Gene expression

By

Dr. Walaa Bayoumie El Gazzar

Gene expression

- **Definition:** Gene expression can be defined as the gene (DNA) undergoes transcription into **mRNA that can translate the encoded genetic information into protein**.

Transcription (RNA synthesis)

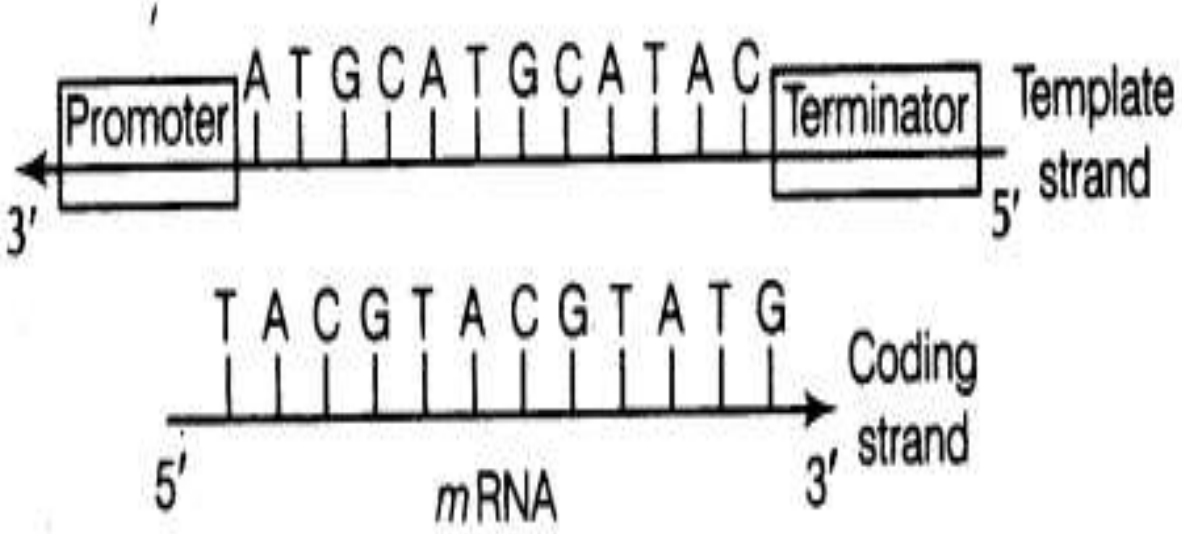
- **Definition:** Transcription is the synthesis of RNA using DNA as a template by **an enzyme DNA –dependent RNA polymerase** or RNA polymerase (**RNAP**)

Features of transcription:

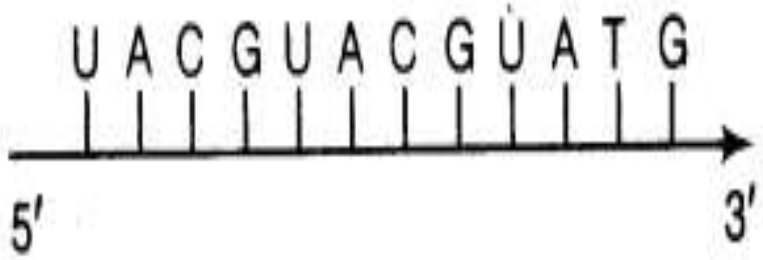
- One strand of the two DNA strands is transcribed only, this strand is called **template strand** (anti-sense), because it provides template for ordering the sequence of nucleotides in an RNA transcript.
- The other strand (non-transcribed) is called **coding strand** (sense strand), because its sequence is the same as the newly synthesized RNA transcript (except for thymine is substituted by uracil)

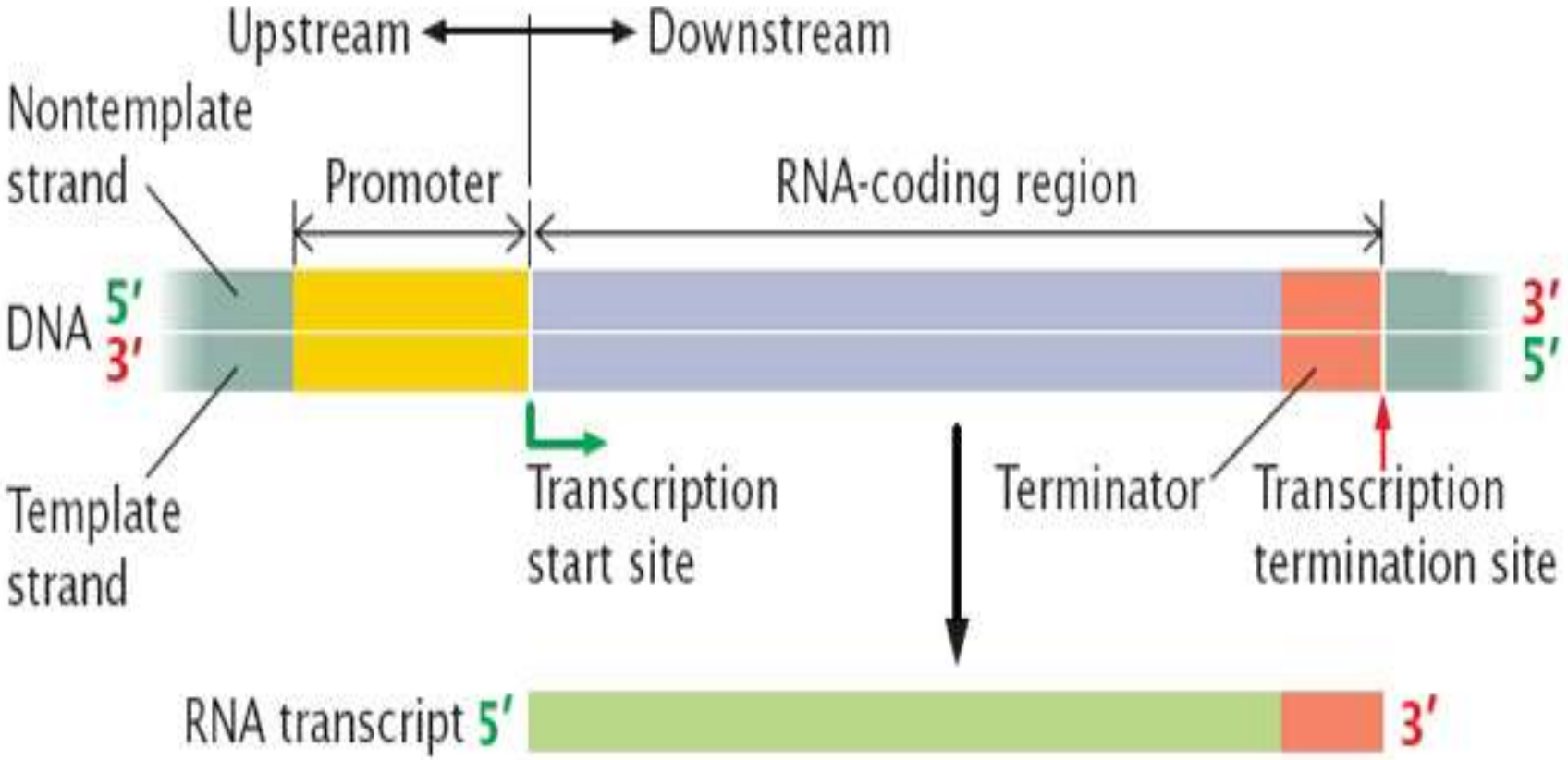
- The DNA template strand is read in 3\to 5\ by RNA polymerase enzyme and the new RNA is synthesized in the direction of 5\ to 3\.
- Upstream means in the 3` direction of the template strand.
- Downstream means in the 5` direction of the template strand.
- A transcription region is the nucleotide sequence transcribed by the enzyme RNAP. It is the region between the promoter and the terminator.

