

6.5 TRANSCRIPTION

3' / 5' Newly synthesized strands

Figure 6.8 Replication

The process of copying genetic information from one strand of the DNA into RNA is termed as **transcription**. Here also, the principle of complementarity governs the process of transcription, except the adenosine complements now forms base pair with uracil instead of thymine. However, unlike in the process of replication, which once set in, the total DNA of an organism gets duplicated, in transcription only a segment of DNA and only one of the strands is copied into RNA. This necessitates defining the boundaries that would demarcate the region and the strand of DNA that would be transcribed.

Why both the strands are not copied during transcription has the simple answer. First, if both strands act as a template, they would code for RNA molecule with different sequences (Remember complementarity does not mean identical), and in turn, if they code for proteins, the sequence of amino acids in the proteins would be different. Hence, one segment of the DNA would be coding for two different proteins, and this would complicate the genetic information transfer machinery. Second, the two RNA molecules if produced simultaneously would be complementary to each other, hence would form a double stranded RNA. This would prevent RNA from being translated into protein and the exercise of transcription would become a futile one.

6.5.1 Transcription Unit

A transcription unit in DNA is defined primarily by the three regions in the DNA:

- (i) A Promoter
- (ii) The Structural gene
- (iii) A Terminator

There is a convention in defining the two strands of the DNA in the structural gene of a transcription unit. Since the two strands have opposite polarity and the **DNA-dependent RNA polymerase** also catalyse the polymerisation in only one direction, that is, $5' \rightarrow 3'$, the strand that has the polarity $3' \rightarrow 5'$ acts as a template, and is also referred to as **template strand**. The other strand which has the polarity ($5' \rightarrow 3'$) and the sequence same as RNA (except thymine at the place of uracil), is displaced during transcription. Strangely, this strand (which does not code for anything) is referred to as **coding strand**. All the reference point while defining a transcription unit is made with coding strand. To explain the point, a hypothetical sequence from a transcription unit is represented below:

3'-ATGCATGCATGCATGCATGC-5' Template Strand

5'-TACGTACGTACGTACGTACGTACG-3' Coding Strand

Can you now write the sequence of RNA transcribed from the above DNA?

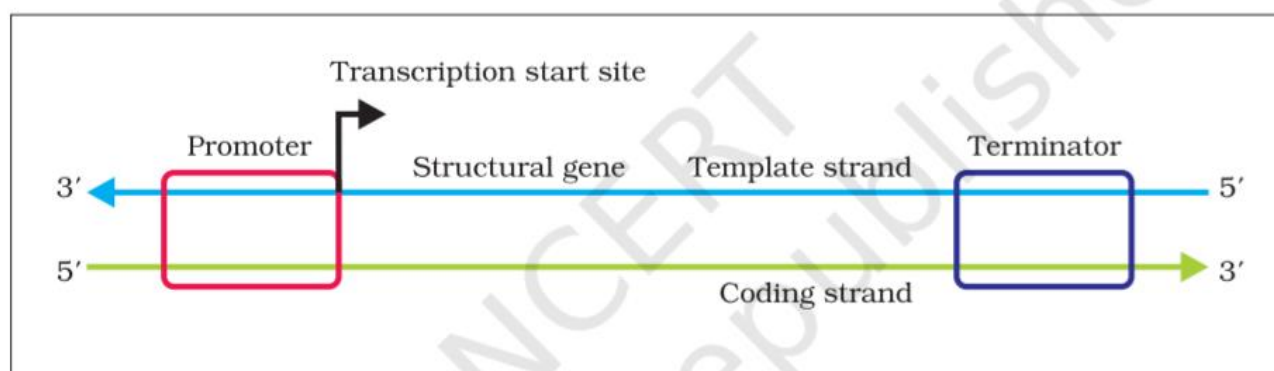


Figure 6.9 Schematic structure of a transcription unit

The **promoter** and **terminator** flank the **structural gene** in a transcription unit. The promoter is said to be located towards 5'-end (upstream) of the structural gene (the reference is made with respect to the polarity of coding strand). It is a DNA sequence that provides binding site for RNA polymerase, and it is the presence of a promoter in a transcription unit that also defines the template and coding strands. By switching its position with terminator, the definition of coding and template strands could be reversed. The terminator is located towards 3'-end (downstream) of the coding strand and it usually defines the end of the process of transcription (Figure 6.9). There are additional regulatory sequences that may be present further upstream or downstream to the promoter. Some of the properties of these sequences shall be discussed while dealing with regulation of gene expression.

