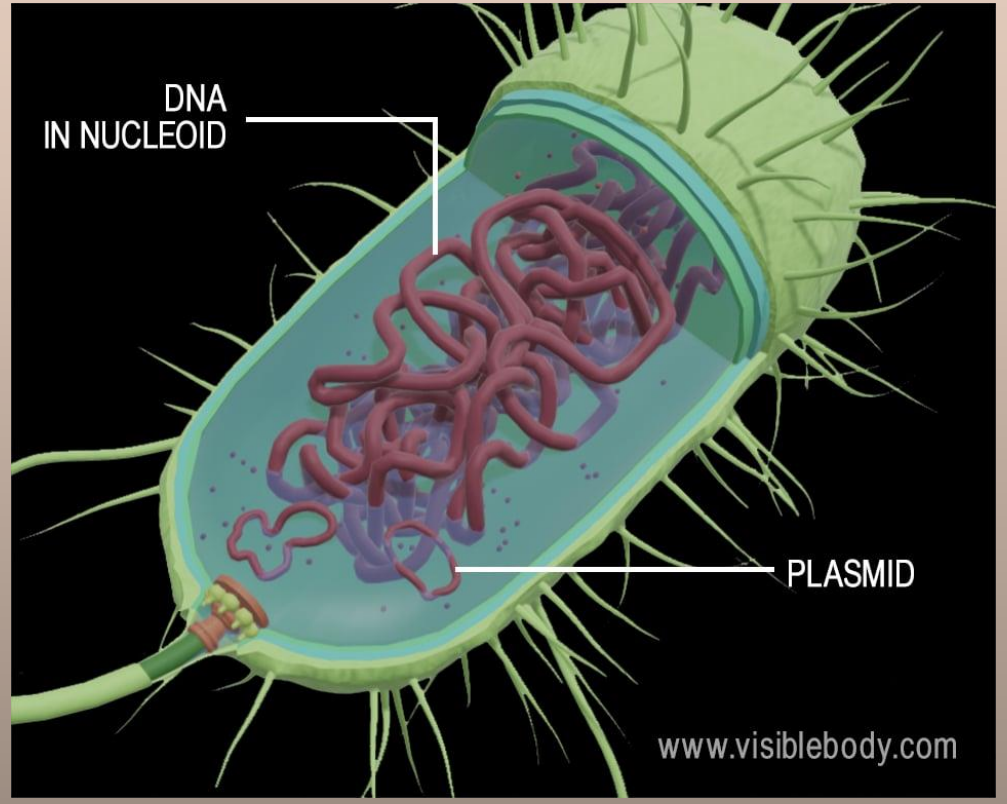
Realization of genetic information. Elements regulating the expression of prokaryotic genes



Lecturer: Senior Lecturer, Department of Molecular Biology and Genetics, PhD, Smekenov I.T.
Subject: Genetic engineering

(Lecture 5)

Lecture Goal:
To understand how genetic information is realized in prokaryotes, focusing on the elements that regulate gene expression and the operon model.

- **Tasks:**

1. Describe the genome of prokaryotes and distinguish between constitutive and inducible enzymes in gene expression.
2. Explain transcriptional regulation in prokaryotes, including the Jacob-Monod operon model, and discuss the value paradox.
3. Compare positive/negative inducible operons with positive/negative repressible operons in regulating gene expression.

Keywords: *Prokaryotic genome, constitutive enzymes, inducible enzymes, transcriptional regulation, Jacob-Monod operon model, operon, value paradox, positive inducible operon, negative inducible operon, positive repressible operon, negative repressible operon*

Summary

1. Prokaryotic genome

- The complete set of genetic material in prokaryotic cells, usually a single circular chromosome and sometimes plasmids.
- Contains all genes necessary for survival, metabolism, and regulation of gene expression.

2. Constitutive enzymes

- Enzymes that are produced continuously, regardless of environmental conditions.
- Responsible for essential metabolic functions, e.g., glycolysis enzymes.

3. Inducible enzymes

- Enzymes synthesized only in response to a specific substrate or environmental signal.
- Example: β-galactosidase in the lac operon is induced by lactose.

4. Transcriptional regulation

- Control of gene expression at the level of mRNA synthesis.
- Achieved via promoters, operators, repressors, activators, and inducers.

5. Jacob-Monod operon model

- A model explaining gene regulation in prokaryotes.
- Genes with related functions are clustered in an **operon**, controlled by a **promoter** and **operator**, with regulation by repressors or activators.

6. Operon

- A cluster of structural genes under control of a single promoter and operator.
- Allows coordinated expression of genes involved in the same pathway.

7. Value paradox

The observation that the presence of a gene does not always result in constant expression; expression is modulated by environmental and regulatory signals.

8. Types of operons

- **Positive inducible operon**: normally off; transcription is activated by an inducer.
- **Negative inducible operon**: normally off; transcription is turned on when the repressor is inactivated by an inducer.
- **Positive repressible operon**: normally on; transcription is turned off when an activator is inactivated by a corepressor.
- **Negative repressible operon**: normally on; transcription is blocked when a repressor is activated by a corepressor.

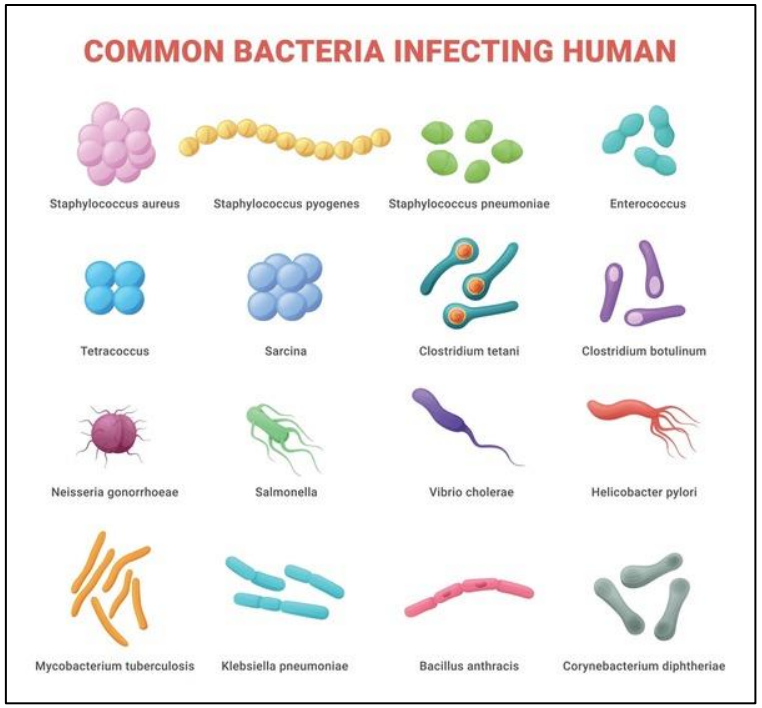
Key questions

1. What are the main differences between **constitutive** and **inducible** enzymes?
2. Describe the **Jacob-Monod operon model** and its components.
3. What is an **operon**, and why is it important for prokaryotic gene regulation?
4. Explain the **value paradox** in gene expression.
5. How do **positive inducible** and **negative inducible** operons differ in mechanism?
6. How do **positive repressible** and **negative repressible** operons regulate transcription?
7. Give an example of a metabolic pathway controlled by an inducible operon in bacteria.

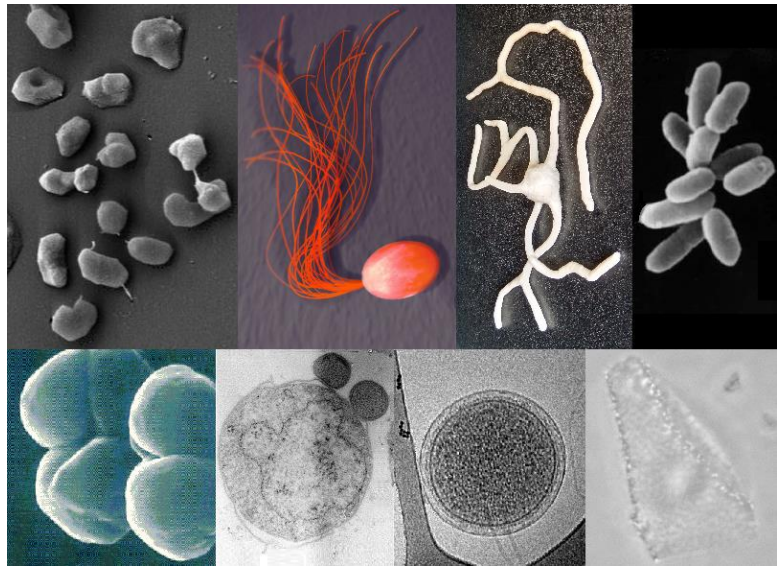
Prokaryotes

- **Prokaryotes** are unicellular organisms having a single cell, for example— bacteria, blue-green algae and archaea.
- Prokaryotes include two major groups — *Bacteria* and *Archaea*.

- Their genetic material (DNA) is located in a region called **the nucleoid**, not enclosed by a membrane.
- The traditional view has been that an entire prokaryotic genome is contained in a **single circular DNA** molecule.
- As well as this single 'chromosome', prokaryotes may also have additional genes on independent smaller, circular or linear DNA molecules called **plasmids.**



Cyanobacteria

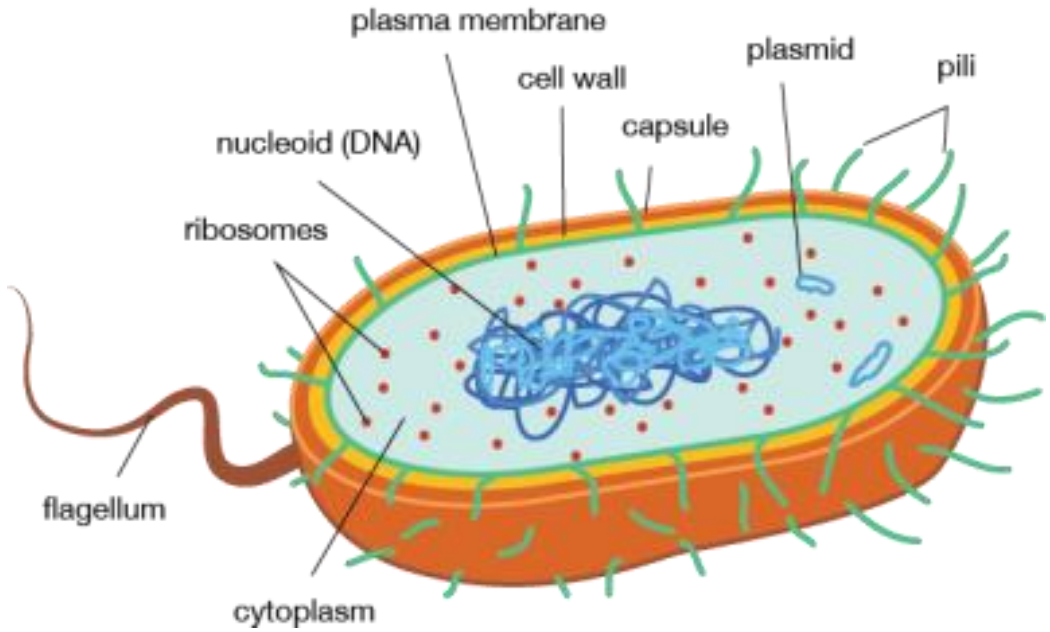


Archaea - The Daily Garden

Archaea are slightly more complex than bacteria, but not by much.

Genome of prokaryotes

- ✓ **Genome** of prokaryotes contains all the genetic information necessary for cell growth, metabolism, and reproduction.
- ✓ The prokaryotic genome is generally smaller and more compact than that of eukaryotes, with few noncoding regions.



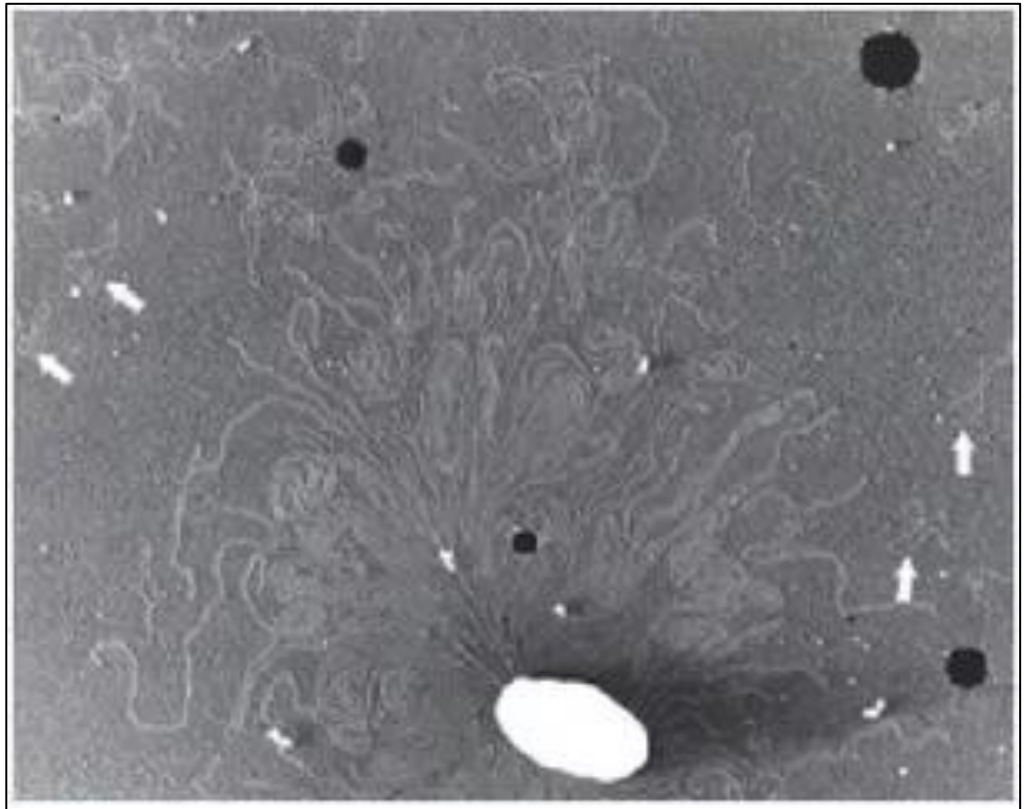
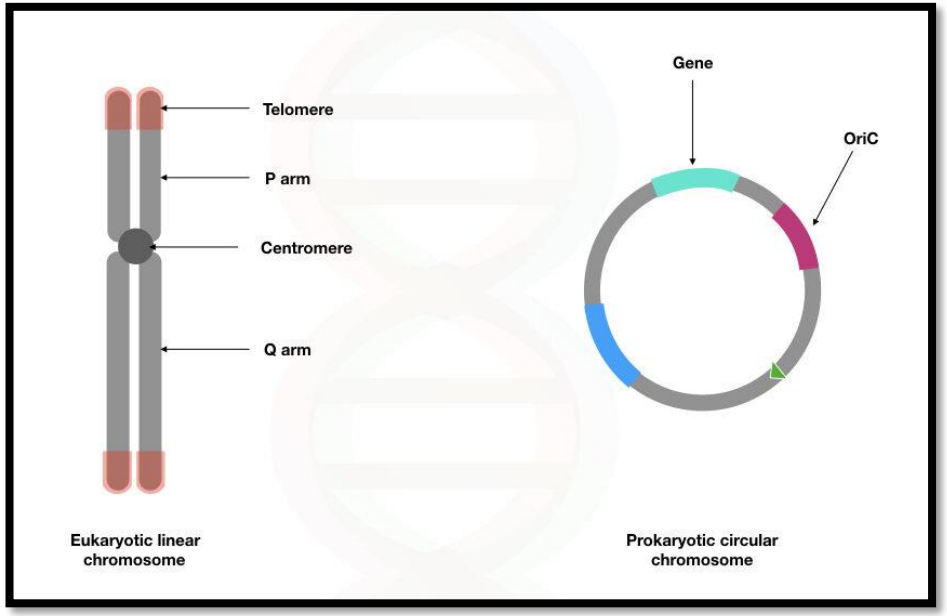


