Chapter 14

Gene Expression: From Gene to Protein

Chapter Focus

This chapter examines 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 nucleotide pairs in DNA, may alter the protein product. A gene may be defined as a DNA sequence whose final product is either a polypeptide or an RNA molecule.

Chapter Review

The DNA-directed synthesis of proteins (or sometimes just RNA) is called **gene expression**.

14.1 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 treated cells of the haploid bread mold, *Neurospora crassa*, with X-rays to produce "nutritional mutants" that were unable to grow on *minimal media*. By transferring samples of these mutants growing on *complete growth medium* to combinations of minimal medium and one added nutrient, they were able to identify the specific metabolic defect for each mutant. Experiments that supplemented different precursors of a particular metabolic pathway indicated that different mutants were blocked at a different enzymatic step in that pathway. Such research supported Beadle and Tatum's *one gene–one enzyme hypothesis*.

Biologists revised this hypothesis to one gene–one protein. Because many proteins consist of more than one polypeptide chain, the axiom changed to *one gene–one polypeptide*. (Exceptions include genes that code for more than one protein due to alternative splicing and genes that code for RNA molecules.)

Basic Principles of Transcription and Translation RNA is the link between a gene and the protein for which it codes.

Transcription is the transfer of information from DNA to **messenger RNA (mRNA)**. **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 sites of the synthesis of a polypeptide.

In prokaryotes, which lack a nucleus, transcription and translation can occur simultaneously. In eukaryotes, the mRNA, called *pre-mRNA*, is processed before it exits the nucleus and enters the cytoplasm, where translation occurs. The initial RNA transcript of any gene is called the **primary transcript**.

FOCUS QUESTION 14.1

a. In what three ways does RNA differ from DNA?

b. Fill in the following sequence in the flow of genetic information, often called the *central dogma*. Above each arrow, write the name of the process involved.

_____ → _____ → _____

The Genetic Code A sequence of three nucleotides provides 4^3 (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 nucleotide base triplets along the **template strand** of a gene are transcribed into complementary mRNA **codons**. One strand of a DNA molecule can serve as the template strand for one gene and the nontemplate strand for another. The mRNA is complementary to the DNA template because its nucleotides follow the same base-pairing rules, with the exception that uracil substitutes for thymine in RNA. The term *codon* can also refer to the DNA nucleotide triplets on the *nontemplate* strand, which are identical to the mRNA, except that they have T instead of U.

During translation, the sequence of codons, read in the 5′ → 3′ direction, determines the sequence of amino acids in the polypeptide.

How did molecular biologists crack the genetic code? 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. By the mid-1960s, molecular biologists had deciphered all 64 codons. 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.

FOCUS QUESTION 14.2

Practice using the codon table in your textbook. Determine the amino acid sequence for a polypeptide coded for by the following mRNA transcript (written 5′ → 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.

14.2 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. RNA polymerase joins RNA nucleotides that are complementary to the DNA template strand in a 5′ → 3′ direction. In bacteria, 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 specific binding of RNA polymerase to the promoter determines where transcription starts and which DNA strand is used as the template. The promoter includes the transcription **start point** and 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 along a 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 signal sequence (coding for AAUAAA), and proteins cut loose the pre-mRNA.

FOCUS QUESTION 14.3

Describe the key steps of transcription as they occur in eukaryotes:

a.

b.

c.

14.3 Eukaryotic cells modify RNA after transcription

In eukaryotes, pre-mRNA is modified by **RNA processing** before it leaves the nucleus.

Alteration of mRNA Ends A modified guanine nucleotide is attached to the 5′ end of a pre-mRNA, and a string of adenine nucleotides, called a **poly-A tail**, is added to the 3′ end. The **5′ cap** and poly-A tail may facilitate transport of mRNA from the nucleus, aid ribosome attachment, and protect the ends of mRNA from hydrolytic enzymes. The cap and tail are attached to the untranslated regions (UTRs) at the 5′ and 3′ ends.

Split Genes and RNA Splicing Long segments of noncoding nucleotide sequences, known as **introns** or intervening sequences, occur within the boundaries of eukaryotic genes. The remaining coding regions are called **exons**, since they are expressed in protein synthesis (with the exception of the 5′ and 3′ UTRs, which are not translated). After a primary transcript is made of a gene, introns are removed and exons joined together before the mRNA leaves the nucleus—a process called **RNA splicing**.