6.5.2 Transcription Unit and the Gene

A gene is defined as the functional unit of inheritance. Though there is no ambiguity that the genes are located on the DNA, it is difficult to literally

define a gene in terms of DNA sequence. The DNA sequence coding for tRNA or rRNA molecule also define a gene. However by defining a **cistron** as a segment of DNA coding for a polypeptide, the structural gene in a transcription unit could be said as **monocistronic** (mostly in eukaryotes) or **polycistronic** (mostly in bacteria or prokaryotes). In eukaryotes, the monocistronic structural genes have interrupted coding sequences – the genes in eukaryotes are split. The coding sequences or expressed sequences are defined as **exons**. Exons are said to be those sequence that appear in mature or processed RNA. The exons are interrupted by **introns**. Introns or intervening sequences do not appear in mature or processed RNA. The split-gene arrangement further complicates the definition of a gene in terms of a DNA segment.

Inheritance of a character is also affected by promoter and regulatory sequences of a structural gene. Hence, sometime the regulatory sequences are loosely defined as regulatory genes, even though these sequences do not code for any RNA or protein.

6.5.3 Types of RNA and the process of Transcription

In bacteria, there are three major types of RNAs: mRNA (messenger RNA), tRNA (transfer RNA), and rRNA (ribosomal RNA). All three RNAs are needed to synthesise a protein in a cell. The mRNA provides the template, tRNA brings aminoacids and reads the genetic code, and rRNAs play structural and catalytic role during translation. There is single DNA-dependent RNA polymerase that catalyses transcription of all types of RNA in bacteria. RNA polymerase binds to promoter and initiates transcription (**Initiation**). It uses nucleoside triphosphates as substrate

