BI 8 LECTURE 6

TRANSLATION: MAKING PROTEIN FROM RNA TEMPLATE

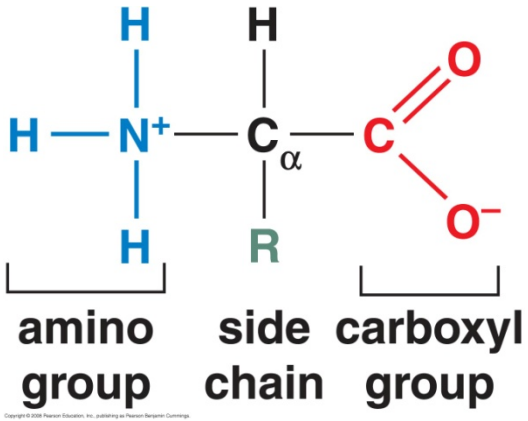
& PROTEIN STRUCTURE INTRO

Ellen Rothenberg

21 January 2016

Chapter 6: pp. 333-366; & begin Chapter 3

Translating from one code to another
Protein basics: subunits are amino acids



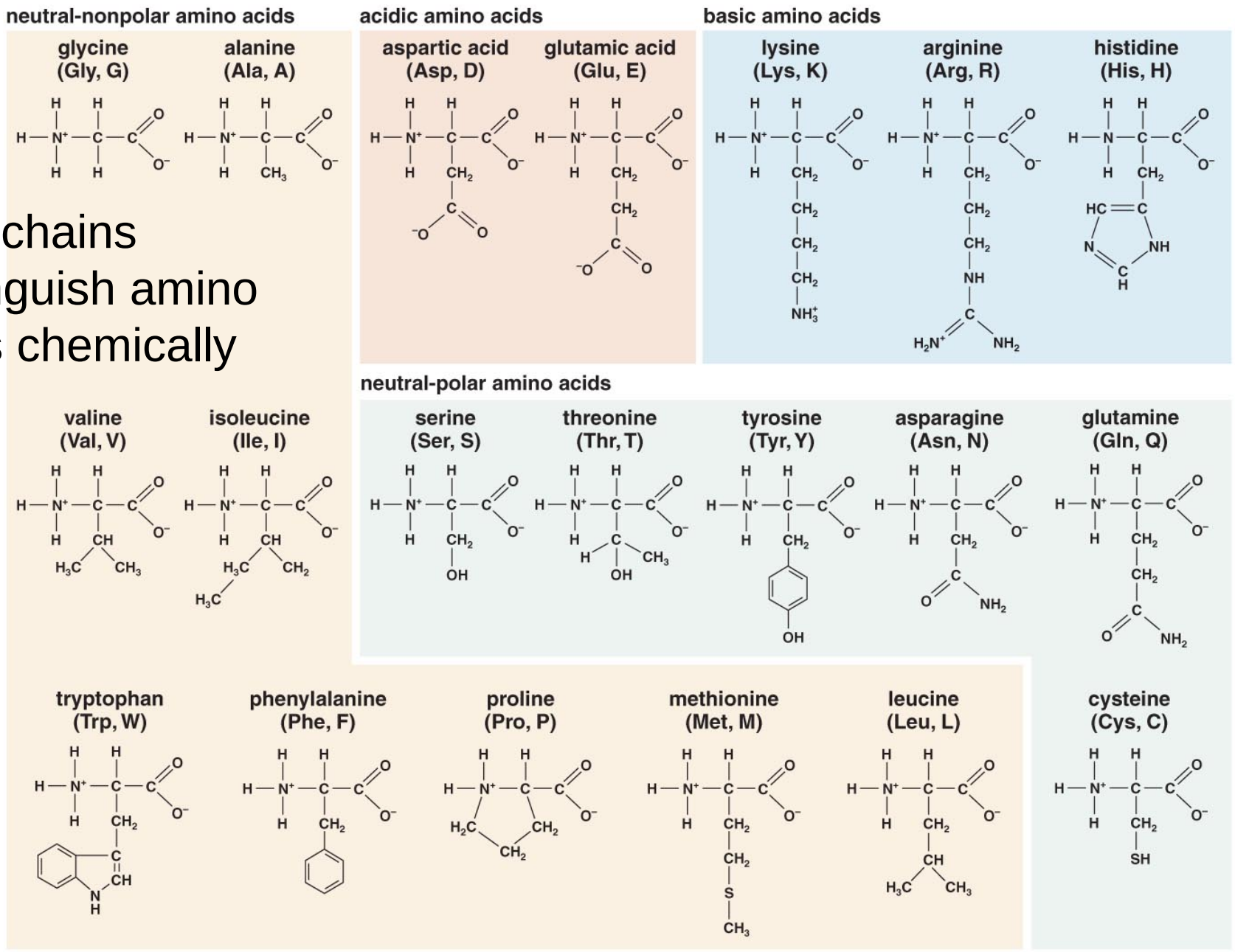
Unlike –NH2 in nucleotides, these *free* amino groups and free carboxyl groups are ionized at neutral pH

Coding is a qualitative transformation, nucleotide triplets translated to amino acids

No structural homology

Only polarity of polymerization is similar: what happens to the end of the chain depends on what comes before it

Side chains distinguish amino acids chemically



Formation of amide bonds between carboxyl groups of one amino acid and amino group of next amino acid polymerizes protein chain… "N" to "C"

Progressive, vectorial polymerization

Read from RNA template, 5' to 3'

N terminal amino acid is first, keeps free NH_3^+

Other residues added sequentially to its carboxy end

Last amino acid added: "C terminal"

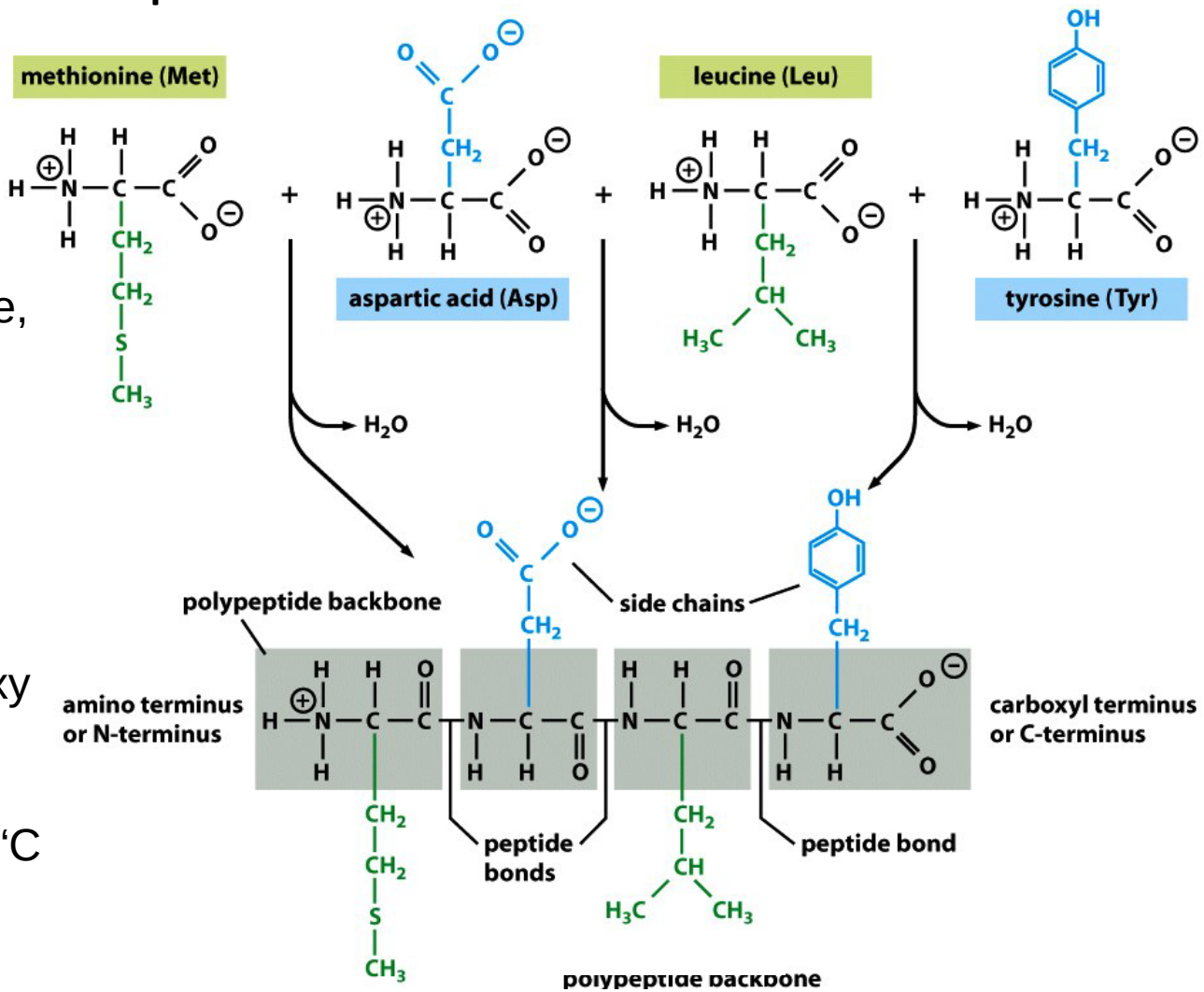


Figure 3-1 *Molecular Biology of the Cell* (© Garland Science 2008)

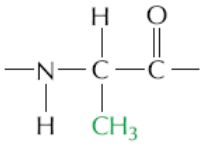
Nonpolar amino acids have "oily" $-CH_3$ or $-CH_2-$ or benzene ring groups in R chains

The longer the hydrocarbon stretch, the more hydrophobic

NONPOLAR SIDE CHAINS

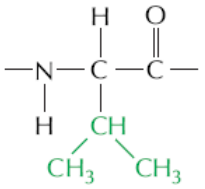
alanine

(Ala, or A)



