Transcription Regulation in Prokaryotes

Dr. R. Prasad
Department of Zoology,
Eastern Karbi Anglong College,
Sarihajan

PRINCIPLES OF TRANSCRIPTIONAL REGULATION

The biological properties of each cell type are largely dependent on the proteins expressed within it. This array of expressed proteins determines much of the cell's architecture, its enzymatic activities, interaction with its environment and many other physiological properties. As the final product of the gene expression is a protein, regulation could be achieved by controlling the transcription of DNA into RNA or the translation of RNA into protein. However, most of the regulation is thought to take place at the level of gene transcription. First, the molecular signals from inside or outside the cell initiate the binding of the regulatory proteins to specific DNA sites outside of protein-encoding regions. Binding of these proteins modulates the rate of transcription. These proteins may directly or indirectly assist *RNA polymerase* in binding to its transcription initiation sites- the promoter; or they may repress transcription by preventing the binding of *RNA polymerase*. The regulation of transcription is a vital process in all living organisms. Most of our information regarding the mechanism of regulation of gene expression has been obtained from *E. coli.* Although bacteria and eukaryotes have much of the logistics of gene regulation in common, there are some fundamental differences in the underlying mechanisms and machinery. Both use sequence-specific DNA-binding proteins to modulate the level of transcription. However, eukaryotic genomes are larger and have a broader range of properties. So naturally their regulation is more complex requiring many more types of regulatory proteins and more extensive interactions with the adjacent regulatory regions in the DNA.

14.3 GENE REGULATION IN PROKARYOTES

Most bacteria are free- living organisms that grow by increasing mass and then divide by binary fission. The growth and division of these bacteria are controlled by genes, the expression of which must be regulated appropriately. *Escherichia coli*, a bacterium that resides in our large intestine, is solely dependent on the nutrients available by us. It must survive under different nutritional environments. *E. coli* makes up for its inability to alter external environment by being internally flexible. For example, if glucose is present, this bacterium utilizes it to generate ATP. In the absence of glucose, it is capable of utilizing lactose, arabinose, maltose, xylose, or any of several other sugars. Similar is the case with amino acids. When amino acids are available, *E. coli* uses them to synthesize proteins. If a particular amino acid is absent, *E. coli* produces enzymes needed to synthesize that amino acid. Thus, *E. coli* has the capability to respond to environmental changes by rapidly regulating gene expression and altering its enzymes. There are two types of genes. Genes whose activity is controlled in response to the needs of a cell or organism are called **regulated genes** and genes whose products are essential to the normal functioning and dividing cells, irrespective of the external conditions are known as **constitutive genes** or **housekeeping genes**. These genes must continue to function under all conditions, e.g., genes that code for enzymes needed for protein synthesis and glucose metabolism.

Bacteria use two different strategies for regulating enzyme synthesis, depending on whether a given enzyme is involved in a catabolic (degradative) or anabolic (synthetic) pathway. The enzymes that catalyze such pathways are often regulated synchronously, synthesis of all the enzymes involved in a particular pathway is turned on and off together. Now we will study two well-understood pathways, one catabolic and one anabolic.

Catabolic Pathways and Substrate Induction

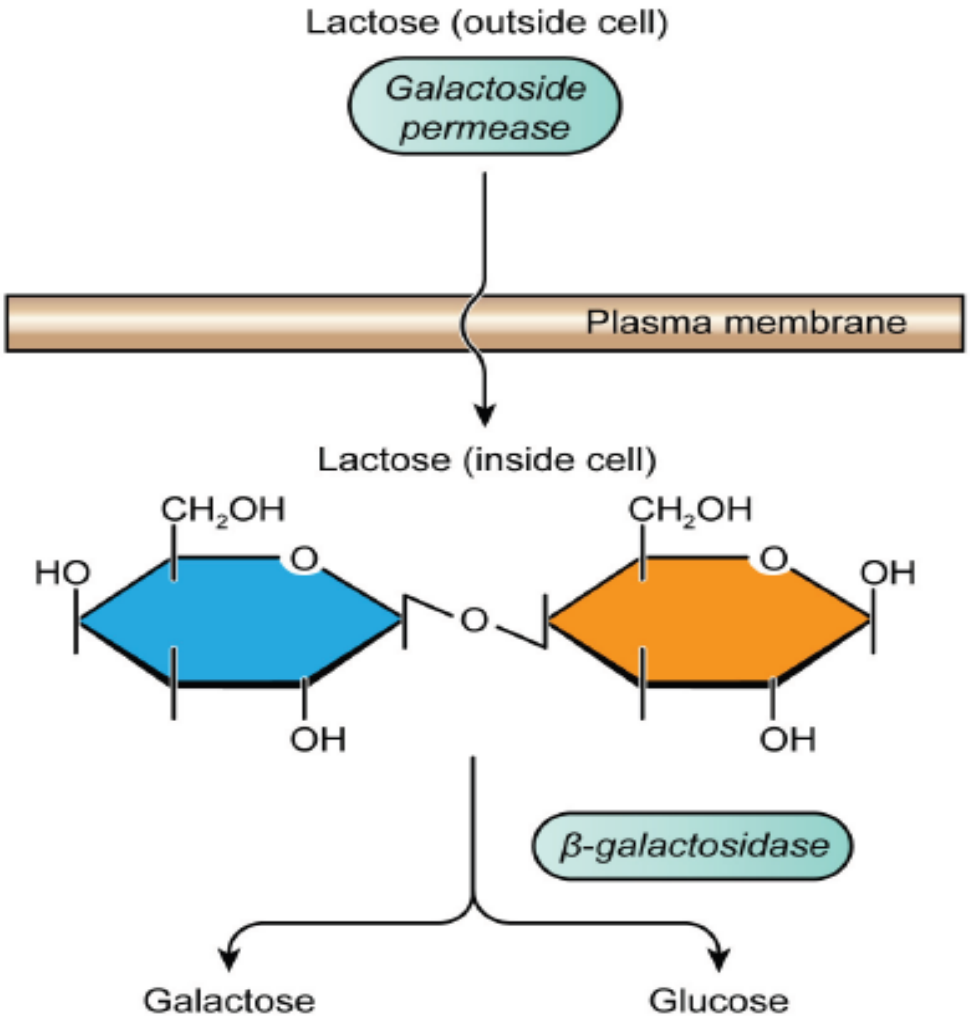


Fig. 14.1: A typical catabolic pathway showing the coordinated regulation of the synthesis of the enzyme *β galactosidase* (for the breakdown of the diasaccharide lactose into glucose and galactose).

