Transcription in Eukaryotes

By

Dr. Walaa Bayoumie El Gazzar

❖Remember:

- **RNA polymerase** (the enzyme that catalyzes RNA synthesis) does not need a primer; rather, it can initiate transcription de novo.

- The RNA product **does not remain base-paired to the template DNA strand**. This displacement is critical for the RNA to perform its functions (e.g., as is most often the case, to be translated to produce its protein product). **Furthermore**, because of this release, multiple RNA polymerase molecules can transcribe the same gene at the same time. Thus, a cell can synthesize large numbers of transcripts from a single gene in a short time.

- Transcription, although very accurate, is **less accurate than replication** (one mistake occurs in 10,000 nucleotides added, compared with one in 10 million for replication). This difference reflects the lack of extensive proofreading mechanisms for transcription, although proofreading for RNA synthesis do exist.

- It makes sense for the cell to worry more about the accuracy of replication than of transcription. DNA is the molecule in which the genetic material is stored, and DNA replication is the process by which that genetic material is passed on. Any mistake that arises during replication can therefore easily be catastrophic: it becomes permanent in the genome of that individual and gets passed on to subsequent generations. **Transcription, in contrast, produces only transient copies and normally several from each transcribed region. Thus, a mistake during transcription will rarely do more harm than render one out of many transient transcripts defective.**

- **The choice of which regions to transcribe is not random:** there are specific DNA sequences that direct the initiation of transcription at the start of each region and others at the end that terminate transcription.

- In different cells, or in the same cell at different times, different sets of genes might be transcribed. Therefore, for example, two genetically identical cells in a human will, in many cases, transcribe different sets of genes, leading to differences in the character and function of those two cells (e.g., one might be a muscle cell and the other a neuron).

- **Bacteria have only one RNA polymerase, all eukaryotes have three different ones (Pol I, II, and III)**. **In addition,** whereas bacteria require only one additional initiation factor (**σ**), several initiation factors are required for efficient and promoter-specific initiation in eukaryotes. **These are called the general transcription factors (GTFs).**

- Once transcribed, **eukaryotic** RNA has to be processed in various ways before being exported from the nucleus where it can be translated.
- These processing events include capping of the 5' end of the RNA, splicing, and polyadenylation of the 3' end of the RNA. The most complicated of these is splicing—the process whereby non-coding introns are removed from RNA to generate the mature mRNA.

Synthesis of RNA in eukaryotes

- Nuclear RNA polymerases of eukaryotic cells:

-There are three types of nuclear polymerases:

1- **RNA pol I:** Transcribes 18S ,5.8 S ,28S ribosomal RNA genes

2- **RNA pol II:** It transcribes mRNA ,and most small nuclear RNAs (snRNA)

3- **RNA pol III:** It transcribes , tRNA and 5S rRNA and some small nuclear RNAs (snRNA)

- **Mitochondrial RNA pol:** Resembles bacterial RNA pol than eukaryotic enzyme. Responsible for mitochondrial gene expression **as well as** for providing RNA primer for initiation of replication of the mitochondrial genome.

❖Transcription phases:

- Similar to prokaryotes, eukaryotic RNA synthesis include three main phases:

1- **Initiation** : involves the binding of RNA polymerase to a region on the DNA which is specific and is known as the promoter region.

2- **Elongation** : after the promoter region is recognized by the RNA polymerase, it starts to synthesize a complementary transcript to the template DNA strand. The RNA polymerase utilizes ribonucleotide triphosphate (ATP, GTP, CTP, UTP) and releases pyrophosphate each time a nucleotide is added to the growing chain.

3- **Termination** : elongation of the RNA chain continues until a termination signal is reached.

Synthesis of mRNA

❖RNA Polymerase II Core Promoters Are Made Up of Combinations of Different Classes of Sequence Element:

- The eukaryotic core promoter refers to the minimal set of sequence elements required for accurate transcription initiation by the Pol II machinery.
- A core promoter is typically about 40–60 nucleotides long, extending either upstream or downstream from the transcription start site.

