Biology from an EE perspective
Lecture 6

Protein Synthesis

Discuss processes that effect

DNA → mRNA→ protein

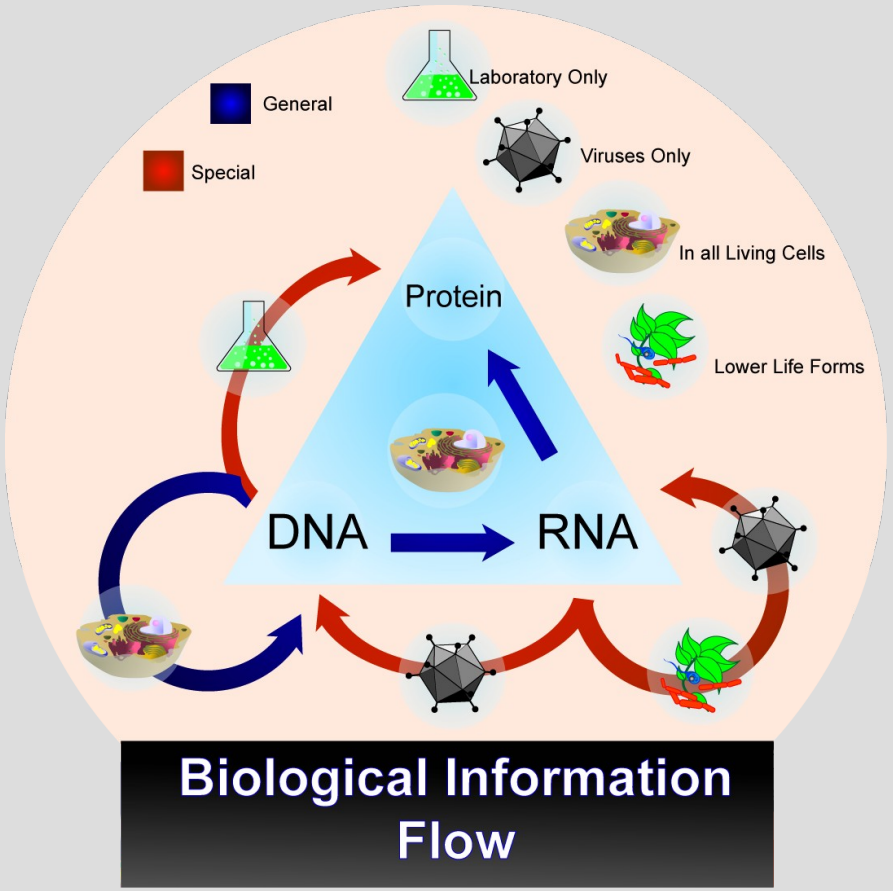
Rakesh K Lal

Lecture Overview

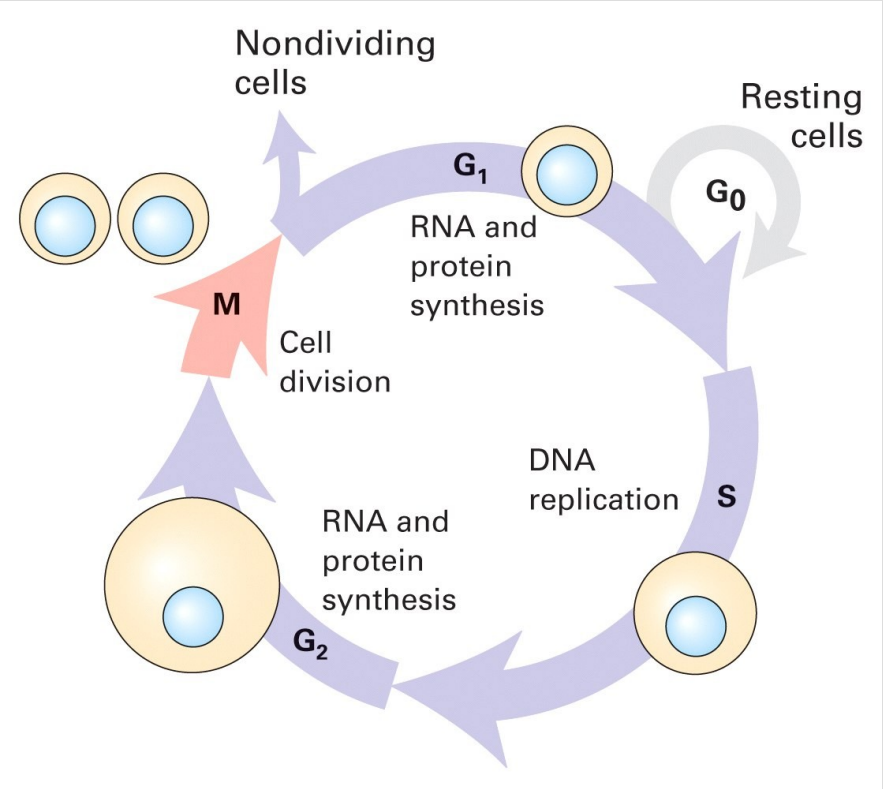
- Look at protein synthesis
- This is an core element of what biologists call the central dogma (I prefer to call it the central principle)
- An important set of processes for all forms of life
- One gets some feel of the complex & accurate synthesis machinery with feedback that exists
- Important for understanding modern literature

The Central Dogma (Principle)

- DNA-->RNA-->protein
- DNA-->DNA
- In a mammalian cell about 1 million peptide bonds formed per sec
- The basic protein synthesis molecular mechanism has many similarities across all species



Cell cycle

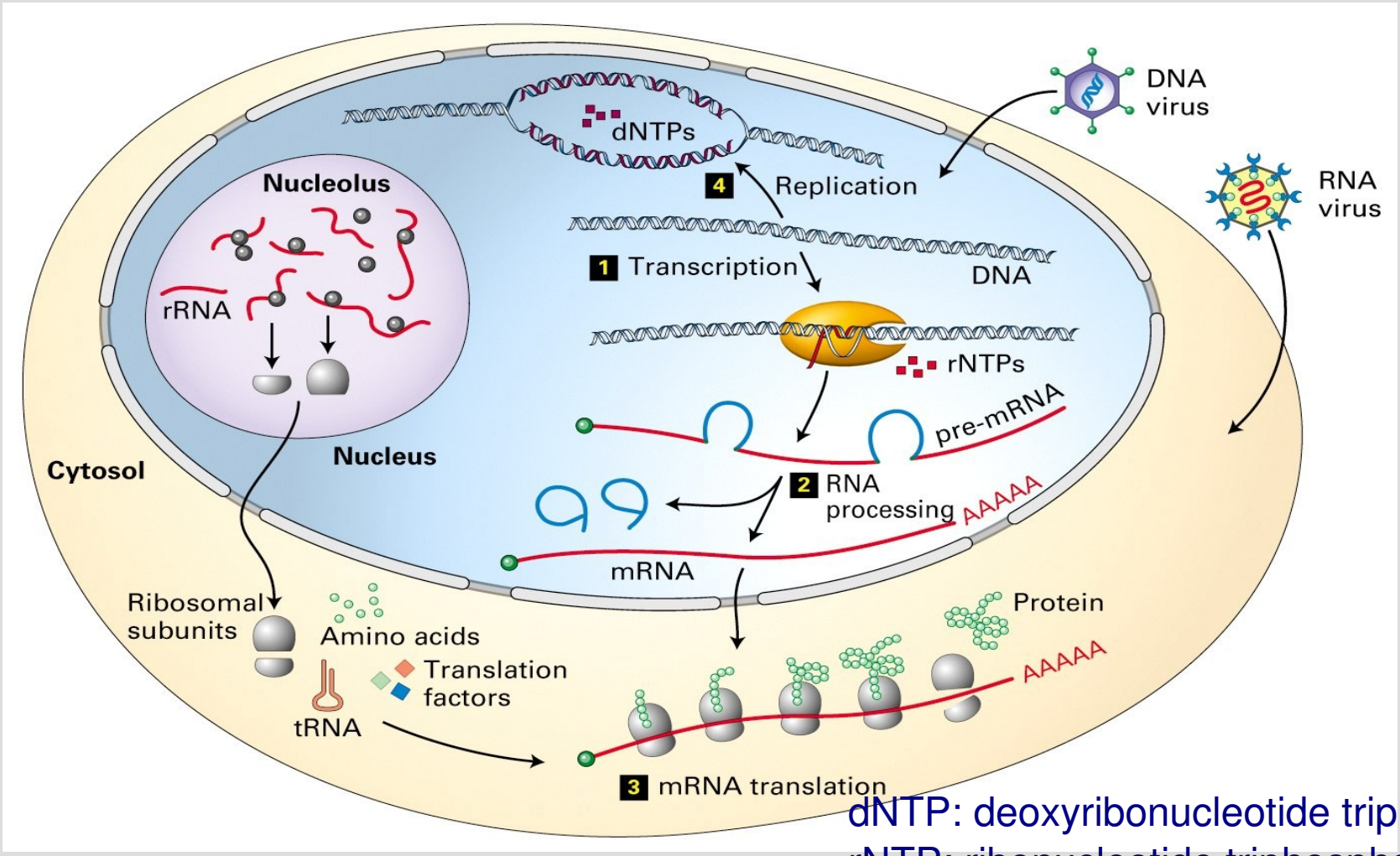


Note major protein synthesis during two phases in the cell cycle:
(a) Prior to cell division, and
(b) During growth and homeostatis phase

One would expect similarities and differences in the proteins being expressed in the two phases

