# Les méthodes de phasage

## *Aspects théoriques*

**Claudine MAYER**
**Université Paris Diderot**
**Institut Pasteur**

université PARIS DIDEROT PARIS 7
USPC Université Sorbonne Paris Cité
Institut Pasteur
Unité de Microbiologie Structurale

# Plan

Introduction et récapitulatifs

1. Principe du phasage
2. La fonction de Patterson
3. Le remplacement isomorphe
4. La diffusion anomale
5. Le remplacement moléculaire

# INTRODUCTION ET RECAPITULATIFS

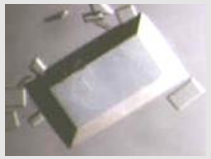
# Introduction

## 1. Les principales étapes

Clonage, expression, purification

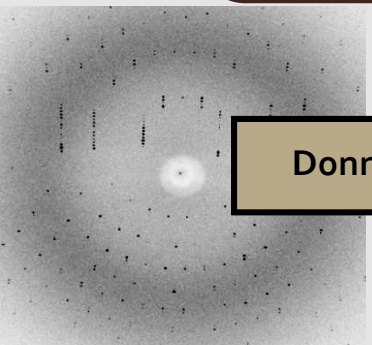
Macromolécules de grande pureté

Cristallisation



Cristaux

Collecte et intégration des données



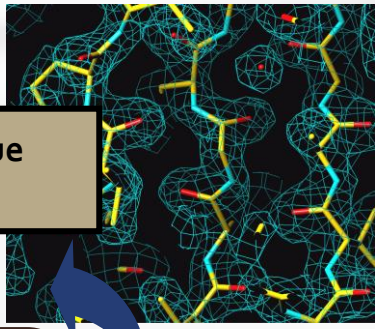
Données de diffraction

Phasage des réflexions

Carte de densité électronique

Affinement cristallographique

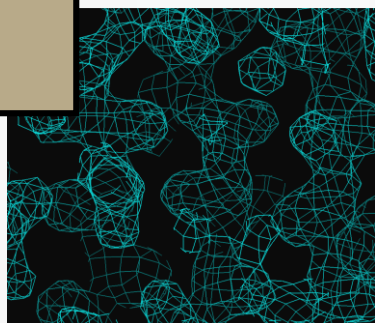
Structure atomique = MODELE



Interprétation biologique

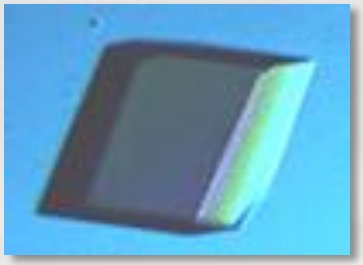
Analyse bioinformatique

Question biologique



# Introduction

## 2. Relations cristal - image de diffraction - structure



cristal

Le cristal placé dans un faisceau de RX réalise une analyse de Fourier

Le cristal est un arrangement 3D ordonné de macromolécules

Impact du rayon diffracté sur le détecteur 2D

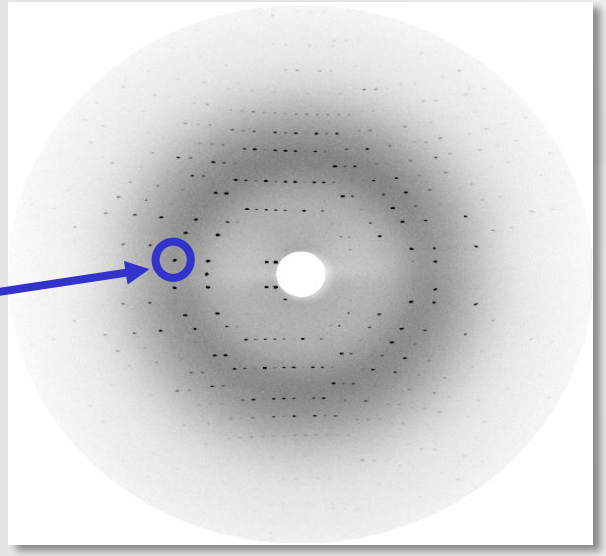
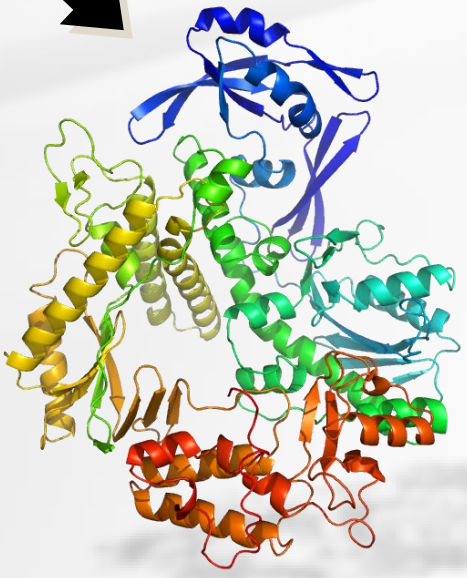


Image de diffraction

?



structure

# 1. PRINCIPE DU PHASAGE

# 1. Principe du phasage

## 1.1. Densité électronique et facteur de structure

On a vu que le facteur de structure s'exprime en fonction de la densité électronique

$$\vec{F}_{hkl} = \int_V \rho(xyz)\ [\exp 2\pi i(hx + ky + lz)]\ dv$$

Inversement, on peut montrer que la densité électronique s'exprime en fonction des facteurs de structure par une relation analogue :

$$\rho_{(x,y,z)} = \frac{1}{V}\sum_h\sum_k\sum_l \vec{F}_{(h,k,l)} \exp[-2\pi \cdot i(hx + ky + lz)]$$

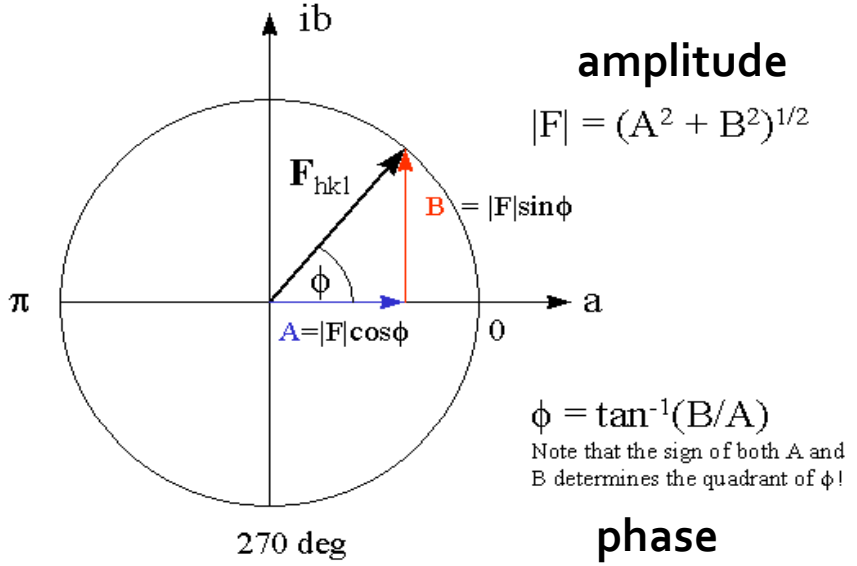
**avec $F_{(h,k,l)}$ défini par une amplitude ET une phase**

**C'est la synthèse de Fourier de la densité électronique**

# 1. Principe du phasage

## 1.1. Densité électronique et facteur de structure

$$\vec{F}_{(h,k,l)} = \sum_{j=1}^{atoms} f_{(j)} \exp\left[2\pi \cdot i(hx_{(j)} + ky_{(j)} + lz_{(j)})\right]$$



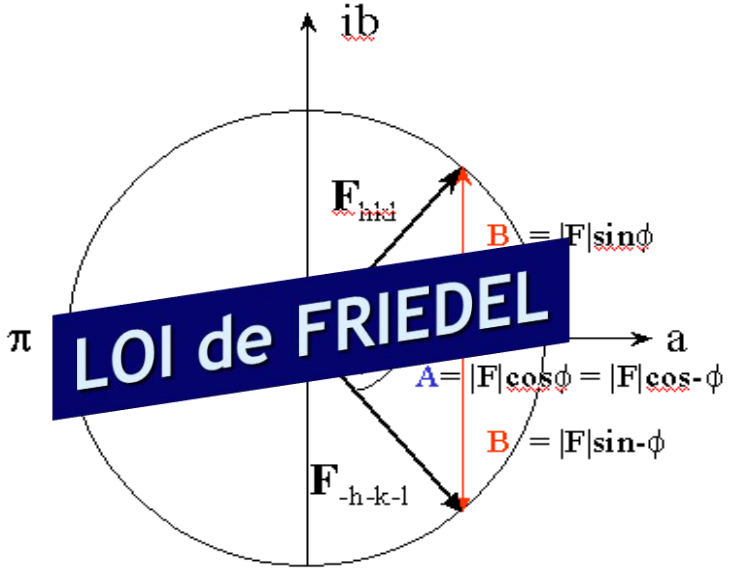

**amplitude**

$$|F| = (A^2 + B^2)^{1/2}$$

$$\phi = \tan^{-1}(B/A)$$

Note that the sign of both A and B determines the quadrant of ϕ!

**phase**




LOI de FRIEDEL

$$\vec{F}_{hkl} = A + iB = |F|e^{i\phi}$$

$$|F(hkl)| = |F(-h-k-l)|$$

# 1. Principe du phasage

## 1.4. Principe des méthodes de phasage

- La connaissance de $F_{hkl}$ en module et en phase est une condition indispensable au calcul de ρ(xyz), donc à la résolution de la structure

- L'expérience conduit au module du facteur de structure (intensité des taches du cliché de diffraction) mais ne conduit pas à la phase

➡ L'objectif des méthodes de résolution est de déterminer les phases des facteurs de structures

# 1. Principe du phasage

## 1.4. Principe des méthodes de phasage

$$\rho(x,y,z) = TF^{-1}(F(h,k,l))$$

$$\rho_{(x,y,z)} = \frac{1}{V}\sum_h\sum_k\sum_l \vec{F}_{(h,k,l)} \exp[-2\pi \cdot i(hx+ky+lz)]$$

avec $\vec{F}_{hkl} = |F_{hkl}|\, e^{i\alpha_{hkl}}$

$F_{obs}$

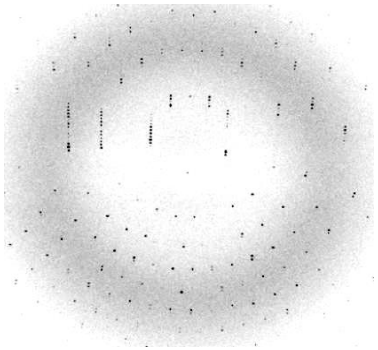
Facteurs de Structure (module, phase)

$I^{1/2}$ ?

*truncate*
*xdsconv*

Données de
diffraction traitées
= liste des réflexions (h,k,l)

| | | | | |
|---|---|---|---|---|
| 1 | 1 | 0 | 157984.0 | 1667.2 |
| 2 | 1 | 0 | 32571.8 | 836.2 |
| 3 | 1 | 0 | 28649.7 | 855.3 |
| 4 | 1 | 0 | 21480.3 | 767.4 |
| 5 | 1 | 0 | 8201.6 | 590.4 |
| 6 | 1 | 0 | 272377.2 | 3366.9 |
| ..... | | | | |



Données de
diffraction brutes

**Méthodes indirectes**
**= estimation**

# 1. Principe du phasage

## 1.4. Principe des méthodes de phasage

Les méthodes reposent sur une propriété remarquable de la série de Fourier : sa convergence

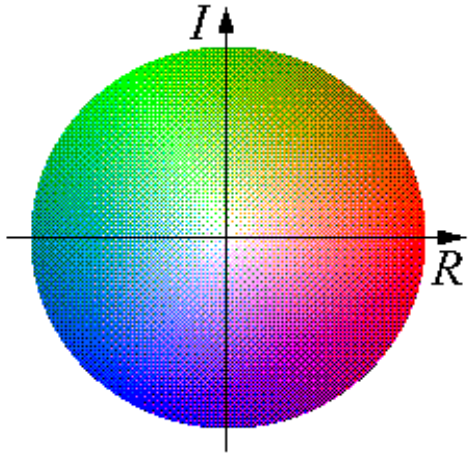
- Connaissant approximativement un certain nombre de phases, une première synthèse de Fourier fait apparaître des atomes de la structure
- On peut ainsi déduire de nouvelles phases et améliorer les phases de départ, une deuxième synthèse de Fourier fait apparaître d'autres atomes
- On recommence le cycle

# 1. Principe du phasage

## 1.5. L'importance des phases

http://www.yorvic.york.ac.uk/~cowtan/fourier/fourier.html

0 ° = rouge
120 ° = vert
240 ° = bleu

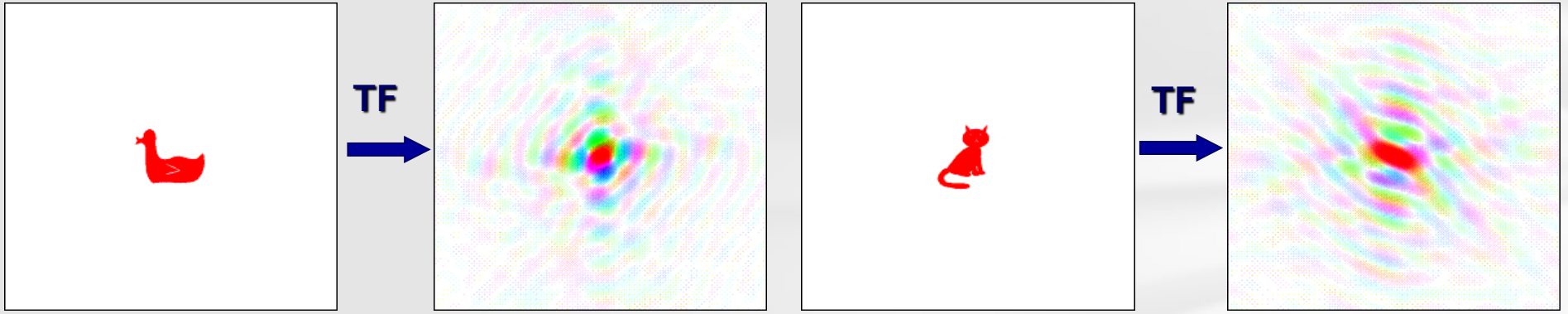



Module = intensité
Phase = couleur

Dr Kevin Cowtan
York Structural Biology Laboratory, Department Of Chemistry
University of York

