

CHAPITRE 0 : BIOCHIMIE STRUCTURALE DES ACIDES NUCLEIQUES

Les nucléotides sont formés par l'association :

- De pentoses (ribose ou désoxyribose),
- De phosphate(s) lié(s) au carbone C5 du pentose et
- D'une base azotée liée au carbone C1 du pentose.

Ce sont des substances biologiques ubiquitaires qui participent à de nombreux processus biologiques :

→ **Molécules énergétiques** : les nucléosides triphosphates et plus particulièrement l'ATP sont des molécules de haut potentiel énergétique.

→ **Rôle régulateur et informationnel** : La plupart des voies métaboliques sont régulées, du moins en partie, par les concentrations cellulaires en ATP, ADP et AMP. De plus, l'AMPc et le GMPc sont des messagers intracellulaires importants.

→ **Précurseurs de coenzymes du métabolisme** : Les nucléotides adényliques sont les précurseurs de coenzymes essentiels tels que NAD⁺, NADP⁺, FMN, FAD et le coenzyme A.

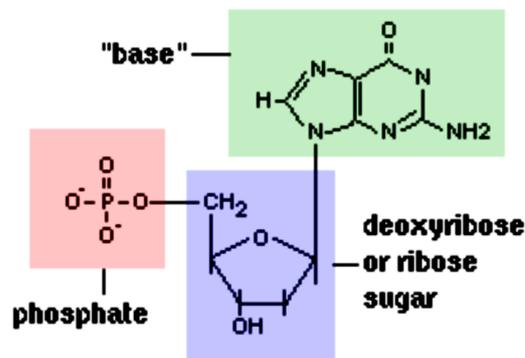
Les acides nucléiques, ADN et ARN, sont le fruit de la polymérisation de ces nucléotides. Ces molécules présentent des structures et des propriétés très caractéristiques. Ce sont notamment les supports de l'information génétique.

1. Nucléotides et désoxynucléotides

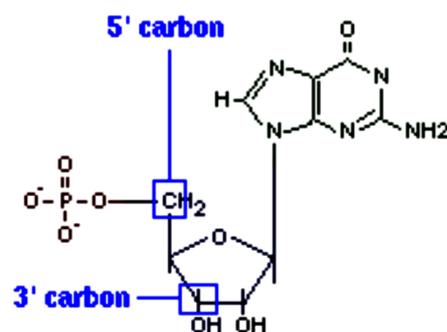
Dans ce cours, nous nous limiterons aux nucléotides impliqués en biologie moléculaire, c'est-à-dire les nucléotides entrant dans la composition de l'ADN et l'ARN. Les nucléotides particuliers (AMPc et GMPc) et les nucléotides entrant dans la composition des cofacteurs nucléotidiques (NAD, NADP, coenzyme A...) ne seront pas présentés.

Les nucléotides sont formés par un pentose (ribose ou désoxyribose), une base azotée (il en existe 5 différentes classées en 2 familles) et un ou plusieurs phosphates estérifiés sur le pentose (1 à 3 phosphates).

deoxyadenosine monophosphate

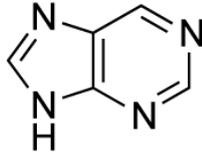


(ribo) adenosine monophosphate

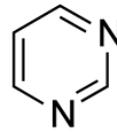


1.1. Les bases azotées

Les bases azotées dérivent d'hétérocycles azotés (cycles formés en partie par des atomes autres que le carbone). Il s'agit de molécules planaires, aromatiques, dérivés de la purine et de la pyrimidine.



purine (R)



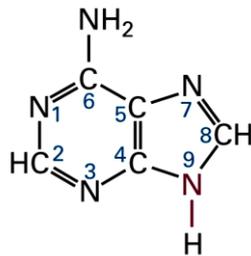
pyrimidine (Y)

On distingue donc deux familles de bases azotées : les bases puriques (A et G) et les bases pyrimidiques (U, T et C).

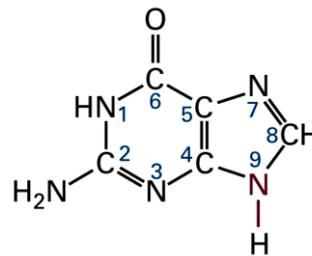
- Les bases puriques

Elles dérivent de la purine, qui comporte deux cycles accolés (dans le détail, un cycle pyrimidique accolé à un noyau imidazole). On distingue dans cette famille deux bases azotées, l'adénine (A) et la guanine (G). Les bases puriques forment des liaisons osidiques avec le ribose via l'atome N9.

PURINES



Adenine (A)

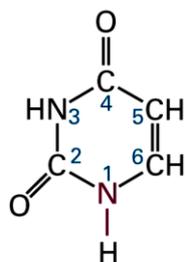


Guanine (G)

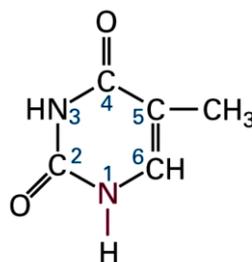
- Les bases pyrimidiques

Elles dérivent de la pyrimidine. Les sommets sont numérotés dans le sens des aiguilles d'une montre à partir du premier azote. Cette famille contient l'uracile (U), la thymine (T) et la cytosine (C). Les bases pyrimidiques se lient au ribose par l'intermédiaire de l'azote N1.

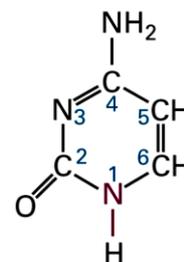
PYRIMIDINES



Uracil (U)



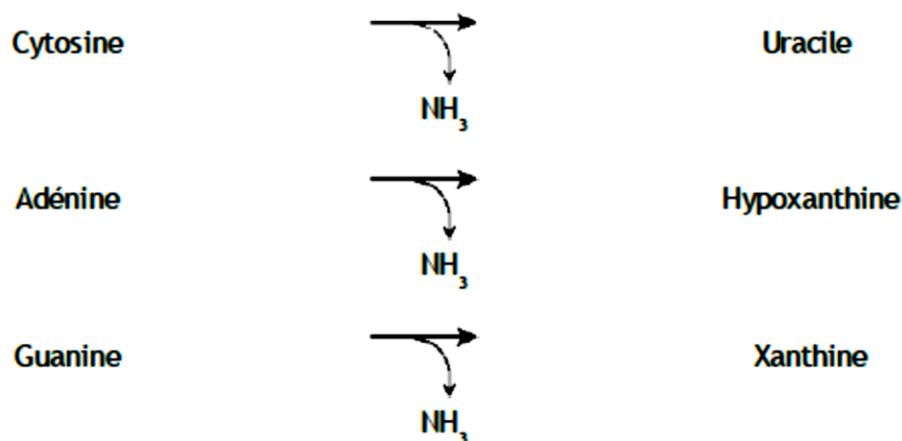
Thymine (T)



Cytosine (C)

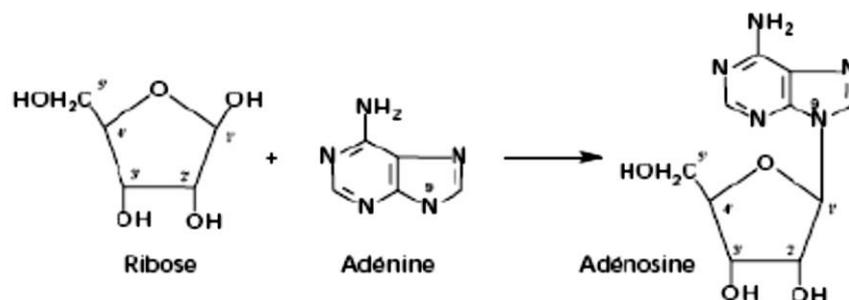
Propriétés physiques des bases azotées : Les purines présentent un spectre d'absorption dans l'UV. Cette propriété optique permet la détection, le dosage et le contrôle de la pureté en nucléotides et acides nucléiques. Pour cela, on fait le rapport des absorbances à 260 nm (acides nucléiques) et 280 nm (protéines contaminantes).

Propriétés chimiques des bases azotées : Les bases azotées sont des bases faibles ; elles peuvent fixer réversiblement des protons sur les azotes. A pH physiologique, les bases azotées sont sous forme basique non protonée donc neutre. Les bases azotées peuvent également subir une désamination oxydative par hydrolyse (réaction d'altération des bases azotées) :

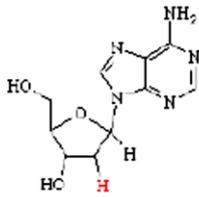


1.2. Les nucléosides

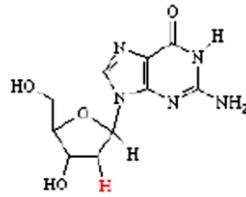
Les nucléosides résultent de la liaison d'une base azotée sur le carbone C1 d'un ribose ou d'un 2'-désoxyribose grâce à une liaison β -N-osidique (N9 pour les bases puriques ; N1 pour les bases pyrimidiques).



