With slides from Silvia Onesti and Alberto Cassetta

X-ray macromolecular crystallography: an overview

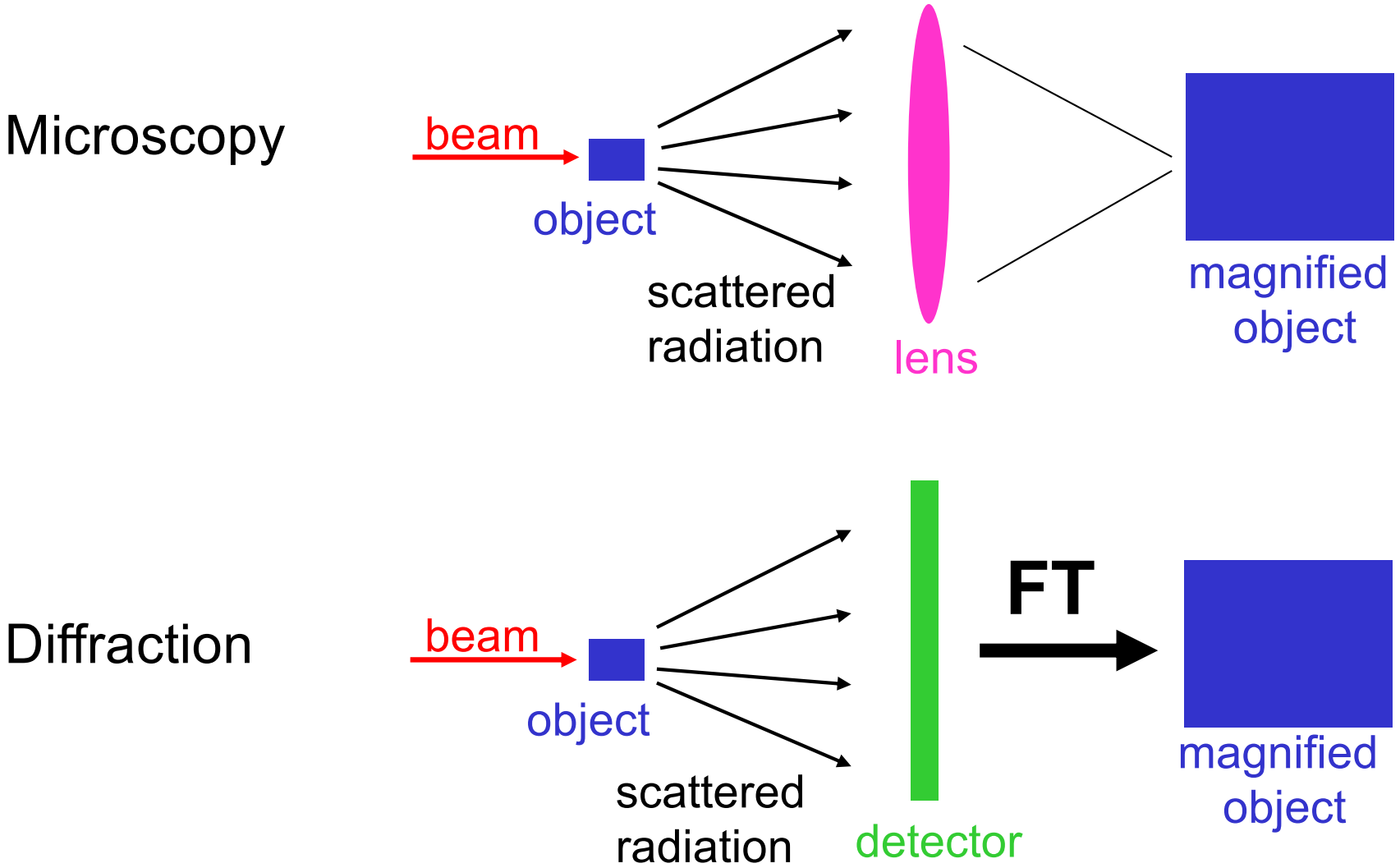
Protein structure determination methods

- X-ray crystallography (protein crystallography)
- X-ray fiber diffraction
- Small angle X-ray scattering (SAXS)

- Scanning electron microscopy (SEM)
- Transmission electron microscopy (TEM):
  - 2D crystals (electron diffraction)
  - Fibers/helices/tubular crystals
  - Single particle electron microscopy
- Atomic Force Microscopy (AFM)

- Nuclear magnetic resonance (NMR)

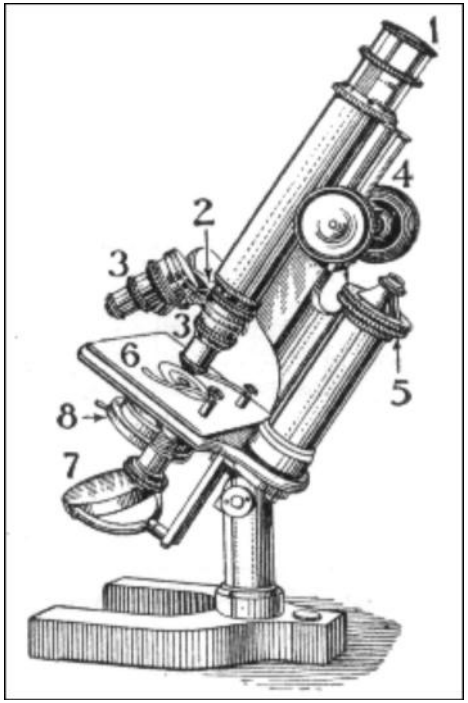
Microscopy vs diffraction

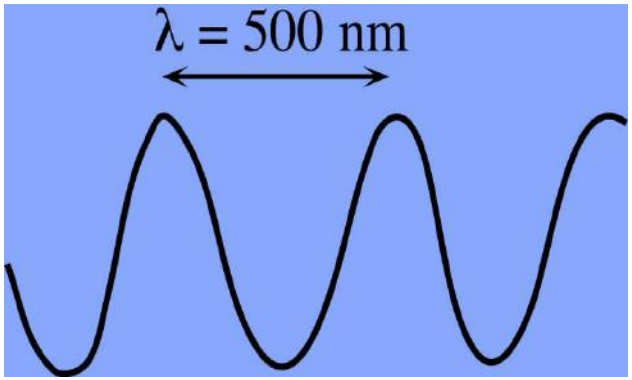


Why can't we use a microscope?

Normally, to look at small objects we use microscopes...

...but they can only provide images of things larger than the wavelength of light





- Protein (10 nm)
- Atom (0.1 nm=1Å)

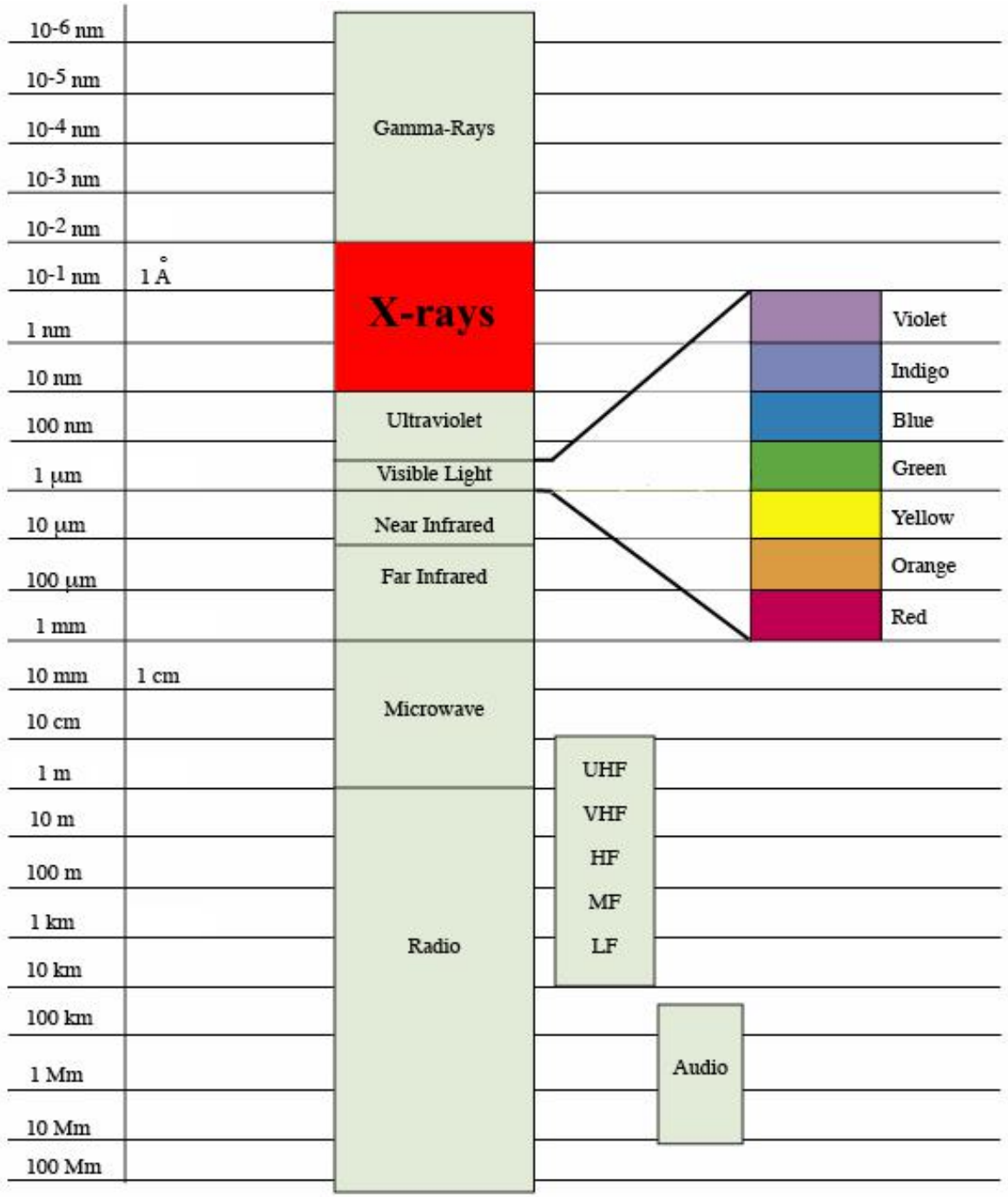
In theory, we could use **X-rays** (light of λ=0.1 nm, the **right size** for looking at atoms)...

X-rays

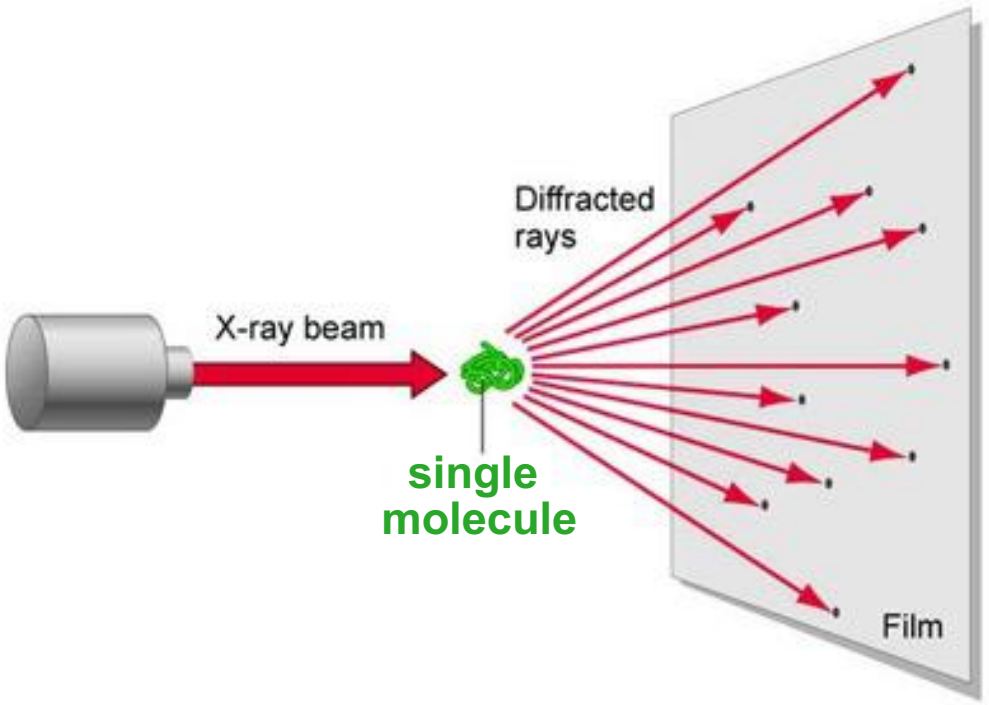
X-rays:
E = 1-100 keV
λ = 10 nM-0.01 nm

Soft X-rays (lower E) are used in medical imaging.
Hard X-rays (higher E) are used in crystallography.

...but we can't make X-ray microscopes as hard X-rays do not reflect or refract easily.



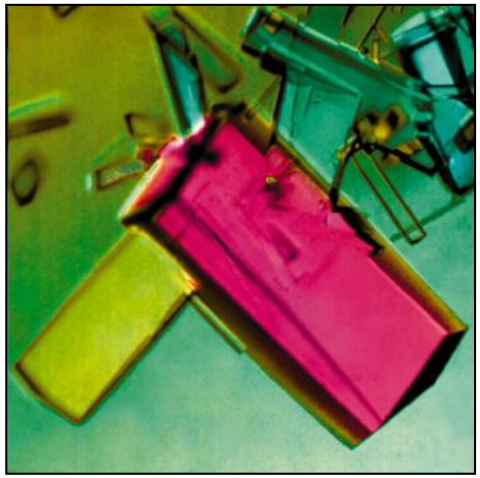
X-ray diffraction with single objects?



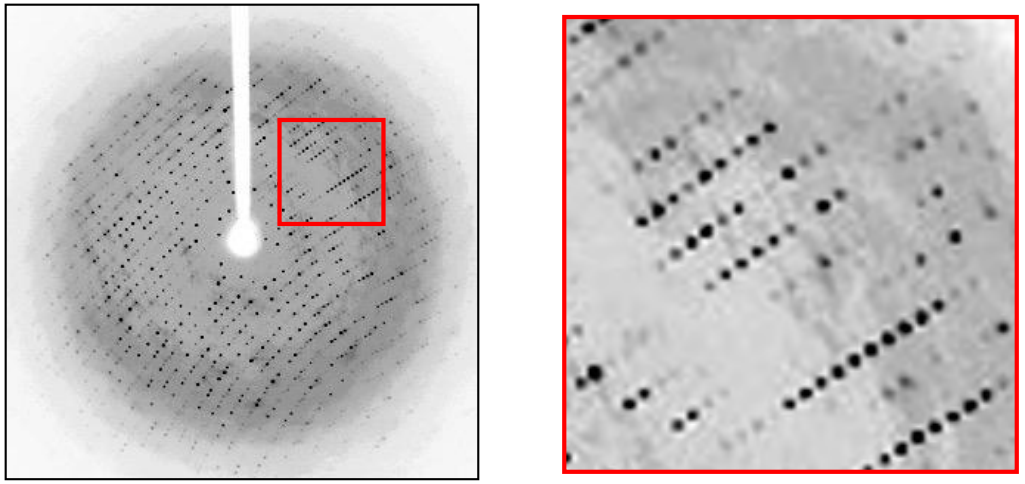
The intensity of the X-ray radiation diffracted by a single biological object (such as a protein or a cell) is very small:

➡ almost impossible to measure with current technology.

But we can use X-rays with crystals

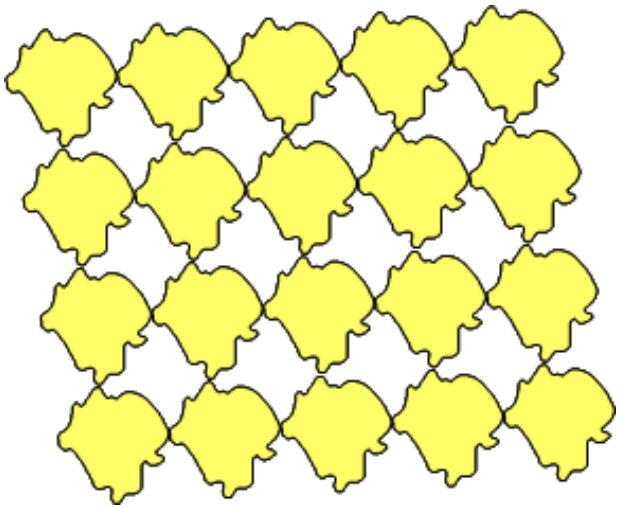


Crystal



X-ray Diffraction pattern

Crystal: an ordered array of molecules



Two effects on diffraction:

- constructive interference in some directions, so that all molecules scatter in phase: strong signal
- destructive interference in most directions, so that the diffraction pattern is non-zero only at a few specific position: discrete spots

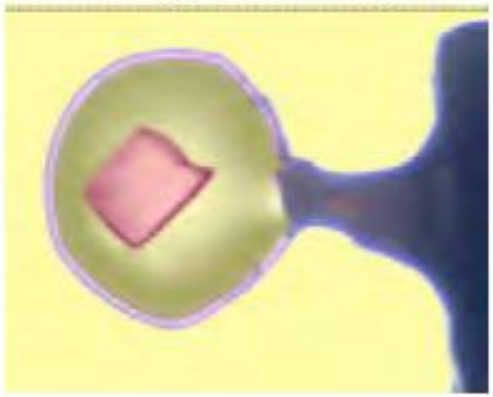
Raccolta dati

Temperatura ambiente
(293-298K)

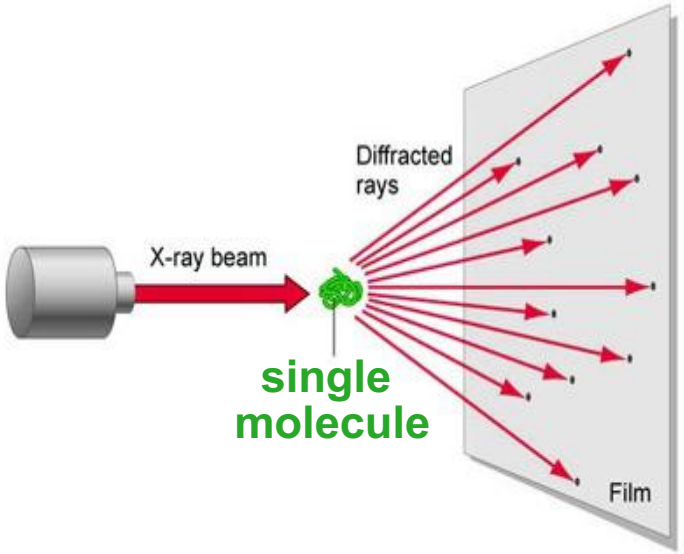
Il cristallo viene
montato in un capillare

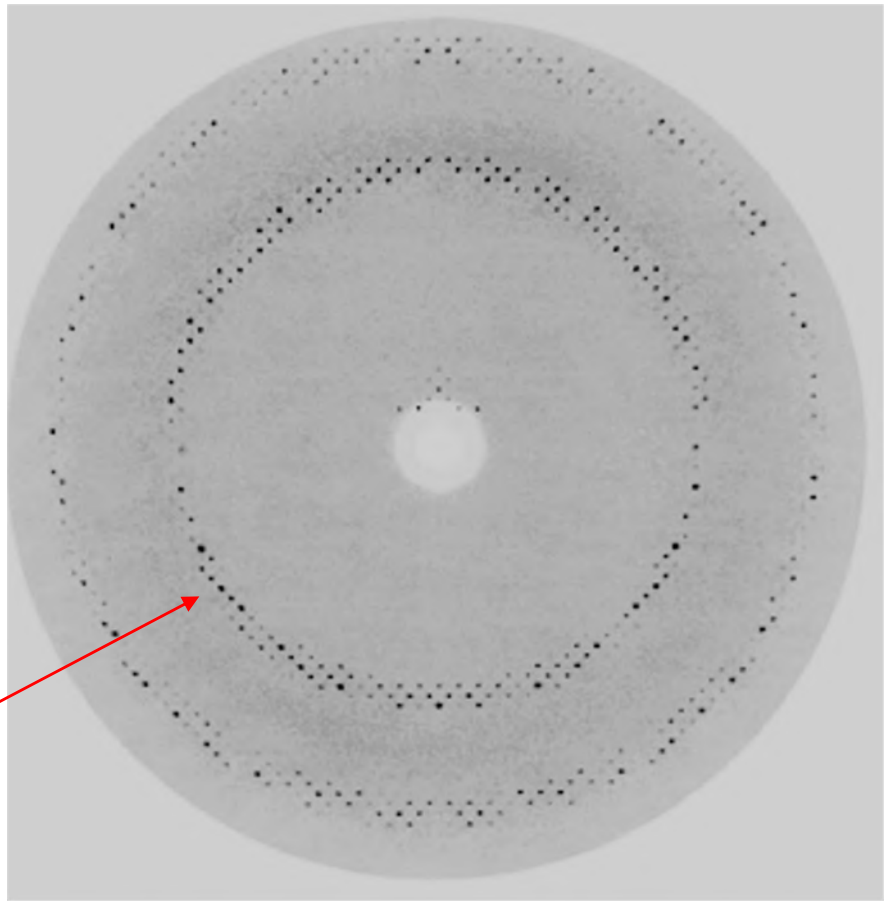
Temperature criogeniche
(50 - 100K)

Il cristallo, protetto con un anticongelante, viene montato in un loop di nylon e raffreddato sotto flusso di N_2 o He liquidi



Scattering from a 3D crystal





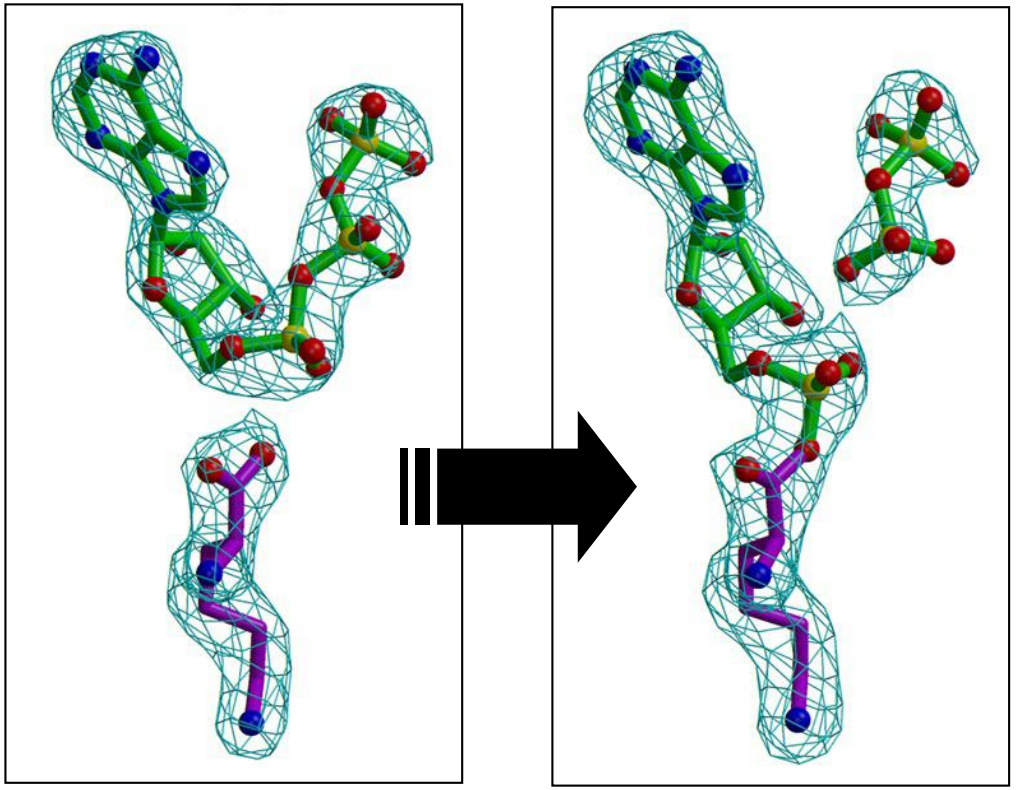
Each of this points (reflection) corresponds to a particular direction **s** and therefore to a particular angle 2θ for which Bragg's condition is satisfied – we can represent each as a set of 3 coordinates (hkl).

Does crystallisation affects the structure?

- The forces that hold molecules in a lattice are very much weaker than those that hold protein structures together, so gross conformational changes triggered by crystallisation are unlikely (but they do occur). However the crystal lattice may favour a conformation that is not dominant in solution.

