

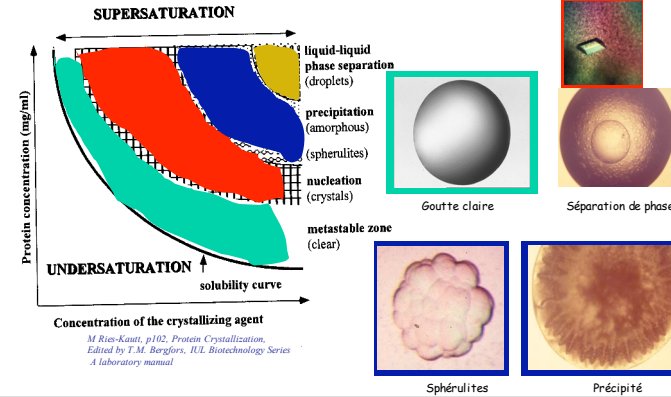
D'où vient l'information structurale et que vaut-elle ?

## La cristallisation

### L'information cristallographique

- Notions de base sur la diffraction
- Qu'est-ce que la résolution d'une structure ?
- La précision des coordonnées atomiques.
- Fiabilité d'une structure.
- L'extension de l'information structurale à des protéines homologues de structure inconnue.

Objectif : cristal tridimensionnel, 100  $\mu\text{m}$  dans la plus petite dimension (à partir de 15  $\mu\text{m}$ ), monocristallin.

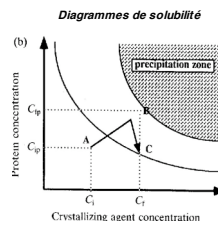
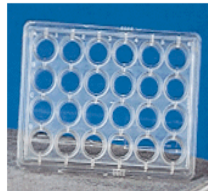


### Méthodes de cristallisation

Diffusion de vapeur

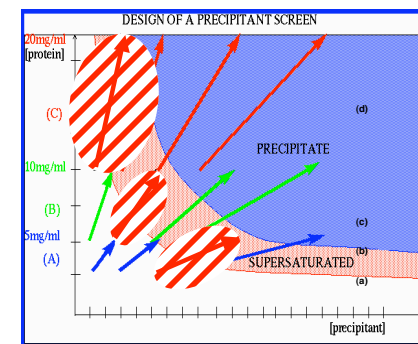
Hanging Drop with Protein

Reservoir with Precipitant

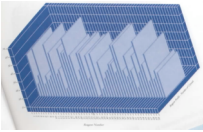



Goutte: typiquement 2  $\mu\text{l}$  protéine + 2  $\mu\text{l}$  de tampon de cristallisation (de 1 à 50  $\mu\text{l}$ ) Puits: 0,5 à 1 ml.  
 Temps d'équilibration: de 24 h (sels) à 10 jours (PEG)  
 Observation des gouttes et manipulation des cristaux aisées

### Cribles matriciels et diagramme de solubilité



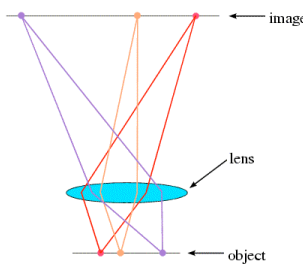
### Cribles statistiques

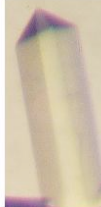
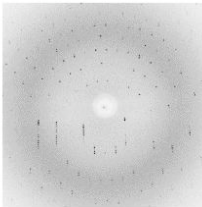
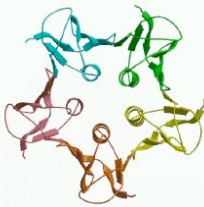



Nombre de cristallisations rapportées pour les différentes conditions du crible de Jancarik & Kim (d'après HamptonResearch)