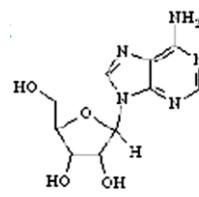
On distingue les désoxyribonucléosides, constitués par un 2'-désoxyribose lié aux bases A, T, C et G et les ribonucléosides constitués par un ribose lié aux bases A, U, C et G.



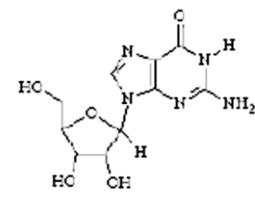
2'-Deoxyadénosine



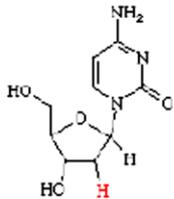
2'-Deoxyguanosine



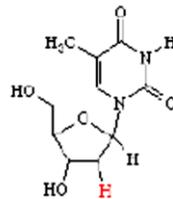
Adénosine



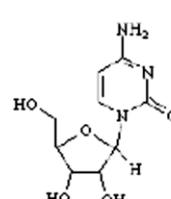
Guanosine



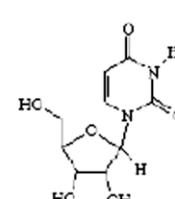
2'-Deoxycytidine



2'-Deoxythymidine



Cytidine



Uridine

Les déoxyribonucléosides

Les ribonucléosides

Chaque nucléoside porte le suffixe -osine si la base est purique et -idine si elle est pyrimidique :

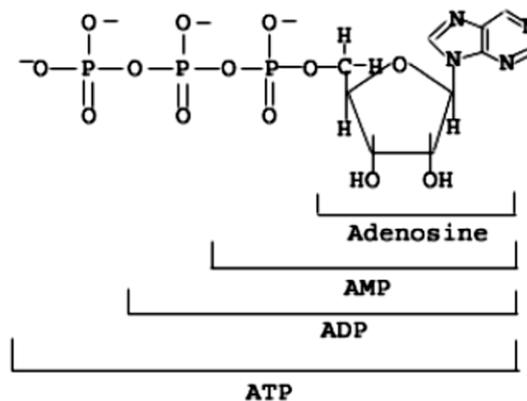
Base	Ribonucléoside	Déoxyribonucléoside
Adénine	Adénosine	Désoxyadénosine
Guanine	Guanosine	Désoxyguanosine
Uracile	Uridine	Désoxyuridine
Cytosine	Cytidine	Désoxycytidine
Thymine	Thymidine (très rare)	Désoxythymidine

NUCLÉOSIDE = OSE + BASE PURIQUE ou PYRIMIDIQUE

1.3. Les nucléotides

Un nucléotide est formé par l'estérification d'un (désoxy)nucléoside et de un ou plusieurs phosphates. On distingue :

- Les (désoxy)nucléotides monophosphate ou (d)XMP
- Les (désoxy)nucléotides diphosphate ou (d)XDP
- Les (désoxy)nucléotides triphosphate ou (d)XTP



Attention aux noms des liaisons

On distinguera les différents types de liaisons impliqués dans la formation d'un nucléotide :

- La liaison β -N-osidique entre le carbone C1 du pentose et la base azotée
- La liaison ester phosphorique entre le C5 du pentose et le premier groupement phosphate
- La liaison anhydride phosphorique entre les groupements phosphates, seule liaison à haut potentiel d'hydrolyse dans un nucléotide.

Les nucléotides sont généralement stabilisés par un ion Mg^{2+} .

Diversité et nomenclature des nucléotides :

Base	Ribonucléoside-5'-P	Désoxyribonucléoside-5'-P
Adénine	Adenosine-5'-monophosphate (AMP)	Désoxyadénosine-5'-monophosphate (dAMP)
Guanine	Guanosine-5'-monophosphate (GMP)	Désoxyguanosine-5'-monophosphate (dGMP)
Uracile	Uridine-5'-monophosphate (UMP)	Désoxyuridine-5'-monophosphate (dUMP)
Cytosine	Cytidine-5'-monophosphate (CMP)	Désoxycytidine-5'-monophosphate (dCMP)
Thymine	Thymidine-5'-monophosphate (TMP)	Désoxythymidine-5'-monophosphate (dTMP)

NUCLÉOTIDE = NUCLÉOSIDE + ACIDE PHOSPHORIQUE

2. Les polynucléotides

Les acides nucléiques sont des composés de masse moléculaire élevée contenant de l'azote et trouvés dans la cellule en association avec des protéines. Les complexes acides nucléiques-protéines sont appelés nucléoprotéines. Ils peuvent être dissociés en acides nucléiques et en protéines en milieu acide ou en présence de concentrations salines élevées.

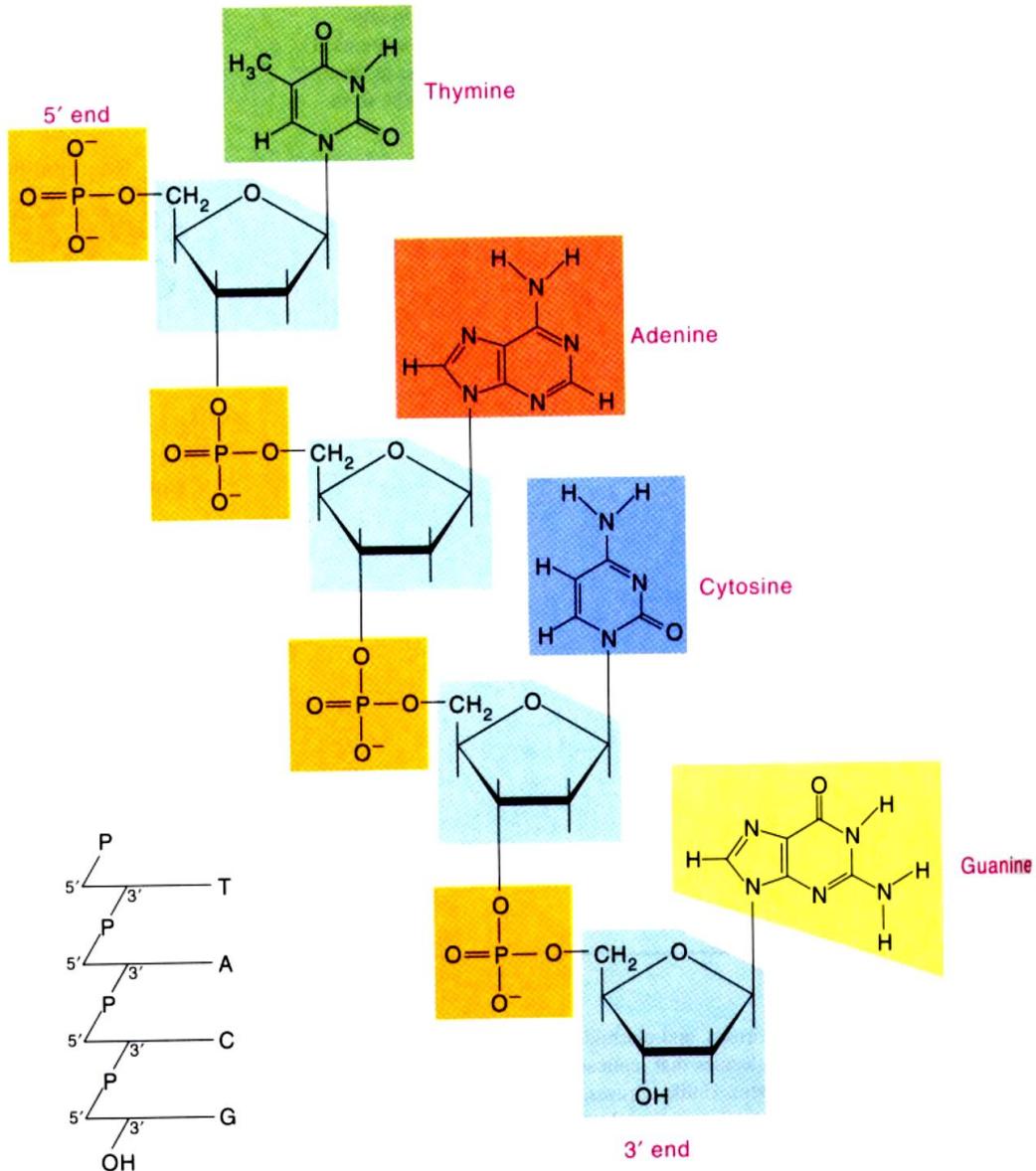
Il existe deux types d'acides nucléiques qui représentent 10 à 20% du poids sec de la matière vivante :

- Les acides ribonucléiques (contenant un ribose) ou ARN
- Les acides désoxyribonucléiques (contenant un 2'-désoxyribose) ou ADN

2.1. Des polymères de nucléotides

2.1.1. Polymérisation par liaison phosphodiester

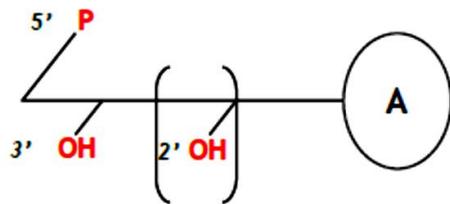
Les acides nucléiques sont des polymères de nucléotides dont les groupes phosphates font le pont entre les positions 3' et 5' de résidus riboses successifs.



