FROM GENE TO PROTEIN

How does DNA carry a code?

- What is a gene?
- What does that code (gene) tell a cell to do?

What do you remember about protein?

How are proteins related to physical traits?

- (This is what we call gene expression)

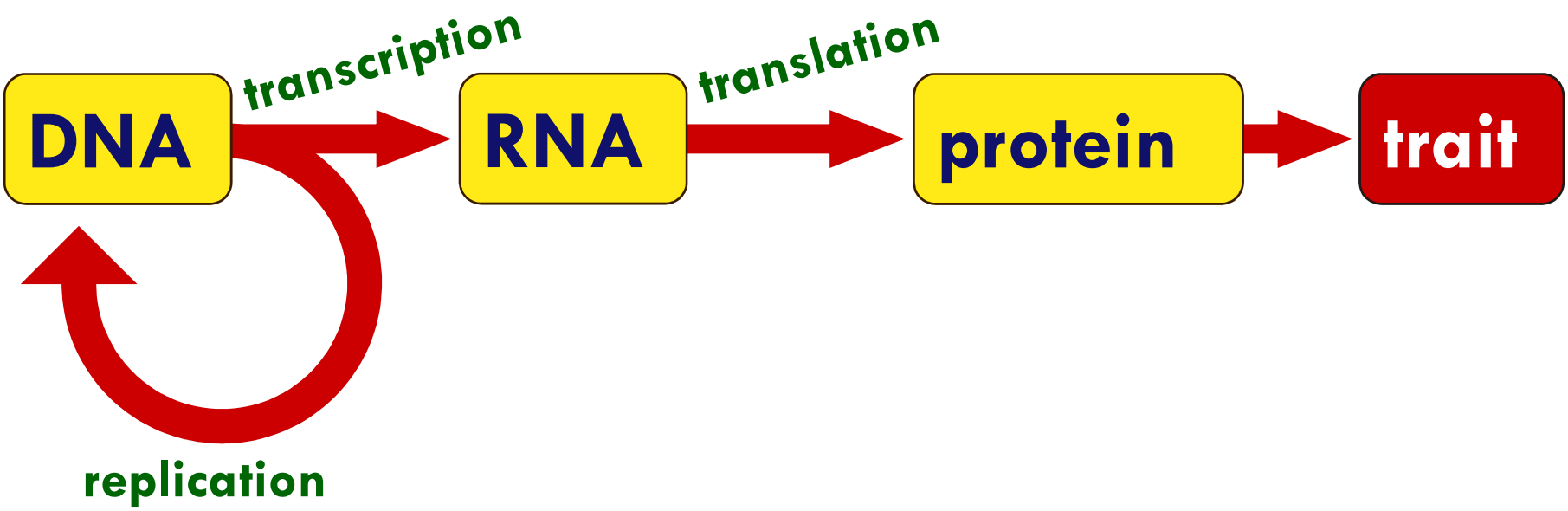
How does this all happen?

