Nigel Sussman (Science (2019), 366, 6467)

Translation

Heeseon An/Alban Ordureau

Translation: from mRNAs to Proteins

1) Components needed for translation

- -mRNA
- -tRNA
- -ribosomes
- -initiation, elongation, and termination factors

2) Roles of these players in ensuring accuracy during the initiation, elongation, termination and recycling steps of the translation process

3) Quality control mechanisms:

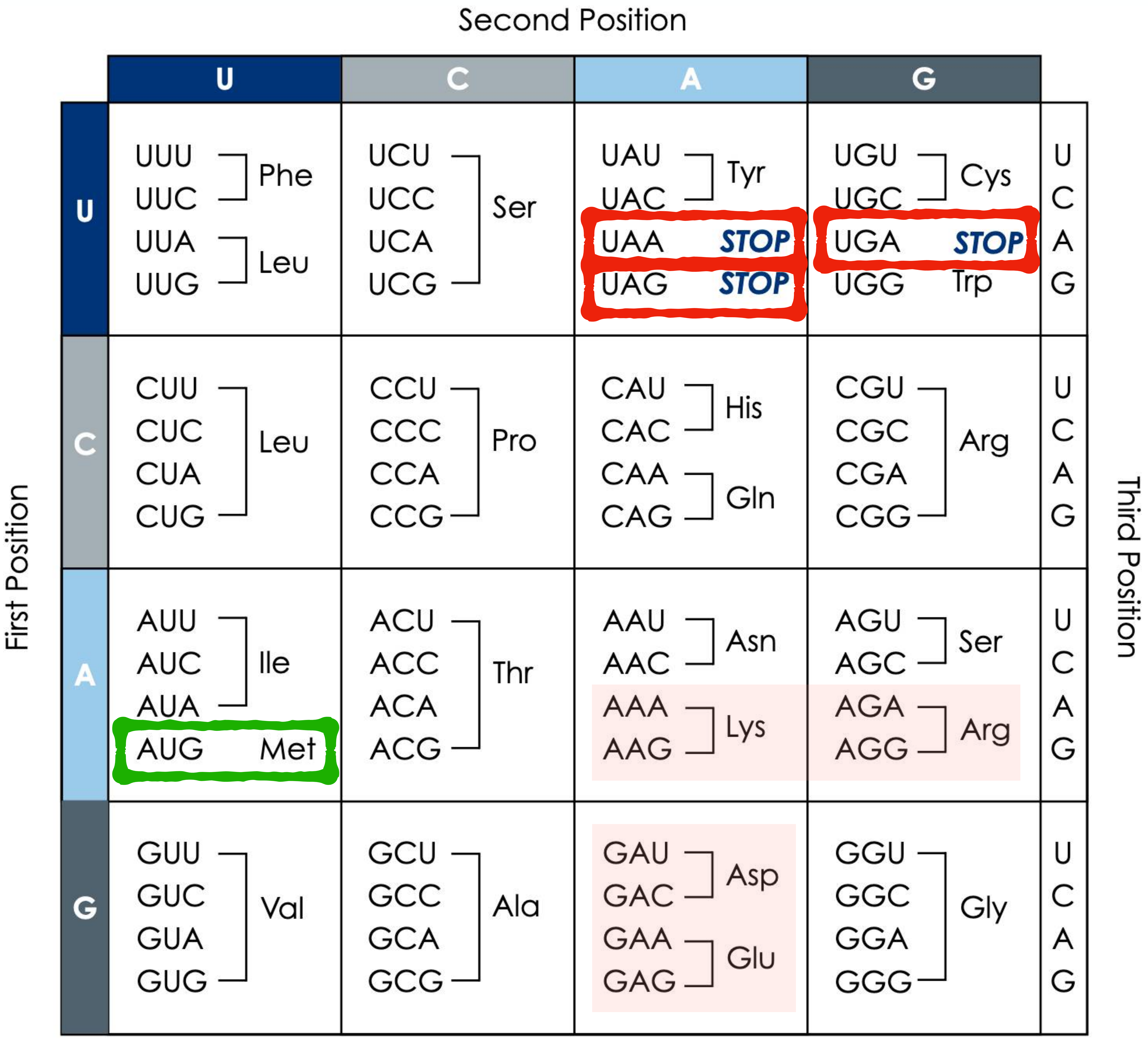
- -How defective mRNAs are detected by ribosomes
- -Degradation of faulty RNA and incompletely translated protein product
- -recycling of the ribosome components

4) Relationship between protein levels and mRNA abundance?

Overview of eukaryotic translation

The Genetic Code

Set of rules that tells us how to go from one alphabet (nucleotides) to another (amino acids)



Second Position

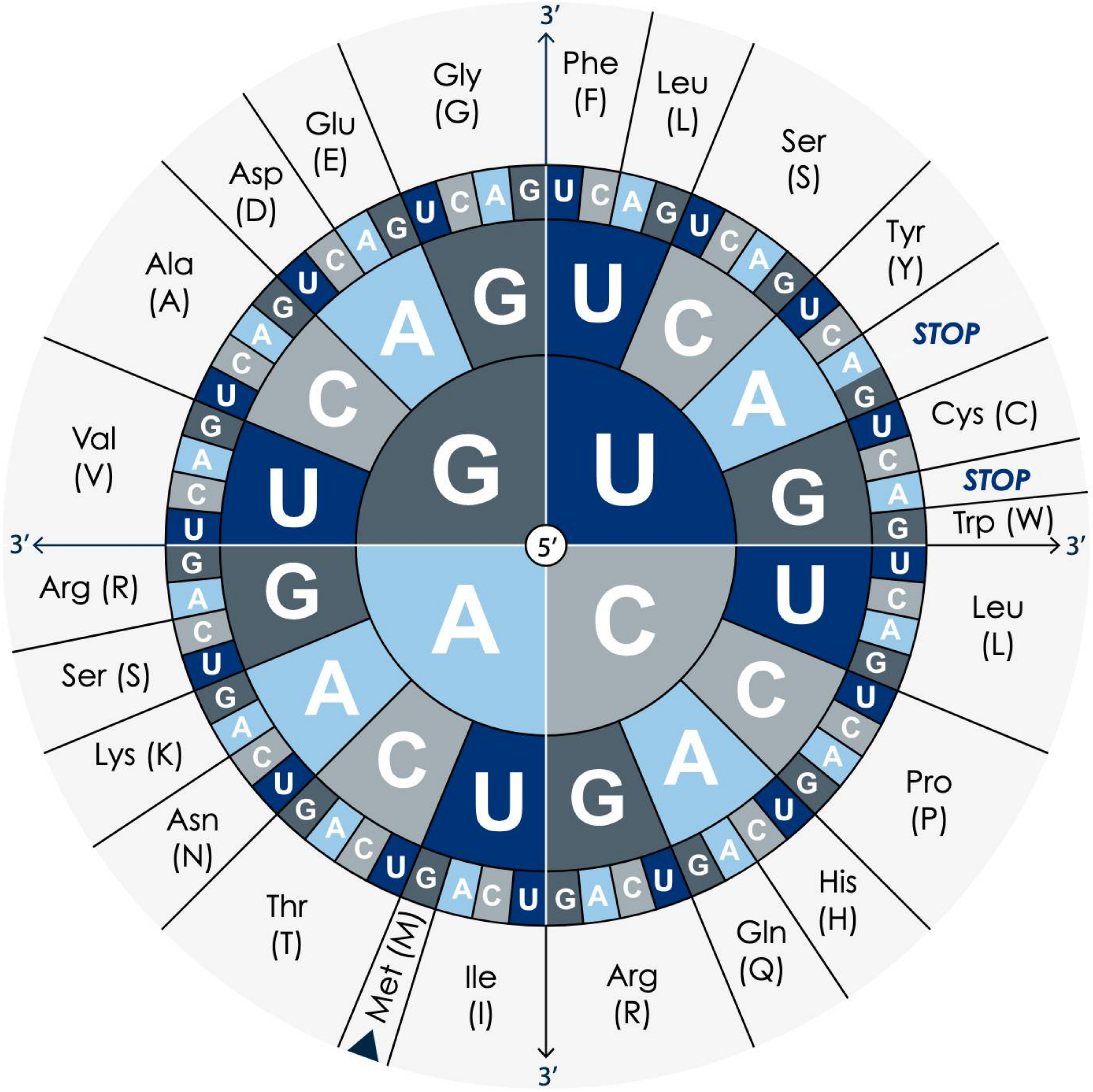
First Position	U	C	A	G	Third Position
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC	UCC	UAC	UGC	C
	UUA Leu	UCA	UAA STOP	UGA STOP	A
	UUG	UCG	UAG STOP	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC	CCC	CAC	CGC	C
	CUA	CCA	CAA Gln	CGA	A
	CUG	CCG	CAG	CGG	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC	ACC	AAC	AGC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

- Need be able to encode a set of 20 amino acids (AA)
- Set of four different building blocks
- 2-letter code does not cover the 20 AA ($4^2 = 16$)
- 3-letter code comprehensively covers the 20 AA complexity ($4^3 = \mathbf{64}$)

- Redundant (AA covered by more than one codon)
- Non-ambiguous (all 64 codons encode something)
- Conservative (mutations can have relatively minor consequences)

The Genetic Code

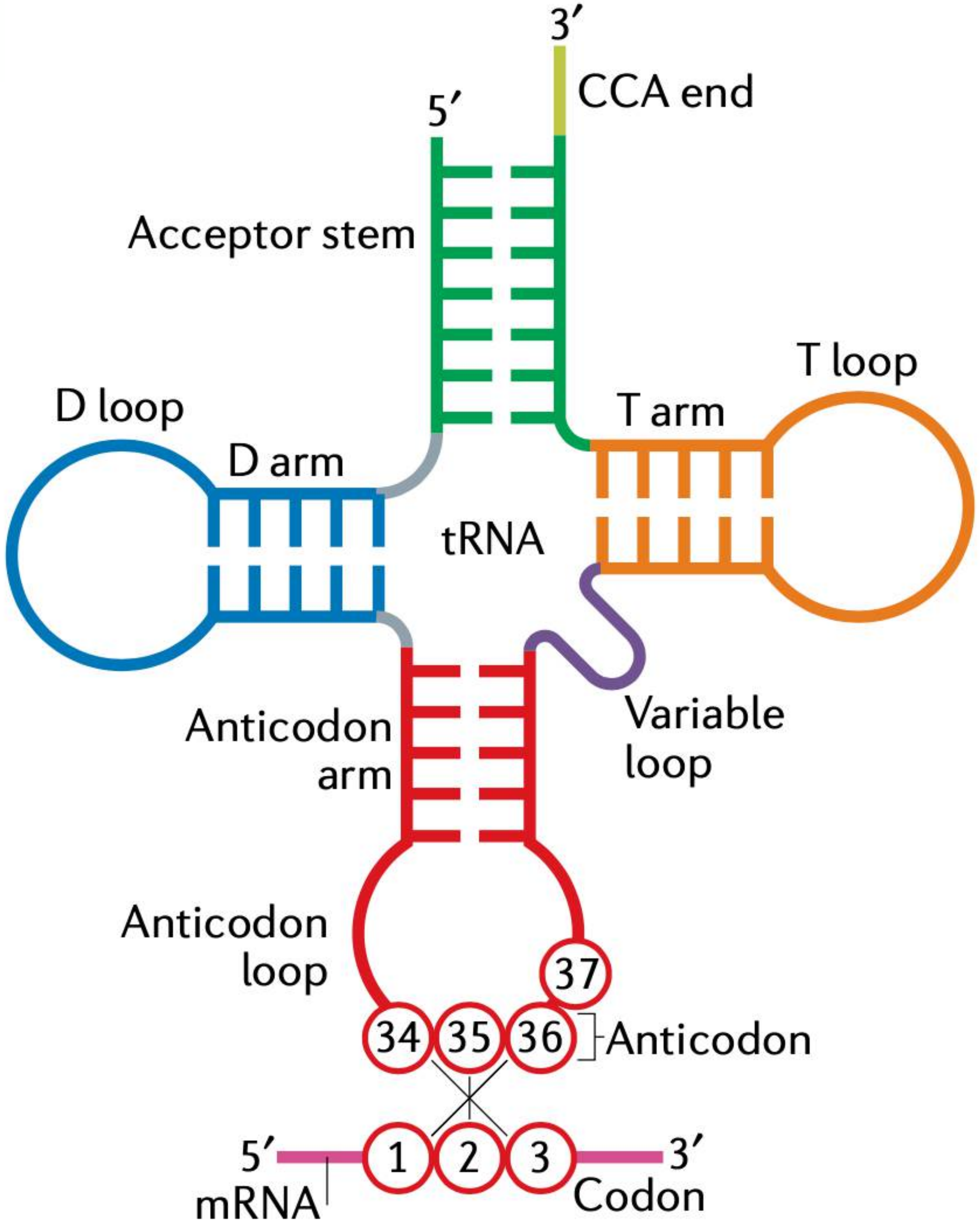
Set of rules that tells us how to go from one alphabet (nucleotides) to another (amino acids)

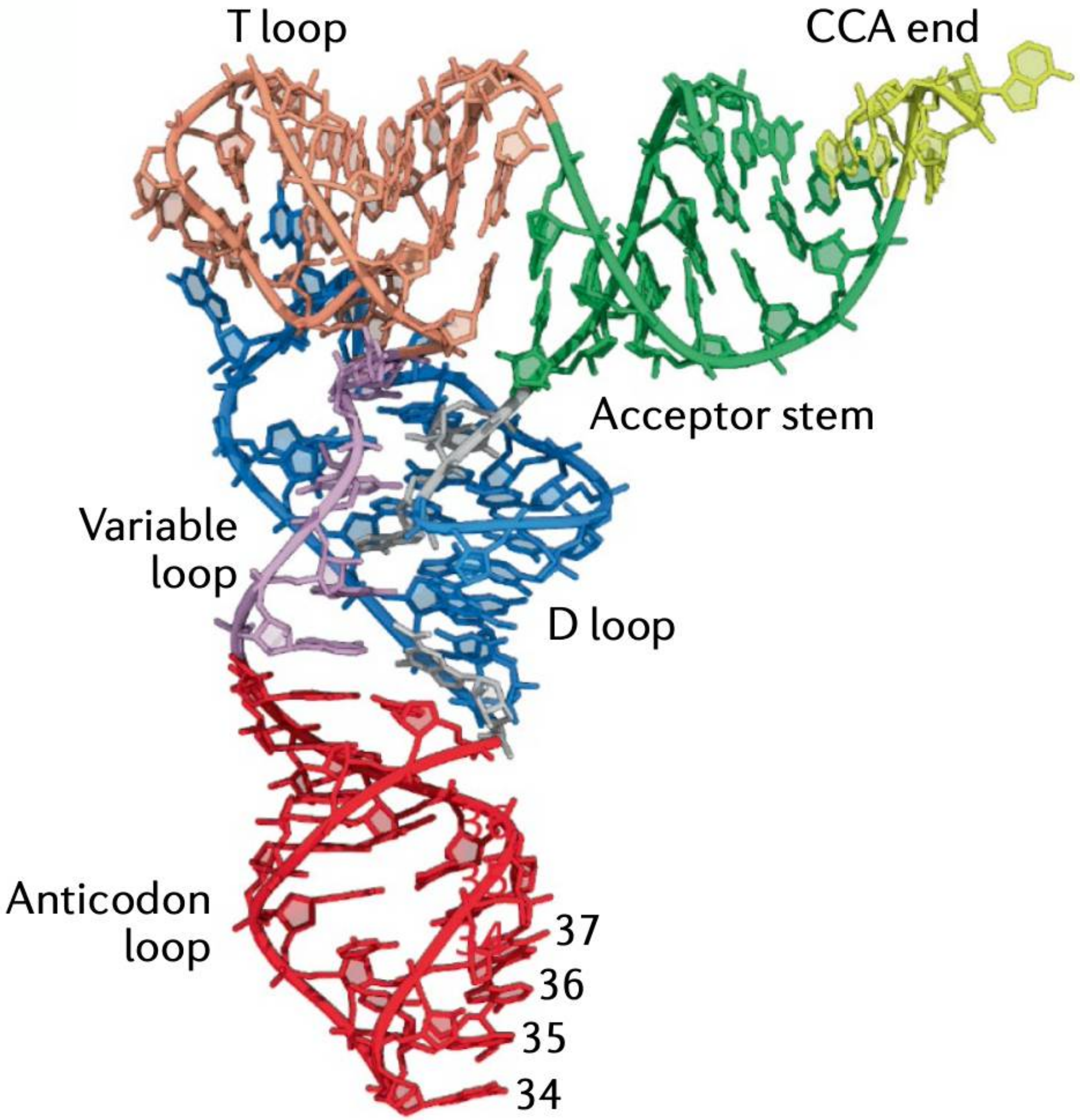


- Need be able to encode a set of 20 amino acids (AA)
- Set of four different building blocks
- 2-letter code does not cover the 20 AA ($4^2 = \mathit{16}$)
- 3-letter code comprehensively covers the 20 AA complexity ($4^3 = \mathbf{64}$)

- Redundant (AA covered by more than one codon)
- Non-ambiguous (all 64 codons encode something)
- Conservative (mutations can have relatively minor consequences)

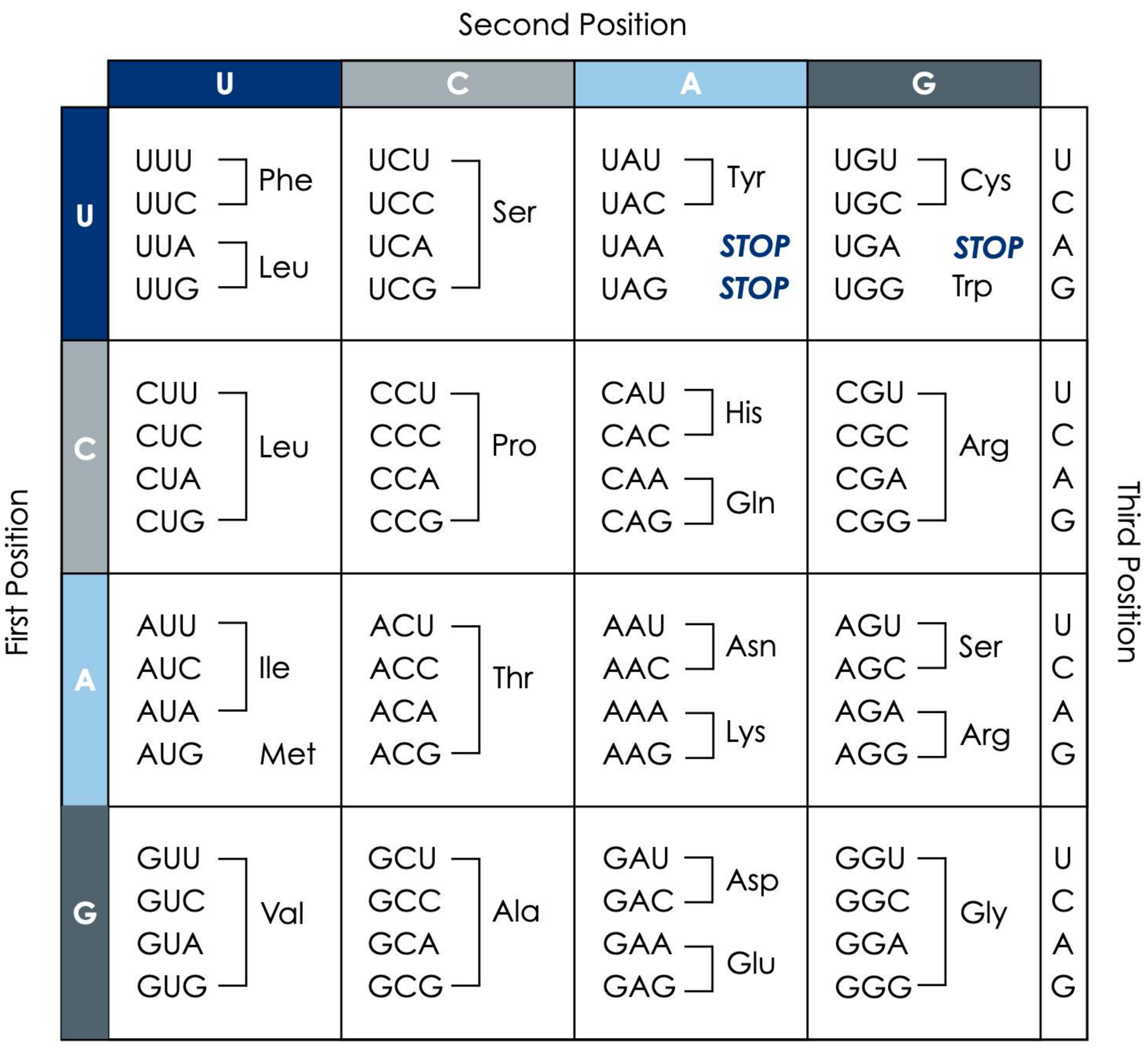
tRNA: the adaptor





https://doi.org/10.1038/s41580-021-00342-0

The Genetic Code



Second Position

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

61 codons needing tRNAs but ~30-40 tRNAs. So certain tRNA must recognize more than one codon.

↓

Wobble pairing

The Genetic Code

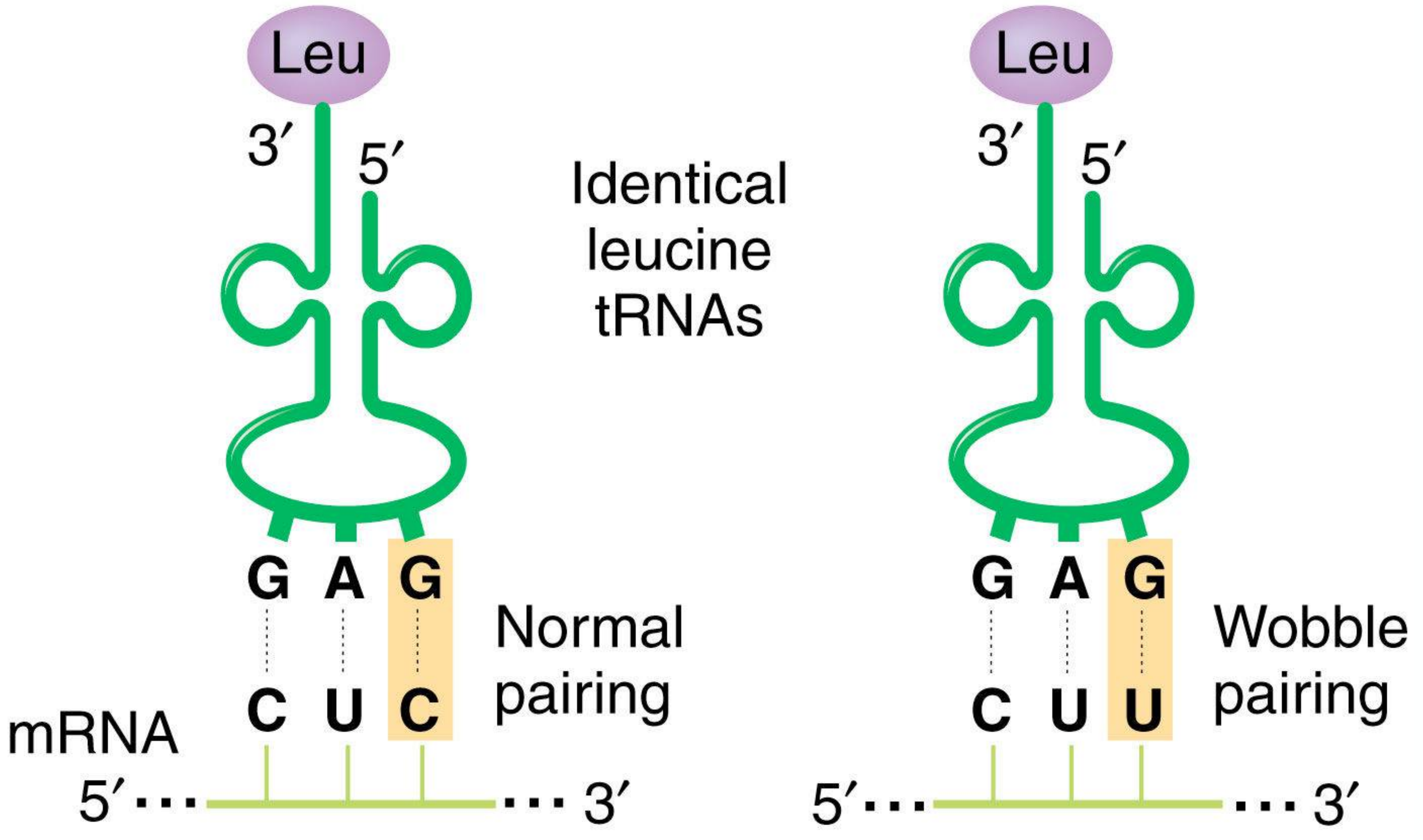
Second Position

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

61 codons needing tRNAs but ~30-40 tRNAs. So certain tRNA must recognize more than one codon.

↓

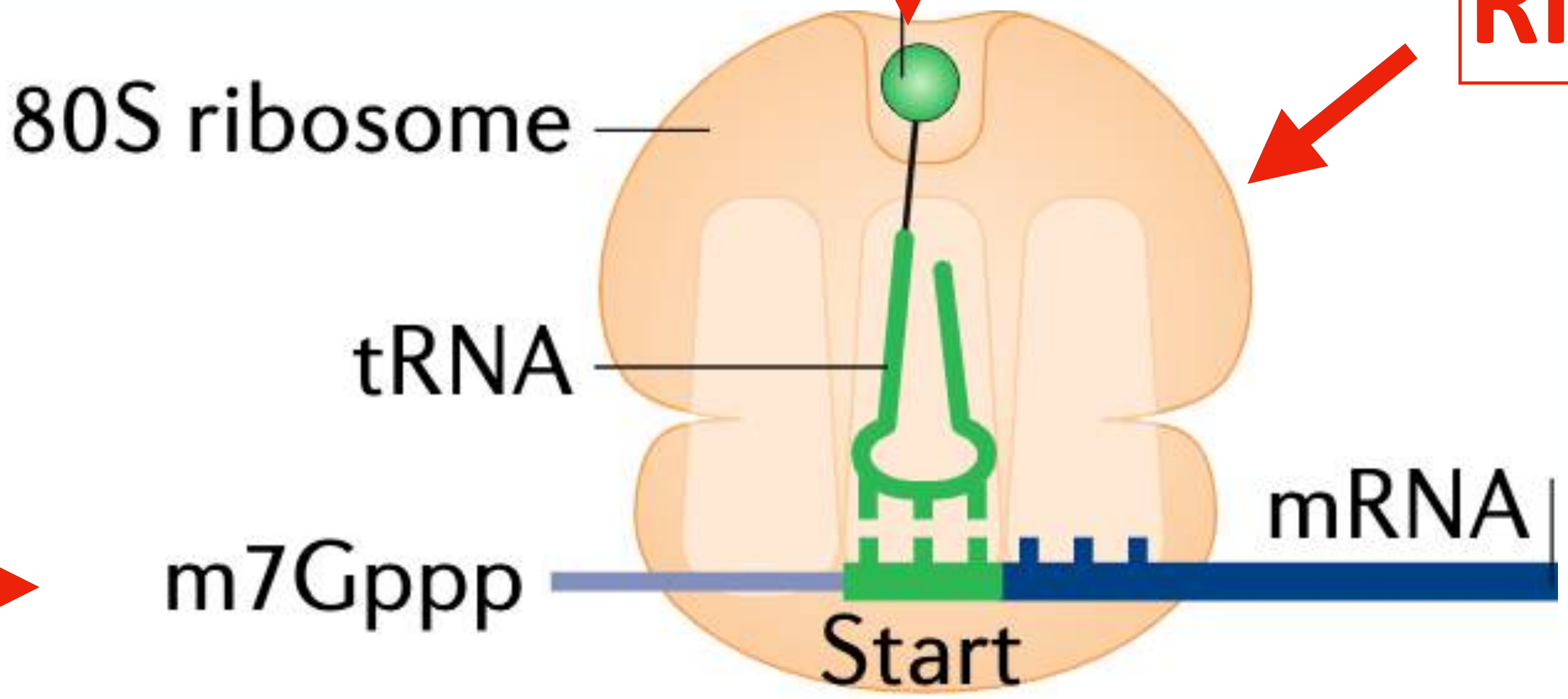
Wobble pairing



Components required for translation

tRNA

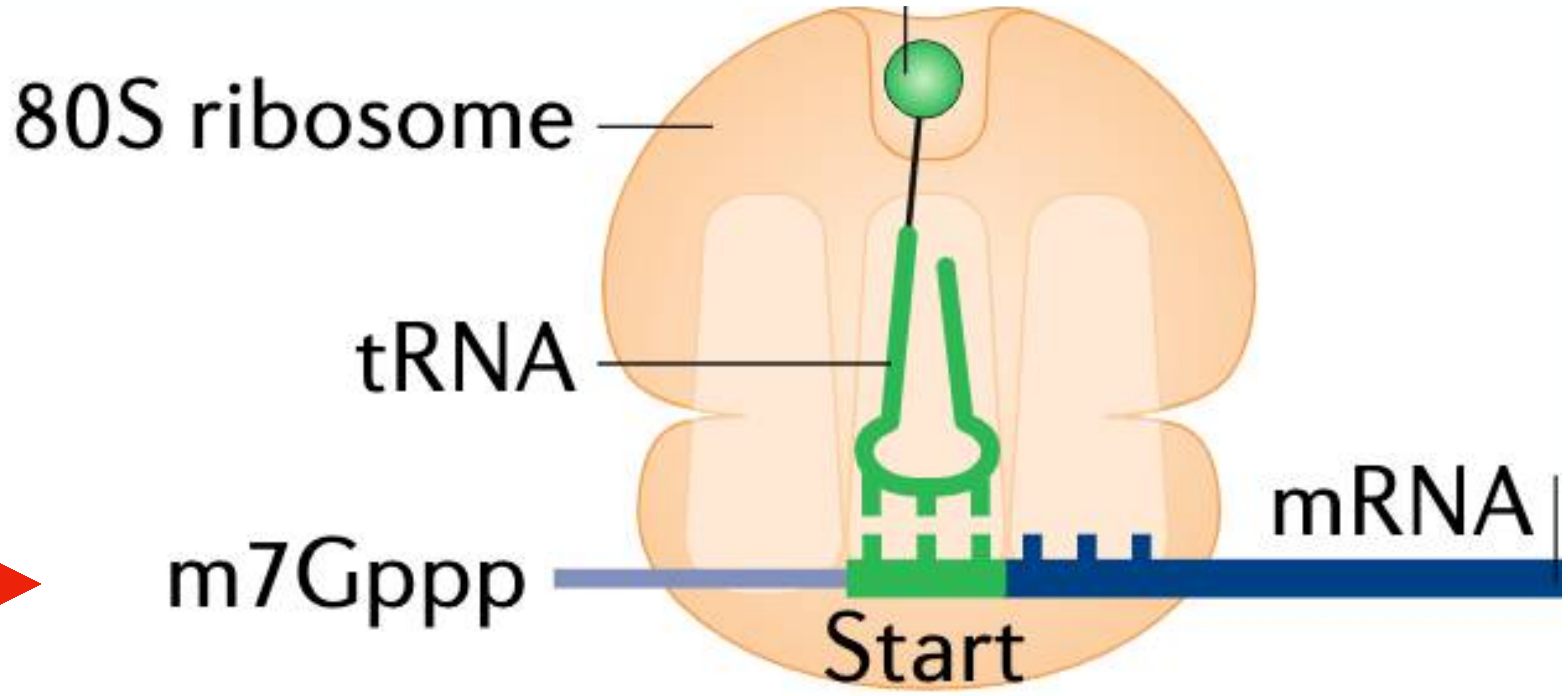
Ribosome



mRNA

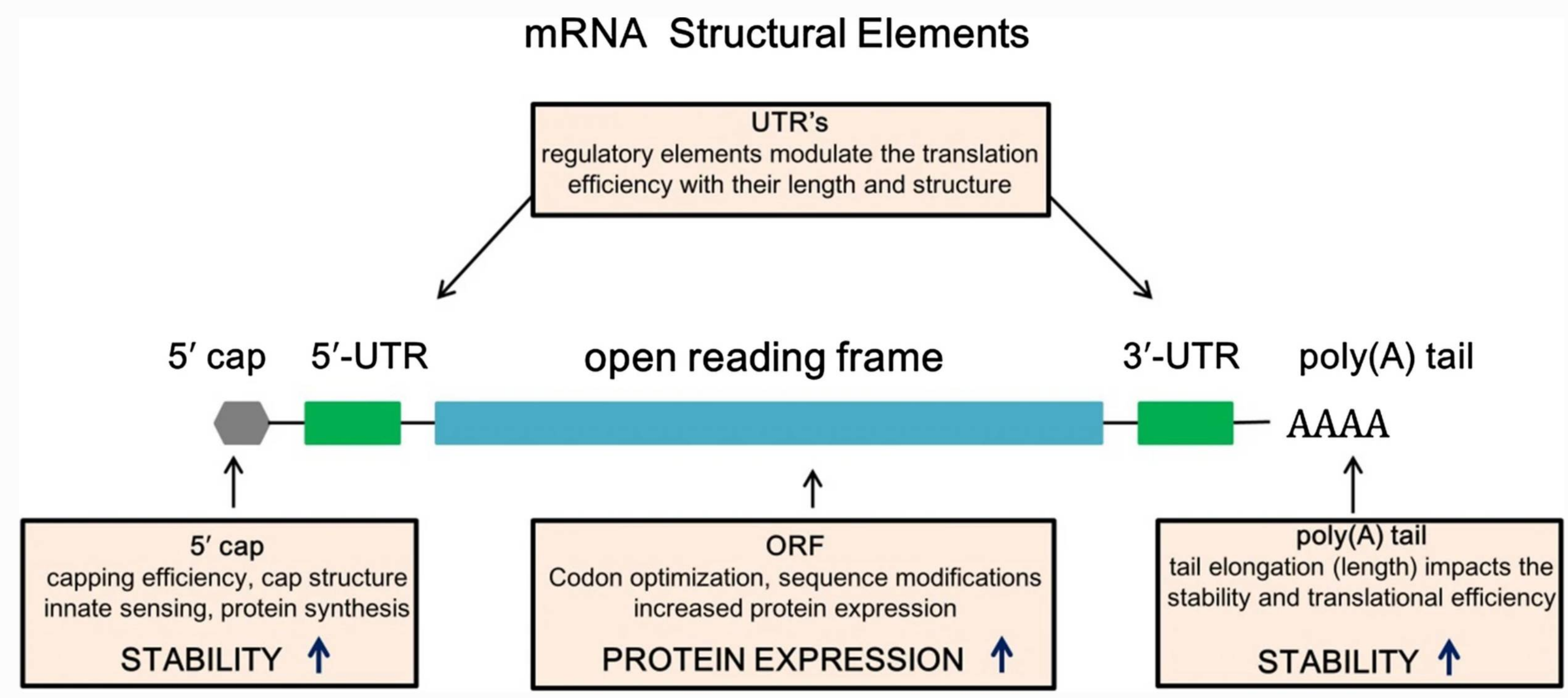
Components required for translation

mRNA →



Components required for translation

mRNA (1~5% of total RNA abundance)



mRNA Structural Elements

UTR's
regulatory elements modulate the translation efficiency with their length and structure

