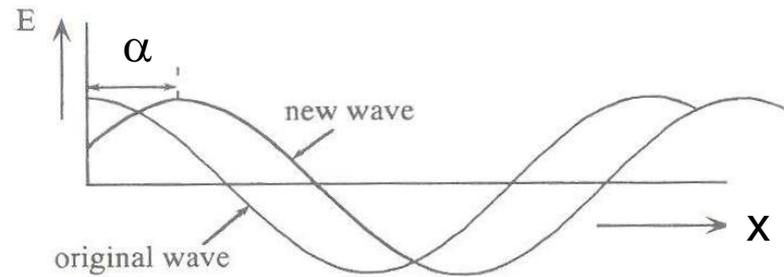


Ondas 1 e 2 com tem o mesmo λ e A mas diferentes α .

Onda 1, de referência, com fase $\alpha = 0$

$$\underline{A} = A \cos(\omega t) \text{ na posição } x = 0$$

$$\underline{A} = A \cos(2\pi x/\lambda) \text{ no tempo } = 0$$



Onda 2, deslocada da onda 1 com fase α

$$\begin{aligned} \underline{A} = A \cos(\omega t + \alpha) &= A \cos(\alpha) \cos(\omega t) - A \sin(\alpha) \sin(\omega t) \\ &= A \cos(\alpha) \cos(\omega t) + A \sin(\alpha) \cos(\omega t + 90^\circ) \end{aligned}$$

parte real
com $\alpha=0$

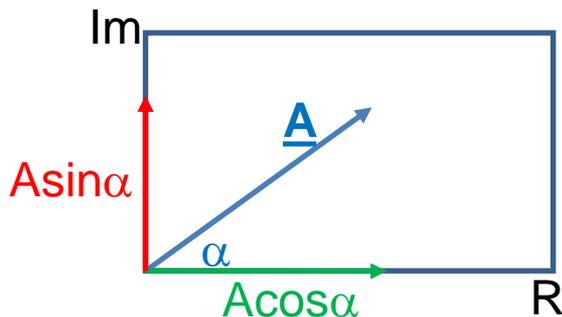
parte imaginária
com $\alpha = 90^\circ$

amplitudes: $A_{\text{real}} = A \cos \alpha$

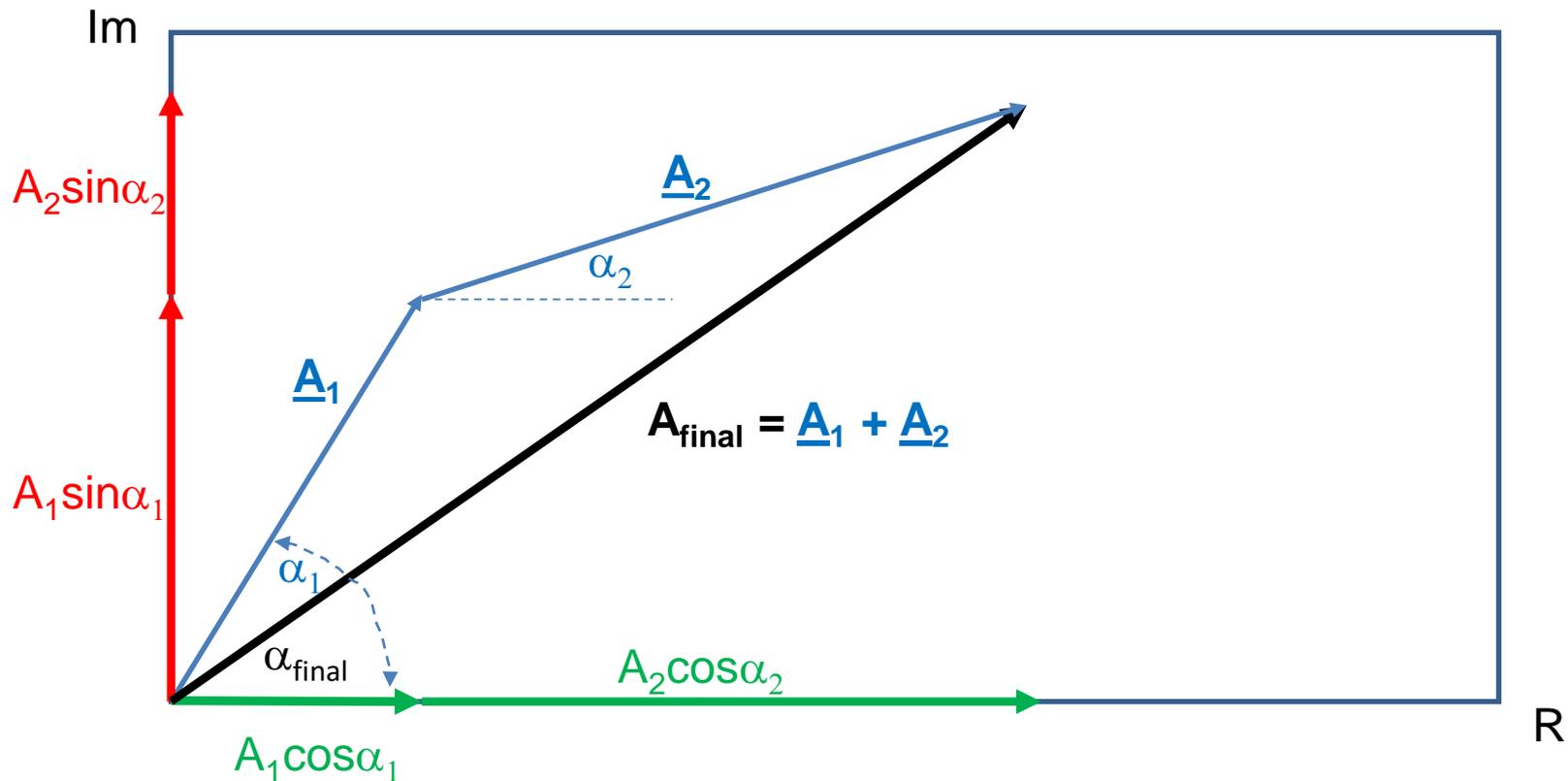
$A_{\text{imag}} = A \sin \alpha$

Logo, qualquer onda com qualquer fase (α) pode ser representada como a soma de duas ondas:

Uma com fase zero e amplitude $A \cos \alpha$
e outra com fase 90° ($\pi/2$ rads) e amplitude $A \sin \alpha$



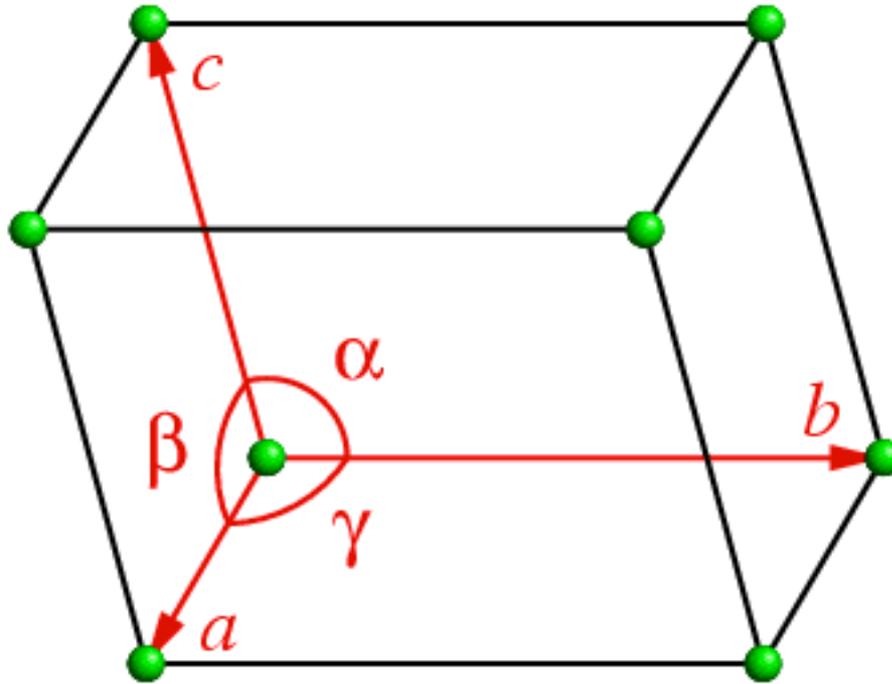
Estes dois componentes podem ser representados por vetores num *plano complexo* (Diagrama de Argand)



Esta representação permite que podemos somar múltiplas ondas simplesmente somando suas partes reais juntas e suas partes imaginárias juntas

$$\underline{A}_{\text{final}} = \underline{A}_1 + \underline{A}_2 = A_1 \cos \alpha_1 + A_2 \cos \alpha_2 + iA_1 \sin \alpha_1 + iA_2 \sin \alpha_2$$

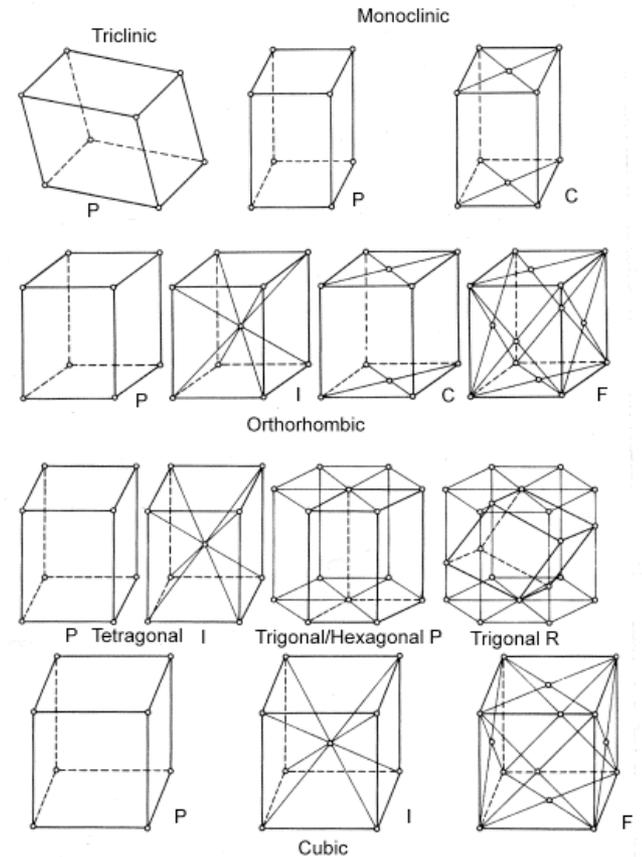
A célula unitária



Podemos expressar a posição de um ponto (x, y, z) dentro da célula empregando coordenadas em Å (Ångstroms) ou em coordenadas fracionais dos vetores a , b e c . Por exemplo, para uma célula com dimensões 40 Å x 50 Å x 100 Å, o ponto (30 Å, 25 Å, 20 Å) é o mesmo do ponto (0.75, 0.5, 0.2) em coordenadas fracionais.

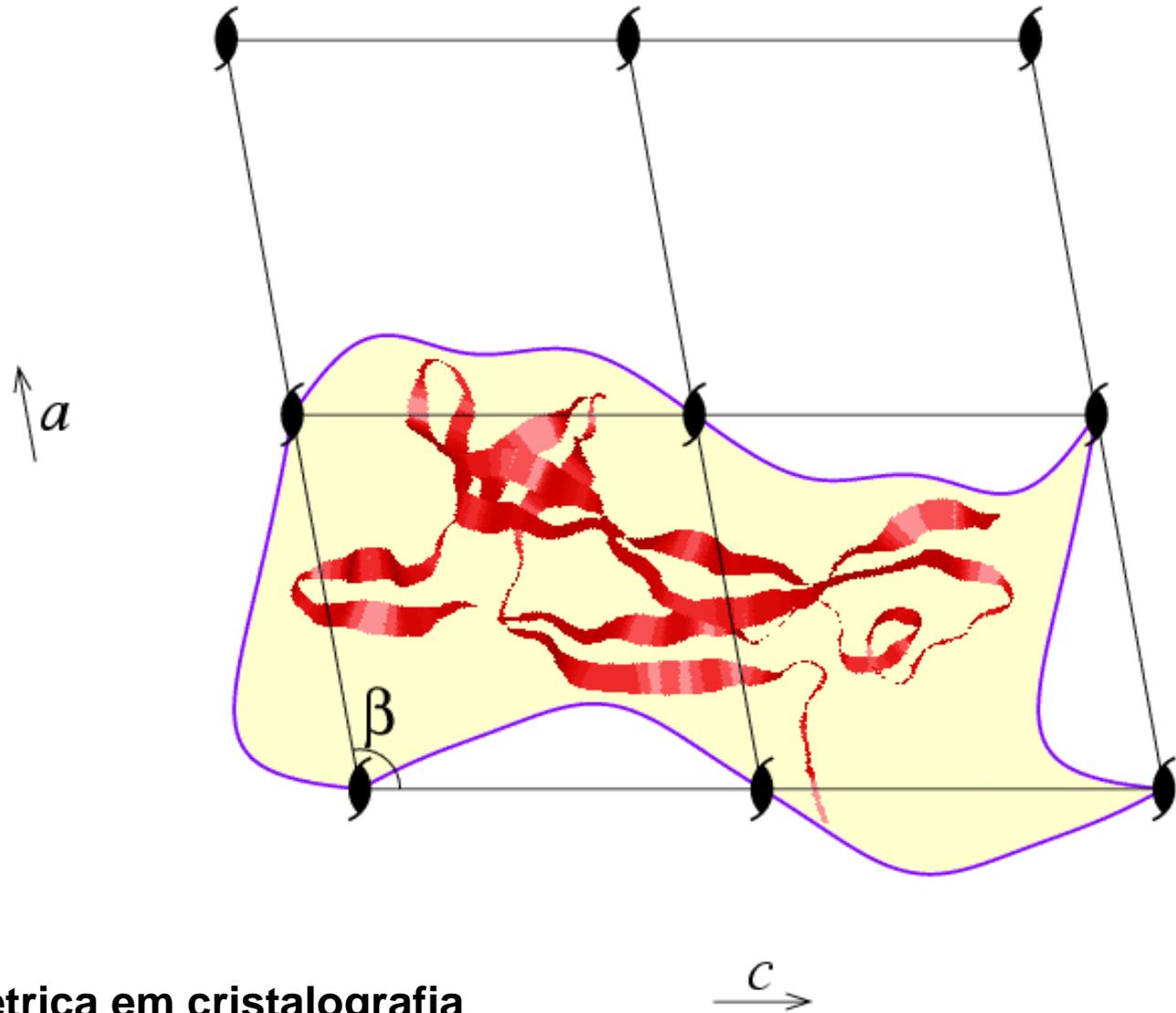
Bravais lattice = crystal system + lattice centering

Crystal family	Lattice system	Centering and 14 Bravais Lattices			
		Primitive	Base-centered	Body-centered	Face-centered
triclinic					
		$\beta \neq 90^\circ$ $a \neq c$	$\beta \neq 90^\circ$ $a \neq c$		
monoclinic					
		$a \neq b \neq c$	$a \neq b \neq c$	$a \neq b \neq c$	$a \neq b \neq c$
orthorhombic					
		$a \neq c$		$a \neq c$	
tetragonal					
		$\alpha \neq 90^\circ$			
hexagonal	rhombohedral				
	hexagonal	$\gamma = 120^\circ$			
cubic					



Crystal System	Laue Class	Crystal Class	Lattice Centring	Enantiomorphic Space Groups	Number of Asymmetric Units per Unit Cell
Triclinic	-1	1	<i>P</i>	<i>P</i> 1	1
Monoclinic	<i>2/m</i>	2	<i>P</i>	<i>P</i> 2, <i>P</i> 2 ₁	2
			<i>C</i>	<i>C</i> 2	4
Orthorhombic	<i>mmm</i>	222	<i>P</i>	<i>P</i> 222, <i>P</i> 222 ₁ , <i>P</i> 2 ₁ 2 ₁ 2, <i>P</i> 2 ₁ 2 ₁ 2 ₁	4
			<i>C</i>	<i>C</i> 222, <i>C</i> 222 ₁	8
			<i>F</i>	<i>F</i> 222	16
			<i>I</i>	<i>I</i> 222, <i>I</i> 2 ₁ 2 ₁ 2 ₁	8
Tetragonal	<i>4/m</i>	4	<i>P</i>	<i>P</i> 4, <i>P</i> 4 ₁ , <i>P</i> 4 ₂ , <i>P</i> 4 ₃	4
			<i>I</i>	<i>I</i> 4, <i>I</i> 4 ₁	8
	<i>4/mmm</i>	422	<i>P</i>	<i>P</i> 422, <i>P</i> 42 ₁ 2, <i>P</i> 4 ₁ 22, <i>P</i> 4 ₁ 2 ₁ 2, <i>P</i> 4 ₂ 22, <i>P</i> 4 ₂ 2 ₁ 2, <i>P</i> 4 ₃ 22, <i>P</i> 4 ₃ 2 ₁ 2	8
			<i>I</i>	<i>I</i> 422, <i>I</i> 42 ₁ 2	16
Trigonal (see note)	-3	3	<i>P</i>	<i>P</i> 3, <i>P</i> 3 ₁ , <i>P</i> 3 ₂	3
			<i>R</i>	<i>R</i> 3	9
	-3 <i>m</i>	312	<i>P</i>	<i>P</i> 312, <i>P</i> 3 ₁ 12, <i>P</i> 3 ₂ 12	6
		321	<i>R</i>	<i>P</i> 321, <i>P</i> 3 ₁ 21, <i>P</i> 3 ₂ 21	6
Hexagonal	<i>6/m</i>	6	<i>P</i>	<i>P</i> 6, <i>P</i> 6 ₁ , <i>P</i> 6 ₂ , <i>P</i> 6 ₃ , <i>P</i> 6 ₄ , <i>P</i> 6 ₅	6
	<i>6/mmm</i>	622	<i>P</i>	<i>P</i> 622, <i>P</i> 6 ₁ 22, <i>P</i> 6 ₂ 22, <i>P</i> 6 ₃ 22, <i>P</i> 6 ₄ 22, <i>P</i> 6 ₅ 22	12
Cubic	<i>m-3</i>	23	<i>P</i>	<i>P</i> 23, <i>P</i> 2 ₁ 3	12
			<i>F</i>	<i>F</i> 23	48
			<i>I</i>	<i>I</i> 23, <i>I</i> 2 ₁ 3	24
			<i>P</i>	<i>P</i> 432, <i>P</i> 4 ₂ 32, <i>P</i> 4 ₃ 32, <i>P</i> 4 ₁ 32	24
			<i>F</i>	<i>F</i> 432, <i>F</i> 4 ₁ 32	96
	<i>I</i>	<i>I</i> 432, <i>I</i> 4 ₁ 32	48		

Crystal System	Laue Class	Crystal Class	Lattice Centring	Enantiomorphic Space Groups	Number of Asymmetric Units per Unit Cell
Triclinic	-1	1	<i>P</i>	<i>P</i> 1	1
Monoclinic	<i>2/m</i>	2	<i>P</i>	<i>P</i> 2, <i>P</i> 2 ₁	2
			<i>C</i>	<i>C</i> 2	4
Orthorhombic	<i>mmm</i>	222	<i>P</i>	<i>P</i> 222, <i>P</i> 222 ₁ , <i>P</i> 2 ₁ 2 ₁ 2, <i>P</i> 2 ₁ 2 ₁ 2 ₁	4
			<i>C</i>	<i>C</i> 222, <i>C</i> 222 ₁	8
			<i>F</i>	<i>F</i> 222	16
			<i>I</i>	<i>I</i> 222, <i>I</i> 2 ₁ 2 ₁ 2 ₁	8
Tetragonal	<i>4/m</i>	4	<i>P</i>	<i>P</i> 4, <i>P</i> 4 ₁ , <i>P</i> 4 ₂ , <i>P</i> 4 ₃	4
			<i>I</i>	<i>I</i> 4, <i>I</i> 4 ₁	8
	<i>4/mmm</i>	422	<i>P</i>	<i>P</i> 422, <i>P</i> 42 ₁ 2, <i>P</i> 4 ₁ 22, <i>P</i> 4 ₁ 2 ₁ 2, <i>P</i> 4 ₂ 22, <i>P</i> 4 ₂ 2 ₁ 2, <i>P</i> 4 ₃ 22, <i>P</i> 4 ₃ 2 ₁ 2	8
			<i>I</i>	<i>I</i> 422, <i>I</i> 42 ₁ 2	16
Trigonal (see note)	-3	3	<i>P</i>	<i>P</i> 3, <i>P</i> 3 ₁ , <i>P</i> 3 ₂	3
			<i>R</i>	<i>R</i> 3	9
	-3 <i>m</i>	312	<i>P</i>	<i>P</i> 312, <i>P</i> 3 ₁ 12, <i>P</i> 3 ₂ 12	6
		321	<i>R</i>	<i>P</i> 321, <i>P</i> 3 ₁ 21, <i>P</i> 3 ₂ 21	6
Hexagonal	<i>6/m</i>	6	<i>P</i>	<i>P</i> 6, <i>P</i> 6 ₁ , <i>P</i> 6 ₂ , <i>P</i> 6 ₃ , <i>P</i> 6 ₄ , <i>P</i> 6 ₅	6
	<i>6/mmm</i>	622	<i>P</i>	<i>P</i> 622, <i>P</i> 6 ₁ 22, <i>P</i> 6 ₂ 22, <i>P</i> 6 ₃ 22, <i>P</i> 6 ₄ 22, <i>P</i> 6 ₅ 22	12
Cubic	<i>m-3</i>	23	<i>P</i>	<i>P</i> 23, <i>P</i> 2 ₁ 3	12
			<i>F</i>	<i>F</i> 23	48
			<i>I</i>	<i>I</i> 23, <i>I</i> 2 ₁ 3	24
			<i>P</i>	<i>P</i> 432, <i>P</i> 4 ₂ 32, <i>P</i> 4 ₃ 32, <i>P</i> 4 ₁ 32	24
			<i>F</i>	<i>F</i> 432, <i>F</i> 4 ₁ 32	96
	<i>I</i>	<i>I</i> 432, <i>I</i> 4 ₁ 32	48		



A unidade assimétrica em cristalografia de proteínas é normalmente determinada pela forma da proteína.

