

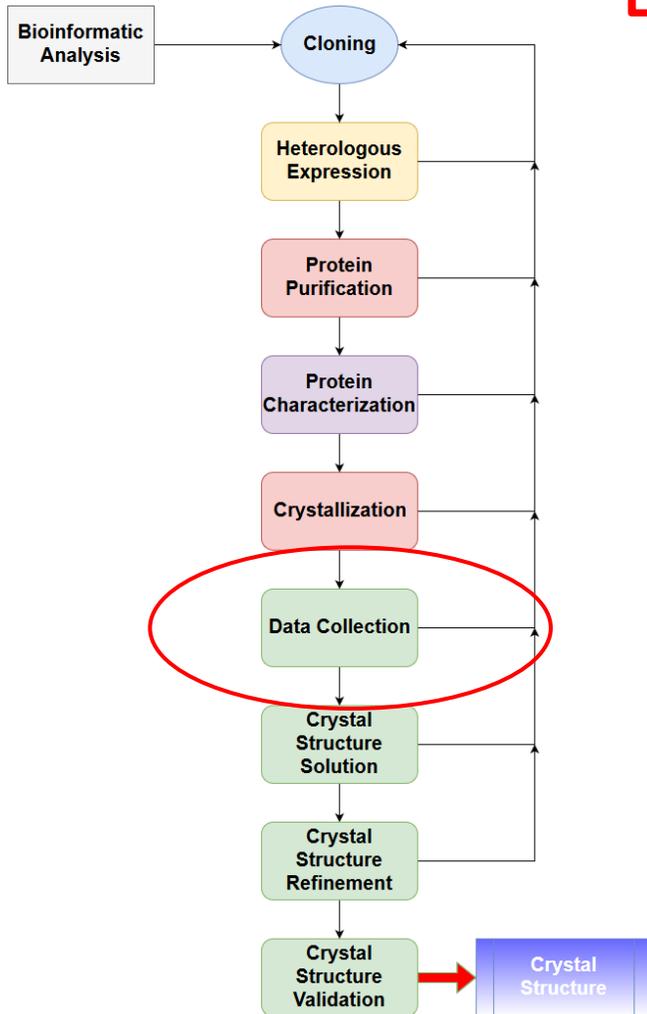
# L'esperimento di diffrazione

Laurea Magistrale in Biotecnologie Mediche  
Curriculum Nanobiotecnologie

A.A. 2022-23

---

# Layout generale



Il protocollo generale per la determinazione strutturale di una macromolecola biologica è piuttosto articolato.

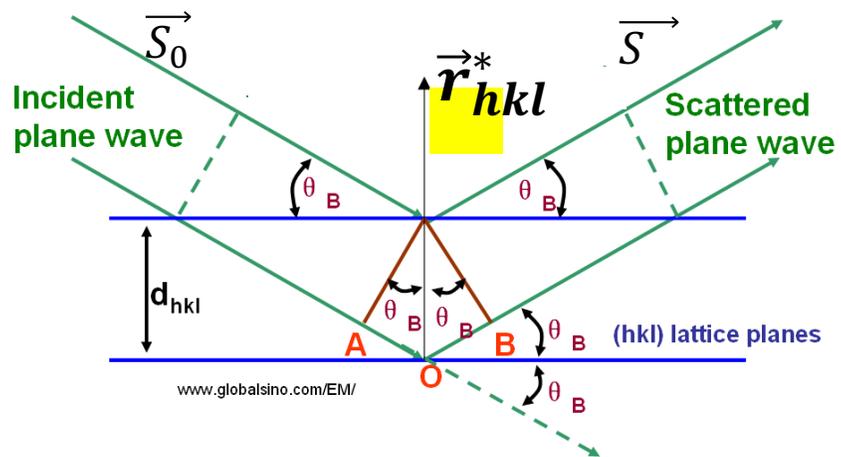
Sono necessarie competenze tra loro molto diverse:

- Biologia Molecolare (clonaggio, espressione)
- Biochimica (cristallizzazione)
- Cristallografia

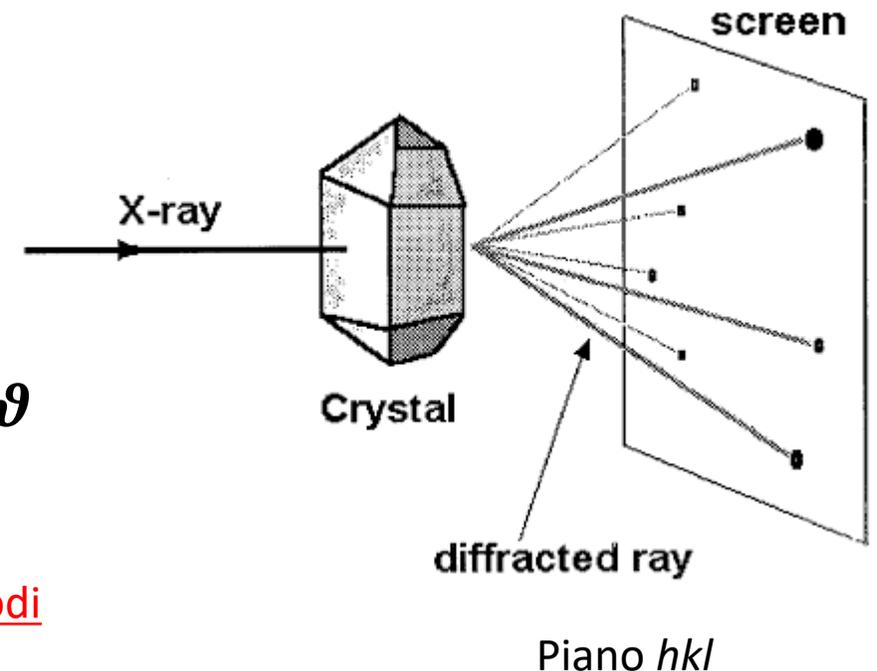
# Introduzione

# Considerazioni preliminari - 1

Quando i raggi-X 'colpiscono' un cristallo, si verifica il fenomeno della diffrazione per cui il cristallo diffonde raggi-X diffratti ad angoli definiti dalla legge di Bragg.



$$n\lambda = 2d_{hkl}\sin\vartheta$$



Parlare di piani cristallini con indici hkl è equivalente a parlare di nodi del reticolo reciproco  $r^*$  con indici hkl.

$r^*$  è la perpendicolare al piano cristallino e ha modulo  $1/d_{hkl}$

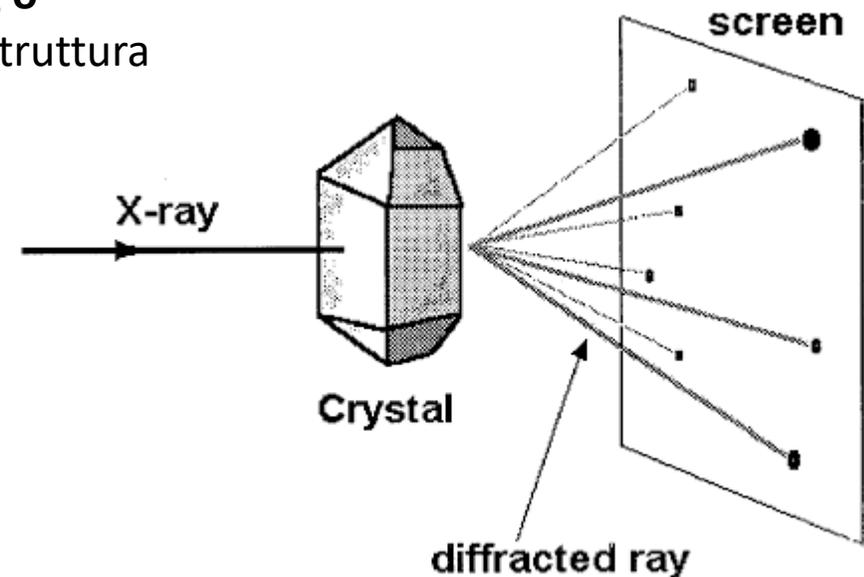
## Considerazioni preliminari - 2

Ogni singolo raggio diffratto, associato ad un preciso piano del reticolo diretto avente indici di Miller (h k l), (detti anche **piani di Bragg** o '**piani di riflessione**') avrà una sua intensità, dipendente dalla struttura del cristallo, ovvero dalla disposizione degli atomi nel cristallo

$$|F_H|^2 \propto I_H$$

**L'intensità diffratta da un piano di Bragg è legata al modulo del fattore di struttura**, indicato come  $F_H$  o anche  $F_{hkl}$  o ancora  $F(hkl)$

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



## Considerazioni preliminari - 3

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

La teoria della diffrazione stabilisce una relazione biunivoca tra le **densità elettronica nella cella elementare** e le **ampiezze delle onde diffratte**, ovvero con i **fattori di struttura**.

Per poter ricostruire la densità elettronica della cella unitaria dobbiamo misurare quante più possibili intensità diffratte.

$$F_{hkl} = |F_{hkl}| \exp(i\varphi_{hkl})$$

Potremo scrivere che:

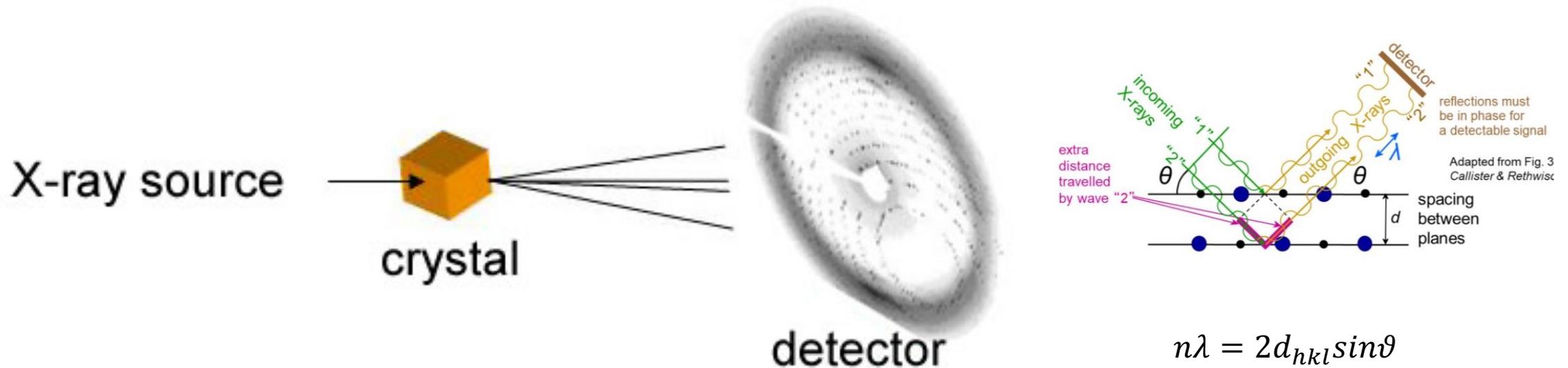
$$\rho_{cell}(r) = \frac{1}{V} \sum_{h,k,l=-\infty}^{\infty} |F_{hkl}| \cdot \exp[-2\pi i (hx + ky + lz) + i\varphi_{hkl}]$$

Misurata sperimentalmente

# Acquisizione dei dati cristallografici (Data Collection)

L'esperimento di diffrazione consiste nell'acquisire quante più intensità diffratte ( $I_{hkl}$ ) dai (potenzialmente) infiniti piani di Bragg del cristallo di indici  $hkl$ .

Le intensità diffratte saranno poi convertire in fattori di struttura  $|F_{hkl}|$  per il calcolo della densità elettronica  $\rho(r)$





# Intensità diffratta

$$I_{hkl} = \frac{\lambda^3}{V_{cell}^2} V_{crys} I_0 PK |F_{hkl}|^2$$

$$|F_H|^2 \propto I_H$$

L'ampiezza dell'onda diffratta è proporzionale alla radice quadrata dell'intensità diffratta **misurata** nel corso di un esperimento di diffrazione

In un esperimento di diffrazione **misuriamo esclusivamente l'intensità diffratta** e quindi l'ampiezza massima dell'onda.

# Intensità diffratta e struttura molecolare

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

Per ricostruire la densità elettronica a partire dai dati di diffrazione dobbiamo conoscere anche  $\varphi_{hkl}$ , ovvero la **fase dell'onda diffratta (fase di  $F_{hkl}$ )**.

**La fase di  $F_{hkl}$  non può essere ottenuta direttamente dall'esperimento di diffrazione ma deve essere determinate in altro modo**, o con altri esperimenti di diffrazione o con metodi computazionali (**problema della fase**)

# Accuratezza della struttura cristallina: Risoluzione e Completezza

# Il cristallo

Caratteristiche dei cristalli di proteine:

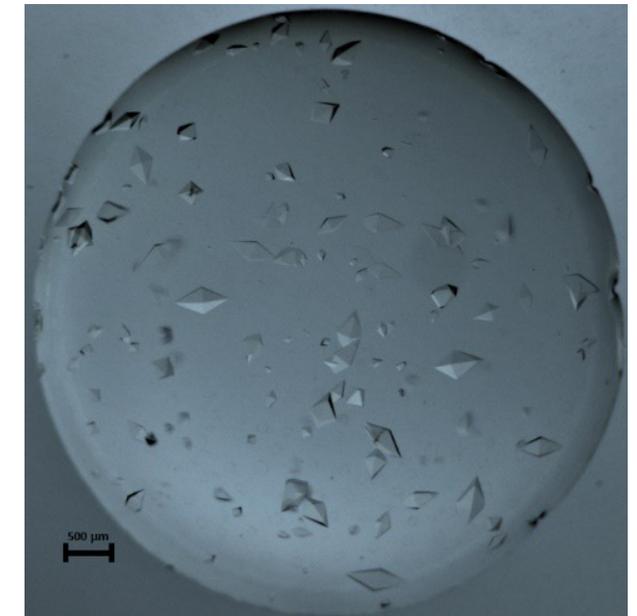
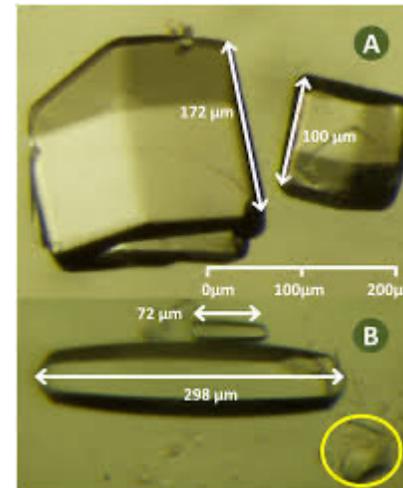
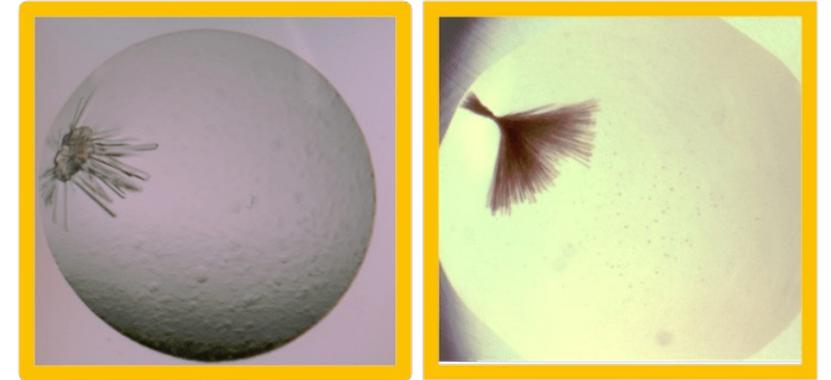
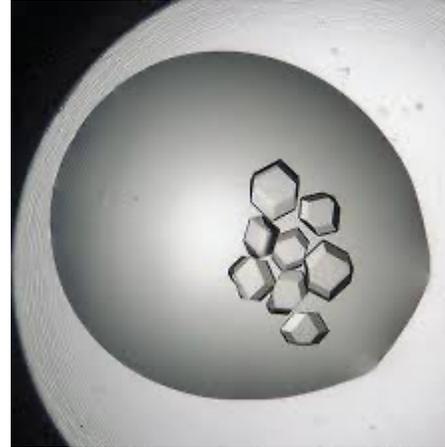
- Fragilità
- Si deteriorano con facilità
- Dimensioni contenute (10 – 100  $\mu\text{m}$ )
- Morfologia differenziata

La forma del cristallo e la sua dimensione sono indicatori preliminari della qualità del cristallo.

La qualità del cristallo viene valutata misurando le proprietà di diffrazione.

