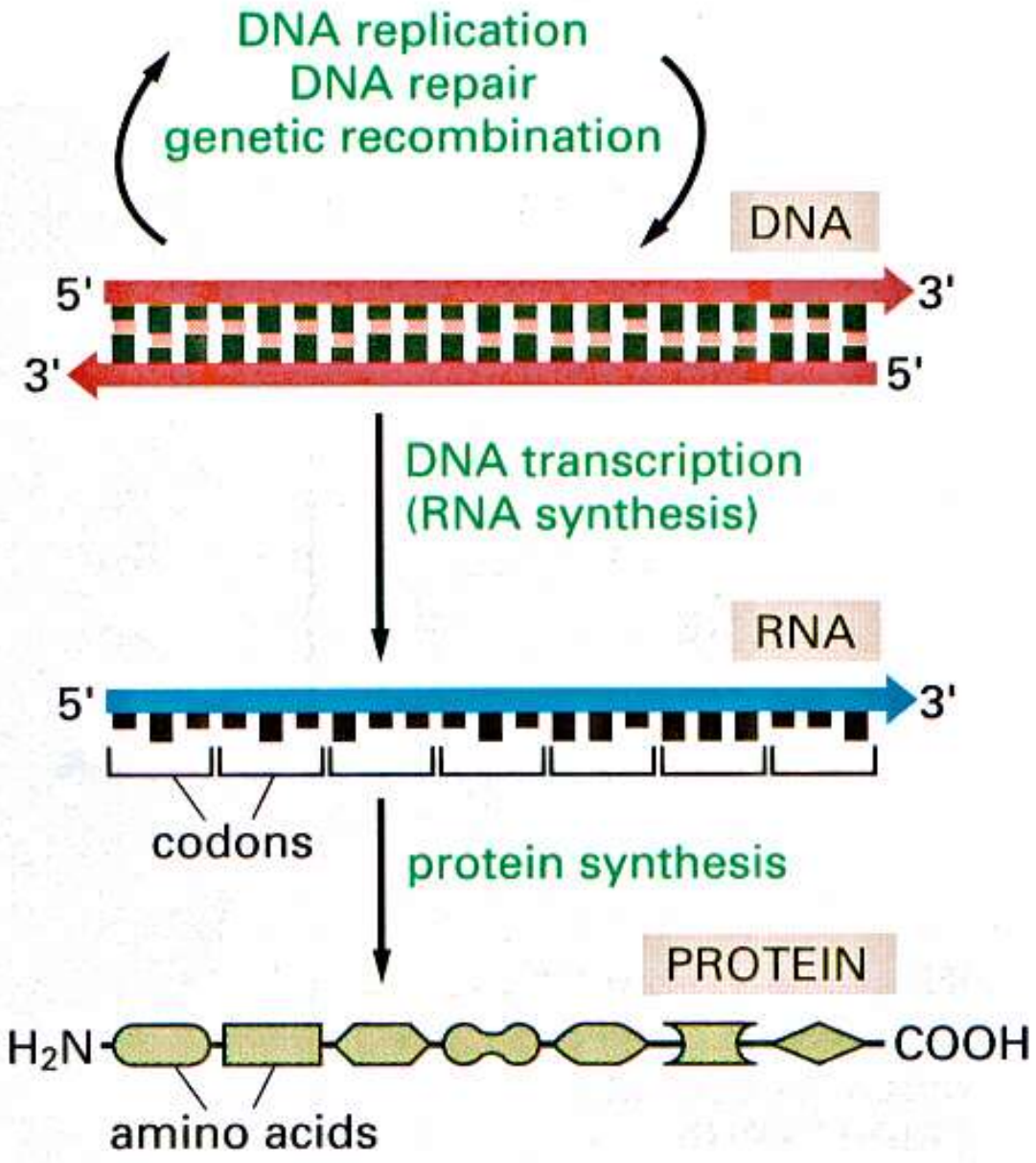
Peter Pristaš

Molecular biology

DNA translation

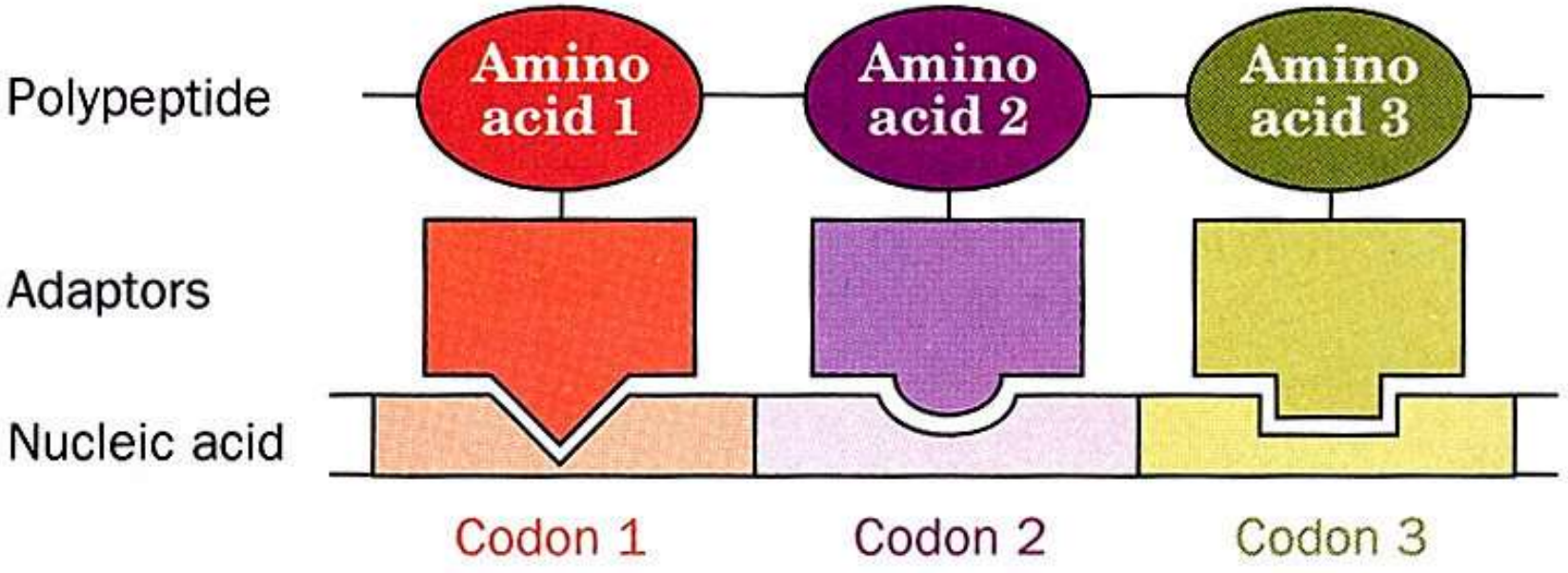
The Central Dogma of Molecular Biology



Adaptor molecule - tRNA

Protein Synthesis

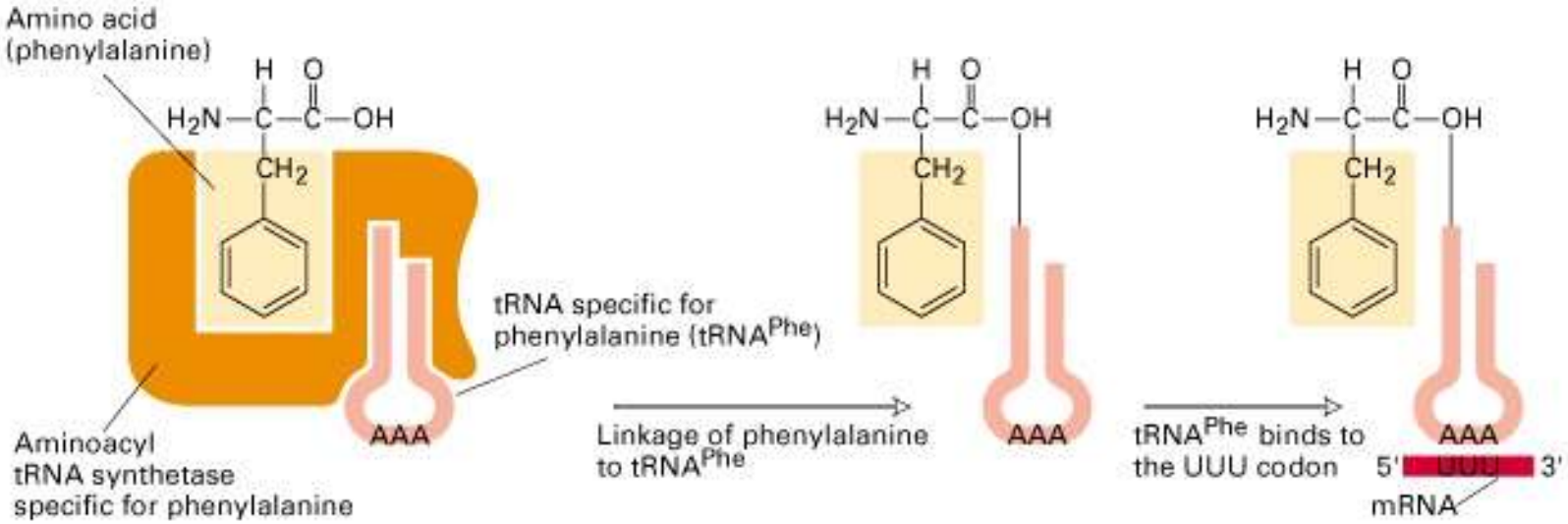
- Adaptor molecule
  - Proteins are coded directly from the mRNA with 3 bases (one codon) for each amino acid.



Genetic code

Deciphering the code *in vivo*

- Correlating a codon on mRNA with an amino acid in a proteins requires 2 levels of specificity
  - Correct amino acid must be attached to correct tRNA
  - tRNA must recognize correct codons

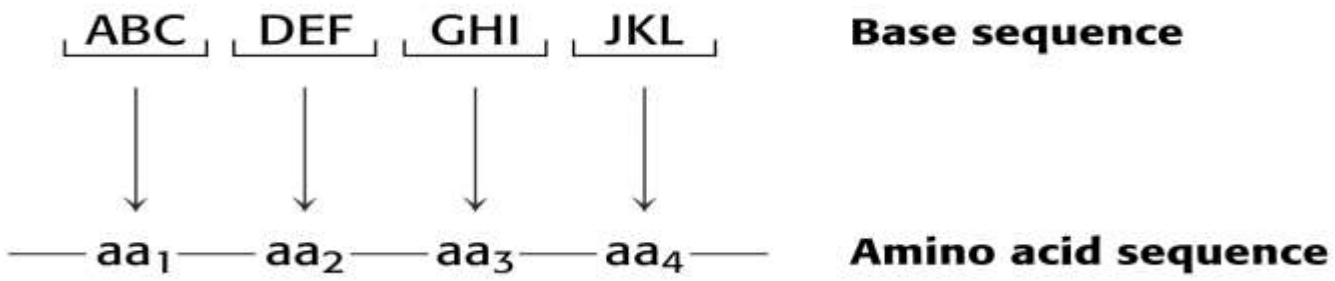


Net Result: Phenylalanine Is Selected by Its Codon

Deciphering the genetic code

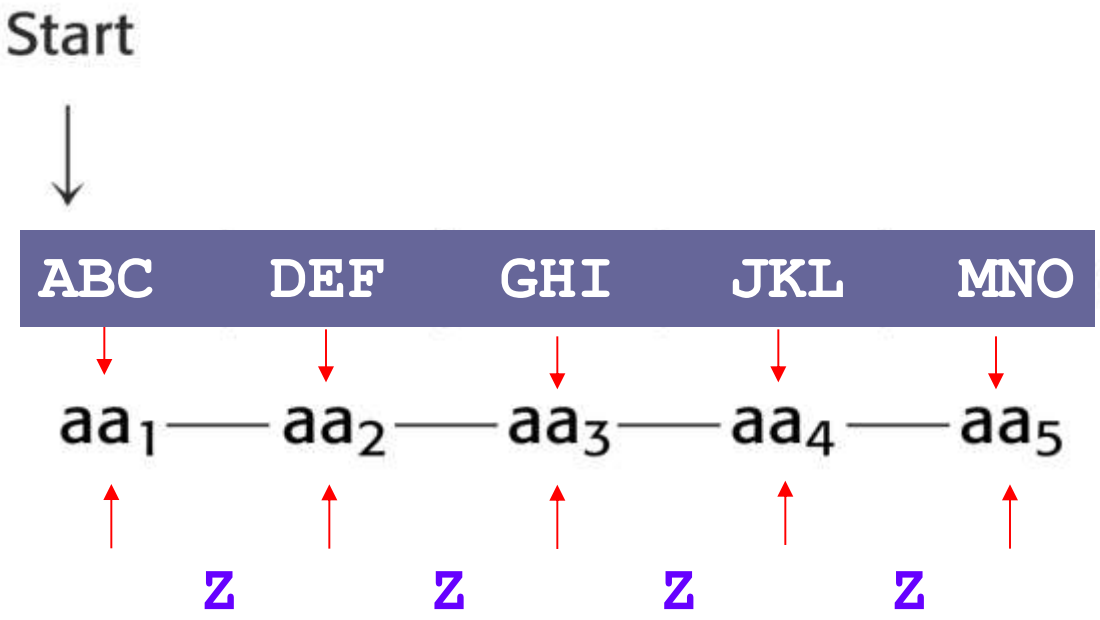
Overlapping versus non-overlapping codons

aa_4



Deciphering the genetic code

A code might have pauses that separate the codons



Deciphering the genetic code

A code may be degenerate or not

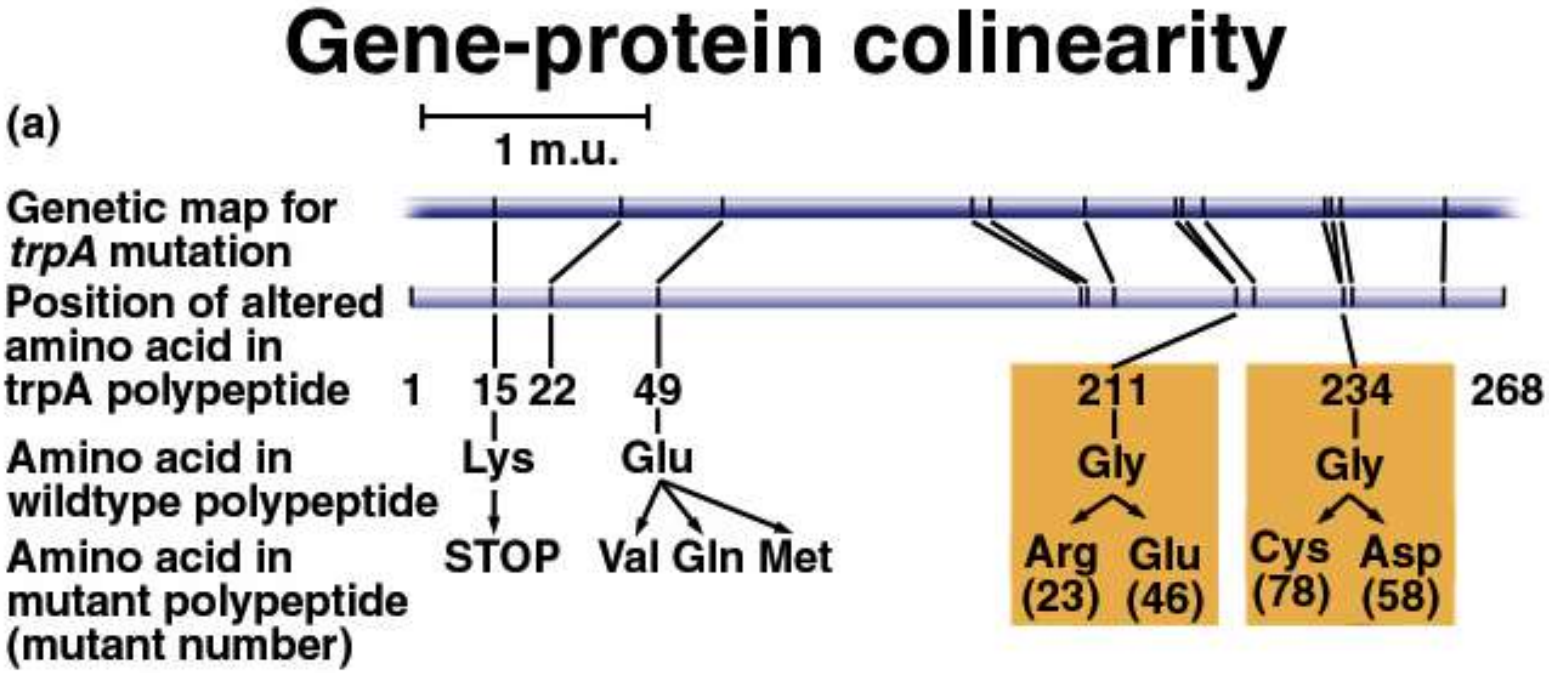
- **In a non-degenerate code, each amino acid is encoded by a single codon.**
- **In a degenerate code, an amino acid can be specified by several different codons**

- **A code is said to be unambiguous if each codon can only encode a specific amino acid. Both degenerate and non-degenerate codes can be unambiguous.**

The information in the gene is colinear with the protein that it encodes

- **Charles Yanofsky made some of the most important early observations concerning the genetic code (1964).**
- **His experiment:**
  - **He first identified a large number of mutations affecting the *trpA* gene (tryptophan synthetase).**
  - **He then constructed an extremely detailed map of these mutations.**
  - **He also purified and sequenced the *trpA* protein produced by many of these mutant alleles.**

Deciphering the genetic code



Deciphering the genetic code

Yanofsky's experiments also showed other important features of the genetic code

- **He observed that each mutation led either to:**
  - **a substitution of one amino acid for another, or**
  - **a truncated protein that was missing a specific region of its carboxy-terminus.**

Deciphering the genetic code

- **These results suggested that:**
  - **the code was non-overlapping.**
  - **the code contains codons that specify the termination of the polypeptide chain (stop codons).**
  - **proteins are probably produced beginning with their amino-terminus.**

How many bases make a codon?

- **Experiments by Francis Crick and Sydney Brenner**
  - **Their experiment:**
    - **They isolated a large number of mutations in the rIIB gene of bacteriophage T4.**
    - **They found that these mutant alleles could not be reverted to wild type by mutagens known to cause point mutations.**
    - **However, proflavin induced mutations could be reverted by inducing proflavin again to introduce a second mutation in the gene.**

- **Brenner and Crick realized that proflavin might intercolate into the DNA and cause single basepair insertions and deletions .**

- **Brenner and Crick used these “frame shift” mutations to determine that the genetic code was read in triplets.**
  - **An important first step was to classify the mutations into ones that either inserted an extra base (+ mutations) or deleted a single base (- mutations).**

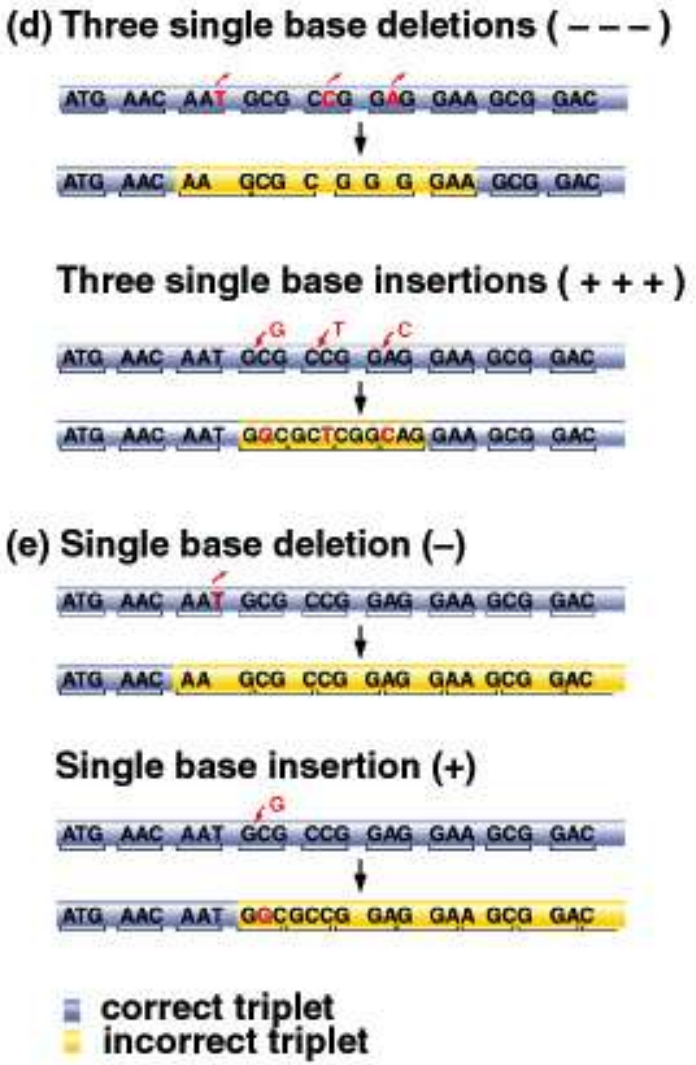
- **Brenner and Crick then recombined together various combinations of + and - mutations and asked whether gene function could be restored.**

(c) Different sets of mutations generate either a mutant or a normal phenotype

Proflavin-induced mutations (+) insertion (–) deletion	Phenotype
– or +	Mutant
– – or + +	Mutant
– – – – or – – – – –	Mutant
– +	Wildtype
– – – or – – – – – – or + + + or + + + + + +	Wildtype

These results indicated that the code must be read as triplets.

The rationale behind Brenner and Crick's conclusion

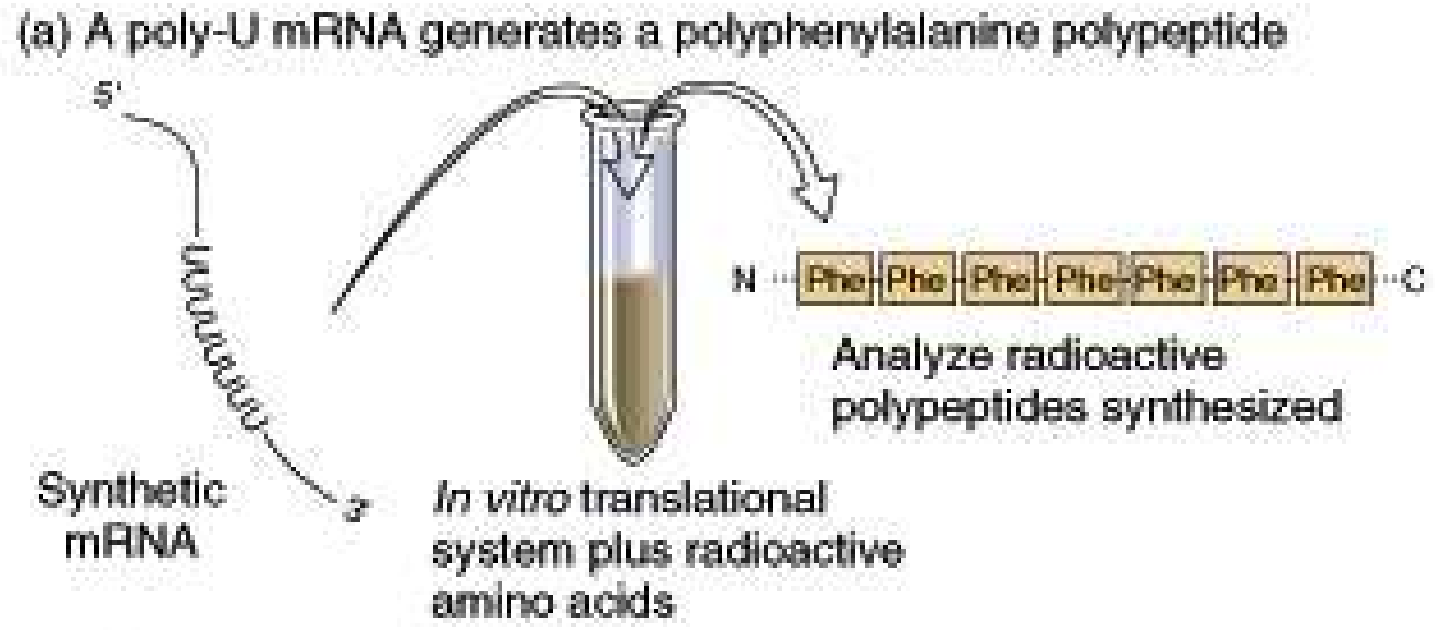


Cracking the triplet code

- The experiments of Yanofsky, Brenner and Crick indicated that the genetic code consisted of non-overlapping triplet codons.
- However, these experiments gave no indication of the actual code.

- The unraveling of the code required the development of cell extracts (called *in vitro* translation systems) capable of synthesizing proteins under the direction of added mRNAs.

The identification of the first codon-”UUU”



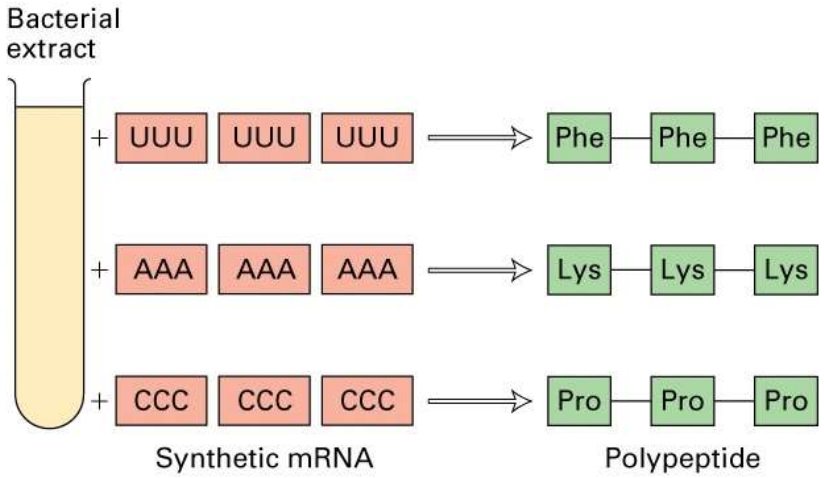
Marshall Nierenberg and Heinrich Matthaei

The identification of the first codon-”UUU”

- **Marshall Nierenberg and Heinrich Matthaei added a chemically synthesized polymer of uridine (5’-UUUUUUUUU...) to an *in vitro* translation system (1961)**
  - **long polymers of phenylalanine were produced**
- **Conclusion:**
  - **The triplet UUU encodes phenylalanine.**

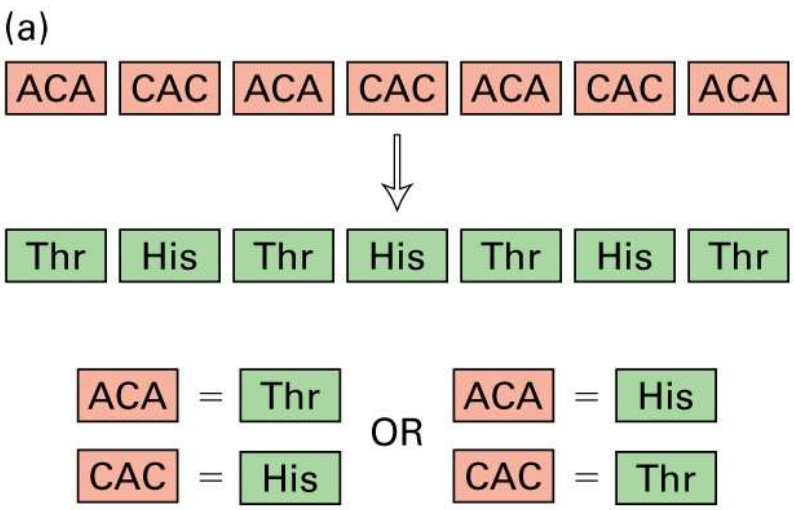
Deciphering the genetic code

- Similar experiments were used to identify the codons:
  - AAA encodes lysine
  - GGG encodes glycine
  - CCC encodes proline



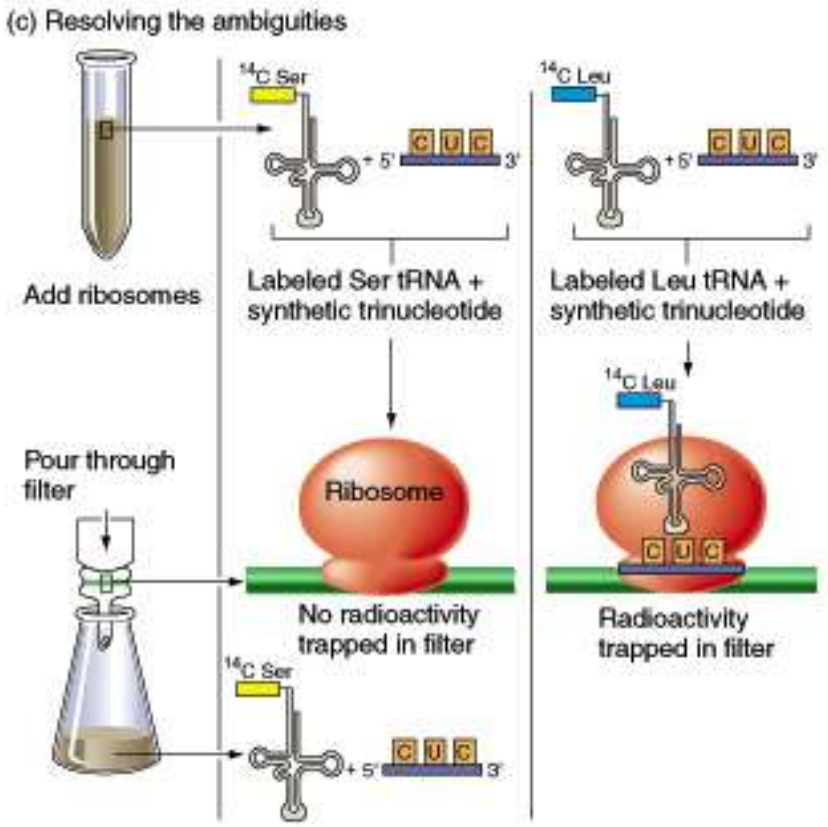
Programming in vitro translation extracts with more complex repetitive RNAs allowed additional codon assignments

```
5′-UCUCUCUC…    Ser-Leu-Ser-Leu...
5′-AGAGAGAG…    Arg-Glu-Arg-Glu...
5′-UGUGUGUG…    Cys-Val-Cys-Leu...
5′-ACACACAC…    Thr-His-Thr-His...
```

