



Cours

Ingénierie des protéines à visée thérapeutique

Niveau M2 Biochimie



Dr. GUENDOUZE A.
2020-2021

CHAPITRE I :
DÉTERMINATION DES STRUCTURES DES PROTÉINES

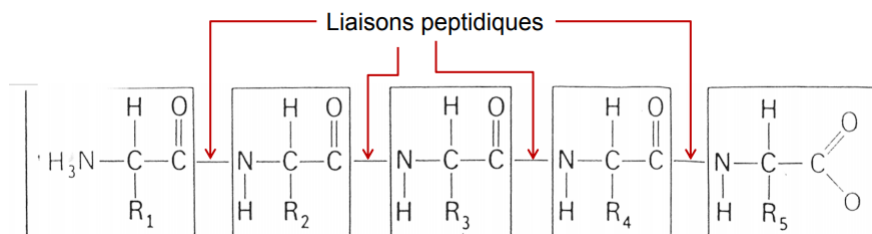
Le structure des protéines

- L'ordre de la séquence des acides aminés distingue une protéine d'une autre.
- La séquence en acides aminés détermine la forme tridimensionnelle de la protéine.
- Une modification de la séquence en acides aminés dans une protéine change sa structure tridimensionnelle.
- La différence entre un polypeptide et une protéine est que le terme polypeptide réfère uniquement à l'enchaînement d'acides aminés alors que le terme protéine s'applique à une chaîne d'acides aminés après son repliement correct et dans certain cas sa modification.
- rôles dans la cellule: catalyse de réactions, intégrité structural, transport de molécules, mouvement, fixation de molécules, ...

1

Structure primaire des peptides et protéines

- C'est la séquence en acides aminés
- Déterminée par les gènes
- Elle commence du côté N-terminale



2

Structure primaire des peptides et protéines

→ Les chaînes non polaires.

Ala (A), Val (V), Leu (L), Ile (I), Met (M), Phe (F), Trp (W), Pro (P)

→ Les chaînes polaires non ioniques

Ser (S), Thr (T), Cys (C), Tyr (Y), Asn (N), Gln (Q)

Rôle clé pour la
structure et la réactivité

→ Les chaînes polaires ioniques

anioniques : Asp (D), Glu (E)

cationiques : His (H), Lys (K), Arg (R)

3

- L'influence de la structure primaire sur la fonction de la protéine est primordiale. **On peut considérer que la séquence en acides aminés est le déterminant principal qui dicte, qui conditionne l'acquisition de la structure tridimensionnelle.**