Anabolic Pathways and End-Product Repression

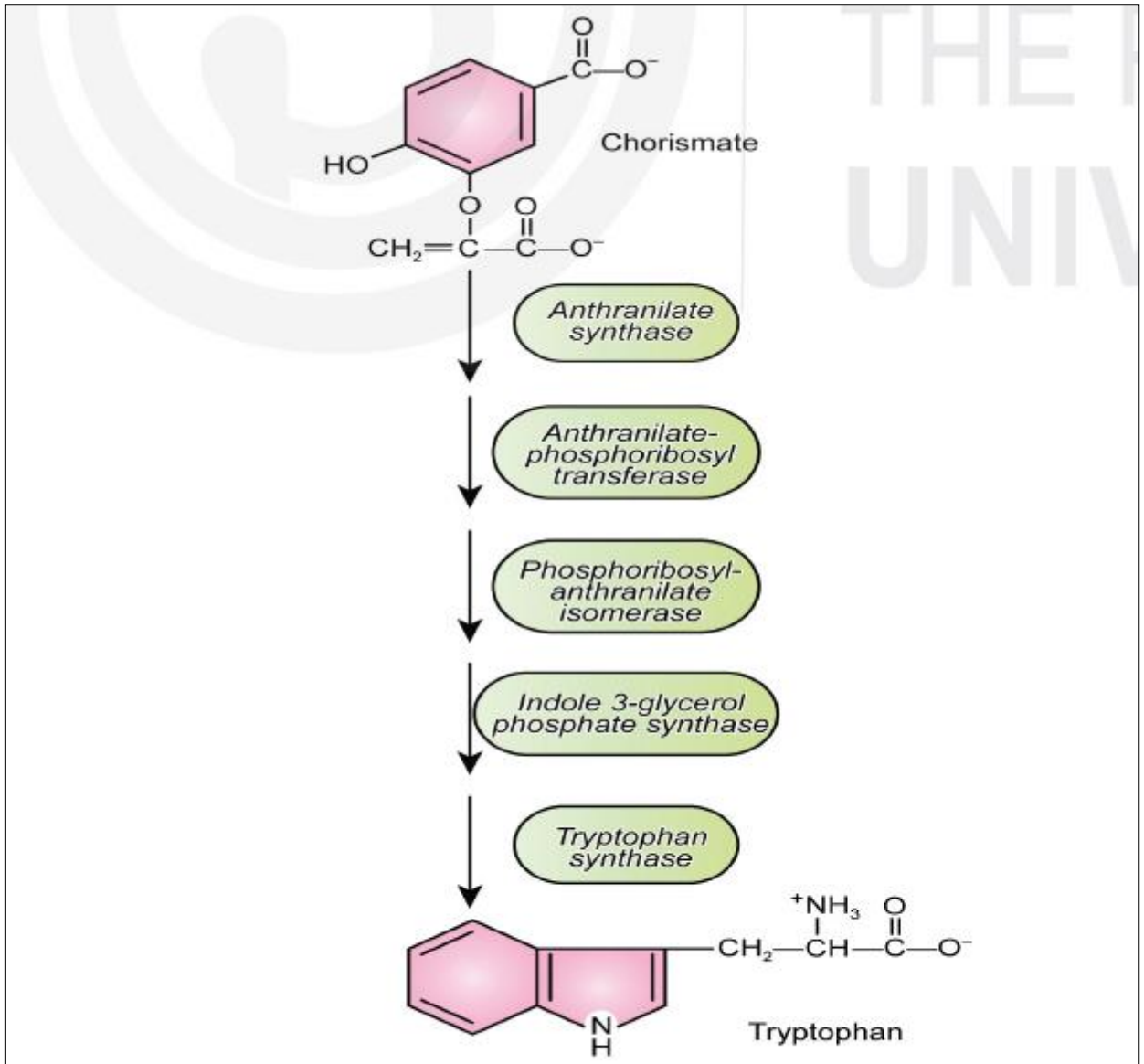


Fig. 14.2: A typical anabolic pathway showing coordinated regulation of enzymes for the synthesis of tryptophan from chorismate.

Mechanisms of Gene Regulation at Transcriptional Level

Operon Control

Much of what we know about the control of gene expression in bacteria, including the vocabulary used to express that knowledge is based on the pioneering studies carried out by French molecular geneticists Francois Jacob and Jacques Monod. The "operon model" described by them in 1961 probably influenced our understanding of gene regulation more than any other work. In prokaryotes the functionally related genes are clustered together, and their DNA sequences are controlled by a single promoter. A significant difference between bacterial and eukaryotic transcriptional control lies in the organization of functionally related genes that are clustered together with DNA sequences under the control of a single promoter that allows the genes to be turned on (forming single mRNA) and off simultaneously as a single unit. Such a group of bacterial structural genes that are transcribed together is called an **operon**. The organization of functionally related genes into operons is commonly observed in prokaryotes but not in eukaryotes. Nonetheless, the operon model established several basic principles that have shaped our understanding of transcriptional regulation in both prokaryotic and eukaryotic systems.

Operon Structure

An operon is a group of structural genes and sequences that control transcription. The organization of a typical operon has been illustrated in Fig. 14.3. A set of **structural genes** such as gene *a*, gene *b* and gene *c* is present at the end of an operon. These structural genes are transcribed into a single mRNA, which is subsequently translated to produce enzymes A, B, and C. These enzymes carry out a series of biochemical reactions that convert the precursor molecule X into product Y. All the three structural genes are transcribed together under the control of a promoter which lies upstream of the first structural gene. The *RNA polymerase* first binds to the promoter and then moves downstream, transcribing the structural genes.

A **regulator gene** helps to control the transcription of the structural genes of the operon. Even though this gene is not considered as part of operon, it significantly affects operon function. The regulator gene possesses its own promoter and is transcribed into a short mRNA segment which is later translated into a small protein (**regulator protein**).This protein binds to a region of the operon called the **operator** and affects transcription. The operator usually overlaps the 3´ end of the **promoter** and in some cases even the 5´end of the first structural gene.

