

Master : Biotechnologie et Pathologie Moléculaire

Matière : Dynamique structurale des macromolécules

Chapitre V : Détermination expérimentale des structures protéiques

I. DETERMINATION DE LA STRUCTURE 3D DES PROTEINES

Déterminer la structure tridimensionnelle des protéines à un niveau de résolution de l'ordre de l'atome est utile pour l'identification de leur fonction, la conception de médicaments basée sur la structure et l'amarrage moléculaire.

RMN : *La spectroscopie par résonance magnétique nucléaire (RMN)* permet d'obtenir des informations sur la structure et la dynamique des protéines. Dans l'analyse par RMN, l'emplacement spatial des atomes est déterminé par leurs déplacements chimiques. Les protéines destinées à une analyse par RMN sont généralement marquées par des isotopes stables (^{15}N , ^{13}C , ^2H) afin d'accroître la sensibilité tout en facilitant la déconvolution structurale. Les marqueurs isotopiques sont généralement introduits par l'apport de nutriments marqués isotopiquement dans le milieu de culture lors de l'expression des protéines.

Radiocristallographie : *La radiocristallographie* permet d'obtenir la structure tridimensionnelle des protéines par *diffraction des rayons X* des protéines cristallisées. La cristallogénèse est réalisée en ensemençant des solutions riches en protéines favorisant la précipitation, avec des cristaux de protéines organisés se formant dans des conditions appropriées. Des rayons X sont envoyés sur les cristaux de protéines, qui les diffusent vers un film ou un détecteur électronique. On fait pivoter les cristaux pour capturer la diffraction en trois dimensions, ce qui permet de calculer la position de chaque atome dans la molécule cristallisée par transformée de Fourier.

II. La cristallographie aux rayons X

II.1. Histoire de bio-cristallographie

- **1912** : Le physicien allemand **Max von Laue** découvre la diffraction des rayons X par les cristaux. Cette avancée montre que les cristaux ont une structure périodique à l'échelle atomique.

- **1913** : **William et Lawrence Bragg** développent les lois mathématiques (loi de Bragg) pour interpréter les motifs de diffraction, jetant les bases de la cristallographie moderne.
- **1934** : **J.D. Bernal et Dorothy Crowfoot Hodgkin** obtiennent les premiers clichés de diffraction sur une enzyme, la **pepsine**, en utilisant des cristaux hydratés. Ils montrent qu'il est possible d'étudier des macromolécules biologiques avec cette méthode
- **1953** : **Watson et Crick**, en s'appuyant notamment sur les clichés de diffraction de **Rosalind Franklin**, élucident la structure de la **double hélice de l'ADN**.
- **1958** : **Hodgkin** détermine la première structure d'une petite protéine : **l'insuline**.
- **1960** : **John Kendrew et Max Perutz** déterminent les structures de la **myoglobine** et de **l'hémoglobine**, premières protéines de grande taille résolues.
- **Années 1970–2000** : Les progrès dans les **ordinateurs**, les **sources de rayons X synchrotron**, la **génétique moléculaire** (production de protéines recombinantes) et la **modélisation informatique** accélèrent les découvertes.
- **2000s et après** : Apparition de banques de données comme la **Protein Data Bank (PDB)**, qui recense aujourd'hui des centaines de milliers de structures.

II.2. Fenêtre sur la structure des macromolécules biologiques

La cristallographie est la méthode de prédilection pour observer les structures des macromolécules. Depuis la découverte de la diffraction des cristaux par des rayons X par Max Von Laue en 1914, de nombreux développements techniques ont permis de résoudre la structure de macromolécules d'intérêt biologique et pharmaceutique.

Les méthodes de cristallographie haut débit ont fait leur apparition dans les années 1990. Celles-ci permettent de trouver de manière semi-automatisée les conditions de cristallisations pour les protéines, afin de produire **des cristaux de protéines** qui pourront être irradiés à l'aide de faisceaux de rayons X produit par des synchrotrons ou des sources de laboratoires pour réaliser des expériences de diffraction.

