

Chapter 17

From Gene to Protein

Key Concepts

- 17.1 Genes specify proteins via transcription and translation
- 17.2 Transcription is the DNA-directed synthesis of RNA: *a closer look*
- 17.3 Eukaryotic cells modify RNA after transcription
- 17.4 Translation is the RNA-directed synthesis of a polypeptide: *a closer look*
- 17.5 RNA plays multiple roles in the cell: *a review*
- 17.6 Comparing gene expression in prokaryotes and eukaryotes reveals key differences
- 17.7 Point mutations can affect protein structure and function

Framework

This chapter deals with the pathway from DNA to RNA to proteins. The instructions of DNA are transcribed to a sequence of codons in mRNA. In eukaryotes, mRNA is processed before it leaves the nucleus. Complexed with ribosomes, mRNA is translated as a sequence of amino acids in a polypeptide as tRNAs match their anticodons to the mRNA codons. Mutations, which alter the base pairs in DNA, usually alter the protein product.

Chapter Review

The information of DNA is contained in specific sequences of nucleotides. These sequences are transcribed and translated to synthesize proteins, which are

responsible for the specific traits of an organism and thus form the link between genotype and phenotype.

17.1 Genes specify proteins via transcription and translation

Evidence from the Study of Metabolic Defects In 1909, A. Garrod first suggested that genes determine phenotype through the action of enzymes, reasoning that inherited diseases were caused by an inability to make certain enzymes.

G. Beadle and E. Tatum, working with mutants of a bread mold, *Neurospora crassa*, demonstrated the relationship between genes and enzymes. They studied several nutritional mutants that could not grow on the minimal medium that sufficed for wild-type mold. By growing nutritional mutants on complete growth medium and then transferring samples to various combinations of minimal medium and one added nutrient, Beadle and Tatum were able to identify the specific metabolic defect for each mutant.

Three classes of *Neurospora* mutants that were unable to synthesize arginine were identified by supplementing different precursors of the pathway. Beadle and Tatum reasoned that the metabolic pathway of each class was blocked at a different enzymatic step. They formulated the *one gene–one enzyme hypothesis*.

Biologists revised Beadle and Tatum's idea to one gene–one protein because not all proteins are enzymes. Because many proteins consist of more than one polypeptide chain, each of which is codified by its own gene, the axiom is now the **one gene–one polypeptide hypothesis**.

Basic Principles of Transcription and Translation

RNA is the link between a gene and the protein for which it codes. RNA differs from DNA in three ways: the sugar component of its nucleotides is ribose, rather than deoxyribose; uracil (U) replaces thymine as one of its nitrogenous bases; and RNA is usually single stranded.

Transcription is the transfer of information from DNA to **messenger RNA (mRNA)** or another type of

RNA, using the “language” of nucleic acids. **Translation** transfers information from mRNA to a polypeptide, changing from the language of nucleotides to that of amino acids. **Ribosomes** are the sites of translation—the synthesis of a polypeptide.

In prokaryotes, which lack a nucleus, transcription and translation can occur simultaneously. In eukaryotes, the mRNA must exit the nucleus before translation can begin. **RNA processing**, the modification of *pre-mRNA* or of the **primary transcript** of any gene (such as one coding for RNA) within the nucleus, occurs only in eukaryotes.

■ INTERACTIVE QUESTION 17.1

Fill in the sequence in the synthesis of proteins. Put the name of the process above each arrow.

_____ → _____ → _____

The Genetic Code A sequence of three nucleotides provides 4^3 , or 64, possible unique sequences of nucleotides, more than enough to code for the 20 amino acids. The translation of nucleotides into amino acids uses a **triplet code** to specify each amino acid.

The base triplets along the **template strand** of a gene are transcribed into mRNA **codons**. The same strand of a DNA molecule can be the template strand for one gene and the complementary strand for another. The mRNA is complementary to the DNA template since its bases follow the same base-pairing rules, with the exception that uracil substitutes for thymine in RNA. Codons are read in the $5' \rightarrow 3'$ direction. Each codon specifies one of the 20 amino acids.

