

INTRODUCTION A LA BIOLOGIE MOLECULAIRE

Michel Morange (ENS) : « La biologie moléculaire est l'ensemble des techniques et découvertes qui ont permis l'analyse moléculaire des processus les plus intimes du vivant, de ceux qui en assurent la pérennité et la reproduction. »

Biologie moléculaire = discipline scientifique au croisement de la génétique, de la biochimie et de la biophysique, dont l'objet est de comprendre les mécanismes de fonctionnement de la cellule au niveau moléculaire. Elle s'intéresse tout particulièrement aux **molécules supportant le message héréditaire** (acides nucléiques, ADN, ARN) et à la façon dont ces molécules commandent la fabrication de protéines (lien génotype / phénotype).

Historique : L'ADN est le support universel de l'information génétique

Bien que les acides nucléiques aient été isolés dès 1869 par **Frédéric MIESCHER** (isolement d'une substance riche en phosphate baptisée nucléine), on a longtemps ignoré le rôle qu'ils pouvaient jouer dans le maintien et la transmission de l'information génétique. En réalité, compte tenu de la relative simplicité de leur composition par rapport à celle des protéines (4 bases pour les uns, 20 acides aminés pour les autres), on préférait leur assigner un rôle structural dans la cellule. Ainsi, même après la publication des travaux d'**AVERY** qui vont être décrits plus bas, le modèle dominant qui prévaudra jusqu'au début des années 1950 sera celui de **Phoebus LEVENE** proposé en 1933 et pour qui l'ADN était la répétition d'un motif élémentaire formé par l'enchaînement des 4 bases A, T, G et C. L'ADN était donc considéré comme une molécule « monotone » dont la fonction était inconnue mais dont on savait qu'elle était associée aux chromosomes. Dans un tel modèle nucléoprotéique, l'ADN constituait le support sur lequel se fixaient les protéines, véritables responsables de la spécificité de l'information génétique.

Cependant, en 1928, **Frederick GRIFFITH** fait une observation capitale en étudiant la pathogénicité de certaines formes de pneumocoques. Le pneumocoque *Diplococcus pneumoniae* existe sous plusieurs formes différant par l'aspect de leurs colonies. Le type III correspondant à la forme S (pour *smooth*, lisse en anglais) est une forme virulente qui s'associe avec les cellules qu'il infecte grâce à un antigène O contenu dans la capsule polysaccharidique qui l'entoure et provoque des infections foudroyantes chez la souris. Le type II correspond à la forme R (pour *rough*, rugueux en anglais), qui en raison d'un déficit enzymatique, ne produit pas d'antigène O et est, de ce fait, non pathogène. Après avoir injecté des souris avec un mélange contenant des bactéries R vivantes et des bactéries S tuées par la chaleur, **GRIFFITH** constate que les souris meurent de pneumonie. Fait plus troublant encore, il retrouve des bactéries S dans le sang des souris mortes, et démontre que cette transformation de la forme R en formes S est permanente et peut être transmise à la descendance. **GRIFFITH** émet l'hypothèse de l'existence d'un **principe transformant**, sans pouvoir en préciser la nature.

En 1931, **Martin DAWSON** reproduit les observations de **GRIFFITH** hors de l'animal par co-culture préalable *in vitro*. L'année suivante, **Lionel HALLOWAY** établit que l'on peut transformer des pneumocoques non virulents de type II *in vitro* par des extraits de pneumocoques virulents de type III préalablement tués. Cette expérience ouvre la voie de la purification du facteur transformant responsable de la transformation d'un type de pneumocoque non virulent en type virulent.

Dès 1935, ces expériences sont reprises par **Oswald AVERY**, **Colin MacLEOD** et **Maclyn McCARTY**. Ces travaux vont durer presque 10 ans et en 1944, ils démontrent que le facteur transformant est l'ADN. Il a les caractéristiques physico-chimiques de l'ADN, est insensible à l'action des protéases et des ribonucléases et est détruit par l'action d'une DNase. En outre, **AVERY** démontre que le facteur ne réagit pas avec des anticorps dirigés contre la capsule des pneumocoques de type III. Le facteur responsable de la formation de la capsule est donc d'une nature chimique différente de celle de la structure cellulaire qu'il spécifie. La prudence des conclusions d'**AVERY** dans son article ne reflète pas sa conviction profonde que le facteur transformant est un gène pur, mais associée aux objections avancées par les partisans de la nature protéique des gènes, elle constitua le principal obstacle à la reconnaissance de sa découverte.