Difração por uma Célula unitária

vs

Difração por um Cristal de N células

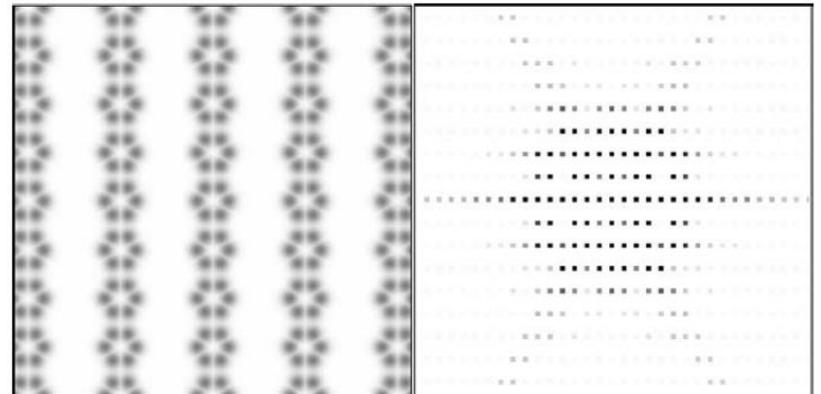
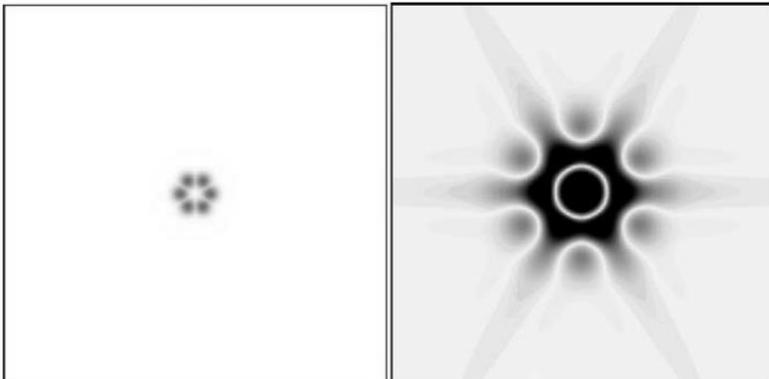
$$\underline{F}_{\text{cell}}(\underline{S}) = \sum f_j e^{2\pi i \underline{r}_j \cdot \underline{S}}$$

$$\underline{F}_{\text{cryst}}(\underline{S}) = N \sum f_j e^{2\pi i \underline{r}_j \cdot \underline{S}}$$

$$\underline{S} \cdot \underline{a} = h$$

$$\underline{S} \cdot \underline{b} = k$$

$$\underline{S} \cdot \underline{c} = l$$



Difração por uma Célula unitária

vs

Difração por um Cristal de N células

$$\underline{\mathbf{F}}_{\text{cell}}(\underline{\mathbf{S}}) = \sum f_j e^{2\pi i \underline{\mathbf{r}}_j \cdot \underline{\mathbf{S}}}$$

Posição dos átomos em
coordenados fracionais

$$\underline{\mathbf{r}}_j = \underline{\mathbf{a}}x_j + \underline{\mathbf{b}}y_j + \underline{\mathbf{c}}z_j$$

$$\underline{\mathbf{r}}_j \cdot \underline{\mathbf{S}} = \underline{\mathbf{S}} \cdot \underline{\mathbf{a}}x_j + \underline{\mathbf{S}} \cdot \underline{\mathbf{b}}y_j + \underline{\mathbf{S}} \cdot \underline{\mathbf{c}}z_j$$

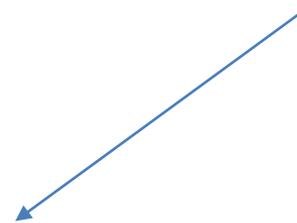
$$\underline{\mathbf{r}}_j \cdot \underline{\mathbf{S}} = hx_j + ky_j + lz_j$$

$$\underline{\mathbf{F}}_{\text{cryst}}(\underline{\mathbf{S}}) = N \sum f_j e^{2\pi i \underline{\mathbf{r}}_j \cdot \underline{\mathbf{S}}}$$

$$\underline{\mathbf{S}} \cdot \underline{\mathbf{a}} = h$$

$$\underline{\mathbf{S}} \cdot \underline{\mathbf{b}} = k$$

$$\underline{\mathbf{S}} \cdot \underline{\mathbf{c}} = l$$



Difração por uma Célula unitária

vs

Difração por um Cristal de N células

$$\underline{\mathbf{F}}_{\text{cell}}(\underline{\mathbf{S}}) = \sum f_j e^{2\pi i \underline{\mathbf{r}}_j \cdot \underline{\mathbf{S}}}$$

Posição dos átomos em
coordenados fracionais

$$\underline{\mathbf{r}}_j = \underline{\mathbf{a}}x_j + \underline{\mathbf{b}}y_j + \underline{\mathbf{c}}z_j$$

$$\underline{\mathbf{r}}_j \cdot \underline{\mathbf{S}} = \underline{\mathbf{S}} \cdot \underline{\mathbf{a}}x_j + \underline{\mathbf{S}} \cdot \underline{\mathbf{b}}y_j + \underline{\mathbf{S}} \cdot \underline{\mathbf{c}}z_j$$

$$\underline{\mathbf{r}}_j \cdot \underline{\mathbf{S}} = hx_j + ky_j + lz_j$$



$$\underline{\mathbf{F}}_{\text{hkl}} = \sum_{\text{atoms}} f_j e^{2\pi i (hx_j + ky_j + lz_j)}$$

$$\underline{\mathbf{F}}_{\text{cryst}}(\underline{\mathbf{S}}) = N \sum f_j e^{2\pi i \underline{\mathbf{r}}_j \cdot \underline{\mathbf{S}}}$$

$$\underline{\mathbf{S}} \cdot \underline{\mathbf{a}} = h$$

$$\underline{\mathbf{S}} \cdot \underline{\mathbf{b}} = k$$

$$\underline{\mathbf{S}} \cdot \underline{\mathbf{c}} = l$$

Logo: cada átomo (j) na célula unitária contribui para cada fator de estrutura (reflexão)

Expressão que descreve a contribuição de cada átomo (j) a cada fator de estrutura (F_{hkl})

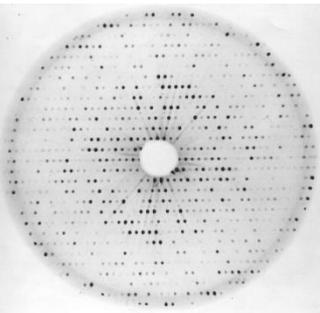
Fator de estrutura
(se manifesta como uma reflexão no padrão de difração)

$$F_{hkl} = \sum_{\text{atoms}} f_j \exp[2\pi i(hx_j + ky_j + lz_j)]$$

Fator de espalhamento do átomo
~ número de elétrons, corrigido pelo fator de temperatura, etc

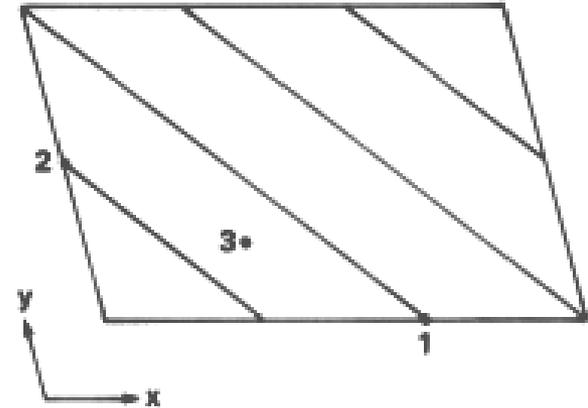
hkl são as índices do plano de reflexão

Fase da contribuição de cada átomo determinada pela índice h k l e as coordenadas fracionais x_j , y_j , e z_j



Por exemplo, imagine uma célula unitária com 3 átomos iguais. Quais são as fases de espalhamento de cada átomo para a reflexão (3,2,0)?

$$f_{hkl} = f_j \exp[2\pi i(hx_j + ky_j + lz_j)]$$



Para átomo 1. $x, y, z = 2/3, 0, 0$:

$$\text{logo fase} = \alpha = 2\pi(hx + ky + lz) = 2\pi(3 \times 2/3 + 2 \times 0 + 0) = 4\pi = 0$$

Notar: O átomo está no plano.

Para átomo 2. $x, y, z = 0, 1/2, 0$:

$$\text{logo fase} = \alpha = 2\pi(hx + ky + lz) = 2\pi(3 \times 0 + 2 \times 1/2 + 0) = 2\pi = 0$$

Notar: O átomo está no plano.

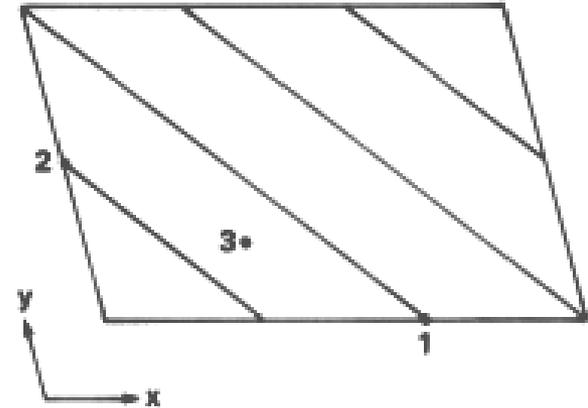
Para átomo 3. $x, y, z = 1/3, 1/4, 0$:

$$\text{logo fase} = \alpha = 2\pi(hx + ky + lz) = 2\pi(3 \times 1/3 + 2 \times 1/4 + 0) = 3\pi = \pi = 180^\circ.$$

Notar: O átomo está na metade de caminho entre os dois planos.

Por exemplo, imagine uma célula unitária com 3 átomos iguais. Quais são as fases de espalhamento de cada átomo para a reflexão (3,2,0)?

$$f_{hkl} = f_j \exp[2\pi i(hx_j + ky_j + lz_j)]$$



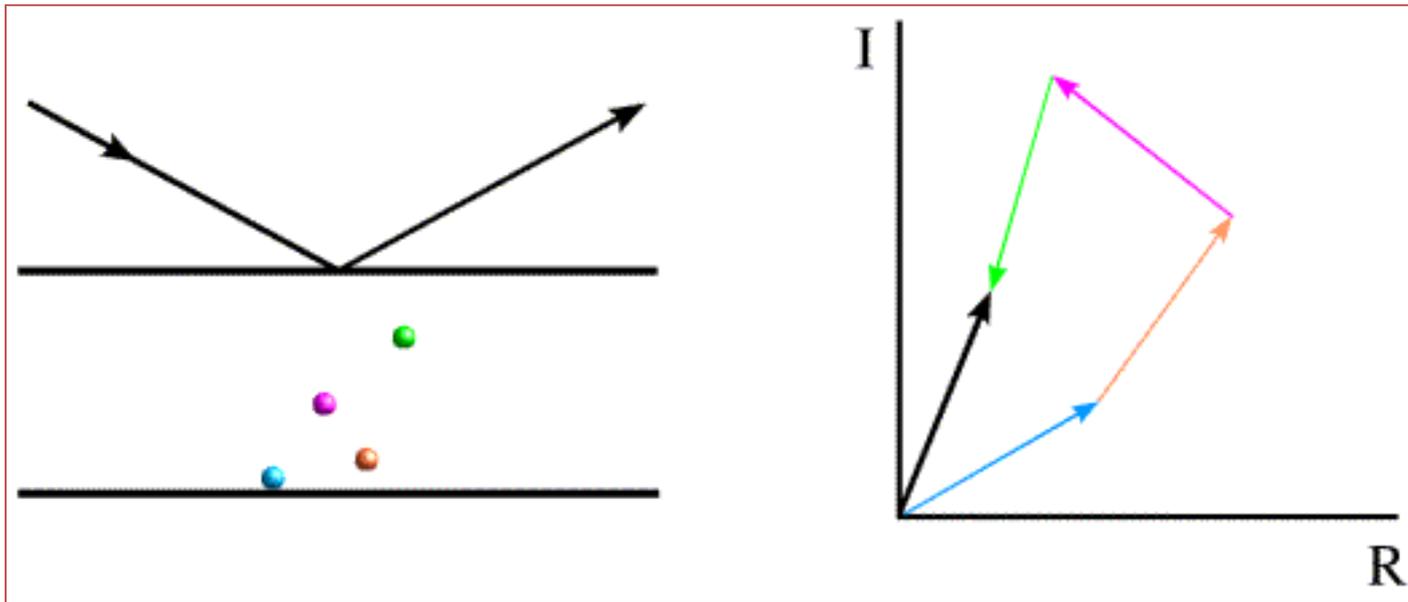
Para átomo 1. $x, y, z = 2/3, 0, 0$: logo $2\pi(hx + ky + lz) = 2\pi(3 \times 2/3 + 2 \times 0 + 0) = 4\pi = 0$
O átomo está no plano.

Para átomo 2. $x, y, z = 0, 1/2, 0$: logo $2\pi(hx + ky + lz) = 2\pi(3 \times 0 + 2 \times 1/2 + 0) = 2\pi = 0$
O átomo está no plano.

Para átomo 3. $x, y, z = 1/3, 1/4, 0$: logo $2\pi(hx + ky + lz) = 2\pi(3 \times 1/3 + 2 \times 1/4 + 0) = 3\pi = \pi$
O átomo está na metade de caminho entre os dois planos.

Desafio: Qual será a amplitude e a fase final do fator de estrutura $F_{3,2,0}$ para esta célula unitária?

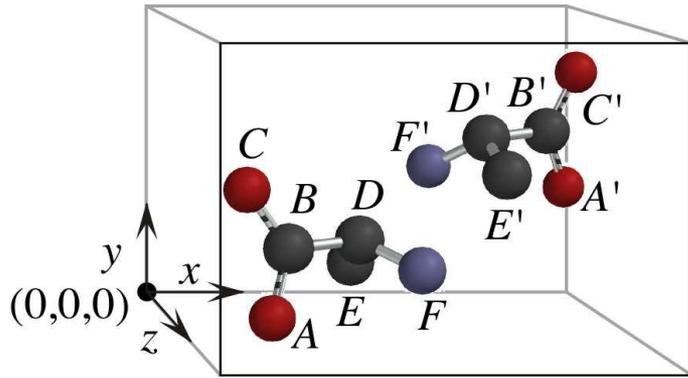
- A amplitude de espalhamento depende de número de eletrons em cada átomo.
- A fase depende da distância fracional que cada atomo se encontra relativo aos planos de reflexão.



Espalhamento de planos de rede

Fatores de estrutura atômicos somam como números complexos ou vetores no plano complexo

Duas maneiras diferentes de descrever o Fator de Estrutura



$$\mathbf{f}_{hkl(j)} = f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$

$$\mathbf{E}_{hkl} = \sum \mathbf{f}_{hkl(j)} = \sum f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$

\mathbf{E}_{hkl} escrito como a soma de contribuições de cada **átomo**

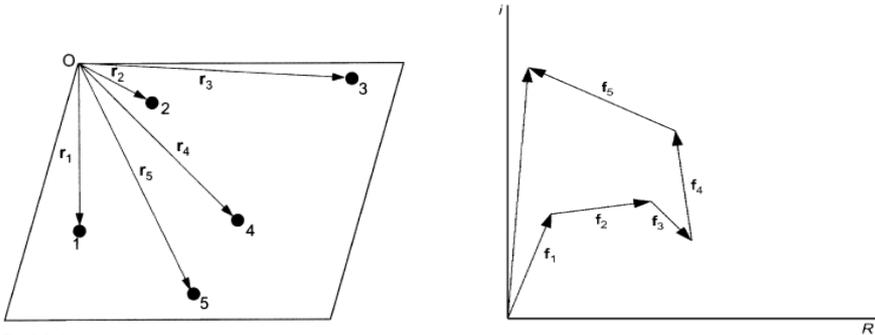
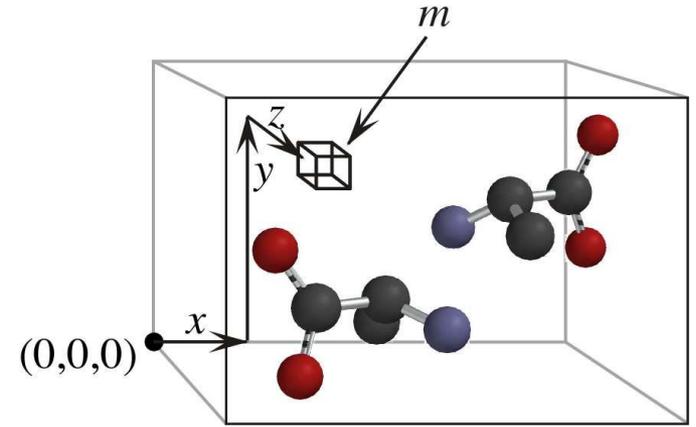


Fig. 3.8 Atomic positions in a unit cell.



$$F_{hkl} = \int_x \int_y \int_z \rho(x, y, z) e^{2\pi i(hx + ky + lz)} dx dy dz$$

$$F_{hkl} = \int_V \rho(x, y, z) e^{2\pi i(hx + ky + lz)} dV,$$

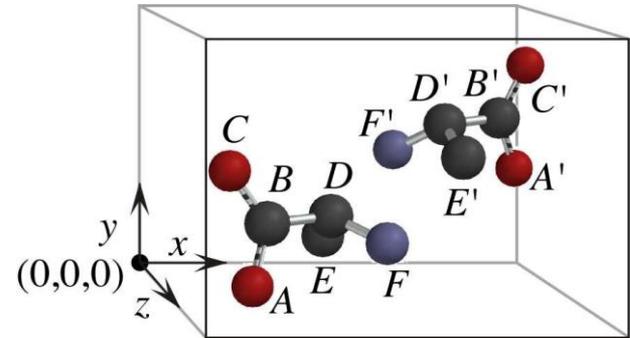
\mathbf{E}_{hkl} pode ser escrito como a soma de contribuições de cada **elemento de volume de densidade eletrônica** da célula unitária.

Assim podemos usar um integral sobre o volume da célula unitária.

\mathbf{E}_{hkl} é a transformada de Fourier de $\rho(x, y, z)$

Por uma célula unitária contendo n átomos, o fator de estrutura F_{hkl} é o soma de todos os valores f_{hkl} para os átomos individuais .

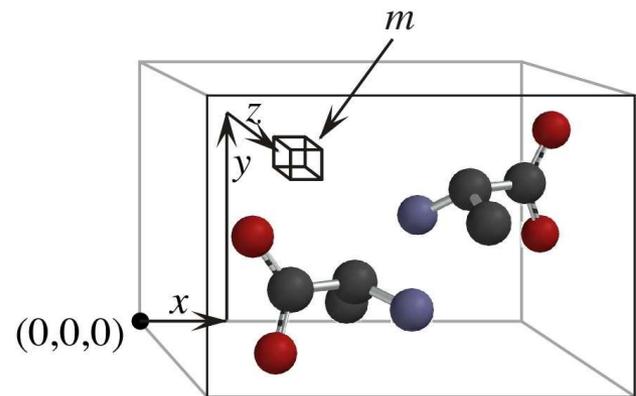
$$F_{hkl} = \sum_{j=1}^n f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$



F_{hkl} pode ser escrito como o soma das contribuições de cada elemento de volume de densidade eletrônica na célula unitária.

$$F_{hkl} = \int_x \int_y \int_z \rho(x, y, z) e^{2\pi i(hx + ky + lz)} dx dy dz$$

Logo: O fator de estrutura (F_{hkl}) é o Transformada de Fourier da densidade eletrônica ($\rho(x,y,z)$)



A transformada de Fourier é reversível.

Logo , a densidade eletrônica é a transformada de Fourier dos fatores de estrutura.

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} e^{-2\pi i(hx + ky + lz)}$$

OBJETO

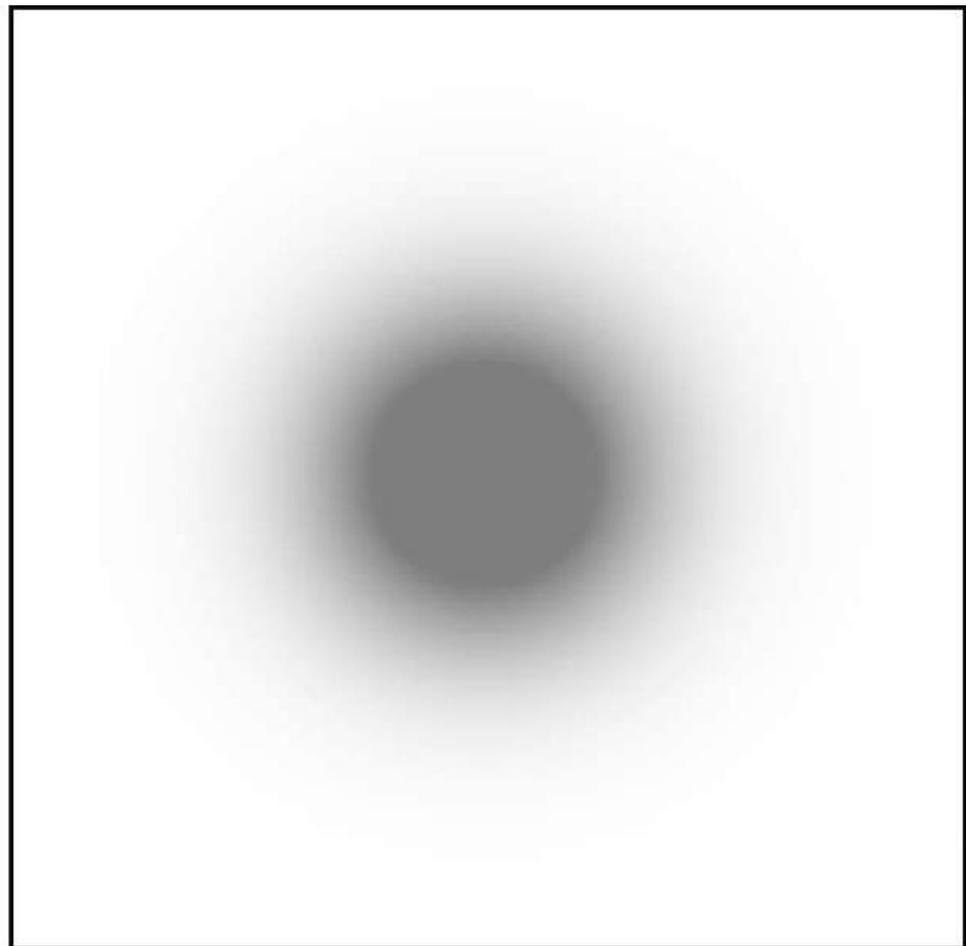
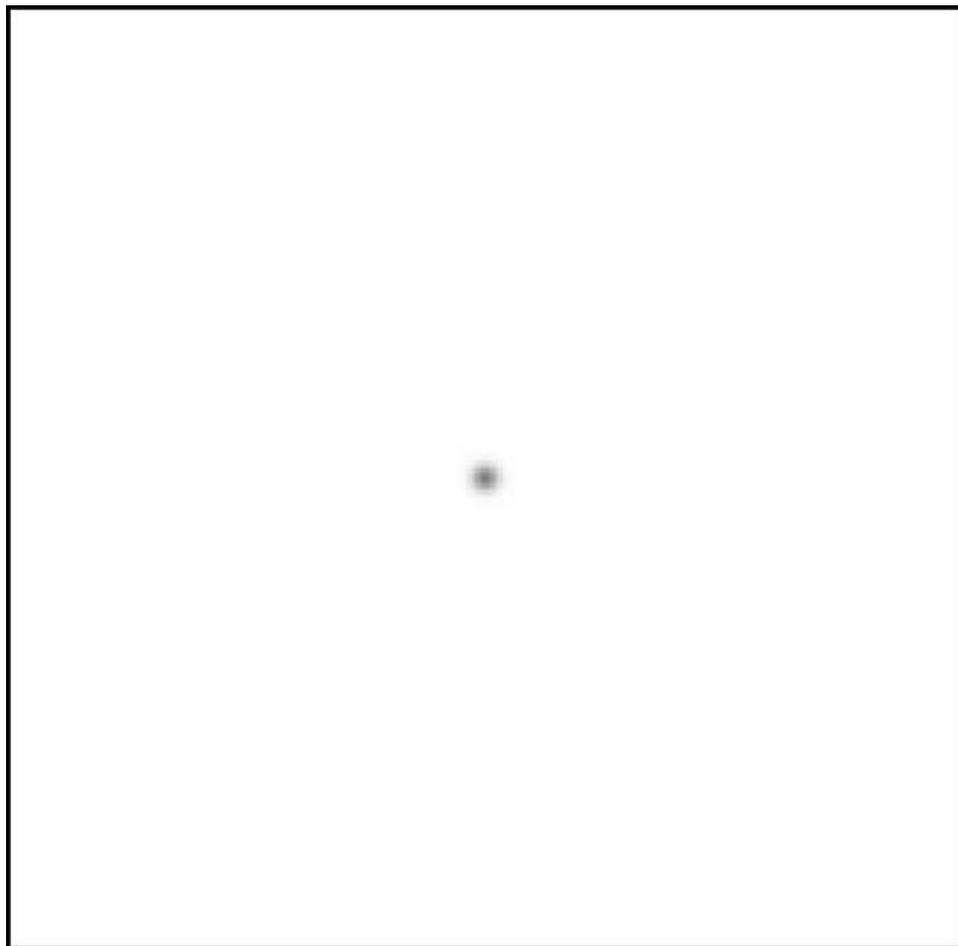
FT