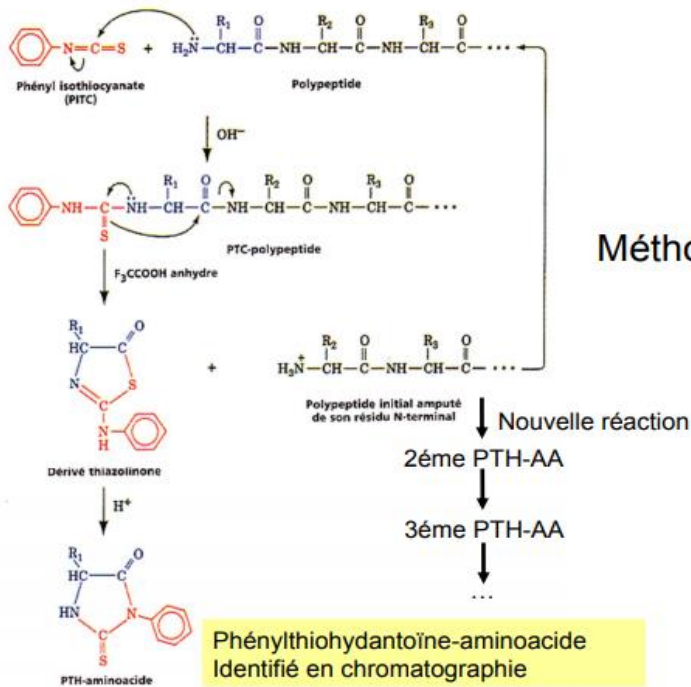
4

Exemple de structure primaire

Grenouille	MALWMQCLPVLVLF	FFSTPNT-EALVNOHLCGSHLVEAI
Poulet	MALWIRSLPLLALV	VFSGPGTSYAAANOHLCGSHLVEAI
Chimpanzé	MALWMRLLPLLVL	LALWGPDPAFAFVNOHLCGSHLVEAI
Homme	MALWMRLLPLLAL	LALWGPDPAFAFVNOHLCGSHLVEAI
Orang-outan	MALWMRLLPLLAL	LALWGPDPAQAFVNOHLCGSHLVEAI
Rat	MALWMRFLPLLAL	VLWEPKPAQAFVKOHLCGPHLVEAI
Souris	MALLVHFLPLLAL	LALWEPKPTQAFVKOHLCGPHLVEAI
Bœuf	MALWTRLRPLLAL	LALWPPPARAFVNOHLCGSHLVEAI
Poisson	MAVWLQAGALLVL	LVSSVSTNPGTP-OHLCGSHLVDAL

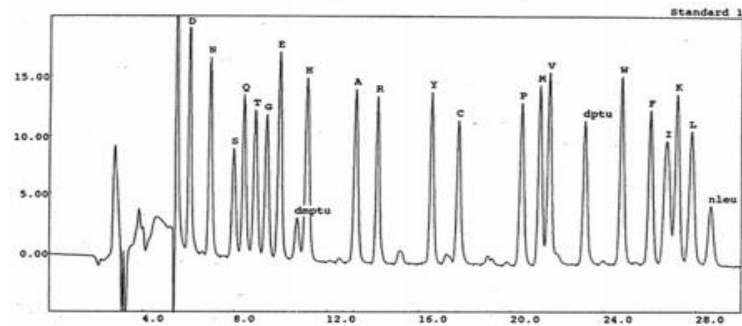
Alignement des séquences primaires de l'insuline de différentes espèces
(TP bio-info)

5



6

Séquenceur automatique



7

Détermination des séquences des acides nucléiques

- Le séquençage génétique de Sanger est un moyen de déterminer l'ordre des quatre nucléotides dans un brin d'ADN. Il est devenu indispensable dans les domaines de la recherche de base, de la biotechnologie, de la médecine légale et du diagnostic médical
- À la fin des années 1970, la biologie voit naître les deux premières méthodes de séquençage d'ADN. L'une de ces méthodes, le séquençage de Maxam-Gilbert, utilise des produits chimiques pour rompre l'ADN et déterminer sa séquence. Frederick Sanger met au point la seconde méthode, pour laquelle il reçoit, avec Maxam et Gilbert, le prix Nobel.

8

La méthode de Sanger : Comment fonctionne-t-elle?

- Pour commencer, il vous faut un brin d'ADN à séquencer. Ensuite, vous ajoutez une séquence d'amorce, les quatre nucléotides (dNTP et ddNTP) et une 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 (**didésoxynucléotide**), 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. Chaque colonne de gel a une base différente à l'extrémité. Par conséquent, chaque colonne représente un brin d'ADN avec cette base à l'extrémité (Figure 1A)

9

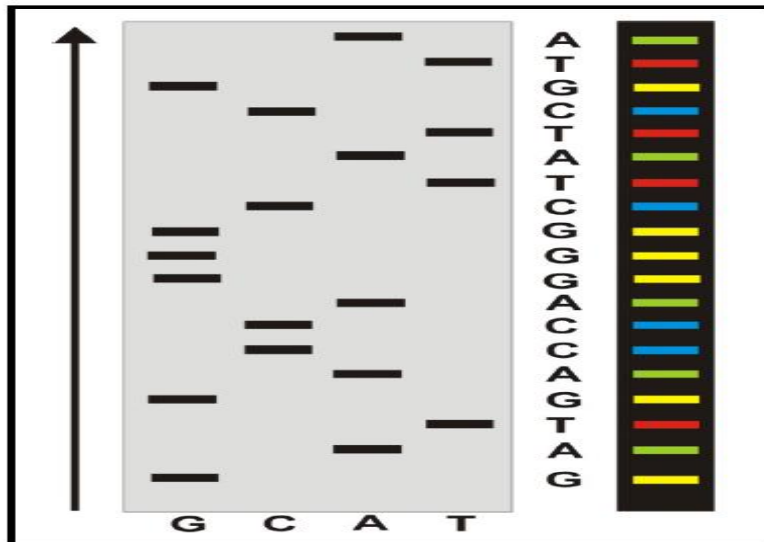
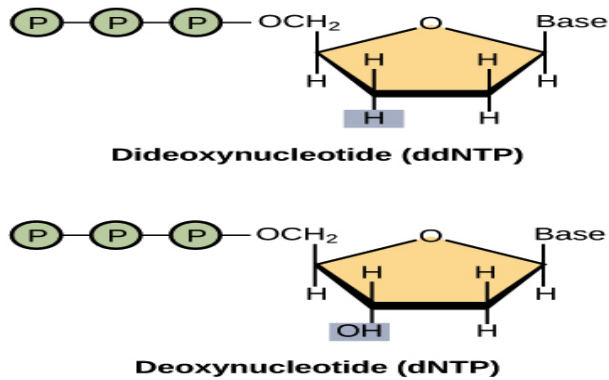