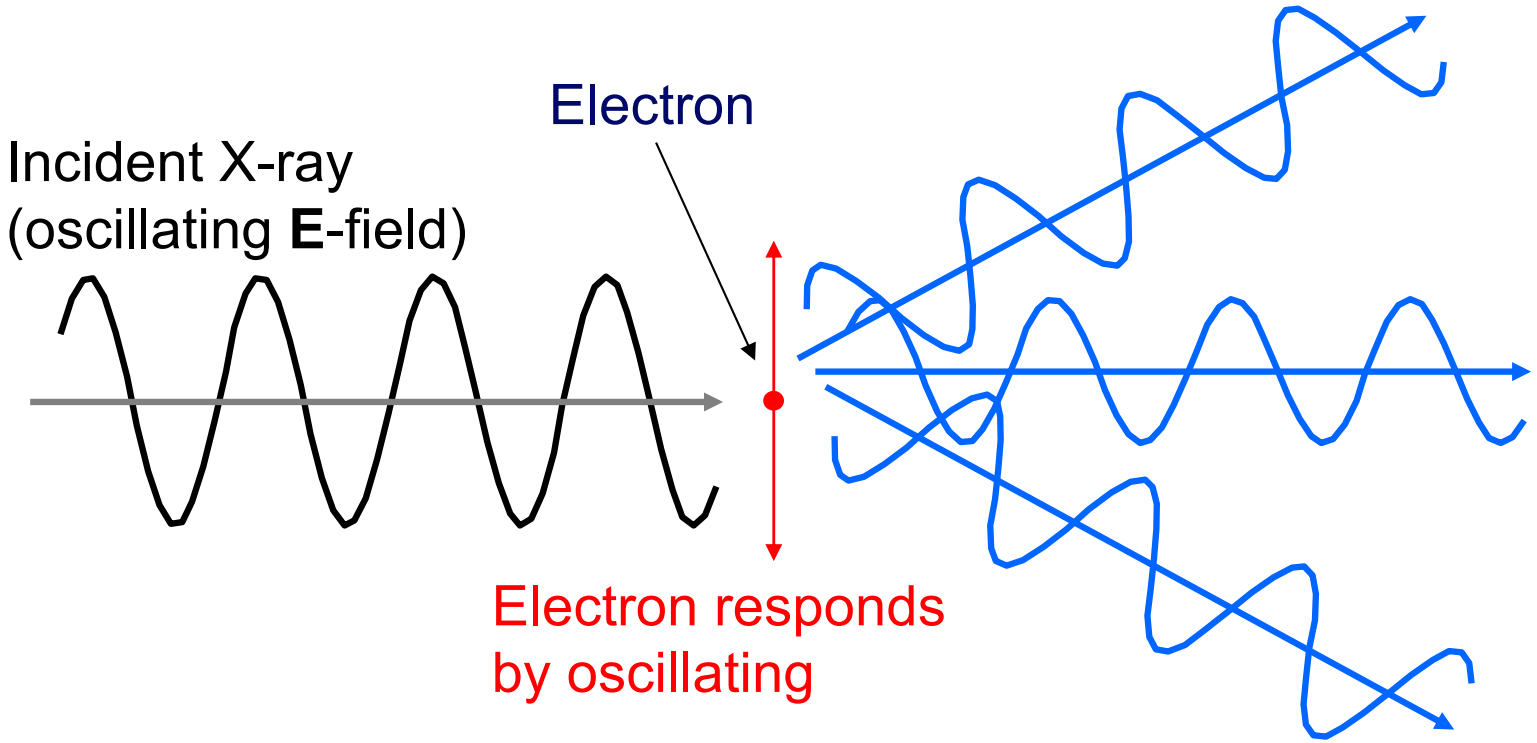
- Proteins crystallised in different crystal forms are often identical or almost identical.

- Some enzymes retain their biological activity within the crystal (a strong indication of a native-like structure).



X-ray scattering by a single electron

➡ the scattering of X-rays is mostly due to electrons



The electron is said to "scatter"
or "diffract" the X-ray

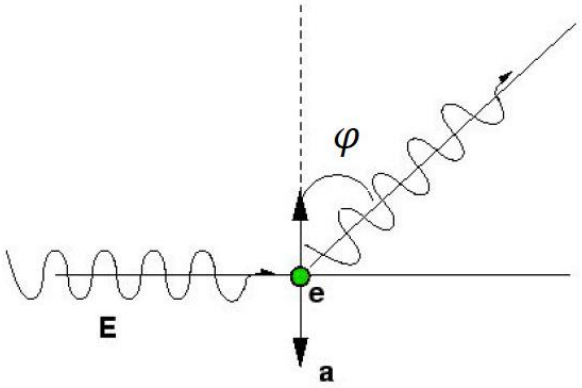
Oscillating electron emits
X-rays over a wide angle

Elastic scattering (Thompson)

Un'onda elettromagnetica incidente, interagisce con una particella di carica ***e*** :

$$\vec{E}_i = E_{i0}\, exp[2\pi\nu i\left(t - \frac{x}{c}\right)]$$

La particella di carica ***e*** è soggetta ad una forza pari a: $\vec{F} = e\overrightarrow{E_i}$



La forza causa un moto oscillatorio di frequenza uguale a quella dell'onda incidente e con accelerazione e pari a:

$$\vec{a} = \frac{\vec{F}}{m} = e\frac{\overrightarrow{E_i}}{m}$$

Una particella carica oscillante è a sua volte fonte di radiazione elettromagnetica, in tutte le direzioni:

$$\vec{E}_d = E_{d0} exp[2\pi\nu i\left(t - \frac{\vec{r}}{c}\right) - i\alpha]$$

(α è un termine di sfasamento ed è uguale a π per l'elettrone)

Elastic scattering (Thompson)

Poicè l'intensità della radiazione è data dalla relazione:

$$I = |E|^2$$

Nel caso della diffusione di Thomson abbiamo:

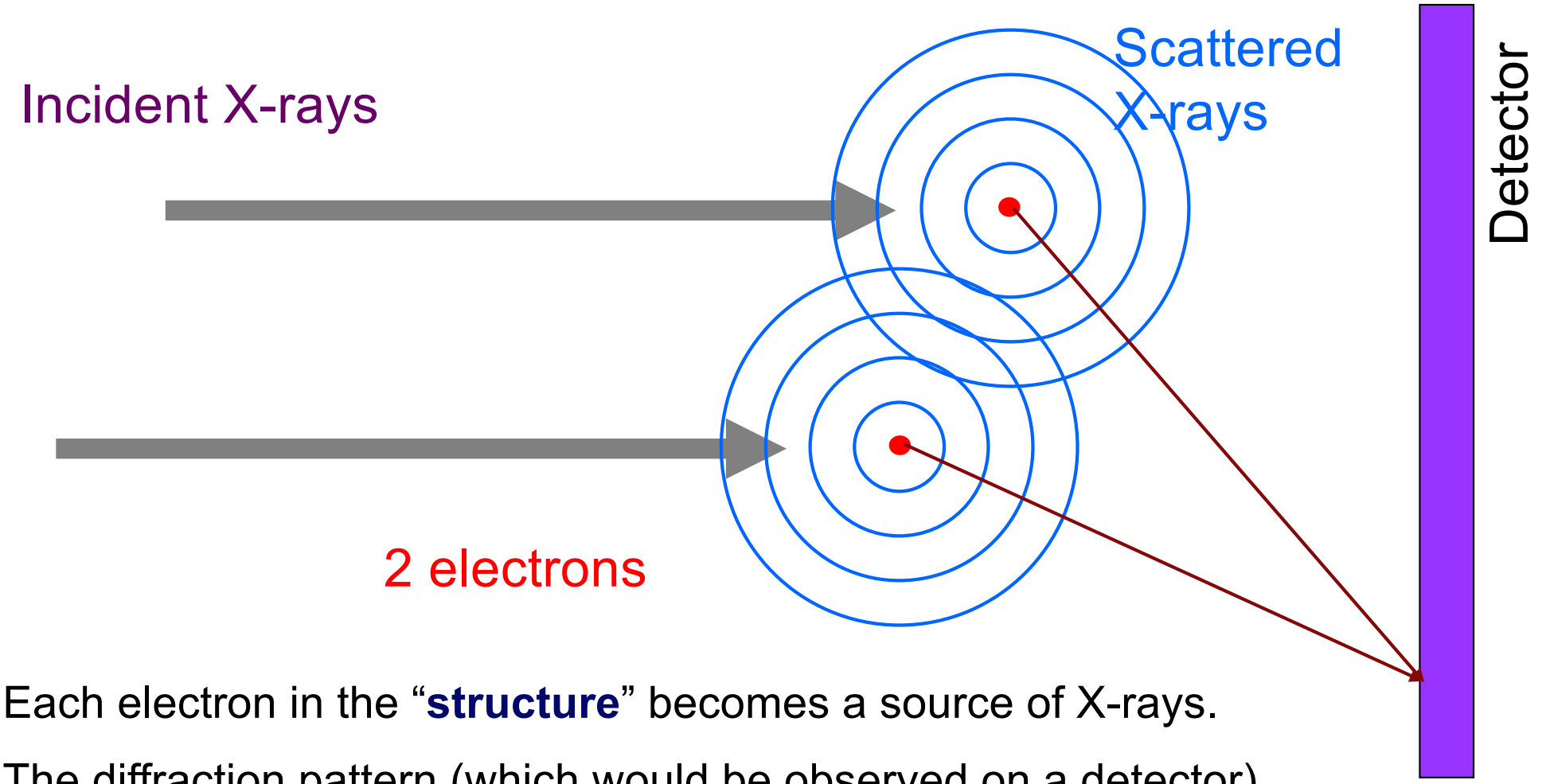
$$I_{eTh} = I_{d0} = |E_{d0}|^2 = \left|\frac{1}{r} E_{i0} \left(\frac{e^2}{mc^2}\right) \sin\varphi\right|^2 = I_{i0} \frac{e^4}{m^2 r^2 c^4} \sin^2\varphi$$

- I_{eTh} = 0 se la particella non è carica (**neutroni non interagiscono con le particelle cariche**)
- I_{eTh} dipende dal rapporto carica/massa, **per un protone il contributo è 1837 volte più debole che per l'elettrone**

L'interazione dei raggi-X con la materia è determinata dagli elettroni!!!.

La diffrazione dei raggi-X vede gli elettroni, non i nuclei!!!

X-ray scattering by two electrons



Each electron in the "**structure**" becomes a source of X-rays.

The diffraction pattern (which would be observed on a detector) is the resultant of **adding** the scattered X-ray waves.

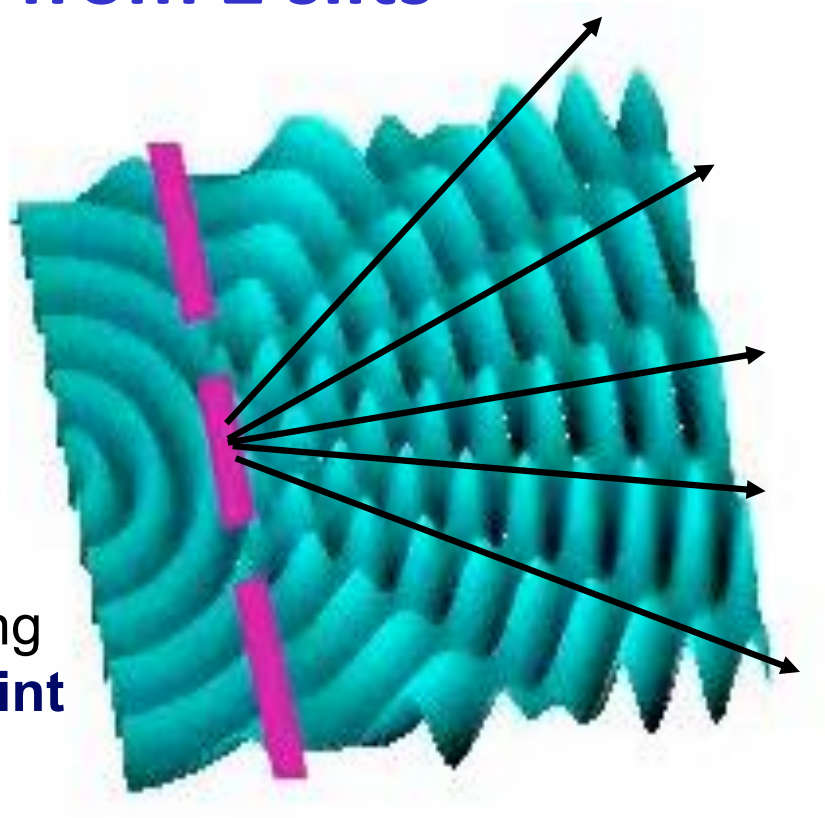
To understand diffraction, we need to know **how to add waves**, i.e. how waves interfere with each other.

Example of diffraction from 2 slits

Diffraction of water waves through 2 slits gives a pattern that **depends** on the slit structure (width, separation) and the wavelength of the waves.

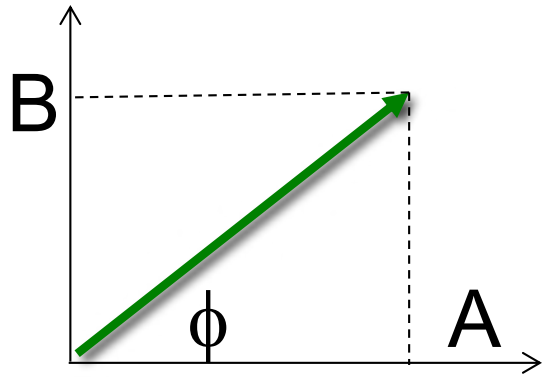
If we can work out this relationship, we can measure the diffraction pattern and figure out the "structure" that gives rise to it.

The observed pattern is the resultant of adding the wave arriving from the 2 slits at **every point** on the detector



Waves can be written as complex numbers/vectors:

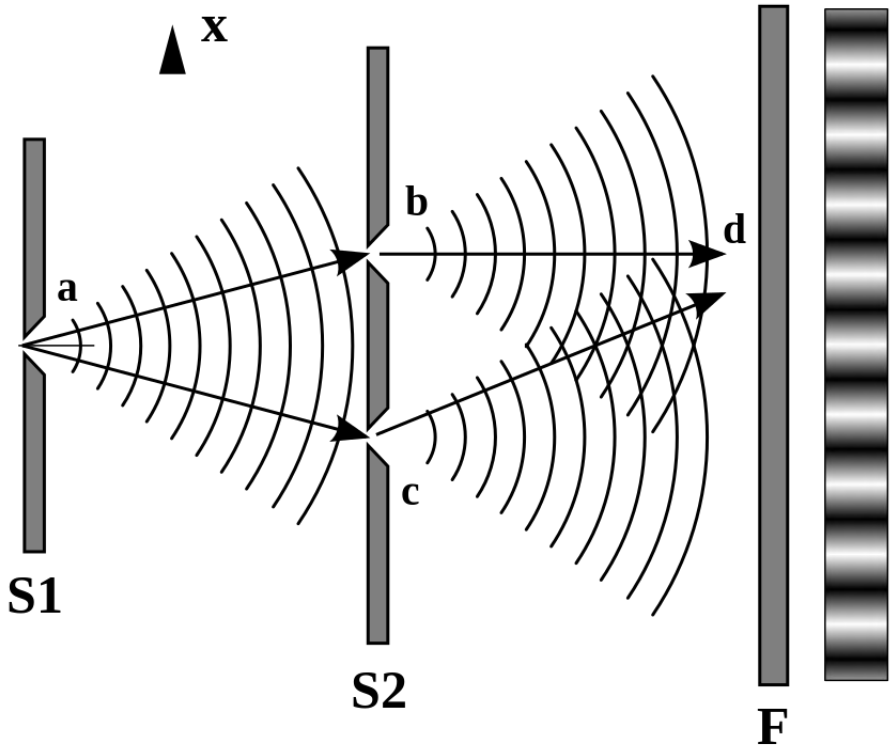
$$\mathbf{Z} = |Z| \exp(i\phi) = |Z|(\cos\phi + i\sin\phi) = A + iB$$



X-ray scattering: interference

•Scattering: ability of an object to change the direction of a wave.

•If two objects (A and B) are hit by a wave they act as a point source of a new wave with same wavelength and velocity (Huygens' principle)



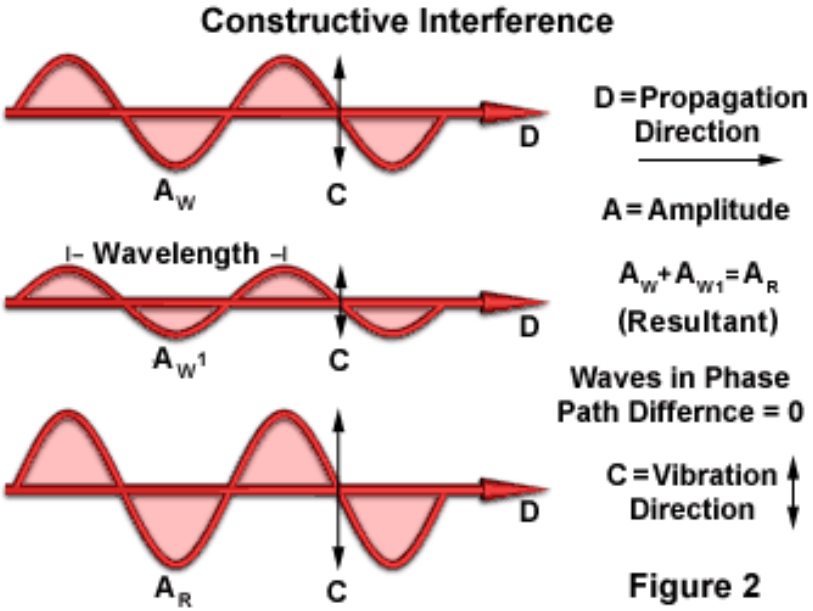


Figure 2

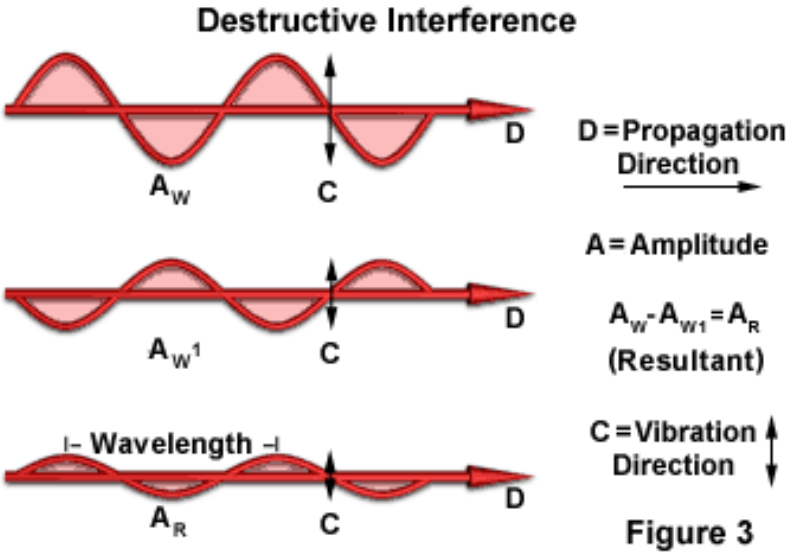


Figure 3

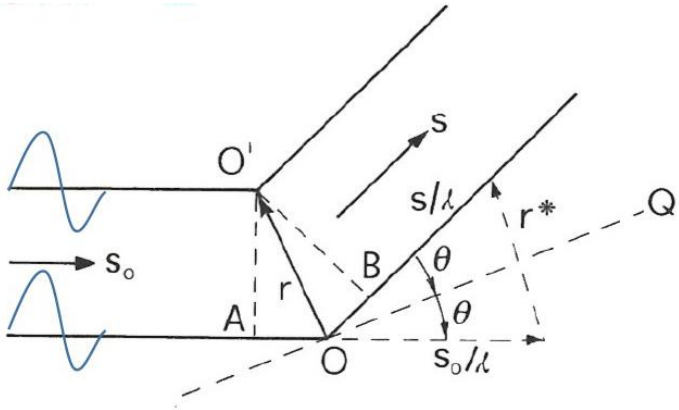
X-ray scattering: interference

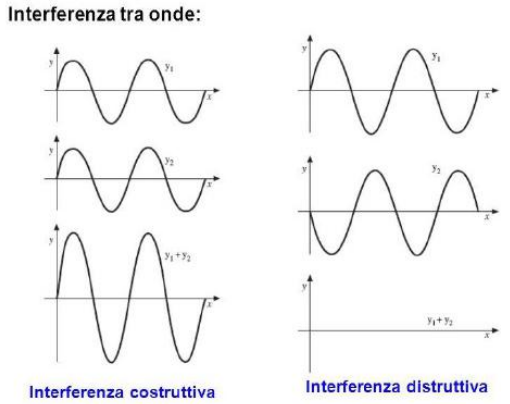
Un'onda (piana) di **lunghezza d'onda** λ interagisce con due elettroni il primo in posizione O e il secondo in posizione O', tra di loro a distanza $\vec{r}$.

Ogni singolo elettrone è sorgente di un'onda diffusa (secondo la descrizione di Thompson).

Le due onde diffuse mantengono una relazione costante di fase (sono coerenti), quindi daranno luogo al fenomeno fisico dell'interferenza, il cui effetto (costruttivo o distruttivo) dipenderà dalla direzione di incidenza dell'onda $\overrightarrow{s_0}$ e dalla direzione di 'osservazione $\vec{s}$ (descritta da vettori di modulo pari a $\frac{1}{\lambda}$).

La condizione di interferenza tra le due onde diffuse è determinate dalla differenza di cammino ottico tra le due onde.





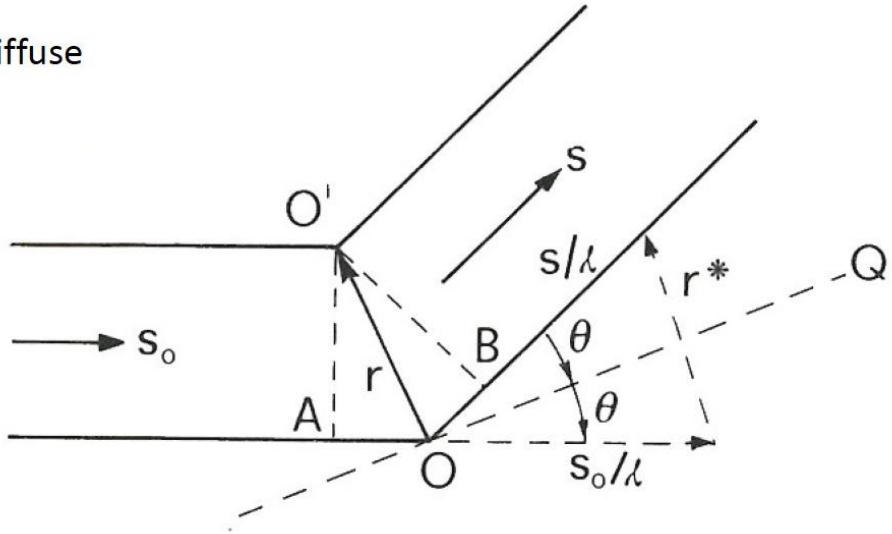
X-ray scattering: interference

Calcoliamo la differenza di cammino ottico delle due onde diffuse

$$\overline{OA} + \overline{OB} = \textbf{differenza di cammino ottico}$$

$\overline{OA} = -\vec{r} \cdot \overrightarrow{s_o}$
$(proiezione\ di\ \vec{r}\ lungo\ \overrightarrow{s_o};\ segno$
$-\ perchè\ opposto\ alla\ direzione\ di\ \vec{s})$

$\overline{OB} = \vec{r} \cdot \vec{s}$ $(proiezione\ di\ \vec{r}\ lungo\ \vec{s})$



Esprimendo la differenza di cammino ottico come **differenza di fase**:

$$\delta = 2\pi \frac{\overline{OA} + \overline{OB}}{\lambda} = 2\pi \frac{\vec{r} \cdot (\vec{s} - \overrightarrow{s_0})}{\lambda}$$

è pari al numero di 'periodi'

X-ray scattering: interference

Abbiamo visto che la differenza di fase tra due onde diffuse a partire dalla medesima onda (piana) è data da:

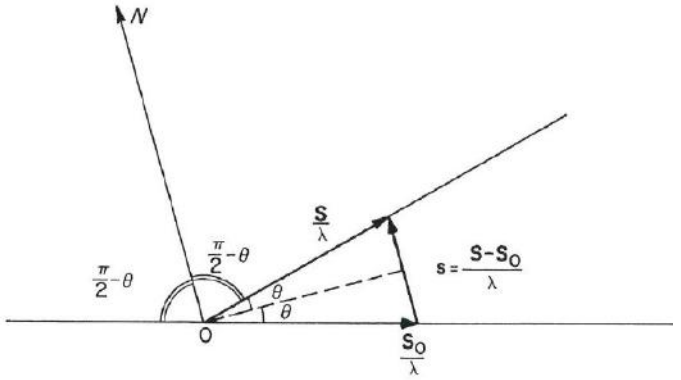
$$2\pi \frac{\vec{r} \cdot (\vec{s} - \vec{s_0})}{\lambda}$$

Se **definiamo**:

$$\vec{r}^* = \frac{(\vec{s} - \vec{s_0})}{\lambda}$$

Quindi:

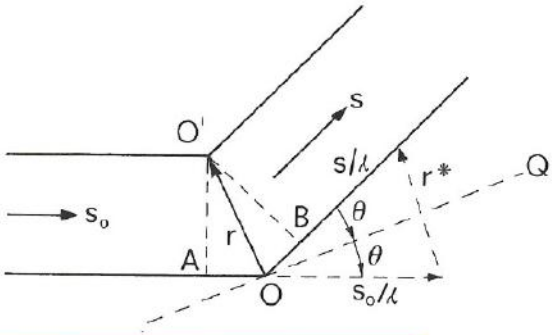
$$\delta = 2\pi \vec{r} \cdot \vec{r^*}$$



Il vettore $\vec{r^*}$ é perpendicolare alla bisettrice di $\vec{s}$ e $\vec{s_0}$ con cui forma un angolo θ

Il modulo del vettore $\vec{r^*}$ è pari a:

$$|\vec{r^*}| = \frac{2sin\theta}{\lambda}$$



X-ray scattering: interference

Condizioni per l'interferenza

la differenza di fase tra due onde diffuse a partire dalla medesima onda (piana) è data da:

$$2\pi \frac{\vec{r} \cdot (\vec{s} - \overrightarrow{s_0})}{\lambda}$$

Che dipende dal rapporto $\boldsymbol{\frac{r}{\lambda}}$

ovvero dal rapporto tra le distanze tra le cariche e la lunghezza d'onda della radiazione incidente.
Se r << λ, non c'è interferenza poichè la differenza di fase (il rapporto r/λ è circa 0)

Se invece r >> λ, perchè $\vec{r} \cdot (\vec{s} - \overrightarrow{s_0})$ sia uguale a $2\pi\lambda$, $\vec{s} - \overrightarrow{s_0}$ deve essere circa 0 (osservo raggi-X diffusi solo a bassissimi angoli (lungo la direzione dei raggi incidenti)

λ deve essere dello stesso ordine di grandezza di r

X-ray scattering: interference

Abbiamo visto in precedenza che l'onda diffusa da un singolo elettrone è pari a:

$$\vec{E}_d^1 = E_{d0} exp[2\pi\nu i\left(t - \frac{\vec{r}}{c}\right) - i\alpha] \qquad \text{Elettrone 1}$$

Se abbiamo un secondo elettrone, questo diffonderà a sua volta un'onda, uguale a quella della prima, a meno del termine di sfasamento δ:

$$\vec{E}_d^2 = E_{d0} exp[2\pi\nu i\left(t - \frac{\vec{r}}{c}\right) - i\alpha + i\delta] \qquad \text{Elettrone 2}$$

Che possiamo scrivere come:

$$\vec{E}_d^2 = E_{d0} exp\left[2\pi\nu i\left(t - \frac{\vec{r}}{c}\right) - i\alpha\right] \exp(i\delta)$$

$\vec{E}_d^1$

Ovvero:

$$\boldsymbol{\vec{E}_d^2 = \vec{E}_d^1 exp(i\delta)}$$

La somma delle onde, ovvero l'onda diffusa risultante potrà essere scritta come:

$$\vec{E}_d = \vec{E}_d^1 + \vec{E}_d^2 = \vec{E}_d^1[1 + \exp(i\delta)]$$

In forma esplicita:

$$\boldsymbol{\vec{E}_d(r^*) = \vec{E}_d^1[1 + exp(2\pi i\vec{r} \cdot \vec{r^*})]}$$

Questo termine additivo determina la somma o sottrazione delle due onde

Diffusione da N elettroni - 1

Se abbiamo N elettroni, tutti uguali nelle posizioni $r_1, r_2, \ldots, r_N$ dal primo elettrone (r_1 sarà 0, origine del sistema di riferimento), e applicando lo stesso procedimento utilizzato per 2 elettroni, otteniamo:

$$\vec{E}_d(r^*) = \vec{E}_d^1 \sum_{j=1}^{N} \exp\left(2\pi i \vec{r_j} \cdot \vec{r^*}\right)$$

Dove r* denota la 'direzione di osservazione', mentre r_j denota la posizione del diffusore j-simo

Più in generale, ammettendo che i diffusori abbiamo una carica generica (non necessariamente quella di 1 elettrone), allora ogni diffusore genererà onde di ampiezza E_0 non necessariamente uguali tra loro (maggiore è la carica maggiore maggiore è l'ampiezza dell'onda diffusa,vedi formula di Thomson) e che indicheremo cone $\boldsymbol{A_j}$
Potremo così scrivere:

$$\vec{E}_d(r^*) = \sum_{j=1}^{N} A_j \exp\left(2\pi i \vec{r_j} \cdot \vec{r^*}\right)$$

Diffusione da N elettroni - 2

$$\vec{E}_d(r^*) = \sum_{j=1}^{N} A_j \exp\left(2\pi i \vec{r_j} \cdot \vec{r^*}\right)$$

Se invece dell'ampiezza dell'onda diffusa usiamo il rapporto tra l'intensità dell'onda diffusa e quella diffusa da un singolo elettrone secondo la teoria di Thomson, che indichiamo come $\mathbf{I_{Th}}$, avremo:

$$\vec{F}(r^*) = \sum_{j=1}^{N} f_j \exp\left(2\pi i \vec{r_j} \cdot \vec{r^*}\right)$$

Dove:

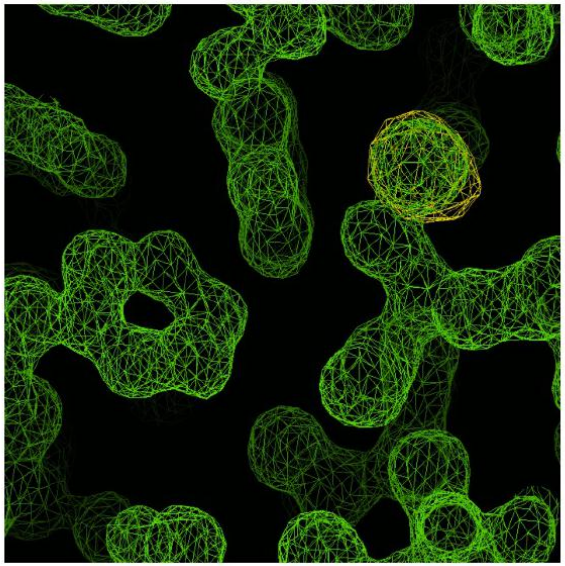
$$f^2{}_j = \frac{I_j}{I_{Th}}$$

f_j è definite come il fattore di struttura del diffusore

Diffusione da un mezzo continuo - 1

Se invece di avere N cariche elettriche puntuali, abbiamo **una distribuzione continua** (non necessariamente omogenea) **di cariche elettriche** $\boldsymbol{\rho(\vec{r})}$,
L'elemento infinitesimo di carica può essere descritto come: $\boldsymbol{\rho(\vec{r})d\vec{r}}$ a cui sarà associate una differenza di fase $\boldsymbol{2\pi i\vec{r_j}\cdot\vec{r^*}}$, **da cui dreiva il termine:**

$$\boldsymbol{\rho(\vec{r})\, exp\left(2\pi i\vec{r_j}\cdot\vec{r^*}\right)}$$



Volendo applicare la formula ottenuta per N elettroni ad un mezzo continuo, carico e di volume V, ovvero sommando su tutto lo spazio, otterremo:

$$\vec{F}\left(\vec{r^*}\right) = \int_V \rho(\vec{r}) \exp\left(2\pi i\vec{r_j}\cdot\vec{r^*}\right) d\vec{r}$$

L'integrazione sostituisce la sommatoria in un mezzo continuo

Diffusione e trasformata di Fourier

$$\vec{F}(\overrightarrow{r^*}) = \int_V \rho(\vec{r}) \exp\left(2\pi i \overrightarrow{r_j} \cdot \overrightarrow{r^*}\right) d\vec{r}$$

Questo integrale altro non è che la Trasformata di Fourier della distribuzione di carica distribuita in un volume V.

Il termine ***F(r*)*** è normalmente chiamato **fattore di struttura** o **fattore di scattering** e rappresenta l'onda elettromagnetica diffusa da una carica elettrica continua.
Lo si indica anche come:

$$\vec{F}(\overrightarrow{r^*}) = \mathbf{TF}[\rho(\vec{r})]$$

Considerazioni sulla diffusione -1

$$\vec{F}(\overrightarrow{r^*}) = \int_V \rho(\vec{r}) \exp\left(2\pi i \overrightarrow{r_j} \cdot \overrightarrow{r^*}\right) d\vec{r}$$

La Trasformata di Fourier stabilisce una relazione tra due spazi, in questo caso abbiamo lo spazio diretto ($\vec{r}$) e lo spazio detto ***spazio reciproco*** ($\overrightarrow{r^*}$).

La densità di carica ρ**(r) (elettronica) è definita nello spazio diretto**, mentre Il **fattore di struttura F(r*) sono definiti nello spazio reciproco**.

L'onda risultante dall'interferenza di un fronte d'onda di raggi-X con una distribuzione di cariche, osservata lungo una certa direzione di osservazione definita da r*, si può ottiene come trasformata di Fourier della distribuzione di cariche stesse.

Considerazioni sulla diffusione - 2

La Trasformata di Fourier ammette una **antitrasformata di Fourier**, descritta come l'operazione inversa. Nel caso del fattore di struttura e della distribuzione di elettroni avremo che:

$$\rho(\vec{r}) = \int_{V^*} \vec{F}(\overrightarrow{r^*}) \exp(-2\pi i \overrightarrow{r_j} \cdot \overrightarrow{r^*}) \, d\vec{r}^*$$

Dalla antitrasformata di Fourier dei fattori di struttura, posso ottenere la distribuzione in un spazio di volume V, della densità elettronica.

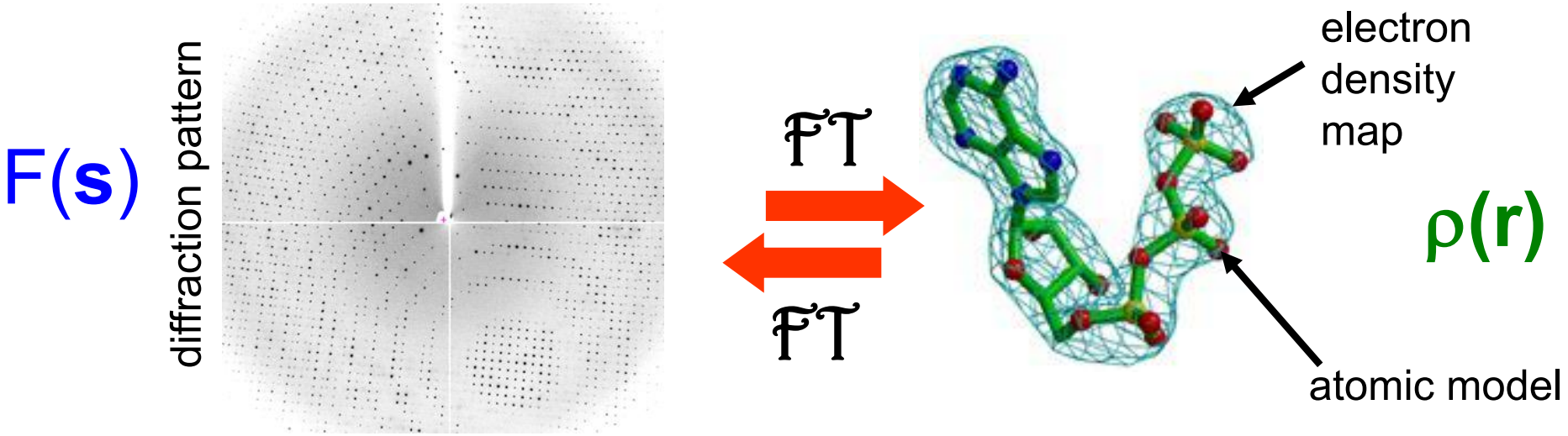
Quindi se noi conosciamo le onde diffuse dalla densità elettronica distribuita in un certo volume (in modulo e fase), noi possiamo ricostruire la distribuzione spaziale dalla densità elettronica stessa a partire dalla antitrasformata di Fourier delle onde diffuse.

Questo è il principio fondamentale della struttristica diffrattometrica

Diffraction and Fourier synthesis

$F(\mathbf{s}) = \mathcal{FT}[\rho(\mathbf{r})]$ the diffraction pattern is the Fourier transform of the electron density

$\rho(\mathbf{r}) = \mathcal{FT}[F(\mathbf{s})]$ the electron density is the Fourier Transform of the diffraction pattern



We can always go back and forward using Fourier transforms:

- if we know the electron density we can calculate the diffraction pattern
- if we know the diffraction pattern we can calculate the electron density

Riassumendo

- **Una carica elettronica è sempre sorgente di onde diffuse**
- **Le onde diffuse interferiscono tra loro**
- L'onda diffusa risultante dipende della **differenza di cammino ottico** determinata dalla posizione relativa delle diverse cariche elettriche
- Considerando una distribuzione di carica elettronica continua, passo da cariche discrete a elementi infinitesimi di carica
- Ogni elemento infinitesimo di carica si comporta come una singola carica elettrica
- Il passaggio da carica discreta a distribuzione continua di carica mette in evidenza come **l'onda diffusa risultante non sia altro che la Trasformata di Fourier della distribuzione continua di carica**

Diffusione da un insieme di atomi (molecola) - 1

Se consideriamo **una molecola**, o comunque **un insieme di atomi quale può essere considerata una cella unitaria di un cristallo**, possiamo definire un **fattore di scattering molecolare**.

Se abbiamo un atomo nell'origine, tale da vere una densità elettronica $\rho(r)$, l'atomo *jesimo* avrà una densità elettronica $\rho(r-r_j)$ dove r_j è la distanza dall'origine.
La densità elettronica della molecola può essere definita come somma delle densità elettroniche dei singoli atomi, ciascuno a distanza r_j dall'atomo nell'origine.

$$\rho_{mol}(r) = \sum_{j=1}^{N} \varrho(r - r_j)$$

Il fattore di scattering molecolare sarà quindi dato da:

$$\rho_{mol}(r) = \sum_{j=1}^{N} \varrho(r - r_j) \quad \Longrightarrow \quad \overrightarrow{F_M}(\overrightarrow{r^*}) = \sum_{j=1}^{N} \int_V \rho_{el}(\vec{r} - \vec{r}_j) \exp(2\pi i \vec{r} \cdot \overrightarrow{r^*})\, d\vec{r}$$

Diffusione da un insieme di atomi (molecola) - 2

$$\overrightarrow{F_M}(\overrightarrow{r^*}) = \sum_{j=1}^{N} \int_V \rho_{el}(\vec{r} - \vec{r}_j) \exp(2\pi i \vec{r} \cdot \overrightarrow{r^*})\, d\vec{r}$$

Ponendo $\mathbf{r}\text{-}\mathbf{r_j}=\mathbf{R_j}$

$$\overrightarrow{F_M}(\overrightarrow{r^*}) = \sum_{j=1}^{N} \int_V \rho_{el}(\overrightarrow{R_j}) \exp[2\pi i(\vec{r}_j + \vec{R}_j) \cdot \overrightarrow{r^*}\ d\overrightarrow{R_j}$$

Che posso riscrivere come:

$$\sum_{j=1}^{N} \left\{ \int_V \rho_{el}(\overrightarrow{R_j}) \exp(2\pi i \vec{R}_j \cdot \overrightarrow{r^*})\, d\overrightarrow{R_j} \right\} \exp(2\pi i \vec{r} \cdot \overrightarrow{r^*})$$

$\exp(2\pi i \vec{r} \cdot \overrightarrow{r^*})$ non dipende dalla variabile $\vec{R}_j$ e quindi vien posto al di fuori dell'intergrale

Diffusione da un insieme di atomi (molecola) - 3

$$\sum_{j=1}^{N} \left\{ \int_V \rho_{el}(\vec{R_j}) \exp(2\pi i \vec{R}_j \cdot \vec{r^*}) \, d\vec{R_j} \right\} \exp(2\pi i \vec{r} \cdot \vec{r^*})$$

Possiamo notare che:

$$\int_V \rho_{el}(\vec{R_j}) \exp(2\pi i \vec{R}_j \cdot \vec{r^*}) \, d\vec{R_j}$$

È uguale a $\boldsymbol{f_j^{at}}$

E quindi:

$$\boldsymbol{\vec{F_M}(\vec{r^*}) = \sum_{j=1}^{N} f_j^{at}(r^*) \, exp(2\pi i \vec{r} \cdot \vec{r^*})}$$

Fattore di struttura molecolare

$$\overrightarrow{F_M}(\overrightarrow{r^*}) = \sum_{j=1}^{N} f_j^{at}(r^*)\, exp\left(2\pi i \vec{r} \cdot \overrightarrow{r^*}\right)$$

Consideriamo l'insieme degli atomi componenti una molecola, Il fattore di struttura della molecola è dato dalla somma dei fattori di struttura dei singoli atomi ognuno moltiplicato per un termine che descrive la differenza di cammino ottico esistente tra onde diffuse da atomi diversi.

Le onde diffuse dai diversi atomi presenteranno delle differenze di fase dovute alla differenza di cammino ottico dipendente dalla geometria della molecola, cioè dalla posizione relativa dei singoli atomi tra loro.

Il concetto di molecola è molto generico e va inteso come 'insieme di atomi'.

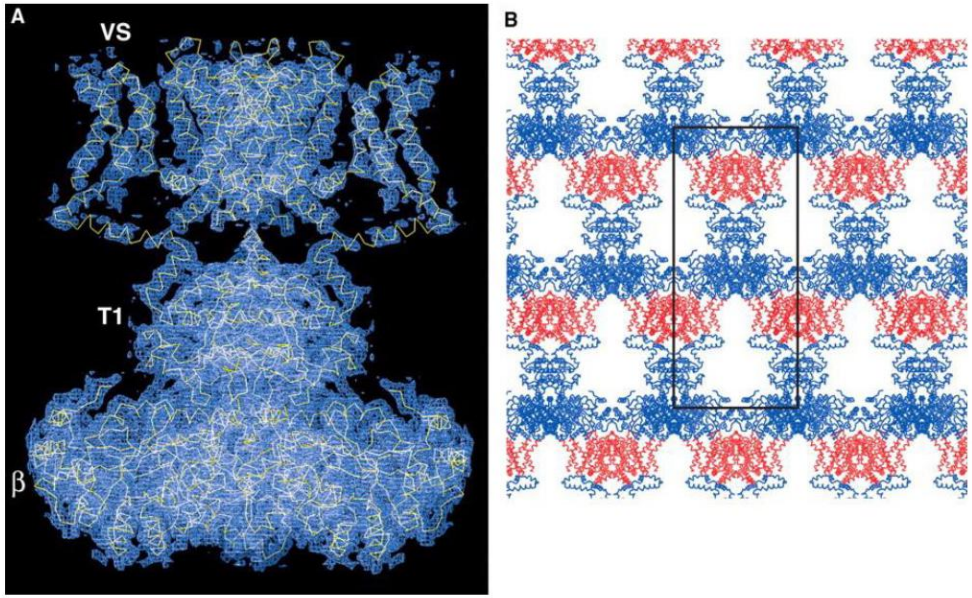
Nella cristallografia diffrattometrica ci riferiamo agli atomi contenuti nella cella unitaria di un sistema cristallino.

Densità elettronica di un cristallo

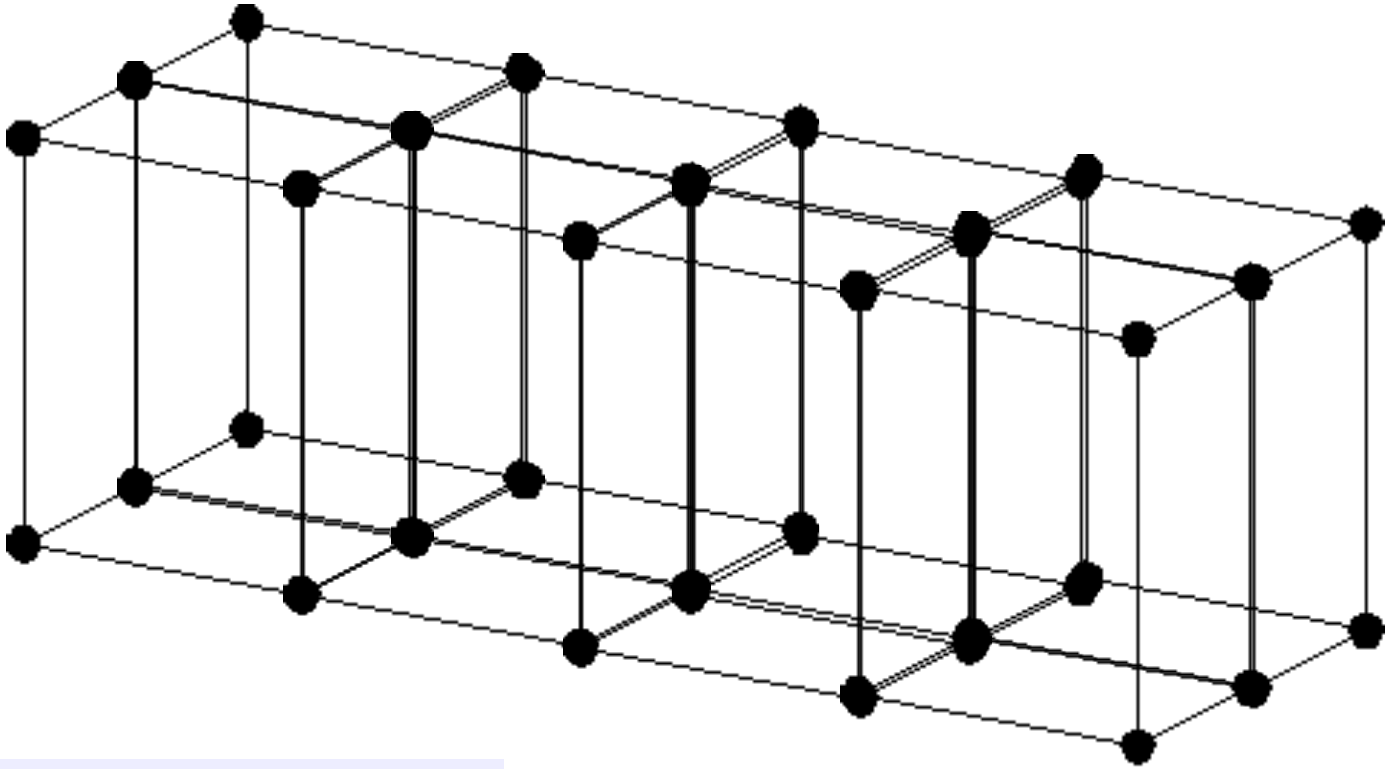
Un cristallo è dato dalla ripetizione nello spazio, secondo le relazioni di simmetria definite dal suo gruppo spaziale, di una parte non ulteriormente riducibile per simmetria (l'**unità asimmetrica**).

Un cristallo può anche essere descritto dalla ripetizione per traslazione lungo tre direzioni ($\vec{a}, \vec{b}, \vec{c}$) di una unità minima, la cella unitaria.

Volendo esprimere la densità elettronica in un cristallo, possiamo considerare la densità elettronica della cella unitaria, traslata lungo le direzioni opportune.



