

TRANSCRIPTION

phD Luu Phuc Loi
Teaching assistants: Phuong Thao & Hoang Kim

Agenda

- **Introduction to Transcription**
- **Key Features of Transcription**
- **Key Components for Transcription**
- **Gene Structure of Prokaryotes and Eukaryotes**
- **Transcription Stages:**
Initiation → Elongation → Termination (→ Post transcriptional modification)



DNA

RNA

Introduction to transcription

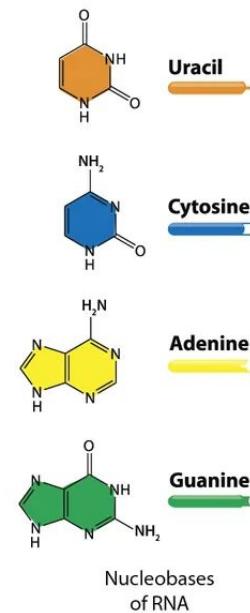
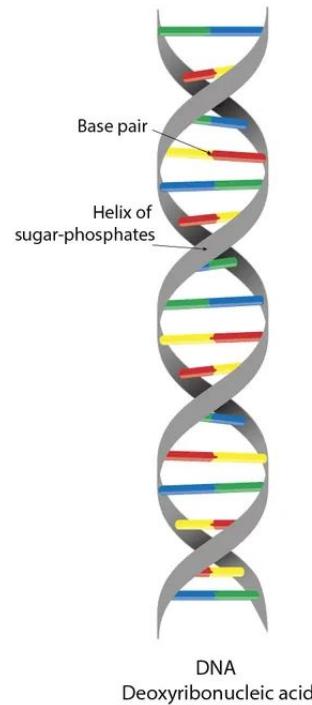
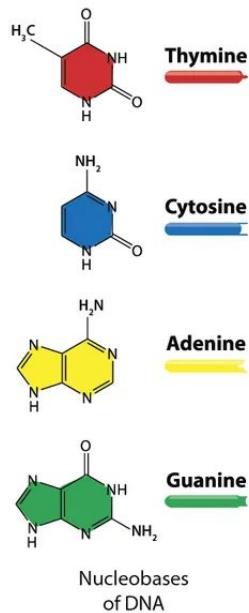
- Transcription is the process of copying a segment of **DNA into RNA**.

Two big questions in transcription:

- 1) **Specificity:** How does RNA polymerase accurately recognize specific promoters to ensure that only the correct genes are transcribed, while the majority of the genome remains silent?
- 1) **Efficiency (frequency):** What determines the transcription frequency of a gene, and why are some genes expressed at low levels whereas others are expressed at high levels?

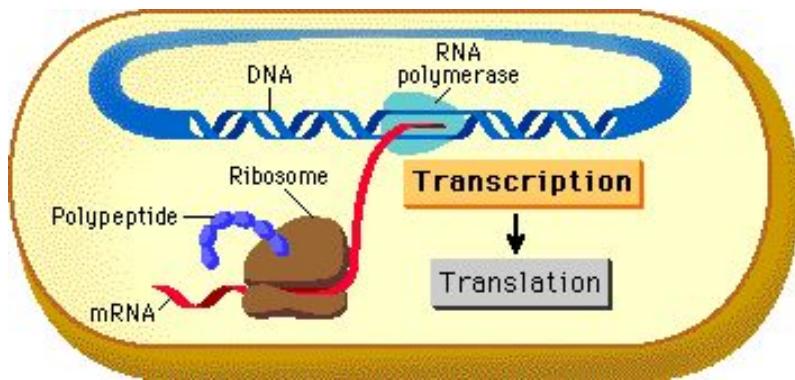


DNA vs RNA

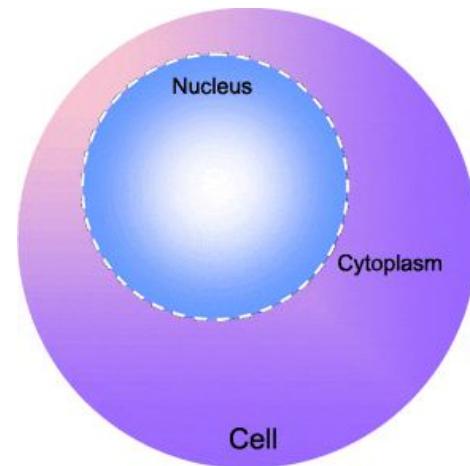


Where Transcription take place in Prokaryotic and Eukaryotic cells

- **Prokaryotes:** Transcription takes place in the **cytoplasm**.
- **Eukaryotes:** Transcription mainly occurs in the **nucleus**. **Mitochondria** (and **chloroplasts** in plants) have their own DNA, so transcription also occurs inside these organelles.



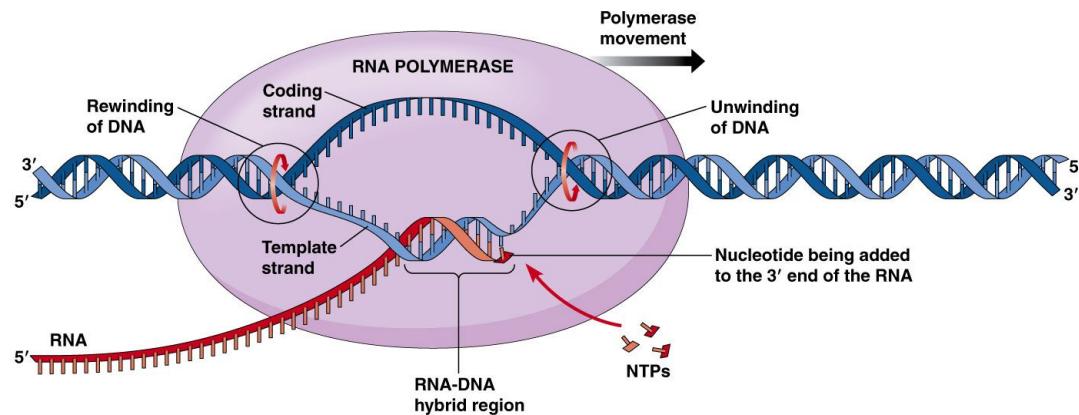
Prokaryote



Eukaryote

Template Strand vs. Coding Strand in Transcription

- **Template strand (antisense):** RNA polymerase reads this strand **$3' \rightarrow 5'$** to synthesize **mRNA $5' \rightarrow 3'$** ; the mRNA sequence is **complementary** to the template strand.
- **Coding strand (sense):** Not used as the template; its sequence **matches the mRNA** (except T in DNA → U in RNA), and the mRNA from **protein-coding genes** is the transcript that can be **translated into protein**.

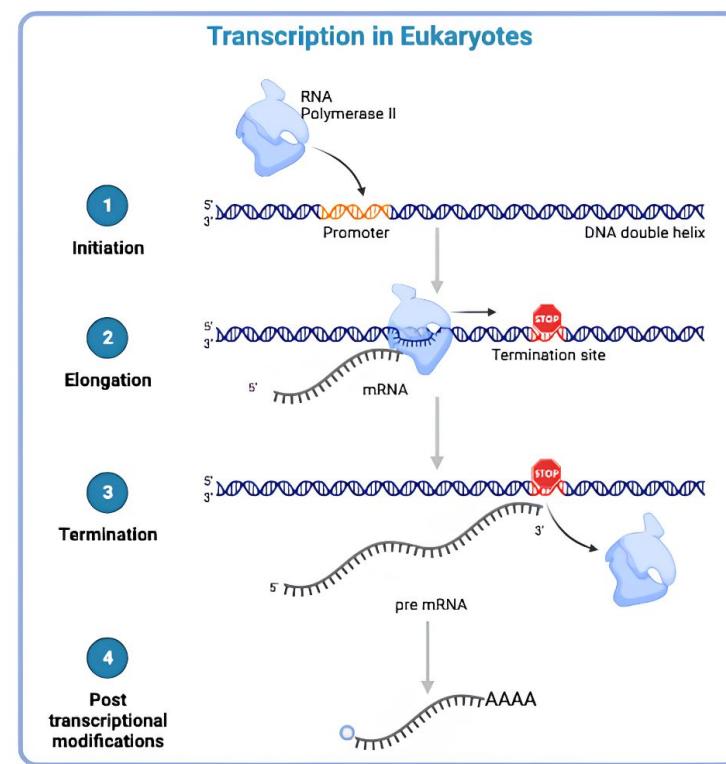
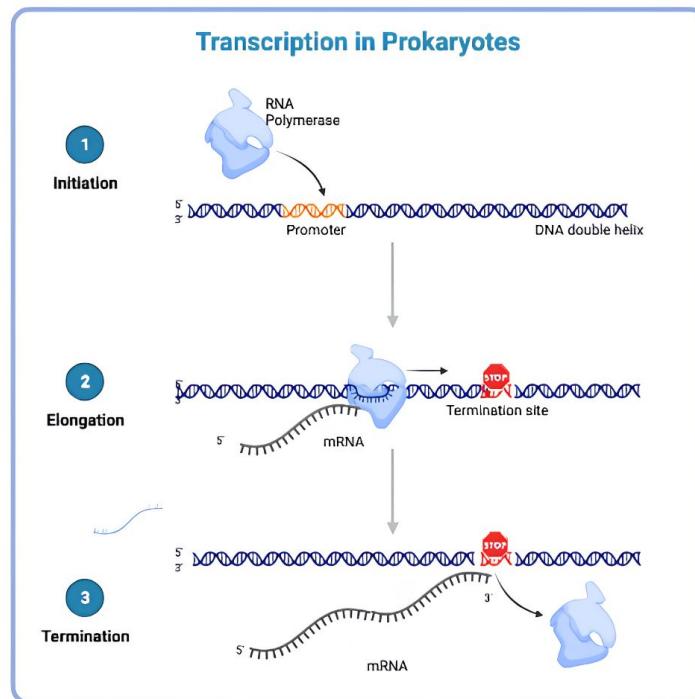


