Lesson 09: Biochemical genetics and the regulation of gene expression

The control of gene expression is essential for the balanced maintenance of cell growth. This control enables the cell to adjust its synthesis to environmental conditions. Genes are regulated at different stages of expression (synthesis of the product of a gene, RNA or proteins), enabling them to be activated or repressed.

1. The proteins of regulation

Regulatory proteins regulate the rate of transcription by binding to DNA at specific sequences near the promoters. They can :

- authorise the initiation of transcription, "turning on" the gene. They are then called inducers (activators) and the regulation is positive (there is transcriptional induction).

- or prohibit the initiation of transcription, "switching off" the gene. They are then called repressors and the regulation is negative (there is transcriptional repression).

Proteins can be :

- active (binding directly to DNA) and can therefore be inactivated by binding to a ligand (L).

- inactive (not binding to DNA) and can therefore be activated by binding to a ligand (L). It is the ligand which, by modifying the activity of regulatory proteins, acts as a "signal". It is therefore the presence of ligand that indicates the level of need or lack of need for the expression of certain genes.

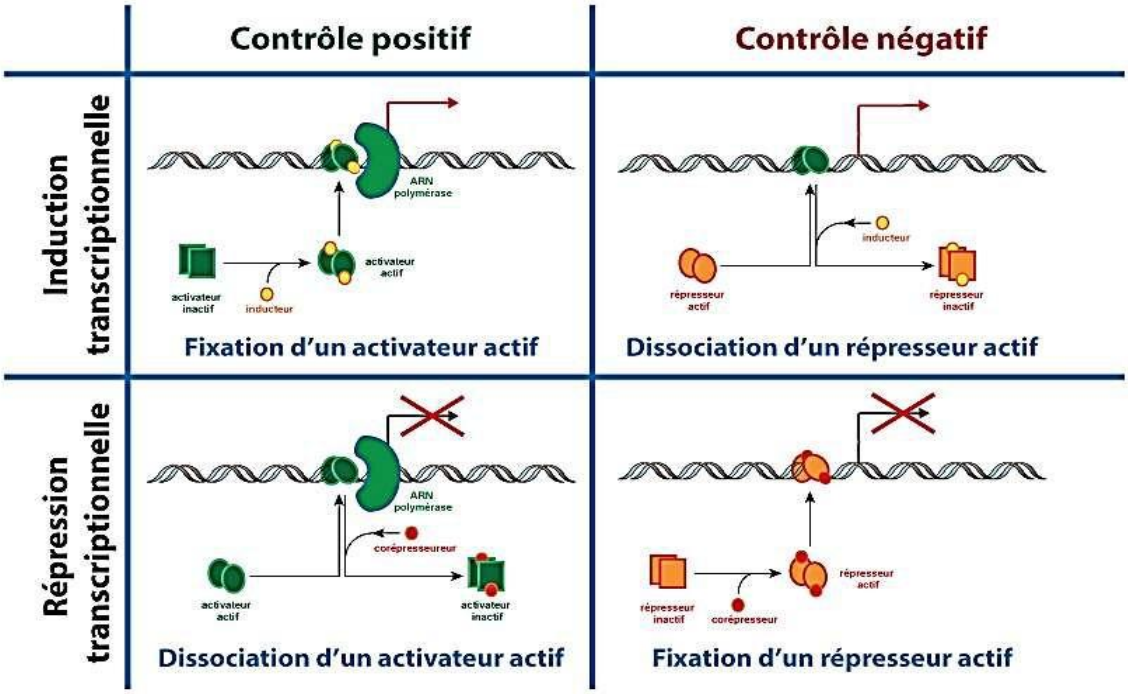

Figure : The functioning of the proteins of regulation (https://rnbio.sorbonne-universite.fr)

2. Regulation of the expression in prokaryotic cell

Regulation can take place at the level of transcription, translation or both. The aim of regulation is to adapt to nutritional conditions as economically as possible. The operon is a unit of expression and regulation of bacterial genes made up of structural genes and control elements recognised by the product(s) of the regulatory genes, i.e. the grouping in space, on the same chromosome, of genes required for the same metabolic function (mRNA is polycistronic, i.e. it contains the information required for the synthesis of different proteins). There are two main types of operon:

-**Inducible operons:** code for enzymes in the catabolic pathway (degradation pathway). Example: lactose operon in the E-coli genome, which is an example of an inducible gene (i.e. genes are expressed when necessary by inactivating a repressor).