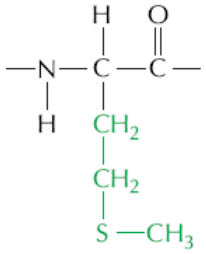
valine

(Val, or V)



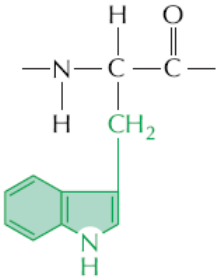
methionine

(Met, or M)



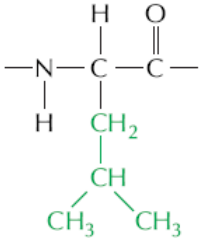
tryptophan

(Trp, or W)



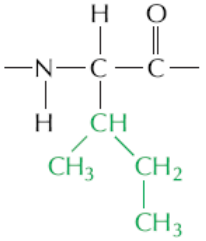
leucine

(Leu, or L)



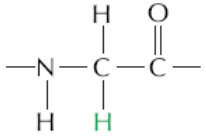
isoleucine

(Ile, or I)



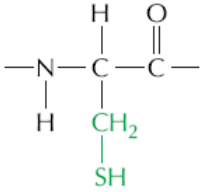
glycine

(Gly, or G)



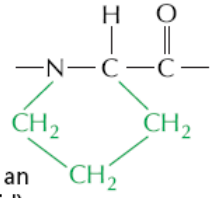
cysteine

(Cys, or C)



proline

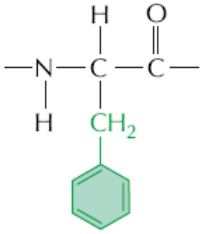
(Pro, or P)



(actually an imino acid)

phenylalanine

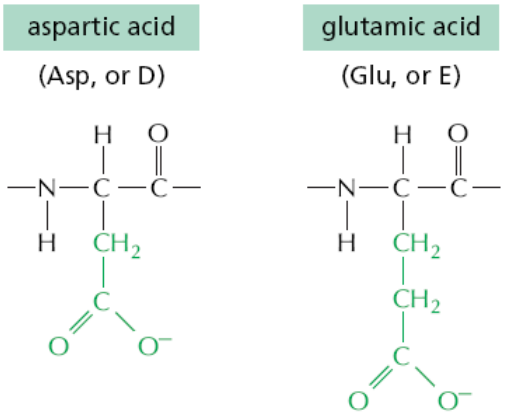
(Phe, or F)



Disulfide bonds can form between two cysteine side chains in proteins.

Highly polar amino acid side chains are charged at neutral pH

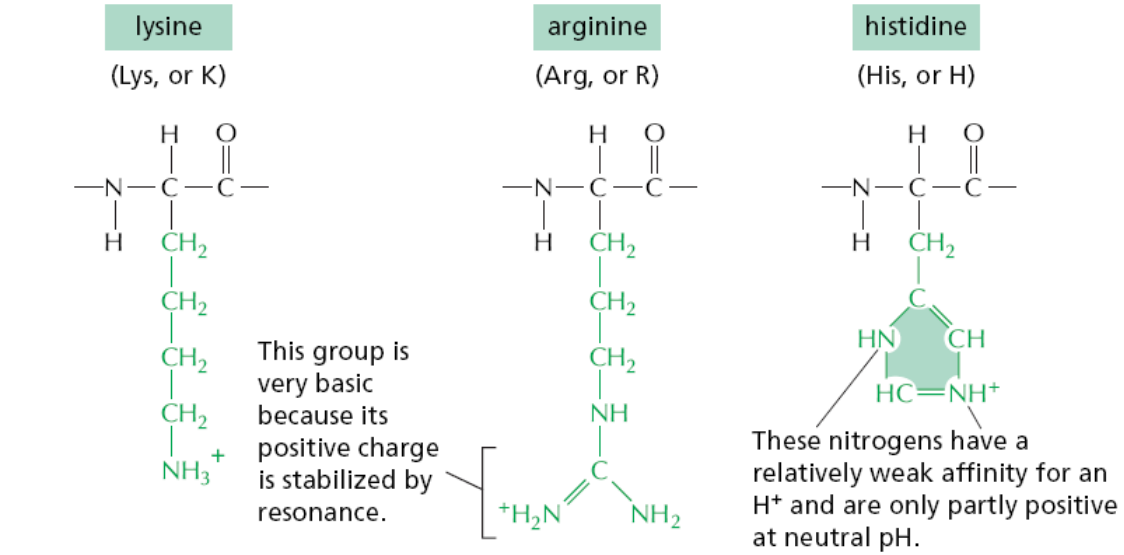
ACIDIC SIDE CHAINS



Acidic residues have given up their hydrogen protons to the water medium… so they have a negative charge left on them

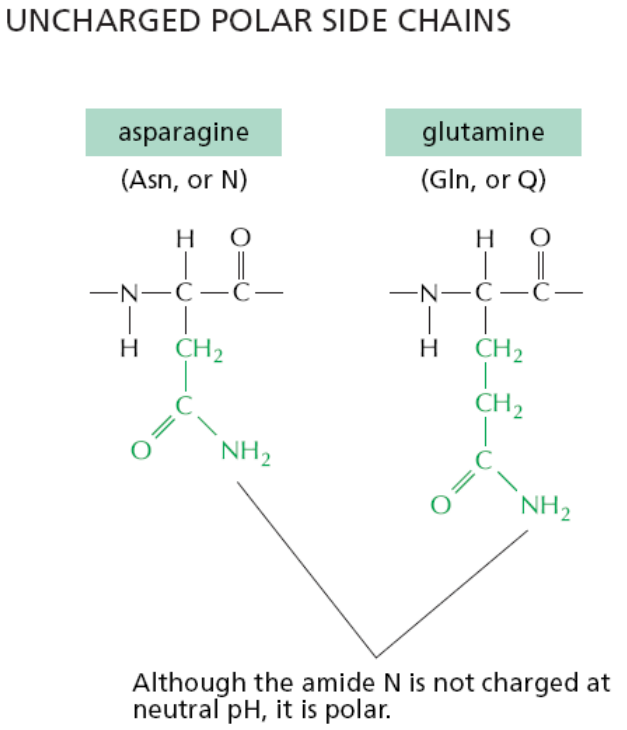
Basic residues have captured hydrogen protons from the water medium… so they have a positive charge on them

BASIC SIDE CHAINS

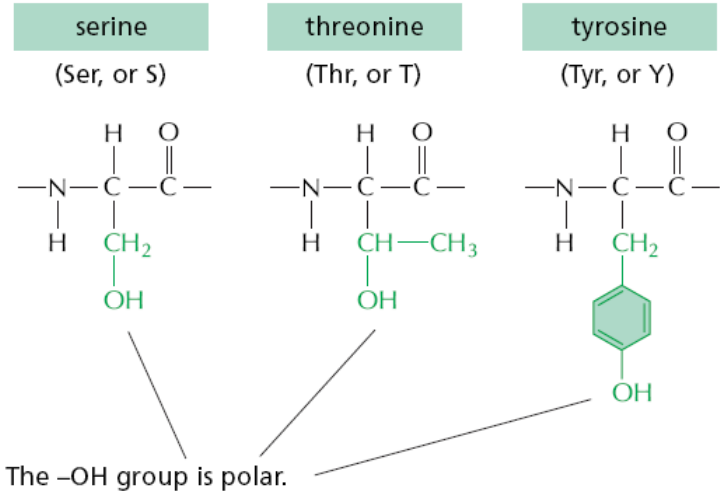


SPECIAL CASE: His

Uncharged but polar: “get along with everyone…” but actually very important for H-bond formation and regulation by post-translational modification



These amino acids can be modified by adding highly ***negatively charged*** phosphate groups at their terminal –OH hydroxyls… frequent response to signaling pathways: radically & ***reversibly*** changes their own chemical activity spectrum



SPECIAL AMINO ACIDS

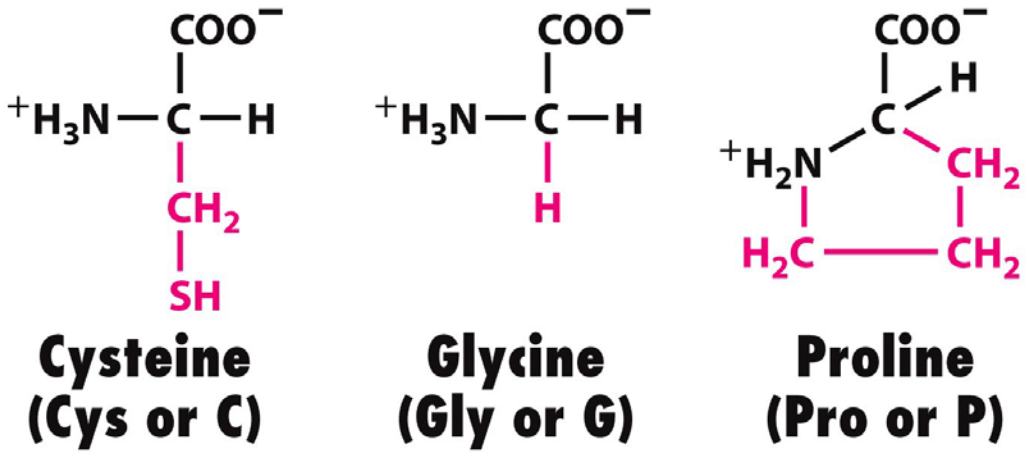
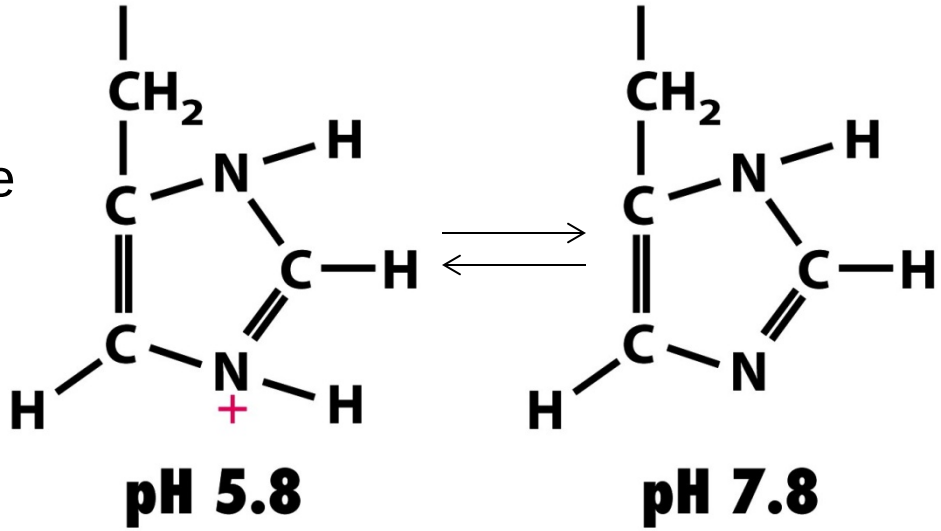