# 1. Principe du phasage

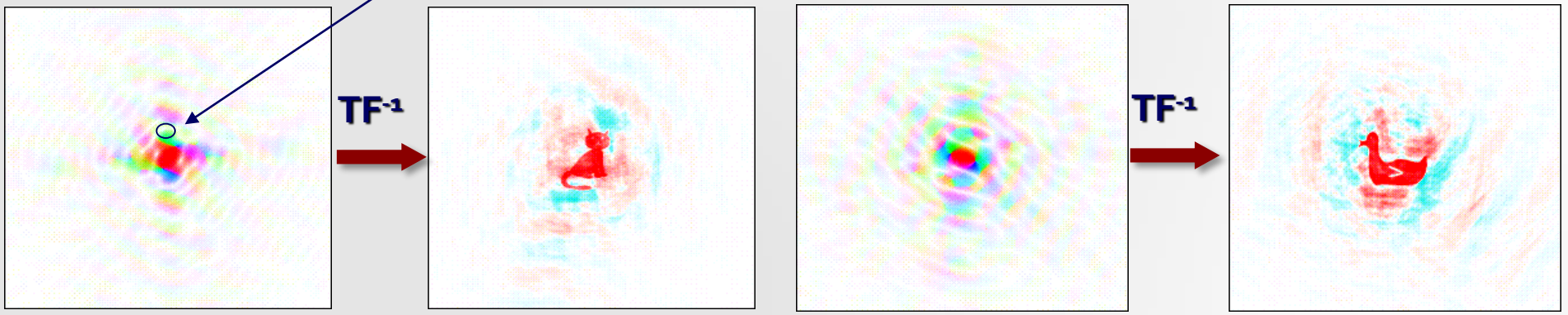
## 1.5. L'importance des phases




Modules du canard et phases du chat

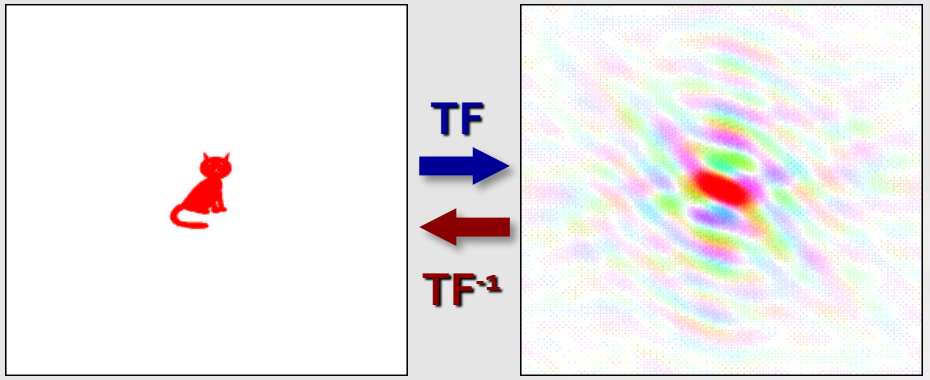
intensité canard
couleur chat

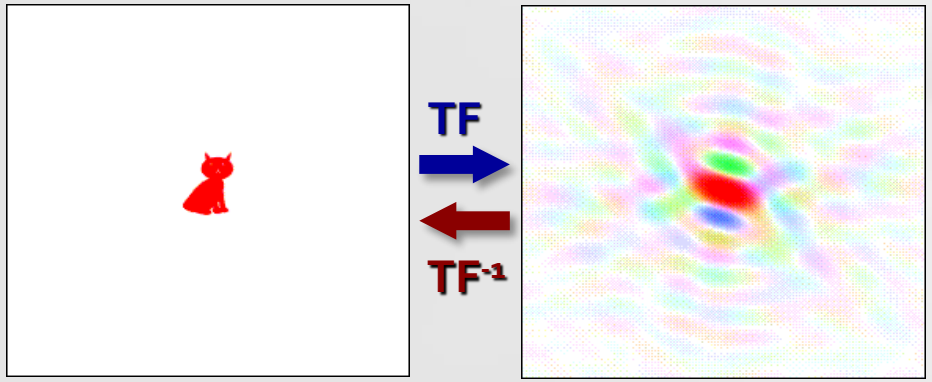
Modules du chat et phases du canard



# 1. Principe du phasage

## 1.5. L'importance des phases




2 objets sont similaires

Les TF sont similaires

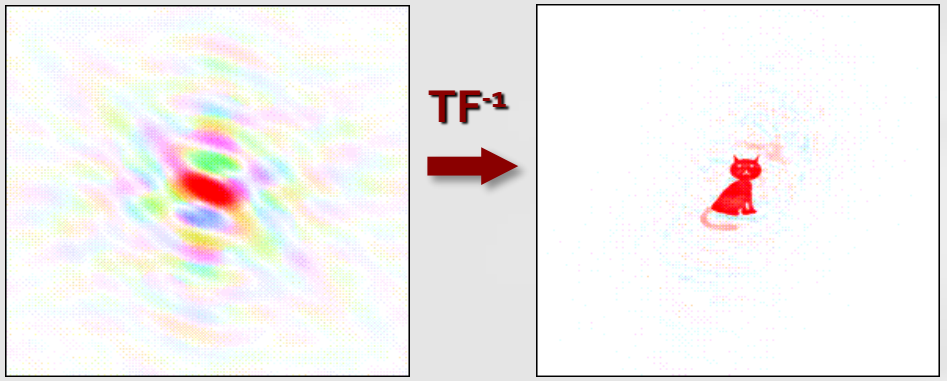
**Dans le cas de cristaux : il faut que les cristaux soient isomorphes**
(mêmes paramètres de maille, même groupe d'espace, même orientation dans l'ua)

Ce principe de base est utilisé :
- dans le phasage par remplacement moléculaire
- dans l'étape d'affinement

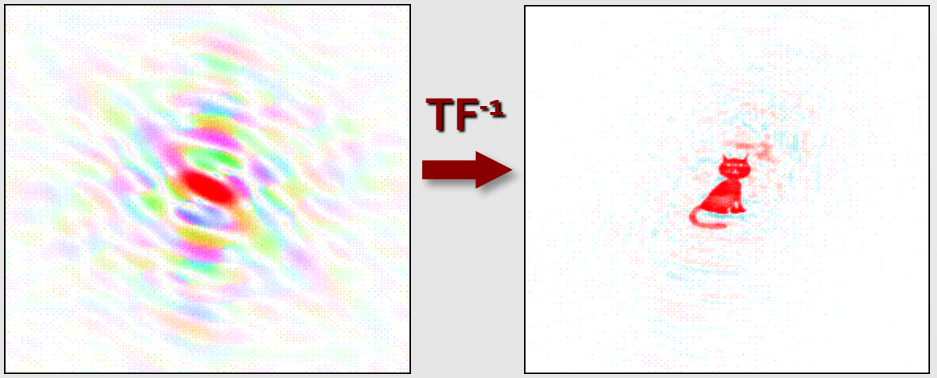
# 1. Principe du phasage

## 1.5. L'importance des phases

Lorsque 2 structures sont similaires, on peut combiner les modules




Modules du chat et
phases du Manx

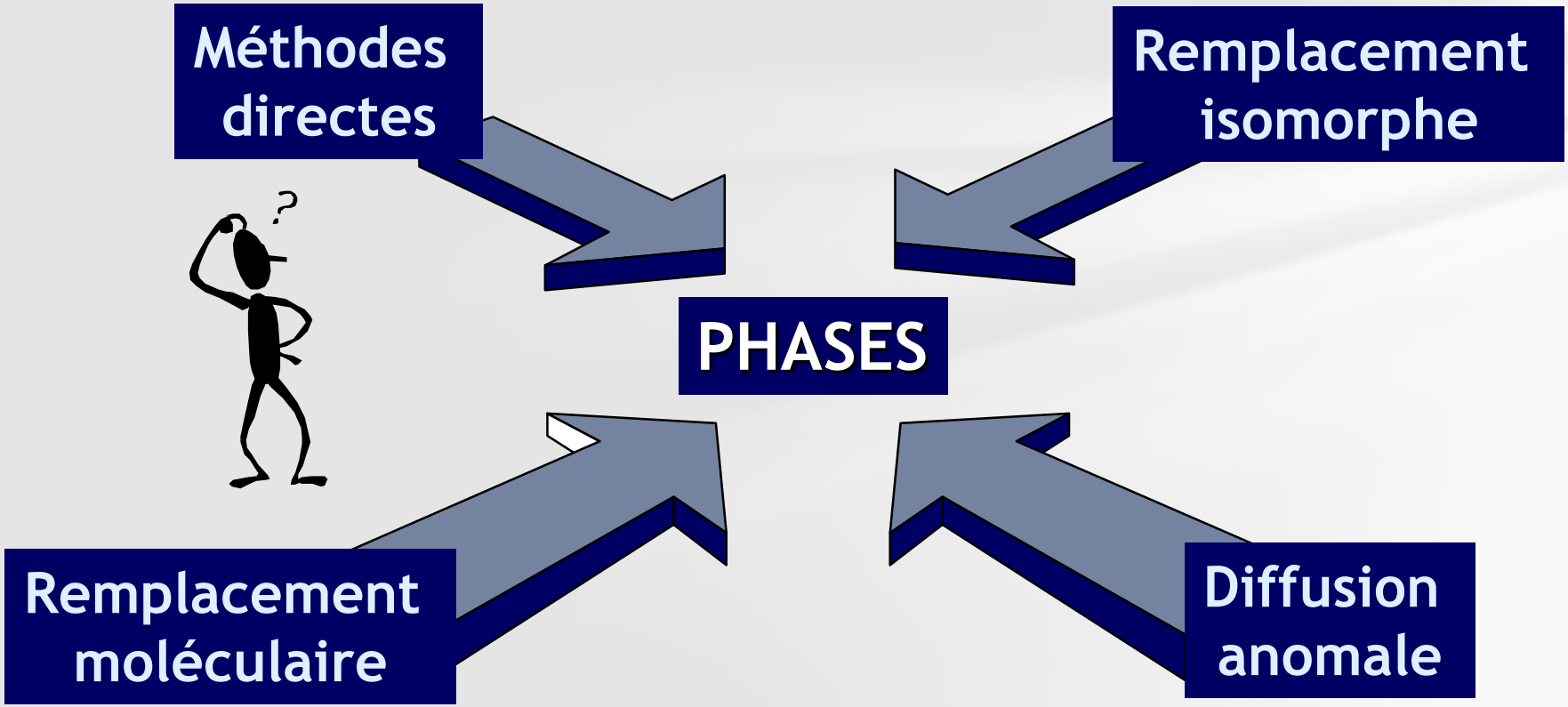



2 x M(chat) - M(manx) et
phases du Manx

M(chat) - M(manx) et
phases du Manx ??

# 1. Principe du phasage

## 1.6. Les différentes méthodes



# 1. Principe du phasage

## 1.6. Les différentes méthodes

**1. Les méthodes directes**
- petites molécules
- en phase de développement pour les macromolécules

**2. Le remplacement isomorphe**
- cristal natif et cristaux de dérivés lourds
- SIR, MIR, SIRAS

**PHASES**

**3. La diffusion anomale**
- cristal avec diffuseurs anomaux
- SAD, MAD

**4. Le remplacement moléculaire**
- nécessité d'une structure homologue disponible (PDB)
- peut être rapide

# 1. Principe du phasage

## 1.7. Choix de la méthode à utiliser

### Recherche d'homologues dans les banques protéiques

**OUI**

- **Remplacement moléculaire**
  succès de la méthode dépend du taux de ressemblance de la structure étudiée par rapport à la structure de la PDB

>50% des structures de la PDB

**NON**

- **Remplacement isomorphe**
- **Diffusion anomale**
- **Combinaison des 2 méthodes**

De plus en plus utilisées

# 1. Principe du phasage

## 1.8. Remarque sur les méthodes directes

Autre nom : méthodes *ab initio*

Méthode encore peu utilisée mais en développement

➡ Applicable pour :

- Des structures d'environ 1000 atomes (max 2000)
- Résolution atomique (d < 1.2 Å)
- Quelques cas à 2.0 Å de résolution

G. Bricogne « phasage *ab initio* »
*Methods in Enzymology*, vol 277 (1997)

Les phases sont directement déterminées à partir des amplitudes observées grâce à des relations mathématiques existant entre les phases (à partir de valeur initiales, déduction d'autres valeurs)

Ces relations existent si la densité électronique est :

(1) partout positive $\rho(r) \geq 0$ positivité

(2) composé d'atomes discrets atomicité

➡ Autre application

Détermination de la « sous-structure » des atomes lourds ou des diffuseurs anomaux

SHELXD SnB ACORN

# 2. LA FONCTION DE PATTERSON

# 2. La fonction de Patterson

## 2.1. Définition

→ **Après le traitement de données, on ne dispose pas des phases**

**Quelle est la transformée de Fourier des intensités mesurées ?**

$$TF(I_{h,k,l}) = TF(\vec{F}_{h,k,l} . \vec{F}_{-h,-k,-l}) = \rho(\vec{r}) * \rho(-\vec{r}) = \text{Fonction de Patterson}$$

$$P(uvw) = 1/v \sum_h \sum_k \sum_l |F_{hkl}|^2 \exp(-2\pi i (hu + kv + lw))$$

Or $|F_{hkl}|^2 = |F_{-h-k-l}|^2$

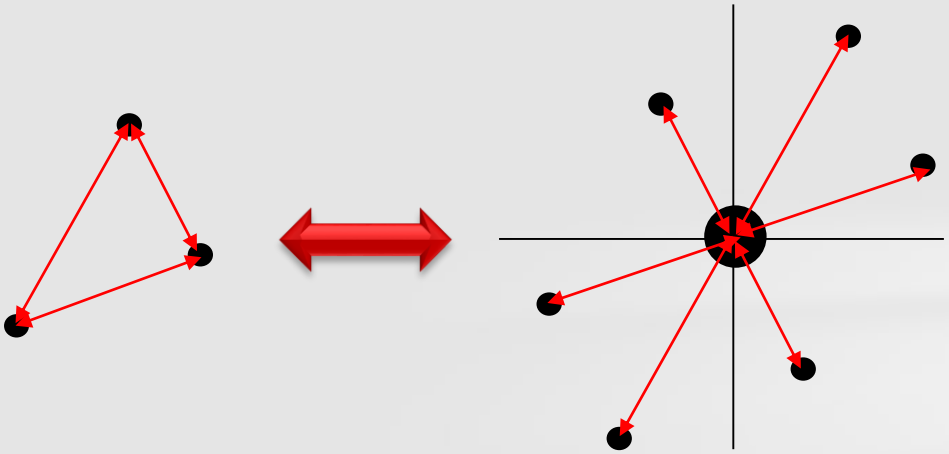
$$P(uvw) = 1/v \sum_h \sum_k \sum_l |F_{hkl}|^2 \cos(2\pi (hu + kv + lw))$$

**La fonction de Patterson est la transformée de Fourier des intensités**

**Fonction 3D dont les maxima correspondent à tous les vecteurs interatomiques définis sur les atomes de la maille (u,v,w)**

# 2. La fonction de Patterson

## 2.1. Définition



L'intensité des pics est proportionnel au numéro atomique des atomes dont provient ce pic ($Z_iZ_j$)

Nombre de pics dans la fonction de Patterson

| N atomes | Nombre total de pics | $N^2$ | 9 |
|---|---|---|---|
| | Nombre de pics à l'origine | N | 3 |
| | Nombre de pics « non-origine » | N(N-1) | 6 |
| | Nombre de pics uniques | N(N-1)/2 | 3 |

| Remarque | 1000 atomes | 499 500 pics !! |
|---|---|---|

# 2. La fonction de Patterson

## 2.1. Définition

Autre écriture

$$P(u,v,w) = \int_r \rho(x,y,z)\, \rho(x+u, y+v, z+w)\, dv$$

- La fonction P(u,v,w) a des valeurs significatives seulement si (u,v,w) est un vecteur interatomique
- Valeur moyenne, sur une période, du produit des densités électroniques séparées par un vecteur (u,v,w)
- Une valeur élevée de P(u,v,w) est observée si, quelque part dans la maille, des maxima de densité sont séparés par un vecteur (u,v,w)
- Inversement, P(u,v,w) atteint un maximum dès que le vecteur (u,v,w) relie 2 maxima de la densité électronique