PADRÃO DE DIFRAÇÃO



FT reverso



OBJETO

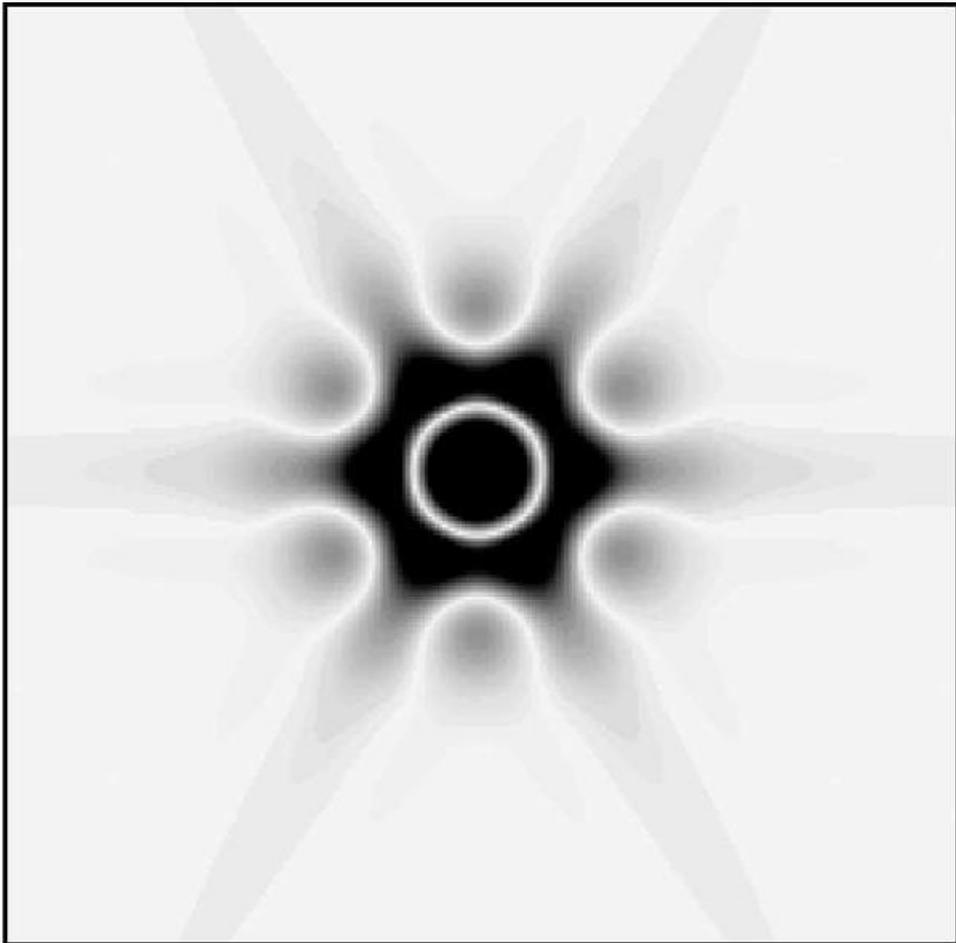
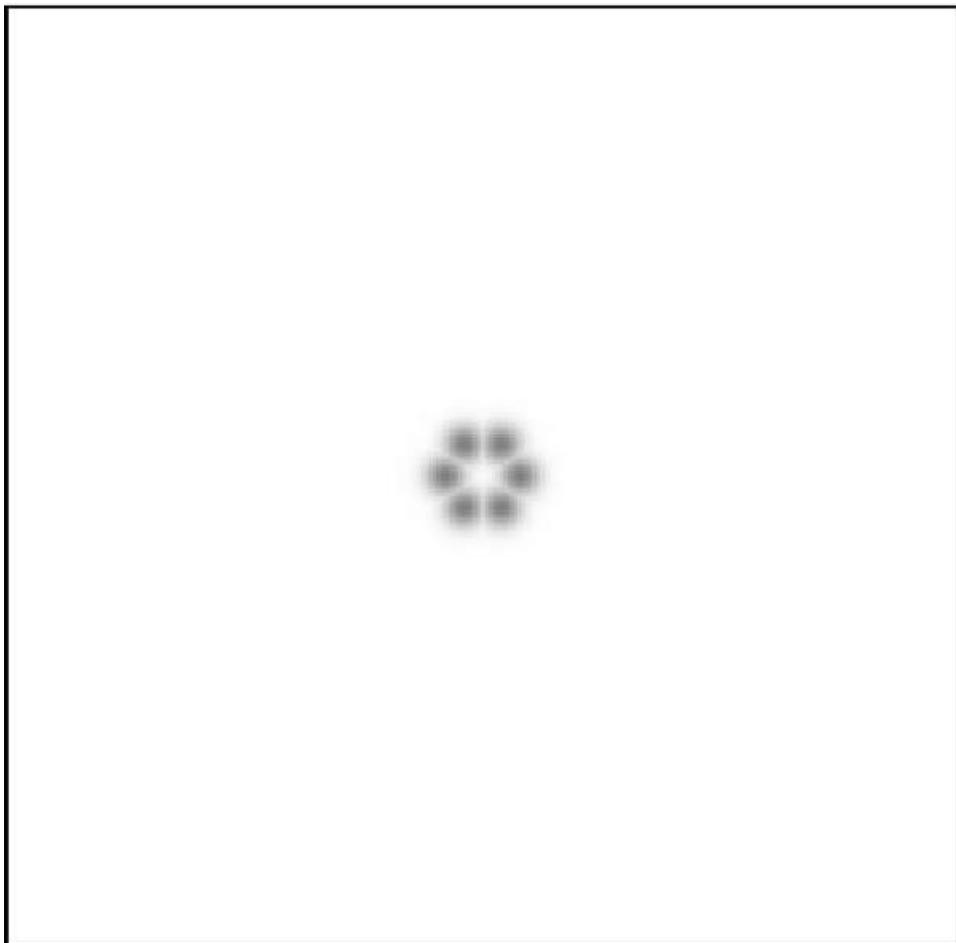
FT



PADRÃO DE DIFRAÇÃO



FT reverso



OBJETO

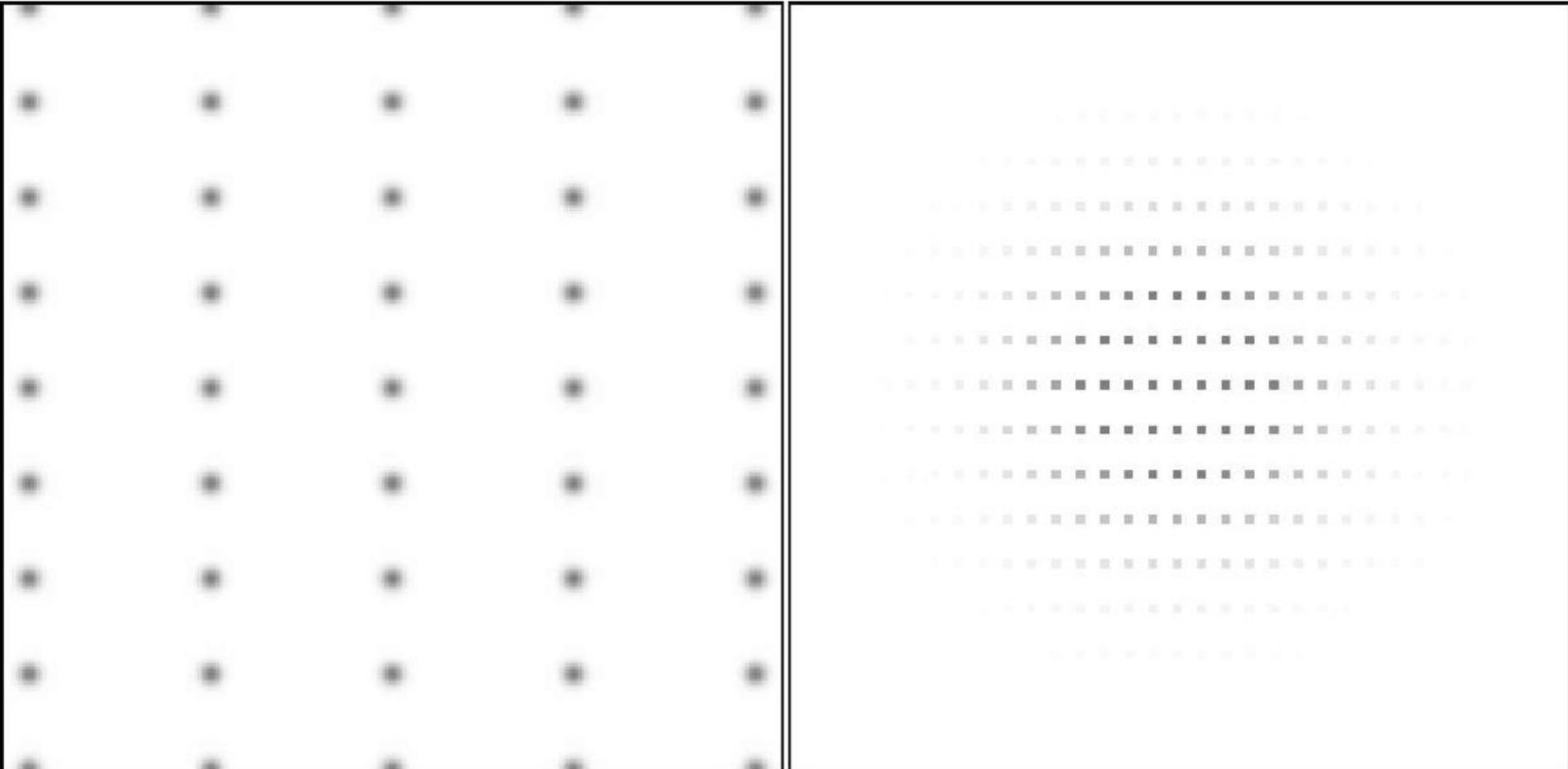
FT

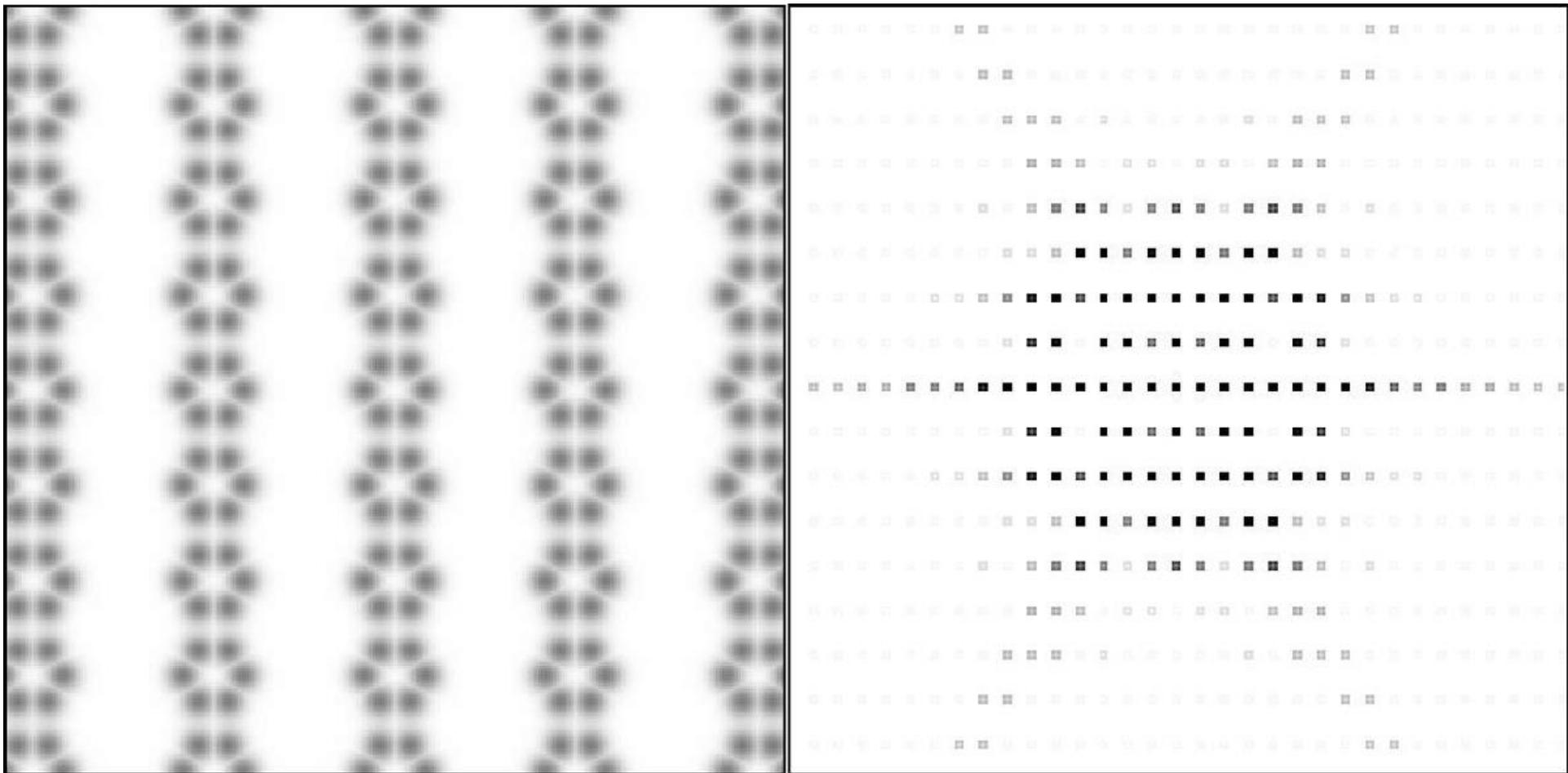
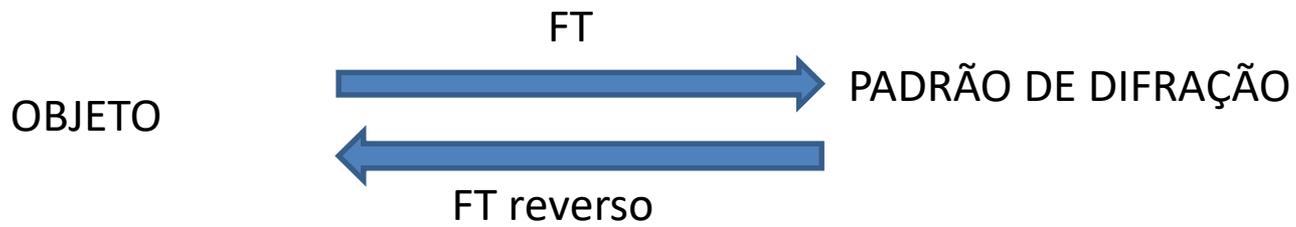


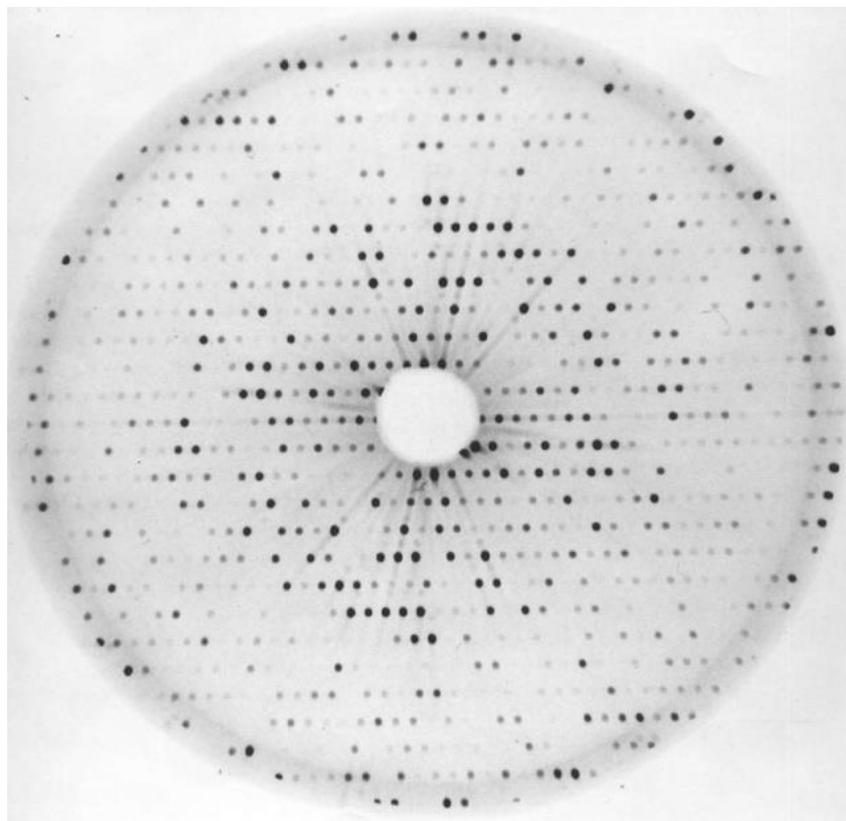
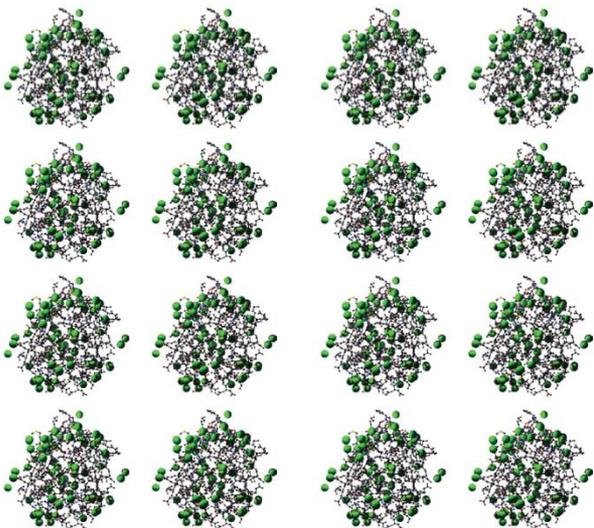
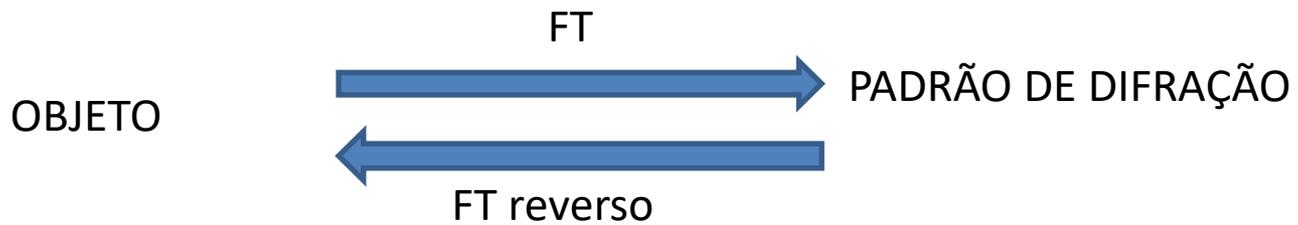
PADRÃO DE DIFRAÇÃO



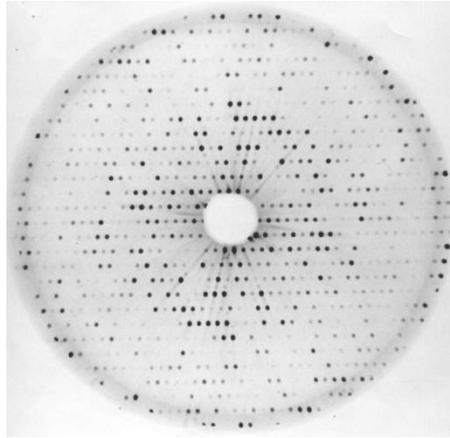
FT reverso







F_{hkl} é a transformada de Fourier de $\rho(x,y,z)$.

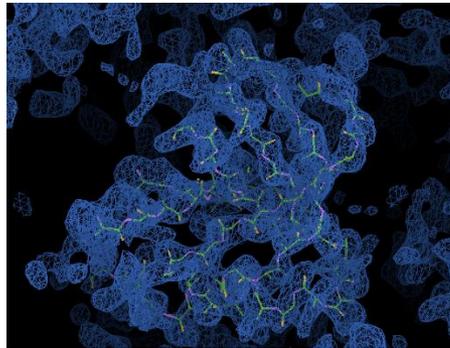


$$F_{hkl} = \sum_{j=1}^n f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$

... é também uma soma de Fourier

$$F_{hkl} = \int_x \int_y \int_z \rho(x, y, z) e^{2\pi i(hx + ky + lz)} dx dy dz,$$

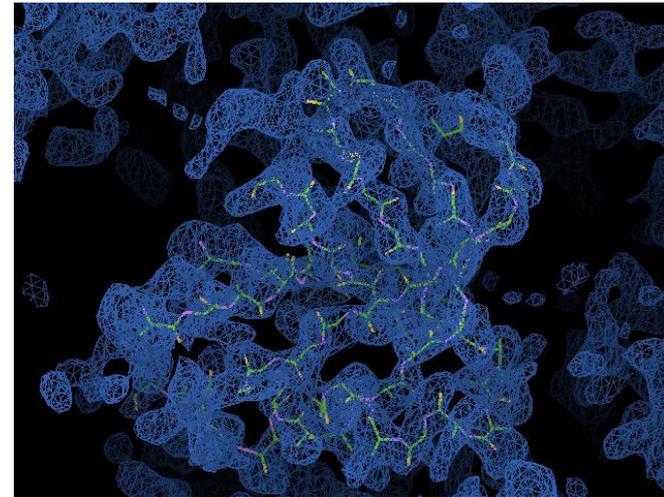
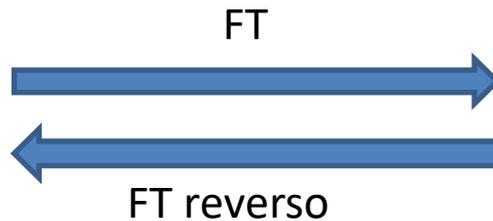
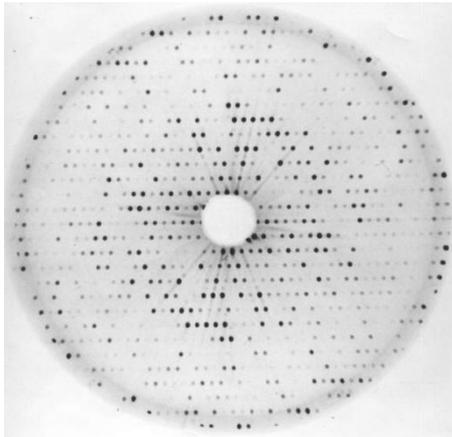
Então, a densidade eletrônica é a FT reversa dos Fatores de Estrutura



$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} e^{-2\pi i(hx + ky + lz)}$$

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} e^{-2\pi i(hx+ky+lz)},$$

Esta equação mostra para nós como determinar a densidade eletrônica em cada ponto (x, y, z) da célula unitária: somar todos os Fatores de estrutura.

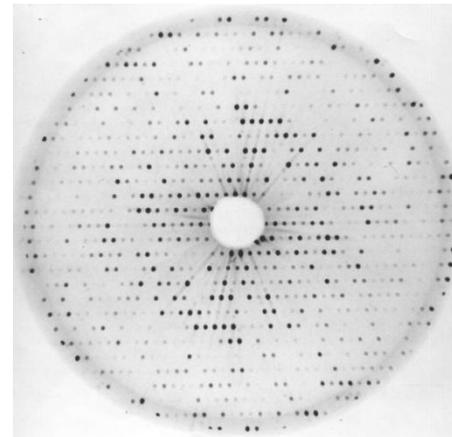


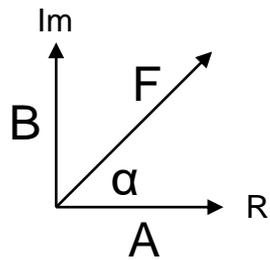
$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} e^{-2\pi i(hx+ky+lz)},$$

Um fator de estrutura (\underline{F}_{hkl}) descreve uma onda que é descrita por 3 parâmetros: amplitude, frequência, e fase.