5′ cap
capping efficiency, cap structure
innate sensing, protein synthesis

STABILITY ↑

ORF
Codon optimization, sequence modifications
increased protein expression

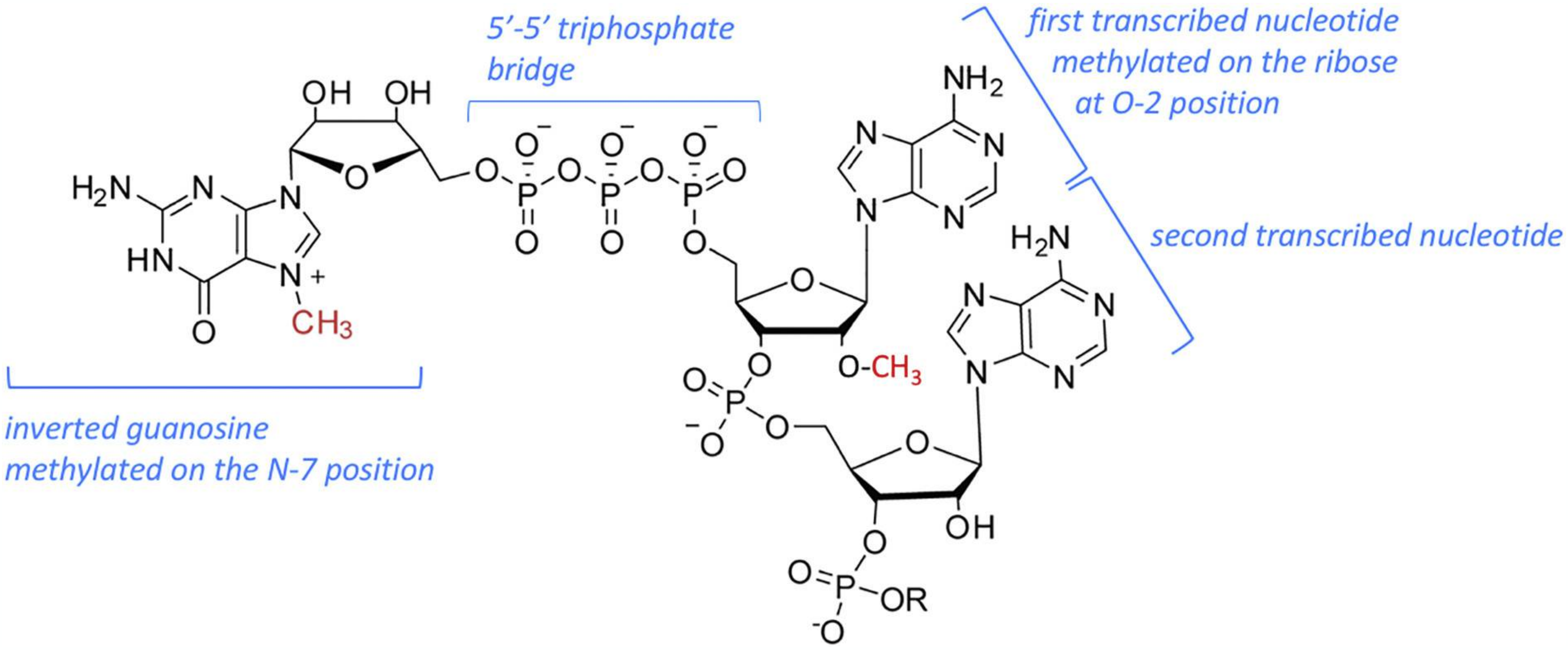
PROTEIN EXPRESSION ↑

poly(A) tail
tail elongation (length) impacts the stability and translational efficiency

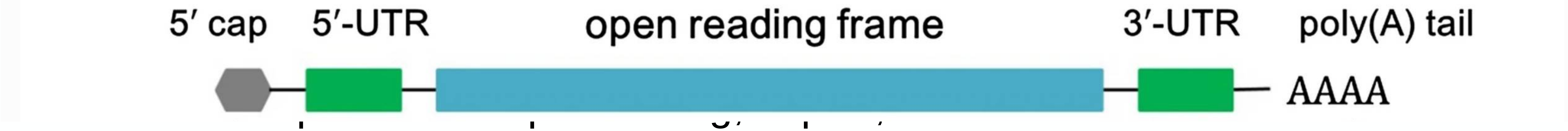
STABILITY ↑

mRNA 5’-Cap

7-Methylguanosine is linked to the first transcribed nucleotide *via* a 5′ to 5′ triphosphate bridge.



Role of the mRNA cap

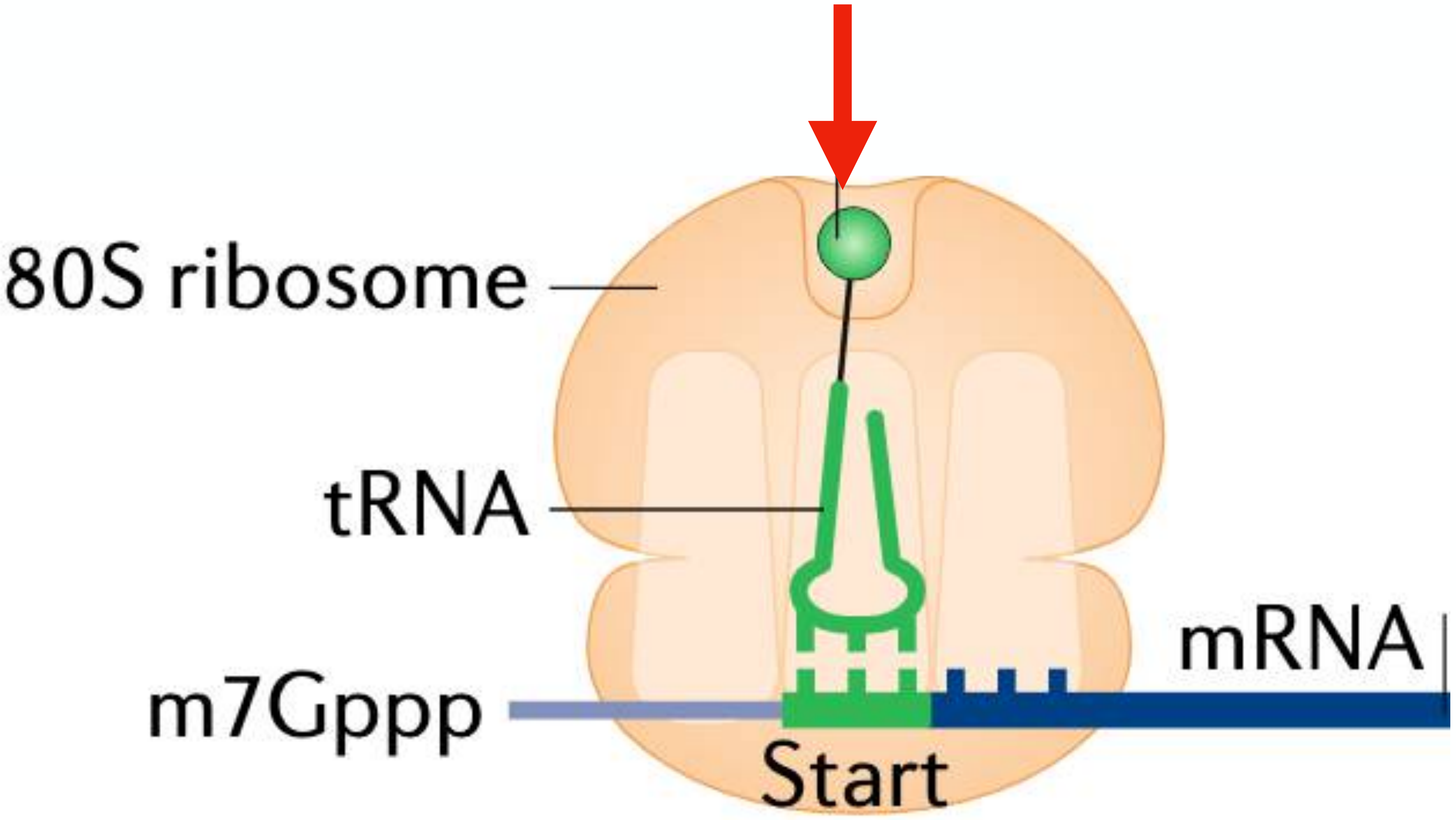


3. Removal of the cap (decapping) initiates degradation of mRNA

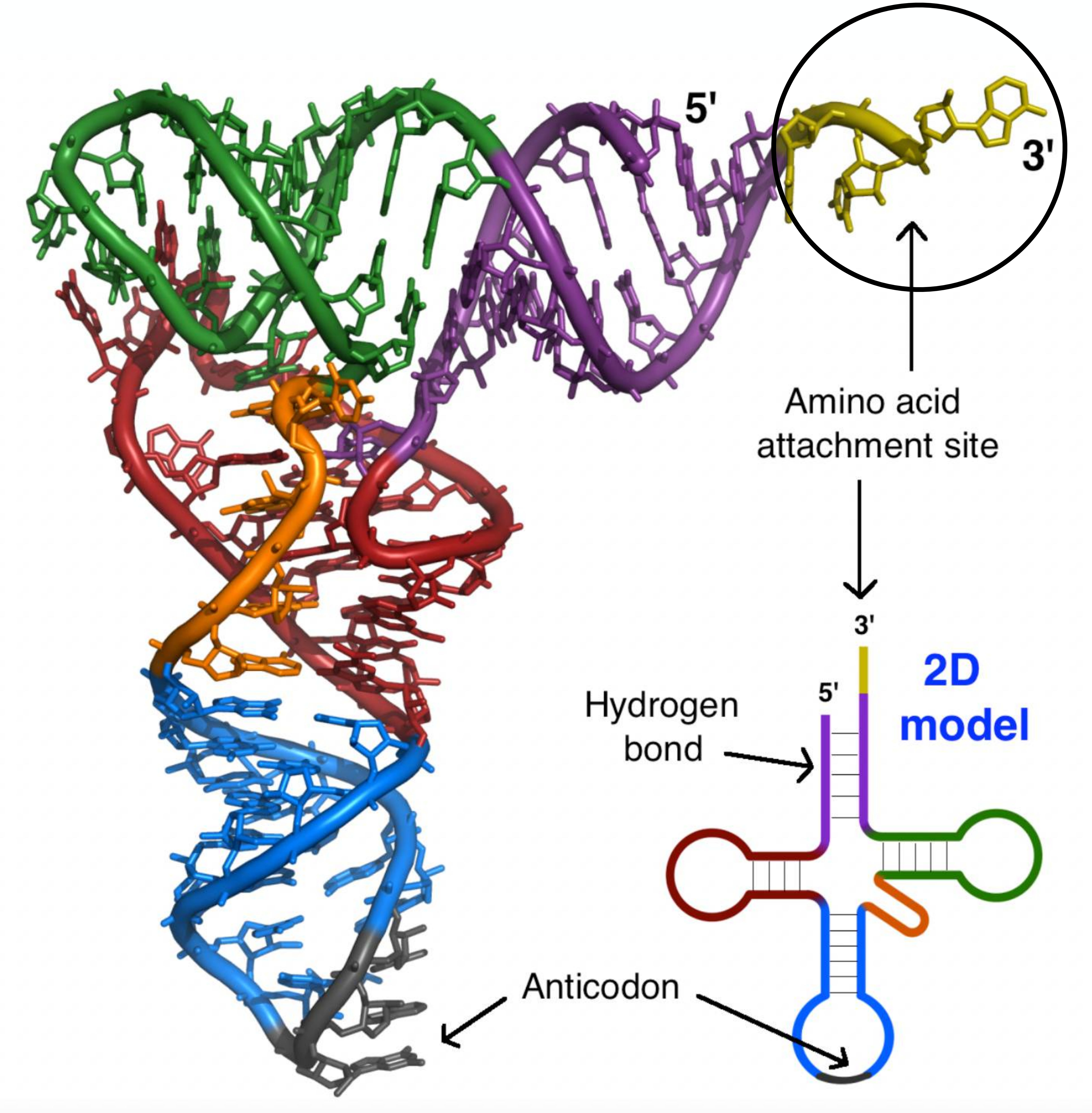
the cap is mechanistically involved in every stage of the mRNA lifecycle

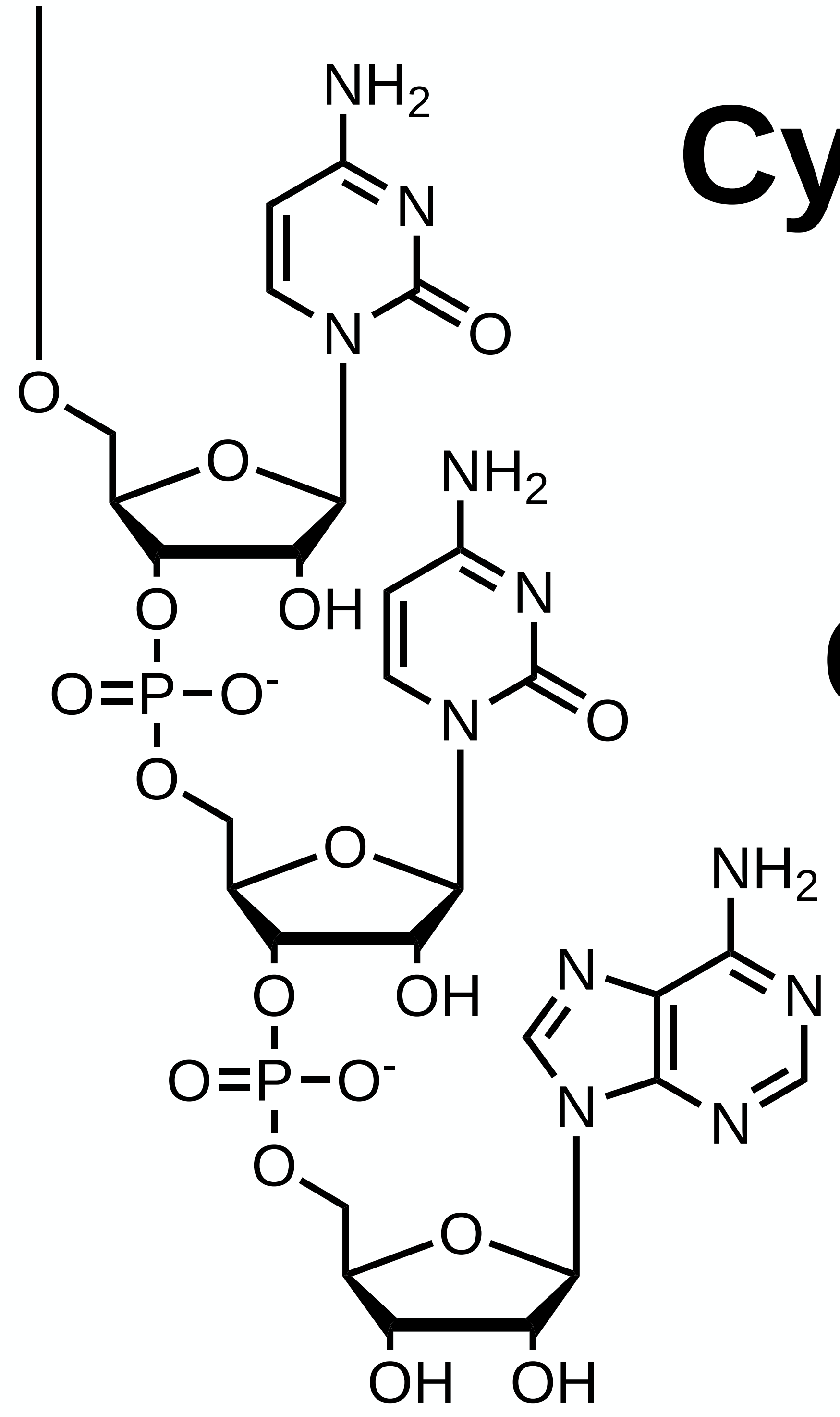
Components required for translation

tRNA



tRNA (~15% of total RNA)



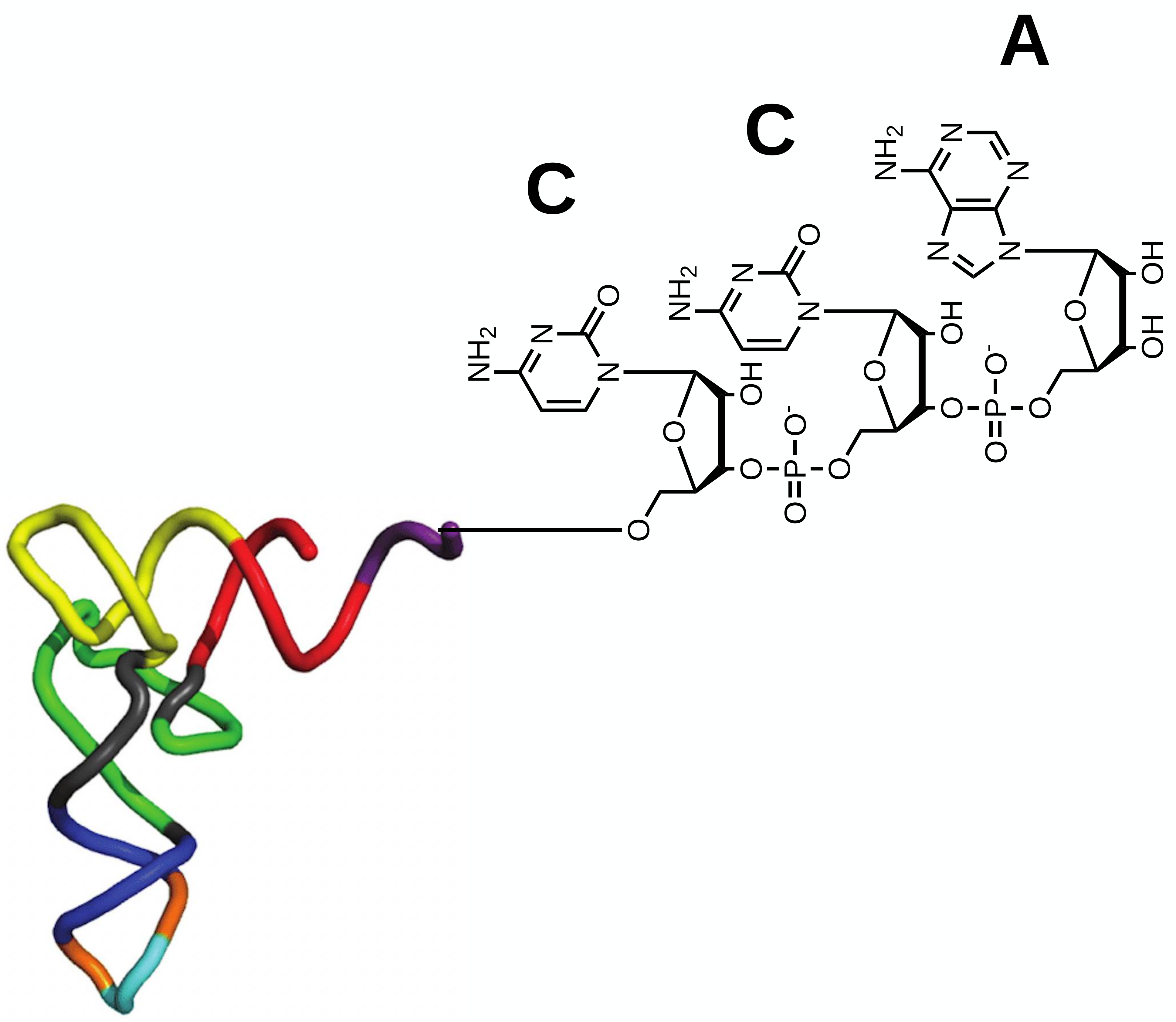


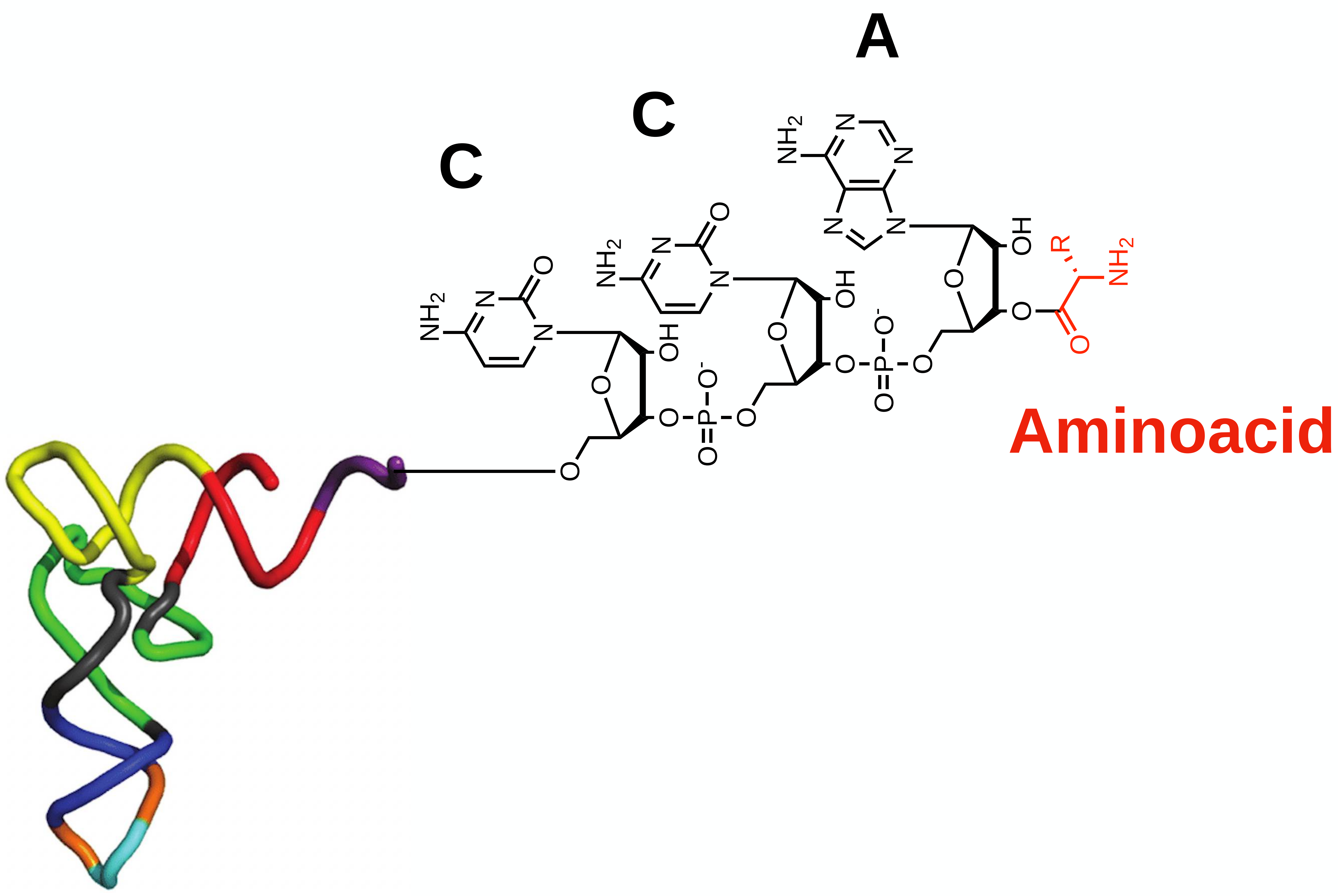
Cytidine

Cytidine

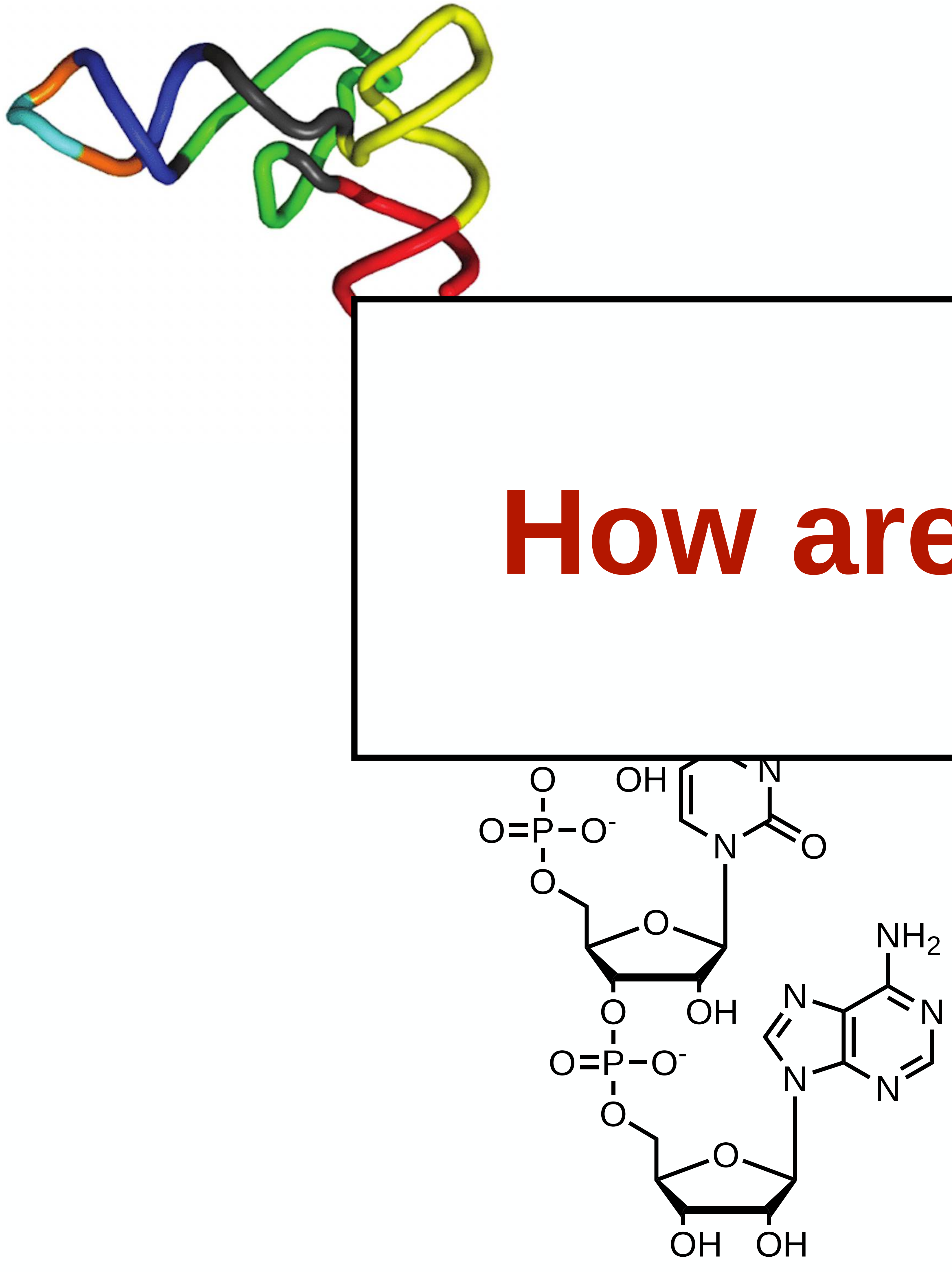
Adenine

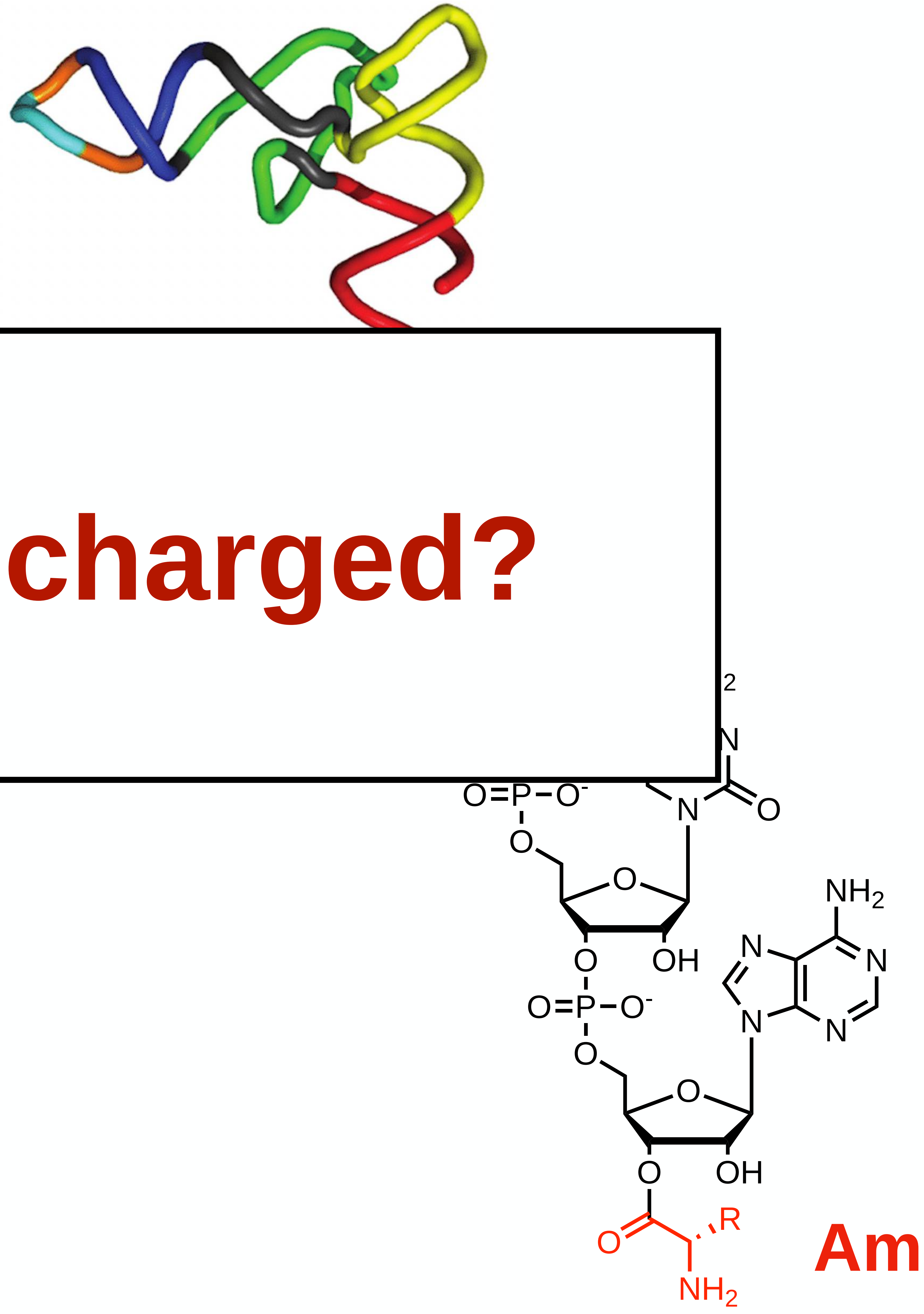
Uncharged tRNA ➔ charged tRNA





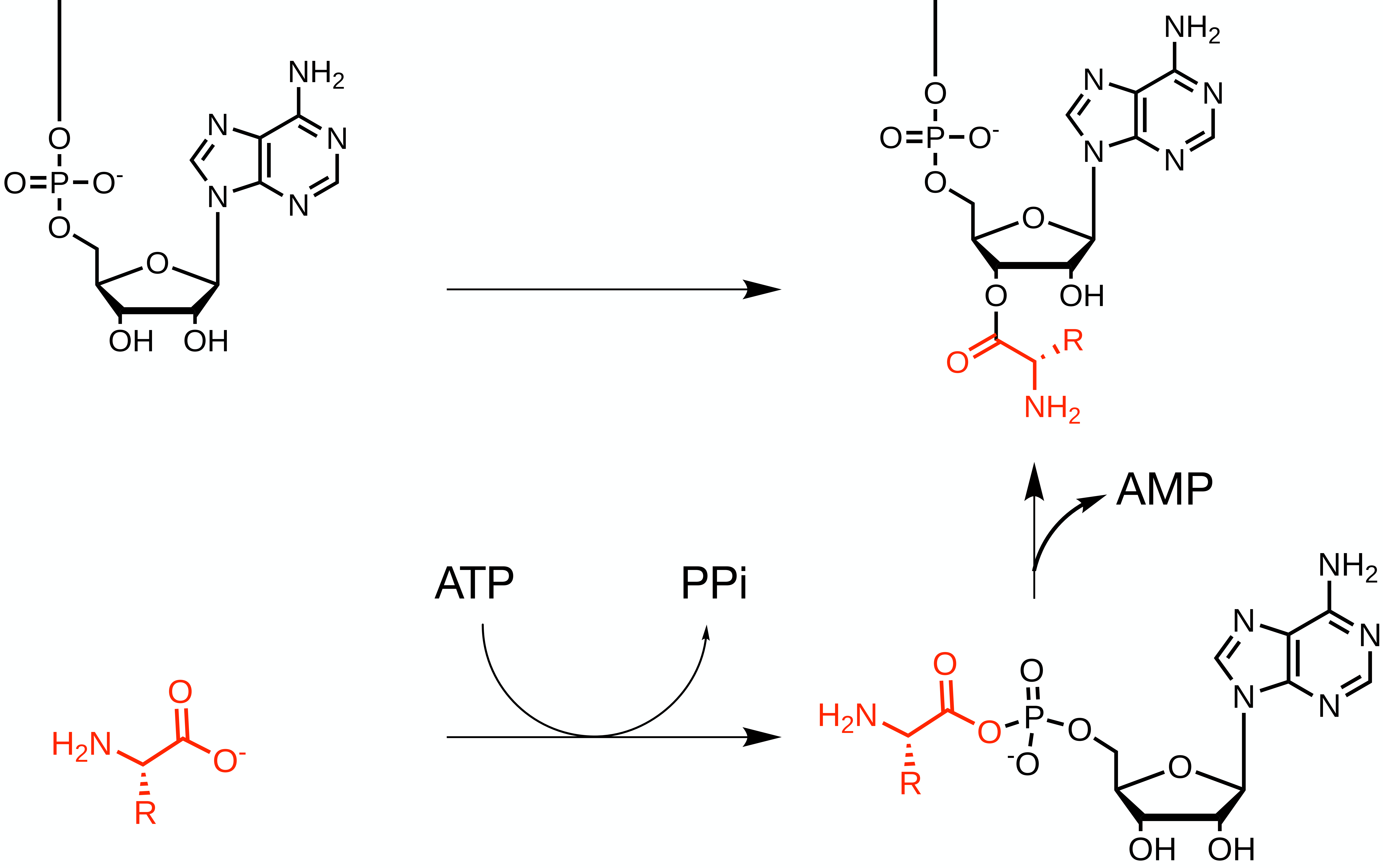
Uncharged tRNA → charged tRNA





Aminoacid

How are tRNAs charged?



