

La génétique

Cours explicatif pour les étudiants BAC1 en faculté de Médecine

random]pLasnLd

Rédigé à partir des notes de cours, *BIOLOGIE (Campbell) et Fundamentals of Biochemistry (D. Voet)*

SECTION 4 – LA GENETIQUE

BERTRAND BOUILLON

La génétique est la science de l'hérédité et des gènes, caractères transmissibles de générations en générations entre individus. La génétique tente également de donner une explication à l'évolution des espèces, les nouveaux caractères et l'émergence de nouvelles espèces ainsi que l'adaptation des organismes. On peut diviser la génétique en deux parties : la génétique mendélienne, qui ne s'intéresse qu'à la transmission des caractères et traits d'un organisme de générations en générations et la génétique moléculaire.

La génétique est une branche de la biologie qui est encore très récente ; Son arrivée en tant qu'une branche scientifique reconnue se doit à Gregor Mendel, moine catholique et botaniste. Gregor Mendel est incontestablement le père de la génétique, il formule les **lois de Mendel** qui définissent la manière dont les gènes se transmettent de générations en générations.

De nombreuses théories sur la transmission des caractères de générations en générations ont émergé au XVIII et XIXe siècle bien qu'il y ait eu des théories qui remontent à la Grèce Antique, notamment Hippocrate et sa théorie des humeurs qui tente d'expliquer la transmission des caractères par l'émergence des humeurs dans les organes génitaux, formant des semences. Cette conception hippocratique sera reprise par Maupertius et Darwin, qui formuleront deux hypothèses concurrentes.

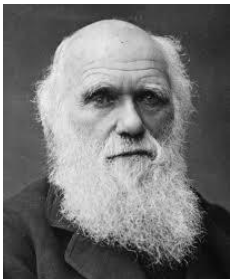


Figure 0-1 - Charles Darwin, père de la théorie de l'évolution

L'une des premières théories de la génétique était la Pangenèse, formulée par Charles Darwin pour expliquer les mécanismes de l'hérédité qui sont au cœur de la **théorie de l'évolution** (De l'origine des espèces, Charles Darwin). Cette hypothèse de Darwin reprend la conception Hippocratique selon laquelle l'ensemble de l'organisme parent participe à l'hérédité en essayant de l'adapter à la **théorie cellulaire** de Schwann. Il présente cette 'hypothèse provisoire' dans son travail « La variation des animaux et des plantes sous domestication ».

Darwin pense que les cellules de l'organisme bourgeonnent en émettant des gemmules, libérées dans le milieu extra-cellulaire et dans la circulation générale. Ces gemmules auraient la capacité de s'agréger en forme de polybourgeon. Il souligne que seules les cellules peuvent régénérer des tissus, ou générer de nouveaux organismes et propose que des gemmules, de taille atomique, s'agrègent et se dispersent dans les organes reproductifs de l'organisme.

Un peu plus tard dans le même siècle, un jésuite et naturaliste français du nom de Jean Baptiste de Lamarck propose une idée nouvelle de la transmission des caractères selon laquelle les organismes se transforment pour s'adapter à leur environnement. Cette théorie transformiste est basée sur plusieurs principes :

1. Les animaux tendent, lors de l'évolution, à se complexifier et à accroître en volume.
2. La diversification des êtres vivants se fait sous l'effet des circonstances variées auxquelles ils sont confrontés dans des milieux variés et auxquels ils sont contraints de s'adapter en modifiant leur comportement ou leurs organes pour répondre à leurs besoins.
3. Le développement de ces organes et leur force d'action sont constamment en raison de l'emploi de ces organes
4. Tout ce qui a été acquis, tracé ou changé, dans l'organisation des individus, pendant le cours de leur vie, est conservé par la génération, et transmis de génération en génération.



Cette théorie de l'évolution des organismes est erronée : Dans son exemple, la girafe n'a pas vu son cou s'allonger par adaptation, mais les mutations génétiques étaient favorables aux girafes dont le cou était allongé. Les mécanismes d'évolution proposés par Lamarck étaient différents de ceux proposés par Darwin, ce dernier s'étant avéré avoir raison.

Jean Baptiste Lamarck, dans son erreur, n'avait pas tout à fait tort ; Sa théorie est à la base de la transmission des caractères acquis, théorie également admise et énoncée par Darwin. Le néo-lamarckisme refait dernièrement surface avec la découverte de la régulation de l'expression génétique en fonction de l'environnement ; on parle de l'épigénétique.

Le travail de Mendel sur la transmission des caractères n'a, au départ, aucun rapport avec celui de Darwin. Mendel ne comprend pas lui-même ce qui est à la base de la transmission des caractères mais il formule simplement les croisements de traits entre des pois et tente de formuler une loi. Le lien entre le travail de Mendel et celui de Darwin, qui sont deux travaux à des échelles différentes, sera établi bien plus tard grâce à l'avancée de la technologie et la découverte de l'ADN.



ADN, gènes et chromosomes

ACIDE NUCLEIQUE, ADN ET ARN

La structure primaire de la protéine est à la base de la conformation entière. Cette structure primaire est déterminée par notre information génétique, un gène qui code pour la séquence primaire de la protéine. Ce gène est lui-même une séquence mais dont les unités monomériques sont des nucléotides. Le polymère entier de nucléotide est un acide nucléique et peut exister sous deux formes : l'acide désoxyribonucléique et l'acide ribonucléique.

Les nucléotides sont des molécules ubiquitaires qui participent à presque tous les processus biochimiques :

- Ce sont les unités monomériques de l'acide nucléique
- Les nucléosides triphosphate, notamment l'ATP, sont les des molécules riches en énergie qui libèrent cette énergie et dont l'utilisation rend la plupart des processus endergoniques possible.
- La plupart des voies métaboliques sont régulées par les concentrations de nucléotides comme l'ATP, l'ADP et l'AMP De nombreux signaux hormonaux sont transmis à l'intérieur de la cellule par l'AMPC et son analogue guanylique le GMPC.

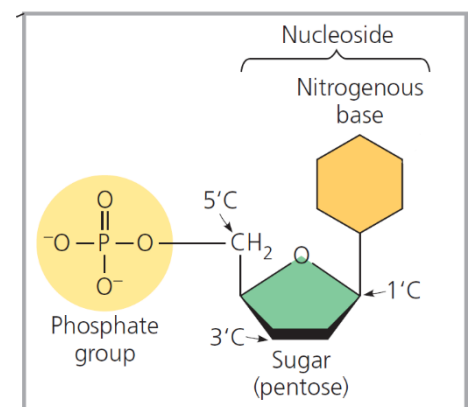
Les deux acides nucléiques (ADN et ARN) sont des molécules complexes qui permettant aux organismes de diriger et mettre le métabolisme de la cellule en marche. L'ADN fournit les directives et est relayé par l'ARN messenger, un ARN qui va être décodé par les ribosomes pour être traduit en une protéine, molécule fonctionnelle du métabolisme. L'ADN est une macromolécule gigantesque, polymère de nucléotides. L'ADN porte des milliers de gènes qui constituent une portion de ce polymère, occupant une portion spécifique le long de la molécule. Quand la cellule se divise, l'ADN est répliqué et transmis aux cellules filles.

NUCLEOTIDE

Les nucléotides participent aux réactions d'oxydo-réduction, transferts d'énergies, signalisation intra-cellulaire et dans les réactions biosynthétiques. Leurs polymères, l'ADN et l'ARN, sont les molécules constituant le stockage de l'information génétique de la cellule ainsi le décodage de cette information. Les nucléotides sont des molécules ubiquitaires qui contiennent une véritable diversité structurale. Un polynucléotide (polymère de nucléotide) possède des monomères (nucléotides) possédant un groupement phosphate. Si un nucléotide ne possède pas de groupement phosphate, il s'agit d'un nucléoside.

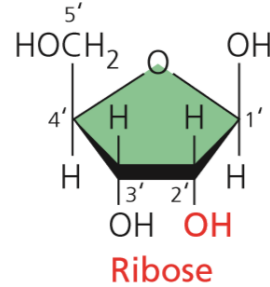
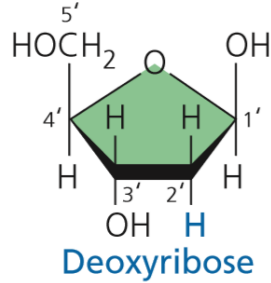
Les nucléotides se lient les uns aux autres pour former des polymères via des liaisons phosphodiesteres. Dans la composition d'un nucléotide, il y a un groupement phosphate, un sucre pentose et une base azotée. Il existe 5 bases azotées différentes et deux pentoses possibles dans la composition du nucléotide, à l'origine de la diversité de ce monomère.

Par convention, on numérote les carbones du pentose des nucléotides. Les atomes de carbones sont numérotés de 1' à 5' en allant dans le sens des aiguilles d'une montre. Le carbone 1' est lié à la base azotée, le carbone 3' est lié à un groupement hydroxyle et le carbone 5' à un groupement phosphate. Un nucléotide va pouvoir se polymériser en une longue chaîne d'acide nucléique au niveau du carbone 3' et 5'. C'est grâce à cette numérotation que l'on définit le sens de l'acide nucléique qui nous servira pour plus tard.

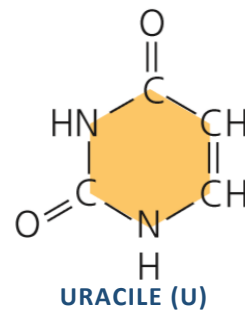
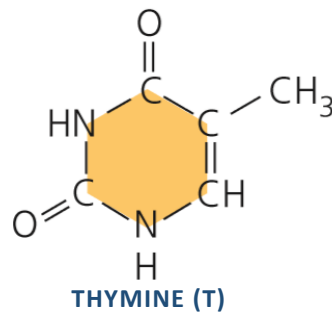
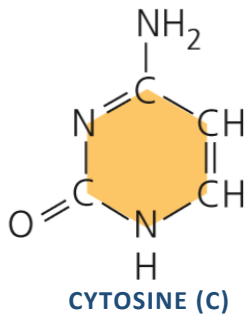


Il existe 8 variétés majeures de nucléotides, chacune composée en 3 parties :

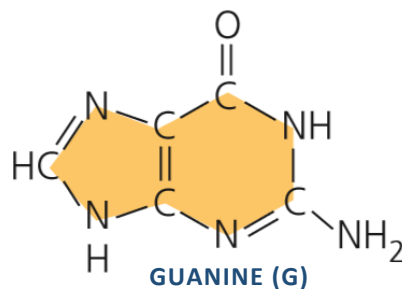
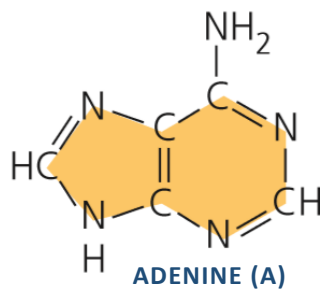
- Un pentose (sucre de 5 carbones) définissant le type de polymère :
 - Ribose : Si le sucre est un ribose alors l'acide nucléique constitué de ces nucléotides est un acide ribonucléique (ARN)
 - Désoxyribose : Il s'agit d'un pentose dont le groupement OH a été dépourvu de son oxygène pour laisser place à l'hydrogène en C2. L'acide nucléique dont les résidus monomériques portent un sucre de désoxyribose est un acide désoxyribonucléique (ADN).



- Un groupement phosphate, lié au pentose par une liaison covalente
- Une base azotée, liée au pentose. Il existe 5 bases azotées qu'on regroupe en deux familles :
 - Pyrimidine : hétérocycle de 2 azotes
 - Cytosine (C)
 - Thymine (T), uniquement présent dans l'ADN
 - Uracile (U), uniquement présent dans l'ARN



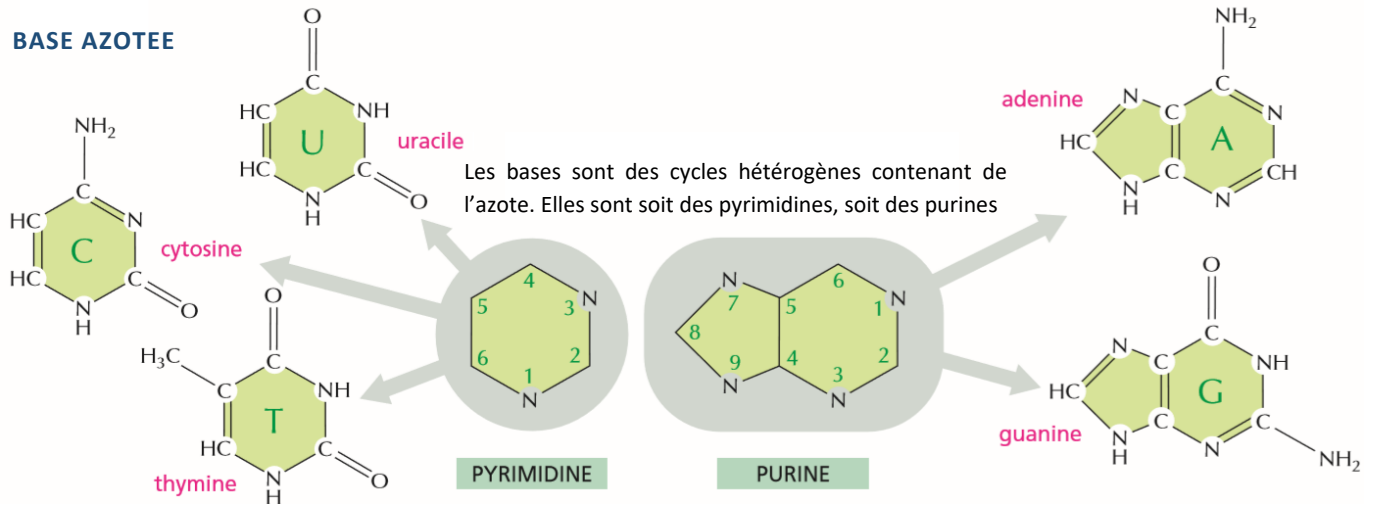
- Purine : composé hétérocyclique azotée constitué d'un cycle de pyrimidine fusionné à un cycle d'imidazole.
 - Adénine (A)
 - Guanine (G)



Les purines forment des liaisons au pentose via leur azote en 9^e position du cycle tandis que la pyrimidine forme une liaison au pentose via l'azote en position 1. Les nucléotides dont le pentose est un ribose sont des ribonucléotides et les nucléotides dont le pentose est un désoxyribose sont des désoxyribonucléotides (ou juste déoxynucléotide)

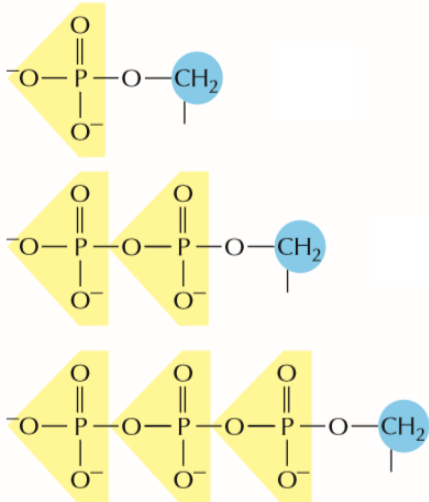
Les nucléotides, quand ils sont résidus de l'acide nucléique, ne possèdent qu'un seul groupement phosphate. Mais le nucléotide peut avoir d'autres fonctions que celle de résidu pour un polynucléotide quand il est sous la forme libre. Le nucléotide peut comporter plusieurs groupements phosphate comme l'adénosine triphosphate qui est un nucléoside lié à 3 groupements phosphates. Ces groupements phosphate sont liés à l'atome C3' ou C5' du pentose pour former un 3'-nucléotide ou un 5'-nucléotide, respectivement.

BASE AZOTEE



PHOSPHATES

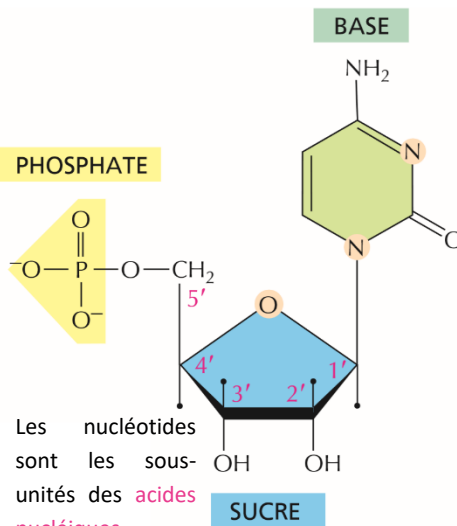
Les phosphates sont liés à l'hydroxyle C5 du sucre ribose ou désoxyribose (désigné 5'). On retrouve des mono, di et triphosphates.



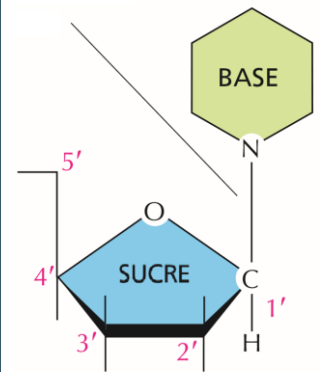
Les phosphates rendent les nucléotides chargés négativement

NUCLEOTIDE

Un nucléotide consiste en une base azotée et d'un sucre à 5 carbone (furopentose) et un ou plusieurs groupes phosphates.



LIAISON BASE-SUCRE

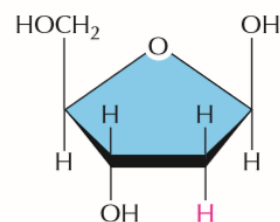
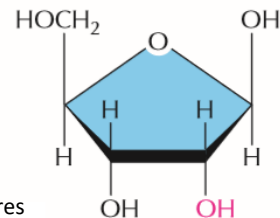
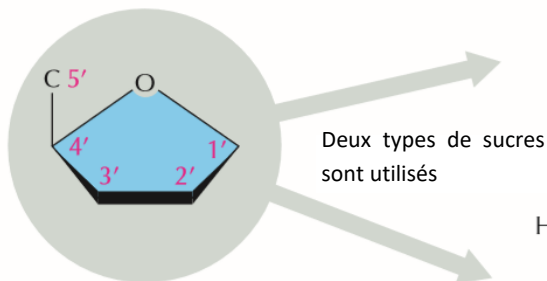


La base est liée au carbone C1 du ribose ou du désoxyribose.