**Il cristallo migliore è quello che diffrange meglio** (non necessariamente il più bello!)

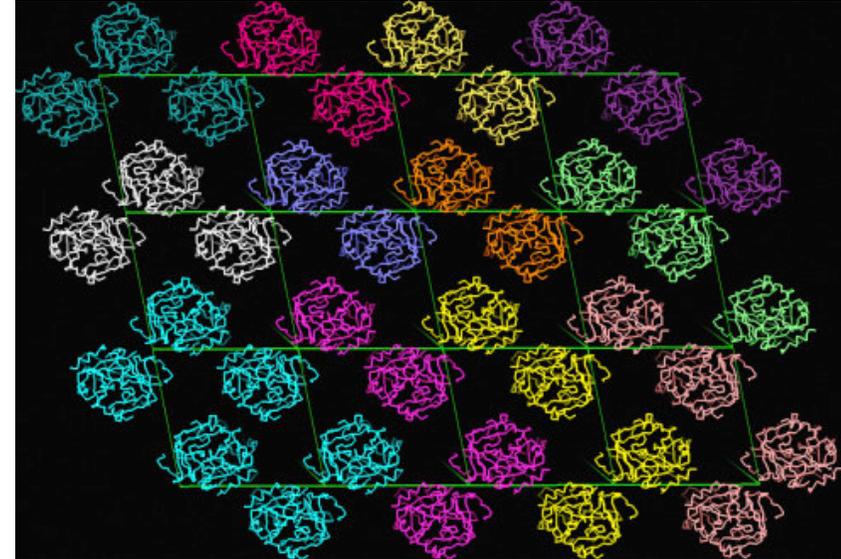
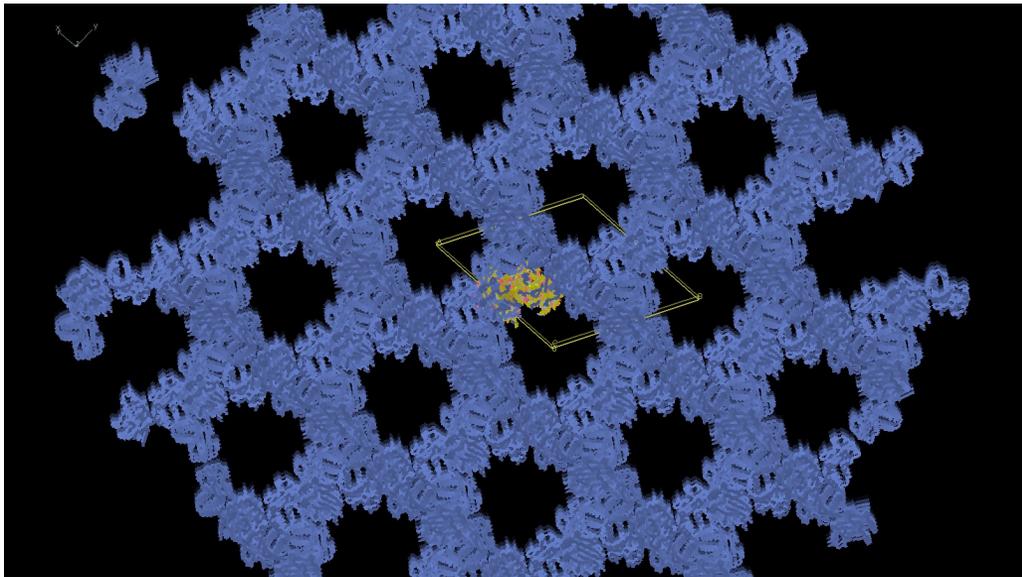
**Basta anche 1 solo cristallo (buono)**



# I cristalli di macromolecole

Un cristallo di 'proteina' è costituito per buona parte di acqua 'disordinata' (30-80 %; 40-50 % in media)

Internamente il cristallo possiede dei 'canali' più o meno larghi occupati da acqua (e dai componenti del buffer)



L'elevato contenuto di acqua è un problema per molti versi: fragilità, diffrazione debole.

Tuttavia l'acqua del cristallo assicura la presenza di un ambiente idoneo alla macromolecola e permette che la proteina sia in uno stato 'funzionale'.

# Struttura molecolare e risoluzione

La formula che lega la densità elettronica alle intensità diffratte ci dice che in linea di principio, per avere una rappresentazione assolutamente fedele della  $\rho(\vec{r})$ , **devo acquisire infiniti raggi diffratti ( $h, k, l \rightarrow \infty$ )**.

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

In pratica questo non è possibile, anche perché i **cristalli reali non diffrangono i raggi-X in modo ideale**.

Più è elevato il numero di termini nella sommatoria ( $F(hkl)$ ) più la mia ricostruzione della densità elettronica sarà fedele alla densità elettronica *vera*.

# Risoluzione

Legge di Bragg

$$n\lambda = 2d_{hkl}\sin\vartheta$$

$$\frac{d_{hkl}^*}{2} = \frac{1}{2d_{hkl}} = \frac{\sin\vartheta}{n\lambda}$$

Nelle condizioni sperimentali oltre un certo valore di  $\frac{\sin\vartheta}{\lambda}$  il rapporto segnale/rumore è troppo basso, ovvero **non distinguo più l'intensità diffratta dal rumore di fondo**. Quindi esiste un angolo  $\vartheta$  al di sopra del quale non ho più un'intensità diffratta significativa perché la diffrazione è così debole da confondersi con il rumore di fondo.

Tale angolo  $\vartheta$  è legato al valore di  $d_{hkl}$  (inversamente proporzionale).

Il valore di  $d_{hkl}$  minimo a cui ho ancora un rapporto segnale/rumore accettabile è nota come **Risoluzione della struttura.**

$$\theta \leftrightarrow \frac{\sin\vartheta}{\lambda} \leftrightarrow d_{hkl}^* \leftrightarrow \frac{1}{d_{hkl}}$$

# Risoluzione

La risoluzione di una struttura cristallografica si indica in Å ed è il valore minimo di  $d_{hkl}$  a cui l'intensità della diffrazione è ancora significativa.

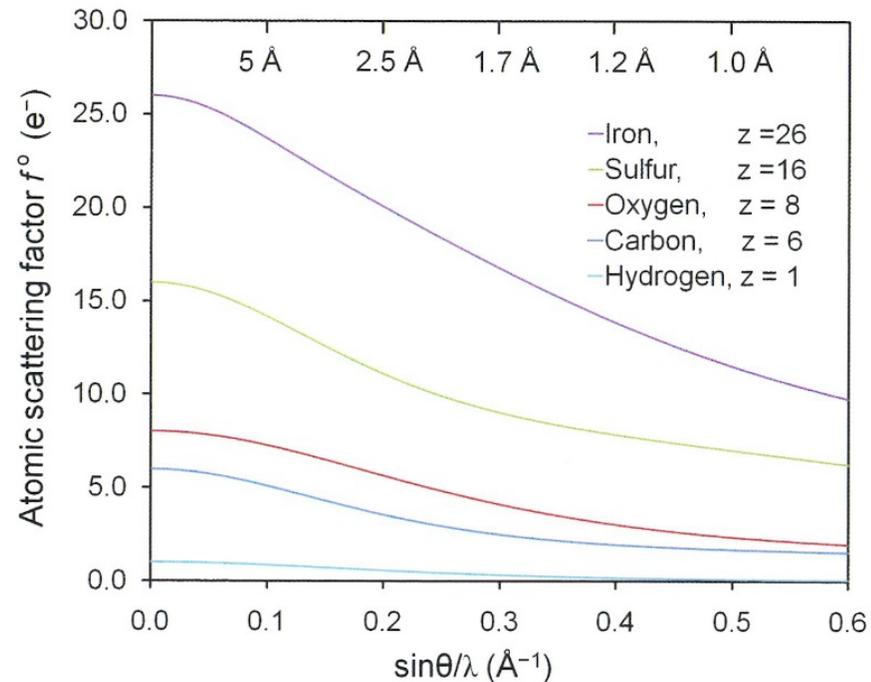
$$\frac{1}{d_{hkl}} = \frac{2 \sin \vartheta}{n\lambda}$$

Più  $d_{hkl}$  è piccolo più la risoluzione è elevata

**E' uno dei parametri più importanti di una struttura cristallografica**

# Parametri che influenzano la risoluzione e $d_{hkl}$

- La **intensità diffusa da un singolo atomo diminuisce all'aumentare dell'angolo di diffusione  $\vartheta$**
- Il **moto termico** attenua ulteriormente l'intensità diffratta al crescere dell'angolo  $\vartheta$
- Il **disordine interno** del cristallo attenua la capacità del cristallo di diffrangere efficacemente



$$f_{at} = f(r^*)$$

O anche

$$f_{at} = f\left(\frac{\sin\vartheta}{\lambda}\right)$$

# Significato della risoluzione

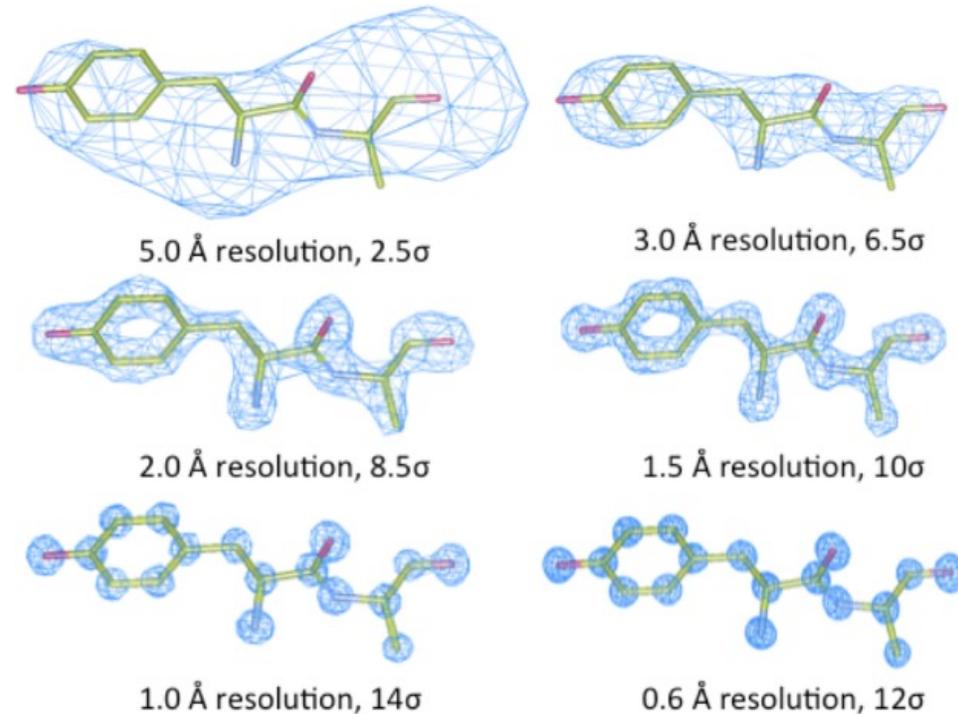
Il significato della risoluzione è uguale a quello descritto in ottica, in sostanza indica la distanza minima a cui posso distinguere due oggetti.

Più la risoluzione è elevata, ovvero  $d_{hkl}$  è piccolo e più la struttura cristallografica sarà precisa, ovvero dettagliata

La diversa precisione nella descrizione della  $\rho(\vec{r})$  dipende direttamente dalla formula che lega fattori di struttura e densità elettronica.

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

Una risoluzione più elevata vuole dire che sto sommando più termini, e avrò così una rappresentazione della  $\rho(\vec{r})$  più fedele alla situazione *vera* (a meno di errori sperimentali).



# Valori di risoluzione

Risoluzione  
[ $d_{hkl}$  in Å]

4-6

Risoluzione

Molto bassa (folding)

3-4

Bassa (Backbone, poco dettaglio)

2-3

Media (Backbone e Side-chains)

1.3- 2.0

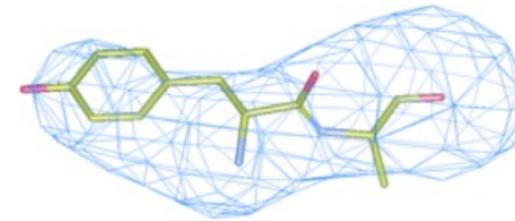
Buona (Mappa dettagliata)

0.9-1.3

Alta (Atomi sferici)

> 0.9

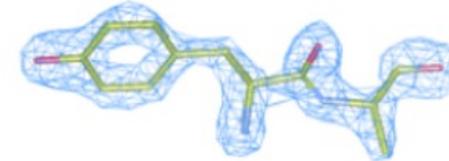
Estremamente Alta (Idrogeni?)



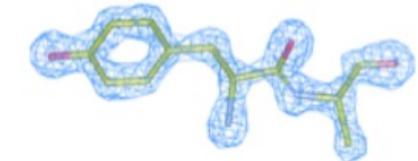
5.0 Å resolution, 2.5 $\sigma$



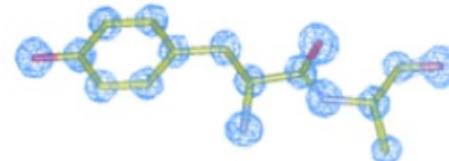
3.0 Å resolution, 6.5 $\sigma$



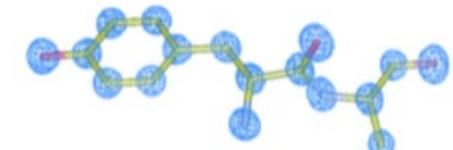
2.0 Å resolution, 8.5 $\sigma$



1.5 Å resolution, 10 $\sigma$



1.0 Å resolution, 14 $\sigma$



0.6 Å resolution, 12 $\sigma$

# Completezza

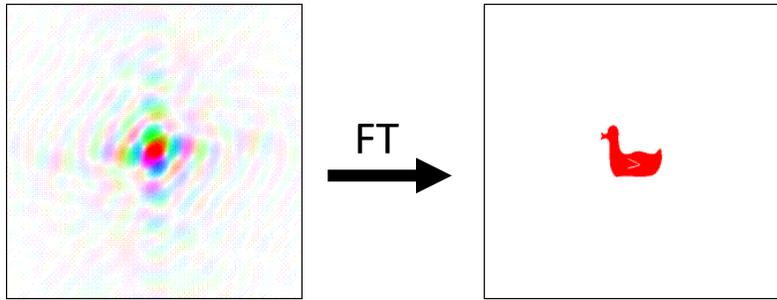
Un altro aspetto importante della formula che lega i fattori di struttura alla densità elettronica, è che non solo la risoluzione deve essere il più elevata possibile, ma che non si deve omettere, per quanto possibile, nessun dato (per la risoluzione data).

L'omissione di fattori di struttura, o gruppi di fattori di struttura, da luogo ad una rappresentazione della  $\rho(\vec{r})$  alterata, a secondo della quantità e delle modalità di omissione dei fattori di struttura.

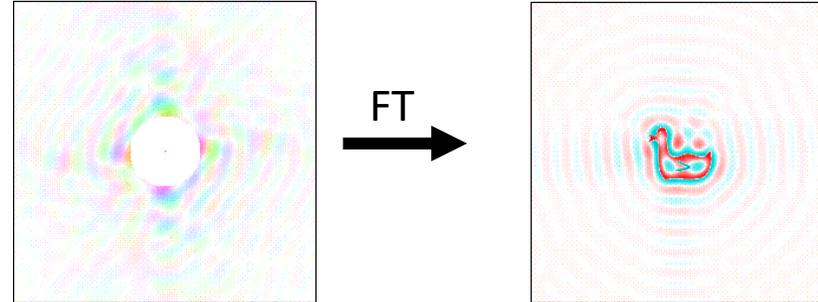
$$\rho_{cell}(r) = \frac{1}{V} \sum_{h,k,l=-\infty}^{\infty} F_H \exp[-2\pi i (hx + ky + lz)]$$

La percentuale di intensità diffratte effettivamente misurate, rispetto alla totalità delle intensità teoricamente misurabili prende il nome di **completezza** del 'data set' e dovrebbe essere superiore al 90% e avvicinarsi al 100%.

