

# HARRY POTUT'RENTREE

## BIOLOGIE MOLECULAIRE



Disclaimer : fiche non complète



### MODULE 2

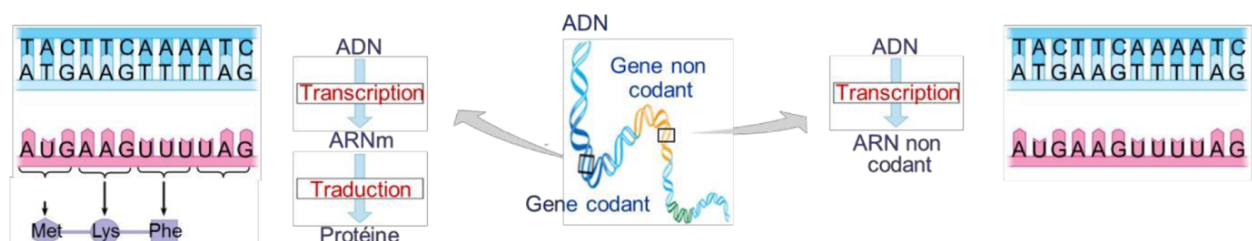
#### I - PRINCIPES GÉNÉRAUX DE L'EXPRESSION D'UN GÈNE : TRANSCRIPTION / TRADUCTION

Le matériel génétique ou génome contient les gènes → un gène contient une information.

**Gène** = enchaînement linéaire de nucléotides formant une séquence d'ADN délimitée par un signal de début "START" et par un signal de fin "STOP"

**Les gènes codants** : ils servent à la synthèse des protéines et leur séquence de désoxyribonucléotides va être tout d'abord transcrite en séquence de ribonucléotides que l'on retrouvera dans l'ARNm, puis traduite en une séquence d'acides aminés pour former une protéine.  
Ils vont subir dans leur expression deux étapes : une étape de transcription, puis une étape de traduction

**Les gènes non codants** : ils servent uniquement à la synthèse d'ARNs non codant comme les ARNs ribosomiaux, les ARNs de transfert, les petits ARNs nucléaires ou nucléolaires.  
Ils vont donc être uniquement transcrits. Il n'y aura pas dans son expression d'étape de traduction.

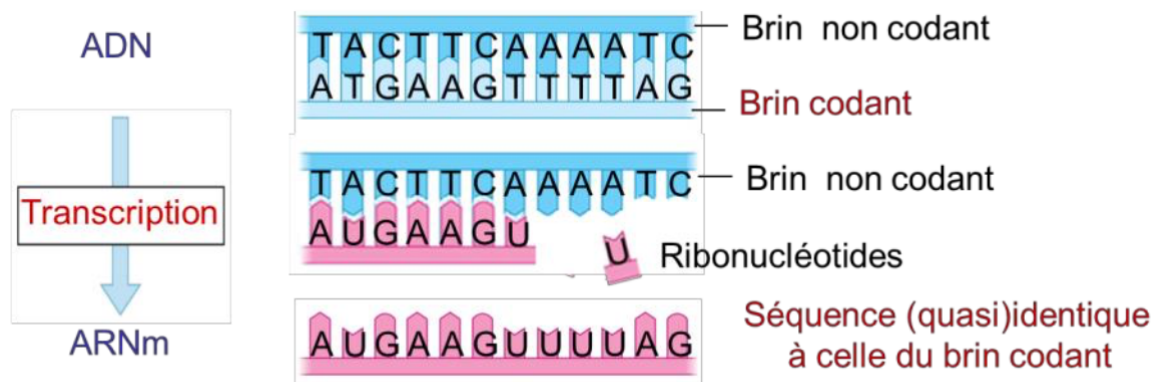


## 1<sup>ère</sup> étape : La transcription

Il s'agit de retranscrire la séquence de désoxyribonucléotides du gène en une séquence de **ribonucléotides** qui sera retrouvée dans l'**ARN messenger**.

Le **brin codant** contient l'**information** qui doit être retranscrite dans l'ARN messenger.

Le **brin non codant** sert de **matrice** pour transcrire quasiment à l'identique l'information du brin codant dans l'ARN messenger.



La transcription est assurée par une **ARN polymérase** : enzyme capable de synthétiser de l'ARN à partir d'ADN.

