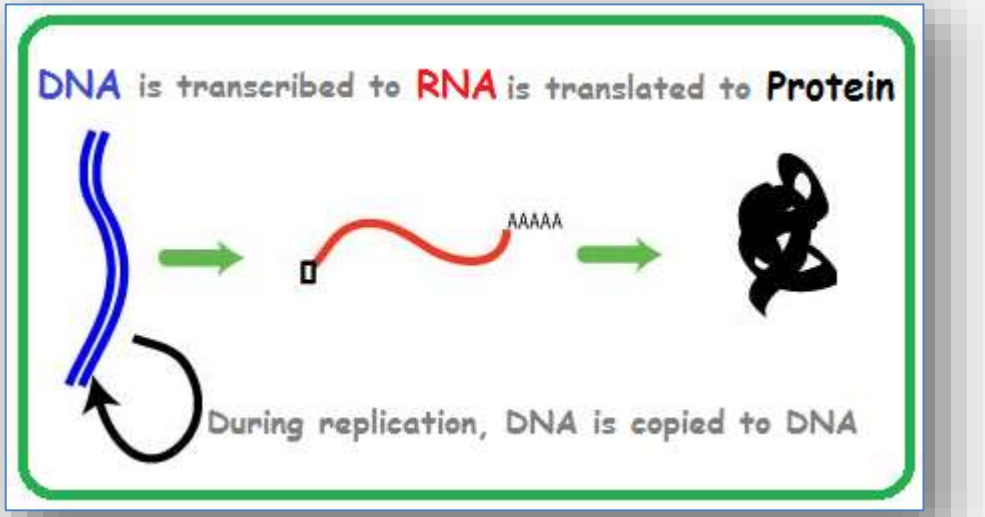
Transcription

Introduction

- DNA stores genetic information in a stable form that can be readily replicated.

- The expression of this genetic information requires its flow from DNA to RNA to protein RNA is the only macromolecule known to have an role both in the storage and transmission of information and in catalysis.



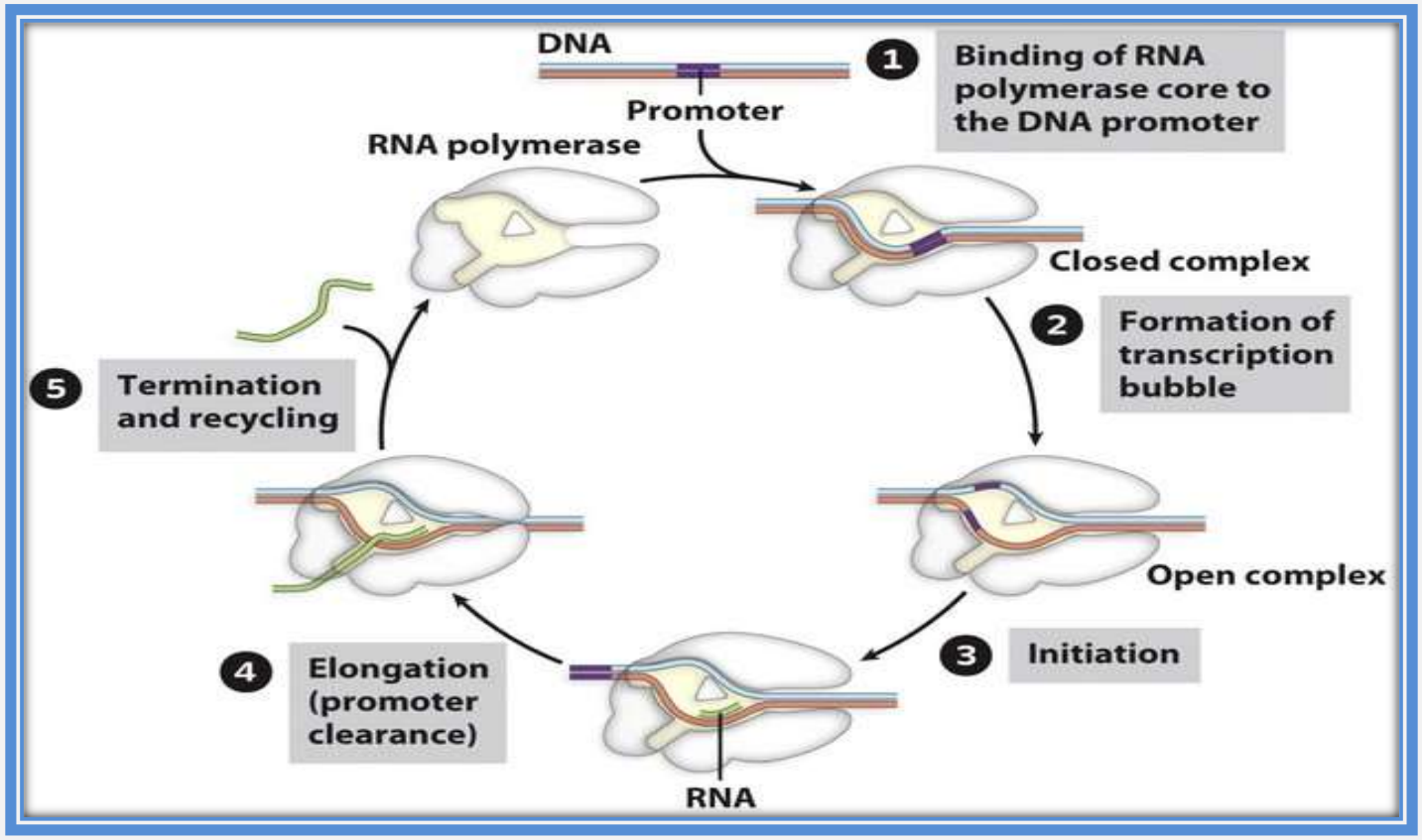
Question:

i. What are the properties of promoters (the DNA sites at which RNA transcription is initiated), and how do the promoters function?

ii. How do RNA polymerase, the DNA template, and the nascent RNA chain interact with one another?

iii. How is transcription terminated?

- The stages of transcription are
  - Initiation
  - Elongation
  - Termination



An overview of transcription. DNA binding at the promoter leads to initiation of transcription by the polymerase holoenzyme, followed by elongation and termination.

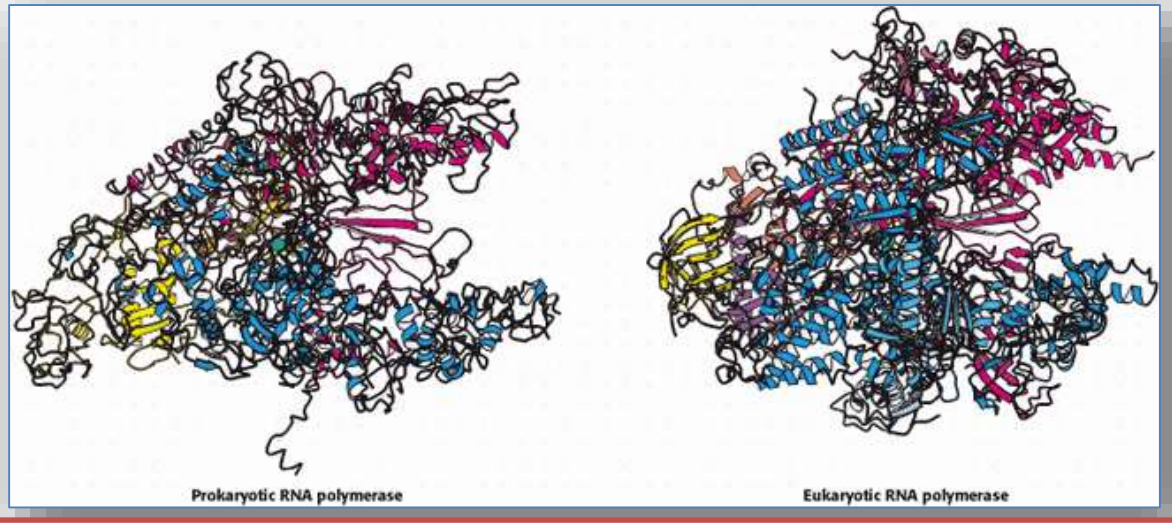
- ✓ During **transcription,** an enzyme system converts the genetic information in a segment of double-stranded DNA into an RNA strand with a base sequence complementary to one of the DNA strands.
- ✓ During replication the entire chromosome is usually copied, but transcription is more selective.
- ✓ Specific regulatory sequences mark the beginning and end of the DNA segments to be transcribed and designate which strand in duplex DNA is to be used as the template

DNA-Dependent Synthesis of RNA

- ✓ Like replication, transcription has initiation, elongation, and termination phases.
- ✓ Transcription differs from replication in that it does not require a primer and, generally, involves only limited segments of a DNA molecule.
- ✓ Additionally, within transcribed segments only one DNA strand serves as a template for a particular RNA molecule.

