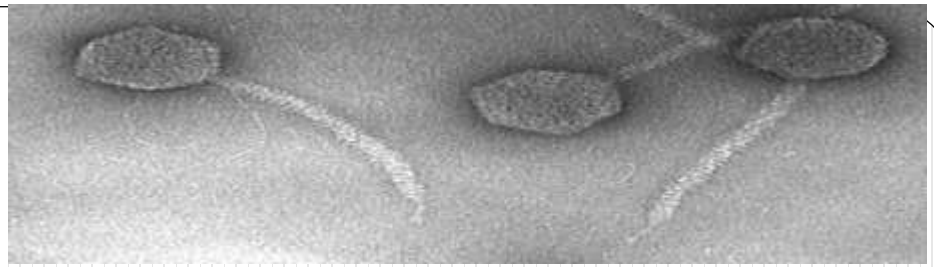
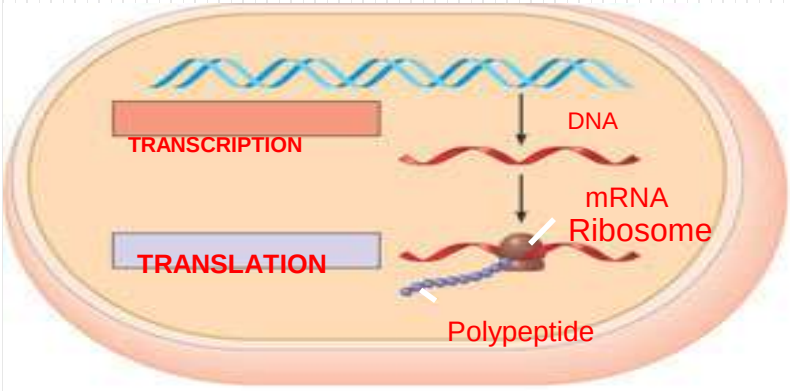


GENE REGULATION-PROKARYOTICS

Dr. Vasant C Gosai



INTRODUCTION

- The fundamental problem of chemical physiology and of embryology is to understand why tissue cells do not all express, all the time, all the potentialities inherent in their genome.

—François Jacob and Jacques Monod..

- Of the 4,000 or so genes in the typical bacterial genome, or the perhaps 1,00,000 genes in the human genome, only a fraction are expressed in a cell at any given time.

Significance of gene Expression

Regulated expression of genes is required for

- Adaptation,
- Differentiation and
- Development

1)ADAPTATION

- Organisms adapt to environmental changes by altering gene expression.

a) Bacteria are highly versatile and responsive organisms: the rate of synthesis of some proteins in bacteria may vary more than a 1000-fold in response to the supply of nutrients or to environmental challenges.

b) Cells of multicellular organisms also respond to varying conditions.

c) Such cells exposed to hormones and growth factors change substantially in shape, growth rate, and other characteristics.

differentiation and development

- The genetic information present in each somatic cell of a metazoan organism is practically **identical.**

- The exceptions in the genetic information are found in those few cells that have amplified or rearranged genes in order to perform specialized cellular functions.
- As the high cost of protein synthesis, regulation of gene expression is essential to making optimal use of the energy.

PRINCIPLE OF GENE REGULATION

- **HOUSEKEEPING GENES:** Genes for products that are required at all times.
- Ex : enzymes of Glycolysis are synthesised by all cells .

- **CONSTITUTIVE GENE EXPRESSION :**Unvarying expression of a gene.

- **REGULATED GENE EXPRESSION:** For other gene products, cellular levels rise and fall in response to molecular signals.

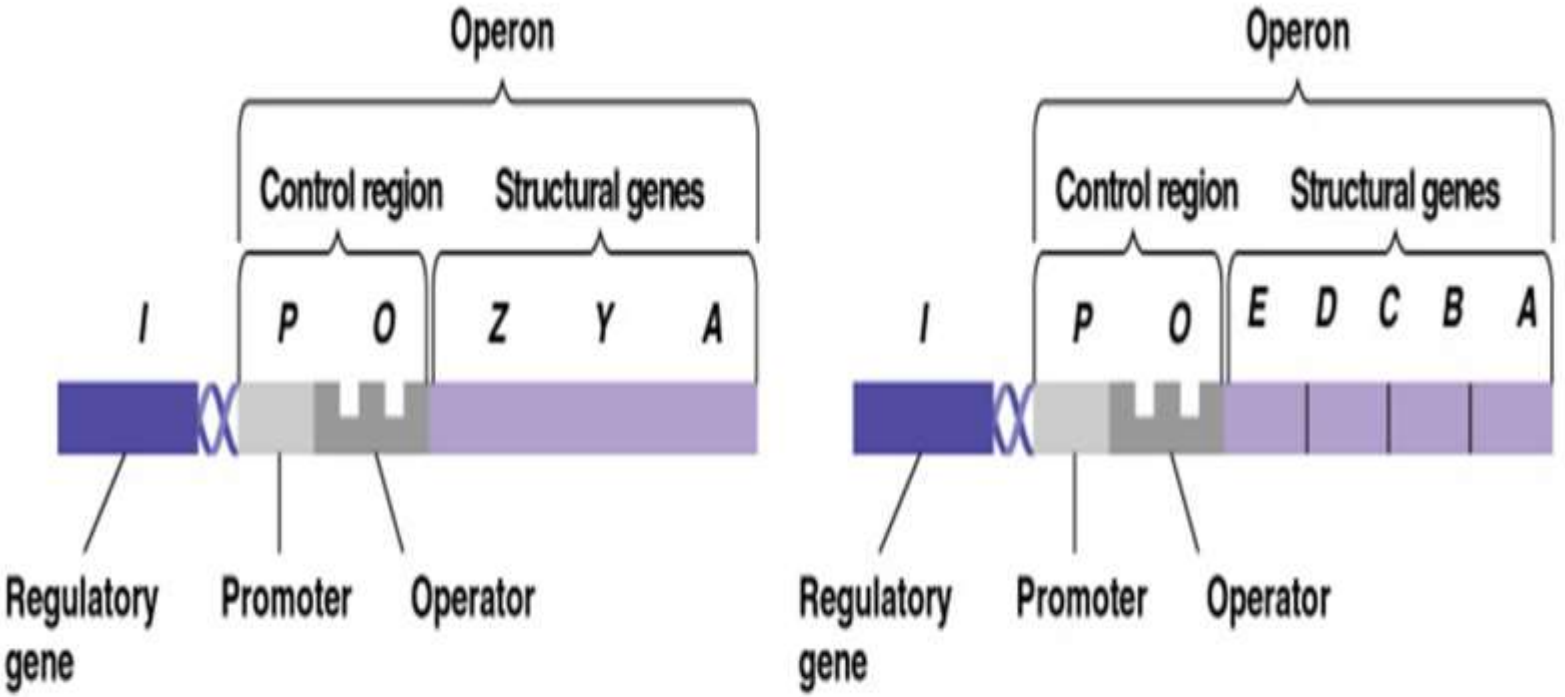
- **INDUCTION :** Induction is turning “on” the switch of the gene.
- A phenomena of increased synthesis of protein or enzyme in response to certain signal.
- Such enzymes are said to be **Inducible.**
- Signals are called **Inducers.**

- **REPRESSION** is turning “off” the gene expression.
- In prokaryotes,the genes involved in a metabolic pathway are often present in a linear array called an **operon** **eg,** the *lac operon.*
- An operon can be regulated by a single promoter or regulatory region.

- **CISTRON :** The smallest unit of genetic expression
- The cistron is the genetic unit coding for protein molecule.
- A single mRNA that encodes more than one type of translated protein is referred to as a **polycistronic mRNA.**

- For example, the polycistronic *lac operon* mRNA is translated into three separate proteins.
- Operons and polycistronic mRNAs are common in bacteria but not in eukaryotes
- Transcription is mediated and regulated by **protein-DNA interactions**, especially those involving the protein components of RNA polymerase.

GENERAL STRUCTURE OF AN OPERON



1 **Structure of the operon.** The operon consists of the promoter (*P*), and operator (*O*) sites, and structural genes which code for the protein. The operon is regulated by the product of the regulatory gene (*I*).

RNA POLYMARASE BINDS TO DNA AT PROMOTERS

- **Promoter site**: near points at RNA synthesis begins.
- RNA polymerases bind to DNA and initiate transcription at promoters.

- **TRANSCRIPTION INITIATION IS REGULATED BY PROTEIN THAT BIND TO OR NEAR PROMOTER**

- At least three type of protein regulate transcription initiation by RNA polymerase.

1. **SPECIFICITY FACTORS**: alter specificity of RNA polymerase for given promoter or to the set of promoter.

   e.g. the σ subunit of the E.coli RNA Polymerase

2. **REPRESSORS :** impede access of RNA Polymerase to the promoter
3. **ACTIVATORS :** enhance the RNA polymerase-promoter interaction.

- **Repressors** bind to specific sites on the DNA. These sites, called **OPERATORS,** are generally near a promoter.
- RNA polymerase movement is **blocked** when the repressor is present.
- Regulation by repressor protein that blocks transcription is **NEGATIVE REGULATION.**

Types of Gene Expression

There are **mainly two types of gene expression and regulation:**

a. Positive regulation
b. Negative regulation.

a. Positive regulation:

- ***When the expression of genetic*** **information** is quantitatively ***increased by the*** presence of specific regulatory element, it is called as positive regulation.
- The element or molecule mediating positive regulation is called **positive regulator.**
- **Ex : Lac Operon**

b. *Negative regulation:*

- ***When the expression of genetic*** information is ***decreased by the presence of a specific*** regulatory element, it is called as negative regulation.
- The element or molecule mediating the negative regulation is called a **negative regulator.**
- **Ex : Tryptophan Operon**

REGULATION IN PROKARYOTES

1. LAC OPERON
2. TRYPTOPHAN OPERON BY ATTENUATION
3. ARABINOSE OPERON
4. SOS RESPONSE
5. RIBOSOMAL PROTEIN SYNTHESIS COORDINATED WITH rRNA SYNTHESIS (STRINGENT RESPONSE)
6. REGULATION IN VIRUS (LAMDA PHAGE)

