15 | GENES AND PROTEINS

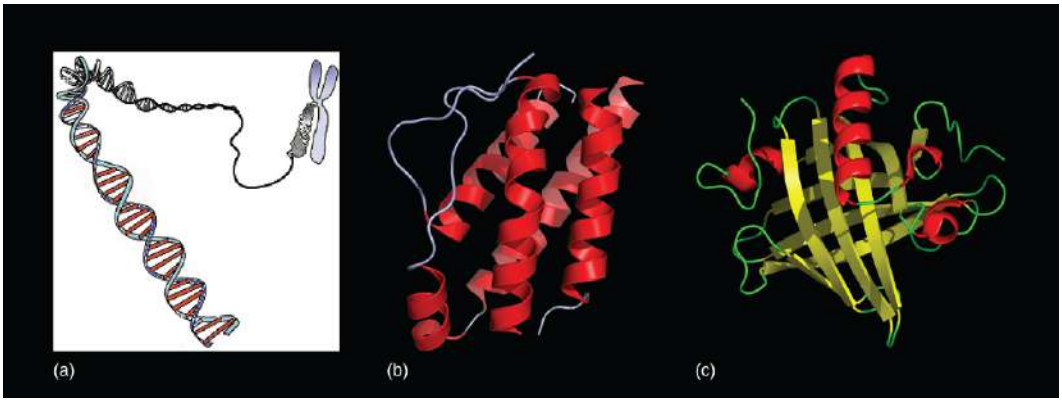

Figure 15.1 Genes, which are carried on (a) chromosomes, are linearly organized instructions for making the RNA and protein molecules that are necessary for all of processes of life. The (b) interleukin-2 protein and (c) alpha-2u-globulin protein are just two examples of the array of different molecular structures that are encoded by genes. (credit "chromosome: National Human Genome Research Institute; credit "interleukin-2": Ramin Herati/Created from PDB 1M47 and rendered with Pymol; credit "alpha-2u-globulin": Darren Logan/rendered with AISMIG)

Chapter Outline

15.1: The Genetic Code

15.2: Prokaryotic Transcription

15.3: Eukaryotic Transcription

15.4: RNA Processing in Eukaryotes

15.5: Ribosomes and Protein Synthesis

Introduction

The definition of gene has progressed from being an abstract unit of heredity in Mendel's time to our current concept of a tangible molecular entity capable of replication, expression, and mutation (**Figure 15.1**). Currently, we can perform tests for many genetic diseases, but these tests create ethical and legal issues. For example, would you want to be tested for a debilitating genetic disease if there was the possibility insurance companies could use that information to deny you coverage? Fortunately, the Genetic Information Nondiscrimination Act of 2008 protects American citizens from discrimination from both insurance companies and employers based on genetic information. More information about policy, legal, and ethical issues in genetic research can be found **here (http://openstaxcollege.org/l/32genomegov)** .

15.1 | The Genetic Code

In this section, you will explore the following questions:

- What is the "Central Dogma" of protein synthesis?
- What is the genetic code, and how does nucleotide sequence prescribe the amino acid and polypeptide sequence?

Connection for AP® Courses

Since the rediscovery of Mendel's work in the 1900s, scientists have learned much about how the genetic blueprints stored in DNA are capable of replication, expression, and mutation. Just as the 26 letters of the English alphabet can be arranged into what seems to be a limitless number of words, with new ones added to the dictionary every year, the four nucleotides of DNA—A, T, C, and G—can generate sequences of DNA called genes that specify tens of thousands of polymers of amino acids. In turn, these sequences can be transcribed into mRNA and translated into proteins which orchestrate nearly every function of the cell. The genetic code refers to the DNA alphabet (A, T, C, G), the RNA alphabet (A, U, C, G), and the polypeptide alphabet (20 amino acids). But how do genes located on a chromosome ultimately produce a polypeptide that can result in a physical phenotype such as hair or eye color—or a disease like cystic fibrosis or hemophilia?

The Central Dogma describes the normal flow of genetic information from DNA to mRNA to protein: DNA in genes specify sequences of mRNA which, in turn, specify amino acid sequences in proteins. The process requires two steps, transcription and translation. During transcription, genes are used to make messenger RNA (mRNA). In turn, the mRNA is used to direct the synthesis of proteins during the process of translation. Translation also requires two other types of RNA: transfer RNA (tRNA) and ribosomal RNA (rRNA). The genetic code is a triplet code, with each RNA codon consisting of three consecutive nucleotides that specify one amino acid or the release of the newly formed polypeptide chain; for example, the mRNA codon CAU specifies the amino acid histidine. The code is degenerate; that is, some amino acids are specified by more than one codon, like synonyms you study in your English class (different word, same meaning). For example, CCU, CCC, CCA, and CCG are all codons for proline. It is important to remember the same genetic code is universal to almost all organisms on Earth. Small variations in codon assignment exist in mitochondria and some microorganisms.

Deviations from the simple scheme of the central dogma are discovered as researchers explore gene expression with new technology. For example the human immunodeficiency virus (HIV) is a retrovirus which stores its genetic information in single stranded RNA molecules. Upon infection of a host cell, RNA is used as a template by the virally encoded enzyme, reverse transcriptase, to synthesize DNA. The viral DNA is later transcribed into mRNA and translated into proteins. Some RNA viruses such as the influenza virus never go through a DNA step. The RNA genome is replicated by an RNA dependent RNA polymerase which is virally encoded.

The content presented in this section supports the Learning Objectives outlined in Big Idea 1 and Big Idea 3 of the AP® Biology Curriculum Framework. The Learning Objectives merge Essential Knowledge content with one or more of the seven Science Practices. These Learning Objectives provide a transparent foundation for the AP® Biology course, along with inquiry-based laboratory experiences, instructional activities, and AP® Exam questions.

Big Idea 1	The process of evolution drives the diversity and unity of life.
Enduring Understanding 1.B	Organisms are linked by lines of descent from common ancestry.
Essential Knowledge	**1.B.1** Organisms share many conserved core processes and features that evolved and are widely distributed among organisms today.
Science Practice	**3.1** The student can pose scientific questions.
Science Practice	**7.2** The student can connect concepts in and across domain(s) to generalize or extrapolate in and/or across enduring understandings and/or big ideas.
Learning Objective	**1.15** The student is able to describe specific examples of conserved core biological processes and features shared by all domains or within one domain of life, and how these shared, conserved core processes and features support the concept of common ancestry for all organisms.
Big Idea 3	Living systems store, retrieve, transmit and respond to information essential to life processes.
Enduring Understanding 3.A	Heritable information provides for continuity of life.

Essential Knowledge	**3.A.1** DNA, and in some cases RNA, is the primary source of heritable information.
Science Practice	**6.5** The student can evaluate alternative scientific explanations.
Learning Objective	**3.1** The student is able to construct scientific explanations that use the structure and functions of DNA and RNA to support the claim that DNA and, in some cases, that RNA are the primary sources of heritable information.

The Science Practice Challenge Questions contain additional test questions for this section that will help you prepare for the AP exam. These questions address the following standards:
[APLO 3.4][APLO 3.25]

The cellular process of transcription generates messenger RNA (mRNA), a mobile molecular copy of one or more genes with an alphabet of A, C, G, and uracil (U). Translation of the mRNA template converts nucleotide-based genetic information into a protein product. Protein sequences consist of 20 commonly occurring amino acids; therefore, it can be said that the protein alphabet consists of 20 letters (**Figure 15.2**). Each amino acid is defined by a three-nucleotide sequence called the triplet codon. Different amino acids have different chemistries (such as acidic versus basic, or polar and nonpolar) and different structural constraints. Variation in amino acid sequence gives rise to enormous variation in protein structure and function.

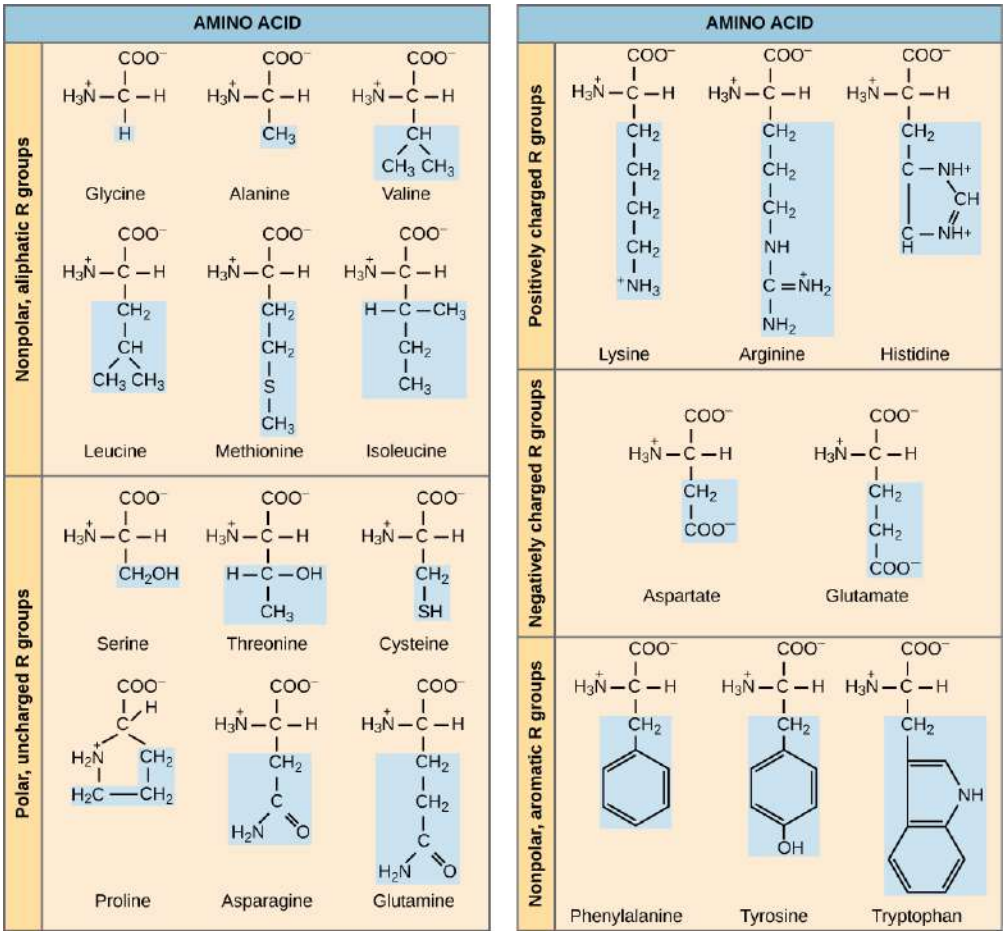

Figure 15.2 Structures of the 20 amino acids found in proteins are shown. Each amino acid is composed of an amino group (NH_3^+), a carboxyl group (COO^-), and a side chain (blue). The side chain may be nonpolar, polar, or charged, as well as large or small. It is the variety of amino acid side chains that gives rise to the incredible variation of protein structure and function.

The Central Dogma: DNA Encodes RNA; RNA Encodes Protein

The flow of genetic information in cells from DNA to mRNA to protein is described by the **Central Dogma** (Figure 15.3), which states that genes specify the sequence of mRNAs, which in turn specify the sequence of proteins. The decoding of one molecule to another is performed by specific proteins and RNAs. Because the information stored in DNA is so central to cellular function, it makes intuitive sense that the cell would make mRNA copies of this information for protein synthesis, while keeping the DNA itself intact and protected. The copying of DNA to RNA is relatively straightforward, with one nucleotide being added to the mRNA strand for every nucleotide read in the DNA strand. The translation to protein is a bit more complex because three mRNA nucleotides correspond to one amino acid in the polypeptide sequence. However, the translation to protein is still systematic and **colinear**, such that nucleotides 1 to 3 correspond to amino acid 1, nucleotides 4 to 6 correspond to amino acid 2, and so on.

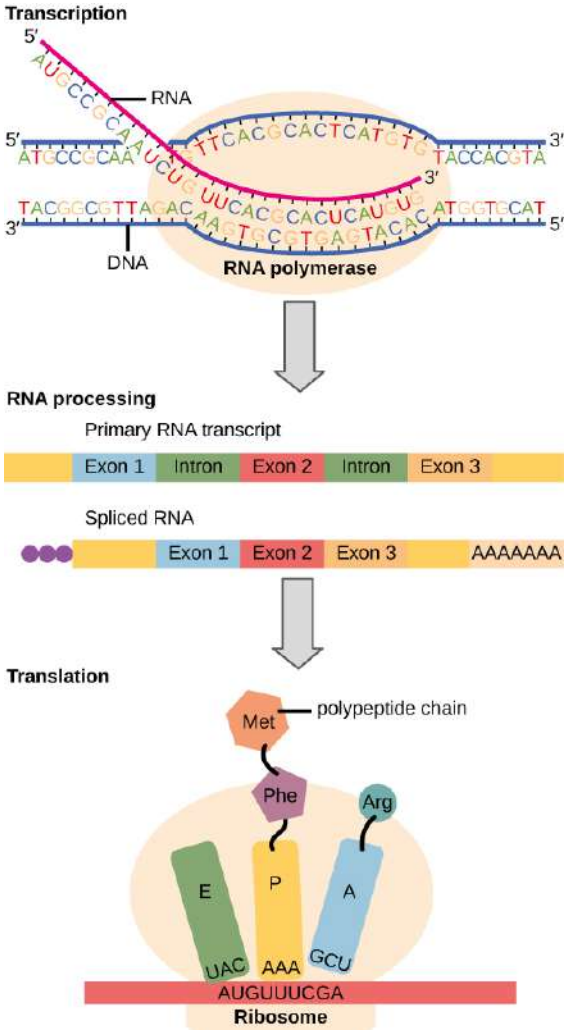

Figure 15.3 Instructions on DNA are transcribed onto messenger RNA. Ribosomes are able to read the genetic information inscribed on a strand of messenger RNA and use this information to string amino acids together into a protein.

The Genetic Code Is Degenerate and Universal

Given the different numbers of "letters" in the mRNA and protein "alphabets," scientists theorized that combinations of nucleotides corresponded to single amino acids. Nucleotide doublets would not be sufficient to specify every amino acid because there are only 16 possible two-nucleotide combinations (4^2). In contrast, there are 64 possible nucleotide triplets (4^3), which is far more than the number of amino acids. Scientists theorized that amino acids were encoded by nucleotide triplets and that the genetic code was **degenerate**. In other words, a given amino acid could be encoded by more than one nucleotide triplet. This was later confirmed experimentally; Francis Crick and Sydney Brenner used the chemical mutagen proflavin to insert one, two, or three nucleotides into the gene of a virus. When one or two nucleotides were inserted, protein synthesis was completely abolished. When three nucleotides were inserted, the protein was synthesized and functional. This

demonstrated that three nucleotides specify each amino acid. These nucleotide triplets are called **codons**. The insertion of one or two nucleotides completely changed the triplet **reading frame**, thereby altering the message for every subsequent amino acid (**Figure 15.4**). Though insertion of three nucleotides caused an extra amino acid to be inserted during translation, the integrity of the rest of the protein was maintained.

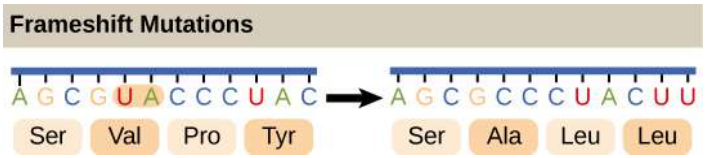

Figure 15.4 The deletion of two nucleotides shifts the reading frame of an mRNA and changes the entire protein message, creating a nonfunctional protein or terminating protein synthesis altogether.

