TRANSCRIPTION
IN
PROKARYOTES

TRANSCRIPTION

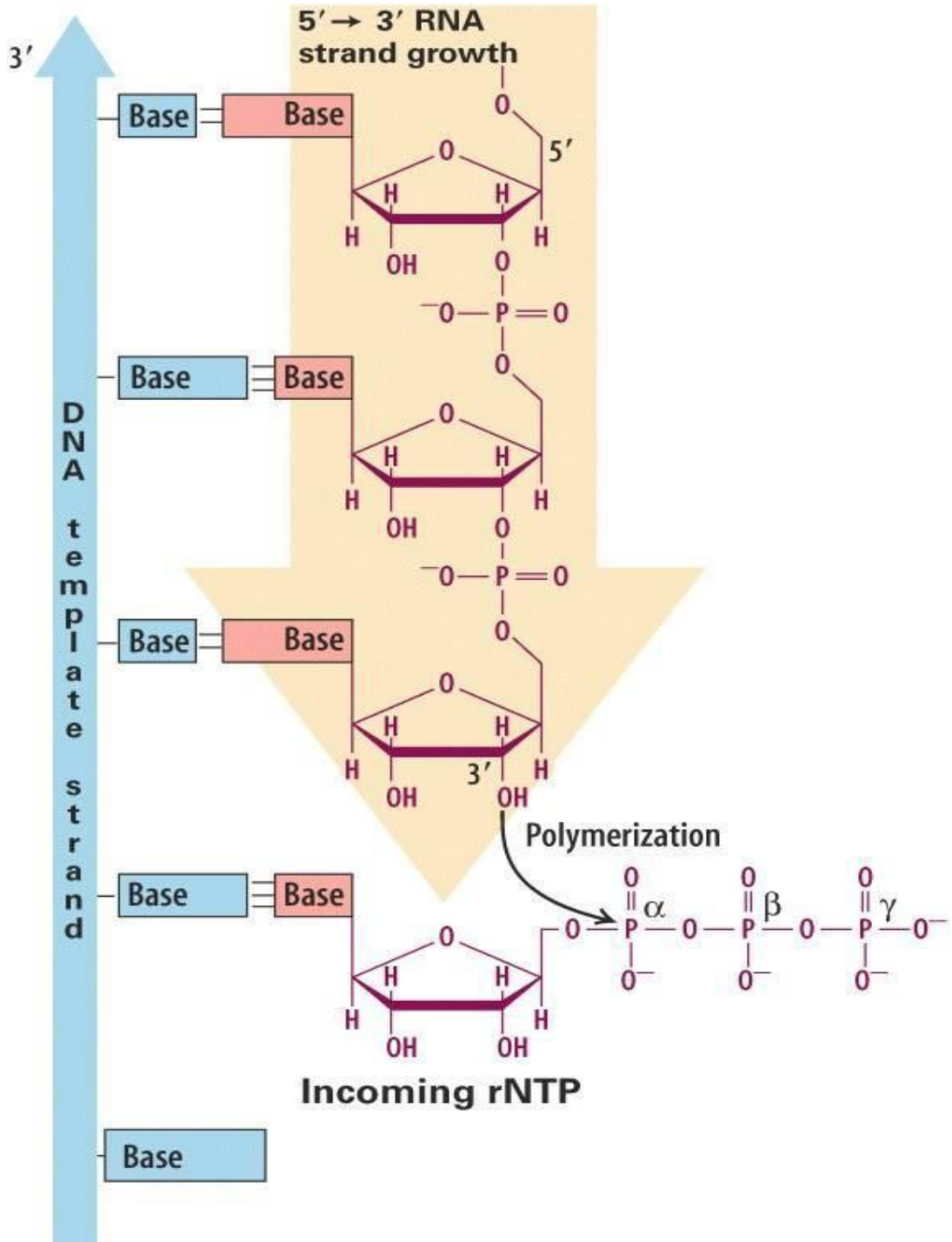
- *"The process of RNA chain initiation, elongation and termination catalyzed by DNA dependent RNA polymerase" is known as Transcription.*
- Transcription is also a polymerization step similar to replication in which Ribo Nucleoside Tri Phosphates are linked to one another forming a single stranded (ss) RNA or polyribonucleotide.

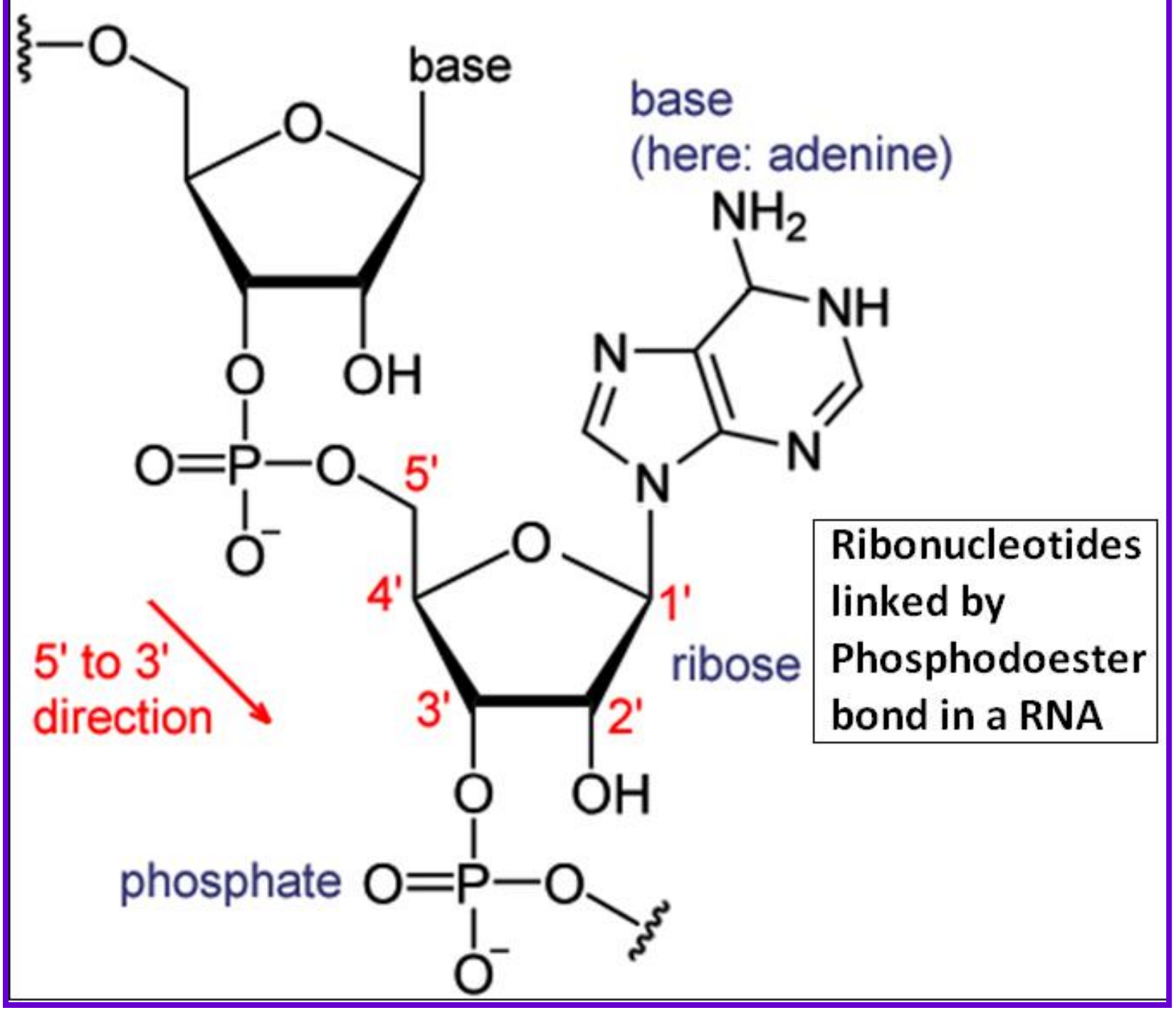
OVERVIEW OF TRANSCRIPTION

- **Requirements**:
  1) Template (ssDNA) - Only one of the 2 strands in DNA code for a protein/ poly peptide at a time and at a place.
  2) RNA polymerase enzyme & Mg^{+2}.
  3) Ribo Nucleoside Tri Phosphates (ATP, CTP, GTP, UTP)
- The overall polymerization of RNA is

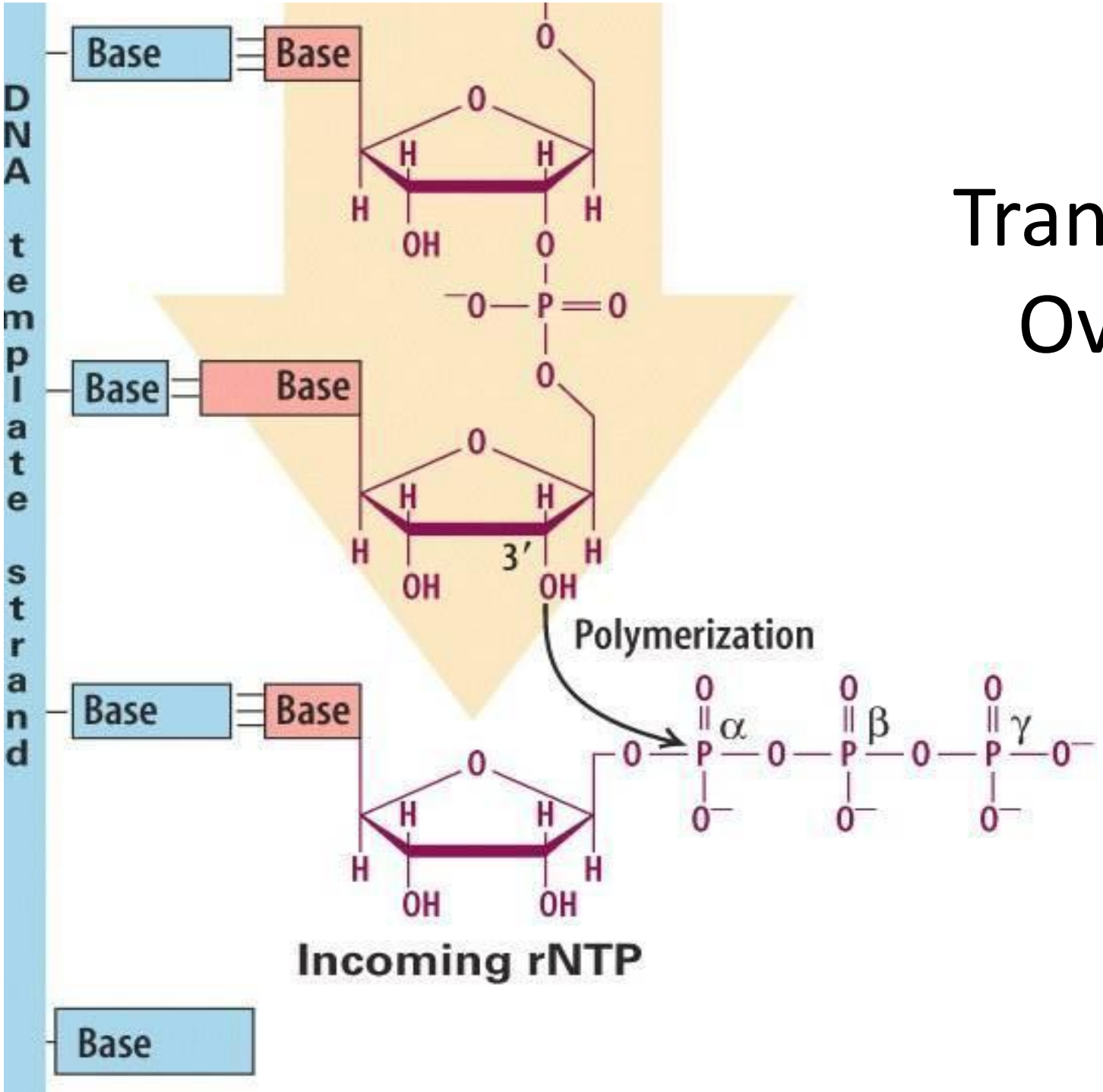
$$nNTP + XTP \rightarrow (NMP)_n\text{-}XTP + nPPi$$

Transcription Overview




base
base
(here: adenine)
NH2
NH
N
N
N
O
OH
O
O=P—O
O⁻
5'
4'
1'
3'
2'
ribose
5' to 3'
direction
phosphate
Ribonucleotides
linked by
Phosphodoester
bond in a RNA

Transcription Overview



OVERVIEW OF TRANSCRIPTION

- RNA polymerases initiate RNA synthesis so no primer is required.
- RNA synthesis also proceeds in 5`→ 3` direction.
- Specificity of base pairing is the basic mechanism for the information transfer.
- Thymine of DNA is replaced by Uracil in RNA.

OVERVIEW OF TRANSCRIPTION

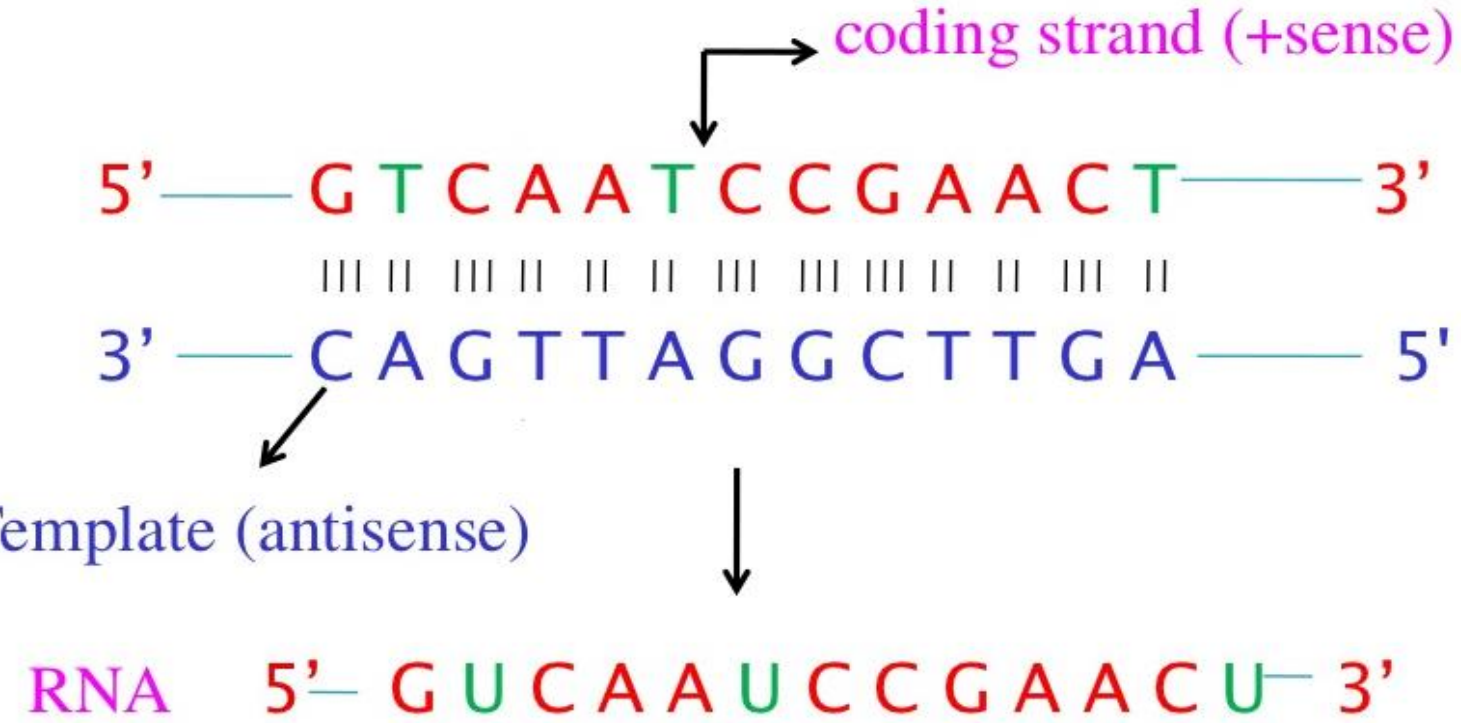
- The strand of DNA which from which mRNA is copied is known as ‘Anti Coding strand’ or ‘Anti Sense strand’ or ‘-ve strand’ or ‘Template strand’.
- The second strand of DNA which resembles RNA except at position of U is known as ‘+ve’ or ‘Coding’ or ‘Sense’ or ‘non Template’ strand.

OVERVIEW OF TRANSCRIPTION

Coding strand	Sense or + strand	5’ GTCAATCCGAACT 3’
Template	Antisense or - strand	3’ CAGTTAGGCTTGA 5
RNA		5’ GUCAAUCCGAACU 3’

OVERVIEW OF TRANSCRIPTION

- The strand of DNA which is transcribed to RNA called as template strand.
- Opposite strand is referred as coding strand.