Transcription key components

- **RNA polymerase:** the enzyme that catalyzes RNA synthesis
- **DNA template:** the gene plus regulatory DNA elements (promoter, regulatory sites, termination signals)
- **NTPs:** ATP, GTP, CTP, UTP (substrates for RNA synthesis)
- **Transcription factors:** proteins that help recruit/position RNA polymerase and regulate transcription (e.g., σ factor in bacteria; general TFs/activators in eukaryotes)
- **Divalent metal ion:** typically Mg²⁺ (required for catalysis)
- **Accessory factors:** termination factors (e.g., Rho in bacteria) and, in eukaryotes, RNA-processing factors (capping, splicing, polyadenylation) linked to transcription

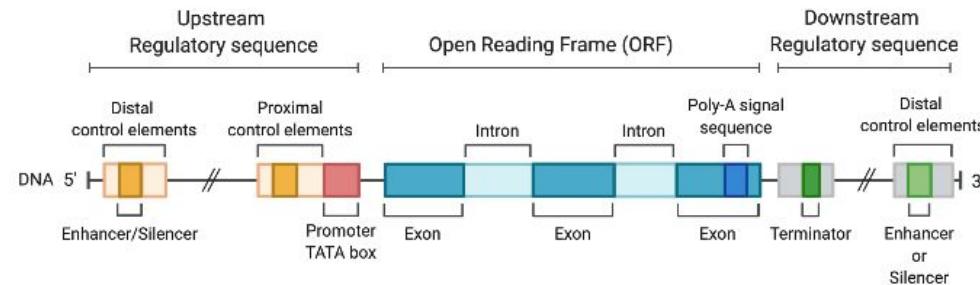


Transcription general process

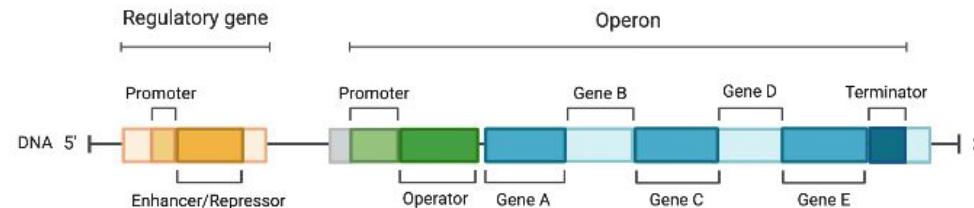


General structure of prokaryotic and eukaryotic gene

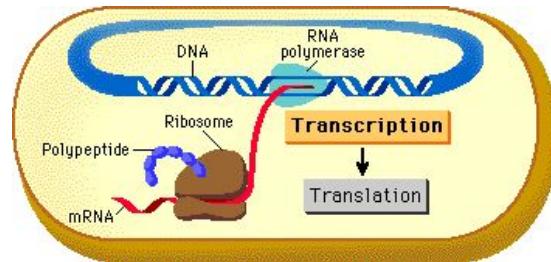
Eukaryotic Gene Structure



Prokaryotic Gene Structure



TRANSCRIPTION IN PROKARYOTES



Initiation in prokaryotes

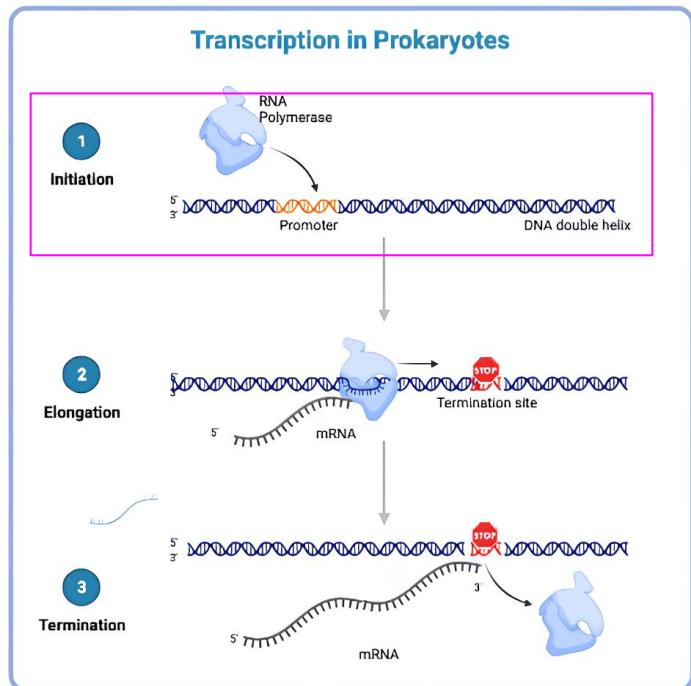
Initiation is the first phase of transcription in which the **RNA polymerase holoenzyme** recognizes a **promoter**, forms an **open complex (transcription bubble)** at the **TSS (+1)**, and synthesizes the first RNA nucleotides before entering **elongation**.

Required components

- Promoter DNA with consensus elements: **-35 (TTGACA)** and **-10/Pribnow box (TATAAT)**
- RNA polymerase holoenzyme = core enzyme ($\alpha_2\beta\beta'\omega$) + σ factor (promoter recognition)
- rNTPs: ATP, GTP, CTP, UTP

Steps

1. **Promoter recognition & binding (closed complex):** σ factor binds **-35/-10** → selects promoter and positions RNAP at **TSS (+1)**.
2. **Open complex formation:** core RNAP melts DNA near TSS (AT-rich -10 helps) → forms **transcription bubble**.
3. **Initial RNA synthesis:** RNAP starts making RNA; may undergo **abortive initiation** (short RNAs ~2–9 nt).
4. **Promoter escape → elongation:** once RNA reaches **~8–10 nt**, RNAP becomes stable and escapes the promoter; σ factor often dissociates



RNA polymerase in prokaryotes

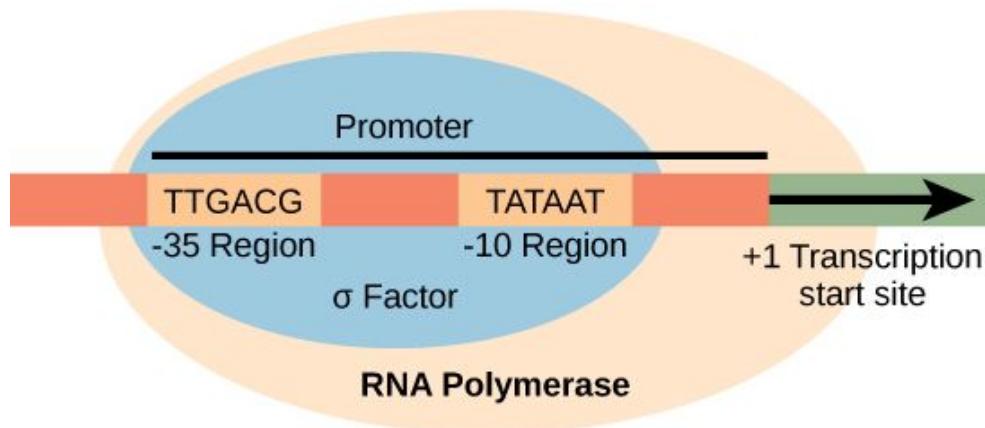
- The main enzyme involved in transcription is RNA polymerase, which uses a single-stranded DNA template to synthesize a complementary strand of RNA. Specifically, RNA polymerase builds an RNA strand in the 5' to 3' direction, adding each new nucleotide to the 3' end of the strand.
- In **prokaryotes**, a **single RNA polymerase holoenzyme** (core enzyme $\alpha_2\beta\beta'\omega$ plus a σ factor) transcribes **all major RNA types**, including **mRNA, rRNA, and tRNA**.

