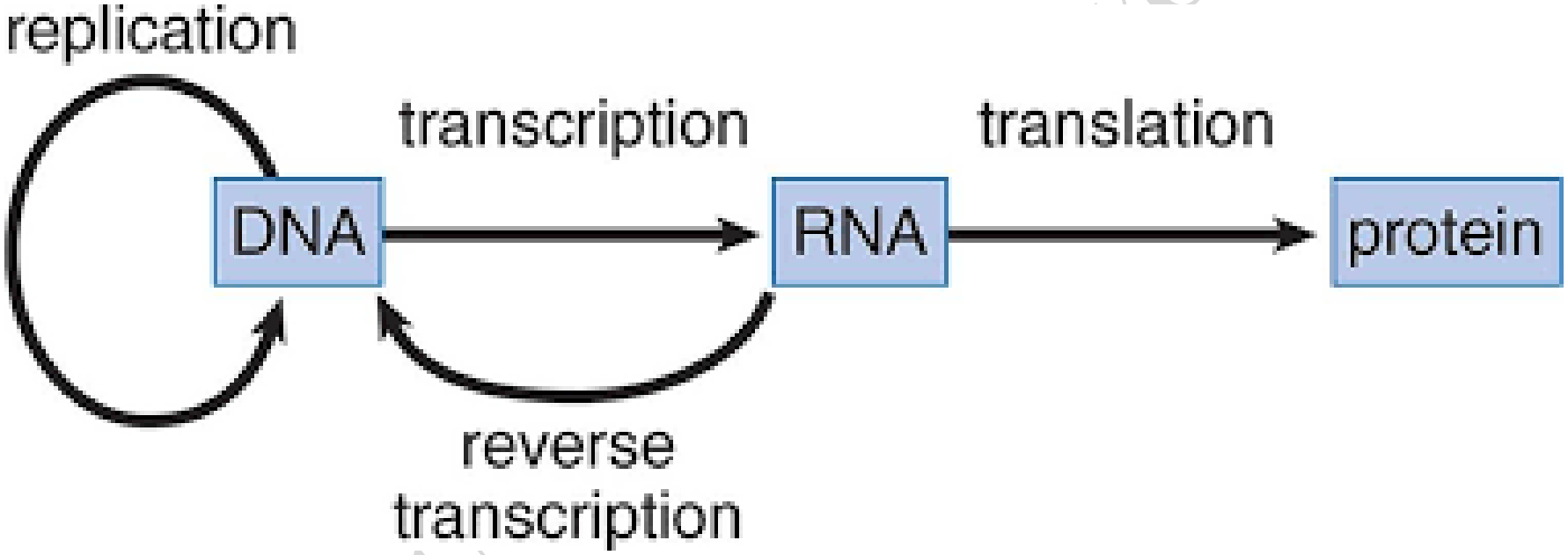
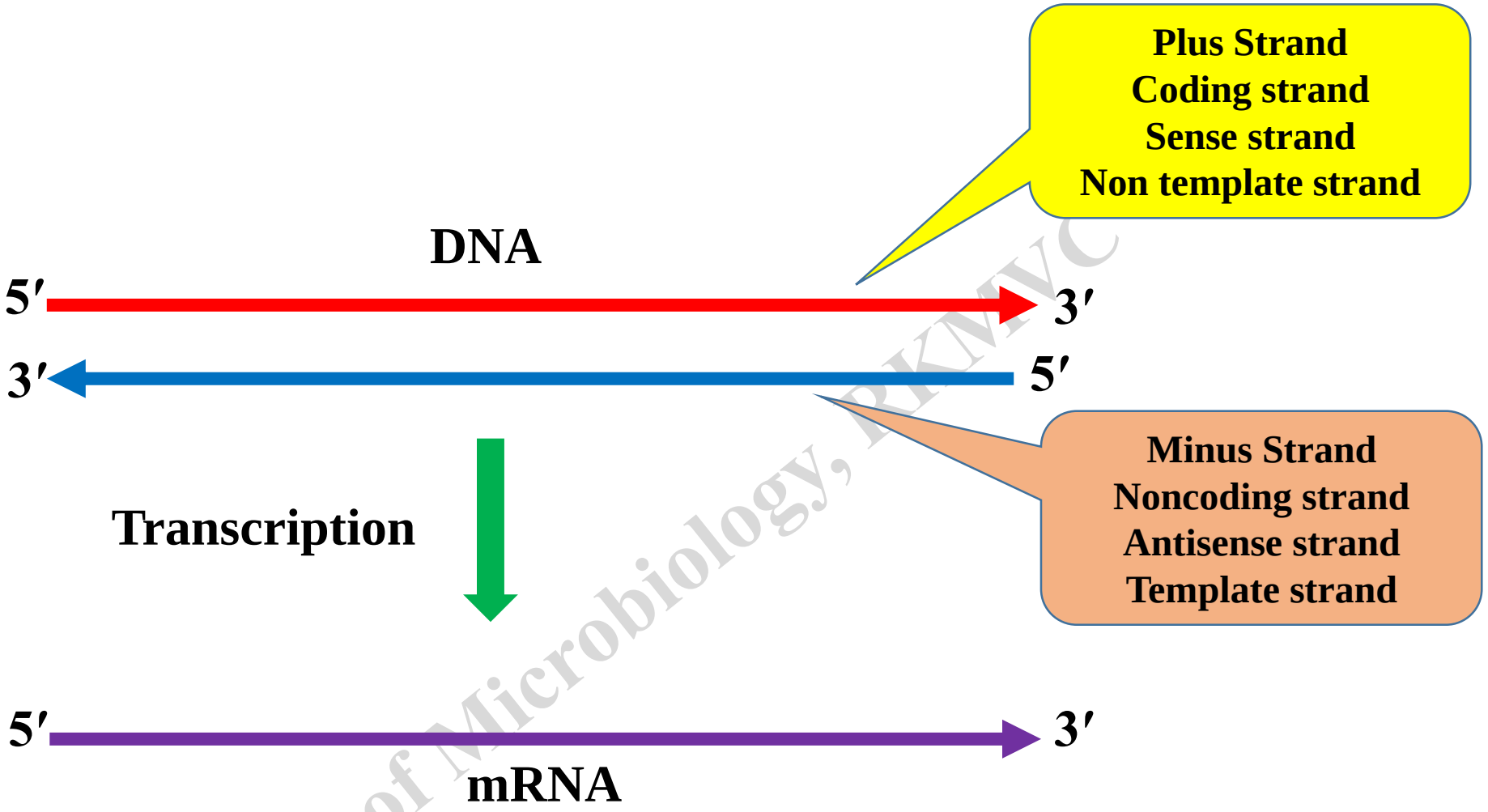
SEMESTER – III (THEORY)