# 2. La fonction de Patterson

## 2.1. Propriétés de la fonction de Patterson

(1) La fonction de Patterson est toujours centrosymétrique

$$U_{i,j} = -U_{j,i}$$

(2) La symétrie de la fonction de Patterson est fonction du groupe d'espace de la structure considérée

Exemple axe 2 $I(h,k,l) = I(-h,k,-l)$

$P(u,v,w) = P(-u,v,-w)$

(3) Les éléments de symétrie translatoire (axes hélicoïdaux) de la maille cristalline engendrent une symétrie de rotation

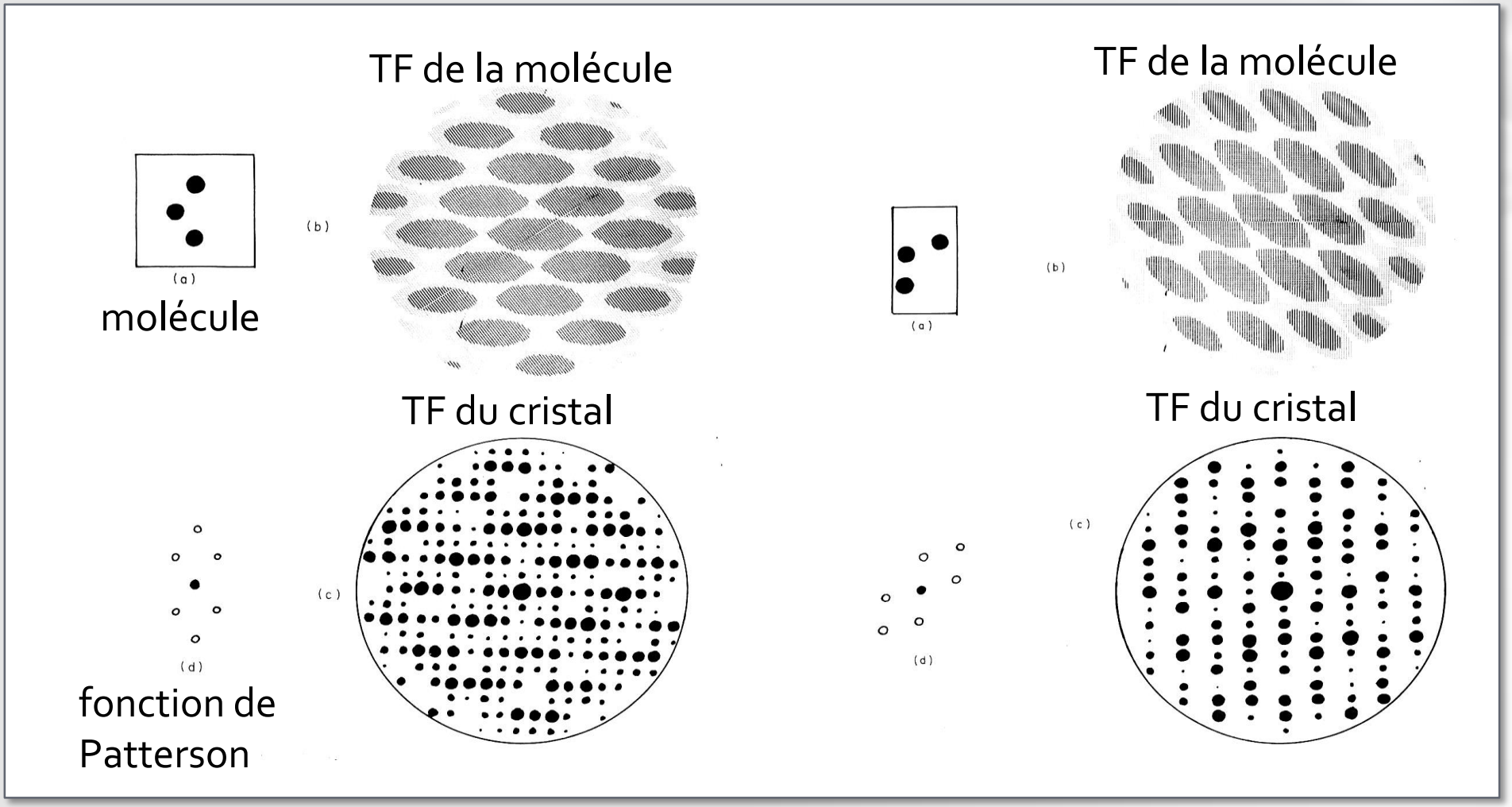
Exemple axe $2_1$ 2 points équivalents dans le RD

(x,y,z) et (-x,y+1/2,-z) $I(h,k,l) = I(-h,k,-l)$

$P(u,v,w) = P(-u,v,-w)$

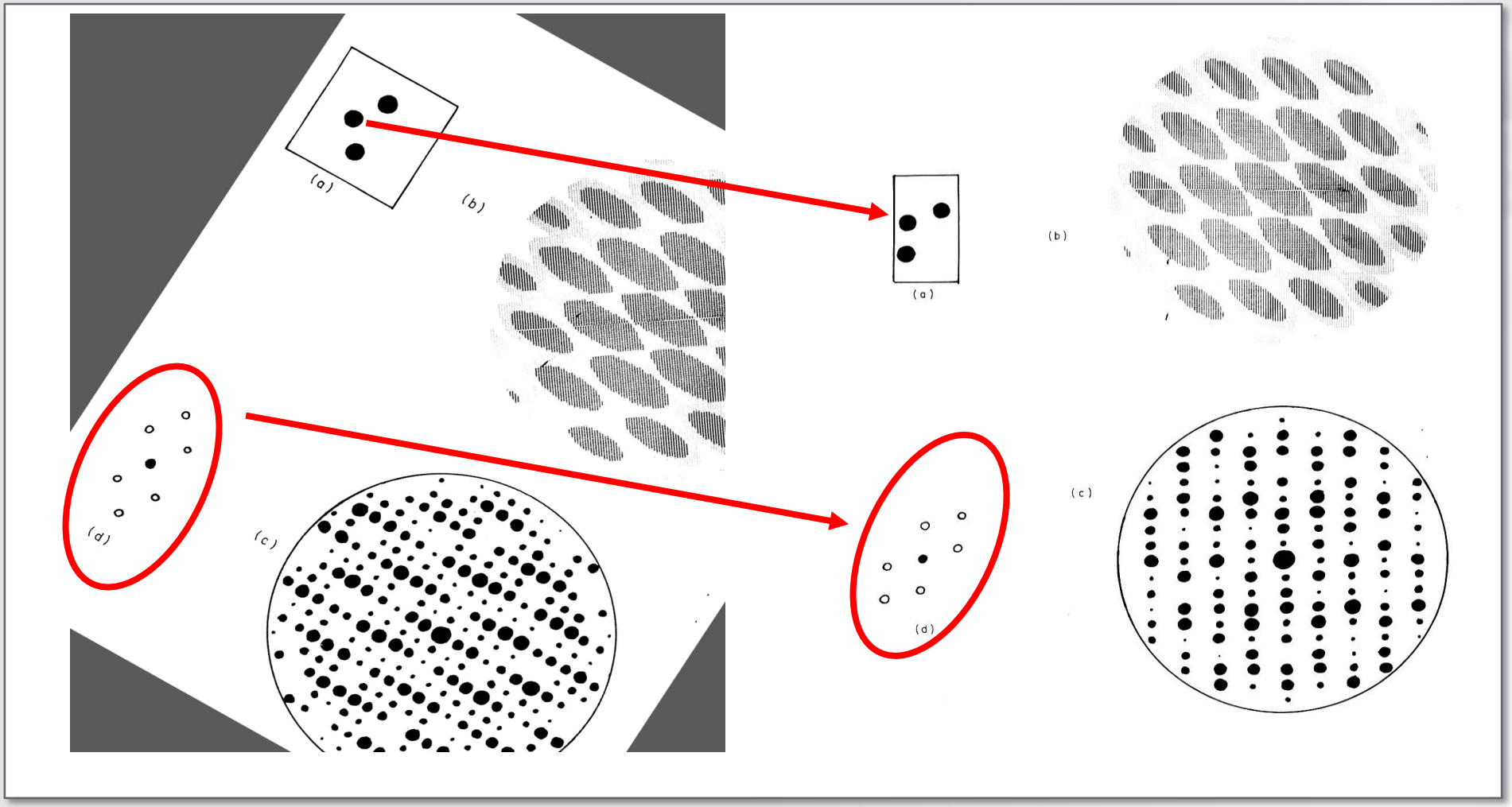
# 2. La fonction de Patterson

## 2.1. Propriétés de la fonction de Patterson



# 2. La fonction de Patterson

## 2.1. Propriétés de la fonction de Patterson



# 2. La fonction de Patterson

## 2.1. Applications de la fonction de Patterson

### (1) Patterson native

On observe un pic dès que 2 molécules sont reliées par translation

- mode de réseau

  exemple mode C  on observe un pic en ½ ½ 0
- si 2 molécules de l'unité asymétrique sont reliées par une translation

### (2) Patterson isomorphe $F^2_{PH} - F^2_P = \Delta F_{iso}$

Localisation des positions des atomes lourds dans la maille (IV.5.1)

### (3) Patterson anomale $(F+ - F-)^2 = \Delta F_{ano}$

Localisation des positions des atomes anomaux dans la maille (IV.5.2)

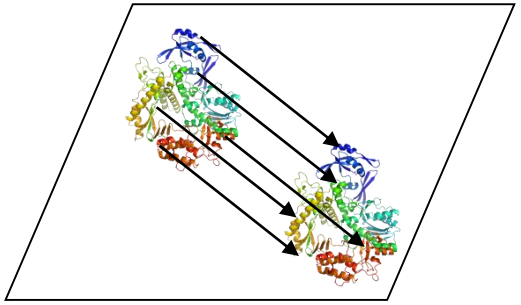
### (4) Fonction de rotation *self-rotation* et *cross-rotation*

Localisation des symétries de rotation (direction)

# 2. La fonction de Patterson

## 2.1. Applications de la fonction de Patterson

### (1) Patterson native

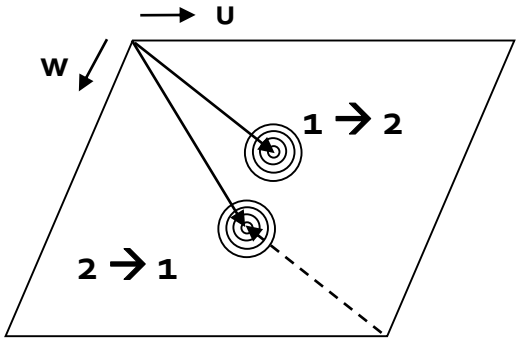


Maille du cristal

2 molécules reliées par translation

⟹ grand nombre de vecteurs tous identiques

2 pics symétriques correspondant à la translation




Carte de Patterson

# 2. La fonction de Patterson

## 2.1. Applications de la fonction de Patterson

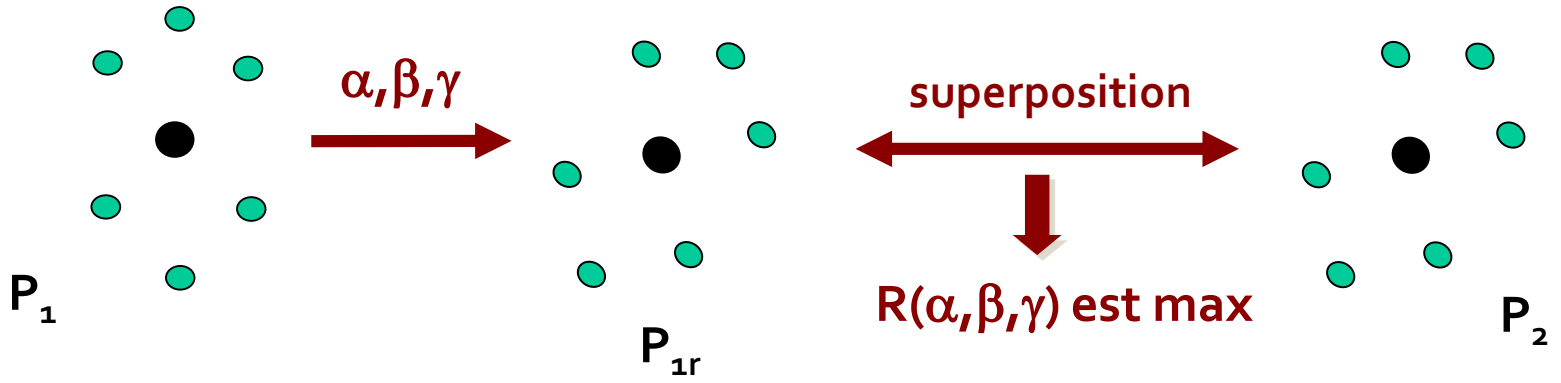
### (4) Fonction de rotation

$$R(\alpha,\beta,\gamma) = \int_U P_2(u)\, P_{1r}(u_r)\, du$$

P(u) fonction de Patterson

$P_r(u_r)$ fonction de Patterson après rotation

**La fonction $R(\alpha,\beta,\gamma)$ présente un maximum lorsque les 2 fonctions de Patterson se superposent**




**(1) Fonction de self-rotation**

$P_1$ et $P_2$ calculées à partir du même cristal

**(2) Fonction de cross-rotation**

$P_1$ calculée à partir du cristal 1

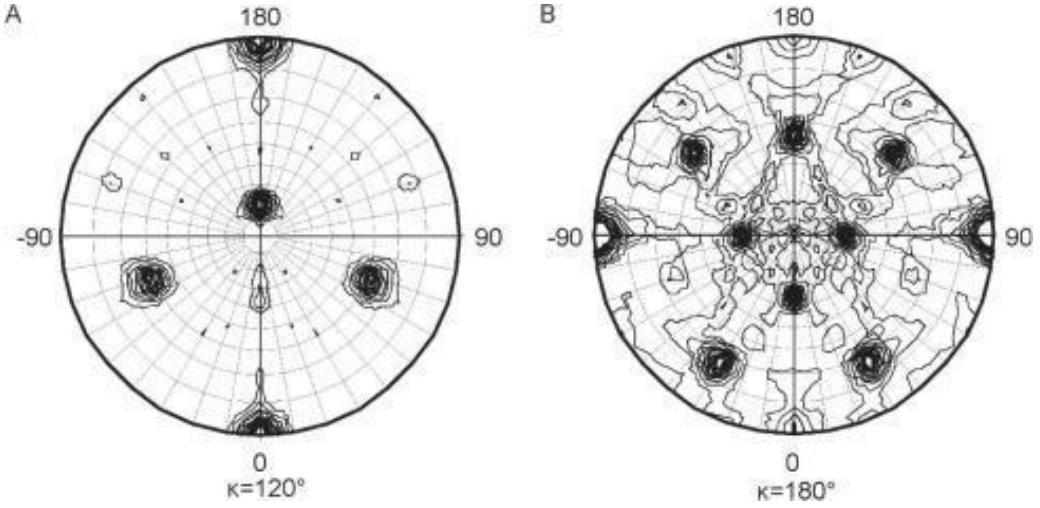
$P_2$ calculée à partir du cristal 2

# 2. La fonction de Patterson

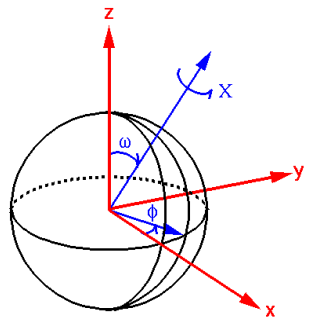
## 2.1. Propriétés de la fonction de Patterson