Experimentalmente somente medimos:

- As indices (h,k,l) – relacionada a frequência que os planos de Miller cortam a célula unitária
- As intensidades (I) – relacionada à amplitude do fator de estrutura : $I = |\underline{F}_{hkl}|^2$
- **Não há informação sobre as fases de \underline{F}_{hkl}**





PROBLEMA DAS FASES

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} e^{-2\pi i(hx+ky+lz)}$$

Lembrando que F_{hkl} é um vetor com fase $\alpha_{hkl} = 2\pi\alpha'_{hkl}$

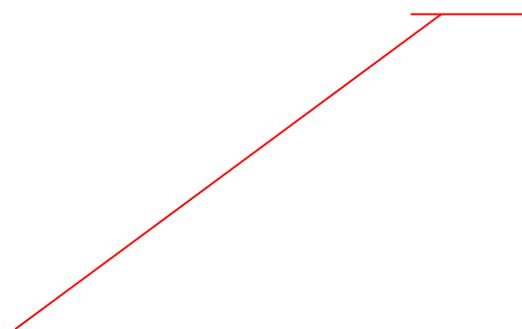
$$F_{hkl} = |F_{hkl}| \times e^{i\alpha_{hkl}} = |F_{hkl}| \times e^{2\pi i\alpha'_{hkl}}$$

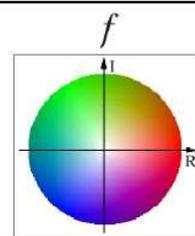
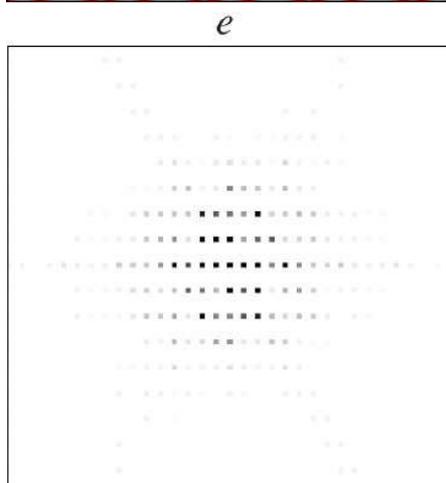
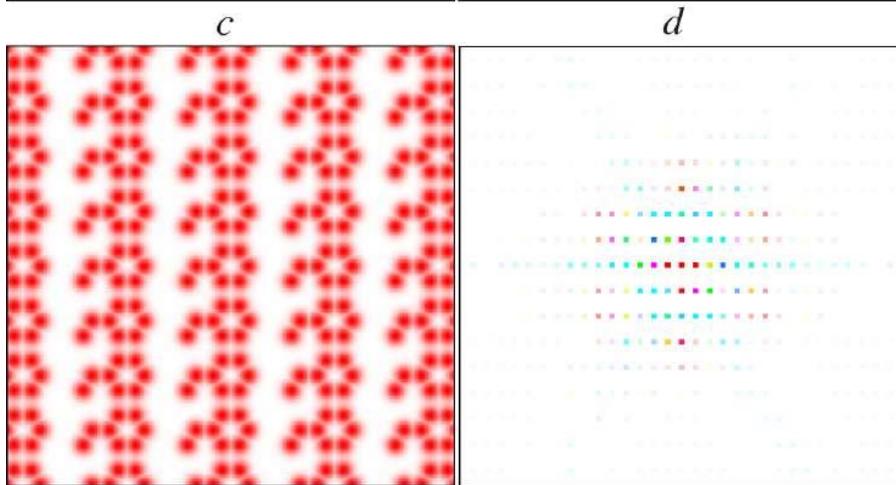
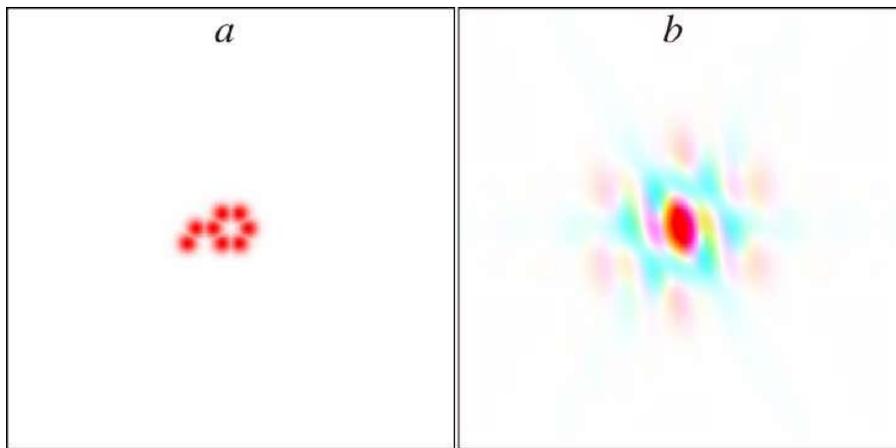
Assim podemos expressar a densidade eletrônica como uma função das amplitudes conhecidas e as fases desconhecidas:

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}| e^{-2\pi i(hx+ky+lz - \alpha'_{hkl})}$$

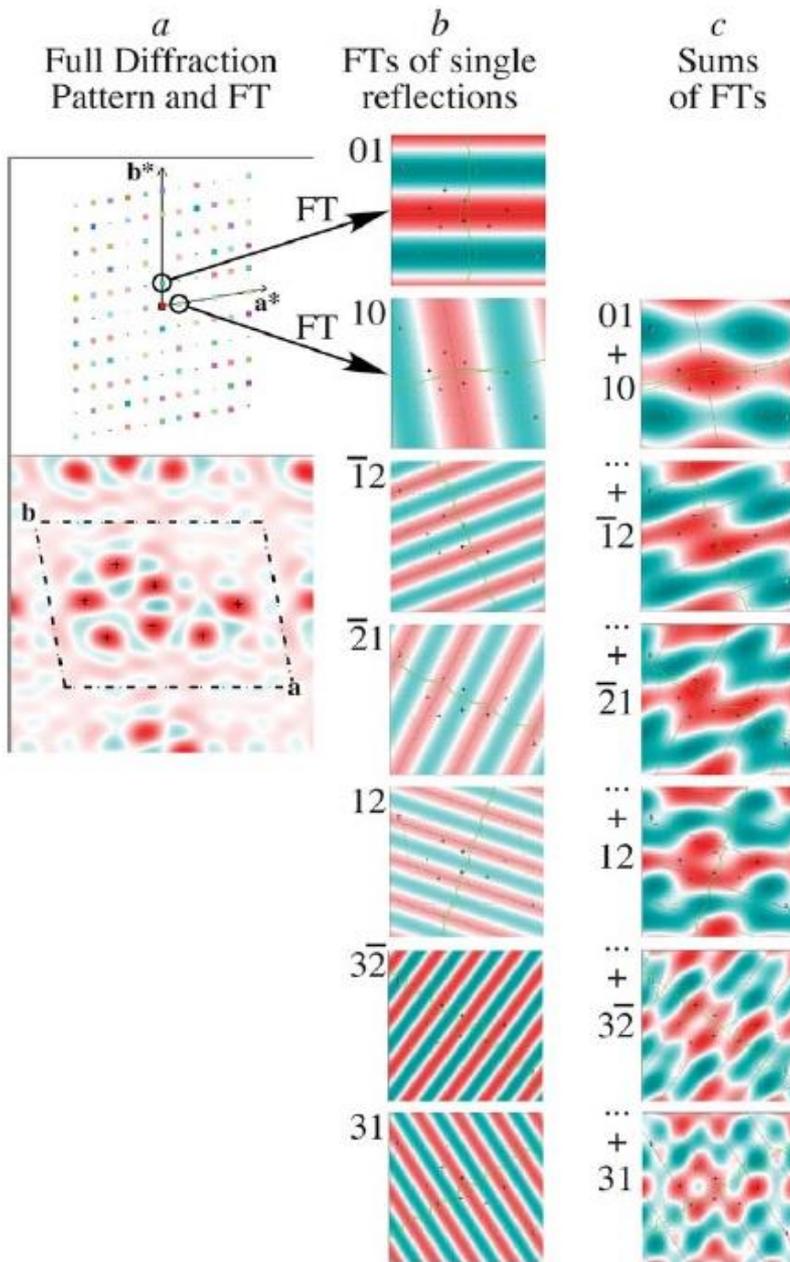
$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}| e^{-2\pi i(hx + ky + lz - \alpha'_{hkl})}$$

O problema das fases





$$F_{hkl} = \sum_{j=1}^N f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$



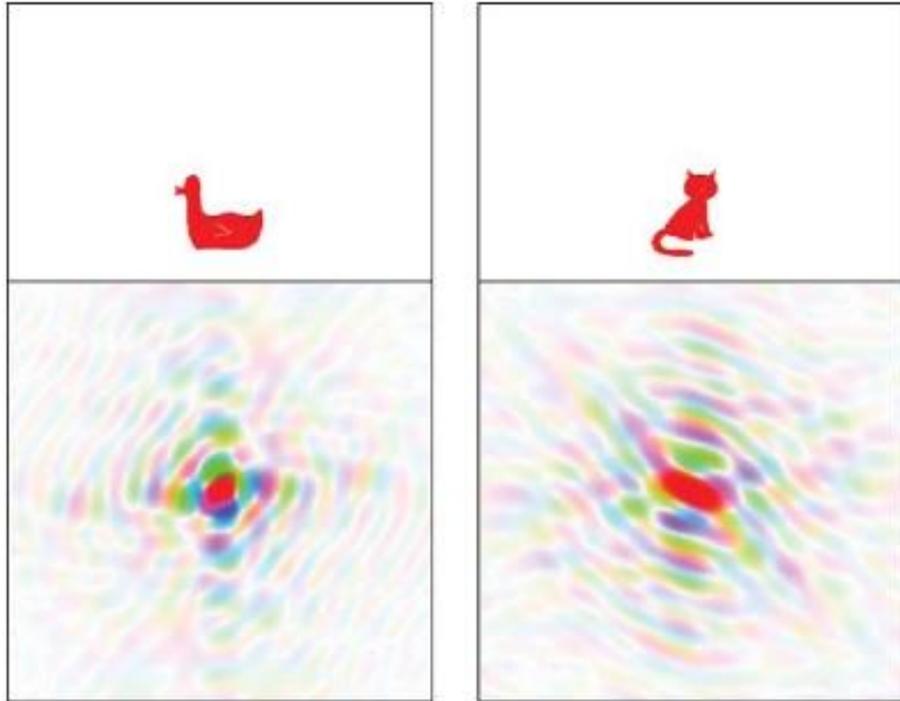
$$F_{hkl} = \sum_{j=1}^n f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}| e^{-2\pi i(hx + ky + lz - \alpha'_{hkl})}$$

A FT de qualquer única reflexão é uma onda simples em espaço 3 dimensional com freqüências em cada direção definidas pela índice hkl.

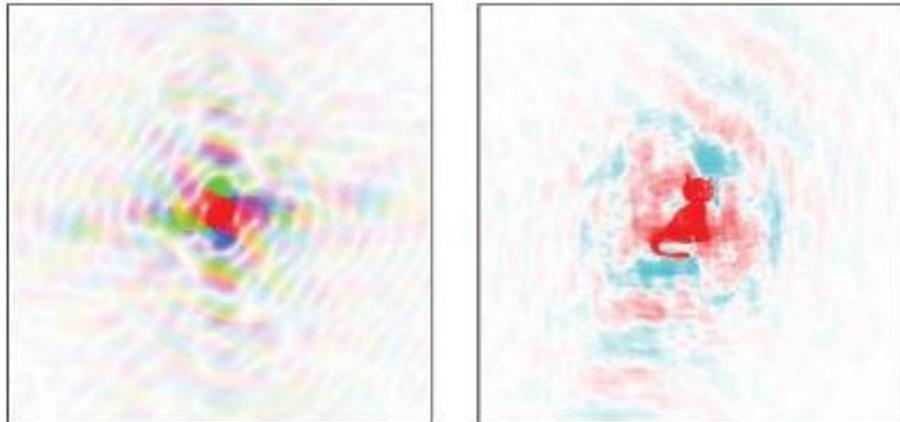
Quando as FTs de 2 ou mais reflexões são somadas, observamos interferência que resulta em um mapa de densidade eletrônica com picos na posições dos átomos.

A IMPORTÂNCIA DAS FASES



a

b

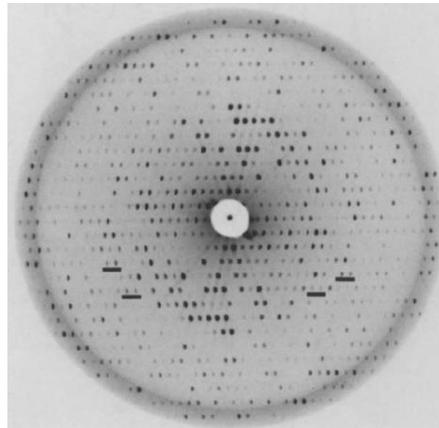
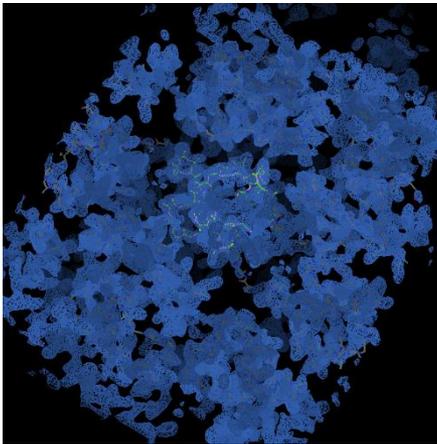


c

d

- a) Pato e FT de um pato
- b) Gato e FT de um gato
- c) Intensidades de FT de pato com fases de FT de gato
- d) FT reverso de (c).

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}| e^{-2\pi i(hx+ky+lz-\alpha'_{hkl})}$$



Esta é a equação que o cristalógrafo tem que solucionar para gerar o mapa de densidade eletrônica

Resolução/Estimativa das fases em cristalografia

Capítulo 6 de Rhodes

Estimação das fases (Capítulo 6 de Rhodes)

- 1) O método de “substituição molecular” (MR)
- 2) O método de “substituição isomorfa múltipla” (MIR)
- 3) O método de “dispersão anômala múltipla” (MAD)

Resolvendo as fases por: “MOLECULAR REPLACEMENT”

Baseado na observação que proteínas que tem seqüências similares também tem estruturas similares

Passo 1: Utilizar a similaridade de sequencias para gerar um modelo molecular da proteína sob estudo baseado numa proteína com estrutura conhecida = “Modelo de homologia”

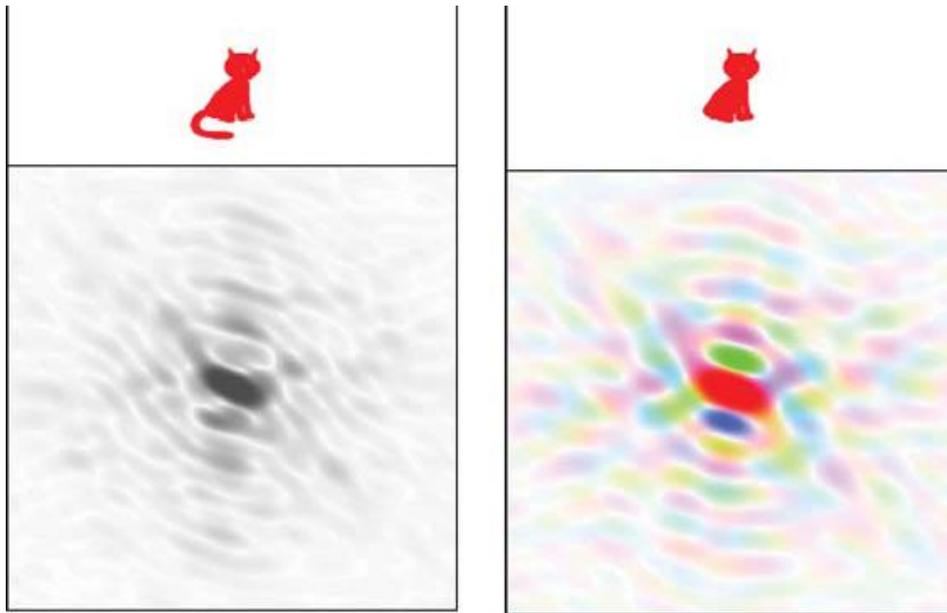
Há vários servidores que podem gerar um modelo inicial com diferentes níveis de sucesso.

Alguns exemplos:

- 1) AlphaFold2 (<https://alphafold.ebi.ac.uk/>)**
- 2) RosettaFold (<https://rosetta.bakerlab.org/>)**
- 3) RaptorX (<http://raptorx.uchicago.edu/>)**
- 4) I-Tasser (<https://zhanglab.dcmf.med.umich.edu/I-TASSER/>)**
- 5) Phyre² (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>)**

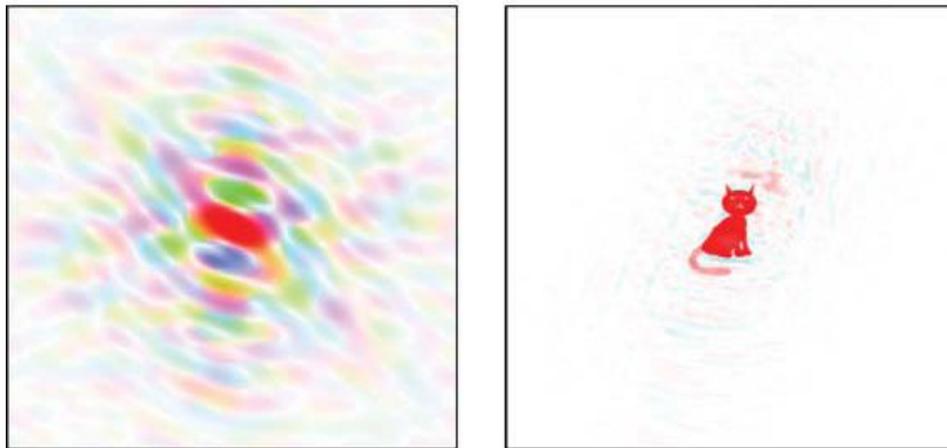
Todos os modelos provavelmente necessitarão alguns ajustes pelo usuário antes de proceder (especialmente para remover pedaços desordenadas com baixa confiabilidade)

IMPORTÂNCIA DAS INTENSIDADES



a

b



c

d

(a) Estrutura desconhecida (gato comum), e sua padrão de difração (não colorida porque as fases não são conhecidas).

(b) Estrutura conhecida (gato Manx, sem cauda) e seu transformada de Fourier calculada (colorida porque o cálculo permite estimar as fases).

(c) Fases do gato Manx combinadas com as intensidades do gato comum.

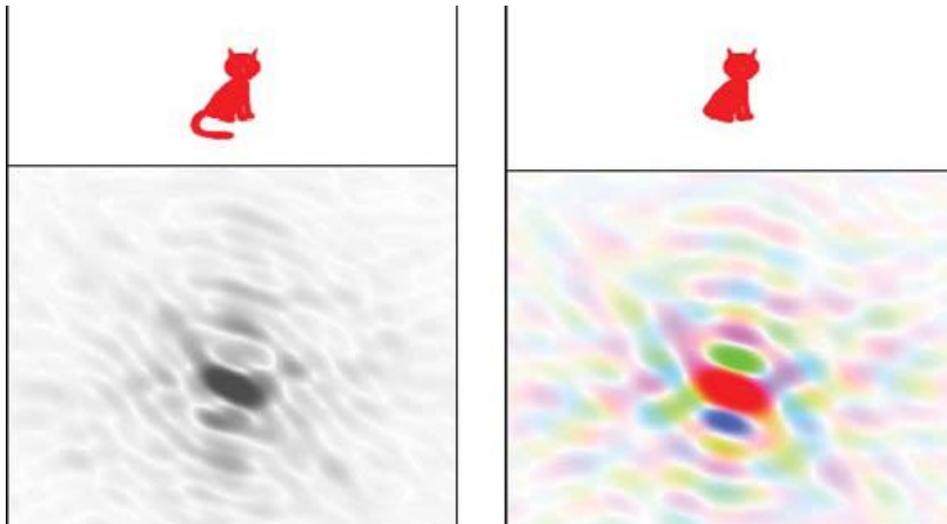
(d) Transformada reversa do (c) Intensidades contêm informação suficiente para revelar diferenças (a cauda) entre o modelo (gato Manx) e a estrutura não conhecida (gato comum).