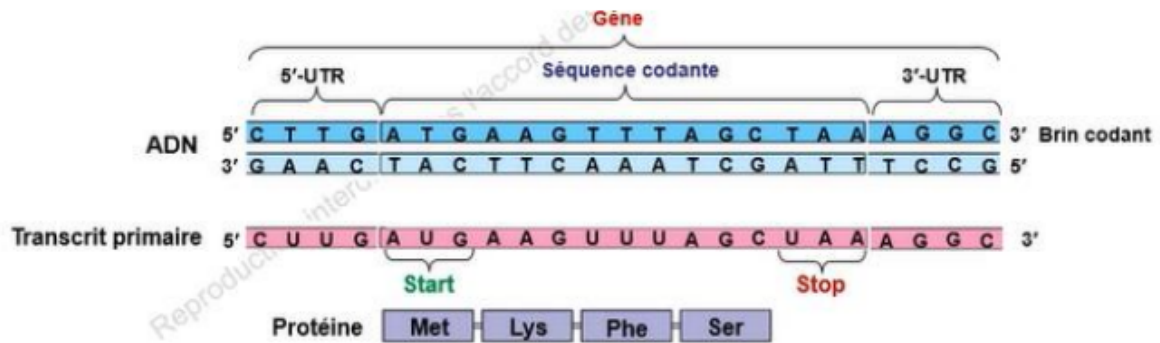
Le transcrit obtenu est appelé **transcrit primaire** et sera utilisé tel quel chez les procaryotes, mais devra subir des étapes de maturation chez les eucaryotes. (vous verrez tout ça plus tard)

Un gène contient aussi des séquences dites « non codantes ».

Ces séquences encadrent en 5' et en 3' la séquence codante. Cependant elles ne seront pas traduites.

Elles sont appelées « **5'-UTR et 3'-UTR (Untranslated)** ».

L'**ARN polymérase** débute la transcription en amont de la séquence codante et l'achève en aval : elle produit un ARN plus grand que celui qui correspond à la séquence codante du gène.



## 2<sup>ème</sup> étape : La traduction

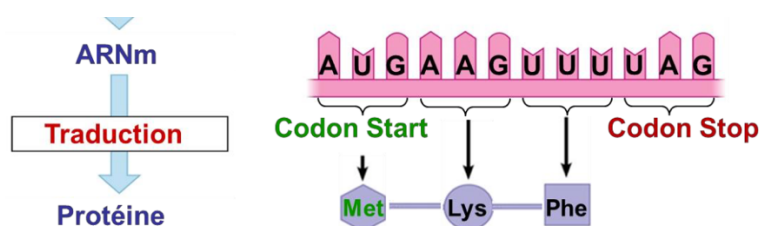
L'expression d'un gène se termine par sa traduction dans le cytoplasme.

Il s'agit de quoi?

**La traduction consiste à décoder le message de l'ARNm afin de former une protéine**

Les **ribonucléotides** sont lus trois par trois → chaque triplet de nucléotides forme un **codon**.

La traduction débute au niveau d'un codon appelé **codon START** pour s'achever au niveau d'un **codon STOP**.



Elle repose sur le code génétique qui indique à quel acide aminé correspond chaque codon de l'ARNm (donc 3 nucléotides codent un acide aminé)

C'est bien le code génétique qui permet de déchiffrer l'information de l'ARNm.

		Second Letter					
		T	C	A	G		
First Letter	T	TTT } TTC } Phe TTA } TTG } Leu	TCT } TCC } TCA } Ser TCG }	TAT } TAC } Tyr TAA } Stop TAG } Stop	TGT } TGC } Cys TGA } Stop TGG } Trp	Third Letter	T
	C	CTT } CTC } CTA } Leu CTG }	CCT } CCC } CCA } Pro CCG }	CAT } CAC } His CAA } Gln CAG }	CGT } CGC } CGA } Arg CGG }		C
	A	ATT } ATC } Ile ATA } ATG } Met	ACT } ACC } ACA } Thr ACG }	AAT } AAC } Asn AAA } Lys AAG }	AGT } AGC } Ser AGA } Arg AGG }		A
	G	GTT } GTC } Val GTA } GTG }	GCT } GCC } GCA } Ala GCG }	GAT } GAC } Asp GAA } Glu GAG }	GGT } GGC } GGA } Gly GGG }		G

Ce tableau liste toutes les possibilités de codons différents et donne leur correspondance.

Au total, il existe  $4^3 = 64$  combinaisons de trois nucléotides pouvant former un codon.

