

L'ESSENTIEL DE

**BIOLOGIE
MOLÉCULAIRE**

TOUT EN FICHES

L'ESSENTIEL DE

BIOLOGIE MOLÉCULAIRE

Philippe LUCHETTA

Maître de conférences

à l'université de Cergy-Pontoise

DUNOD

Tout le catalogue sur
www.dunod.com



Conception graphique de la couverture : Hokus Pokus Créations

Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.

Le Code de la propriété intellectuelle du 1^{er} juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements

d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour

les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée.

Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du Centre français d'exploitation du

droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).



Les précédentes éditions de cet ouvrage sont parues dans la collection Express.

© Dunod, Paris, 2009, 2013, 2018

ISBN 978-2-10-077866-9

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

Table des matières

Partie 1 – Structure

Fiche 1	Constituants	1
Fiche 2	Structure	6
Fiche 3	Organisation	12
Fiche 4	Compaction	16
Fiche 5	Chromatine	20

Partie 2 – Méthodes

Fiche 6	Purification et quantification	28
Fiche 7	Électrophorèse	33
Fiche 8	Enzymes spécifiques	37
Fiche 9	Carte de restriction	42
Fiche 10	PCR et RT-PCR	47
Fiche 11	Clonage	50
Fiche 12	ADNc	55
Fiche 13	Hybridation moléculaire	59
Fiche 14	Séquençage	65
Fiche 15	Gel retard et <i>footprinting</i>	72

Partie 3 – Fonctions

Fiche 16	Gènes	77
Fiche 17	Gènes procaryotes	83

Fiche 18	Gènes eucaryotes	87
Fiche 19	Transcription	97
Fiche 20	Transcription procaryote	102
Fiche 21	Transcription eucaryote	113
Fiche 22	Maturation des ARNm	124
Fiche 23	Épissage alternatif des ARNm	135
Fiche 24	Surveillance et contrôle qualité des ARNm	143
Fiche 25	Régulation de la transcription	148
Fiche 26	Opéron lactose	156
Fiche 27	Opéron tryptophane	169
Fiche 28	Traduction	175
Fiche 29	Traduction procaryote	184
Fiche 30	Traduction eucaryote	194
Fiche 31	Régulation post-transcriptionnelle	202
Fiche 32	Système Ferritine-Tranferrine	205
Fiche 33	Réplication	211
Fiche 34	Réplication procaryote	216
Fiche 35	Réplication eucaryote	224
Fiche 36	Éléments transposables	228
Fiche 37	Évolution moléculaire	238
Index		247

Les organismes vivants possèdent deux types d'acides nucléiques :

- L'ADN (Acide **D**éoxyriboNucléique).
- L'ARN (Acide **R**iboNucléique).

Ces acides nucléiques sont principalement constitués de bases associées à un sucre.

1. Bases

Chaque acide nucléique contient quatre bases. Trois d'entre elles (adénine, cytosine et guanine) sont communes à l'ADN et l'ARN, la quatrième diffère. On trouve la thymine dans l'ADN et l'uracile dans l'ARN. La différence entre ces deux bases porte uniquement sur le carbone n° 5 (avec ou sans CH_3). On trouve deux catégories de bases (Figure 1.1a et 1.1b) :

- les **purines** constituées de deux cycles aromatiques ;
- les **pyrimidines** constituées d'un seul cycle aromatique.

Les atomes de carbone et d'azote des cycles aromatiques sont numérotés de 1 à 9 (bases puriques) et de 1 à 5 (bases pyrimidiques). Les flèches bleues indiquent la liaison qui se produit entre la base et le sucre.