Scientists painstakingly solved the genetic code by translating synthetic mRNAs in vitro and sequencing the proteins they specified (**Figure 15.5**).

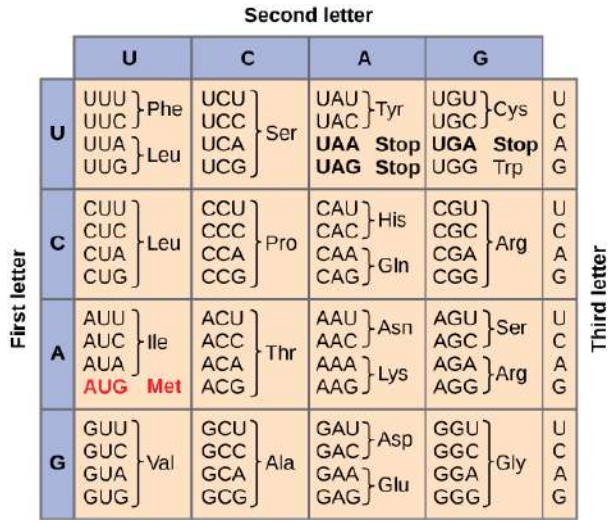

Figure 15.5 This figure shows the genetic code for translating each nucleotide triplet in mRNA into an amino acid or a termination signal in a nascent protein. (credit: modification of work by NIH)

In addition to instructing the addition of a specific amino acid to a polypeptide chain, three of the 64 codons terminate protein synthesis and release the polypeptide from the translation machinery. These triplets are called **nonsense codons**, or stop codons. Another codon, AUG, also has a special function. In addition to specifying the amino acid methionine, it also serves as the start codon to initiate translation. The reading frame for translation is set by the AUG start codon near the 5' end of the mRNA.

The genetic code is universal. With a few exceptions, virtually all species use the same genetic code for protein synthesis. Conservation of codons means that a purified mRNA encoding the globin protein in horses could be transferred to a tulip cell, and the tulip would synthesize horse globin. That there is only one genetic code is powerful evidence that all of life on Earth shares a common origin, especially considering that there are about 10^{84} possible combinations of 20 amino acids and 64 triplet codons.

Transcribe a gene and translate it to protein using complementary pairing and the genetic code at this **site (http://openstaxcollege.org/l/create_protein)** .

Some hereditary and age-related diseases are caused by translation errors. Explain why an error in translation may cause disease.

a. If there is an error in translation, the correct lipids will not be made for signaling, storage of energy or to perform vital functions. This can cause hereditary and age-related diseases.

b. Translation is the process in which a particular segment of DNA is copied into RNA (mRNA) by the enzyme RNA polymerase. Error in such copying can lead to various hereditary and age-related diseases.

c. Translation is the process used by ribosomes to synthesize proteins from amino acids. If there is an error in this process, the correct proteins will not be made to build important body tissue or perform vital functions thus leading to hereditary and age-related diseases.

d. Translation is the process Golgi bodies use to synthesize proteins from amino acids. If there is an error in this process, the correct proteins will not be made to build important body tissue or perform vital functions.

Think About It

- A strand of DNA has the nucleotide sequence 3'……GCT GTC AAA TTC GAT……5'. What is the sequence of mRNA that is complementary to this DNA sequence? Using the chart of codons in the text, determine the sequence of amino acids which can be generated from this strand of DNA.
- How does degeneracy of the genetic code make cells less vulnerable to mutations? What is an advantage of degeneracy with respect to the negative impact of random mutations on natural selection and evolution?

Degeneracy is believed to be a cellular mechanism to reduce the negative impact of random mutations. Codons that specify the same amino acid typically only differ by one nucleotide. In addition, amino acids with chemically similar side chains are encoded by similar codons. This nuance of the genetic code ensures that a single-nucleotide substitution mutation might either specify the same amino acid but have no effect or specify a similar amino acid, preventing the protein from being rendered completely nonfunctional.

Which Has More DNA: A Kiwi or a Strawberry?

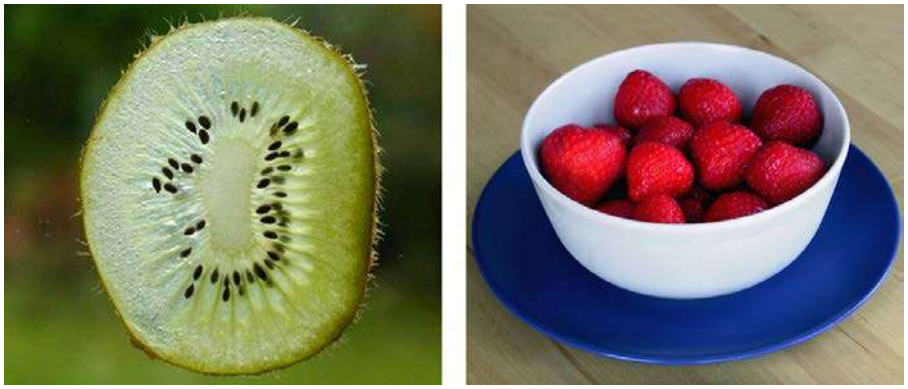

Figure 15.6 Do you think that a kiwi or a strawberry has more DNA per fruit? (credit “kiwi”: "Kelbv"/Flickr; credit: “strawberry”: Alisdair McDiarmid)

Question: Would a kiwifruit and strawberry that are approximately the same size (**Figure 15.6**) also have approximately the same amount of DNA?

Background: Genes are carried on chromosomes and are made of DNA. All mammals are diploid, meaning they have two copies of each chromosome. However, not all plants are diploid. The common strawberry is octoploid (8*n*) and the cultivated kiwi is hexaploid (6*n*). Research the total number of chromosomes in the cells of each of these fruits and think about how this might correspond to the amount of DNA in these fruits’ cell nuclei. Read about the technique of DNA isolation to understand how each step in the isolation protocol helps liberate and precipitate DNA.

Hypothesis: Hypothesize whether you would be able to detect a difference in DNA quantity from similarly sized strawberries and kiwis. Which fruit do you think would yield more DNA?

Test your hypothesis: Isolate the DNA from a strawberry and a kiwi that are similarly sized. Perform the experiment in at least triplicate for each fruit.

1. Prepare a bottle of DNA extraction buffer from 900 mL water, 50 mL dish detergent, and two teaspoons of table salt. Mix by inversion (cap it and turn it upside down a few times).
2. Grind a strawberry and a kiwifruit by hand in a plastic bag, or using a mortar and pestle, or with a metal bowl and the end of a blunt instrument. Grind for at least two minutes per fruit.
3. Add 10 mL of the DNA extraction buffer to each fruit, and mix well for at least one minute.
4. Remove cellular debris by filtering each fruit mixture through cheesecloth or porous cloth and into a funnel placed in a test tube or an appropriate container.
5. Pour ice-cold ethanol or isopropanol (rubbing alcohol) into the test tube. You should observe white, precipitated DNA.
6. Gather the DNA from each fruit by winding it around separate glass rods.

Record your observations: Because you are not quantitatively measuring DNA volume, you can record for each trial whether the two fruits produced the same or different amounts of DNA as observed by eye. If one or the other fruit produced noticeably more DNA, record this as well. Determine whether your observations are consistent with several pieces of each fruit.

Analyze your data: Did you notice an obvious difference in the amount of DNA produced by each fruit? Were your results reproducible?

Draw a conclusion: Given what you know about the number of chromosomes in each fruit, can you conclude that chromosome number necessarily correlates to DNA amount? Can you identify any drawbacks to this procedure? If you had access to a laboratory, how could you standardize your comparison and make it more quantitative?

15.2 | Prokaryotic Transcription

In this section, you will explore the following questions:

- What are the steps, in order, in prokaryotic transcription?
- How and when is transcription terminated?

Connection for AP® Courses

During transcription, the enzyme RNA polymerase moves along the DNA template, reading nucleotides in a 3′ to 5′ direction, with U pairing with A and C with G, and assembling the mRNA transcript in a 5′ to 3′ direction. In prokaryotes, mRNA synthesis is initiated at a promoter sequence on the DNA template. Transcription continues until RNA polymerase reaches a stop or terminator sequence at the end of the gene. Termination frees the mRNA and often occurs by the formation of an mRNA hairpin.

Information presented and the examples highlighted in the section support concepts outlined in Big Idea 3 of the AP® Biology Curriculum Framework. The Learning Objectives listed in the Curriculum Framework provide a transparent foundation for the AP® Biology course, an inquiry-based laboratory experience, instructional activities, and AP® Exam questions. A Learning Objective merges required content with one or more of the seven Science Practices.

Big Idea 3	Living systems store, retrieve, transmit and respond to information essential to life processes.
Enduring Understanding 3.A	Heritable information provides for continuity of life.
Essential Knowledge	**3.A.1** DNA, and in some cases RNA, is the primary source of heritable information.
Science Practice	**6.5** The student can evaluate alternative scientific explanations.
Learning Objective	**3.1** The student is able to construct scientific explanations that use the structures and mechanisms of DNA and RNA to support the claim that DNA and, in some cases, that RNA are the primary sources of heritable information.

The Science Practice Challenge Questions contain additional test questions for this section that will help you prepare for the AP exam. These questions address the following standards:
[APLO 2.23][APLO 3.28][APLO 4.8][APLO 4.24]

The prokaryotes, which include bacteria and archaea, are mostly single-celled organisms that, by definition, lack membrane-bound nuclei and other organelles. A bacterial chromosome is a covalently closed circle that, unlike eukaryotic chromosomes, is not organized around histone proteins. The central region of the cell in which prokaryotic DNA resides is called the nucleoid. In addition, prokaryotes often have abundant **plasmids**, which are shorter circular DNA molecules that may only contain one or a few genes. Plasmids can be transferred independently of the bacterial chromosome during cell division and often carry traits such as antibiotic resistance.

Transcription in prokaryotes (and in eukaryotes) requires the DNA double helix to partially unwind in the region of mRNA synthesis. The region of unwinding is called a **transcription bubble.** Transcription always proceeds from the same DNA strand for each gene, which is called the **template strand**. The mRNA product is complementary to the template strand and is almost identical to the other DNA strand, called the **nontemplate strand**. The only difference is that in mRNA, all of the T nucleotides are replaced with U nucleotides. In an RNA double helix, A can bind U via two hydrogen bonds, just as in A–T pairing in a DNA double helix.

The nucleotide pair in the DNA double helix that corresponds to the site from which the first 5' mRNA nucleotide is transcribed is called the +1 site, or the **initiation site**. Nucleotides preceding the initiation site are given negative numbers and are designated **upstream**. Conversely, nucleotides following the initiation site are denoted with "+" numbering and are

called **downstream** nucleotides.

Initiation of Transcription in Prokaryotes

Prokaryotes do not have membrane-enclosed nuclei. Therefore, the processes of transcription, translation, and mRNA degradation can all occur simultaneously. The intracellular level of a bacterial protein can quickly be amplified by multiple transcription and translation events occurring concurrently on the same DNA template. Prokaryotic transcription often covers more than one gene and produces polycistronic mRNAs that specify more than one protein.

Our discussion here will exemplify transcription by describing this process in *Escherichia coli*, a well-studied bacterial species. Although some differences exist between transcription in *E. coli* and transcription in archaea, an understanding of *E. coli* transcription can be applied to virtually all bacterial species.

Prokaryotic RNA Polymerase

Prokaryotes use the same RNA polymerase to transcribe all of their genes. In *E. coli*, the polymerase is composed of five polypeptide subunits, two of which are identical. Four of these subunits, denoted α, α, β, and β' comprise the polymerase **core enzyme**. These subunits assemble every time a gene is transcribed, and they disassemble once transcription is complete. Each subunit has a unique role; the two α-subunits are necessary to assemble the polymerase on the DNA; the β-subunit binds to the ribonucleoside triphosphate that will become part of the nascent "recently born" mRNA molecule; and the β' binds the DNA template strand. The fifth subunit, σ, is involved only in transcription initiation. It confers transcriptional specificity such that the polymerase begins to synthesize mRNA from an appropriate initiation site. Without σ, the core enzyme would transcribe from random sites and would produce mRNA molecules that specified protein gibberish. The polymerase comprised of all five subunits is called the **holoenzyme**.

Prokaryotic Promoters

A **promoter** is a DNA sequence onto which the transcription machinery binds and initiates transcription. In most cases, promoters exist upstream of the genes they regulate. The specific sequence of a promoter is very important because it determines whether the corresponding gene is transcribed all the time, some of the time, or infrequently. Although promoters vary among prokaryotic genomes, a few elements are conserved. At the -10 and -35 regions upstream of the initiation site, there are two promoter **consensus** sequences, or regions that are similar across all promoters and across various bacterial species (Figure 15.7). The -10 consensus sequence, called the -10 region, is TATAAT. The -35 sequence, TTGACA, is recognized and bound by σ. Once this interaction is made, the subunits of the core enzyme bind to the site. The A–T-rich -10 region facilitates unwinding of the DNA template, and several phosphodiester bonds are made. The transcription initiation phase ends with the production of abortive transcripts, which are polymers of approximately 10 nucleotides that are made and released.

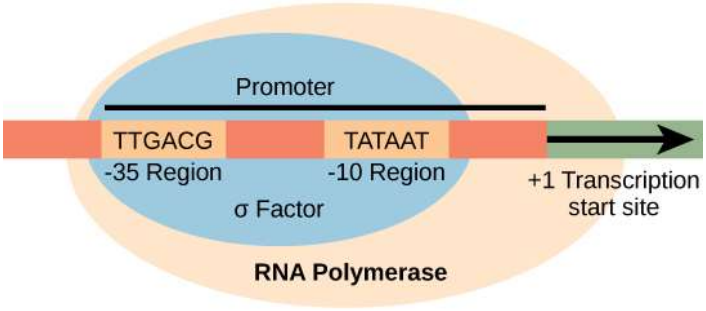

Figure 15.7 The σ subunit of prokaryotic RNA polymerase recognizes consensus sequences found in the promoter region upstream of the transcription start sight. The σ subunit dissociates from the polymerase after transcription has been initiated.

LINK TO LEARNING

View this **MolecularMovies animation (http://openstaxcollege.org/l/transcription)** to see the first part of transcription and the base sequence repetition of the TATA box.

Mutations can occur in any part of the DNA. What can happen if there is a mutation in the promoter sequence?

a. All processes will carry on as usual. Nothing will be affected.

b. RNA polymerase will not be able to attach.

c. RNA polymerase will bind upstream of the promoter sequence.

d. RNA polymerase will bind downstream of the promoter sequence.

Elongation and Termination in Prokaryotes

The transcription elongation phase begins with the release of the σ subunit from the polymerase. The dissociation of σ allows the core enzyme to proceed along the DNA template, synthesizing mRNA in the 5' to 3' direction at a rate of approximately 40 nucleotides per second. As elongation proceeds, the DNA is continuously unwound ahead of the core enzyme and rewound behind it (**Figure 15.8**). The base pairing between DNA and RNA is not stable enough to maintain the stability of the mRNA synthesis components. Instead, the RNA polymerase acts as a stable linker between the DNA template and the nascent RNA strands to ensure that elongation is not interrupted prematurely.