Parmi ces 64 combinaisons, quatre d'entre elles restent assez particulières :

Le **codon AUG** qui code pour la **méthionine** initie toujours la traduction et joue donc le **rôle de codon START**. Il peut également se trouver ailleurs dans la séquence d'un ARNm où il prendra alors le même sens.

Trois codons : **UAA ; UAG ; UGA** ne codent pour aucun acide aminé et indiquent la fin de la traduction de la protéine (**codon STOP**).

Le **code génétique** possède 4 caractéristiques majeures :

<b>Quasi-universel</b>	La plupart des espèces vivantes utilisent la même correspondance entre codons et acides aminés. Cependant, il subsiste de rares exceptions, telles que les mitochondries qui reposent sur le sens de quelques codons.
<b>Non-chevauchant</b>	Chaque nucléotide de l'ARNm ne peut appartenir qu'à un seul codon. L'ARNm est ainsi décodé selon un cadre de lecture fixe et précis.

<b>Non-ambigu</b>	Un codon donné correspond toujours au même acide aminé.
<b>Dégénéré</b>	Comme il existe un excès de codons par rapport au nombre d'acides aminés, la majorité des acides aminés sont spécifiés par plusieurs codons différents SAUF pour la méthionine et le tryptophane.

Toutefois, il existe différentes **mutations** du code génétique qui amènent diverses conséquences :

- 1- La substitution est le **remplacement** d'un nucléotide dans un codon par un autre.  
Cela peut amener à 3 types de mutations :

La mutation synonyme	La mutation faux-sens	La mutation non-sens
Est dite neutre car elle ne change ni l'acide aminé codé ni la protéine synthétisée après la traduction (ex : GGT changé en GGU = Gly dans les deux cas).	Elle change le sens du codon et l'acide aminé dans la séquence de la protéine.	Elle crée un codon qui interrompt la traduction. En effet, elle remplace le codon de départ par un codon STOP prématuré. La protéine est alors dite « tronquée ».

- 2 - Il existe également des **insertions** et des **délétions** dans le code génétique : il s'agit d'ajouter ou supprimer des nucléotides.

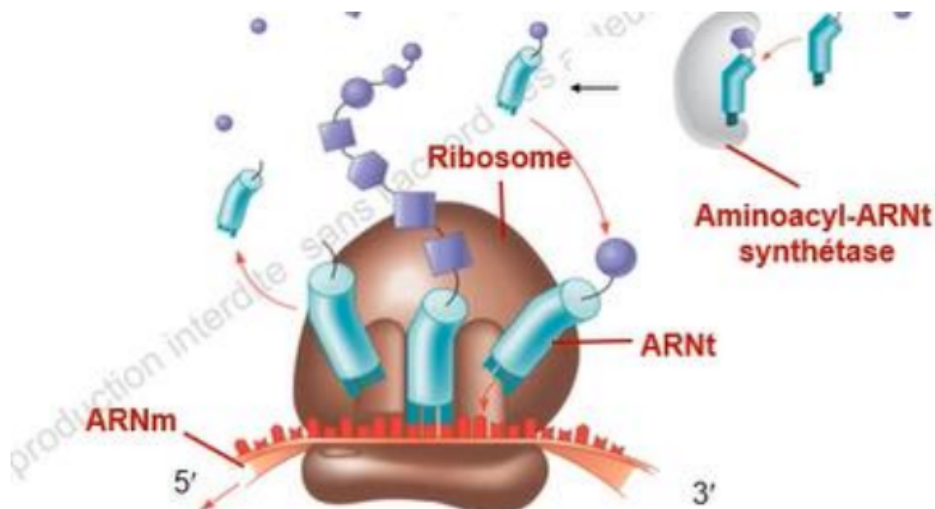
(Rappel : un codon = 3 nucléotides)

Si multiple de 3	Si non multiple de 3
Mutation non décalante : car ajout ou retrait d'un ou plusieurs codon(s) sans effet sur les codons précédant. Le cadre de lecture est respecté. La conséquence est un ajout ou un retrait d'acides aminés sur la protéine.	Le cadre de lecture de l'ARNm sera décalé d'un ou deux nucléotides. Il y aura la présence de faux sens multiples voire la modification de la position du Codon Stop ce qui pourra selon les cas aboutir à une protéine raccourcie ou allongée.