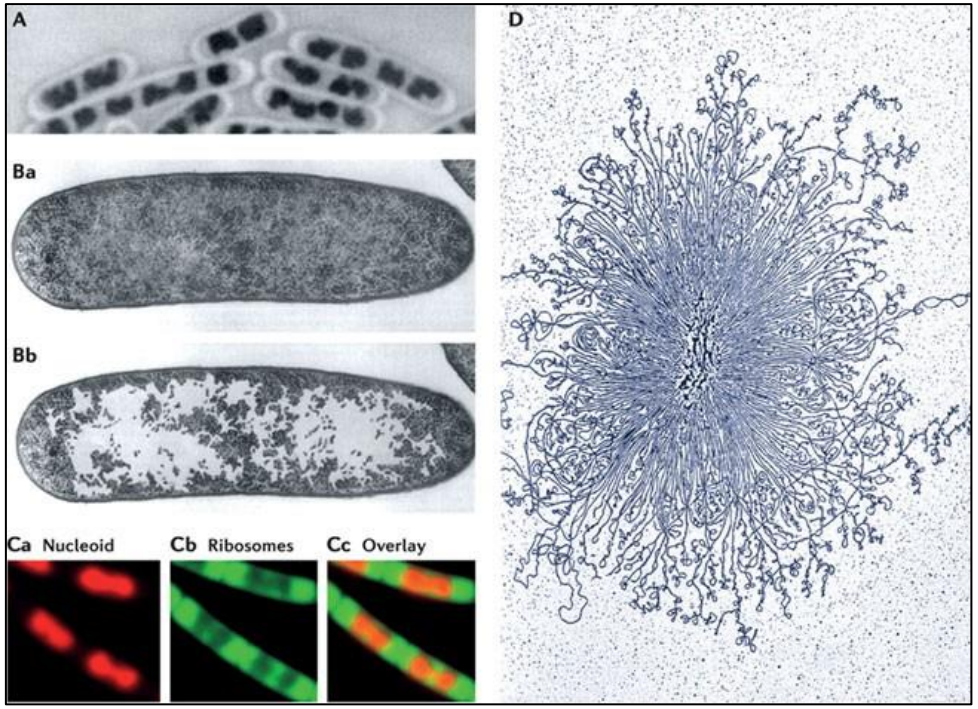
Figure. DNA from a lysed *E. coli* cell. Circular plasmid molecules are indicated by white arrows. The white and black spots are artifacts.

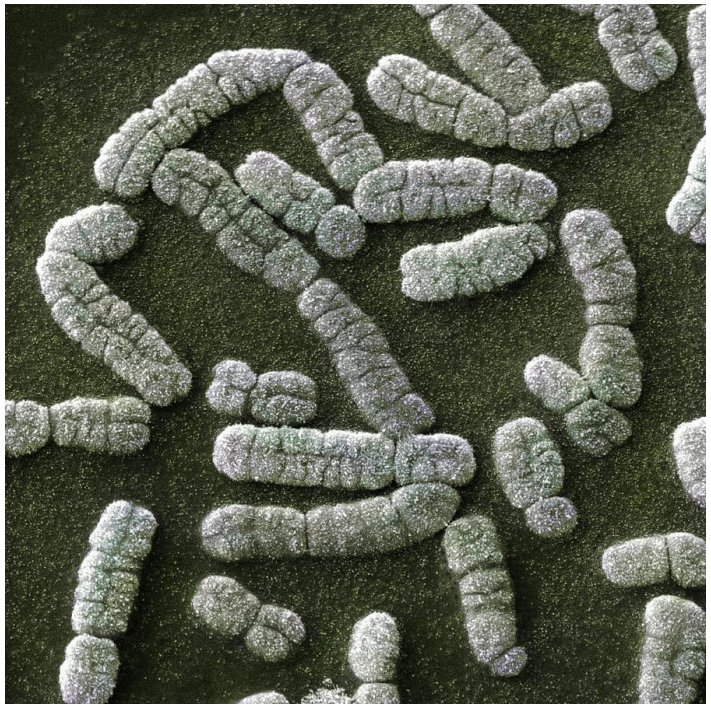
The genomes of prokaryotes and eukaryotes differ significantly in structure, organization, and function. Here's a comparison between the two type of cells:

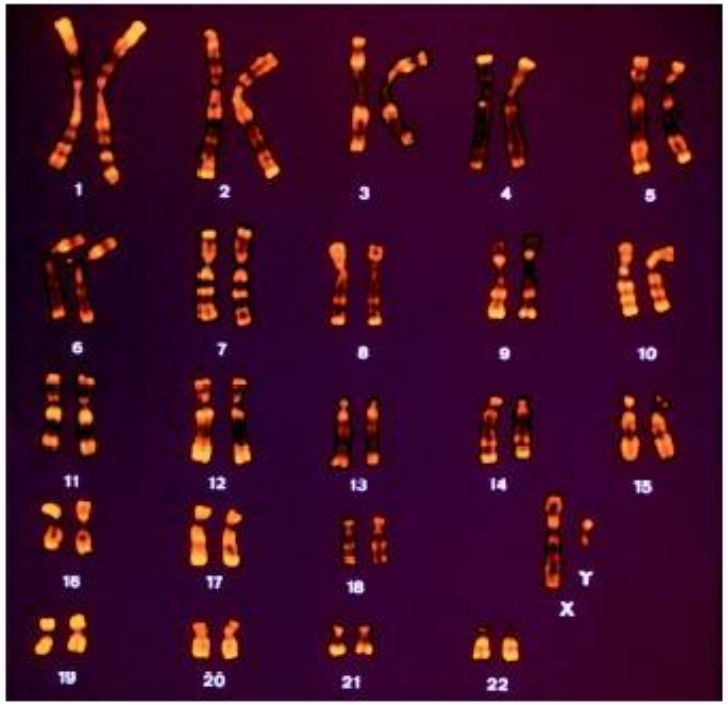
Feature	Prokaryotes	Eukaryotes
Genome Size	Small (thousands to millions of bp)	Large (millions to billions of bp)
Chromosome Structure	Single, circular	Multiple, linear
Plasmids	Common	Rare
Gene Density	High, few introns	Low, many introns and non-coding DNA
Gene Organization	Operons (clusters of genes)	Single genes with introns
Regulation	Mostly at transcriptional level	Complex, at multiple levels
Introns	Rare	Common
Organellar Genomes	None	Mitochondria, chloroplasts
Horizontal Gene Transfer	Frequent	Rare
Replication Origins	Single	Multiple
Telomeres	Not needed (circular chromosomes)	Required (linear chromosomes)

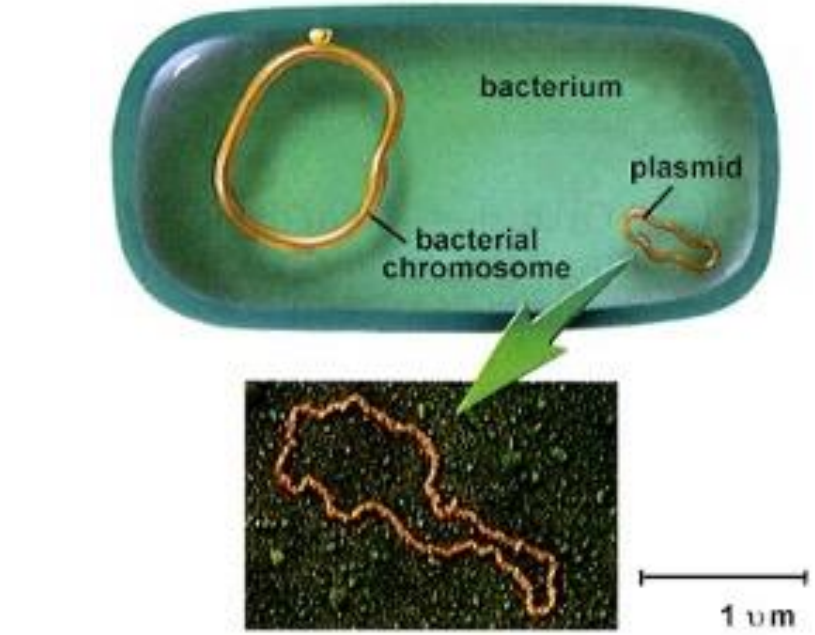


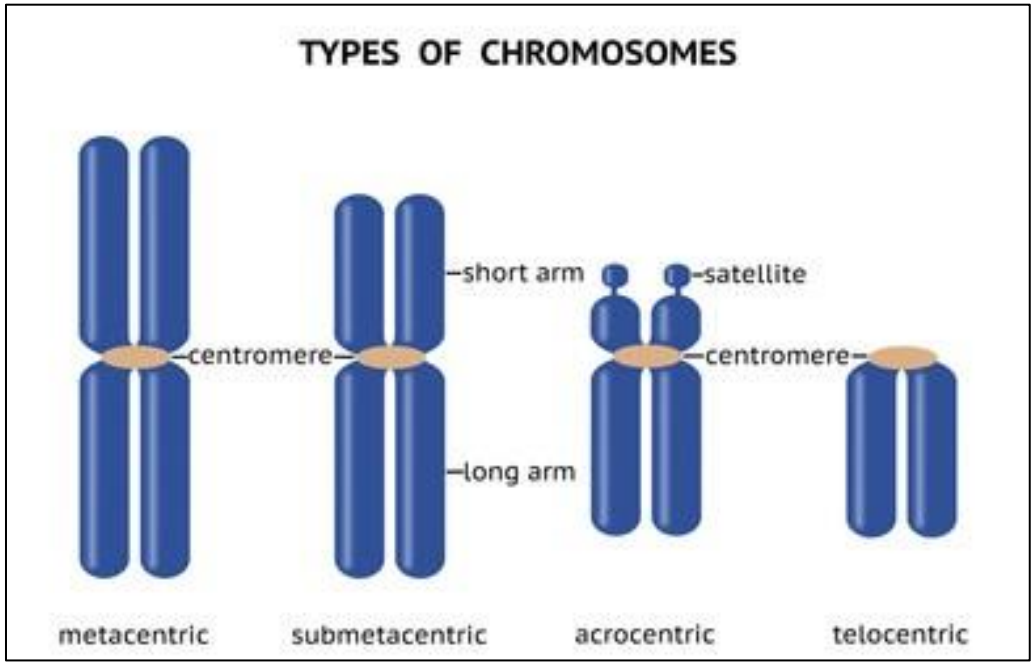
Prokaryotes possess a single chromosome while eukaryotes possess different numbers of chromosomes.


A
Ba
Bb
Ca Nucleoid
Cb Ribosomes
Cc Overlay
D




1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Y
X
19
20
21
22


bacterium
plasmid
bacterial chromosome
1 μm


TYPES OF CHROMOSOMES
centromere
short arm
long arm
satellite
centromere
metacentric
submetacentric
acrocentric
telocentric

- Number of Genes and Genome Size in Prokaryotes:

Prokaryote	No of Genes	Genome size
E. coli K12	4400	4.64 MB
M.tuberculosis	4000	4.41 MB
A.fulgidus (Archaea)	2500	2.18 MB

- Number of Genes and Genome Size in Eukaryotes:

Eukaryotes	No of Genes	Genome size
H.sapiens	20 to 25,000	3200 MB
A.thaliana	25,500	125 MB
D.melanogaster	13,600	180 MB

Sizes of genomes - The C-value paradox

- The **C-value** is the amount of DNA in the haploid genome of an organism.
- The **C-paradox** is the lack of correlation between the physical size of the genome and the complexity of organisms.