- DNA in the nucleus
- Ribosomes are in the cytoplasm

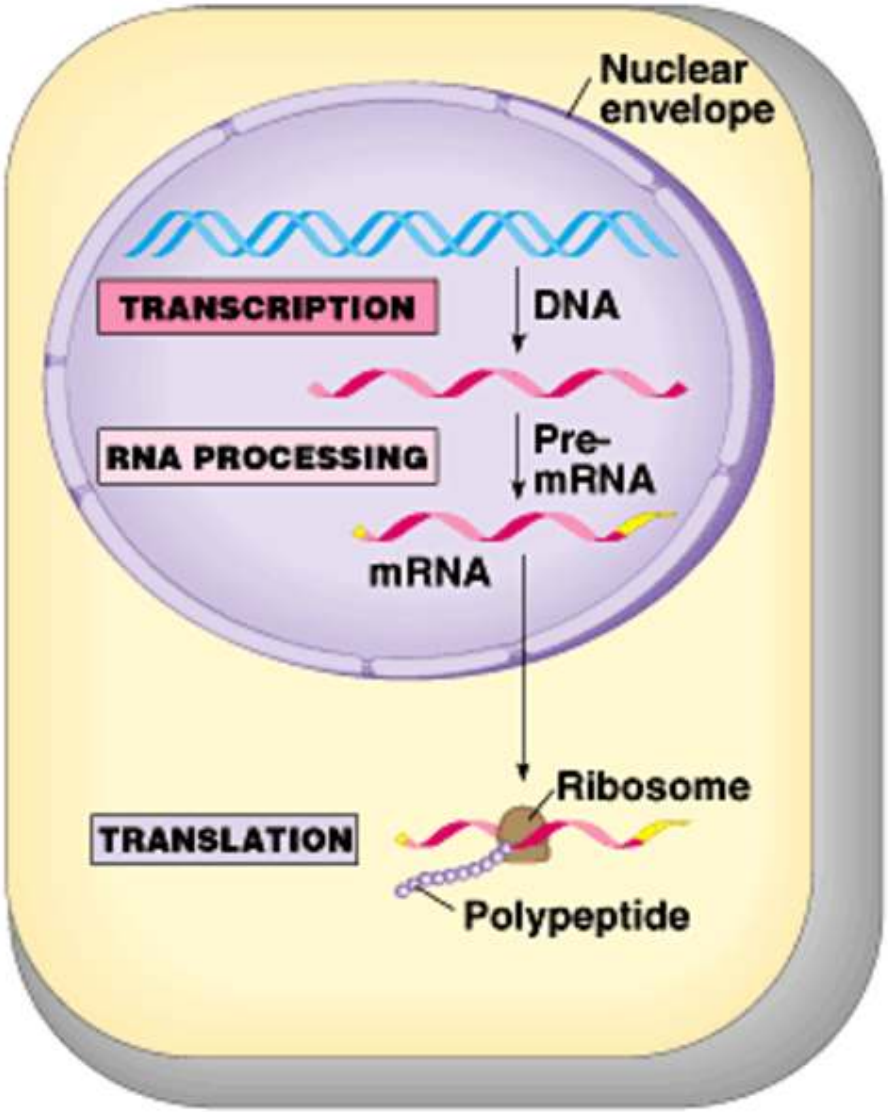
Two important processes

- Transcription
- Translation

The "Central Dogma" of Biology

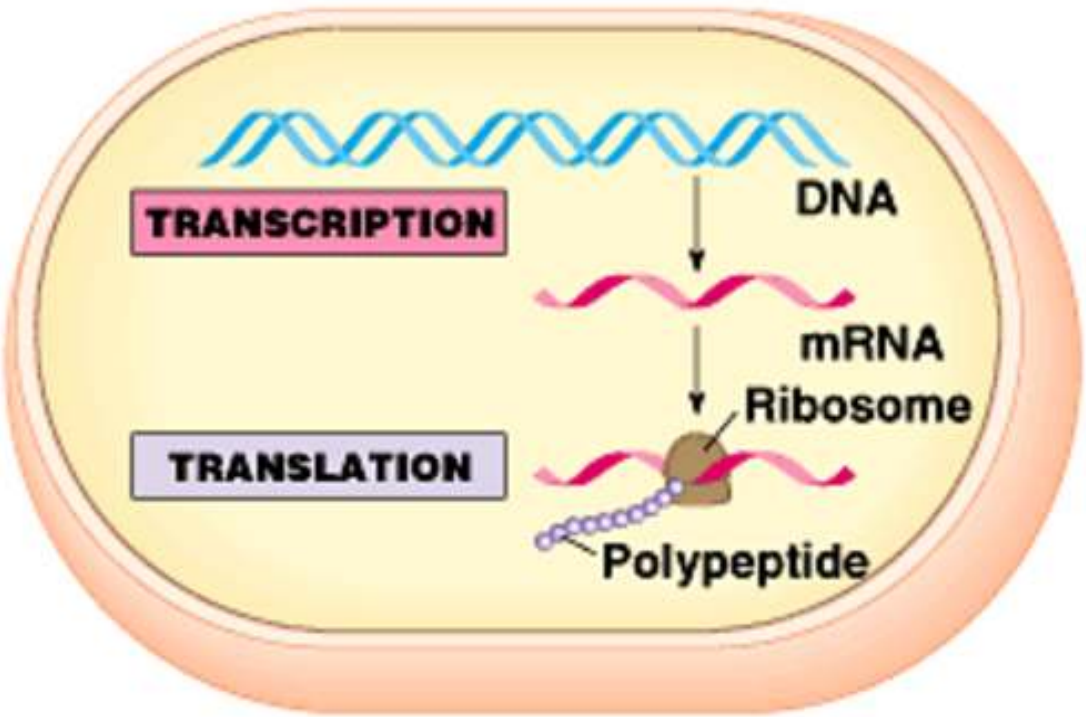


Gene expression overview



(b) Eukaryotic cell

Gene expression in prokaryotes



(a) Prokaryotic cell

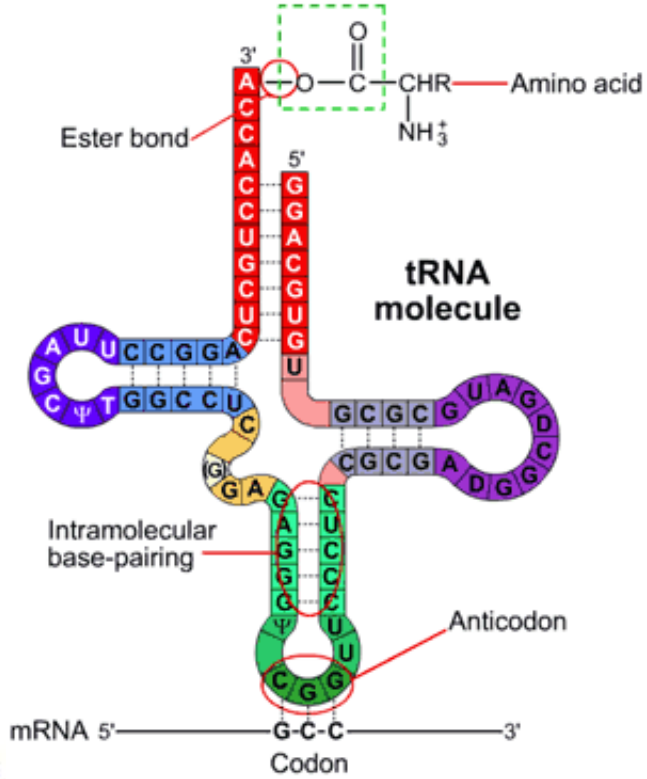
What is the end result?

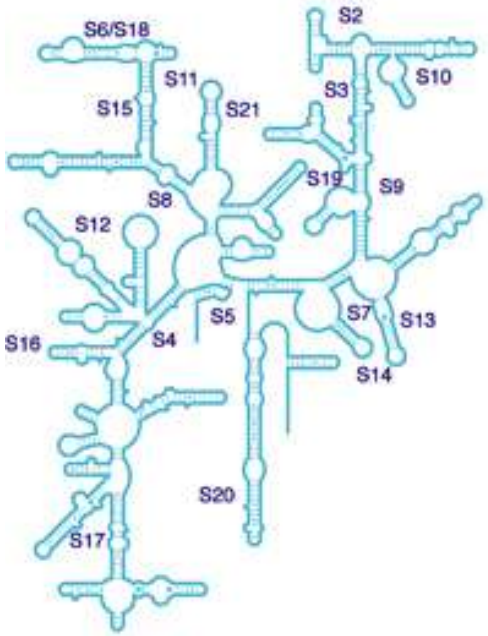
- Cellular chain of command
  - DNA $\rightarrow$ RNA $\rightarrow$ Protein

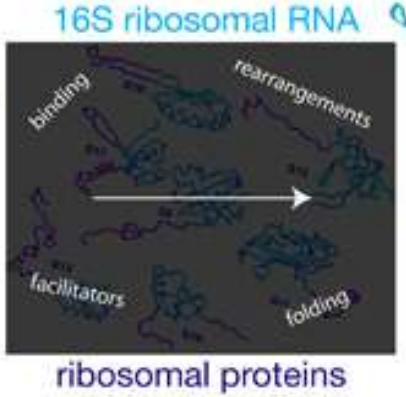
Real quick RNA review

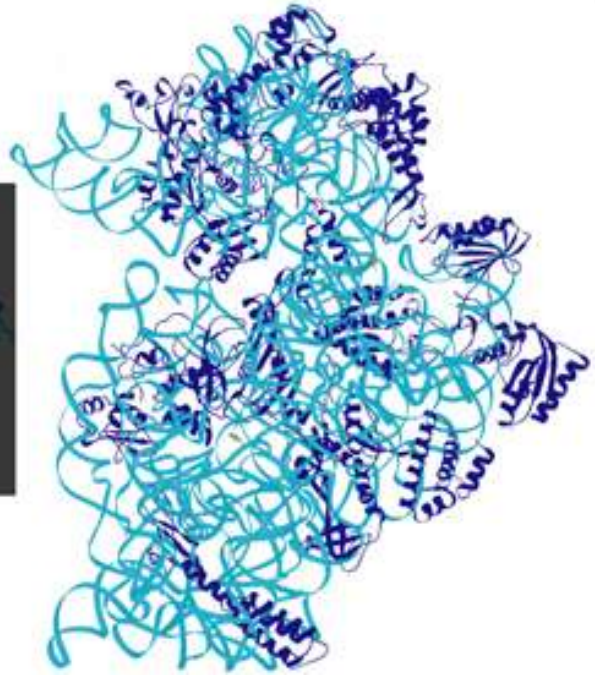
- Sugar and shape
- Bases
- The 3 main types of RNA
  - mRNA (messenger)
  - tRNA (transfer)
  - rRNA (ribosomal)


mRNA
5`
3`
A G C U G A C C U A G C G G A C A A


Amino acid
Ester bond
tRNA molecule
Intramolecular base-pairing
Anticodon
mRNA 5'
G-C-C
3'
Codon


S6/S18
S2
S11
S10
S15
S21
S3
S19
S9
S8
S12
S5
S7
S13
S16
S4
S14
S20
S17


16S ribosomal RNA
binding
rearrangements
facilitators
folding
ribosomal proteins



Why do we need to translate DNA into RNA anyway?