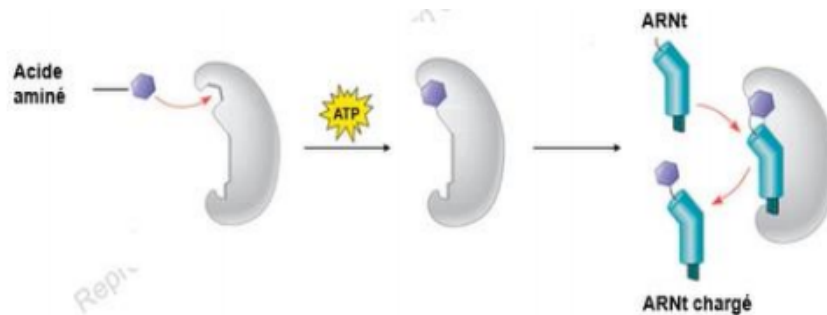
## Les acteurs et le déroulement de la traduction

### 1- Les acteurs de la traduction

<b>L'ARNm</b>	Contient les instructions pour la synthèse de la protéine.
<b>Les ARNt (ARN de transfert)</b>	Portent leur unique acide aminé spécifique, ils se fixent sur le codon de l'ARNm correspondant à ce dernier et apportent l'acide aminé au ribosome.
<b>Les aminoacyl-ARNt synthétases</b>	Fixent les acides aminés sur les ARNt.
<b>Les ribosomes</b> , formés de protéines et d'ARN ribosomaux	Constitués de deux sous-unités - accueillent les ARNt chargés de leur acide aminé. Ensuite, ils relient entre eux ces acides aminés par des liaisons peptidiques afin de former la protéine.



Les **aminoacyl-ARNt synthétases** ont une activité de correction (proofreading), Cela leur permet d'éliminer un acide aminé fixé par erreur avant de libérer l'ARNt, ce qui évite son incorporation erronée et assure donc la fidélité de la traduction.



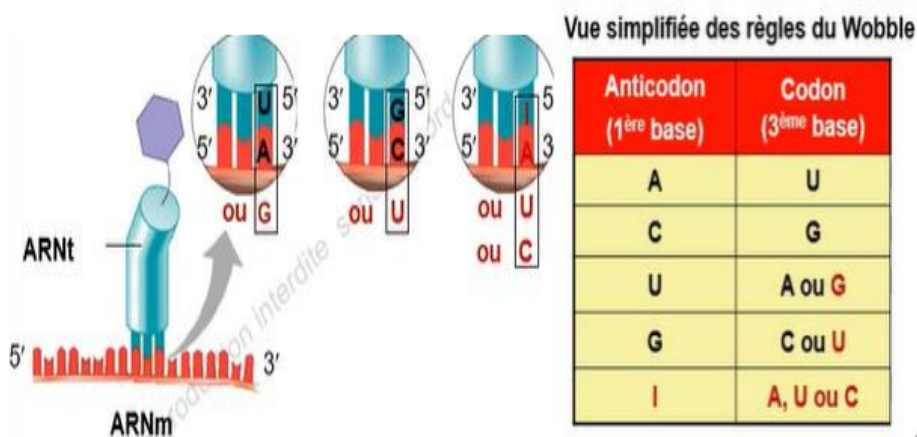
### Un appariement flexible : le Wobble

Le déchiffrement du code utilise le Wobble comme un appariement flexible.

Cet appariement ne respecte pas le principe de complémentarité.

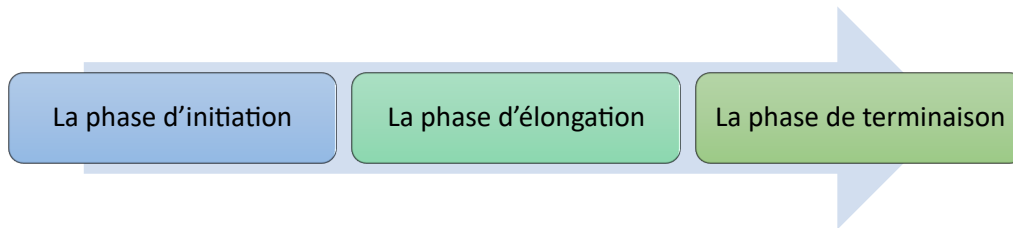
Cependant, la règle de l'appariement entre une purine et une pyrimidine est dans la majorité des cas respectée.

Le Wobble permet à l'anticodon d'un ARNt de s'apparier avec plusieurs codons qui spécifient le même acide aminé (codons synonymes) afin de réduire le nombre d'ARNt nécessaire à la traduction.