La structure des protéines est dynamique et ne saurait se réduire à une conformation figée dans un cristal. La plupart des protéines gardent leurs propriétés catalytiques même sous forme cristalline.

Découverts en 1895 par le physicien allemand Röntgen, les rayons X sont à la base de différentes techniques d'analyse comme la radiographie, la spectroscopie et la diffractométrie.

Ces radiations électromagnétiques ont une longueur d'onde de l'ordre de l'Ångström ($1 \text{ \AA} = 10^{-10} \text{ m}$) ou ($0.1 < \lambda < 10 \text{ nm}$).

Un **crystal** est un agencement d'atomes, d'ions ou de molécules, avec un motif se répétant périodiquement dans les trois dimensions. Les distances interatomiques sont de l'ordre de l'Ångström, du même ordre de grandeur que les longueurs d'onde des rayons X : un cristal constitue donc un réseau 3D qui peut diffracter les rayons X.

Le **dispositif comprend** un tube scellé lié à la cathode qui est un filament de tungstène. La partie de l'anticathode, est une plaquette métallique très pure qui cible des rayons X. En effet, les propriétés des rayons X émis dépendent de la nature de l'anticathode (anode).

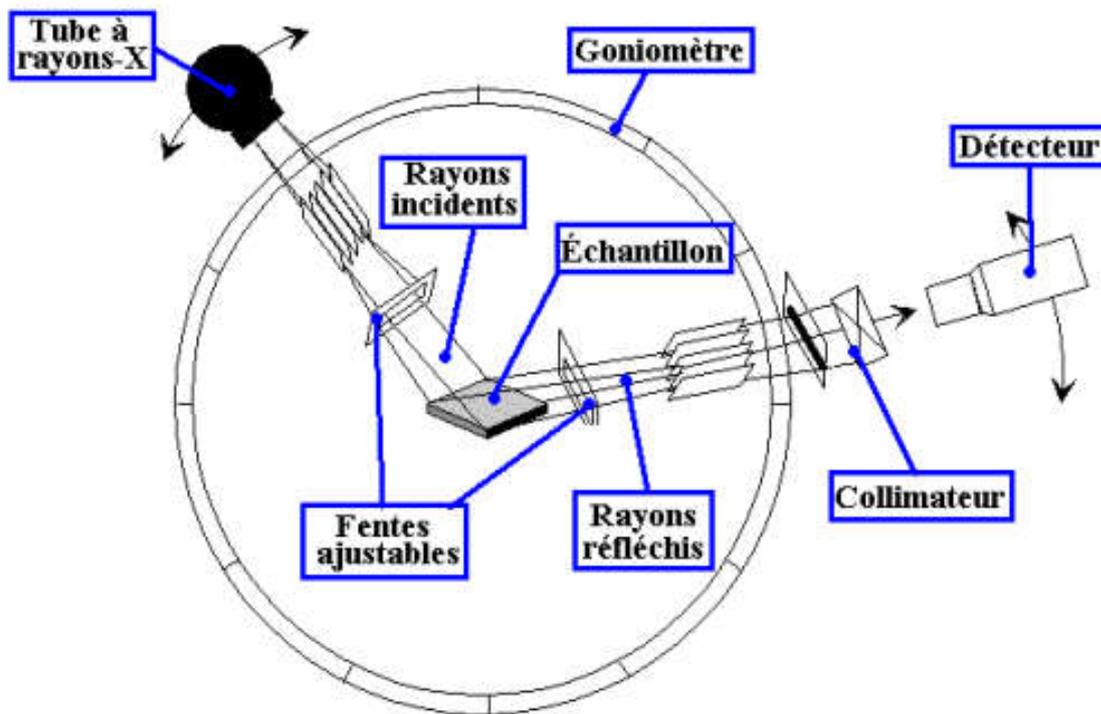


Schéma d'un spectromètre (SIEMENS D500) de diffraction des rayons X.