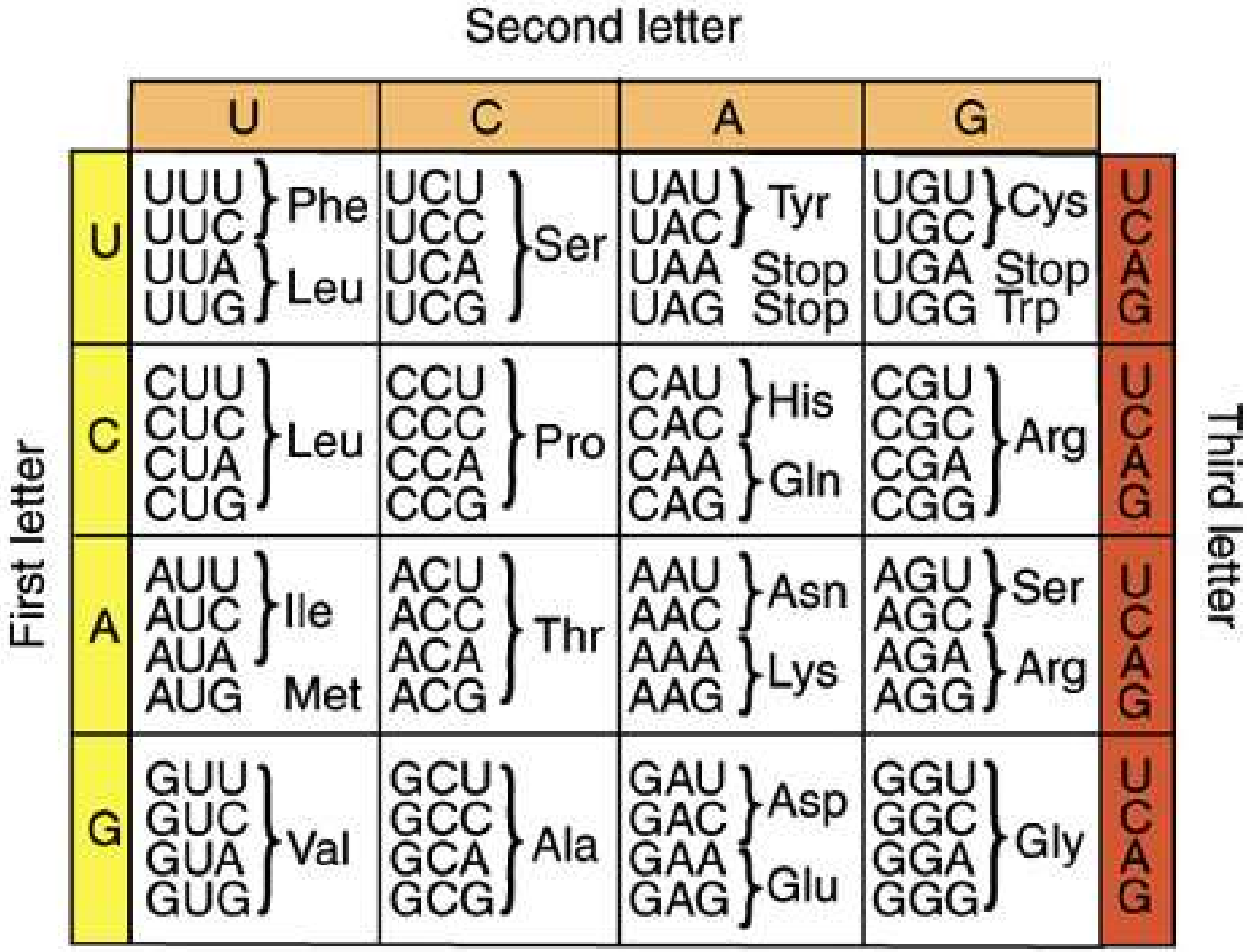


Nierenberg and Matthaei's experiment

The remainder of the code was determined by studying the interaction of charged tRNA molecules with ribosomes



The genetic code

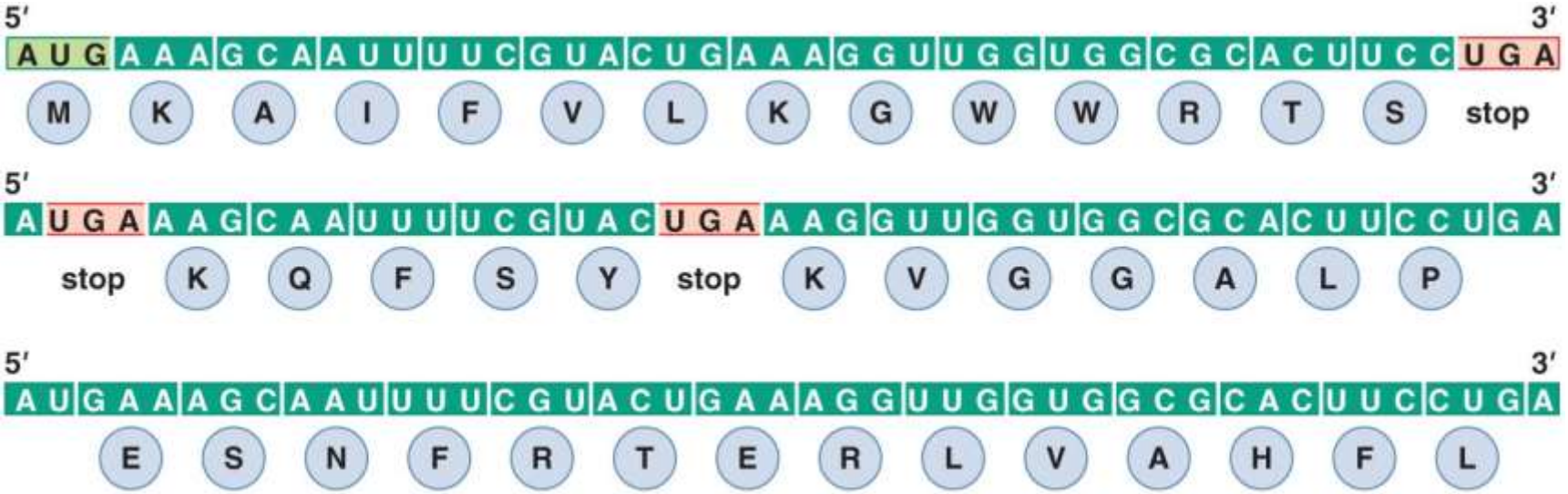


Properties of the genetic code

- **The code consists of triplet codons**
- **The codons are nonoverlapping .**
- **The code includes nonsense or stop codons.**
- **The code is degenerate. Several codons may encode the same amino acid.**
- **The reading frame used to translate a messenger RNA is established by the position of the first codon that is translated.**
  - **The protein synthesis machinery (the ribosome always reads the next three bases..it never skips)**
- **The 5-3' direction of the mRNA corresponds to the amino-carboxy-terminal direction of the protein.**

Reading of the Triplet Code

There are three potential reading frames in all mRNAs. However, only one reading frame is used for translation, and is selected based on the frame in which the AUG start codon appears. Triplet codons are read in a non-overlapping, comma-less manner. Rarely are mRNAs read in more than one frame.



Genetic code

- only strand of DNA is transcribed into mRNA

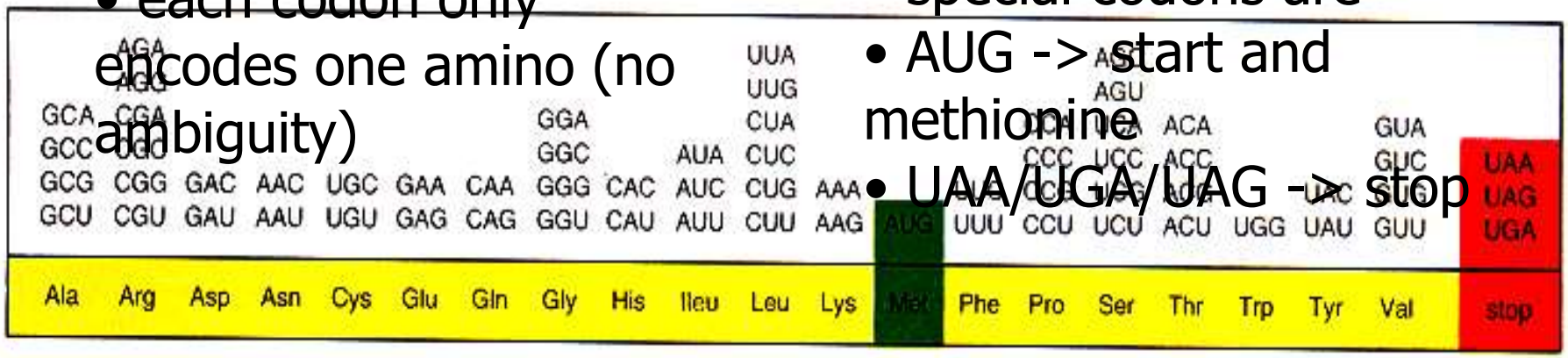
- the mRNA is read in three's (codons) and translated by the ribosome into protein

- each codon only encodes one amino (no ambiguity)

- one amino may be encoded by more than one codon (redundancy)

- variability is in the 3rd codon position (wobble base)

- special codons are
- AUG -> start and methionine
- UAA/UGA/UAG -> stop



Codon Usage

- More than one codon exists for most amino acids (except Met and Trp)
- Organism may have a preferred codon for a particular amino acid
- Codon usage correlates with abundance of tRNAs (preferred codons are represented by abundant tRNAs)
- Rare tRNAs correspond to rarely used codons
- mRNAs containing rare codons experience slow translation

The genetic code is nearly universal

- The codons used in mitochondria and some ciliates differ slightly from the "normal" genetic code

TABLE 5.5 **Distinctive codons of human mitochondria**

Codon	Standard code	Mitochondrial code
UGA	Stop	Trp
UGG	Trp	Trp
AUA	Ile	Met
AUG	Met	Met
AGA	Arg	Stop
AGG	Arg	Stop

Genetics tRNA and rRNA

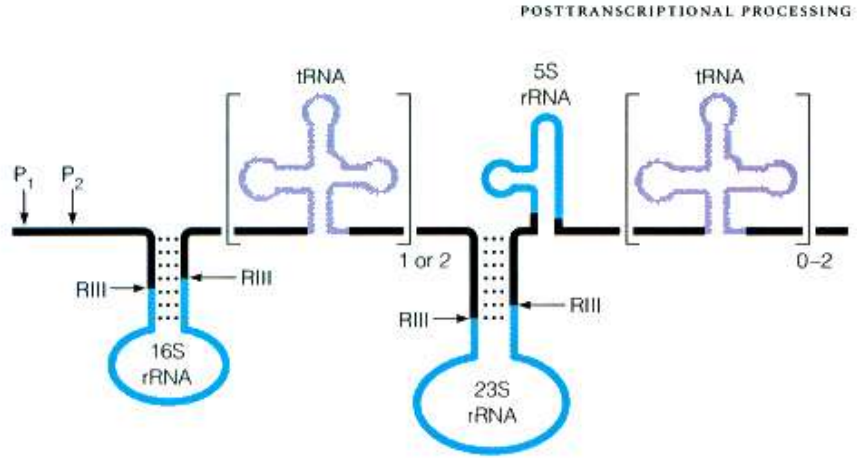
7 different operons in *Escherichia coli*

rrnA, rrnB, rrnC, rrnD, rrnE, rrnF, rrnG* and *rrnG

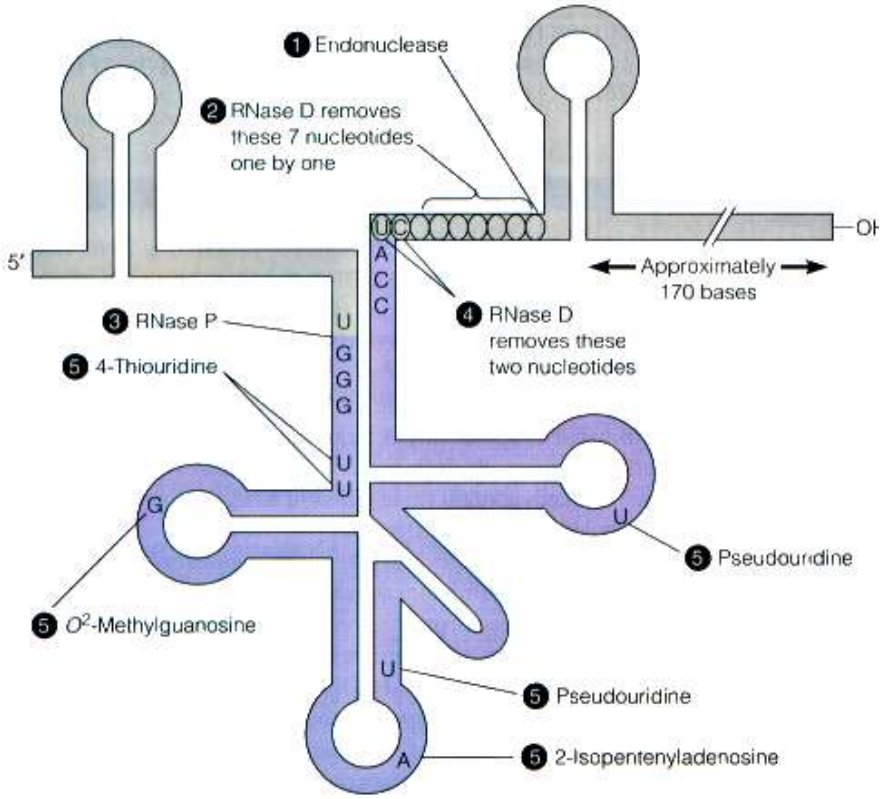
Each operon contains
(i) one copy of the 23S, 16S and 5S rRNA genes
(ii) 1-4 copies tRNA genes

Post-transcriptional Regulation

RNA Processing by RNAases
RNase III, M5, M6 and M23



Processing of *E. coli* 30S pre-rRNA

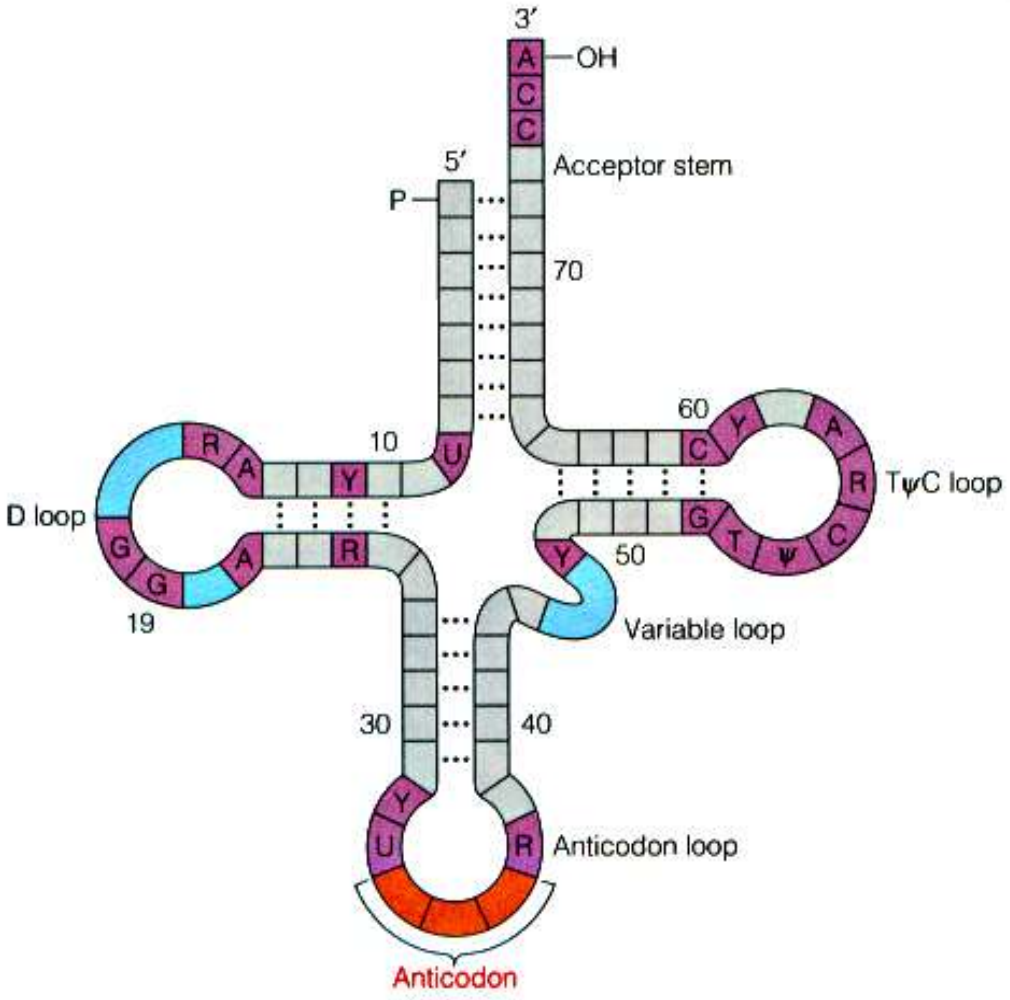


Maturation of *E. coli* $tRNA^{Tyr}$ from its transcript

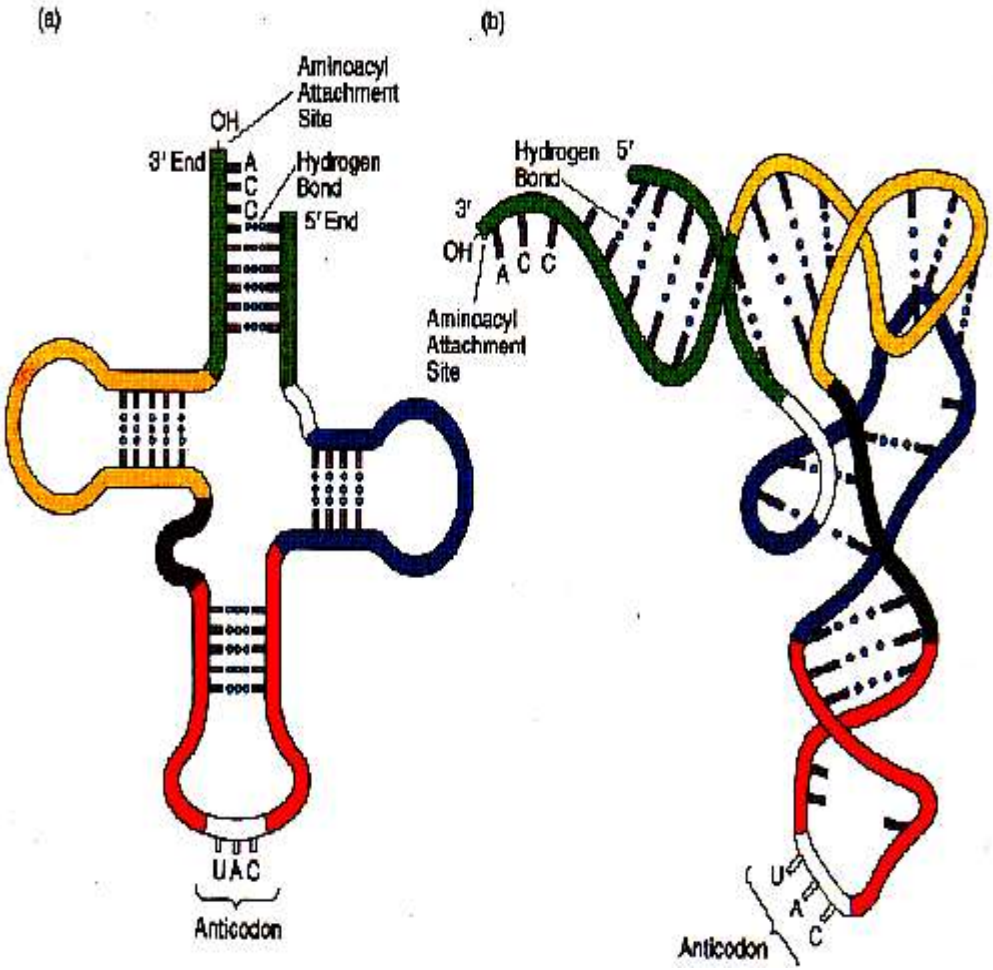
Generalized tRNA structure

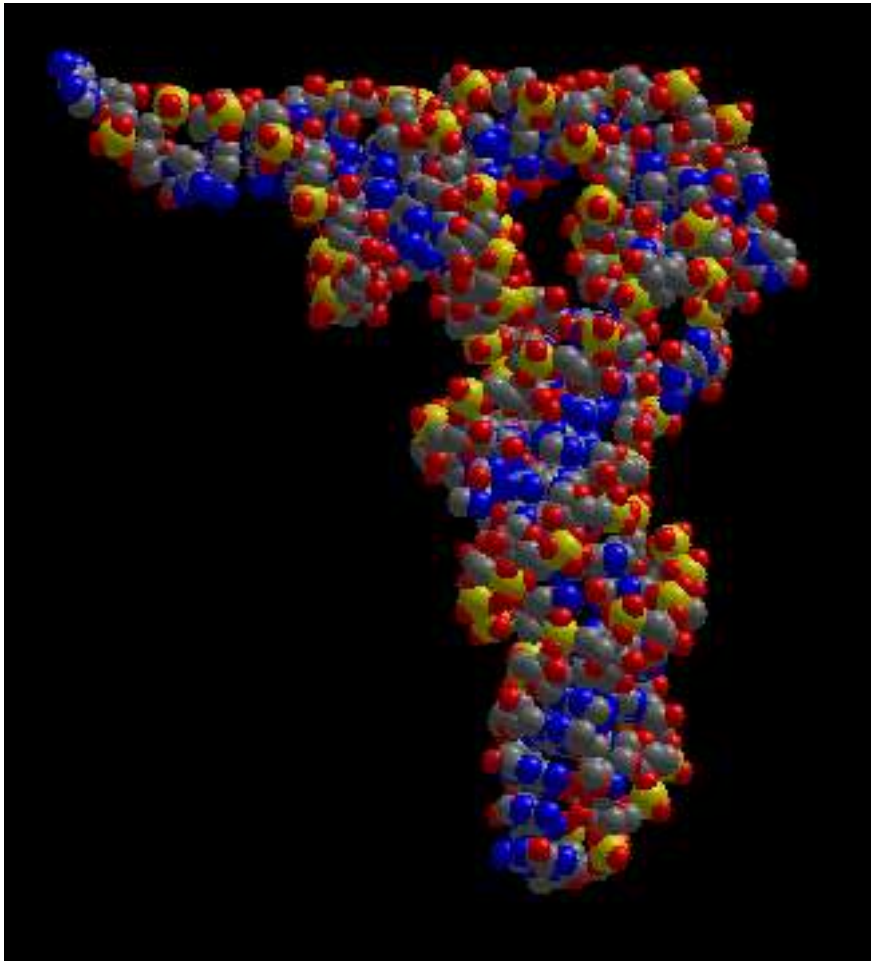
The positions of invariant and rarely varied bases are shown in purple.

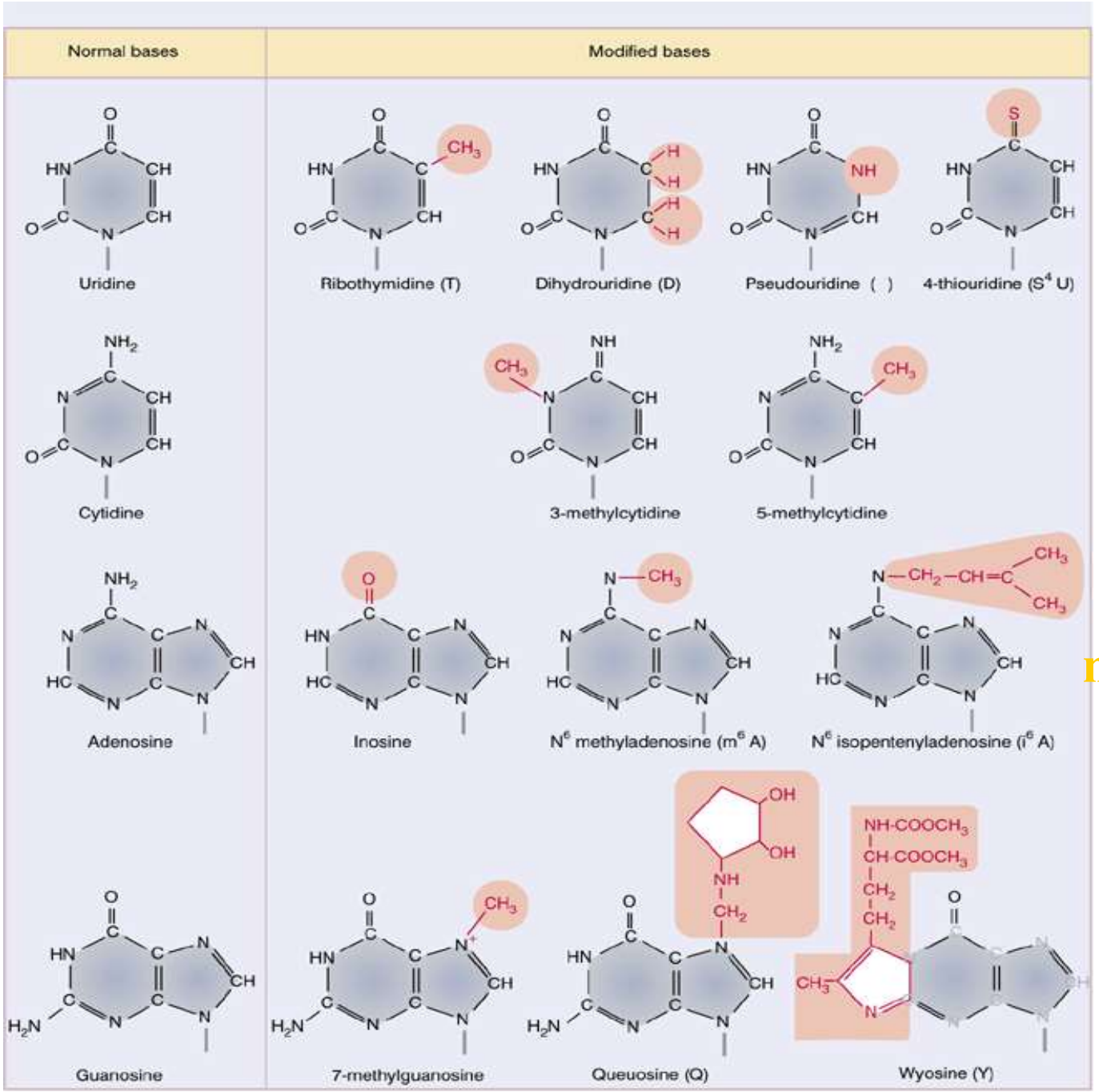
Regions in the D loop and the variable loop that can contain different numbers of nucleotides are shown in blue



t – RNA – adaptor molecule







Examples of t-RNA modifications

Specific tRNA modifying enzymes

The tRNA anticodon loop

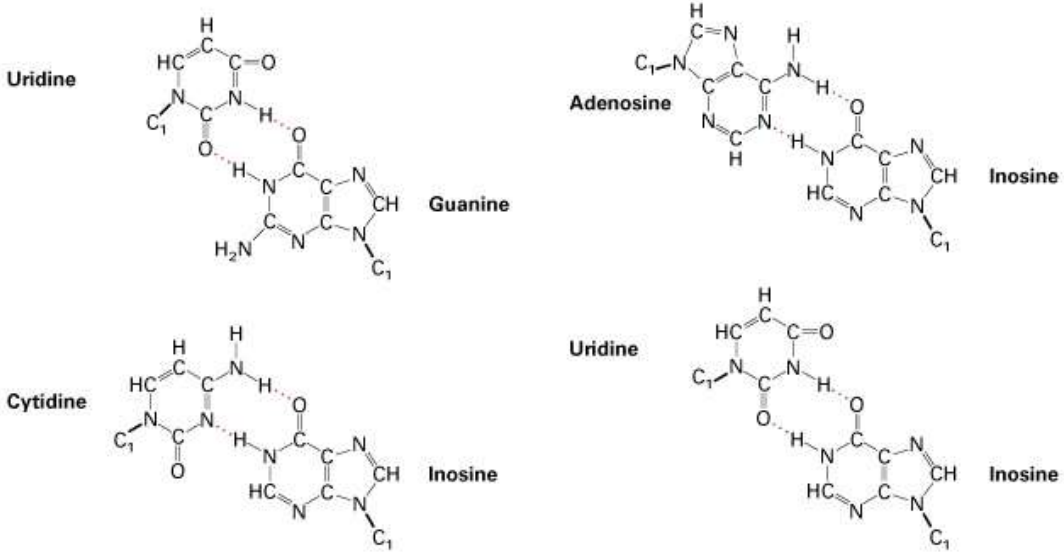
tRNAs often recognize more than one codon due to “wobble” at the 3rd base of a codon – nonstandard base pairing.

If these bases are in first, or wobble, position of anticodon

C	A	G	U	I	
G	U	C U	A G	C A U	then the tRNA may recognize codons in mRNA having these bases in third position

If these bases are in third, or wobble, position of codon of an mRNA

C	A	G	U	
G I	U I	C U	A G I	then the codon may be recognized by a tRNA having these bases in first position of anticodon



Degeneracy: an organized feature of the genetic code

Base pairing between mRNA and tRNA is variable at the 3rd position, and subjected to “wobble”

Codon-Anticodon Pairings Allowed by the Wobble Rules

5' end of anticodon	3' end of codon
G	U or C
C	G only
A	U only
U	A or G
I	U, C, or A

AGC
AGU
UCA
UCC
UCG
UCU

Ser

Different tRNAs that Can Service Codons for Serine

tRNA	Anticodon	Codon
$tRNA^{Ser}_1$	AGG + wobble	UCU UCC
$tRNA^{Ser}_2$	AGU + wobble	UCA UCG
$tRNA^{Ser}_3$	UCG + wobble	AGU AGC

Degeneracy: an organized feature of the genetic code

Numbers of different tRNA used by some organisms

E. coli (Gram- bacteria)	40
Gram+ bacteria	33
Mycoplasmas	28
Mitochondria	22
Chloroplasts	30
Mammals	45

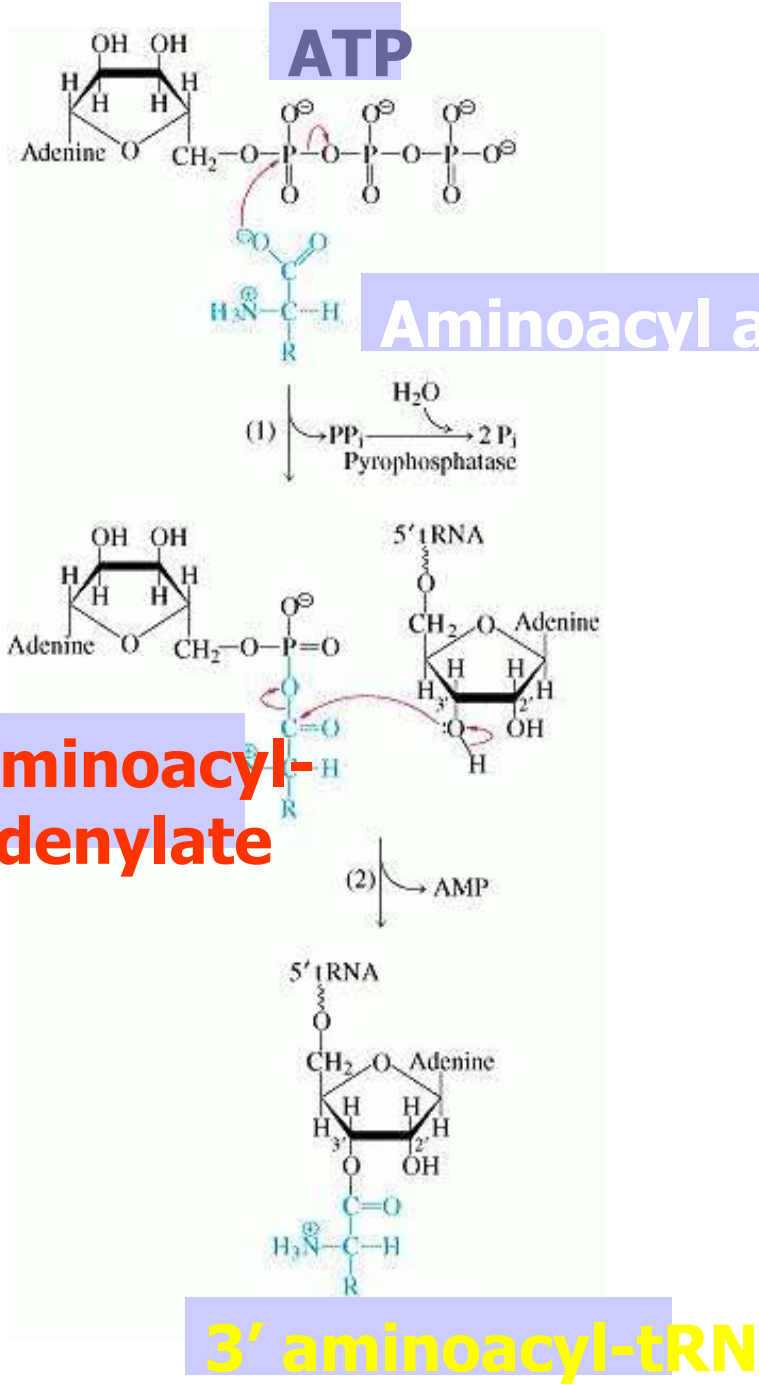
Charging tRNA

Each enzyme must recognize *both* the tRNA specific for an amino acid and the corresponding amino acid.

