

# Extraction d'ADN, PCR et séquençage de l'ADN

## I. Extraction d'ADN

### 1. Introduction

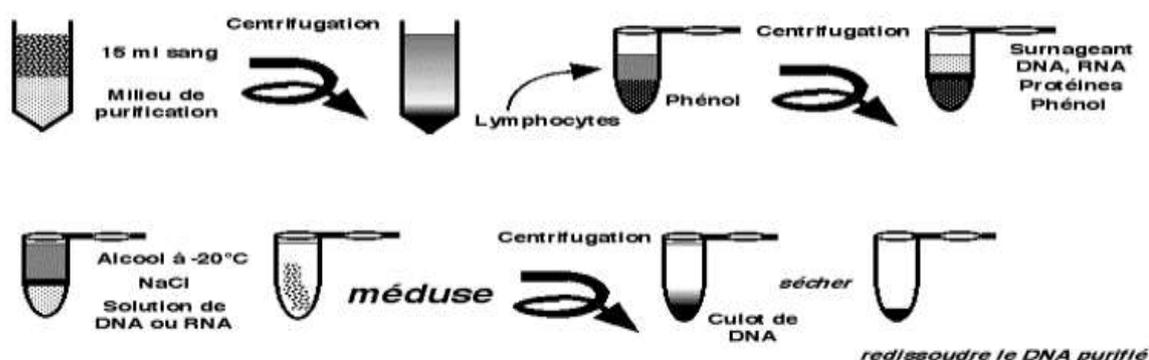
Toute étude de génétique moléculaire implique la disposition d'échantillon d'acides nucléiques. Les techniques d'extraction d'acides nucléiques sont relativement simples. Il convient simplement d'éviter toute destruction enzymatique ou mécanique. En effet les acides nucléiques qui sont stables dans la cellule intacte, deviennent très vulnérables à la digestion par les nucléases endogènes une fois la cellule lysée.

### 2. Préparation de l'ADN à partir du sang total

Les globules blancs (lymphocytes, macrophages...etc.) représentent la source majeure d'ADN en médecine. A partir de 10 à 30 ml de sang en présence d'un anticoagulant (ex. EDTA), il est possible de récupérer une quantité d'ADN ( $\approx 100\mu\text{g}$ ) suffisante, pour des études qualitatives et quantitatives.

### 3. Protocole expérimental

#### **Extraction et purification du DNA**



### 4. Protocole d'extraction d'ADN

1. Prélèvement de 10 à 30 ml de sang dans un tube contenant un anticoagulant (ex. EDTA).
2. Eclatement des globules rouges par une solution hypotonique
3. Récupération des globules blancs par centrifugation.
4. Préparation du lysat cellulaire en utilisant la solution de lyse (détergent comme le SDS ou sarcosyl + Protéinase K). Cette étape permet la libération de l'ADN nucléaire dans le milieu et la digestion des protéines qui lui étaient associées.
5. Extraction phénolique : cette étape permet de récupérer l'ADN dans la phase aqueuse après la centrifugation, car le phénol est un déprotéinisant puissant dans lequel les acides nucléiques ne sont pas solubles.

6. Extraction au chloroforme ou à l'éther : cette étape complète toujours une extraction phénolique, car elle permet d'éliminer les traces de phénol qui aurait pu être importé par la phase aqueuse (le phénol peut inhiber les enzymes utilisées dans l'amplification, la restriction...etc.).
7. Précipitation : cette étape permet la précipitation de l'ADN pour faciliter sa récupération. Il y'a deux types largement utilisés :
- a - Précipitation à l'alcool éthylique : Elle doit être effectuée à haute force ionique (2.5 volumes d'éthanol 95° contre un volume d'échantillon). Pour les faibles concentrations, il faut prolonger le temps de précipitation (> 10h), comme il est possible d'accélérer la précipitation par le froid (-20°C à -70°C). L'ADN est récupéré par centrifugation.
  - b - Précipitation à l'isopropanol : Le principe est le même que la précipitation éthanolique. Ce type de précipitation se fait volume à volume. Mais deux caractéristiques majeures la différencient de la précipitation éthanolique.
    - Le sel n'est pas nécessaire
    - Les très petits fragments de DNA même à hautes concentrations ne sont pas précipités, ce qui permet de les éliminer.
8. Lavage : l'utilisation de l'éthanol 70° permet d'éliminer les sels, et les traces de l'isopropanol. L'ADN est récupéré par centrifugation.
9. Séchage: cette étape permet l'élimination de l'éthanol (s'évapore).
10. Resuspendre dans 10mM tris EDTA, l'eau purifiée de nucléases...etc.