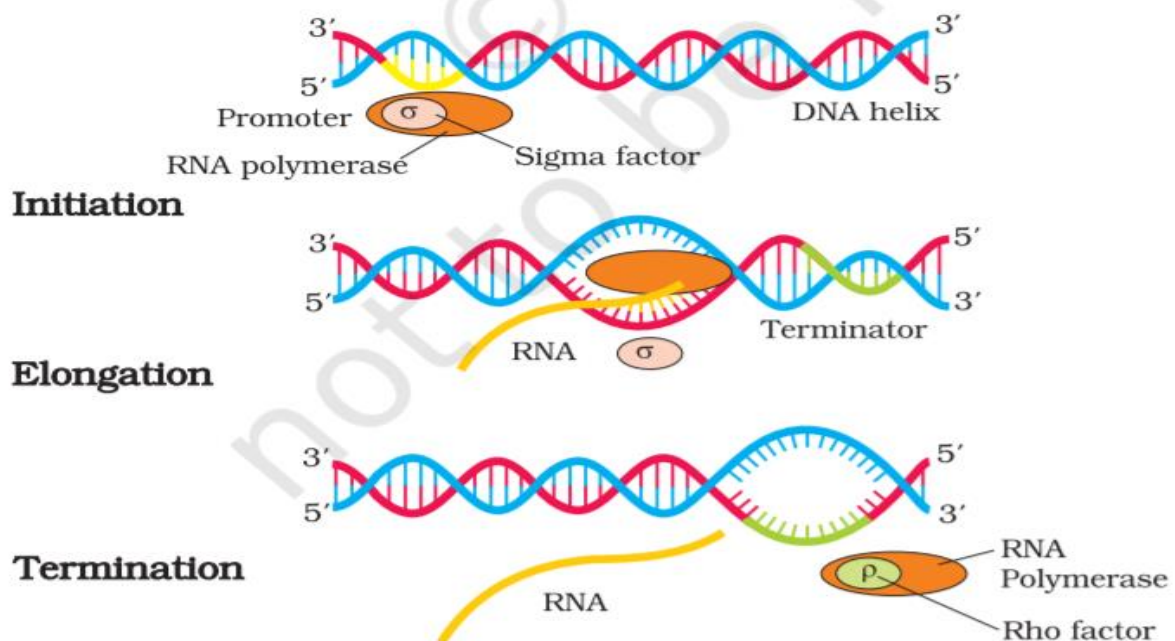


Figure 6.10 Process of Transcription in Bacteria

and polymerises in a template depended fashion following the rule of complementarity. It somehow also facilitates opening of the helix and continues elongation. Only a short stretch of RNA remains bound to the enzyme. Once the polymerases reaches the terminator region, the nascent RNA falls off, so also the RNA polymerase. This results in **termination** of transcription.

An intriguing question is that how is the RNA polymerases able to catalyse all the three steps, which are initiation, elongation and termination. The RNA polymerase is only capable of catalysing the process of elongation. It associates transiently with **initiation-factor** (σ) and **termination-factor** (ρ) to initiate and terminate the transcription, respectively. Association with these factors alter the specificity of the RNA polymerase to either initiate or terminate (Figure 6.10).

In bacteria, since the mRNA does not require any processing to become active, and also since transcription and translation take place in the same compartment (there is no separation of cytosol and nucleus in bacteria), many times the translation can begin much before the mRNA is fully transcribed. Consequently, the transcription and translation can be coupled in bacteria.

In eukaryotes, there are two additional complexities –

- (i) There are at least three RNA polymerases in the nucleus (in addition to the RNA polymerase found in the organelles). There is a clear cut division of labour. The RNA polymerase I transcribes **rRNAs**

