

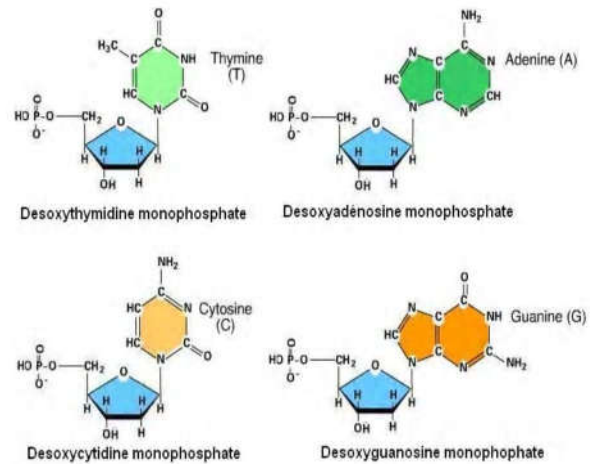
Université Biskra  
 Master Biochimie Fondamentale et Appliquée  
 Matière : Expression des gènes des procaryotes et régulation  
 Semestre : S2  
 Enseignant : M.A. Derradji yacine

# Expression des gènes des procaryotes et régulation

## I- Introduction

### 1- ADN, support de l'information génétique :

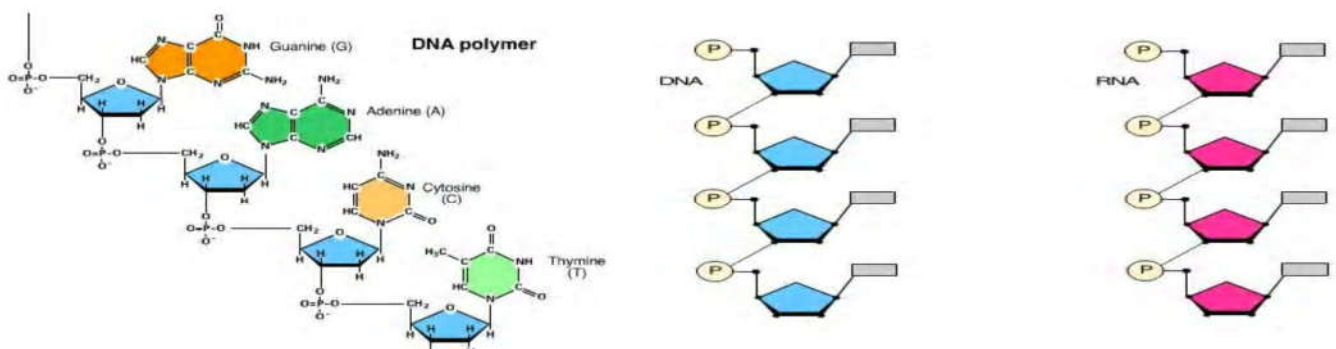
Les gènes sont constitués de brins d'acides nucléiques: acide désoxyribonucléique presque toujours. Les brins d'acides nucléiques sont composés d'une succession de nucléotides unis les uns aux autres par des liaisons phosphodiester. Chaque nucléotide se compose d'une base qui lui confère une identité (A, C, G, T), d'un sucre à 5 atomes de carbone (pentose) le désoxyribose, et d'acide phosphorique lié à l'hydroxyle 5' du sucre.



**Figure 1** : Nucléotide, unité structurale de l'ADN

La synthèse des brins d'ADN dans les cellules se fait toujours par addition d'un nucléotide au niveau de l'extrémité de la chaîne polynucléotidique ayant un radical hydroxyle libre en 3', radical hydroxyle avec lequel réagit le résidu d'acide phosphorique lié à l'hydroxyle 5' du nucléotide ajouté lequel aura lui-même un hydroxyle 3' libre du côté opposé.

L'élongation des brins d'acide nucléique est donc polarisée, du fait que la réaction d'élongation se réalise dans le sens «5'→3'». Cette polarité permet également de désigner les deux extrémités d'un brin d'acide nucléique (extrémité 5', signifiant l'extrémité ayant un hydroxyle 5' libre ou estérifié par un acide phosphorique; extrémité 3', signifiant l'extrémité ayant un hydroxyle libre en 3').

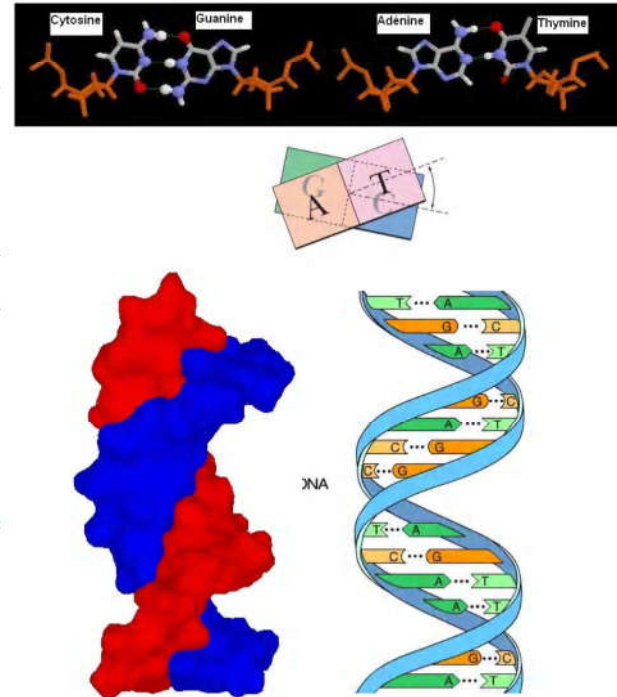


**Figure 2** : Le DNA monocaténaire, polymère polarisé de 5' à 3'

L'ADN est le plus souvent sous la forme d'un double brin s'appariant avec précision grâce à la complémentarité des bases, une base purique s'appariant à une base pyrimidique, A avec T et G avec C.

Les deux brins d'une double hélice d'ADN sont donc antiparallèles, un brin orienté dans le sens 5'→3' de gauche à droite s'appariant avec son brin complémentaire orienté dans le sens 3'→5'.

Dans l'ADN double brin, qui recèle l'information génétique, le squelette des ponts phosphodiester est à l'extérieur; l'axe de l'exosquelette est presque perpendiculaire aux axes des bases (ADN B). Ces dernières, dirigées vers l'intérieur, forment un empilement de structures moléculaires planes et parallèles dont les axes sont décalés de 34,6° entre deux couples de bases : le double brin d'ADN donc hélicoïdal, réalisant un tour complet (360°) toutes les 10,4 bases. Cette structure hélicoïdale comporte deux sillons : grand sillon et petit sillon, qui jouent un rôle important dans la fixation des protéines régulatrices.

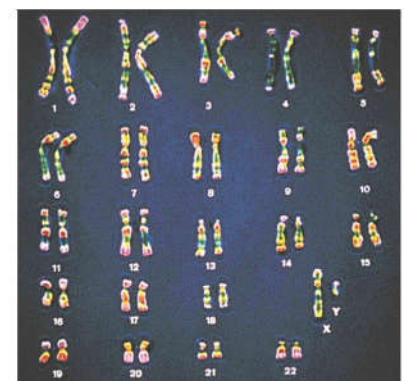
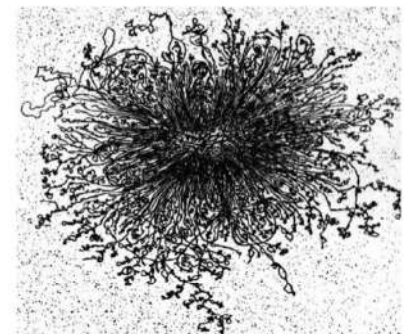


**Figure 3** : L'ADN double hélice

## 2- Le génome eucaryote et procaryote

Au sens strict, le génome est l'ensemble des gènes d'un organisme. Par extension, on utilise ce terme pour désigner l'ensemble des molécules portant les gènes, c'est-à-dire l'ADN.

On dira donc que le génome humain (haploïde) est composé de 3 millions de kilopaires de bases (kpb) ( $3 \times 10^9$ pb) organisé dans 23 chromosomes. La plus grande partie de cet ADN ne code pour aucune séquence protéique et est constituée soit de répétitions plus ou moins monotones de nucléotides (ADN répétitif), soit de séquences uniques non codantes faisant ou non partie des gènes.

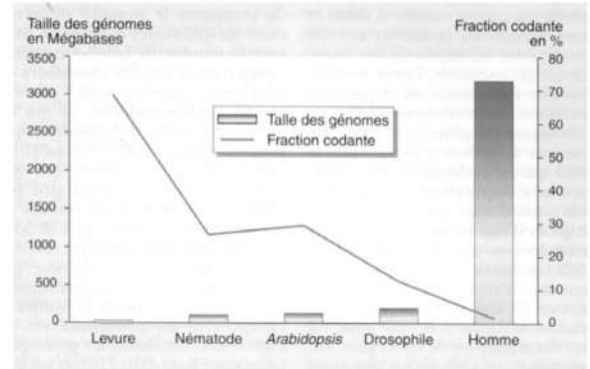


**Figure 4** : génome eucaryote et procaryote



Le génome bactérien est beaucoup plus petit. Chez *Escherichia coli*, par exemple, il est seulement d'environ 5000 kpb ( $5 \times 10^6$ pb) contenus dans un seul chromosome.

Le nombre total de gènes dans un génome donné n'est évidemment pas connu avec précision. On l'évalue, chez l'homme, à environ 30000 à 50000 gènes, qui occupent probablement moins du dixième du génome total. Chez les bactéries, les gènes sont beaucoup plus rapprochés et ne contiennent que des séquences codantes, ce qui rend très élevée ici la proportion d'ADN « signifiant ».



**Figure 5 :** Taille et pourcentage exprimé du génome.

Le rôle, chez les animaux et chez l'homme, des séquences non codantes d'ADN n'est toujours pas connu en détail. Elles contiennent des éléments de contrôle de la transcription des gènes, mais aussi de larges régions qui n'ont probablement aucune fonction et dont la signification précise dans l'évolution des espèces reste l'objet de spéculations.

## II- Notions sur la régulation de l'expression génétique

### 1- L'expression des gènes n'est pas anarchique :

En comparant les différentes cellules spécialisées chez l'homme (neurone et lymphocyte, par exemple), et vu la grande différence morphologique et physiologique, les scientifiques pensaient que la spécialisation conduit à la perte sélective de gènes (la cellule perd les gènes qu'elle n'utilise plus).



**Figure 7 :** Ovule fécondée

Les expérimentations menées au cours des années cinquante sur les ovules énucléés dont on remplaçait le noyau par celui de cellules spécialisées, ont confirmé la conservation de la totalité du génome dans les cellules spécialisées.

La grenouille *Xenopus laevis* a fait l'objet de ces premières expérimentations. Les noyaux prélevés à différents stades de développement embryonnaire (Blastula, Gastrula, Neurula ...) puis transférés dans des ovocytes énucléés ont permis d'obtenir des grenouilles, mais le succès de l'opération diminue en avançant dans le stade développement embryonnaire duquel le noyau est obtenu.



**Figure 8 :** *Xenopus laevis*.



En 1962 John Gurdon (université d'Oxford, prix Nobel 2012 avec Shinya Yamanaka) a réussi à obtenir des grenouilles en transférant des noyaux prélevés de cellules dans leur dernier stade de spécialisation (cellules intestinales) ce qui a confirmé qu'aucun gène n'est perdu dans le processus de spécialisation. Chez les mammifères la première expérience de transferts nucléaire réussite a été faite en 1996 par Ian Wilmut et a donné naissance à la brebis Dolly.



**Figure 9** : La brebis Dolly.

Il est maintenant admis que l'activité des unités génétiques (gènes) ne s'exerce pas de manière indépendante et anarchique, elle est coordonnée par des mécanismes de régulation. Ceux-ci permettent à la cellule de s'adapter au milieu en utilisant de manière variée ses unités génétiques. La régulation intervient également dans la différenciation des cellules des organismes pluricellulaires.

On désigne sous le nom de régulation génétique l'ensemble des mécanismes capables de moduler l'expression du potentiel héréditaire dans une cellule vivante en fonction des signaux qu'elle reçoit, en provenance d'autres cellules ou de l'environnement.

Les modalités de cette régulation sont maintenant très bien connues chez les micro-organismes, tels que bactéries ou levures. Ces cellules, vivant en contact direct avec un milieu extérieur susceptible de subir d'importants changements, ne maintiennent leur intégrité que par l'adaptation de leur métabolisme aux conditions du moment.

Cependant, il existe des gènes qui sont exprimés de façon à peu près identique dans toutes les cellules d'un organisme, quelles que soient les conditions, on les dénomme « gènes constitutifs » ou encore « gènes de ménage » (codant pour les ribosomes, cytosquelette, histones...).

## **2- le contrôle de l'expression des gènes peut être positif ou négatif**

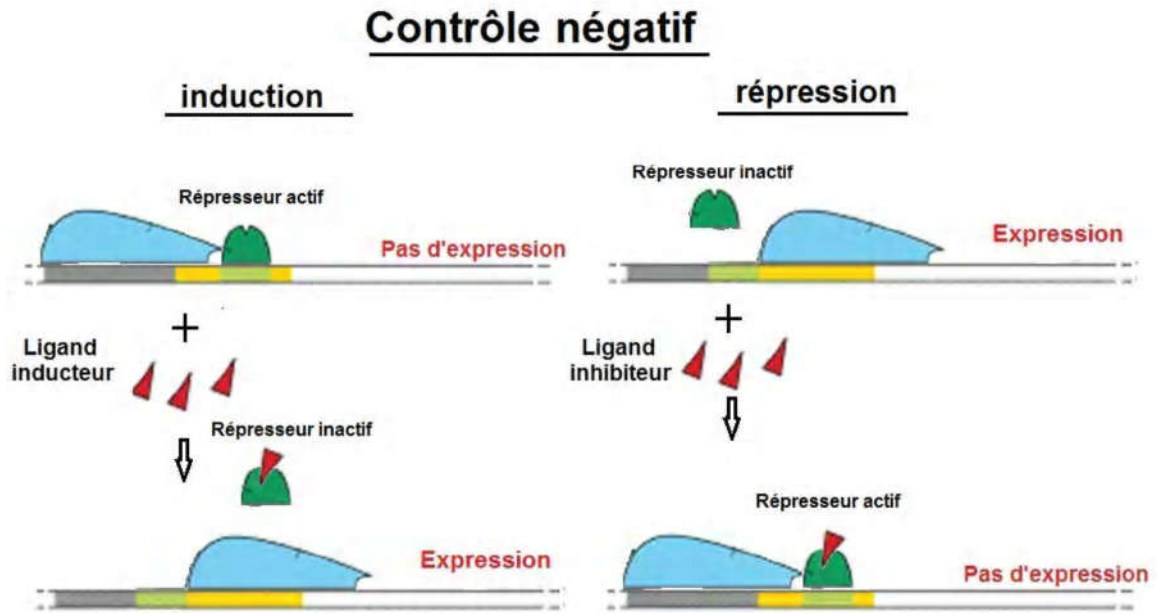
L'expression des gènes peut être contrôlée négativement ou positivement. Le contrôle négatif est exercé par des facteurs appelés répresseurs et le contrôle positif est exercé par des activateurs. L'activateur et le répresseur doivent exister sous deux formes "active" et "inactive", cela est généralement réalisé à l'aide de substances chimiques appelées ligands. Un ligand inducteur est capable d'activer un activateur ou de désactiver un répresseur alors qu'un ligand inhibiteur désactive un activateur ou active un répresseur.

Dans le cas du contrôle négatif on observe deux situations :

- gènes inductibles : dans ce cas les gènes sont sous le contrôle d'un répresseur actif qui empêche leur expression. Un ligand inducteur désactive le répresseur et permet l'expression des gènes.



- gènes répressibles: dans ce cas les gènes sont sous le contrôle d'un répresseur inactif, donc ils sont exprimés de façon normale (pas de répression). Un ligand inhibiteur (corépresseur) active le répresseur ce qui empêche l'expression des gènes.

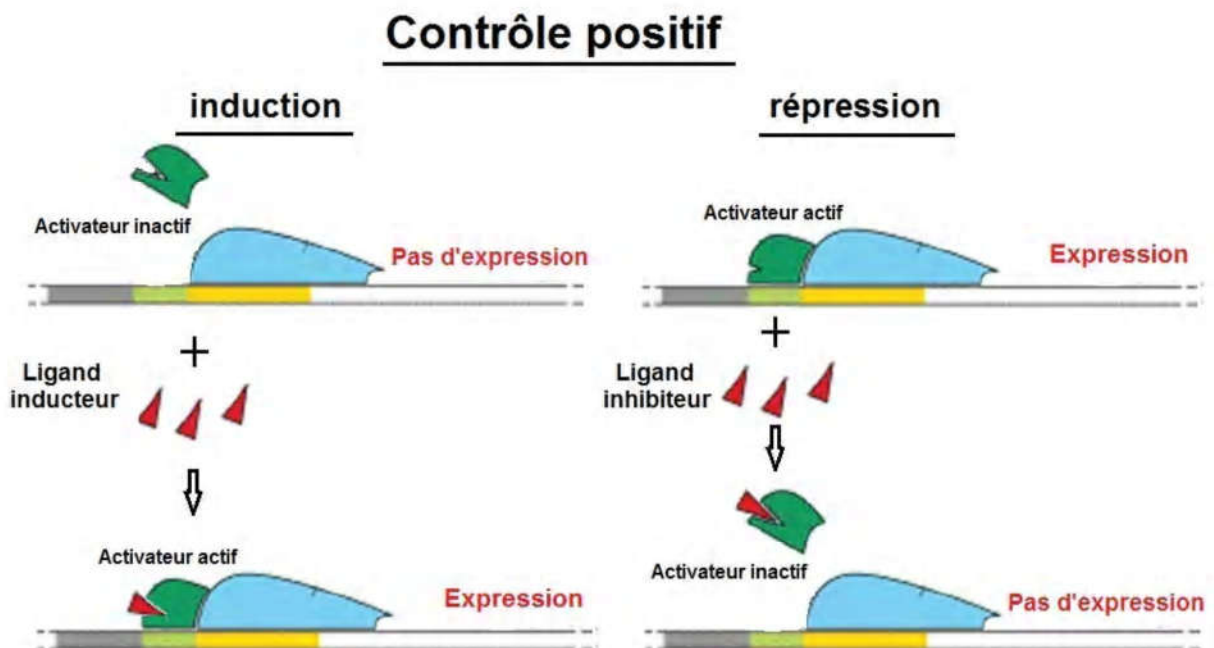


**Figure 10 :** Contrôle négatif.

Dans le cas du contrôle positif on observe deux situations :

- gènes inductibles : dans ce cas les gènes sont sous le contrôle d'un activateur inactif. Un ligand inducteur (coactivateur) active l'activateur et permet l'expression des gènes.

- gènes répressibles: dans ce cas les gènes sont sous le contrôle d'un activateur actif, donc ils sont exprimés. Un ligand inhibiteur désactive l'activateur ce qui empêche l'expression des gènes.