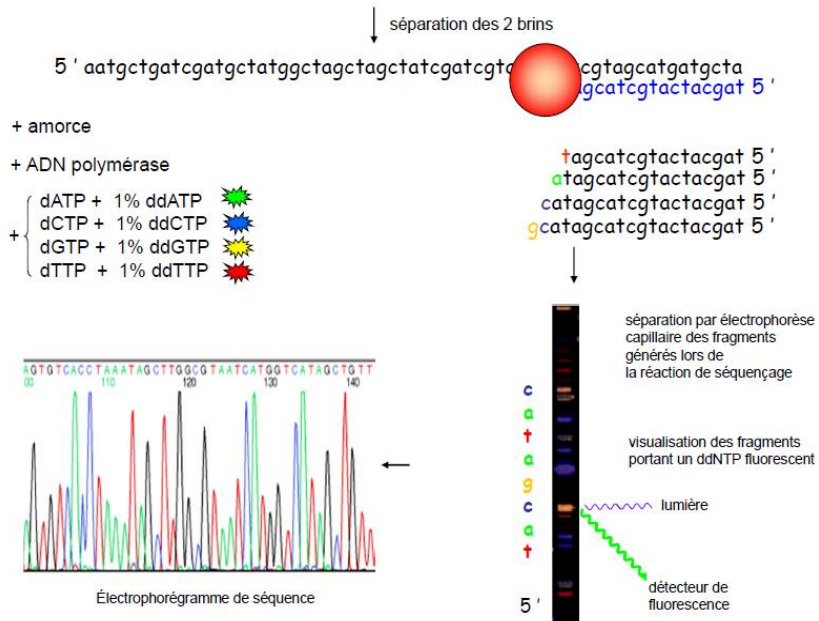
Figure 1A (gauche) : Électrophorèse en gel standard.
Figure 1B (droite) : Séquençage à l'aide de fluorophores.

10

- Des avancées ultérieures dans cette méthode ont incorporé l'utilisation de fluorophores, qui sont de petits composés chimiques dégageant des lumières colorées. En ajoutant un fluorophore coloré différent à chaque nucléotide, le séquençage peut être effectué en une seule réaction avec une seule colonne de gel pour représenter les brins d'ADN; la couleur de la bande indique la base située à l'extrémité du fragment d'ADN (Figure 1B)



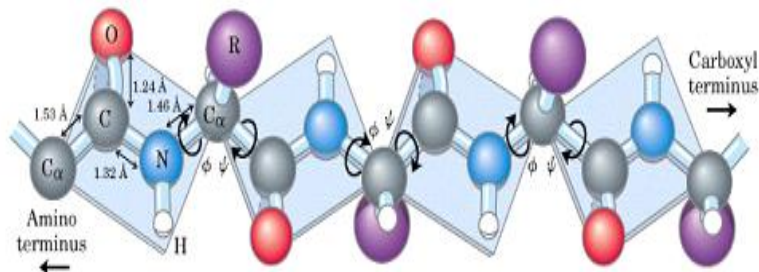
11



12

Structure secondaire

- Repliement local de la chaîne d'acides aminés d'une protéine
- Seules certaines conformations des angles ϕ et ψ sont possibles et déterminent les motifs structuraux. L'angle ϕ correspond à l'angle de rotation N-C α et l'angle ψ correspond à l'angle C α -C.