## **5. Dosage des acides nucléiques**

Les pyrimidines et les purines absorbent fortement les UV à 260 nm. Une unité de densité optique à 260 nm correspond à:

- Une solution de DNA double brin à 50µg/ml
- Une solution de DNA simple brin ou RNA à 25µg/ml

$$C = A_{260} * DF * 100$$

Ces valeurs s'appliquent à des acides nucléiques parfaitement purs et en solution homogène.

Pour vérifier la pureté de l'ADN il faut calculer :

$$P (\text{pureté}) = A_{260}/A_{280}$$

Une solution d'ADN est considérée pure si :

$$1.7 \leq P \leq 2$$

## **II. Réaction de Polymérisation en Chaîne "PCR"**

La synthèse et le séquençage d'ADN ont permis l'émergence d'une méthode d'amplification de l'ADN appelée PCR (Polymerase Chain Reaction), à partir d'un gène (ou fragment) spécifique, de grande quantité d'ADN sont obtenues *in vitro*. *L'ADN polymérase copiant jusqu'à un milliard de fois cette région cible « quantité suffisante pour être révélée ».*

## **1. Historique**

Cette méthode de Biologie Moléculaire a été mise au point en 1985 par Kary Mullis, qui obtint pour ces travaux le prix Nobel de Chimie en 1993.

Aujourd'hui, ce procédé révolutionnaire couplé à l'utilisation d'une ADN polymérase thermorésistante permet d'obtenir, sans clonage, une amplification considérable d'un fragment donné d'ADN.

## **2. Etapes de PCR**

La dénaturation de la double hélice nécessite une haute température (autour de 95°C), donc l'enzyme utilisée doit être résistante aux hautes températures ou bien elle devait être ajoutée lors de chaque cycle. Ce problème a été résolu en utilisant une ADN polymérase thermiquement stable d'une bactérie thermophile : *Thermus aquaticus*, isolée d'une source chaude. Cette enzyme nommée ADN Taq polymérase est stable à 95°C (tab.1) et n'est pas affectée par l'étape de dénaturation. Avec l'utilisation du Taq polymérase, l'ADN étant copié à 72°C et non plus à 37°C, son utilisation augmente la spécificité de la PCR. Car aux hautes températures l'hybridation non spécifique d'amorces à l'ADN non cible est rare. Les produits de l'ADN Taq polymérase sont donc plus homogènes que ceux obtenus avec l'ADN polymérase III d'*E. coli*. En résumé il y'a quatre étapes :

1 - Dénaturation (autour de 95°C) : Sert à dénaturer l'ADN pour obtenir des matrices simple brin.

2 - Hybridation (autour de Th): Lors du refroidissement du mélange, les amorces étant en excès, la plupart des brins d'ADN cibles se fixent à celles et non entre eux.

3 - Extension (72°C) : L'ADN polymérase « prolonge les amorces : côté 3'-OH libre » en utilisant les brins cibles comme matrice.

4 - Après une période appropriée d'incubation, le mélange est chauffé de nouveau pour séparer les brins. Il est alors refroidi pour que les amorces s'hybrident aux régions complémentaires de l'ADN nouvellement synthétisé. Le processus complet est alors répété « cycle ».

	ADN Taq polymérase	ADN polymérase III
Stabilité thermique à 95°C	+++	-
Activité exonucléase	-	+ (5' → 3')

**Tab.1: Les différences entre ADN Taq Polymérase et l'ADN polymérase III**

## **3 .Thermocycleur**

Les trois étapes, constituant un cycle de PCR, sont effectuées à des températures différentes permettant de contrôler l'activité enzymatique. Pour effectuer ces transitions de températures, les microtubes contenant le mélange réactionnel sont placés dans un appareil programmable : un thermocycleur (fig. 1).

Cet appareil permet d'exposer les tubes à des températures choisies et pour des durées déterminées par l'expérimentateur (effectuer automatiquement les cycles de chauffage) (fig. 2).