Ok, time for Transcription

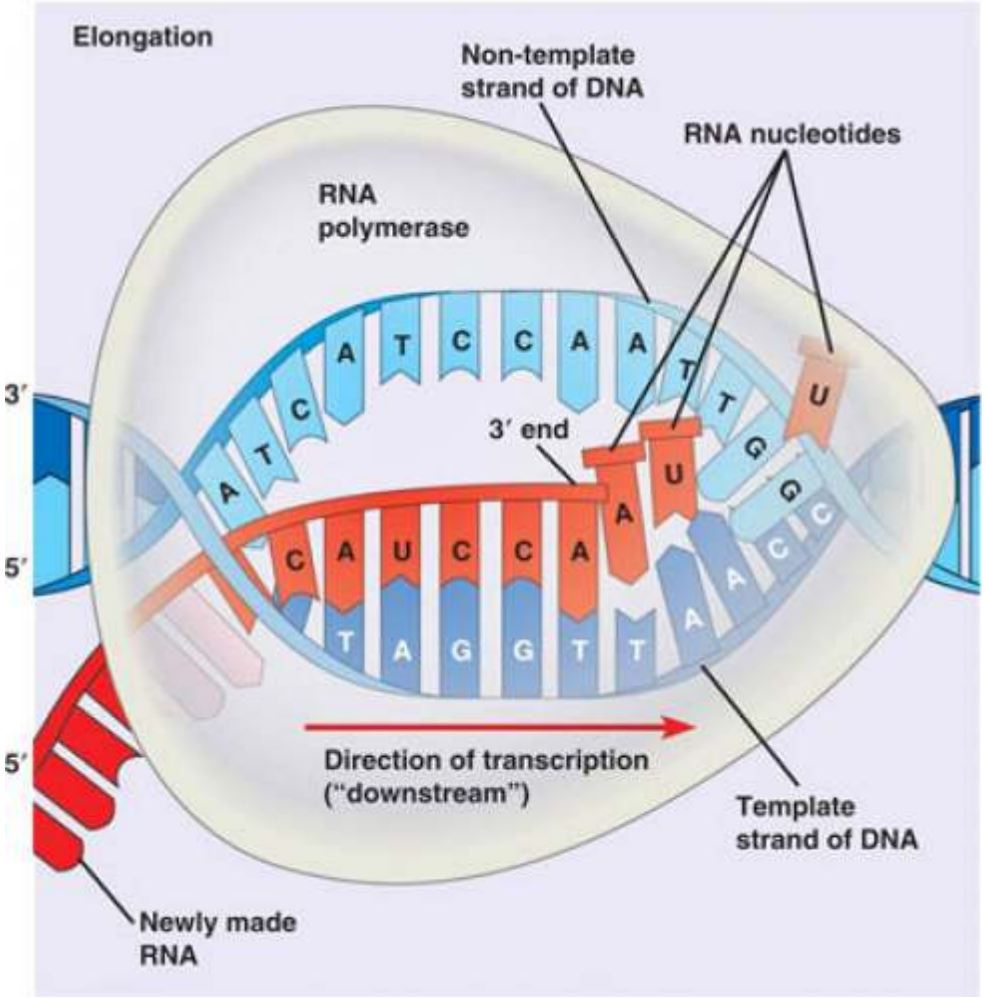
- DNA is read by RNA polymerase
  - What is that?
  - What does it do?
  - What does it read?
  - Which direction does it read?
  - How does it know where to start?

Ok, time for Transcription

- Transcription
  - Information from DNA is copied into RNA

Transcription

- Elongation: RNA polymerase begins moving downstream
  - Unwinds DNA
  - Adds complementary RNA nucleotides (5' to 3')
  - DNA reforms double helix afterwards
  - Approx. 60 nucleotides per second



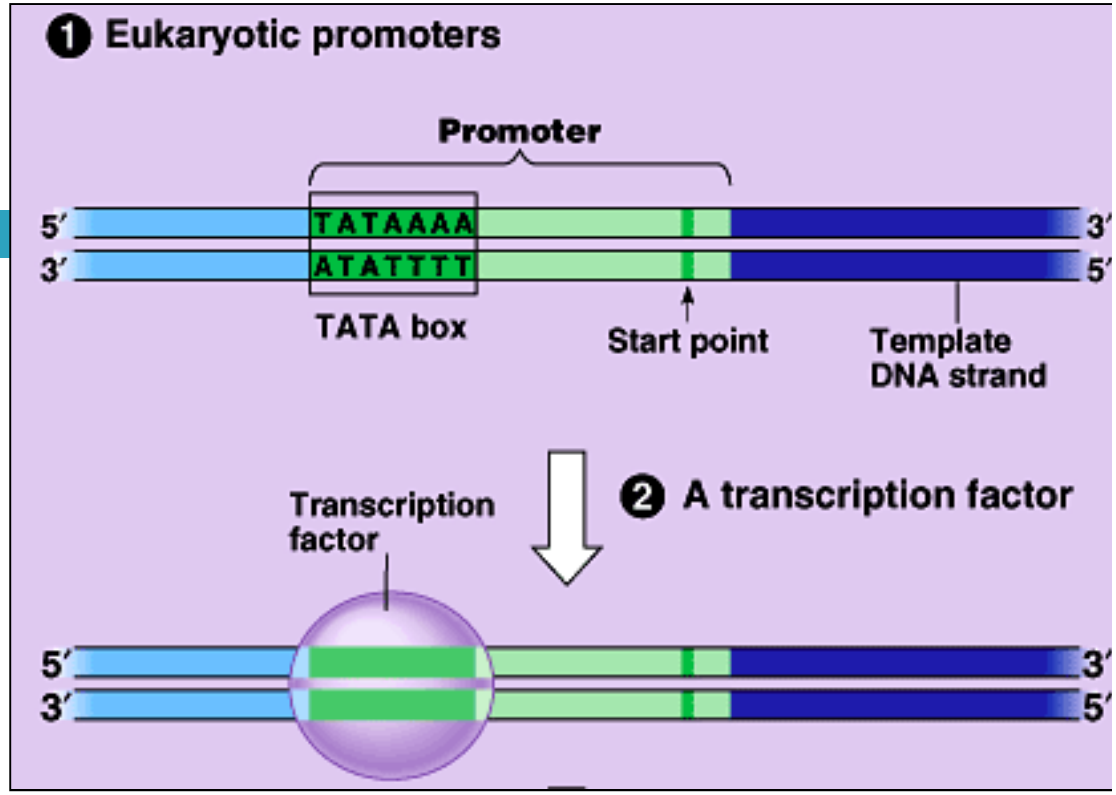
Let's practice

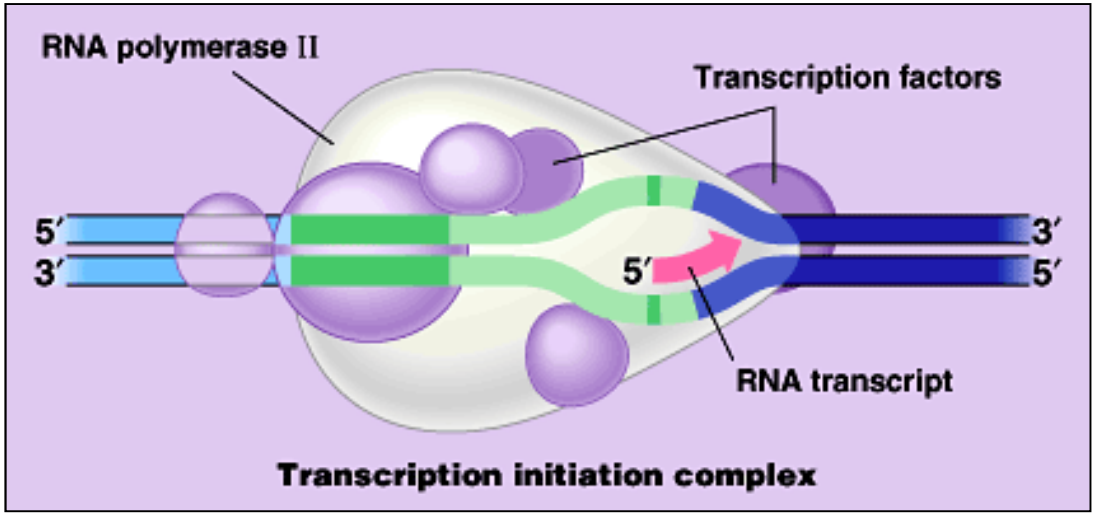
- If this is the DNA code from the template strand, what is the corresponding RNA strand?