C-7: MOLECULAR BIOLOGY

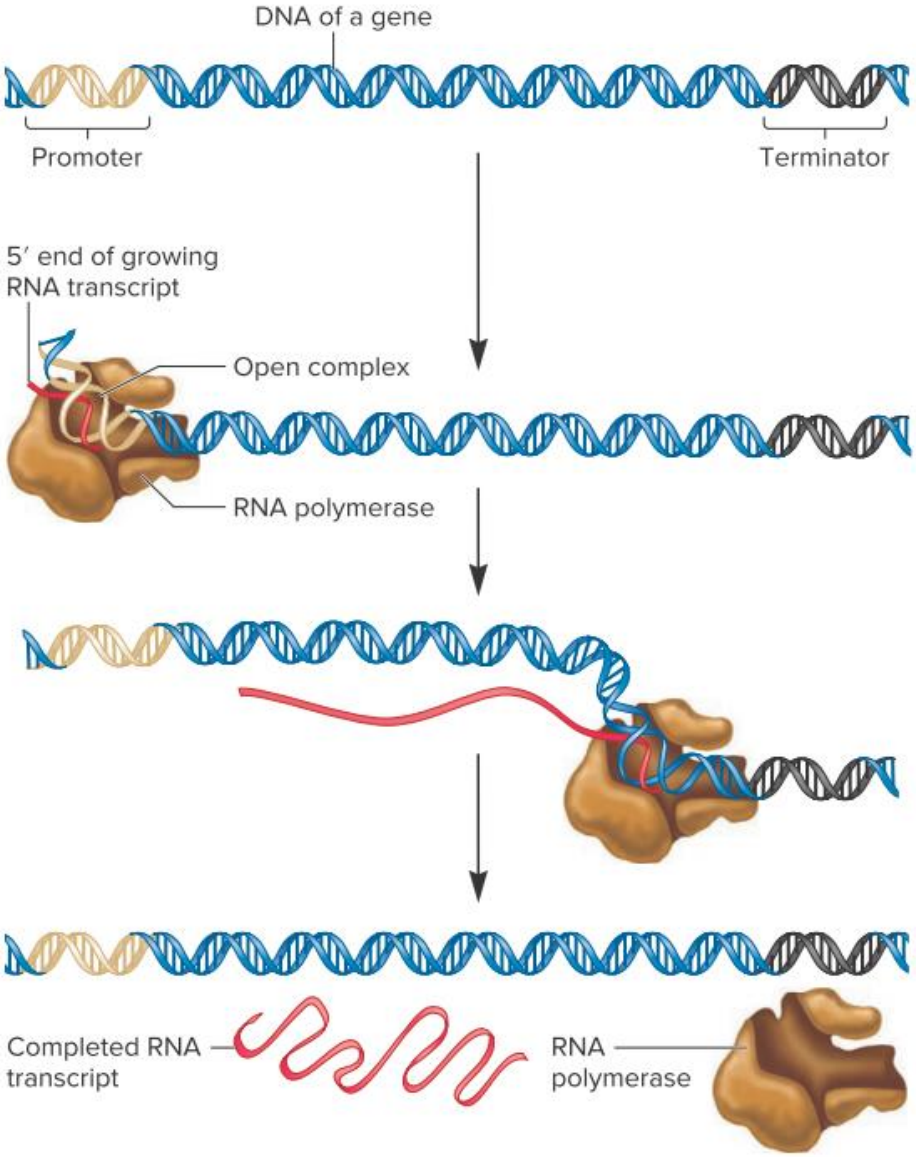
Unit 3 Transcription in Prokaryotes and Eukaryotes

Central dogma of molecular biology





Stages of transcription

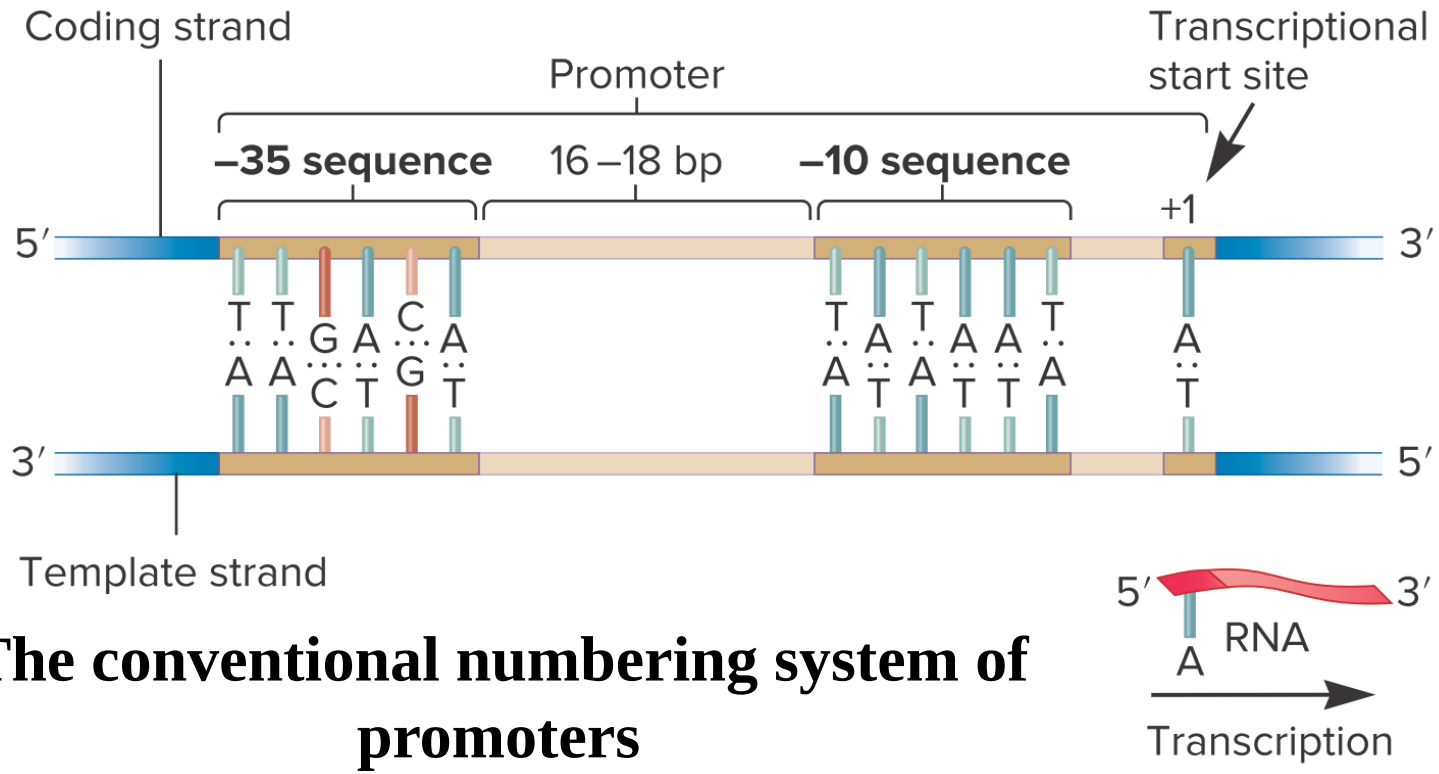


Initiation: The promoter functions as a recognition site for transcription factors (not shown). The transcription factors enable RNA polymerase to bind to the promoter. Following binding, the DNA is denatured into a bubble known as the open complex.