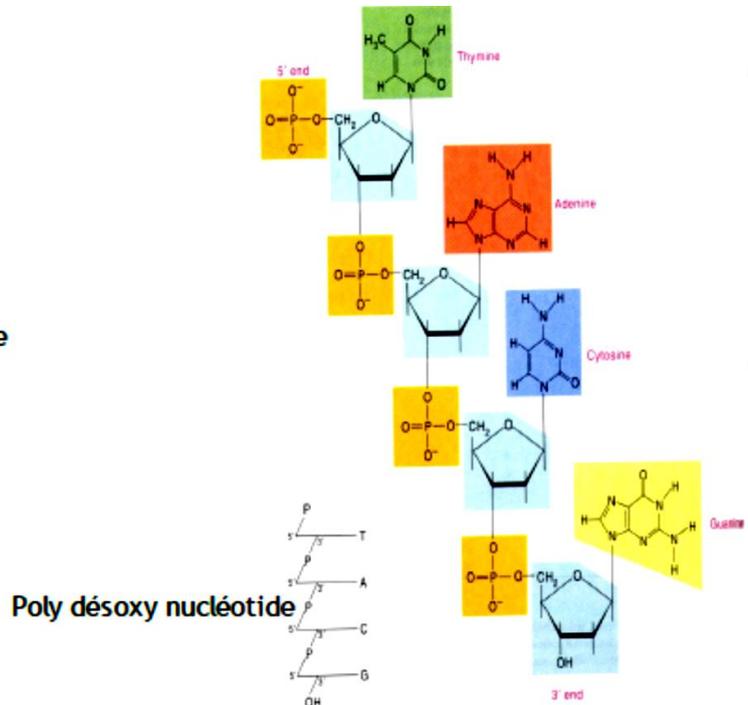
Les groupes phosphates, devenus phosphodiesters, sont ionisés à pH physiologique (pH = 7.4). L'ADN est donc chargé négativement à pH physiologique. En électrophorèse à pH=7, l'ADN migre donc vers l'anode. En général, l'extrémité 5' d'un brin d'ADN est phosphorylée et l'extrémité 3' ne l'est pas (3'-OH).

2.1.2. Symbole et représentation d'un acide nucléique

Pour ne pas représenter les molécules entières, il est possible d'adopter une représentation simplifiée des nucléotides et polynucléotides.



(Désoxy) Adénosine monophosphate



Par convention, on écrit la séquence d'un polynucléotide en partant de l'extrémité 5' et en allant vers l'extrémité 3'. La séquence présentée dans le document s'écrit alors : pTpApCpG ou encore plus simplement -TACG-.

Parfois, la présence de « d » ou « r » en amont de la séquence précise s'il s'agit d'un polymère de nucléotides (r) ou de désoxynucléotides (d).

Remarques :

- Il existe des enzymes capables de rompre les liaisons phosphodiester et qui peuvent ainsi dégrader l'ADN. Ces enzymes, appelées nucléases, DNases ou enzymes de restriction, sont très utilisées en biologie moléculaire. On distingue :
 - Les exonucléases, qui coupent le dernier nucléotide en 5' ou 3'
 - Les endonucléases, qui coupent au milieu de la molécule
- Les enzymes qui coupent les liaisons phosphodiesters dans l'ARN sont appelées RNases.
- Les acides nucléiques présentent également des structures primaires (séquence en nucléotides), secondaires (association en double hélice (ADN) ou repliement local (ARN)) et tertiaires (superenroulement).

2.2. L'ADN

En 1953, James Watson et Francis Crick ont décrit la structure tridimensionnelle de l'ADN. Cette brillante découverte est à placer parmi les plus importantes de l'histoire de la biologie.

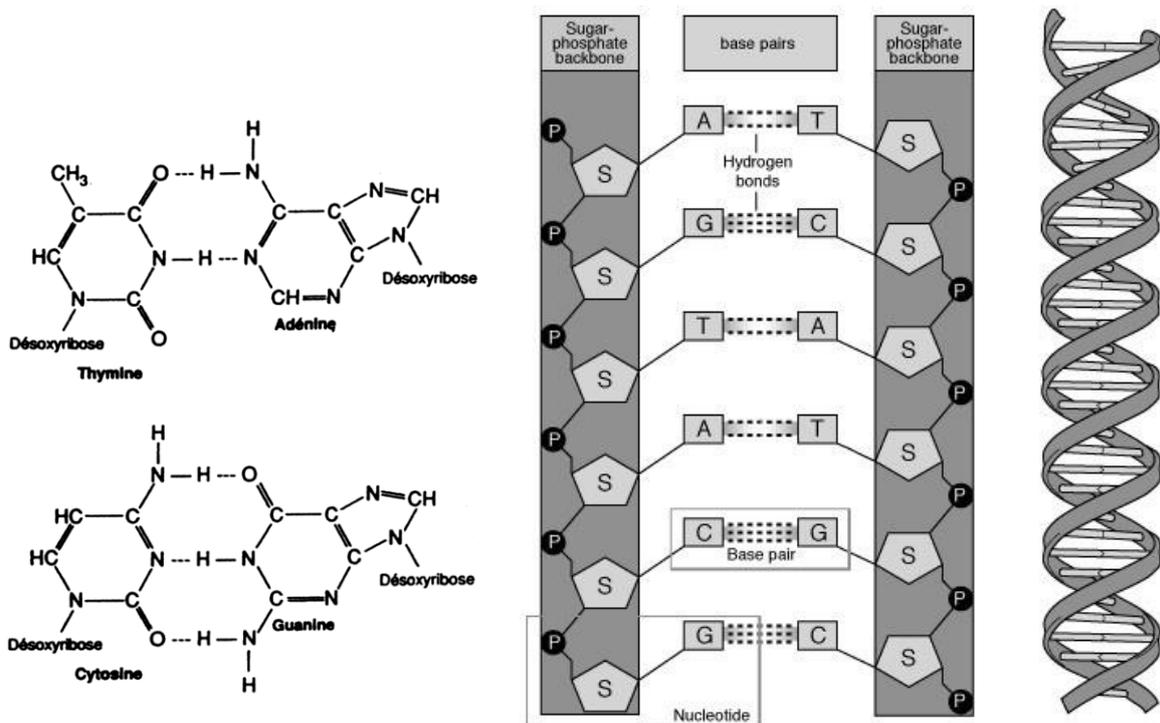
2.2.1. Une organisation en hélice double brin

Appariement des bases grâce aux liaisons H : Les bases azotées sont capables de s'apparier 2 à 2 grâce à des liaisons hydrogènes :

A = T

C ≡ G

Pour s'associer, les deux brins d'ADN sont antiparallèles, c'est-à-dire orientés en sens opposés, une chaîne dans le sens 5' → 3' et l'autre dans le sens 3' → 5'.



Cette structure double brin n'est pas stable. Ainsi, d'autres liaisons faibles vont stabiliser la molécule en lui conférant une structure hélicoïdale. L'hélice la plus rencontrée est l'hélice B mais il en existe d'autres.

La structure de l'hélice B d'ADN a été élucidée en 1953 par diffraction aux rayons X (Prix Nobel de Médecine 1962 – Watson et Crick ET Rosalind Franklin). Deux chaînes hélicoïdales de polynucléotides sont enroulées autour d'un axe commun et forment une double hélice dextrogyre (droite). Les 2

chaînes sont assemblées en paires de base par des liaisons hydrogène. L'appariement des bases doit limiter les encombrements stériques : une base purique (grande) est toujours associée à une base pyrimidique (petite). Des interactions hydrophobes de « stacking » (forces d'empilement) sont responsables d'interactions verticales entre les plans des bases empilées. Cette structure vérifie la règle de Chargaff : A est appariée avec T et G est appariée avec C. Les bases puriques et pyrimidiques sont à l'intérieur de la double hélice : leur plan est perpendiculaire à celui de l'hélice et à celui des molécules de désoxyribose. Le diamètre de l'hélice est de $20 \text{ \AA} = 2 \text{ nm}$. Les bases sont séparées de $34 \text{ \AA} = 3,4 \text{ nm}$. Aucune restriction n'est imposée à la séquence des bases le long de la chaîne polynucléotidique : la séquence précise des bases porte l'information génétique.