**Figure 11 :** Contrôle positif.

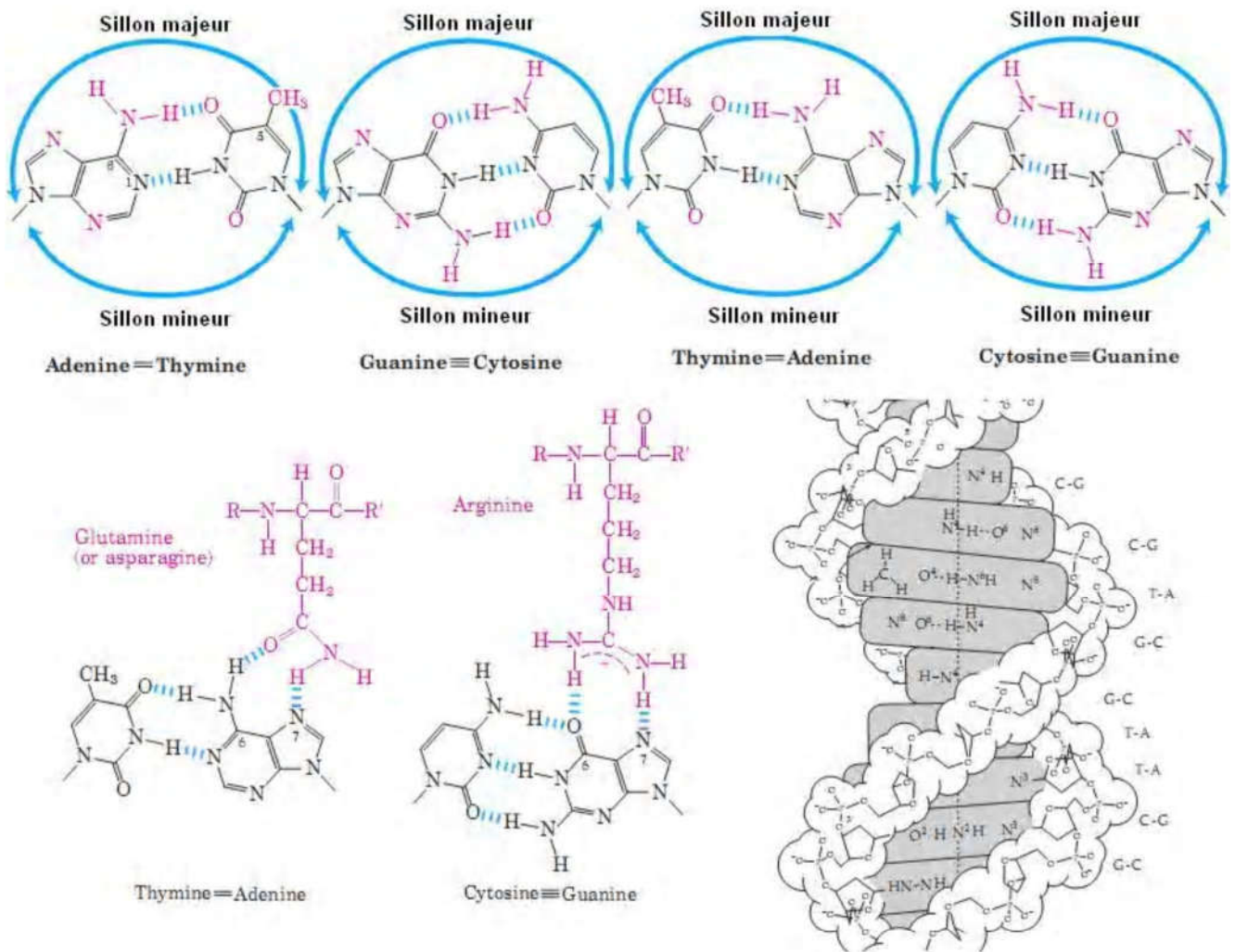


### 3- Des protéines qui se lient à l'ADN participent à la régulation

Différentes protéines (facteur trans) découvertes au départ chez les procaryotes (répresseur lambda, répresseur Lac...), reconnaissent des séquences spécifiques de paire de base sur l'ADN double hélice (éléments cis-régulateur). En se fixant sur ces séquences régulatrices, les protéines participent à la régulation positive ou négative de l'expression génétique.

On pensait que ces protéines devaient avoir accès aux liaisons hydrogènes entre les bases pour connaître les séquences spécifiques, mais maintenant il est prouvé que la structure externe de la double hélice est chargée d'information sur la séquence nucléotidiques, les bords de chaque paire de base sont exposés à la surface de la double hélice (sillon majeur et mineur) présentant un modèle différent de donneur et d'accepteur de liaison hydrogène ainsi que des groupe hydrophobes.

Toutes ses protéines reconnaissant l'ADN partagent des domaines structuraux qui sont responsables de cette reconnaissance spécifique. On distingue quatre domaines : hélice-tour-hélice, doigts de zinc, leucine fermeture éclair, hélice-boucle-hélice.



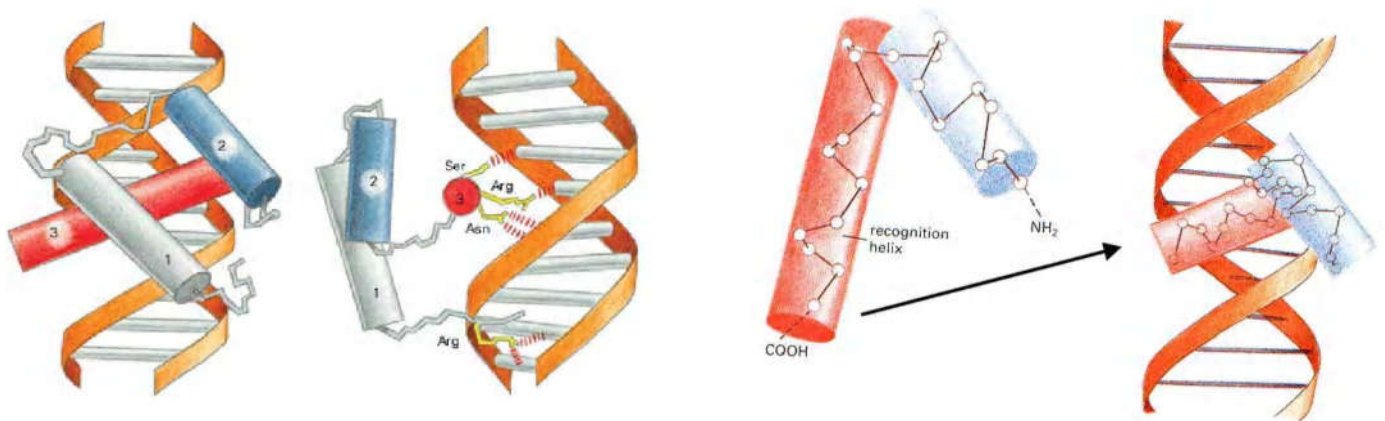
**Figure 10** : Reconnaissance spécifique protéine-ADN



**Domaine hélice-tour-hélice (Helix-Turn- Helix ; HTH) :**

C'est le premier domaine liant l'ADN à être découvert, chez les procaryotes (répresseur Lac et Trp et phage-λ, CRP ...). Il est constitué de deux hélices α, situées à un angle droit, connectées par une courte séquence peptidique (tour), l'hélice la plus proche du carboxyle terminal est appelée hélice de reconnaissance, car elle s'étend en partie dans le sillon majeur et interagit avec l'ADN, sa séquence en acides aminés, différente d'une protéine à l'autre, détermine la séquence ADN à reconnaître.

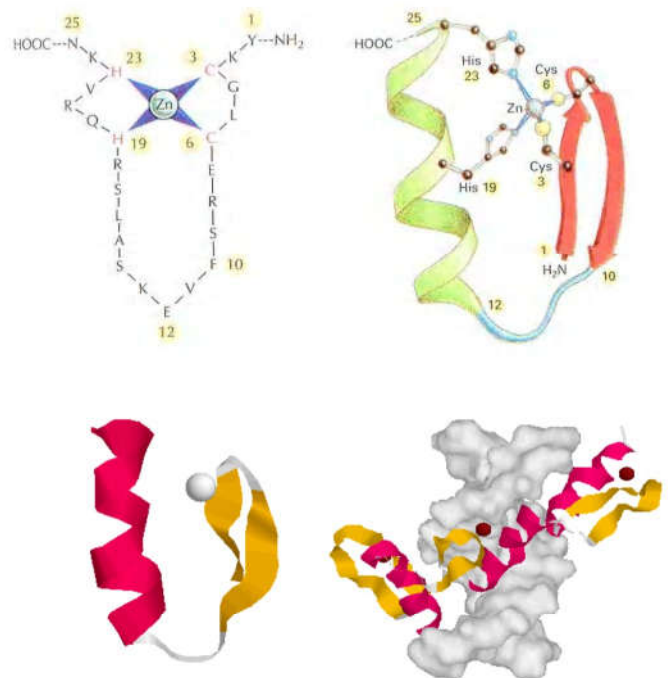
L'homéodomaine est une classe spéciale du domaine HTH avec une troisième hélice α, il a été découvert au départ dans la drosophile, puis dans la plupart des eucaryotes étudiés.



**Figure 11 :** interaction domaine HTH-ADN

**Domaine doigt de zinc (zinc finger) :** ce domaine est constitué de deux feuillets β et une hélice α qui assure l'interaction avec l'ADN, la liaison de l'ion de zinc est coordonnée par deux résidus histidine et deux cystéines, ce domaine se répète trois fois dans le facteur de transcription SPI et neuf fois dans TFIIIA.

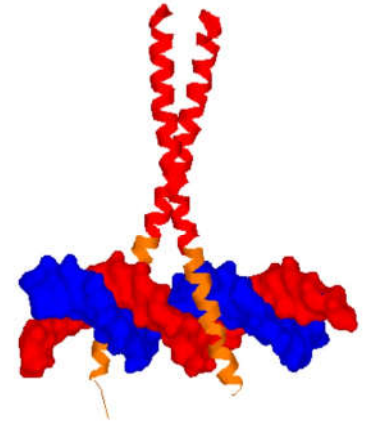
Dans les récepteurs nucléaires des hormones stéroïdiens, l'ion de zinc est coordonné par quatre résidus cystéine.



**Figure 12 :** Interaction domaine doigt de Zn-ADN

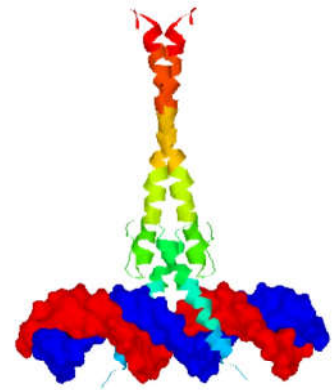


**Domaine Leucines en fermeture éclair (Leucine zipper):** dans cette structure on distingue un domaine de base formé par une l'hélice  $\alpha$  permettant la liaison à l'ADN. A l'extrémité C-terminale de ce domaine s'étend une région hélicoïdale riche en résidus Leucine qui se positionnent tous sur un seul côté de l'hélice formant une surface hydrophobe. Cette surface constitue un domaine de dimérisation entre deux protéines (exemple : GCN4).



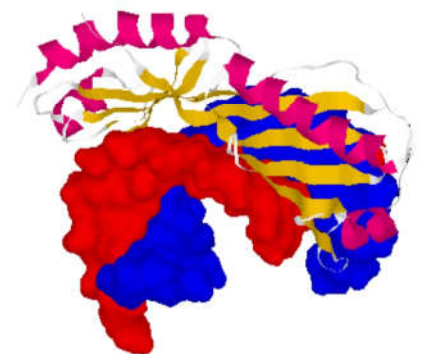
**Figure 13 :** interaction domaine Leucine zipper-ADN

**Domaine Hélice-boucle-Hélice (Helix-loop-Helix : HLH):** c'est un domaine à double fonction comme le domaine leucine zipper : liaison à l'ADN et dimérisation. La seule différence est la présence d'une hélice courte (domaine de base) et une autre long (domaine de dimérisation) reliées par une partie non hélicoïdale (boucle) peptidique (exemple : MAX).



**Figure 14 :** interaction domaine Hélice-boucle-Hélice-ADN

**Les feuillets  $\beta$  peuvent eux aussi reconnaître l'ADN :** dans les quatre domaines précédents l'hélice  $\alpha$  est la structure utilisée pour la liaison à l'ADN. Cependant, il existe des protéines qui adoptent une stratégie différente pour la reconnaissance des séquences d'ADN, en utilisant les feuillets  $\beta$ . On cite le répresseur Met et la TATA-BP (la protéine liant la boîte TATA).

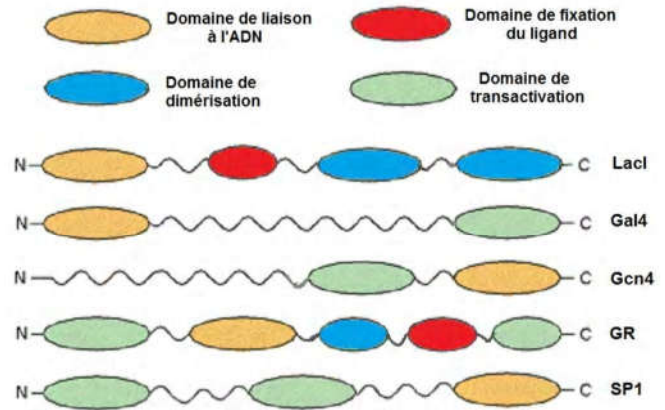


**Figure 15 :** interaction TATABP-ADN



**Les protéines de liaison à l'ADN possèdent d'autres domaines:**

En plus du domaine de liaison à l'ADN et le domaine de dimérisation. Certaines protéines de régulation possèdent un domaine de liaison du ligand qui permet la régulation de l'activité de la protéine en fixant une petite molécule accessoire. On trouve également, des surfaces d'interaction avec d'autres régulateurs de la transcription que l'on appelle domaine de transactivation. Les récepteurs des hormones stéroïdiens sont un exemple des protéines régulatrices contenant ces quatre types de domaines.



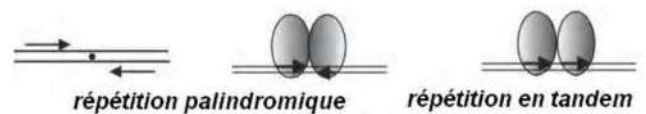
**Figure 18** : les différents domaines des protéines régulatrices.

**4- Certains aspects de l'interaction protéines de régulation-AND**

**\* Répétition des séquences de reconnaissance**

La taille des séquences de reconnaissance est comprise, en générale, entre 3 et 8 pb. Souvent on trouve plusieurs copies de ces séquences sous formes de répétitions palindromiques (la même séquence de gauche à droite ou de droite à gauche "radar", "Laval") ou tandem (directes). Les sites de liaison des protéines qui forment des homodimères sont souvent sous forme d'une répétition palindromique ce qui permet le positionnement idéal des domaines de liaison à l'ADN et de dimérisation permettant une liaison coopérative qui offre plus d'affinité avec l'ADN que la forme monomérique.

Les séquences en tandem sont généralement utilisées par des hétérodimères ou pour l'oligomérisation. La séquence et la position des séquences de reconnaissance sont deux facteurs indispensables à la liaison des protéines régulatrices, une petite variation peut conduire à l'impossibilité de liaison.



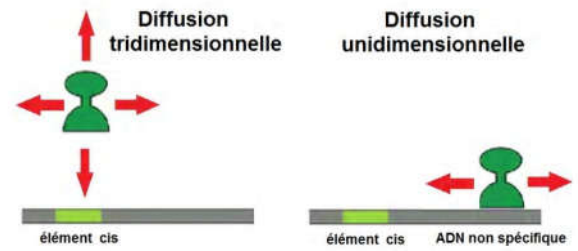
**Figure 19** : répétitions palindromique et en tandem.

**\* Rapidité de rencontre entre protéines de régulation et l'AND:**

La rencontre entre deux molécules en solution est gouvernée par la loi de la diffusion tridimensionnelle ( $t = x^2/6D$  ; t= temps, x= distance parcourue, D= coefficient de diffusion). Les chercheurs ont constaté que les protéines de liaison à l'ADN trouvent leurs sites plus rapidement, car elles se fixent n'importe où sur l'ADN avec une affinité très faible (forces électrostatiques avec

l'exosquelette de l'ADN) et scannent l'ADN (marche sur l'ADN) jusqu'à trouver le site spécifique où elles s'arrêtent car l'affinité est trop élevée.

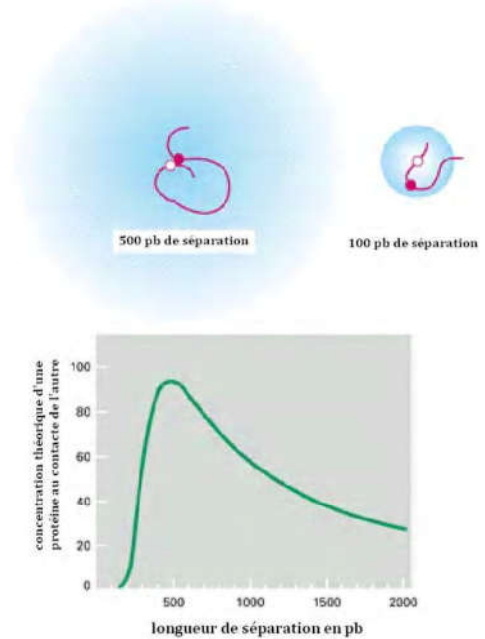
Donc la diffusion est unidimensionnelle ce qui réduit considérablement le temps nécessaire au rencontre. Un exemple est le répresseur : Riggs et ces collaborateurs (1970) ont estimé que sa vitesse réelle de liaison ( $K_{on}$ = constante d'association=  $10^{10} M^{-1}S^{-1}$ ) est 100 fois plus que la vitesse de liaison calculée selon la loi de diffusion 3D ( $K_{on}$ =  $10^8 M^{-1}S^{-1}$ ). Sa vitesse de scan est de  $10^6$  pb/s, donc il parcourt le génome d'E. coli dans 5 secondes.



**Figure 20** : diffusion unidimensionnelle.

**\* La liaison à l'ADN facilite le contact entre protéines:**

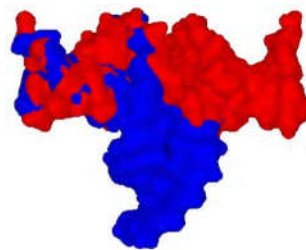
La liaison de deux protéines à l'ADN augmente leurs chances de contact en comparaison avec les protéines non liées. La double hélice d'ADN est flexible, un segment de 200 pb peut faire une courbure de 90° ce qui permet de rapprocher les deux protéines liés, les protéines séparées par moins de 200 pb ont moins de chance de se rejoindre. La longueur de séparation qui donne plus de chance de contact est 500 pb. Puisque l'ADN fait un tour chaque 10pb, la distance séparant deux protéine doit être un multiple de 10pb (exemple:  $20 \times 10 = 200$ ) ce qui place les protéines sur le même côté et favorise leur contacte.



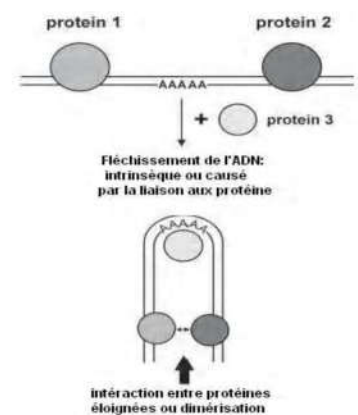
**Figure 21** : probabilité de contact entre protéines liées à l'ADN

**\* Influence de l'interaction avec les protéines sur la structure de l'ADN**

L'interaction protéine ADN n'est pas sans conséquence sur la structure de l'ADN. La liaison à l'ADN peut causer des changements conformationnels (affectant la largeur du sillon, l'inclinaison des paires de bases l'un par rapport à l'autre) et le fléchissement (courbure) de l'ADN (TATA-BP et CRP par exemple). La liaison d'une protéine qui courbe d'ADN à un fragment de moins de 200 pb favorise le contact entre deux protéines fixées aux deux extrémités de ce fragment d'ADN.



**Figure 22** : l'interaction avec TBP Courbe l'ADN de 100°



**Figure 23** : Une protéine qui courbe l'ADN facilite le contact d'autres protéines.



### III- Organisation et expression des gènes procaryotes

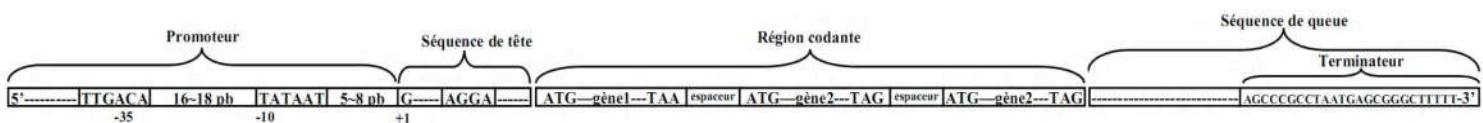
#### 1- Organisation du gène procaryote:

La structure du gène procaryote a été décrite la première fois par François Jacob et Jacques Monod en 1961, suite à leur observation que les gènes impliqués dans l'utilisation du lactose sont exprimés ou réprimés au même temps.



**Figure 24 :** François Jacob et Jacques Monod

Les gènes procaryotes sont généralement organisés sous forme d'opérons où plusieurs gènes effectuant des tâches souvent complémentaires (même voie métabolique par exemple) sont soumis au contrôle d'un seul promoteur. Les gènes d'un opéron sont toujours soit exprimés, soit réprimés au même temps. La similitude de réaction est liée au fait qu'une seule molécule d'ARN messager est synthétisée par l'ensemble. Le messenger polygénique (polycistronique) est ensuite traduit en autant de polypeptides qu'il y a de gènes contigus.



