

## III. Techniques du génie génétique (Séquençage d'ADN)

Le séquençage d'ADN est devenu un outil essentiel en biologie moléculaire tant en médecine que dans de nombreuses autres disciplines des sciences de la vie. Le séquençage a été décrit il y a environ 30 ans et n'a cessé d'évoluer depuis cette période.

Cette méthode est devenue une technique courante dans les laboratoires de biologie moléculaire. Les connaissances acquises grâce à cette méthode et la possibilité de séquencer des génomes de grande taille, tel que le génome humain, ont amené les chercheurs à développer des techniques de séquençage de plus en plus sophistiquées.

### III.1 Historique du séquençage

Les deux premières techniques de séquençage de l'ADN, celle de **Maxam-Gilbert** et celle de **Sanger** ont été décrites en 1977. Le premier organisme qui a été séquencé, il s'agissait du virus bactériophage  $\phi$ X174, possédant un ADN simple brin ne nécessitant donc pas l'étape de dénaturation utilisée dans les méthodes de Sanger et Maxam et Gilbert.

La technique de Sanger a révolutionné le monde de la biologie moléculaire en permettant de décrypter différents génomes, tel que celui du génome humain complètement déchiffré en 2006 ou d'autres génomes, le génome bactérien, par exemple, le premier d'entre eux étant celui d'*Haemophilus influenzae*, complètement décrit en 1995.

### III.2 Définition de séquençage

Le séquençage de l'ADN constitue une méthode dont le but est de déterminer la succession linéaire des bases A, C, G et T prenant part à la structure de l'ADN. La lecture de cette séquence permet d'étudier l'information biologique contenue par celle-ci.

Étant donné l'unicité et la spécificité de la structure de l'ADN chez chaque individu, la séquence de l'ADN permet de nombreuses applications dans le domaine de la médecine, comme, par exemple,

- le diagnostic,
- les études génétiques,
- l'étude de paternité,
- la criminologie,
- la compréhension de mécanismes physiopathologiques,
- la synthèse de médicaments,
- les enquêtes épidémiologiques.

Bien que les techniques de séquençage évoluent, la méthode de Sanger continue d'être la méthode la plus employée dans le monde à l'heure actuelle.

### III.3 Les différents acteurs du séquençage

- ↳ l'ADN provient des organismes dont on souhaite séquencer le génome. L'ADN, le plus souvent sous la forme double-brin, est dénaturé afin de séparer les deux brins. Celui qui sera séquencé s'appelle le brin « matrice »  
Les **nucléotides**, ou plutôt désoxynucléotides, sont les bases de l'ADN (A, C, G ou T). Ils sont attachés les uns aux autres grâce à des liaisons chimiques nécessitant la présence d'un groupement OH particulier (sur le carbone 3' du ribose)
- ↳ les **didésoxynucléotides** sont des nucléotides privés du groupement OH en position 3'. Leur incorporation dans une chaîne d'ADN interrompt définitivement la synthèse de l'ADN (voir figure 1).
- ↳ une **amorce** est un brin d'ADN très court (une vingtaine de nucléotide), qui peut s'hybrider à une séquence complémentaire spécifique
- ↳ l'ADN **polymérase** est une enzyme qui a pour rôle de copier l'ADN, en synthétisant un brin complémentaire au brin matrice. L'ADN polymérase ne fonctionne qu'en ajoutant des nucléotides à

## PREMIERE PARTIE : METHODES D'ETUDE DE LA MORPHOLOGIE DES CELLULES

une amorce déjà présente, en fonction de la succession de nucléotides du brin matrice (un A est toujours positionné en face d'un T et un C toujours en face d'un G).