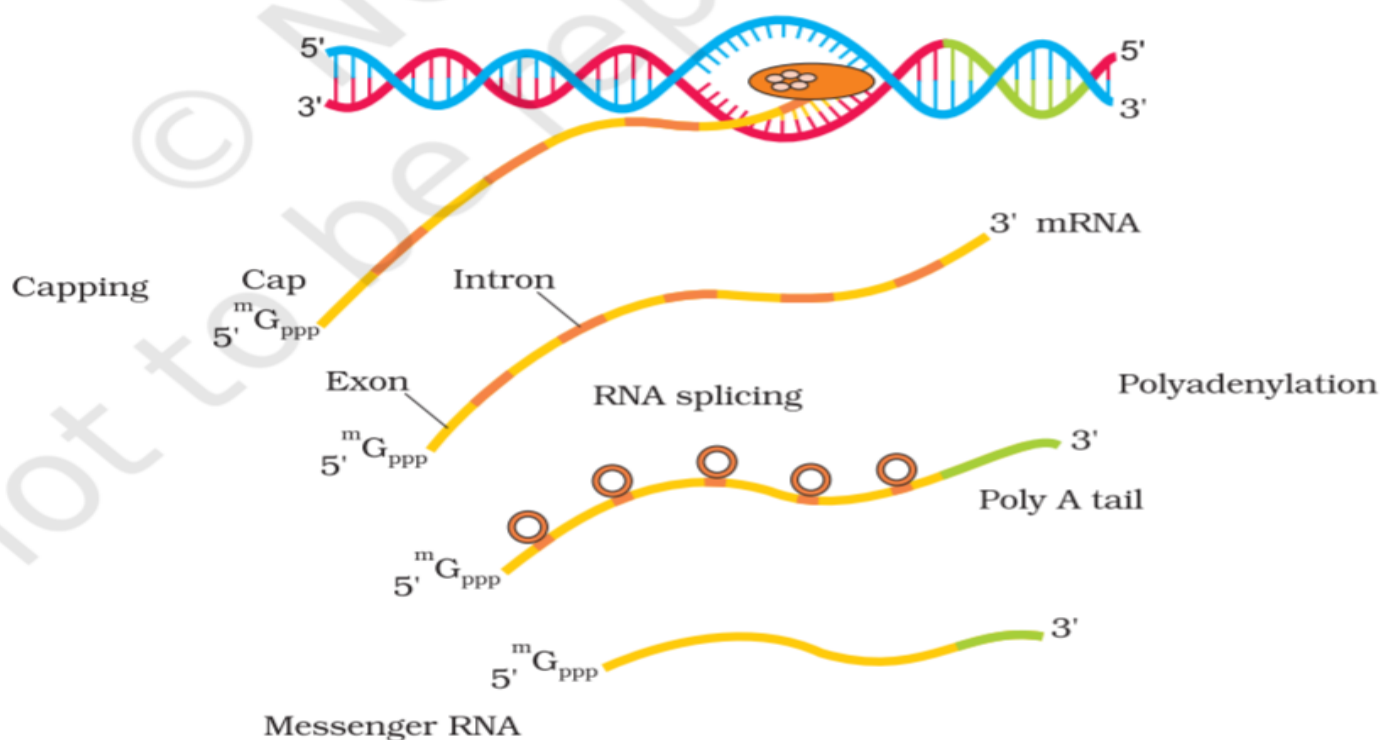


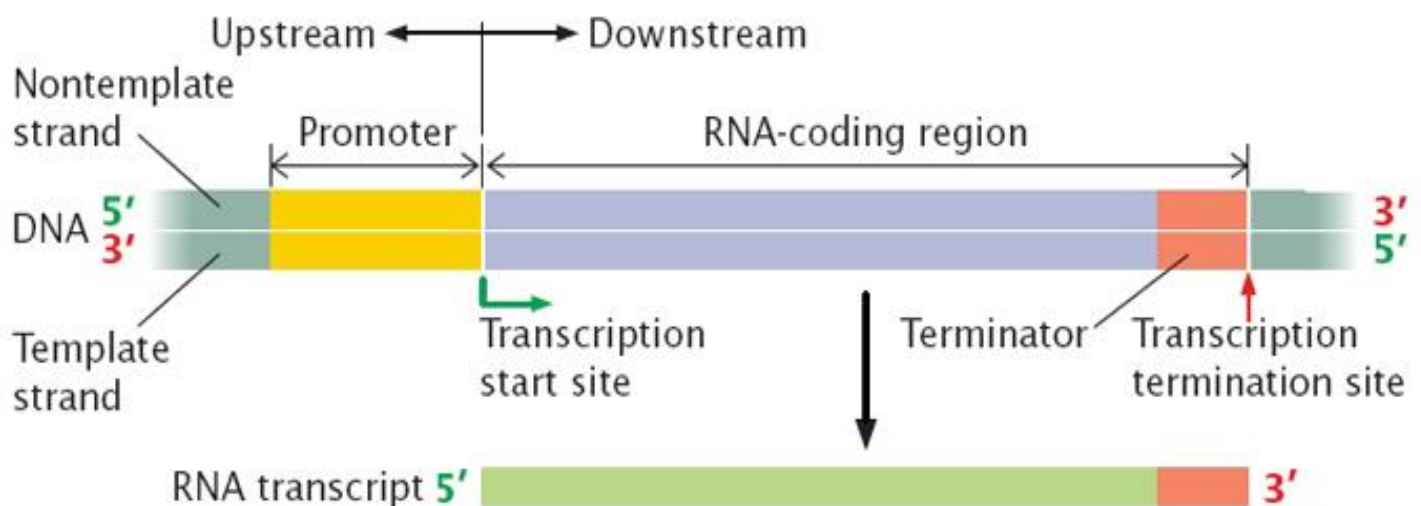
Figure 6.11 Process of Transcription in Eukaryotes

(28S, 18S, and 5.8S), whereas the RNA polymerase III is responsible for transcription of **tRNA**, **5srRNA**, and **snRNAs (small nuclear RNAs)**. The RNA polymerase II transcribes precursor of mRNA, the **heterogeneous nuclear RNA (hnRNA)**.