Elongation/synthesis of the RNA transcript: RNA polymerase slides along the DNA in an open complex to synthesize RNA.

Termination: A terminator is reached that causes RNA polymerase and the RNA transcript to dissociate from the DNA.

A Promoter Is a Short Sequence of DNA That Is Necessary to Initiate Transcription



The conventional numbering system of promoters

RNA Polymerase

- The enzyme that catalyzes the synthesis of RNA is RNA polymerase. In *E. coli*, the core enzyme is composed of five subunits, **$\alpha_2\beta\beta'\omega$**. The association of a sixth subunit, **sigma (σ) factor**, with the core enzyme creates what is referred to as **RNA polymerase holoenzyme**.

- The two α subunits are important in the **proper assembly of the holoenzyme** and in the process of binding to DNA. The β and β′ subunits are also needed for binding to the DNA, and they carry out the **catalytic synthesis of RNA**. The ω (omega) subunit is important for the proper assembly of the core enzyme. The **primary role of σ factor is to recognize the promoter**.

- When the holoenzyme encounters a promoter, σ factor recognizes both the **−35 and −10 sequences**. The σ-factor protein contains a structure called a **helix-turn-helix motif** that can bind tightly to these sequences. **Two α helices of the protein fit within the major groove of the DNA**. Amino acids within the α helices form hydrogen bonds with the bases in the DNA.

Transcription Initiation

- The process of transcription is initiated when σ factor within the holoenzyme has bound to the promoter to form a **closed complex**.

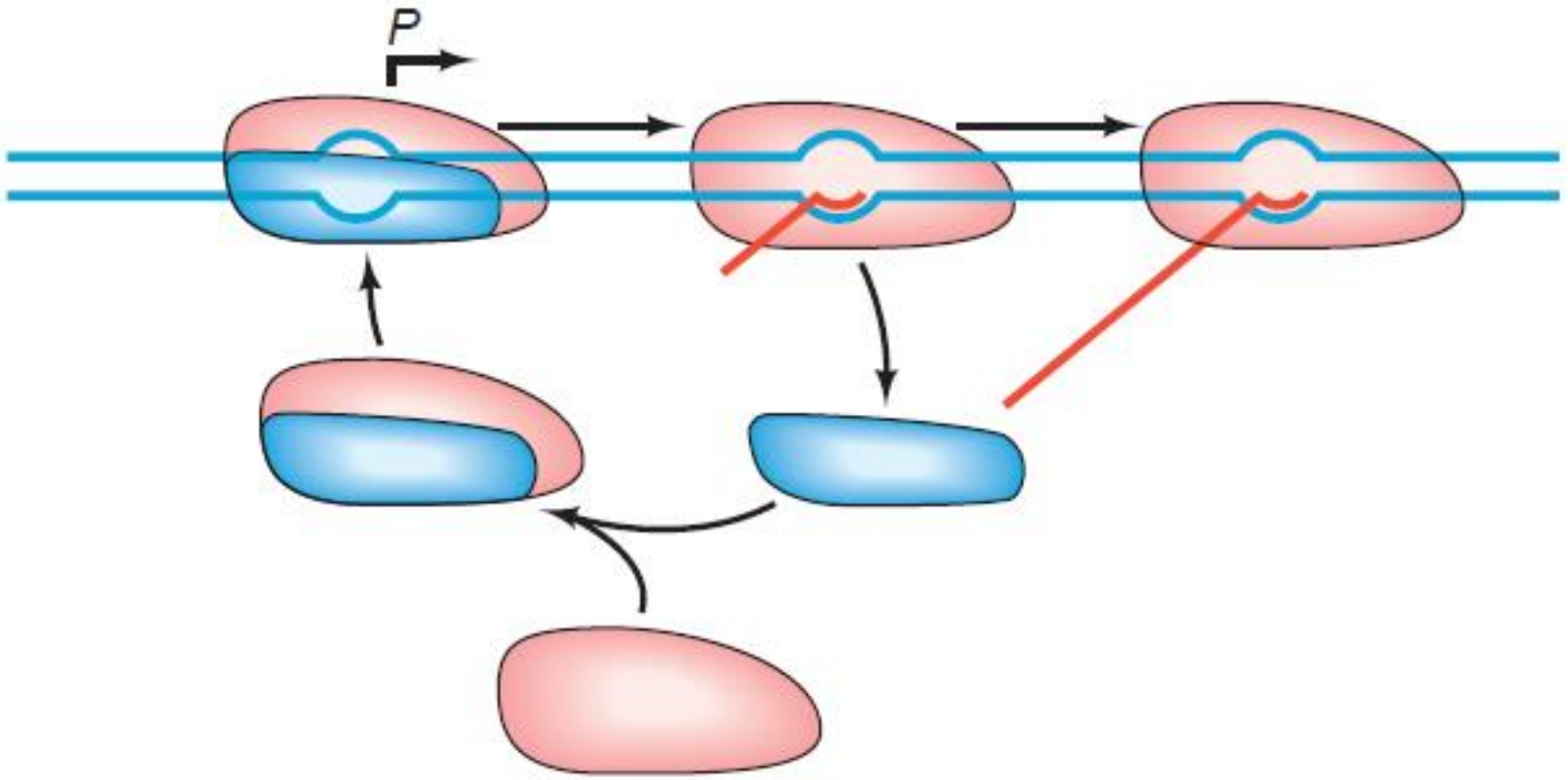
- For transcription to begin, the double-stranded DNA must then be unwound into an **open complex**. This unwinding first occurs at the **TATAAT sequence in the −10 site**, which contains only AT base pairs.

- A short strand of RNA (**abortive transcript**) is made within the open complex, and **then σ factor is released** from the core enzyme. The release of σ factor marks the transition to the elongation phase of transcription. The core enzyme may now slide down the DNA to synthesize a strand of RNA.

