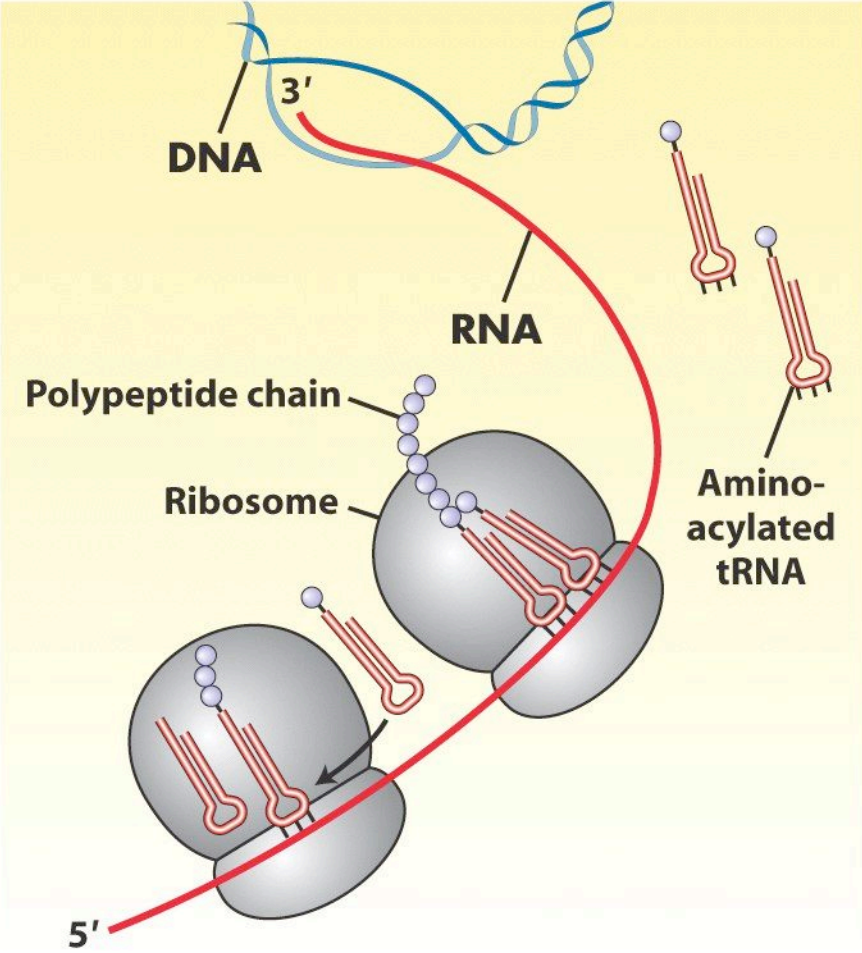


Central Dogma of Genetics

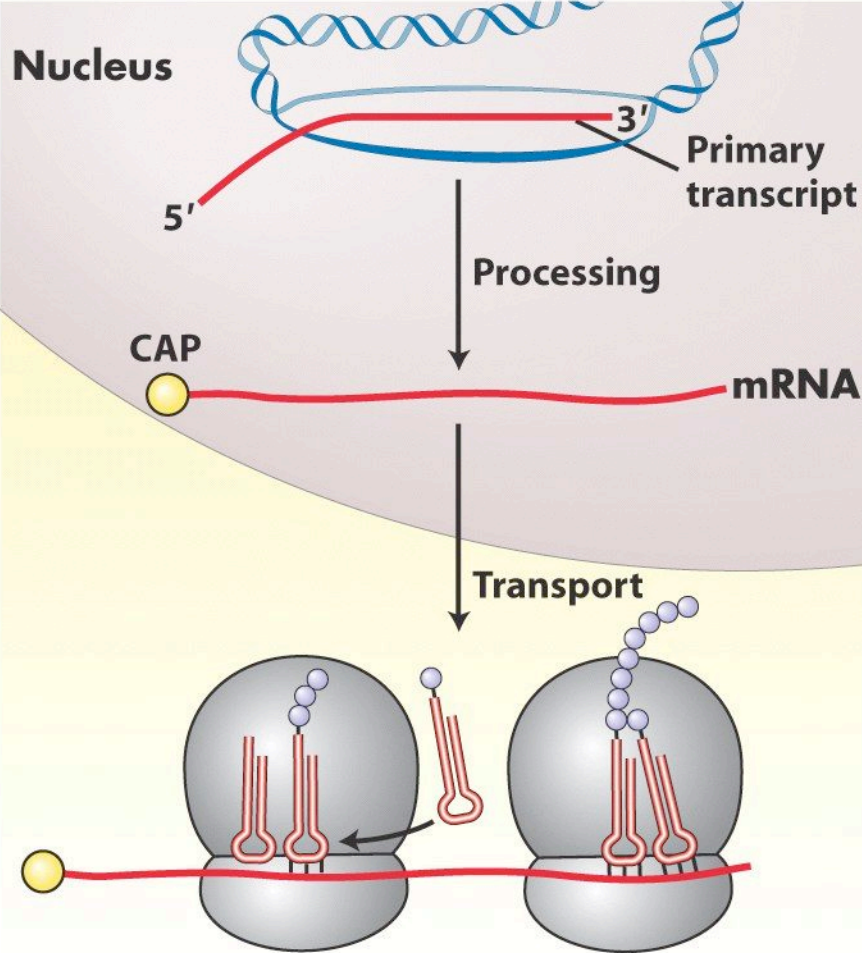
- Within each cell the genetic information flows from
 - **DNA to RNA to protein.**
- This flow of information is unidirectional and irreversible.
- The information carried within the DNA dictates the end product (protein) that will be synthesized.
 - This information is the **genetic code**.
- Conversion of DNA encoded information to RNA
 - is called **transcription**.
- The information from a mRNA is then **translated** to an amino acid sequence in the corresponding protein