IMPORTÂNCIA DAS INTENSIDADES

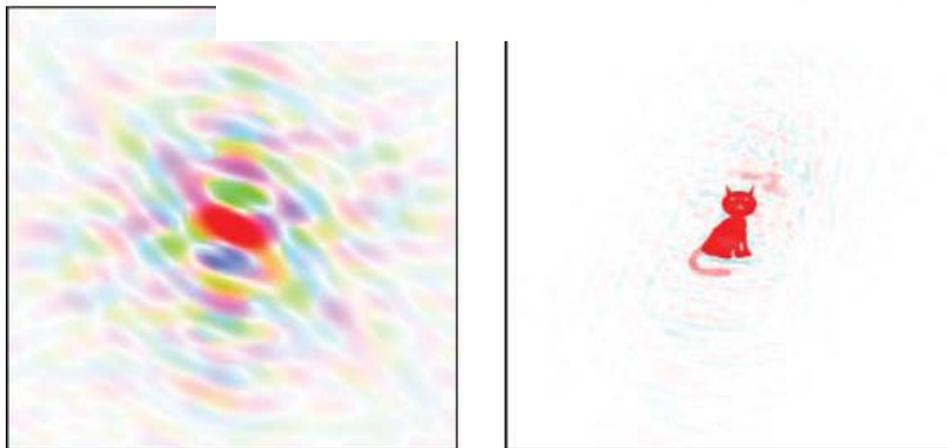
Estimada usando

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

de um modelo de uma proteína com estrutura conhecida e similar



$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l \left| F_{hkl}^{\text{target}} \right| e^{-2\pi i(hx + ky + lz - \alpha_{hkl}^{\text{model}})}$$



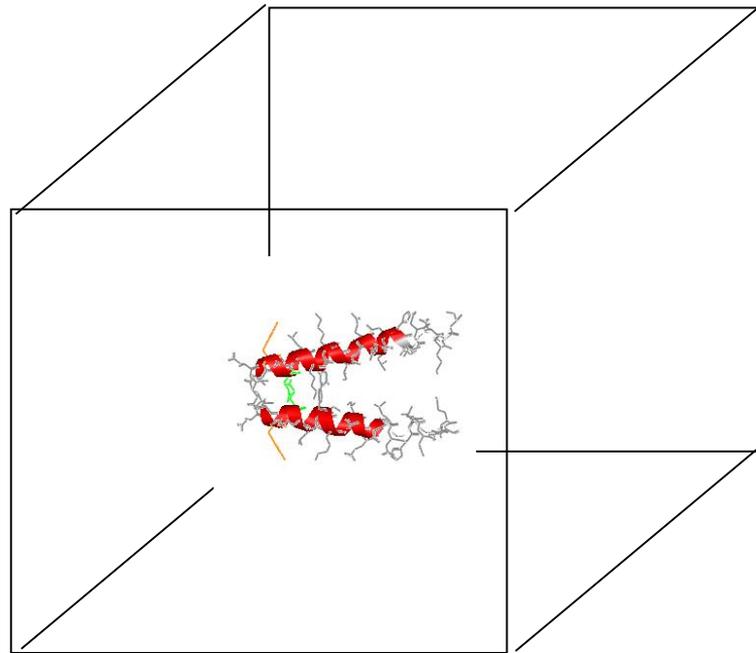
c

d

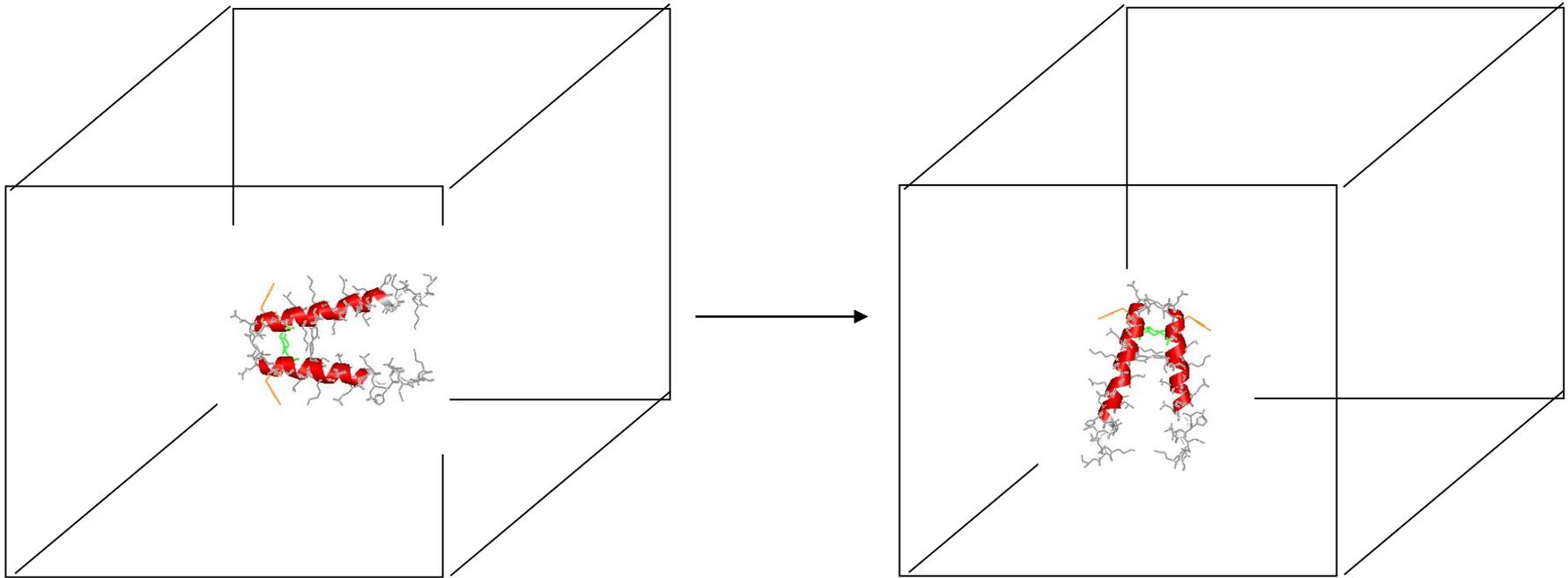
Dados Experimentais
(intensidades das reflexões)^{1/2}

MOLECULAR REPLACEMENT

Passo 2: Colocar o modelo dentro da célula unitária da proteína nova



Passo 3: Procurar achar a posição e a orientação do modelo na célula unitária que produz as $|F_{hkl}|$ mais próximas às da proteína sob estudo. Utilizar este novo modelo para calcular as fases para cada F_{hkl} (estas fases calculadas são chamadas α_{calc} ou ϕ_{calc})

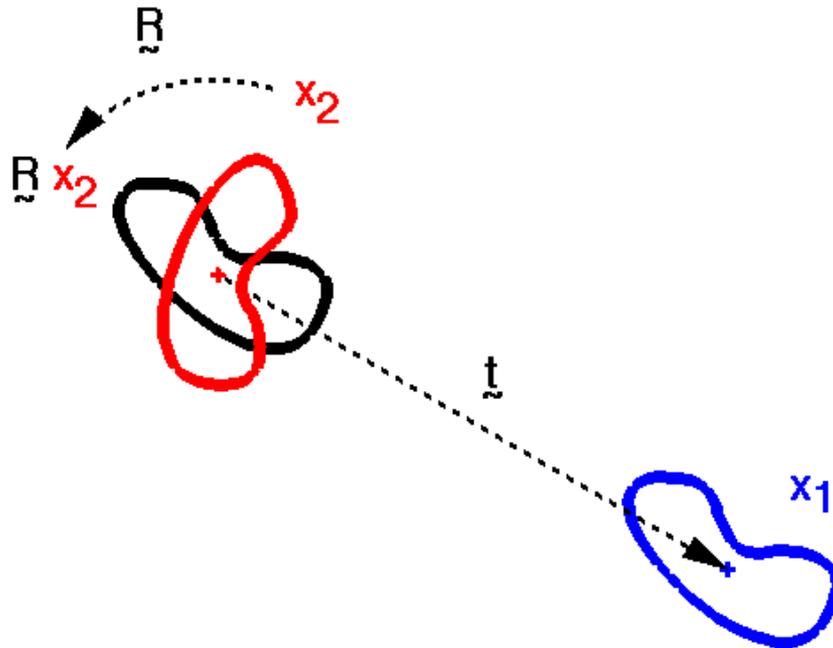


$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}^{\text{tar}}| e^{-2\pi i (hx + ky + lz - \alpha_{hkl}^{\text{calc}})}$$

Há varios programas que podem fazer isso.
Um dos melhores é “Phaser” dentro dos pacotes de CCP4i e Phenix

Em substituição Molecular, separamos o problema 6-dimensional em 2 problemas 3-dimensionais.

- Uma **função de rotação (R)** pode ser computado para encontrar os 3 angulos de rotação para encontrar a correta orientação.
- Depois, o modelo orientado corretamente, pode ser posicionado corretamente dentro da célula com uma **função de translação (t)**.



$$x_1 = \tilde{R} x_2 + \tilde{t}$$

Quando ambos os fatores de estrutura e as fases são conhecidos, um mapa de densidade eletrônica pode ser calculado para cada ponto x, y, z na unidade celular através de um síntese de Fourier.

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}| e^{-2\pi i(hx+ky+lz-\alpha'_{hkl})}$$

(α_{hkl}) ainda são desconhecidas

Função de Patterson

Arthur Lindo Patterson (1935) perguntou, o que resulta se, em vez disso, calculamos um transformada de Fourier das **intensidades (amplitudes quadrada)**, que somente requiere os dados medidos experimentalmente?

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

Fases não requeridas

Esta função produz um “mapa de Patterson” que tem algumas características interessantes.

Função de Patterson

Integral (somatório) dos produtos das densidades eletrônicas de todas as posições (x) com todas as posições (x+u)

Patterson demonstrou que:

$$P(u) = \int \rho(x) \cdot \rho(x+u) du = \frac{1}{V} \sum_h |F_{obs}(h)|^2 \cdot e^{-2\pi i h \cdot u}$$

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

*

Para dois átomos na célula unitária com coordenados x_1, y_1, z_1 e x_2, y_2, z_2 , a função de Patterson (ou sua representação gráfica – o mapa de Patterson) vai apresentar um pico (máximo) no ponto u, v, w seguindo a relação:

$$u = x_1 - x_2$$

$$v = y_1 - y_2$$

$$w = z_1 - z_2$$

E também um pico no

$$u = x_2 - x_1$$

$$v = y_2 - y_1$$

$$w = z_2 - z_1$$

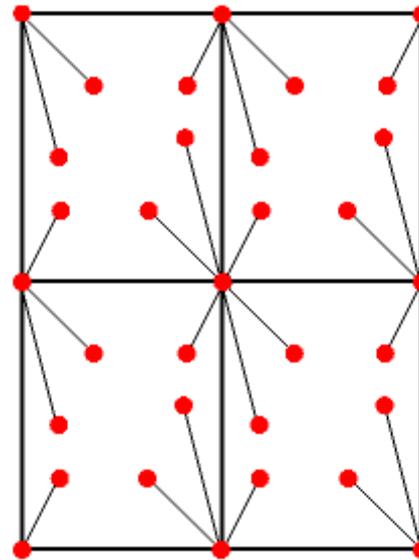
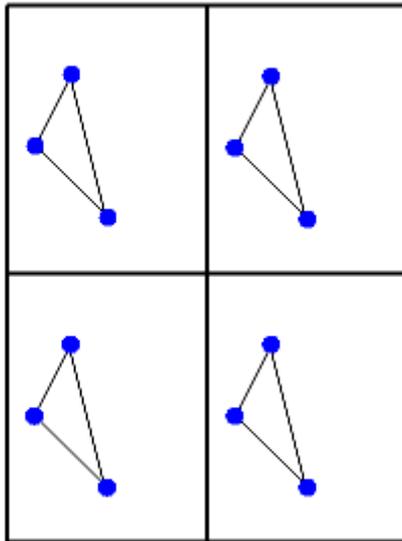
Ou seja: A função de Patterson produz um mapa com “picos” nas pontas dos vetores entre átomos, todos saindo da origem.

***Ver prova desta relação no final deste arquivo**

$$P(u) = \int \rho(x) \cdot \rho(x+u) du = \frac{1}{V} \sum_h |F_{obs}(h)|^2 \cdot e^{-2\pi i h \cdot u}$$

Cada par de átomos gera dois vetores de tamanhos equivalentes mas direção oposta. Logo, o mapa de Patterson tem simetria de inversão.

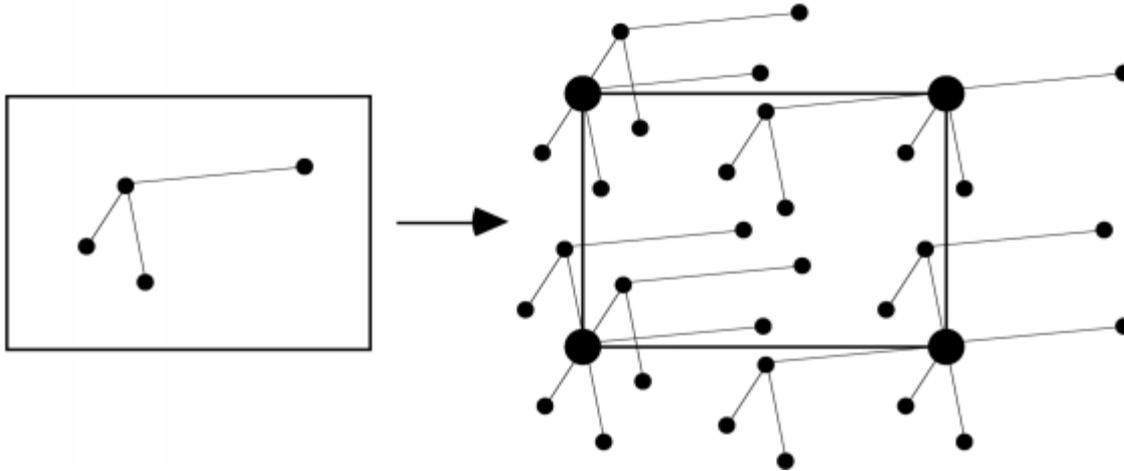
Em duas dimensões



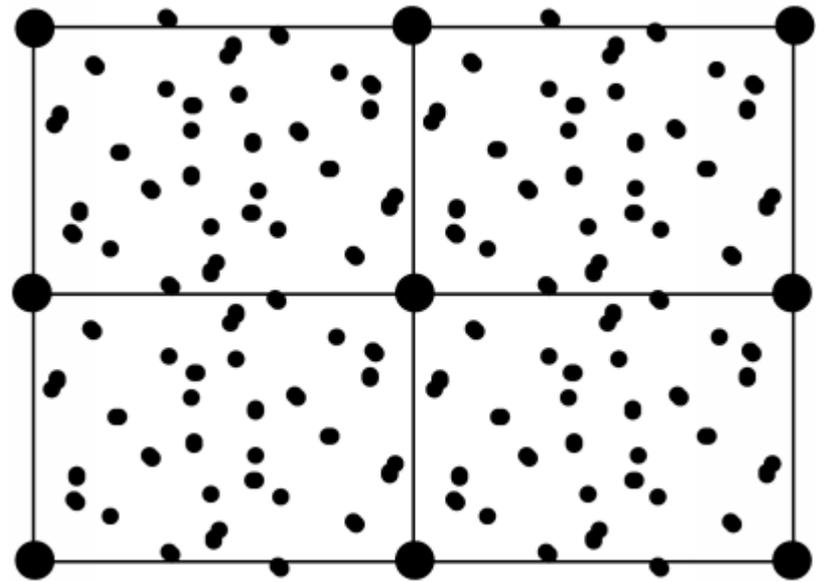
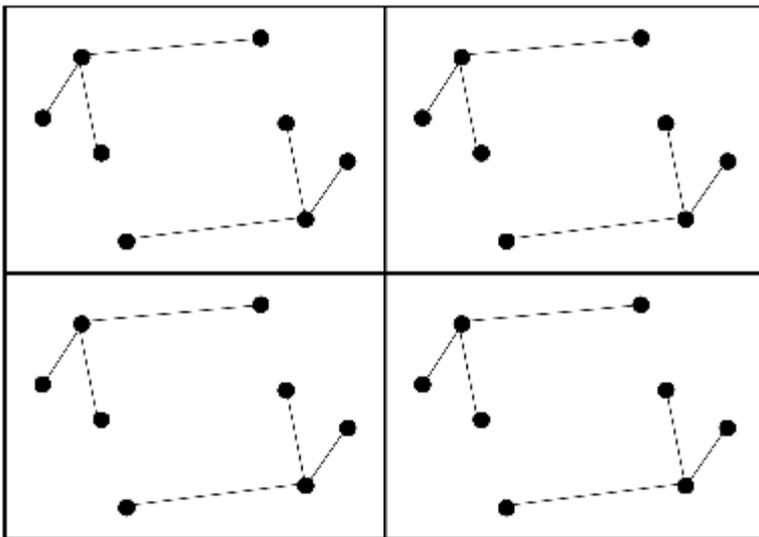
O exemplo apresenta 4 células unitárias de grupo espacial $P1$ com 3 átomos per célula (azul) e 4 células unitárias do grupo especial de Patterson $P-1$ com 6 picos per célula + origem (vermelho).

A altura do pico no mapa de Patterson é proporcional ao produto das Alturas dos dois picos no mapa de densidade eletrônica

Você pode pensar de um mapa de Patterson como a soma de imagens da molécula com cada átomo posicionado na origem. Isso vai gerar todos os vetores entre todos os pares de átomos.



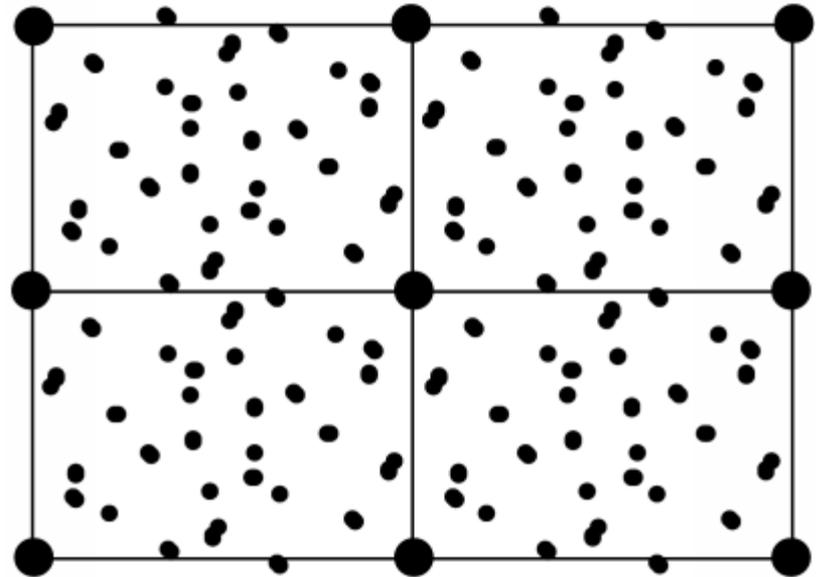
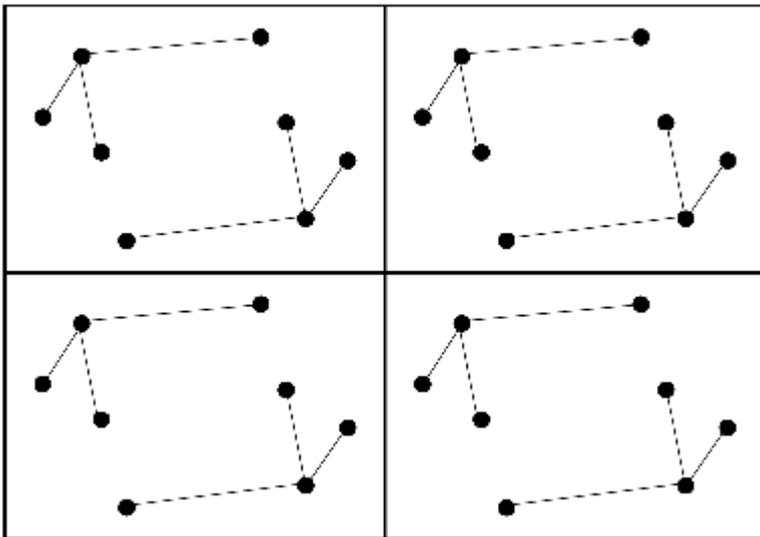
Aumentando o # de átomos (N), aumenta a complexidade do mapa de Patterson



A Função de Patterson e Substituição Molecular

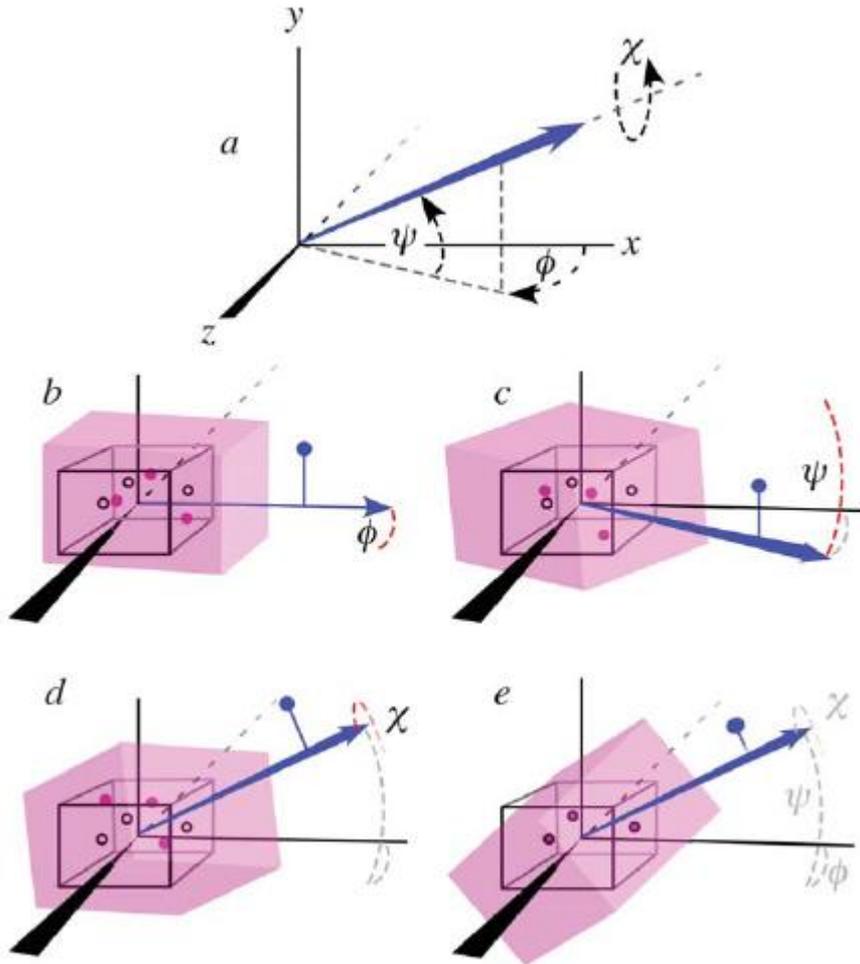
Apesar dos vetores no mapa de Patterson são de extrema difícil interpretação para uma estrutura do tamanho de uma proteína, eles podem ser usados como uma “assinatura digital” de uma proteína.

Vetores entre átomos da mesma molécula somente dependem da orientação da molécula dentro da célula unitária e não na sua posição.



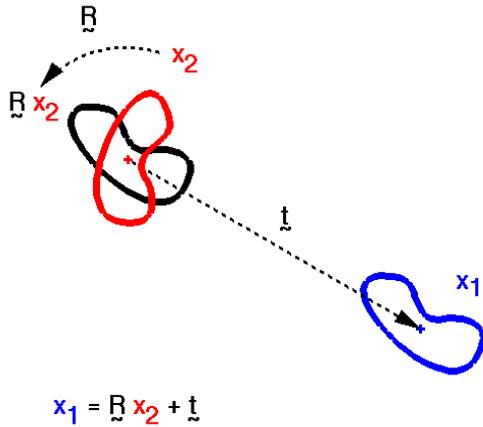
FUNÇÃO de ROTAÇÃO

$$R(\phi, \varphi, \chi) = \int_{u,v,w} P^{\text{target}}(u, v, w) P^{\text{model}}\{(u, v, w) \times [\phi, \varphi, \chi]\} du dv dw$$



- Integração do produto dos dois mapas de Patterson (target(exptal) e modelo) para cada **orientação do modelo (ϕ, φ, χ)**.
- Onde um mapa de Patterson tem um pico e o outro não tem, o produto $P^{\text{target}}P^{\text{model}}$ é zero.
- Onde os mapas de Patterson tem picos coincidentes, o produto $P^{\text{target}}P^{\text{model}}$ é grande.
- Logo um integral do produto $P^{\text{target}}P^{\text{model}}$ somado para todo o espaço (u,v,w) será muito grande se a sobreposição tem muitos picos coincidentes.
- **Procuramos (ϕ, φ, χ) que maximiza $R(\phi, \varphi, \chi)$.**

FUNÇÃO de TRANSLAÇÃO



Finalmente, para encontrar a **posição** correta da molécula, tentamos procurar a **função de translação** que produz a maior correlação entre as magnitudes (amplitudes) dos Fatores de estrutura observados experimentalmente (F_{obs}) e o fatores de estrutura calculados (F_{calc}) para o modelo em cada posição x, y, z na célula unitária.

Para isso, tentamos **minimizar o R-factor** que compara a somatório das diferenças entre F_{obs} e F_{calc}

$$R = \frac{\sum ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum |F_{\text{obs}}|}$$

$|F_{\text{calc}}|$ pode ser computado para cada reflexão hkl para o modelo em cada posição usando a equação:

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

Resumo do método de Substituição Molecular:

Importância das fases e das intensidades

Modelo de Homologia

Função e Mapa de Patterson

Função de Rotação

Função de Translação

Refinamento (ajustes) e validação do modelo

ESPAÇO REAL
(Coot)

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}| e^{-2\pi i (hx + ky + lz - \phi_{hkl})}$$

ESPAÇO RECIPROCO
(Phenix_refine ou Refmac5)

MAPA vs MODELO

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

fases

F_{obs} vs F_{calc}

1ª vez obtida usando fases estimadas por MR, MIR, MAD, etc
Subsequentemente baseado no modelo, minimizando o target function

$$2F_{obs} - F_{calc}$$

$$F_{obs} - F_{calc}$$

$$R_{free}, R_{factor} = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$$

(faz parte do Target function)

Mapa FWT = mapa 2Fo-Fc (azul)

Mapa DELFWT = mapa Fo-Fc (verde/vermelho)

- Geometria - ângulos,**
- ligações,
 - chiralidade,
 - planos
 - Ramachandran plot
 - outros

Target function

$$\Phi = \sum_{hkl} w_{hkl} (|F_o| - |F_c|)_{hkl}^2 \quad \text{Dados vs Modelo}$$

$$+ \sum_i^{\text{bonds}} w_i (d_i^{\text{ideal}} - d_i^{\text{model}})^2$$

Ideal vs Modelo

$$+ \sum_j w_j (\phi_j^{\text{ideal}} - \phi_j^{\text{model}})^2, \quad + \text{etc}$$

Sugestão de fluxograma para Refinamento e Validação do Modelo Cristalográfico

Ciclo 1

Abrir mapa e modelo no Coot e Adicionar todas as cadeias laterais.

