Translation

the base sequence of an mRNA is translated into the amino acid sequence in a protein.

1. THE GENETIC CODE

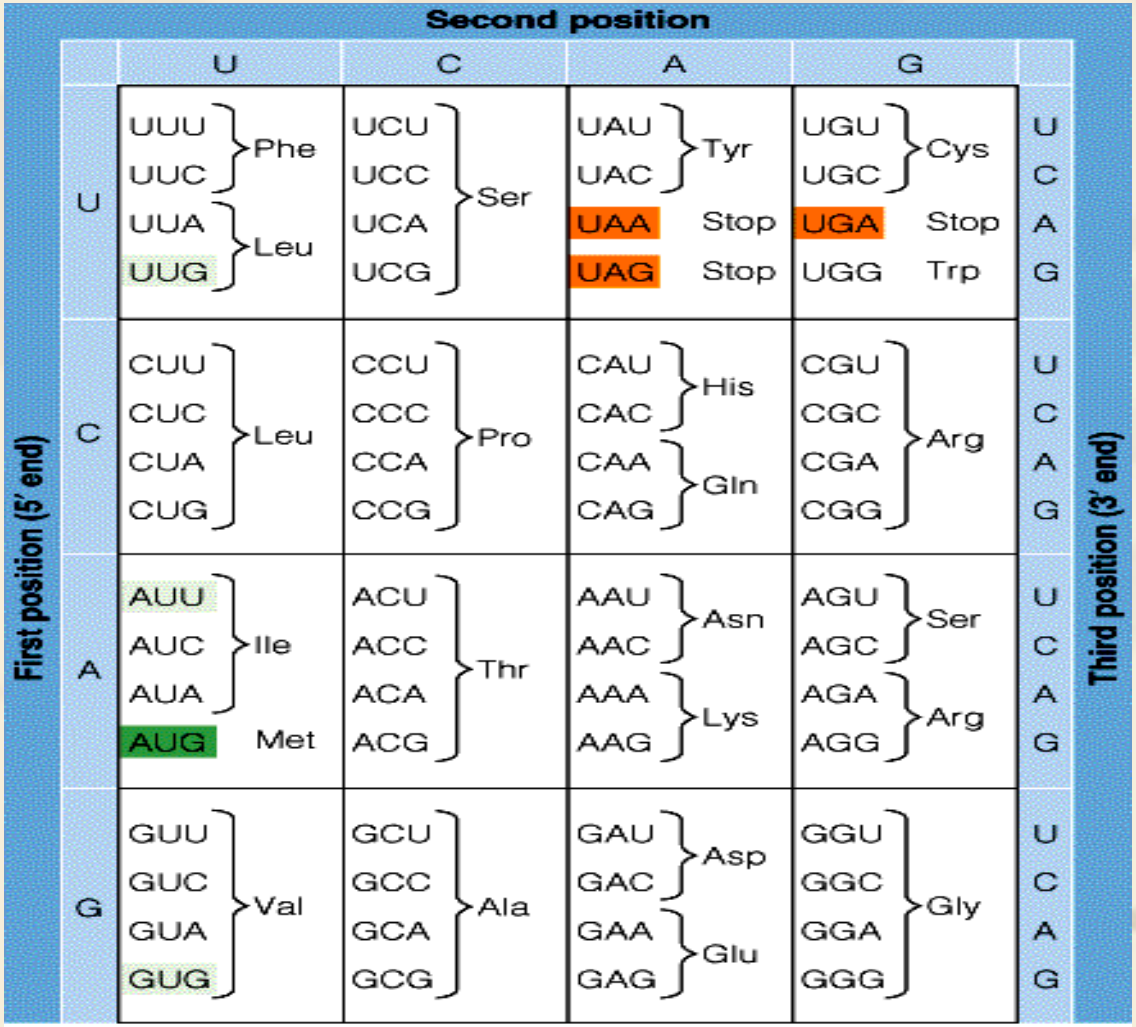
- **Genetic code is a triplet code** in mRNA *(Three nucleotides encode one amino acid)*
- The triplet codons are non-overlapping and comma-less.

---UCU UCC CGU GGU GAA---

- **Genetic code is degenerate :**

  Only 20 amino acids are encoded by 4 nucleotides in triplet codons (4^3 =64 of amino acids could potentially be encoded). Therefore, more than one triplets are used to specify a amino acids, and the genetic codes are said to be degenerate, or to have redundancy.

The ‘Universal’ genetic code



First position (5′ end)	Second position: U	Second position: C	Second position: A	Second position: 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

ORFs and overlapping genes

Open reading frames (ORFs) are suspected coding regions starting with **ATG (GUG)** and end with **UAG**, **TAA** or **UGA**.

Small organisms (viruses and phages) have overlapping reading frames.

- **More than one start codons in a DNA sequence are used for translating different proteins.**
- **A way to maximize the coding capability of a given DNA sequence.**
- **Example of overlapping genes:**
- **Phage ϕX174 with 5386 bp (maximal 200 kDa), in fact, it encodes 11 proteins of 262 kDa in total)**

tRNA

tRNA are the adaptor molecules that deliver amino acids to the ribosome and decode the information in mRNA.

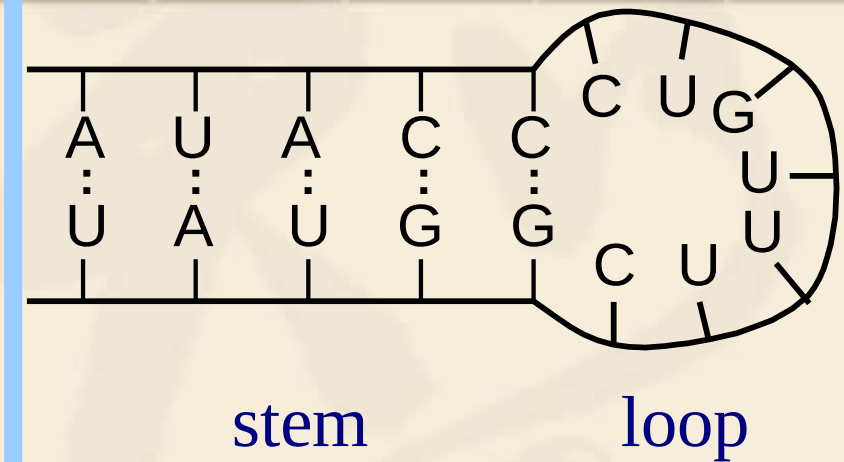
tRNA Structure:

Small, single-stranded nucleic acids ranging in size from 73-75 nucleotides, many modified bases: Di-hydroU; Pseudo-U: Y; modified G's; Thymine ...

The cloverleaf structure is a common secondary structural representation of tRNA molecules which shows the base paring of various regions to form four stems (arms) and three loops.

RNA structure:

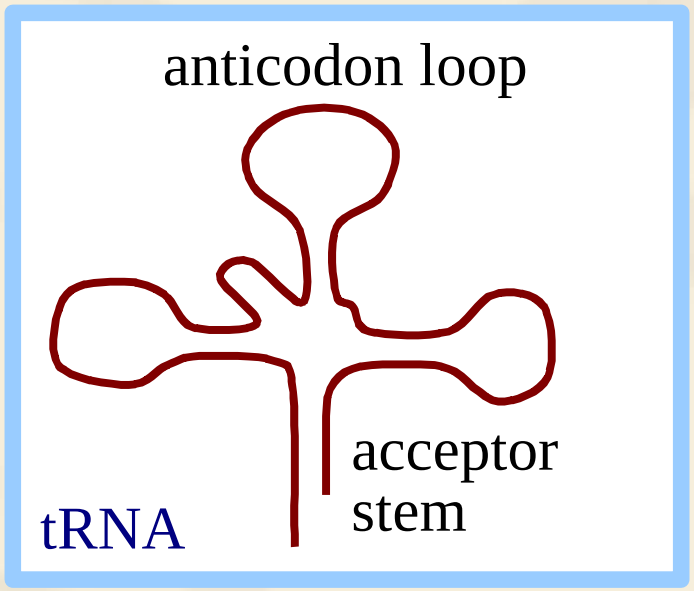
Most RNA molecules have **secondary structure**, consisting of **stem** & **loop** domains.



- **Double helical stem domains** arise from **base pairing** between complementary stretches of bases within the same strand.

  These stem structures are stabilized by stacking interactions as well as base pairing, as in DNA.

- **Loop domains** occur where **lack of complementarity** or the presence of **modified bases** prevents base pairing.



The “**cloverleaf**” model of **tRNA** emphasizes the two major types of secondary structure, **stems & loops**.

tRNAs typically include many **modified bases**, particularly in loop domains.

RNA tertiary structure

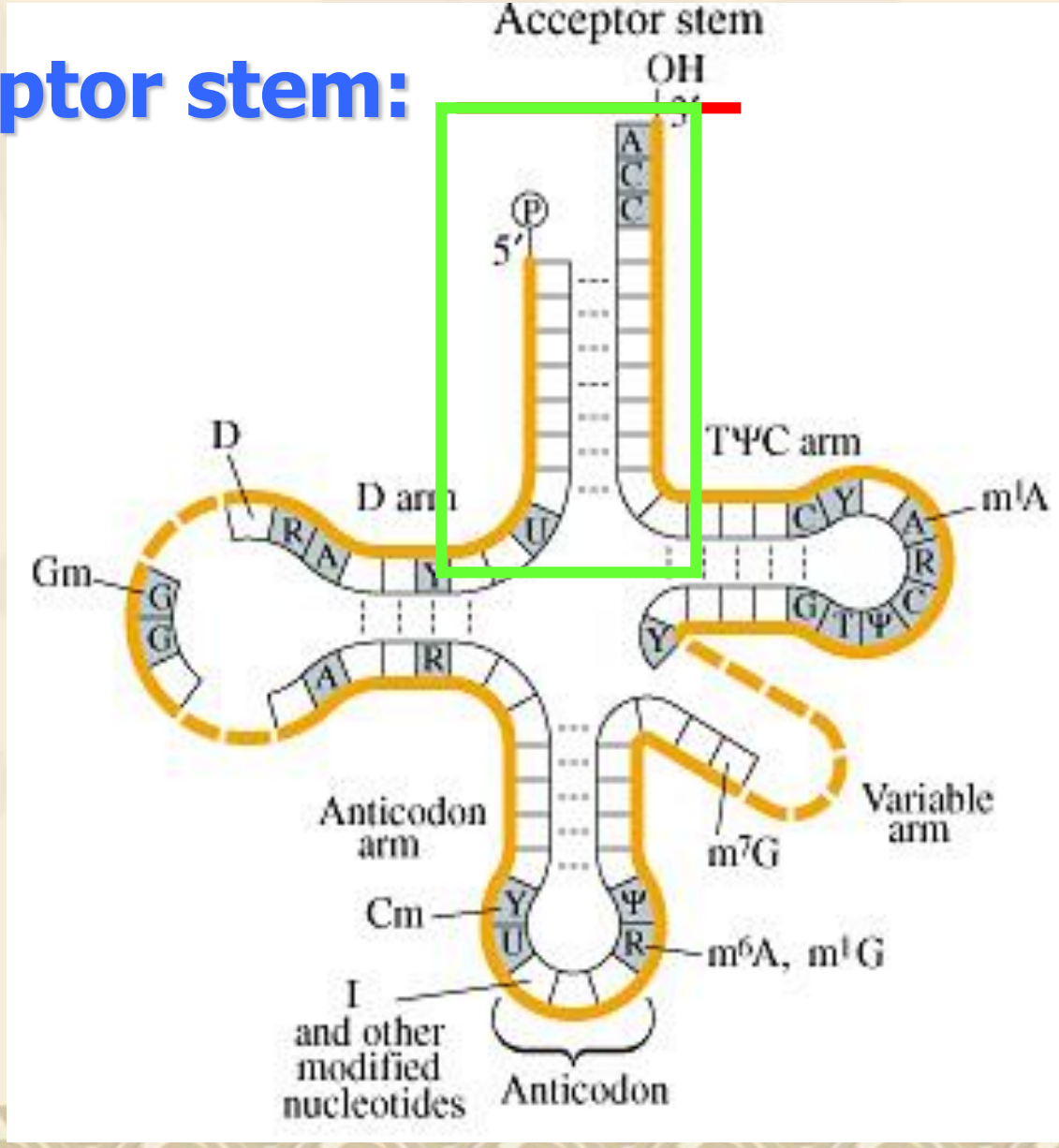
RNA tertiary structure depends on interactions of bases at **distant sites**.

- These interactions generally involve **non-standard** base pairing and/or interactions involving **three** or more bases.
- Unpaired **adenosines** (not involved in conventional base pairing) **predominate** in participating in non-standard interactions that stabilize tertiary RNA structures.

tRNAs have an **L-shaped** tertiary structure.

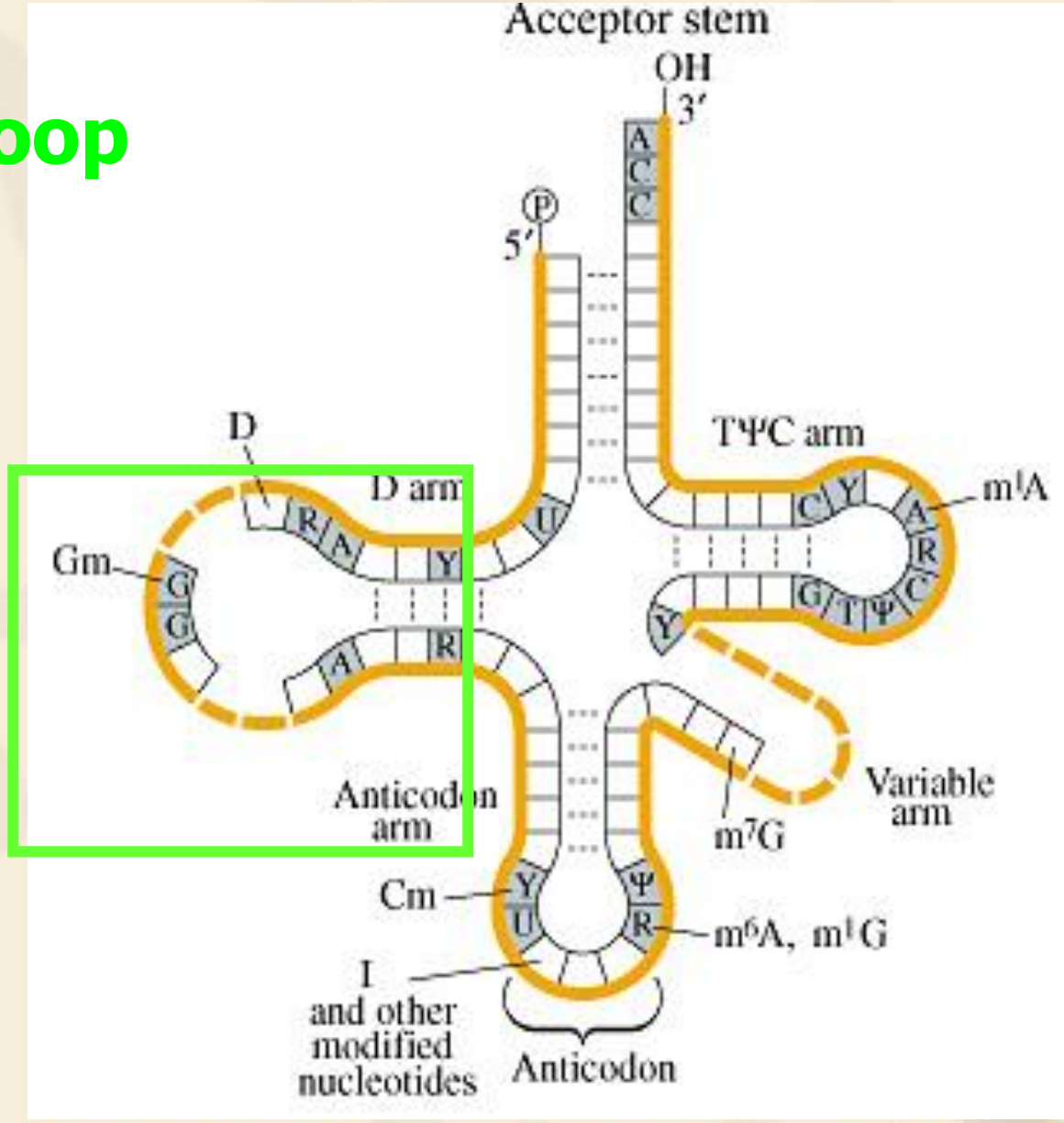
Amino acid acceptor stem:

- The 5’-and 3’-end are largely base-paired to form the amino acid acceptor stem which has no loop.
- Extending from the acceptor stem, the 3' end of each tRNA has the sequence CCA.
- The appropriate amino acid is attached to the ribose of the terminal adenosine (A, in red) at the 3' end.



