10 From DNA to Protein: Gene Expression

Chapter Outline

In Chapter 10, we make the transition from stored information, which is generally in the form of DNA, to the retrieval of that information, usually by the synthesis of "effector" proteins (polypeptides) that change what is happening in and around the cells of living organisms. The overall sequence is:

$$\text{DNA} \xrightarrow{\text{transcription}} \text{RNA} \xrightarrow{\text{splicing}} \text{mRNA} \xrightarrow{\text{translation}} \text{Polypeptide}$$

Once the transcription of the DNA code to the RNA code is complete, RNA modification occurs, and the exon sequences of RNA get spliced back together, producing messenger RNA (mRNA). The mRNA departs from the nucleus and moves to the ribosomes, where it guides protein synthesis. Proteins, or polypeptides, are often modified within cells to become specific effector proteins. These modifications occur after translation has been completed.

Regulating which DNA sequences (i.e., genes) are transcribed to make RNA sequences is most directly controlled in the cells by signals called *transcription factors*. Each of us, male or female, has enough genetic information to make most of the male and female parts of the reproductive system. So why do most individuals have only male *or* female organs? We make only the reproductive parts we need because the expression of genetic information is closely regulated by our sex-specific transcription factors. In the case of reproductive development, hormones and other signals ensure that the correct genes are expressed.

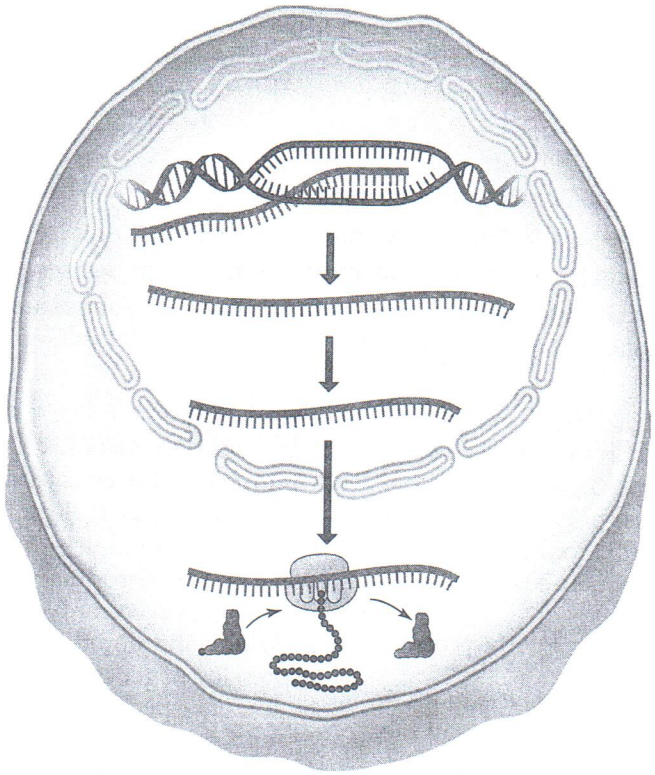
Chapter 10 expands on your understanding of **Big Idea 3**, concerning the nature of information transfer in living organisms. Specifically, Chapter 10 includes:

- **3.A.1:** DNA, and in some cases RNA, is the primary source of heritable information.

Chapter Review

Concept 10.1 *describes how genes code for proteins. Enzymes were an obvious and measurable category of proteins for early researchers, whose observations led to the hypothesis that genetically determined diseases are often based on mutations in the genes that code for enzymes. Subsequent work led to the one gene–one protein hypothesis. Today we understand this better as the one gene–one polypeptide relationship, because many functional proteins (e.g., insulin) are fragments of much larger polypeptides (e.g., preproinsulin).*

1. On the diagram below, label DNA, pre-mRNA, tRNA, ribosome, translation, polypeptide, and transcription.



2. Describe the structure and function of mRNA, rRNA, microRNA, and tRNA.

3. A dog was taken to a veterinary clinic because she had become increasingly lethargic. A blood test determined that the dog had very low levels of the steroid hormones produced by the adrenal glands, especially cortisol. The veterinarian's diagnosis was "a mutation in the cortisol gene." Explain why this characterization is inaccurate, given that steroid hormones are lipids. Speculate on a way that a genetic mutation could result in low levels of steroid hormones.