Differences between replication and transcription

	Replication	Transcription
Template	Both strand whole genome	single strand small portion of genome
Primer	yes	no
Enzyme	DNA polymerase	RNA polymerase
Product	dsDNA	ssRNA
Base pair	A-T, G-C	A-U, T-A, G-C
Proof reading	yes	no

Types of RNAs & Their Functions

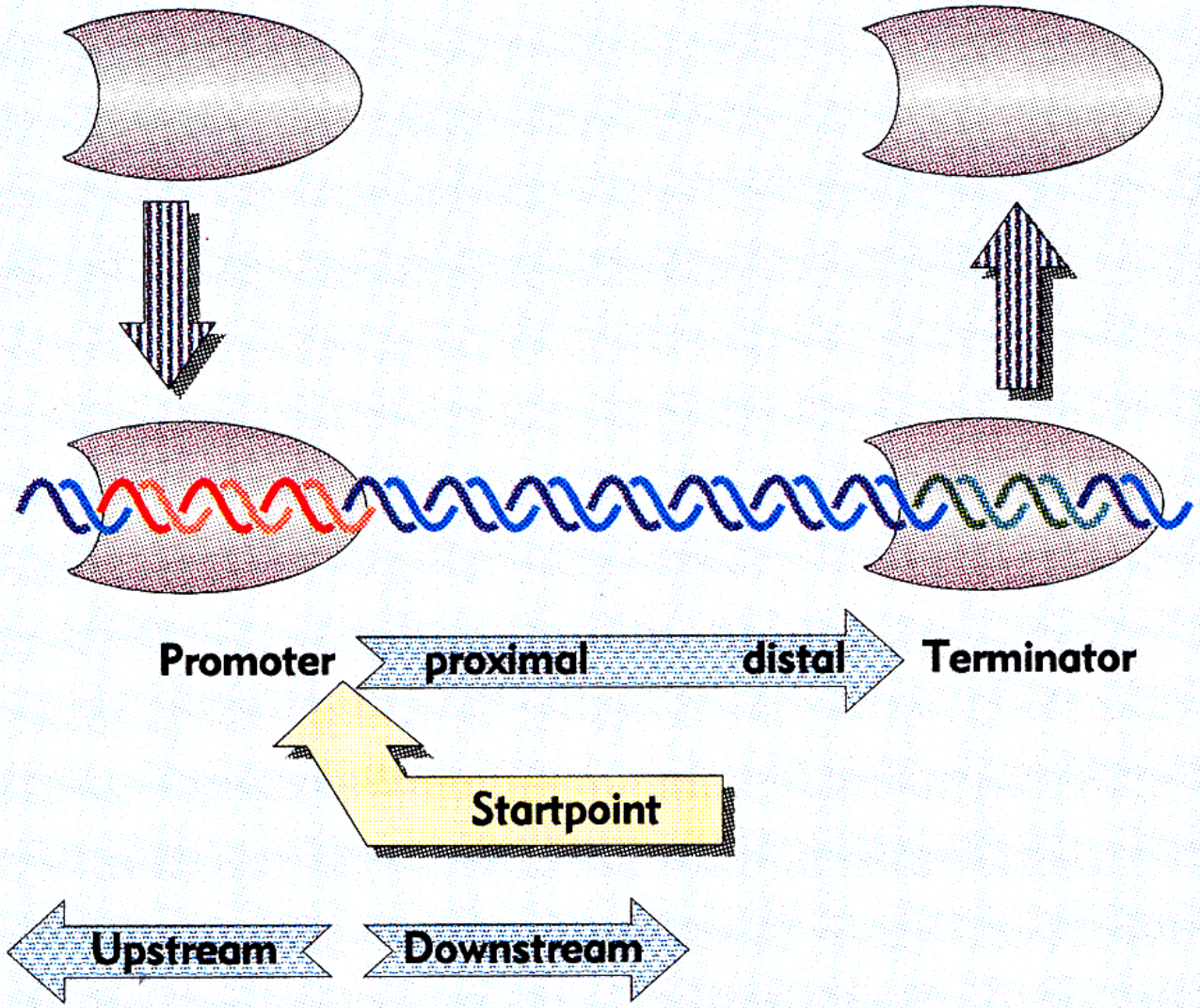
- Generally RNAs consist of single strand polynucleotide chain.
- RNAs are polar molecules, their one end is called as 5’ end and the other end is called as 3’ end.
- Types of RNAs - mRNA, rRNA, tRNA, snRNA, micro RNA.
- These RNAs participate and perform different functions in protein biosynthesis.

Types of RNAs & Their Functions

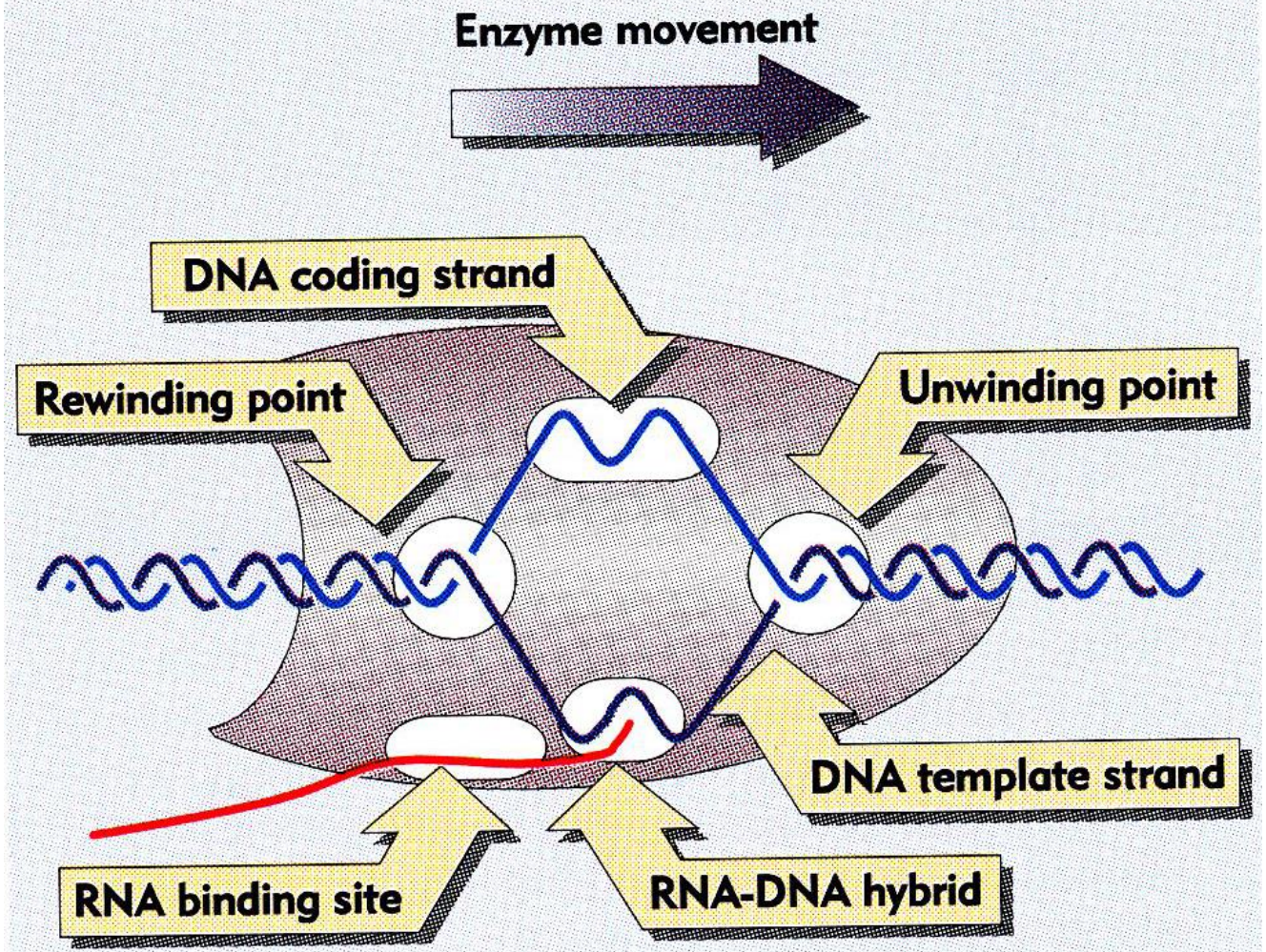
- **2)** Many of the RNA transcripts are synthesized as precursors that is known as primary transcripts.

- Which on modifications & trimming converted into functional RNA .
- **SITE:**
- **Transcription –** Prokaryotes– cytoplasm(all RNAs).
  Eukaryotes– Nucleus & mitochondria
  a) Nucleolus – rRNA
  b) Nucleoplasm –tRNA & mRNA.

Overview: a transcription unit is a sequence of DNA transcribed into a single RNA, starting at the promoter and ending at the terminator.



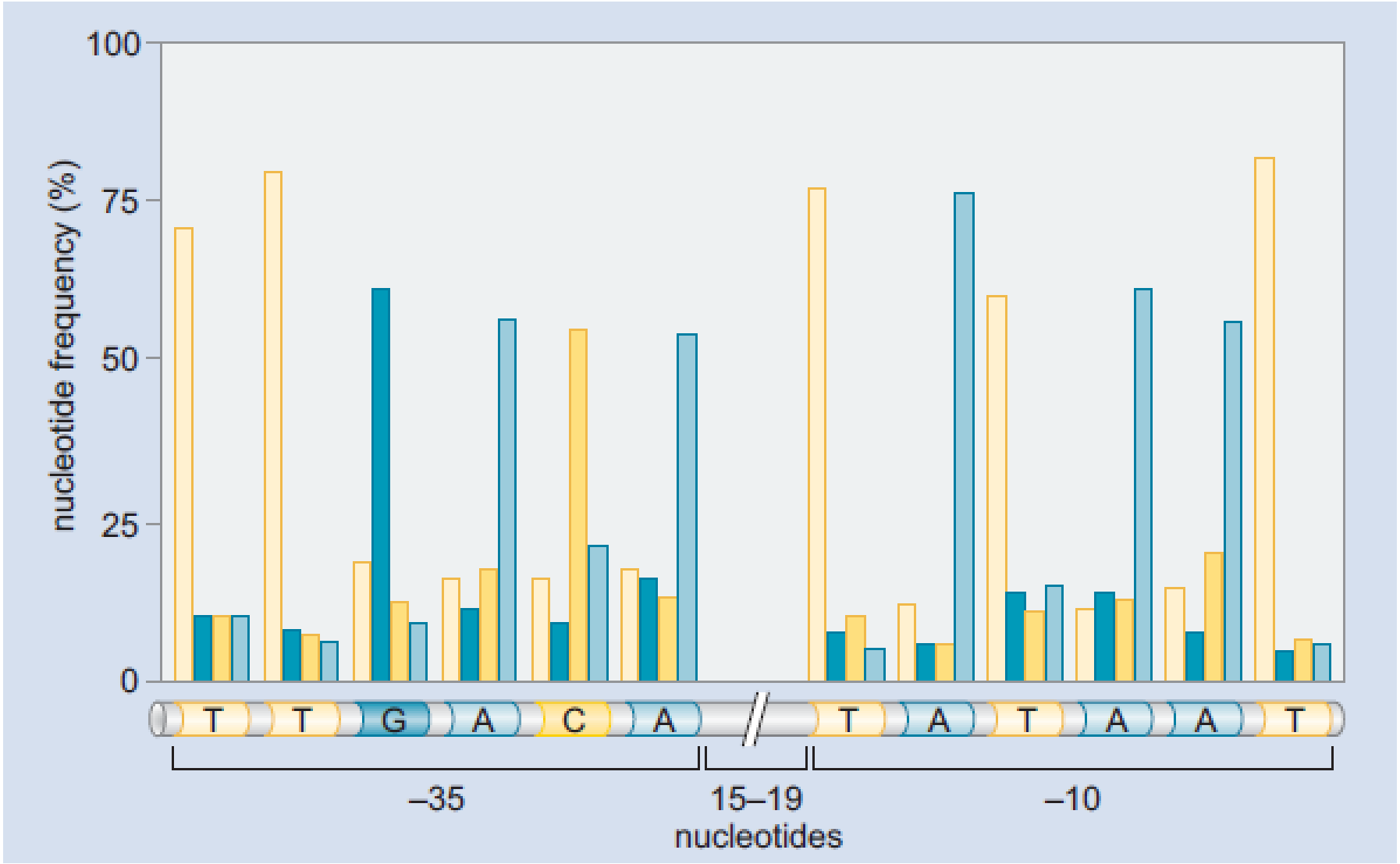
During transcription, the bubble is maintained within bacterial RNA polymerase, which unwinds and rewinds DNA, maintains the conditions of the partner and template DNA strands, and synthesizes RNA.



- **Consensus Sequence**: *A basic sequence derived from a large set of observed similar sequences is called consensus sequence.* It is obtained by comparing a large number of sequences from a particular region.
- Promoter is a DNA sequence where the RNA polymerase binds and transcribes the gene adjacent to it. It is an example for a *consensus* sequence.
- Various bacterial promoters were analyzed & two consensus sequences have been framed - (1) *Pribnow box* or -10 region (2) -35 region

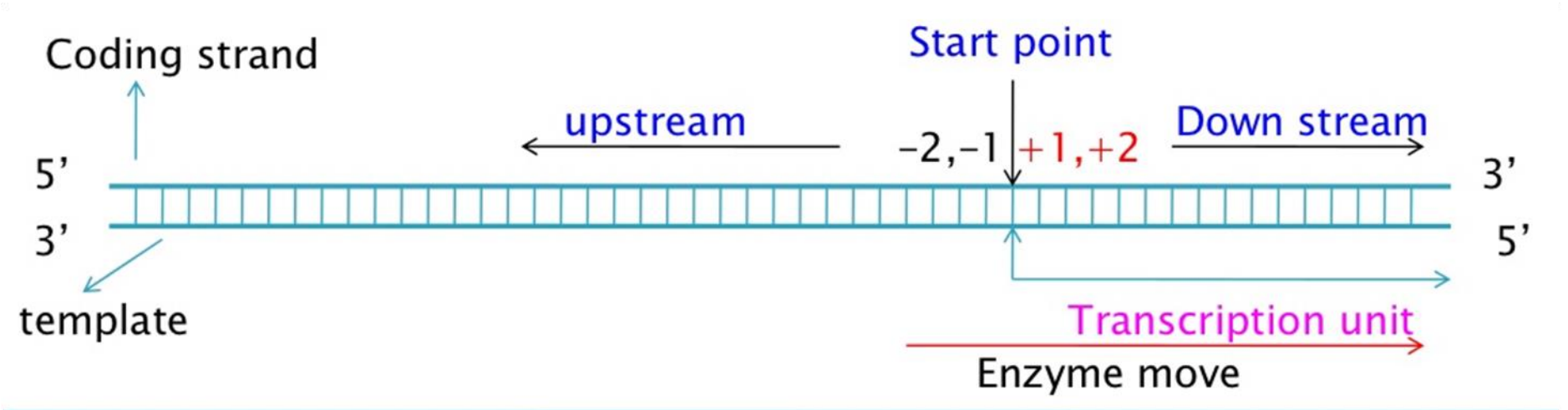
- *Pribnow* box has a consensus sequence of **TATAAT** where as -35 region has **TTGACA**.
- However none of the naturally occurring promoters have complete homology with these consensus sequences and **greater** the **homology** greater is the **strength** of the promoter.

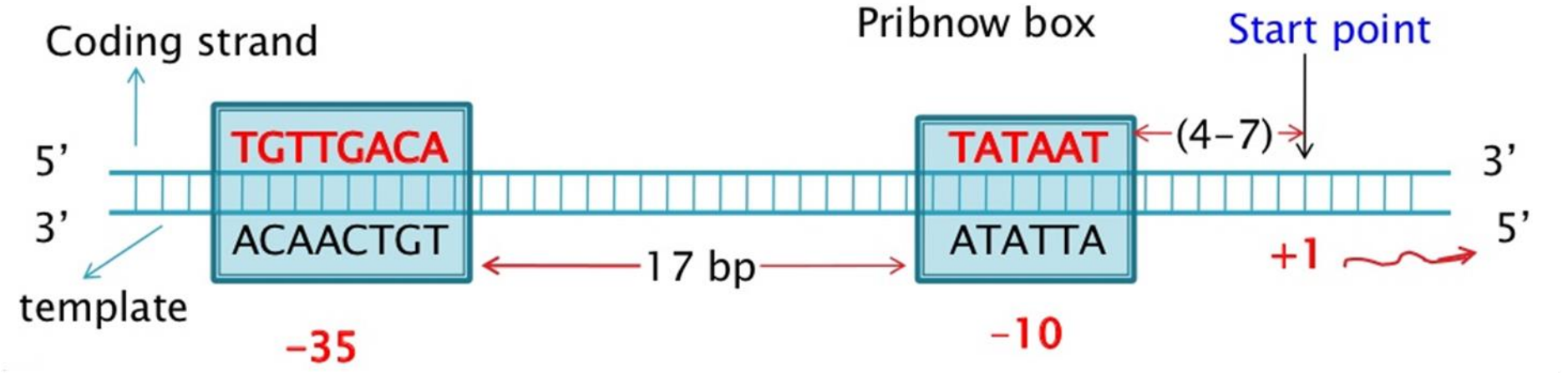
- **NOTE:** The first base to be transcribed is selected as a reference point and is numbered as +1.
- All the bases that follow the first base to be transcribed are called **downstream bases** and are numbered +2, +3, and so on.
- The bases which are not transcribed and are present on the other side (left side) are called **upstream bases** and are numbered -1, -2, -3 and so on.