**Prokaryotic RNA Polymerase:
Holoenzyme Enzyme**

Subunit	Size	#/Molecule	Function
α	36.5 kD	2	chain initiation and interaction with regulatory proteins
β	151 kD	1	chain initiation and elongation
β'	155 kD	1	DNA binding
σ	70 kD	1	promoter recognition



Promoters of Eukaryotes

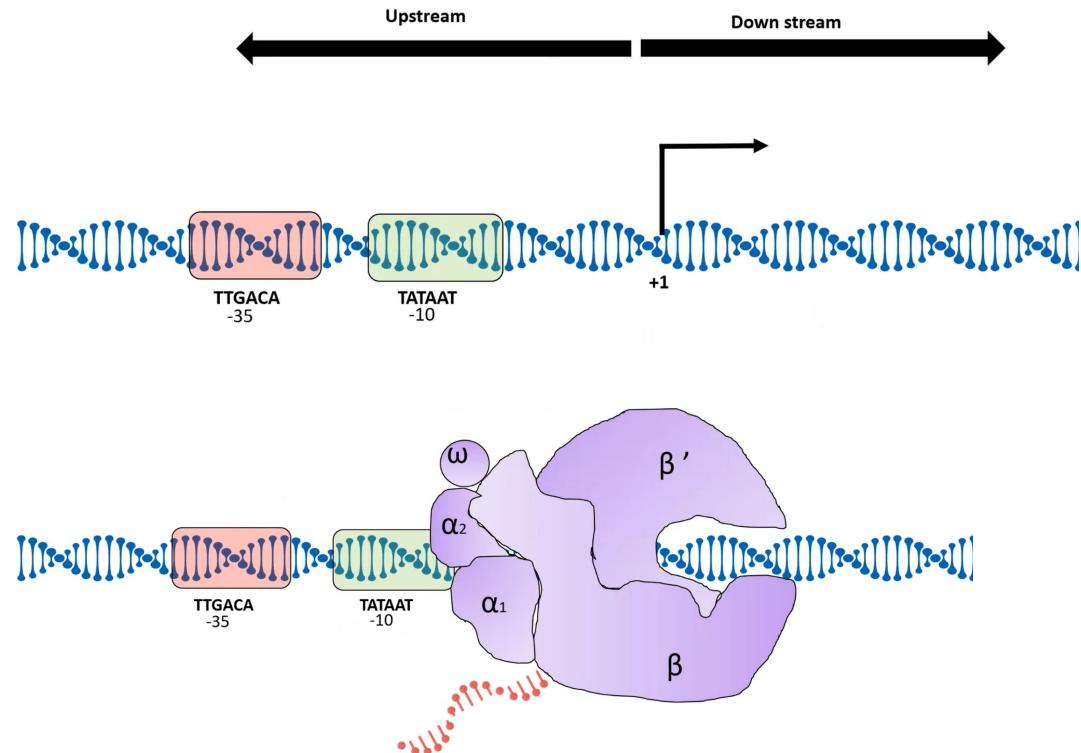
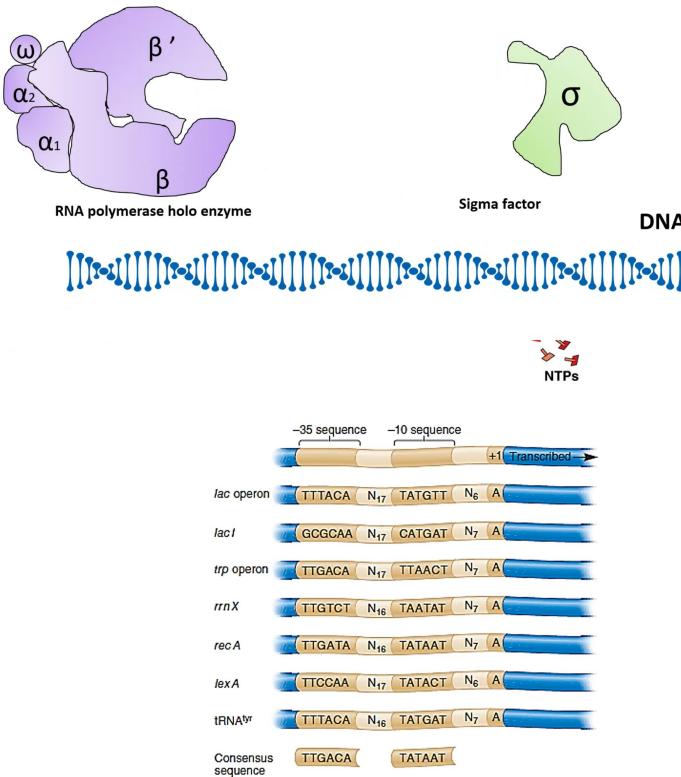


Prokaryote

Prokaryotic promoters

- Prokaryotic promoters have two key elements upstream of the TSS: **-10** and **-35**.
- **-10 (Pribnow box): TATAAT** — essential for transcription initiation.
- **-35: TTGACA** — contributes to strong promoter recognition and higher transcription.
- **TSS = +1**; positions **upstream** are negative (-), **downstream** are positive (+).

Initiation in Prokaryotes



Elongation in Prokaryotes

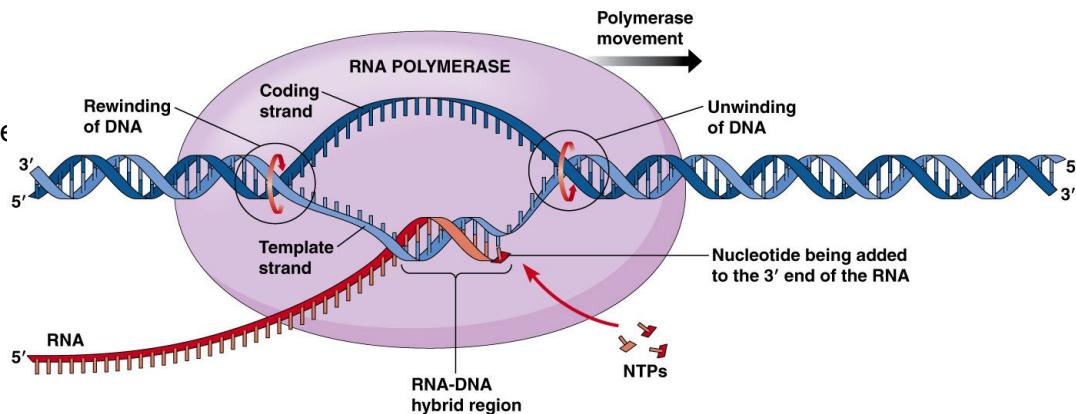
Definition: The phase of transcription where **core RNAP** moves along the **template strand**, continuously adds **rNTPs** to extend RNA in the **5'→3'** direction.

Required components:

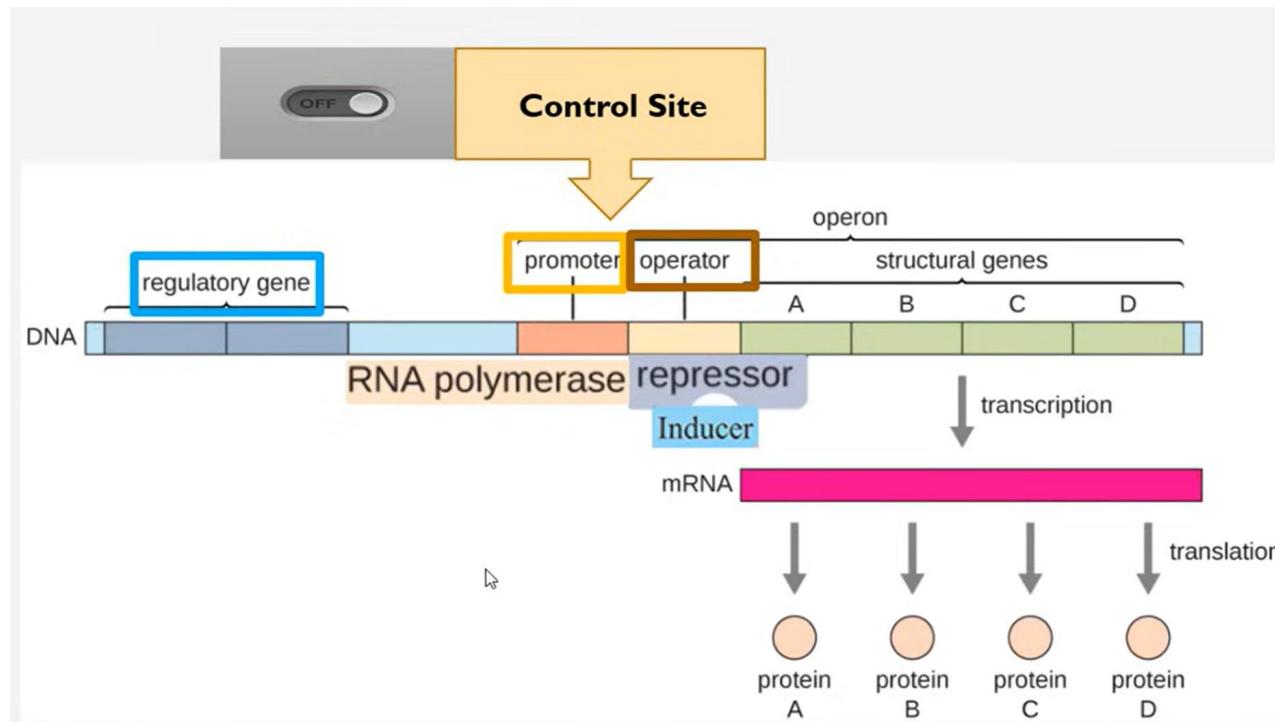
- **Core RNA polymerase** (σ usually not needed after promoter escape)
- **DNA template strand** (read 3'→5')
- **rNTPs** (ATP, GTP, CTP, UTP) + **Mg²⁺** (catalysis)
- **Transcription bubble** (~unwound DNA region)

Key steps:

1. **Template selection:** RNAP reads **template (antisense) strand 3'→5'**
2. **RNA synthesis:** RNA grows 5'→3'; sequence matches **coding strand** except **U replaces T**
Bubble movement: RNAP unwinds ahead and re-anneals behind (positive supercoils ahead, negative behind)



Transcription regulation in prokaryotes



Termination in Prokaryotes

Termination is the final stage of transcription in which termination signals cause:

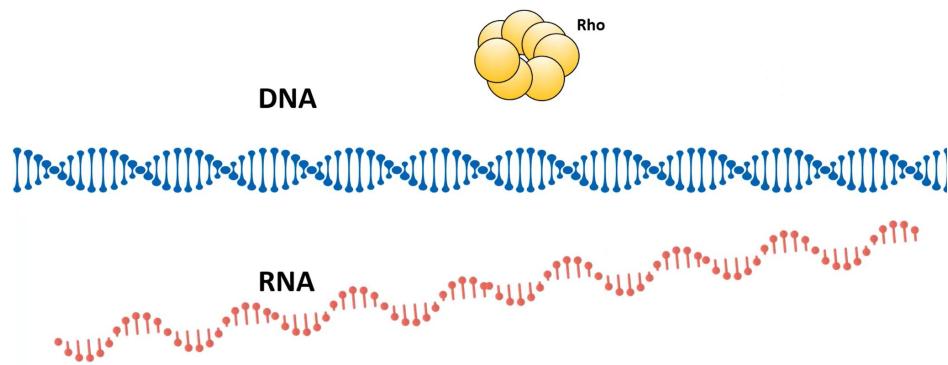
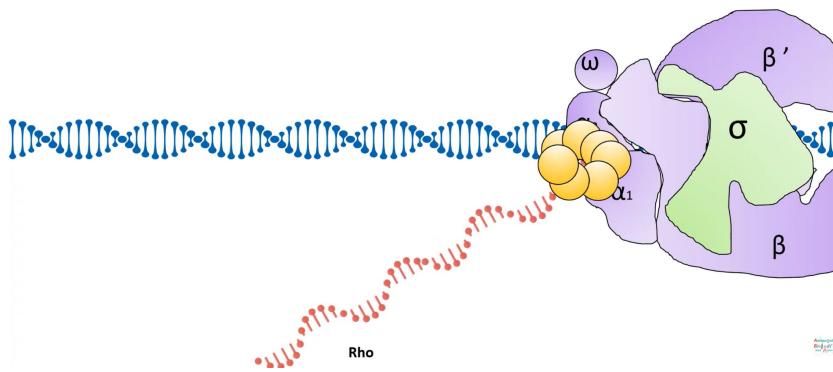
1. **RNA polymerase pauses/stops RNA synthesis**
2. **The RNA transcript is released** from the transcription complex
3. **RNA polymerase dissociates from the DNA**, and the complex disassembles

Required components

- **DNA terminator signals**
- Either:
 - **Rho factor (ATP-dependent RNA helicase) (*Rho-dependent*)**
 - **Intrinsic terminator sequence: GC-rich inverted repeats + poly-U tract (*Rho-independent*)**

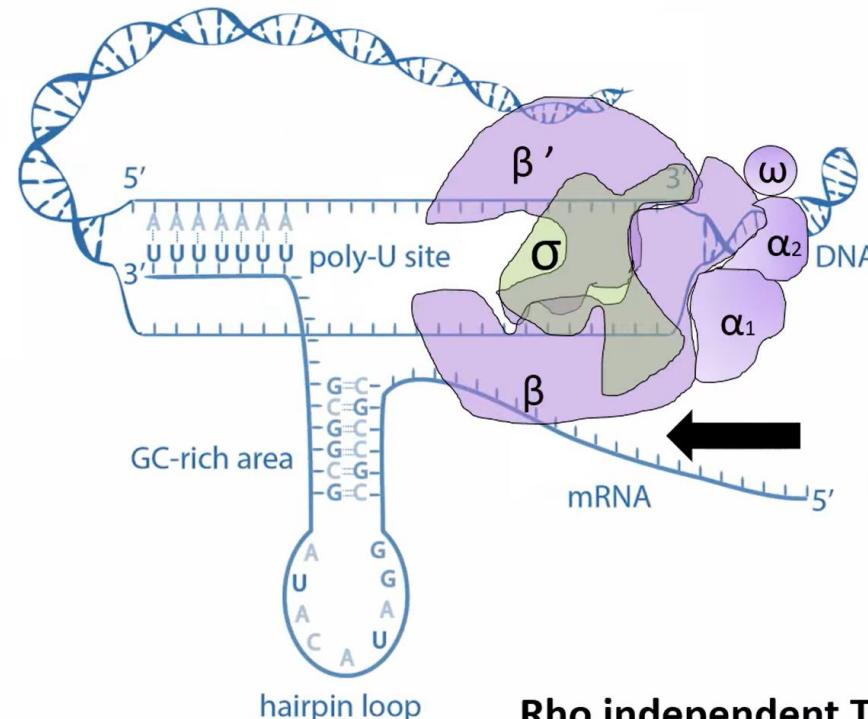
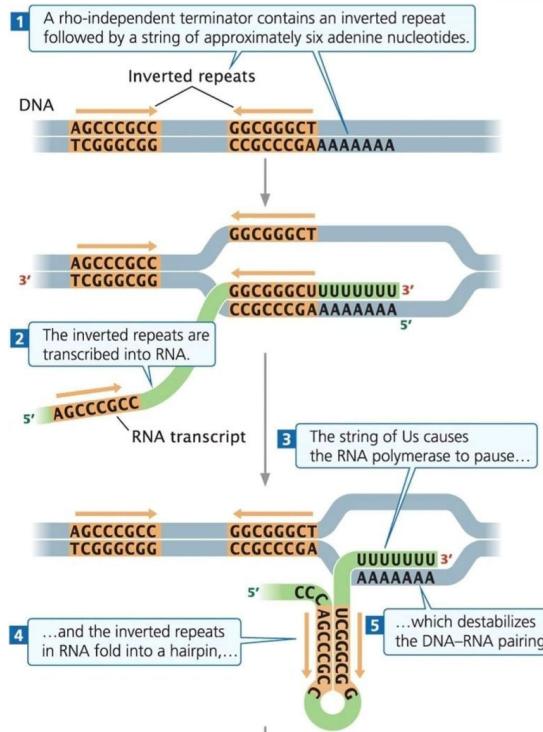
Termination in Prokaryotes