In the early 1960s, M. Nirenberg added artificial “poly U” mRNA to a test tube containing all the biochemical ingredients necessary for protein synthesis and obtained a polypeptide containing a single amino acid. Molecular biologists had deciphered all 64 codons by the mid-1960s. Three codons function as stop signals, or termination codons. The codon AUG both codes for methionine and functions as an initiation codon, a start signal for translation.

The code is often redundant, meaning that more than one codon may specify a single amino acid. The code is never ambiguous; no codon specifies two different amino acids.

The nucleotide sequence on mRNA is read in the correct **reading frame**, starting at a start codon and reading each triplet sequentially.

■ INTERACTIVE QUESTION 17.2

Practice using the dictionary of the genetic code in your textbook. Determine the amino acid sequence for a polypeptide coded for by the following mRNA transcript (written $5' \rightarrow 3'$):

AUGCCUGACUUUAAGUAG

The genetic code of codons and their corresponding amino acids is almost universal. A bacterial cell can translate the genetic messages of human cells. The near universality of a common genetic language provides compelling evidence of the antiquity of the code and the evolutionary connection of all living organisms.

17.2 Transcription is the DNA-directed synthesis of RNA: a closer look

Molecular Components of Transcription The **promoter** is the DNA sequence where **RNA polymerase** attaches and initiates transcription. In prokaryotes, the **terminator** is the sequence that signals the end of transcription. A **transcription unit** is the sequence of DNA that is transcribed into one RNA molecule.

Bacteria have one type of RNA polymerase. Eukaryotes have three types; the one that synthesizes mRNA is called RNA polymerase II.

Synthesis of an RNA Transcript The promoter is the binding site for RNA polymerase and determines where transcription starts and which DNA strand is used as the template. It includes recognition sequences, such as the **TATA box** common in eukaryotes, upstream from the start point. In eukaryotes, **transcription factors** must first recognize and bind to the promoter before RNA polymerase II can attach, at which point the assembly is called the **transcription initiation complex**.

RNA polymerase untwists the double helix, exposing DNA nucleotides for base pairing with RNA nucleotides, and joins the nucleotides to the $3'$ end of the growing polymer. The new RNA peels away from the DNA template and the DNA double helix re-forms. Several molecules of RNA polymerase may be transcribing simultaneously from different parts of a single gene, enabling a cell to produce large quantities of mRNA.

In prokaryotes, transcription ends after RNA polymerase transcribes the terminator sequence. In eukaryotes, polymerase continues past a polyadenylation

■ INTERACTIVE QUESTION 17.5

Using some of the codons and the amino acids you identified in Interactive Question 17.2, fill in the following table.

DNA Triplet 3' → 5'	mRNA Codon 5' → 3'	Anticodon 3' → 5'	Amino Acid
			methionine
		GGA	
TTC			
	UAG		

Each amino acid has a specific **aminoacyl-tRNA synthetase** that attaches it to its appropriate tRNA molecule to create an aminoacyl tRNA. The hydrolysis of ATP drives this process.

Ribosomes facilitate the specific pairing of tRNA anticodons with mRNA codons during protein synthesis. They consist of a large and a small subunit, each composed of proteins and a form of RNA called **ribosomal RNA (rRNA)**. Subunits are constructed in the nucleolus in eukaryotes. Prokaryotic ribosomes are smaller and differ enough in molecular composition that some antibiotics can inhibit them without affecting eukaryotic ribosomes.

A large and small subunit join to form a ribosome when they attach to an mRNA molecule. Ribosomes have a binding site for mRNA, a **P site** (peptidyl-tRNA site) that holds the tRNA carrying the growing polypeptide chain, an **A site** (aminoacyl-tRNA site) that holds the tRNA carrying the next amino acid, and an **E site** (exit site) from which discharged tRNAs leave the ribosome. The attachment of an amino acid to the carboxyl end of the growing polypeptide chain is catalyzed by the ribosome.