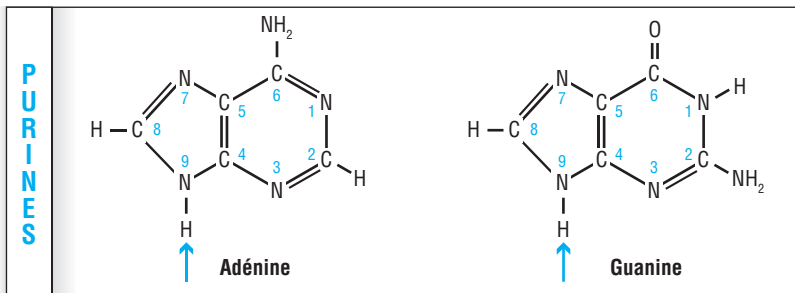


Figure 1.1a

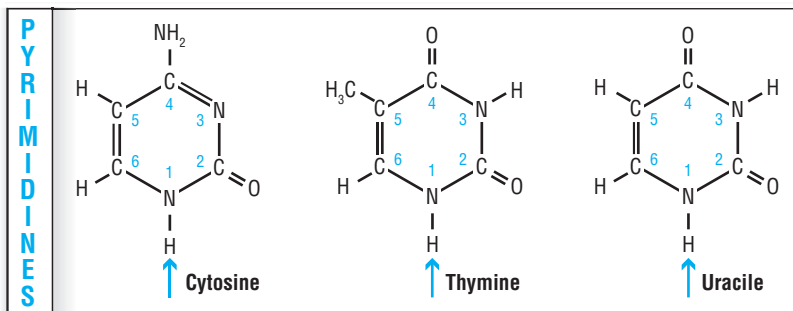


Figure 1.1b

2. Sucre

Le sucre est un pentose composé de 5 carbones numérotés de 1' à 5'. Cette numérotation en « ' » permet de faire la distinction avec celle des bases. L'ARN contient du **ribose** alors que l'ADN contient du **désoxyribose** (Figure 1.2). La différence entre ces deux sucres se situe au niveau du carbone 2'. On trouve un groupement **2'-OH (ARN)** pour le ribose et un **2'-H (ADN)** pour le désoxyribose. Cette différence qui semble peu importante a pourtant une conséquence déterminante sur la stabilité de l'acide nucléique. Ainsi, à pH alcalin (pH 12), l'ARN est hydrolysé alors que l'ADN est stable. Attention à ne pas confondre ici « hydrolyse » et « dénaturation ». On parle de dénaturation lorsque, par exemple à pH alcalin, un ADN double brin se dissocie en deux ADN simples brins par rupture des liaisons hydrogènes. Pour l'ARN, à ce même pH, le mécanisme est différent car il y aura hydrolyse de la molécule par coupure des liaisons phosphodiester entre les carbones 5' et les carbone 3' des nucléotides adjacents.

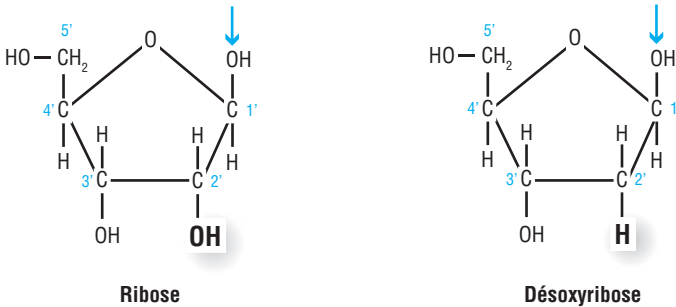


Figure 1.2

3. Nucléosides et Nucléotides

■ Nucléosides

Le carbone 1' du sucre se lie à l'azote de la base (N1 ou N9) pour former un **nucléoside**. Cette liaison est appelée **N-glycosidique**. Un ou plusieurs groupements phosphate (P) peuvent se lier avec le carbone 5' pour former un nucléoside phosphate (Figure 1.3). On établit une nomenclature très précise en fonction de la structure de la molécule. Si le sucre est le ribose (ARN), on a la nomenclature NMP, NDP, NTP en fonction du nombre de groupements phosphate. Si le sucre est un désoxyribose (ADN), on a la même nomenclature précédée d'un « d ». Enfin, il faut connaître la position des groupements P à partir du C5' : 1^{er} P en α , 2^e P en β , 3^e P en γ . Cette nomenclature est importante car, même si elle semble fastidieuse, elle permet de comprendre certaines abréviations de molécules utilisées en biologie. Par exemple, le $\alpha^{32}\text{P}$ -dCTP signifie que le 1^{er} phosphate fixé sur le C en 5' du dCTP est radioactif (^{32}P).