Central Dogma

PROKARYOTE



EUKARYOTE



Proteins

- Building blocks are Amino acids
- AA's are joined between the carboxyl end of one AA to the amino group of the other
 - One water molecule is lost
 - Enzyme called peptidyl transferase is involved

Protein Structure

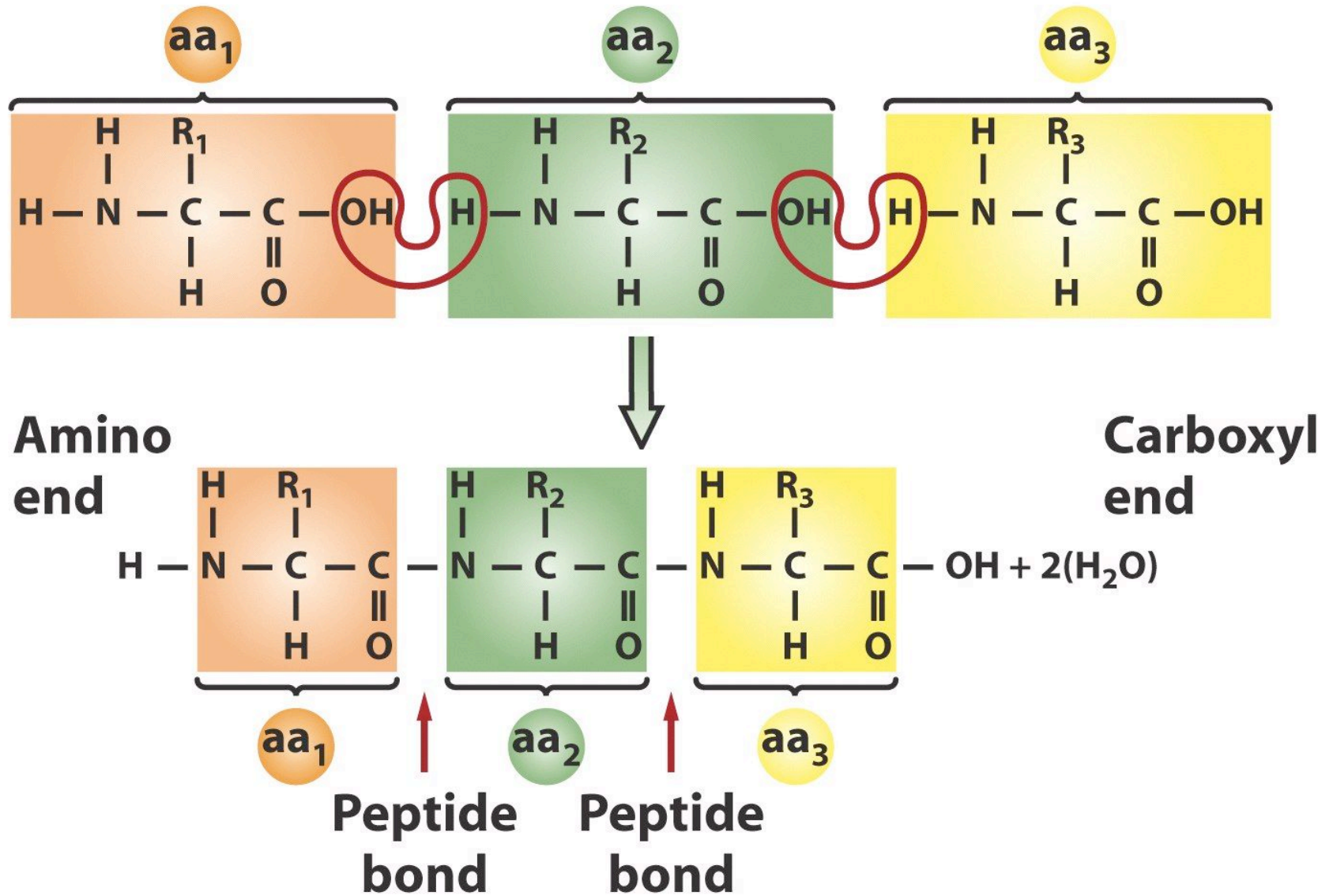
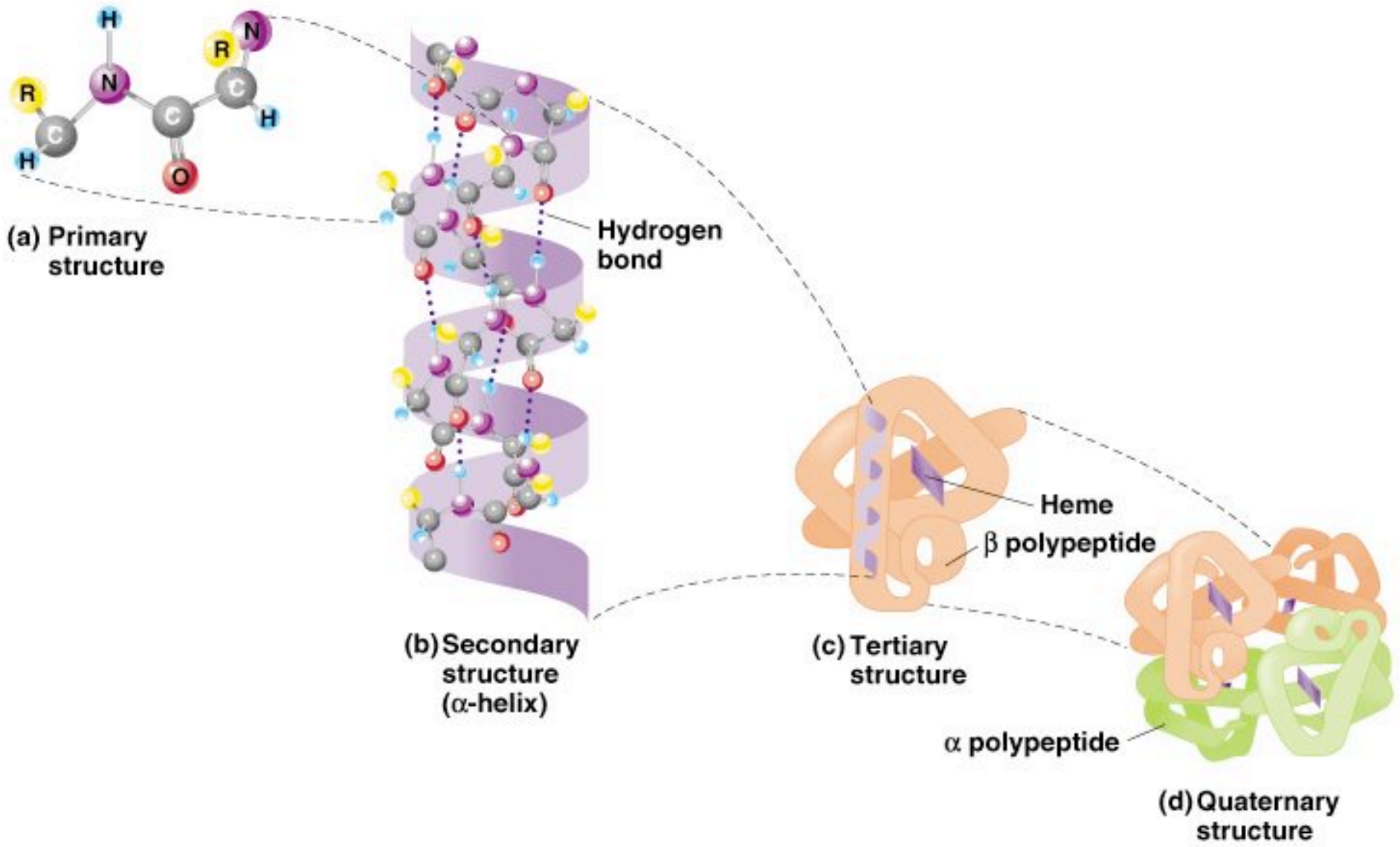


