13

Molecular Biology

Doctor 2019 | Medicine | JU

Sheet Slides

Done By

Leen M. Hadid/Hayat Suwi

Contributed in the Scientific Correction

Ayat B Nabil

Contributed in the Grammatical Correction

Ayat B Nabil

Doctor

Mamoun Ahram

(specifically, we will talk about the lac operon)- lac stand for lactose -

-We will talk about **Regulation of transcription** or in other words **controlling gene activity** or in other words '**how genes are expressed**' (Note the term, so gene expression).

When we say gene expression (بالعربي: التعبير الجيني), it means that we are talking about if genes are active or not that is at the transcriptional level are they transcribed or not, so if they are **transcribed** it means that genes are expressed (active), and if they are **not transcribed**, it means that genes are not active or not expressed.

And again, the reason why we start with prokaryotic systems is because they are easy to understand.

→ Understanding gene expression sort of started in the 1950s.

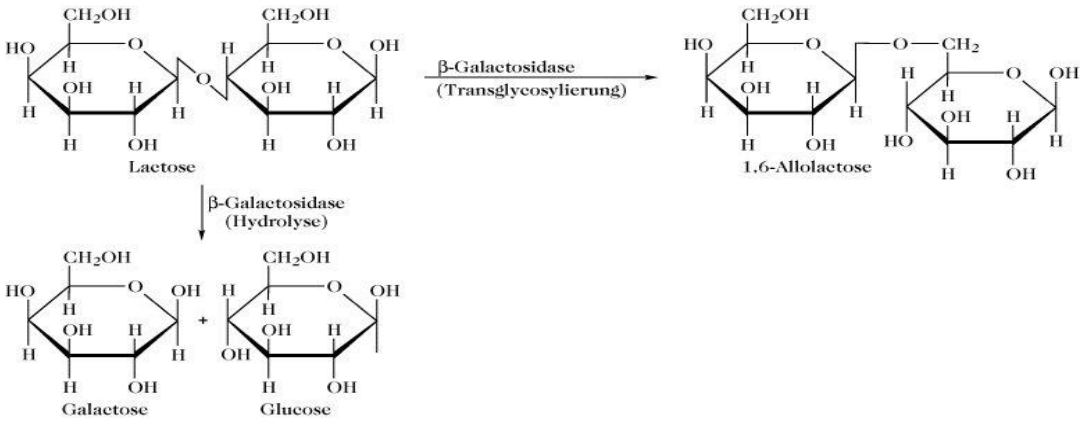
In the 1950s, pioneering experiments were carried out by French scientists François Jacob and Jacques Monod who studied regulation of gene transcription in E. coli by analyzing the expression of enzymes involved in the metabolism of lactose, and they got the Nobel Prize for their work. (they did these experiments showing how bacteria can control metabolism of lactose).

Lactose (a sugar) is a disaccharide which means that it is made of two sugars and the sugars are glucose connected to galactose.

Now lactose is metabolized in two ways:

1- cleaved by an enzyme called **β- galactosidase** which hydrolysis lactose using water "hydro", it "lysis" that is it cleaves lactose (between the two sugar monomers), producing two sugar monomers (monosaccharides) -glucose and galactose- and this is the preferred way. → Hydrolyse
2- or the same enzyme **β- galactosidase** can change the conformation (the structure) of lactose into something known as 1,6-allolactose which is an isomer to the lactose, it is clear on the figure. So, instesd of having the connection (the bond) between the two sugars between the two carbons shown in the figure (in Lactose), the connection is now difference between carbons (in 1,6-Allolactose). → Transglycosylierung

But this is a minor form, the majority of lactose would be hydrolyzed (cleaved) into two monosaccharides, galactose and glucose.



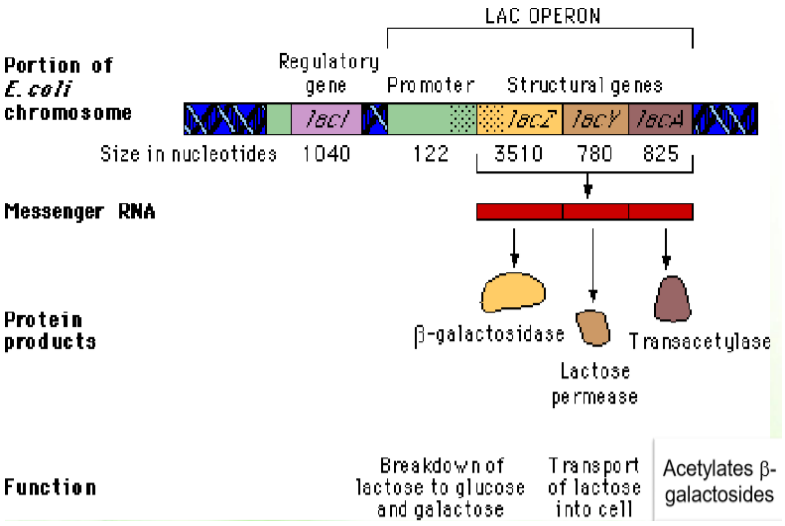
The lac operon:

Remember when we said an **operon** is basically a genetic unit (a cistron) that is **transcribed into one single messenger RNA** (shown in red in the figure) but it is **translated producing different proteins:** 1- each one of them has its own function and 2- they participate in the same process. So, these three proteins in the figure are involved in the metabolism of lactose. (please pay attention to the names of these proteins and its function from the figure.).

Components of lac operon

Notes about the figure:

- The proteins that are produced from the lac operon are: β-galactosidase, lactose permease and transacetylase.
- β-galactosides are sugars just like lactose, when we say acetylate it means that what it does is that it adds an acetyl group to molecules like β-galactosides.



- We know very well what is the functions of permease and galactosidase, but we don't know much about the function of the transacetylase. So we will focus on these two particularly the galactosidase enzyme.

- This lac operon has its own promoter region (which is in green), this is the RNA polymerase binding site and you have transcription of this operon into a

single unit of messenger RNA and each part of the messenger RNA would produce a different protein with a different function.

THE LAC OPERON COMPOSED OF: THE PROMOTER (this is the RNA binding site) & THE STRUCTURAL GENES.

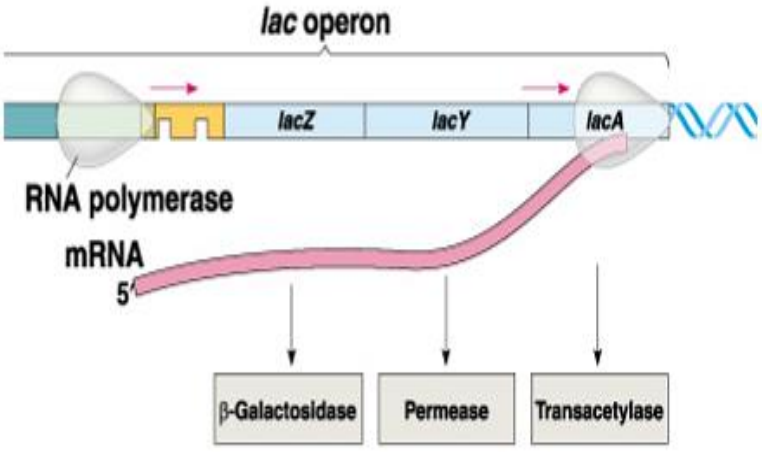
We also have another gene that is located far away from the lac operon (on the same chromosome of course because as you know bacteria have one single chromosome as their genome).

the name of this gene is lac I >> I stand for inhibitor. ---- The lac I has its own promoter.

The product of this gene regulates how the lac operon is expressed (that is if it is expressed or not).

→ Now remember something that bacteria do have operons just like the lac operons but at the same time they do have genes that produce one single protein or one single polypeptide at least. So, they do have monocistronic genes and they also have polycistronic genes as well.

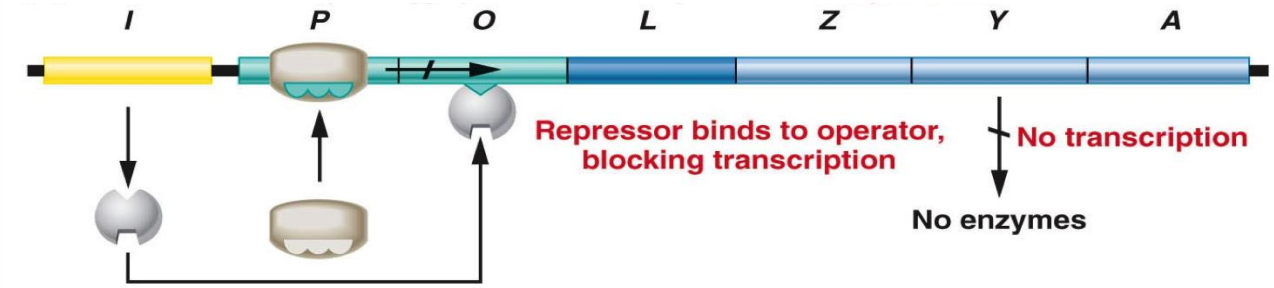
What is an operon?