A C T A C G G C A A C T T T C

*Transcription

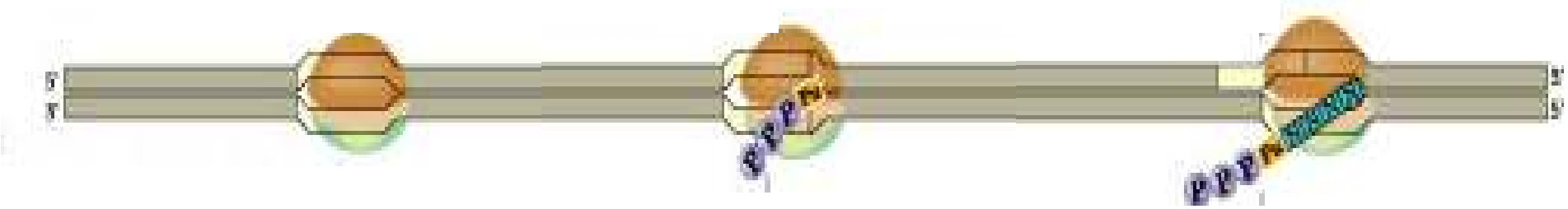
- DNA has a promoter
  - DNA sequence upstream from site of transcription
  - TATA box
    - Template vs. non-template strand
    - Transcription factors





*Transcription

- Multiple RNA polymerases can be working on the same gene



*Transcription Termination

- How does it know where to stop?
- Prokaryotes: termination sequence in DNA
- Eukaryotes: RNA pol. II reaches the polyadenylation sequence (AAUAAA)

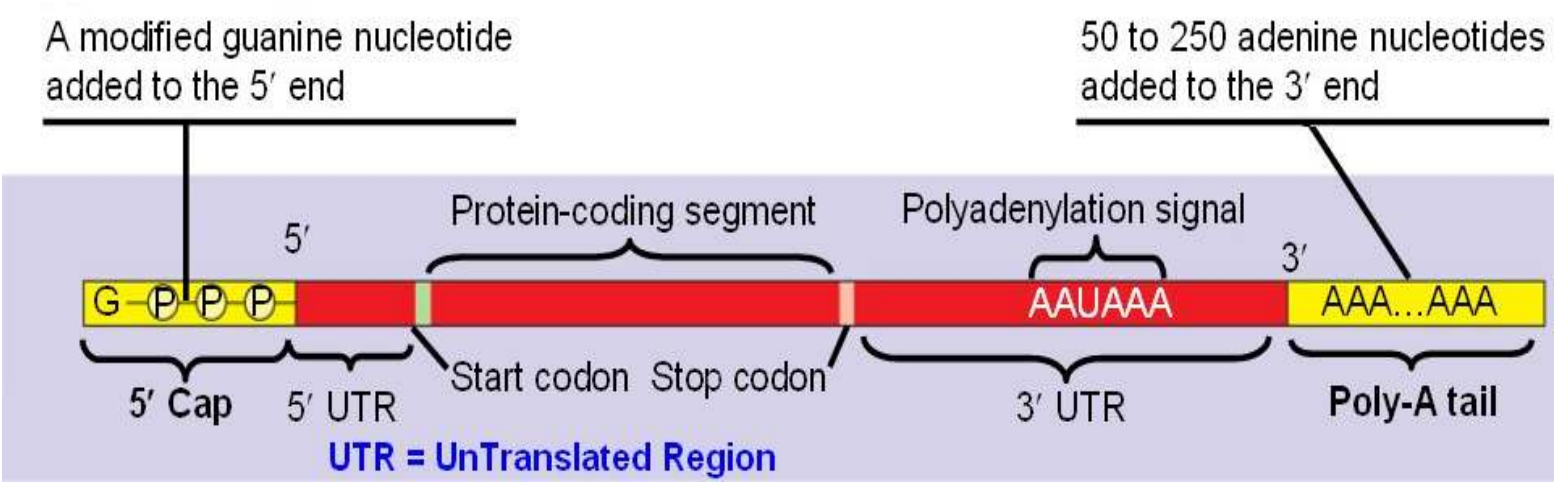
- Done. Transcription has produced a pre-mRNA strand

Now what? Is it a protein yet?

- Nope
- Where did all of this happen?
- Pre-mRNA is ready for processing
  - mRNA is prepared to leave the nucleus for translation
  - Basically, pre-mRNA is physically modified to become mRNA

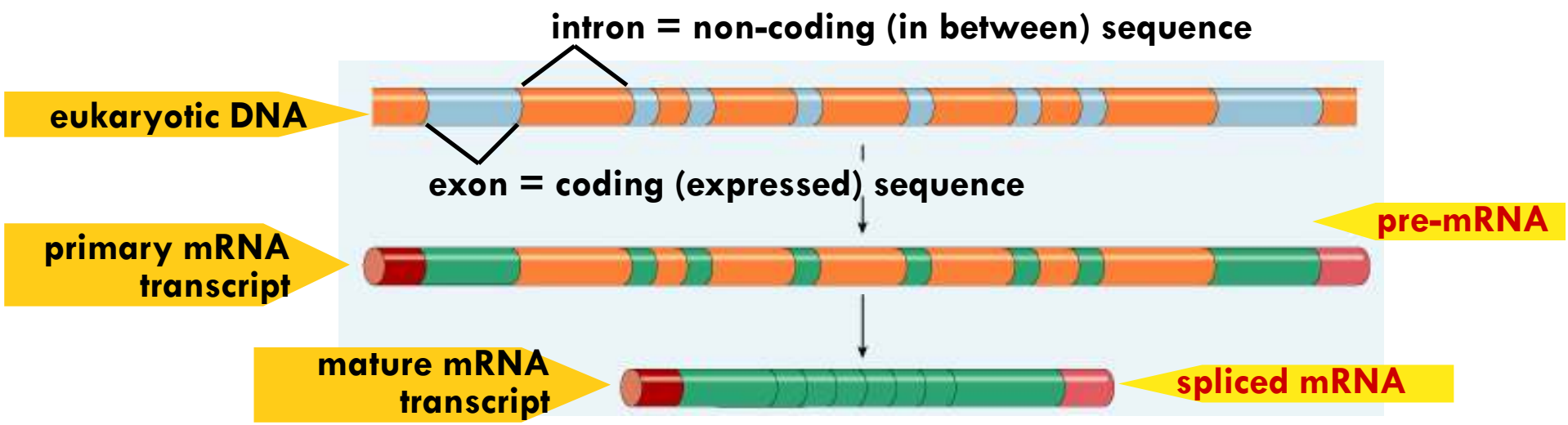
Pre-mRNA processing

- 5' end receives a guanine nucleotide cap
- 3' end receives a poly-A tail (AAAAAAAAA)
- Why?
  - Export
  - Protection from enzymes
  - Attach to ribosome