## 1- Le déroulement de la traduction

La traduction se fait en 3 phases

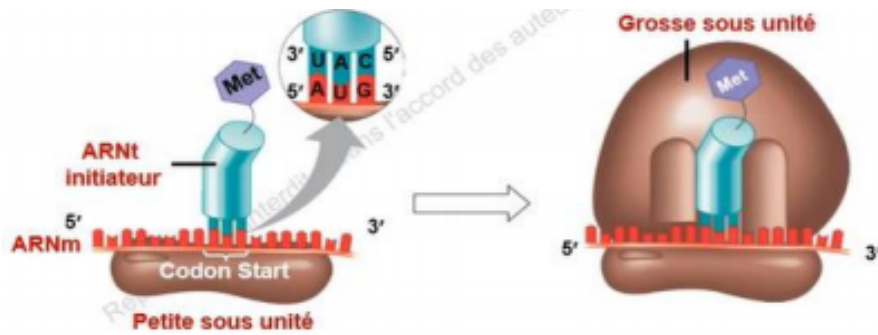


La phase d'initiation	La phase d'élongation	La phase de terminaison
Elle aboutit à l'assemblage du ribosome complet sur l'ARNm au niveau du Codon Start AUG qui indique le début de la séquence codante à traduire.	Elle correspond au déplacement du ribosome sur l'ARNm de codon en codon selon le cadre de lecture jusqu'au Codon Stop d'arrêt de la traduction. A chaque codon, un nouvel acide aminé apporté par un ARNt est incorporé au peptide en cours de synthèse par formation d'une liaison peptidique.	Elle correspond à la fin de la traduction avec libération de la protéine complète.

### 1<sup>ère</sup> étape - La phase d'initiation comprend deux étapes

<b>L'assemblage d'un complexe de pré-initiation sur l'ARNm</b>	Il se forme au niveau du Codon Start chez les procaryotes et en amont chez les eucaryotes. Il comprend notamment la petite sous-unité et l'ARN de transfert initiateur portant la méthionine.
<b>L'assemblage du ribosome complet au niveau du Codon Start</b>	Il nécessite le déplacement du complexe de pré-initiation sur l'ARNm chez les eucaryotes. A ce stade, l'ARN de transfert initiateur et la méthionine sont positionnés au site P du ribosome.



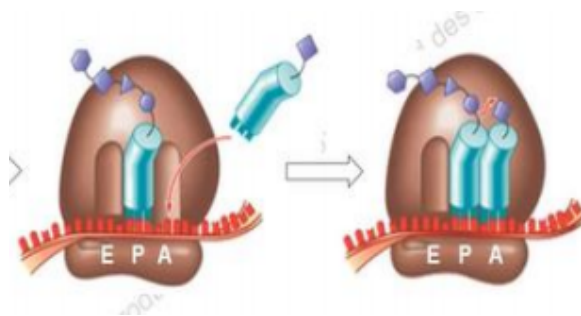


## 2<sup>ème</sup> étape - La phase d'élongation : une succession de cycles.

Si l'appariement codon-anticodon est correct, alors le peptide est transféré sur l'acide aminé par la formation d'une liaison peptidique.

Puis, le ribosome se déplace d'un codon. Le peptide allongé d'un acide aminé revient au site (P) et l'ARNt déchargé passe au site (E) et est éjecté.

Le cycle recommence ainsi par arrivée au site (A) ribosomal d'un nouvel ARNt chargé d'un AA.

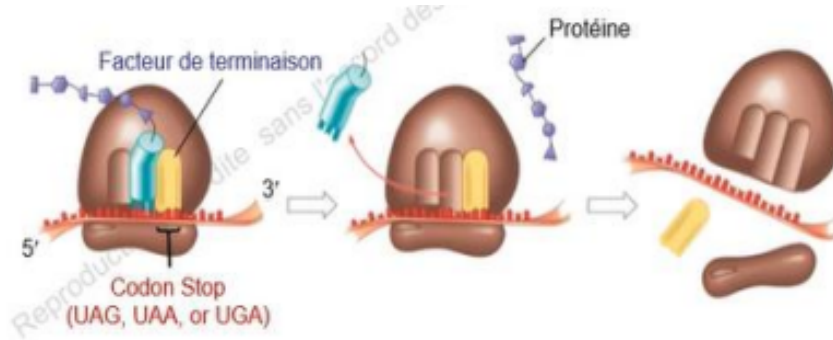


## 3<sup>ème</sup> étape - La phase de terminaison

Elle correspond à l'arrêt de la traduction qui s'achève lorsque le ribosome rencontre un Codon Stop.