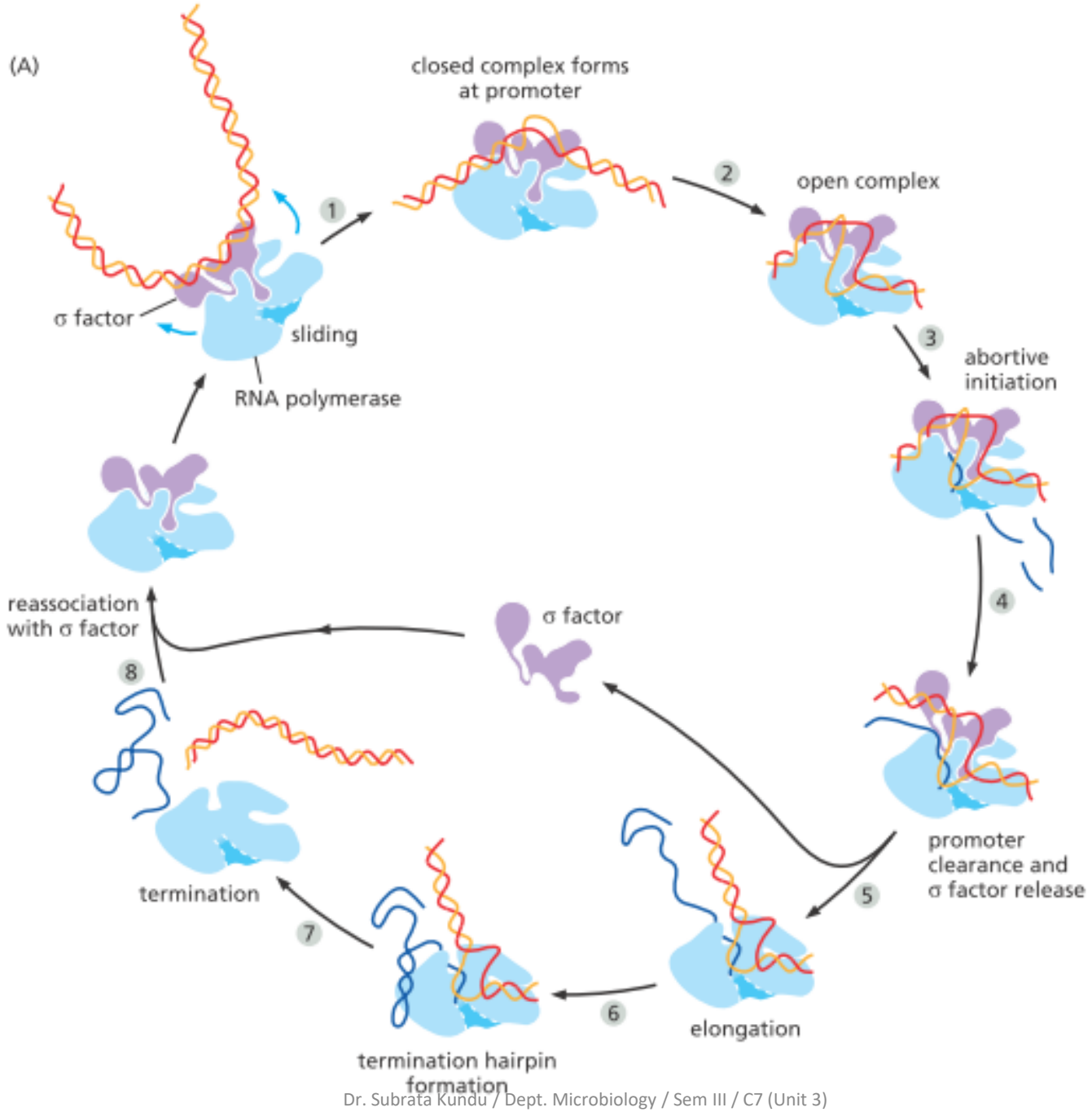
- RNAP enters into **abortive cycles of synthesis and release of short RNA products** as an RNAP-promoter **initial transcribing complex** (RP_{itc}) and, upon synthesis of an RNA product **~9-11 nt in length**, escapes the promoter and enters into productive synthesis of RNA as an **RNAP-DNA elongation complex (RD_e)**.

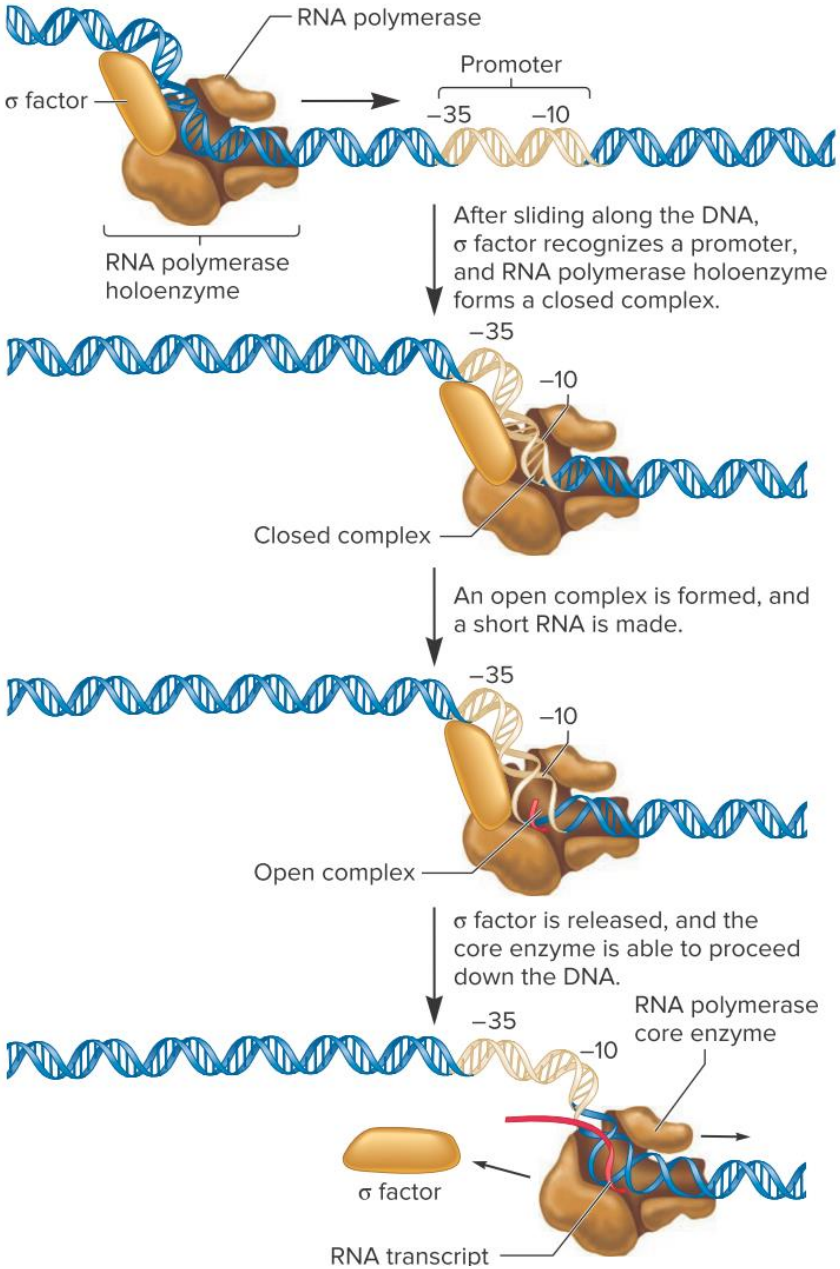
Transcription Initiation

- RNA polymerase (RNAP) binds to promoter DNA to yield an RNAP-promoter closed complex (RP_c).
- RNAP then unwinds ~1 turn of DNA surrounding the transcription start site to yield an RNAP-promoter open complex (RP_o).
- RNAP then enters into **abortive cycles of synthesis and release of short RNA products** as an RNAP-promoter initial transcribing complex (RP_{itc}) and, upon synthesis of an RNA product ~9-11 nt in length, escapes the promoter and enters into productive synthesis of RNA as an RNAP-DNA elongation complex (RD_e).
- According to '**DNA scrunching**' model, in each cycle of abortive initiation, RNAP pulls downstream DNA into itself, pulling in 1 bp per phosphodiester bond formed and accommodating the accumulated DNA as single-stranded bulges within the unwound region; upon release of the abortive RNA, RNAP extrudes the accumulated DNA, regenerating the initial state.