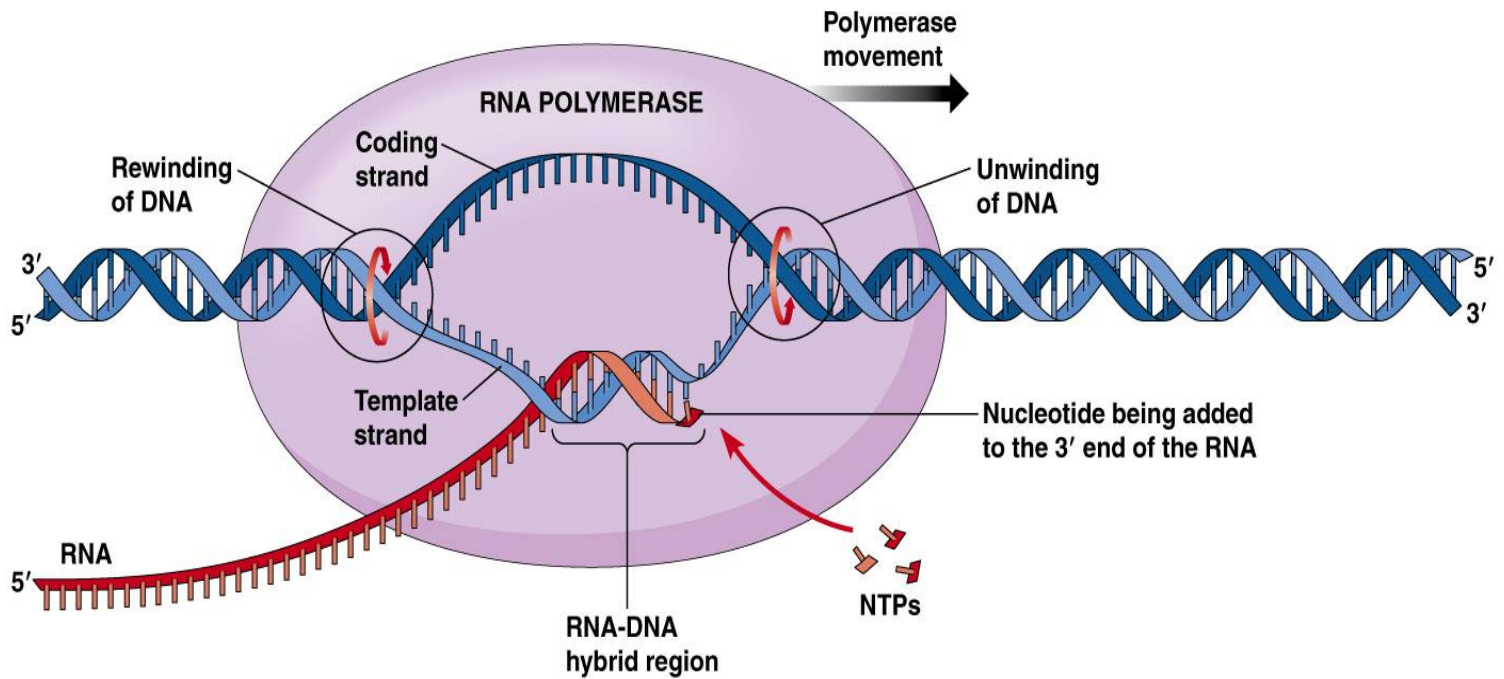
- (ii) The second complexity is that the primary transcripts contain both the exons and the introns and are non-functional. Hence, it is subjected to a process called **splicing** where the introns are removed and exons are joined in a defined order. hnRNA undergoes additional processing called as capping and tailing. In **capping** an unusual nucleotide (methyl guanosine triphosphate) is added to the 5'-end of hnRNA. In **tailing**, adenylate residues (200-300) are added at 3'-end in a template independent manner. It is the fully processed hnRNA, now called mRNA, that is transported out of the nucleus for translation (Figure 6.11).

The significance of such complexities is now beginning to be understood. The split-gene arrangements represent probably an ancient feature of the genome. The presence of introns is reminiscent of antiquity, and the process of splicing represents the dominance of **RNA-world**. In recent times, the understanding of RNA and RNA-dependent processes in the living system have assumed more importance.

