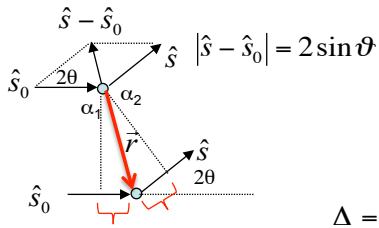


Il segno della trasformata di Fourier



$$\Delta = r(\sin \alpha_1 + \sin \alpha_2) = -\vec{r} \cdot (\hat{s} - \hat{s}_0)$$

$$\frac{2\pi}{\lambda} \Delta = -\frac{2\pi}{\lambda} \vec{r} \cdot (\hat{s} - \hat{s}_0) = -2\pi \vec{r} \cdot \vec{S}$$

$$\vec{r} \cdot \vec{S} = (hx + ky + lz)$$

$$F(h, k, l) = \int_{cell} \rho(x, y, z) \exp[-2\pi i(hx + ky + lz)] dV$$

1

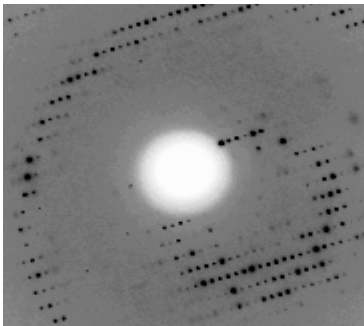
Il problema della fase

misura sperimentale

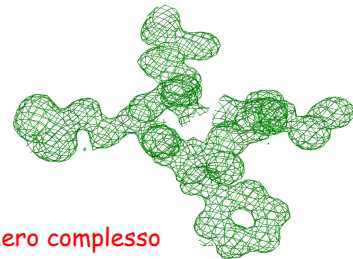
$$I \propto |F(h, k, l)|^2$$

$$F(h, k, l) = \int_{cell} \rho(x, y, z) \exp[-2\pi i(hx + ky + lz)] dV \longleftrightarrow \rho(x, y, z) = \frac{1}{V} \sum_{h, k, l} F(h, k, l) \exp[2\pi i(hx + ky + lz)]$$

F.T.



$$I(h, k, l) \propto |F(h, k, l)|^2$$



Il fattore di struttura e' un numero complesso

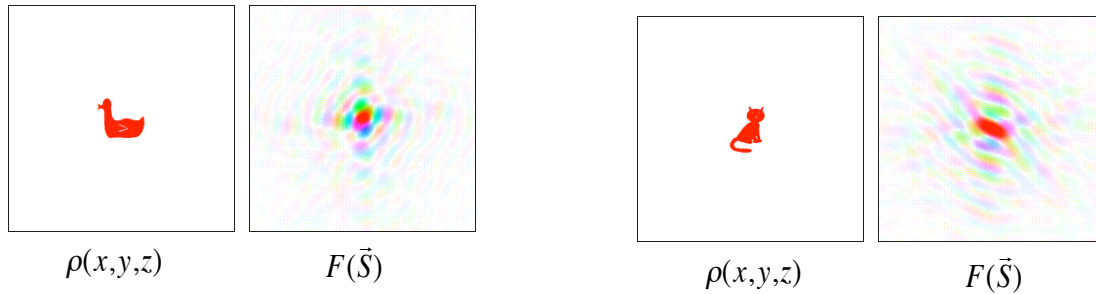
$$F = F' + iF'' = |F(h, k, l)| e^{i\alpha(h, k, l)}$$

$$\rho(x, y, z) = \frac{1}{V} \sum_{h, k, l} |F(h, k, l)| e^{i\alpha(h, k, l)} \exp[2\pi i(hx + ky + lz)]$$

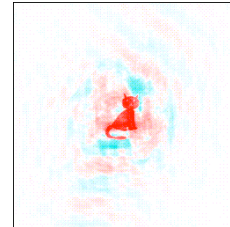


La fase non e' misurabile direttamente

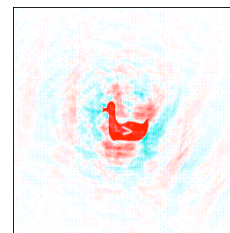
2



$$\rho(x,y,z) = \frac{1}{V} \sum_{h,k,l} |F(h,k,l)|_{duck} e^{i\alpha(h,k,l)_{cat}} \exp[-2\pi i(hx + ky + lz)]$$



$$\rho(x,y,z) = \frac{1}{V} \sum_{h,k,l} |F(h,k,l)|_{duck} e^{i\alpha(h,k,l)_{cat}} \exp[-2\pi i(hx + ky + lz)]$$



the phases contain the bulk of the structural information.