	1	2	3	4	5	6
A	1 MPD 30% 0.02M CaCl2 NaAcO (4.6) pH 5.3	2 Na Tartrate 0.4M pH 6.9	3 NH4PO4 0.4M pH 4.4	4 AmS 2.0M Tris HCl pH 8.5	5 MPD 30% 0.2M NaCitrate Na Hepes 7.5	6 PEG 4K 30% 0.2M MgCl2 Tris HCl 8.5 pH 8.8
B	7 NaAcO 1.4M Na Caco 6.5	8 2-propanol 30% 0.2M NaCitrate Na Caco 6.5 pH 7.2	9 PEG 4K 30% 0.2 M NH4AcO Na Citrate (5.6) pH 6.7	10 PEG 4K 30% 0.2 M NH4AcO NaAcO (4.6) pH 6.1	11 NH4PO4 1.0M Na Citrate (5.6) pH 5.0	12 2-propanol 30% 0.2M MgCl2 Na Hepes 7.5
C	13 PEG 400 30% 0.2M NaCitrate Tris HCl (8.5) pH 9.3	14 PEG 400 28% 0.2M CaCl2 Na Hepes (7.5) pH 6.9	15 PEG 4K 30% 0.2 M AmS Na Caco (6.5) pH 6.9	16 LiSO4 1.5M NaHepes (7.5) pH 7.9	17 PEG 4K 30% 0.2M LiSO4 Tris HCl (8.5) pH 9.2	18 PEG 8K 20% 0.2M MgAcO2 Na Caco (6.5) pH 6.8
D	19 2-propanol 30% 0.2M NH4AcO Tris HCl (8.5) pH 8.9	20 PEG4K 25% 0.2M AmS NaAcO (4.6) pH 5.1	21 MPD 30% 0.2M MgAcO2 Na Caco (6.5) pH 9.2	22 PEG 4K 30% 0.2M NaAcO Tris HCl (8.5) pH 7.4	23 PEG 400 30% 0.2M MgCl2 Na Hepes (7.5) pH 7.4	24 2-propanol 20% 0.2M CaCl2 NaAcO (4.6) pH 4.9

Les 24 premières conditions



## La cristallisation

### L'information cristallographique

- Notions de base sur la diffraction
- Qu'est-ce que la résolution d'une structure ?
- La précision des coordonnées atomiques.
- Fiabilité d'une structure.
- L'extension de l'information structurale à des protéines homologues de structure inconnue.

## Diffraction des Rayons X

- Diffusion par un objet
- Diffraction par un objet périodique
- Résolution

Diffusion des rayons X - 1

$R = R_1 + R_2$

$E = E_0 \exp 2i\pi\nu(t-R/c) = E_0 \exp 2i\pi(\nu t - R/\lambda) \quad \lambda = c/\nu$

$E = E_0 \exp 2i\pi(\nu t - R/\lambda) + E_0 \exp 2i\pi(\nu t - R'/\lambda)$

Diffusion des rayons X - 2

$\Delta R = r \cdot (s_0 - s)$

$\Delta \phi = 2\pi r(s - s_0)/\lambda$

$\vec{h} = (s - s_0)/\lambda$

$\Delta \Phi = 2\pi \vec{r} \cdot \vec{h}$

$\rho(r) \exp(2i\pi \vec{r} \cdot \vec{h}) dV$

$F(\vec{h}) = \iiint_V \rho(r) \exp(2i\pi \vec{h} \cdot \vec{r}) dV$

$\rho(\vec{r}) = \iiint F(\vec{h}) \exp(-2i\pi \vec{h} \cdot \vec{r}) d\vec{h}$

Diffraction à une dimension

N atomes ponctuels

$F(h) = \sum c \exp(2i\pi n_i h r_0)$

$r_0 \cdot b = 1 \quad h = k/b$

$F(h) = \sum c \exp(2i\pi n_i k) = c \cdot \frac{1 - (\exp 2i\pi k)^{N+1}}{1 - \exp 2i\pi k}$

**k ≠ entier    F(h) ≈ c**  
**k entier    F(h) = cN**

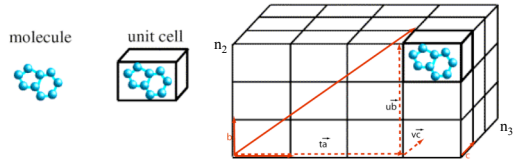
Diffraction à une dimension (fin)

Molécule

Cristal moléculaire

### Diffraction à 3 dimensions

1 dimension :  $F(h) = \sum c \exp(2i\pi n_i k) = c \frac{1 - (\exp 2i\pi k)^{N+1}}{1 - \exp 2i\pi k}$   
 3 dimension



$$F(\vec{h}) = \sum f(\vec{h}) \exp(2i\pi(\tau a + u b + v c) \cdot \vec{h})$$

$$F(\vec{h}) = f(\vec{h}) \sum_{\tau=0}^{n_1} \exp[2i\pi \tau \vec{a} \cdot \vec{h}] \times \sum_{u=0}^{n_2} \exp[2i\pi u \vec{b} \cdot \vec{h}] \times \sum_{v=0}^{n_3} \exp[2i\pi v \vec{c} \cdot \vec{h}]$$

$$\vec{a} \cdot \vec{h} = h$$

$$\vec{b} \cdot \vec{h} = k$$

$$\vec{c} \cdot \vec{h} = l$$

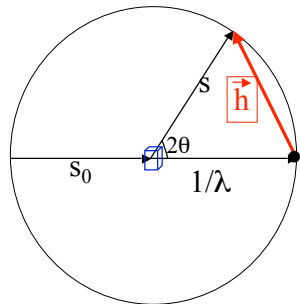
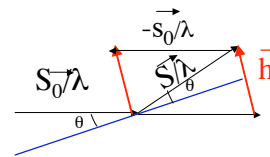
$$\vec{a} \cdot \vec{h} = h$$

$$\vec{b} \cdot \vec{h} = k \Leftrightarrow \vec{h} = h \vec{a}^* + k \vec{b}^* + l \vec{c}^*$$