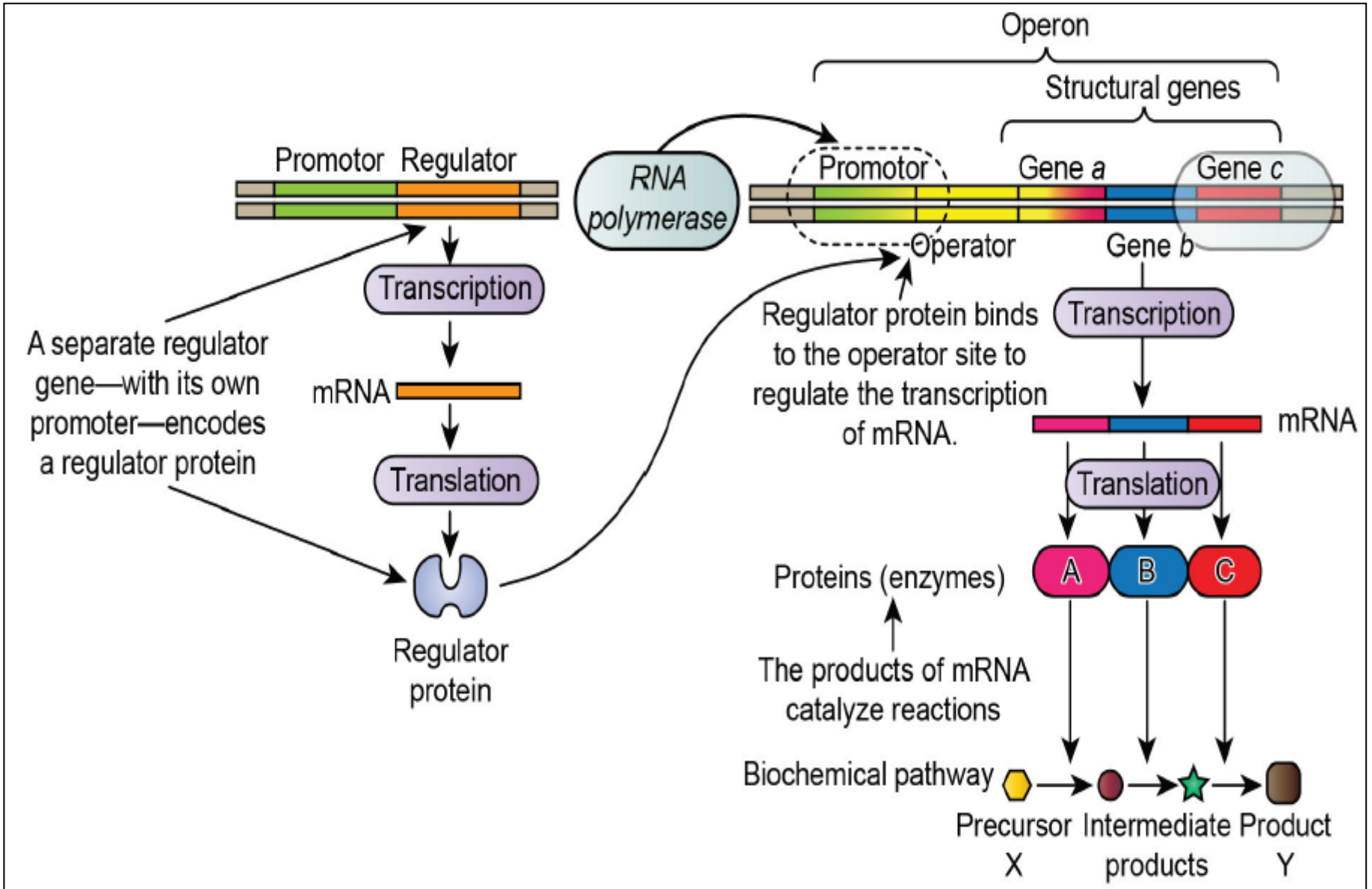


Fig. 14.3: An operon as a single transcriptional unit that includes a series of structural genes, a promoter, and an operator.

Strategies of Gene Regulation

In bacterial cells there are two types of transcriptional regulation: **negative regulation** and **positive regulation** (Fig.14.4).

Negative Regulation: In this type of regulation the regulatory protein present is called an **inhibitor or repressor**. The repressor binds to DNA and inhibiting or suppressing transcription of relevant gene. A signal molecule or inducer is required to react with or remove repressor to allow the initiation of transcription. Therefore, the negative regulation of gene expression is also called repression.

Positive Regulation: In positive regulation an **activator or inducer** molecule activates the promoter which in turn promotes transcription of mRNA. The positive regulation of gene regulation is also called induction and the substance that induces gene action is called inducer.

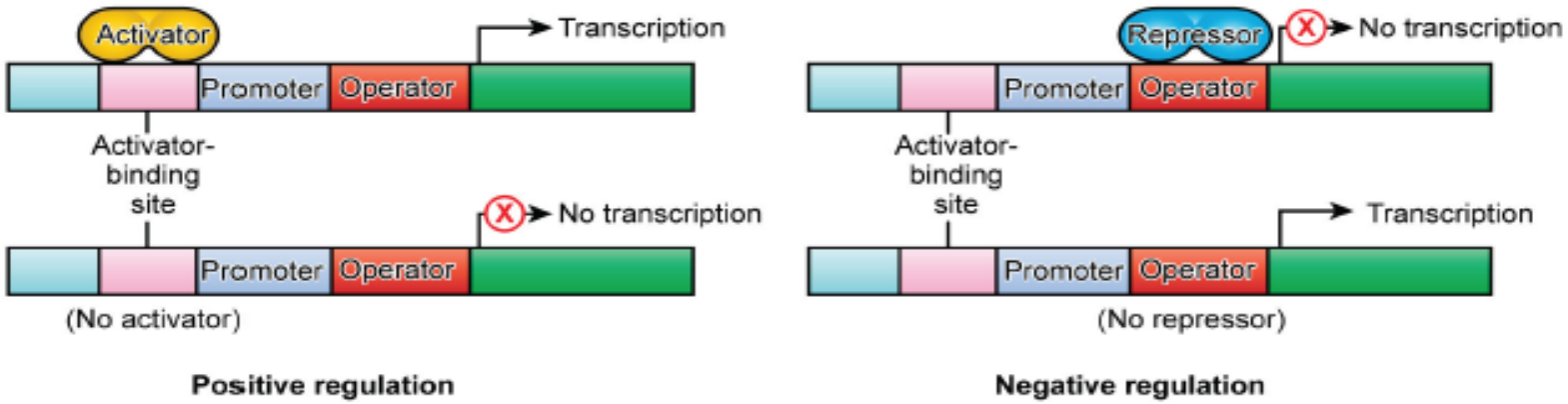


Fig. 14.4: Negative and positive regulation of bacterial gene action by regulatory protein control.

Operons can also be either **inducible** or **repressible**:

Inducible operons: In such systems gene or genes normally remain switched off. They must be induced or 'switched on' to produce mRNA needed for the synthesis of required enzyme or enzymes. The substance which induces the gene for protein synthesis or enzyme production is known as an **inducer**.

Repressible operons: In this system gene or genes are normally switched on. The activity of these genes is suppressed, and the synthesis of specific protein is stopped or reduced. The substance which stops or suppresses the protein synthesis is known as **repressor**.

There are several varieties of basic control mechanisms for regulation of transcription in bacteria:

Negative inducible operons: The **regulator gene** encodes an active repressor that readily binds to the **operator**. In the absence of an inducer, the repressor binds to the operator and physically blocks the binding of *RNA polymerase* to the promoter. This is because the operator sites overlap the promoter site. As a result, the structural genes *d, e*, and *f* (which metabolize precursor V) cannot synthesize (Fig.14.5 a). Hence, there is no transcription.

Transcription can only be turned on when a small molecule, an **inducer**, binds to the repressor (Fig.14.5 b).The regulatory proteins frequently have two binding sites: one that binds to DNA and another that binds to a small molecule called **inducer**. Binding of the inducer (precursor V in Fig. 14.5 b) alters the shape of the repressor, preventing it from binding to DNA. The DNA molecule is now free to transcribe.