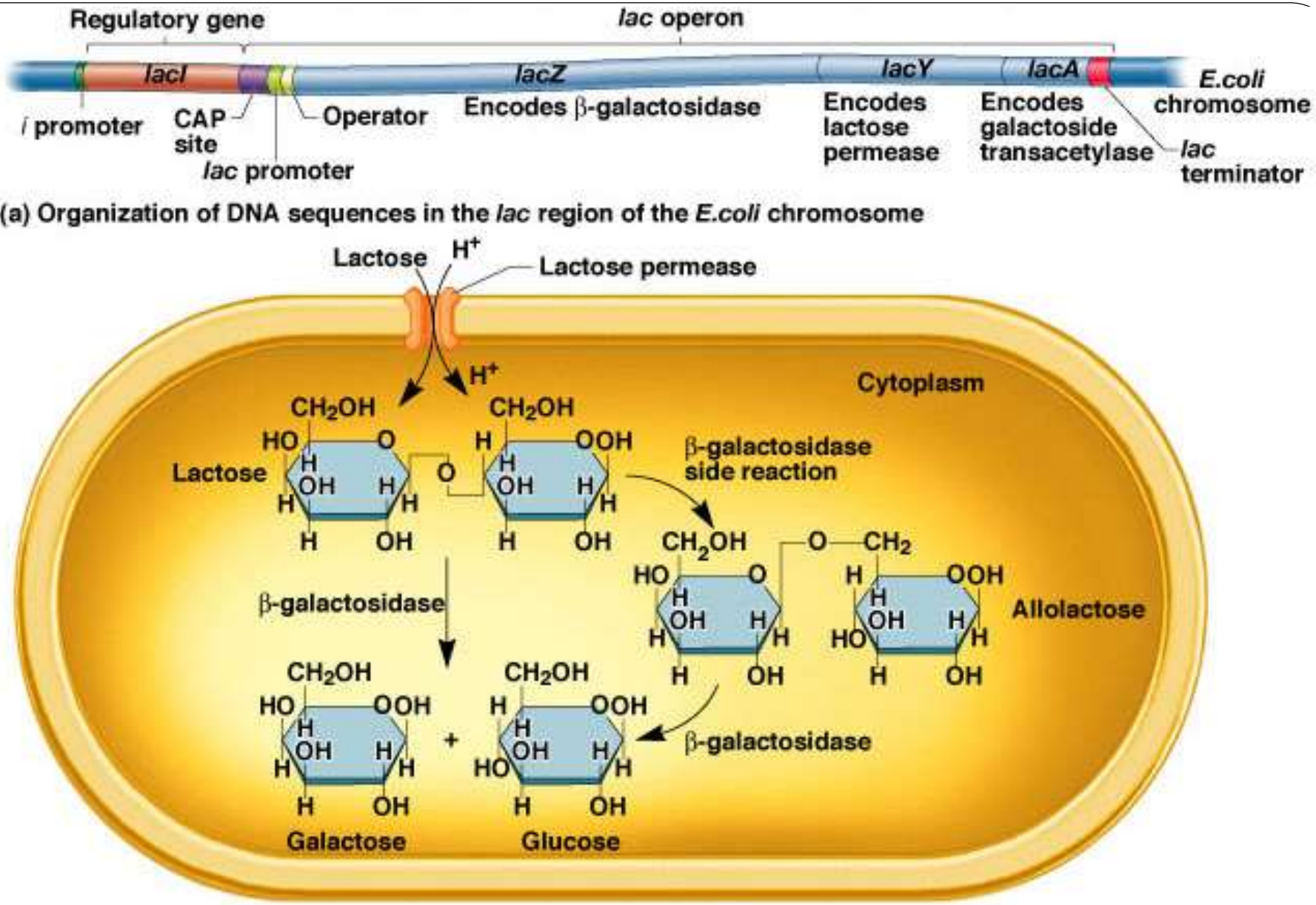
LAC OPERON MODEL

- Jacob and Monod in 1961 described their Operon model in a classic paper.
- Their hypothesis was to a large extent based on observations on the regulation of lactose metabolism by the intestinal bacterium E coli.

François Jacob

Jacques Monod, 1910–1976

- Bacteria such as *E. coli* usually rely on glucose as their source of carbon and energy.
- When glucose is scarce, *E. coli* can use lactose as their carbon source .
- An essential enzyme in the metabolism of lactose is β-*galactosidase*, which hydrolyzes lactose into galactose and glucose.


Regulatory gene
lac operon
lacI
lacZ
lacY
lacA
E.coli chromosome
i promoter
CAP site
Operator
Encodes β-galactosidase
Encodes lactose permease
Encodes galactoside transacetylase
lac terminator
lac promoter
(a) Organization of DNA sequences in the lac region of the E.coli chromosome
Lactose
H+
Lactose permease
Cytoplasm
Lactose
β-galactosidase side reaction
β-galactosidase
Allolactose
β-galactosidase
Galactose
Glucose
(b) Functions of lactose permease and β-galactosidase

- The lactose (lac) operon codes for three proteins involved in the catabolism of the disaccharide, lactose :

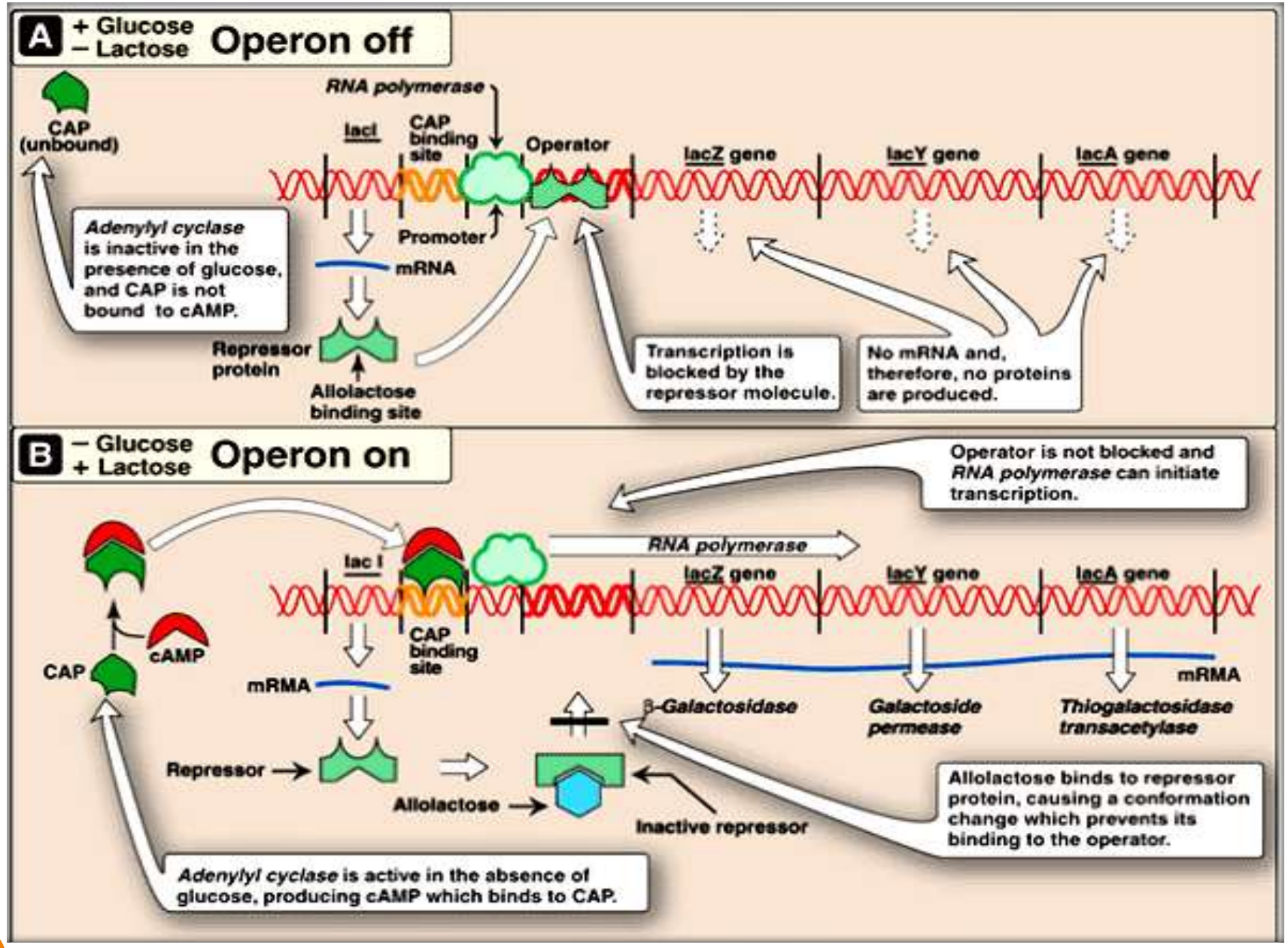
I. lacZ gene codes for **β-galactosidase**, which hydrolyzes lactose to galactose and glucose

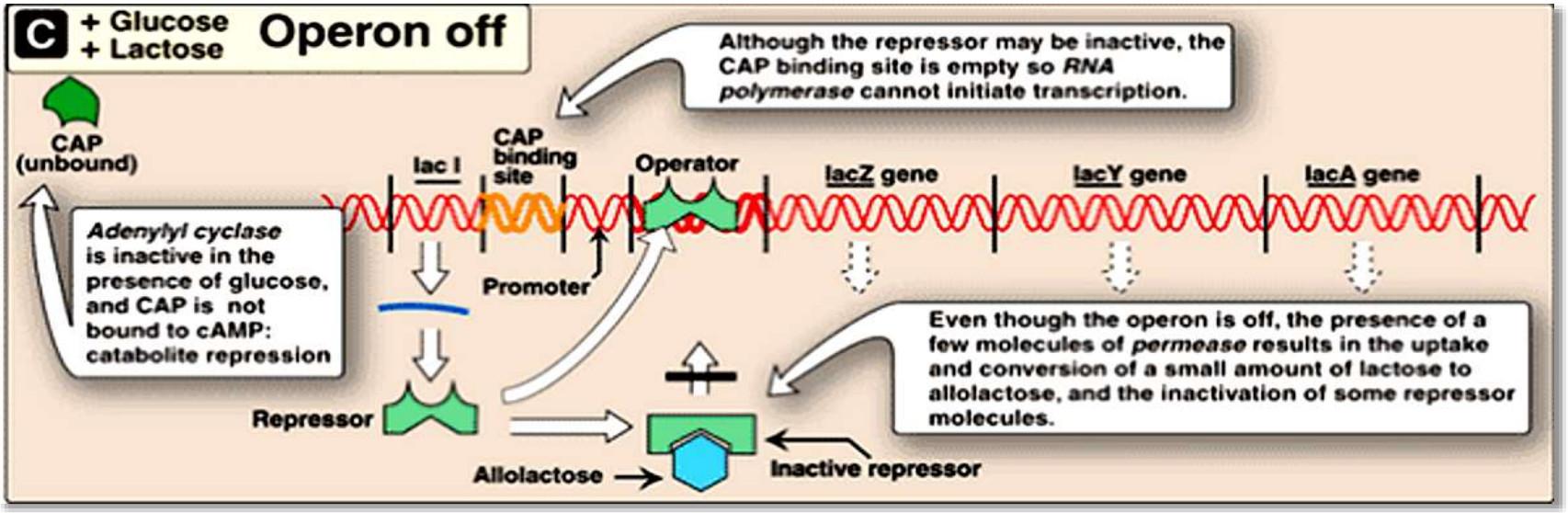
II. The lacY gene, which codes for a **permease** that facilitates the movement of lactose into the cell.

III. lacA gene that codes for **thiogalactoside transacetylase** whose exact physiologic function is unknown.

- All of these proteins are produced when lactose is available to cell.
- The regulatory portion of the operon is upstream of the three structural genes, and consists of the **promoter (P)** region where RNA polymerase binds, and two additional sites, the **operator (O)** site and the **CAP site**, where regulatory proteins bind

- The lacZ, lacY, and lacA genes are expressed only when the O site is empty.
- The CAP site is bound by a complex of cyclic adenosine monophosphate and the catabolite gene activator protein or CAP.
- A regulatory gene, the ***lacI*** gene, codes for the repressor protein that binds to the operator site.


A + Glucose – Lactose Operon off
CAP (unbound)
RNA polymerase
lacI
CAP binding site
Operator
lacZ gene
lacY gene
lacA gene
Promoter
mRNA
Adenylyl cyclase is inactive in the presence of glucose, and CAP is not bound to cAMP.
Repressor protein
Allolactose binding site
Transcription is blocked by the repressor molecule.
No mRNA and, therefore, no proteins are produced.
B – Glucose + Lactose Operon on
Operator is not blocked and RNA polymerase can initiate transcription.
RNA polymerase
lac I
lacZ gene
lacY gene
lacA gene
CAP
cAMP
CAP binding site
mRMA
mRMA
β-Galactosidase
Galactoside permease
Thiogalactosidase transacetylase
Repressor
Allolactose
Inactive repressor
Allolactose binds to repressor protein, causing a conformation change which prevents its binding to the operator.
Adenylyl cyclase is active in the absence of glucose, producing cAMP which binds to CAP.



- When glucose is the **only sugar** available: In this case, the lac operon is repressed (turned off).
- 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** .

- When only lactose is available: In this case, the lac operon is induced (expressed or turned on).
- A small amount of lactose is converted to an isomer, **allolactose.**
- **This compound is an inducer** that binds to the repressor protein, changing its conformation so that it can no longer bind to the operator.

- In the absence of glucose, adenylyl cyclase is active, and sufficient quantities of cAMP are made and bind to the CAP protein.
- The cAMP–CAP complex binds to the CAP-binding site, causing RNA polymerase to more efficiently initiate transcription at the promoter site .
- This is an **example of positive regulation**.