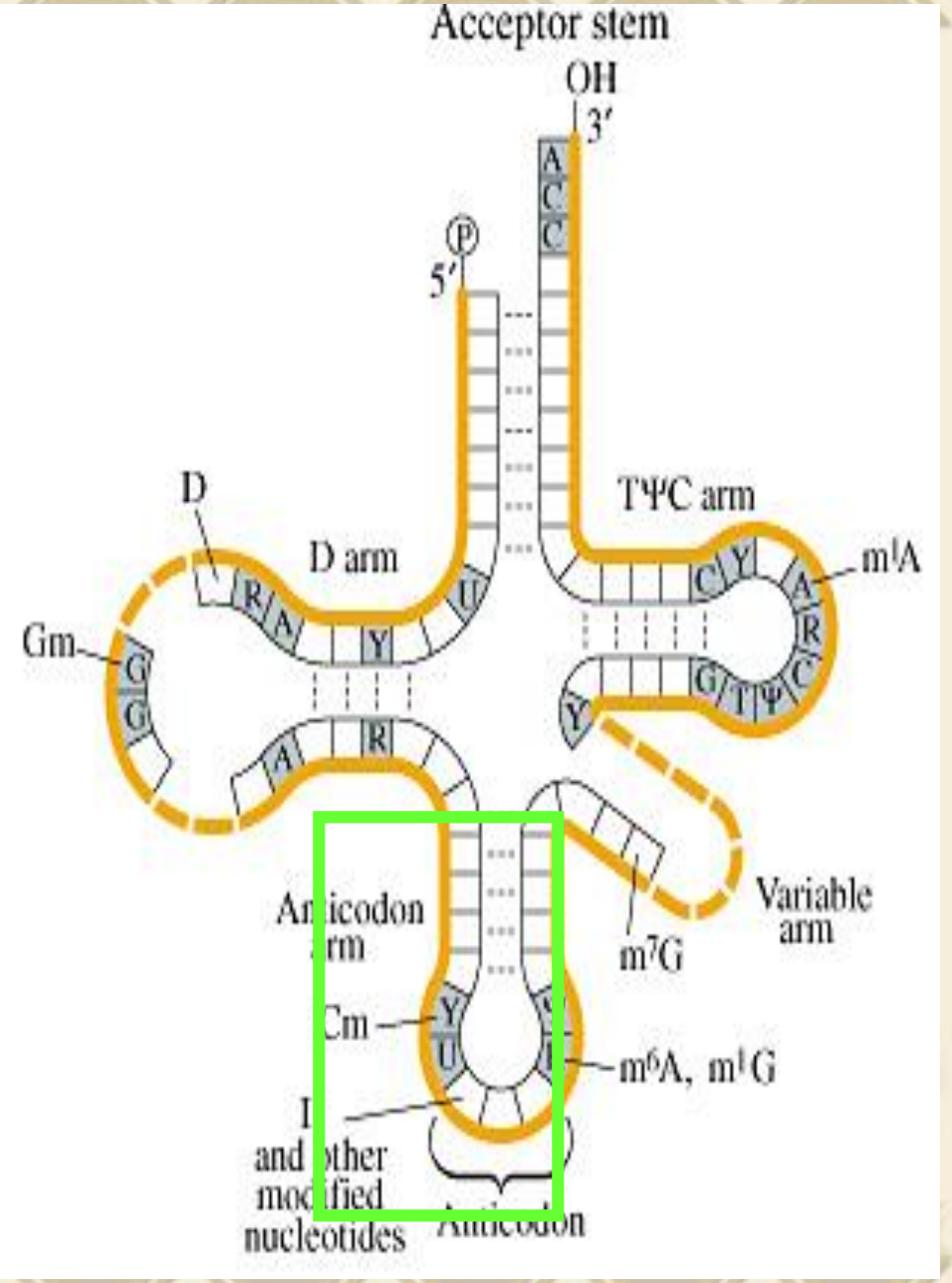
- **D-arm and D-loop**

Composed of 3 or 4 bp stem and a loop called the D-loop (DHU-loop) usually containing the modified base dihydrouracil.

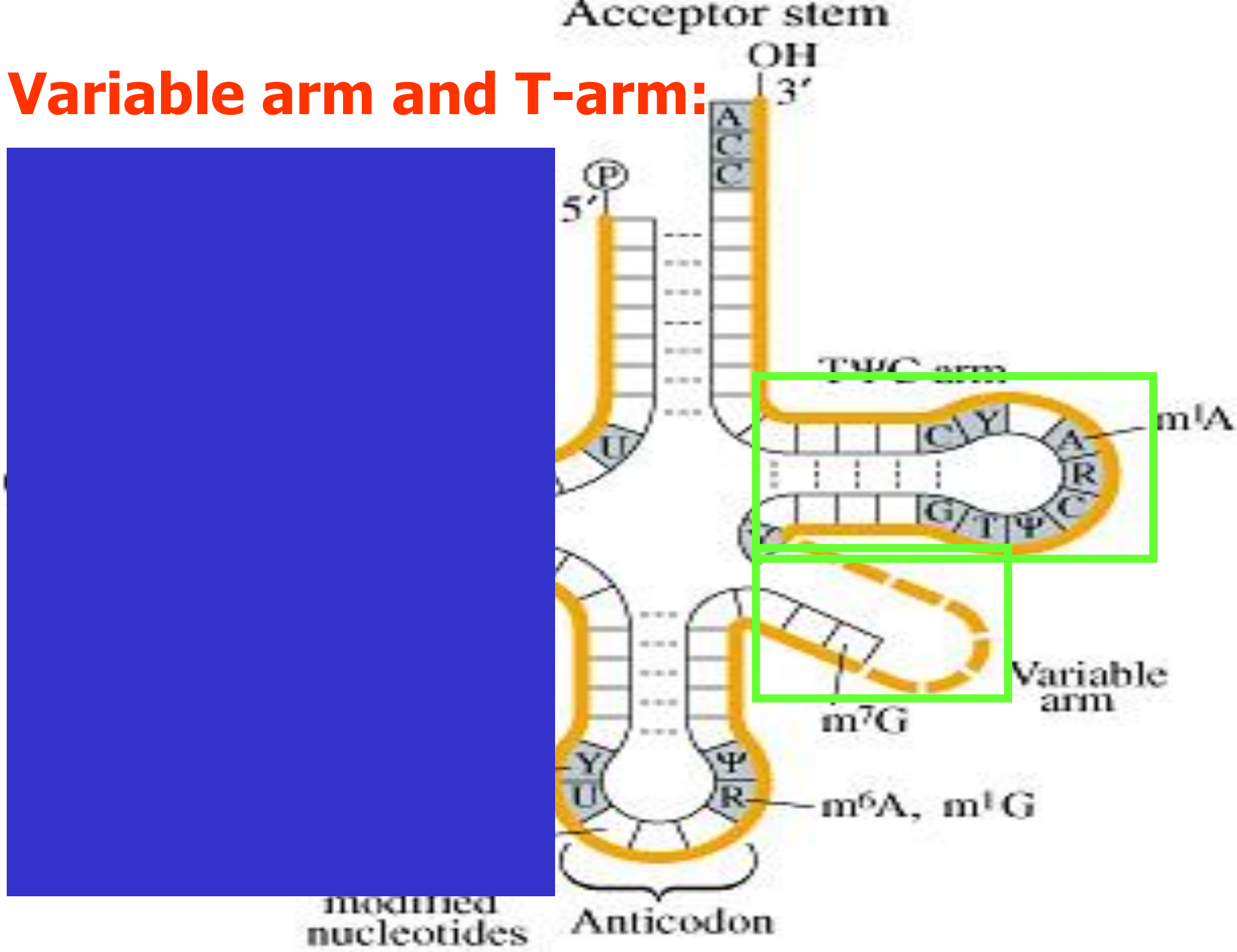


Anticodon loop:

Consisting of a 5 bp stem and a 7 residues loop in which there are three adjacent nucleosides called the anticodon which are complementary to the codon sequence (a triplet in the mRNA) that the tRNA recognize.



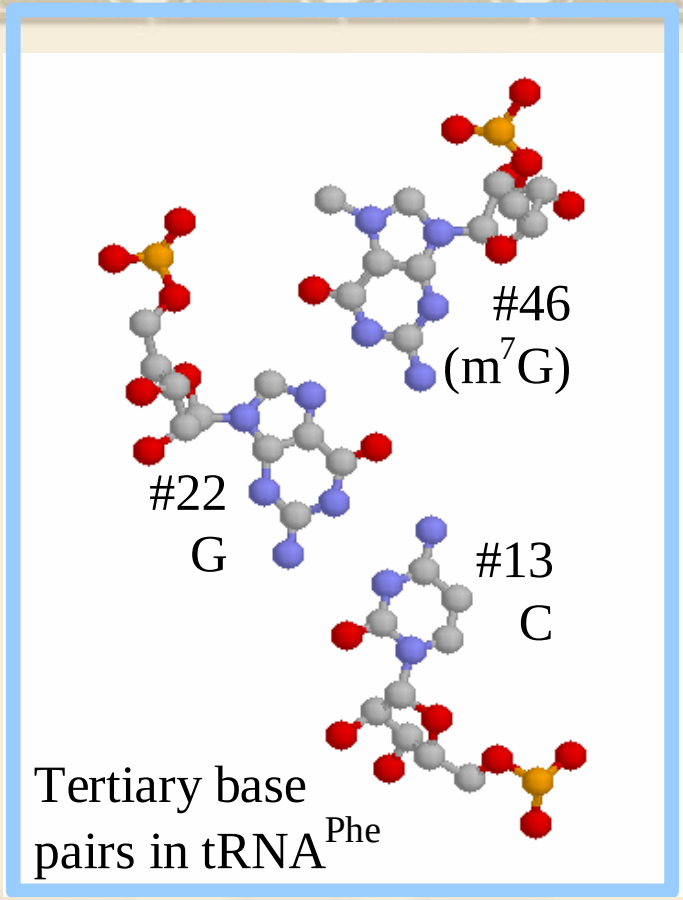
Variable arm and T-arm:

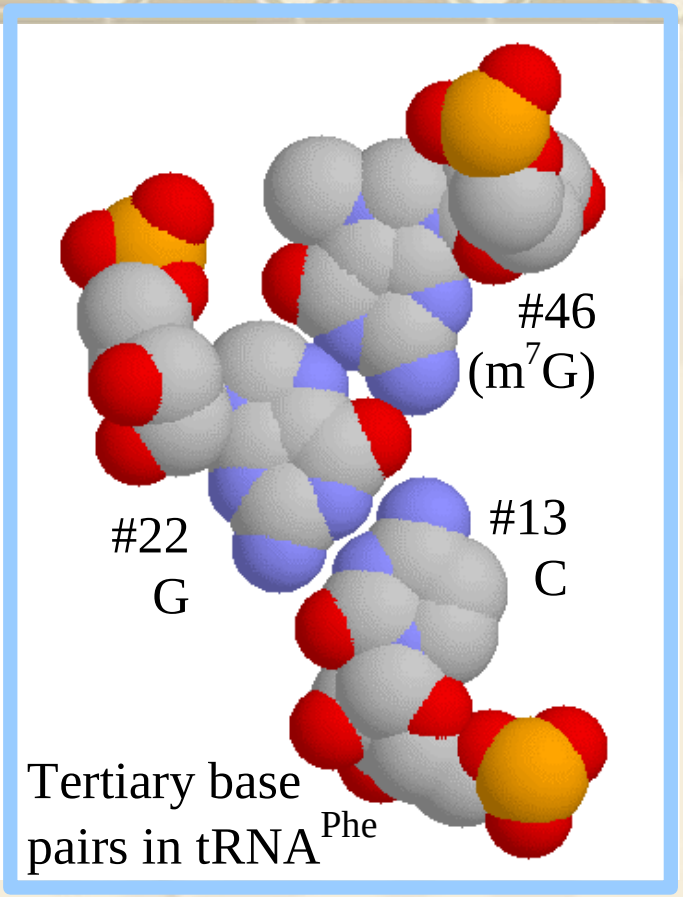


Function of each tRNA arm:

1) AminoAcyl or Acceptor arm: covalently bind the amino acid; interact with aminoacyl tRNA synthetases; provide substrate for peptidyl transferase in the large subunit ribosome.
2) D-arm (Dihydro arm): important in 3D structure of tRNA; interacts with large subunit ribosome in P and A sites

Tertiary structure: L-shape, with formation of additional H-bond interactions, some of which are nonstandard W-C base interactions .



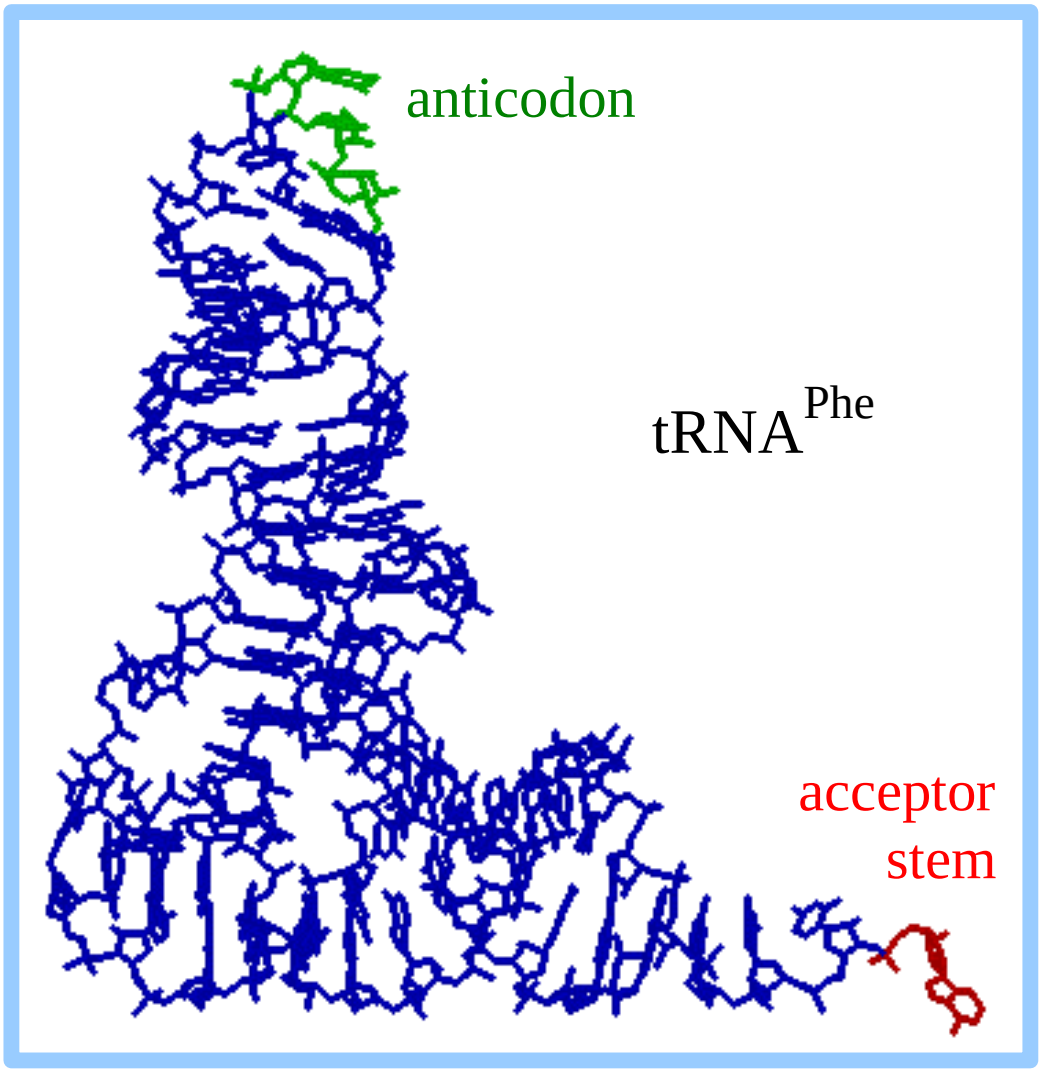


An example of **non-standard H bond interactions** that help to stabilize the L-shaped tertiary structure of a tRNA, in ball & stick & spacefill displays.