Fig. 14.4 Four levels of protein structure



Deducing the genetic code

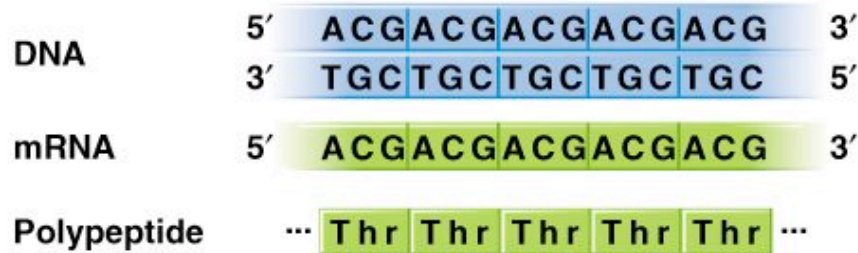
- Before the genetic code was cracked, it was a mystery as to how four nucleotides could encode for 20 amino acids.
- One or two nucleotides representing each amino acid would not suffice
 - A one-letter code could specify four amino acids
 - A two letter code could specify 16 ($4^2 = 16$)
 - To accommodate 20, at least three letters are needed ($4^3 = 64$)
- therefore 3 letters to a codon was the most likely option
- Experimental research carried out on bacteria (figure 6.5) showed that the genetic code is a triplet code.

