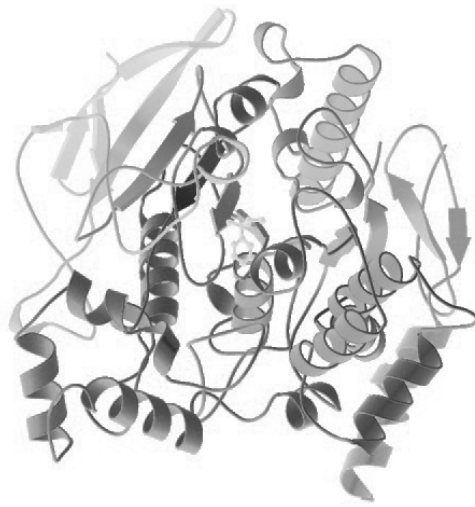


Cristallographie biologique (Seconde partie)



Bioinformatique structurale - X9IB040

M2 – BS / Bio-Info

Master de Biologie-Santé

Année 2013-2014

Michel Evain

Introduction

Ce cours de cristallographie biologique vise à introduire certaines notions sur les cristaux et leur étude par la diffraction des rayons X. En fait, cette enseignement ne se rattache pas particulièrement à la biologie ou à la biochimie, mais concerne en fait tous les domaines scientifiques dans lesquels interviennent des cristaux : la biologie et la biochimie, mais aussi la chimie organique, la chimie inorganique, la pharmacie, la médecine, la géologie, la métallurgie, etc. Les fondements de la cristallographie géométrique ont été posés dans une première partie en M1 (BioInfo-BioStat). Cette seconde partie aborde la diffraction des rayons X par les cristaux et la détermination des structures biologiques (protéines, virus, etc.).

Chapitre 1

Diffraction des rayons X

Introduction

Dans ce premier chapitre nous allons exposer les principes géométriques généraux de la diffraction des rayons X par un cristal, à la fois dans l'espace direct et dans l'espace réciproque. Nous verrons ensuite brièvement comment il est possible, à partir de la géométrie des faisceaux diffractés, de tirer des informations sur la métrique du système. Finalement, nous décrirons succinctement quelques appareils qui permettent de produire et détecter les rayons X et de mesurer avec précision les intensités diffractées.

1.1 – De la diffusion à la diffraction

1.1.1 – Rappel sur les ondes

Une onde électromagnétique monochromatique simple, comme la lumière visible ou les rayons X, peut être décrite par une fonction périodique,[†] par exemple sous les formes (pour un système à une dimension) :

$$f(x) = F_h \cos 2\pi(hx - \alpha) \quad \text{ou} \quad f(x) = F_h \sin 2\pi(hx - \alpha) \quad (1.1.1)$$

Dans ces expressions, $f(x)$ indique la hauteur de l'onde à la position x le long de l'onde. Le produit hx et la constante α sont exprimés en fraction de 2π (équivalent à une longueur d'onde). Ainsi, pour $h = 1$, $x = 1$ dénote une position à 2π radians ou 360° de l'origine, ou de façon équivalente à une répétition entière de la fonction périodique (une longueur d'onde, voir figure 1.1.1a). La constante F_h indique l'amplitude de l'onde (la hauteur des crêtes et la profondeur des creux). La constante h précise la fréquence de l'onde (dimension inverse de celle de x). La figure 1.1.1b présente une onde de

[†] Les notations utilisées ici ne sont pas les notations usuelles : fréquence ν et temps t , elles sont adaptées pour une utilisation ultérieure au chapitre 2. Notez qu'il est équivalent de regarder une onde "à une position donnée en fonction du temps" et "à un temps donné en fonction de la position".

fréquence $h = 3$, c'est à dire avec une longueur d'onde trois fois plus petite que celle de l'onde de la figure 1.1.1a où $h = 1$. Finalement, α précise la phase de l'onde, c'est à dire la position de l'onde par rapport à l'origine. Sur la figure 1.1.1c, $\alpha = 1/4$ et l'onde est décalée de $2\pi/4$, c'est à dire de $\lambda/4$.

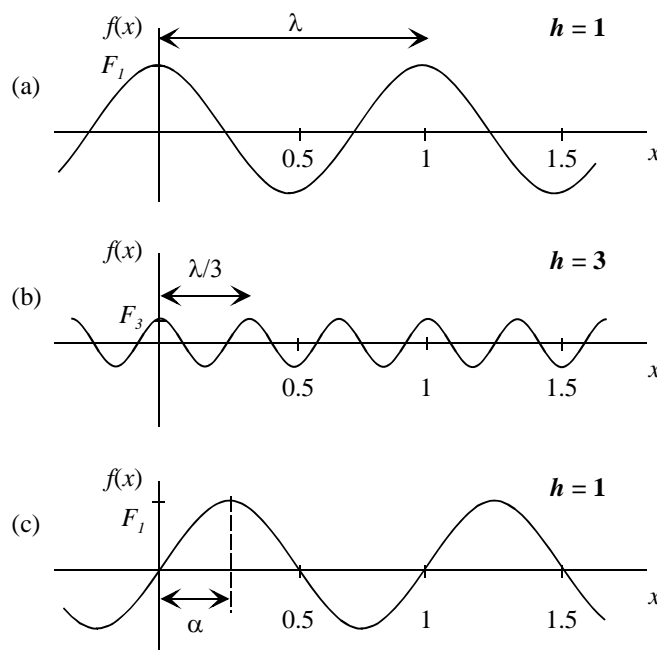


Fig. 1.1.1 – Fonctions d'onde (forme en cosinus)

1.1.2 – Diffusion par un atome

L'interaction des rayons X, ondes électromagnétiques, avec les électrons de la matière conduit à deux effets :

- l'absorption du rayonnement conduisant à des transitions entre les niveaux atomiques
- la diffusion du rayonnement (dispersion dans différentes directions) qui est elle-même de deux types : 1) la diffusion Compton (diffusion dite incohérente, c'est-à-dire avec perte d'énergie) et 2) la diffusion Thomson (diffusion dite cohérente, c'est-à-dire sans perte d'énergie). C'est la diffusion Thomson (voir figure 1.1.2) qui conduit à la diffraction (\equiv diffusion cohérente par une structure périodique), la diffusion Compton n'étant qu'une perturbation fort heureusement très faible.

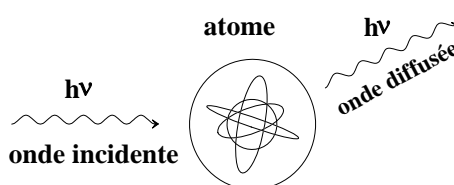


Fig. 1.1.2 – Diffusion Thomson

1.1.3 – Diffusion par deux atomes

Les ondes X diffusées par deux atomes interfèrent (se combinent) de façon constructive ou de façon destructive suivant la direction d'observation choisie. Sur la figure 1.1.3 on note que la différence de marche (= de parcours) entre les deux ondes diffusées est égale à :

$$\delta = AQ - BP \quad (1.1.2)$$

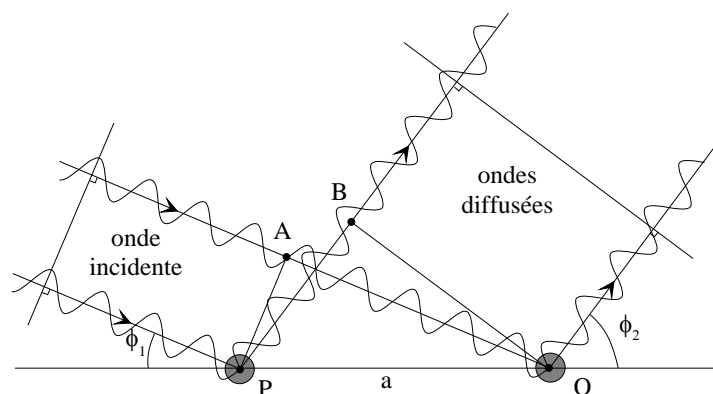


Fig. 1.1.3 – Diffusion par deux atomes P et Q

La combinaison des deux ondes diffusées est constructive (ondes en phase) si la différence de marche est égale à un nombre entier de fois la longueur d'onde λ :

$$\delta = a(\cos \phi_1 - \cos \phi_2) = n\lambda \quad (1.1.3)$$

Elle est destructive (ondes en opposition de phase) pour une différence de marche égale à un nombre entier et impair de fois $\lambda/2$. Dans les cas intermédiaires, il y a à la fois un changement d'amplitude et de phase comme indiquée sur la figure 1.1.4.

2.1.4 – Diffusion par une rangée ou un plan d'atomes

La différence de phase entre deux ondes diffusées par deux atomes P et Q étant liée à la différence de marche entre les deux ondes, il est facile de voir que si l'angle d'incidence est égal à l'angle de réflexion ($\phi_1 = \phi_2$), en d'autres termes si $\delta = 0$, il n'existe aucune différence de phase entre les ondes diffusées par les deux atomes, comme d'ailleurs entre celles diffusées par toute paire d'atomes d'une rangée infinie. L'interférence est alors constructive. Changeons maintenant l'angle de réflexion (ϕ_2) par une quantité infinitésimale. La différence de marche δ devient alors finie et non nulle, et pour un atome P1 donné il est toujours possible de trouver dans la rangée (supposée infinie) un atome P2 tel que δ soit égal à $(2n+1)\lambda/2$, ce qui implique l'annulation de la diffusion. Par conséquent, pour une rangée infinie d'atomes, l'amplitude diffusée est

nulle pour tous les angles, exception faite de celui pour lequel l'angle d'incidence est égal à l'angle de réflexion. On parle alors, non plus de diffusion, mais de diffraction. La même condition s'applique à un plan infini d'atomes.

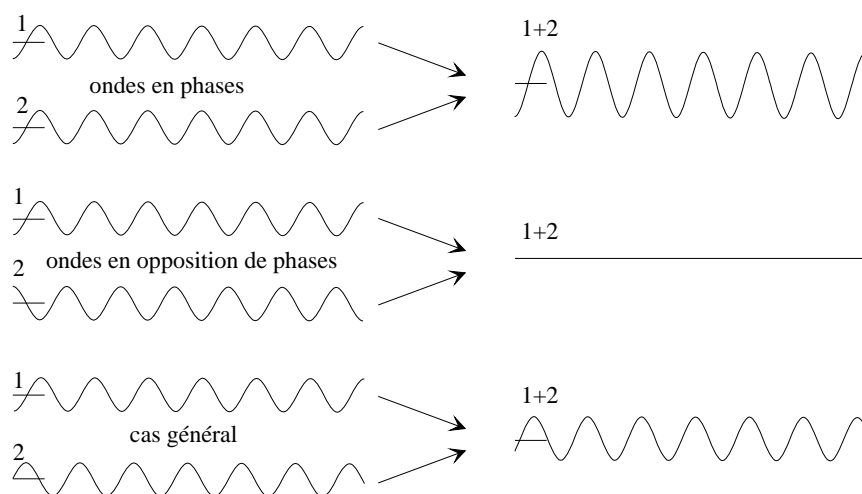


Fig. 1.1.4 – Interférence entre deux ondes

1.1.5 – La loi de Bragg

Le principe développé par Bragg est le plus général et le plus puissant de la théorie de la diffraction. La figure 1.1.5 montre le schéma de diffraction pour deux plans diffusants P1 et P2 distants de d et interceptant un faisceau X, de longueur d'onde λ , avec un angle d'incidence θ . Plaçons-nous dans le cas où chaque plan diffracte, c'est-à-dire pour un angle de réflexion θ . Le paramètre expérimental est en fait l'angle 2θ entre l'onde X diffractée et celle qui n'a pas été déviée. Le diagramme présenté sur la figure 1.1.5 ne montre que deux plans mais il existe implicitement de très nombreux plans identiques (une infinité) séparés par la distance d . Ces plans ne correspondent pas nécessairement avec des plans d'atomes (même si ce sont les atomes qui diffusent), ce sont en fait les plans réticulaires (hkl) introduits dans la première partie. Nous verrons au chapitre 2 que l'intensité du faisceau diffracté par une famille de plans réticulaires dépend de la position des atomes par rapport aux plans, plus les atomes sont proches des plans et plus l'intensité est forte. Pour avoir une interférence constructive entre les plans, la différence de marche δ doit être égale à un nombre entier de longueur d'onde, soit :

$$\delta = O'A + O'B = n\lambda \quad (1.1.4)$$

Un simple calcul géométrique montre que la différence de marche est par ailleurs liée à d_{hkl} par la relation :

$$\delta = 2d_{hkl} \sin \theta \quad (1.1.5)$$

Nous avons alors la relation établie par Bragg (dite loi de Bragg) :

$$2d_{hkl} \sin \theta = n\lambda \quad (1.1.6)$$

où n définit l'ordre de la diffraction.

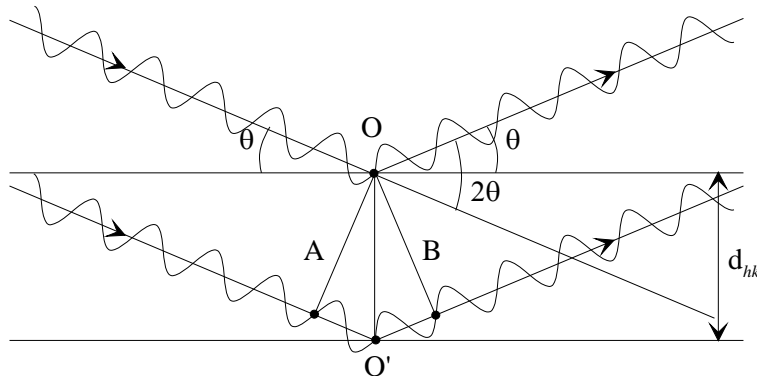


Fig. 1.1.5 – Equation de Bragg, réflexion par des plans réticulaires

Rappelons que la droite nodale du réseau réciproque $\mathbf{T}^*(H,K,L) = H\mathbf{a}^* + K\mathbf{b}^* + L\mathbf{c}^*$, avec $H = nh$, $K = nk$ et $L = nl$ (h , k et l premiers entre eux), est perpendiculaire à la famille de plans réticulaires (hkl) du réseau cristallin. La norme de \mathbf{T}^* est $|\mathbf{T}^*(H,K,L)| = n/d_{hkl}$. Donc $2\sin\theta / \lambda = n/d_{hkl} = 1/d_{HKL}$. La diffraction d'ordre n par des plans réticulaires (hkl) peut donc être interprétée comme une diffraction d'ordre 1 par des plans réticulaires (HKL) :

$$2 d_{hkl} \sin \theta = n \lambda \quad (h, k \text{ et } l \text{ premiers entre eux}) \quad (1.1.7)$$

$$2 d_{HKL} \sin \theta = \lambda \quad (H = nh, K = nk \text{ et } L = nl) \quad (1.1.8)$$

D'après la définition des plans réticulaires, la famille de plans (HKL) n'existe pas car certains plans ne passent pas par des nœuds (il s'agit en effet de la famille (hkl) et de plans vides intercalés). En pratique, il est cependant plus commode de se référer à la réflexion (222), sans en préciser l'ordre (puisque $n = 1$), qu'à la réflexion (111) en précisant qu'elle est du deuxième ordre ($n = 2$).

En résumé, les plans réticulaires se comportent comme des miroirs vis à vis des rayons X, d'où le nom de **réflexion** donné à chaque faisceau diffracté.