The σ cycle. RNA polymerase binds to the promoter at left, causing local melting of the DNA. As the polymerase moves to the right, elongating the RNA, the **σ**-factor dissociates and joins with a new core polymerase to initiate another RNA chain.

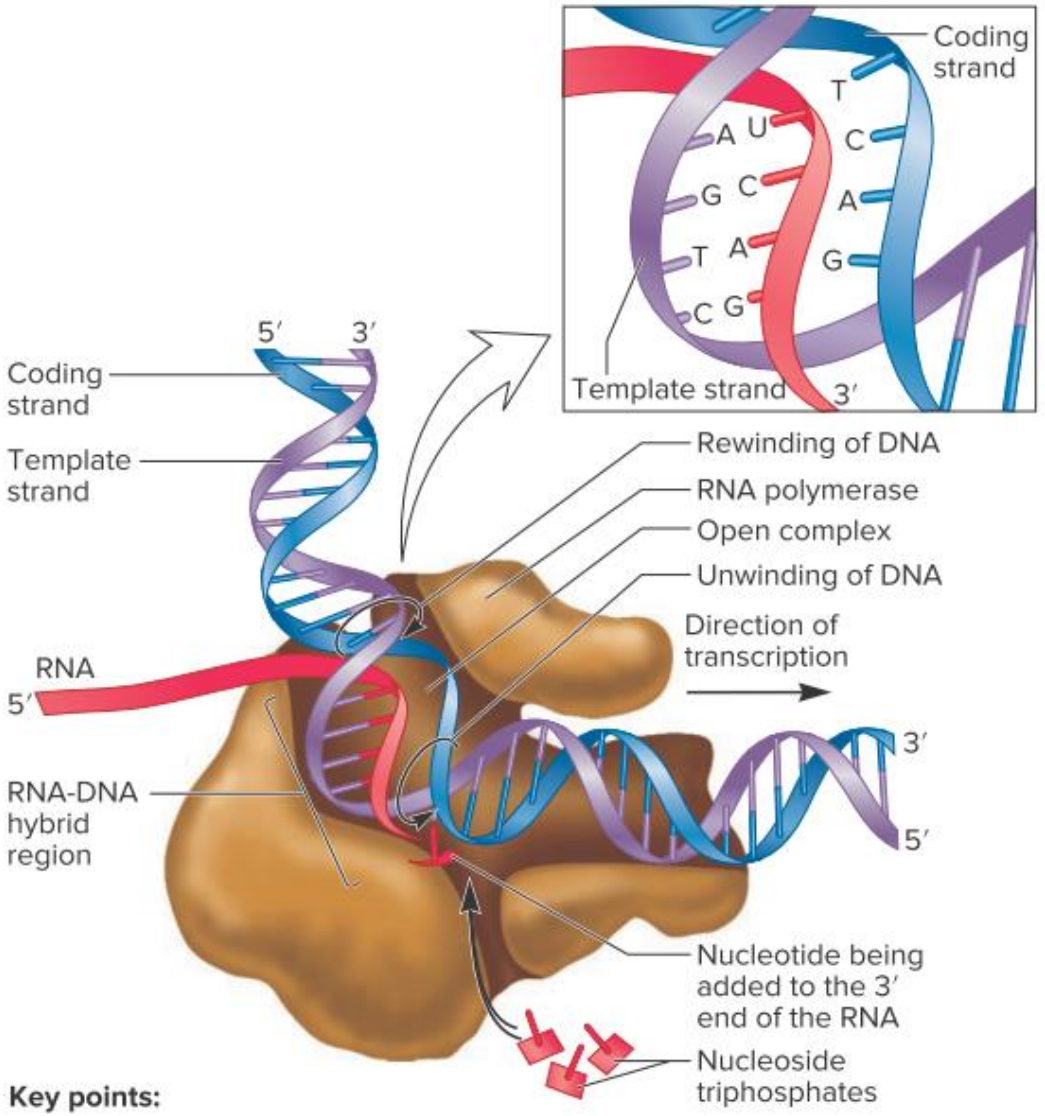

(A)
closed complex forms at promoter
open complex
abortive initiation
promoter clearance and σ factor release
elongation
termination hairpin formation
termination
reassociation with σ factor
σ factor
RNA polymerase
sliding
1
2
3
4
5
6
7
8



The initiation stage of transcription in bacteria

Synthesis of the RNA transcript (elongation)

- RNA polymerase slides along the DNA, creating an open complex as it moves.
- The DNA strand known as the template strand is used to make a complementary copy of RNA, resulting in an RNA-DNA hybrid.
- RNA polymerase moves along the template strand in a 3′ to 5′ direction and RNA is synthesized in a 5′ to 3′ direction using nucleoside triphosphates as precursors.
- The complementarity rule is the same as the AT/GC rule except that U is substituted for T in the RNA.



Termination

- The end of RNA synthesis is referred to as termination. Prior to termination, the hydrogen bonding between the **DNA and RNA within the open complex** is of central importance in preventing dissociation of RNA polymerase from the template strand. Termination occurs when this short **RNA-DNA hybrid region is forced to separate**, thereby releasing RNA polymerase as well as the newly made RNA transcript. In *E. coli*, two different mechanisms for termination have been identified.

- For certain genes, an RNA-binding **protein known as ρ (rho)** is responsible for terminating transcription, in a mechanism called **ρ-dependent termination**. For other genes, termination does not require the involvement of the **ρ protein—and** in these cases, it is referred to as **ρ-independent termination**.