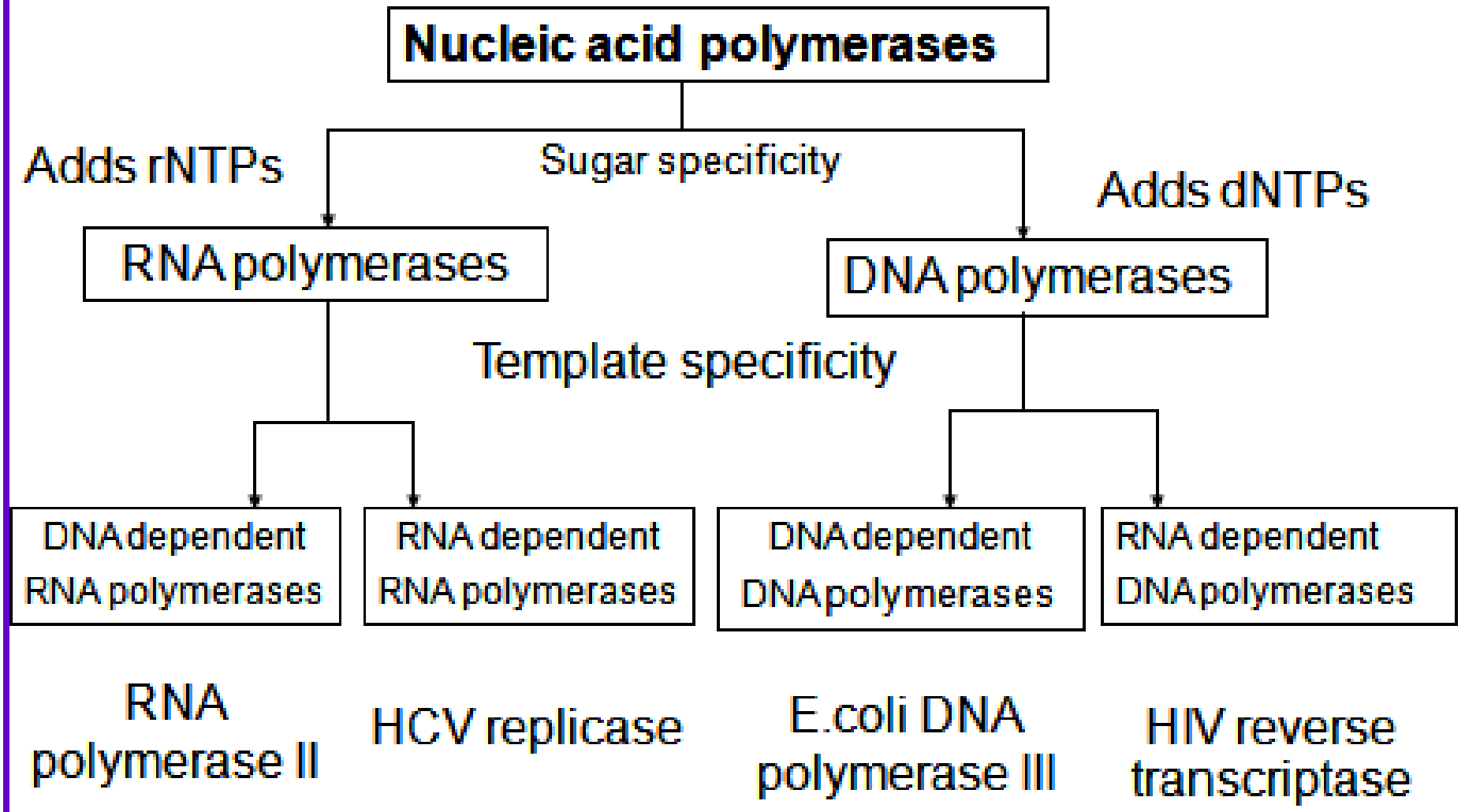
Promoter consensus sequence and spacing consensus.

OVERVIEW OF TRANSCRIPTION





Nucleic acid polymerases: classification



E. coli RNA Polymerase

- *E. coli* RNA Polymerase is a very large enzyme multiple subunits and a molecular weight of 465,000 Daltons.
- This holoenzyme has 6 subunits viz., 2α subunits, 1 each of β, β′, ω and σ subunits. Holoenzyme is required for initiation of transcription.
- σ factor is released from the transcription complex after initiation.

E. coli RNA Polymerase

- Core enzyme lacks σ subunit and catalyzes elongation.
- The core enzyme has $\alpha_2\beta\beta'\omega$ subunits.
- The *E. coli* RNA polymerase requires Mg^{++} for its activity.
- The whole enzyme can bind to 60 base pairs of DNA.

E. coli RNA Polymerase

core

β'

β

ω

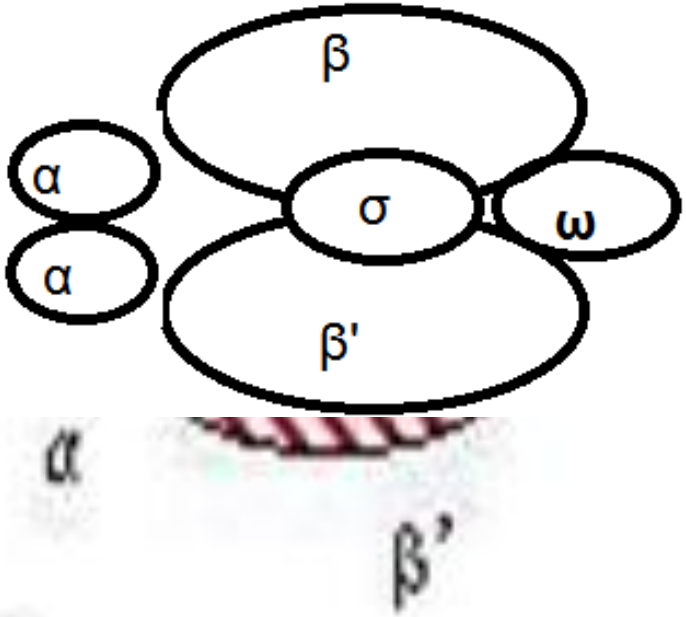
α

Large Subunits

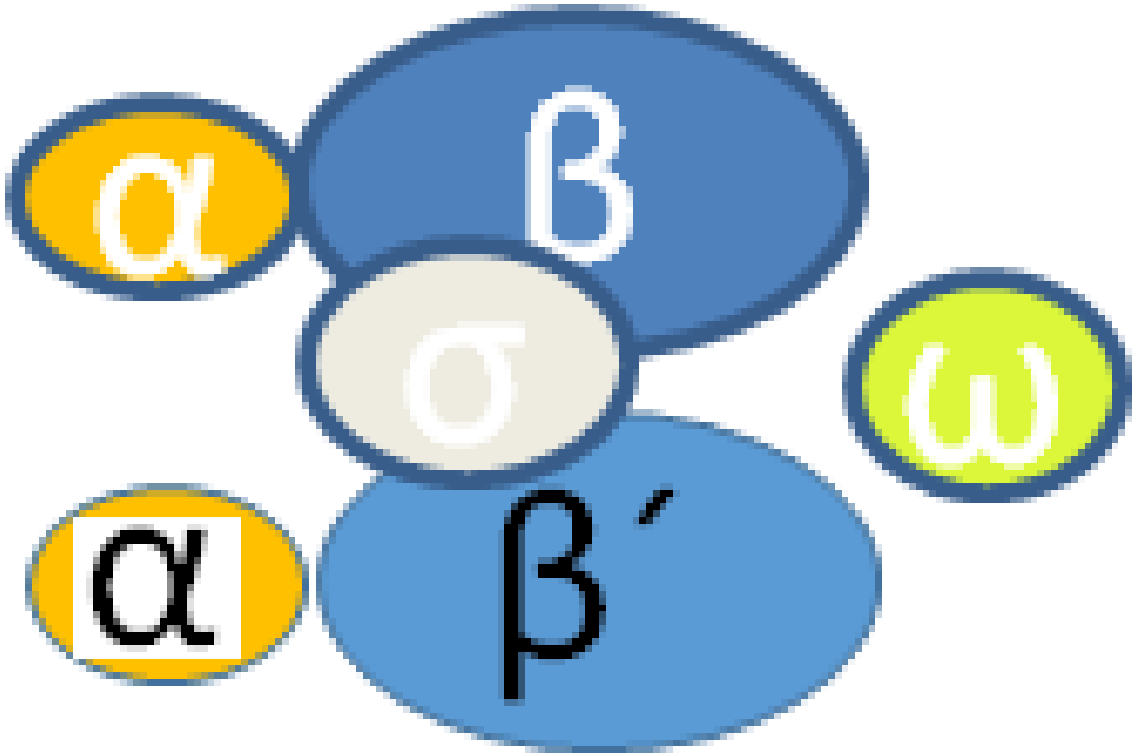
Small Subunit

σ

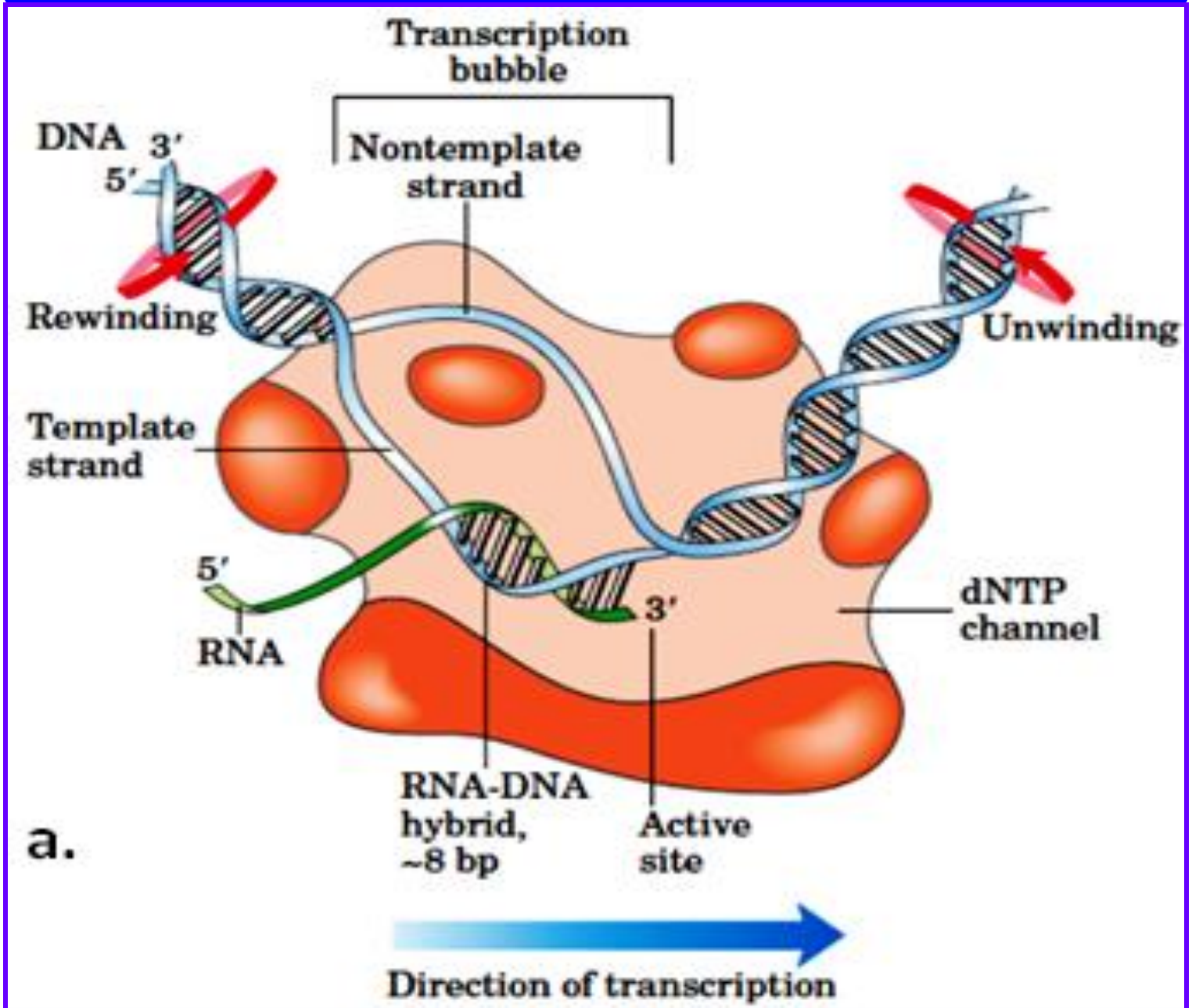
Sigma Factors (various)



E. coli RNA Polymerase



E. coli RNA Polymerase



RNA Polymerase α subunit

- **α** subunit is coded by rpoA gene.
- Two (2) identical α subunits are present in core RNA polymerase.
- **α** subunit binds to DNA in the upstream of -35 region in some promoters.
- **α** subunit may play a role in recognizing the promoter.
- **α** subunit is required for the assembling the core protein.

RNA Polymerase β subunit

- rpoB gene codes for **β** subunit.
- The core enzyme contains one **β** subunit which is a part of catalytic centre of RNA polymerase.
- **β** subunit is useful in both initiation and elongation.
- Rifampicin, (antibiotic) which inhibits initiation of transcription (but not elongation or termination) binds to **β** subunit.
- Streptolydigins, (antibiotic) which inhibits elongation also binds to **β** subunit.

RNA Polymerase β’ subunit

- **β' subunit** is coded by rpoC gene.
- **β'** subunit binds to 2 Zn^{++} ions involved in the catalytic function of **RNA Polymerase**.
- The function of **β'** subunit is to bind to DNA which carries a net –ve charge.
- Heparin, a polyamine binds to **β'** subunit and inhibits transcription by competing with DNA in binding to the RNA polymerase.

RNA Polymerase ω subunit

- Coded by rpoZ gene.
- Restores the fully functional form of denatured RNA polymerase in vitro.

RNA Polymerase σ subunit

Sigma factor	Gene	Function
σ^{70}	*rpoD*	principal sigma factor
σ^{54}	*rpoN* (*ntrA, glnF*)	nitrogen-regulated gene transcription
σ^{32}	*rpoH*	heat-shock gene transcription
σ^{S}	*rpoS*	gene expression in stationary phase cells
σ^{F}	*rpoF*	expression of flagellar operons
σ^{E}	*rpoE*	involved in heat shock and oxidative stress responses; regulates expression of extracytoplasmic proteins
σ^{FecI}	*fecI*	regulates the *fec* genes for iron dictate transport

Anti-sigma factors

- The importance of anti-sigma factors has been established in recent years. These factors form complexes with their **cognate sigma factor**, thereby inhibiting its function.
  1. **FlgM** is an anti-sigma factor for flagellar sigma factor s^F.
  2. **Rsd** is an anti-s^{70} factor synthesized only during stationary phase.
- **Rsd** blocks the activity of s^{70} & allows s^S to associate with the core RNA polymerase and direct expression of stationary phase genes.

Transcription Process

- Transcription may be divided into four stages
  1) Binding of RNA polymerase to DNA
  2) Initiation
  3) Elongation and
  4) Termination.

Binding of RNA polymerase to DNA

- The RNA polymerase binds to the DNA and slides along the DNA till it encounters the -35 region of the promoter. (The σ factor recognizes this region).
- The enzyme then binds to this region and forms a complex called *Closed Promoter Complex* or **CPC**.

