

13 Transcription



The death cap mushroom, *Amanita phalloides*, causes death by inhibiting the process of transcription. [Ron Wolf/Tom Stack & Associates.]

DEATH CAP POISONING

On November 8, 2009, 31-year-old Tomasa was hiking the Lodi Lake nature trail east of San Francisco with her husband and cousin when they came across some large white mushrooms that looked very much like the edible mushrooms that they enjoyed in their native Mexico. They picked the mushrooms and took them home, cooking and consuming them for dinner. Within hours, Tomasa and her family were sick and went to the hospital. They were later transferred to the critical care unit at California Pacific Medical Center in San Francisco, where Tomasa died of liver failure 3 weeks later. Her husband eventually recovered after a lengthy hospitalization; her cousin required a liver transplant to survive.

The mushrooms consumed by Tomasa and her family were *Amanita phalloides*, commonly known as the death cap. A single death cap mushroom contains enough toxin to kill an adult human. The death rate among those who consume death caps is 22%; among children under the age of 10, it's more than 50%. Death cap mushrooms appear to be spreading in California, leading to a recent surge in the number of mushroom poisonings.

Death cap poisoning is insidious. Gastrointestinal symptoms—abdominal pain, cramping, vomiting, diarrhea—begin within 6 to 12 hours of consuming the mushrooms, but these symptoms usually subside within a few hours and the patient seems to recover. Because of this initial remission, the poisoning is often not taken seriously until it's too late to pump the stomach and remove the toxin from the body. After a day or two, serious symptoms begin. Cells in the liver die, often causing permanent liver damage and death within a few days. There is no effective treatment, other than a liver transplant to replace the damaged organ.

How do death caps kill? Their deadly toxin, contained within the fruiting bodies that produce reproductive spores, is the protein α -amanitin, which consists of a short peptide of eight amino acids that forms a circular loop. α -Amanitin is a potent inhibitor of RNA polymerase II, the enzyme that transcribes protein-encoding genes in eukaryotes. RNA polymerase II binds to genes and synthesizes RNA molecules that are complementary to the DNA template. In the process of transcription, the RNA polymerase moves down the DNA template, adding one nucleotide at a time to the growing RNA chain. α -Amanitin binds to RNA polymerase and jams the moving parts of the enzyme, interfering with its ability to move along the DNA template. In the presence of α -amanitin, RNA synthesis

slows from its normal rate of several thousand nucleotides per minute to just a few nucleotides per minute. The results are catastrophic. Without transcription, protein synthesis—required for cellular function—ceases and cells die. The liver, where the toxin accumulates, is irreparably damaged and stops functioning. In severe cases, the patient dies.

Death cap poisoning illustrates the extreme importance of transcription and the central role that RNA polymerase plays in the process. This chapter is about the process of transcription, the first step in the central dogma, the pathway of information transfer from DNA (genotype) to protein (phenotype). Transcription is a complex process that requires precursors to RNA nucleotides, a DNA template, and a number of protein components. As we examine the stages of transcription, try to keep all the details in perspective; focus on understanding how the details relate to the overall purpose of transcription—the selective synthesis of an RNA molecule.

This chapter begins with a brief review of RNA structure and a discussion of the different classes of RNA. We then consider the major components required for transcription. Finally, we explore the process of transcription in eubacteria, eukaryotic cells, and archaea. At several points in the text, we'll pause to absorb some general principles that emerge.

13.1 RNA, Consisting of a Single Strand of Ribonucleotides, Participates in a Variety of Cellular Functions

Before we begin our study of transcription, we will consider the past and present importance of RNA, review the structure of RNA, and examine some of the different types of RNA molecules.

An Early RNA World

Life requires two basic functions. First, living organisms must be able to store and faithfully transmit genetic information during reproduction. Second, they must have the ability to catalyze the chemical transformations that drive life processes. A long-held belief was that the functions of information storage and chemical transformation are handled by two entirely different types of molecules: genetic information is stored in nucleic acids, whereas chemical transformations are catalyzed by protein enzymes. This biochemical dichotomy—nucleic acid for information, proteins for catalysts—created a dilemma. Which came first: proteins or nucleic acids? If nucleic acids carry the coding instructions for proteins, how could proteins be generated without them? Because nucleic acids are unable to copy themselves, how could they be generated without proteins? If DNA and proteins each require the other, how could life begin?

This apparent paradox was given a possible answer in 1981 when Thomas Cech and his colleagues discovered that

RNA can serve as a biological catalyst. They found that RNA from the protozoan *Tetrahymena thermophila* can excise 400 nucleotides from its RNA in the absence of any protein. Other examples of catalytic RNAs have now been discovered in different types of cells. Called **ribozymes**, these catalytic RNA molecules can cut out parts of their own sequences, connect some RNA molecules together, replicate others, and even catalyze the formation of peptide bonds between amino acids. The discovery of ribozymes complements other evidence suggesting that the original genetic material was RNA.

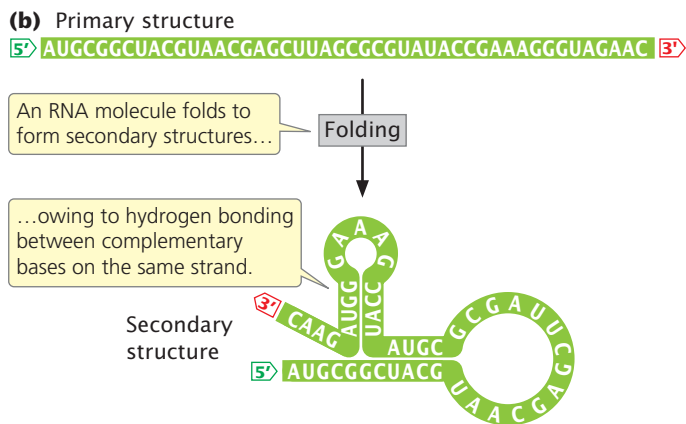
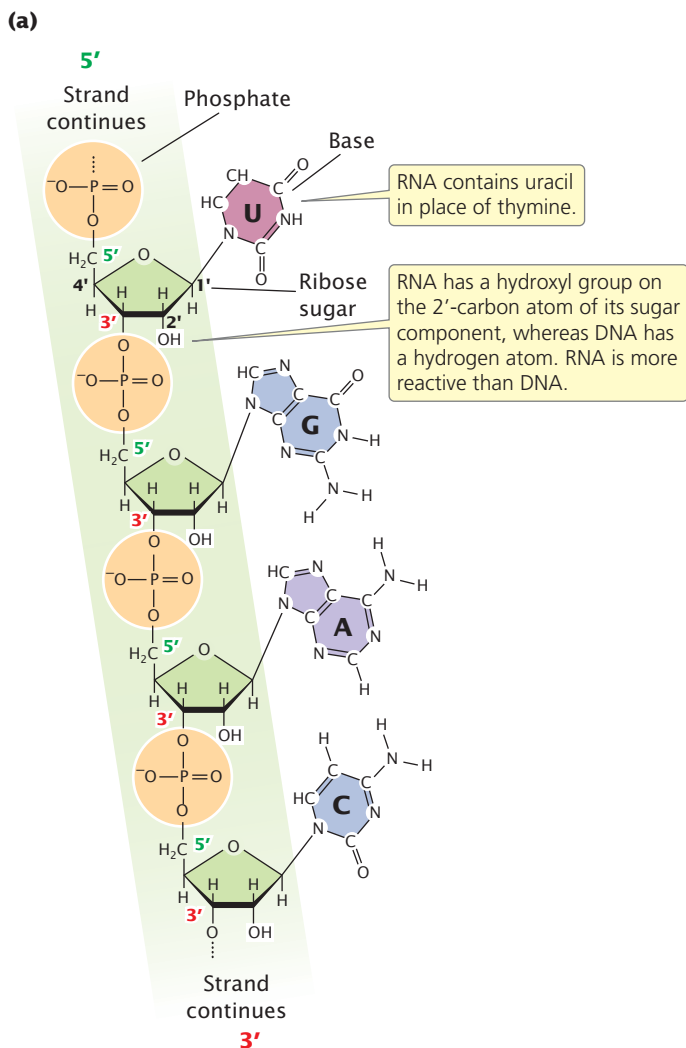
Self-replicating ribozymes probably first arose between 3.5 billion and 4 billion years ago and may have begun the evolution of life on Earth. Early life was probably an RNA world, with RNA molecules serving both as carriers of genetic information and as catalysts that drove the chemical reactions needed to sustain and perpetuate life. These catalytic RNAs may have acquired the ability to synthesize protein-based enzymes, which are more efficient catalysts. With enzymes taking over more and more of the catalytic functions, RNA probably became relegated to the role of information storage and transfer. DNA, with its chemical stability and faithful replication, eventually replaced RNA as the primary carrier of genetic information. Nevertheless, RNA still plays a vital role in many biological processes, including replication, RNA processing, and translation. And research in the past 10 years has determined that newly discovered small RNA molecules play a fundamental role in many basic biological processes, demonstrating that life today is still very much an RNA world. These small RNA molecules will be discussed in more detail in Chapter 14.

CONCEPTS

Early life was probably centered on RNA, which served as the original genetic material and as biological catalysts.

The Structure of RNA

RNA, like DNA, is a polymer consisting of nucleotides joined together by phosphodiester bonds (see Chapter 10 for a discussion of RNA structure). However, there are several important differences in the structures of DNA and RNA. Whereas DNA nucleotides contain deoxyribose sugars, RNA nucleotides have ribose sugars (**Figure 13.1a**). With a free hydroxyl group on the 2'-carbon atom of the ribose sugar, RNA is degraded rapidly under alkaline conditions. The deoxyribose sugar of DNA lacks this free hydroxyl group; so



13.1 RNA has a primary and a secondary structure.

DNA is a more stable molecule. Another important difference is that thymine, one of the two pyrimidines found in DNA, is replaced by uracil in RNA.

A final difference in the structures of DNA and RNA is that RNA is usually single stranded, consisting of a single polynucleotide strand (**Figure 13.1b**), whereas DNA

normally consists of two polynucleotide strands joined by hydrogen bonding between complementary bases. Some viruses contain double-stranded RNA genomes, as discussed in Chapter 8. Although RNA is usually single stranded, short complementary regions within a nucleotide strand can pair and form secondary structures (see **Figure 13.1b**). These RNA secondary structures are often called hairpin-loop or stem-loop structures. When two regions within a single RNA molecule pair up, the strands in those regions must be antiparallel, with pairing between cytosine and guanine and between adenine and uracil (although, occasionally, guanine pairs with uracil).