**Figure 25 :** structure d'un gène procaryote.

Le promoteur procaryote est une séquence du gène qui participe à l'expression mais ne sera pas transcrite. La séquence du promoteur varie d'un gène à l'autre, néanmoins on constate une conservation dans deux segments (séquence consensus), la séquence exacte de ces deux segments détermine l'efficacité du promoteur.

séquence en -10					
T	A	T	A	A	T
77%	76%	60%	61%	56%	82%
séquence en -35					
T	T	G	A	C	A
69%	79%	61%	56%	54%	54%

**Figure 26 :** probabilités d'apparition de bases dans les séquences consensus.

La première séquence consensus se trouve à 35 bases en amont du site d'initiation de la transcription (-35) « 5'TTGACA3' », elle constitue le site de reconnaissance de l'ARN polymérase. La deuxième à -10 base « 5'TATAAT3' », c'est là où l'ARN polymérase commence à dérouler l'ADN

pour une transcription éventuelle (boite Pribnow). En plus des deux boites TTGA et TATA (Pribnow), le promoteur contient une (ou plusieurs) séquence opérateur située de part ou d'autre du site de liaison de l'ARN polymérase.



**Figure 27 :** différentes séquences -35 et -10 pour différents opérons.

La position +1 est appelée site d'initiation de la transcription, c'est là où la polymérase commence à ajouter des nucléotides, cette position est toujours occupée par une purine (G de préférence ou A).

En aval du site d'initiation se trouve une séquence de tête (leader), cette séquence ne fait pas partie du produit du gène (transcrite mais non traduite), elle contient une séquence consensus « 5'AGGA3' » (Shine-Dalgarno) reconnue par le ribosome ce qui permet son orientation sur l'ARNm pour initier la traduction. Dans certains opérons la séquence leader joue un rôle régulateur de la transcription (opéron Trp).

En aval du leader viennent les régions codantes (gènes) qui sont séparées par de courtes séquences (espaceur). Chaque gène est bordé par un codon ATG (initiation de la traduction) en 5' et l'un des trois codons TAA, TAG ou TGA (stop) en 3'.

En aval des régions codantes on trouve une séquence de queue (trailer) qui se termine par une séquence auto-complémentaire riche en G-C nommée terminateur. Une fois transcrit, le terminateur forme une structure en épingle à cheveux responsable de la terminaison de la transcription.

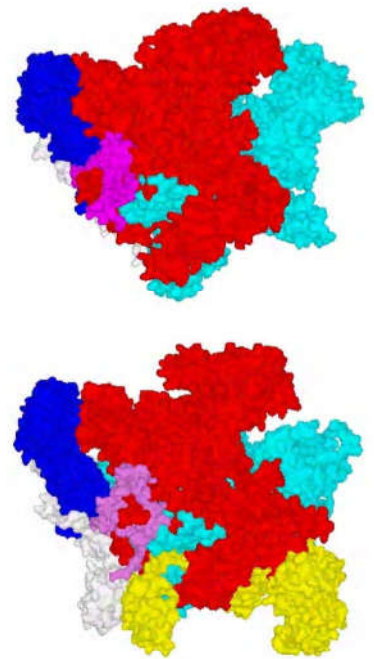


**Figure 28 :** Terminateurs Epingle à cheveux.

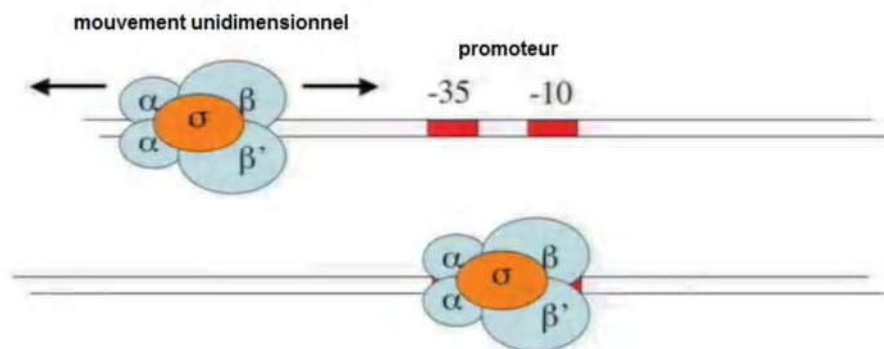


## 2- Expression du gène procaryote :

La transcription de l'ADN est orchestrée par l'ARN polymérase, qui est un complexe enzymatique responsable de la synthèse de l'acide ribonucléique, ou ARN, à partir d'une matrice d'ADN. L'association de cinq sous unités ( $\alpha 2\beta\beta'\omega$ ) forme une enzyme cœur dont chaque sous-unité possède une fonction spécifique :  $\alpha$  (36,5 kDa), la liaison à des séquences ou protéines régulatrices;  $\beta$  (150,6 kDa), contient le site catalytique qui forme les liaisons phosphodiester;  $\beta'$  (155,2 kDa), liaison au brin matrice; oméga (6kDa), assemblage de l'holoenzyme. L'enzyme cœur possède une grande affinité non spécifique à l'ADN mais elle n'arrive pas à reconnaître le promoteur ce qui l'empêche d'initier une transcription optimale. La liaison à un facteur  $\sigma$  (70,2 kDa) (reconnaissance du promoteur), formant l'holoenzyme, diminue d'un facteur de 10000 l'affinité non spécifique de l'ARN polymérase et augmente son affinité pour le promoteur de 1000 fois. L'efficacité de la liaison de l'enzyme et de l'initiation de la transcription dépende en grande partie des séquences -35 et -10 du promoteur.

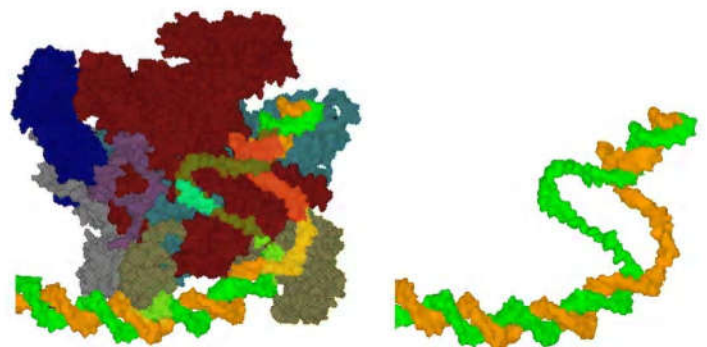


**Figure 29** : Structure de l'ARN polymérase cœur et holoenzyme E. coli



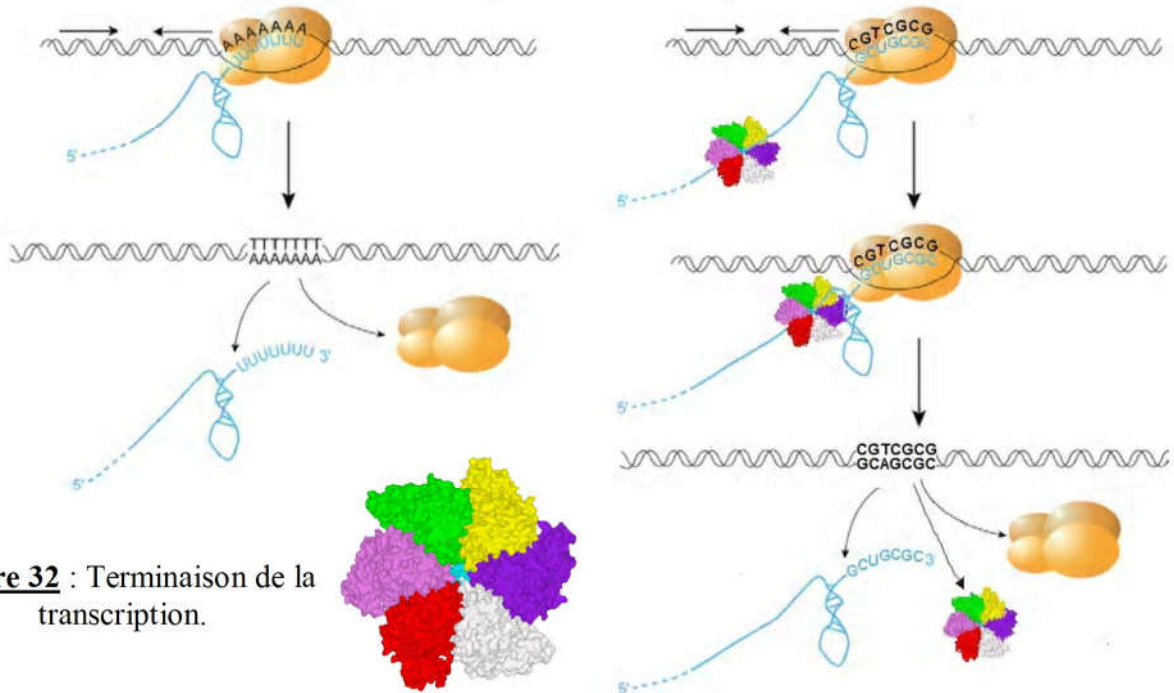
**Figure 30** : liaison de l'ARN pol holoenzyme au promoteur.

Après sa liaison, l'ARN polymérase déroule l'ADN au niveau de la boîte TATA créant une bulle de transcription de 17pb. Elle synthétise une copie du brin sens en utilisant le brin anti-sens comme matrice (complémentarité des paires de base). L'enzyme ajoute 9 nucléotides avant de commencer à bouger. A ce stade le facteur  $\sigma$  se détache et la polymérase cœur poursuit l'élongation du transcrit.



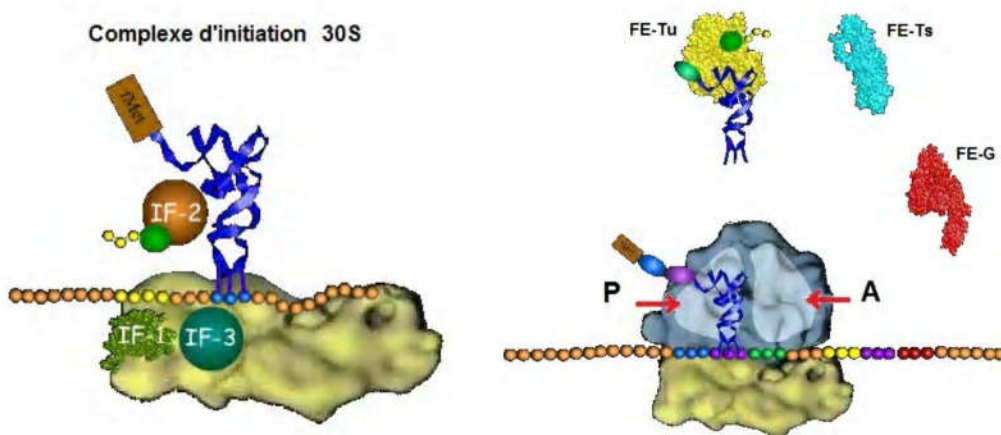
**Figure 31** : ouverture de l'ADN et initiation de la transcription.

Lorsque elle rencontre un terminateur (tige-boucle "épingle à cheveux") l'enzyme arrête la transcription et se détache avec ou sans l'aide d'un facteur rho (hexamère). La structure en épingle à cheveux oblige l'enzyme à faire une pose et tire l'ARN, qui se le détache sans l'aide du facteur rho (ARN possédant des U en 3') ou avec son aide (ARN sans U en 3').



**Figure 32** : Terminaison de la transcription.

La traduction s'effectue au niveau du ribosome, elle se déroule en trois étapes: initiation, élongation et terminaison. Initiation: la petite sous-unité (30S) du ribosome se dissocie de la grosse sous-unité (50S) avec l'aide des facteurs d'initiation 3 et 1 (FI-3, FI-1), ensuite elle se lie en 5' de l'ARNm (avant même la fin de la transcription) et cherche la séquence Shine-Dalgarno (complémentarité avec l'ARNr 16S), la liaison à cette séquence donne à la petite sous-unité une position idéale par rapport au codon Start (AUG). Le facteur d'initiation 2 (FI-2) se lie spécifiquement à l'aminoacyl-ARNt qui correspond au codon start (fMet-ARNt) et l'aide à se lier sur l'ARNm formant ainsi le complexe d'initiation 30S. L'arrivée de la grosse sous-unité et la dissociation des facteurs d'initiation donnent le complexe d'initiation 70S.



**Figure 33** : traduction de l'ARNm.



Elongation et terminaison: le facteur d'élongation Tu (FE-Tu) apporte un à un les aminoacyl-ARNt au site "A" libre sur le ribosome, ces derniers se fixent s'ils sont complémentaires aux codons de l'ARNm, une fois la fixation réussite le FE-Tu se dissocie en hydrolysant une molécule de GTP. La régénération du FE-Tu (fixation d'une nouvelle GTP) est aidée par le facteur d'élongation Ts (EF-Ts). Une réaction de traspeptidation permet le transfère du peptide naissant de l'ARNt de la position "P" vers l'ARNt de la position "A". L'intervention du facteur d'élongation G (FE-G) permet la translocation du ribosome déplaçant le peptidyl-ARNt vers la position "P" et libérant la position "A". L'élongation continue jusqu'à l'apparition d'un codon stop, à ce stade intervient les facteurs de libération 1, 2 et 3 pour libérer le peptide, l'ARNm et le ribosome.

#### **IV- Régulation de l'expression génétique chez les procaryotes**

##### **1- Régulation de l'initiation de la transcription:**

Chez les procaryotes la régulation de l'expression génique s'exerce essentiellement au niveau de l'initiation de la transcription. Cette régulation est dans la majorité des cas négative (utilisation de répresseurs), chez les eucaryotes le mode positif de régulation est le plus courant.

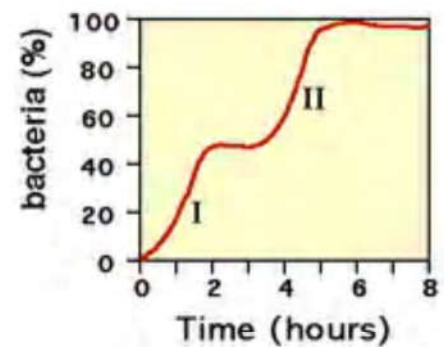
**a) Double contrôle de l'opéron lactose :** pour faire face à toutes les situations de présence ou absence du lactose et du glucose dans son environnement, *E. coli* contrôle négativement et positivement son opéron lac.

##### **- Phénomène de diauxie :**

Lorsque *E. coli* se développe dans un milieu contenant du glucose et du lactose, la bactérie utilise préférentiellement le glucose jusqu'à ce qu'il soit complètement consommé. Ensuite, après une courte latence, la croissance reprend en utilisant le lactose comme source de carbone. Cette croissance est dite diauxique ou biphasique. La latence entre les deux phases de croissance correspond au laps de temps nécessaire à l'expression des gènes codant pour les enzymes de métabolisme du lactose.

##### **- Découverte de l'opéron lactose :**

La description de l'opéron lactose a été le résultat du travail pionnier de François Jacob et Jacques Monod basé sur l'isolement de souches d'*E. coli* mutantes dans l'opéron lac et le test de complémentation pour rechercher les gènes touchés par la mutation. L'utilisation d'inducteurs gratuits essentiellement l'IPTG (isopropylethiogalactoside) et de substrats gratuits essentiellement l'ONPG (orthonitrophényl-β-D-galactoside) (coloration jaune) et le X-Gal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside) (précipité de coloration bleu) a apporté beaucoup d'aide parce qu'elle permet



**Figure 34** : Double croissance de l'*E. coli*



d'isoler la propriété inducteur de la propriété substrat (le lactose est un inducteur et substrat au même temps). Deux importants mutants ont été isolés:

- Les mutants constitutifs : expriment l'opéron Lac indépendamment de l'inducteur. Donnent une coloration bleue (présence de  $\beta$ -galactosidase) dans un milieu de culture qui

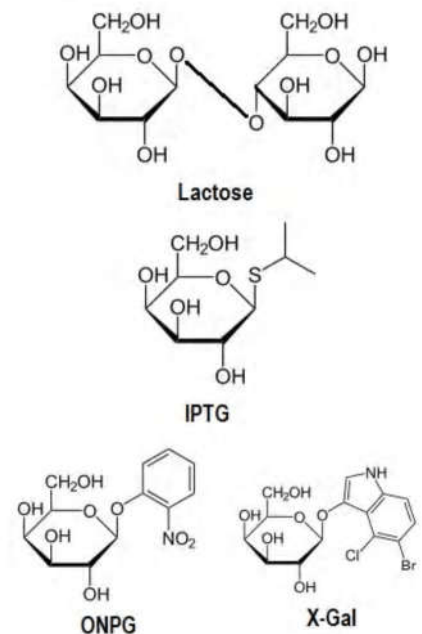
contient le glycérol (source de carbone) et le X-Gal, en présence et en absence de l'IPTG.

- Les mutant non-inductibles: pas d'expression de l'opéron Lac indépendamment de l'inducteur. Pas de coloration bleue (absence de la  $\beta$ -galactosidase) Dans un milieu de culture qui contient le glycérol (source de carbone) et le X-Gal, en présence et en absence de l'IPTG.

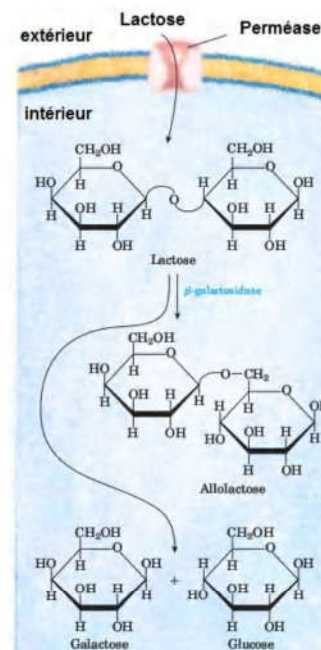
La création de diploïdes partiels entre ces mutants et la souche sauvage a permis de proposer le modèle d'opéron où un répresseur (facteur trans "de l'autre côté") contrôle l'expression d'un groupe de gènes en se fixant sur un opérateur (élément cis "du même côté"). Jacob et Monod pensaient que le répresseur est un ARN (reconnaissance spécifique de l'opérateur), les travaux d'autres chercheurs en utilisant l'IPTG radioactif ont permis d'isoler et de découvrir la nature protéique du répresseur.

**- Organisation de l'opéron Lac :**

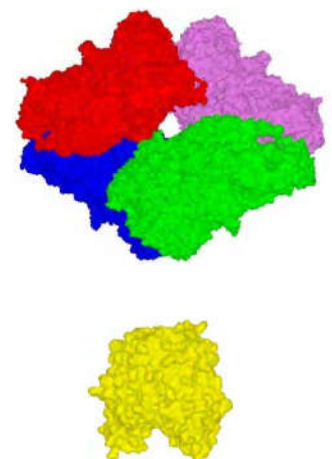
L'opéron lactose est composé de trois gènes structuraux : lacZ ( $\beta$ -Galactosidase : un tétramère de presque 500 kDa, hydrolyse le lactose en galactose et glucose), lacY ( $\beta$ -galactoside perméase : protéine membranaire de 30kDa, transporte le lactose à l'intérieure de la cellule), lacA ( $\beta$ -thiogalactoside acétyltransférase : transfert d'un groupement acétyle de l'acétyl-CoA aux  $\beta$ -galactosides). Ces trois gènes structuraux sont transcrits en une seule unité de transcription lacZYA sous forme d'un seul ARNm polycistronique.



**Figure 35 :** un inducteur et un substrat gratuits



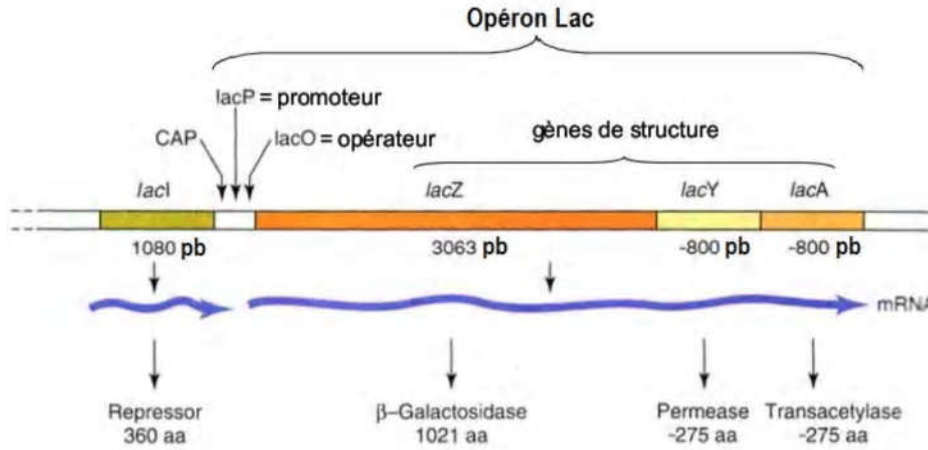
**Figure 37:** Fonctions de la  $\beta$ -galactosidase et la perméase



**Figure 36:** Structure de la  $\beta$ -galactosidase et de la perméase



Les gènes structuraux sont précédés (en amont) par un promoteur unique, qui lui-même est précédé par un gène régulateur lacI qui code pour un répresseur.