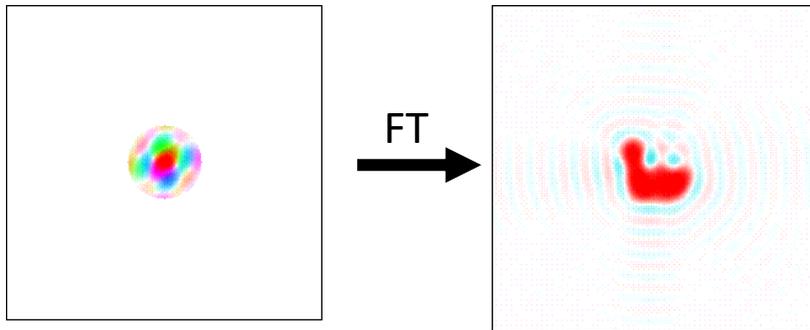
# Trasformata di Fourier di una papera



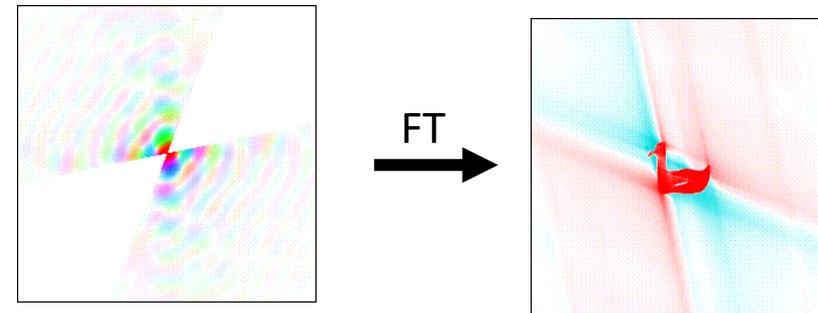
Tutti i dati



Solo i dati ad alta risoluzione

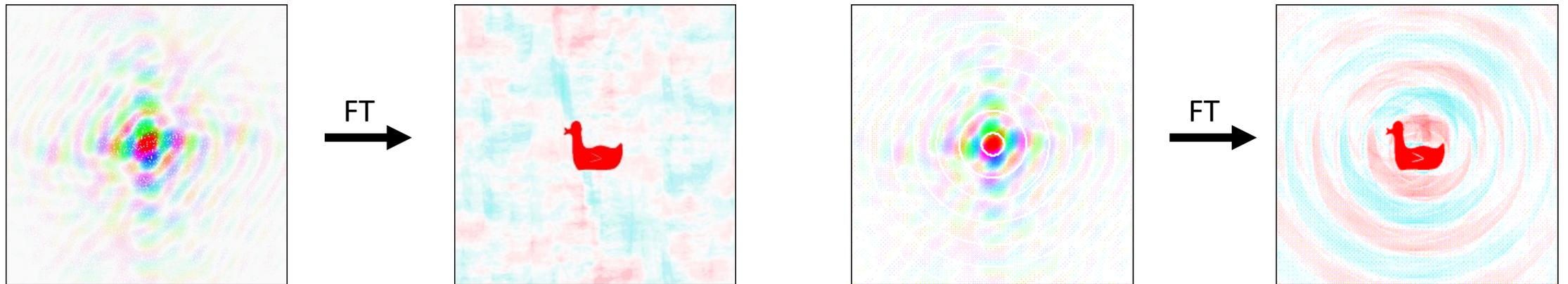


Solo i dati a bassa risoluzione



Manca un intero segmento di dati

# Trasformata di Fourier di una papera



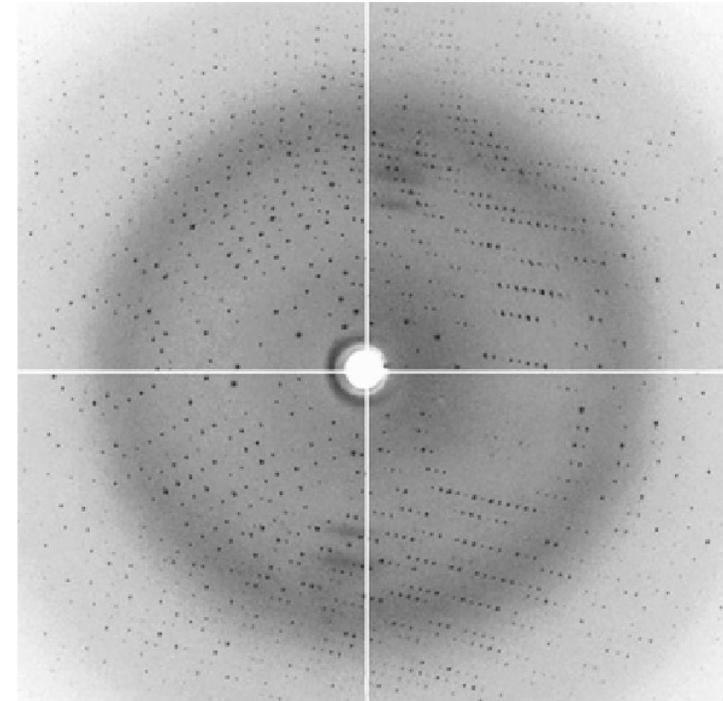
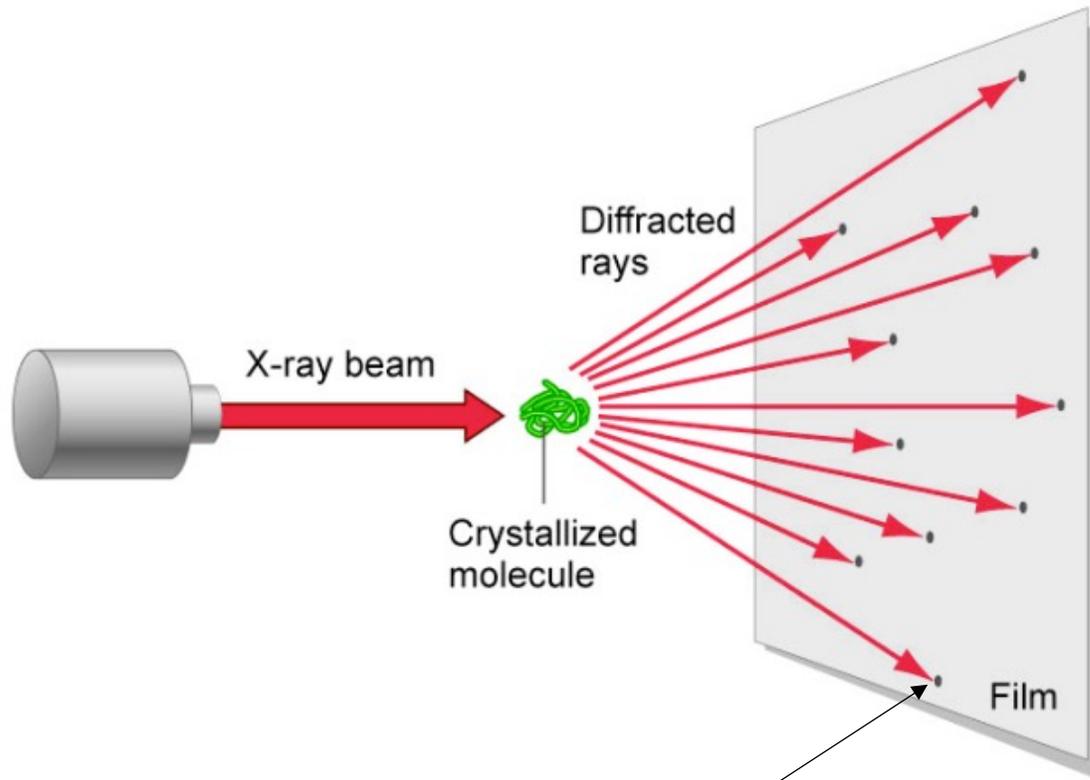
10% dei dati omessi in modo random

10% dei dati omessi in cerchi concentrici

**L'assenza di una parte dei fattori di struttura (dati 'incompleti') causa delle aberrazioni nell'immagine (densità elettronica) ricostruita**

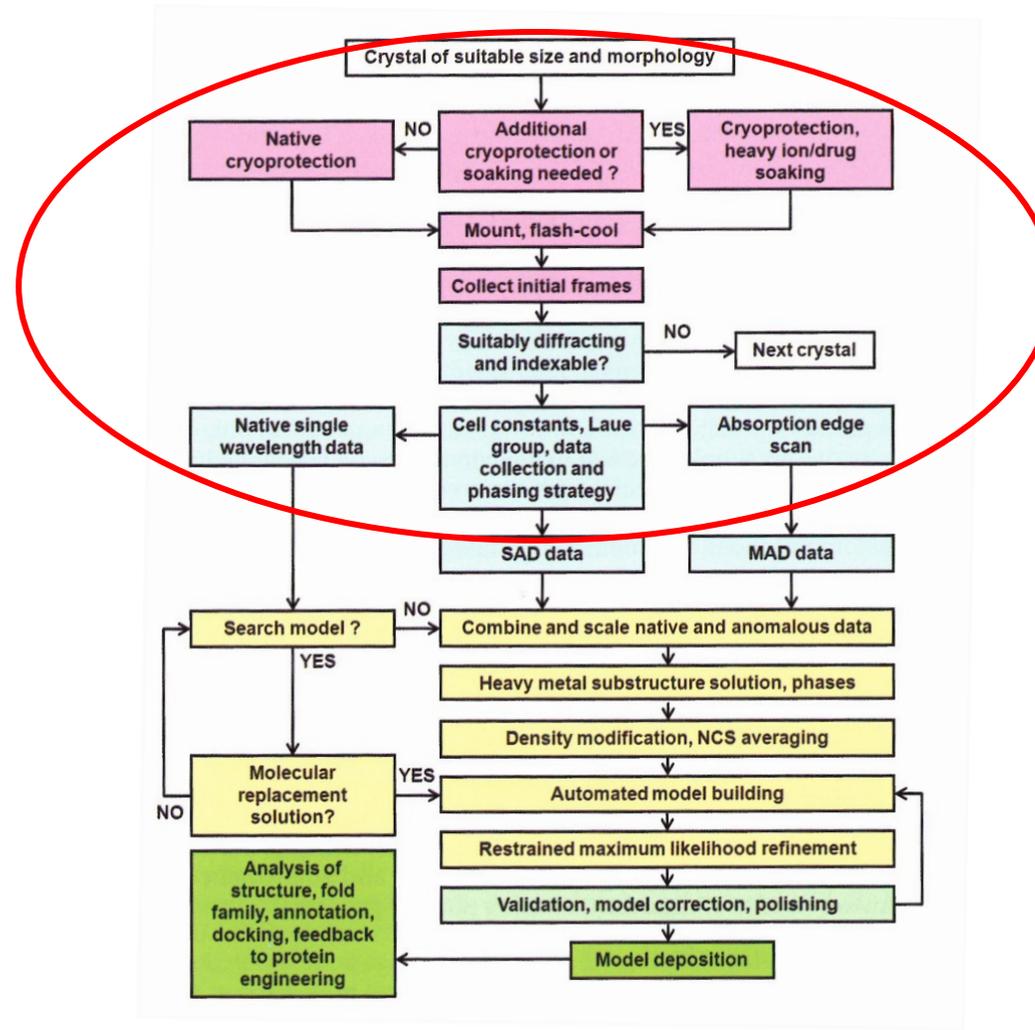
# Acquisizione dei dati di diffrazione: Principio

# Esperimento di diffrazione



Diffrazione dal Piano reticolare (hkl)

# Schema generale



# Montaggio del cristallo

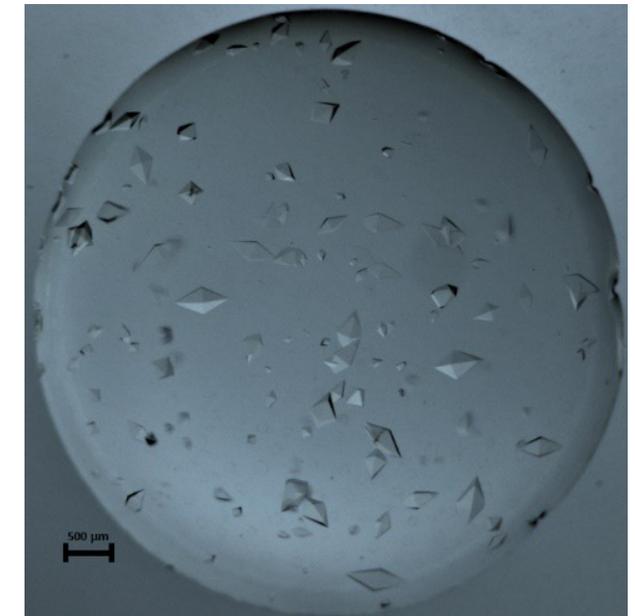
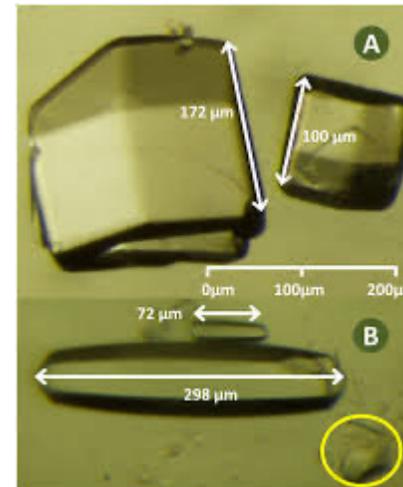
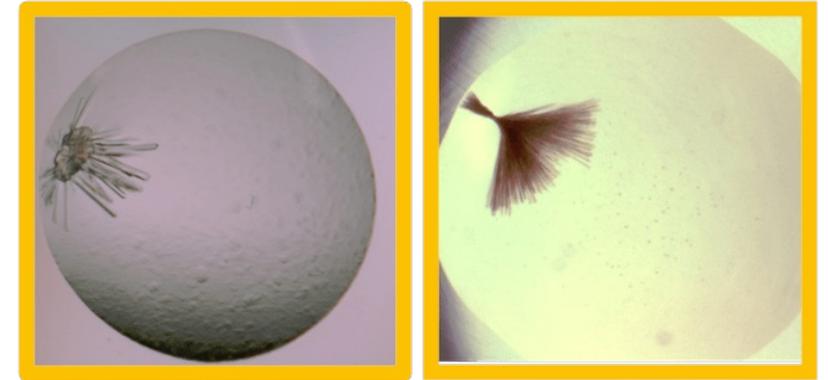
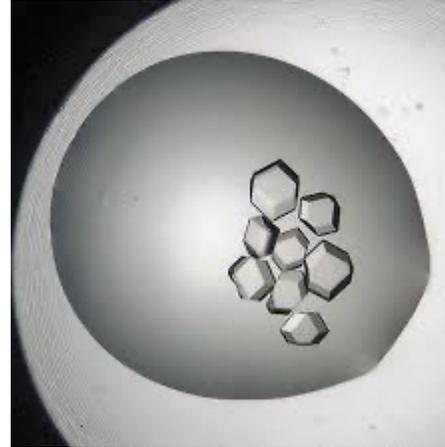
# Il cristallo

Caratteristiche dei cristalli di proteine:

- Fragilità
- Si deteriorano con facilità
- Dimensioni contenute (10 – 100  $\mu\text{m}$ )
- Morfologia differenziata

I cristalli vengono ‘cresciuti’ in volumi di liquido (gocce) molto piccoli (100 nL – 10  $\mu\text{L}$ ) in un sistema chiuso dove si è stabilito un equilibrio.

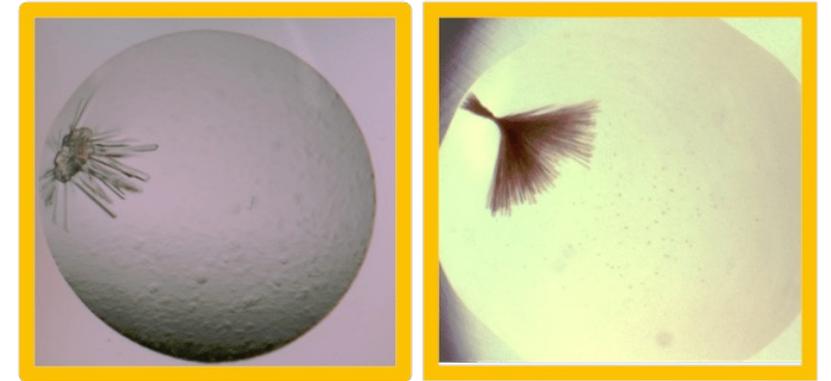
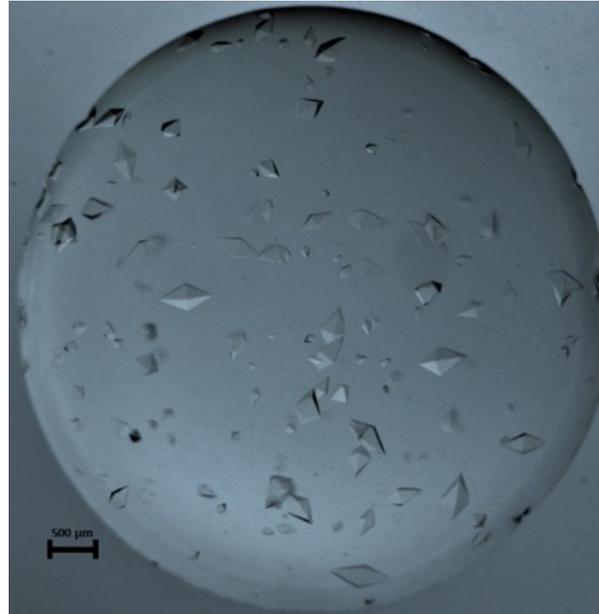
Nel suo ambiente di crescita il cristallo si trova in condizioni chimico-fisiche ben definite (composizione chimica, temperatura, talvolta luce e potenziale redox)



# Osservazione al microscopio

I cristalli devono essere trasferiti in appositi 'contenitori' (loops, capillari) per la misura diffrattometrica, questa operazione è detta '**montaggio**' .

Il montaggio dei cristalli perturba drasticamente l'ambiente chimico-fisico dei cristalli

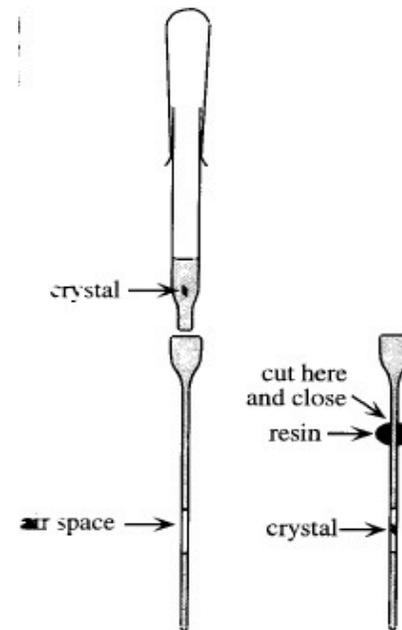
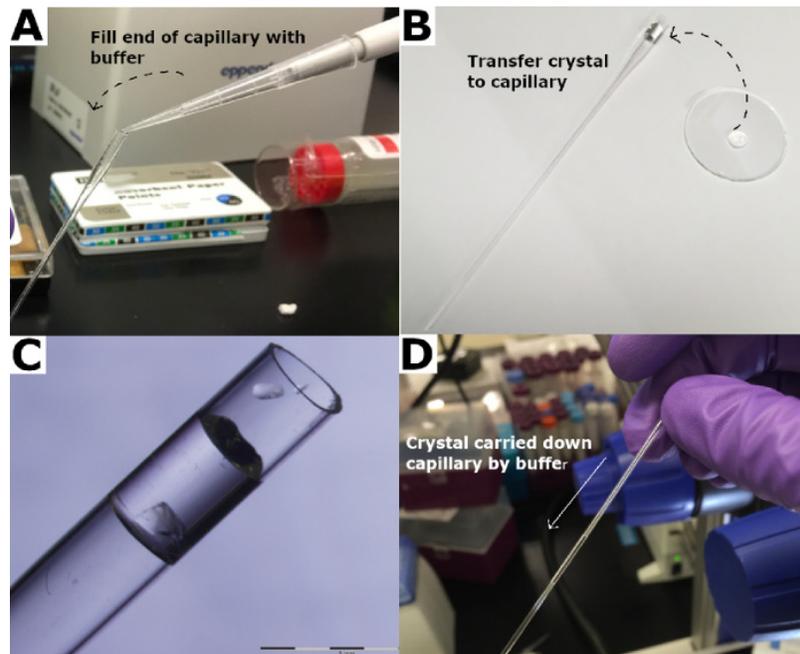


La manipolazione prolungata dei cristalli proteici tende a danneggiarli e a compromettere la qualità della diffrazione.

**Il montaggio dei cristalli e deve essere il più rapida possibile.**

# Montaggio del cristallo - capillari

I cristalli di proteina, se l'esperimento è acquisito a **temperatura ambiente**, possono essere inseriti in **capillari di quarzo** di dimensioni opportune (diametro da 0.1 a 2 mm) a seconda delle dimensioni del cristallo.