Rho dependent Termination



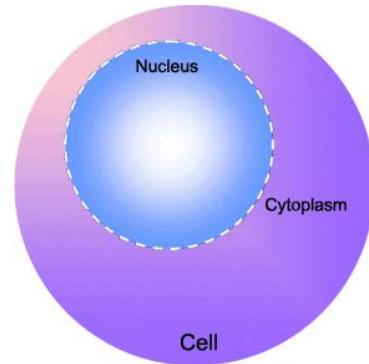
Termination in Prokaryotes

- Inverted repeats → hairpin loop

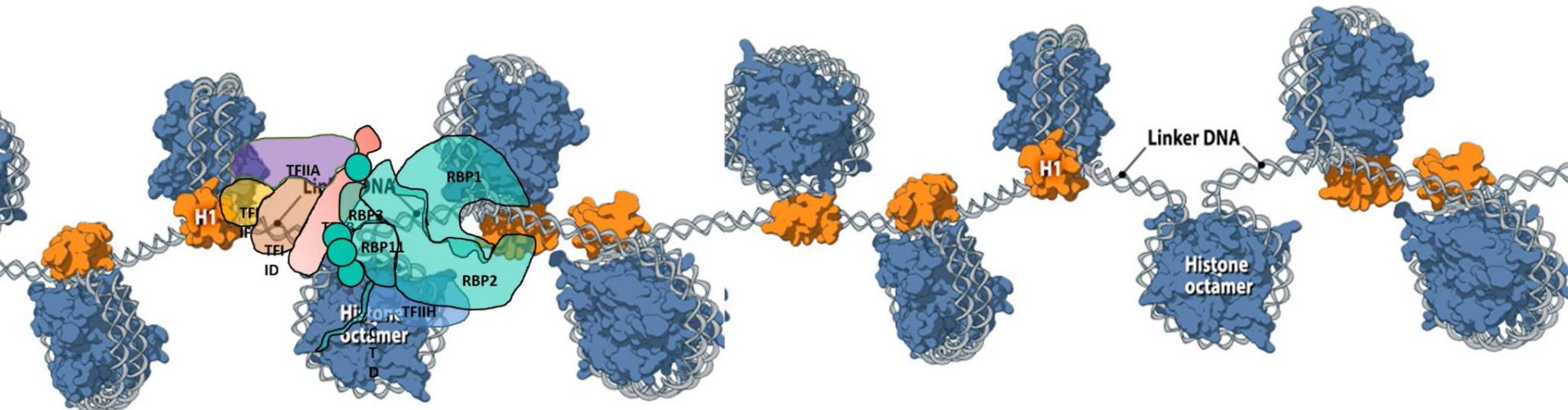


Rho independent Termination

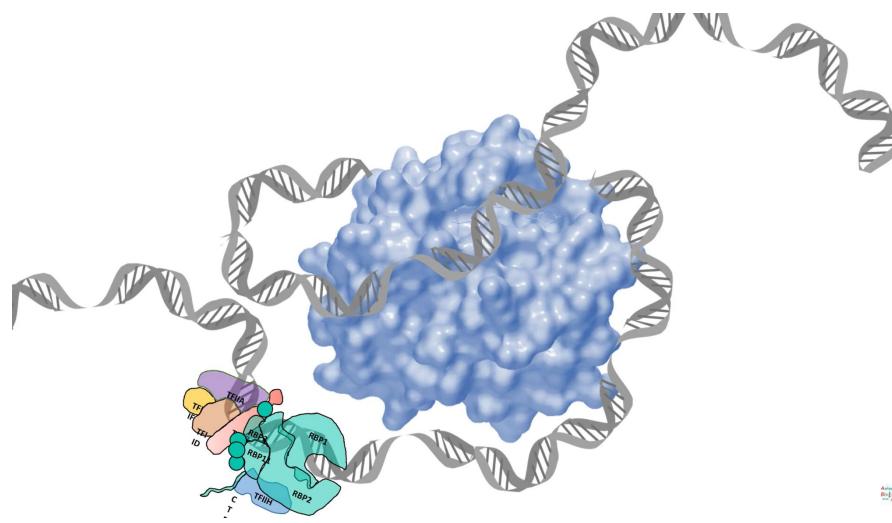
TRANSCRIPTION IN EUKARYOTES



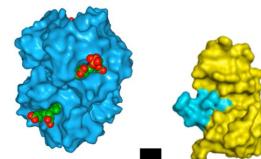
Transcription in eukaryotes



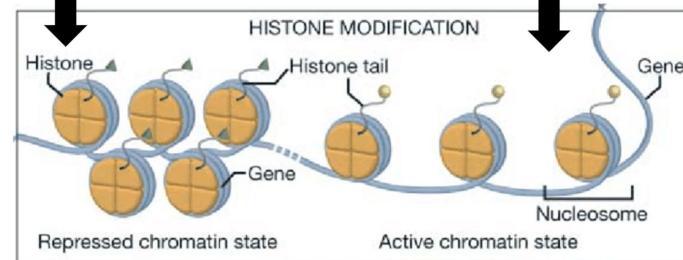
Transcription in eukaryotes



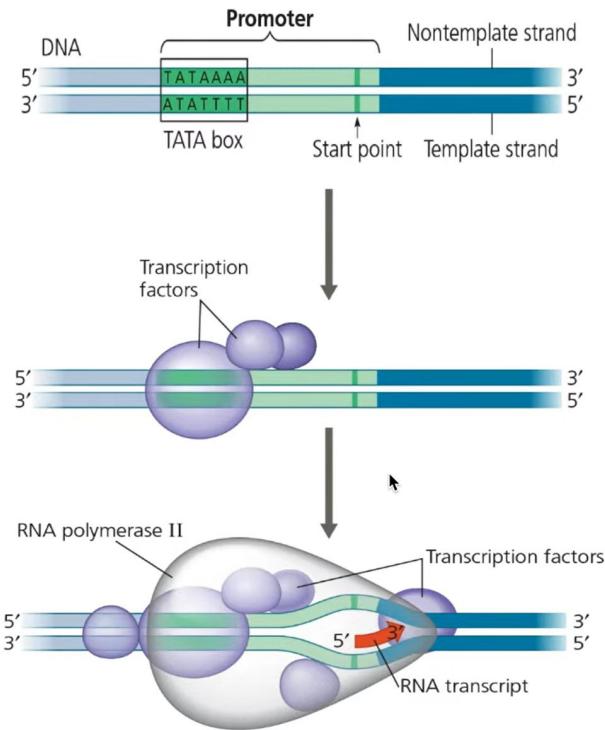
Histone modifiers



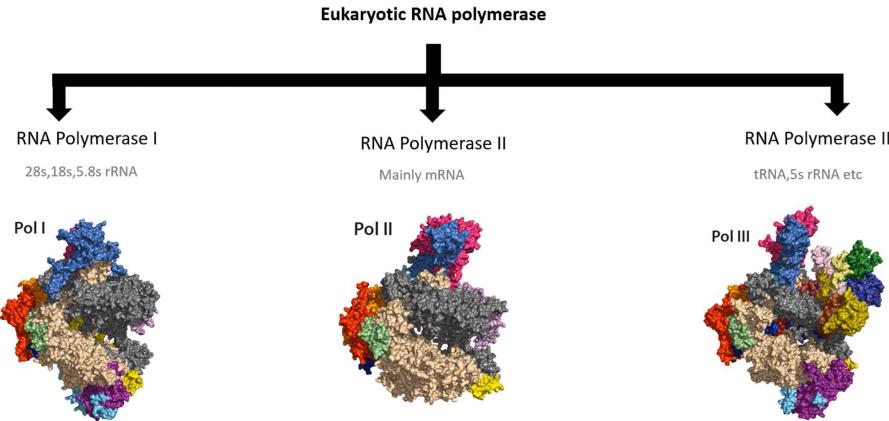
Nucleosome remodelling complex



Initiation in Eukaryotes



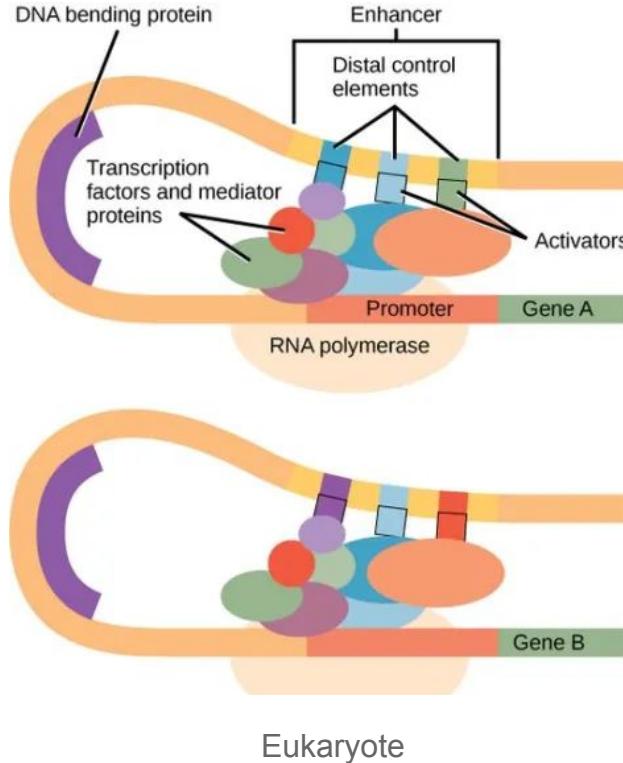
RNA polymerase in eukaryotes



In eukaryotes, there are **three major nuclear RNA polymerases** that synthesize different RNA classes:

- **RNA polymerase I (Pol I):** transcribes most **rRNA** (e.g., **18S, 5.8S, 28S rRNA**)
- **RNA polymerase II (Pol II):** transcribes **mRNA** (protein-coding genes) and many **noncoding RNAs** (e.g., some **snRNAs/miRNA precursors**)
- **RNA polymerase III (Pol III):** transcribes **tRNA**, **5S rRNA**, and other small RNAs (e.g., **U6 snRNA**)

Promoters of Prokaryotes and Eukaryotes

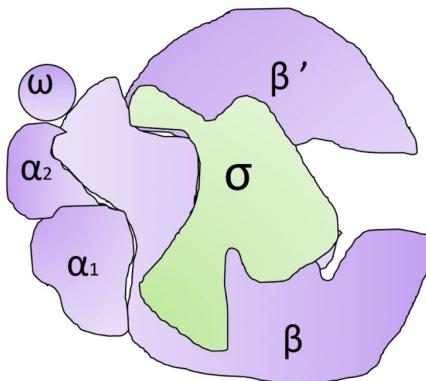


Eukaryotic promoters

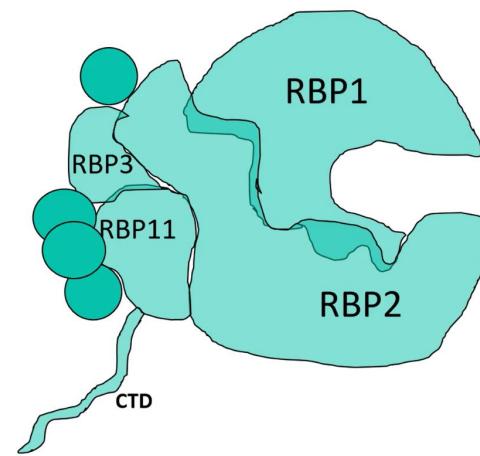
- Eukaryotic promoters are **diverse** and often **more complex** than prokaryotic promoters.
- Regulatory elements can be located **far from the TSS** (often kilobases away).
- Many promoters contain a **TATA box (TATAAA)** that binds **TBP** to help assemble the transcription initiation complex.
- The **TATA box** is usually **close to the TSS** (often within ~50 bp).

RNA polymerase in eukaryotes

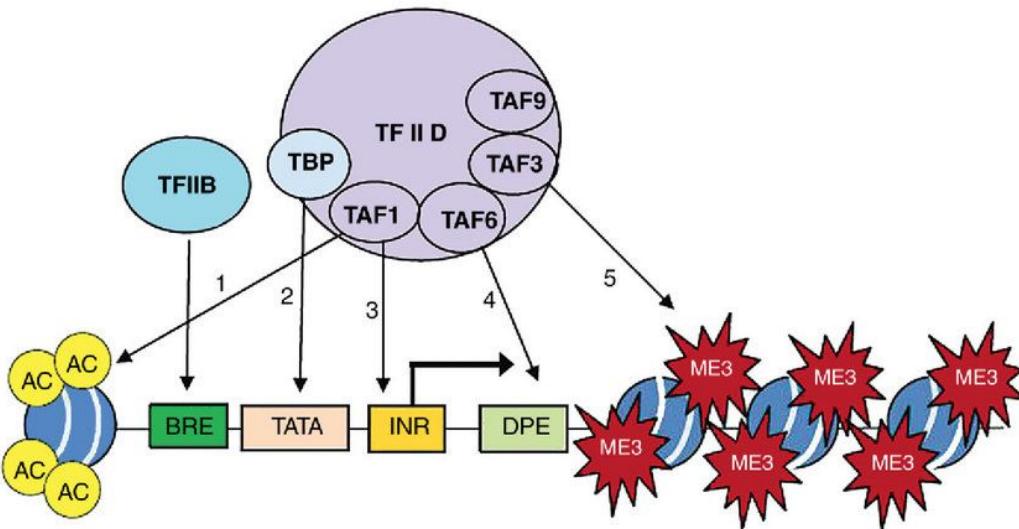
Prokaryotic RNA polymerase



Eukaryotic RNA polymerase



General transcription factor

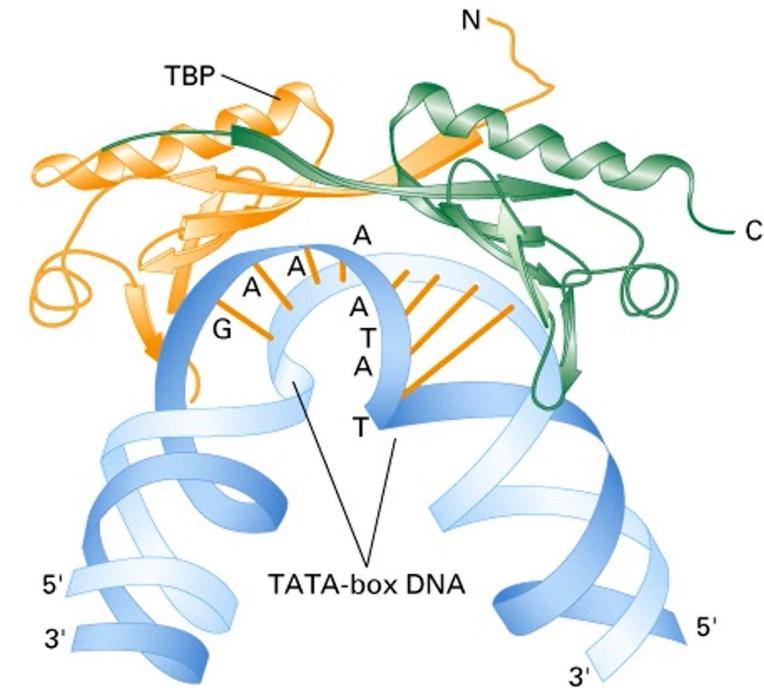
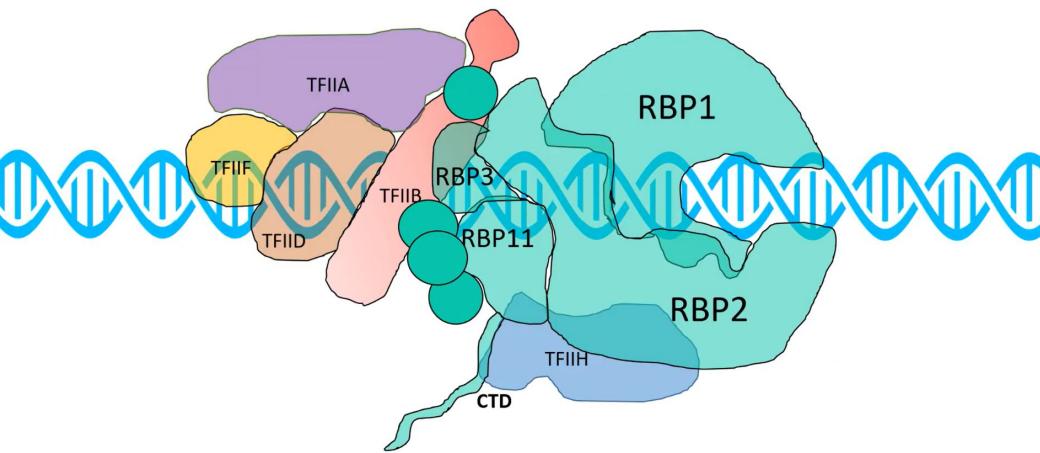


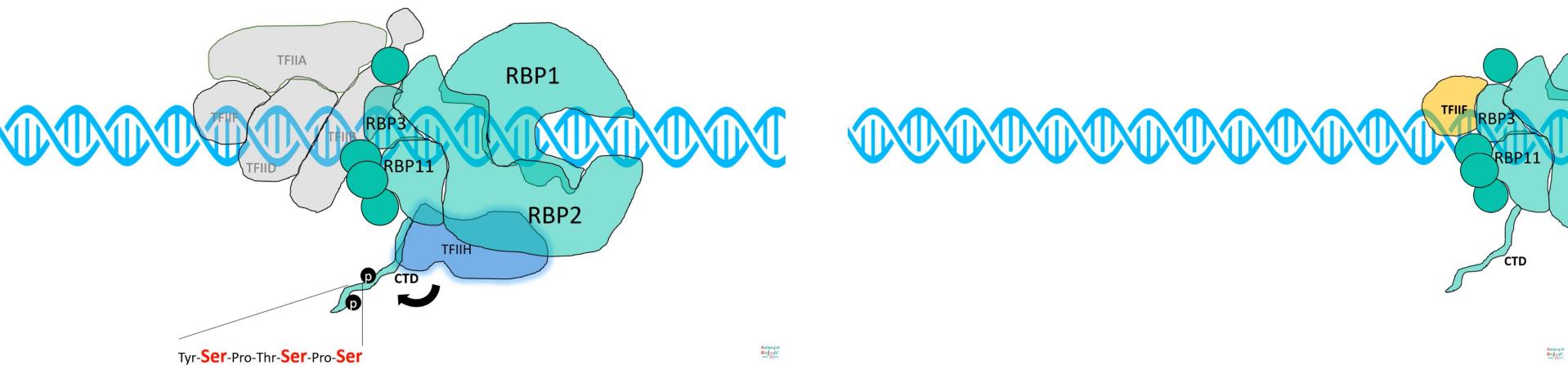
TFIID (TBP + TAFs) is the key *landing platform* for **RNA polymerase II**—often the first GTF to bind the core promoter and initiate **PIC assembly**.