An Overview of RNA Synthesis

- RNA synthesis, or transcription, is the process of transcribing DNA nucleotide sequence information into RNA sequence information.
- RNA synthesis is catalyzed by a large enzyme called RNA polymerase.
- The basic biochemistry of RNA synthesis is common to prokaryotes and eukaryotes, although its regulation is more complex in eukaryotes.



RNA Polymerase Structures: The similarity of these structures reveals that these enzymes have the same evolutionary origin and have many mechanistic features in common.

RNA polymerase performs multiple functions in the RNA synthesis:

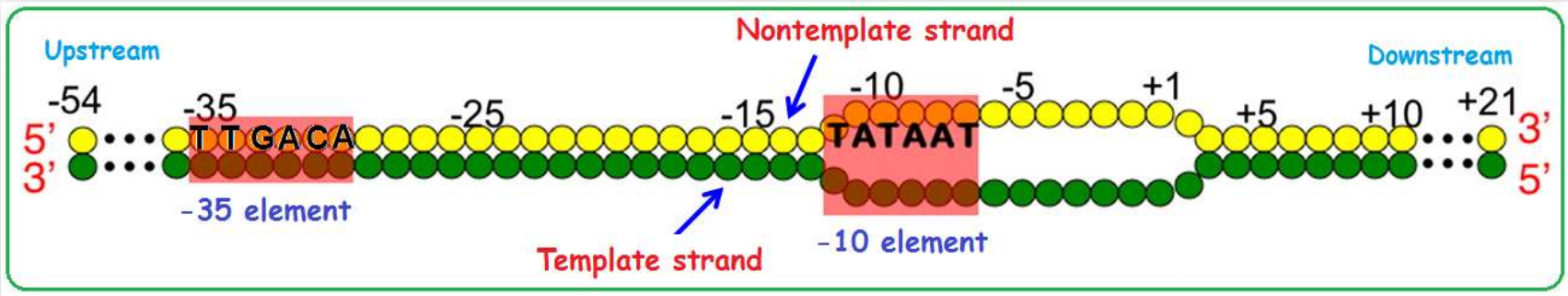
- It searches DNA for initiation sites, called promoter sites.
- It unwinds a short stretch of double-helical DNA to produce a single-stranded DNA template from which it takes instructions.
- It selects the correct ribonucleoside triphosphate and catalyzes the formation of a phosphodiester bond. This process is repeated many times as the enzyme moves unidirectionally along the DNA template.
- It detects termination signals that specify where a transcript ends.
- It interacts with activator and repressor proteins that modulate the rate of transcription initiation. These proteins, which play a more prominent role in eukaryotes than in prokaryotes, are called transcription factors.
- Gene expression is controlled mainly at the level of transcription

Transcription Is Catalyzed by RNA Polymerase

- ✓ RNA polymerase from E. coli is a very large (~400 kd) and complex enzyme consisting of four kinds of subunits .
- ✓ The subunit composition of the entire enzyme, called the **holoenzyme,** is $\alpha_2 \beta \beta \sigma$.
- ✓ The σ subunit helps find a promoter site where transcription begins, participates in the initiation of RNA synthesis, and then dissociates from the rest of the enzyme.
- ✓ RNA polymerase without this subunit ($\alpha_2 \beta \beta$) is called the **core enzyme**.
- ✓ The core enzyme contains the catalytic site.

Transcription Is Initiated at Promoter Sites on the DNA Template

- Transcription starts at promoters on the DNA template. Promoters are sequences of DNA that direct the RNA polymerase to the proper initiation site for transcription.
- Two common motifs are present on the 5 (upstream) side of the start site. They are known as the -10 sequence and the -35 sequence because they are centered at about 10 and 35 nucleotides upstream of the start site. These sequences are each 6 bp long. Their consensus sequences, deduced from analyses of many promoters, are:



Remember:

- The sequence of the template strand of DNA is the complement of that of the RNA transcript .
- In contrast, the coding strand of DNA has the same sequence as that of the RNA transcript except for thymine (T) in place of uracil(U).
- The coding strand is also known as the sense (+) strand, and the template strand as the antisense (-) strand.