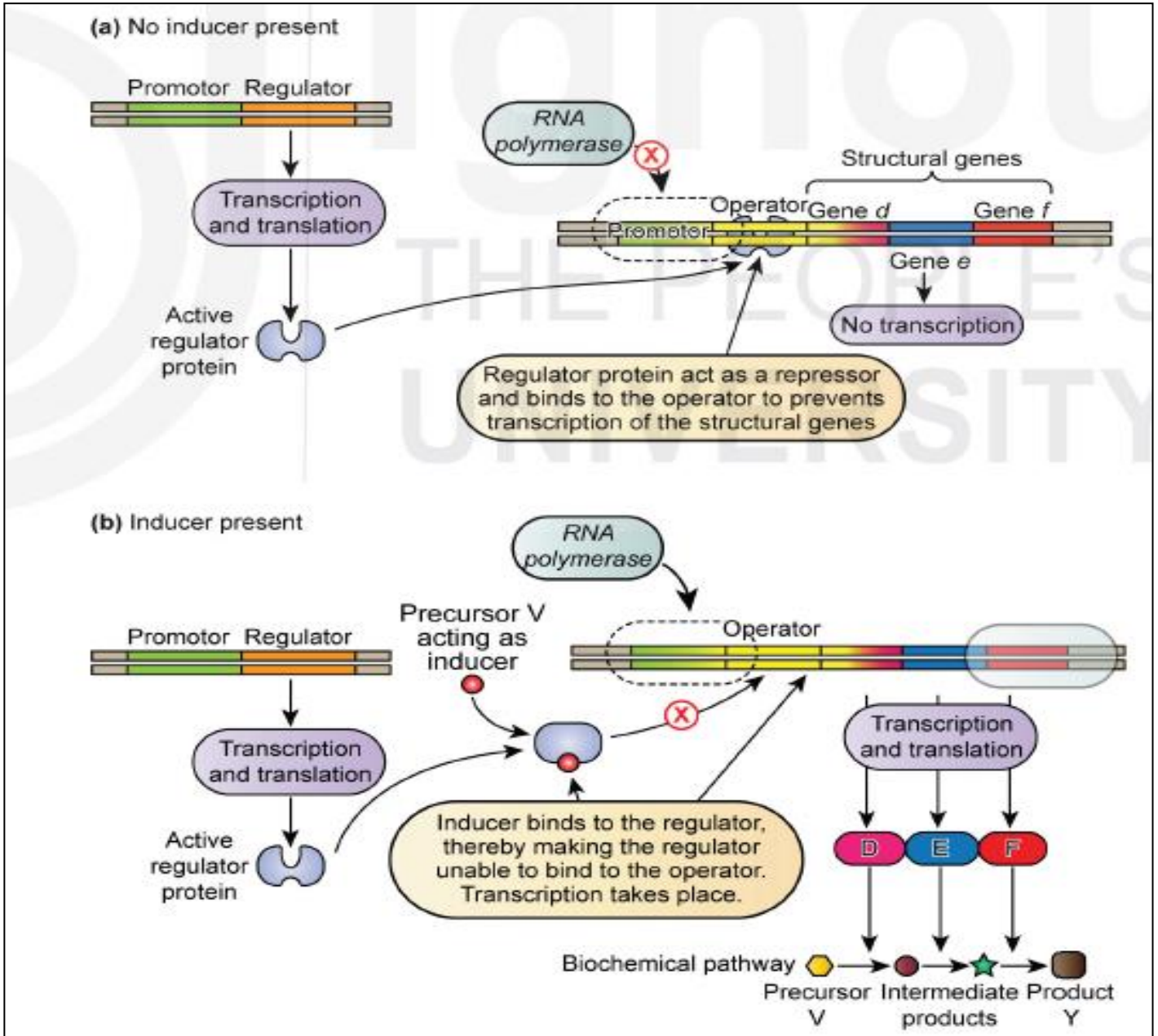


Fig. 14.5: Negative inducible operons. a) Structural genes *d*, *e*, and *f* become switched off when the inducer is absent; b) Genes *d*, *e*, and f and switched on when the inducer is present.

Negative repressible operons: In such type of repressible system, the activity of gene or genes is normally kept suppressed or turned off. This is because the regulator protein in this type of operon is also a repressor. But it is synthesized in an inactive form and cannot bind to the operator by itself. Since there is no repressor bound to the operator, *RNA polymerase* readily binds to the promoter and transcription of structural genes takes place. The activity of these genes can be turned on only if the repressor is activated. This job is done by a small molecule that binds to the repressor and makes it capable of binding to the operator and is called **corepressor**. As illustrated in Figure 14.6 a, the product U of the metabolic reaction is a corepressor. So, the negative repressible operon regulates the gene expression depending on the availability of the product U of the metabolic reaction.