Understanding the Genetic Code

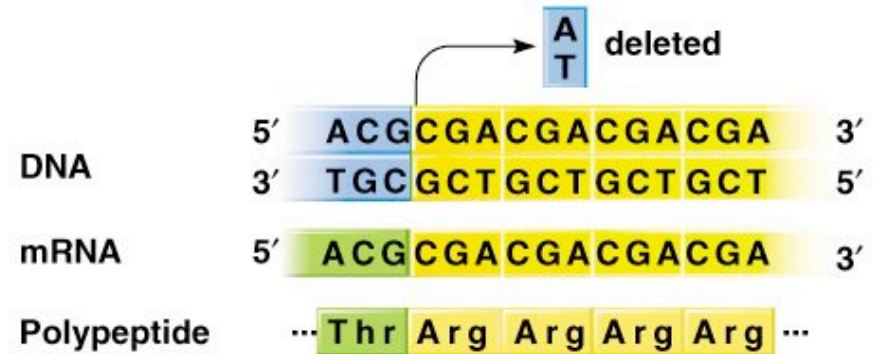
- In the 1950s and 1960s, Charles Yanofsky studied the gene tryptophan synthetase,
 - which synthesizes Tryptophan (Trp),
 - and allows bacteria to grow in Trp free medium.
- He made several substitution mutants (**auxotrophs**),
 - which required Trp in the medium in order to grow.
- Using genetic recombinations, he made a map of the various mutations.
- Using technology available in the 1950s he sequenced the mutant proteins being synthesized to correlate the mutation in the DNA sequence with the corresponding change in the amino acid sequence.

Fig. 14.5 Reversion of a deletion frameshift mutation by a nearby addition mutation

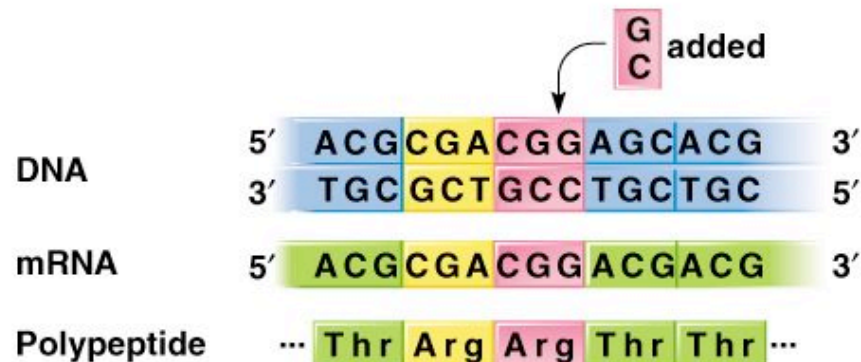
a) Wild type



b) Frameshift mutation by deletion



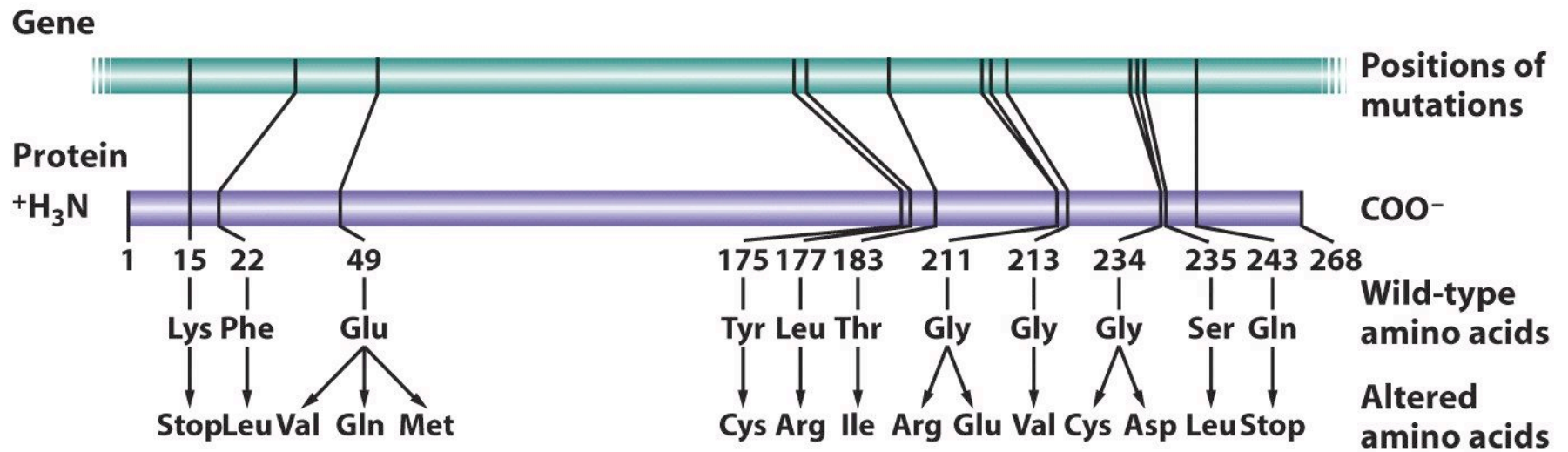
c) Reversion of deletion mutation by addition



Interpretations from Yanofsky's experiments

- The changes were colinear.
- Each mutation resulted in only one amino acid change:
 - Each nucleotide was part of only one codon: the genetic code was non-overlapping.
- Different point mutations in the same position (based on his power of resolution) could result in different amino acids in the product:
 - The nucleotide sequence determines the amino acid sequence.