Our current understanding of complex genomes reveals several factors that help explain the classic C-value paradox:

- ✓ Introns in genes
- ✓ Regulatory elements of genes
- ✓ Pseudogenes
- ✓ Multiple copies of genes
- ✓ Intergenic sequences
- ✓ Repetitive DNA

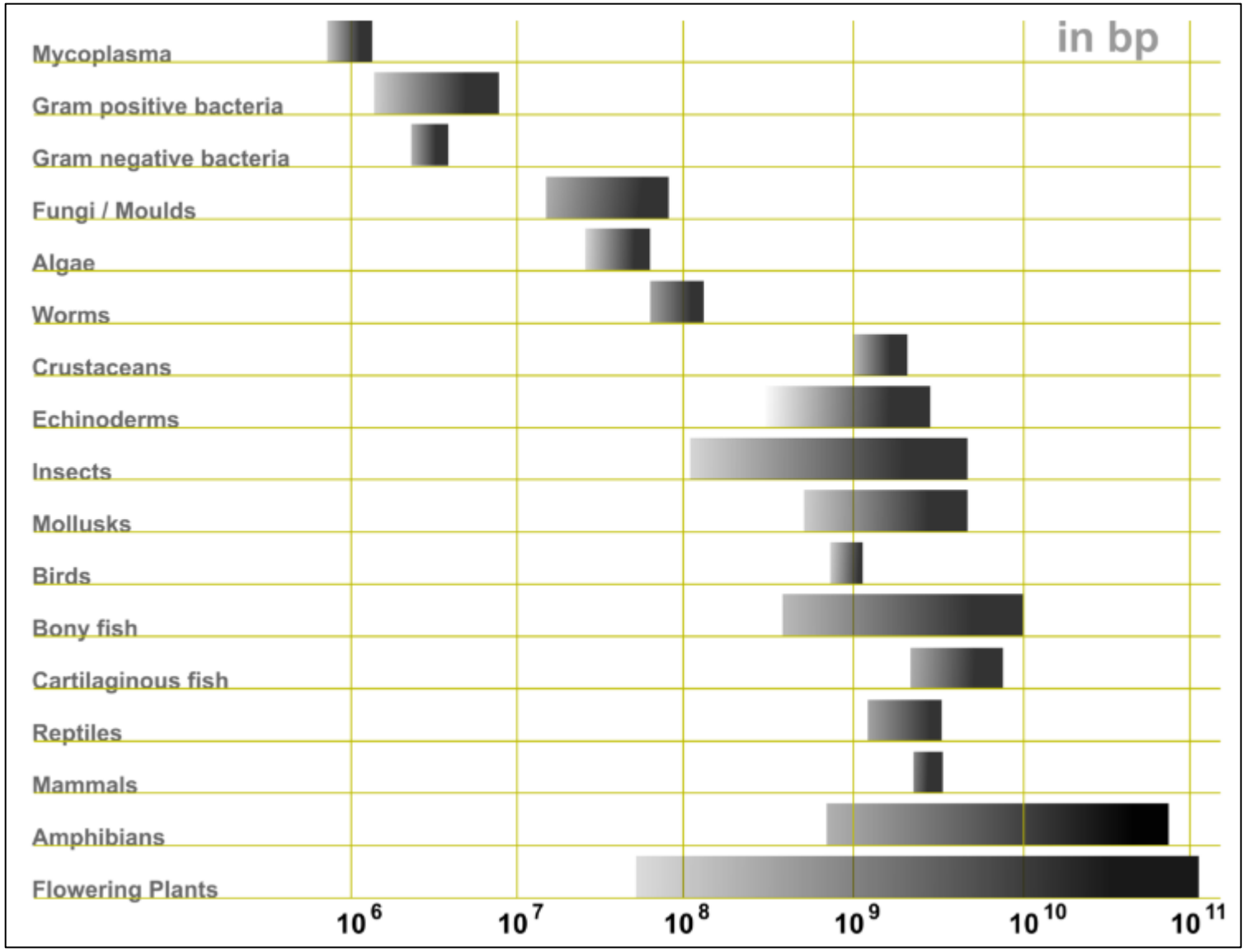
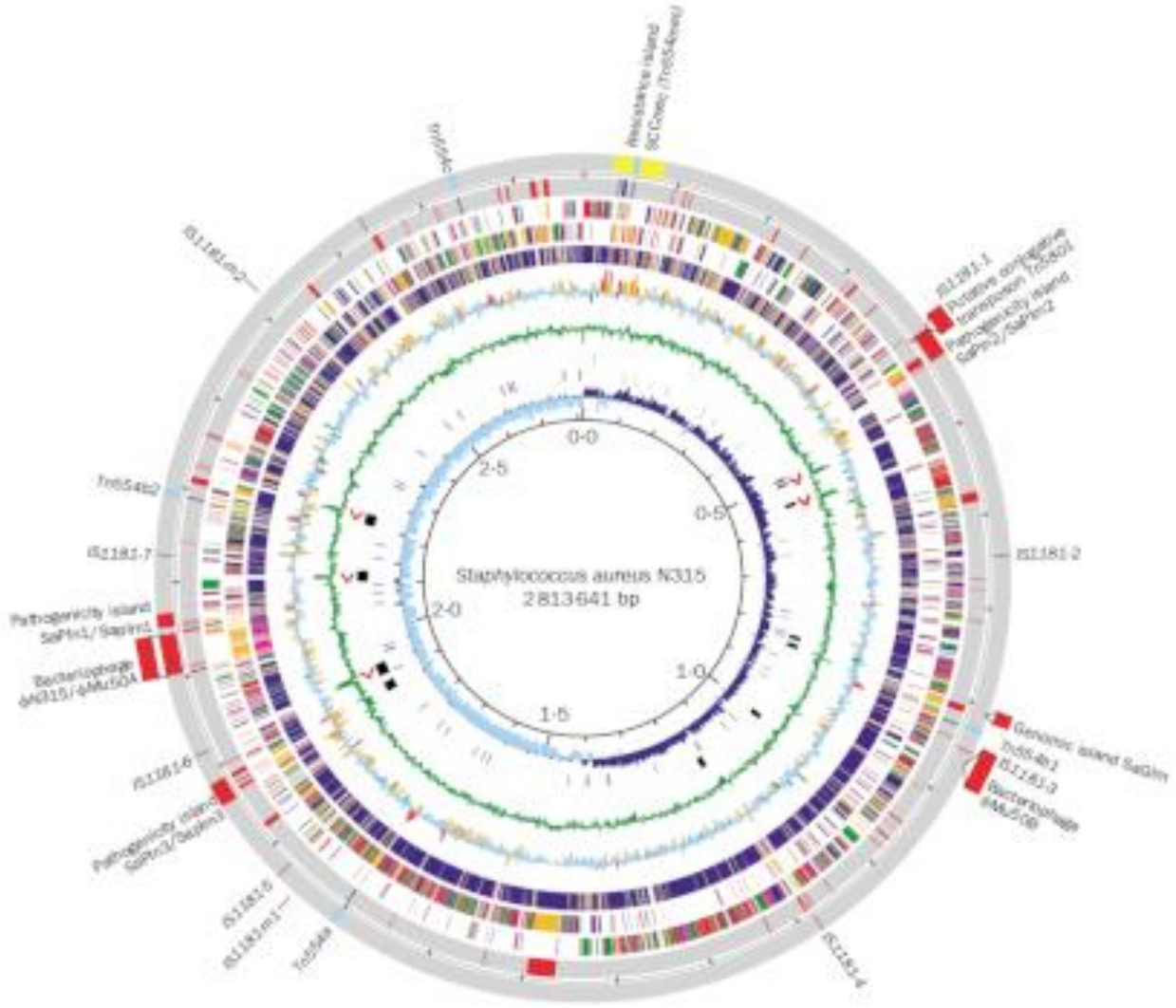


Figure: Genome size ranges (in base pairs) of various life forms.

Structure and Composition of the Bacterial Genome

- The bacterial genome is a highly organized and compact structure that contains all the genetic information necessary for the survival and reproduction of the cell.
- Unlike eukaryotic cells, bacteria lack a membrane-bound nucleus. Instead, their DNA is localized in a specific region of the cytoplasm known as the nucleoid.
- **The nucleoid** is not surrounded by a membrane but is a dynamic, organized structure formed by the folding and supercoiling of DNA, assisted by proteins such as **HU, IHF,** and **H-NS** that compact and stabilize the chromosome.
- Supercoiling, maintained by enzymes like DNA gyrase and topoisomerase I, is essential for fitting the long DNA molecule into the small bacterial cell and for regulating gene expression.



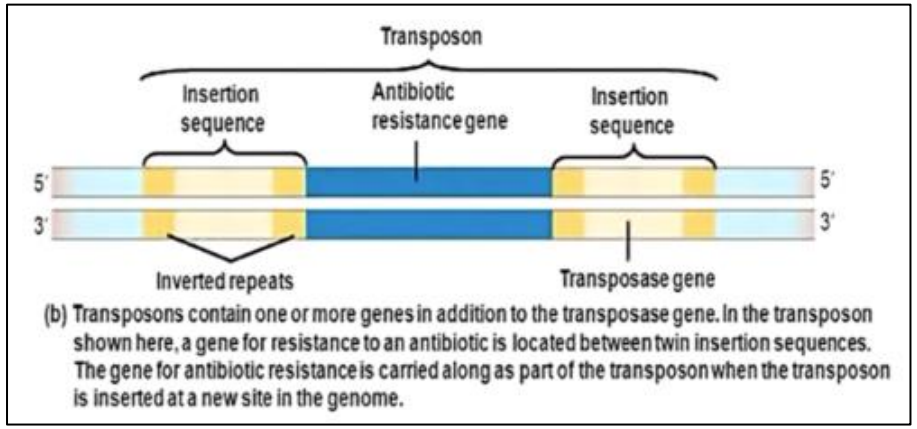
- The bacterial genome is generally gene-dense, meaning that most of its DNA encodes functional products, with very few non-coding or repetitive sequences. Genes are often arranged in operons, which **are clusters** of genes transcribed together as a **single mRNA molecule**.

This organization allows bacteria to efficiently regulate groups of genes involved in the same metabolic pathway or function.

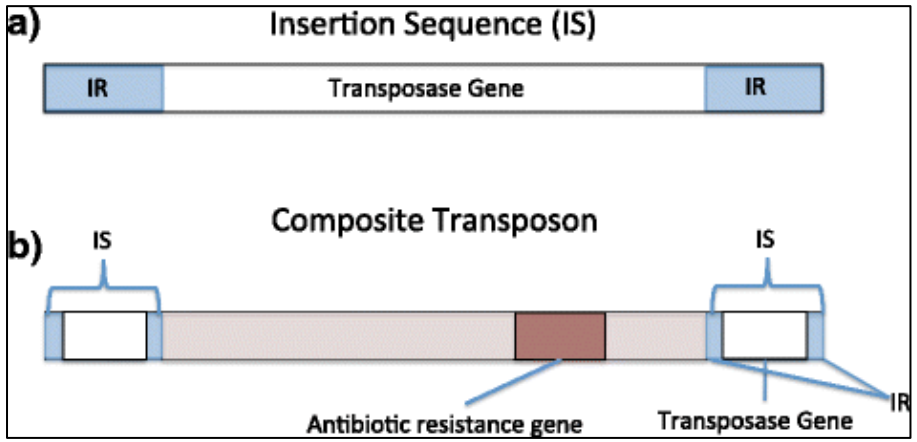
- The bacterial genome also contains mobile genetic elements, including transposons, insertion sequences, and integrons, which can move within or between DNA molecules.

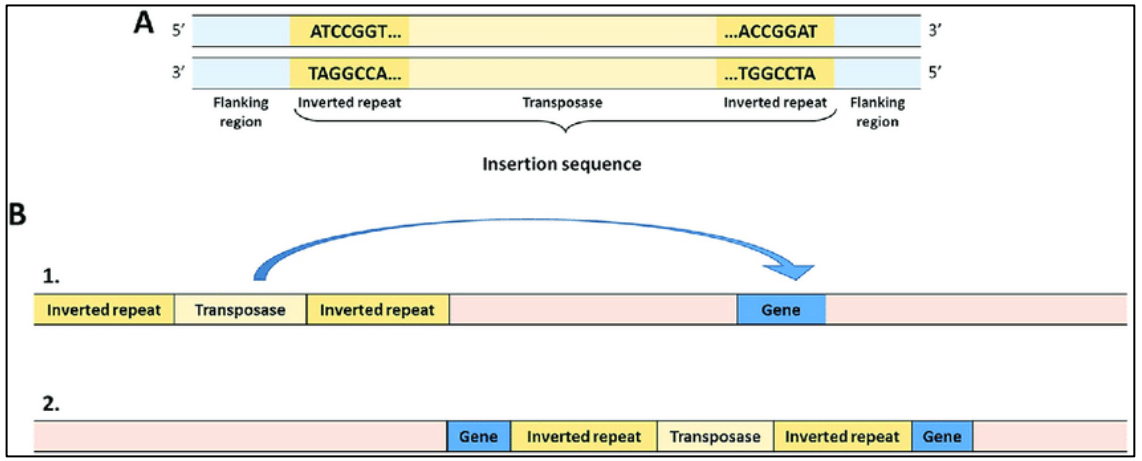
These elements play a significant role in genome evolution by promoting genetic rearrangements, gene duplication, and horizontal gene transfer.

✓ **Overall, the structure of the bacterial genome reflects the evolutionary pressure toward efficiency and adaptability. Its compact organization, operon-based gene regulation, and ability to exchange genetic material enable bacteria to respond rapidly to environmental changes and survive in diverse conditions.**



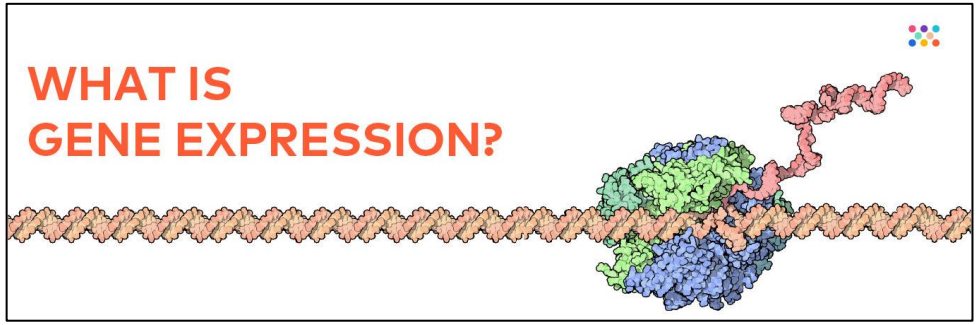
(b) Transposons contain one or more genes in addition to the transposase gene. In the transposon shown here, a gene for resistance to an antibiotic is located between twin insertion sequences. The gene for antibiotic resistance is carried along as part of the transposon when the transposon is inserted at a new site in the genome.