Activation of Carboxylic acid using ATP

~20 tRNA Synthetases:

Extremely high fidelity process 10^{-5} (proof-reading)

CCA hydroxyl group

ATP binding pocket

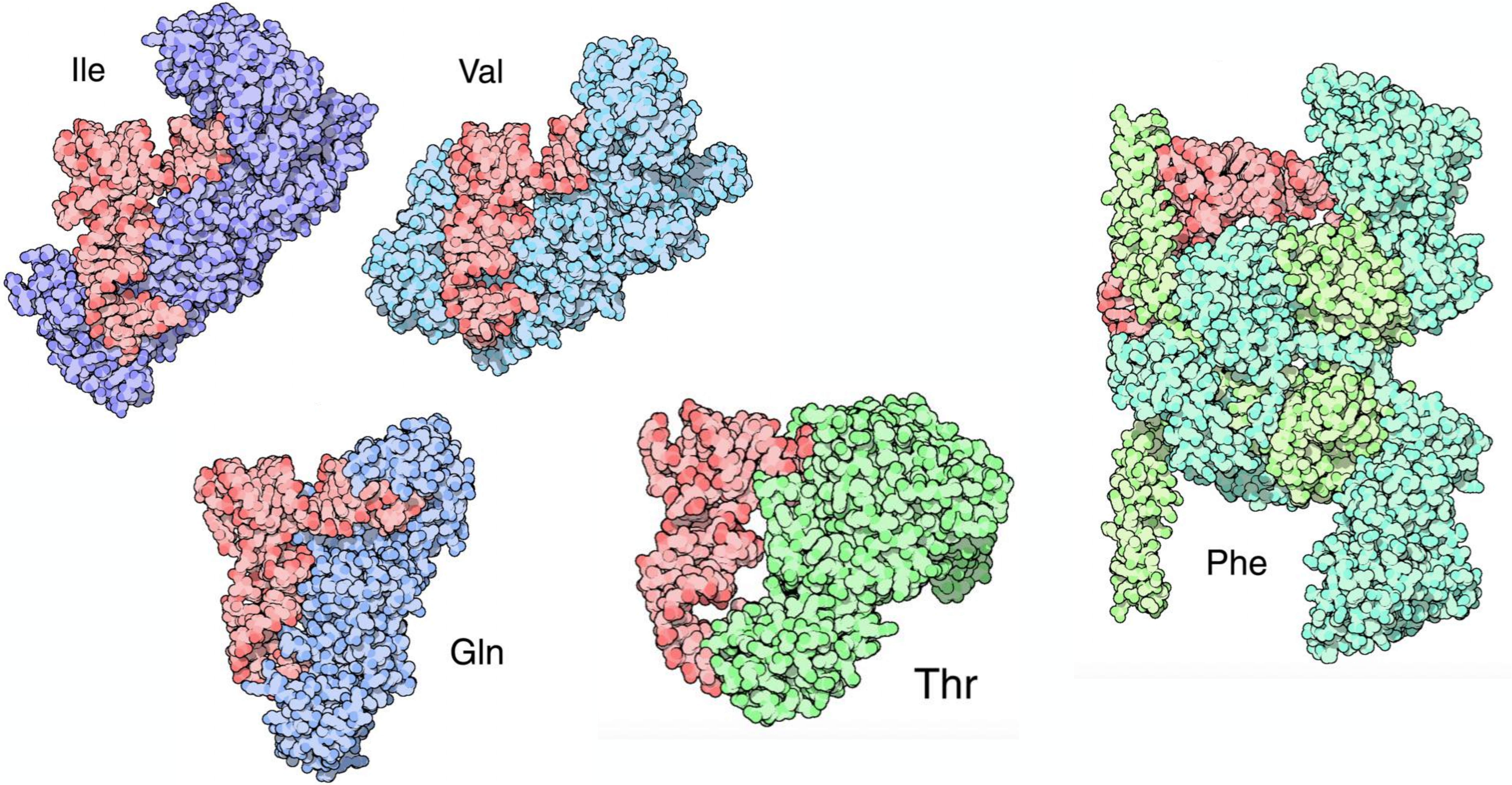
tRNA

Glutamyl-tRNA Synthetase

Science **285**: 1074-1077

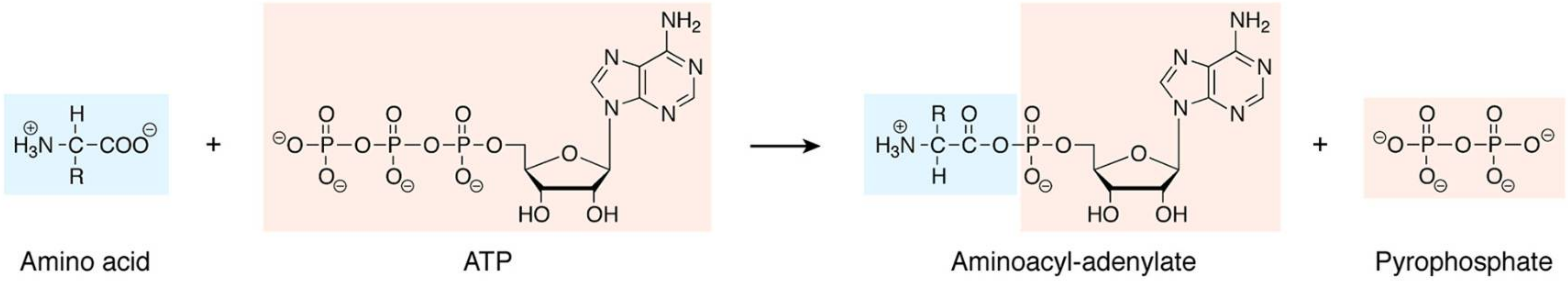
~20 tRNA Synthetases:

Different Approaches to the Same Problem

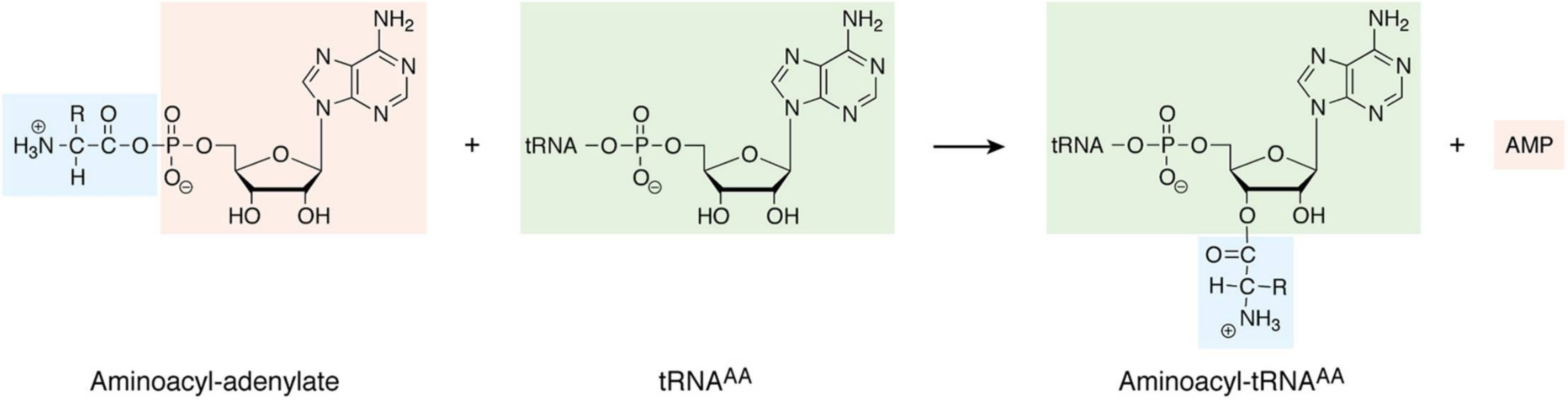


tRNA Synthetases:

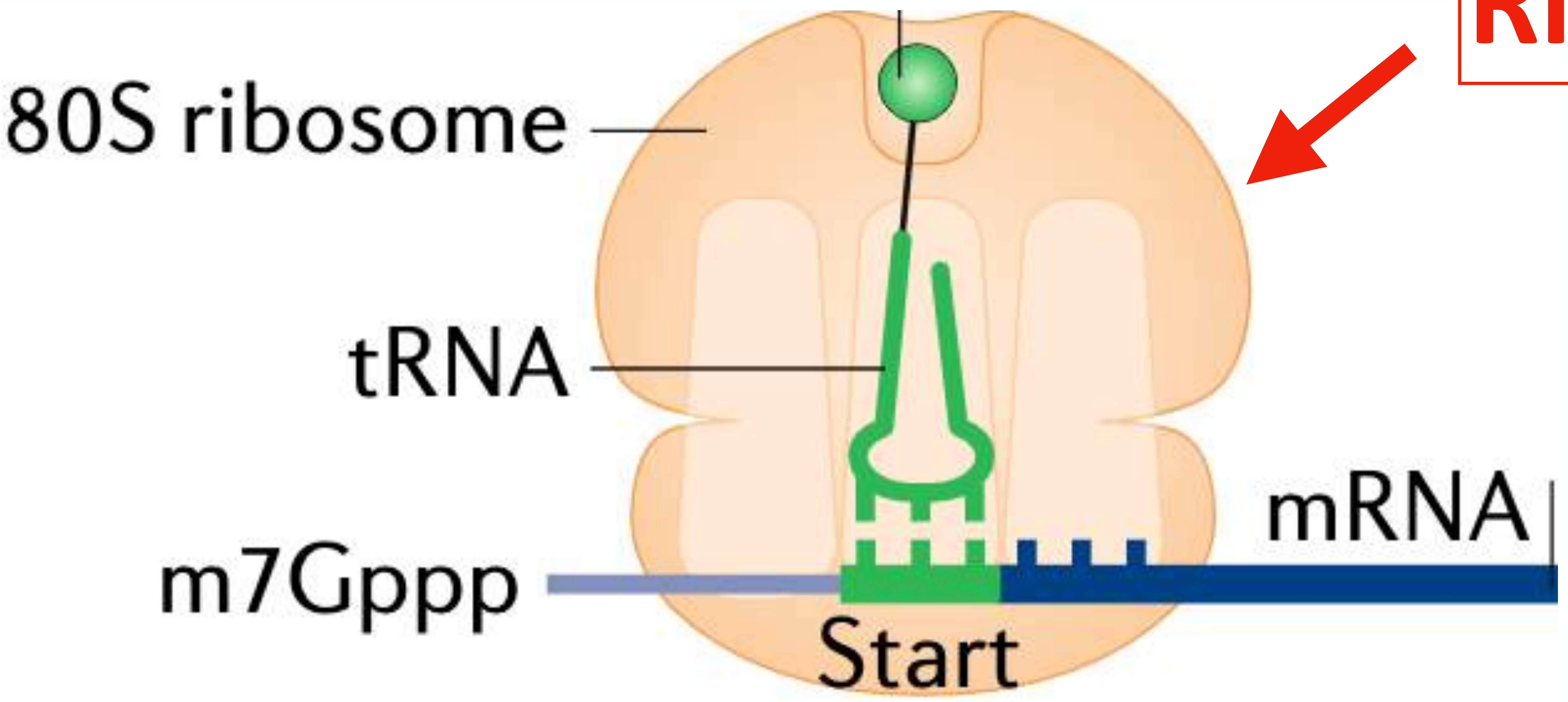
Amino acid activation



Aminoacyl transfer

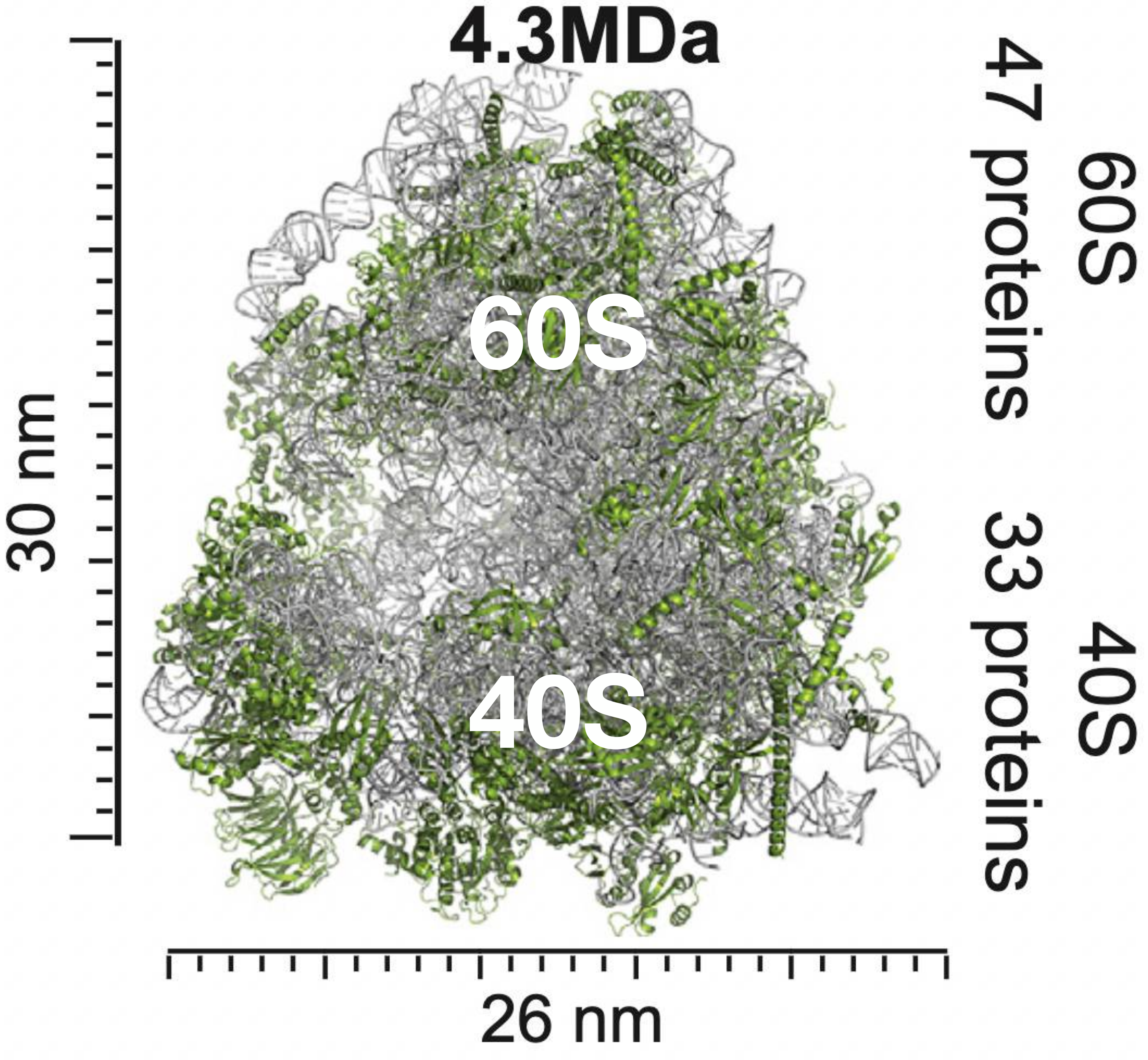


Components required for translation



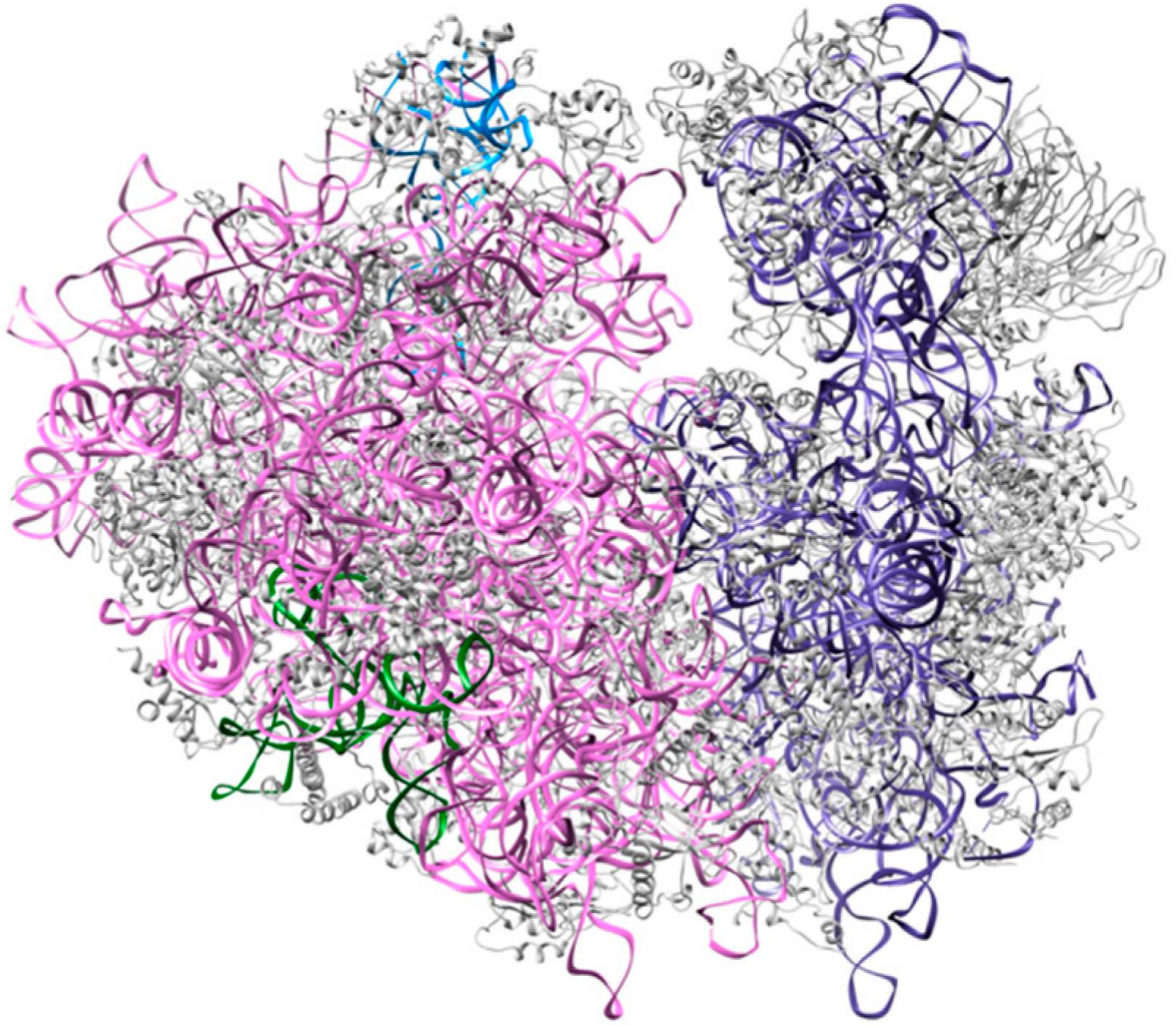
Ribosome

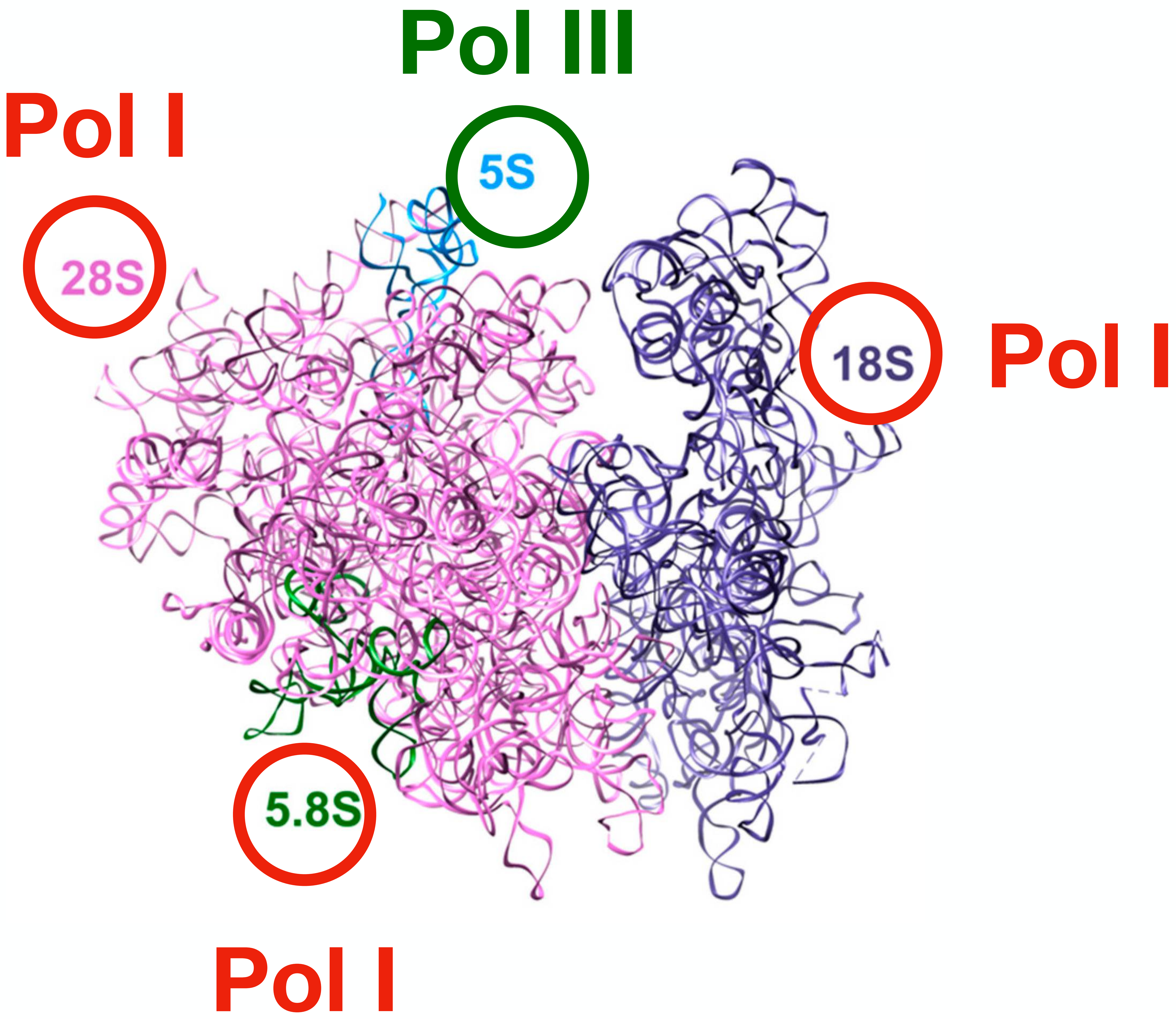
Human ribosomes by the number



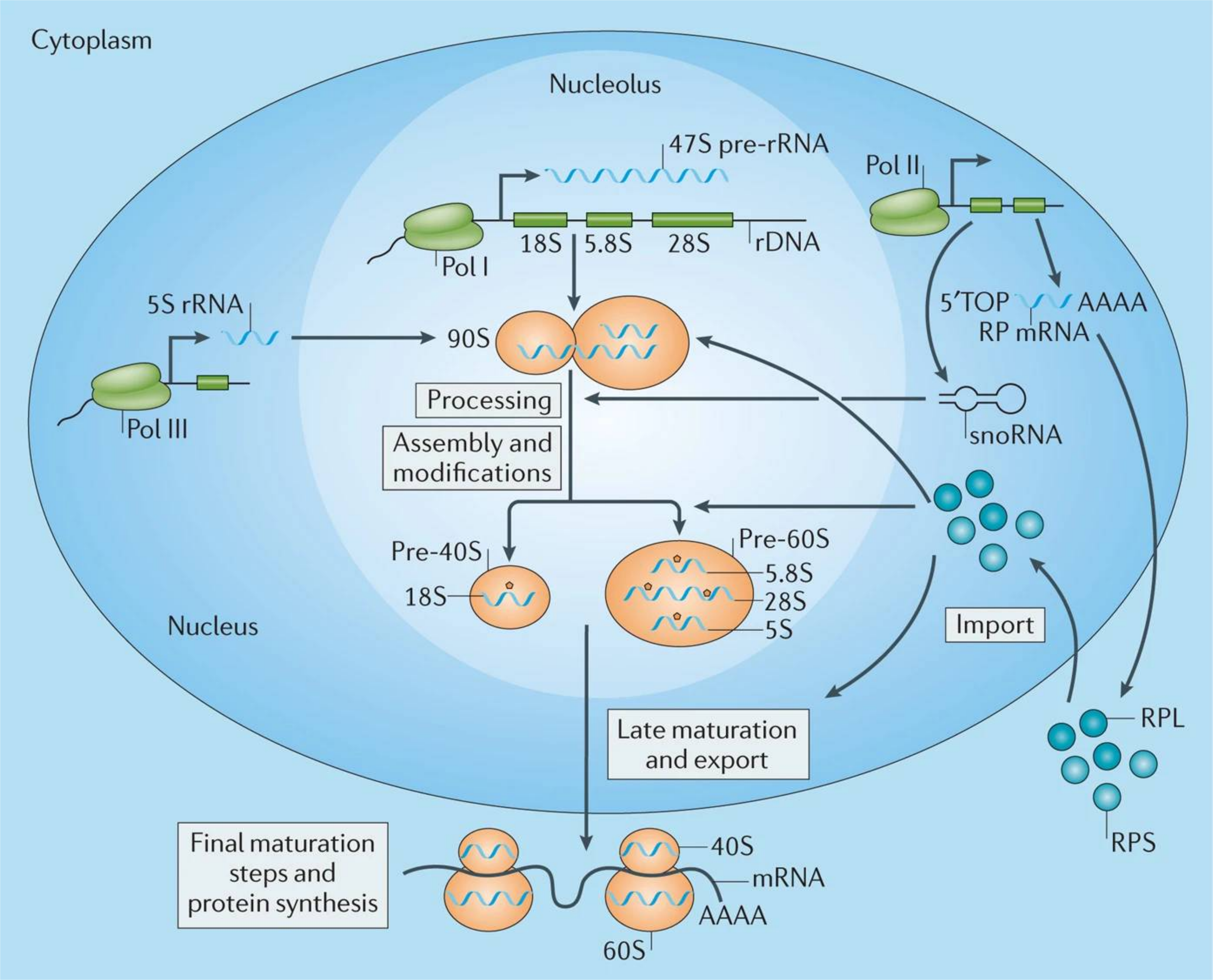
	H. Sapiens
RNA : Protein mass ratio	1:1
r-protein numbers	**80**
rRNA/total RNA	~ 80%
r-protein mass/proteome	4~6 %
r-protein copy number / proteome	**~10 %**
Ribosome number / cell	1-10 X 10^6