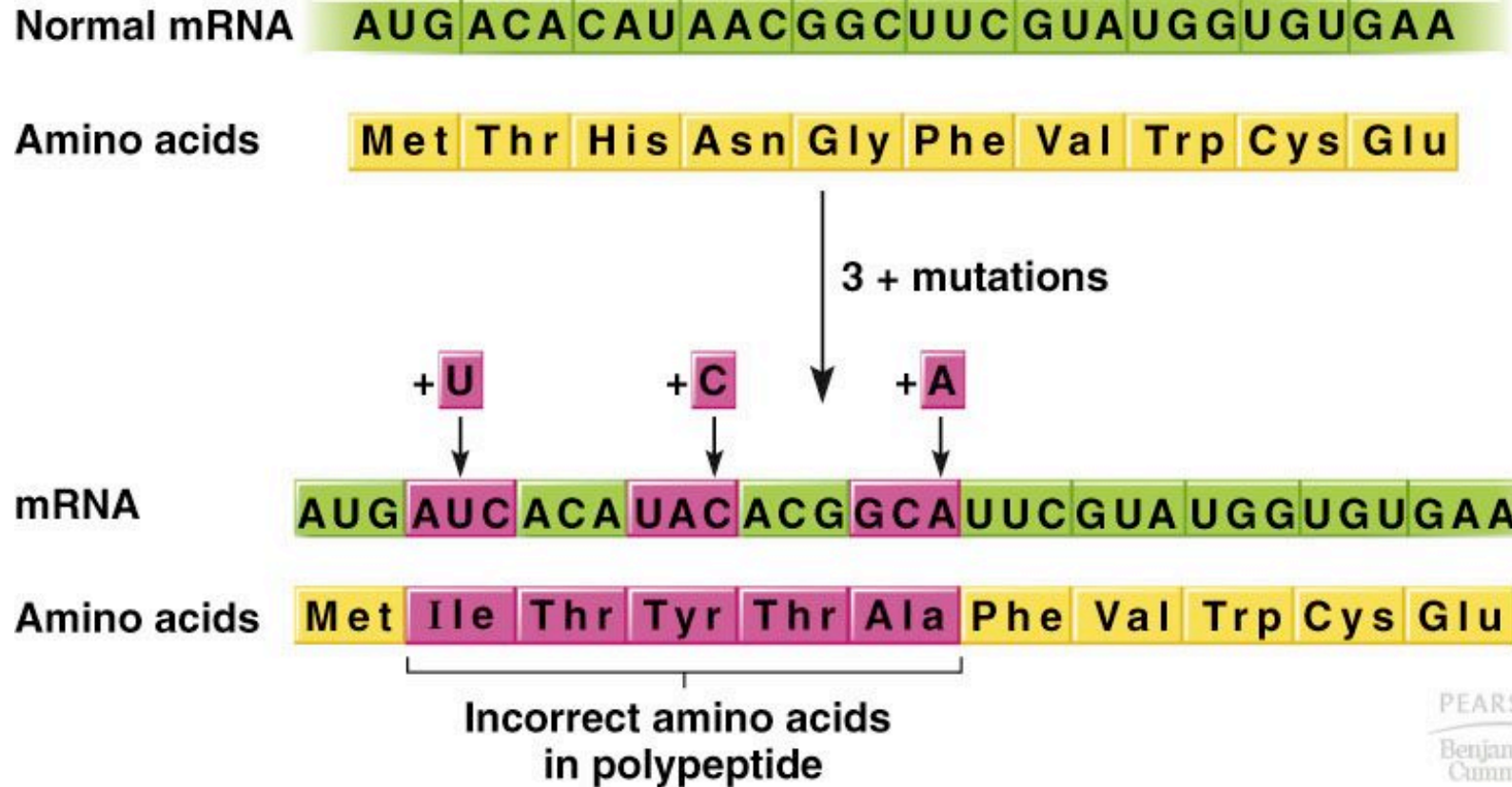
Gene-Protein colinearity



The Genetic Code is a triplet code

- Recombination between two different mutants in the same position resulted in wild-type E.Coli:
 - One codon consists of more than one nucleotide.
- The addition of 3 nucleotides results in 1 additional amino acid in the protein
- The deletion of 3 nucleotides results in 1 less amino acid in the protein

The Genetic Code is a triplet code



Most amino acids are encoded by multiple codons



- There are potentially 64 codons, and only 20 amino acids.
- If each amino acid is encoded by only one codon,
 - there would be 44 codons which would not code for any amino acid.
- This would imply that more than 50% of the time, a frame shift would result in a codon that would be a ‘nonsense’ codon.
 - This was contrary to experimental observations
- Conclusion: *the genetic code is degenerate.*
 - More than one codon can code for each amino acid.
 - Each amino acid can be coded for by more than one codon
 - Each codon only codes for one amino acid

Characteristics of the Genetic Code

- It is a triplet code.
 - Each three-nucleotide codon in the mRNA specifies one amino acid
- It is comma free.
 - mRNA is read three bases at a time without skipping any bases.
- It is non-overlapping/non-ambiguous.
 - Each nucleotide is part of only one codon and is read only once during translation.
- It is almost universal.
 - In nearly all organisms, most codons have the same amino acid meaning.
- It is degenerate.
 - Of 20 amino acids, 18 are encoded by 2 or more codons.
- The code has start and stop signals.
 - AUG is the usual start signal and defines the open reading frame.
- Stop signals are codons with no corresponding tRNA
 - the nonsense or chain-terminating codons.
 - generally three stop codons: UAG, UAA, and UGA.