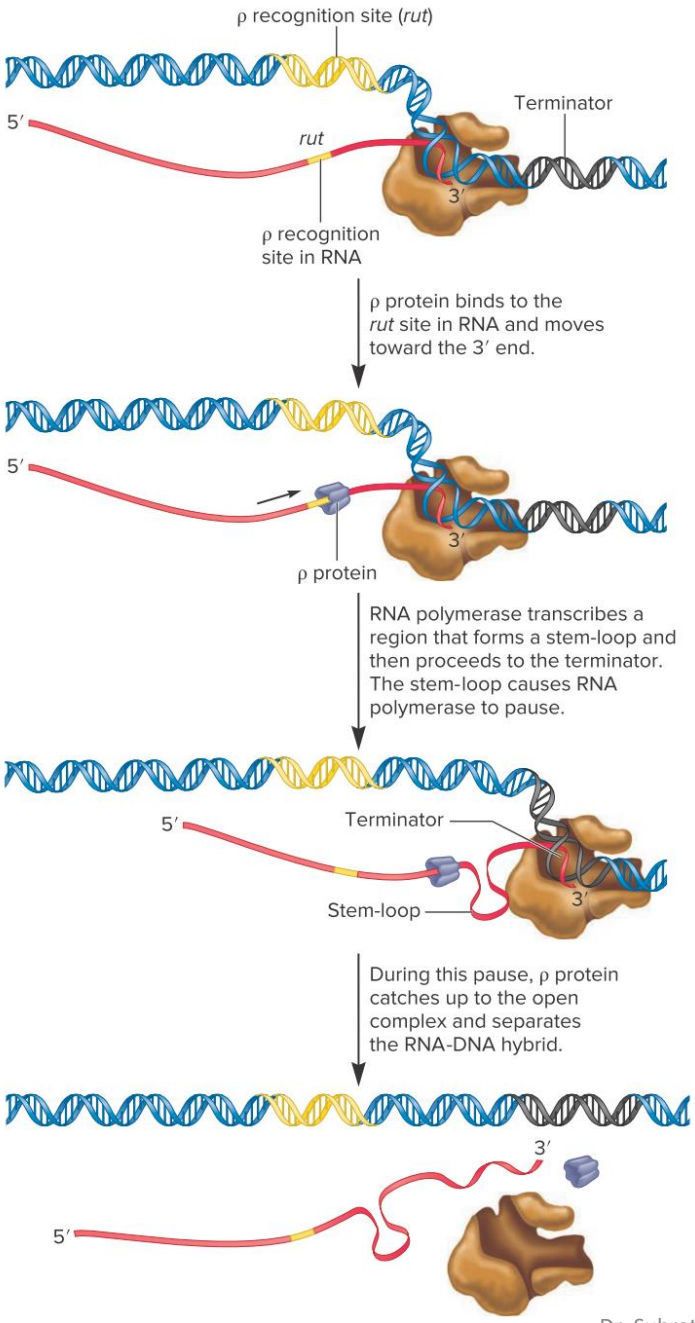
ρ-Dependent Termination

- In ρ-dependent termination, the termination process requires two components. First, a site in the DNA, called the **rut site** (for rho utilization site), encodes a sequence in the RNA that acts as a **recognition site for the binding of the ρ protein**.

- After the rut site is synthesized in the RNA, ρ protein binds to the RNA and moves in the direction of RNA polymerase. The ρ protein functions as a **helicase**, an enzyme that can separate **RNA-DNA hybrid regions**.

- The second component of ρ-dependent termination is the site where termination actually takes place. At this terminator site, the DNA encodes an RNA sequence containing several **GC base pairs that form a stem-loop structure (hairpin structure)**. RNA synthesis terminates several nucleotides beyond this stem-loop.

- This stem-loop forms almost immediately after the RNA sequence is synthesized and quickly binds to RNA polymerase. This binding results in a conformational change that causes RNA polymerase to pause in its synthesis of RNA.



- The pause allows ρ protein to catch up to the stem-loop, pass through it, and break the hydrogen bonds between the DNA and RNA within the open complex. When this occurs, the completed RNA strand is separated from the DNA along with RNA polymerase.

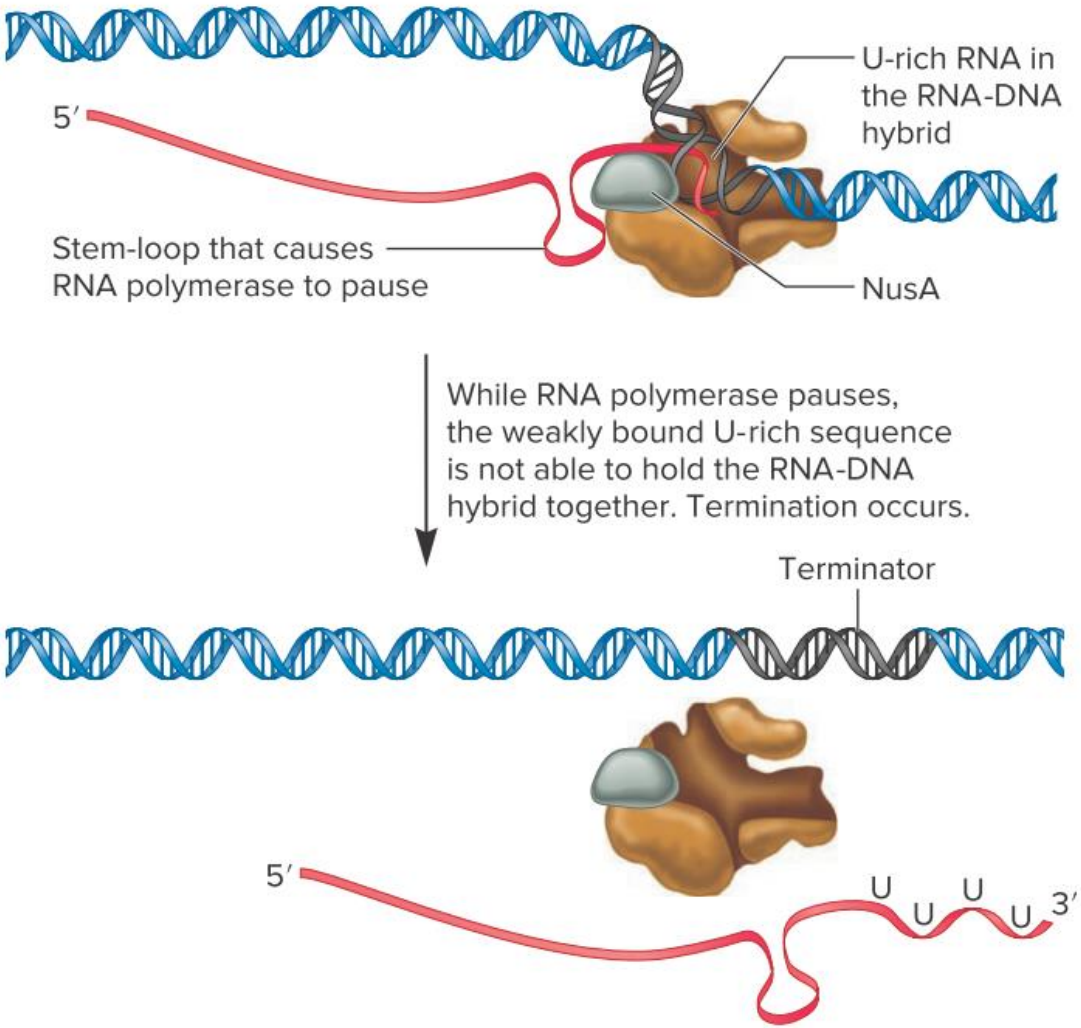
ρ-independent or intrinsic termination

- The process of ρ-independent termination does not require the ρ protein. In this case, the terminator involves two adjacent nucleotide sequences. One sequence promotes the formation of a stem-loop. The second sequence, which is downstream from the stem-loop, is a uracil-rich sequence located at the 3′ end of the RNA.