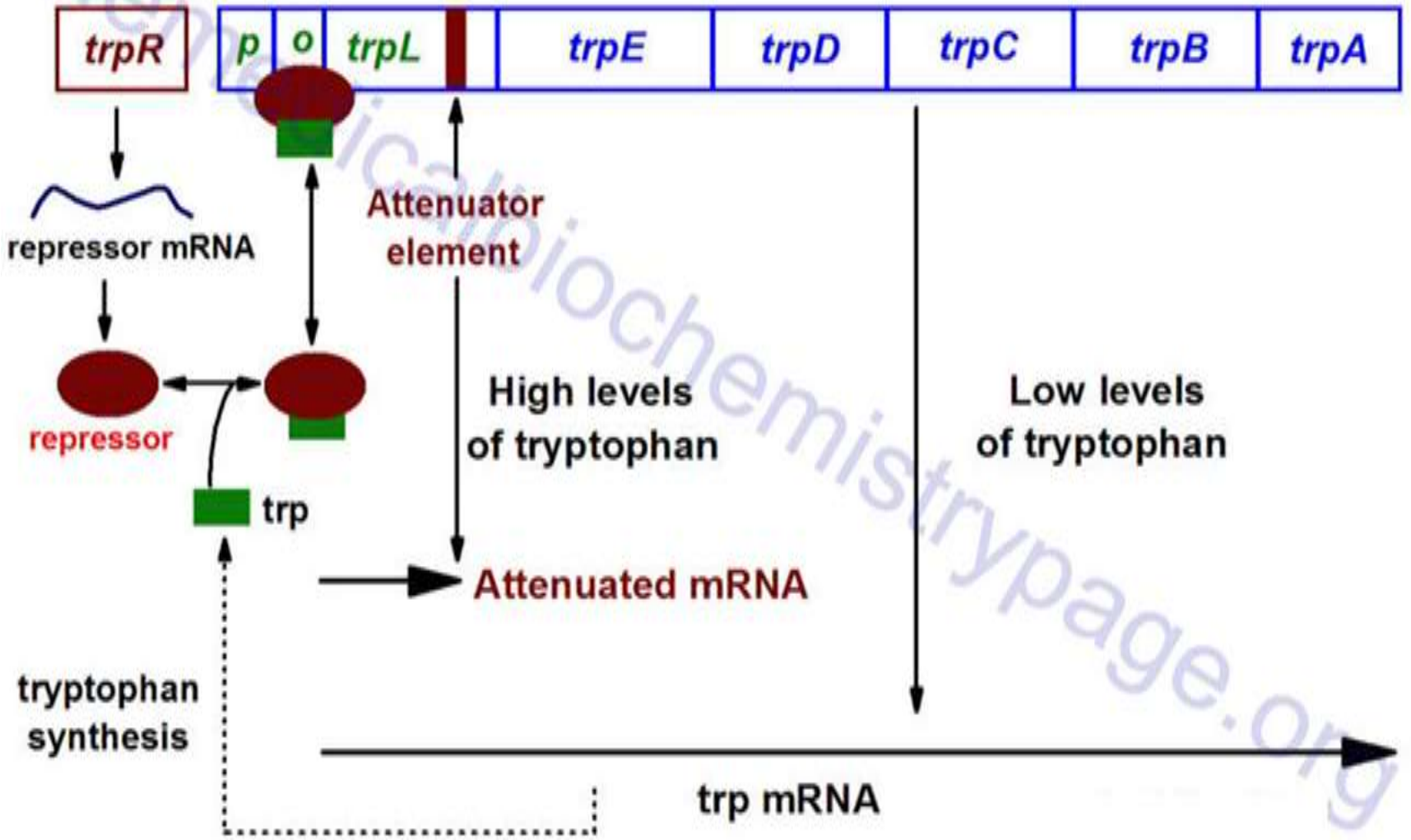
- When both glucose and lactose are available: In this case, transcription of the lac operon is negligible, even if lactose is present at a high concentration.
- Adenylyl cyclase is deactivated in the presence of glucose—a process known as **catabolite repression.**
- So no cAMP–CAP complex forms and the CAP-binding site remains empty.
- RNA polymerase is, therefore, unable to effectively initiate transcription, even though the repressor may not be bound to the operator region.

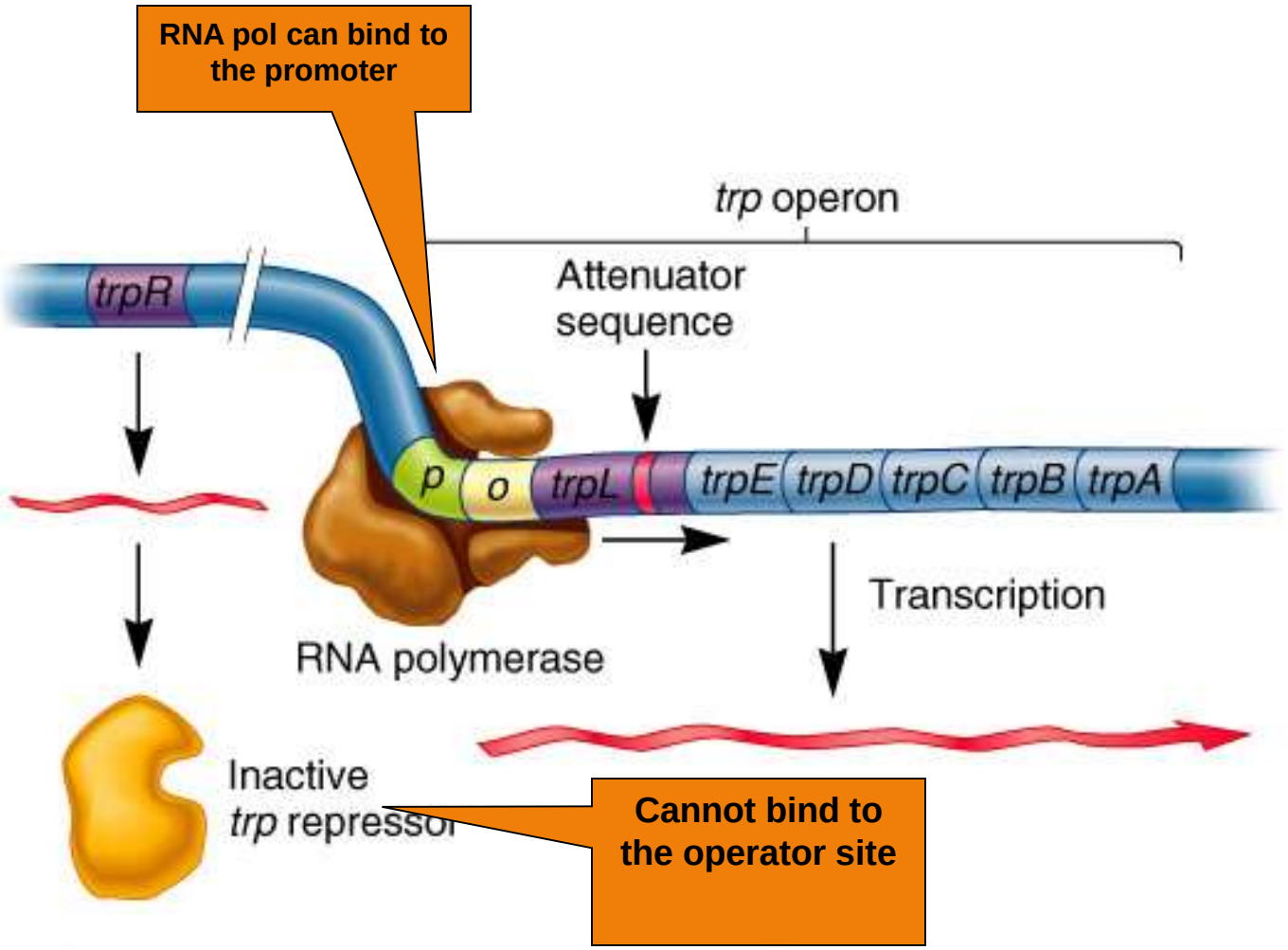
TRYPTOPHAN OPERON

- The tryptophan (trp) operon codes for five proteins that are required for the synthesis of the amino acid, tryptophan.
- This operon contains five structural genes:
  - trp E,
  - trp D,
  - trp C,
  - trp B, and
  - trp A, which encodes tryptophan synthetase.

- E- D codes for Anthranilate synthase I
- C -Codes for N 5’ Antharanilate Isomease (Indole 3- glycerol phophate synthetase)
- B-Codes for Tryptophan syntheatse (Beta subunit)
- A –Codes for Tryptophan syntheatse (Alpha subunit)