(*i*) Transcription unit



(*ii*) RNA transcribed




Upstream
Downstream
Nontemplate strand
Promoter
RNA-coding region
DNA 5′
3′
3′
5′
Template strand
Transcription start site
Terminator
Transcription termination site
RNA transcript 5′
3′

- **Promoter:** A promoter is the DNA sequence that initially binds the RNA polymerase (together with any initiation factors required). i.e. Nucleotide sequence in DNA to which RNA polymerase binds to begin transcription.
- **Terminator:** Sequences trigger the elongating polymerase to dissociate from the DNA and release the RNA chain it has made.
- **The start point** is the nucleotide at the 3` end of the transcription region " that codes for the initial base of the mRNA" . It is designated +1. Adjacent nucleotides are given positive numbers that increase as we go downstream the transcription unit.

- **The nucleotide in the promoter adjacent to the +1 nucleotide is designated -1 and adjacent nucleotides are given negative numbers that increase as we go upstream the promoter.**

- **The transcription unit Sequence of nucleotides in DNA that codes for a single RNA molecule, along with the sequences necessary for its transcription; normally contains a promoter, an RNA-coding sequence, and a terminator. (i.e. includes the promoter, the transcription region, and the terminator)**

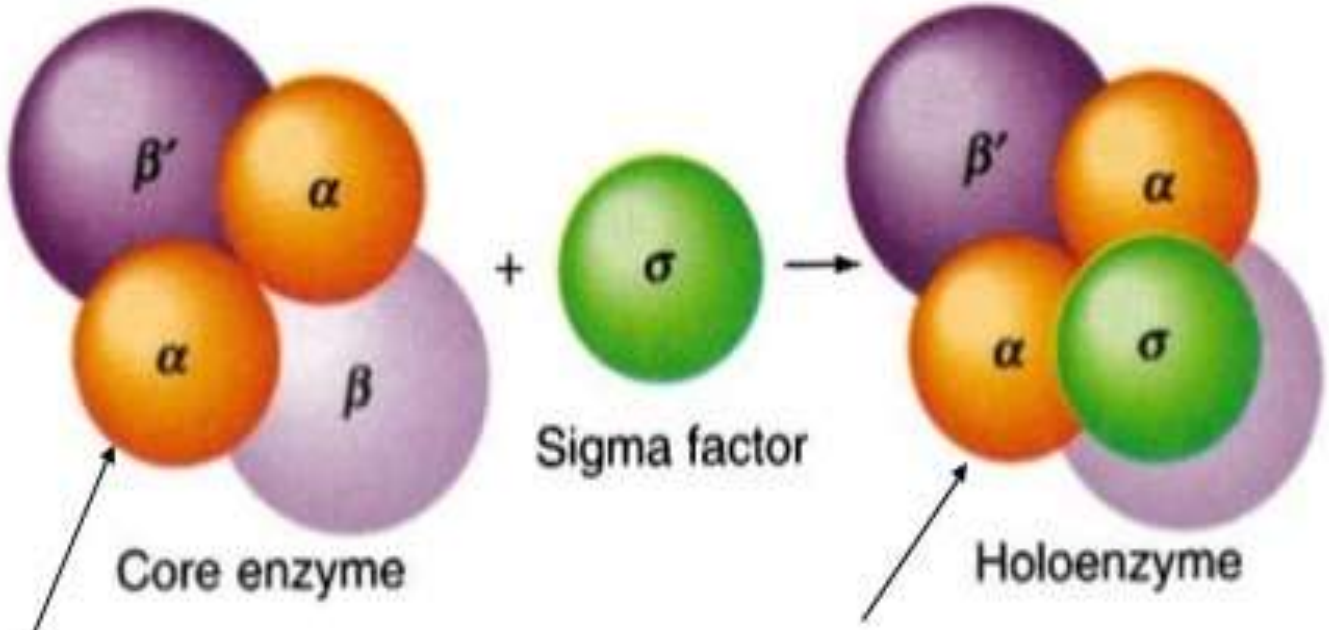
- The DNA nucleotide encoding the beginning of the RNA chain is called the transcription start site and is designated the “+1” position.
- Sequences in the direction in which transcription proceeds are referred to as **downstream** from the start site. Likewise, sequences preceding the start site are referred to as **upstream** sequences.
- When referring to a specific position in the upstream sequence, this is given a negative value. Downstream sequences are allotted positive values.