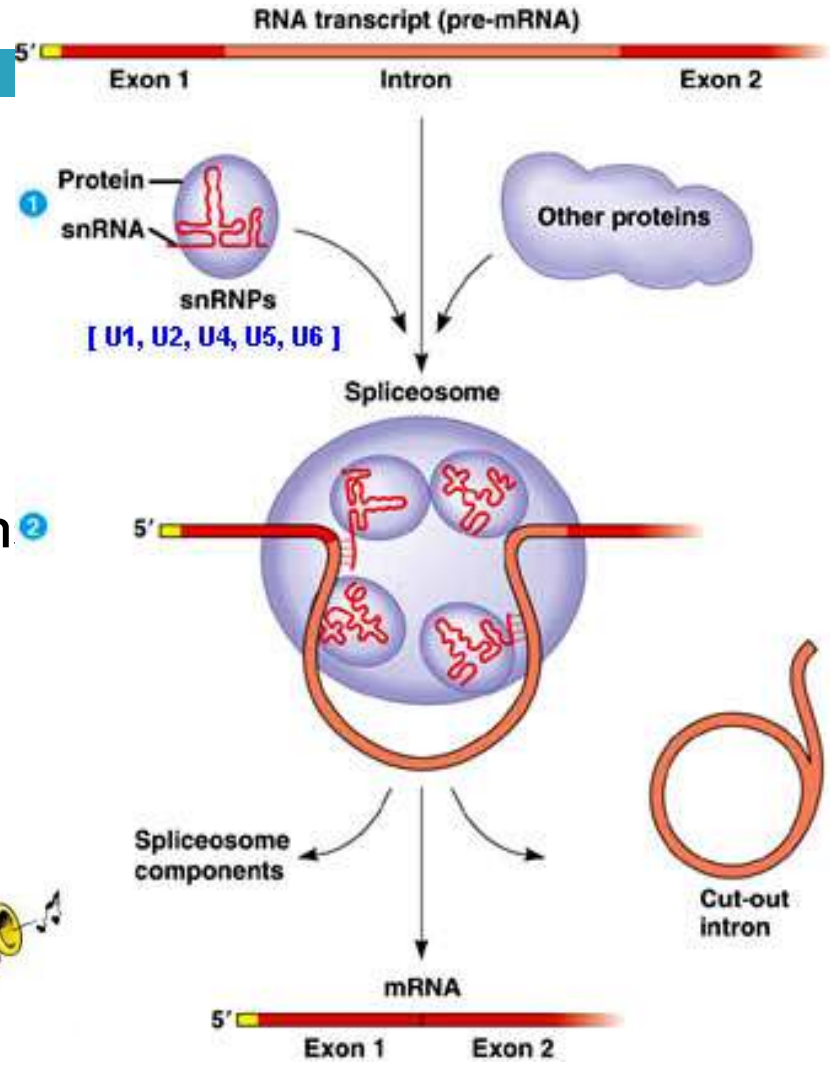
RNA splicing

- 8000 nucleotides transcribed; only 1200 code for an average protein (400 aa)
- Genes are not continuous – filled with "junk"
  - **Introns**- non-coding in between filler
  - **Exons** – expressed, real genes, exit the nucleus



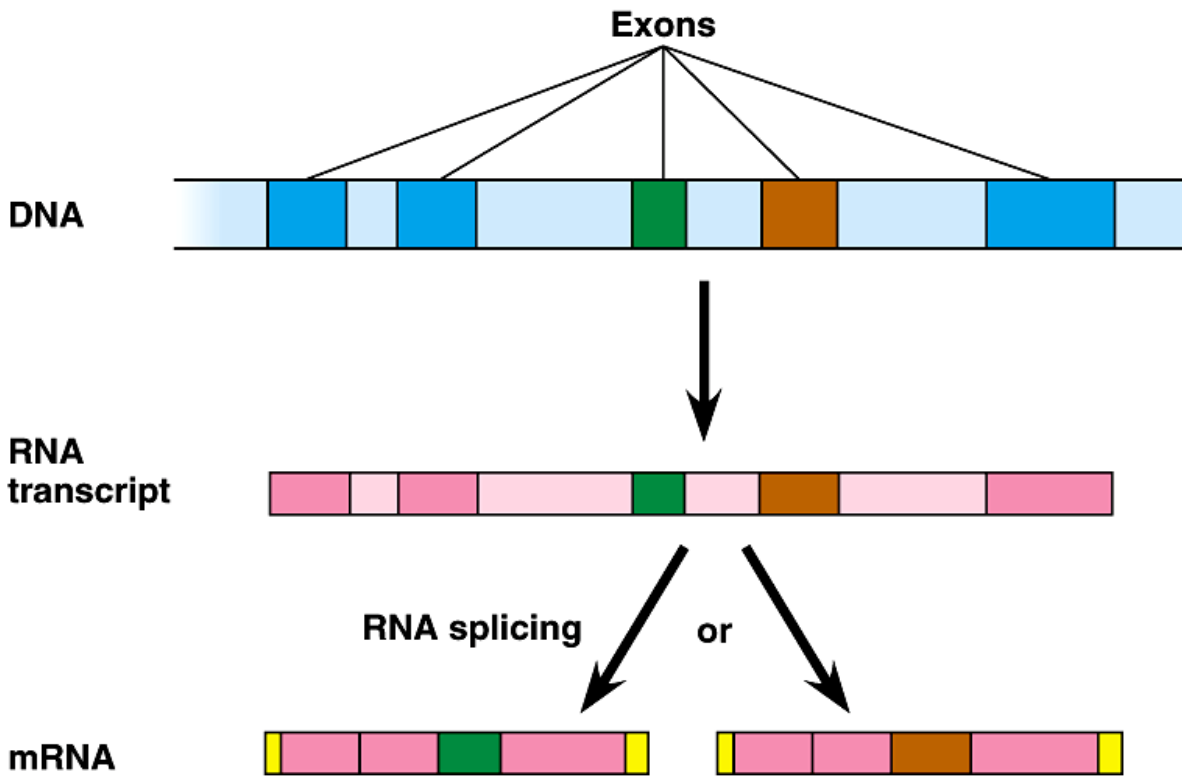
*How does splicing occur?

- snRNP's (small nuclear ribonucleoproteins)
  - Found in the nucleus
  - Make up the spliceosome
  - Spliceosome recognizes intron removes them and reconnects the mRNA
  - Ribozymes



*Alternative RNA splicing

- 1 gene can code for more than one polypeptide
  - Depends on which sequences treated as exons
  - Maybe this is why we have so few genes (only 2x drosophila!)
  - Possibly helpful for evolution



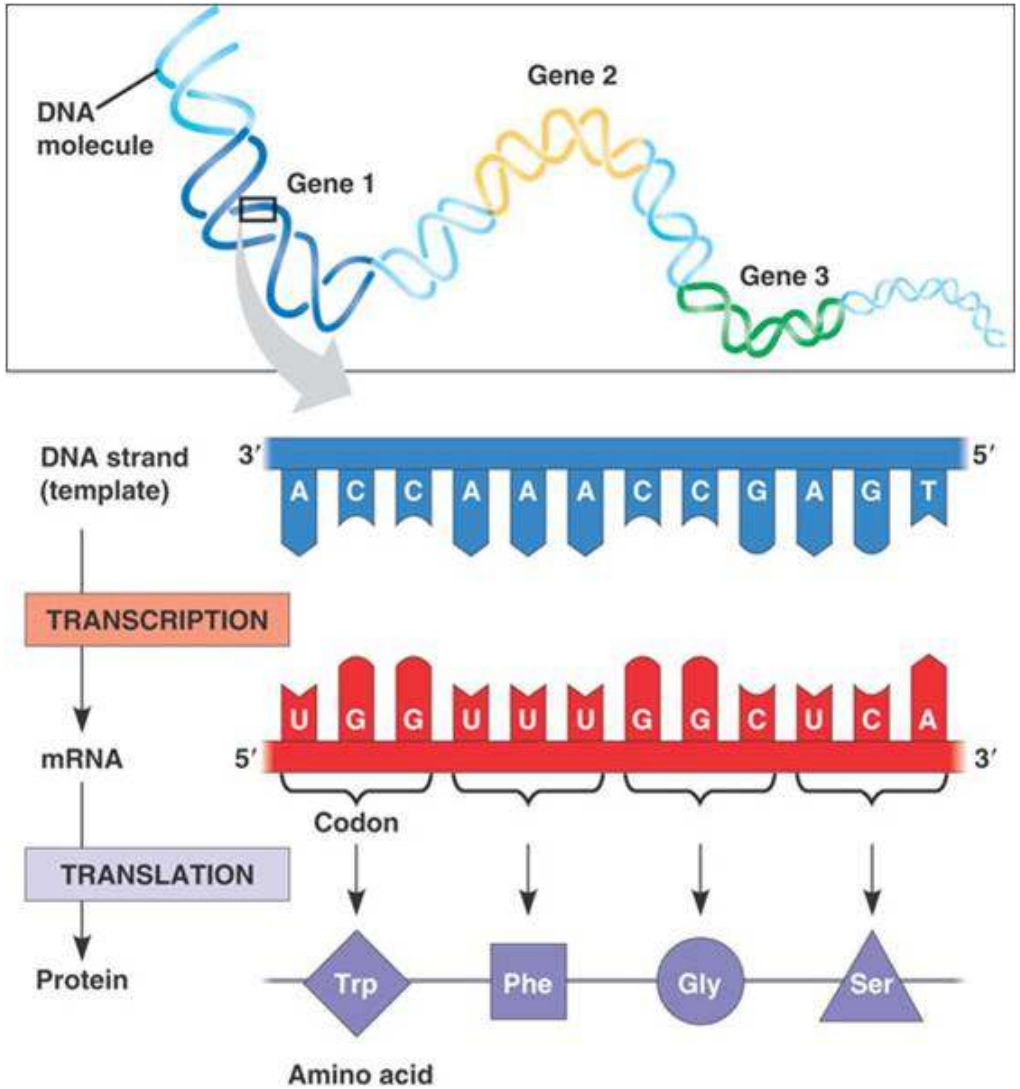
What now? Translation!