SUCRES

PENTOSE

Un sucre à 5 carbones



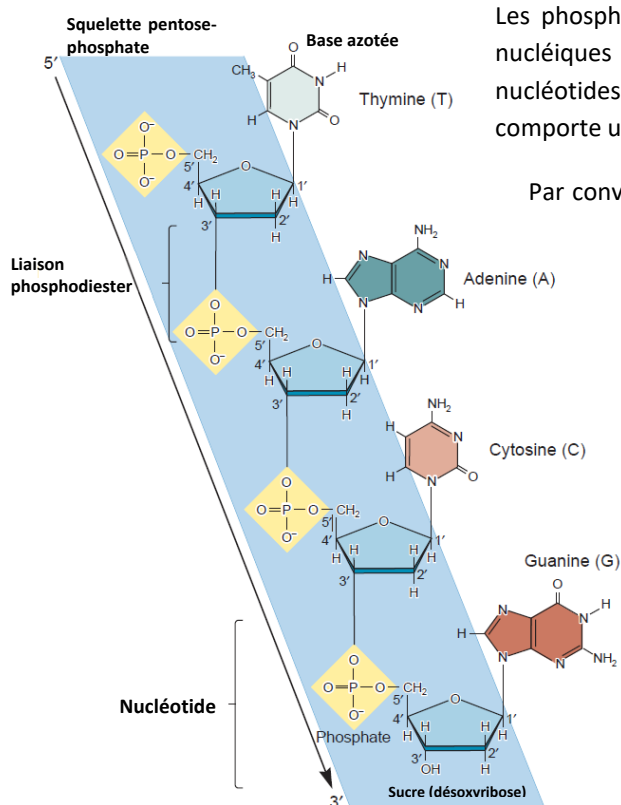
Les groupements phosphate sont le plus souvent liés en C5' et sont appelés nucléoside monophosphates, biphosphate ou triphosphate. Les nucléotides libres sont sous forme anionique (*nucleotide*²⁻) sont presque toujours associés avec un cation Mg^{2+} dans la cellule. Les nucléotides libres et les dérivés de nucléotides réalisent une énorme variété de fonctions métaboliques qui ne sont pas liées à la gestion de l'information génétique.

La polymérisation des nucléotides en acides nucléiques se fait via une liaison phosphodiester : la liaison à lieu entre le groupe 5'-phosphate d'un nucléotide et le groupe hydroxyle 3' d'un autre nucléotide

ACIDE NUCLEIQUE

Les acides nucléiques, ou polynucléotides, sont des polymères dont les résidus monomériques sont des nucléotides liés par des liaisons phosphodiester. Les résidus nucléotides sont constitués d'un nucléoside lié à un seul groupement phosphate. La liaison phosphodiester lie les nucléotides via le carbone en position 3' du pentose d'un nucléotide au groupement phosphate d'un autre nucléotide. Ce phosphate est lié au C5 du pentose du nucléoside auquel il est lié : le sens du nucléotide est alors 5' → 3'.

Les phosphates de ces polynucléotides sont acides : à pH physiologique, les acides nucléiques sont des polyanions. Chaque phosphate est estérifié à deux pentoses de nucléotides voisins, excepté aux extrémités de la séquence polynucléotidique qui comporte une extrémité 5' et une extrémité 3'.



Par convention, la séquence de résidus de nucléotides dans un acide nucléique est écrite, de gauche à droite, de l'extrémité 5' à l'extrémité 3'. Le motif résultant de la répétition des liaisons phosphodiester entre les résidus pour former l'acide nucléique est nommé le squelette pentose-phosphate. Le long de ce squelette, se trouve une chaîne de bases azotées portées par chaque résidu de nucléotide.

La séquence de bases le long de l'ADN est unique pour chaque gène et fournit une information très spécifique à la cellule qui sera retranscrit sur un ARN messager (ARNm) qui servira de matrice aux ribosomes qui ont pour rôle de traduire cette séquence nucléotidique spécifique en une protéine fonctionnelle.

Le gène est une séquence de nucléotide de l'ADN, un polynucléotide de taille énorme qui renferme une immense quantité de gènes. Un gène représente plusieurs centaines à plusieurs milliers de résidus monomériques. Sachant que chaque monomère peut se différencier par une des quatre bases azotées possibles, le nombre de possibilités de

séquences différentes pour un gène est sans limite. Pour un gène à seulement 100 résidus nucléotidiques par exemple, il existe 4^{100} ($1,607 \cdot 10^{60}$) combinaisons possibles.

Chaque séquence a une signification différente à la cellule, qui sera traduite lors de la protéosynthèse. Une séquence de 3 ribonucléotides forme un codon, ce codon a une signification pour les ribosomes. Un ou plusieurs codons correspondent à un acide aminé spécifique, le ribosome se sert de l'ARNm qui n'est que la retranscription du gène spécifique qui sert de matrice pour former les protéines et assembler les acides aminés bouts à bouts.

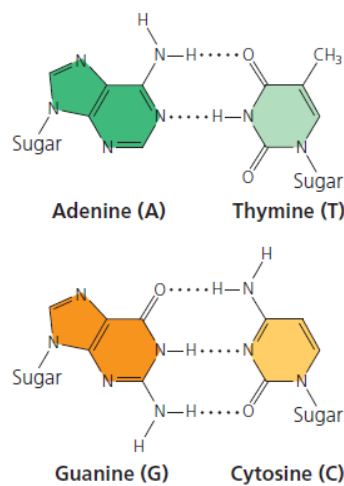
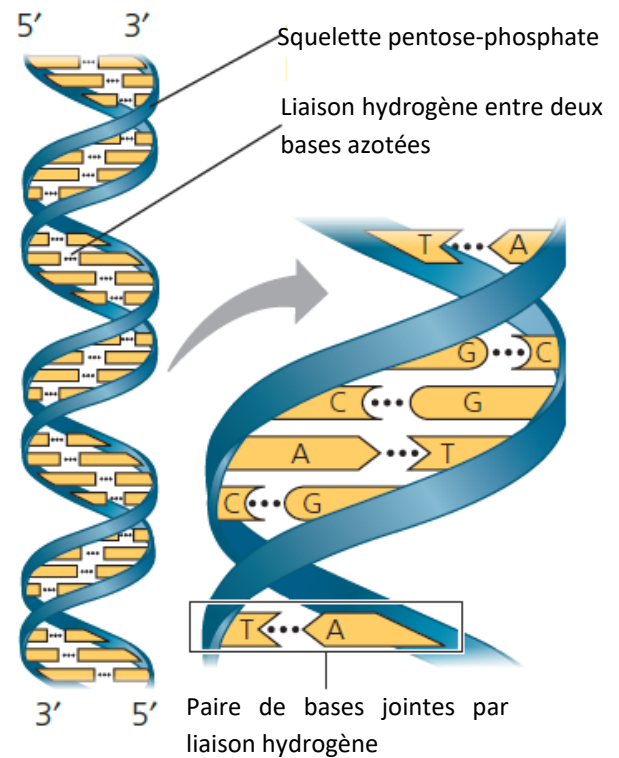
Le sens 3' et 5' est défini par rapport au sucre des nucléotides : chaque carbone du pentose de chaque nucléotide est identifié. Le carbone C1, noté 1', est lié à la base azotée, le carbone 3' à un hydroxyle et le carbone 5' à un phosphate. Les nucléotides se polymérisent par une réaction de condensation, en fixant le phosphate d'un nucléotide voisin, lié à son pentose en 5', au groupement hydroxyle situé en 3' du pentose. La condensation libère une molécule d'eau et aboutit à la formation d'une liaison phosphodiester. Nous verrons plus en détail ce processus dans le cadre de la répllication de l'ADN. Un acide nucléique a donc une extrémité 3'(OH) et une extrémité 5'(PO₄²⁻).

L'ACIDE DESOXYRIBONUCLEIQUE (ADN)

L'ADN, renfermé dans le noyau chez les eucaryotes, est une macromolécule qui renferme deux polynucléotides. Chaque polynucléotide est un « brin », les deux brins s'enroulent en spirale autour d'un axe imaginaire pour former une double hélice dont les deux squelettes pentose-phosphate se retrouvent sur les bordures extérieures de l'hélice tandis que les bases azotées sont confinées à l'intérieur et se font face à l'intérieur et s'apparient via des liaisons hydrogènes.

Chaque base azotée est appariée à une base homologue du brin qui lui fait face via une liaison hydrogène. L'Adénine a pour base homologue la Thymine, la Cytosine à la Guanine pour homologue. Ainsi, il est possible de prédire la séquence de bases azotées d'un brin à partir du brin qui lui fait face. Par exemple, si on donne la séquence suivante : 5'-AGGTCCG-3', le brin en face aura la séquence 3'-TCCAGGC-5'. Les paires de bases dans la double hélice établissent des liaisons hydrogènes : L'adénine et la Thymine forment ensemble deux liaisons hydrogènes tandis que la Guanine en forme trois avec la Cytosine.

Le squelette de pentose-phosphate d'un brin est en sens 5' → 3' tandis que le deuxième brin a un squelette pentose-phosphate sens 3' → 5'. Cet arrangement est dit antiparallèle. La règle des bases homologues est la **règle de Chargaff** qui stipule que dans l'ADN, l'ADN possède en nombre égaux les résidus adénine et thymine et en nombres égaux ceux de la guanine et de la cytosine. Les deux polynucléotides sont maintenus ensemble grâce aux liaisons hydrogènes entre les bases azotées ; la liaison hydrogène est une liaison faible mais l'ensemble de ces liaisons entre les deux brins créé une force collective semblable à celle des dents d'une fermeture éclair.

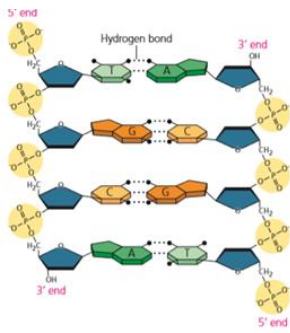


Les deux brins de la double hélice sont complémentaires l'un de l'autre. C'est cette particularité de l'ADN qui rend possible de générer deux copies identiques de chaque molécule d'ADN dans une cellule qui se prépare à la division cellulaire. Quand la cellule se divise, les copies sont distribuées aux cellules filles, les rendant génétiquement identiques aux cellules parentes (sauf lors de mutations non corrigées). De cette façon, la structure de l'ADN justifie sa fonction de transmettre l'information génétique quand une cellule se sépare.

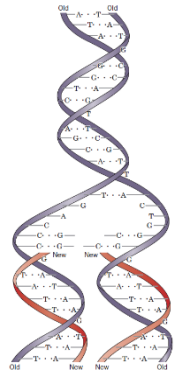
Quand un polynucléotide n'est composé que d'un seul brin, on parle d'un polynucléotide monocaténaire. Quand il y a deux brins, il s'agit d'un polynucléotide bicaténaire. L'ADN monocaténaire est rare mais il se retrouve dans certains groupes de cellules ou organismes ou encore dans les virus. Il existe plusieurs variations possibles de la structure spatiale de l'ADN ; la plus courante est l'ADN-B, structure en double hélice tournant vers la droite en effectuant un tour complet tous les 10 nucléotides. L'ADN-B est considéré comme la forme native parce que son mode de diffraction des rayons X ressemble à celui de l'ADN des têtes intactes des spermatozoïdes.

L'ADN, PORTEUR DE L'INFORMATION GENETIQUE

Jusque fin des années 1940, il était pensé que l'information génétique était portée par des complexes protéiques ou que cette information devait être portée par une molécule très complexe dans sa structure et dans ses propriétés. L'acide nucléique, isolé pour la première fois en 1869 par Friedrich Miescher, était vu comme une molécule polymérique dont la séquence de nucléotide était trop répétitive (seulement 4 différents nucléotides contre les protéines qui ont 20 résidus d'acides aminés différents conférant d'énormes combinaisons différentes). En 1944, Oswald Avery, Colin MacLeod et Maclyn McCarty ont démontré que



L'ADN est la molécule porteuse de l'information génétique. Leurs expériences ont montré que l'ADN – qui n'est pas une protéine-extrait d'un brin pathogénique d'une bactérie *Diplococcus pneumoniae* – était la substance qui a transformé un brin non-pathogénique de l'organisme au brin virulent. La découverte d'Avery a été sujette à un fort scepticisme de la part de la communauté scientifique. Cette expérience a néanmoins influencé Erwin Chargaff, dont les règles ont mené à des modèles conséquents de la structure et de la fonction de l'ADN.



La structure en double hélice de l'ADN facilite sa réplication : Quand une cellule se divise, chaque brin d'ADN se sépare et agit comme un modèle à partir duquel un brin complémentaire va se former. Par conséquent, chaque cellule fille contient une molécule d'ADN complète (ou une collection complète de molécules d'ADN dans les organismes dont le génome contient plus d'un chromosome). Chaque molécule d'ADN consiste en un brin « parent », issu de la cellule mère et un « brin fille », formé lors de la réplication : On dit que la réplication est semi-conservative. Les nouveaux brins formés sont synthétisés par la polymérisation de nucléotides, bouts à bouts, et spécifiquement au brin « parent », de manière à ce que le nouveau brin formé soit composé de nucléotides homologues au brin qui lui fait face ; ainsi, pour un brin 5' – *ATTGCA* – 3', le brin répliqué sera 3' – *TAAGCT* – 5'.

ACIDE RIBONUCLEIQUE (ARN)

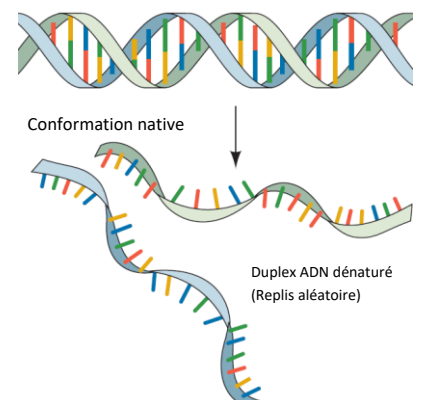
L'ARN est un polynucléotide composé d'un seul brin qui présente généralement une structure compacte plutôt que de se laisser en une chaîne étendue. Il existe néanmoins certains virus qui ont pour matériel génétique un ARN bicaténaire. Dans l'ARN, il n'y a pas de thymine mais celle-ci est la base équivalente de l'Uracile, base unique à l'ARN et homologue à l'adénine. Un brin d'ARN comporte, à l'instar de l'ADN, un squelette pento-phosphate mais le pentose est un ribose. Bien que l'ARN se retrouve généralement sous forme d'un seul brin, les bases azotées forment des liaisons hydrogènes entre elles sur le même brin ARN, donnant des structures « stem-loop » (boucle de brin) ainsi que des structures encore plus complexes quand plusieurs de ces boucles interagissent entre elles.



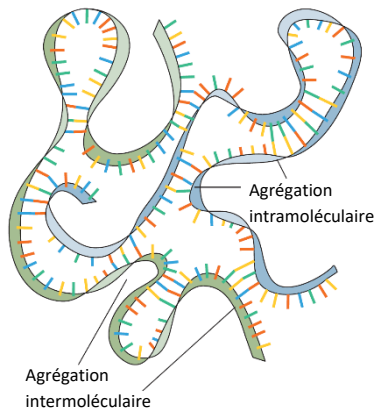
STABILISATION DES STRUCTURES DES ACIDES NUCLEIQUES

La forme native de l'ADN, forme de double hélice selon le modèle de Watson et Crick (ADN-B), apparaît lorsque l'ADN est dans un environnement favorable à sa conformation. Néanmoins, l'ADN ne présente pas une grande complexité structurale comme les protéines : les nucléotides étant répétitifs, il n'y a pas autant d'interactions possible qu'une chaîne polypeptidique ; l'ADN ne peut donc pas adopter énormément de structures secondaires.

La conformation de l'ADN est stabilisée par les paires de bases azotées homologues des deux brins de la double hélice ; les paires forment des liaisons hydrogènes. Cette paire de base agit comme une « colle » qui maintient les deux brins polynucléotidiques ensemble. Bien que l'ADN peut prendre d'autres conformations que celle proposée par Watson et Crick (modèle de l'ADN-B), il existe d'autres conformations que l'ADN peut prendre, de structures plus ou moins semblable. Néanmoins, la structure de l'ADN-B est la plus stable. Quand une solution d'ADN duplex (par duplex, on entend la structure en double hélice) est chauffée au-dessus d'une certaine température, les deux chaînes polynucléotidiques se séparent et prennent une conformation aléatoire ondulante : il s'agit de la **dénaturation**, il est accompagné d'une modification qualitative des propriétés physiques de l'ADN. La viscosité, résultant de la résistance à la déformation du duplex d'ADN, chute rapidement dès que l'ADN prend la forme de simples chaînes polynucléotidiques libres.



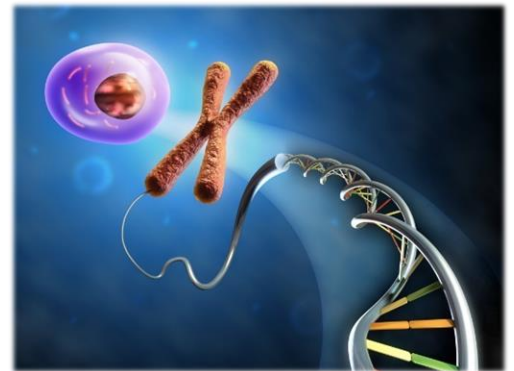
La stabilité de la conformation ADN-B (ADN duplex) dépend de plusieurs facteurs, dont la nature du solvant, la composition ionique de la solution et le pH. Cette stabilité est indiquée par la température de fusion T_f ; au-delà de cette température, la conformation du duplex ADN est dénaturée. Cette température de fusion T_f s'accroît linéairement en fonction de la composition molaire des bases G et C. Ces bases azotées établissent trois liaisons hydrogènes pour deux seulement entre la thymine et l'adénine, conférant une plus grande stabilité à la molécule.



Si la solution d'ADN dénaturé est refroidie rapidement en dessous de sa température de fusion T_f , l'ADN résultant ne sera que très partiellement apparié car les brins complémentaires n'ont pas assez de temps pour se réassocier complètement avant l'établissement d'appariements partiels aléatoires. Si on refroidit la solution lentement d'ADN dénaturé, et qu'on l'a maintenu à une température stable d'environ 25°C en dessous de T_f , l'énergie thermique du système sera suffisante pour que des liaisons d'hydrogènes se forment et que des appariements de bases s'établissent sur de courtes distances grâce à des dénaturations et renaturations successives mais insuffisante pour qu'il y ait dénaturation sur de longues séquences complémentaires. On parle de conditions d'annélation et l'ADN dénaturé pourra se renaturer complètement dans ces conditions. De la même manière, grâce aux conditions d'annélation, il est possible d'hybrider des brins complémentaires d'ADN et ARN pour donner une structure d'une double hélice hybride à peine moins stable que les ADN en double hélices.