The formation of secondary structures plays an important role in RNA function. Secondary structure is determined by the base sequence of the nucleotide strand; so different RNA molecules can assume different structures. Because their structure determines their function, RNA molecules have the potential for tremendous variation in function. With its two complementary strands forming a helix, DNA is much more restricted in the range of secondary structures that it can assume and so has fewer functional roles in the cell. Similarities and differences in DNA and RNA structures are summarized in **Table 13.1**. **TRY PROBLEM 15** →

Classes of RNA

RNA molecules perform a variety of functions in the cell. **Ribosomal RNA (rRNA)** and ribosomal protein subunits make up the ribosome, the site of protein assembly. We'll take a more detailed look at the ribosome in Chapter 14. **Messenger RNA (mRNA)** carries the coding instructions for polypeptide chains from DNA to a ribosome. After attaching to the ribosome, an mRNA molecule specifies the sequence of the amino acids in a polypeptide chain and provides a template for joining amino acids. Large precursor molecules, which are termed **pre-messenger RNAs (pre-mRNAs)**, are the immediate products of transcription

Table 13.1 The structures of DNA and RNA compared

Characteristic	DNA	RNA
Composed of nucleotides	Yes	Yes
Type of sugar	Deoxyribose	Ribose
Presence of 2'-OH group	No	Yes
Bases	A, G, C, T	A, G, C, U
Nucleotides joined by phosphodiester bonds	Yes	Yes
Double or single stranded	Usually double	Usually single
Secondary structure	Double helix	Many types
Stability	Stable	Easily degraded

Table 13.2 Location and functions of different classes of RNA molecules

Class of RNA	Cell Type	Location of Function in Eukaryotic Cells*	Function
Ribosomal RNA (rRNA)	Bacterial and eukaryotic	Cytoplasm	Structural and functional components of the ribosome
Messenger RNA (mRNA)	Bacterial and eukaryotic	Nucleus and cytoplasm	Carries genetic code for proteins
Transfer RNA (tRNA)	Bacterial and eukaryotic	Cytoplasm	Helps incorporate amino acids into polypeptide chain
Small nuclear RNA (snRNA)	Eukaryotic	Nucleus	Processing of pre-mRNA
Small nucleolar RNA (snoRNA)	Eukaryotic	Nucleus	Processing and assembly of rRNA
Small cytoplasmic RNA (scRNA)	Eukaryotic	Cytoplasm	Variable
MicroRNA (miRNA)	Eukaryotic	Cytoplasm	Inhibits translation of mRNA
Small interfering RNA (siRNA)	Eukaryotic	Cytoplasm	Triggers degradation of other RNA molecules
Piwi-interacting RNA (piRNA)	Eukaryotic	Cytoplasm	Thought to regulate gametogenesis, but function poorly defined

*All eukaryotic RNAs are transcribed in the nucleus.

in eukaryotic cells. Pre-mRNAs are modified extensively before becoming mRNA and exiting the nucleus for translation into protein. Bacterial cells do not possess pre-mRNA; in these cells, transcription takes place concurrently with translation.

Transfer RNA (tRNA) serves as the link between the coding sequence of nucleotides in the mRNA and the amino acid sequence of a polypeptide chain. Each tRNA attaches to one particular type of amino acid and helps to incorporate that amino acid into a polypeptide chain (discussed in Chapter 15).

Additional classes of RNA molecules are found in the nuclei of eukaryotic cells. **Small nuclear RNAs** (snRNAs) combine with small protein subunits to form **small nuclear ribonucleoproteins** (snRNPs, affectionately known as “snurps”). Some snRNAs participate in the processing of RNA, converting pre-mRNA into mRNA. **Small nucleolar RNAs** (snoRNAs) take part in the processing of rRNA. Small RNA molecules are also found in the cytoplasm of eukaryotic cells; these molecules, called **small cytoplasmic RNAs** (scRNAs), have varied and often unknown function.

A class of very small and abundant RNA molecules, termed **microRNAs** (miRNAs) and **small interfering RNAs** (siRNAs), are found in bacteria and eukaryotic cells and carry out RNA interference (RNAi), a process in which these small RNA molecules help trigger the degradation of mRNA or inhibit its translation into protein. More will be said about RNA interference in Chapter 14. Recent research has uncovered another class of small RNA molecules called **Piwi-interacting RNAs** (piRNAs; named after Piwi proteins,

with which they interact). Found in mammalian testes, these RNA molecules are similar to miRNAs and siRNAs; they are thought to have a role in the regulation of sperm development. The different classes of RNA molecules are summarized in **Table 13.2**.

CONCEPTS

RNA differs from DNA in that RNA possesses a hydroxyl group on the 2'-carbon atom of its sugar, contains uracil instead of thymine, and is normally single stranded. Several classes of RNA exist within bacterial and eukaryotic cells.

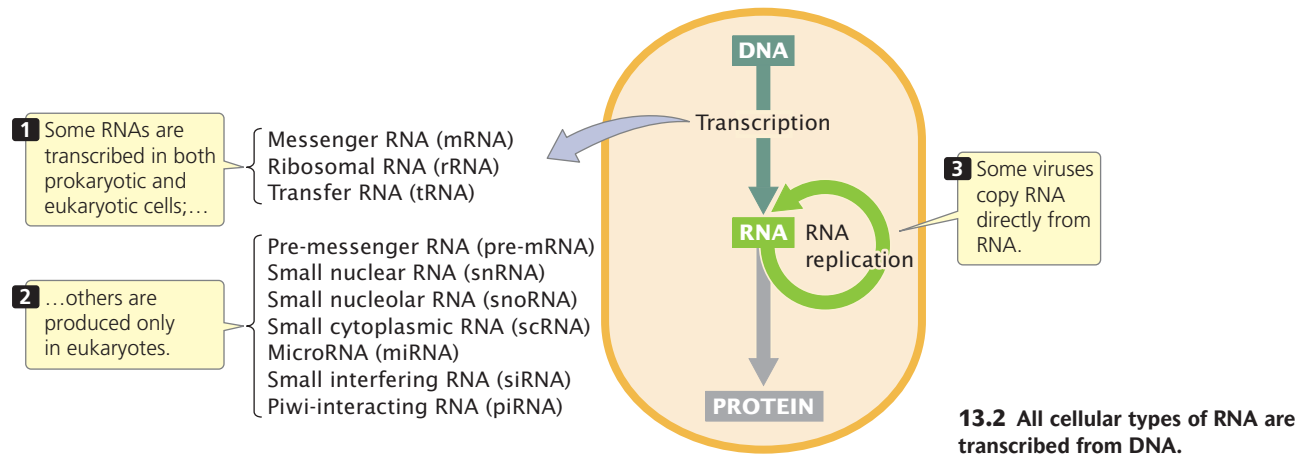
✓ CONCEPT CHECK 1

Which class of RNA is correctly paired with its function?

- Small nuclear RNA (snRNA): processes rRNA
- Transfer RNA (tRNA): attaches to an amino acid
- MicroRNA (miRNA): carries information for the amino acid sequence of a protein
- Ribosomal RNA (rRNA): carries out RNA interference

13.2 Transcription Is the Synthesis of an RNA Molecule from a DNA Template

All cellular RNAs are synthesized from DNA templates through the process of transcription (**Figure 13.2**). Transcription is in many ways similar to the process of replica-



tion, but a fundamental difference relates to the length of the template used. In replication, all the nucleotides in the DNA template are copied, but, in transcription, only small parts of the DNA molecule—usually a single gene or, at most, a few genes—are transcribed into RNA. Because not all gene products are needed at the same time or in the same cell, the constant transcription of all of a cell's genes would be highly inefficient. Furthermore, much of the DNA does not encode a functional product, and transcription of such sequences would be pointless. Transcription is, in fact, a highly selective process: individual genes are transcribed only as their products are needed. But this selectivity imposes a fundamental problem on the cell—the problem of how to recognize individual genes and transcribe them at the proper time and place.

Like replication, transcription requires three major components:

1. a DNA template;
2. the raw materials (substrates) needed to build a new RNA molecule; and
3. the transcription apparatus, consisting of the proteins necessary to catalyze the synthesis of RNA.

The Template

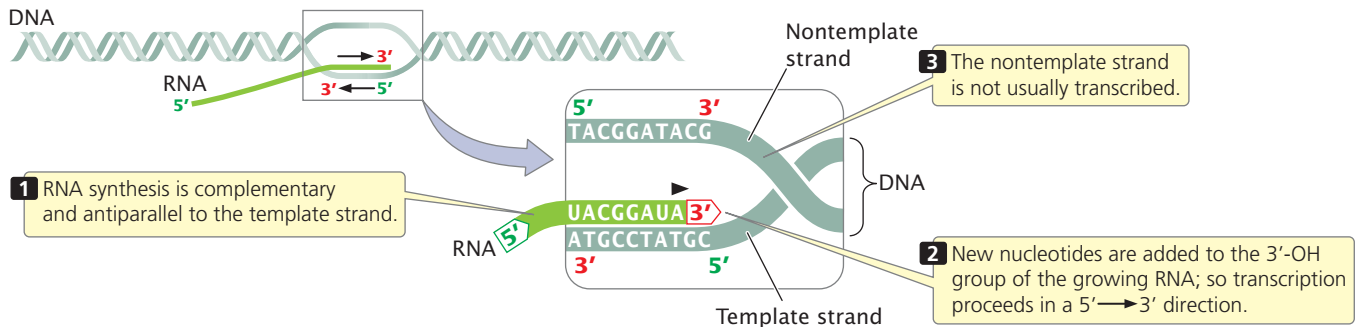
In 1970, Oscar Miller, Jr., Barbara Hamkalo, and Charles Thomas used electron microscopy to examine cellular contents and demonstrate that RNA is transcribed from a DNA template. They saw within the cell Christmas-tree-like structures: thin central fibers (the trunk of the tree), to which were attached strings (the branches) with granules (**Figure 13.3**). The addition of deoxyribonuclease (an enzyme that degrades DNA) caused the central fibers to disappear, indicating that the “tree trunks” were DNA molecules. Ribonuclease (an enzyme that degrades RNA) removed the granular strings, indicating that the branches were RNA. Their conclusion was that each “Christmas tree” represented a gene undergoing transcription. The transcription of each

gene begins at the top of the tree; there, little of the DNA has been transcribed and the RNA branches are short. As the transcription apparatus moves down the tree, transcribing more of the template, the RNA molecules lengthen, producing the long branches at the bottom.

The transcribed strand The template for RNA synthesis, as for DNA synthesis, is a single strand of the DNA double helix. Unlike replication, however, the transcription of a gene takes place on only one of the two nucleotide strands of DNA (**Figure 13.4**). The nucleotide strand used for transcription is termed the **template strand**. The other strand, called the **nontemplate strand**, is not ordinarily transcribed. Thus, within a gene, only one of the nucleotide strands is normally transcribed into RNA (there are some exceptions to this rule).