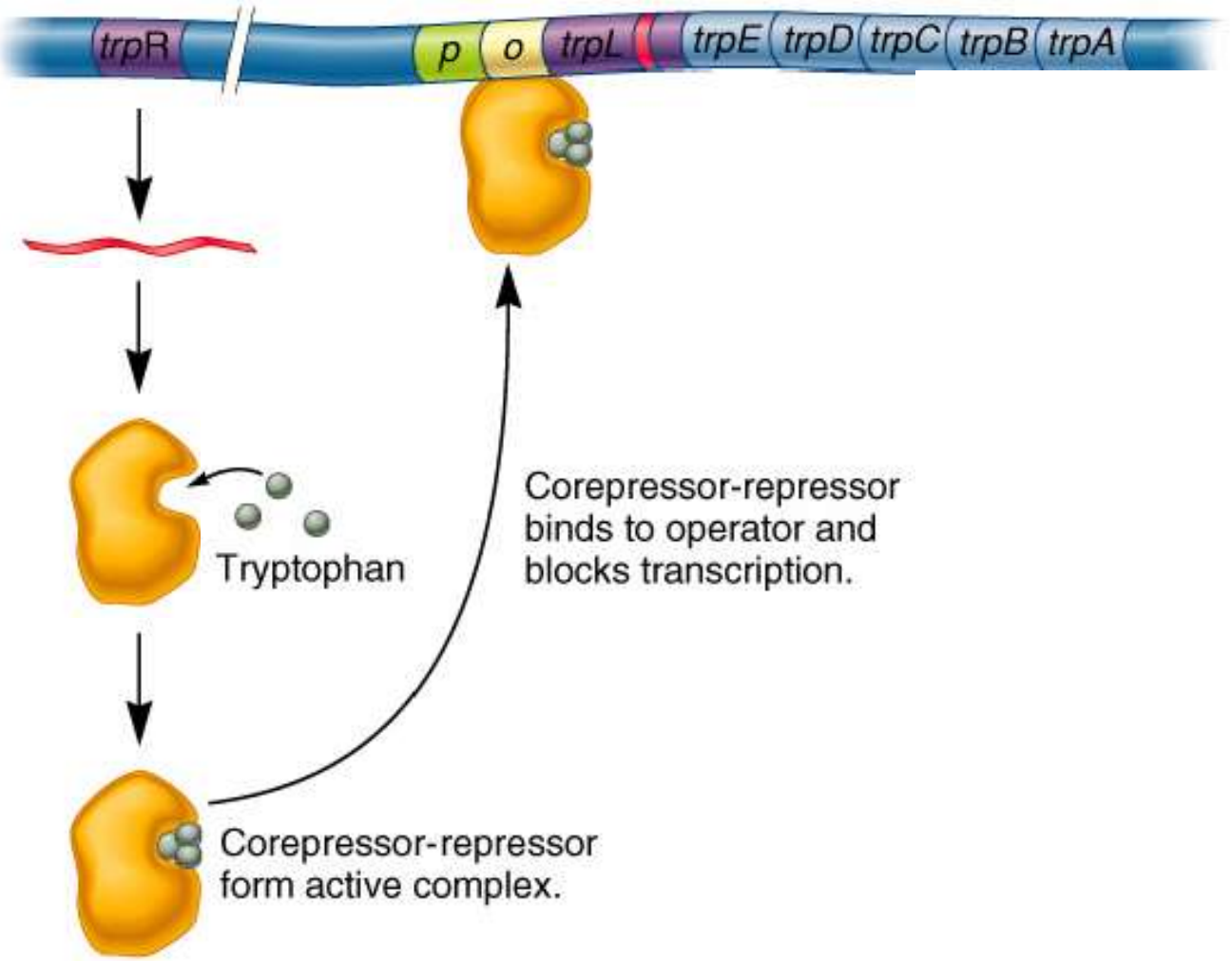
Structure of the *trp* Operon





(a) Low tryptophan levels, transcription of the entire *trp* operon occurs

Organization of the *trp* operon and regulation via the *trp* repressor protein



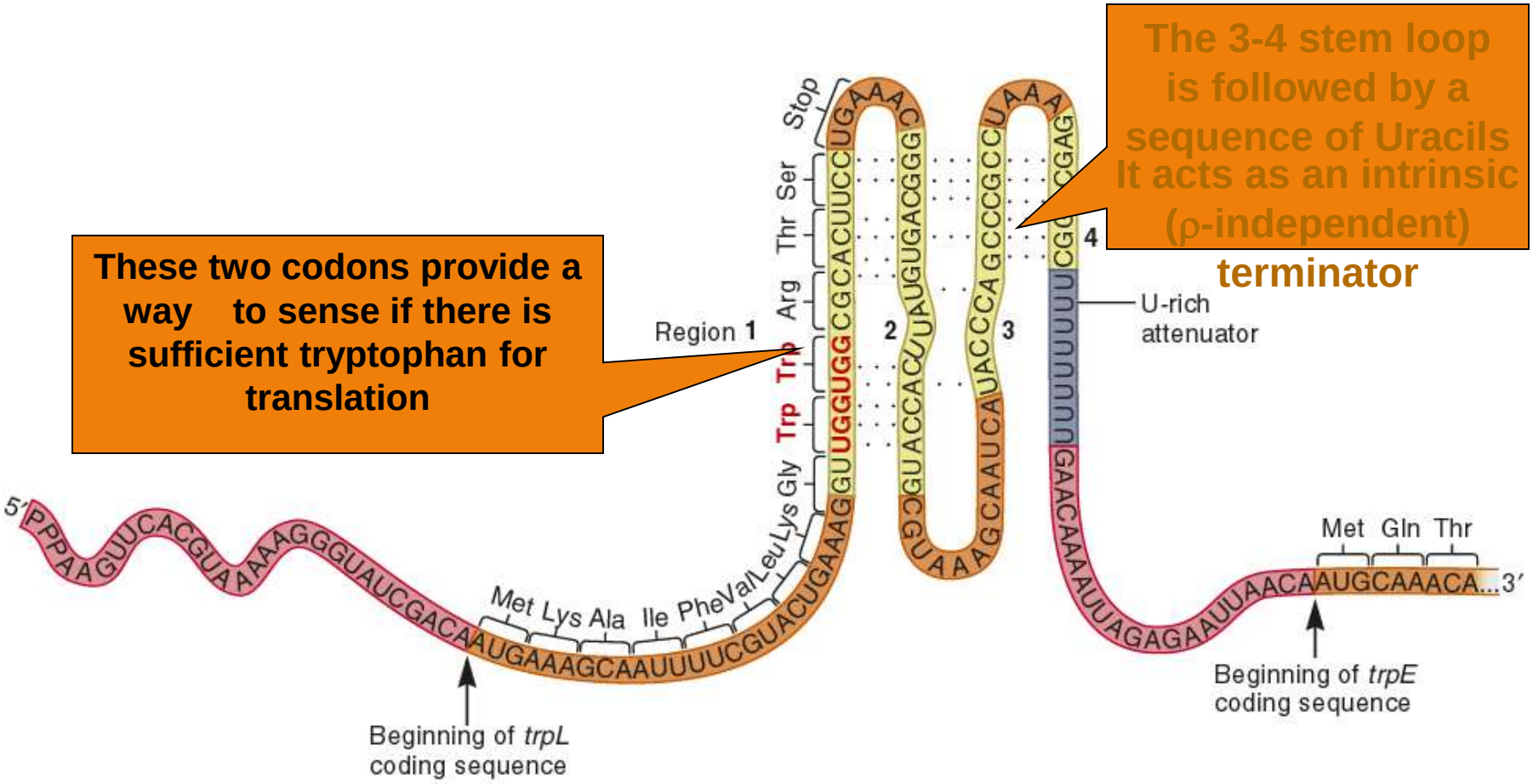
(b) High tryptophan levels, repression occurs

Organization of the *trp* operon and regulation via the *trp* repressor protein

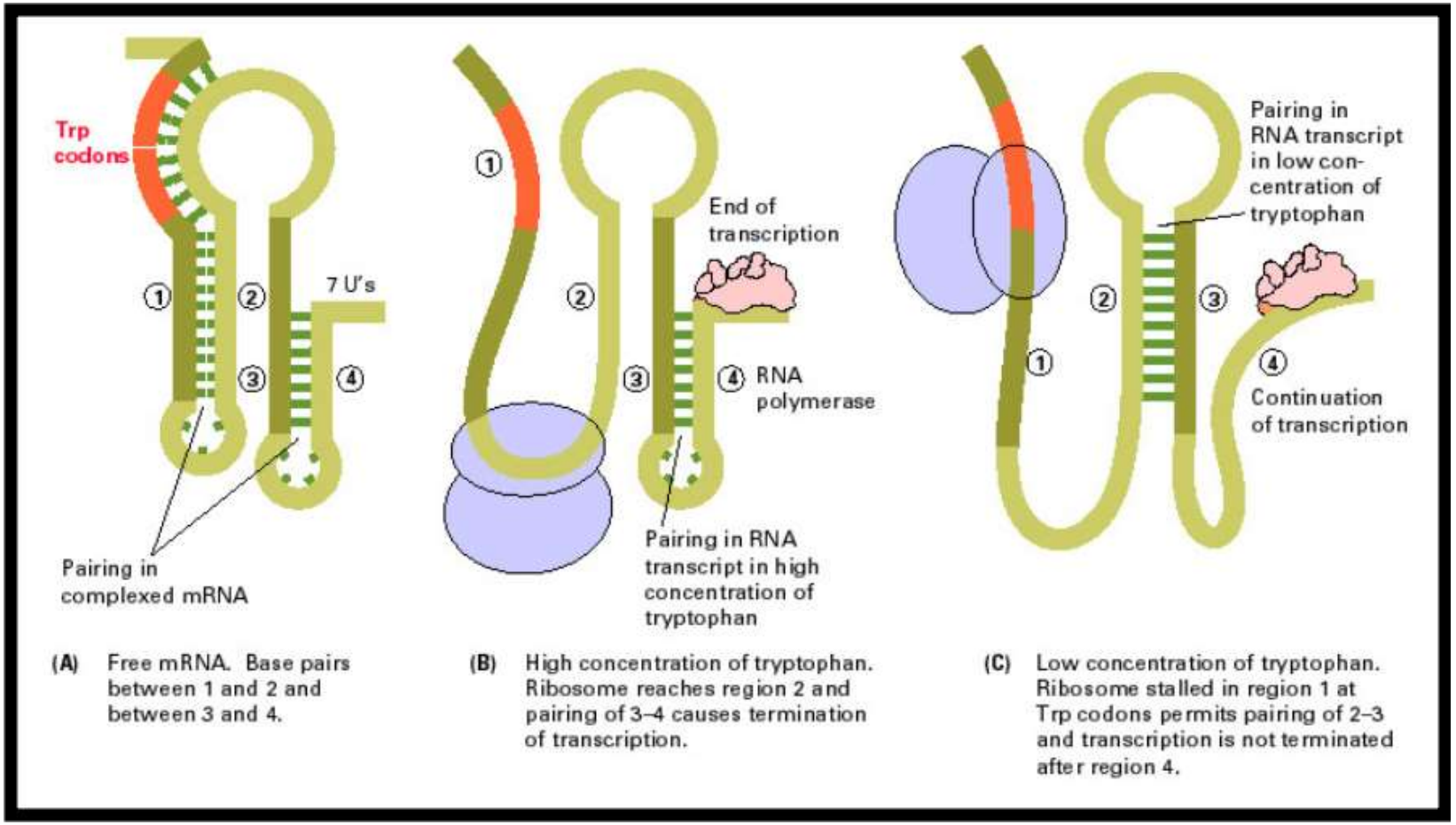
- **Negative control** includes trp itself binding to the repressor protein and facilitating the binding of the repressor to the operator.
- Repression by trp is not always complete, however, the trp operon is also regulated by a process known **as attenuation**.
- With attenuation, transcription is initiated but is terminated well before completion.

- If trp is **plentiful**, transcription initiation that escaped repression by trp is attenuated (stopped) by the formation at the 5′-end of the mRNA of a hairpin (stem-loop) structure like that seen in ρ-independent termination
- Transcription and translation are coupled processes in prokaryotes , therefore attenuation also results in the formation of a truncated, nonfunctional peptide product that is rapidly degraded.

- Region 2 is complementary to regions 1 and 3
- Region 3 is complementary to regions 2 and 4
  - Therefore several stem-loops structures are possible



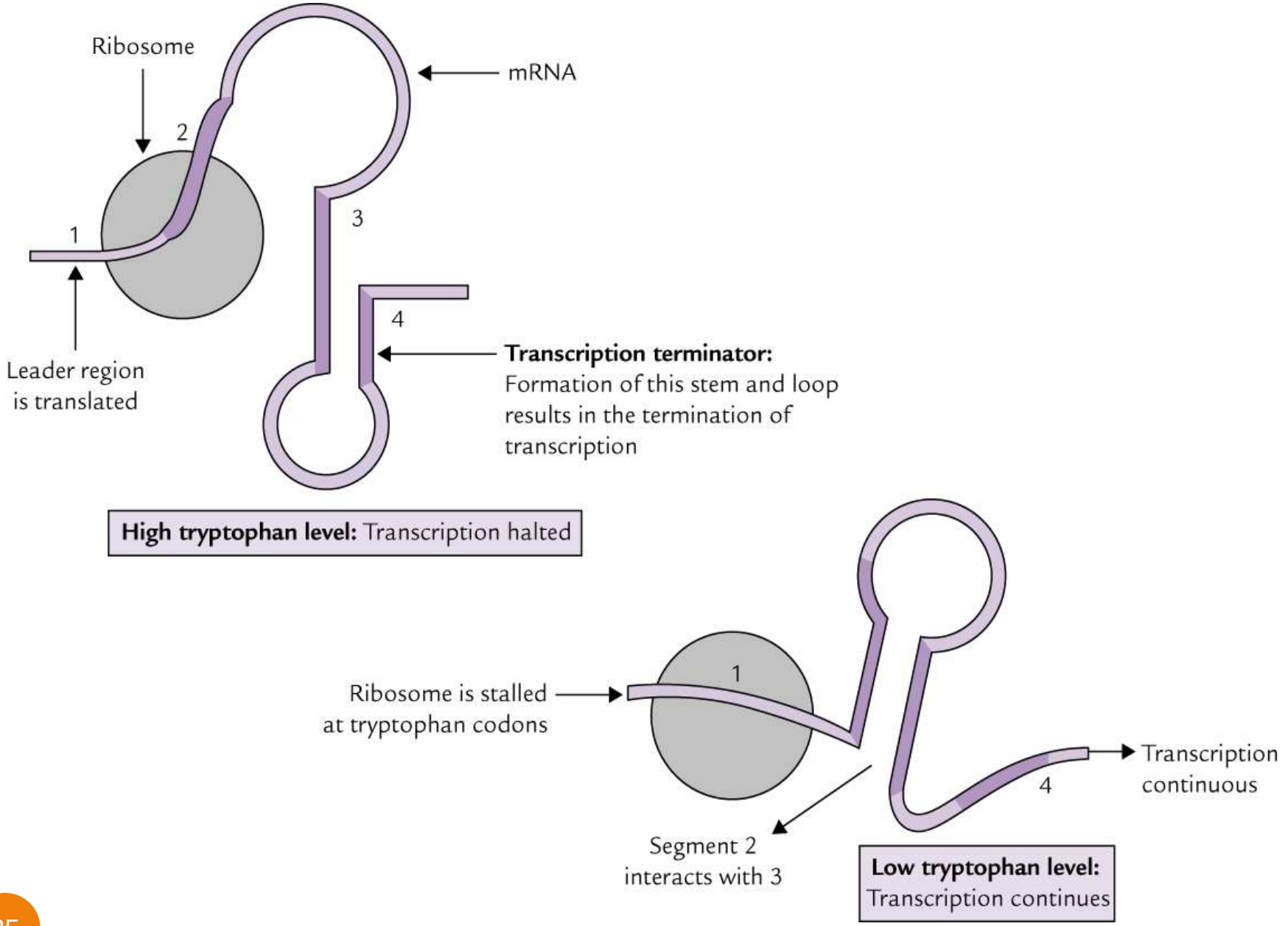
Sequence of the *trpL* mRNA produced during attenuation



(A) Free mRNA. Base pairs between 1 and 2 and between 3 and 4.

