

DFGSP2 – UE4

Cours de Biologie Moléculaire

Document de travail personnel

Ce document est à l'usage exclusif des étudiants de DFGSP2 de l'UFR de Pharmacie de l'Université Paris-Sud Paris Saclay.

Ce document est destiné à être diffusé dans le strict cercle de la DFGSP2, certains objets d'iconographie étant protégés par des droits d'auteurs.

Il ne s'agit que de rappels de connaissances considérées comme acquises en fin de PACES et qui ne nécessitent pas d'être reprises en cours magistral.

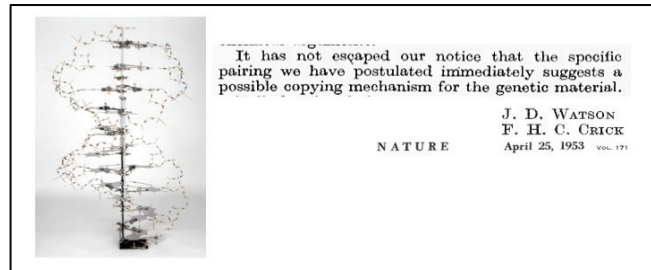
Il s'agit principalement de mécanismes liés à la Biologie Moléculaire. Ces mécanismes doivent être connus.

Les étudiants sont invités à travailler cette partie du cours par eux même et pourront poser des questions sur cette partie afin qu'elles soient traitées lors de l'amphi ED organisé à la fin des cours.

I- Structure de l'ADN et Réplication

En 1953 Watson et Crick démontrent la structure hélicoïdale de la molécule d'Acide Désoxyribonucléique (ADN). L'ADN, dans les conditions physiologiques est une double hélice composée d'un squelette sucre-phosphate au centre de laquelle se trouvent les bases azotées.

Le squelette sucre-phosphate est composé de liaison phosphodiester. Il existe une **complémentarité** des bases d'un brin à l'autre par l'intermédiaire de liaisons hydrogènes (2 liaisons hydrogènes entre A et T, 3 liaisons hydrogènes entre G et C). Enfin, les brins sont orientés par des extrémités 5' et 3' (numérotation des atomes de carbone du désoxyribose). Les 2 brins composant une double hélice sont **antiparallèles** et complémentaires.



Il n'avait pas échappé à Watson et Crick que cette structure en double hélice pouvait également expliquer le mode de duplication de la molécule.

Il est ainsi possible d'imaginer un mécanisme de déshybridation de chaque brin composant la double hélice et de duplication de chacun des brins "matrice" par le jeu de la complémentarité et de l'antiparallélisme.

Trois modèles théoriques peuvent rendre compte de la réplication de l'ADN.

- un modèle conservatif
- un modèle semi-conservatif
- un modèle dispersif

La double hélice est composée de deux brins complémentaires et antiparallèles.

Question au milieu des années 50:
Comment ces deux brins sont-ils utilisés pour générer deux nouveaux brins lors de la réplication ?

3 hypothèses possibles:

Modèle conservatif Modèle semi-conservatif Modèle dispersif

Les travaux de Meselson et Stahl, à l'aide de culture de levure en présence de radio isotopes légers ou lourds de l'azote ont permis de démontrer que la réplication de l'ADN se réalise selon un **modèle semi-conservatif**.

Les levures, au cours de leur croissance ont incorporé les atomes d'azote dans les cycles azotés composant les nucléotides. Ces nucléotides sont les précurseurs de la molécule d'ADN et seront incorporés dans la molécule néosynthétisée.

Après un cycle de réplication, l'observation d'une seule population moléculaire de densité intermédiaire indique que la réplication se réalise selon un modèle semi-conservatif.

La réplication est semi conservative

1958: Meselson et Stahl mettent en évidence le principe **semi-conservatif** de la réplication.
Mise au point du protocole de séparation de molécules d'ADN sur gradient de CsCl.

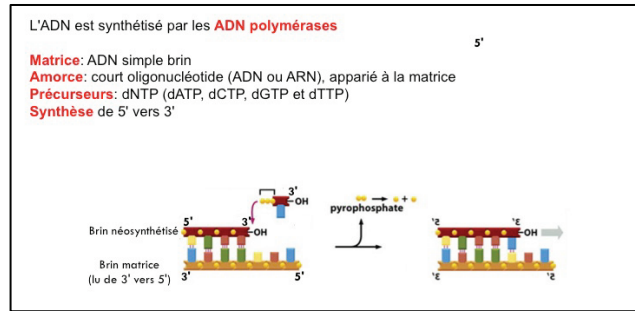
Original DNA Parental strand New strand



Au cours de l'année 1955 Arthur et Sylvie Kornberg publient leurs travaux sur l'étude de la synthèse d'ADN à partir d'un extrait brut de bactérie. Ils démontrent que l'ADN polymérase lit le brin matrice de 3' vers 5'. Elle permet à un désoxyribonucléotide "entrant" de s'apparier par complémentarité à la première base présente sur l'ADN simple brin et faisant immédiatement suite au dernier nucléotide incorporé. Enfin, la polymérase catalyse la formation de la **liaison phosphodiester 5'P-3'OH** du ribose n+1 sur le ribose n.

Ainsi ils démontrent les grandes propriétés des ADN polymérases.

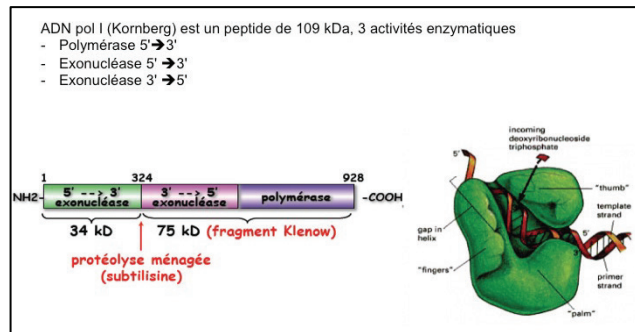
Ce sont des polymérases ADN dépendantes (le brin matrice doit être un brin d'ADN), elles nécessitent une amorce (ADN ou ARN).



L'ADN polymérase étudiée par Kornberg et l'ADN pol-I, elle possède trois activités enzymatiques :

- polymérase 5'-3'
- exonucléase 5'-3'
- exonucléase 3'-5'.

Ces trois activités permettent la néosynthèse ainsi que l'activité de correction.



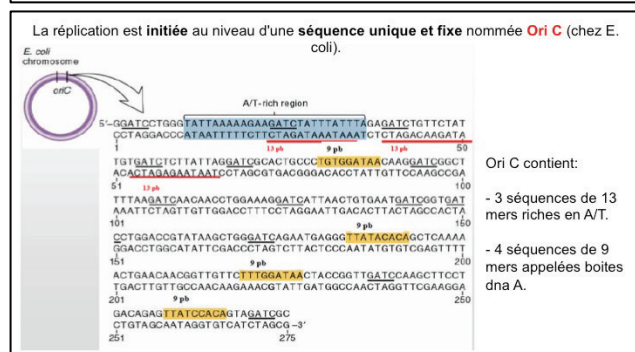
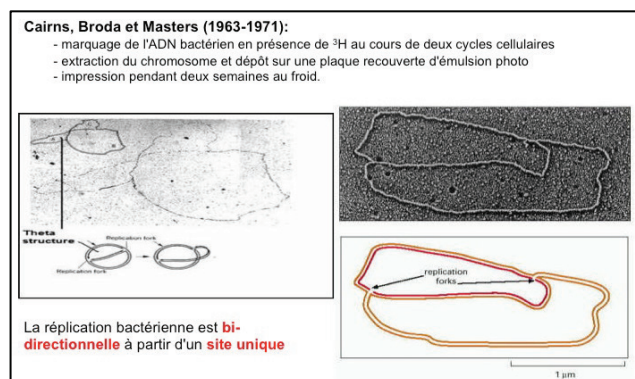
La réplication, qu'elle soit eucaryote ou procaryote se divise en trois étapes :

- **initiation** : ouverture de la double hélice et maintien sous forme simple brin
- **élongation** : Recopie des deux brins de manière coordonnée, et finalisation des brins retardés
- **terminaison** : Finalisation de la copie. Décaténation des moléculaires circulaires. Résolution des problèmes liés à la structure selon que l'ADN est circulaire (procaryote) ou linéaire (eucaryote).

Sur le chromosome bactérien, circulaire, l'origine de réplication est fixe et unique. Les origines de réplication sont multiples sur les chromosomes eucaryotes linéaires.

La réplication est **bi-directionnelle**.

La boucle de réplication, telle que décrite par Cairns, est composée de 2 fourches de réplication qui progressent en sens opposé.



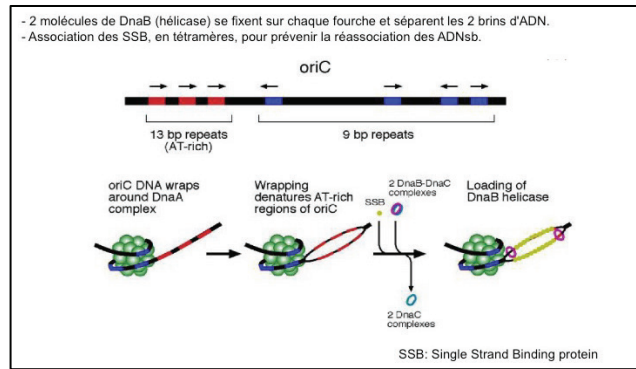
L'**origine** de réplication procaryote est **unique** et **fixe**. Elle est nommée OriC chez *E. coli*.

Il s'agit d'une séquence consensus qui permet le recrutement d'un complexe protéique dnaA au niveau de 4 séquences répétées (9mers).

La double hélice s'enroule autour du complexe dnaA et subit une contrainte physique qui conduit à l'ouverture des 3 séquences répétées riches en AT (13 mers).