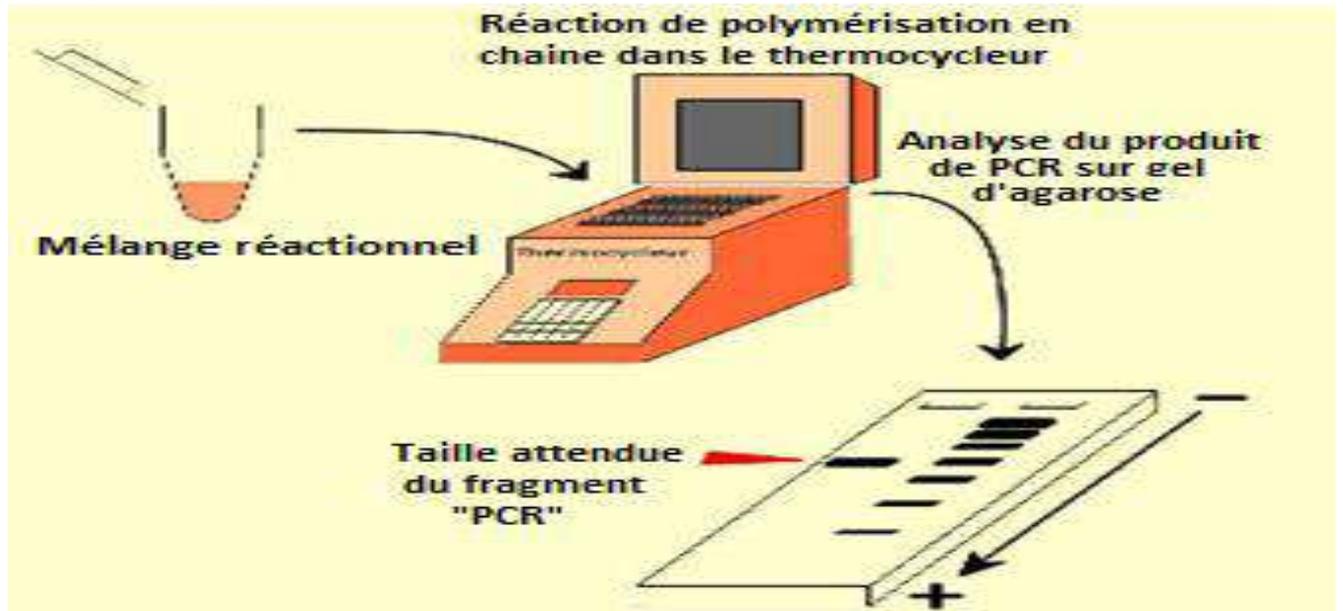
Le secret de la PCR est que le produit d'un cycle d'extension sert de matrice pour le suivant (fig. 2), ainsi la technique PCR est remarquable par le fait que l'ADN cible se double à

chaque cycle. En pratique 20 à 40 cycles sont habituellement réalisés, ils génèrent 106 à 109 fois la séquence cible. « Elle repose sur la succession de plusieurs cycles » (fig. 3).

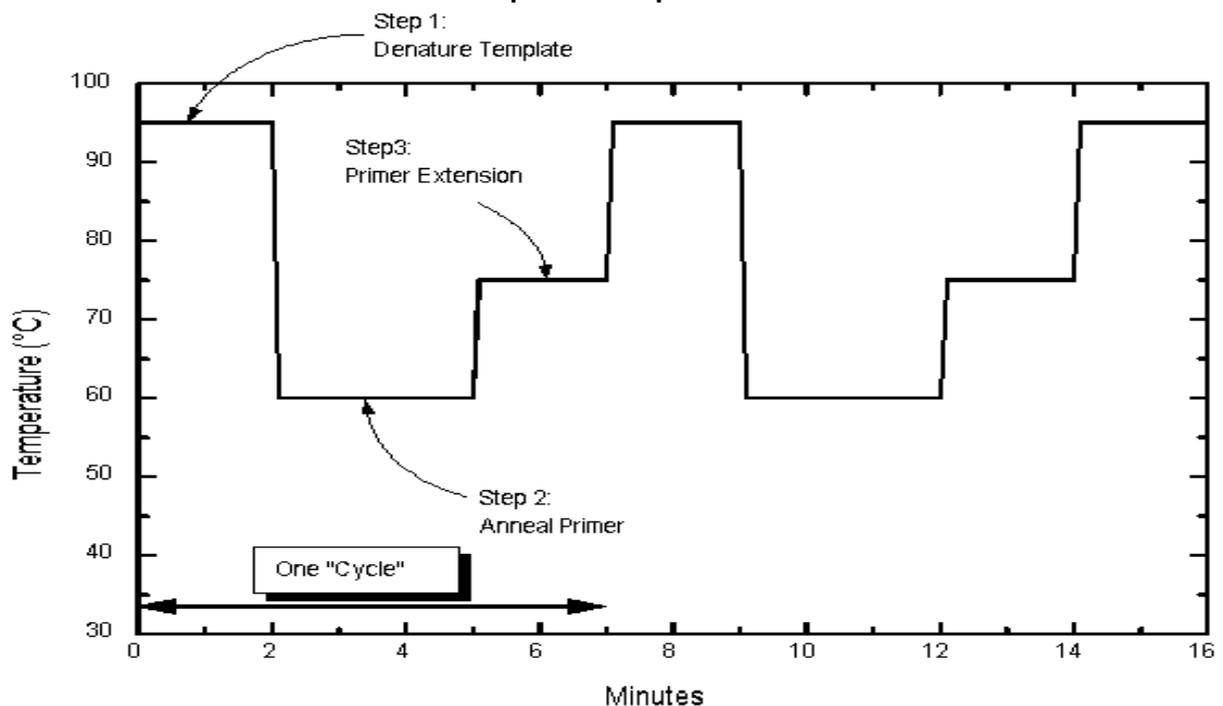
La progression géométrique de raison 2 (tab. 2) permet d'établir une relation pour calculer le nombre des copies totale et cibles.