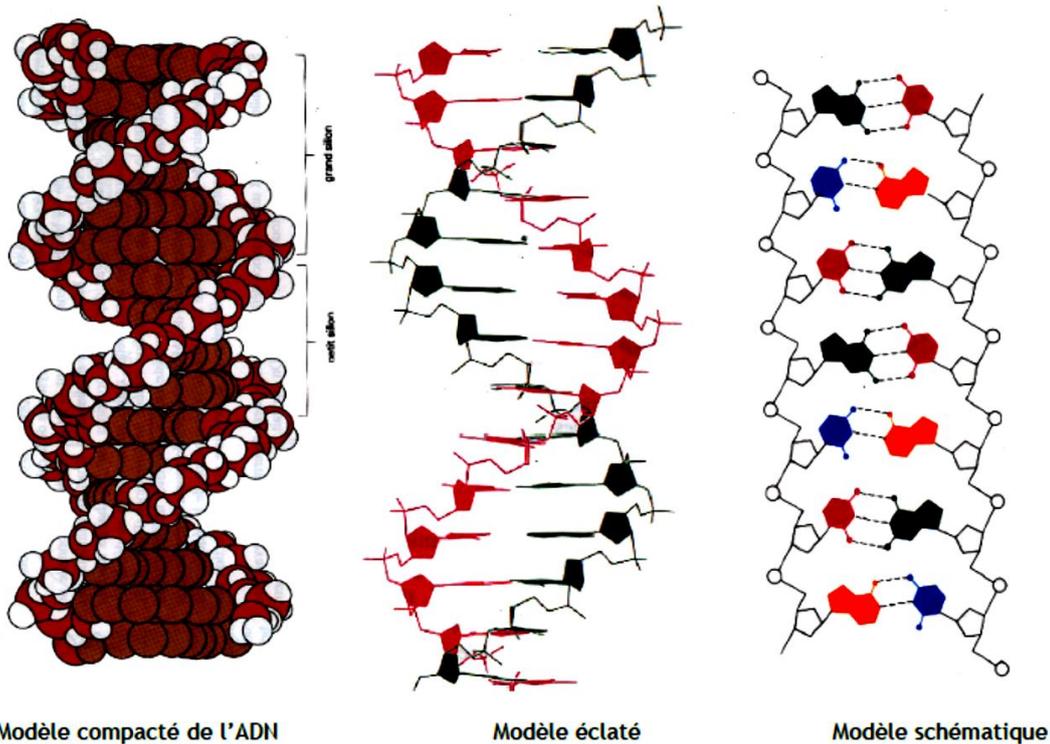


FIGURE 8-14 X-ray diffraction pattern of DNA. The spots forming a cross in the center denote a helical structure. The heavy bands at the left and right arise from the recurring bases.

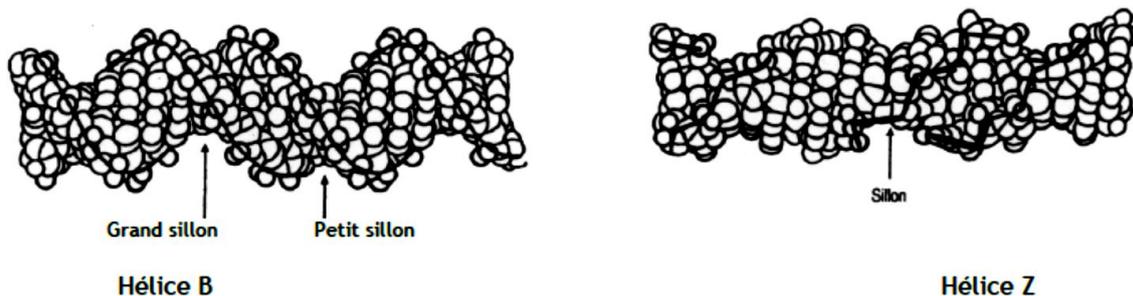


Rosalind Franklin,
1920-1958

La conformation B, celle décrite par Watson et Crick, est la plus stable dans les conditions physiologiques et donc considérée comme la forme native de référence des ADN. Mais la flexibilité relative de la structure lui permet d'adopter des conformations en duplex différentes dont deux sont bien caractérisées par cristallographie.

On décrit notamment les hélices A et Z :

	ADN B	ADN A	ADN Z
Hélice	droite	droite	gauche
Diamètre (°A)	19	26	18
Pb / tour	10,4	11	12
Pas de l'hélice	34	28	45
Remarque	<i>La plus stable</i>	<i>bactéries Gram- sporulant - résiste aux UV</i>	<i>complexes ADN / Protéine</i>

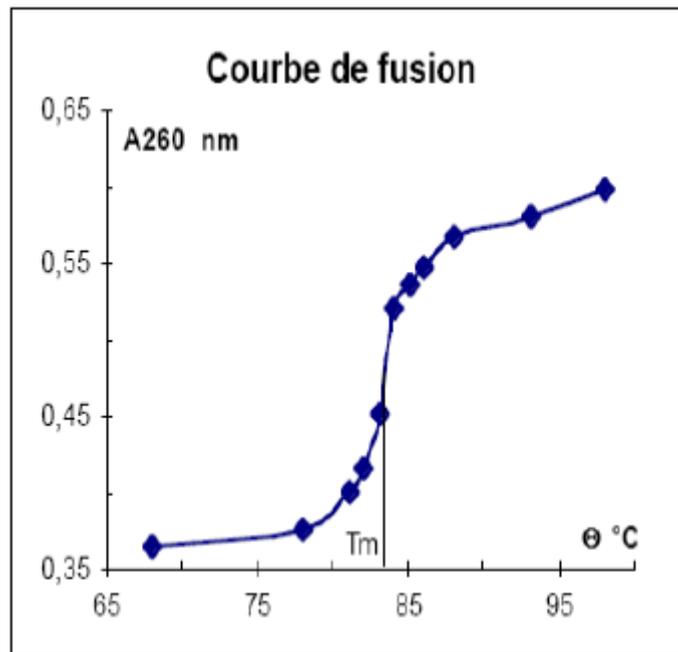


2.2.2. Propriétés de la molécule d'ADN

Solubilité : La molécule d'ADN est soluble en milieu salé. Les solutions d'ADN ont une très grande viscosité (en raison de la rigidité de la double hélice et de sa longueur). L'ADN est précipité par l'éthanol en longues fibres blanchâtres élastiques.

Absorption dans l'UV : L'ADN absorbe vers 260 nm du fait de la présence des bases azotées. La dégradation de la double hélice (chaleur, urée...) entraîne une augmentation de l'absorption (effet hyperchrome, car les bases azotées non impliquées dans des liaisons hydrogène absorbent davantage la lumière).

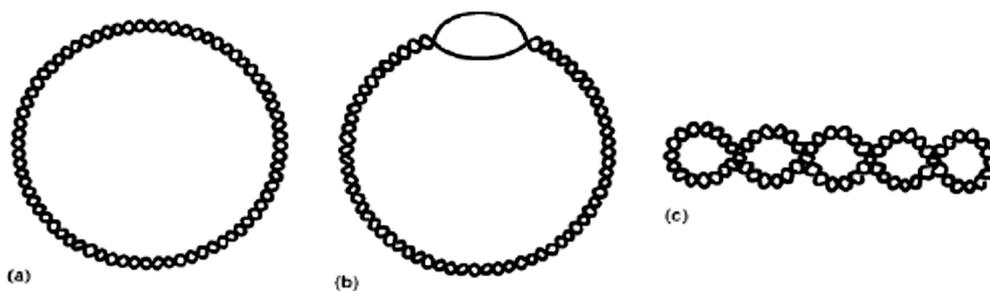
Dénaturation par la chaleur : Quand une solution diluée d'ADN est chauffée, on observe une augmentation de 30 à 40% de l'absorbance (effet hyperchrome) et une diminution de la viscosité. Cela correspond à la séparation des 2 brins de l'hélice bicaténaire, due à la rupture des liaisons hydrogène. Il n'y a pas de rupture des liaisons phosphodiester au sein de la chaîne polynucléotidique. La température de fusion (T_m pour melting) correspond à la température pour laquelle 50% des molécules d'ADN se sont séparées (fusion de la double hélice). Cette température augmente avec le pourcentage de paires GC, plus solidement appariées que les paires AT. Pour un fragment de moins de 50 paires de bases, la détermination approximative de T_m est donnée par l'équation $T_m = 2*(AT)+4*(GC)$. Cette dénaturation de l'ADN est réversible ; si le refroidissement est réalisé lentement, les brins se réappariaient et reconstituent l'hélice bicaténaire (ADN natif). Dans le cas d'un refroidissement rapide, la dénaturation est irréversible, la viscosité est diminuée, la masse molaire des molécules d'ADN formées est diminuée de moitié et l'effet hyperchrome subsiste. Remarque : la dénaturation peut aussi s'observer en présence d'agents chimiques comme l'urée.



2.2.3. Localisation et structure supramoléculaire de l'ADN

Chez les procaryotes : Chez les procaryotes, l'ADN est présent sous deux formes : le chromosome bactérien et les plasmides. Le chromosome bactérien est une longue molécule d'ADN circulaire double brin d'environ 10^6 paires de bases. Les plasmides sont des molécules d'ADN circulaires mais de plus petite taille (de 10^3 à 10^4 paires de bases).