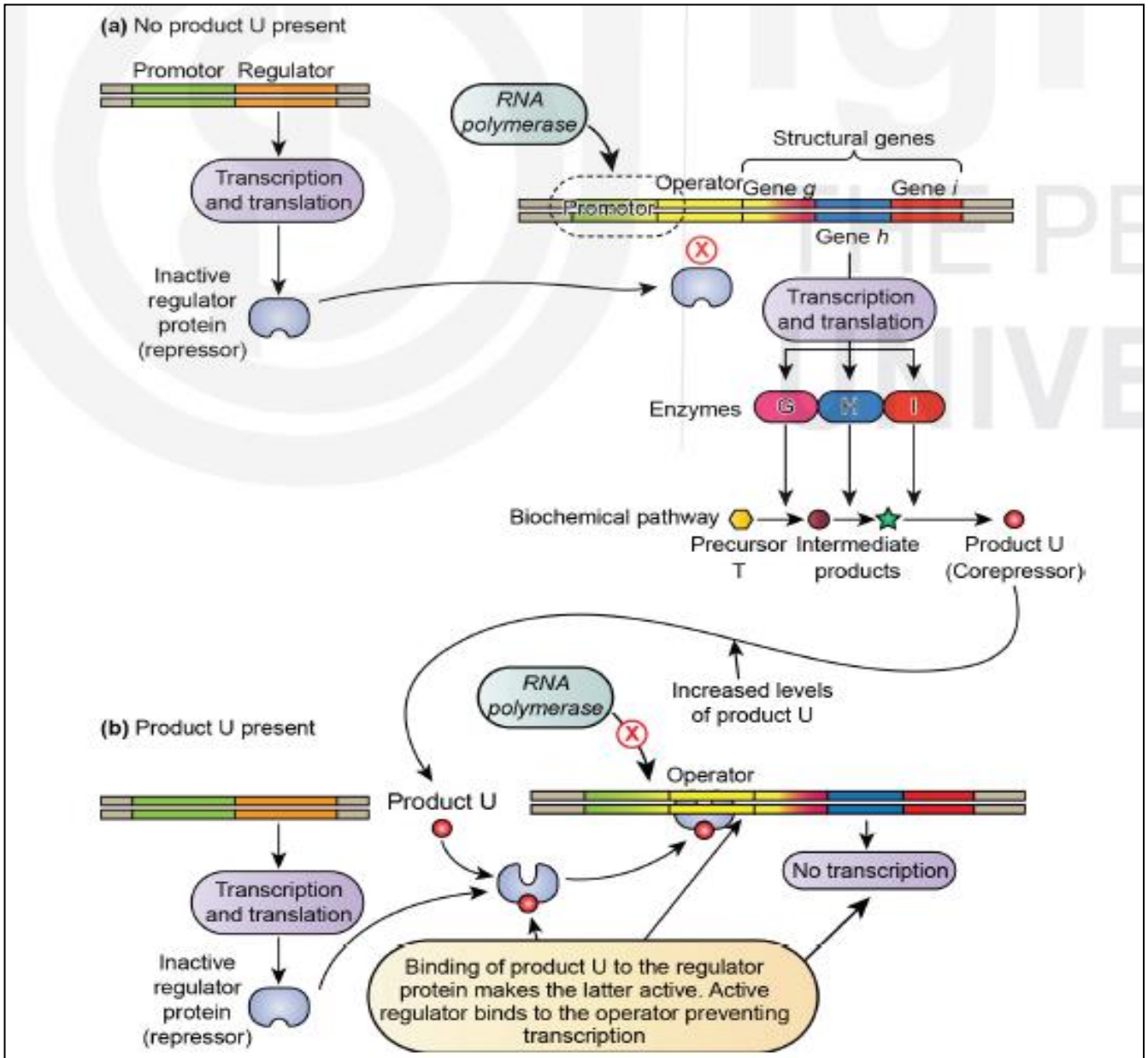


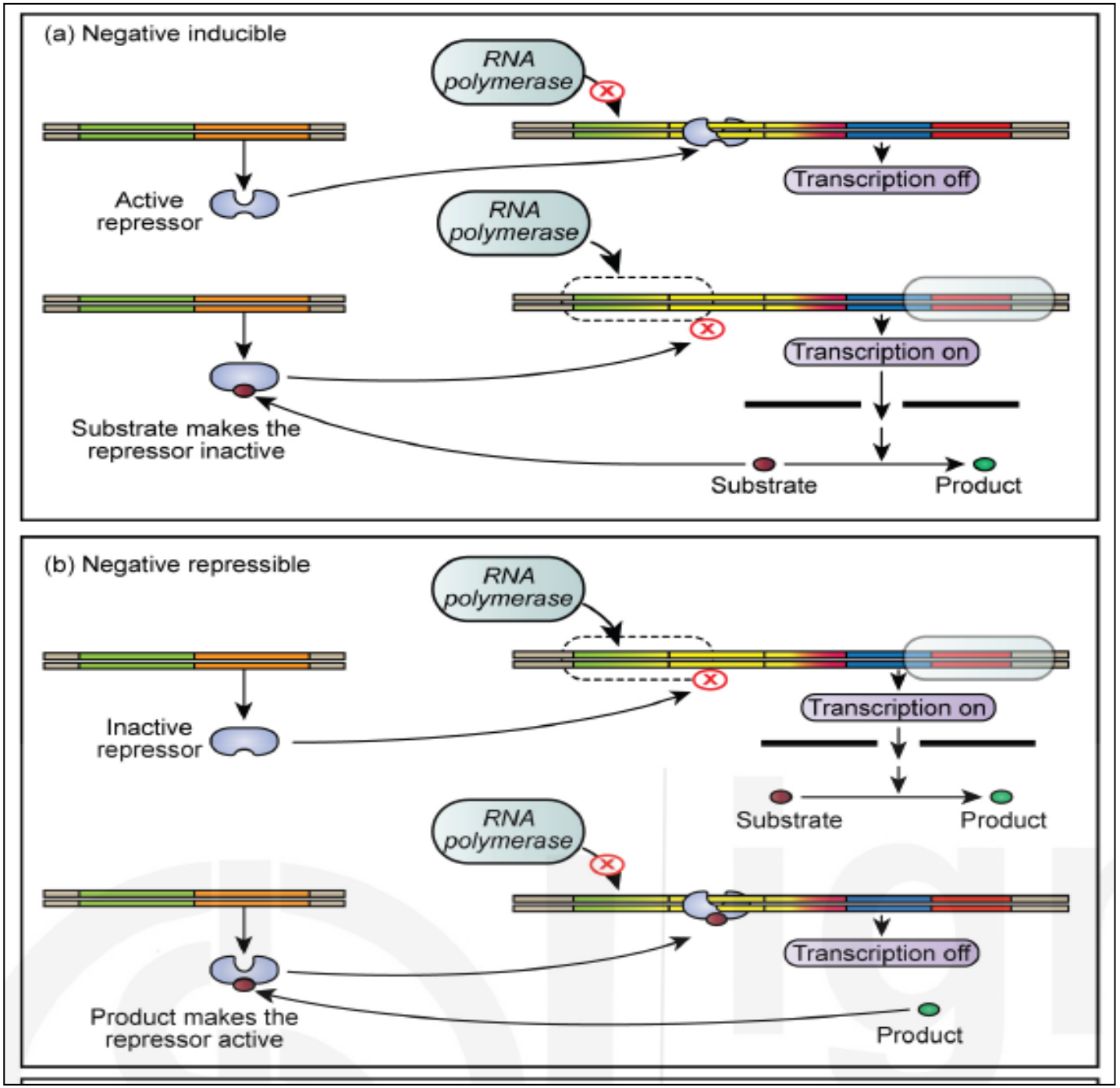
Fig.14.6: Negative repressible operons. a) Structural genes *g*, *h*, and *i* become switched on when the metabolic product U is absent; b) Structural genes *g*, *h*, and *i* become switched off when the product U is present.

As mentioned above, both the inducible and the repressible systems represent forms of negative control, where the regulatory protein functions as a repressor. You will now study the positive control in which the regulator protein itself stimulates transcription.

Positive control: In this mode, a regulatory protein is an activator with a positive control. The activator binds to the DNA usually at a site other than the operator and stimulates transcription. Positive control can be inducible or repressible.

Inducible positive operon: In this system transcription normally remains turned off because the regulator protein (an activator) is produced in an inactive form. Transcription can take place only when the regulatory protein is activated through attachment of an inducer. It is, therefore, logical to assume that the inducer should necessarily be the precursor of the reaction which is controlled by the operon so that the necessary enzymes would be produced only when the substrate for their reaction was present (Fig.14.7c).

Repressible positive operon: In this operon, the regulatory protein is produced in an active form that readily binds to DNA to switch on the transcription. Transcription can be stopped/ inhibited only when a substance becomes attached to the activator and makes it unable to bind to the DNA. It means that the product of the reaction controlled by the operon would logically be the repressing substance, as shown in the Fig.14.7d. It is economical for the cell to prevent the transcription of genes that allow the synthesis of the product when these products are already available in excess. The characteristics of positive and negative control in inducible and repressible operons have been summarized in Fig.14.7 (a, b).