Salvar modelo.

Rodar Phenix_refine no Phenix (ou Refmac no CCP4i)

Anotar Rwork, Rfree, RMSBond, RMSAngle, estatísticas do Ramachandran e outros

Ciclo 2

Abrir mapa e modelo no Coot e ajustar o modelo

Adicionar águas

Salvar o modelo

Rodar Phenix_refine no Phenix (ou Refmac no CCP4i)

Anotar Rwork, Rfree, RMSBond, RMSAngle, estatísticas do Ramachandran e outros

Ciclo 3

Abrir mapa e modelo no Coot e ajustar o modelo usando ferramentas de validação

Salvar o modelo

Rodar Phenix_refine no Phenix (ou Refmac no CCP4i)

Anotar Rwork, Rfree, RMSBond, RMSAngle, estatísticas do Ramachandran e outros

Ciclos 4, 5, 6 etc

Repetir ciclo 3

Até Rwork e Rfree não diminuem mais; gráfico Ramachandran não tem *outliers*,

Diminuir ou eliminar a maioria dos outros *outliers* de parâmetros geométricos

VALIDAÇÃO DO MODELO MOLECULAR

Critérios de Qualidade do Modelo
Um exemplo usando ProCheck

CAPÍTULO 8 de RHODES

topical reviews

Acta Crystallographica Section D
**Biological
Crystallography**

ISSN 0907-4449

Validation of protein crystal structures

Gerard J. Kleywegt

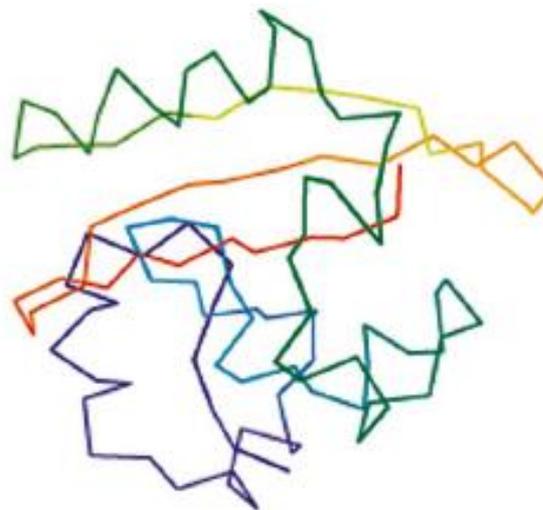
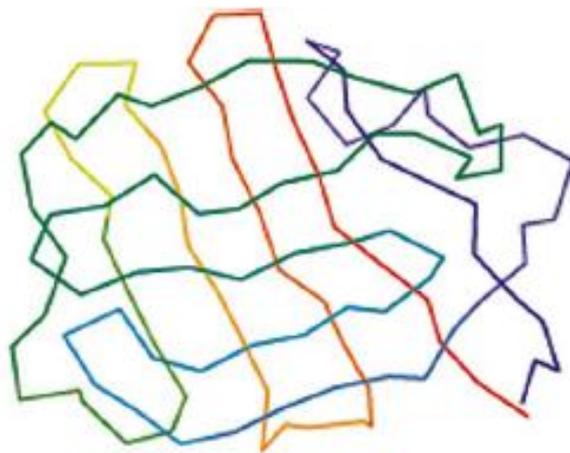
[Acta Crystallogr D Biol Crystallogr.](#)
2000 Mar; 56(Pt 3):249-65.

Department of Cell and Molecular Biology,
Uppsala University, Biomedical Centre,
Box 596, SE-751 24 Uppsala, Sweden

Incorrect structures

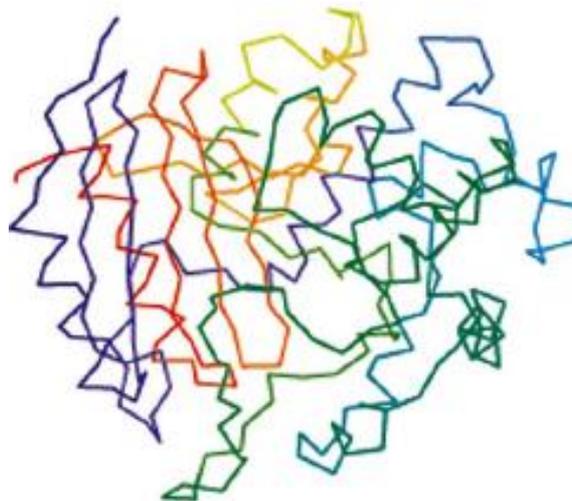
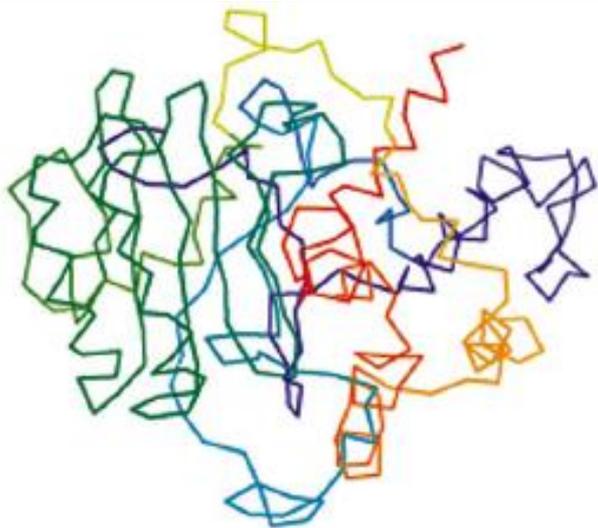
Correct structures

1phy



2phy
(photoactive
yellow
Protein)

1pte



3pte
(D-ala-D-ala
carboxypep-
tidase/trans-
Peptidase)

Cr terias de Qualidade

- R
- Completeza
- Redund ncia/Multiplicidade
- Sinal/ru do (I/σ)
- Resolu o
- Simetria
- Geometria e estereoqu mica
 - comprimento de liga o
 -  ngulo de liga o
 - quiralidade
-  ngulo de tor o
 - Diagrama de Ramachandran
- Contatos e ambientes
- Non-crystallographic symmetry (NCS)
- Mol culas de solvente
- Fatores de temperatura (B-factors)

Supplementary Table 1: X-ray data collection and refinement statistics for X-Tff^{XAC2610}.

	Native	Hg(II) acetate derivative	
Data collection			
Space group	C2	C2	C2
Cell dimensions			
<i>a, b, c</i> (�)	160.46, 37.70, 43.25	159.99, 37.51, 43.01	159.91, 37.52, 43.00
α, β, γ (�)	90, 92.61, 90	90, 92.76, 90	90, 92.75, 90
		<i>Peak</i>	<i>Inflection</i>
Wavelength (�)	1.54187	1.00798	1.00912
Resolution (�)	38.78 – 2.00 (2.07 – 2.00)	40.00 – 2.23 (2.31–2.23)	40.00 – 2.52 (2.70 – 2.5
R_{merge}	0.065 (0.497)	0.064 (0.230)	0.047 (0.108)
$I/\sigma I$	8.8 (2.1)	17.2 (4.7)	13.9 (5.3)
Completeness (%)	95.1 (89.9)	94.7 (78.9)	94.2 (79.9)
Redundancy	3.7 (3.7)	3.8 (3.4)	1.9 (1.8)
Refinement			
Resolution (�)	38.78 – 2.00 (2.07 – 2.00)		
No. reflections	16,929		
$R_{\text{work}}/ R_{\text{free}}$	0.200 / 0.241		
No. atoms	1,789		
Protein	1,608		
Ligand/ion	1		
Water	180		
<i>B</i> -factors (� ²)	51.7		
Protein	51.4		
Ligand/ion	41.1		
Water	54.2		
R.m.s deviations			
Bond lengths (�)	0.007		
Bond angles (�)	1.186		
MolProbity score	1.15 - 100 th percentile		
Ramachandran, favored / allowed / disallowed regions (%)	97.2 / 2.8 / 0.0		

Each data set was collected from a single crystal. Values in parentheses are for highest-resolution shell. R_{free} calculated from 5% subset of randomly selected reflections.

Biological Assembly 1 ?



3D View: [Structure](#) | [Electron Density](#) | [Ligand Interaction](#)

Standalone Viewers

[Protein Workshop](#) | [Ligand Explorer](#)

Global Symmetry: Asymmetric - C1 ⓘ

Global Stoichiometry: Monomer - A ⓘ

Biological assembly 1 assigned by authors and generated by PISA (software)

Macromolecule Content

- Total Structure Weight: 23357.08 ⓘ
- Atom Count: 1609 ⓘ
- Residue Count: 214 ⓘ
- Unique protein chains: 1

4QTQ

Structure of a Xanthomonas Type IV Secretion System related protein

DOI: [10.2210/pdb4QTQ/pdb](https://doi.org/10.2210/pdb4QTQ/pdb)

Classification: [HYDROLASE INHIBITOR](#)

Organism(s): [Xanthomonas axonopodis pv. citri \(strain 306\)](#)

Expression System: [Escherichia coli](#)

Deposited: 2014-07-08 Released: 2015-06-17

Deposition Author(s): [Souza, D.P.](#), [Guzzo, C.R.](#), [Farah, C.S.](#)

Experimental Data Snapshot

Method: X-RAY DIFFRACTION

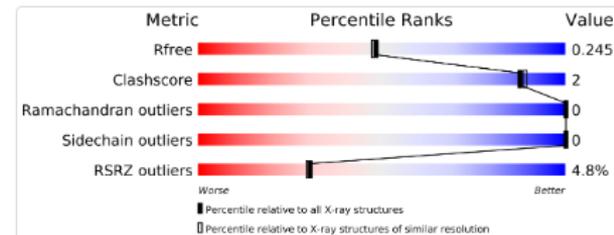
Resolution: 2 Å

R-Value Free: 0.241

R-Value Work: 0.200

wwPDB Validation

[3D Report](#) [Full Report](#)



This is version 1.0 of the entry. See complete [history](#).

Literature

[Download Primary Citation](#)

Bacterial killing via a type IV secretion system.

[Souza, D.P.](#), [Oka, G.U.](#), [Alvarez-Martinez, C.E.](#), [Bisson-Filho, A.W.](#), [Dunger, G.](#), [Hobeika, L.](#), [Cavalcante, N.S.](#), [Alegria, M.C.](#), [Barbosa, L.R.](#), [Salinas, R.K.](#), [Guzzo, C.R.](#), [Farah, C.S.](#)

(2015) Nat Commun **6**: 6453-6453

PubMed: [25743609](#) [Search on PubMed](#)

DOI: [10.1038/ncomms7453](https://doi.org/10.1038/ncomms7453)

PubMed Abstract:

Type IV secretion systems (T4SSs) are multiprotein complexes that transport effector proteins and protein-DNA complexes through bacterial membranes to the extracellular milieu or directly into the cytoplasm of other cells. Many bacteria of the family ...

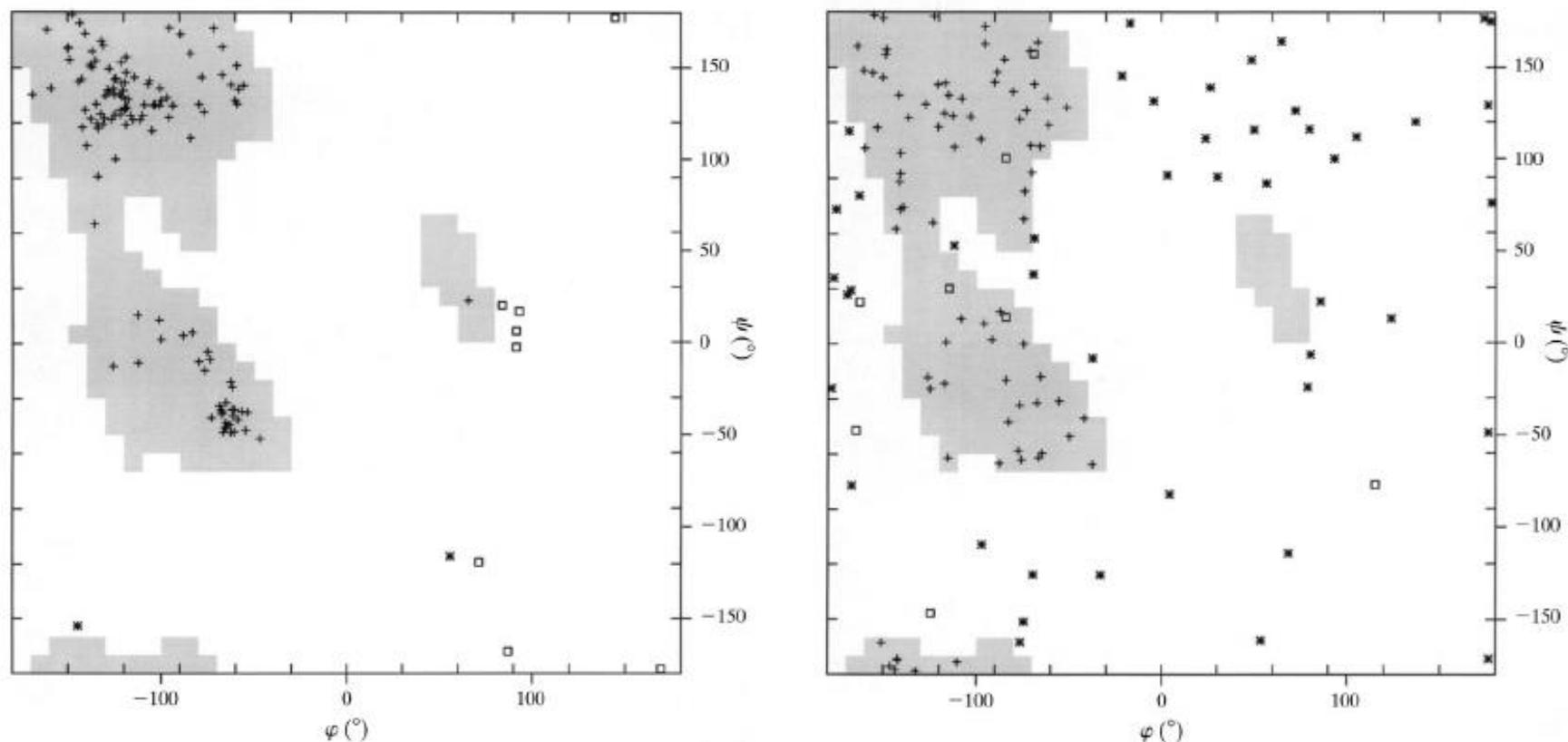
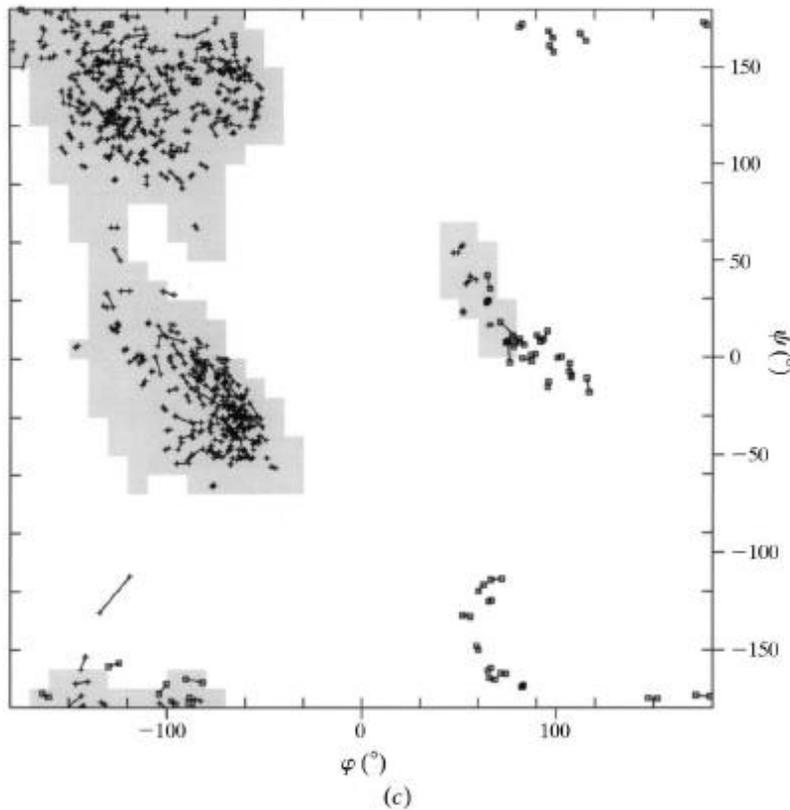


Figure 3

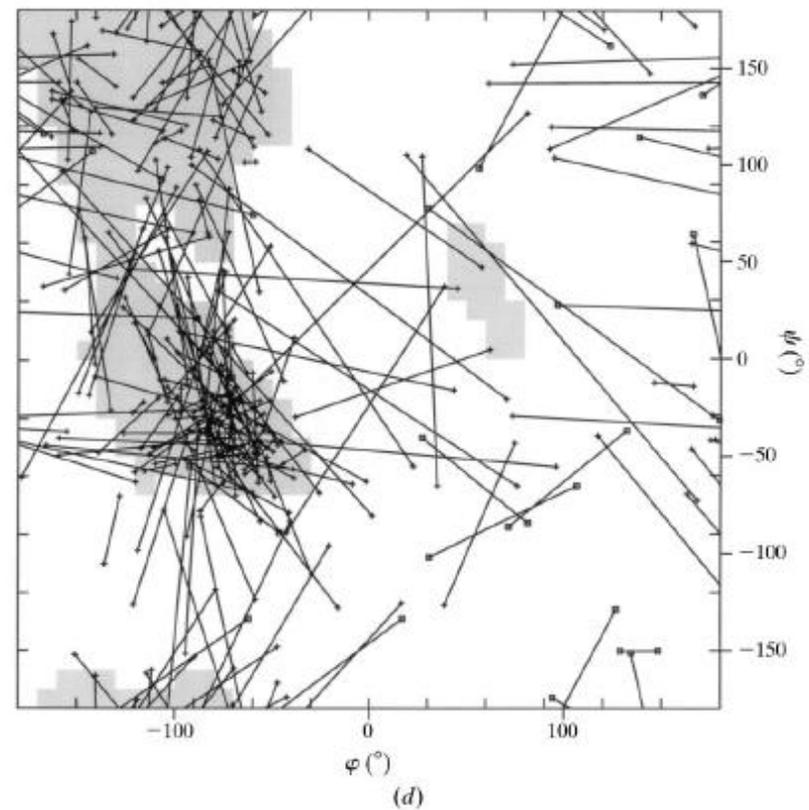
Example of good and poor Ramachandran plots. In general, a Ramachandran plot of a good model has both a tight clustering of residues in the most favoured regions of the plot and simultaneously a very small number of residues in unfavourable regions. (a) Ramachandran plot of cellular retinoic acid binding protein type II (Kleywegt *et al.*, 1994) refined to 1.8 Å resolution, which reveals only two outliers (1.6%). The shaded areas comprise the core regions of the Ramachandran plot as defined in Kleywegt & Jones (1996b). (b) Ramachandran plot of an intentionally backwards-traced model of the same protein, which was subsequently refined to 3.0 Å (Kleywegt & Jones, 1995b). No fewer than 46 residues (36.2%) are outliers in the plot and the remaining residues do not show a tight clustering in the core regions.

Comparando estruturas de múltiplas subunidades na unidade assimétrica (NCS)

Multiple model Ramachandran Plot



BOM



RUIM



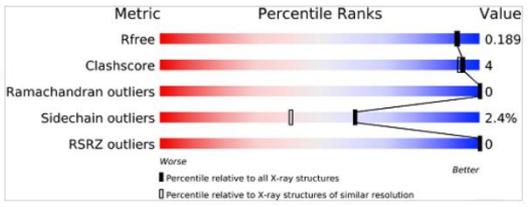
Developers: Join the RCSB PDB Team | Explore Open Positions

- Welcome
- Deposit
- Search
- Visualize
- Analyze
- Download
- Learn

- Deposit Options
 - Prepare Data
 - Validate Data
 - Deposit Data
 - Deposition Help
- Documentation
 - PDBx/mmCIF Dictionary Resources
 - Chemical Component Dictionary
 - Biologically Interesting Molecule Reference Dictionary (BIRD)
 - PDB Format Guide

Validation Services

Validation reports contain an assessment of the quality of a structure and highlight specific concerns by considering the coordinates of the model, the experimental data and the fit between the two. Easily interpretable summary information that compares the quality of a model with that of other models in the archive will help users of PDB data to critically assess archived entries and to select the most appropriate structural models for their needs. These reports are developed using the recommendations of the wwPDB Validation Task Forces.



Reports for released entries are available from Structure Summary pages.
Validation reports for manuscript reviewers are created during annotation of deposited structures.
Information and example Validation Reports (at wwpdb.org).

Check your X-ray, NMR, or EM structures before depositing (standalone server).