Fig. 14.8
The genetic code

		Second letter				
		U	C	A	G	
U	U	UUU Phe (F)	UCU	UAU Tyr (Y)	UGU Cys (C)	U
		UUC	UCC Ser (S)	UAC	UGC (C)	C
		UUA Leu (L)	UCA	UAA Stop	UGA Stop	A
		UUG	UCG	UAG Stop	UGG Trp (W)	G
C	C	CUU	CCU	CAU His (H)	CGU	U
		CUC Leu (L)	CCC Pro (P)	CAC	CGC Arg (R)	C
		CUA	CCA	CAA Gln (Q)	CGA	A
		CUG	CCG	CAG	CGG	G
A	A	AUU	ACU	AAU Asn (N)	AGU Ser (S)	U
		AUC Ile (I)	ACC Thr (T)	AAC	AGC (S)	C
		AUA	ACA	AAA Lys (K)	AGA Arg (R)	A
		AUG Met (M)	ACG	AAG	AGG	G
G	G	GUU	GCU	GAU Asp (D)	GGU	U
		GUC Val (V)	GCC Ala (A)	GAC	GGC Gly (G)	C
		GUA	GCA	GAA Glu (E)	GGA	A
		GUG	GCG	GAG	GGG	G

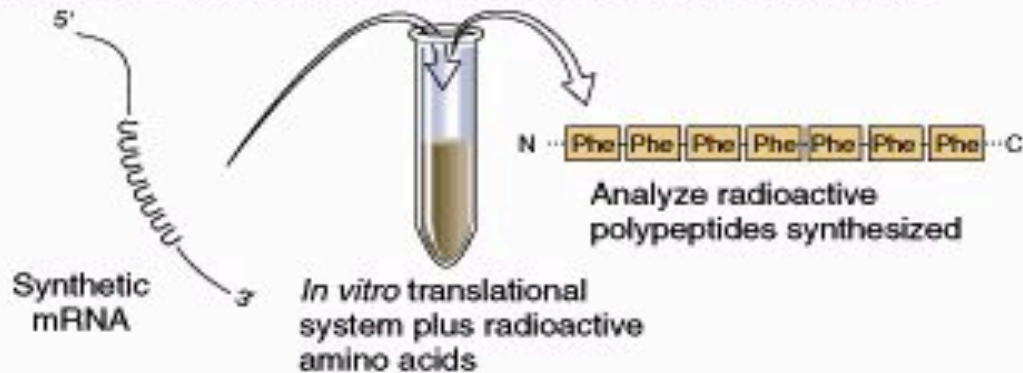
 = Chain termination codon (stop)
 = Initiation codon

Cracking the Genetic Code-I

- Discovery of mRNA as an intermediate molecule helped crack the genetic code.
- Scientists developed cell extracts which when mixed with mRNA, could synthesize polypeptide chains:
 - *in vitro* translation.
- Scientists were also able to synthesize short segments of mRNA of defined sequence.
- Using defined mRNA sequences, scientists began cracking the genetic code in a systematic fashion:
 - UUUUUU--- = Phe-Phe---,
 - UCUCUC--- = Ser-Leu---- , etc.
- Some ambiguities were resolved by further experimentation.