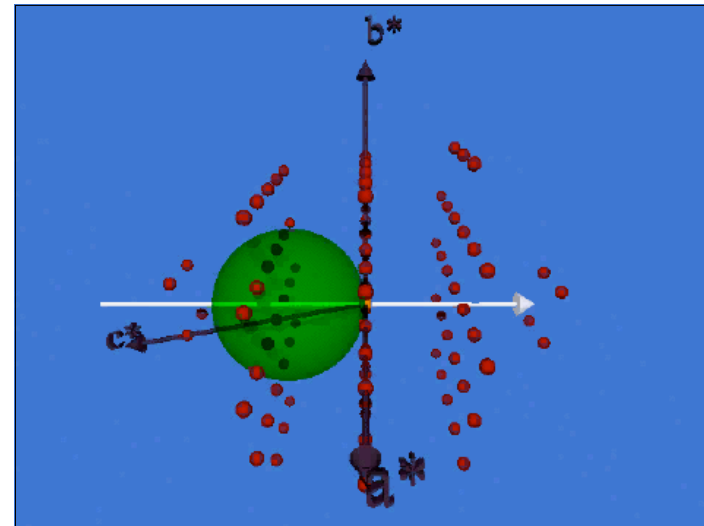
$$\vec{c} \cdot \vec{h} = l \quad \vec{a} \cdot \vec{a}^* = 1 \quad \vec{b} \cdot \vec{a}^* = 0 \quad \vec{c} \cdot \vec{a}^* = 0$$

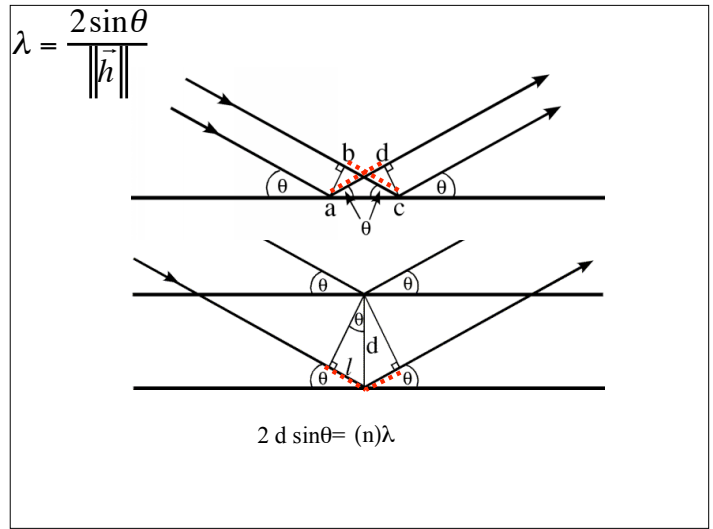
$$\vec{a} \cdot \vec{b}^* = 0 \quad \vec{b} \cdot \vec{b}^* = 1 \dots \dots$$