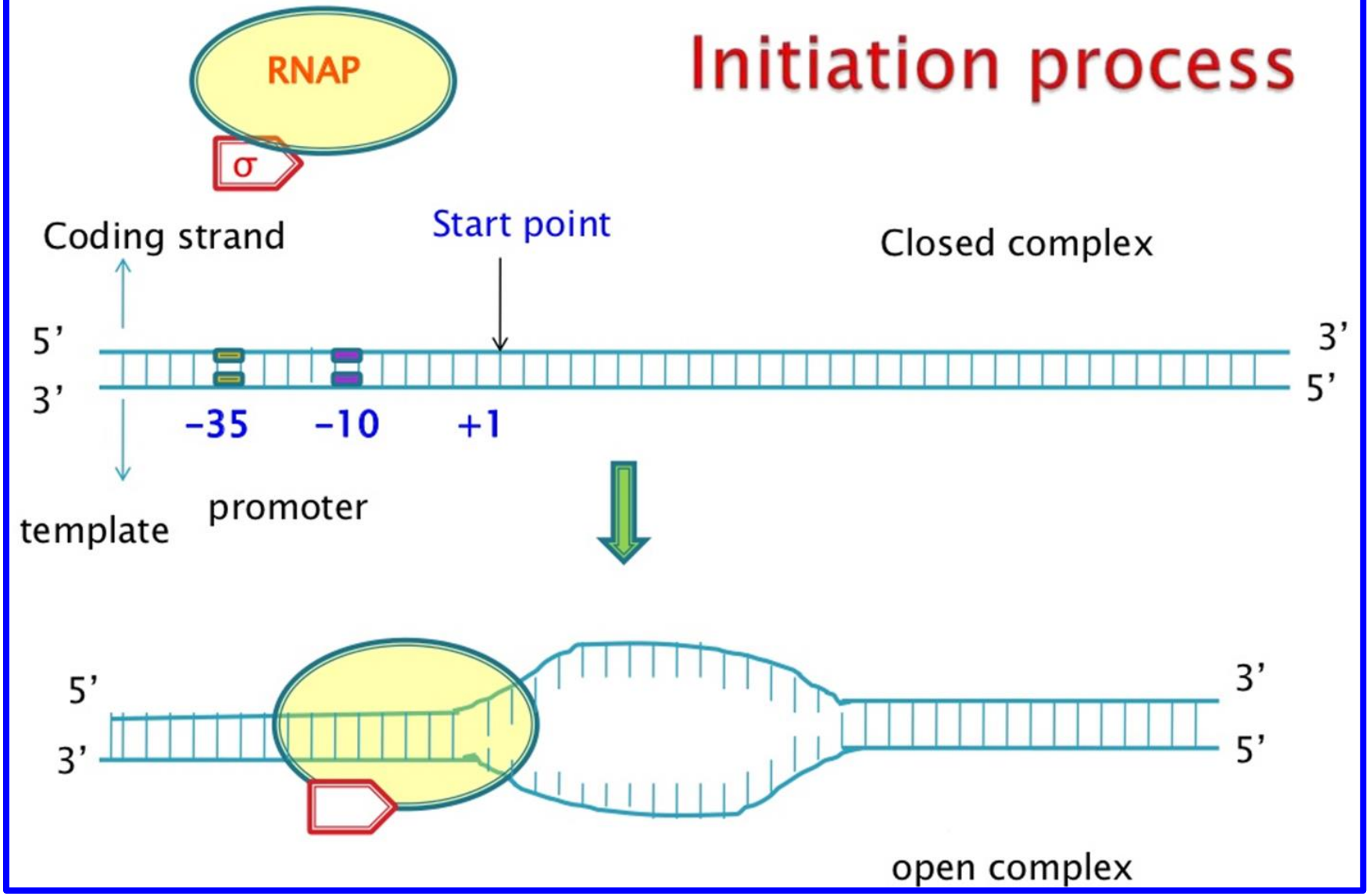
Binding of RNA polymerase to DNA

- Then the σ factor converts this closed promoter complex into an *Open Promoter Complex* or OPC in which the hydrogen bonds that link up the complementary bases are broken down thereby yielding ssDNA (template) required for transcription.
- DNA unwinding starts from -10 region and extends up to a few bases in the downstream. This leads to the formation of Transcription Bubble.

Binding of RNA polymerase to DNA

- Transcription bubble is a dynamic, transiently melted structure representing the site of transcription.
- In **transcription bubble** roughly 12 RNA bases are paired with DNA.
- Approximately 18 bases of the coding strand (+) remain **single stranded**.

Binding of RNA polymerase to DNA



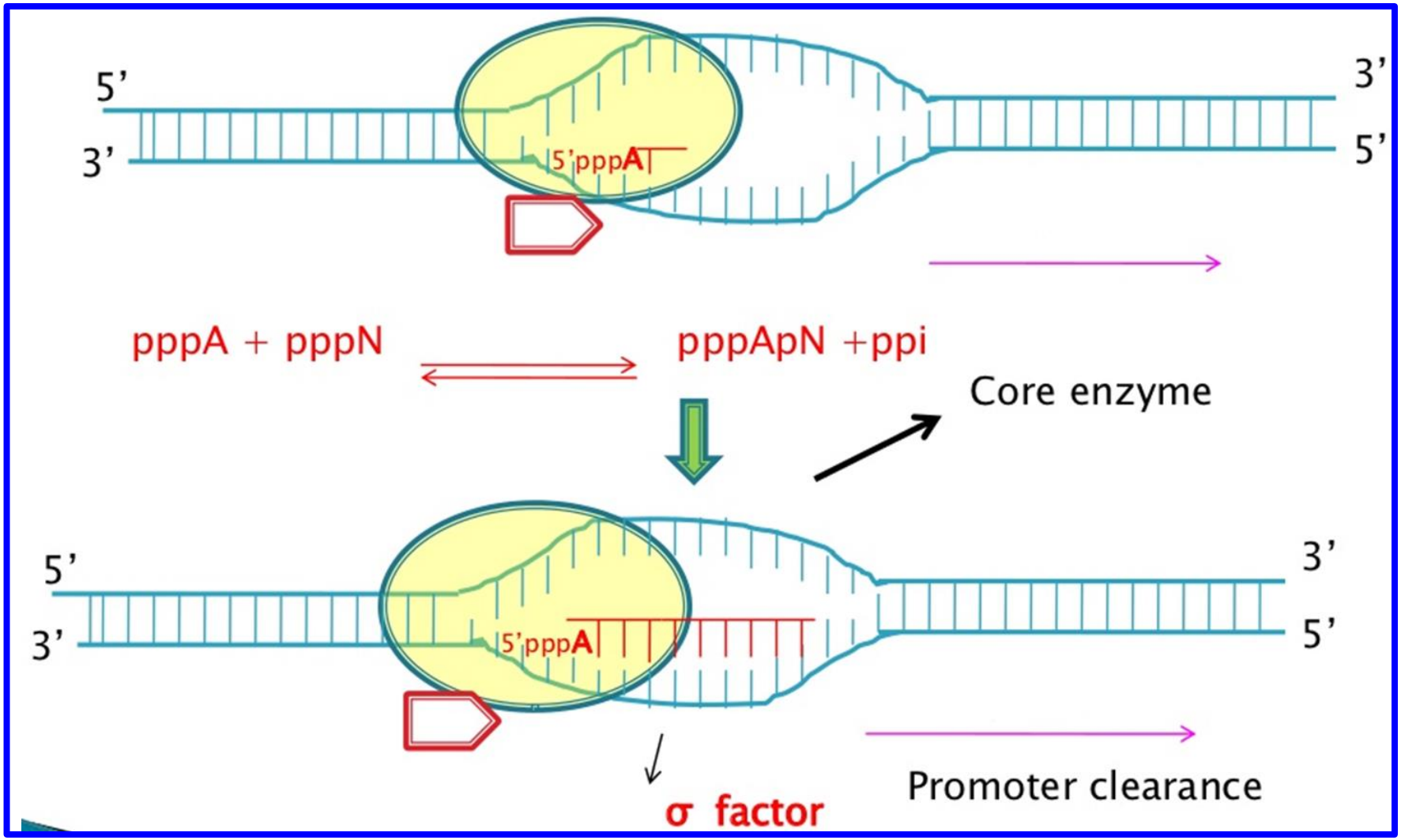
Initiation

- RNA polymerase contains 2 nucleotide binding sites called initiation site and the elongation or catalytic site.
- Usually a ribo purine triphosphate binds to the initiation site and hence the first base of the DNA that is transcribed is a pyrimidine.

Initiation

- Initially the first NTP binds to the enzyme in the open promoter complex and forms hydrogen a bond with complementary DNA base.
- Then the elongation site is filled with a nucleoside triphosphate that is selected by its ability to hydrogen bond with the next base in the DNA strand.
- The two nucleotides are then joined together by phosphodiester bond and the first base is released from the initiation site.

Initiation


5'
3'
5'pppA
3'
5'
pppA + pppN
pppApN +ppi
Core enzyme
5'
3'
5'pppA
3'
5'
σ factor
Promoter clearance

Initiation

- The enzyme (RNA polymerase) moves along the template such that the two binding sites are exactly shifted by one nucleotide.
- The dinucleotide remains hydrogen bonded to the DNA. The next nucleoside triphosphate complementary to the DNA nitrogen base comes and binds in the elongation site and then forms hydrogen bonds with the DNA base.
- Then phosphodiester bond is established and the enzyme moves a distance of exactly one nucleotide so that the elongation site is emptied and the initiation site contains the newly added nucleotide linked to the dinculeotide already formed.

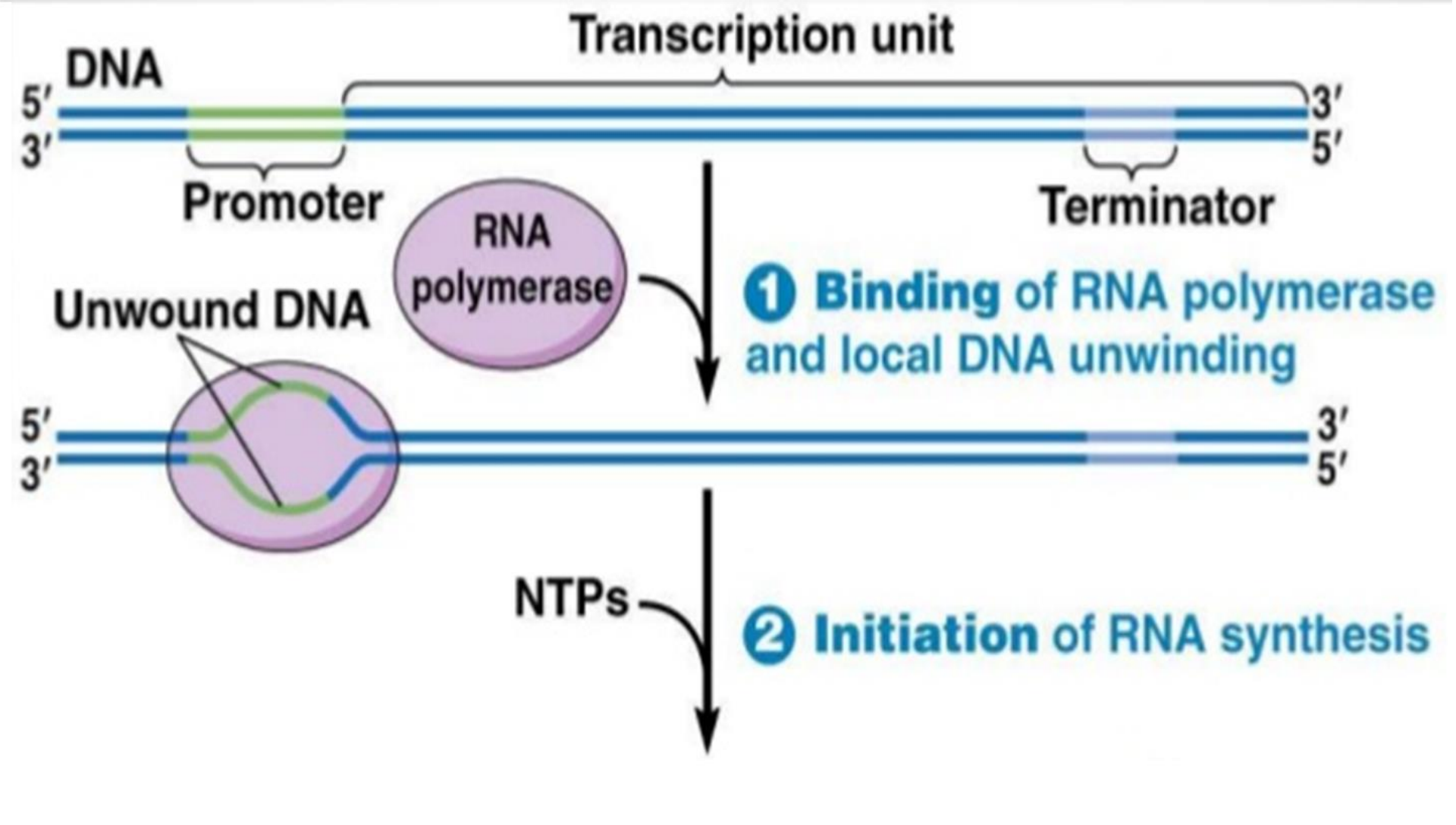
Initiation

- However the initiation is not complete until the polymerization of first 6-10 nucleotides is completed.
- Usually during initiation stage after synthesizing an oligonucleotide (<10 nucleotides long) the RNA polymerase terminates the RNA synthesis prematurely and starts the initiation afresh.
- Such initiations which terminate prematurely are known as *abortive initiations.* This is a regulatory mechanism.

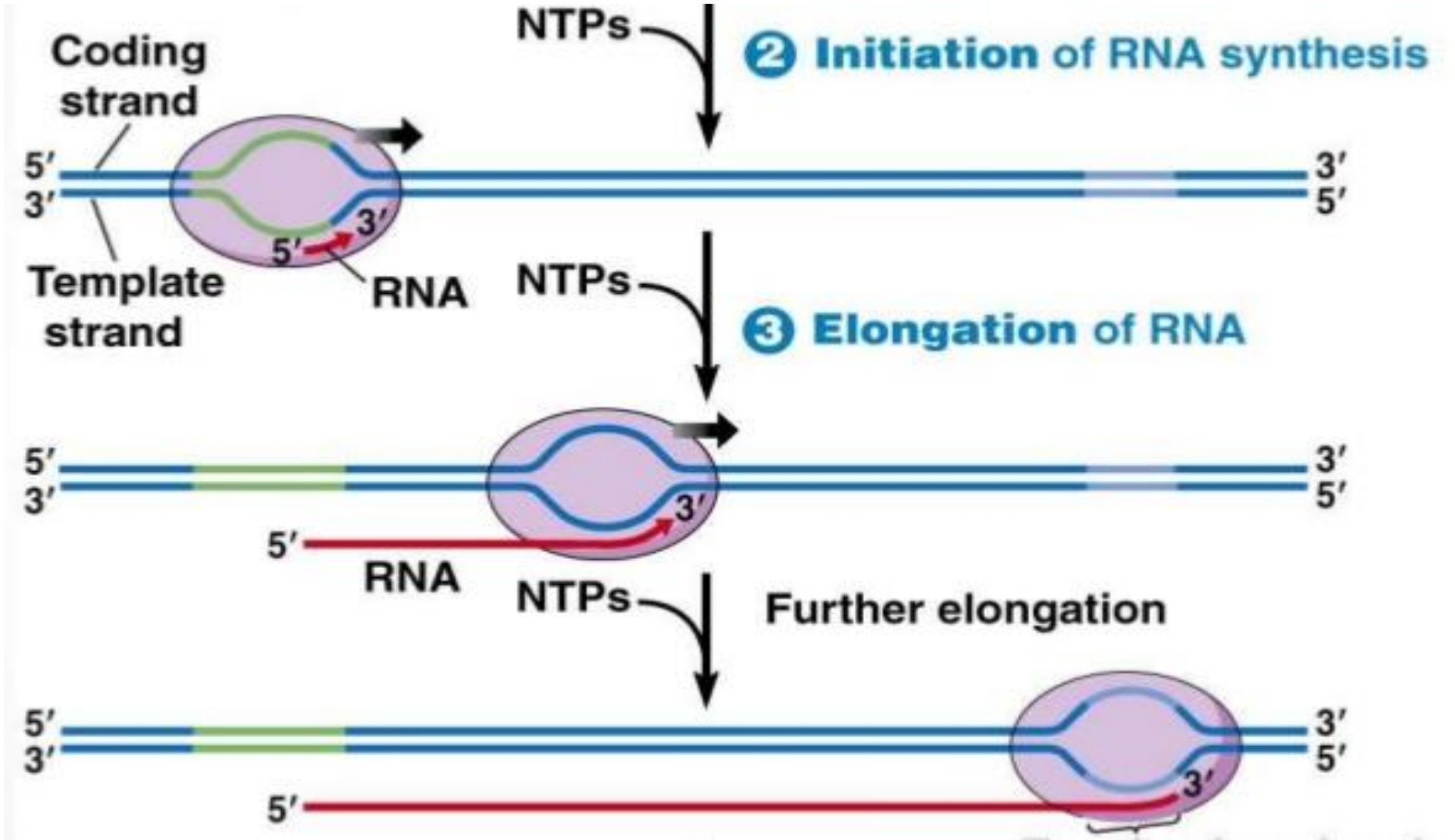
Initiation

- In an unknown way finally the RNA polymerase becomes locked into forward motion and the elongation phase starts. (**Promoter Clearance**).
- After some nucleotides are added (8 in most cases) to the growing RNA chain the RNA polymerase holoenzyme undergoes a conformational change, looses the σ subunit and forms core enzyme that catalyzes elongation .