The efficiency or strength of a promoter sequence serves to regulate transcription.

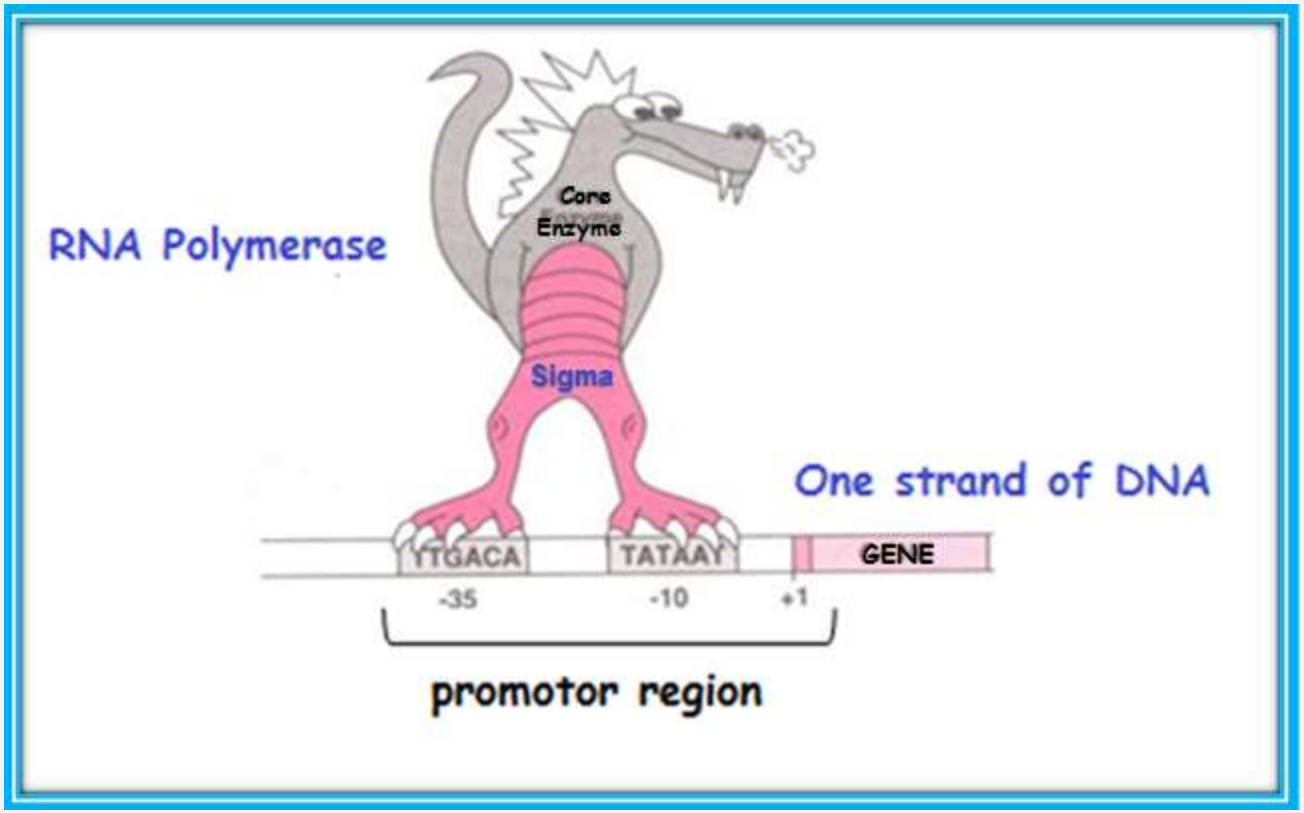
How?

- Genes with strong promoters are transcribed frequently as often as every 2 seconds in E. coli.
- In contrast, genes with very weak promoters are transcribed about once in 10 minutes.
- The -10 and -35 regions of most strong promoters have sequences that correspond closely to the consensus sequences, whereas weak promoters tend to have multiple substitutions at these sites.
- Indeed, mutation of a single base in either the -10 sequence or the -35 sequence can diminish promoter activity.
- The distance between these conserved sequences also is important; a separation of 17 nucleotides is optimal.