Cycle	1	2	3	4	5	6	-----n	n : nombre de cycle
Copies totale (N)	2	4	8	16	32	64	-----	$N = 2^n$
Copies parasites	2	4	6	8	10	12	-----	2n
Copies cibles	0	0	2	8	22	52	-----	$2^n - 2n$

**Tab. 2:** Progression géométrique de raison 2 des cycles de PCR et nombre de copies.

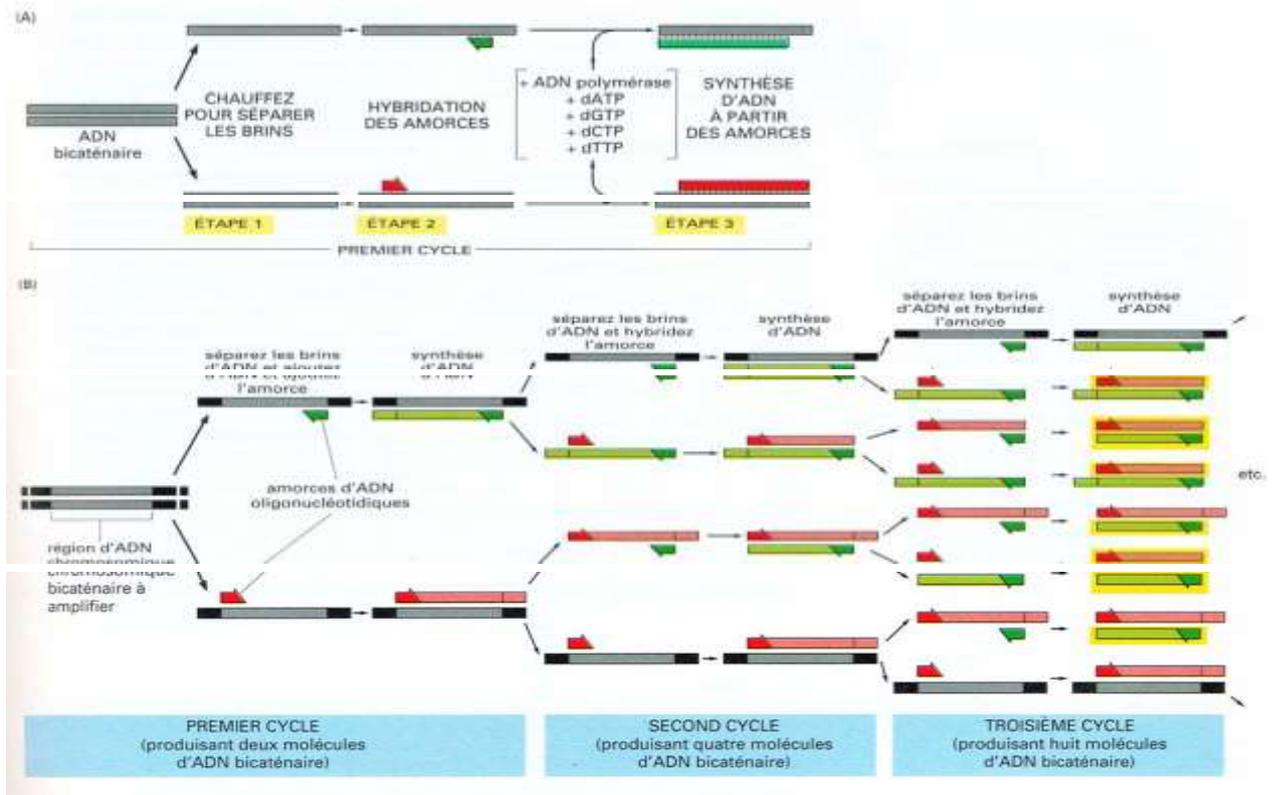


**Fig. 1:** Réaction d'amplification d'ADN dans le thermocycleur et visualisation du produit de PCR par électrophorèse



**Fig. 2:** Profile de cycles de la température de la PCR

#### 4. Aspects quantitatifs de la réaction PCR



**Fig. 3 :** Amplification d'ADN par PCR. Le produit d'un cycle sert de matrice pour le cycle qui suit.

#### 5. Le désigne des amorces

Le désigne des amorces (primers) repose sur les critères suivants:

- Taille entre 17 à 31 nucléotides
- Contenu GC :  $\approx 50\%$
- Th proches entre les deux amorces (F et R).
- Absence de répétition d'un même nucléotide
- Absence de complémentarité entre les deux amorces (pas de risque de formation de dimères)
- Faible risque de formation des structures secondaires (Ex. Epingle à cheveux).

Pour calculer le Th d'une amorce on utilise la formule suivante:

$$Th = 4(C+G) + 2(A+T) - 5$$

#### 6. Limites et Applications

##### Limite :

- Taille : on ne peut amplifier que des petits fragments (taille inférieure à 2 kb)
- Connaissance d'une partie de l'ADN à amplifier (pour l'amorce)
- ADN pol thermorésistante est dépourvue de la fonction proof-reading d'où un risque d'erreur.
- Amplification parasite : mauvaise hybridation et contamination.

### **Applications :**

- *Génération* de sondes
- Recherche de mutations
- RT-PCR : étude des ARN
- Clonage
- PCR quantitative
- PCR en temps réel : quantification précise.

## **III. Séquençage**

Le séquençage de l'ADN, consiste à déterminer l'ordre d'enchaînement des nucléotides d'un fragment d'ADN donné. Le séquençage d'un ADN, c'est-à-dire la détermination de la succession des nucléotides le composant.

### **1. Historique**

Les premières techniques de séquençage ont, été développées en parallèle au milieu des années 1970. Les méthodes de Sanger (Grande-Bretagne) et Gilbert (Etats-Unis) ont toutes deux été récompensées d'un prix Nobel de chimie en 1980.

Le premier organisme a été séquencé en 1977. Il s'agissait du virus bactériophage  $\Phi$ X174, possédant un ADN simple brin ne nécessitant donc pas l'étape de dénaturation utilisé dans les méthodes de Sanger et Maxam et Gilbert.

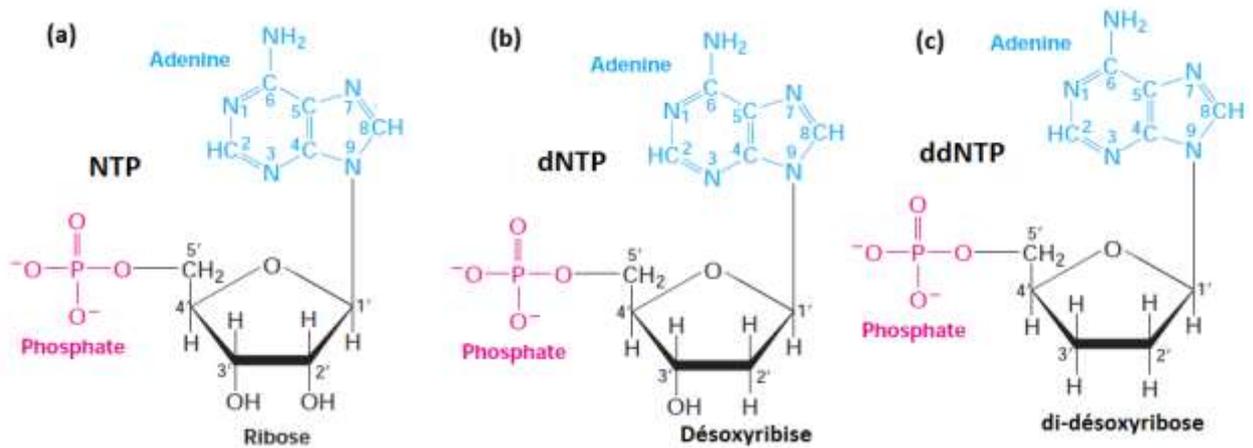
L'apparition des séquenceurs automatiques a notamment permis l'automatisation de ces technologies et ont contribué à la finalisation du génotypage humain en 2003.

### **2. Méthode de Sanger**

Le principe de cette méthode consiste à initier la polymérisation de l'ADN à l'aide d'un petit oligonucléotide (amorce) complémentaire à une partie du fragment d'ADN à séquencer.