Alternative RNA splicing allows different polypeptides to be produced from a single gene, depending on which exons are combined during RNA processing.

The signals for RNA splicing are sets of a few nucleotides at either end of each intron. Small RNAs within a large protein complex called a **spliceosome** snip an intron out of the RNA transcript and connect the adjoining exons.

Such RNA molecules that act as enzymes are called **ribozymes**. In some cases of RNA splicing, intron RNA catalyzes its own removal. What three properties of RNA relate to its ability to function as an enzyme? RNA is single-stranded and can base-pair with itself, forming a specific three-dimensional structure; some of its bases contain functional groups that can participate in catalysis; and it can hydrogen-bond with other nucleic acid molecules, allowing it to precisely locate splicing regions.

FOCUS QUESTION 14.4

How does the mRNA that leaves the nucleus differ from pre-mRNA?

14.4 Translation is the RNA-directed synthesis of a polypeptide: a closer look

Molecular Components of Translation **Transfer RNA (tRNA)** molecules carry amino acids to ribosomes, where they are added to a growing polypeptide. Each tRNA carries a specific amino acid and has a nucleotide triplet, called an **anticodon**, that base-pairs with a complementary codon on mRNA, thus assuring that amino acids are arranged in the sequence prescribed by the transcription from DNA.

Transfer RNA is transcribed in the nucleus of a eukaryote and moves into the cytoplasm, where it can be used repeatedly. Due to hydrogen bonding between complementary nucleotide base sequences, these single-stranded, short RNA molecules fold into a three-dimensional, roughly L-shaped structure. The anticodon is at one end of the L; the 3′ end is the attachment site for its specific amino acid.

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

Sixty-one codons for amino acids can be read from mRNA, but there are only about 45 different tRNA molecules. A phenomenon known as **wobble** enables the third nucleotide of some tRNA anticodons to pair with more than one kind of nucleotide in the codon. Thus, a tRNA may recognize more than one mRNA codon, all of which code for the same amino acid carried by that tRNA.

FOCUS QUESTION 14.5

Using some of the codons and the amino acids you identified in Focus Question 14.2, fill in the following table.

DNA Triplet 3′→ 5′	mRNA Codon 5′→ 3′	Anticodon 3′→ 5′	Amino Acid
			methionine
		GCA	
TTC			
	UAG		

Ribosomes facilitate the specific pairing of tRNA anticodons with mRNA codons during protein synthesis. Ribosomes 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 differ enough in molecular composition that some

antibiotics can inhibit them without affecting eukaryotic ribosomes.

A large and a small subunit join to form a ribosome when they attach to an mRNA molecule. Ribosomes have a binding site for mRNA and three tRNA binding sites: a **P site** (*p*eptidyl tRNA-binding site) that holds the tRNA carrying the growing polypeptide chain, an **A site** (*a*minoacyl tRNA-binding site) that holds the tRNA carrying the amino acid to be added next, and an **E site** (*e*xit site) from which discharged tRNAs leave the ribosome. A ribosome also has an *exit tunnel* through which the growing polypeptide passes out of the ribosome.

Building a Polypeptide Each of the three stages of protein synthesis—initiation, elongation, and termination—requires the aid of protein "factors." The first two stages also require energy 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 hydrolysis of 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 GTP.