- **TBP** binds the **TATA box** (if present), **bends DNA**, and helps position the transcription start region near the **TSS**.
- **TAFs** enable recognition of **TATA-less promoters** (e.g., via **Inr/DPE**) and help recruit other GTFs (especially **TFIIB**) and **Pol II**.

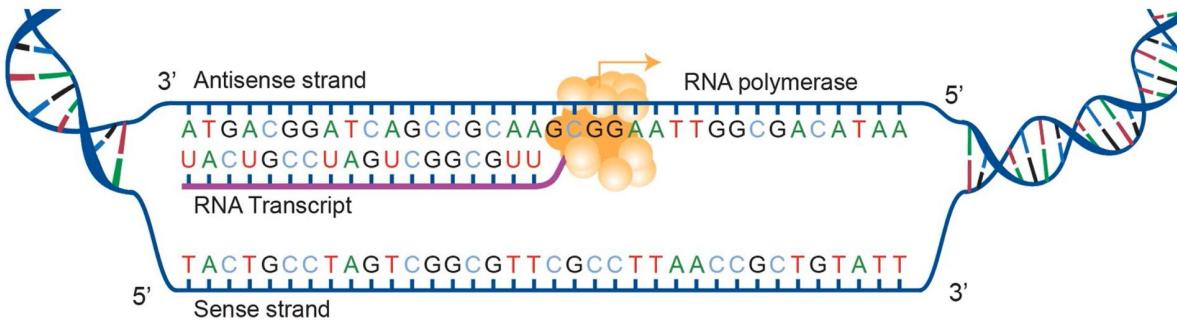
Other general TFs: **TFIIA**, **TFIIB**, **TFIIF**, **TFIIE**, **TFIIFH**.

General transcription factor in eukaryotes





Elongation in prokaryotes



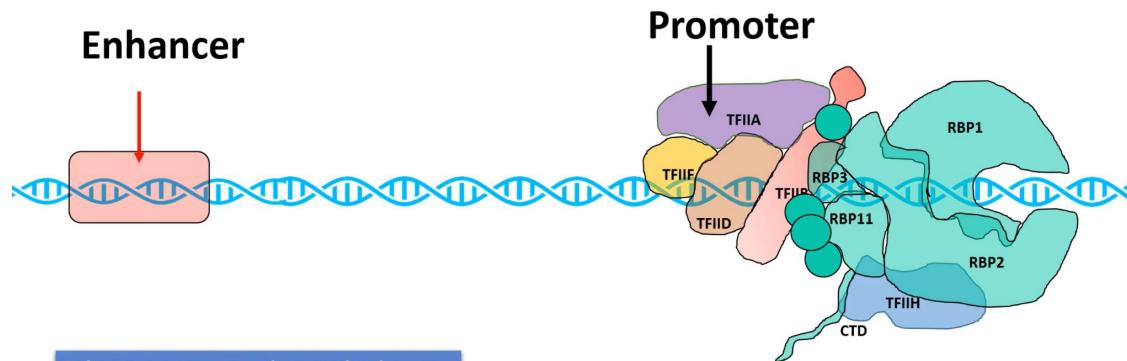
Definition: The phase where **RNA polymerase II (Pol II)** leaves the promoter and **extends the RNA transcript** by adding rNTPs in the $5' \rightarrow 3'$ direction while moving along the **template strand ($3' \rightarrow 5'$)**.

Required components:

- **RNA polymerase II + DNA template strand**
 - **rNTPs (ATP, GTP, CTP, UTP) + Mg²⁺**
 - **CTD (C-terminal domain) of Pol II** (key regulatory “platform”)
- Elongation/processing factors** (e.g., **TFIIF** for promoter escape; **P-TEFb** for productive elongation)

Key steps:

1. **Early elongation:** After ~20–25 nt, most **general TFs** dissociate; Pol II continues (often with **TFIIF**).
2. **Processive RNA synthesis:** Pol II reads **template 3'→5'**, synthesizes **RNA 5'→3'** (Mg^{2+} -dependent catalysis).
3. **Coupling to RNA processing:** **CTD phosphorylation recruits enzymes** for **5' capping, splicing, and later 3'-end processing** while transcription is still ongoing.
4. **Regulation by enhancers:** **Enhancers + activator TFs** can act from far away via **DNA looping**, increasing elongation efficiency.

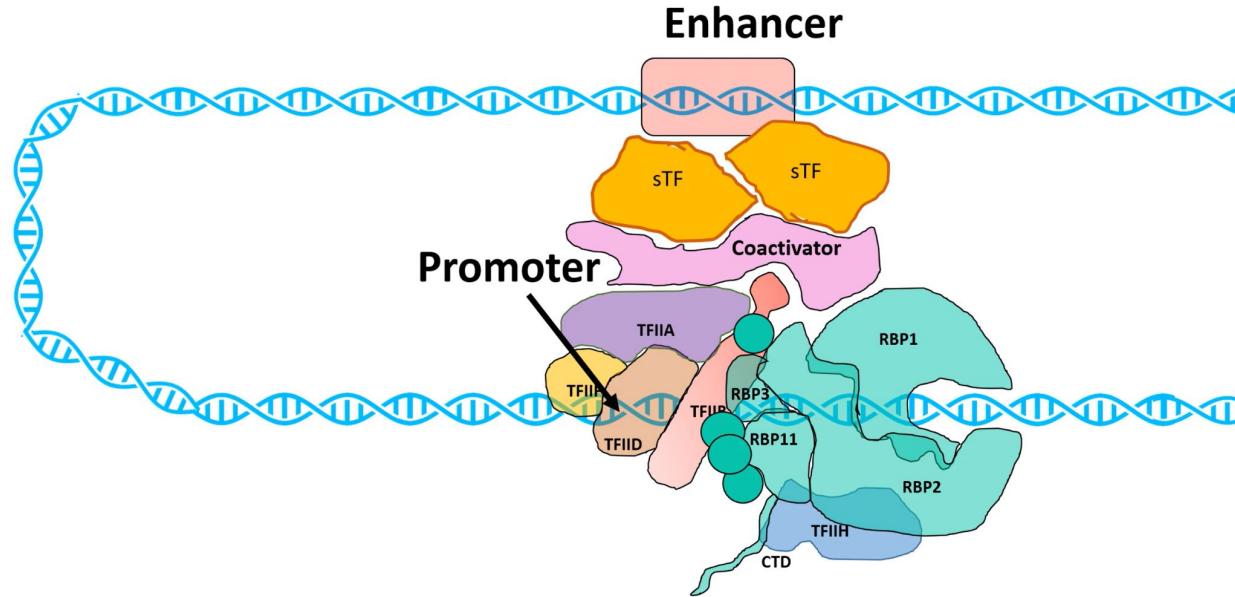


Enhancer: a sequence element that boosts transcription from a target core promoter in a cell type-specific manner and independently of the relative orientation and distance of the enhancer. Known enhancers are several tens to several hundred bp in length.

a short sequence around the transcription start-site (TSS) that can direct the recruitment of RNA polymerase II (Pol II) and transcription initiation. Core promoters typically have low basal activities in the absence of enhancers and are also termed minimal promoters

Transcription regulation in Eukaryotes

- **Enhancer regulation (long-range):** Enhancers can be **kilobases away**; they bind **specific TFs** and, via DNA looping + co-activators, increase promoter escape and boost elongation efficiency/stability.
- Silencer: (similar mechanism)



