

I – Des molécules du vivant à la cellule : organisation fonctionnelle

I – A Organisation fonctionnelle des molécules du vivant

Chapitre 5 : Les nucléotides et les acides nucléiques

Les **nucléotides** sont des molécules organiques composées d'un pentose estérifié par un à trois groupements phosphoryle d'une part et lié à une base azotée d'autre part.

Ce sont les constituants de base des **acides nucléiques**, ADN et ARN.

Quelles sont les fonctions des différents types de nucléotides au sein des cellules ?

Comment se fait leur assemblage sous forme d'acides nucléiques ?

En quoi cette polymérisation permet-elle de faire apparaître des propriétés biologiques supplémentaires ?

Nous étudierons les nucléotides et leurs propriétés en tant que monomères, puis en tant que polymères.

I) Les nucléotides : des molécules solubles de petite taille à rôles informationnels ou énergétiques

1) **Le nucléotide : une base azotée reliée à un pentose phosphorylé**

On appelle **nucléoside** un pentose type **ribose** ou **désoxyribose** relié de manière covalente à une base azotée. Quand le pentose est aussi estérifié par un ou plusieurs groupements phosphoryle, on parle de **nucléotide (ou de nucléoside mono, di ou triphosphate)**.

Les bases azotées sont de deux types :

* les **purines** présentent 2 hétérocycles constitués de 5 atomes de carbone et 4 d'azote ; **l'adénine** et **la guanine** sont des bases puriques les plus fréquentes. Dans les ARN, on retrouve parfois une base particulière, l'inosine.

* les **pyrimidines** présentent un seul hétérocycle constitué de 4 atomes de carbone et 2 d'azote ; **la thymine**, **la cytosine** et **l'uracile** sont des bases pyrimidiques. La thymine est présente seulement dans l'ADN alors que l'uracile est présent seulement dans l'ARN. **Nucléosides = adénosine, guanosine, thymidine, cytidine, uridine.**

Les hétérocycles présentent des doubles liaisons conjuguées qui leur permettent d'absorber dans l'UV (à 250nm) et donc d'être dosés par **spectrophotométrie**.

Le pentose peut être relié à 1, 2 ou 3 **groupements phosphates**. Par exemple, on peut avoir de l'AMP, de l'ADP et de l'ATP. Les nucléotides puriques à 3 groupements phosphate peuvent libérer de l'énergie par hydrolyse (ATP ou plus rarement GTP), on les appelle **des nucléotides activés**.

2) **Les nucléotides peuvent s'associer à des protéines : exemple de la communication cellulaire**

Les nucléotides peuvent servir de **ligands** à d'autres molécules, par exemple les enzymes. Ils interviennent en particulier au cours de la communication cellulaire.

Prenons l'exemple de la voie de transduction du glucagon (voir chapitre 4 sur les protéines), une hormone libérée par le pancréas et ayant un effet hyperglycémiant sur les cellules musculaires et hépatiques :

- la fixation du glucagon sur son récepteur active (transition allostérique) côté intracellulaire une **protéine G (en réalité trimérique – voir deuxième année)**. La forme inactive de la protéine G est liée à du GDP, l'activation (transition allostérique) consiste en un **remplacement du GDP par du GTP**
- la protéine G active une enzyme membranaire (transition allostérique), l'adénylate cyclase, qui produit de **l'AMP cyclique, un second messenger**. L'AMPc se lie à une protéine appelée PKA, l'active (transition allostérique) et entraîne une chaîne de réactions moléculaires aboutissant à la lyse des réserves de glycogène et à la libération de glucose circulant, ce qui fait augmenter la glycémie.
- L'ATP sert aussi de ligand pour certaines enzymes, il régule leur activité (activation ou inhibition).

Toutes ces interactions moléculaires font intervenir des liaisons faibles et des changements de conformation (transitions allostériques) du fait de ces liaisons qui modifient les liaisons faibles entre radicaux au sein des protéines.

