Transcription: DNA-Directed RNA Synthesis

7.2

During transcription, the DNA code is chemically rewritten as an RNA code. This occurs within the nucleus. Transcription is divided into three sequential processes: initiation, elongation, and termination. Refer to **Figure 1** (next page) as you read the next three sections describing these three processes of transcription.

Initiation

In both prokaryotes and eukaryotes, the process of transcription begins when the enzyme RNA polymerase binds to the DNA and unwinds it near the beginning of a gene (Figure 1, Step 1). The binding occurs at a **promoter**: a specialized sequence on one strand of DNA, located just upstream from the start of the gene. A key element of the promoter in eukaryotes is the **TATA box**: a section of DNA with a high percentage of thymine and adenine bases, which is recognized by RNA polymerase. Prokaryotes have a TATAAT sequence instead of a TATA box for this purpose. Adenine and thymine share only two hydrogen bonds, whereas guanine and cytosine share three hydrogen bonds. Since less energy is needed to break two bonds, the RNA polymerase expends less energy opening up the DNA helix if it possesses a high concentration of adenine and thymine base pairs. The part of the gene that is to be transcribed into RNA is called the transcription unit.

promoter a nucleotide sequence that lies just before a gene and allows for the binding of RNA polymerase

TATA box a region of the promoter that enables the binding of RNA polymerase

Elongation

Once the RNA polymerase binds to the promoter and opens the DNA double helix, it starts to build the single-stranded RNA molecule. RNA polymerase, unlike DNA polymerase, can begin making the complementary copy without needing a primer to be already in place. RNA is made in the 5′→3′ direction, using the 3′→5′ DNA strand as a template strand. The opposite strand of DNA—the strand that is not being copied—is known as the **coding strand**, since it contains the same base-pair sequence as the new RNA molecule, except for the absence of uracil and the presence of thymine. Remember that the template strand contains the sequence that is complementary to the sequence that is going to be transcribed. Therefore, the beginning of the RNA strand is the 5′ end, and the other end is the 3′ end.

coding strand the DNA strand that is not being copied but contains the same sequence as the new RNA molecule

As RNA polymerase moves along the DNA, it unwinds the DNA at the forward end of the enzyme. The new RNA molecule elongates as nucleotides are added, one by one (Figure 1, Step 2). The new RNA molecule winds temporarily with the template strand of the DNA into a hybrid RNA–DNA double helix. Beyond this short region of pairing, the growing RNA strand unwinds from the DNA and extends from the RNA polymerase as a single nucleotide chain. As the RNA polymerase passes, the DNA double helix reforms.

Once an RNA polymerase molecule has started transcription and progressed past the beginning of a gene, another molecule of RNA polymerase may start producing another RNA molecule if there is room at the promoter. Most genes undergoing transcription have many RNA polymerase molecules spaced closely along them, and each molecule makes an RNA transcript. When cells require a particular protein, they usually need to produce thousands or even millions of copies. For example, a single red blood cell contains 375 million hemoglobin molecules. The process of making hemoglobin would be very slow if the gene had only one RNA polymerase enzyme making one mRNA molecule at a time. Many copies of mRNA are made so that many ribosomes can mass-produce the protein required.

Termination

termination sequence a sequence of bases at the end of a gene that signals the RNA polymerase to stop transcribing

The transcription of a protein-coding gene is terminated when RNA polymerase recognizes a **termination sequence** (Figure 1, Step 3). In prokaryotes, one termination mechanism involves a protein binding to the mRNA and stopping transcription.