L'élongation de l'amorce est réalisée par le fragment de Klenow (une ADN polymérase I dépourvue d'activité exonucléase 5'→3') et maintenue par des ADN polymérases thermostables, celles qui sont utilisées pour la PCR. Les quatre désoxyribonucléotides (dATP, dCTP, dGTP, dTTP) sont ajoutés, ainsi qu'une faible concentration de l'un des quatre didésoxyribonucléotides (ddATP, ddCTP, ddGTP ou ddTTP).

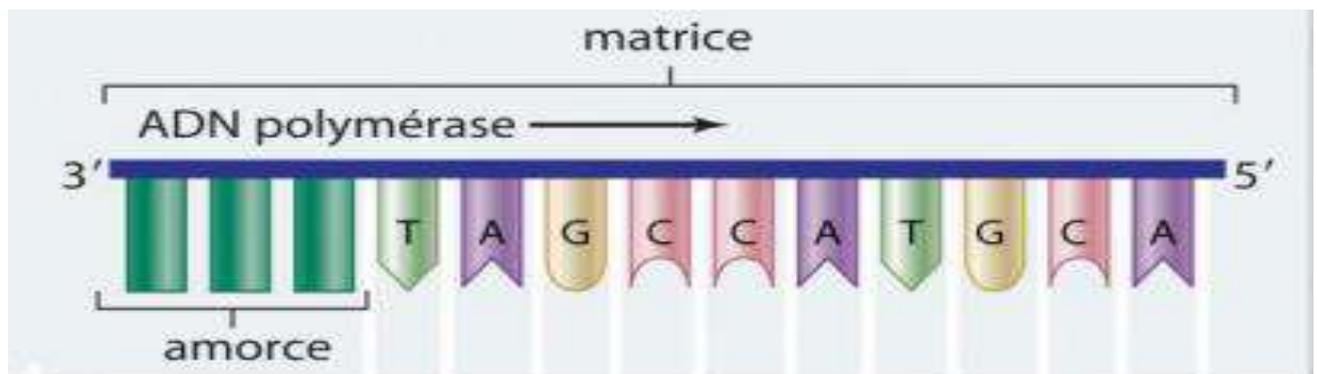
Ces didésoxyribonucléotides agissent comme des «poisons» terminateurs de chaîne: une fois incorporés dans le nouveau brin synthétisé, ils empêchent la poursuite de l'élongation. Cette terminaison se fait spécifiquement au niveau des nucléotides correspondant au didésoxyribonucléotide incorporé dans la réaction. Pour le séquençage complet d'un même fragment d'ADN, on répète cette réaction quatre fois en parallèle, avec les quatre didésoxyribonucléotides différents.

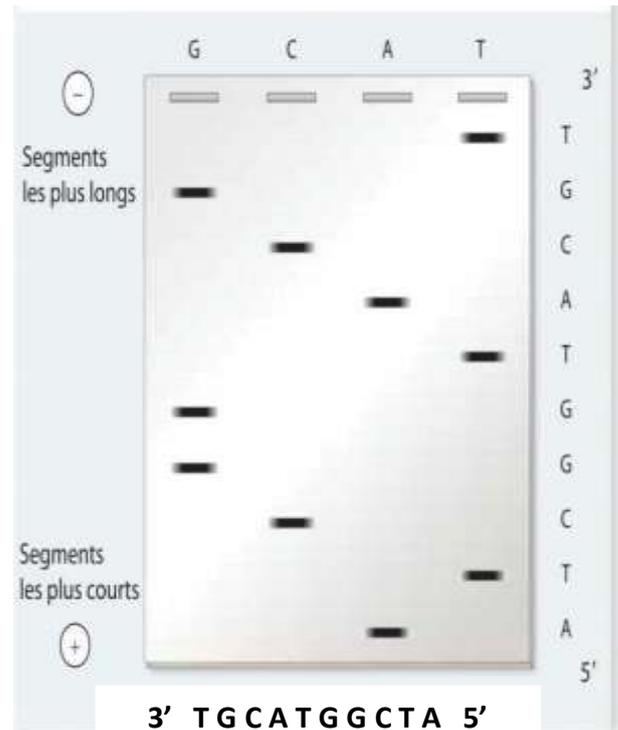
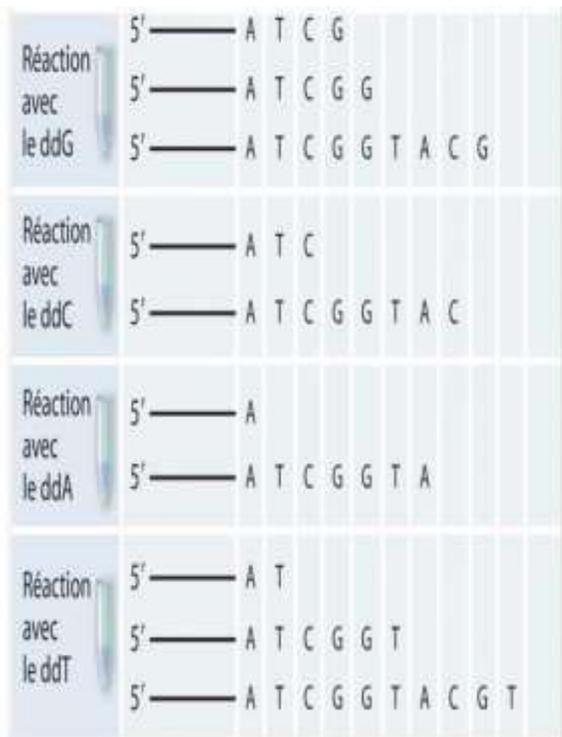