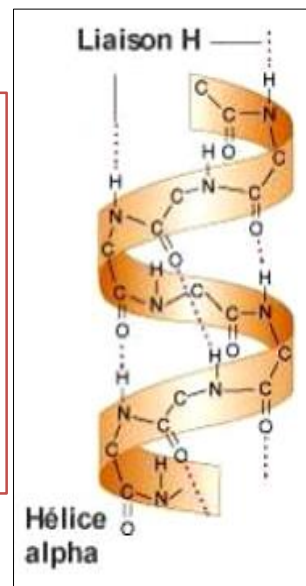


13

Structure secondaire

L'hélice α

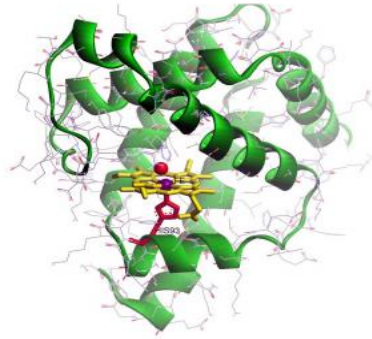
Dans une structure en hélice, chaque oxygène carbonyle forme une liaison H avec l'hydrogène de l'amide situé 4 résidus plus loin. Donc tous les groupements CO et NH de la chaîne polypeptidique sont liés les uns avec les autres, formant une structure très stable.



14

Structure secondaire

Exemples de protéines riches en hélices alpha



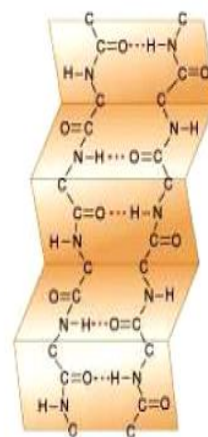
Myoglobine

15

Structure secondaire

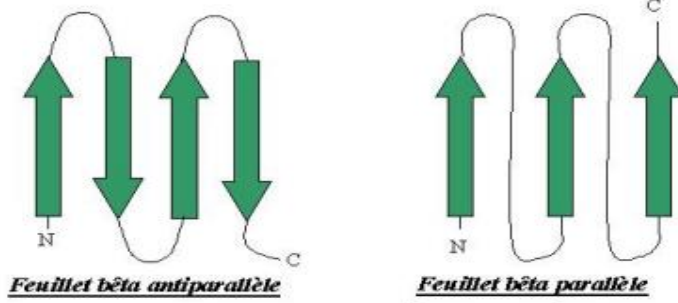
Le feuillet plissé β

Ici, le squelette peptidique ne s'enroule pas autour d'un axe mais se déplie en accordéon, pour former une structure étirée. **C'est en se superposant que ces structures étirées vont établir des liaisons H qui vont les stabiliser.**



Feuillet bêta

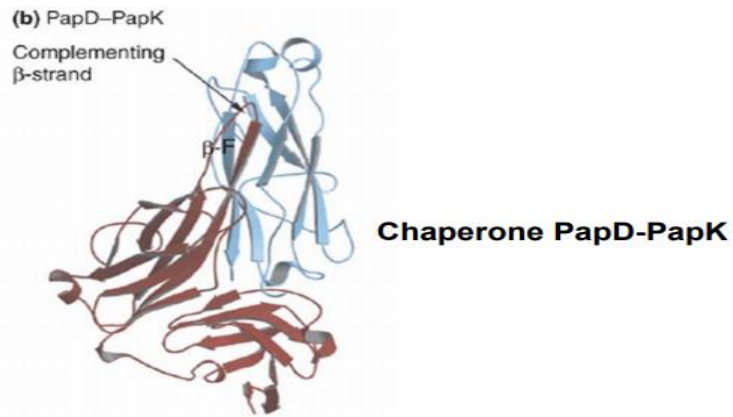
16



17

Structure secondaire

Exemple de protéines riches en feuillets bêta



18

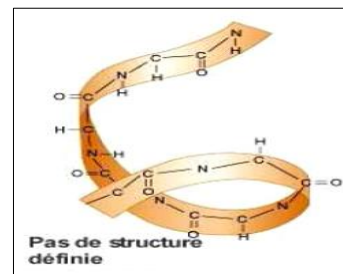
Structure secondaire

- Dans certaines protéines, la quasi-totalité de la séquence adopte une structure en hélice α . Dans d'autres protéines, la quasi-totalité est en feuillet plissé β .
- Mais dans la majorité des protéines, on trouve au sein de la même chaîne polypeptidique à la fois des structures en hélice α et en feuillet plissé β .

19

L'absence de structure secondaire

- Toutes les protéines n'adoptent **pas forcément une structure secondaire sur toute la longueur de la chaîne polypeptidique**. Les structures qui ont une séquence en feuillet ou en hélice alternent avec des séquences qui n'ont pas de structure secondaire.
- La plupart du temps, les séquences sans structures secondaires permettent d'unir entre eux des segments de chaînes qui ont une structure secondaire et vont donc constituer des zones flexibles nécessaires au repliement de la protéine.
D'une façon générale, les séquences sans structure secondaire comportent en moyenne entre 2 et 15 acides aminés qui sont très souvent des acides aminés hydrophiles et qui auront tendance à se localiser à la surface de la protéine.