ATP-dependent and results in the cleavage of two high-energy phosphate bonds:

Activated aa is trasferred to the tRNA with the release of AMP.

Most tRNA synthetases must be able to recognize more than one type of tRNA (i.e. 6 codons for Arg).



Aminoacyl-tRNA synthetases

They all carry out very similar tasks, but they vary greatly in size (40-100 kDalton).

2 classes - structure of the active site

Class I		Class II	
Glu	α	Gly	α2β2
Gln	α	Ala	α4
Arg	α	Pro	α2
Cys	α2	Ser	α2
Met	α2	Thr	α2
Val	α	His	α2
Ile	α	Asp	α2
Leu	α	Asn	α2
Tyr	α2	Lys	α2
Trp	α2	Phe	α2β2

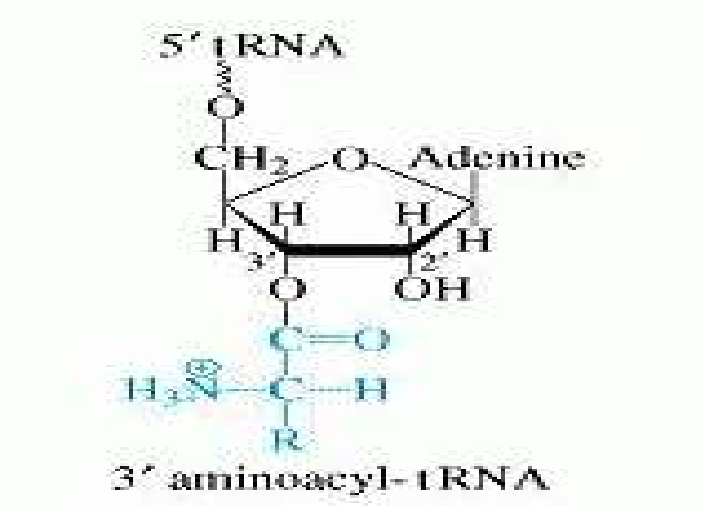
Class I enzymes

Generally (though not always) monomeric
Attach the carboxyl of their target amino acid to the 2' OH of adenosine 76 in the tRNA molecule.

Class II enzymes

Generally dimeric or tetrameric
Attach their amino acid to the 3' OH of their tRNA

Exception:
Phe-tRNA synthetase which uses the 2' OH.



Aminoacyl-tRNA synthetases

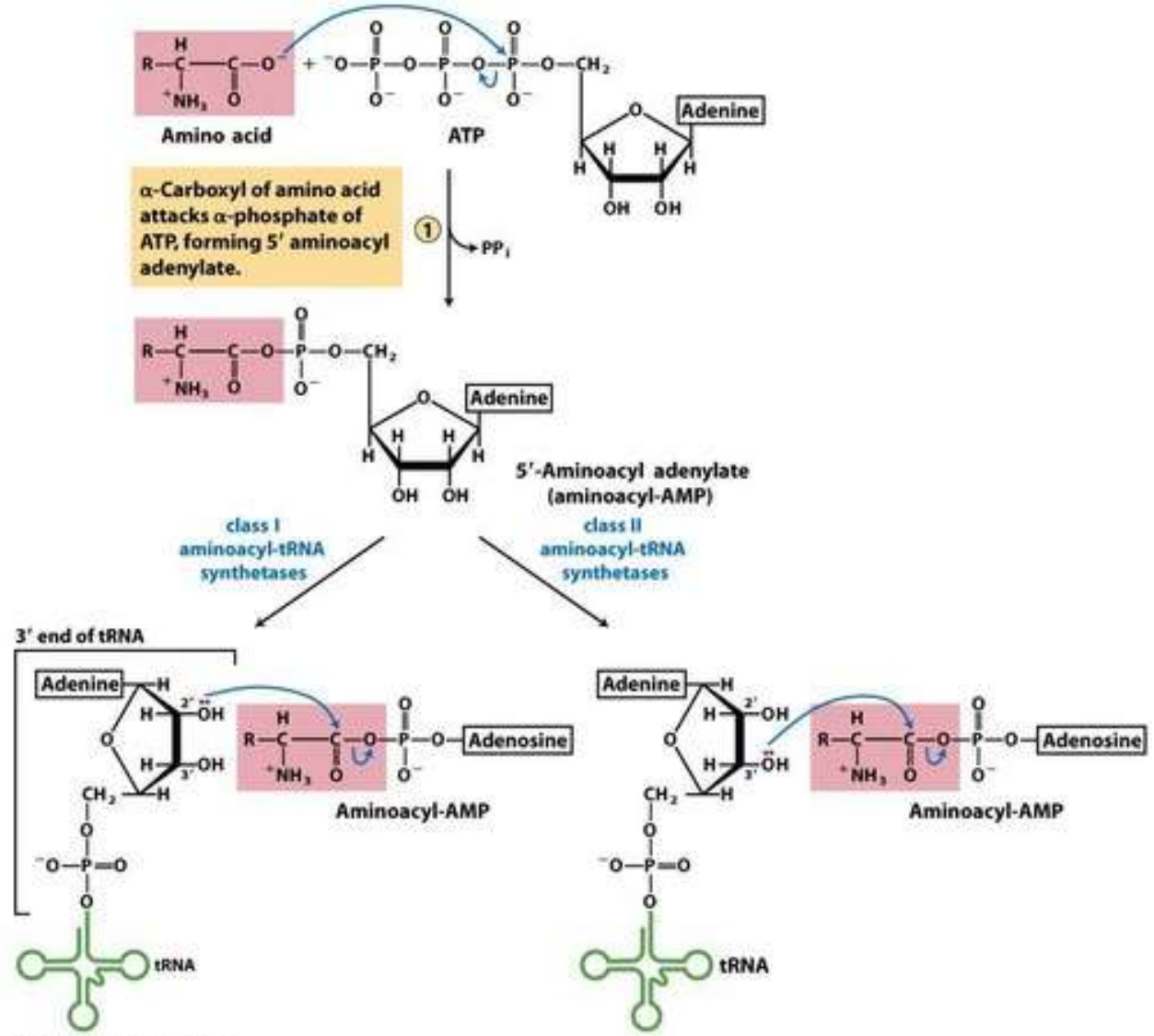


Figure 27-19 part 1
Lehninger Principles of Biochemistry, Fifth Edition

Aminoacyl-tRNA synthetases

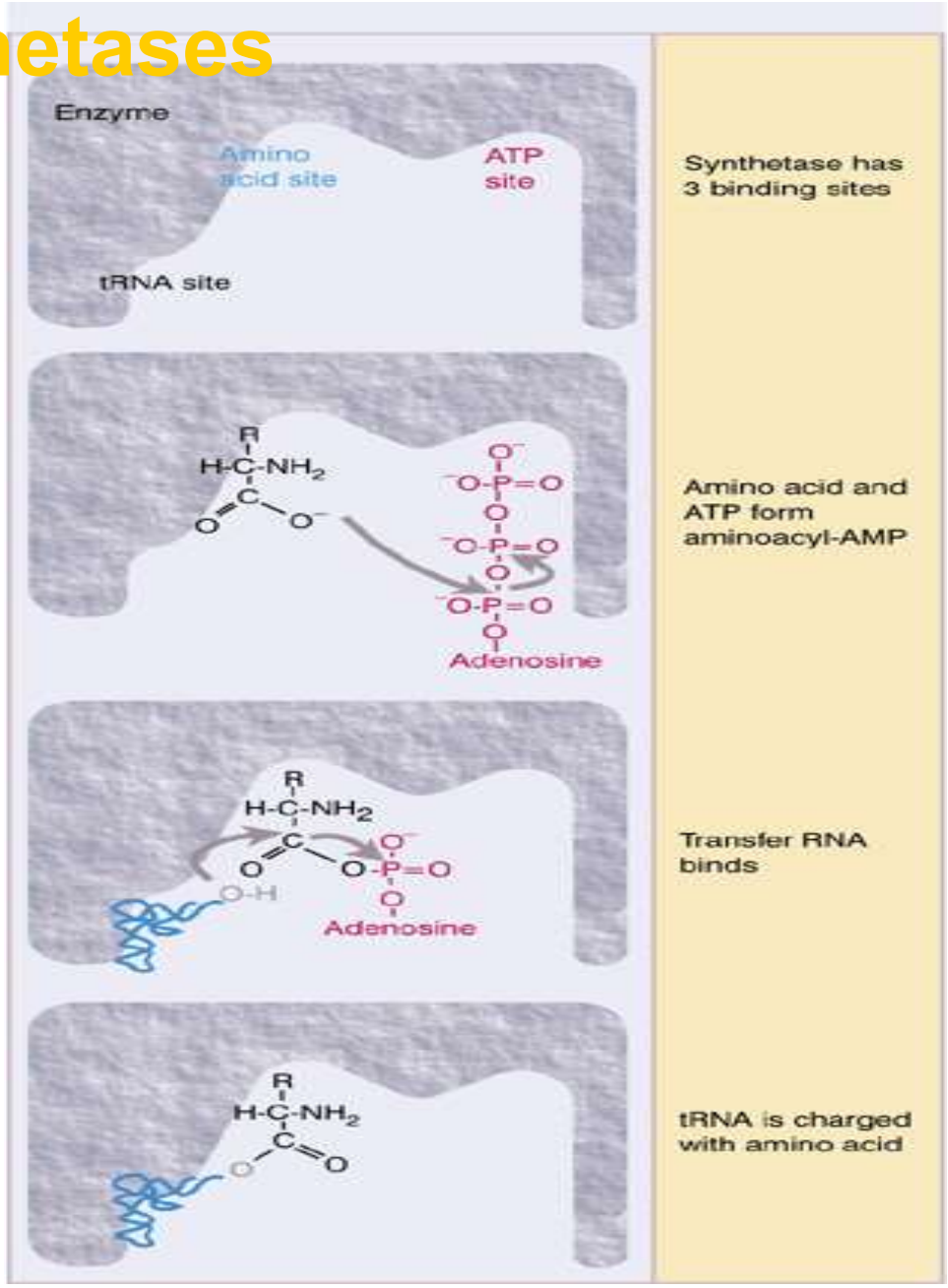
Activation occurs in two steps and is catalyzed by aminoacyl tRNA synthetase:

1. The amino acid is covalently linked to AMP (acid anhydride linkage)

2. The activated amino acid is then transferred to its cognate tRNA, and AMP is released.

Note that the reaction consumes one ATP.

There is one aminoacyl tRNA synthetase for each amino acid. Because of the degeneracy of the genetic code, several tRNAs, all coding for the same amino acid, can be recognized by one aminoacyl tRNA synthetase enzyme.



Aminoacyl-tRNA synthetases

Each aminoacyl-tRNA synthetase attaches a single amino acid to one or more tRNAs

Most organisms have 20 different tRNA synthetases.

An exception in some bacteria: amination of Glu-$tRNA^{Gln}$ to Gln-$tRNA^{Gln}$

Aminoacyl-tRNA formation is very accurate

Seletion of the correct amino acid is more difficult.
Nevertheless, less than 1 in 1,000 tRNAs is charged with the incorrect amino acid.

Aminoacyl-tRNA synthetases

The ribosome is unable to discriminate between correctly and incorrectly charged tRNAs

A mutant tRNA with a nucleotide s

the anticodon delivers its usual cc

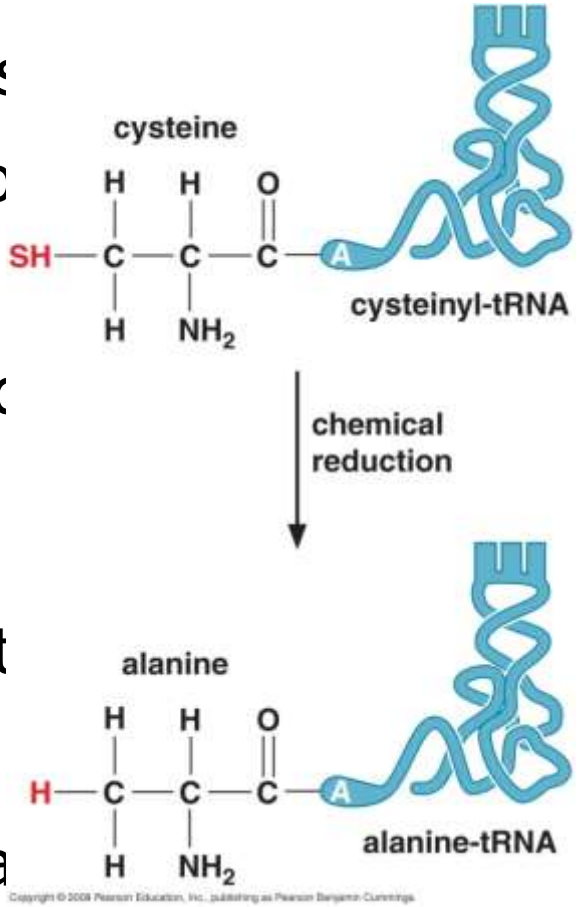
amino acid to the wrong codon.

Cysteinyl-tRNA^{Cys} can be converte
tRNA^{Cys} by chemical reduction. Thi
results in

introduction of alanines at the cyst

codons.

High fidelity of tRNA synthetases a
accurate decoding of mRNAs.



Aminoacyl-tRNA synthetases

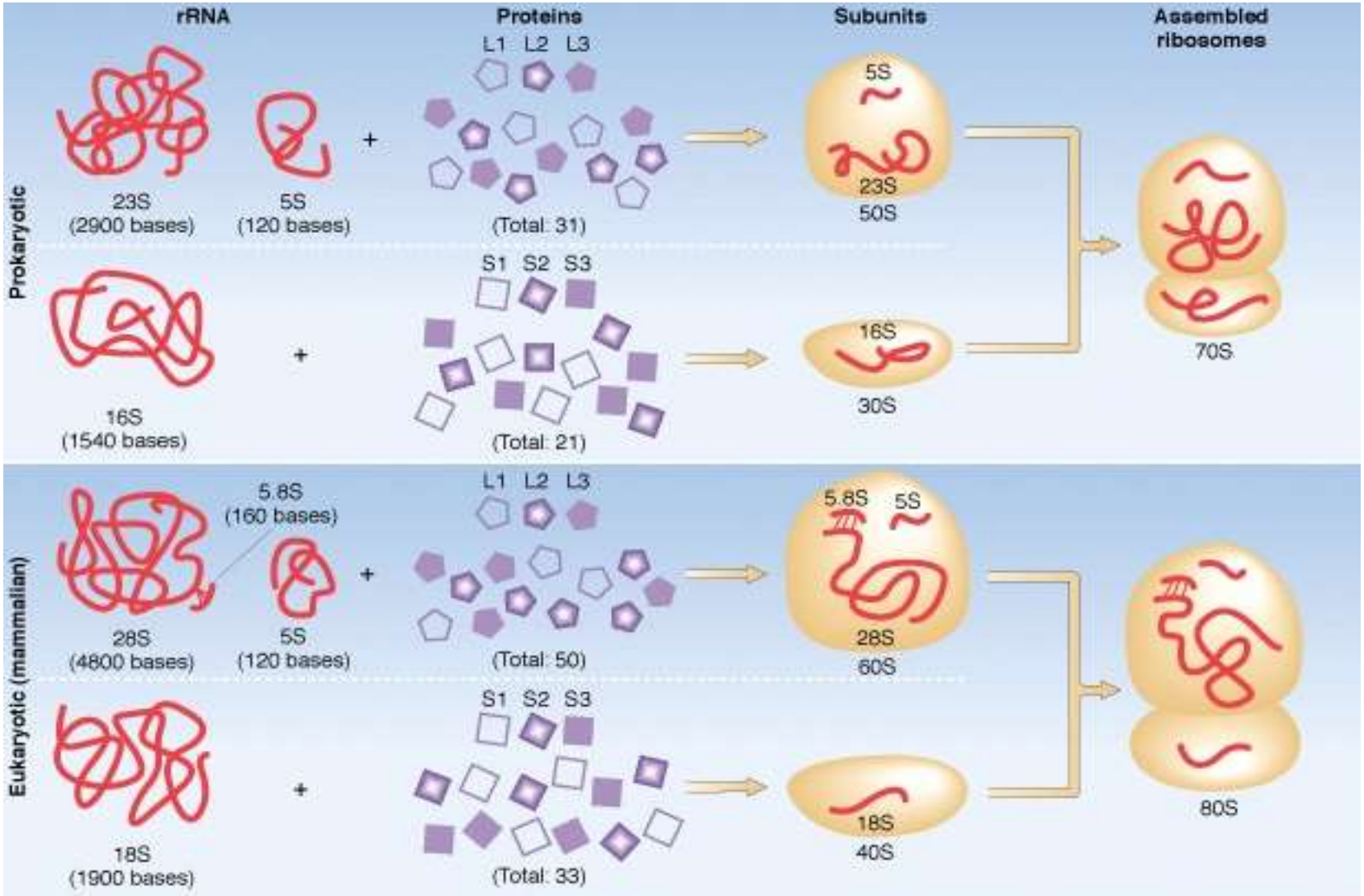
Specificity determinants: the acceptor stem and the anticodon loop

Discriminator in the acceptor stem: Changing a particular base pair in the acceptor stem converts the recognition specificity of a tRNA from one synthetase to another.

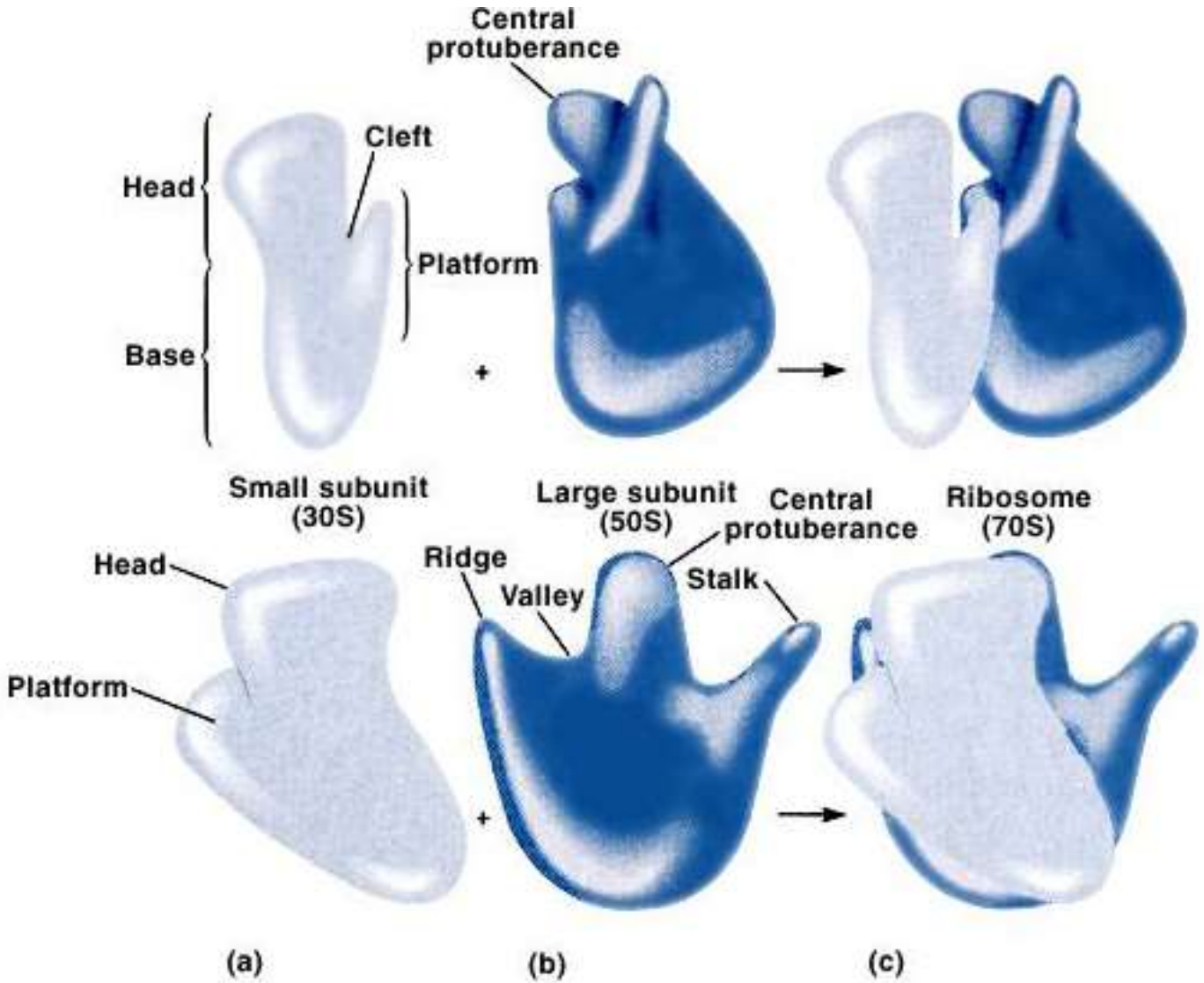
The anticodon loop, including recognition of the anticodon itself, contributes to discrimination. However, in case of serine, AGC and UCA are completely different from each other. The specificity determinants should be outside of the anticodon.

The set of tRNA determinants that enable

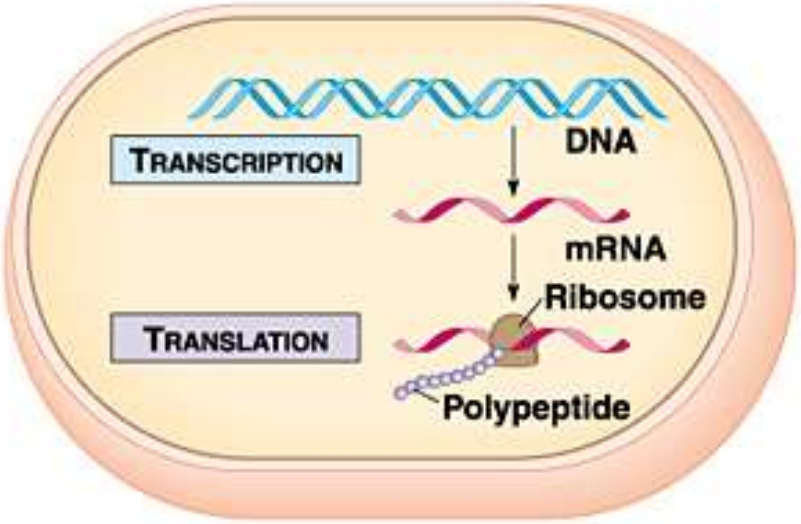
Prokaryotic and eukaryotic ribosomes differ in size and subunit structure.



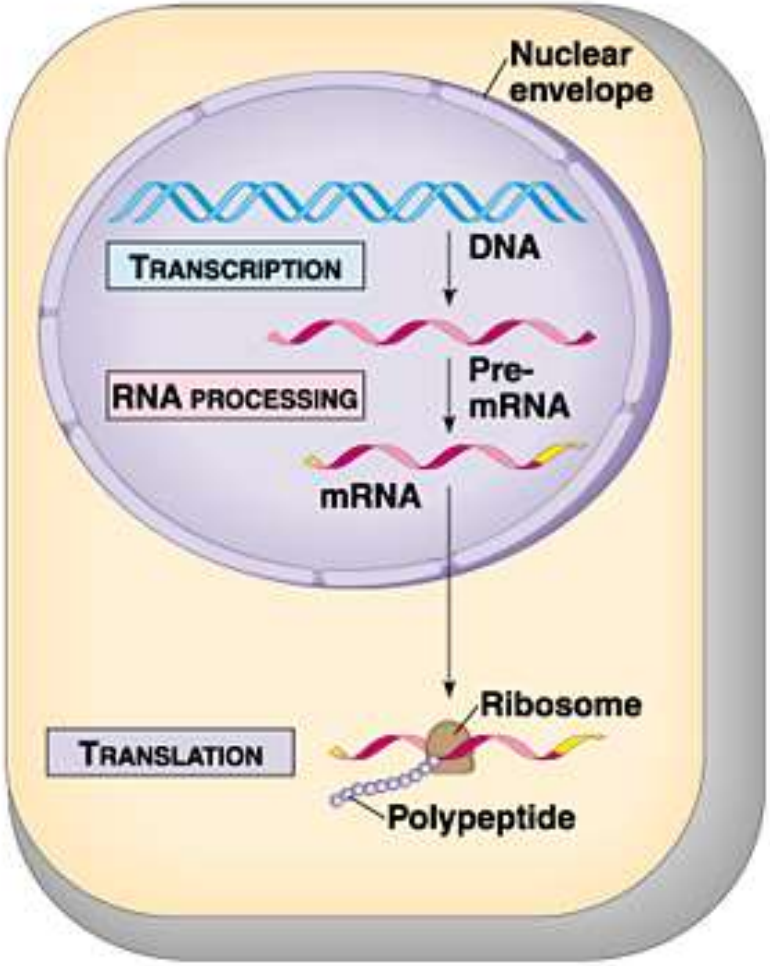
Escherichia coli ribosome



Prokaryotes and Eukaryotes



(a) Prokaryotic cell

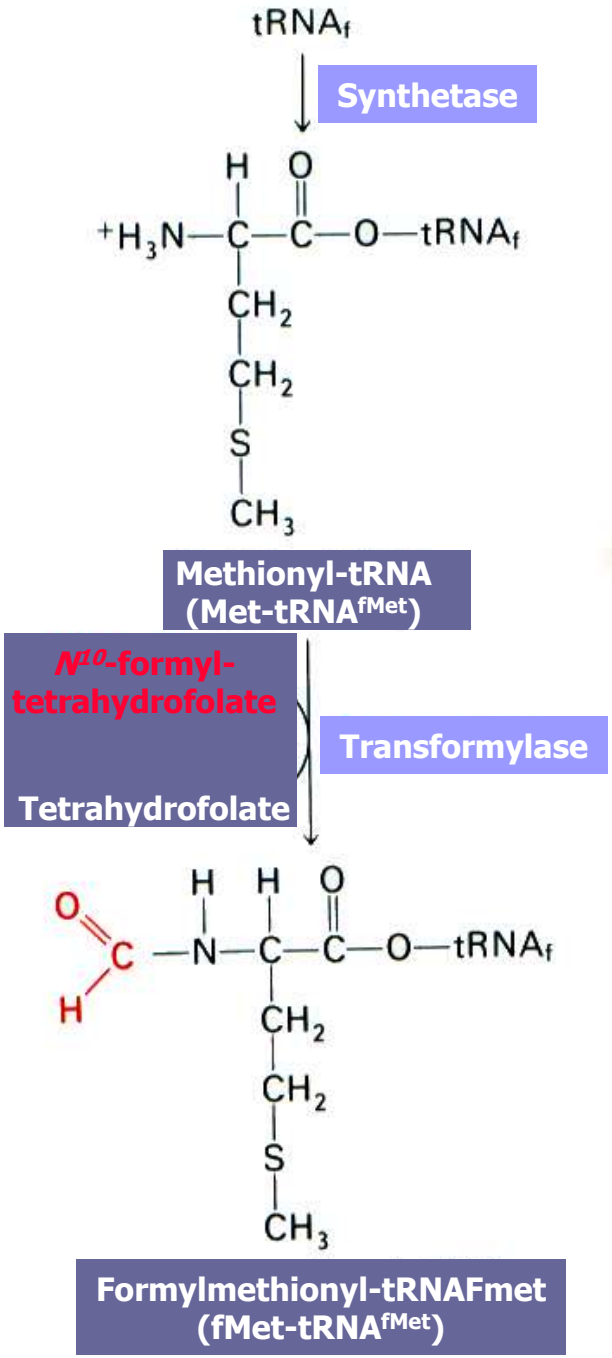


(b) Eukaryotic cell

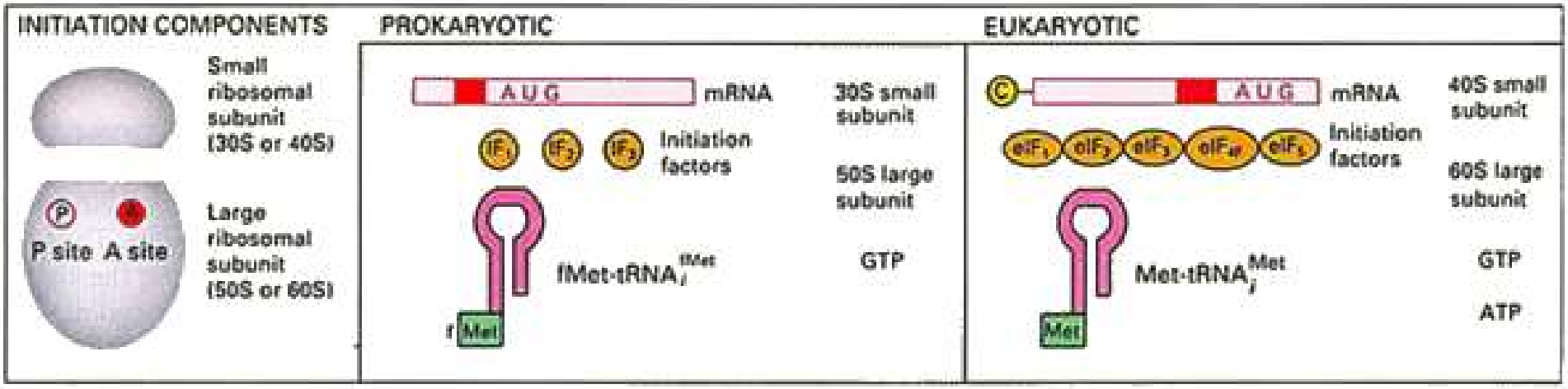
- prokaryotes (bacteria) do not have nuclei
- eukaryotes segrege transcription in the nucleus. mRNA is also preprocessed prior to translation in eukaryotes