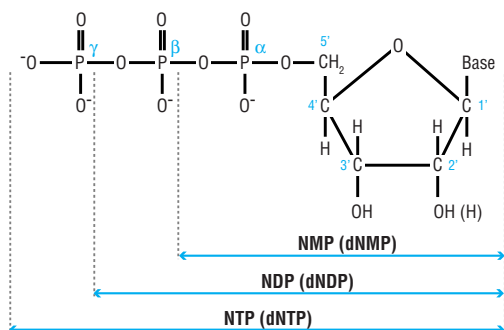


Figure 1.3

(désoxy)nucléoside		(désoxy)ribose	+ Base	
(désoxy)nucléoside monophosphate	(d)NMP	(désoxy)ribose	+ Base	+ 1P
(désoxy)nucléoside diphosphate	(d)NDP	(désoxy)ribose	+ Base	+ 2P
(désoxy)nucléoside triphosphate	(d)NTP	(désoxy)ribose	+ Base	+ 3P

Nucléotides

Le nucléoside monophosphate est aussi appelé **nucléotide** car l'association de plusieurs nucléosides monophosphate entre eux porte le nom de **chaîne polynucléotidique** (Fiche 2). Le nom donné aux nucléosides ou aux nucléotides en fonction de la base présente est différent du nom de la base seule. Par exemple dans l'ATP, l'adénine devient adénosine, dans l'UMP, l'uracile devient uridine, etc.

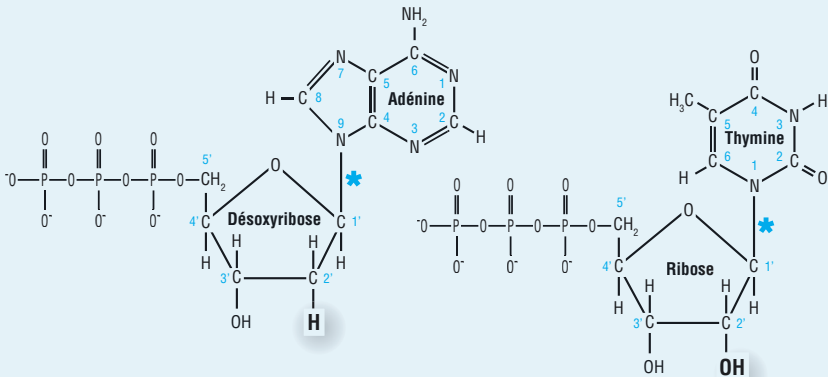
Base	Nom du nucléoside
Adénine (A)	Adénosine
Cytosine (C)	Cytidine
Guanine (G)	Guanosine
Thymine (T)	Thymidine
Uracile (U)	Uridine

EXERCICE 1

1. Écrivez la structure à pH 7 du désoxyribonucléoside triphosphate contenant de l'adénine et du ribonucléoside triphosphate contenant de la thymine. Vous indiquerez sur le schéma la numérotation des carbones du sucre et de la base et la position de la liaison N-glycosidique.
2. Comment appelle-t-on ces deux nucléosides ? Quelle est leur abréviation ?
3. Ces deux molécules sont-elles présentes dans l'ADN ou dans l'ARN ?

Solution

1. Voir la Figure 1.4. La liaison N-glycosidique est indiquée par une étoile bleue.

**Figure 1.4**

2. À gauche, le désoxyribo adénosine triphosphate (dATP). À droite, le ribo thymidine triphosphate (TTP).
3. Le dATP est présent dans l'ADN mais sous forme monophosphate (dAMP) tandis que le TTP n'est présent ni dans l'ADN, ni dans l'ARN. Dans l'ADN, on aurait le dTTP (sous forme dTMP) et dans l'ARN, on aurait l'UTP (sous forme UMP).