ORGANISATION DU MATERIEL GENETIQUE CHEZ LES EUCARYOTES

L'homme a un génotype d'environ 3 Mb (3 milliards de nucléotides). Si on déroulait l'ADN d'une cellule, la molécule ferait 2m de long. Selon une étude récente, un corps humain de taille moyenne contient environ $3,72 \times 10^{13}$ cellules ; **Si on déroulait tout l'ADN de chacune de ces cellules et qu'on les mettait bout à bout, il atteindrait jusqu'à 400 fois la distance Terre-Soleil !** Le défi de la cellule est d'arriver à compacter cette molécule de 2m dans un noyau cellulaire d'environ 6µm à 20 µm de diamètre ; ce compactage se fait grâce à des interactions entre l'acide nucléique et des protéines nucléaires. On distingue deux types de protéines nucléaires : les protéines histoniques et les protéines non-histoniques.



La fonction la plus importante de l'ADN est de porter des gènes, l'information qui spécifie toutes les protéines et molécules d'ARN que l'organisme synthétise, dont l'information sur la quantité de chaque protéine et son endroit spécifique (tissu, compartiment cellulaire spécifique...). Le génome des eucaryotes est divisé en chromosomes : Au lieu d'avoir une seule macromolécule d'acide désoxyribonucléique, la cellule comprend plusieurs molécules appelées chromosomes, comportant des gènes différents. Le génotype d'un organisme définit l'ensemble de ses gènes et le caryotype définit l'ensemble des chromosomes d'un organisme. Le chromosome est formé d'un complexe d'ADN, ARN et de protéines appelé la chromatine. Les chromosomes n'apparaissent que lors de la mitose, quand la cellule se divise en deux cellules filles (de patrimoine génétique semblable). Lors de l'interphase, les chromosomes sont si diffus qu'ils sont indiscernables individuellement. Dans la cellule eucaryote, l'ADN est donc toujours sous forme de chromatine, associé à de l'ARN et des protéines. On peut distinguer deux types de chromatines : L'euchromatine, qui est la partie décondensée et permettant l'expression génétique et l'hétérochromatine, régions d'ADN condensé où la transcription d'un gène ne peut pas se faire.

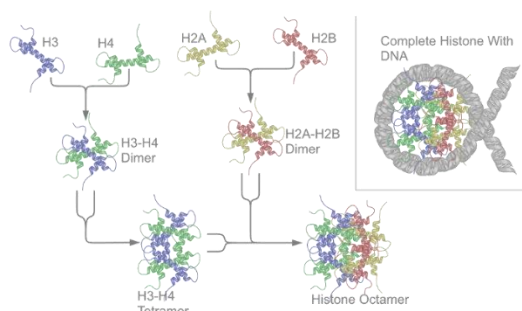
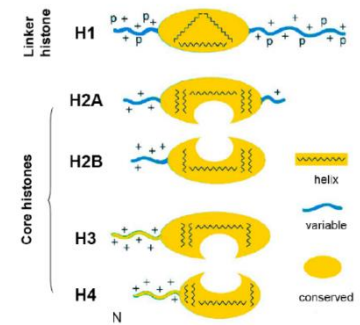
HISTONES, NUCLEOSOMES ET FIBRE DE CHROMATINE

Placer 46 chromosomes humains dans un noyau de 20 µm de diamètre est un défi auquel la cellule doit répondre par un mécanisme de compaction de l'ADN. Les chromosomes sont compactés selon plusieurs niveaux à différentes échelles et impliquent une association intime avec des protéines permettant le repliement du brin bicaténaire afin de réduire la longueur de l'ADN. Chaque chromosome est une longue molécule d'ADN bicaténaire qui ne porte aucune interruption sur toute sa longueur et qui est associée à des protéines de nature basique permettant une association étroite. On distingue deux types de protéines associées à l'ADN et responsables de sa compaction : les protéines histoniques et les protéines non-histoniques. Les protéines histoniques, ou histones, sont des protéines majoritairement constituées d'acides aminés basiques (Lysine et Arginine, notamment). Les histones sont intimement associées à l'ADN et permettent sa compaction en formant des **nucléosomes**. Le caractère basique des histones confère à la protéine une charge globale positive, permettant de fortes interactions avec l'ADN (de nature acide), au niveau des groupements phosphates des nucléotides qui portent des charges négatives. Ces protéines sont réparties en 5 classes principales : Les histones **H1, H2A, H2B, H3 et H4**. Il existe également des histones H5 qui appartiennent à la famille des H1.

Les histones sont une classe de protéines qui ont été hautement conservées lors de l'évolution : Au cours de l'évolution des eucaryotes, la structure primaire n'a de ces protéines n'a été que très peu modifiée ; Par exemple, si on compare les histones H4 entre une cellule issue d'un Bovin et d'un petit pois, espèces qui ont divergé il y a 1,2 milliards d'années, on remarque qu'ils ne diffèrent que par deux changements de résidus. L'histone H4 est la protéine la plus conservatrice connue. *Une telle stabilité au cours de l'évolution implique que ces quatre histones ont des fonctions essentielles pour lesquelles leurs structures sont si bien adaptées qu'elles ne tolèrent aucun changement.*

Si la structure primaire des histones a été hautement conservée d'une espèce à une autre, en revanche, les histones peuvent subir des modifications réversibles post-traductionnelles au niveau de résidus spécifiques qui varient fortement d'une espèce à une autre, du tissu ou de la phase du cycle cellulaire. Les histones comprennent un domaine C-terminal globulaire, un domaine histone-fold et un domaine N-terminale. Le domaine histone-fold est le domaine qui est hautement conservé et intolérant aux mutations. Ce domaine consiste en plusieurs hélices α , dont trois placées plus ou moins perpendiculairement chez les histones H2A/B, H3 et H4 de manière à s'attacher à d'autres histones avec un motif appelé « poignée de main » et permettre la dimérisation.

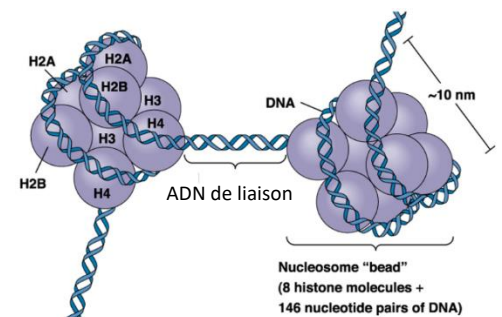
Principales protéines histones



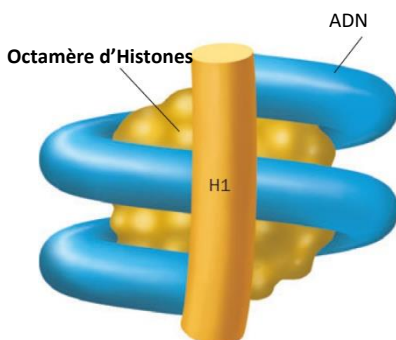
Les histones H3 et H4 s'associent et forment un dimère, de la même manière que le H2A et le H2B se dimérisent, toujours par un motif « poignée de main ». Ces dimères s'associent ensuite entre eux pour former un tétramère H3-H4 et H2A-H2B. Ces deux tétramères s'associent pour donner un octamère d'histones. Quand le duplex ADN vient s'enrouler autour d'un octamère d'histone, on parle d'un nucléosome ; le nucléosome est le premier niveau d'organisation de la chromatine. Les histones de classe H2A, H2B, H3 et H4 interviennent donc dans le premier niveau en formant une structure sphérique, on parle d'histones « de cœur » (core-histone) car elles constituent le cœur du

nucléosome. L'ADN s'enroule deux fois autour d'un nucléosome.

Au sens strict du terme, le nucléosome définit l'ensemble formé par une particule de cœur (octamère d'histone) et l'ADN de liaison adjacent, bien qu'on l'emploie généralement uniquement pour désigner la particule de ce cœur. Au sein de la chromatine, deux nucléosomes sont séparés par des segments d'ADN dit *de liaison*. Entre deux nucléosomes, il y a un espace dans lequel l'ADN se trouve « à nu », la longueur (et donc le nombre de paires de bases) varie d'une espèce à une autre. Le degré de condensation varie le long des chromosomes, dans la chromatine. Dans l'euchromatine, cette condensation est faible pour pouvoir laisser l'accès à l'ARN polymérase pour la transcription ADN \rightarrow ARN. Dans l'hétérochromatine, le degré de condensation est fort : ces zones ne peuvent pas être transcrites.



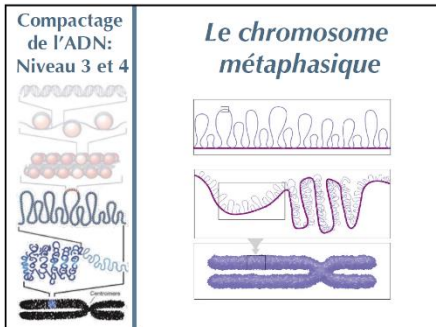
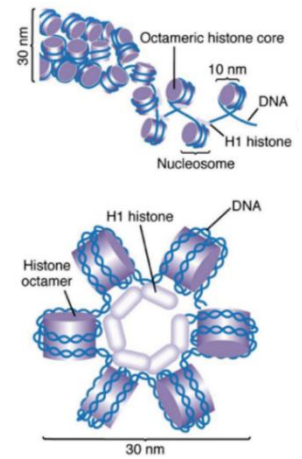
Le domaine N-terminale des histones n'est pas structuré : il apparaît de manière linéaire et constitue la « queue » des histones. C'est sur ce domaine que peuvent avoir lieu des modifications post-traductionnelles réversibles telles que la phosphorylation, l'acétylation, la méthylation, ubiquitination etc... Ces modifications font intervenir des enzymes spécifiques, on parle de protéines non-histoniques et elles régulent le degré de condensation ; ces modifications jouent un rôle dans la régulation épigénétique.



Les histones de la classe H1/H5 sont des histones de « liaison ». L'histone H1 scelle le nucléosome en liant l'ADN à l'endroit où il rentre et l'endroit où il sort de l'octamère : on parle d'un **chromatosome**. L'histone H1 ne se situe pas dans le nucléosome mais bien à l'extérieur, en stabilisant la fixation de l'ADN autour du nucléosome. En plus de sceller les nucléosomes, les histones H1 se composent d'un domaine globulaire hautement conservé flanqué de deux bras N- et C-terminaux riches en résidus basiques. Ces bras basiques interagissent avec l'ADN et liaison et réunissent les nucléosomes adjacents. L'enchaînement de nucléosomes forme un nucléofilament (ou fibre de chromatine) qui prend l'apparence d'un collier de perle. Les nucléosomes y apparaissent séparés par des segments d'ADN de liaison : cette apparence de « collier de perle » ne s'observe jamais dans des conditions physiologiques, ce n'est qu'un

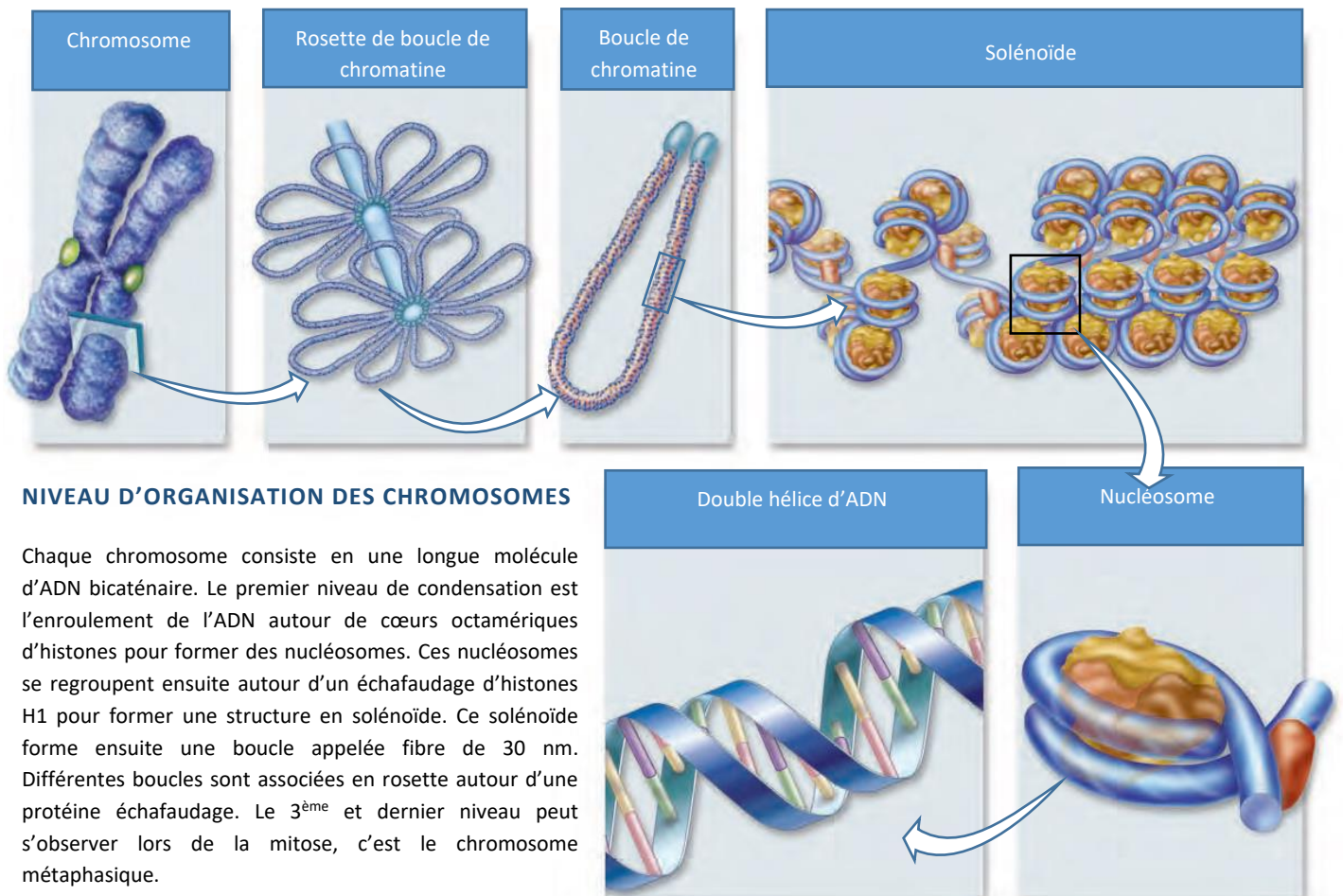
artéfact expérimental qui résulte de l'exposition de la chromatine à concentration saline augmentée. Cette forme « ouverte » de la chromatine est libre à la transcription. Les nucléosomes, en particulier les queues amino-terminales, jouent un rôle dans la régulation de l'expression génétique (régulation épigénétique).

Quand on atteint la concentration saline physiologique, on peut observer que les nucléosomes se replient entre eux grâce à des interactions entre les histones H1. Ces histones H1 vont enrôler les nucléosomes, avec 6 nucléosomes par tour, pour donner une fibre épaisse de 30 nm appelée la fibre chromosomique. Ce repliement des nucléosomes permet de compacter l'ADN en donnant une structure hélicoïdale qu'on appelle également solénoïde. Quand l'ADN est compacté en solénoïde (ou fibre de chromatine 30 nm), il n'est pas actif à la transcription. Le solénoïde est stabilisé par les histones H1 dont les bras N- et C-terminaux entrent en contact avec les nucléosomes adjacents. La fibre de chromatine 30 nm est le deuxième niveau de compactage de l'ADN.



Chromosome métaphasique : Le troisième niveau est caractérisé par des boucles radiales de fibre de chromatine 30nm sur un échafaudage composé de protéines non-histones : ces boucles radiales sont appelées *domaines en boucles* et n'ont lieu que lorsque la cellule est en **phase mitotique**. Ces domaines en boucles s'enroulent et se replient de sorte à compacter encore plus l'ADN et former une fibre de 300 nm. C'est ce repliement qui donne l'aspect des chromosomes en X lors de la **métaphase**. La condensation de la chromatine en métaphase est telle que l'on peut nettement distinguer chaque chromosome individuellement, permettant d'établir le caryotype de la cellule. Au cours de l'interphase, entre deux divisions cellulaires, les chromosomes sont étendus et ne prennent pas la structure de fibre 300 nm (domaines en boucles). On ne peut donc pas distinguer chaque chromosome individuellement dans le noyau. Le chromosome métaphasique est constitué de deux chromatides (Cht). Un chromatide est un nucléofilament en forme de bâtonnet (=fibre de chromatine) associé à des protéines d'histones et des protéines non-histones (PNH). Les nucléofilaments forment des boucles qui entrent et sortent de l'échafaudage à peu près au même endroit.

On ne peut donc pas distinguer chaque chromosome individuellement dans le noyau. Le chromosome métaphasique est constitué de deux chromatides (Cht). Un chromatide est un nucléofilament en forme de bâtonnet (=fibre de chromatine) associé à des protéines d'histones et des protéines non-histones (PNH). Les nucléofilaments forment des boucles qui entrent et sortent de l'échafaudage à peu près au même endroit.

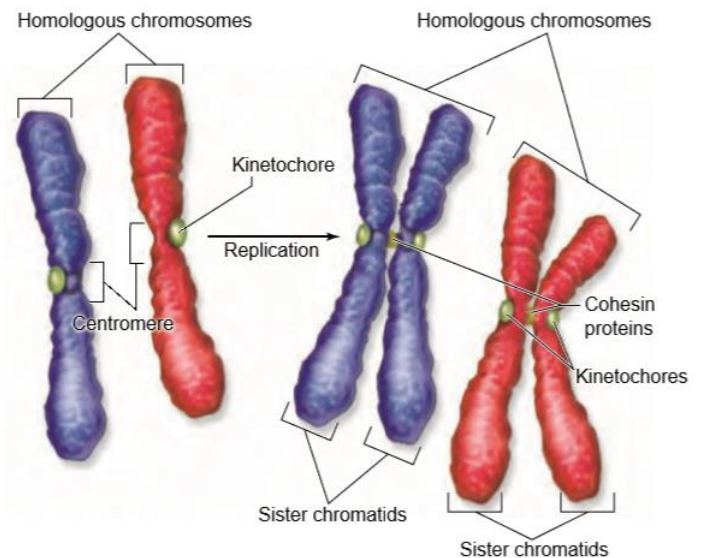


NIVEAU D'ORGANISATION DES CHROMOSOMES

Chaque chromosome consiste en une longue molécule d'ADN bicaténaire. Le premier niveau de condensation est l'enroulement de l'ADN autour de cœurs octamériques d'histones pour former des nucléosomes. Ces nucléosomes se regroupent ensuite autour d'un échafaudage d'histones H1 pour former une structure en solénoïde. Ce solénoïde forme ensuite une boucle appelée fibre de 30 nm. Différentes boucles sont associées en rosette autour d'une protéine échafaudage. Le 3^{ème} et dernier niveau peut s'observer lors de la mitose, c'est le chromosome métaphasique.