Chez les bactéries, l'ADN présente un surenroulement (structure tertiaire). Il s'agit d'un enroulement de la molécule en sens inverse de celui de la double hélice. Cette conformation permet la compaction et la stabilisation de la molécule d'ADN.



Surenroulement d'un ADN circulaire

(a) Molécule d'ADN relâchée à 36 tours d'hélice. (b) Molécule d'ADN sous-enroulée à 32 tours d'hélice, ce sous-enroulement produit une séparation des brins. (c) même molécule que (b) mais avec 4 supertours.

Chez les eucaryotes : Chez les eucaryotes, on trouve de l'ADN principalement dans le noyau cellulaire (ADN nucléaire), mais également dans des organites dits semi-autonomes (mitochondries, chloroplastes). Les ADN mitochondriaux et chloroplastiques sont semblables aux plasmides bactériens.

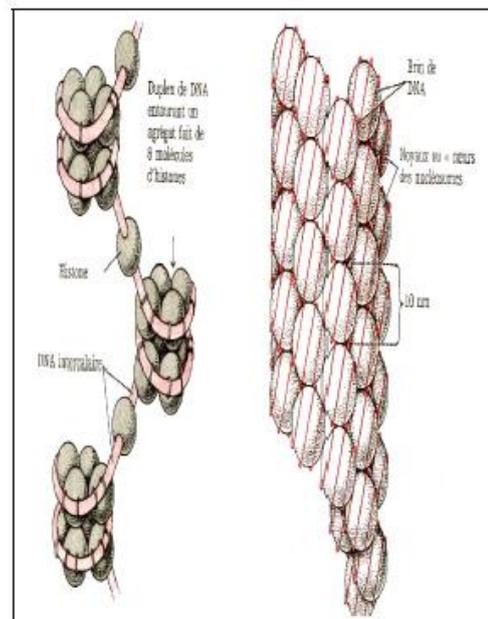
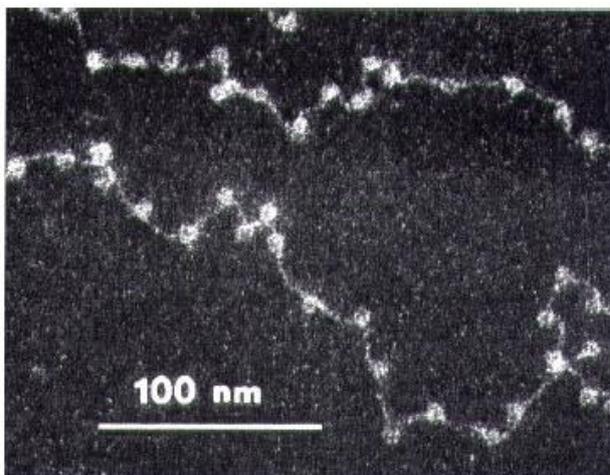
Concernant l'ADN nucléaire, plusieurs molécules d'ADN sont contenues dans le noyau cellulaire. Le nombre de ces molécules varie avec l'espèce :

Nombre normal de chromosomes dans différentes espèces	
Drosophile	8
Abeille	16
Maïs	20
Rat	42
Lapin	44
Homme	46
poulet	78

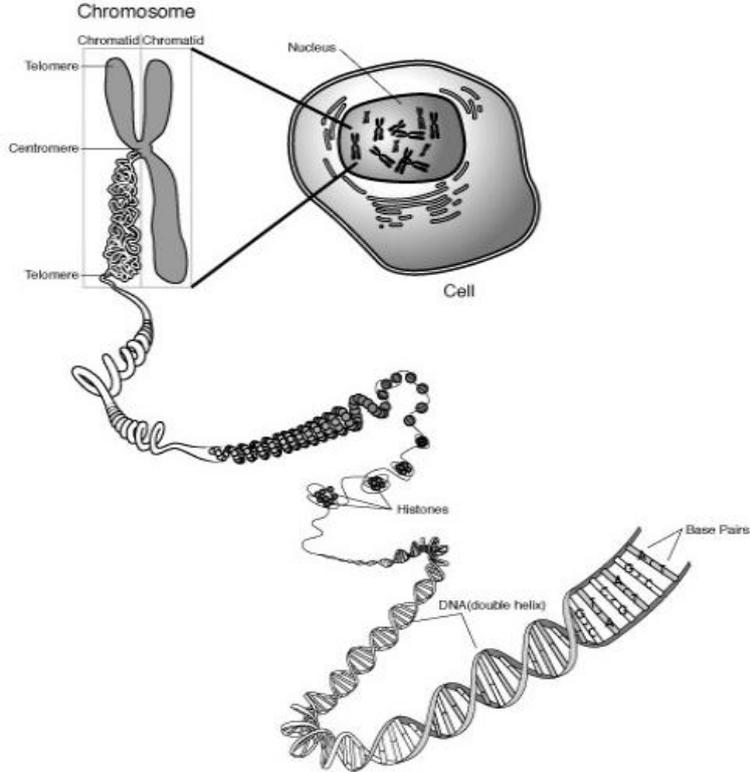
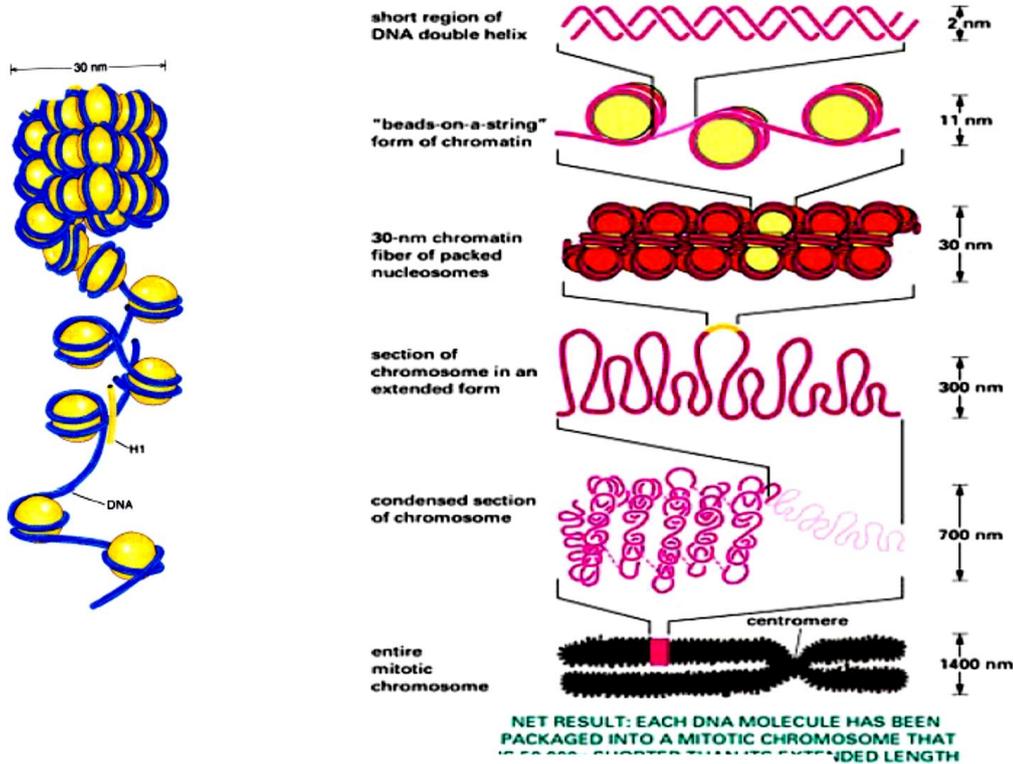
Chez les eucaryotes, l'ADN se présente la plupart du temps sous la forme de chromatine. Il s'agit d'un filament de nature nucléoprotéique. La chromatine est en fait composée de 60% de protéines, de 35% d'ADN et de 5% d'ARN.

L'ADN est associé, entre autres, à des protéines basiques appelées histones qui jouent un rôle dans la compaction de la chromatine et dans la régulation de l'expression génétique. L'association du double brin d'ADN et des complexes d'histones forme des unités structurales appelées nucléosomes.

Chromatine en microscopie électronique



Lors de la mitose, ces histones vont s'associer et permettre la compaction de la chromatine en chromosomes mitotiques visibles au microscope.



2.3. L'ARN

Le groupe des ARN est très hétérogène. Ce sont des molécules monocaténaire d'enchaînement de nucléotides dans le sens 5'→3'. Le désoxyribose est remplacé par le ribose et la thymine par l'uracile ; les ARN contiennent également quelques bases modifiées assez rares (ex. hypoxanthine). La chaîne peut se replier dans l'espace, et certaines bases s'apparient (régions en double hélice) mais la complémentarité est moins complète que dans l'ADN.