Diffraction  $\Leftrightarrow$  h, k et l entiers



$$\|\vec{h}\| = \frac{2 \sin \theta}{\lambda}$$

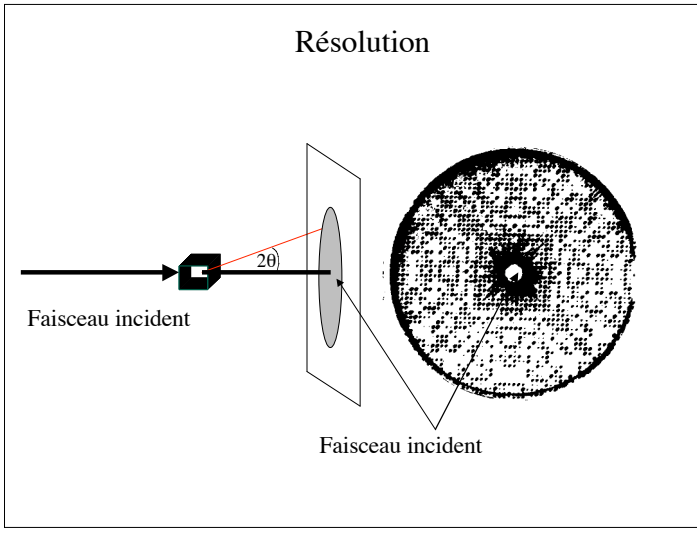
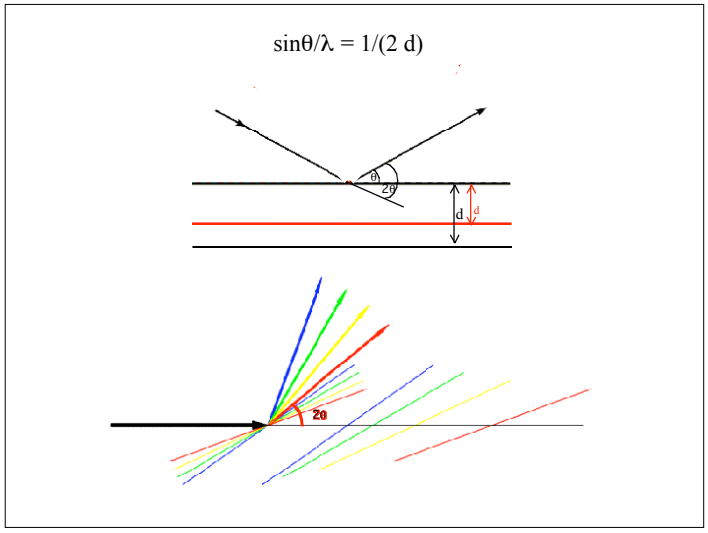




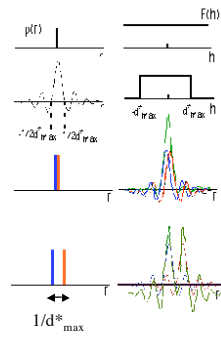
## La cristallisation

### L'information cristallographique

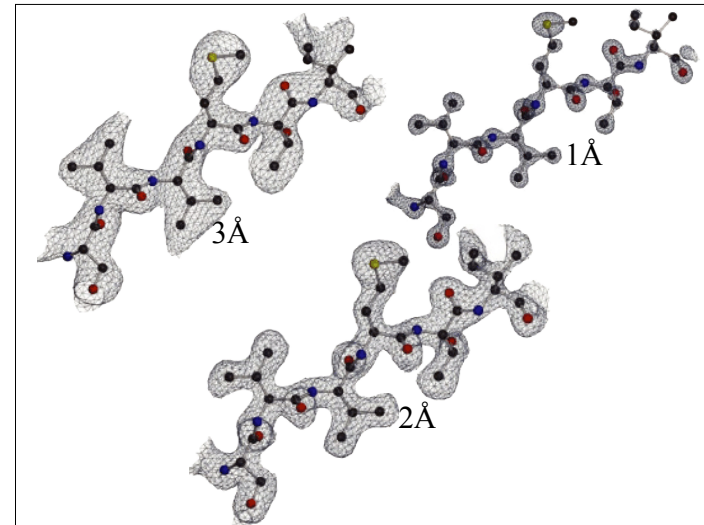
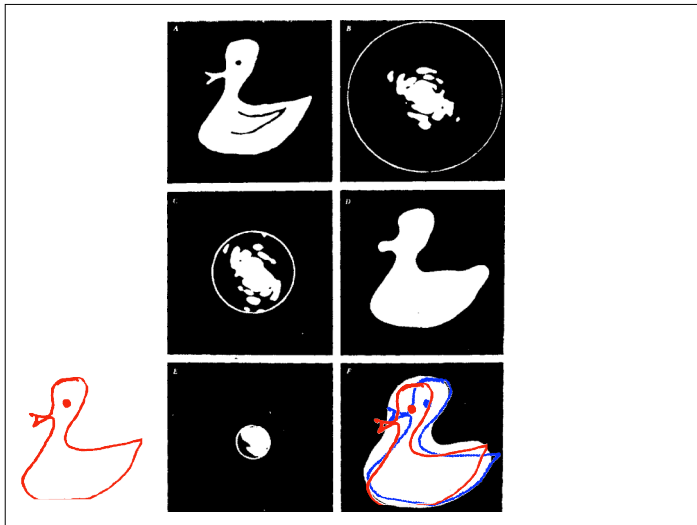
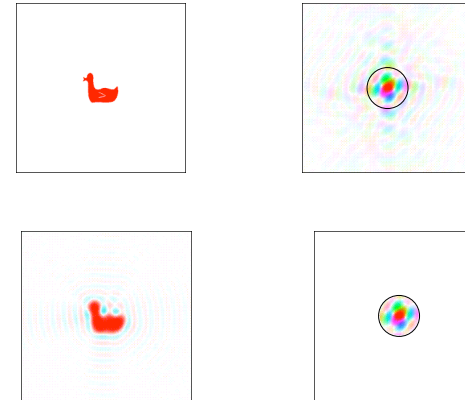
- Notions de base sur la diffraction
- Qu'est-ce que la résolution d'une structure ?
- La précision des coordonnées atomiques.
- Fiabilité d'une structure.
- L'extension de l'information structurale à des protéines homologues de structure inconnue.