Crystals: unit cells

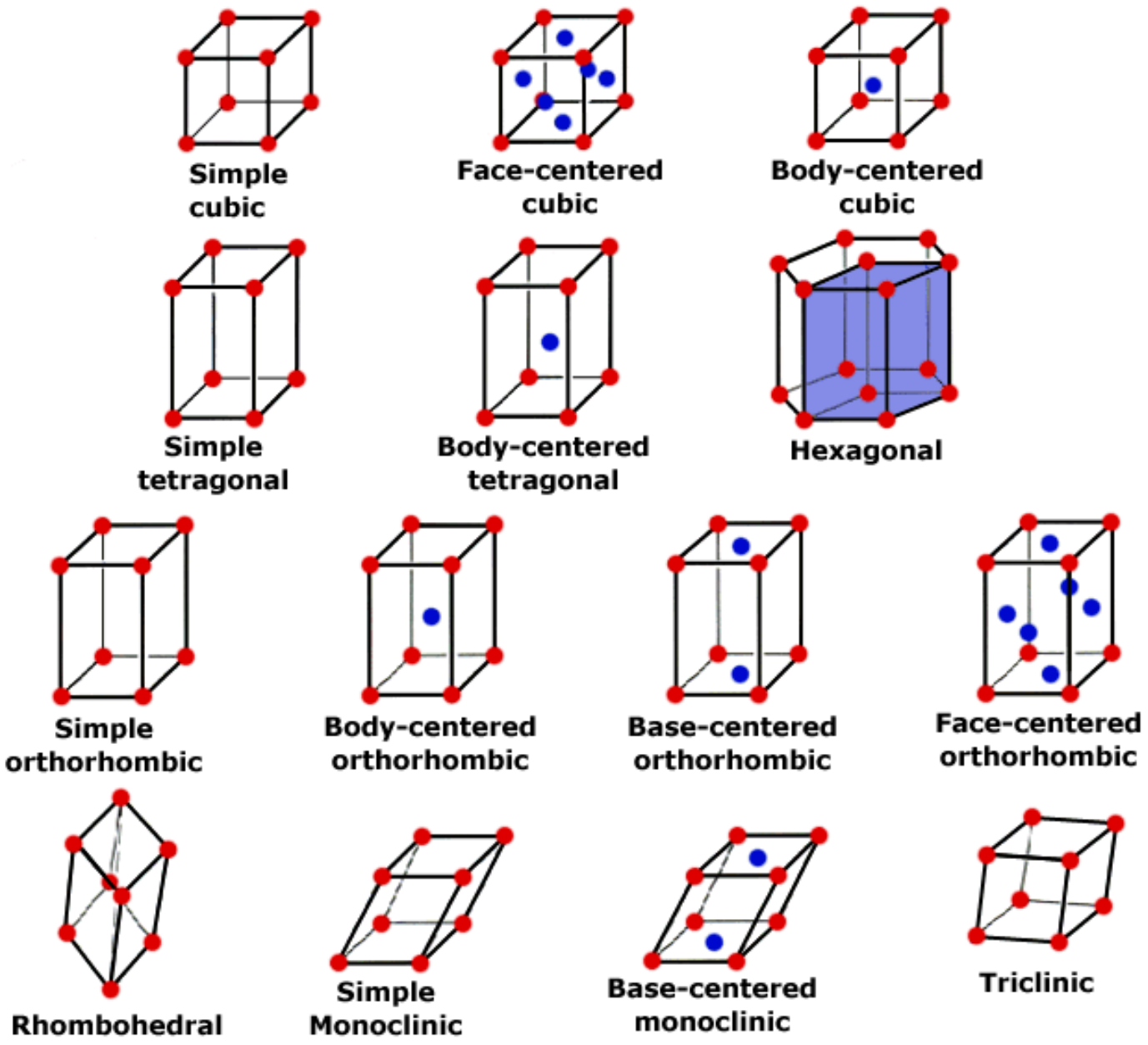


$$\vec{T} = u\vec{a} + v\vec{b} + w\vec{c}$$

u, v, w interi

$\vec{T}$ is a vector of the Bravais Lattice

The 14 Bravais Lattices

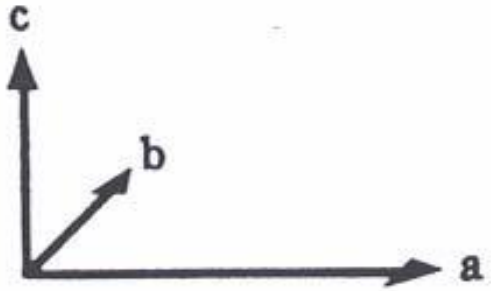


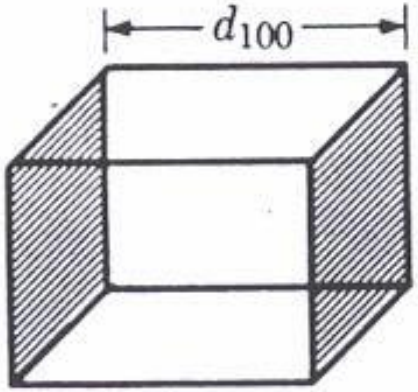
The space groups

Table 1.2 *Symmetry elements*

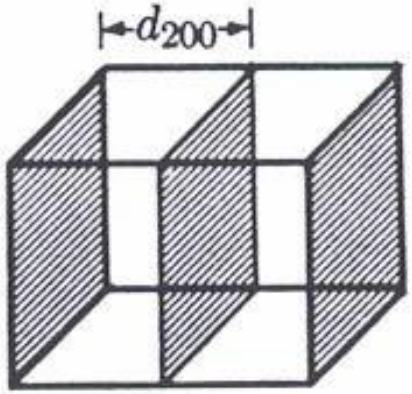
	Symmetry element	Hermann–Mauguin symbols (crystallography)	Schönflies symbols (spectroscopy)
Point symmetry	Mirror plane	m	σ_v, σ_h
	Rotation axis	$n = 2, 3, 4, 6$	$C_n (C_2, C_3,$ etc.)
	Inversion axis	$\bar{n} (= 1, 2,$ etc.)	—
	Alternating axis*	—	$S_n (S_1, S_2,$ etc.),
	Centre of symmetry	$\bar{1}$	i
Space symmetry	Glide plane	a, b, c, d, n	—
	Screw axis	$2_1, 3_1,$ etc.	—

* The *alternating axis* is a combination of rotation (n-fold) and reflection perpendicular to the rotation axis. It is little used in crystallography.

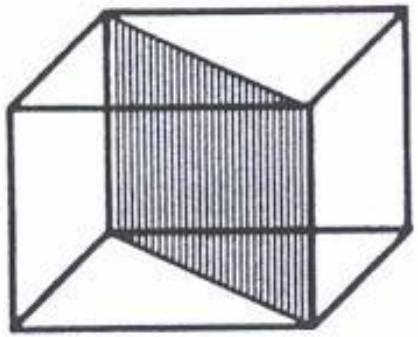




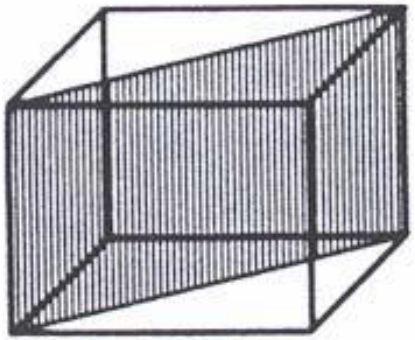
(100)



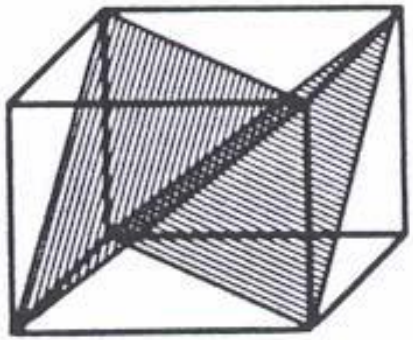
(200)



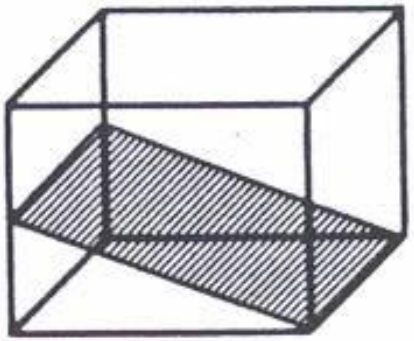
(110)



($\bar{1}10$)



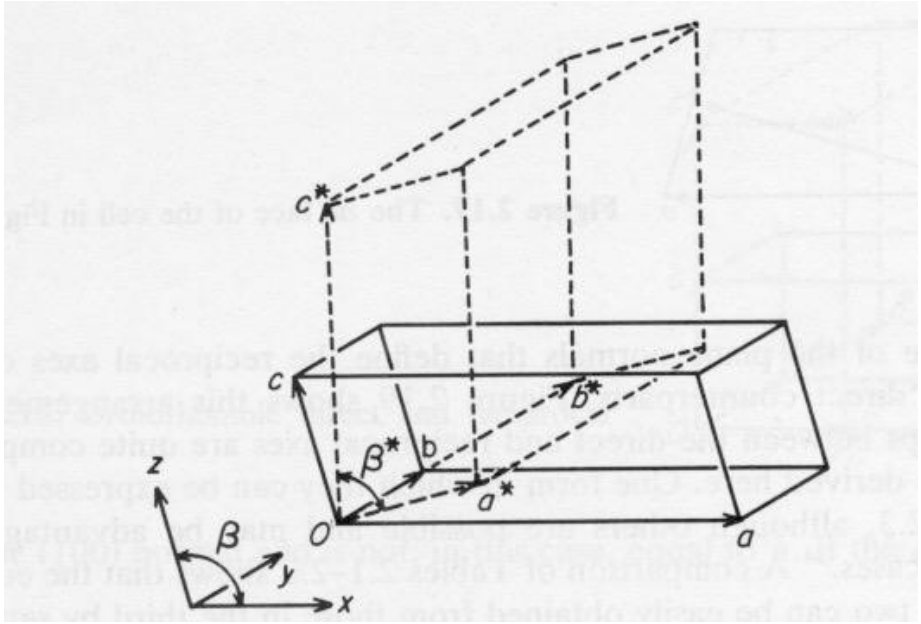
($11\bar{1}$)



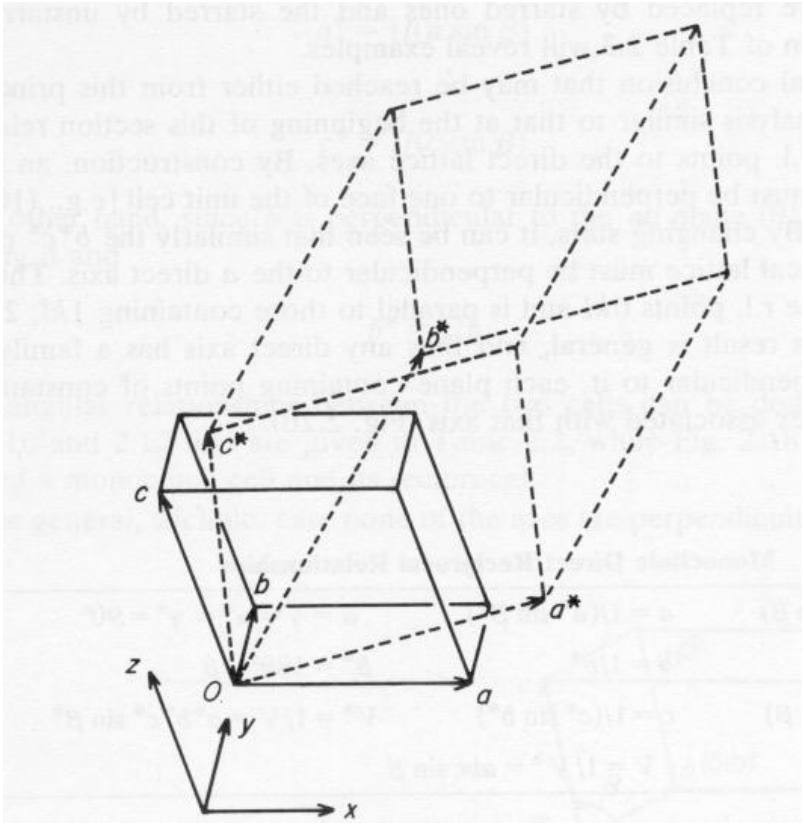
(102)

Examples of direct and reciprocal lattices

(Figures from Jensen and Stout "*X-Ray structure determination. A pratical guide*"

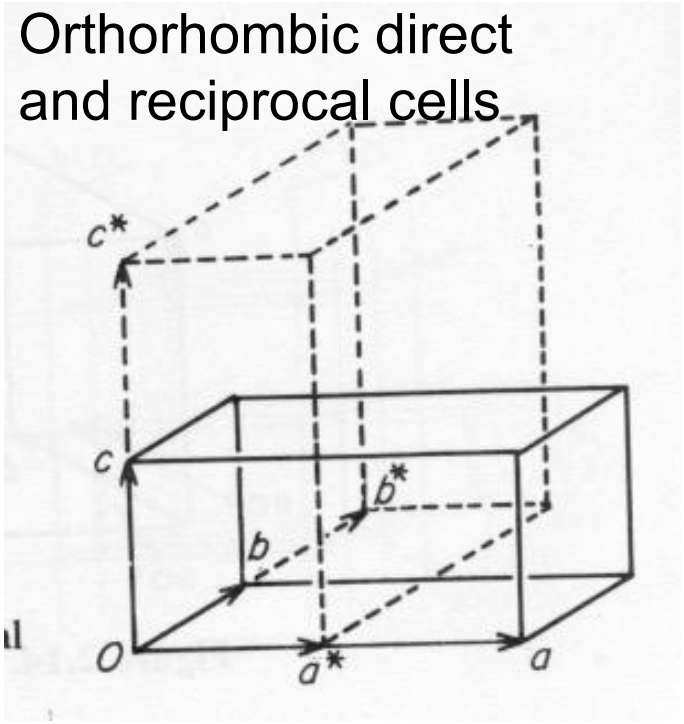


Monoclinic direct and reciprocal ce



Triclinic direct and reciprocal cells

Orthorhombic direct and reciprocal cells

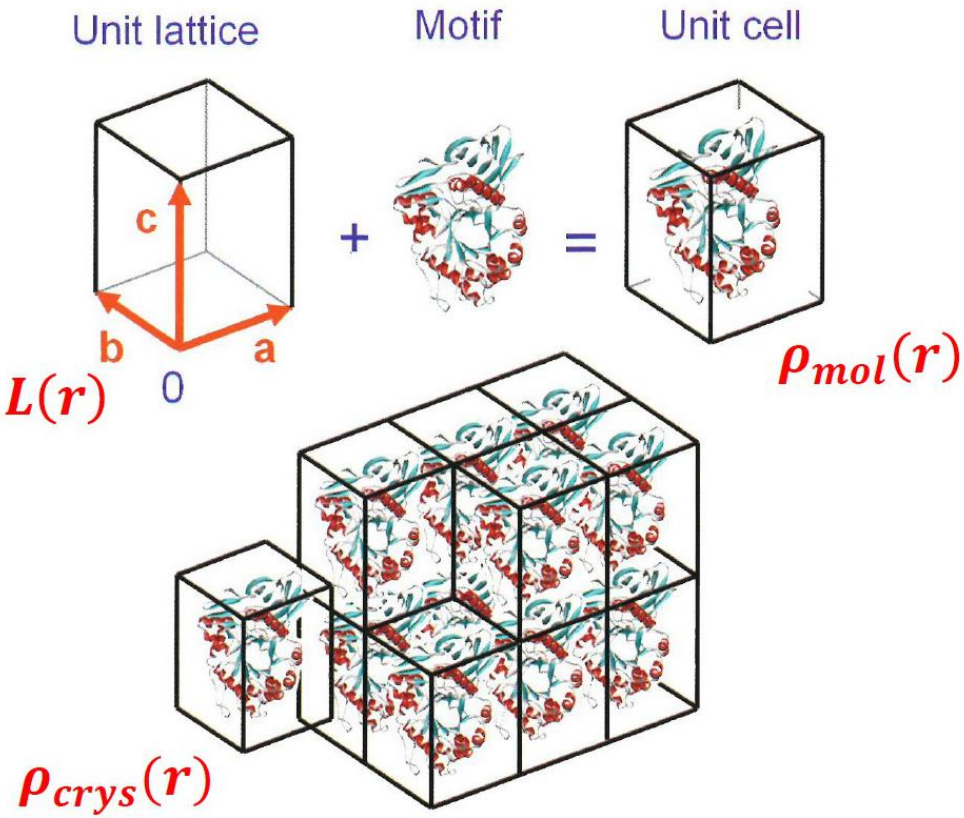


Descrizione della densità elettronica in un cristallo

Matematicamente questo si può ottenere operando la **convoluzione** tra la funzione che esprime la densità elettronica nella cella unitaria $\rho_{mol}(r)$ e una funzione matematica che descrive il reticolo cristallino, ovvero la sua ripetizione nello spazio, $L(r)$:

$$\rho_{crys}(r) = \rho_{mol}(r) * L(r)$$

$\rho_{crys}(r)$ è la densità elettronica nel cristallo



Scattering da un cristallo - 1

Come descritto in precedenza, le onde diffuse dagli elettroni e gli elettroni stessi, o meglio la loro densità elettronica sono legati dall'operazione matematica della Trasformata di Fourier.

$$\overrightarrow{F_{crys}}(\overrightarrow{r^*}) = TF[\rho_{crys}(r)]$$

Alla luce delle espressioni precedenti avremo che:

$$TF[\rho_{crys}(r)] = TF[\rho_{mol}(r) * L(r)]$$

Per la nota proprietà delle convoluzioni e della trasformata di Fourier, quest'ultima relazione può essere scritta come:

$$TF[\rho_{mol}(r) * L(r)] = TF[\rho_{mol}(r)] \cdot TF[L(r)]$$

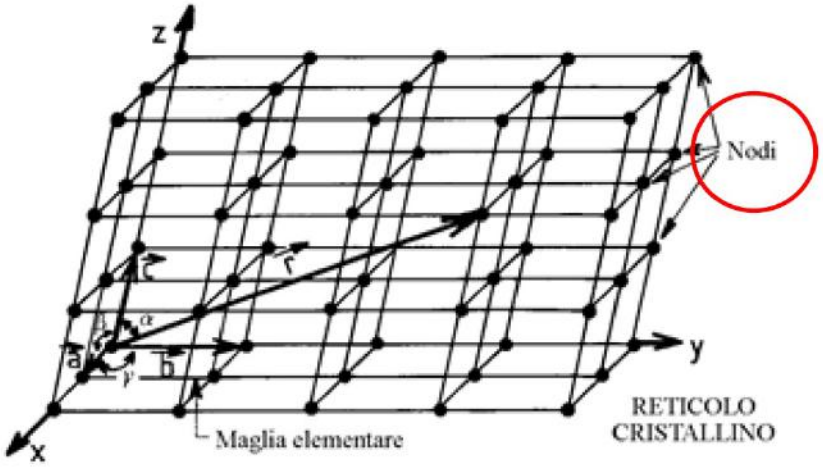
Diffrazione da un cristallo - 1

$$\overrightarrow{F_{crys}}(\overrightarrow{r^*}) = TF[\rho_{mol}(r)] \cdot TF[L(r)]$$

$TF[\rho_{mol}(r)]$ è noto ed è uguale al fattore di struttura della molecola/cella unitaria

$TF[L(r)]$ dipende da come è 'fatta' la funzione $L(r)$, la funzione reticolo.

La funzione reticolo $L(r)$ descrive la struttura del reticolo e quindi $L(r)$ sarà diversa da zero solo in corrispondenza dei **nodi del reticolo**.



La Trasformata di Fourier della funzione reticolo, TF[L(r)], si comporta allo stesso modo della L(r), sarà diversa da zero ed avrà **valore 1 solo in corrispondenza dei nodi del reticolo reciproco** (sono passato dallo spazio reale a quello reciproco e quindi dal reticolo reale a quello reciproco, tramite la trasformata di Fourier)

Scattering from a 3D crystal

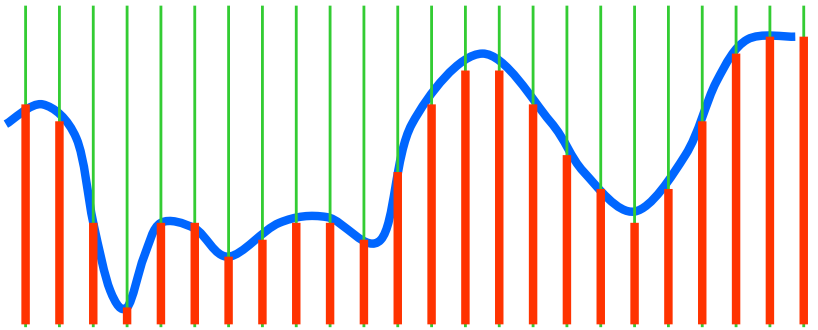
$$F_{3D}(s) = F(s) \cdot \left[\frac{\sin(2N+1)\pi(a\cdot s)}{\sin \pi(a\cdot s)}\right]\left[\frac{\sin(2N+1)\pi(b\cdot s)}{\sin \pi(b\cdot s)}\right]\left[\frac{\sin(2N+1)\pi(c\cdot s)}{\sin \pi(c\cdot s)}\right]$$

scattering from one molecule

3D fringe function that depends on the lattice spacing **a, b, c**

The fringe function has two effects:
- makes the pattern “discrete”
- amplifies the signal

The diffraction pattern from a crystal is the diffraction from the molecule **sampled and amplified** according to the fringe function



Solving macromolecular structures by X-ray diffraction

- Now assume that we have more than one atom in a unit cell→ the diffraction angle θ will remain the same!! The diffraction angle is still related to the unit cell dimension by Bragg's law.

- **However**, with one atom, the phase and amplitude of the resultant wave from each plane (atom) was the same. With more than one atom, the phase and amplitude coming from each plane of different atoms may be different, resulting in **different intensities** of the reflections.
- → Deconvolute each reflection into the phase and amplitude contributions from each of the individual reflections from each atom in the molecule.

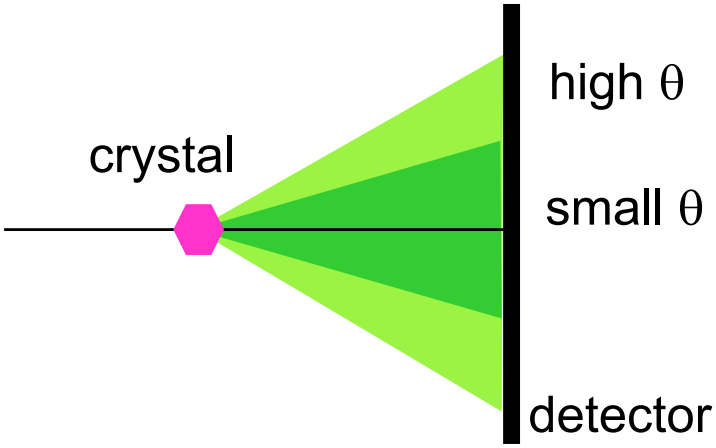
# of atoms in unit cell does NOT affect reflection angle (Bragg's law).

# of atoms (electron density) in unit cell DOES affect **intensity** of reflection spots.

The electron density equation: problem # 1:resolution limits

Theoretical limit

There are theoretical limits that depends on the wavelength of the X-ray radiation. Typically one uses a wavelength of roughly 1 A, which would allow for resolution of 0.5 A.



Practical limit

In practice for macromolecules the resolution limit is usually set by the intrinsic degree of order of the crystal typically one sees diffraction to 2.0-3.5 A – this is even more true of membrane protein crystals which tend to be more disordered due to the less directional nature of hydrophobic interactions.

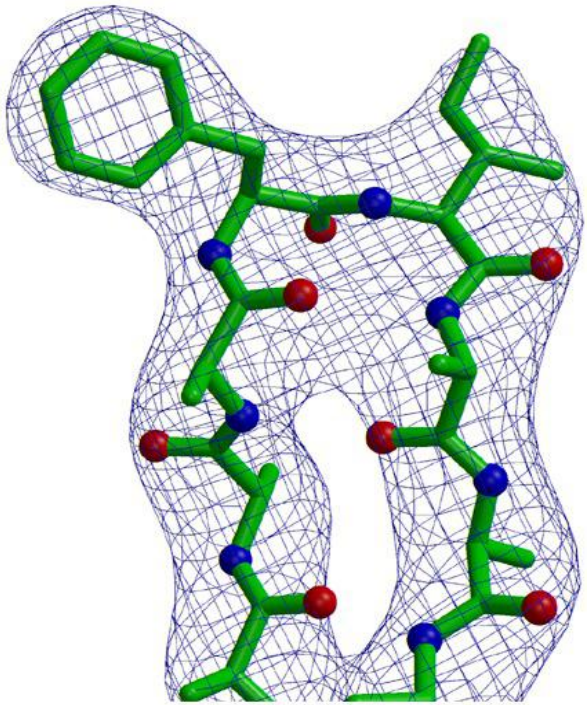
Why is resolution important?