Important facts on protein synthesis in both prokaryotes and eukaryotes

1. Proteins are synthesized from amino- to carboxyl-terminus.
2. The mRNA is read in the 5' to 3
3. All proteins start with methioni first amino acid.
4. There are two methionine-spec $tRNA^{fMet}$ and $tRNA^{Met}$.
5. The amino group in Met-$tRNA^{fM}$ formylated in prokaryotes. In eukaryotes, no formylation t
6. Specific deformylases later ren formyl group from the protein.
7. Methionine-specific amino-pep cleave the terminal Met from so



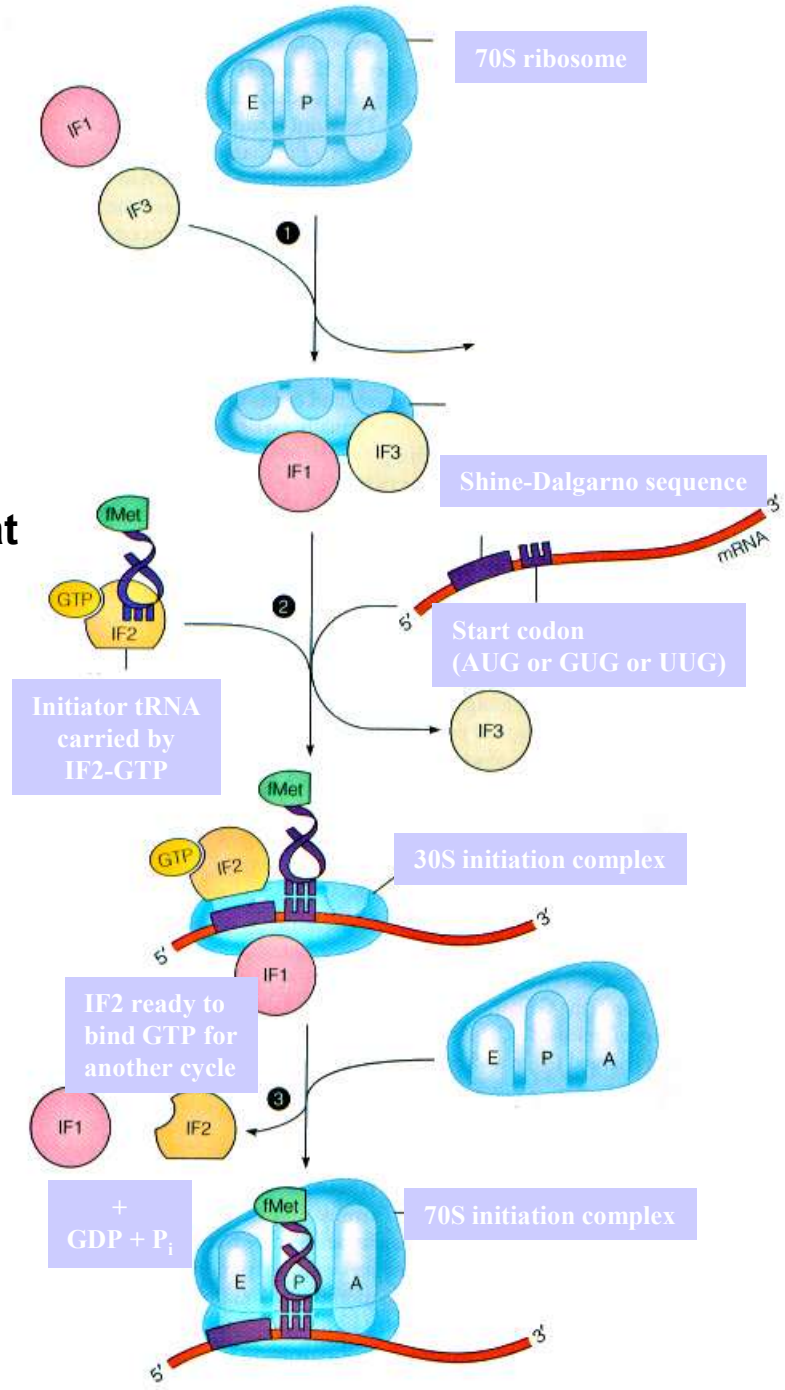
Differences between translation initiation in prokaryotes and eukaryotes



Initiation of protein synthesis in prokaryotes

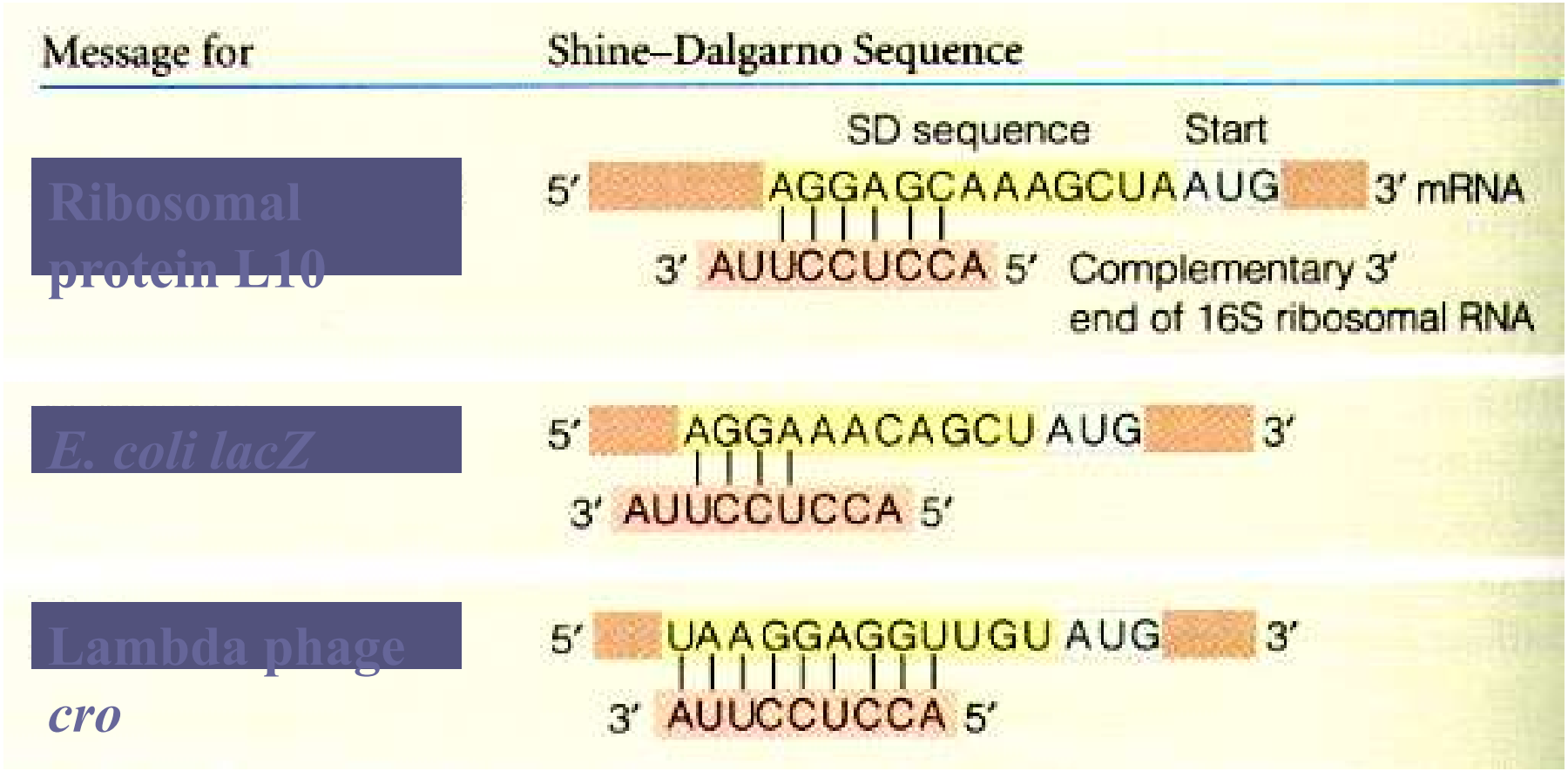
1. IF1 and IF3 promote dissociation of a 70S ribosome.
2. IF2 binds to initiator tRNA together with GTP.
3. Together with the 30S ribosomal subunit this complex then attaches to a start codon that is preceded by a Shine-Dalgarno sequence. IF1 remains bound but IF3 is released.
4. The resulting 30S initiation complex is then joined by the larger 50S subunit.
5. This causes GTP to be hydrolyzed to GDP and P_i, which initiates a conformational change in IF2. As a result of this change, IF2 is released together with IF1.
6. The 70S initiation complex is now ready for the next step of translation, elongation.

The 50S subunit has three sites for tRNA binding, called the P (petidyl) site, the A (aminoacyl) site, and the E (exit) site. The initiator tRNA is bound to the P site

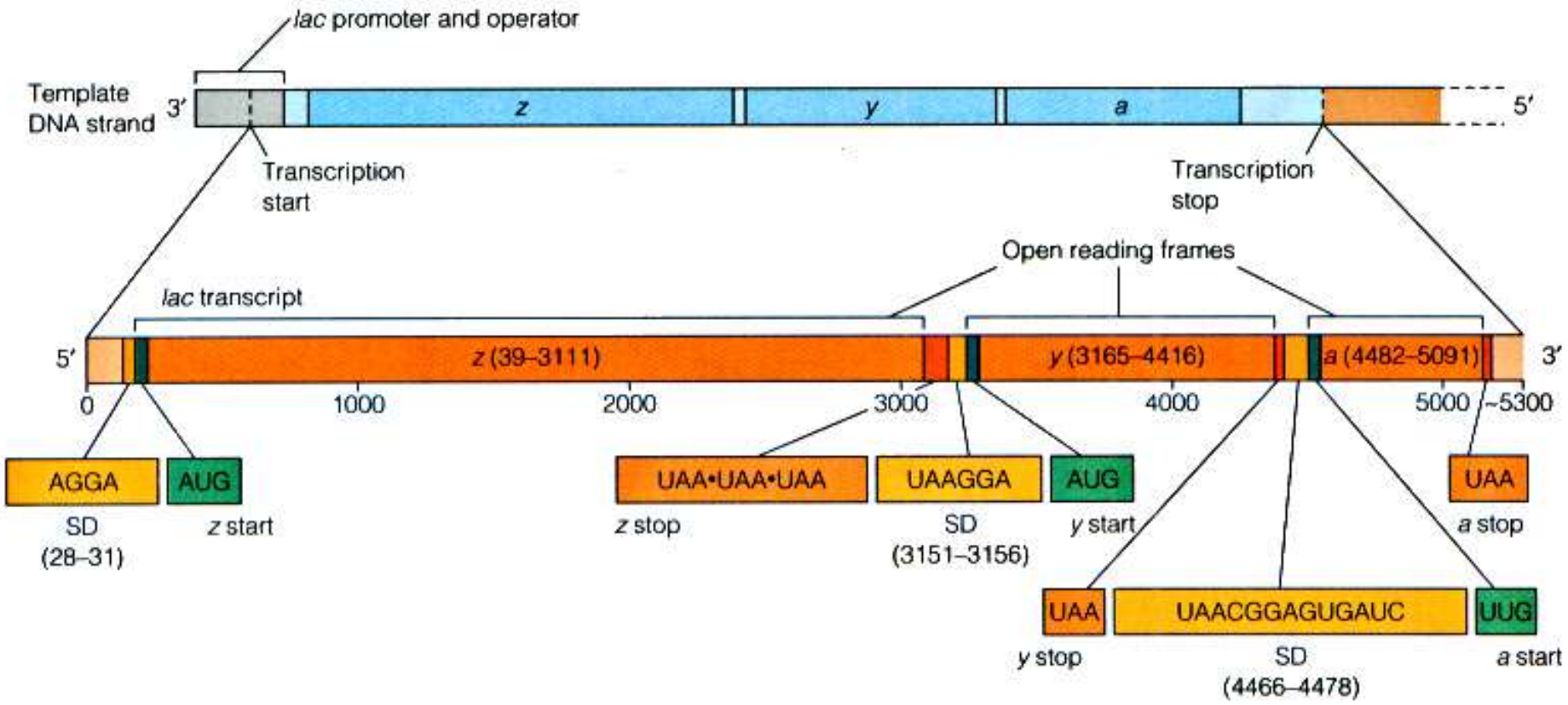


Ribosomal binding site or Shine-Dalgarno sequence

Polypurine sequence AGGAGG located on mRNA just prior to an initiation codon
Complementary to the sequence at the 3' end of 16S rRNA
Involved in binding of ribosome to mRNA.



Multiple Shine-Dalgarno sequences in the polycistronic *lac* mRNA



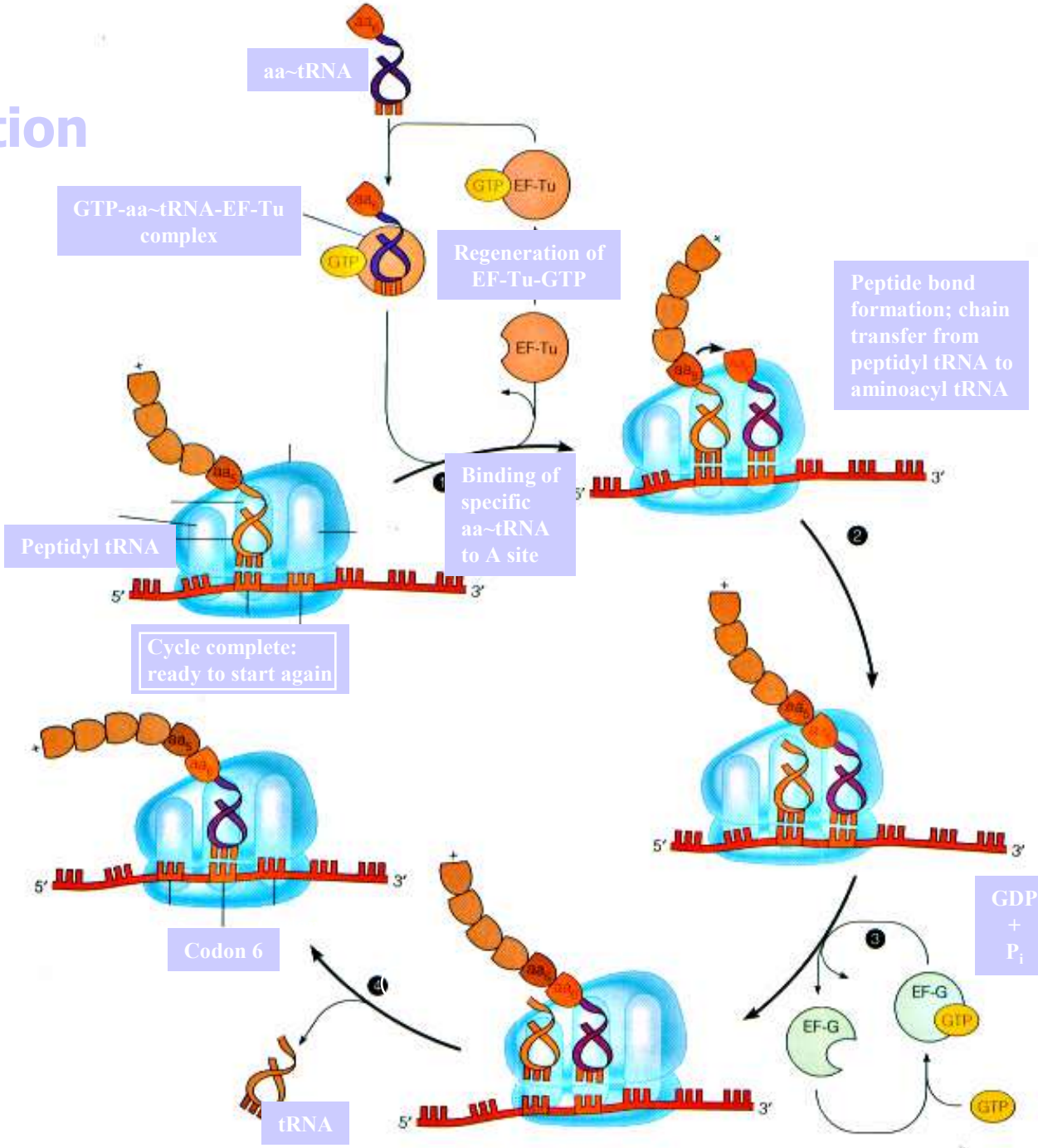
Prokaryotic initiation factors

Factor	GTP Binding	Role
IF-1	No	Blocks A site on 30S subunit so fmet-tRNAf binds in P site
IF-2	Yes	Binds only fmet-tRNA (initiator tRNA)
IF-3	No	Stabilizes 30S subunit dissociated from inactive 70S ribosome and aids mRNA

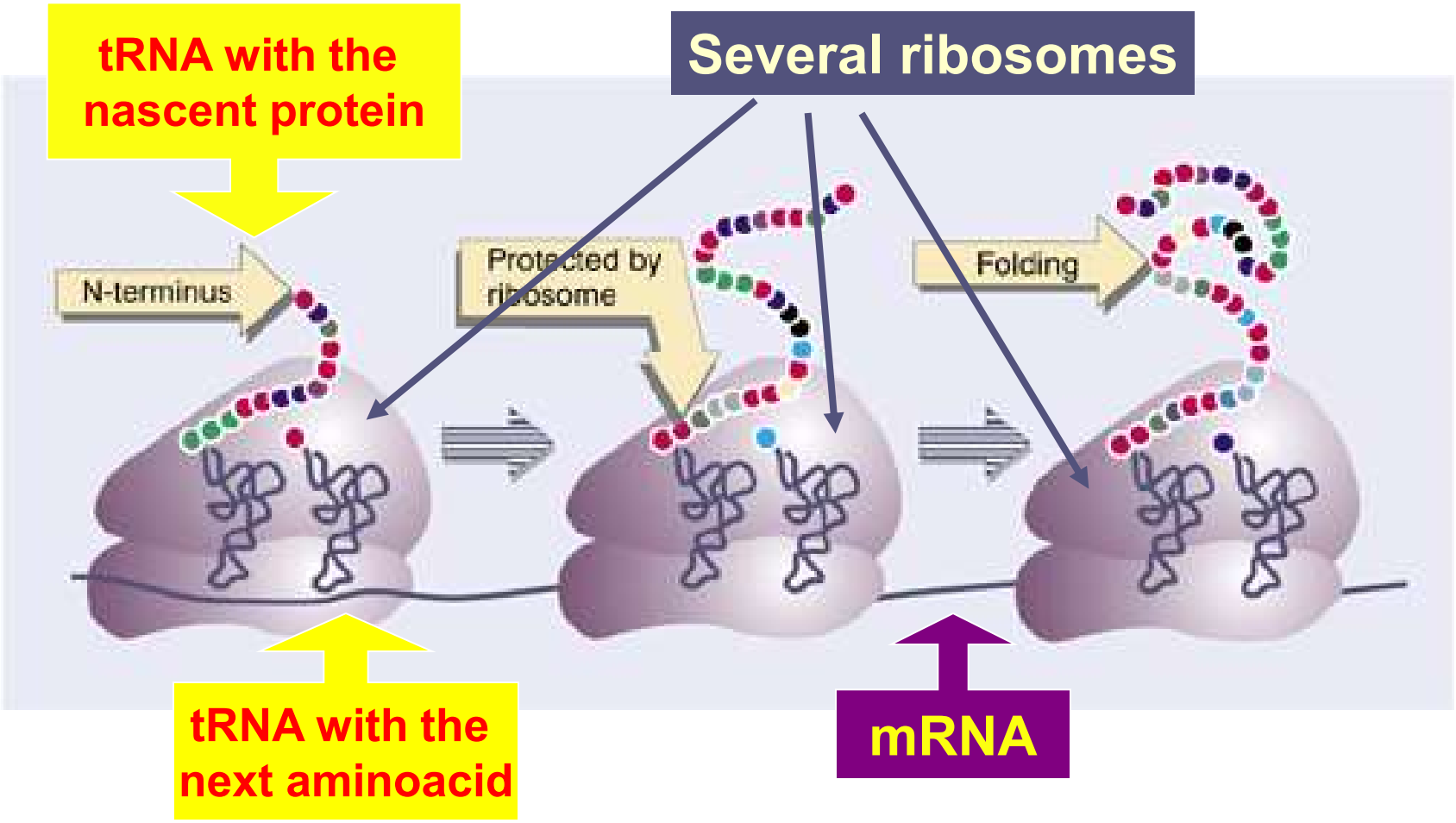
Chain elongation in prokaryotic translation

1. An aminoacyl tRNA in a complex with EF-Tu and GTP enters the A site of the ribosome. Upon codon-anticodon matching, GTP is hydrolyzed and EF-Tu-GDP is released
2. A new peptide bond is formed by transfer of peptide chain from P site to aa-tRNA at A site. The uncharged tRNA moves from the P site to the E site.
3. Translocation of peptidyl tRNA from A site to P site is catalyzed by EF-G. GTP is cleaved and the ribosome moves along mRNA by one codon.
4. As soon as the A site is vacated the uncharge tRNA is released from the E site.

EF-Tu-GTP is regenerated in a separate cycle catalyzed by EF-Ts.



Polyribosome



- both eukaryotes and prokaryotes

Moving in the direction of 5′ to 3′

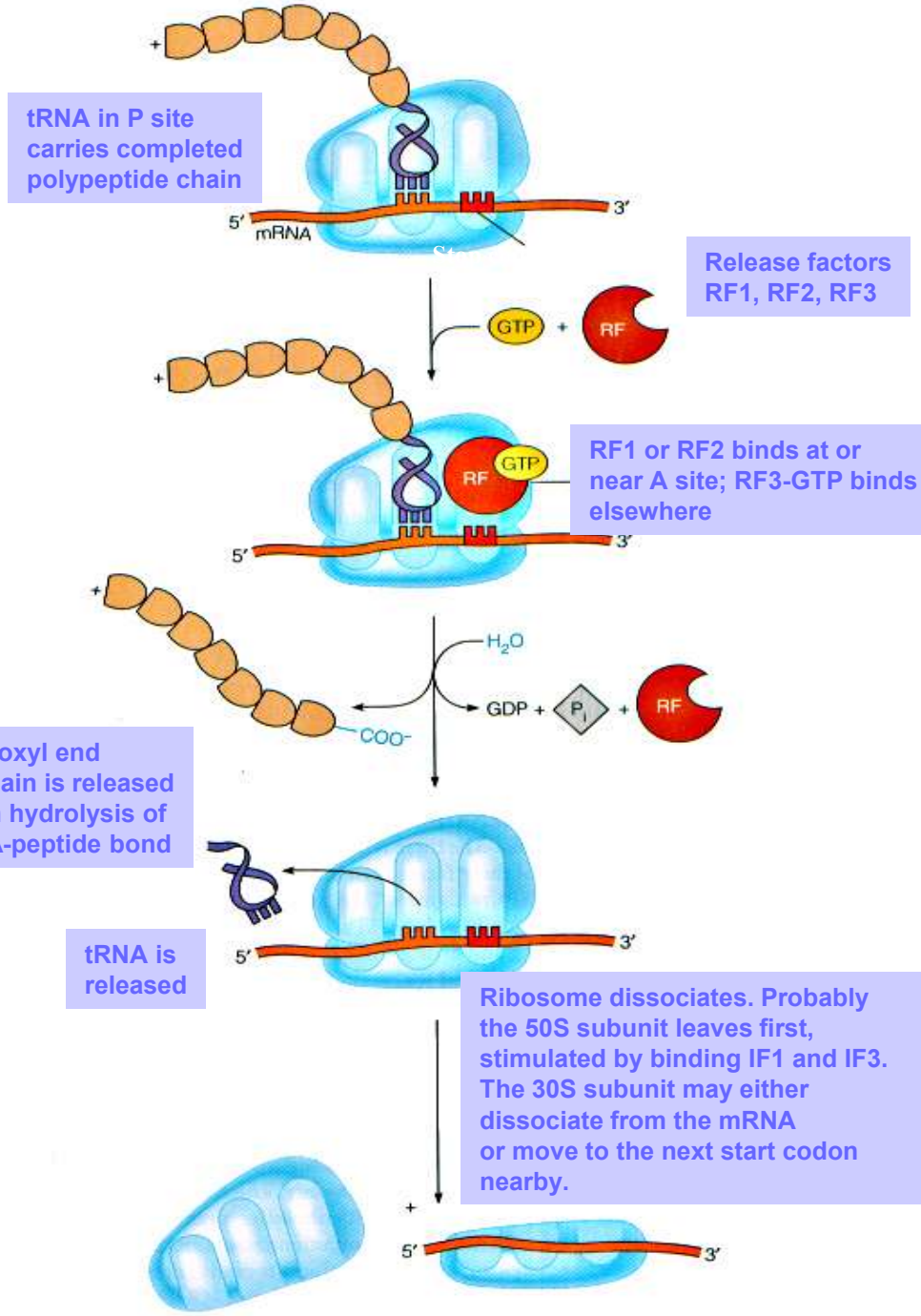
Prokaryotic elongation factors

Factors	GTP Binding	Role
EF-Tu	Yes	Binds all aminoacylated tRNAs (but not fmet-tRNA nor met-tRNAi) Most abundant protein in *E. coli* 1 copy per ribosome; 20,000 molecules per cell
EF-Ts	No	Displaces GDP from EF-Tu
EF-G	Yes	Translocation

Only active when bound to GTP not GDP

Termination of translation in prokaryotes

1. tRNA in the P site carries a completed polypeptide chain. There is a stop codon in the mRNA facing the A site.
2. RF1 or RF2 binds at or near the A site, RF3-GTP binds somewhere else.
3. Peptidyl transferase transfers the poly-peptide chain from the P site to a wate molecule. This releases the protein fro ribosome.
4. The RF factors are then released, followed by the uncharged tRNA.
5. The ribosome is now unstable. Its instability is accentuated by the presence of the initiation factors IF1 and IF3. It dissociates into its subunits.