-**Repressible operons:** code for enzymes in the anabolic pathway (biosynthesis). Example: tryptophan operon.

2.1. At the level transcriptional

Transcription is often regulated by protein factors that bind to DNA. Genes subject to these regulations are involved in different metabolic pathways

2.2.1. Regulation of the expression of genes involved in catabolic pathways (Example: the Lactose operon)

- **Metabolism of lactose**

The bacterium finds its carbon source in the catabolism of sugars. While glucose is the "preferred" carbon source, lactose (which is a β-galactoside) can also be consumed by the bacteria and metabolised into galactose and glucose. The enzymes needed to use lactose will only be synthesised in the presence of this substrate. The following animation shows lactose permease, which allows lactose to enter the cell with a flow of protons, and β-galactosidase, which hydrolyses lactose into hexoses, which follow other metabolic pathways.

- **Organization of the lactose operon**

At the level of the lactose operon, we can highlight three structural genes essential to the degradation of lactose: the Lac Z, Lac Y and Lac A genes, coding for different proteins. The genes of these proteins being part of the same transcription unit, we then speak of a **polycistronic unit** . They code:

- **β-galactosidase (lacZ gene)** which catalyses hydrolysis of the β1-4 osidic bond of β- galactosides. In this way, it hydrolyses lactose into galactose and glucose,
- **- Lactose permease (lacY gene).** This membrane protein allows lactose to enter the cell.
- **- Thiogalactoside transacetylasc (gene /acA).** Its role is not well known. It acetylates non-metabolizable P-galactosides, which can then be eliminated from the cell by diffusion across the plasma membrane.
- These three structural genes are preceded by a region responsible for regulating their expression. This regulatory region comprises the promoter and the operator, known as the regulatory elements:
- **- A cis-active element,** the operator.
- **- A trans-active factor,** the repressor, which is the result of the Lac I regulatory gene and acts at the level of the operator. Lac I has its own (non-inducible) promoter and continuously expresses a repressor which blocks expression of the lactose operon.

- - in the absence of lactose, is in its active form (tetrameric) and binds to operator site 0, preventing the RNA polymerase from binding and giving expression to the lactose operon (Repression)
- - in the presence of lactose, the repressor complexes with allolactose (a lactose

isomer). The repressor bound to the allolactose changes confrontation, loses its affinity for the operon and dissociates from the lac operon. The lactose operon can then be expressed (Induction).

- **Regulation negative of transcription**

In the absence of lactose, the repressor is active. It will bind specifically to the operon of the lactose operon, blocking RNA polymerase access to the transcription initiation site.

As a result, transcription of the lactose operon genes is negatively regulated. The enzymes required for lactose metabolism are not synthesised because they are useless in the absence of lactose.

- **Regulation positive of transcription**

In the presence of lactose, allolactose (/J-D-galactopyranosyl-(1-6)-P-D-glucopyranose), an isomer of lactose, acts as an inducer, binding to the repressor and inactivating it. This binding leads to a conformational change in the repressor, which then loses its affinity for the operator.

With the operator site released, the RNA polymerase can reach the transcription initiation site and synthesise polycistronic RNA. Production of the enzymes required for lactose metabolism is therefore dependent on the presence of the substrate.

The introduction of the lactose operon requires two conditions: the presence of lactose and the absence of glucose.

If both lactose and glucose are present, the operon is only weakly activated (basal activity). In this case, the CAP protein is inactive because the level of cAMP is low. So, in the absence of active CAP, RNA polymerase binding to the promoter is low, even if the repressor has been inactivated. cAMP: Cyclic adenosine 5'-monophosphate, an intracellular second messenger formed by adenylate cyclase.

CAP (or CRP): Catabolite gene Activator Protein, or CRP for "cAMP Receptor Protein". This protein controls the initiation of transcription of the catabolite genes which will enable the use of other nutrient molecules when glucose is absent.