**Fig. 4: Les différentes formes des nucléotides**

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 6, ce qui permet ainsi de repérer la position des G dans la séquence. La détection des fragments ainsi synthétisés se fait en incorporant un traceur dans l'ADN synthétisé. Initialement ce traceur était radioactif; aujourd'hui, on utilise des traceurs fluorescents, attachés soit à l'oligonucléotide, soit au didésoxyribonucléotide.

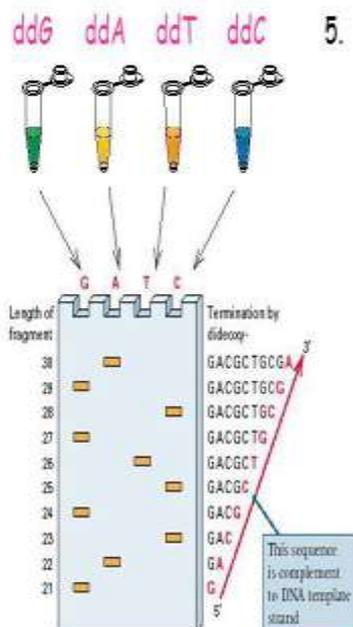
Pour commencer, il vous faut un brin d'ADN à séquencer. Ensuite, vous ajoutez une séquence d'amorce, les quatre nucléotides et un enzyme appelé ADN polymérase qui incorpore de nouvelles bases de nucléotide, faisant un nouveau brin d'ADN conforme à l'original. Dans la méthode originale de Sanger, quatre différentes réactions de séquençage sont effectuées. Chaque réaction comprend un nucléotide modifié différent qui, une fois incorporé, constitue la fin d'une chaîne d'ADN, ce qui permet d'identifier la base finale. Ces échantillons sont alors soumis à l'électrophorèse en gel, méthode qui permet de séparer les nouveaux brins d'ADN sur une base en gel à l'aide de courant électrique. Les brins d'ADN peuvent alors être vus à l'aide de rayons X ou de lumière ultraviolette. Pour lire le gel, vous commencez par le bas et regardez les bandes (tirets noirs) afin de déterminer le séquençage du fragment d'ADN.