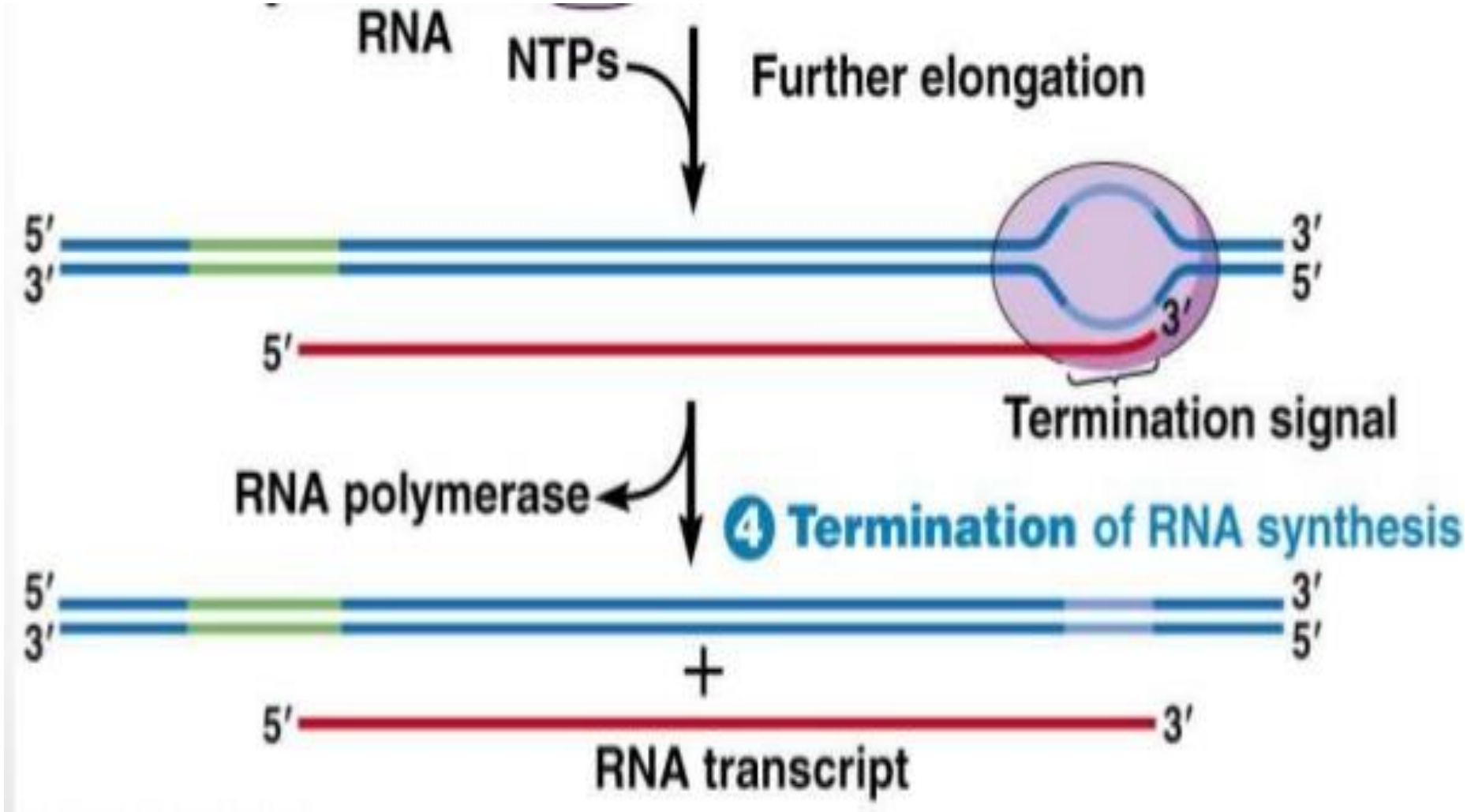
Transcription



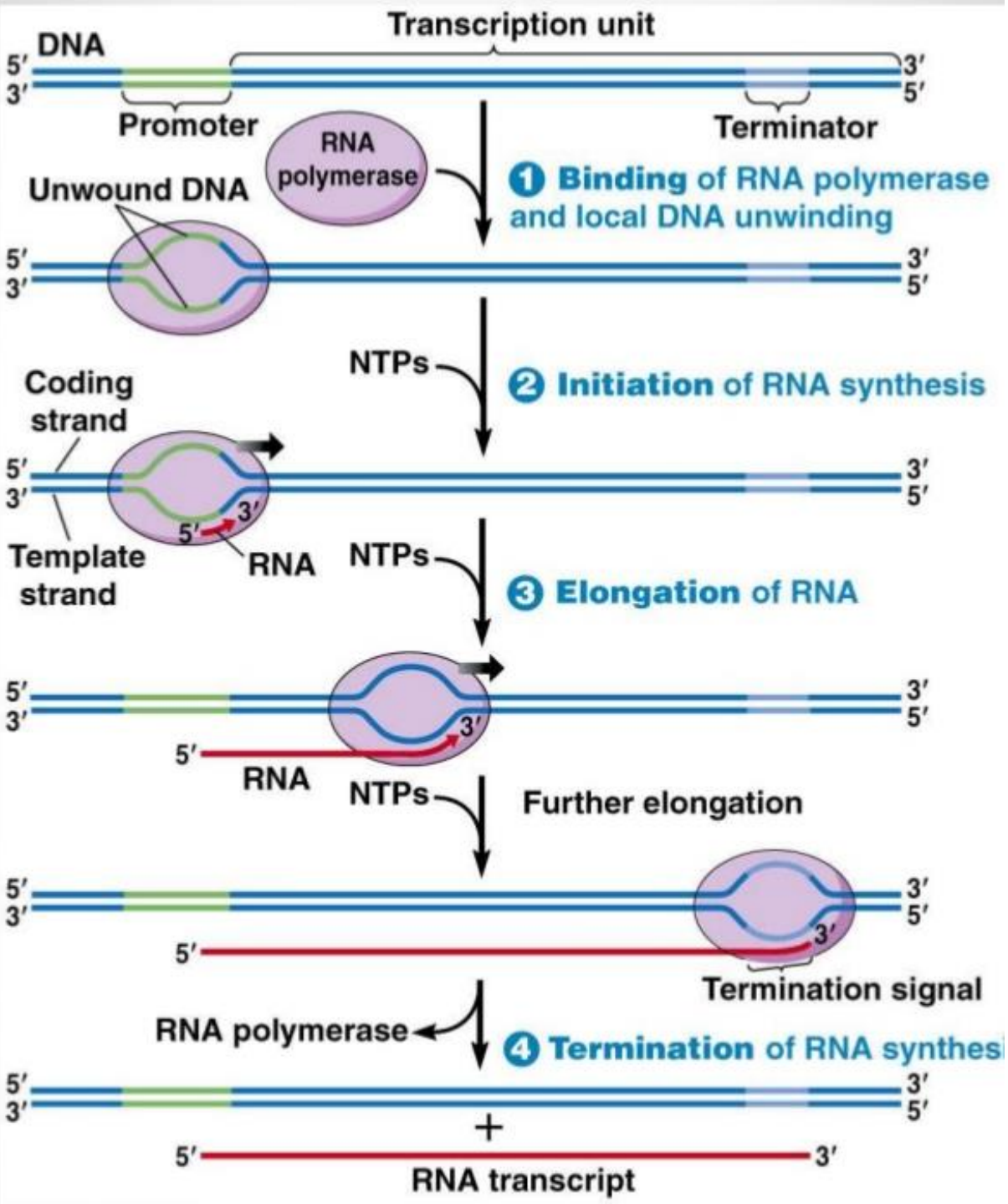
Transcription



Transcription



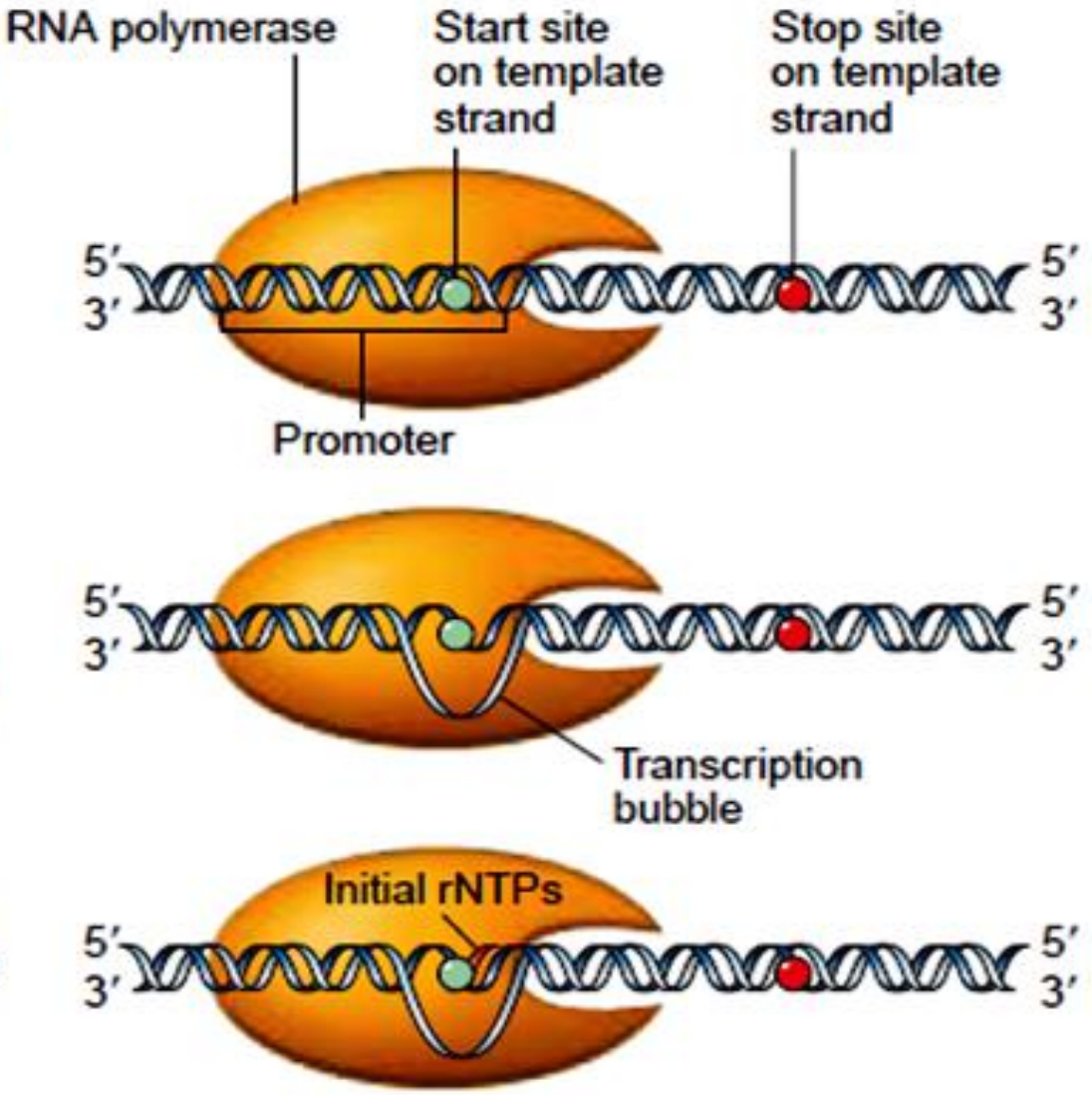
Transcription



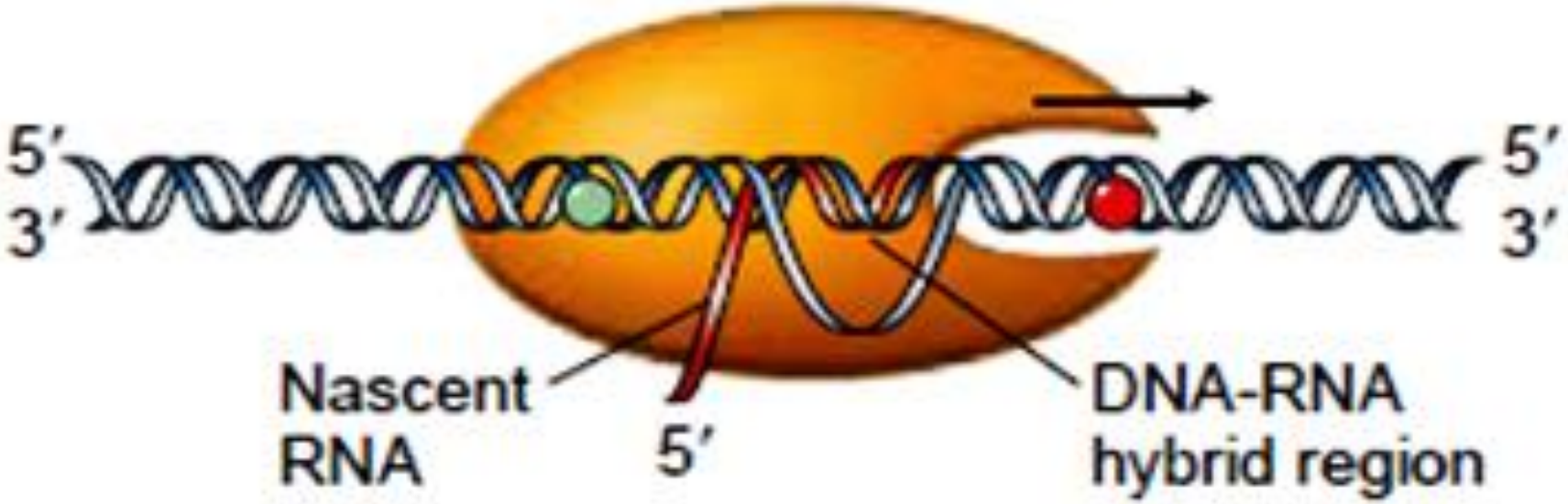
Binding of RNA polymerase to DNA

Binding of RNA pol to DNA & Initiation

1. Polymerase binds to promoter sequence in duplex DNA. "Closed complex"
2. Polymerase melts duplex DNA near transcription start site, forming a transcription bubble. "Open complex"
3. Polymerase catalyzes phosphodiester linkage of two initial rNTPs.

