

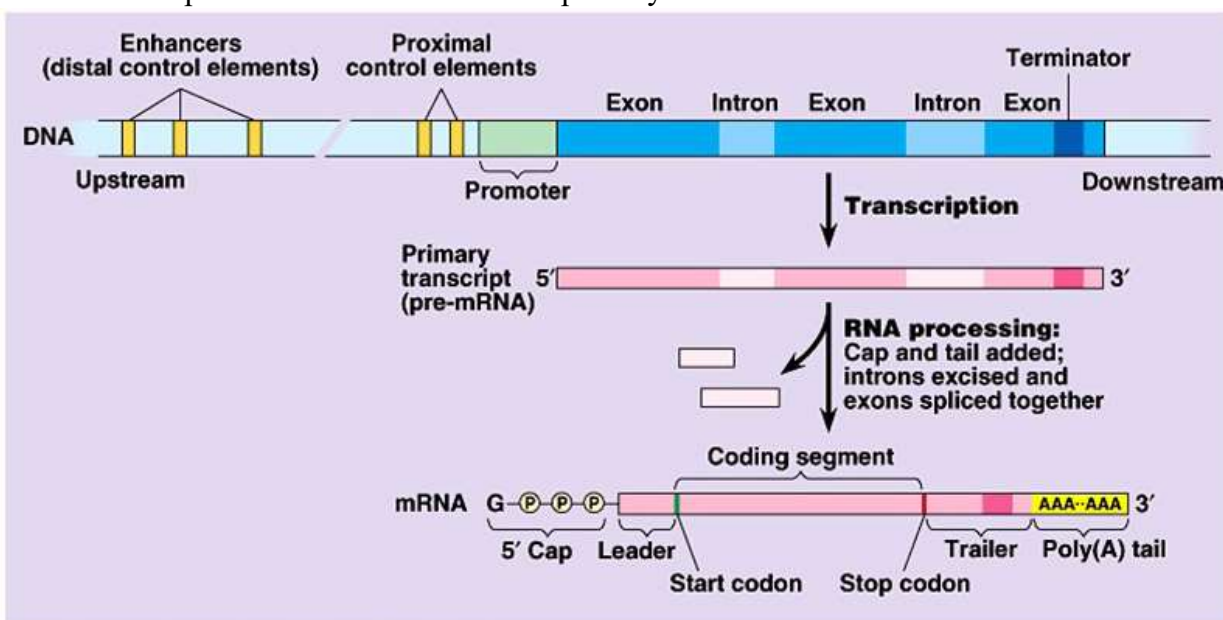
Eukaryotic Transcription

- Transcription is the process by which the information in a strand of DNA is copied into a new molecule of **RNA**.
- It is the first step of gene expression, in which a particular segment of DNA is copied into RNA (especially mRNA) by the enzyme RNA polymerase.
- It results in a complementary, antiparallel RNA strand called a primary transcript.

Transcription in Eukaryotes

Transcription occurs in eukaryotes in a way that is similar to prokaryotes with reference to the basic steps involved. However, some major differences between them include:

- Initiation is more complex.
- Termination does not involve stem-loop structures.
- Transcription is carried out by three enzymes (RNA polymerases I, II and III).
- The regulation of transcription is more extensive than prokaryotes.



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

Enzyme(s) Involved in Eukaryotic Transcription

Unlike prokaryotes where all RNA is synthesized by a single RNA polymerase, the nucleus of a eukaryotic cell has three RNA polymerases responsible for transcribing different types of RNA.

- **RNA polymerase I (RNA Pol I)** is located in the nucleolus and transcribes the 28S, 18S, and 5.8S rRNA genes.
- **RNA polymerase II (RNA Pol II)** is located in the nucleoplasm and transcribes protein-coding genes, to yield pre-mRNA, and also the genes encoding small nucleolar RNAs (snoRNAs) involved in rRNA processing and small nuclear RNAs (snRNAs) involved in mRNA processing, except for U6 snRNA.
- **RNA polymerase III (RNA Pol III)** is also located in the nucleoplasm. It transcribes the genes for tRNA, 5S rRNA, U6 snRNA, and the 7S RNA associated with the signal recognition particle (SRP) involved in the translocation of proteins across the endoplasmic reticulum membrane.
- Each of the three eukaryotic RNA polymerases contains 12 or more subunits and so these are large complex enzymes.
- The genes encoding some of the subunits of each eukaryotic enzyme show DNA sequence similarities to genes encoding subunits of the core enzyme of *E. coli* RNA polymerase.
- However, four to seven other subunits of each eukaryotic RNA polymerase are unique in that they show no similarity either with bacterial RNA polymerase subunits or with the subunits of other eukaryotic RNA polymerases.