1. ARN

■ Chaîne d'ARN simple brin

L'ARN est constitué d'un polymère de nucléotides. Les nucléotides sont liés par une **liaison phosphodiester**. Dans cette liaison, le groupement phosphate entre les deux nucléotides est relié d'un côté par une liaison ester au carbone 5' et de l'autre côté par une seconde liaison ester au carbone 3'. Cette chaîne polynucléotidique est orientée et possède deux extrémités, l'**extrémité 5'P** et l'**extrémité 3'OH** (Figure 2.1)

■ Structures secondaires de l'ARN

Selon les types d'ARN, la chaîne peut être simple brin (ARNm) ou partiellement double brin (ARNr et ARNt). On observe des associations doubles brins si deux séquences sont complémentaires au sein d'une même chaîne d'ARN. Ces associations se font comme dans l'ADN, entre A et U (2 liaisons hydrogène) et entre C et G (3 liaisons hydrogène). Les structures les plus communes sont l'« épingle à cheveux » et la « tige boucle » (Figure 2.2). La différence vient seulement de la taille de la région entre les deux séquences appariées. Il existe d'autres structures secondaires plus complexes comme le « pseudo-nœud ».

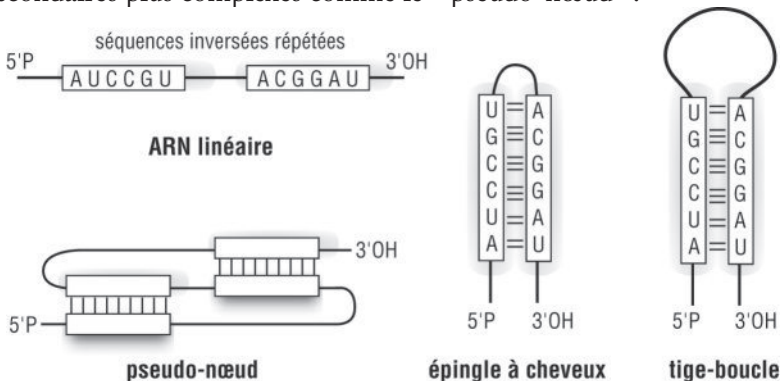


Figure 2.2

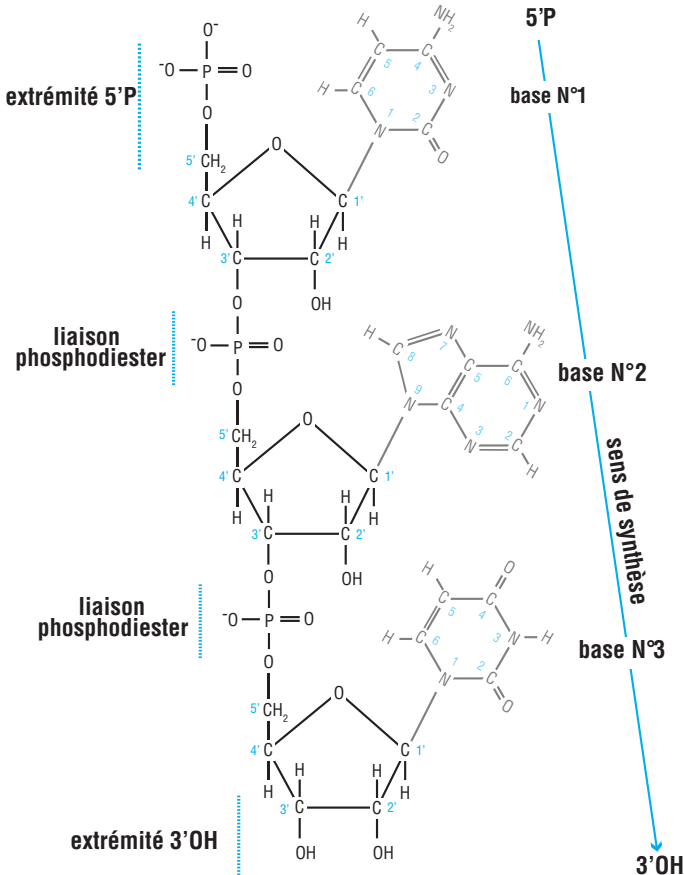


Figure 2.1

2. ADN

On trouve plusieurs structures en double hélice mais la plus fréquente est celle qui a été décrite en 1953 par James Watson et Francis Crick. Cette structure est appelée **hélice Watson/Crick** ou **hélice B**. Elle possède les caractéristiques suivantes.

Les deux brins sont antiparallèles et associés en paire de bases

Deux molécules d'ADN simple brin peuvent s'associer entre elles par complémentarité de leurs bases pour former une molécule d'ADN double brin. Cette structure porte aussi le nom d'**ADN bicaténaire**.