Prokaryotic termination factors

TERMINATION CODONS UAA Ochre - most common

UAG Amber

UGA Opal

No tRNA molecules to recognize

release or termination factors

Respond to termination codons

1 RF per ribosome

Cause release of the completed polypeptide chain and the ribosome

Activate ribosome to hydrolyze

RF-1 UAA and UAG

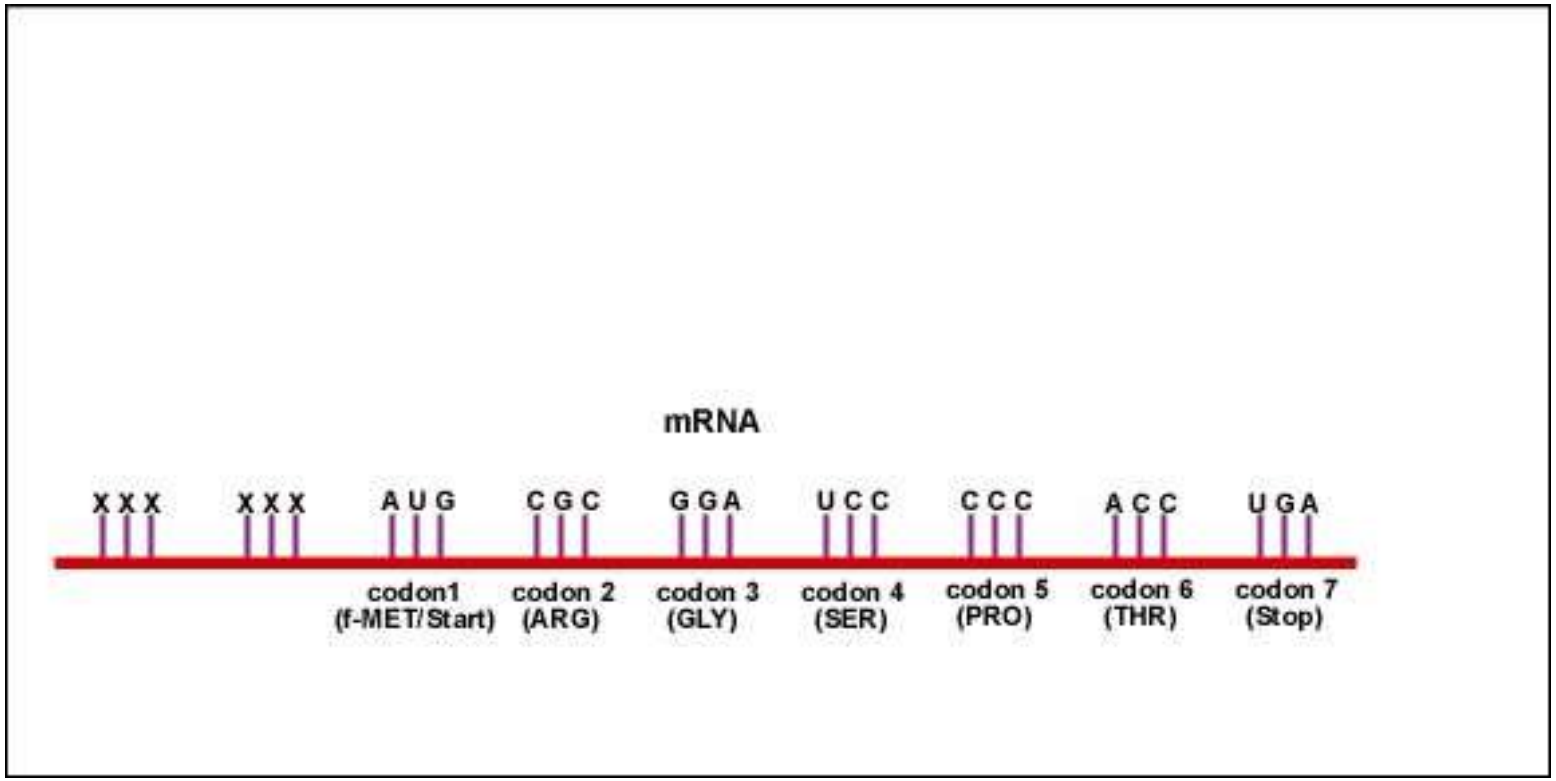
RF-2 UAA and UGA

Dissociation of complex

Release of peptide

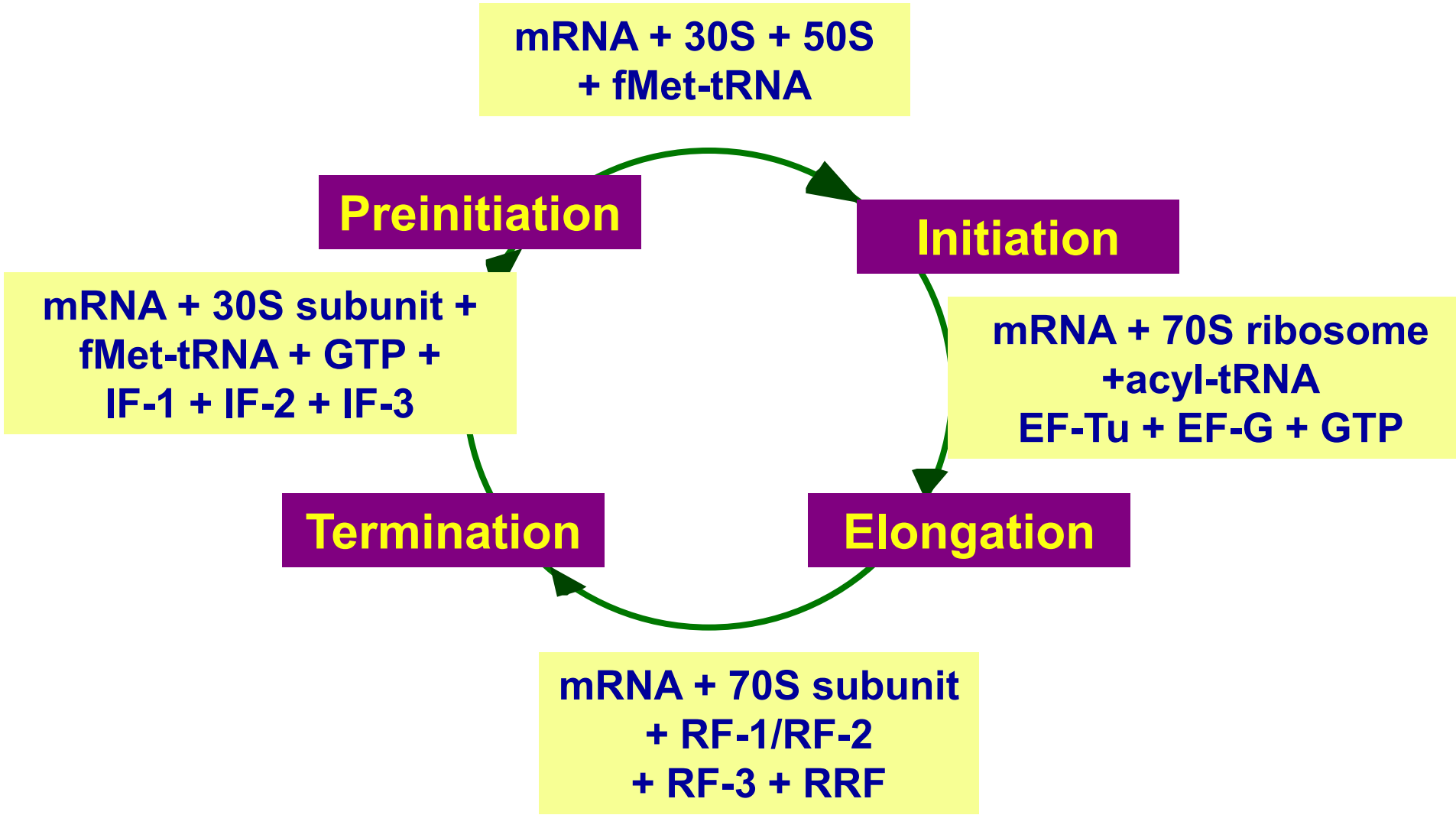
Expulsion of tRNA

Translation Cycle

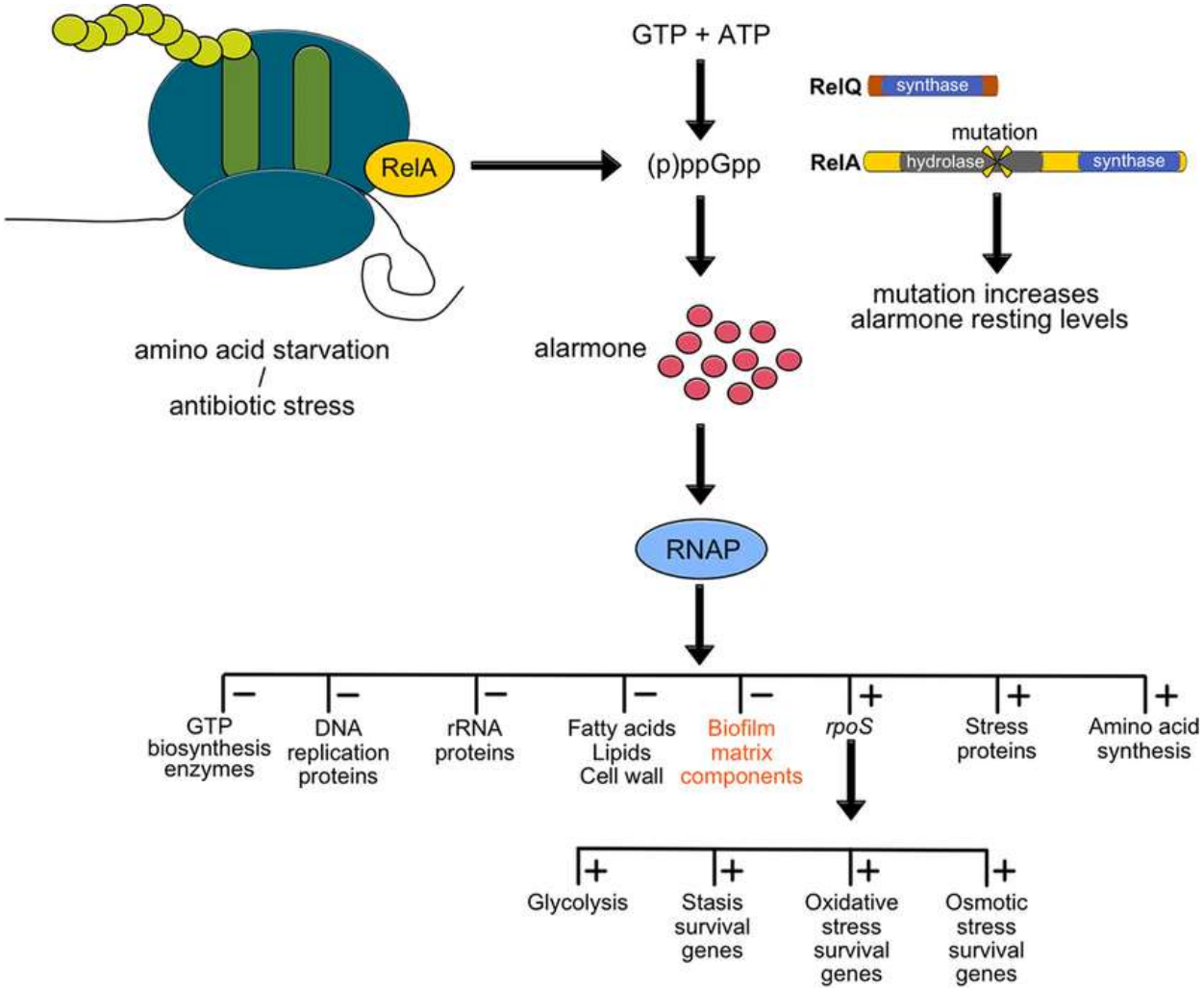


High energy cost of translation - 90% of energy produced in *E. coli* is used for protein synthesis

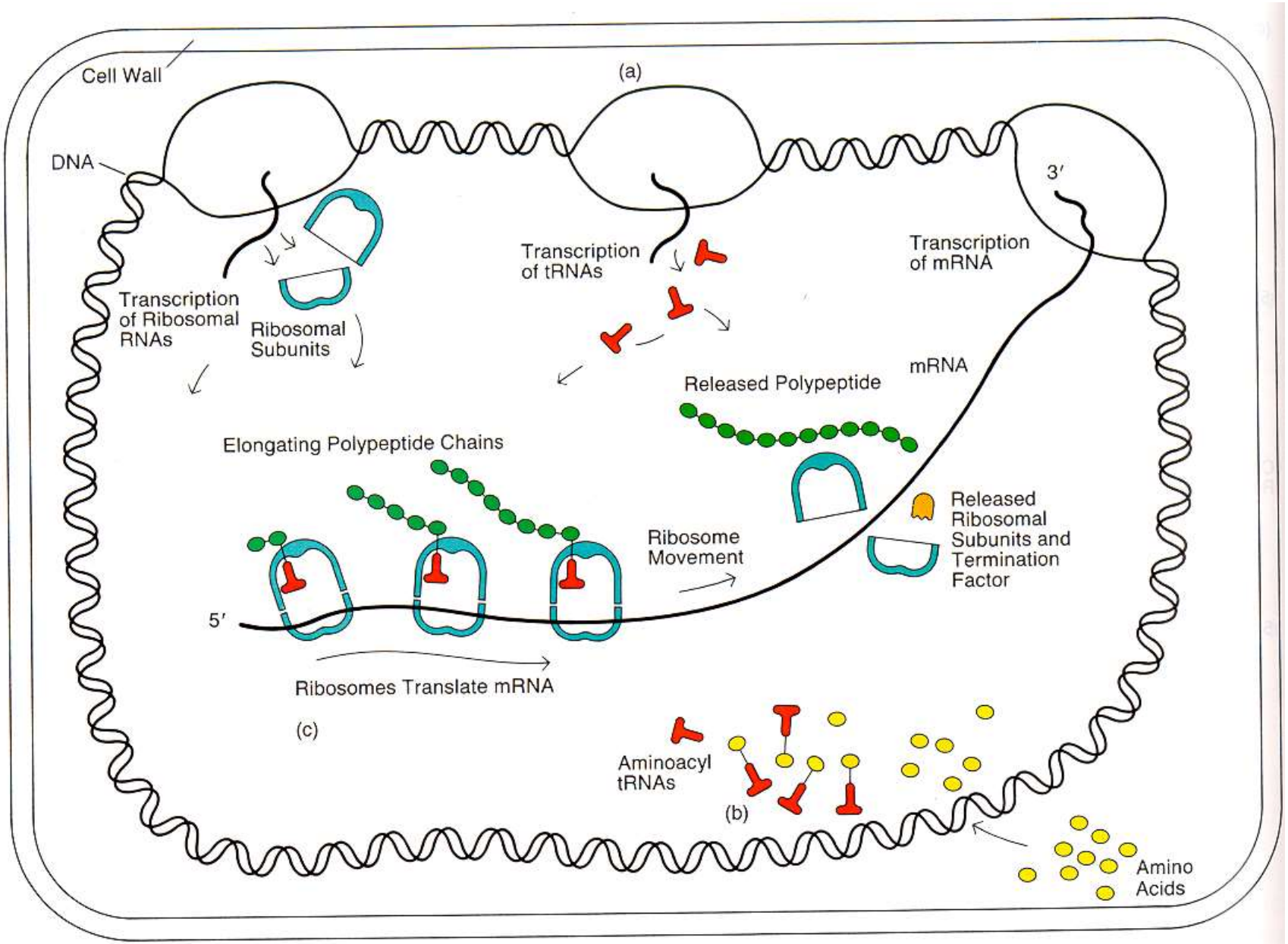
Translation Cycle



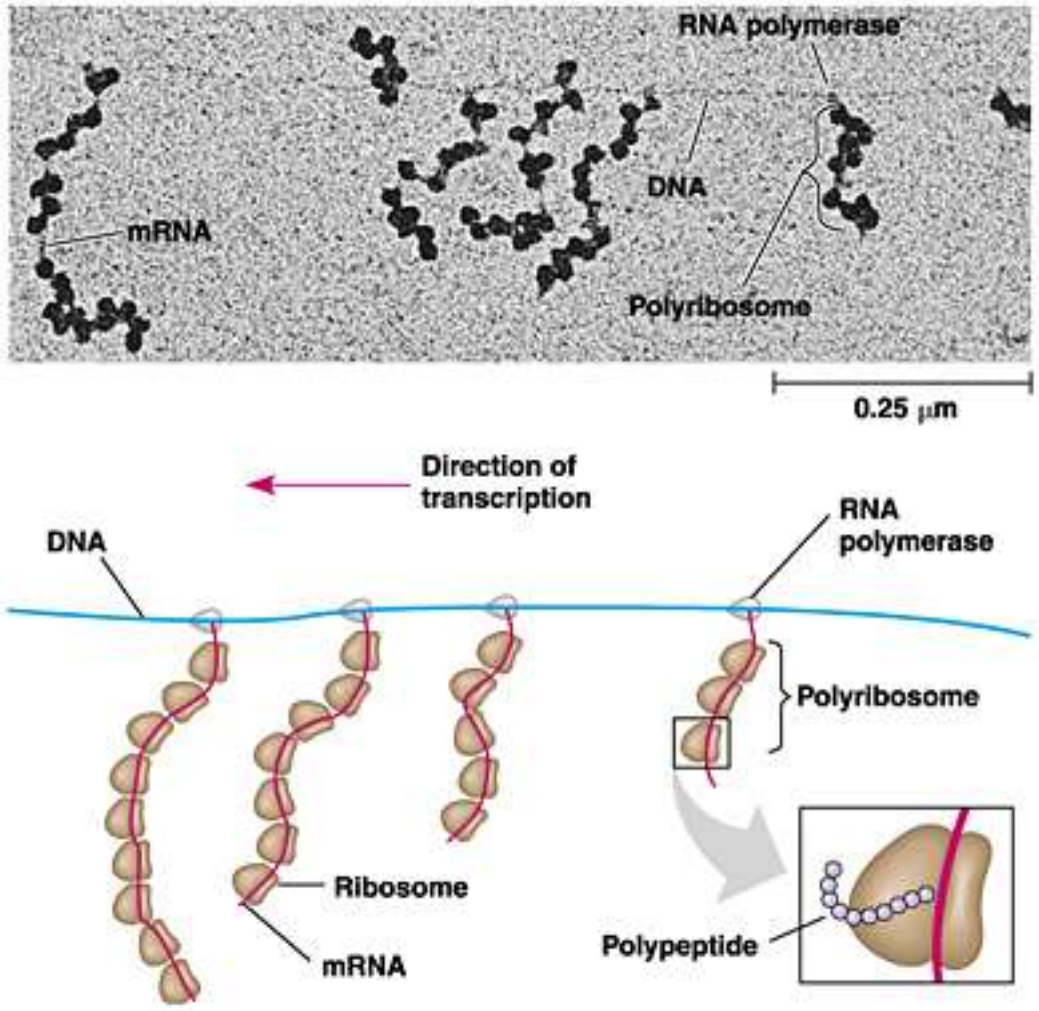
Stringent response



Overview of transcription and translation process in prokaryotes



Transcription and translation are coupled in prokaryotes



Key points of translation in eukaryotes

The requirement for translation factors is considerably more complex than in prokaryotes. At least 11 proteins are required.

Some of the factors bind to mRNA rather than the ribosome.

The major initiation factor eIF2 forms a complex with Met-tRNA^F and GTP. After binding to the 60S subunit, and formation of the 80S ribosome, eIF2-GDP is recycled as in prokaryotes.

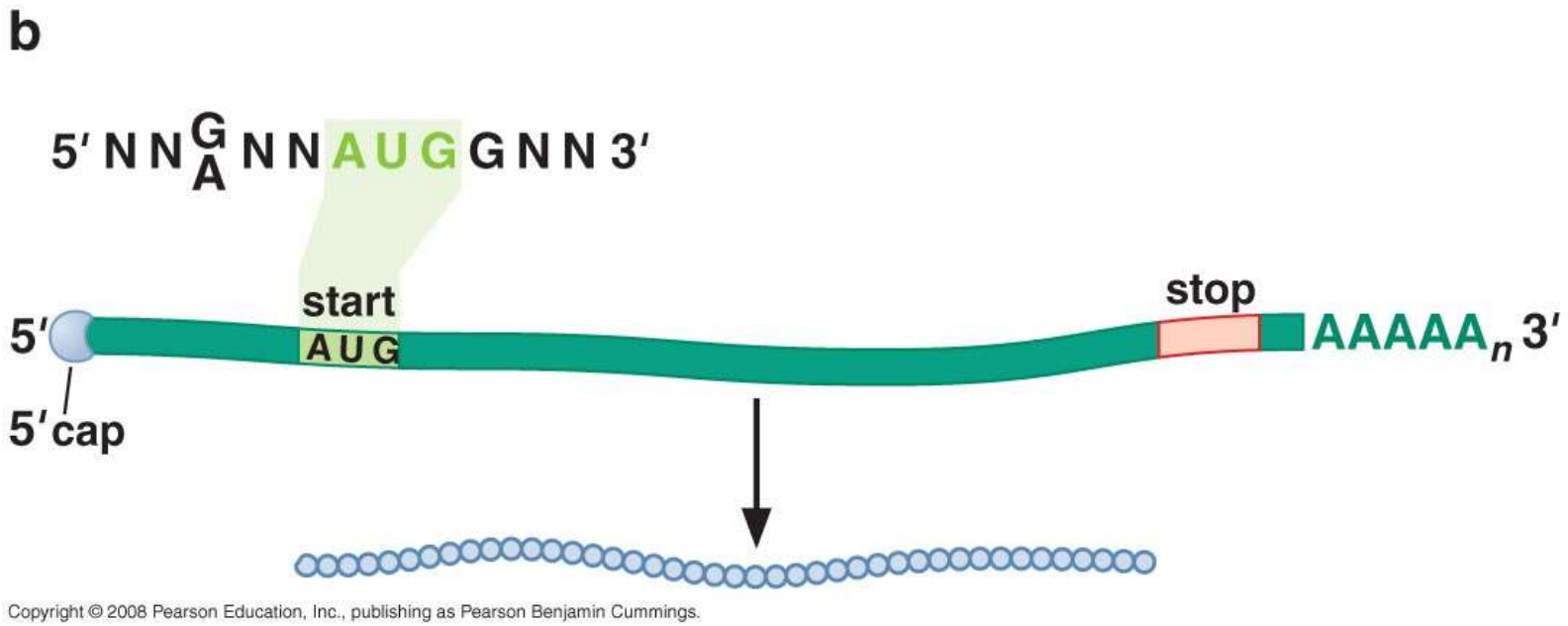
The initiator Met-tRNA^F is not formylated.

There is no Shine-Dalgarno sequence. Instead the 40S subunit binds to the cap and scans the mRNA until it finds the first AUG. This requires ATP.

Elongation is very similar to the prokaryotic process.

Termination is very similar to the prokaryotic process, but requires only a single release factor.

The Kozak sequence (PuNNAUGG) interacts with initiator tRNA. Poly-A tail promotes efficient recycling of ribosomes



Summary of Translation Initiation

Translation initiation is the process whereby the ribosome and the initiator methionyl tRNA is recruited to the start codon
The Process requires:

i) The Ribosome-40S and 60S subunits

ii) protein factors-eukaryotic initiation factors

eIF1, 1A, 2, 2B, 3, 4A, 4B, 4E, 4G, 5, 5B.

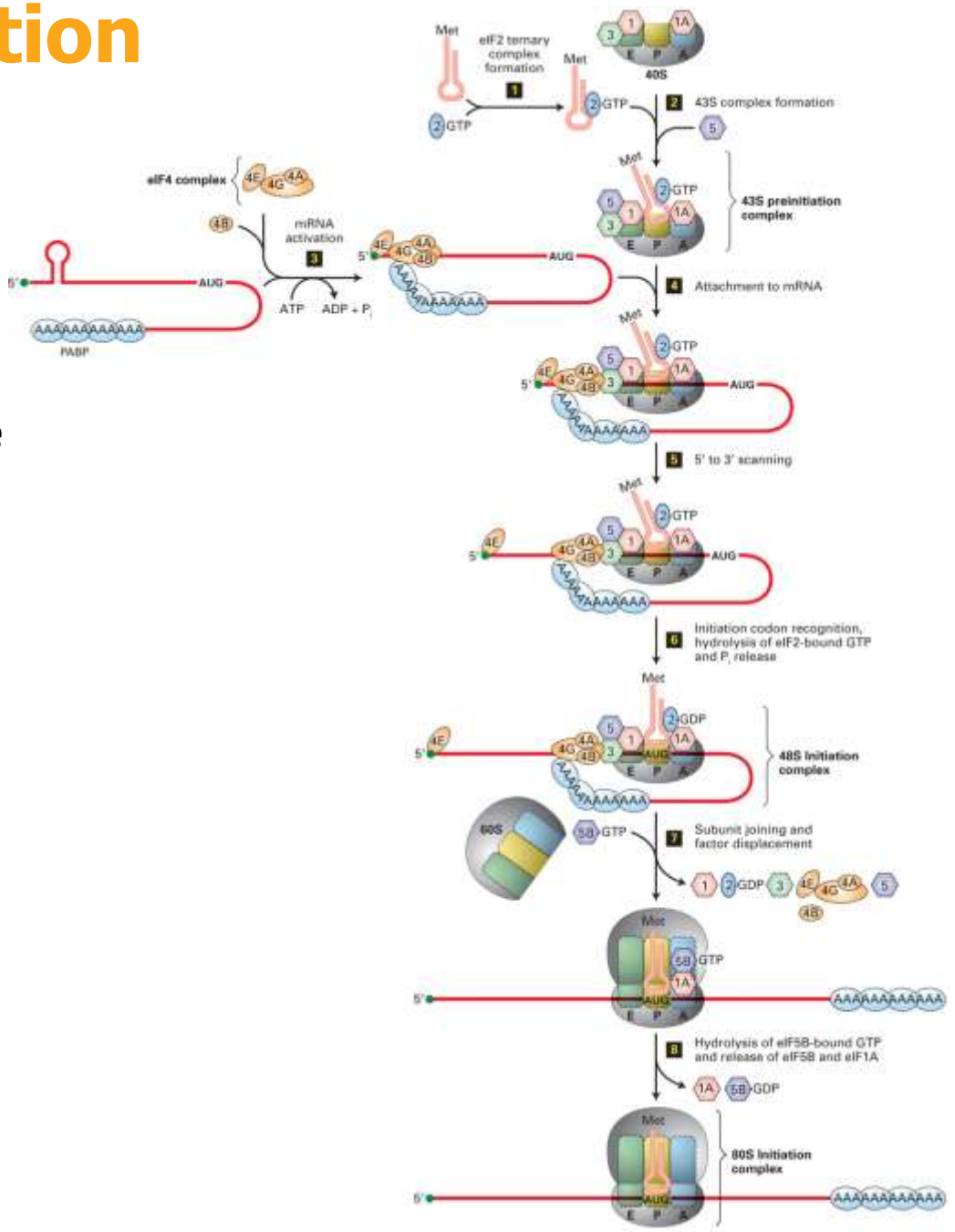
iii) aminoacylated initiator methionyl tRNA (Met-$tRNA_i^{Met}$)

This process can be broken down into 5 steps:-

1. 40S ribosomal subunit and $tRNA_i^{Met}$ preparation
2. mRNA selection and preparation
3. 40S/ mRNA binding,
4. scanning and AUG recognition
5. 60S ribosomal subunit joining

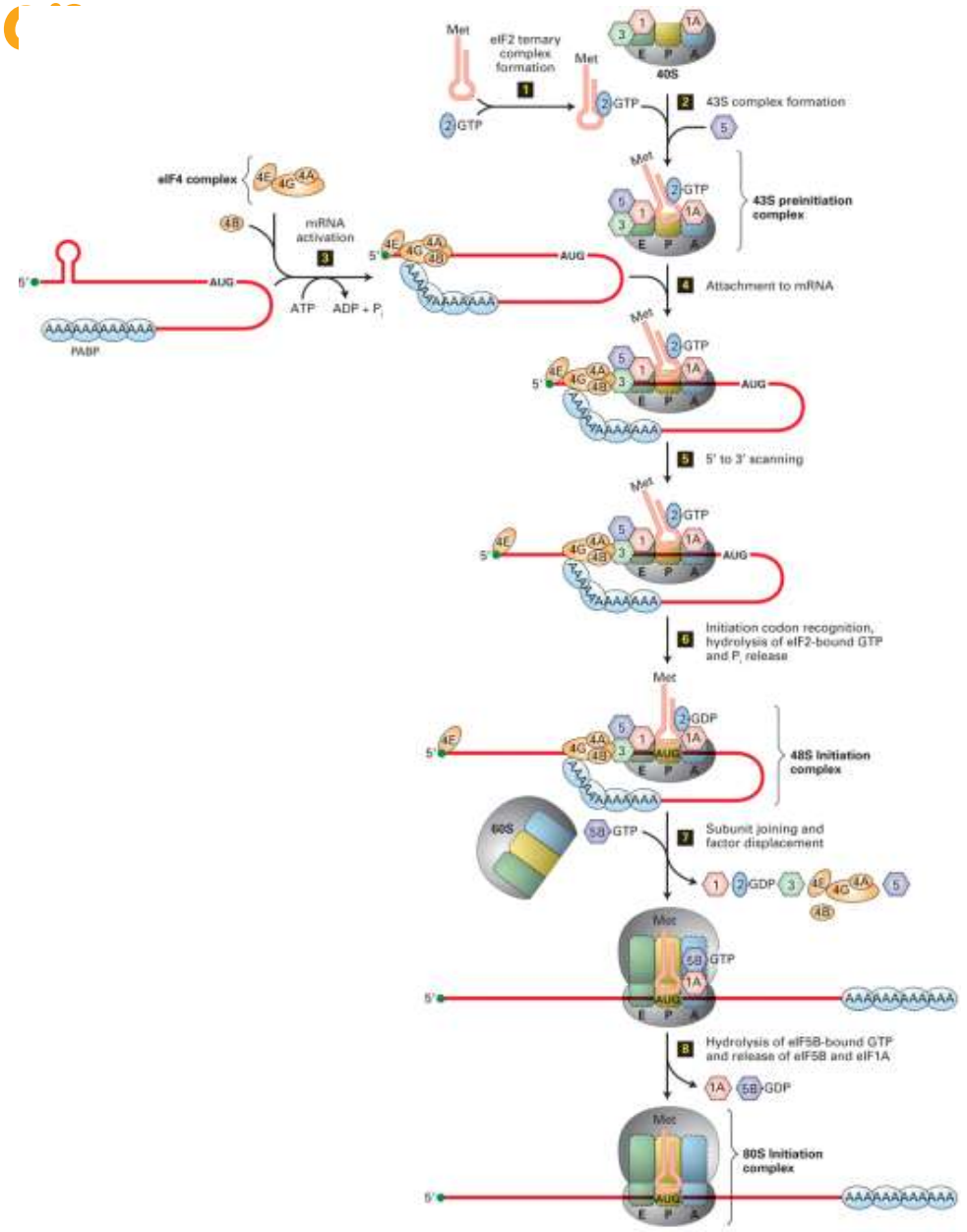
Translation Initiation in Eukaryotes I

Translation initiation in eukaryotes begins with three components/complexes. These are 1) the 40S ribosomal subunit, to which the eIF1, eIF1A, and eIF3 initiation factors are bound; 2) the eIF2·GTP + Met-tRNA$_i^{Met}$ ternary complex; and 3) a circular mRNA formed by the binding of the eIF4 cap-binding complex at the 5' end of the mRNA to poly(A) binding protein (PABP) associated with the 3' end of the mRNA. These components associate in Steps 2 and 4 of the diagram, placing Met-tRNA$_i^{Met}$ in the P site of the 40S subunit.



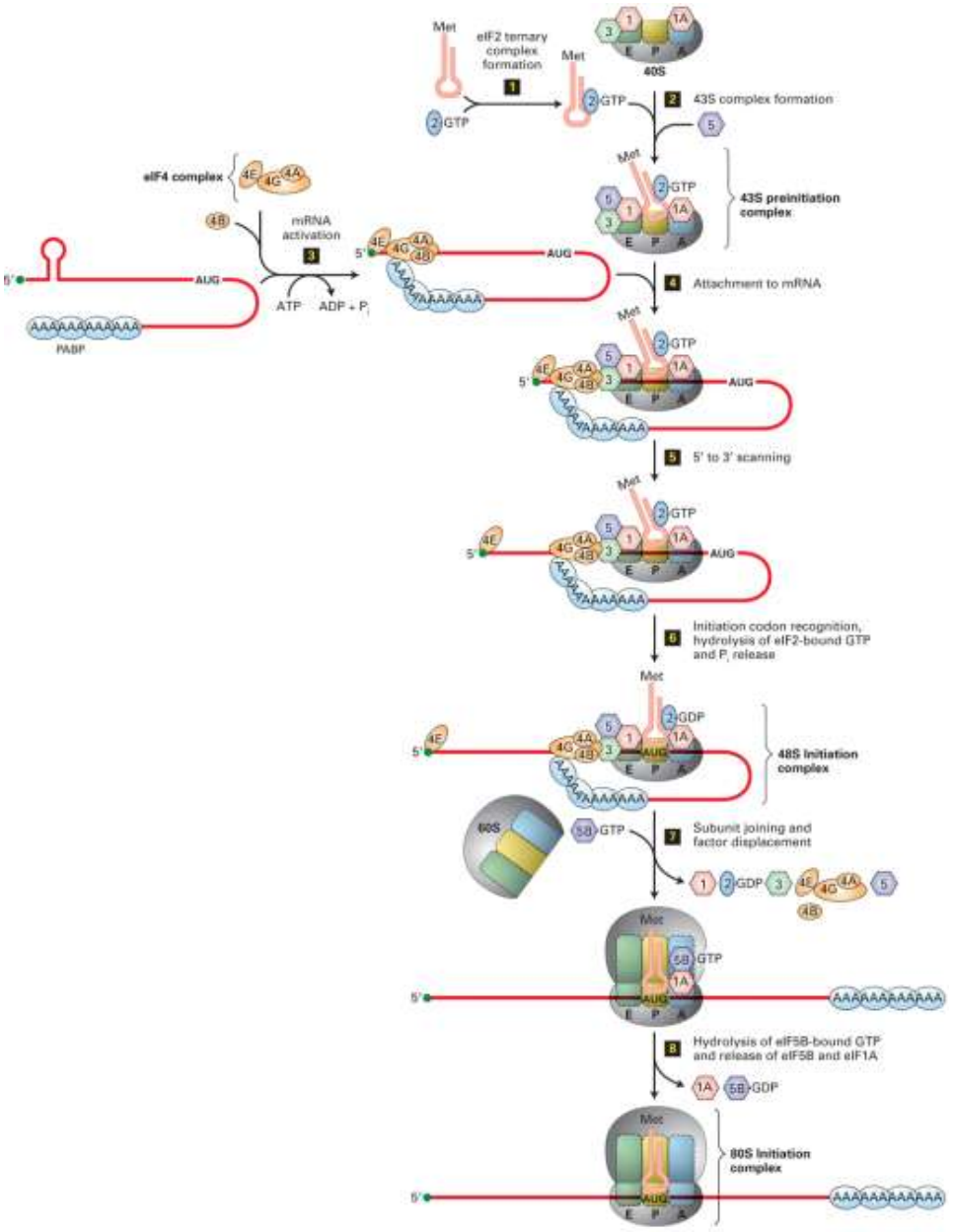
Translation Initiation in Eukaryotes II

In the next stage of initiation, the mRNA is scanned in the 5' to 3' direction until the first AUG start codon is brought into the P site (Steps 5 & 6). Then the hydrolysis of GTP by eIF2 generates a stable <u>48S initiation complex</u> in which the initiator tRNA (Met-tRNA$_i^{Met}$) is H-bonded to the AUG codon.