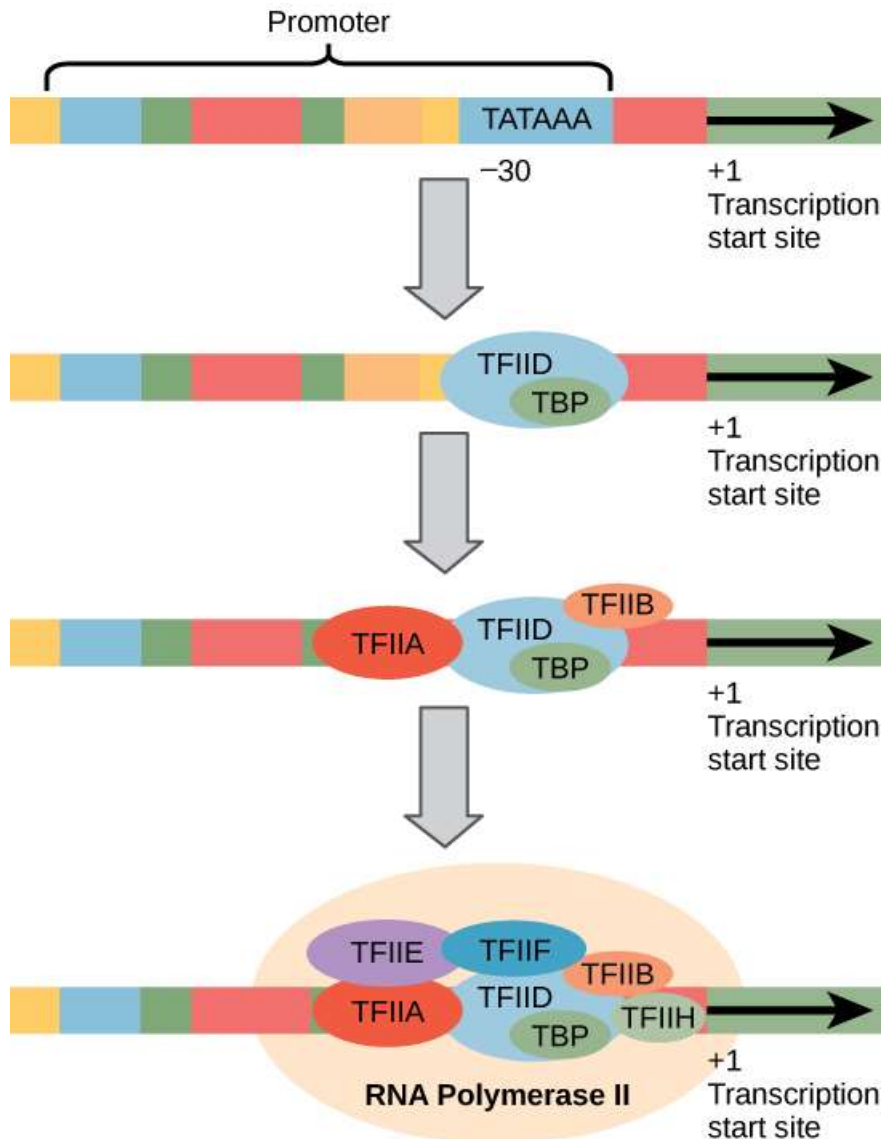
Features of Eukaryotic Transcription

- Transcription in eukaryotes occurs within the nucleus and mRNA moves out of the nucleus into the cytoplasm for translation.
- The initiation of RNA synthesis by RNA polymerase is directed by the presence of a promoter site on the 5' side of the transcriptional start site.
- The RNA polymerase transcribes one strand, the antisense (-) strand, of the DNA template.
- RNA synthesis does not require a primer.
- RNA synthesis occurs in the 5' → 3' direction with the RNA polymerase catalyzing a nucleophilic attack by the 3-OH of the growing RNA chain on the alpha-phosphorus atom on an incoming ribonucleoside 5-triphosphate.
- mRNA in eukaryotes is processed from the primary RNA transcript, a process called maturation.