3

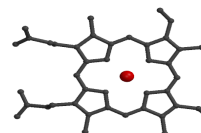
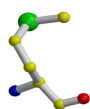
Soluzione del problema della fase

- Molecular Replacement (MR)
- Atomi pesanti: Se, Fe, Hg, Pt, Sm, ...
- Metodi diretti (...)

MIR
Multiple Isomorphous Replacement (IR)

S/MIRAS
Single/Multiple IR with Anomalous Signal

MAD
Multiple Anomalous Dispersion



4

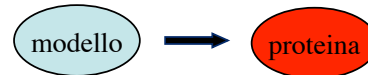
Molecular Replacement

Prerequisito:

modello atomico da cui possono essere calcolate delle **fasi approssimate** per la **proteina** di cui si vuole risolvere la struttura

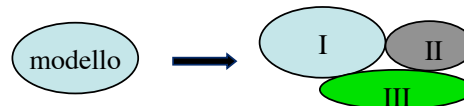
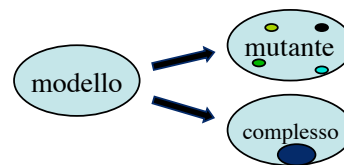
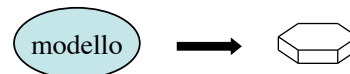


proteina omologa



l'omologia strutturale e' correlata col livello di identità in sequenza (MR se > **25-30% identità di amminoacidi**)

- stessa proteina in una diversa forma cristallina
- stessa proteina in forma 'nativa' (studi di mutanti o complessi)
- parte di una proteina multidominio



5

Molecular Replacement

- Ricerca di un modello adeguato: allineamento di sequenza contro il PDB, **editare il modello**

proteina	... YKTQAGKTVVDYINAAIGG----SADGAGL ...
modello(i)	... YQTQASKTVVDYITAA L AGSR N VSADAAGL ...
	* *

- **orientare** il modello nella cella elementare del cristallo così come è orientata la proteina (**ricerca di 3 parametri rotazionali**). Si utilizza la **funzione di Patterson**

$$P(\mathbf{u}) = P(u,v,w) \equiv \frac{1}{V} \sum_{\mathbf{h}} |F_{\mathbf{h}}|^2 \exp(-2\pi i \mathbf{h} \cdot \mathbf{u}) = \int_{\text{cella}} \rho(\mathbf{r}) \rho(\mathbf{r} + \mathbf{u}) d^3 r$$

- **posizionare** il modello correttamente orientato nella posizione in cui si trova la proteina nella cella elementare del cristallo (**ricerca di 3 parametri traslazionali**)

$$R = \frac{\sum_{\mathbf{h}} \left| |F_{\mathbf{h}}(\text{obs})| - k |F_{\mathbf{h}}(\text{calc})| \right|}{\sum_{\mathbf{h}} |F_{\mathbf{h}}(\text{obs})|}$$

6

Quante sono le molecole nell'unità asimmetrica?

Numero di Matthews (densità): $V_M \equiv \frac{V_{cella}}{Z_{cella} \times Mw} \in \left[1.7, 3.5 \right] \frac{\text{\AA}^3}{Da} \longrightarrow \bar{V}_M \approx 2.15 \frac{\text{\AA}^3}{Da}$

Esempio: gruppo spaziale C2 \longrightarrow 4 unità asimmetriche nella cella elementare $\longrightarrow Z = 4z$

$V = 319 \cdot 10^3 \text{\AA}^3$
 $Mw = 32.1 \cdot 10^3 Da$

$\longrightarrow V_M = \begin{cases} z = 1, & 2.48 \frac{\text{\AA}^3}{Da} \\ z = 2, & 1.24 \frac{\text{\AA}^3}{Da} \end{cases}$

$z = \text{numero di proteine nell'unità asimmetrica (incognita)}$

% di solvente nella cella elementare ?