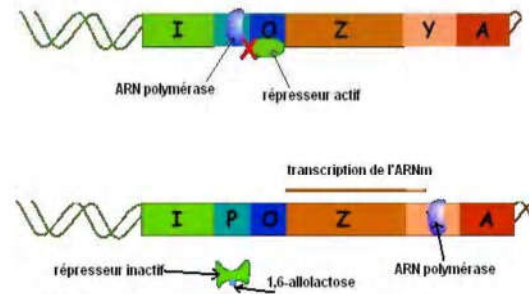
**Figure 38 :** Organisation de l'opéron Lac

**- Contrôle négatif :**

**\* Mécanisme :**

Le répresseur lac est le produit du gène de régulation lacI, il est transcrit continuellement à un taux faible (40 molécule par cellule). Le répresseur est actif et peut se fixer sur le site opérateur ce qui empêche l'ARN polymérase d'avancer dans la transcription.

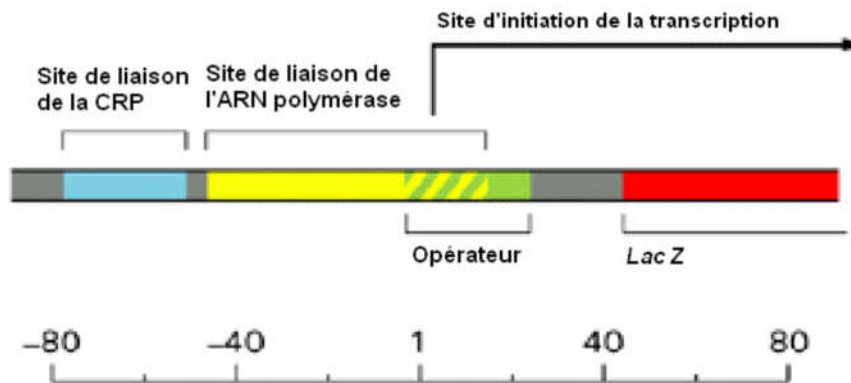
En présence du lactose, les perméases existantes permettent son absorption et la β-galactosidase catalyse la conversion de certaines molécules en 1,6-allolactose (produit secondaire de l'action de la β-galactosidase). L'allolactose joue le rôle d'inducteur, sa fixation sur le répresseur conduit à des modifications de sa conformation spatiale, réduisant ainsi son affinité à l'opérateur. Le détachement du répresseur permet l'initiation de la transcription par l'ARN polymérase.



**Figure 39 :** contrôle négatif de l'opéron Lac

**\* Site de fixation des différentes composantes :**

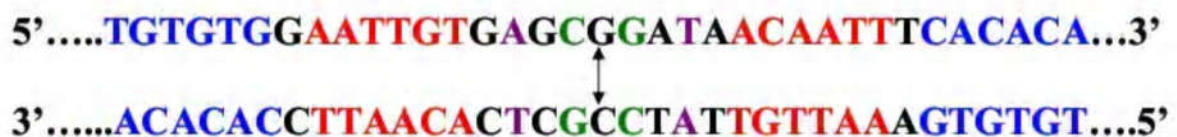
Le site de fixation de l'ARN polymérase couvre une région de -55 à +20, le site de fixation du répresseur (O1) s'étend de -5 à +21. Donc le répresseur occupe une partie du site de fixation de la polymérase, ce qui n'empêche pas sa fixation (en parle même d'augmentation de l'affinité de la polymérase avec l'ADN de 100 fois) mais bloque son initiation de la transcription. L'induction (désactivation du répresseur) permet un démarrage immédiat de la transcription, puisque la polymérase est déjà présente sur le site d'initiation. La CRP se fixe en amont (-72 à -50) du site de fixation de la polymérase.



**Figure 40** : Sites de fixation des régulateurs de l'opéron Lac

**\* Séquence reconnue par le répresseur Lac:**

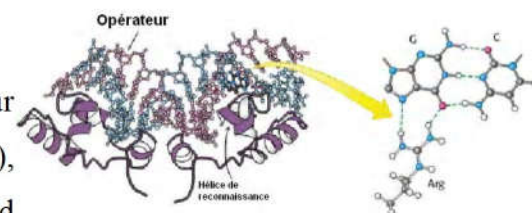
Les répétitions palindromiques jouent un rôle important dans la dimérisation. La proximité de deux séquences cis reconnues par le répresseur et la disposition inversée d'une séquence vis-à-vis de l'autre permettent le positionnement idéal des domaines de liaison à l'ADN et de dimérisation pour former un dimère qui possède plus d'affinité à l'ADN que la forme monomérique. Une petite variation dans la position de ces séquences répétées peut conduire à l'impossibilité de fixation du répresseur.



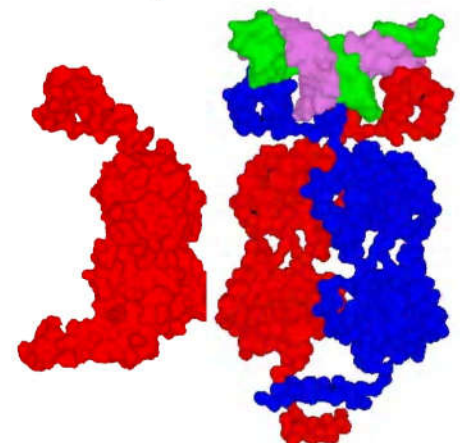
**Figure 41** : répétition palindromique de l'opérateur lac.

**\* Monomère et dimère répresseur Lac:**

La région N-terminale de type Hélice-tour-hélice du répresseur lac constitue le domaine de liaison à l'ADN (résidus 1 à 58), l'interaction se fait donc entre l'hélice de reconnaissance et le grand sillon de l'ADN qui comporte la séquence reconnue. La région cœur (résidus 62 à 330) forme le domaine de fixation du ligand (inducteur) et le domaine de dimérisation. La fixation d'une molécule inducteur (allolactose) sur chaque monomère répresseur conduit à des changements de conformation qui seront répercutés sur le positionnement exact des deux domaines de liaison à l'ADN rendant leur fixation simultanée impossible, ce qui réduit dramatiquement l'affinité du dimère à l'ADN. La région C-terminale (résidus 334 à 360) forme le domaine de tétramérisation (oligomérisation).



**Figure 42** : Interaction répresseur-opérateur lac.



**Figure 43** : monomère et dimère répresseur lac.



**\* Le tétramère répresseur lac:**

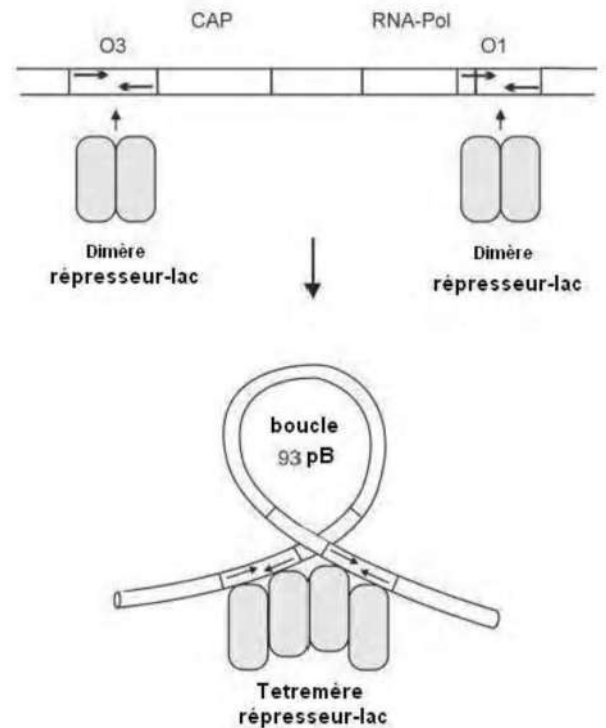
Le répresseur lac ne donne une répression maximale (1000 fois) qu'en travaillant sous forme de tétramère, où le domaine C-terminal (résidus 340 à 360) de chaque monomère participe à la tétramérisation (oligomérisation). La formation de cette structure tétramérique implique la présence de plus d'un site de fixation palindromique (un site opérateur ne fixe pas plus de deux monomères).

En plus de l'opérateur principal O1, le répresseur lac possède deux autres sites de fixation en amont et en aval. O2 centré à +410 à l'intérieur de la séquence codante, et O3 centré à -83 (en amont du site de fixation de la CRP). L'élimination de l'un des deux opérateurs (2 ou 3) conduit à la réduction de la répression de 2 à 4 fois, la délétions des deux opérateurs (2 et 3) au même temps conduit à une chute de la répression de 100 fois.

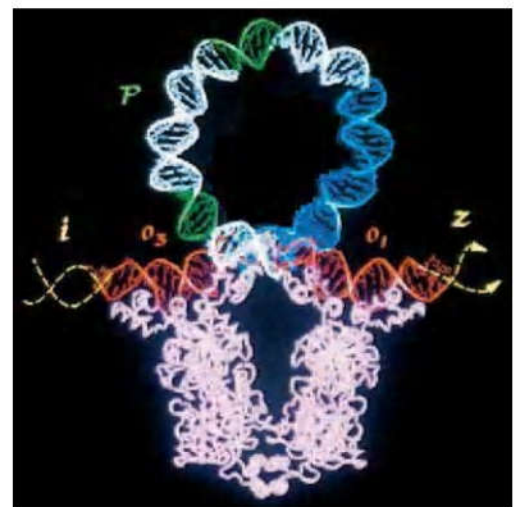
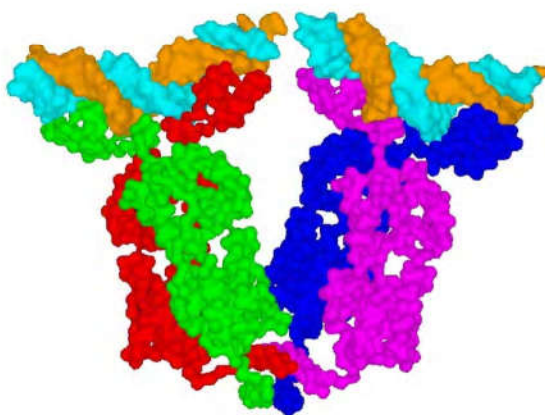
La formation du tétramère implique la courbure du fragment d'ADN qui sépare les deux opérateurs impliqués, ce qui donne une boucle de 93pb (entre -82 et +11) dont les deux extrémités sont liées par le tétramère répresseur.



**Figure 44 :** les trois sites opérateurs.



**Figure 45 :** boucle 93 pb.



**Figure 46 :** Tétramère répresseur lac.

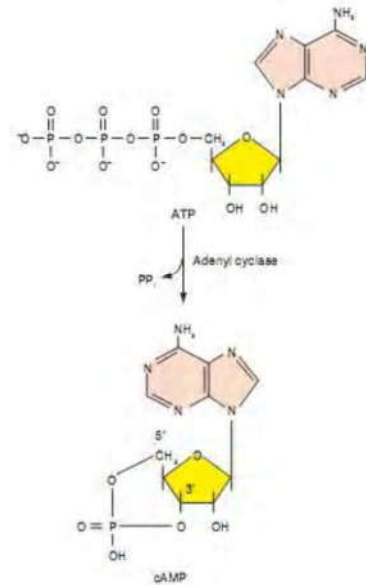
**- contrôle positif :**

Le contrôle négatif seul n'est pas suffisant lorsque la bactérie se trouve en présence du lactose et du glucose. Elle choisit d'utiliser le glucose (le sucre le plus simple), donc l'opéron lac doit rester réprimé en présence des deux sucres en même temps.

Le promoteur de l'opéron lac ne fait pas partie des promoteurs puissants (pas efficace), l'ARN polymérase ne peut pas initier une transcription efficace à partir de ce promoteur sans l'aide d'une protéine accessoire la CRP (cAMP receptor protein ou CAP catabolite activator protein).

**\*AMPc signale de carence chez les bactéries :**

L'AMPc est un signal de carence chez E. coli, son taux augmente avec la diminution du taux du glucose. Un mécanisme complexe permet l'activation d'une enzyme, l'adénylate cyclase, qui transforme l'ATP en AMPc en l'absence du glucose. En présence du glucose l'enzyme est inactive et le taux de l'AMPc chute.

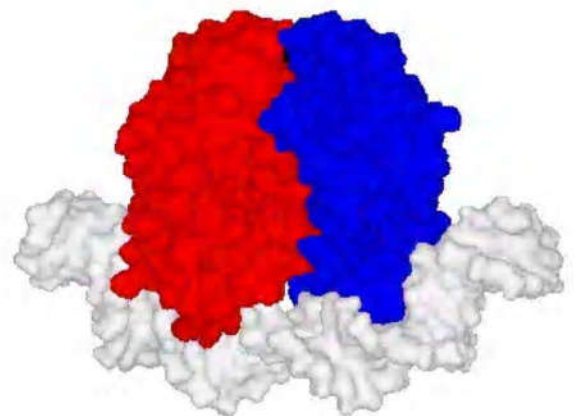


**Figure 47** : AMP cyclique.

**\* la CRP se fixe en dimère :**

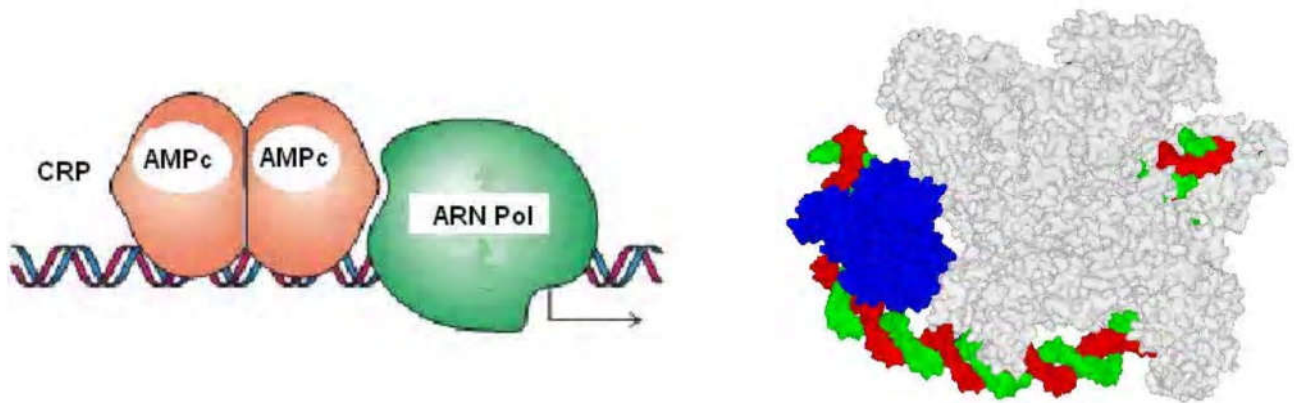
Comme le répresseur lac la CRP (homodimère de deux unités de 22.5 kDa ) possède un domaine de liaison à l'ADN HTH, un domaine de fixation du ligand et un domaine de dimérisation. La protéine activée par l'AMPc forme un dimère et se lie à une répétition palindromique de -50 à -72.

La fixation en dimère induit une inclinaison de l'ADN (courbure) de 90°. En plus, la CRP interagit, via son domaine de transactivation, avec l'ARN polymérase (sous-unité  $\alpha$ ). Le résultat de la fixation de la CRP est la stimulation de la transcription de l'ordre de 50 fois.



**Figure 48** : Dimère CRP.

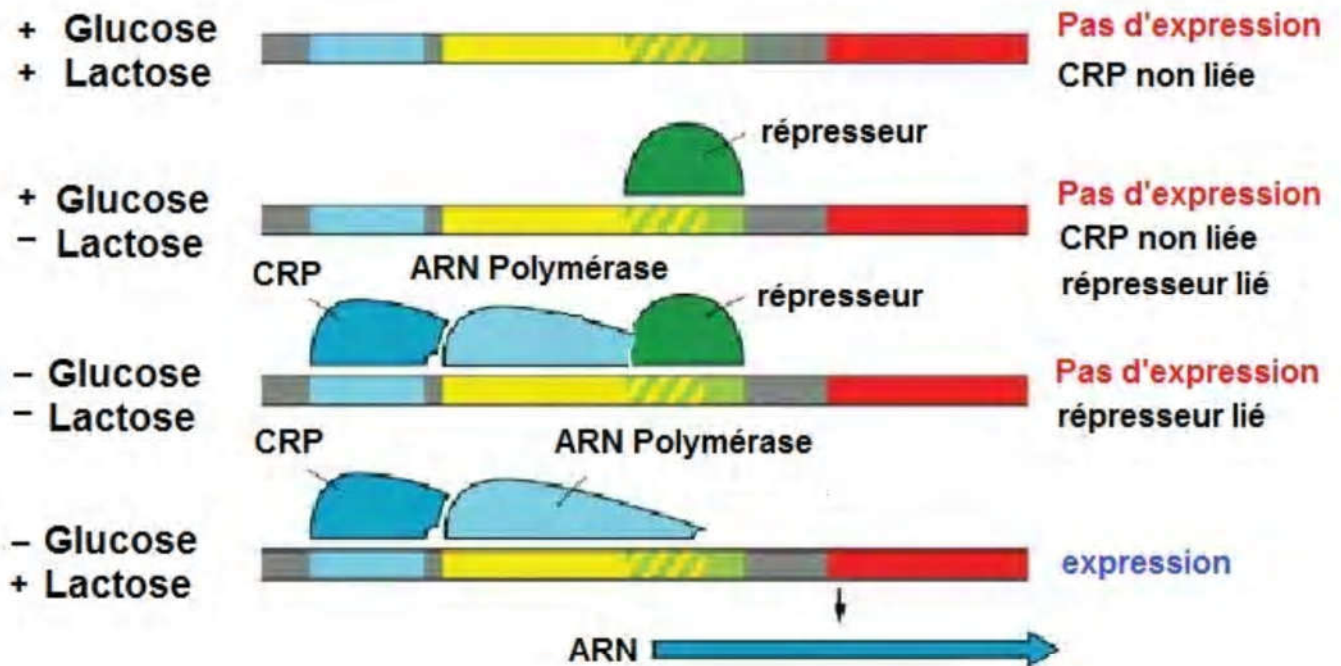




**Figure 49 :** Interaction CRP-ARN polymérase.

**- mécanisme du double contrôle :**

Une transcription optimale ne peut avoir lieu que si deux conditions sont réunies : la présence du lactose (désactivation du répresseur) et l'absence du glucose (activation de la CRP par l'AMPc). Toutes les autres combinaisons ne permettent pas une expression optimale de l'opéron lac.



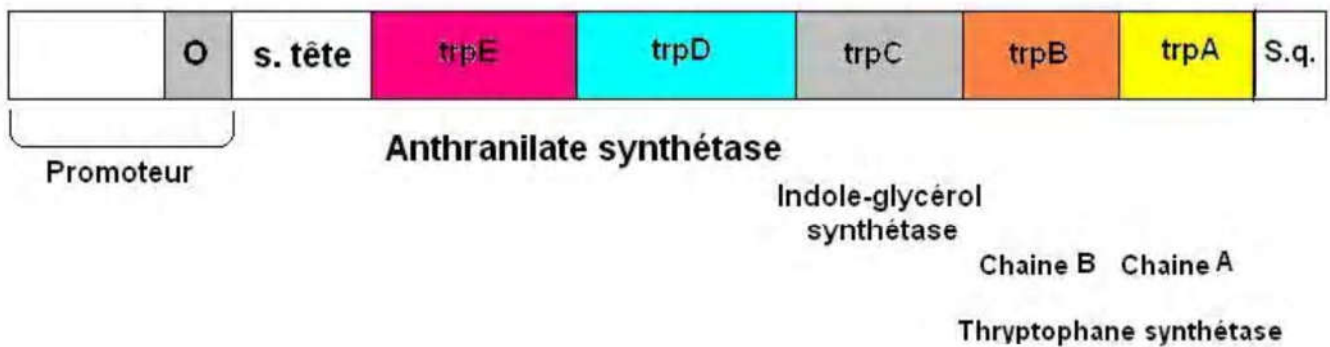
**Figure 50 :** Mécanisme du double contrôle de l'opéron lac.