Figure 2-14 part 3
Molecular Cell Biology, Sixth Edition

Histidine

(R chain only)



Un 2-2
Molecular Cell Biology, Sixth Edition

Cys: for redox-controlled covalent crosslinks

Gly: max flexibility

Pro: rigid locked phi bond→helix breaker

His: champion of proton exchange in physiological pH's

A major contribution to inter- and intradomain linkages comes from Cys-Cys disulfide bridges

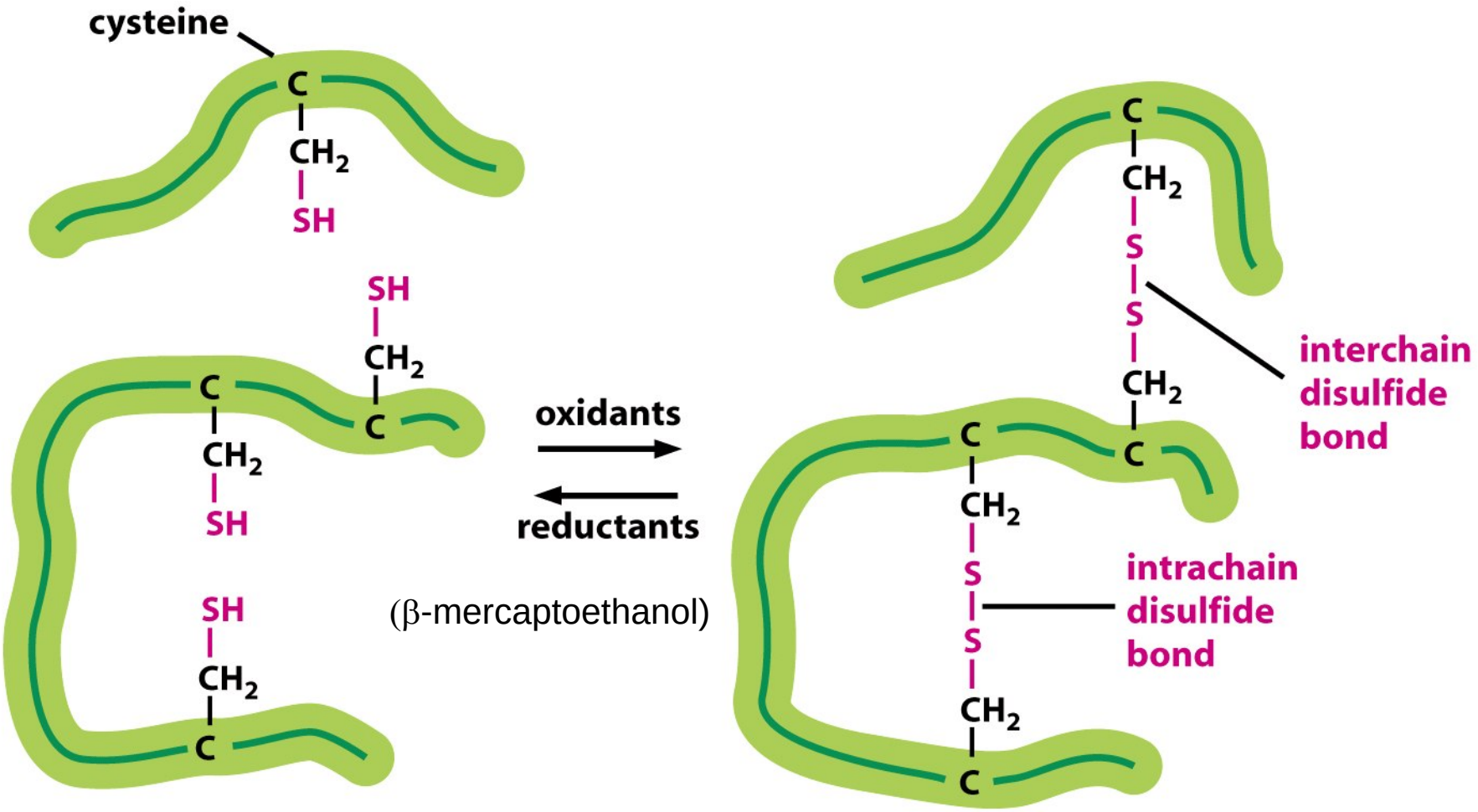


Figure 3-28 *Molecular Biology of the Cell* (© Garland Science 2008)

Crucial: get to know the amino acids

AMINO ACID			SIDE CHAIN
Aspartic acid	Asp	D	negative
Glutamic acid	Glu	E	negative
Arginine	Arg	R	positive
Lysine	Lys	K	positive
Histidine	His	H	positive
Asparagine	Asn	N	uncharged polar
Glutamine	Gln	Q	uncharged polar
Serine	Ser	S	uncharged polar
Threonine	Thr	T	uncharged polar
Tyrosine	Tyr	Y	uncharged polar

POLAR AMINO ACIDS

AMINO ACID			SIDE CHAIN
Alanine	Ala	A	nonpolar
Glycine	Gly	G	nonpolar
Valine	Val	V	nonpolar
Leucine	Leu	L	nonpolar
Isoleucine	Ile	I	nonpolar
Proline	Pro	P	nonpolar
Phenylalanine	Phe	F	nonpolar
Methionine	Met	M	nonpolar
Tryptophan	Trp	W	nonpolar
Cysteine	Cys	C	nonpolar

NONPOLAR AMINO ACIDS

At least in terms of general properties

Figure 3-2 *Molecular Biology of the Cell* (© Garland Science 2008)

THREE MATTERS… in protein coding regions

You must know where to start counting and stay in register

4^3 = 64 triplets
20 amino acids + stop signals

Unique codon for starting protein chains; also used for internal Met

GAUAU=?

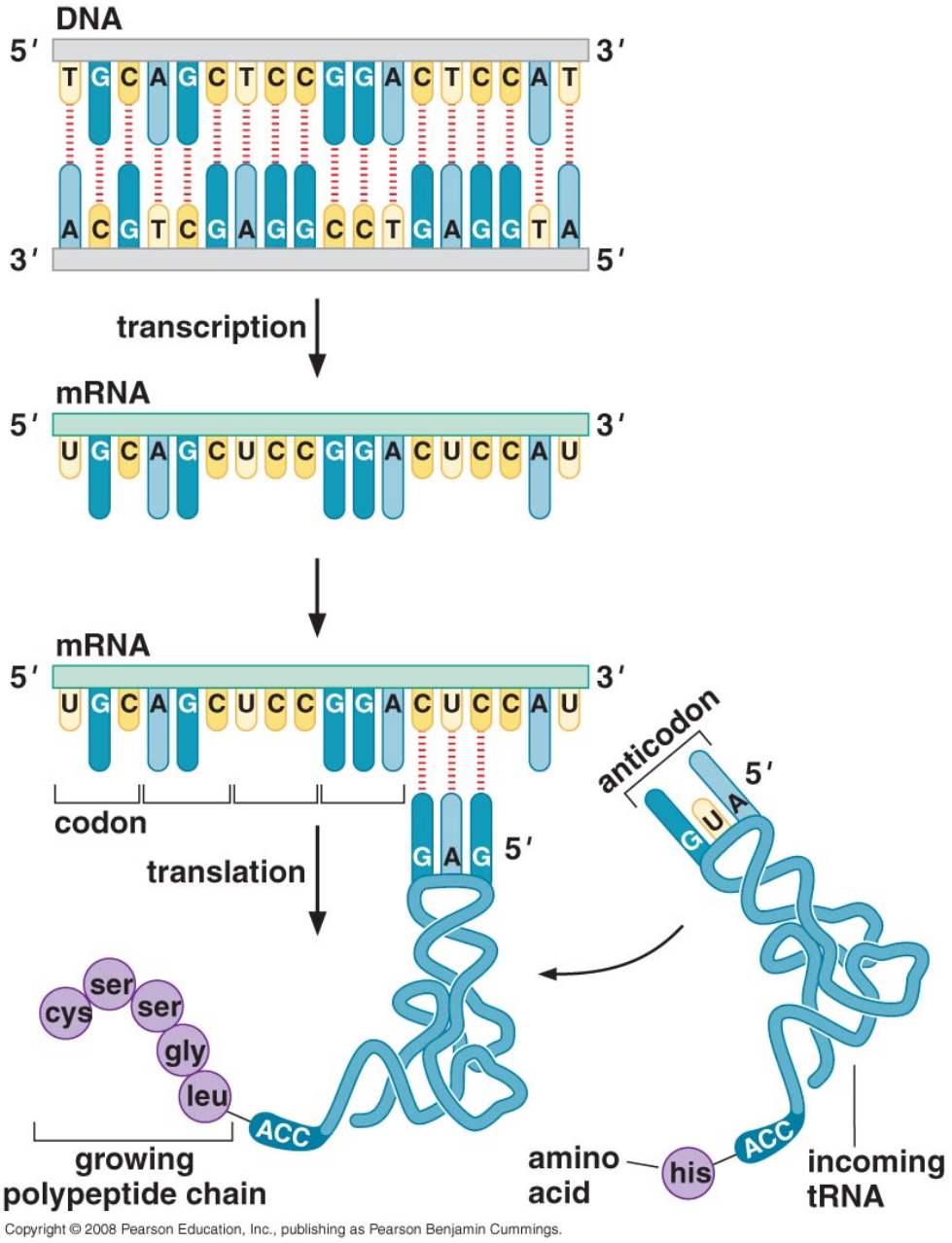
TABLE 15-1 The Genetic Code

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

* Chain-terminating or "nonsense" codons.