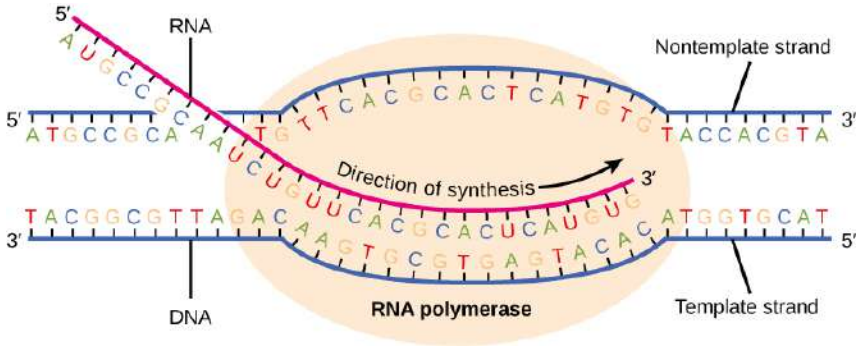

Figure 15.8 During elongation, the prokaryotic RNA polymerase tracks along the DNA template, synthesizes mRNA in the 5' to 3' direction, and unwinds and rewinds the DNA as it is read.

Prokaryotic Termination Signals

Once a gene is transcribed, the prokaryotic polymerase needs to be instructed to dissociate from the DNA template and liberate the newly made mRNA. Depending on the gene being transcribed, there are two kinds of termination signals. One is protein-based and the other is RNA-based. **Rho-dependent termination** is controlled by the rho protein, which tracks along behind the polymerase on the growing mRNA chain. Near the end of the gene, the polymerase encounters a run of G nucleotides on the DNA template and it stalls. As a result, the rho protein collides with the polymerase. The interaction with rho releases the mRNA from the transcription bubble.

Rho-independent termination is controlled by specific sequences in the DNA template strand. As the polymerase nears the end of the gene being transcribed, it encounters a region rich in C–G nucleotides. The mRNA folds back on itself, and the complementary C–G nucleotides bind together. The result is a stable **hairpin** that causes the polymerase to stall as soon as it begins to transcribe a region rich in A–T nucleotides. The complementary U–A region of the mRNA transcript forms only a weak interaction with the template DNA. This, coupled with the stalled polymerase, induces enough instability for the core enzyme to break away and liberate the new mRNA transcript.

Upon termination, the process of transcription is complete. By the time termination occurs, the prokaryotic transcript would already have been used to begin synthesis of numerous copies of the encoded protein because these processes can occur concurrently. The unification of transcription, translation, and even mRNA degradation is possible because all of these processes occur in the same 5' to 3' direction, and because there is no membranous compartmentalization in the prokaryotic cell (**Figure 15.9**). In contrast, the presence of a nucleus in eukaryotic cells precludes simultaneous transcription and translation.

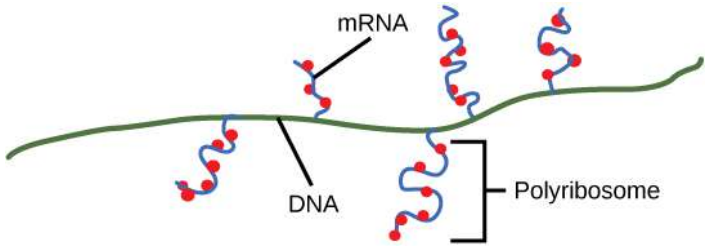

Figure 15.9 Multiple polymerases can transcribe a single bacterial gene while numerous ribosomes concurrently translate the mRNA transcripts into polypeptides. In this way, a specific protein can rapidly reach a high concentration in the bacterial cell.

LINK TO LEARNING

Visit this **BioStudio animation (http://openstaxcollege.org/l/transcription2)** to see the process of prokaryotic transcription.

Why is the stop codon necessary for translation?

a. The stop codon is the first step in a series of steps to end translation.

b. The stop codon is necessary to initiate translation.

c. The stop codon ends translation which allows the polypeptide strand to be released.

d. The stop codon ends translation in order to initiate transcription.

science practices CONNECTION for AP® Courses

Activity

Working in small groups, use a model of DNA to demonstrate synthesis transcription of mRNA to other groups in your class. In your demonstration, be sure to distinguish the differences between DNA and RNA, the template and non-template strands of the DNA, the directionality of transcription, and the significance of promoters.

Think About It

If mRNA is complementary to the DNA template strand, and the DNA template strand is complementary to the DNA non-template strand, are the base sequences of mRNA and the DNA non-template strand ever identical? Justify your answer.

15.3 | Eukaryotic Transcription

In this section, you will explore the following questions:

- What are the steps in eukaryotic transcription?
- What are the structural and functional similarities and differences among the three RNA polymerases?

Connection for AP® Courses

As expected, transcription in eukaryotes is more complex than transcription in prokaryotes. First, transcription in eukaryotes involves one of three types of RNA polymerase, depending on the gene being transcribed. Second, the initiation of transcription involves the binding of several transcription factors to complex promoters which are usually located upstream of the gene being copied. Transcription factors can either activate or inhibit gene expression. Termination of transcription involves the RNA polymerases.

Information presented and the examples highlighted in the section support concepts outlined in Big Idea 3 of the AP® Biology Curriculum Framework. The Learning Objectives listed in the Curriculum Framework provide a transparent foundation for the AP® Biology course, an inquiry-based laboratory experience, instructional activities, and AP® Exam questions. A Learning Objective merges required content with one or more of the seven Science Practices.

Big Idea 3	Living systems store, retrieve, transmit and respond to information essential to life processes.
Enduring Understanding 3.A	Heritable information provides for continuity of life.
Essential Knowledge	**3.A.1** DNA, and in some cases RNA, is the primary source of heritable information.
Science Practice	**6.5** The student can evaluate alternative scientific explanations.
Learning Objective	**3.1** The student is able to construct scientific explanations that use the structures and mechanisms of DNA and RNA to support the claim that DNA and, in some cases, that RNA are the primary sources of heritable information.

The Science Practice Challenge Questions contain additional test questions for this section that will help you prepare for the AP exam. These questions address the following standards:
[APLO 3.3][APLO 3.22][APLO 2.36][APLO 1.14][APLO 2.22][APLO 4.5]

Prokaryotes and eukaryotes perform fundamentally the same process of transcription, with a few key differences. The most important difference between prokaryotes and eukaryotes is the latter's membrane-bound nucleus and organelles. With the genes bound in a nucleus, the eukaryotic cell must be able to transport its mRNA to the cytoplasm and must protect its mRNA from degrading before it is translated. Eukaryotes also employ three different polymerases that each transcribe a different subset of genes. Eukaryotic mRNAs are usually monogenic, meaning that they specify a single protein.

Initiation of Transcription in Eukaryotes

Unlike the prokaryotic polymerase that can bind to a DNA template on its own, eukaryotes require several other proteins, called transcription factors, to first bind to the promoter region and then help recruit the appropriate polymerase.

The Three Eukaryotic RNA Polymerases

The features of eukaryotic mRNA synthesis are markedly more complex those of prokaryotes. Instead of a single polymerase comprising five subunits, the eukaryotes have three polymerases that are each made up of 10 subunits or more. Each eukaryotic polymerase also requires a distinct set of transcription factors to bring it to the DNA template.

RNA polymerase I is located in the nucleolus, a specialized nuclear substructure in which ribosomal RNA (rRNA) is

transcribed, processed, and assembled into ribosomes (Table 15.1). The rRNA molecules are considered structural RNAs because they have a cellular role but are not translated into protein. The rRNAs are components of the ribosome and are essential to the process of translation. RNA polymerase I synthesizes all of the rRNAs except for the 5S rRNA molecule. The "S" designation applies to "Svedberg" units, a nonadditive value that characterizes the speed at which a particle sediments during centrifugation.

Locations, Products, and Sensitivities of the Three Eukaryotic RNA Polymerases

RNA Polymerase	Cellular Compartment	Product of Transcription	α-Amanitin Sensitivity
I	Nucleolus	All rRNAs except 5S rRNA	Insensitive
II	Nucleus	All protein-coding nuclear pre-mRNAs	Extremely sensitive
III	Nucleus	5S rRNA, tRNAs, and small nuclear RNAs	Moderately sensitive

Table 15.1

RNA polymerase II is located in the nucleus and synthesizes all protein-coding nuclear pre-mRNAs. Eukaryotic pre-mRNAs undergo extensive processing after transcription but before translation. For clarity, this module's discussion of transcription and translation in eukaryotes will use the term "mRNAs" to describe only the mature, processed molecules that are ready to be translated. RNA polymerase II is responsible for transcribing the overwhelming majority of eukaryotic genes.

RNA polymerase III is also located in the nucleus. This polymerase transcribes a variety of structural RNAs that includes the 5S pre-rRNA, transfer pre-RNAs (pre-tRNAs), and **small nuclear** pre- **RNAs**. The tRNAs have a critical role in translation; they serve as the adaptor molecules between the mRNA template and the growing polypeptide chain. Small nuclear RNAs have a variety of functions, including "splicing" pre-mRNAs and regulating transcription factors.

A scientist characterizing a new gene can determine which polymerase transcribes it by testing whether the gene is expressed in the presence of a particular mushroom poison, α-amanitin (Table 15.1). Interestingly, α-amanitin produced by *Amanita phalloides*, the Death Cap mushroom, affects the three polymerases very differently. RNA polymerase I is completely insensitive to α-amanitin, meaning that the polymerase can transcribe DNA in vitro in the presence of this poison. In contrast, RNA polymerase II is extremely sensitive to α-amanitin, and RNA polymerase III is moderately sensitive. Knowing the transcribing polymerase can clue a researcher into the general function of the gene being studied. Because RNA polymerase II transcribes the vast majority of genes, we will focus on this polymerase in our subsequent discussions about eukaryotic transcription factors and promoters.

Structure of an RNA Polymerase II Promoter

Eukaryotic promoters are much larger and more complex than prokaryotic promoters, but both have a TATA box. For example, in the mouse thymidine kinase gene, the TATA box is located at approximately -30 relative to the initiation (+1) site (Figure 15.10). For this gene, the exact TATA box sequence is TATAAAA, as read in the 5' to 3' direction on the nontemplate strand. This sequence is not identical to the *E. coli* TATA box, but it conserves the A–T rich element. The thermostability of A–T bonds is low and this helps the DNA template to locally unwind in preparation for transcription.

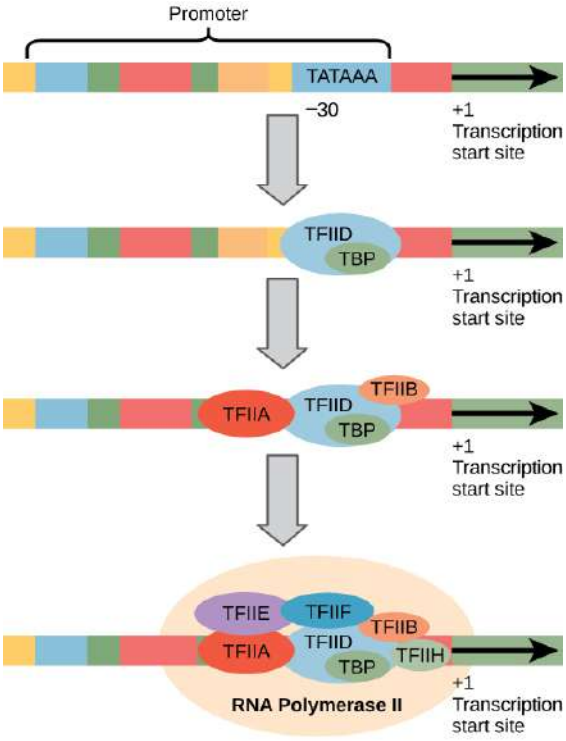

Figure 15.10 A generalized promoter of a gene transcribed by RNA polymerase II is shown. Transcription factors recognize the promoter. RNA polymerase II then binds and forms the transcription initiation complex.

visual CONNECTION

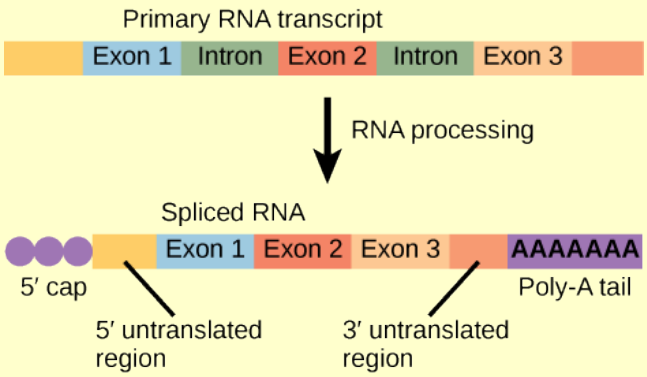

Figure 15.11 Eukaryotic mRNA contains introns that must be spliced out. A 5' cap and 3' poly-A tail are also added.

A scientist splices a eukaryotic promoter in front of a bacterial gene and inserts the gene in a bacterial chromosome. Would you expect the bacterium to transcribe the gene?

a. Initially the bacterium will not transcribe it, but will transcribe after 5’ capping.
b. It will not transcribe it because there is no poly A tail in prokaryotes.
c. No, prokaryotes use different promoters than eukaryotes.
d. Yes, the bacterium would transcribe the eukaryotic gene.

The mouse genome includes one gene and two pseudogenes for cytoplasmic thymidine kinase. Pseudogenes are genes that have lost their protein-coding ability or are no longer expressed by the cell. These pseudogenes are copied from mRNA and

incorporated into the chromosome. For example, the mouse thymidine kinase promoter also has a conserved **CAAT box** (GGCCAATCT) at approximately -80. This sequence is essential and is involved in binding transcription factors. Further upstream of the TATA box, eukaryotic promoters may also contain one or more **GC-rich boxes** (GGCG) or **octamer boxes** (ATTTGCAT). These elements bind cellular factors that increase the efficiency of transcription initiation and are often identified in more "active" genes that are constantly being expressed by the cell.

Transcription Factors for RNA Polymerase II

The complexity of eukaryotic transcription does not end with the polymerases and promoters. An army of basal transcription factors, enhancers, and silencers also help to regulate the frequency with which pre-mRNA is synthesized from a gene. Enhancers and silencers affect the efficiency of transcription but are not necessary for transcription to proceed. Basal transcription factors are crucial in the formation of a **preinitiation complex** on the DNA template that subsequently recruits RNA polymerase II for transcription initiation.

The names of the basal transcription factors begin with "TFII" (this is the transcription factor for RNA polymerase II) and are specified with the letters A–J. The transcription factors systematically fall into place on the DNA template, with each one further stabilizing the preinitiation complex and contributing to the recruitment of RNA polymerase II.

The processes of bringing RNA polymerases I and III to the DNA template involve slightly less complex collections of transcription factors, but the general theme is the same. Eukaryotic transcription is a tightly regulated process that requires a variety of proteins to interact with each other and with the DNA strand. Although the process of transcription in eukaryotes involves a greater metabolic investment than in prokaryotes, it ensures that the cell transcribes precisely the pre-mRNAs that it needs for protein synthesis.

everyday CONNECTION for AP® Courses

During human embryonic development, a transcription factor encoded by the *SRY* gene starts a chain of events, causing the embryo to develop male sex characteristics. This gene is on the Y chromosome in humans and many other mammals. A deletion or mutation of the *SRY* gene can cause the human embryo to not develop into a male even though the individual has an XY genotype, a condition called Swyer syndrome.