- The formation of the stem-loop causes RNA polymerase to pause in its synthesis of RNA. This pausing is stabilized by other proteins that bind to RNA polymerase. For example, a protein called **NusA binds to RNA polymerase and promotes pausing at stem-loop sequences**. At the precise time that RNA polymerase pauses, the uracil-rich sequence in the RNA transcript is bound to the DNA template strand.

- The hydrogen bonding of RNA to DNA keeps RNA polymerase clamped onto the DNA. However, the binding of this uracil-rich sequence to the DNA template strand is relatively weak, causing the RNA transcript to spontaneously dissociate from the DNA and cease further transcription.

ρ-independent or intrinsic termination



Because this process does not require the ρ protein to physically remove the RNA transcript from the DNA, it is also referred to as ***intrinsic termination***.

Transcription in eukaryotes

- The genetic material within the nucleus of a eukaryotic cell is transcribed by three different RNA polymerase enzymes, designated RNA polymerase I, II, and III.

- RNA polymerase I: transcribes all of the genes for ribosomal RNA (rRNA) except for the 5S rRNA.

- RNA polymerase II: transcribes all protein-encoding genes. Therefore, it is responsible for the synthesis of all mRNAs. It also transcribes the genes for most snRNAs which are needed for RNA splicing. In addition, it transcribes several types of genes that produce other non-coding RNAs, including most long non-coding RNAs, microRNAs, and snoRNAs.

- RNA polymerase III: transcribes all tRNA genes and the 5S rRNA gene. To a much lesser extent than RNA polymerase II, it also transcribes a few genes that produce other non-coding RNAs, such as snRNAs, long non-coding RNAs, microRNAs, and snoRNAs.

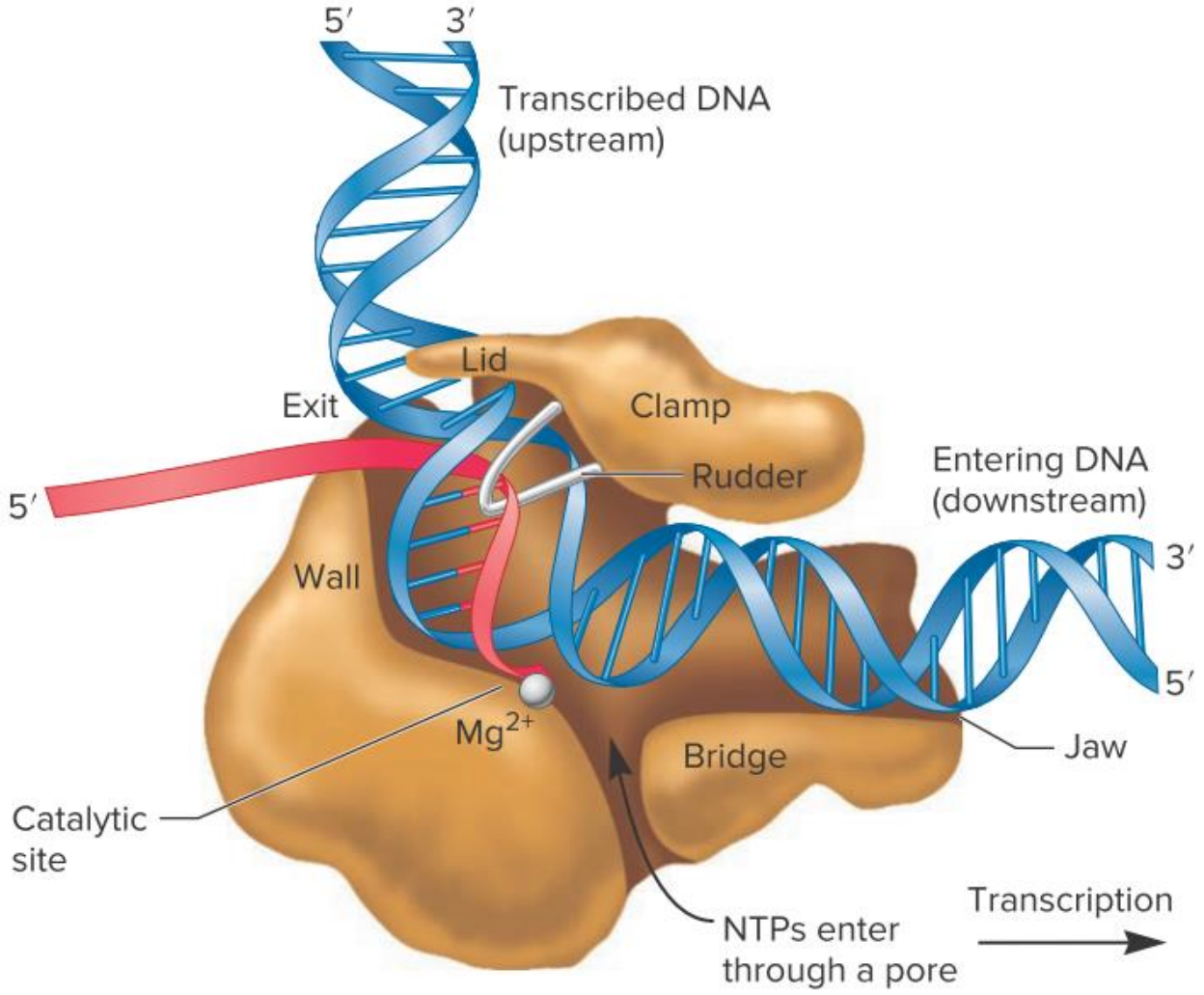
- The DNA enters the enzyme through the jaw and lies on a surface within RNA polymerase termed the bridge. The part of the enzyme called the clamp is thought to control the movement of the DNA through RNA polymerase.

- A wall in the enzyme forces the RNA-DNA hybrid to make a right-angle turn. This bend facilitates the ability of nucleotides to bind to the template strand. Mg^{2+} is located at the catalytic site, which is precisely at the 3′ end of the growing RNA strand.

- Nucleoside triphosphates (NTPs) enter the catalytic site via a pore region. The correct nucleotide binds to the template DNA and is covalently attached to the 3′ end.

- As RNA polymerase slides down the template, a rudder, which is about 9 bp away from the 3′ end of the RNA, forces the RNA-DNA hybrid apart. The DNA and the single-stranded RNA then exit under a small lid.