An operon is a cluster of genes transcribed from one promoter (yellow in the picture above) producing a polycistronic mRNA that is used to make proteins that are totally different in structure and function, but they participate in the same pathway (purpose) or the same mechanism (In this case, it's the metabolism of lactose).

The operator

- Another part of the lac operon which is the operator – a regulatory region (an element) that exist between the promotor and the transcription start site- this region is the binding site of the protein product of the lac I gene.
- The lac I gene produces a protein and this protein is known as the lac repressor then this protein binds in the promoter preventing the RNA polymerase from starting the transcription so it blocks the transcription – So, if the lac repressor is bound to the operator the RNA polymerase cannot start transcription, it is an inhibitor for the lac operon.

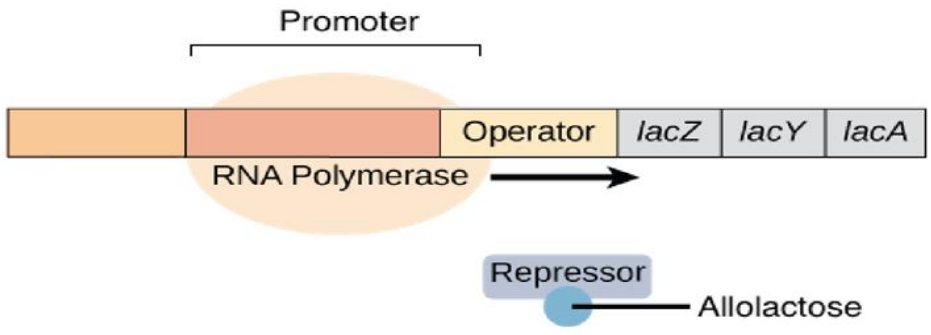


To sum up:

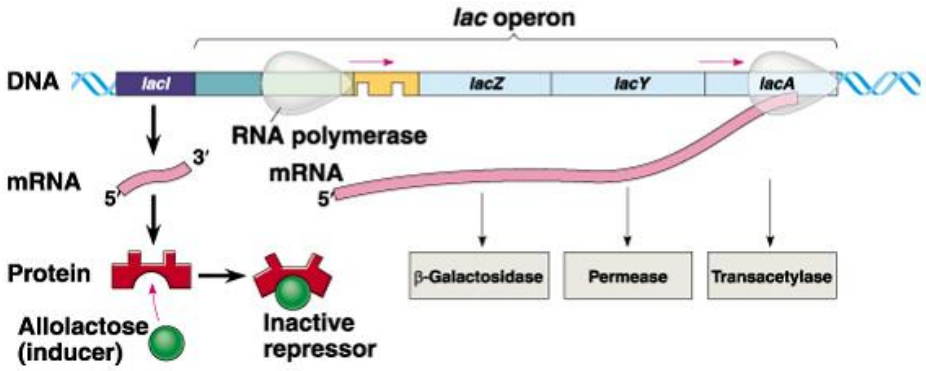
- The promoter region includes the operator region, which is a binding site of a protein called the lac repressor.
- The lac repressor blocks transcription by preventing the RNA polymerase from binding to the promoter.

Now in order for the RNA polymerase to start transcription of the lac operon the repressor must be released, and it is released when the molecule allolactose binds to it. So when there is lactose some of it is converted to allolactose and this allolactose would bind to the repressor releasing it from the operator and the RNA polymerase can start transcription and it make sense because you don't want transcription of the lac operon to start unless there is lactose otherwise it would be a waste of time and resources.

Glucose present, lactose present:



Regulation by lactose (positive)



(b) Lactose present, repressor inactive, operon on

When ever lactose is present the repressor is inactive, so the lac I would make the repressor (Note in the figure below: the mRNA of the lac I gene producing the repressor protein) and if there is allolactose (if there is lactose some of it would be converted to allolactose), it would bind to the repressor preventing its binding to the operator so the RNA polymerase can start transcription producing the polycistronic mRNA producing the 3 different proteins and metabolism of lactose would start whereby galactosidase would cleave lactose into two monosaccharide and these two monosaccharide can be metabolized -THIS IS KNOWN AS POSITIVE REGULATION OF TRANSCRIPTION- because the present of lactose positively regulate transcription meaning that it activates transcription of the gene.