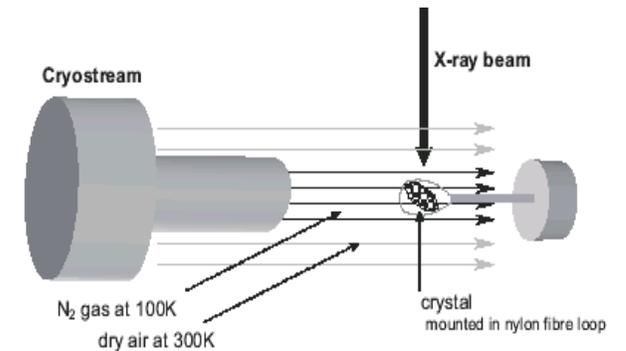


# Loops

Al giorno d'oggi è molto raro che le misure siano fatte a T ambiente poiché i cristalli si danneggerebbero immediatamente a causa del *danno da radiazione*. Le misure sono quasi sempre fatte a 100 K

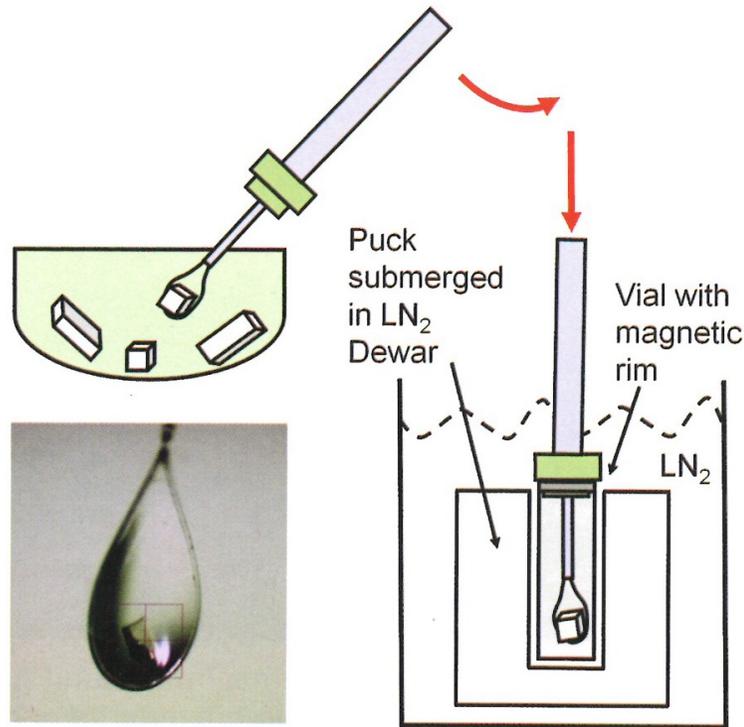
Per questo motivo si utilizzano dei 'cappi' (**loops**) di nylon con cui pescare i cristalli che poi sono congelati e mantenuti a T di 100 K

I loops hanno dimensioni e forme variabili, a seconda delle dimensioni e forma del cristallo.



# Cryocooling

Il modo usato per preservare la vita dei cristalli al fine di poter acquisire un dataset completo, è quello di congelare i cristalli in azoto liquido e di eseguire tutto l'esperimento di diffrazione a 100 K



I cristalli vengono 'pescati' con un piccolo cappio di nylon (**loop**) dalla goccia dove sono cresciuti (liquido madre), immersi brevemente in una soluzione di crioprotezione e quindi congelati.

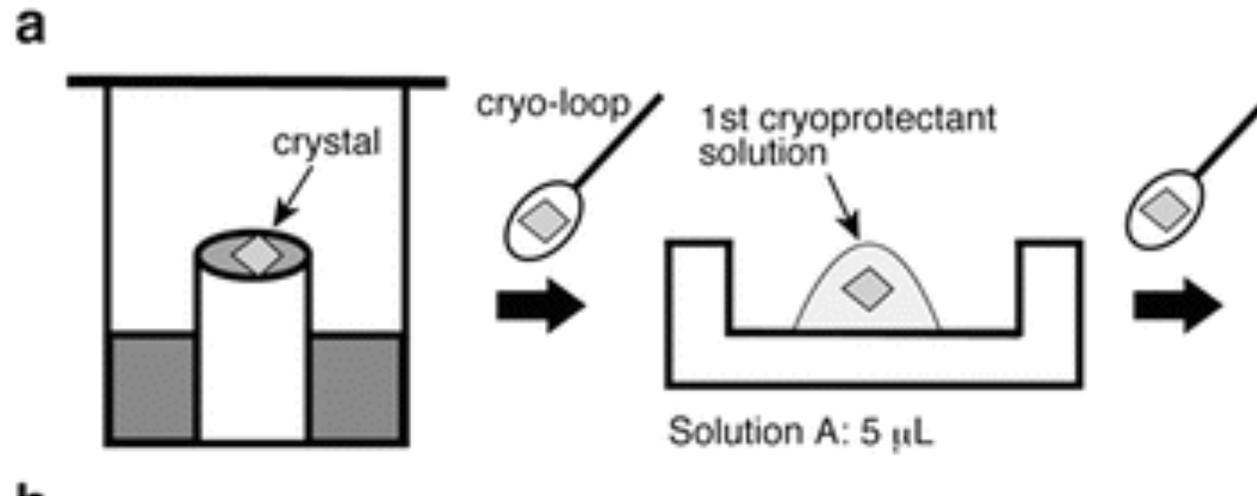
Successivamente il 'loop' viene montato sul goniometro e mantenuto ad una T di 100 K per mezzo di un criostato a flusso aperto.

Con i sincrotroni più brillanti, la vita media del cristallo è comunque limitata e bisogna attenuare il fascio di raggi-X

# Crioprotezione

A meno che il liquido madre non contenga già sostanze chimiche appropriate, la soluzione di cristallizzazione presente nel loop forma ghiaccio e danneggia il cristallo.

Per prevenire la formazione di ghiaccio, prima di essere congelato il cristallo deve essere immerso brevemente in una **soluzione di crioprotezione** che impedirà la formazione di ghiaccio.



# Crioprotettori

La composizione della soluzione di crioprotezione è variabile, difficilmente predicibile, anche se si usa un numero limitato di composti 'tipici'.

Il più delle volte si può partire dal liquido madre stesso a cui è addizionata in proporzione variabile (15 -40 %) un **crioprotettore** opportuno:

- **Glicerolo (20 – 40 %)**
- **glicol etilenico (10 -40 %)**
- **Saccarosio**
- **Trealosio**
- **PEG a basso peso molecolare (200 - 600)**
- **NaCl e LiCl ad alta concentrazione (3-5 M)**
- ....

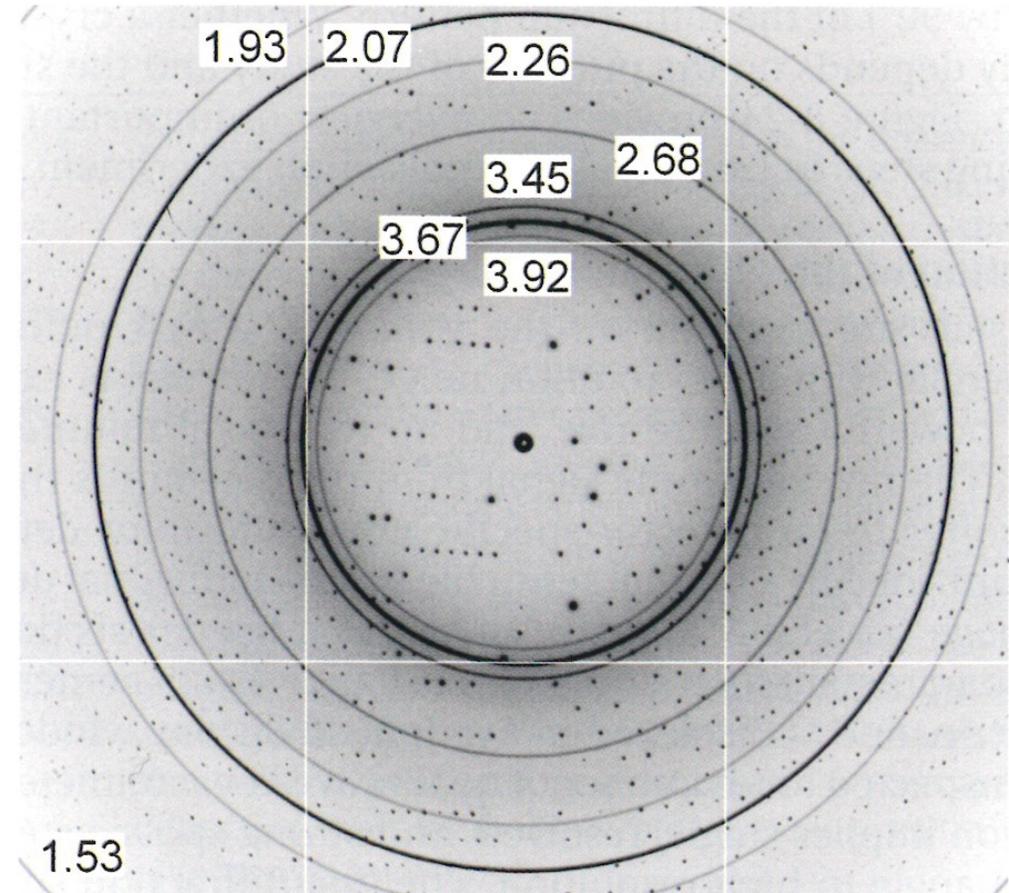
Si procede per *trial and error* a partire da scelte 'classiche' (Glicerolo o Glicol Etilenico al 20%)

# Problemi con la crioprotezione

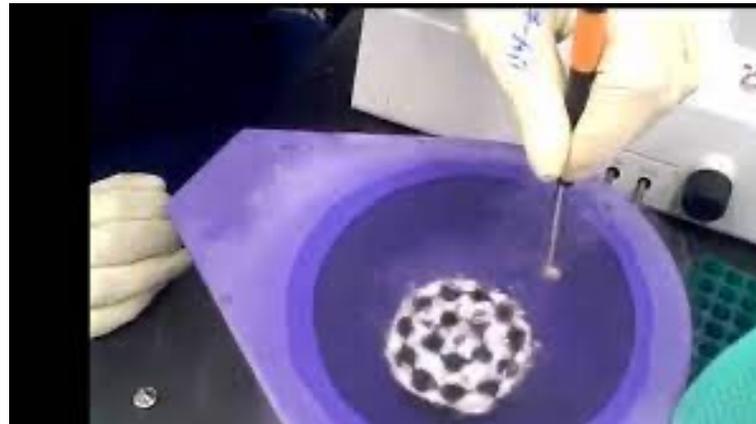
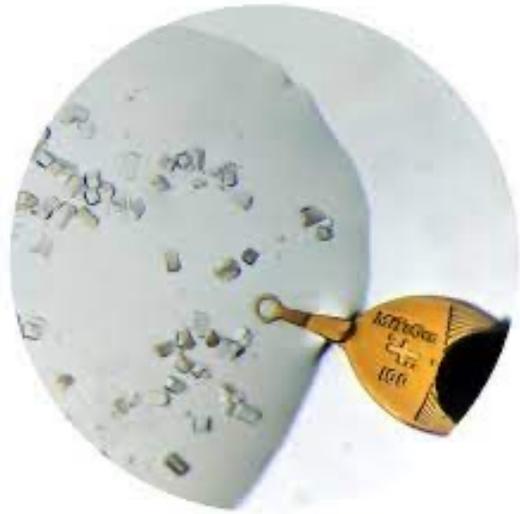
Una scelta errata del crioprotettore si manifesta con 'anelli' nell'immagine dovuti alla diffrazione di micro-cristalli di ghiaccio

Una crioprotezione non opportuna può peggiorare la capacità dei cristalli di diffrangere i raggi-X

La scelta di una crioprotezione opportuna può essere complessa

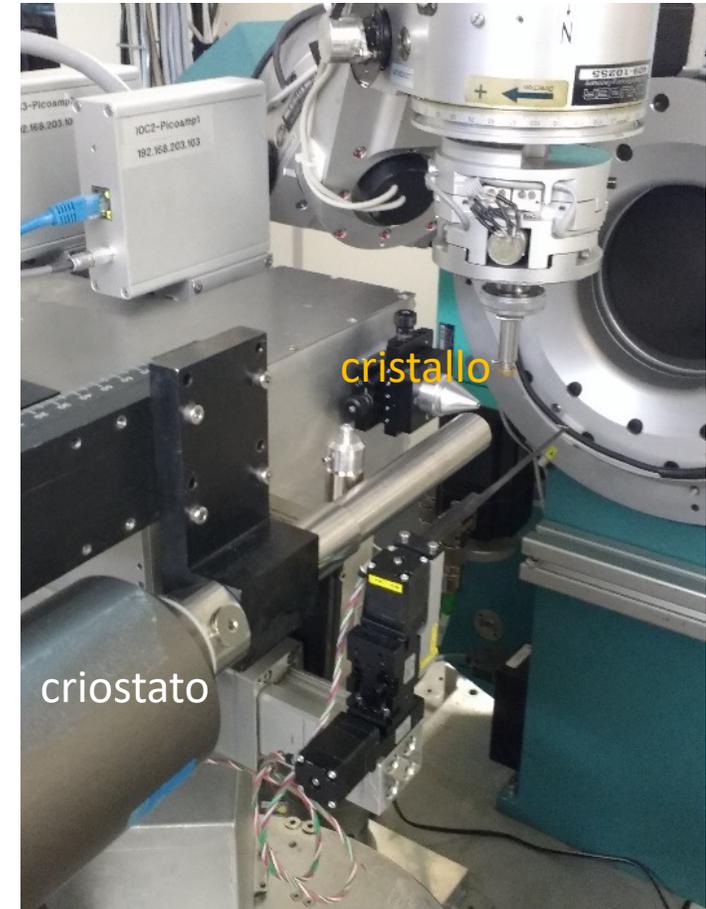


# Conservazione dei cristalli congelati



# Montaggio automatico

I cristalli congelati sono mantenuti a temperatura criogenica e sono montati sul goniometro per mezzo di un sistema robotico controllato remotamente



# Modalità di acquisizione dati

# Modalità di acquisizione

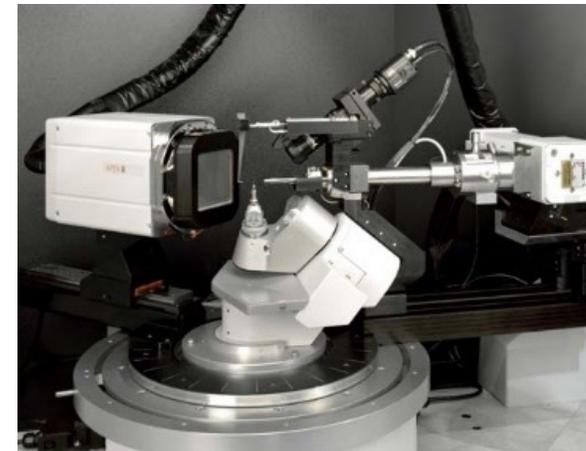
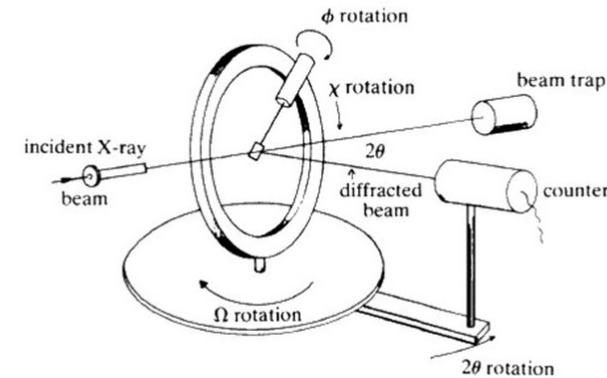
Il cristallo va ruotato in modo opportuno per portare i piani di Bragg in condizioni di diffrazione.

$$n\lambda = 2d_{hkl}\sin\vartheta \quad d_{hkl} \rightarrow \vartheta_{hkl}$$

Possiamo operare in due modi:

1. Misurare una intensità alla volta con un rivelatore monodimensionale (rivelatore puntuale)
2. Misurare più intensità 'contemporaneamente' utilizzando un rivelatore bidimensionale

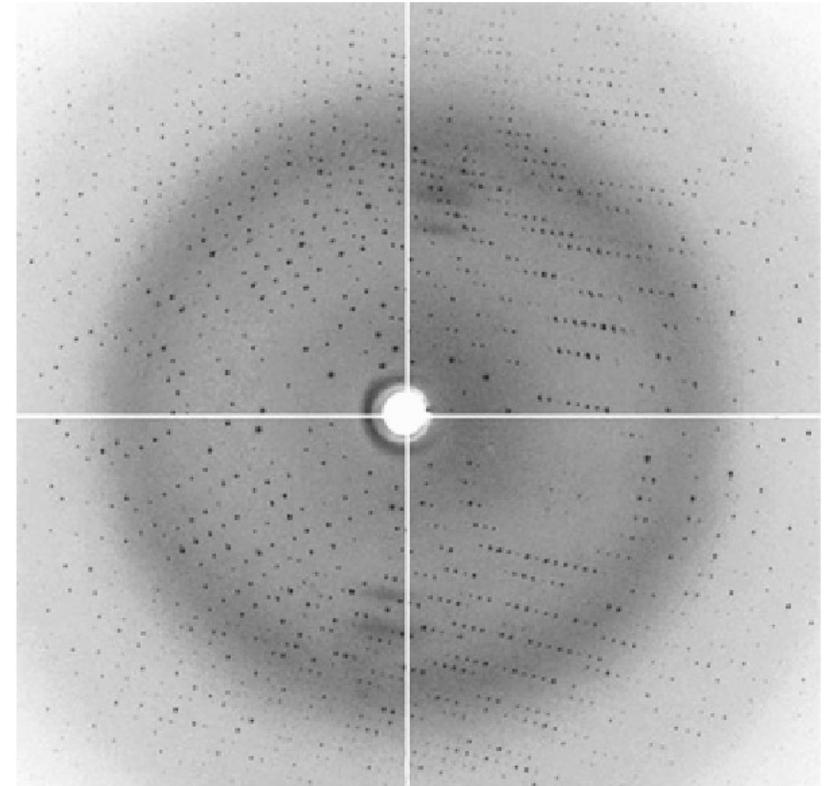
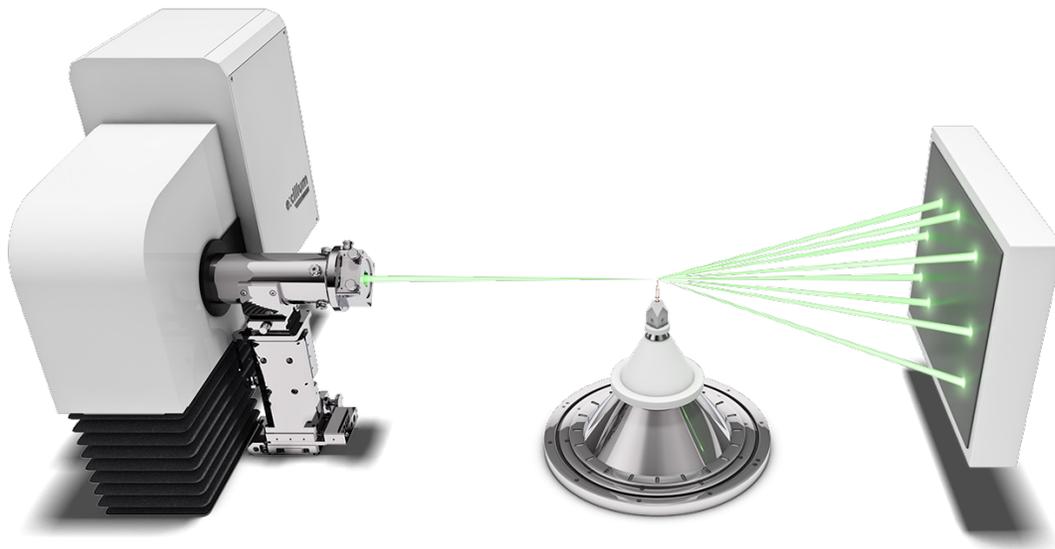
L'approccio 1 è più accurato, l'approccio 2 è più rapido



# Acquisizione dati per le proteine

I cristalli di proteine hanno celle molto grandi e di conseguenza vanno misurate molte intensità.