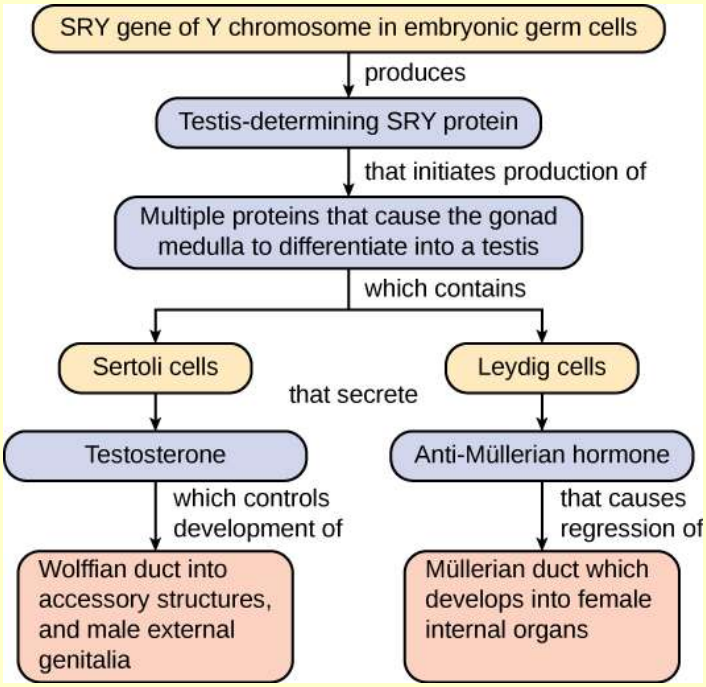

Figure 15.12 The SYR gene of the Y chromosome produces proteins that lead to the expression of primary sex characteristics, as shown.

The protein product of the SRY gene is a DNA binding protein. Together with a protein called SF1, the SRY protein acts as a transcription factor that "turns on" certain genes. Which of the following statements best describes how a change in these two proteins would affect male sexual development?

a. A mutation that abolished activity of SF1 would increase the effect of a SRY mutation, making the person more feminine.

b. A mutation that abolished activity of SF1 would cancel out a mutation in SRY, so if both mutations occur together male sex characteristics would develop normally.

c. A mutation in the SRY protein that abolished activity would result in abnormal development of male sex characteristics but a mutation of SF1 would not.

d. Both a mutation in the SRY protein and a mutation in SF1 that abolished activity would result in a lack of development of male sex characteristics.

evolution CONNECTION

The Evolution of Promoters

The evolution of genes may be a familiar concept. Mutations can occur in genes during DNA replication, and the result may or may not be beneficial to the cell. By altering an enzyme, structural protein, or some other factor, the process of mutation can transform functions or physical features. However, eukaryotic promoters and other gene regulatory sequences may evolve as well. For instance, consider a gene that, over many generations, becomes more valuable to the cell. Maybe the gene encodes a structural protein that the cell needs to synthesize in abundance for a certain function. If this is the case, it would be beneficial to the cell for that gene's promoter to recruit transcription factors more efficiently and increase gene expression.

Scientists examining the evolution of promoter sequences have reported varying results. In part, this is because it is difficult to infer exactly where a eukaryotic promoter begins and ends. Some promoters occur within genes; others are located very far upstream, or even downstream, of the genes they are regulating. However, when researchers limited their examination to human core promoter sequences that were defined experimentally as sequences that bind the preinitiation complex, they found that promoters evolve even faster than protein-coding genes.

It is still unclear how promoter evolution might correspond to the evolution of humans or other higher organisms. However, the evolution of a promoter to effectively make more or less of a given gene product is an intriguing alternative to the evolution of the genes themselves.[1]

According to this passage, which of the following has been shown to evolve faster than protein-coding genes?

a. core promoters that bind the preinitiation complex
b. core promoters that occur within genes
c. promoters that occur far upstream of the gene
d. promoters that occur downstream of a gene

Promoter Structures for RNA Polymerases I and III

In eukaryotes, the conserved promoter elements differ for genes transcribed by RNA polymerases I, II, and III. RNA polymerase I transcribes genes that have two GC-rich promoter sequences in the -45 to +20 region. These sequences alone are sufficient for transcription initiation to occur, but promoters with additional sequences in the region from -180 to -105 upstream of the initiation site will further enhance initiation. Genes that are transcribed by RNA polymerase III have upstream promoters or promoters that occur within the genes themselves.

Eukaryotic Elongation and Termination

Following the formation of the preinitiation complex, the polymerase is released from the other transcription factors, and elongation is allowed to proceed as it does in prokaryotes with the polymerase synthesizing pre-mRNA in the 5' to 3' direction. As discussed previously, RNA polymerase II transcribes the major share of eukaryotic genes, so this section will focus on how this polymerase accomplishes elongation and termination.

Although the enzymatic process of elongation is essentially the same in eukaryotes and prokaryotes, the DNA template is more complex. When eukaryotic cells are not dividing, their genes exist as a diffuse mass of DNA and proteins called chromatin. The DNA is tightly packaged around charged histone proteins at repeated intervals. These DNA–histone complexes, collectively called nucleosomes, are regularly spaced and include 146 nucleotides of DNA wound around eight histones like thread around a spool.

For polynucleotide synthesis to occur, the transcription machinery needs to move histones out of the way every time it encounters a nucleosome. This is accomplished by a special protein complex called **FACT**, which stands for "facilitates chromatin transcription." This complex pulls histones away from the DNA template as the polymerase moves along it. Once the pre-mRNA is synthesized, the FACT complex replaces the histones to recreate the nucleosomes.

The termination of transcription is different for the different polymerases. Unlike in prokaryotes, elongation by RNA polymerase II in eukaryotes takes place 1,000–2,000 nucleotides beyond the end of the gene being transcribed. This pre-mRNA tail is subsequently removed by cleavage during mRNA processing. On the other hand, RNA polymerases I and

1. H Liang et al., "Fast evolution of core promoters in primate genomes," *Molecular Biology and Evolution* 25 (2008): 1239–44.

III require termination signals. Genes transcribed by RNA polymerase I contain a specific 18-nucleotide sequence that is recognized by a termination protein. The process of termination in RNA polymerase III involves an mRNA hairpin similar to rho-independent termination of transcription in prokaryotes.

15.4 | RNA Processing in Eukaryotes

In this section, you will explore the following questions:

- What are the steps in eukaryotic transcription?
- What are the structural and functional similarities and differences among the three RNA polymerases?

Connection for AP® Courses

Scientists discovered a strand of mRNA translated into a sequence of amino acids (polypeptide) shorter than the mRNA molecule transcribed from DNA. Before the information in eukaryotic mRNA is translated into protein, it is modified or edited in several ways. A 5′ methylguanosine (or GTP) cap and a 3′ poly-A tail are added to protect mature mRNA from degradation and allow its export from the nucleus. Pre-mRNAs also undergo splicing, in which introns are removed and exons are reconnected. Exons can be reconnected in different sequences, a phenomenon referred to as alternative gene splicing, which allows a single eukaryotic gene to code for different proteins. (We will explore the significance of alternative gene splicing in more detail in other chapters.)

Information presented and the examples highlighted in the section support concepts outlined in Big Idea 3 of the AP® Biology Curriculum Framework. The Learning Objectives listed in the Curriculum Framework provide a transparent foundation for the AP® Biology course, an inquiry-based laboratory experience, instructional activities, and AP® Exam questions. A Learning Objective merges required content with one or more of the seven Science Practices.

Big Idea 3	Living systems store, retrieve, transmit and respond to information essential to life processes.
Enduring Understanding 3.A	Heritable information provides for continuity of life.
Essential Knowledge	**3.A.1** DNA, and in some cases RNA, is the primary source of heritable information.
Science Practice	**6.5** The student can evaluate alternative scientific explanations.
Learning Objective	**3.1** The student is able to construct scientific explanations that use the structures and mechanisms of DNA and RNA to support the claim that DNA and, in some cases, that RNA are the primary sources of heritable information.

After transcription, eukaryotic pre-mRNAs must undergo several processing steps before they can be translated. Eukaryotic (and prokaryotic) tRNAs and rRNAs also undergo processing before they can function as components in the protein synthesis machinery.

mRNA Processing

The eukaryotic pre-mRNA undergoes extensive processing before it is ready to be translated. The additional steps involved in eukaryotic mRNA maturation create a molecule with a much longer half-life than a prokaryotic mRNA. Eukaryotic mRNAs last for several hours, whereas the typical *E. coli* mRNA lasts no more than five seconds.

Pre-mRNAs are first coated in RNA-stabilizing proteins; these protect the pre-mRNA from degradation while it is processed and exported out of the nucleus. The three most important steps of pre-mRNA processing are the addition of stabilizing and signaling factors at the 5' and 3' ends of the molecule, and the removal of intervening sequences that do not specify the appropriate amino acids. In rare cases, the mRNA transcript can be “edited” after it is transcribed.

evolution CONNECTION

RNA Editing in Trypanosomes

The trypanosomes are a group of protozoa that include the pathogen *Trypanosoma brucei*, which causes sleeping sickness in humans (**Figure 15.13**). Trypanosomes, and virtually all other eukaryotes, have organelles called mitochondria that supply the cell with chemical energy. Mitochondria are organelles that express their own DNA and are believed to be the remnants of a symbiotic relationship between a eukaryote and an engulfed prokaryote. The mitochondrial DNA of trypanosomes exhibit an interesting exception to The Central Dogma: their pre-mRNAs do not have the correct information to specify a functional protein. Usually, this is because the mRNA is missing several U nucleotides. The cell performs an additional RNA processing step called **RNA editing** to remedy this.

Figure 15.13 *Trypanosoma brucei* is the causative agent of sleeping sickness in humans. The mRNAs of this pathogen must be modified by the addition of nucleotides before protein synthesis can occur. (credit: modification of work by Torsten Ochsenreiter)

Other genes in the mitochondrial genome encode 40- to 80-nucleotide guide RNAs. One or more of these molecules interacts by complementary base pairing with some of the nucleotides in the pre-mRNA transcript. However, the guide RNA has more A nucleotides than the pre-mRNA has U nucleotides to bind with. In these regions, the guide RNA loops out. The 3' ends of guide RNAs have a long poly-U tail, and these U bases are inserted in regions of the pre-mRNA transcript at which the guide RNAs are looped. This process is entirely mediated by RNA molecules. That is, guide RNAs—rather than proteins—serve as the catalysts in RNA editing.

RNA editing is not just a phenomenon of trypanosomes. In the mitochondria of some plants, almost all pre-mRNAs are edited. RNA editing has also been identified in mammals such as rats, rabbits, and even humans. What could be the evolutionary reason for this additional step in pre-mRNA processing? One possibility is that the mitochondria, being remnants of ancient prokaryotes, have an equally ancient RNA-based method for regulating gene expression. In support of this hypothesis, edits made to pre-mRNAs differ depending on cellular conditions. Although speculative, the process of RNA editing may be a holdover from a primordial time when RNA molecules, instead of proteins, were responsible for catalyzing reactions.

In eukaryotes, pre-mRNAs are processed to form mature mRNAs. How does the mRNA editing that occurs in *Trypanosoma brucei* differ from mRNA processing that occurs in all eukaryotes?

a. mRNA editing changes the coding sequence of the mRNA, but mRNA processing does not.
b. mRNA editing splices out noncoding RNA, but mRNA processing does not.
c. mRNA editing adds a cap of 5’-methylguanosine to the mRNA, but mRNA processing does not.
d. mRNA editing adds a 3’ poly-A tail, but mRNA processing does not.

5' Capping

While the pre-mRNA is still being synthesized, a **7-methylguanosine cap** is added to the 5' end of the growing transcript by a phosphate linkage. This moiety (functional group) protects the nascent mRNA from degradation. In addition, factors involved in protein synthesis recognize the cap to help initiate translation by ribosomes.

3' Poly-A Tail

Once elongation is complete, the pre-mRNA is cleaved by an endonuclease between an AAUAAA consensus sequence and a GU-rich sequence, leaving the AAUAAA sequence on the pre-mRNA. An enzyme called poly-A polymerase then adds a string of approximately 200 A residues, called the **poly-A tail**. This modification further protects the pre-mRNA from degradation and signals the export of the cellular factors that the transcript needs to the cytoplasm.

Pre-mRNA Splicing

Eukaryotic genes are composed of **exons**, which correspond to protein-coding sequences (*ex*-on signifies that they are *ex*pressed), and *int*ervening sequences called **introns** (*int*-ron denotes their *int*ervening role), which may be involved in gene regulation but are removed from the pre-mRNA during processing. Intron sequences in mRNA do not encode functional proteins.

The discovery of introns came as a surprise to researchers in the 1970s who expected that pre-mRNAs would specify protein sequences without further processing, as they had observed in prokaryotes. The genes of higher eukaryotes very often contain one or more introns. These regions may correspond to regulatory sequences; however, the biological significance of having many introns or having very long introns in a gene is unclear. It is possible that introns slow down gene expression because it takes longer to transcribe pre-mRNAs with lots of introns. Alternatively, introns may be nonfunctional sequence remnants left over from the fusion of ancient genes throughout evolution. This is supported by the fact that separate exons often encode separate protein subunits or domains. For the most part, the sequences of introns can be mutated without ultimately affecting the protein product.

All of a pre-mRNA's introns must be completely and precisely removed before protein synthesis. If the process errs by even a single nucleotide, the reading frame of the rejoined exons would shift, and the resulting protein would be dysfunctional. The process of removing introns and reconnecting exons is called **splicing** (**Figure 15.14**). Introns are removed and degraded while the pre-mRNA is still in the nucleus. Splicing occurs by a sequence-specific mechanism that ensures introns will be removed and exons rejoined with the accuracy and precision of a single nucleotide. The splicing of pre-mRNAs is conducted by complexes of proteins and RNA molecules called spliceosomes.

visual CONNECTION

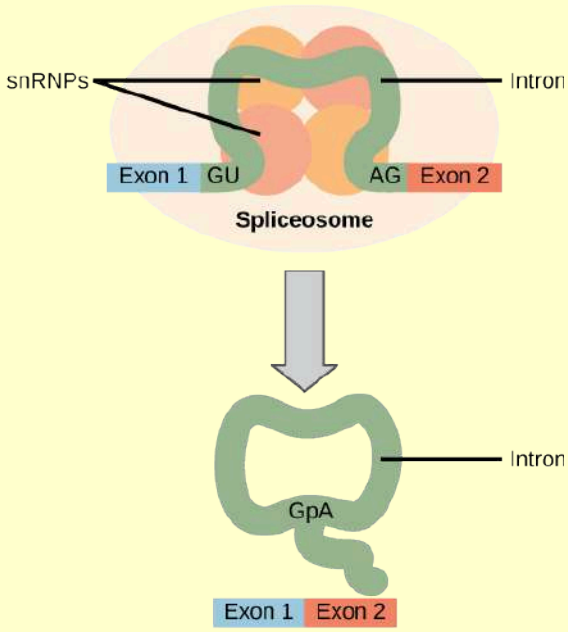

Figure 15.14 Pre-mRNA splicing involves the precise removal of introns from the primary RNA transcript. The splicing process is catalyzed by protein complexes called spliceosomes that are composed of proteins and RNA molecules called snRNAs. Spliceosomes recognize sequences at the 5' and 3' end of the intron.

Errors in splicing are implicated in cancers and other human diseases. What kinds of mutations might lead to splicing errors? Think of different possible outcomes if splicing errors occur.

a. Mutations in the spliceosome recognition sequence at each end of an intron, or in the proteins and RNAs that make up the spliceosome, may occur. Mutations may also add new spliceosome recognition sites.

b. Mutations in the spliceosome recognition sequence at each end of an exon, or in the proteins and RNAs that make up the spliceosome, may occur. Mutations may also add new spliceosome recognition sites.

c. Mutations in the spliceosome recognition sequence at each end of an intron, or in the proteins and RNAs that make up the spliceosome, may occur. Mutations may also delete existing spliceosome recognition sites.

d. Mutations at the each end of intron and exon, or in the proteins and RNAs that make up the spliceosome, may occur. Mutations may also add new spliceosome recognition sites and delete existing sites.