Autres exemples : polymérisation de l'actine et ATP, de la tubuline et GTP. Dans ce cas, aussi rôle énergétique

3) Rôle de l'ATP dans le métabolisme : une monnaie énergétique

Calcul de l'enthalpie libre de réaction et spontanéité de la réaction

$$A+B \rightleftharpoons C+D$$
$$\Delta rG = \Delta rG'_0 + RT \ln \frac{[C][D]}{[A][B]}***$$

Pour les chimistes, les conditions de l'état standard d'un système sont : une pression de 1 atmosphère ; une température de 25°C, soit 298 degrés Kelvin ; une concentration des solutés de 1 M ; en conséquence : pH = 0. Noté $\Delta rG'_0$
Les conditions standard pour les biochimistes sont modifiées ainsi : un pH de 7 et donc une concentration $[H^+] = 10^{-7}$ M, une concentration de l'eau (55,5 M) qui est considérée comme constante et dont le terme n'apparaît pas dans l'expression des constantes d'équilibre. Noté $\Delta rG'_0***$

L'ATP est une molécule carrefour du métabolisme énergétique. En effet, son dernier groupe phosphoryle est hydrolysé selon une réaction fortement **exergonique** : $\Delta rG'_0 = -30,5 \text{ kJ/mol}$.

Dans les conditions physiologiques, le rapport $[ADP].[Pi]/[ATP]$ étant d'environ 1/500, il est possible de vérifier que le ΔrG est négatif : $ATP + H_2O \rightarrow ADP + Pi$ $\Delta rG^{0'} = -30.5 \text{ kJ/mol}$ Donc : $\Delta rG = -30500 + 8,32.300.\ln(1/500)$
 $= -46 \text{ kJ/mol}$

Dans les conditions cellulaires, le ΔrG vaut en fait entre -46 kJ/mol et -55 kJ/mol

L'ATP est ainsi un **très bon vecteur d'énergie**, car :

- Son **hydrolyse est favorable** (voir document 4')
- Dans le même temps, il est stable car sa vitesse d'hydrolyse est faible (énergie activation élevée): l'hydrolyse **nécessite une enzyme**, ce qui évite un gaspillage d'ATP cellulaire.
- Il a une **position énergétique « intermédiaire »**

L'ATP est produite dans les organites semi-autonomes, **mitochondrie** et **chloroplaste**.

- Le chloroplaste le produit grâce à l'énergie lumineuse et l'utilise pour ses propres besoins (fixation du CO_2 et fabrication de glucides). **Il ne peut pas sortir du chloroplaste.**
- La mitochondrie produit l'ATP grâce à la respiration cellulaire, c'est-à-dire au **catabolisme oxydatif** des nutriments ; elle libère de l'ATP dans toute la cellule.

L'ATP est utilisé pour de nombreux travaux cellulaires, comme les **synthèses** de l'anabolisme (voir synthèse amidon et glycogène dans chap 3 Les glucides ou synthèse des protéines) et les **déplacements** à l'intérieur de la cellule : contraction musculaire, moteurs moléculaires, fonctionnement des transports actifs primaires etc. Le **GTP** est également hydrolysé lors de certaines réactions, par exemple lors de l'ajout des dimères de tubuline lors de la « polymérisation » des microtubules ou dans la traduction (voir cours biosynthèses).

4) Rôle des coenzymes redox dans le métabolisme énergétique

Au cours des réactions du métabolisme, il y a souvent intervention de molécules appelées **coenzymes** qui aident les enzymes (biocatalyseurs) à assurer leurs fonctions.

Parmi ces coenzymes, on distingue les **coenzymes d'oxydoréduction**, qui sont des **dérivés de nucléotides à adénine** pouvant exister sous deux formes, une forme oxydée (pauvre en électrons) et une forme réduite (riche en électrons). Les couples de coenzymes redox sont $NAD^+/NADH, H^+$; $NADP^+/NADPH, H^+$; $FAD/FADH_2$. Notons R les coenzymes d'oxydo-réduction :

L'oxydation des coenzymes réduites libère de l'énergie : $RH_2 = R + 2H^+ + 2e^- + \text{énergie}$

L'oxydation des coenzymes est donc exergonique.