### (4) Fonction de rotation

**Self-rotation**




**Représentation par des projections stéréographiques**




**ω et Φ positionnent l'axe**
**K définit l'angle de rotation**

**Localisation des axes de symétrie**

Un code d'orthogonalisation positionne (xyz) par rapport à la maille

(x y z) selon (a, c* Λ a, c*)

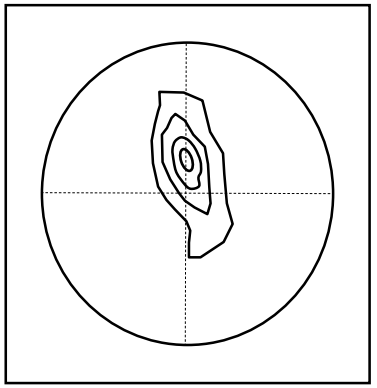
**Cross-rotation** **Remplacement moléculaire : orientation du modèle**

# 2. La fonction de Patterson

## 2.1. Propriétés de la fonction de Patterson

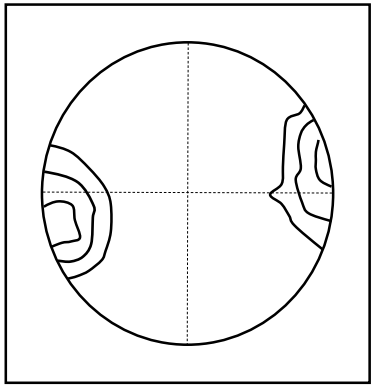
### (4) Fonction de rotation

Exemple



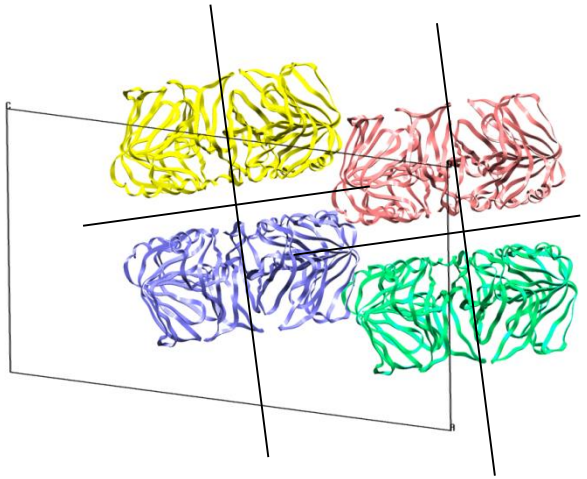
Kappa = 180 °

Axe 2



Axe 7

Kappa = 52 °



**Groupe d'espace P1**
**4 molécules dans l'unité asymétrique**

# 3. LE REMPLACEMENT ISOMORPHE

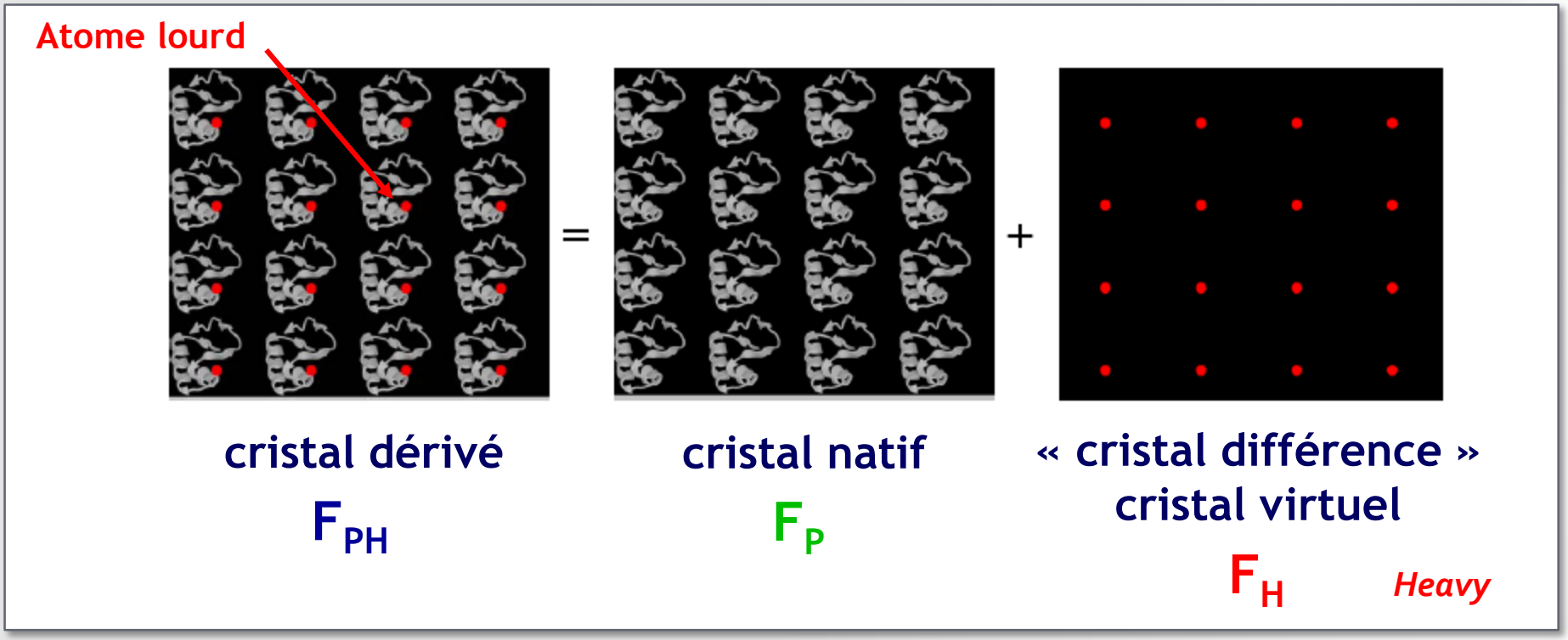
# 3. Le remplacement isomorphe

## 3.1. Principe de la méthode

**Il est basé sur le changement d'intensité causé par la présence d'un atome lourd dans la molécule ($F_H$, $\varphi_H$)**

**La méthode consiste à « marquer » la macromolécule dans le cristal par des atomes lourds**

- Brome ou iode pour les acides nucléiques
- Gaz rares (Xe, Kr), Plomb, mercure, lanthanides… pour les protéines

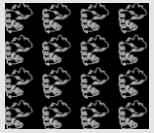


# 3. Le remplacement isomorphe

## 3.1. Principe de la méthode

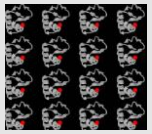
**Hypothèse de base** les cristaux de la structure native et du dérivé lourd doivent être **isomorphes**

Pour chaque réflexion hkl



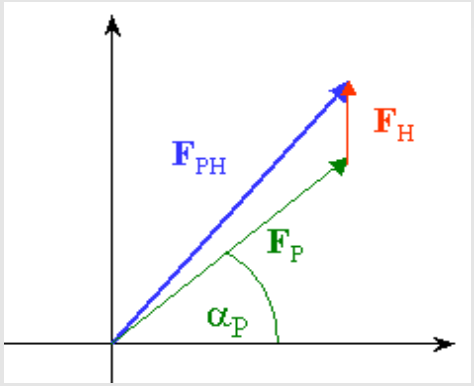
$$\vec{F}_{hkl} = \underbrace{\Sigma_P f_P \, e^{2\pi i (hx_P + ky_P + lz_P)}}_{|F_P| \, e^{i\phi_P}}$$

$\vec{F}_P$



$$\vec{F}_{hkl} = \underbrace{\Sigma_P f_P \, e^{2\pi i (hx_P + ky_P + lz_P)}}_{|F_P| \, e^{i\phi_P}} + \underbrace{\Sigma_H f_H \, e^{2\pi i (hx_H + ky_H + lz_H)}}_{|F_H| \, e^{i\phi_H}}$$

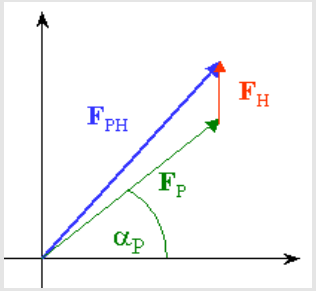
$\vec{F}_{PH} = \vec{F}_P + \vec{F}_H$




$F_{PH}$ = facteur de structure du cristal dérivé
$F_H$ = facteur de structure des atomes lourds
$F_P$ = facteur de structure de la protéine

# 3. Le remplacement isomorphe

## 3.2. Principe de la détermination des phases

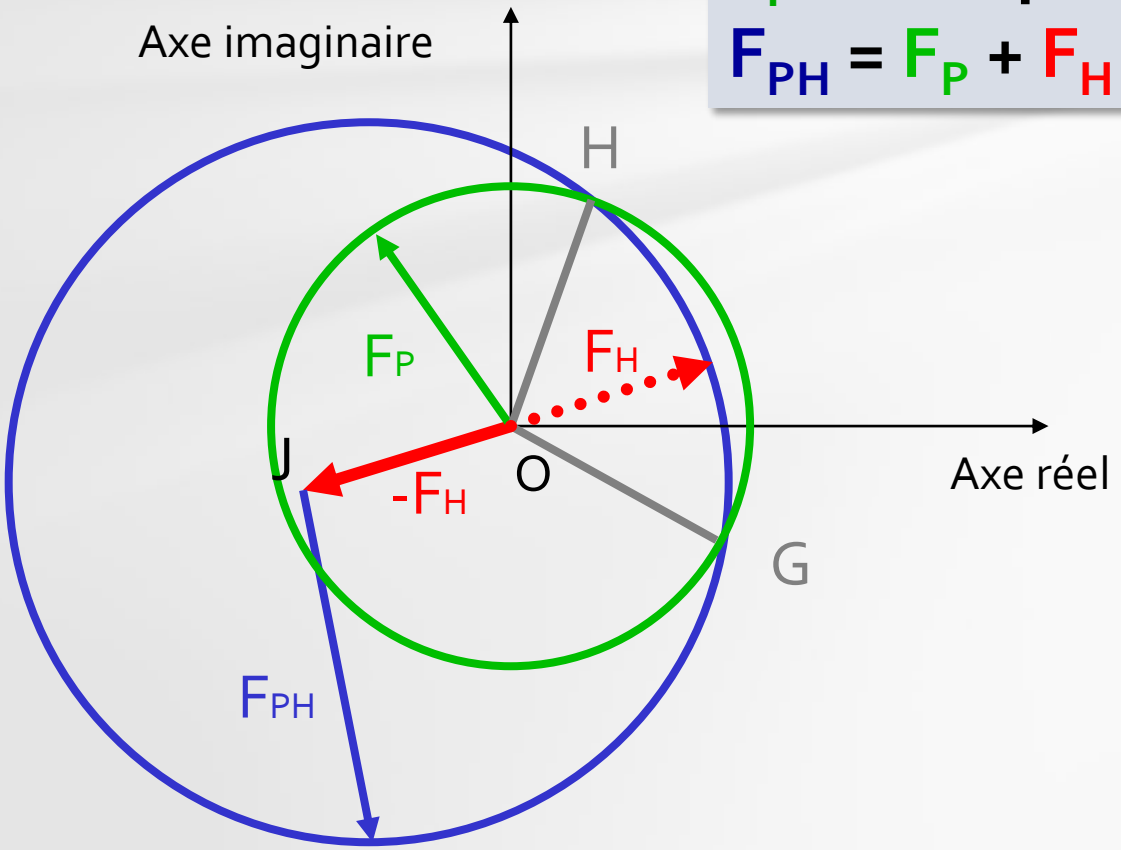



On dispose de

- module de $F_P$
- module et phase de $F_H$
- module de $F_{PH}$

**Construction de Harker**

$F_P$ ? tels que
$F_{PH} = F_P + F_H$



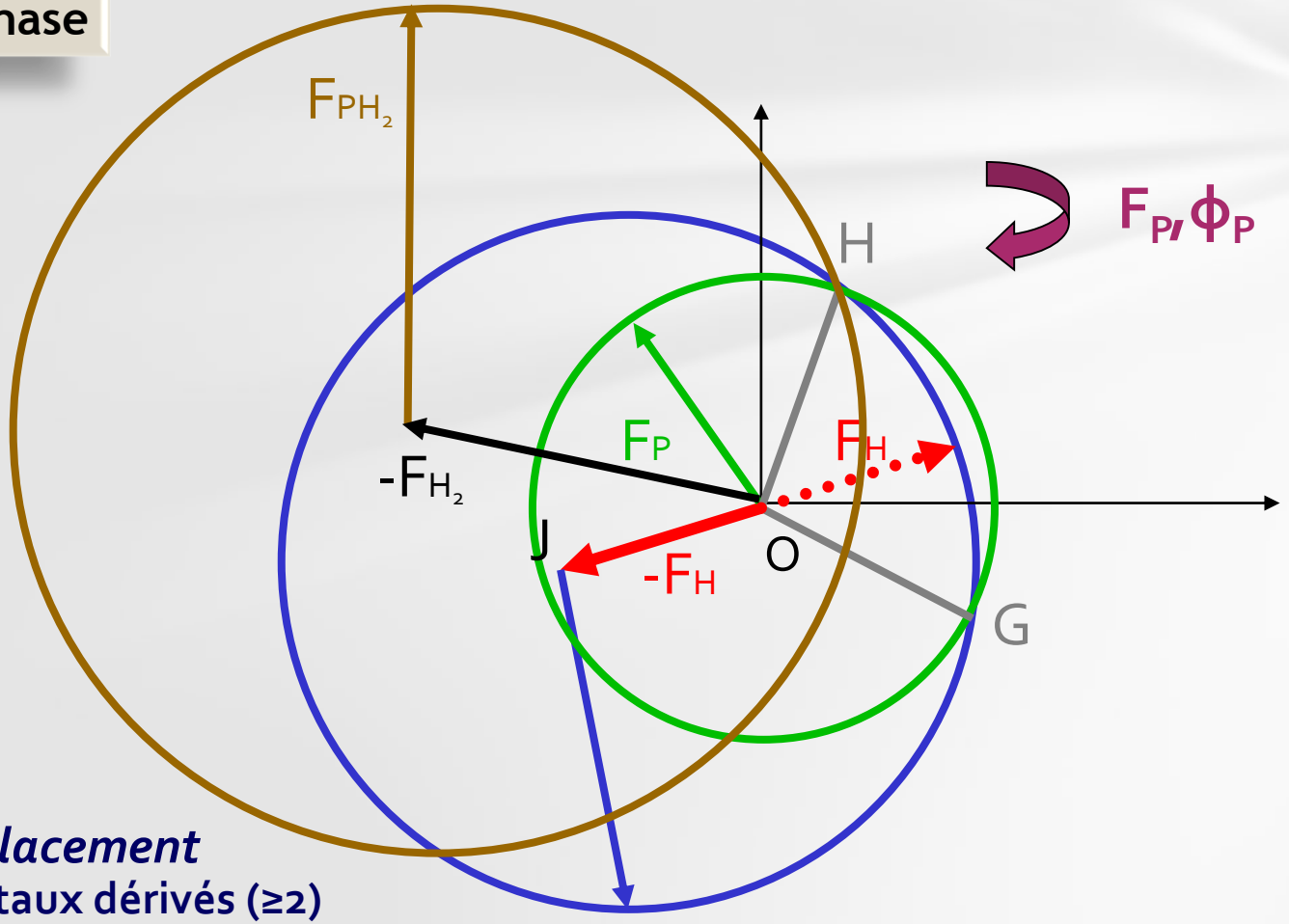

**SIR**

***Single Isomorphous Replacement***
**1 cristal natif + 1 seul cristal dérivé**

# 3. Le remplacement isomorphe

## 3.2. Principe de la détermination des phases

**Avec un deuxième dérivé, on lève l'ambiguïté de phase**




**MIR**

***Multiple Isomorphous Replacement***

**1 cristal natif + plusieurs cristaux dérivés (≥2)**

# 3. Le remplacement isomorphe

## 3.3. Les différentes étapes

- Préparation des cristaux de dérivés lourds
- Enregistrement espaces complets de diffraction
  - de la structure native
  - du (des) dérivé(s) lourd(s)
- Mise à l'échelle des jeux natif/dérivé(s)
- Détermination de la position des atomes lourds
- Affinement des paramètres (x, y ,z ,q ,B) des atomes lourds
- Calcul des phases (probabilité de phases)
- Calcul d'une carte de densité électronique