1.1.6 – Construction d'Ewald

Dans la première partie du cours nous avons indiqué que le réseau réciproque facilitait la visualisation des plans réticulaires et leur orientation. C'est la **construction d'Ewald** qui permet cette simplification. La construction d'Ewald s'effectue de la façon suivante (voir figure 1.2.6) :

- on dessine en premier lieu le réseau réciproque ;

- on positionne ensuite le faisceau X incident, caractérisé par un vecteur \mathbf{S}_0 de module $1/\lambda$, de sorte que la pointe de \mathbf{S}_0 coïncide avec l'origine du réseau réciproque (nœud 000) ;
- on trace alors une sphère autour du point M marquant le début de \mathbf{S}_0 , et passant par l'origine (la sphère, dite d'Ewald, a donc pour rayon $1/\lambda$) ;
- on place enfin le cristal au point M (même si les origines des réseaux direct et réciproque sont souvent confondues).

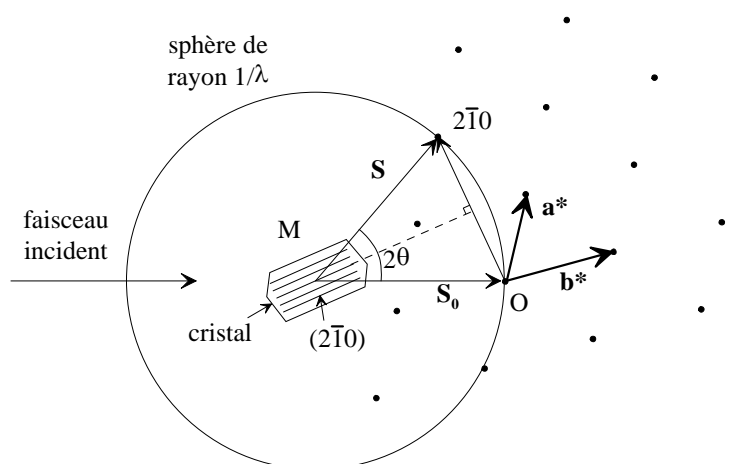


Fig. 1.2.6 – Construction d'Ewald

Par construction, tout nœud du réseau réciproque situé sur la sphère d'Ewald est en condition de diffraction (ou plus exactement les plans réticulaires associés sont en condition de diffraction) et le vecteur \mathbf{S} reliant M et le nœud situé sur la sphère représente le rayon réfléchi. On remarque en effet que la relation de Bragg est satisfaite pour un nœud hkl sur la sphère car alors :

$$2 \sin \theta / \lambda = T_{hkl}^*, \text{ soit } 2d_{hkl} \sin \theta = \lambda \quad (1.1.9)$$

Selon la loi de Bragg, il faut orienter les plans réticulaires pour réfléchir le faisceau X incident : la réflexion est sélective. Il faut donc placer judicieusement le cristal par rapport au faisceau X pour obtenir une réflexion particulière (un cristal orienté de manière aléatoire dans un faisceau X ne produit pas nécessairement un rayon diffracté). Cette rotation se retrouve sur la construction d'Ewald. En faisant tourner une famille de plans réticulaires (plans $(2\bar{1}0)$ sur la figure 1.2.6) pour amener ces plans en condition de diffraction, on fait aussi tourner le réseau réciproque et le nœud $2\bar{1}0$ vient se placer sur la sphère d'Ewald. Inversement, en amenant le nœud $2\bar{1}0$ sur la sphère d'Ewald, on oriente conjointement les plans réticulaires $(2\bar{1}0)$ en condition de diffraction.

La figure 1.2.7 montre ce qu'on obtient sur un détecteur bidimensionnel lorsqu'on fait tourner un cristal de protéine d'environ 1° . De très nombreux nœuds de l'espace réciproque traversent la sphère d'Ewald pendant la rotation et entrent en condition de diffraction car les paramètres du réseau réciproque sont très petits (ceux du réseau direct sont très grands) et les nœuds très proches les uns des autres.

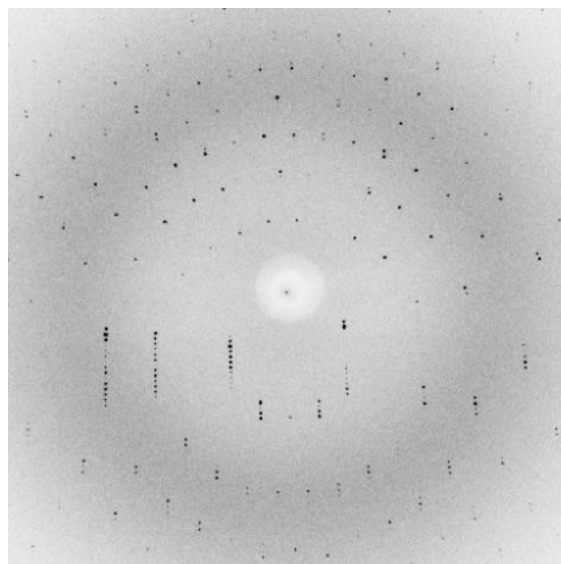


Fig. 1.2.7 – Cliché de diffraction pour un protéine (rotation d'environ 1°)

Notez au passage que tous les nœuds du réseau réciproque ne peuvent pas être amenés sur la sphère d'Ewald qui est de dimension finie. Pour des nœuds qui seraient trop éloignés, pour une taille de sphère donnée, pour être amenés en condition de diffraction, il faut changer (réduire) la longueur d'onde pour augmenter le rayon de la sphère. La sphère contenant tous les nœuds qui peuvent être amenés en condition de diffraction pour une longueur d'onde donnée s'appelle la *sphère de diffraction*.

1.1.7 – Détermination des paramètres de maille

Puisque l'espacement des nœuds du réseau réciproque détermine les angles de diffraction, la position des « taches » sur le détecteur est directement reliée aux paramètres du réseau réciproque (la relation exacte dépend de la géométrie utilisée pour enregistrer les réflexions : distance du détecteur, mouvement éventuel du détecteur, etc.). Les paramètres du réseau réciproque sont à leur tour reliés à ceux du réseau direct. Il est donc possible de déterminer les paramètres de maille directement à partir des positions des réflexions mesurées sur le détecteur. Pour les protéines, un seul cliché comme celui de la figure 1.2.7 permet de trouver les paramètres de maille. Pour les petites mailles ce n'est pas le cas car il y a très peu de « taches » sur chaque cliché, il est alors nécessaire d'enregistrer plusieurs clichés correspondant à plusieurs orientations du cristal.

1.2 – Production des rayons X

Pour enregistrer un très grand nombre de réflexions le plus rapidement possible et avec le plus de précision possible, il est important de disposer de sources de rayon X très intenses. Dans un laboratoire de recherche classique, il est possible d'utiliser des sources dites conventionnelles de faible ou moyenne intensité. Pour pouvoir disposer de sources beaucoup plus puissantes, il faut se déplacer sur un "grand instrument" : le synchrotron.

1.2.1 – Les sources conventionnelles

Dans un tube à rayons X conventionnel, les rayons X sont produits en bombardant une cible métallique (en général de cuivre ou de molybdène) par des électrons, émis par un filament chauffé et accélérés dans un champ électrique (voir figure 1.2.8). Les électrons entrent en collision avec les électrons des couches profondes du métal et les éjectent de leurs orbitales. Des électrons de niveaux de plus hautes énergies viennent alors combler les lacunes en émettant des photons X (transition électronique entre deux niveaux d'énergie).

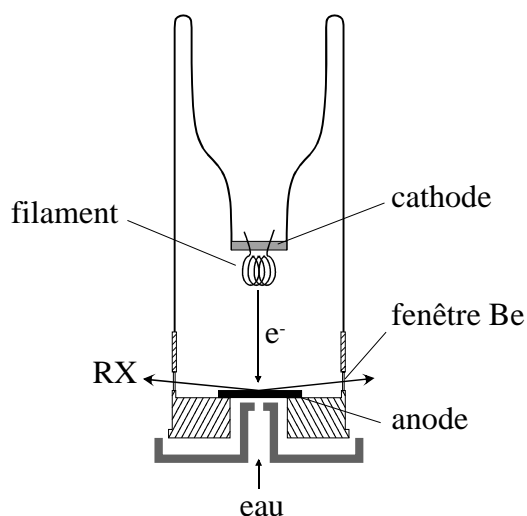


Fig. 1.2.8 – Schéma d'un tube à rayons X

Les problèmes liés aux sources conventionnelles sont doubles : 1) Les tubes dégagent une importante quantité de chaleur qu'il faut évacuer par une circulation d'eau (ce qui limite la puissance maximale délivrée par les tubes) ; 2) les rayons X émis ont plusieurs longueurs d'onde caractéristiques (correspondant à plusieurs transitions dans le métal), sans compter un rayonnement moins intense mais continu en longueur d'onde (rayonnement blanc lié au freinage des électrons dans le métal, voir figure 1.2.9). Il est alors nécessaire d'utiliser des filtres ou des monochromateurs pour sélectionner une seule longueur d'onde.

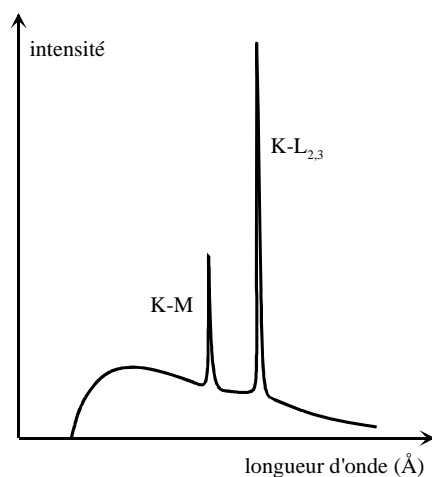


Fig. 1.2.9 – Caractéristique du spectre X pour un tube au cuivre

1.2.2 – Le rayonnement synchrotron

Les rayons X, comme d'autres radiations électromagnétiques, peuvent aussi être générés par des sources synchrotron. Le rayonnement synchrotron est produit par des électrons ou positons relativistes accélérés dans des anneaux de stockage de grande taille (jusqu'à plusieurs centaines de mètres). Ce rayonnement couvre continûment tout le spectre, de l'infra rouge aux rayons X, ce qui permet de choisir la longueur d'onde de travail, et offre une brillance environ mille milliards de fois plus grande que celle des tubes conventionnels à rayons X ! En France, la source synchrotron la plus intense se situe à Grenoble (European Synchrotron Radiation Facility, figure 1.2.10).



Fig. 1.2.10 – ESRF (European Synchrotron Radiation Facility), Grenoble

1.3 – Détection des rayons X

Les intensités des réflexions sont mesurées avec des compteurs à scintillation, qui comptent les photons X et ainsi donnent une intensité assez précise sur un large domaine. De façon simplifiée, un compteur à scintillation est constitué d'un matériau qui produit un photon lumineux lorsqu'il reçoit un photon X (scintillation). Le signal lumineux est alors amplifié et mesuré.

Il existe des compteurs ponctuels qui permettent de mesurer individuellement l'intensité de chaque réflexion. Ces compteurs ont été largement supplantés par des détecteurs bidimensionnels qui couvrent un assez large domaine angulaire et permettent de mesurer simultanément un très grand nombre de réflexions. Parmi ces détecteurs on trouve les films photographiques, mais ils ont presque disparus et ont été remplacés par les "plaques images" et plus récemment les caméras CCD.

1.4 – Les diffractomètres

Entre la source de rayons X et le détecteur se tient un appareil qui permet de précisément orienter le cristal pour l'amener en condition de diffraction. Cet appareil peut prendre plusieurs formes. Nous ne mentionnerons ici que le diffractomètre à 4 cercles (figure 1.2.11), qui est le plus pratique d'utilisation. Quelque soit l'appareil utilisé, le cristal est toujours préalablement monté sur une tête goniométrique qui permet au cristallographe de positionner le cristal dans une position particulière.

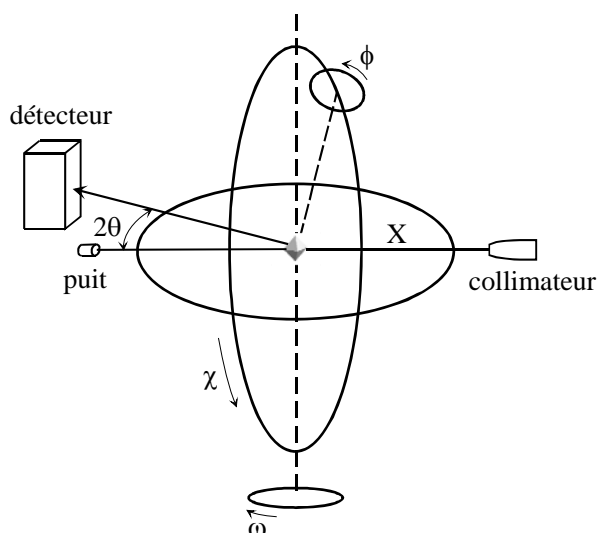


Fig. 1.2.11 – Schéma d'un diffractomètre à 4 cercles

Dans un diffractomètre, la tête goniométrique et le cristal sont placés dans un système de cercles motorisés, appelé goniomètre, qui permet de positionner le cristal

dans a priori n'importe quelle position par rapport au faisceau X incident et au détecteur (voir figures 1.2.11 à 1.2.13).



Fig. 1.2.12 – Collimateur, tête goniométrique et puit

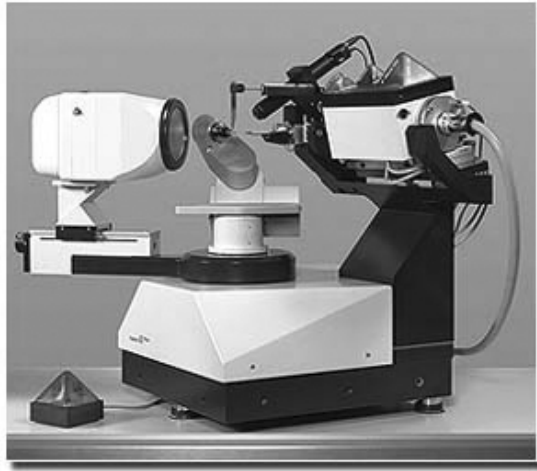


Fig. 1.2.13 – Diffractomètre Nonius équipé d'une caméra CCD (à gauche)

Chapitre 2

Détermination structurale

Introduction

Dans le chapitre 1 nous avons tiré avantage des clichés de diffraction pour remonter à la métrique du réseau cristallin. Pour cela nous avons exploité les positions des réflexions sur ces clichés. Nous allons maintenant exploiter l'autre information présente sur les clichés, à savoir les intensités des réflexions, pour déterminer le contenu du motif.

2.1 – La transformation de Fourier

Nous avons vu dans l'introduction générale du cours qu'il n'était pas possible d'avoir des lentilles pour les rayons X et qu'il était nécessaire de simuler par calcul ces lentilles pour reconstruire, à partir d'une image de l'espace réciproque (les réflexions enregistrées avec le détecteur), une image de l'espace direct (la structure cristalline). La transformation s'appelle une **transformation de Fourier**. Elle permet le passage de l'espace direct à l'espace réciproque, c'est à dire d'un objet réel à l'image de diffraction, et réciproquement, c'est à dire de l'image de diffraction à l'objet réel. Etudions qualitativement cette transformation de Fourier à partir de la décomposition en série de Fourier d'un objet fictif à une seule dimension.

Soit une structure hypothétique constituée de trois atomes par maille : deux atomes de carbone et un atome de fluor (voir figure 1.1.1). C'est un objet de l'espace direct. Essayons de modéliser la densité électronique de cet objet par une série d'ondes simples (de type sinus ou cosinus). On réalise qu'il y a en gros deux massifs : un premier massif à gauche associé aux deux atomes de carbone et un massif à droite correspondant à l'atome de fluor. Une simulation grossière de la structure est alors obtenue par une onde de fréquence 2, avec une amplitude ajustée à la densité électronique des atomes et une phase qui assure que les maxima de l'onde coïncident avec les maxima de densité électronique (voir figure 1.1.2).