H atoms are not shown. (From NDB file 1TN2).

Some other RNAs, including viral RNAs & segments of ribosomal RNA, fold in **pseudoknots**, tertiary structures that mimic the 3D structure of tRNA.

Pseudoknots are similarly stabilized by non-standard H-bond interactions.

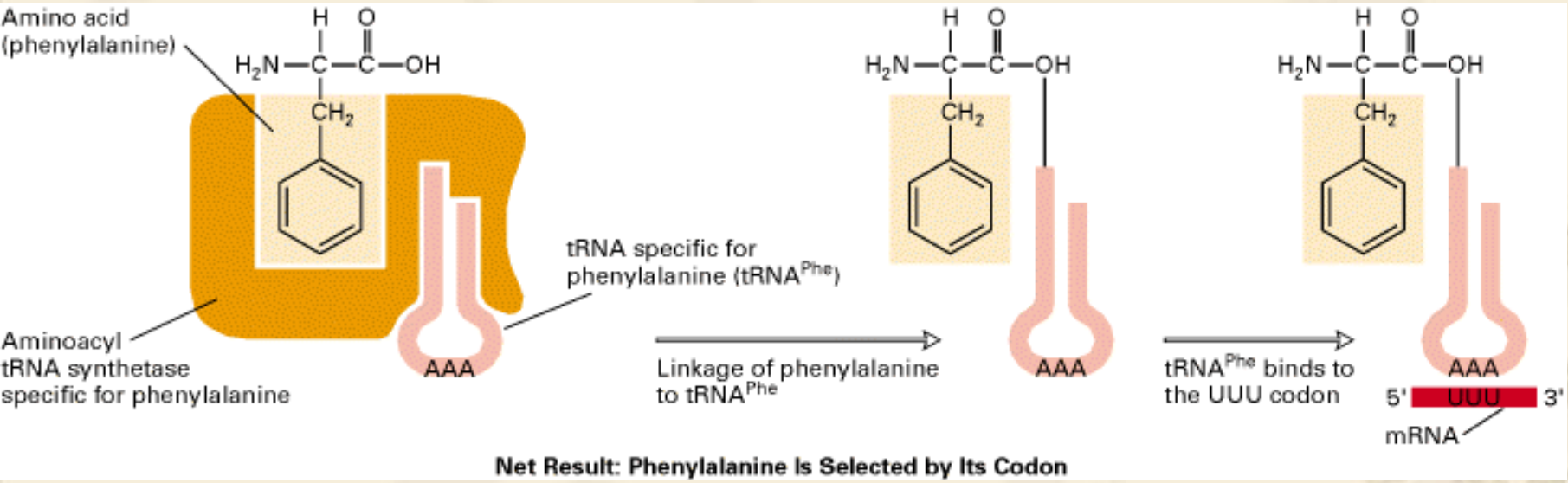


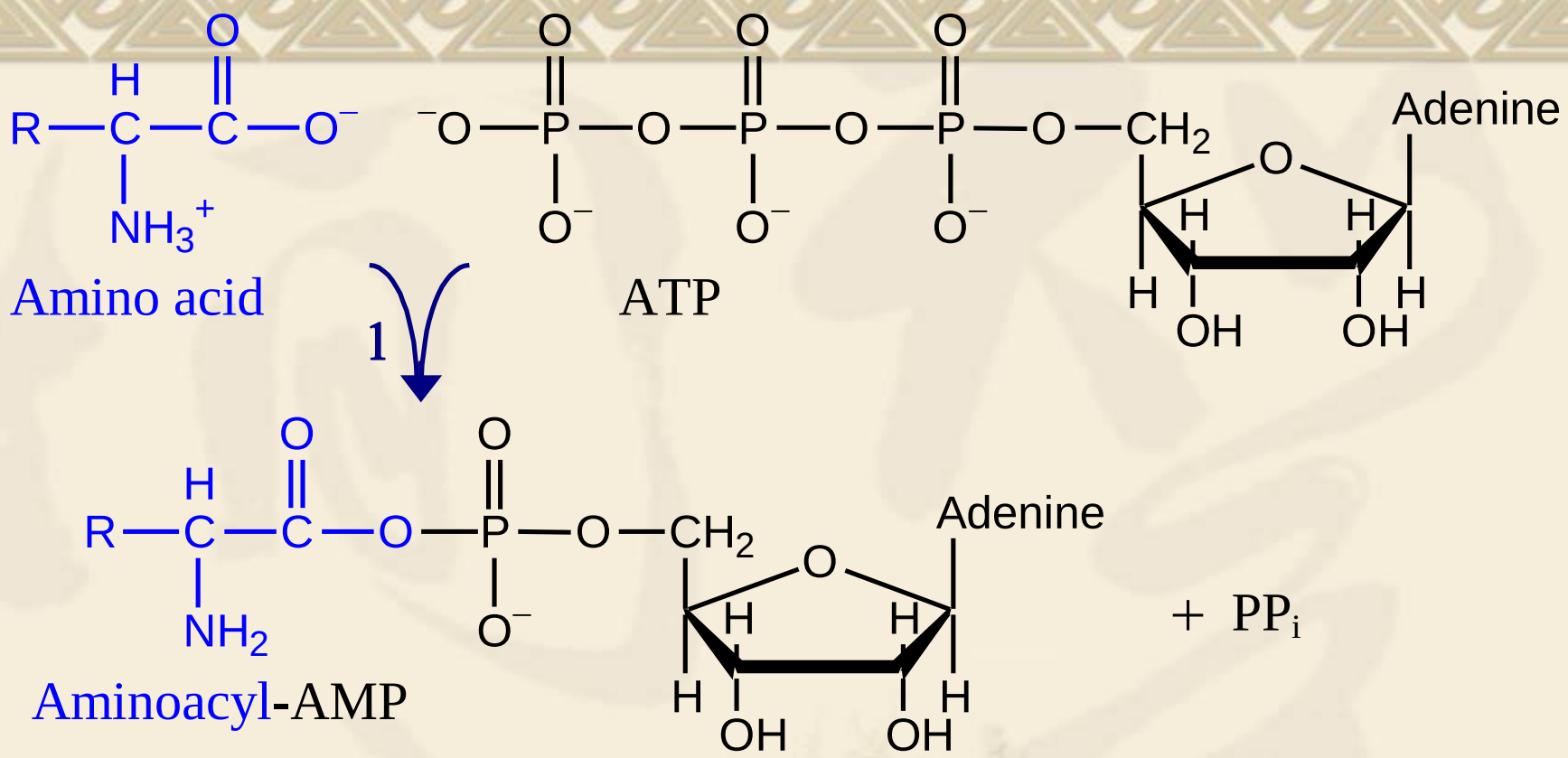
2.2. tRNA Function:

- bring amino acids to the ribosomes for protein synthesis via covalent linkage of amino acid to the tRNA: aminoacyl-tRNA species

2.3. Aminoacyl-tRNA Synthetases

- Each tRNA molecule is recognized by one and only one of the 20 aminoacyl-tRNA synthetases.



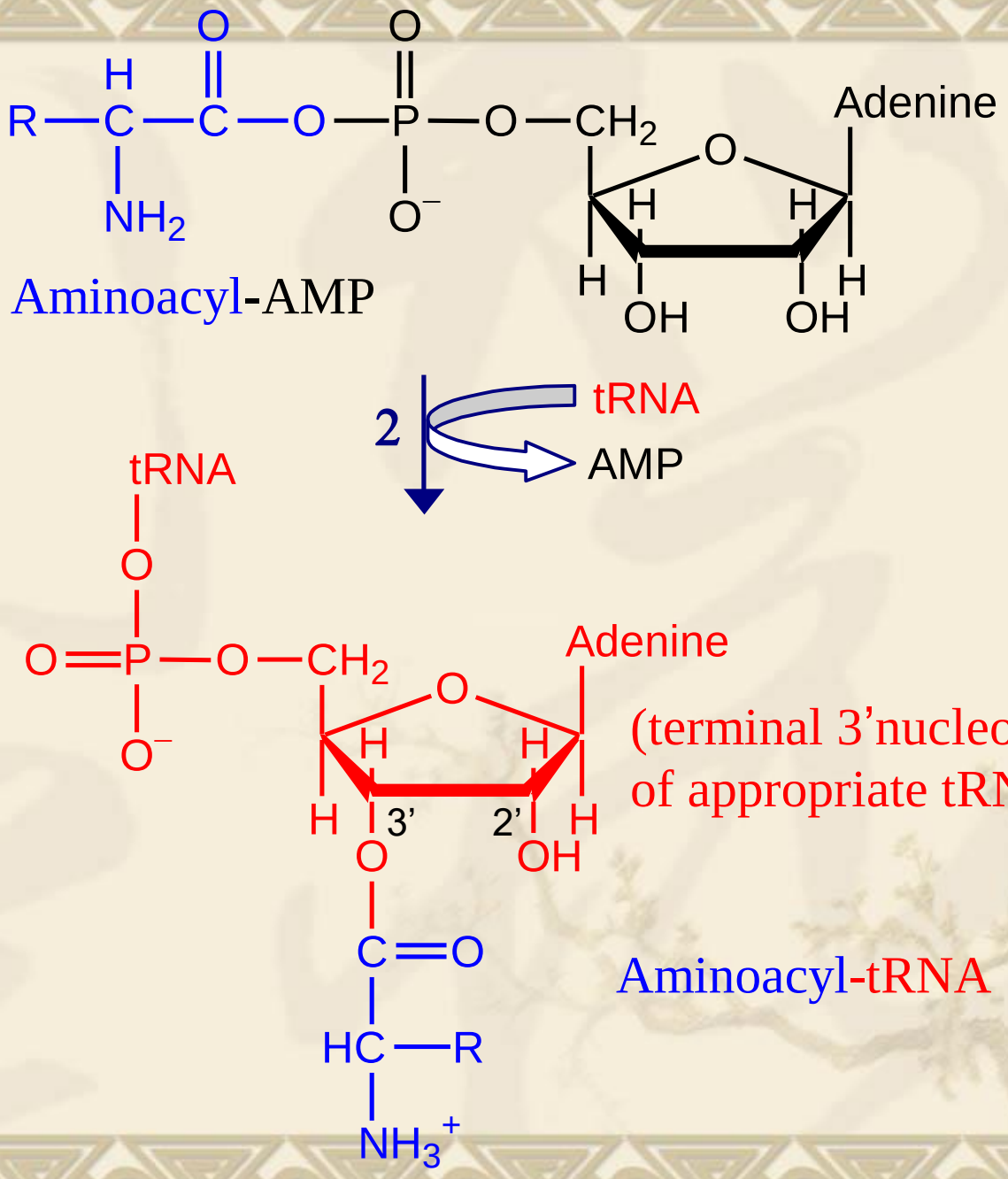


Aminoacyl-tRNA Synthetases catalyze linkage of the appropriate amino acid to each tRNA.

The reaction occurs in 2 steps.

In **step 1**, an O atom of the amino acid α-carboxyl attacks the P atom of the initial phosphate of ATP.

In **step 2**, the 2' or 3' OH of the terminal adenosine of tRNA attacks the amino acid carbonyl C atom.



Aminoacyl-tRNA Synthetase - summary:

1. **amino acid + ATP → aminoacyl-AMP + PP_i**
2. **aminoacyl-AMP + tRNA → aminoacyl-tRNA + AMP**

The 2-step reaction is **spontaneous** overall, because the concentration of **PP_i** is kept low by its hydrolysis, catalyzed by Pyrophosphatase.

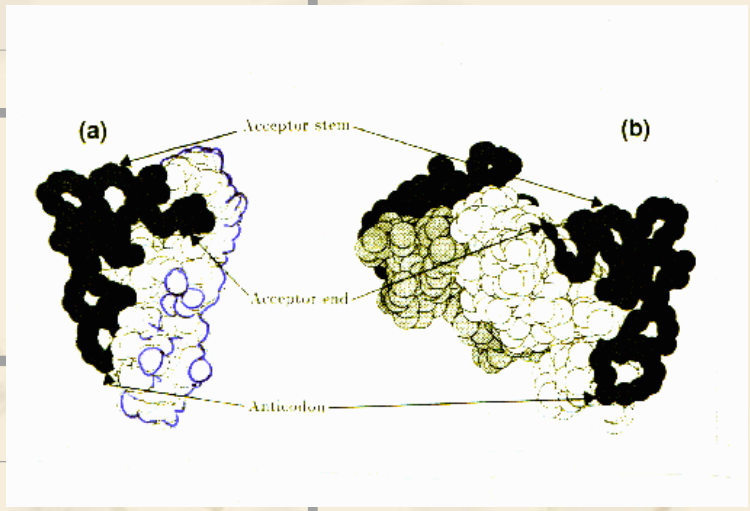
Two classes of Aminoacyl-tRNA Synthetases:

bind to opposite "sides" of the tRNA

- **Class I**: binds minor groves of acceptor and anticodon stems; ATP binding domain a nucleotide binding domain at N-term end of enzyme; anticodon arm binding domain at C term end; synthetase attaches aa to the **2'-OH** of the ribose of the 3 terminal nucleotide (the A of the -CCA) of the tRNA...
- **Class II**: binds major groves of acceptor and anticodon stems; anticodon arm bound in major groove by N-term domain; ATP binding domain more C-term; synthetase attaches aa to the **3'-OH** of the ribose of the 3 terminal nucleotide (the A of the -CCA) of the tRNA..

Members of Class I

Ia	Ib	Ic
Leu (alpha)	Tyr (alpha 2)	Arg (alpha)
Ile (alpha)	Trp (alpha 2)	Gln (alpha)
Val (alpha)		Glu (alpha)
Cys (alpha 2)		
Met (alpha 2)		



Members of Class II

IIa	IIb	IIc
His (alpha2)	Asp(alpha2)	Gly (alpha2/beta2))
Pro (alpha2)	Asn (alpha2)	Ala (alpha4)
Ser (alpha2)	Lys (alpha2)	Phe (alpha2/beta2)
Thr (alpha2)		

tRNA "Identity" Elements

- Most organisms encode close to **One Synthetase per amino acid**
- **Problem:** If there are roughly 1 tRNA species per codon but only one Synthetase per amino acid, then a Synthetase for an amino acid which has more than one codon must be able to specifically recognize and bind each of the tRNA species for each of these codons.
- Answer (in part ... much is still unknown): tRNA species that encode a given amino acid contain **Identity Elements**. These are nucleotides found in common between these tRNA species, but which are, at least in part, different in all other tRNA species.

Identity Element Nucleotides

- These are found primarily in two locations:
  1) the Anticodon
  2) the 5'-terminal nucleotides and the 3'-terminal nucleotides adjacent to the -CCA residues
  Some are also found in the D-loop stem-loop

Class I aaRSs:

Identity elements usually include residues of the anticodon loop & acceptor stem.