4. Describe the genetic condition of a person who has sickle-cell disease. Provide specific details about which gene is mutated, and include information about the polypeptide that has been altered by the mutation. Compare the effects of being heterozygous versus homozygous for the mutation.

Concept 10.2 *explains how gene expression begins with the transcription of DNA to make RNA. Transcription of DNA begins with the association of the enzyme RNA polymerase with the DNA template. The necessary "ingredients" for transcription include free nucleoside triphosphates (ATP, GTP, CTP, and UTP) that are incorporated into the RNA strand being built by RNA polymerase moving along the DNA sequence. Transcription follows this sequence: initiation → elongation → termination. As transcription proceeds, each base in the DNA template strand pairs with its complementary nucleoside phosphate, which is then incorporated into the RNA. Here is a short example of that complementarity:*

3′-T-C-A-A-G-T-5′ in DNA results in 5′-A-G-U-U-C-A-3′ in RNA

The start of transcription along a strand of DNA is determined by the presence of a promoter sequence in the DNA, and RNA polymerase binds here. The next part of the sequence specifies the "start" codon UGA in the RNA. Elongation of RNA proceeds until a "stop" codon (UAA or UAG) is specified. While still in the nucleus, pre-mRNA receives a GTP-cap on its 5′ end, and a poly-A tail on its 3′ end; these modifications enhance RNA stability and later facilitate mRNA exiting from the nucleus and binding to ribosomes to start protein synthesis.

5. Mutations can have many different effects on the outcome of a gene's expression.

   a. Describe two possible impacts on a gene's expression resulting from a mutation in the promoter region of that gene.

   b. Describe two possible impacts on a gene's expression resulting from a mutation in the stop codon of that gene.

6. Describe what happens to "intron" and "exon" sequences as pre-mRNA is processed.

7. Describe the functions of the GTP cap and the poly-A tail on mRNA.

Concept 10.3 *describes how the genetic code of RNA is translated, driving the specific sequence of amino acids that are incorporated into newly synthesized proteins. The ribosomes "translate" the mRNA code by using it to determine which amino acid gets placed at which position in newly synthesized proteins. More specifically, the mRNA code consists of sequences of triplets of ribonucleotides (three bases in length) called codons. Each mRNA codon determines which amino acid will be added onto the linear sequence of the amino-acid chain that is the backbone of the growing polypeptide.*

8. Use the table below to characterize the peptide product by translating the mRNA fragment shown below the table.

First letter	Second letter: U	Second letter: C	Second letter: A	Second letter: G	Third letter
U	UUU, UUC: Phenylalanine UUA, UUG: Leucine	UCU, UCC, UCA, UCG: Serine	UAU, UAC: Tyrosine UAA: Stop codon UAG: Stop codon	UGU, UGC: Cysteine UGA: Stop codon UGG: Tryptophan	U, C, A, G
C	CUU, CUC, CUA, CUG: Leucine	CCU, CCC, CCA, CCG: Proline	CAU, CAC: Histidine CAA, CAG: Glutamine	CGU, CGC, CGA, CGG: Arginine	U, C, A, G
A	AUU, AUC, AUA: Isoleucine AUG: Methionine; start codon	ACU, ACC, ACA, ACG: Threonine	AAU, AAC: Asparagine AAA, AAG: Lysine	AGU, AGC: Serine AGA, AGG: Arginine	U, C, A, G
G	GUU, GUC, GUA, GUG: Valine	GCU, GCC, GCA, GCG: Alanine	GAU, GAC: Aspartic acid GAA, GAG: Glutamic acid	GGU, GGC, GGA, GGG: Glycine	U, C, A, G

5′ AUG UUU CAG CGA GGA UGA 3′

9. Describe how the result would change if the above sequence were altered so that the sixth base from the 5′ end was switched to G.

10. Describe how the result would change if the above sequence were altered so that 19th base from the 5′ end was switched to A.