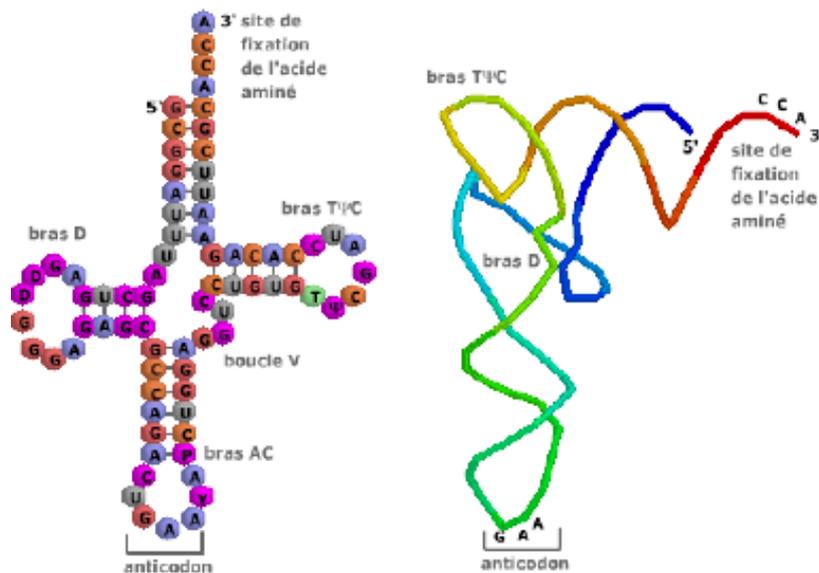
2.3.1. Les ARN ribosomaux représentent 80% des ARN totaux

Les ribosomes des eucaryotes sont formés d'ARNr associé à des protéines. Ils sont formés d'une grande sous-unité de 60S (coefficient de sédimentation) et d'une petite sous-unité de 40S qui forment ensemble une particule de 80S possédant une masse de 4,2 kDa (2,7 kDa pour les ribosomes de procaryotes).

La sous-unité 40S contient un ARN 18S qui est homologue à l'ARN 16S des procaryotes. La sous-unité 60S contient 3 ARN synthétisés à partir d'un même précurseur : 5S et 28S possédant des homologues procaryotes, et 5,8S spécifique des eucaryotes.

2.3.2. Les ARN de transfert

La structure des ARN de transfert a été résolue en 1974 grâce à la cristallographie. Elle comporte des segments de double hélice impliquant une dizaine d'appariements de bases et plusieurs sites qui ont la capacité de créer des liaisons hydrogène. Ils sont également formés à partir d'ARN précurseurs de plus grande taille. Il existe au moins 32 ARNt différents (ils représentent 15% de l'ARN total).



2.3.3. Les ARN messagers

Au contraire des ARNm des procaryotes, ceux des eucaryotes subissent une maturation très complexe. Ils ont 3 caractéristiques qui les différencient des ARNm procaryotes :

- Ils sont généralement monogéniques (ceux des procaryotes sont polygéniques)
- Ils possèdent à leur extrémité une queue de 100 à 200 résidus A successifs appelée queue poly-A en 3'
- Ils possèdent une coiffe en 5' (5' cap), constituée par un résidu méthyl-7-guanosine triphosphate

La queue poly A et la coiffe interviennent dans la stabilisation de la molécule d'ARN ainsi que dans le contrôle de la traduction.

Les ARNm eucaryotes sont épissés suite à leur biosynthèse. L'épissage consiste à retirer des parties d'ARNm qui ne serviront pas par la suite, appelé introns..

2.3.4. Les autres ARN

ARNsn : small nuclear RNA

ARNsno : small nucléolar RNA

ARNmi : microARN

Etc....

Comparaison ADN/ARN

L'ADN est un polymère de nucléotides qui constitue le support de l'information génétique au sein des cellules. Les ARN sont des polymères de nucléotides assurant l'expression de l'information génétique portée par l'ADN ainsi que sa régulation. **Quels sont les points communs entre ADN et ARN ? En quoi leurs structures leur permettent-elles d'assurer leurs fonctions biologiques complémentaires au sein des organismes ?**

1. ADN et ARN, des molécules séquencées

1.1. Des polymères de nucléotides

Les ADN et ARN sont des hétéropolymères séquencés, c'est-à-dire formés de plusieurs monomères différents et dont l'ordre est porteur d'un message. Ils sont constitués de 4 types de nucléotides. Un nucléotide comprend un sucre, ribose (ARN) ou désoxyribose (ADN), relié à un ou plusieurs groupements phosphates et à une base azotée de nature purique (adénine, guanine) ou pyrimidique : cytosine et soit uracile (ARN) soit thymine (ADN). L'ADN est bicaténaire et adopte une structure hélicoïdale, l'ARN est monocaténaire et adopte différents repliements spatiaux.

1.2. Localisation cellulaire de ces polymères

Le suivi de ces molécules peut s'effectuer par autoradiographie en utilisant de l'uridine ou de la thymidine tritiée. Chez les eucaryotes, l'ADN se trouve dans le noyau et dans les organites semi-autonomes (mitochondrie, chloroplaste) ; les ARN se trouvent dans le cytoplasme et dans le noyau.

1.3. Importance de l'ordre des nucléotides dans la réalisation de leurs fonctions

L'ADN constitue le support de l'IG : l'ordre des nucléotides (la séquence) au sein d'un gène détermine l'ordre des acides aminés au sein de la protéine correspondante. Les ARN peuvent être de différentes natures et coopèrent pour assurer l'expression de l'information génétique. L'ARNm constitue une copie transitoire du message de l'ADN, il est produit dans le noyau par complémentarité avec le brin transcrit grâce à l'ARN polymérase. Pour les autres ARN, la séquence détermine leur repliement spatial et leurs propriétés : prise en charge des AA pour les ARNt, catalyse de la liaison peptidique pour les ARNr, régulation de l'expression génétique pour les petits ARN.

2. ADN et ARN, des molécules de taille et de stabilité différente

2.1. L'ADN, une molécule pérenne et peu mobile

Le patrimoine génétique est le même pour toutes les cellules de l'organisme. La production continue mais régulée de protéines à partir du message contenu dans l'ADN assure la survie de chaque cellule. La nature bicaténaire de l'ADN assure sa stabilité et des protéines basiques, les histones, participent à sa compaction au sein du noyau chez les eucaryotes. L'ADN est une molécule de grande taille (plusieurs millions de paires de base pour chaque chromosome), qui ne peut pas sortir par les pores nucléaires : elle reste au sein du noyau. Il existe des protéines appelées facteurs de transcription qui peuvent se fixer dans le petit sillon de l'ADN et modifier l'expression génétique. Les pores nucléaires contrôlent les entrées de facteurs de transcription dans le noyau et donc régulent la production de protéines.

2.2. Les ARN, des molécules mobiles et de courte durée de vie

Les ARNm sont les copies mobiles de l'ADN : longues seulement de quelques milliers de nucléotides, elles peuvent sortir par les pores nucléaires. Les modifications des ARNm (épissage), leur sortie du noyau et leur durée de vie sont des processus étroitement contrôlés, qui vont déterminer les quantités de protéines produites. La durée de vie d'un ARNm est tout au plus de quelques heures.

Les ARNt sont mobiles dans le cytoplasme : ils sont chargés par voie enzymatique par les AA-ARNt-transférase, qui leur fixent des acides aminés en 3', et véhiculent ces AA jusqu'à leur lieu de traduction. Les ARNr sont assemblés avec des protéines au sein de gros complexes appelés ribosomes qui sont également mobiles dans le cytoplasme : ils se fixent sur les ARNm et assurent leur traduction. Les petits ARN sont mobiles dans le noyau ou dans le cytoplasme, ils vont participer à la régulation de la transcription (sn ou sno) et de la traduction (micro et ARNi).

3. ADN et ARN, des molécules comportant des bases complémentaires

3.1. Règle de Chargaff : les appariements AT et GC structurent l'ADN et l'ARN

L'appariement par liaison hydrogène des bases complémentaires AT et GC est à l'origine de la nature bicaténaire de l'ADN : ces appariements assurent la cohésion des brins de la double hélice. Ils sont aussi à l'origine de la structure des ARNt, par exemple : la structure en feuille de trèfle est permise par l'établissement de liaisons hydrogène entre bases complémentaires. Il en va de même pour les ARNr (structure très complexe).

3.2. Complémentarité des bases et coopération de l'ADN et des ARN

La nature bicaténaire de l'ADN assure sa stabilité, sa réparation et permet sa réplication avant mitose. Tout ceci est essentiel pour sa fonction de « stockage » d'information. La complémentarité des bases permet la création d'ARNm à partir de l'ADN : l'ARN polymérase fabrique l'ARNm par complémentarité avec le brin transcrit. La complémentarité des bases permet également la reconnaissance, au moment de la traduction, entre le codon de l'ARNm et l'anticodon de l'ARNt. Ainsi, les AA sont positionnés en face des codons des ARNm respectant le code génétique. La complémentarité des bases permet aux ARN interférents, de se fixer sur les ARNm et d'empêcher leur traduction au sein du cytoplasme.