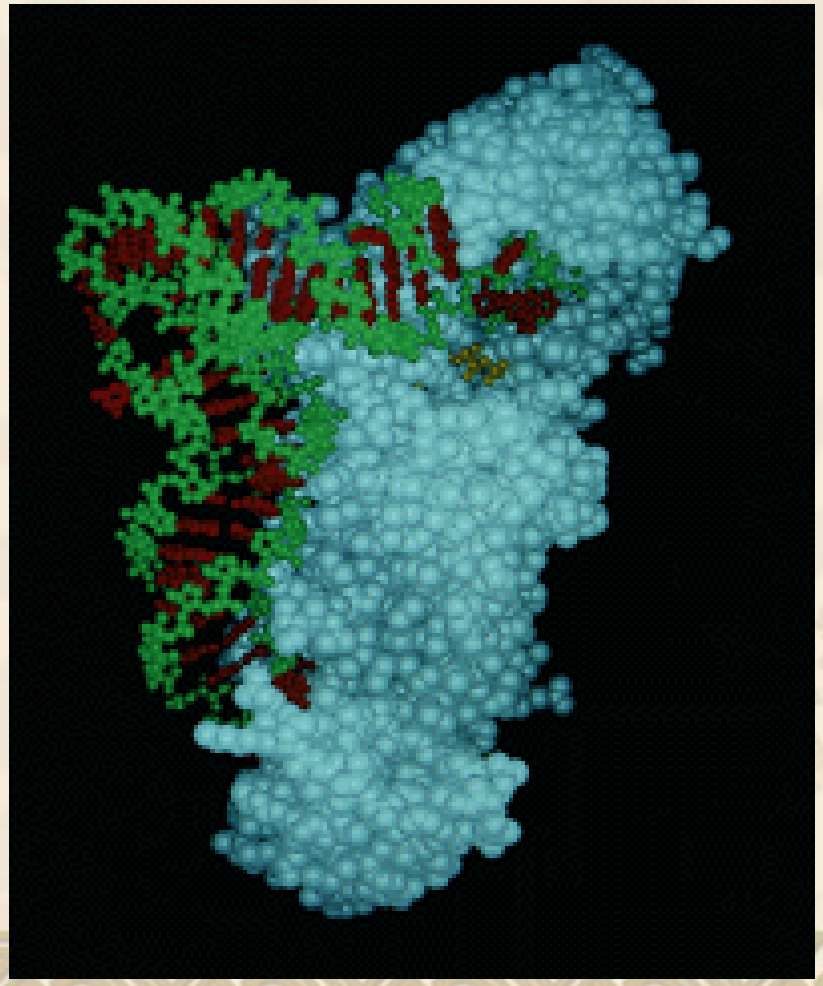
Class I aaRSs aminoacylate the **2'-OH** of adenosine at their 3' end.

Class II aaRSs:

Identity elements for some Class II enzymes do not include the anticodon domain.

Class II aaRSs tend to aminoacylate the **3'-OH** of adenosine at their 3' end.

The recognition of a tRNA molecule by its aminoacyl-tRNA synthetase.



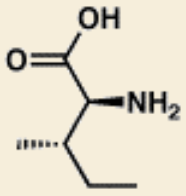
Proofreading/quality control:

Some Aminoacyl-tRNA Synthetases are known to have separate catalytic sites that **release by hydrolysis inappropriate amino acids** that are misacylated or mis-transferred to tRNA.

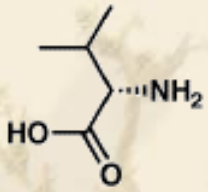
E.g., the aa-tRNA Synthetase for isoleucine (IleRS) occationally activates the closely related valine.

After **valine** is transferred to $tRNA^{Ile}$, to form Val-$tRNA^{Ile}$, it is removed by hydrolysis at a **separate active site** of IleRS that accommodates Val but not the larger Ile.

In some bacteria, editing of some misacylated tRNAs is carried out by **separate proteins**.



isoleucine



valine

Some amino acids are **modified** after being linked to a tRNA. Examples:

- In prokaryotes the initiator **$tRNA^{fMet}$** is first charged with **methionine**.

  Methionyl-tRNA formyltransferase then catalyzes formylation of the methionine, using **tetrahydrofolate** as formyl donor, to yield **formylmethionyl-$tRNA^{fMet}$**.

- In some prokaryotes, a **non-discriminating aaRS loads aspartate** onto **$tRNA^{Asn}$**.

  The aspartate moiety is then **converted** by an amido-transferase **to asparagine**, yielding Asn-$tRNA^{Asn}$.

  Glu-$tRNA^{Gln}$ is similarly formed and converted to **Gln-$tRNA^{Gln}$** in such organisms.

- Some proteins contain the unusual amino acid **selenocysteine** (**Sec**), with selenium substituting for the sulfur atom in cysteine.

  The **$tRNA^{Sec}$** is first loaded with **serine** via Seryl-tRNA Synthetase.

  The serine moiety is then **converted to selenocysteine** by another enzyme.

  Utilization of Sec-$tRNA^{Sec}$ during protein synthesis also requires **special elongation factors**.

Other roles of aminoacyl-tRNA synthetases:

- In some organisms, **Aminoacyl-tRNA Synthetases** (**aaRSs**) have evolved to take on **signaling roles** in addition to the catalytic role of joining an amino acid to the correct tRNA.
- Examples have been identified of particular aaRSs that regulate **transcription**, **translation** or **intron splicing** through binding to DNA or RNA.

- Proteolytic **cleavage** of the human **aaRS^{Tyr}** yields a **cytokine** that stimulates **angiogenesis**.

  A **truncated** form of the human **aaRS^{Trp}** inhibits angiogenesis.

- Regulation of **apoptosis** by the human **aaRS^{Gln}** is dependent on the concentration of its substrate glutamine.

- Several mammalian Aminoacyl-tRNA Synthetases associate with other proteins to form large **macromolecular complexes** whose roles are actively being investigated.

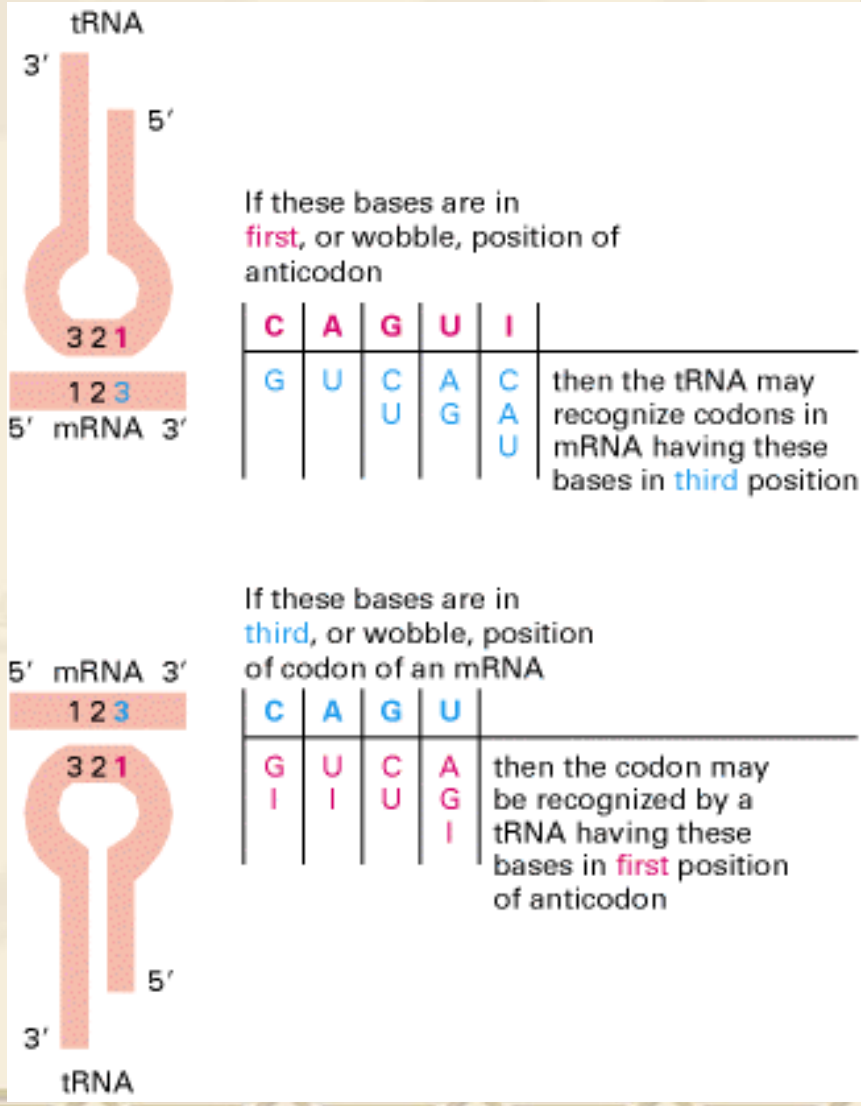
Wobble hypothesis: codon-anticodon recognition

- As studies on tRNA proceeded, 30 – 40 different tRNAs were identified in bacterial cells and as many as 50 – 100 in animal and plant cells. Thus the number of tRNAs in most cells is more than the number of amino acids found in proteins (20) and also differs from the number of codons in the genetic code (61).
- Consequently, many amino acids have more than one tRNA to which they can attach (explaining how there can be more tRNAs than amino acids); in addition, **many tRNAs can attach to more than one codon** (explaining how there can be more codons than tRNAs). As noted previously, most amino acids are encoded by more than one codon, requiring some tRNAs to recognize more than one codon.

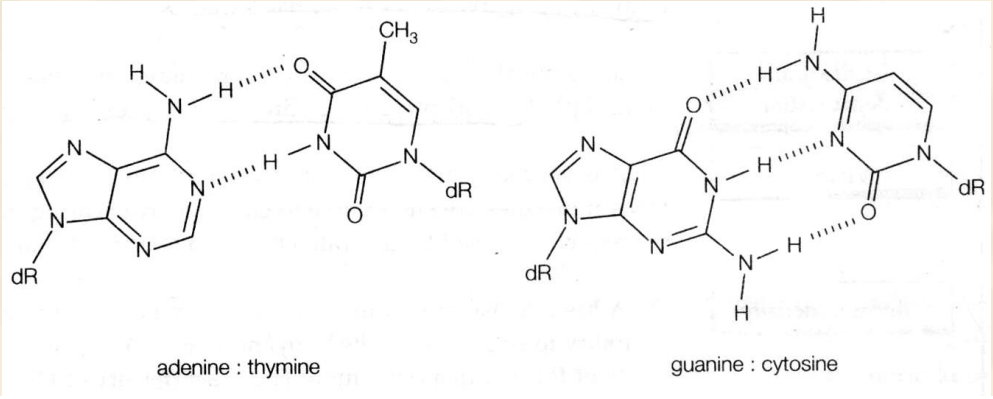
- Consequently, many amino acids have more than one tRNA to which they can attach (explaining how there can be more tRNAs than amino acids); in addition, many tRNAs can attach to more than one codon (explaining how there can be more codons than tRNAs). **As noted previously, most amino acids are encoded by more than one codon, requiring some tRNAs to recognize more than one codon.**

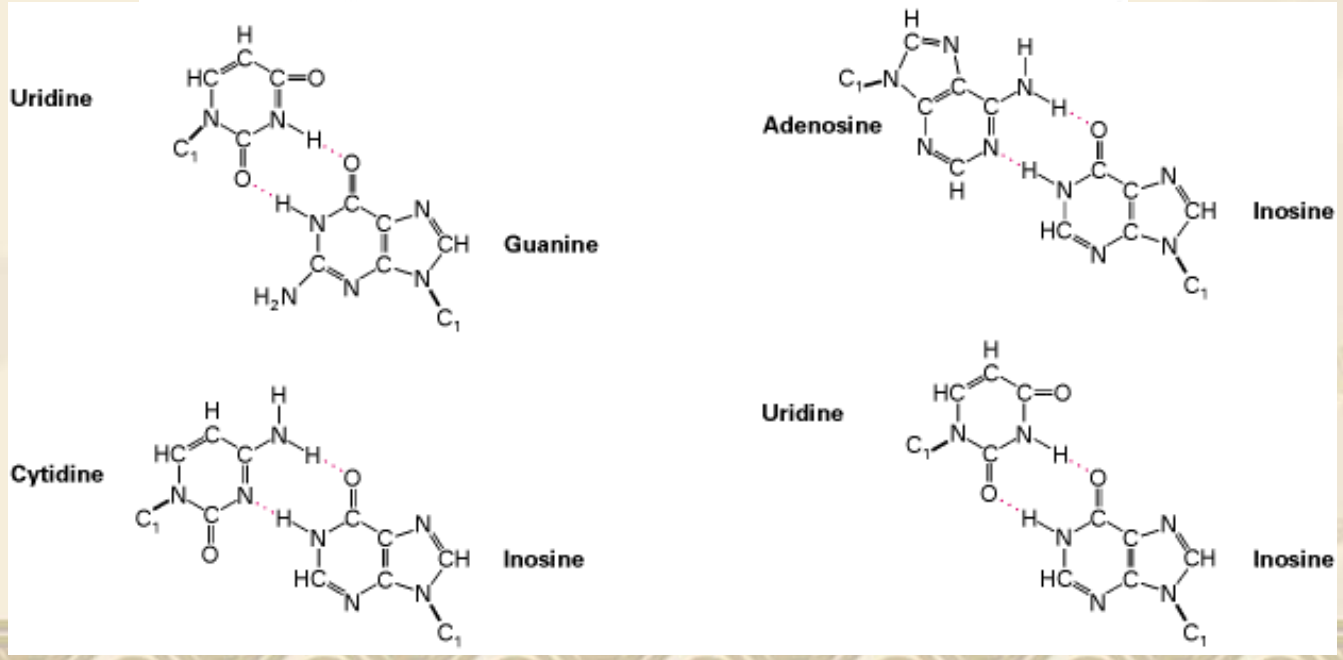
- Degeneracy or Redundancy in the genetic code occurs primarily in the 3rd position of the codon.
- To account for this, Crick proposed the Wobble Hypothesis.
- This hypothesis states that base pairing of bases at the 3rd position of the codon and the 1st position of the anticodon can be nonstandard due to "wobble" of the nucleotide at the 1st position of the anticodon. Such wobble is due to the curvature of the anticodon in the loop of the anticodon arm of the tRNA.

- **The first and second bases in an mRNA codon form Watson-Crick base pairs with the third and second bases, respectively, of a tRNA anticodon.**



The nonstandard, wobble base pairs U·G, C·I, A·I, and U·I.