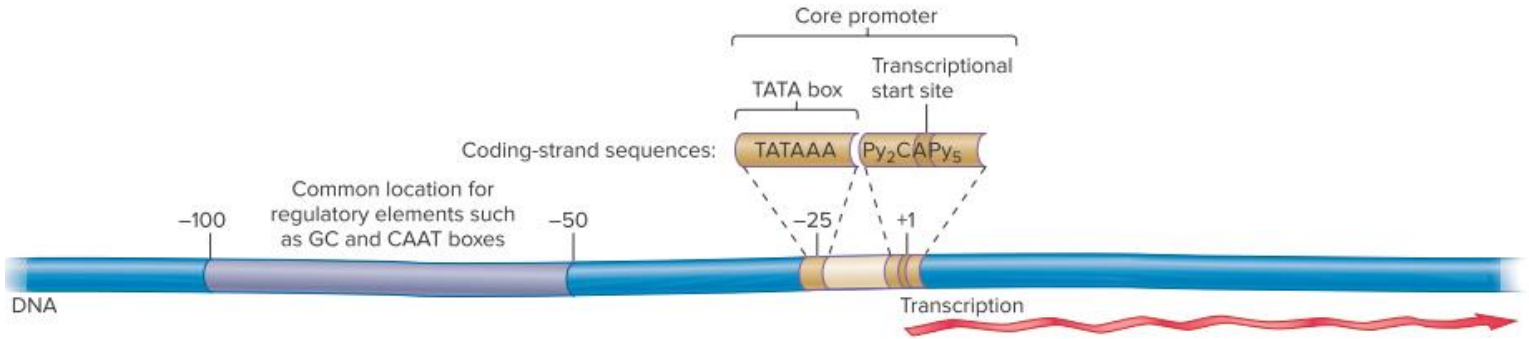
Structure and molecular function of RNA polymerase



- For transcription to occur at an appropriate rate, eukaryotic genes have two features: **a core promoter** and **regulatory elements**.

- The core promoter is a relatively short DNA sequence that is necessary for transcription to take place. It typically consists of a **TATAAA sequence** called the **TATA box** and the **transcriptional start site**. The TATA box, which is usually about **25 bp upstream** from a transcriptional start site, is important in **determining the precise starting point** for transcription. The core promoter, by itself, produces a low level of transcription. This is termed **basal transcription**.

- There are two categories of regulatory elements. Activating sequences, known as **enhancers**, are needed to stimulate transcription and **silencers** (DNA sequences that are recognized by transcription factors that inhibit transcription).



- Three categories of proteins are needed for basal transcription at the core promoter: RNA polymerase II, general transcription factors, and a complex called mediator.

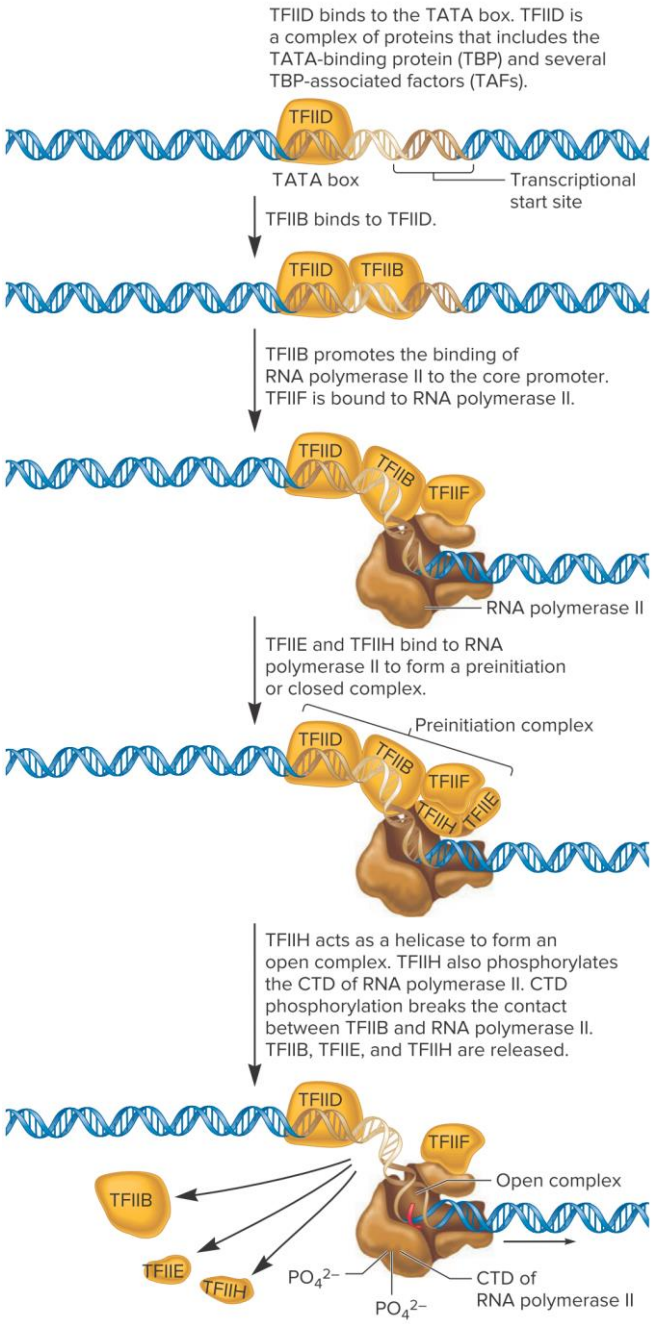
- RNA polymerase II: The enzyme that catalyzes the linkage of nucleotides in the 5′ to 3′ direction, using DNA as a template. Most eukaryotic RNA polymerase II proteins are composed of 12 subunits. The two largest subunits are structurally similar to the β and β′ subunits found in E. coli RNA polymerase.

General transcription factors:

- **TFIID:** Composed of TATA-binding protein (TBP) and other TBP-associated factors (TAFs). Recognizes the TATA box of eukaryotic protein-encoding gene promoters.
- **TFIIB:** Binds to TFIID and then enables RNA polymerase II to bind to the core promoter. Also promotes TFIIF binding.
- **TFIIF:** Binds to RNA polymerase II and plays a role in its ability to bind to TFIIB and the core promoter. Also plays a role in the ability of TFIIE and TFIIH to bind to RNA polymerase II.
- **TFIIE:** Plays a role in the formation or the maintenance (or both) of the open complex. It may exert its effects by facilitating the binding of TFIIH to RNA polymerase II and regulating the activity of TFIIH. .

- **TFIIH**: A multisubunit protein that has multiple roles. First, certain subunits act as helicases and promote the formation of the open complex. Other subunits phosphorylate the carboxyl terminal domain (CTD) of RNA polymerase II, which releases its interaction with TFIIB, thereby allowing RNA polymerase II to proceed to the elongation phase.

- **Mediator:** A multisubunit complex that mediates the effects of regulatory transcription factors on the function of RNA polymerase II. Though mediator typically has certain core subunits, many of its subunits vary, depending on the cell type and environmental conditions. The ability of mediator to affect the function of RNA polymerase II is thought to occur via the CTD of RNA polymerase II. Mediator can influence the ability of TFIIH to phosphorylate CTD, and subunits within mediator itself have the ability to phosphorylate CTD. Because CTD phosphorylation is needed to release RNA polymerase II from TFIIB, mediator plays a key role in the ability of RNA polymerase II to switch from the initiation to the elongation stage of transcription.



Steps leading to the formation of the open complex

Possible mechanisms for transcriptional termination of a protein-encoding gene.