→ Lactose binds to the repressor, thereby preventing it from binding to the operator DNA.

→ This is Known as positive regulation.

Cis vs. trans regulatory elements

- Now there is a few terms you need to know and understand, there is something in molecular biology about gene expression and transcription is that you have what is known as cis-acting elements -whenever I say elements we are talking about DNA or RNA sequences like the iron response element binding protein, so it is an element that exists in RNA and it controls (it regulates) transcription somehow- .
    - Cis means same level.
    - Cis-acting element is basically a region that exists in DNA for example, and this region controls the expression of this particular gene.

- If we take this region and we move it (change its location), if we put for example the operator somewhere else it would not be active (it would not be able to regulate transcription), because in this case the operator would bind to repressor but it doesn't block the RNA polymerase.
- ✓ it is simply a DNA regulatory sequences like the operator 'called **cis-acting elements**' because they affect the expression of only genes linked on the same DNA molecule or close-by.

→ Mention other examples of cis-acting elements from the previous lecture on transcription.

Hint: enhancers for example these are cis-acting elements, except that there is a twist, there is something with enhancers that is if you take these enhancers and you change their location they can still be functional, but if you put them really far away from where the gene is they wouldn't be functional anymore. If you take them and put them on a different chromosome they would not be functional anymore.

So, think about other cis-acting elements.

- The second term is trans-acting factors,- when we say factors we are talking about products of transcription so we talking about RNA molecules or protein molecules so we not talking about sequences we are rather talking about products of genes like again RNA or proteins-. because they can affect the expression of genes located on other chromosomes within the cell. They are produced from trans-acting elements --- trans-acting elements are DNA sequences that encode trans acting factors (often proteins such as transcription factors).
  - Trans means different level.
  - Why we called it trans or different level? Because if you take a gene that produces a certain protein and you put it on different region of the chromosome that's far away from where the gene that regulate it is or put it on different chromosome it can still be functional. Why? Because it would produce a protein and this protein can bind on a sequence of another chromosome, so it's like floating it, it swims in the nucleus and it can bind anywhere on DNA, so it's known as transacting elements.
  - An example of the transacting elements is the repressor. That is if you take the lac I gene and you put it somewhere else far away from the lac operon it can still be transcribed if it has its own promoter of course, and it can still

produce a protein (a repressor), and this protein would bind to the operator.

- ✓ Proteins like the repressor are called **transacting factors** because they can affect the expression of genes located on other chromosomes within the cell. They are produced from **trans-acting elements** --- trans-acting elements are DNA sequences that encode trans acting factors (often proteins such as transcription factors).

→Try to think of other types of trans-acting elements.

Effect of mutations

So why did Jacob and Monod got the Nobel prize, because what they did is that they created a number of mutations (so they induced mutations themselves) and in different region and try to understand what these mutations would do and based on these mutations they were able to understand the lac operon and how it is regulated.

Some mutation take place in our genome can result (and they did it):

A- Some mutations result in **constitutive expression** (a mutation that results in the gene being always on and it's not regulated any more so it doesn't get turned off). Examples:

1- mutation in the lac I which produces a defective lac repressor.

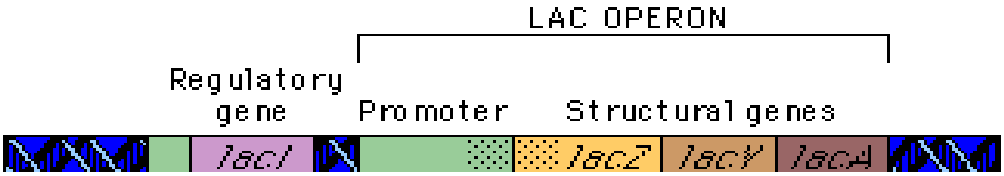
Defective lac repressor can't bind to the operator, so nothing prevents the binding of RNA polymerase to the promoter. The gene expression is always ON.

2- mutation in the operator:

- ➢ It changes the DNA sequence of the operator, so that the lac I product that is the repressor cannot bind to the operator, this means that the RNA polymerase can bind freely to the promoter and it can transcribe the lac operon all the time.

the repressor is functional, but it can't bind to the operator due to mutation in the operator, so nothing prevents the binding of RNA polymerase to the promoter. The gene expression is always ON.

→So think of other mutations that can result in constitutive expression.