"Resolution" is related to the level of details that can be visualised: a high resolution map gives a far more detailed picture of the atomic structure:

low resolution → **high resolution**

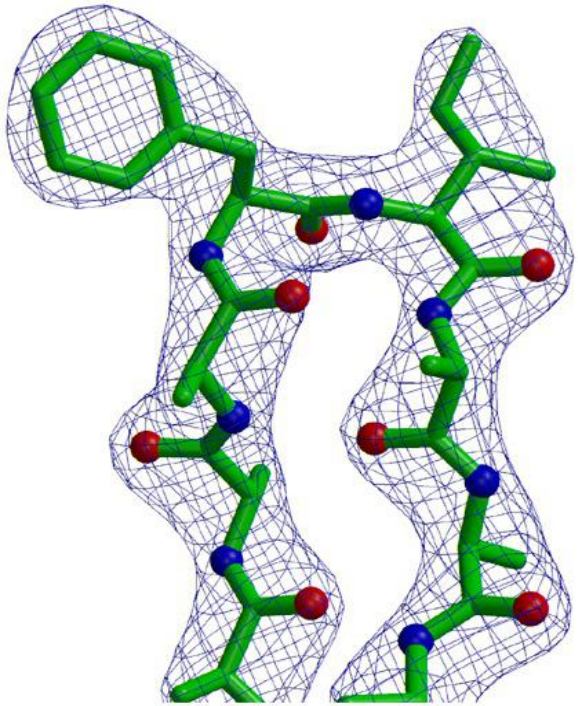
low s, small θ, large d

high s, large θ, small d

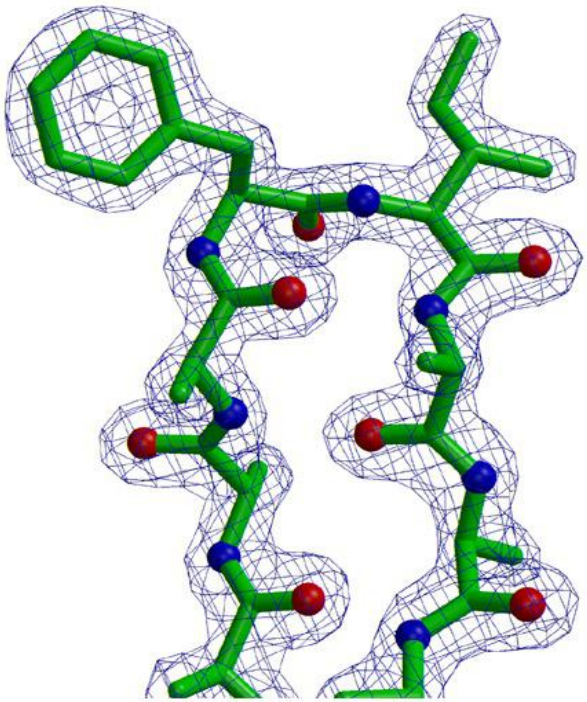


Here is a 4 Å map

things which are less than 4Å apart tend to be blurred



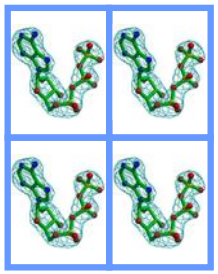
Here is a 2.9 Å map



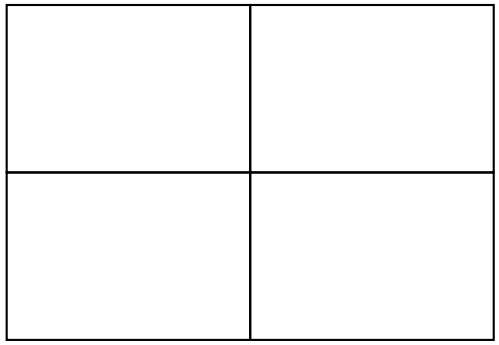
Here is a 2 Å map

almost atomic detail

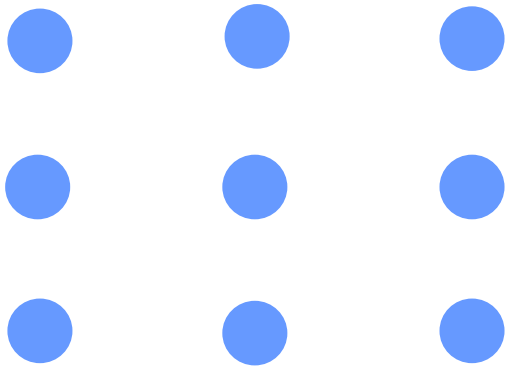
Small molecule crystal



small unit cell

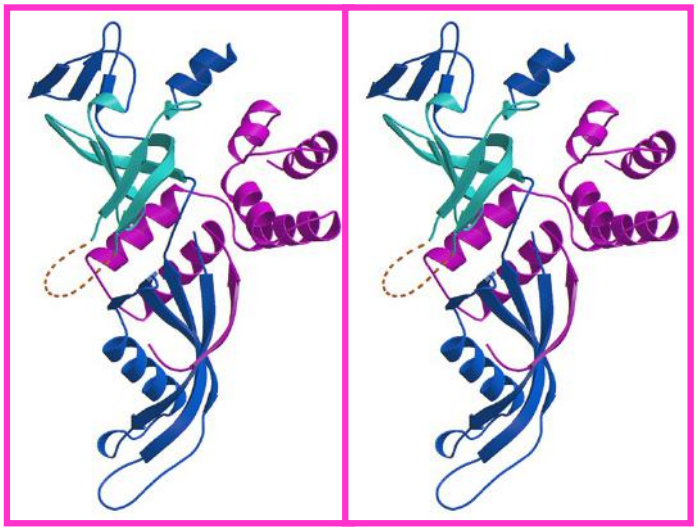


large “reciprocal lattice”
(coarse fringe function)

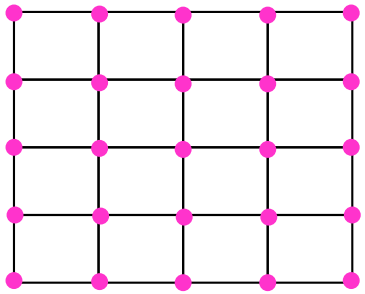


a few, very strong reflections;
diffract to high resolution

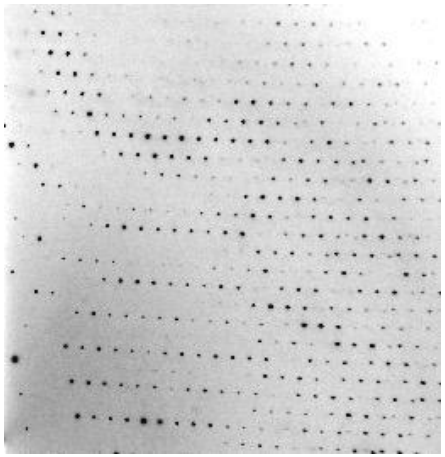
Protein crystal



large unit cell



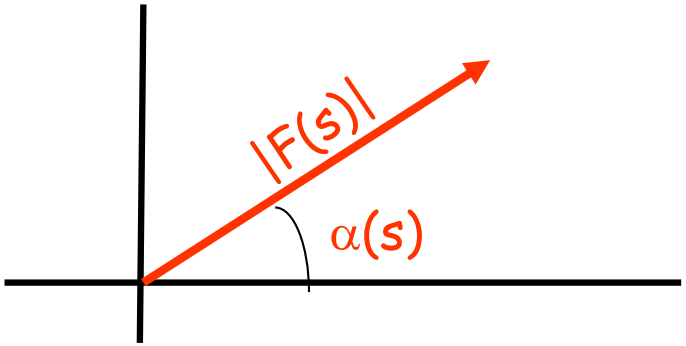
small “reciprocal lattice”
(fine fringe function)

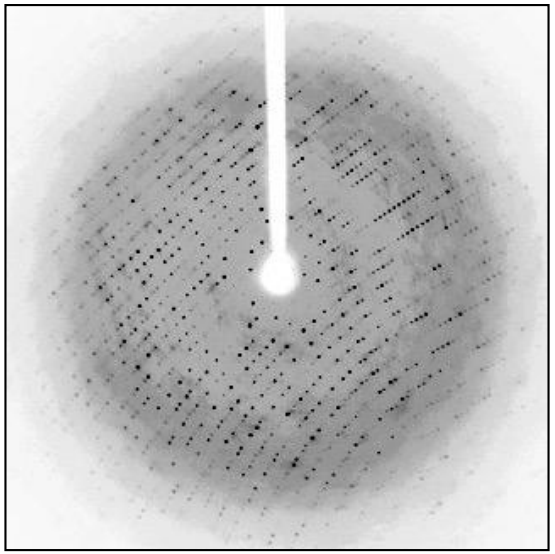


many, very weak reflections;
diffract to low resolution

The electron density equation: problem # 2:the phase problem

- **F**(**s**) is a complex number with modulus |F(**s**)| and phase α





In a diffraction experiment we measure the intensity of each spot I(**s**)

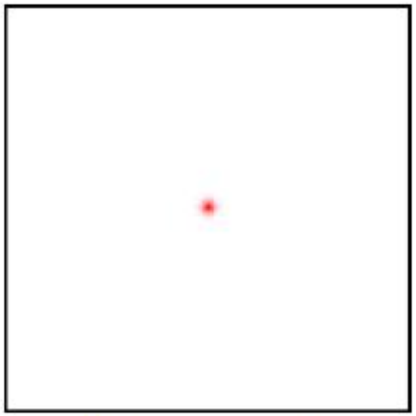
$I(\mathbf{s}) = |F(\mathbf{s})|^2$

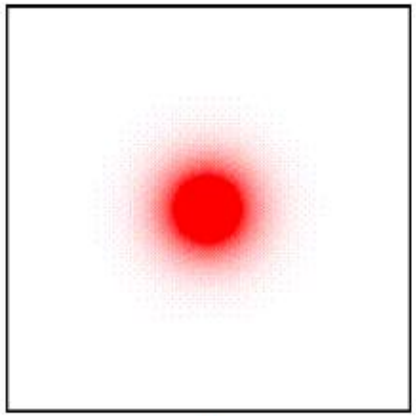
We can derive the amplitude F(**s**) but we have lost the information about the relative phase

the phase problem!!

Universal problem in crystallography – also for small molecules.

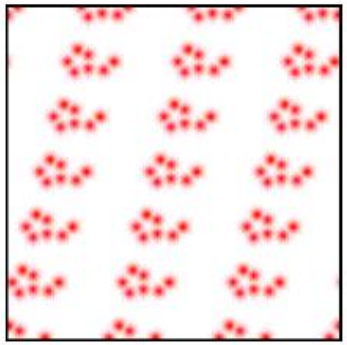
- An atom, and its Fourier Transform:





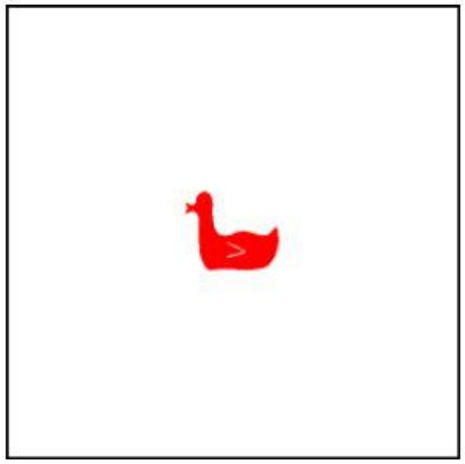
- Note the both functions have circular symmetry. The atom is a sharp feature, whereas its transform is a broad smooth function. This illustrates the reciprocal relationship between a function and its Fourier transform.

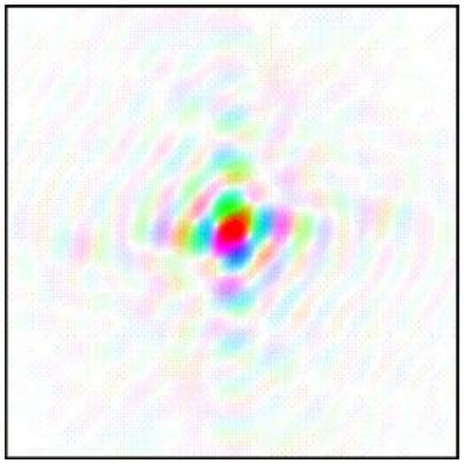
- A crystal, and its Fourier Transform:



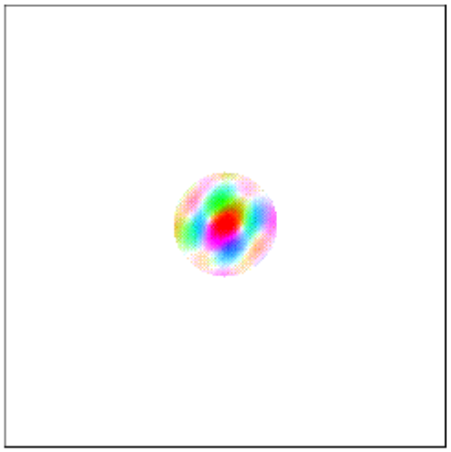
- Finally, we build up a crystal by convoluting the *molecule* with the *grid*. The result is a crystal structure. The Fourier transform of the crystal is thus the product of the *molecular transform* and the *reciprocal lattice*. This is the *diffraction pattern*.

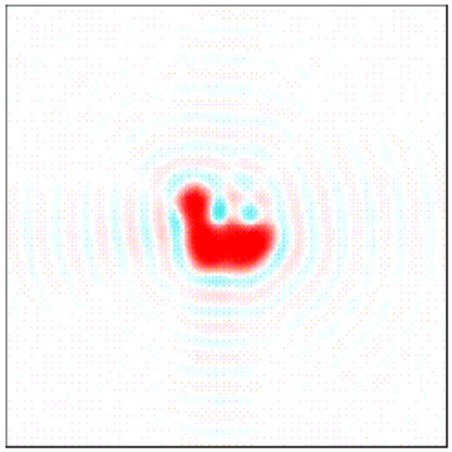
- A duck and its Fourier Transform





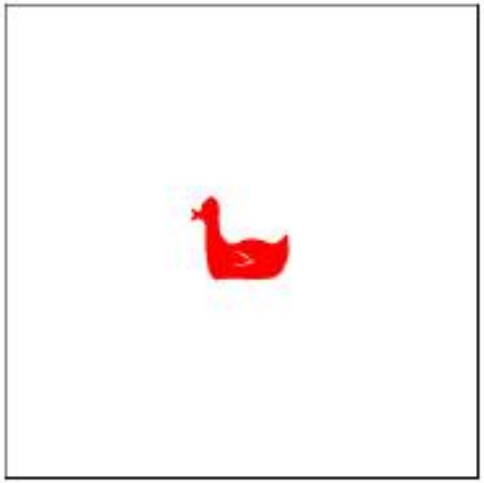
- If we only have the low resolution terms of the diffraction pattern, we only get a low resolution duck:

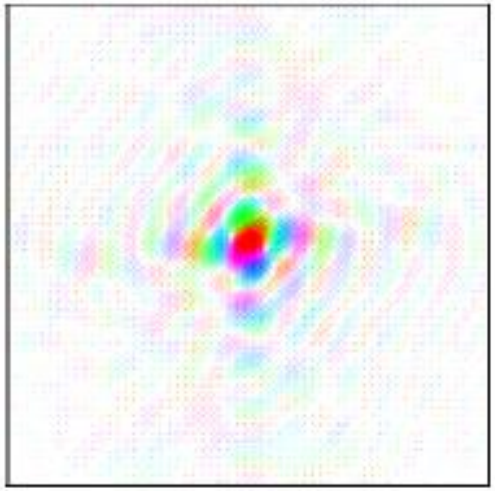




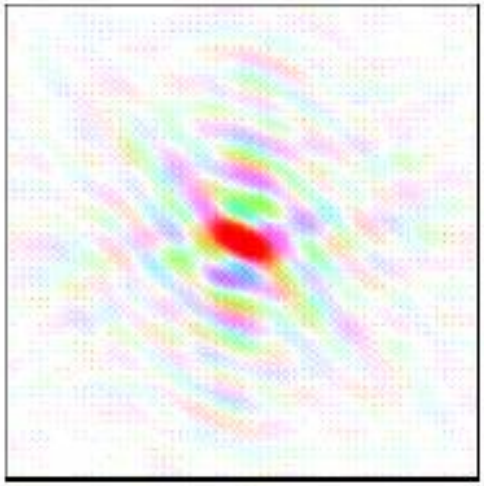
- **Crystallographic Interpretation:**
- There is considerable loss of detail. At low resolution, your atomic model may reflect more what you expect to see than what is actually there.

- Here is our old friend; the Fourier Duck, and his Fourier transform:

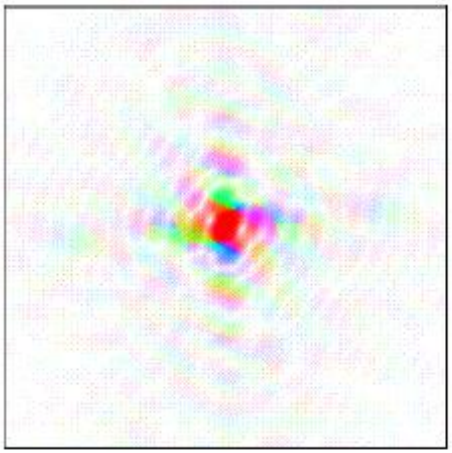


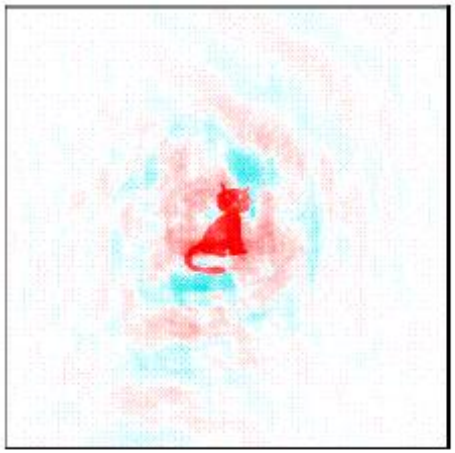


- And here is a new friend; the Fourier Cat and *his* Fourier transform:

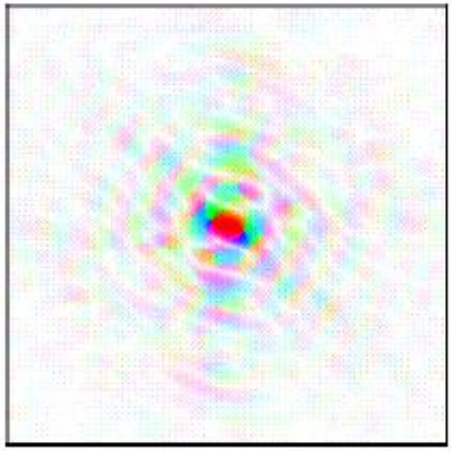


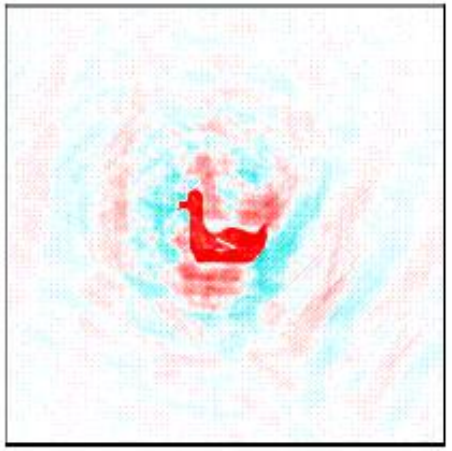
- Now we will mix them up. Let us combine the the magnitudes from the Duck transform with the phases from the Cat transform. (You can see the brightness from the duck and the colours from the cat). If we then transform the mixture, we get the following:





- We can do the same thing the other way round. Using the magnitudes from the Cat transform and the phases from the Duck transform, we get:



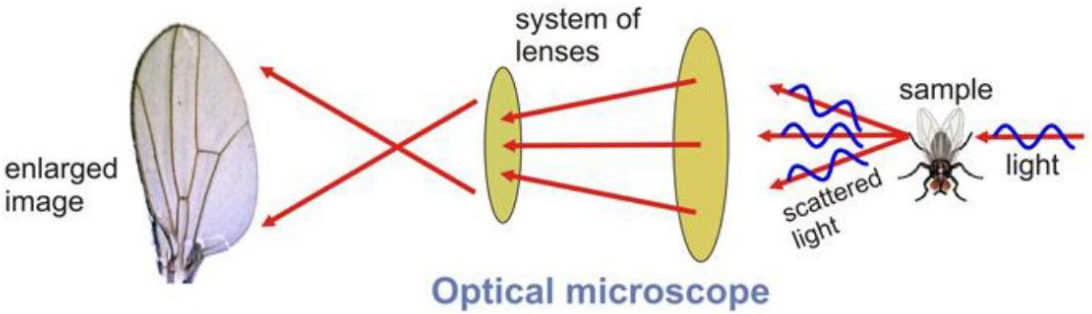


Il problema della fase

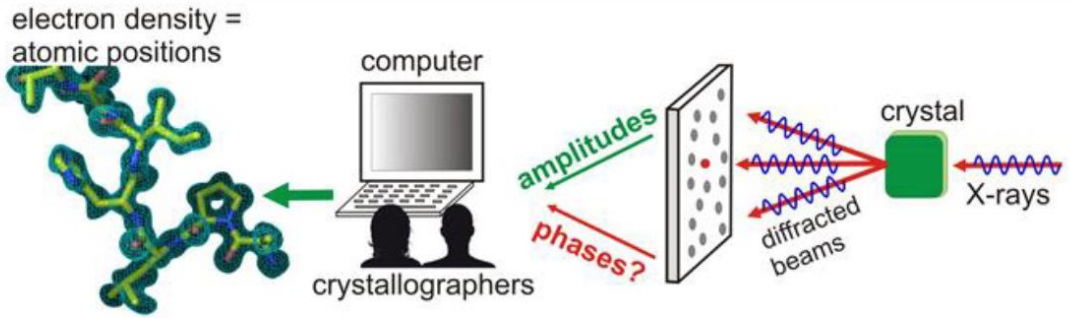
Per ricostruire l'immagine (la struttura molecolare nel cristallo) abbiamo bisogno di **Ampiezze** e **Fasi** delle onde diffratte dal cristallo.

In cristallografia non abbiamo un sistema di lenti.

L'informazione di fase va 'recuperata' in qualche altro modo.



Optical microscope



"Impossible" X-ray microscope

Il problema della Fase in cristallografia

Un esperimento di diffrazione permette di ottenere le **intensità** e quindi il **modulo** dei fattori di struttura $|F_{hkl}|$

$$\rho_{cell}(r) = \frac{1}{V} \sum_{h,k,l=-\infty}^{\infty} \boldsymbol{F_H} \exp[-2\pi i\,(hx + ky + lz)]$$

Per la ricostruzione della densità elettronica dobbiamo conoscere i fattori di struttura in **modulo** e **fase**

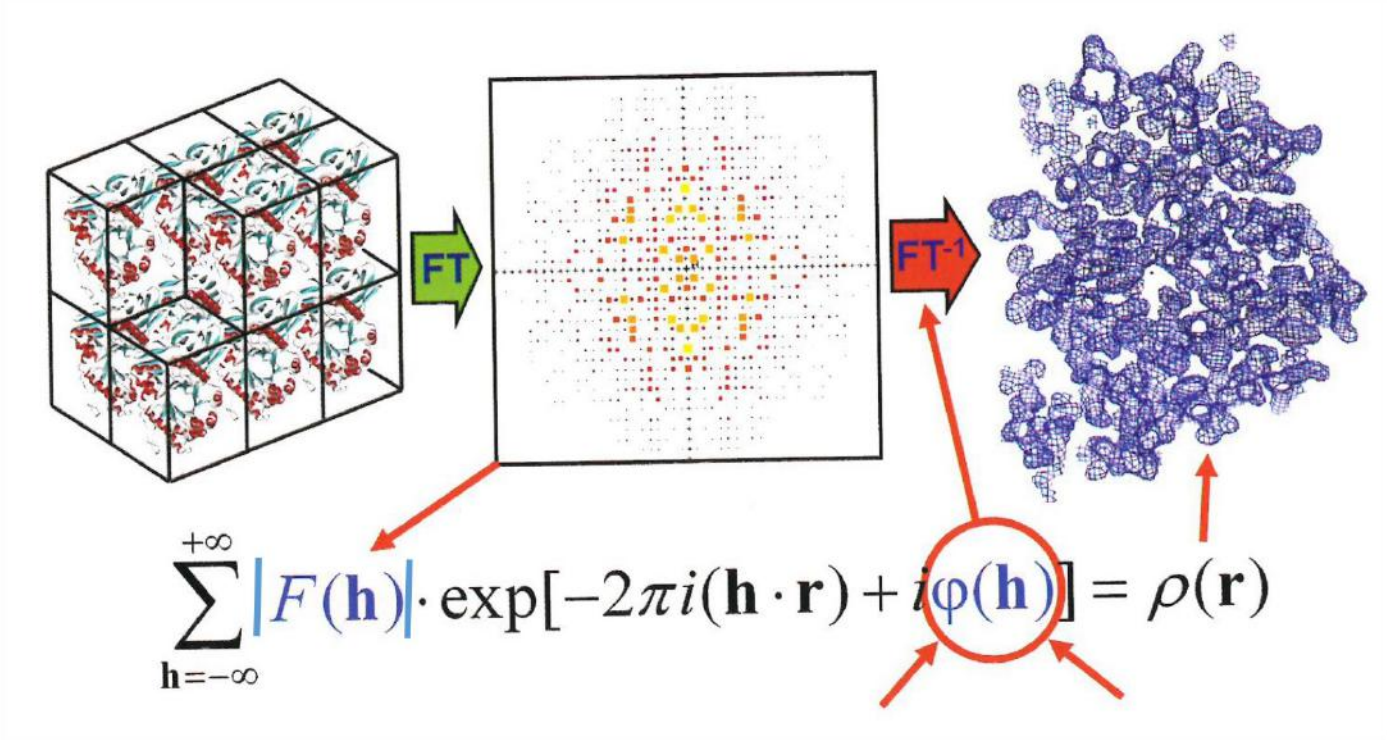
$$\rho_{cell}(r) = \frac{1}{V} \sum_{h,k,l=-\infty}^{\infty} |\boldsymbol{F_{hkl}}| \cdot \exp[-2\pi i\,(hx + ky + lz) + i\boldsymbol{\varphi_{hkl}}]$$

Determinato dall'esperimento

Non è nota!

L'assenza di informazione sulle fasi delle onde diffratte costituisce il **problema della fase** in cristallografia

Problema della fase



Fase di un fattore di struttura

$$F_{hkl} = \sum_{j=1}^{N} f_j(r^*)\, exp\left(2\pi i \vec{r_j} \cdot \overrightarrow{r^*_{h,k,l}}\right)$$

Il fattore di struttura F_{hkl} può essere visto come la somma di tanti contributi atomici, ognuno con una sua fase definita dalle coordinate dell'atomo, all'interno della cella unitaria.