III. La résonance magnétique nucléaire

La RMN est une **méthode spectroscopique** d'analyse de la matière, fondée sur les *propriétés magnétiques* de certains noyaux atomiques. L'échantillon à étudier, placé dans un champ magnétique très intense, acquiert *une aimantation nucléaire* qui est détectée par sa mise en résonance avec un champ électromagnétique.

La RMN repose sur le fait que les noyaux de certains atomes, à « *nombre de spin* » non nul, existent dans plusieurs états dont les énergies sont différentes en présence d'un champ magnétique. La différence d'énergie entre ces états est proportionnelle au champ magnétique affectant le noyau et que ce champ est la somme du champ externe, créé par le spectromètre, et d'un champ local, créé par l'environnement du noyau.

Le processus de détermination de structures par RMN se poursuit par l'enregistrement et le traitement de spectres permettant d'attribuer une valeur de déplacement chimique à chacun des atomes de la molécule, puis par l'enregistrement de spectres permettant l'extraction d'informations structurales. La modélisation de structures sous contraintes et la validation des modèles constituent les étapes finales du processus.

III.1. Principe de base

Tous les noyaux atomiques possèdent *une charge en rotation*, identifiée sous le nom de *spin nucléaire* (ils sont assimilables à des petits aimants et de ce fait peuvent présenter un moment magnétique nucléaire).

Remarque : Certains noyaux ne sont pas observables en RMN, car ils n'ont pas de propriétés magnétiques. Sous l'action d'un champ magnétique externe uniforme, le noyau atomique (son moment magnétique nucléaire) peut prendre différentes orientations.

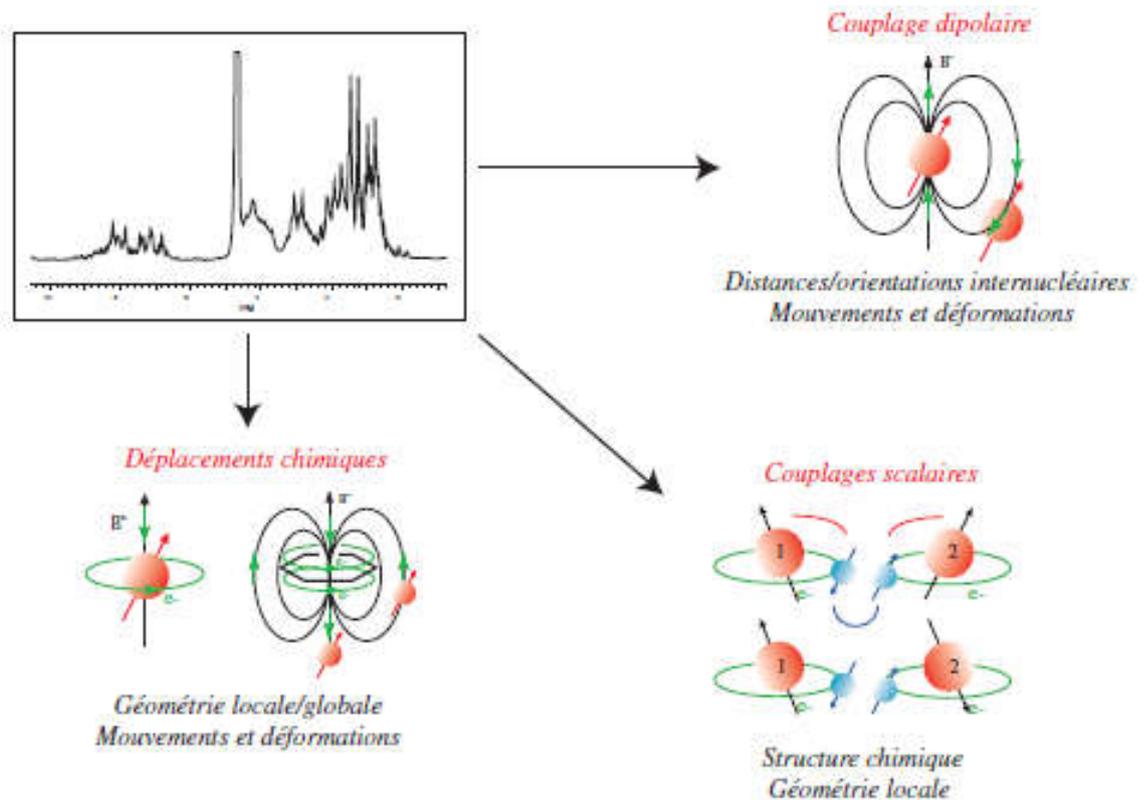
A ces différentes orientations, correspondent différents niveaux d'énergie :