Le nombre de chromosomes varie très fortement d'une espèce à une autre. Chez l'homme, le caryotype est de 23 paires presque identiques de chromosomes : 22 paires de chromosomes autosomiques, et une paire de chromosomes sexuels. Lorsqu'une cellule contient un seul jeu de chromosomes, elle est dite haploïde : la cellule contient alors 23 chromosomes. Lorsqu'elle contient deux jeux de chromosomes, elle est dite diploïde : la cellule contient alors 23 paires de chromosomes, soit 46 chromosomes au total. Les chromosomes qui appartiennent à une même paire sont dits des chromosomes homologues. Les cellules somatiques chez les mammifères sont des cellules diploïdes, notées $2n$ tandis que les cellules germinales (servant à la reproduction) sont haploïdes, notées n . Lorsque deux cellules sexuelles fusionnent pour donner un zygote, les cellules haploïdes forment alors une cellule diploïde : chaque paire de chromosome est donc constituée d'un chromosome maternel et paternel.

Chaque chromosome d'une cellule peut être constitué d'une ou de deux chromatides selon son état : juste après la mitose, chaque chromosome de la cellule ne contient qu'un chromatide. Les chromosomes de la cellule mère étaient constitués de deux chromatides qui ont été séparés pour donner deux cellules filles de chromosome à un seul chromatide. Les deux chromatides contiennent exactement la même information génétique et sont reliés par un centromère. Deux chromatides d'un même chromosome sont des chromatides sœurs, ce sont deux copies identiques formées par la réplication. Lors de l'interphase, les chromosomes à un seul chromatide reforment le chromatide sœur pour le prochain cycle de division cellulaire. Les deux chromatides sœurs sont séparés lors de la mitose ainsi que durant la seconde division de la méiose (division des cellules sexuelles dites gamètes).



Lorsque la cellule rentre en division cellulaire, l'ADN atteint son niveau de compaction maximale en organisant les chromosomes métaphasiques. Cette condensation permet de séparer les chromosomes pour la mitose et est effectuée à l'aide de complexes protéiques appelés les condensines. Les condensines font partie d'une famille appelées les protéines de maintenance structurelle des chromosomes (protéines SMC) et qui interviennent dans l'organisation et la dynamique des chromosomes au cours du cycle de la cellule. Les SMCs sont toutes des ATPases : elles consomment donc de l'ATP pour pouvoir fonctionner. L'activité des complexes de condensines sont hautement régulées pendant le cycle cellulaire et certaines ne sont activées que lorsque la cellule rentre en division, tandis que d'autres sont actives en interphases et se chargent de la condensation des chromosomes aux premiers niveaux d'organisation. Le complexe de condensine I est présent dans le cytoplasme durant l'interphase et ne peut condenser les chromosomes seulement lorsque la cellule rentre en division et que le noyau cellulaire commence à se désintégrer.

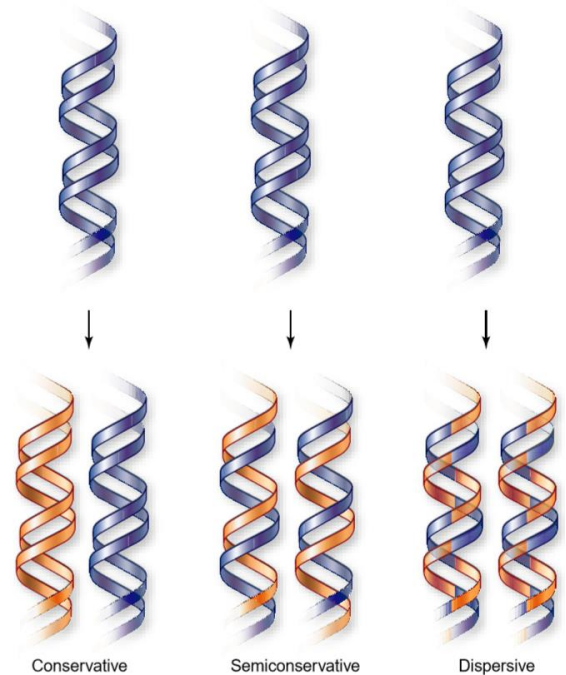
Avant de pouvoir rentrer en division, la cellule doit nécessairement répliquer l'information génétique en double pour pouvoir la redistribuer aux deux cellules filles. Les chromosomes répliquent donc leur chromatide pour former des chromosomes de deux chromatides. Lorsqu'un chromosome ne possède qu'un seul chromatide, il est dit $1c$. Quand l'ADN est répliqué, chaque chromosome possède deux chromatides sœurs, notés $2c$. Par conséquent, une cellule haploïde qui va rentrer en division a un génome $2n+4c$ (deux jeux de chromosomes de deux chromatides, donnant au total 4 chromatides). La cellule haploïde qui vient de se diviser quant à elle a un génome $2n+2c$.

Lors de la réplication de l'ADN, la cellule possède deux jeux de chromosomes homologues et deux chromatides. Chaque chromosome est alors en double : les chromatides sœurs ne sont que des copies identiques l'un de l'autre. Lors de la condensation des chromosomes en chromosome métaphasique, les chromosomes apparaissent comme deux chromatides sœurs attachés l'un à l'autre. Les chromatides sœurs d'un chromosome sont rattachés entre eux au niveau d'une zone centrale appelée le centromère. La cohésine est un complexe protéique SMC qui se charge du rattachement des chromatides sœurs ainsi qu'à leur séparation lors de la division cellulaire. Il est important que les chromatides restent unis pour la division cellulaire afin d'assurer une bonne redistribution de chaque chromatide sœur de manière égale aux deux cellules filles, de manière à ce qu'elles aient toutes les deux la même information génétique. La cellule s'assure qu'un exemplaire de chaque chromosome se dirige vers chaque cellule fille en unissant les chromatides entre eux au niveau des centromères. Nous reviendrons un peu plus en détail sur le chromosome métaphasique dans le chapitre sur la division cellulaire.

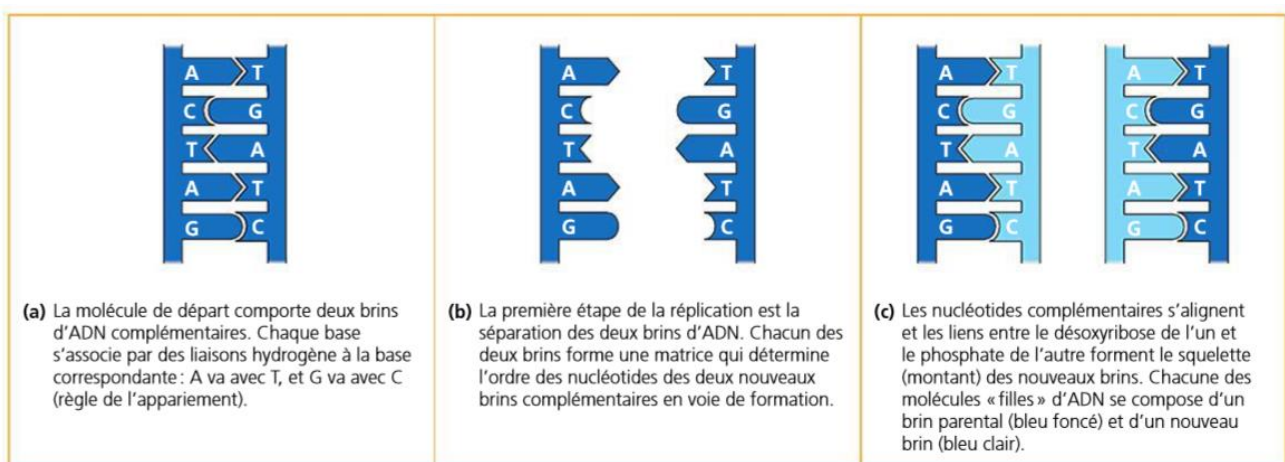
Lorsqu'on a élucidé la structure tridimensionnelle de l'ADN (par Crick et Watson et leur modèle hélicoïdal à double brin), on a aussitôt pensé que la complémentarité des bases azotées était à la base même de la réplication de l'information génétique. En effet, avec un brin d'ADN on peut facilement deviner la composition de son brin homologue puisque chaque nucléotide a son nucléotide complémentaire sur le brin d'en face : un nucléotide AMP (nucléotide de base azotée adénine) a pour nucléotide complémentaire le nucléotide TMP, et la GMP a pour nucléotide complémentaire la CMP. La séquence de bases azotées d'un brin détermine totalement la séquence de son brin partenaire.

Lorsque la cellule se prépare à la division, elle passe automatiquement par une phase obligatoire de réplication de l'information génétique : chaque chromosome est répliqué à l'identique. Pour ce faire, il faut donc que les chromosomes soient lisibles. Pour être lisible, il faut que la partie du chromosome en train d'être répliquée soit lâche (non condensée). Lors de la réplication, les deux brins de l'hélice d'ADN d'un chromosome, les brins parentaux, vont être répliqués pour donner au total 4 brins, soit deux hélices d'ADN (2 chromatides). Les deux chromatides alors formés à partir d'un seul vont se séparer et rester unis au niveau de leur centromère. Néanmoins, on ne savait pas exactement comment la réplication de l'ADN s'effectuait ainsi que la répartition des brins. On a donc proposé trois modèles de réplication :

1. **Un modèle conservatif** : Les deux brins « parentaux » de l'hélice d'ADN restent intacts après la réplication et s'apparient de nouveau pour former la même hélice et donnent deux brins fils identiques qui forment une nouvelle molécule d'ADN. On a donc deux chromatides sœurs identiques mais l'un vient d'être formé à l'identique et est une molécule d'ADN entièrement nouvelle : on a donc conservation de la molécule d'ADN de départ.
2. **Un modèle semi-conservatif** : L'hélice se sépare et un nouveau brin vient se former à partir d'un brin parent. L'hélice de départ se sépare alors complètement, donnant deux hélices d'ADN comportant chacune un brin parent et un brin fils : on a une semi-conservation de la molécule de départ.
3. **Le modèle dispersif** : les deux nouvelles molécules d'ADN formées sont un mélange de brins parentaux et de brins fils. Après réplication, l'ADN serait dispersé dans les deux brins de chaque molécule.



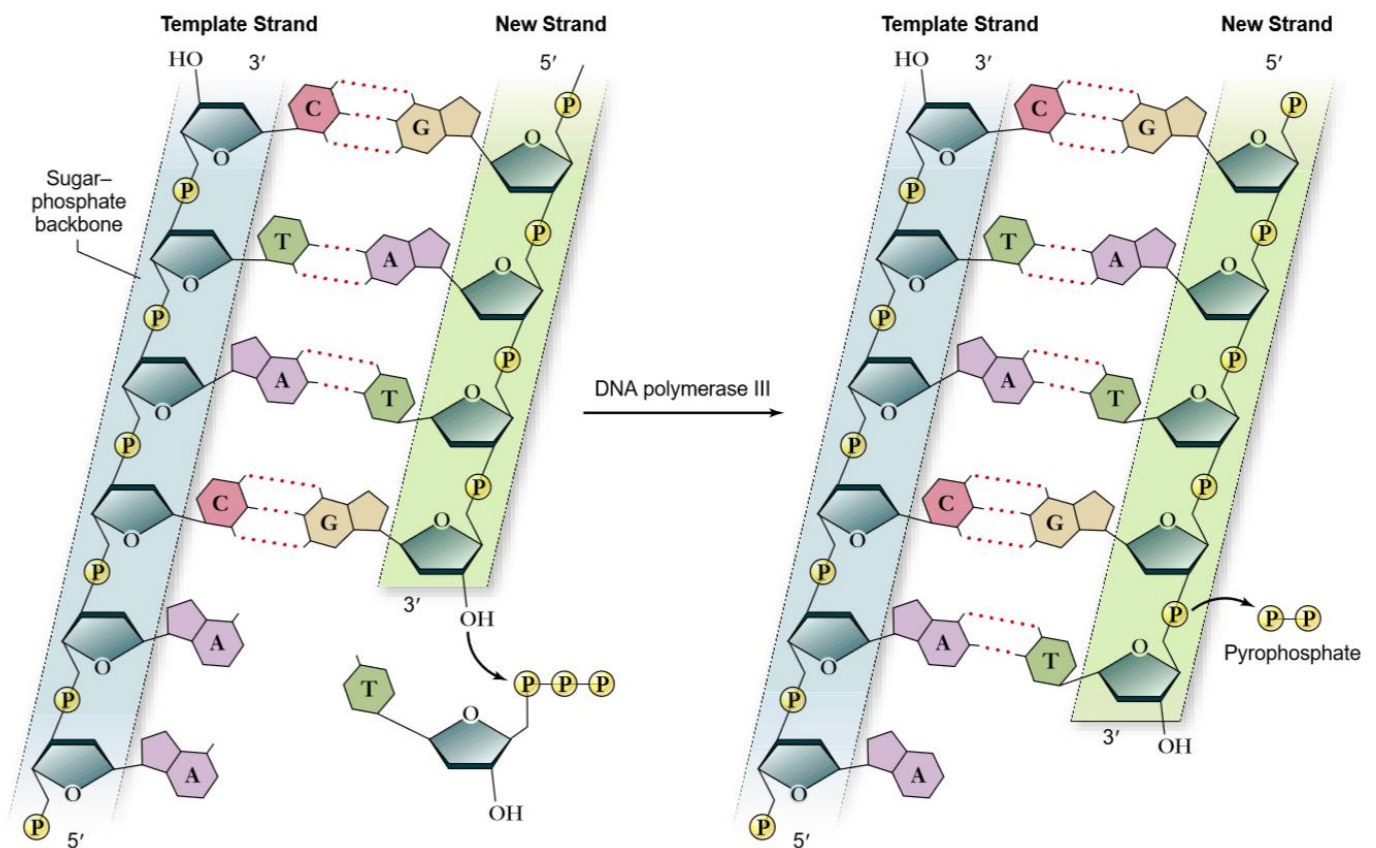
En réalité, la réplication de l'ADN suit le modèle semi-conservatif. Ce modèle de réplication a été prouvé par l'expérience de Meselson-Stahl en traçant l'ADN d'une bactérie avec de l'azote radioactif ^{15}N . Puisque les nouveaux nucléotides sont des azotes ^{14}N (azote non-radioactif), il a regardé la répartition des azotes radioactifs dans les deux nouvelles cellules pour voir comment l'ADN s'était redistribué. L'ADN se réplique par ouverture de son hélice : chaque brin parental sert alors de matrice pour la synthèse d'un brin complémentaire et aboutit à deux hélices filles possédant chacune un brin parental et un brin nouveau.



Chez le procaryote, la réplication de l'ADN démarre au niveau d'une séquence spécifique appelé le site d'origine. Le matériel génétique du procaryote consiste en une seule molécule d'ADN circulaire. La synthèse début au site d'origine et termine au site de terminaison. Chez les eucaryotes, on peut supposer que l'ADN comporte également une origine quand ce processus démarre. Néanmoins, en regardant la vitesse de synthèse de l'ADN, la réplication de l'ADN d'une cellule eucaryote humaine pourrait prendre plusieurs jours pour être entièrement répliqué. Ce n'est pas possible puisque certaines cellules embryonnaires peuvent se répliquer avec moins d'une heure d'intervalle.

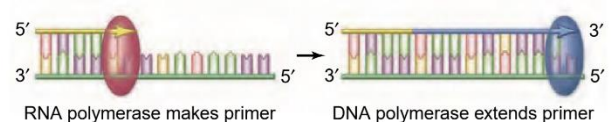
A l'instar de la réplication des procaryote, la réplication se déroule par 3 processus : initiation, élongation et terminaison. Plusieurs enzymes sont impliquées dans la réplication de l'ADN : elles collaborent ensemble à assembler un nouveau brin à partir d'un brin matrice (le brin parental) et de nucléotides complémentaires. Un chromosome eucaryote comporte plusieurs origines de réplication (séquence nucléotidique qui initie la réplication) afin de pouvoir répliquer l'ADN plus rapidement.

Dans un premier temps, nous allons voir la réplication de l'ADN chez les procaryotes en raison de sa simplicité par rapport à la réplication des eucaryotes. L'ADN des procaryotes est circulaire et possède une seule origine de réplication appelée OriC. L'origine de réplication est reconnue par des protéines de réplication chargées de reconnaître la séquence d'ADN et amorcer la réplication de l'ADN. L'enzyme responsable de la synthèse du nouveau brin en faisant correspondre les nucléotides du brin matrice avec des nucléotides complémentaires est l'ADN polymérase. Il existe plusieurs ADN polymérases différentes et toutes ont la particularité de synthétiser dans le sens 5'→3' car elles ne savent lire que les brins dans le sens 3'→5'.



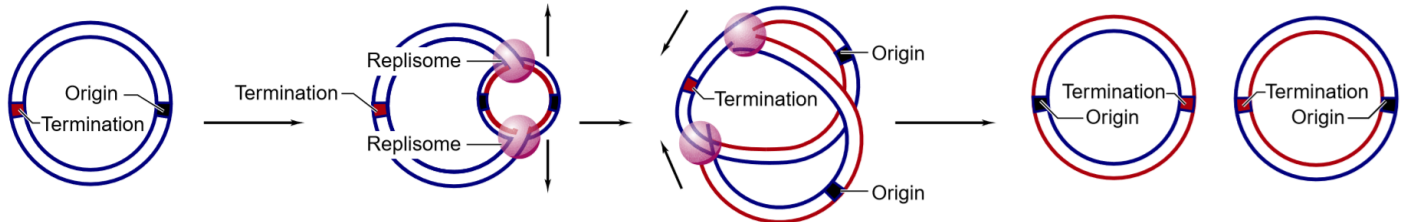
De plus, les ADN polymérases ont besoin d'une amorce pour entamer la réplication de l'ADN ; il faut que le brin matrice ait un petit bout de brin complémentaire à partir duquel l'ADN polymérase va continuer la réplication. L'ADN polymérase est incapable de répliquer un simple brin d'ADN s'il ne possède pas d'amorce. Cette amorce peut être un segment d'ADN ou d'ARN (un ARN évidemment complémentaire au brin, où les TMP sont remplacés par les UMPs).