(B) High concentration of tryptophan. Ribosome reaches region 2 and pairing of 3–4 causes termination of transcription.

(C) Low concentration of tryptophan. Ribosome stalled in region 1 at Trp codons permits pairing of 2–3 and transcription is not terminated after region 4.


Ribosome
mRNA
2
1
3
4
Leader region
is translated
Transcription terminator:
Formation of this stem and loop
results in the termination of
transcription
High tryptophan level: Transcription halted
Ribosome is stalled
at tryptophan codons
1
4
Transcription
continuous
Segment 2
interacts with 3
Low tryptophan level:
Transcription continues

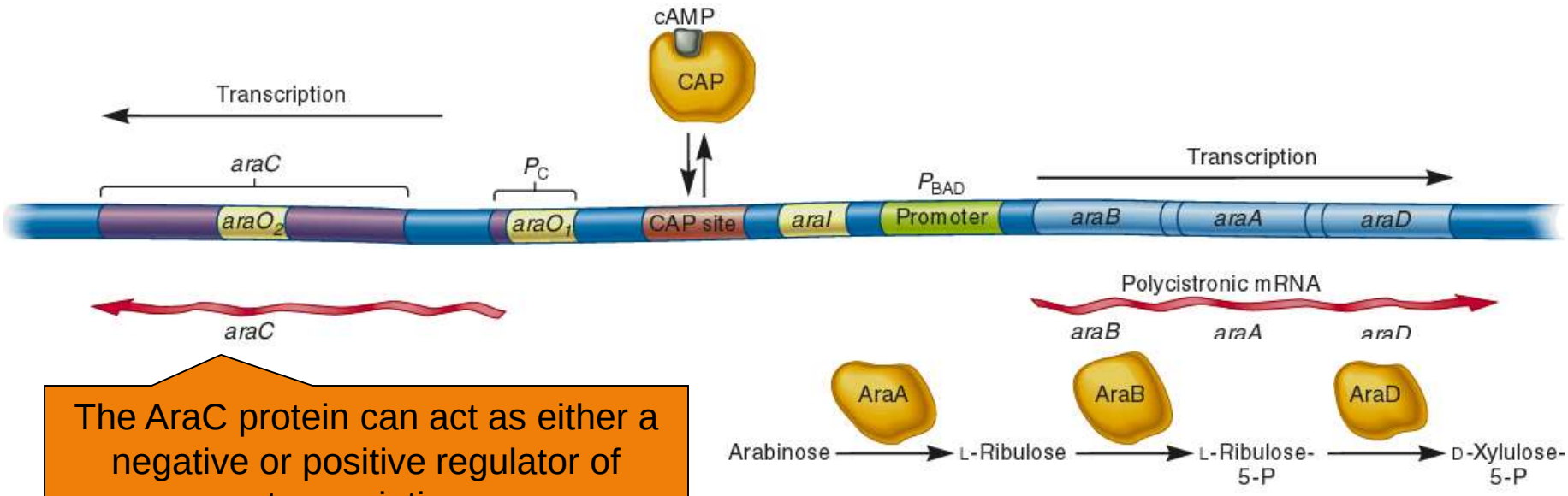
- The lack of trp causes ribosomes to stall at these codons, covering regions of the mRNA required for formation of the attenuation hairpin.
- This prevents attenuation and thus allows transcription to continue.

ARABINOSE OPERON

- The three structural genes (*araB*, *araA*, *araD*) encode for enzymes needed for the metabolism of the sugar arabinose in bacterial cells.
- *araB*,*araA*, and *araD* encode for the enzymes kinase, isomerase, and epimerase.
- Isomerase converts arabinose to ribulose.
- Kinase converts ribulose to ribulose-5-phosphate.

- The arabinose operon also contains the arabinose C gene which produce Ara c protein
- The *araC* gene regulates the expression of the structural genes and the *araC* protein.
- **Thus, the *araC* gene is auto regulated.**
- The presence of both arabinose and the *araC* gene product activates the expression of the BAD genes.

- The *araC* gene is adjacent to the *ara* operon
  - It has its own promoter, P_C
  - It encodes a regulatory protein, AraC
  - AraC can bind to **three different operator** sites
    - Designated *araI*, $araO_1$ and $araO_2$



The AraC protein can act as either a negative or positive regulator of transcription

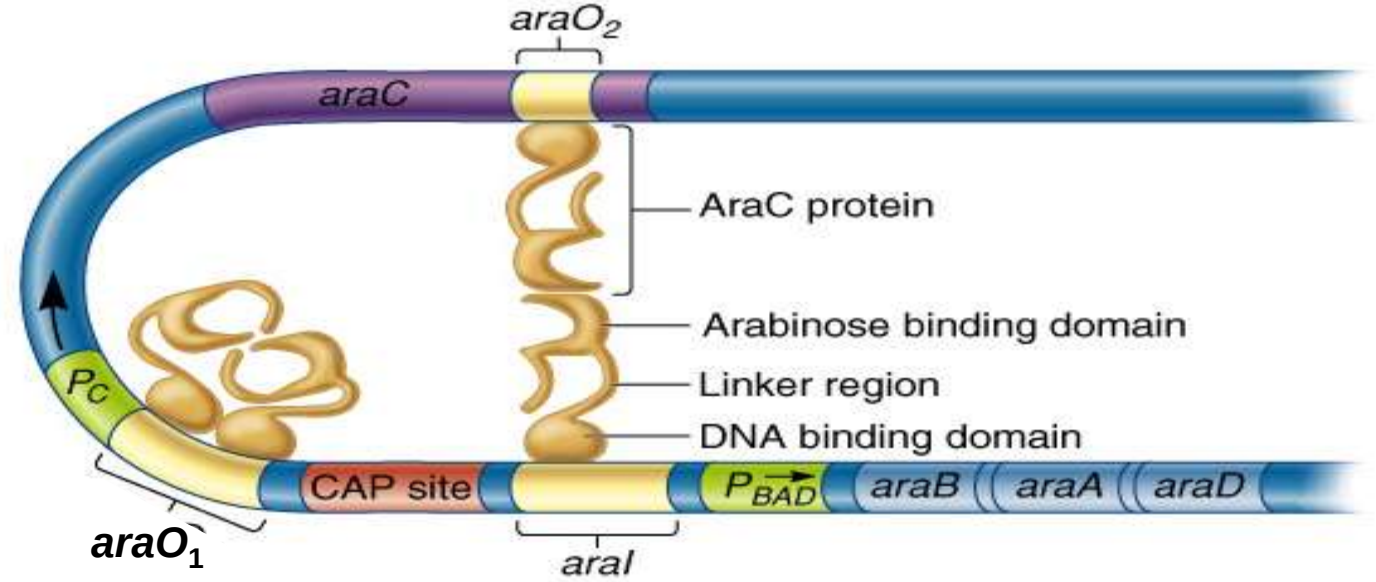
Depending on whether or not arabinose is present

POSITIVE REGULATION

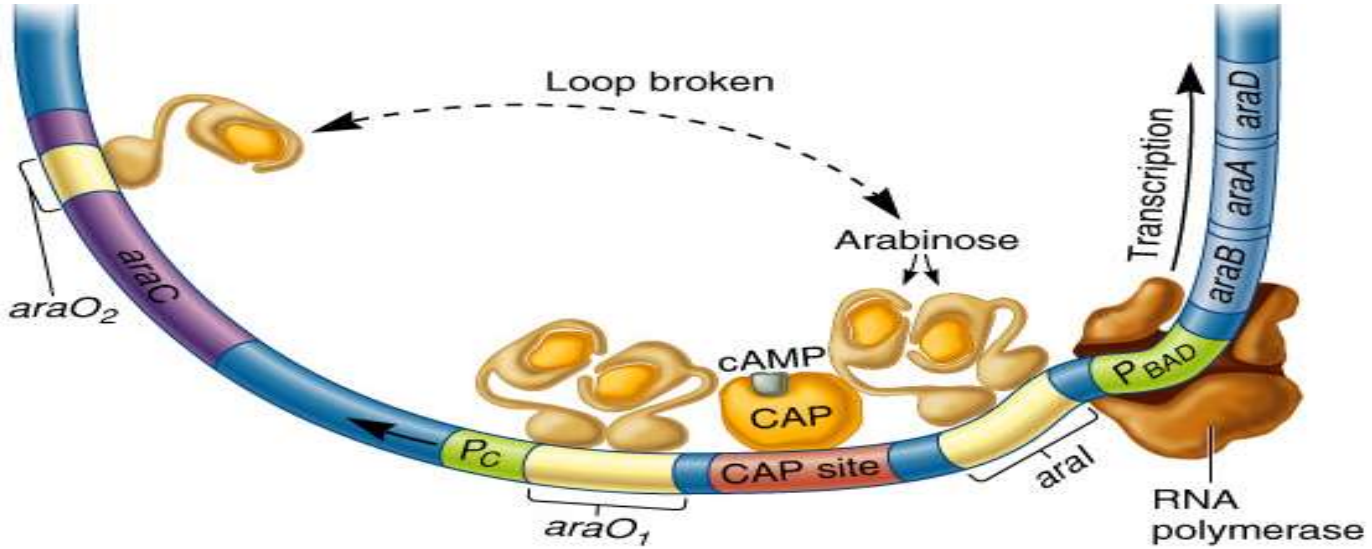
- When arabinose is present, it binds to AraC protein and changes AraC conformation.
- An arabinose- AraC dimer complex binds preferentially to AraI and NOT to *araO*$_2$ which causes ‘opening ‘ of the loop.
- This allows RNA pol to bind to $P_{BAD.}$
- If glucose levels are low, cAMP-CAP complex binds to Pc & active the transcription

NEGATIVE REGULATION

- When arabinose is absent, the AraC protein acts as a negative regulator.
- AraC acts as a dimer and causes the DNA loop.
- Looping brings the AraI and *araO*$_2$ sites in proximity to one another.
- One *AraC* monomer binds to araI and a second monomer binds to *araO*$_2$.
- Binding of *araC* prevents RNA pol from binding to the P_{BAD} promoter.
- Transcription not occur.