### Résolution (1 dimension)



### Résolution (2 dimensions)



## La cristallisation

### L'information cristallographique

- Notions de base sur la diffraction
- Qu'est-ce que la résolution d'une structure ?
- La précision des coordonnées atomiques.
- Fiabilité d'une structure.
- L'extension de l'information structurale à des protéines homologues de structure inconnue.

### Pour chaque atome :

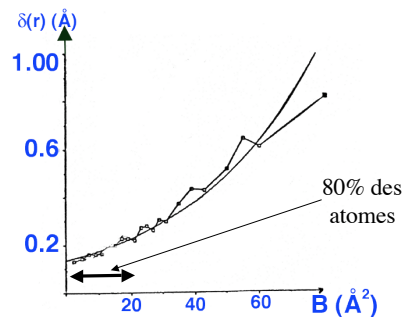
- 3 coordonnées : x, y, z
- 1 facteur d'agitation thermique

$$B = 8 \pi^2 \langle u^2 \rangle$$

Amplitude des vibrations de l'atome

### Précision des structures de protéines

Résolution	3 Å	1.7 Å
Précision moyenne	0.6 Å	0.2 Å



## La cristallisation

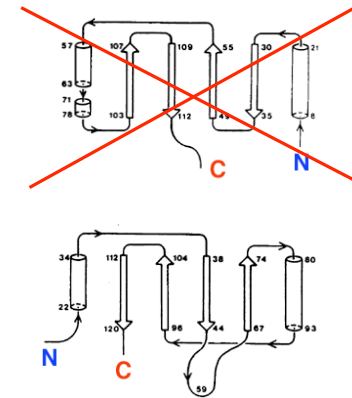
### L'information cristallographique

- Notions de base sur la diffraction
- Qu'est-ce que la résolution d'une structure ?
- La précision des coordonnées atomiques.
- Fiabilité d'une structure.
- L'extension de l'information structurale à des protéines homologues de structure inconnue.

Le problème cristallographique est sous déterminé

Resolution	Observations/paramètres (x,y,z, B)
3.5Å	0.5
3.0Å	0.8
2.5Å	1.4
2.0Å	2.8
1.5Å	6.2

La petite sous-unité de la RuBisCO 8-(15 kD +55 kD)  
Résolution : 2.5 Å



Fiabilité d'une structure  
Facteur d'accord cristallographique

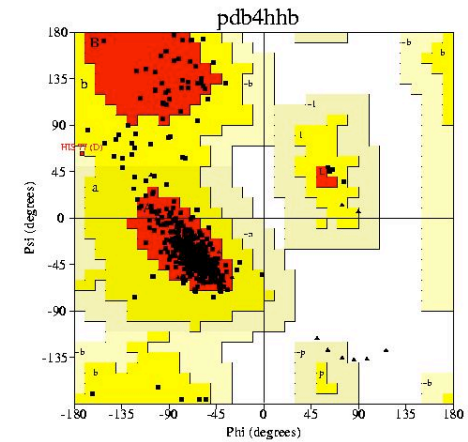
$$\rho(\vec{r}) = \sum_{\text{atomes}} \rho(\vec{r}_i)$$

$$F_{\text{calc}}(\vec{h}) = \mathcal{F}\left(\sum_{\text{atomes}} \rho(\vec{r}_i)\right) = \sum_{\text{atomes}} \mathcal{F}(\rho(\vec{r}_i))$$

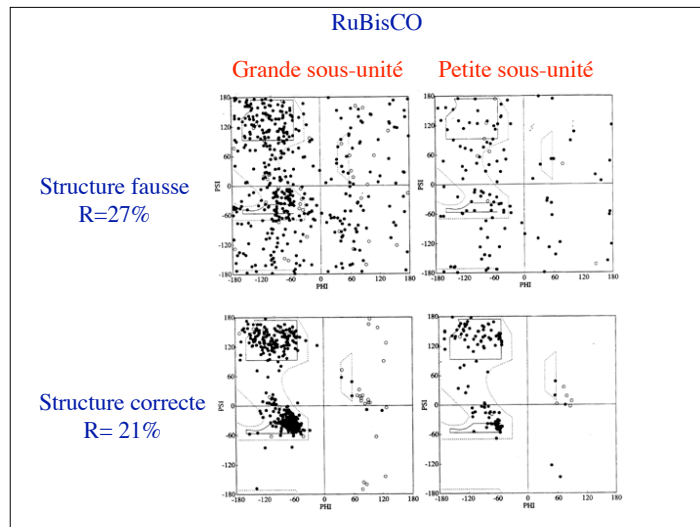
$$F_{\text{calc}}(\vec{h}) = \sum_{\text{atomes}} f(\vec{h}) \exp 2i\pi\vec{h}\vec{r}_i$$