- l'un de basse énergie, si le moment magnétique est parallèle et de même sens que le champ extérieur,
- l'autre d'énergie plus élevée, si le sens est contraire.

La différence d'énergie ΔE entre ces deux états est proportionnelle au champ extérieur. La transition du niveau bas au niveau haut peut avoir lieu par absorption d'une radiation de fréquence ν telle que $\Delta E = h\nu$.

Les fluctuations peuvent être créées par le spectromètre (sous forme d'impulsions de champ), mais sont aussi dues aux mouvements et aux déformations de la molécule. Finalement, la différence d'énergie entre les états d'un noyau dépend aussi de l'état de spin des noyaux qui lui sont chimiquement liés et/ou qui lui sont spatialement proches.

Chaque raie d'un spectre RMN (ici un spectre proton enregistré à 600 MHz sur une protéine) renseigne sur l'environnement local d'un proton ou d'un groupe de protons.



Les mouvements moléculaires sont à l'origine de la richesse des observations procurées par la RMN, que ce soit en haute résolution, en imagerie médicale ou industrielle. Cette dernière propriété en fait une méthode parfaitement complémentaire de la cristallographie par RX pour la détermination de la structure 3D de molécules biologiques. L'accroissement exponentiel du nombre de structures obtenues par RMN en solution démontre l'intérêt de cette approche.

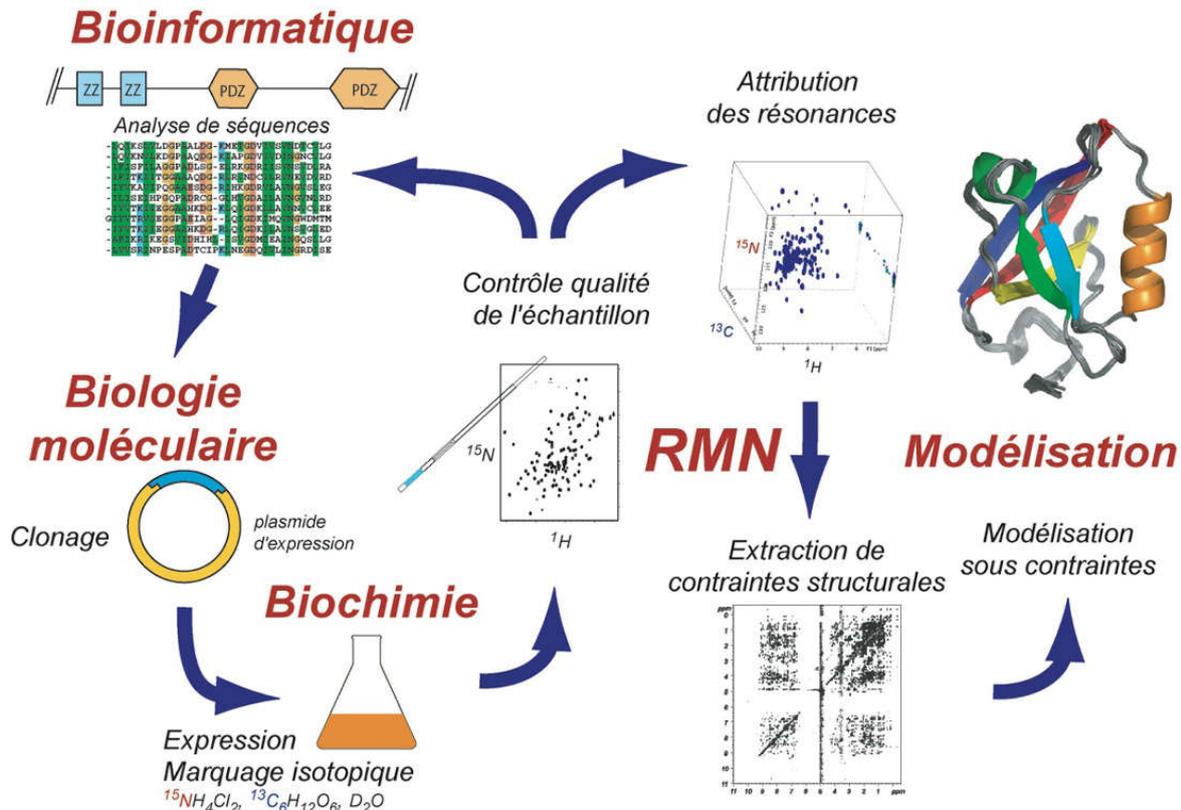
III.2. Comparaison entre RMN et cristallographie