Process of Eukaryotic Transcription

The basic mechanism of RNA synthesis by these eukaryotic RNA polymerases can be divided into the following phases:

Initiation Phase



- During initiation, RNA polymerase recognizes a specific site on the DNA, upstream from the gene that will be

transcribed, called a **promoter site** and then unwinds the DNA locally.

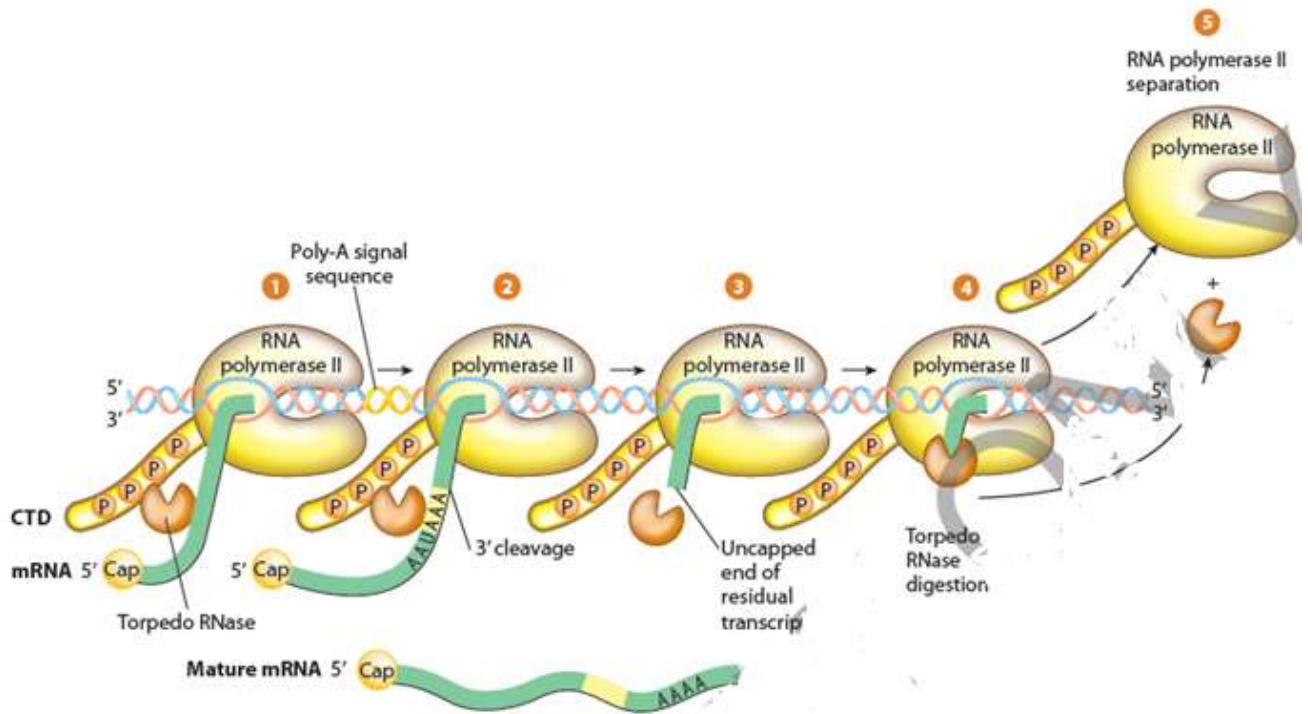
- Most promoter sites for RNA polymerase II include a highly conserved sequence located about 25–35 bp upstream (i.e. to the 5' side) of the start site which has the consensus TATA(A/T)A(A/T) and is called the TATA box.
- Since the start site is denoted as position +1, the TATA box position is said to be located at about position -25.
- The TATA box sequence resembles the -10 sequence in prokaryotes (TATAAT) except that it is located further upstream.
- Both elements have essentially the same function, namely recognition by the RNA polymerase in order to position the enzyme at the correct location to initiate transcription.
- The sequence around the TATA box is also important in that it influences the efficiency of initiation. Transcription is also regulated by upstream control elements that lie 5' to the TATA box.
- Some eukaryotic protein-coding genes lack a TATA box and have an initiator element instead, centered around the transcriptional initiation site.
- In order to initiate transcription, RNA polymerase II requires the assistance of several other proteins or protein complexes, called general (or basal) transcription factors, which must assemble into a complex on the promoter in order for RNA polymerase to bind and start transcription.
- These all have the generic name of TFII (for Transcription Factor for RNA polymerase II).
- The first event in initiation is the binding of the transcription factor IID (TFIID) protein complex to the TATA box via one of its subunits called TBP (TATA box binding protein).
- As soon as the TFIID complex has bound, TFIIA binds and stabilizes the TFIID-TATA box interaction. Next, TFIIB binds to TFIID.
- However, TFIIB can also bind to RNA polymerase II and so acts as a bridging protein. Thus,
- RNA polymerase II, which has already complexed with TFIIF, now binds.
- This is followed by the binding of TFIIE and H. This final protein complex contains at least 40 polypeptides and is called the **transcription initiation complex**.
- Those protein-coding genes that have an initiator element instead of a TATA box appear to need another protein(s) that binds to the initiator element.
- The other transcription factors then bind to form the transcription initiation complex in a similar manner to that described above for genes possessing a TATA box promoter.