- Translation – Information from RNA is converted into a protein
  - The language of nucleotides is translated into the language of amino acids

- Basically, information found on mRNA is read by the ribosome and used to create a sequence of amino acids, otherwise known as a polypepetide or protein

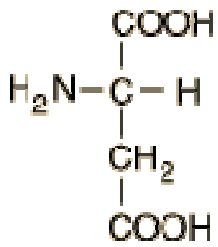
How does it work- Summary

- The ribosome reads 3 nucleotide bases at a time. (triplet)
  - Codon
  - mRNA is read from 5' to 3' in segments of 3 bases
- A codon signals for a specific amino acid.

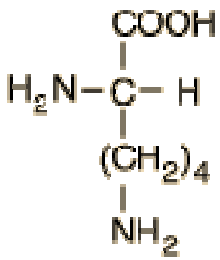


*Hold on a second…

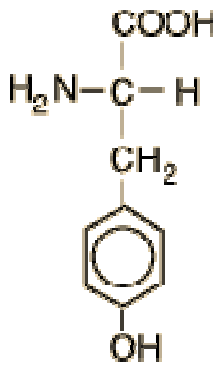
- How do the right amino acids end up in the right locations? (And why does that even matter?)
  - 20 AA but only 4 bases



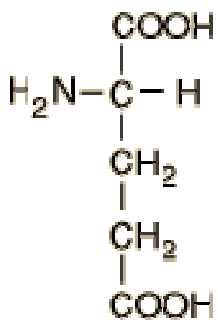
Aspartic acid (asp)



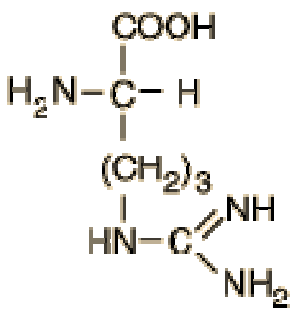
Lysine (lys)*



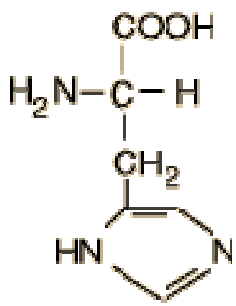
Tyrosine (tyr)



Glutamic acid (glu)



Arginine (arg)*



Histidine (his)*

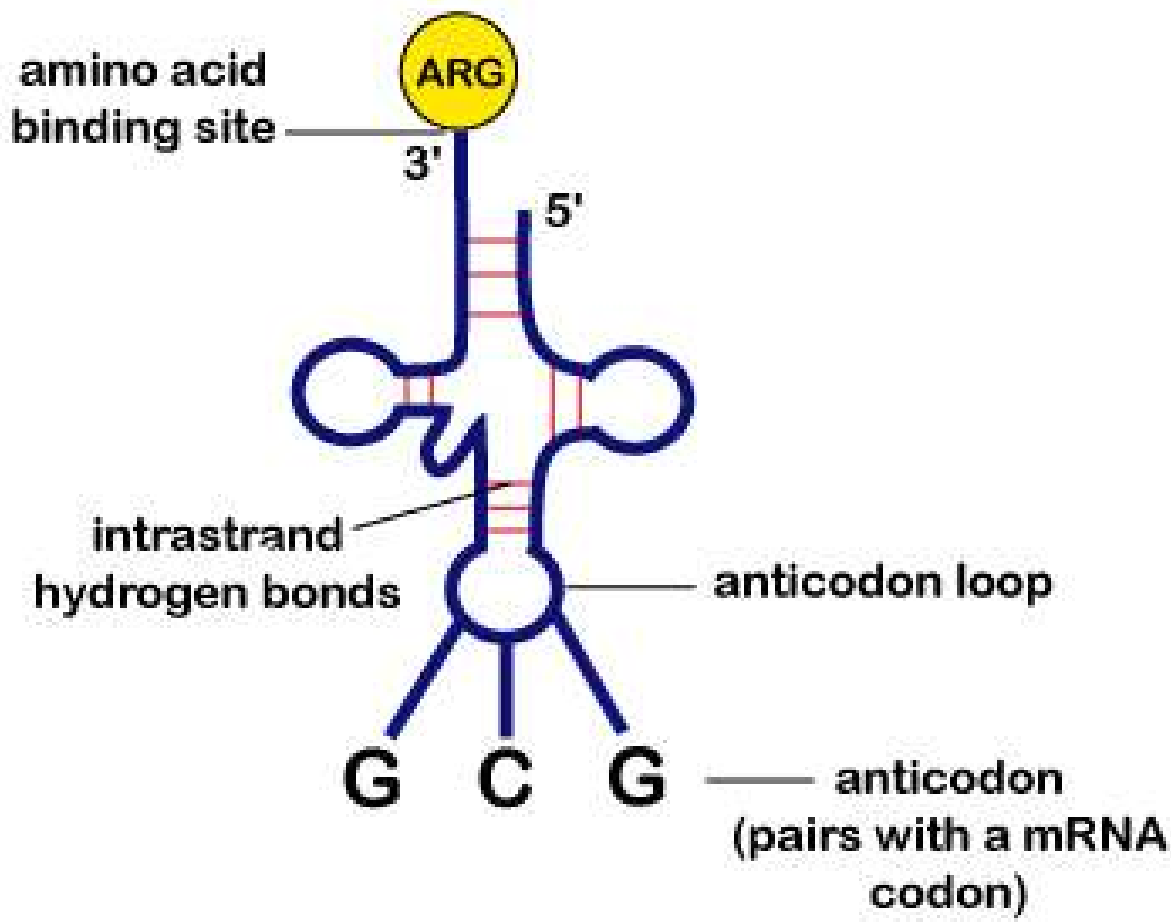
The Genetic Code

- Observations?
  - Start and stop
  - Multiple ways for one AA
  - Reading frame matters
- If a gene codes for an mRNA molecule with 1236 nucleotides, how many amino acids are found in the protein?