- The Figure shows the location, relative to the transcription start site, of elements found in Pol II core promoters. These are the TFIIB recognition element (BRE), the TATA element (or box), the initiator (Inr), and the downstream promoter elements (known as DPE, DCE, and MTE).
- Typically, a promoter includes some subset of these elements. Thus, for example, promoters typically have either a TATA element or a DPE element, not both. Often, a TATA-containing promoter also contains a DCE.
- The Inr is the most common element, found in combination with both TATA and DPEs.
- The core promoter serves as a binding platform for the transcription machinery, which comprises Pol II and its associated **general transcription factors** (GTFs)

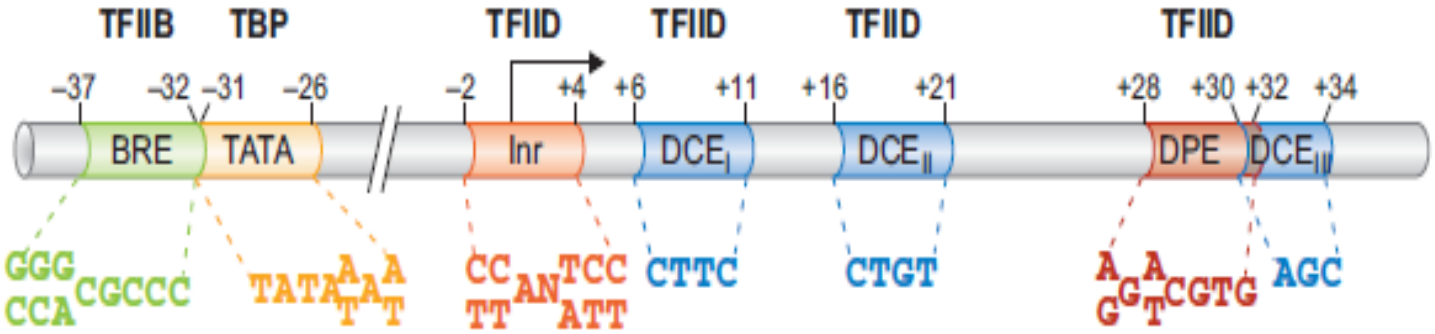


FIGURE 13-15 Pol II core promoter. The figure shows the positions of various DNA elements relative to the transcription start site (indicated by the arrow above the DNA). These elements, described in the text, are as follows: (BRE) TFIIB recognition element; (TATA) TATA box; (Inr) initiator element; (DPE) downstream promoter element; and (DCE) downstream core element. Another element, MTE (motif ten element), described in the text, is not shown in this figure but is located just upstream of the DPE. Also shown are the consensus sequences for each element (determined in the same way as described for the bacterial promoter elements; see Box 13-1) and (above) the name of the general transcription factor that recognizes each element.

❖ RNA Polymerase II Forms a Preinitiation Complex with General Transcription Factors at the Promoter:

- The general transcription factors help polymerase bind to the promoter and melt the DNA. They also help polymerase escape from the promoter and embark on the elongation phase. The complete set of general transcription factors and polymerase, bound together at the promoter and poised for initiation, is called the **preinitiation complex.**
- As we described above, many Pol II promoters contain a so-called TATA element (some 30 bp upstream of the transcription start site). This is where preinitiation complex formation begins. **The TATA element is recognized by the general transcription factor called TFIID.** (The nomenclature "TFII" denotes a transcription factor for Pol II, with individual factors distinguished as A, B, and so on.).

- Like many of the general transcription factors, TFIID is, in fact, a multi-subunit complex. The component of TFIID that binds to the TATA DNA sequence is called TBP (TATA-binding protein).
- The other subunits in this complex are called **TAFs**, for TBP-associated factors. Some TAFs recognize other core promoter elements such as the Inr, DPE, and DCE, although the strongest binding is between TBP and TATA. Thus, **TFIID is a critical factor in promoter recognition and preinitiation complex establishment.**

- The resulting TBP–DNA complex provides a platform to recruit other general transcription factors and polymerase itself to the promoter. These proteins assemble at the promoter in the following order: TFIIA, TFIIB, TFIIF together with polymerase, and then TFIIE and TFIIH.
- Formation of the preinitiation complex containing these components is followed by promoter melting. **In contrast to the situation in bacteria, promoter melting in eukaryotes requires hydrolysis of ATP and is mediated by TFIIH.**

❖ Promoter Escape Requires Phosphorylation of the Polymerase "Tail"

- In eukaryotes, promoter escape involves two steps not seen in bacteria: **one** is ATP hydrolysis (in addition to the earlier ATP hydrolysis needed for DNA melting), and **the other** is phosphorylation of the polymerase.