# 3. Le remplacement isomorphe

## 3.4. Obtention des dérivés lourds

- Par trempage des cristaux
  Les composés lourds diffusent dans le cristal à travers les canaux de solvant
  Ils peuvent se fixer sur des sites précis
  - Ex : composés Hg réagissent avec les Cys accessibles et réduites à pH neutre

  Cette étape est souvent longue et fastidieuse

- Co-cristallisation

**cf. COURS ALAIN**

- Modification de la macromolécule
  - Brome ou iode pour les acides nucléiques
  - Sélénométhionine

cf diffusion anomale

Pike AC et al, An overview of heavy-atom derivatization of protein crystals *Acta Cryst D*, 72, 303-18 (2016)

# 3. Le remplacement isomorphe

## 3.5. Défaut d'isomorphisme

**L'incorporation de l'atome lourd ne doit pas trop perturber l'empilement cristallin**

$$R_{iso} = \frac{\left\langle \left\| F_{PH} \right| - \left| F_P \right\| \right\rangle}{\left\langle \left| F_P \right| \right\rangle} = \sqrt{\frac{N_H}{2}} \frac{f_H}{\left\langle \left| F_P \right| \right\rangle}$$

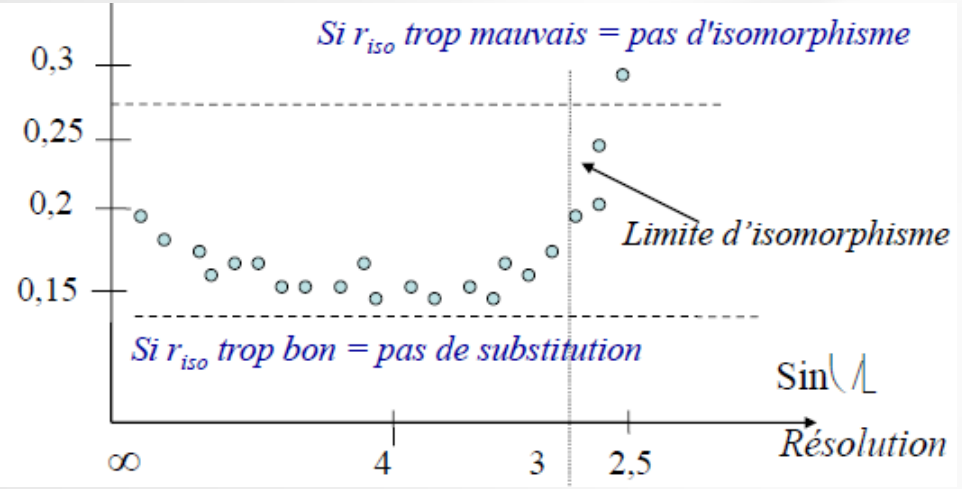
$$\left\langle \frac{\Delta I}{I} \right\rangle = \sqrt{\frac{N_H}{2N_P}} \frac{f_H}{f_P}$$

Pour 1000 atomes, l'incorporation d'un atome Hg change l'intensité de 25%
Mais un changement d'isomorphe par un changement de 0.5% des paramètres de mailles ou de la rotation de 0.5° de la molécule dans l'unité asymétrique change les intensités de 15% ...

Pour une protéine, une valeur approchée de $<|F_P|>$ en e⁻ peut être calculée par:

$\frac{6{,}7 \times \sqrt{NA}}{\sqrt{346 \times NR}}$ ou $\sqrt{3{,}14 \times MM}$

NA: nombre atomes
NR: nombre de résidus
MM: masse moléculaire




**Le défaut d'isomorphisme augmente avec la résolution**

**En général : Le cristal dérivé est isomorphe dans la zone de résolution où $R_{iso}$ reste constant**

# 3. Le remplacement isomorphe

## 3.6. Détermination de la position des atomes lourds

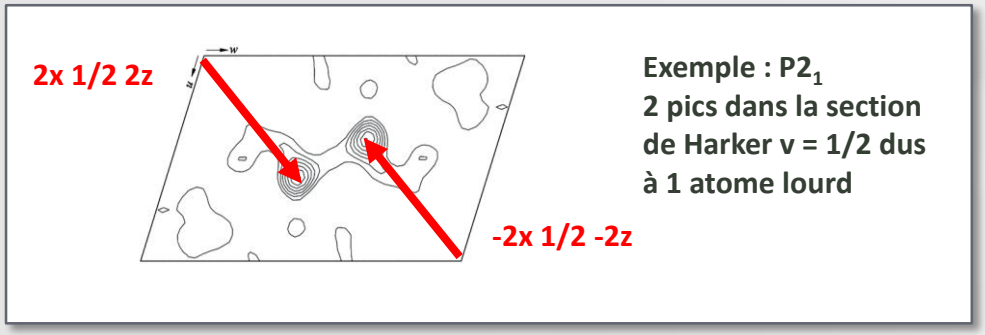
### Fonction de Patterson

$$P(uvw) = \Sigma (\Delta |F|_{iso})^2 \cos (2\pi (hu + kv + lw))$$

avec $\Delta |F|_{iso} = |F_{PH}| - |F_P|$ $(\Delta |F|_{iso})^2 \sim |F_H|^2$

Calculée sur les données de diffraction issues de 2 cristaux (natif et dérivé)

Fonction de Patterson définissant une carte dont les pics sont les vecteurs interatomiques entre atomes lourds « uniquement »




Exemple : $P2_1$
2 pics dans la section de Harker v = 1/2 dus à 1 atome lourd

**Section de Harker**

### Méthodes directes

- Calcul des phases (probabilité de phases)
- Calcul de la densité électronique

**SHELXD**

# 3. Le remplacement isomorphe

## 3.6. Détermination de la position des atomes lourds

**Section de Harker**

Les éléments de symétrie de la maille cristalline peuvent engendrer des pics dans certains plans ou sections particulières

➡ **Coordonnées des vecteurs interatomiques entre 2 mêmes atomes reliés par symétrie**

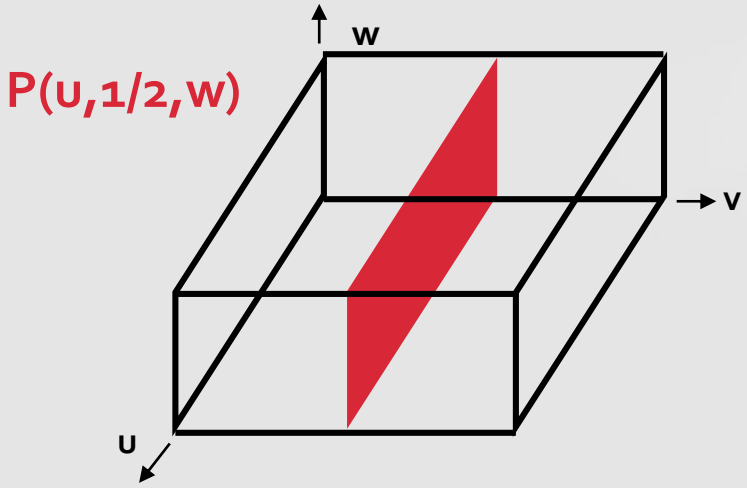
Exemple : **groupe d'espace $P2_1$ (2 positions équivalentes)**

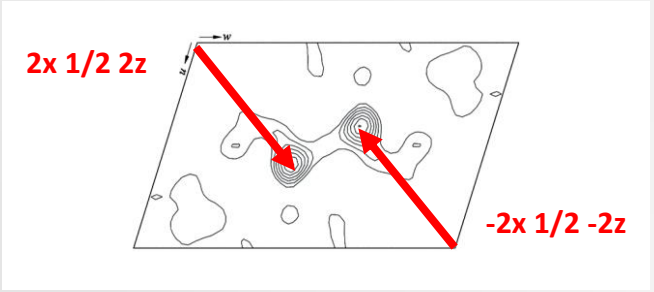
| | x y z | -x y+1/2 -z |
|---|---|---|
| x y z | 0 0 0 | -2x +1/2 -2z |
| -x y+1/2 -z | 2x -1/2 2z | 0 0 0 |

**U = ±2X V = ±1/2 W ± 2Z**

**Les pics correspondants appartiennent tous à un plan**

**V = 1/2** **SECTION DE HARKER**




**2 pics**

**1 atome**

**Carte des vecteurs interatomiques en V = +1/2**

# 3. Le remplacement isomorphe

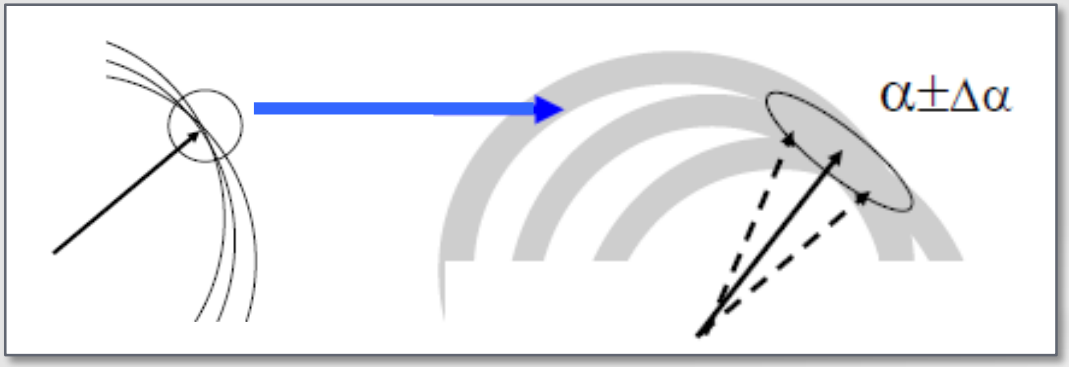
## 3.6. Détermination de la position des atomes lourds

La section de Harker dépend du groupe d'espace

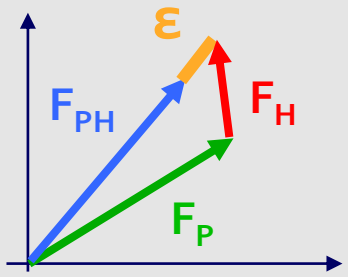
| | |
|---|---|
| pas de symétrie | pas de section de Harker |
| axe de rotation selon b | section de Harker en v = 0 |
| axe de rotation selon c | section de Harker en w = 0 |
| axe hélicoïdal $2_1$ selon b | section de Harker en v = 1/2 |
| axe hélicoïdal $3_1$ selon c | section de Harker en w = 1/3 |
| axe hélicoïdal $4_1$ selon c | section de Harker en w = 1/4 |
| etc … | |

# 3. Le remplacement isomorphe

## 3.7. Probabilité de phases




Sources d'erreurs
- Paramètres (x,y,z,q,B) des atomes lourds
- Amplitudes des F mesurés ($F_P$ et $F_{PH}$)
- Mise à l'échelle
- Défaut d'isomorphisme




**Erreur de fermeture** **hypothèse : $F_{PH} + \varepsilon$**

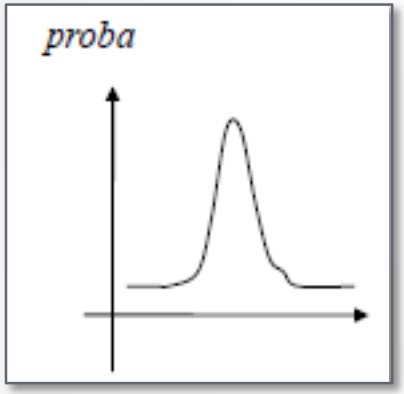
$$\varepsilon = |F_{PHobs}| - |F_{PHcalc}| = |F_{PHobs}| - |[|F_P|e^{i\phi_P} + |F_H|e^{i\phi_H}]$$

**Les erreurs sont traitées avec des méthodes statistiques**

Distribution de **probabilité de phase**

$$P(\alpha) = Ne^{\frac{-\varepsilon^2}{2E^2}} \quad \text{avec} \quad E = \left\langle \left| F_{PH(obs)} - F_{PH(calc)} \right|^2 \right\rangle$$

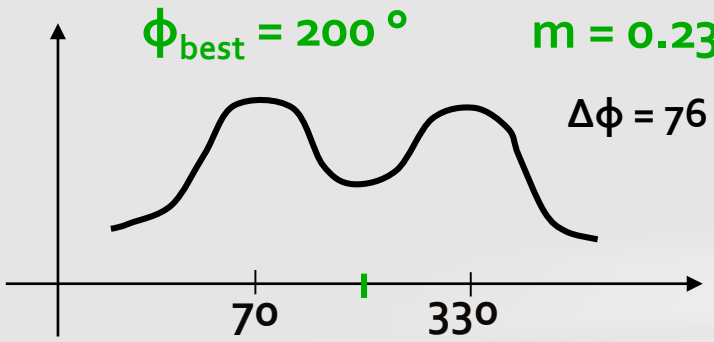
Hypothèse



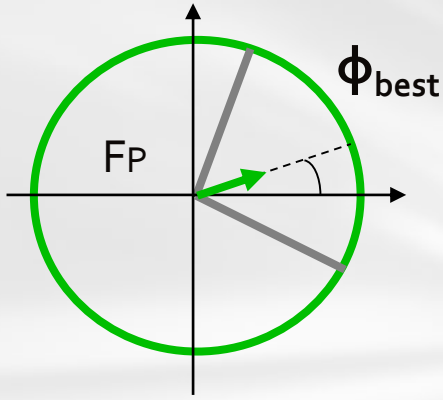
# 3. Le remplacement isomorphe

## 3.7. Probabilité de phases

**SIR**



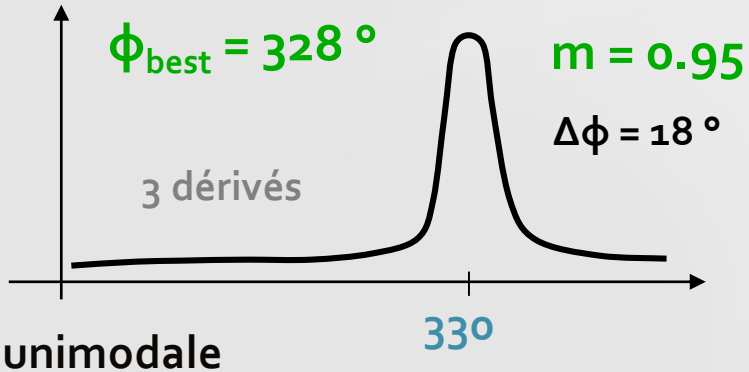

$m = \cos \Delta\phi$




La qualité de l'information sur chaque phase est mesurée par la figure de mérite (m ou fom) qui est le reflet de l'étroitesse de la distribution de probabilité