First base (5′ end)	Second base: U	Second base: C	Second base: A	Second base: G	Third base (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 or start	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

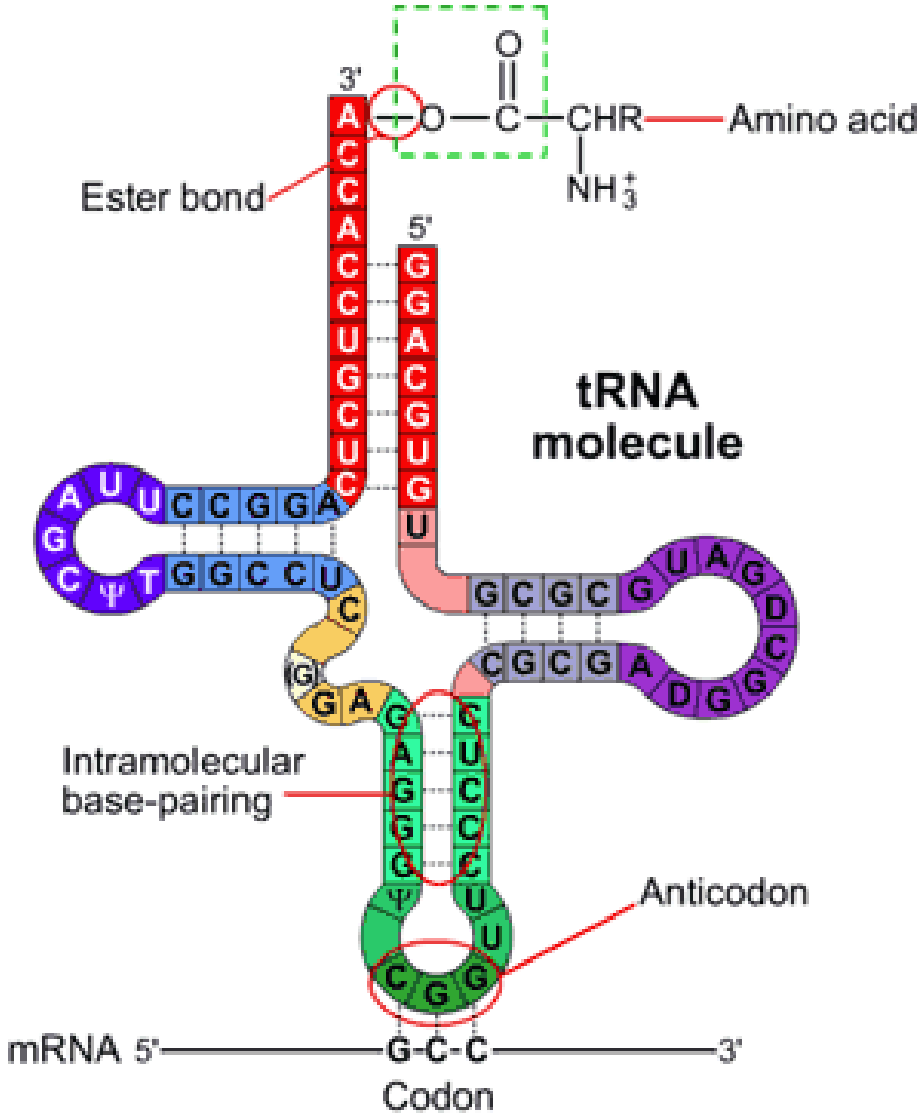
How does translation work?

- Amino acids are brought to the ribosome by tRNA
- tRNA have 3 nucleotide bases that match up with the codon
  - Anticodon
- Recycled
  - Picks up AA in cytoplasm



*tRNA

- Each tRNA molecule only carries one specific type of AA which matches the anticodon
- tRNA anticodons match up with mRNA codons
- This make sure the right AA are put in the right place



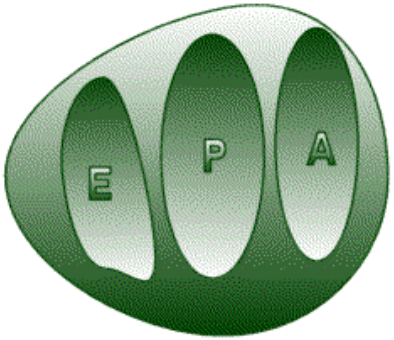
How does translation work?

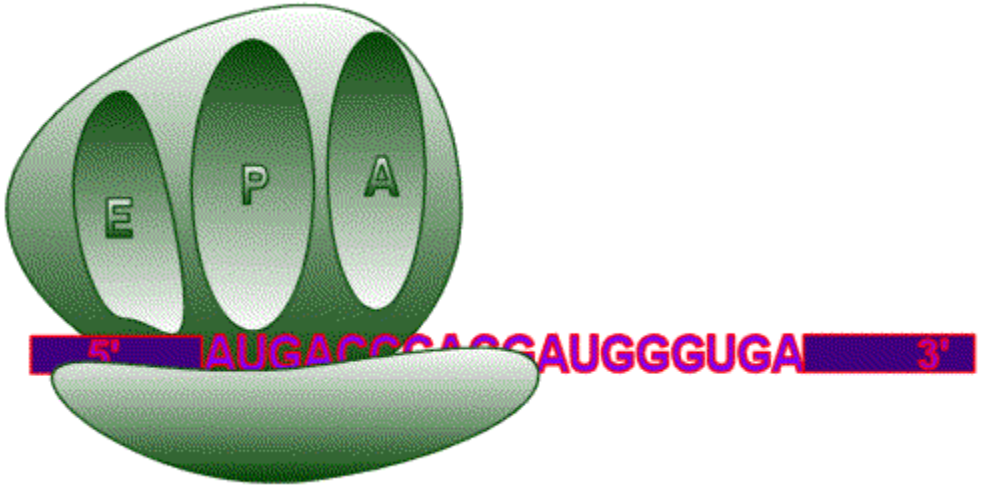
- The ribosome helps match up the bases from the mRNA with the bases from the tRNA
  - Codon with anticodon