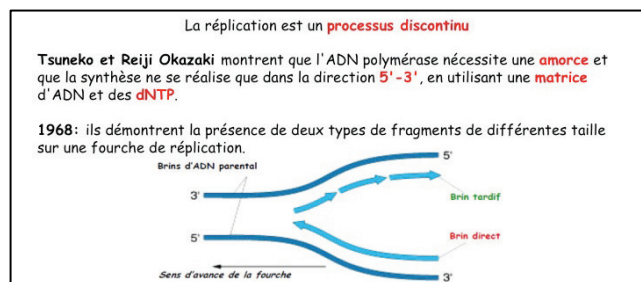
L'ouverture de la double hélice permet le recrutement de 2 dimères dnaB-dnaC. Les protéines dnaC sont libérées. dnaB sont les hélicases qui vont continuer d'ouvrir la double hélice, de chaque côté de la boucle de réplication. Ces hélicases provoquent une déshybridation des brins d'ADN appariés et consomment de l'ATP.

L'ADN simple brin est alors recouvert de protéines SSB qui protège l'ADN et retardent la possibilité de ré-appariement.



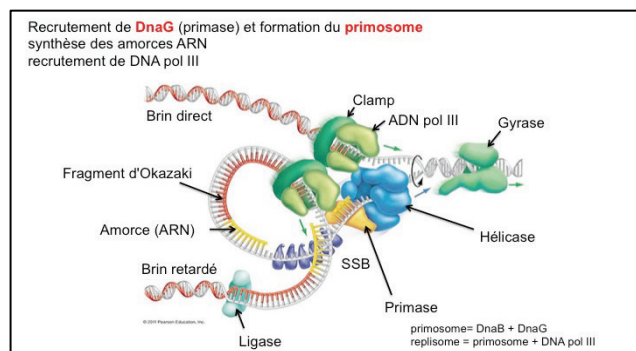
La réplication est un **processus discontinu**.

Dans les années 60, le couple Tsuneko et Reiji **Okazaki** montrent que les ADN polymérases synthétisent l'ADN exclusivement dans le sens 5'-3', à partir d'une amorce et en utilisant une matrice ADN et des précurseurs sous forme de dNTP.



Ils démontrent également qu'au niveau d'une fourche de réplication les fragments néo-synthétisés présentent des tailles différentes. Il existe un brin direct (continu) et un brin retardé (discontinu).

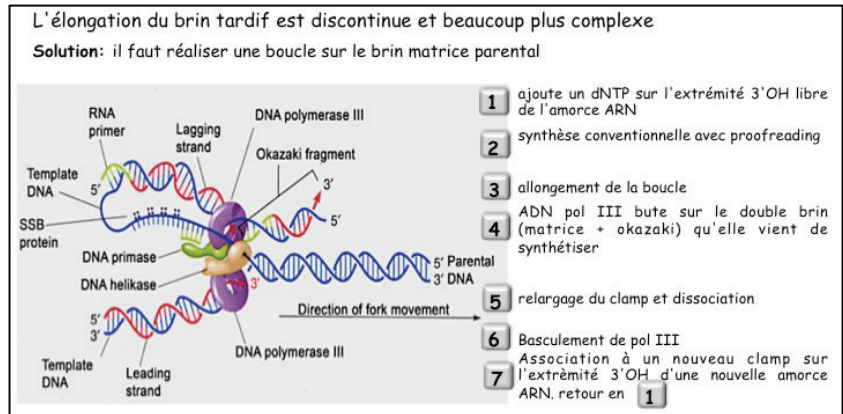
A l'issue de la phase d'initiation, la protéine dnaG s'associe à dnaB et forme le primosome. La fonction de dnaG est de synthétiser de courts brins d'ADN en utilisant le brin parental comme matrice. dnaG est une primase (ARN polymérase ADN dépendante). Ces courts fragments d'ARN sont les amorces (primers) nécessaires à l'activité de l'ADN polymérase.



Sur chaque fourche de réplication une holoenzyme composée de deux ADN-pol III, d'un clamp loader et de deux clamps avance dans le sens de progression de la fourche. Une polymérase synthétise de 5'-3' dans le sens de progression de la fourche (brin direct) tandis que l'autre polymérase synthétise de 5'-3' dans le sens opposé au sens de progression de la fourche (brin retardé).

Ainsi, sur le brin retardé, la polymérase devra régulièrement se libérer du clamp pour se dissocier de l'ADN parental et reprendre son travail de synthèse un peu plus loin. La polymérase reprend son travail au niveau d'une nouvelle amorce synthétisée par la primase. Elle se décrochera à nouveau en venant rencontrer le fragment qu'elle a synthétisé précédemment. Ces courts fragments d'ADN néosynthétisés sont les fragments d'Okazaki. Je vous conseille de regarder ce document pour faciliter la compréhension:

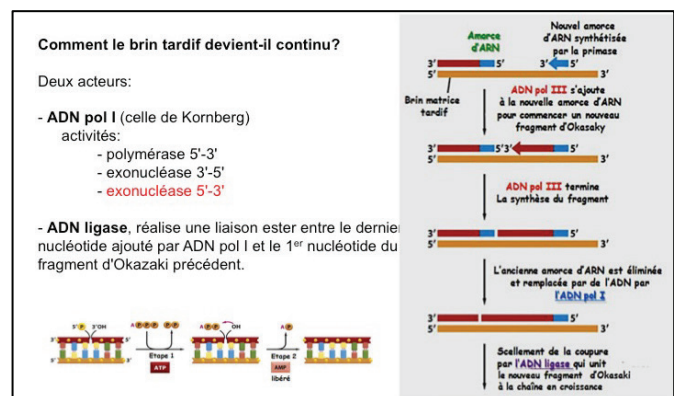
<https://www.dnalc.org/resources/3d/04-mechanism-of-replication-advanced.html>



Le brin tardif, composé des fragments d'Okazaki nouvellement synthétisés, des amorces ARN et apparié au brin d'ADN parental doit subir un certain nombre d'événement de maturation afin de devenir un brin continu.

L'ADN polymérase I, par son activité exonucléase 5'-3' hydrolyse les amorces ARN. Elle utilise le brin parental comme matrice et le fragment d'Okazaki précédent comme amorce pour synthétiser un fragment d'ADN à la place de l'amorce hydrolysée.

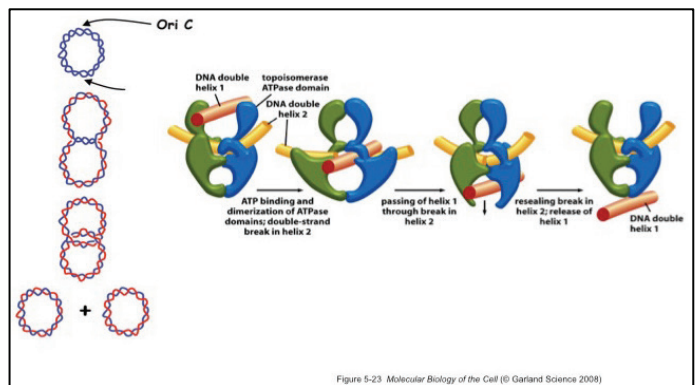
La dernière liaison phosphodiester est réalisée par une ADN ligase.



La terminaison est un phénomène complexe contrôlé par le complexe protéique Tus en interaction avec des séquences spécifiques de terminaison présentes sur le chromosome. Les 2 fourches de réplication doivent pouvoir se rencontrer et terminer proprement la réplication jusqu'à l'obtention de 2 chromosomes circulaires complets. Ces 2 chromosomes sont associés comme les maillons d'une chaîne. Il va falloir ouvrir un maillon pour séparer ces deux molécules, c'est la déconcaténation qui fait intervenir, entre autre une topoisomérase II.

Les topoisomérases sont des moléculaires majeures de la réplication de l'ADN. La gyrase, qui avance en amont de chaque fourche de réplication est une topoisomérase de type II. Elle permet de relâcher l'ADN superenroulé qui se compacte devant chaque replisome.

Ce sont des enzymes responsables de la topologie de l'ADN. Elles "font et défont" les



La réplication - topoisomérases

Les topoisomérases:

- enzymes responsables du contrôle de la topologie des molécules d'ADN.
- Font et défont des supertours.
- Parmi les protéines les plus conservées dans le monde ADN

- 2 catégories:

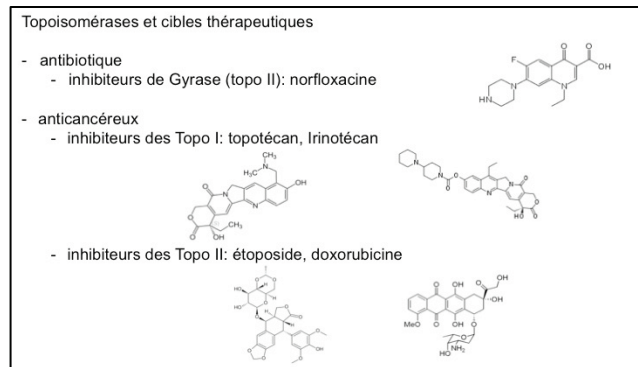
- Type I: responsables de cassures simple brin. Topo IA, IB, IC, III, V
- Type II: responsables de cassures double brin. Topo IIA, IIB, IV

- Fonctions variées:

- compaction
- réplication, ségrégation des chromosomes
- Transcription
- Recombinaison et réparation

supertours et elles sont parmi les protéines les plus conservées dans le monde ADN.

Les topoisomérases représentent des cibles pharmacologiques majeures. Leur inhibition est une stratégie antibiotique et anticancéreuse particulièrement efficace. Cependant, leur conservation dans le monde vivant rend délicate l'utilisation d'inhibiteur.



La réplication du génome nucléaire eucaryote est globalement similaire à la réplication procaryote.