- The large subunit of Pol II has a carboxy-terminal domain (CTD), which is referred to as the "tail". The CTD contains a series of repeats of the heptapeptide sequence: Tyr-Ser-Pro-Thr-Ser-Pro-Ser. There are 52 of these repeats in humans. Each repeat contains sites for phosphorylation by **specific kinases, including one that is a subunit of TFIIH.**

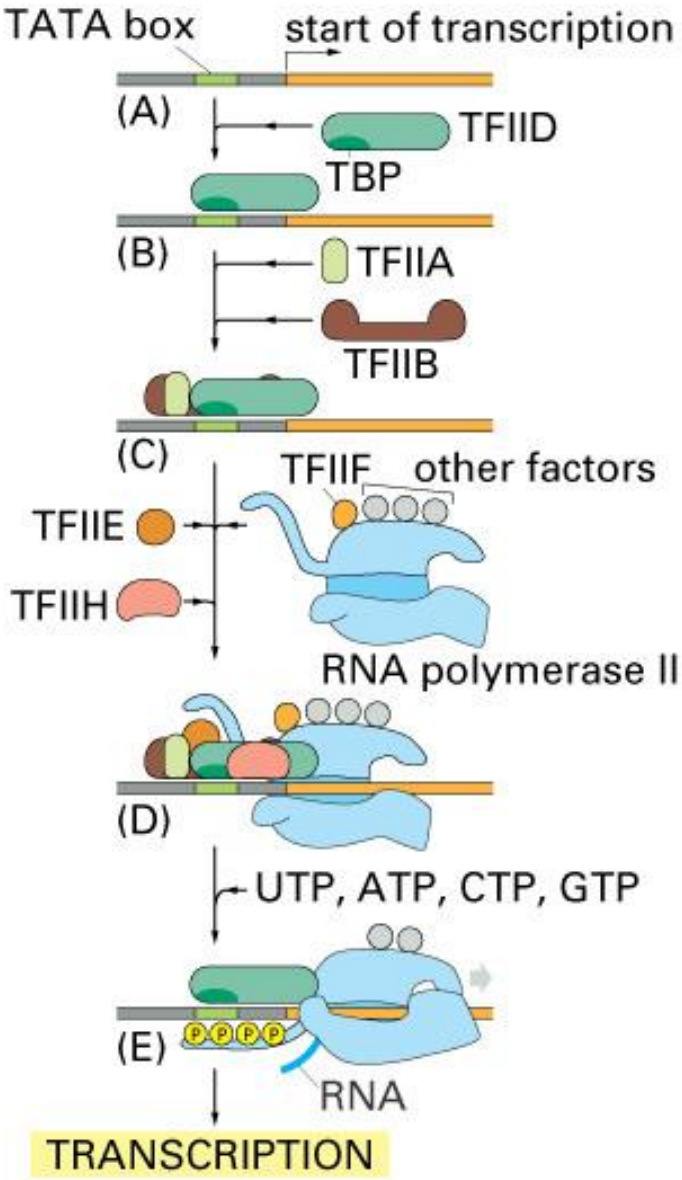


Figure 8-10 Essential Cell Biology, 2/e. (© 2004 Garland Science)

1-The first step is binding of the **TFII D** (contains TATA binding protein, TBP)to the TATA box.

2- Binding of **TFII A&B**, followed by binding of **RNA polymerase II-TFII F complex** (TFII F brings the RNAP II to the promoter site).

3- Binding of **TFII E&H** to form **preinitiation complex (PIC).**

4- Phosphorylation by a kinase produces activation of the polymerase II.

✓ *TFII F brings the RNAP II to the promoter site, while TFIIH activates it by phosphorylation.*

- For pol II-transcribed genes, and unlike bacterial RNA polymerase, **promoter melting requires hydrolysis of ATP and is mediated by TFIIH.**

- ***TFIIH is a ten-subunit protein, including both ATPase and protein kinase activities.***

5- Release of TFII A, B, E,&H

6- Pol II-TF IIF complex leaves the promoter, and starts transcription.

7- Transcription proceeds till the termination signal is reached.

8- Pol II-TF IIF complex is dissociated.