11. Even though the amino acid serine is specified by UCU, UCC, UCA, UCG, AGU, and AGC, is it said that the genetic code is not ambiguous. Explain.

12. Explain the comment "the genetic code is nearly universal" in terms of evolutionary ancestry.

13. Explain how the human gene for insulin can be inserted into an *E. coli* bacterium to produce human insulin for medicinal purposes.

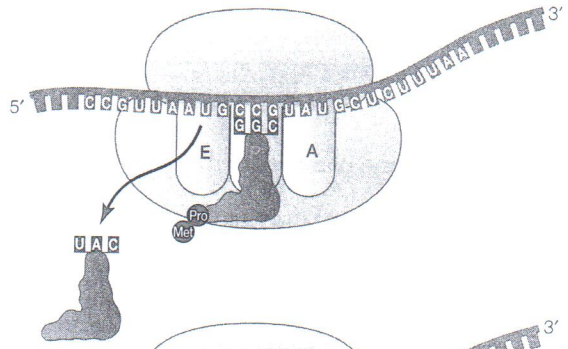
14. What distinguishes the impacts of silent mutations from those of nonsense mutations?

Concept 10.4 *describes that the translation of the genetic code occurs when transfer RNA (tRNA) molecules deliver amino acids to ribosomes. Amino acids are carried on specialized RNA carriers known as tRNA. Each tRNA molecule includes a triplet sequence of RNA that is complementary to, and will bind with, a particular codon of mRNA, triggering the delivery of the tRNA's amino-acid cargo to the growing polypeptide. After that delivery is complete, the remnants of the tRNA leave the ribosome, and the mRNA code is then moved along to the next codon, which specifies the next tRNA to bind and leave its amino acid for the growing polypeptide.*

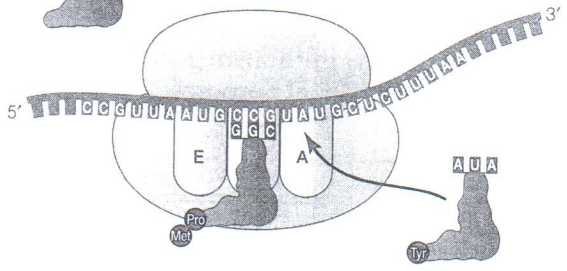
15. Explain the relationship between codons and anticodons, and describe where they interact.

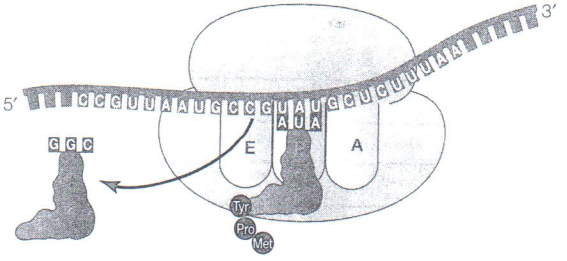
16. Describe how more than one ribosome can be active in the process of translating a piece of mRNA.

17. For the ribosome shown at the right, state the anticodon sequence for the next tRNA to bind (at the bottom of the figure), then use the table in Question 8 to determine which amino acid will be incorporated into the peptide next.