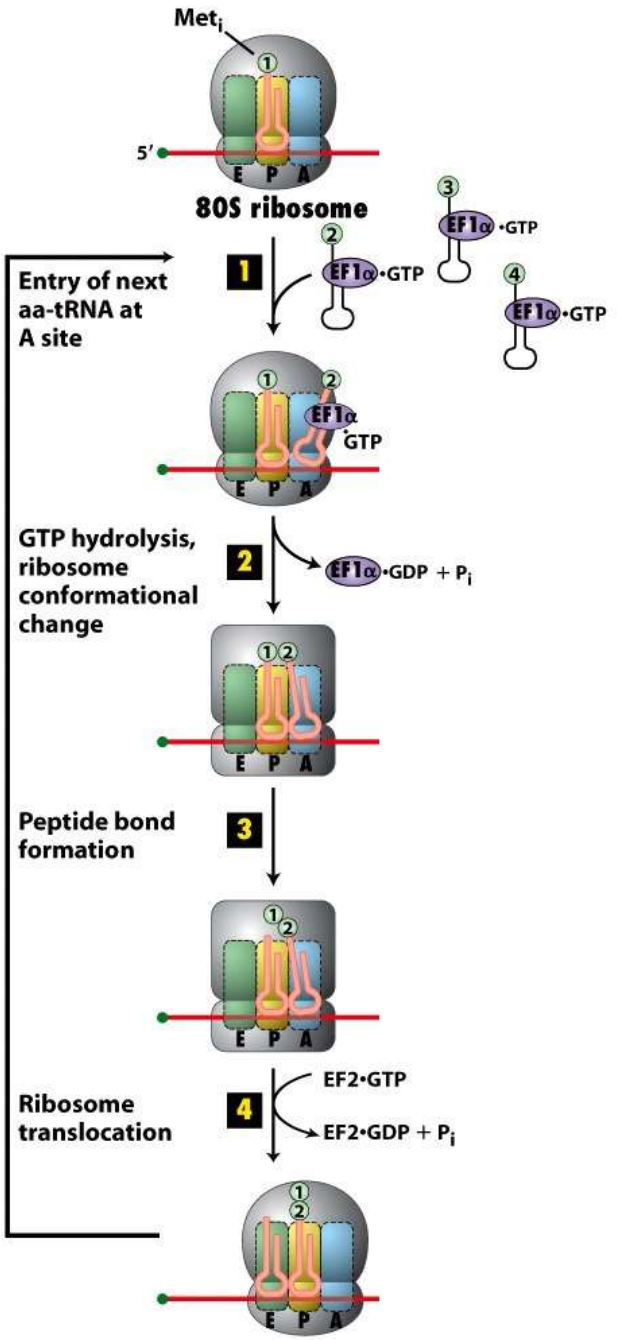
Translation Initiation in Eukaryotes III

In the final stages of initiation, all initiation factors except eIF1A dissociate from the 48S initiation complex and the 80S subunit and eIF5B·GTP complex add on (Step 7). After eIF5B hydrolyzes GTP, the last initiation factors depart, and the stable 80S initiation complex is created (Step 8). This complex contains the complete E (exit), P (peptidyl-tRNA), and A (aminoacyl-tRNA) binding sites, with Met-tRNA$_i^{Met}$ bound to the P site.



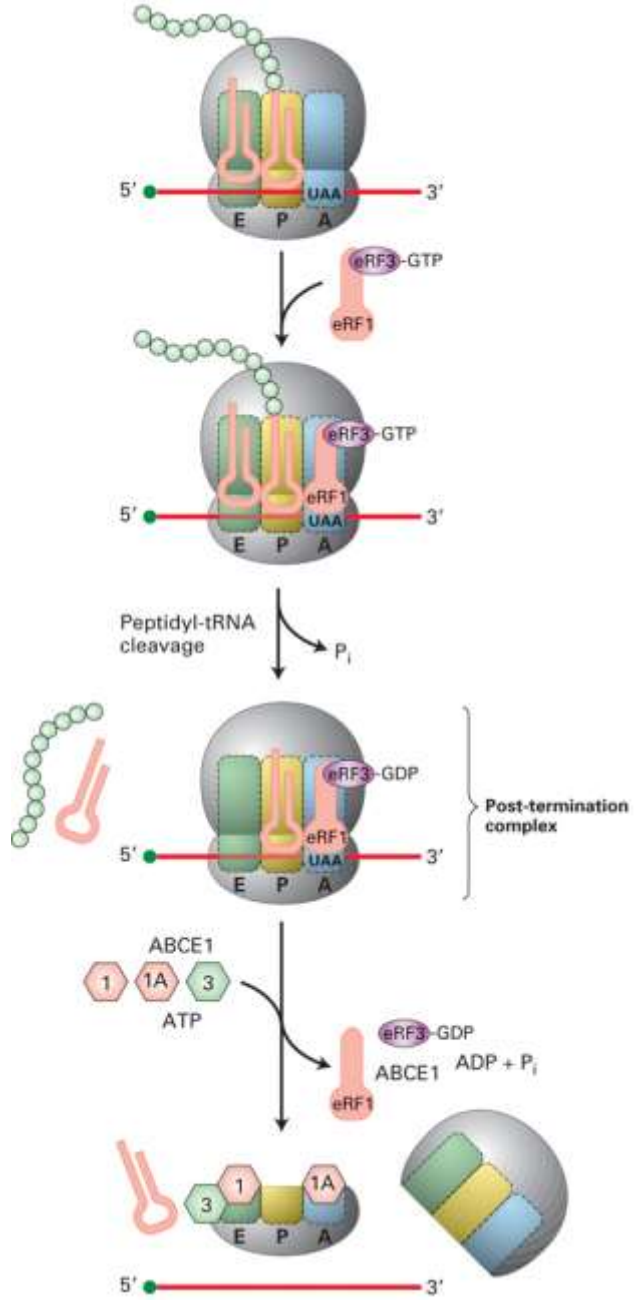
Translation Elongation in Eukaryotes

Translation elongation requires the assistance of elongation factors. In Step 1 of elongation, the second amino acid of the polypeptide is carried to the A site of the ribosome by an EF1α·GTP complex. It binds to the mRNA via the anticodon located in the A site. In Step 2, GTP is hydrolyzed and EF1α departs. In Step 3, the 28S rRNA of the 60S subunit catalyzes peptide bond formation (see Fig. 4.17), resulting in a dipeptidyl-tRNA residing in the A site. In Step 4, the factor EF2·GTP



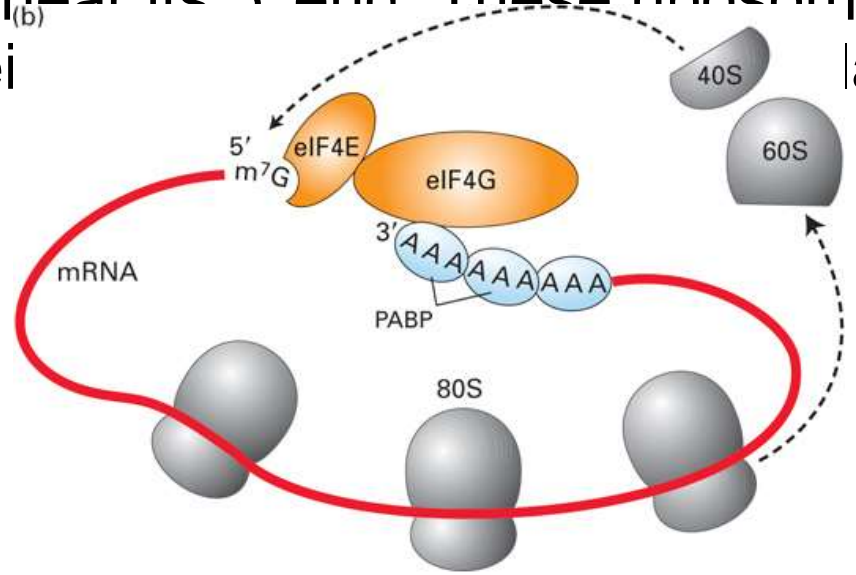
Translation Termination in Eukaryotes

When a stop codon (UAA, UAG, UGA) enters the A site, it is recognized and bound by the eRF1 release factor (Fig. 4.27). eRF1 forms a complex with eRF3·GTP. Hydrolysis of GTP by eRF3 results in cleavage of the linkage between the polypeptide and peptidyl-tRNA and release of the protein from the ribosomal post-termination complex. A protein called ABCE1 then binds to the complex, and via ABCE1 hydrolysis of ATP, the 40S and 60S subunits are separated. The 40S subunit recombines with the eIF1, eIF1A, and eIF3 factors making it ready for another round



Polysomes & Ribosome Recycling

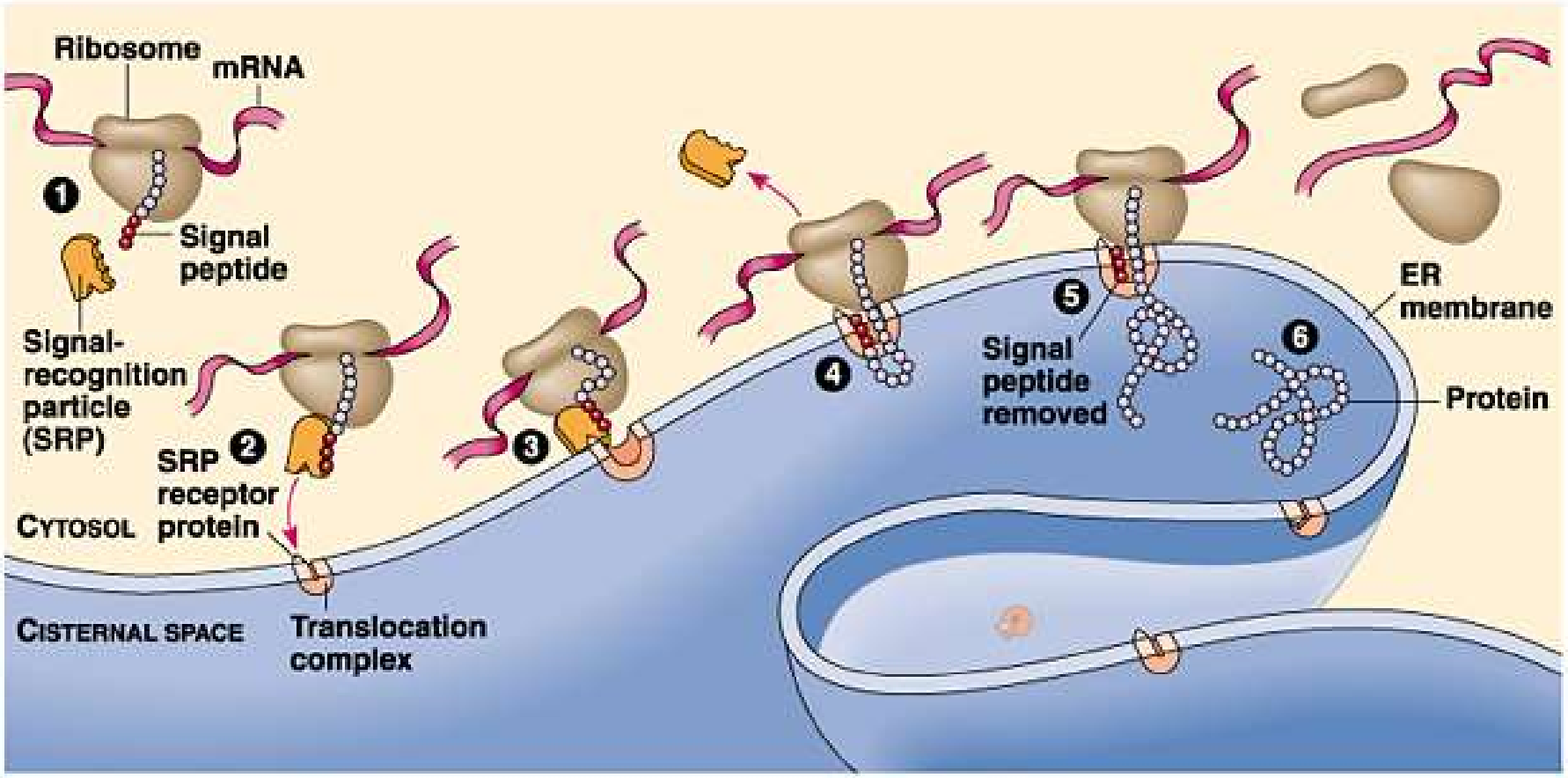
Polypeptide chain elongation proceeds at a rate of 3-5 amino acids per second. The efficiency of translation is increased via the binding of multiple ribosomes (polysomes) to the mRNA at a given time (Fig. 4.28b). Translation efficiency is further increased due to the complex between poly(A)-binding protein (PABP) and the eIF4-mRNA 5'-cap that occurs in mRNA (Fig. 4.28b). This circular complex positions ribosomes that have just terminated translation of the message near its 5' end. These ribosomes are recycled and rapidly rei ation.



Eukaryotic initiation factors

Factor	Role
eIF-1A	binds and stimulates 43S complex formation
eIF-2	binds met-tRNA, regulates ternary complex formation
eIF-2A	binding of ternary and 43S complex
eIF-2B	GTP to GDP conversion
eIF-3	ribosome subunit dissociation and stabilizes 40S subunit
eIF-3A	ribosome subunit dissociation and stabilizes 60S subunit

Translation and translocation are coupled in eukaryotes



Translation and translocation are coupled in eukaryotes

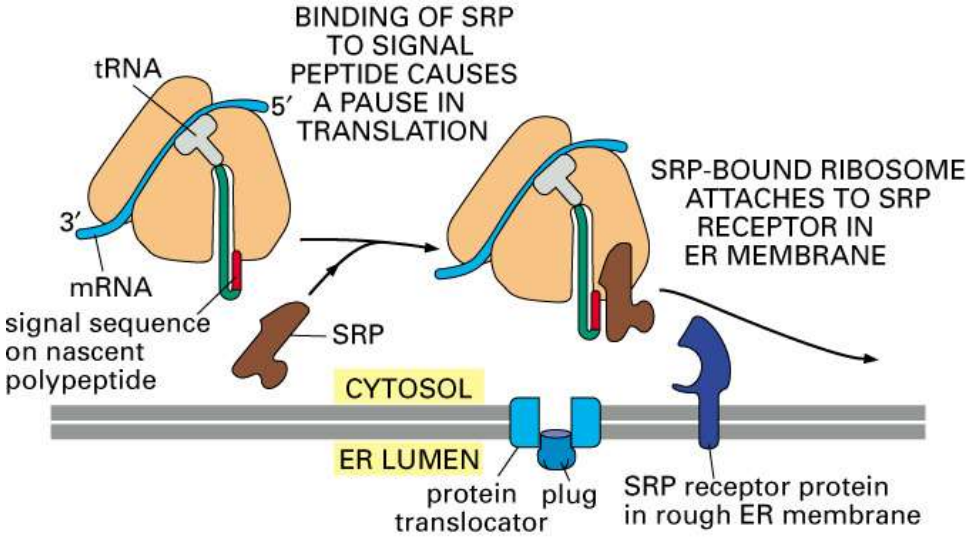


Figure 12–42 part 1 of 2. Molecular Biology of the Cell, 4th Edition.

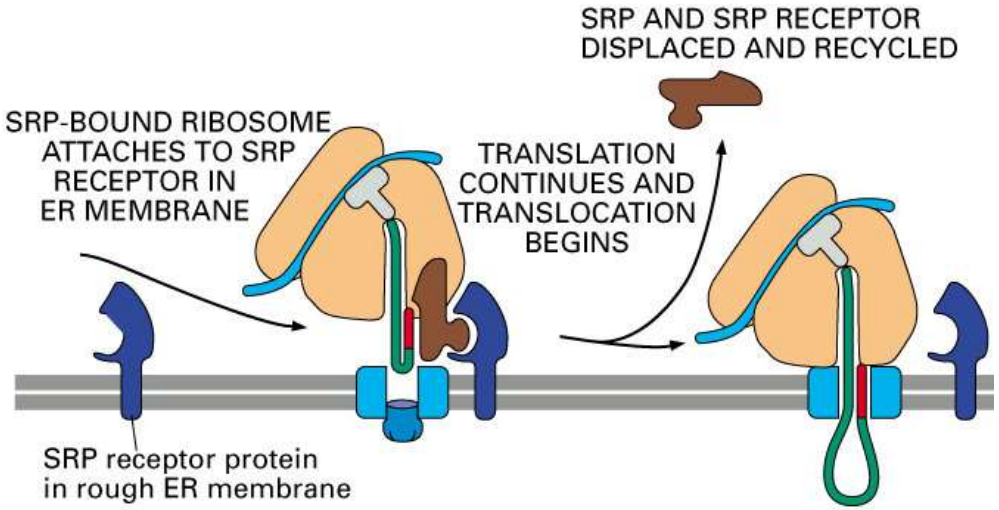


Figure 12–42 part 2 of 2. Molecular Biology of the Cell, 4th Edition.

- A signal sequence of approximately 20 amino acids and rich with hydrophobic amino acids is often located at the N-terminus.
- Since the ribosome masks about 30 amino acids, the signal sequence isn't fully exposed until the nascent polypeptide is about 50 amino acids long.
- SRP-ribosome attaches to SRP receptor and then docks on a protein translocator.
- SRP and receptor dissociate.
- Translation and translocation proceed in unison - co-translational transport.
- The energy for transport is provided by the translation process - as the polypeptide grows, it is pushed through the protein translocator.

SRP: signal-recognition particle
SRP receptor

Translation and translocation are coupled in eukaryotes

The signal sequence of secreted proteins is cleaved by a signal peptidase. In the literature, the signal sequence of secreted proteins is often called a "leader peptide".

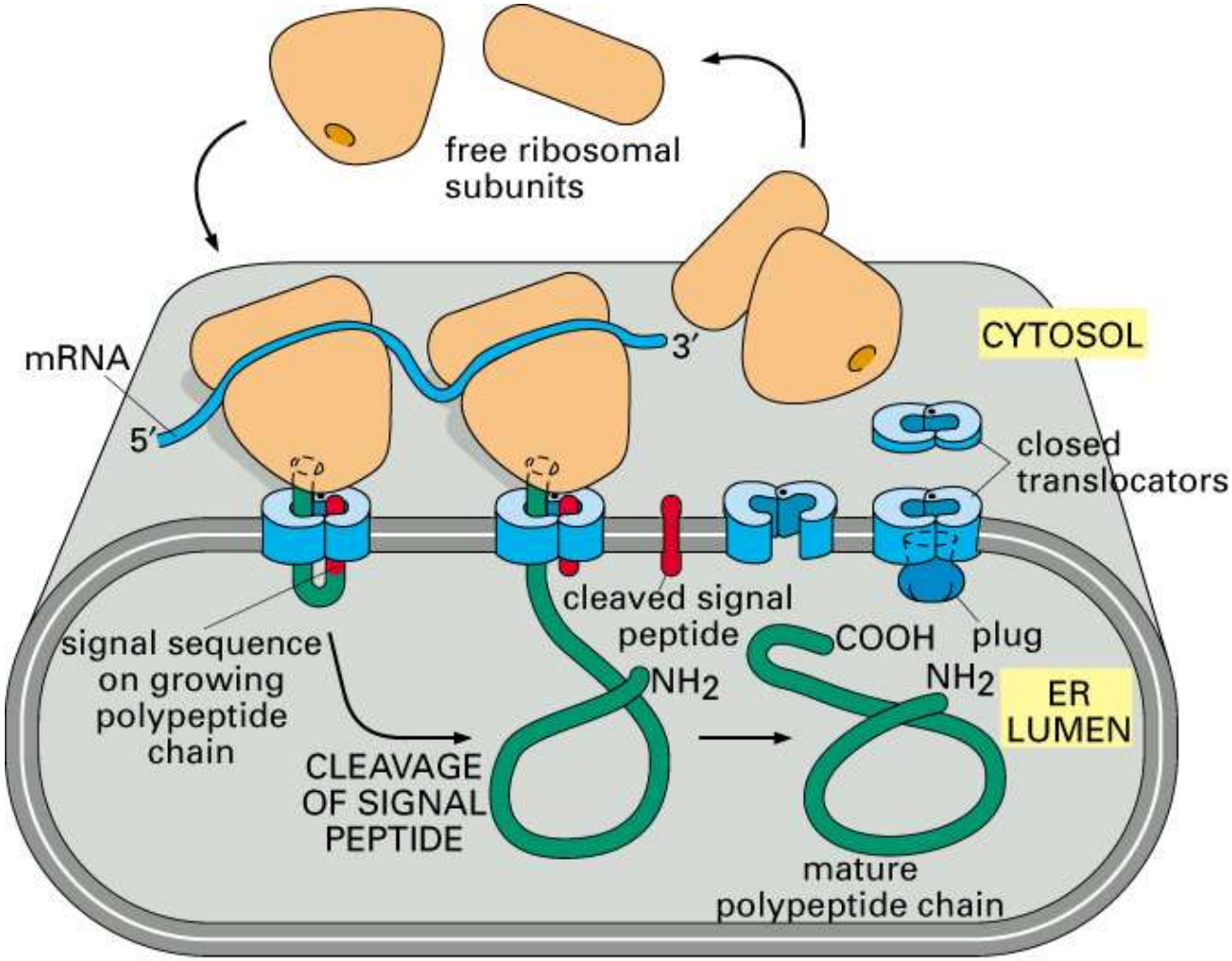


Figure 12–40. Molecular Biology of the Cell, 4th Edition.

Signal sequences

Table 14-3 Some Typical Signal Sequences

Function of Signal	Example of Signal Sequence
Import into ER	$^{+}H_3N$-Met-Met-Ser-Phe-Val-Ser-**Leu-Leu-Leu-Val-Gly-Ile-Leu-Phe-Trp-Ala**-Thr-**Glu**-Ala-**Glu**-Gln-Leu-Thr-**Lys**-Cys-**Glu**-Val-Phe-Gln-
Retention in lumen of ER	-**Lys**-**Asp**-**Glu**-Leu-COO^{-}
Import into mitochondria	$^{+}H_3N$-Met-Leu-Ser-Leu-**Arg**-Gln-Ser-Ile-**Arg**-Phe-Phe-**Lys**-Pro-Ala-Thr-**Arg**-Thr-Leu-Cys-Ser-Ser-**Arg**-Tyr-Leu-Leu-
Import into nucleus	-Pro-Pro-**Lys**-**Lys**-**Lys**-**Arg**-**Lys**-Val-
Import into peroxisomes	-Ser-**Lys**-Leu-

Positively charged amino acids are shown in ***red***, and negatively charged amino acids in ***green***. An extended block of hydrophobic amino acids is shown in ***blue***. $^{+}H_3N$ indicates the amino terminus of a protein; COO^{-} indicates the carboxyl terminus.

Organelles have DNA

Figure 3.13 Human mitochondrial DNA has 22 tRNA genes, 2 rRNA genes, and 13 protein-coding regions. 14 of the 15 protein-coding or rRNA-coding regions are transcribed in the same direction. 14 of the tRNA genes are expressed in the clockwise direction and 8 are read counter clockwise.

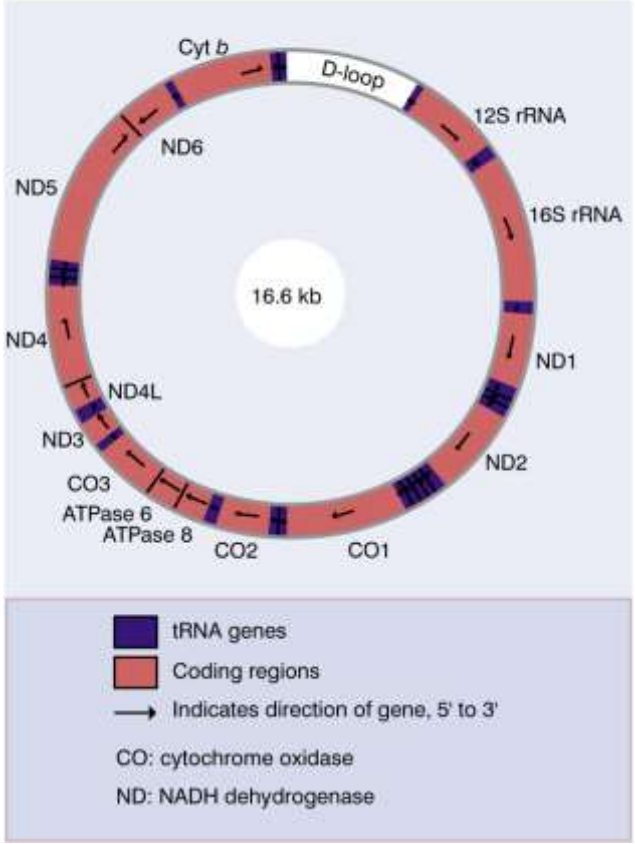


Figure 3.14 The mitochondrial genome of *S. cerevisiae* contains both interrupted and uninterrupted protein-coding genes, rRNA genes, and tRNA genes (positions not indicated). Arrows indicate direction of transcription.

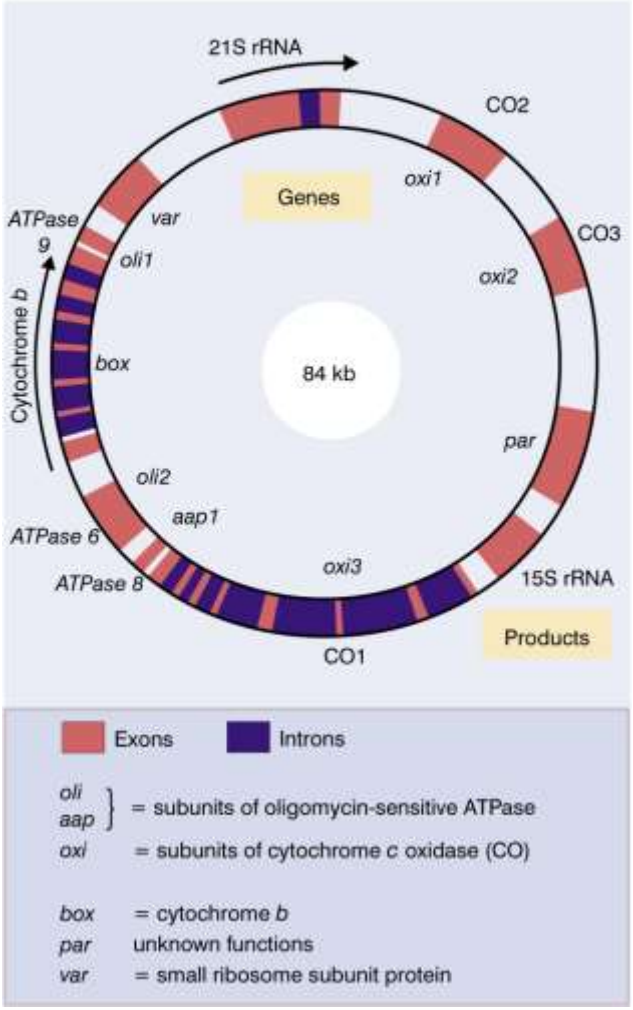
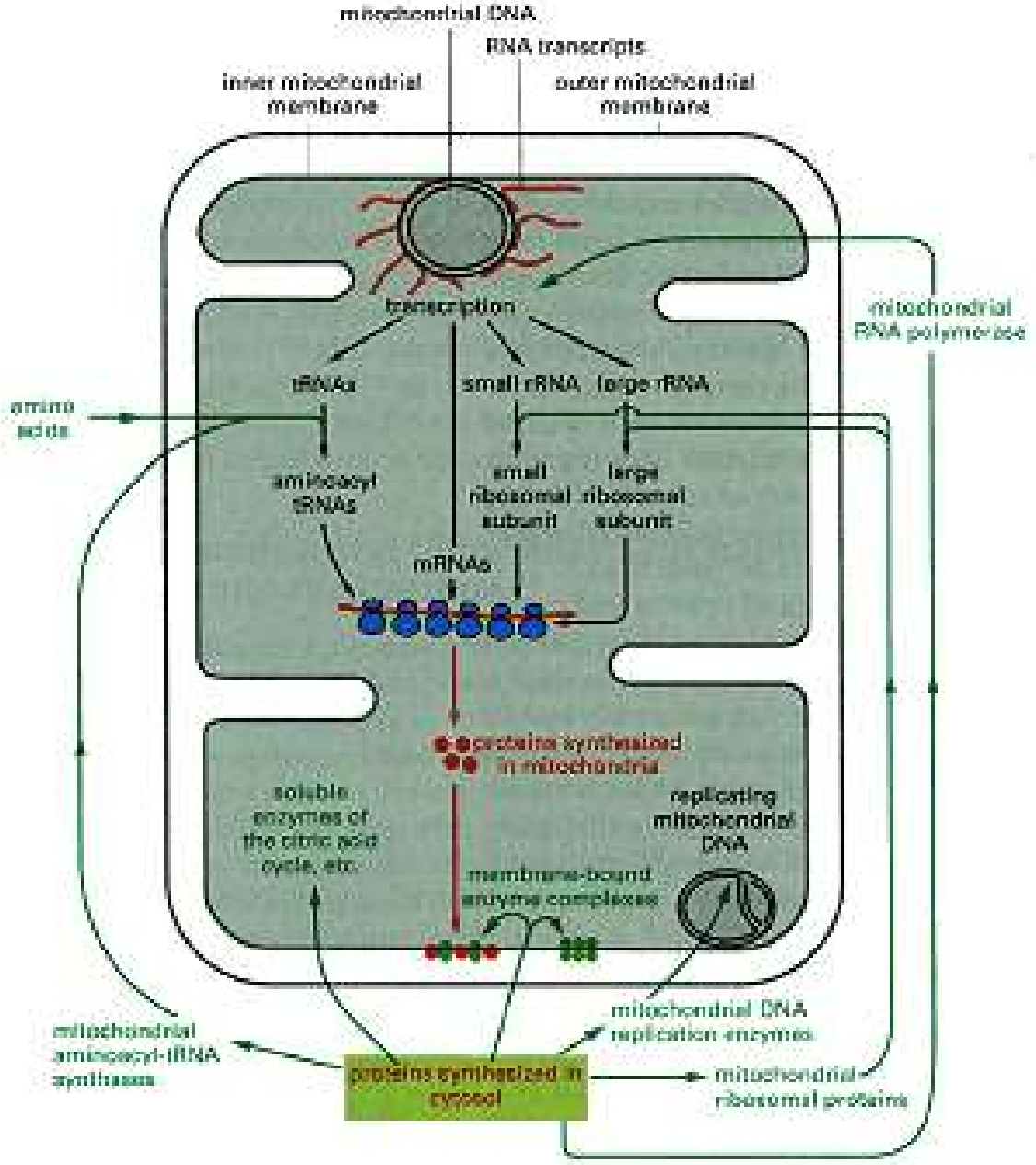


Figure 3.15 The chloroplast genome codes for 4 rRNAs, 30 tRNAs, and ~50 proteins.