B- Other mutations that can cause **non-inducible or repressed expression** (meaning that the gene is always off. In other words, whether there is lactose or not the gene is always off, it cannot be induced). Examples:

1- a mutation in the promoter:

RNA polymerase can't bind even if the operator is free of lac repressor. Transcription is always OFF.

2- a mutation in lacZ :

non-functional β-galactosidase is produced, so it won't convert Lactose to Allolactose. Transcription is always OFF.

3- a mutation in lacI gene:

when a mutation happens in lacI it may produce a defective repressor (as I mentioned earlier in A-1 constitutive expression). Or it can produce a repressor with altered Allolactose-binding region, so Allolactose can't bind anymore to the repressor, so this mutation would mean that the lac repressor would always be bound to the operator and will still bound so prevent RNA polymerase from transcription, so that would make the gene non-induced or repressed, even if there is lactose inside the cells it cannot bind to the repressor and the repressor would still be bound preventing the RNA polymerase from transcribing the gene. Transcription is always OFF.

WE CAN APLIED THESE CONSEPTS TO OTHER GENES WHETHER IN BACTRIA OR IN OUR CELLS.

Another level of regulation

It is mediated by cAMP

cAMP: a small molecule that is produce from ATP (adenosine triphosphate) which can converted to cAMP (it has one phosphate) and has a structure that looks like more cyclic so it's called cyclic AMP (cAMP), which binds to another protein known as catabolite activating protein (CAP).

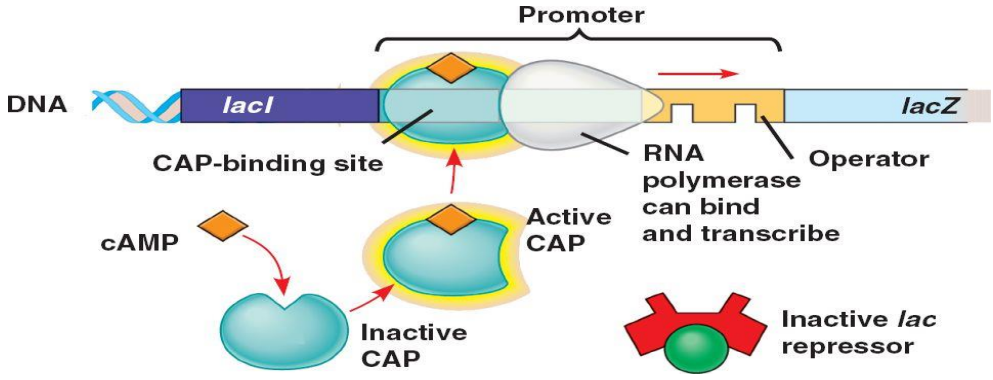
→ Another regulator is cAMP, which binds to a protein known as catabolite activating protein (CAP) and stimulates its binding to regulatory sequences upstream of the promoter.

. What happens is, cAMP binds to CAP and activates it. Then CAP binds to DNA region that is upstream of the promoter, so when it binds to this region it interacts with the RNA polymerase and it activates it, and this cap protein binds to this region of DNA when there is a cAMP.

→ CAP interacts with the RNA polymerase to facilitate the binding of polymerase to the promoter (P) and this cap protein binds to this region of DNA when there is a cAMP.

There is cAMP → bind to CAP → CAP becomes active → it binds upstream of the promoter → it interacts with the RNA polymerase → it stimulates (induces/ activates) the RNA polymerase → RNA polymerase is more efficient (more active), so it keeps on transcribing the lac operon very efficiently and very quickly producing a lot of mRNA from the lac operon, producing a lot of β-galactosidase, lactose permease and transacetylase.

VERY IMPORTANT NOTE: without CAP, RNA polymerase can still do transcription.



cAMP and CAP influence on lac operon transcription

How is cAMP produced? Why is it produced? In what circumstances?

Again, if you have CAP and CAP is not bound upstream of the promoter region, RNA polymerase (and if there is no repressor) can start transcription. But, if there is cAMP, cAMP would bind to CAP, CAP can then bind to DNA inducing the RNA polymerase producing a lot of polycistronic mRNA from the operon (lac operon),so you have a lot of transcription going on.