Ribosome biogenesis





Ribosome biogenesis



Processing

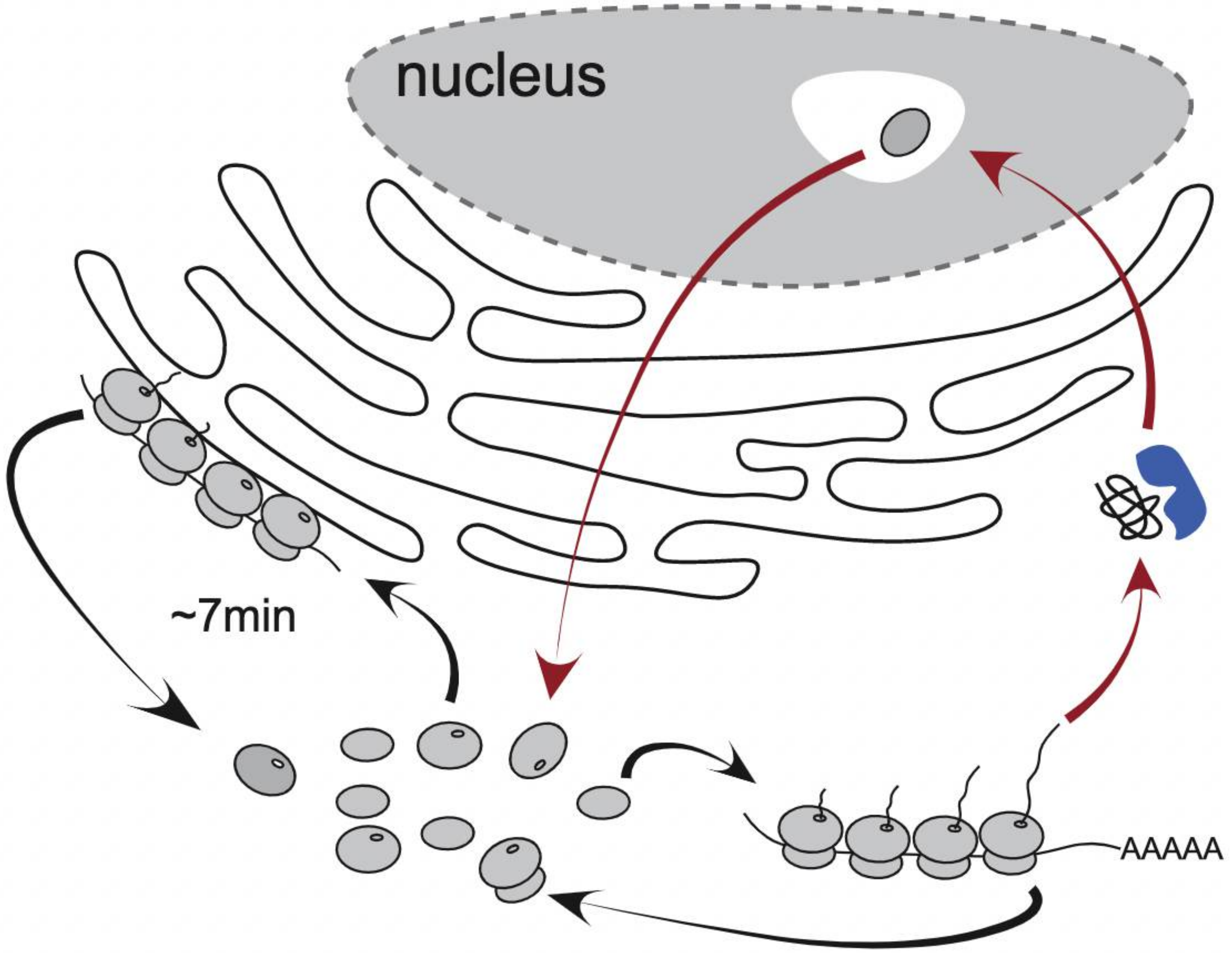
Assembly and modifications

Import

Late maturation and export

Final maturation steps and protein synthesis

Ribosome biogenesis

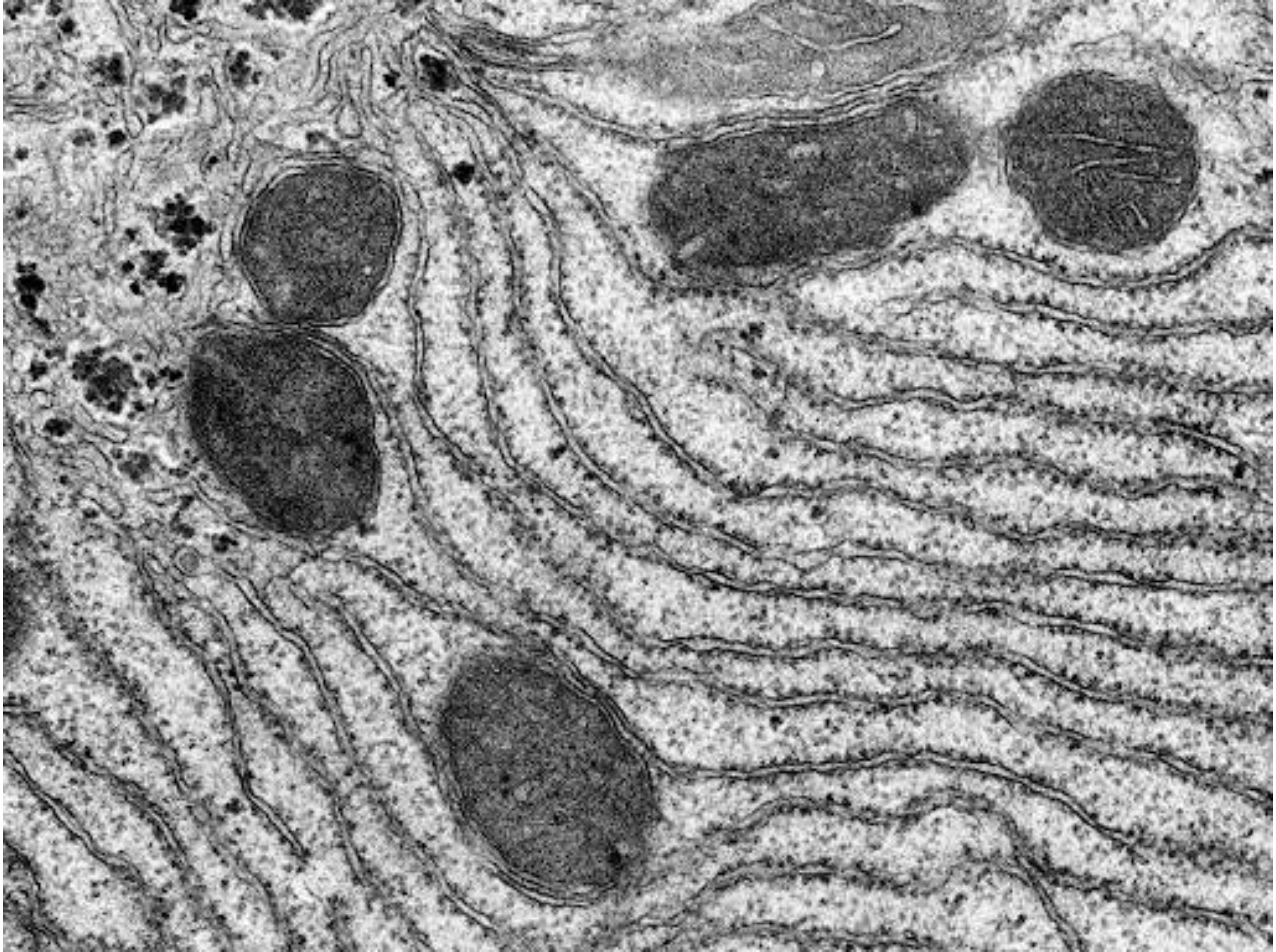


Cytosol-ER exchange

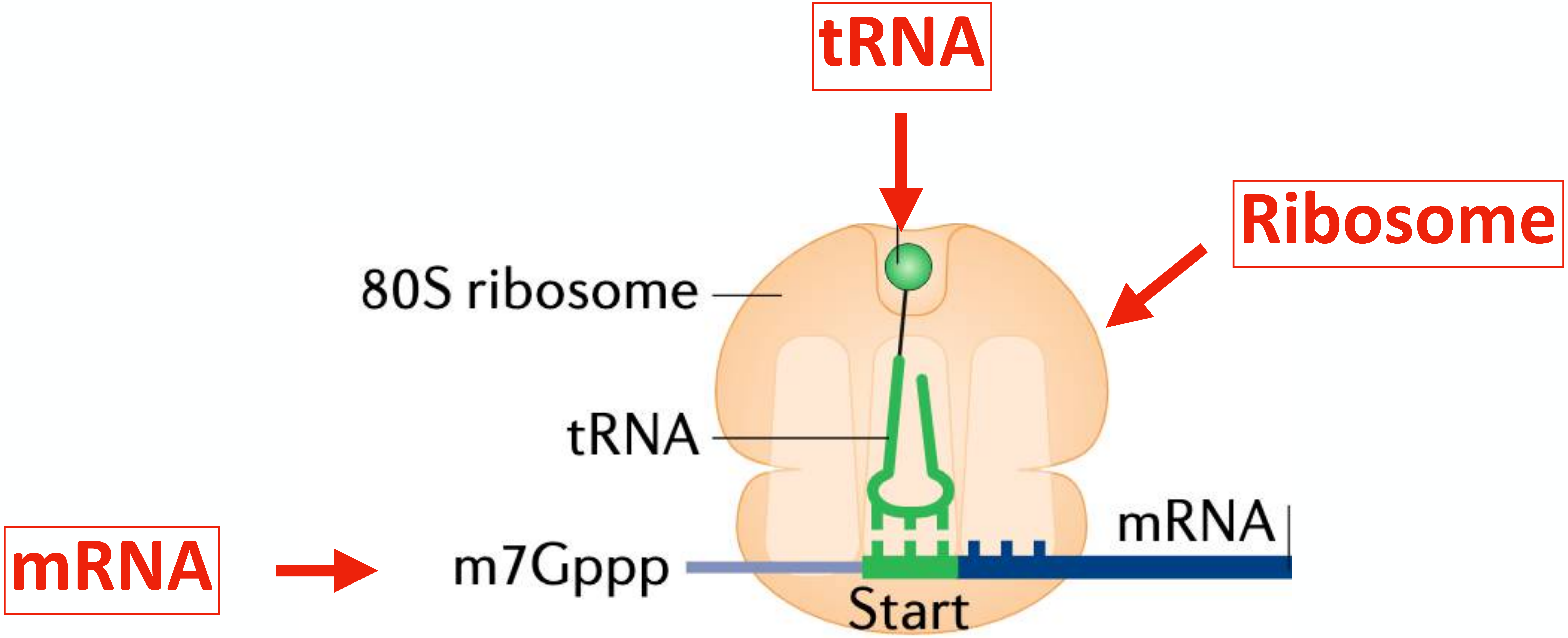
biogenesis

Rate = ~ 4000 ribosomes/min

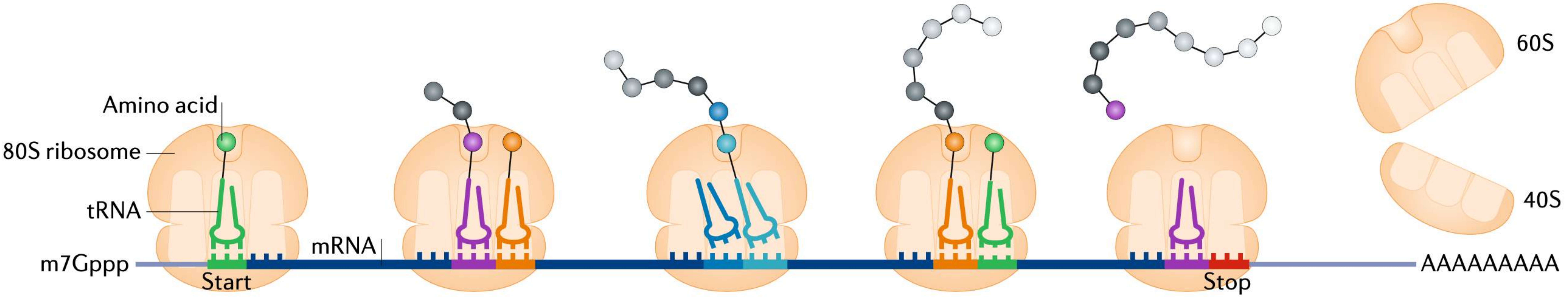
Average r-protein size = 21.5 KDa



Components required for translation



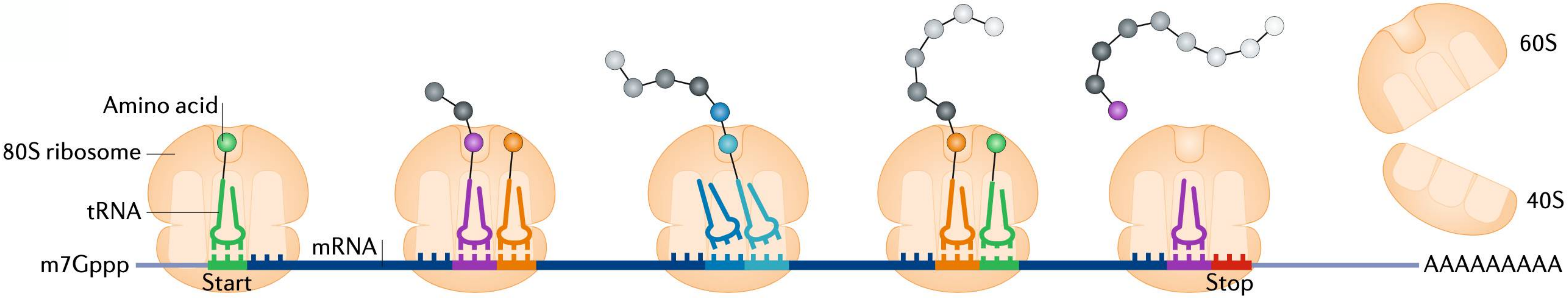
Overview of eukaryotic translation



Phase:	Initiation	Elongation	Termination	Recycling
Factors:	eIFs $\text{Met-tRNA}^{\text{iMet}}$	eEFs aa-tRNAs	eRFs	ABCE1

Initiation **Elongation** **Termination**

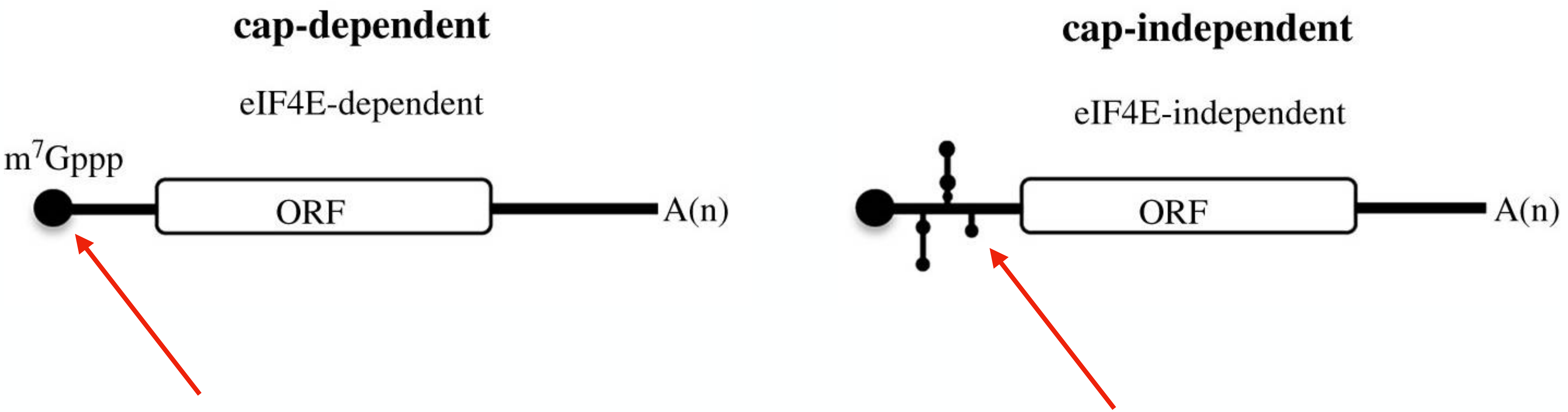
Overview of eukaryotic translation



Phase:	Initiation	Elongation	Termination	Recycling
Factors:	eIFs Met-tRNA^{iMet}	eEFs aa-tRNAs	eRFs	ABCE1

Initiation **Elongation** **Termination**

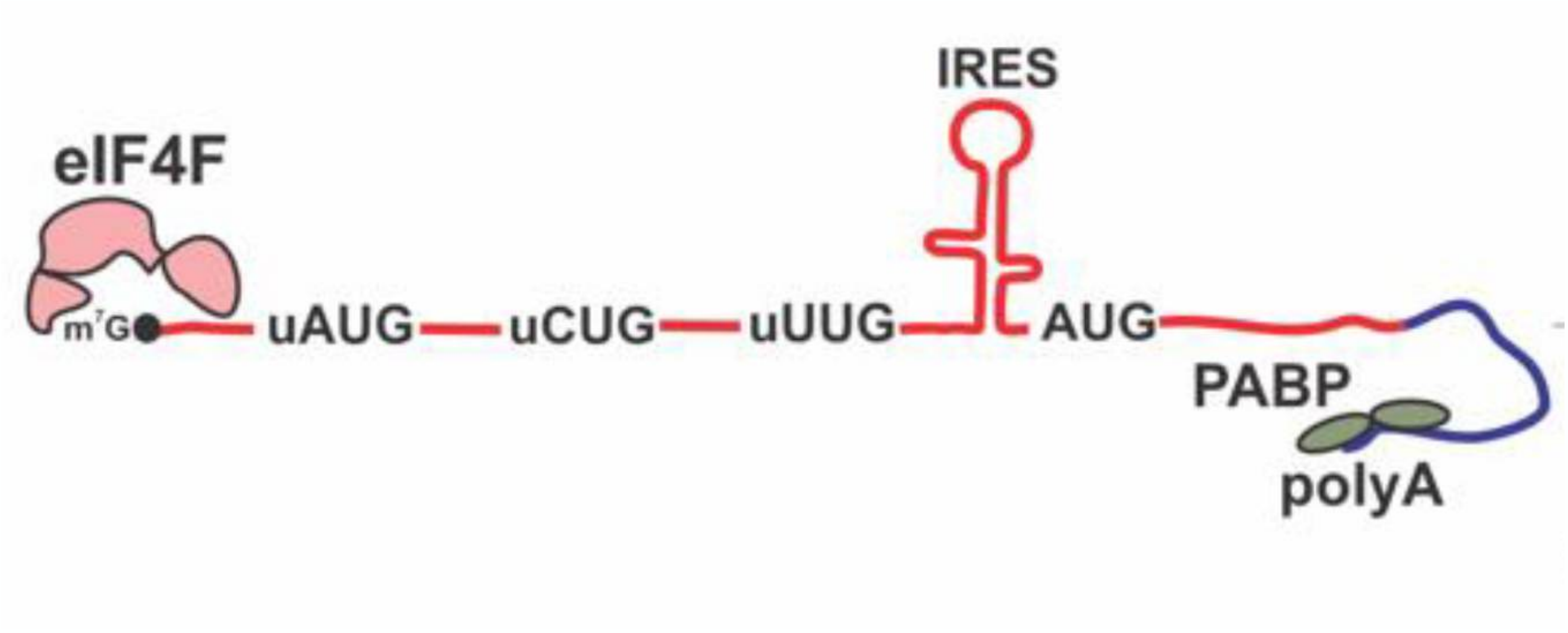
Initiation: Cap dependent vs. Cap independent



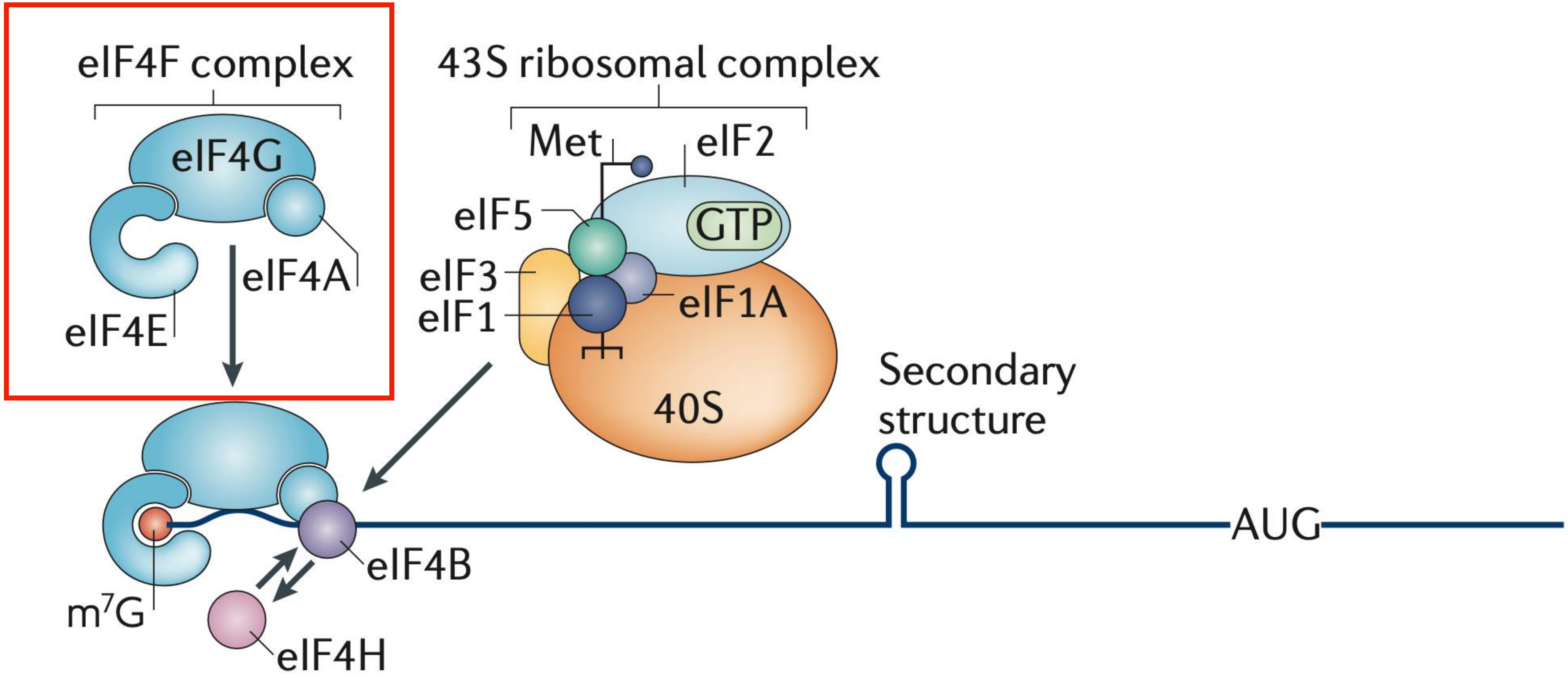
eIF4F complex
recognizes 5’ Cap, then recruits 43S complex

mRNA secondary structure
directly recruits ribosome
(ex. IRES: internal ribosome entry site)

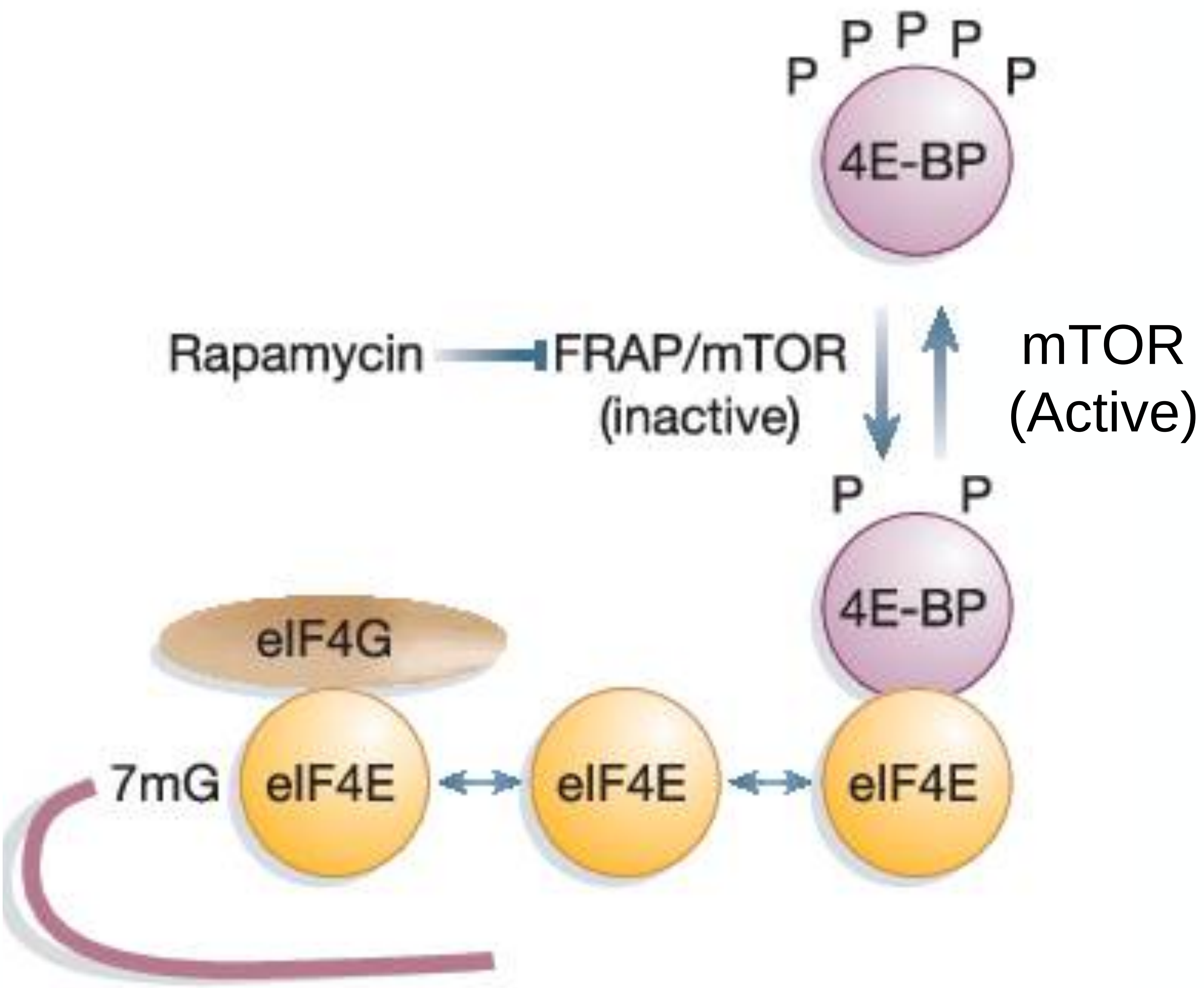
Many recombinant plasmids contain both Cap dependent initiation and IRES initiation sites



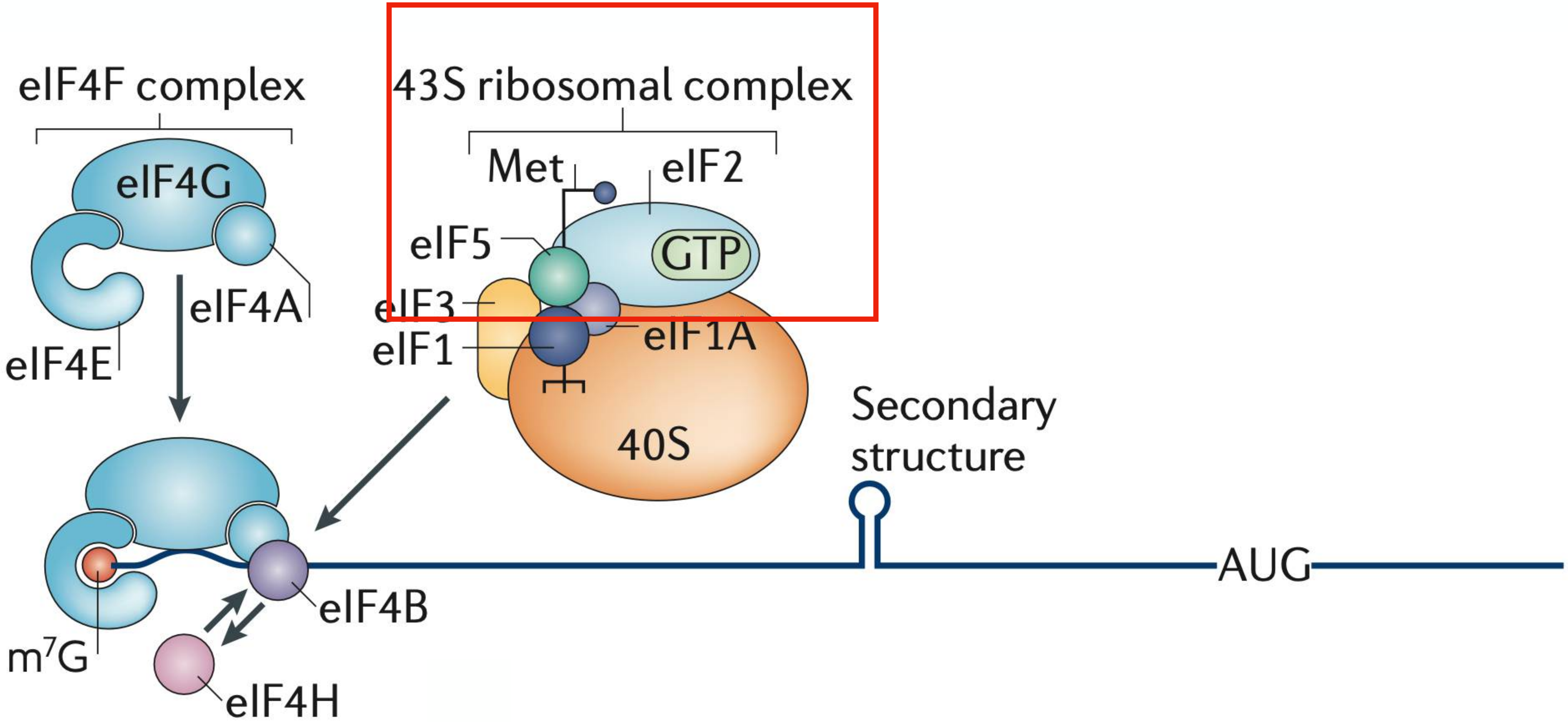
Cap dependent translational initiation



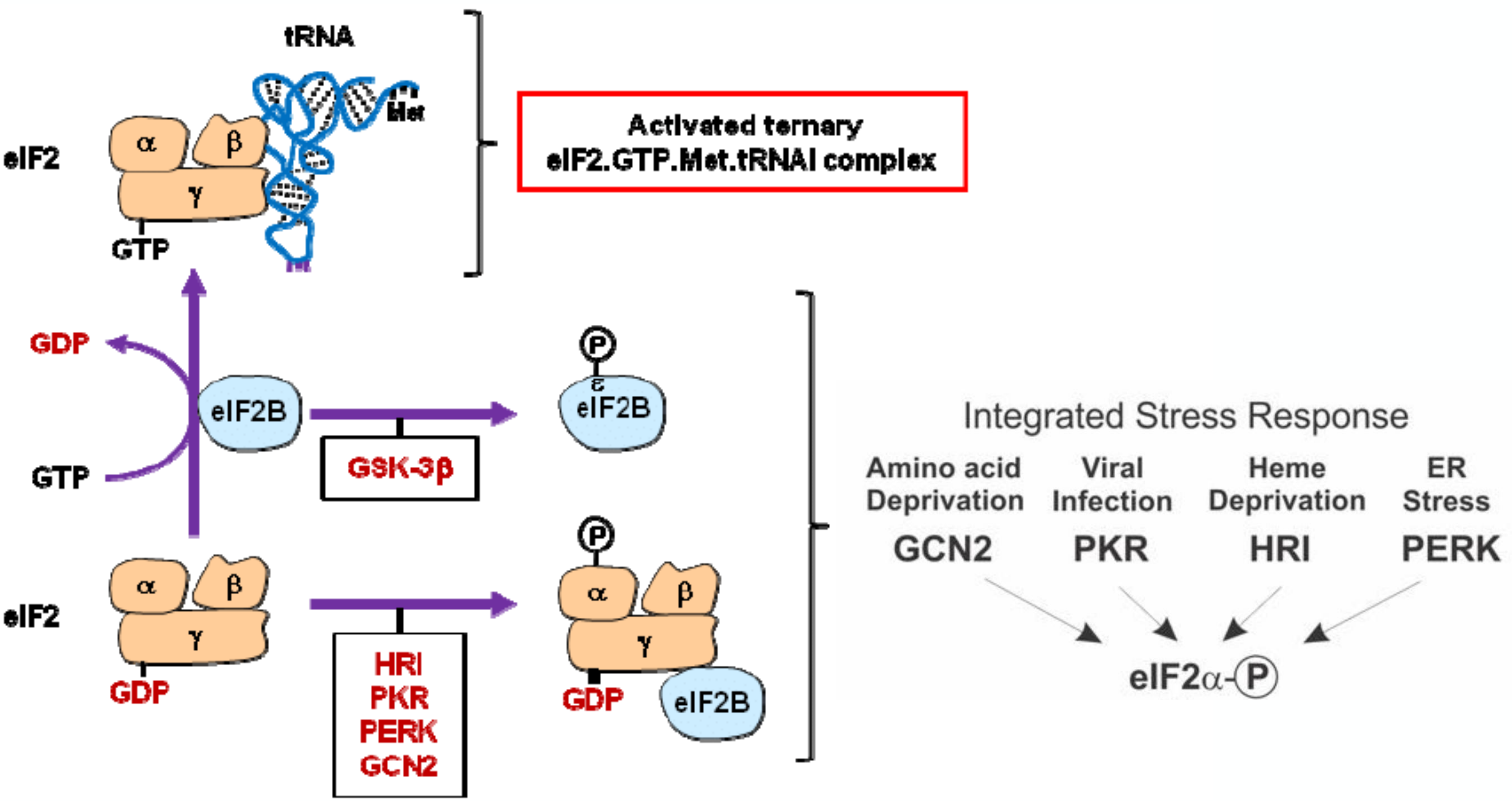
mTOR controls Cap dependent translation through eIF4F