**Misuriamo più intensità contemporaneamente**

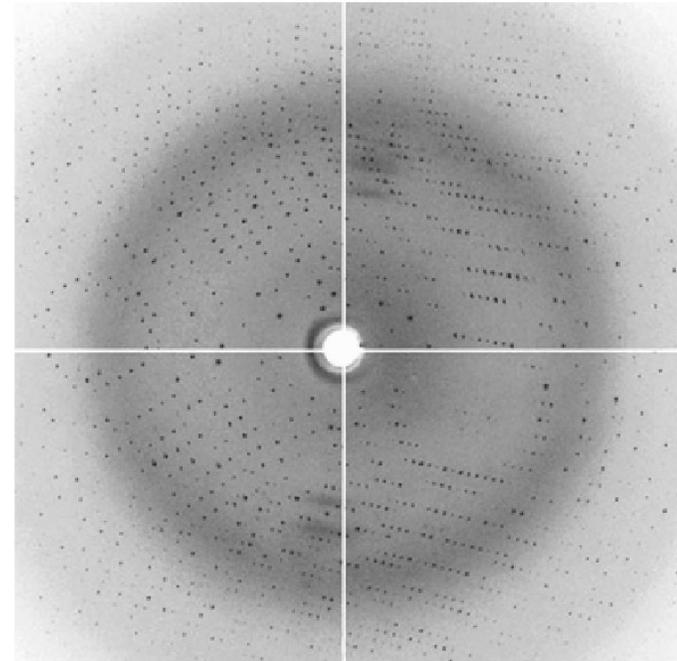
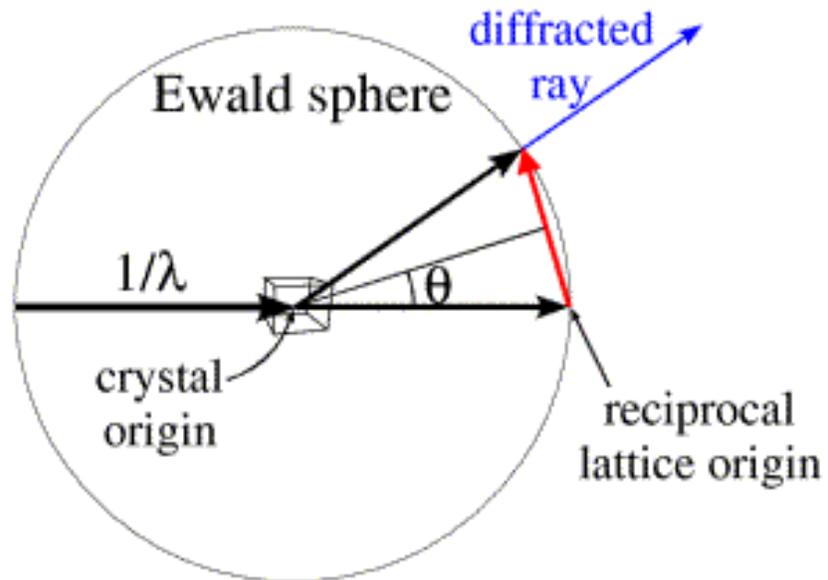




# Sfera di Ewald - 3

Se ruotiamo cristallo portiamo i diversi piani di bragg in condizioni di diffrazione

In un cristallo di macromolecole biologiche, con una piccola rotazione, dell'ordine di  $0.1-1^\circ$ , portiamo in condizioni di diffrazione molti piani di Bragg. I raggi diffratti saranno 'registrati' da un opportuno detector

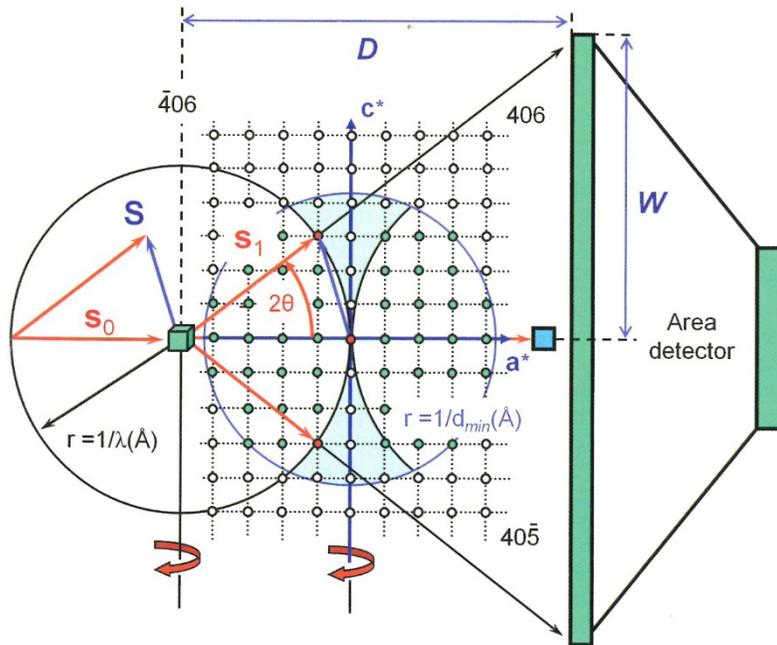
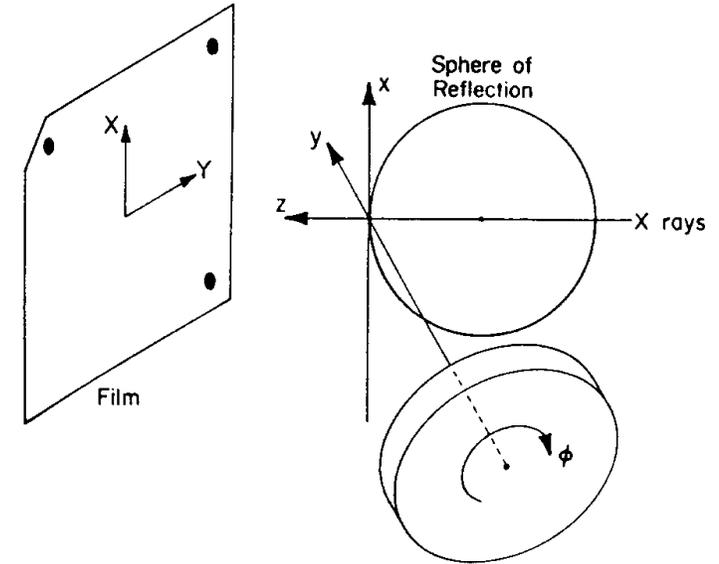


# Metodo del cristallo oscillante

# Metodo del cristallo oscillante - 1

Il cristallo è montato e centrato su un goniometro, perpendicolarmente alla direzione dei raggi-X e parallelamente al piano del rivelatore.

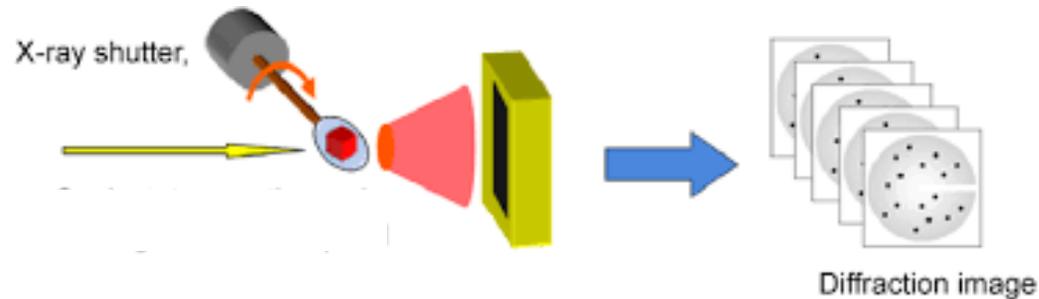
Il goniometro viene fatto **ruotare di un certa entità ( $\Delta\phi$ )**, tipicamente  $0.1-1^\circ$  e contemporaneamente viene esposto ai raggi-



**Tutti i piani cristallini che nel corso della rotazione soddisferanno la legge di Bragg e daranno luogo a diffrazione.**

I raggi diffratti saranno intercettati dal rivelatore bidimensionale e al termine della rotazione verrà registrata una **immagine di diffrazione**.

# Metodo del cristallo oscillante - 2

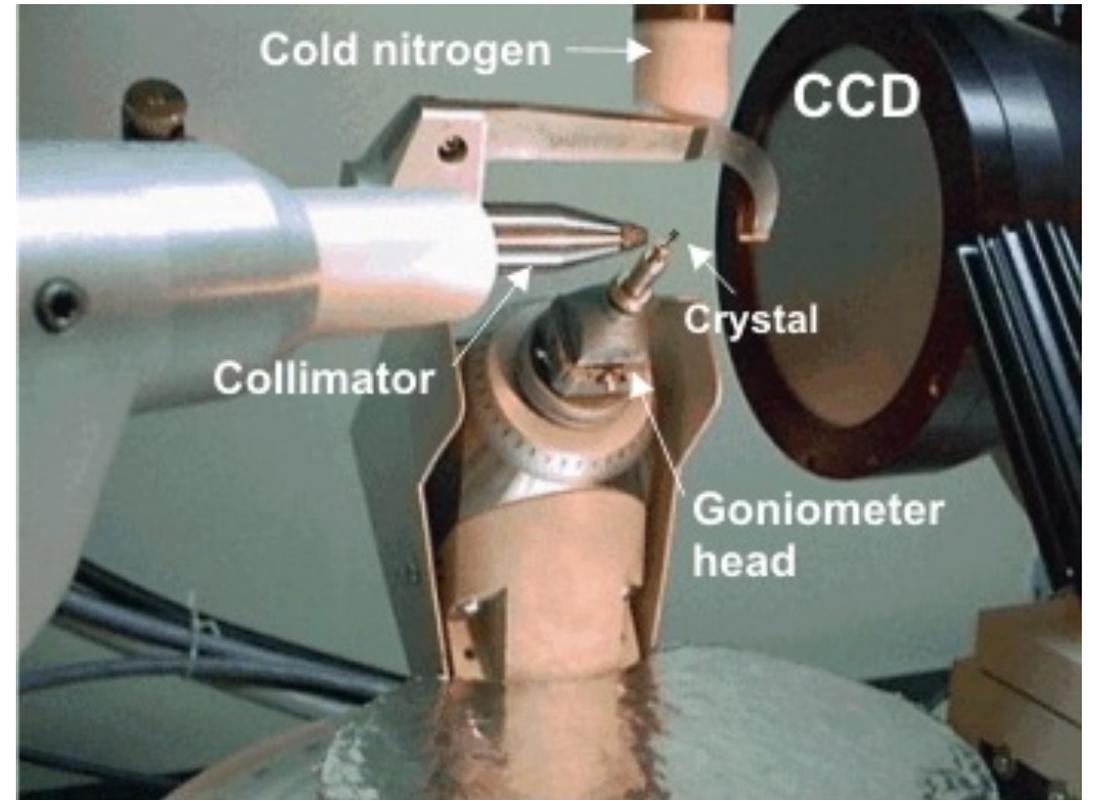
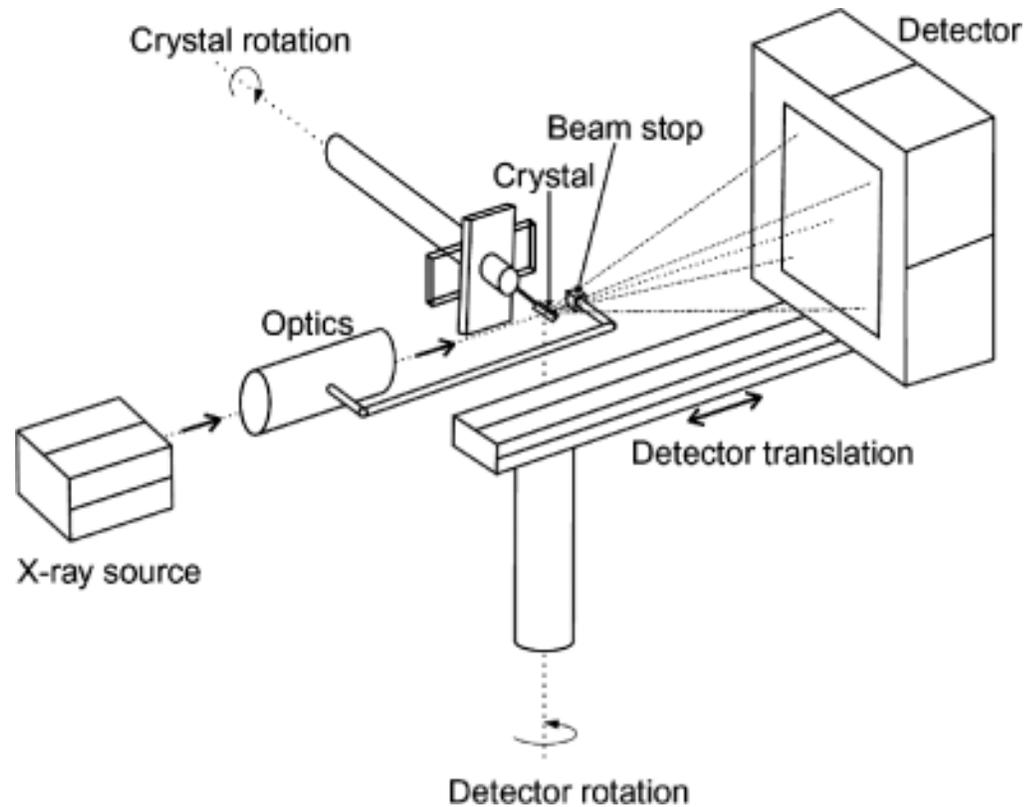


Al termine della prima rotazione il cristallo viene nuovamente esposto ai raggi-X e il cristallo ruotato nuovamente di  $\Delta\varphi$ , al termine della nuova rotazione verrà prodotta una nuova immagine.