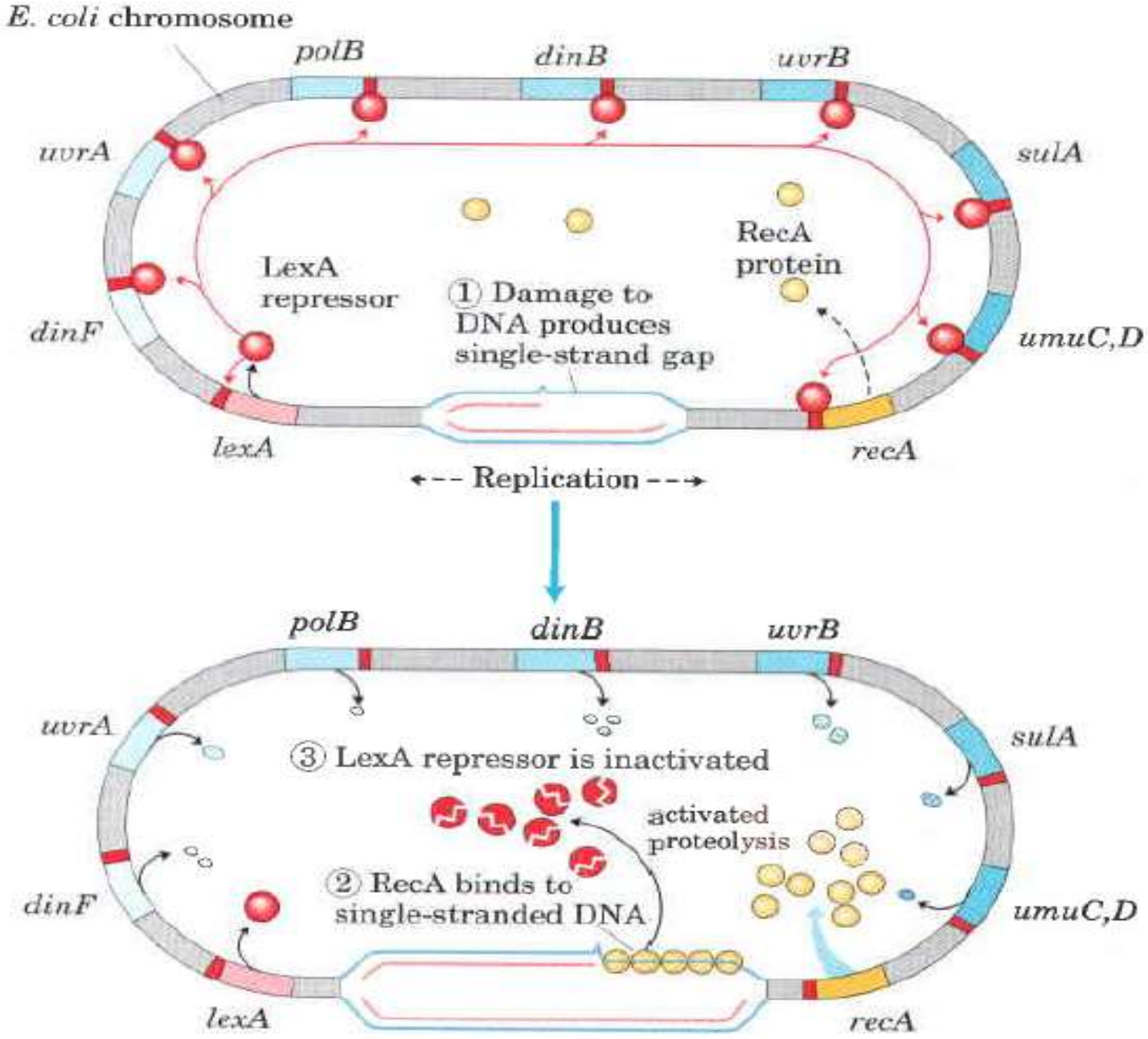
(a) Operon inhibited in the absence of arabinose



(b) Operon activated in the presence of arabinose

INDUCTION OF SOS RESPONSE

- Extensive DNA damage in the bacterial chromosome triggers the induction of many distantly located genes.
- This response, called the **SOS response**.
- When DNA is extensively damaged(such as by UV light)DNA replication is halted and the number of single-strand in the DNA increases.
- **RecA protein** binds to this damaged single-stranded DNA activating the protein's co protease activity.


E. coli chromosome
polB
dinB
uvrB
uvrA
sulA
LexA
repressor
RecA
protein
① Damage to
DNA produces
single-strand gap
dinF
umuC,D
lexA
recA
Replication
polB
dinB
uvrB
uvrA
sulA
③ LexA repressor is inactivated
activated
proteolysis
② RecA binds to
single-stranded DNA
dinF
umuC,D
lexA
recA

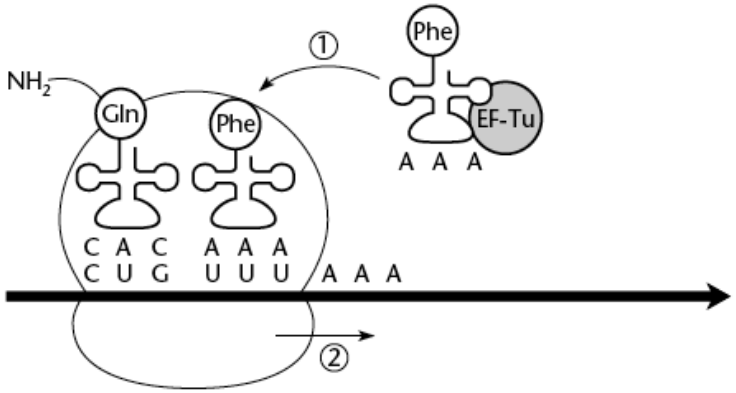
- While bound to DNA, the RecA protein facilitates cleavage and inactivation of the **LexA repressor** when the repressor is inactivated the SOS genes including recA, are induced;
- RecA levels increase 50 to 100 fold.

RIBOSOMAL PROTEIN WITH rRNA SYNTHESIS

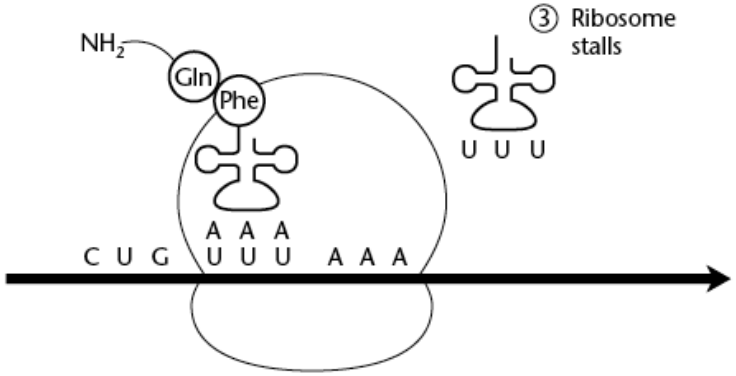
- Protein synthesis is major consumer of energy in bacteria.
- Because the number of ribosomes is primary determinants of level of translation and ribosome synthesis itself is an enrgy-intensive process.
- When cells are in a condition where there is an insufficient supply of amino acids to sustain protein, the stringent response is activated.

STRINGENT RESPONSE

- Stringent response causes a 10-20 times reduction in the synthesis of rRNA and tRNA. This causes a reduction of about 10% of the mRNA in the cell.
- The stringent response is accompanied by the increase of the alarmones ppGpp and pppGpp: guanosine tetraphosphate with diphosphates attached to the 5’ and 3’ ends of guanosin.
- Also guanosine pentaphosphate with a 5’ triphosphate group and 3’ diphosphate.



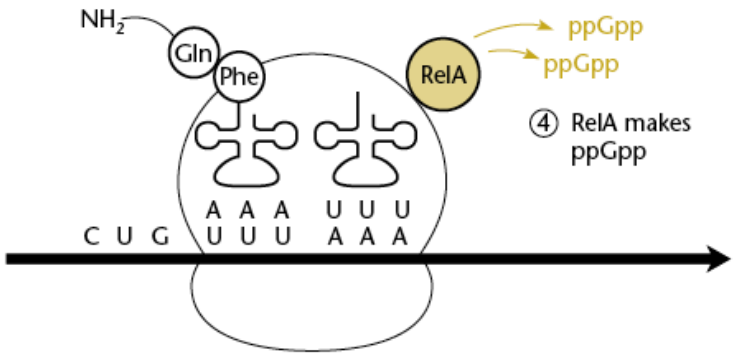
Charged tRNA binds to the A-site in riboso



Amino acid starvation results in an uncharged tRNA will bind to the A-site em

deacylated

tRNA



An uncharged tRNA is a signal for RelA binding to the ribosome and synthesizes ppGpp guanosine 3’5’ bisphophate, also pppGpp

ppGpp now interacs with RNA polymerase at key metabolic pathways

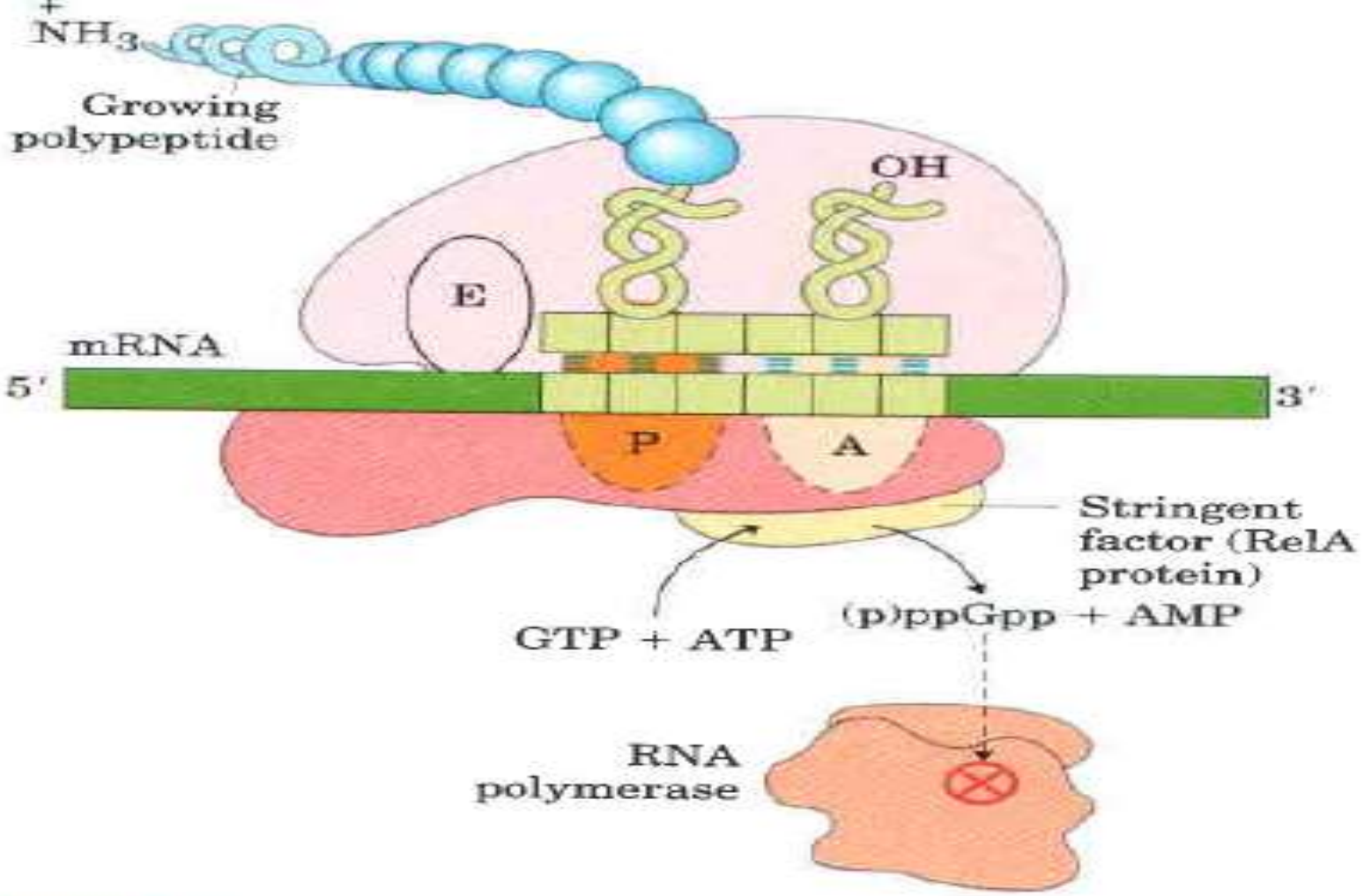


FIGURE 28–24 Stringent response in *E. coli*. This response to amino acid starvation is triggered by binding of an uncharged tRNA in the ribosomal A site. A protein called stringent factor binds to the ribosome and catalyzes the synthesis of pppGpp, which is converted by a phosphohydrolase to ppGpp. The signal ppGpp reduces transcription of some genes and increases that of others, in part by binding to the β subunit of RNA polymerase and altering the enzyme's promoter specificity. Synthesis of rRNA is reduced when ppGpp levels increase.