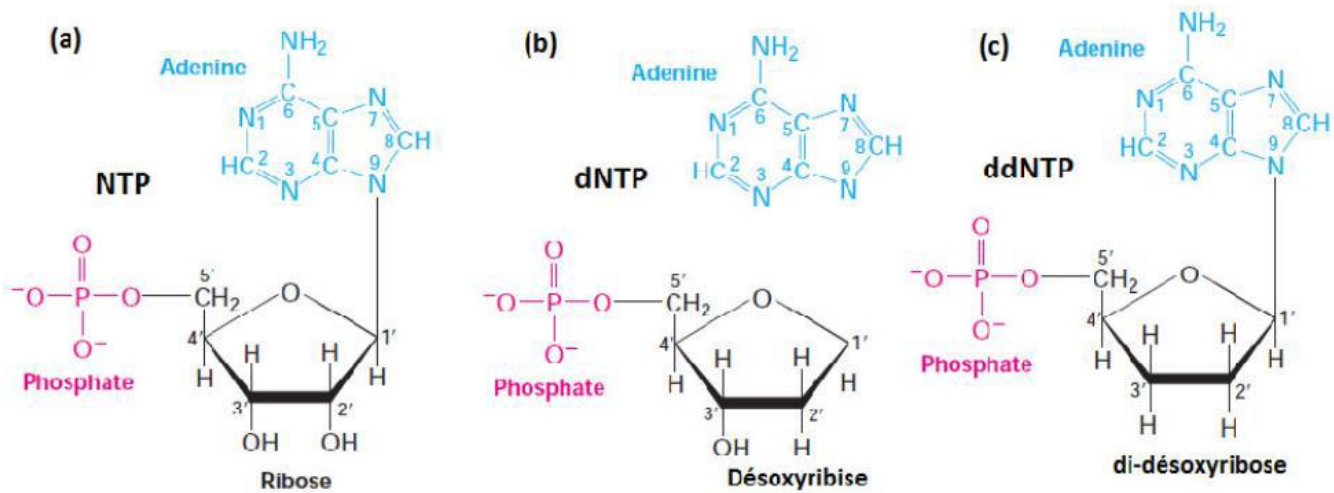


Figure 01 : les différentes formes des nucléotides.

### III.4 La technique de Sanger

La méthode de Sanger a en effet rapidement dépassé la méthode de Maxam-Gilbert pour la remplacer et reste à ce jour la principale méthode de séquençage utilisée dans les laboratoires. Son principe est le suivant ;

Dans un premier temps, il est nécessaire d'amplifier l'ADN cible par PCR, puis de le dénaturer afin d'obtenir un ADN simple brin. À l'aide d'une amorce spécifique et complémentaire du brin étudié (sens ou antisens), identique ou différente de celle utilisée pour la PCR, une ADN polymérase effectue alors la synthèse de l'ADN complémentaire à partir de cette amorce. De l'extrémité 5' vers l'extrémité 3', cette enzyme ajoute les désoxyribonucléotides-triphosphates (dNTP) complémentaires et de manière aléatoire et inconstante des didéoxyribonucléotides triphosphates (ddNTP), par exemple un ddGTP sera parfois ajouté à la place d'un dGTP.

La réaction se faisant dans un seul tube, les ddNTP (ddATP, ddGTP, ddCTP et ddTTP) sont marqués à l'aide de fluorophores différents pour chaque ddNTP (fluorophores « quatre couleurs »).

Lorsqu'un ddNTP est incorporé à la place d'un dNTP, l'ADN polymérase ne peut plus continuer sa polymérisation. La réaction d'extension s'arrête (en effet, le didéoxynucléotide ne possède pas de groupe 3'-hydroxyle indispensable à la réaction de polymérisation de l'enzyme). Statistiquement, au cours de la réaction, pour chaque « base » de l'ADN cible, au moins une fois, un ddNTP complémentaire sera incorporé à la place d'un dNTP. Par conséquent, à la fin de la réaction, nous obtiendrons des fragments de taille différente. L'analyse de la réaction est ensuite effectuée.

Par exemple, dans la réaction où on a ajouté du ddGTP, la synthèse s'arrête au niveau des G. Le mélange réactionnel contenant, à la fois du dGTP et un peu de ddGTP, la terminaison se fait de manière statistique suivant que l'ADN polymérase utilise l'un ou l'autre de ces nucléotides. Il en résulte un mélange de fragments d'ADN de tailles croissantes, qui se terminent tous au niveau d'un des G dans la séquence. Ces fragments sont ensuite séparés par électrophorèse sur un gel de polyacrylamide, ce qui permet ainsi de repérer la position des G dans la séquence.

*Différentes méthodes d'analyse sont possibles ;*

Aujourd'hui, l'électrophorèse capillaire réalisée sur un automate de séquençage est la méthode de choix. Lors de la migration, chaque fragment (contenant un ddNTP marqué par un fluorophore) sera excité par un

## PREMIERE PARTIE : METHODES D'ETUDE DE LA MORPHOLOGIE DES CELLULES

laser et le signal obtenu analysé par un logiciel spécifique. L'analyse informatique des signaux permet d'obtenir la séquence étudiée, par exemple, sous forme d'un électrophorégramme, de lecture manuelle aisée mais souvent fastidieuse (Figure 02).

Des logiciels d'analyse des séquences peuvent être utilisés. Dans tous les cas, l'analyse d'un fragment d'ADN après PCR se fait toujours à l'aide d'une amorce sens et antisens afin de confirmer la séquence (et une éventuelle anomalie de séquence). En général, cette technique permet d'obtenir des séquences de longueur comprise entre 400 et 850 pb.