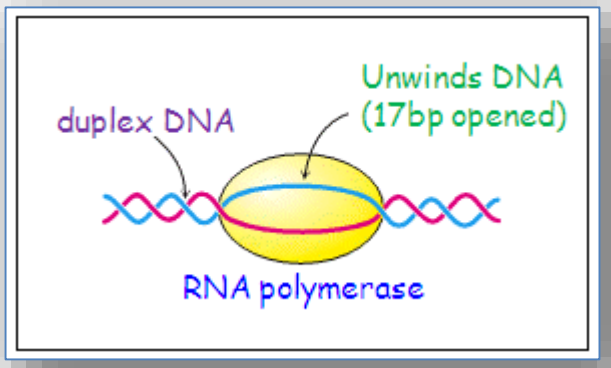
Sigma Subunits of RNA Polymerase Recognize Promoter Sites

- The α_2 β β core of RNA polymerase is unable to start transcription at promoter sites. Rather, the complete α_2 β β σ holoenzyme is essential for initiation at the correct start site.

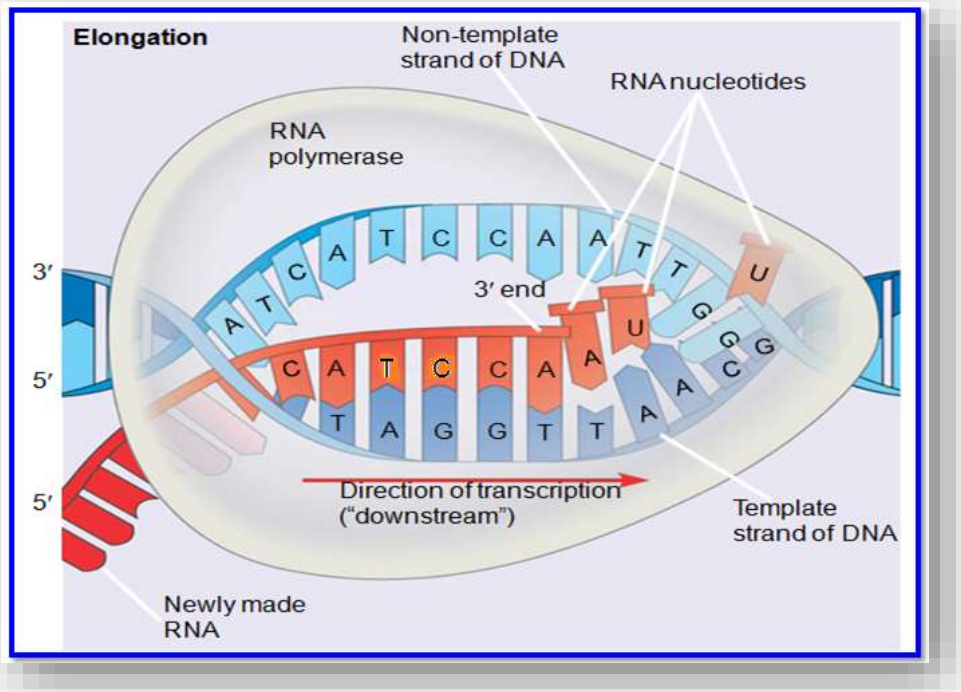
- The σ subunit contributes to specific initiation in two ways:
  - First, it decreases the affinity of RNA polymerase for general regions of DNA by a factor of 10^4 . In its absence, the core enzyme binds DNA indiscriminately and tightly.
  - Second, the s subunit enables RNA polymerase to recognize promoter sites.


RNA Polymerase
Core Enzyme
Sigma
One strand of DNA
TTGACA
TATAAT
GENE
-35
-10
+1
promotor region