(a) Negative inducible
RNA polymerase
Transcription off
Active repressor
RNA polymerase
Transcription on
Substrate makes the repressor inactive
Substrate
Product
(b) Negative repressible
RNA polymerase
Transcription on
Inactive repressor
Substrate
Product
RNA polymerase
Transcription off
Product makes the repressor active
Product

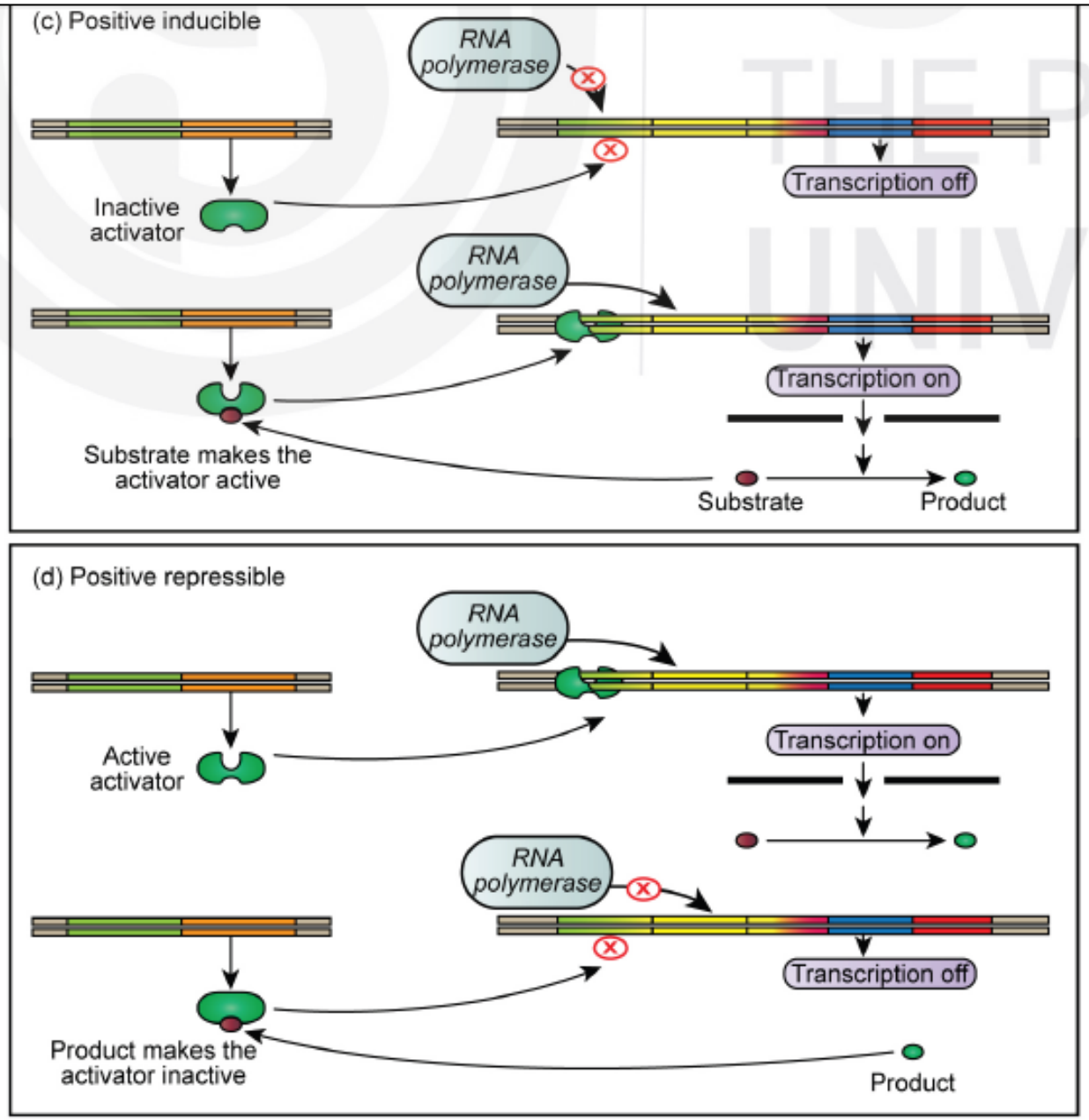


Fig. 14.7: Summary of the characteristics of negative and positive controls in inducible and repressible systems.

Model Systems for Gene Regulation in Prokaryotes

Lac operon and *trp* operon are model systems for understanding of gene regulation in prokaryotic organisms.

The *lac* operon of *E. coli*: An example of a negative inducible operon

The operon model proposed by Jacob and Monod proved to be a basic model of gene regulation with far-reaching implications. A key feature of the operon model is the idea that genes with metabolically related functions are clustered together so their transcription can be regulated as a single unit. But how is this regulation accomplished? Jacob and Monod addressed this question by studying genetic control of lactose metabolism in *E. coli*.

Lactose metabolism and regulation of *lac* operon: Lactose is a major carbohydrate found in milk. It is metabolized by *E. coli* residing in mammalian gut. Lactose does not easily diffuse across the *E. coli* cell membrane and be actively transported into the cell by the protein *permease*. Since lactose cannot be used directly as an energy source. Transport of lactose into the cell is an active process and is mediated by the enzyme *permease*. *E. coli* first breaks it into glucose and galactose. This reaction is catalyzed by the enzyme *β −galactosidase*. In addition, this enzyme also converts lactose into allolactose, a compound that plays important role in regulating lactose metabolism itself. A third enzyme, *thiogalactoside transacetylase* is also produced by *lac* operon, but its function in lactose metabolism is not yet fully understood (Fig. 14.8).