Il n'existe pas d'ARNt spécifique aux Codons Stop.

Une protéine appelée facteur de terminaison se positionne au niveau du site (A), la protéine est alors libérée et le ribosome se dissocie pour participer éventuellement à un autre cycle de traduction.



A noter : Un ARNm est traduit simultanément par plusieurs ribosomes.

L'ensemble de l'ARNm – ribosomes forme un **polyribosome**. Il contient en permanence de nombreuses molécules en cours de synthèse. Ainsi, l'efficacité et la rapidité de la traduction sont accrues (bravo les ribosomes).

## 2 - EXPRESSION GÉNIQUE ET RÉGULATION CHEZ LES PROCARYOTES

En absence de noyau, la **transcription** et la **traduction** se font de manière **simultanée** chez les êtres procaryotes.

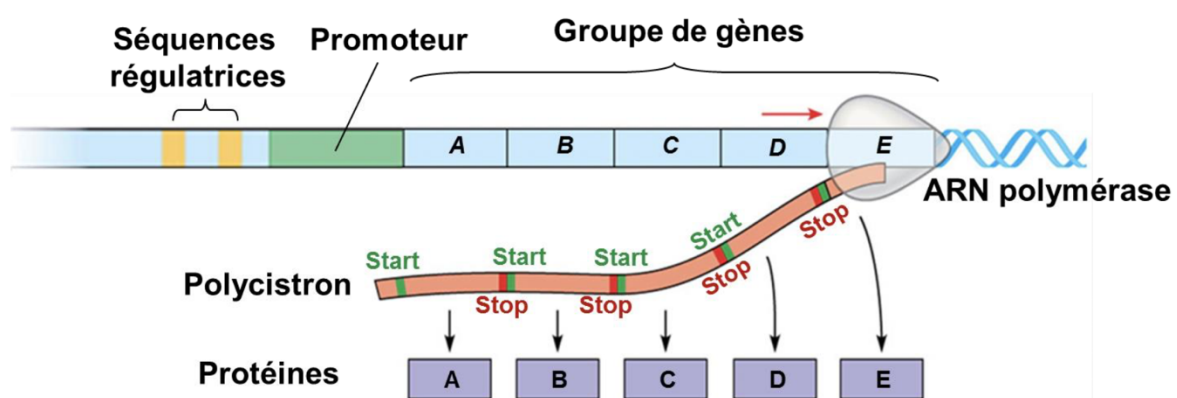
Les ribosomes peuvent s'associer à l'ARNm dès le début de sa synthèse et commencer à le traduire en protéine.

Les gènes procaryotes sont organisés en **opérons** : un **ensemble de gènes soumis à une régulation commune**.

- ◆ Cette régulation est assurée par un promoteur et d'autres séquences régulatrices situées en amont.
- ◆ L'intérêt de cette régulation commune est de pouvoir activer ou réprimer simultanément l'expression de gènes impliqués dans une même fonction.
- ◆ Un opéron contient sous une forme compacte, la séquence codante de plusieurs gènes
- ◆ Ces séquences codantes sont mises bout à bout et ininterrompues.

L'opéron est alors transcrit sous la forme d'un **unique et long ARNm (polycistron)** immédiatement mature.

⇒ C'est cette organisation qui explique et autorise la simultanéité de la transcription et de la traduction



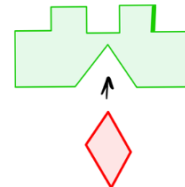
### Ça ressemble à quoi un opéron ??



#### Ensemble de gènes

**Promoteur** → Séquence régulatrice reconnue par ARN polymérase pour commencer la transcription (souvent séquence appelée TATA box)

**Opérateur** → Fixation possible par une protéine régulatrice



**Protéine TRANSrégulatrice** → le gène qui code pour cette protéine se trouve en amont

⇒ On parle de régulation en **CIS**, quand la régulation se trouve proche de l'opéron (c'est le cas du promoteur et de l'opérateur)

⇒ Et en **TRANS**, quand à l'inverse la régulation se fait en amont

**Ligand** → molécule qui modifie la conformation de la protéine régulatrice et donc l'expression de l'opéron

⇒ Le promoteur et l'opérateur **font partie de l'opéron**

(Schema repris à notre vieille Stabilo'drey <3)

La régulation d'un opéron fait intervenir 2 types d'éléments :

#### ♥ Des séquences « cis-régulatrices »

⇒ On parle de régulation en « cis » (*proche*) car ces éléments sont formés de séquences d'ADN contenues dans l'opéron lui-même.