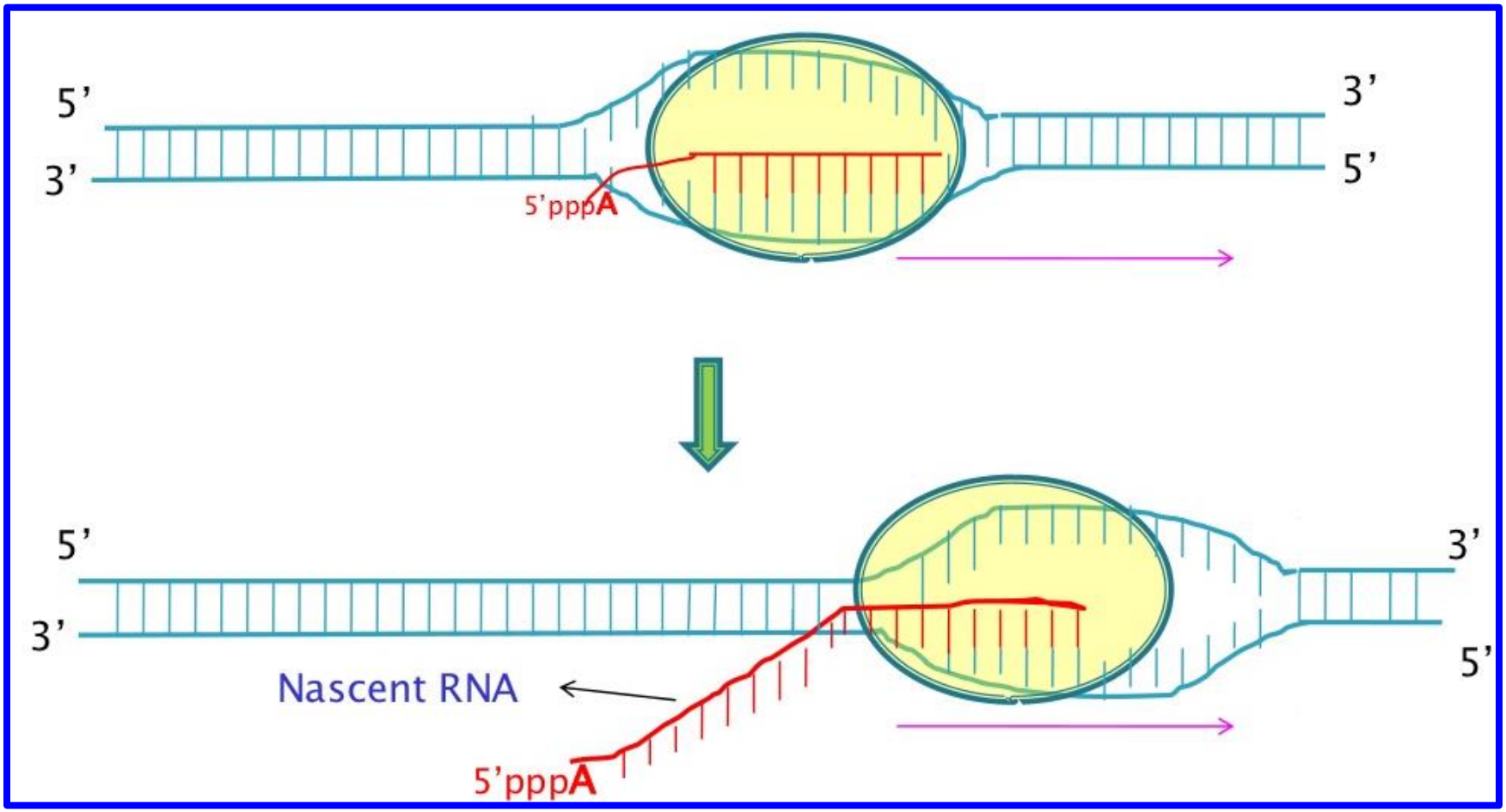


Elongation



Elongation: Polymerase advances 3’ ⟶ 5’ down template strand, melting duplex DNA and adding rNTPs to growing RNA

Elongation



Elongation: Polymerase advances 3'⟶ 5' down template strand, melting duplex DNA and adding rNTPs to growing RNA

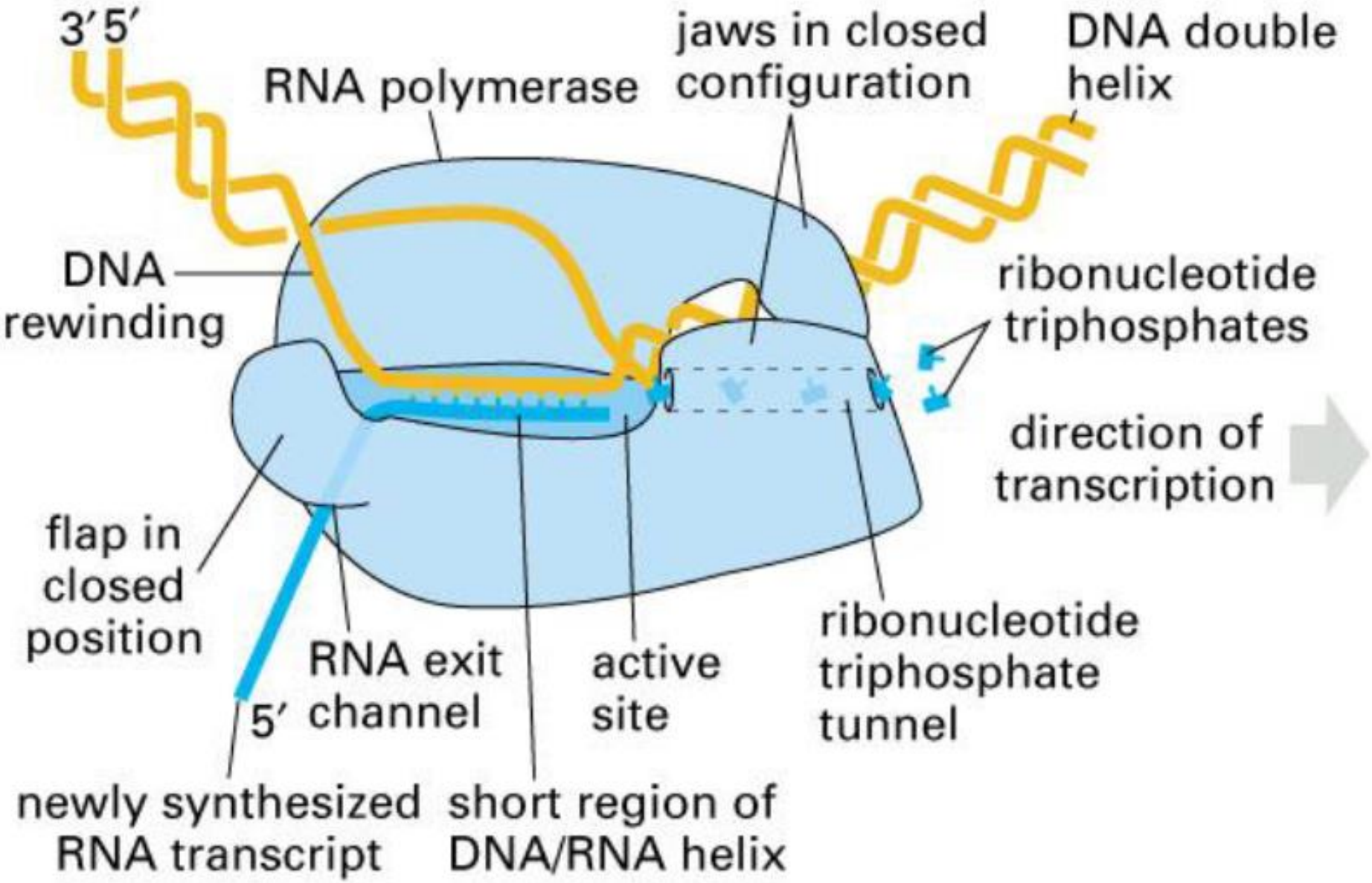
Elongation

- The core enzyme moves along the DNA, during which it performs.
  1) Opening of DNA helix as it moves along DNA.
  2) A nucleoside triphosphate that can pair with the next DNA base comes and binds to core enzyme.

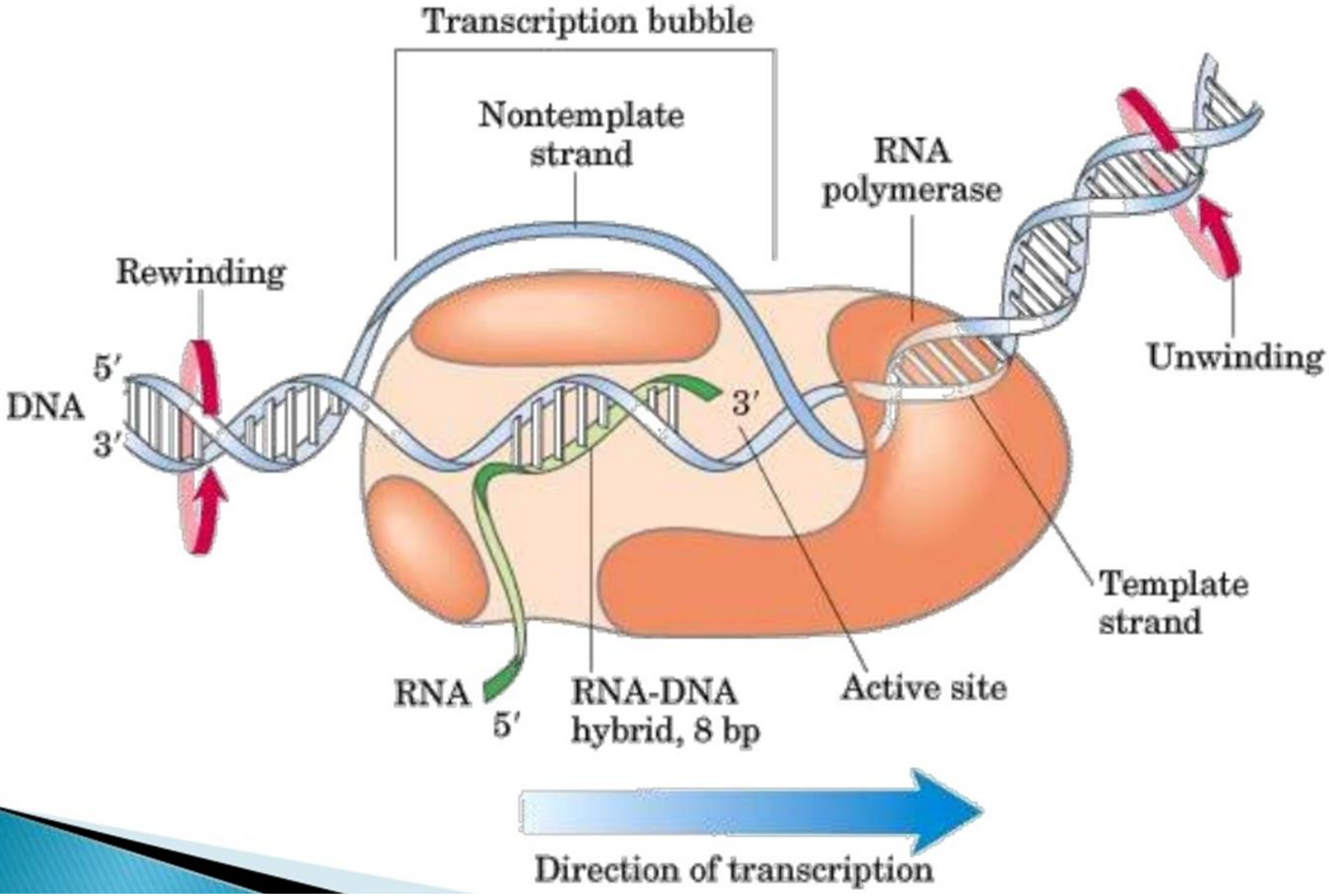
Elongation

- As the enzyme moves away, the newly synthesized RNA is released from the DNA breaking the hydrogen bonds with the bases of DNA and the re-closure of ssDNA to form a dsDNA.
- Roughly 12 RNA bases are paired to DNA in the open region which comprises of a transcription bubble.
- Chain elongation of RNA does not occur at a constant rate the synthesis markedly slow down at some regions of DNA, accelerates again, then continues in the normal rate, slows down and so forth. This reduction in the rate of RNA synthesis is called a *Pause.*

Model of elongating RNA polymerase



Transcription bubble



Termination of newly synthesized RNA

- Termination of RNA synthesis occurs at specific base sequences within the DNA molecules.
- These sequences are of 2 types, (1) simple terminators or **Rho independent** mechanism and (2) those that require auxiliary termination factors or **Rho dependent** mechanism.

Rho independent Termination of RNA

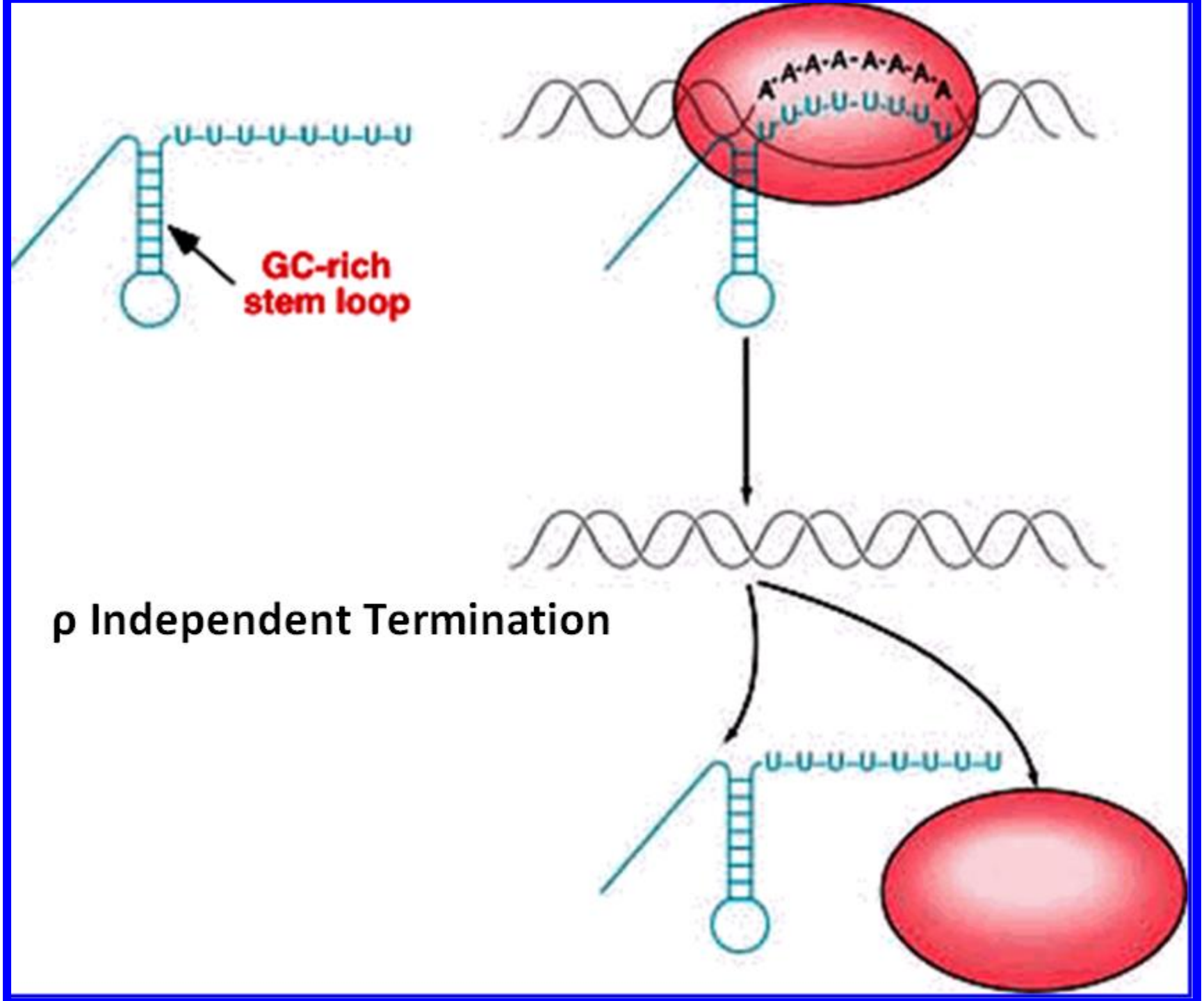
- **Simple terminator sequence mechanism or ρ independent termination** is dependent only on DNA base sequence. In this type 3 important regions are seen, they are:
  - There is an inverted repeat containing a central non-repeating segment which forms a stem and loop in the RNA transcript and possibly a cruciform structure in DNA strands.
  - (Reason may be to protect the RNA from RNase II, an intracellular RNase which is inactive against dsRNA)

Rho independent Termination of RNA

- A high GC region is seen near the loop end of the stem. RNA polymerase slows down during the synthesis of corresponding RNA.
- The GC rich region is followed by a sequence of ‘A’s in the template DNA strand. As the template has a stretch of ‘A’s, the mRNA has 6-8 uridylates.