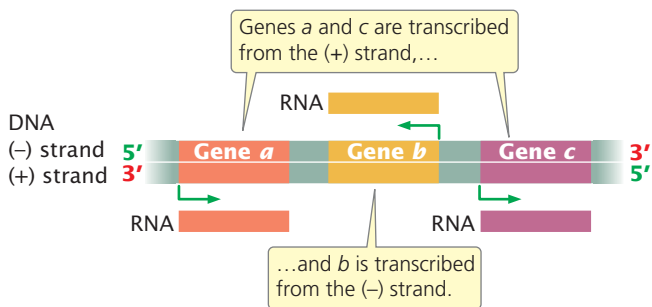
13.3 Under the electron microscope, DNA molecules undergoing transcription exhibit Christmas-tree-like structures. The trunk of each “Christmas tree” (a transcription unit) represents a DNA molecule; the tree branches (granular strings attached to the DNA) are RNA molecules that have been transcribed from the DNA. As the transcription apparatus moves down the DNA, transcribing more of the template, the RNA molecules become longer and longer. [Dr. Thomas Broker/Phototake.]



13.4 RNA molecules are synthesized that are complementary and antiparallel to one of the two nucleotide strands of DNA, the template strand.

During transcription, an RNA molecule that is complementary and antiparallel to the DNA template strand is synthesized (see Figure 13.4). The RNA transcript has the same polarity and base sequence as that of the nontemplate strand, with the exception that RNA contains U rather than T. In most organisms, each gene is transcribed from a single strand, but different genes may be transcribed from different strands, as shown in **Figure 13.5**. **TRY PROBLEM 16** →

fig 13.5



13.5 RNA is transcribed from one DNA strand. In most organisms, each gene is transcribed from a single DNA strand, but different genes may be transcribed from one or the other of the two DNA strands.

CONCEPTS

Within a single gene, only one of the two DNA strands, the template strand, is usually transcribed into RNA.

✓ CONCEPT CHECK 2

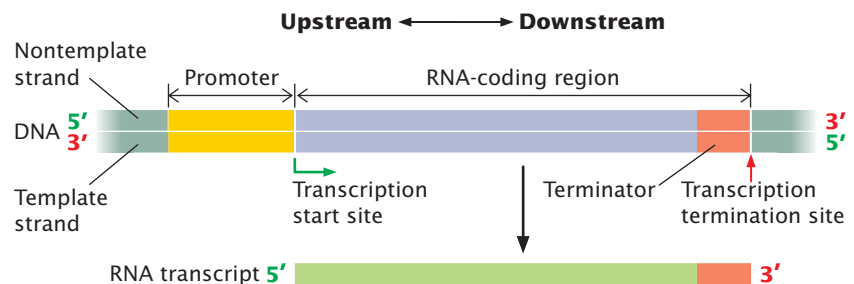
What is the difference between the template strand and the nontemplate strand?

The transcription unit A **transcription unit** is a stretch of DNA that encodes an RNA molecule and the sequences necessary for its transcription. How does the complex of enzymes and proteins that performs transcription—the transcription apparatus—recognize a transcription unit? How does it know which DNA strand to read and where to start and stop? This information is encoded by the DNA sequence.

Included within a transcription unit are three critical regions: a promoter, an RNA-coding sequence, and a terminator (**Figure 13.6**). The **promoter** is a DNA sequence that the transcription apparatus recognizes and binds. It indicates which of the two DNA strands is to be read as the template and the direction of transcription. The promoter also determines the transcription start site, the first nucleotide that will be transcribed into RNA. In most transcription units, the promoter is located next to the transcription start site but is not, itself, transcribed.

The second critical region of the transcription unit is the **RNA-coding region**, a sequence of DNA nucleotides that is copied into an RNA molecule. The third component of the transcription unit is the **terminator**, a sequence of nucleotides that signals where transcription is to end. Terminators are usually part of the RNA-coding sequence; that is, transcription stops only after the terminator has been copied into RNA.

Molecular biologists often use the terms *upstream* and *downstream* to refer to the direction of transcription and the location of nucleotide sequences surrounding the RNA-coding sequence. The transcription apparatus is said to move downstream during transcription: it binds to the



13.6 A transcription unit includes a promoter, an RNA-coding region, and a terminator.

promoter (which is usually upstream of the start site) and moves toward the terminator (which is downstream of the start site).

When DNA sequences are written out, often the sequence of only one of the two strands is listed. Molecular biologists typically write the sequence of the nontemplate strand, because it will be the same as the sequence of the RNA transcribed from the template (with the exception that U in RNA replaces T in DNA). By convention, the sequence on the nontemplate strand is written with the 5' end on the left and the 3' end on the right. The first nucleotide transcribed (the transcription start site) is numbered +1; nucleotides downstream of the start site are assigned positive numbers, and nucleotides upstream of the start site are assigned negative numbers. So, nucleotide +34 would be 34 nucleotides downstream of the start site, whereas nucleotide -75 would be 75 nucleotides upstream of the start site. There is no nucleotide numbered 0.

CONCEPTS

A transcription unit is a piece of DNA that encodes an RNA molecule and the sequences necessary for its proper transcription. Each transcription unit includes a promoter, an RNA-coding region, and a terminator.

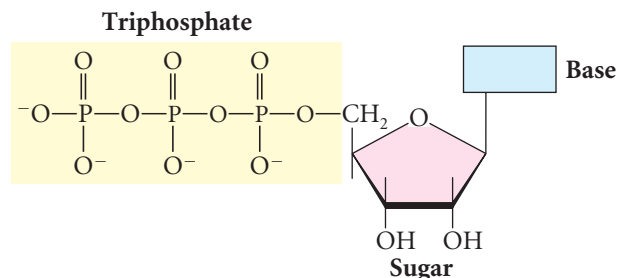
✓ CONCEPT CHECK 3

Which of the following phrases does *not* describe a function of the promoter?

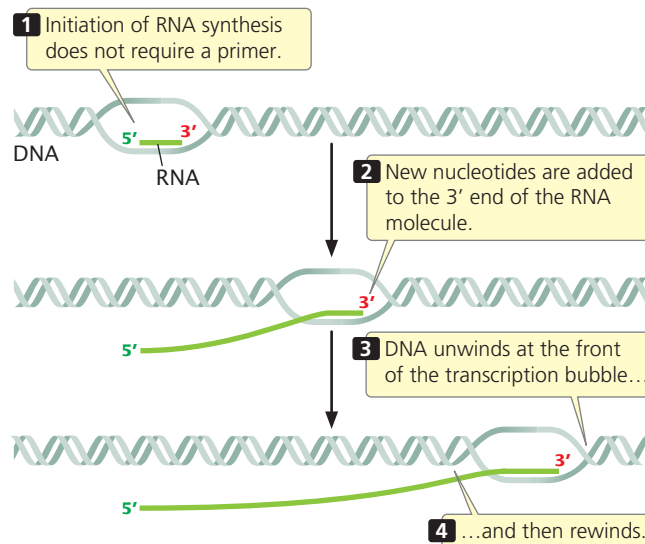
- Serves as sequence to which transcription apparatus binds
- Determines the first nucleotide that is transcribed into RNA
- Determines which DNA strand is template
- Signals where transcription ends

The Substrate for Transcription

RNA is synthesized from **ribonucleoside triphosphates** (rNTPs; **Figure 13.7**). In synthesis, nucleotides are added one at a time to the 3'-OH group of the growing RNA molecule. Two phosphate groups are cleaved from the incoming ribonucleoside triphosphate; the remaining phosphate group participates in a phosphodiester bond that connects

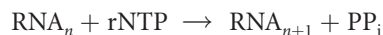


13.7 Ribonucleoside triphosphates are substrates used in RNA synthesis.



13.8 In transcription, nucleotides are always added to the 3' end of the RNA molecule.

the nucleotide to the growing RNA molecule. The overall chemical reaction for the addition of each nucleotide is:



where PP_i represents pyrophosphate. Nucleotides are always added to the 3' end of the RNA molecule, and the direction of transcription is therefore 5'→3' (**Figure 13.8**), the same as the direction of DNA synthesis during replication. The synthesis of RNA is complementary and antiparallel to one of the DNA strands (the template strand). Unlike DNA synthesis, RNA synthesis does not require a primer.

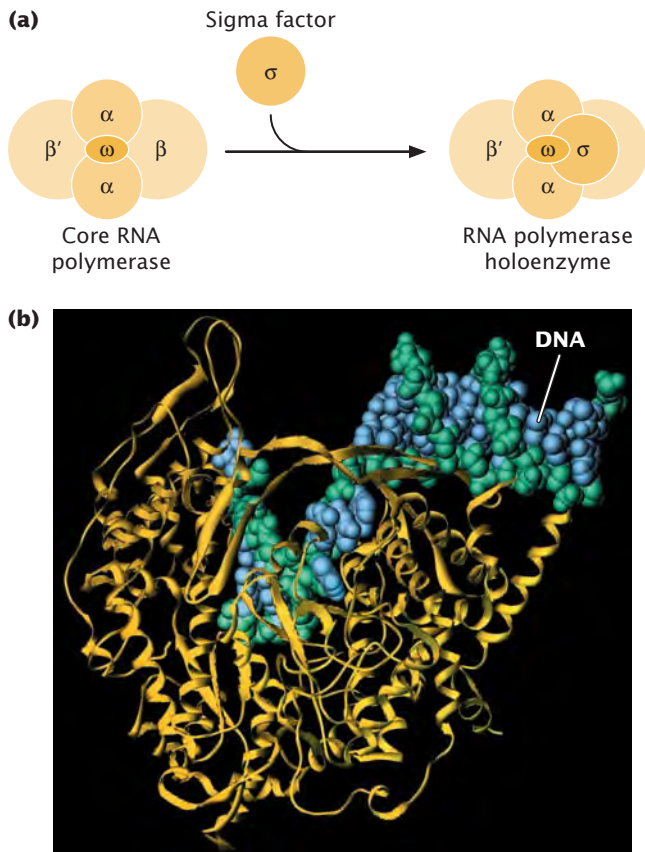
CONCEPTS

RNA is synthesized from ribonucleoside triphosphates. Transcription is 5'→3': each new nucleotide is joined to the 3'-OH group of the last nucleotide added to the growing RNA molecule.

The Transcription Apparatus

Recall that DNA replication requires a number of different enzymes and proteins. Although transcription might initially appear to be quite different because a single enzyme—**RNA polymerase**—carries out all the required steps of transcription, on closer inspection, the processes are actually similar. The action of RNA polymerase is enhanced by a number of accessory proteins that join and leave the polymerase at different stages of the process. Each accessory protein is responsible for providing or regulating a special function. Thus, transcription, like replication, requires an array of proteins.