Genes	Types
RNA-coding	
16S rRNA	1
23S rRNA	1
4.5S rRNA	1
5S rRNA	1
tRNA	30
Gene Expression	
r-proteins	19
RNA polymerase	3
Others	2
Thylakoid Membranes	
Photosystem I	2
Photosystem II	7
Cytochrome *b/f*	3
H^+-ATPase	6
Others	
NADH dehydrogenase	6
Ferredoxin	3
Ribulose BP Cblase	1
Unidentified	29
Total	110

Translation in mitochondria



Translation in mitochondria

prokaryotic type of ribosomes

dependent upon proteins synthetised in cytoplasm (amino-acyl-tRNA synthetases)

Differences in genetic code used

Mitochondria	UGA	AUA	AGA/G
Mammals	**Trp**	**Met**	**Stop**
Yeasts	**Trp**	**Met**	**Arg**
Drosophila	**Trp**	**Met**	**Ser**
Standard	Stop	Ile	Arg

Translation in mitochondria

Human Mitochondrial mRNA Translation

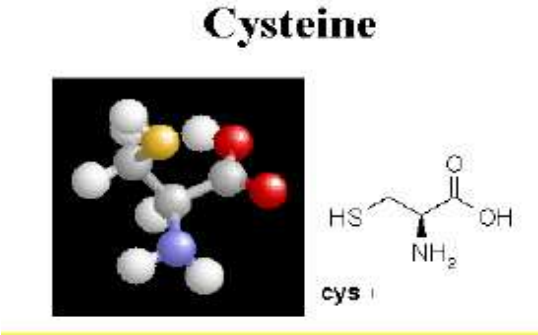
Codon	Nuclear	Mitochondrial	Codon	Nuclear	Mitochondrial
AUU	Ile	Ile	AAU	Asn	Asn
AUC	Ile	Ile	AAC	Asn	Asn
AUA	Ile	**Met**	AAA	Lys	Lys
AUG	Met	Met	AAG	Lys	Lys
GUU	Val	Val	GAU	Asp	Asp
GUC	Val	Val	GAC	Asp	Asp
GUA	Val	Val	GAA	Glu	Glu
GUG	Val	Val	GAG	Glu	Glu
UCU	Ser	Ser	UGU	Cys	Cys
UCC	Ser	Ser	UGC	Cys	Cys
UCA	Ser	Ser	UGA	Stop	**Trp**
UCG	Ser	Ser	UGG	Trp	Trp
CCU	Pro	Pro	CGU	Arg	Arg
CCC	Pro	Pro	CGC	Arg	Arg
CCA	Pro	Pro	CGA	Arg	Arg
CCG	Pro	Pro	CGG	Arg	Arg
ACU	Thr	Thr	AGU	Ser	Ser
ACC	Thr	Thr	AGC	Ser	Ser
ACA	Thr	Thr	AGA	Arg	**Stop**
ACG	Thr	Thr	AGG	Arg	**Stop**

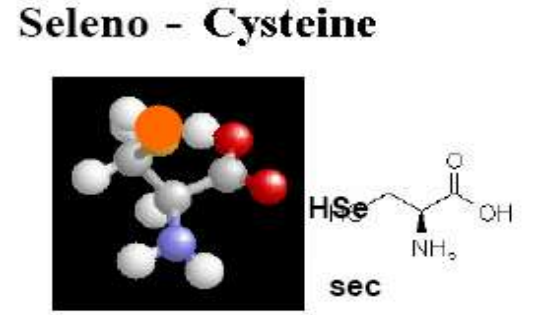
Comparison of prokaryotic and eukaryotic protein synthesis factors

Prokaryotic factor	Eukaryotic factor	Function
Initiation factors		
IF1		Involved in forming
IF2	eIF2	initiation complex
IF3	eIF3, eIF4C	"
	CBPI	Involved in cap binding
	eIF4A, eIF4B, eIF4F	Search for first AUG
	eIF5	Helps dissociate eIF2, eIF3 eIF4C
	eIF6	Helps dissociate 60S subunit from inactive ribosomes
Elongation factors		
EF-Tu	eEF1α	Delivery of aatRNA to ribosomes
EF-Ts	eEF1βγ	Aids in recycling factor above
EF-G	eEF2	Translocation factor
Release factors		
RF1	eRF	Release of completed
RF2		Polypeptide chain
RF3		"

Multifunctional codons

- AUG – methionine
- AUG – start
  - UGA – stop
  - UGA – selenocysteine
    - UGU – cysteine
    - UGC - cysteine





Incorporation of selenocysteine in polypeptide

1- Reduce selinate to selinide

2- Phosphorylate selinide

3- Transfer selenyl from phosphate to methoxy of Ser t-RNA ^{sec}

4a- Form EF^{sec} - Sec-tRNA^{sec} complex (specialized elongation complex)

4b- From SECIS/SBP2 complex

5- Bind SBP2/SECIS /EF^{sec} - Sec-tRNA^{sec} complexes

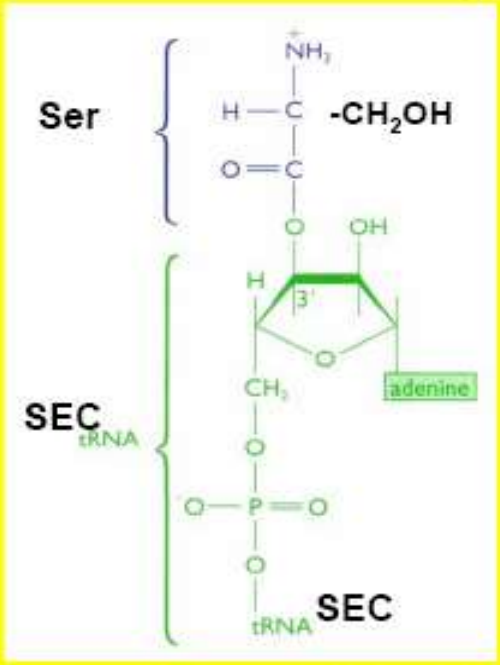
6- Bind Sec-t-RNA ^{sec} to mRNA UGA codon & stimulate peptidyl transferase

Incorporation of selenocysteine in polypeptide

$SeO_3^{2-} + GSH + ATP \rightarrow HSe^- + GSSH + ADP$ Step 1

↓ SEL-D (selenophosphate synthase) Step 2

$Se\text{-}PO_3^{-2}$



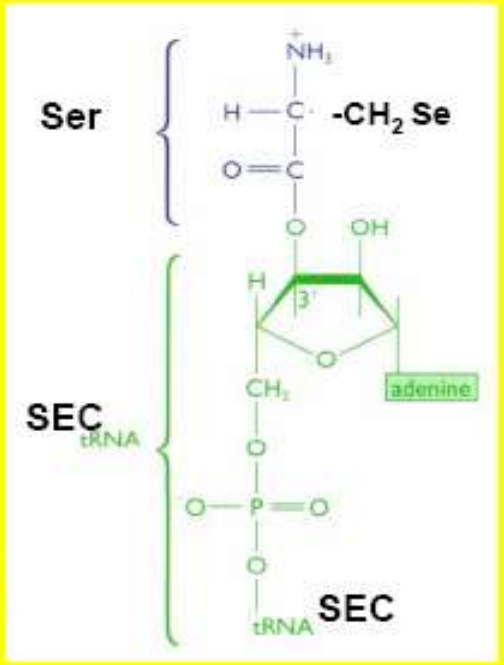
t-RNA ^{Ser} synthetases recognizes both t-RNA^{Sec} and T-RNA ^{Ser}

$Se\text{-}PO_3^{-2}$

Sel-A*

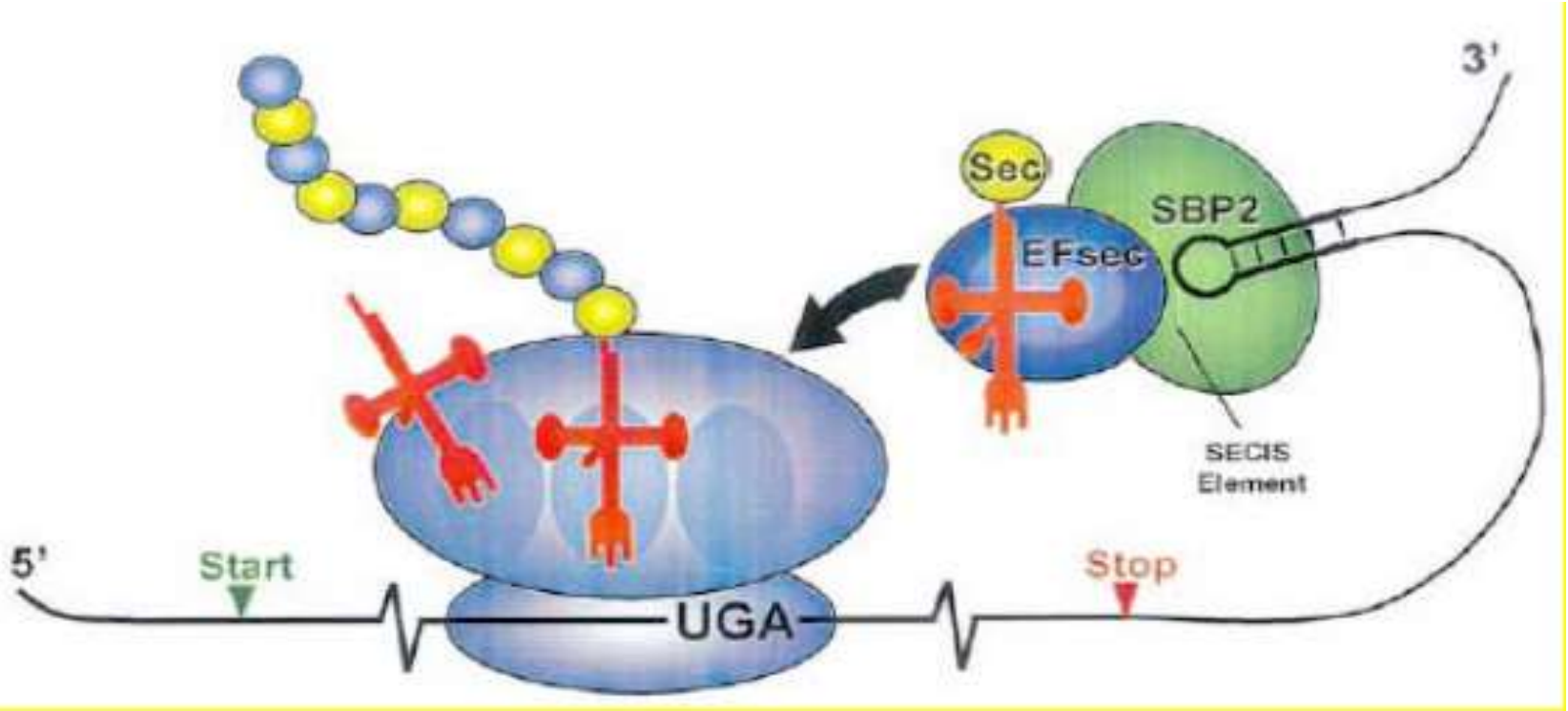
Step 3

*Mammalian protein unidentified



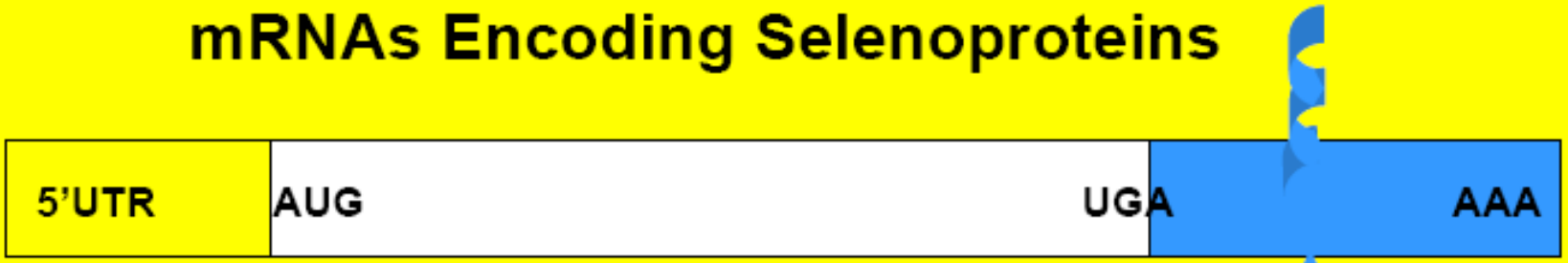
Sec t-RNA ^{Sec}; longest t-RNA known-90 nt

Incorporation of selenocysteine in polypeptide



Structure of SECIS elements

mRNAs Encoding Selenoproteins



SECIS element in the 3'UTR

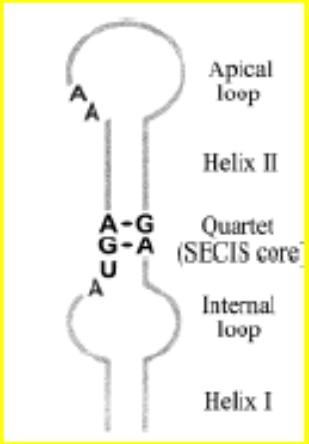
Position: Up to 4000 nt from the UGA

Secondary structure: Stem loop

- 50- 60 nucleotides, mostly helix
- AA in loop/AGUA and GA helix base
- 3D structure unknown

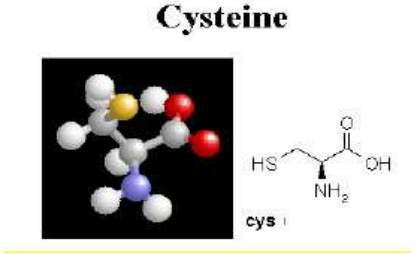
Each protein has distinctive mRNA 3'UTR SECIS

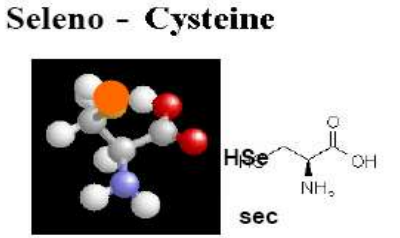
- Conserved-specific SECIS sequence
- Specific 3'UTR location
- Different UGA contexts

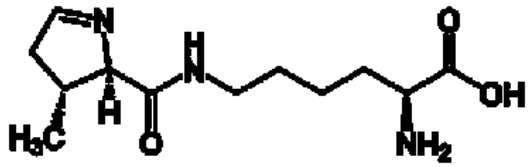


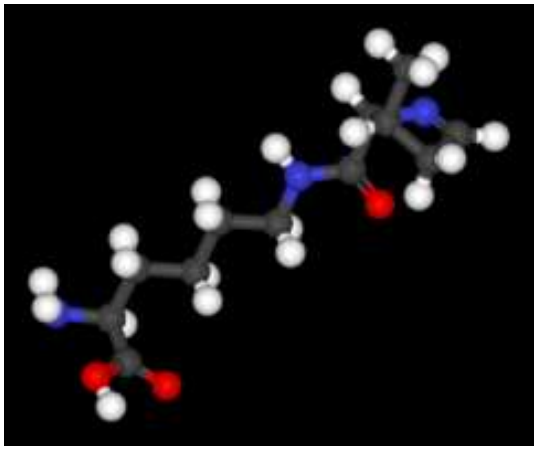
Alternative reading of stop codons

- UGA – stop
- UGA – selenocysteine
  - UGU - cysteine
  - UGC - cysteine
- UAG - stop
- UAG - pyrrolysine
  - only in Bacteria and Archea

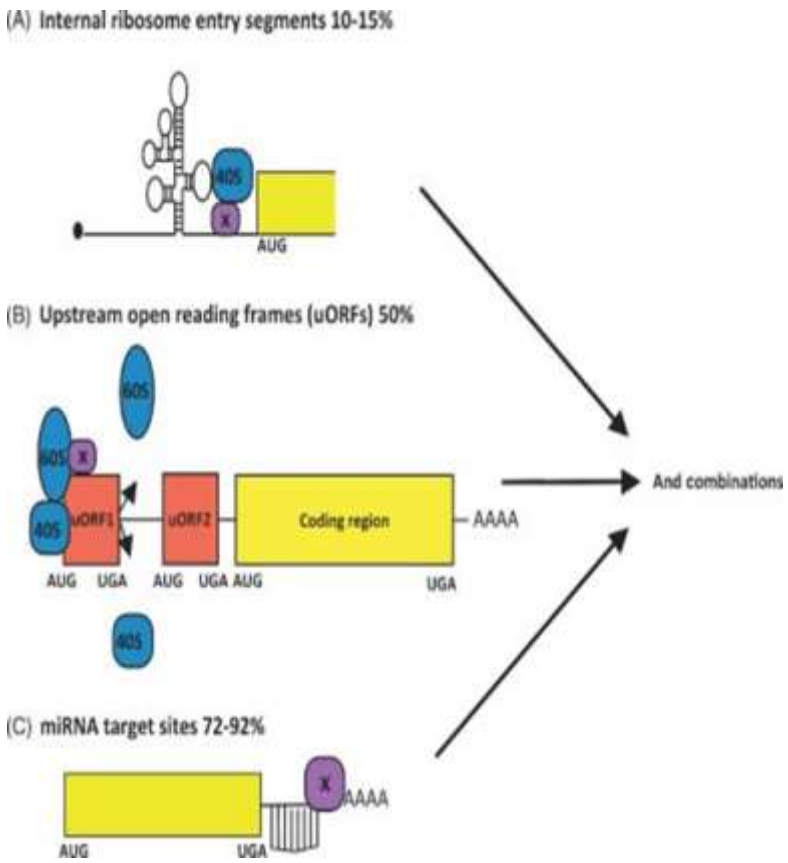








mRNA- specific regulation of translation



- Most examples of message-specific regulation are dependent upon **sequence elements**, which may or may not be structured, **within the 5' and 3' UTRs**.

- These are

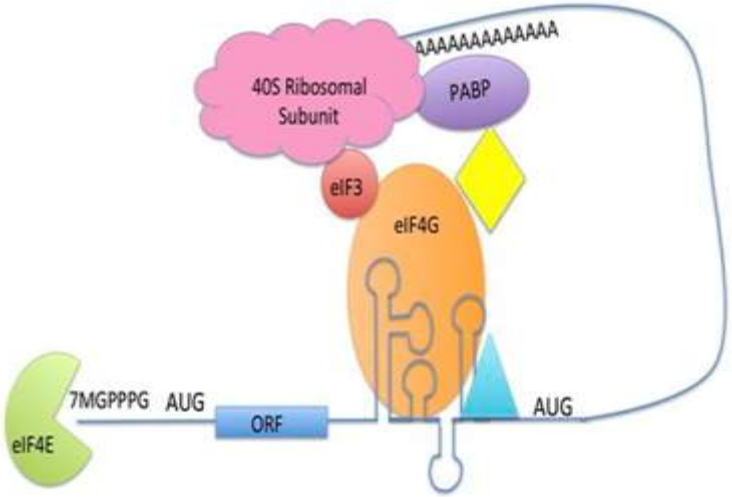
  - ✓ **I**nternal **R**ibosome **E**ntry **S**egments**- IRESs**,

  - ✓ **U**pstream **O**pen **R**eading **F**rames **– uORFs** and

Internal ribosome entry segments (IRES)

- **IRESs** are typically **highly structured RNA elements in the 5' UTR** that allow **binding of ribosomes** at or near the AUG start codon**, independent of cap recognition**.

- Have been shown to activate or maintain translation following a range of cellular stresses that compromise the cap-binding complex.



- They are assisted by proteins, **IRES *trans*-acting factors (ITAFs),** which bind to the IRES and modify its structure and/or interact with other elements of the translation machinery.

- It is estimated that 10% of mRNAs contain IRES elements within their 5' UTRs.

Post-translational modifications

- During translation, about 30-40 polypeptide residues are relatively protected by the ribosome. Once the polypeptide chain emerges from the ribosome it starts to fold and can be subject to post-translational modifications.

Why post-translational processing?

- adds functionality
- effects targeting
- regulates activity
- increases mechanical strength

Post-translational modifications

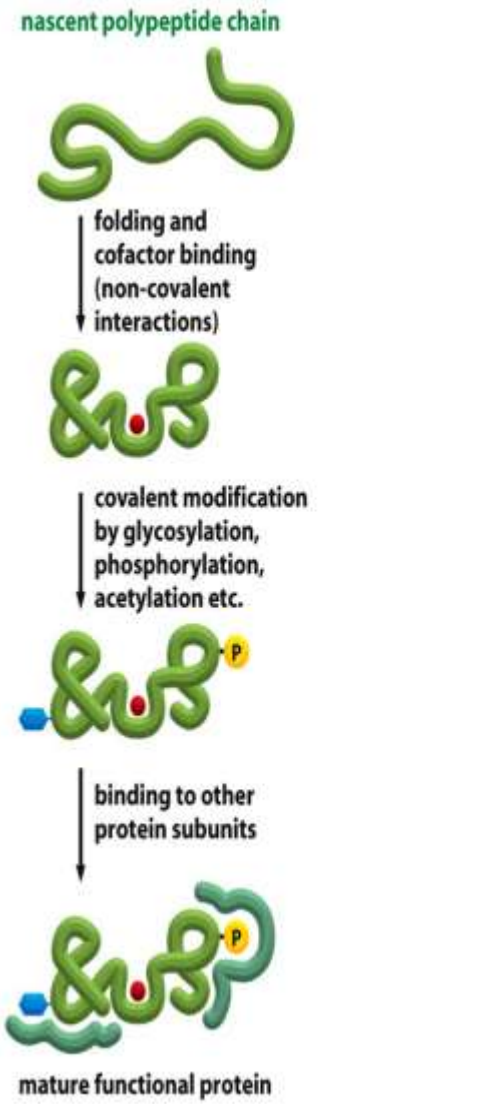


Figure 6-82 Molecular Biology of the Cell 5/e (© Garland Science 2008)

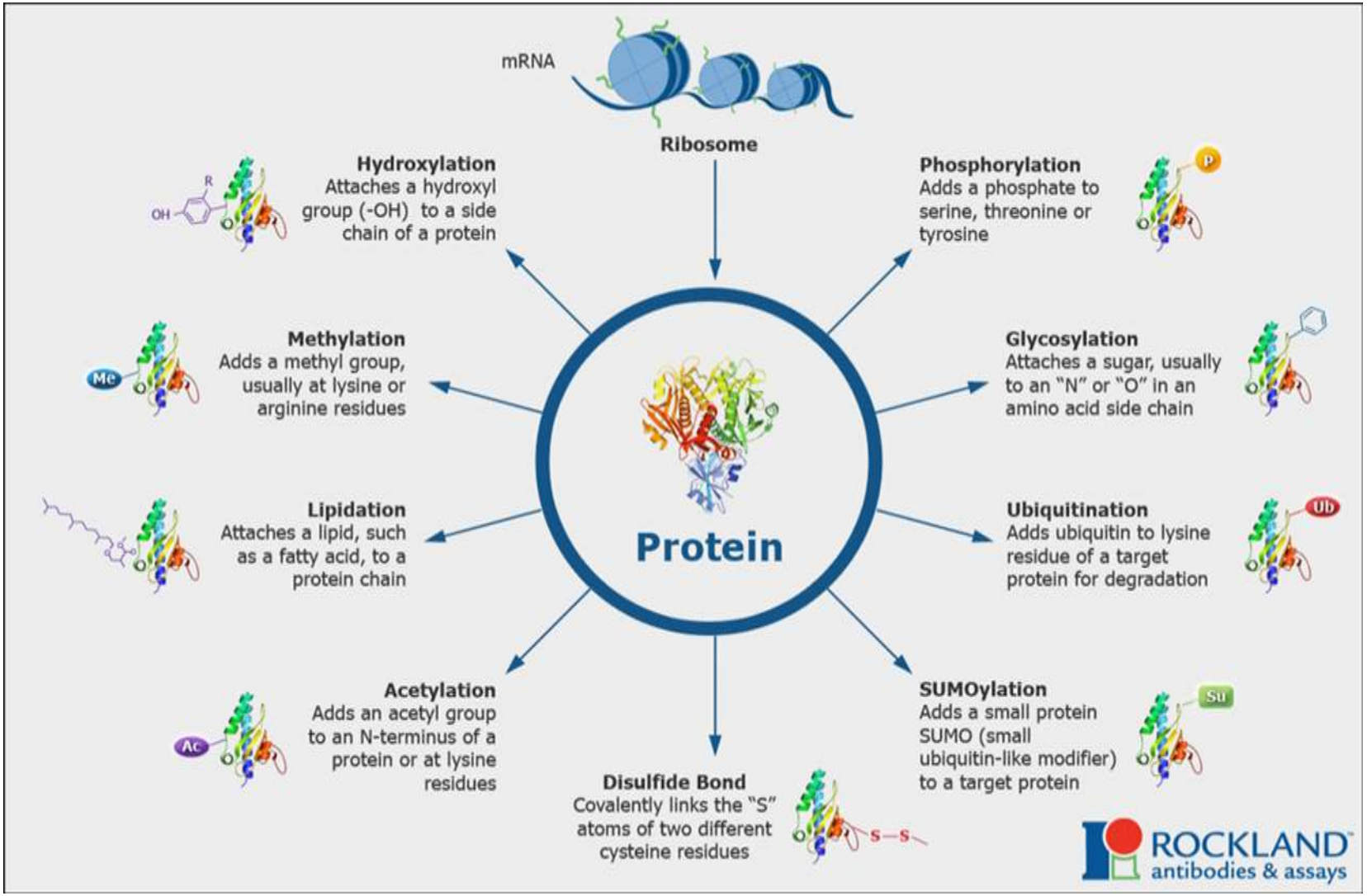
The translation of an mRNA sequence into an amino acid sequence on the ribosome is not the end of the process of forming a protein.

To function, the completed polypeptide chain must fold correctly into its three-dimensional conformation, bind any cofactors required.

Many proteins also require covalent modifications of selected amino acids.

Although the most frequent modifications are protein glycosylation and protein phosphorylation, more than 100 different types of covalent modifications are known.