Conclusion : dans les deux cas, on a un polymère séquencé de nucléotides pouvant former des appariements complémentaires. La stabilité de l'ADN lui permet de conserver le patrimoine génétique de la cellule. Le caractère éphémère et mobile des ARN leur permet de produire des protéines en continu, tout en s'adaptant aux conditions environnementales (signaux hormonaux, etc.). Ouverture : autres rôles des nucléotides (ATP, AMPc, coenzymes).

Conclusion : interactions ADN – protéines

D'après l'article « La riche union de l'ADN et des protéines », *Medecine/Sciences* 1987

La régulation de la transcription des gènes implique, dans toutes les cellules vivantes, des interactions entre l'ADN des séquences régulatrices et des protéines qui contrôlent, in fine, la fixation sur le promoteur de l'ARN polymérase. Plus généralement, d'ailleurs, de tels contacts entre des protéines et des acides nucléiques interviennent aussi dans la réplication de l'ADN, les réactions d'excision-épissage, la maturation des extrémités 5' et 3' des transcrits, le stockage et la traduction des ARN messagers, la formation des ribosomes et des autres particules ribonucléoprotéiques . . . c'est-à-dire dans la totalité des réactions mettant en jeu des acides nucléiques.

La mieux étudiée des interactions entre une protéine et un double brin d'ADN concerne le répresseur et la protéine Cro du phage lambda et leurs sites opérateurs de fixation. La fixation alternative de ces deux protéines aux sites opérateurs du phage contrôle soit le maintien de l'ADN phagique à l'état intégré dans le chromosome bactérien (phénomène de lysogénie), soit le déclenchement du cycle lytique.

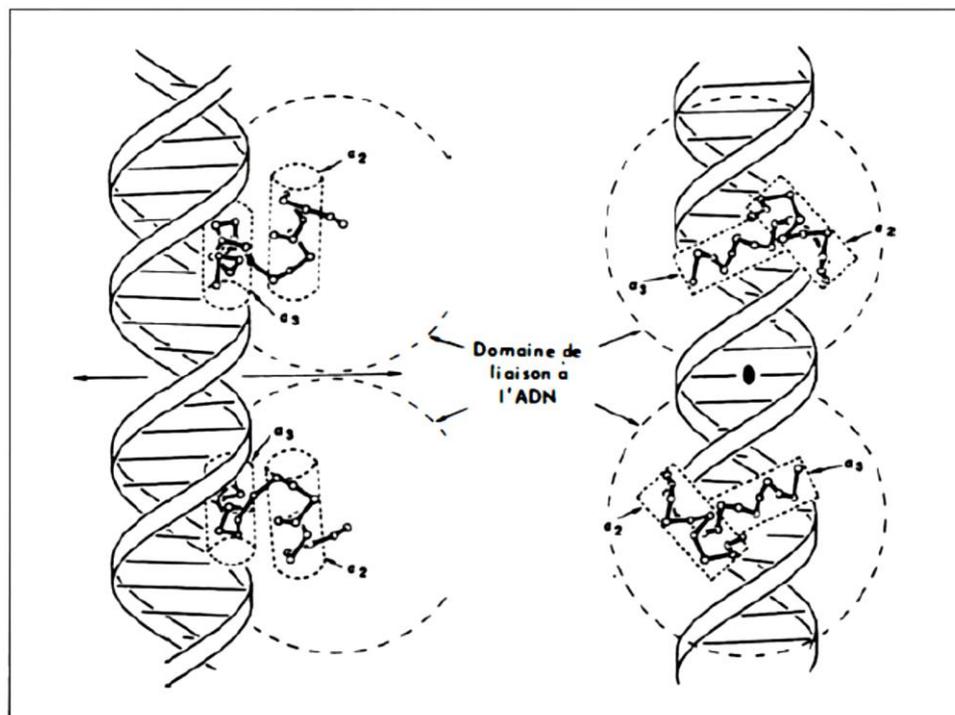


Figure 1. **Interactions entre le répresseur et un opérateur du phage lambda : modèle « hélice-tour-hélice ».** Deux des quatre hélices α de chaque monomère, α_2 et α_3 , établissent des contacts avec, respectivement, le squelette de ponts phosphodiester et les bases du grand sillon du double brin d'ADN. Deux monomères sont ainsi fixés sur une même face de l'ADN, liés à des séquences situées à un tour d'hélice de distance. La flèche (schéma de gauche) et l'ovale (schéma de droite) indiquent le centre de symétrie des sites de fixation. Les deux monomères sont liés entre eux par des interactions hydrogènes et hydrophobes.

m/s n° 7 vol. 3, septembre 87

La figure 1 montre un schéma des interactions entre un dimère de répresseur et la double hélice d'ADN : chaque monomère de répresseur est composé de quatre hélices α reliées par des feuilletts β . Une hélice α se positionne dans le grand sillon de la double hélice d'ADN, parallèlement à

ses berges, établissant des contacts étroits et spécifiques avec certaines des bases. Une autre hélice α , perpendiculaire à la précédente, établit des contacts avec le squelette de ponts phosphodiester de l'ADN. Les deux monomères sont fixés de façon identique sur la même face de l'ADN, à un tour de spire de distance de la double hélice (c'est-à-dire 10 paires de bases) (figures 1 et 2A).

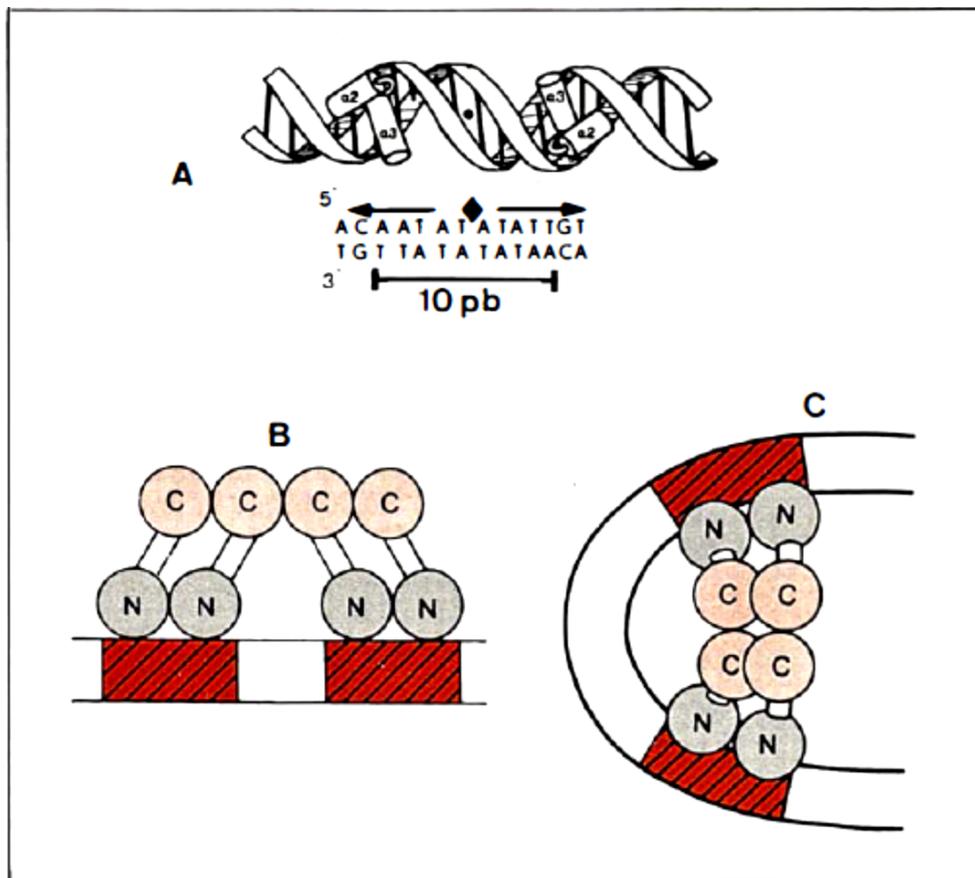


Figure 2. Coopération dans la fixation du répresseur du phage lambda à l'opérateur et induction de courbures de l'ADN. La figure 2A montre un opérateur « canonique », constitué d'une séquence palindromique, dont les centres de chaque héli-site sont espacés de 10 paires de bases, c'est-à-dire d'un tour d'hélice. Les séquences des héli-sites de fixation sont symétriques, de même que les monomères interagissant avec eux. Le centre de symétrie est symbolisé par un point noir au niveau du double brin et un losange au niveau de la séquence. La figure 2B montre la fixation de deux dimères à des sites proches (en rouge hachuré) du double brin d'ADN : les extrémités N-terminales des monomères forment la région de liaison à l'ADN, alors que les extrémités C-terminales interagissent entre elles, entraînant un phénomène de coopération dans la fixation des molécules. En 2C, si les sites opérateurs sont éloignés l'un de l'autre tout en demeurant sur une même face de l'ADN, l'interaction entre les molécules de répresseur entraînera une courbure du double brin... et la fixation restera coopérative.