According to the recently determined detailed structure of the small and large subunits of bacterial ribosomes, the protein components are mostly on the outside and the rRNA is in the interior at the interface

between the subunits and at the A and P sites, supporting the hypothesis that rRNA performs a ribosome's catalytic functions.

Building a Polypeptide The three stages of protein synthesis—initiation, elongation, and termination—all require the aid of protein “factors.” The first two stages also require energy, which is provided by the hydrolysis of GTP (guanosine triphosphate).

The initiation stage begins as the small subunit of the ribosome binds to an mRNA and an initiator tRNA carrying methionine, which attaches to the start codon AUG on the mRNA. With the aid of proteins called *initiation factors* and the expenditure of a GTP, the large subunit of the ribosome attaches to the small one, forming a translation initiation complex. The initiator tRNA fits into the P site.

The addition of amino acids in the elongation stage involves several proteins called *elongation factors* and occurs in a three-step cycle.

In codon recognition, an aminoacyl-tRNA base-pairs with the mRNA codon in the A binding site. This step requires energy from the hydrolysis of 2 GTP.

In the peptide bond formation step, an RNA molecule of the large subunit catalyzes the formation of a peptide bond between the carboxyl end of the polypeptide held in the P site and the amino acid in the A site. The polypeptide is now held by the amino acid in the A site.

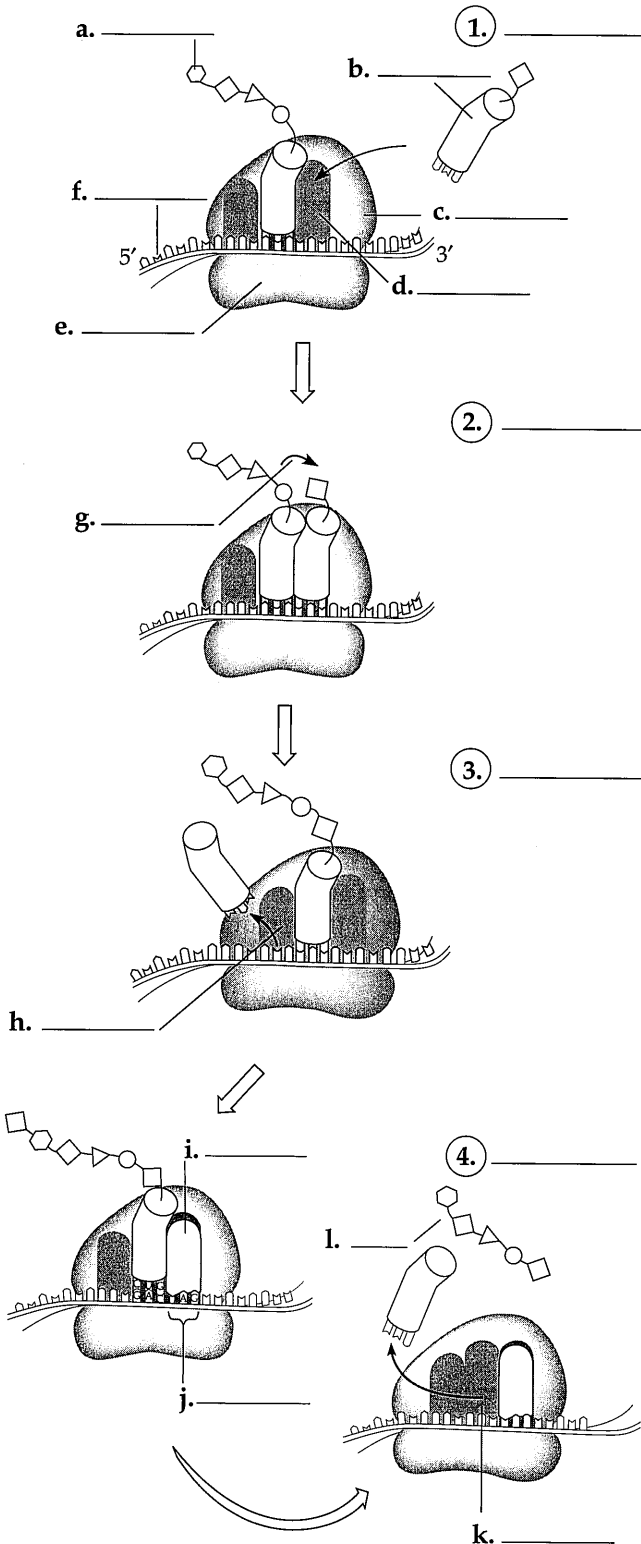
In translocation, tRNA carrying the growing polypeptide is translocated to the P site, a process requiring energy from the hydrolysis of a GTP molecule. The empty tRNA from the P site is moved to the E site and released. The next mRNA codon moves into the A site as the mRNA moves through the ribosome.