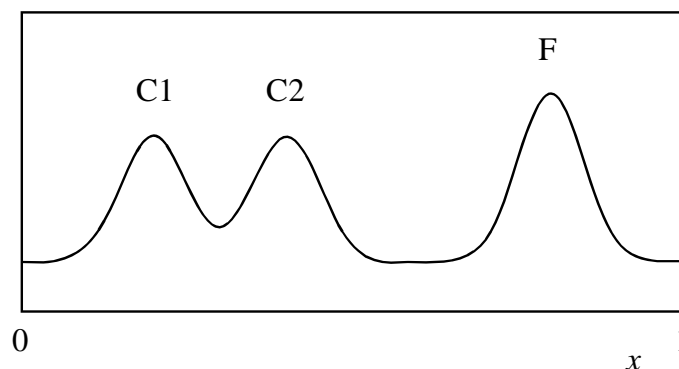


Fig. 2.1.1 – Maille d'une structure hypothétique 1D

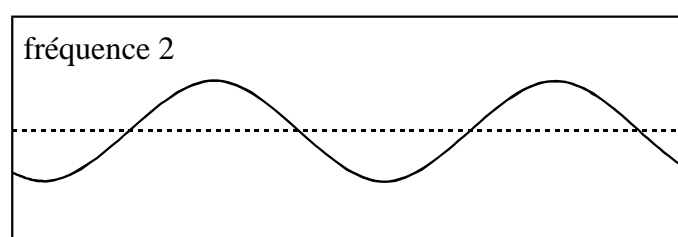


Fig. 2.1.2 – Simulation grossière de la densité électronique avec une onde de fréquence 2

Affinons la simulation en ajoutant des ondes de fréquences supérieures, ajustées en amplitudes et en phases (voir figure 2.1.3). On remarque qu'en ajoutant deux ondes de fréquences 3 et 5, on obtient une assez bonne modélisation des trois atomes. Il reste cependant quelques imperfections, en particulier entre les atomes. Nous venons de réaliser une décomposition en série de Fourier, partielle car il manque des termes.

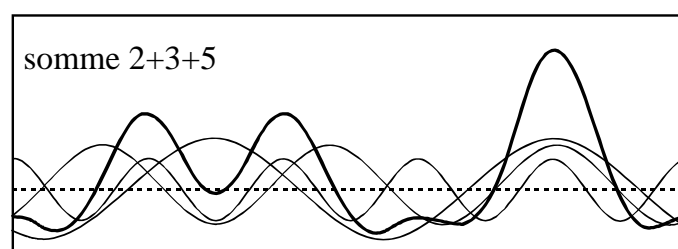


Fig. 2.1.3 – Simulation de la densité électronique de trois atomes par une somme (trait épais) de trois ondes (traits fins) de fréquences 2, 3 et 5.

Mathématiquement, la sommation (série) de Fourier s'écrit :

$$f(x) = \sum_{h=0}^n F_h \cos(2\pi(hx + \alpha_h)) \quad (2.1.1)$$

La théorie de Fourier indique que toute fonction périodique, même très compliquée, peut être approximée par cette série, en choisissant les bonnes valeurs pour les fréquences h , les amplitudes F_h et les phases α_h .

Une forme de fonction d'onde de base très utile en cristallographie est la suivante :

$$f(x) = [\cos 2\pi(hx) + i \sin 2\pi(hx)]$$

C'est une combinaison d'un cosinus et d'un sinus pour donner un nombre complexe (notez le $i = \sqrt{-1}$ des nombres complexes). Bien que la phase n'apparaisse pas dans cette expression, elle est implicite et dépend de h et x comme nous le verrons un peu plus loin. Avec cette forme, la série de Fourier s'écrit :

$$f(x) = \sum_{h=0}^n F_h [\cos(2\pi hx) + i \sin(2\pi hx)] \quad (2.1.2)$$

En utilisant la notation exponentielle des nombres complexes et en omettant les limites de la sommation, on obtient :

$$f(x) = \sum_h F_h \exp 2\pi i(hx) \quad (2.1.3)$$

Pour une fonction des trois coordonnées de l'espace (x , y et z), la série de Fourier implique trois sommations, une par fréquence dans chacune des directions de l'espace :

$$f(x, y, z) = \sum_h \sum_k \sum_l F_{hkl} \exp 2\pi i(hx + ky + lz) \quad (2.1.4)$$

Joseph Fourier (mathématicien et physicien français) a démontré que pour toute fonction $f(x)$ il existait une fonction $F(h)$ telle que :

$$F(h) = \int_{-\infty}^{+\infty} f(x) \exp 2\pi i(hx) dx \quad (2.1.5)$$

où $F(h)$ est appelée la transformée de Fourier (TF) de $f(x)$ et où les dimensions de h sont inverses de celles de x . Dans le cas tridimensionnel, la TF s'écrit :

$$F(h, k, l) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y, z) \exp 2\pi i(hx + ky + lz) dx dy dz \quad (2.1.6)$$

Cette transformation de Fourier est naturellement très adaptée pour relier l'espace direct et l'espace réciproque qui ont des dimensions inverses. En calculant la TF de la structure fictive précédente (figure 2.1.1) on obtient les résultats présentés sur la figure 2.1.4.

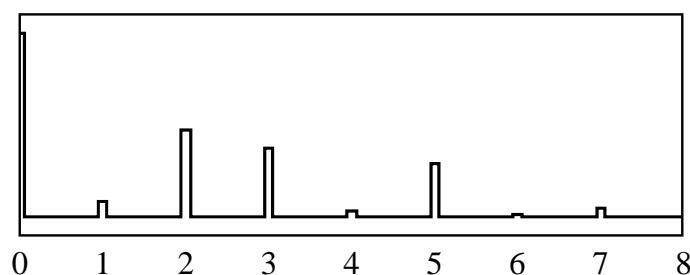


Fig. 2.1.4 – Transformée de Fourier de la structure hypothétique de la figure 2.1.1

On remarque que la TF est constituée d'une série de pics, dont les plus intenses sont ceux de fréquences 2, 3 et 5 (ce que nous avons précédemment trouvé dans notre simulation). Notez que toutes les informations ne sont pas présentes sur cette figure, en particulier les phases (car la TF est une fonction complexe) et les fréquences négatives.

2.2 – L'intensité diffractée

2.2.1 – Le facteur de forme atomique et le facteur de structure

La transformée de Fourier de la densité électronique $\rho_j(\mathbf{r})$ d'un atome j s'appelle le *facteur de forme atomique*, $f_j(\mathbf{s})$:

$$f_j(\mathbf{s}) = \int \rho_j(\mathbf{r}) \exp 2\pi i(\mathbf{s} \cdot \mathbf{r}) dV_r \quad (2.2.1)$$

Dans cette écriture, nous avons utilisé la notation vectorielle pour simplifier l'écriture dans le cas d'un espace tridimensionnel ($\mathbf{r} = (x, y, z)$ pour l'espace direct et $\mathbf{s} = (h, k, l)$ pour l'espace réciproque). Le facteur de forme atomique de l'atome j , $f_j(\mathbf{s})$, est une fonction continue sur l'espace réciproque (à ce stade h , k et l sont donc réels). $f_j(\mathbf{s})$ est en unité d'électrons. Sur les figures 2.2.1 et 2.2.2 sont présentées respectivement la densité d'un atome $\rho_j(\mathbf{r})$ et sa transformation de Fourier $f_j(\mathbf{s})$.

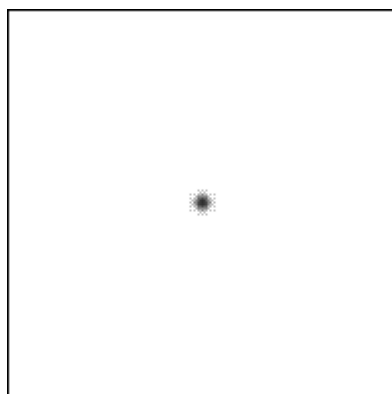


Fig. 2.2.1 – Densité électronique d'un atome

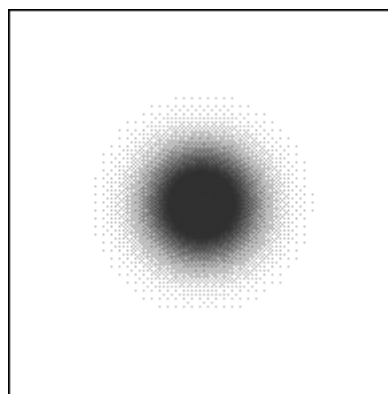


Fig. 2.2.2 – TF de la densité électronique d'un atome = facteur de forme atomique

Notez que le facteur de forme atomique décroît en fonction de l'angle comme indiqué sur la figure 1.2.3. Le maximum à $\theta = 0$ est égal au numéro atomique Z de l'atome considéré, c'est à dire au nombre d'électrons diffusants. Cette décroissance en fonction de l'angle se retrouve sur les clichés de diffraction (voir figure 1.2.7, Ch1 ou figure 2.2.6 ci-après).

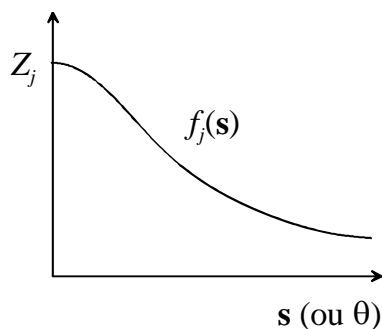


Fig. 2.2.3 – Facteur de forme atomique

Pour un ensemble de n atomes (motif) repérés par les coordonnées \mathbf{r}_j , la densité électronique totale $\rho(\mathbf{r})$ est la somme des densités électroniques centrées sur les atomes. La transformée de Fourier de cette densité électronique totale se réduit alors à une série de Fourier selon :

$$F(\mathbf{s}) = \int_{\text{maille}} \rho(\mathbf{r}) \exp 2\pi i(\mathbf{s} \cdot \mathbf{r}) d\mathbf{r} = \sum_{j=1}^n f_j(\mathbf{s}) \exp 2\pi i(\mathbf{s} \cdot \mathbf{r}_j) \quad (2.2.2)$$

Chaque terme de la somme de la série (membre de droite) correspond à la contribution de l'atome j à la diffusion. Une illustration de cette transformation de Fourier[†] est donnée sur les figures 2.2.4 et 2.2.5.

Pour un cristal, l'introduction de la périodicité conduit simplement à limiter la diffusion à certaines directions de l'espace (diffusion \rightarrow diffraction), comme nous l'avons vu dans le chapitre précédent et comme l'illustrent les figures 2.2.6 et 2.2.7. La transformée de Fourier précédente devient alors discrète et s'écrit :

$$F_{hkl} = \sum_{j=1}^n f_j \exp 2\pi i(hx_j + ky_j + lz_j) \quad (2.2.3)$$

F_{hkl} s'appelle le **facteur de structure**. Chaque onde diffractée résulte des contributions des atomes de la maille. F_{000} est égale au nombre d'électrons du motif.

[†] Une TF est une fonction complexe avec une amplitude et une phase pour chaque point de l'espace. Une représentation correcte de cette fonction nécessite donc deux paramètres, par exemple une intensité lumineuse pour l'amplitude et une couleur pour la phase. Ceci n'est pas possible sur un polycopié noir et blanc. Ceci n'est pas très important car les images de l'espace réciproque sont difficilement visuellement interprétables !

Notez que la contribution au facteur de structure est maximale lorsque le produit $(hx_j + ky_j + lz_j)$ est égal à 1 (ou 0), c'est à dire lorsque l'atome est sur le plan réticulaire (hkl) . Par exemple un atome en $(2/3, 1/3, 0)$ a une contribution maximale au facteur de F_{210} car il est situé sur le plan réticulaire (210) .

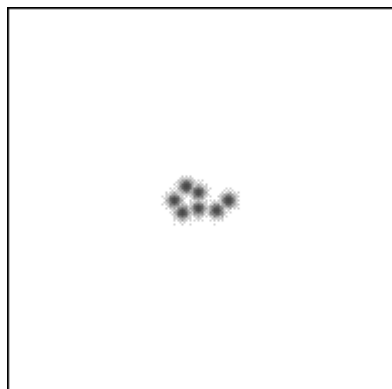


Fig. 2.2.4 – Densité électronique d'un motif

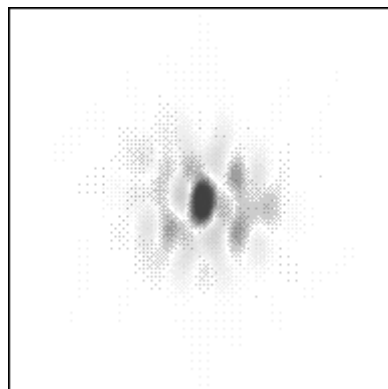


Fig. 2.2.5 – TF de la densité électronique du motif de la figure 2.2.4

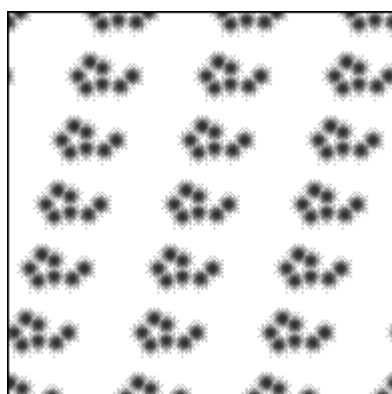


Fig. 2.2.6 – Densité électronique d'un cristal

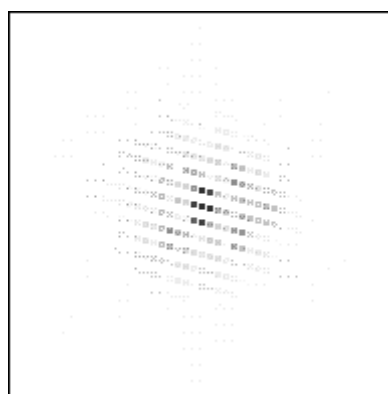


Fig. 2.2.7 – TF de la densité électronique du cristal de la figure 2.2.6

On montre que la transformée de Fourier d'un cristal ($TF(\mathcal{C})$) est égale à la transformée de Fourier du motif ($TF(\mathcal{M})$), que multiplie la transformée de Fourier du réseau direct ($TF(\mathcal{R})$), c'est à dire le réseau réciproque (\mathcal{R}^*) :

$$TF(\mathcal{C}) = TF(\mathcal{M}) \times TF(\mathcal{R}) = TF(\mathcal{M}) \times \mathcal{R}^* \quad (2.2.4)$$

Le réseau direct du cristal de la figure 2.2.6 est présenté sur la figure 2.2.8, sa transformée de Fourier est donnée sur la figure 2.2.9. C'est tout simplement le réseau réciproque du réseau direct. Notez que la transformée de Fourier du cristal présentée figure 2.2.7 est bien le produit des contenus des figures 2.2.5 et 2.2.9.

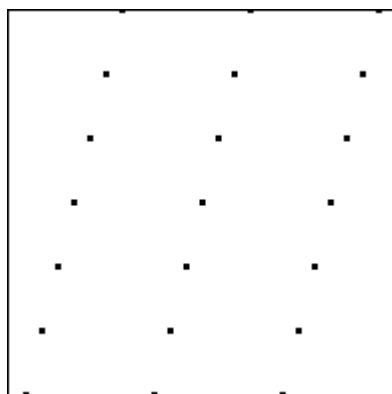


Fig. 2.2.8 – Réseau direct du cristal présenté figure 2.2.6

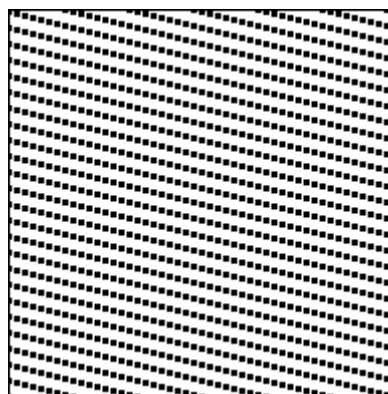


Fig. 2.2.9 – Réseau réciproque du cristal présenté figure 2.2.6

2.2.2 – L'intensité diffractée et la loi de Friedel

L'intensité diffractée (celle que l'on mesure sur un détecteur) est égale à :

$$I_{hkl} \propto |F_{hkl}|^2 = F_{hkl} \times F_{hkl}^* \quad (2.2.5)$$

Le coefficient de proportionnalité dépend de l'intensité de la source, du temps d'exposition, de la grosseur du cristal, etc. L'intensité diffractée pour une réflexion hkl nous donne donc accès à l'amplitude du facteur de structure (à un coefficient près identique pour toutes les réflexions) par l'intermédiaire de la relation :

$$|F_{hkl}| \propto \sqrt{I_{hkl}} \quad (2.2.6)$$

mais elle ne nous donne pas d'information sur la phase du facteur de structure !