**b) Répression de l'opéron tryptophane :**

De la même manière qu'avec la régulation de la transcription des gènes entrant dans le catabolisme du lactose. La bactérie régule l'expression des gènes entrant dans la synthèse de certains acides aminés en fonction de leur disponibilité dans le milieu. L'opéron tryptophane est sous un contrôle négatif exercé par un répresseur.

**\* Structure de l'opéron trp :**

Dans l'opéron trp, cinq gènes structuraux, qui codent pour les enzymes participant à la synthèse du tryptophane: trpE, trpD, trpC, trpB, et trpA, sont placés en aval d'un seul promoteur. Ce dernier contient un site opérateur.



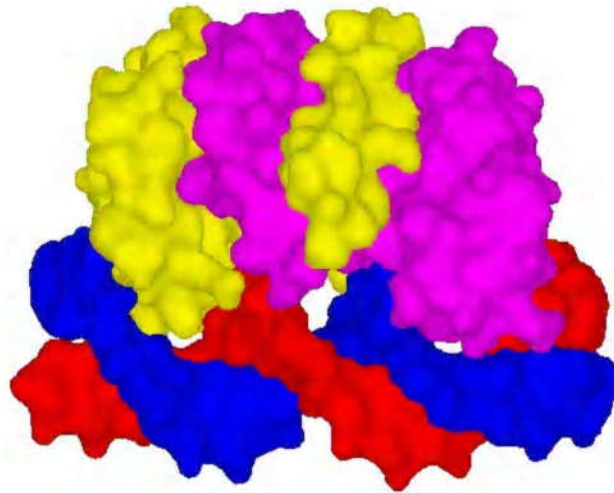
**Figure 51 :** Organisation de l'opéron trp.

**\* Répresseur de l'opéron tryptophane :**

Le répresseur trp est une petite protéine produite d'un opéron distinct, trpR. Comme pour le répresseur lac, le répresseur trp possède un domaine de liaison à l'ADN HTH, un domaine de dimérisation et un domaine de fixation du ligand (tryptophane). La seule différence est que la fixation du ligand active le répresseur.

Le site de fixation sur le promoteur de l'opéron trp est une répétition palindromique de 20pb, qui s'étend de -23 à -3. Cette position correspond au milieu de site de fixation de l'ARN polymérase, donc la fixation de l'un empêche la fixation de l'autre.

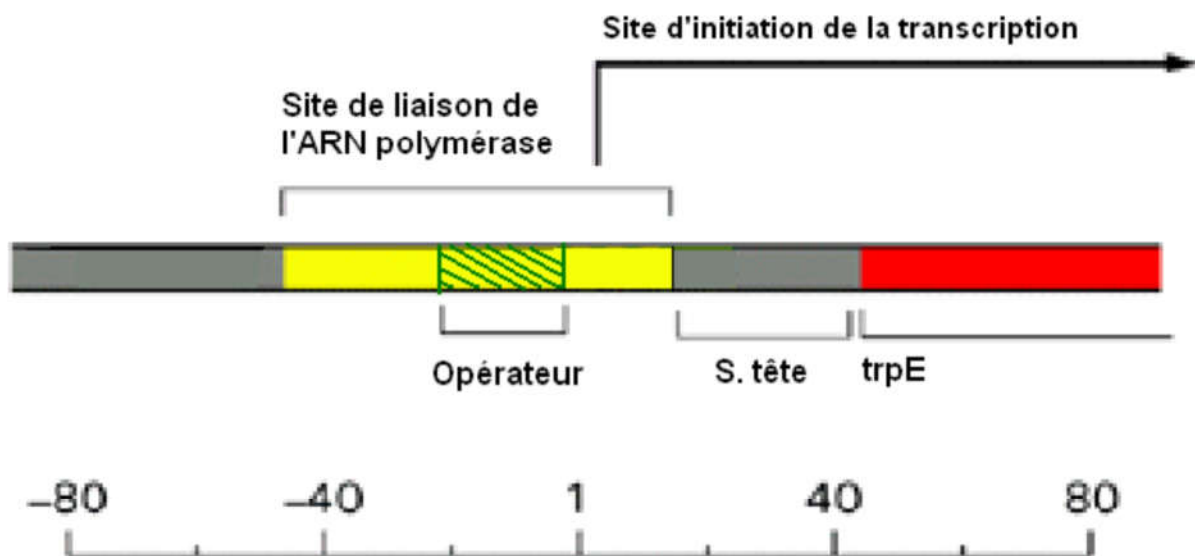




**Figure 52 :** Dimère répresseur trp-ADN.

5'.....GAACTAGTAACTAGTAC.....3'  
3'.....CTGATCAATTGATCATG.....5'

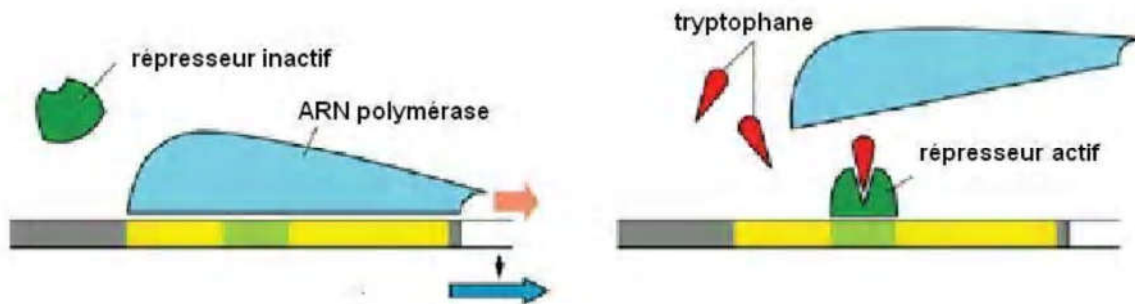
**Figure 53 :** Séquence de l'opérateur trp.



**Figure 54 :** Position de l'opérateur trp.

**\* Mécanisme de la répression :**

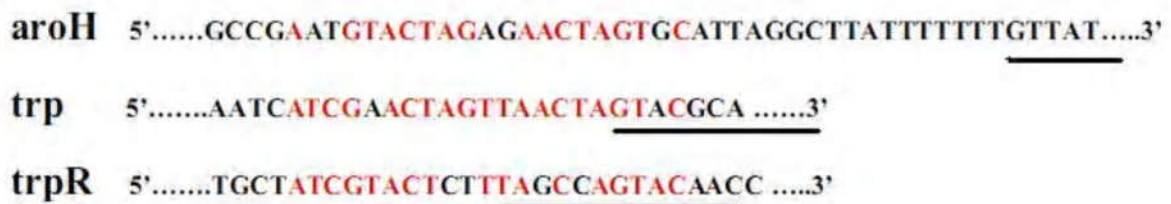
Le répresseur trp est inactif à l'état naturel, il ne peut se lier à l'ADN qu'en fixant une molécule de tryptophane (co-répresseur) et en formant un dimère, la configuration spatiale des domaines liant l'ADN sera alors correspondante à la séquence de l'opérateur. Ce qui permet la liaison et empêche l'ARN polymérase de s'attacher à l'ADN et de commencer la transcription. Dans un milieu pauvre en tryptophane le répresseur est inactif et la polymérase peut travailler normalement. Il faut noter que l'inhibition ici est de 70 fois donc moins importante que pour l'opéron lac, ce qui nécessite un autre mécanisme de contrôle (atténuation).



**Figure 55** : Mécanisme de répression de l'opéron trp.

**\* Répression de plusieurs gènes :**

Contrairement au répresseur lac, le répresseur trp possède un opérateur sur trois opérons différents. Cela permet le contrôle des trois opérons au même temps. Il existe une différence dans la séquence des trois opérateurs mais les points de contacts « paire de base de reconnaissance » sont les mêmes. Le répresseur réprime l'opéron trpR (opérateur -12+9) (auto-inhibition), l'opéron trp (opérateur -23-3), et l'opéron aroH (opérateur -49-29). Ce dernier code pour une des trois enzymes qui catalysent la réaction initiale dans la synthèse des acides aminés aromatiques. On note ici que la différence de séquence opérateur implique une différence de l'affinité donc une différence dans le taux d'inhibition.



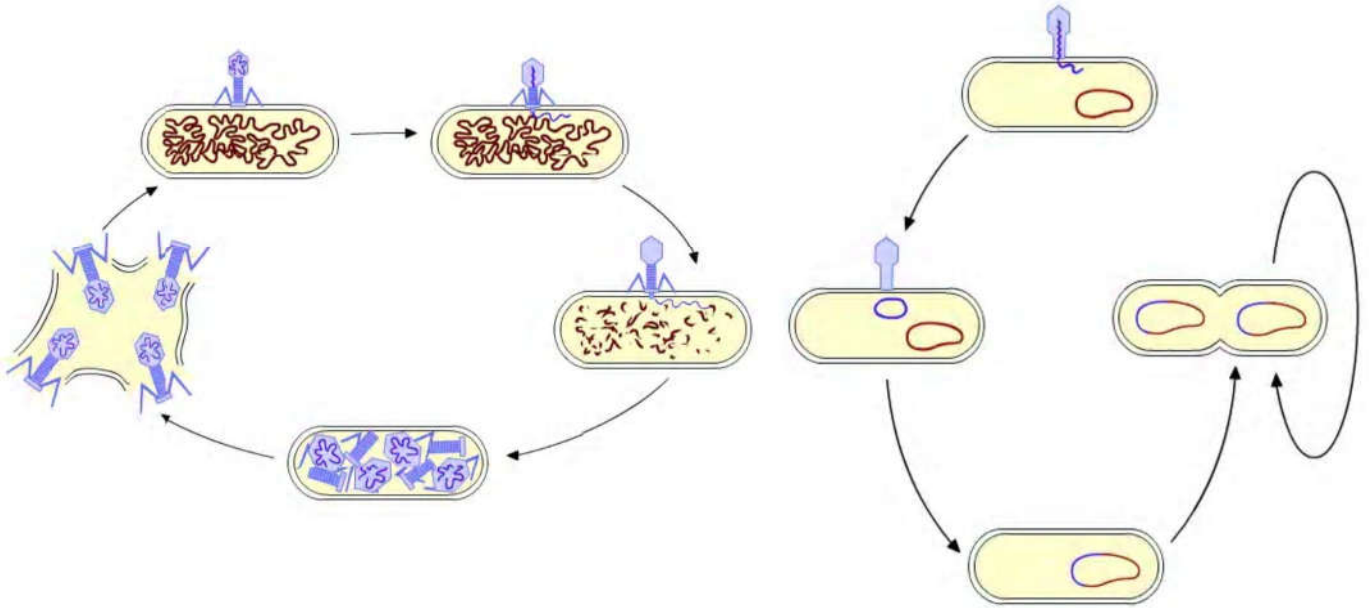
**Figure 56** : Séquence des trois différents opérateurs.



**c) Cycles lytique/lysogénique du phage  $\lambda$  : la compétition entre deux répresseurs**

**\* Cycles lytique et lysogénique :**

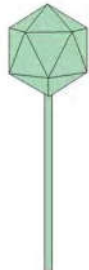
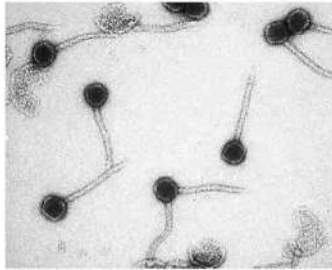
Les bactériophages existent dans le monde bactérien sous deux états distincts : en tant que phage virulent (qui se réplique dans une cellule bactérienne réceptive) ou sous forme lysogène (phage tempéré) (inséré dans le génome sous la forme d'un prophage, il devient partie intégrante du génome de l'hôte).



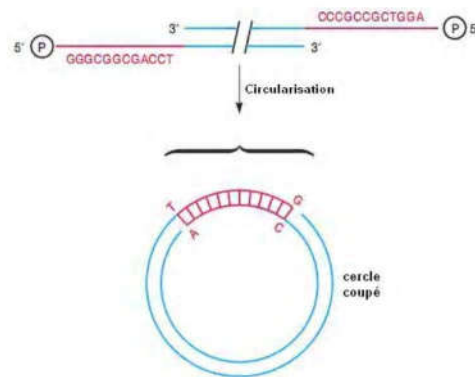
**Figure 57 :** cycles lytique et lysogénique des phages.

Tous les virus bactériophages ont un cycle lytique (infectieux) pendant lequel le virus, incapable de se reproduire par ses propres moyens, injecte son matériel génétique dans la bactérie. Grâce aux enzymes et aux ribosomes de l'hôte, le virus peut être répliqué à plus de cent exemplaires avant que l'hôte n'éclate. Mais parfois, certains bactériophages se comportent autrement. Leur matériel génétique s'intègre au chromosome de la bactérie. Une fois réalisée la fusion de l'acide nucléique du phage avec celui de la bactérie hôte, l'ADN phagique sera transmis à la génération suivante grâce à la division cellulaire. Dans un cas pour cent mille, l'ADN viral est activé et entame un cycle lytique. Donc une bactérie est dite « lysogène » si son génome inclut un génome viral. (La souche est dite « multilyso-gène » si elle en contient plusieurs). Ce caractère peut expliquer certaines caractéristiques pathogènes de certaines souches, par exemple de *Pseudomonas*. La bactérie lysogénisée est prémunie (immunisée) contre une deuxième l'infection par le même phage et certains phages apparentés.

\* **Le bactériophage  $\lambda$**  : ce bactériophage de la famille *Siphoviridae* est le phage tempéré le mieux étudié. Le virion est constitué d'une tête écosaèdre contenant un ADN double brin linéaire de 48,5 Kpb. Une fois injecté dans l'*E. coli*, l'ADN prend une forme circulaire (circularisation) suite à la liaison de ses deux extrémités monocaténaire (12b) complémentaires Cos (cohésives) (avec la contribution de la ligase bactérienne).



**Figure 58** : bactériophage  $\lambda$ .



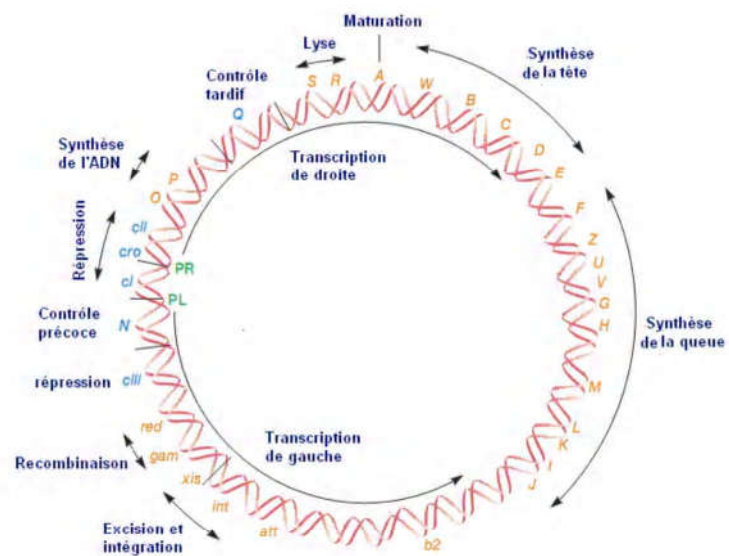
**Figure 59** : Circularisation du génome du bactériophage  $\lambda$ .

\* **Le génome du bactériophage  $\lambda$**  : le génome du phage est entièrement séquencé, les gènes sont divisés en trois catégories suivant la chronologie de leur expression :

**Gènes Précoces-immédiats** : N (facteur d'antitermination. En se fixant sur les *nut*, il permet la poursuite de la transcription des gènes précoces-tardifs) et Cro (réprime la transcription du répresseur lambda et les gènes précoces-immédiats)

**Gènes Précoces-retardés** : deux gènes impliqués dans la répllication de l'ADN, sept gènes de recombinaison (deux pour l'intégration du génome dans l'ADN bactérien et les autres pour la recombinaison pendant le cycle lytique) et trois gènes de régulation : *cII* et *cIII* (activation du promoteur  $P_{RE}$  pour la transcription du répresseur  $\lambda$ ) et *Q* facteur d'antitermination qui permet la transcription des gènes tardifs.

**Gènes Tardifs** : gènes impliqués dans la synthèse de la tête, la queue, la maturation et la lyse de la bactérie hôte.



**Figure 60** : cartographie du génome du bactériophage  $\lambda$ .



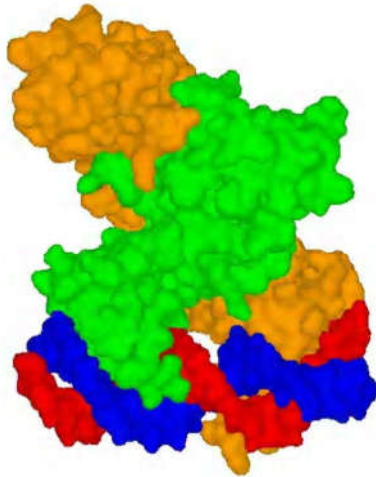
**Figure 61** : Transcription des différents groupes de gènes.



**\* Les deux répresseurs ( $\lambda$  et Cro) :**

Le répresseur  $\lambda$  est une protéine de petite taille (27KDa) constituée de deux domaines : un domaine hélice-tour-hélice N-terminal assurant la liaison à l'ADN, et un domaine C-terminal de dimérisation. Les deux domaines sont reliés par un connecteur de 40 résidus.

Le répresseur lambda assure deux fonction : l'entré et le maintien du cycle lysogénique, et l'immunité de la bactérie contre l'infection par un autre phage lambda.

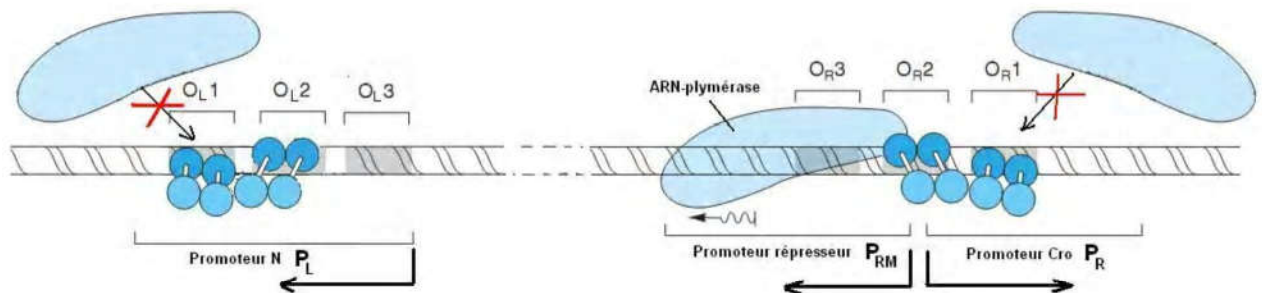


**Figure 62 :** Dimère répresseur  $\lambda$ -ADN.



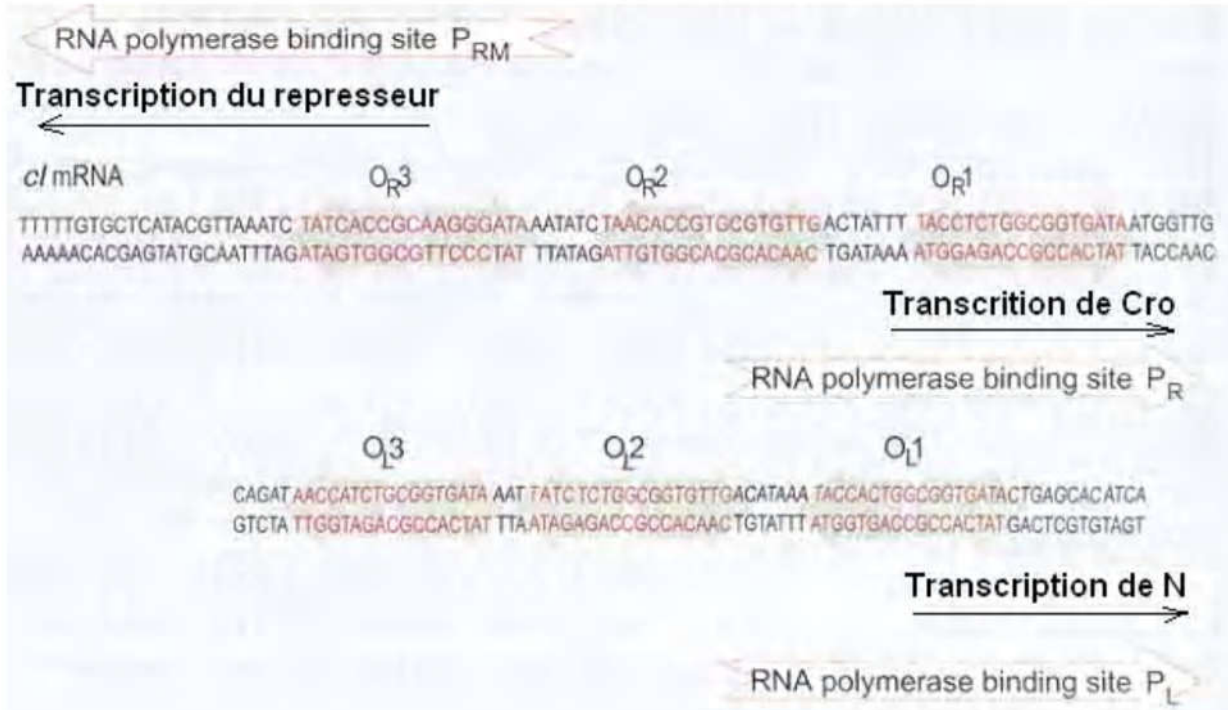
**Figure 63 :** liaison coopérative de deux dimères répresseur  $\lambda$ -ADN.