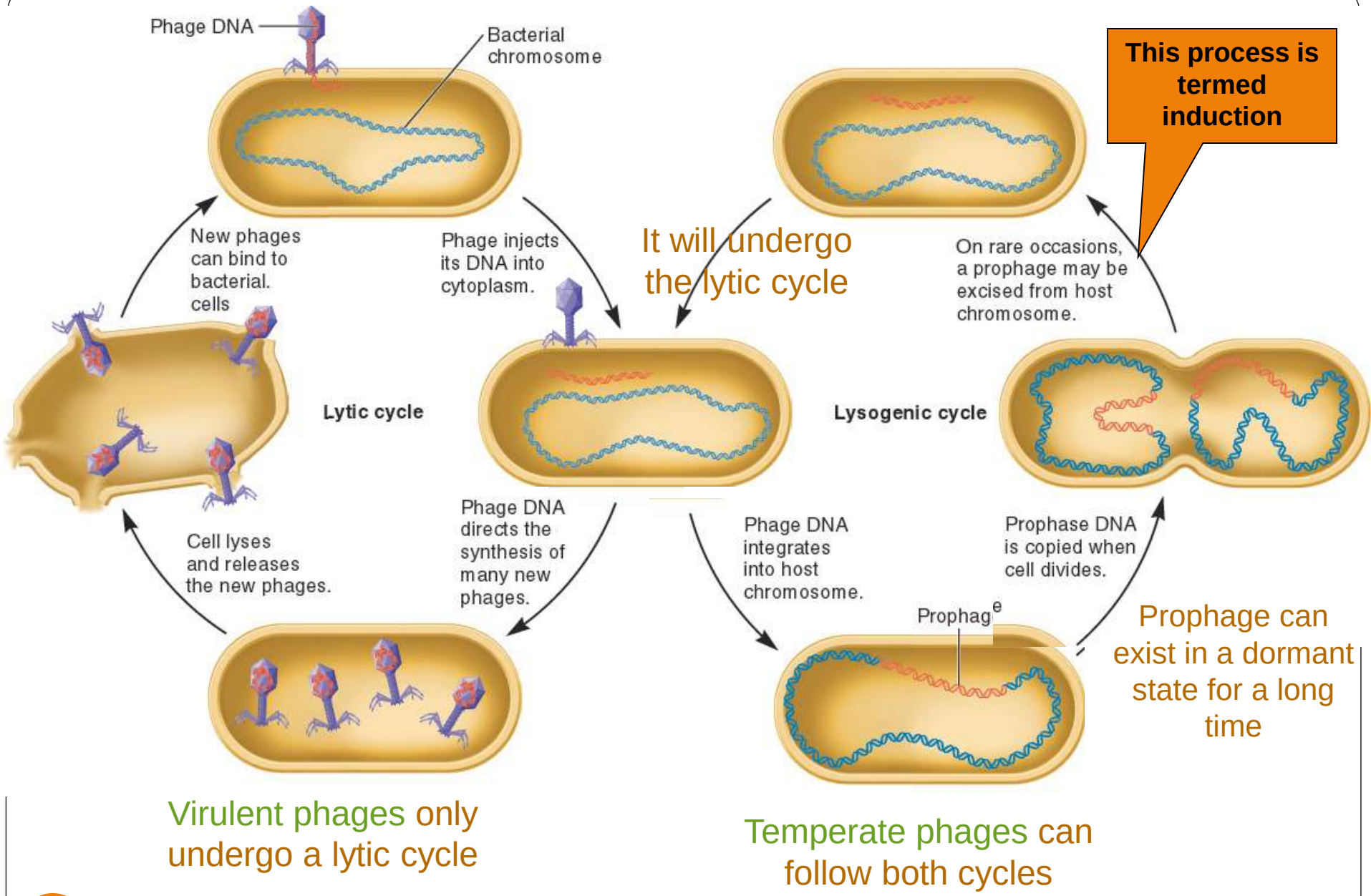
- The ppGpp molecule binds to the site where the polymerase would form the open complex.
- The mechanism of ppGpp then is to prevent the open complex formation of the transcription.

REGULATORY MECHANISM IN VIRUS (TRANSDUCTION)

❖ BACTERIOPHASE LAMDA:

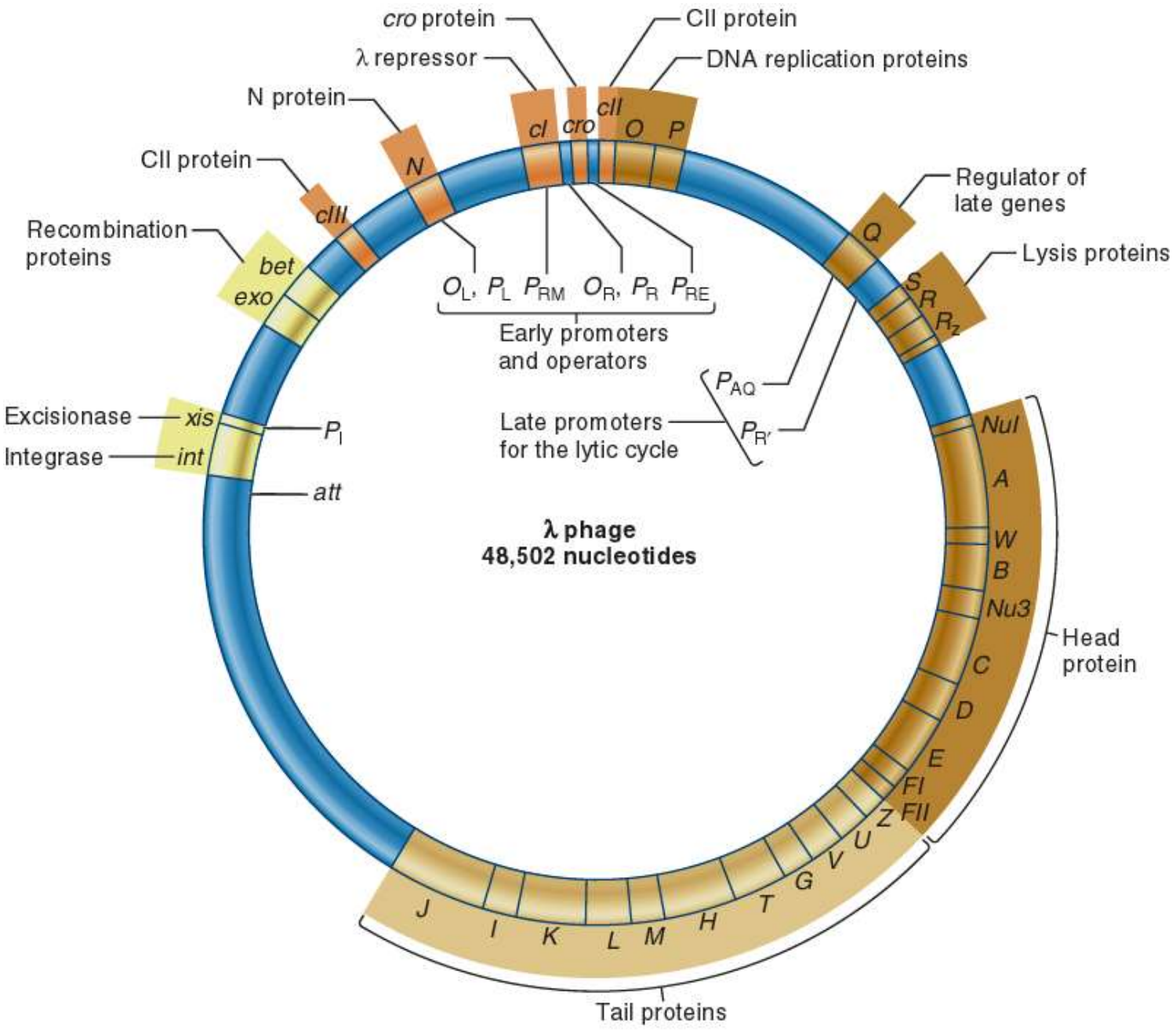
- some bacterial viruses can either reside in a dormant state within the host chromosomes or can replicate within the bacterium and eventually lead to lysis and killing of the bacterial host.
- When lambda infects an organism of that species it injects its 45,000-bp, double-stranded, linear DNA genome into the cell.

- Depending upon the nutritional state of the cell, the lambda DNA will either **integrate** into the host genome **(lysogenic pathway) and remain** dormant until activated , or it will commence **replicating until it has made about 100 copies** of complete, protein-packaged virus, at which point it causes lysis of its host **(lytic pathway).**


Phage DNA
Bacterial chromosome
This process is termed induction
New phages can bind to bacterial. cells
Phage injects its DNA into cytoplasm.
It will undergo the lytic cycle
On rare occasions, a prophage may be excised from host chromosome.
Lytic cycle
Lysogenic cycle
Phage DNA directs the synthesis of many new phages.
Phage DNA integrates into host chromosome.
Prophase DNA is copied when cell divides.
Cell lyses and releases the new phages.
Prophage
Prophage can exist in a dormant state for a long time
Virulent phages only undergo a lytic cycle
Temperate phages can follow both cycles

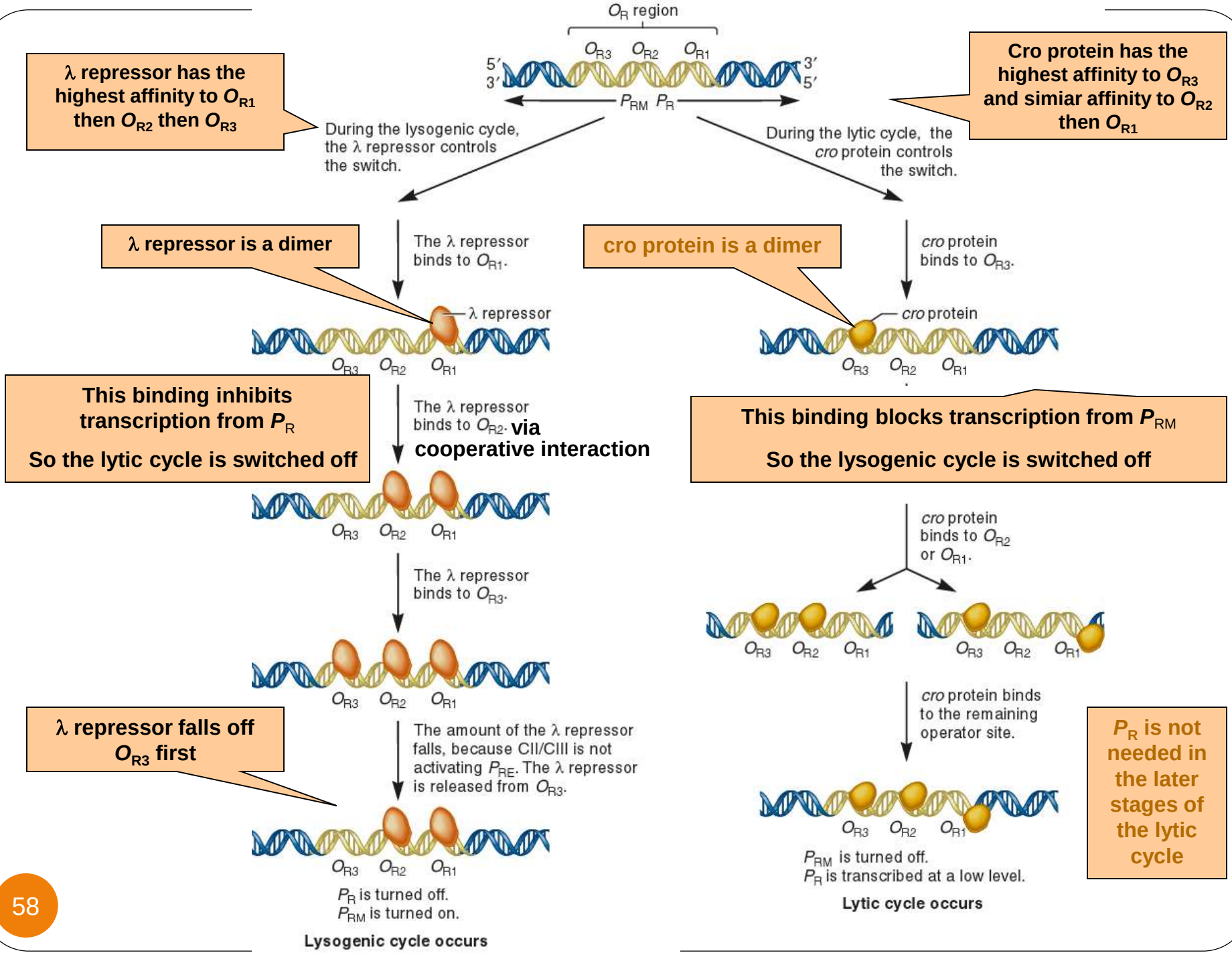
- Inside the viral head, phage λ DNA is linear.
- After injection into the bacterium, the two ends attach covalently to each other forming a circle.
- The organization of the genes within this circular structure reflects the two alternative life cycles of the virus.