Cependant :

- i) le nombre de protéines impliquées est plus important
- ii) la molécule d'ADN à dupliquer est plus longue
- iii) les chromosomes sont structurés en nucléosomes
- iv) les chromosomes sont linéaires.

Il existe un plus grand nombre d'ADN polymérases eucaryotes. Une quinzaine contre 5 chez les procaryotes. Seules les polymérases α , β et δ sont impliquées dans la réplication. Les autres ont des fonctions spécifiques dans les mécanismes de réparation ou dans des mécanismes cellulaires spécifiques.

Le mécanisme global de réplication est assez similaire à celui des procaryotes et les détails les différenciant ne rentrent pas dans le cadre de cette formation.

La réplication eucaryote est bidirectionnelle, mais elle est initiée en de multiples sites.

ADN pol α , associée à une primase. Synthétise d'abord une amorce ARN puis active ADN pol α prend le relai pour une dizaine de nucléotides

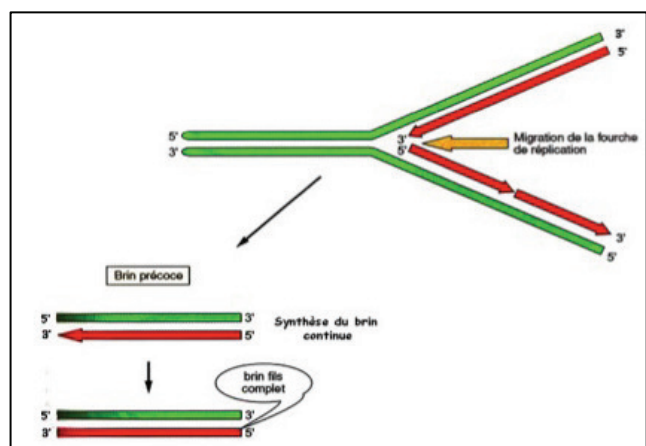
ADN pol δ est l'enzyme principale de polymérisation

élimination des amorces ARN par une **RNase H**

remplacement de l'ARN par **ADN pol δ**
Ligation par **ADN ligase**

- Vitesse de ADN pol eucaryotes est inférieure à celle des ADN pol procaryotes
- Nombre plus important
- Fonctions plus séparées entre réplication et réparation

Un point important, en ce qui nous concerne, porte sur l'extrémité des chromosomes eucaryotes. Comme cela a déjà été mentionné, ces chromosomes sont linéaires. Cette caractéristique pose une difficulté lors de la terminaison de la réplication. En effet, la polymérase positionnée sur le brin direct avance dans le sens de la fourche et elle pourra donc terminer sa synthèse en extrémité de brin sans difficulté.



Par contre la polymérase positionnée sur le brin discontinu ne pourra pas terminer la synthèse du brin néosynthétisé, faute d'amorce. De ce fait, le brin retard va conduire à un raccourcissement du chromosome.

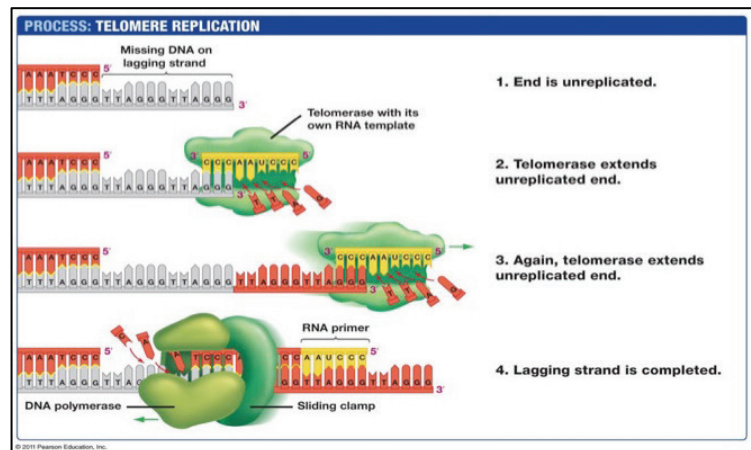
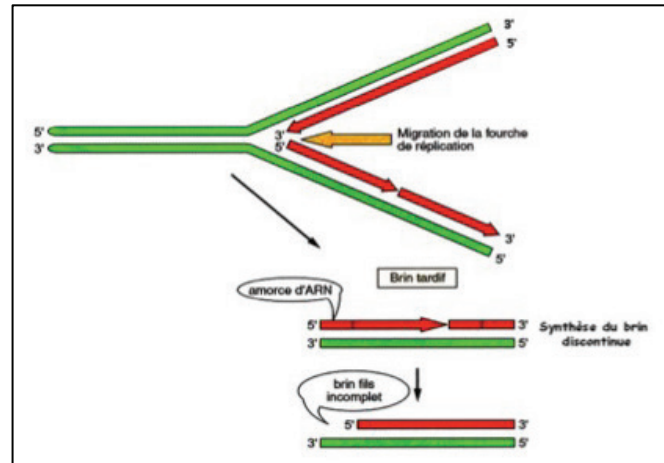
Un mécanisme additionnel va permettre d'allonger la double hélice d'ADN en ajoutant une courte séquence d'ADN répétée une multitude de fois. Cette séquence répétée en extrémité de chromosome est appelée télomère et est ajoutée à l'issue de chaque réplication par une télomérase. Les télomérases sont des ADN polymérases ARN dépendantes. Il s'agit d'un complexe nucléo-protéique renfermant un brin d'ARN qui sert de matrice à la synthèse d'une séquence ADN.

Ce brin d'ARN simple brin porté par les télomérases renferme une séquence 5'-CCCUAACCC-3'.

Ainsi, tout chromosomes se termine par un télomère qui est une répétition de l'héxanucléotide 5'-TTAGGG-3'.

Par complémentarité et antiparallélisme, comme pour toute hybridation entre acides nucléiques, la télomérase va s'associer à cette séquence 5'-TTAGGG-3' par sa molécule d'ARN 3'-CCCAAUCCC-5'. La séquence 3'-AAUCCC-5' simple brin sert alors de matrice pour la synthèse d'une nouvelle répétition 5'-TTAGGG-3'. La télomérase se décroche et recommence un peu plus loin.

La polymérase δ va ainsi reprendre son travail et allongé le brin discontinu en synthétisé une séquence complémentaire à ces multiples répétition. Lors de la terminaison l'extrémité du chromosome sera raccourci mais seule une séquence télomérique sera éliminée. Ainsi, aucune séquence génomique n'est perdue.



L'activité des télomérases a tendance à diminuer avec l'âge des cellules et la longueur des télomères peut être un indicateur du vieillissement cellulaire. Ceci est valable pour certains organismes (dont l'Homme) et pas pour d'autre (comme la Souris).

II- Mutations et mécanismes de réparation de l'ADN.

Comme nous l'avons vu, un chromosome est, dans la majorité des cas, constitué d'une molécule d'ADN hélicoïdale double brin. Cette structure lui confère une certaine stabilité physique et informative. En effet, il existe assez peu de risque de perdre exactement le même message deux fois.

Les ADN polymérases responsables de la réplication sont des enzymes douées de fidélité et d'activité de correction d'erreur. Ce qui conduit à des taux d'erreur globale relativement bas, de l'ordre de 10^{-6} à 10^{-8} .

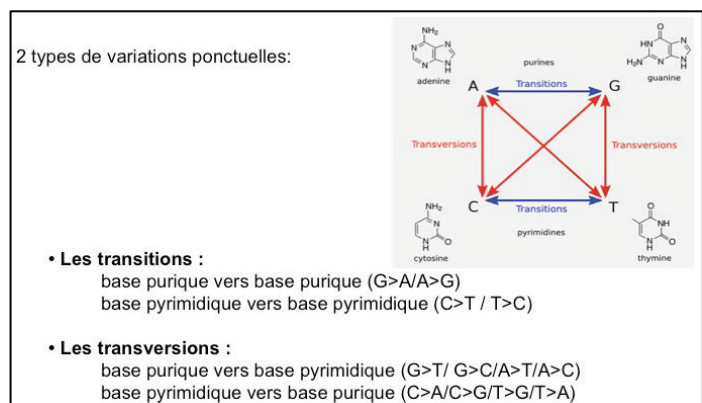
L'ADN peut aussi subir des dommages qui vont conduire à la modification de certaines bases ou encore provoquer des cassures simple ou double brin. Ces dommages doivent être détectés et réparés par des machineries cellulaires spécialisées. Il s'agit d'un processus particulier, spécifique à l'ADN qui est la seule macromolécule qui puisse être réparée.

Il arrive cependant que certaines erreurs ou dommages ne puissent pas être corrigés et ces défauts de réparation sont à l'origine d'un grand nombre de pathologies.

Les mutations peuvent avoir des causes diverses, certaines sont citées ci-dessous mais cette liste n'est en rien exhaustive.

MUTAGÈNES PHYSIQUES	
Chaleur	désamination, dépurination
Rayonnement UV	dimères de thymine
Radiations ionisantes	cassures simple et double brin
MUTAGÈNES CHIMIQUES	
Analogues de bases	erreur réplication
Modificateurs de bases	mésappariement
Alkylants	mésappariement
Intercalants	erreur réplication
Agents de Pontage	liaison covalentes inter-brin
MUTAGÈNES BIOLOGIQUES	
Virus, Transposons	insertion-délétion (Indel)
Activité Biologique	ADN pol, ROS

Les mutations dites ponctuelles portent sur la modification d'une base et une seule. Ce sont des transitions lorsque la nature, purique ou pyrimidique, de la base est conservée ou des transversions si la nature de la base est modifiée.



Les mutations peuvent se présenter sous diverses formes et affecteront l'ADN tant dans la séquence que dans la taille de la molécule. Les mutations ponctuelles ou de substitution ainsi que les inversions ne conduisent pas à une modification de taille. Les délétions, insertion, duplication, amplification conduiront à une modification de la taille.