Note that more than 70 individual introns can be present, and each has to undergo the process of splicing—in addition to 5' capping and the addition of a poly-A tail—just to generate a single, translatable mRNA molecule.

See how introns are removed during RNA splicing at **this website (http://openstaxcollege.org/l/RNA_splicing)** .

Explain why helper proteins are necessary for the formation of the final protein during RNA splicing in higher organisms.

a. Helper proteins attach themselves to the ends of introns so that they can be spliced out during RNA splicing and coded areas are spliced together to form mRNA which then codes for the final protein.

b. Helper proteins attach themselves to the ends of exons so that they can be spliced out during RNA splicing and coded areas are spliced together to form mRNA which encodes the final protein.

c. Helper proteins attach themselves to mRNA in order to remove the non-coded areas and thus form the pre-mRNA which codes for the final protein.

d. Helper proteins help the pre-mRNA to recruit various other components which splice out the non-coded regions and form mRNA which codes for the final protein.

Processing of tRNAs and rRNAs

The tRNAs and rRNAs are structural molecules that have roles in protein synthesis; however, these RNAs are not themselves translated. Pre-rRNAs are transcribed, processed, and assembled into ribosomes in the nucleolus. Pre-tRNAs are transcribed and processed in the nucleus and then released into the cytoplasm where they are linked to free amino acids for protein synthesis.

Most of the tRNAs and rRNAs in eukaryotes and prokaryotes are first transcribed as a long precursor molecule that spans multiple rRNAs or tRNAs. Enzymes then cleave the precursors into subunits corresponding to each structural RNA. Some of the bases of pre-rRNAs are methylated; that is, a $–CH_3$ moiety (methyl functional group) is added for stability. Pre-tRNA molecules also undergo methylation. As with pre-mRNAs, subunit excision occurs in eukaryotic pre-RNAs destined to become tRNAs or rRNAs.

Mature rRNAs make up approximately 50 percent of each ribosome. Some of a ribosome's RNA molecules are purely structural, whereas others have catalytic or binding activities. Mature tRNAs take on a three-dimensional structure through intramolecular hydrogen bonding to position the amino acid binding site at one end and the **anticodon** at the other end (**Figure 15.15**). The anticodon is a three-nucleotide sequence in a tRNA that interacts with an mRNA codon through complementary base pairing.

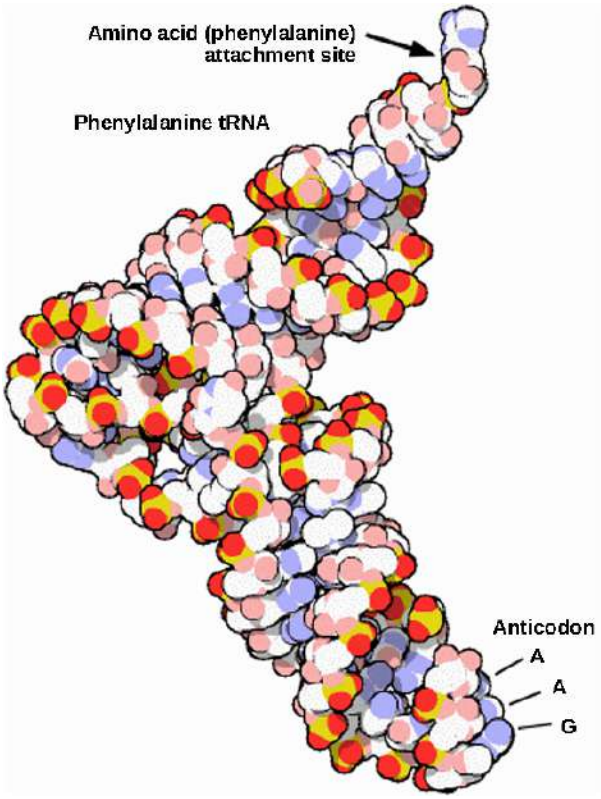

Figure 15.15 This is a space-filling model of a tRNA molecule that adds the amino acid phenylalanine to a growing polypeptide chain. The anticodon AAG binds the Codon UUC on the mRNA. The amino acid phenylalanine is attached to the other end of the tRNA.

15.5 | Ribosomes and Protein Synthesis

In this section, you will explore the following questions:

- What are the different sequential steps in protein synthesis?
- What is the role of ribosomes in protein synthesis?

Connection for AP® Courses

After the information in the gene has been transcribed to mRNA, it is ready to be translated to polypeptide. The players in translation include the mRNA template, ribosomes, tRNA molecules, amino acids, and various enzymes. Ribosomes consist of small and large subunits of protein and rRNA which bind with mRNA; many ribosomes can move along the same mRNA at a time. Translation begins at the initiating AUG on mRNA, specifying methionine, the first amino acid in any polypeptide. Each amino acid is carried to the ribosome by attaching to a specific molecule of tRNA. A tRNA molecule often is depicted as a cloverleaf, with an anticodon on one end, and the amino acid attachment site at the other. Amino-acid charging enzymes ensure that the correct amino acid is attached to the correct tRNA. The anticodons on tRNA are complementary to the codons on mRNA; for example, the anticodon AAA on tRNA corresponds to TTT on mRNA. Sequential amino acids are linked by peptide bonds. The mRNA is translated, elongating the polypeptide, until a STOP or nonsense codon is reached. When this happens, a release factor dissociates the components and frees the new polypeptide. Folding of the protein occurs during and after translation. Once a polypeptide is synthesized, its role as a protein is established, such as determining a physical phenotype of an organism.

Information presented and the examples highlighted in the section support concepts outlined in Big Idea 3 of the AP® Biology Curriculum Framework. The Learning Objectives listed in the Curriculum Framework provide a transparent foundation for the AP® Biology course, an inquiry-based laboratory experience, instructional activities, and AP® Exam questions. A Learning Objective merges required content with one or more of the seven Science Practices.

Big Idea 3	Living systems store, retrieve, transmit and respond to information essential to life processes.
Enduring Understanding 3.A	Heritable information provides for continuity of life.
Essential Knowledge	**3.A.1** DNA, and in some cases RNA, is the primary source of heritable information.
Science Practice	**1.2** The student can describe representations and models of natural or man-made phenomena and systems in the domain.
Learning Objective	**3.4** The student is able to describe representations and models illustrating how genetic information is translated into polypeptides.
Essential Knowledge	**3.A.1** DNA, and in some cases RNA, is the primary source of heritable information.
Science Practice	**6.4** The student can make claims and predictions about natural phenomena based on scientific theories and models.
Learning Objective	**3.6** The student can predict how a change in a specific DNA or RNA sequence can result in changes in gene expression.

The Science Practice Challenge Questions contain additional test questions for this section that will help you prepare for the AP exam. These questions address the following standards:
[APLO 1.16][APLO 4.22][APLO 3.6]

The synthesis of proteins consumes more of a cell's energy than any other metabolic process. In turn, proteins account for more mass than any other component of living organisms (with the exception of water), and proteins perform virtually every function of a cell. The process of translation, or protein synthesis, involves the decoding of an mRNA message into a polypeptide product. Amino acids are covalently strung together by interlinking peptide bonds in lengths ranging from approximately 50 amino acid residues to more than 1,000. Each individual amino acid has an amino group (NH_2) and a carboxyl (COOH) group. Polypeptides are formed when the amino group of one amino acid forms an amide (i.e., peptide) bond with the carboxyl group of another amino acid (**Figure 15.16**). This reaction is catalyzed by ribosomes and generates one water molecule.

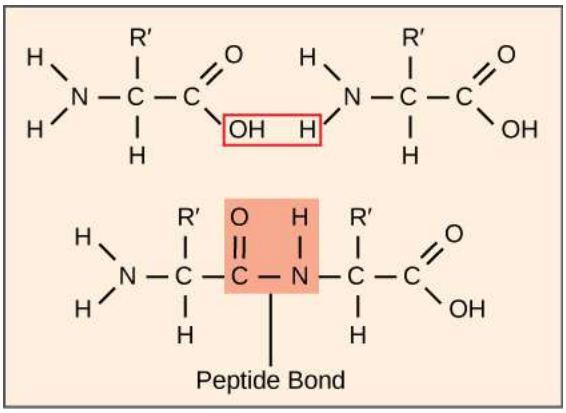

Figure 15.16 A peptide bond links the carboxyl end of one amino acid with the amino end of another, expelling one water molecule. For simplicity in this image, only the functional groups involved in the peptide bond are shown. The R and R' designations refer to the rest of each amino acid structure.

The Protein Synthesis Machinery

In addition to the mRNA template, many molecules and macromolecules contribute to the process of translation. The composition of each component may vary across species; for instance, ribosomes may consist of different numbers of rRNAs and polypeptides depending on the organism. However, the general structures and functions of the protein synthesis machinery are comparable from bacteria to human cells. Translation requires the input of an mRNA template, ribosomes, tRNAs, and various enzymatic factors.

Click through the steps of this **PBS interactive (http://openstaxcollege.org/l/prokary_protein)** to see protein synthesis in action.

A lack of protein in the diet can cause hair loss. Explain why this occurs.

a. Due to lack of protein in the diet, our body will not be able to form other proteins; thus, it will conserve the protein it has for critical use, leading to hair loss.

b. Lack of protein in the diet can weaken the immune system, thus leading to hair loss.

c. Due to lack of protein in the diet, energy will be lost, thus leading to hair loss.

d. Lack of protein in the diet will lead to breakage of disulfide bonds between proteins, thus leading to hair loss.

Ribosomes

Even before an mRNA is translated, a cell must invest energy to build each of its ribosomes. In *E. coli*, there are between 10,000 and 70,000 ribosomes present in each cell at any given time. A ribosome is a complex macromolecule composed of structural and catalytic rRNAs, and many distinct polypeptides. In eukaryotes, the nucleolus is completely specialized for the synthesis and assembly of rRNAs.

Ribosomes exist in the cytoplasm in prokaryotes and in the cytoplasm and rough endoplasmic reticulum in eukaryotes. Mitochondria and chloroplasts also have their own ribosomes in the matrix and stroma, which look more similar to prokaryotic ribosomes (and have similar drug sensitivities) than the ribosomes just outside their outer membranes in the cytoplasm. Ribosomes dissociate into large and small subunits when they are not synthesizing proteins and reassociate during the initiation of translation. In *E. coli*, the small subunit is described as 30S, and the large subunit is 50S, for a total of 70S (recall that Svedberg units are not additive). Mammalian ribosomes have a small 40S subunit and a large 60S subunit, for a total of 80S. The small subunit is responsible for binding the mRNA template, whereas the large subunit sequentially binds tRNAs. Each mRNA molecule is simultaneously translated by many ribosomes, all synthesizing protein in the same direction: reading the mRNA from 5' to 3' and synthesizing the polypeptide from the N terminus to the C terminus. The complete mRNA/poly-ribosome structure is called a **polysome**.

tRNAs

The tRNAs are structural RNA molecules that were transcribed from genes by RNA polymerase III. Depending on the species, 40 to 60 types of tRNAs exist in the cytoplasm. Serving as adaptors, specific tRNAs bind to sequences on the mRNA template and add the corresponding amino acid to the polypeptide chain. Therefore, tRNAs are the molecules that actually “translate” the language of RNA into the language of proteins.

Of the 64 possible mRNA codons—or triplet combinations of A, U, G, and C—three specify the termination of protein synthesis and 61 specify the addition of amino acids to the polypeptide chain. Of these 61, one codon (AUG) also encodes the initiation of translation. Each tRNA anticodon can base pair with one of the mRNA codons and add an amino acid or terminate translation, according to the genetic code. For instance, if the sequence CUA occurred on an mRNA template in the proper reading frame, it would bind a tRNA expressing the complementary sequence, GAU, which would be linked to the amino acid leucine.

As the adaptor molecules of translation, it is surprising that tRNAs can fit so much specificity into such a small package. Consider that tRNAs need to interact with three factors: 1) they must be recognized by the correct aminoacyl synthetase (see below); 2) they must be recognized by ribosomes; and 3) they must bind to the correct sequence in mRNA.

Aminoacyl tRNA Synthetases

The process of pre-tRNA synthesis by RNA polymerase III only creates the RNA portion of the adaptor molecule. The corresponding amino acid must be added later, once the tRNA is processed and exported to the cytoplasm. Through the process of tRNA “charging,” each tRNA molecule is linked to its correct amino acid by a group of enzymes called

aminoacyl tRNA synthetases. At least one type of aminoacyl tRNA synthetase exists for each of the 20 amino acids; the exact number of aminoacyl tRNA synthetases varies by species. These enzymes first bind and hydrolyze ATP to catalyze a high-energy bond between an amino acid and adenosine monophosphate (AMP); a pyrophosphate molecule is expelled in this reaction. The activated amino acid is then transferred to the tRNA, and AMP is released.

The Mechanism of Protein Synthesis

As with mRNA synthesis, protein synthesis can be divided into three phases: initiation, elongation, and termination. The process of translation is similar in prokaryotes and eukaryotes. Here we'll explore how translation occurs in *E. coli*, a representative prokaryote, and specify any differences between prokaryotic and eukaryotic translation.

Initiation of Translation

Protein synthesis begins with the formation of an initiation complex. In *E. coli*, this complex involves the small 30S ribosome, the mRNA template, three initiation factors (IFs; IF-1, IF-2, and IF-3), and a special **initiator tRNA**, called $\text{tRNA}_{\text{f}}^{\text{Met}}$. The initiator tRNA interacts with the **start codon** AUG (or rarely, GUG), links to a formylated methionine called fMet, and can also bind IF-2. Formylated methionine is inserted by $\text{fMet} - \text{tRNA}_{\text{f}}^{\text{Met}}$ at the beginning of every polypeptide chain synthesized by *E. coli*, but it is usually clipped off after translation is complete. When an in-frame AUG is encountered during translation elongation, a non-formylated methionine is inserted by a regular Met-tRNA^{Met}.

In *E. coli* mRNA, a sequence upstream of the first AUG codon, called the **Shine-Dalgarno sequence** (AGGAGG), interacts with the rRNA molecules that compose the ribosome. This interaction anchors the 30S ribosomal subunit at the correct location on the mRNA template. Guanosine triphosphate (GTP), which is a purine nucleotide triphosphate, acts as an energy source during translation—both at the start of elongation and during the ribosome's translocation.

In eukaryotes, a similar initiation complex forms, comprising mRNA, the 40S small ribosomal subunit, IFs, and nucleoside triphosphates (GTP and ATP). The charged initiator tRNA, called Met-tRNA_{i}, does not bind fMet in eukaryotes, but is distinct from other Met-tRNAs in that it can bind IFs.

Instead of depositing at the Shine-Dalgarno sequence, the eukaryotic initiation complex recognizes the 7-methylguanosine cap at the 5' end of the mRNA. A cap-binding protein (CBP) and several other IFs assist the movement of the ribosome to the 5' cap. Once at the cap, the initiation complex tracks along the mRNA in the 5' to 3' direction, searching for the AUG start codon. Many eukaryotic mRNAs are translated from the first AUG, but this is not always the case. According to **Kozak's rules**, the nucleotides around the AUG indicate whether it is the correct start codon. Kozak's rules state that the following consensus sequence must appear around the AUG of vertebrate genes: 5'-gccRccAUGG-3'. The R (for purine) indicates a site that can be either A or G, but cannot be C or U. Essentially, the closer the sequence is to this consensus, the higher the efficiency of translation.

Once the appropriate AUG is identified, the other proteins and CBP dissociate, and the 60S subunit binds to the complex of Met-tRNA_{i}, mRNA, and the 40S subunit. This step completes the initiation of translation in eukaryotes.