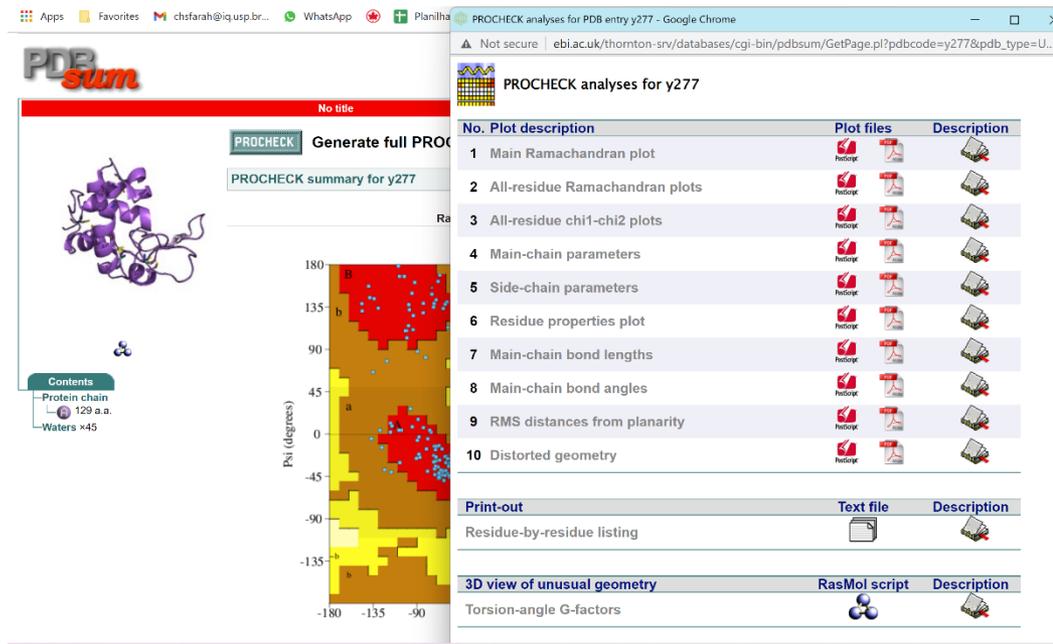
Latest Entries | As of Tue May 10 2022 | Features & Highlights | News | Publications

https://www.rcsb.org/#Subcategory-deposit_validate | New Ways to Explore Similar Proteins | Curated Files for 3D Printing

Validação

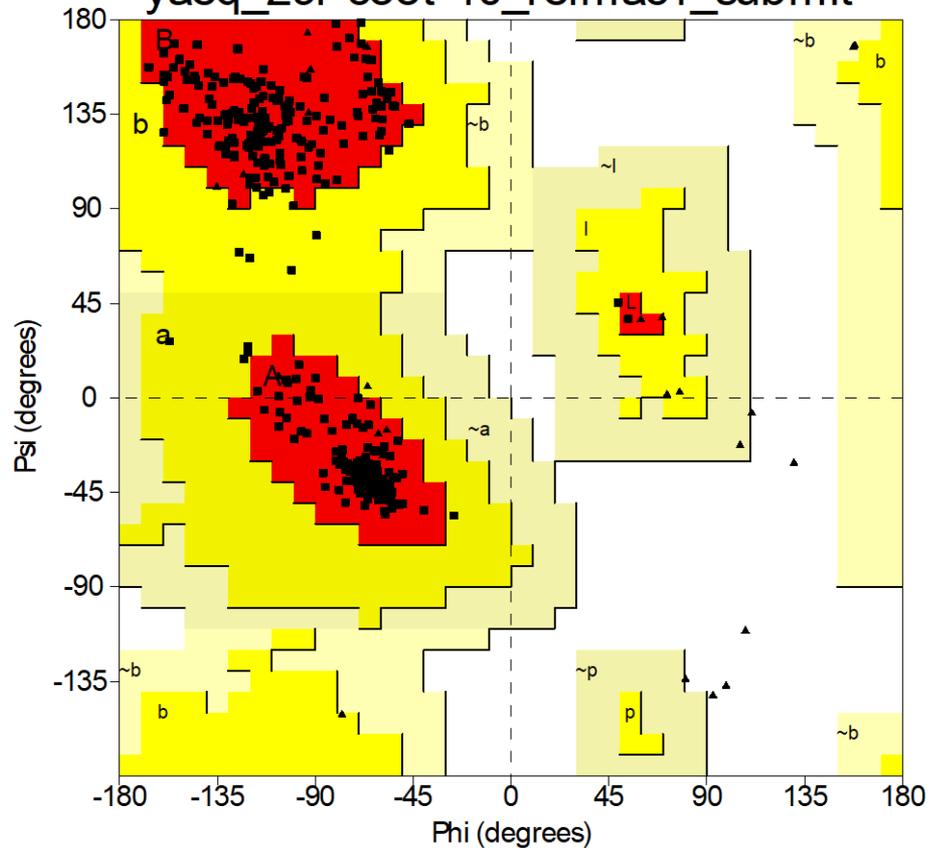
Um exemplo usando ProCheck dentro do pacote PDBsum

(<http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html>)



Ramachandran Plot

yaeq_29r-coot-10_refmac1_submit



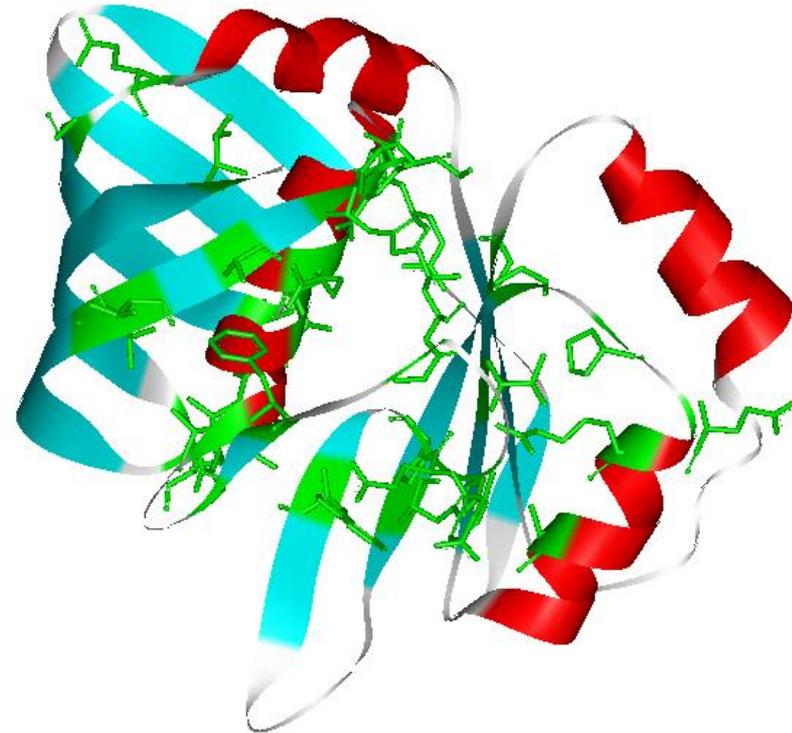
Plot statistics

Residues in most favoured regions [A,B,L]	297	95.2%
Residues in additional allowed regions [a,b,l,p]	15	4.8%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%

Number of non-glycine and non-proline residues	312	100.0%
Number of end-residues (excl. Gly and Pro)	226	
Number of glycine residues (shown as triangles)	26	
Number of proline residues	12	

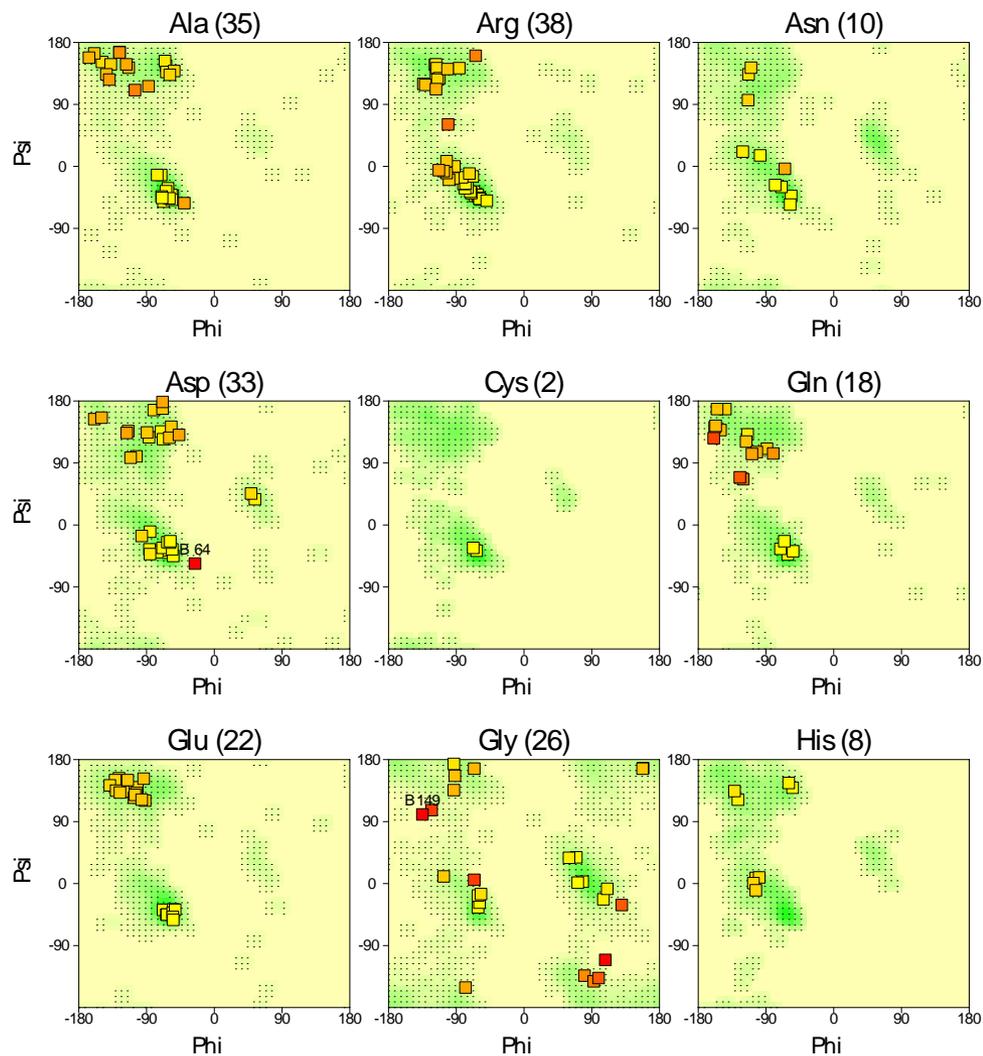
Total number of residues	576	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.



Ramachandran plots for all residue types

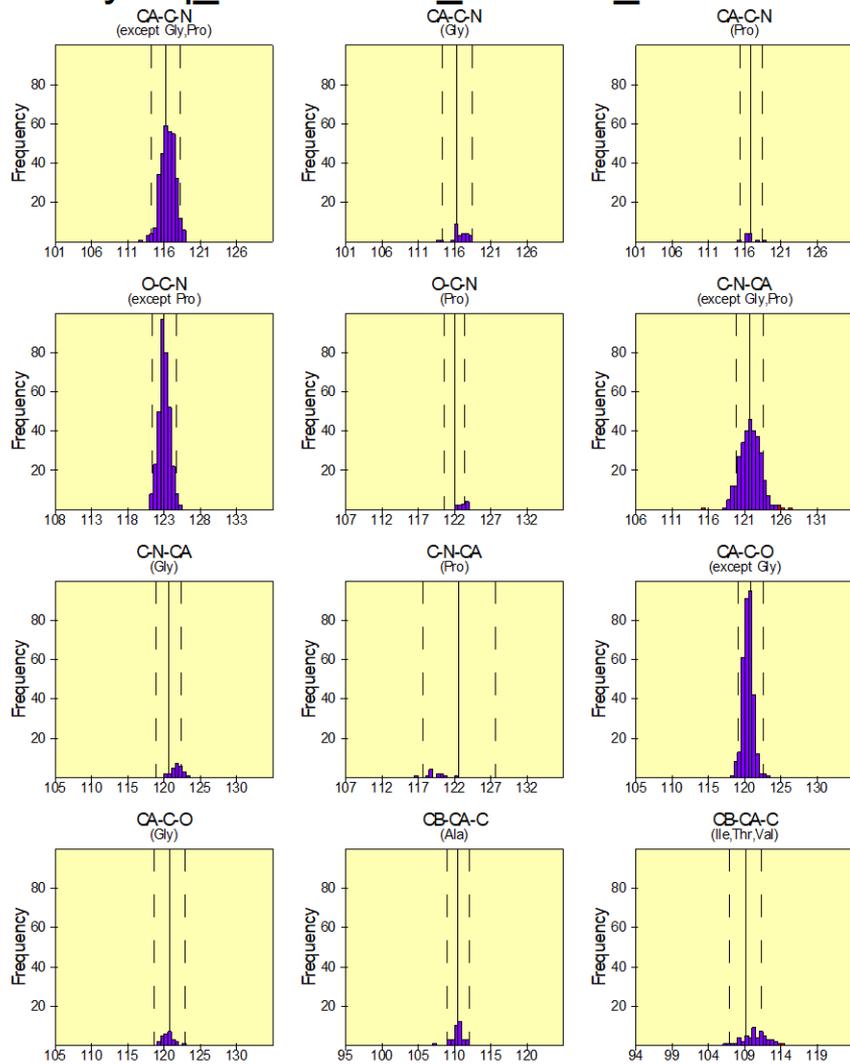
yaeq_29r-coot-10_refmac1_submit



Numbers of residues are shown in brackets. Those in unfavourable conformations (score < -3.00) are labelled. Shading shows favourable conformations as obtained from an analysis of 163 structures at resolution 2.0A or better.

Main-chain bond angles

yaeq_29r-coot-10_refmac1_submit

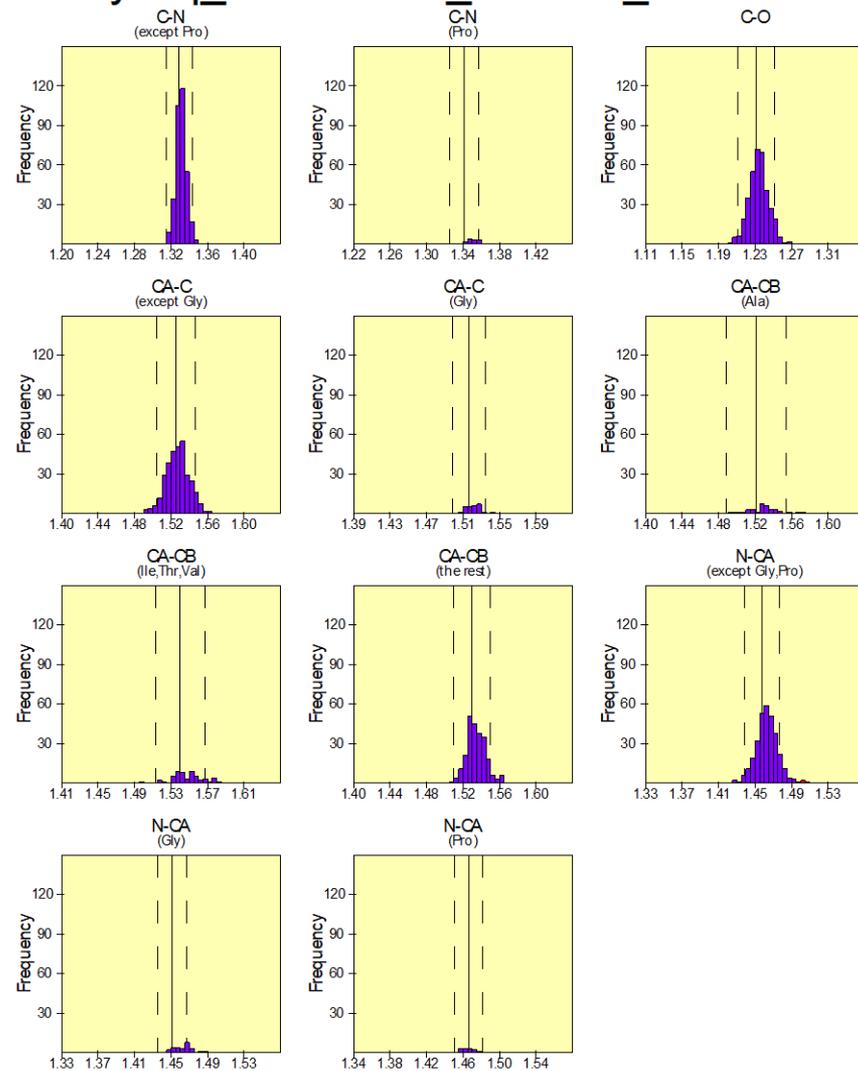


Black bars > 2.0 st. devs. from mean.

Solid and dashed lines represent the mean and standard deviation values as per Engh & Huber small-molecule data.

Main-chain bond lengths

yaeq_29r-coot-10_refmac1_submit

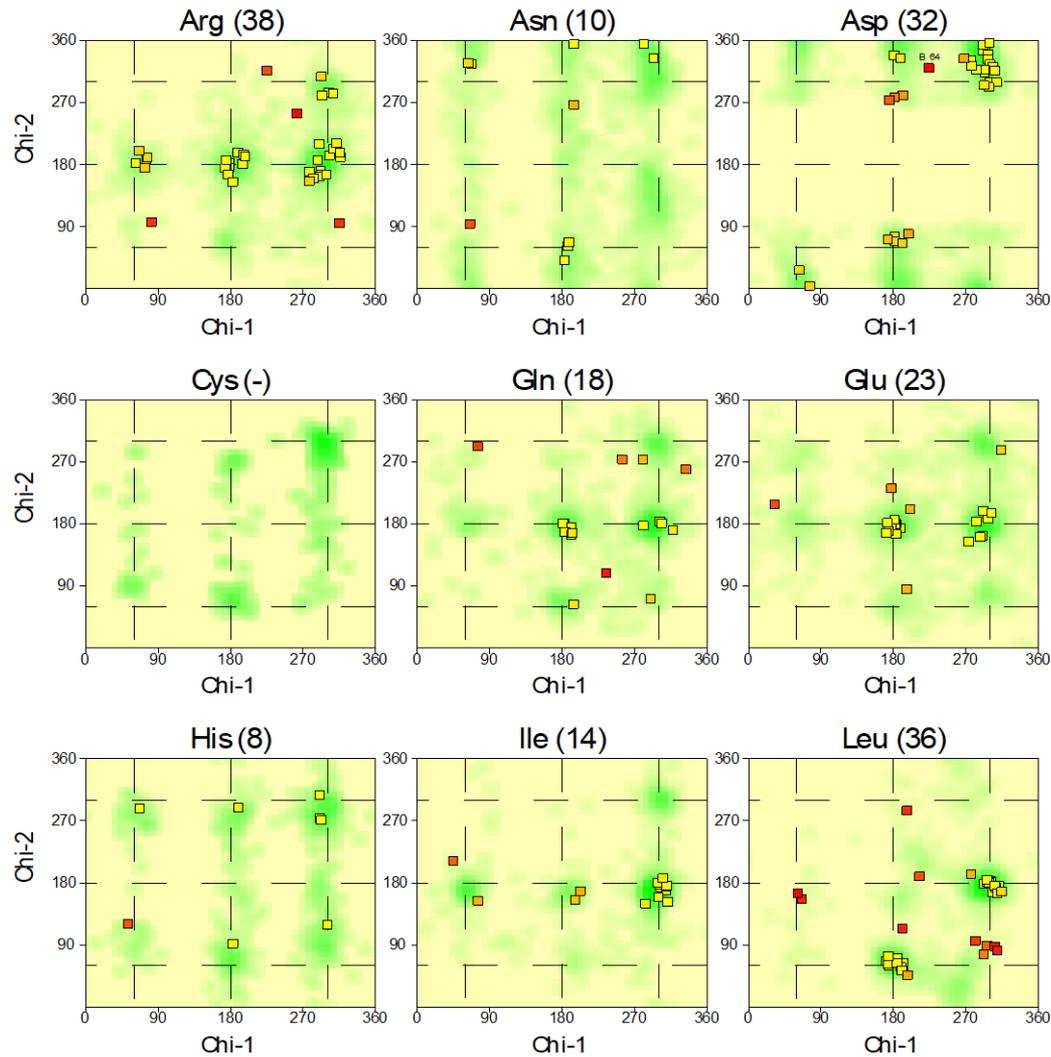


Black bars > 2.0 st. devs. from mean.

Solid and dashed lines represent the mean and standard deviation values as per Engh & Huber small-molecule data.

Chi1-Chi2 plots

yaeq_29r-coot-10_refmac1_submit

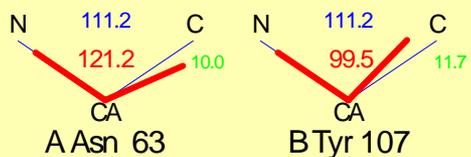


Numbers of residues are shown in brackets. Those in unfavourable conformations (score < -3.00) are labelled. Shading shows favourable conformations as obtained from an analysis of 163 structures at resolution 2.0Å or better.

Distorted geometry

yaeq_29r-coot-10_refmac1_submit

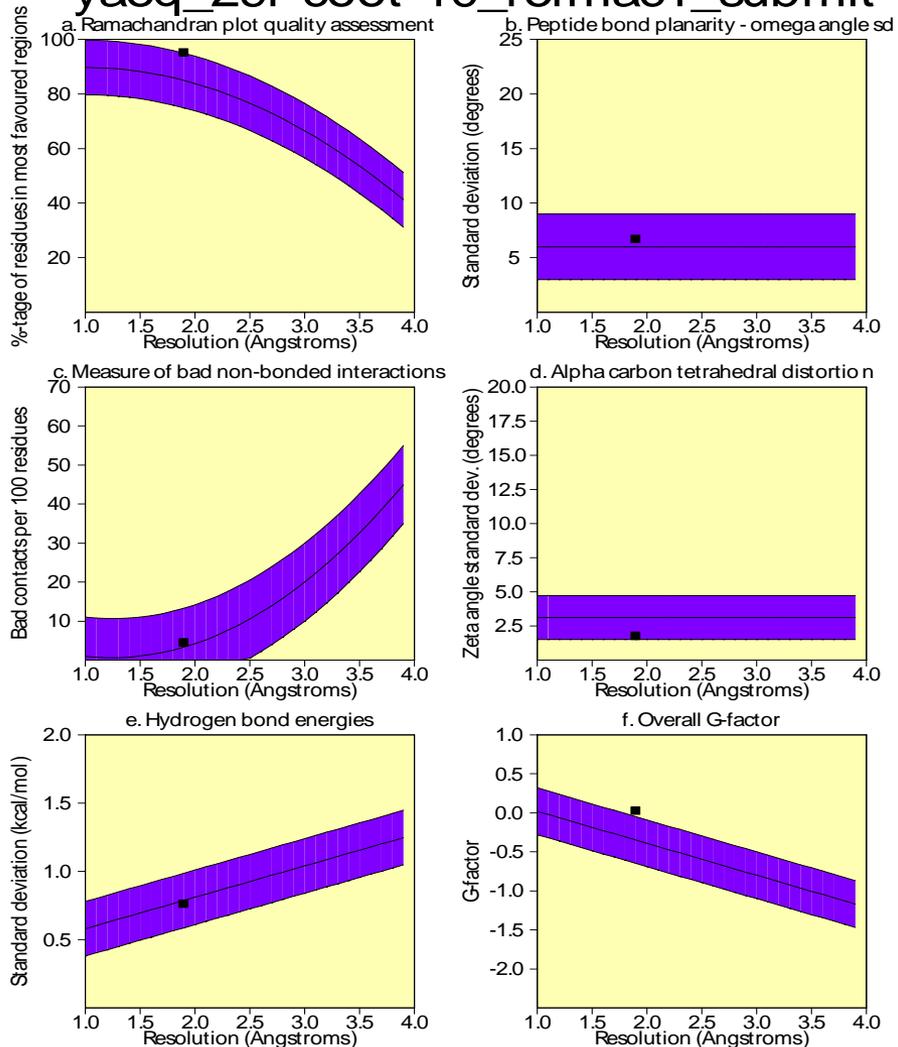
Main-chain bond angles



Bond angles differing by > 10.0 degrees from small-molec values. Values shown: "ideal", actual, diff.

Main-chain parameters

yaeq_29r-coot-10_refmac1_submit



Plot statistics

Stereochemical parameter	No. of data pts	Parameter value	Comparison values		No. of band widths from mean	
			Typical value	Band width		
a. %age residues in A, B, L	312	95.2	85.0	10.0	1.0	BETTER
b. Omega angle st dev	350	6.7	6.0	3.0	0.2	Inside
c. Bad contacts / 100 residues	26	4.5	3.3	10.0	0.1	Inside
d. Zeta angle st dev	328	1.8	3.1	1.6	-0.8	Inside
e. H-bond energy st dev	233	0.8	0.8	0.2	-0.1	Inside
f. Overall G-factor	576	0.0	-0.3	0.3	1.2	BETTER

Appendix E - G-factors

The **G-factor** provides a measure of how "normal", or alternatively how "unusual", a given stereochemical property is.

In **PROCHECK** it is computed for the following properties:-

Torsion angles:-

- **phi-psi** combination
- **chi1-chi2** combination
- **chi1** torsion for those residues that do not have a **chi-2**
- combined **chi-3** and **chi-4** torsion angles
- **omega** torsion angles

Covalent geometry:-

- main-chain **bond lengths**
- main-chain **bond angles**

The **G-factor** is essentially just a **log-odds score** based on the observed distributions of these stereochemical parameters.