La réduction des coenzymes oxydées libère de l'énergie. $R + 2H^+ + 2e^- + \text{énergie} = RH_2$

La réduction des coenzymes est donc endergonique.

Ils interviennent dans les réactions d'oxydoréduction au sein des mitochondries et des chloroplastes. Ils permettent des transferts d'électrons mais aussi d'énergie chimique (**pouvoir réducteur**). Voir cours métabolisme.

5) Rôle de la coenzyme A dans le métabolisme énergétique

Le **coenzyme A** est aussi un dérivé de nucléotide à adénine mais elle n'est pas dotée de propriétés redox. Il sert par exemple à importer les acides gras dans la mitochondrie (acide gras lié à coenzyme A = acylCoA aussi appelé acide gras activé). Le **Coenzyme A** est une molécule organique permettant de réaliser le transfert de groupements acétyl $\text{CH}_3\text{-CO-}$ ou acyl $\text{CH}_3\text{-(CH}_2\text{)}_n\text{-CO}$. La rupture de la liaison thioester étant énergétique, elle permet d'assurer la mise en place d'une nouvelle liaison covalente (ou une production d'énergie).

La dégradation des nucléotides, qui sont très riches en azote, libère de l'acide urique. Celui-ci représente environ 5% des déchets azotés chez les Mammifères.

Comment les nucléotides peuvent-ils polymériser pour former des acides nucléiques (ADN, ARN) ?

6) Polymérisation des nucléotides et formation d'une liaison 3'5' phosphodiester

Les liaisons entre nucléotides au sein de l'ADN sont des **condensations** qui s'effectuent par réaction entre l'extrémité 3'OH libre d'un nucléotide et le 1^{er} groupement phosphoryle d'un autre nucléotide (porté par le C5). Il y a donc **deux liaisons ester** de part et d'autre du groupement phosphoryle : on parle de liaison **3'5'phosphodiester**.

Les polymères de nucléotides sont des **macromolécules** appelées acides nucléiques : ADN et ARN.

II) La double hélice d'ADN, support de l'information génétique

Plusieurs expériences permettent d'appréhender la structure de la molécule d'ADN :

- Des expériences de dénaturation-renaturation montre que l'ADN est formé de deux brins associés par des liaisons faibles ; voir cours sur l'organisation du génome.
- Les proportions des différentes bases azotées (rapports de Chargaff) ont conduit Francis Crick à proposer l'existence de paires de bases complémentaires, associées par deux liaisons hydrogènes (AT) ou trois (CG) ;
- L'étude par cristallographie et diffraction aux rayons X a permis à James Watson et Francis Crick (avec Wilkins ; Nobel en 1962) et Rosalind Franklin) d'aboutir en 1953 à la structure complète de la molécule.

1) L'ADN : 2 chaînes de nucléotides complémentaires

Chaque brin est constitué d'une séquence de nucléotides reliés par des **liaisons covalentes**.

Le dosage des quantités respectives des 4 types de bases azotées montre qu'on a toujours exactement le même pourcentage pour A et T d'une part, pour C et G d'autre part (**règle de Chargaff**). On peut expliquer ceci par des appariements spécifiques A-T et G-C par des **liaisons hydrogène** : 2 liaisons pour AT, 3 pour GC. L'ADN est une **macromolécule bicaténaire**, constituée de deux brins (=chaînes) **complémentaires** qui s'enroulent en **double hélice droite autour d'un même axe**.

Les deux brins sont orientés de manière **antiparallèle** : l'extrémité 3' libre d'un brin fait face à l'extrémité 5' de l'autre => *n'oubliez pas d'orienter vos brins dans vos schémas d'ADN.*

L'alternance sucre-phosphate constitue les « montants de l'échelle », les bases, disposées perpendiculairement à ces montants, sont les barreaux de l'échelle. Les bases sont moins hydrophiles que le squelette sucre-phosphate : le fait qu'elles soient moins accessibles à l'eau, au cœur de la molécule, est donc thermodynamiquement stable. La double hélice fait **2nm** de diamètre et il y a **10 paires de bases** par tour d'hélice (un tour = 3,4nm). L'ADN est globalement chargé négativement, des ions Mg^{2+} stabilisent sa structure. On distingue le **petit sillon** (entre les 2 brins complémentaires – 0,6nm) et le **grand sillon** (entre 2 tours de spire – 1,2 nm). Certaines protéines peuvent se fixer à l'ADN et modifier l'expression génétique : on les appelle de **facteurs de transcription**.