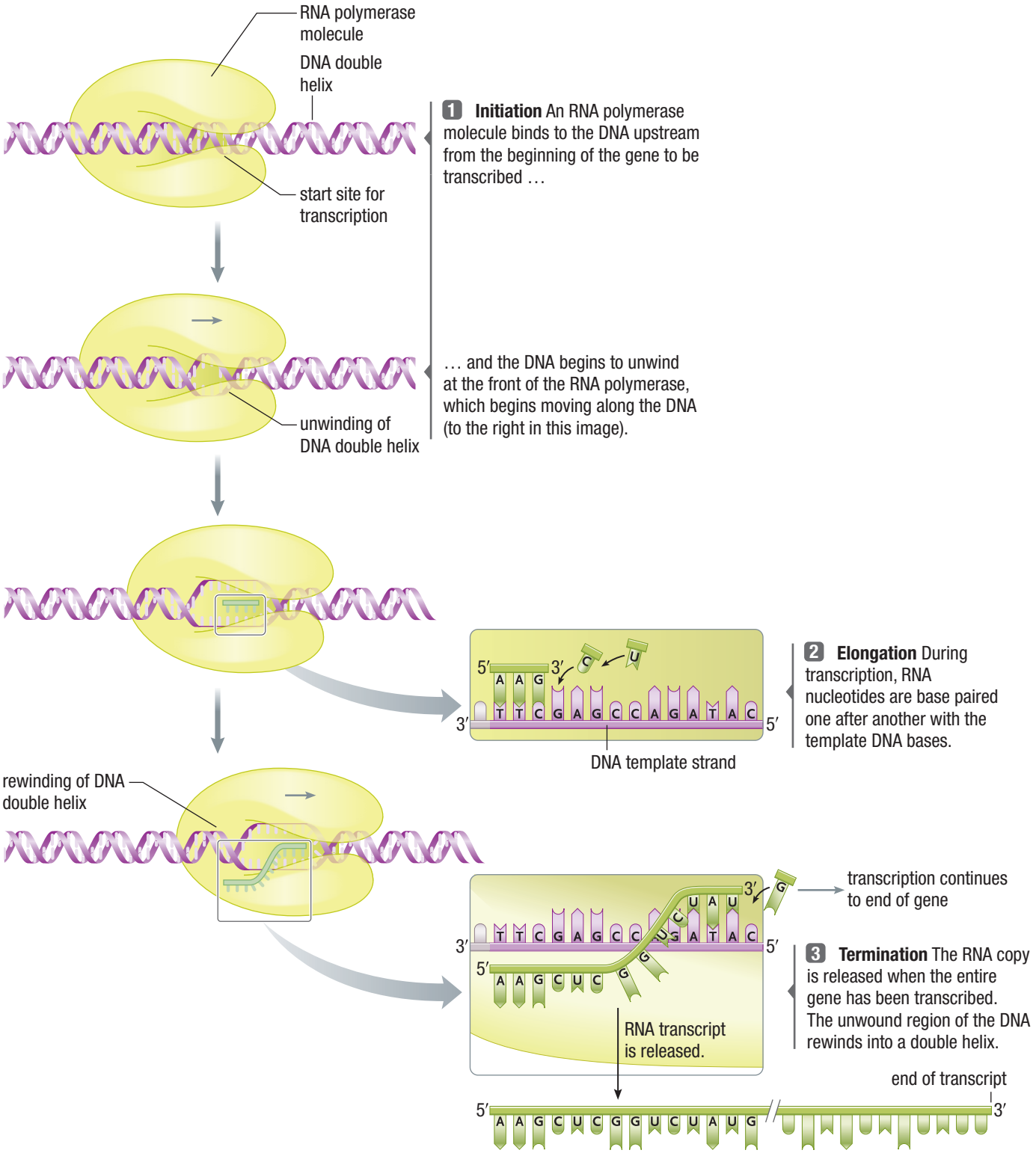

Figure 1 Transcribing a gene into precursor mRNA in eukaryotes. There are three stages: initiation, elongation, and termination. RNA polymerase moves along the gene, separating the two DNA strands to allow RNA synthesis in the 5′ → 3′ direction. The 3′ → 5′ DNA strand is used as a template.

Another termination mechanism involves the mRNA binding with itself in a hairpin loop and stopping transcription. In eukaryotes, one termination sequence is a string of adenines, which are transcribed as a string of uracils on the RNA. Nuclear proteins bind to the polyuracil site and stop transcription. The newly synthesized RNA then dissociates from the DNA template strand. Transcription ceases, and the RNA polymerase is free to bind to another promoter region and transcribe another gene.

Post-transcriptional Modifications

At this point, the newly transcribed eukaryotic RNA, known as the primary transcript or precursor mRNA (or pre-mRNA), is vulnerable to the enzymes and conditions outside the cell nucleus. The pre-mRNA must undergo additional modifications before it can exit the nucleus and reach the ribosome.

One modification is the addition of a chain of 50 to 250 adenine nucleotides, one nucleotide at a time, to the 3′ end by an enzyme called poly-A polymerase (**Figure 2**). The chain of adenine nucleotides, called the **poly(A) tail**, enables mRNA to be translated efficiently and protects it from attack by RNA-digesting enzymes in the cytosol.

poly(A) tail a chain of adenine nucleotides that are added to the 3′ end of the pre-mRNA molecule to protect it from enzymes in the cytosol