Translation, Elongation, and Termination

In prokaryotes and eukaryotes, the basics of elongation are the same, so we will review elongation from the perspective of *E. coli*. The 50S ribosomal subunit of *E. coli* consists of three compartments: the A (aminoacyl) site binds incoming charged aminoacyl tRNAs. The P (peptidyl) site binds charged tRNAs carrying amino acids that have formed peptide bonds with the growing polypeptide chain but have not yet dissociated from their corresponding tRNA. The E (exit) site releases dissociated tRNAs so that they can be recharged with free amino acids. There is one exception to this assembly line of tRNAs: in *E. coli*, $\text{fMet} - \text{tRNA}_{\text{f}}^{\text{Met}}$ is capable of entering the P site directly without first entering the A site. Similarly, the eukaryotic Met-tRNA_{i}, with help from other proteins of the initiation complex, binds directly to the P site. In both cases, this creates an initiation complex with a free A site ready to accept the tRNA corresponding to the first codon after the AUG.

During translation elongation, the mRNA template provides specificity. As the ribosome moves along the mRNA, each mRNA codon comes into register, and specific binding with the corresponding charged tRNA anticodon is ensured. If mRNA were not present in the elongation complex, the ribosome would bind tRNAs nonspecifically.

Elongation proceeds with charged tRNAs entering the A site and then shifting to the P site followed by the E site with each single-codon "step" of the ribosome. Ribosomal steps are induced by conformational changes that advance the ribosome by three bases in the 3' direction. The energy for each step of the ribosome is donated by an elongation factor that hydrolyzes GTP. Peptide bonds form between the amino group of the amino acid attached to the A-site tRNA and the carboxyl group of the amino acid attached to the P-site tRNA. The formation of each peptide bond is catalyzed by **peptidyl transferase**, an RNA-based enzyme that is integrated into the 50S ribosomal subunit. The energy for each peptide bond formation is

derived from GTP hydrolysis, which is catalyzed by a separate elongation factor. The amino acid bound to the P-site tRNA is also linked to the growing polypeptide chain. As the ribosome steps across the mRNA, the former P-site tRNA enters the E site, detaches from the amino acid, and is expelled (**Figure 15.17**). Amazingly, the *E. coli* translation apparatus takes only 0.05 seconds to add each amino acid, meaning that a 200-amino acid protein can be translated in just 10 seconds.

visual CONNECTION

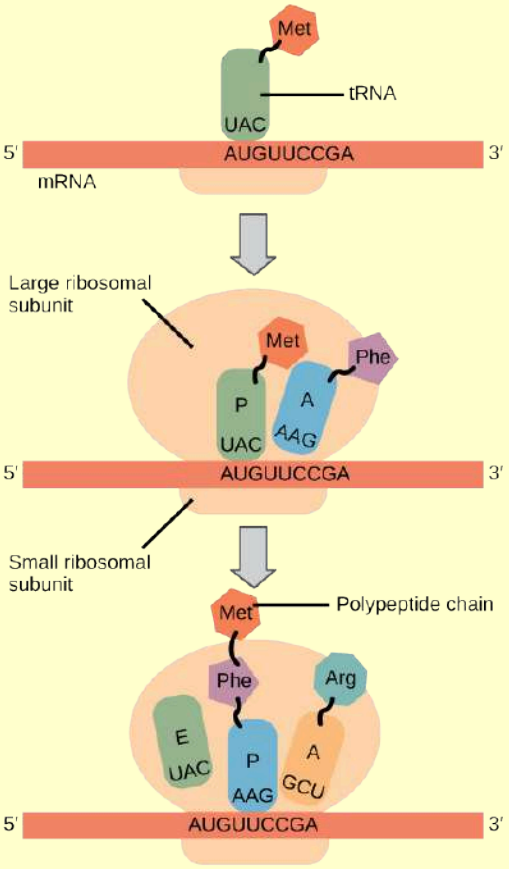

Figure 15.17 Translation begins when an initiator tRNA anticodon recognizes a codon on mRNA. The large ribosomal subunit joins the small subunit, and a second tRNA is recruited. As the mRNA moves relative to the ribosome, the polypeptide chain is formed. Entry of a release factor into the A site terminates translation and the components dissociate.

Many antibiotics inhibit bacterial protein synthesis. For example, tetracycline blocks the A site on the bacterial ribosome, and chloramphenicol blocks peptidyl transfer. What specific effect would you expect each of these antibiotics to have on protein synthesis?

Tetracycline would directly affect:

1. tRNA binding to the ribosome
2. Ribosome assembly
3. Growth of the protein chain

Chloramphenicol would directly affect:

1. tRNA binding to the ribosome
2. Ribosome assembly
3. Growth of the protein chain
    a. Tetracycline would directly affect tRNA binding to the ribosome. Chloramphenicol would affect the growth of the protein chain.
    b. Tetracycline would directly affect ribosome assembly. Chloramphenicol would affect the growth of the protein chain.

c. Tetracycline would directly affect the growth of the protein chain. Chloramphenicol would affect the tRNA binding to the ribosome.

d. Tetracycline would directly affect mRNA binding to the ribosome. Chloramphenicol would affect the ribosome assembly.

Termination of translation occurs when a nonsense codon (UAA, UAG, or UGA) is encountered. Upon aligning with the A site, these nonsense codons are recognized by release factors in prokaryotes and eukaryotes that instruct peptidyl transferase to add a water molecule to the carboxyl end of the P-site amino acid. This reaction forces the P-site amino acid to detach from its tRNA, and the newly made protein is released. The small and large ribosomal subunits dissociate from the mRNA and from each other; they are recruited almost immediately into another translation initiation complex. After many ribosomes have completed translation, the mRNA is degraded so the nucleotides can be reused in another transcription reaction.

Protein Folding, Modification, and Targeting

During and after translation, individual amino acids may be chemically modified, signal sequences may be appended, and the new protein "folds" into a distinct three-dimensional structure as a result of intramolecular interactions. A **signal sequence** is a short tail of amino acids that directs a protein to a specific cellular compartment. These sequences at the amino end or the carboxyl end of the protein can be thought of as the protein's "train ticket" to its ultimate destination. Other cellular factors recognize each signal sequence and help transport the protein from the cytoplasm to its correct compartment. For instance, a specific sequence at the amino terminus will direct a protein to the mitochondria or chloroplasts (in plants). Once the protein reaches its cellular destination, the signal sequence is usually clipped off.

Many proteins fold spontaneously, but some proteins require helper molecules, called chaperones, to prevent them from aggregating during the complicated process of folding. Even if a protein is properly specified by its corresponding mRNA, it could take on a completely dysfunctional shape if abnormal temperature or pH conditions prevent it from folding correctly.

science practices CONNECTION for AP® Courses

Activity

- Working in a small group, create a simple board game to model the key steps of transcription and translation and have classmates spend ten minutes playing the game.
- Provided with incomplete or incorrect diagrams illustrating transcription and translation in prokaryotes, have students refine or revise the diagrams and share the edited versions with classmates for critical review.

Think About It

- Many antibiotics inhibit protein synthesis. For example, tetracycline blocks the A site on the ribosome. What is the likely effect of tetracycline on protein synthesis?
- Using a chart of codons, transcribe and translate the following DNA sequence (non-template strand): 5′-ATGGCCGGTTATTAAGCA-3′. How can a single nucleotide change affect the protein produced from this sequence and its function?

KEY TERMS

7-methylguanosine cap modification added to the 5' end of pre-mRNAs to protect mRNA from degradation and assist translation

aminoacyl tRNA synthetase enzyme that "charges" tRNA molecules by catalyzing a bond between the tRNA and a corresponding amino acid

anticodon three-nucleotide sequence in a tRNA molecule that corresponds to an mRNA codon

CAAT box (GGCCAATCT) essential eukaryotic promoter sequence involved in binding transcription factors

Central Dogma states that genes specify the sequence of mRNAs, which in turn specify the sequence of proteins

codon three consecutive nucleotides in mRNA that specify the insertion of an amino acid or the release of a polypeptide chain during translation

colinear in terms of RNA and protein, three "units" of RNA (nucleotides) specify one "unit" of protein (amino acid) in a consecutive fashion

consensus DNA sequence that is used by many species to perform the same or similar functions

core enzyme prokaryotic RNA polymerase consisting of α, α, β, and β' but missing σ; this complex performs elongation

degeneracy (of the genetic code) describes that a given amino acid can be encoded by more than one nucleotide triplet; the code is degenerate, but not ambiguous

downstream nucleotides following the initiation site in the direction of mRNA transcription; in general, sequences that are toward the 3' end relative to a site on the mRNA

exon sequence present in protein-coding mRNA after completion of pre-mRNA splicing

FACT complex that "facilitates chromatin transcription" by disassembling nucleosomes ahead of a transcribing RNA polymerase II and reassembling them after the polymerase passes by

GC-rich box (GGCG) nonessential eukaryotic promoter sequence that binds cellular factors to increase the efficiency of transcription; may be present several times in a promoter

hairpin structure of RNA when it folds back on itself and forms intramolecular hydrogen bonds between complementary nucleotides

holoenzyme prokaryotic RNA polymerase consisting of α, α, β, β', and σ; this complex is responsible for transcription initiation

initiation site nucleotide from which mRNA synthesis proceeds in the 5' to 3' direction; denoted with a "+1"

initiator tRNA in prokaryotes, called $tRNA_f^{Met}$; in eukaryotes, called $tRNA_i$; a tRNA that interacts with a start codon, binds directly to the ribosome P site, and links to a special methionine to begin a polypeptide chain

intron non–protein-coding intervening sequences that are spliced from mRNA during processing

Kozak's rules determines the correct initiation AUG in a eukaryotic mRNA; the following consensus sequence must appear around the AUG: 5'-GCC(**purine**)CCAUG**G**-3'; the bolded bases are most important

nonsense codon one of the three mRNA codons that specifies termination of translation

nontemplate strand strand of DNA that is not used to transcribe mRNA; this strand is identical to the mRNA except that T nucleotides in the DNA are replaced by U nucleotides in the mRNA

Octamer box (ATTTGCAT) nonessential eukaryotic promoter sequence that binds cellular factors to increase the efficiency of transcription; may be present several times in a promoter

peptidyl transferase RNA-based enzyme that is integrated into the 50S ribosomal subunit and catalyzes the formation of peptide bonds

plasmid extrachromosomal, covalently closed, circular DNA molecule that may only contain one or a few genes; common in prokaryotes

poly-A tail modification added to the 3' end of pre-mRNAs to protect mRNA from degradation and assist mRNA export from the nucleus

polysome mRNA molecule simultaneously being translated by many ribosomes all going in the same direction

preinitiation complex cluster of transcription factors and other proteins that recruit RNA polymerase II for transcription of a DNA template

promoter DNA sequence to which RNA polymerase and associated factors bind and initiate transcription

reading frame sequence of triplet codons in mRNA that specify a particular protein; a ribosome shift of one or two nucleotides in either direction completely abolishes synthesis of that protein

Rho-dependent termination in prokaryotes, termination of transcription by an interaction between RNA polymerase and the rho protein at a run of G nucleotides on the DNA template

Rho-independent termination sequence-dependent termination of prokaryotic mRNA synthesis; caused by hairpin formation in the mRNA that stalls the polymerase

RNA editing direct alteration of one or more nucleotides in an mRNA that has already been synthesized

Shine-Dalgarno sequence (AGGAGG); initiates prokaryotic translation by interacting with rRNA molecules comprising the 30S ribosome

signal sequence short tail of amino acids that directs a protein to a specific cellular compartment

small nuclear RNA molecules synthesized by RNA polymerase III that have a variety of functions, including splicing pre-mRNAs and regulating transcription factors

splicing process of removing introns and reconnecting exons in a pre-mRNA

start codon AUG (or rarely, GUG) on an mRNA from which translation begins; always specifies methionine

TATA box conserved promoter sequence in eukaryotes and prokaryotes that helps to establish the initiation site for transcription

template strand strand of DNA that specifies the complementary mRNA molecule

transcription bubble region of locally unwound DNA that allows for transcription of mRNA

upstream nucleotides preceding the initiation site; in general, sequences toward the 5' end relative to a site on the mRNA

CHAPTER SUMMARY

15.1 The Genetic Code

The genetic code refers to the DNA alphabet (A, T, C, G), the RNA alphabet (A, U, C, G), and the polypeptide alphabet (20 amino acids). The Central Dogma describes the flow of genetic information in the cell from genes to mRNA to proteins. Genes are used to make mRNA by the process of transcription; mRNA is used to synthesize proteins by the process of translation. The genetic code is degenerate because 64 triplet codons in mRNA specify only 20 amino acids and three nonsense codons. Almost every species on the planet uses the same genetic code.

15.2 Prokaryotic Transcription

In prokaryotes, mRNA synthesis is initiated at a promoter sequence on the DNA template comprising two consensus sequences that recruit RNA polymerase. The prokaryotic polymerase consists of a core enzyme of four protein subunits

and a σ protein that assists only with initiation. Elongation synthesizes mRNA in the 5' to 3' direction at a rate of 40 nucleotides per second. Termination liberates the mRNA and occurs either by rho protein interaction or by the formation of an mRNA hairpin.

15.3 Eukaryotic Transcription

Transcription in eukaryotes involves one of three types of polymerases, depending on the gene being transcribed. RNA polymerase II transcribes all of the protein-coding genes, whereas RNA polymerase I transcribes rRNA genes, and RNA polymerase III transcribes rRNA, tRNA, and small nuclear RNA genes. The initiation of transcription in eukaryotes involves the binding of several transcription factors to complex promoter sequences that are usually located upstream of the gene being copied. The mRNA is synthesized in the 5' to 3' direction, and the FACT complex moves and reassembles nucleosomes as the polymerase passes by. Whereas RNA polymerases I and III terminate transcription by protein- or RNA hairpin-dependent methods, RNA polymerase II transcribes for 1,000 or more nucleotides beyond the gene template and cleaves the excess during pre-mRNA processing.

15.4 RNA Processing in Eukaryotes

Eukaryotic pre-mRNAs are modified with a 5' methylguanosine cap and a poly-A tail. These structures protect the mature mRNA from degradation and help export it from the nucleus. Pre-mRNAs also undergo splicing, in which introns are removed and exons are reconnected with single-nucleotide accuracy. Only finished mRNAs that have undergone 5' capping, 3' polyadenylation, and intron splicing are exported from the nucleus to the cytoplasm. Pre-rRNAs and pre-tRNAs may be processed by intramolecular cleavage, splicing, methylation, and chemical conversion of nucleotides. Rarely, RNA editing is also performed to insert missing bases after an mRNA has been synthesized.

15.5 Ribosomes and Protein Synthesis

The players in translation include the mRNA template, ribosomes, tRNAs, and various enzymatic factors. The small ribosomal subunit forms on the mRNA template either at the Shine-Dalgarno sequence (prokaryotes) or the 5' cap (eukaryotes). Translation begins at the initiating AUG on the mRNA, specifying methionine. The formation of peptide bonds occurs between sequential amino acids specified by the mRNA template according to the genetic code. Charged tRNAs enter the ribosomal A site, and their amino acid bonds with the amino acid at the P site. The entire mRNA is translated in three-nucleotide “steps” of the ribosome. When a nonsense codon is encountered, a release factor binds and dissociates the components and frees the new protein. Folding of the protein occurs during and after translation.