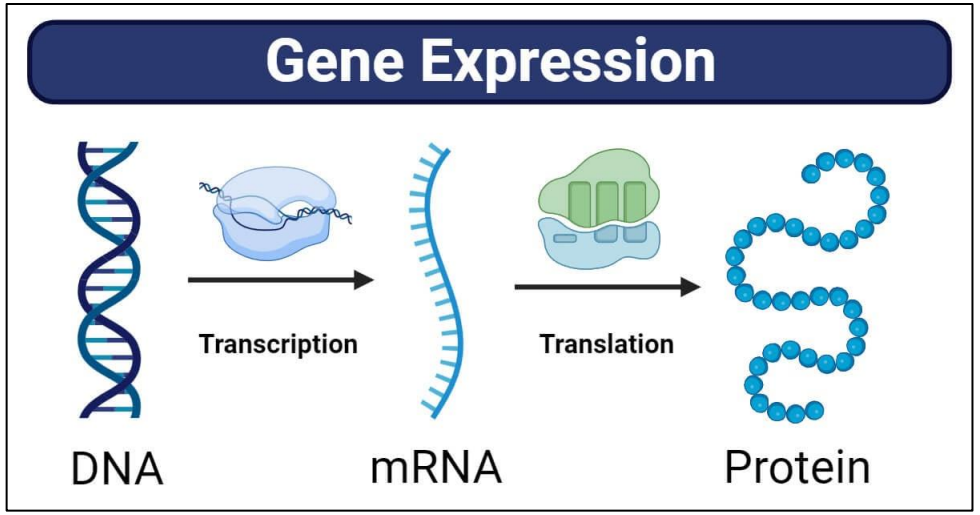




Prokaryotic Gene Regulation

- ***Gene expression*** refers to the sum of processes that result in a particular level of a specified mRNA and protein in the cell.
- The ***gene expression*** represents the information present in the gene. This is the initial step that produces the mRNAs (protein-coding RNAs) and other functional RNAs, such as rRNA or tRNA. Gene expression is a very basic process and appears in all organisms, which build the macromolecular machinery for an organism.
- ***Gene expression*** is the process by which the information encoded into genomic DNA is used by cells to synthesize a functional gene product through a cascade of steps consisting of transcription, post-transcriptional control, mRNA splicing, translation, and post-translational modifications.





- To understand how **gene expression** is regulated, we must first understand how a gene codes for a functional protein in a cell. The process occurs in both prokaryotic and eukaryotic cells, just in slightly different manners.

- **Prokaryotes** are also characterized by the **conjugation** of *transcription* and *translation processes.*

In **prokaryotic cells**, translation often begins while the mRNA is still being synthesized by RNA polymerase — a process known as **coupled transcription and translation**.

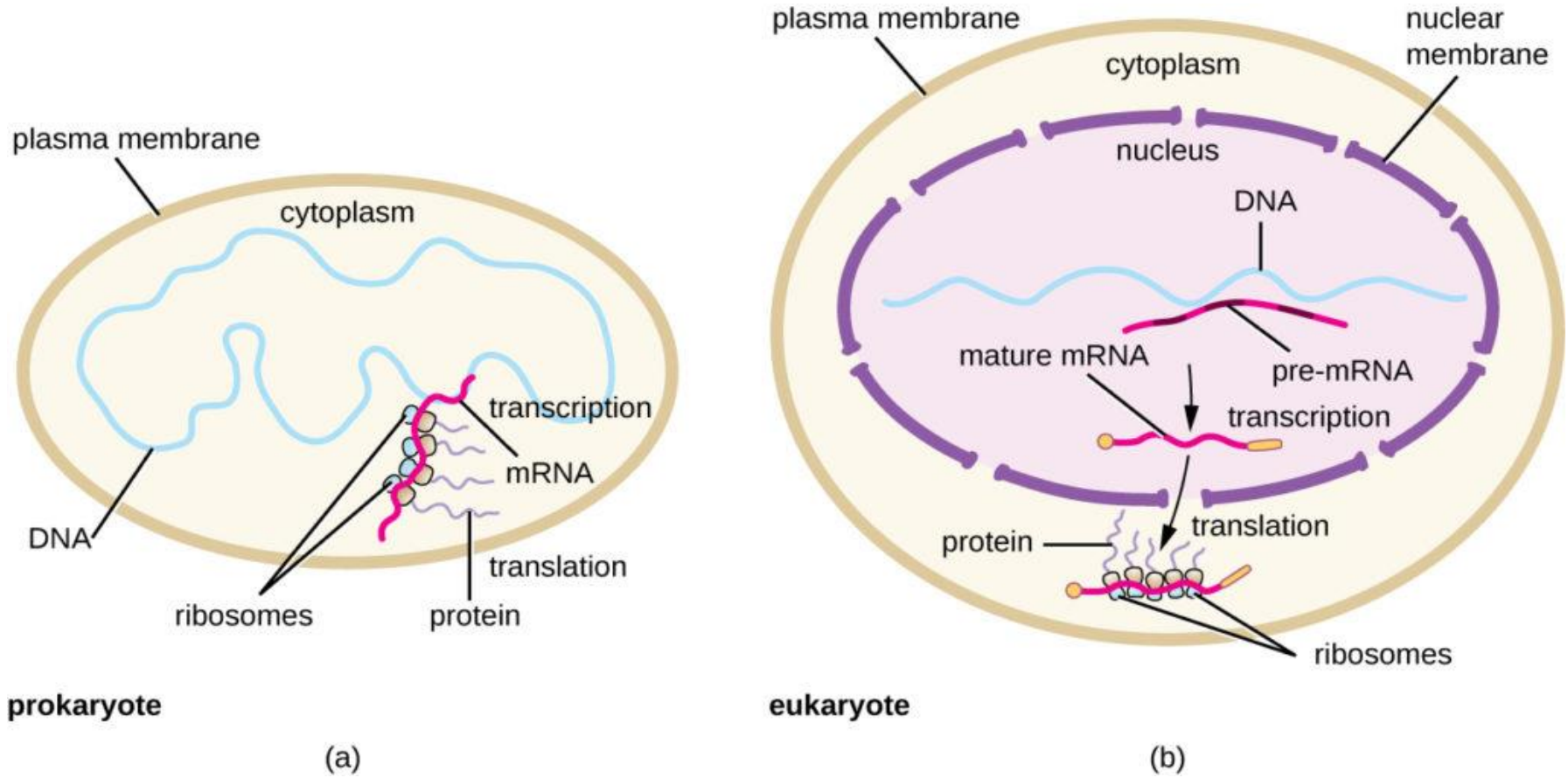
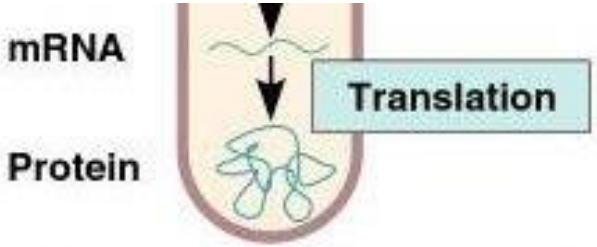
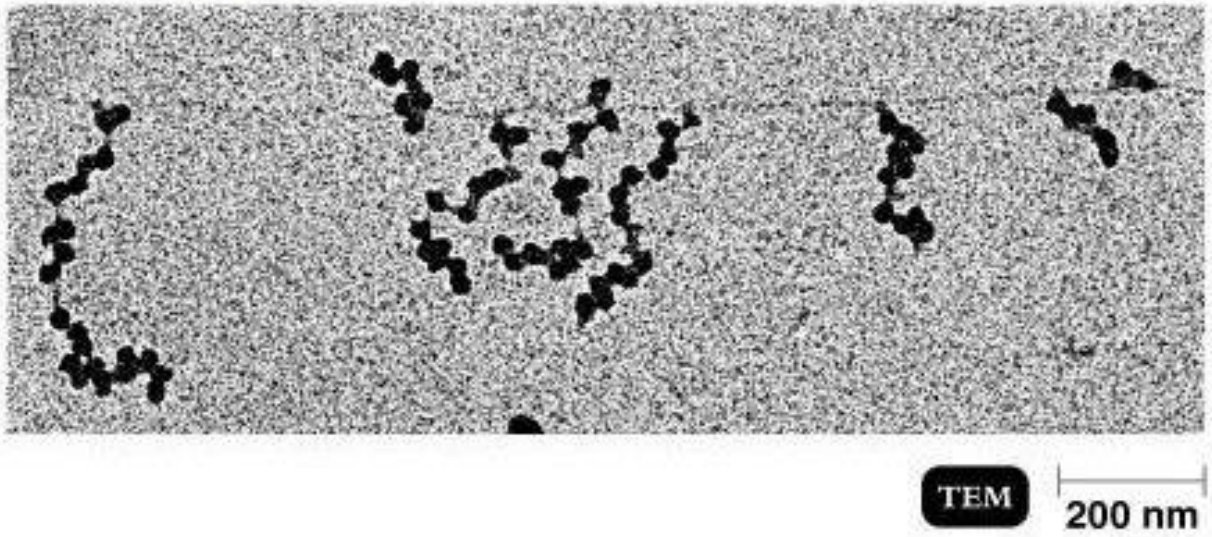
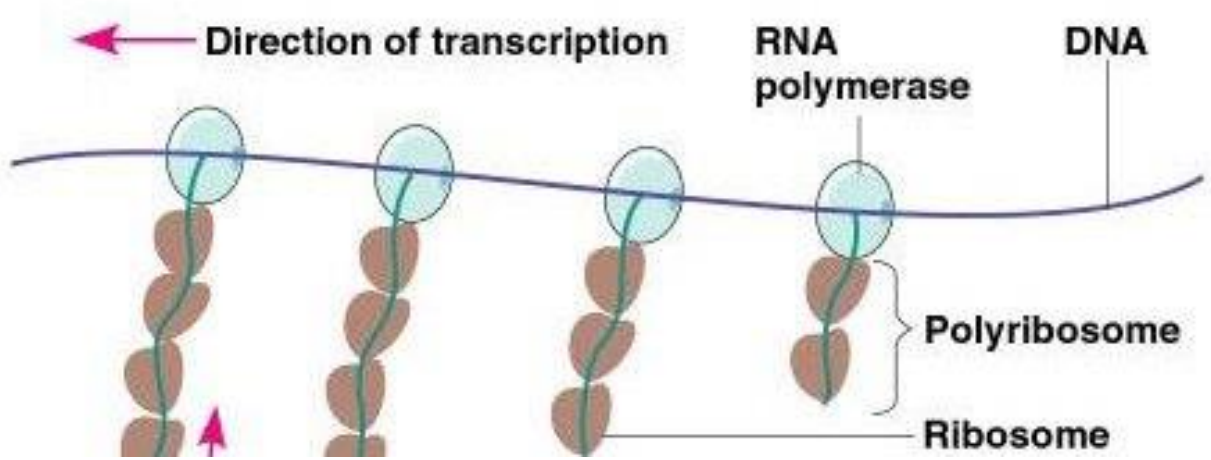


Figure: Prokaryotic transcription and translation occur simultaneously in the cytoplasm, and regulation occurs at the transcriptional level. Eukaryotic gene expression is regulated during transcription and RNA processing, which take place in the nucleus, and during protein translation, which takes place in the cytoplasm. Further regulation may occur through post-translational modifications of proteins.



In **electron micrographs**, polyribosomes appear as bead-like or spiral structures, with each "bead" representing a ribosome. The formation of polyribosomes greatly enhances the efficiency of protein synthesis, as many protein molecules can be synthesized from one mRNA before it degrades.





Polyribosome (Polysome)

- A **polyribosome**, also known as a **polysome**, is a complex formed when multiple ribosomes simultaneously translate a single messenger RNA (mRNA) molecule. This structure allows the cell to produce many copies of the same protein efficiently and rapidly from one mRNA transcript.
- As soon as the ribosome-binding site on the nascent mRNA becomes available, ribosomes attach and start translation. Multiple ribosomes can attach at different positions along the same mRNA, forming a chain-like structure called a polyribosome.
- Each ribosome on the mRNA moves independently along the transcript, translating the nucleotide sequence into a growing polypeptide chain. The ribosome closest to the 3′ end of the mRNA is usually the farthest along in translation, while those near the 5′ end are just beginning synthesis.
- In **eukaryotic cells**, polyribosomes also exist, either free in the cytoplasm or attached to the endoplasmic reticulum, forming **rough endoplasmic reticulum (RER)**. However, transcription and translation are spatially separated in eukaryotes — transcription occurs in the nucleus, and translation occurs in the cytoplasm.

Thus, the polyribosome is an essential feature of cellular protein synthesis, reflecting the high efficiency and coordinated control of gene expression in both prokaryotic and eukaryotic systems.

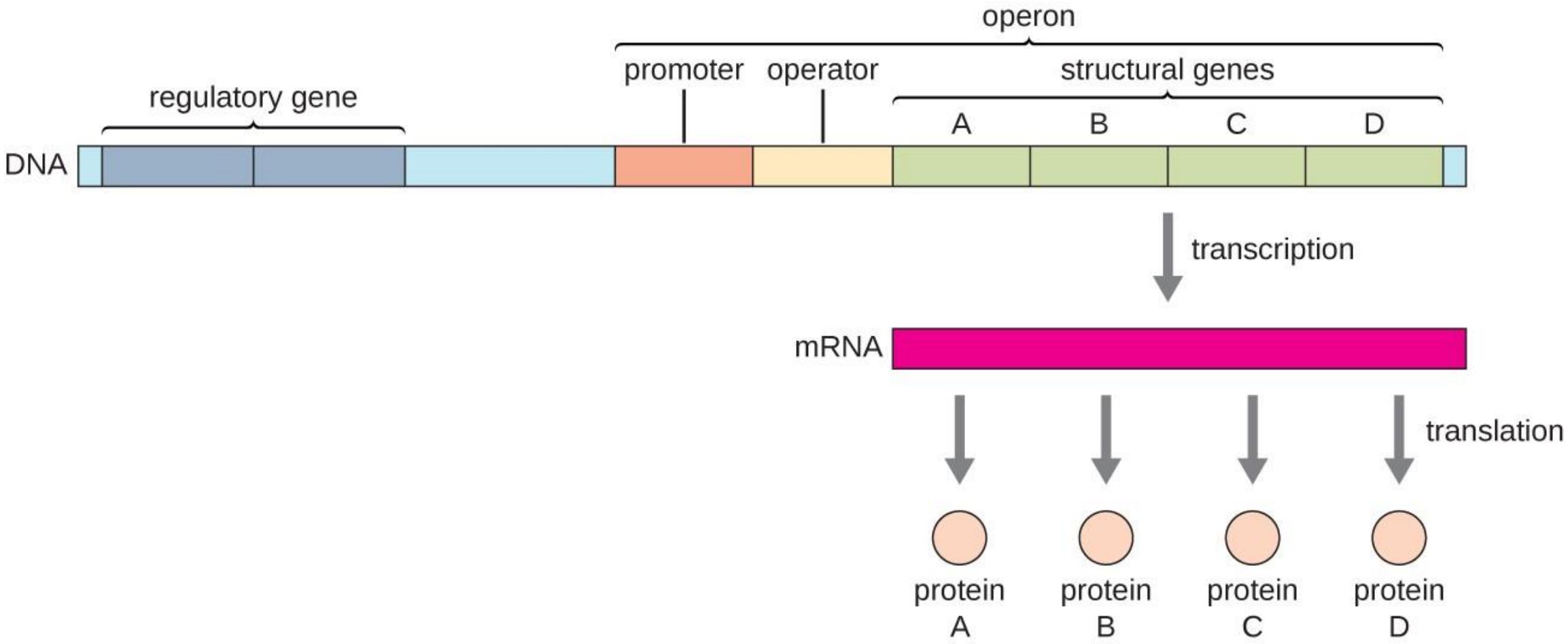
Transcriptional regulation in prokaryotes

- The DNA of prokaryotes is organized into a circular chromosome supercoiled in the nucleoid region of the cell cytoplasm. Proteins that are needed for a specific function, or that are involved in the same biochemical pathway, are encoded together in blocks called **operons**.
- An **operon -** a functional unit of the genome in prokaryotes, which includes cistrons that encode jointly or sequentially working proteins and are united under one (or several) promoters.
- **The concept** of an operon for prokaryotes was proposed in 1961 by the French scientists Jacob and Monod, for which they received the Nobel Prize in 1965.