Elongation Phase

TFIIH has two functions:

1. It is a helicase, which means that it can use ATP to unwind the DNA helix, allowing transcription to begin.
2. In addition, it phosphorylates RNA polymerase II which causes this enzyme to change its conformation and dissociate from other proteins in the initiation complex.
 - The key phosphorylation occurs on a long C-terminal tail called the C-terminal domain (CTD) of the RNA polymerase II molecule.
 - Interestingly, only RNA polymerase II that has a non-phosphorylated CTD can initiate transcription but only an RNA polymerase II with a phosphorylated CTD can elongate RNA.
 - RNA polymerase II now starts moving along the DNA template, synthesizing RNA, that is, the process enters the elongation phase.
 - RNA synthesis occurs in the 5' → 3' direction with the RNA polymerase catalyzing a nucleophilic attack by the 3'-OH of the growing RNA chain on the alpha-phosphorus atom on an incoming ribonucleoside 5-triphosphate.
 - The RNA molecule made from a protein-coding gene by RNA polymerase II is called a primary transcript.

Termination Phase



- Elongation of the RNA chain continues until termination occurs.
- Unlike RNA polymerase in prokaryotes, RNA polymerase II does not terminate transcription at a specific site but rather transcription can stop at varying distances downstream of the gene.
- RNA genes transcribed by RNA Polymerase II lack any specific signals or sequences that direct RNA Polymerase II to terminate at specific locations.
- RNA Polymerase II can continue to transcribe RNA anywhere from a few bp to thousands of bp past the actual end of the gene.
- The transcript is cleaved at an internal site before RNA Polymerase II finishes transcribing. This releases the upstream portion of the transcript, which will serve as the initial RNA prior to further processing (the pre-mRNA in the case of protein-encoding genes.)
- This cleavage site is considered the “end” of the gene. The remainder of the transcript is digested by a 5'-exonuclease (called Xrn2 in humans) while it is still being transcribed by the RNA Polymerase II.
- When the 5'-exonuclease “catches up” to RNA Polymerase II by digesting away all the overhanging RNA, it helps disengage the polymerase from its DNA template strand, finally terminating that round of transcription.

RNA processing

The primary eukaryotic mRNA transcript is much longer and localised into the nucleus, when it is also called heterogenous nuclear RNA (hnRNA) or pre- mRNA.

It undergoes various processing steps to change into a mature RNA:

Cleavage

- Larger RNA precursors are cleaved to form smaller RNAs.
- Primary transcript is cleaved by ribonuclease-P (an RNA enzyme) to form 5-7 tRNA precursors.

Capping and Tailing

- Initially at the 5' end a cap (consisting of 7-methyl guanosine or 7 mG) and a tail of poly A at the 3' end are added.
- The cap is a chemically modified molecule of guanosine triphosphate (GTP).

Splicing

- The eukaryotic primary mRNAs are made up of two types of segments; non-coding introns and the coding exons.
- The introns are removed by a process called RNA splicing where ATP is used to cut the RNA, releasing the introns

and joining two adjacent exons to produce mature mRNA.

Nucleotide Modifications

- They are most common in tRNA-methylation (e.g., methyl cytosine, methyl guanosine), deamination (e.g., inosine from adenine), dihydrouracil, pseudouracil, etc.

Post-transcription processing is required to convert primary transcript into functional RNAs.

Significance

- Transcription of DNA is the method for regulating gene expression.
- It occurs in preparation for and is necessary for protein translation.