REVIEW QUESTIONS

1. What is the flow of information for the synthesis of proteins according to the central dogma?

a. DNA to mRNA to protein
b. DNA to mRNA to tRNA to protein
c. DNA to protein to mRNA to protein
d. mRNA to DNA to mRNA to protein

2. The DNA of virus A is inserted into the protein coat of virus B. The combination virus is used to infect *E. coli*. The virus particles produced by the infection are analyzed for DNA and protein contents. What results would you expect?

a. DNA and protein from B
b. DNA and protein from A
c. DNA from A and protein from B
d. DNA from B and protein from A

3. The AUC and AUA codons in mRNA both specify isoleucine. What feature of the genetic code explains this?

a. Complementarity
b. Degeneracy
c. Nonsense codons
d. Universality

4. How many nucleotides are in 12 mRNA codons?

a. 12
b. 24
c. 36
d. 48

5. Which of the following molecules does not contain genetic information?

a. DNA
b. mRNA
c. Protein
d. RNA

6. Which molecule in the central dogma can be compared to a disposable photocopy of a book kept on reserve in the

library?

a. DNA
b. mRNA
c. Protein
d. tRNA

7. Which subunit of the *E. coli* polymerase confers specificity to transcription?

a. α
b. β
c. β'
d. σ

8. Why are the -10 and -35 regions of prokaryotic promoters called consensus sequences?

a. They are identical in all bacterial species.
b. They are similar in all bacterial species.
c. They exist in all organisms.
d. They have the same function in all organisms.

9. The sequence that signals the end of transcription is called the:

a. promoter
b. stop codon
c. TATA box
d. terminator

10. If the ρ protein is missing, will a prokaryotic gene be terminated?

a. It depends on the gene.
b. No, the rho protein is essential.
c. Transcription termination is not required.
d. Yes, the rho protein is not involved in transcription.

11. Which feature of promoters can be found in both prokaryotes and eukaryotes?

a. GC box
b. octamer box
c. TATA box
d. -10 and -35 sequences

12. At what stage in the transcription of a eukaryotic gene would TFII factors be active?

a. elongation
b. initiation
c. processing
d. termination

13. Which polymerase is responsible for the synthesis of 5S rRNA?

a. polymerase I
b. polymerase II
c. polymerase III
d. ribonuclease I

14. What transcripts will be most affected by low levels of α-amanitin?

a. 18S and 28S rRNAs
b. 5S rRNAs and tRNAs
c. other small nuclear RNAs
d. pre-mRNAs

15. Which of the following features distinguishes eukaryotic transcription from bacterial transcription?

a. Eukaryotic transcription does not start at a consensus sequence.
b. Eukaryotic transcription does not require an initiation complex.
c. Eukaryotic transcription and translation do not take place at the same time.
d. Eukaryotic transcription does not require a termination sequence.

16. A poly-A sequence is added at the:

a. 5’ end of a transcript in the nucleus
b. 3’-end of a transcript in the nucleus
c. 5’ end of a transcript in the cytoplasm
d. 3’-end of a transcript in the cytoplasm

17. Which pre-mRNA processing step is important for initiating translation?

a. poly-A tail
b. RNA editing
c. splicing
d. 7-methylguanosine cap

18. Where are the RNA components of ribosomes synthesized?

a. cytoplasm
b. endoplasmic reticulum
c. nucleus
d. nucleolus

19. What processing step enhances the stability of pre-tRNAs and pre-rRNAs?

a. cleavage
b. methylation
c. nucleotide modification
d. splicing

20. What are introns?

a. DNA sequences to which polymerases bind
b. the processed mRNA
c. translated DNA sequences in a gene
d. untranslated DNA sequences in a gene

21. What is often the first amino acid added to a polypeptide chain?

a. adenine
b. leucine
c. methionine
d. thymine

22. In any given species, there are at least how many types of aminoacyl tRNA synthetases?

a. 20
b. 40
c. 100
d. 200

23. In prokaryotic cells, ribosomes are found in/on the:

a. cytoplasm
b. mitochondrion
c. nucleus
d. endoplasmic reticulum

24. The peptide bond synthesis in prokaryotic translation is catalyzed by:

a. a ribosomal protein
b. a cytoplasmic protein
c. mRNA itself
d. ribosomal RNA

25. What would happen if the 5’ methyl guanosine was not added to an mRNA?

a. The transcript would degrade when the mRNA moves out of the nucleus to the cytoplasm.
b. The mRNA molecule would stabilize and start the process of translation within the nucleus of the cell.
c. The mRNA molecule would move out of the nucleus and create more copies of the mRNA molecule.
d. The mRNA molecule would not be able to add the poly-A tail on its strand at the 5’ end.

26. Which of the following is associated with the docking of mRNA on a ribosome in eukaryotic cells?

a. Kozak’s sequence
b. poly-A sequence
c. Shine-Dalgarno sequence
d. TATA box

CRITICAL THINKING QUESTIONS

27. If mRNA is complementary to the DNA template strand and the DNA template stand is complementary to the DNA non-template strand, why are base sequences of mRNA and the DNA non-template strand not identical? Could they ever be?

a. No, they cannot be identical because the T nucleotide in DNA is replaced with U nucleotide in RNA and AUG is the start codon.
b. No, they cannot be identical because the T nucleotide in RNA is replaced with U nucleotide in DNA.
c. They can be identical if methylation of the U nucleotide in RNA occurs and gives T nucleotide.
d. They can be identical if de-methylation of the U nucleotide in RNA occurs and gives T nucleotide.

28. Imagine if there were 200 commonly occurring amino acids instead of 20. Given what you know about the genetic code, what would be the shortest possible codon length? Explain.

a. Four
b. Five
c. Two
d. Three

29. What part of central dogma is not always followed in viruses?

a. The flow of information in HIV is from RNA to DNA, then back to RNA to proteins. Influenza viruses never go through DNA.
b. The flow of information is from protein to RNA in HIV virus, while the influenza virus converts DNA to RNA.
c. The flow of information is similar, but nucleic acids are synthesized as a result of translation in HIV and influenza viruses.
d. The flow of information is from RNA to protein. This protein is used to synthesize the DNA of the viruses in HIV and influenza.

30. Suppose a gene has the sequence ATGCGTTATCGGGAGTAG. A point mutation changes the gene to read ATGCGTTATGGGGAGTAG. How

would the polypeptide product of this gene change?

31. Explain the initiation of transcription in prokaryotes. Include all proteins involved.

a. In prokaryotes the polymerase is composed of five polypeptide subunits, two of which are identical. Four of these subunits, denoted α, α, β, and β', comprise the polymerase core enzyme. The fifth subunit, σ, is involved only in transcription initiation. The polymerase comprised of all five subunits is called the holoenzyme.

b. In prokaryotes the polymerase is composed of four polypeptide subunits, two of which are identical. These subunits, denoted α, α, β, and β', comprise the polymerase core enzyme. There is a fifth subunit that is involved in translation initiation. The polymerase comprised of all four subunits is called the holoenzyme.

c. In prokaryotes the polymerase is composed of five polypeptide subunits, two of which are identical. Four of these subunits, denoted α, α, β, and β', comprise the polymerase holoenzyme. The fifth subunit, σ, is involved only in transcription initiation. The polymerase comprised of all five subunits is called the core enzyme.

d. In prokaryotes the polymerase is composed of five polypeptide subunits, two of which are identical. Four of these subunits, denoted α, α α, β, and β', comprise the polymerase core enzyme. The fifth subunit, σ, is involved only in termination. The polymerase comprised of all five subunits is called the holoenzyme.

32. In your own words, describe the difference between ρ-dependent and ρ-independent termination of transcription in prokaryotes.

a. Rho-dependent termination is controlled by rho protein and the polymerase stalls near the end of the gene at a run of G nucleotides on the DNA template. In rho-independent termination, when the polymerase encounters a region rich in C-G nucleotides the mRNA folds into a hairpin loop that causes the polymerase to stall.

b. Rho-independent termination is controlled by rho protein and the polymerase stalls near the end of the gene at a run of G nucleotides on the DNA template. In rho-dependent termination, when the polymerase encounters a region rich in C-G nucleotides, the mRNA folds into a hairpin loop that causes polymerase to stall.

c. Rho-dependent termination is controlled by rho protein and the polymerase begins near the end of the gene at a run of G nucleotides on the DNA template. In rho-independent termination, when the polymerase encounters a region rich in C-G nucleotides, the mRNA creates a hairpin loop that causes polymerase to stall.

d. Rho-dependent termination is controlled by rho protein and the polymerase stalls near the end of the gene at a run of G nucleotides on the DNA template. In rho-independent termination, when the polymerase encounters a region rich in A-T nucleotides, the mRNA creates a hairpin loop that causes polymerase to stall.

33. What is the main structure that differentiates between ρ-dependent and ρ-independent termination in prokaryotes?

a. Rho-independent termination involves the formation of a hairpin.

b. Rho-dependent termination involves the formation of a hairpin.

c. Rho-dependent termination stalls when the polymerase begins to transcribe a region rich in A-T nucleotides.

d. Rho-independent termination stalls when the polymerase begins to transcribe a region rich in G nucleotides.

34. Which step in the transcription of eukaryotic RNA differs the most from its prokaryotic counterpart?

a. The initiation step in eukaryotes requires an initiation complex with enhancers and transcription factors. Also, the separation of the DNA strand is different as histones are involved.

b. The initiation step in prokaryotes requires an initiation complex with enhancers and transcription factors. Also, the separation of the DNA strand is different as histones are involved.

c. The elongation step in eukaryotes requires an initiation complex with enhancers and transcription factors. Also, the separation of the DNA strand is different as histones are involved.

d. The initiation step in eukaryotes requires an initiation complex with enhancers and transcription factors. Also, the separation of the DNA strand is different as histones are not involved.

35. Would you be able to determine which RNA polymerase you isolated from a eukaryotic cell without analyzing its products?

a. No, because they have the same α-amanitin sensitivity in all products.

b. No, quantitative analysis of products is done to determine the type of polymerase.

c. Yes, they can be determined as they differ in α-amanitin sensitivity.

d. Yes, they can be determined by the number of molecules that bind to DNA.

36. Can you predict how alternative splicing may lead to an economy of genes? Do you need a different gene for every protein that the cell can produce?

a. Alternative splicing can lead to the synthesis of several polypeptides from a single gene.

b. Alternative splicing can lead to the synthesis of several forms of mRNA from a single gene.

c. Alternative splicing can lead to the synthesis of several forms of codons from a set of genes.

d. Alternative splicing can lead to the synthesis of several forms of ribosomes from a set of genes.

37. What is the major challenge in the production of RNA in eukaryotes compared to prokaryotes?

a. exporting the mRNA across the nuclear membrane

b. importing the mRNA across the nuclear membrane

c. the mRNA staying inside the nuclear membrane

d. the mRNA translating into proteins within seconds

38. What would happen if the 5’ methyl guanosine was not added to an mRNA?

a. The transcript would degrade when the mRNA moves out of the nucleus to the cytoplasm.

b. The mRNA molecule would stabilize and start the process of translation within the nucleus of the cell.

c. The mRNA molecule would move out of the nucleus and create more copies of the mRNA molecule.

d. The mRNA molecule would not be able to add the poly-A tail on its strand at the 5’ end.

39. Transcribe and translate the following DNA sequence (nontemplate strand): 5’-ATGGCCGGTTATTAAGCA-3’

a. The mRNA would be 5’-AUGGCCGGUUAUUAAGCA-3’ and the protein will be MAGY.

b. The mRNA would be 3’-AUGGCCGGUUAUUAAGCA-5’ and the protein will be MAGY.

c. The mRNA would be 5’-ATGGCCGGTTATTAAGCA-3’ and the protein will be MAGY.

d. The mRNA would be 5’-AUGGCCGGUUAUUAAGCA-3’ and the protein will be MACY.

40. The RNA world hypothesis proposes that the first complex molecule was RNA and it preceded protein formation. Which major function of the ribosomal RNA supports the hypothesis?

a. rRNA has catalytic properties in the large subunit and it assembles proteins.

b. rRNA is a protein molecule that helps in the synthesis of other proteins.

c. rRNA is essential for the transcription process.

d. rRNA plays a major role in post-translational processes.

41. A tRNA is chemically modified so that the amino acid bound is different than the one specified by its anticodon. Which codon in the mRNA would the tRNA recognize: the one specified by its anticodon or the one that matches the modified amino acid it carries?

a. The anticodon will match the codon in mRNA.

b. The anticodon will match with the modified amino acid it carries.

c. The anticodon will lose the specificity for the tRNA molecule.

d. The enzyme amino acyl tRNA synthetase would lose control over the amino acid.

TEST PREP FOR AP® COURSES

42. What characteristic of the genetic code points to a common ancestry for all organisms?

a. The code is degenerate
b. The code contains 64 codons.
c. The genetic code is almost universal.
d. The code contains stop codons

43. What process transfers heritable material to the next generation?

a. replication
b. splicing
c. transcription
d. translation

44. When comparing transcription of heritable information in prokaryotes and eukaryotes, which events are the same?

a. Transcription by polymerase, recognition of a consensus sequence in the promoter, and termination by a hairpin loop are conserved.
b. Translation by polymerase, recognition of a consensus sequence in the promoter, and termination by a hairpin loop are conserved.
c. Transcription by polymerase, recognition of a highly variable sequence in the promoter, and termination by a hairpin loop are conserved.
d. Transcription by polymerase, recognition of a consensus sequence in the promoter, and elongation by a hairpin loop are conserved.

45. Which of the following cell structures does not contain heritable information?

a. chloroplast
b. cytoplasmic membrane
c. mitochondria
d. nucleus

46. How does the enzyme reverse transcriptase violate the central dogma of molecular biology in HIV?

a. The enzyme reverse transcriptase reverse transcribes the RNA in the genome of HIV to DNA.
b. The enzyme reverse transcriptase translates the RNA of the HIV into protein and then back to DNA.
c. The enzyme reverse transcriptase transcribes the DNA straight into the protein molecules.
d. The enzyme reverse transcriptase transcribes DNA to RNA, then again to DNA. There is no protein synthesis.

47. Radioactive deoxythymidine triphosphate is supplied to the protist *Euglena*. After an interval of time, the cells are homogenized and different fractions are analyzed for radioactivity content in large nucleic acid molecules. Which fraction will not be labeled?

a. nucleus
b. mitochondrion
c. chloroplast
d. plasma membrane

48. You sequence a gene of interest and isolate the matching mRNA. You find that the mRNA is considerably shorter than the DNA sequence. Why is that?

a. There was an experimental mistake. The mRNA should have the same length as the gene.
b. The mRNA should be longer than the DNA sequence because the promoter is also transcribed.
c. The processed mRNA is shorter because introns were removed.
d. The mRNA is shorter because the signal sequence to cross the nuclear membrane was removed.

49. A mutation in the promoter region of the gene for the beta-globin can cause beta-thalassemia, a hereditary condition which causes anemia. Why would mutations in the promoter region lead to low levels of hemoglobin?

a. The globin chains produced are too long to form functional hemoglobin.
b. The globin chains are too short to form functional hemoglobin.
c. Fewer globin chains are synthesized because less mRNA is transcribed.
d. Globin chains do not fold properly and are non-functional.

50.