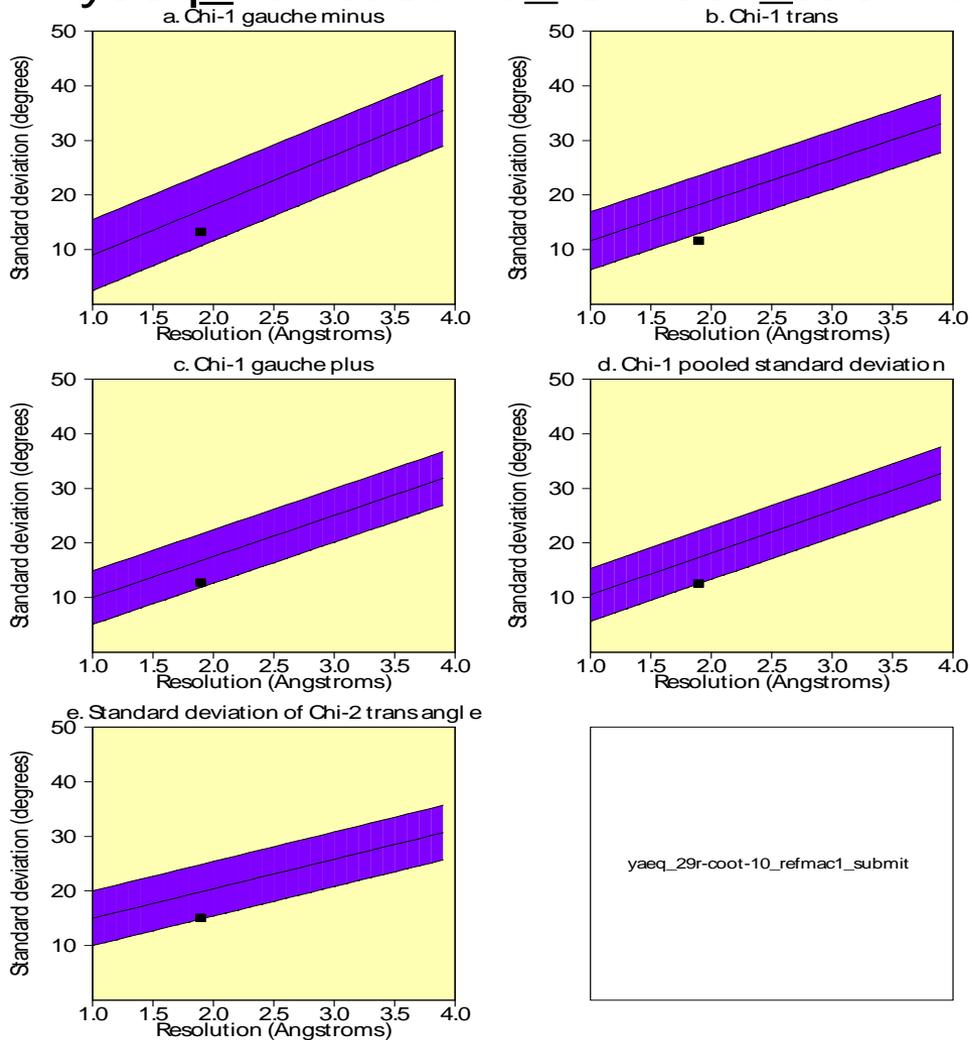
When applied to a given residue, a **low G-factor** indicates that the property corresponds to a low-probability conformation. So, for example, residues falling in the disallowed regions of the **Ramachandran plot** will have a **low** (or very negative) **G-factor**.

Similarly for unfavourable **chi1-chi2** and **chi1** values.

Thus, if a protein has many residues with low **G-factors** it suggests that something may be amiss with its overall geometry.

Side-chain parameters

yaeq_29r-coot-10_refmac1_submit



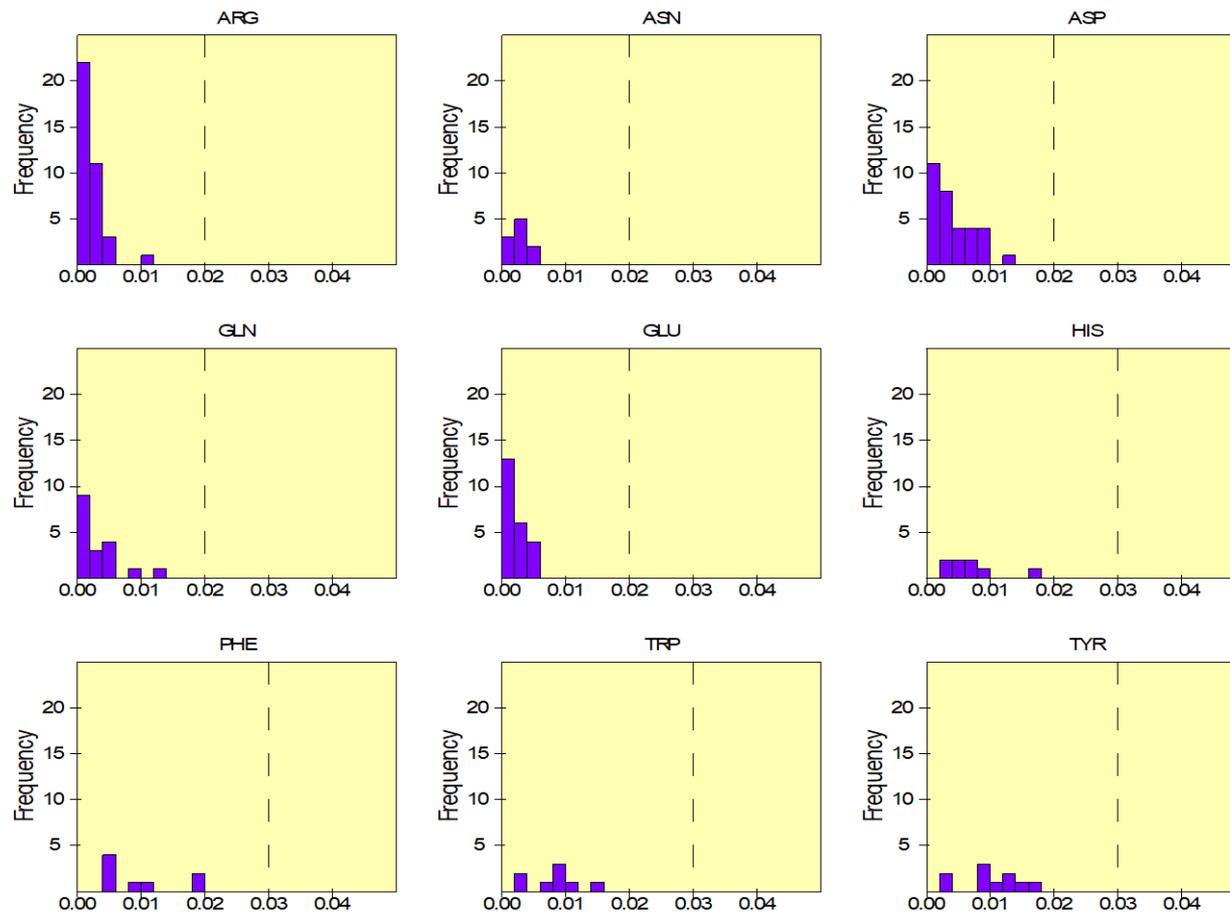
yaeq_29r-coot-10_refmac1_submit

Plot statistics

Stereochemical parameter	No. of data pts	Parameter value	Comparison values		No. of band widths from mean	
			Typical value	Band width		
a. Chi-1 gauche minus st dev	37	13.3	17.2	6.5	-0.6	Inside
b. Chi-1 trans st dev	105	11.6	18.2	5.3	-1.3	BETTER
c. Chi-1 gauche plus st dev	125	12.7	16.7	4.9	-0.8	Inside
d. Chi-1 pooled st dev	267	12.6	17.4	4.8	-1.0	Inside
e. Chi-2 trans st dev	98	15.1	19.8	5.0	-1.0	Inside

RMS distances from planarity

yaeq_29r-coot-10_refmac1_submit

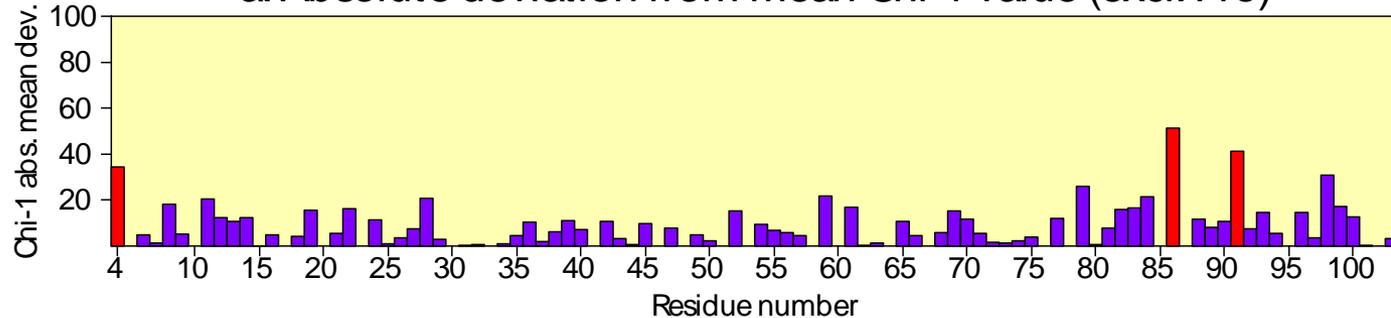


Histograms showing RMS distances of planar atoms from best-fit plane.
Black bars indicate large deviations from planarity: RMSdist > 0.03 for rings, and > 0.02 otherwise .

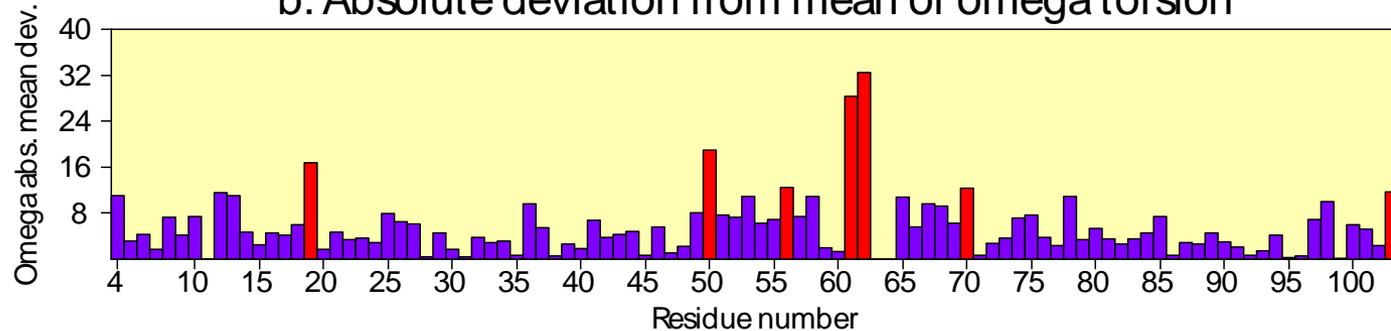
Residue properties

yaeq_29r-coot-10_refmac1_submit

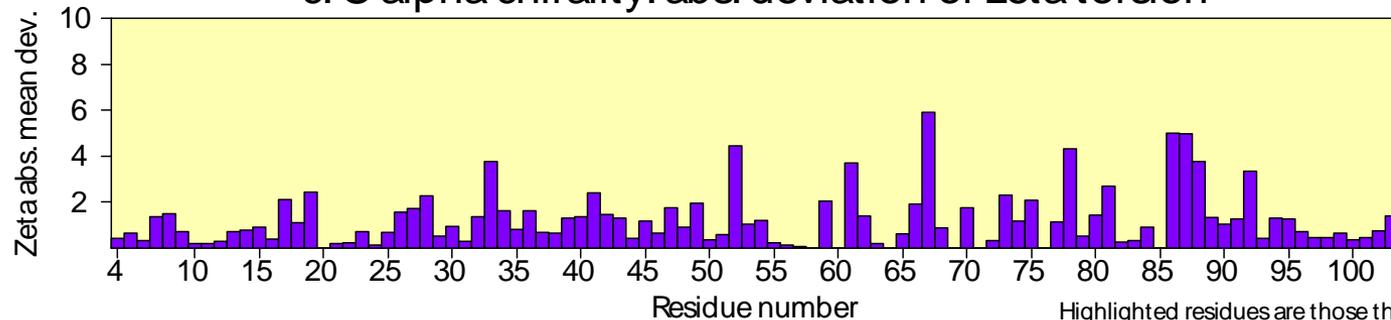
a. Absolute deviation from mean Chi-1 value (excl. Pro)



b. Absolute deviation from mean of omega torsion

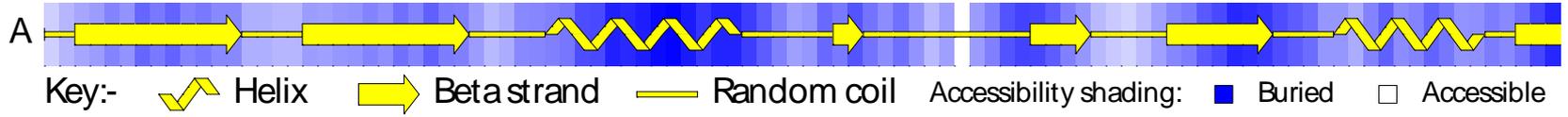


c. C-alpha chirality: abs. deviation of zeta torsion



Highlighted residues are those that deviate by more than 2.0 st. devs. from ideal

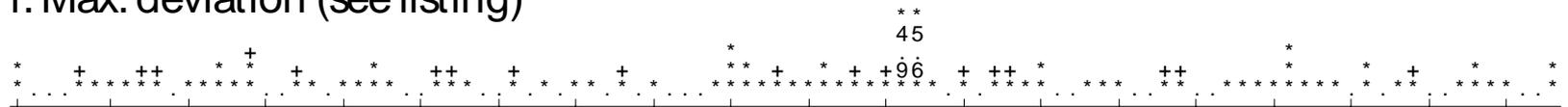
d. Secondary structure & estimated accessibility



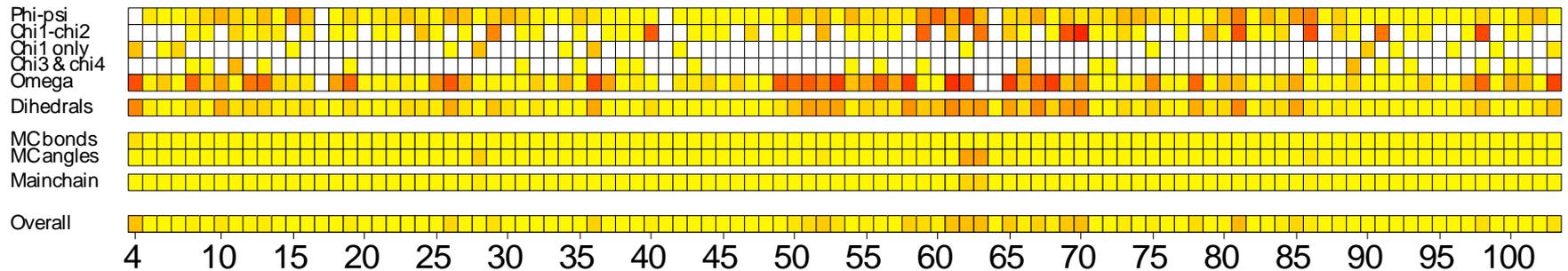
e. Sequence & Ramachandran regions



f. Max. deviation (see listing)



g. G-factors



Refinamento e Validação do modelo

ESPAÇO REAL

(Coot)

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}| e^{-2\pi i (hx + ky + lz - \phi_{hkl})}$$

ESPAÇO RECIPROCO

(Phenix_refine ou Refmac5)

MAPA vs MODELO

fases

F_{obs} vs F_{calc}

1ª vez obtida usando fases estimadas por MR, MIR, MAD, etc
Subsequentemente baseado no modelo, minimizando o target function

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

$2F_{obs} - F_{calc}$

$F_{obs} - F_{calc}$

$$R_{free}, R_{factor} = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$$

(faz parte do Target function)

Mapa FWT = mapa 2Fo-Fc (azul)

Mapa DELFWT = mapa Fo-Fc (verde/vermelho)

- Geometria - ângulos,**
- ligações,
 - chiralidade,
 - planos
 - Ramachandran plot
 - outros

Target function

$$\Phi = \sum_{hkl} w_{hkl} (|F_o| - |F_c|)_{hkl}^2 \quad \text{Dados vs Modelo}$$

$$+ \sum_i^{\text{bonds}} w_i (d_i^{\text{ideal}} - d_i^{\text{model}})^2$$

Ideal vs Modelo

$$+ \sum_j w_j (\phi_j^{\text{ideal}} - \phi_j^{\text{model}})^2, \quad + \text{etc}$$

Projeto para entregar até

 Determinação da Estrutura da Lisozima por Cristalografia_Phenix

 lysozyme_intensities

 lysozyme_intensities.ref

 modelo_polyA

REPRESENTANDO MOLÉCULAS RELACIONADA POR SIMETRIA CRISTALOGRÁFICA

WinCoot 0.8.6.1

File Edit Calculate Draw Measures Validate HID About Ligand Extensions

Reset View

- Go To Atom... F6
- Display Manager... F7
- Undo Last Navigation U
- Centre Atom Label...
- Sequence View
- Anisotropic Atoms...
- Cell & Symmetry...**
- Additional Representation...
- NCS Ghost Control...
- Spin View On/Off
- Rock View On/Off
- Screenshot
- Generic Display Objects...
- Stereo...
- Clipping...
- Crosshairs...
- Zoom...

R/R/C
Map

CB 7367 7HEVA

(mol. no: 3) CB 1/A/382 PHE occ: 1.00 bf: 23.38 ele: C pos: (-25.69,-50.57,-13.75)

Show Symmet...

Symmetry & Unit Cell

Symmetry

Master Switch: Show Symmetry Atoms?

Yes

No

Symmetry by Molecule...

Symmetry Atom Display Radius:

Radius: 13.0 A

Symmetry Colour

Sym Colour

Colour Merge: 0.5

Expanded Symmetry Atom Labels?

Show Unit Cells?

Yes

No

Apply OK

Symmetry Controller

Display as CAs Colour by Molecule

Molecule 6 ZikaEnvelope_2.1.pdb

Show Symmetry?

Display Options

Display Sphere Standard Colouring

Display Near Chains Colour by Symop

Display as CAs Colour by Molecule

OK

VALIDAÇÃO DENTRO DE COOT

The image shows the WinCoot 0.8.6.1 software interface. The main window displays a 3D molecular model with a yellow mesh representing the electron density map and a stick representation of the protein backbone in green and blue. A menu is open over the 'Validate' option in the top menu bar. The menu items are:

- Ramachandran Plot
- Kleywegt Plot...
- Incorrect Chiral Volumes...
- Unmodelled blobs...
- Difference Map Peaks...
- Check/Delete Waters...
- Geometry analysis
- Peptide omega analysis
- Temp. fact. variance analysis
- Average Temp. fact. analysis
- GLN and ASN B-factor Outliers
- Rotamer analysis
- Density fit analysis
- Probe clashes
- NCS Differences
- Highly coordinated waters...
- Pukka Puckers...?
- Alignment vs PIR...

The status bar at the bottom of the window displays the following information: (mol. no: 6) CA /1/A/43 ILE occ: 1.00 bf: 23.97 ele: C pos: (-12.78, -63.93, 5.77)

MUDANDO MÚLTIPLOS RESIDUOS NO COOT

The image shows the WinCoot 0.8.6.1 software interface. The main window displays a 3D molecular model with a yellow mesh and colored sticks (green, blue, pink). A 'Mutate sequence' dialog box is open in the center, allowing the user to specify the molecule to mutate, the chain ID, and the residue range. The dialog box contains the following fields and options:

- Mutate Molecule: 2 ...pe/ZikaEnvelope_2.1-coot-2.pdb
- Chain ID: A
- Residue numbers: [] to [] (inclusive)
- to (single letter) sequence: []
- Autofit the mutated residues?
- Buttons: Mutate (green checkmark), Cancel (red X)

The 'Calculate' menu is open, showing options such as Model/Fit/Refine..., Other Modelling Tools..., SSM Superpose..., LSQ Superpose..., Merge Molecules..., Mutate Residue Range..., Align & Mutate..., Renumber Residues..., Change Chain IDs..., Move Molecule Here..., Fit Loop, Map Sharpening..., Map Skeleton..., NCS Maps..., Frames/Sec..., Scripting..., Run Script..., and Ligand Builder.

At the bottom of the window, the status bar displays: (mol. no: 6) CA /1/A/43 ILE occ: 1.00 bf: 23.97 ele: C pos: (-12.78,-63.93, 5.77)

ENCONTRANDO MOLÉCULAS DE ÁGUA

WinCoot 0.8.6.1

File Edit **Calculate** Draw Measures Validate HID About Ligand Extensions

Model/Fit/Refine... F5

- Other Modelling Tools...
- SSM Superpose...
- LSQ Superpose...
- Merge Molecules...
- Mutate Residue Range...
- Align & Mutate...
- Renumber Residues...
- Change Chain IDs...
- Move Molecule Here...
- Fit Loop
- Map Sharpening...
- Map Skeleton...
- NCS Maps...
- Frames/Sec...
- Scripting...
- Run Script...
- Ligand Builder

Other Modelling Tools

- Find Waters...
- Find Ligands...
- Find Secondary Structure...
- Build Nucleic Acid...
- Cis <-> Trans
- C-alpha Baton Mode...
- Ca Zone -> Mainchain
- Add OXT to Residue...
- Multi-Residue Torsion
- Reverse Direction...
- Place Helix Here
- Place Strand Here
- Ideal DNA/RNA...
- Base Pair...
- Choose "Undo" Molecule...

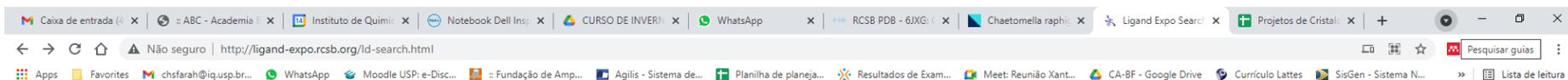
Close

(mol. no: 6) CA /1/A/43 ILE occ: 1.00 bf: 23.97 ele: C pos: (-12.78,-63.93, 5.77)

The image shows the WinCoot software interface. The main window displays a 3D molecular model with a yellow mesh surface and colored sticks (green, blue, pink) representing atoms. The 'Calculate' menu is open, showing various modeling tools. The 'Other Modelling Tools' dialog box is also open, listing specific actions like 'Find Waters...', 'Find Ligands...', and 'Find Secondary Structure...'. The status bar at the bottom provides coordinates for a specific atom: (mol. no: 6) CA /1/A/43 ILE occ: 1.00 bf: 23.97 ele: C pos: (-12.78,-63.93, 5.77).

Para encontrar o código de 3 letras para moléculas pequenas

<http://ligand-expo.rcsb.org/ld-search.html>



Ligand Expo

Home Search Browse Download Ligand Expo Help

Chemical Component Search Tools

Use the forms below to search for chemical components within the PDB Component Dictionary.

- Search for chemical components by *3-letter component identifier code, molecular name, molecular formula, SMILES description, or InChi/InChiKey chemical description*. You can also check to see if a 3-letter code is being held by a deposition in progress. Either start with a SMILES description or chemical data file (see drop-menu for acceptable formats), or draw a 2D chemical structure from scratch (**Launch** without input). It can also generate chemical component definitions from your 2D structure.
- Search for *instances* of a chemical component throughout the PDB. The **Display** option allows you to simply see a list of PDB codes, or to download these coordinates in PDB, MOL/SDF and mmCIF formats.
- You can also search for analogs to the standard amino acids, nucleotides, popular drugs, and common aromatic ring systems by using the *Browse* feature in the top menu bar.

Your query results are also searchable! Each hit from your initial query will contain links to continue searching by similar name, chemical formula, or structure (SMILES).

MOLECULAR NAME, FORMULA, AND DESCRIPTOR SEARCH OPTIONS ?

Search term Search type

Go

SEARCH FOR INSTANCES OF CHEMICAL COMPONENTS BY 3-LETTER ID CODE ?

Component ID code Display

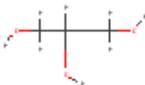
Go

Ligand Expo Search Result Summary

Query: glycerol

Query type: Molecular name (exact)

Result count: 2

	ID	View Options	Description
	CRY	Chemical details Coordinates files	Name: propane-1,2,3-triol Synonyms: RCSB OBS glycerol propane-1,2,3-triol propane-1,2,3-triol SMILES: C(C(CO)O)O Formula: C3 H8 O3
	GOL	Chemical details Coordinates files	Name: glycerol Synonyms: EBI REL glycerin; propane-1,2,3-triol glycerin; propane-1,2,3-triol propane-1,2,3-triol propane-1,2,3-triol SMILES: C(C(CO)O)O Formula: C3 H8 O3