18. Compare the actions of RNA polymerase and ribosomes.



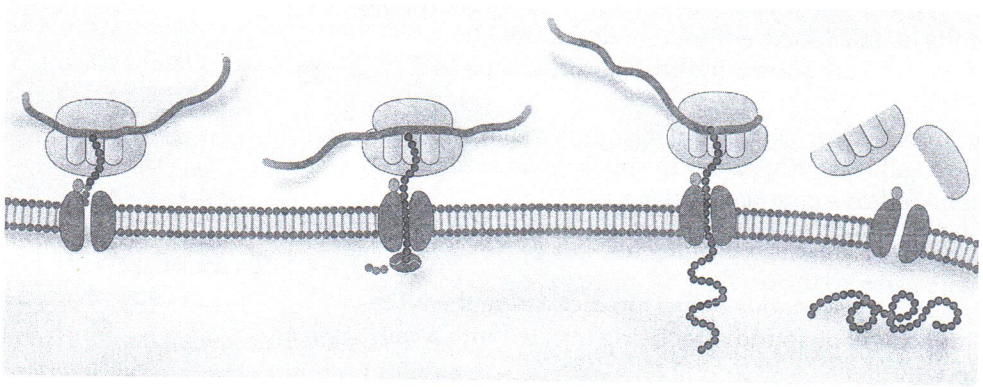


Concept 10.5 *describes how proteins are modified after translation. When the mRNA has been translated up to a "stop" codon (UAA, UAG, or UGA), the polypeptide is released from the ribosome and either used as is or modified by other parts of the cell. Three major categories of posttranslational modification can take place. First, proteolytic enzymes could cut a long polypeptide into smaller proteins. Second, carbohydrates might be added onto the polypeptide, especially in proteins that play a role in cellular identity. Third, phosphorylation could covalently modify the protein, thus altering its shape and function.*

19. Secreted proteins, including insulin and other hormones, typically interact with at least two membrane-bound organelles prior to their secretion from the source cell. Describe the activities of these two organelles in terms of preparing the polypeptide for secretion out of the source cell.

20. Describe the possible roles of RNA processing and posttranslational modification in explaining the observation that human insulin, with only 51 amino acids, is in fact a product of the INS gene, which has nearly 40,000 base pairs.

21. On the diagram below, label RER, phospholipid bilayer, mRNA, polypeptide, and ribosome.



Science Practices & Inquiry

In the AP Biology Curriculum Framework, there are seven **Science Practices**. In this chapter, we focus on **Science Practice 1:** The student can use representations and models to communicate scientific phenomena and solve scientific problems. More specifically, we focus on **Practice 1.2:** The student can describe representations and models of natural or man-made phenomena and systems in the domain. We also cover **Science Practice 6:** The student can work with scientific explanations and theories. More specifically, we focus on **Practice 6.4:** The student can make claims and predictions about natural phenomena based on scientific theories and models.

Question 22 asks you to describe representations and models illustrating how genetic information is translated into polypeptides (**Learning Objective 3.4**), while question 23 asks you to predict how a change in a specific DNA or RNA sequence can result in changes in gene expression (**Learning Objective 3.6**).

22. Researchers have determined that a short-chain polypeptide signal (nuclear localization signal, or NLS) comprised of eight amino acids must be part of the central sequence of a protein if it is to enter the nucleus. When the NLS-protein complex docks with a pore in the nuclear membrane, the signal causes the pore to open. The sequence for this NLS peptide is: -Pro-Pro-Lys-Lys-Lys-Arg-Lys-Val-.

a. Explain why methionine could not be part of NLS.

b. Write out one mRNA strand that could produce this sequence.

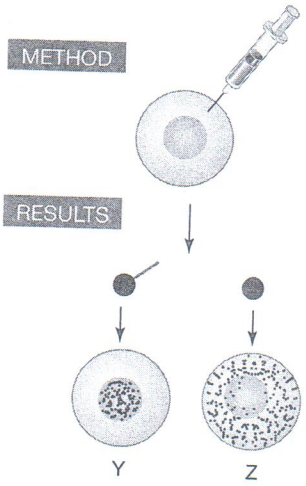
c. Write out the DNA sequence that this mRNA sequence comes from.

d. Is this the only possible DNA sequence that could code for this NLS? Explain your answer.

23. The results of an experiment used to determine the function of the NLS are shown in the diagram at the right.

In scenario Y, an NLS-protein–red-dye complex that was injected into a cell is later found in the nucleus. In scenario Z, the protein–red-dye complex, lacking the NLS peptide, is injected into the cell, and it is later found only in the cytoplasm.

For each of the three scenarios below, predict where the NLS-protein complex will be found after being injected into a cell, and explain why.



a. A cytosolic protein, normally found only in the cytoplasm, is bound to an NLS.

b. A nuclear protein is attached to a mutated form of an NLS that is missing its final valine.

c. A nuclear protein is attached to an NLS that was produced from the DNA sequence TAC-GGG-GGT-TTT-TTC-TTC-GCT-TAC-CAC-stop.