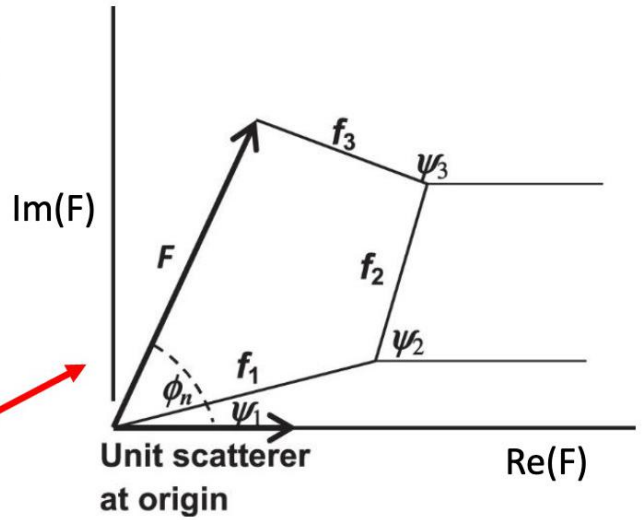
$$F_{hkl} = f_1\, exp[2\pi i(hx_1 + ky_1 + lz_1)] + \ \ldots + f_N\, exp[2\pi i(hx_N + ky_N + lz_N)]$$

Che può essere anche scritto come

$$F_{hkl} = |F_{hkl}|exp(i\phi_{hkl}) = f_1\, exp(i\psi_1) + \ \ldots + f_N\, exp(i\psi_N)$$

Che per tre atomi può essere scritto (e rappresentato sul piano di Argand)

$$F_{hkl} = |F_{hkl}|exp(i\phi_{hkl}) = f_1\, exp(i\psi_1) + f_2\, exp(i\psi_2) + f_3\, exp(i\psi_3)$$

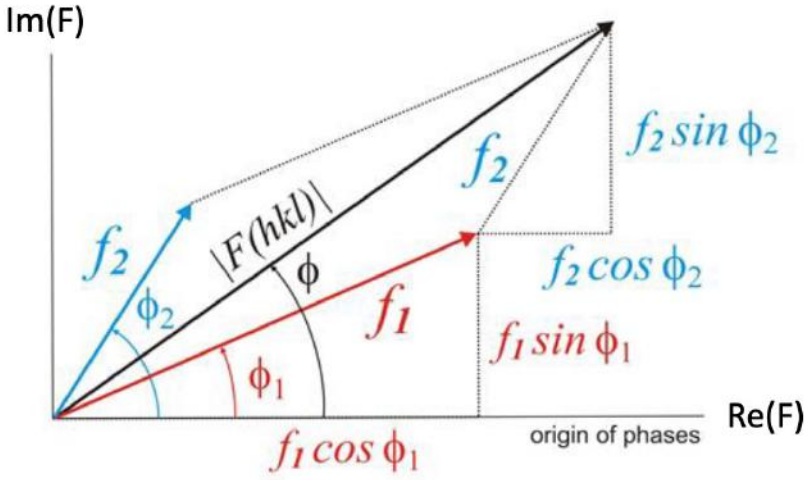


Fase di un fattore di struttura

$$F_H = |F_{hkl}|exp(i\phi_{hkl}) = \sum_{j=1}^{N} f_j(r^*)\cos[2\pi(hx_j + ky_j + lz_j)] + i\sum_{j=1}^{N} f_j(r^*)sin[2\pi(hx_j + ky_j + lz_j)]$$

La fase di un fattore di struttura è quindi determinata dalla posizione degli atomi nella cella unitaria.

Conoscendo le posizioni degli atomi nella cella unitaria e il loro tipo (C, N, Fe, …), il fattore di struttura è completamente determinato



Intensità e struttura cristallina

$$I_{hkl} = |F_{hkl}|^2 = [\sum_{i=1}^{N} f_j(r^*)\, exp\left(2\pi i \overrightarrow{r_j} \cdot \overrightarrow{r^*_{h,k,l}}\right)] \cdot [\sum_{j=1}^{N} f_j(r^*)\, exp\left(-2\pi i \overrightarrow{r_j} \cdot \overrightarrow{r^*_{h,k,l}}\right)]$$

Da cui si ottiene:

$$|F_{hkl}|^2 = F_{hkl} \cdot {F_{hkl}}^*$$

$$|F_{hkl}|^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} f_i f_j exp[2\pi i r^*_{hkl}(r_i - r_j)]$$

Le intensità diffratte contengono una informazione sulla struttura cristallina, in particolare**, le intensità diffratte dipendono dalle distanze tra gli atomi nella cella unitaria** (vettori interatomici)

Assioma fondamentale per la soluzione del problema della fase

Per una qualsiasi struttura cristallina il problema della fase è risolto quando sono determinate le fasi ϕ_{hkl} di tutti i fattori di struttura F_{hkl}, ovvero quando sono note le posizioni degli atomi nella cella unitaria (cella elementare).

La soluzione dell'equazione che lega intensità e posizioni atomiche **non è necessariamente univoca**. In teoria esiste più di una soluzione che lega intensità a posizione atomica (l'intensità dipende da vettori interatomici, non dalle posizioni 'assolute' dei singoli atomi!).

Tuttavia nella risoluzione del problema della fase si ammette e si considera sempre valido il presente assioma:

Esiste una sola soluzione al problema della fase compatibile con una struttura molecolare chimicamente corretta

Questo vuol dire che esiste una sola soluzione che permetta di ricostruire delle molecole dotate di senso chimico. Il 'senso' chimico della struttura ottenuta (lunghezze, angoli di legame, distanze interatomiche, stereochimica), ci permette di distinguere tra una soluzione corretta e una sbagliata.

La risoluzione del problema della fase

Il problema della fase è stato chiaro fin dagli inizi della cristallografia
Nel corso degli anni sono stati sviluppati metodi teorici e sperimentali per affrontare questo problema.

Inizialmente si procedeva 'trial and error', un approccio possibile con strutture semplici, ma sicuramente non adeguato a strutture (relativamente) complesse.

Metodi Sperimentali:
Richiedono ulteriori esperimenti (di diffrazione).

Metodi Computazionali:
Non richiedono ulteriori esperimenti di diffrazione, ma eventualmente delle informazioni aggiuntive

Metodi Sperimentali

- Single/**Multiple Isomorphous Replacement** (SIR/MIR)
- Multiwavelength Anomalous Diffraction (MAD)
- **Single Wavelength Anomalous Diffraction** (SAD)

Metodi Computazionali

- **Molecular Replacement**
- Metodi Diretti

How to solve macromolecular structures:

MIR (multiple isomorphous replacement)

Older method (Cambridge, 60') – relies on binding "heavy" atoms to the crystal and compare he diffraction pattern to the native. Trial and error search for good heavy atoms, it may take longer to get it right

MR (molecular replacement)

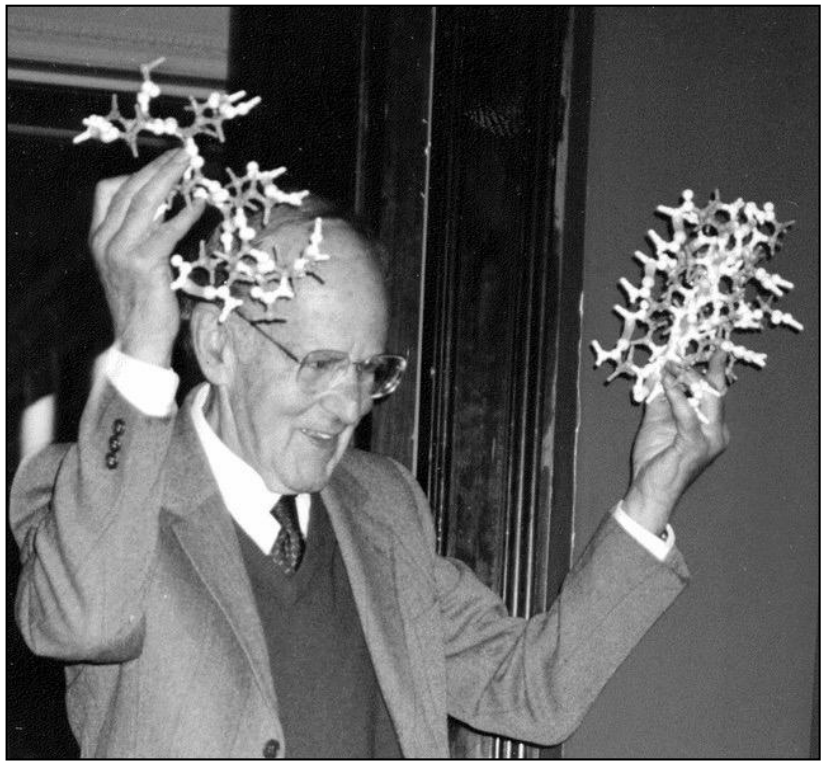
Older method (Cambridge, 70'-80') – relies on the expected similarity between the protein and another whose structure is known. Cannot solve de novo structures. Requires high homology (30% sequence identity?)

MAD (multiwavelength anomalous dispersion)

Relies on the absorption of specific wavelengths due to electronic transitions within the atom core. Similar to MIR but generally far quicker and more accurate. Requires high specification synchrotron radiation.

Some of the players...

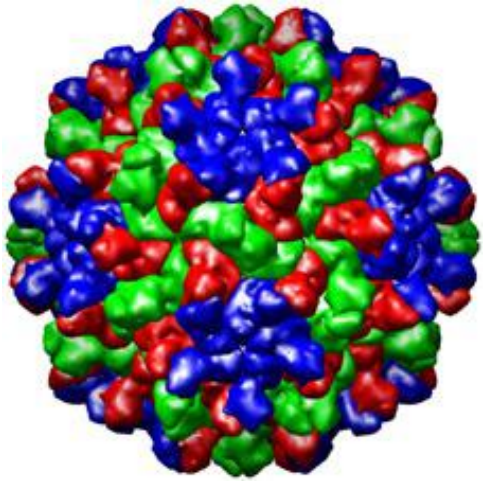
Max Perutz (1914-2002), the inventor of MIR



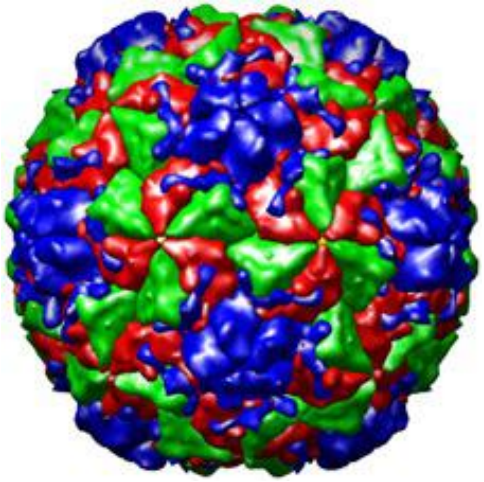
In 1953, Perutz showed that the diffracted X-rays from protein crystals could be phased by comparing the patterns from crystals of the protein with and without heavy atoms attached.
In 1959, he employed this method to determine the molecular structure of hemoglobin. This work resulted in his sharing with John Kendrew the 1962 Nobel Prize for Chemistry.

Molecular Replacement

The Molecular replacement method was mostly developed by Michael Rossmann. He used the structure of Tomato Bushy Stunt Virus, a plant virus, to determine the crystal structure of the Human Rhinovirus 14 (the common cold virus) in the early 80's.



Tomato Bushy Stunt Virus



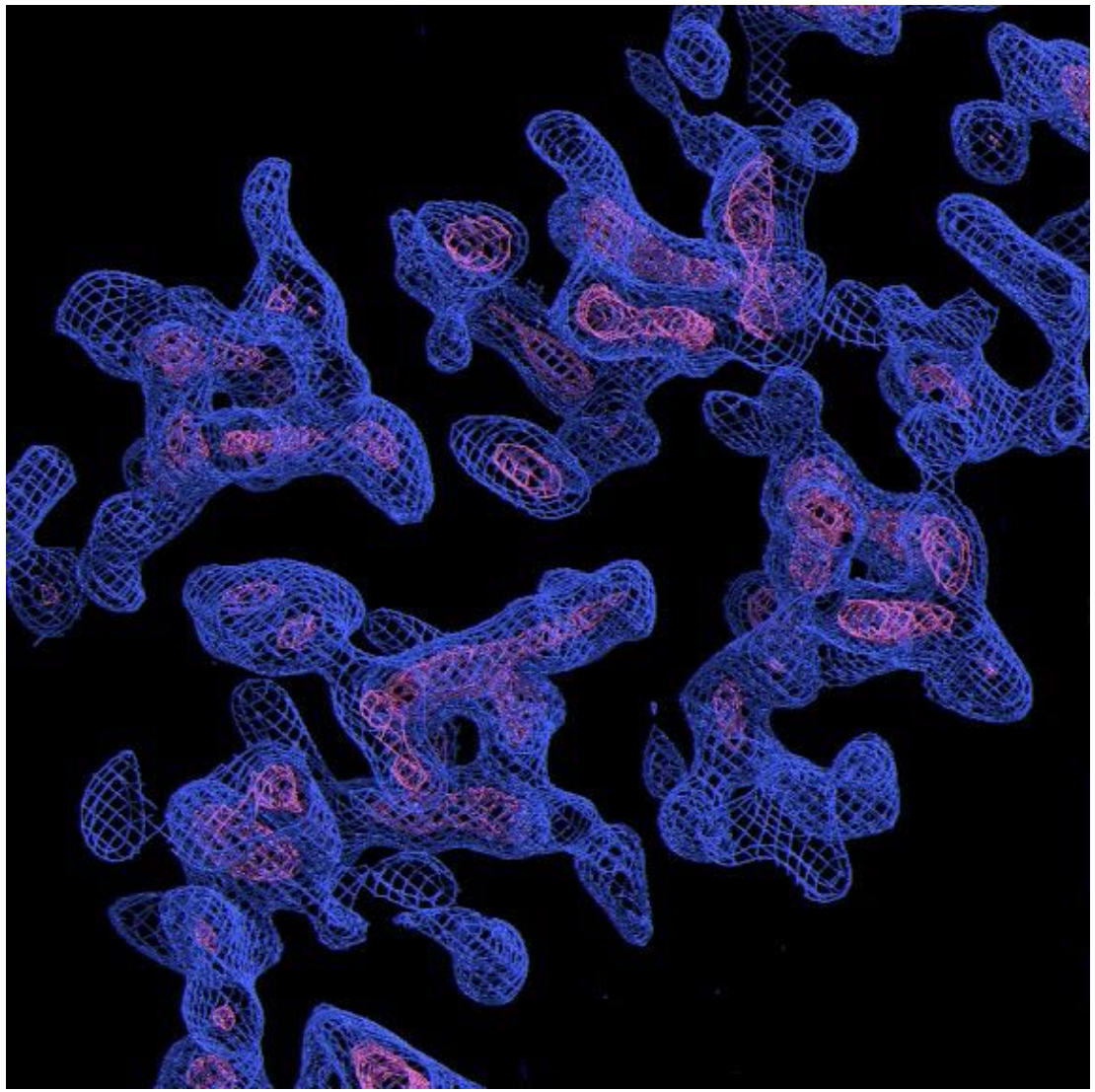
Human Rhinovirus 14

However the theoretical basis were developed much earlier:

Rossman, M. G. and Blow, D. M. (1962). *ActaCryst*. 15:24-31.

Electron density maps

After all this effort, we have a 3D map showing the shape of the protein:

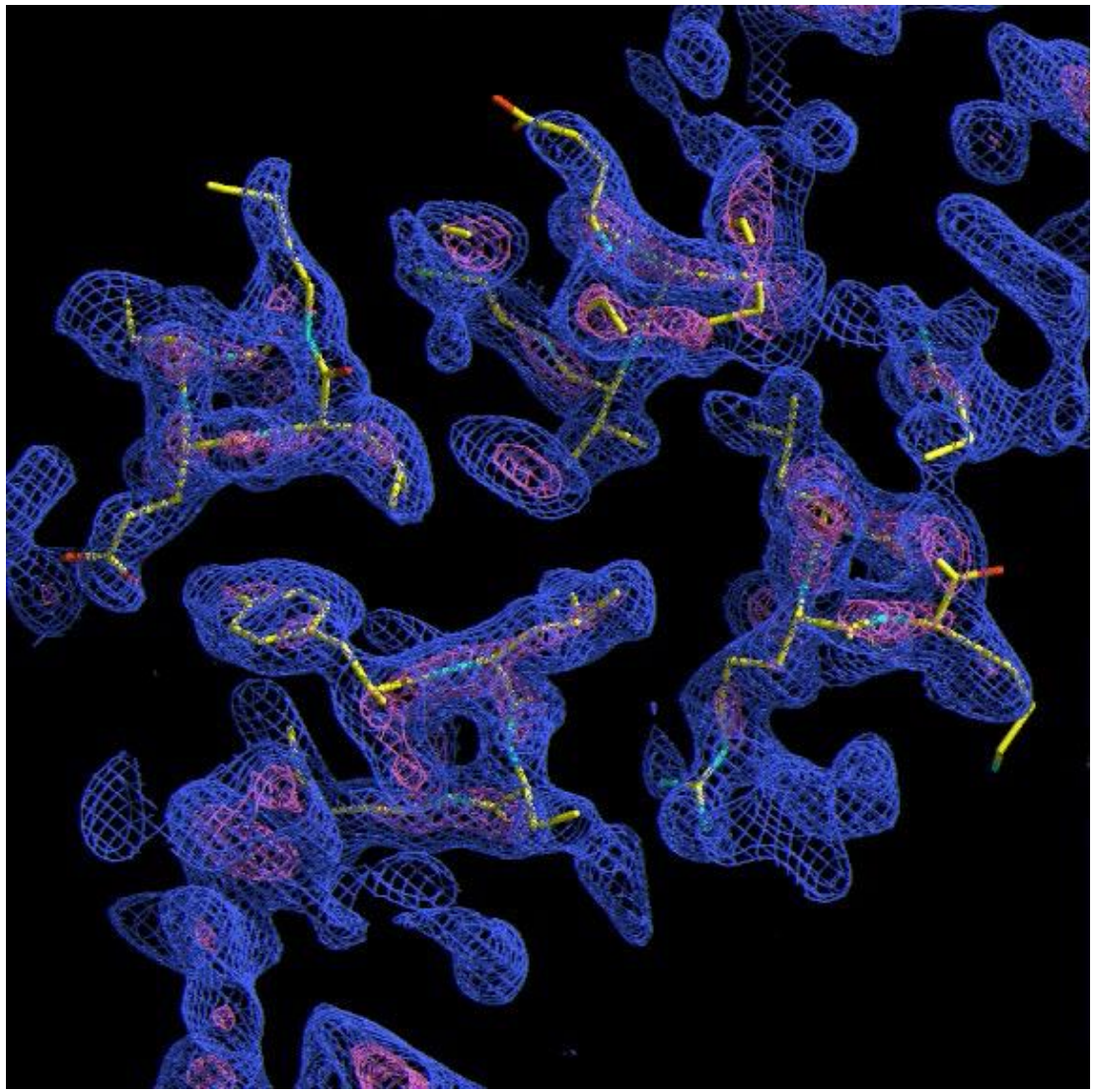


Electron density map displayed at two contour levels:

blue = 1 r.m.s (1σ);
magenta = 3 r.m.s (3σ)

Electron density maps

The task now is to try to fit an atomic model of the protein to the map…



Electron density map displayed at two contour levels:

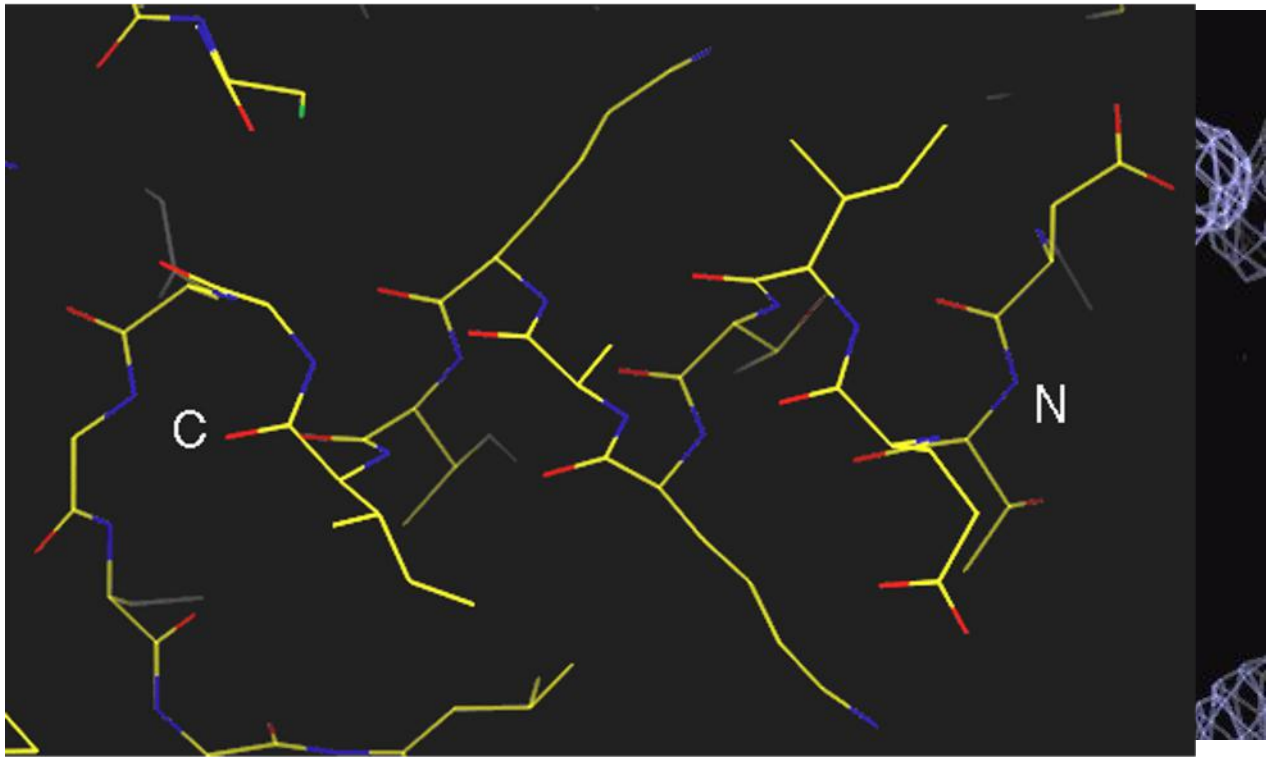
blue = 1 r.m.s (1σ);
magenta = 3 r.m.s (3σ)

Atomic model fitted to the map (in yellow)

Displaying the electron density map

Visualising a 3D electron density map over the entire unit cell can be daunting.

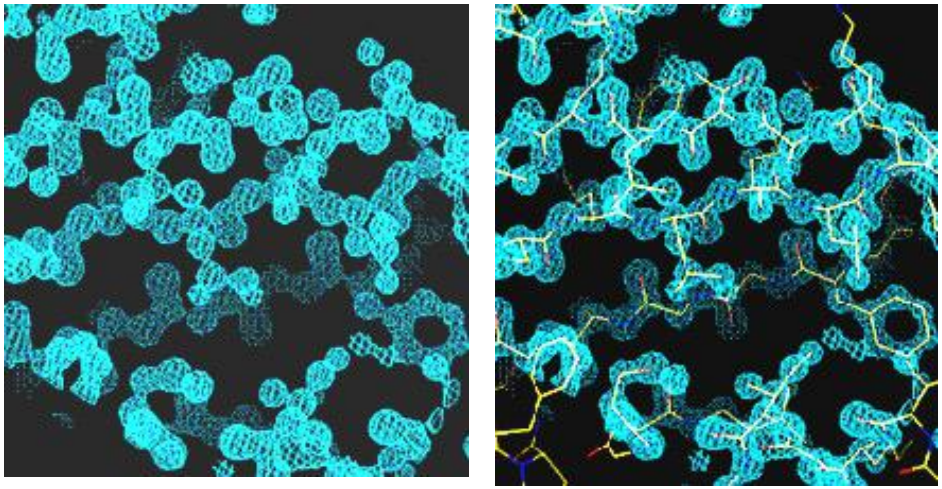
To visualise electron density maps we use high resolution interactive graphics, which display the map as a chicken wire. A map is a 4-dimensional object (each point x,y,z has an associated value) and to display it in 3D we have to choose a contour level that allows us to see about 70 % of the van der Waals radius of the atoms (typically one contours at the 1σ level)



Here is a rather good bit of density (much better than the density you are likely to obtain from the initial phases) and still it is not easy for the “untrained eye” to see what is going on...