Questa procedura verrà replicata  $N$  volte fino a coprire un angolo totale  $\varphi_{tot}$  producendo alla fine  $N$  immagini di diffrazione

$$N = \frac{\varphi_{tot}}{\Delta\varphi}$$

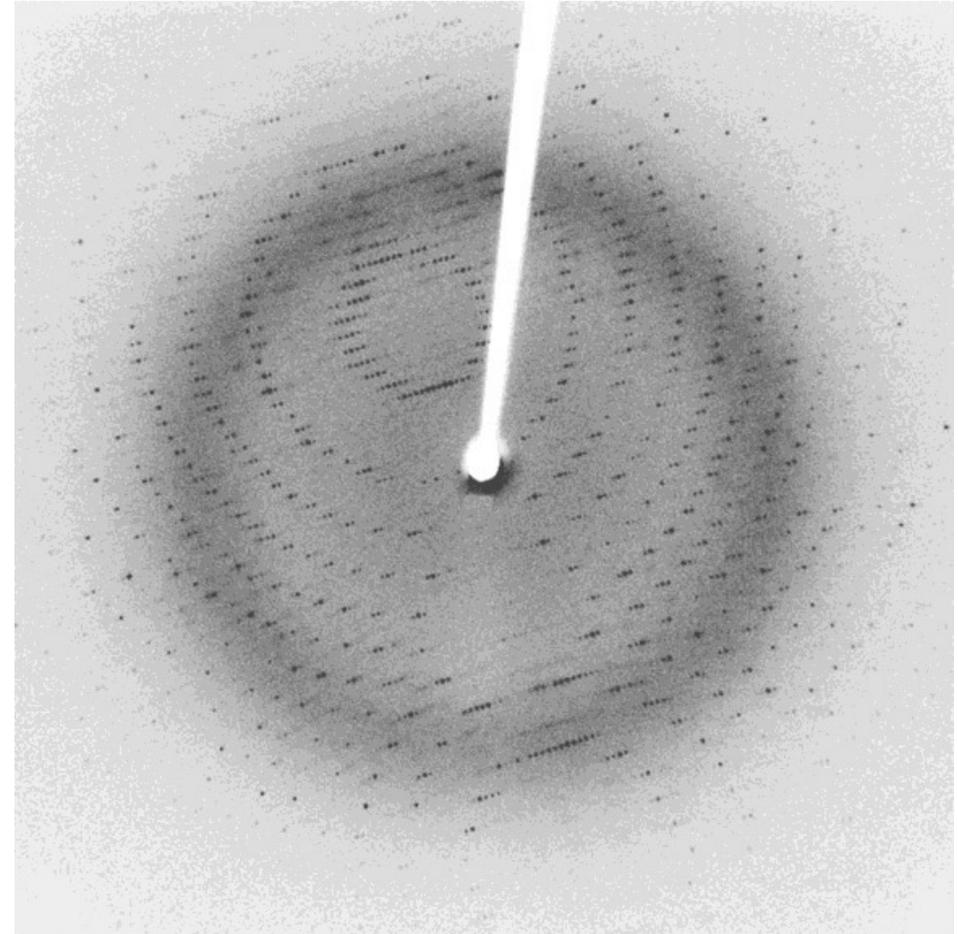
# Metodo del cristallo oscillante - 3



# Immagine di diffrazione

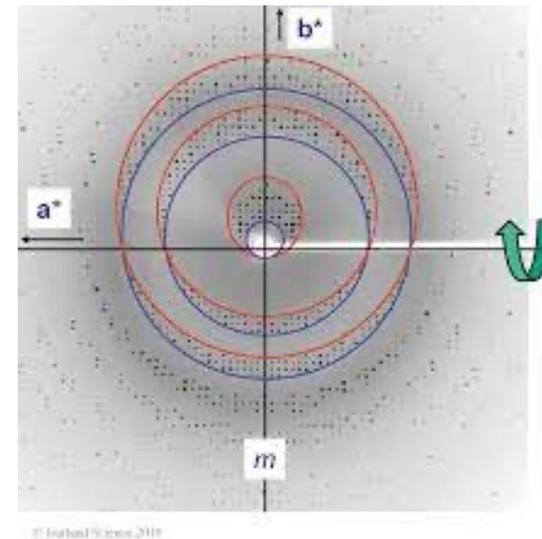
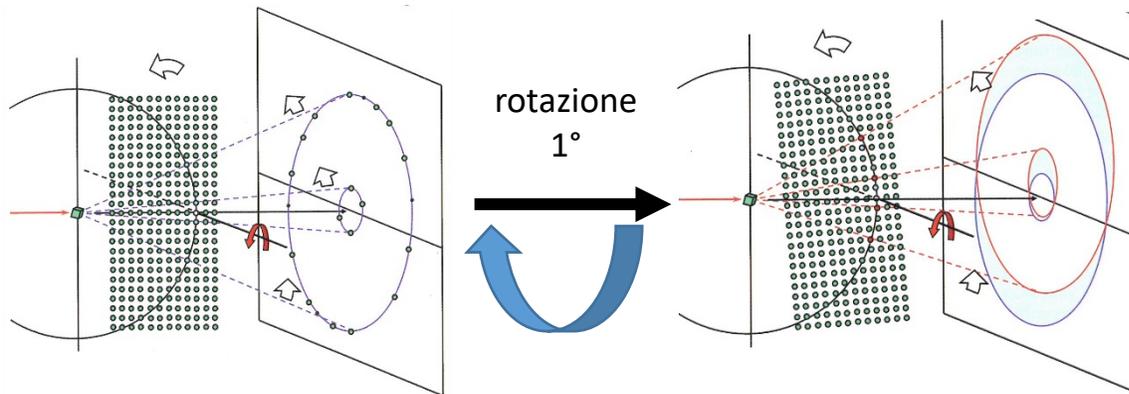
Ogni singola 'macchia' corrisponde ad un raggio diffratto da un determinato piano di Bragg di indici (h k l), o in modo equivalente da un nodo del reticolo reciproco con gli stessi indici (h k l), secondo la costruzione di Ewald.

La distribuzione delle intensità diffratte sull'area del rivelatore, dipenderà dall'orientamento del sistema di riferimento del cristallo, ovvero la cella unitaria, rispetto al sistema di riferimento del laboratorio. Dipenderà inoltre dalla simmetria del reticolo cristallino



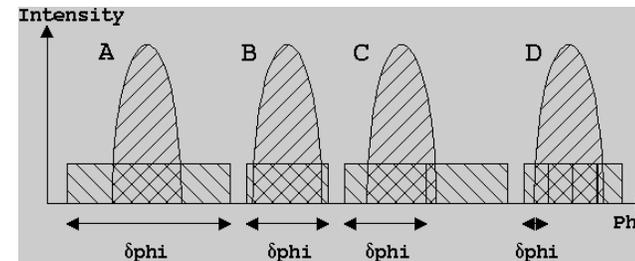
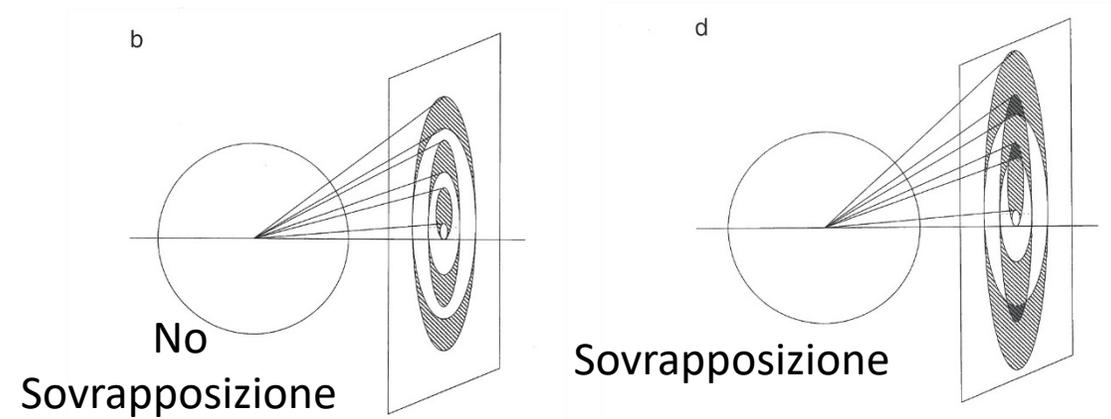
# Metodo del cristallo oscillante - 3

Poiché la cella unitaria dei cristalli di macromolecole è piuttosto grande (50-100 Å), una rotazione di circa  $1^\circ$  farà in modo che un numero piuttosto elevato di piani di Bragg siano portati in condizioni di diffrazione nel corso della rotazione.



# Scelta dell'angolo di rotazione ( $\Delta\phi$ )

L'angolo di rotazione, normalmente indicato come  $\Delta\phi$ , non deve essere eccessivo per evitare che ci siano intensità («spots») sovrapposti o eccessivamente vicini (difficili da integrare)

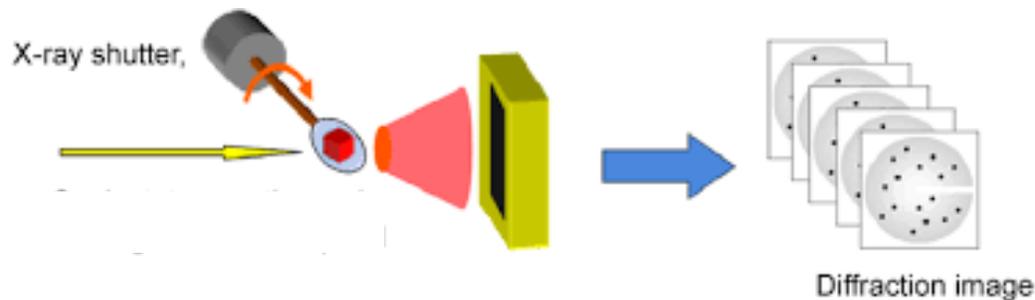


# Angolo totale ( $\varphi_{\text{tot}}$ )

La piccola rotazione di valore  $\Delta\varphi$  genererà una immagine di diffrazione ma 'esplorerà' solo una piccola parte del reticolo cristallino.

**Affinché si acquisiscano tutte le intensità necessarie alla ricostruzione della  $\rho(r)$ , dobbiamo acquisire quante più intensità possibili, diffratte dai rispettivi piani di Bragg. Cioè devo portare in condizioni di diffrazione tutti i piani cristallini (o nodi del reticolo reciproco) possibili (risoluzione e completezza!)**

Per fare questo **la rotazione verrà ripetuta con lo stesso intervallo angolare fino ad arrivare ad un valore massimo stabilito**, dipendente dalla simmetria del cristallo e dalla sua orientazione.



Es: voglio ruotare complessivamente di  $60^\circ$  ad intervalli di  $0.5^\circ$ : 120 immagini  
( $0^\circ \Rightarrow 0.5^\circ$ ;  $0.5^\circ \Rightarrow 1^\circ$ ; ...,  $119.5^\circ \Rightarrow 120^\circ$ )

# Angolo totale ( $\varphi_{\text{tot}}$ ) e simmetria

L'entità dell'angolo totale da 'ruotare' per ottenere tutte le possibili diffrazioni è legato alla simmetria del cristallo.

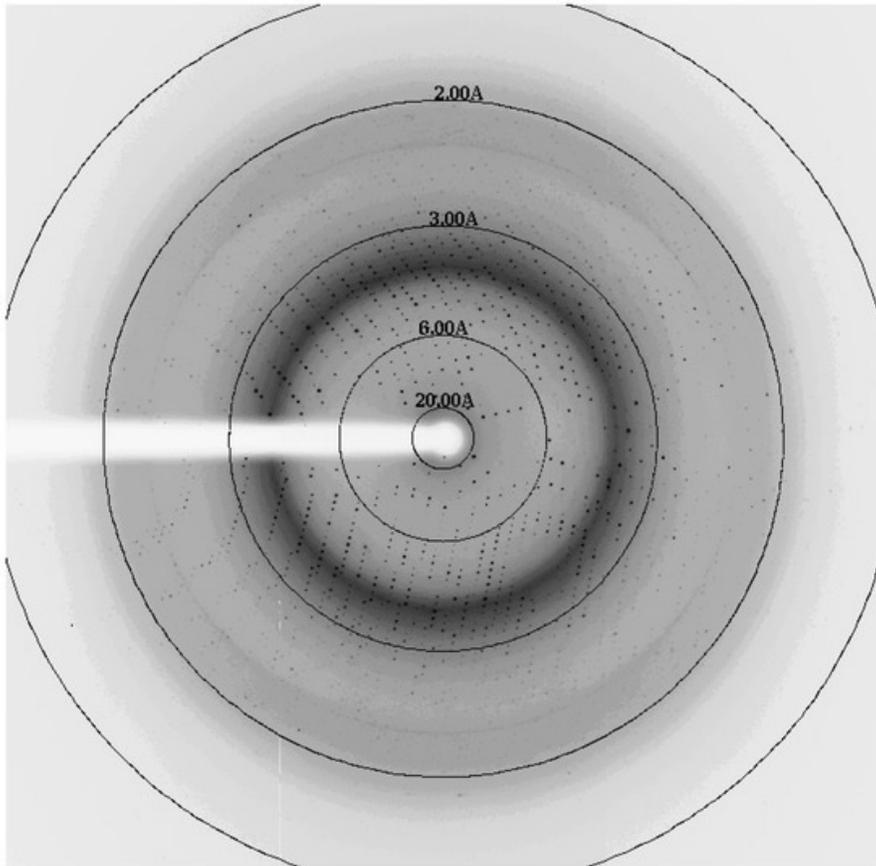
Poiché, in virtù della simmetria cristallina, le intensità di alcune diffrazioni sono identiche ad altre, l'angolo totale necessario, può essere inferiore a  $180^\circ$ .

**Se ruoto di  $360^\circ$  ottengo comunque tutti i dati possibili per una data risoluzione, lunghezza d'onda e orientazione**

TABLE II  
TOTAL ROTATION REQUIRED FOR COMPLETE DATA IN CASE OF  
SYMMETRIC DETECTOR POSITION

Crystal class	Point group	Rotation required for	
		Standard data	Anomalous data
Triclinic	<i>1</i>	$180^\circ$	$360^\circ$
Monoclinic	<i>2</i>	$180^\circ (b^*), 90^\circ (a^*, c^*)$	$180^\circ (a^*, b^*, c^*)$
Orthorhombic	<i>222</i>	$90^\circ (a^*, b^*, c^*)$	$90^\circ (a^*, b^*, c^*)$
Tetragonal	<i>4</i>	$90^\circ (a^*, b^*, c^*)$	$90^\circ (c^*), 180^\circ (a^*, b^*)$
	<i>422</i>	$45^\circ (c^*), 90^\circ (a^*, b^*)$	$45^\circ (c^*), 90^\circ (a^*, b^*)$
Trigonal	<i>3</i>	$60^\circ (c^*), 90^\circ (a^*, b^*)$	$120^\circ (c^*), 180^\circ (a^*, b^*)$
	<i>321</i>	$30^\circ (c^*), 90^\circ (a^*, b^*)$	$60^\circ (c^*), 180^\circ (a^*, b^*)$
	<i>312</i>	$30^\circ (c^*), 90^\circ (a^*, b^*)$	$60^\circ (c^*), 180^\circ (a^*, b^*)$
Hexagonal	<i>6</i>	$60^\circ (c^*), 90^\circ (a^*, b^*)$	$60^\circ (c^*), 180^\circ (a^*, b^*)$
	<i>622</i>	$30^\circ (c^*), 90^\circ (a^*, b^*)$	$30^\circ (c^*), 90^\circ (a^*, b^*)$
Cubic	<i>23</i>	About $60^\circ$	About $70^\circ$
	<i>432</i>	About $35^\circ$	About $45^\circ$

# Posizione del rivelatore - 1

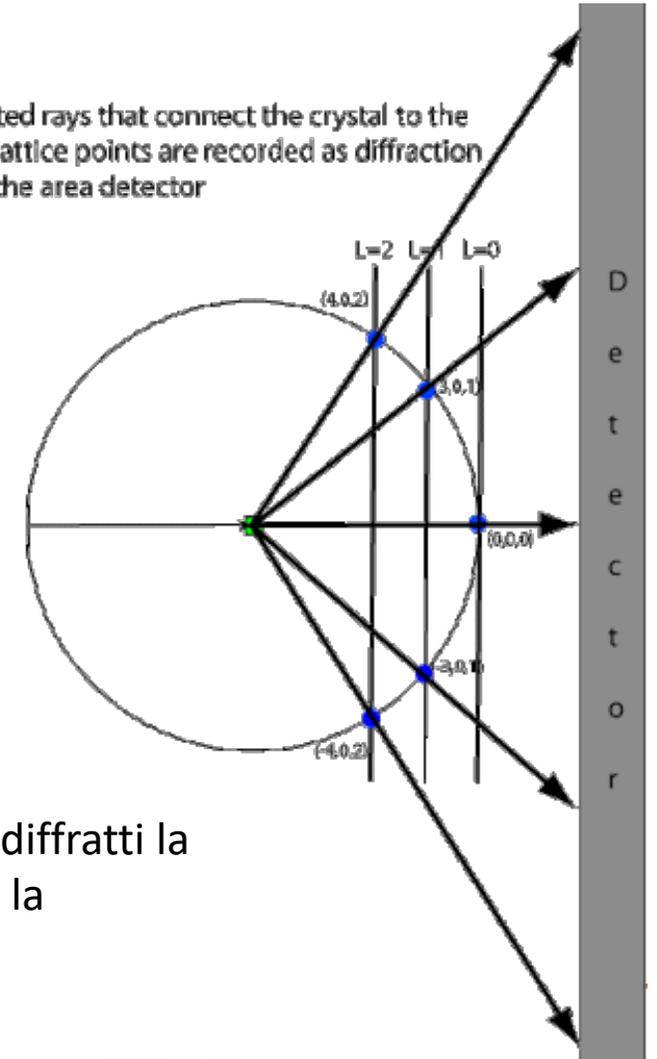


Più il detector è vicino al cristallo, maggiore sarà il valore massimo dell'angolo di diffrazione accettato e quindi la risoluzione massima ottenibile.

La posizione del rivelatore dovrà essere tale da distribuire i raggi diffratti su tutta la superficie del rivelatore.