- m = 1 : distribution étroite
- m = 0 : distribution plate

**MIR**




Distribution unimodale

les probabilités de phase des différents dérivés sont multipliées

Calcul de la carte de densité électronique contenant le moins d'erreur

- module $F_{best} = m F_P$
- phase $\phi_{best}$

# 4. LA DIFFUSION ANOMALE

# 4. La diffusion anomale

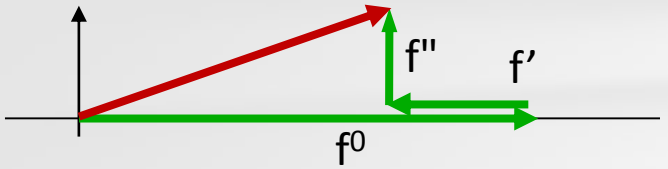
## 4.1. Principe de la méthode

Il est basé sur la modification de la diffraction lorsqu'un ou plusieurs diffuseurs anomaux sont présents dans le cristal

Différence d'intensité d'une paire de Bijvoet lorsque la loi de Friedel n'est plus respectée

Correction par rapport à la diffusion d'un électron libre

*Voir diapo suivante*



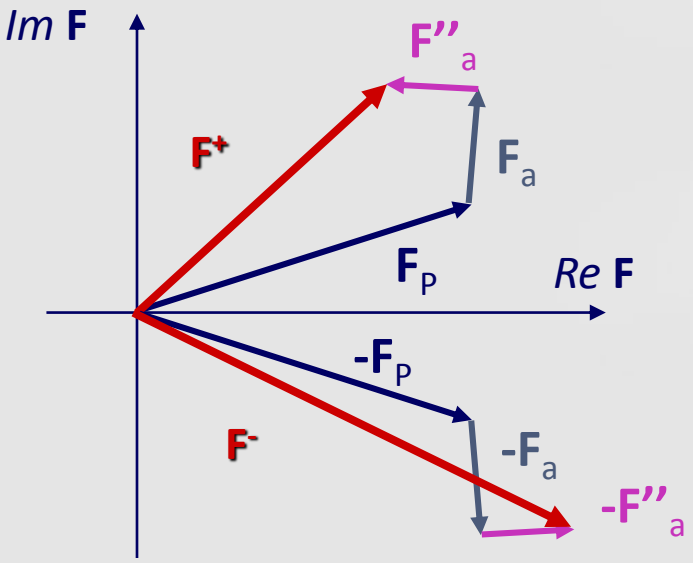

facteur de diffusion $f_j$

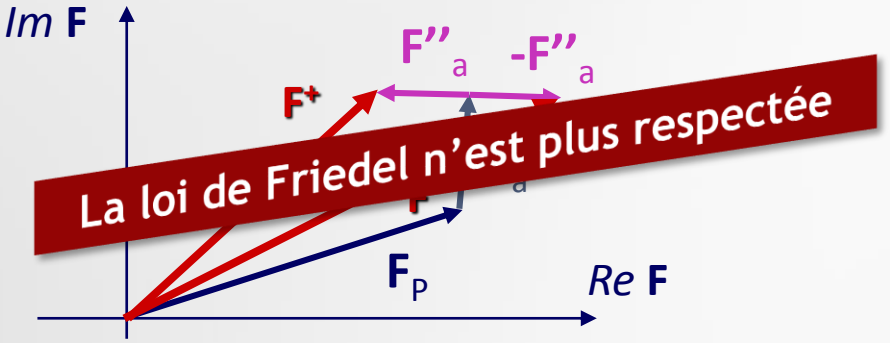
**Pour les atomes anomaux**

$f_a = f°_a + f'_a + i\, f''_a$

**On pose $f_a = f°_a + f'_a$**

$$F = \Sigma f_P\, e^{i\phi P} + \Sigma(f_a + if''_a)\, e^{i\phi a}$$
$$F = \Sigma f_P\, e^{i\phi P} + \Sigma f_a\, e^{i\phi a} + i\Sigma f''_a\, e^{i\phi a}$$
$$F = F_p + F_a + iF''_a$$



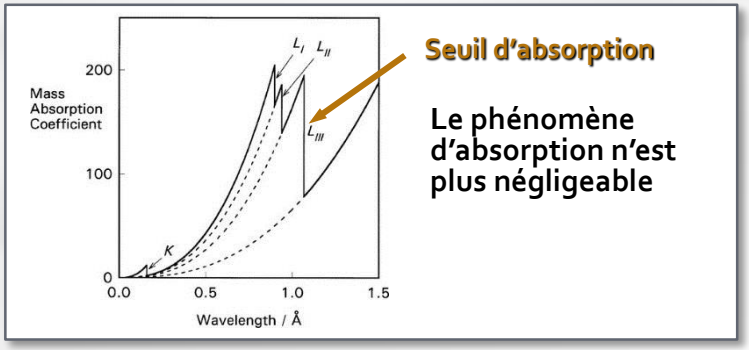
# 4. La diffusion anomale

## 4.1. Principe de la méthode

Dans la plupart des cas, f' et f '' sont négligeables (modèle de l'électron libre)

SAUF lorsque $\lambda_{incident} \approx \lambda_{seuil}$ cad aux longueurs d'ondes correspondant à l'énergie nécessaire pour extraire des électrons d'une couche électronique donnée (K, $L_I$, $L_{II}$, $L_{III}$)

**Pour $\lambda_{incident} \approx 1$ Å, seuls les atomes lourds présentent une absorption non négligeable**



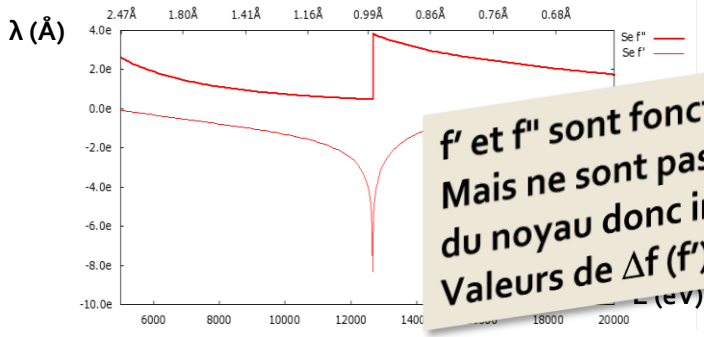

**Seuil d'absorption**

Le phénomène d'absorption n'est plus négligeable

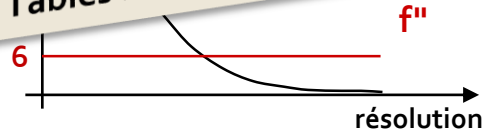
**Le facteur de diffusion devient un nombre complexe**

$$f_a = f^\circ_a + f'_a + i\, f''_a$$

Variation de f en fonction de la longueur d'onde



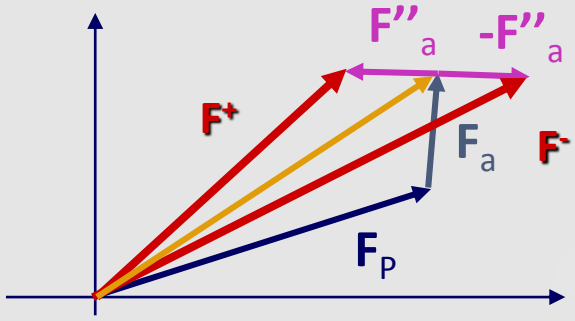

Variation de f en fonction de la résolution




f' et f'' sont fonction de la longueur d'onde
Mais ne sont pas fonction de la résolution (concerne les e⁻ proches du noyau donc indépendants de sinθ)
Valeurs de Δf (f') et f'' tabulées dans les Tables Internationales

# 4. La diffusion anomale

## 4.2. Principe de la détermination des phases

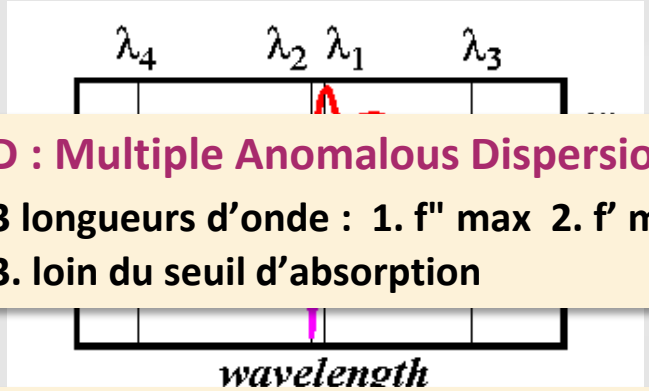



Les termes dépendants et indépendant de la longueur d'onde peuvent être séparés

$$\vec{F}_T(\lambda) = \vec{F}_T(h) + \vec{F}_A(\lambda)$$

**Le problème des phases a une solution analytique**

Résolution de N équations à 4 inconnues à plusieurs longueurs d'ondes

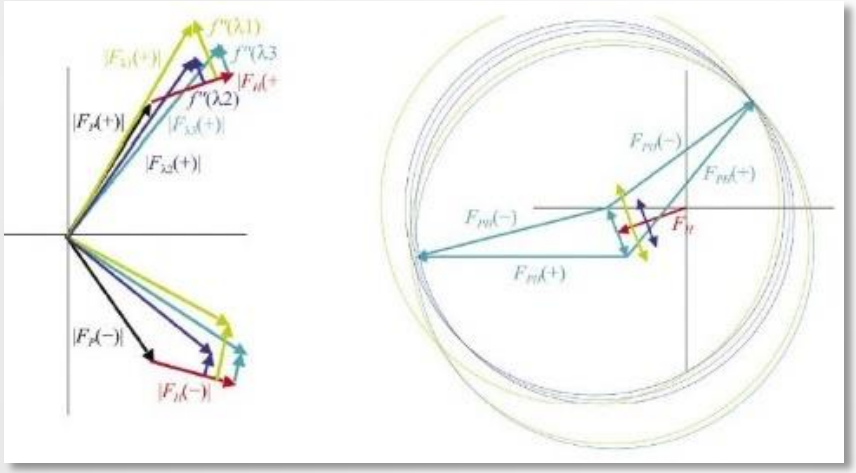



**MAD : Multiple Anomalous Dispersion**

3 longueurs d'onde : 1. f" max 2. f' max
3. loin du seuil d'absorption

**SAD : Single Anomalous Dispersion**

1 longueur d'onde (f" maximal)



# 4. La diffusion anomale

## 4.3. Les différentes étapes

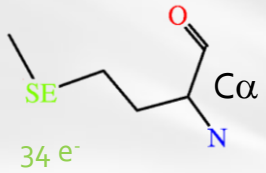
- **Production et cristallisation de la protéine modifiée**
- **Détermination de la longueur d'onde pour laquelle le signal anomal est maximal**

  Mesure d'un spectre de fluorescence du cristal
- **Enregistrement d'espaces complets ET redondants**
  - A partir d'un seul cristal
- **Détermination de la position des diffuseurs anomaux**
- **Affinement des paramètres des diffuseurs anomaux**
- **Calcul des phases sous la forme de probabilité de phases**
- **Calcul d'une carte de densité électronique**

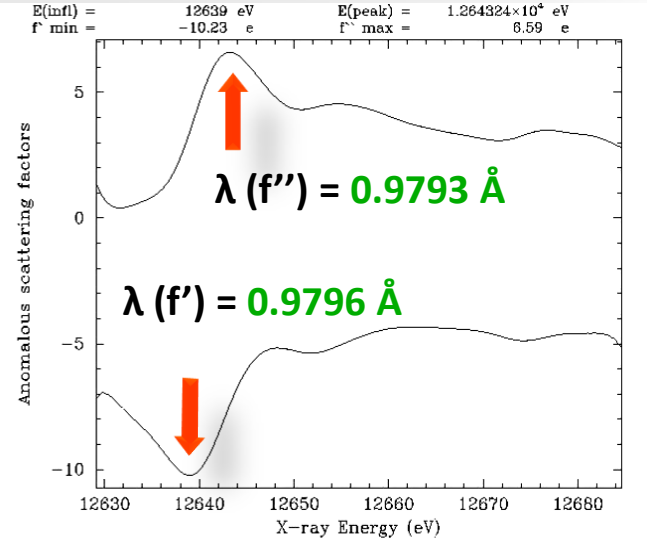
# 4. La diffusion anomale

## 4.4. Méthodologie

- Introduction de diffuseurs anomaux dans le cristal ou directement dans la protéine
  - Atomes lourds
  - Développements de cette technique : utilisation de la sélénométhionine