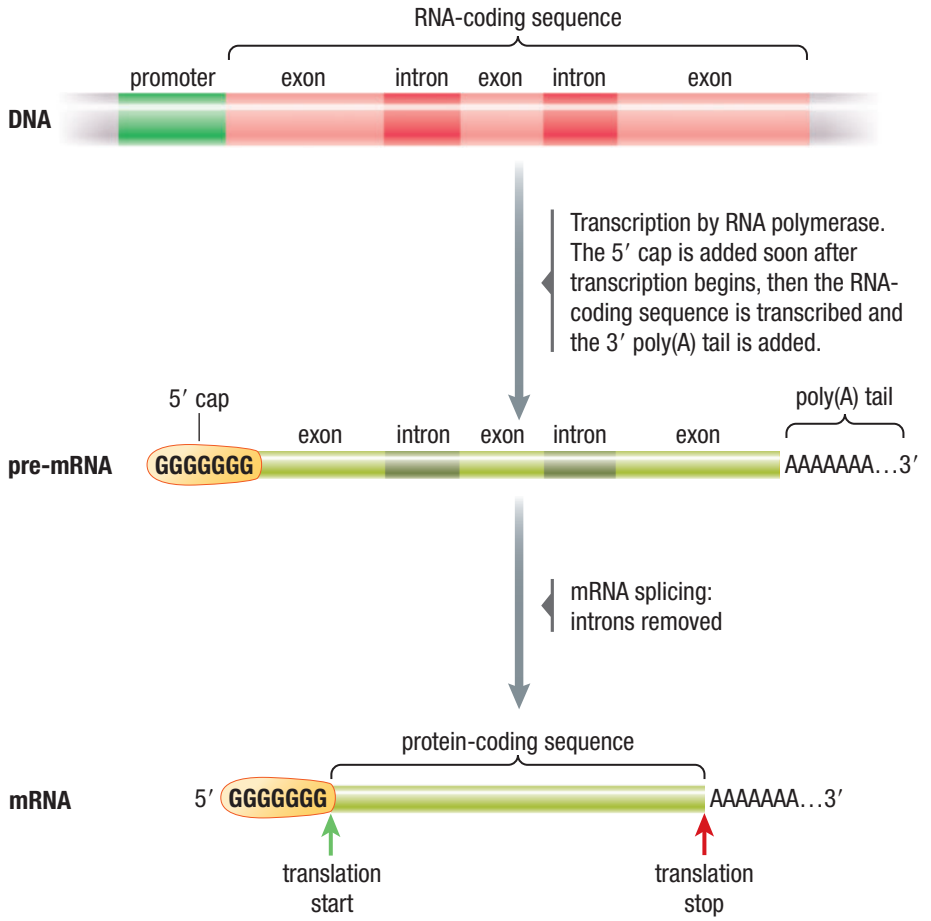

Figure 2 The relationship between a eukaryotic protein-coding gene, the pre-mRNA transcribed from it, and the mRNA processed from the pre-mRNA

Modifications are also made at the beginning of the pre-mRNA transcript, where a **5′ cap**, consisting of seven Gs, is added by a different enzyme complex. The 5′ cap functions as the initial attachment site for mRNAs to ribosomes, to allow for translation. This whole process is known as capping and tailing.

5′ cap a sequence of seven Gs that is added to the start of a pre-mRNA molecule; ribosomes recognize this site and use it as the site of initial attachment

exon a sequence of DNA or RNA that codes for part of a gene

intron a non-coding sequence of DNA or RNA

The precursor mRNA is still not ready to exit the nucleus. Further modifications need to be made. The DNA of a eukaryotic gene is composed of coding regions known as **exons** and non-coding regions known as **introns**. The introns are interspersed among the exons and are transcribed into pre-mRNAs (Figure 2). However, the introns do not code for part of the protein. If they were left in the mRNA, they would alter the sequence of the amino acids that are used to build the protein. This would result in additional amino acids and a protein that would not fold as it should and therefore would not function correctly. Therefore, the introns are deleted and the exons are retained in fully processed mRNA. The majority of known eukaryotic genes contain at least one intron, and some contain more than 60. Prokaryrote DNA does not contain any introns.

spliceosome an enzyme-protein complex that removes introns from the mRNA

small ribonucleoprotein (snRNP) a protein that binds to introns and signals them for removal

In the nucleus, a process called mRNA splicing removes the introns from pre-mRNAs and joins the exons together. mRNA splicing occurs in a **spliceosome**: a complex formed between the pre-mRNA and a handful of **small ribonucleoproteins** called **snRNPs** (pronounced snurps). The snRNPs bind in a particular order to an intron in the pre-mRNA (**Figure 3**). The first snRNPs are those that recognize and form complementary base pairs with mRNA sequences at the junctions of the intron and adjacent exons. Other snRNPs are then recruited, causing the intron to loop out and bring the two exon ends close together. At this point, an active spliceosome has been formed, releasing the intron and joining together the two exons. The cutting and splicing are so exact that not a single base of an intron is retained in the finished mRNA, and not a single base is removed from the exons.