Termination occurs when a stop codon—UAA, UAG, or UGA—reaches the A site of the ribosome. A *release factor* binds to the stop codon and hydrolyzes the bond between the polypeptide and the tRNA in the P site. The completed polypeptide leaves through the exit tunnel of the large subunit. The two ribosomal subunits and other components dissociate.

An mRNA may be translated simultaneously by several ribosomes in strings called **polyribosomes** (or **polysomes**).

INTERACTIVE QUESTION 17.6

In the following diagrams of polypeptide synthesis, name the stages (1–4), identify the components (a–l), and then briefly describe what happens in each stage.



Completing and Targeting the Functional Protein
 During and following translation, a polypeptide folds spontaneously into its secondary and tertiary structures. Chaperone proteins often facilitate the correct folding. The protein may need to undergo *post-translational modifications*: Amino acids may be chemically modified; one or more amino acids at the beginning of the chain may be enzymatically removed; segments of the polypeptide may be excised; or several polypeptides may associate into a quaternary structure.

All ribosomes are identical, whether they are free ribosomes that synthesize cytosolic proteins or ER-bound ribosomes that make membrane and secretory proteins. Polypeptide synthesis begins in the cytoplasm. If a protein is destined for the endomembrane system or for secretion, its polypeptide chain will begin with a **signal peptide** that is recognized by a protein-RNA complex called a **signal-recognition particle (SRP)**, which attaches the ribosome to a receptor protein that is part of a multiprotein complex on the ER membrane. As the growing polypeptide threads into the ER, the signal peptide is usually removed. Other signal peptides direct some proteins made in the cytosol to specific sites such as mitochondria, chloroplasts, or the interior of the nucleus.

INTERACTIVE QUESTION 17.7

What determines if a ribosome becomes bound to the ER?

17.5 RNA plays multiple roles in the cell: a review

The versatility of RNA molecules stems from their ability to hydrogen-bond to other RNA or DNA molecules, to form specific three-dimensional shapes by forming intramolecular hydrogen bonds, and to act as a catalyst.

■ INTERACTIVE QUESTION 17.8

Fill in the functions for the following types of RNA molecules.

- a. mRNA
 - b. tRNA
 - c. rRNA
 - d. snRNA
 - e. SRP RNA
 - f. snoRNA (small nucleolar RNA)
 - g. siRNA (small interfering RNA) and miRNA (microRNA)
-

17.6 Comparing gene expression in prokaryotes and eukaryotes reveals key differences

The basic processes of transcription and translation are the same in prokaryotes and eukaryotes, although the RNA polymerases and ribosomes differ. Transcription factors are required in eukaryotes for RNA polymerase to bind, and transcription is terminated differently. In bacteria, transcription and translation occur almost simultaneously, whereas in eukaryotes, transcription is physically separated from translation by the nuclear envelope, allowing extensive RNA processing to occur before RNA leaves the nucleus. Eukaryotes also can target proteins to particular organelles.