Bacterial RNA polymerase Bacterial cells typically possess only one type of RNA polymerase, which catalyzes the synthesis of all classes of bacterial RNA: mRNA, tRNA, and rRNA.



13.9 In bacterial RNA polymerase, the core enzyme consists of five subunits: two copies of alpha (α), a single copy of beta (β), a single copy of beta prime (β'), and a single copy of omega (ω). The core enzyme catalyzes the elongation of the RNA molecule by the addition of RNA nucleotides. (a) The sigma factor (σ) joins the core to form the holoenzyme, which is capable of binding to a promoter and initiating transcription. (b) The molecular model shows RNA polymerase (in yellow) binding DNA.

Bacterial RNA polymerase is a large, multimeric enzyme (meaning that it consists of several polypeptide chains).

At the heart of most bacterial RNA polymerases are five subunits (individual polypeptide chains) that make up the **core enzyme**: two copies of a subunit called alpha (α) and single copies of subunits beta (β), beta prime (β'), and omega (ω) (Figure 13.9). The ω subunit is not essential for transcription, but it helps stabilize the enzyme. The core enzyme catalyzes the elongation of the RNA molecule by the addition of RNA nucleotides. Other functional subunits join and leave the core enzyme at particular stages of the transcription process. The **sigma** (σ) **factor** controls the binding of RNA polymerase to the promoter. Without sigma, RNA polymerase will initiate transcription at a random point along the DNA. After sigma has associated with the core enzyme (forming a **holoenzyme**), RNA polymerase binds stably only to the promoter region and initiates transcription at the proper start site. Sigma is required only for promoter binding and initiation; when

a few RNA nucleotides have been joined together, sigma usually detaches from the core enzyme. Many bacteria have multiple types of sigma factors; each type of sigma initiates the binding of RNA polymerase to a particular set of promoters.

Rifamycins are a group of antibiotics that kill bacterial cells by inhibiting RNA polymerase. These antibiotics are widely used to treat tuberculosis, a disease that kills almost 2 million people worldwide each year. The structures of bacterial and eukaryotic RNA polymerases are sufficiently different that rifamycins inhibit bacterial RNA polymerases without interfering with eukaryotic RNA polymerases. Recent research has demonstrated that several rifamycins inhibit RNA polymerase by binding to the part of the RNA polymerase that clamps on to DNA and jamming it, thus preventing the RNA polymerase from interacting with the promoter on the DNA.

Eukaryotic RNA polymerases Most eukaryotic cells possess three distinct types of RNA polymerase, each of which is responsible for transcribing a different class of RNA: **RNA polymerase I** transcribes rRNA; **RNA polymerase II** transcribes pre-mRNAs, snoRNAs, some miRNAs, and some snRNAs; and **RNA polymerase III** transcribes other small RNA molecules—specifically tRNAs, small rRNA, some miRNAs, and some snRNAs (Table 13.3). RNA polymerases I, II, and III are found in all eukaryotes. Two additional RNA polymerases, **RNA polymerase IV** and **RNA polymerase V**, have been found in plants. RNA polymerases IV and V transcribe RNAs that play a role in DNA methylation and chromatin structure.

All eukaryotic polymerases are large, multimeric enzymes, typically consisting of more than a dozen subunits. Some sub-

Table 13.3 Eukaryotic RNA polymerases

Type	Present in	Transcribes
RNA polymerase I	All eukaryotes	Large rRNAs
RNA polymerase II	All eukaryotes	Pre-mRNA, some snRNAs, snoRNAs, some miRNAs
RNA polymerase III	All eukaryotes	tRNAs, small rRNAs, some snRNAs, some miRNAs
RNA polymerase IV	Plants	Some siRNAs
RNA polymerase V	Plants	RNA molecules taking part in heterochromatin formation

units are common to all RNA polymerases, whereas others are limited to one of the polymerases. As in bacterial cells, a number of accessory proteins bind to the core enzyme and affect its function.

CONCEPTS

Bacterial cells possess a single type of RNA polymerase, consisting of a core enzyme and other subunits that participate in various stages of transcription. Eukaryotic cells possess several distinct types of RNA polymerase that transcribe different kinds of RNA molecules.

✓ CONCEPT CHECK 4

What is the function of the sigma factor?

13.3 The Process of Bacterial Transcription Consists of Initiation, Elongation, and Termination

Now that we've considered some of the major components of transcription, we're ready to take a detailed look at the process. Transcription can be conveniently divided into three stages:

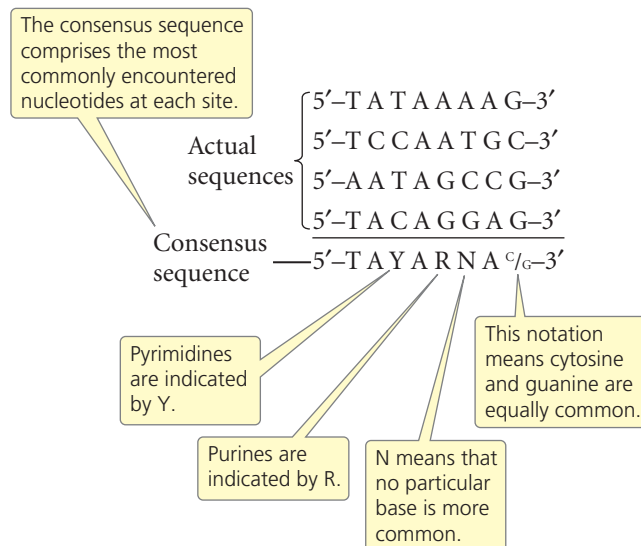
1. initiation, in which the transcription apparatus assembles on the promoter and begins the synthesis of RNA;
2. elongation, in which DNA is threaded through RNA polymerase, the polymerase unwinding the DNA and adding new nucleotides, one at a time, to the 3' end of the growing RNA strand; and
3. termination, the recognition of the end of the transcription unit and the separation of the RNA molecule from the DNA template.

We will first examine each of these steps in bacterial cells, where the process is best understood; then we will consider eukaryotic and archaeal transcription.

Initiation

Initiation comprises all the steps necessary to begin RNA synthesis, including (1) promoter recognition, (2) formation of the transcription bubble, (3) creation of the first bonds between rNTPs, and (4) escape of the transcription apparatus from the promoter.

Transcription initiation requires that the transcription apparatus recognize and bind to the promoter. At this step, the selectivity of transcription is enforced; the binding of RNA polymerase to the promoter determines which parts of the DNA template are to be transcribed and how often. Different genes are transcribed with different frequencies, and promoter binding is primarily responsible for determining the frequency of transcription for a particular gene.



13.10 A consensus sequence consists of the most commonly encountered bases at each position in a group of related sequences.

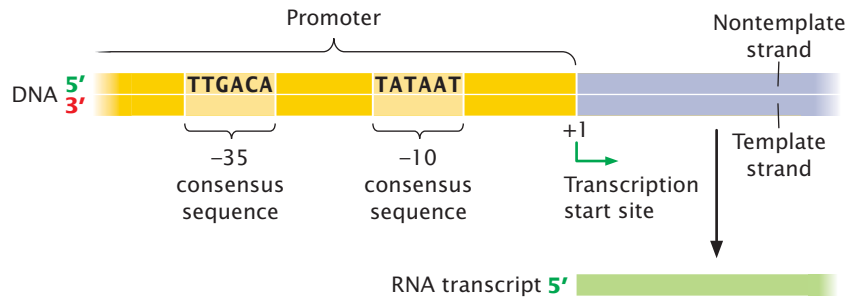
Promoters also have different affinities for RNA polymerase. Even within a single promoter, the affinity can vary with the passage of time, depending on the promoter's interaction with RNA polymerase and a number of other factors.

Bacterial promoters Essential information for the transcription unit—where it will start transcribing, which strand is to be read, and in what direction the RNA polymerase will move—is imbedded in the nucleotide sequence of the promoter. Promoters are DNA sequences that are recognized by the transcription apparatus and are required for transcription to take place. In bacterial cells, promoters are usually adjacent to an RNA-coding sequence.

An examination of many promoters in *E. coli* and other bacteria reveals a general feature: although most of the nucleotides within the promoters vary in sequence, short stretches of nucleotides are common to many. Furthermore, the spacing and location of these nucleotides relative to the transcription start site are similar in most promoters. These short stretches of common nucleotides are called **consensus sequences**; “consensus sequence” refers to sequences that possess considerable similarity, or consensus (**Figure 13.10**). The presence of consensus in a set of nucleotides usually implies that the sequence is associated with an important function. **TRY PROBLEM 19** →

The most commonly encountered consensus sequence, found in almost all bacterial promoters, is centered about 10 bp upstream of the start site. Called the **−10 consensus sequence** or, sometimes, the Pribnow box, its consensus sequence is





13.11 In bacterial promoters, consensus sequences are found upstream of the start site, approximately at positions -10 and -35 .

and is often written simply as TATAAT (Figure 13.11). Remember that TATAAT is just the *consensus* sequence—representing the most commonly encountered nucleotides at each of these positions. In most prokaryotic promoters, the actual sequence is not TATAAT.

Another consensus sequence common to most bacterial promoters is TTGACA, which lies approximately 35 nucleotides upstream of the start site and is termed the **-35 consensus sequence** (see Figure 13.11). The nucleotides on either side of the -10 and -35 consensus sequences and those between them vary greatly from promoter to promoter, suggesting that these nucleotides are not very important in promoter recognition.

The function of these consensus sequences in bacterial promoters has been studied by inducing mutations at various positions within the consensus sequences and observing the effect of the changes on transcription. The results of these studies reveal that most base substitutions within the -10 and -35 consensus sequences reduce the rate of transcription; these substitutions are termed *down mutations* because they slow down the rate of transcription. Occasionally, a particular change in a consensus sequence increases the rate of transcription; such a change is called an *up mutation*.

As mentioned earlier, the sigma factor associates with the core enzyme (Figure 13.12a) to form a holoenzyme, which binds to the -35 and -10 consensus sequences in the DNA promoter (Figure 13.12b). Although it binds only the nucleotides of consensus sequences, the enzyme extends from -50 to $+20$ when bound to the promoter. The holoenzyme initially binds weakly to the promoter but then undergoes a change in structure that allows it to bind more tightly and unwind the double-stranded DNA (Figure 13.12c). Unwinding begins within the -10 consensus sequence and extends downstream for about 14 nucleotides, including the start site (from nucleotides -12 to $+2$).

Some bacterial promoters contain a third consensus sequence that also takes part in the initiation of transcription. Called the **upstream element**, this sequence contains a number of A–T pairs and is found at about -40 to -60 . A number of proteins may bind to sequences in and near the promoter; some stimulate the rate of transcription and others repress it. We will consider these proteins, which regulate gene expression, in Chapter 16. **TRY PROBLEM 22** →

CONCEPTS

A promoter is a DNA sequence that is adjacent to a gene and required for transcription. Promoters contain short consensus sequences that are important in the initiation of transcription.