Transcription in prokaryotes:

- **All types of RNA is synthesized by a specific enzyme called RNA polymerase except for the short RNA primers needed for DNA replication are synthesized by a primase enzyme.**

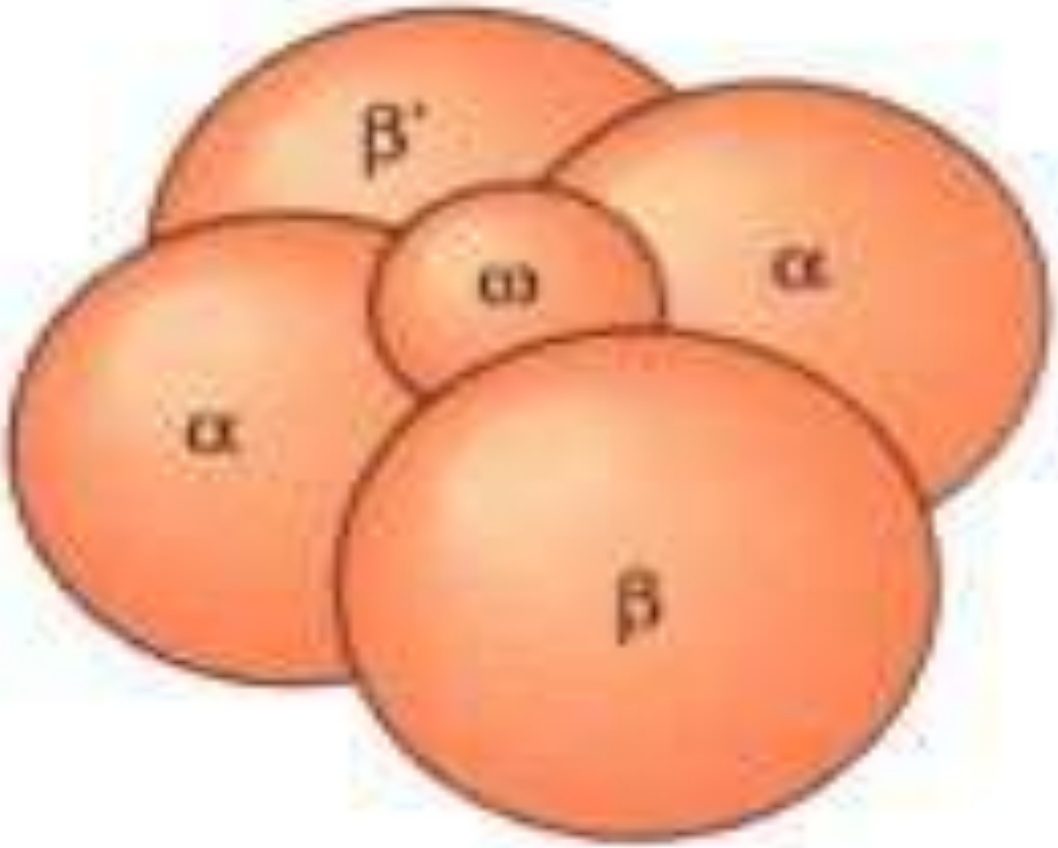
Structure of prokaryotic RNA polymerase:

- **It is a multi-subunit enzyme formed of core enzyme and sigma factor**



Required for polymerization activity

Required for correct initiation of transcription: ***binding to promoter***



Core enzyme

Subunit

- **Core enzyme:** two identical α subunits (regulatory subunits) and two β not identical (β & β') and one ω chain. One of the β subunits (β) binds to the DNA and the other (β‘) is responsible for the formation of phosphodiester bond.

- *RNA polymerase enzyme lacks specificity, that is, it cannot recognize the promoter region on the DNA template.*

- **The σ subunit (“sigma factor”):** It enables RNA polymerase to recognize promoter regions on the DNA. The σ subunit plus the core enzyme make up the holoenzyme. [Note: Different σ factors recognize different groups of genes.]

- **N.B.: The antibiotic rifampicin binds to the β subunits of RNA polymerase and inhibits RNA synthesis in prokaryotes as it interferes with the formation of the first phosphodiester bond. Rifampicin is useful in the treatment of tuberculosis.**

Steps of RNA synthesis in prokaryotes:

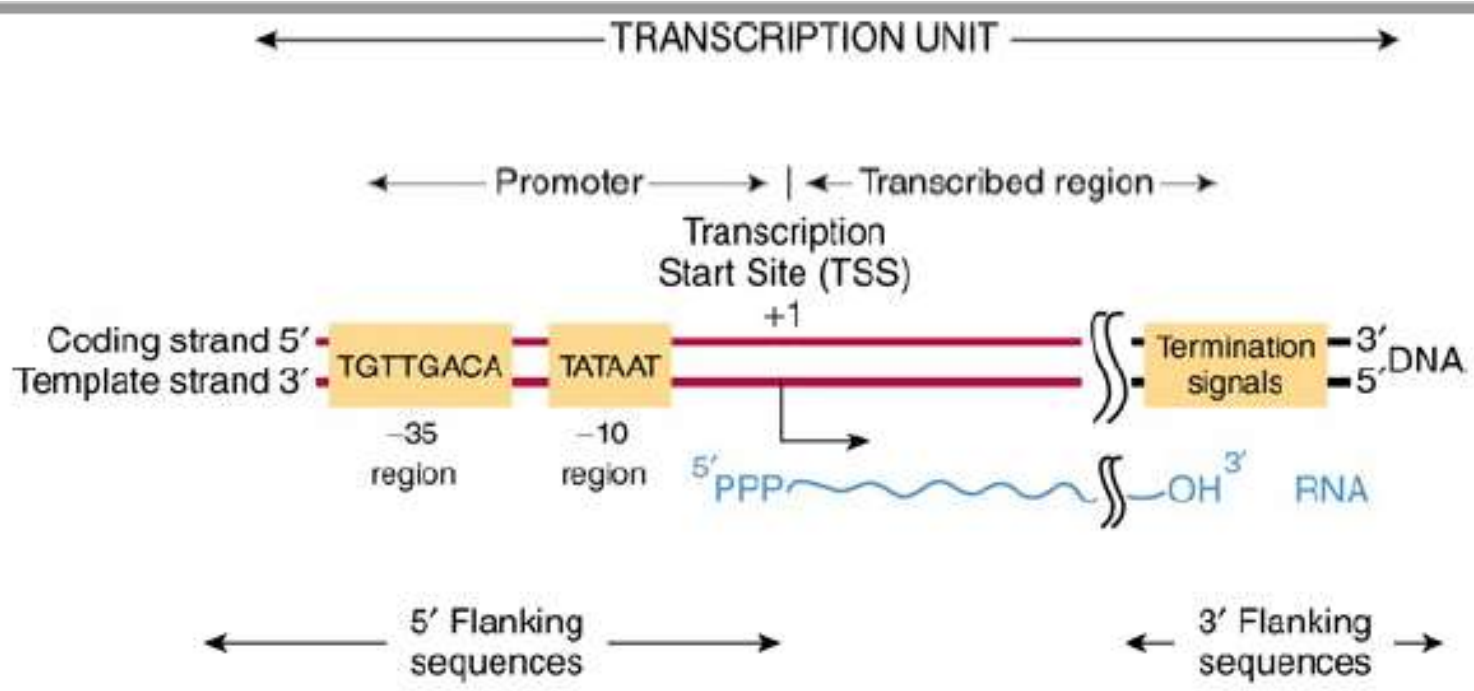
- It is divided into three phases: **initiation**, **elongation** and **termination**.

Initiation:

- It involves the binding of RNA polymerase to a specific region on the DNA known as the promoter region formed of specific base sequence. It needs a specific protein factor called sigma factor (σ) that recognizes and **binds to the promoter region at the TATA box** then RNA polymerase starts transcription at the start point (+1) it is the first base transcribed as RNA.

- The characteristic nucleotide sequences of the prokaryotic promoter region (**as indicated in the coding strand in the 5` to 3 ` direction**) include:

- **TATA box:** It is formed of six nucleotides (TATAAT) and is located 10 bases upstream (i.e. usually occurs around base-10)to the start point (+1 point).**It determines where transcription starts.**
- **The (TTGACA) box:** this sequence is 35 bases upstream to the start point (located at -35 base i.e. centered about 35 bases to the left of the transcription start site) .**It determines the frequency of transcription**



Source: Murray RK, Bender DA, Botham KM, Kennelly PJ, Rodwell VW, Weil PA: *Harper's Illustrated Biochemistry, 29th Edition*: www.accessmedicine.com

Elongation:

- In prokaryotes only **one** type of RNA polymerase synthesizes the three types of prokaryotic RNA.
- The binding of RNA polymerase to DNA template produce **local unwinding** of the DNA double helix to expose the bases.
- The enzyme begins to synthesize RNA in the direction of 5\to 3\with the base sequence complementary to that of the DNA template strand. Sigma factor is released after initiation of transcription.
- The core enzyme moves along the DNA template uses ribonucleoside triphosphate (ATP, GTP, CTP& UTP) and releases pyrophosphate

- **Unlike DNA polymerase, RNA polymerase does not require a primer and has intrinsic helicase activity, therefore no separate enzyme is needed to unwind the DNA (in contrast to DNA polymerase).**
- **RNAP not only initiates RNA transcription, it also guides the nucleotides into position, facilitates attachment and elongation, has intrinsic proofreading (It doesn't not posses a proof reading feature as efficient as the DNA polymerase but it posses the capability of correct some misadded nucleotide as well) and replacement capabilities, and termination recognition capability.**

Termination:

- RNA polymerase recognizes a termination signal at the end of the DNA sequence to be transcribed (termination sequence). Then RNA polymerase stop transcription and releases RNA molecule.
- ***There are two mechanisms for transcription termination:***

1-Rho factor dependent termination (ATP dependent)

2-Rho factor independent termination (intrinsic termination)