Nous avons vu dans la première partie du cours que les plans réticulaires (hkl) étaient identiques aux plans ($\bar{h}\bar{k}\bar{l}$). L'intensité de la réflexion hkl est donc égale à l'intensité de la réflexion $\bar{h}\bar{k}\bar{l}$, c'est la **loi de Friedel** :

$$I_{hkl} = I_{\bar{h}\bar{k}\bar{l}} \quad (2.2.7)$$

En fait les réflexions hkl et $\bar{h}\bar{k}\bar{l}$ sont deux réflexions par une même famille de plans, mais d'un côté et de l'autre par rapport à ces plans, ce qui explique pourquoi les intensités sont identiques (nous modulerons ce propos un peu plus loin).

2.2.3 - Extinctions systématiques

Il est très facile de montrer que si certaines opérations de symétrie du groupe d'espace d'un cristal comportent des translations non primitives (c'est-à-dire autres que les translations entières de vecteurs de maille), les intensités de certaines réflexions hkl sont nulles. Ceci constitue les **extinctions systématiques**, qui permettent de déduire la

présence de certains éléments de symétrie à partir des clichés de diffraction. Illustrons ceci avec deux exemples.

Considérons en premier lieu le cas d'une maille centrée de type C. A chaque atome m de coordonnées (x_m, y_m, z_m) correspond un atome m' de coordonnées $(x_m+1/2, y_m+1/2, z_m)$, donc :

$$\begin{aligned} F_{hkl} &= \sum_{\substack{\text{atome } m \\ 1/2 \text{ maille}}}^{1/2 \text{ maille}} f_m \{ \exp 2\pi i (hx_m + ky_m + lz_m) + \exp 2\pi i (h(x_m + \frac{1}{2}) + k(y_m + \frac{1}{2}) + lz_m) \} \\ &= \sum_{\substack{\text{atome } m \\ 1/2 \text{ maille}}}^{1/2 \text{ maille}} f_m \exp 2\pi i (hx_m + ky_m + lz_m) \{ 1 + \exp 2\pi i (\frac{h}{2} + \frac{k}{2}) \} \\ &= \sum_{\substack{\text{atome } m \\ 1/2 \text{ maille}}}^{1/2 \text{ maille}} f_m \exp 2\pi i (hx_m + ky_m + lz_m) \{ 1 + \cos(\pi(h+k)) + i \sin(\pi(h+k)) \} \end{aligned} \quad (2.2.8)$$

On remarque que si $h+k=2n+1$, alors $\cos(\pi(h+k)) = -1$ et $\sin(\pi(h+k)) = 0$, donc $F_{hkl} = 0$. Il n'y a pas d'intensité aux nœuds hkl du réseau réciproque tels que $h+k=2n+1$. C'est une extinction systématique liée au centrage du réseau. Les conditions imposées par les autres centrages (condition de présence) sont résumées sur le tableau 2.2.1.

Tableau 2.2.1 – Les extinctions systématiques liées au mode de réseau

Condition de présence sur hkl	Interprétation
$k + l = 2n$	mode de réseau A
$h + l = 2n$	mode de réseau B
$h + k = 2n$	mode de réseau C
$h + k + l = 2n$	mode de réseau I
h, k, l tous pairs ou tous impairs	mode de réseau F

Considérons maintenant une translation non primitive lié à un axe hélicoïdal 2_1 coïncidant avec l'axe a. A chaque atome m de coordonnées (x_m, y_m, z_m) correspond un atome m' de coordonnées $(x_m+1/2, -y_m, -z_m)$, donc :

$$F_{hkl} = \sum_{\substack{\text{atome } m \\ 1/2 \text{ maille}}}^{1/2 \text{ maille}} f_m \{ \exp 2\pi i (hx_m + ky_m + lz_m) + \exp 2\pi i (h(x_m + \frac{1}{2}) - ky_m - lz_m) \} \quad (2.2.9)$$

On remarque que pour les réflexions du type $h00$:

$$\begin{aligned} F_{h00} &= \sum_{\substack{\text{atome } m \\ 1/2 \text{ maille}}}^{1/2 \text{ maille}} f_m \{ \exp 2\pi i hx_m + \exp 2\pi i h(x_m + \frac{1}{2}) \} \\ &= \sum_{\substack{\text{atome } m \\ 1/2 \text{ maille}}}^{1/2 \text{ maille}} f_m \exp 2\pi i hx_m \{ 1 + \exp \pi i h \} \\ &= \sum_{\substack{\text{atome } m \\ 1/2 \text{ maille}}}^{1/2 \text{ maille}} f_m \exp 2\pi i hx_m \{ 1 + \cos \pi h + i \sin \pi h \} \end{aligned} \quad (2.2.10)$$

Si $h=2n+1$, alors $\cos(\pi h) = -1$ et $\sin(\pi h) = 0$, et donc $F_{hkl} = 0$. Il n'y a pas d'intensité aux nœuds $h00$ du réseau réciproque tels que $h=2n+1$. C'est une extinction systématique liée à la translation $1/2$ de l'axe hélicoïdal 2_1 . Il existe une condition d'extinction de ce type pour chaque opération de symétrie impliquant une translation non primitive. Ces extinctions sont répertoriées dans les Tables Internationales de Cristallographie.

2.3 – Le problème des phases

Calculons la transformée de Fourier d'un objet particulier : un chat (figures 2.3.1 et 2.3.2). Notez que le calcul, contrairement à une expérience de diffraction, donne toute l'information : amplitude et phase. Opérons la même transformation sur un autre objet : un canard (figures 2.3.3 et 2.3.4).



Fig. 2.3.1 – Chat dans l'espace direct

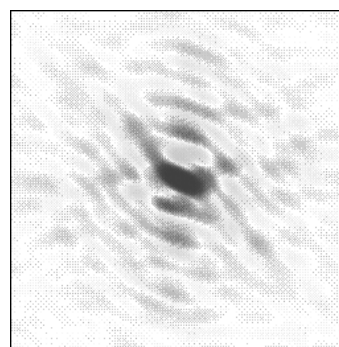


Fig. 2.3.2 – Transformée de Fourier du chat

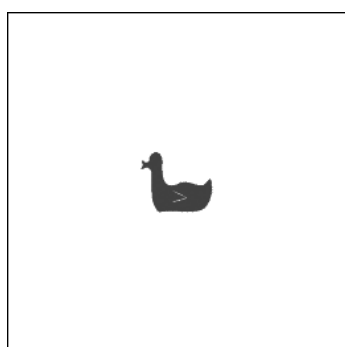


Fig. 2.3.3 – Canard dans l'espace direct

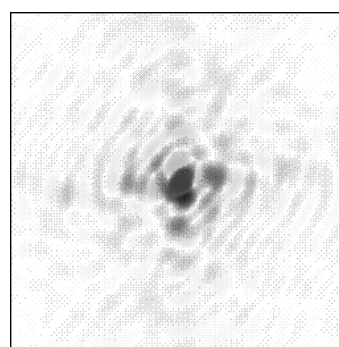


Fig. 2.3.4 – Transformée de Fourier du canard

En opérant des transformations de Fourier inverses à partir des images 2.3.2 ou 2.3.4, on retrouve les images réelles 2.3.1 et 2.3.3. Opérons une opération très particulière. Combinons, dans l'espace réciproque, les amplitudes de la transformée de Fourier du canard avec les phases de la transformée de Fourier du chat et opérons une

transformation de Fourier sur ce "mélange", figure 2.3.5. Qu'obtient-on ? La réponse est donnée sur la figure 2.3.6. Le résultat peut paraître surprenant : le mélange ne donne pas naissance à un animal hybride, mais bel et bien aux contours bien visibles d'un chat. Ce sont les phases qui sont primordiales pour retrouver l'objet de l'espace direct, et pas les amplitudes.

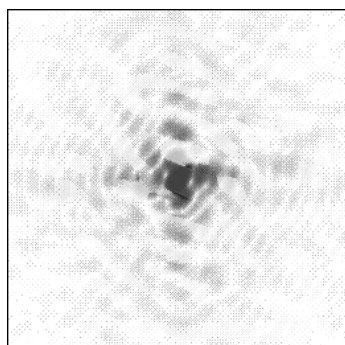


Fig. 2.3.5 – "Mélange" phases/amplitudes dans l'espace réciproque

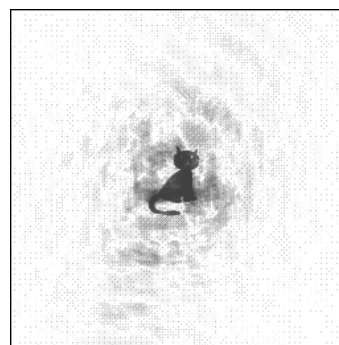


Fig. 2.3.6 – TF inverse du mélange (espace direct)

Rappelons que chaque facteur de structure, correspondant à l'onde diffractée dans une direction particulière de l'espace, est une fonction périodique avec une amplitude et une phase. L'amplitude est proportionnelle à l'intensité du faisceau diffracté, elle est donc mesurable. Seule la phase ne peut pas être obtenue, et nous venons de montrer que la phase est le facteur prépondérant. C'est le **problème des phases**, déjà évoqué, que nous allons maintenant essayer de contourner (avec succès).

2.4 – Obtenir les phases

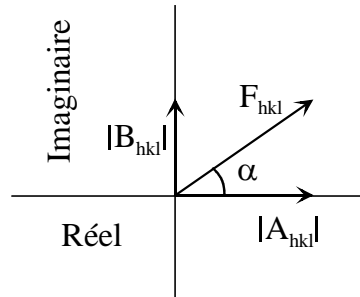
2.4.1 – Représentation graphique des facteurs de structure

Comme nous venons de le rappeler, le facteur de structure F_{hkl} donne une description complète d'une réflexion (onde diffractée), avec son amplitude et sa phase. Pour simplifier les discussions qui vont suivre sur l'obtention des phases, il est pratique de représenter les facteurs de structure sous la forme de vecteurs dans le plan des nombres complexes, comme indiqué sur la figure 2.4.1. \mathbf{F}_{hkl} (notez le changement d'écriture pour désigner un vecteur) s'écrit sous la forme :

$$\mathbf{F}_{hkl} = |\mathbf{A}_{hkl}| + i|\mathbf{B}_{hkl}| = |\mathbf{F}_{hkl}|(\cos \alpha + i \sin \alpha) = |\mathbf{F}_{hkl}| \exp i\alpha_{hkl} \quad (2.4.1)$$

avec :

$$\begin{aligned} \mathbf{A}_{hkl} &= \sum_{j=1}^n f_j \cos 2\pi(hx_j + ky_j + lz_j) \\ \mathbf{B}_{hkl} &= \sum_{j=1}^n f_j \sin 2\pi(hx_j + ky_j + lz_j) \end{aligned} \quad (2.4.2)$$

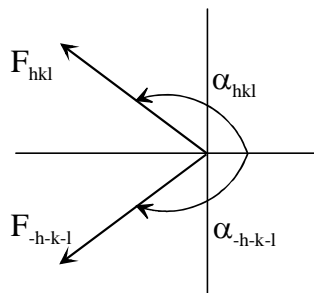

 Fig 2.4.1 – Représentation vectorielle de \mathbf{F} dans le plan complexe

Nous avons les relations :

$$\begin{aligned} \tan \alpha_{hkl} &= \frac{|\mathbf{B}_{hkl}|}{|\mathbf{A}_{hkl}|}, \quad |\mathbf{F}_{hkl}| = \sqrt{|\mathbf{A}_{hkl}|^2 + |\mathbf{B}_{hkl}|^2} \\ |\mathbf{A}_{hkl}| &= |\mathbf{F}_{hkl}| \cos \alpha_{hkl}, \quad |\mathbf{B}_{hkl}| = |\mathbf{F}_{hkl}| \sin \alpha_{hkl} \end{aligned} \quad (2.4.3)$$

où nous retrouvons explicitement la phase α_{hkl} , que nous avons indiqué (§2.1.1) apparaître de façon implicite dans l'expression de \mathbf{F} .

Nous avons énoncé la loi de Friedel au §2.2.2. Il sera utile un peu plus loin dans ce chapitre d'utiliser les représentations, données figure 2.4.2, des facteurs de structure \mathbf{F}_{hkl} et $\mathbf{F}_{\bar{h}\bar{k}\bar{l}}$, appelés *paires de Friedel*. Notez que même si les intensités I_{hkl} et $I_{\bar{h}\bar{k}\bar{l}}$ sont égales, les facteurs de structure ne le sont pas. Les facteurs de structures d'une paire de Friedel ont des phases opposées, $\mathbf{F}_{\bar{h}\bar{k}\bar{l}}$ est l'image miroir de \mathbf{F}_{hkl} .


 Fig. 2.4.2 – Représentation vectorielle de la paire de Friedel \mathbf{F}_{hkl} et \mathbf{F}_{-h-k-l}

La représentation vectorielle sera aussi utilisée dans les paragraphes suivants pour combiner des facteurs de structure. Ainsi par exemple, la somme de trois facteurs de structure \mathbf{f}_1 , \mathbf{f}_2 et \mathbf{f}_3 est facilement obtenue en réalisant une somme vectorielle, comme indiqué sur la figure 2.4.3.

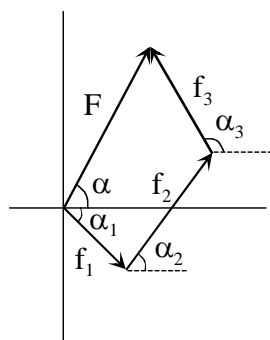


Fig. 2.4.3 – Combinaison de facteurs de structure

2.4.2 – La densité électronique du cristal

Nous avons vu au §2.2.1 que le facteur de structure était la transformée de Fourier de la densité électronique du contenu de la maille :

$$\mathbf{F}(\mathbf{s}) = \int_{\text{maille}} \rho(\mathbf{r}) \exp 2\pi i(\mathbf{s} \cdot \mathbf{r}) d\mathbf{r} \quad (2.4.4)$$

ou, de façon équivalente :

$$\mathbf{F}_{hkl} = \iiint \rho(x, y, z) \exp 2\pi i(hx + ky + lz) dx dy dz \quad (2.4.5)$$

Pour obtenir la densité électronique de la maille, et donc la structure cristalline recherchée, il est nécessaire d'effectuer une transformation de Fourier inverse des facteurs de structure :

$$\begin{aligned} \rho(x, y, z) &= \frac{1}{V} \sum_h \sum_k \sum_l \mathbf{F}_{hkl} \exp\{-2\pi i(hx + ky + lz)\} \\ &= \frac{1}{V} \sum_{hkl} |\mathbf{F}_{hkl}| \exp\{-2\pi i(hx + ky + lz) + i\alpha_{hkl}\} \end{aligned} \quad (2.4.6)$$

où V est le volume de la maille. Le signe $-$ dans l'exponentielle indique le fait que nous avons à faire à une transformation de Fourier inverse. Noter le passage d'une intégrale triple (sur la maille) à une triple sommation (sur hkl , en notation abrégée dans la deuxième formulation) du fait du caractère discret des facteurs de structure. Dans la série, chaque terme est une onde tridimensionnelle d'amplitude $|\mathbf{F}_{hkl}|$, de fréquence hkl (h le long de a , k le long de b et l le long de c) et de phase α_{hkl} . Lorsque nous aurons les phases, ce "simple" calcul mathématique nous donnera la densité électronique de la maille et donc la structure recherchée (chaque atome correspondant à un maximum de la densité électronique).