La liaison coopérative du répresseur à l'opérateur  $O_L$  ( $O_{L1}$  et  $O_{L2}$ ) réprime l'expression du gène  $N$  et tous les gènes précoces de gauche. La liaison à l'opérateur  $O_R$  ( $O_{R1}$  et  $O_{R2}$ ) a deux fonctions : elle réprime l'expression du gène  $Cro$  et tous les gènes précoces de droite et, au même temps, elle active son propre expression à cause de l'interaction avec l'ARN polymérase (le répresseur est considéré comme activateur de son propre gène  $cI$ ) ce qui explique la maintenance de l'état lysogénique.



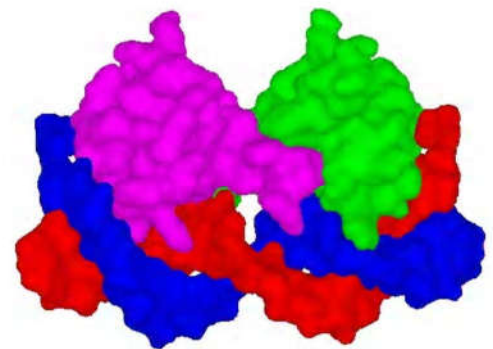
**Figure 64 :** Interaction répresseur  $\lambda$ -opérateurs-ARN polymérase.

L'affinité du répresseur à  $O_{R1}$  est 10X supérieure qu'à  $O_{R2}$  et  $O_{R3}$ . Ce qui implique une liaison préférentielle à  $O_{R1}$ , ensuite par coopération un autre dimère se lie à  $O_{R2}$ . Le répresseur ne se lie à  $O_{R3}$  sauf si sa concentration est très élevée, empêchant ainsi l'activité de l'ARN polymérase jusqu'à le retour à une concentration normale (contrôle positif à faible concentration et négatif à forte concentration).



**Figure 65 :** Séquence des opérateurs de droite et de gauche.

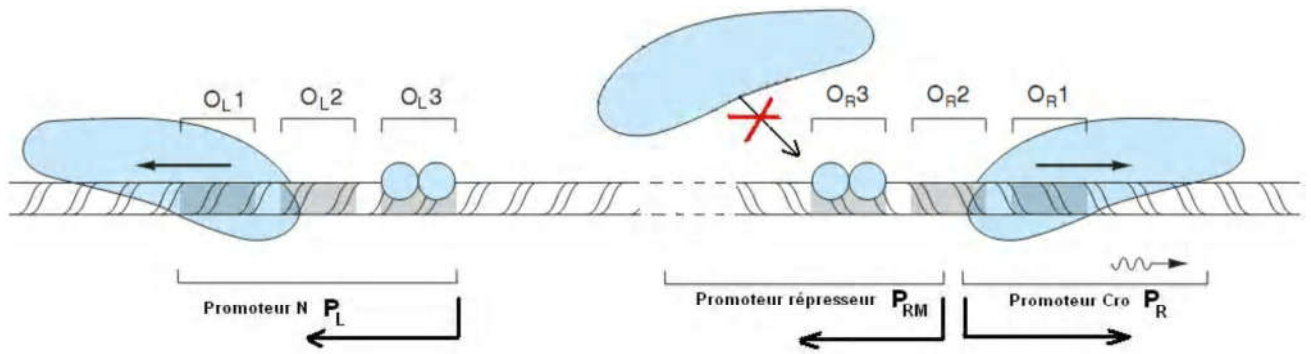
La protéine Cro est une protéine de petite taille (9KDa), elle possède un domaine de liaison à l'ADN de type Hélice-tour-hélice. Cro fonctionne en dimère et permet une fois liée aux opérateurs de bloquer la transcription du circuit de maintenance du répresseur  $\lambda$ , et à des fortes concentrations de réprimer la transcription des gènes précoces (elle-même comprise).



**Figure 66 :** Dimère Cro-ADN.

L'affinité de la protéine Cro envers les différents opérateurs est semblable à celle du répresseur  $\lambda$  mais en sens inverse :  $O_{R3} > O_{R2} > O_{R1}$ . Ce qui permet l'arrêt de la transcription du répresseur en premier lieu et si la concentration de Cro devient élevée elle arrête la transcription des autres gènes précoces en arrêtant au même temps sa propre transcription.

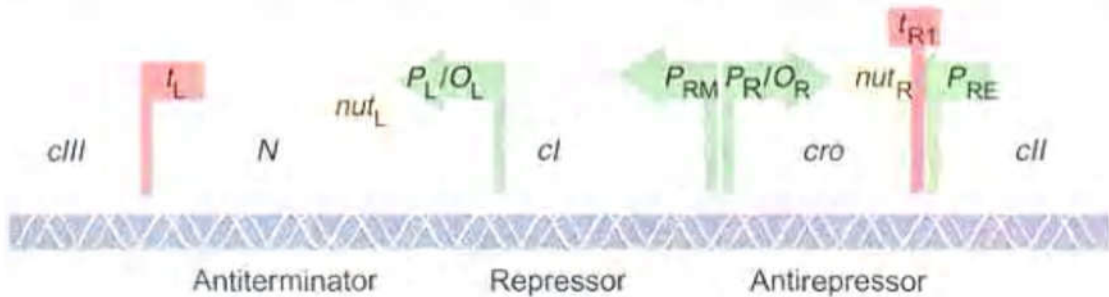




**Figure 67** : Interaction Cro-opérateurs.

**\* Les antiterminateurs (facteurs d'antitermination) :**

Découvertes au départ chez le phage  $\lambda$  puis chez les bactéries, ces protéines jouent un rôle régulateur de la transcription. Un antiterminateur reconnaît spécifiquement une séquence dans l'ARNm (différente du terminateur) (appelées *nut* (N utilisation) et *qut* (Q utilisation) chez le phage  $\lambda$ ) et en interagissant avec l'ARN polymérase il l'empêche de s'arrêter au niveau du terminateur et de se détacher, ce qui permet la poursuite de la transcription au-delà du terminateur.



**Figure 68** : Termineur-antitermineur.

**\* Le cycle lysogénique/lytique :**

Le phage  $\lambda$  possède deux cycles ; lysogénique et lytique. Le passage d'un cycle à l'autre est régulé par plusieurs facteurs dont la bactérie elle-même. On assiste à une compétition entre les deux protéines : répresseur  $\lambda$  et anti-répresseur *Cro*, si l'expression est en faveur du  $\lambda$  le phage entame un cycle lysogénique, si elle est en faveur *Cro* le phage entame un cycle lytique.

Après injection du génome virale dans la bactérie et sa circularisation, la ARN polymérase commence la transcription des gènes précoces-immédiats et s'arrête au niveau des deux terminateurs rho dépendants  $t_L$  et  $t_{RI}$  donnant naissance à deux protéines  $N$  ( $gpN$ ) et  $Cro$  ( $gpCro$ ).

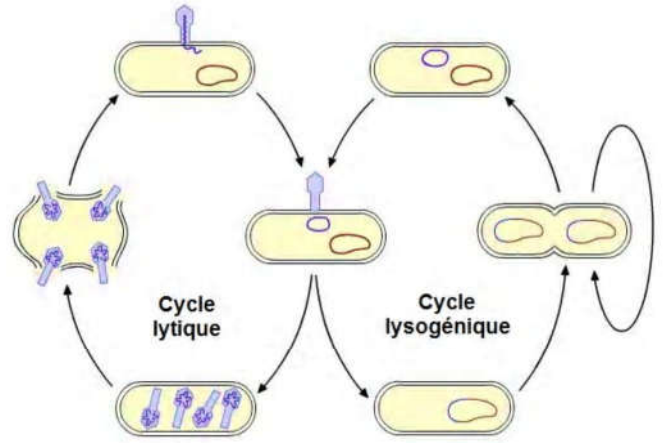


Figure 69 : Cycles lytique et lysogénique du phage  $\lambda$ .

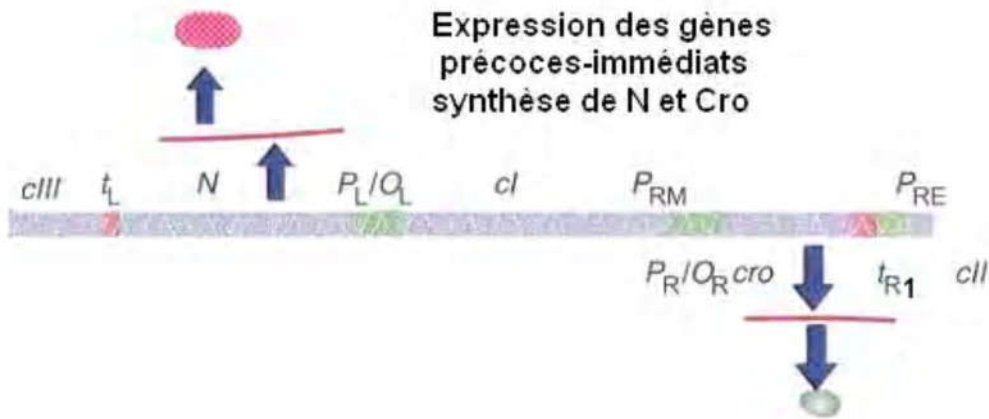


Figure 70 : Expression des gènes précoces-immédiats.

Le produit du gène  $N$  ( $gpN$  ou simplement protéine  $N$ ) est un antiterminateur. Maintenant présent, il aide la ARN polymérase à aller au-delà des deux terminateurs  $t_L$  et  $t_{RI}$  pour donner naissance aux protéines régulatrices  $cII$ ,  $cIII$  et  $Q$  (et aussi d'autres protéines de réplication de l'ADN et sa recombinaison). La transcription droite s'arrête après  $Q$  sur le terminateur  $t_{R2}$ .

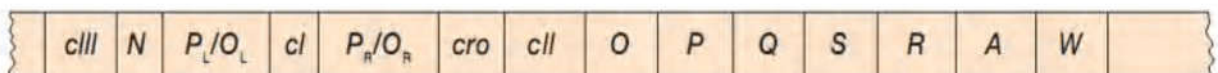
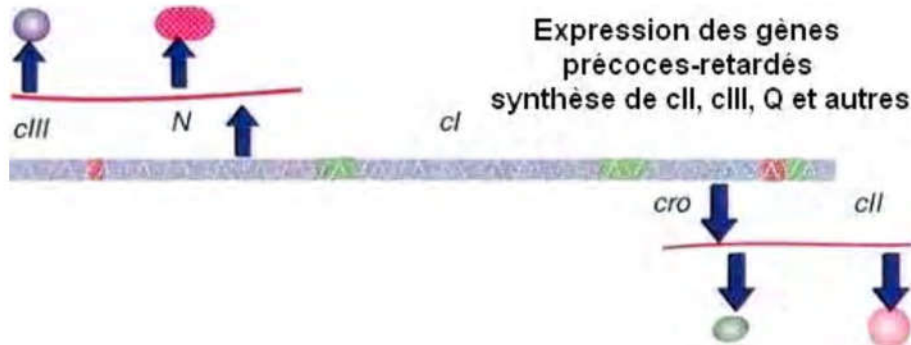
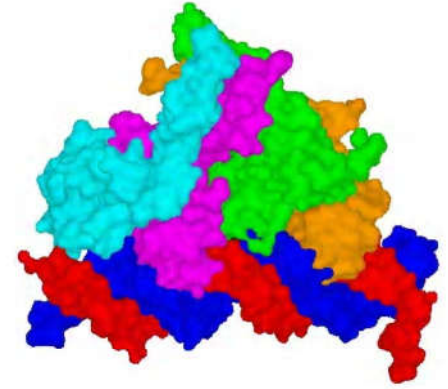


Figure 71 : Expression des gènes précoces-retardés.

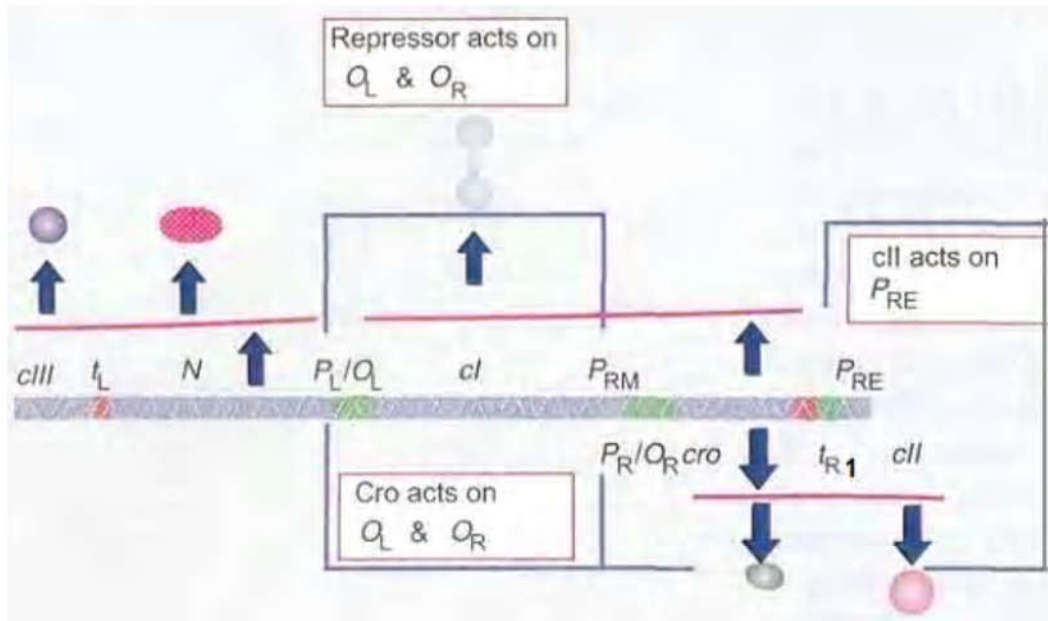


Les deux étapes précédentes sont communes aux deux cycles (lytique et lysogénique). Le phage arrive maintenant à l'étape critique qui détermine le cycle à suivre : Le produit du gène *cII* (formant un tétramère) active la transcription du gène *cI* à partir du promoteur  $P_{RE}$  (establishment) ce qui produit le répresseur lambda et en même temps réprime la transcription du gène *Cro* en agissant comme ARN-antisens. Cela favorise la synthèse du répresseur  $\lambda$ , donc le cycle sera lysogénique.



**Figure 72 :** Tétramère CII-ADN

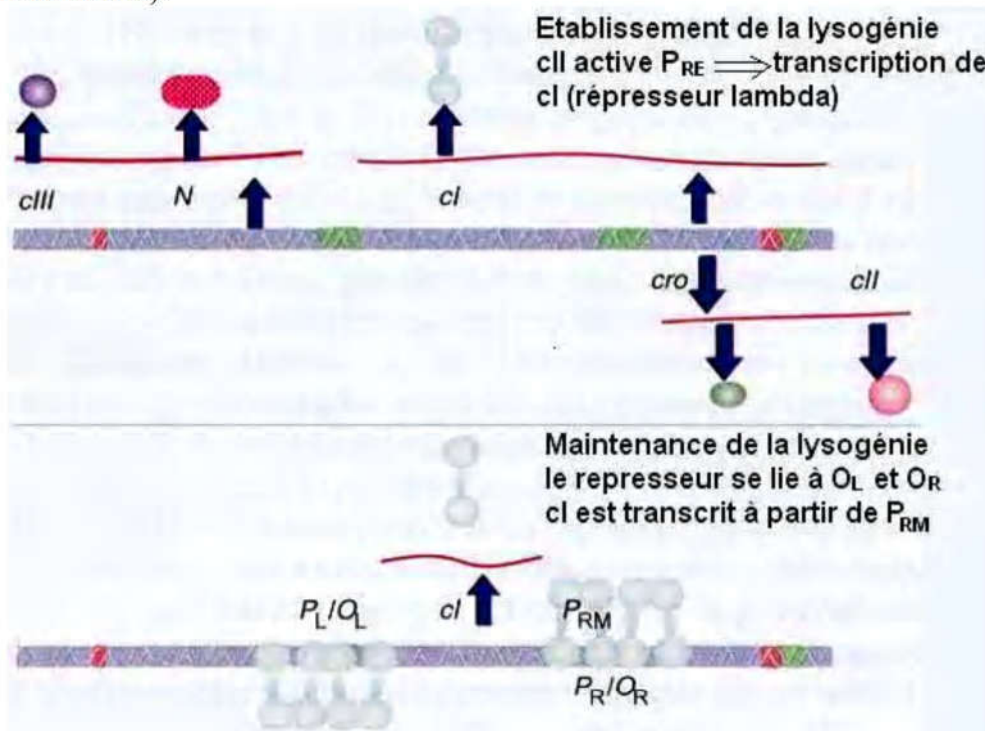
Mais cette protéine est très rapidement dégradée par une protéase bactérienne HflA (high frequency lysogenization). La protéine *cIII* se lie à *cII* et la protège partiellement de la dégradation (augmente sa demi-vie). Donc le taux de la protéine *cII* détermine le cycle à suivre. Certaines bactéries possèdent une activité protéolytique élevée donc dégradent rapidement *cII*, ces bactéries seront lysées. Les bactéries à faible activité protéolytique seront lysogénisées.



**Figure 73 :** *cII* et l'étape critique.

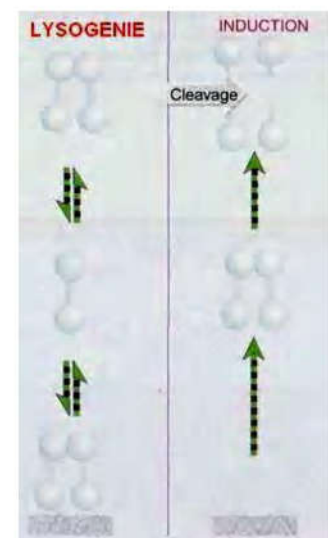
En considérant que la balance a été en faveur du cycle lysogénique (production du répresseur  $\lambda$ ). On assiste à la production du répresseur à partir du promoteur  $P_{RE}$  avec un taux élevé, lui permettant d'occuper les opérateurs 1 et 2 ( $O_{R1}$ ,  $O_{R2}$ ,  $O_{L1}$  et  $O_{L2}$ ) et arrêter l'expression des gènes précoces (*Cro*, *N*, *cII*, *cIII*, *Q* et les autres).

L'arrêt de la transcription de *cII* en plus de la dégradation de la protéine *cII* déjà produite conduisent à l'arrêt de la transcription à partir de  $P_{RE}$ , mais le répresseur déjà présent active sa propre transcription à partir de  $P_{RM}$  (maintenance) (interaction avec l'ARN polymérase sur  $O_{R2}$ ). Ce circuit de maintenance permet la production continue du répresseur  $\lambda$  et le maintien de l'état lysogène où la seule protéine produite est le répresseur  $\lambda$ . L'intégrase déjà produite avant l'arrêt de la transcription intègre le prophage dans le génome bactérien entre les opérons gal (utilisation du galactose) et bio (synthèse de la biotine).