Salvage synthesis widely used for DNA, RNA & proteins

Protein synthesis – a bird's eye view

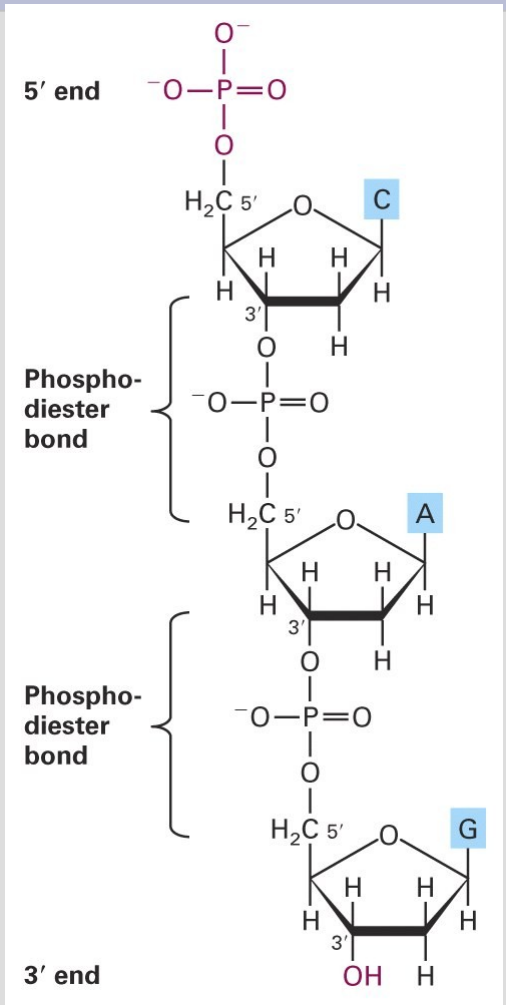


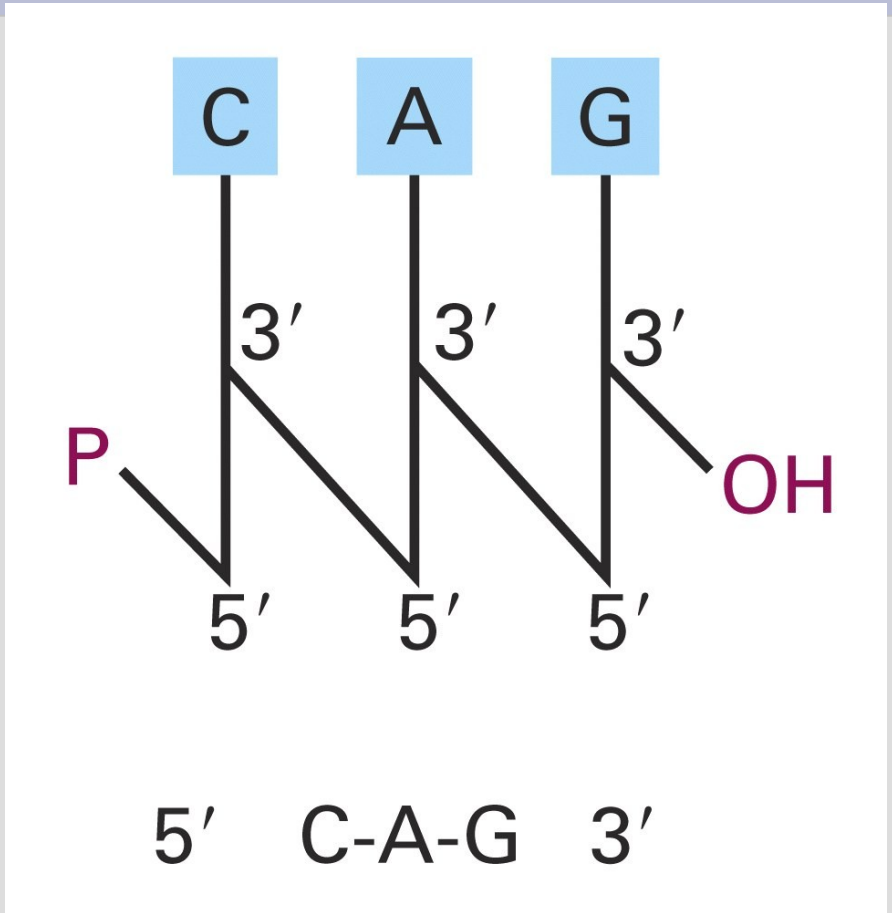
dNTP: deoxyribonucleotide triphosphate
rNTP: ribonucleotide triphosphate

Does it help to not have protein synthesized directly?

- Cell is able to isolate DNA from many enzymes in the cell
- One DNA can serve as template for many mRNA copies – so protein synthesis rate can be modified
- Possibly different segments of the DNA can be used for simultaneous transcription
- RNA requires lower energy for degradation

DNA – two representations

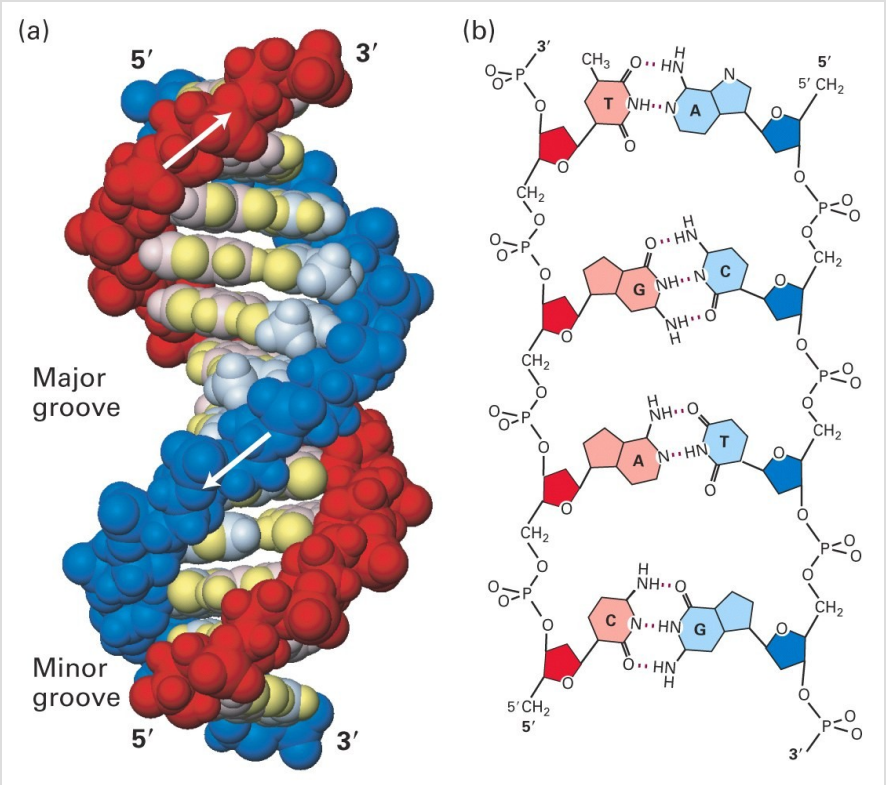




Explain 3' & 5' ends

DNA - cartoon

- Carries genetic code for development & function
- Two strands entwined in a helical fashion with hydrogen bonds between adenine (A) and thymine (T) & guanine (G) & cytosine (C)
- Strands un-entwined for copying



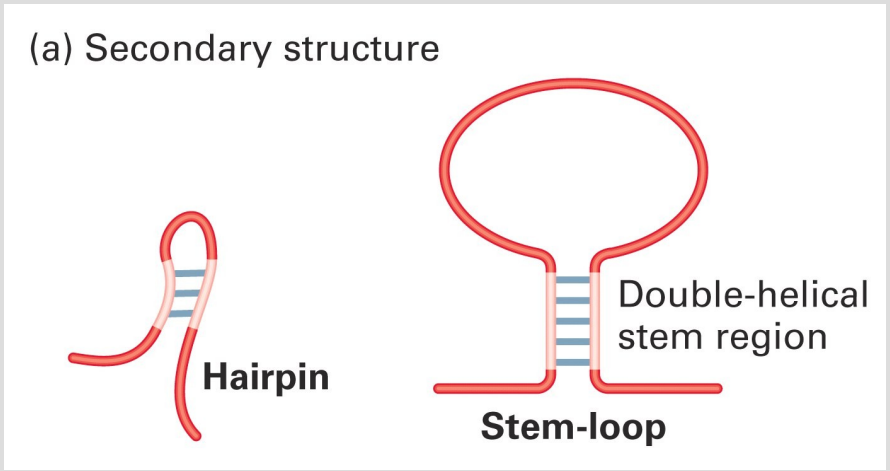
RNA

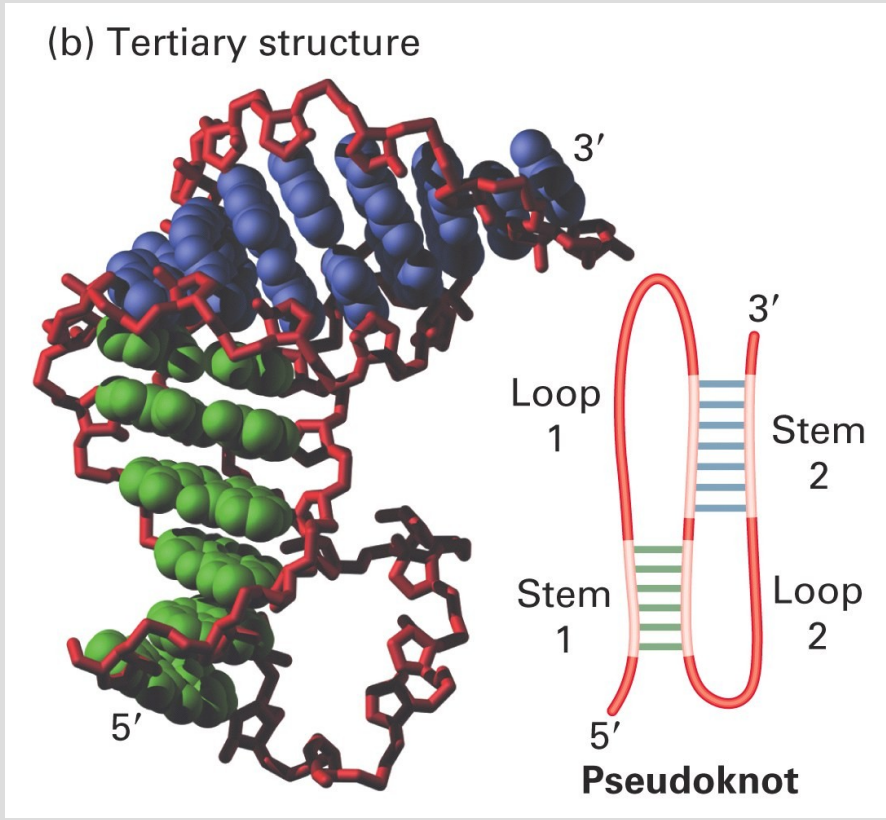
- RNA has several important roles to play in DNA replication and protein expression processes
  - Acts as a messenger (mRNA)
  - Helps to decode the codon (tRNA)
  - Acts as primer to start polymerization in DNA synthesis
  - Short RNA segments called microRNA regulate post transcriptional mRNA

Some comparison between DNA & RNA

- The deoxyribose is replaced by ribose
- Uracil is the pyrimidine in place of thyamine & it hydrogen bonds with adenosine (so one has A-U bonds rather than A-T bonds via hydrogen bonding (two hydrogen bonds again))
- Doesn't form stable double helix
- However hydrogen bonding can form hairpins and stem loops
- The RNA polymer requires a lower energy for degradation than the DNA polymer – more energy efficient for salvage synthesis

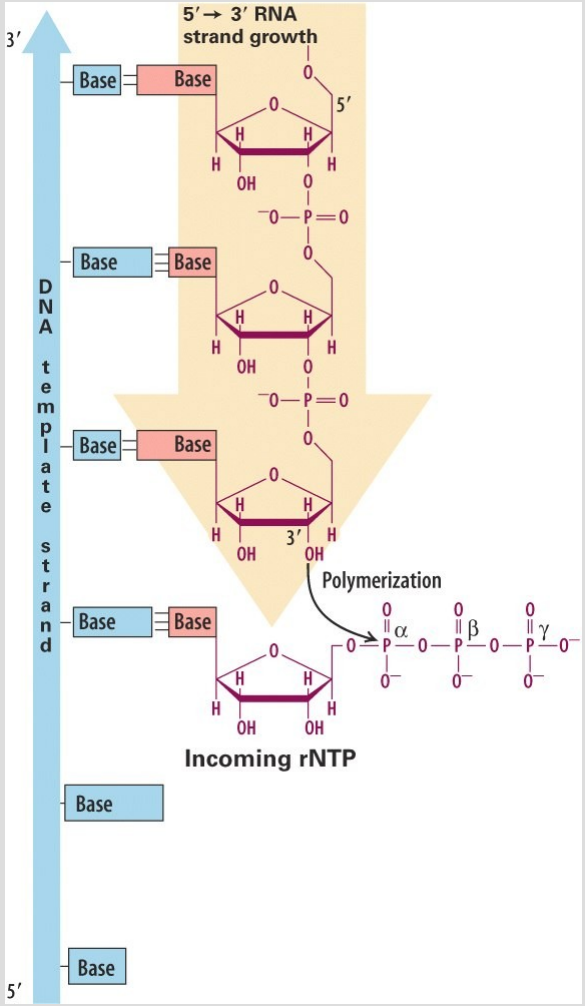
Secondary & tertiary structures





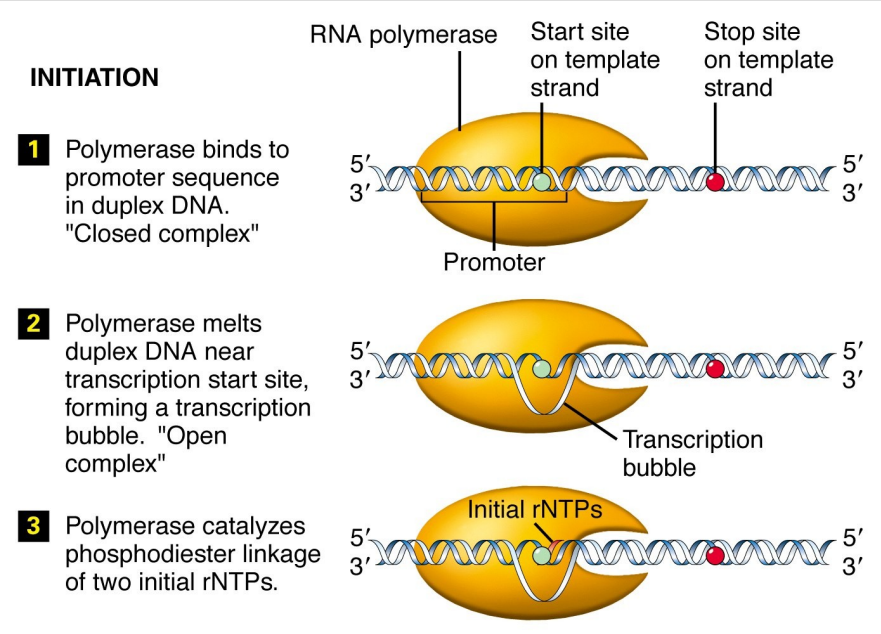
Transcription -1

- Transcription from
  - 3'-5' end of the template
  - 5'->3' end of the strand being synthesized
- How does the process start?
- How does this synthesis done?
- How does it end?