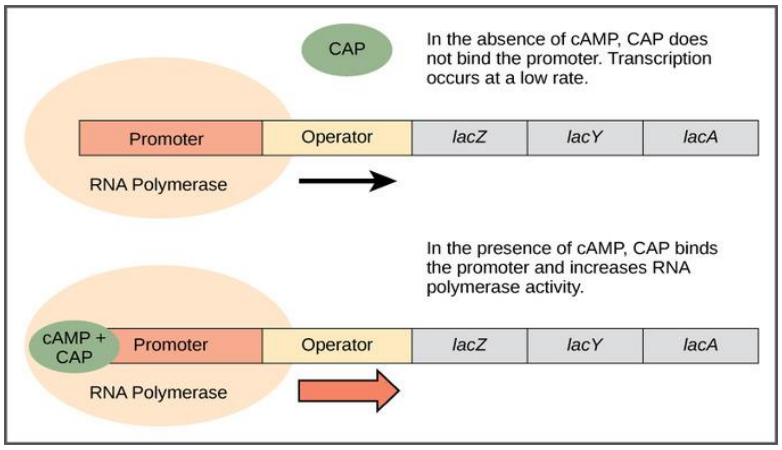
When do cells produce cAMP? And why?

It's all regulated by glucose, and glucose induces what is known as negative regulation that's the opposite of lactose. Meaning that if there is glucose it means that lac operon should not be active, and it makes sense , if bacteria have both lactose and glucose ,what do you think bacteria would prefer to metabolize? One glucose, why? Because first you have to hydrolyze lactose into two monosaccharides and these are glucose and galactose, if there is glucose already in the cells why do cells have to bother and produce β-galactosidase? or there is no need to produce β-galactosidase a lot.

*How dose glucose regulate transcription? -A SUMMARY-

-Glucose binds to and inhibits a protein and enzyme known as adenylyl cyclase.

-ATP is converted to cAMP by an enzyme known as adenylyl cyclase (or adenylate cyclase), this enzyme is regulated by glucose.



- If there is glucose , it would inhibit adenylyl cyclase meaning that there is no more conversion of ATP to cAMP, meaning that CAP would not be able to bind upstream of the promoter region and it would not be able to activate the RNA polymerase, meaning that you may have some transcription but it is not very efficient compered with if CAP bound to cAMP and bound to the DNA region.

Regulation by glucose

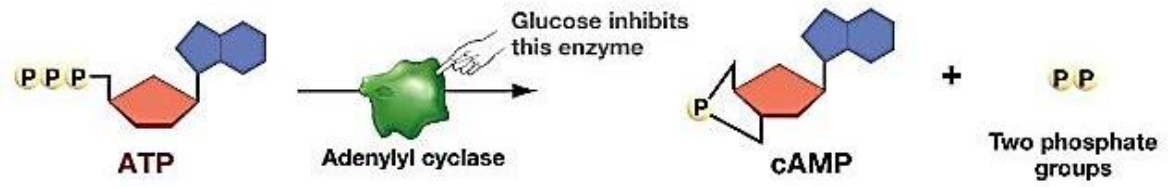
A Negative regulation (inactivates gene expression) that is opposite to lactose , if we have glucose it means lac operon should be not active.

→The ability of CAP to bind to the promoter is influenced by how much cAMP is in the cell is produced by adenylyl cyclase, which is inhibited by high level of glucose.

→Glucose is preferentially utilized by bacterial cells and it represses the lac operon even in the presence of the normal inducer (lactose).why? because you need to hydrolyze the lactose into two monosaccharide.

→This is known as negative regulation.

*Summary: if there is glucose ◊ it inhibits adenylyl cyclase ◊ no conversion of ATP to cAMP ◊ no activation of CAP ◊ no activate of RNA polymerase ◊ transcription is slow and not efficient.