Les ARN polymérases n'ont pas cette exigence et sont capable de synthétiser de l'ARN à partir d'un brin matrice et sans amorce. Les ARN polymérases se chargent alors de synthétiser une amorce d'ARN pour que l'ADN polymérase puis faire son travail.



REPLICATION DES PROCARYOTES

Dans cette section, nous allons prendre le cas de la réplication de l'ADN chez *E. Coli*. Cette bactérie possède un ADN circulaire et bicaténaire contenant l'origine de réplication OriC ainsi qu'un site de terminaison. La réplication débute au niveau d'OriC et se poursuit dans les deux sens jusqu'à arriver au site de terminaison. L'OriC est une séquence d'ADN répétée riche en bases azotées A et T. Si cette séquence est si riche en adénine et en thymine c'est pour une raison bien particulière : les adénines et thymines sont des bases complémentaires qui ne forment que 2 ponts d'hydrogène avec leur base complémentaire. En diminuant le nombre de ponts d'hydrogène, on facilite l'ouverture de l'hélice au niveau d'OriC.

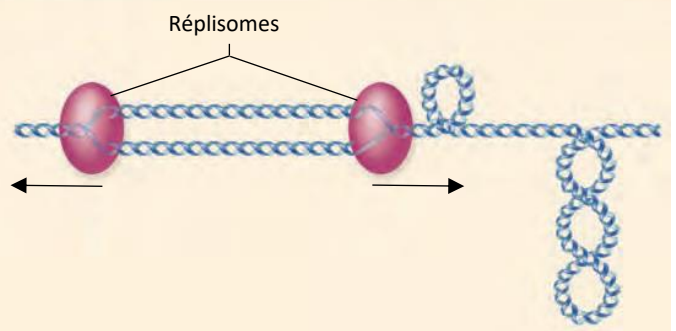


Le procaryote ne possède qu'une seule origine de réplication dans son unique chromosome (OriC). Cette séquence est reconnue par une protéine initiatrice qui va recruter les enzymes nécessaires à l'ouverture et la réplication du duplex hélicoïdal. Chez l'Eucaryote, l'ADN est linéaire et comporte plusieurs milliers de sites d'origines (environ 100 000 chez l'Homme). L'ADN répliqué à partir d'un site d'origine forme une unité appelé le réplicon : la bactérie ne comportant qu'un ADN et un seul site d'origine, il contient un seul réplicon. La séparation des deux brins parentaux au niveau de l'OriC forme un « œil de réplication », l'œil s'agrandit au fur et à mesure que l'hélice s'ouvre et que la réplication se poursuit. Lorsqu'un chromosome comporte plusieurs origines de réplifications (chez les eucaryotes exclusivement), les réplicons finissent par fusionner pour former au final deux molécules d'ADN complètes sans interruptions.

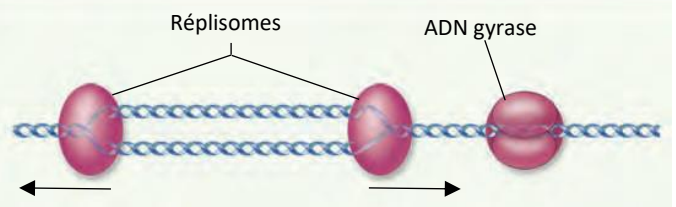
Une fois que les protéines de réplication ont trouvé l'origine de réplication OriC, elles recrutent les enzymes responsables de la réplication de l'ADN. La réplication débute une fois que le brin est ouvert : l'ouverture se fait grâce à une hélicase, enzyme qui rompt les liaisons d'hydrogène entre les bases complémentaires. L'hélicase nécessite de l'ATP pour séparer les deux brins parentaux. Certaines ADN polymérases sont capables de dérouler l'ADN pendant la réplication de l'ADN mais le déroulement par une hélicase est bien plus efficace et intervient quasi toujours. Le déroulement est un processus endergonique et les brins « nu » d'ADN produits par l'hélicase sont instables en raison de l'exposition des bases hydrophobes à l'eau. Des protéines fixatrices d'ADN monocaténaire (le brin est monocaténaire une fois déparié), appelées protéines SSB (pour single-strand binding proteins), vont s'attacher au brin d'ADN non appariés et l'empêcher de s'enrouler sur lui-même. Le déroulement de l'hélice provoque des tensions de torsions importantes de la même manière que la tire deux rubans de caoutchouc torsadés l'un autre de l'autre : on parle de surenroulement. En mathématiques, les torsions et les surenroulements sont étudiés dans une branche appelée la topologie.

L'enroulement de la double hélice suite à sa séparation des brins est un état topologique de l'ADN. Le problème avec le surenroulement de l'ADN est que la partie surenroulée de l'ADN ne peut être ouverte par une hélicase : il faut donc une enzyme capable de dérouler les torsions de l'ADN en diminuant les tensions de torsions afin de pouvoir déparier les deux brins. On a donc recours à des enzymes capables de modifier l'état topologique de l'ADN : des topoisomérases. La topoisomérase chargée de diminuer les tensions de torsions dans la réplication de l'ADN est l'ADN gyrase. L'ADN gyrase diminue les tensions de torsion en coupant l'ADN et en faisant pivoter les brins d'ADN puis en réparant ces coupures. L'hélicase peut alors procéder et séparer les brins parentaux qui pourront servir de matrice aux ADN

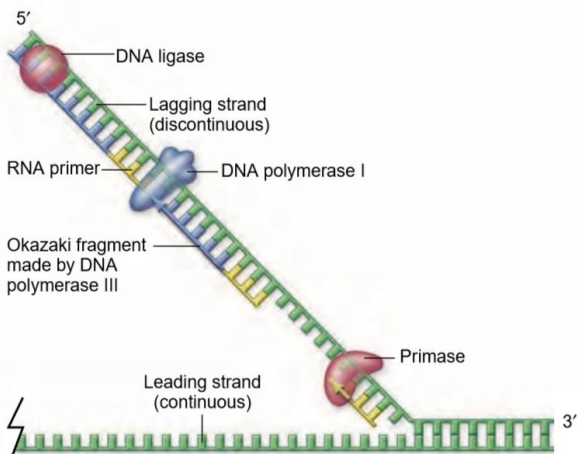
SURENROULEMENT DU AUX TENSIONS DE TORSIONS



PAS DE SURENROULEMENT AVEC UN ADN GYRASE



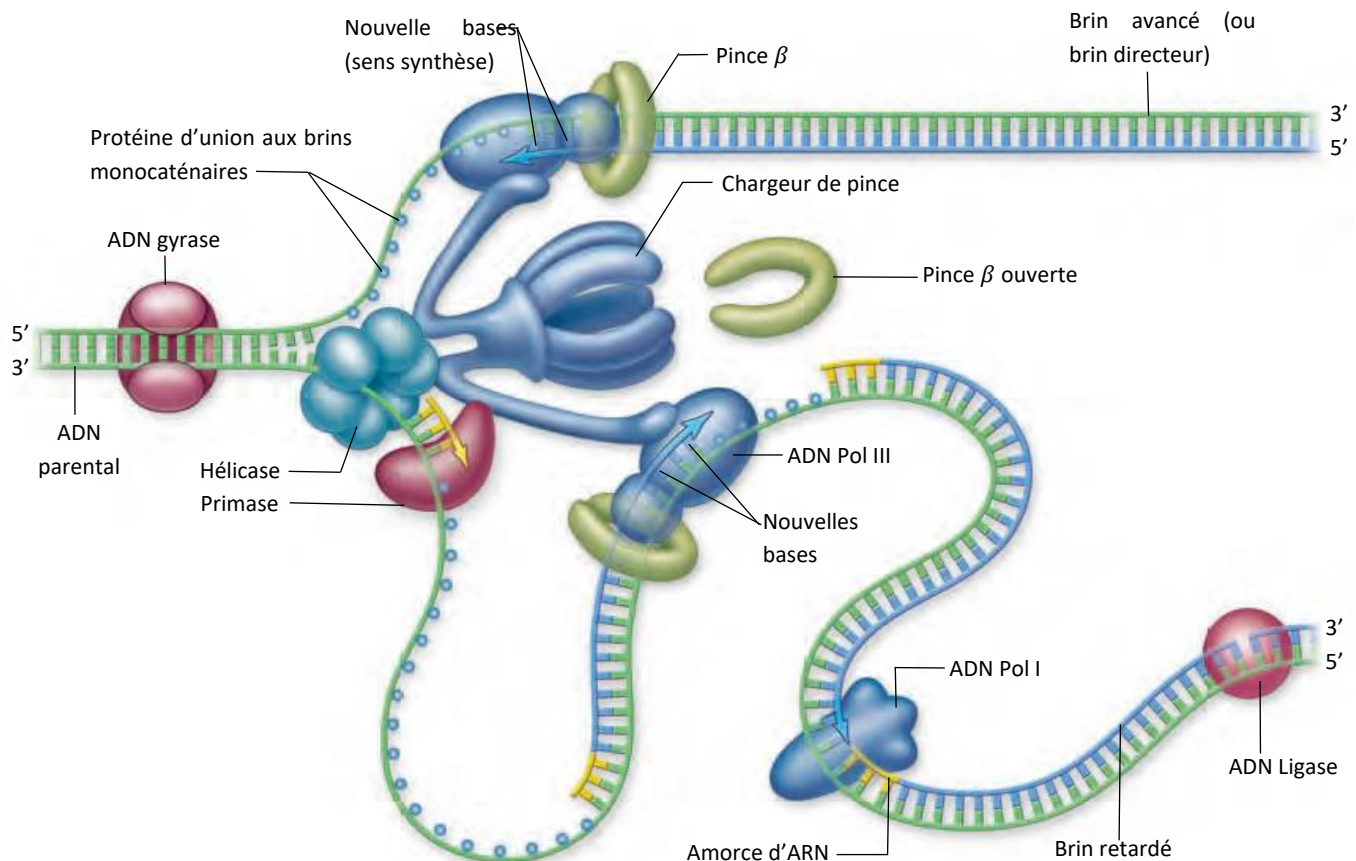
La synthèse du brin avancé est assez simple puisque l'ADN polymérase peut directement commencer à synthétiser le nouveau brin à partir d'une amorce déposée par une primase au niveau de la fourche de réplication. Ainsi, un réplicon nécessite une seule amorce d'ARN pour la synthèse du brin avancé (brin 5'-3'). Dans le cas de E. Coli qui ne possède qu'un réplicon, l'ADN Pol III (polymérase principale) peut synthétiser tout un brin en continu et donc répliquer tout le chromosome circulaire.

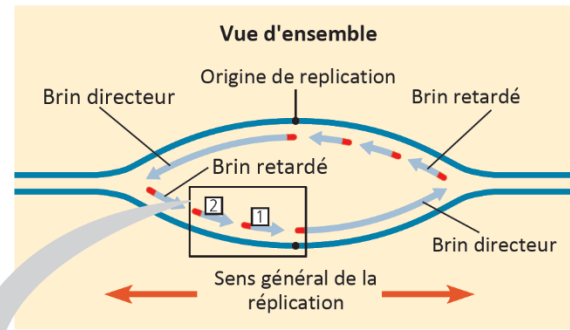
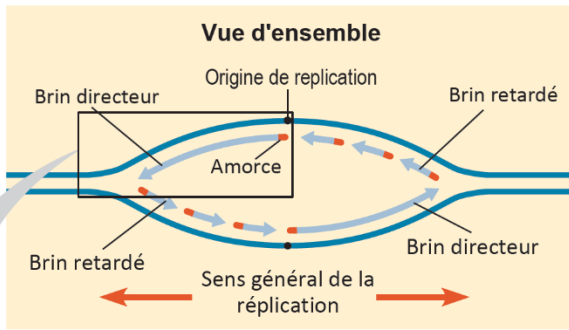


La synthèse du brin retardé se fait de manière discontinue. Contrairement au brin avancé, la synthèse n'est pas continue : l'enzyme ADN polymérase ne pouvant lire que dans le sens 5'-3', elle lire le brin 3'-5' dans le sens inverse. Pour ce faire, il va lire le brin matrice en s'éloignant de la fourche de réplication (lecture en 5'-3'). L'ADN primase va synthétiser plusieurs amorces ARN qui serviront à la synthèse des fragments d'Okazaki par une ADN polymérase. Ces amorces devront être ensuite remplacées par leur séquence d'ADN correspondante. Enfin, les fragments doivent être soudés entre eux. La longueur des fragments d'Okazaki varie fortement entre Eucaryotes et Procaryotes : Chez E. Coli, les fragments sont d'une longueur de 1.000 à 2.000 nucléotides tandis qu'elles ne sont que de 100 à 200 nucléotides chez la plupart des Eucaryotes.

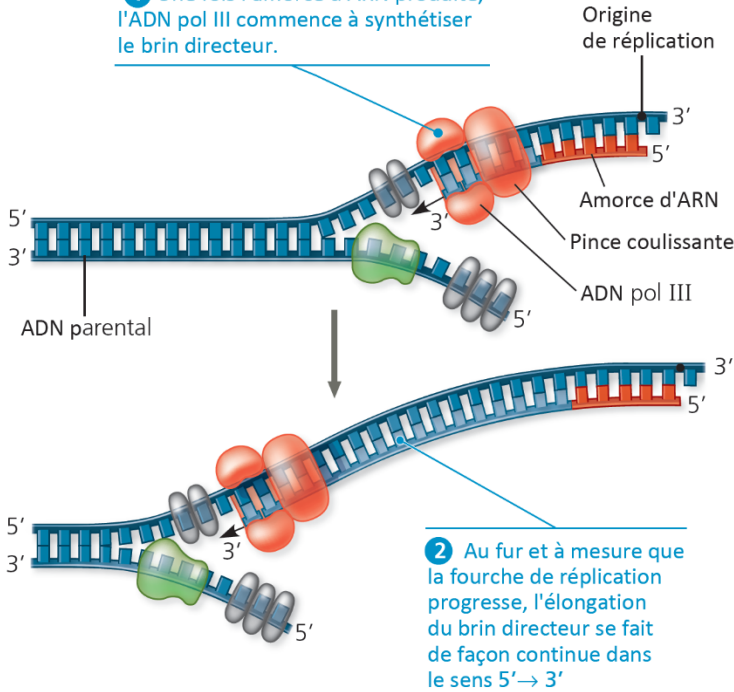
L'ADN pol III va synthétiser les fragments d'Okazaki à partir des amorces tandis que l'ADN Pol I va se charger de remplacer les amorces par leur séquence ADN correspondante, toujours dans le sens de synthèse 5'-3'. L'ADN Pol I ne peut commencer à remplacer une amorce qu'une fois que le fragment d'Okazaki précédent a été synthétisé par l'ADN Pol III car l'ADN Pol I a également besoin de son amorce pour fonctionner, et c'est l'extrémité du fragment d'Okazaki qui sert d'amorce. Les différents fragments sont ensuite reliés par une ADN ligase qui effectue des liaisons phosphodiester entre les différents fragments. La synthèse du brin retardé est détaillée à la page suivante.

La réplication du réplicon de E. Coli se termine au site de terminaison, situé à l'opposé d'oriC. Les enzymes qui interviennent dans la réplication de l'ADN (les ADN polymérases, la primase, l'ADN ligase et l'hélicase) forment un énorme complexe appelé le réplisome.

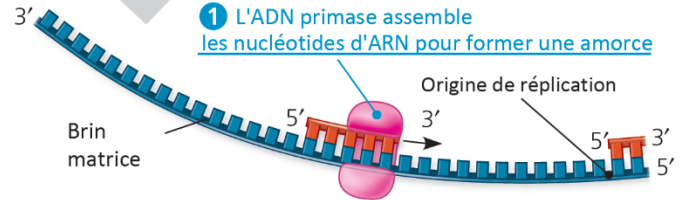




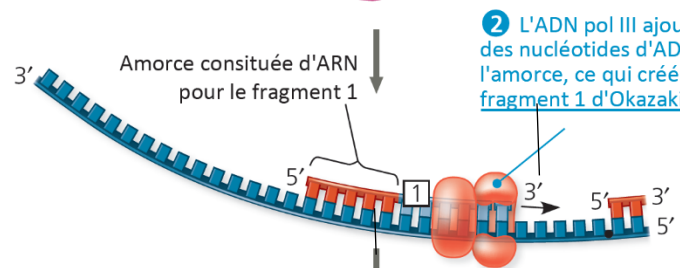
1 Une fois l'amorce d'ARN produite, l'ADN pol III commence à synthétiser le brin directeur.



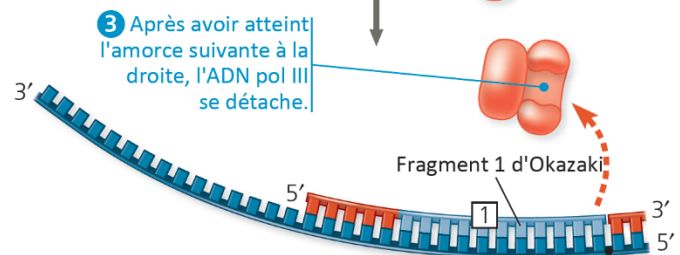
1 L'ADN primase assemble les nucléotides d'ARN pour former une amorce



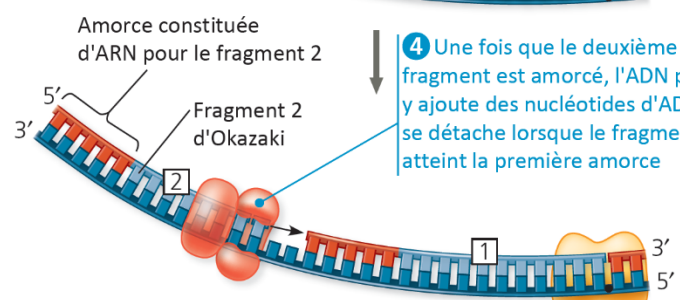
2 L'ADN pol III ajoute des nucléotides d'ADN à l'amorce, ce qui crée le fragment 1 d'Okazaki.