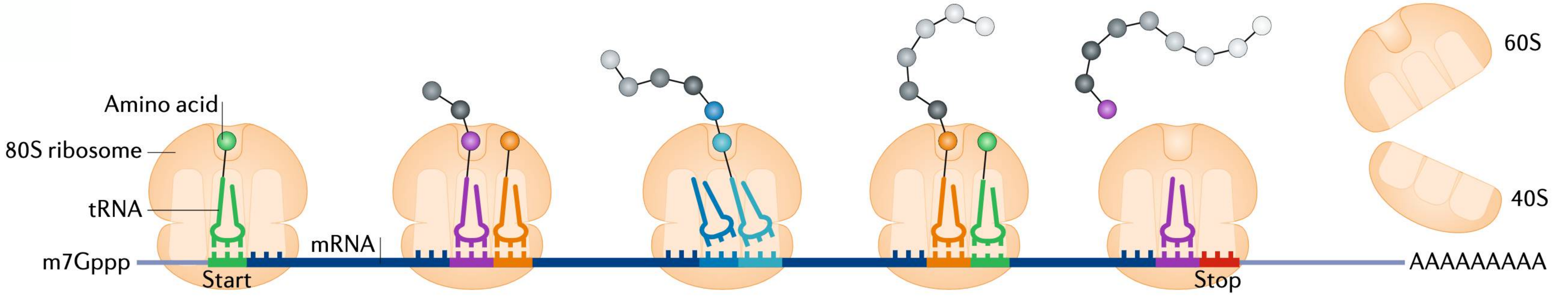
Cap dependent translational initiation



Cellular stress controls global translation through eIF2



Overview of eukaryotic translation

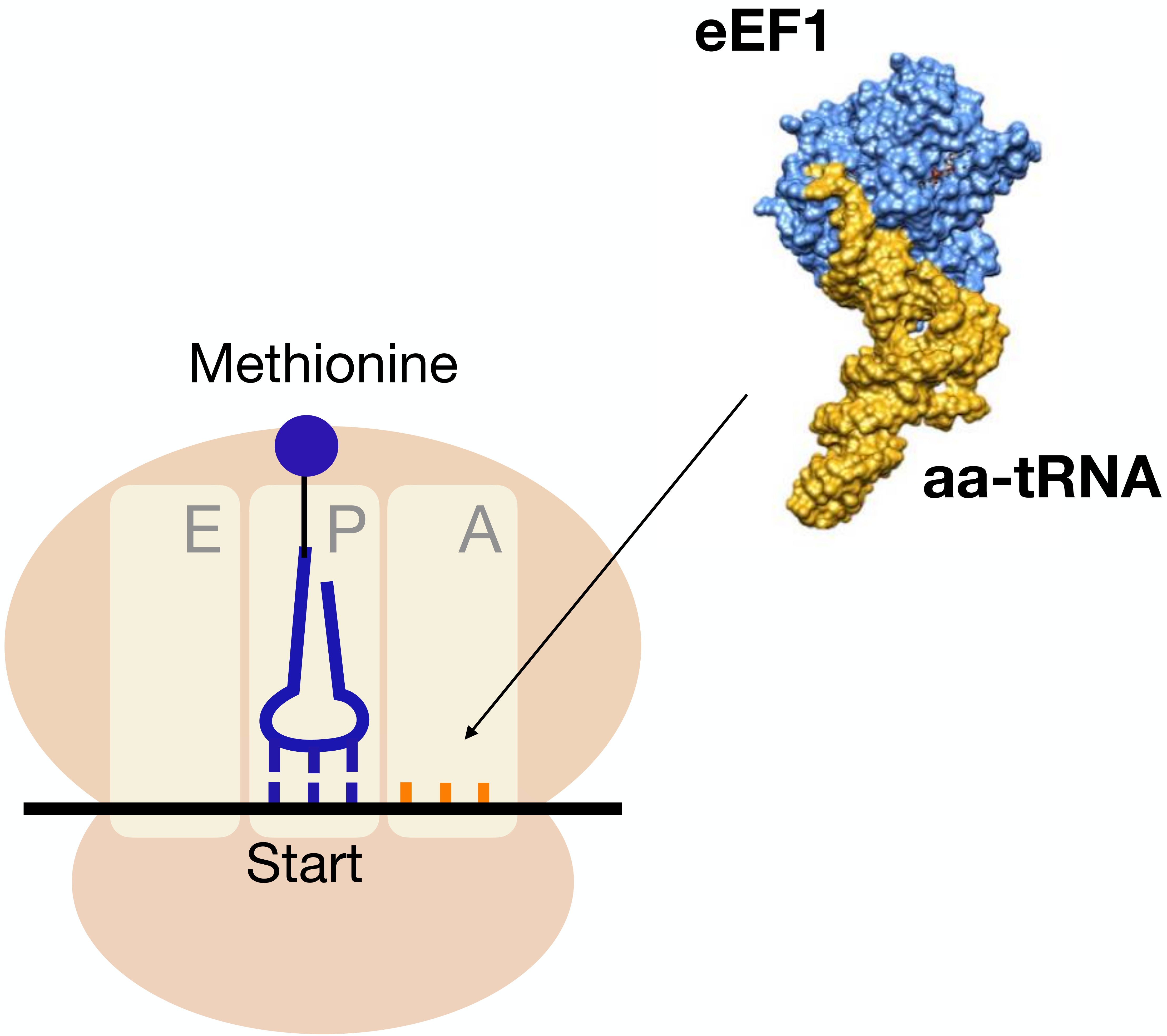


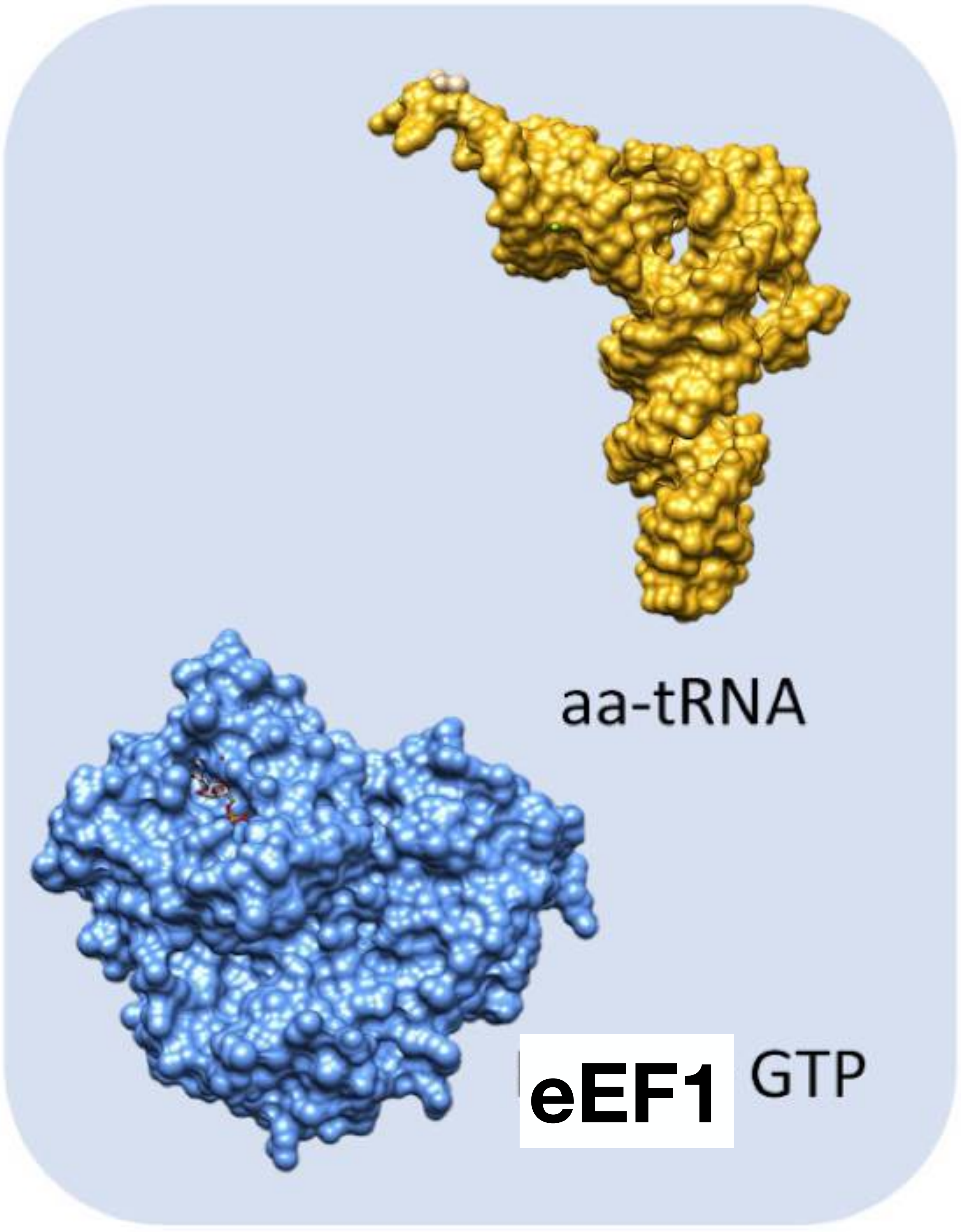
Phase:	Initiation	Elongation	Termination	Recycling
Factors:	eIFs Met-tRNA^{iMet}	eEFs aa-tRNAs	eRFs	ABCE1

Initiation

Elongation

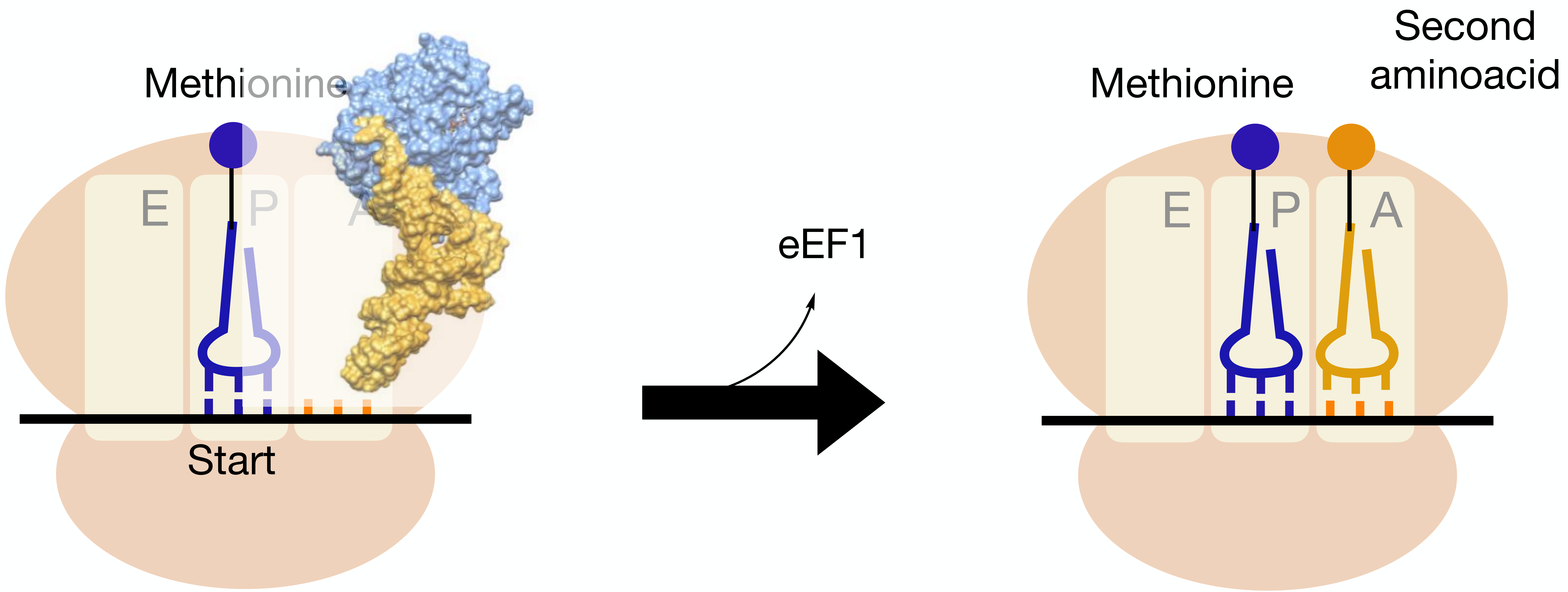
Termination



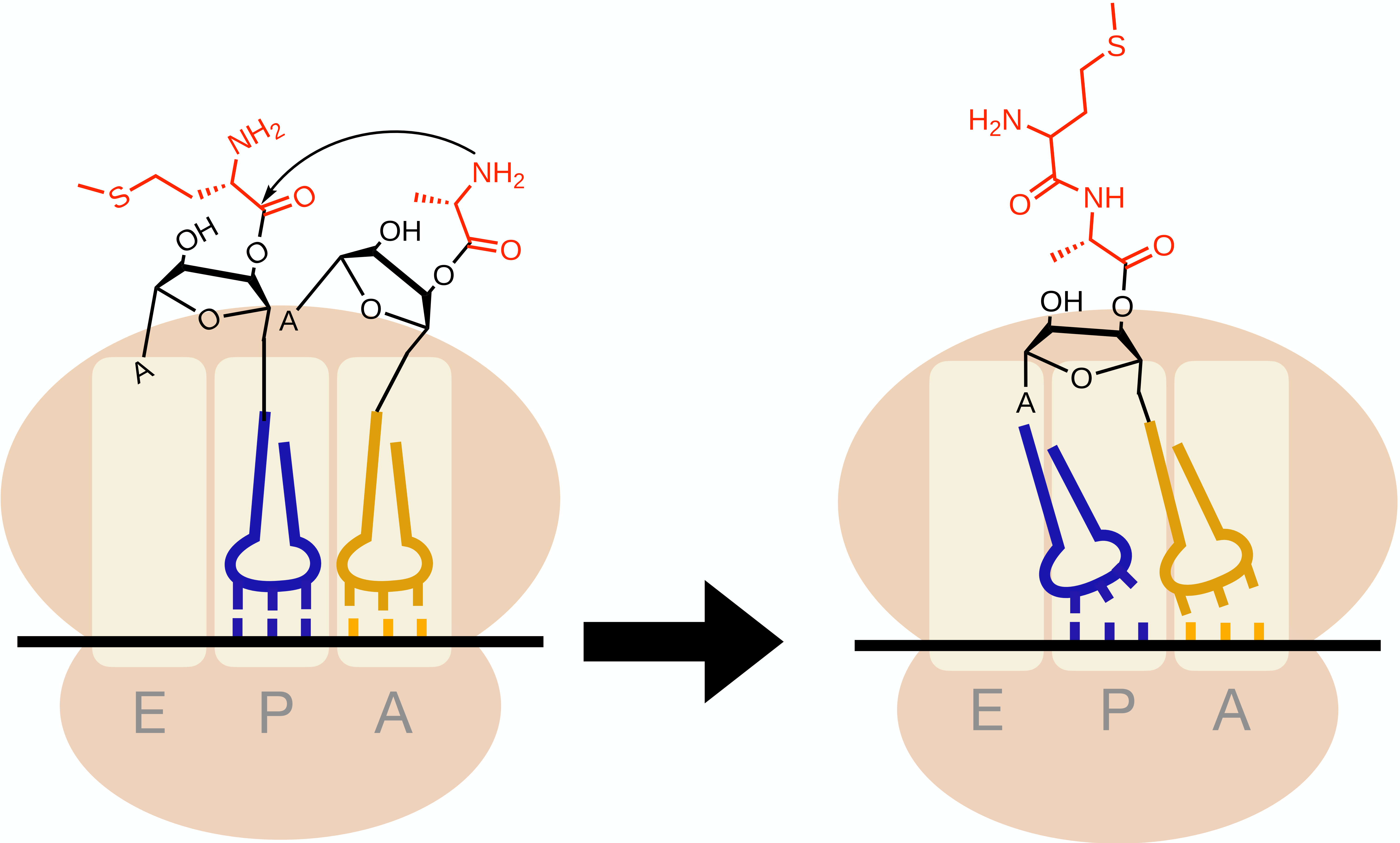


E (exit) P (peptidyl) A (aminoacyl)

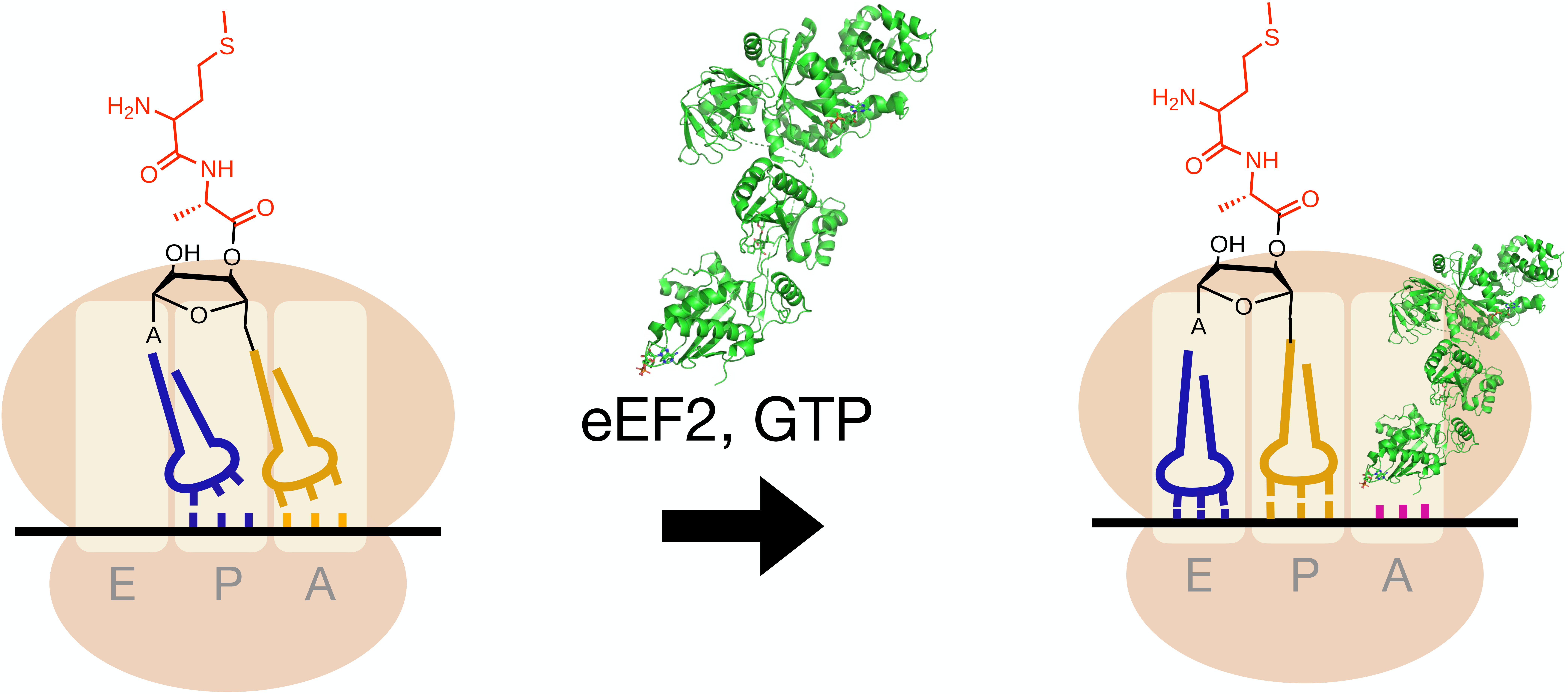
Elongation : if codon-anticodon matches, GTP is hydrolyzed in eEF1, and eEF1 leaves



Elongation : peptide bond formation

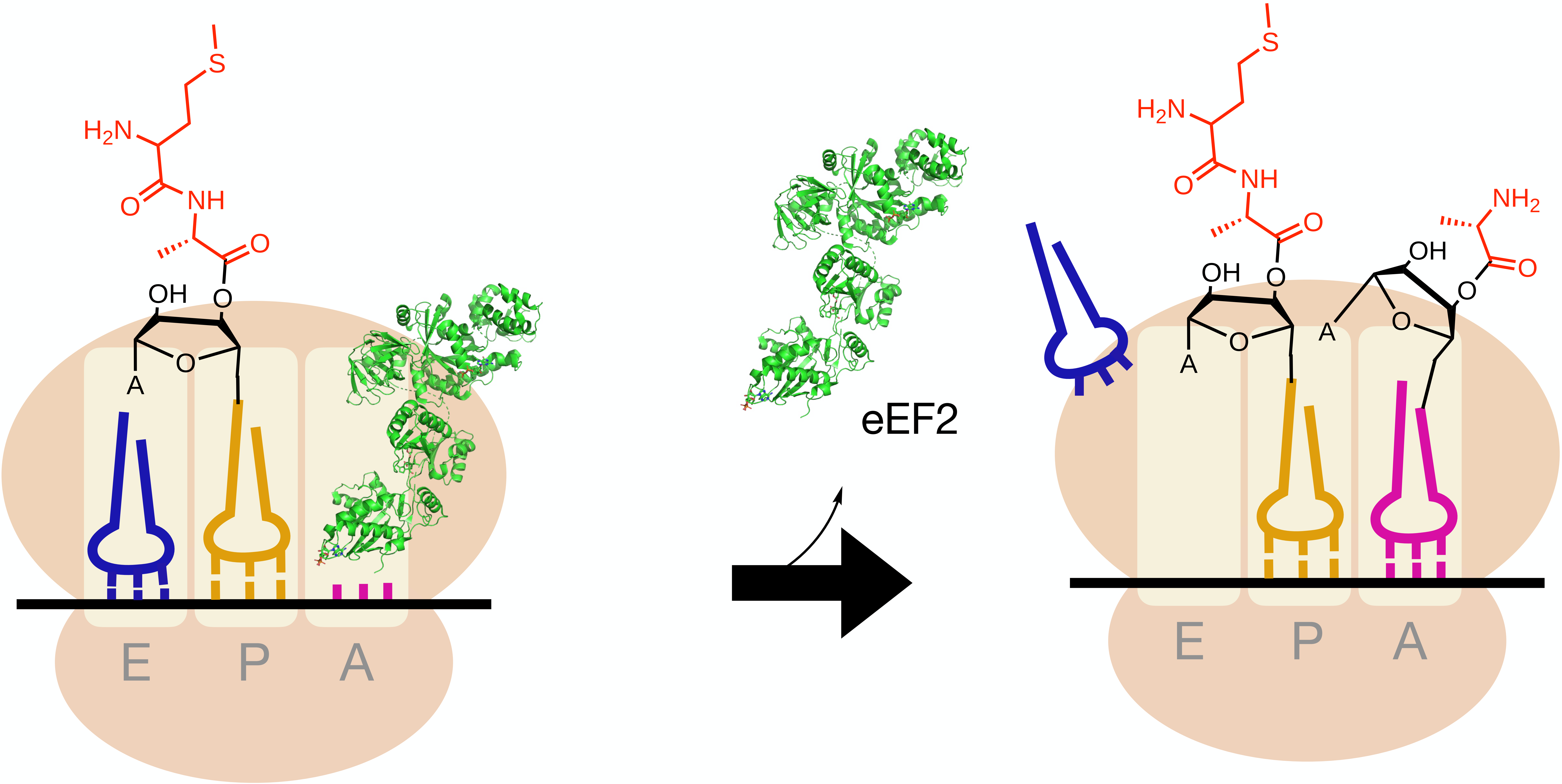


Elongation : tRNA translocation

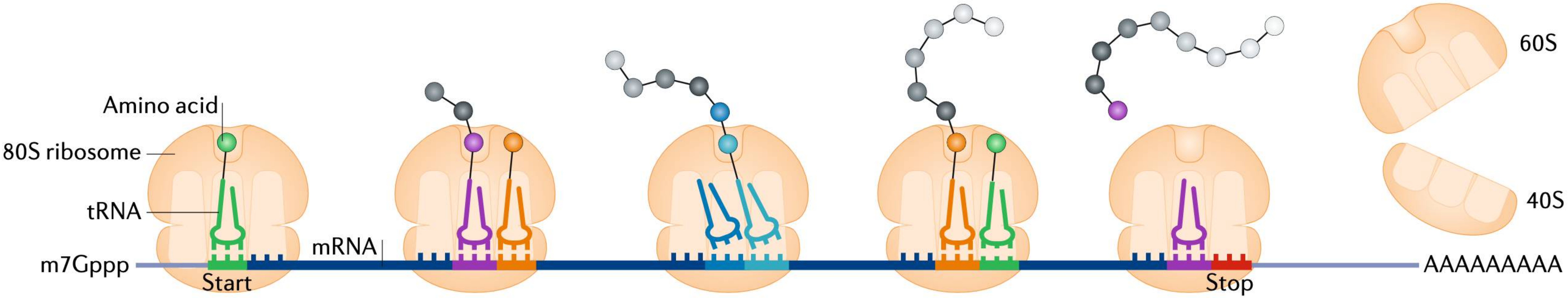


eEF2, GTP

Elongation : peptide bond formation



Overview of eukaryotic translation



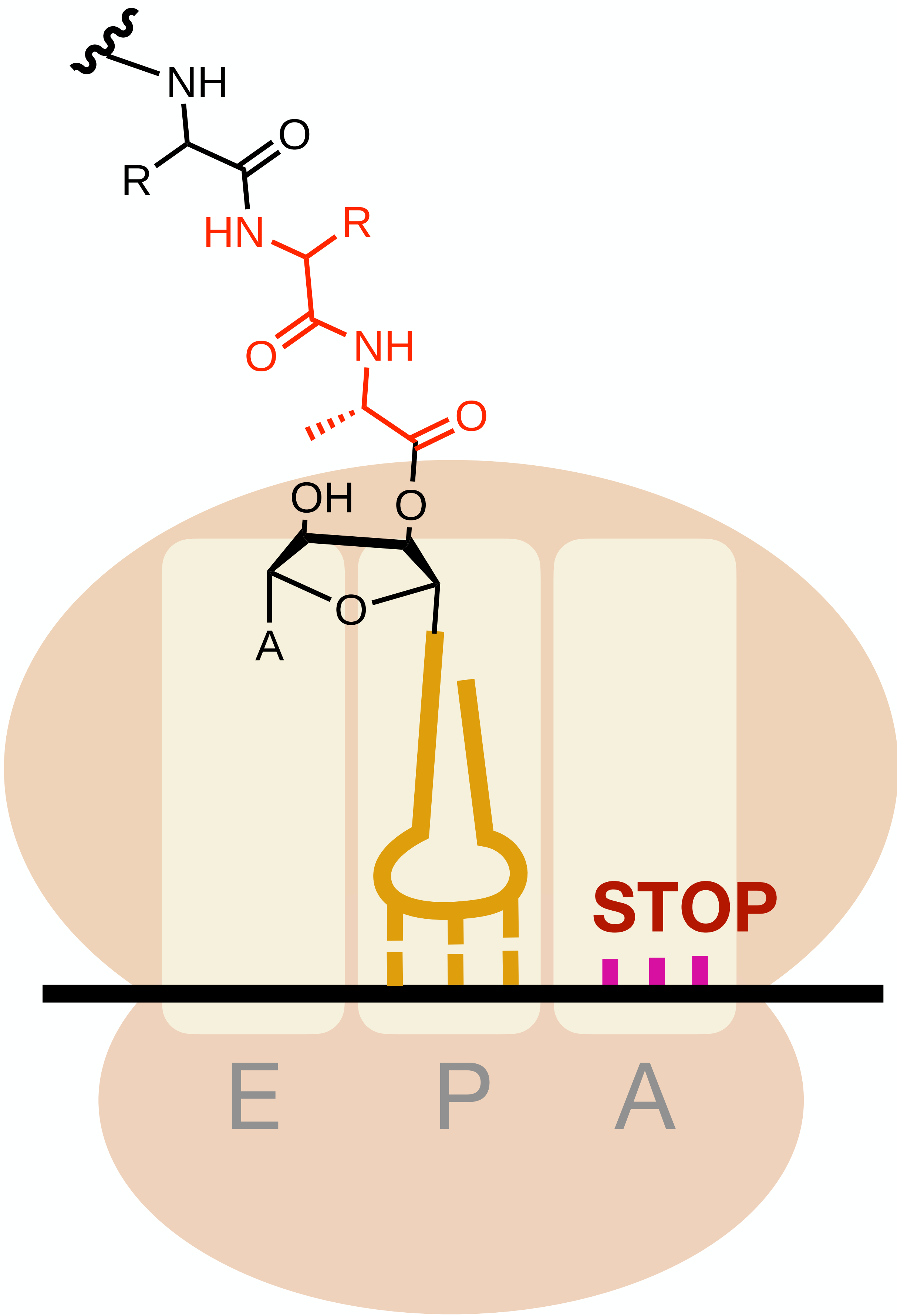
Phase:	Initiation	Elongation	Termination	Recycling
Factors:	eIFs Met-tRNA^{iMet}	eEFs aa-tRNAs	eRFs	ABCE1