Pour des raisons structurales, les deux brins sont orientés de manière opposée. Ils sont dits **antiparallèles**.

Il y a toujours complémentarité entre une purine et une pyrimidine. Il y a 2 liaisons hydrogène entre A et T et 3 liaisons hydrogène entre C et G (Figure 2.3a et 2.3b, les flèches bleues indiquent la liaison avec le sucre). Cette association a une conséquence directe sur la composition en nucléotides de l'ADN double brin. À la fin des années 1940, le biochimiste **Erwan Chargaff** a analysé la composition en nucléotides de génomes de différents organismes vivants (procaryotes) contenant de l'ADN double brin. Il en a déduit une règle appelée « **règle de Chargaff** » : la quantité d'adénine est identique à celle de thymine ($A = T$) et la quantité de cytosine est identique à celle de guanine ($C = G$), 2) le rapport de purines sur pyrimidines est égal à 1. Cette règle s'applique uniquement pour les molécules d'acides nucléiques double brin et n'est bien sûr valable pour de l'ADN simple brin ou de l'ARN simple brin.

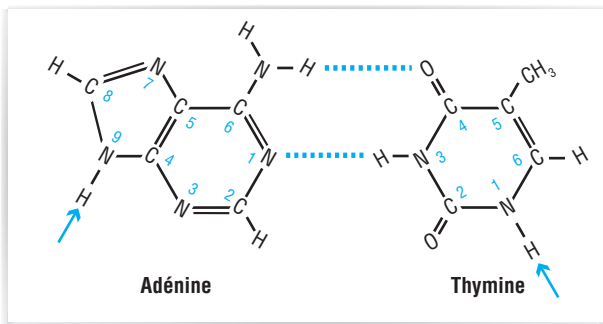
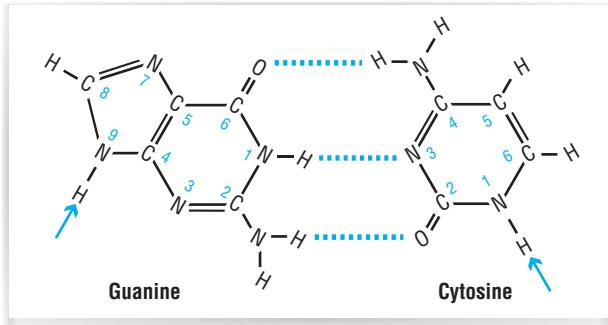
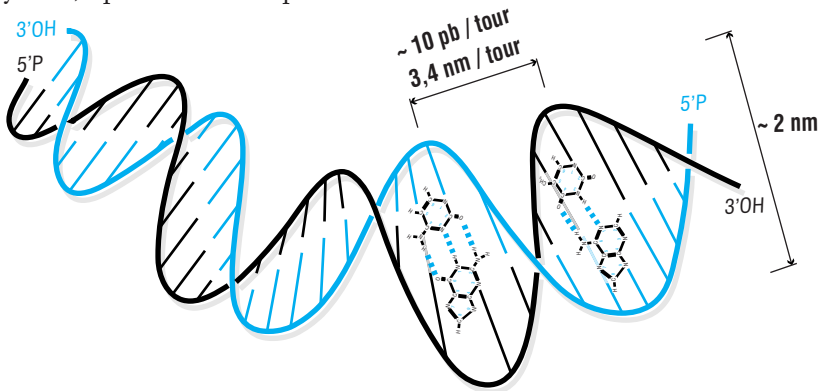


Figure 2.3a

**Figure 2.3b**

■ La double hélice B est « droite »

L'hélice B est dite « droite » car le sens du pas de l'hélice est à droite (Figure 2.4). L'angle de rotation entre deux bases consécutives est de 36° . Le « pas » de l'hélice est de 3,4 nm et son diamètre de 2,37 nm. Il y a 10,4 paires de bases par tour.