⇒ Le motif qui est formé par ces séquences régulatrices va constituer un signal de fixation pour des protéines régulatrices impliquées, selon les cas, dans l'activation ou dans la répression de la transcription (*on explique tout ça plus bas*)

⇒ Promoteur = type de séquence régulatrice qui va être reconnue par l'ARN polymérase et au niveau de laquelle elle va se fixer pour initier la transcription. Le plus fréquent est la TATA Box (séquence TATAA)

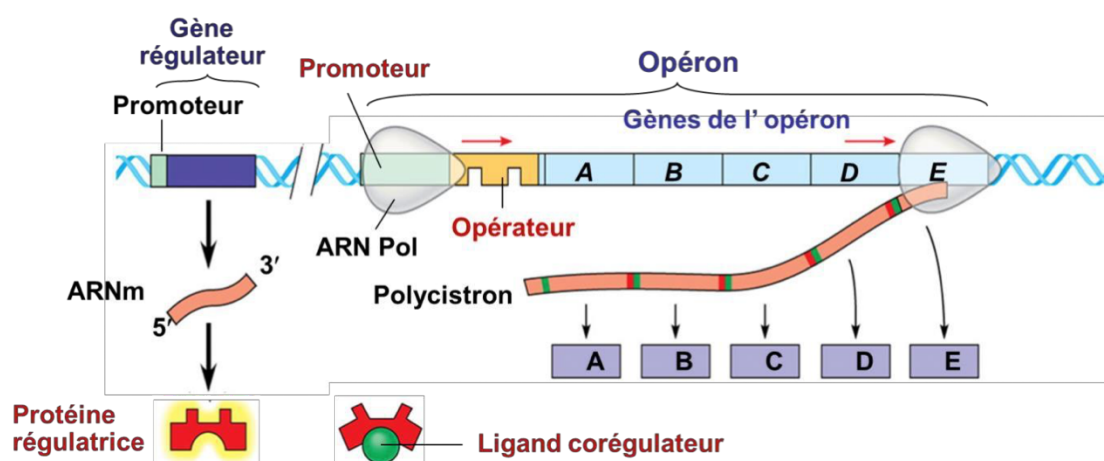
⇒ Opérateur = séquence plus ou moins éloignées du promoteur qui participe à la régulation de l'opéron

## ♥ Des protéines appelées facteurs transrégulateurs

Elles vont venir activer ou inhiber la transcription en se fixant à l'ADN au niveau de la séquence régulatrice qui leur est spécifique

On parle de régulation en « Trans » car le gène codant pour une protéine régulatrice est situé à distance de l'opéron et possède lui-même son promoteur et ses séquences régulatrices propres

En plus d'un domaine de liaison à l'ADN, ces protéines possèdent un domaine de liaison à de petites molécules appelées ligands et dont la fixation modifie leur conformation et leur activité



## Types d'opérons

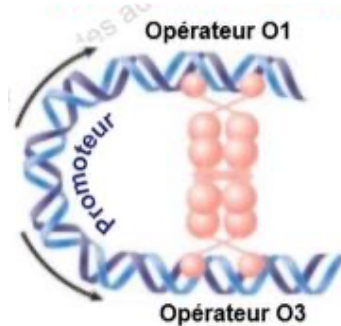
Opéron dit répressible <b>s'exprime quand il n'y a pas de ligand</b>  s'exprime de façon « constitutive »	Voies <b>anaboliques</b>	Ex. synthèse du <b>tryptophane</b> – rôle de ligand. Si absent, faut en synthétiser.	Quand la molécule est disponible pour la cellule, elle joue le rôle de <b>ligand corépresseur</b> en se fixant sur une protéine régulatrice répressive et en l'activant.
Opéron dit inductible <b>s'exprime quand il y a le ligand</b>  peut être réprimé de façon « constitutive »	Voies <b>cataboliques</b>	Ex. dégradation du <b>lactose</b> - rôle de ligand. Si présent, faut le dégrader.	En présence de la molécule de lactose, cette molécule va jouer le rôle de ligand inducteur en se fixant sur la protéine répressive et en l'inactivant. Ainsi l'opéron peut s'exprimer.