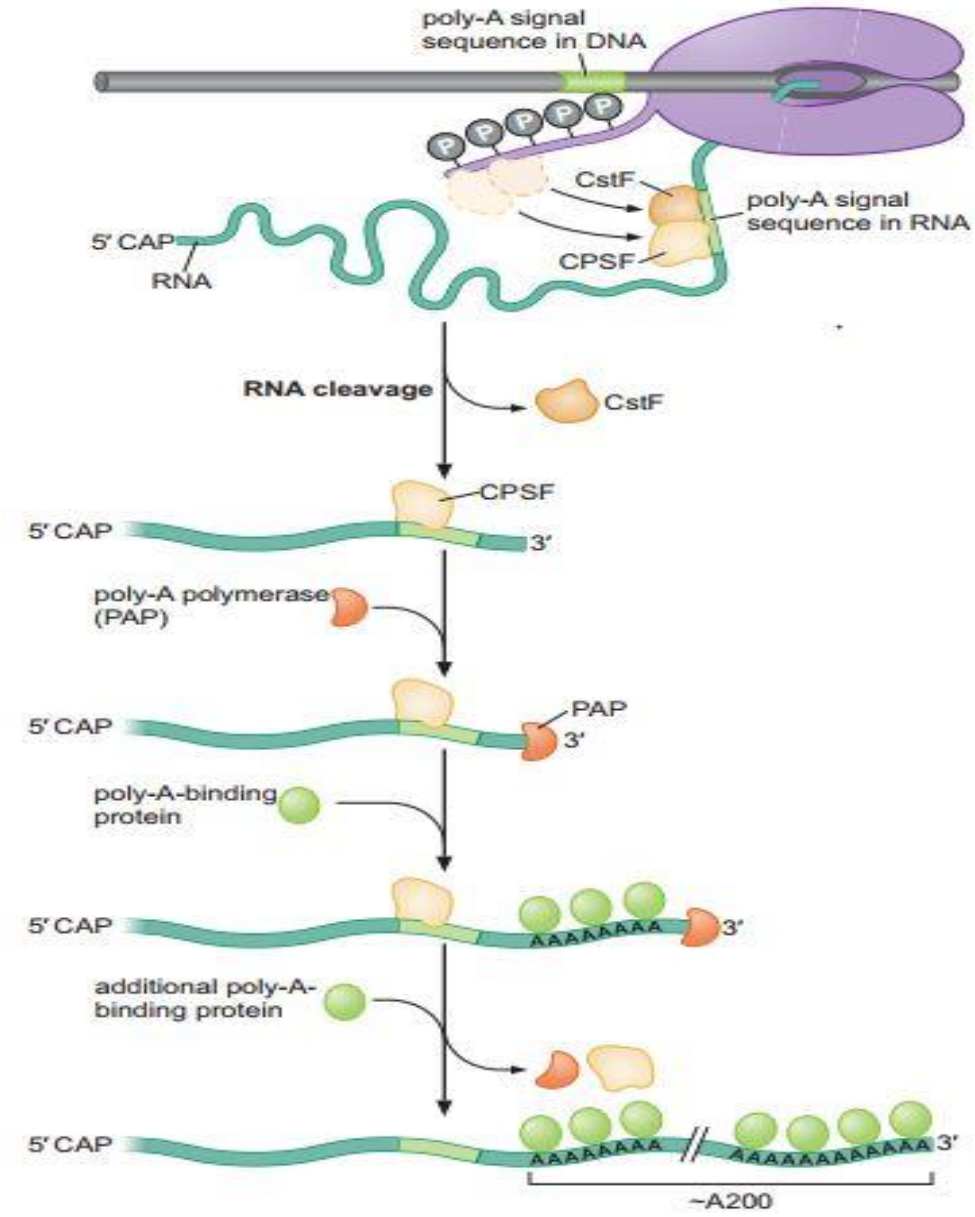
9- Pol II-TF IIF complex is dephosphorylated by a phosphatase.

10- A new cycle of transcription may start again.

Termination in eukaryotes:

- Leads to the dissociation of the complete transcript and the release of RNA polymerase from the template DNA. The process differs for each of the three RNA polymerases.
- As Pol II reaches the end of a gene, two protein complexes carried by the CTD (carboxy terminal domain), CPSF (cleavage and polyadenylation specificity factor) and CSTF (cleavage stimulation factor), recognize the poly-A signal (polyadenylation signal sequence AAUAAA) in the transcribed RNA.
- The sequences that, once transcribed into RNA, trigger transfer of these factors to the RNA are called poly-A signals

- Poly-A-bound CPSF and CSTF recruit other proteins to carry out RNA cleavage and then polyadenylation. **Poly-A polymerase** adds approximately 200 adenines to the cleaved 3' end of the RNA without a template. The long poly-A tail is unique to transcripts made by Pol II.
- The RNA molecule made by RNA pol II is called a **primary transcript**, which needs extensive RNA processing in order to produce a mature mRNA for translation & protein synthesis.


poly-A signal sequence in DNA
P
CstF
poly-A signal sequence in RNA
5′ CAP
RNA
CPSF
RNA cleavage
CstF
CPSF
5′ CAP
3′
poly-A polymerase (PAP)
5′ CAP
PAP
3′
poly-A-binding protein
5′ CAP
AAAAAAAA
3′
additional poly-A-binding protein
5′ CAP
AAAAAAAA
AAAAAAAAAAAA
3′
~A200