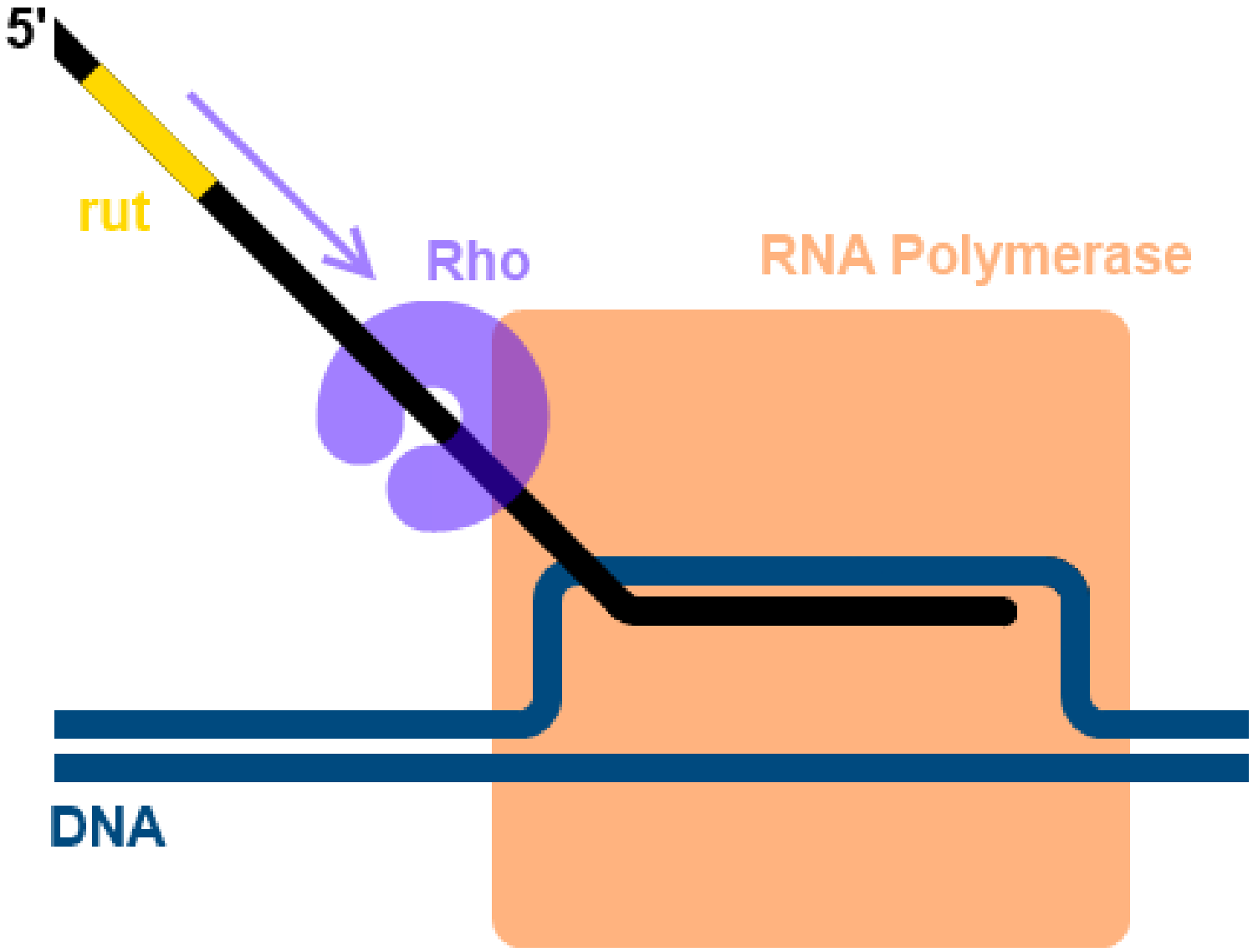
1-Rho-dependent termination:

- It uses a **termination factor called ρ factor** (rho factor) which is a protein that binds at a **rho utilization site** (*rut*) on the nascent RNA strand (cytosine-rich sequence) and runs along the mRNA towards the RNAP (in a 5'-3' direction).

(The rut serves as a mRNA loading site and as an activator for Rho)

- The rho protein is an **ATP dependent RNA-DNA helicase**
- when ρ-factor reaches the RNAP, it causes RNAP to dissociate from the DNA, terminating transcription.

- **Rho is able to catch up with the RNA polymerase. Contact between Rho and the RNA polymerase complex stimulates dissociation of the transcriptional complex through a mechanism involving allosteric effects of Rho on RNA polymerase.**


5'
rut
Rho
RNA Polymerase
DNA

2-Rho factor independent termination (intrinsic termination):

- The termination sequences is self-complementary sequences rich in GC that are present at the 3\ end of mRNA. These complementary bases join each other forming hairpin loop like structure that leads to dissociation of RNA from the DNA and release the RNA polymerase enzyme.

- **The termination site is characterized by the presence of two regions that are separated by a few bases (4-6) in the form of a palindrome.**
- **DNA palindrome:** A **palindromic sequence** is a nucleic acid sequence on double-stranded DNA wherein reading 5‘ to 3‘ forward on one strand matches the sequence reading 5' to 3' on the complementary strand with which it forms a double helix. **(form symmetrical inverted repeat).**
- **When the RNA is created, the inverted repeates can loop back on themselves to form a hairpin loop, which acts as a termination signal.**