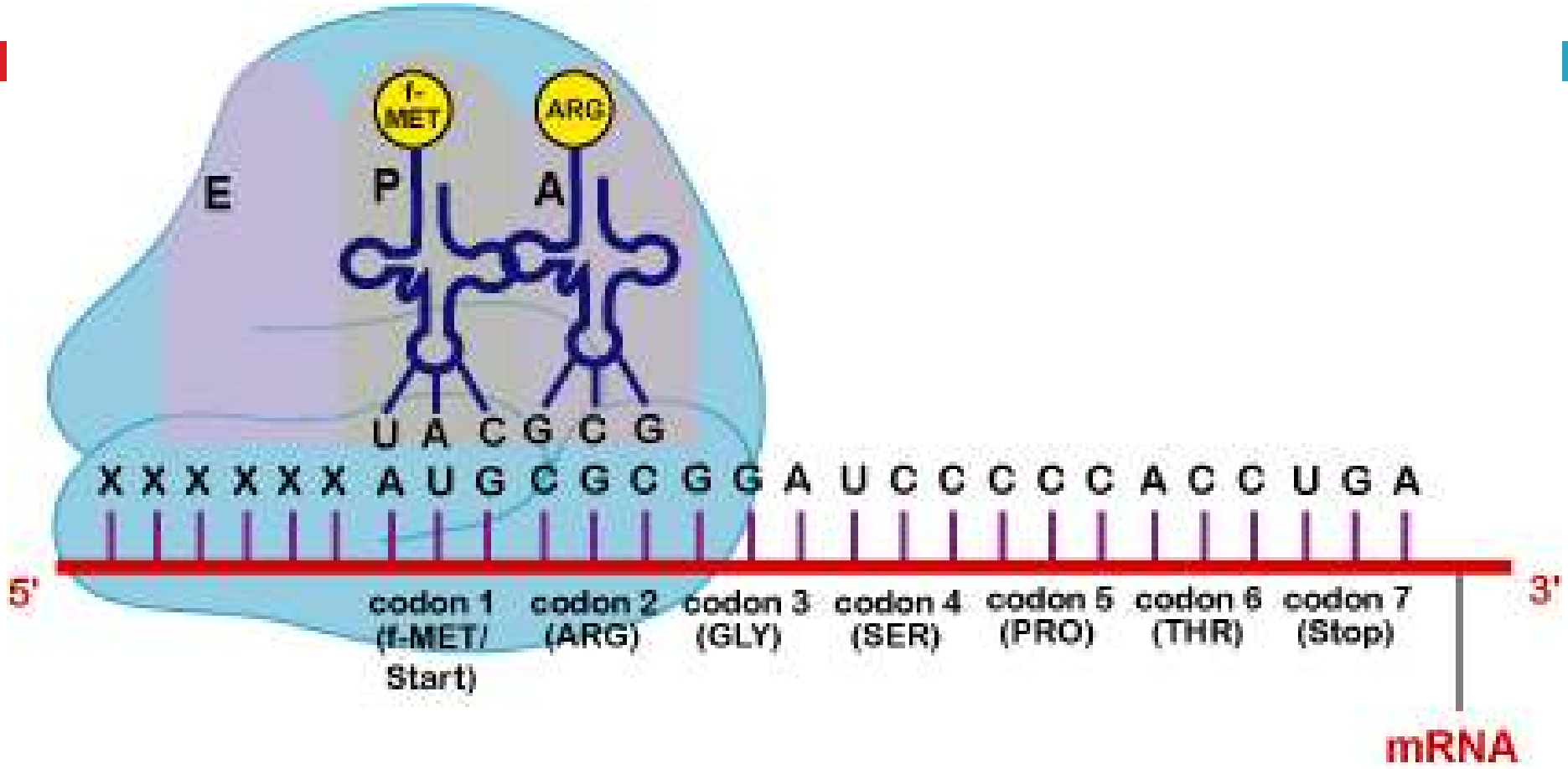
- The ribosome helps connect the amino acids together to form a chain- future protein!
  - Covalent peptide bond

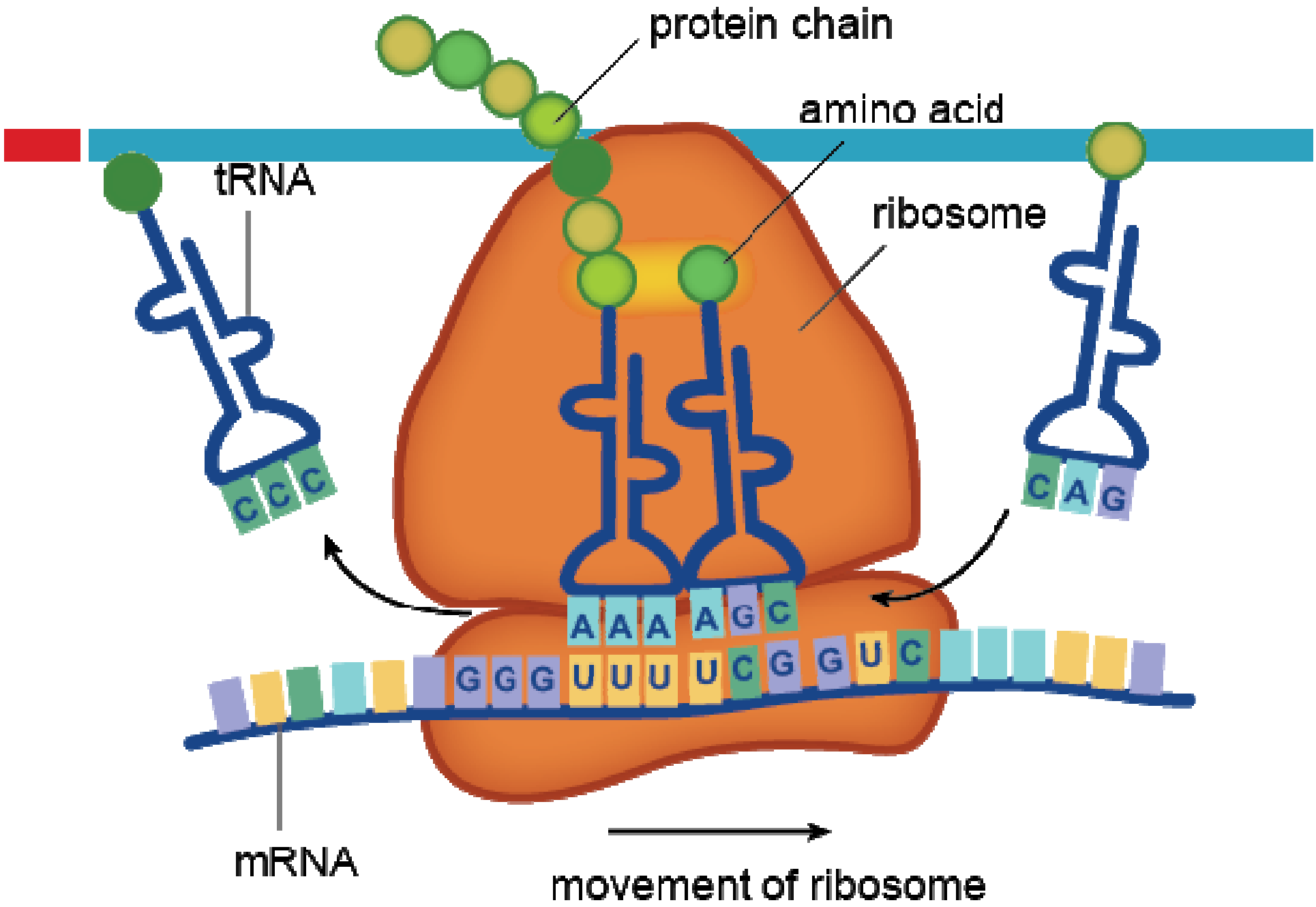
*Ribosomes

- Made from rRNA
- 2 subunits, made in the nucleolus
  - Large
  - Small
- Subunits join only when attaching with mRNA






f-MET
ARG
E
P
A
U A C G C G
X X X X X X A U G C G C G G A U C C C C C A C C U G A
5'
3'
codon 1 (f-MET/ Start)
codon 2 (ARG)
codon 3 (GLY)
codon 4 (SER)
codon 5 (PRO)
codon 6 (THR)
codon 7 (Stop)
mRNA


protein chain
amino acid
tRNA
ribosome
C C C
C A G
A A A A G C
G G G U U U U C G G U C
mRNA
movement of ribosome

First base (5′ end)	Second base: U	Second base: C	Second base: A	Second base: G	Third base (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 or start	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

Practice

- DNA

  ACC CGA TAC

- mRNA

- Amino acid code

Practice

- DNA

  TTC GAC CGT

- mRNA

- Amino acid code

Translation overview

- Ribosome binds to mRNA
- tRNA comes with Amino Acid and binds to mRNA
  - tRNA anticodon binds with mRNA codon
- Ribosome (rRNA) facilitates formation of a peptide bond between AA
- Ribosome moves downstream on mRNA allowing free tRNA to leave
- The process continues until a “stop” codon is reached
- A polypeptide chain is released
  - Ribosome disassembles
- Forms a protein- folds and coils as it is made

Translation

- http://highered.mcgraw-hill.com/olcweb/cgi/pluginpop.cgi?it=swf::535::535::/sites/dl/free/0072437316/120077/micro06.swf::Protein%20Synthesis

- Summarize!