Cracking the Code

(a) A poly-U mRNA generates a polyphenylalanine polypeptide



(b) Analyzing the coding possibilities

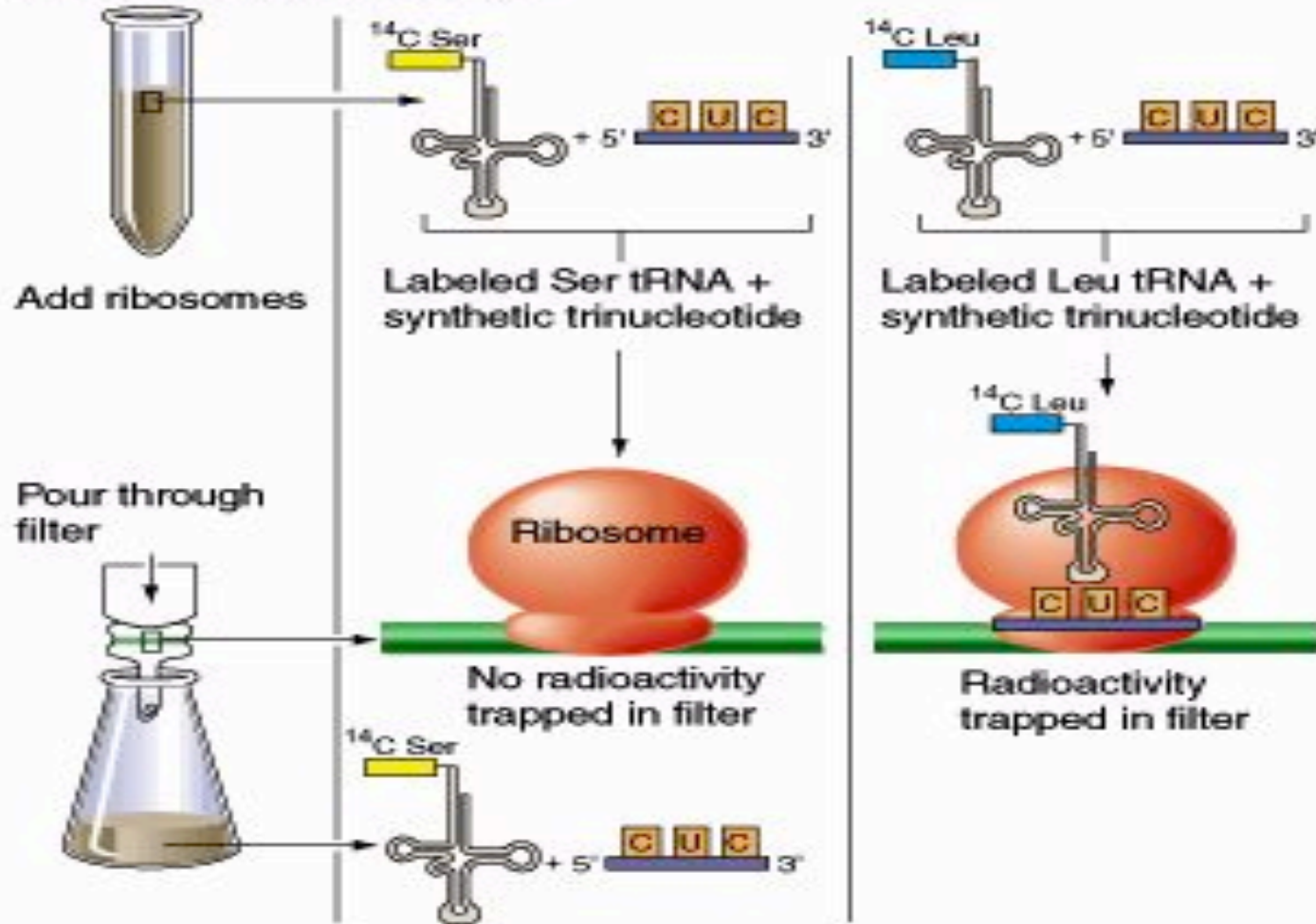
Synthetic mRNA	Polypeptides synthesized
Polypeptides with one amino acid	
poly-U UUUU ...	Phe-Phe-Phe ...
poly-C CCCC ...	Pro-Pro-Pro ...
poly-A AAAA ...	Lys-Lys-Lys ...
poly-G GGGG ...	Gly-Gly-Gly ...
Repeating dinucleotides	
poly-UC UCUC ...	Ser-Leu-Ser-Leu ...
poly-AG AGAG ...	Arg-Glu-Arg-Glu ...
poly-UG UGUG ...	Cys-Val-Cys-Val ...
poly-AC ACAC ...	Thr-His-Thr-His ...
Repeating trinucleotides	
poly-UUC UUCUUCUUC ...	Phe-Phe... and Ser-Ser... and Leu-Leu...
poly-AAG AAGAAGAAG ...	Lys-Lys... and Arg-Arg... and Glu-Glu...
poly-UUG UUGUUGUUG ...	Leu-Leu... and Cys-Cys... and Val-Val...
poly-UAC UACUACUAC ...	Tyr-Tyr... and Thr-Thr... and Leu-Leu...
Repeating tetranucleotides	
poly-UAUC UAUCUAUC ...	Tyr-Leu-Ser-Ile-Tyr-Leu-Ser-Ile...
poly-UUAC UUACUUAC ...	Leu-Leu-Thr-Tyr-Leu-Leu-Thr-Tyr...
poly-GUAA GUAAGUAA ...	none
poly-GAUA GAUAGAUA ...	none

Cracking the Genetic Code-II

- By 1965 it was known
 - that polypeptide chains were synthesized on ribosomes,
 - and that different tRNA molecules carry the appropriate amino acid to the ribosome
 - depending on the genetic codon on the mRNA.
- Using radiolabeled amino acids and defined mRNA sequences,
 - it was determined which codon corresponded to which amino acid.