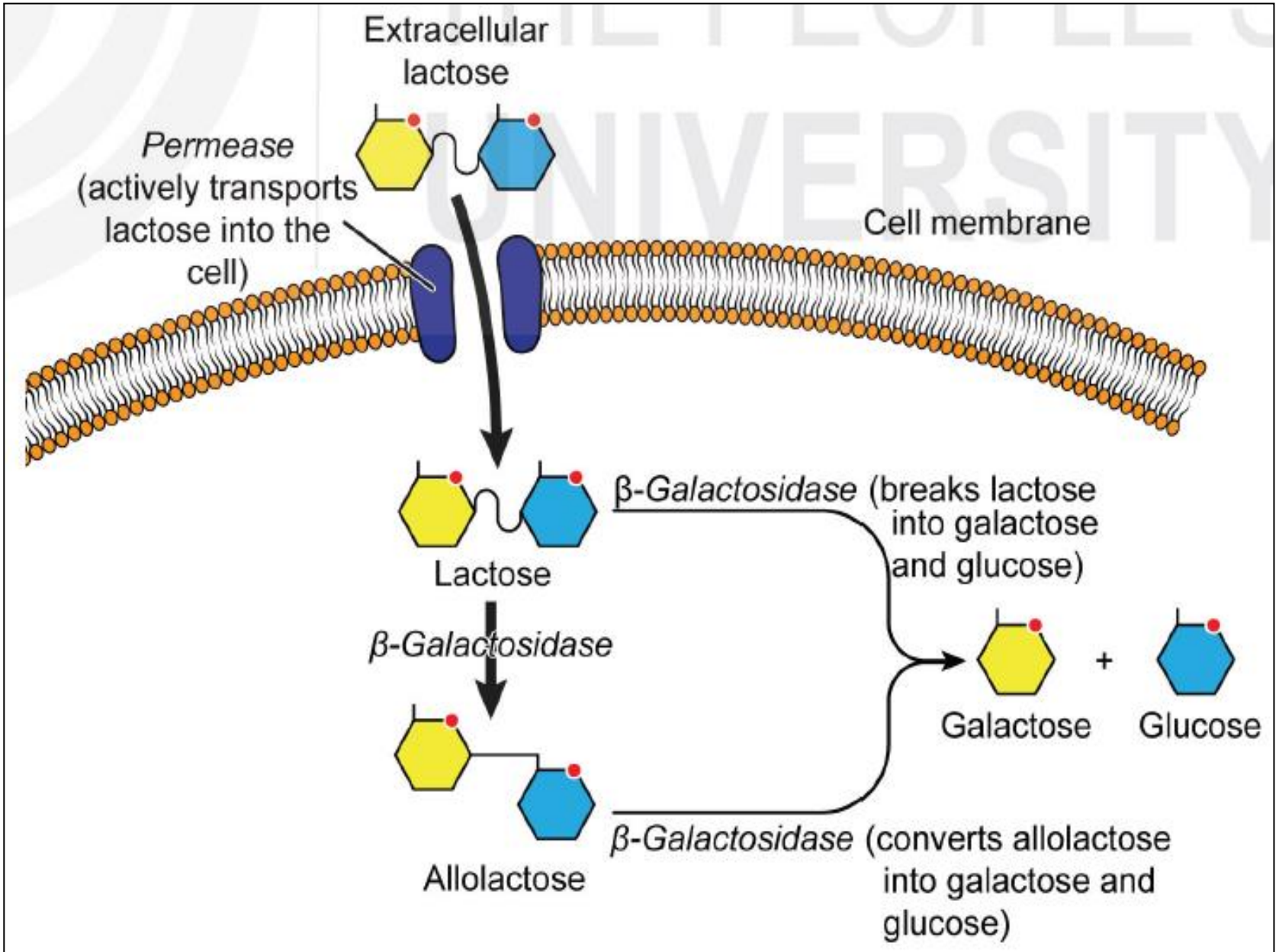


Fig.14.8: Enzymes involved in lactose metabolism in *E. coli* cells.

The *lac* operon is an example of an inducible operon. The adjacently placed structural genes in *lac* operon in *E. coli*, along with a common promoter *lac P* encode the enzymes *β –galactosidase*, *permease*, and *transacetylase* respectively (Fig. 14.9). *β –galactosidase* is encoded by *lac Z* gene, *permease* by *lacY* gene and *thiogalactoside transacetylase* by the *lac A* gene. As you have already studied about the negative inducible operon earlier, lactose seems to be inducer here and allolactose is responsible for induction. There are two situations in which lactose is metabolized by *lac* operon.

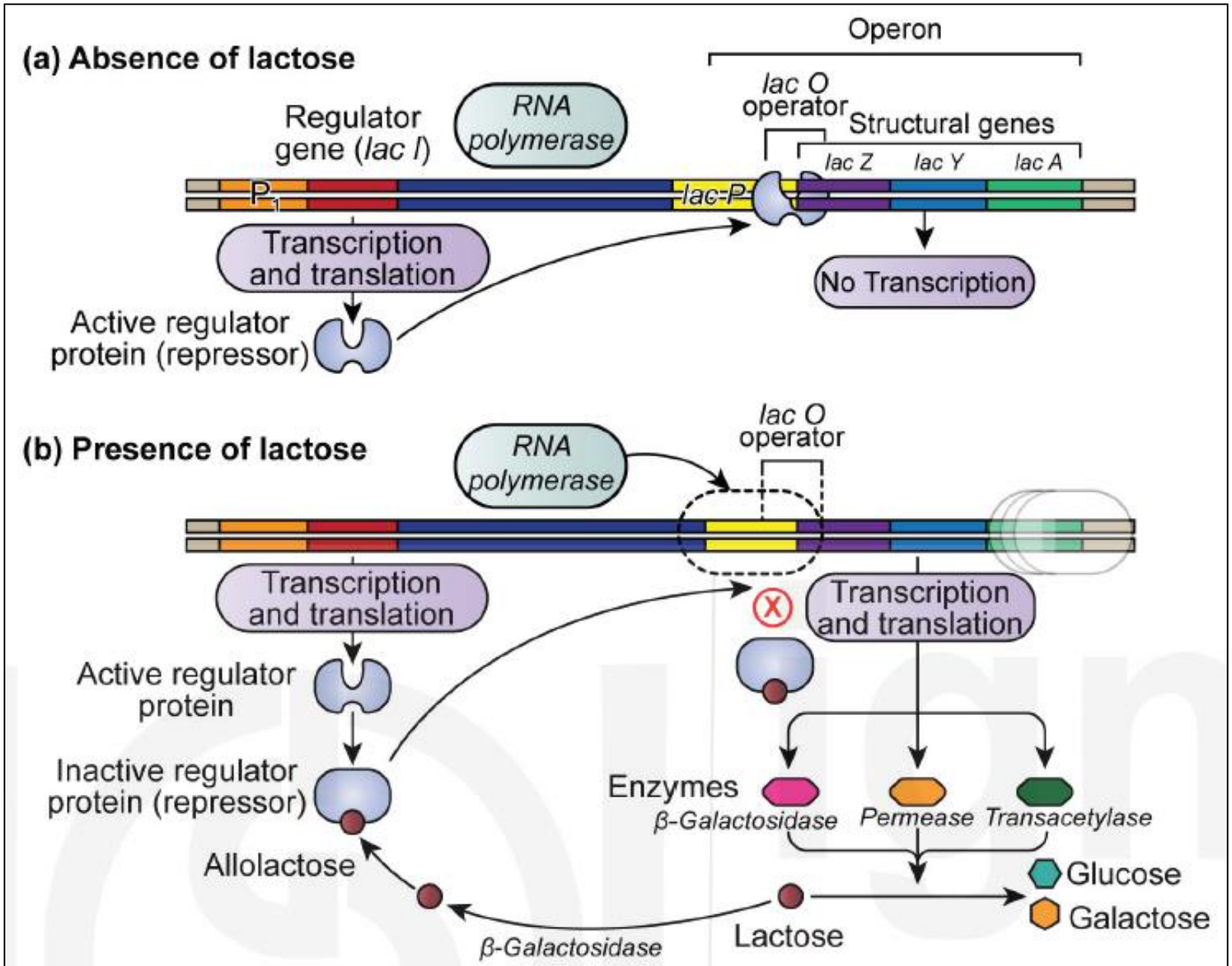


Fig. 14.9: The *lac* operon regulates lactose metabolism.

a) In the absence of lactose, transcription is inhibited because the regulator protein (repressor) binds to the operator which blocks the binding of *RNA polymerase*.

b) When lactose is present, the structural genes are transcribed and translated. This is because some of the lactose gets converted into allolactose which then binds to the regulator protein, making the protein inactive. Since the regulator protein cannot bind to the operator the structural genes get freely transcribed and translated.