Termination in Eukaryotes

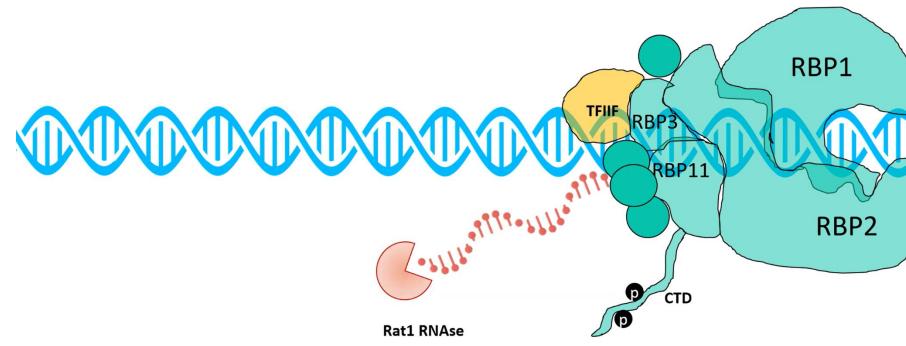
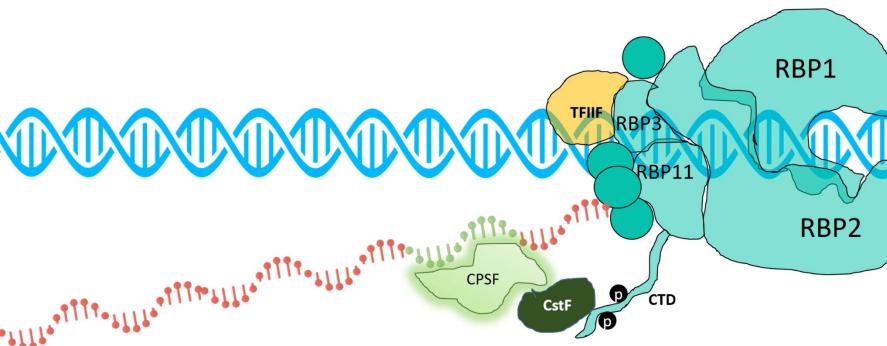
Termination is the final stage of transcription in which termination signals cause:

1. **RNA polymerase pauses/stops RNA synthesis**
2. **The RNA transcript is released** from the transcription complex
3. **RNA polymerase dissociates from the DNA**, and the complex disassembles

Required components

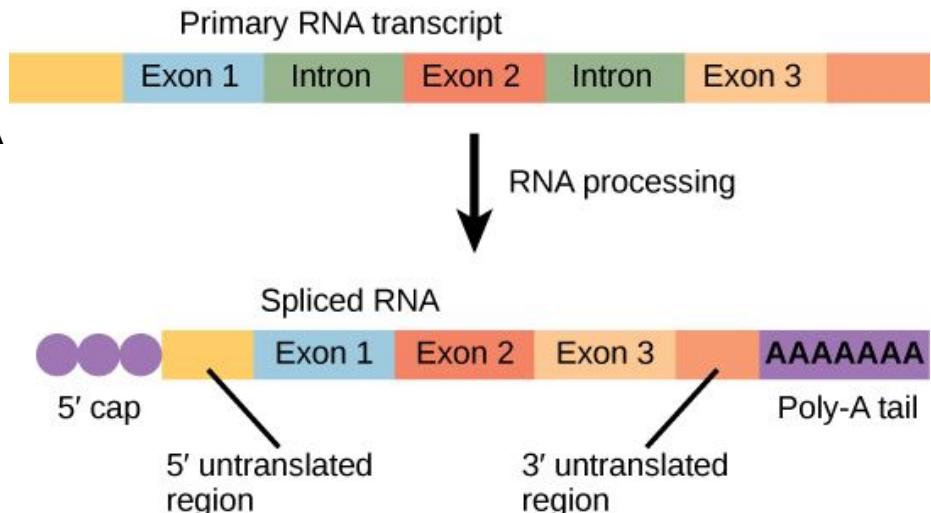
- **Polyadenylation signal (PAS)** in the transcript: **AAUAAA** (or variant) + downstream elements
- **Cleavage and polyadenylation factors** (protein complex)
- Often a **5'→3' exonuclease** for termination after cleavage (conceptually important)

- No clear “termination sequence” like in prokaryotes: RNA polymerase II often continues transcribing past the end of the gene.
- Polyadenylation signal appears in the nascent RNA: typically AAUAAA.
- Cleavage factors recognize the signal: CPSF and CstF bind near AAUAAA and recruit the cleavage machinery.
- mRNA is cleaved downstream of the signal: the pre-mRNA is cut to generate the mature 3' end (then poly(A) addition follows).
- “Torpedo” model: a 5'→3' exonuclease (e.g., Rat1/Xrn2) degrades the remaining RNA and helps dislodge Pol II from DNA.
- Note: alternative models exist, but cleavage at the poly(A) site + exonuclease “chase” is one of the best-supported mechanisms.



Post transcriptional modification

- The RNA made directly from transcription is the **primary transcript**.
- In eukaryotes, **RNA polymerase II** produces **hnRNA (pre-mRNA)**.
- hnRNA is processed in the **nucleus** to become **mature mRNA** (*mRNA processing / post-transcriptional modification*).
- Mature mRNA** is then exported out of the nucleus.
- Major processing steps:
 - 5' capping**
 - 3' poly(A) tailing (polyadenylation)**
 - Splicing (remove introns, join exons)**



Capping and Tailing

1) Capping

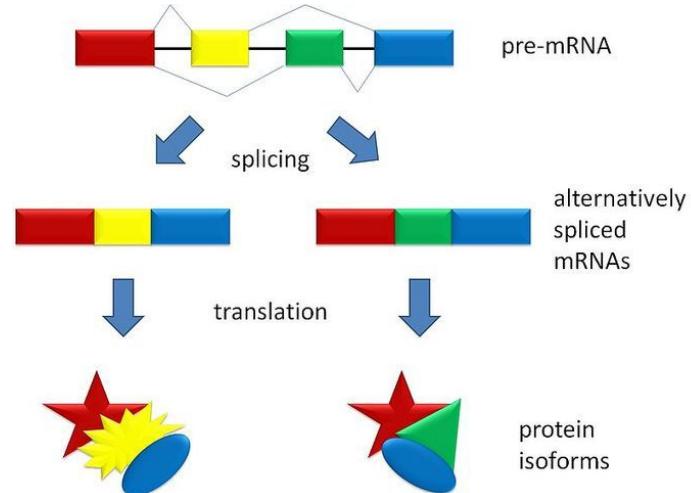
- Capping is the formation of a “protective cap” at the 5’end of pre-mRNA by the addition of 7-methyl guanosine with 5’→ 5’ triphosphate linkage.
- It stabilizes the structure of mRNA.

2) Tailing Polyadenylation

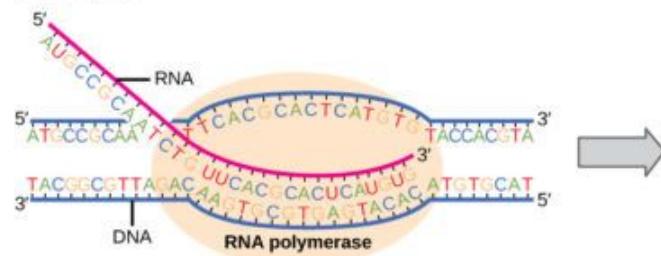
- To the 3’end of pre-mRNA about ~200-250 adenine nucleotides are added with the help of enzyme adenylic acid transferase. This is known as poly(A)tail.
- It protects the mRNA from rapid enzymatic degradation.
- It allows the passage of mRNA through the nuclear pore to the cytoplasm.

Splicing

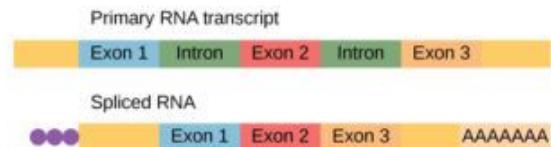
- The primary transcript contains **alternating exons** (coding/expressing) and **introns** (intervening).
- Exons** carry the genetic code for **protein synthesis**; introns are **not translated**.
- RNA splicing** removes introns from the primary transcript to produce mature mRNA.
- Splice junction signals: **5' GU** and **3' AG** define the exon–intron boundaries.
- snRNPs** recognize splice sites to ensure accuracy (**U1** recognizes the **5' splice site**, **U5** recognizes the **3' splice site**).
- snRNPs + hnRNA form the **spliceosome**.
- The splicing reaction is catalyzed by RNA (a **ribozyme**).



Transcription



RNA processing



Translation