✓ CONCEPT CHECK 5

What binds to the -10 consensus sequence found in most bacterial promoters?

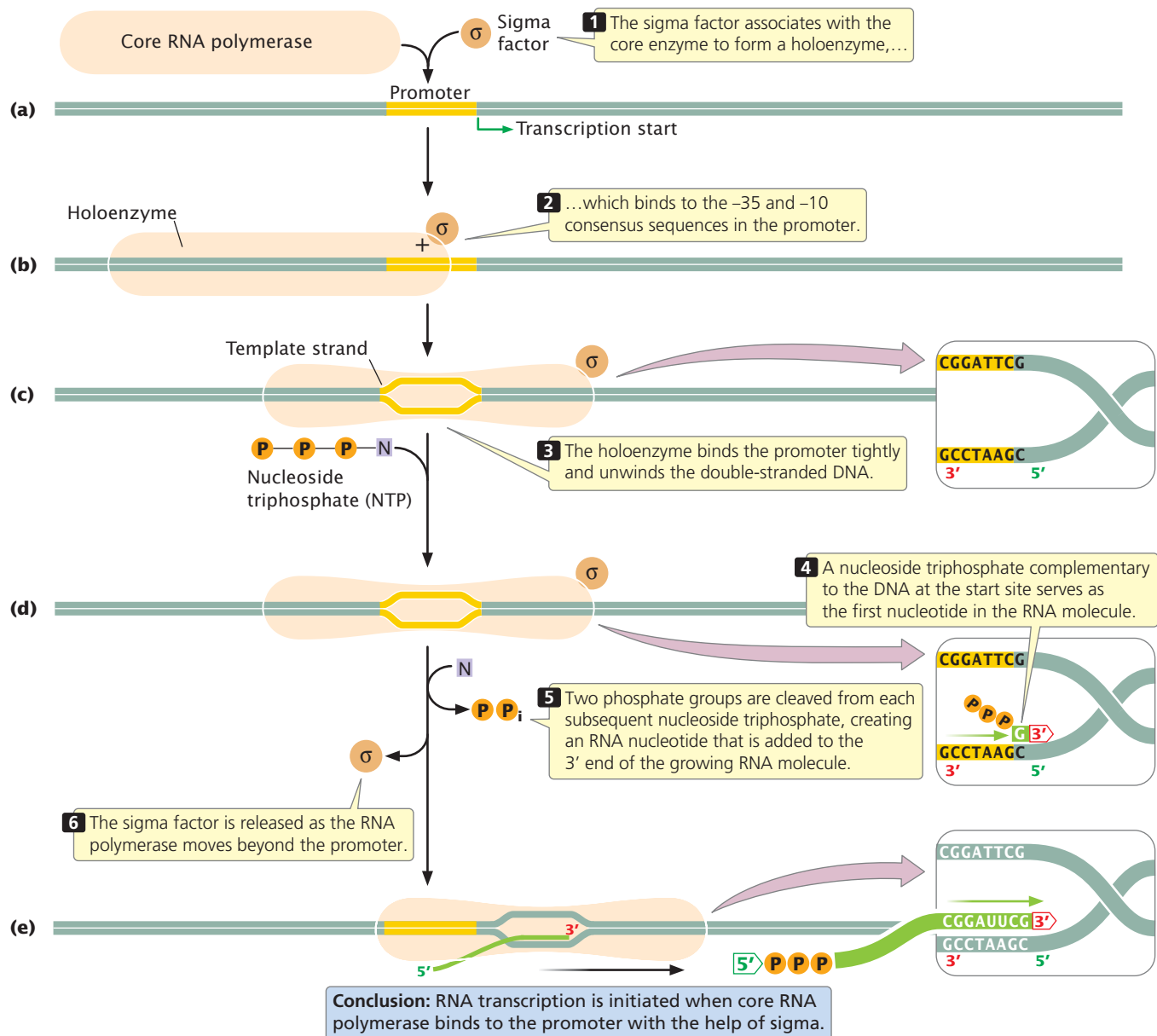
- The holoenzyme (core enzyme + sigma)
- The sigma factor alone
- The core enzyme alone
- mRNA

Initial RNA synthesis After the holoenzyme has attached to the promoter, RNA polymerase is positioned over the start site for transcription (at position $+1$) and has unwound the DNA to produce a single-stranded template. The orientation and spacing of consensus sequences on a DNA strand determine which strand will be the template for transcription and thereby determine the direction of transcription.

The position of the start site is determined not by the sequences located there but by the location of the consensus sequences, which positions RNA polymerase so that the enzyme's active site is aligned for the initiation of transcription at $+1$. If the consensus sequences are artificially moved upstream or downstream, the location of the starting point of transcription correspondingly changes.

To begin the synthesis of an RNA molecule, RNA polymerase pairs the base on a ribonucleoside triphosphate with its complementary base at the start site on the DNA template strand (Figure 13.12d). No primer is required to initiate the synthesis of the $5'$ end of the RNA molecule. Two of the three phosphate groups are cleaved from the ribonucleoside triphosphate as the nucleotide is added to the $3'$ end of the growing RNA molecule. However, because the $5'$ end of the first ribonucleoside triphosphate does not take part in the formation of a phosphodiester bond, all three of its phosphate groups remain. An RNA molecule therefore possesses, at least initially, three phosphate groups at its $5'$ end (Figure 13.12e).

Often, in the course of initiation, RNA polymerase repeatedly generates and releases short transcripts, from 2 to 6 nucleotides in length, while still bound to the promoter. After



13.12 Transcription in bacteria is carried out by RNA polymerase, which must bind to the sigma factor to initiate transcription.

several abortive attempts, the polymerase synthesizes an RNA molecule from 9 to 12 nucleotides in length, which allows the RNA polymerase to transition to the elongation stage.

Elongation

At the end of initiation, RNA polymerase undergoes a change in conformation (shape) and thereafter is no longer able to bind to the consensus sequences in the promoter. This change allows the polymerase to escape from the promoter and begin transcribing downstream. The sigma subunit is usually released after initiation, although some populations of RNA polymerase may retain sigma throughout elongation.

As it moves downstream along the template, RNA polymerase progressively unwinds the DNA at the leading (downstream) edge of the transcription bubble, joining

nucleotides to the RNA molecule according to the sequence on the template, and rewinds the DNA at the trailing (upstream) edge of the bubble. In bacterial cells at 37°C, about 40 nucleotides are added per second. This rate of RNA synthesis is much lower than that of DNA synthesis, which is 1000 to 2000 nucleotides per second in bacterial cells.

The transcription bubble

Transcription takes place within a short stretch of about 18 nucleotides of unwound DNA—the transcription bubble. Within this region, RNA is continuously synthesized, with single-stranded DNA used as a template. About 8 nucleotides of newly synthesized RNA are paired with the DNA-template nucleotides at any one time. As the transcription apparatus moves down the DNA template, it generates positive supercoiling ahead of the transcription bubble and negative supercoiling behind

it. Topoisomerase enzymes probably relieve the stress associated with the unwinding and rewinding of DNA in transcription, as they do in DNA replication.

Transcriptional pausing A number of features of RNA or DNA, such as secondary structures and specific sequences, can cause RNA polymerase to pause for a second or more in the elongation stage of transcription. Transitory pauses in transcription are important in the coordination of transcription and translation in bacteria and in the

coordination of RNA processing in eukaryotes. Pausing also affects the rates of RNA synthesis. Sometimes a pause may be stabilized by sequences in the DNA that ultimately lead to the termination of transcription (see the section on termination next).

Accuracy of transcription Although RNA polymerase is quite accurate in incorporating nucleotides into the growing RNA chain, errors do occasionally arise. Recent research has demonstrated that RNA polymerase is capable of a type of proofreading in the course of transcription. When RNA polymerase incorporates a nucleotide that does not match the DNA template, it backs up and cleaves the last two nucleotides (including the misincorporated nucleotide) from the growing RNA chain. RNA polymerase then proceeds forward, transcribing the DNA template again.

CONCEPTS

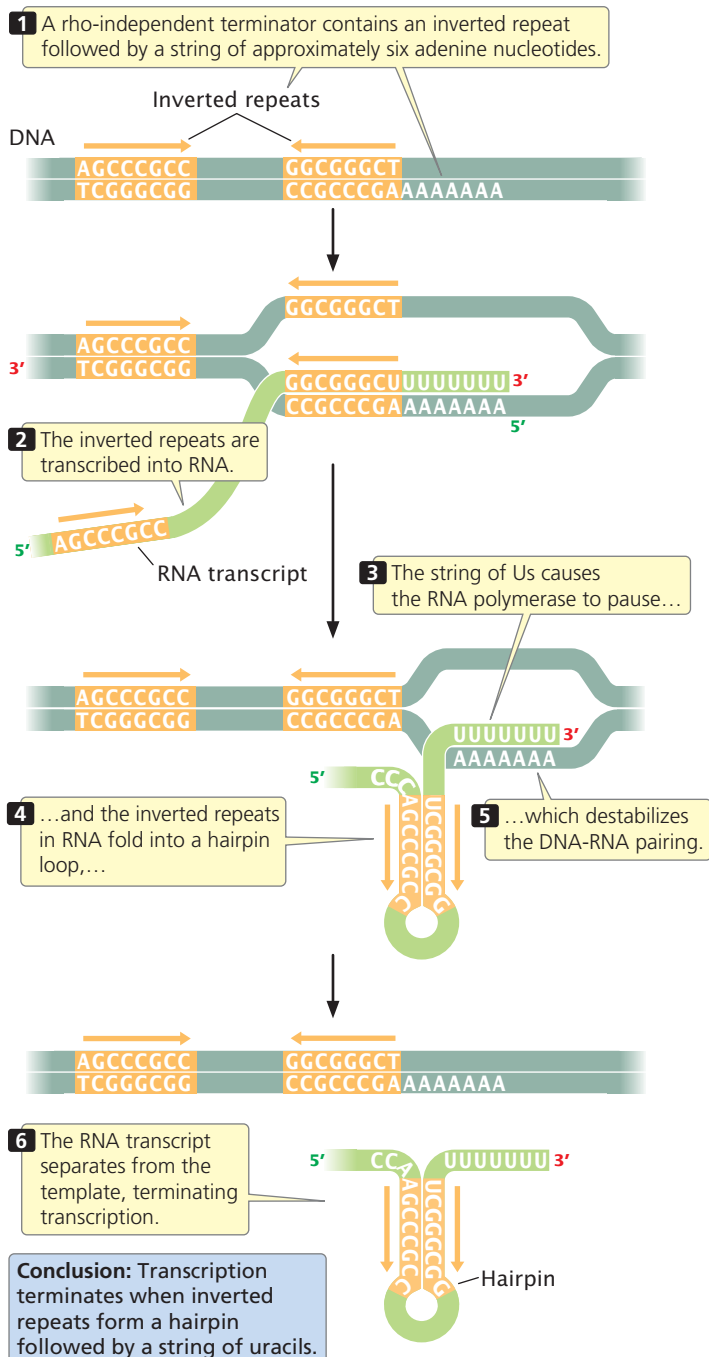
Transcription is initiated at the start site, which, in bacterial cells, is set by the binding of RNA polymerase to the consensus sequences of the promoter. No primer is required. Transcription takes place within the transcription bubble. DNA is unwound ahead of the bubble and rewound behind it. There are frequent pauses in the process of transcription.