Des dimères liés à des sites proches sur la même face de l'ADN établissent entre eux de puissantes liaisons hydrogènes et hydrophobes, de telle sorte que leur fixation sur l'ADN est hautement coopérative (figure 2B) : la fixation d'une molécule de répresseur interagissant avec une autre molécule augmente la concentration locale de ces protéines et donc l'affinité pour elles du site d'ADN non encore occupé. Cette coopération se manifeste encore lorsque les sites de reconnaissance sont éloignés l'un de l'autre, à condition cependant qu'ils soient toujours situés sur une même face de l'ADN, c'est-à-dire qu'ils restent séparés d'un nombre entier de tours de spire.

La figure 2C montre que la conséquence d'une telle fixation est une courbure imposée au double brin d'ADN, amenant un « bâillement », du grand sillon à l'extérieur de la courbure et son pincement à l'intérieur. Une telle courbure de l'ADN n'est d'ailleurs pas forcément la conséquence de la fixation d'une protéine : elle peut au contraire la précéder et la faciliter. Le site d'ADN reconnu par la protéine est alors un double brin incurvé. Les séquences séparant deux points de contact de la protéine et de l'ADN, séquences n'ayant elles-mêmes aucune interaction directe avec la protéine, peuvent ainsi influencer sur l'affinité pour cette protéine (figure 2). Il existe probablement une correspondance assez étroite entre les acides aminés et les bases au niveau des sites de reconnaissance et, dans des cas très privilégiés, on a pu rétablir l'affinité d'un site muté pour sa protéine spécifique en mutant également celle-ci au niveau du résidu interagissant avec la base modifiée de l'ADN.

Le modèle d'interaction que l'on vient de décrire, appelé « hélice - tour- hélice » se retrouve dans bien d'autres systèmes que celui du phage lambda, dans la levure comme chez des eucaryotes supérieurs. L'interaction de « l'homéobox » des gènes homéotiques avec l'ADN serait de ce type. La particularité des sites d'ADN participant à ce type de contact est d'être « palindromiques », c'est-à-dire d'avoir une symétrie en miroir et de comporter ainsi des séquences identiques sur les deux brins (figure 2A) interagissant avec deux monomères de la protéine spécifique, également symétriques l'un par rapport à l'autre.

Un second modèle de liaison d'une protéine à une séquence spécifique d'ADN est celui des protéines dactyles dont l'archétype est TFIII-A, un facteur de transcription se fixant à un promoteur spécifique de l'ARN polymérase III et initialement caractérisé chez une espèce d'amphibien, le Xénope (*Xenopus laevis*). Le promoteur des gènes transcrits par l'ARN polymérase III a la particularité remarquable d'être localisé dans la séquence transcrite elle-même. La séquence de TFIII-A comporte un domaine N-terminal de 30 kDa, composé de 9 répétitions de 30 acides aminés, parmi lesquels 2 cystéines et 2 histidines placées en des positions invariantes. Miller et al. ont suggéré que ces répétitions correspondaient à des structures en doigts, stabilisées par un atome de zinc échangeant 4 liaisons de coordination, 2 avec les cystéines et 2 avec les histidines (figure 3). Différentes méthodes de protection du complexe TFIII-A/ADN contre la digestion par des nucléases ont indiqué que la zone d'interaction était de 50 bases et n'intéressait qu'un seul des deux brins, avec des points de contact toutes les 5 bases, soit tous les 1/2 tours de spire de l'ADN, selon l'un des deux modèles proposés par Fairall et al. (figure 3) . Les particularités de la liaison de TFIII-A à l'ADN apportent un élément d'explication à la manière dont cette protéine, fixée dans le gène lui-même, peut ne pas interrompre la transcription par l'ARN polymérase III : on peut supposer que les doigts proximaux se détachent de leur site lors du « passage » de la polymérase, puis se fixent à nouveau, alors que la dissociation des doigts distaux laisse le champ libre à cette polymérase.

L'intérêt pour les protéines dactyles a été considérablement renforcé par la découverte récente de possibles structures en doigts au niveau de nombreuses protéines nucléaires, notamment des produits de gènes intervenant dans le développement de la drosophile et au niveau des récepteurs des hormones stéroïdes et de facteurs transcriptionnels ubiquitaires, tel Sp1. Il existe, cependant, d'importantes différences entre ces dernières protéines et TFIII-A, touchant au nombre et à la nature des « doigts ». De plus, les récepteurs des hormones stéroïdes se fixent en général en amont des gènes transcrits par l'ARN polymérase II et n'ont pas à être « déplacés » par le passage de la polymérase, si bien que leur ressemblance structurale et fonctionnelle avec TFIII-A pourrait être en fait assez lointaine. Le fait que ces récepteurs, contrairement à TFIII-A, interagissent avec les deux brins d'ADN, va dans le même sens.

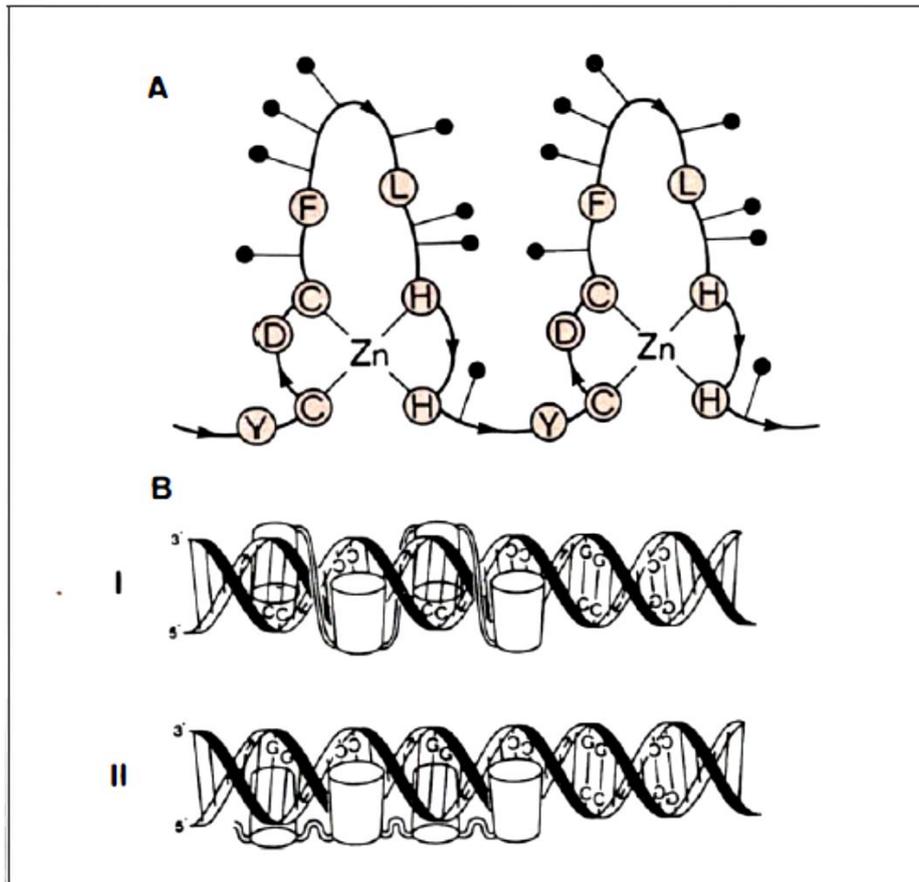


Figure 3. Protéines dactyles et modèle de leur interaction avec l'ADN. En 3A sont représentés, selon Miller et al. [7], deux « doigts » protéiques stabilisés par des liaisons de coordination entre un atome de zinc d'une part, 2 cystéines (C) et 2 histidines (H) d'autre part. Ces doigts sont constitués au niveau d'éléments répétitifs contenant, outre les cystéines et histidines invariantes, des acides aminés acides (D = acides aspartiques) et hydrophobes (F = phénylalanine, L = leucine, Y = tyrosine). Les points noirs indiquent les contacts probables avec l'ADN. La figure 3B présente, selon Fairall et al. [10] deux modèles possibles d'interaction entre le facteur TFIIIA et l'ADN. Les doigts sont représentés par les cylindres. Selon le modèle I, la protéine est enroulée le long d'un des brins d'ADN, les doigts établissant des contacts avec lui tous les 1/2 tours de spire. Selon le modèle II, la protéine est parallèle au double brin, les doigts établissant des contacts alternativement sur l'une et l'autre face du double brin.

De ce qui précède, il faut retenir qu'il existe des interactions très spécifiques entre l'ADN et des protéines jouant un rôle essentiel de régulation de la réplication et de la transcription. L'anatomie des complexes ADN-protéine est connue dans le cas du modèle « hélice-tour-hélice », dont un exemple, l'interaction du répresseur et des opérateurs du phage lambda, a pu bénéficier de l'apport de la cristallographie des complexes. Des courbures de l'ADN interviennent dans la constitution du site de reconnaissance de la protéine et sont provoquées ou amplifiées par la fixation de cette protéine.