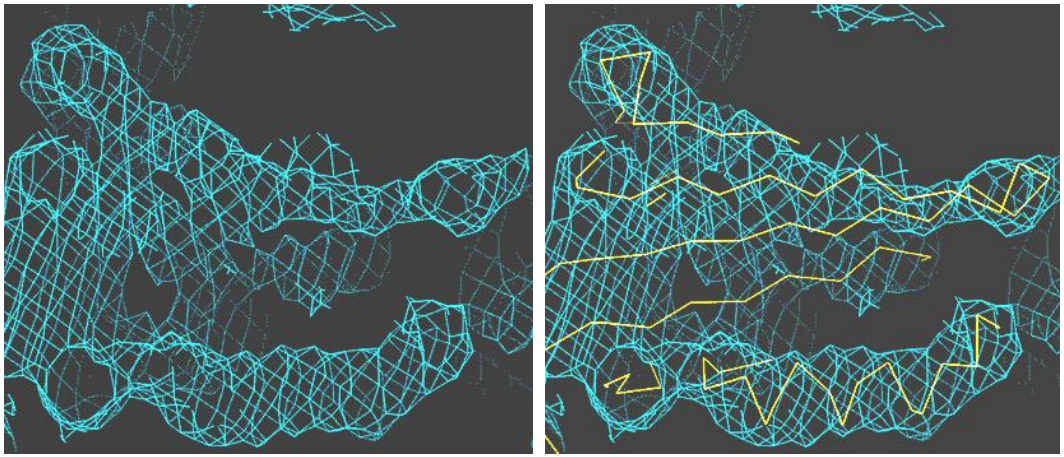
Maps and resolution

The task of model building is to interpret the electron density maps in light of chemical knowledge, basic stereochemistry, chemical sequence, etc...
The level of interpretation depends on the resolution of the map:



Here is a 1 Å map

At very high resolution, individual atoms can be seen and fitted in the electron density blobs: the problem therefore is reduced to 'join-the-dots'



Here is a 6 Å map

At very low resolution only large features can be seen - for example helices look like rods and β-sheets can barely be detected.

Easy/difficult to crystallise?

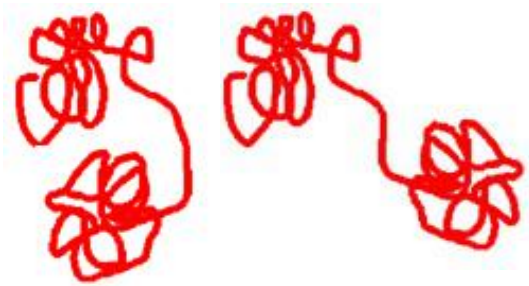
A lot of interesting proteins are difficult to crystallise.

Best cases:

- single proteins
- rigid domain structure
- one dominant conformation

Worst cases:

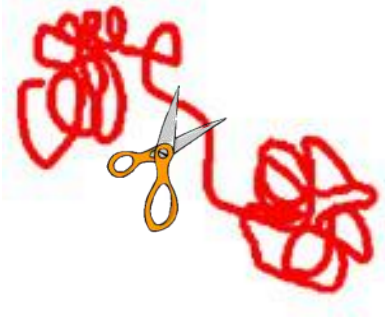
- multi-domain proteins with flexible unstructured linkers
- proteins with flexible N- and/or C-termini
- proteins that are part of large macromolecular complexes
- presence of posttranslational modifications



“Optimise” proteins for crystallisation

Use bioinformatics (database searches, sequence alignments) to identify “core domains” that can be expressed in a soluble form.

Use limited proteolysis to identify compact domains.



beware:
you can
cut flexible
loops).

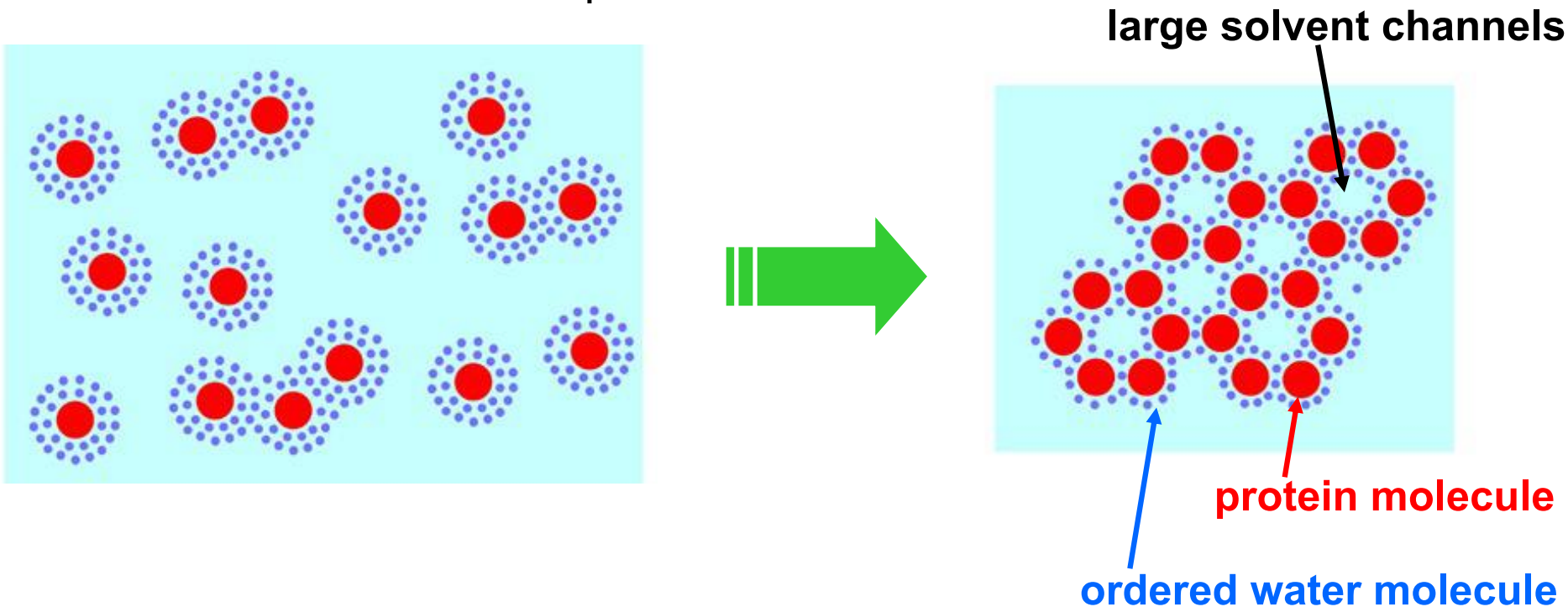
Co-express proteins that are part of the same complex.

Add ligands/inhibitors/cofactors/metals
to stabilise one conformation.

Avoid/encourage post-translational modifications (mutagenesis of target residues, mimic modifications by mutagenesis, change expression system/cell lines, so that the process does/does not occur).

Crystallisation

Protein crystallisation is the transfer of protein molecules from an aqueous solution to an ordered solid phase.



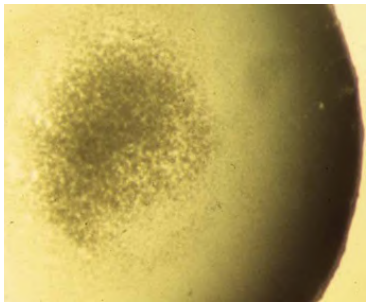
Protein crystallisation is controlled by:
- thermodynamic factors governing the solubility
- kinetic factors governing nucleation and growth

Solubility and crystals

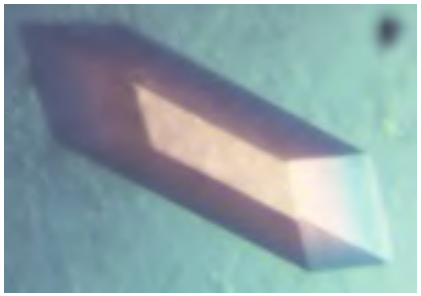
To coax a protein into forming crystals, one has to decrease its solubility to the point where the solution becomes saturated. This is done by changing pH, ionic strength or temperature, or by adding organic solvents.

In a saturated solution the protein in solution is in thermodynamic equilibrium with one or more solid states:

- amorphous precipitate →

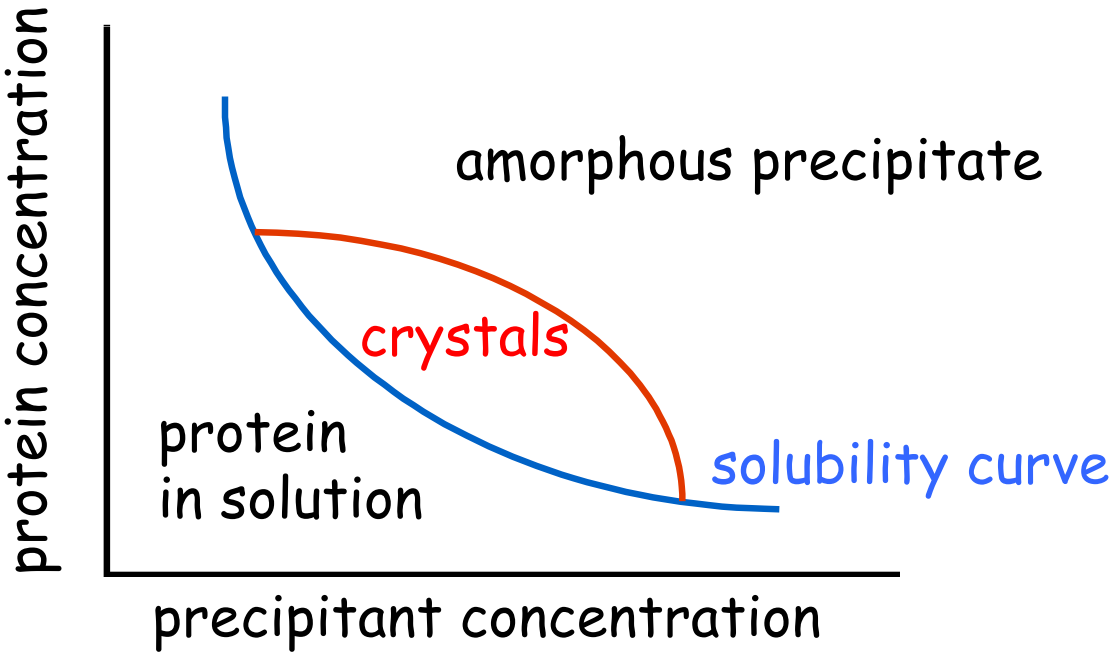


- crystalline state
  - micro-crystals
  - needles
  - plates
  - large single crystals →



Phase diagram

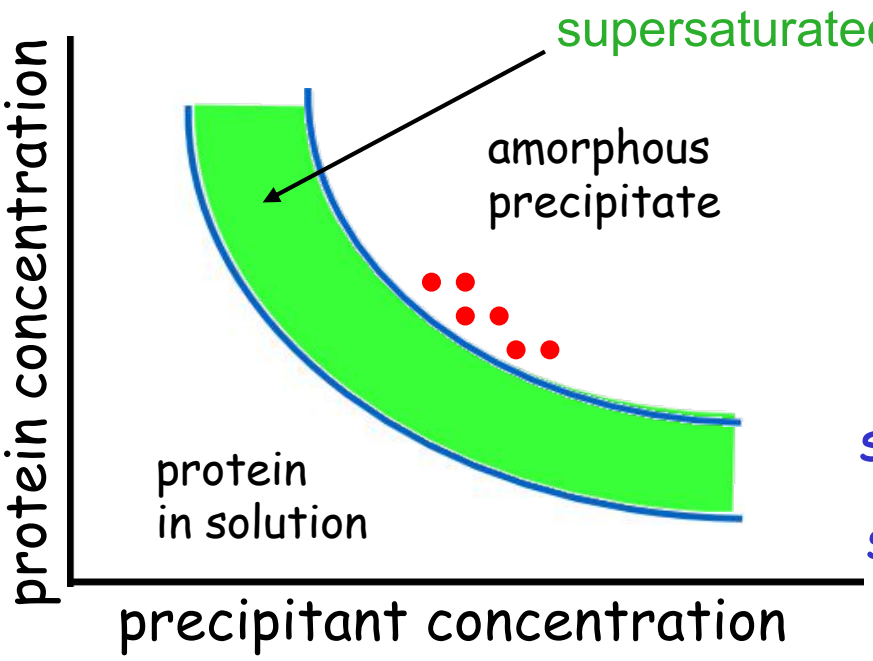
The process of crystal formation can be understood by using a phase diagram. Because several factors influence protein solubility, the phase diagram has many dimensions. A very important factor is the precipitant concentration, and a two-dimensional diagram is shown here:



Supersaturation

Above the solubility curve the solid phase is thermodynamically favoured; yet, because of kinetic barriers to the formation of aggregates, the protein may remain in solution, forming a supersaturated solution.

A supersaturated solution is thermodynamically metastable: its conversion to a stable solid state is kinetically controlled and can be triggered by the presence of nucleation centres.



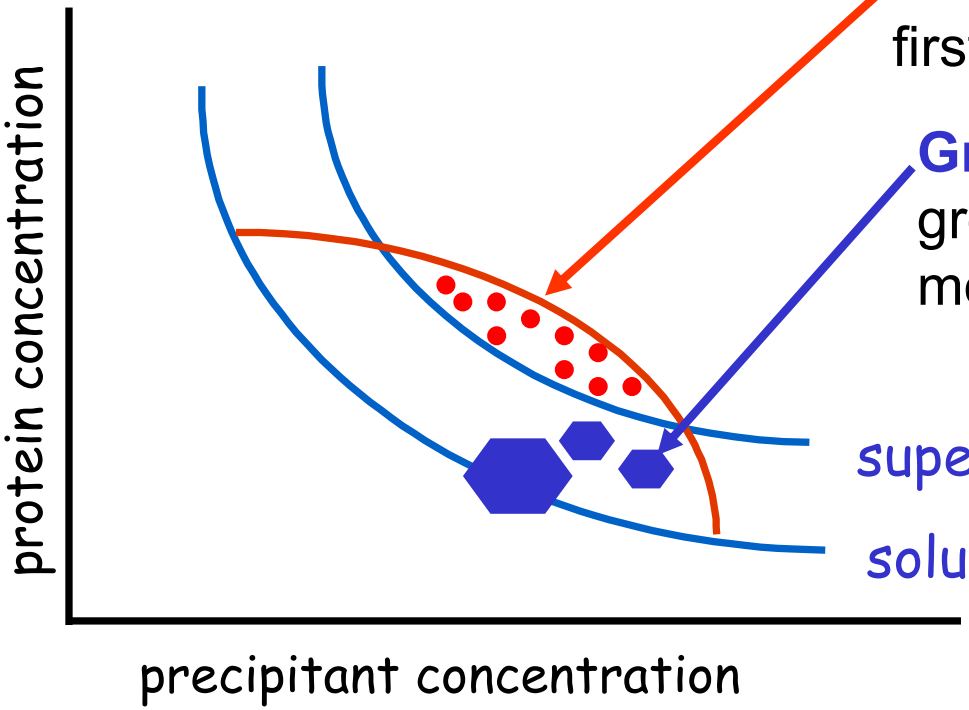
Above a certain level a supersaturated solution will spontaneously give rise to the formation of nuclei, without any need of 'external' nucleation centres.

supersolubility curve

solubility curve

Nucleation and growth

The formation of crystals requires aggregates of a critical size (nuclei). Aggregates smaller than the critical size are unstable and will dissolve, while aggregates larger than the critical size will grow.
The process of crystallisation can be divided into two distinct stages:



Nucleation: the formation of the first few ordered aggregates

Growth: the "most fit" aggregates grow in size by addition of protein molecules from the solution

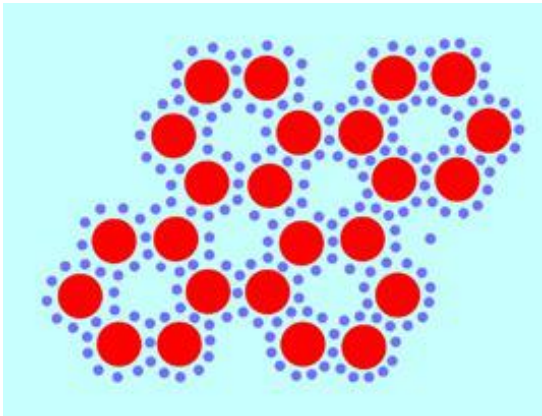
Crystallising proteins

Why proteins are difficult to crystallise:

- most proteins are labile and easily denatured
- large proteins often exist in multiple conformations
- complex behaviour -> polymorphism
- many proteins are difficult to obtain in large amounts
- proteins needs to be highly purified for crystallisation

Why protein crystals are difficult to handle:

- high solvent content (30-80%)
- mechanically fragile
- not well ordered -> resolution limits
- sensitive to radiation damage



Crystallisation of membrane proteins may presents additional problems such as homogeneity of the purified samples, choice of detergents, presence of micelles, tendency to form hydrophobic interactions which are less directional and ordered, etc..

Crystallising proteins

Purity is not an absolute requirement for crystallisation, but often crystals obtained from impure solutions are small or of poor quality.

The protein solution has to be homogeneous not only in terms of composition, but also in terms of structure and conformation.

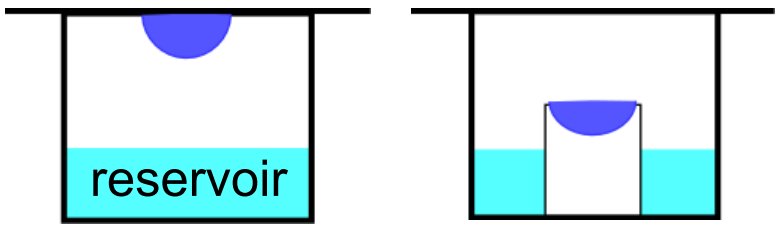
There are many factors to consider. Some of the more important ones are:
- presence of ligands (substrates, cofactors, metal ions etc)
- flexible domains and loops
- post-translational modifications (glycosylation, phosphorylation etc)
- aggregation

Highly homologous proteins will typically crystallise in very different conditions.

One needs large amounts of very pure protein that can be concentrated to several mg/ml without aggregation.
Robotic screening allows testing of ~1000 conditions/mg of protein.

Vapour diffusion

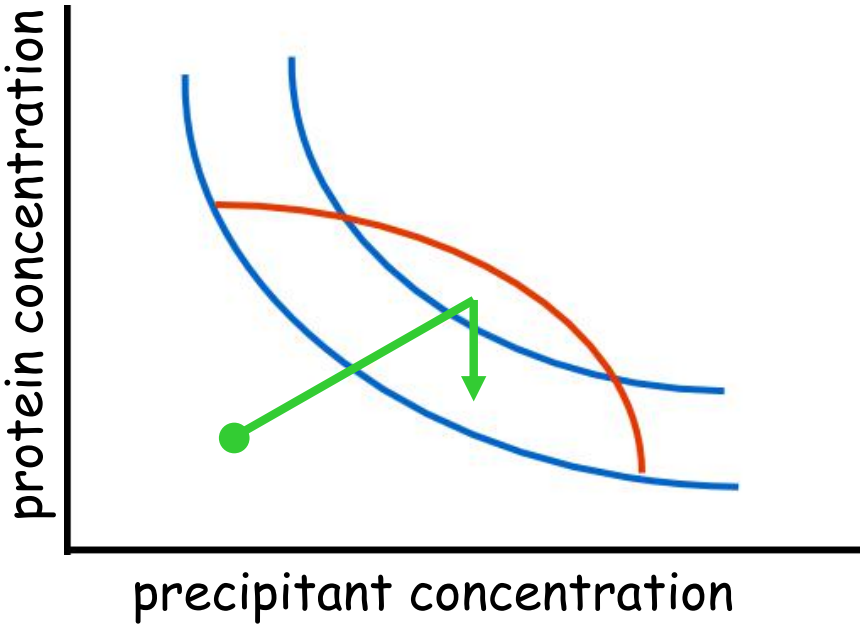
A droplet containing the protein and the precipitating agent is equilibrated against a reservoir containing the precipitant at higher concentration. Equilibrium is reached through the vapour phase, usually by loss of water from the droplet.



hanging drop

sitting drop

Typically the drop is made by mixing equal volumes of protein and reservoir solution.

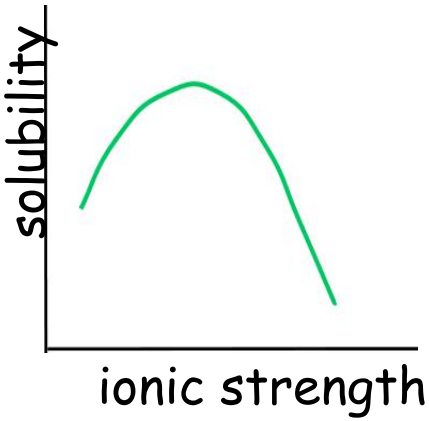


Vapour diffusion is best suited for robotic nanolitre crystallisation.

Precipitants

SALTS

Solubility changes with the ionic strength of the solution, causing salting in/salting out effects
Common ionic precipitants include $(NH4)_2SO_4$, various phosphates, NaCl, LiCl, and many others.



ORGANIC SOLVENTS

Reduce the dielectric constant of the medium, thereby enhancing the electrostatic interactions between the protein molecules and lowering the solubility. May denature proteins.
Commonly used organic solvents include ethanol and isopropanol.

POLYETHYLENE GLYCOLS (PEGs)

Believed to act by a combination of effects: volume exclusion, modification of the dielectric properties of the solution, electrostatic interaction with proteins, depletion of the hydration sphere of the protein. Very effective

Salts and PEGs are the most common precipitants

High throughput crystallisation

Automated methods for crystallisation (and crystal visualization) are now routinely used by most labs. These development are driven by the needs of pharmaceutical companies and structural genomics projects.

Robotic crystallization systems relies on the same principles described so far, but carry out each step of the procedure quickly, accurately, in smaller volumes.

Manual crystallisation:

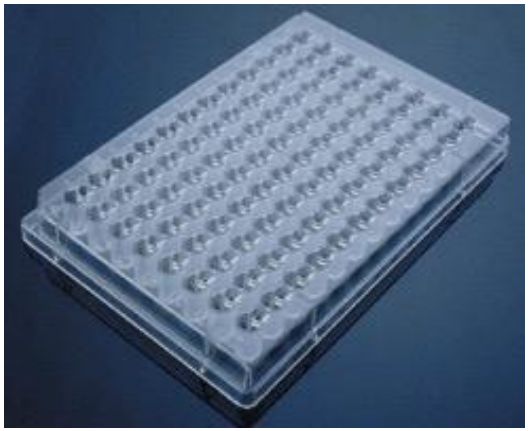
- slow and time consuming
- error prone and not always reproducible
- expensive in terms of amount of purified protein (drops: 1-2 μl)

Robotic crystallisation:

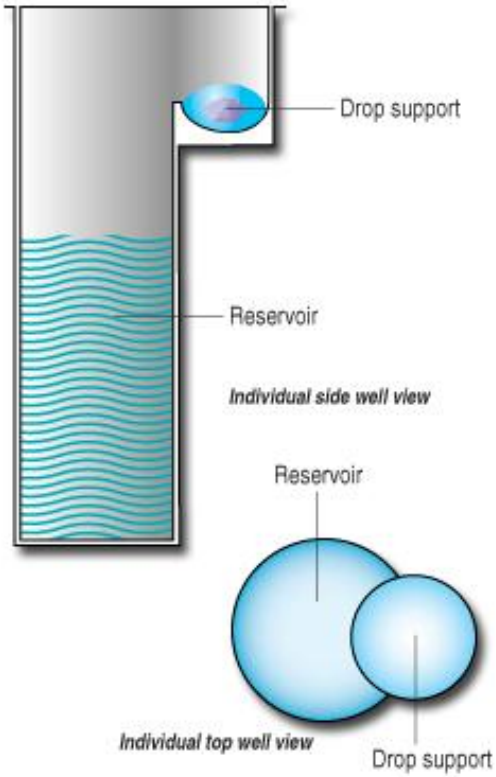
- faster and more efficient
- more accurate and reproducible
- smaller sample sizes (down to 100 nl drops) cut down on expenditure of purified protein

96-wells plates

The standard format for automated crystallisation are plates with 96 reservoir wells. These plates are designed for the sitting drop vapor diffusion method, and are sealed with clear tape.