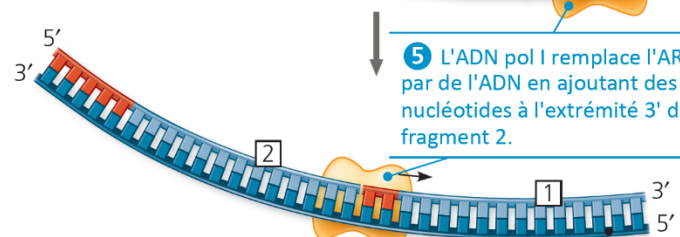
3 Après avoir atteint l'amorce suivante à la droite, l'ADN pol III se détache.



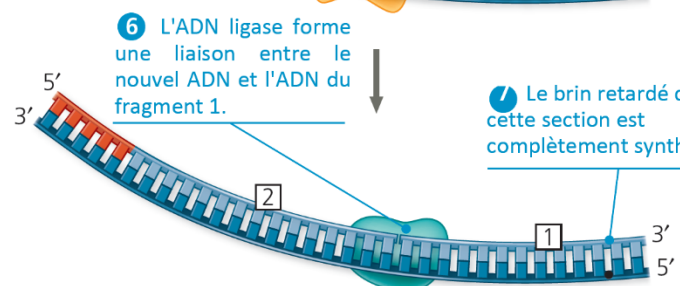
4 Une fois que le deuxième fragment est amorcé, l'ADN pol III y ajoute des nucléotides d'ADN et se détache lorsque le fragment atteint la première amorce



5 L'ADN pol I remplace l'ARN par de l'ADN en ajoutant des nucléotides à l'extrémité 3' du fragment 2.



6 L'ADN ligase forme une liaison entre le nouvel ADN et l'ADN du fragment 1.

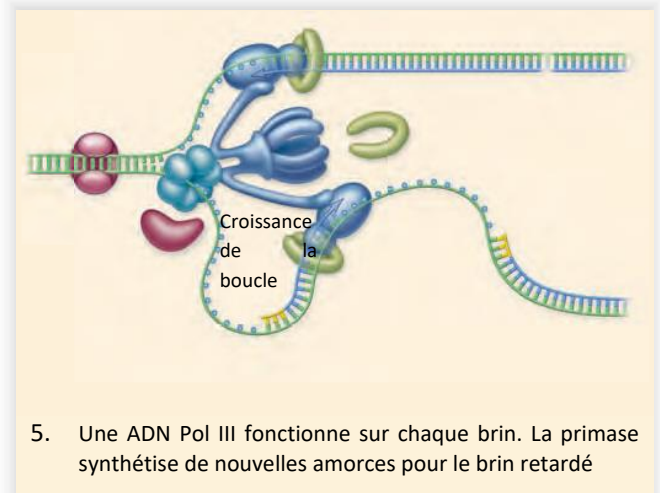
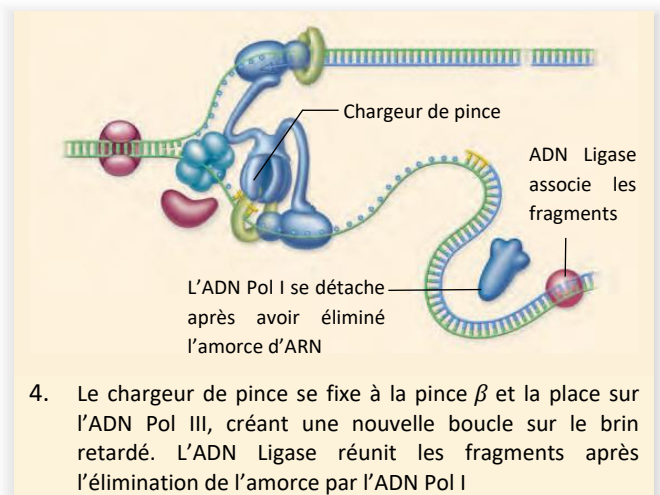
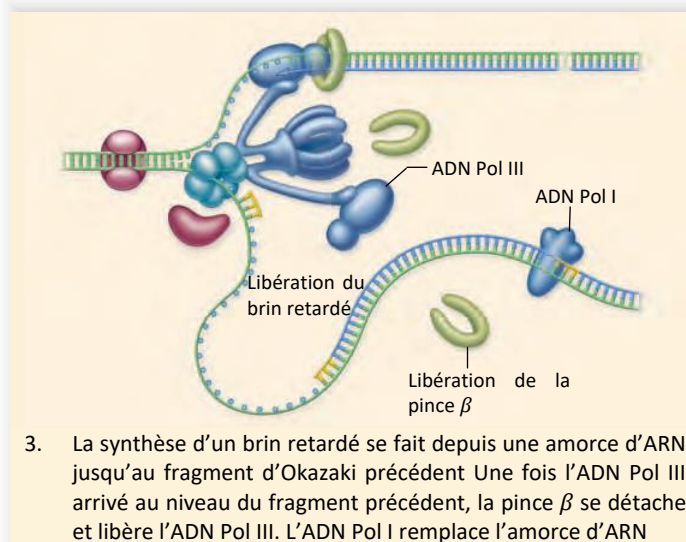
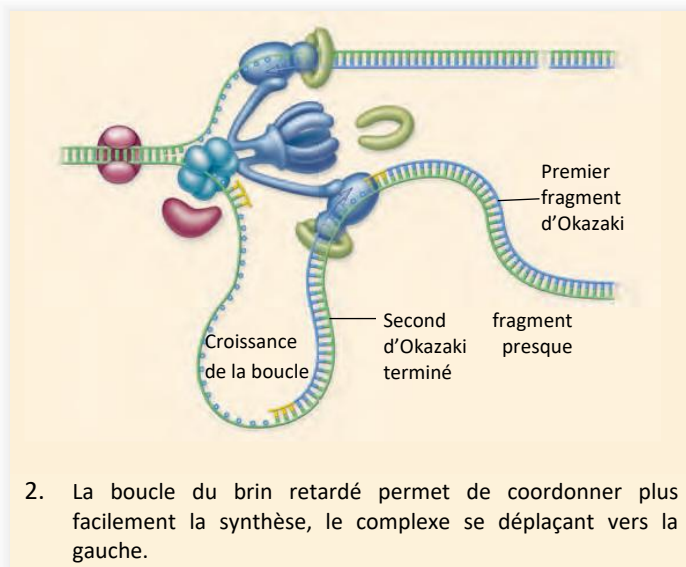
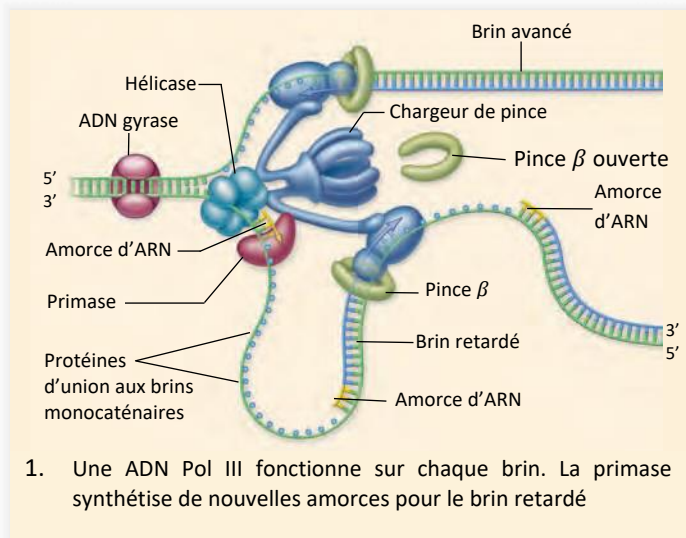


Le brin retardé de cette section est complètement synthétisé

Sens général de répllication

Protéines de répllication procaryote et leurs fonctions

PROTEINE	FONCTION
Hélicase	Sépare les deux brins parentaux de l'hélice au niveau de la fourche de répllication.
Protéine de liaison au brin monocaténaire (SSB)	Se lie au brin parental nu et le stabilise jusqu'à ce qu'il soit utilisé comme matrice par l'ADN pol
Topoisomérase	Diminue les tensions de torsion en aval de la fourche de répllication.
Primase	Synthétise une amorce d'ARN à l'extrémité 5' du brin directeur et à l'extrémité 5' de chaque fragment d'Okazaki du brin retardé
ADN Pol III	Polymérase responsable de la synthèse du brin directeur et des fragments d'Okazaki à partir d'un brin parental matrice.
ADN pol I	Enlève les amorces d'ARN et les remplace par leur équivalent ADN
ADN ligase	Lie les fragments d'Okazaki du brin retardé.



Les ADN polymérase III sont fixées au brin matrice grâce à des pinces β qui viennent s'enrouler autour de l'ADN polymérase. La pince est initialement ouverte et se referme autour du complexe ADN pol III – brin matrice pour former un anneau, également appelé la pince coulissante. Un chargeur de pinces est lié aux deux ADN Pol III et se charge de fixer les anneaux au niveau des amorces. Ce chargeur de pince sert également de protéine d'infrastructure pour lier toutes les autres enzymes en un complexe unique, le réplisome.

Les ADN Pol III de chaque brin sont coordonnés grâce au complexe enzymatique du réplisome qui dimérise les deux ADN Pol III grâce à des protéines auxiliaires de la réplication. Pour permettre une réplication coordonnée, le brin retardé se replie et forme une boucle pour chaque fragment à synthétiser qui permet à l'ADN Pol III de glisser.

Les polymérase sont des enzymes douées d'une capacité appelée la processivité. La processivité se définit comme la capacité à catalyser des réactions successives sur une même molécule sans la relâcher : les ADN pol III sont capable d'ajouter des dNTP en continu sans lâcher le brin matrice et réitérent l'allongement du brin en synthèse sans lâcher le brin matrice. La Pol III peut donc synthétiser en continu à partir d'une amorce sans dans le cas d'un brin retardé puisque l'ADN Pol III de ce brin est lié au complexe qui se déplace dans le sens de l'ADN Pol III du brin avancé (directeur), et doit donc faire par petite séquences en s'éloignant de la fourche de réplication. La processivité des ADN Pol III sont assurée par la pince β qui a pour rôle l'attachement ferme au brin

matrice. La terminaison de la transcription se fait lorsque deux fourches de réplication se rencontrent. Dans le cas de l'ADN circulaire, la fourche de réplication de droite va finir par rencontrer la fourche de réplication de gauche au site de terminaison. Ces deux fourches appartiennent au même œil de réplication puisqu'il n'y a qu'une seule origine de réplication, l'oriC.

Une animation montre la condensation de l'ADN et la réplication étapes par étapes (Anglais) : <https://www.youtube.com/watch?v=dKubylRiN84> ou <https://www.youtube.com/watch?v=TNKWgcFPHqw> Explication avec le réplisome (et chargeur de pince) : <https://www.youtube.com/watch?v=JZXT2uOcd2w>

Add video of DNA Replication here.

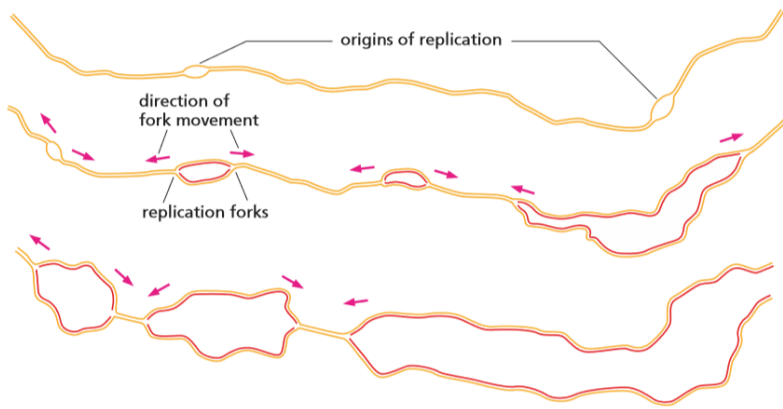
Link 3 : <https://www.youtube.com/watch?v=JZXT2uOcd2w>

REPLICATION CHEZ LES EUCARYOTES

Contrairement aux procaryotes, l'ADN des eucaryotes est linéaire et organisé en plusieurs chromosomes qui résultent de la condensation de la chromatine grâce à l'association de protéines liées à l'ADN. Puisque les chromosomes sont linéaires, il n'y a pas de sites de terminaison aux extrémités permettant de terminer la réplication : il faut donc de nouvelles activités enzymatiques pour prendre en compte les extrémités.

En dehors de ce problème, le processus de réplication est quasi identique pour l'eucaryote comme pour le procaryote. Les chromosomes eucaryotes possèdent plusieurs origines de réplifications afin d'accélérer la synthèse. Chaque œil de réplication finit par fusionner avec un autre et marque l'arrêt de la synthèse du nouveau brin pour chaque fusion de fourches de réplifications. Il peut également y avoir un arrêt de réplication par un signal de terminaison de réplication : les signaux « ter ». Chaque œil de réplication dans un chromosome suit le même schéma de réplication que chez le Procaryote :

1. **Initiation** : ouverture de l'hélice au niveau de l'origine de réplication formant un œil de réplication. Une ADN gyrase se charge de diminuer les tensions de torsions, l'hélicase ouvre l'hélice et sépare les deux brins matrice (un brin directeur et un brin retardé) tandis que des protéines stabilisatrices vont se fixer au brin matrice monocaténaire (le brin complémentaire n'étant pas encore synthétisé) et le stabiliser.
2. **Elongation** : le réplisome se charge de synthétiser deux nouveaux brins à partir des brins parentaux selon le modèle semi-conservatif. Un brin est le brin directeur et est synthétisé en continu par une ADN polymérase à partir d'une seule amorce d'ARN (qui a été faite par une primase) tandis que l'autre brin est un brin retardé qui doit être synthétisé par fragments d'Okazaki. Les amorces d'ARN sont ensuite remplacées par leur séquence ADN correspondante par une autre polymérase. L'ADN ligase va ensuite lier les fragments pour former un seul brin continu.
3. **Terminaison** : le réplisome quitte la fourche de réplication lorsqu'il fusionne avec une autre fourche ou qu'il reçoit un signal.



La quantité d'ADN de chaque chromosome pose un problème pour la réplication de l'information génétique : si chaque chromosome ne possédait qu'une origine de réplication, le temps qu'il faudrait pour répliquer le génome entier de la cellule serait prohibitif. De plus, l'empaquetage de l'ADN est une cause de ralentissement de l'étape d'élongation : chez les procaryotes, le réplisome synthétise un nouveau brin avec une moyenne de 1000 nucléotides par seconde. Chez l'Homme, la synthèse se fait uniquement à 100 nucléotides par seconde. Pour éviter que la réplication ne prenne trop de temps, il

faut donc plusieurs origines de réplifications. Chaque origine de réplication forme un œil constitué de deux fourches : la synthèse est donc bidirectionnelle. Une fourche d'un œil de réplication finit par fusionner avec une fourche d'un œil adjacent (fusion des réplicons).

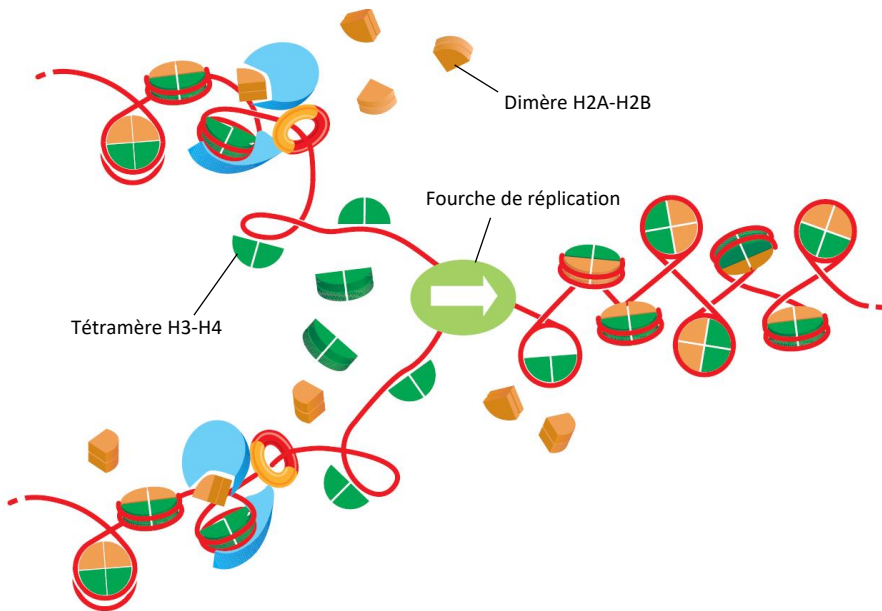
La condensation de la chromatine n'est pas la seule raison du ralentissement de la synthèse d'un nouveau brin par le réplisome. Le réplisome lui-même chez les eucaryotes est en fait plus complexe que chez les procaryotes. Nous avons vu que le réplisome des procaryotes utilisait principalement l'ADN polymérase III pour la synthèse des brins, tandis qu'une autre polymérase se chargeait uniquement de remplacer les amorces d'ARN (l'ADN Pol I). Chez les eucaryotes, plusieurs polymérases sont impliquées dans le réplisome et chacune a un rôle particulier pour effectuer la tâche que l'ADN Pol III accomplit toute seule. Les trois polymérases principales chez le mammifère sont l'ADN polymérase α , δ et ϵ . L'ADN polymérase ϵ synthétise le brin directeur en continu de la même façon que l'ADN pol III des procaryotes.

L'ADN polymérase α s'associe en un complexe avec une primase et forme des amorces ARN-ADN à partir desquelles l'ADN polymérase δ synthétise les fragments d'Okazaki du brin retardé. Les fragments d'Okazaki ont des longueurs allant de 100 à 400 nucléotides. Les origines de réplifications ne sont pas aussi spécifiques chez l'Eucaryote que l'origine *oriC* de *E. Coli*. La reconnaissance des origines de réplifications semble dépendre à la fois de la séquence et de la structure de la chromatine : au cours du développement de l'organisme, la condensation de la chromatine peut déterminer le nombre d'origines de réplifications afin d'augmenter la vitesse de synthèse lorsque la division cellulaire doit être rapide. Le niveau de complexité supérieur du réplisome des eucaryotes reflète l'existence de mécanismes de contrôles plus élaborés : la division cellulaire de certains tissus doit être lente tandis que pour d'autres, elle doit être rapide. Le réplisome est un complexe hautement régulé et surveillé par la cellule. De plus, la réplication de l'ADN doit être coordonnée avec un processus de division cellulaire. Nous verrons plus tard que la division cellulaire principale chez les Eucaryotes est la mitose. Toutes les cellules sont caractérisées par un cycle cellulaire comportant plusieurs phases : la cellule a un stade de division et un stade de préparation à la division. L'ADN nucléaire est répliqué lors du stade de préparation à la division, appelé interphase, à la phase S (pour DNA-Synthesis Phase).