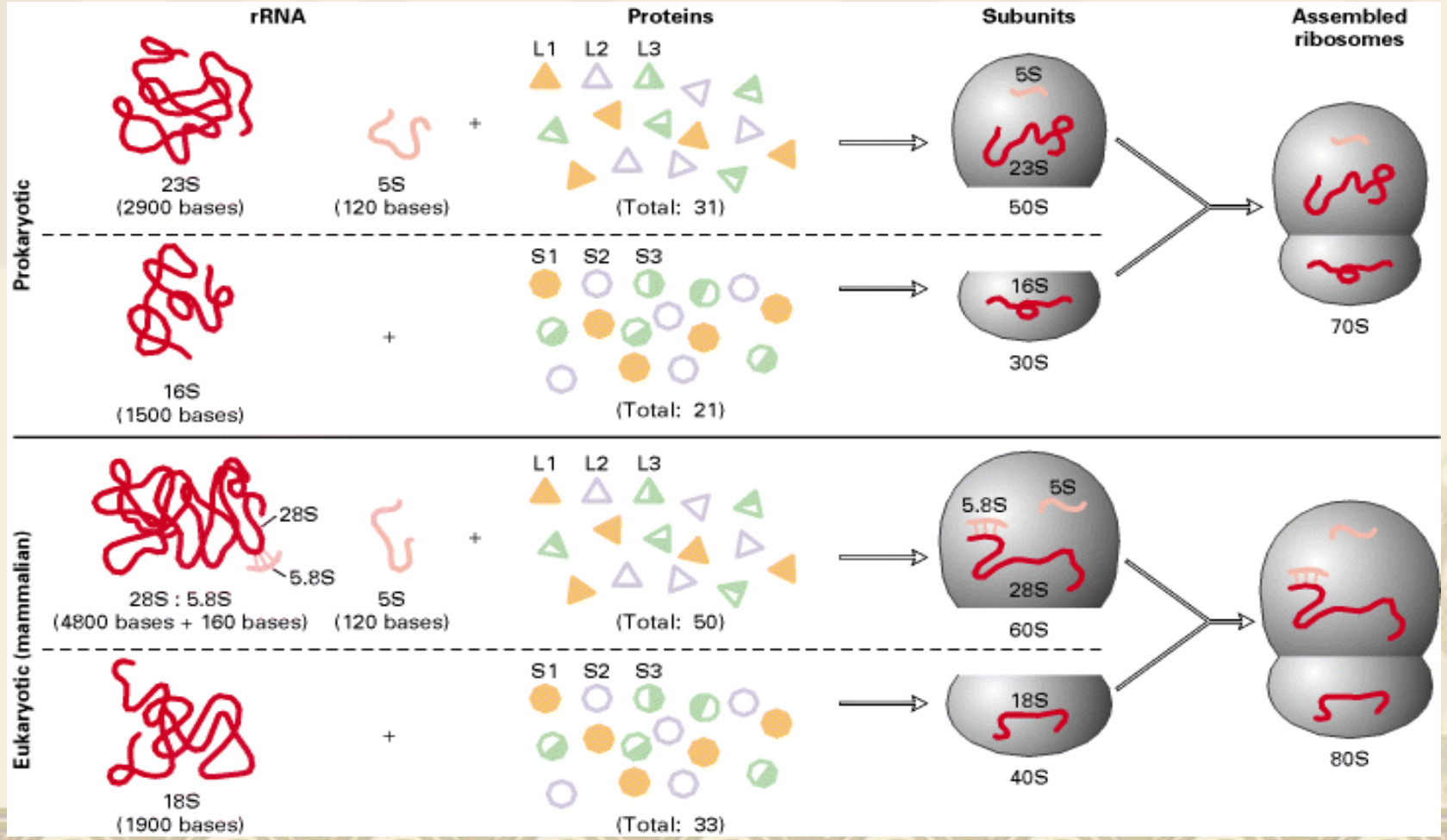




3. Ribosomes

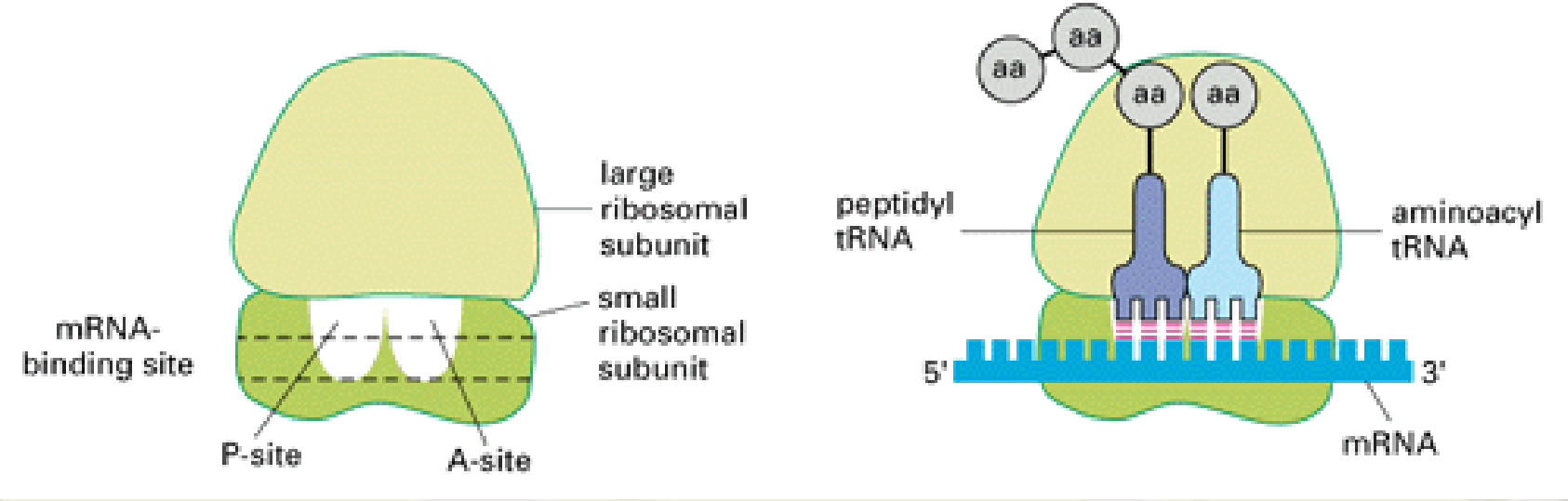
- **3.1. Ribosome Structure**:
  Prokaryotes: 30S small subunit + 50S large subunit -> 70S complete ribosome
  Higher Eukaryotes: 40S small subunit + 60S large subunit -> 80S complete ribosome

The general structure of ribosomes in prokaryotes and eukaryotes.

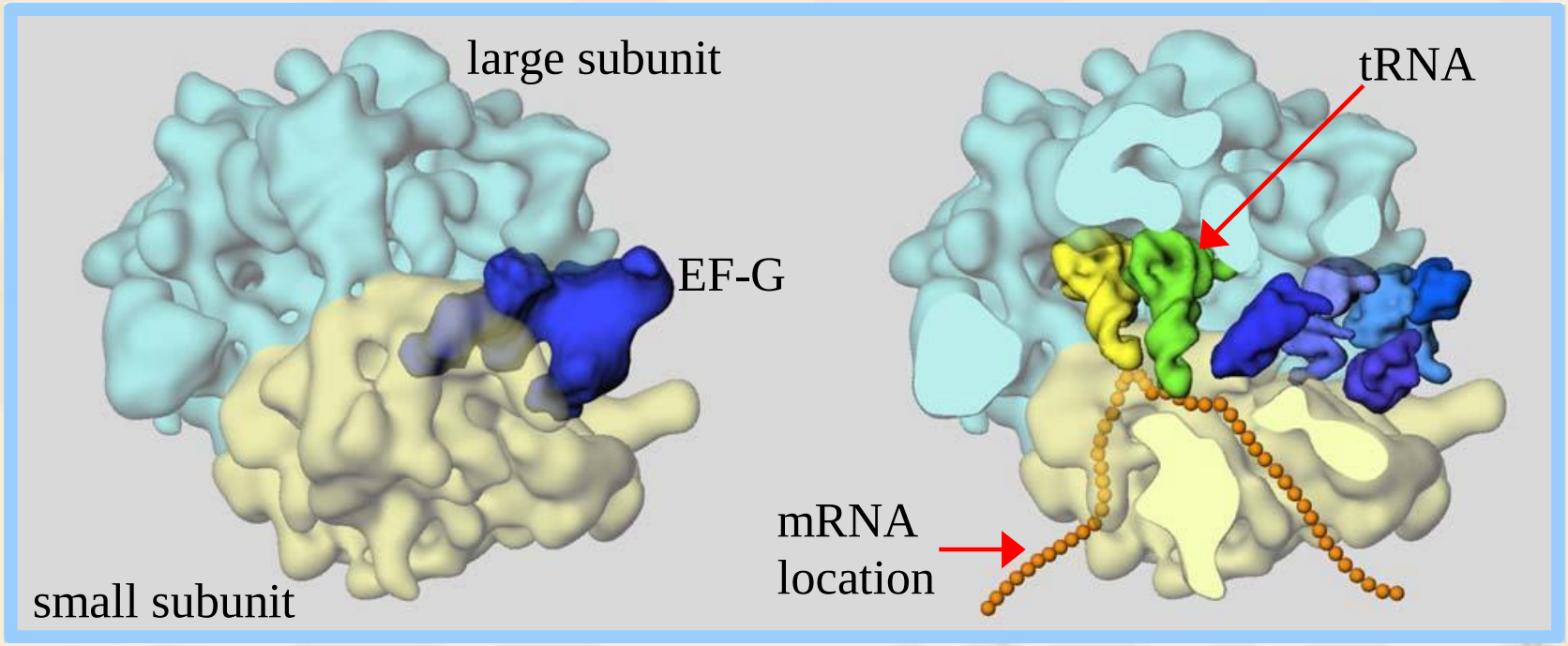


Ribosome Binding Sites

- 3 sites for tRNA: A site (entry site), P site (aa transfer), E site (exit)
  mRNA site on 30S ribosome
  Peptidyl transferase site on 50S ribosome



Structure of the *E. coli* **Ribosome**



The cutaway view at right shows positions of **tRNA** (P, E sites) & **mRNA** (as orange beads). EF-G will be discussed later. This figure was provided by Joachim Frank, whose lab at the Wadsworth Center carried out the cryo-EM and 3D image reconstruction on which the images are based.

Initiation of Translation Process

- During the initial stage of protein synthesis in all cells, a ribosome assembles, complexed with a mRNA and an activated initiator tRNA, which is correctly positioned at the start codon. Because the details of initiation and the mechanism for locating the translation start site differ in bacteria and eukaryotes, we discuss the two systems separately.

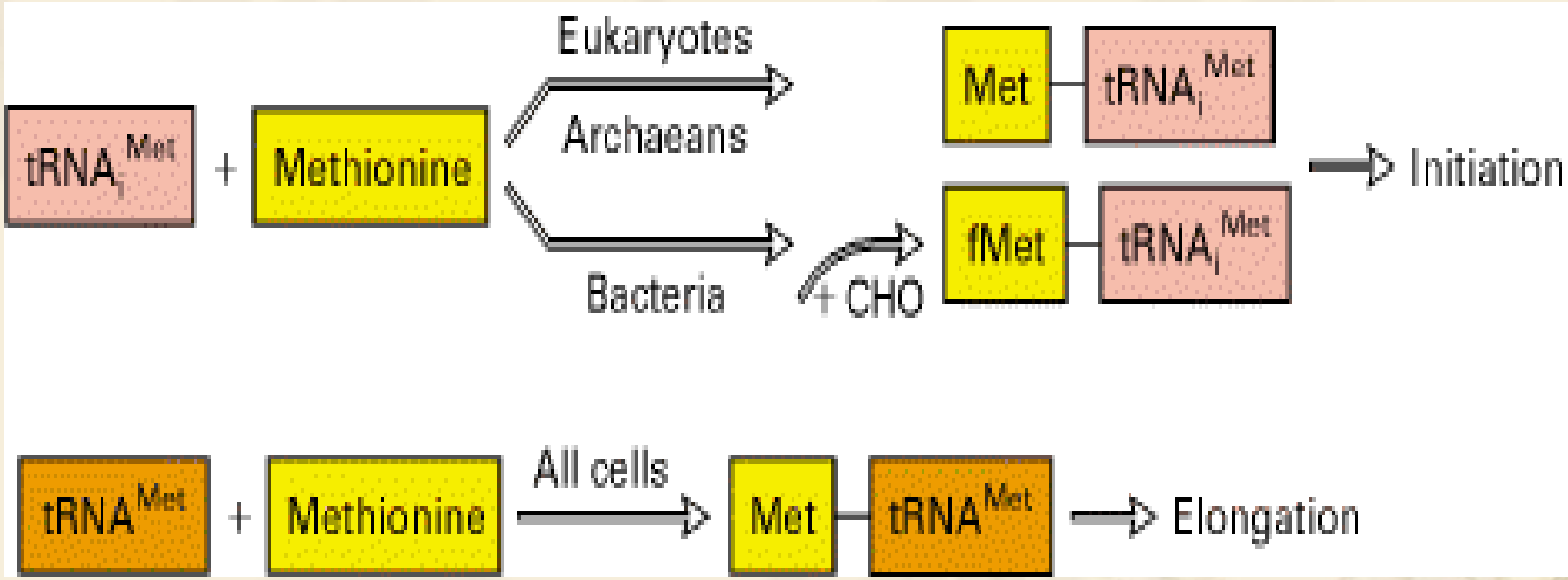
Prokaryotic mRNA has two recognition sites:

- 1). The **Shine-Delgarno ribosome binding site** AGGAGG, complementary in sequence to a CCUCCU sequence on 16S rRNA on the 30S ribosome
- 2). The translation **Start Codon**, usually AUG, sometimes GUG

fMet-tRNAf

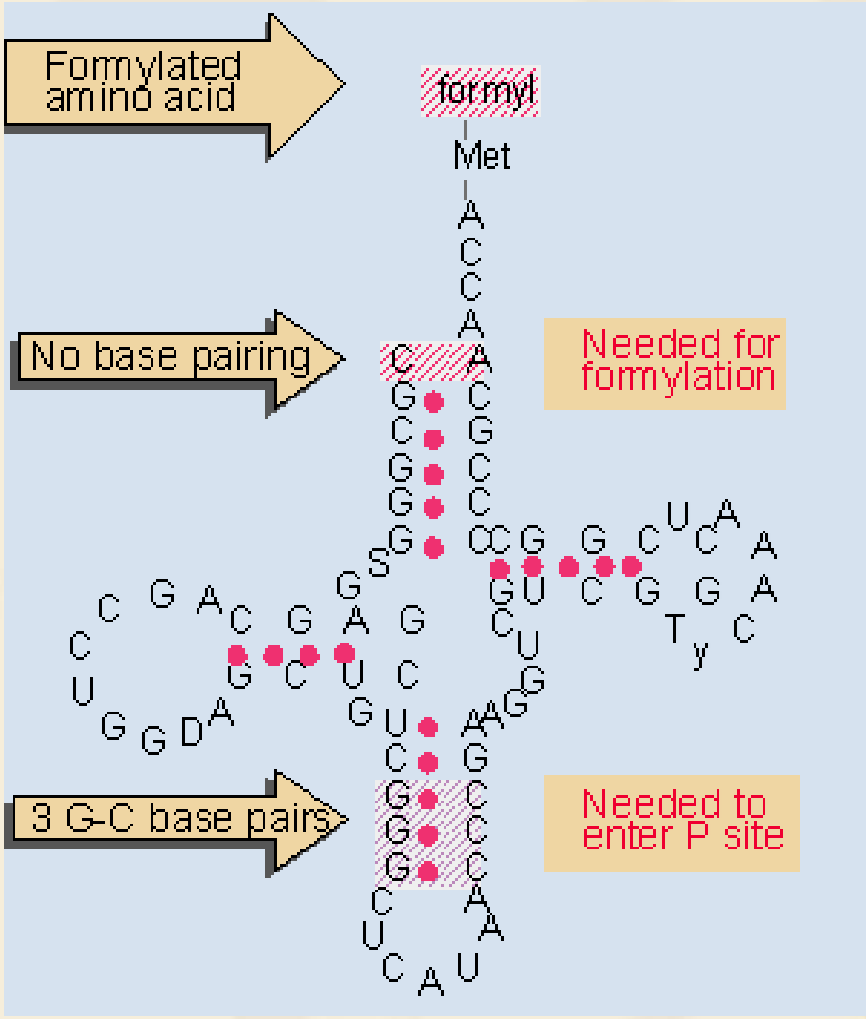
- A specific aminoacyl-tRNA used in Initiation : The formylation reaction occurs after tRNAf is charged with Methionine
  This reaction uses formyl-tetrahydrofolate as formyl donor.
- A deformylase removes the Formyl group during the Termination process; the Met is also sometimes removed

Two types of methionine tRNA are found in all cells.