Il existe plusieurs méthodes qui permettent d'obtenir une estimation des phases. Cependant la résolution d'une structure de protéine ne peut pas être réalisée comme une

résolution de structure de matériaux inorganiques, ou organiques avec de petites molécules. En effet, à cause du trop grand nombre d'atomes, les méthodes usuelles d'obtention des phases ne donnent pas les résultats escomptés. Il faut utiliser des techniques spécifiques.

2.4.3 – La méthode de l'atome lourd (remplacement isomorphe)

Pour pouvoir résoudre une structure de protéine, le cristallographe doit préparer un ou plusieurs dérivés d'atomes lourds en plongeant les cristaux de la protéine dans des solutions d'atomes lourds (ions ou complexes ioniques de Hg, Pt ou Au en général). Dans la plupart des cas, les ions lourds se lient à la protéine sans modifier sa conformation ou l'arrangement cristallin. Si c'est effectivement le cas, il peut espérer remonter aux phases, et donc à la structure de la protéine.

Nous avons vu que chaque atome de la maille contribue à chaque réflexion hkl d'un cliché de diffraction (si la réflexion n'est pas éteinte à cause de la symétrie). Nous avons aussi vu §2.2.1 que la contribution à l'intensité dépend du nombre d'électrons de l'atome : plus l'atome est lourd (= Z élevé) et plus la contribution est importante, et de la position de l'atome vis à vis de la famille de plans (hkl) en condition de diffraction : plus l'atome est proche d'un plan de la famille et plus la contribution est importante (elle est maximale lorsque l'atome est sur un plan). Un changement même minime de la structure au niveau d'un atome ou d'un groupe d'atomes occasionne donc une perturbation du cliché de diffraction. Cette perturbation peut être utilisée pour remonter aux phases.

Considérons deux jeux de données, obtenus pour deux cristaux : 1) un cristal natif avec seulement la protéine (P) et 2) un cristal dérivé comprenant la protéine et un atome lourd (PL). Pour une réflexion hkl particulière, l'amplitude mesurée est $|\mathbf{F}_P|$ pour le cristal natif et $|\mathbf{F}_{PL}|$ pour le cristal substitué. Comme les contributions des atomes à une réflexion sont additives, la différence d'amplitude ($|\mathbf{F}_{PL}| - |\mathbf{F}_P|$) est la contribution de l'atome lourd seul (si on admet que les phases α_{PL} et α_P sont voisines, ce qui n'est qu'une approximation assez grossière). Les intensités $(|\mathbf{F}_{PL}| - |\mathbf{F}_P|)^2$ correspondent alors aux intensités qui seraient obtenues si l'atome lourd était seul dans la maille. La structure hypothétique avec seulement cet atome lourd est suffisamment simple pour pouvoir être résolue par les méthodes classiques (voir ci-dessous). Si cette structure hypothétique est effectivement déterminée, tout est connu pour les facteurs de structure \mathbf{F}_L correspondants, y compris les phases.

Considérons à nouveau une réflexion hkl . La figure 2.4.4 montre la relation entre les facteurs de structure \mathbf{F}_L , \mathbf{F}_P et \mathbf{F}_{PL} . Comme les contributions sont additives, on a :

$$\mathbf{F}_{PL} = \mathbf{F}_L + \mathbf{F}_P \quad (2.4.7)$$

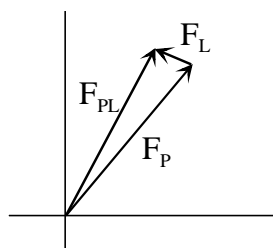
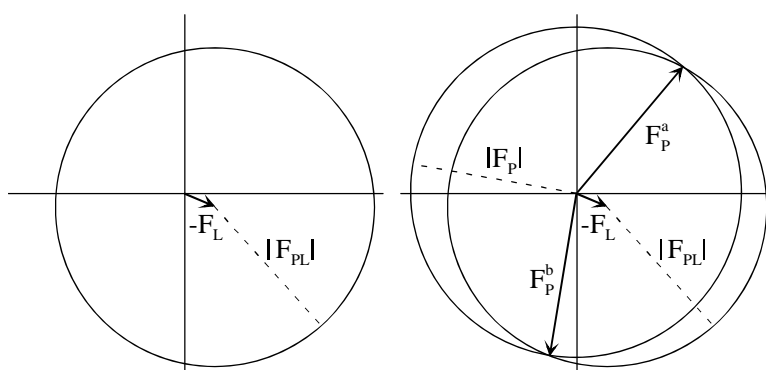


Fig. 2.4.4 – Le facteur de structure d'un dérivé d'atome lourd

Pour chaque réflexion hkl , on veut connaître \mathbf{F}_P . En fait, on connaît déjà les amplitudes $|\mathbf{F}_P|$ et $|\mathbf{F}_{PL}|$ par les mesures d'intensités pour les deux cristaux. Il nous manque les phases de \mathbf{F}_P et \mathbf{F}_{PL} , et donc les orientations dans l'espace des vecteurs \mathbf{F}_P et \mathbf{F}_{PL} . Pour résoudre le problème, observons la figure 2.4.5. Sur cette figure (partie gauche) est porté le vecteur $-\mathbf{F}_L$. A partir de l'extrémité de ce vecteur est dessiné un cercle de rayon $|\mathbf{F}_{PL}|$. Tous les points sur ce cercle correspondent à des solutions pour la différence $\mathbf{F}_P = \mathbf{F}_{PL} - \mathbf{F}_L$, ce qui veut dire que l'extrémité du vecteur \mathbf{F}_P se tient quelque part sur ce cercle. Comme on connaît le module de \mathbf{F}_P , un cercle de rayon $|\mathbf{F}_P|$ est dessiné à partir de l'origine (partie droite de la figure 2.4.5). L'intersection entre les deux cercles donne deux solutions, a et b, pour \mathbf{F}_P , et par conséquent deux phases possibles α_a et α_b pour \mathbf{F}_P .

Fig. 3.4.5 – Détermination de la phase de \mathbf{F}_P

Comment savoir quelle est la bonne phase ? Si les phases sont très différentes, il est nécessaire de réaliser une deuxième substitution (obtenir un second cristal substitué) qui va conduire à deux autres phases (α_c et α_d), la solution étant la phase commune entre les deux systèmes (ici $\alpha_a = \alpha_c$, voir figure 2.4.6). C'est la méthode du remplacement isomorphe multiple (MIR : multiple isomorphous replacement). Pour que la deuxième substitution apporte une information supplémentaire sur la phase, il est nécessaire que l'atome lourd se fixe sur une position différente de la première car deux substitutions avec des atomes différents mais à la même position conduiraient aux mêmes phases (seules les amplitudes seraient différentes, voir expression du facteur de structure).

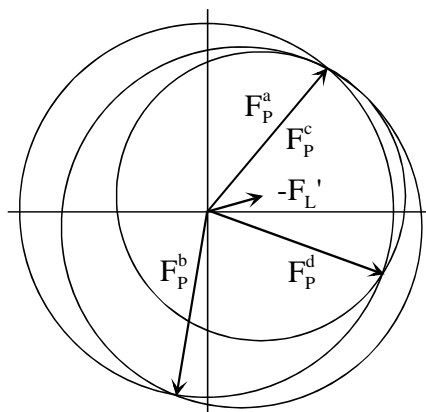


Fig. 2.4.6 – Détermination de la phase de F_P avec deux substitutions isomorphes (MIR)

Si par contre les deux phases sont voisines, la moyenne constitue une assez bonne approximation. C'est la méthode du remplacement isomorphe unique (SIR : single isomorphous replacement).

2.4.4 – La localisation de l'atome lourd (fonction de Patterson)

Une méthode pour déterminer la structure hypothétique simple ne contenant que l'atome lourd est l'utilisation de la fonction de Patterson. Cette technique est couramment utilisée pour les structures simples (et réelles) des composés inorganiques, en particulier ceux présentant quelques atomes beaucoup plus lourds que les autres.

La fonction de Patterson est une série de Fourier sans les phases. Elle a pour expression :

$$P(u, v, w) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}|^2 \exp\{-2\pi i(hu + kv + lw)\} \quad (2.4.8)$$

dans laquelle les coefficients sont les modules au carré des facteurs de structure. Pour obtenir la fonction de Patterson de la structure hypothétique avec l'atome lourd, il faut prendre les coefficients $\Delta F^2 = (|F_{PL}| - |F_P|)^2$. Comme la fonction de Patterson n'implique pas les phases, elle peut être directement calculée à partir des données enregistrées. Une carte de densité électronique (série de Fourier avec pour coefficient les facteurs de structure) révèle les positions atomiques, qui coïncident avec les maxima de la carte. Que révèle la fonction de Patterson ?

La fonction de Patterson présente des pics à des positions correspondant aux vecteurs[†] entre les atomes, la hauteur des pics étant proportionnelle au produit des numéros atomiques ($Z_j \times Z_k$). Les coordonnées d'un pic de Patterson correspondent au vecteur entre un atome 1 situé à (x_1, y_1, z_1) et un atome 2 situé à (x_2, y_2, z_2) , soit le vecteur

[†] en fait, les pics sont positionnés à la pointe des vecteurs, l'origine du vecteur coïncidant avec l'origine de la maille.

$(x_1-x_2, y_1-y_2, z_1-z_2)$. Il y a évidemment beaucoup plus de vecteurs interatomiques que d'atomes et la fonction de Patterson est en général très compliquée. C'est pourquoi elle n'est utilisée que dans les cas simples avec peu d'atomes, ou avec quelques atomes très lourds (Z élevés) par rapport aux autres.

Analysons à partir d'un exemple comment il est possible de déduire la position des atomes à partir des vecteurs entre ces atomes. Soit la structure hypothétique bidimensionnelle présentée sur la figure 2.4.7. Il existe six vecteurs interatomiques entre les trois atomes (sans compter les vecteurs nuls connectant les atomes avec eux-mêmes) qui peuvent être combinés avec une quelconque translation de maille. La fonction de Patterson de cette structure est donc celle présentée sur la figure 2.4.8.

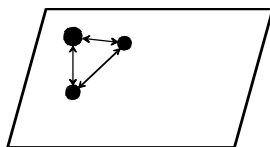


Fig. 2.4.7 – Structure hypothétique 2D (trois atomes représentés par des disques noirs)

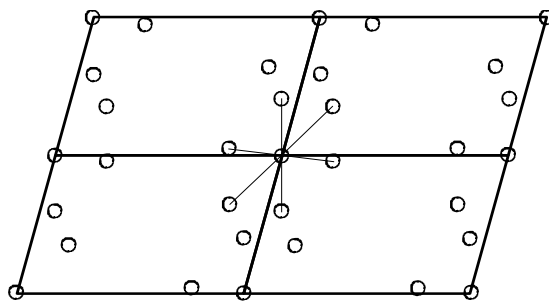


Fig. 2.4.8 – Fonction de Patterson de la structure hypothétique 2D de la figure 2.4.7 (pics de Patterson représentés par des disques blancs)

Partons maintenant de la carte de Patterson de la figure 2.4.8 pour remonter à la structure. Pour n atomes dans une maille, il y a $n(n-1)$ pics de Patterson, sans compter le pic origine. Comme il y a 6 pics de Patterson dans la maille, en dehors du pic origine, on en déduit qu'il y a 3 atomes dans la maille. Pour avoir les positions des atomes, au moins respectivement les uns par rapport aux autres, il nous faut utiliser deux vecteurs. Lesquels ? La démarche consiste à choisir deux vecteurs arbitraires, ce qui conduit à une structure d'essai, puis à calculer la fonction de Patterson à partir de cette structure d'essai. Si les deux fonctions de Patterson sont identiques, la structure d'essai est la bonne. Si elles sont différentes, la structure n'est pas la bonne et il faut choisir d'autres vecteurs. Il est clair que la complexité de l'interprétation de la fonction de Patterson augmente très rapidement avec le nombre d'atomes.

La symétrie du contenu de la maille peut simplifier la tâche (et positionner les atomes par rapport aux éléments de symétrie). Par exemple, la présence d'un axe hélicoïdal 2_1 selon b fait correspondre à un atome 1 en (x_1, y_1, z_1) un atome 2 en

$(-x_1, y_1+1/2, -z_1)$, et donc un vecteur interatomique $(2x, 1/2, 2z)$. La détection d'un tel pic dans la carte de Patterson indique la présence d'un atome sur un axe 2_1 .

2.4.5 – La méthode utilisant la dispersion anormale sur des cristaux dérivés

La loi de Friedel (voir §2.2.2) indiquant que $I_{hkl} = I_{\bar{h}\bar{k}\bar{l}}$ n'est en fait qu'une approximation. Cette loi n'est pas toujours respectée, à cause d'une absorption partielle des rayons X par les atomes pour certaines longueurs d'onde. Cette absorption implique un changement de phase dans le facteur de diffusion atomique qui devient complexe, les termes correctifs étant appelés la **diffusion anormale** :

$$\mathbf{f} = \mathbf{f}_0 + \Delta\mathbf{f}' + i\Delta\mathbf{f}'' \quad (2.4.9)$$

Dans cette expression, \mathbf{f}_0 est le facteur atomique non corrigé et $\Delta\mathbf{f}'$ et $\Delta\mathbf{f}''$ sont respectivement les termes réel et imaginaire de la correction anormale. $\Delta\mathbf{f}'$ est aligné avec \mathbf{f}_0 et $\Delta\mathbf{f}''$ est tourné dans le plan complexe de 90° dans le sens inverse des aiguilles d'une montre par rapport à \mathbf{f}_0 .

L'absorption des rayons X par un atome, et donc l'effet de la diffusion anormale, devient importante lorsque la longueur d'onde des rayons X incidents est proche d'un seuil d'absorption de l'atome. Les longueurs d'onde habituellement utilisées en cristallographie sont très éloignées des seuils d'absorption des éléments légers. Par conséquent, les atomes tels que C, N ou O ne contribuent pratiquement pas à la dispersion anormale. Par contre, ces longueurs d'onde sont proches des seuils d'absorption des éléments lourds, qui contribuent alors de façon non négligeable à la diffusion anormale. En ajustant la longueur d'onde (ce qui est possible avec un rayonnement synchrotron), l'effet peut être à volonté minimisé ou renforcé pour un atome lourd donné.

L'effet de la dispersion anormale sur un facteur de structure particulier \mathbf{F}_{PL} des données obtenues pour un cristal substitué est illustré sur la figure 2.4.9 (il est important de souligner que cet effet est en réalité relativement faible et qu'il a été exagéré sur la figure). $\mathbf{F}_{\text{PL}}^{\lambda_1}$ dénote un facteur de structure pour une longueur d'onde λ_1 qui ne donne pas, ou pratiquement pas, de dispersion anormale et $\mathbf{F}_{\text{PL}}^{\lambda_2}$ le même facteur de structure pour une longueur d'onde λ_2 occasionnant une importante dispersion anormale au niveau de l'atome lourd. Nous avons la relation :

$$\mathbf{F}_{\text{PL}}^{\lambda_2} = \mathbf{F}_{\text{PL}}^{\lambda_1} + \Delta\mathbf{F}_r + \Delta\mathbf{F}_i \quad (2.4.10)$$

Comme dans la méthode de l'atome lourd, le but est d'obtenir la phase de $\mathbf{F}_{\text{PL}}^{\lambda_1}$ pour ensuite en déduire la phase de \mathbf{F}_P , connaissant la phase de \mathbf{F}_L (déterminer en résolvant la structure hypothétique avec l'atome lourd seul, voir ci-dessous). Ici, cependant, l'ambiguïté sur la phase, consécutive à l'application de la méthode SIR, va

être levée non pas en utilisant un deuxième cristal substitué comme dans la méthode MIR, mais en réalisant deux enregistrements à deux longueurs d'onde différentes pour le même cristal substitué.