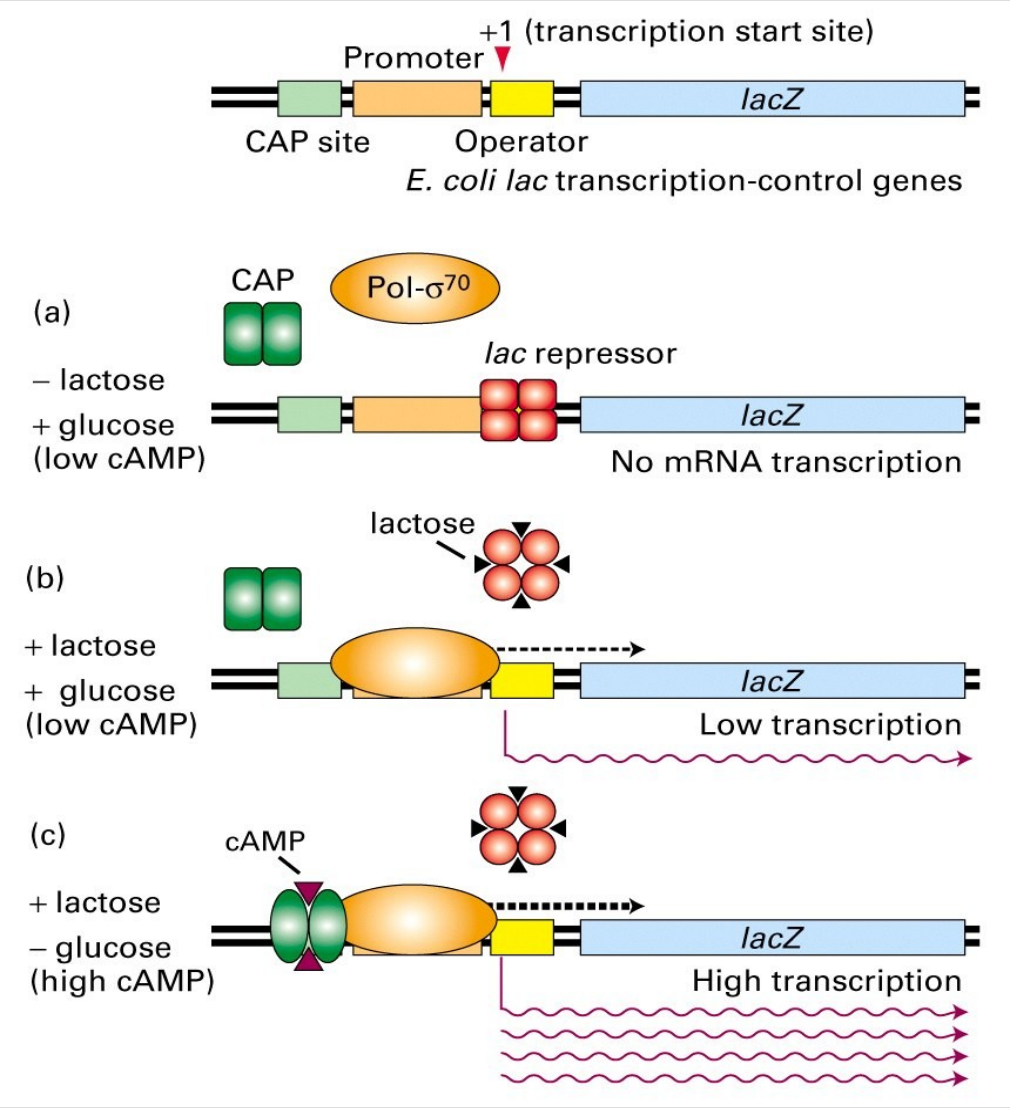


Transcription -2

- Process begins at a promoter
- Only one strand transcribed
- Process catalyzed by RNA polymerase after is binds to the dsDNA
- Polymerization proceeds till the stop site

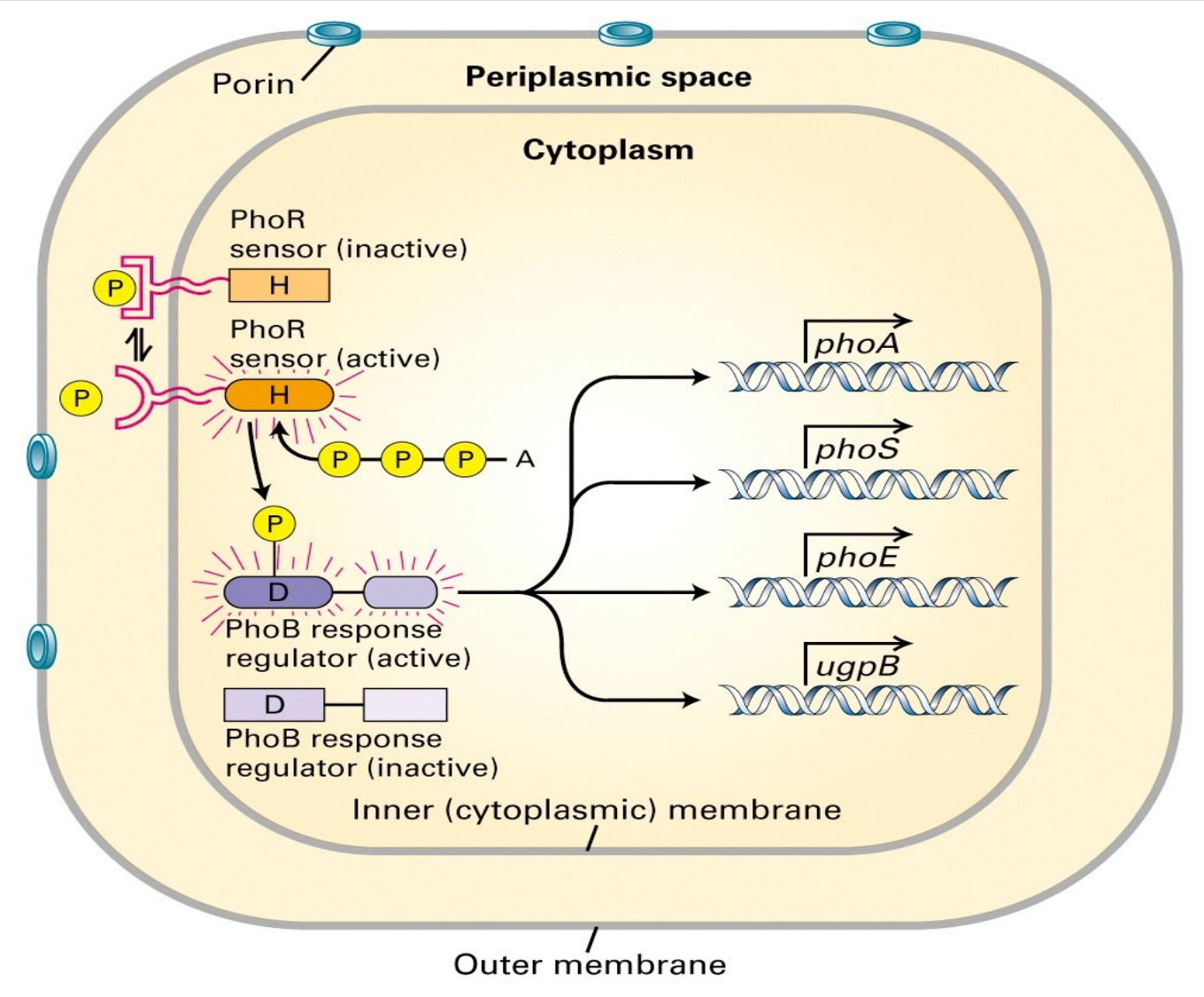


Transcription control



Will do this in more detail when we look at RNA & DNA synthesis in more detail

How process initiated



Again expression control for later

Codons -1

- The amino acid protein sequence stored as a sequence and coding in sets of three nucleotides – the minimum needed to code for 20 amino acids using four nucleotides for coding
- mRNA carries the sequence information as tri-nucleotide codes called codons
- These codons are decoded at the time of protein synthesis

Codons -2

- A codon set of 3 nucleotides could code for 64 amino acids – however only twenty are coded for
- So there is a degeneracy, i.e. more than one code corresponds to one amino acid
- How is first codon identified? And how is a sequence terminated?
- AUG is a codon for methionine and is also the “start” codon
- UAA, UAG and UGA do not code for any specific amino acid but act a “stop” codons
- Therefore there are 61 codons that code for amino acids

Codons -3

TABLE 4-1 The Genetic Code (RNA to Amino Acids)*

First Position (5′ end)	Second Position U	Second Position C	Second Position A	Second Position G	Third Position (3′ end)
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu (Met)*	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met (start)	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val (Met)*	Ala	Glu	Gly	G

*AUG is the most common initiator codon; GUG usually codes for valine, and CUG for leucine, but, rarely, these codons can also code for methionine to initiate a protein chain.

Codons conserved across species

Codons -4

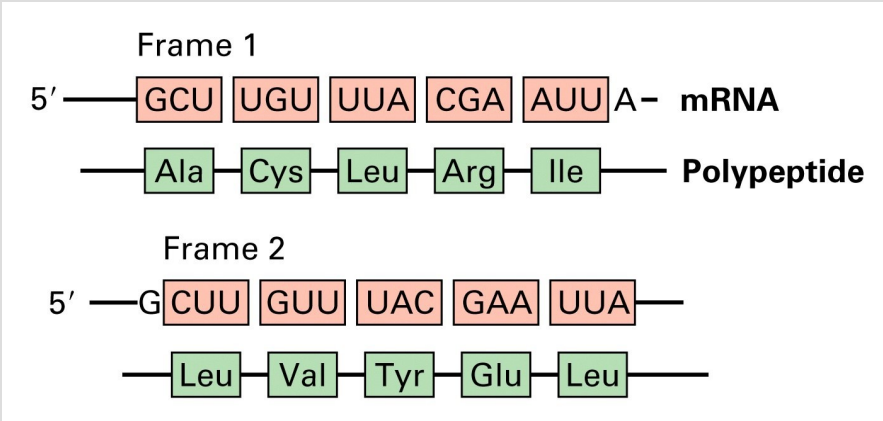
TABLE 4-2 Known Deviations from the Universal Genetic Code

Codon	Universal Code	Unusual Code*	Occurrence
UGA	Stop	Trp	*Mycoplasma, Spiroplasma,* mitochondria of many species
CUG	Leu	Thr	Mitochondria in yeasts
UAA, UAG	Stop	Gln	*Acetabularia, Tetrahymena,* Paramecium, etc.
UGA	Stop	Cys	*Euplotes*

*"Unusual code" is used in nuclear genes of the listed organisms and in mitochondrial genes as indicated.
SOURCE: S. Osawa et al., 1992, *Microbiol. Rev.* **56**:229.