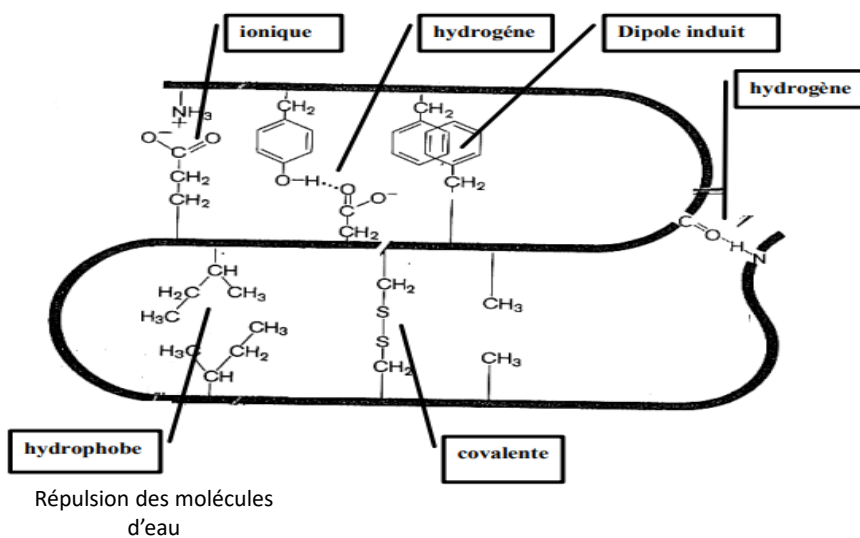


20

La structure tertiaire

La structure tertiaire d'une protéine décrit la façon dont une protéine globulaire se replie dans l'espace à 3 dimensions et décrit l'ensemble des interactions susceptibles de s'établir entre les acides aminés de la chaîne.

21



22

La structure tertiaire

Les grosses protéines globulaires (+ de 200 acides aminés) sont souvent composées de plusieurs régions appelées **domaines**, qui se replient de façon indépendantes et qui constituent souvent des unités fonctionnelles. Par exemple, un domaine protéique peut contenir le site de fixation d'une enzyme, un autre un site de régulation... le tout sur la même protéine, grâce à un repliement indépendant.

23

La structure tertiaire

Existence de différents modes de repliements:


 $(\alpha/\alpha)_6$

 β -jelly roll

 $(\beta/\alpha)_8$

 β -propeller

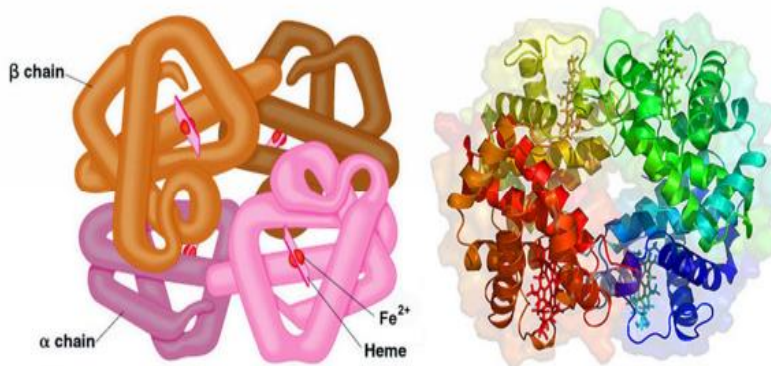
24

La structure quaternaire

- Beaucoup de protéines sont constituées de plusieurs chaînes polypeptidiques indépendantes qui s'associent entre elles de façon non covalentes.
- Les sous unités qui constituent la protéine quaternaire peuvent être identiques ou différentes.
- Leur nombre pourra être variable selon les protéines : l'hémoglobine (première protéine oligomérique dont la structure tridimensionnelle a été publiée en 1959 : 4 sous unités en 2 types = 2 sous unités alpha et 2 sous unités bêta).