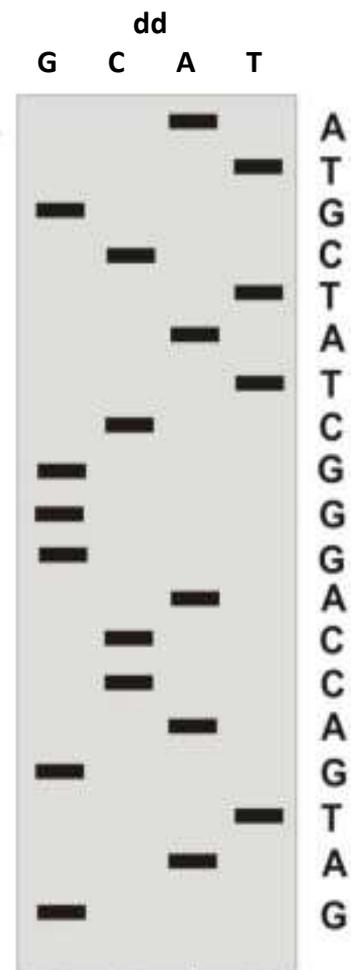


3' T G C A T G G C T A 5'  
Le brin complémentaire

## Méthode de Sanger



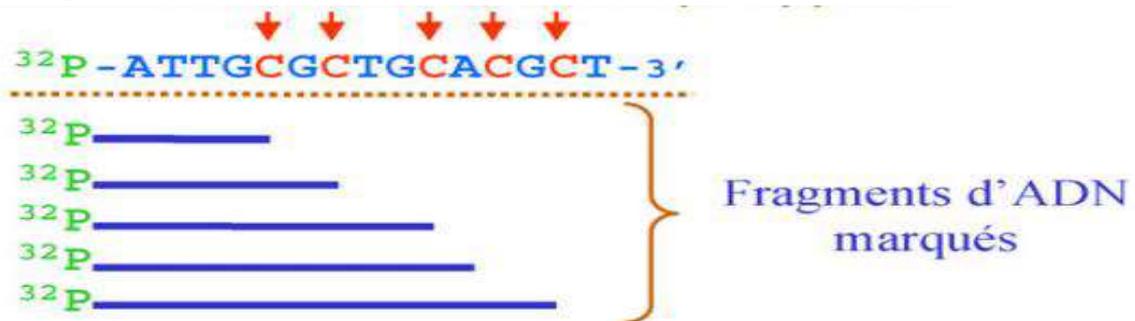
5. Préparation du gel du polyacrylamide  
Dénaturation et dépôt sur gel  
Migration par électrophorèse



### 3. Méthode de Maxam et Gilbert

Il s'agit, d'une méthode chimique de séquençage, est basée sur une dégradation chimique de l'ADN et utilise les réactivités différentes des quatre bases G, C, [A+G] et [C+T] pour réaliser des coupures sélectives. Cette technique permettait d'analyser des fragments allant jusqu'à 500 pb. En reconstituant l'ordre des coupures, on peut remonter à la séquence des nucléotides de l'ADN correspondant. On peut décomposer ce séquençage chimique en six étapes successives:

- **Marquage** : Les extrémités des deux brins d'ADN à séquencer sont marquées par un traceur radioactif  $^{32}\text{P}$ . Cette réaction se fait en général au moyen d'ATP radioactif et de polynucléotide kinase.
- **Coupure** : les ADNs sont soumis à des réactions chimiques spécifiques des différents types de base, l'ADN est clivé au niveau de la modification par réaction avec une base, par exemple, une réaction pour les G (alkylation par le sulfate de diméthyle), une réaction pour les G et les A (dépurination), une réaction pour les C, ainsi qu'une réaction pour les C et les T (hydrolyse alcaline).
- Le produit de séquence est déposé sur un gel d'acrylamide, puis la séquence lue après autoradiographie. Cette analyse est analogue à celle que l'on effectue pour la méthode de Sanger.



G	A+G	C+T	C	Séquence	3'
—	—	—	—	C	↑
	—			G	
	—	—		A	
		—		T	
		—		T	
		—		T	
		—	—	C	
—	—			G	
—	—	—		G	
	—	—	—	A	
	—		—	T	
	—			C	
	—			A	
	—			A	

Séquence: 5' -AACTAGGCTTTAGC-3'

#### **4. Automatisation du séquençage**

Aujourd'hui, la plupart des séquençages sont réalisés par des séquenceurs industriels entièrement automatisés. Ceux-ci utilisent la technique de Sanger mais avec des méthodes de révélation différentes.

Les fragments d'ADN sont marqués par des marqueurs fluorescents; leur taille est ensuite déterminée par chromatographie ou électrophorèse assistée par ordinateur.

Avec ces techniques, on peut séquencer jusqu'à 1000 bases avec les meilleurs séquenceurs contre 200 à 300 via une méthode manuelle comme celles exposées ci-dessus. En effet, lors de l'électrophorèse manuelle, le nombre de bases est limité afin de ne pas rendre le chromatogramme illisible par la surcharge des bandes et ainsi ne plus permettre une lecture horizontale.

C'est notamment grâce à la rapidité de ces appareils que le séquençage du génome humain fut réalisé en un temps record par rapport aux prévisions effectuées lors du démarrage du projet.

#### **5. Performances & Limitations**

##### **Performances des séquenceurs Sanger modernes**

- ◆ Plusieurs centaines d'échantillons simultanément et un séquençage par heure.
- ◆ Séquences de longueur 300-1000 nucléotides max

##### **Limitations**

- ◆ Si amplification (PCR) avant séquençage: des parties de la séquence du vecteur d'amplification sont retrouvées dans le séquençage Sanger.
- ◆ Erreurs en début de séquence: reconnaissance imparfaite de l'amorce
- ◆ Faible pouvoir de résolution entre séquences qui ne diffèrent en longueur que d'un nucléotide