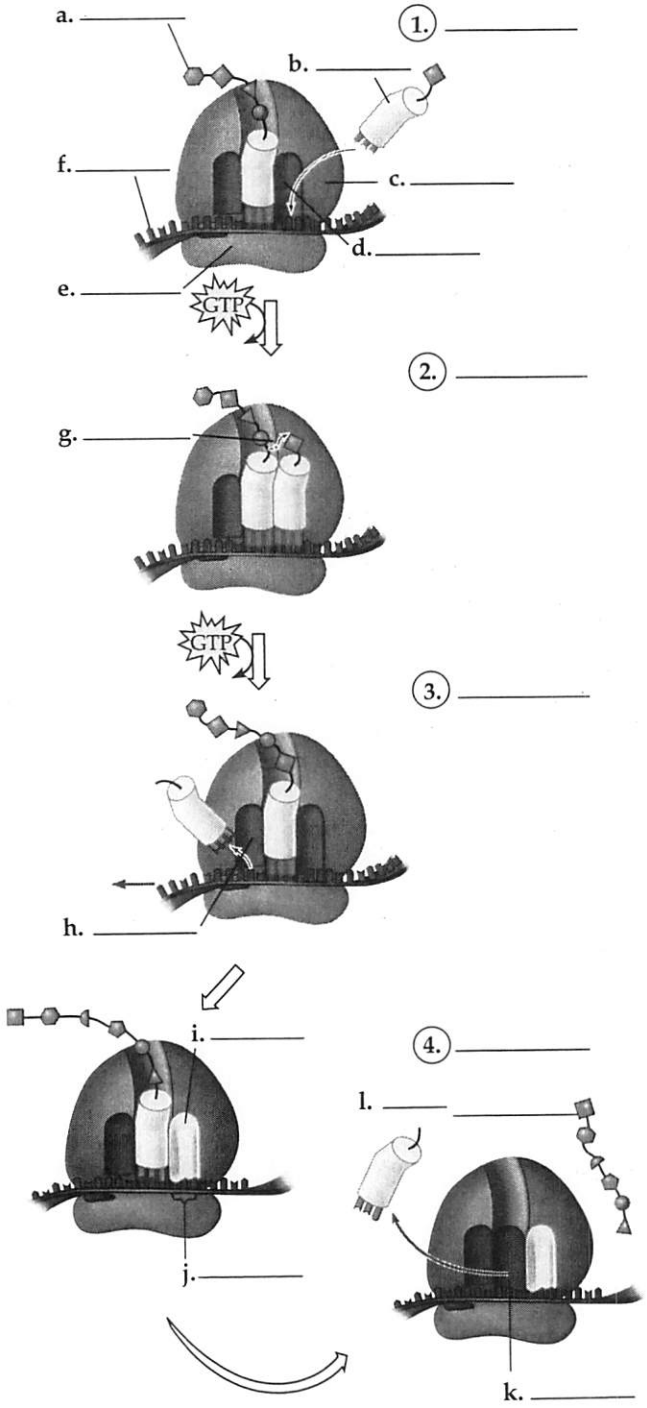
In the peptide bond formation step, an RNA molecule of the large subunit catalyzes the formation of a peptide bond between the carboxyl end of the polypeptide held in the P site and the amino group of the new amino acid in the A site. The polypeptide is now held by the tRNA in the A site.

In translocation, the tRNA carrying the growing polypeptide is translocated to the P site, a process requiring energy from the hydrolysis of GTP. 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. With the hydrolysis of GTP and the aid of other protein factors, the two ribosomal subunits and other components dissociate. Review translation by completing Focus Question 14.6.

FOCUS QUESTION 14.6

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. (This diagram does not include the initiation 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 removed; segments of the polypeptide may be excised; or several polypeptides may associate into a protein with 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**. This short sequence of amino acids 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 translocation 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.

FOCUS QUESTION 14.7

What determines if a ribosome becomes bound to the ER?

Making Multiple Polypeptides in Bacteria and Eukaryotes An mRNA may be translated simultaneously by several ribosomes in strings called polyribosomes (or *polysomes*).

14.5 Mutations of one or a few nucleotides can affect protein structure and function

Mutations, changes in the genetic information of a cell (or virus), may be either large scale (involving long segments of a chromosome) or small scale (such as **point mutations**, which affect just one nucleotide pair). If the mutation occurs in a cell that gives rise to a gamete, it may be passed on to offspring.

Types of Small-Scale Mutations This first type of small-scale mutation is a **nucleotide-pair substitution**, in which one nucleotide and its complementary partner is replaced with another pair of nucleotides. Due to the redundancy of the genetic code, some substitutions do not change the amino acid translation and are called **silent mutations**. A **missense mutation**, which results in the insertion of a different amino acid, may not alter 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 nucleotide-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. Occasionally such a change proves beneficial.

Nonsense mutations occur when a 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.

The second type of small-scale mutations includes nucleotide-pair **insertions** or **deletions**. If these are not in multiples of three, they will alter the reading frame. All nucleotides downstream from the mutation will be improperly grouped into codons, creating extensive missense mutations and usually ending in nonsense. Such **frameshift mutations** almost always produce nonfunctional proteins.