25

Exemples de structures quaternaires « L'hémoglobine »



26

Modifications post-traductionnelles des protéines

- Modification covalente des protéines qui change la structure et donc les propriétés des protéines
- Différentes MPT opérées par des enzymes dédiées :
 - clivage protéolytique
 - addition de groupements
- Incidences sur l'activité, la stabilité, le repliement, les interactions avec d'autres protéines

27

Modifications post-traductionnelles des protéines

Modification	Site	Exemple
Acétylation	K	Histones
Gamma-carboxylation	carbone en position gamma de E	prothrombine
Amidation	G terminale	gastrine
Biotinylation	K	pyruvate carboxylase
Carboxylation	K, D, E	Prothrombine
Glycosylation	O-glycosylation sur S, T et hydroxylysine; N-glycosylation sur N	Groupes sanguins Collagène
Hydroxylation	P, K, D et Y	Collagène
Prénylation (farnesylation, geranylgeranylation)	C en C-terminal	Ras
S-acylation	C	rhodopsine
Myristoylation	G en N-terminal	NADH-cytochrome b5
Méthylation	K, H, R, D, E	Actine
Phosphorylation	Y, S, T	MAP kinases
Sulfatage	Y	APP, thyroglobuline
Ubiquitination	K	Histones H2A, H2B

28

La "Protein DataBank"

- La base de données mondiale recueil des structures tri-dimensionnelles des macromolécules biologiques est la [Protein DataBank](#) (PDB).
- Elle contient plus de 131.000 structures de protéines, d'acides nucléiques (ADN et ARN) et de complexes ribonucléo-protéiques, déterminées par différentes techniques :
 - ✓ cristallographie - diffraction des rayons X : plus de 117 000 structures
 - ✓ résonance magnétique nucléaire : plus de 12 000 structures
 - ✓ cryo-microscopie électronique (technique la plus récemment utilisée) : plus de 1600 structures
- Les organismes les plus représentés sont : *Homo sapiens*, *Escherichia coli*, *Mus musculus*, *Saccharomyces cerevisiae*.....

29

Les techniques pour déterminer - analyser les structures des protéines

- Il n'y a pas une technique meilleure qu'une autre. Elles ont toutes leur spécificité avec leurs avantages et leurs inconvénients.
- Les meilleurs modèles structuraux de macromolécules biologiques ou de complexes biologiques résultent de combinaisons de données obtenues par plusieurs de ces techniques.
 - ✓ La diffraction des rayons X / Cristallographie aux rayons X/ radiocristallographie
 - ✓ La résonance magnétique nucléaire
 - ✓ La cryo-microscopie électronique
 - ✓ Autres techniques moins fréquemment utilisées

30

a. Cristallographie aux rayons X

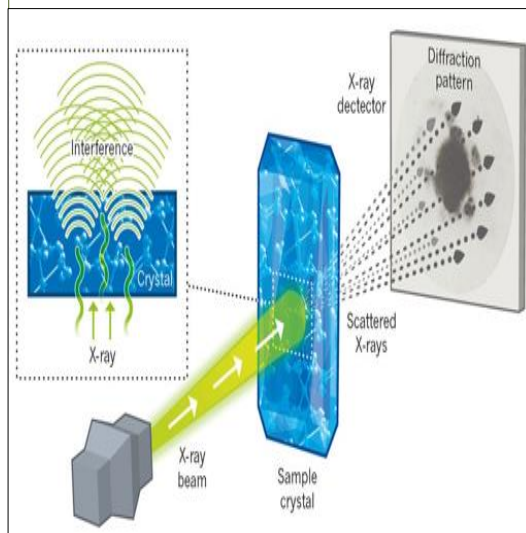
- Technique la plus utilisée
- 1^{ères} structures “hémoglobine” (M. Perutz) et «myoglobine » (J. C. Kendrew) (1958). Ils ont partagé le prix Nobel en chimie en 1962.
- Quantités nécessaires : dizaine(s) de mg de protéine **pure** en solution
- Cristal = organisation des molécules en un réseau ordonné
- Obtention du cristal: protéine pure+ solution de cristallisation on laisse la goutte s'évaporer. Un mécanisme provoque un réagencement des molécules qui compose cette goutte. Avec du temps et les bonnes conditions de cristallisations, nous obtiendrons des cristaux de protéine.