- Mesure d'un spectre de fluorescence
  - Peak, inflection




$\lambda_{peak}$

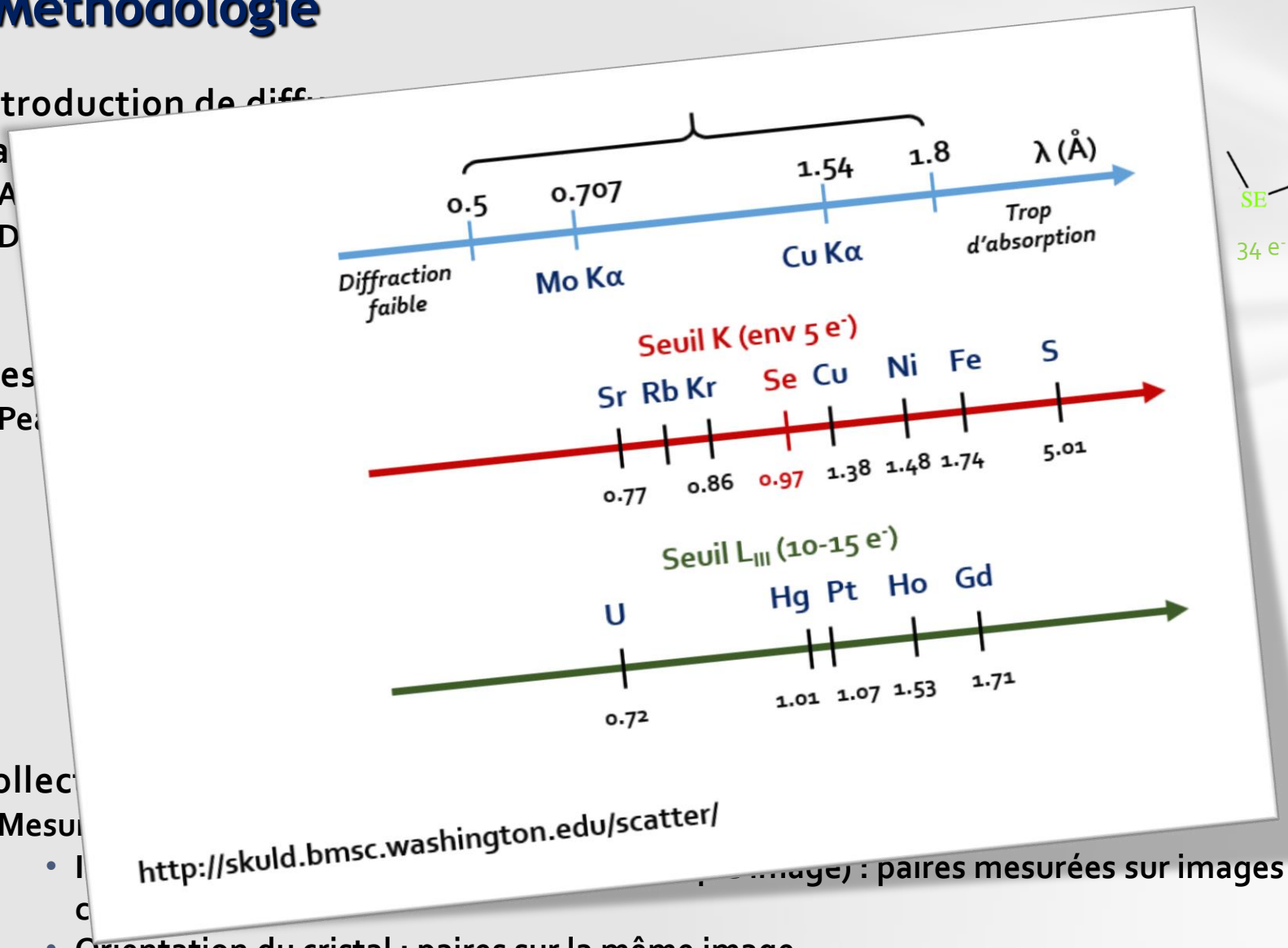
$\lambda_{inflection}$

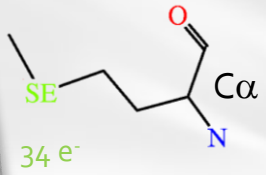
- Collecte des données anomales
  - Mesurer avec précision les paires de Bijvoet
    - Inverse beam (rotation de 180° entre chaque image) : paires mesurées sur images consécutives
    - Orientation du cristal : paires sur la même image

# 4. La diffusion anomale

## 4.4. Méthodologie

- Introduction de diff
  da
  - A
  - D
- Mes
  - Pe
- Collec
  - Mesu
    - I
    - Orientation du cristal : paires sur la même image



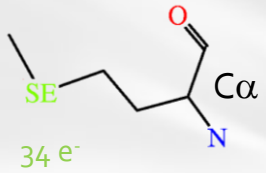

$\lambda_{peak}$

$\lambda_{inflection}$

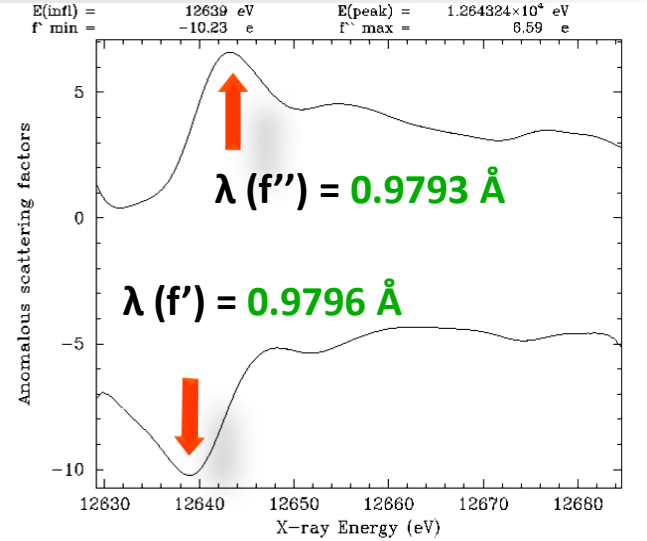
# 4. La diffusion anomale

## 4.4. Méthodologie

- Introduction de diffuseurs anomaux dans le cristal ou directement dans la protéine
  - Atomes lourds
  - Développements de cette technique : utilisation de la sélénométhionine




- Mesure d'un spectre de fluorescence
  - Peak, inflection




$\lambda_{peak}$

$\lambda_{inflection}$

- Collecte des données anomales
  - Mesurer avec précision les paires de Bijvoet
    - Inverse beam (rotation de 180° entre chaque image) : paires mesurées sur images consécutives
    - Orientation du cristal : paires sur la même image

# 4. La diffusion anomale

## 4.5. Détermination de la position des atomes anomaux

### Fonction de Patterson

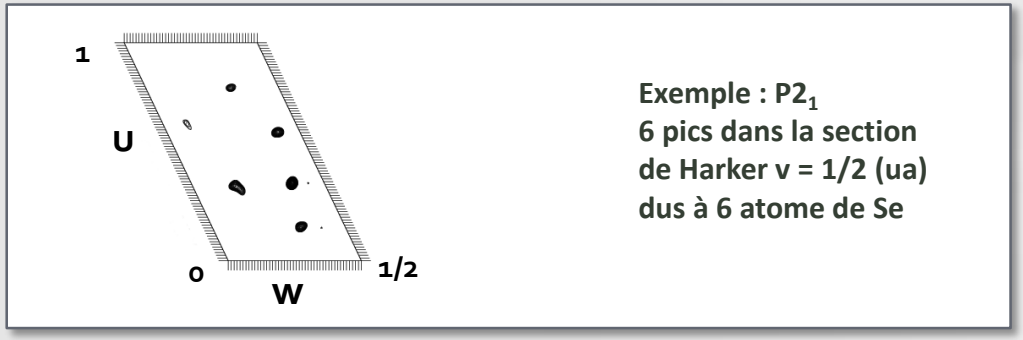
$$P(uvw) = \Sigma (\Delta |F|_{ano})^2 \cos(2\pi(hu + kv + lw))$$

avec $\Delta |F|_{ano} = |F_{hkl}| - |F_{-h-k-l}|$

$\Delta F_{ano} \approx \Sigma \Delta F''$

Calculée sur les données de diffraction issue d'un seul cristal

Fonction de Patterson définissant une carte dont les pics sont les vecteurs interatomiques entre atomes anomaux « uniquement »




Exemple : $P2_1$
6 pics dans la section de Harker v = 1/2 (ua)
dus à 6 atome de Se

**Section de Harker**

### Méthodes directes

- Calcul des phases (probabilité de phases)
- Calcul de la densité électronique

**SHELXD**

# 4. La diffusion anomale

## 4.6. Avantages et inconvenients

### Avantages

- Pas de défaut d'isomorphisme (1 seul cristal nécessaire)
- pas de problème d'occupation (sélénométhionine, ou tout diffuseur anomal lié covalemment, est un point d'ancrage pour la construction de la structure)
- Phasage à haute résolution

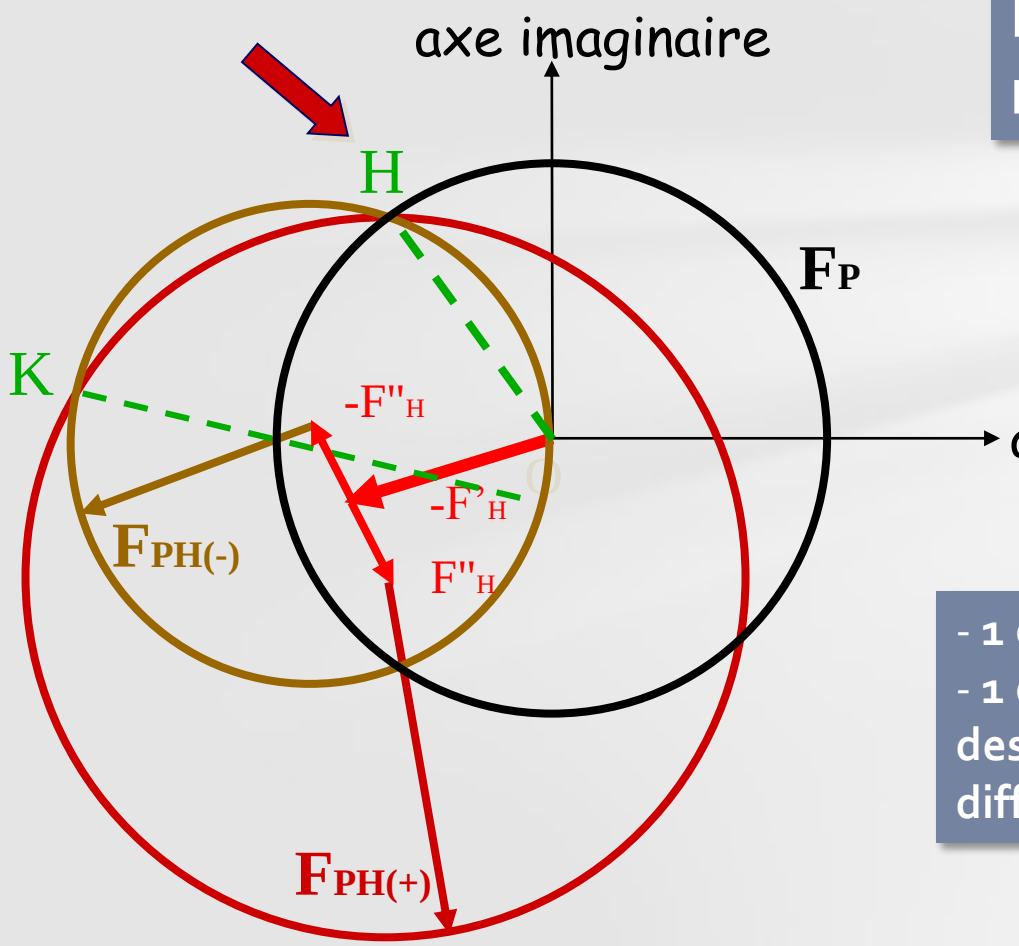
### Inconvénients

- Signal anomal faible par rapport aux intensités collectées
  - il faut 1 Me pour 8 kDa
- L'accès aux sources synchrotron est nécessaire
  - utilisation de Gd sur anode tournante

# 4. La diffusion anomale

## 4.3. Combinaison remplacement isomorphe et diffusion anomale

SIRAS




**Principe de la méthode SIRAS**

- 1 cristal natif
- 1 cristal dérivé contenant des atomes lourds qui sont diffuseurs anomaux

MIRAS

# 5. LE REMPLACEMENT MOLECULAIRE

# 5. Le remplacement moléculaire

## 5.1. Principe de la méthode

**RAPPEL** **Dans le cas des protéines**
Lorsque l'identité de séquence > 25 % les séquences sont dites homologues
cad elles possèdent un ancêtre commun
**DONC** les structures sont similaires

**Hypothèse de base** **La structure d'un homologue est connue**

**Propriétés de la Transformée de Fourier**

**Si le modèle et la structure inconnue sont similaires et sont placés et orientés de la même manière dans une maille présentant les mêmes caractéristiques (paramètres, symétrie)**

**ALORS** $F_{mod}, \varphi_{mod} \approx F_{struct}, \varphi_{struct}$

**manx** **chat**

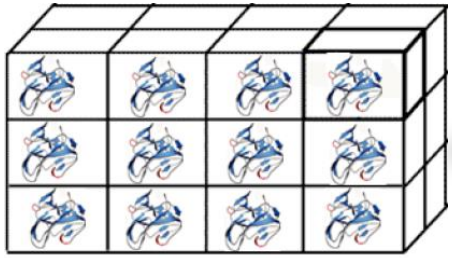
**On peut utiliser les phases $\varphi_{mod}$ à la place de $\varphi_{struct}$**

# 5. Le remplacement moléculaire

## 5.1. Principe de la méthode

**Homologue de structure connue = Modèle de départ**

**Orienter et positionner cette structure modèle dans la maille du cristal étudié**

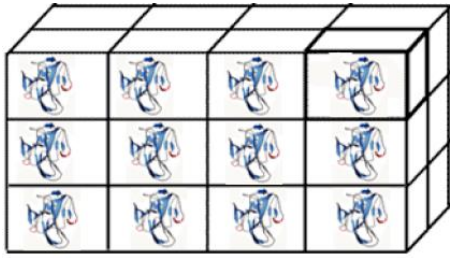


Structure connue (x,y,z) issue de la PDB

1. orienter

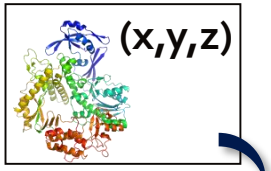
2. positionner

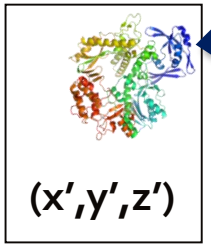
?