La structure décrite par Watson et Crick correspond à la forme B de l'ADN, nettement majoritaire in vivo. Si l'humidité relative passe sous 75% (à peu près jamais dans une cellule...) l'ADN passe sous forme A, plus ramassée avec un diamètre supérieur. Si la position des bases azotées par rapport aux riboses est inversée (passant de anti à syn), on obtient un positionnement plus externe des bases, ce qui conduit à une hélice gauche : l'ADN Z (la composition nucléotidique – poly GC et les conditions physico-chimique favorisent parfois cette forme)

2) Le message génétique est contenu dans la séquence nucléotidique

L'ADN est le **support de l'information génétique** au sein des cellules. La molécule d'ADN peut être libre dans le cytoplasme (procaryotes) ou bien enfermée dans un noyau (eucaryotes) et est aussi présente dans les organites semi-autonomes (mitochondrie, chloroplaste).

Chez les eucaryotes, l'**ADN nucléaire** se trouve sous forme de **chromosomes**. Dans un chromosome, l'ADN est étroitement compacté et associé à des protéines basiques : cet assemblage s'appelle la **chromatine**. En tout, chez l'homme, 1,80m d'ADN tient dans un noyau de quelques microns. Ces 1,80 m sont constitués de **3 milliards de paires** de nucléotides et contiennent un peu moins de **25 000 gènes**.

Les acides nucléiques sont des **hétéropolymères séquencés**, c'est-à-dire formés d'un enchaînement linéaire de monomères différents dont l'ordre est porteur d'une information. C'est l'ordre des nucléotides au sein d'une molécule d'ADN qui détermine l'ordre des acides aminés au sein de la protéine qu'il code.

En première approximation, on appelle **gène** un fragment d'ADN qui peut être transcrit en ARN. Un gène donné occupe un certain emplacement sur un chromosome : cet emplacement est son **locus**.

Certains gènes existent sous forme de plusieurs variantes (qui diffèrent par un ou quelques nucléotides dans la séquence) appelées les **allèles**. Environ 10-15% des gènes existent sous forme d'au moins deux allèles, ce sont des gènes dits **polymorphes**. Au sein d'une espèce, tous les individus possèdent les mêmes gènes mais pas forcément les mêmes allèles. C'est le polymorphisme des allèles qui entraîne les différences de phénotypes observées au sein des populations.

L'expression génétique se fait en deux étapes :

- **la transcription** : dans le noyau, l'ADN est copié en une copie courte et mobile, l'ARNm, qui sort ensuite du noyau ;
- **la traduction** : dans le cytoplasme, l'ARNm est lu par les ribosomes et sert à la synthèse protéique.

Au sein de l'ADN, chaque ensemble de 3 nucléotides constitue un **codon** ou **triplet** et détermine la mise en place d'un acide aminé au sein de la chaîne polypeptidique (**code génétique**). Un codon code le début de la synthèse (codon initiation), trois codons codent l'arrêt de la chaîne (codons stop).

Plusieurs codons codent le même acide aminé : le code génétique est **dégénéré** (= redondant).

De plus, si l'on insère un fragment d'ADN provenant d'une espèce dans le matériel génétique d'une autre espèce, la protéine codée par ce fragment est normalement produite. Ce type d'expérience s'appelle la **transgène** et prouve l'**universalité** du code génétique.