Palindrome dependent Termination

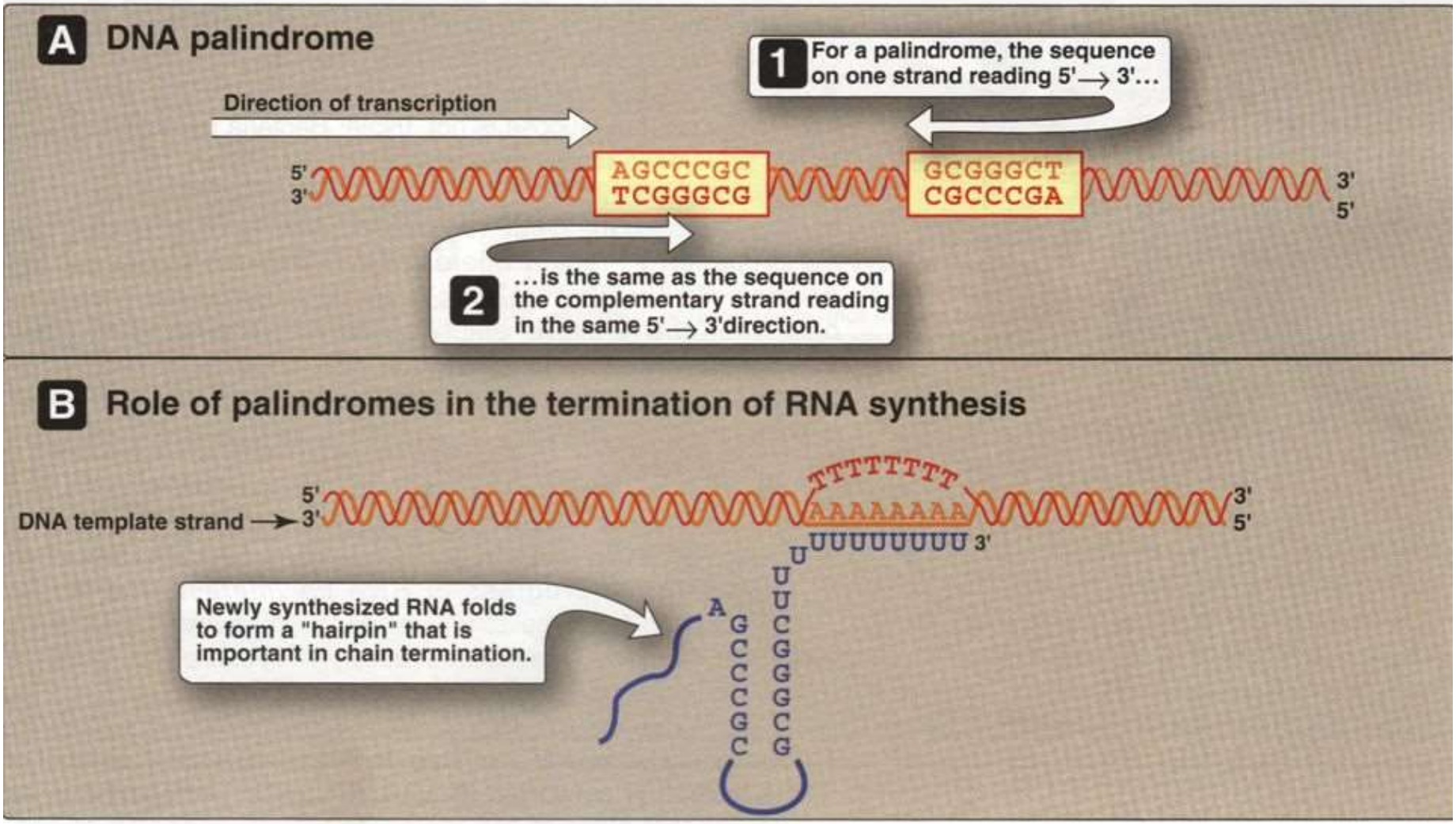
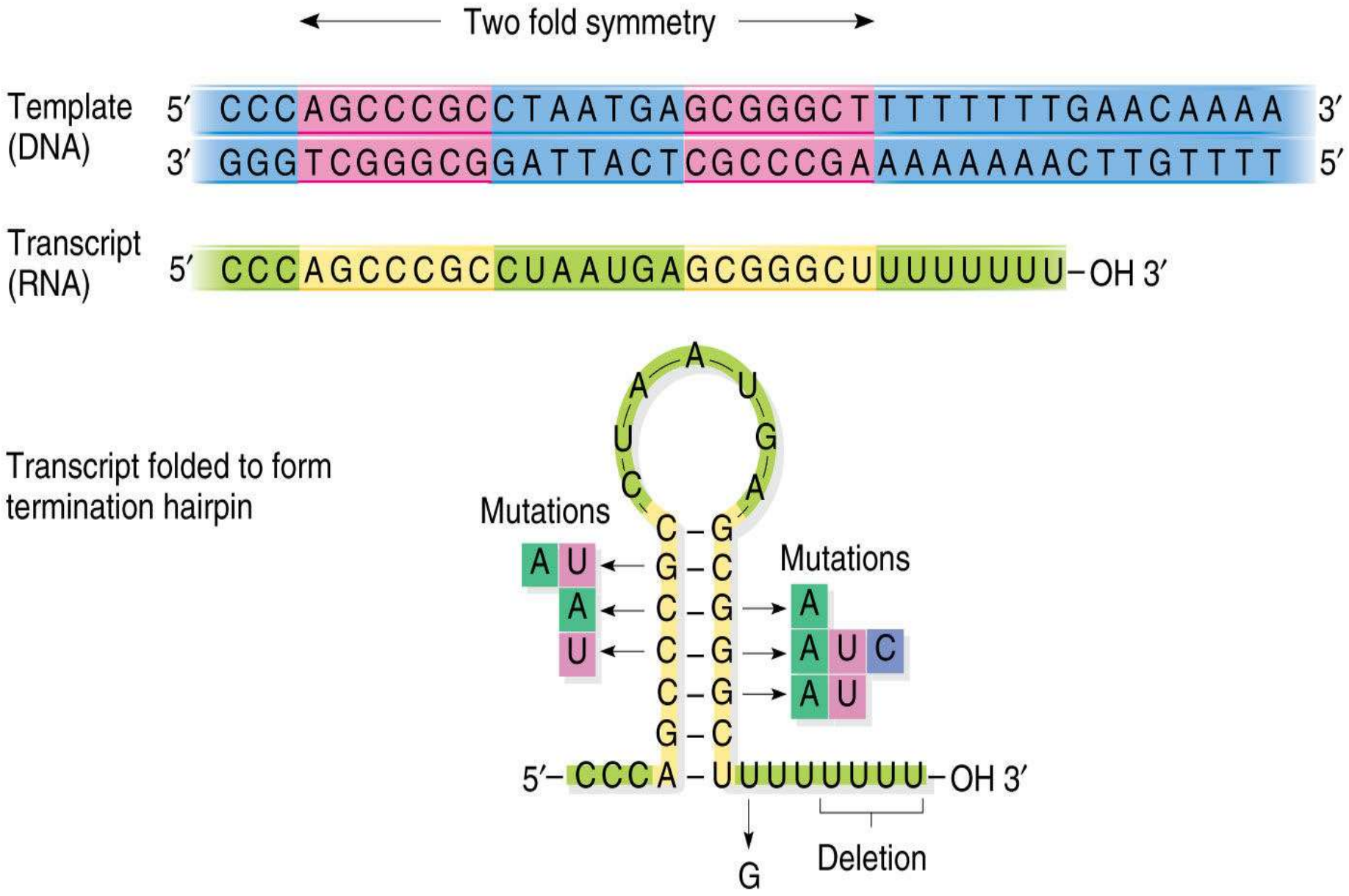


Figure 30.9

Rho-independent termination of transcription. A. An example of a palindrome in double-stranded DNA. B. A transcribed DNA palindrome codes for RNA that can form a hairpin turn.



Regulation of prokaryotic gene expression

❖ ***There are two types of gene according to their expression:***

1-Constitutive genes:

- These genes are not regulated ,
- They code their protein products which are required for the basic cellular functions and so, they are continuously expressed at a low rate;
- They are also known as "**housekeeping**" genes.

2-Inducible genes:

- They express their protein product only in the presence of an **inducer or derepressor**.
- They are negatively regulated by specific proteins termed **repressors**.
- **The inducer produces inactivation of the repressor.**

- **In bacteria**, the **structural genes** that code for the enzymes of a metabolic pathway are often found grouped together on the chromosome together with the **regulatory genes** that determine their transcription as a single long piece of mRNA.
- **This entire package is referred to as an operon.**
- So operon is a linear array of the genes that are involved in a metabolic pathway.
- One of the best understood examples is the **lactose operon of E. coli**.