Volume specifico medio di una proteina $\rho = 0.74 \frac{cm^3}{gr} = 0.74 \frac{10^{24}}{6 \cdot 10^{23}} = 1.23 \frac{\text{\AA}^3}{Da}$

$$\begin{cases} 1 \text{ cm} = 10^8 \text{\AA} \\ 1 \text{ gr} = 1 Da \times N_A = 1 Da \times 6 \cdot 10^{23} \end{cases}$$

volume della cella elementare

\downarrow
 $V_C = V_P + V_S = MwZ\rho + V_S$

volume proteina

volume solvente

$$\%V_S = \frac{V_S}{V_C} = \frac{V_C - V_P}{V_C} = 1 - \frac{ZMw\rho}{V_C} = 1 - \frac{1.23}{V_M}$$

$$\%V_S \in \left[28\%, 65\% \right] \longrightarrow \langle \%V_S \rangle \approx 43\%$$

Funzione di Patterson

$$P(\mathbf{u}) = P(u,v,w) = \frac{1}{V} \sum_{\mathbf{h}} |F_{\mathbf{h}}|^2 \exp(-2\pi i \mathbf{h} \cdot \mathbf{u}) = \int_V \rho(\mathbf{r}) \rho(\mathbf{r} + \mathbf{u}) d^3 r$$

$$F_{\mathbf{h}} = \int_V d^3 r \rho(\mathbf{r}) \exp(2\pi i \mathbf{h} \cdot \mathbf{r})$$

$$F_{\mathbf{h}}^* = \int_V d^3 r \rho(\mathbf{r}) \exp(-2\pi i \mathbf{h} \cdot \mathbf{r}) \quad \rho(\mathbf{r}) = \rho^*(\mathbf{r}) \in \mathfrak{R}$$

$$\begin{aligned} P(\mathbf{u}) &= \frac{1}{V} \sum_{\mathbf{h}} \int_V \int_V d\mathbf{r} d\mathbf{r}' \rho(\mathbf{r}) \rho(\mathbf{r}') \exp(2\pi i \mathbf{h} \cdot (\mathbf{r}' - \mathbf{r})) \exp(-2\pi i \mathbf{h} \cdot \mathbf{u}) = \\ &= \frac{1}{V} \sum_{\mathbf{h}} \int_V \int_V d\mathbf{r} d\mathbf{r}' \rho(\mathbf{r}) \rho(\mathbf{r}') \exp(-2\pi i \mathbf{h} \cdot (\mathbf{u} - \mathbf{r}' + \mathbf{r})) = \\ &= \int_V d\mathbf{r} \rho(\mathbf{r}) \rho(\mathbf{r} + \mathbf{u}) \end{aligned}$$

\downarrow \downarrow
 $\mathbf{r}' = \mathbf{r} + \mathbf{u}$

La funzione di Patterson e' diversa da 0 quando \mathbf{u} e' un vettore che congiunge 2 atomi

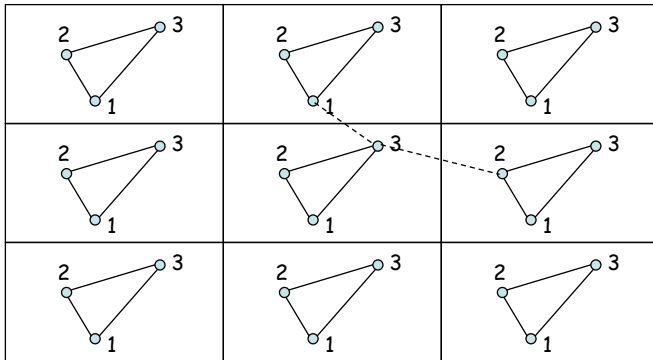
9

Esempio di funzione di Patterson

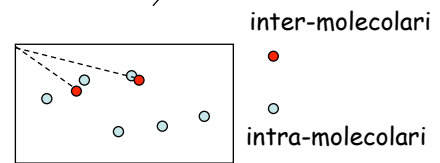
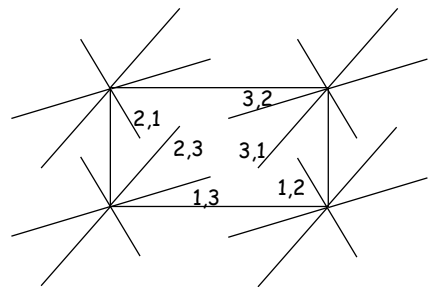
$$P(\mathbf{u}) = \int_V d\mathbf{r} \rho(\mathbf{r}) \rho(\mathbf{r} + \mathbf{u})$$

Spazio Patterson: $N^2 - N = N(N-1)$ picchi
 prot. di 200 aa. \sim 2000 atomi \sim 4×10^6 picchi