Mutagens *Spontaneous mutations* include nucleotide-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 a variety of chemical agents can cause mutations and are called **mutagens**. Tests can measure the mutagenic effects of chemicals and thus their potential carcinogenic risk.

FOCUS QUESTION 14.8

Define the following terms, and explain what type of small-scale mutation could cause each of these types of 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 one gene–one polypeptide. Research continually refines our understanding of the structural and functional aspects of genes, which now include introns, promoters, and other regulatory regions. Currently, 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 nucleotide triplet at one end of a tRNA molecule that base-pairs with a particular complementary codon on an mRNA molecule)

exo- = out, outside, without (*exon:* a sequence within a primary transcript that remains in the RNA after RNA processing; also the region of DNA from which this sequence was transcribed)

intro- = within (*intron:* a noncoding, intervening sequence within a primary transcript that is removed during RNA processing; also the region of DNA from which this sequence was transcribed)

muta- = change; **-gen** = producing (*mutagen:* a chemical or physical agent that interacts with DNA and can cause a mutation) .

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

Structure Your Knowledge

1. You have been introduced to several types of RNA in this chapter. List three of these types and their functions. (You may also recall a fourth type that functions in spliceosomes.) What explains the functional versatility of RNA molecules?

2. Make sure you understand and can explain the processes of transcription and translation. To help you review these processes, fill in the following table describing various aspects of eukaryotic gene expression (either by yourself or in a study group).

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

3. 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?

4. Prepare a concept map showing the types and consequences of small-scale mutations.

Test Your Knowledge

MULTIPLE CHOICE: *Choose the one best answer.*

1. A series of studies on mutants of *Neurospora* identified 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) have to be supplied to the mutants that had a defective enzyme for the ornithine → citrulline step in order for them to grow?
   a. the precursor
   b. ornithine
   c. citrulline
   d. either ornithine or citrulline
   e. the 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 nontemplate DNA strand is CAT, what is the corresponding codon on mRNA?
   a. UAC
   b. CAU
   c. GUA
   d. GTA
   e. CAT

4. A bacterial gene 600 nucleotides long can code at most for a polypeptide of how many amino acids?
   a. 100
   b. 200
   c. 300
   d. 600
   e. 1,800

5. 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.
   e. moves along the template strand of DNA, elongating an RNA molecule in a 5′ → 3′ direction.

6. 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 is established by the promoter.
   d. It doesn't matter which strand is the template because they are complementary and will produce the same mRNA.
   e. It is signaled by a polyadenylation signal sequence.

7. Which enzyme synthesizes tRNA?
   a. DNA polymerase
   b. RNA polymerase
   c. reverse transcriptase
   d. aminoacyl-tRNA synthetase
   e. ribosomal RNA

8. Which of the following is *not* involved in the formation of a eukaryotic transcription initiation complex?
   a. TATA box
   b. transcription factors
   c. small RNA molecules
   d. RNA polymerase II
   e. promoter

9. 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 small RNAs, 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.

10. All of the following are transcribed from DNA *except*
    a. exons.
    b. introns.
    c. tRNA.
    d. 3′ and 5′ UTRs.
    e. promoter.

11. 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 exons caused by alternative RNA splicing.
    d. an RNA molecule that functions as an enzyme.
    e. an enzyme that produces both small and large ribosomal subunits.

12. Which of the following would *not* be found in a bacterial cell?
    a. mRNA
    b. rRNA
    c. small RNAs in spliceosome
    d. RNA polymerase
    e. simultaneous transcription and translation

13. Which of the following is transcribed and then translated to form a protein product?
    a. a gene for tRNA
    b. an intron
    c. a gene for a transcription factor
    d. 5′ and 3′ UTRs
    e. a gene for rRNA

14. Transfer RNA
    a. translocates a growing polypeptide destined for export to the endoplasmic reticulum.
    b. binds to its specific amino acid in the active site of an aminoacyl-tRNA synthetase.
    c. has catalytic activity and is thus a ribozyme.
    d. is translated from mRNA.
    e. is produced in the nucleolus.