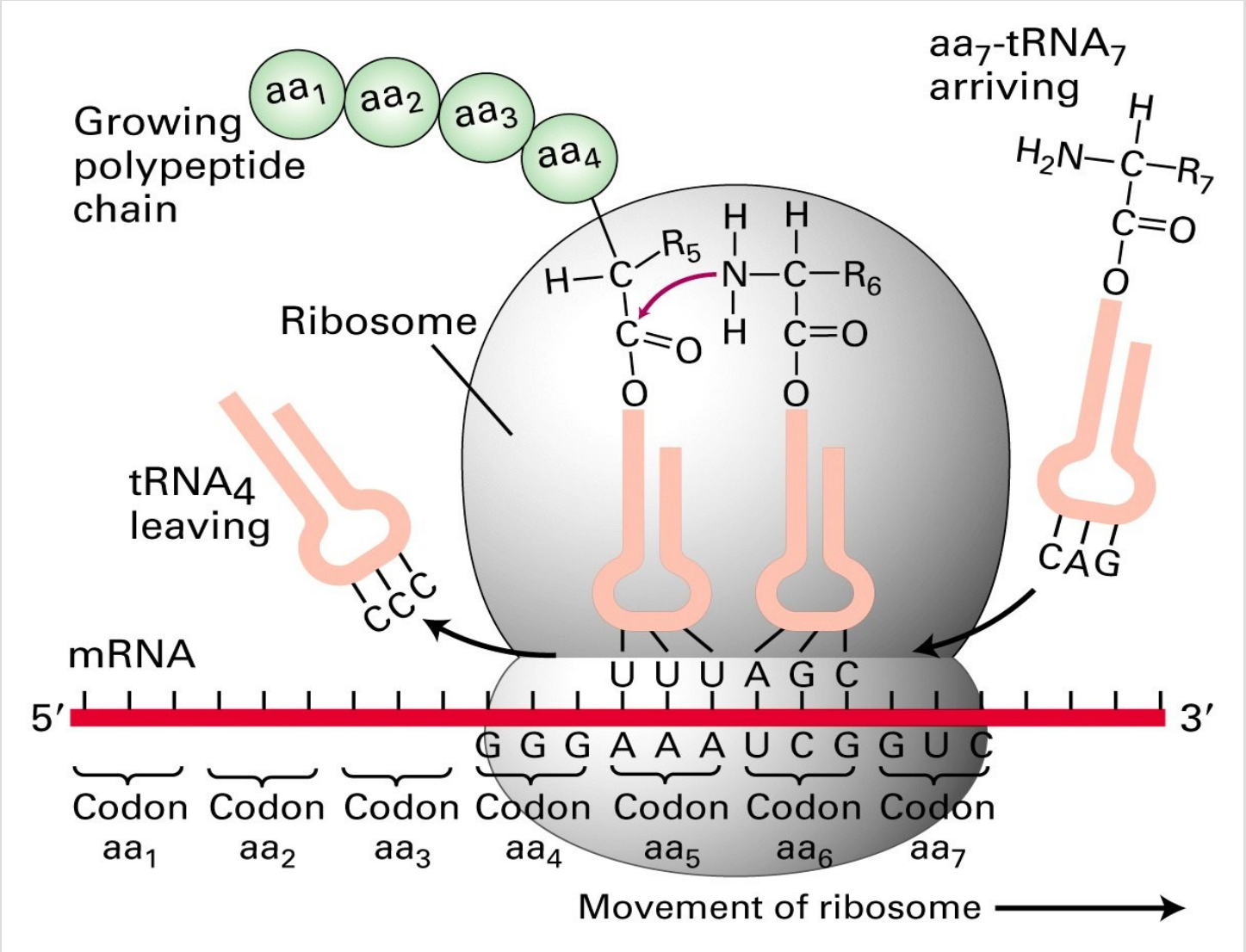
Codon needs to be translated -- how?

- mRNA acts as the messenger template carrying the code for the protein to be synthesized
- The translation is done in ribosomes in which tRNA decode the sequence and attach the appropriate amino acid



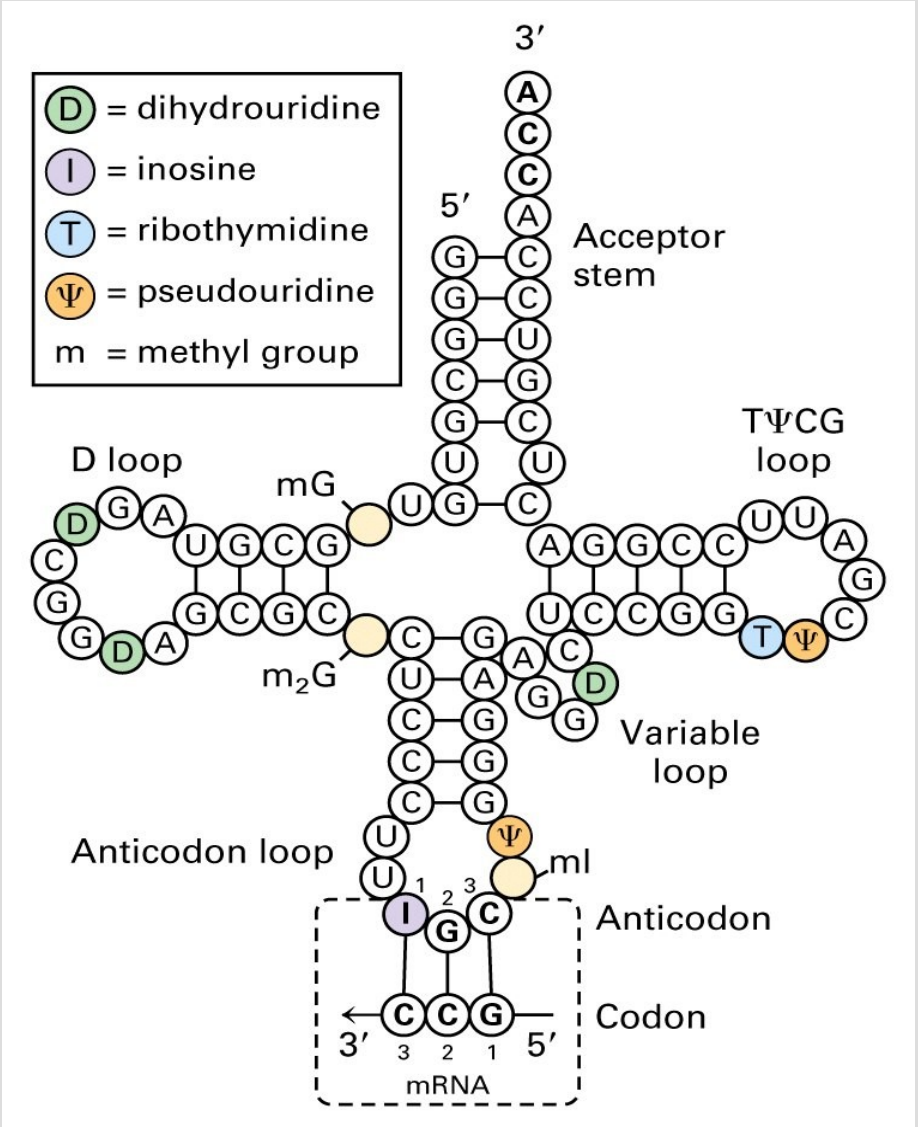
Ribosomes

After ribosome attachment around mRNA, tRNA with complementary anti-codon & appropriate amino acid attached attach and the next amino acid in peptide is fused



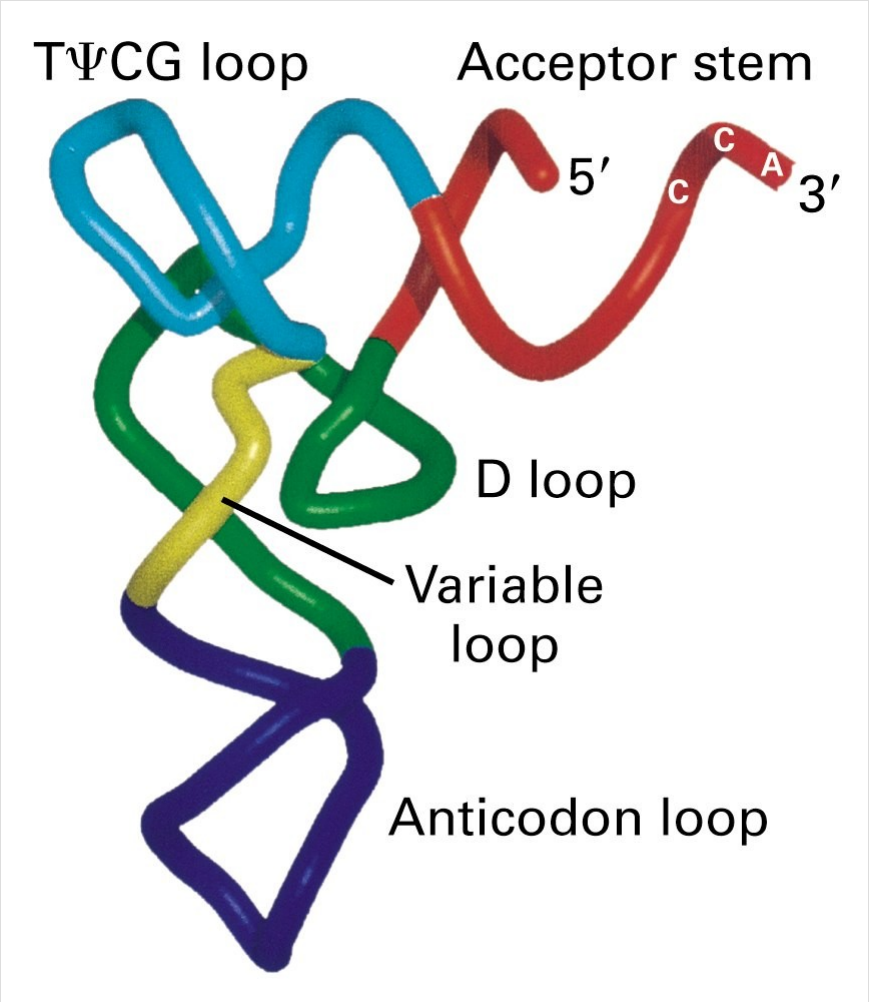
tRNA

Codon deocded by base pairing interactions of codon in the mRNA with anti-codons in tRNA



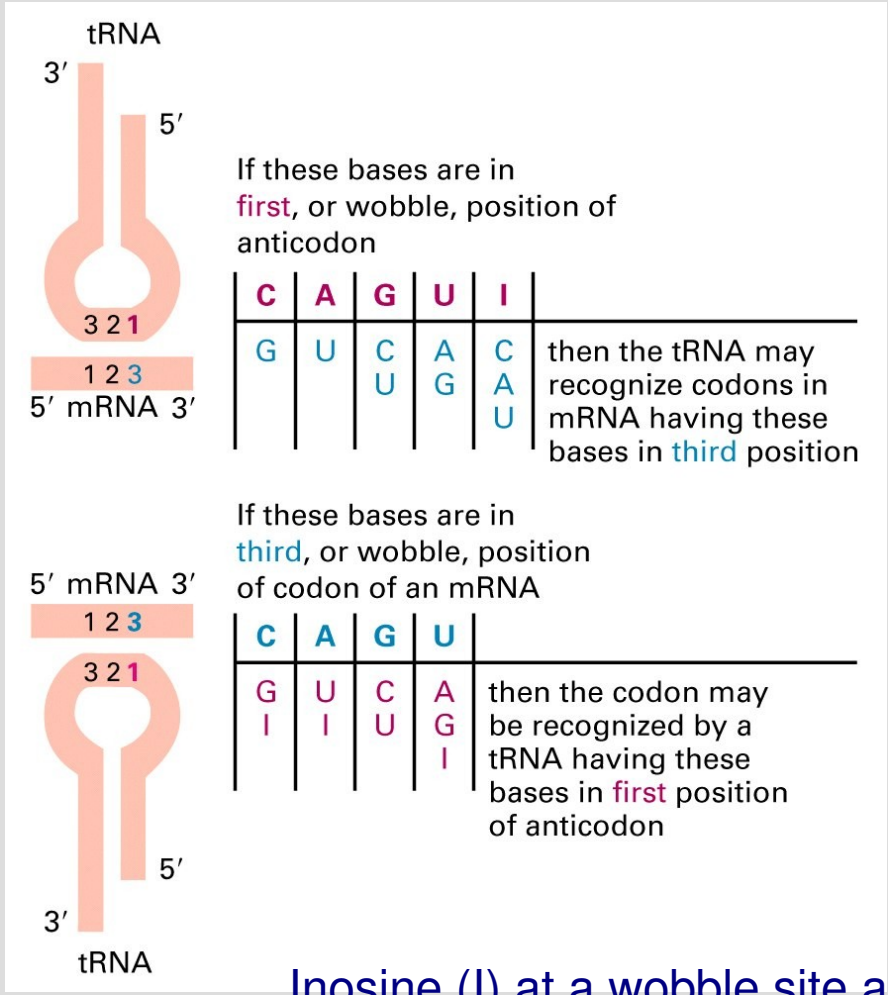
tRNA

- Note the anti-codon is a three dimensional surface – this enables more functional flexibility in codon recognition