Post-translational modifications



Post-translational modifications

- Covalent modification of
  - a: peptide bonds
  - b: the N-terminus
  - c: the C-terminus
  - d: amino acid residues (side chains).
- Noncovalent modifications: folding, addition of co-factors.
- Translocation: compartment selection and transport (Trafficking/Targeting).
- Involvement of molecular chaperones

Modifications involving the peptide bond (*peptide bond cleavage* or *limited proteolysis*):

usually carried out by enzymes called ***peptidases*** or ***proteases:***

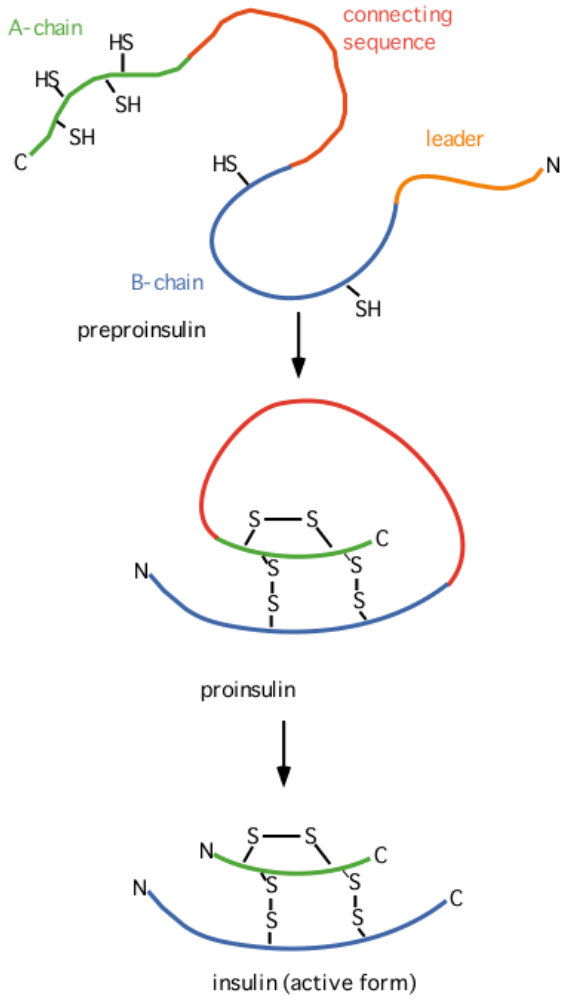
activation of proenzymes (digestive enzymes, blood clotting cascade, complement activation etc.) and prohormones (insulin)

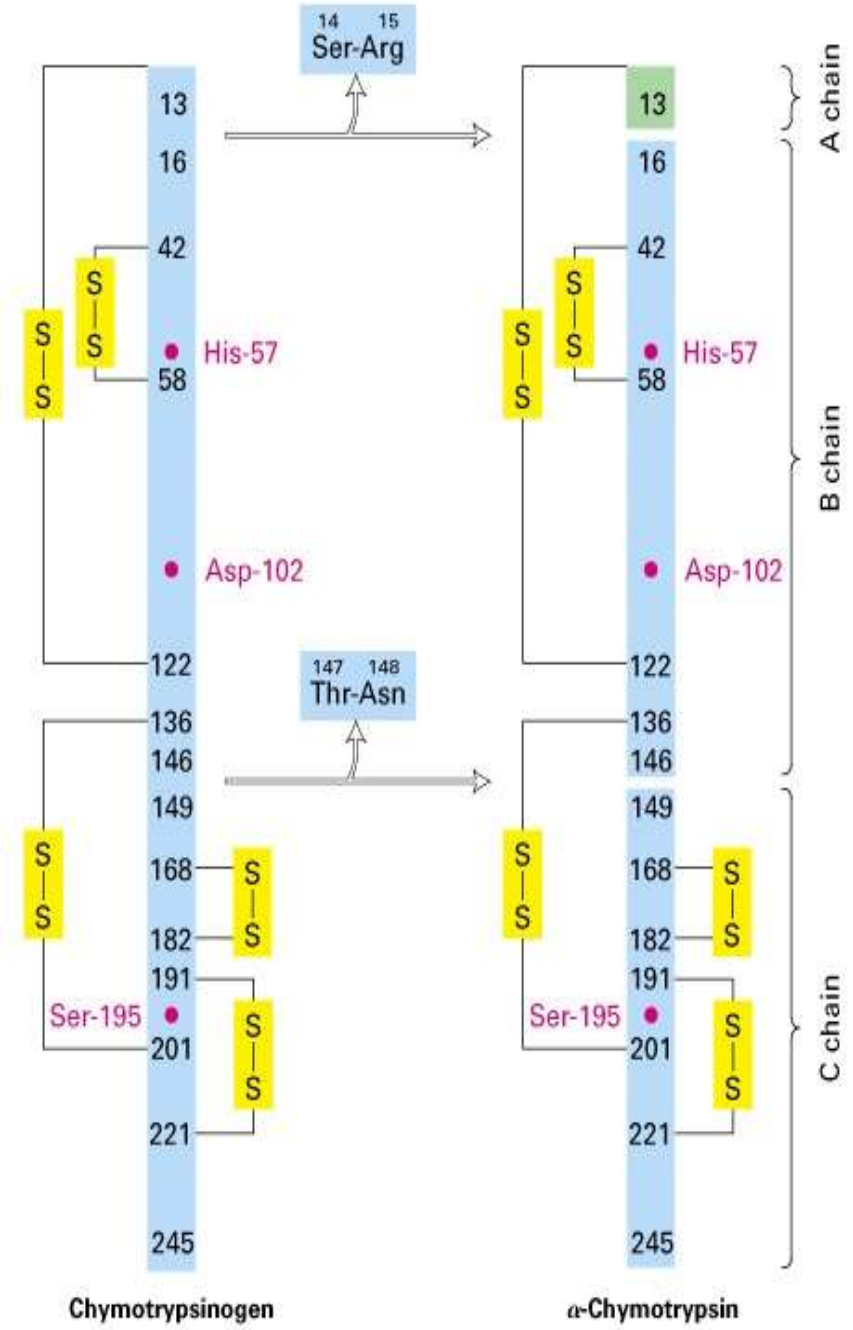
production of active neuropeptides and peptide hormones from high molecular weight precursors

macromolecular assembly in virus particles (e.g. HIV protease)

removal of signal sequences

Covalent modifications





Covalent modification of proteins

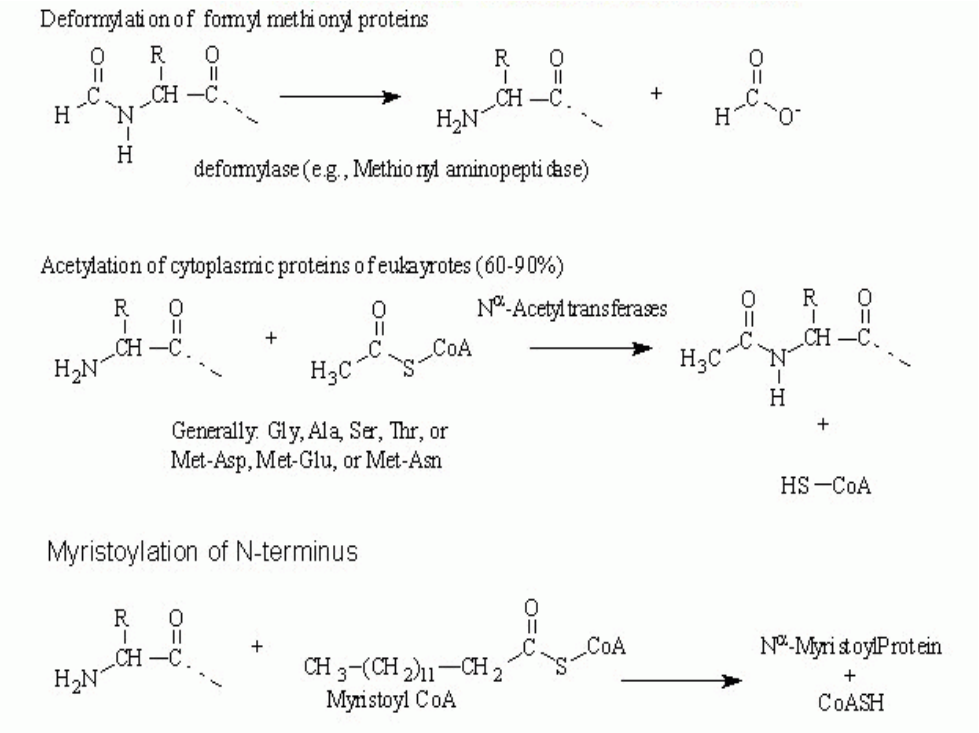
Modifications involving the amino terminus:

trimming of formyl group from formyl-Met

proteolytic removal of N-terminal Met by aminopeptidases

acetylation

lipidation (myristoyl

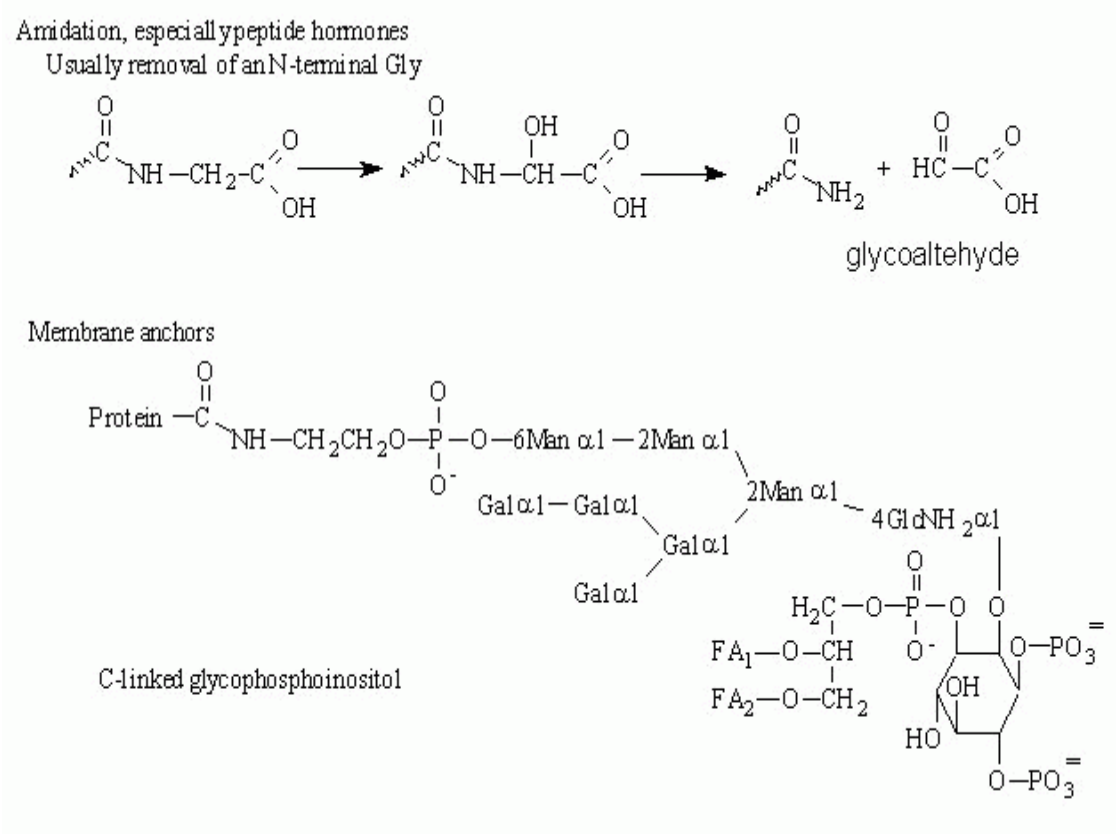


Covalent modification of proteins

Modifications involving the carboxy terminus:

amidation of C-terminal glycine

attachment of membrane anchors



Covalent modification of proteins

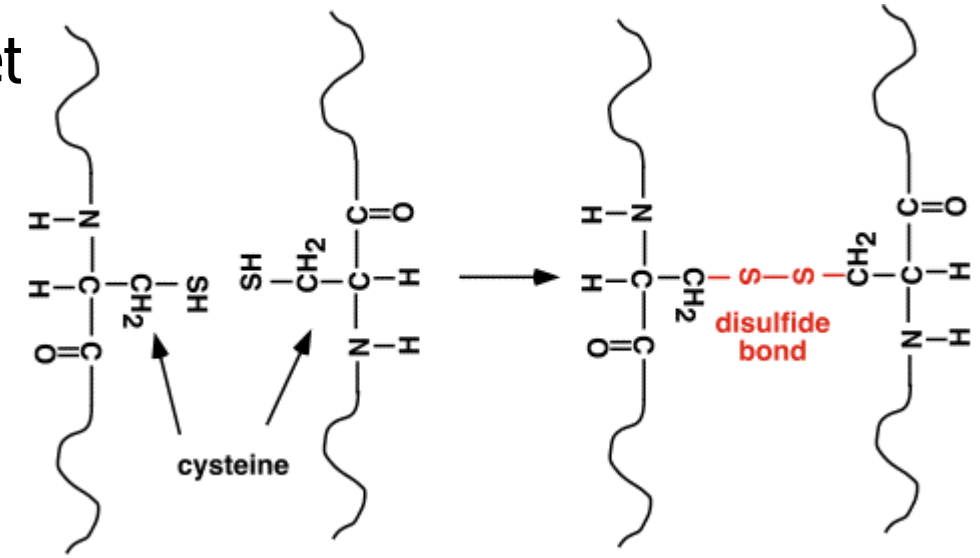
Modifications involving amino acid side chains:

Disulfide cross-linking

Phosphorylation of hydroxyls by kinases (serine, threonine, tyrosine)

Glycosylation

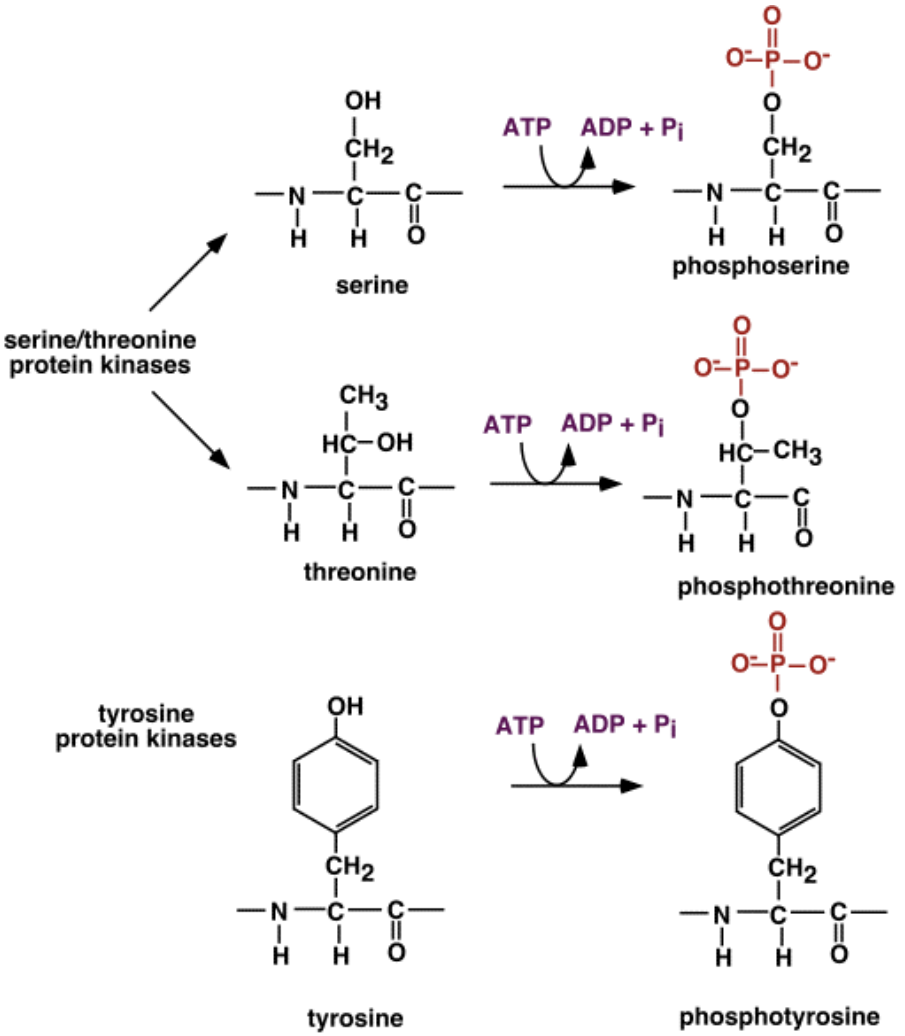
Prosthet al etc.)



Protein fosforylation

control activity

Enzymes that add a phosphate to a hydroxyl side chain are commonly called kinases. Enzymes that remove a phosphate from a phosphorylated side chain are called phosphatases.



Protein glycosylation
(mainly eukaryotes !!)

There are two basic types of glycosylation which occur on:

asparagines (N-linked)

serines and threonines (O-linked)

Covalently attached to the polypeptide as oligosaccharide chains containing 4 to 15 sugars

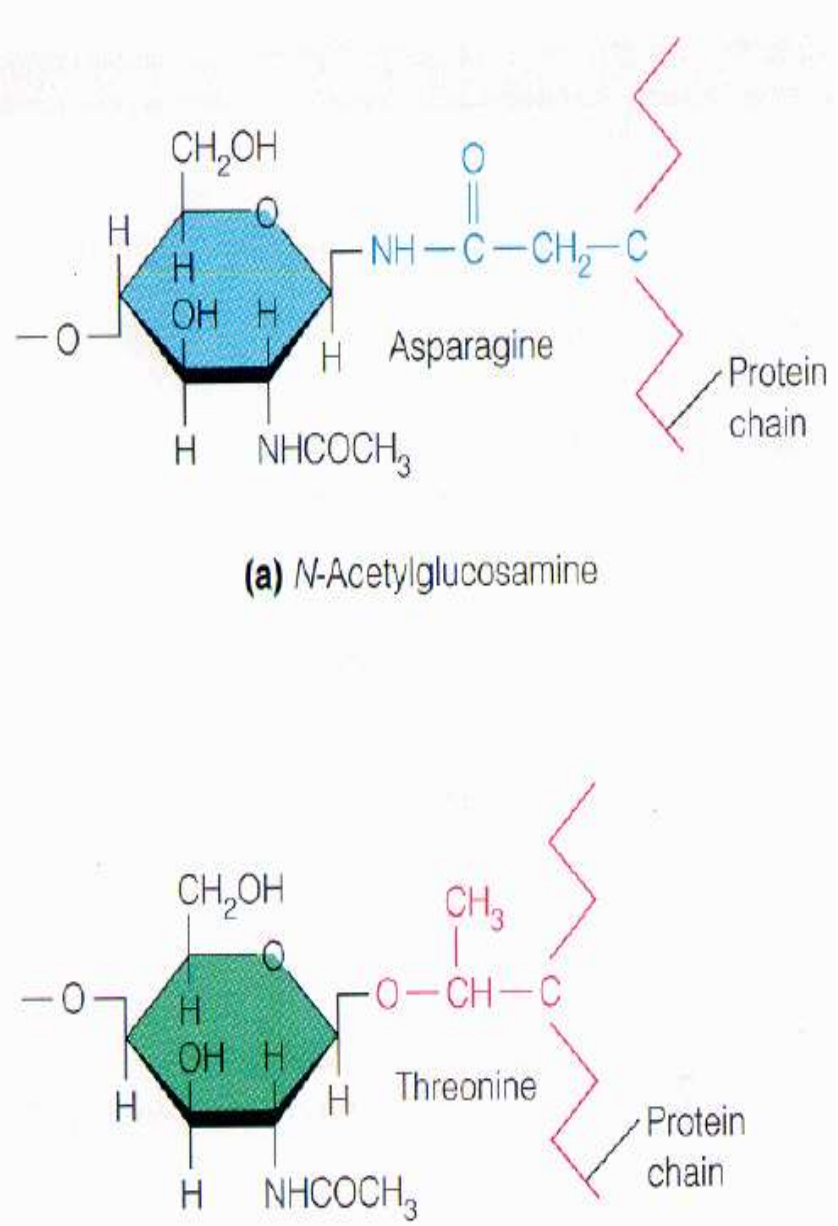
Sugars frequently comprise 50% or more of the total molecular weight of a glycoprotein

Most glycosylated proteins are either secreted or remain membrane-bound

Glycosylation is the most abundant form of post-translational modification

Glycosylation confers resistance to protease digestion by steric protection

Important in cell-cell recognition



Non-covalent modification of proteins

Addition of metal ions and co-factors

Nearly 50% of all proteins contain metal ions

Metal ions play regulatory as well as structural roles

Modifications involving tertiary structure (protein fold)

Enzymes called molecular chaperones are responsible for detecting mis-folded proteins.

Chaperones only bind mis-folded proteins that exhibit large hydrophobic patches on their surfaces.

Subunit multimerization

Many enzymes are only functional as multimeric units, either as homo- or hetero-oligomers. Example: ribosomes!

Protein folding - chaperons

Chaperones:

Mediate folding and assembly.

Do not convey steric information.

Do not form part of the final structure.

Suppress non-productive interactions by binding to transiently exposed portions of the polypeptide chain.

First identified as heat shock proteins (Hsp).

Hsp expression is elevated when cells are grown at higher-than-normal temperatures.

Stabilize proteins during synthesis.

Assist in protein folding by binding and releasing unfolded/mis-folded proteins.

Use an ATP-dependent mechanism.

Major types of chaperones:

Hsp70 (cytoplasm, ER, chloroplasts, mitochondria)

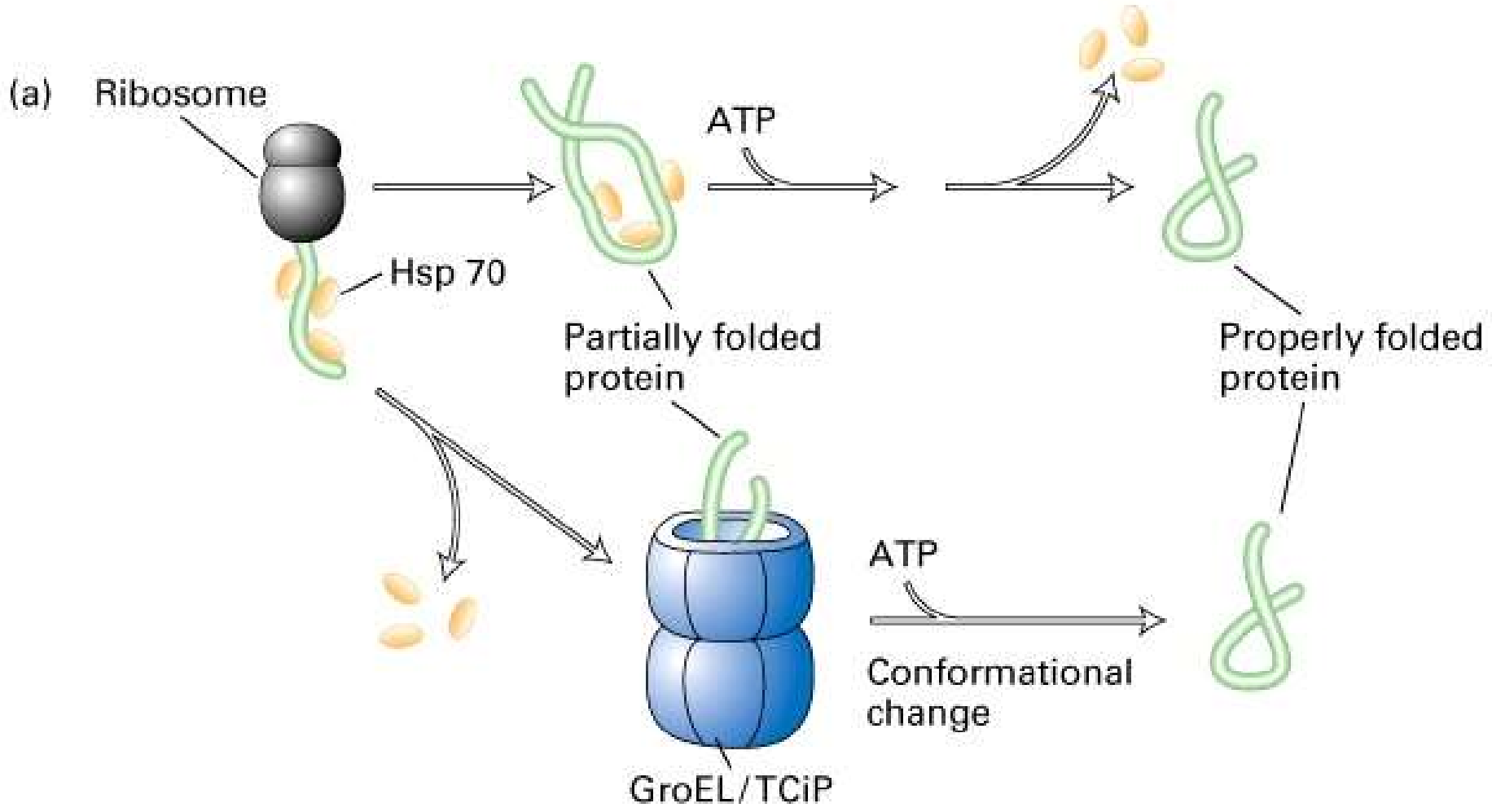
thought to bind and stabilize the nascent polypeptide chain as it is being extruded from the ribosome

also involved in "pulling" newly synthesized polypeptide into ER lumen.

Hsp60 (mitochondria, chloroplasts):

forms large 28-subunit complexes called GroEL

Protein folding - chaperons



Protein Degradation

Turnover of protein is NOT constant

Half lives of proteins vary from minutes to infinity

"Normal" proteins – 100-200 hrs

Short-lived proteins
regulatory proteins
enzymes that catalyze committed steps
transcription factors

Long-lived proteins
Special cases (structural proteins, crystallins)

Protein Degradation

- **May depend on tissue distribution**

  Example: Lactic Acid Dehydrogenase

Tissue	Half-life
Heart	1.6 days
Muscle	31 days
Liver	16 days

- **Protein degradation is a regulated process**

  Example: Acetyl CoA carboxylase

Nutritional state	Half-life
Fed	48 hours
Fasted	18 hours

Protein Degradation

- **Ubiquitin/Proteasome Pathway**

  80-90%

  Most intracellular proteins

- **Lysosomal processes**

  10-20%

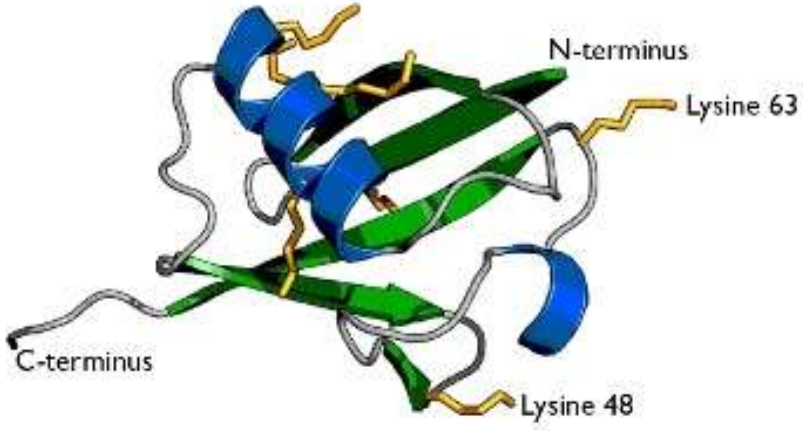
  Extracellular proteins

  Cell organelles

  Some intracellular proteins

UBIQUITIN

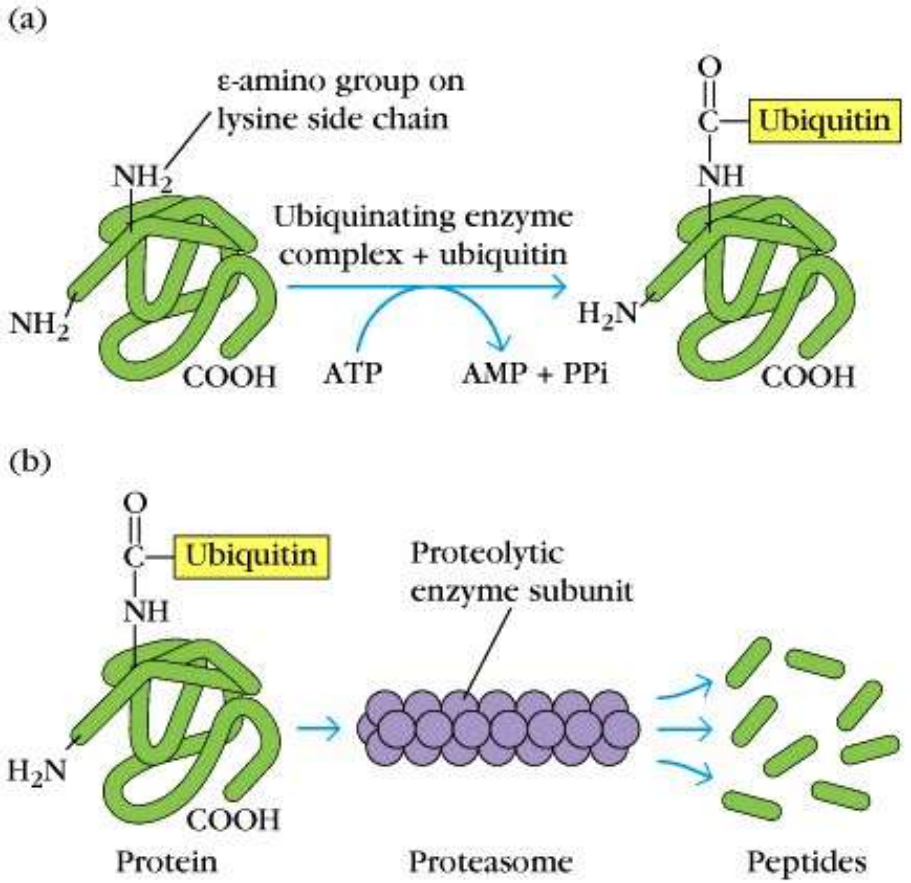
- Small peptide that is a “TAG”
- 76 amino acids
- C-terminal glycine - isopeptide bond with the e-amino group of lysine residues on the substrate
- Attached as monoubiquitin or polyubiquitin chains



Protein degradation – ubiquitin system

Covalent conjugation
to Ubiquitin

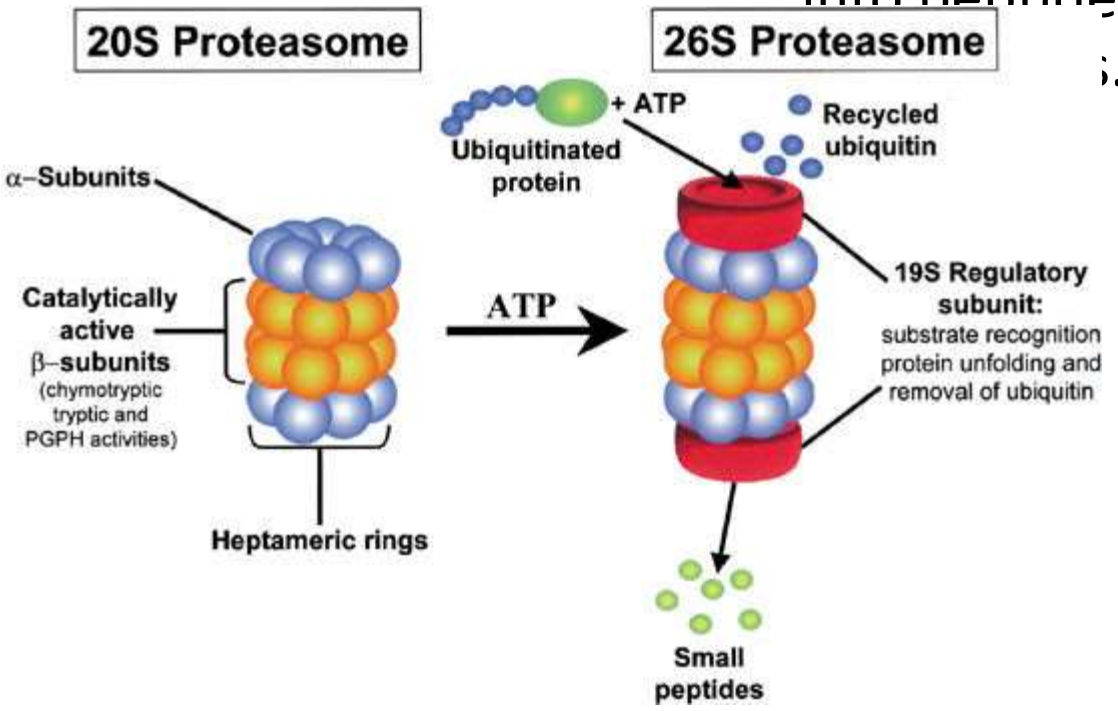
Ubiquitin
targets proteins
to Proteasome



Proteasome

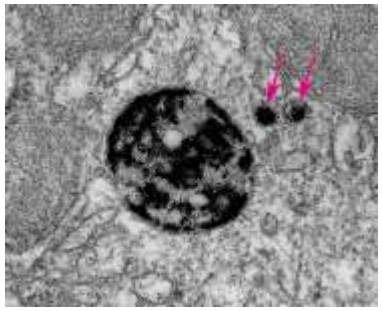
The **proteasome** is a cylindrical shaped catalytic protease complex of 28 subunits for cytosolic protein degradation.

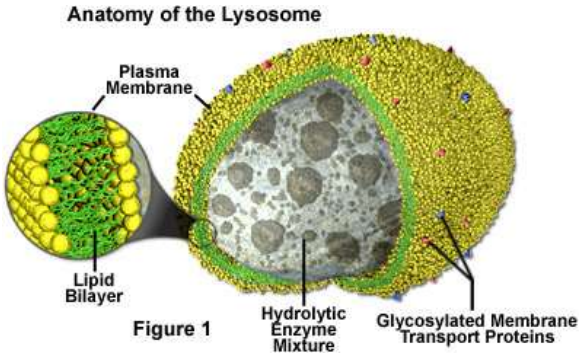
The proteasome unfolds proteins and then cleaves proteins into peptides and [illegible].



Conserved throughout the eukaryotes and the archaebacteria

Lysosomes degrade and recycle macromolecules including proteins





ECB 15-34

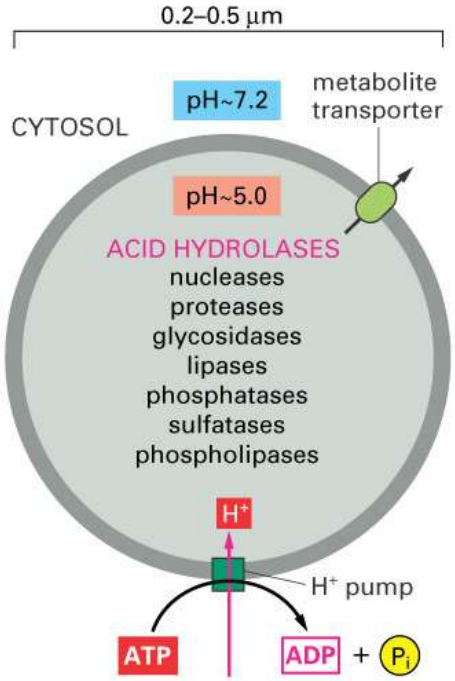


Figure 15-34 Essential Cell Biology, 2/e. (© 2004 Garland Science)

Lysosomes acid hydrolases