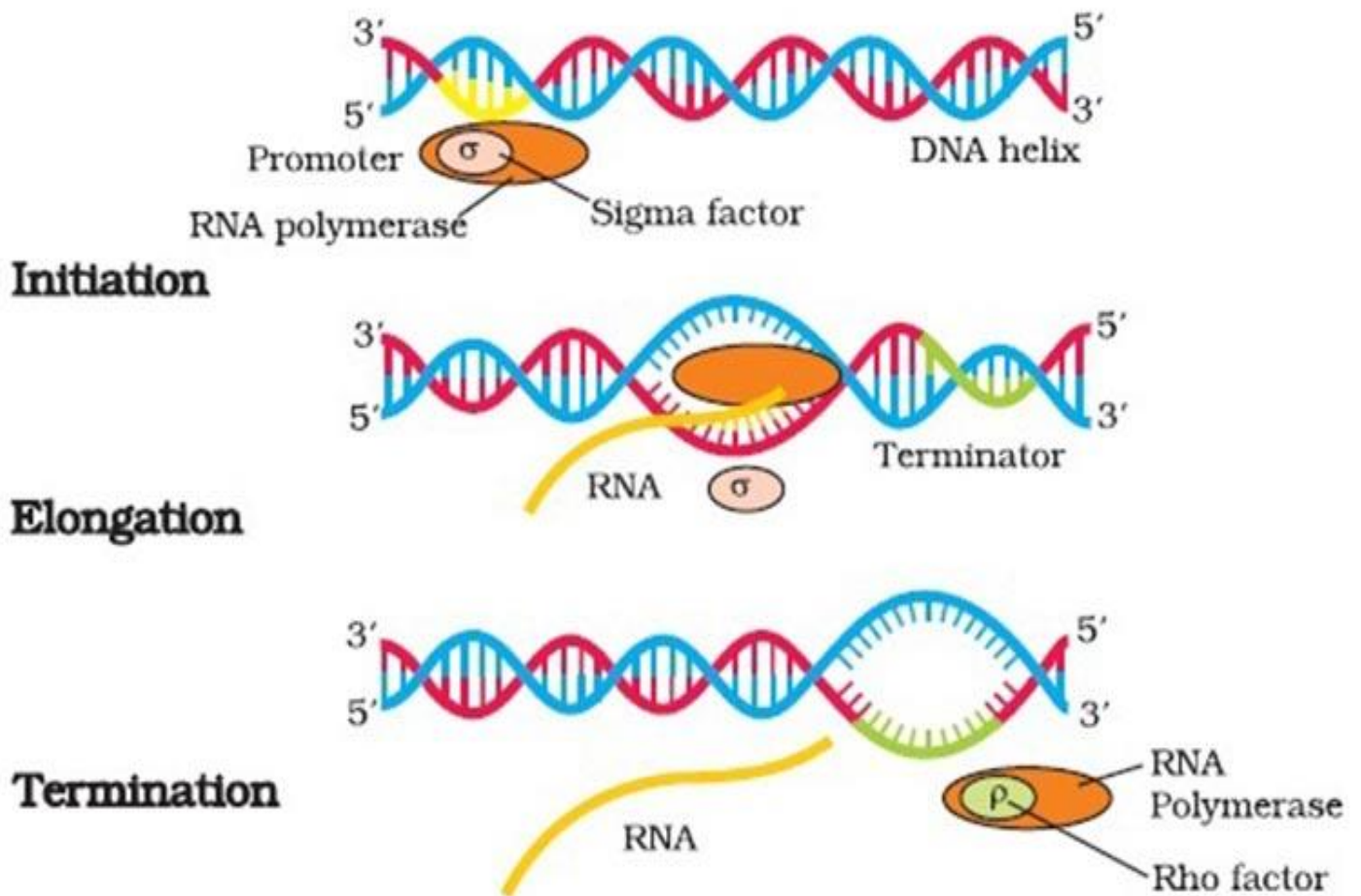
TRANSCRIPTION UNIT



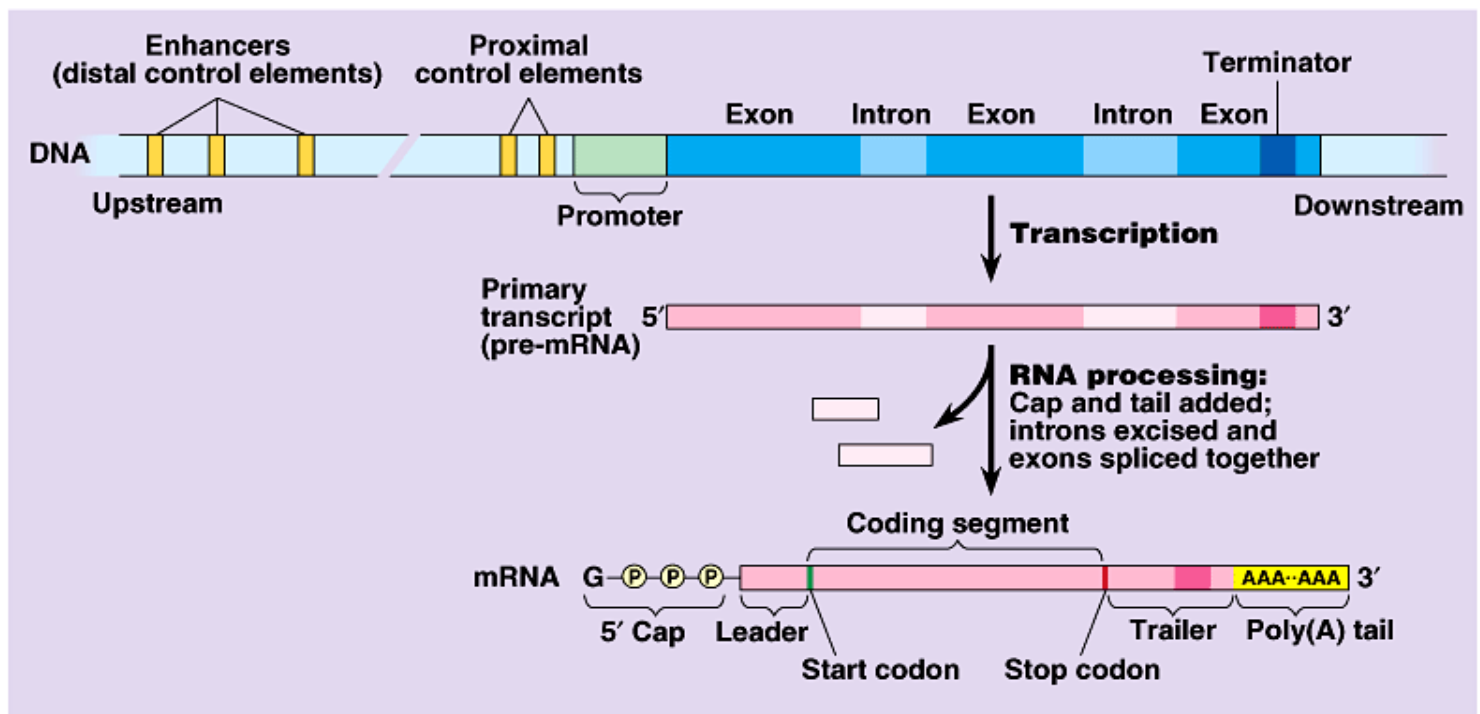
TRANSCRIPTION IN PROKARYOTES



© 2012 Pearson Education, Inc.