François Jacob

Jacques Lucien Monod

REGULATORY ELEMENTS OF THE PROKARYOT GENOME



• **Operons** according to the number of **cistrons** are divided into **mono-, oligo- and polycistronic**, containing, respectively, only one, several or many cistrons (genes).

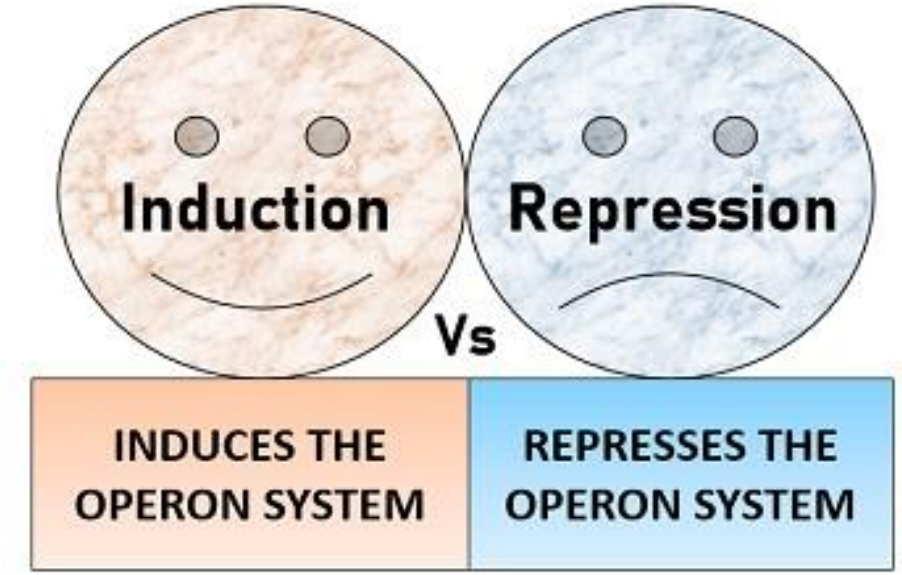
The operon principle of gene organization in prokaryotes

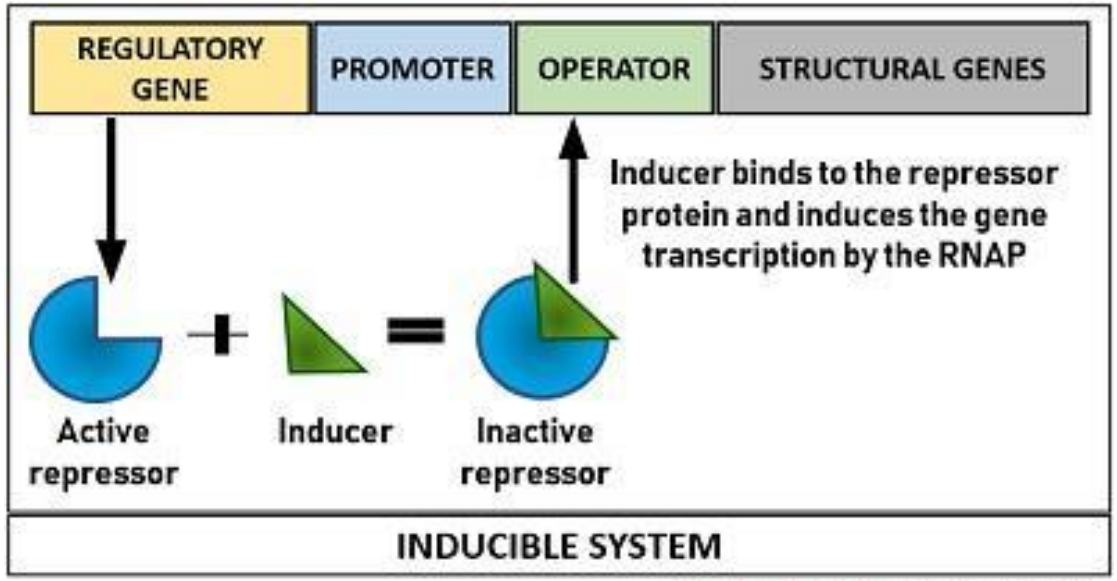
Advantages of the operon organization of prokaryotic genes::

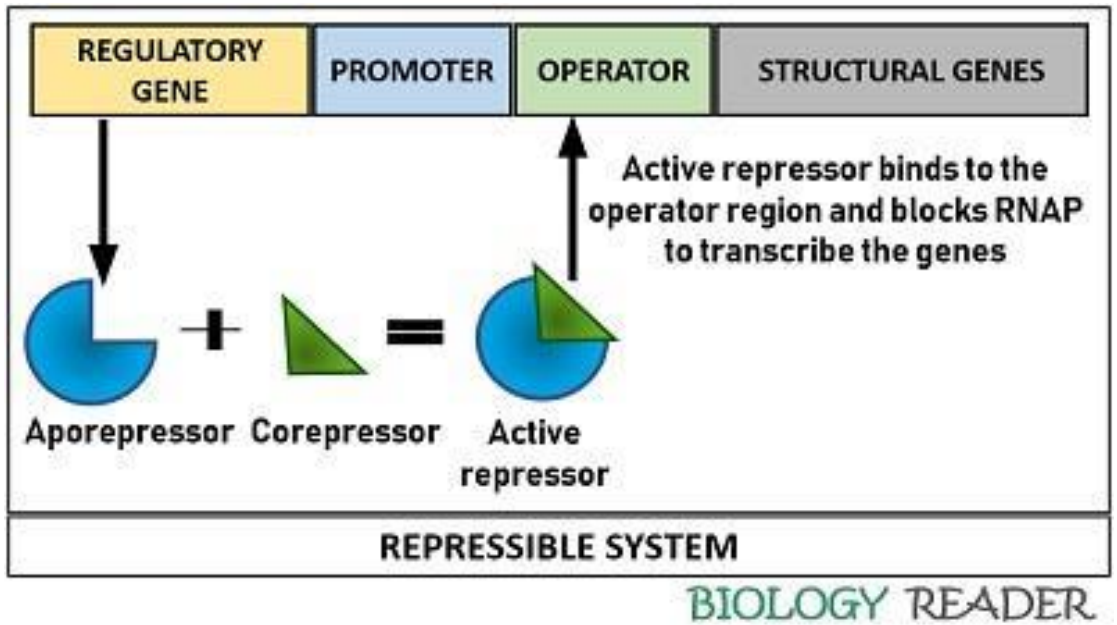
- compactness
- quick response to environmental changes: the synthesis of essential enzymes starts and stops at any moment.
- activity regulation coordination: all genes are expressed or not expressed in unison

Types of regulation of gene activity in prokaryotes

- ***Repression*** and ***induction*** of protein synthesis in prokaryotes implement the principles of adaptation to changing conditions of existence and cellular economy: enzymes appear in cells when there is a need for them, and cease to be produced if the need disappears.



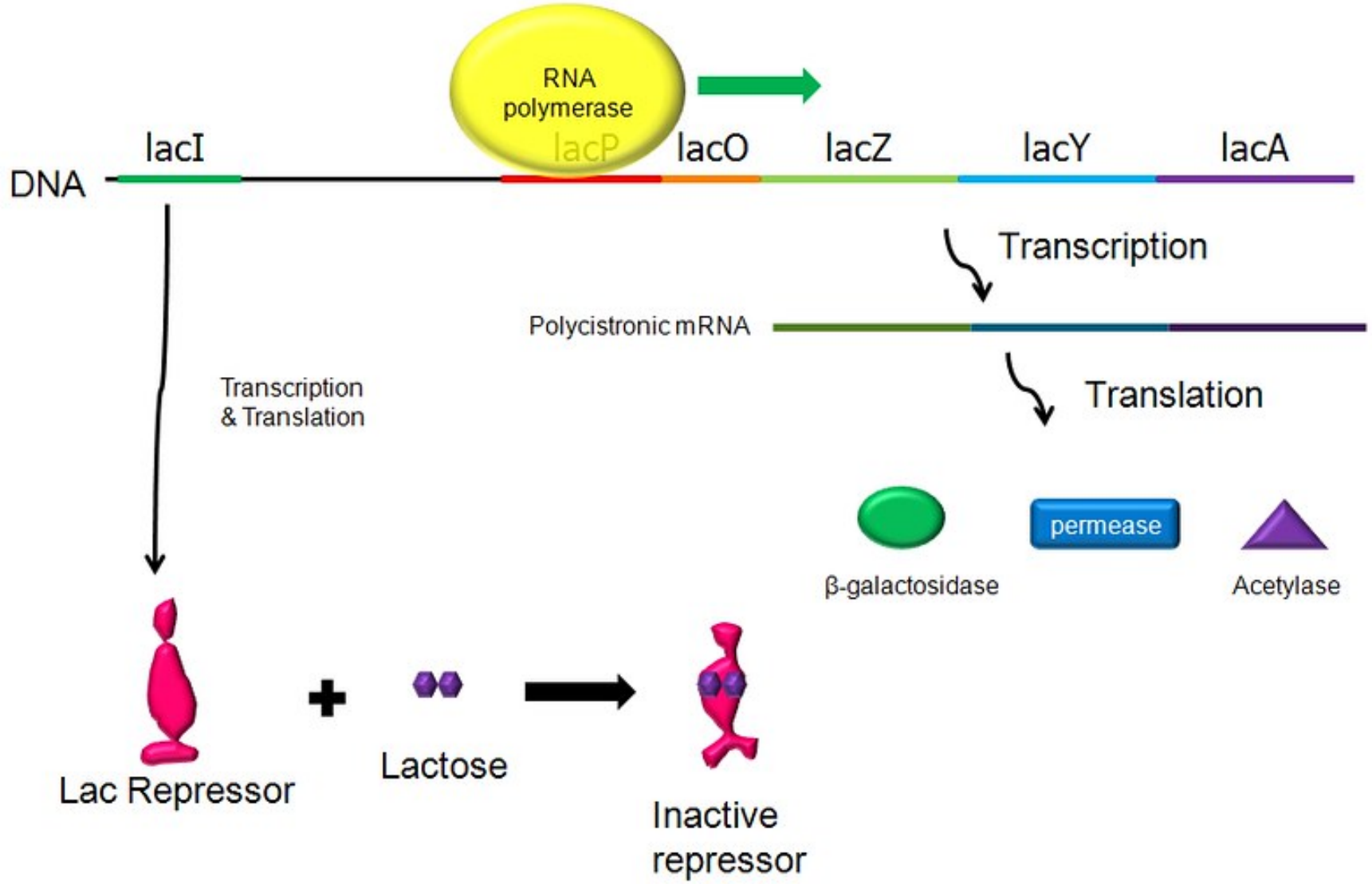

REGULATORY GENE
PROMOTER
OPERATOR
STRUCTURAL GENES
Inducer binds to the repressor protein and induces the gene transcription by the RNAP
Active repressor
Inducer
Inactive repressor
INDUCIBLE SYSTEM
BIOLOGY READER


REGULATORY GENE
PROMOTER
OPERATOR
STRUCTURAL GENES
Active repressor binds to the operator region and blocks RNAP to transcribe the genes
Aporepressor
Corepressor
Active repressor
REPRESSIBLE SYSTEM
BIOLOGY READER