$$R = \frac{\sum_{\vec{h}} \left| |F_{\vec{h}}^{\text{obs}}| - |F_{\vec{h}}^{\text{calc}}| \right|}{\sum_{\vec{h}} |F_{\vec{h}}^{\text{obs}}|}$$

$$R \approx 0.18 \leftrightarrow 0.22$$





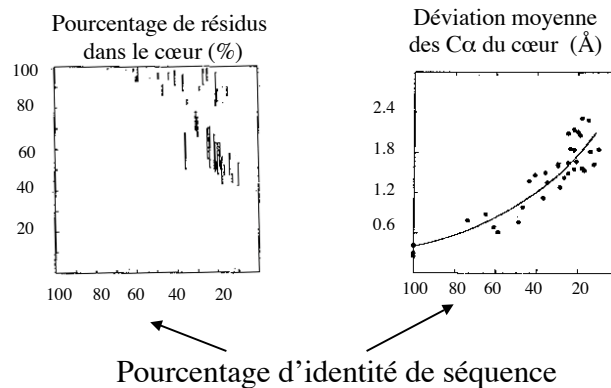


## La cristallisation

### L'information cristallographique

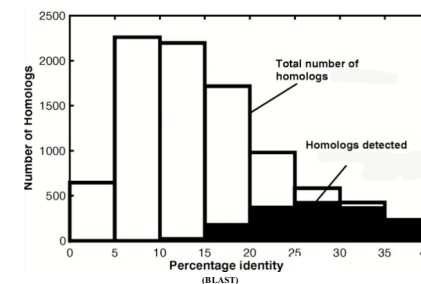
- Notions de base sur la diffraction
- Qu'est-ce que la résolution d'une structure ?
- La précision des coordonnées atomiques.
- Fiabilité d'une structure.
- L'extension de l'information structurale à des protéines homologues de structure inconnue.

### Similitude de séquences et ressemblance structurale



Chothia et Lesk (1986) EMBO J., pp 823-6

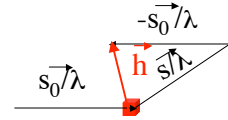
### Reconnaissance d'une homologie



S. E. Brenner et al. PNAS Vol 95 pp 6073-6078 (1998)

## Comment on détermine la structure d'une protéine

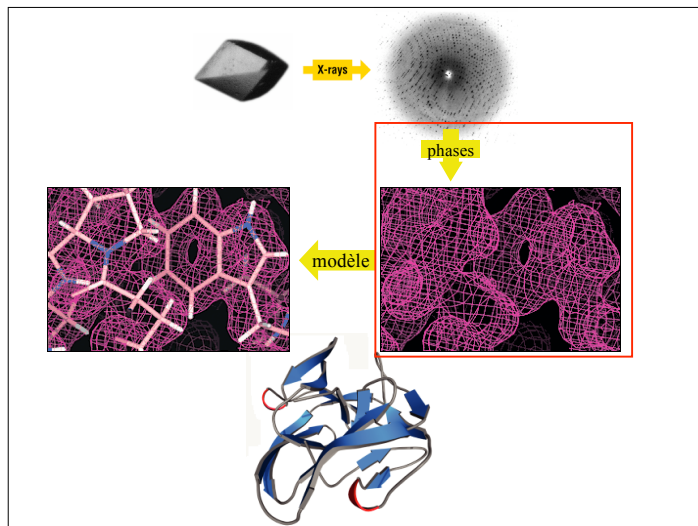
- Remplacement Moléculaire
- Remplacement Isomorphe
- Construction du modèle
- Affinement