Les dommages à l'ADN peuvent avoir des origines multiples :

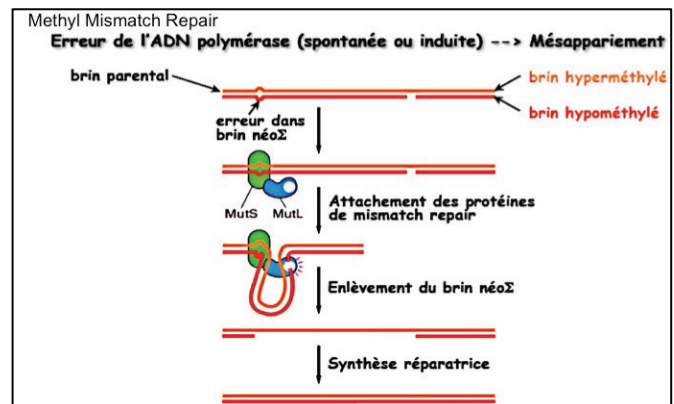
- Cassures simples brins (très fréquentes)
- cassures double brin (peu fréquentes)
- modification du métabolisme cellulaire
- Erreurs de réplication
- Exposition aux UV
- Exposition aux radiations ionisantes
- Exposition à des agents chimiques

L'ADN subit des cassures simples et doubles brin un nombre multiple de fois au cours d'un cycle cellulaire (entre 2 réplifications). Ces cassures sont réparées par des machineries spécialisées.

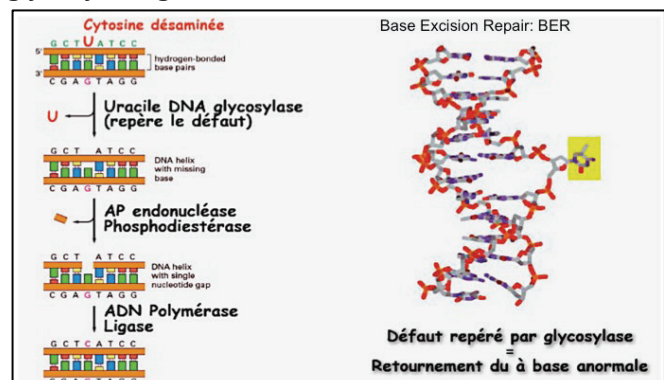
La protéine PARP (Poly-ADP Ripose Polymérase) est une molécule qui détecte et signale les cassures simple brin. Elle détecte une cassure simple brin et permet le recrutement de la machinerie de réparation des cassures simple brin (Single Strand Breaks). Le BER qui nous verrons plus loin est un de ces mécanismes de réparation. Si la protéine PARP est défectueuse ou si elle est inhibée, les SSB ne sont pas réparés et conduisent très souvent à des DSB (Cassures Double Brin) au cours de la réplication suivante.

Les cassures double brin sont prise en charge par des mécanismes complexes qui feront appel à des processus de réparation par recombinaison homologue (HRR) ou non homologue (NHEJ). Dans le premier cas la réparation ne conduit pas à une modification du génome, dans le second cas, la réparation est forcée mais de l'information génétique est perdue (avec des conséquences variables).

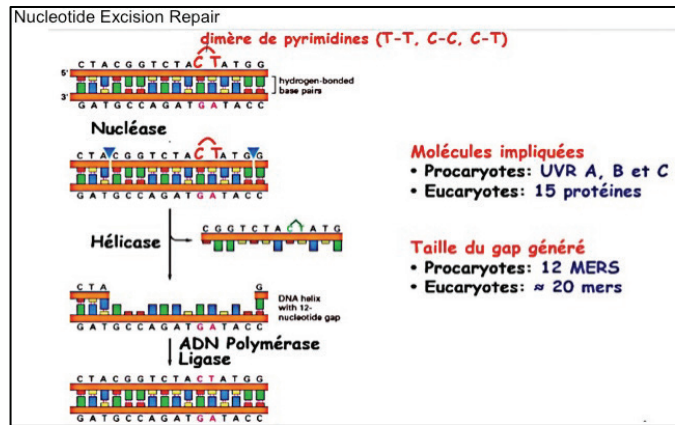
Lors d'une erreur d'appariement par l'ADN polymérase au cours de la réplication, le mésappariement est reconnu par le système MMR (Methyl Mismatch Repair). Ce système reconnaît le site altéré par la modification de structure engendrée par le Mismatch et reconnaît le brin néosynthétisé par son état d'hypométhylation.



La modification d'une base, comme la présence d'un Uracile par désamination d'une Cytosine, sera prise en charge par le système BER. L'Uracile DNA glycosylase génère un site AP. ce site AP est reconnu par l'AP endonucléase phosphodiesterase qui va rompre les liaisons phosphodiester entourant le site AP, libérant ainsi le désoxyribose et générant une cassure simple brin. Cette cassure, détectée par PARP est comblée par une ADN polymérase et la dernière liaison phosphodiester est réalisée par une ADN ligase.

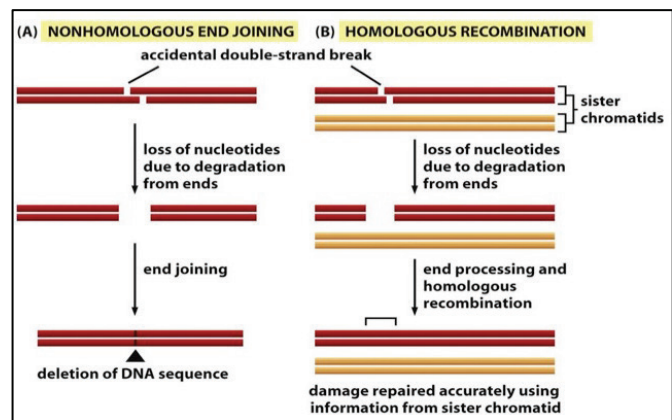


Les dimères de pyrimides, générés principalement par une exposition aux UV, sont pris en charge par le système NER (Nucleotide excision Repair). Un fragment de 12 à 20 nucléotides de longueur, encadrant le dommage est excisé par l'action d'une nucléase et d'un hélicase, laissant une région simple brin qui va être comblée par une ADN polymérase et une ligase.



Les cassures doubles brins peuvent être provoquées par des agents oxydants, des produits métaboliques ou des radiations ionisantes. Ces cassures doivent être réparées afin de maintenir l'intégrité des chromosomes. Ces cassures sont résolues par deux grands mécanismes :

- 1- le HRR (Homology Recombination Repair) qui aboutit à une réparation parfaite de la cassure
- 2- le NHEJR (Non Homologous End Joining Repair) qui aboutit à une perte de nucléotides.



Des défauts dans ces mécanismes de réparation ont été identifiés et associés à la survenue d'un certain nombre de pathologies. Une liste non exhaustive est donnée ci-contre à titre d'exemple.

Maladie	Symptômes	Défaut moléculaire	Gènes impliqués
Xeroderma Pigmentosum	Photosensibilité, Cancer de la peau	BER, synthèse translesionnelle	XPA, B, C, D, E, F, G, XPV (pol h)
Trichothiodystrophie	Retard Physique et mental	BER, Transcription	XPD, XPB, TTDA
Syndrome de Cockayne	Nanisme, déficience mentale, photosensibilité	Réparation préférentielle de l'ADN transcrit	CSA, CSB, XPB, D, G
Syndrome de Bloom	Prédisposition au cancer, remaniements chromosomiques, instabilité génétique	Hélicase, Réparation par recombinaison	BLM
Ataxia Telangiectasia	troubles neurologiques, dilatation des vaisseaux, prédisposition au cancer	absence d'arrêt en G1/S	ATM, ATR
Syndrome de Li Fraumeni	prédisposition au cancer	absence d'arrêt en G1/S	P53
Anémie de Fanconi	anémie, prédisposition aux leucémies, hypersensibilité aux agents pontants, instabilité chromosomique	Gènes de réparation	FANCA, C, D2, E, F, GA
Syndrome de Werner	vieillesse précoce, instabilité génomique	Hélicase	WRN
HNPCC	cancer du colon, instabilité microsatellitaire	déficience du MMR	MSH2, MLH1

III LA TRANSCRIPTION

L'ADN génomique est structuré et dynamique comme vous l'avez vu en cours. Un certain nombre d'éléments seront transcrits en ARN. Parmi ces ARN, environ 2 à 3% sont représentés par les ARNm qui seront, le cas échéant, traduits en polypeptides.

Les ARN

Structures moléculaires ordonnées et linéaires
 Molécules simples brins
 Orientées 5'-3'
 Possibilités d'un grand nombre de structures secondaires

Les principaux types d'ARN

	% des ARN totaux (cellule eucaryote)
ARNm: codent pour les protéines	3%
ARNr: participent à la structure des ribosomes et à la traduction	70%
ARNt: participent à la traduction	15%
ARNsn: petits ARN nucléaires	} 10%
ARNsno: petits ARN nucléolaires	
ARNi: ARN interférents, régulation de la traduction	
mi-ARN: micro ARN, régulation de la traduction ou de la demie-vie des ARNm	

La transcription représente l'ensemble des mécanismes qui permettent à la machinerie de transcription de reconnaître une séquence à transcrire (sur un brin), de lire ce brin et de synthétiser une séquence complémentaire et antiparallèle en utilisant des Ribonucléotides. Les ARN, sauf cas exceptionnels, seront donc composés de A, U, G et C.

TRANSCRIPTION

Transcription:

- synthèse d'ARN à partir d'une matrice ADN

- Nécessite:

- Une matrice ADN
- une ARN polymérase
- des NTP (ATP, GTP, CTP et UTP)

Les enzymes responsables de la synthèse d'ARN sont des ARN polymérase ADN dépendantes. Ce sont des enzymes composées de plusieurs sous unités. Il existe une seule ARN polymérase chez E.coli et il en existe 5 différentes chez les eucaryotes.