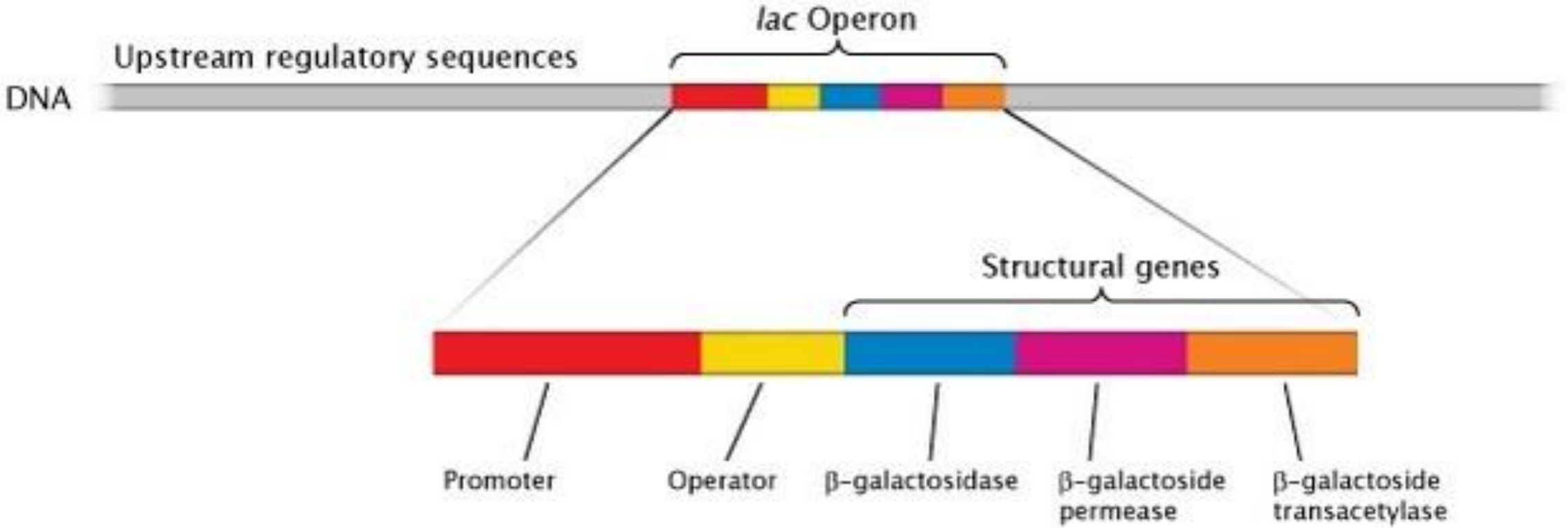
Comparison Chart

Properties	Induction	Repression
Meaning	The meaning of induction system states that it will induce the gene expression via an inducer	The meaning of repression system states that it will suppress the gene expression via a corepressor
Regulation	The operon system regulates the synthesis of enzymes that are stimulated by the addition of inducer	The operon system suppresses the enzyme synthesis, which is facilitated by the existing end-product or corepressor molecules
Key element	Inducer or anti-repressor	Corepressor or effector molecule
Mechanism	Inducer inactivates the repressor protein by preventing the attachment of a repressor to the operator region	Corepressor activates the apo-repressor protein (inactive) and allows the attachment of an active repressor to the operator region
Effect on the control system	Mediates movement of the RNA polymerase along the control system, i.e. promoter and operator region	Blocks the movement of RNA polymerase along the promoter and operator region
Effect on structural genes	It initiates the expression of structural genes	It inhibits the expression of structural genes
Overall impact	An induction system activates or turns on the whole operon system through an adequate supply of the inducer metabolites	A repression system terminates or switches off the entire operon system under an adequate level of corepressor molecules
Operon system	The genetic system regulated by the presence or absence of an inducer is called inducible operon	The genetic system regulated by the presence or absence of a corepressor is called repressible operon
Enzyme	Enzymes synthesis stimulated by the addition of inducer metabolites are termed as inducible enzymes	Enzymes synthesis inhibited by the addition of corepressor molecules are termed as repressible enzymes
Example	Lactose operon	Tryptophan operon
Metabolic pathway	It operates a catabolic synthesis	It operates an anabolic synthesis

Structure of the lactose operon



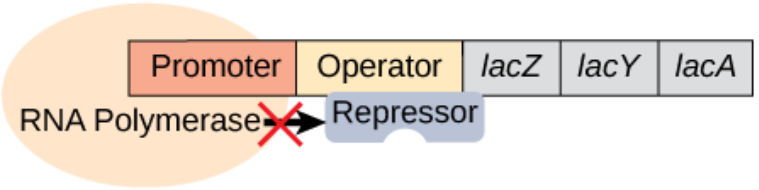
Operon structure



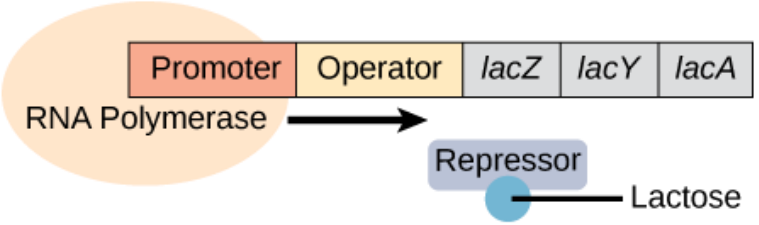
Expressed genes can be divided into the following categories:

- **constitutive**, present in cells in constant quantities, regardless of the metabolic state of the body;
- **induced**, their concentration under normal conditions is low, but can increase by a factor of 100 or more if, for example, a substrate of such an enzyme is added to the cell culture medium;
- **repressed**, i.e. enzymes of metabolic pathways, the synthesis of which stops when the end product of these pathways is added to the growth medium.

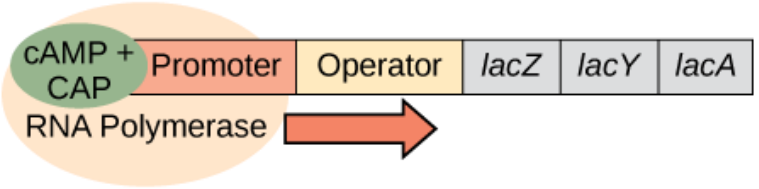
In the absence of lactose, the lac repressor binds the operator, and transcription is blocked.



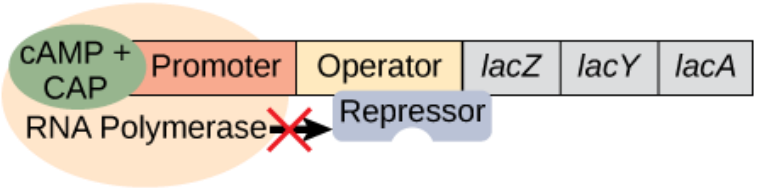
In the presence of lactose, the lac repressor is released from the operator, and transcription proceeds at a slow rate.



cAMP-CAP complex stimulates RNA Polymerase activity and increases RNA synthesis.



However, even in the presence of cAMP-CAP complex, RNA synthesis is blocked when repressor is bound to the operator.



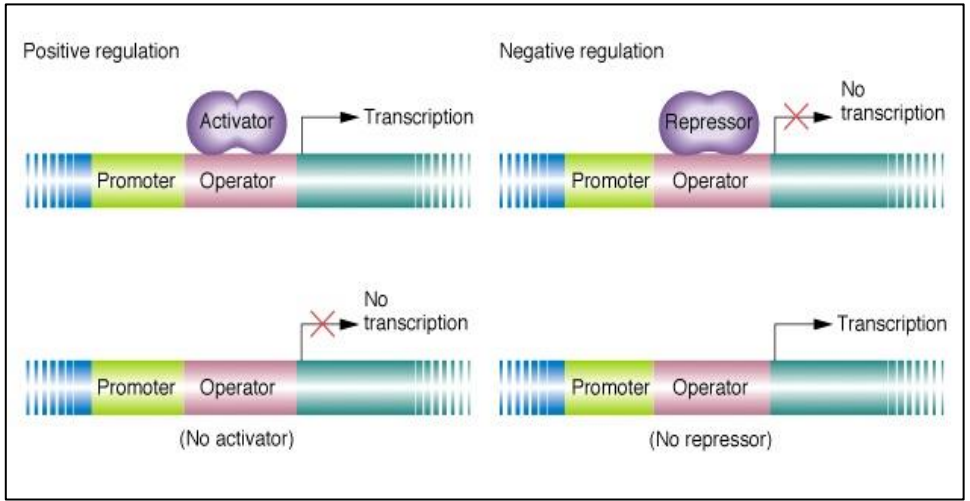
In addition to control by the **lac repressor**, the lac operon is also regulated by a **catabolite activator protein (CAP)**, also known as the **cAMP receptor protein (CRP)**. CAP regulation is an example of **positive control** and ensures that the lac operon is expressed only when glucose, the preferred energy source, is scarce.

Conversely, when glucose is abundant, the intracellular concentration of cAMP decreases. As a result, CAP cannot bind to the promoter, and RNA polymerase binding becomes inefficient, even if lactose is present. This phenomenon, known as **catabolite repression**, ensures that the cell preferentially uses glucose before switching to alternative sugars like lactose.

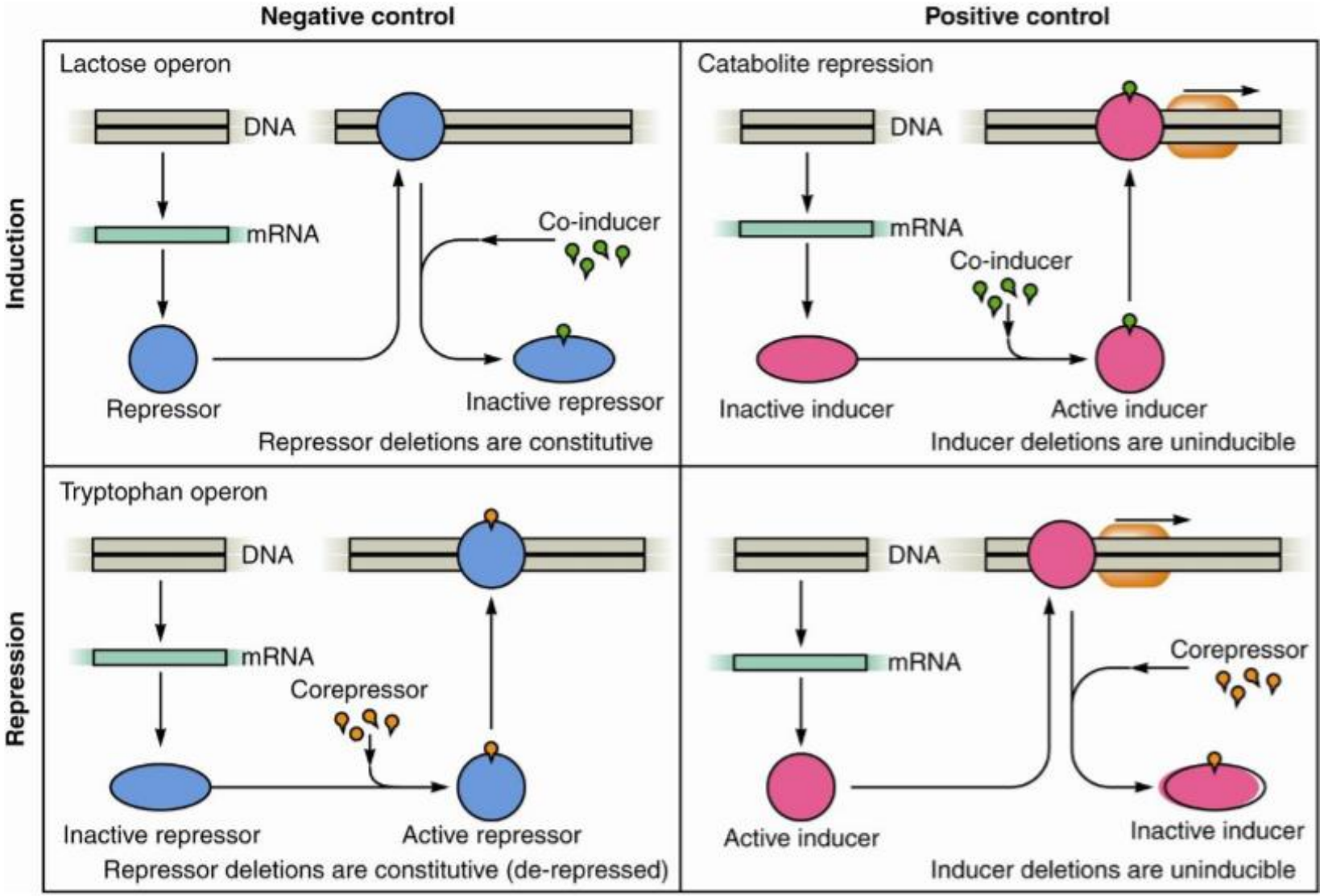
The following schemes of transcriptional regulation in prokaryotes are distinguished:

1) negative induction (Jacobeau and Monod);
2) positive induction;
3) positive repression;
4) negative repression.

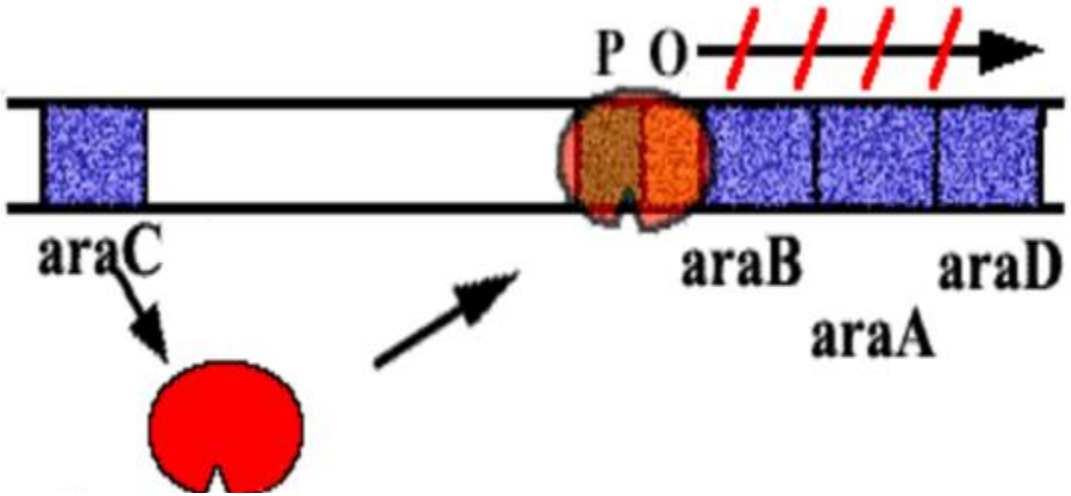
	Regulatory protein is present	Example of regulatory protein	Mutate regulatory gene to lose function
Positive control	Operon ON	Activator	Operon OFF
Negative control	Operon OFF	Repressor	Operon ON