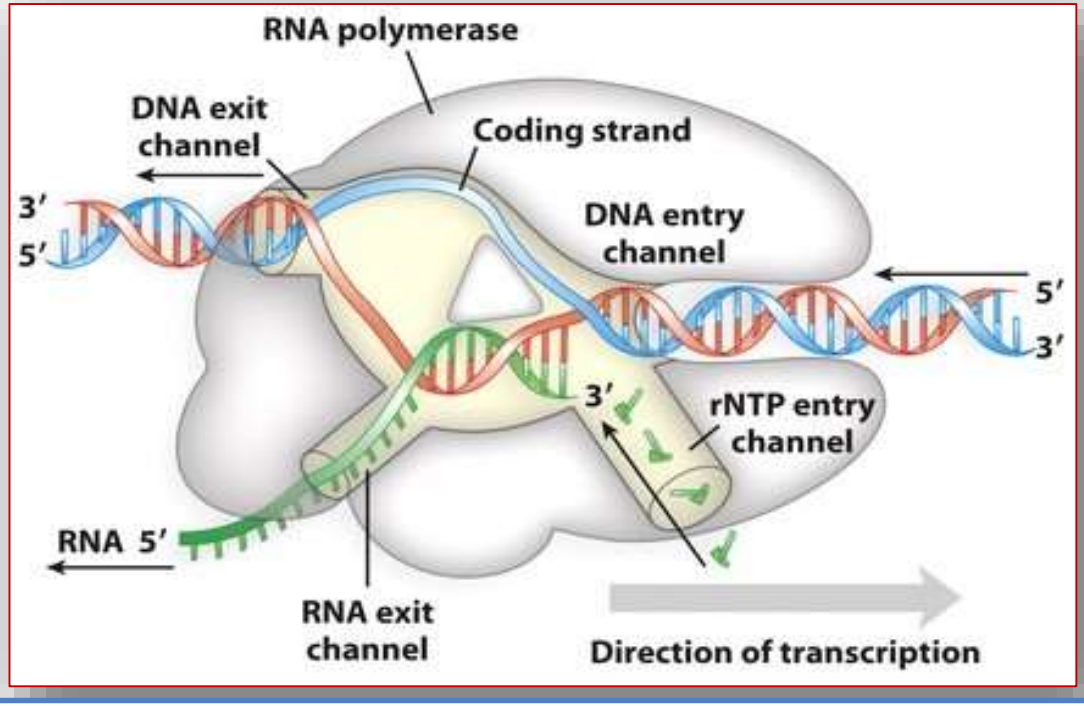
- A region of duplex DNA must be unpaired so that nucleotides on one of its strands become accessible for base-pairing with incoming ribonucleoside triphosphates.
- The DNA template strand selects the correct ribonucleoside triphosphate by forming a Watson-Crick base pair with it.
- Because unwinding increases the negative supercoiling of the DNA, the degree of negative supercoiling increased in proportion to the number of RNA polymerase molecules bound per template DNA, showing that the enzyme unwinds DNA.



The second stage of Transcription :Elongation



- Once RNA polymerase enters the elongation phase, the enzyme does not release the DNA template until it encounters a termination sequence.
- During transcript elongation, the DNA moves through the polymerase active site, as observed in the polymerase open complex