17.7 Point mutations can affect protein structure and function

Mutations, changes in the genetic information of a cell (or virus) may involve large portions of a chromosome or affect just one base pair of nucleotides, as in a **point mutation**. If the mutation is in a cell that gives rise to a gamete, it may be passed on to offspring.

Types of Point Mutations A **base-pair substitution** replaces one nucleotide and its complementary partner with another pair of nucleotides. Due to the redundancy of the genetic code, some base-pair substitutions in the third nucleotide of a codon do not change the amino acid translation and are called *silent mutations*. A substitution may result in the insertion of a different amino acid without altering the character of the protein if the new amino acid has similar properties or is not located in a region crucial to that protein's function.

A base-pair substitution that results in a different amino acid in a critical portion of a protein, such as the active site of an enzyme, may significantly impair protein function.

A substitution that results in an incorrectly coded amino acid is called a **missense mutation**. **Nonsense mutations** occur when the point mutation changes an amino acid codon into a stop codon, prematurely halting the translation of the polypeptide chain and usually creating a nonfunctional protein.

Base-pair **insertions** or **deletions** that are not in multiples of three nucleotides alter the reading frame. All nucleotides downstream from the mutation will be improperly grouped into codons, creating extensive missense and usually prematurely ending in nonsense. These **frameshift mutations** almost always produce nonfunctional proteins.

Mutagens Mutations can occur in a number of ways. *Spontaneous mutations* include base-pair substitutions, insertions, deletions, and longer mutations that occur during DNA replication, repair, or recombination. Physical agents, such as X-rays and UV light, and various chemical agents that cause mutations are called **mutagens**. Chemical mutagens include **base analogs** that substitute for normal bases and then pair incorrectly in DNA synthesis, chemicals that insert into and distort the double helix, and other agents that chemically change DNA bases. Tests can measure the mutagenic effects of chemicals and thus their potential carcinogenic risk.

■ INTERACTIVE QUESTION 17.9

Define the following, and explain what type of point mutation could cause each of these mutations.

- a. silent mutation
 - b. missense mutation
 - c. nonsense mutation
 - d. frameshift mutation
-

What is a gene? Revisiting the question

Our definition of a gene has evolved from Mendel's heritable factors, to Morgan's loci along chromosomes, to the one gene–one polypeptide axiom.

Research continually refines our understanding of the structural and functional aspects of genes, which now include introns, promoters, and other regulatory regions. The best working definition of a gene is that it is a region of DNA whose final product is either a polypeptide or an RNA molecule.

Word Roots

- anti-** = opposite (*anticodon*: a specialized base triplet on one end of a tRNA molecule that recognizes a particular complementary codon on an mRNA molecule)
- exo-** = out, outside, without (*exon*: a coding region of a eukaryotic gene that is expressed)
- intro-** = within (*intron*: a noncoding, intervening sequence within a eukaryotic gene)
- muta-** = change; **-gen** = producing (*mutagen*: a physical or chemical agent that causes mutations)

poly- = many (*poly-A tail*: the modified end of the 3' end of an mRNA molecule consisting of the addition of some 50 to 250 adenine nucleotides)

trans- = across; **-script** = write (*transcription*: the synthesis of RNA on a DNA template)

Structure Your Knowledge

1. Make sure you understand and can explain the processes of transcription and translation. You may find that filling in the table below (for eukaryotic gene expression) in a study group helps you to review these processes.
2. What is the genetic code? Explain redundancy and the wobble phenomenon. What is the significance of the fact that the genetic code is nearly universal?
3. Prepare a concept map showing the types and consequences of point mutations.