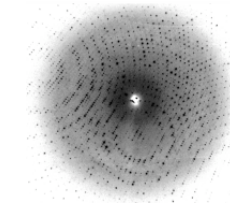
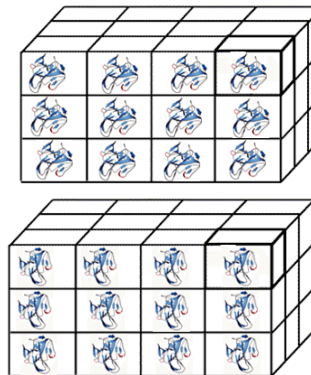
$$\vec{h} = (\vec{s} - \vec{s}_0) / \lambda$$

$$F(\vec{h}) = \iiint_V \rho(r) \exp(2i\pi\vec{h}\vec{r}) dV$$

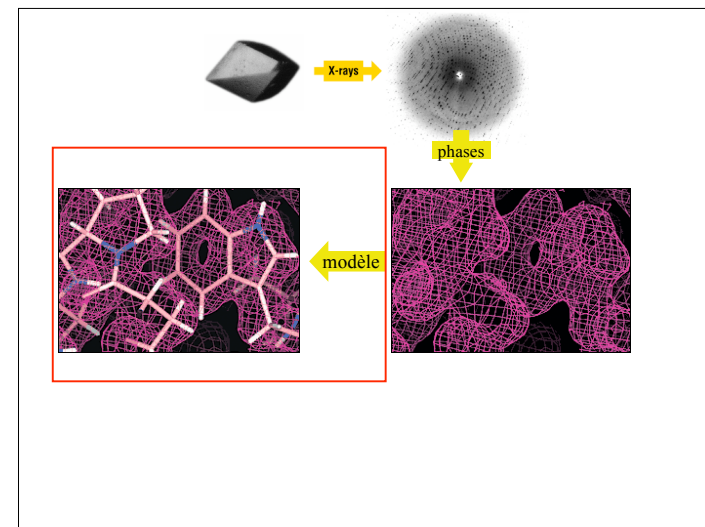
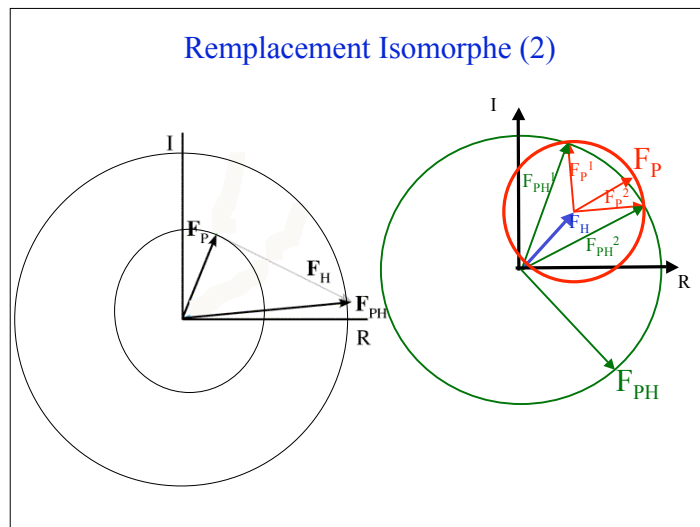
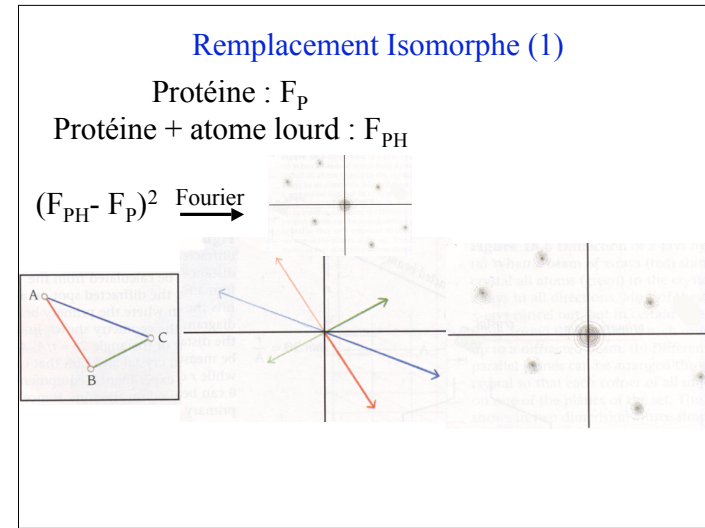
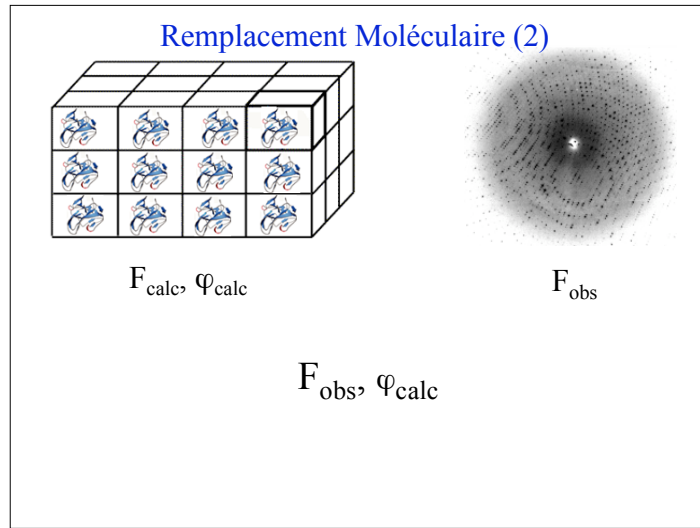
$$\rho(\vec{r}) = \iiint F(\vec{h}) \exp(-2i\pi\vec{h}\vec{r}) d\tau$$