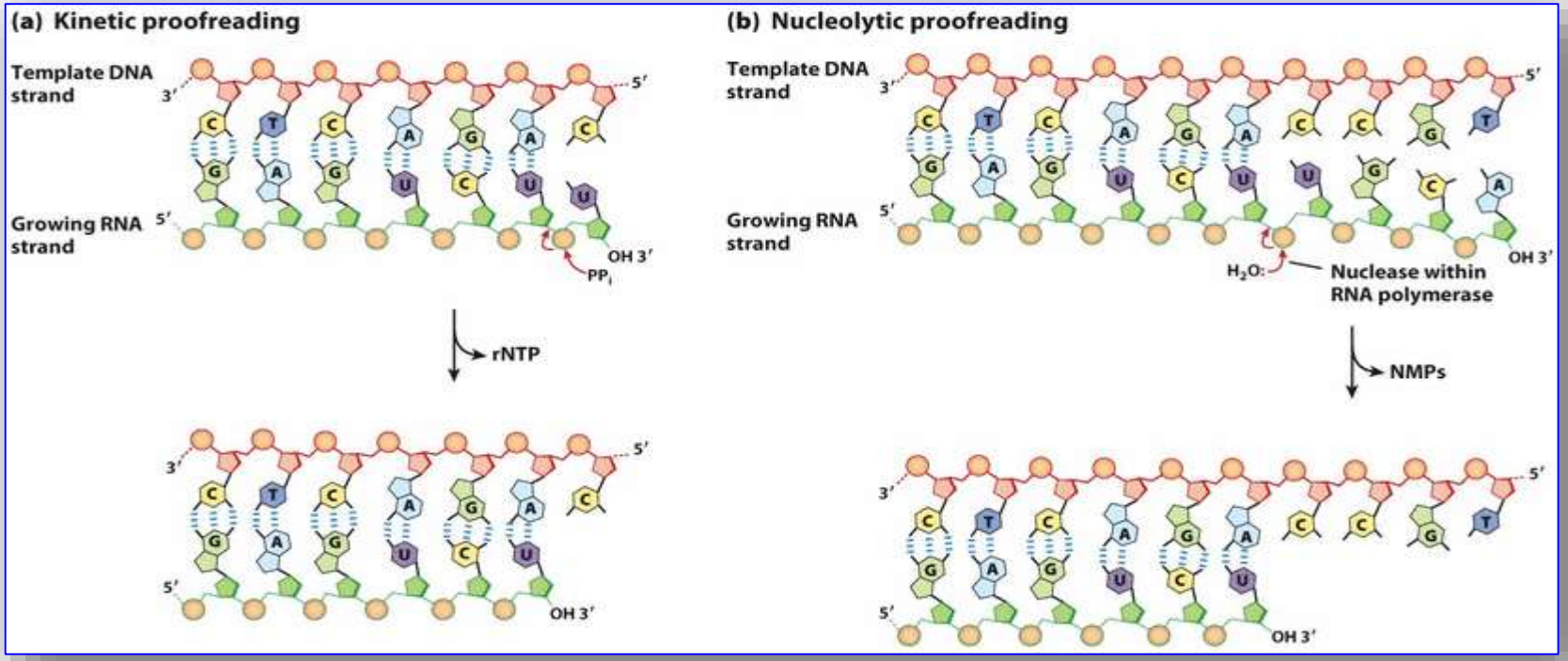


RNA polymerase channels. Distinct channels in RNA polymerase allow the DNA to enter as double-stranded DNA and to peel apart within the polymerase so that 8 bp form between the template strand and the growing RNA transcript. Two other channels provide entry for rNTPs and an exit for the transcript

- During elongation, the polymerase attempts to ensure the accuracy of transcription by pyrophosphorolysis, in which the catalytic reaction runs in reverse whenever the polymerase stalls along the DNA.

- This process, known as kinetic proofreading, works because the polymerase tends to stall after incorporating a mismatched base into the growing RNA chain, thus enabling pyrophosphorolysis to remove the incorrect base.