- The genes on the *left* side of the viral genome encode proteins that are responsible for the lysogenic infection.
- The genes on the *right* side of the viral genome encode proteins that are responsible for the lytic infection.


cro protein
CII protein
λ repressor
DNA replication proteins
N protein
cI
cro
cII
O
P
CII protein
N
cIII
Regulator of late genes
Recombination proteins
Q
Lysis proteins
bet
exo
S
R
Rz
OL, PL PRM OR, PR PRE
Early promoters and operators
PAQ
PR'
Late promoters for the lytic cycle
Excisionase
xis
PI
Integrase
int
att
λ phage
48,502 nucleotides
NuI
A
W
B
Nu3
Head protein
C
D
E
FI
FII
Z
U
V
G
T
H
M
L
K
I
J
Tail proteins

The O_R Region Provides a Genetic Switch Between the Two Cycles

- The O_R region contains three operator sites, designated O_{R1}, O_{R2}, and O_{R3} .
- These operator sites control two promoters, P_R and P_{RM}, which transcribe in opposite directions.
- The l repressor protein or the cro protein can bind to any or all of the three operator sites.
- This binding governs the switch between the lysogenic and the lytic cycles.


λ repressor has the highest affinity to O_{R1} then O_{R2} then O_{R3}
O_R region
O_{R3} O_{R2} O_{R1}
5′ 3′
3′ 5′
P_{RM} P_R
Cro protein has the highest affinity to O_{R3} and simiar affinity to O_{R2} then O_{R1}
During the lysogenic cycle, the λ repressor controls the switch.
During the lytic cycle, the cro protein controls the switch.
λ repressor is a dimer
The λ repressor binds to O_{R1}.
λ repressor
cro protein is a dimer
cro protein binds to O_{R3}.
cro protein
This binding inhibits transcription from P_R
So the lytic cycle is switched off
The λ repressor binds to O_{R2}. via cooperative interaction
This binding blocks transcription from P_{RM}
So the lysogenic cycle is switched off
cro protein binds to O_{R2} or O_{R1}.
The λ repressor binds to O_{R3}.
cro protein binds to the remaining operator site.
λ repressor falls off O_{R3} first
The amount of the λ repressor falls, because CII/CIII is not activating P_{RE}. The λ repressor is released from O_{R3}.
P_R is turned off.
P_{RM} is turned on.
Lysogenic cycle occurs
P_{RM} is turned off.
P_R is transcribed at a low level.
Lytic cycle occurs
P_R is not needed in the later stages of the lytic cycle

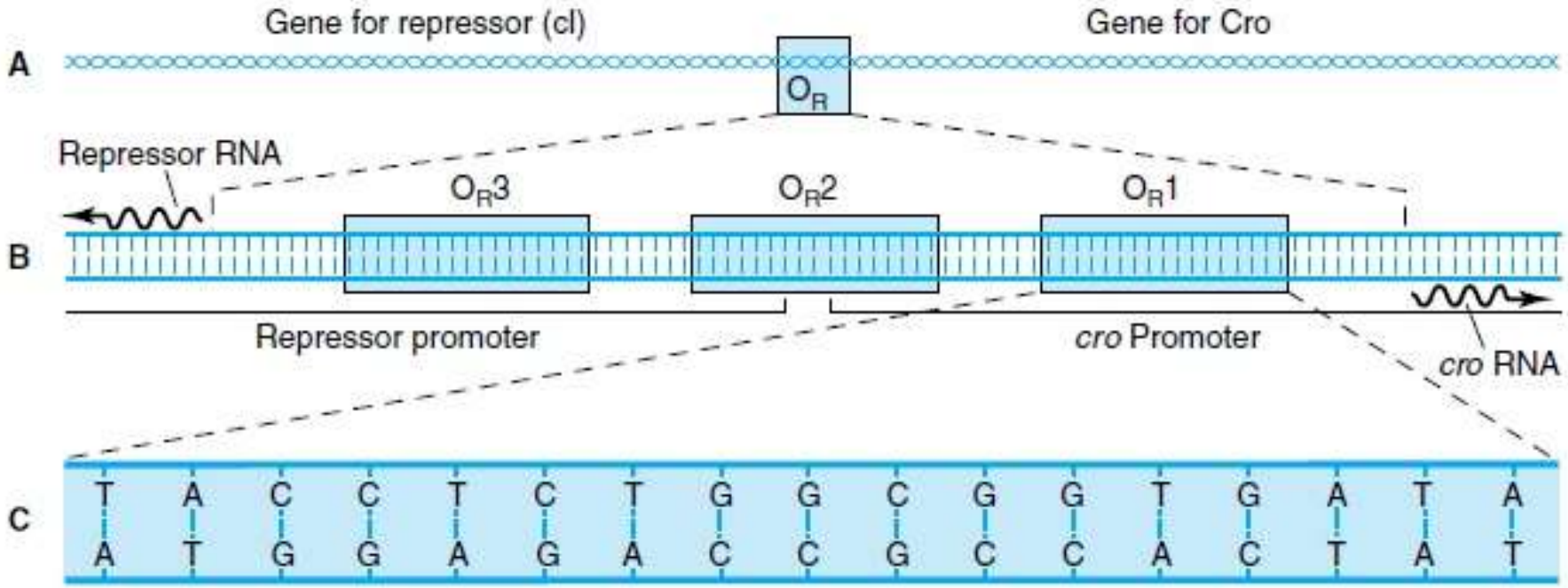
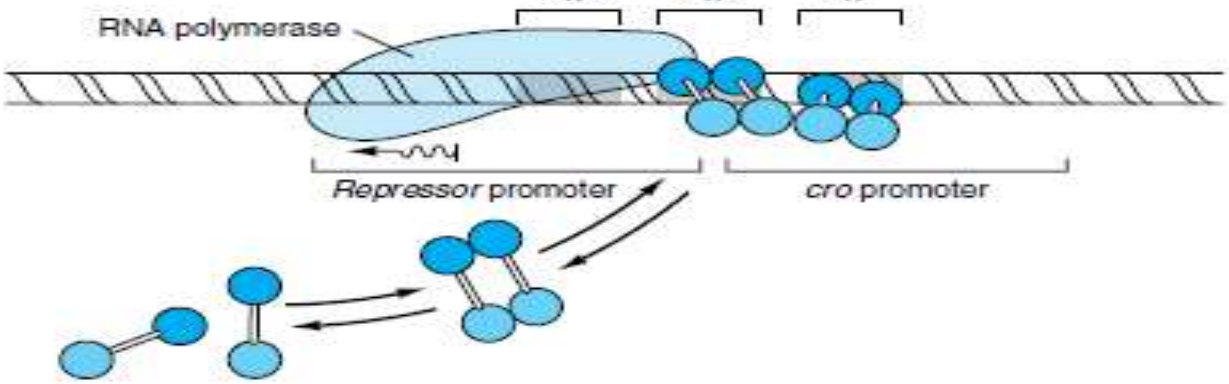
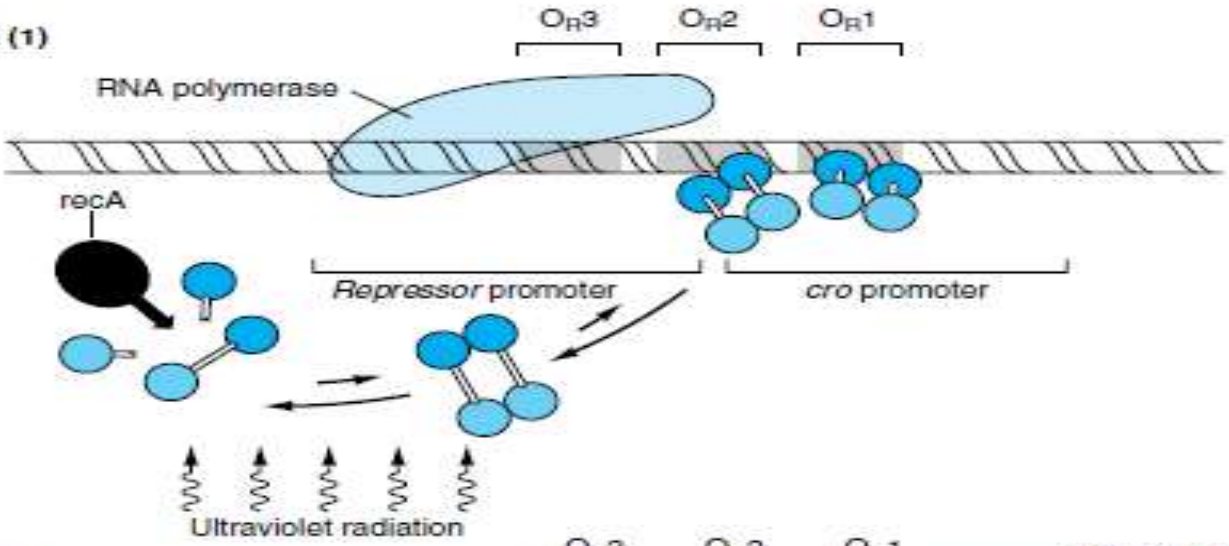


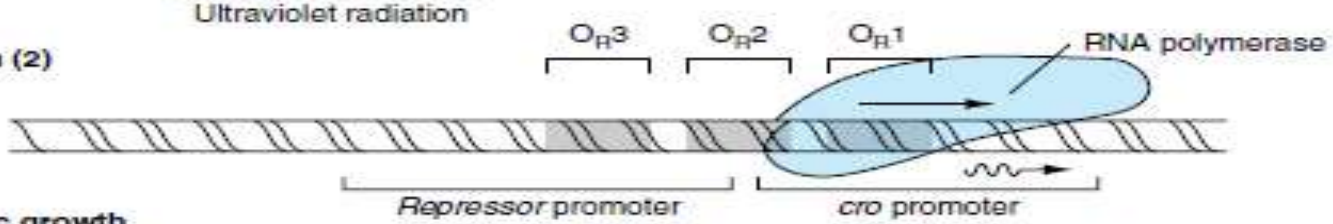
Figure 39–5. Right operator (O_R) is shown in increasing detail in this series of drawings. The operator is a region of the viral DNA some 80 base pairs long **(A).** To its left lies the gene encoding lambda repressor (cI), to its right the gene *(cro)* encoding the regulator protein Cro. When the operator region is enlarged **(B),** it is seen to include three subregions, O_R1, O_R2, and O_R3, each 17 base pairs long. They are recognition sites to which both repressor and Cro can bind. The recognition sites overlap two promoters—sequences of bases to which RNA polymerase binds in order to transcribe these genes into mRNA (wavy lines), that are translated into protein. Site O_R1 is enlarged **(C)** to show its base sequence. Note that in this region of the λ chromosome, both strands of DNA act as a template for transcription **(Chapter 39).** (Reproduced, with permission, from Ptashne M, Johnson AD, Pabo CO: A genetic switch in a bacterial virus. Sci Am [Nov] 1982;247:128.)



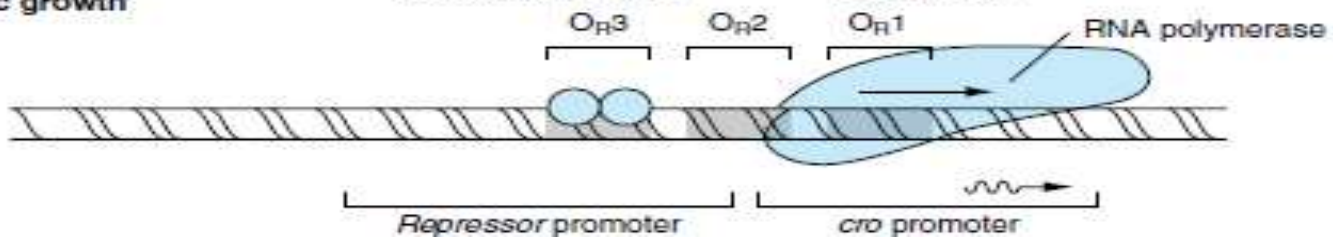
Induction (1)



Induction (2)



Early lytic growth



Reference

- Lehninger text book of biochemistry 5th edition
- Harper,book of biochemistry,29th edition
- Dinesh puri,text book of medical biochemistry,3rd edition.
- Google images.