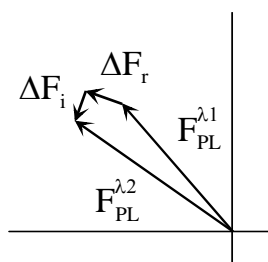


Fig. 2.4.9 – Contribution réelle et imaginaire de la dispersion anormale au facteur de structure

Les amplitudes des termes réel et imaginaire de la correction anormale pour un élément donné sont constantes et quasi indépendantes de l'angle θ de la réflexion (donc des indices hkl). Elles sont tabulées dans les Tables Internationales de Cristallographie. Les phases de ces mêmes termes ne dépendent que de la position de l'atome et sont donc connues dès que la structure hypothétique de l'atome lourd seul est résolue. Ceci permet de trouver la phase de $\mathbf{F}_{\text{PL}}^{\lambda 1}$, en considérant l'effet de la diffusion anormale sur la paire de Friedel $\mathbf{F}_{\text{PL}}^{\lambda 1+}$ (hkl) et $\mathbf{F}_{\text{PL}}^{\lambda 1-}$ ($\bar{h}\bar{k}\bar{l}$). Cet effet est illustré sur la figure 2.4.10. On remarque que, comme nous l'avons déjà signalé, la loi de Friedel n'est plus vérifiée en présence de diffusion anormale.

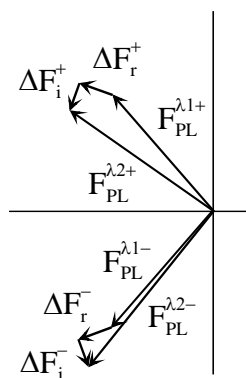


Fig. 2.4.10 – Effet de la diffusion anormale sur une paire de Friedel

La construction géométrique que nous allons utiliser est semblable à celle que nous avons employé pour la méthode MIR. Partons de l'équation suivante impliquant le facteur de structure, $\mathbf{F}_{\text{PL}}^{\lambda 1+}$, d'une des composantes d'une paire de Friedel :

$$\mathbf{F}_{\text{PL}}^{\lambda 1+} = \mathbf{F}_{\text{PL}}^{\lambda 2+} - \Delta \mathbf{F}_r - \Delta \mathbf{F}_i \quad (2.4.11)$$

Ne connaissant que les modules $|\mathbf{F}_{\text{PL}}^{\lambda 1+}|$ et $|\mathbf{F}_{\text{PL}}^{\lambda 2+}|$ de $\mathbf{F}_{\text{PL}}^{\lambda 1+}$ et $\mathbf{F}_{\text{PL}}^{\lambda 2+}$, deux solutions sont possibles, \mathbf{F}_a et \mathbf{F}_b , comme indiqué sur la figure 2.4.11.

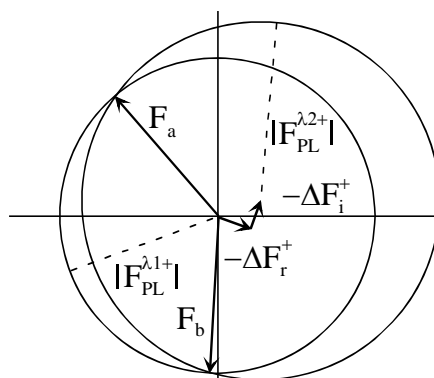


Fig. 2.4.11 – Vecteurs solutions de l'équation 2.4.11

Le second facteur de structure, $\mathbf{F}_{\text{PL}}^{\lambda 2-}$, de la seconde composante de la paire de Friedel va nous permettre de lever l'ambiguïté. Appliquons tout d'abord une symétrie par rapport à l'axe des réels de toutes les contributions à $\mathbf{F}_{\text{PL}}^{\lambda 2-}$, comme indiqué sur la partie gauche de la figure 2.4.11. Notez que l'image par rapport à l'axe des réels de $\Delta \mathbf{F}_r^-$ est $\Delta \mathbf{F}_r^+$ et l'image de $\Delta \mathbf{F}_i^-$ est $-\Delta \mathbf{F}_i^+$ (voir figure 2.4.10). De même, l'image de $\mathbf{F}_{\text{PL}}^{\lambda 1-}$ est $\mathbf{F}_{\text{PL}}^{\lambda 1+}$. Par contre l'image de $\mathbf{F}_{\text{PL}}^{\lambda 2-}$ n'est pas $\mathbf{F}_{\text{PL}}^{\lambda 2+}$, seul le module de $\mathbf{F}_{\text{PL}}^{\lambda 2-}$ est conservé. A partir de cette relation, ne connaissant pas les phases de $\mathbf{F}_{\text{PL}}^{\lambda 1+}$ et $\mathbf{F}_{\text{PL}}^{\lambda 2-}$, deux solutions sont possibles (\mathbf{F}_c et \mathbf{F}_d , voir partie droite de la figure 3.4.12). Même si ces solutions sont très imprécises à cause de la quasi tangence des cercles, \mathbf{F}_c correspond sans ambiguïté à la solution \mathbf{F}_a précédemment trouvée.

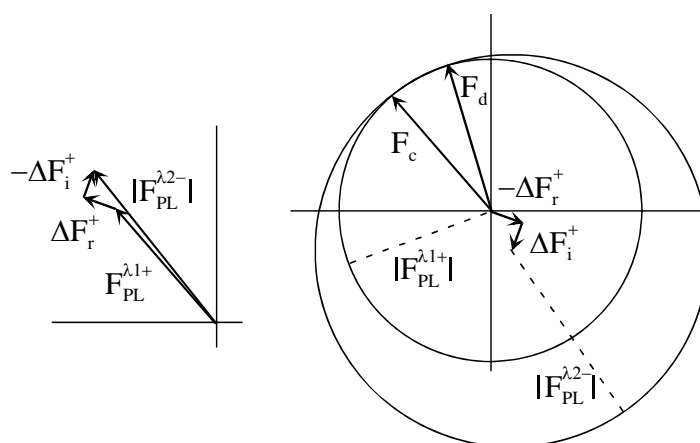


Fig. 2.4.12 – Vecteurs solutions pour le second facteur de structure de la paire de Friedel

2.4.6 – La méthode utilisant la diffusion anormale sans cristaux dérivés

Certaines protéines contiennent naturellement un atome lourd (par exemple Fe dans une globine ou un cytochrome). D'autres peuvent être modifiées en remplaçant la méthionine à base de soufre par la sélénométhionine et donc contenir un atome "lourd", le sélénium. Pour ces protéines, il n'est pas nécessaire de préparer des dérivés. En effet avec un seul cristal, en jouant sur la dispersion anormale, des enregistrements à plusieurs

longueurs d'onde permettent de remonter aux phases. C'est la méthode MAD, "Multiwavelength Anomalous Dispersion".

2.4.7 – Le remplacement moléculaire

Lorsque deux protéines sont supposées être très proches en termes de structure et conformation, il est parfois possible d'utiliser les phases déterminées pour une des protéines (de structure connue) comme base pour la détermination structurale de l'autre protéine. C'est la méthode du remplacement moléculaire. Ce remplacement moléculaire est illustré sur les figures 2.4.13 à 2.4.17.

La transformée de Fourier (amplitudes et phases) du chat sans queue présenté sur la figure 2.4.13 est donnée sur la figure 2.4.14. Les amplitudes (et seulement les amplitudes) de la transformée de Fourier d'un objet inconnu, de forme supposée assez proche du chat sans queue, sont présentées sur la figure 2.4.15.



Fig. 2.4.13 – Objet connu

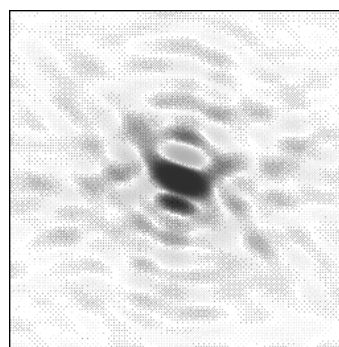


Fig. 2.4.14 – TF de l'objet connu

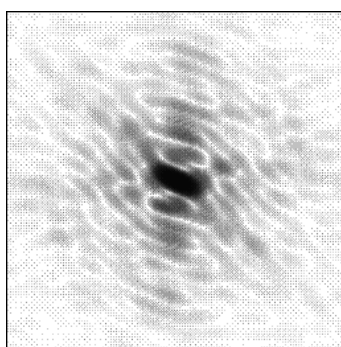
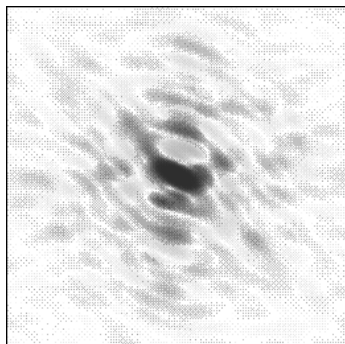


Fig. 2.4.15 – Amplitudes de la TF de l'objet inconnu

En combinant (figure 2.4.15) les amplitudes de la TF de l'objet à découvrir et les phases de la TF de l'objet connu, on découvre par transformée de Fourier inverse un chat entier, certes avec une queue d'intensité moitié (figure 2.4.17). Ce chat (quelque peu mutilé et bruité) peut servir de nouveau modèle pour poursuivre l'affinement de l'objet.



**Fig. 2.4.16 – Mélange phase objet connu /
amplitude objet inconnu**

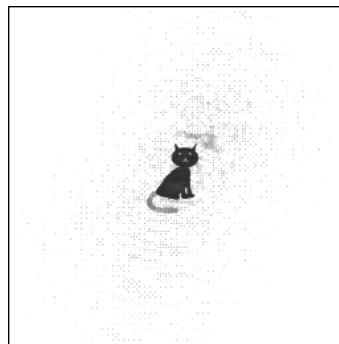


Fig. 2.4.17 – TF inverse du mélange

2.4.8 – Les méthodes directes

Les méthodes impliquant un atome lourd ne s'appliquent qu'aux très grosses molécules (plus de 500 atomes, sans compter les atomes d'hydrogène). Pour les plus petites molécules, les phases peuvent être déterminées par ce qu'on appelle les "méthodes directes". Ces méthodes font appel à des traitements mathématiques et statistiques relativement compliqués. Le principe général est de générer des relations statistiques (entre les phases), à partir desquelles le choix d'un nombre limité de phases permet de déterminer l'ensemble des phases.

Chapitre 3

Compléter et affiner une structure

Introduction

Les modèles structuraux obtenus par les méthodes présentées dans le chapitre précédent sont très souvent incomplets (il manque beaucoup d'atomes) et peu précis (les positions et la nature des atomes sont approximatives). Ceci est lié au fait que les phases obtenues ne sont que des estimations grossières des phases réelles. Il est alors nécessaire de compléter et affiner les structures. Si environ 50 à 60% de la densité électronique a pu être localisée avec une assez grande précision, il est assez facile de terminer la structure par des synthèses de Fourier et des affinements par moindres carrés. Si le modèle initial est plus restreint, il est quelquefois possible de finir la structure, en prenant quelques précautions. Dans ce chapitre, nous limiterons notre étude au cas des petites molécules, plus facile à comprendre, mais les mêmes principes s'appliquent pour les protéines.

3.1 – Compléter une structure

A partir d'un modèle structural il est possible de calculer les facteurs de structure, selon :

$$\mathbf{F}_{calc}(hkl) = \sum_{j=1}^n f_j \exp 2\pi i(hx_j + ky_j + lz_j) \quad (3.1.1)$$

Ces facteurs de structure comprennent une amplitude et une phase :

$$\mathbf{F}_{calc}(hkl) = |\mathbf{F}_{vrai}(hkl)| \exp 2\pi i \alpha_{calc}(hkl) \quad (3.1.2)$$

Ici la phase est à nouveau exprimée en fraction de 2π , contrairement à la notation de l'équation 3.4.1.

L'enregistrement des données de diffraction donne accès aux modules "observés" des facteurs de structure de la structure réelle. Ces modules sont légèrement différent des modules vrais, $|\mathbf{F}_{vrai}(hkl)|$, puisque la mesure introduit nécessairement une erreur. Notons ces modules $|\mathbf{F}_{obs}(hkl)|$.

3.1.1 – Les synthèses de Fourier \mathbf{F}_{calc}

Une synthèse de Fourier (= transformation de Fourier) avec les coefficients $\mathbf{F}_{calc}(hkl)$ par le biais de l'équation :

$$\begin{aligned}\rho_{calc}(x, y, z) &= \frac{1}{V} \sum_h \sum_k \sum_l \mathbf{F}_{calc}(hkl) \exp\{-2\pi i(hx + ky + lz)\} \\ &= \frac{1}{V} \sum_h \sum_k \sum_l |\mathbf{F}_{calc}(hkl)| \exp\{-2\pi i(hx + ky + lz) + i\alpha_{calc}(hkl)\}\end{aligned} \quad (3.1.3)$$

donne une densité électronique avec des maxima aux positions des atomes. Puisque c'est la transformation de Fourier inverse de celle de l'équation 3.1.1, on retrouve intégralement le modèle structural qui a initialement servi à calculer les facteurs de structure.

3.1.2 – Les synthèses de Fourier \mathbf{F}_{obs}

Une synthèse de Fourier avec les modules $|\mathbf{F}_{obs}(hkl)|$ (mesurés) et les phases vrai $\alpha_{vrai}(hkl)$ (inconnues) par le biais de l'équation :

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |\mathbf{F}_{obs}(hkl)| \exp\{-2\pi i(hx + ky + lz) + i\alpha_{vrai}(hkl)\} \quad (3.1.4)$$

donnerait une densité électronique avec des maxima aux positions des atomes de la structure réelle (et un bruit de fond lié aux erreurs sur les $|\mathbf{F}_{obs}(hkl)|$). Si le modèle est proche de la structure réelle, les phases $\alpha_{calc}(hkl)$ doivent être voisines des phase vraies, $\alpha_{vrai}(hkl)$, et le calcul de synthèse de Fourier \mathbf{F}_{obs} suivant :

$$\rho_{obs}(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |\mathbf{F}_{obs}(hkl)| \exp\{-2\pi i(hx + ky + lz) + i\alpha_{calc}(hkl)\} \quad (3.1.5)$$

doit donner une densité électronique proche de celle de la structure réelle.

3.1.3 – Les synthèses de Fourier différence $\mathbf{F}_{obs} - \mathbf{F}_{calc}$

Pour connaître la différence entre le modèle et la structure réelle, il faut calculer la synthèse de Fourier différence $\mathbf{F}_{obs} - \mathbf{F}_{calc}$:

$$\begin{aligned}
 \Delta\rho(x, y, z) &= \rho_{obs}(x, y, z) - \rho_{calc}(x, y, z) \\
 &= \frac{1}{V} \sum_h \sum_k \sum_l (\mathbf{F}_{obs}(hkl) - \mathbf{F}_{calc}(hkl)) \exp\{-2\pi i(hx + ky + lz)\} \\
 &= \frac{1}{V} \sum_h \sum_k \sum_l (|\mathbf{F}_{obs}(hkl)| - |\mathbf{F}_{calc}(hkl)|) \exp\{-2\pi i(hx + ky + lz) + i\alpha_{calc}(hkl)\}
 \end{aligned}
 \tag{4.1.6}$$

S'il manque des informations dans le modèle, celles-ci vont apparaître dans les cartes de Fourier différence sous la forme de résidus positifs. Inversement, si le modèle contient des informations en trop, celles-ci vont apparaître sous la forme de résidus négatifs.