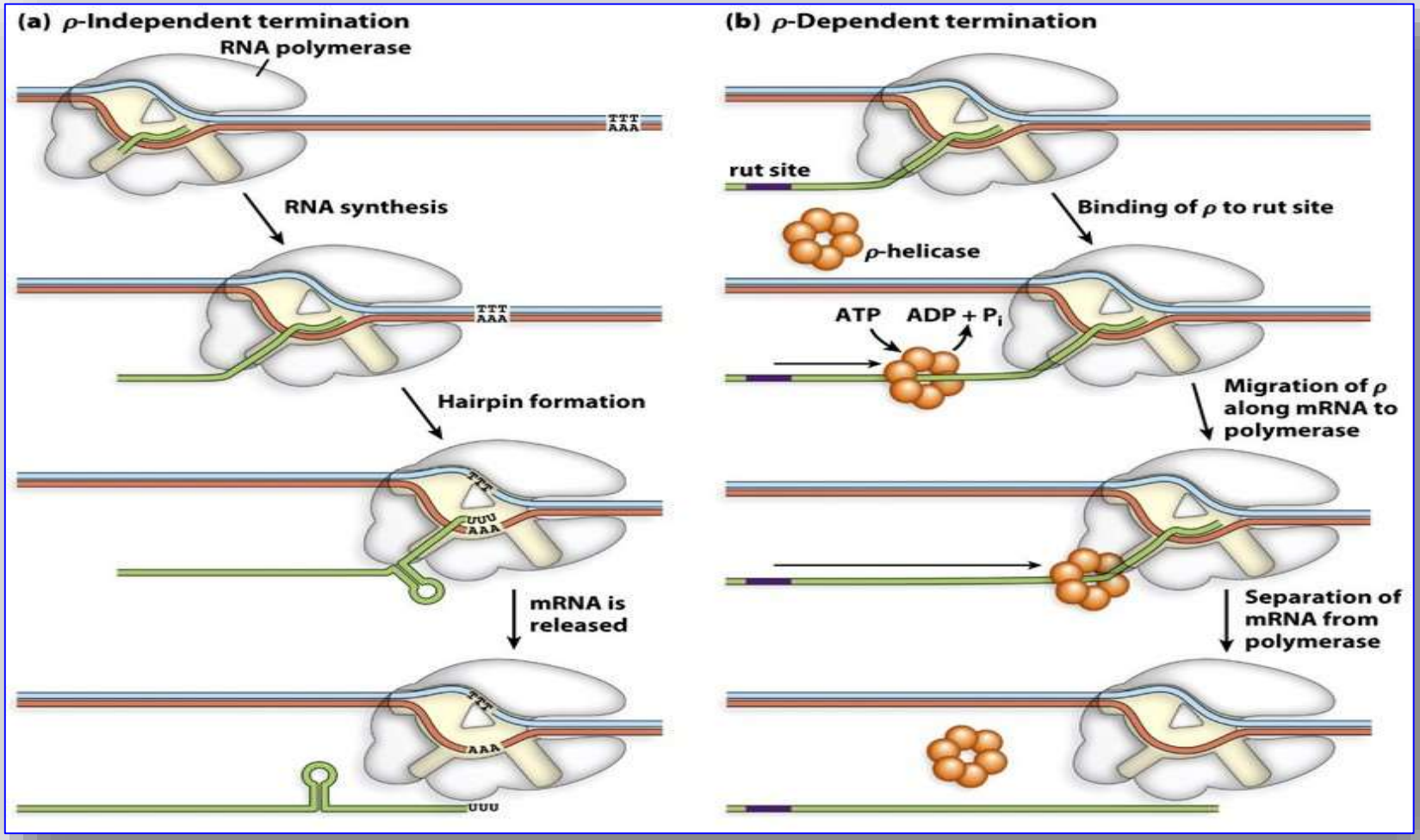
- Pyrophosphorolysis is also used in the proofreading that occurs during DNA synthesis.



- Proofreading by RNA polymerase. (a) In kinetic proofreading, the polymerase stalls after incorporating a mismatched base into the growing RNA chain, enabling pyrophosphorolysis to remove the incorrect base. (b) In nucleolytic proofreading, the polymerase backtracks on the DNA, melting several nucleotides of the RNA (i.e., breaking the DNA-RNA base pairs), then an intrinsic nuclease removes the section of melted RNA.

The third stage of Transcription :Termination

- Transcription stops when the RNA polymerase transcribes through certain sequences in the DNA template.
- At this point, the polymerase releases the finished transcript and dissociates from the template.
- E. coli DNA has at least two classes of such termination sequences, one class that relies primarily on structures that form in the RNA transcript and another that requires an accessory protein factor called rho (ρ).
- Most ρ-independent termination sequences have two distinguishing features.



Termination of transcription.
(a) In ρ-independent termination, an mRNA sequence forms a hairpin, followed by Us' residues, stalling the polymerase and separating it from the mRNA.
(b) RNAs that include a rut site (purple) recruit the ρ helicase, which migrates in the 5'→3' direction along the mRNA and separates it from the polymerase.

- Transcription begins at specific promoter sequences upstream from the coding sequence in the DNA template. A sigma factor, of which there are several classes in bacteria, binds to the polymerase holoenzyme and recruits it to a particular type of promoter, enabling transcription at subsets of genes in response to environmental stimuli and the needs of the cell.
- RNA polymerase first forms a closed complex on promoter DNA, a readily reversible state that is not yet capable of transcription.
- Transcription initiation requires promoter clearance, in which the RNA polymerase moves beyond the promoter region of the DNA to begin rapid elongation of the transcript.
- During elongation, the RNA polymerase is highly processive, synthesizing transcripts without dissociating from the DNA template.
- RNA polymerase corrects errors in newly synthesized transcripts through the use of nucleolytic proofreading, in which the polymerase reverses direction by one or a few nucleotides and hydrolyzes the RNA phosphodiester bond upstream of a mismatched base, removing the error-containing strand.
- Termination occurs when the polymerase transcribes through certain DNA sequences in a process that sometimes requires an accessory factor, ρ.