Ces ARN polymérase lisent un brin d'ADN matrice (le brin transcrit) de 3'-5' et synthétisent donc un brin d'ARN de 5'-3'. L'ARN synthétisé aura donc exactement la même séquence et la même orientation que le brin codant, au U près (voir les conventions dans le document de cours).

Les ARN polymérase

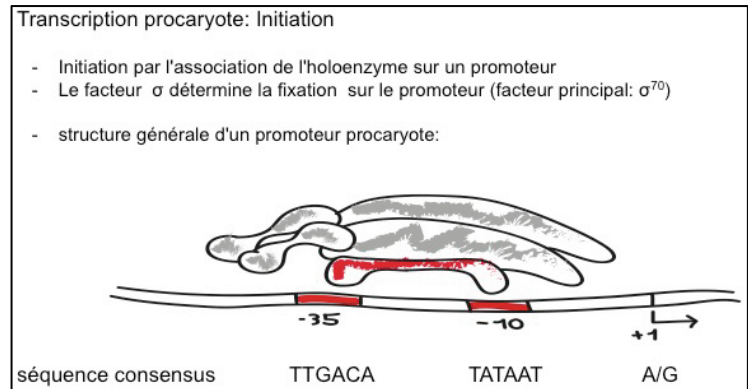
Les ARN polymérase connues sont formées de plusieurs sous unités

Procaryotes	Eucaryotes
Une seule ARN polymérase pour tous les gènes. 5 sous unités (α , α' , β , β' , ω) + σ <div style="text-align: center; margin-top: 5px;"> </div> cœre holoenzyme	ARN pol I : ARNr 18s; 28s; 5,8s (nucléole) ARN pol II: ARNm (noyau), miARN ARN pol III: ARNt, ARNr (5s), μ ARN (noyau) ARN pol IV: siARN (plantes) ARN pol V: petits ARN (plantes) de 12 à 16 sous unités ARN pol mitochondriale est un cas particulier

Les différentes ARN polymérase eucaryotes sont responsables de la transcription des différents types d'ARN.

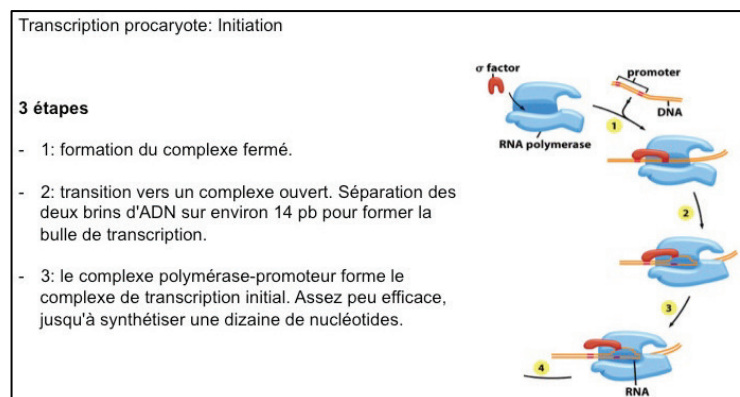
Toute transcription débute par une étape de reconnaissance d'une séquence particulière sur la double hélice d'ADN spécifiant un site d'initiation de la transcription. Cette séquence contient des éléments conservés d'un organisme à l'autre, ce sont des séquences consensus. Ces séquences définissent un promoteur. Le promoteur est toujours situé en amont, en 5' de la séquence à transcrire (Attention aux conventions ici!).

L'initiation de la transcription chez les procaryotes débute par l'association de l'holoenzyme sur un site promoteur. Cette association est déterminés par le facteur σ . (la nature du facteur σ spécifiera le type de promoteur). L'holoenzyme procaryote reconnaîtra plus particulièrement une séquence consensus en -35 du site d'initiation et la TATA box en -10 du site d'initiation.



La phase d'initiation est une phase relativement peu efficace et peu processive. Le facteur σ favorise la transition vers un complexe ouvert et permet la séparation de la double hélice d'ADN sur environ 14 pb. La bulle de transcription est ainsi formée.

Le complexe de transcription initial (CTI) synthétise une dizaine de nucléotides de manière peu processive.

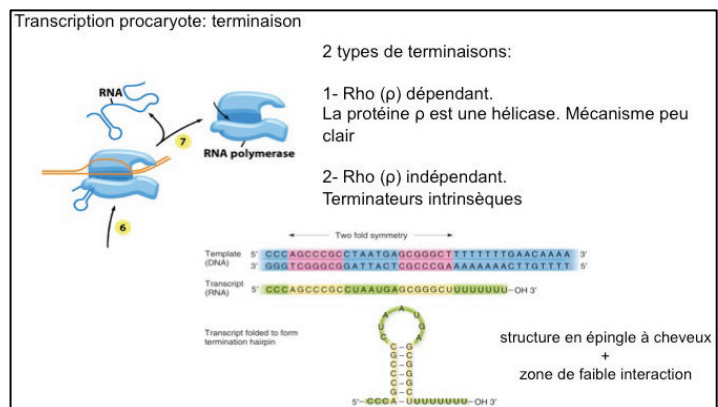


Par la suite, l'ARN en cours de synthèse favorise le départ du facteur σ . Le CTI se dissocie du promoteur et la formation d'un complexe ternaire stable permet de procéder vers la phase d'élongation qui est processive et capable de correction d'erreurs de synthèse.

La terminaison de la transcription procaryote peut être réalisée selon deux modes :

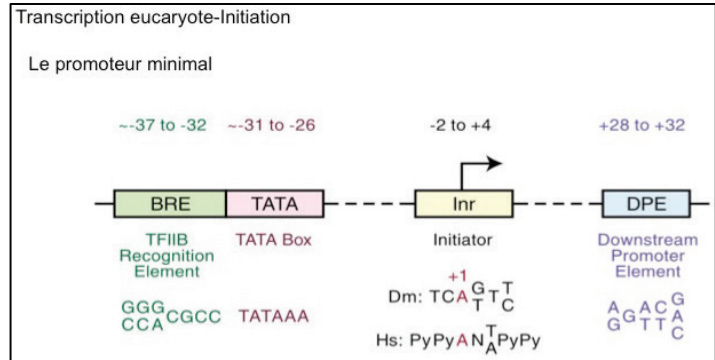
- Terminaison ρ -dépendante. La protéine ρ est une hélicase qui va favoriser la dissociation du complexe ternaire, le mécanisme mis en jeu est encore peu clair.

- Terminaison ρ -indépendante : dépend de structures secondaires se formant sur le transcrit en cours de synthèse, ce sont des terminateurs intrinsèques.



La transcription eucaryote est légèrement plus complexe et les mécanismes de régulation sont plus fins.

La notion de promoteur est plus étendue. Nous parlerons dans un premier de promoteur minimal. Il s'agit de séquences consensus situées dans une zone allant de 40 pb en 5' à 30 pb en 3' du site d'initiation du gène à transcrire.



Nous simplifierons ici au maximum. Il faut cependant retenir qu'en plus de ce promoteur minimal existe d'autres séquences consensus qui vont permettre de recruter des facteurs de transcriptions et des facteurs de la régulation de la transcription. Ces différents acteurs protéiques participent à la régulation transcriptionnelle en fonction de signaux issus de l'environnement cellulaire. L'ensemble de ces facteurs viendra réguler l'activité transcriptionnelle.

Nous décrierons ici les grandes étapes de l'initiation de la transcription d'un gène transcrit par l'ARN polymérase II (donc transcription d'un ARNm ou d'un miARN).

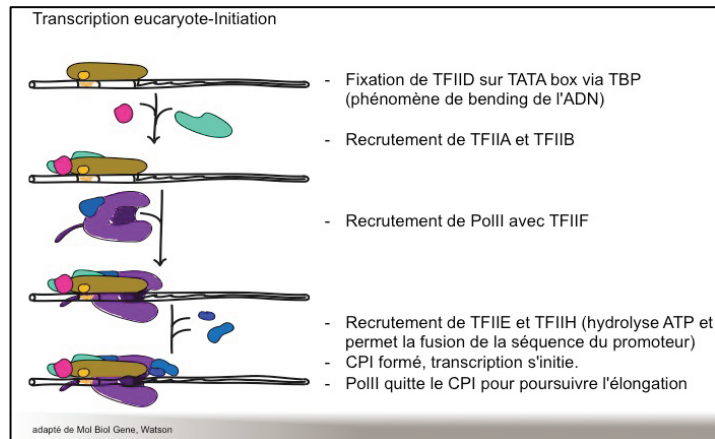
Le complexe de transcription eucaryote est composé de l'ARN polymérase proprement dite et d'un certain nombre de Facteurs Généraux de la Transcription.

- Le facteur TFIID, en coopération avec TBP, s'associe à la séquence TATA et conduit au recrutement de TFIIA et TFIIB.

- L'ARN pol-II et TFIIF s'associent à ce complexe et recrutent TFIIE et TFIIH.

- Le complexe de pré-initiation est formé et initie la transcription

- L'ARN polII poursuit l'élongation de la transcription en se dissociant du CPI.



Les mécanismes d'élongation de la transcription eucaryotes sont, dans les grandes lignes, assez semblables aux mécanismes procaryotes.

Toutefois, en cours d'élongation le complexe s'associera à des facteurs protéiques spécifiques qui faciliteront l'élongation en présence d'histones (le complexe FACT), qui permettront des modifications de l'ARN (coiffe et maturation du transcrit).