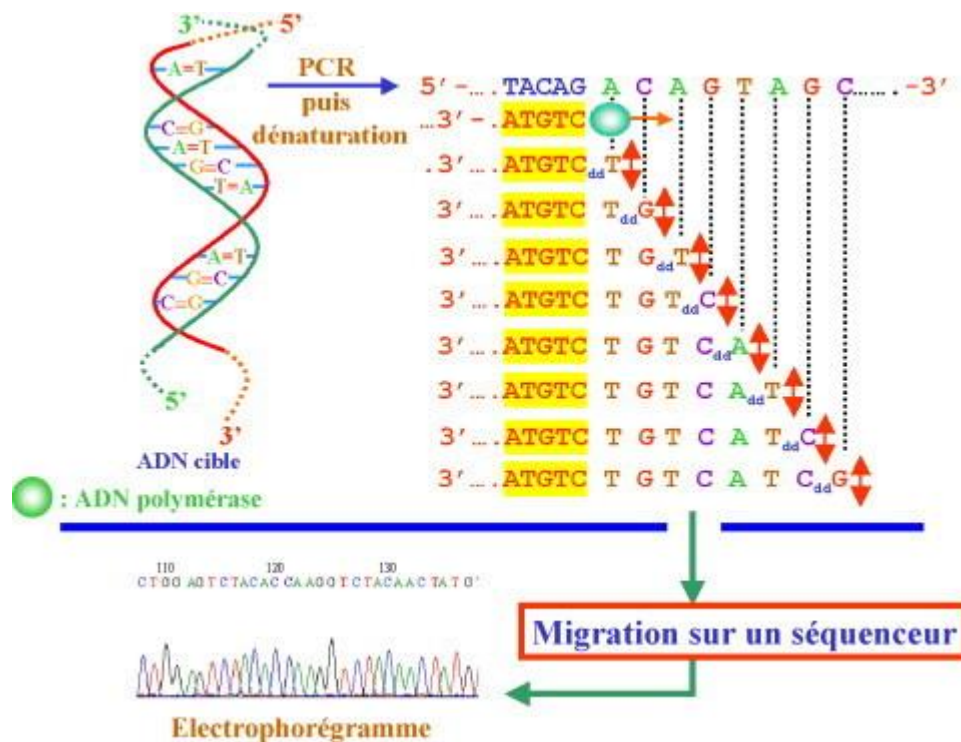


Figure 01 : Principe du séquençage d'ADN par la méthode enzymatique de Sanger

### Pour la simplicité du schéma ;

Après dénaturation du produit amplifié par séquençage, l'un des deux brins (ici, le brin sens) s'hybride à une amorce spécifique. Nous avons pris une amorce de 5 pb, la taille habituelle des amorces étant de 20 pb environ. Le mélange réactionnel contient, outre les tampons et l'ADN polymérase, des déoxynucléotides triphosphates (dNTP, dA-, dC-, dG-, dT-TP) mais aussi des didéoxynucléotides triphosphates (ddNTP, ddA-, ddC-, ddG-, ddT-TP).

L'incorporation aléatoire d'un ddNTP à la place d'un dNTP ne permet plus la polymérisation par l'ADN polymérase. L'extension s'arrête. À la fin de la réaction de séquence effectuée selon cycles thermiques identiques à ceux de la PCR (on parle de PCR asymétrique, une seule amorce étant utilisée au lieu de deux), nous avons des fragments de taille différente. Ces fragments sont soumis à migration dans un champ électrique. Il s'agit le plus souvent d'une électrophorèse capillaire.

Chaque ddNTP étant marqué par un **fluorophore** différent, un signal lumineux sera généré, spécifique de la base didéoxy incorporée. Les fragments étant de taille différente et la résolution allant jusqu'à une base de différence, il sera simple de recueillir ce signal et en déduire la séquence. Les signaux lumineux sont analysés par un logiciel spécifique, et le résultat de l'analyse peut être lu, par exemple, sous forme d'un électrophorégramme de lecture facile.

Des logiciels d'interprétation des séquences sont également disponibles. Pour confirmer un résultat, toute réaction de séquence d'un fragment d'ADN est systématiquement faite sur le brin sens et le brin antisens.

**Références bibliographiques :**

1. Lamoril, J., Ameziane, N., Deybach, J. C., Bouizegarène, P., & Bogard, M. (2008). Les techniques de séquençage de l'ADN : une révolution en marche. Première partie [DNA sequencing technologies: A revolution in motion. Part one]. *Immuno-analyse & Biologie Spécialisée*, 23(5), 260–279. <https://doi.org/10.1016/j.immbio.2008.07.016>
2. Gilbert W., Maxam A. The nucleotide sequence of the lac operator. *Proc Natl Acad Sci U S A*. 1973;70:3581–3584.
3. Sanger F., Nicklen S. Coulson AR DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. 1977;74:5463–5467.
4. Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*. 1995;269:496–512
5. Hall N. Advanced sequencing technologies and their wider impact in microbiology. *J Exp Biol*. 2007;210:1518–1525.