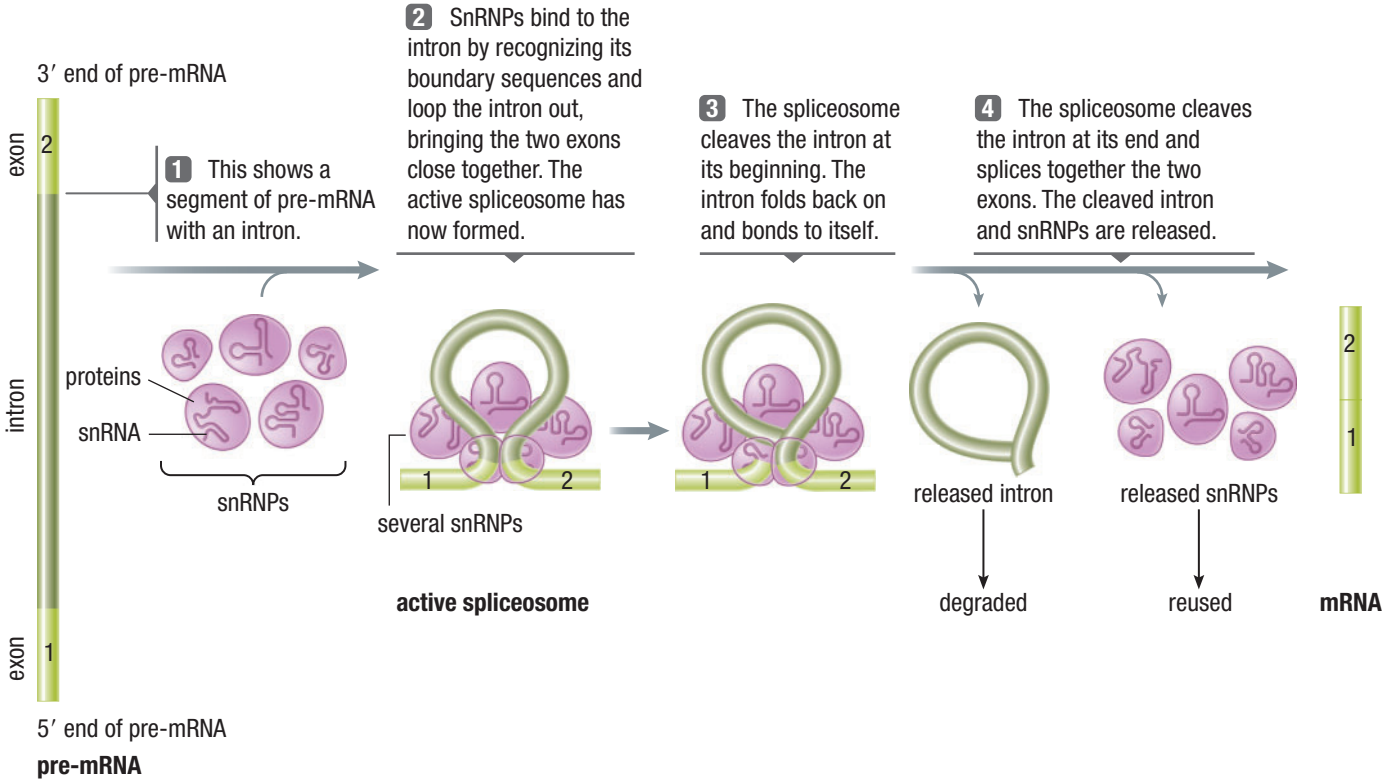

Figure 3 Introns are removed from eukaryotic precursor mRNA, and exons are joined together.

alternative splicing a process that produces different mRNAs from pre-mRNA (exons and introns), allowing more than one possible polypeptide to be made from a single gene

Exons may be joined in different combinations to produce different mRNAs from a single DNA gene sequence. A mechanism called **alternative splicing** greatly increases the number and variety of proteins encoded by a single gene. According to current estimates, three-quarters of all human pre-mRNAs are subjected to alternative splicing. In each case, the different mRNAs that are produced from the parent pre-mRNA are translated to produce a family of related proteins with various combinations of amino

acid sequences derived from the exons. Each protein in the family, then, varies in its function. Alternative splicing helps us understand why humans with only about 20 000 genes can produce approximately 100 000 proteins. After the final mRNA has been produced, it is ready to leave the nucleus and be translated by a ribosome.