8 x 12 matrix

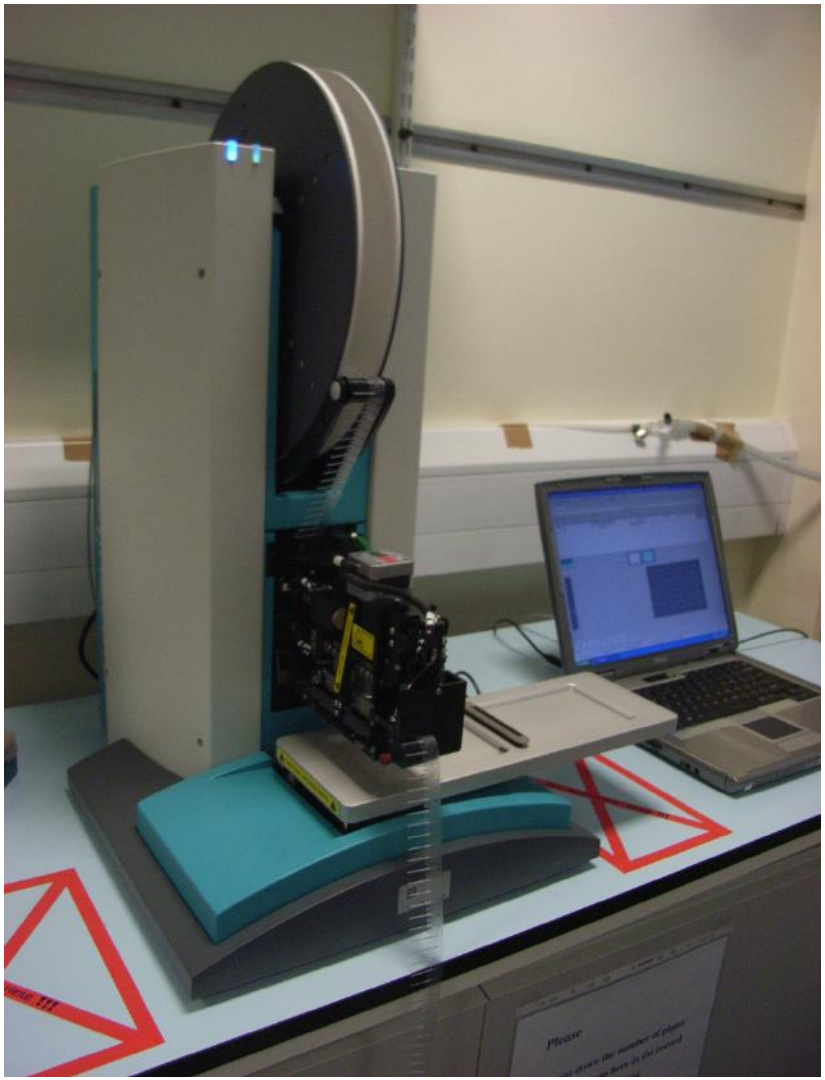


Mosquito crystallisation robot

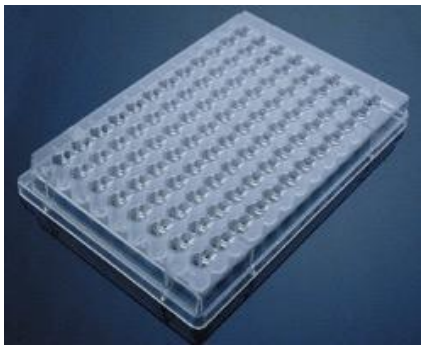
To set up the crystallisation drops containing nanolitre volumes of protein and well solution.

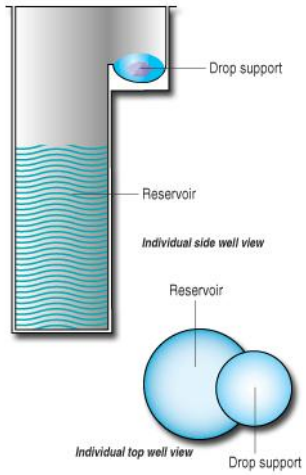
Employs disposable tips and pipettes
100-200 nl drops.

Takes only 2 min to set up 96 drops.



Use of specialised 96-well plates





Structural Biology: Other techniques

For EM: some figs borrowed from Helen Saibil (Birkbeck College)
For STED: some slides borrowed from Stefan Hell (MPI, Goettingen)

Small angle X-rays scattering (SAXS)

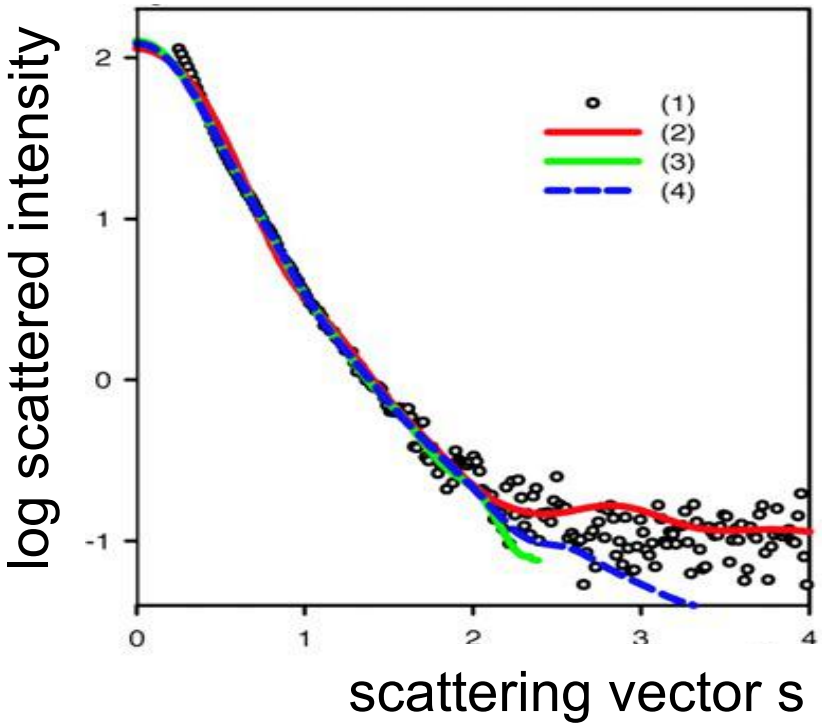
Crystals:

- diffracted intensity concentrates around specific directions
- high resolution (~ λ)

Molecules in solution:

- intensity is isotropic and concentrates around the beam (small-angle)
- low resolution (>>>λ)
- scattering is proportional to that of a single particle averaged over all orientations
- allows to determine size and shape to 1-10 nm resolution (?)
- not a lot of data (1D scattering pattern, radially averaged)
- no unique solution: multiple 3D structures can give rise to same pattern
- need concentration of ~ 1mg/ml, with no aggregation
- need accurate estimate of conc.

Small angle X-rays scattering

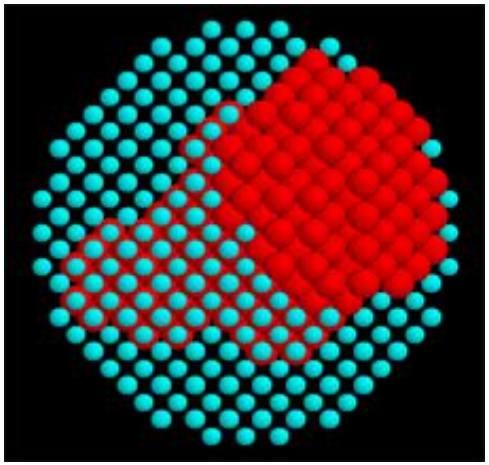


o experimental scattering curve

curve calculated from PX model

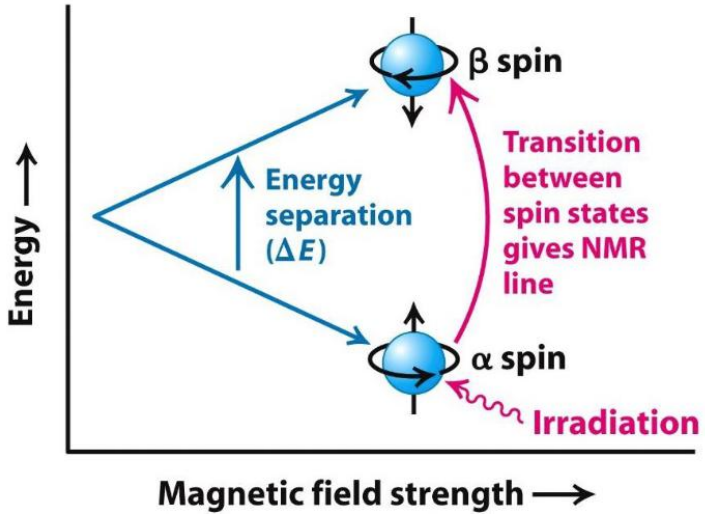
curves calculated from ab initio models

- compare data with known structure/s, to analyse:
  - conformational changes
  - shape/position of unknown parts
- *ab initio* reconstruction by modelling spheres and compare with experimental curve (?)



Nuclear magnetic resonance (NMR)

Nuclei with spin ½ (^{1}H, ^{13}C, ^{15}N, ^{31}P) have angular momentum. In high magnetic fields we can separate the energy levels of nuclei with different spin angular momentum.



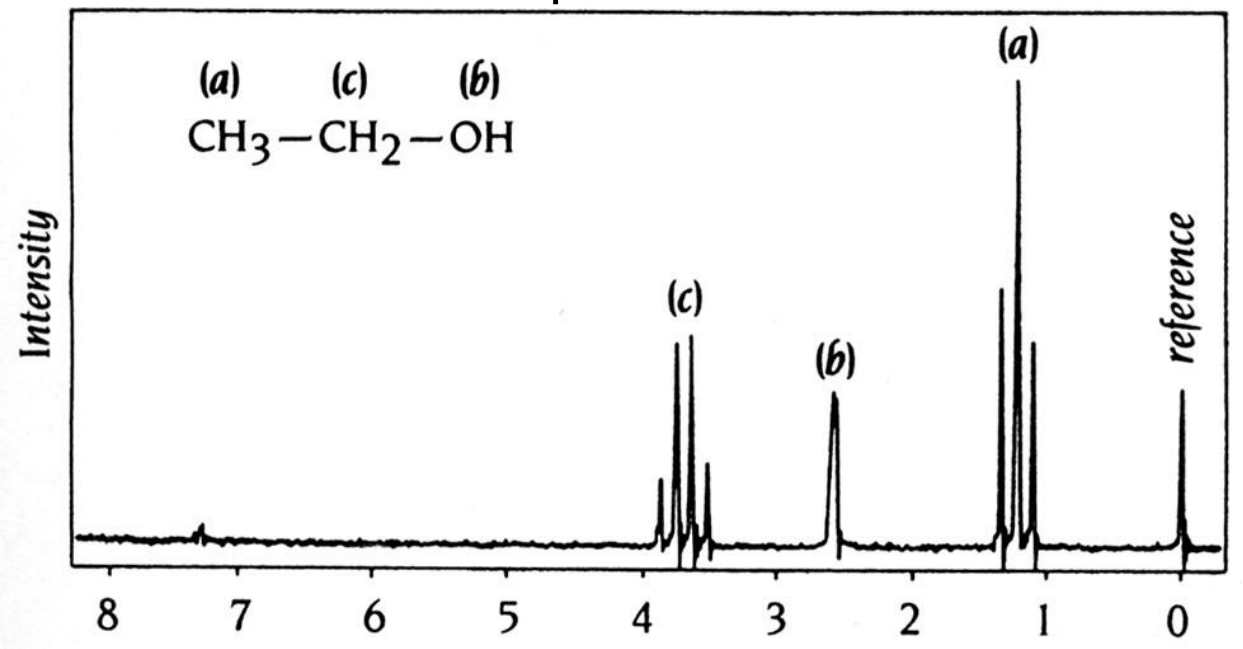
the spin will align along the field, but absorption of electromagnetic radiation of appropriate frequency (radio) induces a transition.

When the nuclei revert to their equilibrium state they emit radiation that can be measured. The exact frequency of the emitted radiation depends on the environment of the nucleus.

NMR: chemical shifts

The exact frequency of the emitted radiation depends on the environment of the nucleus. Nuclei in different chemical environment will absorb at slightly different frequencies. These different frequencies (relative to a reference signal) are called chemical shifts.

^{1}H-NMR spectrum of ethanol

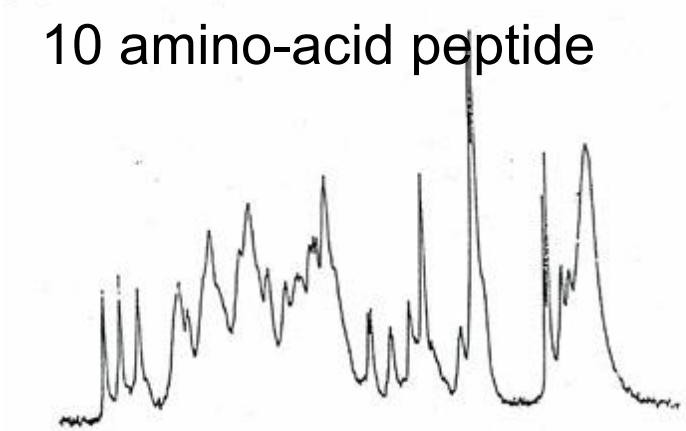


Chemical shifts for all the hydrogens in this small molecule: hydrogen that are part of the CH_3, or CH_2, or OH group have different shifts.

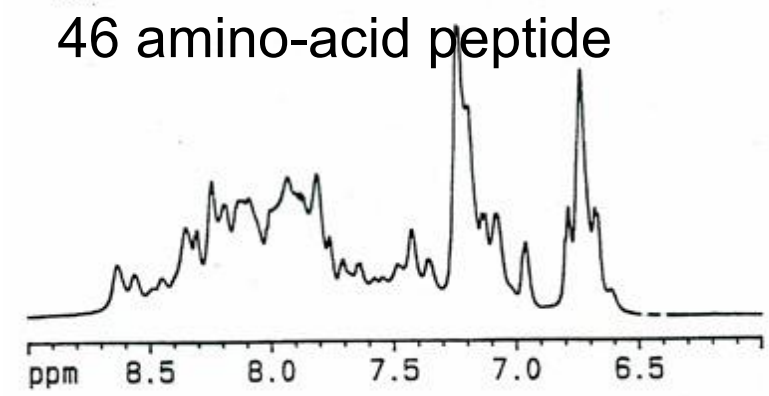
NMR and peptides

1D NMR spectra of proteins contain such a large number of overlapping peaks as to be uninterpretable.

10 amino-acid peptide

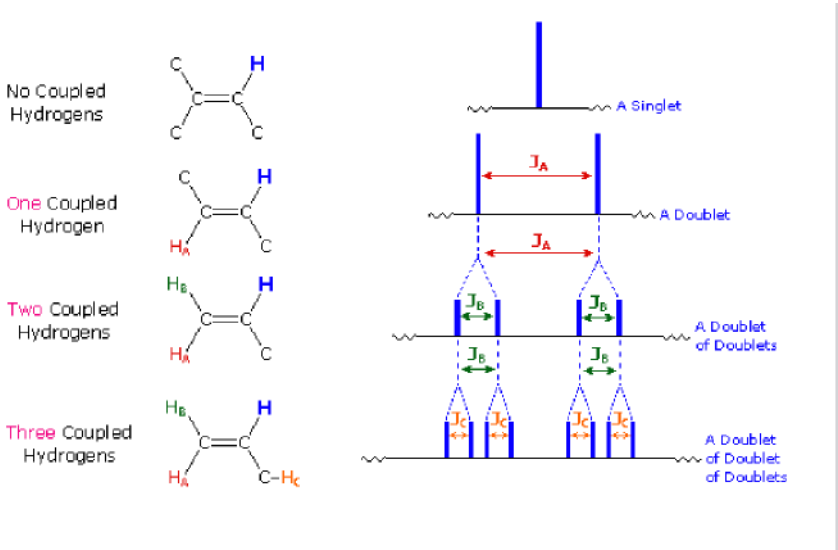


46 amino-acid peptide



153 amino-acid peptide

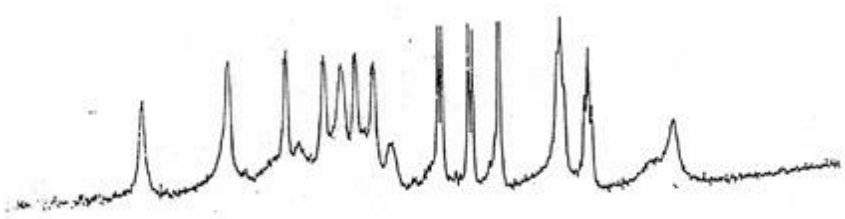
J-coupling: scalar coupling interaction through chemical bonds (electrons) between close nuclei



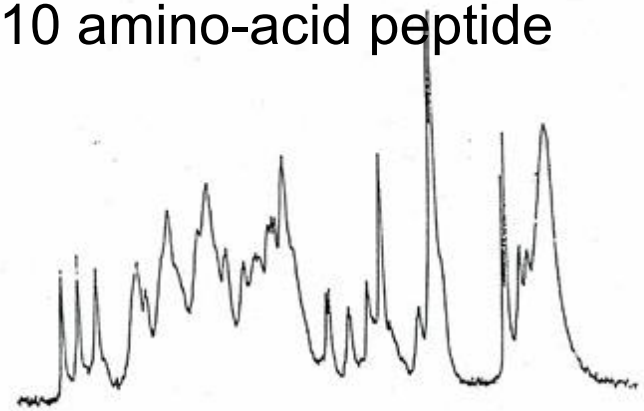
Nuclear Overhauser Effect (NOE) or dipolar coupling, through space—magnetization transfer bewtween close nuclei

NMR and peptides

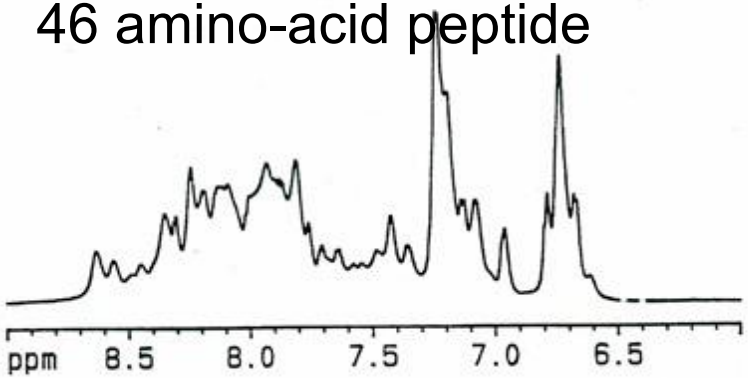
1D NMR spectra of proteins contain such a large number of overlapping peaks as to be uninterpretable.



10 amino-acid peptide

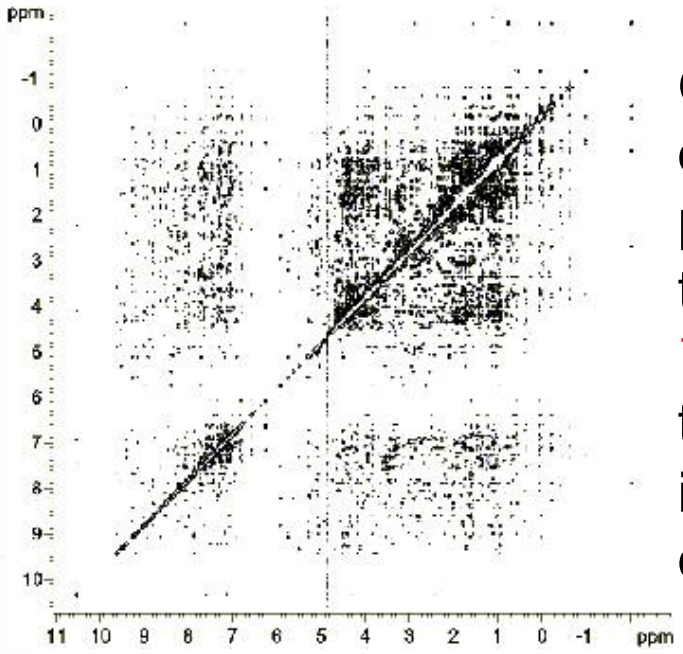


46 amino-acid peptide



153 amino-acid peptide

We need to use *pulse sequences* with specific shapes and time intervals to allow magnetization transfer between nuclei and, therefore the detection of nuclear-nuclear interactions through space (NOESY).



One problem is overlap between peaks: labelling the protein with ^{13}C and ^{15}N help to resolve overlap in the proton dimension.

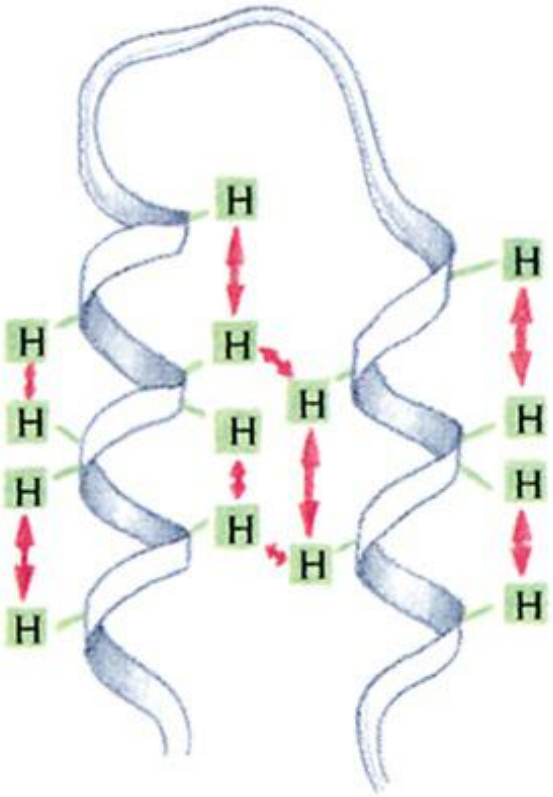
NMR: mapping distances

The tricky thing here is doing the assignment, i.e establishing the identity of the 2 H that give rise to that particular peak. Very labour-intensive. (Wüthrich made a major contribution - Nobel price in 2002).

One problem is overlap between peaks: labelling the protein with ^{13}C and ^{15}N help to resolve overlap in the proton dimension.

The intensity of a NOESY peak is $\propto r^{-6}$, so the distance is determined according to intensity of the peak. The intensity-distance relationship is not exact, so usually a distance range is used.

We can basically map distances between pairs of atoms: the NMR spectra by specifying which pairs of atoms are close together in space, contain three-dimensional information about the protein molecules.

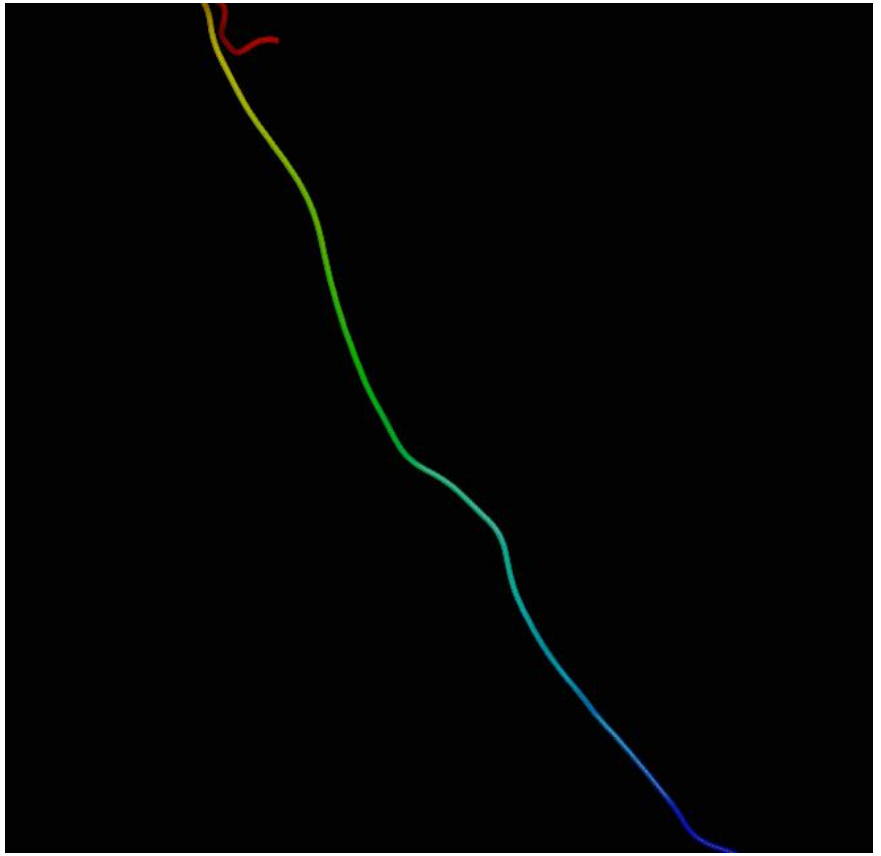


NMR refinement

The experimentally determined distance restraints (together with the protein stereochemical constraints) can be used as input for the structure determination.

Algorithms start from an unfolded polypeptide and attempt to satisfy as many of the restraints as possible, in addition to general properties of proteins such as bond lengths and angles.
As in crystallography, this is done by converting them into energy terms and thus minimizing the energy.

Repeat many times and compare results.



NMR and protein size

Traditionally NMR is limited to relatively small proteins or protein domains.

Problems resolving overlapping peaks - alleviated by the introduction of isotope labelling and multidimensional experiments.

A more serious problem is the fact that in large proteins the magnetization relaxes faster, which means there is less time to detect the signal, causing the peaks to become broader and weaker, and eventually disappear.

Two techniques have been introduced to attenuate the relaxation: transverse relaxation optimized spectroscopy (TROSY) and deuteration of proteins.

No labelling:	< 15 kDa
$^{13}C/^{15}N$:	< 20 kDa
$^{13}C/^{15}N$ and deuteration:	< 40 kDa
Deuteration + TROSY:	up to 100 kDa??