Termination

RNA polymerase adds nucleotides to the 3' end of the growing RNA molecule until it transcribes a terminator. Most terminators are found upstream of the site at which termination actually takes place. Transcription therefore does not suddenly stop when polymerase reaches a terminator, as does a car stopping at a stop sign. Rather, transcription stops after the terminator has been transcribed, like a car that stops only after running over a speed bump. At the terminator, several overlapping events are needed to bring an end to transcription: RNA polymerase must stop synthesizing RNA, the RNA molecule must be released from RNA polymerase, the newly made RNA molecule must dissociate fully from the DNA, and RNA polymerase must detach from the DNA template.

Bacterial cells possess two major types of terminators. **Rho-dependent terminators** are able to cause the termination of transcription only in the presence of an ancillary protein called the **rho factor**. **Rho-independent terminators** (also known as intrinsic terminators) are able to cause the end of transcription in the absence of rho.

Rho-independent terminators Rho-independent terminators, which make up about 50 percent of all terminators in prokaryotes, have two common features. First, they contain inverted repeats (sequences of nucleotides on one strand that are inverted and complementary). When inverted repeats have been transcribed into RNA, a hairpin secondary structure forms (Figure 13.13). Second, in rho-independent terminators, a string of seven to nine adenine nucleotides follows the second inverted repeat in the tem-

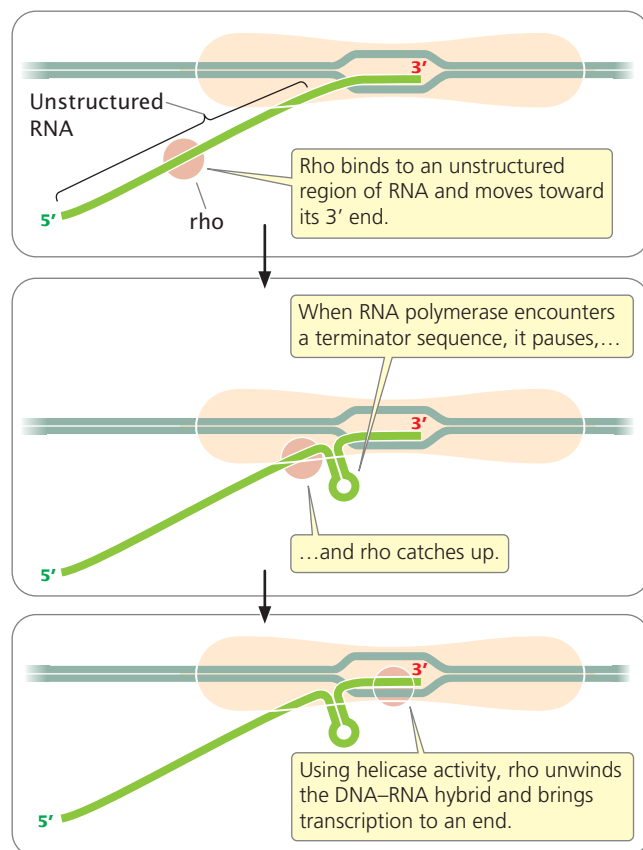


13.13 Rho-independent termination in bacteria is a multistep process.

plate DNA. Their transcription produces a string of uracil nucleotides after the hairpin in the transcribed RNA.

The string of uracils in the RNA molecule causes the RNA polymerase to pause, allowing time for the hairpin structure to form. Evidence suggests that the formation of the hairpin destabilizes the DNA–RNA pairing, causing the RNA molecule to separate from its DNA template. Separation may be facilitated by the adenine–uracil base pairings, which are relatively weak compared with other types of base pairings. When the RNA transcript has separated from the template, RNA synthesis can no longer continue (see Figure 13.13).

Rho-dependent terminators Rho-dependent terminators have two features: (1) DNA sequences that produce a pause in transcription and (2) a DNA sequence that encodes a stretch of RNA upstream of the terminator that is devoid of any secondary structures. This unstructured RNA serves as a binding site for the rho protein, which binds the RNA and moves toward its 3' end, following the RNA polymerase (Figure 13.14). When RNA polymerase encounters the terminator, it pauses, allowing rho to catch up. The rho protein has helicase activity, which it uses to unwind the RNA–DNA hybrid in the transcription bubble, bringing transcription to an end. **TRY PROBLEM 26** →



13.14 The termination of transcription in some bacterial genes requires the presence of the rho protein.

Polycistronic mRNA In bacteria, a group of genes is often transcribed into a single RNA molecule, which is termed a **polycistronic RNA**. Thus, polycistronic RNA is produced when a single terminator is present at the end of a group of several genes that are transcribed together, instead of each gene having its own terminator. Typically, each eukaryotic gene is transcribed and terminated separately, and so polycistronic mRNA is uncommon in eukaryotes.



www.whfreeman.com/pierce4e

The process of transcriptional initiation, elongation, and termination are illustrated in Animation 13.1, which shows how the different parts of the transcriptional unit interact to bring about the complete synthesis of an RNA molecule.

CONCEPTS

Transcription ends after RNA polymerase transcribes a terminator. Bacterial cells possess two types of terminator: a rho-independent terminator, which RNA polymerase can recognize by itself; and a rho-dependent terminator, which RNA polymerase can recognize only with the help of the rho protein.

✓ CONCEPT CHECK 6

What characteristics are most commonly found in rho-independent terminators?

CONNECTING CONCEPTS

The Basic Rules of Transcription

Before we examine the process of eukaryotic transcription, let's pause to summarize some of the general principles of bacterial transcription.

1. Transcription is a selective process; only certain parts of the DNA are transcribed at any one time.
2. RNA is transcribed from single-stranded DNA. Within a gene, only one of the two DNA strands—the template strand—is normally copied into RNA.
3. Ribonucleoside triphosphates are used as the substrates in RNA synthesis. Two phosphate groups are cleaved from a ribonucleoside triphosphate, and the resulting nucleotide is joined to the 3'-OH group of the growing RNA strand.
4. RNA molecules are antiparallel and complementary to the DNA template strand. Transcription is always in the 5'→3' direction, meaning that the RNA molecule grows at the 3' end.
5. Transcription depends on RNA polymerase—a complex, multimeric enzyme. RNA polymerase consists of a core enzyme, which is capable of synthesizing RNA, and other subunits that may join transiently to perform additional functions.
6. A sigma factor enables the core enzyme of RNA polymerase to bind to a promoter and initiate transcription.
7. Promoters contain short sequences crucial in the binding of RNA polymerase to DNA; these consensus sequences are interspersed with nucleotides that play no known role in transcription.
8. RNA polymerase binds to DNA at a promoter, begins transcribing at the start site of the gene, and ends transcription after a terminator has been transcribed.

13.4 Eukaryotic Transcription Is Similar to Bacterial Transcription but Has Some Important Differences

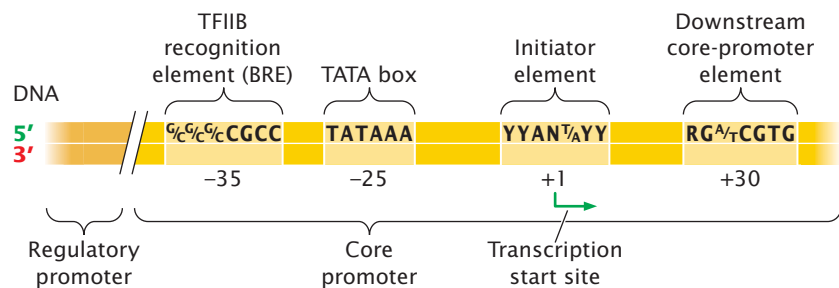
Transcription in eukaryotes is similar to bacterial transcription in that it includes initiation, elongation, and termination, and the basic principles of transcription already outlined apply to eukaryotic transcription. However, there are some important differences. Eukaryotic cells possess three different RNA polymerases, each of which transcribes a different class of RNA and recognizes a different type of promoter. Thus, a generic promoter cannot be described for eukaryotic cells, as was done for bacterial cells; rather, a promoter's description depends on whether the promoter is recognized by RNA polymerase I, II, or III. Another difference is in the nature of promoter recognition and initiation. Many proteins take part in the binding of eukaryotic RNA polymerases to DNA templates, and the different types of promoters require different proteins.

Transcription and Nucleosome Structure

Transcription requires sequences on DNA to be accessible to RNA polymerase and other proteins. However, in eukaryotic cells, DNA is complexed with histone proteins in highly compressed chromatin (see Figure 11.4). How can the proteins necessary for transcription gain access to eukaryotic DNA when it is complexed with histones?

The answer to this question is that chromatin structure is modified before transcription so that the DNA is in a more open configuration and is more accessible to the transcription machinery. Several types of proteins have roles in chromatin modification. Acetyltransferases add acetyl groups to amino acids at the ends of the histone proteins, which destabilizes nucleosome structure and makes the DNA more accessible. Other types of histone modification also can affect chromatin packing. In addition, proteins called chromatin-remodeling proteins may bind to the chromatin and displace nucleosomes from promoters and other regions important for transcription. We will take a closer look at the role of changes in chromatin structure associated with gene expression in Chapter 17.

13.15 The promoters of genes transcribed by RNA polymerase II consist of a core promoter and a regulatory promoter that contain consensus sequences. Not all the consensus sequences shown are found in all promoters.



CONCEPTS

The initiation of transcription requires modification of chromatin structure so that DNA is accessible to the transcriptional machinery.

Promoters

A significant difference between bacterial and eukaryotic transcription is the existence of three different eukaryotic RNA polymerases, which recognize different types of promoters. In bacterial cells, the holoenzyme (RNA polymerase plus the sigma factor) recognizes and binds directly to sequences in the promoter. In eukaryotic cells, promoter recognition is carried out by accessory proteins that bind to the promoter and then recruit a specific RNA polymerase (I, II, or III) to the promoter.