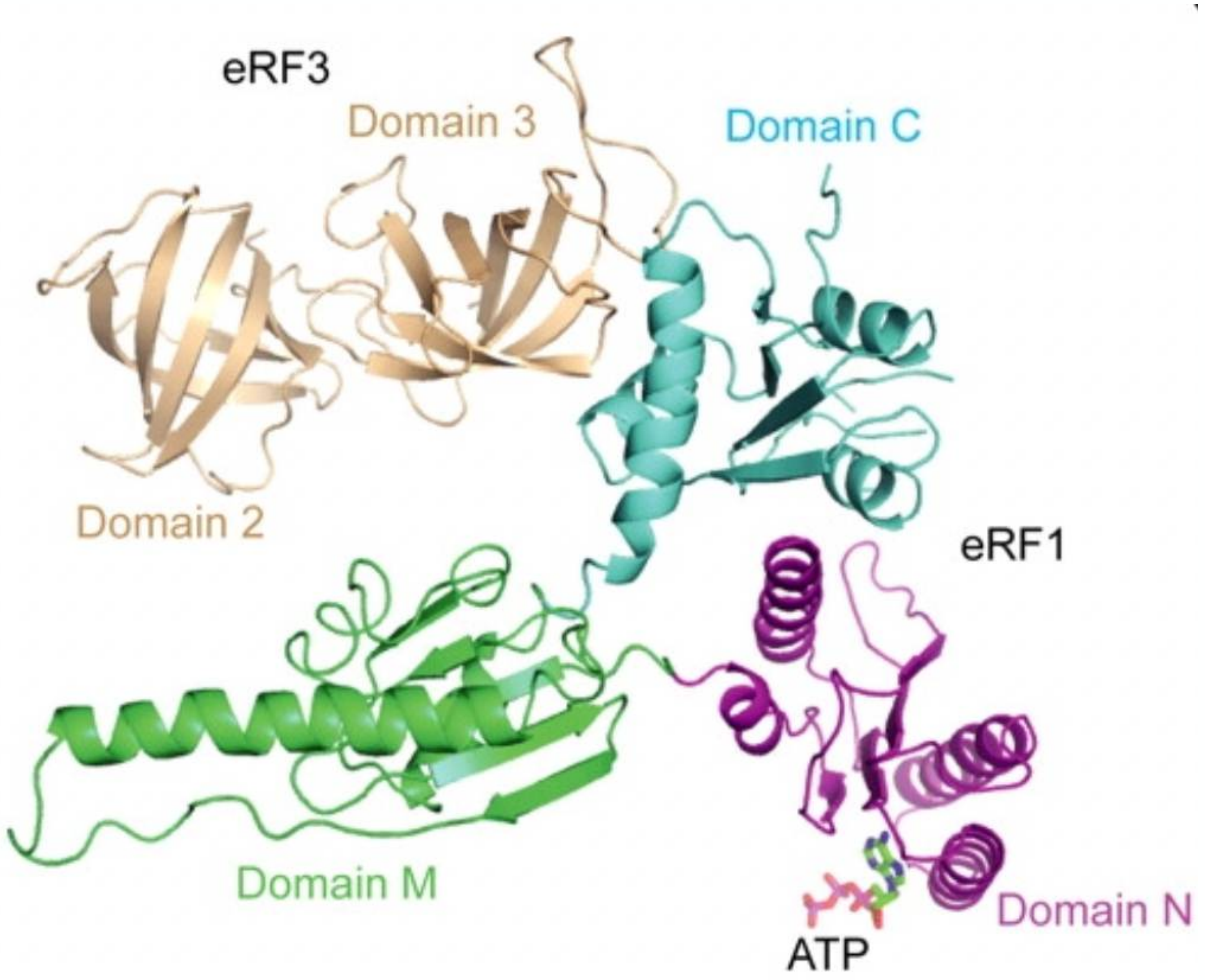
Initiation

Elongation

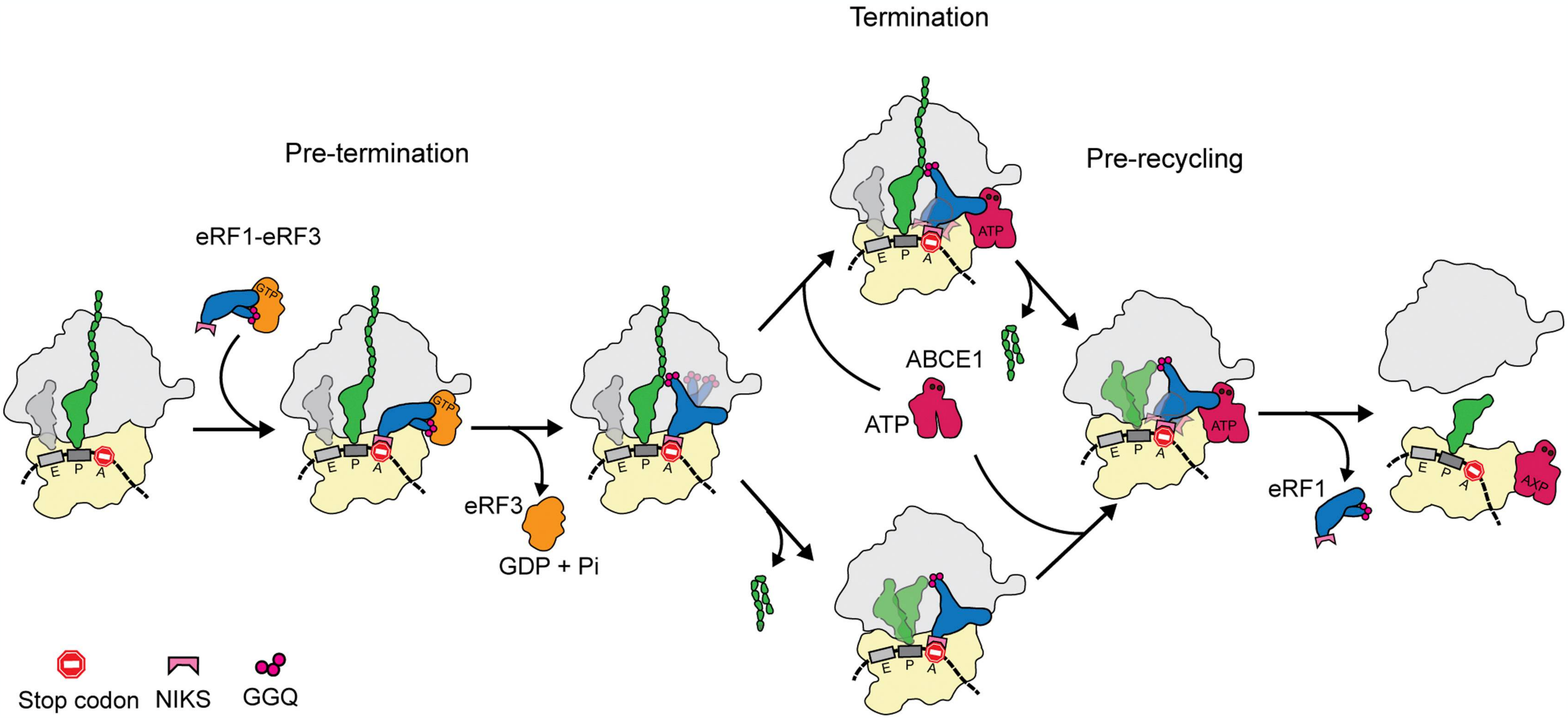
Termination

Termination: eRF1/eRF3 complex

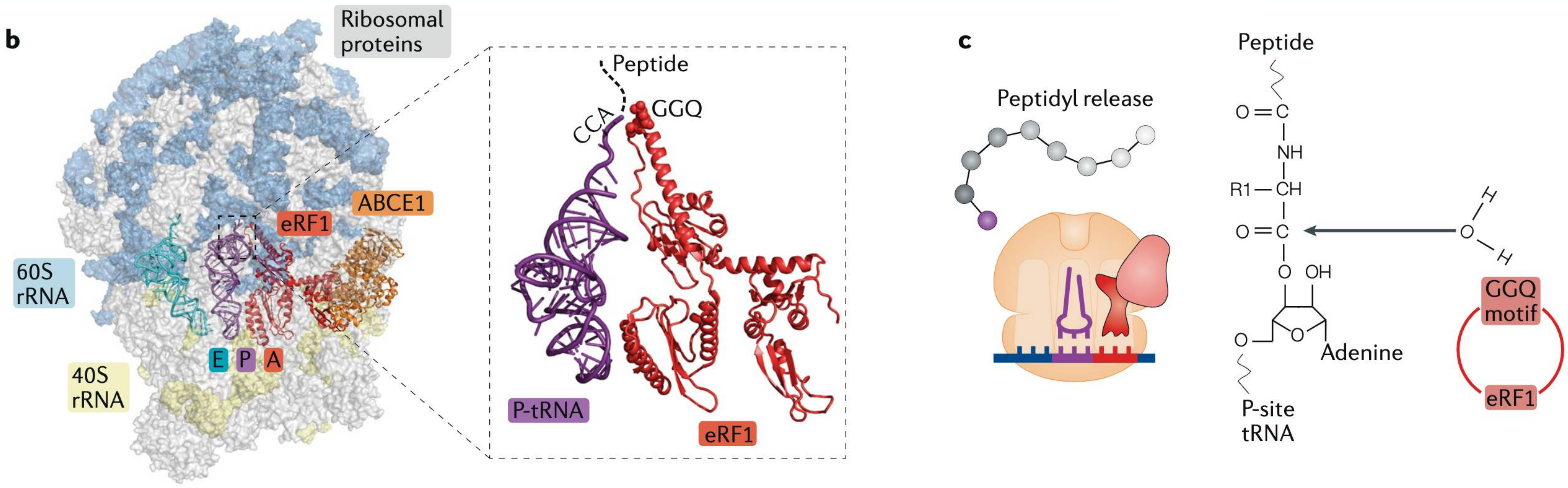




Termination: eRF1/eRF3 complex

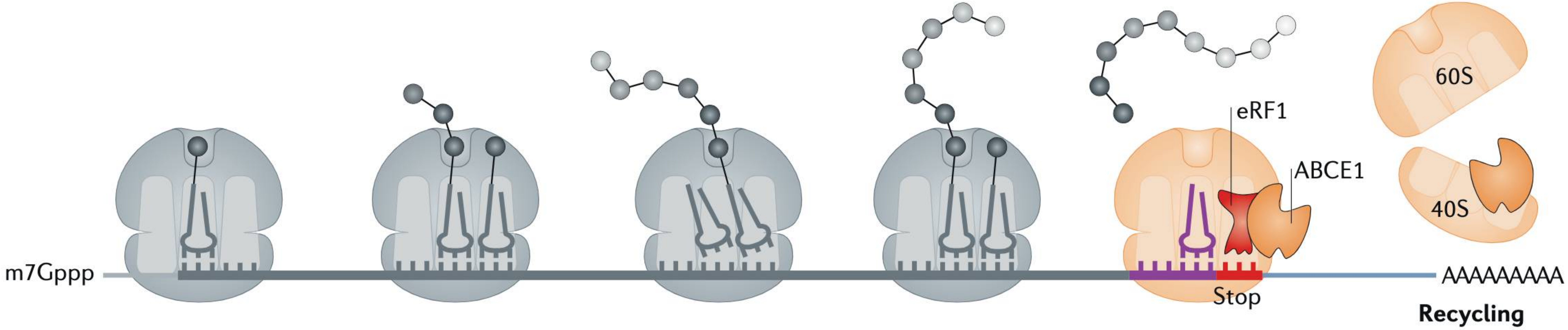


Translation termination and the role of mRNP context



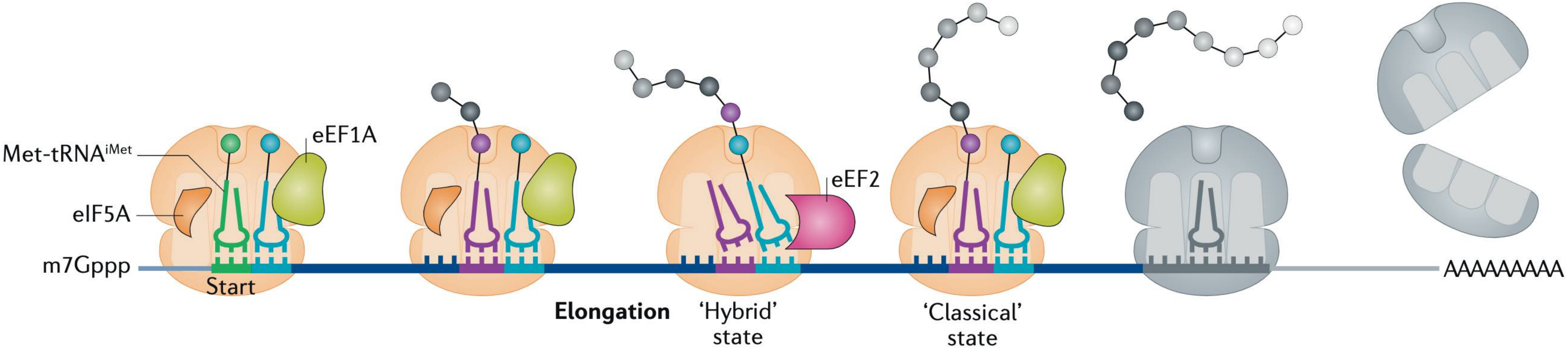
Schematic of peptidyl- release reaction coordinated by eRF1.

Ribosome recycling



Overview of ribosome recycling by ATP-binding cassette subfamily E member 1 (ABCE1). ABCE1 binds to 80S ribosomes loaded with eukaryotic peptide chain release factor subunit 1 (eRF1) and uses the power generated from ATP binding and hydrolysis to dissociate the ribosomal subunits. ABCE1 remains bound to the 40S subunit to stimulate subsequent translation initiation steps.

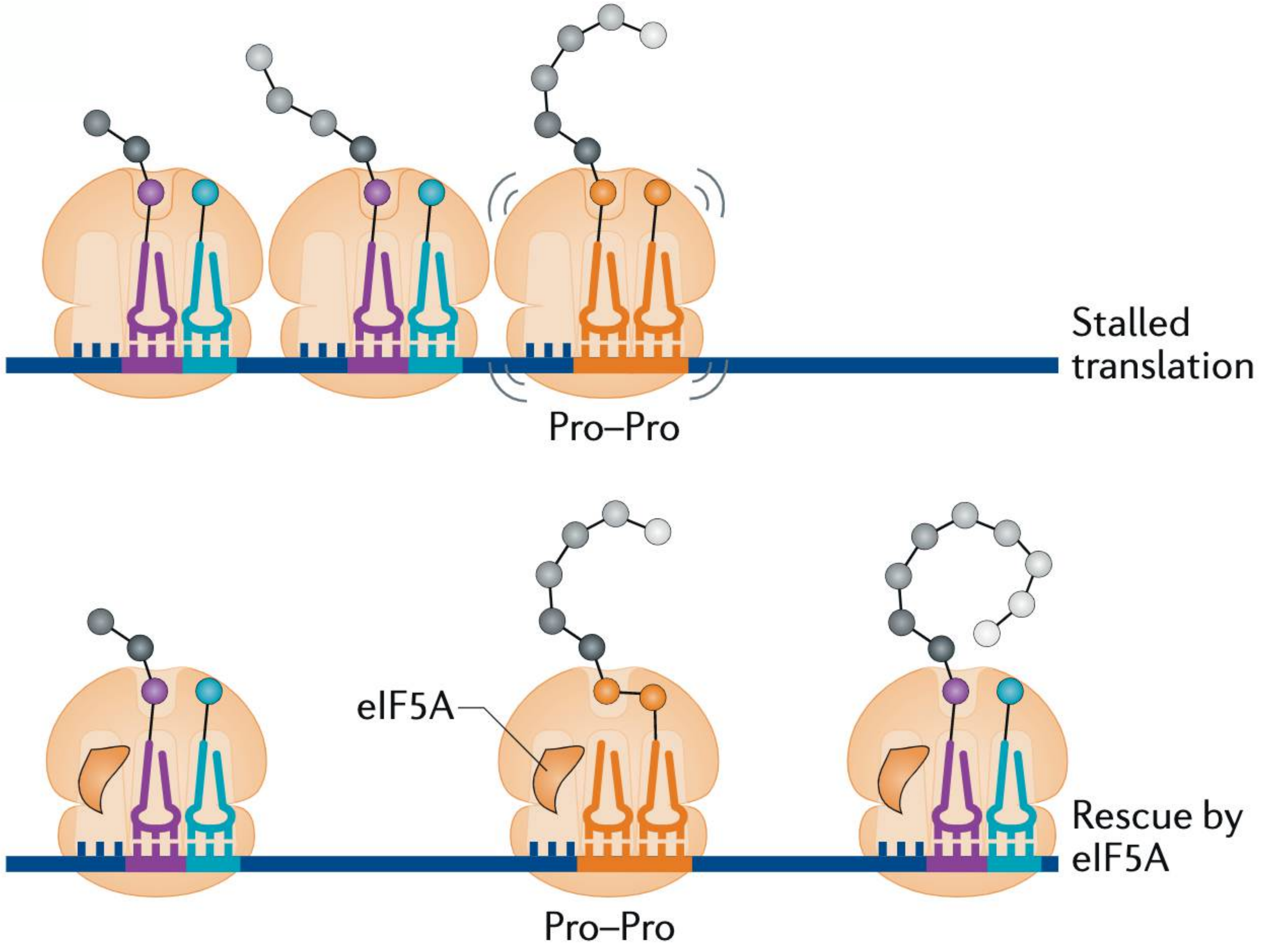
Translation elongation and resolution of ribosome stalling



What happens when something goes wrong in this highly conserved protein machinery pathway?

What can cause trouble and how to resolve it?

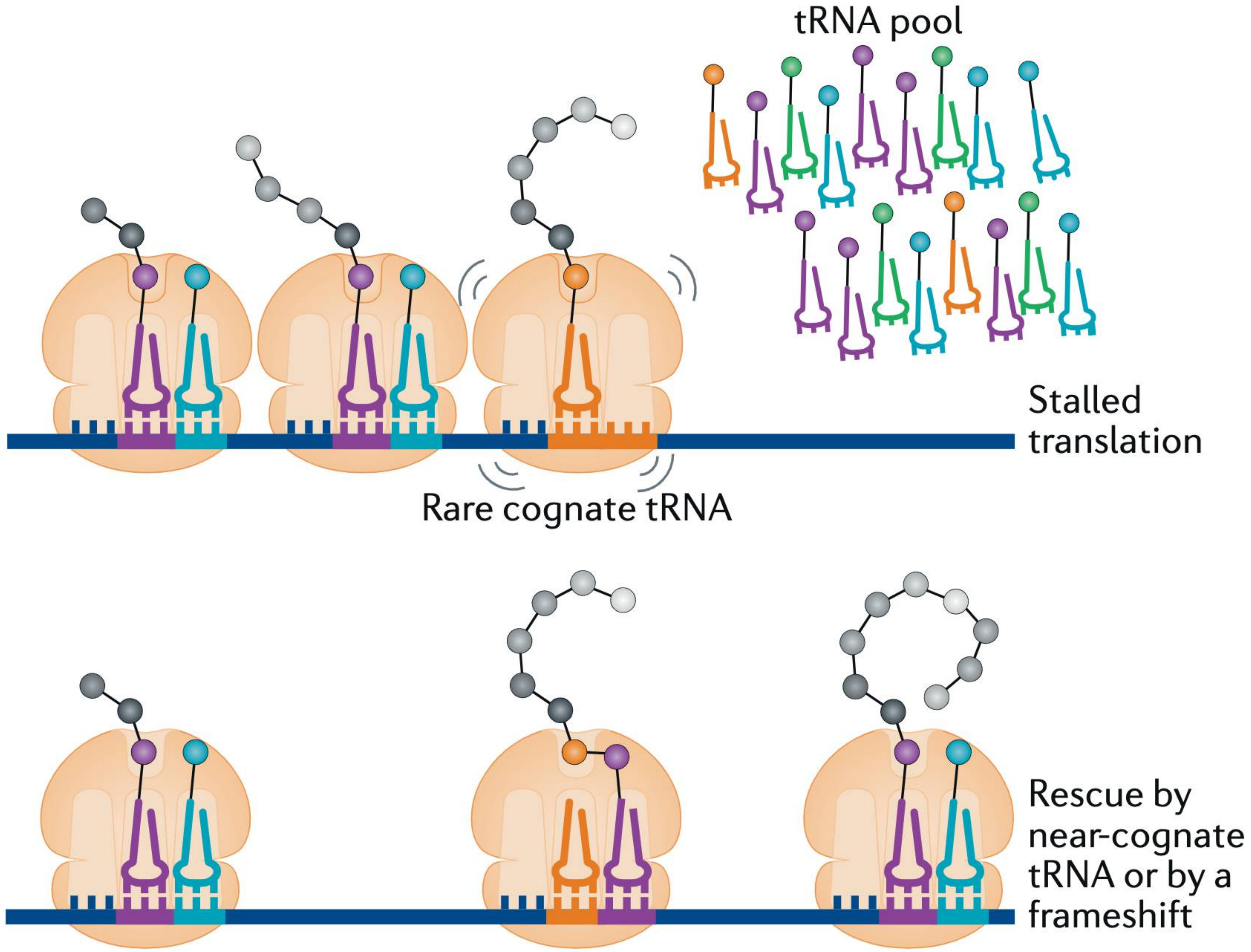
Translation elongation and resolution of ribosome stalling



Example 1:

Ribosome stalling due to slow peptidyl-transfer kinetics (such as during the formation of Pro–Pro) is rescued by eIF5A, which promotes peptide-bond formation.

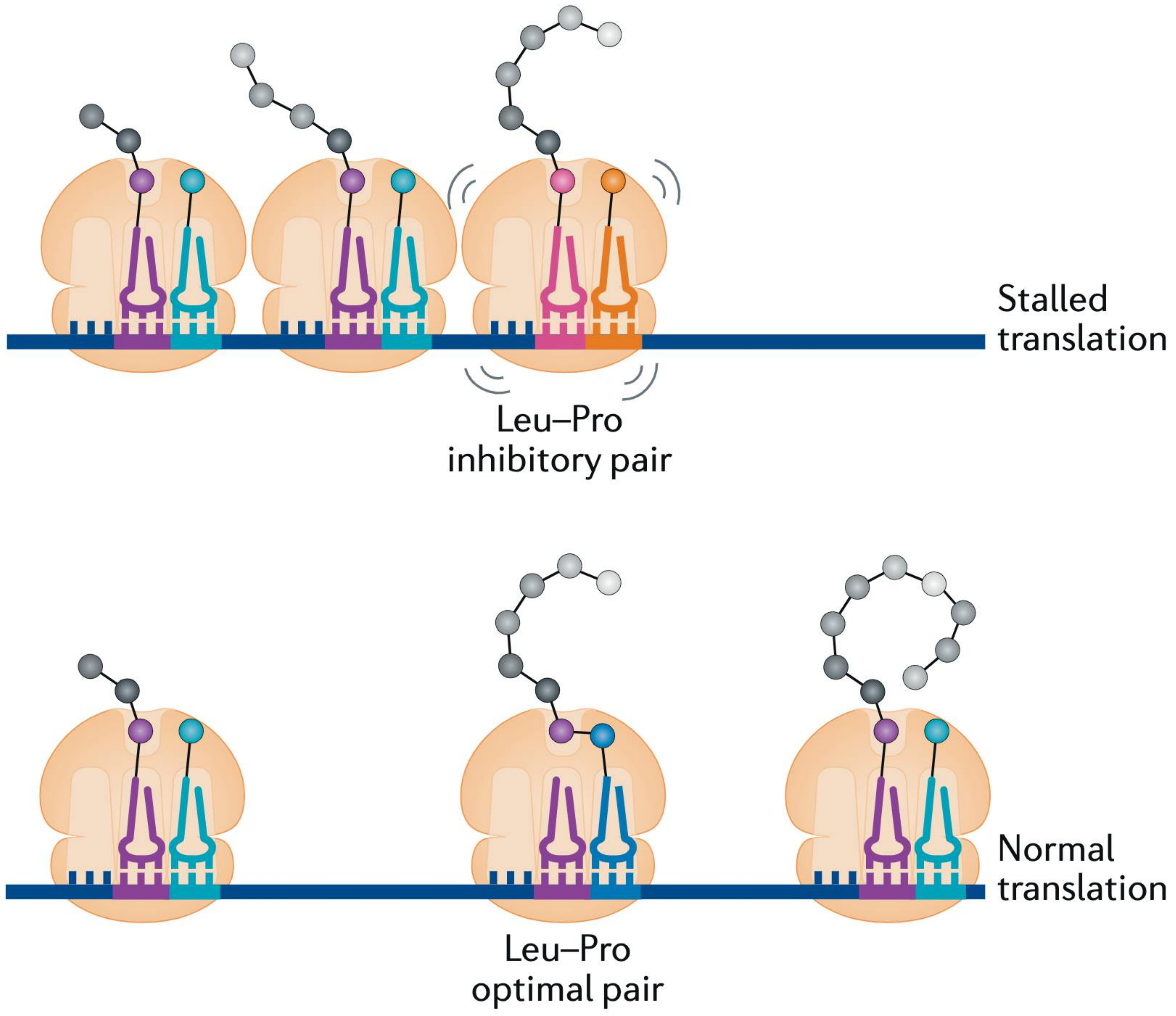
Translation elongation and resolution of ribosome stalling



Example 2:

Ribosome stalling caused by poor A-site occupancy resulting from poorly expressed tRNAs or from poor tRNA aminoacylation can be rescued by mis-incorporation of near-cognate tRNAs or by frame shifting (represented as conversion of orange to purple).

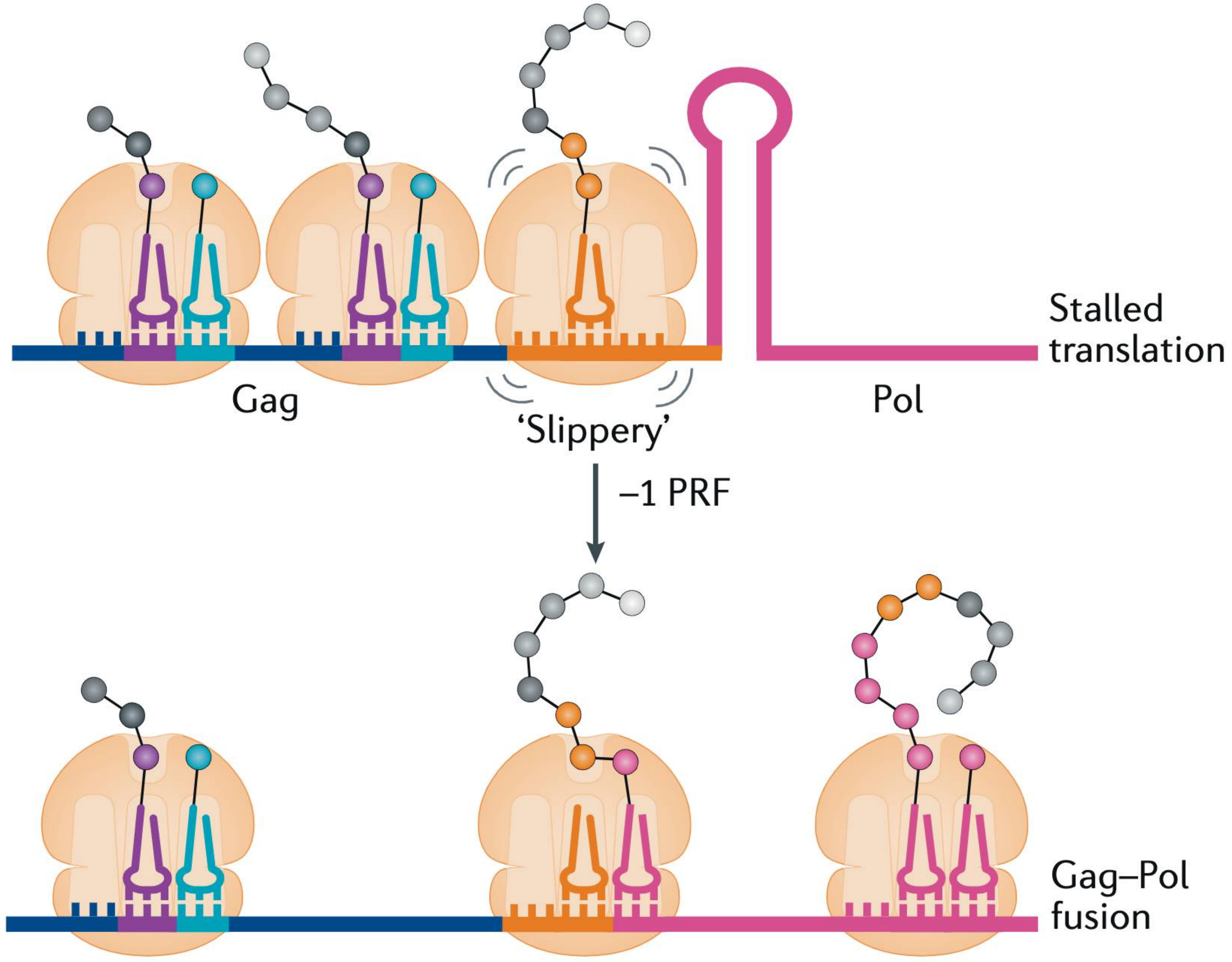
Translation elongation and resolution of ribosome stalling



Example 3:

Ribosome stalling can be caused by certain consecutive tRNA–codon pair orders that are suboptimal (pink and orange) relative to synonymous pairs (purple and blue).

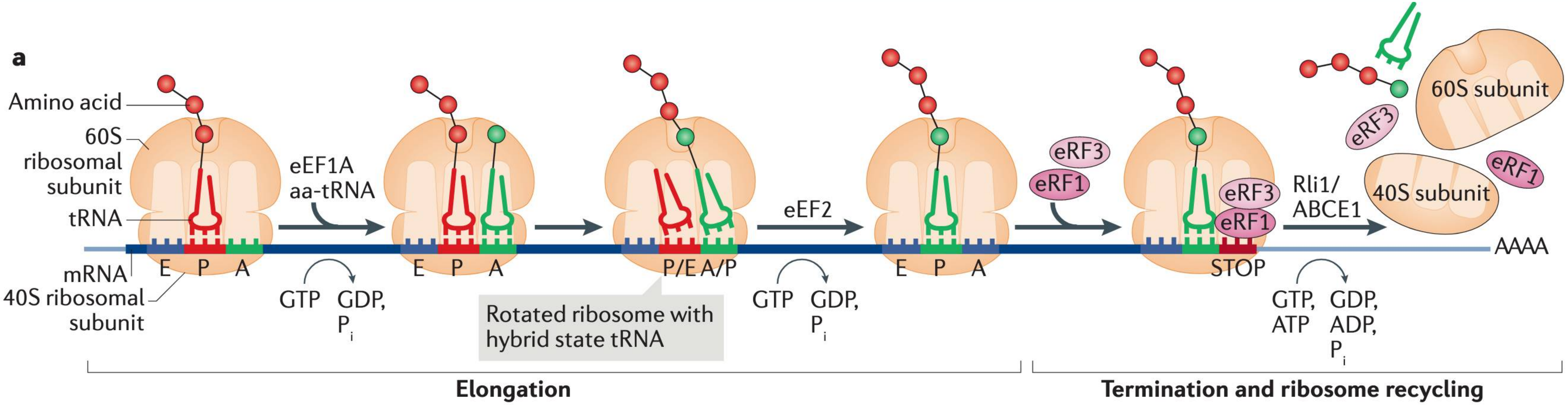
Translation elongation and resolution of ribosome stalling



Example 4:

Ribosome stalling caused by mRNA secondary structures can be resolved by programmed ribosomal frame shifting (PRF) at adjacent 'slippery' sequences. The example illustrates the –1 PRF that is required for translation of the Gag–Pol fusion protein of HIV

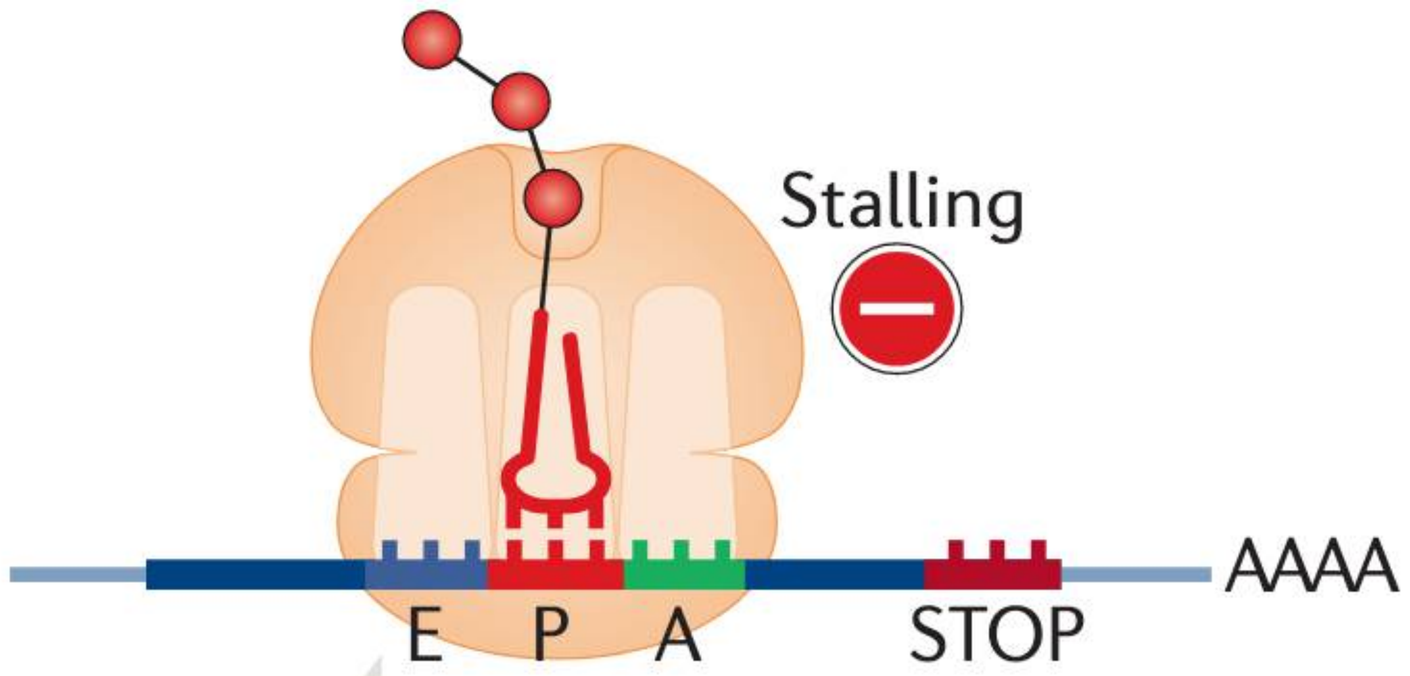
Ribosome Quality control (and rescue)



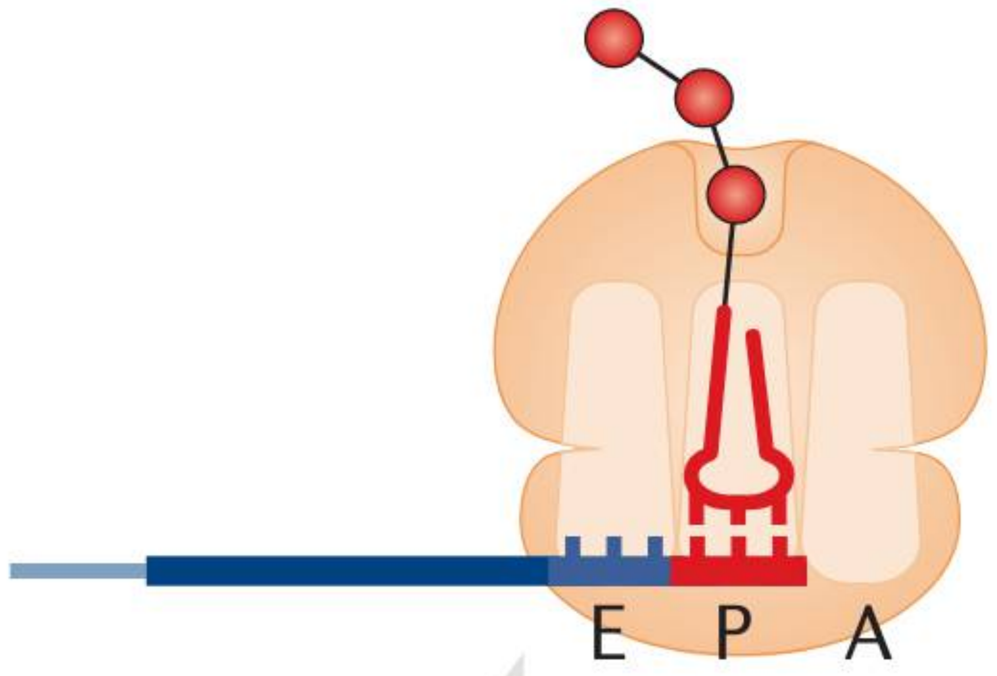
During the translation cycle, ribosome can stall for different reasons at different steps of the cycle

What happens when stalling can't easily be resolved?