3) Des interactions acides nucléiques-protéines

Certaines protéines peuvent interagir avec l'ADN. Il peut s'agir d'**enzymes**, intervenant dans des synthèses d'acide nucléique ou dans divers phénomènes (réparations de l'ADN, etc.), mais aussi d'**autres protéines intervenant dans la structure du chromosome, la réplication ou la transcription de l'ADN, le contrôle de l'expression de l'information génétique...**

On peut ainsi observer des protéines capables de se lier à l'ADN indépendamment de la séquence de ce dernier. C'est le cas en particulier des protéines histones, associées en permanence à l'ADN des eucaryotes (ce qui forme la structure élémentaire de la chromatine, dans le noyau), grâce à la présence de nombreuses charges positives (acides aminés chargés : lysine, arginine en particulier) interagissant avec les charges négatives du squelette désoxyribose-phosphate.

D'autres protéines peuvent reconnaître une séquence spécifique de l'ADN. C'est le cas en particulier des facteurs de transcription, qui interviennent dans la régulation de l'expression génétique chez les eucaryotes. Ces protéines présentent des structures particulières, leur permettant de se positionner généralement au niveau du grand sillon de l'ADN (plus rarement du petit sillon) et d'établir des liaisons faibles avec les paires de bases.

III) La coopération fonctionnelle des ARN au cours de l'expression génétique

Voir aussi cours sur les biosynthèse

Différents ARN coopèrent pour assurer la synthèse protéique : non seulement les ARN messagers mais aussi les ARN de transfert qui apportent les acides aminés sur le lieu de la traduction et les ARN ribosomiques qui catalysent la formation de la liaison peptidique.

Tous les ARN sont des polymères de nucléotides de 4 types différents : A, G, C et l'**uracile** U qui remplace la thymine. L'ose constituant les nucléotides des ARN est un **ribose** et non un désoxyribose.

1) **Les ARNm, des intermédiaires mobiles entre l'ADN et la synthèse protéique**

Dans le noyau, l'ADN est **transcrit** par l'enzyme **ARN polymérase** : l'un des brins de l'ADN est copié en une copie mobile et de petite taille, à courte durée de vie : l'ARNm. Le brin de l'ADN dont la séquence est identique à celle de l'ARNm est le **brin codant**, l'autre brin (qui sert de matrice à l'ARN polymérase) est le **brin transcrit**. Chaque ARNm est donc la copie fidèle d'un gène. L'export des ARNm est étroitement contrôlé par les **poros nucléaires**. Après excision épissage pour les gènes morcelés des eucaryotes (exon/intron) – voir génétique.

Une fois dans le cytoplasme, les ARNm sont **traduits** en protéines : c'est là qu'interviennent les autres ARN. La durée de vie des ARNm est courte : quelques minutes à quelques heures. Cette durée conditionne la quantité de protéines traduites. Les ARNm sont dégradés par des **nucléases**.

2) **Les ARNt transportent les acides aminés jusqu'au lieu de la traduction**

Les ARNt sont constitués d'environ 80 nucléotides, ils ont une forme dépliée (structure secondaire) en **feuille de trèfle**. Les branches du trèfle tiennent ainsi par liaisons hydrogènes établies entre bases complémentaires, qui stabilisent l'ARNt. L'extrémité **3'OH** libre est capable de fixer un **acide aminé** spécifique. La liaison entre AA et extrémité 3'OH se fait dans le cytosol, **indépendamment de la traduction**, par une enzyme spécifique : **l'aminoacyl ARNt transférase**.

La boucle centrale du trèfle porte un ensemble de 3 nucléotides appelé **anticodon** et capable de reconnaître un codon d'un ARNm. Il y a une **complémentarité** codon-anticodon.

L'AA fixé par l'extrémité 3'OH dépend donc de la nature de l'**anticodon** porté par la boucle centrale et ce grâce aux aminoacyl ARNt transférase qui reconnaissent simultanément les structures 3D de l'ARNt et de l'acide aminé. **Les aminoacyl ARNt transférase sont donc des acteurs clés dans la réalisation du code génétique.**

Un ARNt utilisé pour la traduction porte **déjà** un acide aminé qui est lié à son extrémité 3' par liaison covalente (on dit que cet ARNt est **chargé**).