Spazio reale: 3 atomi



spazio patterson: 6 picchi



Ricerca rotazionale: convoluzione di 2 funzioni di Patterson

$$R(\alpha, \beta, \gamma) = \int_U d^3 u P^{esp}(\mathbf{u}) P_R^{mod}(\mathbf{u}_R)$$

volume di integrazione U

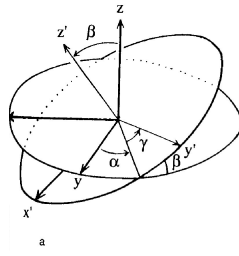
$$I \propto z^2$$

(Atomi pesanti)

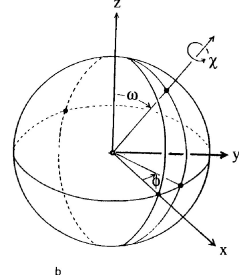
10

Ricerca rotazionale

Angoli di Eulero



Angoli polari



Funzione di sovrapposizione delle funzioni di Patterson sperimentale e del modello

$$\mathbf{x}_R = \hat{C}\mathbf{x} \quad \hat{C}^T = \hat{C}^{-1} \quad C = \text{operatore di rotazione}$$

$$R(\alpha, \beta, \gamma) = \int_U d^3u P^{exp}(\mathbf{u}) P_R^{mod}(\mathbf{u}_R) = \frac{1}{V^2} \int_U d^3u \sum_{\mathbf{h}, \mathbf{h}'} |F_{\mathbf{h}}|^2 |F_{\mathbf{h}'}|^2 \exp(-2\pi i(\mathbf{h} \cdot \mathbf{u} + \mathbf{h}' \cdot \hat{C}\mathbf{u})) =$$

$$= \frac{1}{V^2} \int_U d^3u \sum_{\mathbf{h}, \mathbf{h}'} |F_{\mathbf{h}}|^2 |F_{\mathbf{h}'}|^2 \exp(-2\pi i(\mathbf{h} + \hat{C}^{-1}\mathbf{h}') \cdot \mathbf{u})$$

$$\downarrow \quad \mathbf{H} \equiv \mathbf{h} + \hat{C}^{-1}\mathbf{h}'$$

$$\int_U d^3u \exp(-2\pi i\mathbf{H} \cdot \mathbf{u})$$

11

$$\int_U d^3u \exp(-2\pi i\mathbf{H} \cdot \mathbf{u}) = \int_U du d\varphi \sin\varphi d\vartheta u^2 \exp(-2\pi iHu \cos\varphi) =$$

$$= -2\pi \int du d(\cos\varphi) u^2 \exp(-2\pi iHu \cos\varphi) =$$

$$u \in [0, U]$$

$$\varphi \in [0, \pi]$$

$$\vartheta \in [0, 2\pi]$$

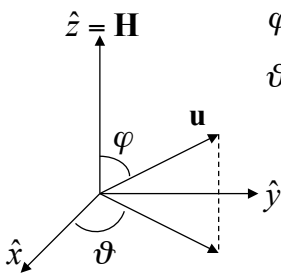
$$= 2\pi \int du u^2 \frac{1}{2\pi iHu} 2i \sin(2\pi Hu) =$$

$$= \frac{2}{H} \int du u \sin(2\pi Hu)$$

$$\begin{aligned} 2\pi Hu &\equiv x \\ 2\pi HU &\equiv X \end{aligned}$$

$$\downarrow$$

$$\frac{2}{(2\pi H)^2 H} \int dx x \sin(x)$$



Integrazione per parti

$$\int_0^X dx x \sin(x) = -x \cos x \Big|_0^X + \int_0^X dx \cos(x) = (-x \cos x + \sin x) \Big|_0^X = \sin X - X \cos X$$

12

$$\int_U d^3u \exp(-2\pi i \mathbf{H} \cdot \mathbf{u}) = \frac{2}{(2\pi H)^2 H} [\sin(2\pi H U) - 2\pi H U \cos(2\pi H U)] = F(H)$$

$F(H)$ è massima quando $H=0$ e va a 0 quando H si allontana da 0



$F(H)$ seleziona valori di H vicini a 0!

$$R(\alpha, \beta, \gamma) = \int_U d^3u P(\mathbf{u}) P_R(\mathbf{u}_R) = \frac{1}{V^2} \sum_{\mathbf{h}, \mathbf{h}'} |F_{\mathbf{h}}|^2 |F_{\mathbf{h}'}|^2 F(H)$$

$$\mathbf{H} \equiv \mathbf{h} + \hat{C}^{-1} \mathbf{h}' \longrightarrow \mathbf{h} \approx -\hat{C}^{-1} \mathbf{h}'$$