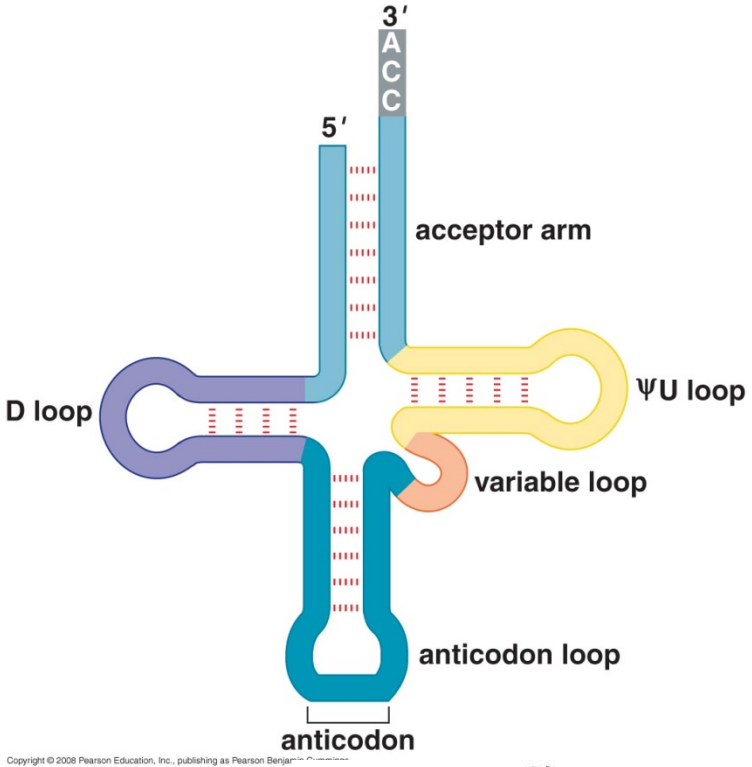
† Also used in bacteria to specify the initiator formyl-Met-tRNA^{fMet}.

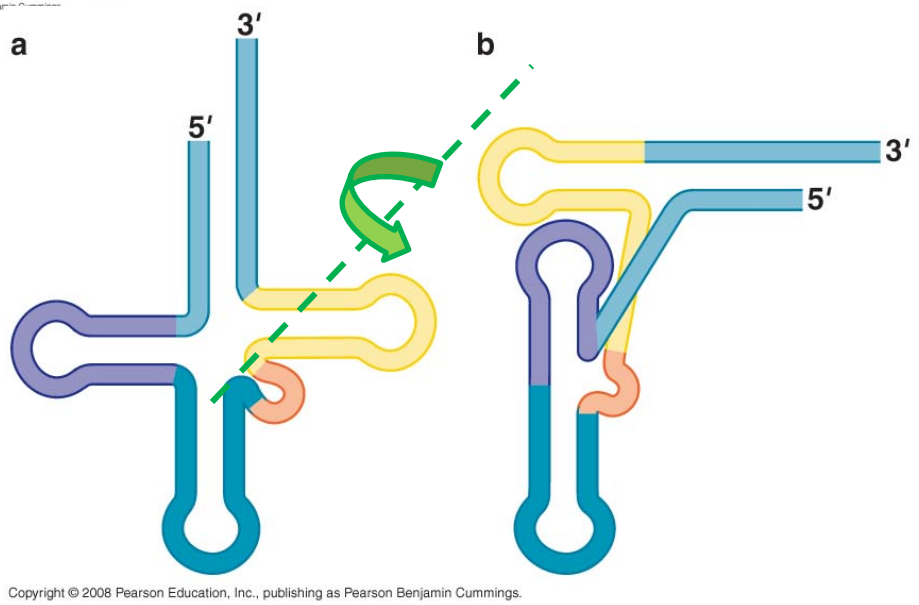
Converting mRNA sequence to protein sequence requires triplet decoding adaptor molecules: tRNA

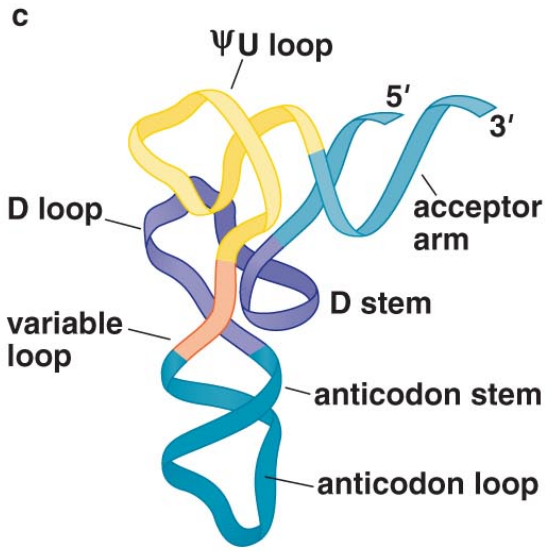


tRNA

a highly modified, folded small RNA (~75 nt) is the "translation adaptor" to interpret genetic code

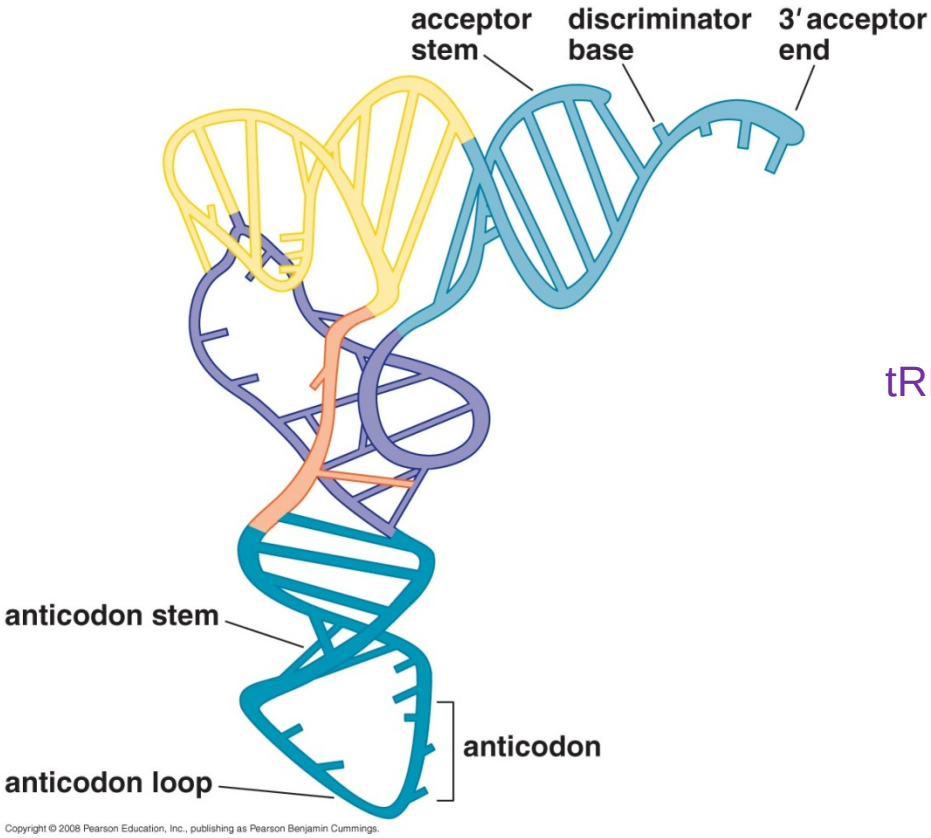


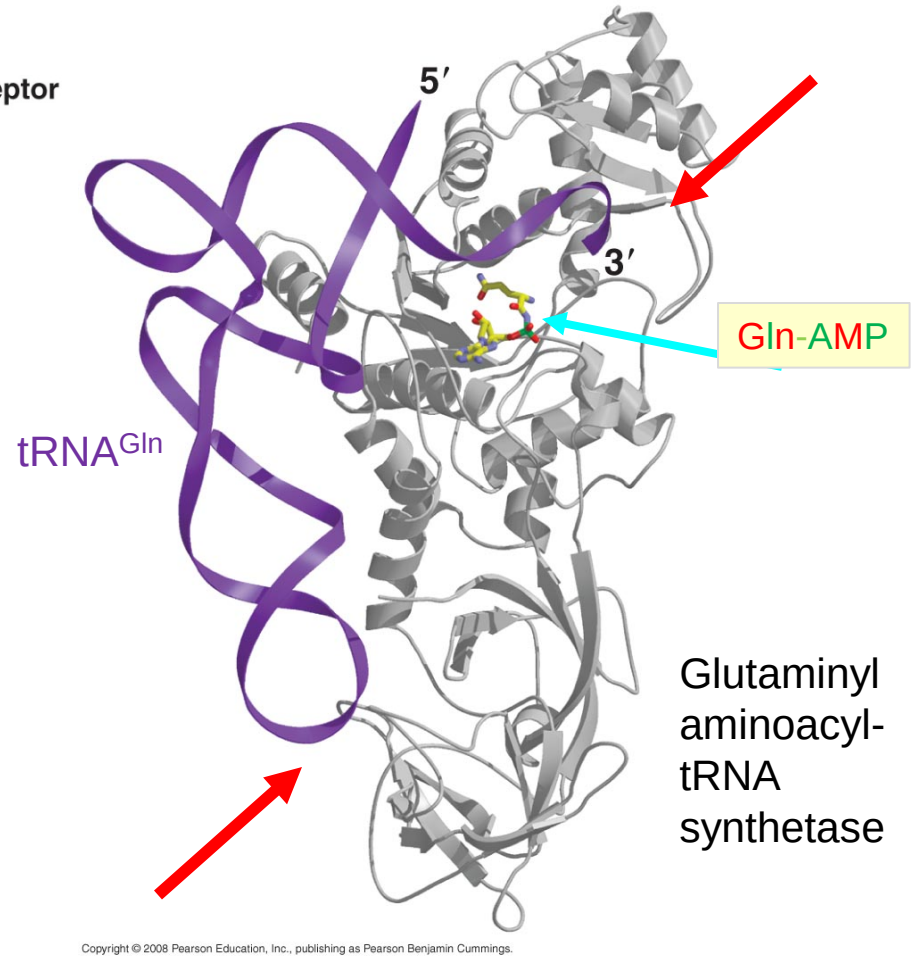




Aminoacyl-tRNA synthetases: enzymes that interpret specificity of genetic code

recognize anticodon *and* acceptor ends of tRNA and covalently link the right amino acid





Glutaminyl aminoacyl-tRNA synthetase

The problem of coding specificity

4^3 = 64 triplets

20 amino acids
+ stop signals

Unique codon for starting protein chains; also used for internal Met

TABLE 15-1 The 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

* Chain-terminating or "nonsense" codons.