## Exemple de l'opéron Lactose

- ♦ L'opéron lactose est l'opéron qui permet la **dégradation du lactose**.
- ♦ Se trouve dans la bactérie E. Coli.
- ♦ Prolifère en présence de glucose et de lactose. Mais si les 2 nutriments sont présents, sa préférence est le glucose.
- L'opéron lactose est constitué de 3 gènes, et du gène **Lac I** qui code pour la protéine régulatrice Lac I.



La particularité de cet opéron est qu'il possède **3 opérateurs** dont O3 et O1 qui **encadrent le promoteur**

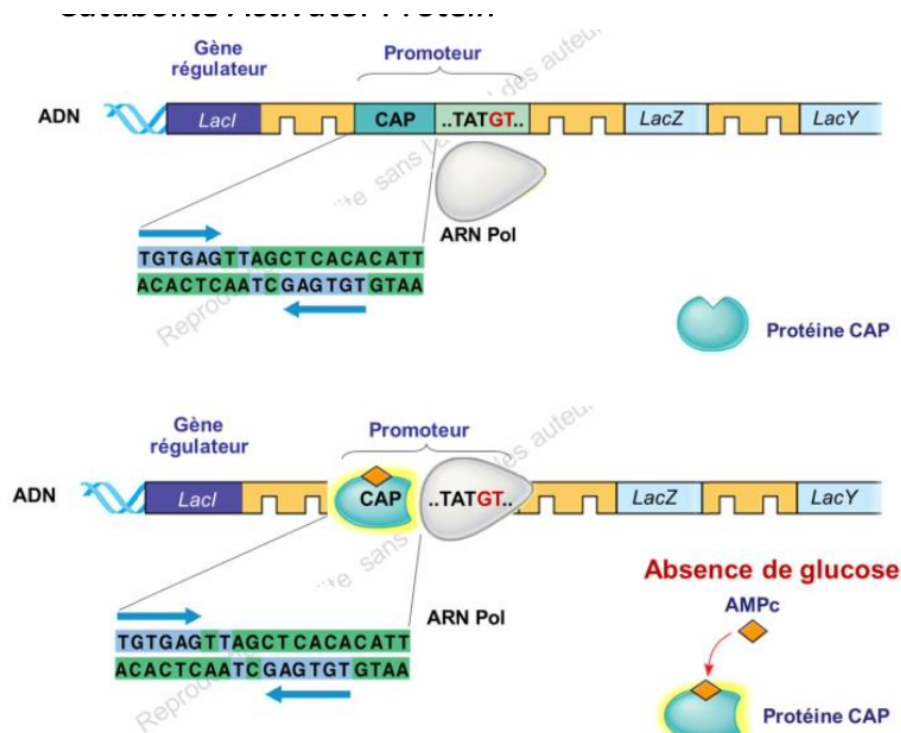


Si le lactose (= le ligand) n'est pas présent, Lac I est sous la forme d'un homotétramère et il enferme le promoteur. L'ARN polymérase ne peut pas accéder donc **aucune expression** car la transcription est impossible.

Mais s'il y a **présence de lactose**, alors **Lac I** modifie sa conformation et **libère le promoteur** et **l'expression est de nouveau possible**.

**Mais** le **glucose** a aussi une influence sur l'expression de l'opéron lactose

- Juste devant le promoteur se trouve la **séquence CAP**. C'est un site de fixation pour la **protéine CAP**. Lorsque cette protéine est fixée, la liaison de l'ARN polymérase sur le promoteur est facilitée.
- La protéine CAP possède un domaine de liaison pour **l'AMPc**. Quand l'AMPc se fixe à la protéine, cette dernière est en conformation active et peut se fixer à la séquence.
- Mais la production d'AMPc se fait **inversement** à la présence de glucose . L'AMPc se fixe à la protéine CAP qui se fixe à la séquence CAP uniquement **en absence** de **glucose ++**



Au final on peut distinguer 3 états transcriptionnels de l'opéron lactose

Absence de lactose	Présence de lactose et de glucose	Présence uniquement de lactose
Protéine Lac I renferme le promoteur – aucune expression possible	Lactose : rôle inducteur  Glucose : rôle répresseur – empêchant production d'AMPc  Fixation ARN polymérase instable sur le promoteur	Promoteur libre  Présence AMPc, activation protéine CAP  Stabilisation ARN polymérase
<b>Etat réprimé</b>	<b>Etat permissif</b>	<b>Etat activé</b>