**Figure 74** : Etablissement et maintien de la lysogénie.

Alors que chez certains phages le passage du cycle lytique est spontané (1/100000), chez le phage lambda ce passage est induit par les facteurs qui provoquent des lyses au niveau de l'ADN (UV, agents mutagènes chimiques.....). Les lésions induisent la transcription de la protéine *RecA* et son activation. *RecA* possède une fois activée une activité co-protéolytique ce qui clive le répresseur (auto-clivage assisté par *RecA*) séparant les deux domaines et réduisant l'affinité du répresseur aux opérons. Cela permet le retour de l'expression des gènes précoces (*Cro*, *N*, *cII*, *cIII*, *Q*, protéines de réplication de l'ADN) et le passage vers le cycle lytique.

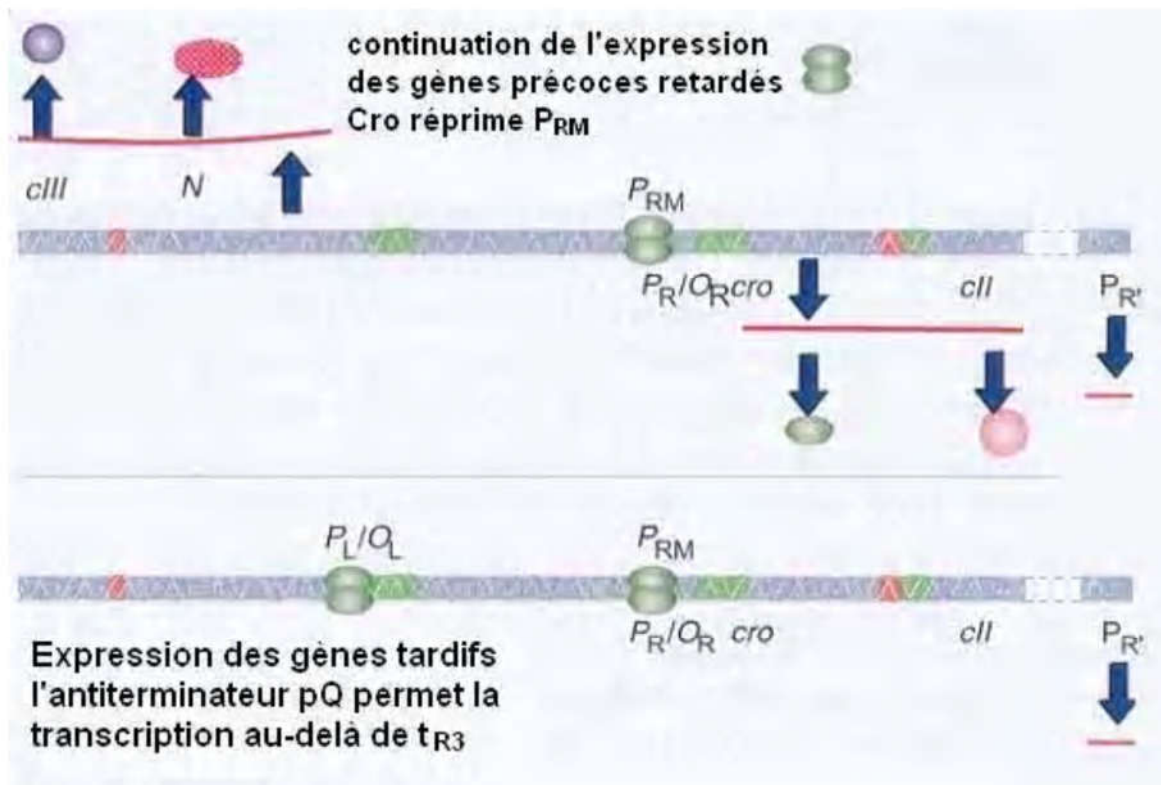


**Figure 75** : Induction du cycle lytique.



Cro se lie en premier lieu sur  $O_{R3}$  conduisant à la répression du gène  $cI$  (répresseur  $\lambda$ ), lorsque le taux de Cro devient très élevé elle se lie à  $O_{R2}$  puis  $O_{R1}$  conduisant à l'arrêt de l'expression des gènes précoces (Cro, N, cII, cIII, Q, protéines de réplication de l'ADN ....). Mais une quantité suffisante de protéines de réplication de l'ADN est déjà présente ce qui permet la production de copies du génome du phage.

Les gènes tardifs sont exprimés à partir promoteur  $P_R$  mais la transcription s'arrête au niveau d'un terminateur  $t_{R3}$  générant un ARNm de 194 B, la présence de l'antitermineur Q permet la poursuite de la transcription de tous les gènes tardifs (tête, queue, maturation, lyse). A la fin de l'assemblage des virions la paroi bactérienne est lysée et le phage se trouve dans le milieu extérieur capable d'infecter d'autres bactéries.

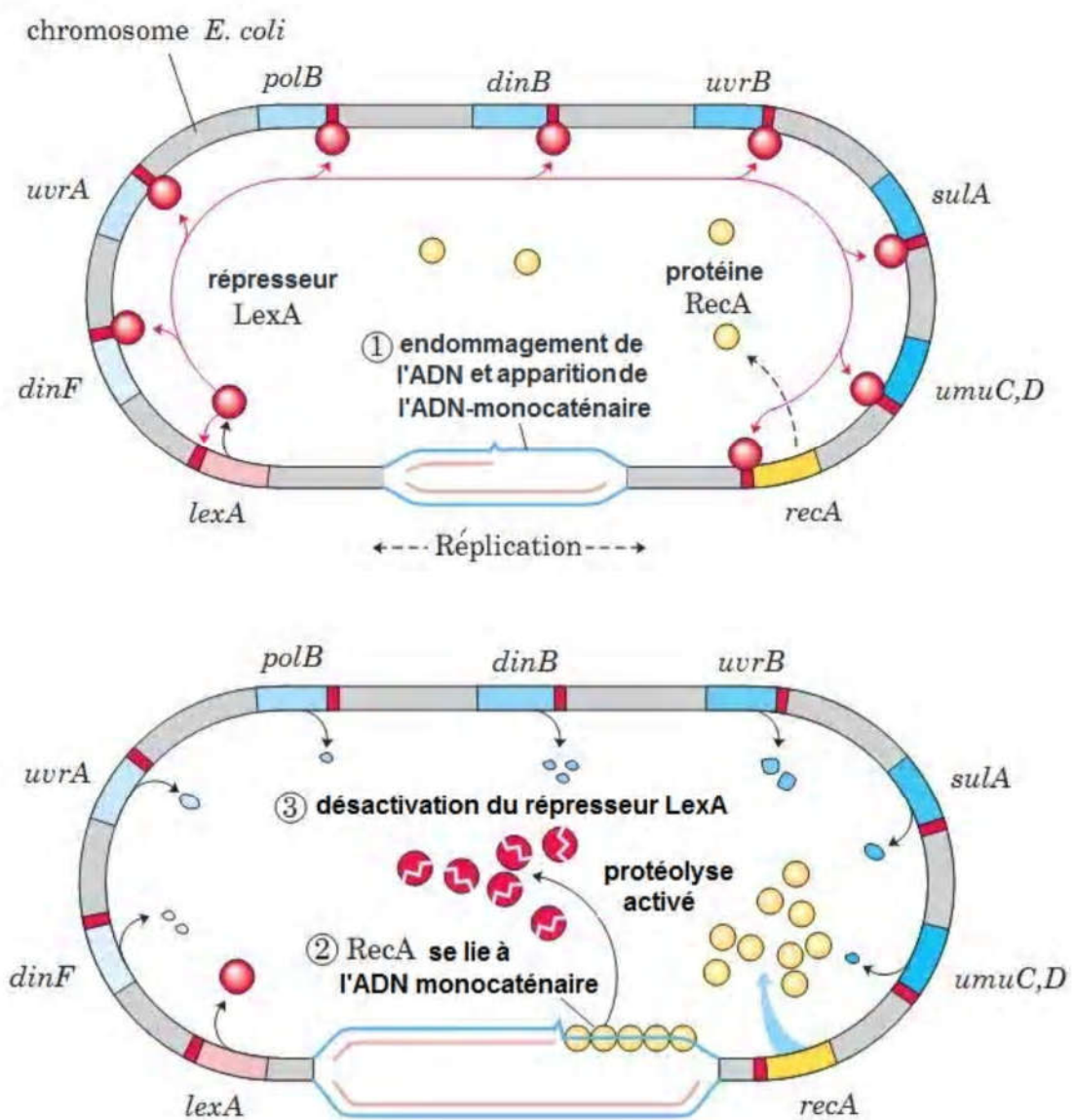


**Figure 76 :** Passage vers le cycle lytique.

**d) Le système SOS : un seul répresseur pour plusieurs gènes**

Plusieurs opérons peuvent partager le même répresseur. La désactivation de ce répresseur induit l'expression de tous ces opérons simultanément. C'est le cas du répresseur LexA qui contrôle l'expression de plusieurs gènes éloignés (différents opérons) impliqués dans la réparation de l'ADN.

Comme pour le répresseur du phage λ, LexA est inactivé par la protéine RecA (Rec=recombinaison) qui à son tour est activée par l'apparition d'ADN monocaténaire suite aux lésions provoquées par les UV et autres mutagènes (chimiques ou physiques). RecA se lie à l'ADN monocaténaire et devient active ce qui permet son interaction avec le répresseur LexA facilitant ainsi l'auto-clivage de ce dernier (activité co-protéolytique).

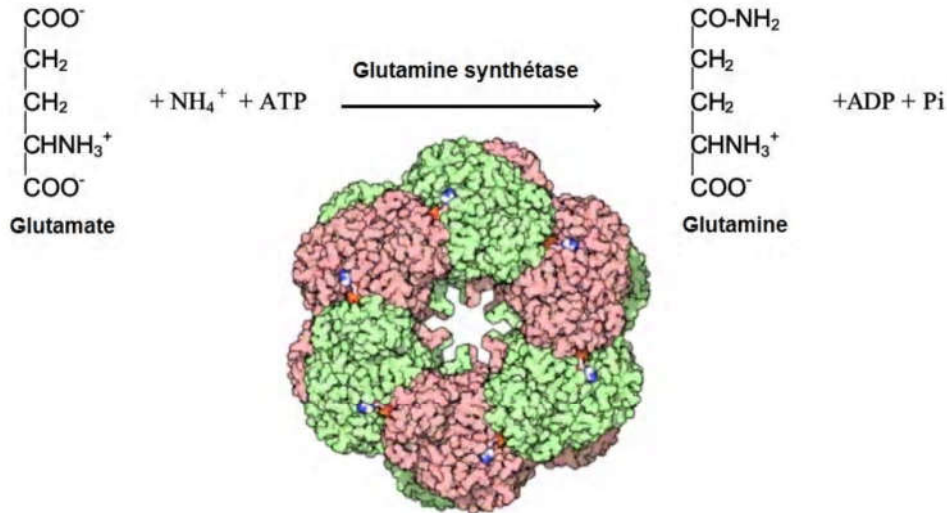


**Figure 77** : Induction des gènes SOS.



**e) Régulation de la transcription à distance: notion d'enhancer**

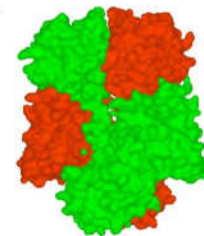
La glutamine (Gln) est un acide aminé utilisé par *E. coli* pour le stockage de l'azote en plus de son rôle dans la synthèse des protéines. Cet acide aminé est produit par la glutamine synthétase (enzyme composée de 12 sous-unités) qui utilise l'ammoniac pour ajouter un groupement amine au glutamate (Glu) (assimilation de l'ammoniac).



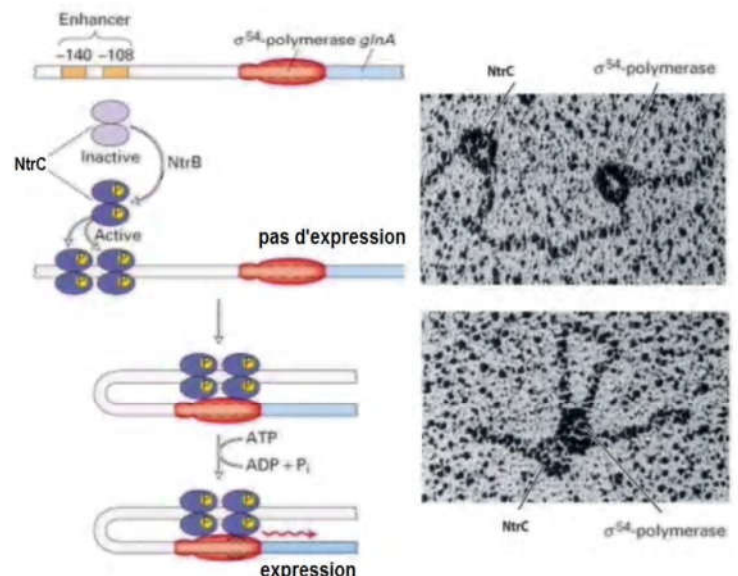
**Figure 78** : transformation du glutamate en glutamine

Le gène qui code pour la glutamine synthétase (*glnA*) est sous le contrôle d'un promoteur reconnu par un facteur  $\sigma^{54}$ , l'ARN polymérase (avec  $\sigma^{54}$ ) se lie sans pouvoir ouvrir l'ADN. Une initiation efficace de la transcription nécessite un activateur, la NtrC (nitrogen regulatory protein C), cette protéine forme un dimère inactif qui possède une faible affinité à l'ADN.

La diminution du stock d'azote chez *E. coli* (taux de la glutamine) conduit à l'autophosphorylation de la NtrB, cette dernière transfère le phosphate à l'NtrC la rendant active. La NtrC-P (active) possède une affinité très élevée pour son site de fixation sur l'ADN que la forme non phosphorylée. Deux dimères se fixent sur deux sites à plus de 100 pB en amont du promoteur *glnA* l'ADN (-108 et -140), la formation d'une boucle (visible en microscope électronique) permet l'interaction entre NtrC et ARN polymérase (facteur  $\sigma^{54}$ ), l'hydrolyse de l'ATP conduit à l'activation de l'ARN polymérase qui ouvre l'ADN et commence la transcription.



**Figure 79** : Dimère NtrC.



**Figure 80** : Activation de l'ARNpol par l'NtrC.

L'utilisation de séquences de contrôle loin du promoteur (enhancer) a été observée au départ chez les eucaryotes où elle est très courante, chez les procaryotes cela est très rare mais mérite d'être cité.

**f) Régulation de la transcription par des facteurs sigma alternatifs :**

L'ARN polymérase cœur a besoin d'un facteur sigma qui l'assiste dans l'initiation de la transcription en reconnaissant spécifiquement les séquences des promoteurs et en permettant la fixation de l'enzyme. Certaines bactéries utilisent ce facteur comme point de contrôle pour permettre la régulation de l'expression de plusieurs opérons au même temps en utilisant seulement le facteur sigma qui reconnaît exclusivement les promoteurs de ces opérons (différence entre les séquences consensus). Deux exemples permettent de comprendre ce mode de régulation : l'expression des protéines du choc thermique chez *E. coli*, et l'expression en cascade des gènes du phage SPO1.

\* **Expression des protéines du choc thermique** : le passage de la température du milieu de culture de l'*E. coli* de 37°C à 42°C induit l'expression transitoire de 17 protéines du choc thermique. Si la température augmente jusqu'à 50°C, la bactérie arrête la synthèse de toute autre protéine et se consacre à la synthèse des protéines du choc thermique. Ces protéines jouent un rôle dans la protection de la bactérie contre l'augmentation de la température. Les promoteurs des opérons codants pour ces protéines partagent des séquences consensus (-35 et -10) différentes de celles des promoteurs généraux reconnus par le facteur sigma standard ( $\sigma^{70}$ ). Un facteur  $\sigma$  alternatif  $\sigma^{32}$ , dont l'expression est induite par l'augmentation de la température, reconnaît spécifiquement les promoteurs de ces opérons et permet l'initiation de leur transcription.

Gene	Factor	-35 Sequence	Separation	-10 Sequence
<i>rpoD</i>	$\sigma^{70}$	TTGACA	16-18 bp	TATAAT
<i>rpoH</i>	$\sigma^{32}$	CCCTTGAA	13-15 bp	CCCGATNT
<i>rpoN</i>	$\sigma^{54}$	CTGGNA	6 bp	TTGCA
<i>fliA</i>	$\sigma^{28}$ ( $\sigma^F$ )	CTAAA	15 bp	GCCGATAA
<i>sigH</i>	$\sigma^H$	AGGANPuPu	11-12 bp	GCTGAATCA

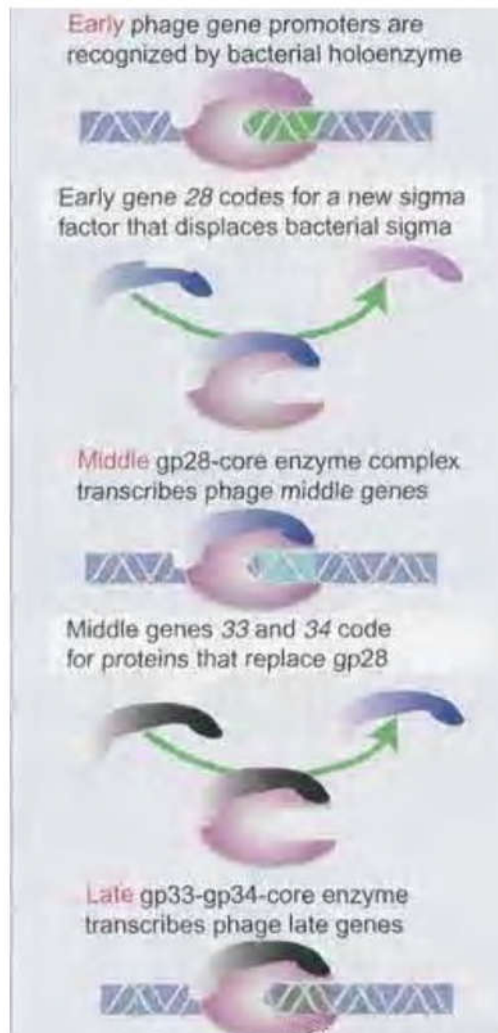
**Figure 81 :** Différentes séquences consensus pour différents facteurs sigma.



**\* Expression en cascade des gènes du phage SPO1 :**

Suite à son infection de *Bacillus subtilis*, les gènes précoces du phage SPO1 commence à être exprimés immédiatement. Après 4-5 minutes, l'expression des gènes précoces s'arrête et commence celle des gènes moyens. 8-12 minute après infection, les gènes moyens arrêtent l'expression et commence celle des gènes tardifs.

Les gènes précoces sont reconnus par la ARN polymérase de la bactérie (avec un facteur  $\sigma_{43}$ ) de la même façon que les gènes de la bactérie. Un des gènes précoce code pour un facteur sigma 28 ( $\sigma_{28}$ ) qui remplace, grâce à son affinité élevée, le  $\sigma_{43}$  sur la ARN polymérase permettant la reconnaissance des promoteurs des gènes moyens du phage et arrétant au même temps l'expression des gènes précoces et les gènes de la bactérie. Deux gènes moyens codent pour deux sous unités (33 et 34) d'un facteur sigma capable de reconnaître les promoteurs des gènes tardifs. Ce dernier facteur sigma remplace le  $\sigma_{28}$  déclenchant l'expression des gènes tardifs et arrétant l'expression des gènes moyens.



**Figure 82 :** Différents facteurs sigma exprimés en cascade.

**2- Régulation de l'élongation de la transcription par la traduction (Atténuation) :**

Ce mode de régulation n'est possible que chez les procaryotes en raison de l'absence de membrane nucléaire, ce qui permet le début de la traduction avant même la fin de la transcription.