The *trp* operon model of *E. coli*: An example of a negative repressible operon

Although the operon concept was formulated by the lac operon of *E. coli*, many other bacterial regulatory systems are now known to follow the same general pattern. As discussed earlier, operons coding for the enzymes involved in catabolic pathway resemble the *lac* operon in being inducible. Here the operons are turned on by a specific allosteric effector, which is usually substrate itself for the pathway involved. As against the catabolic pathway, operons that regulate enzymes involved in anabolic (biosynthetic) pathways resemble tryptophan (*trp*) operon in being repressible operons. They can turn off allosterically by an effector that is itself the product of the reaction. The tryptophan (*trp*) operon contains genes coding for enzymes involved in tryptophan biosynthesis, along with DNA sequences that regulate the production of these enzymes. The effector molecule in this case is the end-

product of the biosynthetic pathway, the amino acid tryptophan. The *trp* operon contains five structural genes *trp*E, *trp*D, *trp*C, *trpB* and *trpA* that produce the component of three enzymes (two of the enzymes consists of two polypeptide chains). These enzymes convert chorismate (the key intermediate of the tryptophan biosynthesis) into tryptophan (Fig. 14.10).

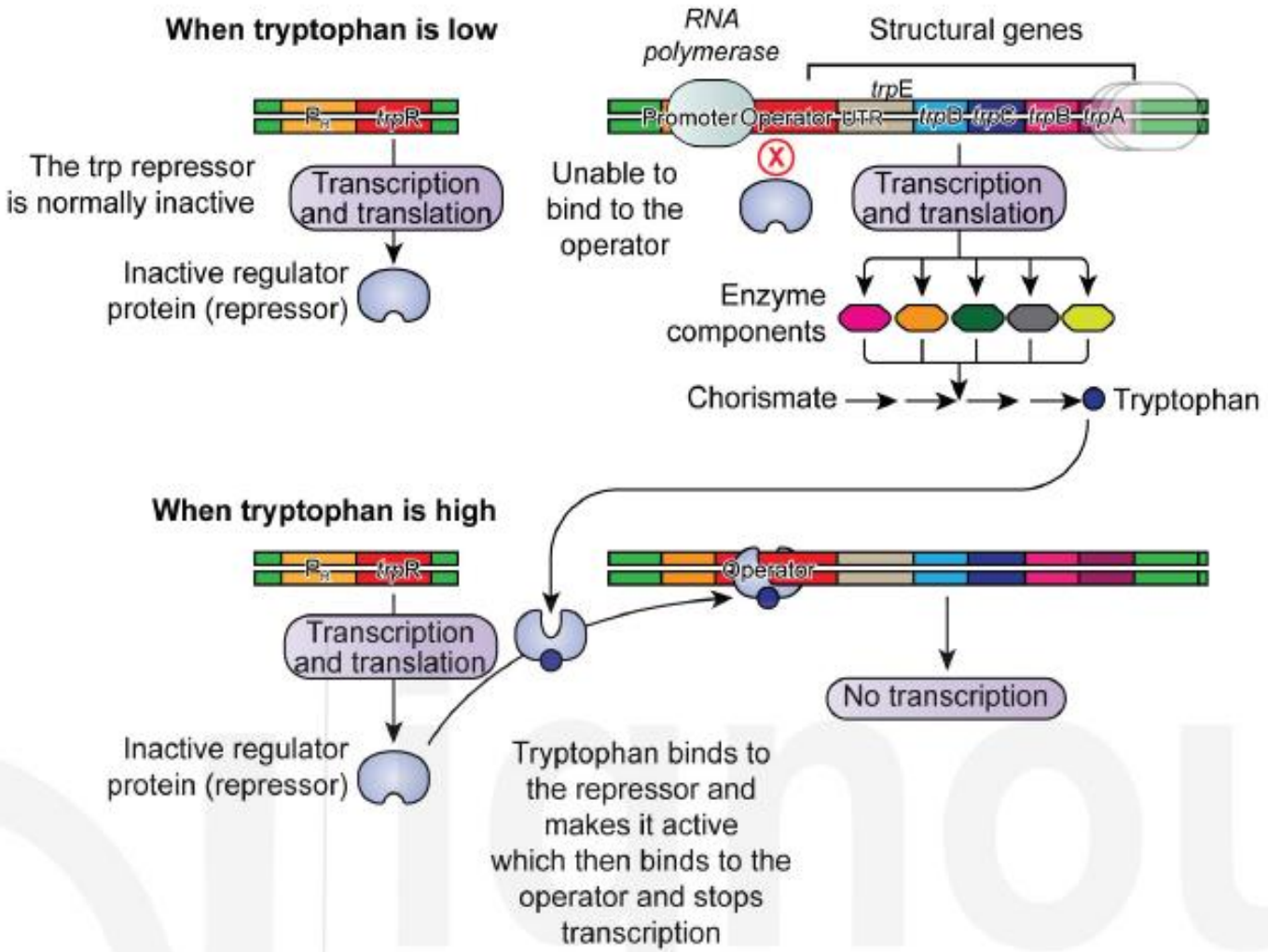


Fig. 14.10: The *trp* operon in *E. coli* controlling the biosynthesis of the amino

The first structural gene, *trp e*, contains a long 5´untranslated region (5´UTR) that is transcribed but does not encode any of these enzymes. Instead, this 5´UTR plays an important role in other regulatory mechanisms. The *trp* promoter is located upstream of the 5´ UTR.Tryptophan is synthesized by this operon in two situations.

a) **Low tryptophan levels**: *RNA polymerase* binds to the promoter and transcribes the five structural genes into a single mRNA, which is then translated into enzymes that convert chorismate into tryptophan.

b) **High tryptophan levels:** Some distance from the *trp* operon is a regulator gene, *trpR* which encodes a repressor that alone cannot bind DNA. Like the *lac* repressor, the tryptophan repressor has two binding sites, one that binds to DNA at the operator site and another which binds the tryptophan (the activator). Binding with tryptophan results in a conformational change in the repressor. This change makes the repressor capable of binding with DNA at the operator site. This site overlaps the promoter. When the operator is occupied by the tryptophan repressor, *RNA polymerase* cannot bind to the promoter and the structural genes cannot be transcribed.

Thus, when the cellular levels of tryptophan are low, transcription of *trp* operon takes place and more tryptophan is synthesized but when the cellular levels of tryptophan are high, transcription of the *trp* operon is inhibited and the synthesis of more tryptophan does not take place (Fig.14.10).