The lactose operon of E coli
(as a model of prokaryotic gene regulation)

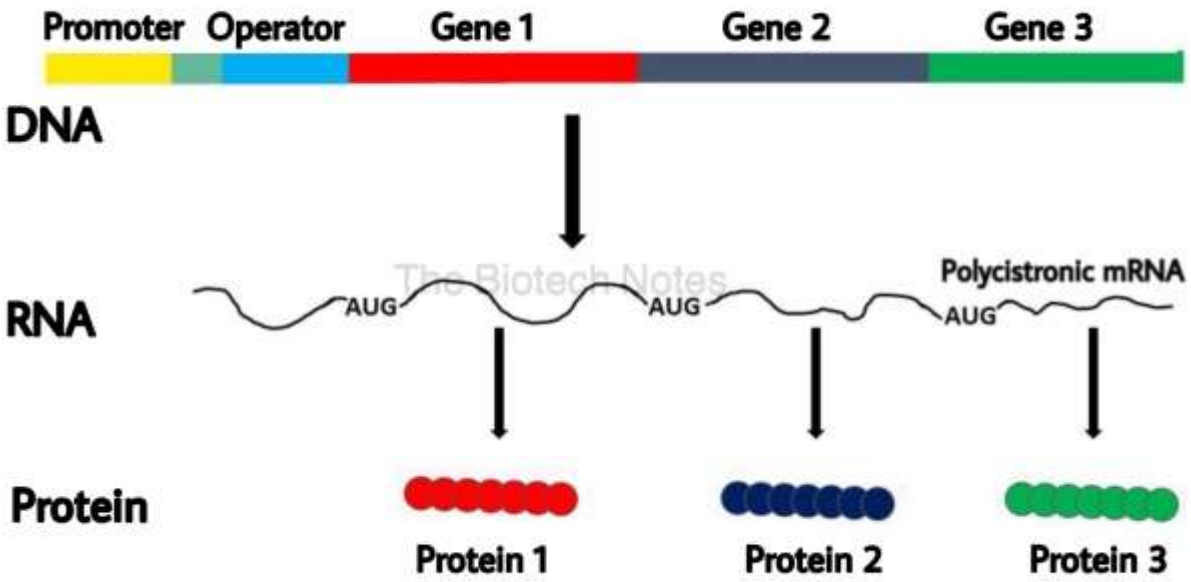
- Lac operon contains the genes **responsible for lactose metabolism by E coli bacteria** when lactose is available to the cell but glucose is not.
- [**Note:** Bacteria use glucose as a fuel in preference to any other sugar.]

The Lac operon of E-coli is formed of:

1-Structural genes: They are three linked **inducible** genes as follows:

- ***-Lac Z gene***; encodes **β-galactosidase** that hydrolyses lactose to glucose and galactose.
- ***-Lac Y gene:*** encodes **permease** enzyme that allows lactose transport into the cells.
- ***-Lac A gene:*** encodes **thiogalactoside transacetylase** of unknown function.
- *The lacA gene encodes thiogalactoside transacetylase, which rids the cell of toxic thiogalactosides that also get transported in by lacY. (i.e. cellular detoxification)*

- **The three linked genes are transcribed into one large polycistronic mRNA molecule that contains multiple independent translation start and stop codons for each cistron.**
- **Thus, each protein is translated separately and they are not processed from a single large precursor protein.**


Promoter
Operator
Gene 1
Gene 2
Gene 3
DNA
RNA
AUG
AUG
AUG
Polycistronic mRNA
Protein
Protein 1
Protein 2
Protein 3

- **A gene** is a part of DNA that gets transcribed into an RNA(mRNA, tRNA, rRNA or any other form of rna).
- **Cistron** is a part of mRNA that begins with a start codon, ends with a stop codon and in between these codons lies the series of codon which code for a single polypeptide.
- You can say that cistron is the part of mRNA that gets translated into polypeptide.

2-Regulatory gene or lac I gene: It is **constitutive gene** and codes for the regulatory protein (**Lac repressor**).

3-Operator region: At which Lac repressor binds and inhibits gene expression.

4-A single common Promoter: It is the site where RNA polymerase binds to it and start transcription of the structural gene.