† Also used in bacteria to specify the initiator formyl-Met-tRNA^{fMet}.

Degeneracy in the genetic code

- There are only 20 aminoacyl-tRNA synthetases, each dedicated for one of the 20 amino acids: here is the specificity

- Several aminoacyl-tRNA synthetases can recognize two or more different tRNAs as substrates for transfer of “their” amino acid

- Charging of tRNAs is entirely quality controlled by the aminoacyl-tRNA synthetase…
  - Ribosome quality-controls only match tRNA “anticodon” triplet to mRNA “codon” triplet
  - Ribosome ***cannot*** tell wrong amino acid from right one if it is charged on the tRNA that is right for the mRNA codon

Degeneracy in the genetic code

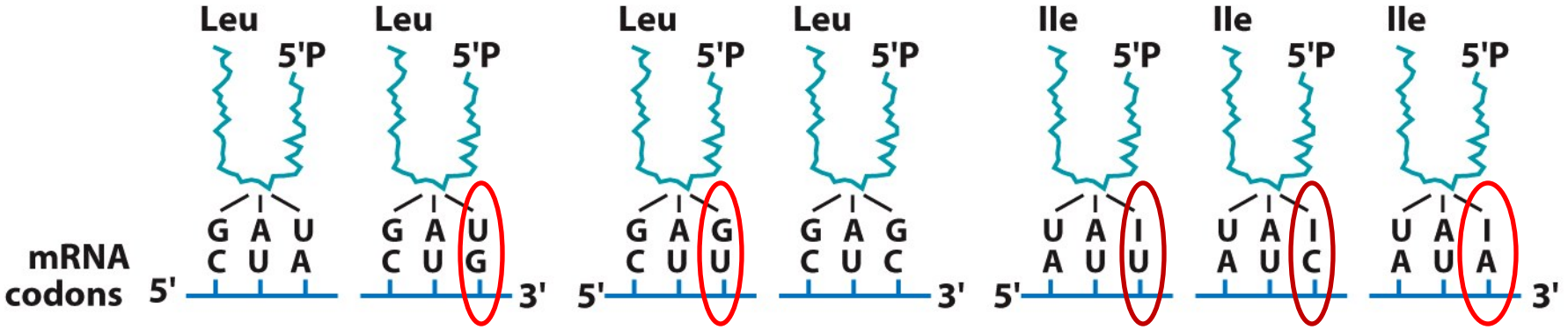
- There are many tRNAs
  - More than number of amino acids
  - Similar but not equal to number of mRNA codons

- Some tRNAs themselves have structural feature that allows them to read more than one sequence in mRNA codon

- Use of modified purine base INOSINE in anticodon

- Post-transcriptional modifications of tRNA include conversion of adenine to inosine… this is key for full decoding of mRNA

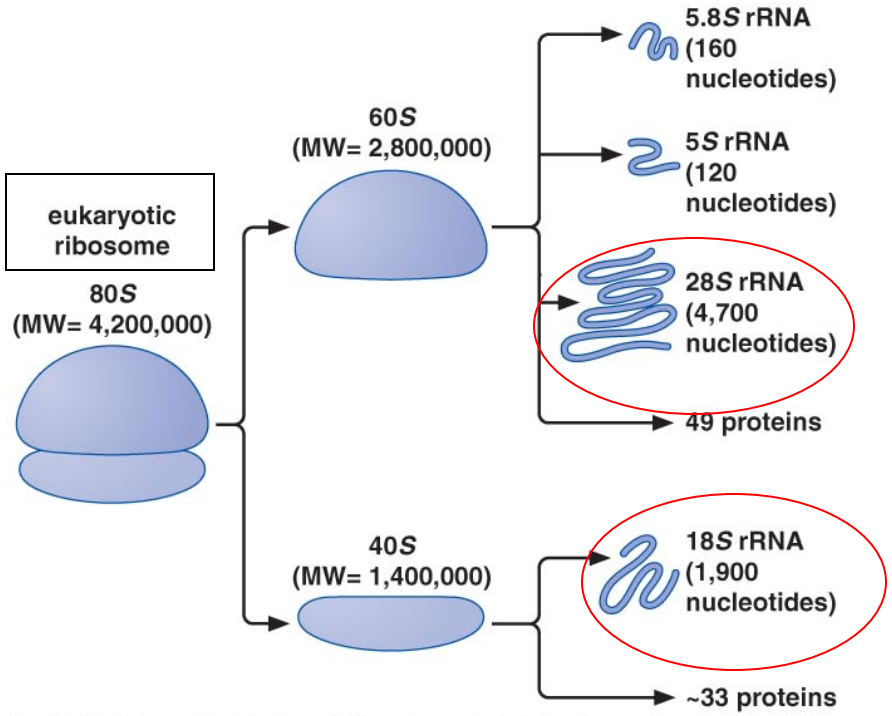
The wobble in the interaction between codon 3rd position and anticodon 1st position accounts for much of the degeneracy of genetic code

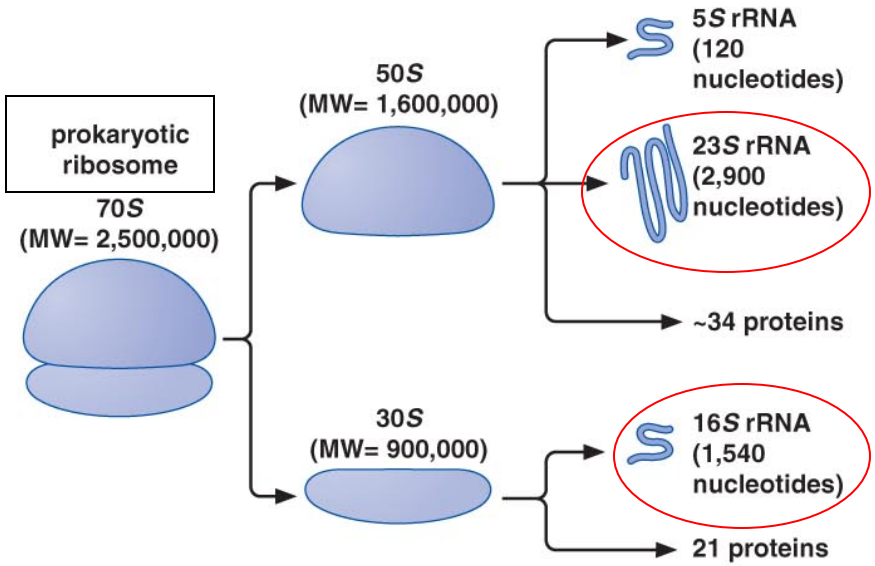


Corollary: 3rd position substitutions in protein-coding regions are under relatively weak evolutionary selection, and mutations affecting these bases can be tracked as evolutionary “clocks”

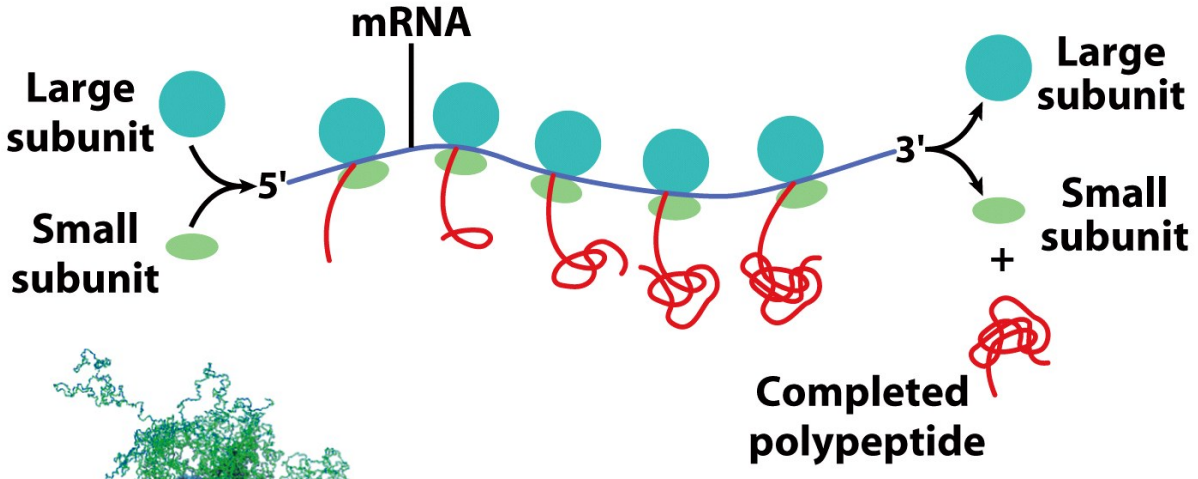
Ribosomes: massive nucleoprotein complexes that read mRNA sequence and build protein