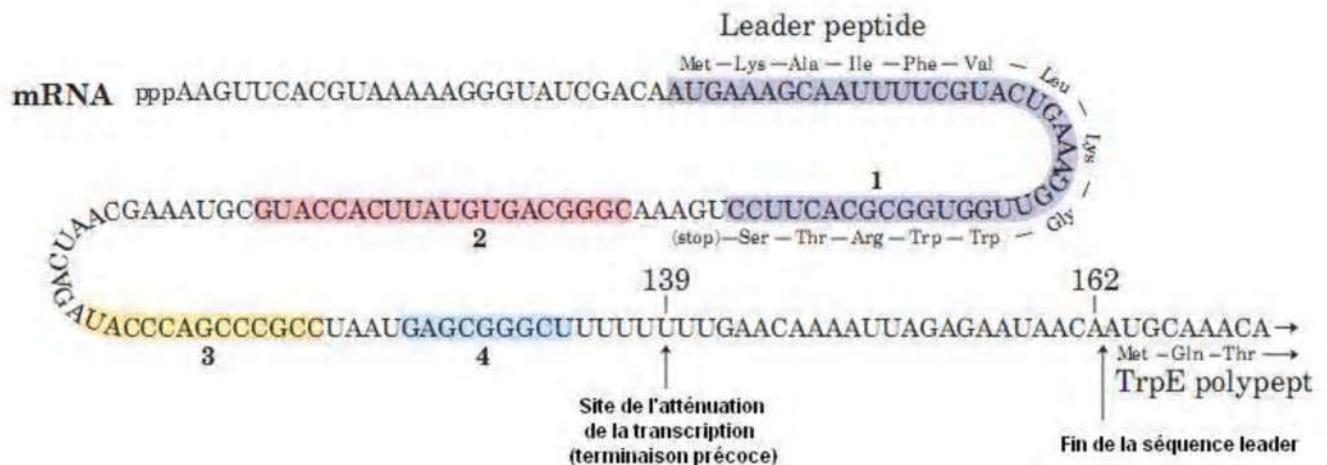
En plus du répresseur qui contrôle l'initiation de la transcription, l'opéron trp possède un autre mécanisme qui contrôle l'élongation de la transcription, l'atténuation. Le contrôle négatif par répresseur réduit le niveau de la transcription de 70 fois, l'atténuation de 8 à 10 fois. L'association des deux modes de contrôle permet un ralentissement de la transcription de 600 à 700 fois.

**\* Séquence de tête :**

L'augmentation du niveau de la transcription de l'opéron trp après délétion de la séquence qui sépare le promoteur du codon d'initiation du gène trpE (séquence de tête), en présence du répresseur actif ou inactif, est une preuve du rôle régulateur de cette séquence. Cette séquence est de 162 nucléotides, elle possède quatre régions complémentaires 1, 2, 3 et 4. Il existe trois possibilités d'appariement entre ces séquences qui donnent trois structures en épingle à cheveux : 1:2, 2:3, 3:4. L'appariement 3:4 donne un terminateur rho indépendant (appelé aussi atténuateur), l'appariement 2:3 (moins favorable que les deux autres, se forme lorsque 1 est 4 ne sont pas disponible) empêche la formation du terminateur et permet la continuation de la transcription (antiterminateur).

**\* Le peptide principal :**

La région 1 (+27 à +68) commence par un codon d'initiation (AUG) et elle code pour un peptide de 14 acides aminés. Ce peptide ne sert qu'au contrôle de la transcription il sera dégradé ensuite. Les deux codons 10 et 11 codent pour le tryptophane (UGG), ce qui donne au peptide un rôle de détecteur de la disponibilité du tryptophane dans le milieu.

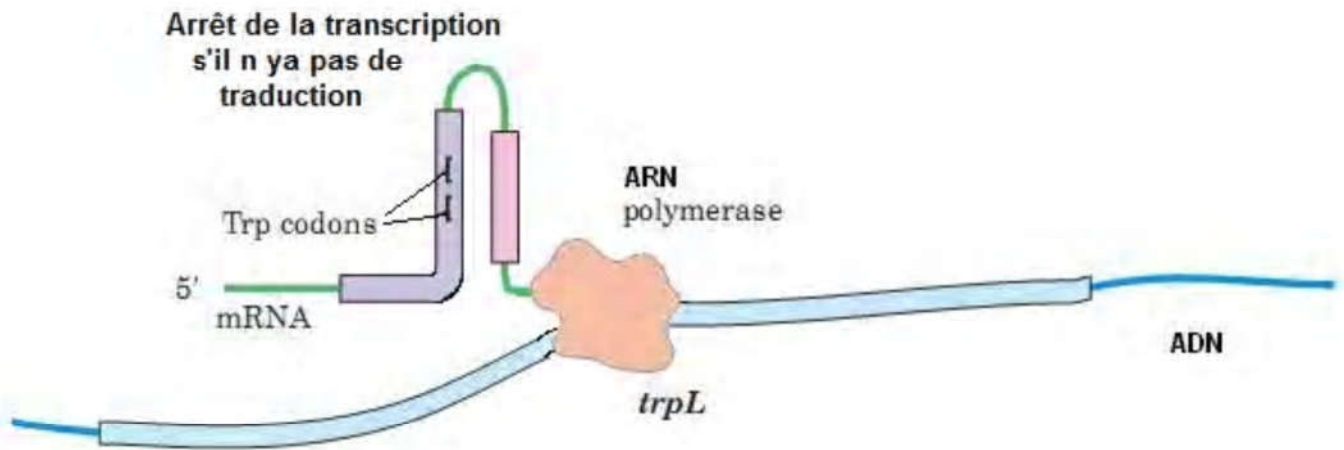


**Figure 83 :** Séquence de tête de l'opéron trp.



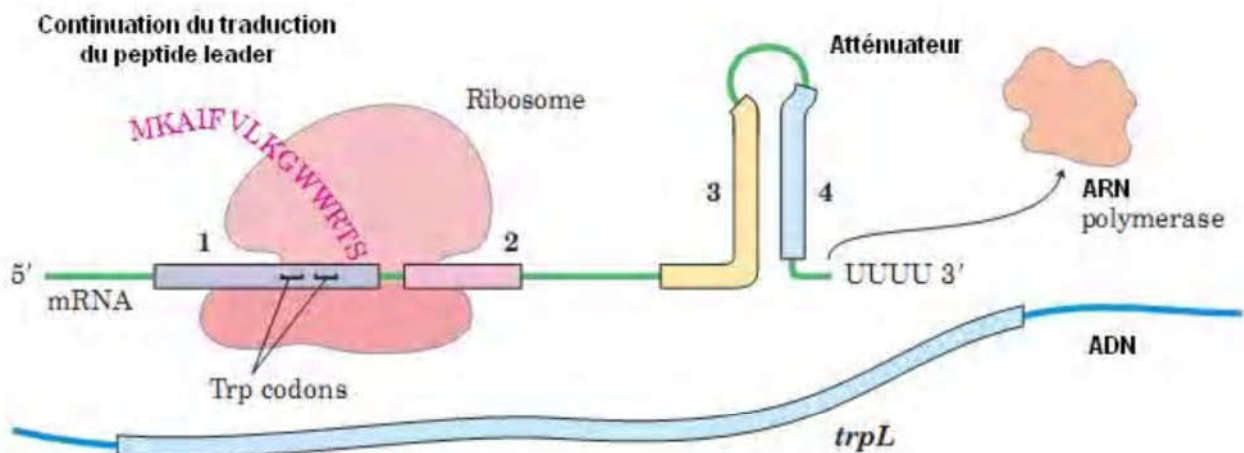
**\* Atténuation en présence du tryptophane:**

En effectuant la transcription l'ARN polymérase marque une pause après la première structure en épingle à cheveux formée entre 1 et 2. Il faut qu'un ribosome commence la traduction pour permettre à l'ARN polymérase de continuer la transcription.



**Figure 84 :** l'ARN polymérase marque une pause après la structure 1:2.

Si un ribosome commence à traduire l'ARNm, il incorpore les acides aminés et en raison de la présence de tryptophane il peut arriver jusqu'au codon stop occupant ainsi la région 1 et le début de 2. Ce qui permet à l'ARN polymérase de poursuivre la transcription et empêche l'appariement entre 2 et 3. Lorsque la transcription dépasse la région 4 le terminateur se forme (appariement 3:4) et l'ARN polymérase se détache indépendamment de rho. On obtient alors un ARNm de 140b seulement.

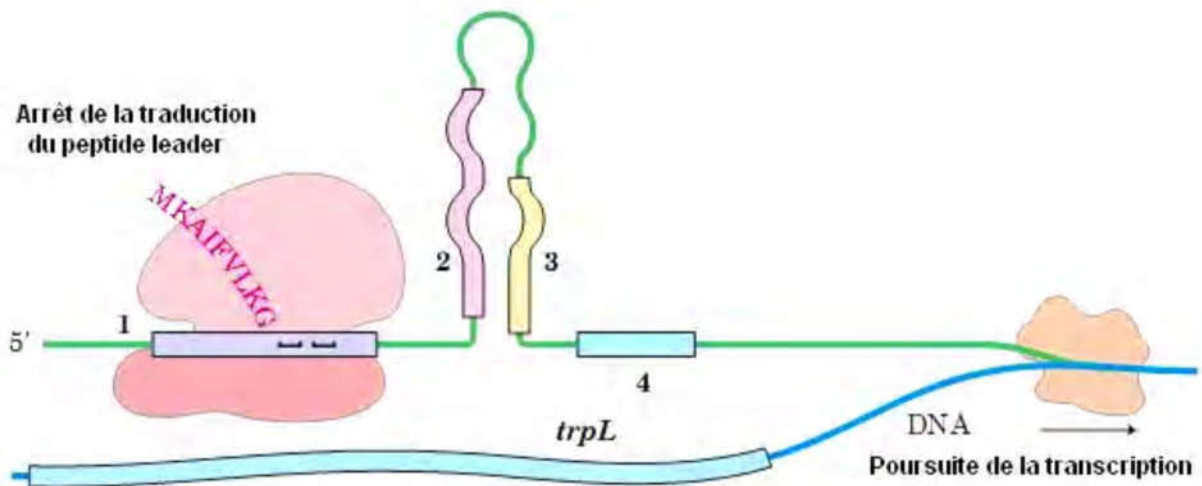


**Figure 85 :** Atténuation en présence du tryptophane.

**\* Continuation en absence du tryptophane :**

En absence du tryptophane le ribosome s'arrête au niveau des codons 10 et 11 occupant la région 1 mais laissant libre la région 2 qui peut former une épingle à cheveux avec la région 3 (antiterminateur), la structure 2:3 n'oblige pas l'ARN polymérase à marquer une pause. Donc le terminateur ne pourra pas se former et l'ARN polymérase peut continuer la transcription sans soucis.

L'atténuation existe au moins dans six opérons codants pour des enzymes impliquées dans la synthèse des acides aminés (tryptophane, histidine, phénylalanine, leucine, isoleucine, thréonine), l'opéron histidine par exemple possède un peptide de tête avec 7 codons histidine successifs (mais il n'est pas contrôlé par un répresseur).



**Figure 86 :** Poursuite de la transcription en absence du tryptophane.

**3) Régulation de la traduction par les ARN antisens :**

Les mécanismes de contrôle de l'expression génétique étudiés jusqu'ici sont orchestrés par des protéines. D'autres molécules non protéiques participent aussi à la régulation, on cite ici les ARN antisens.

Les ARN anti-sens participent à la régulation de l'expression génétique en bloquant la traduction de certains ARNm. L'hybridation entre la région d'initiation de la traduction d'un ARNm et des séquences complémentaires de l'ARN anti-sens empêche l'accès du ribosome et donc empêche la traduction. Le premier exemple est l'inhibition de la traduction de l'ARNm du gène Cro par l'ARN anti-sens issu de la transcription à partir du promoteur PRE.

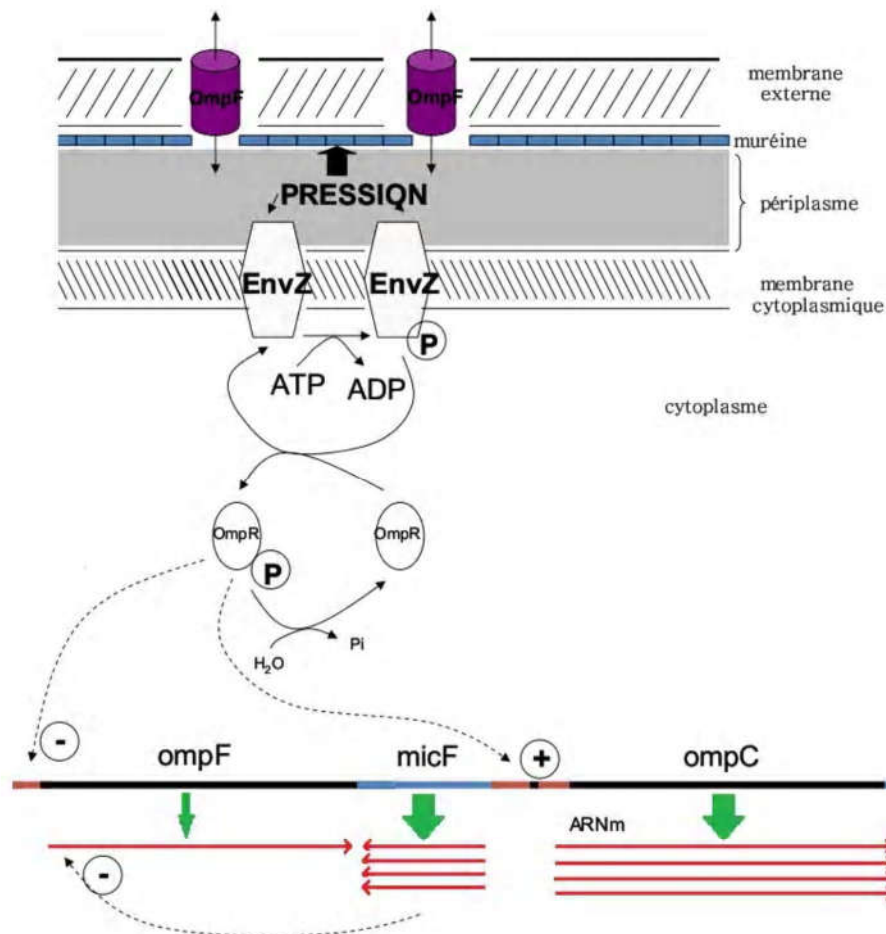
Le deuxième exemple est la régulation de l'expression des porines de la membrane externe de l'E. coli. Les deux porines les plus importantes sont la OmpF (grande taille de pores) et la OmpC (petite taille de pores). Les OmpC sont synthétisés si la bactérie se trouve dans un milieu à pression osmotique élevée ou dans un milieu toxique (empêche l'entrée des substances toxiques), les OmpF sont synthétisés dans un milieu à pression osmotique faible (permet une diffusion plus facile des substances). L'expression des deux gènes est régulée par une protéine OmpR qui à pression osmotique



faible active l'expression d'OmpF seulement et à pression osmotique élevée réprime l'expression de l'OmpF et active celle de l'OmpC.

OmpR possède des opérateurs à forte affinité (liaison avec une concentration faible) et d'autres à faible affinité au niveau du promoteur de OmpF. Les opérateurs au niveau du promoteur d'OmpC sont de faible affinité. A pression osmotique faible la concentration de la forme phosphorylée de l'OmpR est faible, la liaison se fait seulement avec les opérateurs à affinité élevée d'OmpF ce qui conduit à l'activation de la transcription de ce gène et la production des pores à grand taille. L'augmentation de la pression conduit à l'autophosphorylation de la protéine EnvZ (osmocapteur : sensible aux changements de l'osmolarité. Avec activité kinase et phosphatase) qui transfère le phosphate à OmpR augmentant la concentration de la forme phosphorylée. L'OmpR-P peut maintenant se fixer sur les opérateurs à faible affinité au niveau des promoteurs des deux gènes OmpF et OmpC réprimant la transcription du premier et activant celle du dernier ce qui donne des pores à petite taille. Lorsque la pression redevient faible l'EnvZ enlève le phosphate de l'OmpR via son activité phosphatase.

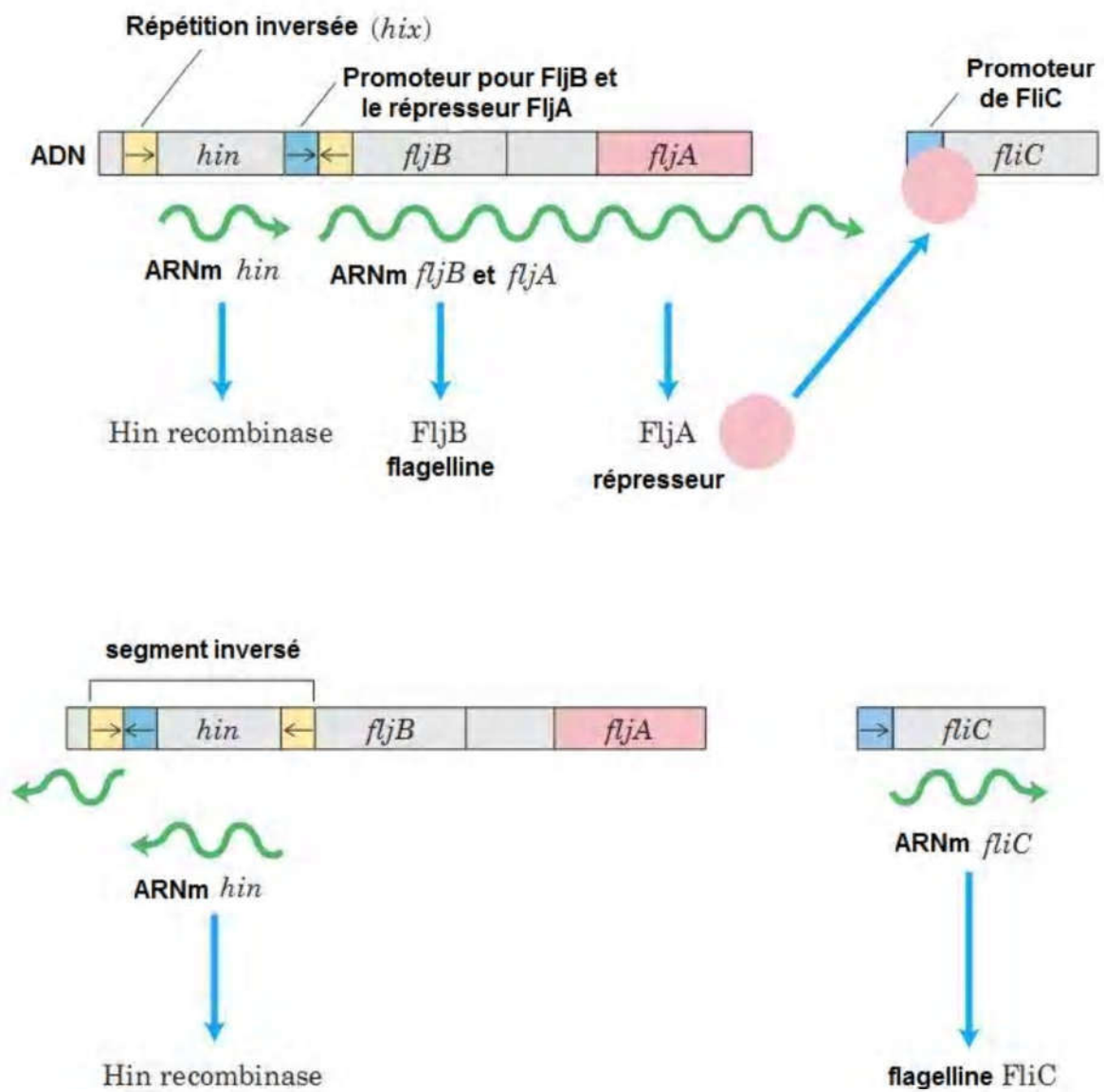
En plus, la transcription d'un autre gène micF (mRNA-interfering complementary), activée par par l'augmentation de la concentration de la forme phosphorylée de l'OmpR, produit un ARN complémentaire avec la région 5' de l'ARNm de l'OmpF. L'hybridation entre ces deux ARN empêche l'accès des ribosomes au site d'initiation de la traduction de l'ARNm de l'OmpF et inhibe sa traduction.



**Figure 87 : Régulation de l'expression des porines.**

**4- Régulation par recombinaison génétique (réarrangement génomique):**

Salmonella typhimurium utilise la recombinaison génétique spécifique au site (réarrangement) pour produire deux types de protéines de flagelle (fljB et fliC) en essayant d'échapper au système immunitaire des mammifères. Ce changement de phase a lieu chaque 1000 génération. La *hin* recombinase reconnaît une séquence de 14 paires de base (*hix*) de part et d'autre de son gène et inverse le segment. Ce qui conduit dans une situation (phase 1) à placer un promoteur en amont de deux gènes *fljB* et *fljA* (répresseur de *fliC*) conduisant à leur expression. Dans l'autre situation (phase 2) le promoteur est déplacé et les deux gènes *fljB* et *fljA* ne sont pas exprimés, seul le gène *fliC* est exprimé (pas de répression).



**Figure 88 :** Régulation de l'expression des protéines du flagelle.