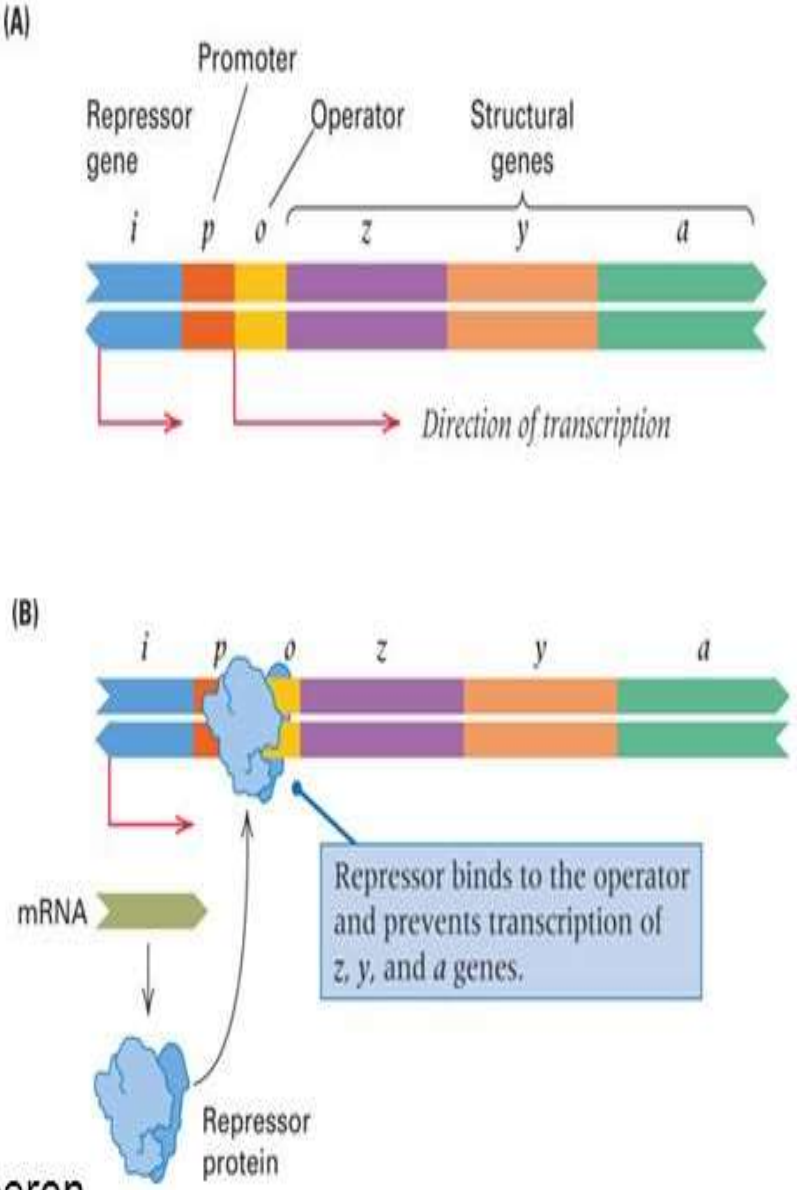
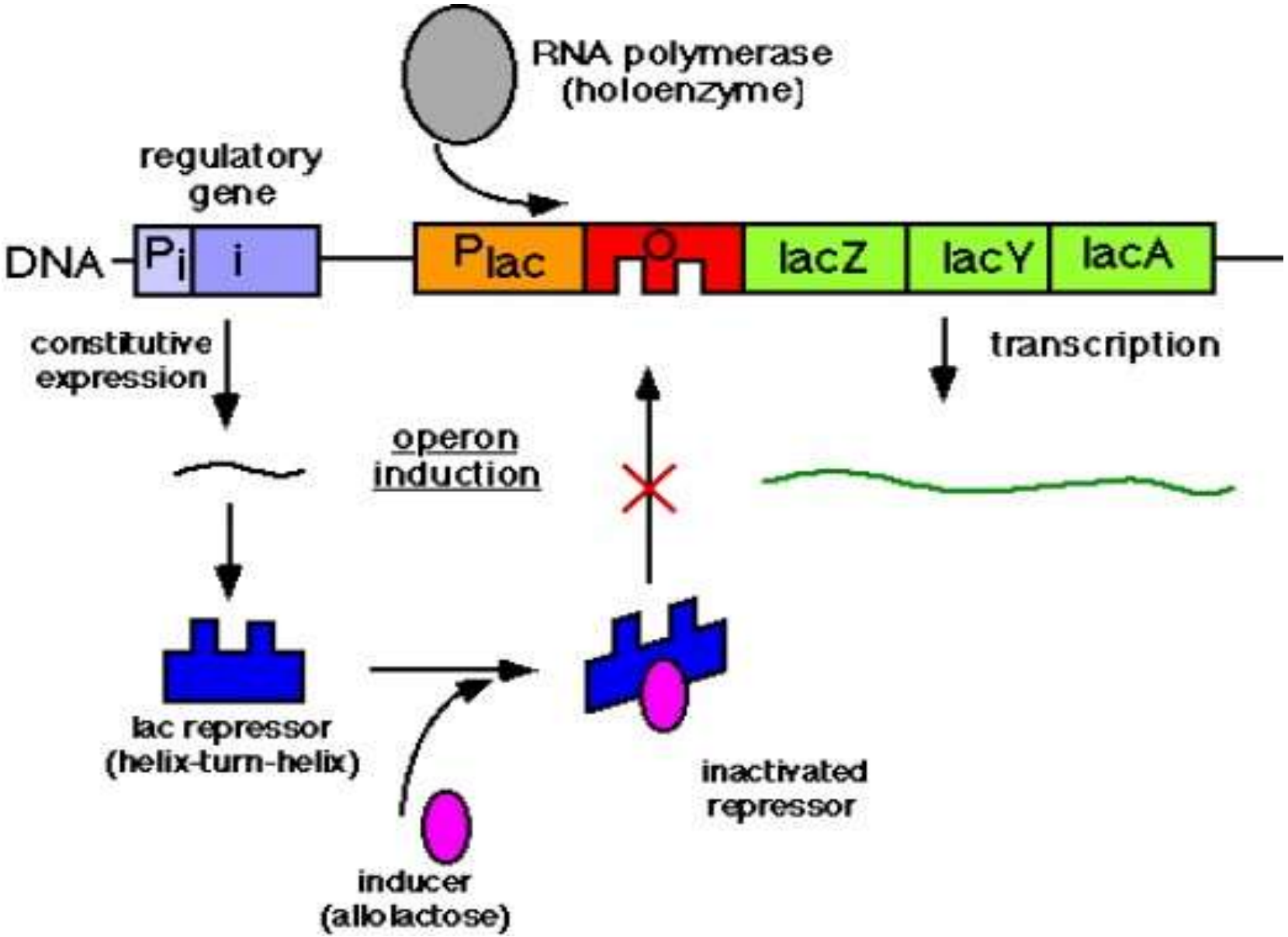


Figure 9.4a/b: *lac* operon


RNA polymerase
(holoenzyme)
regulatory
gene
DNA
Pi
i
Plac
O
lacZ
lacY
lacA
constitutive
expression
transcription
operon
induction
lac repressor
(helix-turn-helix)
inactivated
repressor
inducer
(allolactose)

When glucose is the only sugar available (Gene repression):

- In this case, the lac operon is repressed (**turned off**).
- Repression is mediated by the repressor protein binding to the operator site, **which is downstream of the promoter region.**
- Binding of the repressor interferes with the progress of RNA polymerase, and blocks transcription of the structural genes.
- This is an example of **negative regulation**.