Codon on mRNA	Amino Acid
GCA	alanine
AAG	lysine
GUU	valine
AAU	asparagine
UGC	cysteine
UCG	serine
UCU	serine
UUA	leucine
UAA	stop

You are given three mRNA sequences:

1. 5’-UCG-GCA- AAU-UUA -GUU-3’
2. 5’-UCU-GCA- AAU-UUA -GUU-3’
3. 5’-UCU-GCA- AAU-UAA -GUU-3’

Using the table, write the peptide encoded by each of the mRNA sequences.

a. 1. Serine-alanine-asparagine-leucine-valine
   2. Serine-alanine-asparagine-leucine-valine
   3. Serine-alanine-asparagine(-stop)

b. 1. Serine-phenylalanine-asparagine-leucine-valine
   2. Serine-alanine-asparagine-leucine-valine
   3. Serine-alanine-asparagine (-stop)

c. 1. Serine-alanine-asparagine-leucine-valine
   2. Serine-alanine-asparagine (-stop)
   3. Serine-alanine-asparagine-leucine-valine

d. 1. Serine-alanine-asparagine-leucine-valine
   2. Serine-arginine-asparagine-leucine-valine
   3. Serine-alanine-asparagine(-stop)

51.

Codon on mRNA	Amino Acid
GCA	alanine
AAG	lysine
GUU	valine
AAU	asparagine
UGC	cysteine
UCG	serine
UCU	serine
UUA	leucine
UAA	stop

You are given three mRNA sequences:

1. 5’-UCG-GCA- AAU-UUA -GUU-3’
2. 5’-UCU-GCA- AAU-UUA -GUU-3’
3. 5’-UCU-GCA- AAU-UAA -GUU-3’

Using the peptide encoded by each of the above, compare the three peptides obtained. How are peptides 2 and 3 different from 1? What would be the consequence for the cell in each case?

a. There is a silent mutation in peptide 2 and peptide 3 has a stop codon due to mutation.
b. There is a silent mutation in peptide 3 and peptide 2 has a stop codon due to mutation.
c. There is a different amino acid in peptide 2 and peptide 3 has a stop codon due to mutation.
d. There isn’t a mutation in peptide 2 and peptide 3 has a stop codon due to mutation.

SCIENCE PRACTICE CHALLENGE QUESTIONS

52. Gamow (1954) proposed that the structure of DNA deduced by Watson and Crick (1953) could be interpreted as a way of forming roughly 20 "words" of the common amino acids from the four "letters" A, T, C, and G that represent DNA nucleotides.

Crick and coworkers (1961) used a method developed by Benzer to induce mutations in the DNA of a virus by the insertion of a single nucleotide. The mutant could not infect the bacterium *Escherichia coli* and neither could viruses with a second insertion of a second DNA nucleotide. However, a third nucleotide insertion restored the ability of the virus to infect the bacterium.

In 1961, Nirenberg and Matthaei conducted a series of experiments to better understand the flow of genetic information from gene to protein. They discovered that in solutions containing the contents of ruptured *E. coli* bacterial cells from which DNA had been removed, polymers containing only one repeating amino acid, phenylalanine, would be synthesized if synthetic mRNA composed of only the single nucleotide, uracil (U), was added to the solution in which phenylalanine was also present. In solutions containing mRNA with only adenine (A) or cytosine (C) and the amino acids lysine or proline, polymers containing only these amino acids would be synthesized. The researchers found that when ribosomes were removed by filtration, these polymers did not form. Nirenberg and Leder (1964) extended this work to include other nucleotides.

A. Summarize the conclusions regarding the encoding and decoding of heritable information supported by these studies. Explain how these studies provided evidence to support the Triplet Code.

Khorana (1960) developed a technique for synthesizing RNA composed of predictable distributions of repeated pairs or triplets of nucleotides. He found, for example, that RNA synthesized when A and U were present in relative concentrations of 4:1, respectively, will produce RNA sequences with these distributions determined by their relative probabilities: AAU:AAA, AUA:AAA, and UAA:AAA; $0.8^2 \times 0.2/0.8^3 = 1/4$ [calculated as follows: i) 4/5 of the bases are A, so the likelihood of selecting A is 0.8; ii) the selection is repeated to determine the second letter of the three-letter codon; iii) the likelihood of selecting a U is 1 in 5; iv) the probability of selecting the

set AUU is the product; v) similarly, the probability of AAA is $(4/5)^3$; and vi) the ratio of these probabilities is their relative likelihood]: AUU:AAA, UUA:AAA, and UAU:AAA; $0.8 \times 0.2^2/0.8^3 = 1/16$; and UUU:AAA; $0.2^3/0.8^3 = 1/64$.

B. Based on Khorana's findings, calculate the relative distributions of the following ratios of concentrations of RNA triplet sequences from mixtures in which the relative concentrations of guanine and cytosine, G:C, are 5:1.

Ratio	Relative Probabilities
GGC:GGG GCG:GGG CGG:GGG	
GCC:GGG CGC:GGG CCG:GGG	
CCC:GGG	

Table 15.2

C. Based on the work of Nirenberg, Matthaei, Leder, and Khorana, the following table was constructed (taken from Khorana's Nobel Prize address):

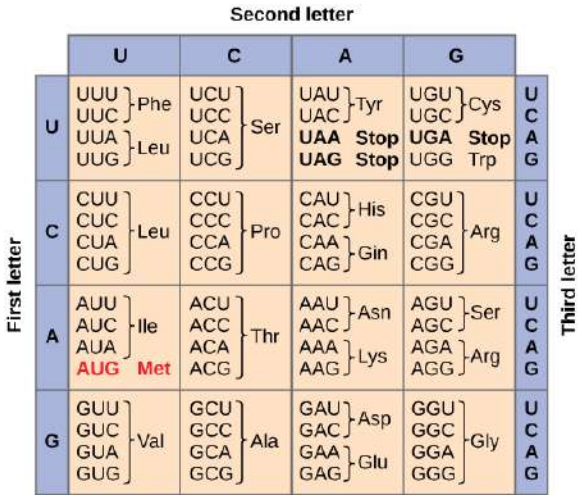

Figure 15.18

A solution containing the amino acids shown in the table above and equal concentrations of the two nucleotides C and G is prepared. Predict the proteins that can be synthesized from this mixture in terms of each possible codon and their relative concentrations in terms of their amino acid repeat sequences.

D. Describe the effects of the codons UAA, UAG, and UGA on protein synthesis.

53. The yeast life cycle is usually dominated by haploid cells, each with a single set of unpaired chromosomes. The cell propagates asexually, and the genetic material is replicated through mitosis. Cell division occurs every 2–4 hours, leading to 60–100 generations in a single day. Yeast also reproduce sexually, particularly under adverse environmental conditions. When two haploid cells—with DNA containing complementary mating-type alleles—conjugate, a diploid zygote results. The diploid zygote can then complete the sexual segment of the life cycle through meiosis. After meiosis, four haploid spores are produced, which can germinate.

Researchers can grow yeast easily on nutrient-containing plates. Because both asexual and sexual reproduction is rapid, yeast has become an important organism for the experimental investigation of mutagenesis and evolution among eukaryotes. Environmental factors, such as chemicals or radiation, induce mutations. High-energy UV-c radiation of less than 1 minute in duration will result in many mutated yeast cells. UV-c can be used to mutate a strain of yeast in which the synthesis of adenine is blocked. This mutation is observable because the *ade-2* mutant has a red color when cultured on nutrient-containing plates. Exposure to uv-c also can result in additional mutations. In particular, one mutant, *ade-7*, changes the color of the *ade-2* mutant to white.

A. You have a uv-c lamp, culture plates, and growth chambers at 23 °C and 37 °C. You also have available known haploid strains that are (*ade-2*,+,+), where + denotes the wild type. **Design** a plan to determine the rate of uv-c-induced mutations in nutrient-containing plates inoculated with yeast.

Earth's ozone layer removes high-energy ultraviolet radiation, uv-c, from the solar radiation received at the surface. Lower-energy ultraviolet radiation, uv-b, strikes Earth's surface. Damage to DNA induced by ultraviolet radiation occurs with the formation of bonds between an adjacent pair of pyrimidine nucleotides, thymine and cytosine, on the same strand of DNA. A repair enzyme, photolyase, which is activated by visible light, is present in plants and most animals, but not in humans. In characterizing the relationship between environmental mutagens and cell damage, a useful assumption is often made and referred to as the linear hypothesis. This assumption states that the extent of damage is proportional to the amount of radiation received.

Mutation rates for a strain (*preac*) that does not produce photolyase and a wild-type (+) strain were studied. Cultures of the two strains of yeast were diluted, and nutrient-containing plates were inoculated in triplicate at 23 °C. The plates were exposed to bright sunlight for varying time intervals. After exposure, the plates were incubated in the dark at 23 °C. After incubation between 1 and 8 hours, data shown in the table below were collected by counting the density of living cells relative to the control, and averaging these among replicates.

B. Using the data table below, **graph** the average survival fraction, relative to the wild-type control. Predict the number of mutations in a sample of 1,000 cells of the *preac* type that are exposed to bright sunlight for 15

seconds.

Incu-bation Time(hr)	10-S Exposure	20-S Exposure	30-S Exposure	40-S Exposure	50-S Exposure
1	0.83	0.58	0.33	0.17	0.08
2	1.00	0.43	0.17	0.09	0.04
3	0.92	0.38	0.12	0.03	0.01
4	0.75	0.35	0.08	0.01	0.00
5	0.99	0.49	0.11	0.01	0.00
6	0.81	0.42	0.12	0.01	0.00
7	0.80	0.32	0.09	0.01	0.00
8	1.05	0.59	0.11	0.01	0.00
Mean	0.89	0.45	0.14	0.04	0.02
Standard Deviation	0.11	0.10	0.08	0.06	0.03

Figure 15.19 This is a 5 column table, showing Incubation time, in hours in the left most column, ranging from 1 to 8. A 10 second exposure has the following values for an incubation time of 1 to 8: 0.83, 1.00, 0.92, 0.75, 0.99, 0.81, 0.80, 1.05 and 0.89 with a standard deviation of 0.11. A 20 second exposure has the following values for an incubation time of 1 to 8: 0.58, 0.43, 0.38, 0.35, 0.49, 0.42, 0.32, 0.59, 0.45, with a standard deviation of 0.10. A 30 second exposure has the following values for an incubation time of 1 to 8: 0.33, 0.17, 0.12, 0.08, 0.11, 0.12, 0.09, 0.11, 0.14, with a standard deviation of 0.08. A 40 second exposure has the following values for an incubation time of 1 to 8: 0.17, 0.09, 0.03, 0.01, 0.01, 0.01, 0.01, 0.01, 0.04, with a standard deviation of .0.06. A 50 second exposure has the following values for an incubation time of 1 to 8: 0.08, 0.04, 0.01, 0.00, 0.00, 0.00, 0.00, 0.00, 0.02 with a standard deviation of 0.03

Yeast can also be used to study sexual reproduction, a somewhat puzzling phenomenon. Cloning of cells through mitosis is molecularly much less complex than meiosis, consumes less energy, and is less risky. Two alternative explanations for the evolution of sexual reproduction are popular. In one model, through assortment of genes, meiosis leads to an increase in the frequency of beneficial mutations. In the second model, detrimental mutations are purged from a population through sex. Studies using yeast (Gray and Goddard, *Evol. Biol.*, 2012 and McDonald et al., *Nature* 2012) have provided a mechanism to study these models. As shown below, the fitness (defined as the log of the ratio of the number of cells in successive generations) of yeast is graphed as a function of number of mitotic reproductions in yeast grown in low-stress and high-stress environments, and with and without alternating induction of sexual reproduction.

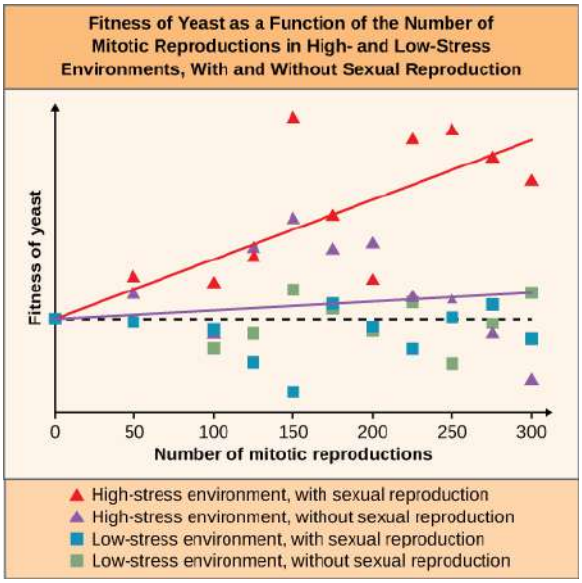

Figure 15.20

C. Based on these data, **evaluate** the merits of the alternative theories of the adaptive advantage provided by sexual reproduction.

54. A. **Describe** the storage and retrieval of genetic information with the following model. Use the list to fill in the blanks with the letter corresponding to the correct term.

a. amino acid
b. tRNA
c. DNA
d. transcription
e. mRNA
f. translation
g. protein
h. RNA polymerase
i. rRNA

Within the cytoplasm, __ is synthesized from __ bound to __ in a sequence that corresponds to information provided by __. This process is called __.

Within the nucleus, information originating in __ is encoded as a sequence of bases in __, which is synthesized by the enzyme __ that is embedded in the __. This process is called __.

B. During development, cell differentiation occurs, and the expression of genes is permanently switched off. Using the model summarized above, **explain** where information flow is most effectively blocked.

C. A chemical message is received by the cell regulating the timing of events controlled by gene expression. Using the model summarized above, **explain** where information flow is most effectively managed.

55. Structure and function in biology result from both the presence of genetic information and the expression of that information. Some genes are continually expressed, whereas the expression of most genes is regulated, commonly at the level of transcription. At the initiation of transcription, the TATA-binding protein (TBP) provides access to the DNA strand to be transcribed. The 5'TATAAA3' sequence called the TATA box is found in prokaryotes, archaebacteria, and eukaryotes. Even among eukarya, when the TATA box is not present among eukaryotes, the initiation of transcription involves TBP. Scientists attribute this common characteristic to the relative thermostability of the A-T interaction. Hydrogen bonds hold the two strands of the DNA double helix together. This type of bond has the smallest interaction energy of all intermolecular forces; as temperature increases, these bonds are broken.

A. **Explain** the advantage, in terms of the energy required, which is provided by an AT-rich region in the sequence where transcription is initiated.

B. The fact that the TATA box or the associated TBP are common to all domains provides evidence of common ancestry among all life. **Pose a scientific question** that would need to be addressed by a valid alternative explanation of this fact.

C. A whole-genome survey of prokaryotes (Zheng and Wu, *BMC Bioinformatics*, 2010) showed that the relative amounts of guanine and cytosine in DNA poorly predicted the temperature range conditions that are suitable for an organism. **Refine the question** posed in part B, taking this result into account.

56. Only a fraction of DNA encodes proteins. The noncoding portion of a gene is referred to as the intron. The intron fraction depends upon the gene. Introns are rare in prokaryotic and mitochondrial DNA; in human nuclear DNA, this fraction is about 95%. The intron is transcribed into mRNA, but this noncoding mRNA is edited out before translation of the coding portion, or exon, of a gene. The edited exon segments are then spliced together by a spliceosome, a very large and complex collection of RNAs and proteins.

Although introns do not encode proteins, they have functions. In particular, they amplify expression of the exon, although the mechanism is unknown. When introns are very long, which is common among mammalian genes with roles in development, they can significantly extend the time required to complete transcription. Analysis of genes common to different plant and animal species shows many shared intronic positions and base sequences, although in some organisms, such as yeast, many introns have been deleted. Because introns do not encode proteins, mutations can remain silent and accumulate.

A. As described above, introns are ancestral remnants that are replicated because they do not disadvantage the organism. Consider the claim that introns are "junk DNA." **Evaluate the claim** with supporting evidence.

B. Introns may be retained during transcription. **Explain** how the retention of a transcribed intron between two transcribed exons within a gene could do the following:

- block expression of one polypeptide sequence
- increase expression of a polypeptide
- alter the polypeptide expressed