E' importante acquisire tutti i raggi diffratti la cui intensità è misurabile (definisce la risoluzione)

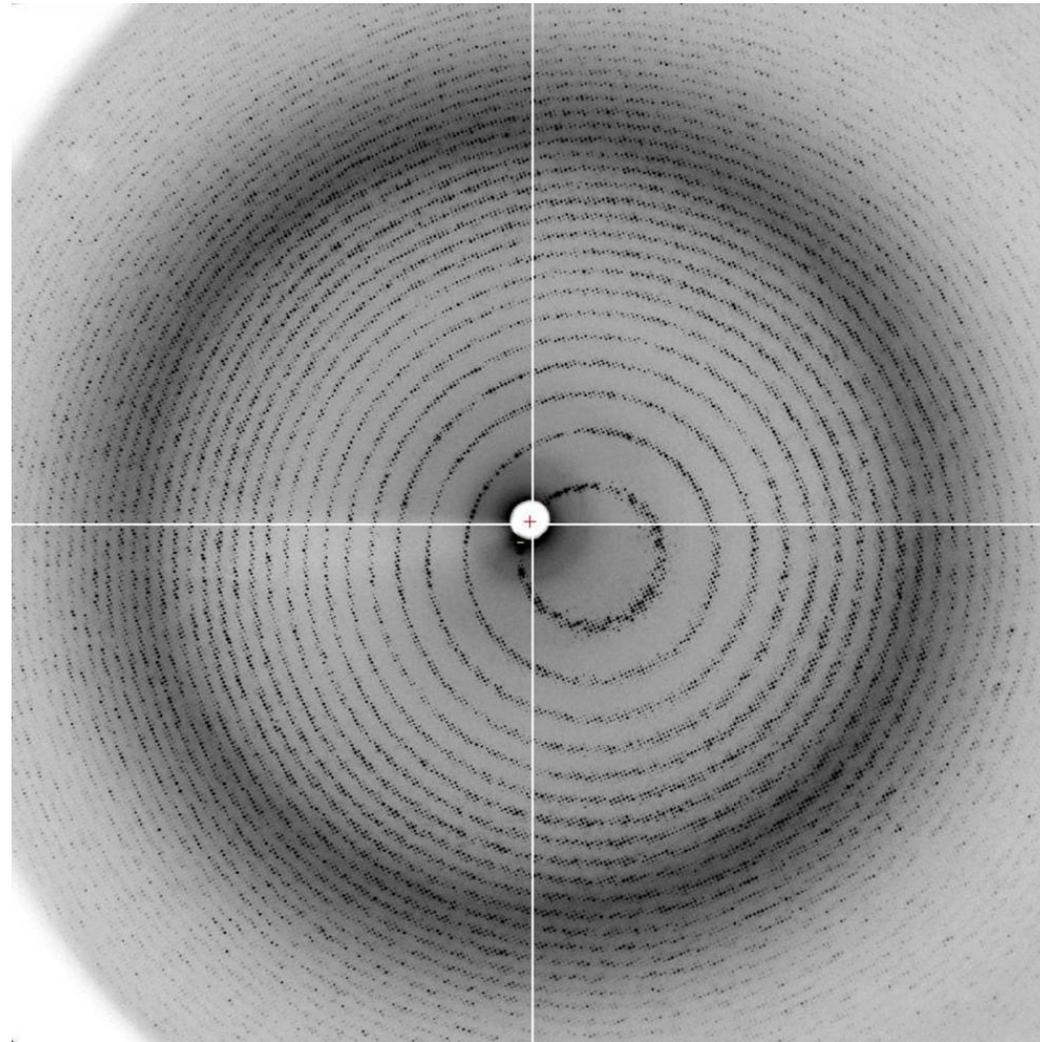
The diffracted rays that connect the crystal to the reciprocal lattice points are recorded as diffraction "spots" on the area detector



## Posizione del rivelatore - 2

Non sempre è necessario avere un detector molto vicino al cristallo, perché il cristallo può diffrangere solo ad angoli relativamente piccoli (bassa risoluzione)

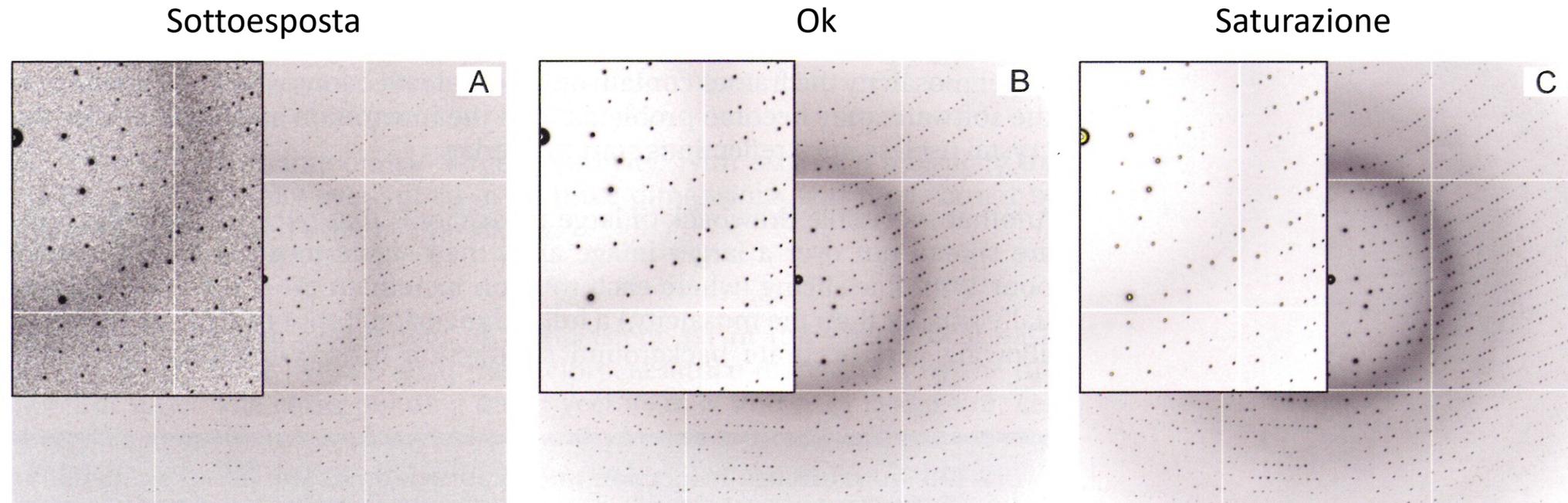
Inoltre se la struttura cristallina ha una cella unitaria molto grande (esempio virus, ribosomi...) l'immagine di diffrazione sarà densamente popolata e le diffrazioni rischiano di 'sovrapporsi'



# Tempo di esposizione

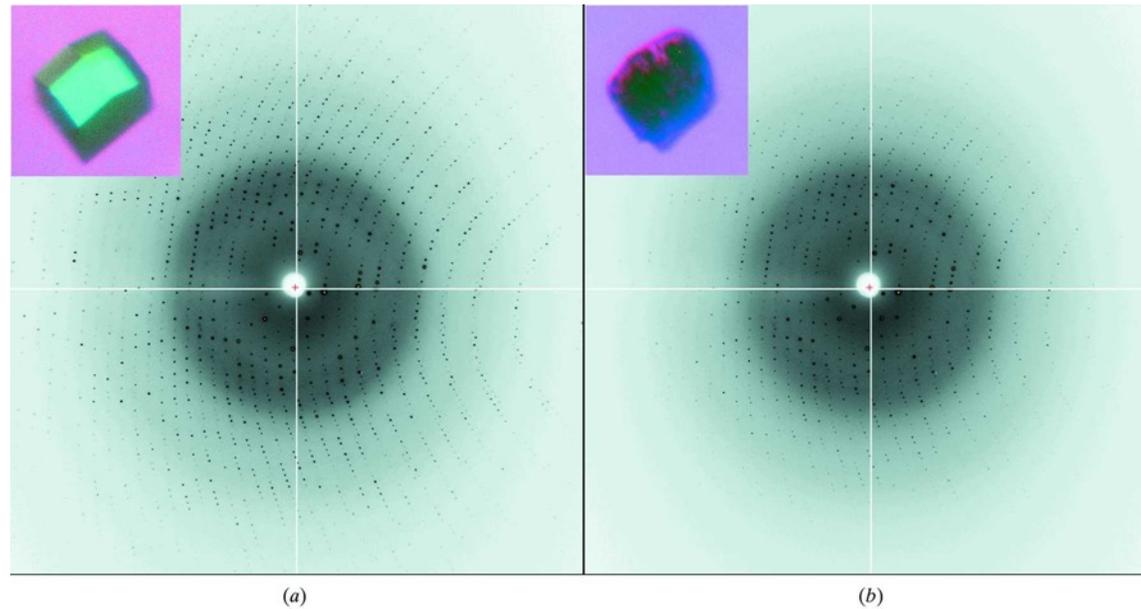
Il tempo di esposizione deve essere adeguato ma non deve saturare il detector.  
Va inoltre considerato il 'danno da radiazione'.

Il tempo di esposizione è un parametro largamente empirico.



# Tempo di esposizione e danno da radiazione

Un'esposizione eccessiva può causare un deterioramento 'anticipato' del cristallo, prima che siano stati acquisiti i dati necessari



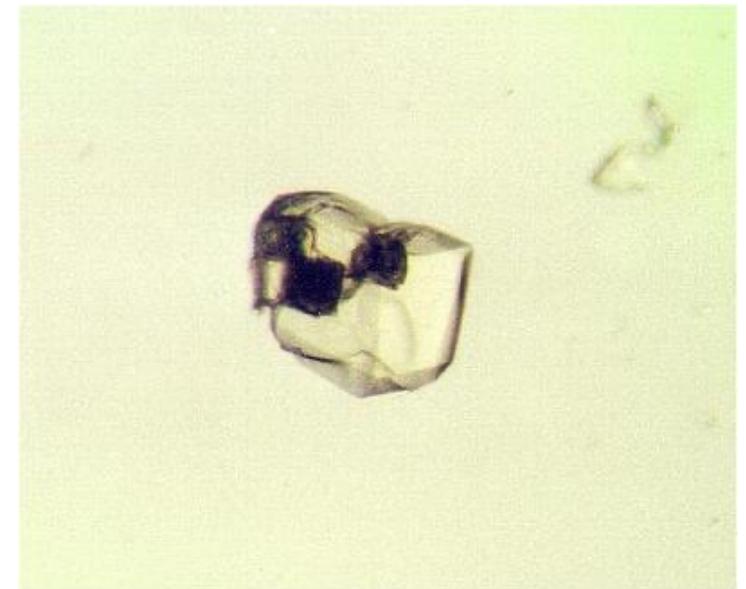
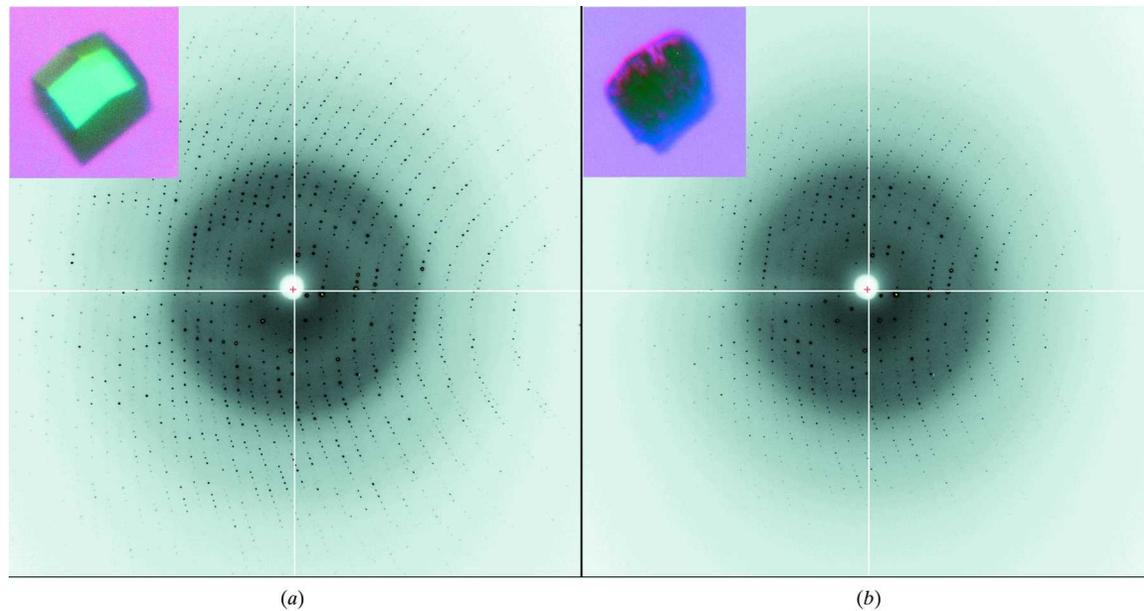
# Danno da Radiazione

# Danno da Radiazione

I moderni sincrotroni di 3° generazione producono fasci di raggi-X estremamente intensi.

Sui cristalli proteici è trasferita una notevole quantità di energia, che causa un deterioramento del cristallo e quindi della sua capacità di diffrangere i raggi-X.

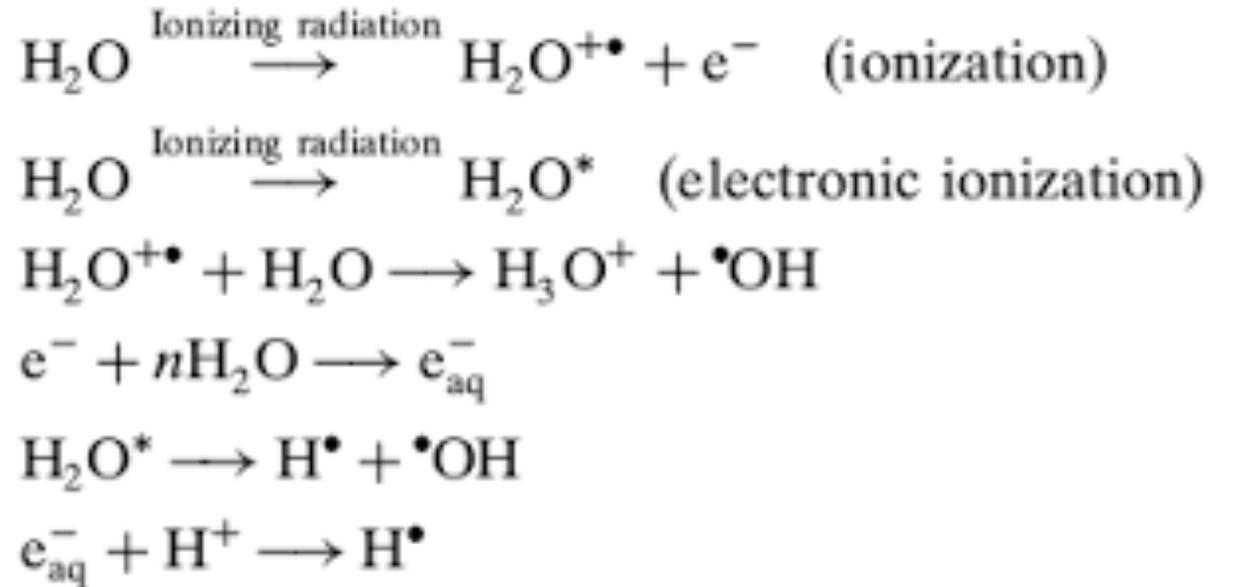
Questo fenomeno è noto come **danno da radiazione** (Radiation Damage).



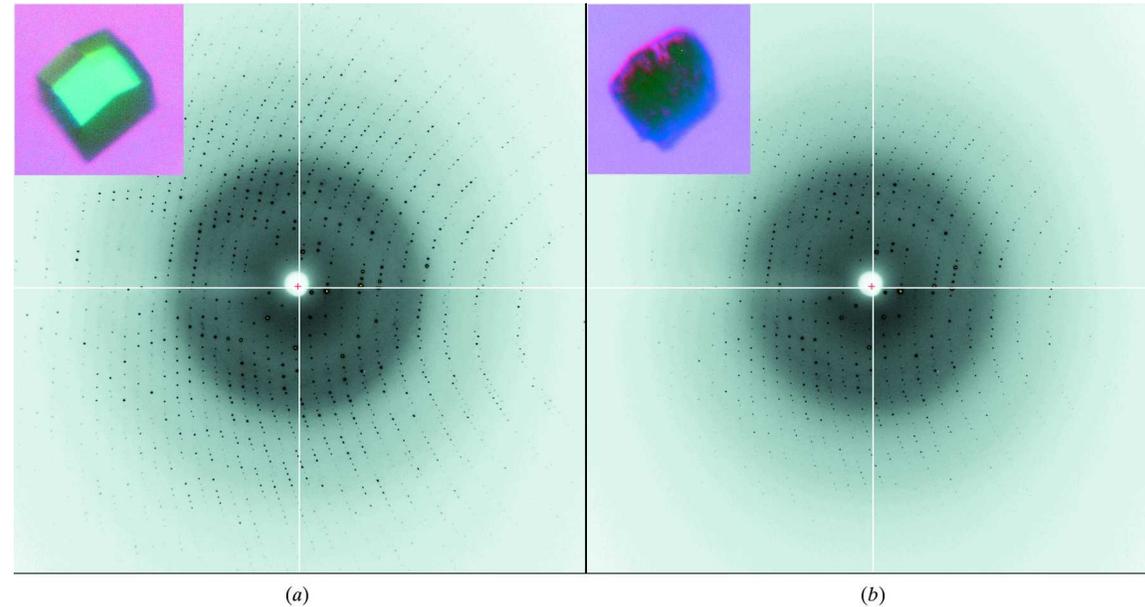
# Meccanismo del danno da radiazioni

Il danno da radiazioni è sia diretto (ionizzazione degli atomi) che dovuto alla formazione di radicali dovuti a radiolisi dell'acqua.