	Transcription	Translation
Template		
Location		
Molecules involved		
Enzymes involved		
Control—start and stop		
Product		
Product processing		
Energy source		

Test Your Knowledge

MULTIPLE CHOICE: Choose the one best answer.

1. In Beadle and Tatum's study of *Neurospora*, they were able to identify three classes of mutants that needed arginine added to minimal media in order to grow. The production of arginine includes the following steps: precursor → ornithine → citrulline → arginine. What nutrient(s) had to be supplied for the mutants with a defective enzyme for the precursor → ornithine step to grow?
 - a. precursor only
 - b. ornithine only
 - c. citrulline only
 - d. ornithine or citrulline
 - e. precursor, ornithine, and citrulline
2. Transcription involves the transfer of information from
 - a. DNA to RNA.
 - b. RNA to DNA.
 - c. mRNA to an amino acid sequence.
 - d. DNA to an amino acid sequence.
 - e. the nucleus to the cytoplasm.
3. If the 5' → 3' nucleotide sequence on the complementary (noncoding) DNA strand is CAT, what is the corresponding codon on mRNA?
 - a. UAC
 - b. CAU
 - c. GUA
 - d. GTA
 - e. CAT

4. RNA polymerase
 - a. is the protein responsible for the production of ribonucleotides.
 - b. is the enzyme that creates hydrogen bonds between nucleotides on the DNA template strand and their complementary RNA nucleotides.
 - c. is the enzyme that transcribes exons but does not transcribe introns.
 - d. is a ribozyme composed of snRNPs.
 - e. begins transcription at a promoter sequence and moves along the template strand of DNA, elongating an RNA molecule in a 5' → 3' direction.
5. How is the template strand for a particular gene determined?
 - a. It is the DNA strand that runs from the 5' → 3' direction.
 - b. It is the DNA strand that runs from the 3' → 5' direction.
 - c. It depends on the orientation of RNA polymerase, whose position is determined by particular sequences of nucleotides within the promoter.
 - d. It doesn't matter which strand is the template because they are complementary and will produce the same mRNA.
 - e. The template strand always contains the TATA box.
6. Which enzyme synthesizes tRNA?
 - a. RNA replicase
 - b. RNA polymerase
 - c. aminoacyl-tRNA synthetase
 - d. ribosomal enzymes
 - e. ribozymes
7. Which of the following is true of RNA processing?
 - a. Exons are excised before the mRNA is translated.
 - b. The RNA transcript that leaves the nucleus may be much longer than the original transcript.
 - c. Assemblies of protein and snRNPs, called spliceosomes, may catalyze splicing.
 - d. Large quantities of rRNA are assembled into ribosomes.
 - e. Signal peptides are added to the 5' end of the transcript.
8. Which of the following is *not* involved in the formation of a eukaryotic transcription initiation complex?
 - a. TATA box
 - b. transcription factors
 - c. snRNA
 - d. RNA polymerase II
 - e. promoter
9. A prokaryotic gene 600 nucleotides long can code for a polypeptide chain of how many amino acids (at most)?

a. 100	c. 300	e. 1,800
b. 200	d. 600	
10. All of the following are transcribed from DNA *except*

a. exons.	c. tRNA.	e. promoter.
b. introns.	d. rRNA.	
11. What might introns have to do with the evolution of new proteins?
 - a. The excised introns are transcribed and translated as new proteins by themselves.
 - b. Introns are more likely to accumulate mutations than exons, and these mutations then result in the production of novel proteins.
 - c. Introns that are self-excising may also function as hydrolytic enzymes for other processes.
 - d. Introns provide more area where crossing over may occur (without interfering with the coding sequences) and thus increase the probability of exon shuffling between alleles.
 - e. Introns often correspond to domains in proteins that fold independently and have specific functions. Changing domains between nonallelic genes could produce novel proteins.
12. A ribozyme is
 - a. an exception to the one gene–one RNA molecule axiom.
 - b. an enzyme that adds the 5' cap and poly-A tail to mRNA.
 - c. an example of rearrangement of protein domains caused by RNA splicing.
 - d. an RNA molecule that functions as an enzyme.
 - e. an enzyme that produces both small and large ribosomal subunits.
13. All of the following would be found in a prokaryotic cell *except*