La procédure est alors la suivante. A partir de la première estimation des phases, une densité électronique est établie par une synthèse de Fourier \mathbf{F}_{obs} . Si cette densité a un quelconque sens (on y reconnaît des fragments de structure), elle sert de modèle de départ. Le modèle peut être amélioré de deux façons : 1) directement par le cristallographe qui peut supprimer, modifier ou ajouter des atomes en fonction de ses connaissances, 2) en calculant une synthèse de Fourier différence qui va donner des informations supplémentaires. Un problème cependant : la synthèse de Fourier différence est effectuée avec les phases calculées, issues du modèle. Si le modèle est très éloigné de la structure réelle, ces phases sont fausses et les informations tirées de la synthèse erronées (souvenez vous du mélange chat/canard, ce sont les phases qui priment sur les amplitudes). La détermination structurale est alors vouée à l'échec... il faut re-préciser les phases initiales (autre cristal dérivé, etc.). Si le modèle est voisin de la structure réelle, les phases sont assez proches des phases réelles et les informations tirées de la synthèse de Fourier différence vont permettre d'améliorer le modèle, ce qui va permettre de calculer de nouvelles phases encore plus proches des phases réelles, etc. La structure est déterminée par itérations jusqu'à épuisement des informations tirées des cartes de Fourier différence qui ne contiennent alors que du bruit.

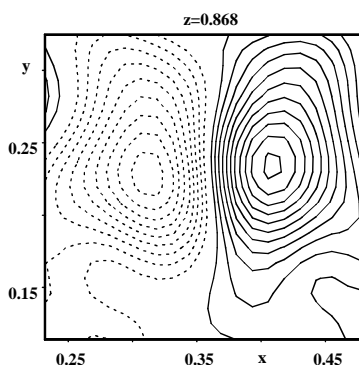


Fig. 2.1.1 – Carte de Fourier différence montrant une position atomique erronée

Non seulement les synthèses de Fourier différence permettent de trouver les atomes manquants, mais elles permettent aussi de repérer des anomalies dans le modèle initial. En effet, des positions erronées vont se traduire par un excès de densité à

proximité d'un défaut de densité (voir figure 3.1.1 ci-dessus), des erreurs sur l'attribution du type d'atome vont conduire à des excès (atome trop léger) ou des défauts (atome trop lourd) de densité électronique au niveau de l'atome concerné (voir figures 3.1.2a et 3.1.2b), des erreurs sur les paramètres de déplacement atomiques ("vibrations") vont être révélées par un anneau positif autour d'un pic négatif (déplacement trop petit) ou inversement (déplacement trop grand) (voir figures 3.1.3a et 3.1.3b).

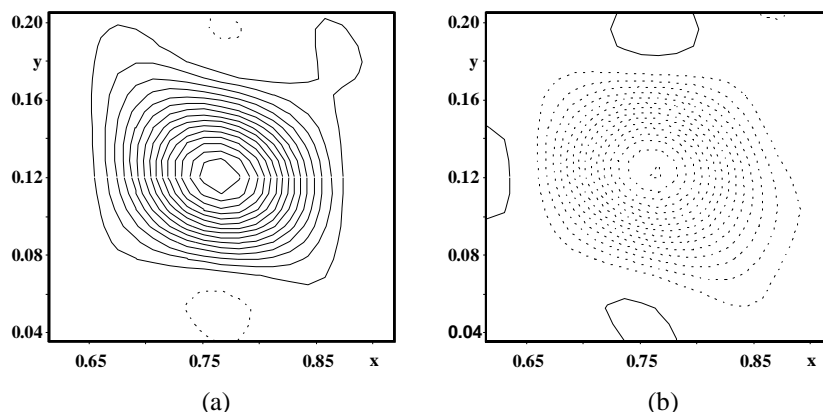


Fig. 3.1.2 – Cartes de Fourier différence montrant (a) un défaut et (b) un excès de densité électronique

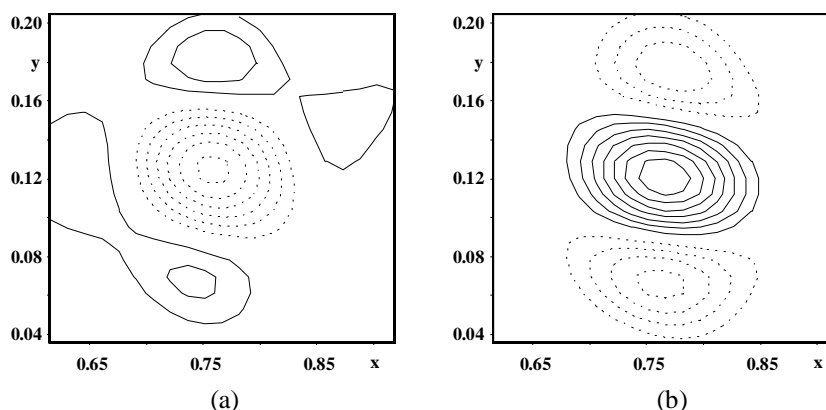


Fig. 3.1.3 – Cartes de Fourier différence montrant un paramètre de déplacement atomique (a) trop petit et (b) trop grand

3.1.4 – Les synthèses de Fourier différence $2\mathbf{F}_{obs} - \mathbf{F}_{calc}$

En cristallographie des protéines, les cartes de Fourier différence sont difficiles à interpréter car la résolution est souvent loin d'être de niveau atomique (voir ci-dessous). Les cristallographes préfèrent alors utiliser une carte $2\mathbf{F}_{obs} - \mathbf{F}_{calc}$ (ou même $3\mathbf{F}_{obs} - 2\mathbf{F}_{calc}$) :

$$\begin{aligned} \Delta\rho(x, y, z) &= 2\rho_{obs}(x, y, z) - \rho_{calc}(x, y, z) \\ &= \frac{1}{V} \sum_h \sum_k \sum_l (2|\mathbf{F}_{obs}(hkl)| - |\mathbf{F}_{calc}(hkl)|) \exp\{-2\pi i(hx + ky + lz) + i\alpha_{calc}(hkl)\} \end{aligned} \quad (3.1.7)$$

Cette carte a l'avantage d'être en général partout positive et de montrer, si le modèle n'est pas totalement erroné, les défauts du modèle en superposition avec la véritable structure. Par exemple, si un oxygène d'un groupe carbonyle du modèle pointe à 180° de sa véritable position dans la structure, la protubérance correspondante dans la carte $2F_{obs} - F_{calc}$ va apparaître moins grosse qu'elle le devait et une faible protubérance sera probablement visible de l'autre côté du carbone du groupe carbonyle. Dans une carte $F_{obs} - F_{calc}$ il y aurait une zone négative d'un côté et une zone positive de l'autre, mais sans la structure, ce qui complique l'interprétation lorsque la résolution est faible.

3.2 – Affiner une structure

3.2.1 – La méthode des moindres carrés

Les cycles de calculs de synthèses de Fourier différence et d'élaboration de modèles sont ponctués de tentatives du cristallographe d'amélioration du modèle par des calculs d'optimisation.

Le facteur de structure peut être exprimé sous une forme qui précise tous les paramètres inclus dans les affinements :

$$F_{hkl} = k \sum_{j=1}^n n_j f_j \exp \left\{ -B_j \left(\frac{\sin \theta}{\lambda} \right)^2 \right\} \exp 2\pi i (hx_j + ky_j + lz_j) \quad (3.2.1)$$

où k est un facteur d'échelle absolu (facteur qui, nous l'avons déjà mentionné, dépend de l'intensité de la source des rayons X, des temps de pose, de la taille du cristal, etc.), n_j est l'occupation de l'atome j (un atome peut être présent dans certaines mailles mais pas dans d'autres, ce qui, en moyenne, conduit à un taux d'occupation inférieur à 1.0) et B_j est le facteur de déplacement atomique de l'atome j . Ce dernier facteur traduit à la fois l'agitation de l'atome (qui peut être considérée dans un premier temps comme isotrope, mais qui est en réalité anisotrope) et les désordres statiques et dynamiques. Notez que le déplacement atomique a pour effet de réduire l'amplitude du facteur de structure, particulièrement aux grands angles.

Les paramètres, p_i , qu'il est possible d'affiner pour un modèle sont alors :

- le facteur d'échelle
- les positions des atomes (coordonnées atomiques)
- les paramètres de déplacement atomique (vibration isotropes, puis anisotropes)
- les coefficients d'occupation des sites (dans le cas de non stœchiométries ou de désordre)

L'affinement de structure est, dans la plupart des programmes, basé sur la minimisation de la fonction de F :

$$S(F) = \sum_{hkl} w(hkl) \left(k \left| \mathbf{F}_{obs}(hkl) \right| - \left| \mathbf{F}_{calc}(hkl) \right| \right)^2 \quad (3.2.2)$$

ou bien sur la minimisation de la fonction de F^2 :

$$S(F^2) = \sum_{hkl} w(hkl) \left(k \left| \mathbf{F}_{obs}(hkl) \right|^2 - \left| \mathbf{F}_{calc}(hkl) \right|^2 \right)^2 \quad (3.2.3)$$

où $w(hkl)$ est un facteur de pondération qui permet d'accorder plus de poids aux informations fiables (intensités fortes, bien mesurées) qu'aux informations douteuses (intensités faibles, mal mesurées). C'est la méthode des moindres carrés. Lorsque la fonction S est minimale, les dérivés de S par rapport aux paramètres sont nulles et le modèle est le plus proche possible de la structure réelle. Dans le cas contraire, des variations de paramètres sont calculées et appliquées dans un processus itératif.

La validité du modèle est habituellement testée à partir de l'indice résiduel :

$$R = \sum_{hkl} \left| \mathbf{F}_{obs}(hkl) \right| - \left| \mathbf{F}_{calc}(hkl) \right| / \sum_{hkl} \left| \mathbf{F}_{obs}(hkl) \right| \quad (3.2.4)$$

et des critères suivants, respectivement pour des affinement en F et F^2 :

$$wR(F) = \left\{ \sum_{hkl} w \left| \mathbf{F}_{obs}(hkl) \right| - \left| \mathbf{F}_{calc}(hkl) \right| \right|^2 / \sum_{hkl} w \left| \mathbf{F}_{obs}(hkl) \right|^2 \right\}^{1/2} \quad (3.2.5)$$

$$wR(F^2) = \left\{ \sum_{hkl} w \left| \mathbf{F}_{obs}(hkl) \right|^2 - \left| \mathbf{F}_{calc}(hkl) \right|^2 \right|^2 / \sum_{hkl} w \left| \mathbf{F}_{obs}(hkl) \right|^2 \right\}^{1/2} \quad (3.2.6)$$

qui sont plus appropriés pour évaluer la signification des petits changements de paramètres en fin d'affinement.

Il est important de noter que l'affinement ne converge pas vers la bonne solution si le modèle initial est trop éloigné de la structure réelle ou ne permet pas d'aboutir à la structure réelle (s'il manque par exemple des atomes).

3.2.2 – Résolution et précision

La précision sur les positions atomiques affinées dépend du nombre de réflexions (et de la qualité des réflexions) incluses dans l'affinement. Ceci est illustré sur les figures 3.2.1 où le fait de limiter la résolution dans l'espace réciproque conduit à une faible précision dans l'espace direct. En cristallographie, la résolution est donnée sous deux formes :

- 1) une distance exprimée en Å (par exemple : résolution à 2Å). Cette distance indique que les mesures ont été effectuées jusqu'à une distance dans l'espace réciproque de $1/2\text{Å}$.

2) la valeur de $\sin \theta / \lambda$ maximale (par exemple : 0.25 \AA^{-1}).

La loi de Bragg ($2d \sin \theta = \lambda$) donne la correspondance entre les deux formes $0.25 \text{ \AA}^{-1} \leftrightarrow 2 \text{ \AA}$. Il n'est pas recommandé de donner l'angle θ maximum, car la résolution dépend alors de la longueur utilisée (qui doit être dans ce cas précisée).

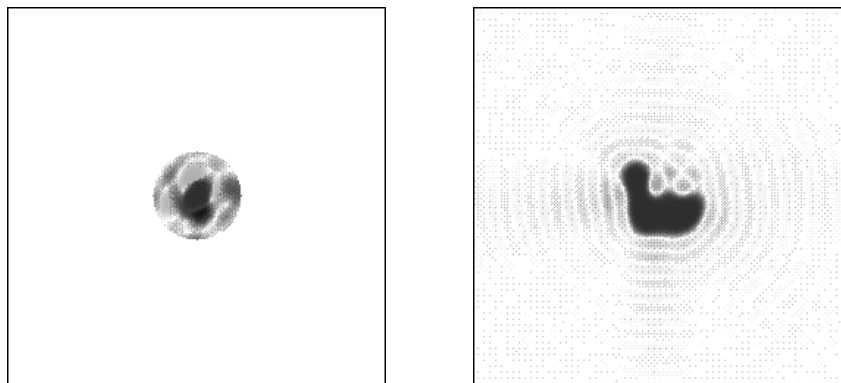


Fig. 3.2.1 – Effet d'une limitation aux termes de basse résolution (partie gauche, espace réciproque) sur la précision des formes du canard obtenues par TF (partie droite, espace direct)

L'effet de la résolution est illustré sur les figures 3.2.2 à 3.2.5. A 1.0 \AA on peut placer facilement les atomes et on distingue bien l'atome d'azote qui est plus gros que les atomes de carbone. A 2.5 \AA , on peut ajuster les cycles. A 3.0 \AA cela devient plus difficile et à 4.0 \AA on est dans le brouillard... Alors pourquoi ne pas toujours enregistrer des données pour les protéines avec une très bonne résolution ? Deux raisons pour cela :

- 1) le nombre de réflexions théoriquement mesurables est approximativement égal à :

$$N = \frac{(4\pi/3) \cdot (2/\lambda)^3}{V^*} = \frac{33.5 \cdot V}{\lambda^3} \quad (3.2.7)$$

(obtenu en divisant le volume de la sphère de diffraction, de rayon $2/\lambda$, par le volume de maille réciproque). Pour une protéine très modeste de maille orthorhombique $40 \times 60 \times 80 \text{ \AA}^3$ et une longueur d'onde $\lambda = 1.54 \text{ \AA}$, le nombre de réflexions possibles est 1.76×10^6 !

- 2) les protéines et les molécules de solvant (eau) de cristallisation sont souvent très agitées et/ou désordonnées, ce qui conduit au même effet : une atténuation très rapide des facteurs de structure, donc des intensités, avec l'angle de diffraction (voir équation 3.2.1). A partir d'une certaine distance dans l'espace réciproque, il n'y a alors plus rien qui sorte du bruit. La résolution réelle est alors bien inférieure à la résolution théorique. Une technique, la cryocristallographie, permet d'étendre le nombre de réflexions mesurables; elle consiste à refroidir le cristal à des températures inférieures à 130 K (-143°C) pour limiter l'agitation thermique et éventuellement résoudre les désordres.