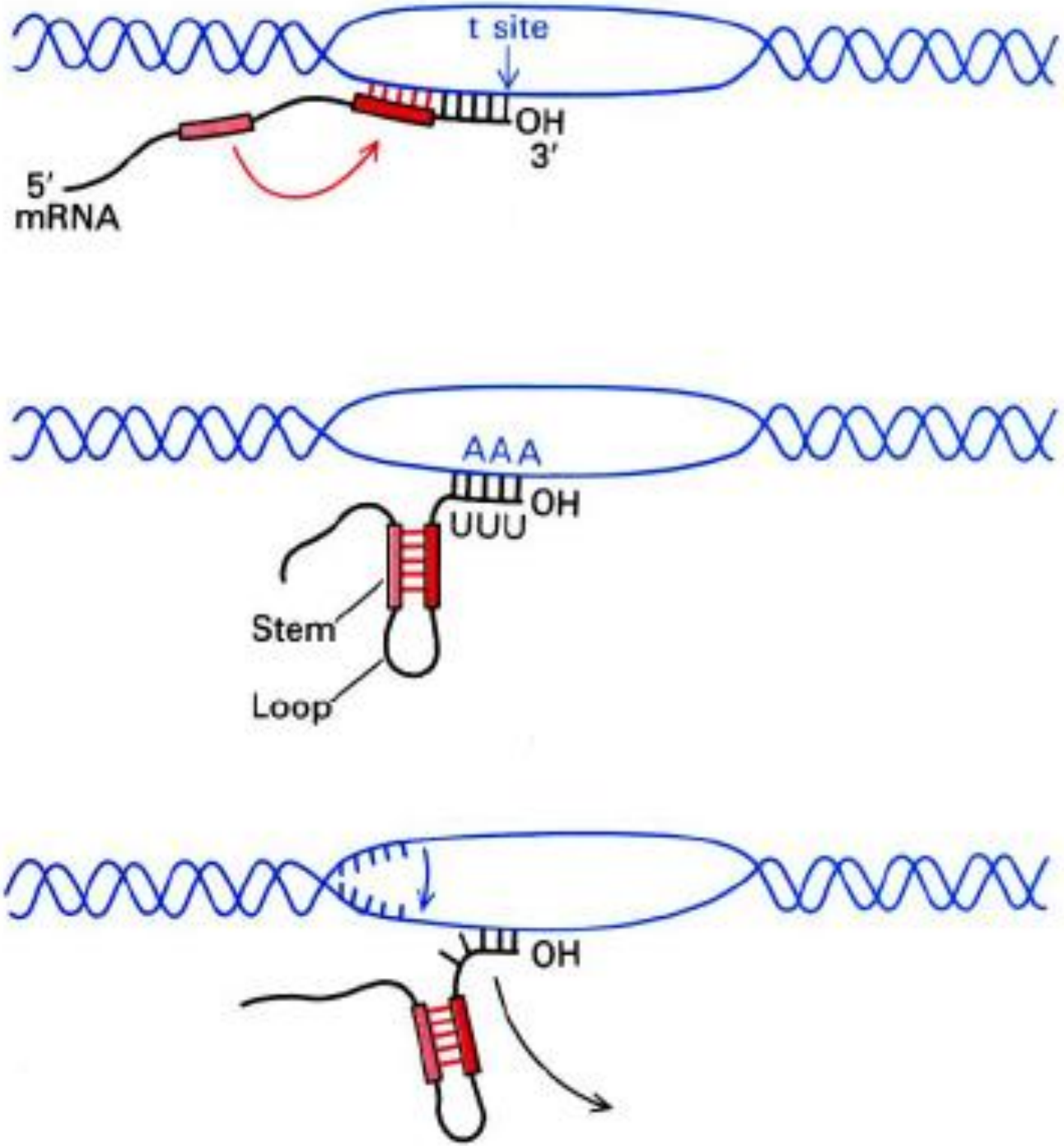
Rho independent Termination of RNA

- The **Rho independent Termination of RNA** is not clearly understood.
- Transcription does not **terminate a unique** site. Probably RNA containing the hairpin loop followed by uridylates is bound to Adenines of DNA only by a double bond.
- This may result in dissociation of RNA polymerase from DNA.
- The number of uridylates, varies from organism to organism.



ρ Independent Termination

Rho Independent Termination



Rho dependent Termination of RNA

- In **Termination which requires auxiliary factors or ρ dependent termination** a hair loop similar to ρ independent termination is formed in the transcript.
- However this hair pin loop is a shorter one and it lacks the uridylates after GC rich region.
- When the enzyme RNA polymerase encounters such a sequence in DNA it pauses.
- During this pause, if ρ protein is present then the termination occurs.
- If ρ protein is not present then transcription continues.

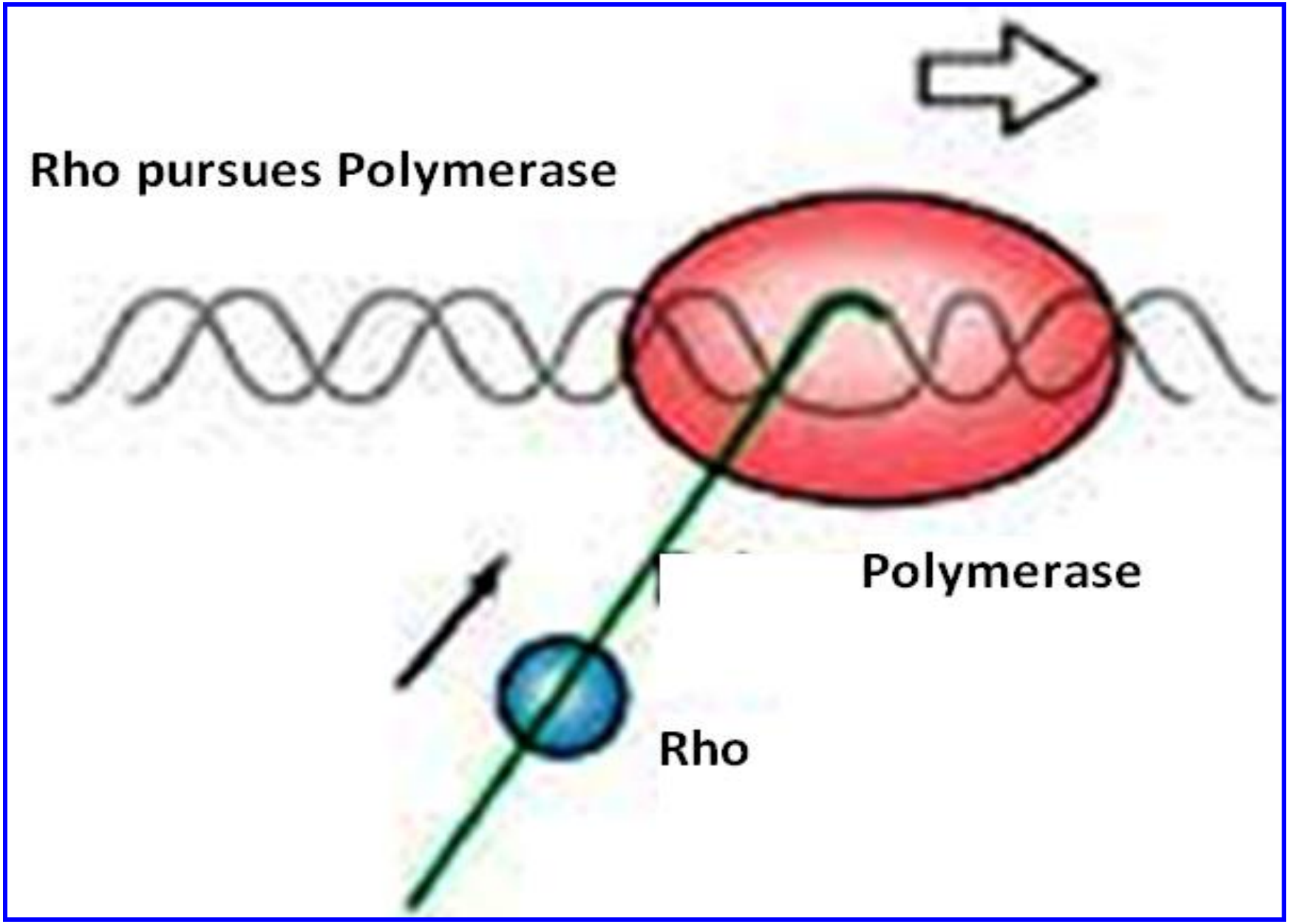
Rho dependent Termination of RNA

- Rho (ρ) protein binds to RNA rich in GC bases and acquires ATP cleaving activity which is essential for termination.
- The mechanism of ρ dependent termination is not very clear. Somehow the ρ protein and RNA polymerase interact and terminate transcription.
- ρ has multifaceted activities. Some sequences which do not require ρ for termination require ρ for the release of RNA.

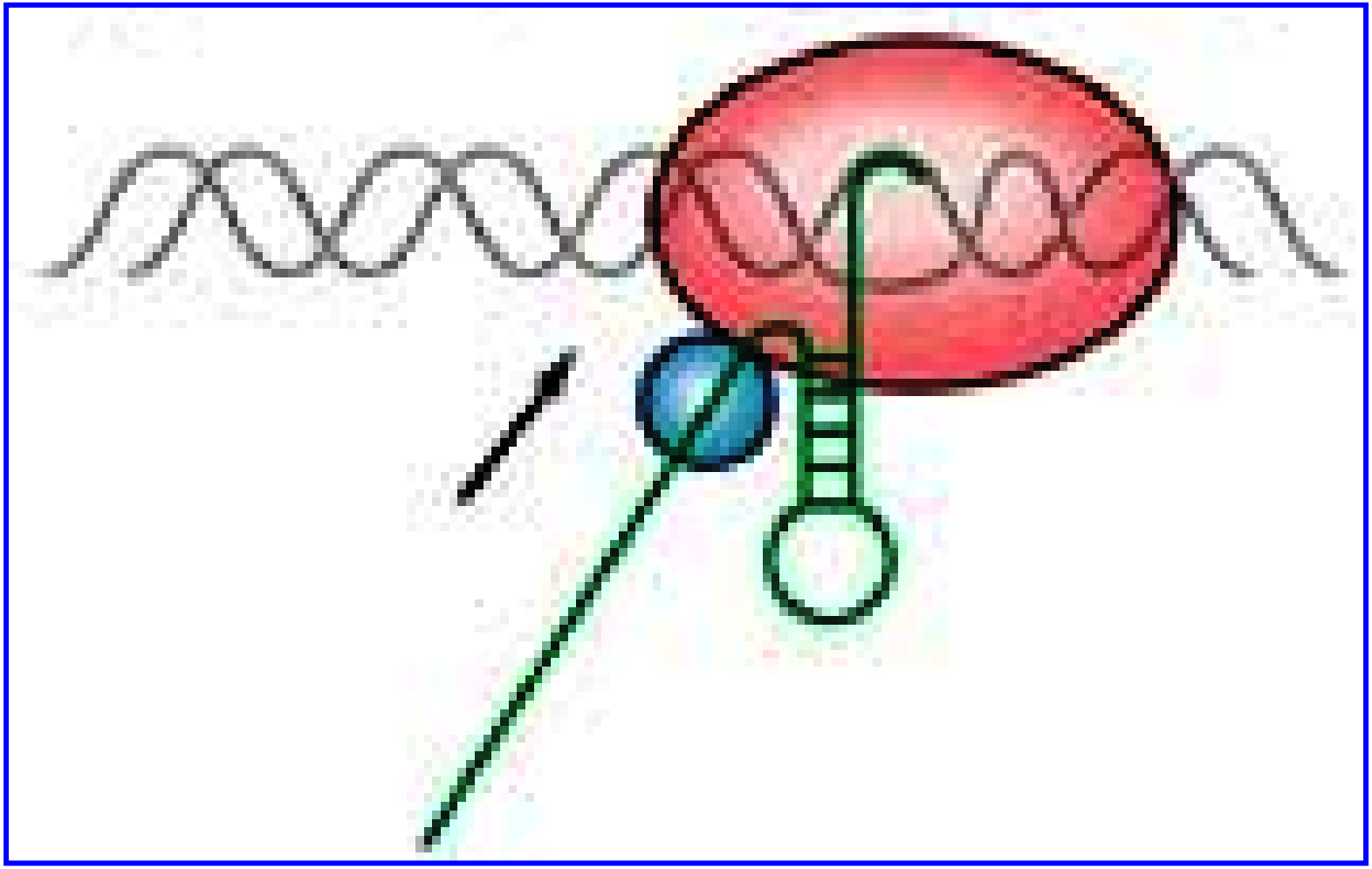
Rho dependent Termination of RNA

- Later studies have revealed that for this termination hair pin loop is not essential.
- ρ protein binds to 72 nucleotide Rho Utilization (***rut***) stretch in RNA.
- ρ moves along nascent RNA towards transcription complex and terminates transcription when it encounters the 72 nucleotide '***rut***' sequence.

Rho dependent Termination of RNA



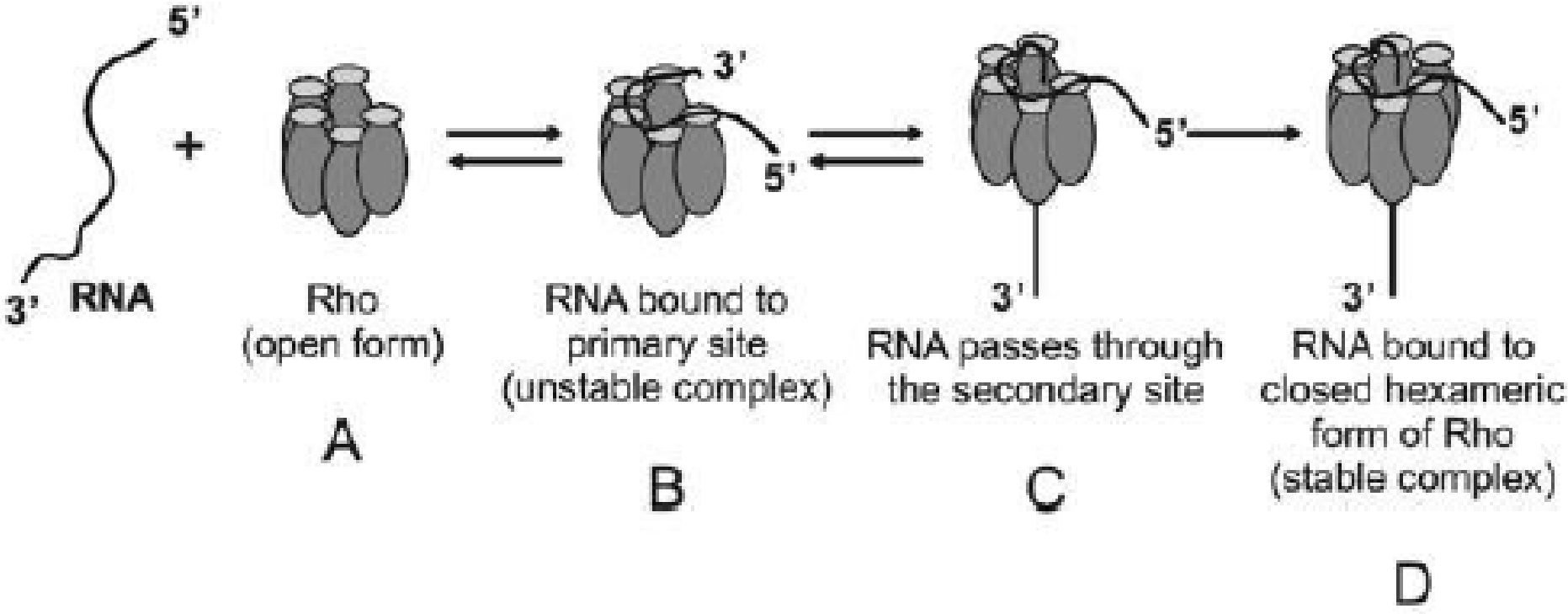
Rho dependent Termination



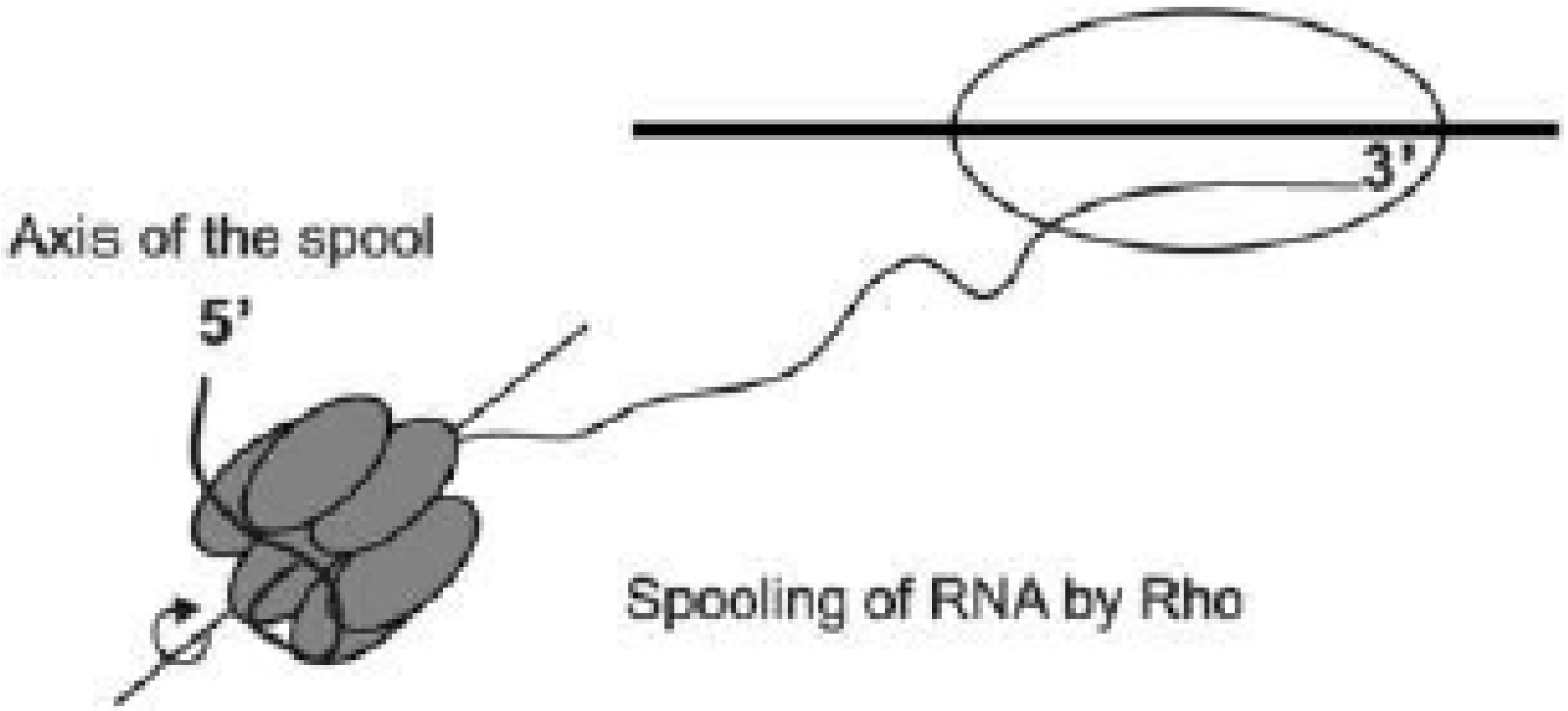
Hairpin forms; Polymerase pauses; Rho catches up.