a. mRNA.
b. rRNA.
c. simultaneous transcription and translation.
d. snRNA.
e. RNA polymerase.
14. Which of the following is transcribed and then translated to form a protein product?
 - a. gene for tRNA
 - b. intron
 - c. gene for a transcription factor
 - d. 5' and 3' UTRs
 - e. gene for rRNA

15. Transfer RNA
- forms hydrogen bonds between its codon and the anticodon of an mRNA in the A site of a ribosome.
 - binds to its specific amino acid in the active site of an aminoacyl-tRNA synthetase.
 - uses GTP as the energy source to bind its amino acid.
 - is translated from mRNA.
 - is produced in the nucleolus.
16. Place the following events in the synthesis of a polypeptide in the proper order.
- A peptide bond forms.
 - An aminoacyl tRNA matches its anticodon to the codon in the A site.
 - A tRNA translocates from the A to the P site, and an unattached tRNA leaves the ribosome from the E site.
 - The large subunit attaches to the small subunit, with the initiator tRNA in the P site.
 - A small subunit binds to an mRNA and an initiator tRNA.
- 4-5-3-2-1
 - 4-5-2-1-3
 - 5-4-3-2-1
 - 5-4-1-2-3
 - 5-4-2-1-3
17. Translocation in the process of translation involves
- the hydrolysis of a GTP molecule.
 - the movement of the tRNA in the A site to the P site.
 - the movement of the mRNA strand one triplet length.
 - the release of the unattached tRNA from the E site.
 - all of the above.
18. Which of the following type of molecule catalyzes the formation of a peptide bond?
- RNA polymerase
 - rRNA
 - mRNA
 - aminoacyl-tRNA synthetase
 - protein ribosomal enzyme
19. Which of the following is *not* true of an anticodon?
- It consists of three nucleotides.
 - It lines up in the 5' → 3' direction along the 5' → 3' mRNA strand.
 - It extends from one loop of a tRNA molecule.
 - It may pair with more than one codon.
 - Its base uracil base-pairs with adenine.
20. Changes in a polypeptide following translation may involve
- the addition of sugars or lipids to certain amino acids.
 - the action of enzymes to add amino acids at the beginning of the chain.
 - the removal of poly-A from the end of the chain.
 - the addition of a 5' cap of a modified guanosine residue.
 - all of the above.
21. Several proteins may be produced at the same time from a single mRNA by
- the action of several ribosomes in a string, called a polyribosome.
 - several RNA polymerase molecules working sequentially.
 - signal peptides that associate ribosomes with rough ER.
 - containing several promoter regions.
 - the involvement of multiple spliceosomes.
22. A signal peptide
- is most likely to be found on cytosolic proteins produced by bacterial cells.
 - directs an mRNA molecule into the cisternal space of the ER.
 - is a sign to help bind the small ribosomal unit at the initiation codon.
 - would be the first 20 or so amino acids of a protein destined for a membrane location or secretion from the cell.
 - is part of the 5' cap.
23. A base deletion early in the coding sequence of a gene may result in
- a nonsense mutation.
 - a frameshift mutation.
 - multiple missense mutations.
 - a nonfunctional protein.
 - all of the above.
24. Base-pair substitutions may have little effect on the resulting protein for all of the following reasons *except* which one?
- The redundancy of the code may result in a silent mutation.
 - The substitution must involve three nucleotide pairs, otherwise the reading frame is altered.
 - The missense mutation may not occur in a critical part of the protein.
 - The new amino acid may have similar properties to the replaced one.
 - The wobble phenomenon could result in no change in translation.