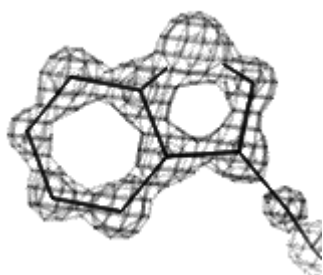


Fig. 3.2.2 – Résolution de 1.0 Å

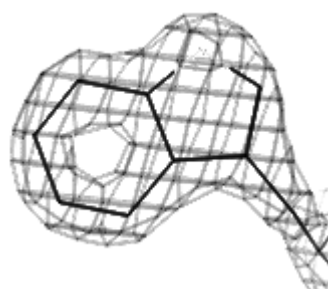


Fig. 3.2.3 – Résolution de 2.5 Å

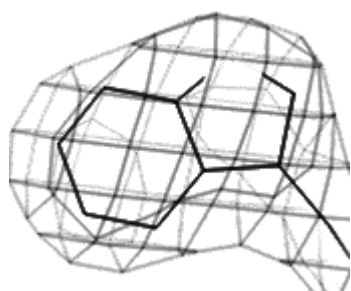


Fig. 3.2.4 – Résolution de 3.0 Å



Fig. 3.2.5 – Résolution de 4.0 Å

Pour les petites molécules organiques, il est assez facile d'avoir en routine de très bonne résolution (<1.0 Å) et de distinguer les atomes d'hydrogène sur les cartes de Fourier (voir figure 3.2.6). Avec quelques efforts, il est possible d'atteindre des résolutions de 0.6 à 0.5 Å, voire encore plus faibles, et d'avoir accès à des informations très précises sur les vibrations, les désordres ou même les doublets libres où les électrons de liaisons !

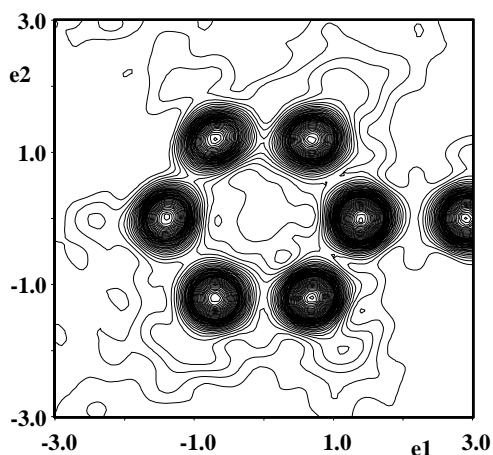


Fig. 3.2.6 – Petite molécule, résolution de 0.77 Å

La présentation finale des structures dépend de la nature des composés. Pour les petites molécules on visualise en général les atomes du modèle comme sur la figure 3.2.7. Pour les protéines, étant donnée la complexité des structures, on préfère en

général donner des enveloppes ou des vues plus schématiques comme celle de la figure 3.2.8.

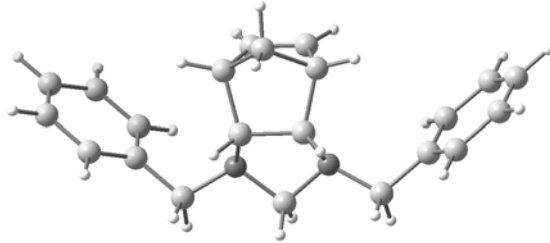


Fig. 3.2.7 – La molécule 8,10-Dibenzylimidazonorbornylene

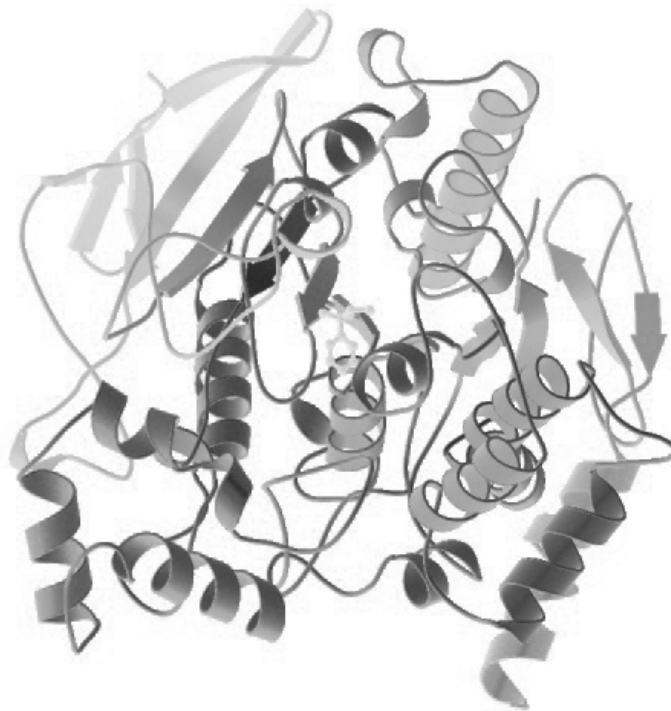


Fig. 3.2.8 – Dessin "en ruban" de l'Acetylcholinesterase

Bibliographie

Niveau élémentaire :

Crystallography Made Crystal Clear, 2nd Edition (2000)

Gale Rhodes

Academic Press

Excellent ouvrage introductif. Quelques simplifications auraient néanmoins sans doute pu être évitées. Attention à quelques erreurs de frappe, en particulier dans les équations.

Les cristaux, fenêtres sur l'invisible (1999)

Alain Pénicaud

Ellipses

Pour savoir comment sont nés les concepts de la cristallographie. Une seule formule mathématique (loi de Bragg) !

Structure determination by X-ray Crystallography, 2nd Edition (1988)

M.F.C. Ladd & R.A. Palmer

Plenum Press

Livre un peu plus ancien qui a servi d'ouvrage de base à de nombreux étudiants.

Niveau avancé :

Practical Protein Crystallography, 2nd Edition (1999)

Duncan E. McRee

Academic Press

Beaucoup de points intéressants, mais dans un ordre un peu déroutant. Enormément (trop ?) de références au programme écrit par l'auteur...

Crystallography (1993)

Walter Borchardt-Ott

Springer-Verlag

Essentiellement un ouvrage de cristallographie géométrique

Fundamentals of Crystals, 2nd Edition (1994)

Boris K. Vainsthein

Springer

Ouvrage de cristallographie "standard". Ne mentionne pas la cristallographie des protéines.

Cristallographie (1993)

Dieter Schwarzenbach

Presses Polytechn. et Univ. Romandes

Essentiellement de la cristallographie géométrique. Très rigoureux.

Space Groups for Solid State Scientists (1990)

G. Burns & A.M. Glazer

Academic Press

Excellent ouvrage sur les groupes d'espace.

Fundamentals of Crystallography (1992)

C. Giacovazzo

Oxford Science Publications

Ouvrage de cristallographie "standard". Assez complexe dans les notations.

Cristallographie géométrique et radiocristallographie (1999)

Jean-Jacques Rousseau

Dunod

Cristallographie pour "chimiste inorganicien"

Quelques sites internet intéressants

Quelques sites ont été utilisés pour écrire ce cours (illustrations, images, etc.), en particulier le site de Kevin Cowtan sur les transformations de Fourier.

Transformation de Fourier :

<http://www.ytbl.york.ac.uk/~cowtan/fourier/fourier.html> (*utilisé dans ce cours*)

Facteur de structure (interactif) :

<http://www.ytbl.york.ac.uk/~cowtan/sfapplet/sfintro.html> (*à voir !*)

Cristallographie des protéines :

<http://www-structure.llnl.gov/Xray/101index.html>

<http://www-structmed.cimr.cam.ac.uk/course.html>

Les principales constantes[†]

Célérité de la lumière dans le vide	c	$2,997\,924\,58\,10^8\text{ m.s}^{-1}$ (exacte)
Charge élémentaire	e	$1,602\,176\,53(14)\,10^{-19}\text{ C}$
Constante de Planck	h	$6,626\,0693(11)\,10^{-34}\text{ J.s}$
Rayon de Bohr	a_0	$0,529\,177\,2108(18)\,10^{-10}\text{ m}$
Rydberg (Unité d'énergie)	$E_0, R_y, R_\infty hc$	$13,605\,6923(12)\text{ eV}$
Constante d'Avogadro	N_A, L	$6,022\,1415(10)\,10^{23}\text{ mol}^{-1}$
Constante de Faraday	$F = N_A e$	$9,648\,533\,83(83)\,10^4\text{ C.mol}^{-1}$
Debye	D	$3.337\,868\,10^{-30}\text{ C.m}$
Perméabilité du vide	$\mu_0 = 4\pi\,10^{-7}$	$12,566\,370\,614\dots\,10^{-7}\text{ H/m}$ (exacte)
Permittivité du vide	$\varepsilon_0 = 1/\mu_0 c^2$	$8,854\,187\,817\dots\,10^{-12}\text{ F.m}^{-1}$ (exacte)
Masse de l'électron	m_e	$9,109\,3826(16)\,10^{-31}\text{ kg}$
Masse du proton	m_p	$1,672\,621\,71(29)\,10^{-27}\text{ kg}$
Masse du neutron	m_n	$1,674\,927\,28(29)\,10^{-27}\text{ kg}$
Constante de Rydberg	$R_H = R_\infty$	$1,097\,373\,156\,8525(73)\,10^7\text{ m}^{-1}$

[†] Constantes physiques fondamentales recommandées par le "Codata task group on fundamental constants" (voir <http://physics.nist.gov/cuu/Constants/index.html> pour la liste de toutes les constantes)

1	1,008	H^g Hydrogène $1s^1$
non métaux \longrightarrow		
2	4,00	He^g Hélium $1s^2$

non métaux

18

1	2
3 Li Lithium 2 ^{s1}	4 Be Béryllium 2 ^{s2}
11 Na Sodium 3 ^{s1}	12 Mg Magnésium 3 ^{s2}
19 K Potassium 4 ^{s1}	20 Ca Calcium 4 ^{s2}
37 Rb Rubidium 5 ^{s1}	38 Sr Strontium 5 ^{s2}
55 Cs Césium 6 ^{s1}	56 Ba Barium 6 ^{s2}
87 *Fr Francium 7 ^{s1}	88 *Ra Radium 7 ^{s2}

X élément gazeux à 25°C, 1 bar

élément liquide à 25°C, 1 bar

élément radioactif naturel

élément radioactif artificiel

Z	A	X	Élément	configuration
---	---	---	---------	---------------

5	B Bore 2,6,2 2p1	6	C Carbone 2,5,2 2p2	7	N Azote 2,5,2 2p3	8	O Oxygène 2,5,2 2p4	9	F Fluor 2,5,2 2p5	10	Ne Néon 2,5,2 2p6	20,18
13	Al Aluminium 3,5,2 3p1	14	Si Silicium 3,5,2 3p2	15	P Phosphore 3,5,2 3p3	16	S Soufre 3,5,2 3p4	17	Cl Chlore 3,5,2 3p5	18	Ar Argon 3,5,2 3p6	39,95
31	Ga Gallium 3d10 4,5,2 4p1	32	Ge Germanium 3d10 4,5,2 4p2	33	As Arsenic 3d10 4,5,2 4p3	34	Se Sélénium 3d10 4,5,2 4p4	35	Br Brome 3d10 4,5,2 4p5	36	Kr Krypton 3d10 4,5,2 4p6	83,80
49	In Indium 4d10 5,5,2 5p1	50	Sn Étain 4d10 5,5,2 5p2	51	Sb Antimoine 4d10 5,5,2 5p3	52	Te Tellure 4d10 5,5,2 5p4	53	I Iode 4d10 5,5,2 5p5	54	Xe Xénon 4d10 5,5,2 5p6	131,3
81	Tl Thallium 4p1 5d10 6,5,2 6p1	82	Pb Plomb 4p1 5d10 6,5,2 6p2	83	Bi Bismuth 4p1 5d10 6,5,2 6p3	84	Po Polonium 4p1 5d10 6,5,2 6p4	85	At Astaté 4p1 5d10 6,5,2 6p5	86	Rn Radon 4p1 5d10 6,5,2 6p6	(222)

métaux

21	Sc Scandium <i>3d1 4s2</i>	22	Ti Titan <i>3d2 4s2</i>	23	V Vanadium <i>3d3 4s2</i>	24	Cr Chrome <i>3d5 4s1</i>	25	Mn Manganese <i>3d5 4s2</i>	26	Fe Fer <i>3d6 4s2</i>	27	Co Cobalt <i>3d7 4s2</i>	28	Ni Nickel <i>3d8 4s2</i>	29	Cu Cuivre <i>3d10 4s1</i>	30	Zn Zinc <i>3d10 4s2</i>
39	Y Yttrium <i>4d1 5s2</i>	40	Zr Zirconium <i>4d2 5s2</i>	41	Nb Niobium <i>4d4 5s1</i>	42	Mo Molybdene <i>4d5 5s1</i>	43	<i>Tc</i> Technetium <i>4d5 5s2</i>	44	Ru Ruthenium <i>4d7 5s1</i>	45	Rh Rhodium <i>4d8 5s1</i>	46	Pd Palladium <i>4d10 5s0</i>	47	Ag Argent <i>4d10 5s1</i>	48	Cd Cadmium <i>4d10 5s2</i>
71	Lu Lutécium <i>4f14 5d1 6s2</i>	72	Hf Hafnium <i>4f14 5d2 6s2</i>	73	Ta Tantale <i>4f14 5d3 6s2</i>	74	W Tungstène <i>4f14 5d4 6s2</i>	75	Re Rhenium <i>4f14 5d5 6s2</i>	76	Os Osmium <i>4f14 5d6 6s2</i>	77	Ir Iridium <i>4f14 5d7 6s2</i>	78	Pt Platine <i>4f14 5d8 6s1</i>	79	Au Or <i>4f14 5d10 6s1</i>	80	Hg Mercure <i>4f14 5d10 6s2</i>
103	Lr Lawrencium <i>5f14 6d1 7s2</i>	104	Rf Rutherfordium <i>5f14 6d2 7s2</i>	105	Db Dubnium <i>5f14 6d3 7s2</i>	106	Sg Seaborgium <i>5f14 6d4 7s2</i>	107	Bh Bohrium <i>5f14 6d5 7s2</i>	108	Hs Hassium <i>5f14 6d6 7s2</i>	109	Mt Meitnerium <i>5f14 6d7 7s2</i>	110	? ? <i>5f14 6d8 7s2</i>	111	? ? <i>5f14 6d9 7s2</i>	112	? ? <i>5f14 6d10 7s2</i>

57	La	138.9	58	Ce	140.1	59	Pr	140.9	60	Nd	144.2	61	Pm	(145)	62	Sm	150.4	63	Eu	152.0	64	Gd	157.3	65	Tb	158.9	66	Dy	162.5	67	Ho	164.9	68	Er	167.3	69	Tm	168.9	70	Yb	173.0
	Lanthane			Cérium			Prasodyme			Néodyme			Prométhium			Samarium			Europium			Gadolinium			Terbium			Dysprosium			Holmium			Thulium			Ytterbium				
	4f0 5d1 6s2			4f1 5d1 6s2			4f3 5d0 6s2			4f4 5d0 6s2			4f5 5d0 6s2			4f6 5d0 6s2			4f7 5d0 6s2			4f7 5d1 6s2			4f9 5d0 6s2			4f10 5d0 6s2			4f11 5d0 6s2			4f13 5d0 6s2			4f14 6s2				
89	Ac	227.0	90	Th	232.0	91	Pa	231.0	92	U	238.0	93	Np	237.1	94	Pu	244.0	95	Am	243.0	96	Cm	247.0	97	Bk	247.0	98	Cf	251.0	99	Es	252.0	100	Fm	257.0	101	Md	258.0	102	No	259.0
	Actinium			Thorium			Protactinium			Uranium			Neptunium			Plutonium			Américium			Curium			Berkélium			Californium			Einsteinium			Fermium			Mendéléïevium			Nobelium	
	5f0 6d1 7s2			5f0 6d2 7s2			5f2 6d1 7s2			5f3 6d1 7s2			5f4 6d1 7s2			5f6 6d0 7s2			5f7 6d0 7s2			5f7 6d1 7s2			5f9 6d0 7s2			5f10 6d0 7s2			5f11 6d0 7s2			5f13 6d0 7s2			5f14 6d0 7s2				