## Remplacement Moléculaire (1)



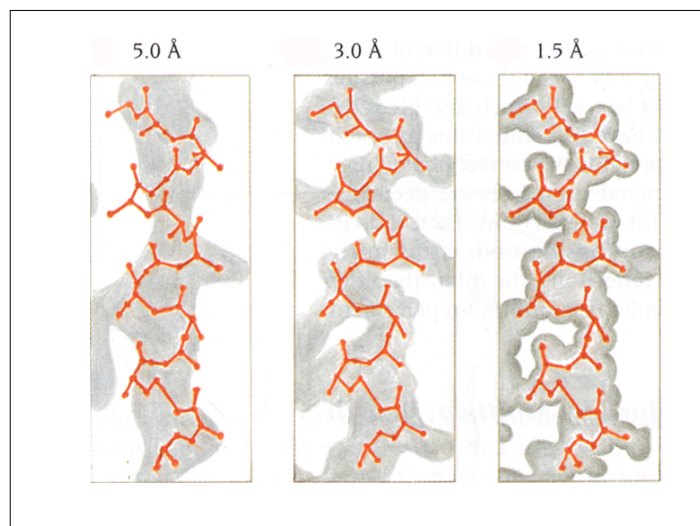
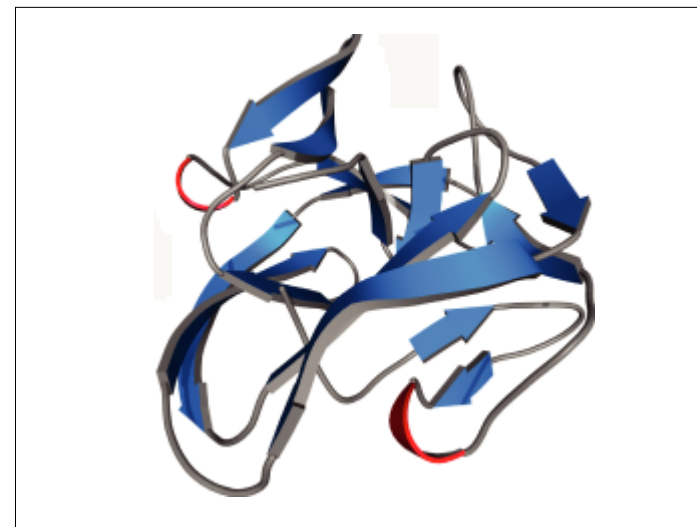
$$\alpha, \beta, \gamma, t_x, t_y, t_z$$



## Affinement

$$F_h^{calc}(\vec{r}_i) = \mathcal{F}\left(\sum_{atoms} \rho(\vec{r}_i)\right) = \sum_{atoms} \mathcal{F}(\rho(\vec{r}_i))$$

$$R(\vec{r}_i) = \frac{\sum_h ||F_h^{obs}| - |F_h^{calc}(\vec{r}_i)||}{\sum_h |F_h^{obs}|}$$



## Ce qu'il faut "absolument" retenir (1)

- L'observation de la densité électronique nécessite de séparer diffusion des rayons X et reconstruction de l'image
- La diffusion des RX par des molécules n'est observable que via l'interférence constructive de la diffusion de nombreuses molécules
- La diffraction s'interprète comme la réflexion par une série de plans dont la distance est d'autant plus petite que l'angle entre rayons incident et diffracté est grand

## Ce qu'il faut "absolument" retenir (2)

- La résolution détermine la précision des coordonnées atomiques et la fiabilité du modèle
- Deux critères objectifs permettent de préjuger de la justesse d'un modèle atomique : facteur R et diagramme de Ramachandran
- La structure d'une protéine est semblable à celles des protéines de séquence semblable
- On détecte de façon fiable (faux positifs < 1%) par alignement de séquences des protéines de même structure et qui ont plus que 20% - 25% d'identité de séquence