31

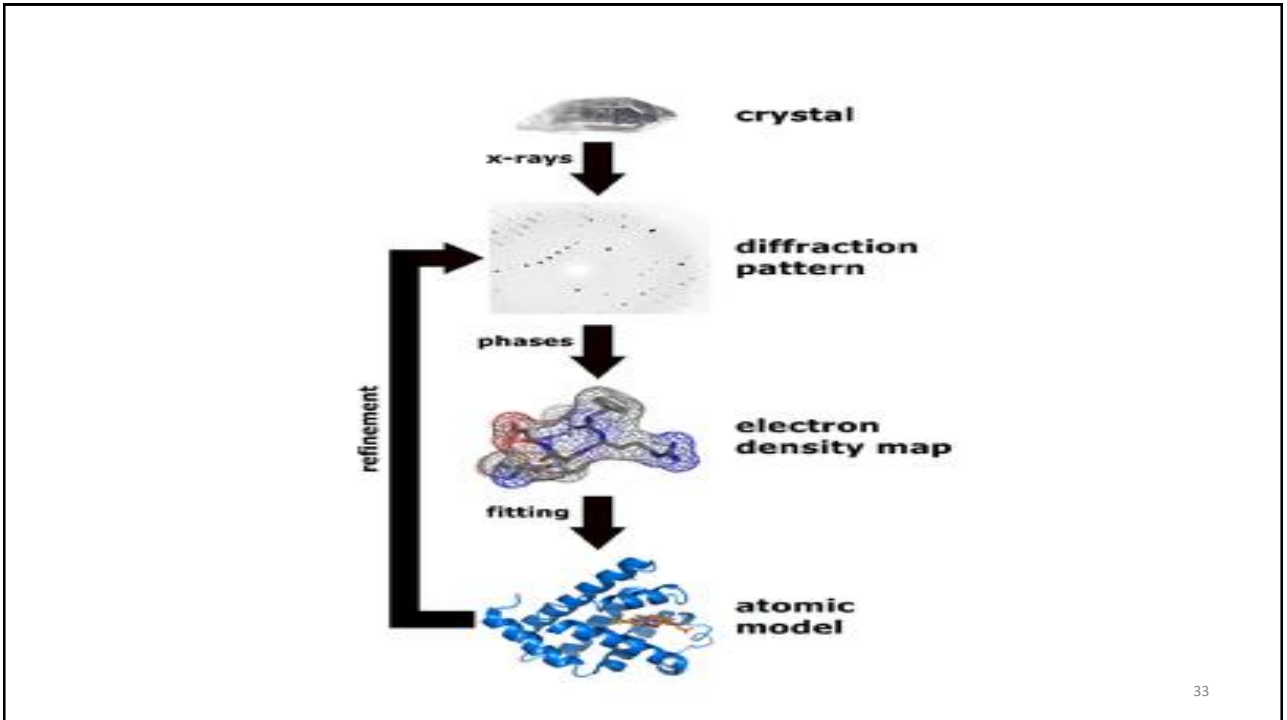
a. Cristallographie aux rayons X

Cette méthode utilise un faisceau de rayons X qui, rencontrant un cristal, est renvoyé dans des directions spécifiques déterminées par la longueur d'onde des rayons X et par les dimensions et l'orientation du réseau cristallin.

À partir des intensités diffractées il est possible, à partir d'une série d'images de diffraction, de déterminer l'arrangement tridimensionnel des atomes d'une structure cristalline.



32



33

b. La résonance magnétique nucléaire

- En biologie structurale, la RMN s'est fait une place à côté de l'autre grande technique qu'est la Cristallographie aux Rayons X:
 - Parce que certaines protéines ne cristallisent pas
 - Parce que l'échantillon de RMN est plus proche des conditions natives (solution aqueuse, température ambiante, interaction avec des ligands ou des récepteurs, dynamique...)
- Prix Nobel dans le domaine de la biochimie, le prix Nobel 2002 a été attribué à K. Wüthrich pour ses travaux concernant la structure tridimensionnelle des macromolécules

34

b. La résonance magnétique nucléaire

- La résonance magnétique nucléaire (RMN) des protéines étudie principalement le **repliement de la chaîne polypeptidique** dans l'espace.
- Une caractéristique importante des spectres RMN de protéines est leur extrême complexité, due au grand nombre de noyaux observables dans l'échantillon. Les différentes étapes de la détermination de structure de protéine par RMN sont : la préparation de l'échantillon, l'acquisition et le traitement du signal RMN produit par les noyaux observés, l'analyse et l'attribution des spectres, le calcul des coordonnées des atomes à partir des paramètres mesurés sur les spectres RMN.