CAP-AMPc: complex formed after the interaction of the CAP (or CRP) protein with cAMP.

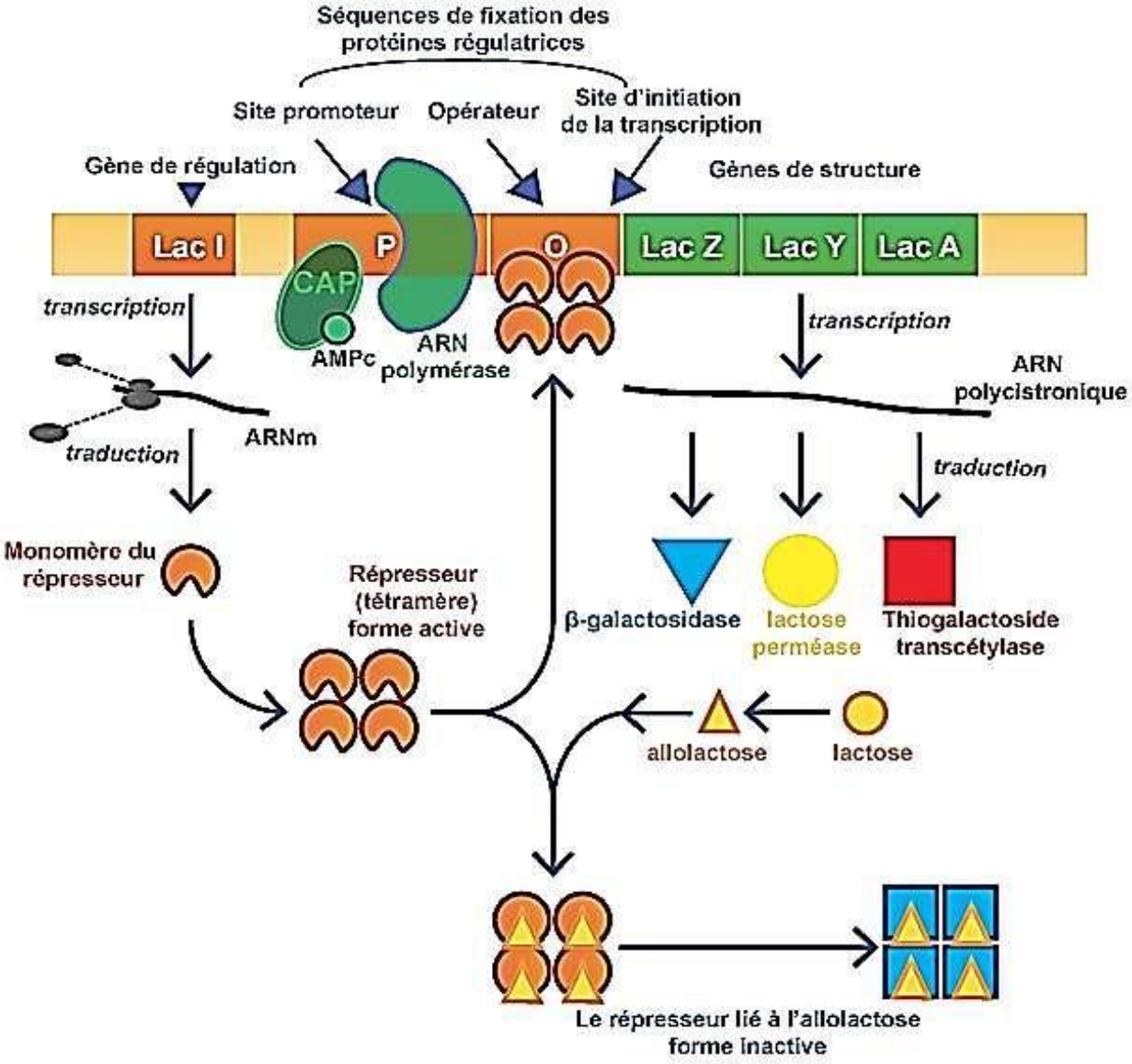

Figure: Regulation of the expression of the Genoa involved In THE ways catabolic (Example : the Lactose operon) (https://rnbio.sorbonne-universite.fr)

- **In summary :**

Transcription of the lactose operon is therefore under the control of two regulatory proteins:

- - the lacI repressor binds to the operator in the absence of lactose and blocks RNA polymerase, resulting in negative regulation.
- - the CAP (or CRP) protein, active in the form of a complex with cAMP, which binds to the DNA and increases the affinity of the RNA polymerase for the promoter, and is positively regulated in the presence of lactose.