**Figure 2.4**

■ Il y a deux sillons différents

Les deux brins de la double hélice d'ADN ne sont pas symétriques par rapport à l'axe central. Sur un plan de coupe, on voit que les deux brins sont plus proches d'un côté que de l'autre (Figure 2.5). La face

où les deux brins sont les plus proches est appelée **petit sillon** (ou **sillon mineur**), l'autre est appelée **grand sillon** (ou **sillon majeur**). Cette particularité a une conséquence importante puisque les bases présentes dans le petit sillon seront masquées et ne pourront pas interagir avec les protéines de liaison à l'ADN comme, par exemple, les protéines de régulation. Les interactions avec ces protéines se feront essentiellement avec les bases accessibles du côté du grand sillon.

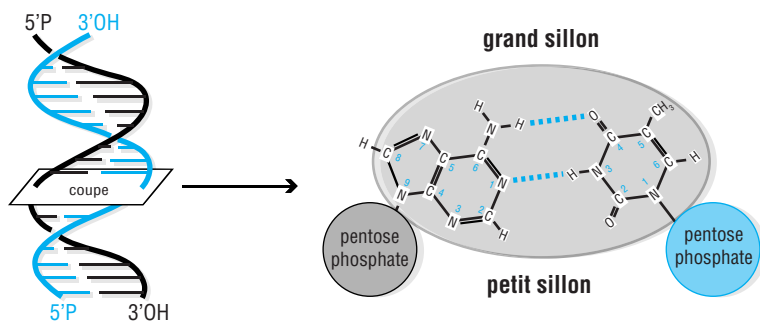


Figure 2.5

■ L'ADN est chargé négativement à pH 7

Les groupements phosphates présents entre deux nucléotides portent des atomes d'oxygène chargés négativement à pH 7. Cela confère une charge globale négative de l'ADN à ce pH.

■ Les autres conformations de la double hélice

L'hélice B que nous venons de décrire est la forme majoritaire trouvée dans les cellules *in vivo*. Deux autres conformations de la double hélice d'ADN sont aussi décrites. L'**hélice A** que l'on trouve dans des solutions où il y a peu d'eau (après extraction) ou *in vivo* lorsque l'ADN est associé à des complexes protéiques. Cette forme A est plus « resserrée » que la forme B avec un diamètre plus important (2,6 nm) et un nombre de paire de bases par tour plus important (environ 11 pb/tour). L'**hélice Z** peut se trouver aussi *in vivo* dans certaines conditions particulières et en très faible proportion. C'est une hélice « gauche » qui apparaît lorsqu'il

Il y a alternance entre bases puriques et pyrimidiques. Cette hélice est plus étirée que l'hélice B avec un diamètre de l'ordre de 1,8 nm et un nombre de paire de bases par tour d'environ 12 pb/tour.

EXERCICE 1

Le pourcentage molaire de la somme Guanine + Cytosine de l'ADN purifié de la bactérie *Pseudomonas fluorescens* est 67 %.

1. Quel est le pourcentage molaire de chacune des 4 bases de l'ADN ?
2. Quel est le rapport bases puriques/bases pyrimidiques ?
3. Quelle indication donne le rapport calculé précédemment sur la structure de l'acide nucléique ?

Solution

1. Les bases G et C étant associées dans l'ADN, il y a autant de chacune de ces deux bases (règle de Chargaff). Il y a donc $67/2 = 33,5$ % de G et 33,5 % de C. Le pourcentage de A + T = $100 - 67 = 33$ %. Sur le même principe, il y a donc $33/2 = 16,5$ % de A et 16,5 % de T.
2. Les bases puriques sont G et A, soit $33,5 + 16,5 = 50$ %. Les bases pyrimidiques sont C et T, soit $33,5 + 16,5 = 50$ %. Le rapport des deux est de 1. C'est aussi un des principes de la règle de Chargaff puisque les purines sont associées aux pyrimidines dans un acide nucléique double brin.
3. Cela montre que l'acide nucléique est double brin. S'il s'agissait d'un ARN ou d'un ADN simple brin, le rapport serait différent de 1.

Tous les organismes vivants possèdent des acides nucléiques comme support d'information génétique. À cela, il faut ajouter les bactériophages et les virus qui, si l'on s'en tient à une définition stricte, ne font pas partie des organismes vivants. Cependant, la notion du « vivant » est complexe et encore très discutée.

1. Domaines du vivant

Aujourd'hui, on distingue trois **domaines** dans le vivant : les **Eubactéries** (Procaryotes), les **Archées** (Procaryotes) et les **Eucaryotes** (Figure 3.1).