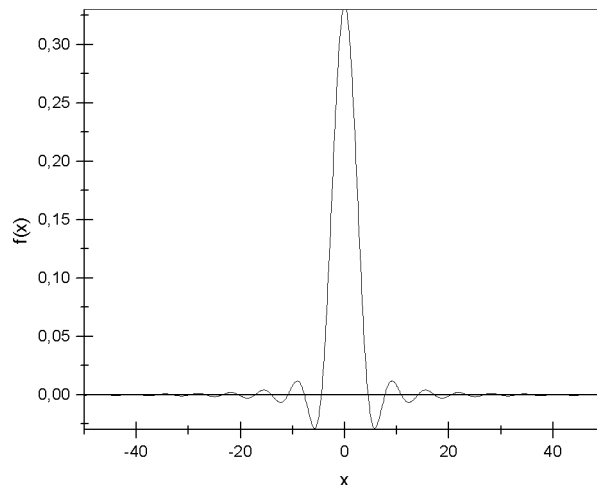
$F(H)$ limita la doppia sommatoria!

13

$$\lim_{H \rightarrow 0} F(H) = \lim_{x \rightarrow 0} \left(4\pi U^3 \frac{\sin x - x \cos x}{x^3} \right) = 4\pi U^3 \lim_{x \rightarrow 0} \frac{x \sin x}{3x^2} =$$

$$= 4\pi U^3 \lim_{x \rightarrow 0} \frac{\sin x + x \cos x}{6x} = 4\pi U^3 \lim_{x \rightarrow 0} \frac{\cos x + \cos x - x \sin x}{6} = \frac{4\pi U^3}{3}$$

$$\lim_{H \rightarrow \pm\infty} F(H) = 0$$



14

Ricerca traslazionale:

Minimizzazione di R:

$$R = \frac{\sum_h |F_h(obs) - k F_h(calc)|}{\sum_h |F_h(obs)|}$$

Massimizzazione del coefficiente di correlazione C:

$$C = \frac{\sum_h \left(|F_h(obs)|^2 - \overline{|F_h(obs)|^2} \right) \sum_k \left(|F_k(calc)|^2 - \overline{|F_k(calc)|^2} \right)}{\left[\sum_h \left(|F_h(obs)|^2 - \overline{|F_h(obs)|^2} \right)^2 \sum_k \left(|F_k(calc)|^2 - \overline{|F_k(calc)|^2} \right)^2 \right]^{1/2}}$$

Significato del coefficiente di correlazione:

$$\sum_i (x_i - \bar{x}) \sum_j (y_j - \bar{y}) = \sum_{i,j} x_i y_j - N \bar{y} \sum_i x_i - N \bar{x} \sum_i y_i + N^2 \bar{x} \bar{y} = \sum_{i,j} x_i y_j - N^2 \bar{x} \bar{y}$$

Se x e y sono grandezze scorrelate: \downarrow = 0

15

Sommario molecular replacement

Ricerca di una struttura da usare come modello nella banca dati pdb

Numero di Matthews: quante sono le molecole nell'unità asimmetrica (u.a.) ?

\downarrow z

Funzione di **self rotation**: eventuale(i) asse(i) di **simmetria non cristallografica (NCS)**

(i.e. z=3, ci potrebbe essere 1 asse di simmetria di ordine 3 = rot. di 120°)

$$\downarrow \quad R(\alpha, \beta, \gamma) = \int_U d^3 u P^{esp}(\mathbf{u}) P^{esp}_R(\mathbf{u}_R)$$

Cross rotation: ricerca dell'orientazione del modello nel cristallo: m > z picchi rotazionali

$$\downarrow \quad R(\alpha, \beta, \gamma) = \int_U d^3 u P^{esp}(\mathbf{u}) P^{mod}_R(\mathbf{u}_R)$$

Translation: partendo da ciascuna delle m orientazioni ricerca delle

z posizioni del modello nell'u.a.

\downarrow

massimizzo: $C = \frac{\sum_h \left(|F_h(obs)|^2 - \overline{|F_h(obs)|^2} \right) \sum_k \left(|F_k(calc)|^2 - \overline{|F_k(calc)|^2} \right)}{\left[\sum_h \left(|F_h(obs)|^2 - \overline{|F_h(obs)|^2} \right)^2 \sum_k \left(|F_k(calc)|^2 - \overline{|F_k(calc)|^2} \right)^2 \right]^{1/2}}$

SOLUZIONE

16