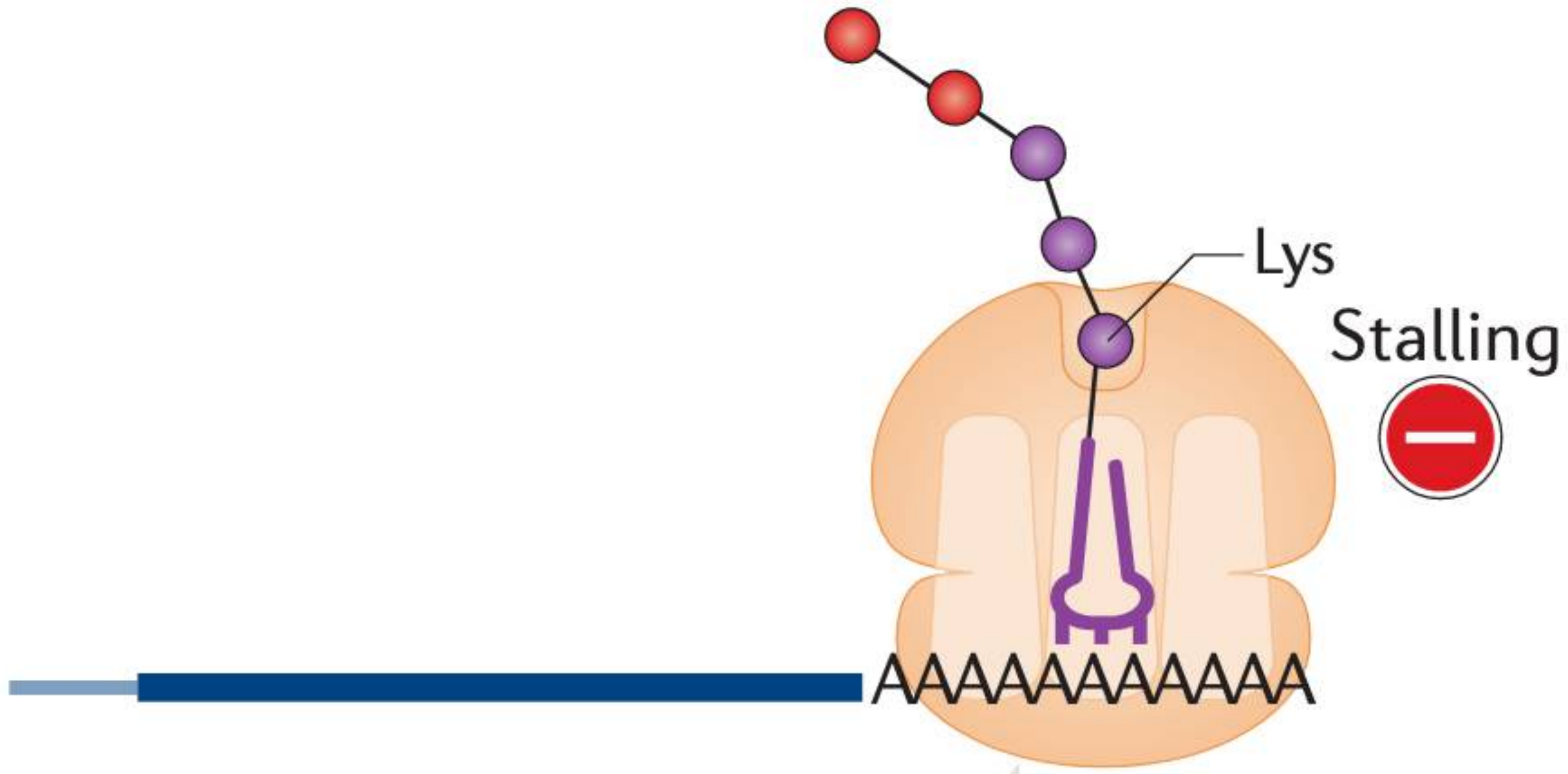
Ribosome Quality Control Upon Stalling



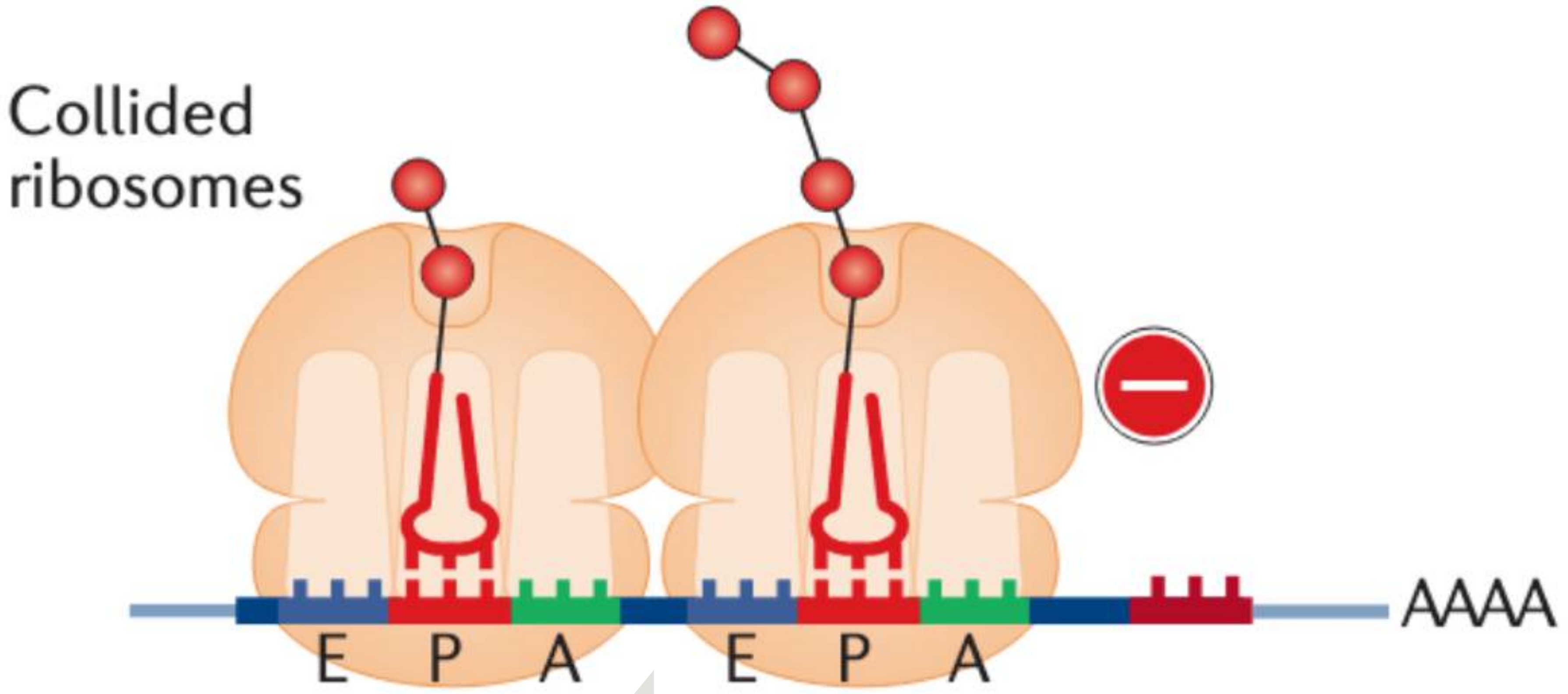
- mRNA or tRNA with abnormal features or damage
- Immature ribosome produced by defective 60S subunit biogenesis



Nonstop mRNA generated by endonuclease cleavage

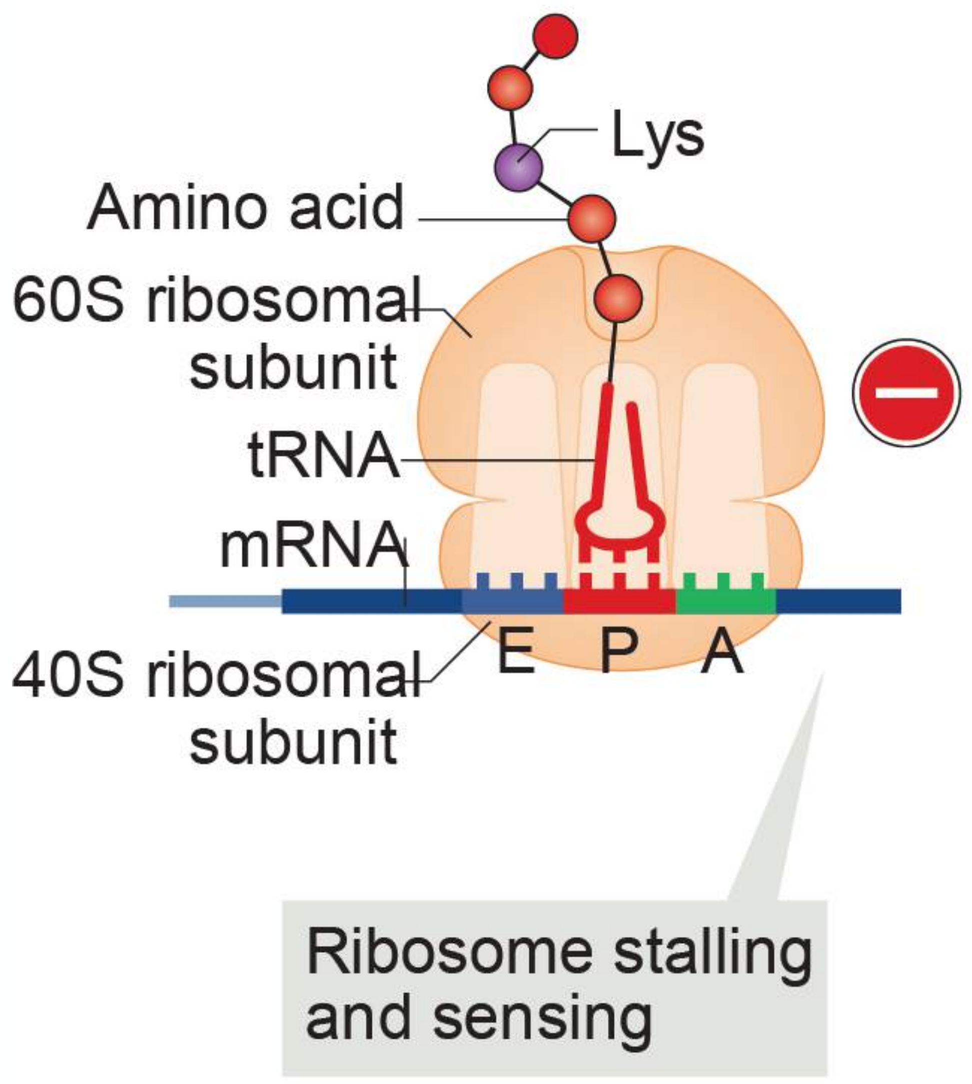


Nonstop mRNA generated by premature mRNA polyadenylation



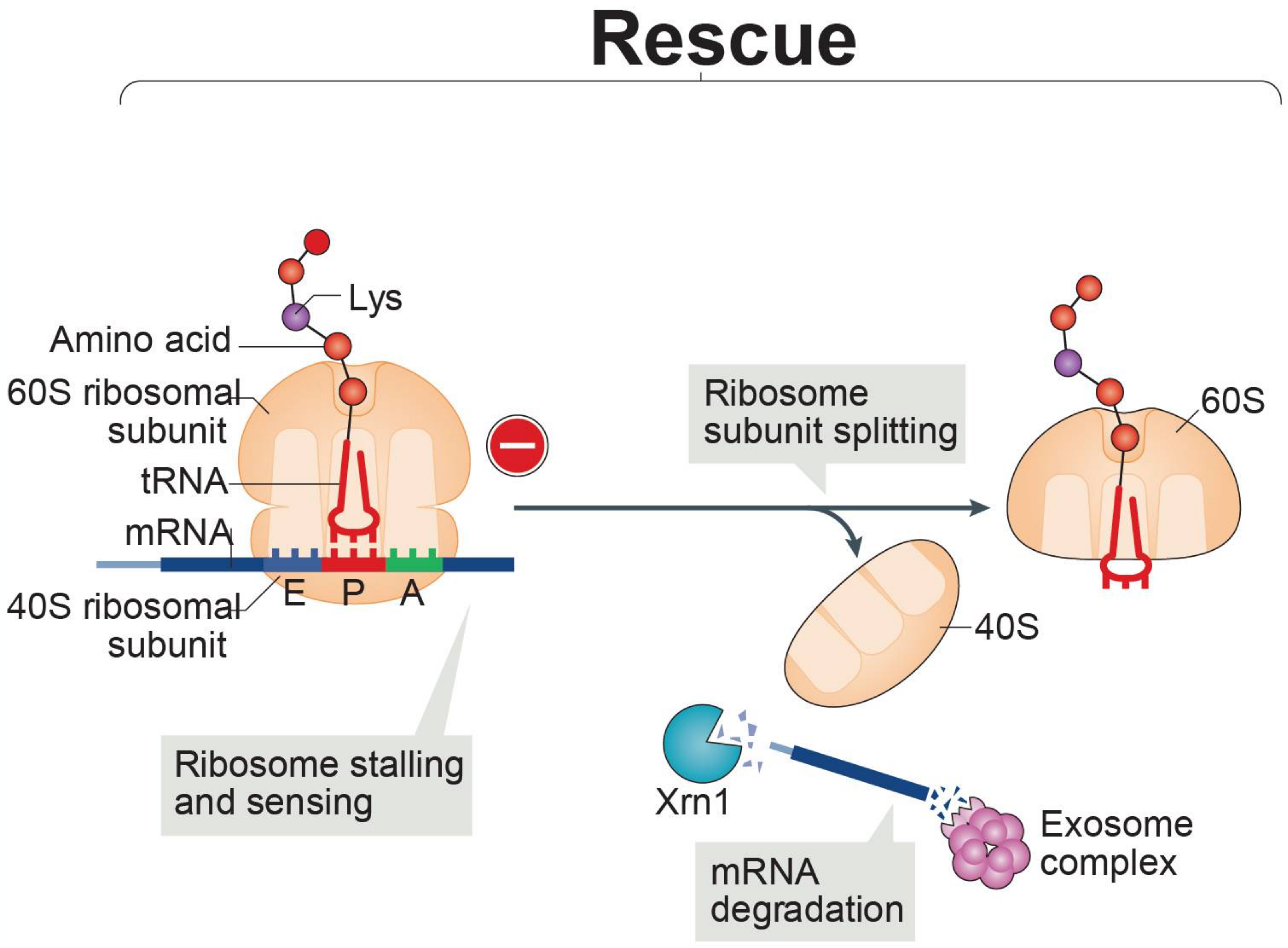
Traffic jam as consequence of non-optimal or rare codons, or chemical damage inducing roadblock

Ribosome Quality Control Upon Stalling



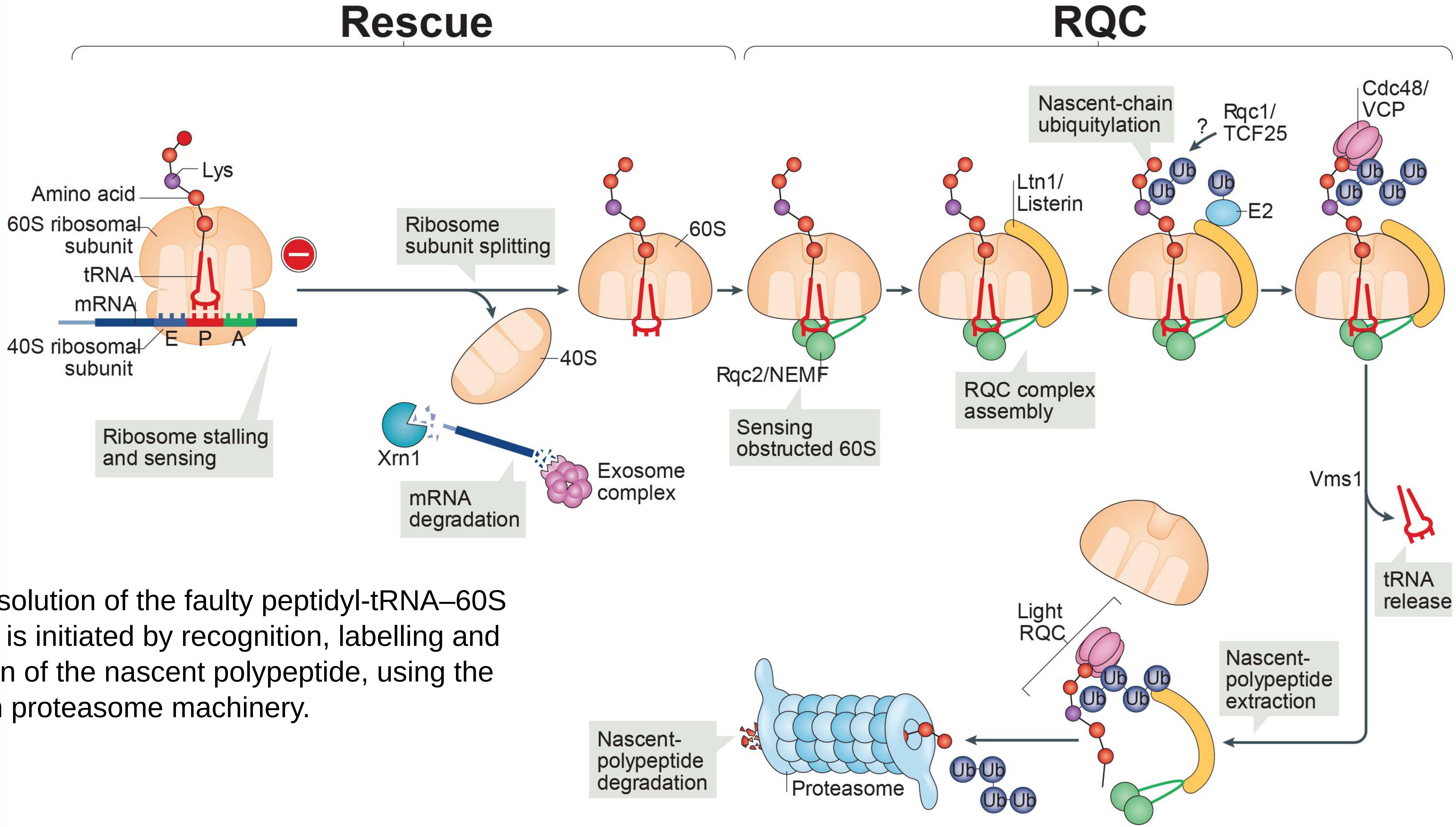
Ribosome stalling and sensing

Ribosome Quality Control Upon Stalling



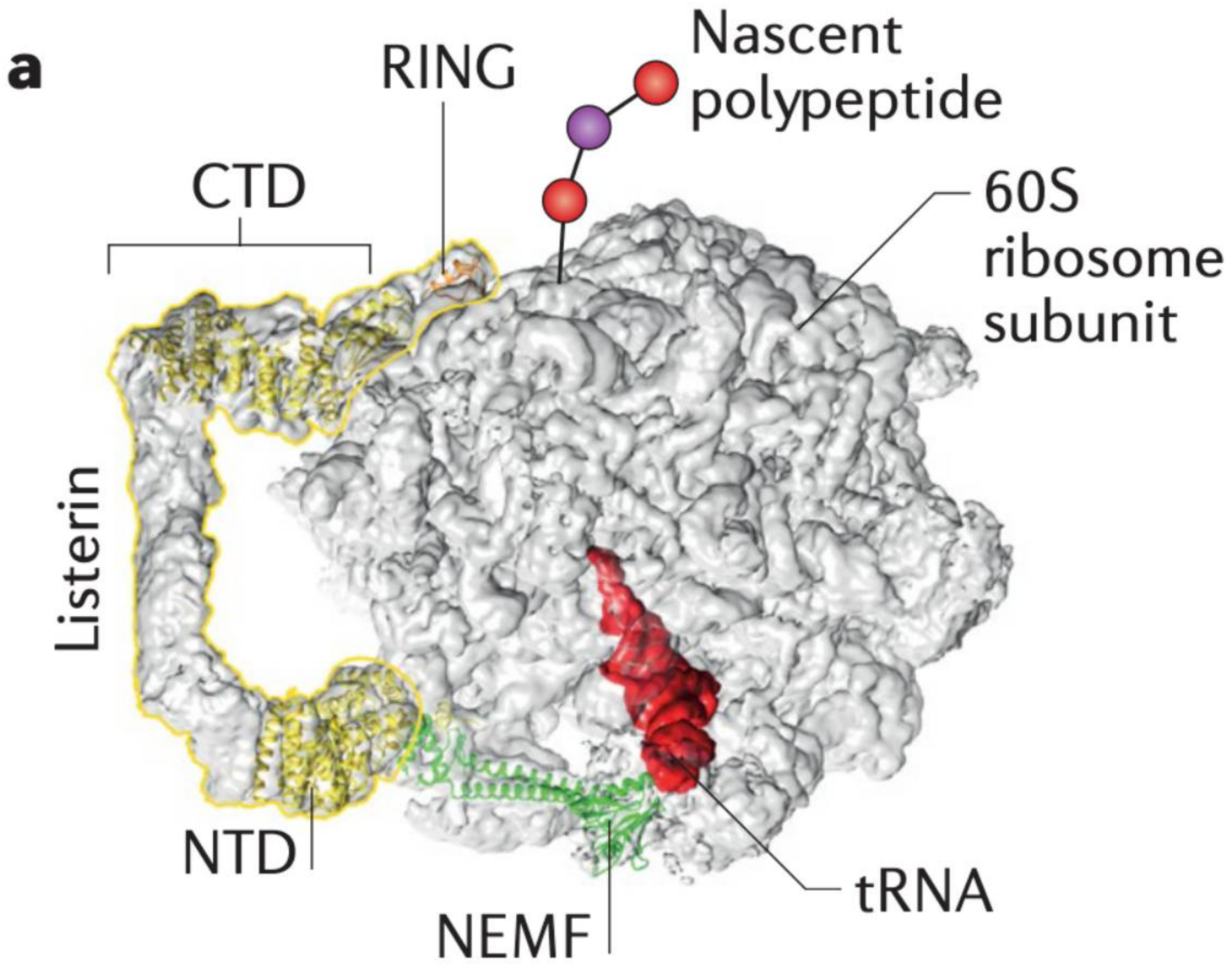
First step to any stalling,
is Ribosome splitting
(large from small subunit)

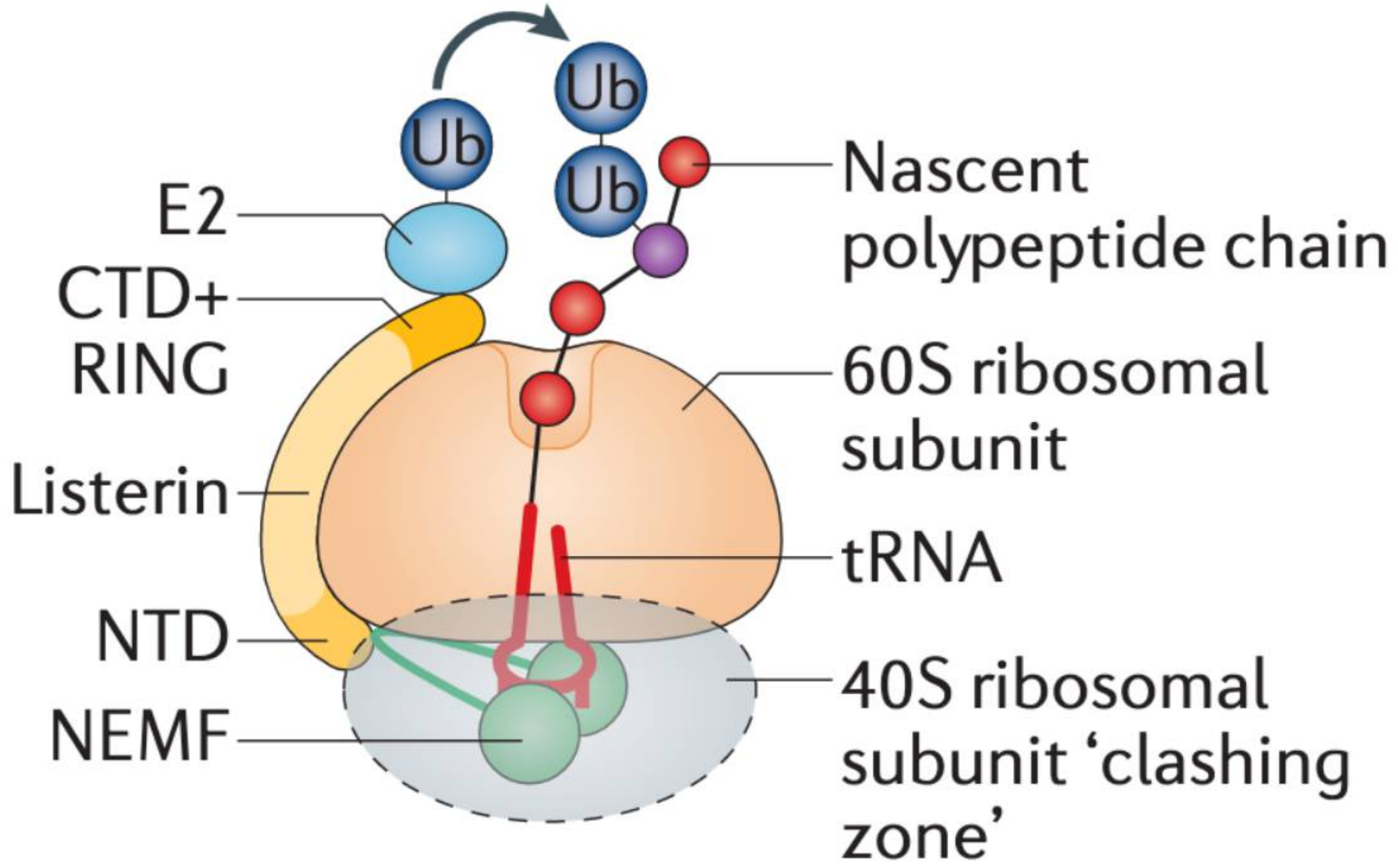
Ribosome Quality Control Upon Stalling



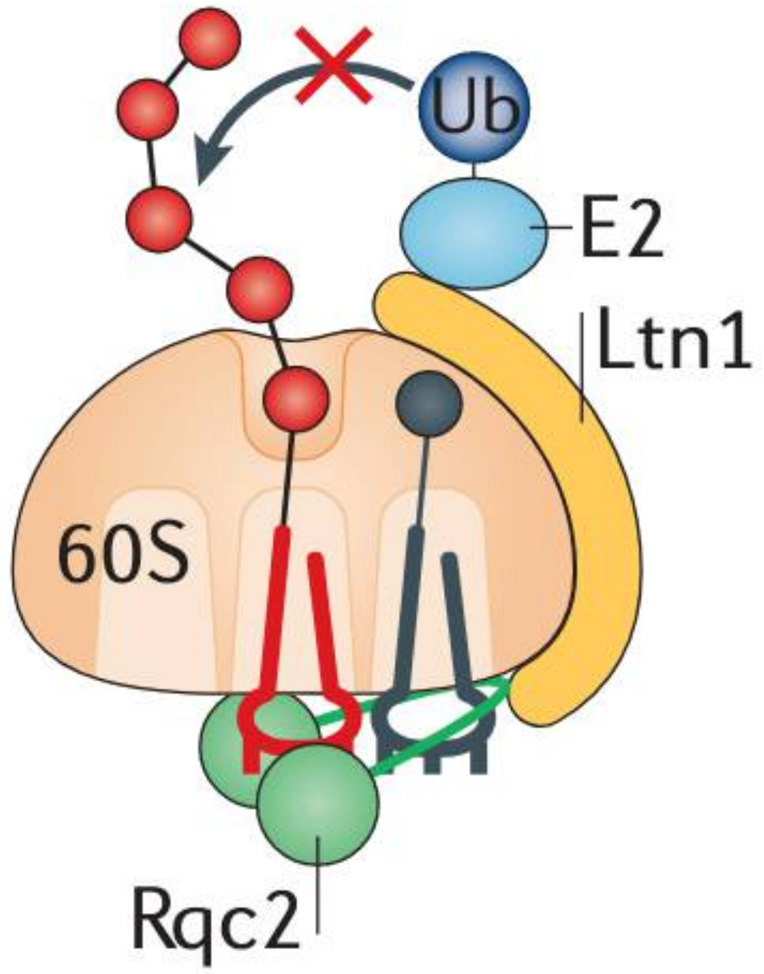
RQC: resolution of the faulty peptidyl-tRNA–60S complex is initiated by recognition, labelling and extraction of the nascent polypeptide, using the Ubiquitin proteasome machinery.

Ribosome Quality control



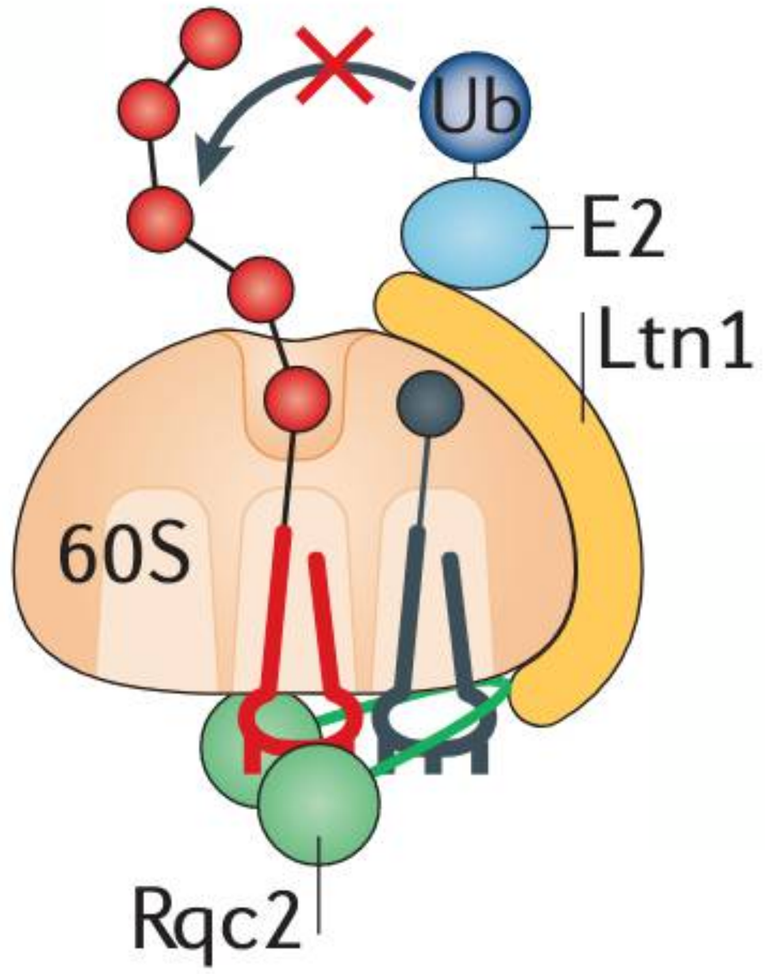


Ribosome Quality control



What happens if the nascent poly-peptide does not contain exposed Lysine residues??