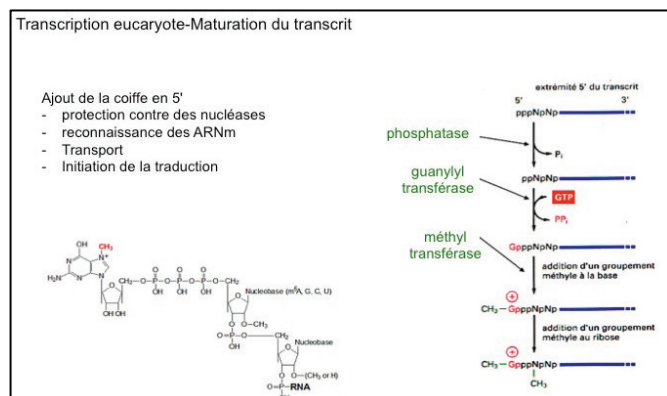
Les mécanismes de terminaisons sont moins bien définis que chez les procaryotes. La terminaison est concomitante à un clivage du transcrit et à sa poly-adénylation, indiquée par une séquence 5'-AAAUAAA-3'.

La maturation des ARNm eucaryotes est réalisée en plusieurs étapes.

En tout début de transcription, une coiffe, formée d'un 7-methyl-guanylate est ajoutée à l'extrémité 5'-phosphate libre du transcrit par une liaison 5'-5'.

Cette coiffe a différentes fonctions :

- protection

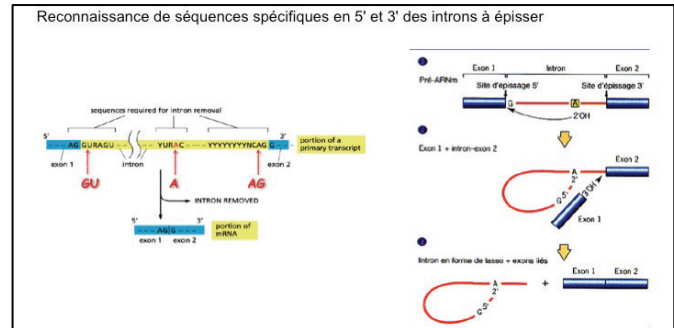


- identification des ARNm
- transport nucléo-cytoplasmique
- initiation de la traduction.

Beaucoup de gènes eucaryotes sont des gènes morcelés composés de séquences exoniques (codantes) et de séquences introniques (non codantes). En fin de transcription, le transcrit primaire, coiffé et poly-adiénylé, contient les exons et les introns. La maturation du messager consistera à éliminer les introns par épissage.

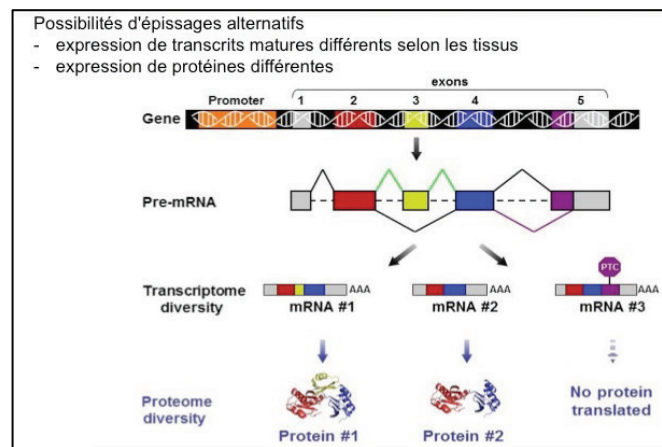
2 mécanismes d'épissages sont décrits:

- auto-épissage, activité enzymatique portée par le transcrit lui-même, c'est une activité ribozyme.
- Epissage par le Spliceosome, important complexe macro-moléculaire faisant intervenir des Small Nuclear Ribo Particles (SNURPs).



Une séquence GU en 5' de l'intron constituera le site donneur d'épissage et une séquence AG en 3' de l'intron formera un site accepteur d'épissage.

Selon l'environnement cellulaire et la transmission de divers signaux que nous ne développerons pas, l'épissage peut conduire à la maturation de différents messagers à partir d'une seule unité de transcription. Ainsi, selon le contexte, certains exons peuvent être épissés en même temps que les introns en utilisant ou en sautant des sites d'épissage. Nous parlons alors d'épissage alternatif. Ce mécanisme permet l'expression de transcrits et donc de protéines différents selon le contexte cellulaire.



IV LA TRADUCTION

Le code génétique

Les mécanismes de traduction sont basés

sur un code génétique dont la clé est parfaitement définie. Ce code génétique suit des règles strictes.

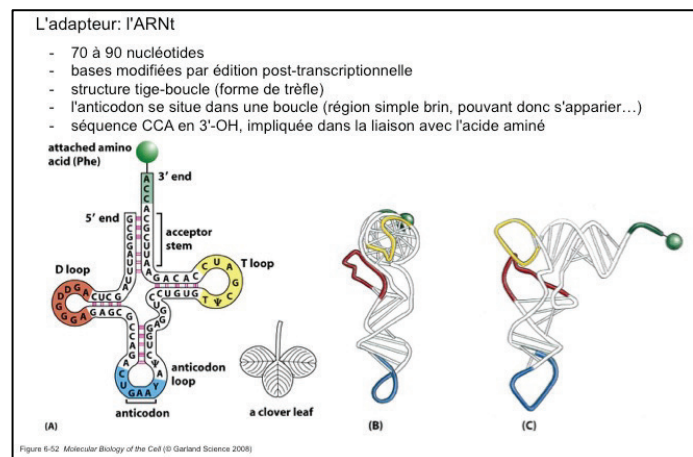
- 1- le code est lu sur l'ARNm de 5'-3'
- 2- Le code est lu par triplets
- 3- Le code est non chevauchant
- 4- Le code est non-ambigu
- 5- Le code est dégénéré
- 6- Le code est universel
- 7- le code comporte des signaux de début et de fin de lecture.

Toute molécule d'ARNm contient donc 3 cadres possibles de lecture. C'est le codon d'initiation (AUG) suivi de codons puis d'un codon stop qui spécifiera un cadre ouvert de lecture (Open Reading Frame = ORF).

le code génétique												
		Deuxième lettre										
		U		C		A		G				
U	Première lettre (côté 5')	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U	C	A
		UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys	C	C	A
		UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop	A	C	A
		UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp	G	C	A
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U	C	A	
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg	C	C	A	
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg	A	C	A	
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg	G	C	A	
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U	C	A	
	AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser	C	C	A	
	AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg	A	C	A	
	AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg	G	C	A	
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U	C	A	
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly	C	C	A	
	GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly	A	C	A	
	GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly	G	C	A	
		codon d'initiation				codon de terminaison						

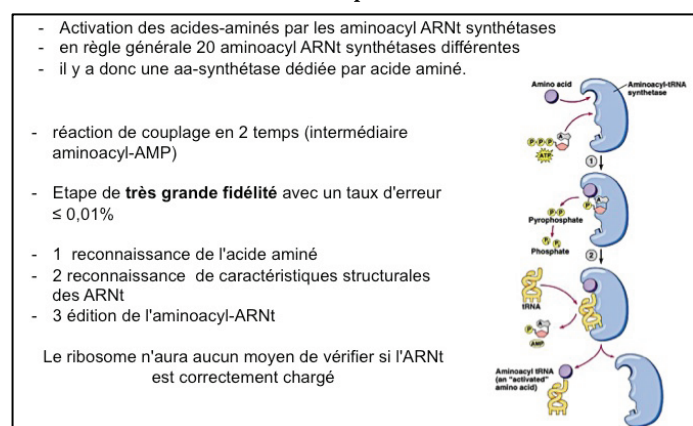
La molécule d'ARN de transfert joue le rôle d'adaptateur entre l'ARNm et le ribosome qui synthétise un polypeptide. La molécule d'ARNt adopte une conformation tridimensionnelle particulière à base de tiges dues à des appariements intramoléculaires et de boucles simples brins. La deuxième boucle porte l'anticodon, complémentaire et antiparallèle du codon porté par l'ARNm.

Enfin, l'ARNt est couplé covalamment à la fonction acide d'un acide aminé, par sa séquence CCA en 3'.



Les amino acyl ARNt synthétase sont les éléments centraux dans le respect et la fidélité du code génétique. Ces enzymes sont responsables de la réaction de couplage d'un acide aminé sur un ARNt. Cette réaction se fait en deux temps. Dans un premier temps l'Amino Acylt ARNt synthétase lie spécifiquement un acide aminé et en présence d'ATP forme un intermédiaire Amino-Acyl-AMP et libère un P_i. Il existe en général à peu près autant d'Amino Acyl ARNt synthétase que d'Acides aminés (20). Chaque AA-ARNt-synthétase est donc spécifique d'un Acide Aminé.

Dans un second temps, l'AA-ARNt-Synthétase associe un ARNt selon des caractéristiques structurales. Des capacités d'édition sont requises. Plusieurs ARNt peuvent être reconnus du fait de la

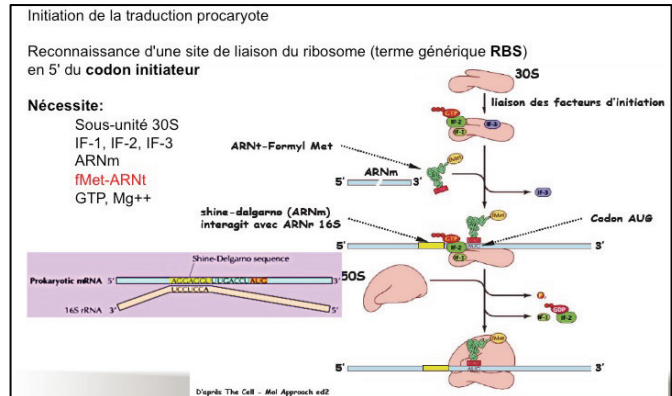


dégénérescence du code. L'enzyme catalyse le couple de l'amino-acyl-AMP sur le 3'OH de l' ARNt et libère un AMP et un Amino-Acyl-ARNt.

La clé du code génétique est donc en grande partie détenue par les amino-acyl ARNt synthétases.