Positive/negative inducible operons and positive/ negative repressible operons



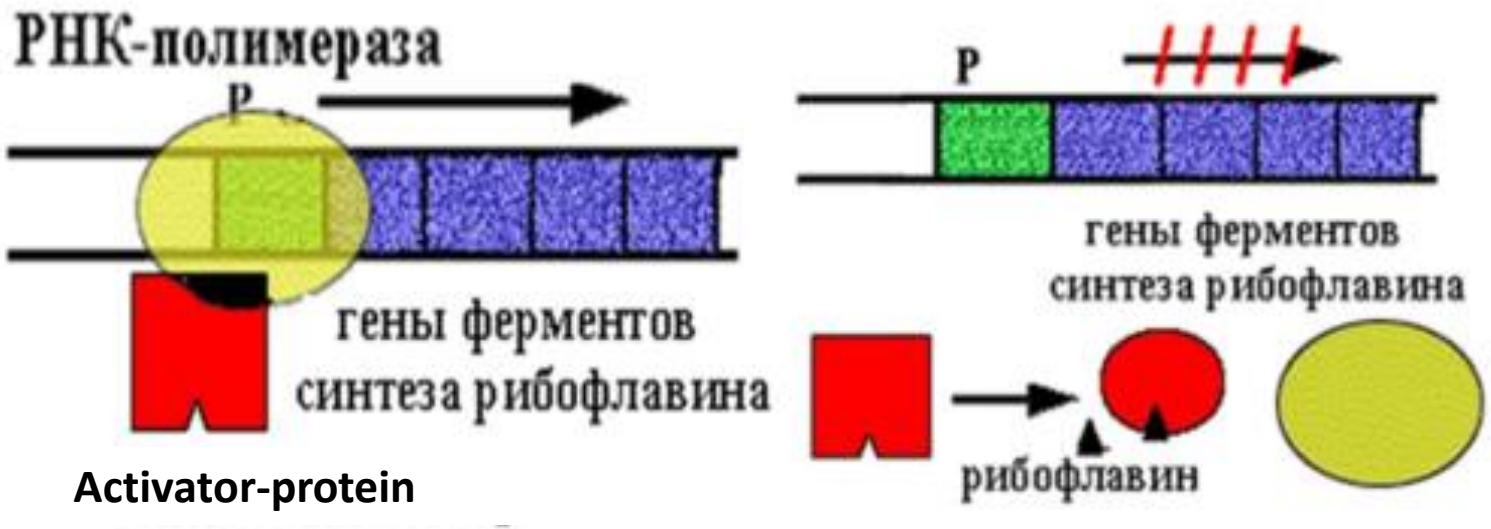
Positive induction scheme - arabinose opera (Aga-operon) Escherichia coli



Repressor-protein

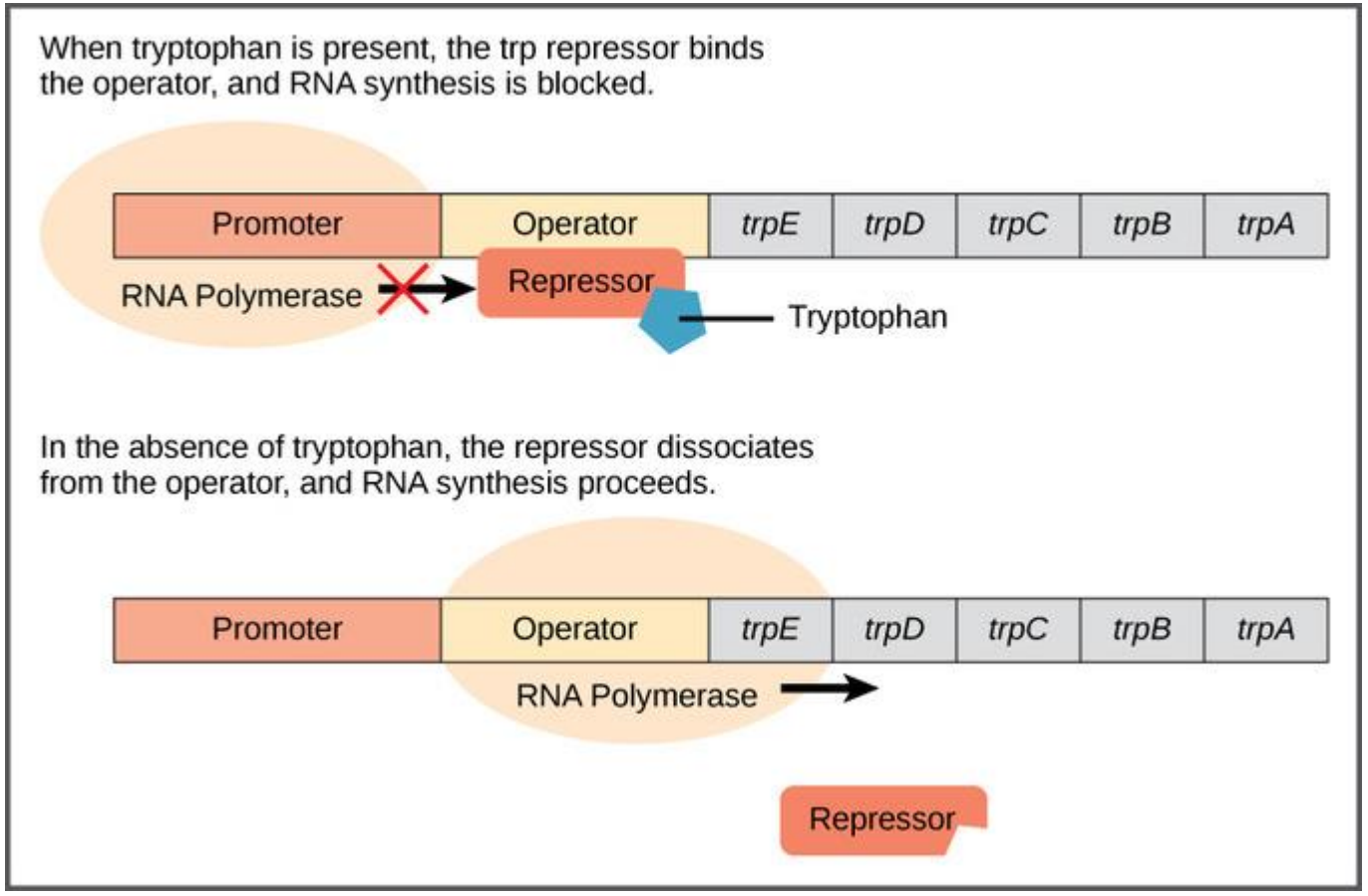
This operon has 3 cistrons that code for enzymes that break down arabinose sugar. Normally, the operon is closed. The repressor protein is linked to the operator. When arabinose enters the cell, it interacts with the repressor protein. The repressor protein changes conformation and turns from a repressor into an activator that interacts with the promoter and facilitates the landing of RNA polymerase on the promoter.

Positive repression scheme — riboflavin synthesis operon in *Bacilus subtilis*



The operon contains cistrons of riboflavin synthesis enzymes. There is an activator protein that ensures the landing of RNA polymerase on the promoter. Normally, the operon is open. N molecules of riboflavin are formed. The N + 1st molecule (extra) interacts with the activator, and it loses its ability to activate the binding of RNA polymerase to the promoter.

Scheme of negative repression by the operation of tryptophan synthesis in *Escherichia coli*

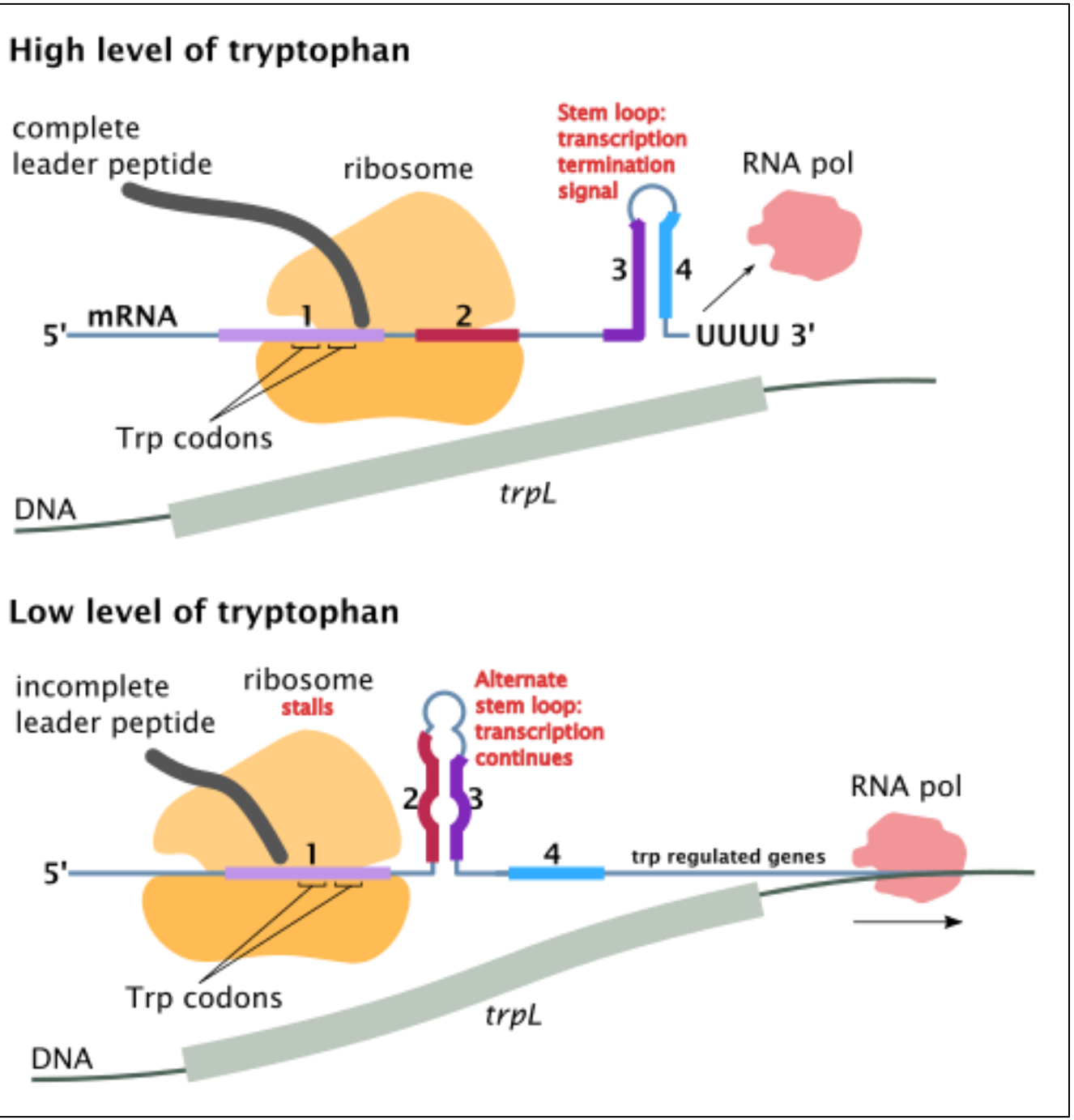


The **tryptophan (trp) operon** in *Escherichia coli* is a classic example of a gene regulation mechanism that controls amino acid biosynthesis through both **repression** and **attenuation**. This dual control allows the cell to fine-tune the production of enzymes involved in tryptophan synthesis according to the availability of this amino acid.

The **repressor mechanism** controls transcription initiation: when intracellular tryptophan levels are high, tryptophan acts as a **corepressor**, binding to the **trp repressor** protein. The activated repressor binds to the operator region of the operon, preventing RNA polymerase from initiating transcription.

Negative repression, because the repressor protein "turns off" the operon. There are 5 cistrons in the operon, which encode the enzymes of a sequential chain of tryptophan synthesis reactions. Normally, the operon is turned on. The repressor protein is inactive (in the form of an apo-repressor), it is not able to sit on the operator. A cell needs N molecules of tryptophan. The N+1 molecule interacts with the apo-repressor. It changes conformation, sits on the operator and RNA synthesis stops.

Attenuation of the Tryptophan Operon



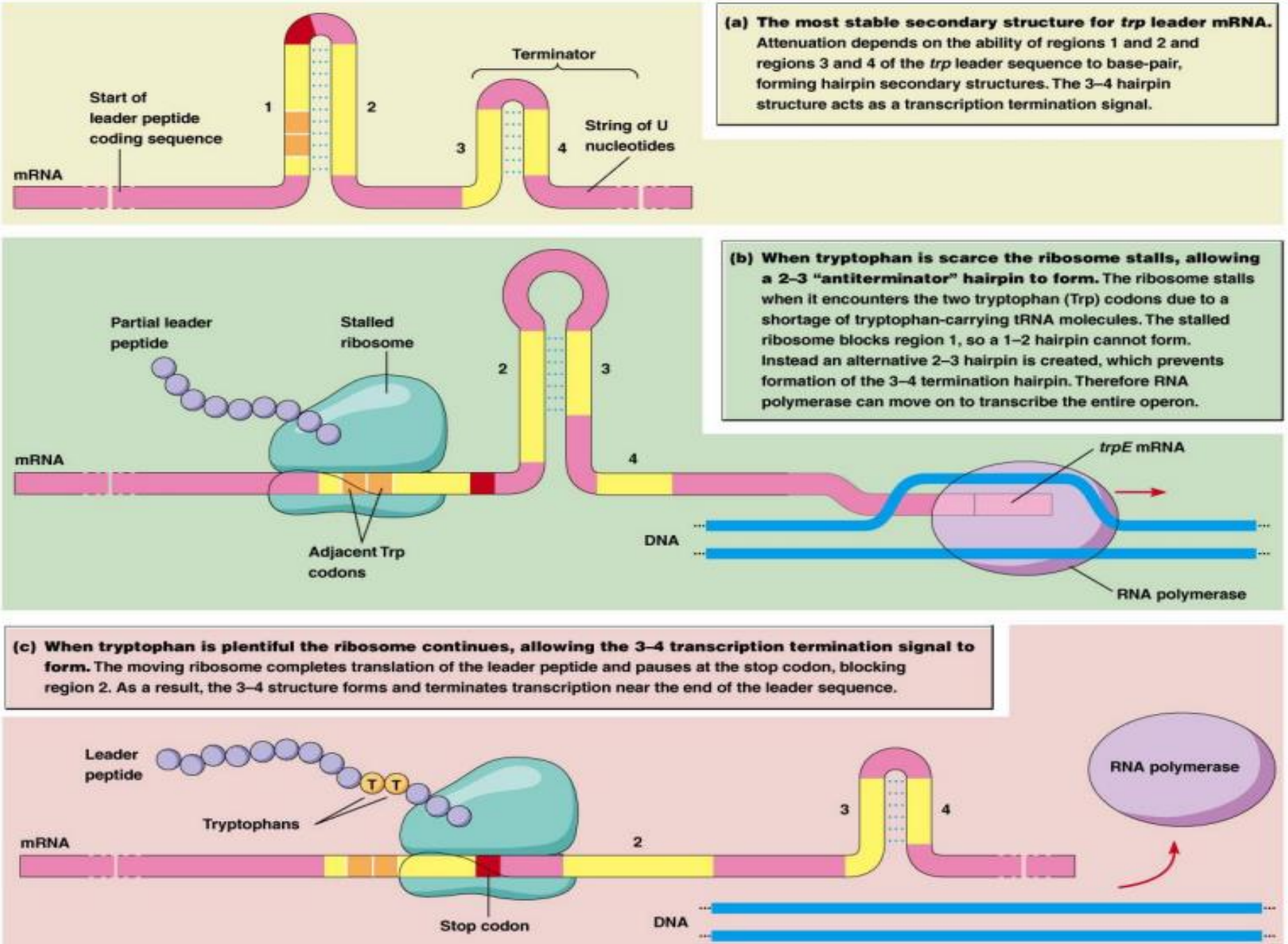
- However, even when some transcription begins, a second level of control — **attenuation** — regulates whether the RNA polymerase continues to transcribe the entire operon. This attenuation mechanism depends on the coupling of **transcription and translation** in prokaryotes.