Model for rho Dependent Termination

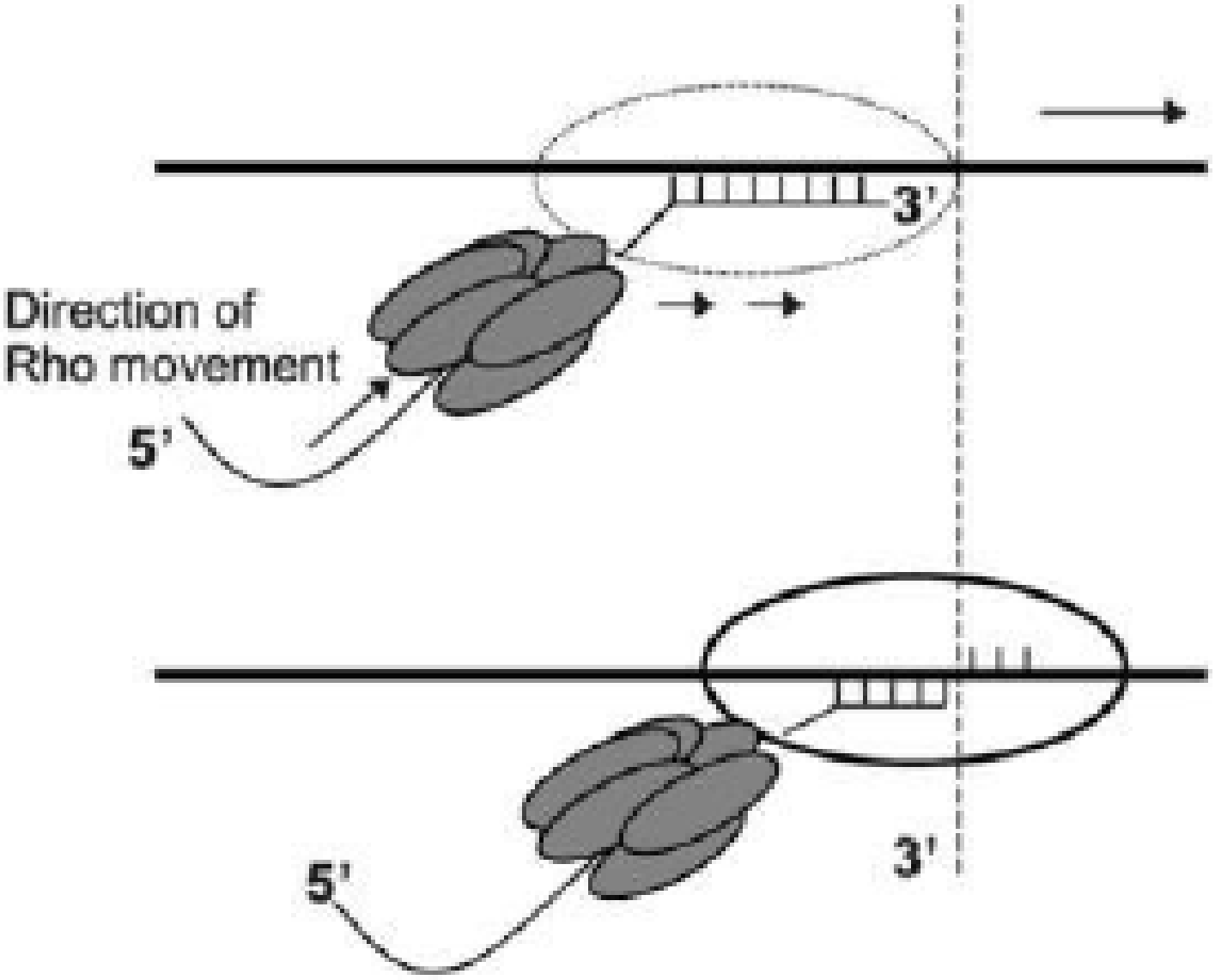
Rho binds as a hexameric protein complex to specific sequences called RUT (rho utilization) sites. The complex also binds ATP and moves along the RNA ultimately disrupting the interactions between the RNA polymerase and the RNA.



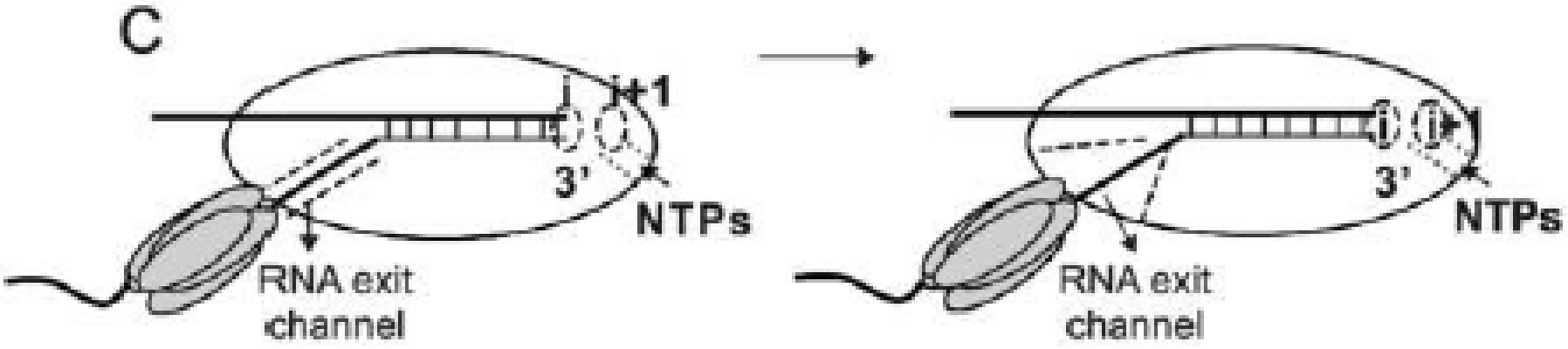
A



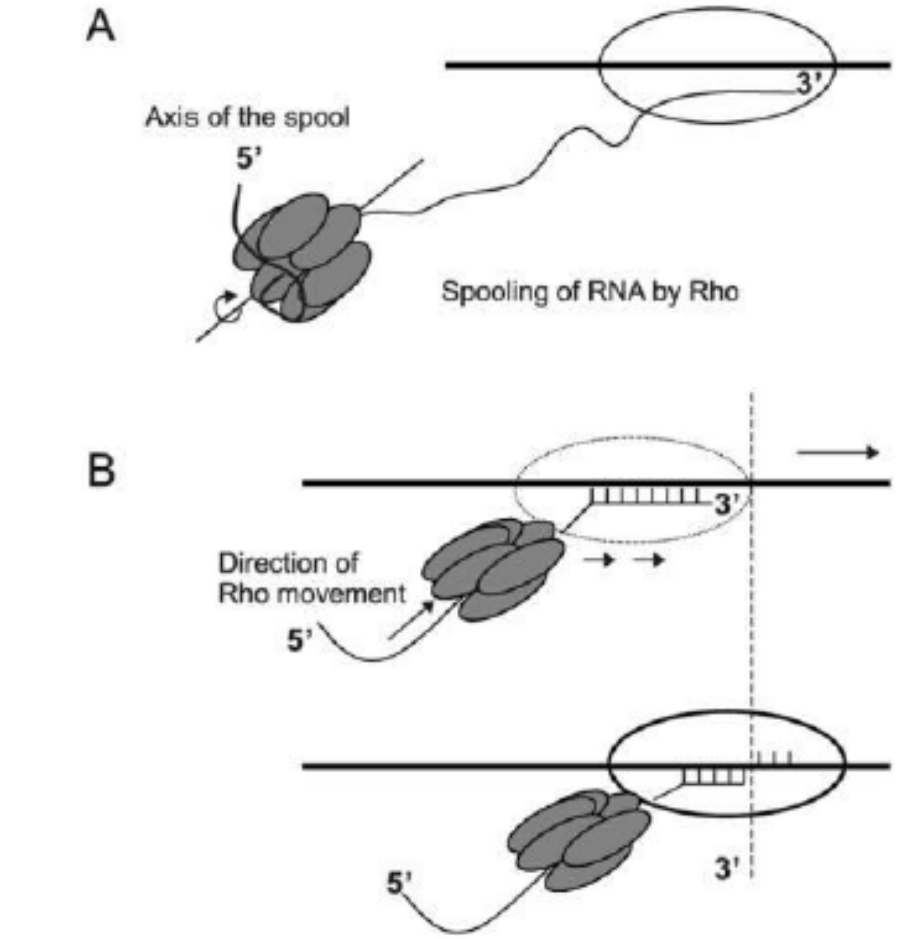
B



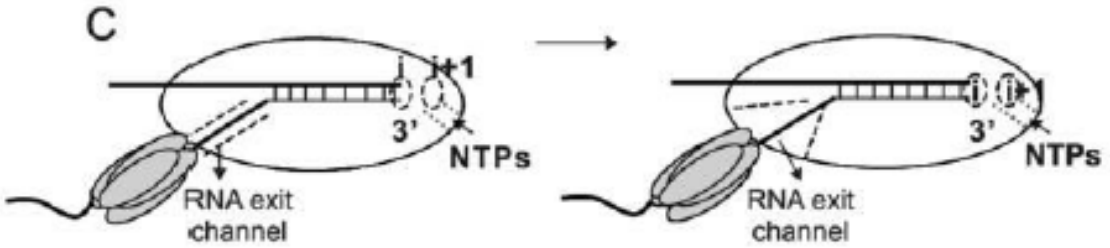
The translocating Rho pushes the RNA polymerase



Rho interaction with RNA polymerase changes the conformation of the RNA exit channel



The translocating Rho pushes the RNA polymerase



Rho interaction with RNA polymerase changes the conformation of the RNA exit channel

Release of newly synthesized RNA

- The final step in termination is disassociation of both the core enzyme and the RNA from the DNA a poorly understood step.
- Following the disassociation of core enzyme and RNA from the DNA, the core enzyme interacts with free σ subunit to reform holoenzyme which again binds to DNA and initiates transcription.
- Thus transcription includes a **σ cycle**.

Antibiotic inhibitors of Transcription

- **Rifampicin:**
- Antibiotics are highly specific inhibitors of biological processes.
- Rifamycins derived from *Streptomyces* and Rifampicin a semi synthetic derivative specifically inhibit the initiation of RNA synthesis.
- Rifampicin does not block the binding of RNA polymerase to the DNA template, but interferes with the formation of first few phosphodiester bonds in the RNA chain.

Antibiotic inhibitors of Transcription

- **Rifampicin:**
- The structure of a complex between prokaryotes RNA polymerase and rifampicin reveals that the antibiotic blocks the channel into which the RNA-DNA hybrid generated by enzyme must pass.
- Rifampicin does not hinder chain elongation once initiated because the RNA-DNA hybrid present in the enzyme prevents the antibiotic from binding.

Antibiotic inhibitors of Transcription

- **Actinomycin D:**
- It is a polypeptide containing antibiotic form is isolated from different strain of S*treptomyces*.
- Actinomycin D binds tightly and specifically to double helical DNA and thereby prevents it from being an effective template for RNA synthesis.

Antibiotic inhibitors of Transcription

- **Actinomycin D:**
- Actinomycin D does not bind to ssDNA or RNA or dsRNA or DNA-RNA hybrids.
- The studies (spectroscopic and hydrodynamic studies) of complexes of Actinomycin D and DNA have suggested that the phenoxazone ring of Actinomycin slips in between neighboring base pairs in DNA (intercalation).

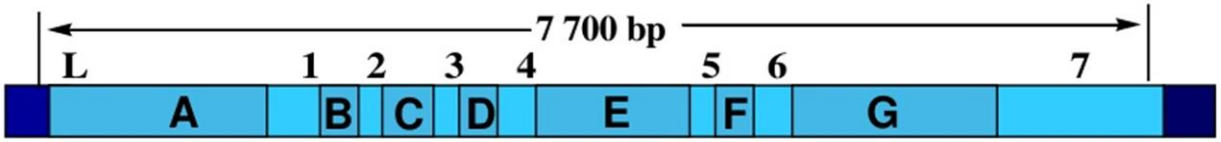
Antibiotic inhibitors of Transcription

- **Actinomycin D:**
- At low concentrations Actinomycin D inhibits transcription without significantly affecting DNA replication or protein synthesis.
- Hence Actinomycin D is used as a highly specific inhibitor of new RNA formation in both prokaryotes and eukaryotes.
- It is also used in cancer because of its ability to inhibit the growth of rapidly dividing cells.

THANK YOU

Splicing

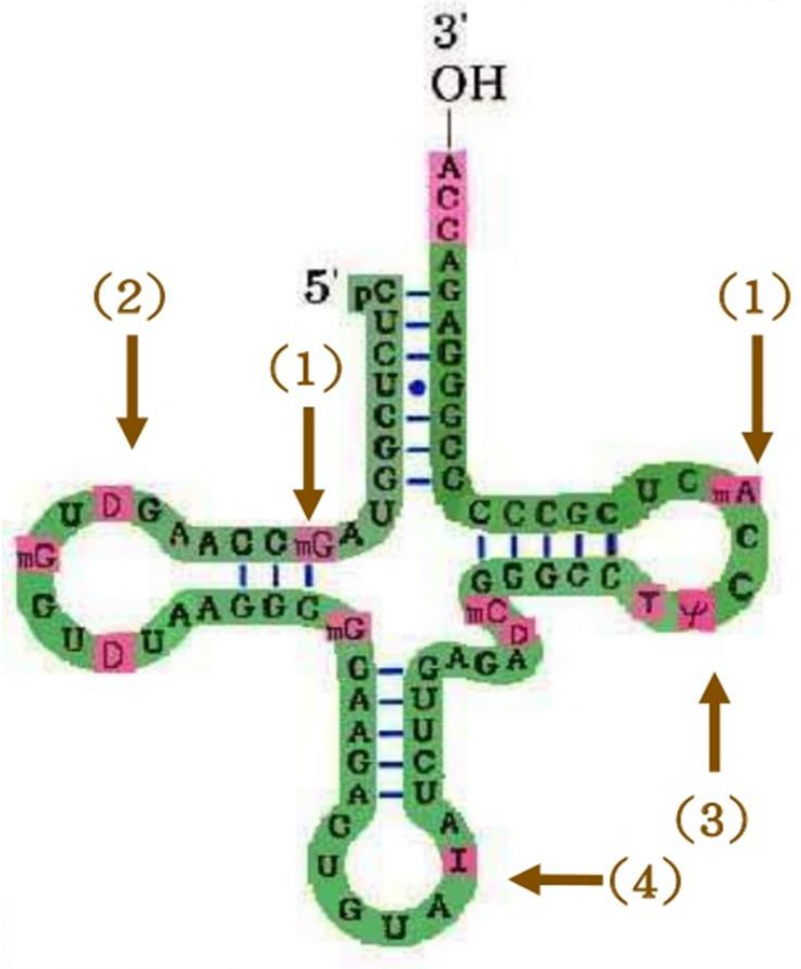
The structural genes are composed of coding and non-coding regions that are alternatively separated.



A-G non-coding region 1-7 coding region

Noncoding sequences called intervening sequences or Introns & coding sequences called Exons

Base modification



1. **Methylation**
   A→mA, G→mG
2. **Reduction**
   U→DHU
3. **Transversion**
   U→ψ
4. **Deamination**
   A→I