Cracking the Code

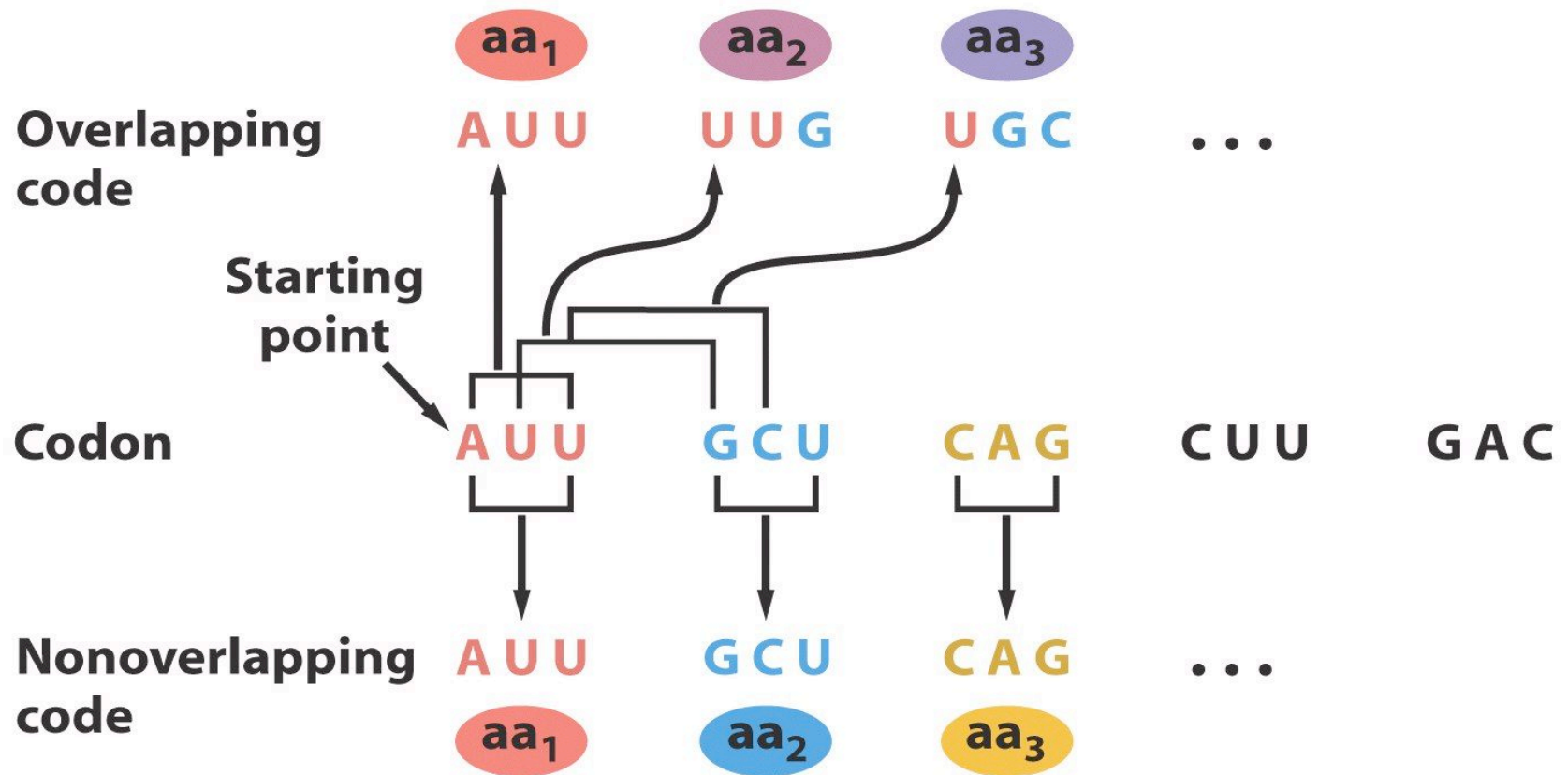
(c) Resolving the ambiguities



Cracking the Genetic Code-III

- Further experiments determined that the direction of 5' to 3' corresponded to the Amino-terminal to Carboxy-terminal direction in the polypeptide.
- The starting point is always defined by the codon **AUG**,
 - which codes for **methionine**.
- It was found that three codons: **UAA, UAG and UGA** did not code for any amino acid:
 - they were **nonsense** or **stop codons**.

The Genetic Code Is Non-Overlapping



Universality of the Genetic Code

- All living beings use the same genetic code.
- Genetic code evolved early in life, and has remained constant over billions of years because of the lack of tolerance for change.
- Some exceptions exist:
 - in some ciliates, there is only one nonsense codon.
- In eukaryotic mitochondria, there are some changes:
 - e.g.: in yeast, CUA codes for Thr instead of Leu.
- The impact of this change in these organisms/organelles may not be as drastic, since there are very few proteins encoded by these systems, allowing evolution of the genetic code.

Question

- A codon is:
 - 1. A 3 base sequence of mRNA that codes for an amino acid
 - 2. A 3 base sequence of rRNA that codes for an amino acid
 - 3. A 3 base sequence of tRNA that codes for an amino acid
 - 4. A 3 base sequence of DNA that codes for an amino acid

Question

- An anti-codon is:
 - 1. A 3 base sequence of mRNA that compliments the codon
 - 2. A 3 base sequence of tRNA that compliments the codon
 - 3. A 3 base sequence of tRNA that codes for an amino acid
 - 4. A 3 base sequence of DNA that codes for an amino acid

Question

- Which change does NOT cause a frameshift mutation?
 - 1) 1 basepair deletion
 - 2) 2 basepair insertion
 - 3) 2 basepair deletion
 - 4) 3 basepair insertion

Question

34. The sequence of a complete eukaryotic gene encoding the small protein Met Arg Val Tyr Ala is shown. All of the written sequences on the template strand are transcribed into RNA. (14 pts total)

strand A 5' ACGCATGCGGGTGTGCCCCAGGTCTACGCGTGAAAT 3'
 strand B 3' TGCGTACGCCACACGGGGTCCAGATGCGCACTTTA 5'

- Which strand (A or B) is the template strand? (2 pts)
- What is the sequence of the nucleotides in the **processed** mRNA molecule for this gene? Indicate 5' and 3' directions of this gene. (4 pts)

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

Homework Problems

Chapter 14

1, 8, 10, 14, 20, 21