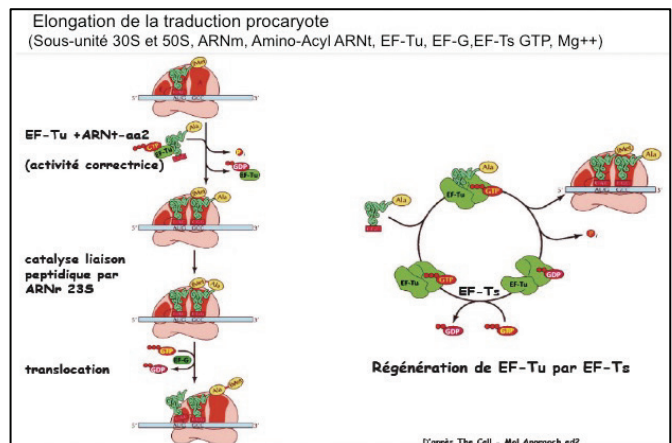
Le ribosome est un imposant complexe macromoléculaire composé de protéines et d'ARNr (la composition est décrite dans les documents de cours).

L'initiation de la traduction procaryote est réalisée par la petite sous unité 30s associée aux facteurs d'initiation IF1, IF2-GTP et IF3. Ce complexe s'associe à un ARNm sur une séquence consensus nommée RBS (Shine-Delgarno chez E. coli) située en 5' du codon initiateur. Un fMet-ARNtMet s'associe au complexe sur l'AUG à proximité du RBS concomitamment au départ de IF3. La grosse sous unité 50s vient se positionner alors que IF1 se dissocie. IF2 hydrolyse le GTP et le complexe IF2-GDP quitte le complexe. La phase d'initiation est terminée avec un ribosome placé sur le codon initiateur et comportant un fMet-ARNtMet chargé dans le site P, le site A est libre.



Lors de l'élongation, un complexe (amino-acyl-ARNt)-(EFTu-GTP), vient se placer dans le site A du ribosome. Cette association est rendue possible par la complémentarité codon-anticodon. L'hydrolyse du GTP et le départ de EFTu-GDP permet la stabilisation de l'association de l'amino-acyl ARNt.

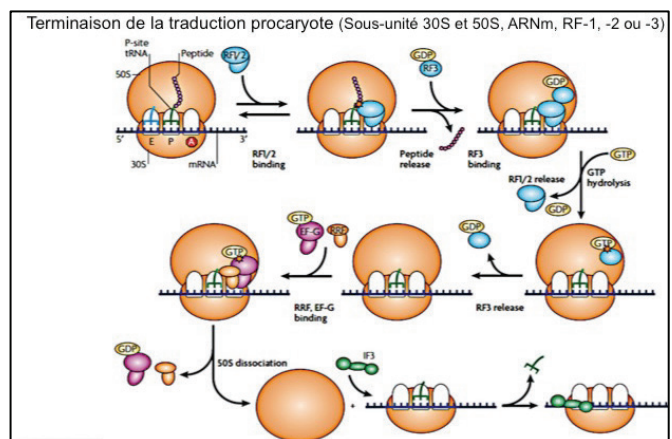
L'ARNr 23s catalyse la formation de la liaison peptidique. Le peptide est désormais lié à l'ARNt placé dans le site A, l'ARNt placé dans le site P est libre. Le facteur EF-G-GTP hydrolyse le GTP et favorise le déplacement du ribosome sur le transcrite vers le codon suivant. L'ARNt libre se situe alors dans le site E, le peptidyl-ARNt est dans le site P et le site A est à nouveau libre. Cette étape d'élongation se répétera jusqu'à ce qu'un codon stop soit dans le site A.



Il faut noter que le facteur EFTu-GDP est à nouveau chargé d'un GTP grâce à l'action du facteur d'échange EF-Ts.

La catalyse de la liaison peptidique, par l'ARNr 28s, est une activité ribozyme. Cette réaction utilise la liaison riche en énergie portée par l'aminoacyl-ARNt.

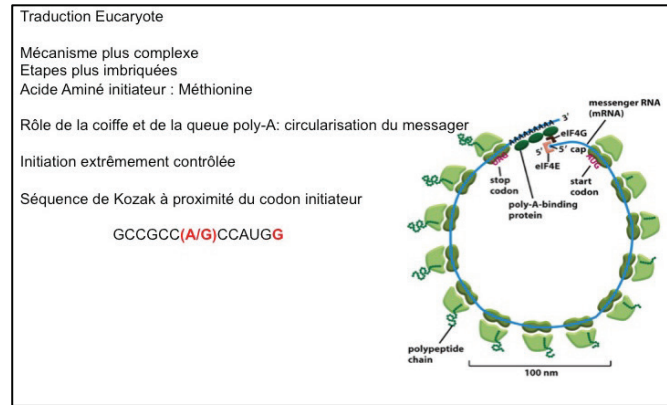
La terminaison est catalysée par de petits facteurs peptidiques appelés "Release Factors" dont les facteurs RF1 et RF2. Ces facteurs présentent une structure leur permettant de mimer un ARNt et entrent dans le site A libre, en opposition au codon stop. Le facteur RF3-GTP est alors chargé et hydrolyse son GTP. Le facteur EF-G-GTP hydrolyse à son tour un GTP et va conduire à la dissociation de l'ensemble du complexe. Le transcrite est libéré, ainsi que le peptide neosynthétisé.



Chez les eucaryotes, l'initiation est légèrement plus complexe et moins bien définie. La coiffe et la queue poly A jouent un rôle important dans l'initiation.

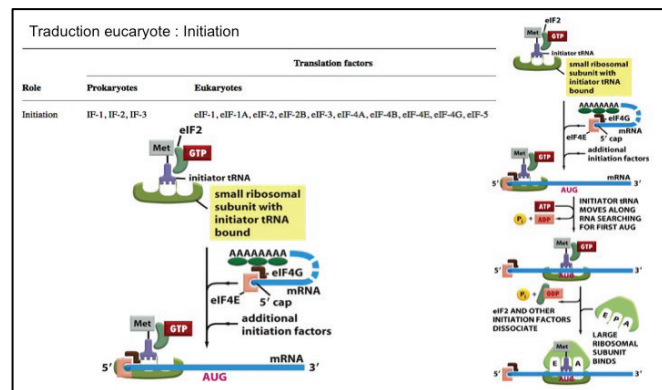
La méthionine est l'acide aminé initiateur (à la différence de la formyl-méthionine chez les procaryotes).

La petite sous unité 40s se place sur une séquence consensus nommée Kozak en association avec un méthionyl-ARNtMet et un facteur eIF2-GTP. Ce complexe va "scanner" le transcrit jusqu'à rencontrer un codon AUG en 3' de la séquence Kozak.



L'initiation s'achèvera par l'assemblage de la sous unité 60s.

Le nombre de facteur d'initiation est bien plus important chez les eucaryotes et cette étape est très finement régulée.



Nous considérerons que les mécanismes d'élongation et de terminaison sont similaires entre eucaryotes et procaryotes.

V EXPRESSION DES GENES

La régulation de l'expression des gènes est multifactorielle et est régulée en faisant appel à la fois à ce que nous nommons des éléments CIS et à des éléments TRANS.

- éléments CIS: séquences d'ADN permettant une interaction avec des facteurs protéiques (TRANS) et participant à la régulation de gènes situés à proximité ou à distance. Eléments également nommés "**opérateur**".
- éléments TRANS: facteurs protéiques se fixant à l'ADN et participant à la régulation de l'expression de gènes, par activation ou inhibition de l'activité de l'ARN polymérase.
Ces facteurs peuvent activer ou inhiber l'activité transcriptionnelle du complexe ARN pol.

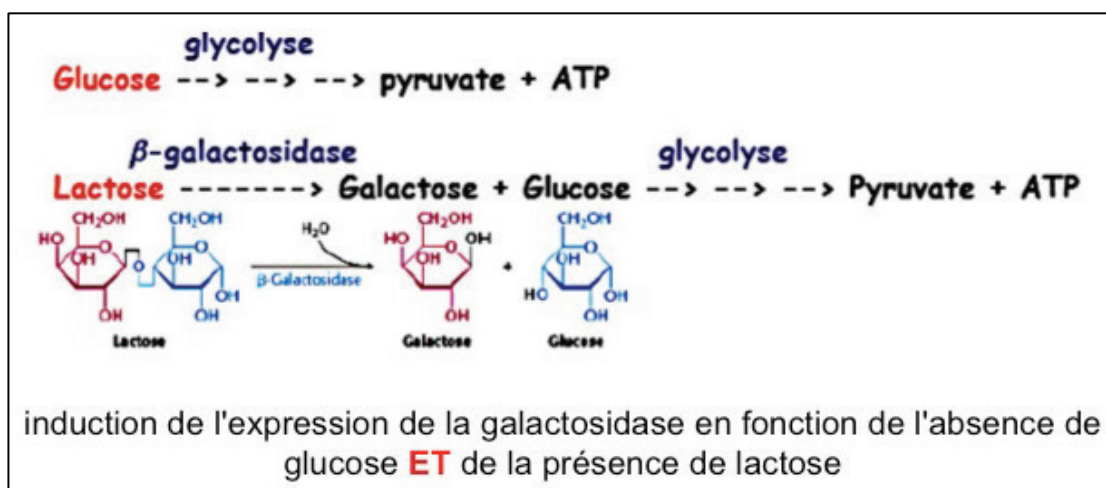
Exemples de régulation d'expression chez les procaryotes.

Dans le cadre de ce cours, 2 mécanismes seront à retenir:

- 1- l'opéron lactose
- 2- l'opéron tryptophane

Un opéron est une région du génome regroupant des gènes placés sous le contrôle d'un même promoteur. La transcription de ces opérons génère des messagers polycistroniques. Ce sont des messagers qui portent plusieurs Ribosome Binding Sites et permettent ainsi la traduction de plusieurs polypeptides.