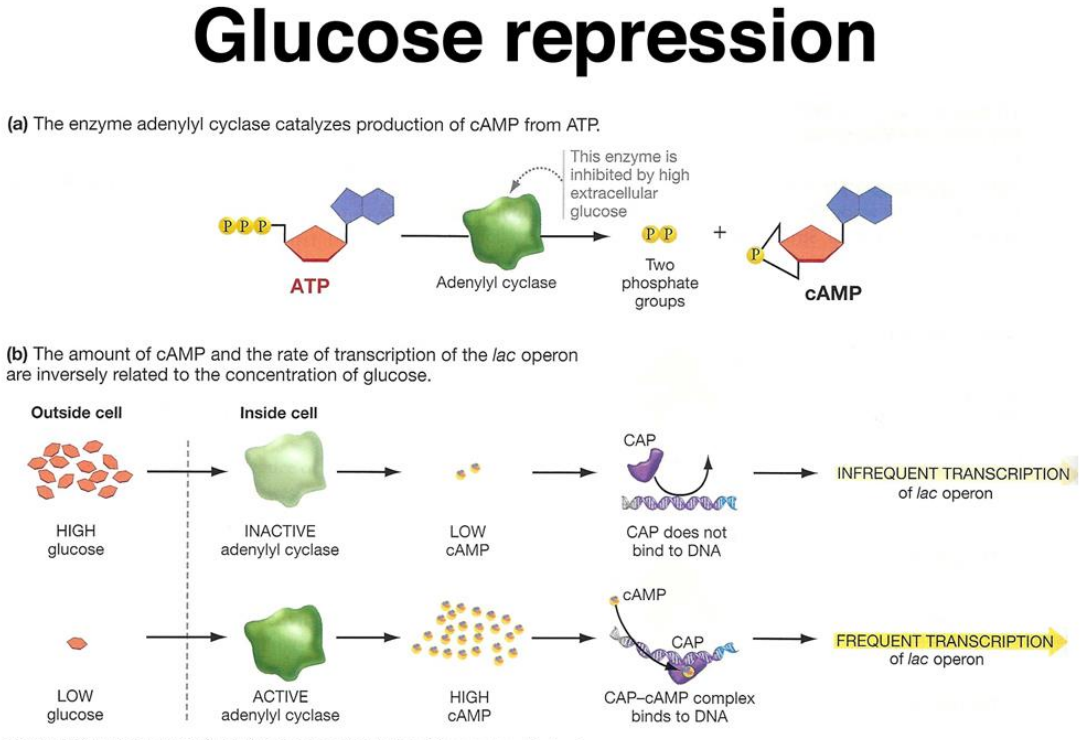


FIGURE 17.11 Cyclic AMP (cAMP) Is Synthesized When Glucose Levels Are Low.

High glucose ◊ bind to adenylyl cyclase◊ inactivate it◊ result have low level of cAMP◊ CAP can't bind to DNA◊ very little transcription occur

Low glucose◊ adenylyl cyclase is vary active◊ producing a lot of cAMP binds to CAP◊ CAP can then bind to the DNA (upstream of the RNA polymerase binding site that promoter) ◊ we have a lot of transcription.

Until now, we have taken 2 substances that regulate transcription of lac operon

Note that when we increase the concentration of lactose, transcription is activated (positive regulation), and when we increase the concentration of glucose, transcription is inactivated (negative regulation).

Since lactose and glucose regulate lac operon transcription, there are 4 scenarios of what will happen in E.coli bacteria: (and the figure bellow illustrate)

1- first case: NO Lactose, NO Glucose

NO transcription because the repressor is always bound to the operator and RNA polymerase can't do transcription.

2- second case: NO Lactose, YES Glucose:

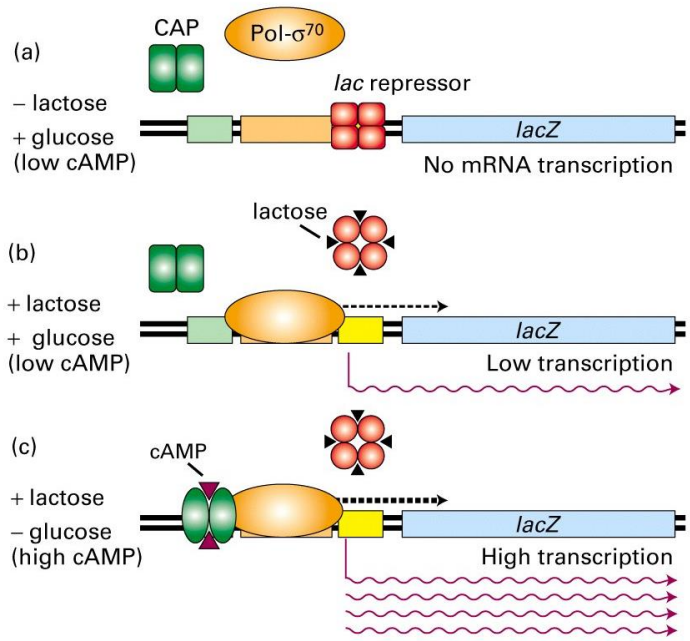
NO transcription because the repressor is always bound to the operator and RNA polymerase can't do transcription.(and of course CAP is not bound, but it does not matter because if there is no lactose the repressor would always be bound)

3- third case: YES Lactose, YES Glucose:

THERE IS TRANSCRIPTION, but it's slow.(there is preference for metabolizing glucose and you can have some transcription of the lac operon to metabolize lactose but not a whole lot. So there is lactose, lactose binds to their repressor, the repressor is released from the operator, the RNA polymerase can transcribe(it can move forward and it can transcribe) but it's not very efficient. Why? Because CAP is not bound. Why the CAP is not bound? Because there is glucose, there is little cAMP, so CAP is not able to bind upstream of the promoter activating the RNA polymerase.