35

c-La cryo-microscopie électronique

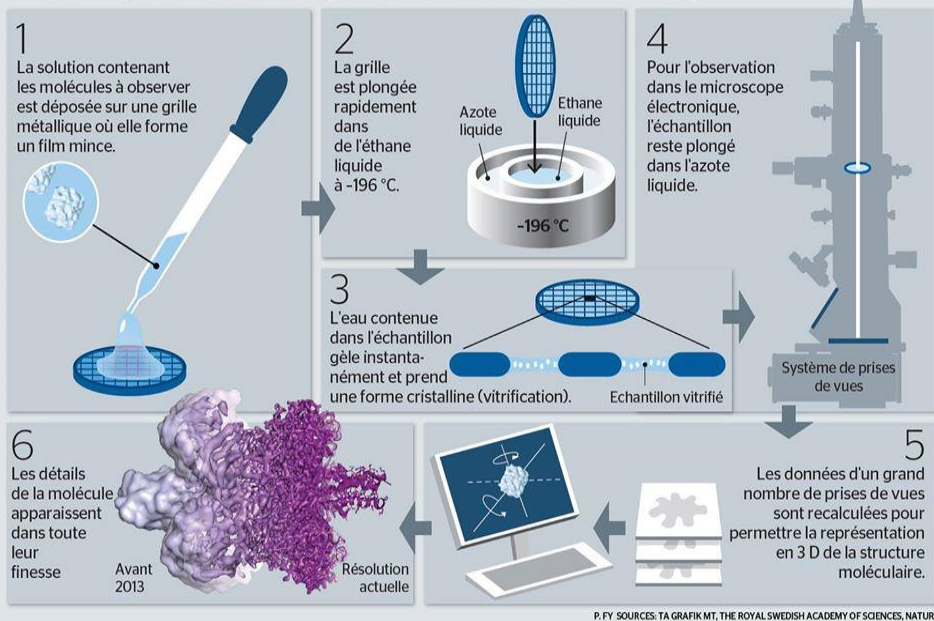
single-particle cryo-electron microscopy" - Cryo-EM

- est l'une des techniques de microscopie électronique. Le préfixe cryo- signifie «froid» en grec ancien. Les techniques de cryomicroscopie utilisent le froid lors de l'étape de fixation.
- Le prix noble en chimie 2017 est décerné au Suisse Jacques Dubochet, à l'Allemand Joachim Frank et à l'Écossais Richard Henderson, pour le développement de la cryo-microscopie électronique, qui simplifie et améliore l'imagerie des biomolécules.
- Grâce à elle, il est désormais possible, non seulement de distinguer les protéines mais aussi de voir les détails les plus intimes de la structure moléculaire.
- Étapes dans la figure suivante + vidéo

36

Comment observer chaque **molécule** séparément






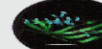
La cryo-microscopie électronique, inventée par le nouveau **Prix Nobel Jacques Dubochet** et ses équipes, permet d'observer la structure complexe des échantillons biologiques (virus, protéines, enzymes). Lire aussi en page 17.



37

- Le processus expérimental pour l'élaboration d'une structure 3D est résumé à l'aide du graphique suivant: (cas cristallographie)

38

 <p>Target Selection</p>	<p>Sélection de la cible : Pour chaque protéine cible, plusieurs séquences homologues de plusieurs espèces (ou de constructions) différentes sont testées.</p>
 <p>Cloning & Expression</p>	<p>Clonage et expression : après sélection, les gènes sont amplifiés et clonés dans des vecteurs d'expressions puis les protéines sont produites en masse dans des bactéries ou des cellules d'insectes. Cette étape est délicate à mettre en œuvre.</p>
 <p>Protein Purification</p>	<p>Purification des protéines : les protéines sont purifiées par des techniques de chromatographie d'affinité, de gel-filtration et d'échange d'ions.</p>
 <p>Crystallization</p>	<p>Cristallisation : Les protéines purifiées sont mises en condition pour cristalliser. Les différentes techniques de cristallisation nécessitent un très grand nombre d'essais.</p>
 <p>X-Ray Diffraction</p>	<p>Diffraction aux rayons X : Une fois les cristaux obtenus, ils sont exposés à un rayonnement intense de rayons X (une source synchrotron) dans un appareillage complexe servant à collecter des clichés de diffraction.</p>
 <p>Structure Determination</p>	<p>Détermination de la structure : Les structures sont résolues par différentes techniques de phasage (MIR, MAD,...) et des techniques informatiques sophistiquées.</p>