- At the beginning of the trp operon lies a short **leader sequence (trpL)** that encodes a small peptide called the **leader peptide**, which contains two adjacent tryptophan codons. The mRNA of this leader region can fold into alternative **secondary structures** that determine whether transcription continues or terminates prematurely.

- When **tryptophan levels are high**, charged **$tRNA^{Trp}$** molecules are abundant. The ribosome quickly translates the leader peptide, reaching a stop codon before the RNA polymerase has transcribed the downstream genes. This allows the formation of a **terminator hairpin** structure (regions 3–4) in the mRNA, which causes **premature termination** of transcription.

- When **tryptophan levels are low**, the ribosome stalls at the two tryptophan codons in the leader peptide because of a shortage of charged **$tRNA^{Trp}$**. This stalling prevents the formation of the terminator hairpin and instead promotes the formation of an **antiterminator structure** (regions 2–3), allowing RNA polymerase to continue transcription of the full operon.

Thus, **attenuation** acts as a **fine-tuning mechanism** that adjusts transcription according to the current tryptophan concentration in the cell. Together with the repressor system, it ensures that the enzymes responsible for tryptophan biosynthesis are produced only when needed, maintaining metabolic efficiency and balance.



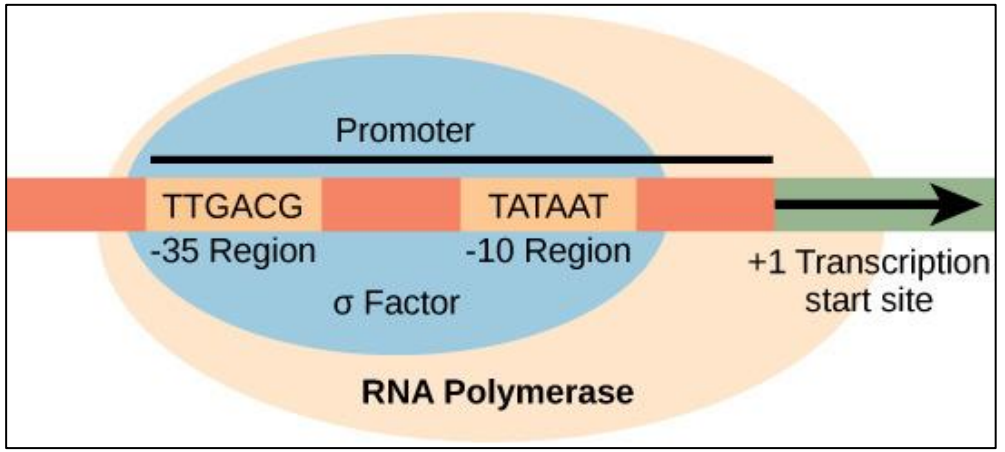
(a) The most stable secondary structure for *trp* leader mRNA. Attenuation depends on the ability of regions 1 and 2 and regions 3 and 4 of the *trp* leader sequence to base-pair, forming hairpin secondary structures. The 3–4 hairpin structure acts as a transcription termination signal.

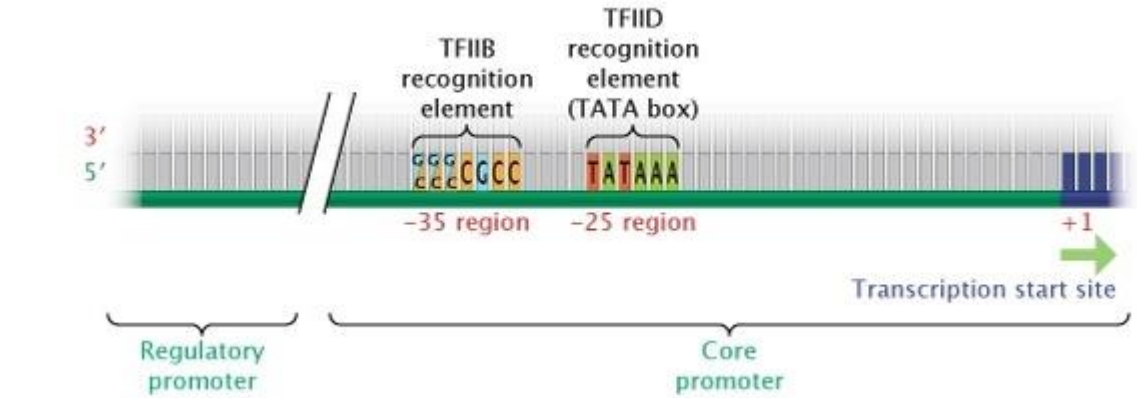
(b) When tryptophan is scarce the ribosome stalls, allowing a 2–3 "antiterminator" hairpin to form. The ribosome stalls when it encounters the two tryptophan (Trp) codons due to a shortage of tryptophan-carrying tRNA molecules. The stalled ribosome blocks region 1, so a 1–2 hairpin cannot form. Instead an alternative 2–3 hairpin is created, which prevents formation of the 3–4 termination hairpin. Therefore RNA polymerase can move on to transcribe the entire operon.

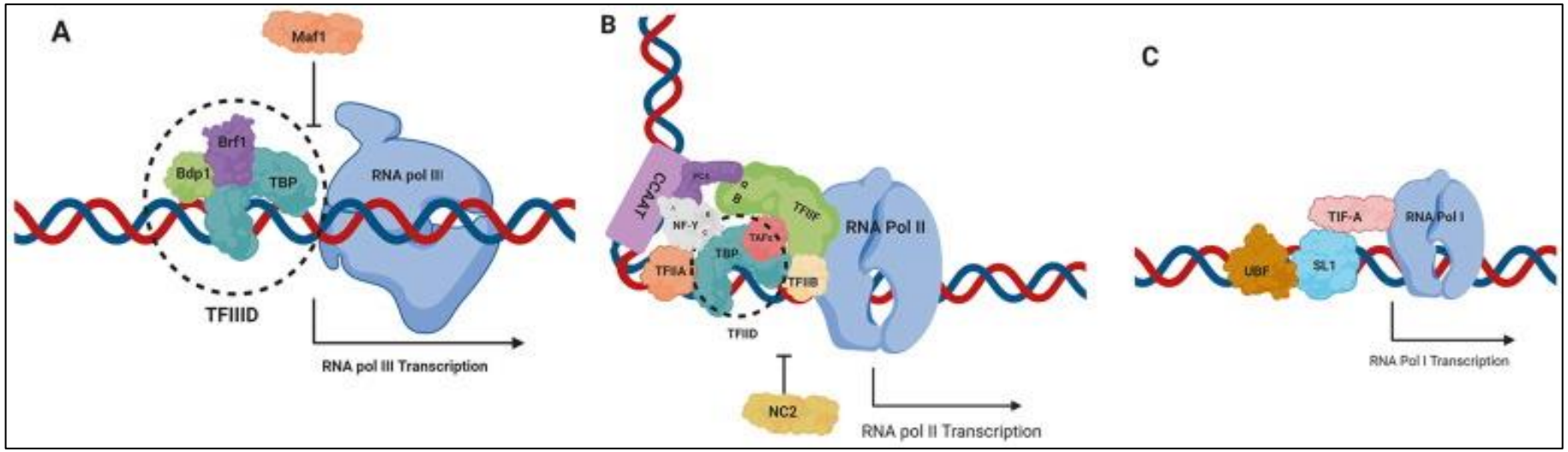
(c) When tryptophan is plentiful the ribosome continues, allowing the 3–4 transcription termination signal to form. The moving ribosome completes translation of the leader peptide and pauses at the stop codon, blocking region 2. As a result, the 3–4 structure forms and terminates transcription near the end of the leader sequence.

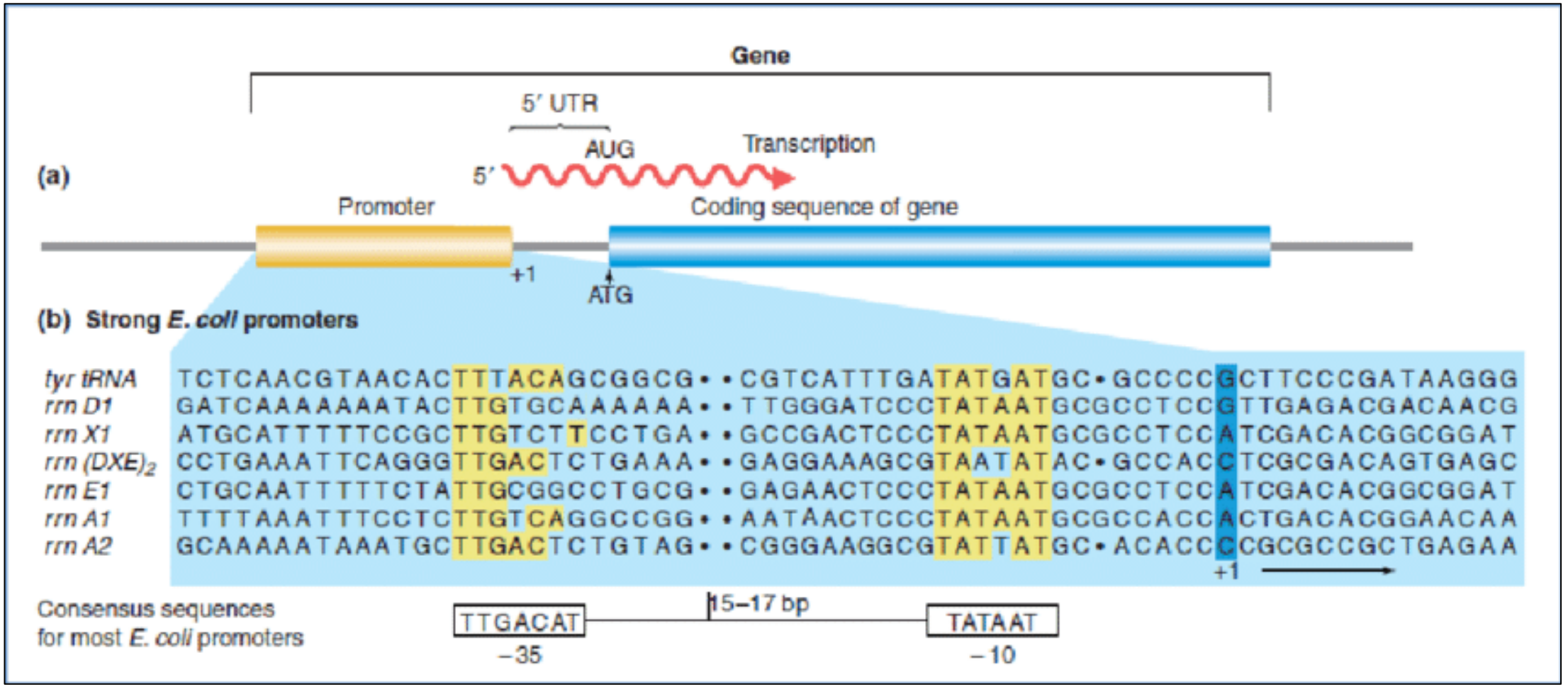
Most E. coli structural genes have two binding sites for RNA polymerase. One of them is usually a nucleotide sequence
TATAAT
ATATTA (TATA-box, or Pribnov box),

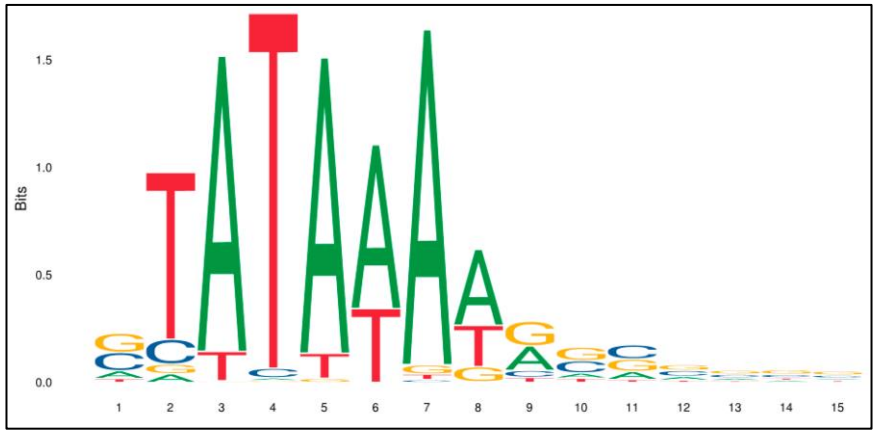
and the other –
TTGAC
AACTG.
The TATA box and the TTGAC sequence are located 10 (region -10) and 35 (region -35) nucleotides before the site of transcription initiation, respectively (nucleotide +1)











A frequency matrix for a TATA box [57] presented as a sequence logo.

References

1) Shchelkunov, S. N.Genetic Engineering: A Reference Manual.3rd Edition. Novosibirsk: Siberian University of Information, 2008.
2) Zhimulev, I. F.General and Molecular Genetics: A Textbook.Novosibirsk: Siberian University of Information, 2007.
3) Sambrook, J., & Russell, D. W.Molecular Cloning: A Laboratory Manual.3rd Edition. Cold Spring Harbor Laboratory Press, New York, 2001.
4) Madigan, M. T., Martinko, J. M., Bender, K., Buckley, D., & Stahl, D.Brock Biology of Microorganisms. 15th Edition. Pearson, 2017.
5) Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., & Walter, P.Molecular Biology of the Cell. 7th Edition. Garland Science, 2015.