La RMN appliquée aux macromolécules biologiques se distingue de la cristallographie par plusieurs points forts :

- les mesures de structures ne sont pas dommageables pour les échantillons protéiques et elles sont réalisées en solution, alors que la radiocristallographie des protéines nécessite au préalable de disposer d'un cristal de protéine de très bonne qualité.

- La RMN permet également de mettre en évidence les fluctuations conformationnelles des protéines, qui jouent souvent un rôle majeur dans la reconnaissance entre molécules ou dans le fonctionnement intime et dynamique des enzymes

De la séquence d'une protéine à sa structure



IV. BASE DE DONNÉES PDB

Les structures tridimensionnelles sont déposées dans une base de données spécialisée, appelée « *Protein Data Bank* » (PDB, RSCB), située actuellement aux États-Unis (Rutgers, New Jersey) (<http://www.rcsb.org/pdb/>). Cette base de données est d'accès libre et gratuit et permet à la communauté internationale d'avoir accès aux coordonnées des structures résolues. Le dépôt des coordonnées d'une macromolécule est devenu une condition préalable à la publication des travaux de recherche dans une revue internationale de haut niveau. Il faut se rappeler que les publications sont le critère premier de jugement du travail scientifique d'un chercheur du secteur public. Le dépôt de coordonnées est une pratique moins courante dans le monde industriel où, au contraire, on ne souhaite pas communiquer aux concurrents les données si précieuses que constituent les coordonnées atomiques d'une structure.

Protein Data Bank (PDB) est la banque de référence des structures protéiques obtenues expérimentalement par cristallographie aux rayons X ou spectroscopie RMN.

En 2007, 7 272 structures nouvelles ont été déposées, soit en moyenne près de 20 structures par jour. À la date de remise à jour de cette revue (octobre 2008), la PDB contenait 53 521 structures dont 85 % de structures cristallographiques.

Les coordonnées des atomes formant la structure d'une protéine, le détail de la séquence, les conditions de cristallisation sont les principales informations disponibles pour chaque structure de la banque PDB.

C'est à partir de cette banque que sont détectés *les homologues structuraux*. La majorité des séquences ont une homologie structurale inférieure à 20% ; environ 1000 structures protéiques originales peuvent modéliser la quasi-totalité des protéines connues (Modélisation Moléculaire).