Transcription in Eukaryotes versus Prokaryotes

Transcription is similar in eukaryotes and prokaryotes, but not identical. For example, a single type of RNA polymerase is used in prokaryotes to transcribe all the genes, both those that encode proteins and those that code for non-protein molecules such as tRNA and rRNA. In contrast, eukaryotic cells use different types of RNA polymerase enzymes depending on what is being transcribed. RNA polymerase type II transcribes protein-coding genes, while RNA polymerase types I and III transcribe the non-protein–coding genes (that produce rRNA and tRNA). The key differences in transcription between eukaryotes and prokaryotes are summarized in **Table 1**.

Table 1 Comparison of Transcription in Eukaryotes and Prokaryotes

Variable	Prokaryotes	Eukaryotes
location	Transcription occurs throughout the cell.	Transcription takes place in the nucleus.
enzymes	A single type of RNA polymerase transcribes all types of genes.	Different RNA polymerases are used to transcribe genes that encode protein (RNA polymerase II) and genes that do not encode protein (RNA polymerase I, III).
elongation	Bases are added quickly (15 to 20 nucleotides per second).	Bases are added slowly (5 to 8 nucleotides per second).
promoters	The promoters are less complex than those in eukaryotes.	The promoters are immediately upstream of protein-coding genes, and they are more complex than those in prokaryotes.
termination	A protein binds to the mRNA and cleaves it, or the mRNA binds with itself.	Nuclear proteins bind to the polyuracil site and terminate transcription.
introns and exons	There are no introns.	There are both introns and exons.
product	Transcription results in mRNA ready to be translated into protein by ribosomes.	Transcription results in pre-mRNA, which must be modified to protect the final mRNA from degradation in the cytosol and to remove introns.

7.2 Review

Summary

- Transcription has three stages: initiation, elongation, and termination. One strand of the double-stranded DNA is used as a template for the synthesis of a complementary single-stranded RNA molecule.
- Initiation begins when RNA polymerase binds to a promoter region, which is upstream of the gene to be transcribed, and begins to unwind the DNA molecule.
- During elongation, a complementary RNA strand is synthesized in the 5′ to 3′ direction. Adenine in the DNA is paired with uracil in RNA.
- When RNA polymerase reaches a termination sequence, elongation ceases.
- In eukaryotes, post-transcriptional modifications include a 5′ cap of seven Gs added to the 5′ end, a string of adenines added to the 3′ end, and introns excised by spliceosomes. There are no modifications to prokaryotic mRNA.
- Transcription differs in eukaryotes and prokaryotes with respect to location, speed, post-transcriptional modifications, and types of RNA polymerase enzymes.

Questions

1. List and describe the three stages of transcription. K/U
2. If the DNA template strand has the sequence 3′-CAAATTGGCTTATTACCGGATG-5′, what would be the sequence of an RNA molecule transcribed from it? T/I
3. Explain the role of each of the following in transcription. K/U
   (a) promoter
   (b) RNA polymerase
   (c) spliceosomes
4. Differentiate between introns and exons. K/U
5. What are the key differences between transcription in eukaryotes and prokaryotes? K/U
6. Compare and contrast DNA replication and transcription. How are they similar? How are they different? Present your answer in table form. T/I C
7. How is it possible for an organism to produce more proteins than it has genes for? K/U
8. As a graduate student in a university laboratory, you have been challenged with the problem of determining whether a sample of mRNA is from a eukaryotic cell or a prokaryotic cell. You have been provided with a nucleotide sequencer, which will help you determine the DNA sequence. What features in the sequence will you look for to determine whether the mRNA is eukaryotic or prokaryotic? T/I
9. Suppose that you are provided with a sample of eukaryotic DNA. You divide the sample into three separate reaction mixtures and perform an experiment. Once transcription is complete, you analyze the base composition of mRNA from each mixture. You obtain the results in **Table 2**. Based on these results, answer the questions below. T/I

Table 2 Experimental Results

	A	G	C	T	U
DNA Strand I	19.1	26.0	31.0	23.9	0
DNA Strand II	24.2	30.8	25.7	19.3	0
mRNA Strand A	19.7	25.9	30.8	0	24.0
mRNA Strand B	24.1	30.9	25.9	0	19.0

   (a) Which strand of DNA served as the template for the synthesis of mRNA strand A? Which strand served as the template for the synthesis of mRNA strand B? Explain your reasoning.
   (b) Explain why the percentage of adenine is higher in the mRNA strands than in the DNA strands.
10. How does the absence of a nucleus in prokaryotes prevent prokaryotes from controlling gene expression by modifying RNA after transcription? T/I