Ribosome Quality control



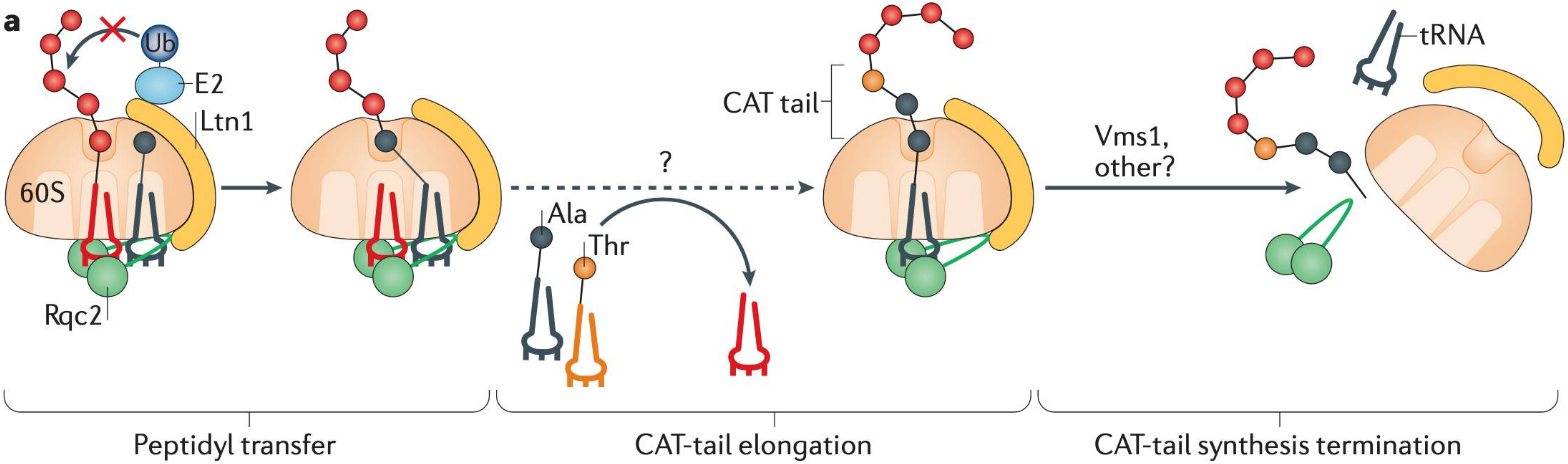
What happens if the nascent poly-peptide does not contain exposed Lysine residues??

↓

CAT-tail backup mechanism

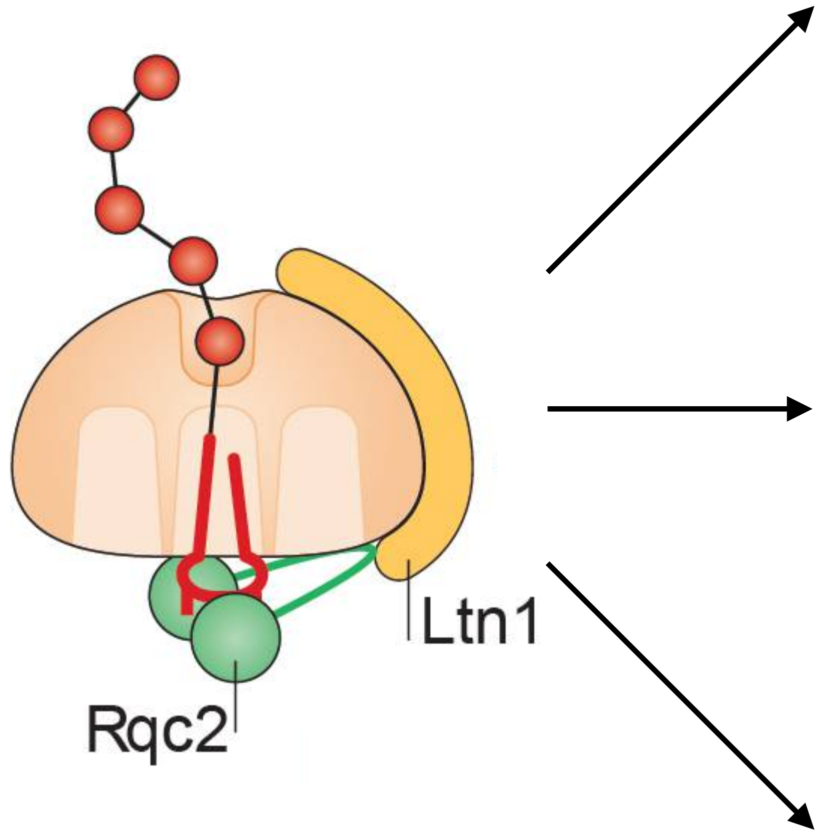
Ribosome Quality control

Carboxy-terminal alanine and threonine tail (CAT tail) synthesis and functions



Ribosome Quality control

Overview of the possible RQC resolution pathway depending on polypeptide context



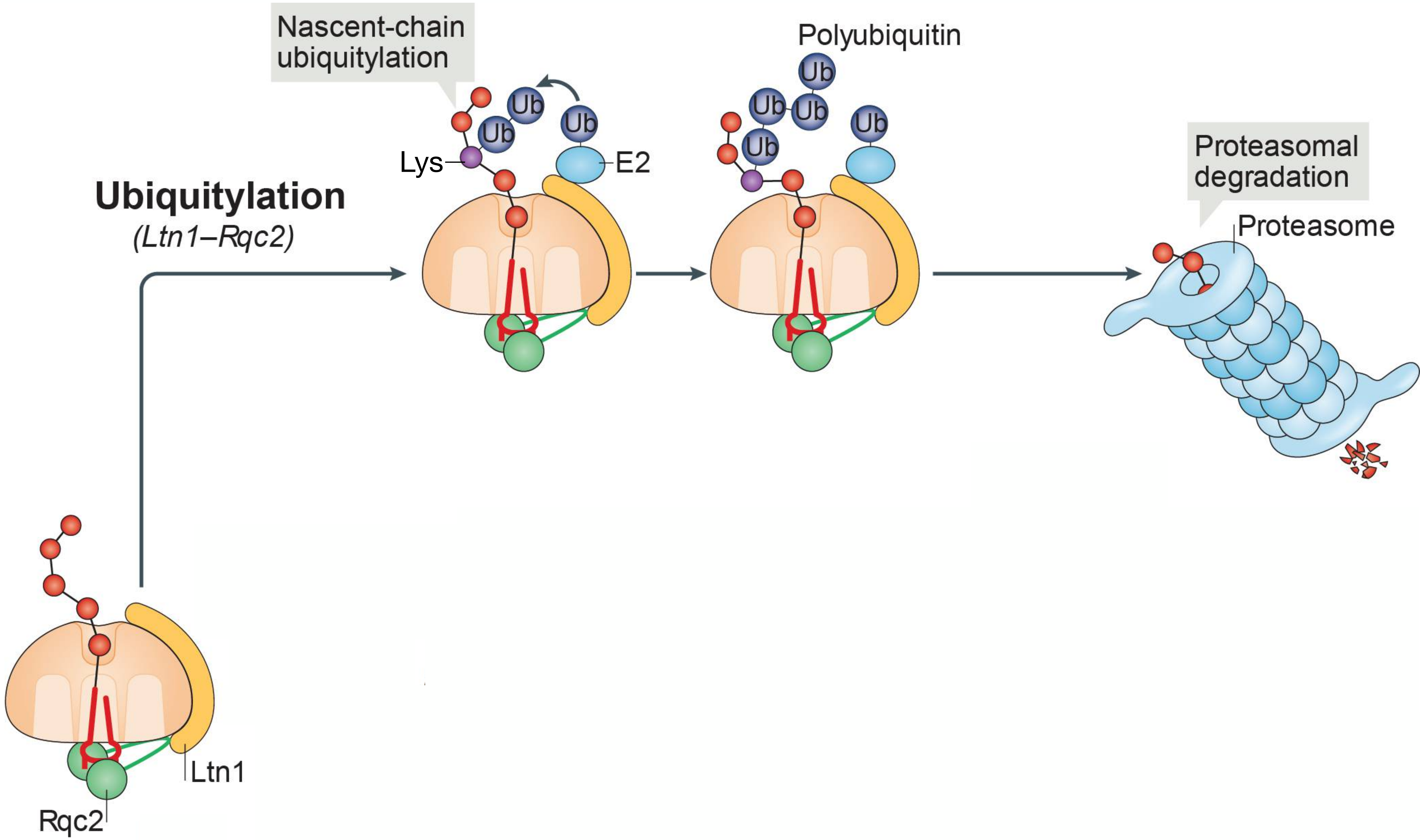
Lysine present and exposed

Lysine present but hidden

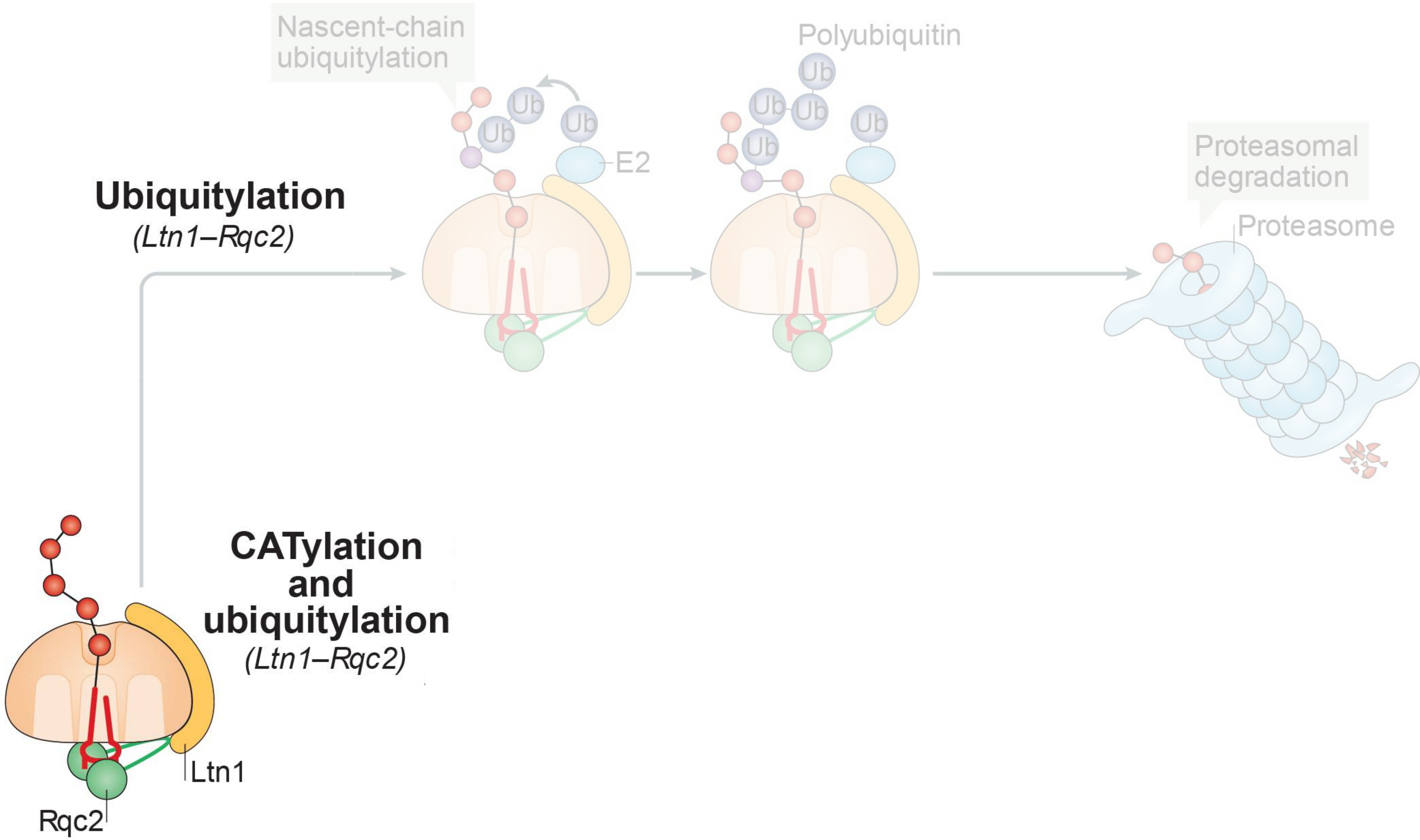
Lysine absent

Ubiquitin needs a Lysine to attache on a polypeptide.

Ribosome Quality control



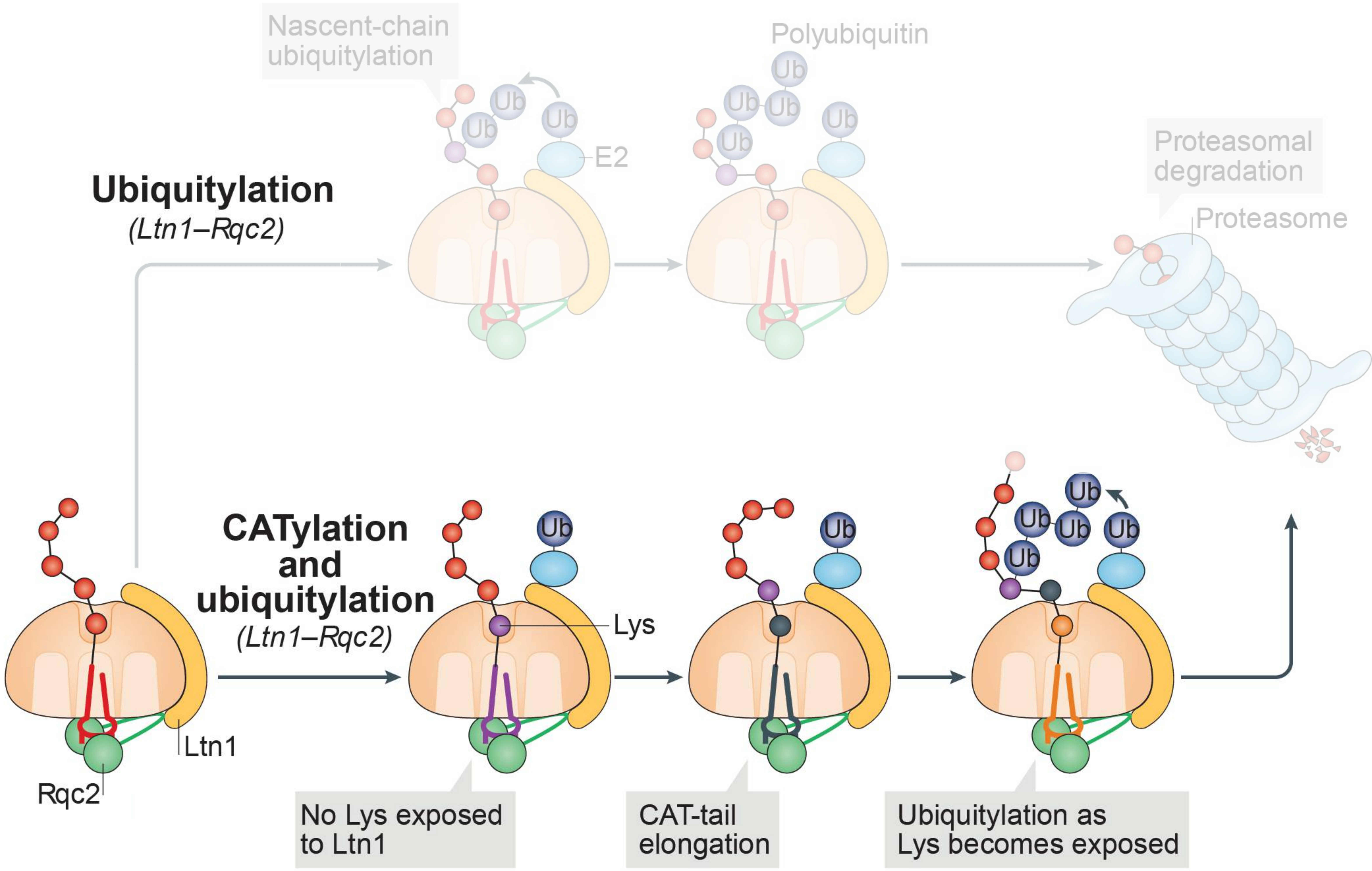
Ribosome Quality control



Ubiquitylation
(Ltn1–Rqc2)

CATylation and ubiquitylation
(Ltn1–Rqc2)

Ribosome Quality control



Ubiquitylation
(Ltn1–Rqc2)

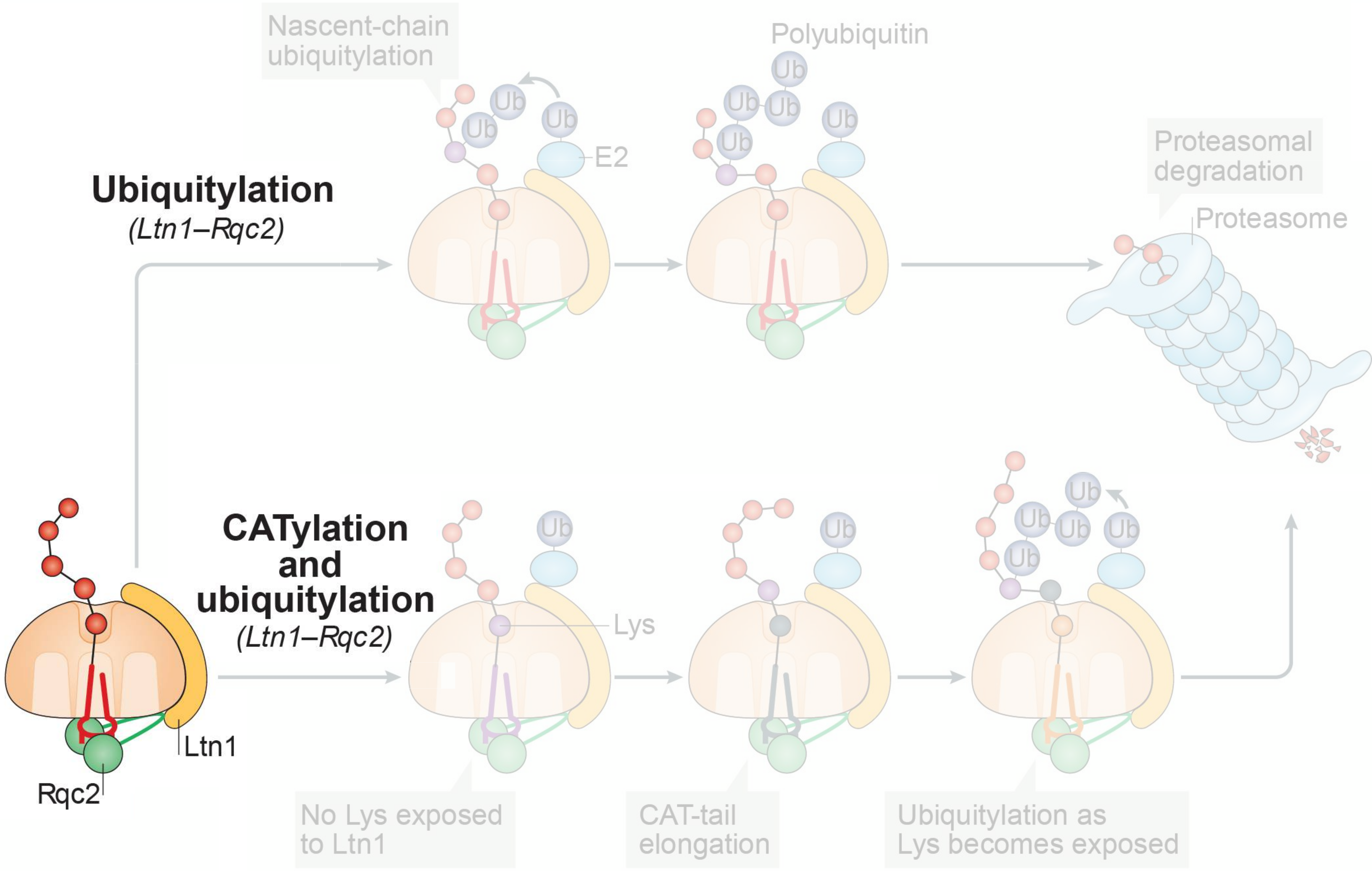
CATylation and ubiquitylation
(Ltn1–Rqc2)

No Lys exposed
to Ltn1

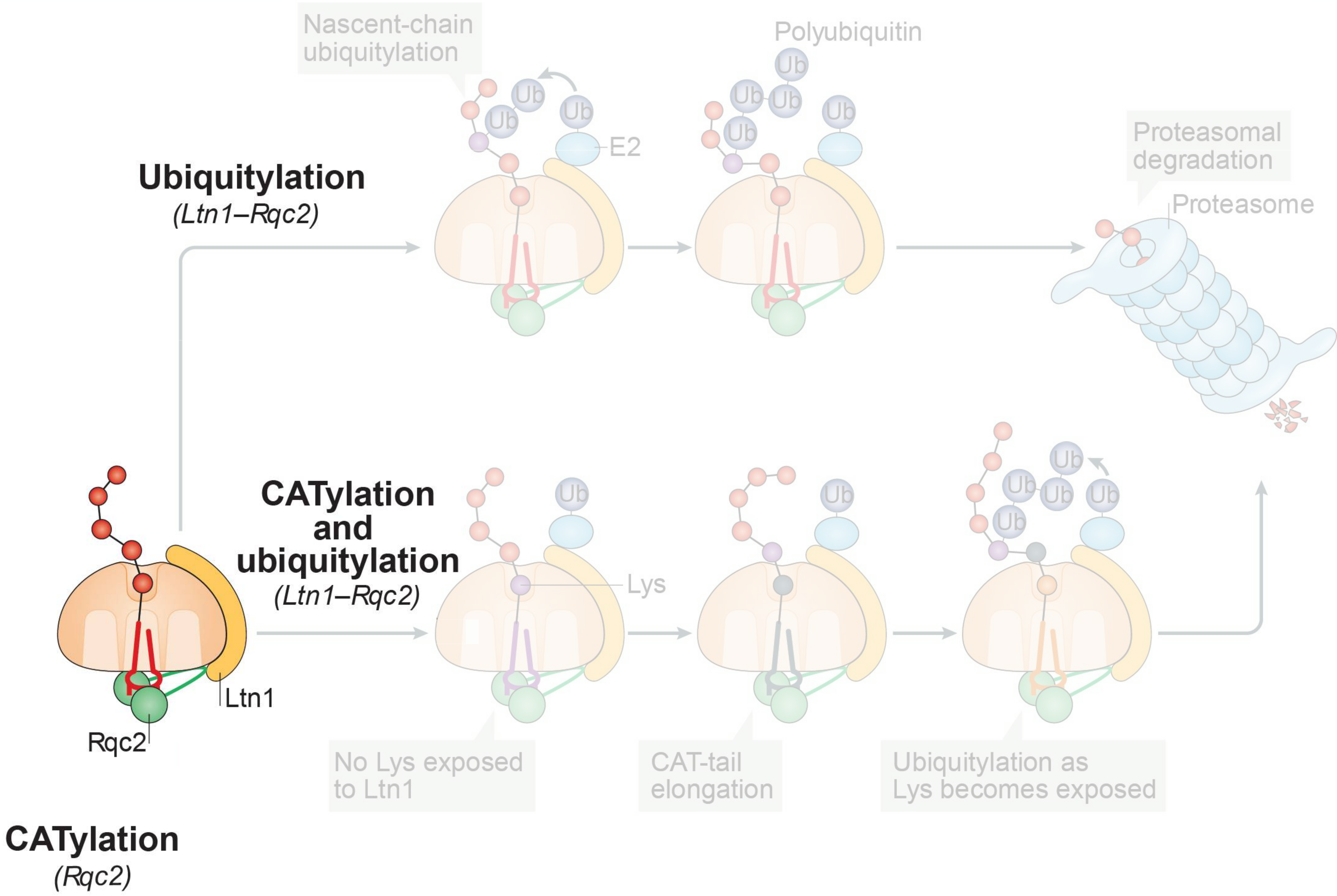
CAT-tail
elongation

Ubiquitylation as
Lys becomes exposed

Ribosome Quality control

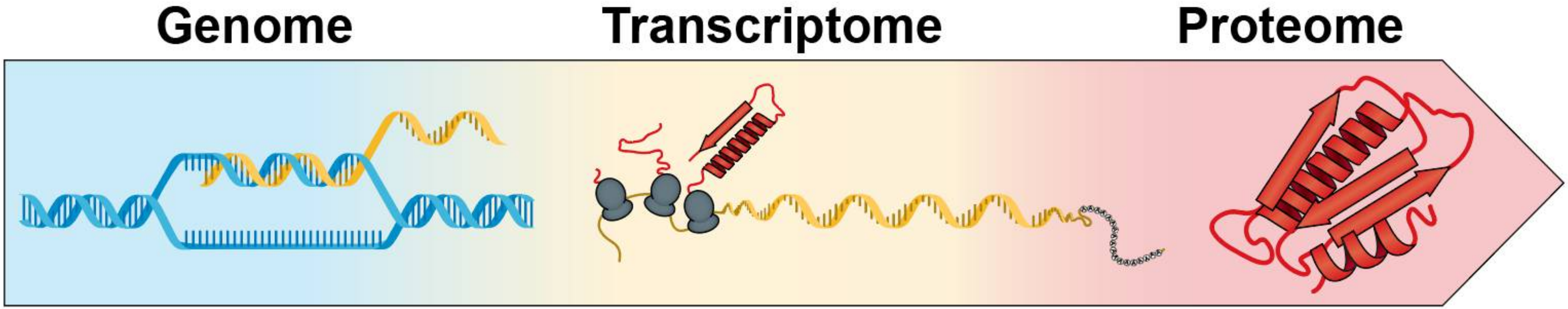


Ribosome Quality control



Ribosome Quality control

Relationship between protein levels and mRNA abundance?



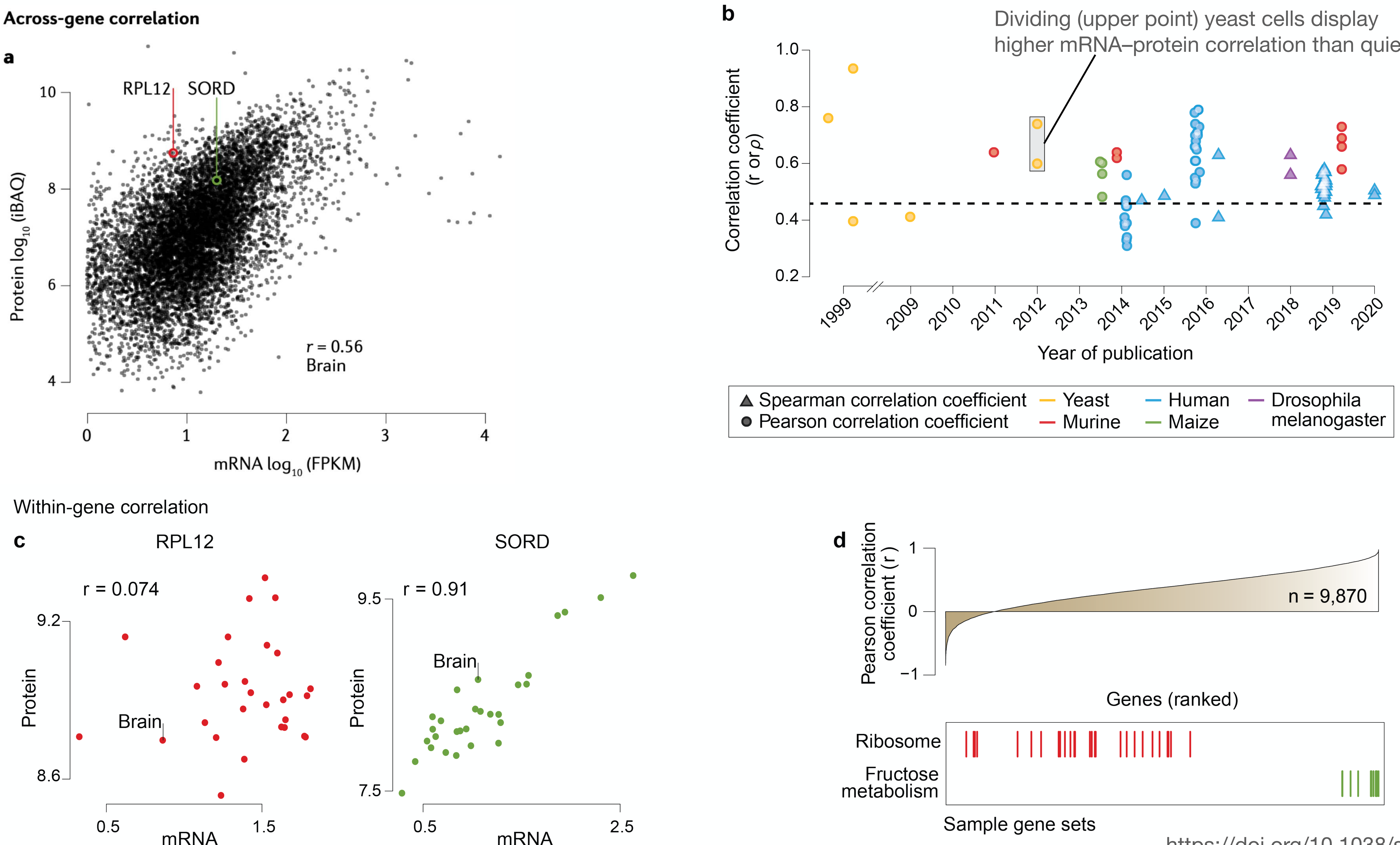
The central dogma of biology tightly links the molecular species DNA, RNA, and protein.

Can we predict protein levels from mRNAs levels?

Protein abundance depends on four factors:
A) transcription rates
B) mRNA half-lives
C) translation rate constants
D) protein half-lives

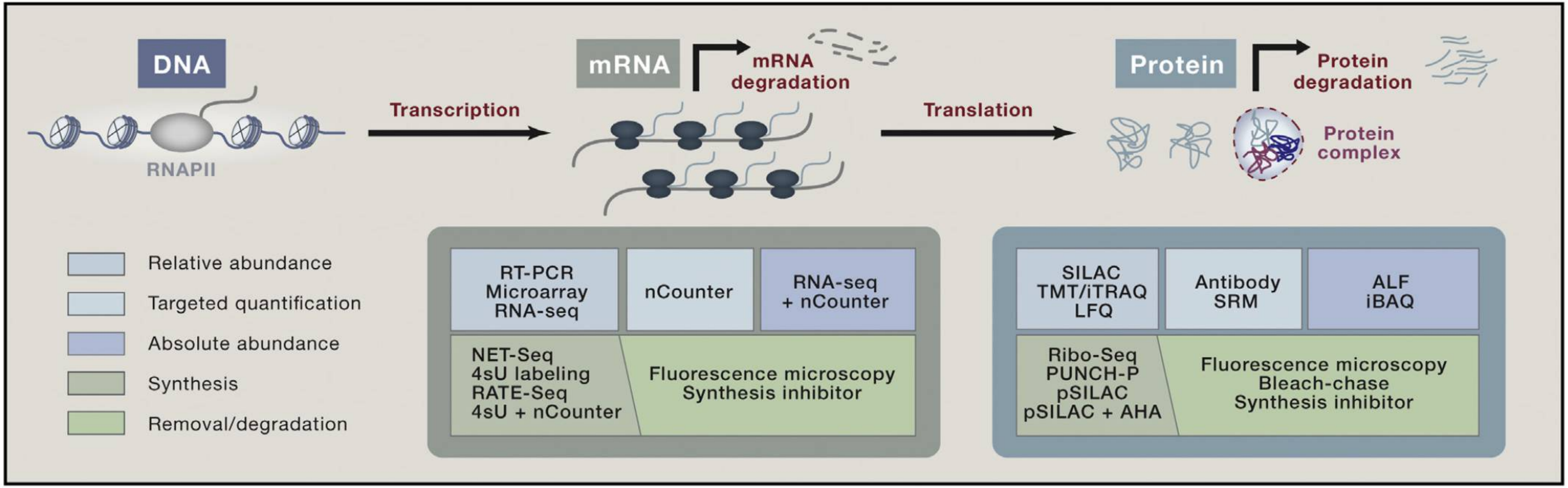
Whereas the nucleotide sequence of a gene determines the sequence of its mRNA product, and whereas an mRNA's sequence determines the amino acid sequence of the resulting polypeptide, **there is *no trivial relationship between the concentration of a transcript and the concentration(s) of the protein(s)*** derived from a particular locus.

Relationship between protein levels and mRNA abundance?



https://doi.org/10.1038/s41576-020-0258-4R

Relationship between protein levels and mRNA abundance?



So when performing Seq experiments and data analysis always keep in mind that quite often this may not correlate with the protein level (end product).

Orthogonal validation at the protein is always a good idea before moving forward with hypothesis.

Relationship between protein levels and mRNA abundance?

Both mRNA-level (transcriptomic) and protein-level (proteomics) measurements provide unique insights into biological systems.

mRNA levels should not be interpreted as the final output of gene expression. Instead, it is more instructive to **think of mRNAs as** what they mechanistically are: **the templates for protein synthesis**.

The presence of mRNAs is required for protein synthesis. However, due to differential translation, protein degradation, contextual confounders and pervasive protein-level buffering, this does not imply that proteins are actually being made or are present in proportional quantities.

Additional resources to explore the subject

Detection and Degradation of Stalled Nascent Chains via Ribosome-Associated Quality Control
https://doi.org/10.1146/annurev-biochem-013118110729

Roadblocks and resolutions in eukaryotic translation
https://doi.org/10.1038/s41580-018-0011-4

The Organizing Principles of Eukaryotic Ribosome Recruitment
https://doi.org/10.1146/annurev-biochem-013118111042

Mechanisms and functions of ribosome-associated protein quality control
https://doi.org/10.1038/s41580-019-0118-2

mRNAs, proteins and the emerging principles of gene expression control
https://doi.org/10.1038/s41576-020-0258-4R