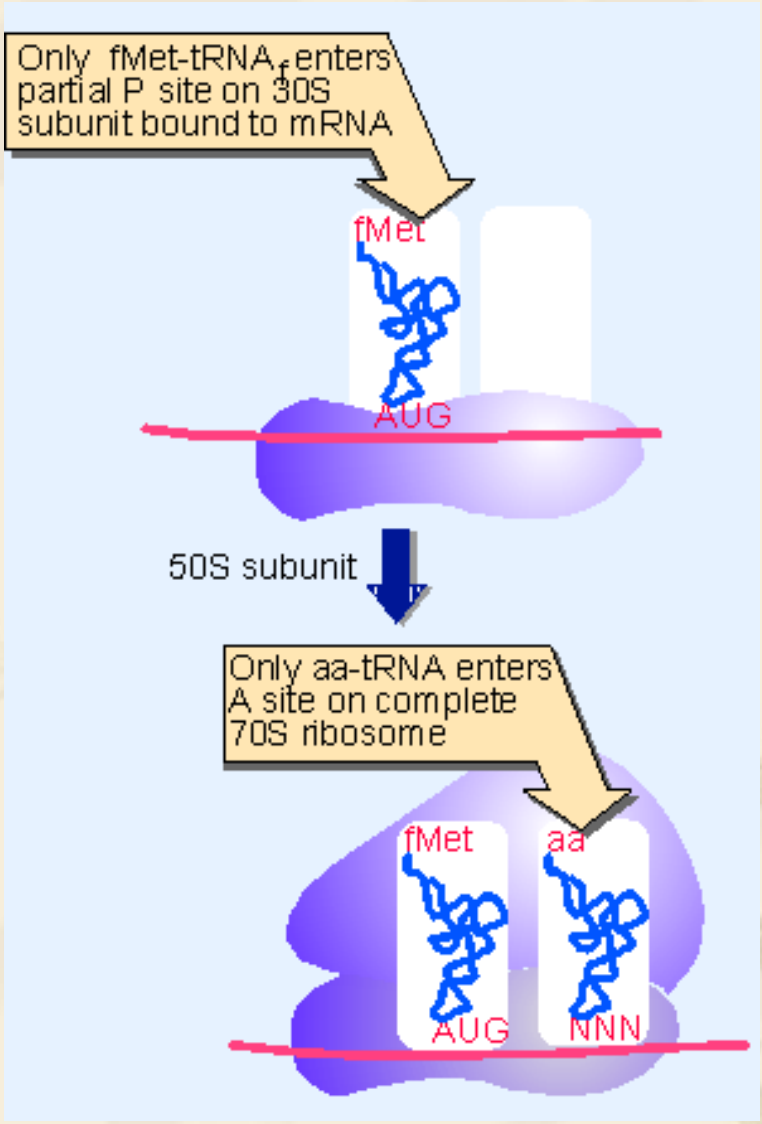


tRNAf differs from tRNAMet and other tRNAs as follows:

- 1) CA unpaired bases in Acceptor Stem => no function in elongation
  2) Three G:C pairs in Anticodon Stem => can enter P site in 30S ribosome

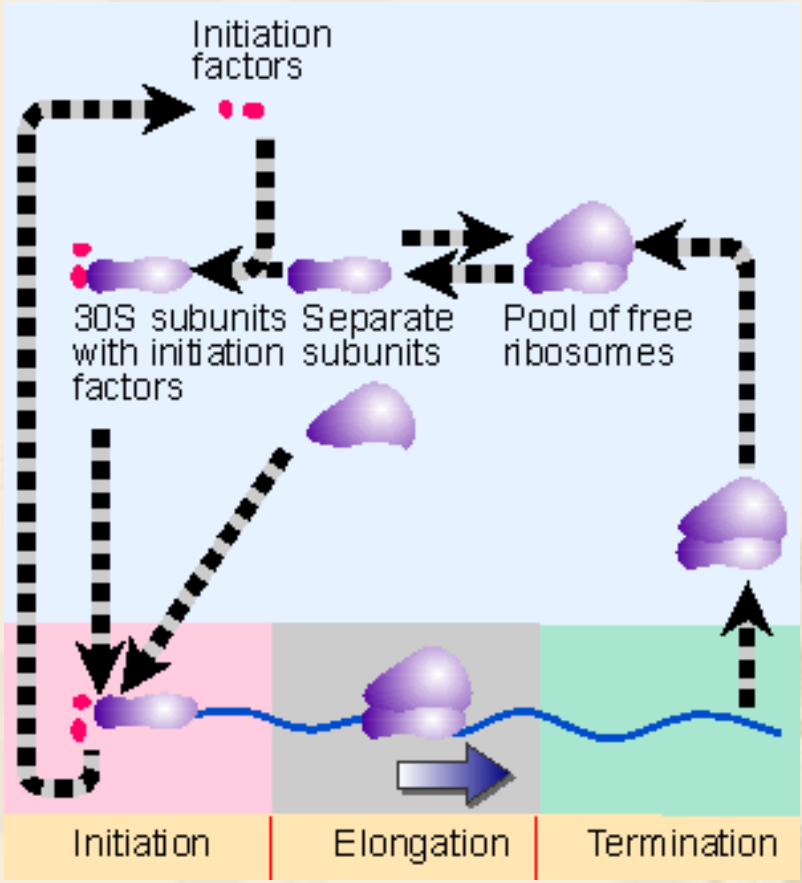


- Only fMet-tRNA$_f$ can be used for initiation by 30S subunits ; only other aminoacyl-tRNAs (aa-tRNA) can be used for elongation by 70S ribosomes.



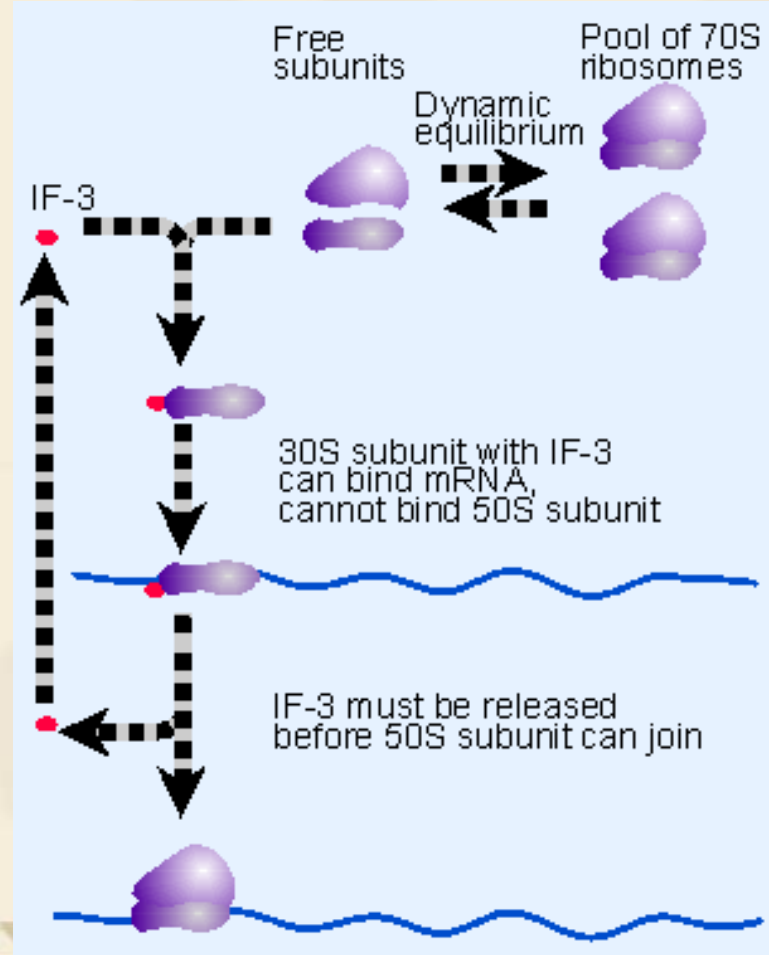
Elongation

- Ribosomes exist in cytoplasm as pools of 30S, 50S, and 70S ribosomes
- ONLY 30S and 50S ribosomes participate in Initiation of Translation
- Three **Initiation Factors** (IF-1, IF-2, IF-3) are required for Initiation



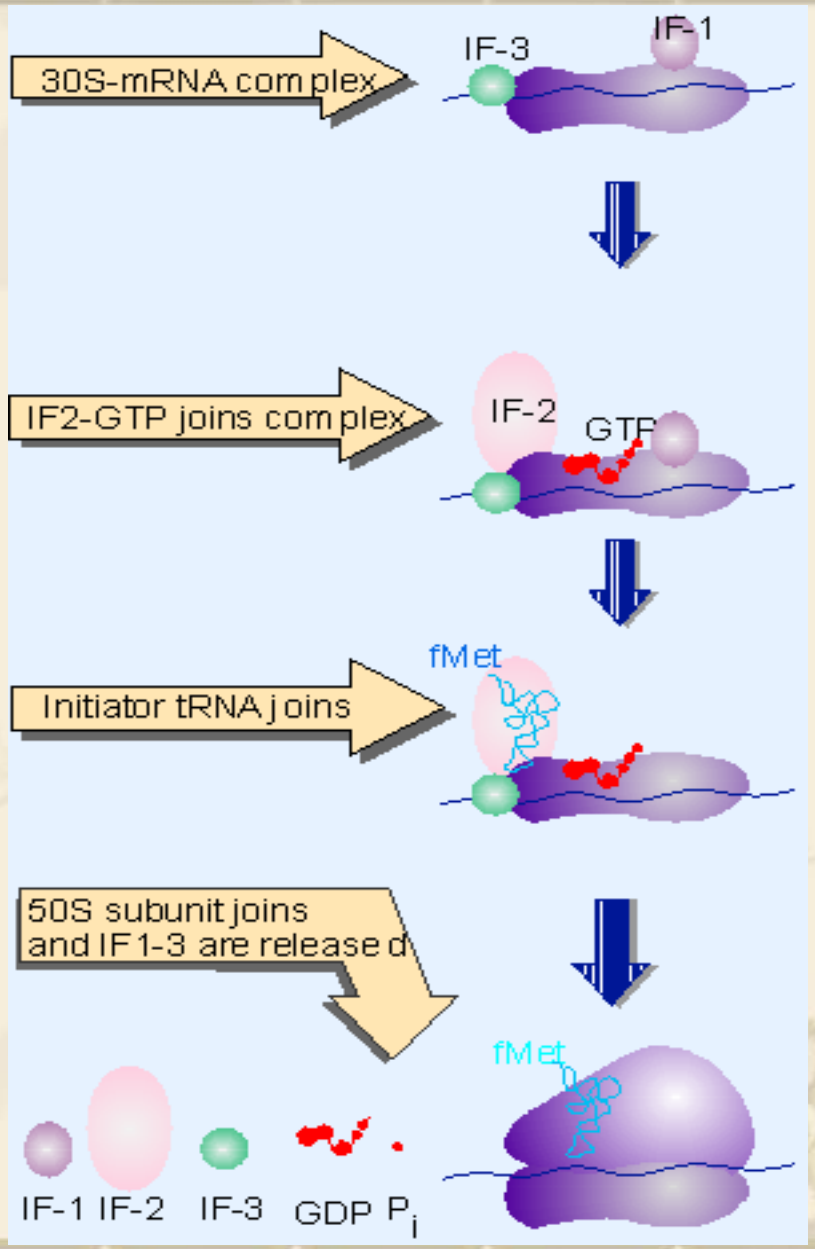
Three **Initiation Factors** :

- 1. IF-3 binds 30S ribosomes, enabling 30S ribosome to bind mRNA and preventing 50S ribosome from binding to 30S ribosome ... This forms the **30S-mRNA-IF3 complex**

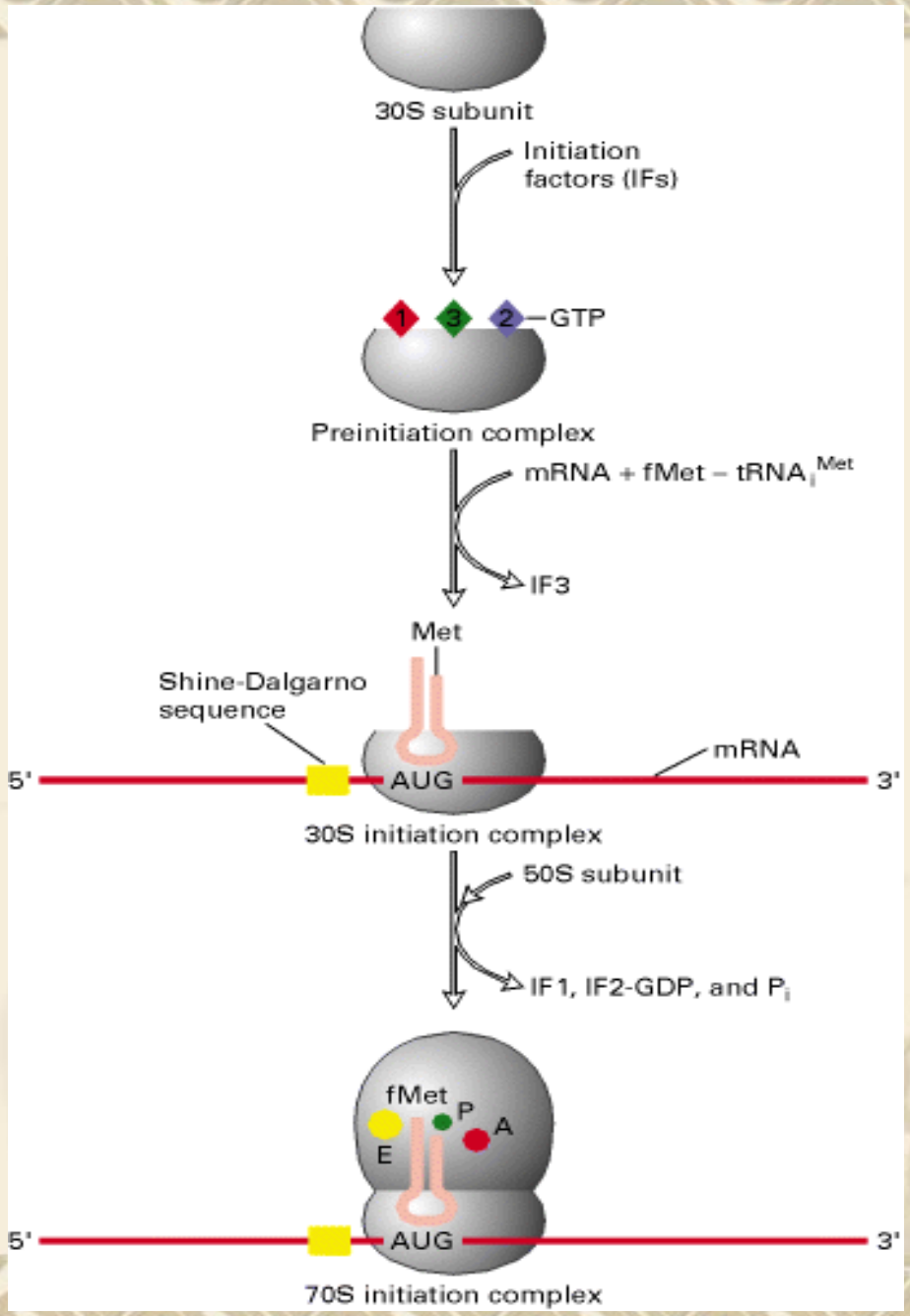


- 2. IF-1 binds 30S ribosomes; function probably is to stablize the Initiation Complex.
- 3. IF-2 binds fMet-tRNAf, brings this activated Initiation tRNA to the 30S-mRNA-IF3 Complex; facilitates insertion of the fMet-tRNAf into P site of 30S ribosome; and activates a GTPase activity to facilitate final joining of 50S ribosome to the Initiation complex.

  IF-2 and IF-3 (and probably IF-1) dissociate when 50S ribosome binds.



Bacterial initiation of protein synthesis.