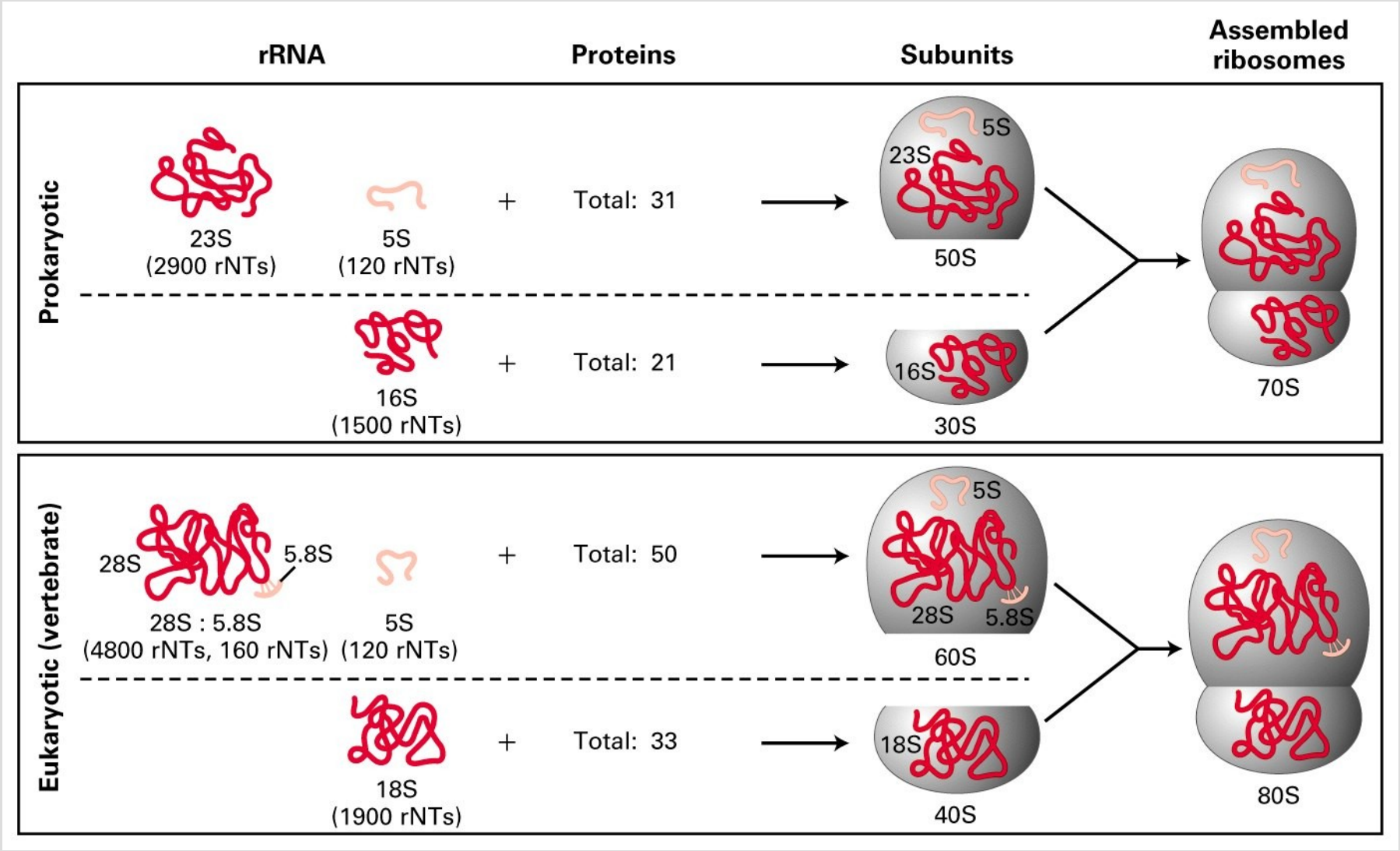
tRNA & wobble

- How the degenerate code is mapped use of wobble reduces the need to have 61 different tRNA sets
- The third base in the codon and the first in the anti-codon are called the wobble position and G,U & I enable more than one base to be recognized in the wobble position in the codon