$F_{calc}$ $\varphi_{calc}$

6 paramètres pour orienter et positionner le modèle connu



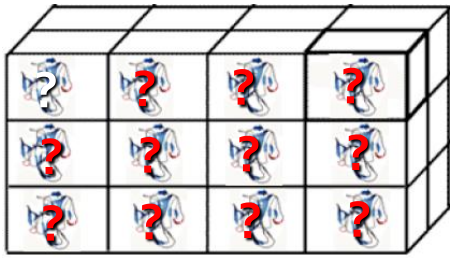




Transformation 6D

$(xyz)_{struct} \approx (xyz)'_{mod} = [C].(xyz)_{mod} + t$

rotation 3D

translation 3D



Structure étudiée

cristal

$F_{obs}$

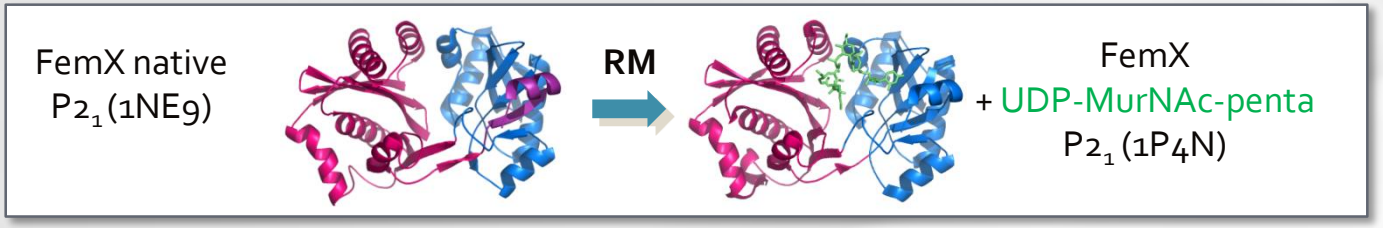
# 5. Le remplacement moléculaire

## 5.2. Pré-requis

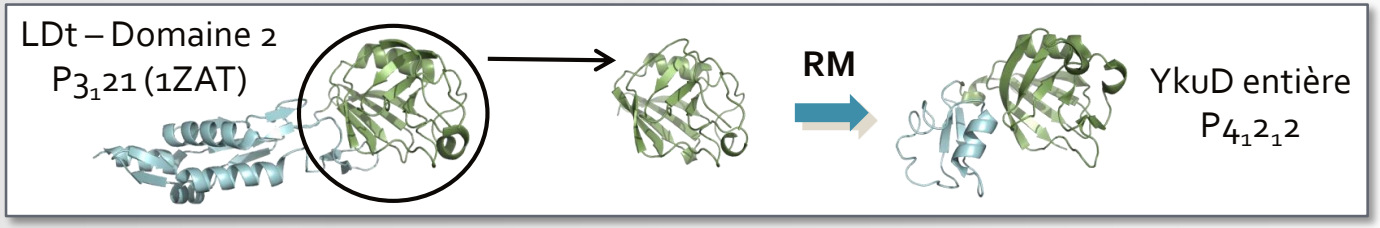
### Concernant le modèle

#### Plusieurs origines possibles

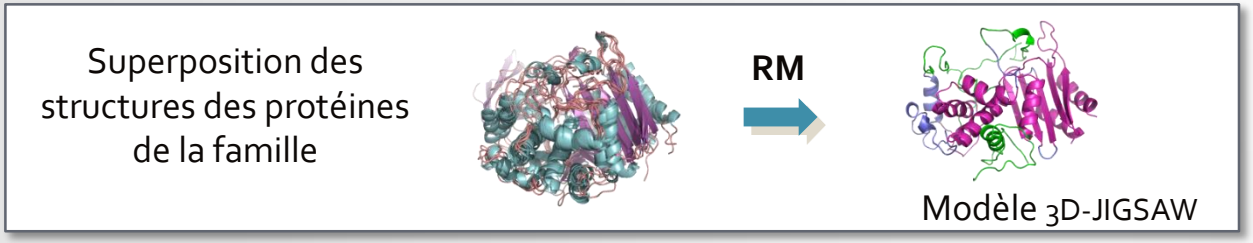
- la même protéine résolue dans un autre groupe d'espace
- la protéine mutée ou en complexe, et vice versa




- une protéine homologue
- un domaine de la protéine




- un domaine homologue à l'un des domaines de la protéine
- un modèle construit à partir de prédiction de structure




- une structure « poly-Ala »

# 5. Le remplacement moléculaire

## 5.2. Pré-requis

### Concernant la composition de la maille

➡ **Coefficient de Matthews $V_m$**

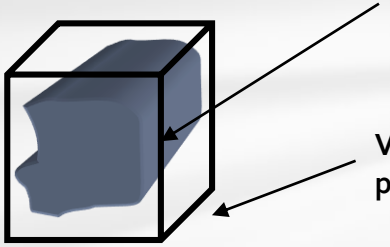
$$V_m = \frac{\text{Volume de la maille}}{Z \times N \times MW}$$

**Volume total occupé par 1 Da (Å³/Da)**

➡ **Pourcentage de solvant**

$$\% \text{ de solvant} = 1 - 1.23 / V_m$$

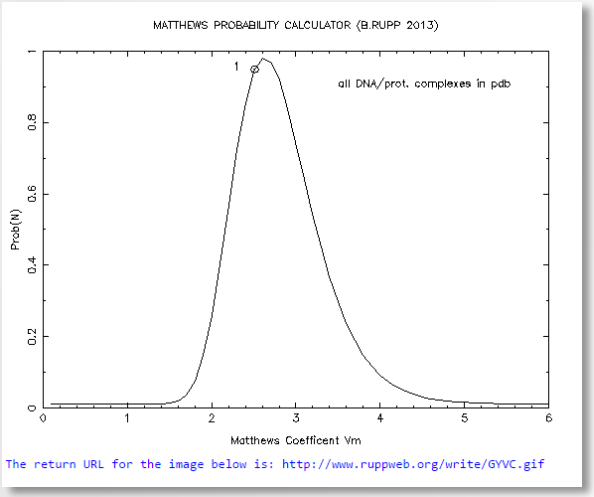
**% de volume occupé par les protéines**




**1,23 Å³/Da = Volume moyen occupé par 1 Da de protéine**

**http://www.ruppweb.org/mattprob/default.html**

```
+-------------------------------------------------------------------------+
| MATTPROB - MATTHEWS PROBABILITY and SOLVENT CONTENT ESTIMATES            |
+-------------------------------------------------------------------------+
| The selected method is parameterized fit to binned 2013 data            |
| (C) 2003-2013 C.X. Weichenberger, K. Kantardjieff and B. Rupp           |
+-------------------------------------------------------------------------+
| Unit cell dimensions in A, deg(o) and cubic A :                         |
+-------------------------------------------------------------------------+
|      a         b         c       alpha      beta       gamma      volume |
|    84.30     98.78     43.90      90.00     90.00      90.00    365562.1 |
+-------------------------------------------------------------------------+
| Space group P 21 21 21 , space group number  19, Laue class mmm  , z=  4 |
+-------------------------------------------------------------------------+
| System as decoded from unit cell dimensions : orthorhombic              |
| is consistent with space group  P 21 21 21  : orthorhombic              |
| Unit cell volume is       :     365562.1 A**3                           |
| Asymmetric unit volume is :      91390.5 A**3                           |
+-------------------------------------------------------------------------+
+-------------------------------------------------------------------------+
| Found  1 possible solutions :                                           |
+-------------------------------------------------------------------------+
| Vm of protein (Matthews coeff.) is  2.51 A**3/Dalton                    |
| Estimated solvent content is       54.94 %                              |
+-------------------------------------------------------------------------+
+-------------------------------------------------------------------------+
|  N(mol)      Prob(N)     Prob(N)        Vm          Vs         Mw       |
|          for resolution  overall      A**3/Da   % solvent     Da        |
+-------------------------------------------------------------------------+
|   1          1.0000      1.0000        2.51       54.94    36440.00     |
+-------------------------------------------------------------------------+
The return URL for the image below is: http://www.ruppweb.org/write/GYV.gif
```



# 5. Le remplacement moléculaire

## 5.3. Procédure

⟹ **Approche utilisant la fonction de Patterson**

1. Calculer les fonctions de Patterson
   - pour le modèle (maille P1 avec des paramètres de maille très grands — $P_1$
   - des données de diffraction ($F_{obs}$) — $P_2$
2. Orienter (rotation) le modèle en utilisant les vecteurs interatomiques intramoléculaires pour les fonctions de Patterson
   - maximum de corrélation de la fonction de rotation
3. Positionner la molécule orientée en utilisant les vecteurs interatomiques intermoléculaires pour les fonctions de Patterson
   - maximum de corrélation pour la fonction de translation
4. La position avec le meilleur score est la solution

MOLREP AMORE

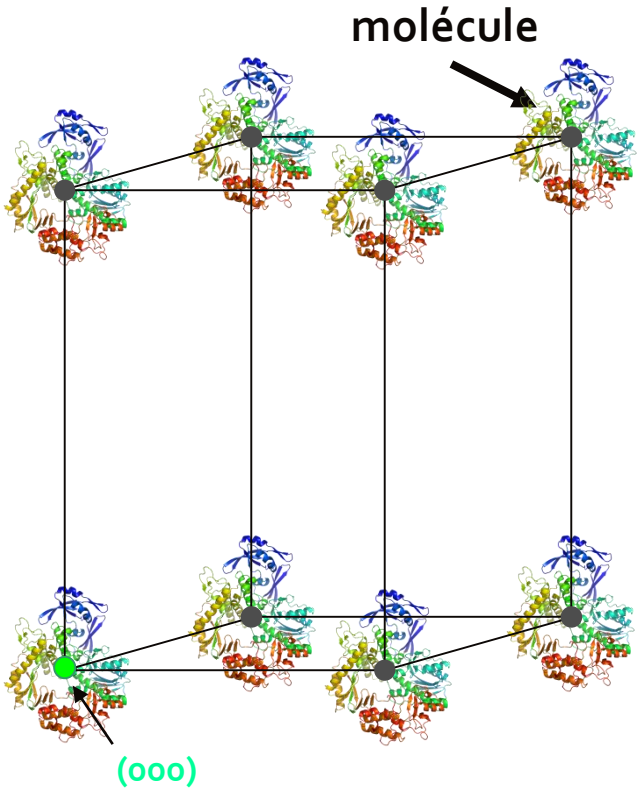
PHASER

Fonction de score basée sur la méthode de maximum de vraisemblance

# 5. Le remplacement moléculaire

## 5.3. Procédure

### Fonction de rotation




Maille P1 = pas de symétrie

Maille P1 très grande :

Distinction vecteurs intermoléculaires et intramoléculaires dans le calcul de la fonction de Patterson

$$R(\alpha,\beta,\gamma) = \Sigma_{intra}\ P_2(u)\ P_{1r}(u_r)$$

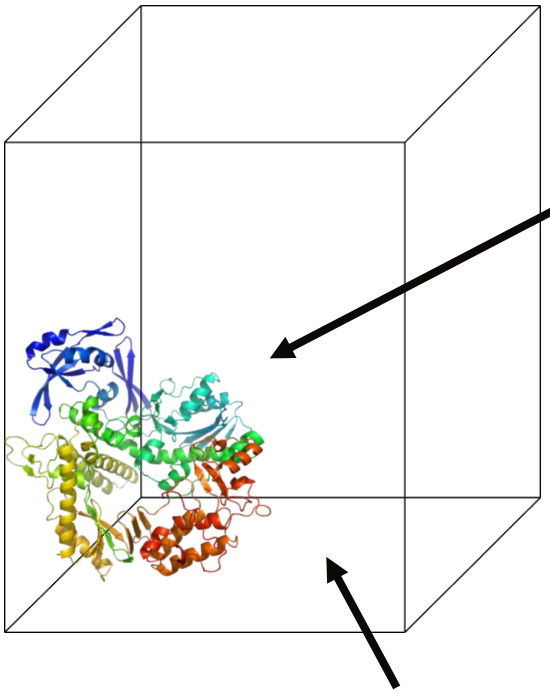
$R(\alpha,\beta,\gamma)$ est maximale lorsque les 2 fonctions de Patterson se superposent

# 5. Le remplacement moléculaire

## 5.3. Procédure

### Fonction de translation

(xyz)struct ≈ (xyz)'mod = [C].(xyz)mod + t



Molécule orientée placée à l'origine selon (α,β,γ) trouvées par la fonction de rotation

$$T(u,v,w) = \Sigma_{inter}\, P_2(u)\, P_{1rt}(u_t)$$

Les vecteurs intramoléculaires de P2 et P1r sont identiques et sont éliminés par soustraction

T(u,v,w) est maximale lorsque les 2 fonctions de Patterson se superposent

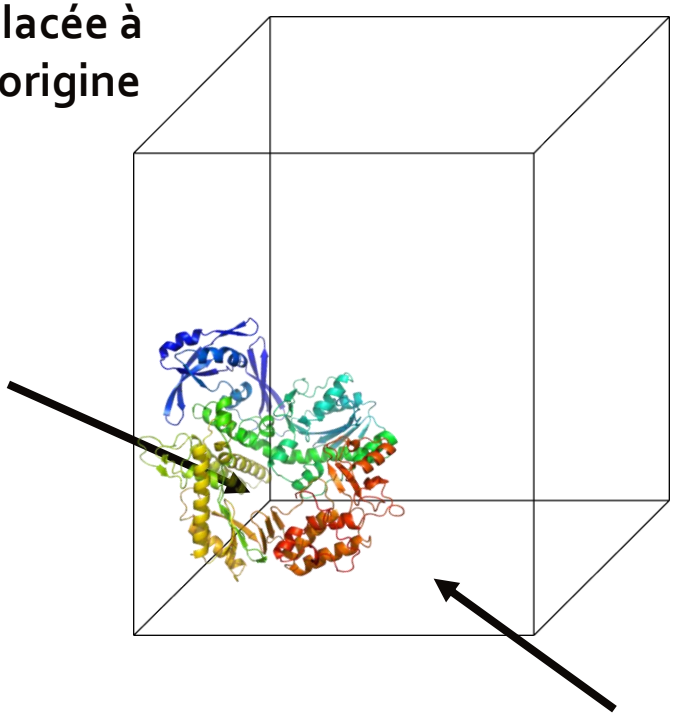
Maille cristalline de la structure inconnue

# 5. Le remplacement moléculaire

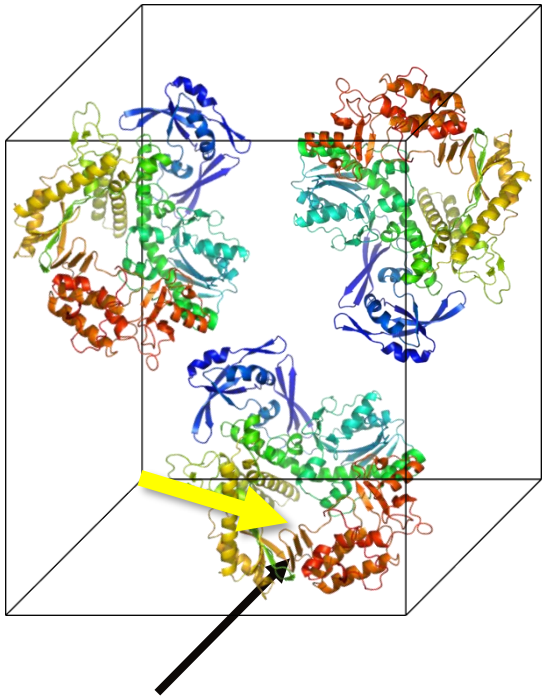
## 5.3. Procédure

### Fonction de rotation

**Molécule orientée placée à l'origine**



**Molécule orientée et positionnée**



Affinement par corps rigides de l'orientation et de la position
Molécule orientée et positionnée $(x,y,z)_{mod}$

$TF^{-1}$

$F_{calc}$, $\phi_{calc}$

**Validation du modèle**

Empilement cristallin correct
pas de collision entre les symétriques

**Maille cristalline de la structure inconnue**

# 5. Le remplacement moléculaire

## 5.6. Avantages et inconvénients

### Avantages

- ne nécessitant qu'un seul espace de diffraction
- ne nécessitant pas l'étape de construction du modèle
- rapide en cas de succès

### Inconvénients

- le succès de la méthode dépend
  - du taux de similarité structurale entre la protéine connue et la protéine étudiée
  - de la « forme » de la protéine (globulaire ou allongée)
  - très sensible aux basses et hautes résolution

# Conclusion

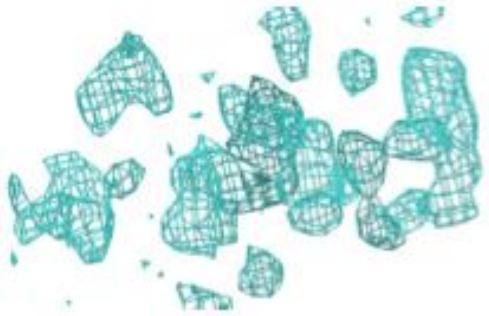
$|^{1/2}$ → F(h,k,l) ϕ(h,k,l)

Données de diffraction

Estimées par l'une des méthodes

ρ(x,y,z)

Calculée sur l'unité asymétrique



m $F_P$ , $\phi_{best}$

**Interprétable ?? Améliorable ??**