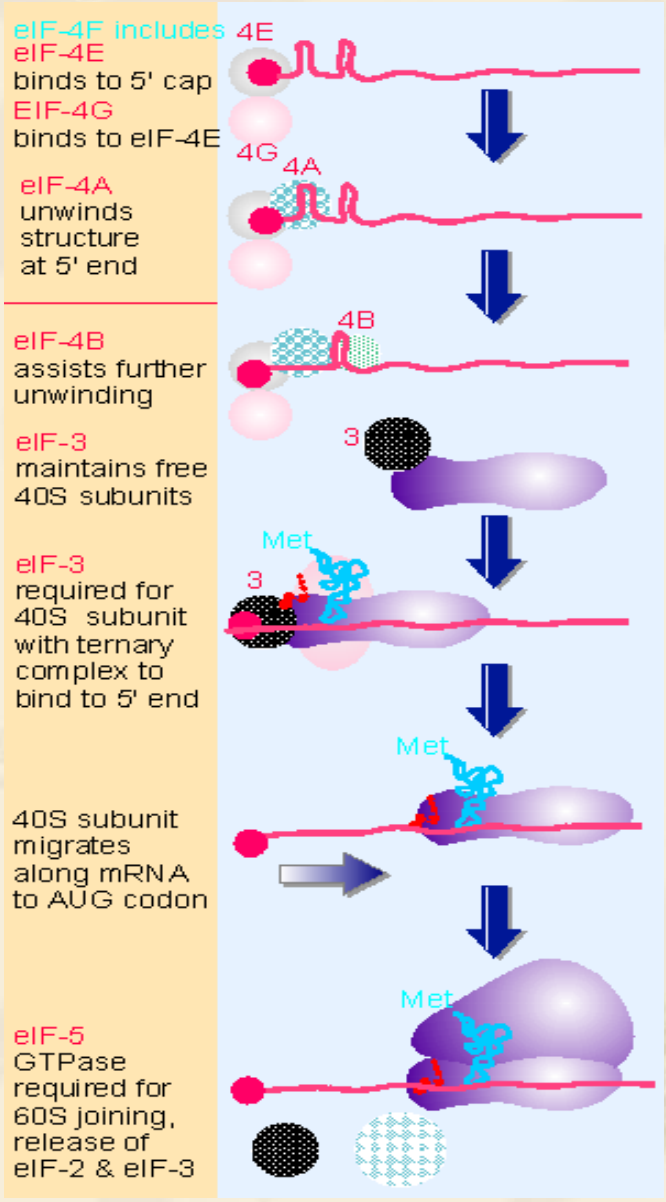


Steps in initiation

- **Two things happen in parallel:**
- **• preparation of the mRNA for 40S subunit binding**
- **• preparation of the 40S subunit for mRNA binding**

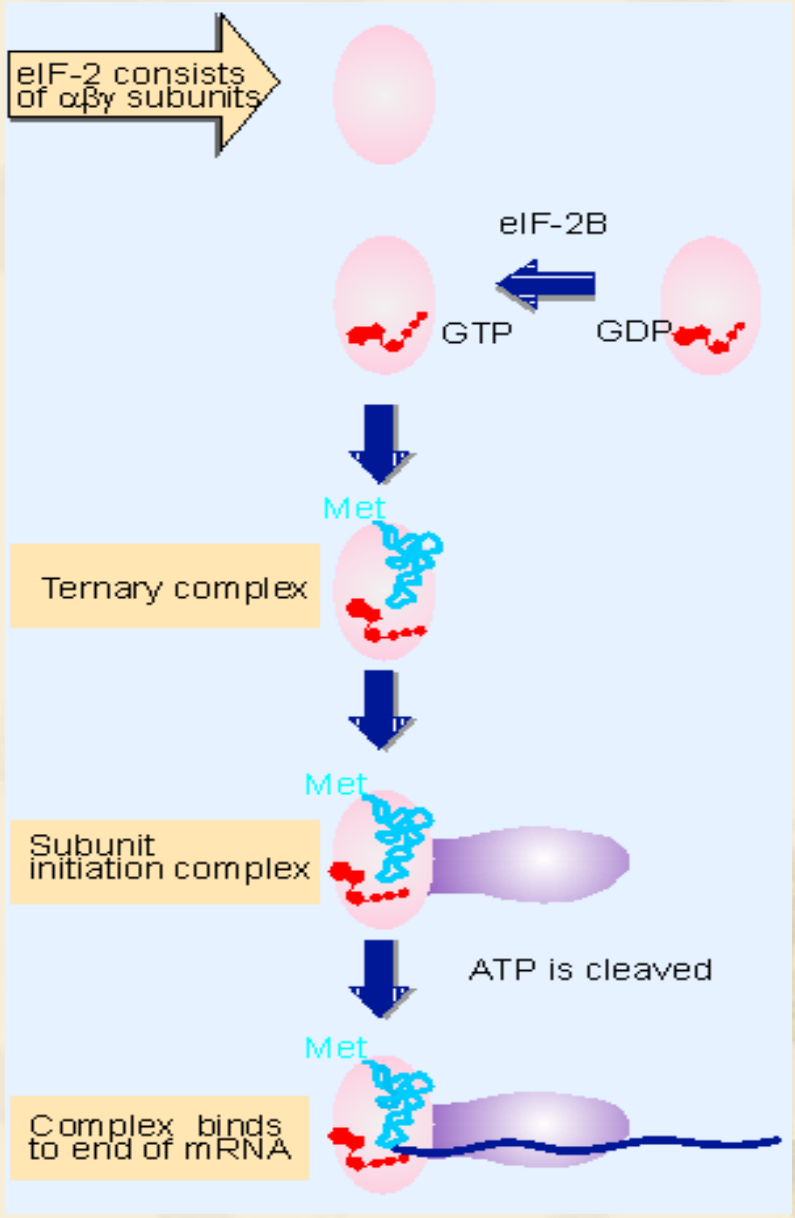
Preparation of the mRNA

- 1. eIF4E (cap binding protein, CBP) binds the 5’ cap; (in mammals as part of eIF4F complex)
- eIF4G(the scaffolding subunit) is to link other components of the initiation complex.
- 2. eIF4A unwinds mRNA structure near the 5’ cap (ATP)
- 3. eIF4B assists further unwinding



Preparation of the 40S subunit

- 1. formation of ternary complex - Met tRNAi, eIF2, GTP
- • eIF2 binds GTP (Met tRNAi affinity increases)
- • eIF2/GTP binds Met tRNAi
- 2. ternary complex binds free 40S subunits
- (increases 40S affinity for mRNA)



Ribosome association with the mRNA

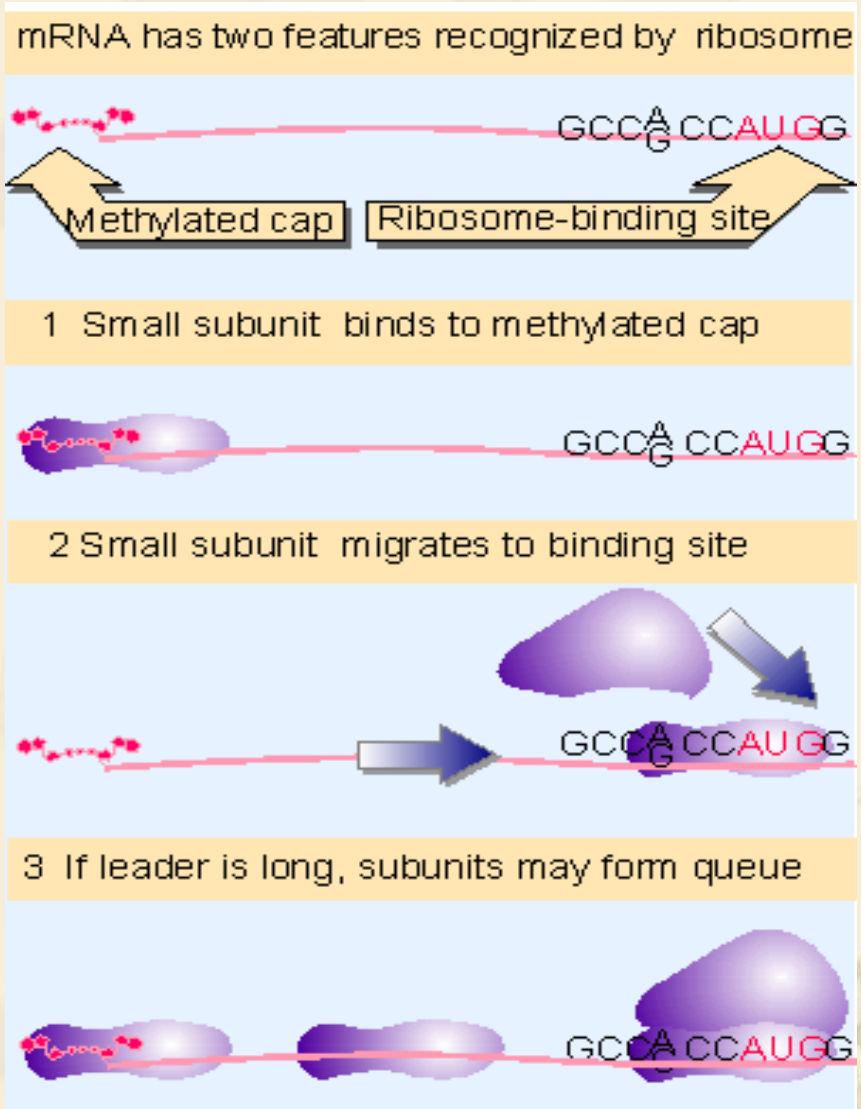
- • 40S binds first
- • 40S scans to first appropriate AUG

  Ribosome scanning:

  ribosome binds 5’ cap, moves down RNA to start AUG

  ATP required for translocation

  mRNA kept unstructured by eIF4A, eIF4B
- • 60S binds



- With the 40S subunit and the mRNA prepared, eIF3 directs their binding to each other (ATP used).
- 40S subunit scans along the mRNA, stopping at the first AUG (ATP used).

recognition of the start AUG

context dependent, especially in higher eukaryotes (GCCACCAUGG) (-3/+4 conserved)

Kozak sequences (for Marilyn Kozak, who defined them)

complementary region in 18S rRNA

no other sequences required, minimum distance between cap and AUG (15 nts.)

stable stem-loops can inhibit translation 20-1000 fold.

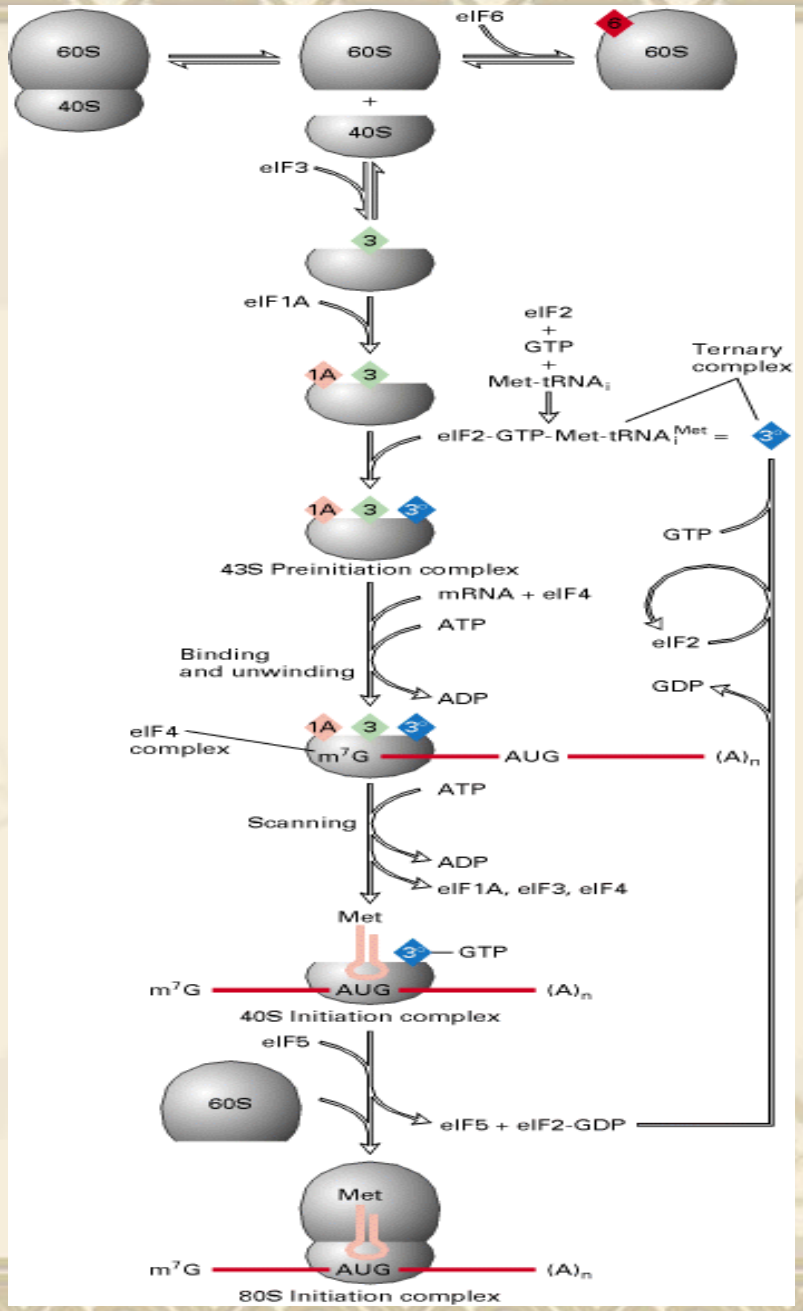
- eIF5 反式作用因子 binds to the 40S subunit stalled at an AUG, triggers the hydrolysis of the GTP bound to eIF2.

  eIF2 loses its affinity for 40S, falls of。

  eIF2 dissociation allows 60S subunit binding，assisted by eIF4C;

  eIF6 is released from 60S, eIF3 is released from 40S.

Eukaryotic initiation of protein synthesis.



factor	size	function
eIF2		binding tRNAiMet
eIF4F		multimer mRNA binding / unwinding
		contains eIF4E, eIF4A, p220
eIF4E (cdc33)	24,000	5’-cap binding
eIF4A		mRNA binding, ATP binding,
		prototype ‘DEAD’ box protein
		putative RNA helicase
eIF4B		mRNA binding / unwinding
eIF3	500,000	mRNA binding
eIF5	150,000	releasing eIF2, eIF3
eIF4C		binding 60S subunit
eIF6	23,000	prevents 40S-60S joining
eIF1	15,000	mRNA binding accessory
eIF4D		unknown
eIF5A		promote 1st peptide bond?

Elongation of Translation Process

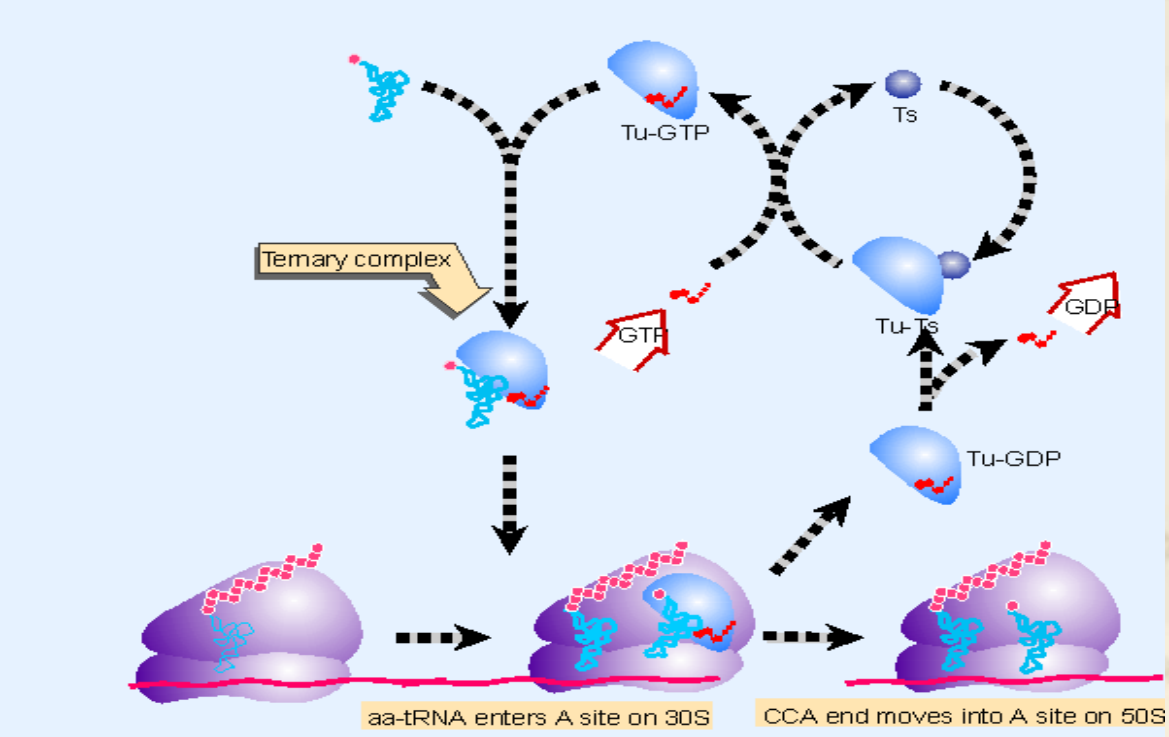
- **Loading Ribosome with AminoAcyl-tRNA molecules**:

  Elongation factor **EF-Tu** 延伸因子 uses energy from GTP to bring Activated tRNA molecules to the A site of 70S ribosomes whose P site is already occupied

  EF-Tu - GDP is then released; this is the Inactive form of EF-Tu

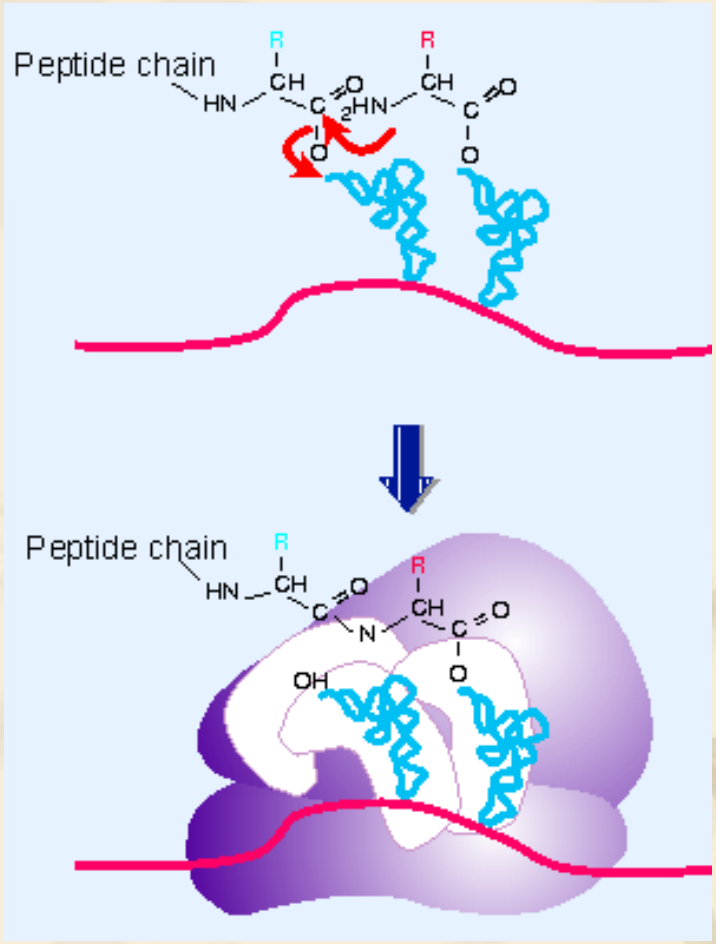
EF-Tu and EF-Ts 鸟苷交换因子

- **Tu-Ts**: GDP is then released by Tu; Tu complexes with elongation factor Ts
  The Tu-Ts complex remains intact until GTP reactivates Tu, releasing Ts
- 70,000 Tu molecules per cell, nearly 1 per tRNA, vs only 10,000 Ts molecules per cell 大致相当于 tRNA的分子量

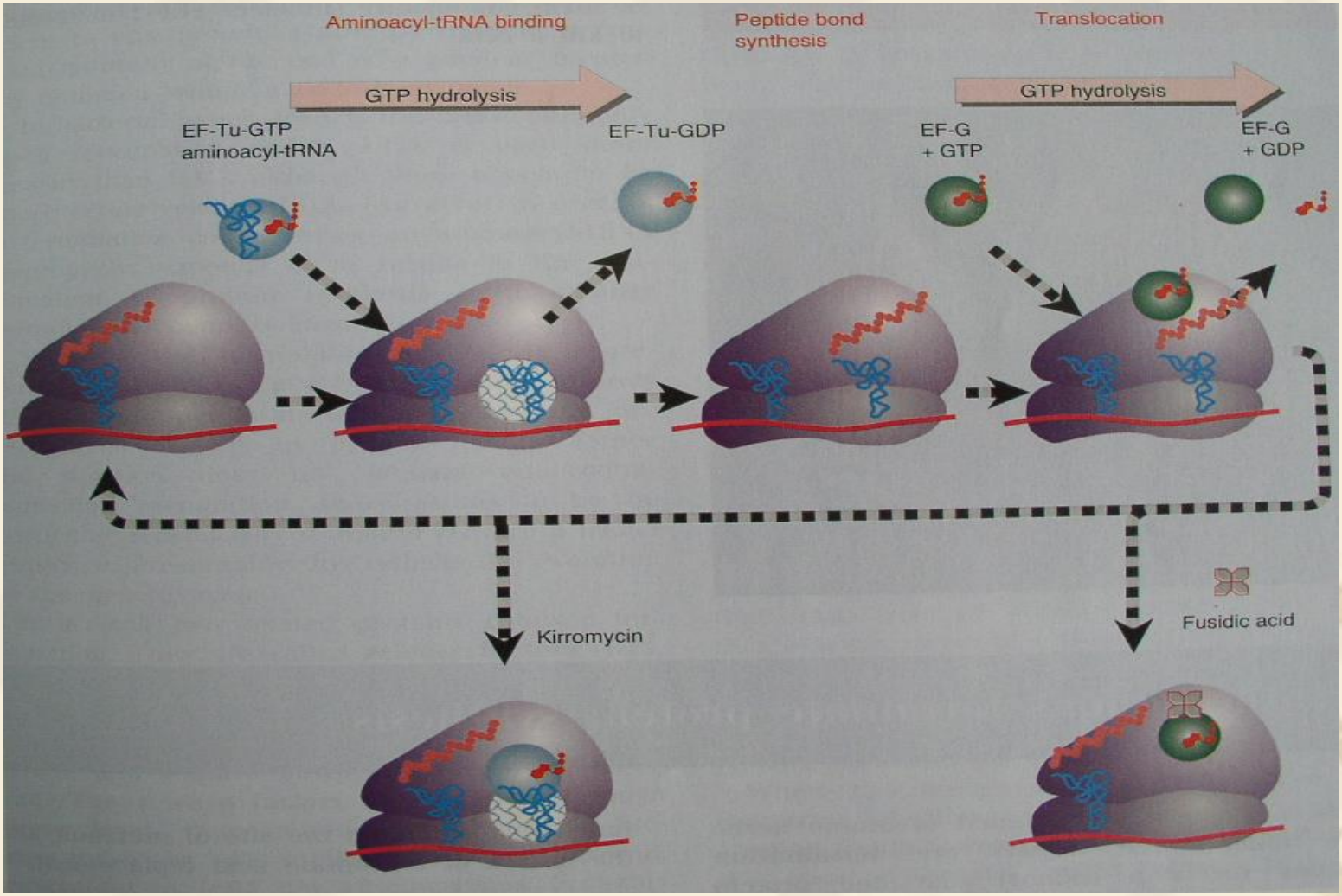


Translocation and Peptide Chain Elongation

- **Peptidyl transferase** 肽酰转移酶
- catalyzes transfer of aa from tRNA to the polypeptide chain; this reaction may well be catalyzed mainly by 23S rRNA 整个P位都空缺了
- Translocation then occurs: factor **EF-G** and GTP
- 20,000 EF-G per cell, about 1 per ribosome; released upon GTP hydrolysis
- RNA也可能是酶

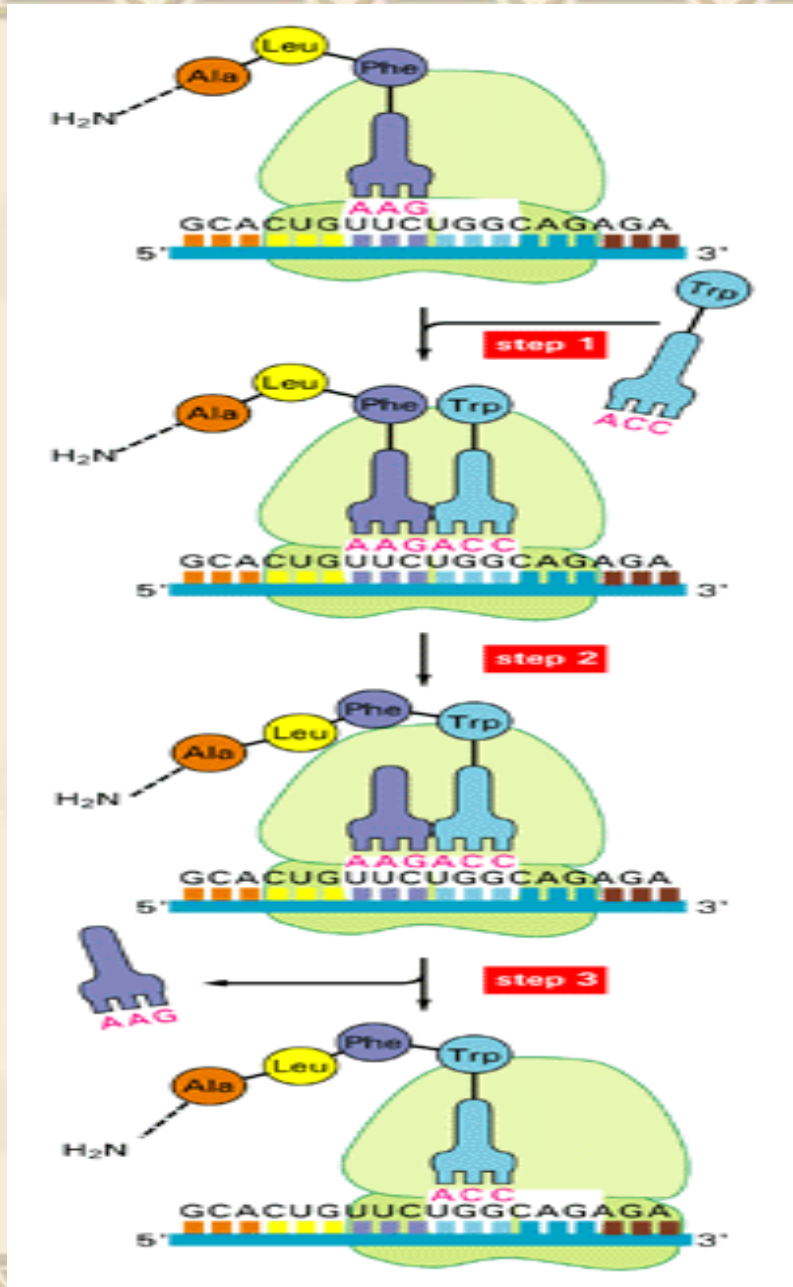


Ribosomes bind EF-Tu and EF-G separately, resulting in a Cyclic, Sequential process:



The elongation phase of protein synthesis on a

ribosome.



Eukaryotic

- Elongation Factor **eEF-1** is similar to EF-Tu, and is often called **EF-T** ...
  with GTP, brings AminoAcyl-tRNA to A site
- Elongation Factor **eEF-2** is similar to EF-G
  转移酶：a GTP-dependent translocase

6. Termination of Translation Process

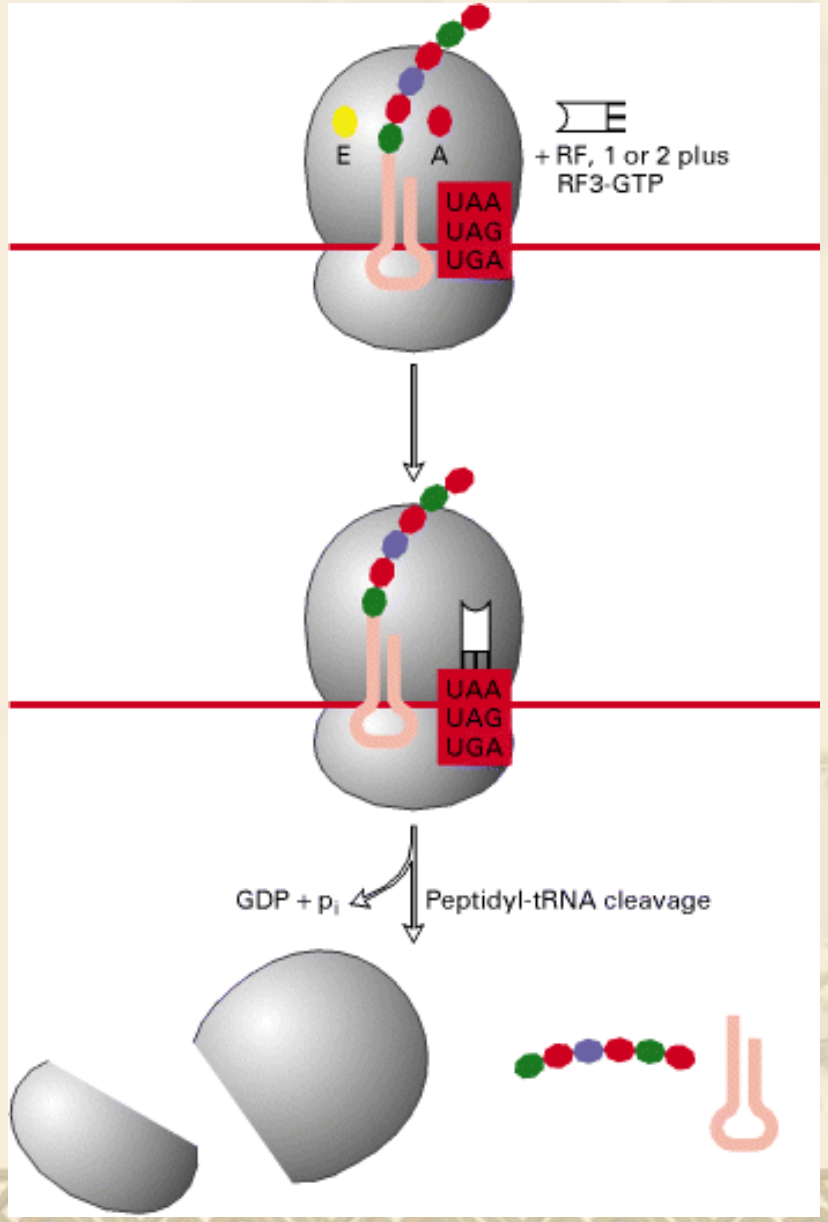
- Stop Codons (Nonsense Codons): universally used - Prokaryotes; Eukaryotes UAG (amber), UAA (ochre), UGA (opal)

6.1. Releasing Factors

- **Prokaryotes:**
- **RF1**: recognizes UAG and UAA; **RF2**: recognizes UAA and UGA
  only ~600 molecules per cell, 1 per 50 ribosomes
  require Peptidyl-tRNA in P site; act at the A site

- **Eukaryotes**:
- **eRF**: recognizes all 3 stop codons ... requires GTP for release

Termination of translation



Introns are found in the

a) in the mRNA in the cytoplasm

b) in the promoter structure

c) initiation complex for transcription

d) processed during transcription

e) none of the above

the intiation of transcription in eukaryotes requires

a) shine dalgarno sequences

b) TATA box

c) transcriptional factors

d) all of the above

e) b and c

The peptidyl transferase is involved in

a) transfer of the peptide to the A site amino acid

b) making the peptide bond between amino acids

c) elongation step of translation

d) all of the above

e) a and b

True or false

1. Sigma factor is essential for initiation in eukaryotic promoters.
2. The tRNA-Amino acid complex recognizes the Shine Dalgarno sequence during the elongation step of translation.
3. Before a eukaryotic mRNA can be translated, it must be modified by the addition of a 5' cap and a 3' poly-A tail.

An assay

DNA sequence
TACCATGTATTCGTCACTGTA
ATGGTACATAAGCAGTGACAT
Amino Acid Sequence
Met-Val-His-Lys-Gln

- Figure out the 5' and 3' ends of each strand of DNA
- Figure out which strand is coding and which strand is template
- Figure out what Amino Acid Sequences would be made with the following mutations and what type of mutations they are

Mutation 1
TACCAATGTATTCGTCACTGTA
ATGGTTACATAAGCAGTGACAT
Mutant #2
TACCATGTAATCGTCACTGTA
ATGGTACATTAGCAGTGACAT

Drill question

A chemical inhibits the synthesis of GTP. What would be the affects of on Replication , Transcription, and Translation?