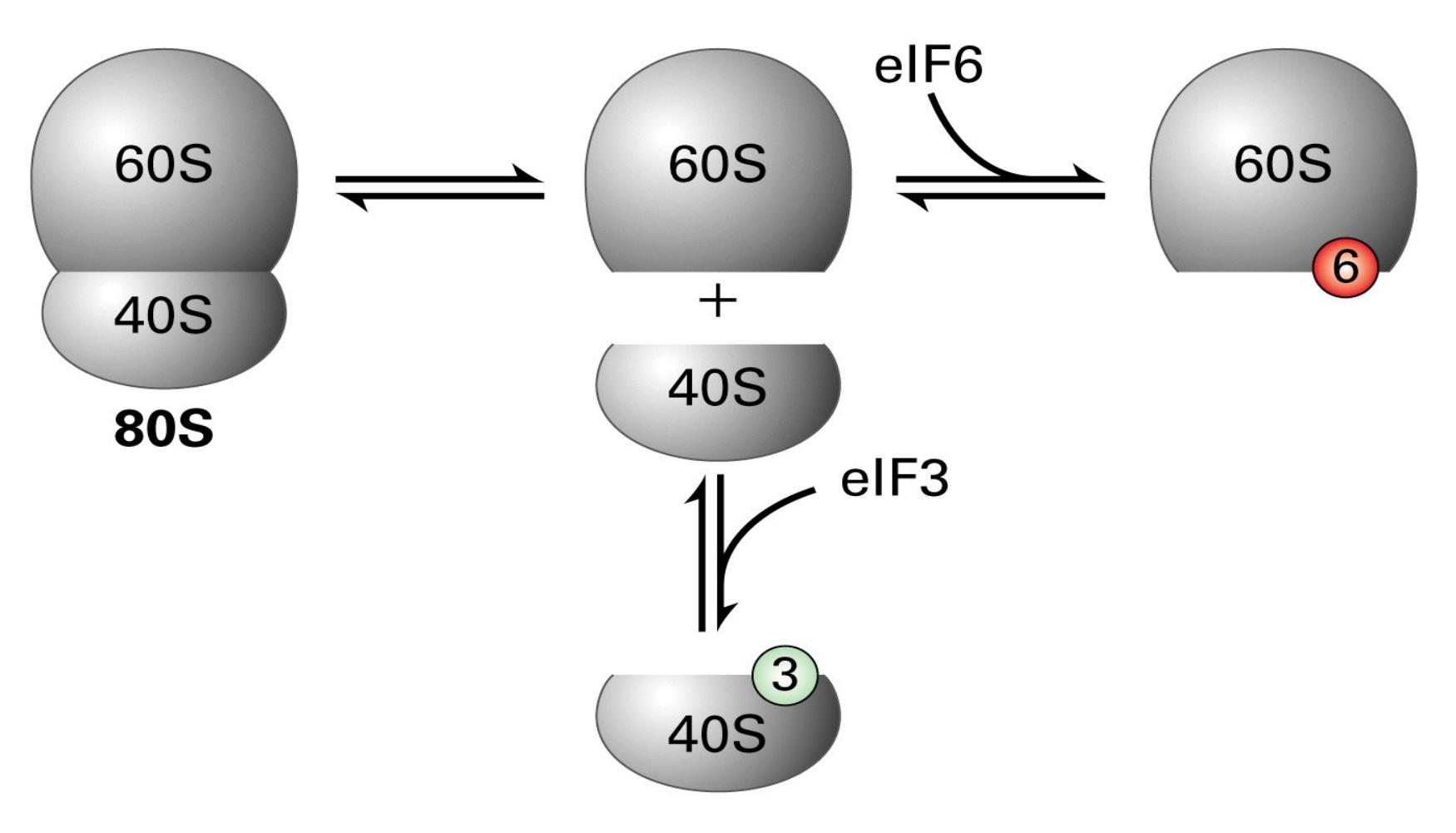


Inosine (I) at a wobble site acts as a “don't care” recognition site

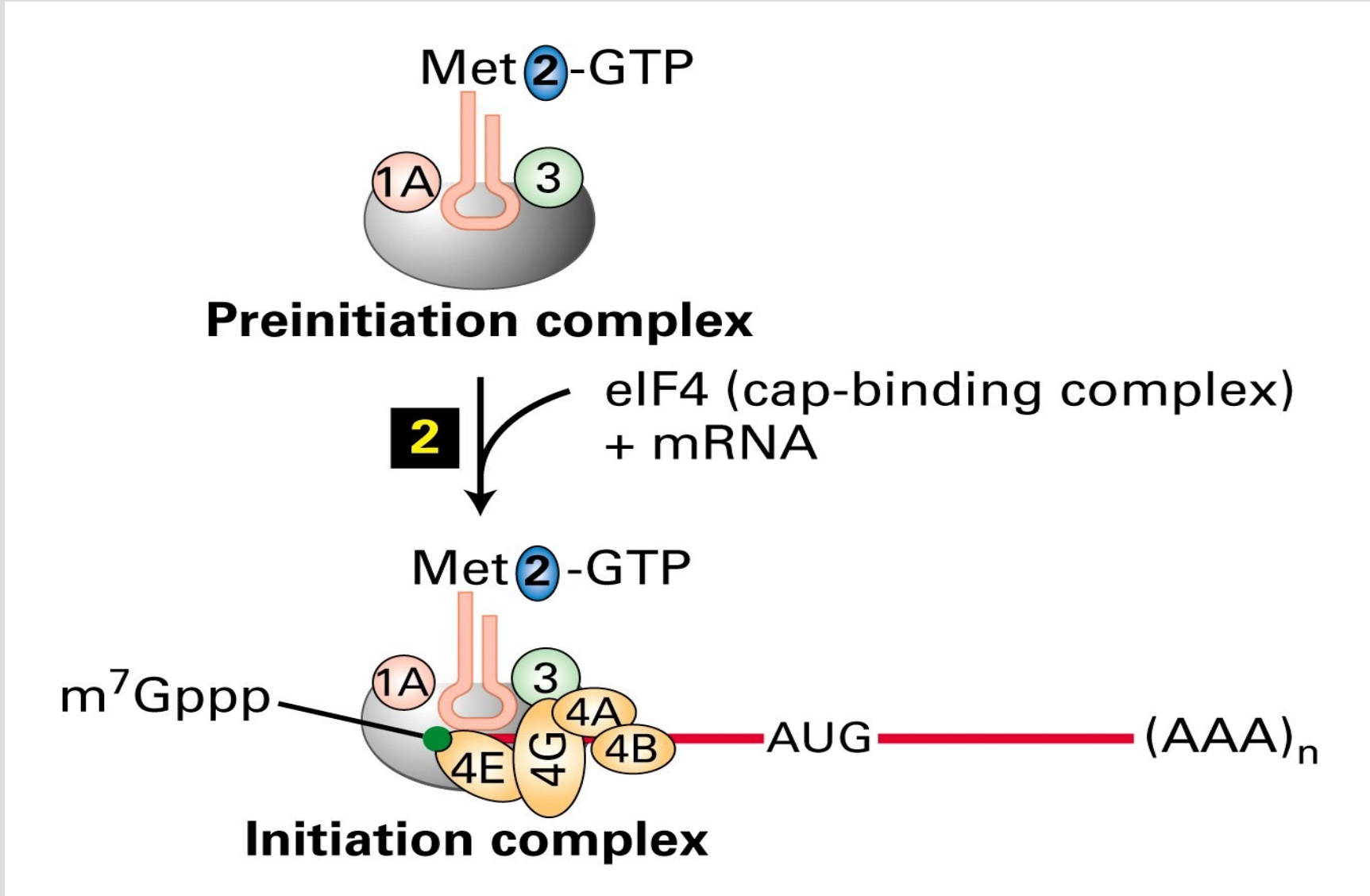
Ribosome complexity



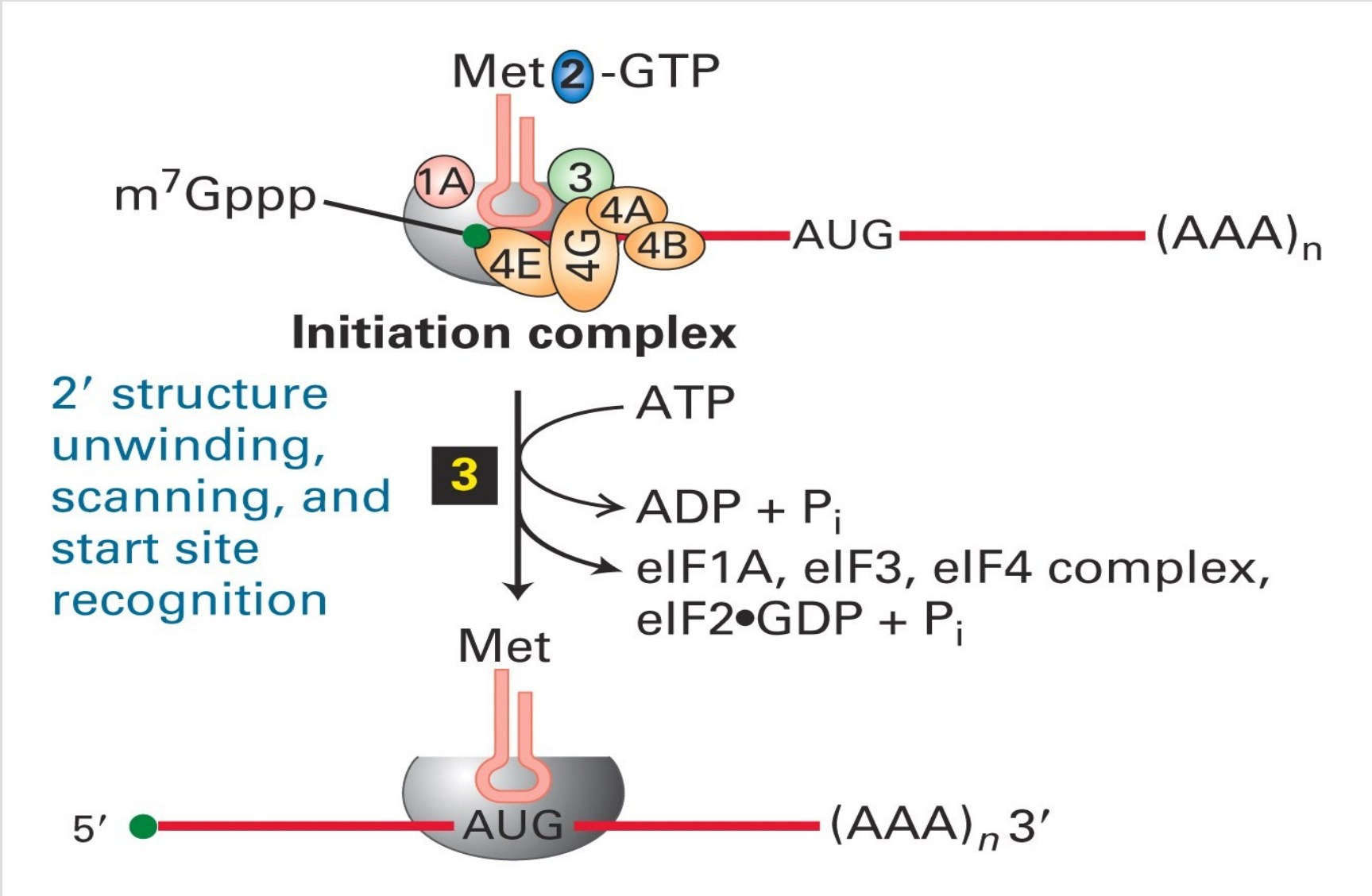
How do we know there are these parts?