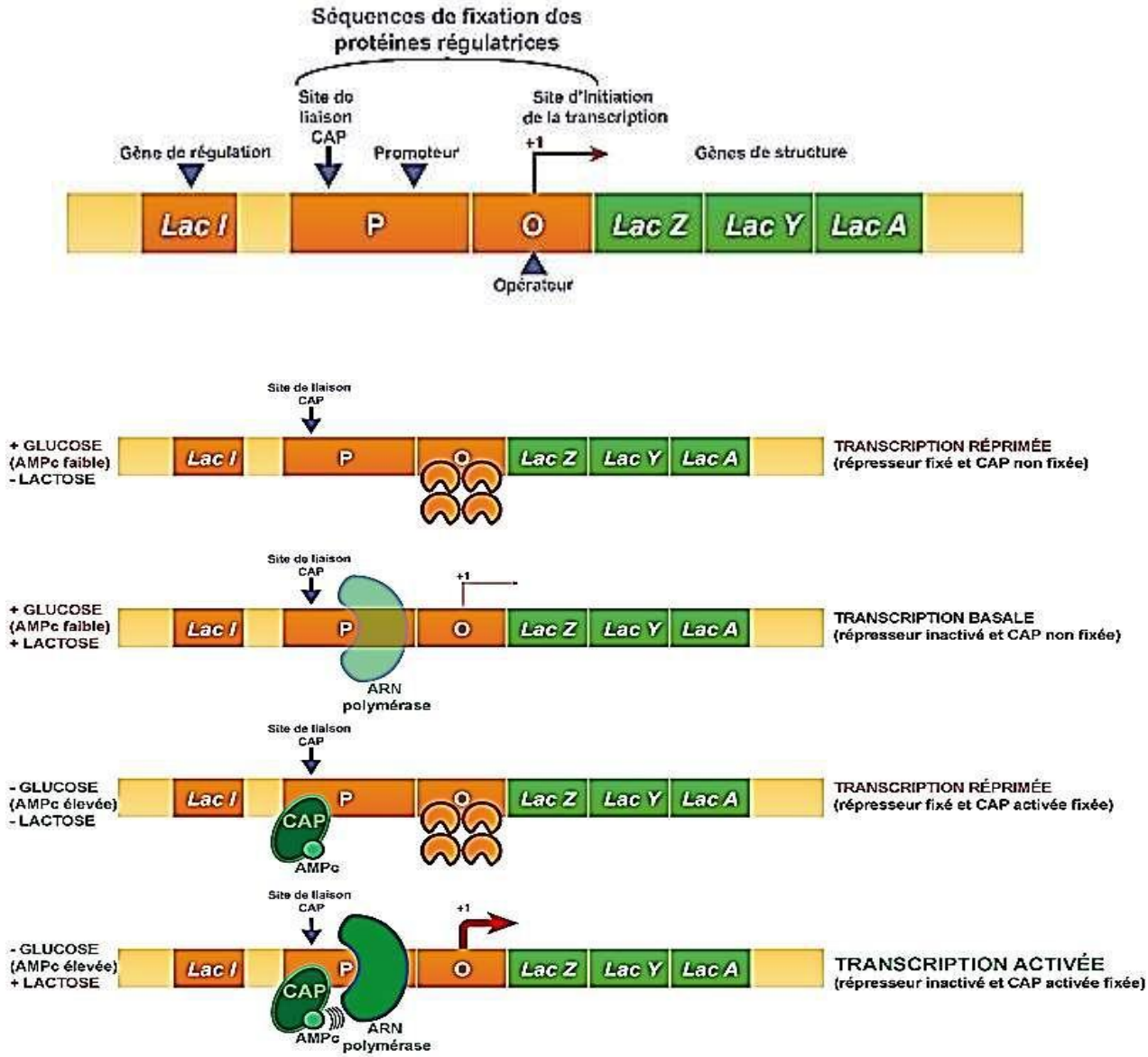

Figure: Effects combined of glucose And of lactose on the expression of the operon lactose (https://rnbio.sorbonne-universite.fr)