L'opéron lactose est induit par une levée de répression catabolique et permet l'utilisation du lactose comme source de carbone si et seulement si le glucose est absent.



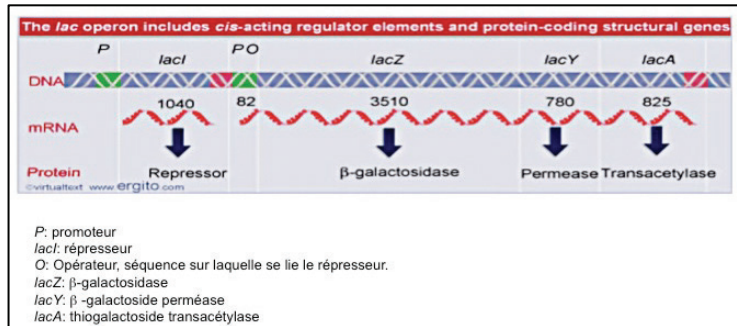
L'opéron lactose est organisé dans le génome procaryote comme schématisé ci-dessous. L'expression du gène lacI est constitutive et placée sous le contrôle d'un promoteur constitutif (qui fonctionne en permanence). Le produit du gène lacI est une protéine "répresseur". C'est un élément TRANS.

L'opéron lactose à proprement parler comporte un promoteur P, régulé par une séquence "Opérateur" O, c'est un élément CIS.

Le répresseur se lie à l'opérateur en absence de lactose et inhibe la transcription. En présence de lactose, celui-ci se lie au répresseur qui subit une modification de conformation allostérique. Le répresseur ne se lie plus à l'opérateur et la transcription peut avoir lieu.

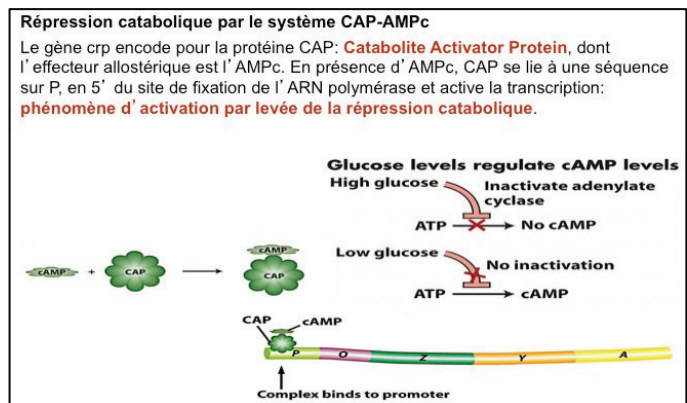
Le message transcrit permet de coder pour trois protéines qui permettent de métaboliser le lactose.

Le système est donc induit par une levée d'inhibition.

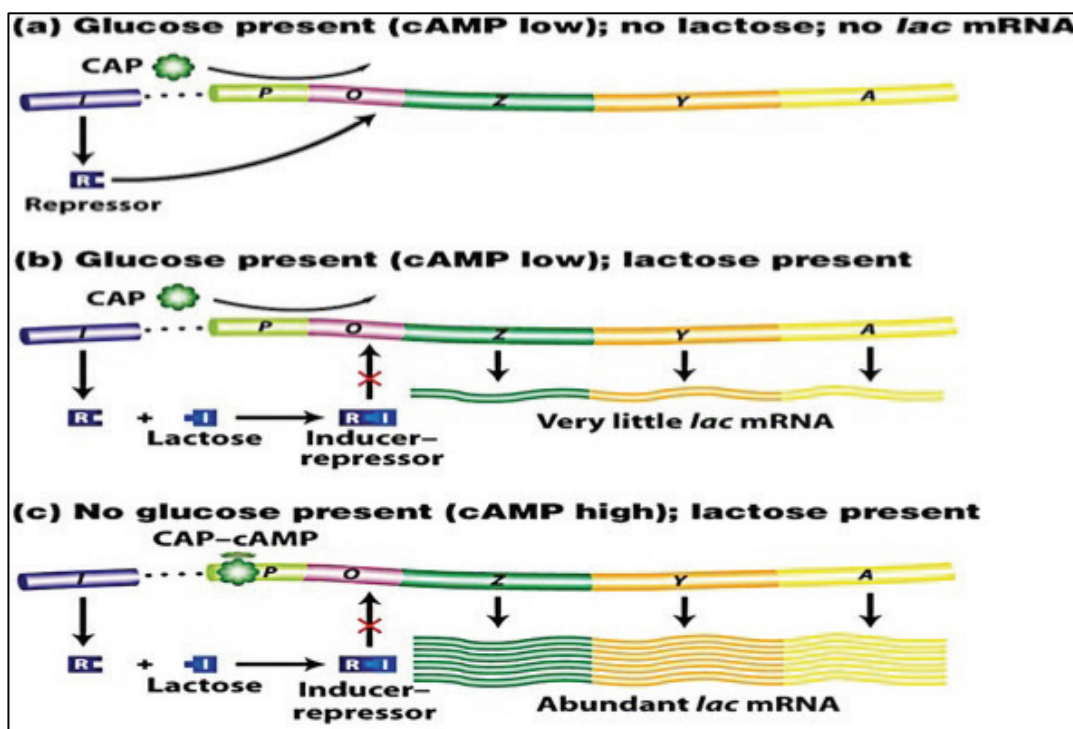


Un autre mécanisme, partagé avec d'autres opérons, vient s'ajouter à ce mécanisme de régulation. En effet, dans un but d'économie d'énergie, l'opéron lactose ne doit être induit que si le glucose vient à manquer. Il existe un système de répression catabolique basé sur le système CAP-AMPC qui joue le rôle de "senseur" pour le glucose.

En présence de glucose l'adénylate cyclase est inhibée. Par conséquent la concentration intracellulaire d'AMPC est faible. La protéine CAP (Catabolite Activator Protein), produit du gène *crp*, ne peut pas se lier au promoteur et ne peut pas activer la transcription au niveau du promoteur. En absence de glucose l'adénylate cyclase est activée. Par conséquent la concentration d'AMPC augmente. Cet AMPC se lie à CAP et le complexe CAP-AMPC active la transcription.

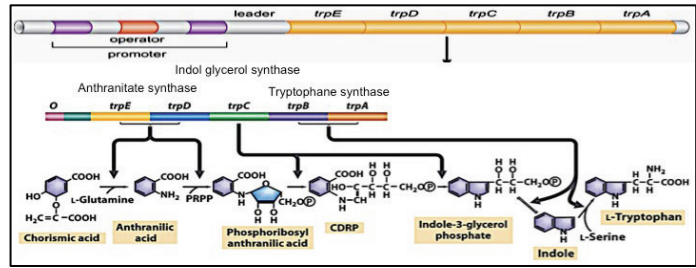


Ainsi, l'organisme procaryote peut se trouver dans les trois situations décrites ci-dessous:

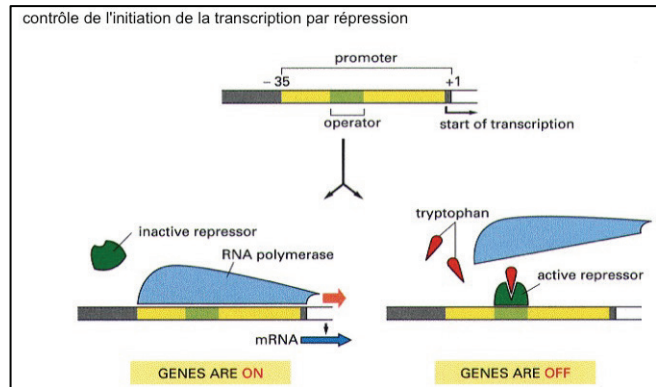


L'opéron tryptophane est un autre type d'opéron qui permettra l'expression de 5 enzymes différentes impliquées dans la synthèse de tryptophane à partir d'acide chorismique lorsque le tryptophane vient à manquer.

Cet opéron est contrôlé par un système de répression de l'initiation de la transcription et par un mécanisme d'atténuation qui contrôle la terminaison de la transcription.



Le promoteur de l'opéron tryptophane contient une séquence régulatrice sur laquelle peut se fixer un répresseur qui est actif lorsque lié à du tryptophane. Dans ce cas, en présence de tryptophane, le répresseur empêche la fixation de l'ARN polymérase et l'initiation de la transcription est inhibée.



En absence de tryptophane, la transcription est initiée et le transcrit est synthétisé. La région 5' du transcrit comporte une séquence appelée "leader" qui contient 2 codons tryptophane et une région appelée "atténuatrice". Cette région comporte 4 courtes séquences complémentaires permettant des repliements par hybridation intra-moléculaire.

Plusieurs combinaisons sont possibles qui permettront d'adopter différentes conformations.

Comme vous le savez, la traduction a lieu avant même la terminaison de la transcription chez les procaryotes. La compartimentation nucléo-cytoplasmique interdit cette traduction co-transcriptionnelle chez les eucaryotes.

Ainsi, le ribosome commence à traduire le transcrit en cours de synthèse et 2 possibilités se présentent:

1- En présence de tryptophane en quantité suffisante, le ribosome passe rapidement les 2 codons "Trp" rencontrés dans la séquence leader. Le ribosome avance donc rapidement et empêche l'hybridation de la séquence 2 avec la séquence 3. La séquence 3 s'hybride donc avec la séquence 4 et forme une structure "tige-boucle" qui va terminer la transcription en chassant l'ARN polymérase. Donc, en présence de tryptophane la transcription de l'opéron tryptophane est inhibée, logique.

2- En absence de tryptophane, le ribosome va ralentir (caller) sur les deux codons tryptophane. Dans ce cas, la séquence 2 s'hybride à la séquence 3 et la transcription se poursuit.