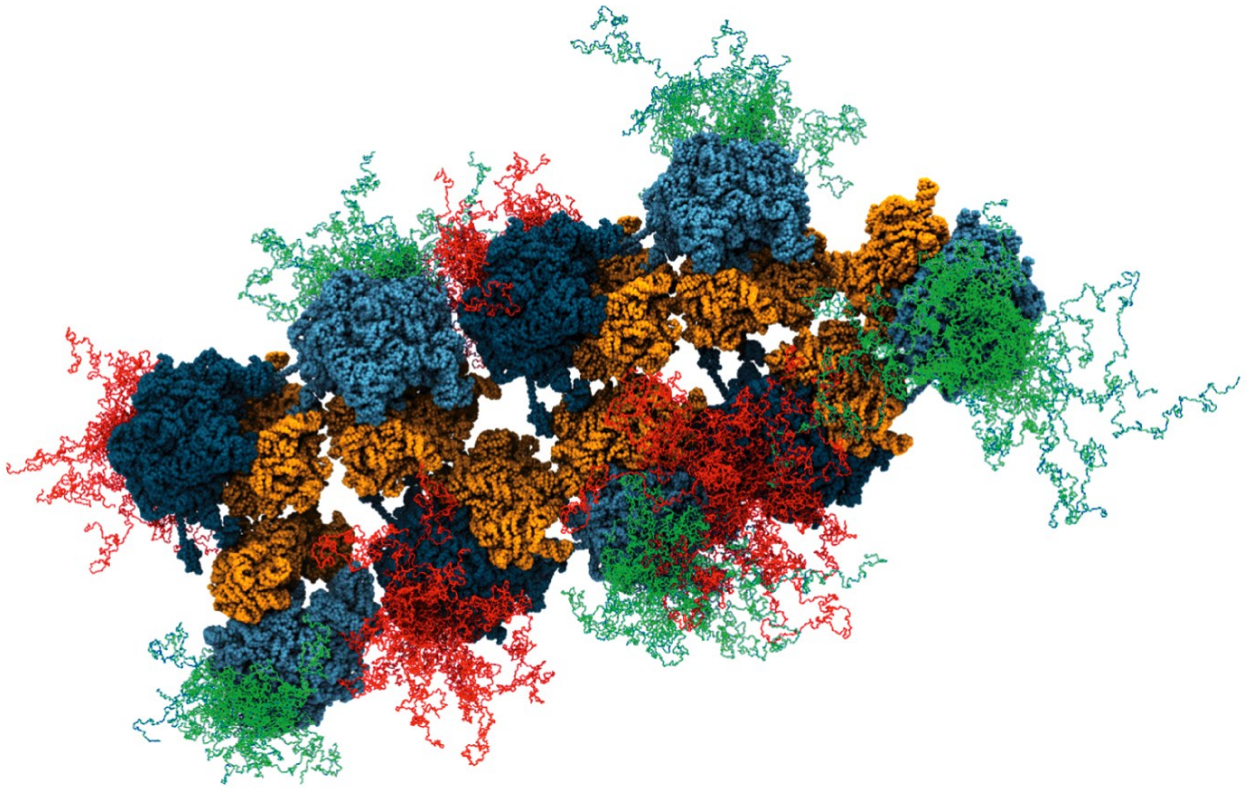
The two main subunit ribosomal RNAs constitute the great majority of total RNA mass expressed in any cell





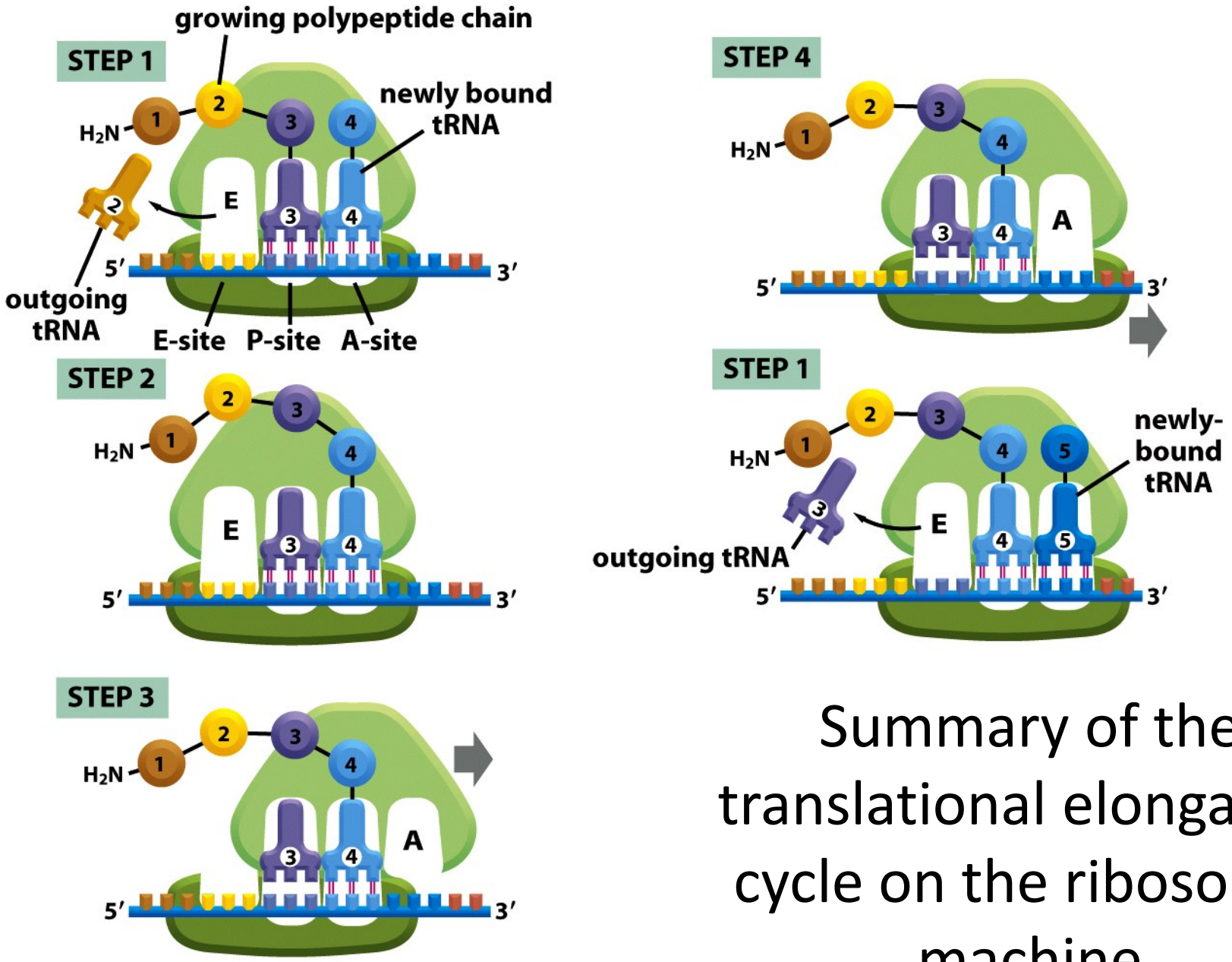
Repeated loading of the same mRNA with new translational initiations → polysomes (polyribosomes)





(Karp, 6th ed.)

From Florian Brandt et al., courtesy of Wolfgang Baumeister, *Cell* 136:267, 2009; by permission of Cell Press.



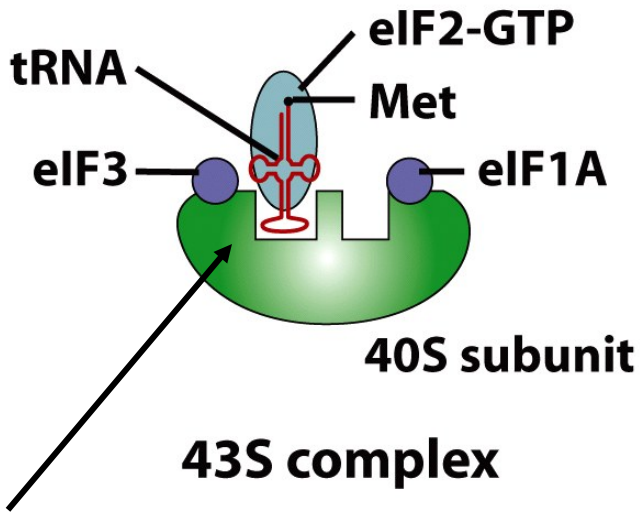
Summary of the translational elongation cycle on the ribosomal machine

Figure 6-66 *Molecular Biology of the Cell* (© Garland Science 2008)

Initiation of protein synthesis in eukaryotes: separate complexes to (1) bring initiator tRNA to the small ribosomal subunit; and to (2) validate mRNA quality

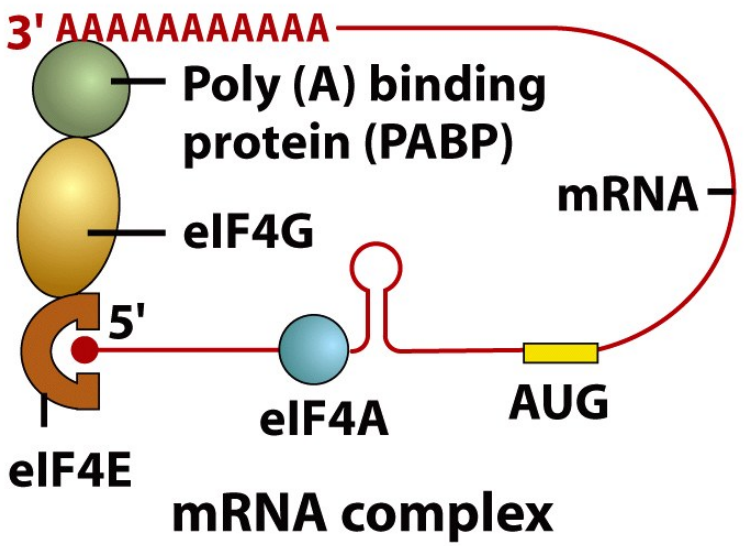
"eIF" = eukaryotic initiation factor (IF in prokaryotes)