A l'instar d'*OriC*, les origines de réplication chez Eukaryota sont des séquences riches en adénine (AMP) et en thymine (TMP) afin que la séparation des brins soit facilitée (l'adénine et la thymine ne font que deux ponts d'hydrogènes). Nous avons vu que l'origine de réplication des procaryotes par des protéines initiatrices pour débiter la réplication ; Chez les Eucaryotes, la phase d'initiation est plus complexe. Les séquences d'ADN servant d'origine de réplication chez les Eucaryotes contiennent un site de liaison pour un **complexe de reconnaissance d'origine** (ORC), une portion riche en bases A et T ainsi qu'un site de liaison pour d'autres protéines responsables de la régulation de la réplication en ajustant la structure de la chromatine.

Lorsque la cellule eucaryote n'est pas en phase de réplication de l'ADN nucléaire (phase S), le complexe de reconnaissance d'origine ORC est lié à l'origine de réplication mais n'est pas actif. Il sera activé lors de l'entrée en phase S une fois phosphorylé par une kinase. A l'état inactif, il forme un complexe de pré-réplication (pre-RC) avec des hélicases. Une fois activé, les hélicases peuvent alors ouvrir l'hélice et le complexe recrute les réplisomes au niveau des fourches du réplicon. Le complexe pre-RC reste au niveau de l'origine de réplication lors de l'élongation pour éviter qu'un autre ORC vienne synthétiser une deuxième fois un même réplicon. Nous reviendrons en détail sur ces mécanismes dans le chapitre de la division cellulaire.

Pour qu'un réplicon puisse être répliqué, il doit passer par une relaxation de la chromatine. De plus, la réplication des chromosomes n'implique pas seulement la réplication de l'ADN mais également la synthèse et l'assemblage de nouveaux nucléosomes derrière chaque fourche de réplication de manière à ce que les deux nouveaux chromosomes soient identiquement



condensés (nous verrons plus tard que les octamères ne sont pas assemblés aléatoirement et jouent un rôle majeur dans la régulation de la transcription de l'ADN). Le réassemblage des nucléosomes est un processus endergonique faisant intervenir différents complexes enzymatiques ainsi que des protéines chaperonnes chargées de remodeler les sous-unités des cœurs octamériques d'histones.

Le processus de réassemblage des nucléosomes ainsi que l'ensemble des régulations épigénétiques forment une discipline nouvelle et encore peu connue. Il les données actuelles sur l'épigénétique portent à croire qu'elle aura une place très importante dans la médecine de demain.

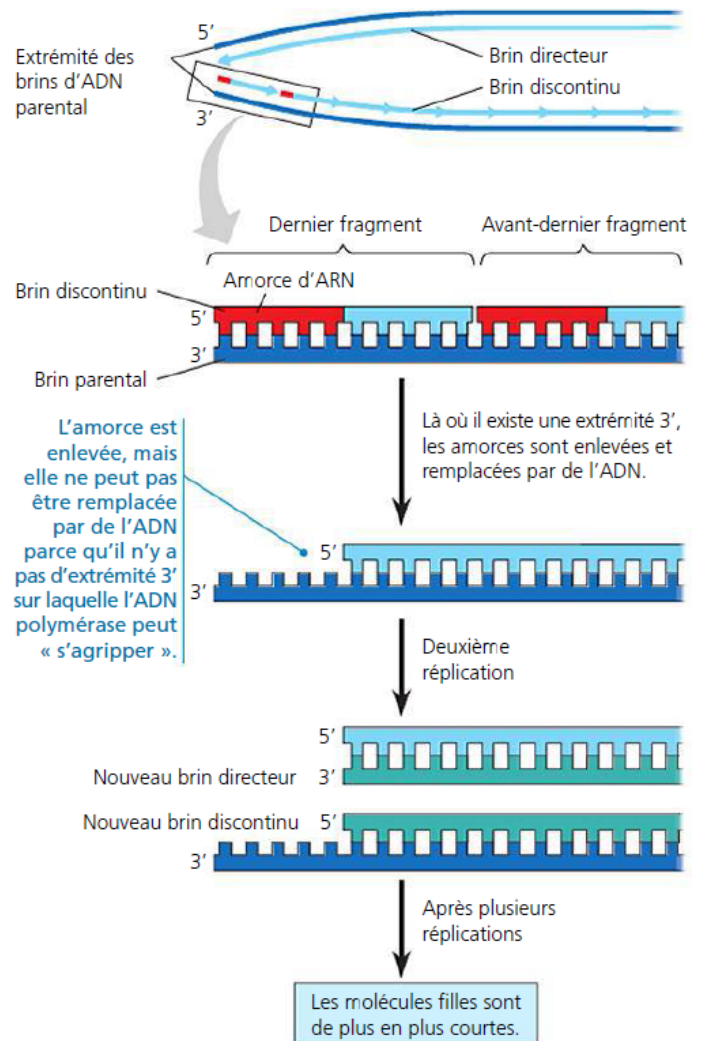
Alors que la fourche de réplication progresse le long de la chromatine, les histones sont déplacés transitoirement de manière à laisser environ 600 paires de nucléotides « à nu » pouvant être répliqué par le réplisome. Le rétablissement des nucléosomes se fait derrière la fourche de réplication par un processus encore peu connu.

TELOMERES

Outre le fait que la cellule Eucaryote a recours à différentes polymérases spécialisées et qu'elle utilise des amorces ARN-ADN, la réplication de l'ADN chez Eukaryota est un processus très semblable aux procaryotes. En revanche, la réplication de l'ADN rencontre une difficulté à laquelle elle doit impérativement pouvoir faire face : les chromosomes sont des molécules d'ADN linéaire et non circulaire comme chez *E. Coli*. Le fait qu'il soit linéaire pose un problème puisque dans le cas de *E. Coli* ou son ADN est circulaire, la terminaison de la phase d'élongation intervient lorsque les deux fourches de réplifications finissent par se rencontrer. Or, le chromosome n'a aucune autre fourche à rencontrer au niveau de ses extrémités qui pourrait lui indiquer la fin de la réplication.

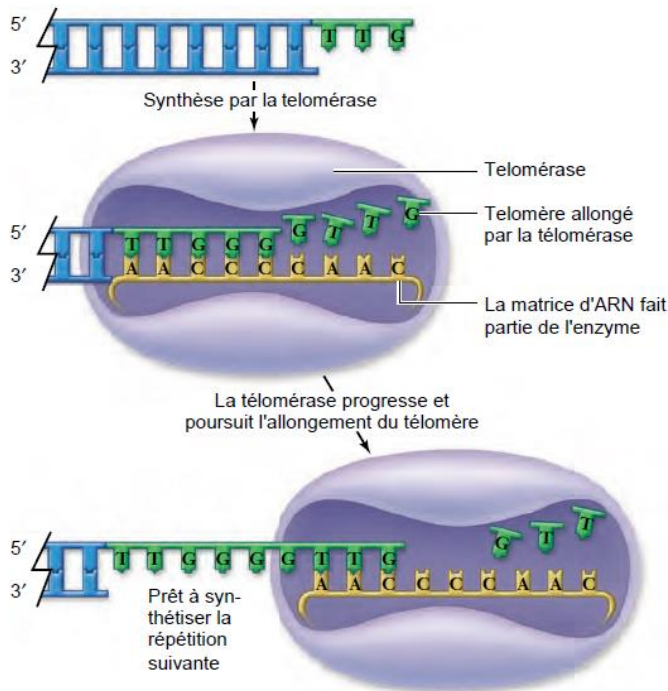
De plus, nous avons vu que la synthèse du brin retardé se fait de manière discontinue par un mécanisme de retrait qui produit des fragments courts d'ADN (les fragments d'Okazaki). Ce mécanisme rencontre un problème quand la fourche de réplication atteint l'extrémité d'un chromosome linéaire. Prenons le cas d'une simple molécule d'ADN linéaire bicaténaire contenant une seule origine de réplication. L'origine est reconnue par un ORC et initie la réplication. Nous avons vu que pour que l'ADN puisse être répliqué, les polymérases nécessitent une amorce -synthétisée par une primase – et ne pouvaient synthétiser que dans une direction (5'-3').

Pour le brin directeur, il ne suffira que d'une seule amorce à partir de laquelle l'ADN pol ϵ va synthétiser le brin en continu. Lorsque la polymérase arrive à l'extrémité, elle finit par arriver au bout de la molécule et s'arrêter, voyant qu'elle n'a plus rien à synthétiser. En revanche, le brin retardé va poser problème au



niveau de l'extrémité : en effet, le brin retardé est synthétisé via plusieurs amorces de manière discontinue et se fait dans le sens contraire au brin continu en plusieurs petits fragments. Il faut donc une amorce finale d'ARN à l'extrémité pour que l'ADN pol δ puisse synthétiser le dernier fragment. Néanmoins, cette amorce finale d'ARN ne peut être remplacée par son équivalent ADN par une polymérase qu'à partir d'un fragment précédent servant d'amorce à cette polymérase. Puisque cette amorce finale n'a pas de fragment derrière elle, elle ne peut être remplacée. Une fois l'amorce ARN clivée, l'ADN répliqué aura un bout non répliqué au niveau du brin 3'-5'.

Par conséquent, à chaque cycle de division, chaque chromosome perd une petite séquence d'ADN au niveau de ses extrémités. Si la cellule n'avait pas trouvé un moyen efficace pour remédier à ce problème, le raccourcissement graduel des chromosomes serait une catastrophe à la vie : chaque génération verrait son génome raccourcir de plus en plus, jusqu'à altérer les gènes fondamentaux à la survie de l'organisme. Les eucaryotes résolvent ce problème grâce à des séquences spécialisées aux extrémités de leurs chromosomes appelées les télomères. Les télomères sont des séquences qui ne sont pas codantes et dont la fonction est d'avoir un bout sacrifié à chaque cycle de division de manière à maintenir l'intégrité du chromosome et de ses gènes codants.



Les télomères sont caractérisés par des séquences courtes et répétitives. Chez l'Homme, cette unité de répétition est GGGTTA et est répétée environ une centaine de fois à chaque télomère. La cellule dispose donc d'une marge de division avant que le raccourcissement ne touche réellement le chromosome. Ainsi, au fur et à mesure qu'une cellule somatique se divise, ses télomères se raccourcissent graduellement.

La télomérase est l'enzyme qui permet de conserver la longueur du chromosome en ajoutant les unités de répétitions télomériques aux extrémités au cours de division cellulaire. Les unités télomériques du chromosome sont reconnues par la télomérase qui s'y lie et synthétise de nouvelles séquences à partir d'une matrice ARN faisant partie de l'enzyme. Cet ajout se fait sur le brin retardé pour conserver sa longueur, tandis que la polymérase ϵ du brin directeur s'occupe de recopier intégralement son brin parent sans soucis. La télomérase est donc l'enzyme qui permet de conserver la longueur des chromosomes et les télomères servent à maintenir l'intégrité du chromosome.

La télomérase n'agit pas sur toutes les cellules. En réalité, elle ne concerne que les cellules souches et les cellules germinales. En effet, si nos cellules germinales ne disposaient pas de télomérase, au cours de chaque cycle cellulaire, nos cellules reproductrices -et nos descendants- auraient des télomères de plus en plus courts jusqu'à ne plus pouvoir maintenir l'intégrité du chromosome. Les autres cellules somatiques ne possèdent pas - ou que très peu - d'activité télomérase. On a testé l'effet de l'inhibition des télomérases chez des souris ; Pendant 6 générations, les souris semblaient normales mais la longueur des télomères diminuant elles présentaient un vieillissement plus rapide. La 7^e génération n'était plus viable. Il y a donc une relation entre le vieillissement des cellules et la longueur de leurs télomères.

En fait, les télomères vont être une limite à la prolifération de la cellule. Les cellules saines ne semblent ne pouvoir se diviser qu'un nombre spécifique de fois (50 à 80 fois) avant de ne plus pouvoir se diviser. Cette limite de division est liée à la longueur des télomères et servirait de mécanisme contre l'apparition de tumeurs : au fur et à mesure qu'une cellule vit et prolifère, elle accumule inexorablement au cours de ses cycles des mutations. Bien que la plupart des mutations ne sont pas dramatiques ou mènent à la mort de la cellule, certaines mutations peuvent changer la machinerie cellulaire et déboucher sur une tumeur. Dans ces cas-là, lorsque la tumeur est maligne, elle ne cesse de proliférer de manière incontrôlée. Pour éviter d'accroître les chances de devenir anormale et dangereuse à l'organisme, la cellule se limite à un nombre spécifique de divisions.

La sénescence progressive des cellules est donc naturellement induite au fur et à mesure qu'elles se divisent et constitue l'un des mécanismes antinéoplasiques (= « anti-tumoral ») de notre organisme. La télomérase fonctionne un peu comme la rétrotranscriptase (ou transcriptase inverse), que nous avons introduit dans la section sur les virus. Cette découverte de la

possibilité de synthèse d'ADN à partir d'une matrice d'ARN a bouleversé la communauté scientifique et a marqué la fin d'un dogme. Il semblerait que les cellules cancéreuses disposent d'une forte activité télomérase leur conférant leur immortalité. On peut donc allonger également la durée de vie et le nombre de cycles de division d'une cellule grâce aux télomérases. La télomérase n'induit en aucun cas la malignité d'une cellule, elle lui permet juste de continuer à proliférer plus de fois que sans son activité.

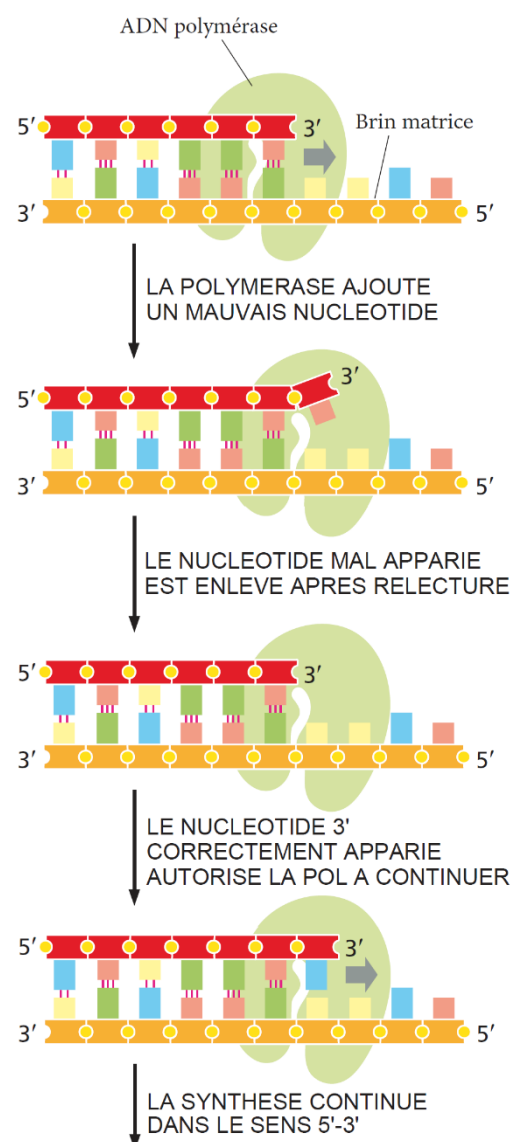
FIDELITE DE REPLICATION ET REPARATION DE L'ADN

Jusqu'ici, nous avons introduit le dispositif de réplication de l'ADN sans parler de la fidélité de réplication. Les polymérases peuvent être conceptualisées des enzymes « imprimantes », chargées de recopier notre information génétique avant chaque division. Cette imprimante est incontestablement l'une des meilleures qui existe grâce à sa rapidité mais aussi grâce à sa fidélité. En effet, la polymérase se trompe rarement mais sachant qu'une mutation peut affecter la conformation d'une protéine et sa fonction, certaines mutations peuvent avoir des conséquences dramatiques. Il faut impérativement maintenir une stabilité génétique pour la survie des organismes. Ce besoin de maintenir le génome intègre n'implique pas seulement un mécanisme de réplication extrêmement fidèle mais également de mécanismes de réparation lorsque des lésions surviennent. La cellule dispose de plusieurs mécanismes de réparation de l'ADN lorsqu'une erreur survient : elle peut réparer lors d'une erreur de réplication où suite à des dommages provoqués par des facteurs externes (exposition à des agents mutagènes).

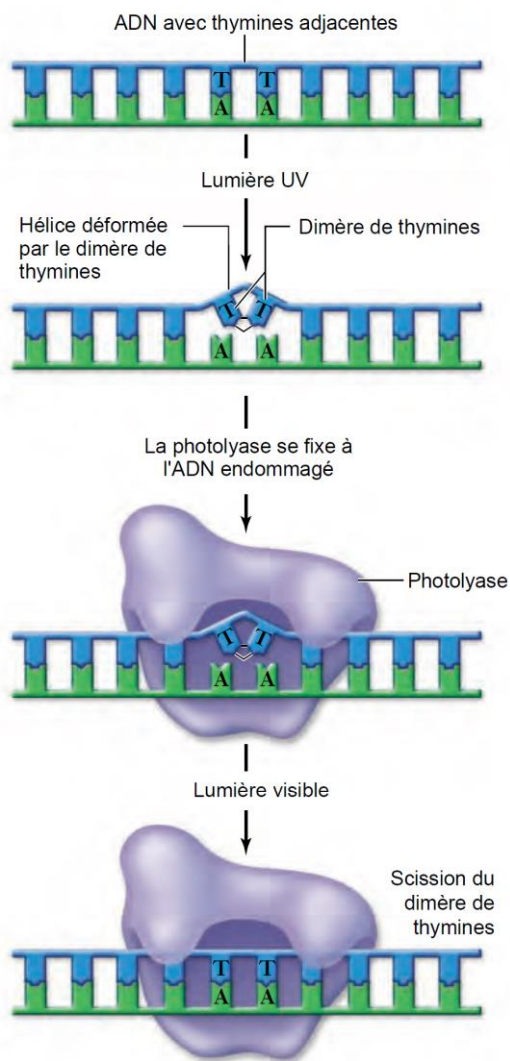
Des erreurs surviennent spontanément et sont inévitables lors de la réplication de l'ADN. On a constaté que les ADN polymérases font une erreur en appariant le mauvais nucléotide au brin matrice environ toutes les 10^7 paires de bases (soit 1 mauvais nucléotide pour 10 millions appariés). Sachant que l'Homme a un génome de 3400 Mpb (million de paires de bases, soit 3,4 milliards de paires de bases), l'ADN polymérase va commettre environ 300 erreurs lors de la réplication. Si ce nombre est petit comparé à la taille du génome, il représente néanmoins une menace non négligeable. Heureusement, certaines ADN polymérases sont douées d'une capacité d'autocorrection ; lorsqu'une erreur survient en phase d'élongation, elles sont capables de revenir en arrière et de corriger l'erreur par excision du mauvais nucléotide puis en remplaçant par le nucléotide complémentaire.