3) **Les ARNr catalysent la formation de la liaison peptidique**

Les ribosomes sont constitués de deux sous-unités : la grosse sous-unité contient 34 protéines et deux ARNr, la petite sous-unité contient 21 protéines et un ARNr (chez les Procaryotes). Tant qu'ils ne traduisent pas d'ARNm, les sous-unités sont dissociées et libres dans le cytoplasme. Au total, les ribosomes des procaryotes ont une taille de 25nm de diamètre et un poids moléculaire de 2700 kDa. Ils contiennent 3 cavités dans lesquelles vont venir se loger les ARNt : **A (aminoacyl)**, **P (peptidyl)** et **E(exit)**. Ces sous-unités s'assemblent au moment de commencer la traduction d'un ARNm.

Voyons comment se déroule la synthèse d'une chaîne polypeptidique (traduction). Celle-ci a lieu de 5' en 3'.

- dans le site P se trouve un ARNt qui porte la chaîne en croissance, de longueur n, en face du codon n
- dans le site A, en face du codon n+1 arrive un ARNt chargé, porteur d'un nouvel acide aminé
- l'ARN ribosomique catalyse la formation de la liaison peptidique entre le nouvel AA (n+1) et la chaîne en croissance : cette chaîne est en quelque sorte « transférée » sur l'AA n+1
- le ribosome se décale d'un codon vers l'extrémité 3' : l'ARNt maintenant porteur de la chaîne est dans le site P et ainsi de suite.

Les acides aminés sont ajoutés de l'extrémité **N** vers l'extrémité **C**.

Remarque : Il existe une grande variété d'ARN intracellulaires comme les ARNmi (micro ARN) d'une vingtaine de nucléotides qui permettent la répression de l'expression de gènes, les ARNsn (sn : small nuclear) qui interviennent dans l'excision épissage – voir cours de génétique. Il existe aussi des ARN double brins (ex : virus)...

BILAN

Les nucléotides sont des monomères pouvant véhiculer de l'énergie chimique (comme l'ATP ou les coenzymes d'oxydoréduction) ou intervenir dans la communication cellulaire, notamment en interagissant avec des protéines (AMPC).

Ils polymérisent grâce à des enzymes appelées polymérases pour former les acides nucléiques.

Les acides nucléiques sont des macromolécules séquencées : l'ordre des nucléotides au sein de ces molécules est porteur de l'information génétique. L'ADN est un long polymère stable, responsable de la conservation de l'information génétique. Les ARN sont des polymères courts, mobiles et éphémères qui participent à la synthèse des protéines à partir de l'information génétique portée par l'ADN.

La complémentarité des bases azotées a de nombreuses conséquences :

- Elle permet l'association des deux brins dans un acide nucléique double brin ;
- Elle permet un repliement d'un brin, en formant des motifs en « épingle à cheveux » comme pour les ARNt intervenant dans leur structure 3D (et cette structure 3D est primordiale pour leur rôle – voir cours biosynthèses) ;
- Elle permet l'association de deux acides nucléiques : par exemple, elle permet le positionnement d'un AA-ARNt sur trois nucléotides (= triplet, ou codon) d'un ARNm ;
- Elle permet la synthèse d'un brin par complémentarité de séquence d'un brin matrice : synthèse d'ADN lors de la réplication, synthèse d'ARN lors de la transcription.

Le fait que l'ADN soit constitué de deux brins complémentaires est donc notamment essentiel à sa réplication et la complémentarité des bases est aussi essentielle à l'expression de l'information génétique.

*L'ARN ribosomique est le seul exemple de catalyseur biologique (= d'enzyme) qui ne soit pas une protéine mais un acide nucléique. On pense qu'il y a 3 milliards d'années, les premières cellules fonctionnaient en utilisant l'ARN comme support de leur information génétique (**monde à ARN**). Cet ARN aurait à la fois servi de « code » pour la synthèse des protéines mais aussi catalysé lui-même les liaisons peptidiques.*

D'autres indices étaiet cette théorie : le désoxyribose est obtenu enzymatiquement à partir du ribose et la thymine peut être obtenue enzymatiquement à partir des cytosines. De plus, certains virus utilisent l'ARN comme support de leur patrimoine génétique (virus à ARN).