Correct initiator tRNA



+

Poly(A) recognition

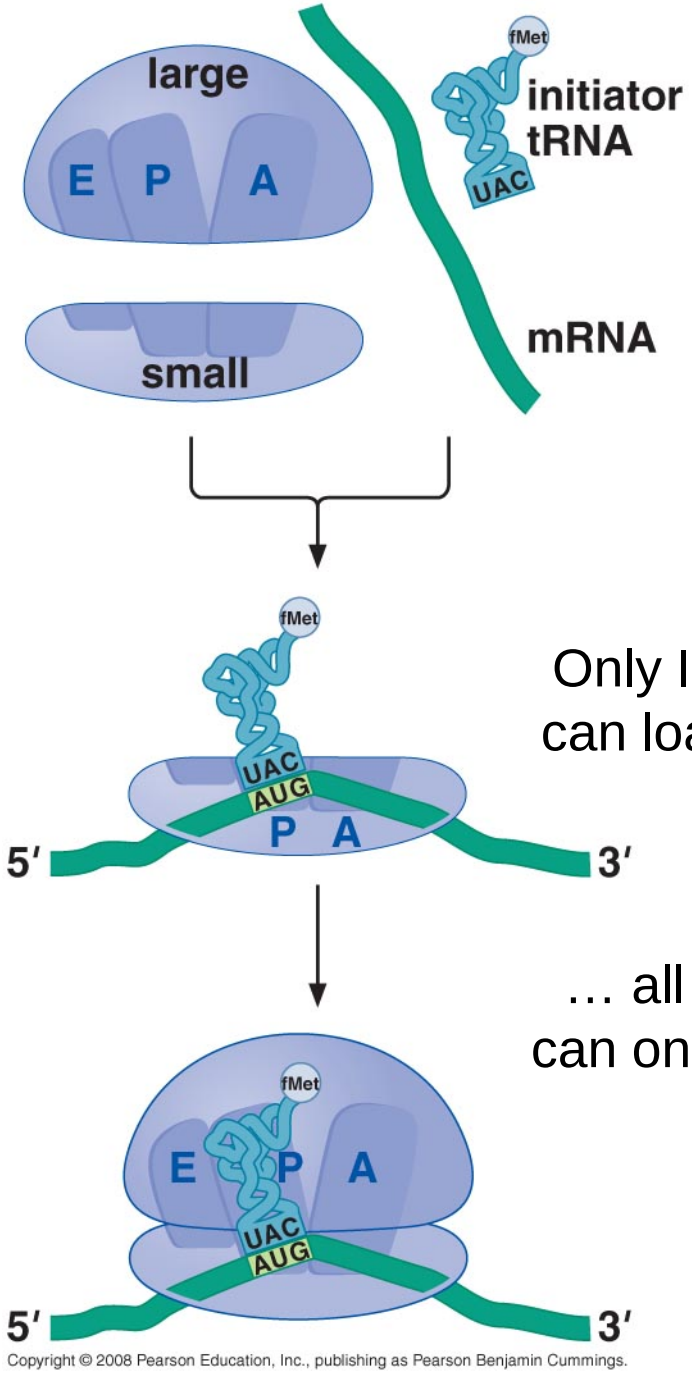


5'-Cap recognition

Note: tRNA is brought in ***first*** by a GTP-bound factor (eIF2)… cleavage of this GTP will be used soon to make AUG recognition irreversible

(Karp, 6th edition)

Starting translation: the ribosomal small subunit captures mRNA, special "initiator $tRNA^{Met}$", and makes complex



Three sites: E, P, A

Only Initiator tRNA can load into P site

… all other tRNAs can only load at "A"

Complex then recruits large subunit to begin adding & linking amino acids

Small ribosomal subunit scanning up & down mRNA takes energy ($ATP \rightarrow ADP + P_i$)

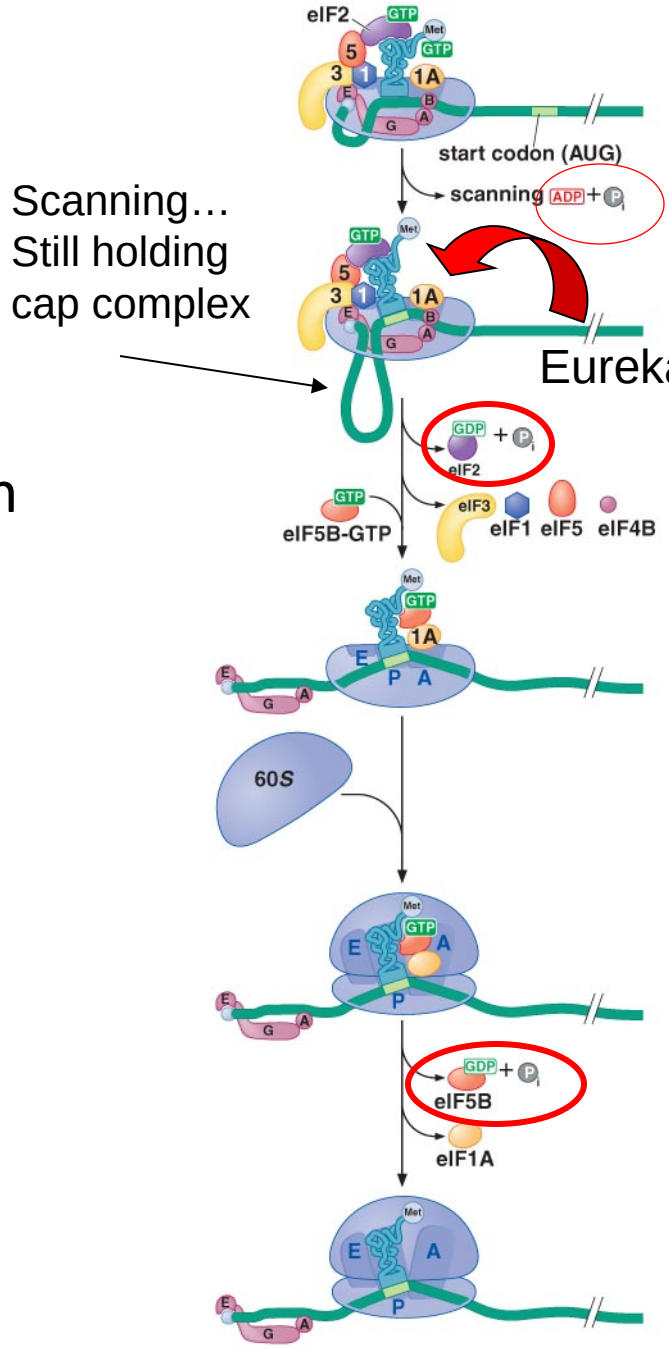
Bound $tRNA_i^{Met}$ does the recognition

Then: $2\ GTP \rightarrow 2\ GDP + 2\ P_i$ to lock down two initiation events:

(1) complex formation of $tRNA_i^{Met}$ with AUG

(2) new complex formation of initiation complex on 40S subunit with large 60S subunit

(prokaryotes combine these steps)

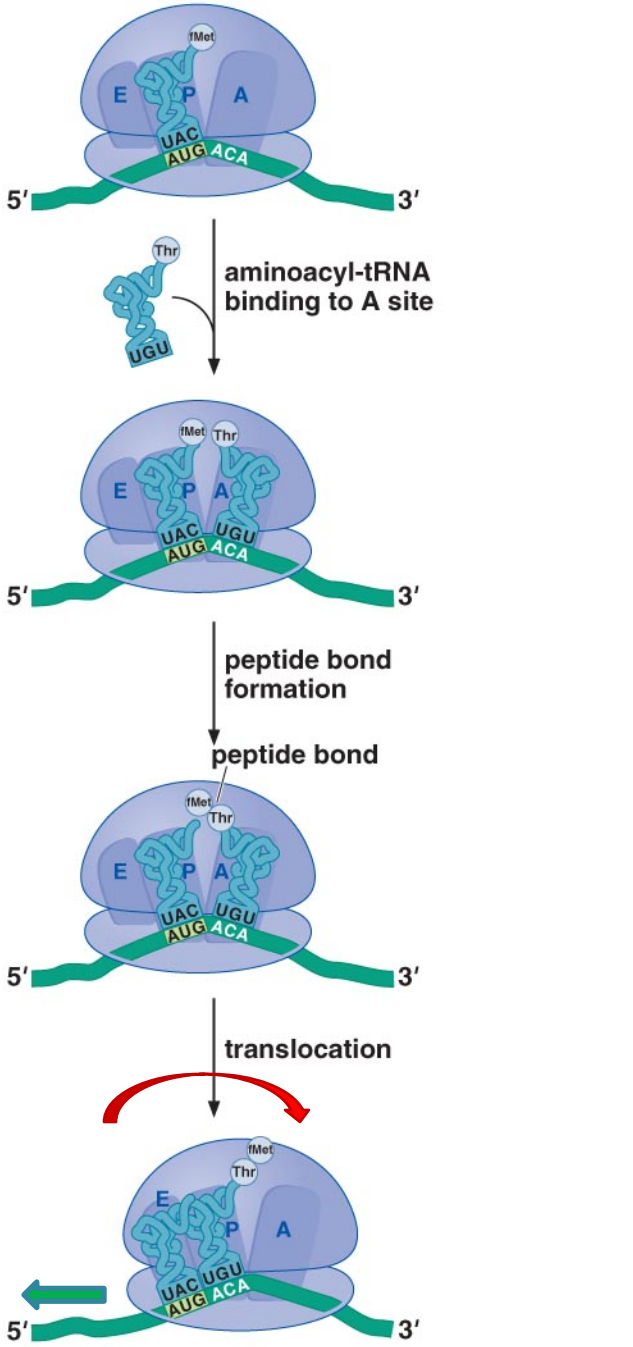


eIF5B: 2nd loading catalyst powered by GTP --> GDP+Pi

Elongation

After initiation… all subsequent aminoacyl-tRNAs have to load into "A" site, not "P" site, of fully formed ribosome complex