Ces travaux, bien qu'on ne sache pas trop comment les interpréter, vont influencer les recherches de nombreux scientifiques.

En 1943, **Joachim HÄMMERLING** étudie les acétabulaires, des algues unicellulaires de grande taille, porteuses de chapeaux pouvant être crénelés (*Acetabularia crenulata*) ou lisses (*Acetabularia mediterranea*). Il observe que si l'on coupe les chapeaux et que l'on permute les noyaux de ces deux espèces, l'algue qui a hérité du noyau de *A. crenulata* reforme un chapeau crénelé et que celle qui a hérité du noyau de *A. mediterranea* reforme un chapeau lisse. Cela montre que l'information nécessaire à l'établissement du phénotype est contenue dans le noyau cellulaire.

En 1945, **Erwin CHARGAFF**, intrigué par les observations d'**AVERY** et leur incompatibilité avec le modèle du tétranucléotide de **LEVENE**, reprend les mesures effectuées par ce dernier. En 1951, il établit que dans l'ADN, les proportions respectives des bases sont variables selon les espèces, ce qui remet en cause le **modèle de LEVENE**. La finesse des analyses qu'il effectue lui permet de montrer que dans tous les cas, les proportions des différents nucléotides sont telles que A=T et G=C. Cette règle qui porte son nom (**règle de CHARGAFF**) joua un rôle déterminant dans la découverte de la structure en double hélice de l'ADN.

En 1953, **James WATSON** et **Francis CRICK** vont déterminer cette structure en double hélice de l'ADN. Exploitant les diagrammes de diffraction aux rayons X établis par l'équipe de cristallographes de **Maurice WIKINS**, et plus particulièrement ceux de **Rosalind FRANKLIN** à Cambridge, ils proposèrent un certain nombre de modèles de doubles ou triples hélices dans lesquels les bases étaient situées à l'extérieur de la molécule avant de se tourner vers des modèles inverses dans lesquels les bases s'appariaient avec elles-mêmes. En tentant d'apparier une base purique avec une base pyrimidique, ils constatèrent que les paires A=T et G=C avaient la même structure spatiale, ce qui permettait de construire une hélice parfaitement régulière et d'expliquer les résultats de **CHARGAFF**. Ils obtiendront le prix Nobel pour ces travaux en 1962, prix qu'ils partageront avec **Maurice WILKINS**. **Rosalind FRANKLIN**, décédée peu de temps auparavant, ne pourra bénéficier du prix...

Entre temps, **Alfred HERSHEY** et **Martha CHASE** vont redémontrer, par une approche différente de celle employée par **AVERY**, le rôle de support de l'information génétique de l'ADN. En 1952, ils infectent des bactéries *Escherichia coli* par le bactériophage T2 en présence de ^{32}P et de ^{35}S dans le milieu de culture, deux radio-isotopes qui vont s'incorporer respectivement dans l'ADN et les protéines du bactériophage. Les particules de phage marquées au ^{32}P et au ^{35}S sont utilisées pour infecter des bactéries cultivées dans un milieu exempt de toute radioactivité. Par une série d'expériences, **HERSHEY** et **CHASE** démontrent que seul le ^{32}P se retrouve à l'intérieur des bactéries et dès lors, que l'ADN du phage est nécessaire et suffisant pour que de nouvelles particules virales soient produites. Contrairement à ce qui s'était passé pour **AVERY**, et malgré la faiblesse de certaines expériences, ces travaux sont vite acceptés par la communauté scientifique... Les travaux d'**AVERY** obtenus 8 années

auparavant avaient fini par pénétrer les esprits, et combinés à d'autres observations telles que celles de **CHARGAFF**, ils rendaient plausibles un rôle génétique pour l'ADN.

Plus récemment, avec la naissance du génie génétique, les scientifiques ont été capables d'introduire de l'ADN d'une espèce dans une autre (expériences de transgénèse ; ex. la GFP produite par une méduse utilisée comme outil de laboratoire pour des marquages cellulaires ou le suivi spécifique d'une protéine dans différents types de cellules de différentes espèces). Cela démontre que la cellule receveuse est capable de lire correctement cet ADN étranger et de produire les protéines qu'il code. Le code génétique, c'est-à-dire la correspondance entre la séquence de nucléotides de l'ADN et la séquence en acides aminés de la protéine, est le même chez tous les êtres vivants : on dit qu'il est universel. Par extension, on dit que l'ADN est le support universel de l'information génétique.

L'ADN est donc le support moléculaire universel de l'IG. Comment s'organise cette molécule ? En quoi sa structure lui permet-elle de réaliser sa fonction biologique ?