One class of accessory proteins comprises **general transcription factors**, which, along with RNA polymerase, form the **basal transcription apparatus**—a group of proteins that assemble near the start site and are sufficient to initiate minimal levels of transcription. Another class of accessory proteins consists of **transcriptional activator proteins**, which bind to specific DNA sequences and bring about higher levels of transcription by stimulating the assembly of the basal transcription apparatus at the start site.

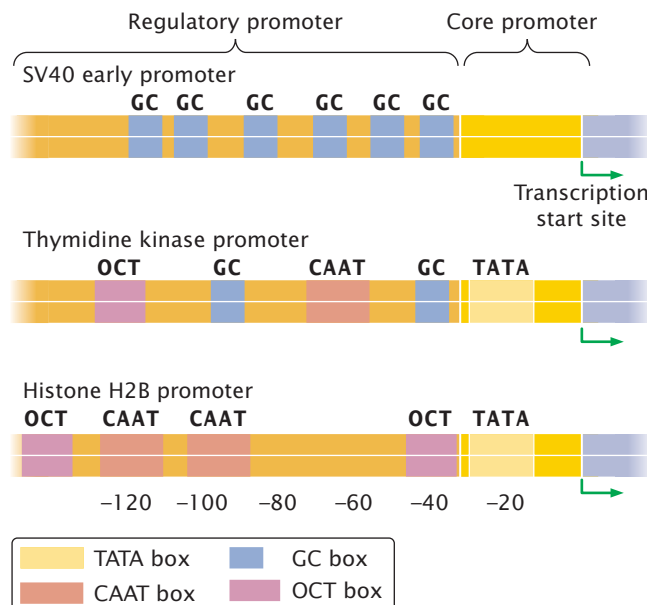
We will focus our attention on promoters recognized by RNA polymerase II, which transcribes the genes that encode proteins. A promoter for a gene transcribed by RNA polymerase II typically consists of two primary parts: the core promoter and the regulatory promoter.

Core promoter The **core promoter** is located immediately upstream of the gene (Figure 13.15) and is the site to which the basal transcription apparatus binds. The core promoter typically includes one or more consensus sequences. One of the most common of these sequences is the **TATA box**, which has the consensus sequence TATAAA and is located from -25 to -30 bp upstream of the start site. Additional consensus sequences that may be found in the core promoters of genes transcribed by RNA polymerase II are

shown in Figure 13.15. These consensus sequences are recognized by transcription factors that bind to them and serve as a platform for the assembly of the basal transcription apparatus.

Regulatory promoter The **regulatory promoter** is located immediately upstream of the core promoter. A variety of different consensus sequences can be found in the regulatory promoters, and they can be mixed and matched in different combinations (Figure 13.16). Transcriptional activator proteins bind to these sequences and either directly or indirectly make contact with the basal transcription apparatus and affect the rate at which transcription is initiated. Transcriptional activator proteins also regulate transcription by binding to more-distant sequences called **enhancers**. The DNA between an enhancer and the promoter loops out, and so transcriptional activator proteins bound to the enhancer can interact with the basal transcription machinery at the core promoter. Enhancers will be discussed in more detail in Chapter 17.

Polymerase I and III promoters RNA polymerase I and RNA polymerase III each recognize promoters that are distinct from those recognized by RNA polymerase II. For example, promoters for small rRNA and tRNA genes, transcribed by RNA polymerase III, contain **internal promoters** that are downstream of the start site and are transcribed into the RNA.



13.16 Consensus sequences in the promoters of three eukaryotic genes. These promoters illustrate the principle that consensus sequences can be mixed and matched in different combinations to yield a functional eukaryotic promoter.

CONCEPTS

General transcription factors and RNA polymerase assemble into the basal transcription apparatus, which binds to DNA near the start site and is necessary for transcription to take place at minimal levels. Additional proteins called transcriptional activators bind to other consensus sequences in promoters and enhancers and affect the rate of transcription.

✓ CONCEPT CHECK 7

What is the difference between the core promoter and the regulatory promoter?

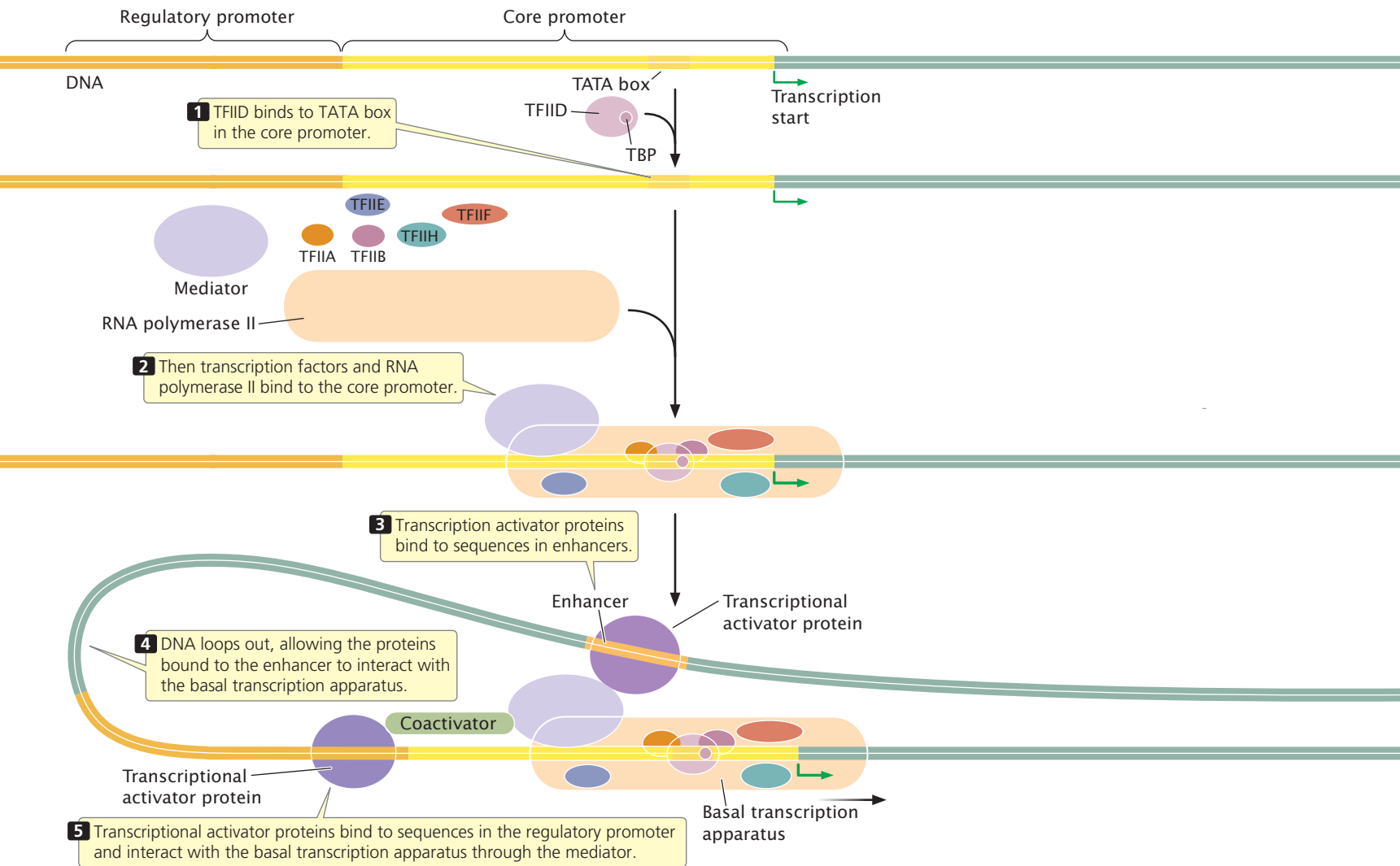
- Only the core promoter has consensus sequences.
- The regulatory promoter is farther upstream of the gene.
- Transcription factors bind to the core promoter; transcriptional activator proteins bind to the regulatory promoter.
- Both b and c.

Initiation

Transcription in eukaryotes is initiated through the assembly of the transcriptional machinery on the promoter. This machinery consists of RNA polymerase II and a series of transcription factors that form a giant complex consisting of 50 or more polypeptides. Assembly of the transcription machinery begins when regulatory proteins bind DNA near the promoter and modify the chromatin structure so that transcription can take place. These proteins and other regulatory proteins then recruit the basal transcriptional apparatus to the core promoter.

The basal transcription apparatus consists of RNA polymerase, a series of general transcription factors, and a complex of proteins known as the mediator (Figure 13.17). The general transcription factors include TFIIA, TFIIB, TFIID, TFIIE, TFIIIF, and TFIIH, in which TFII stands for transcription factor for RNA polymerase II and the final letter designates the individual factor.

RNA polymerase II and the general transcription factors assemble at the core promoter, forming a pre-initiation complex that is analogous to the closed complex seen in bacterial initiation. Recall that, in bacteria, the sigma factor recognizes and binds to the promoter sequence. In eukaryotes, the function of sigma is replaced by that of the general transcription factors. A first step in initiation is the binding of TFIID to the TATA box on the DNA template. TFIID consists of at least nine polypeptides. One of them is the **TATA-binding protein** (TBP), which recognizes and binds to the TATA consensus sequence. The TATA-binding protein binds to the minor groove and straddles the DNA as a

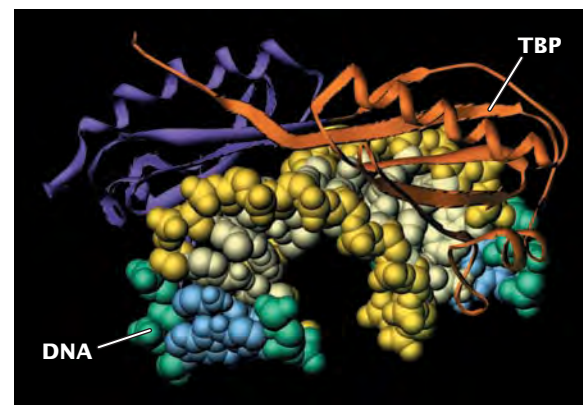


13.17 Transcription is initiated at RNA polymerase II promoters. Transcription is initiated when the TFIID transcription factor binds to the TATA box, followed by the binding of a preassembled holoenzyme containing general transcription factors, RNA polymerase II, and the mediator. TBP stands for TATA-binding protein.

molecular saddle (**Figure 13.18**), bending the DNA and partly unwinding it. Other transcription factors bind to additional consensus sequences in the core promoter and to RNA polymerase and position the polymerase over the transcription start site.

After the RNA polymerase and transcription factors have assembled on the core promoter, conformational changes take place in both the DNA and the polymerase. These changes cause from 11 to 15 bp of DNA surrounding the transcription start site to separate, producing the single-stranded DNA that will serve as a template for transcription.