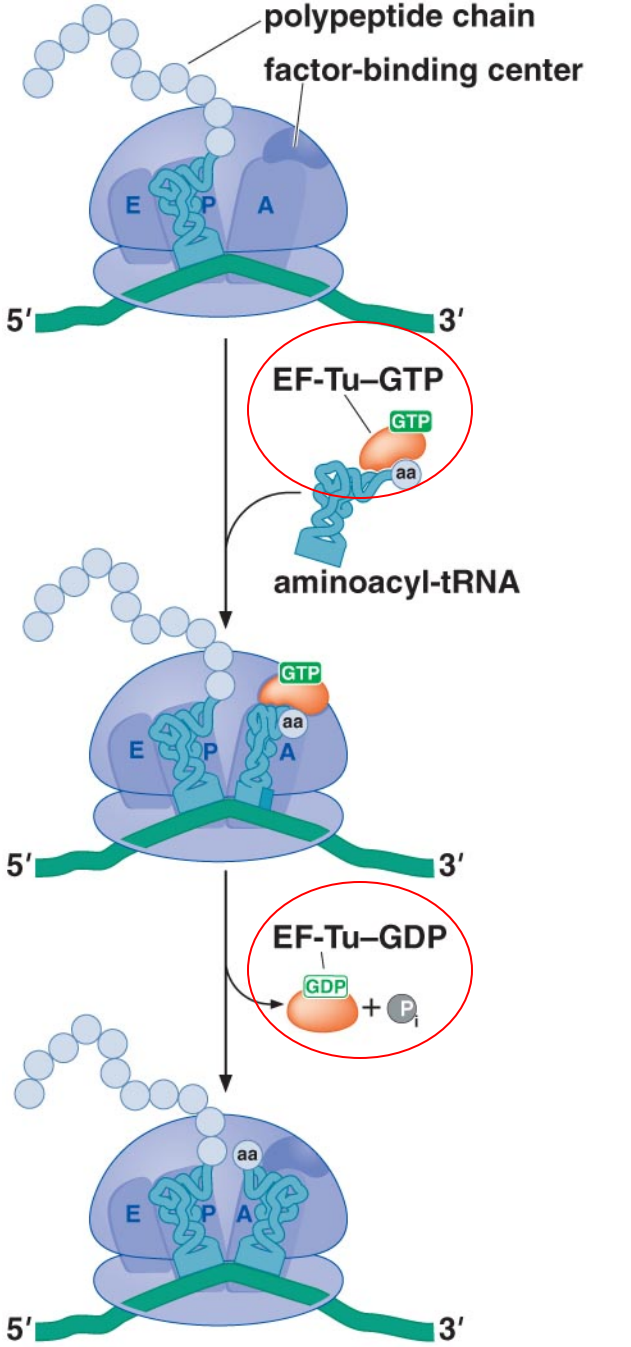
Once peptide bond is formed, whole unit translocates



For every new amino acid…

EF-Tu elongation factors bring in candidate aminoacyl tRNAs to audition for "fit" with mRNA at the A site

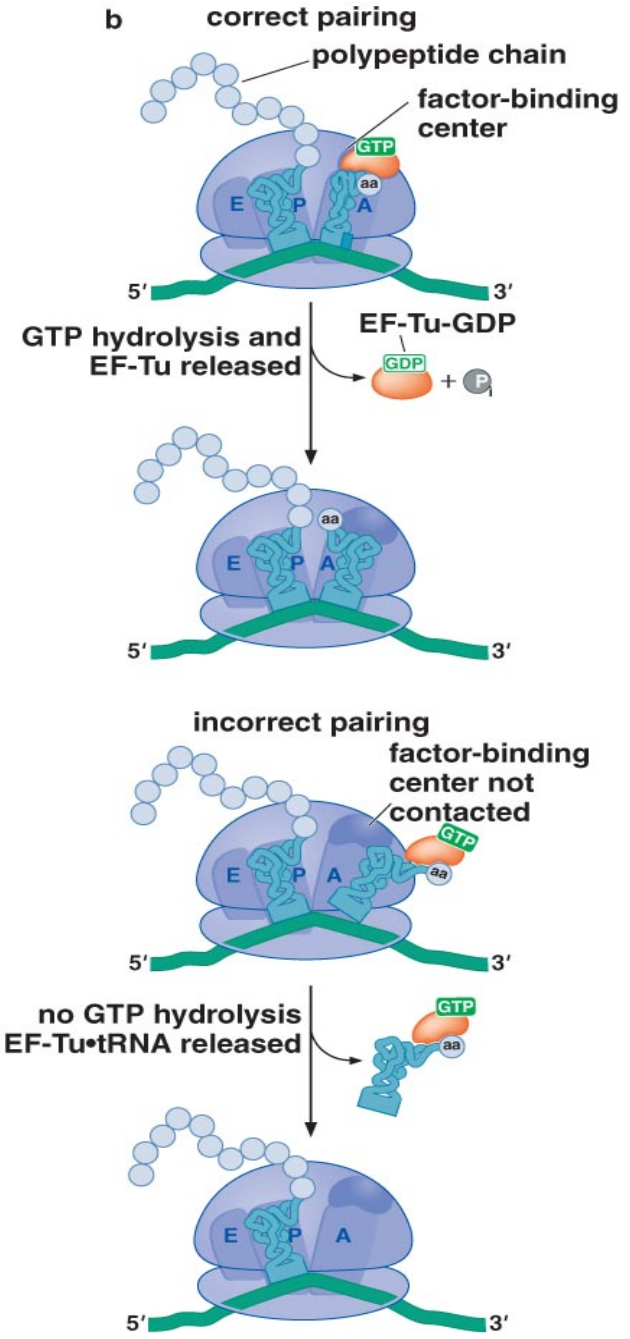
Again – each successful "fit" causes GTP hydrolysis… for aminoacyl-tRNA release and energetic lockdown

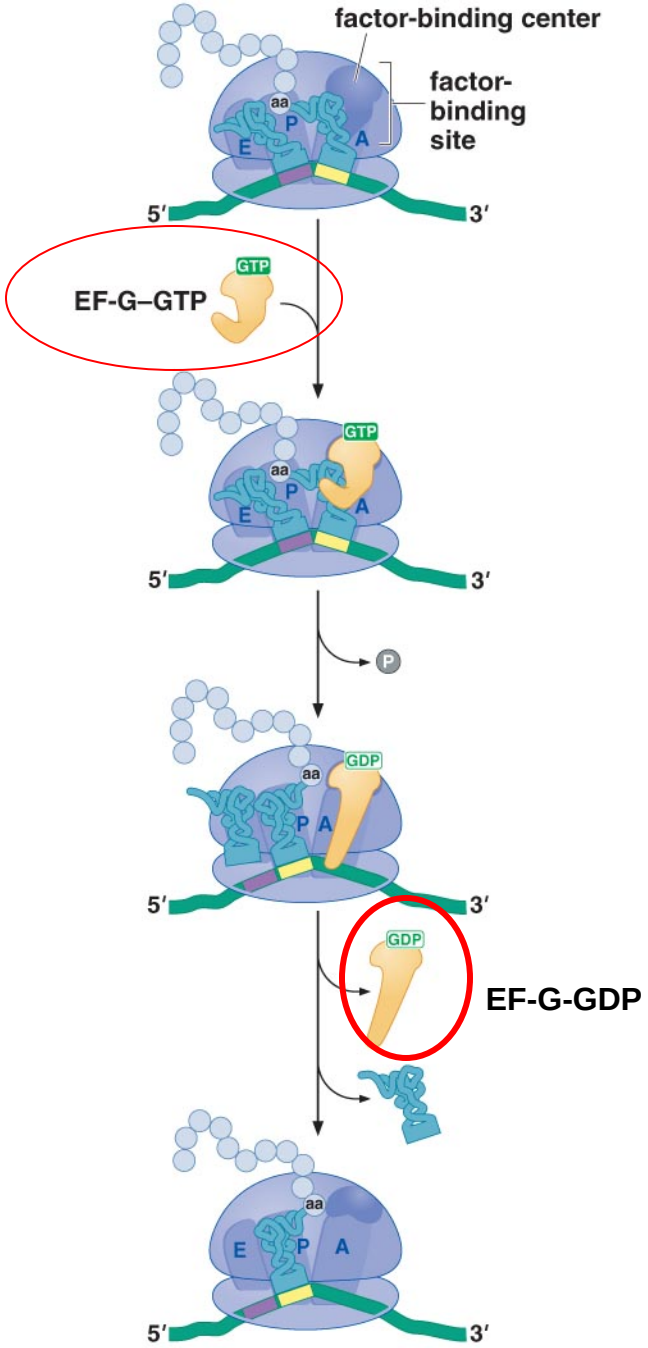


Flush pairing of anticodon with codon is needed to bring complex into position for GTP cleavage

Quality control detector

- No GTP cleavage, no peptide bond can be formed
- No GTP cleavage, aa-tRNA-EF-Tu is free to leave





Translocation of the ribosome, tRNA, peptide complex along the RNA is needed for addition of the next amino acid

Again, GTP to GDP hydrolysis fuels the machine – this time, for translocation and evicting the redundant tRNA from the "E" site

Termination of translation

- No “terminator tRNAs”
- Instead, protein “release factors” enter A site and interact with stop codons
- Some release factors clip finished peptide off the last tRNA
- Other release factors interact with EF-G to kick used-up tRNAs and mRNA out of ribosome

Constant quality control of poly(A) tail and 5’caps as long as translation is going on