15. Place the following events in the synthesis of a polypeptide in the proper order.
    1. A peptide bond forms.
    2. An aminoacyl tRNA matches its anticodon to the codon in the A site.
    3. A tRNA translocates from the A site to the P site, and an unattached tRNA exits from the E site.
    4. The large subunit attaches to the small subunit, with the initiator tRNA in the P site.
    5. A small subunit binds to an mRNA and an initiator tRNA.

    a. 4-5-3-2-1
    b. 4-5-2-1-3
    c. 5-4-3-2-1
    d. 5-4-1-2-3
    e. 5-4-2-1-3

16. Translocation in the process of translation involves
    a. the hydrolysis of GTP.
    b. movement of the tRNA in the A site to the P site.
    c. movement along the mRNA a distance of one triplet.
    d. the release of the unattached tRNA from the E site.
    e. all of the above.

17. Which of the following types of molecules catalyzes the formation of a peptide bond?
    a. RNA polymerase
    b. rRNA
    c. mRNA
    d. aminoacyl-tRNA synthetase
    e. tRNA

18. Which of the following is *not* true of an anticodon?
    a. It consists of three nucleotides.
    b. It lines up in the 5′ → 3′ direction along the 5′ → 3′ mRNA strand.
    c. It extends from one loop of a tRNA molecule.
    d. It may pair with more than one codon.
    e. Its base uracil base-pairs with adenine.

19. Changes in a polypeptide following translation may involve
    a. the addition of sugars or lipids to certain amino acids.
    b. the enzymatic addition of amino acids at the beginning of the chain.
    c. the removal of poly-A from the end of the chain.
    d. the addition of a 5′ cap of a modified guanosine.
    e. all of the above.

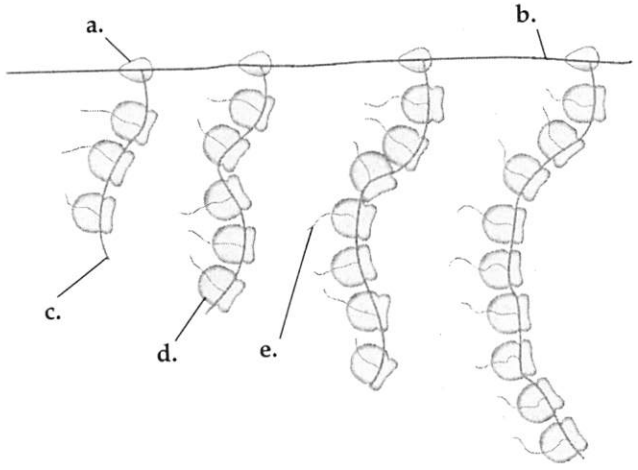
20. Several proteins may be produced at the same time from a single mRNA by
    a. the action of several ribosomes in a string, called a polyribosome.
    b. several RNA polymerase molecules working sequentially.
    c. signal peptides that associate ribosomes with rough ER.
    d. the action of several promoter regions.
    e. the involvement of multiple spliceosomes.

21. A signal peptide
    a. is most likely to be found on cytosolic proteins produced by bacterial cells.
    b. directs an mRNA molecule into the lumen of the ER.
    c. is a sign to bind the small ribosomal unit at the initiation codon.
    d. would be the first 20 or so amino acids of a protein destined for a membrane location or for secretion from the cell.
    e. is part of the UTR following the 5′ cap.

22. A nucleotide deletion early in the coding sequence of a gene would most likely result in
    a. a nonsense mutation.
    b. a frameshift mutation.
    c. multiple missense mutations.
    d. a nonfunctional protein.
    e. all of the above.

23. The type of mutation responsible for sickle-cell anemia is
    a. a silent mutation.
    b. a nucleotide-pair insertion.
    c. a point mutation.
    d. a nucleotide-pair substitution.
    e. Both **c** and **d** describe the type of mutation.

Use the following diagram of coupled transcription and translation in bacteria to answer questions 24 through 28.



24. Which letter refers to an mRNA molecule?

25. Which letter refers to a forming polypeptide?

26. Which letter refers to RNA polymerase?

27. Which letter refers to a ribosome?

28. Which letters indicate structures or molecules containing nucleotides?
    a. **a** and **b**
    b. **a**, **b**, and **d**
    c. **b**, **c**, and **d**
    d. **b**, **d**, and **e**
    e. **c** and **d**