2.2.2. Operon tryptophan : operon repressible

In prokaryotes, biosynthetic enzyme systems are another example of transcription regulation. The tryptophan operon

The tryptophan operon contains 5 genes encoding enzymes involved in tryptophan biosynthesis. Transcription is regulated by the level of tryptophan in the cell.

Upstream of the structural genes is a regulatory sequence coding for a repressor. If tryptophan is present in the cell, it binds to the repressor, which is then activated and can bind to the operator. This will prevent the RNA polymerase from carrying out transcription.

Tryptophan acts as a corepressor; if tryptophan is absent, the repressor cannot bind to the operator and transcription will therefore take place.

2.2. At the level translational

Regulation at the translational level is little used in prokaryotes. One example is the regulation of the synthesis of ribosomal proteins whose genes are organised into operons. An excess of these leads to inhibition of their own translation.

3. Regulation of gene expression in eukaryotic cells

In multicellular organisms, the expression of different genes leads to cellular specialisation.

Precise regulation of gene expression is essential for the production and maintenance of the many cell and tissue types in a multicellular organism.

Cells differentiate into specific cell types through combinations of expressed and repressed genes.

There are several levels of control:

The expression of eukaryotic genes is regulated at the chromatin level, at the transcriptional level, at the post-transcriptional level, at the translational level and at the post-translational level.level.

3.1. At the chromatin level

For a gene to be transcribed, it must be activated. An activated gene is located in non-compacted regions of the chromatin (euchromatin). Otherwise, the gene will not be accessible to polymerases. The other condition for the gene to be transcribed is that it must not be methylated. Base methylation is recognised by enzymes and triggers DNA condensation, leading to gene inactivation.

Chromatin structure is therefore regulated by cytosine methylation or histone acetylation, as well as by chromatin remodelling complexes (CRCs) which break the bonds between

histones and DNA.

3.2 At the level transcriptional

- The main regulatory phenomena concern this stage in particular

-Regulation generally concerns the initiation phase, involving the various TRANSCRIPTION FACTORS:

- cis-regulatory regions: promoter and enhancer (a region of DNA that can bind proteins to stimulate transcription). Promoters and the genes they control are generally adjacent; the promoter is the DNA sequence to which the RNA polymerase binds to begin transcription. The additional sites, called enhancers, may be several hundred or thousands of base pairs upstream or downstream of the promoters they stimulate.
- Trans-regulatory factors: proteins that bind to the enhancer and have NEGATIVE (repressor or silencer) or POSITIVE (activator or amplifier) effects on transcription (depending on the needs of the cell concerned).

- Transcription can also be regulated by alternative promoters: the existence of several promoters in the same DNA region for the same transcribed sequence and therefore for the same gene.

- Transcription can also be regulated by extracellular signals (steroid hormones, thyroid hormones, etc. which act on nuclear receptors).

3.3 At the post- transcriptional level

- **Alternative splicing:** See maturation of primary transcripts in the DNA Transcription chapter.

- **RNA editing:** Like alternative splicing, RNA editing involves obtaining two different proteins from the same gene, but the mechanism is different. This is possible because the same gene is treated differently depending on the organ in which it is found.

3.4. At level translational And post- translational

Regulation can be achieved by modulating the lifespan of mRNAs. In general, mRNAs have a fairly short lifespan, but some mRNAs have a longer lifespan, such as the haemoglobin chain.

Regulation can also be achieved by siRNAs and microRNAs, which block translation. This

post-transcriptional inhibition of gene expression is achieved by interference using anti-sense RNA (siRNA or microRNA) which pair with the sense RNA by complementary sequences. This inhibition can be achieved artificially through the use of trans-genes, which correspond to artificially supplied genes.