4- fourth case: YES Lactose, NO Glucose:

THERE IS TRANSCRIPTION, and it's high.(adenylyl cyclase is very active producing a lot of cAMP, cAMP binds to CAP, RNA polymerase is very active, and because there is lactose the repressor is not bound to the operator and you have production of a lot of messenger RNA)



This concept of regulation we can ably it to human genes as well.

Take-home message

Gene expression is regulated by regulatory proteins (transcriptional factors for example like CAP, like the repressor) that would ultimately:

1-Guide the RNA polymerase (or other regulatory proteins) to the promoter.

2-Strengthen/stabilize RNA polymerase binding to the promoter.

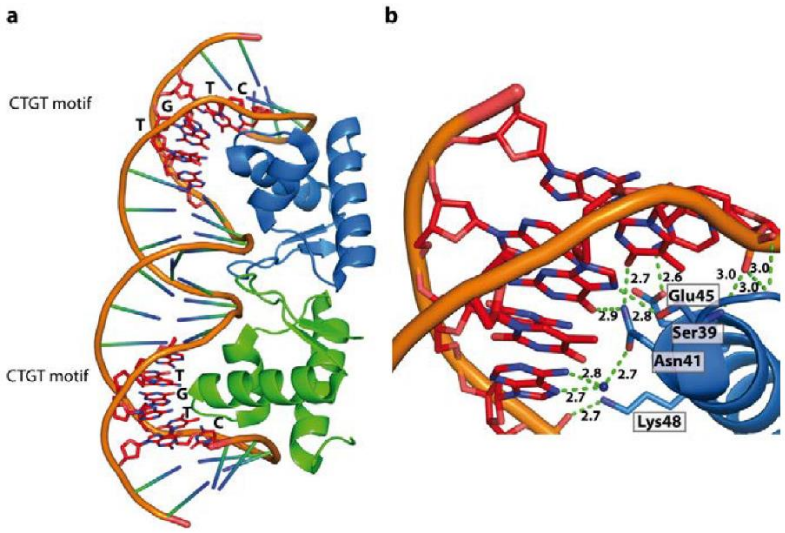
3- Activate the RNA polymerase.

4- Open up the DNA for the RNA polymerase (facilitating the transcribtion) OR the opposite of the above in case of repressors (that is repressors for example they can prevent the binding of RNA polymerase to the promoter, they can inactivate RNA polymerases, they can keep the DNA shot preventing the formation of the open promoter complex, so they keep it in the form of closed promoter complex so it's not open anymore).

- So that's how transcriptional regulatory proteins work, and this can be applied to eukaryotic systems as well.

All of the above effects are mediated via modulating non-covalent interactions between the amino acids of proteins (regulatory proteins) and specific sequences of DNA.

How do proteins recognize/interact with DNA sequences specifically?



Here in the figure you have a DNA and this just to show you how proteins would recognize DNA sequences specifically. That is in other words, why is it that the CAP protein cannot bind somewhere else on DNA nonspecifically? Why is it that the repressor would bind to the operator specifically?

The thing is the operator or the CAP binding site they have certain DNA sequences (certain order of nitrogenous bases of nucleotides like for example in the figure CTGT, so it's a certain sequence.) In blue and green (the two structures are the same, same order of amino acids within the protein interacting on both sides with the same sequence CTGT – the protein having the same structure twice so it's sort of like duplicated) is a regulatory protein that it does interact specifically with these bases, the interaction occurs between a specific sequence or order of amino acids of the protein non-covalently with the bases in the specific order.

REMEMBER and NOTICE:

When we talked about the major groove and minor groove. Remember that proteins preferably like to interact with bases of DNA in the major groove, because they can insert themselves within the major groove and they can become close to the bases of DNA in the major groove.

-- الله معكم –

(يعلم ما بين أيديهم وما خلفهم ولا يحيطون بشيء من علمه إلا بما شاء) {البقرة : 255}

كل الملاجئ دون الله كاذبة❤