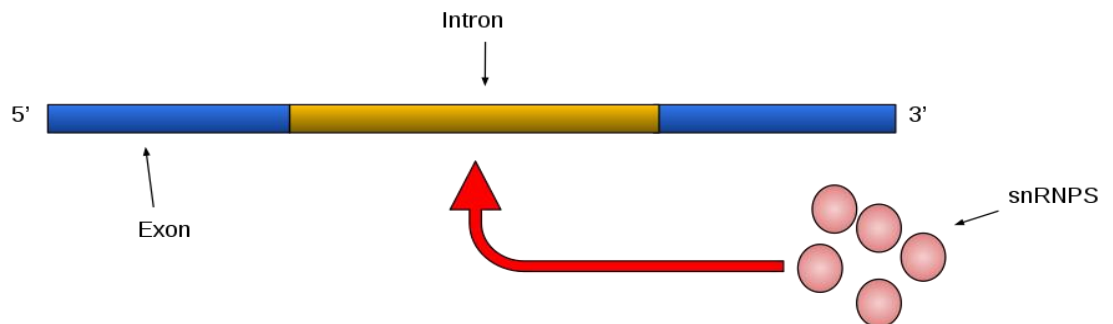


TRANSCRIPTION IN EUKARYOTES

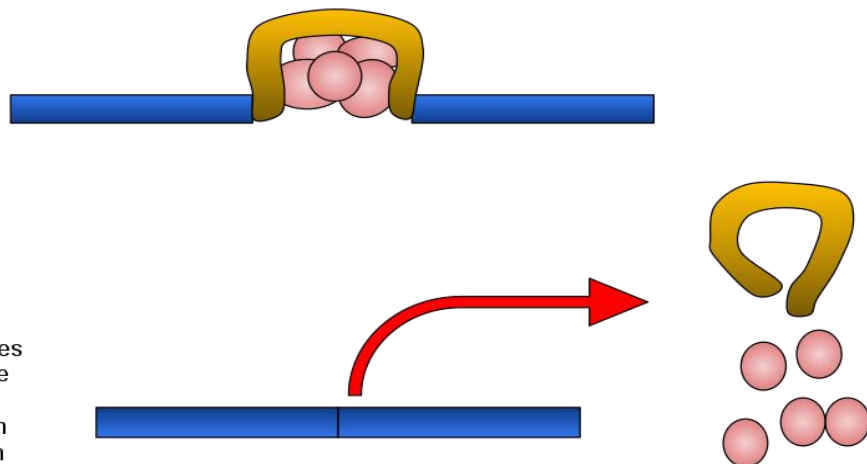


Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.

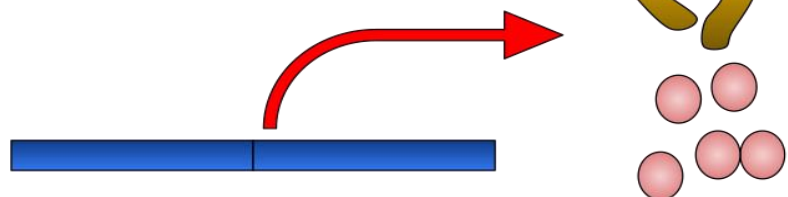
Step 1.
A group of five snRNPS's, or ribonucleoproteins, are needed to bind to the intron of pre-mRNA and remove it to leave only the exons.



Step 2.
The snRNP's bind to the intron and cause it to fold into bring the 5' and 3' ends of the intron closer together, making a loop. The ends of the exons also move closer together to eventually join together.



Step 3.
The intron detaches and the splice sites connect to make a mature mRNA. The introns were previously thought to be "junk" afterwards but most are used in other processes. The snRNP's detach from the intron and are used for more splicing.



ASSIGNMENT 6 (MOLECULAR BASIS OF INHERITANCE):

1. NAME THREE REGIONS PRESENT IN A TRANSCRIPTION UNIT ON DNA.
2. DRAW A LABELED DIAGRAM OF A TRANSCRIPTION UNIT ON DNA.
3. A HYPOTHETICAL SEQUENCE OF A TEMPLATE STRAND OF A TRANSCRIPTION UNIT IS GIVEN BELOW. WRITE THE SEQUENCE OF CODING STRAND OF THIS UNIT.
3' – AGGTTTCGCGCGTAAGGTCA - 5'
4. DESCRIBE THE PROCESS OF TRANSCRIPTION IN PROKARYOTIC CELL WITH A WELL LABELED DIAGRAM.
5. WHAT DO YOU MEAN BY MONOCISTRONIC AND POLYCISTRONIC TRANSCRIPTION UNIT? WHERE ARE THEY FOUND?
6. STATE THE FUNCTION OF m RNA, rRNA , t RNA IN PROTIEN SYNTHESIS.
7. HOW MANY TYPES OF RNA POLYMERASE ARE FOUND IN PRAKARYOTES AND EUKARYOTES. DRAW A COMPARISON BETWEEN THEM IN RESPECT TO THEIR FUNCTION.
8. DESCRIBE THE PROCESS OF TRANSCRIPTION IN EUKARYOTES.
9. COMPARE THE PROCESS OF TRANSCRIPTION IN PROKARYOTES AND EUKARYOTES.