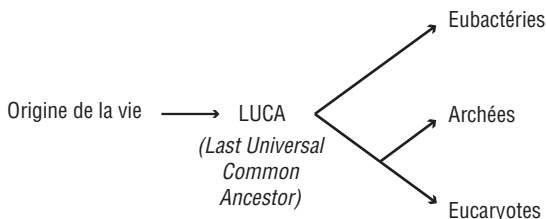


Figure 3.1

Cette nouvelle « classification » qui date du début des années 2000 est parfaitement admise par la communauté scientifique. Certaines notions (le règne) ou certains groupes (protozoaires, invertébrés, etc.) n'existent plus. Dans le domaine des Eubactéries, on trouve toutes les bactéries qui nous sont communes car elles vivent dans notre environnement (par exemple *Escherichia coli*). Dans le domaine des Archées, on trouve des bactéries à métabolismes particuliers (méthanogènes) ou qui vivent dans certains environnements tels que des sources hydrothermales profondes, des geysers, etc. Notons que dans cet ouvrage, toutes les parties faisant références aux Procaryotes correspondent essentiellement aux Eubacteries, les Archées ne sont pas décrites. Dans le domaine très vaste des Eucaryotes, on trouve tous les organismes uni- et pluricellulaires

qui possèdent un noyau cellulaire (en opposition aux Procaryotes qui n'ont pas de noyau).

2. Procaryotes

L'ADN des Procaryotes est **circulaire double brin**. Certaines molécules sont obligatoires et d'autres facultatives. Chaque type d'ADN a une fonction particulière.

	Présence	Fonction	Taille
ADN génomique « chromosome »	Obligatoire	Information génétique de la bactérie	10^5 à 10^7 pb
ADN plasmidique	Facultative	Exemple : résistance aux antibiotiques	10^3 à 10^5 pb
Episome (facteur F)	Facultative	Conjugaison	$\sim 100 \cdot 10^3$ pb

3. Eucaryotes

L'ADN des Eucaryotes est **double brin** mais selon sa localisation il peut être **linéaire** ou **circulaire**.

		Fonction	Structure	Taille
ADN génomique nucléaire « chromosomes »		Information génétique de la cellule ⁽¹⁾	Linéaire	10^7 à 10^{11} pb
ADN extra-chromosomique	ADN mitochondrial	Information génétique de la mitochondrie	Circulaire	$15 \cdot 10^3$ à $30 \cdot 10^3$ pb
	ADN chloroplastique	Information génétique du chloroplaste	Circulaire	$120 \cdot 10^3$ à $220 \cdot 10^3$ pb

⁽¹⁾ L'ADN nucléaire contient les gènes nécessaires à la synthèse des protéines cellulaires mais aussi de nombreux gènes codants des protéines qui seront exportées vers les mitochondries et les chloroplastes.

4. Bactériophages et Virus

Leur matériel génétique est très hétérogène car on peut trouver de l'ADN et de l'ARN. De plus, ces acides nucléiques peuvent être linéaire ou circulaire. Quelques exemples sont regroupés dans le tableau ci-après :

Bactériophages	
øx174 et M13	ADN simple brin circulaire
T4, T7, λ	ADN double brin linéaire
R17	ARN simple brin linéaire

Virus	
herpes	ADN double brin linéaire
SV 40	ADN double brin circulaire
réovirus	ARN double brin linéaire

EXERCICE 1

Deux ADN double brin A et B d'origine différente ont une longueur de 647 paires de bases (pb).

1. Calculez la masse moléculaire (en kDa) de ces ADN ?
2. Donnez la longueur de ces ADN en μm .
3. Sachant que l'ADN A est insensible à l'action des exonucléases et que l'ADN B est sensible aux exonucléases, que pouvez-vous en déduire quant à la structure de ces ADN ? (voir Fiche 8 pour l'action des nucléases).

On considère maintenant un troisième ADN = ADN C. Celui-ci est linéaire, il mesure $0,33 \mu\text{m}$ et a une masse moléculaire de 320 000 Da.

4. Calculez la longueur de cet ADN en pb à partir de la longueur donnée en μm .
5. Même calcul à partir de la masse moléculaire.