Les enzymes ADN polymérases sont caractérisées par une activité dite **nucléase** : elles peuvent cliver des liaisons phosphodiester d'un acide nucléique. On distingue deux types de nucléases : les endonucléases et les exonucléases. Les endonucléases vont couper l'ADN de l'intérieur tandis que les exonucléases vont couper au niveau des extrémités d'un fragment. L'autocorrection fait intervenir l'activité nucléase des polymérases en excisant le nucléotide mal apparié et en le remplaçant. Par exemple, l'ADN pol I chez E. Coli possède une activité exonucléase en excisant l'amorce ARN à l'extrémité du fragment et en la remplaçant par la séquence ADN équivalente. Grâce à leur activité exonucléase 3'-5', les ADN polymérases peuvent corriger les nucléotides mal appariés et ainsi maintenir le génome stable au cours du temps.

Lors de l'élongation, l'ADN polymérase veille constamment à ce que le nucléotide qu'elle vient d'apparier correspond au bon nucléotide complémentaire au brin matrice avant de procéder et ajouter le nucléotide suivant. Si le nucléotide n'est pas le bon, résultant en une paire de base A-C, G-T, T-C ou encore G-A, la polymérase revient en arrière et excise le nucléotide. La polymérisation (élongation) et la relecture sont deux activités étroitement coordonnées et chacune ayant lieu dans différents sites catalytiques de la polymérase. Cette capacité de relecture et de correction fait drastiquement chuter le nombre d'erreurs à la fin de réplication (environ 1 erreur pour 10^9 pb) mais il reste encore des erreurs qui non détectées. En plus des rares mutations persistantes causées lors de la réplication, les cellules sont constamment exposées



à des agents capables de produire des mutations, appelés agents mutagènes. Les radiations (UV et rayons X), ainsi qu'une large gamme d'agents chimique, sont capable de produire des mutations dans l'ADN et ainsi modifier notre information génétique.



La cellule dispose de mécanismes de réparation de l'ADN pouvant être spécifique ou non-spécifique. Avant d'introduire ces mécanismes, il est utile de savoir comment l'ADN peut être endommagé. La plupart des dommages de l'ADN sont provoqués par les rayons UV du soleil. La couche d'ozone agit comme un écran nous protégeant de la plupart des UV mais elle n'empêche pas tous les rayons de passer. De plus, il existe des zones géographiques présentant un « trou » dans la couche d'ozone où l'on voit une prévalence assez haute de cancer de la peau. Lorsque les UV - rayonnement de haute énergie - viennent léser l'ADN, ils provoquent la dimérisation des bases pyrimidines (thymine ou cytosine), notamment la dimérisation des thymines. Les dimères de thymine sont caractérisés par la formation de liaisons covalentes entre deux TMP (Thymine monophosphate) contiguës.

Cette dimérisation déforme l'hélice et empêche la polymérase de lire le dimère et de le répliquer (ou de le transcrire). La photoréactivation est un mécanisme de réparation des lésions de l'ADN spécifique aux dimères de pyrimidines. La photolyase est l'enzyme responsable de ce mécanisme de photoréactivation et puise son énergie justement grâce à la lumière (dans le spectre bleu/violet) et un cofacteur (la flavine réduite, FADH⁻). L'énergie lumineuse va permettre à la photolyase de transférer un électron (issu du cofacteur FADH⁻) au dimère de pyrimidine et provoquer la rupture de la liaison σ entre les deux bases pyrimidines et ainsi rétablir l'ADN. La photoréactivation est présente chez les Procaryotes et beaucoup d'Eucaryotes (dont la plupart des animaux) mais a été perdu chez les mammifères placentaires, dont l'Homme. L'Homme passe alors par un mécanisme de réparation non-spécifique (réparation par excision de nucléotide). Les UV ne sont pas uniquement la cause de dommages exogènes de l'ADN par la dimérisation des bases pyrimidiques : les UV sont des rayons ionisants pouvant provoquer des dommages indirects à l'ADN par la formation de radicaux libres ROS (dont nous avons parlé au chapitre 4). La réparation de

l'ADN pour ces dommages passe par un mécanisme de réparation non-spécifique.

La plupart des autres mécanismes spécifiques sont des renversements directs des lésions ; Lorsque l'ADN est lésé par agencement d'un groupement chimique, certains mécanismes spécifiques peuvent faire la réaction inverse et retirer le groupement fonctionnel. Par exemple, la base azotée guanine peut subir une réaction d'alkylation par agencement d'un groupement méthyl. La méthylation de l'ADN est un processus normal de régulation mais il arrive que l'alkylation se fasse sur une mauvaise position (en C6 au lieu de C5). Cette lésion peut être réparée par un mécanisme de réparation direct faisant intervenir une enzyme méthyltransférase qui va se charger de récupérer ce groupement et restituer la GMP.

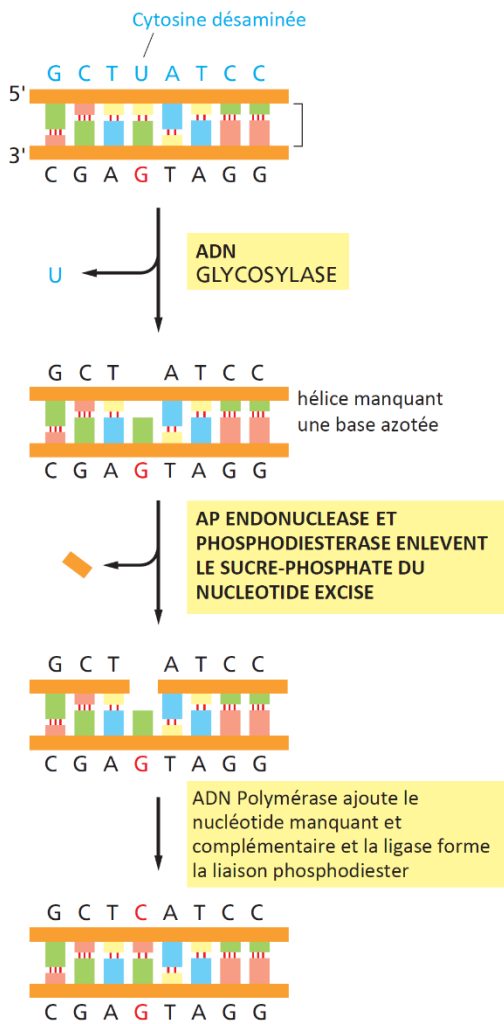
Les mécanismes de réparation de l'ADN non-spécifiques concernent les lésions ne pouvant être réparées par renversement direct (mécanisme spécifique) mais peut également s'opérer sur des lésions pouvant être réparées spécifiquement. Ces mécanismes fonctionnent par méthode d'excision : lorsque l'erreur est détectée, la base ou la séquence endommagée est excisée et remplacée en utilisant le brin complémentaire comme matrice. Il existe 3 principaux mécanismes de réparation non-spécifiques :

1. **Réparation par excision de base** : excision d'une base endommagée par des agents chimiques.
2. **Réparation par excision de nucléotide** : excision de plusieurs nucléotides.
3. **Élimination de mésappariement** : fait intervenir un système de relecture de l'ADN puis resynthèse de la base mésappariée suite à son excision.

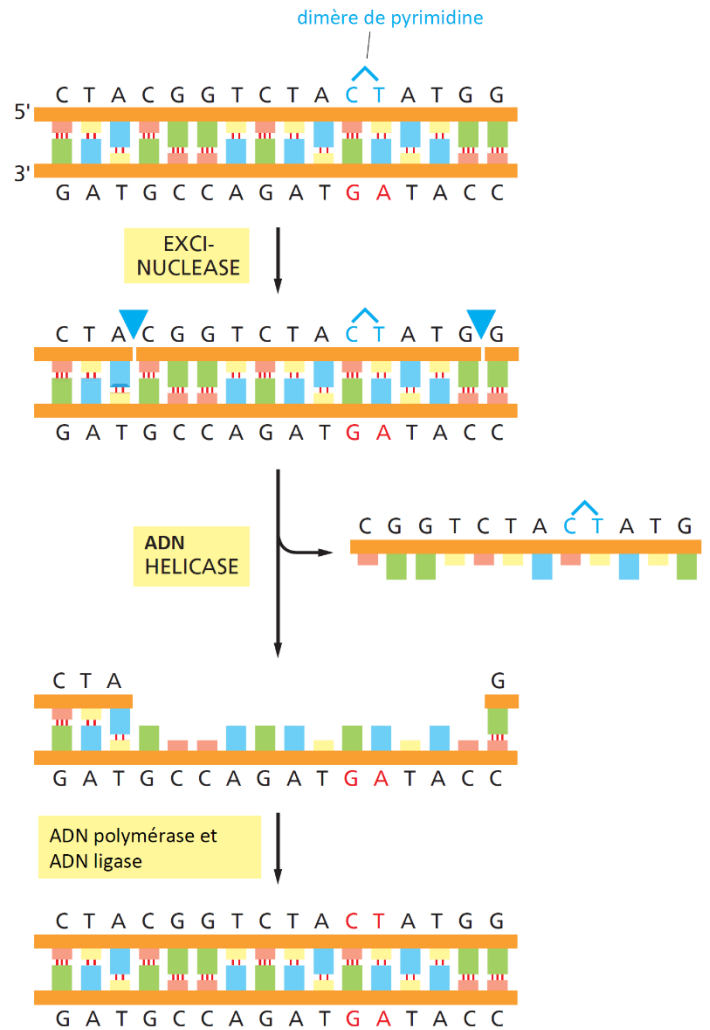
La désamination de la cytosine va transformer cette base azotée en une base Uracile. La réparation de cette lésion fait intervenir la réparation par excision de base. La réparation commence par l'excision de la base Uracile par un ADN glycosylase, enzyme qui rompt la liaison entre une base azotée et le désoxyribose du squelette pentose-phosphate. La réaction libère la base azotée et

laisse un site apurinique ou apyrimidique (site AP). Les sites AP sont réparés par une AP endonucléase qui coupe le désoxyribose-phosphate et excise le site AP et laisse l'hélice scindée avec un nucléotide manquant. Le nucléotide est ensuite remplacé par une ADN polymérase et l'hélice est ressoudé par une ADN ligase.

(A) REPARATION PAR EXCISION DE BASE



(B) REPARATION PAR EXCISION DE NUCLEOTIDE



La seconde voie principale de correction non-spécifique est la réparation par excision de nucléotide. Ce mécanisme peut réparer une lésion causée par dimérisation de pyrimidines ou encore par la liaison d'agents mutagènes provoquant une distorsion de l'hélice. On ne sait pas encore très bien comment la cellule choisit entre l'excision de base ou de nucléotide mais on pense que plusieurs facteurs influencent cette décision : le type de lésion, le cycle cellulaire et l'état fonctionnel de la cellule. La réparation par excision de nucléotide (REN) est en mode de réparation plus général que la réparation par excision de base (REB) car l'excision d'un oligonucléotide n'implique pas des enzymes spécifiques à la lésion comme c'est le cas de la glycosylase dans la REB.

Chez Procaryota, la réparation par excision de nucléotide (REN) commence par la détection de la lésion. Une fois détectées, des facteurs de reconnaissance se lient à la région endommagée et des hélicases viennent ouvrir le duplexe d'ADN dans la région endommagée. Des endonucléases viennent couper en amont et en aval de la lésion de l'ADN en 5' et en 3' pour exciser un oligonucléotide d'une longueur d'environ 12 nucléotides. Après l'excision, une ADN polymérase vient resynthétiser l'oligonucléotide manquant et la ligase soude le fragment pour reformer l'hélice. Chez les Procaryotes, l'endonucléase chargée de la REN est un complexe enzymatique appelé excinuclease qui consiste en trois enzymes *uvrA*, *uvrB* et *uvrC* (formant le complexe *uvrABC*). Chez les Eucaryotes, la REN est légèrement plus complexe mais le mécanisme est globalement le même.

Une mutation dans l'une des enzymes du complexe excinuclease peut rendre la réparation par excision de nucléotide inopérante et disposer l'individu à un plus grand nombre de mutations. Certains individus sont atteints d'une mutation touchant l'une des enzymes de la REN et ne peuvent s'exposer à la lumière du soleil : on les appelle les enfants de la lune. Cette maladie porte le nom

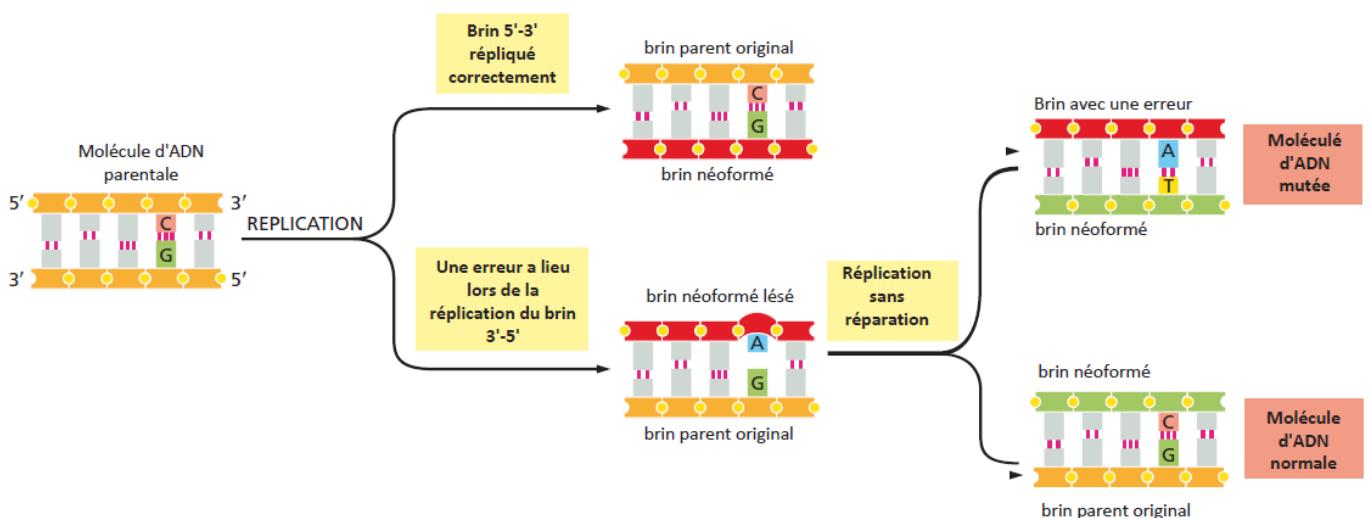
de Xeroderma Pigmentosum et se caractérise par l'accumulation de mutations exogènes, notamment en raison des dimères de thymine causés par les rayons UV. Le seul traitement actuel consiste à s'exposer le moins à la lumière du soleil et aux différents agents mutagènes. Il y a quelques années, la NASA avait développé des combinaisons intégrales de protection anti-UV pour ces personnes mais ces mesures sont très coûteuses et contraignantes.

Le troisième mécanisme de réparation non spécifique est l'élimination des mésappariements. La plupart des mésappariements sont éliminés lors de la correction à l'épreuve grâce à la relecture des ADN polymérases mais il arrive que des mésappariements ne soient pas détectés par l'ADN polymérase. L'ADN néoformé est scruté par un système de relecture afin d'identifier les mésappariements. Lorsque la base mal appariée est identifiée, elle est excisée puis remplacée par une ADN polymérase. Comment la cellule peut-elle savoir lequel des deux brins contient la base mal appariée ? Chez E. Coli, la méthylation de résidus d'adénine de séquences GATC permet reconnaître le brin parental et de distinguer brin parental et brin néoformé. Le système d'élimination des mésappariements peut alors éliminer les bases mal appariées du brin néoformé. Dans la plupart des organismes, le système de reconnaissance brin parent/brin néoformé exact n'est pas bien connu.

MUTATIONS ET REPOSE GLOBALE

Lorsque des dommages persistent malgré les mécanismes de réparation, les lésions sont alors des agents mutagènes. Il faut noter que l'ADN endommagé et une mutation sont deux choses distinctes : lorsque l'ADN endommagé n'est pas réparé, une mutation surviendra à la génération suivante de la cellule. En effet, l'ADN endommagé est en raison d'une anomalie dans la conformation de l'ADN (mésappariement, dimère T-T, alkylation, désamination, etc...) et peut être détecté par les mécanismes de correction. En revanche, une mutation est silencieuse et fait suite à la non-correction d'une erreur de l'ADN : une fois qu'elle a échappé à la relecture et l'ensemble des mécanismes de correction, le brin contenant le nucléotide désapparié servira de brin matrice lors de la prochaine réplication et ne sera pas détecté comme une erreur. On passe donc d'une erreur dans l'ADN à une mutation. La mutation est un phénomène qui est irréversible, contrairement aux dommages dans l'ADN qui sont détectables et réversibles. La mutation d'un gène peut entraver la fonction d'une protéine pour laquelle elle code tandis qu'un dommage dans l'ADN peut entraver le processus de réplication et de transcription (mécanisme que nous verrons au chapitre suivant).

La lésion devient une mutation lorsqu'elle n'est pas réparée et rentre en phase de division. Lors de la division, les chromatides sont séparés et les deux cellules devront à leur tour répliquer leur information génétique pour le prochain cycle de division. Le brin lésé servira alors de matrice (brin parental) lors de la réplication : si l'hélice est distordue, le brin néoformé ne peut être proprement répliqué et il en résulte des mutations. De même, s'il y a un mésappariement persistant, le brin matrice sera normalement répliqué mais l'hélice finale contient une séquence différente. Une mutation est une modification permanente dans l'ADN et peut avoir différentes conséquences selon le type de mutation.



Lorsque la cellule détecte une lésion mais n'a pas les outils pour la réparer, elle dispose de 3 choix possibles :

1. Suicide cellulaire : la cellule ne sait pas faire face aux lésions accumulées et choisit la mort pour préserver l'organisme pluricellulaire. Cette mort programmée porte le nom d'apoptose.