The single-stranded DNA template is positioned within the active site of RNA polymerase, creating a structure called the open complex. After the open complex has formed, the synthesis of RNA begins as phosphate groups are cleaved off nucleoside triphosphates and nucleotides are joined



13.18 The TATA-binding protein (TBP) binds to the minor groove of DNA, straddling the double helix of DNA like a saddle.

together to form an RNA molecule. As in bacterial transcription, RNA polymerase may generate and release several short RNA molecules in abortive transcription before the polymerase initiates the synthesis of a full-length RNA molecule.

TRY PROBLEM 32 →

CONCEPTS

Transcription is initiated when the basal transcription apparatus, consisting of RNA polymerase and transcription factors, assembles on the core promoter and becomes an open complex.

✓ CONCEPT CHECK 8

What is the role of TFIID in transcription initiation?

Elongation

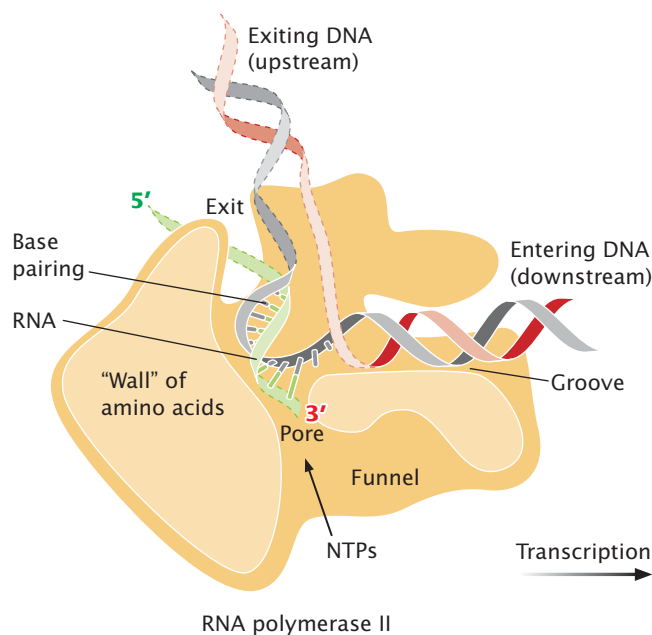
After about 30 bp of RNA have been synthesized, the RNA polymerase leaves the promoter and enters the elongation stage of transcription. Many of the transcription factors are left behind at the promoter and can serve to quickly reinitiate transcription with another RNA polymerase enzyme.

The molecular structure of eukaryotic RNA polymerase II and how it functions during elongation have been revealed through the work of Roger Kornberg and his colleagues, for which Kornberg was awarded a Nobel Prize in chemistry in 2006. The RNA polymerase maintains a transcription bubble during elongation, in which about eight nucleotides of RNA remain base paired with the DNA template strand. The DNA double helix enters a cleft in the polymerase and is gripped by jawlike extensions of the enzyme (**Figure 13.19**). The two strands of the DNA are unwound and RNA nucleotides that are complementary to the template strand are added to the growing 3' end of the RNA molecule. As it funnels through the polymerase, the DNA–RNA hybrid hits a wall of amino acids and bends at almost a right angle; this bend positions the end of the DNA–RNA hybrid at the active site of the polymerase, and new nucleotides are added to the 3' end of the growing RNA molecule. The newly synthesized RNA is separated from the DNA and runs through another groove before exiting from the polymerase.

Termination

The three eukaryotic RNA polymerases use different mechanisms for termination. RNA polymerase I requires a termination factor like the rho factor utilized in the termination of some bacterial genes. Unlike rho, which binds to the newly transcribed RNA molecule, the termination factor for RNA polymerase I binds to a DNA sequence downstream of the termination site.

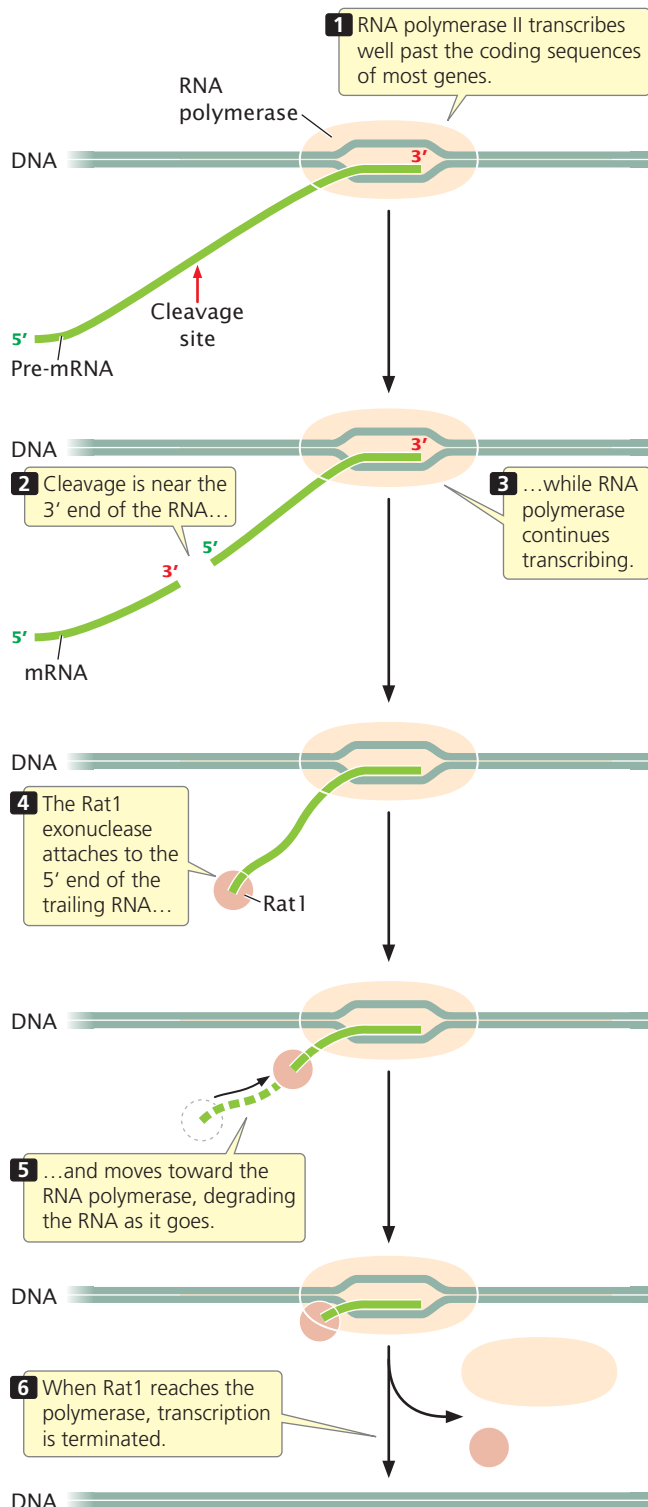
RNA polymerase III ends transcription after transcribing a terminator sequence that produces a string of uracil



13.19 The structure of RNA polymerase II is a source of insight into its function. The DNA double helix enters the polymerase through a groove and unwinds. The DNA–RNA duplex is bent at a right angle, which positions the 3' end of the RNA at the active site of the enzyme. New nucleotides are added to the 3' end of the RNA.

nucleotides in the RNA molecule, like that produced by the rho-independent terminators of bacteria. Unlike rho-independent terminators in bacterial cells, however, RNA polymerase III does not require that a hairpin structure precede the string of Us.

The termination of transcription by RNA polymerase II is not at specific sequences. Instead, RNA polymerase II often continues to synthesize RNA hundreds or even thousands of nucleotides past the coding sequence necessary to produce the mRNA. As we will see in Chapter 14, the end of pre-mRNA is cleaved at a specific site, designated by a consensus sequence, while transcription is still taking place at the 3' end of the molecule. Cleavage cuts the pre-mRNA into two pieces: the mRNA that will eventually encode the protein and another piece of RNA that has its 5' end trailing out of the RNA polymerase (**Figure 13.20**). An enzyme called Rat1 attaches to the 5' end of this RNA and moves toward the 3' end where RNA polymerase continues the transcription of RNA. Rat1 is a 5'→3' exonuclease—an enzyme capable of degrading RNA in the 5'→3' direction. Like a guided torpedo, Rat1 homes in on the polymerase, chewing up the RNA as it moves. When Rat1 reaches the transcriptional machinery, transcription terminates. Note that this mechanism is similar to that of rho-dependent termination in bacteria (see Figure 13.14), except that rho does not degrade the RNA molecule.



13.20 Termination of transcription by RNA polymerase II requires the Rat1 exonuclease. Cleavage of the pre-mRNA produces a 5' end to which Rat1 attaches. Rat1 degrades the RNA molecule in the 5→3' direction. When Rat1 reaches the polymerase, transcription is halted.

CONCEPTS

The different eukaryotic RNA polymerases use different mechanisms of termination. Transcription at genes transcribed by RNA polymerase II is terminated when an exonuclease enzyme attaches to the cleaved 5' end of the RNA, moves down the RNA, and reaches the polymerase enzyme.

✓ CONCEPT CHECK 9

How are the processes of RNA polymerase II termination and rho-dependent termination in bacteria similar and how are they different?

13.5 Transcription in Archaea Is More Similar to Transcription in Eukaryotes Than to Transcription in Eubacteria

Some 2 billion to 3 billion years ago, life diverged into three lines of evolutionary descent: the eubacteria, the archaea, and the eukaryotes (see Chapter 2). Although eubacteria and archaea are superficially similar—both are unicellular and lack a nucleus—the results of studies of their DNA sequences and other biochemical properties indicate that they are as distantly related to each other as they are to eukaryotes. The evolutionary distinction between archaea, eubacteria, and eukaryotes is clear. However, did eukaryotes first diverge from an ancestral prokaryote, with the later separation of prokaryotes into eubacteria and archaea, or did the archaea and the eubacteria split first, with the eukaryotes later evolving from one of these groups?

Studies of transcription in eubacteria, archaea, and eukaryotes have yielded important findings about the evolutionary relationships of these organisms. Archaea, like eubacteria, have a single RNA polymerase, but this enzyme is most similar to the RNA polymerases of eukaryotes. Archaea possess a TATA-binding protein, a critical transcription factor in all three of the eukaryotic polymerases but not present in eubacteria. TBP binds the TATA box in archaea with the help of another transcription factor, TFIIB, which also is found in eukaryotes but not in eubacteria.

Transcription, one of the most basic of life processes, has strong similarities in eukaryotes and archaea, suggesting that these two groups are more closely related to each other than either is to the eubacteria. This conclusion is supported by other data, including those obtained from a comparison of gene sequences.

CONCEPTS

The process of transcription in archaea has many similarities to transcription in eukaryotes.