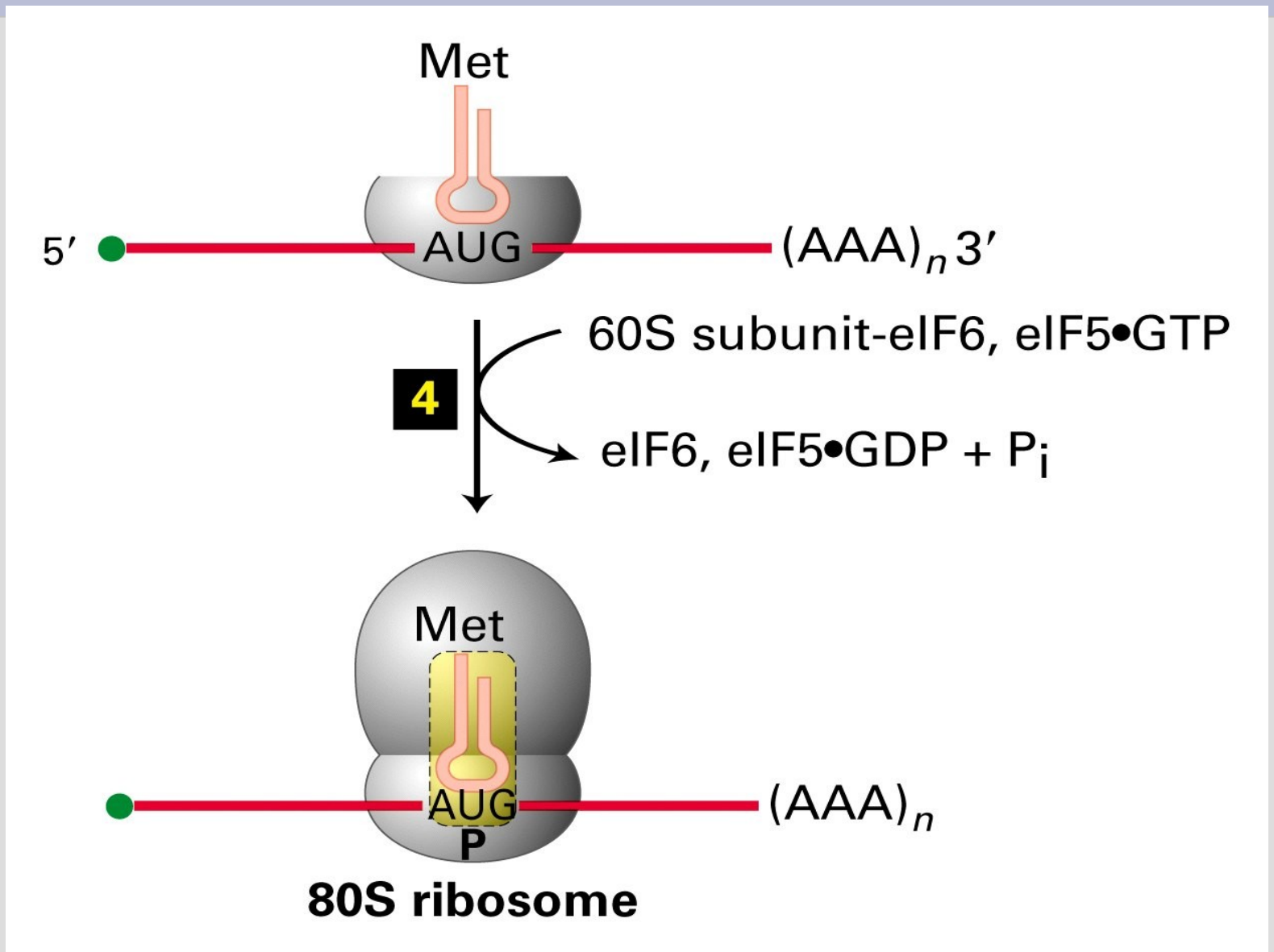
Ribosome attachment process



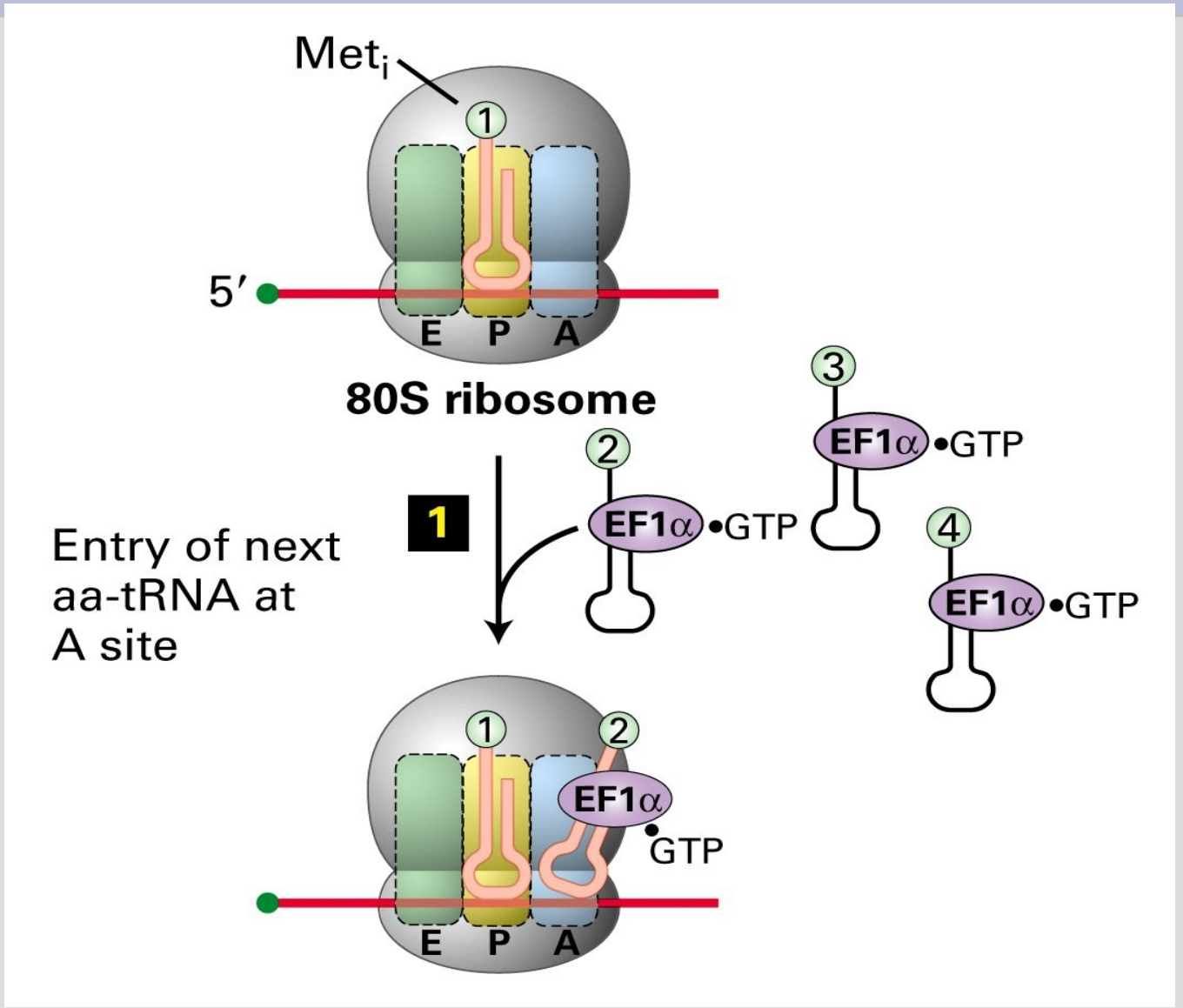
Ribosome attachment



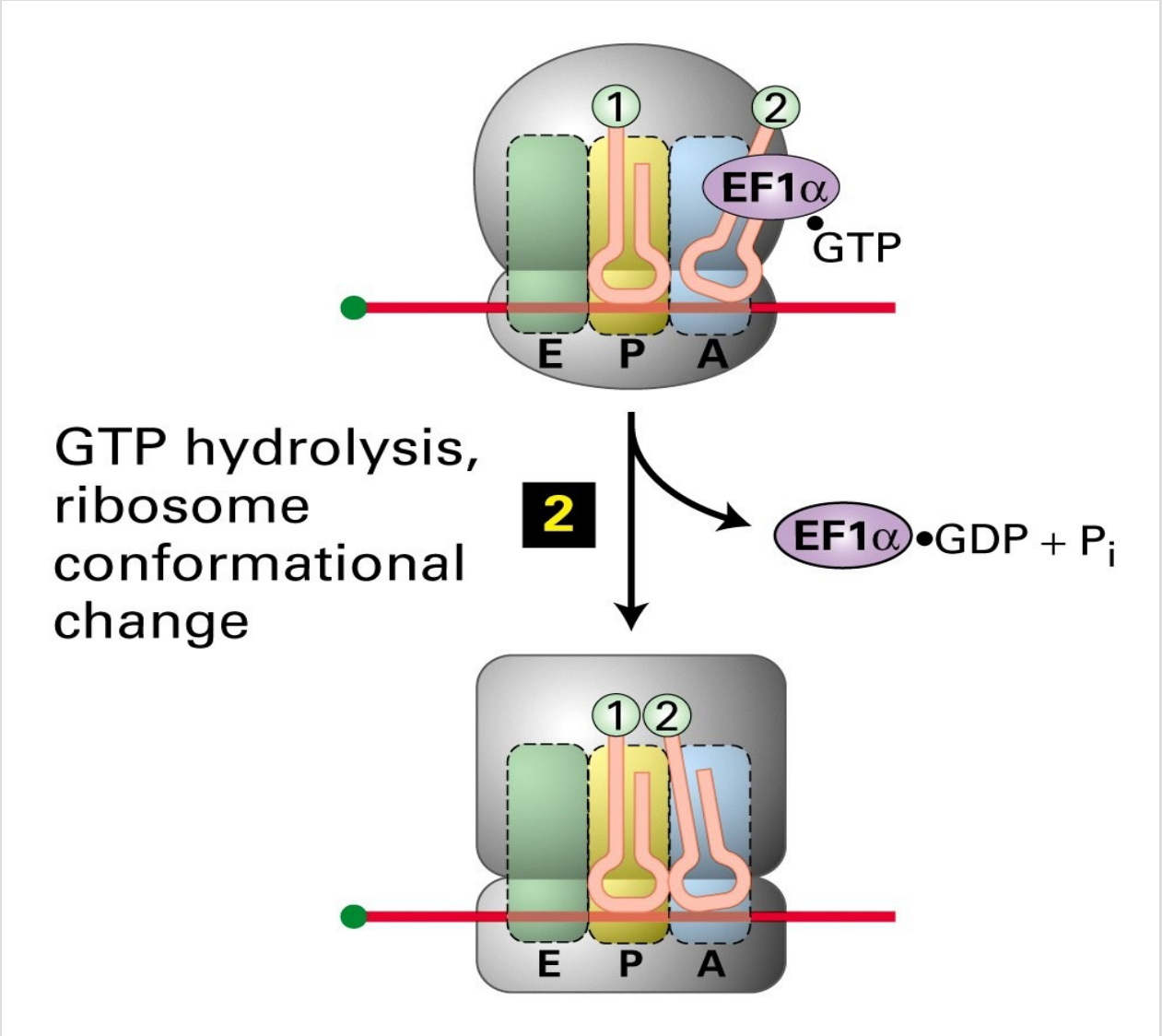
Ribosome complete attachment



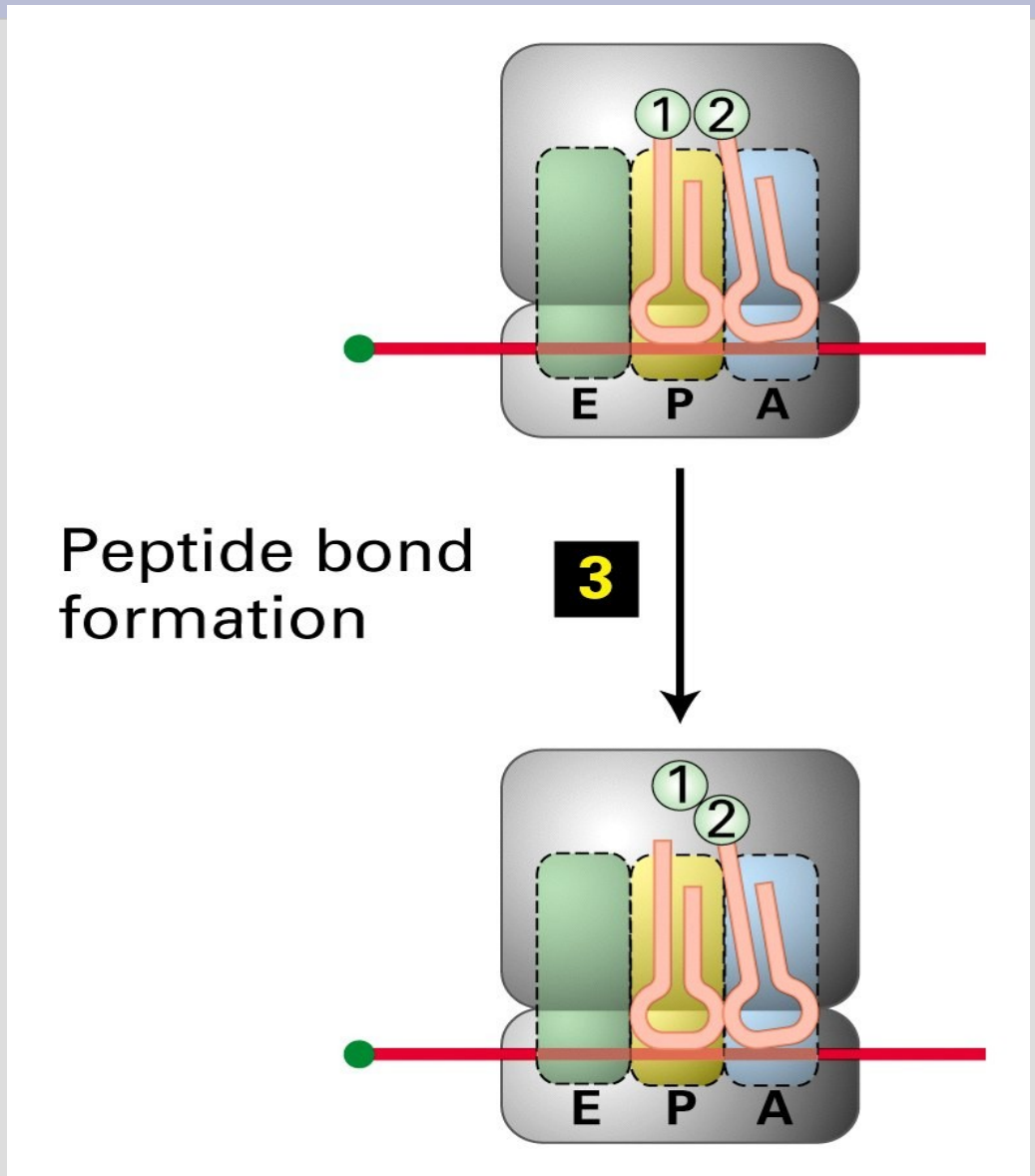
Translation process -1



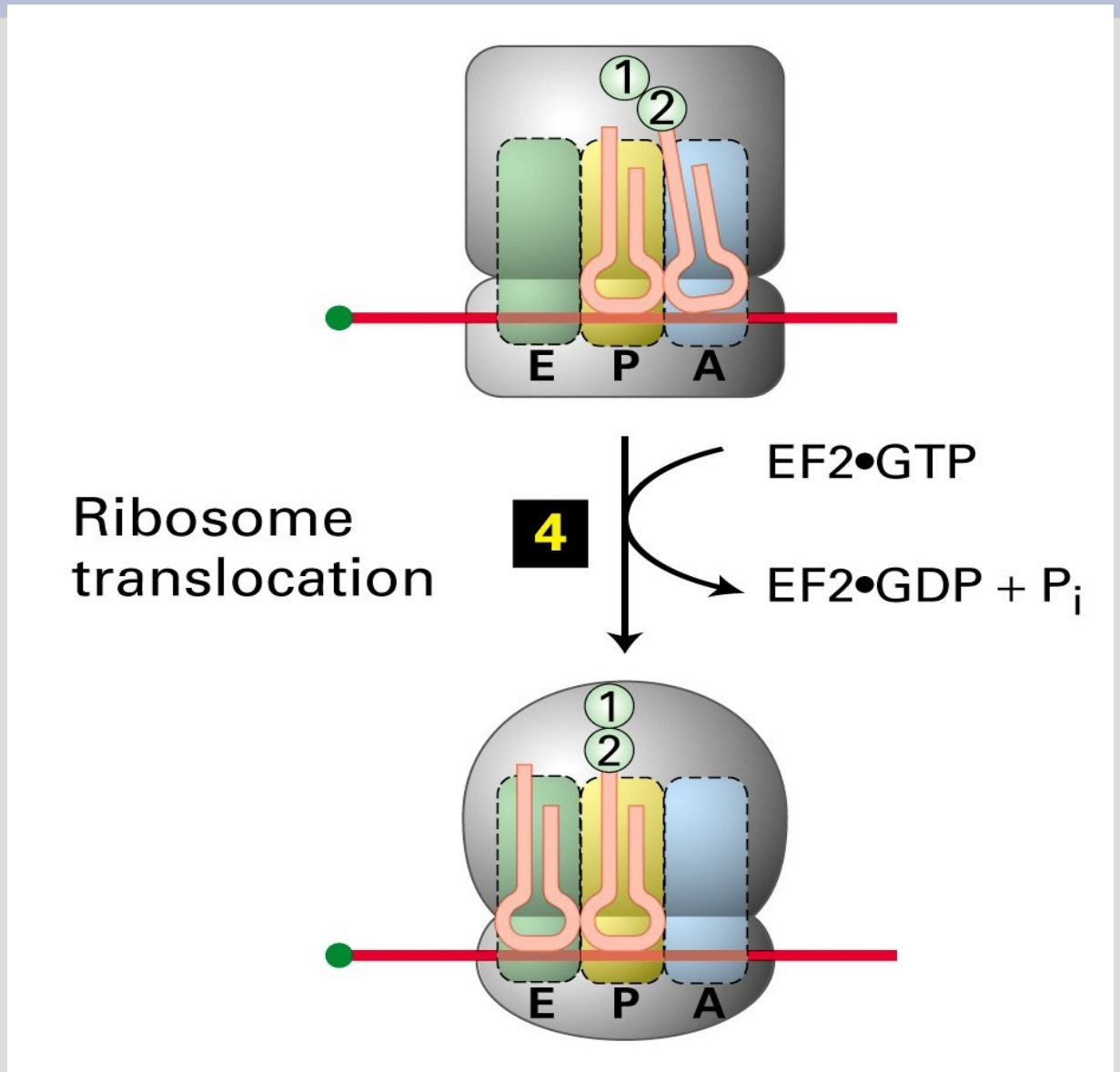
Translation -2



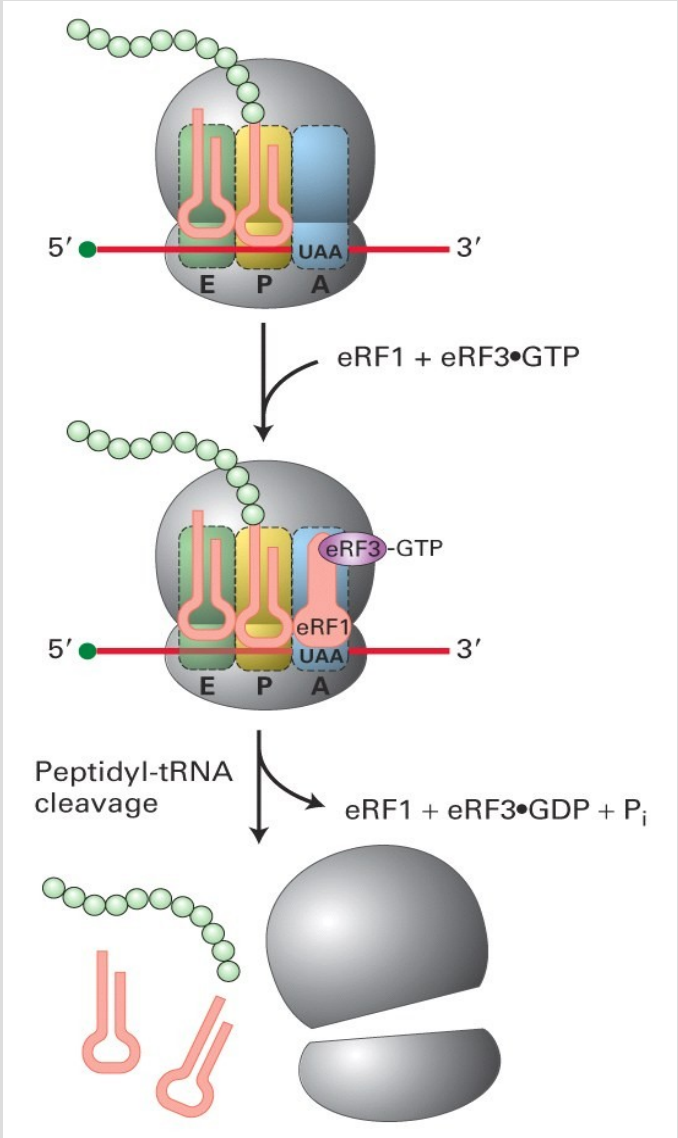
Translation -3



Translation -4

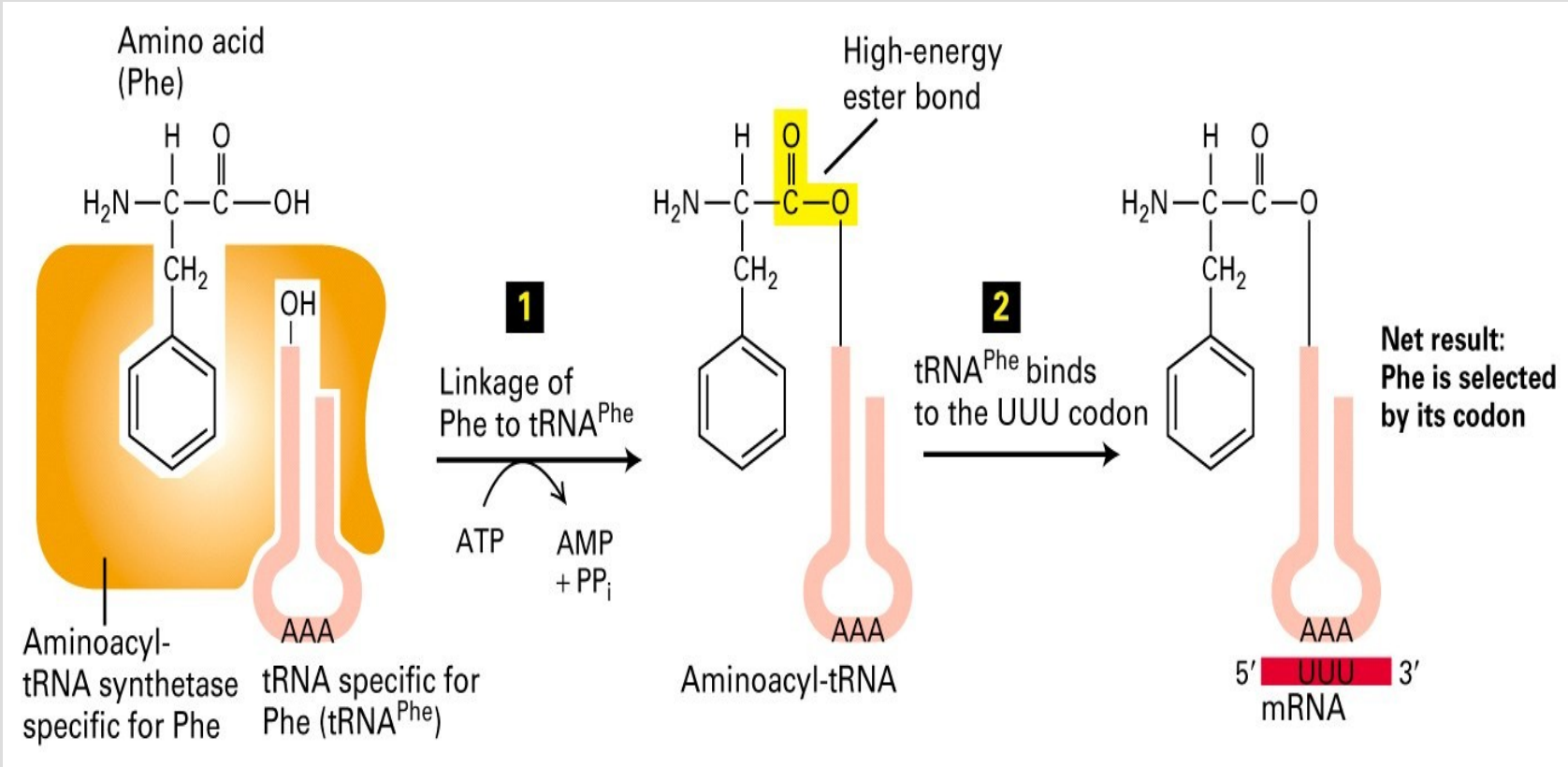


Ribosome detachment



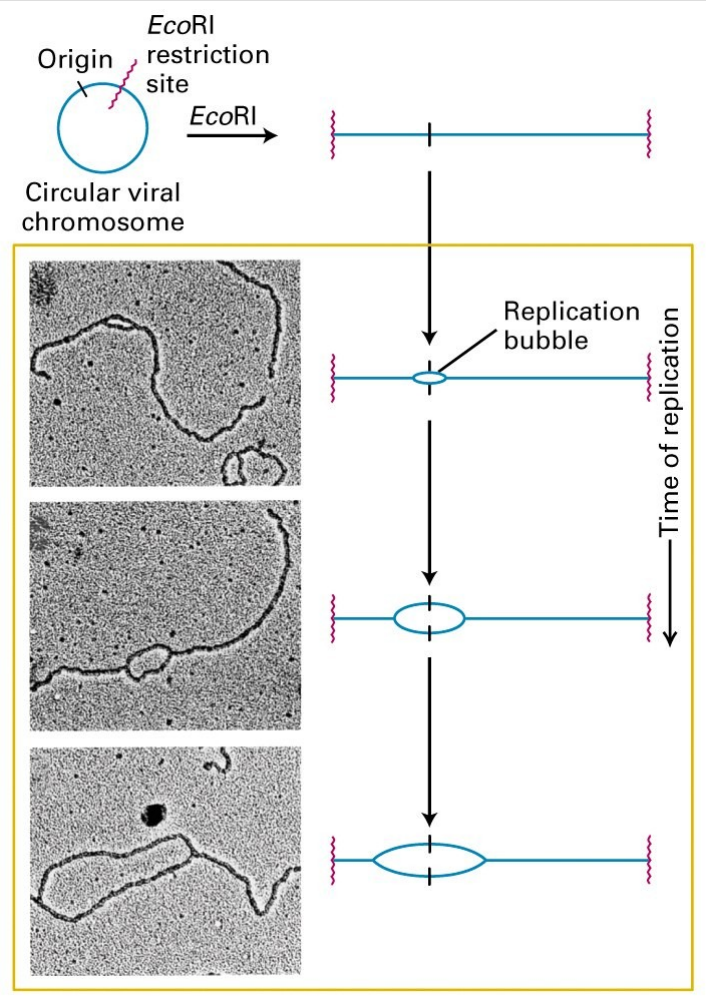
tRNA synthesis

[matched attachment codon-amino acid via tRNA synthetase]



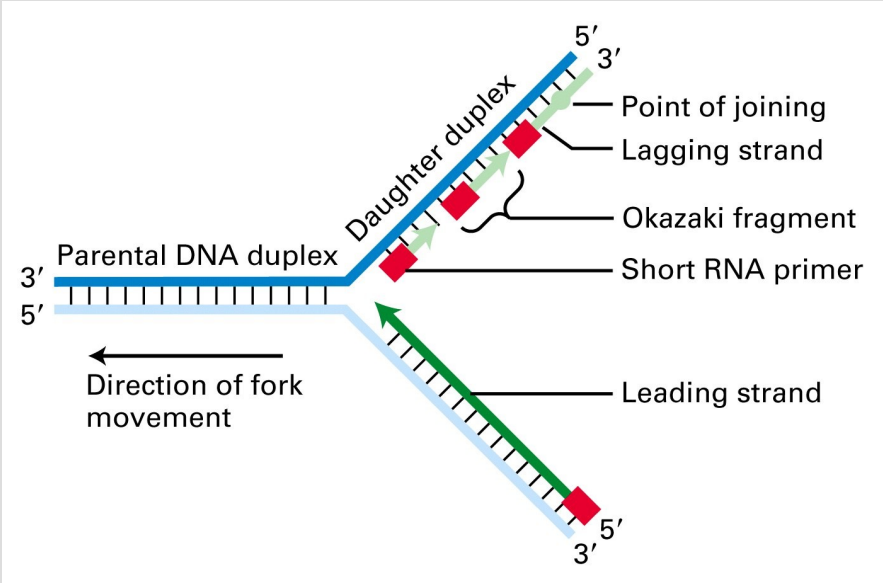
Quick mention of DNA replication

- Replication a more complex process than transcription
- Replication begins at replication bubbles and moves through the genome
- Unpacking & packing of DNA important for eukaryotic cells



Replication fork & Okazaki fragments

- Replication always 3'-5' along the template
- Need of a primer and RNA serve as primer and need to be excised
- So the two daughter strands have slightly different mechanisms for formation – the lagging strand forming via a series of short segments that are subsequently joined



Summary

- Protein synthesis & formation essentially via six important sets of molecular machines
  - Transcription of mRNA (or pre-mRNA -- only for eukaryotes)
  - Splicing of pre-mRNA to form mRNA (only eukaryotes)
  - microRNA based regulation
  - Diffusion of mRNA to ribosomes and decoding of codons & synthesis of proteins there in
  - Chaperon enabled folding of proteins
  - Packaging and transport of proteins
  - Degradation of malformed proteins and mRNA after use

Summary

- Need to still understand, how is the protein synthesis process triggered?
- How are the appropriate segments identified for transcription?
- How are the transcription and translation processes terminated?