I radicali reagiscono rapidamente con la proteina, danneggiandola e degradando la qualità del cristallo



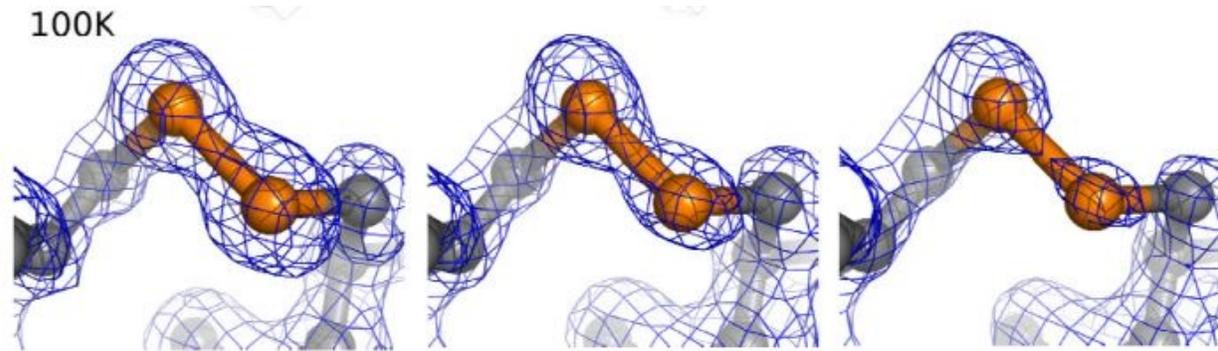
# Danno da radiazione - diffrazione



Il primo effetto visibile del danno da radiazione è il deterioramento della qualità delle immagini di diffrazione

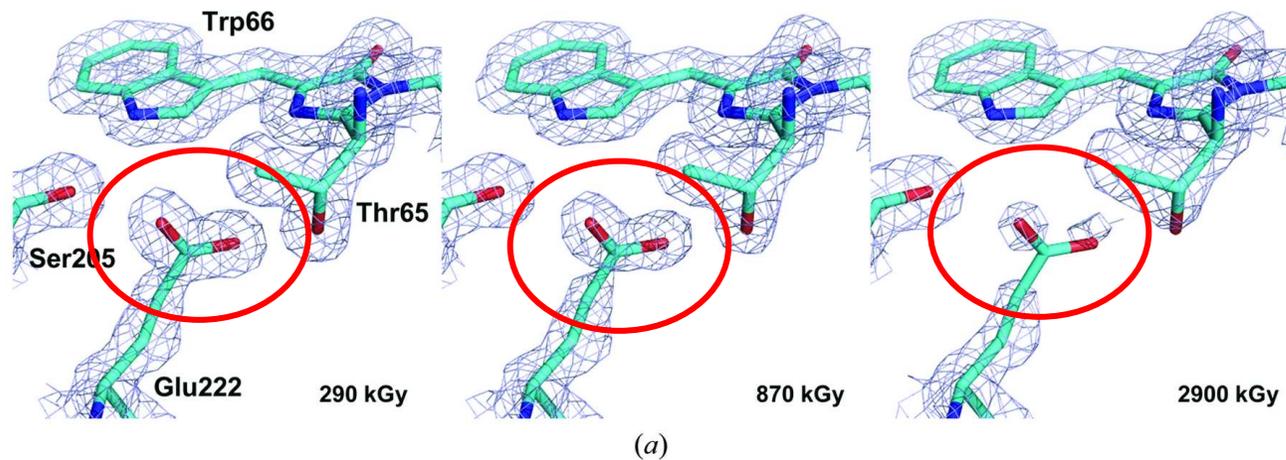
**Un cristallo danneggiato diffrange 'peggio'** cioè a più bassa risoluzione e con intensità inferiore

# Effetti del danno da radiazione



Un effetto tipico del danno da radiazioni è la rottura dei ponti di solfuro

Dose di raggi-X

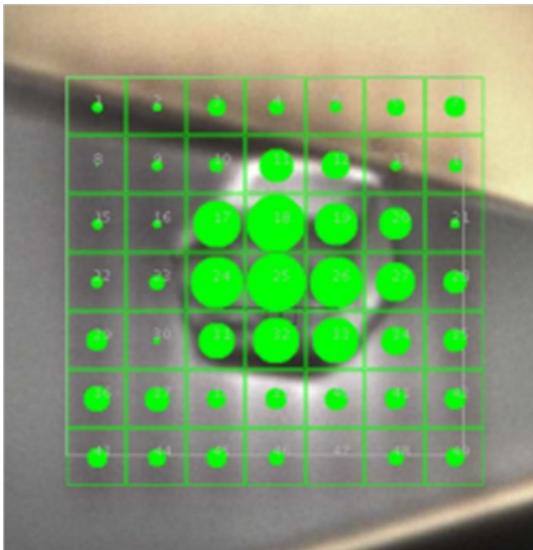


Un altro effetto tipico è la decarbossilazione degli Acidi Aspartici e Glutammici.

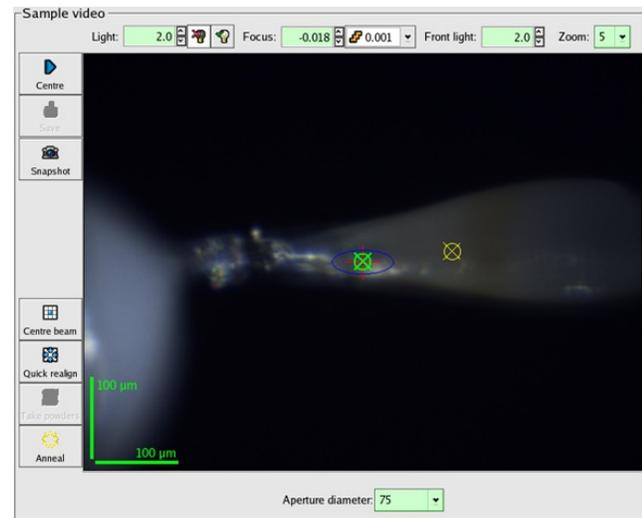
# Acquisizione con luce di sincrotrone

Con le moderne sorgenti di luce di sincrotrone, alla luce delle dimensioni molto contenute del fascio di raggi-X (pochi  $\mu\text{m}$ , microbeam) è possibile mettere in atto nuovi metodi di acquisizione.

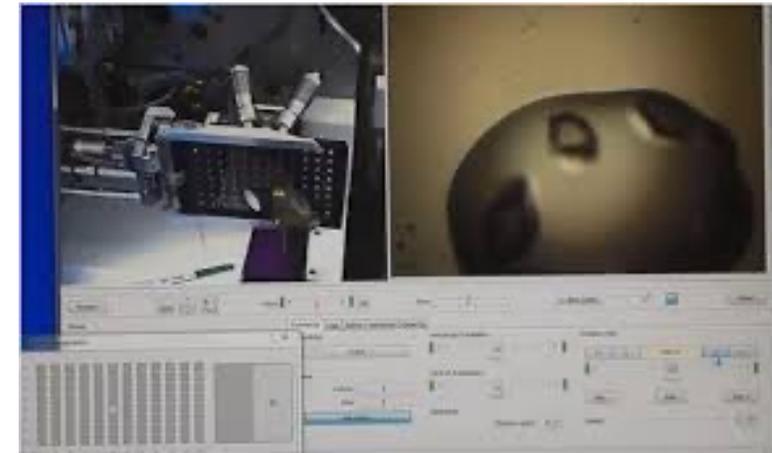
Ad esempio è possibile fare uno scan di un cristallo per centrare il punto di miglior diffrazione, o acquisire dati direttamente dal contenitore (piastra multiwell) di crescita dei cristalli.



Grid Scan



Helical Scan



In-situ Scan

# Riduzione dei dati acquisiti

# Data reduction

Con il termine **Data reduction** intendiamo un processo piuttosto complesso che permette di **trasformare le immagini di diffrazione in un file numerico contenente le intensità diffratte dai piani di Bragg**.

Per ogni immagine saranno identificate le diffrazioni e assegnati i relativi valori **h**, **k** e **l**, corrispondenti al piano di Bragg, o nodo del reticolo reciproco (è lo stesso), che l'ha generato.

Saranno successivamente valutate (*estratte*) le intensità corrispondenti.

Questa operazione viene svolta da programmi automatici che a partire dalle immagini di diffrazione generano un file con i valori di  $F(h\ k\ l)$ :

- **Mosflm**
  - **XDS**
  - **AutoPROC**
  - **HKL3000**
  - **DIALS**
-

# Analisi dei dati

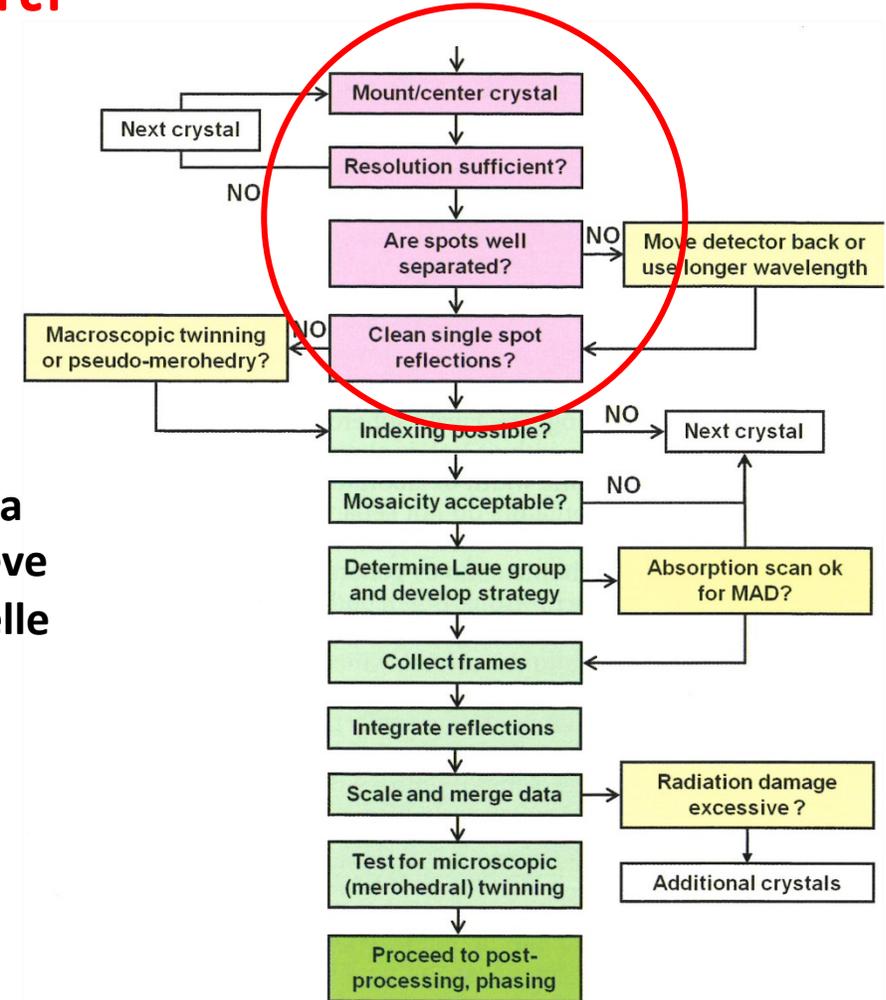
**Una volta acquisite un certo numero di immagini si fa una prima valutazione della qualità del pattern di diffrazione.**

Eventualmente si possono cambiare i parametri impostati per l'acquisizione (tempo di esposizione, angolo di oscillazione, distanza del detector, lunghezza d'onda) o addirittura passare ad un nuovo cristallo.

**Con una prima e immediata analisi cristallografica si determina la simmetria del cristallo per poter decidere, per esempio quale deve essere l'angolo totale da acquisire per avere una completezza delle intensità diffratte sufficiente.**

L'analisi completa invece si compone di:

- **Indicizzazione**
- **Integrazione**
- **Postrefinement/scalatura**



# Indicizzazione

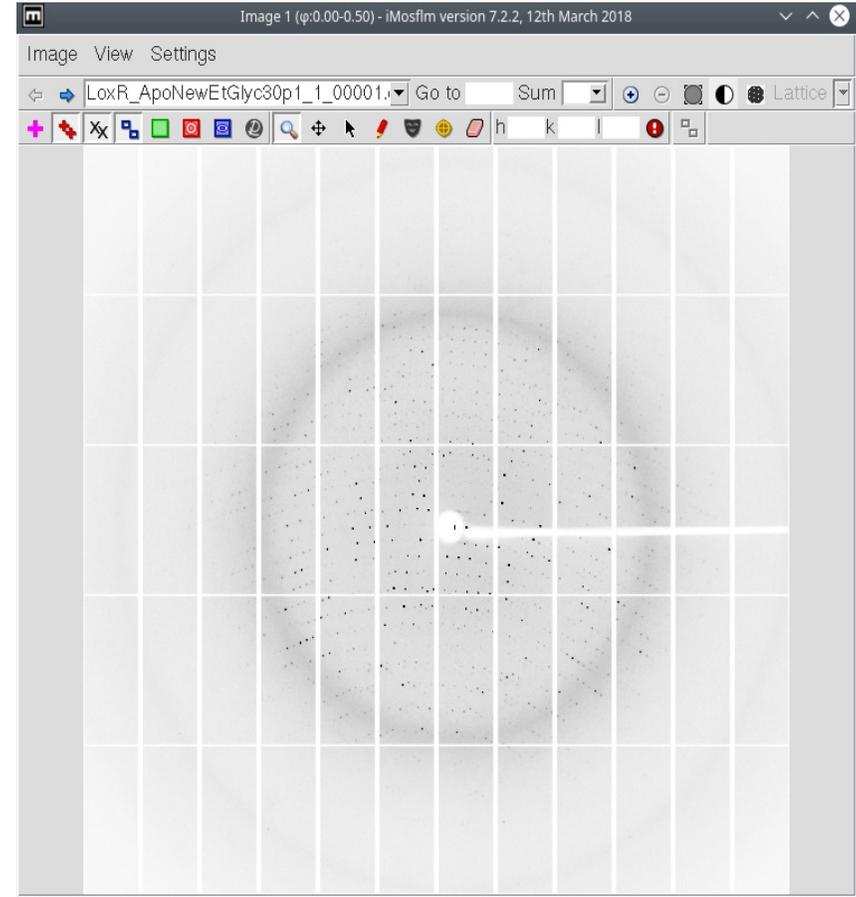
**L'indicizzazione consiste nel determinare il gruppo puntuale a cui appartiene il cristallo, ovvero**

- Il sistema cristallino (triclino, monoclino...)
- Il gruppo puntuale (in realtà è il gruppo di Laue)

1) Gli spots più intensi sono 'estratti' dalle immagini e per ogni spot sono determinate delle coordinate (x, y, z).

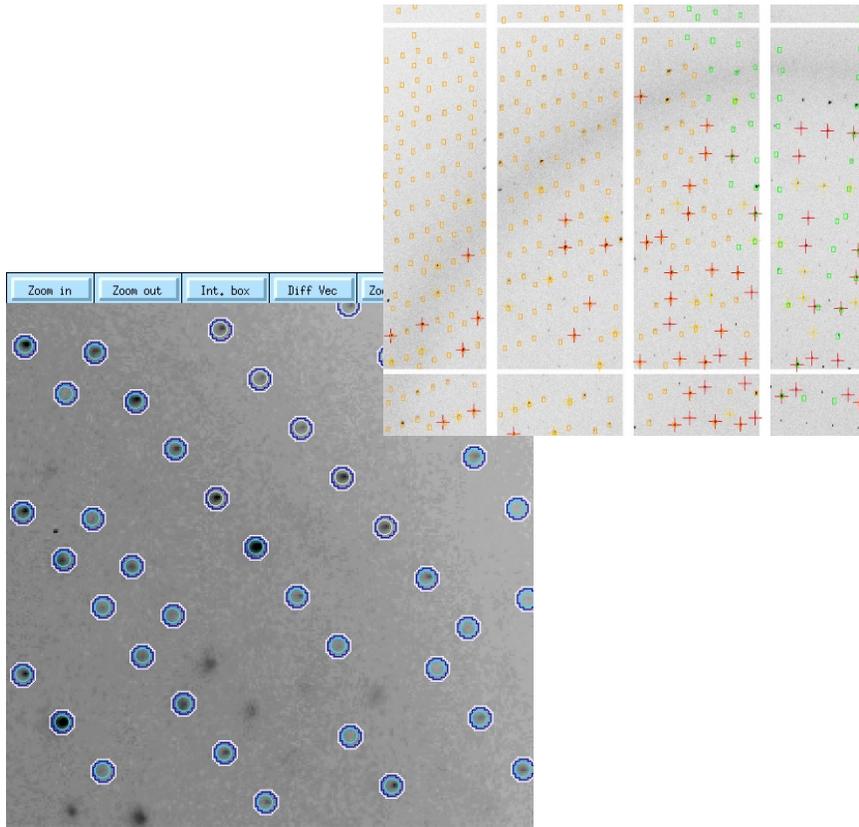
2) Viene determinata la cella elementare e la sua orientazione rispetto al sistema di riferimento in modo da poter predire la posizione degli spots (diffrazioni) osservati nell'immagine.

3) Le predizioni vengono confrontate con le posizioni sperimentali. Se la cella è corretta vi sarà accordo tra previsione e esperimento



# Indicizzazione

Si cerca la cella che predice al meglio il pattern di diffrazione osservato



iMosflm version 7.2.2, 12th March 2018

Session Settings Help

228.93 215.51 429.00 5.00 10.0 0.86 1.03 0.00 10 434

### Autoindexing

LoxR\_ApoNewEtGlyc30p1\_1\_####.cbf | 1, 180

Image	$\phi$ range	Auto	Man	Del	> I / $\sigma$ (I)	Find	Use
1	0.00 - 0.50	458	0	0	281		<input checked="" type="checkbox"/>
180	89.50 - 90.00	278	0	0	159		<input checked="" type="checkbox"/>

Total 736 0 0 440

### Lattice 1

Solution	Lat.	Pen.	a	b	c	$\alpha$	$\beta$	$\gamma$	$\sigma$ (xy)	Nref	$\delta$ beam
1 (ref)	aP	0	83.8	124.0	124.0	120.0	90.0	90.0	0.13	403	0.35 ( 0.1)
2 (ref)	mC	0	124.0	214.8	83.8	90.0	90.0	90.0	0.14	404	0.35 ( 0.2)
3 (ref)	aP	0	83.8	124.0	124.0	60.0	90.0	90.0	0.13	403	0.35 ( 0.1)
4 (ref)	mP	0	124.0	83.8	124.0	90.0	60.0	90.0	0.13	402	0.35 ( 0.1)
5 (ref)	oC	0	124.0	214.7	83.8	90.0	90.0	90.0	0.13	403	0.35 ( 0.2)
6 (ref)	mC	0	214.8	124.0	83.8	90.0	90.0	90.0	0.14	405	0.35 ( 0.2)
7 (ref)	mC	0	214.8	124.0	83.8	90.0	90.0	90.0	0.14	405	0.35 ( 0.2)
8 (ref)	mC	1	124.1	214.7	83.8	90.0	90.0	90.0	0.13	404	0.35 ( 0.2)
9 (ref)	oC	1	124.0	214.7	83.8	90.0	90.0	90.0	0.13	403	0.35 ( 0.2)
10 (ref)	hP	1	124.0	124.0	83.8	90.0	90.0	120.0	0.13	401	0.35 ( 0.1)
11 (reg)	mC	182	261.7	83.8	124.0	90.0	118.3	90.0	-	-	-
12 (reg)	mI	182	149.7	230.5	83.8	90.0	124.0	90.0	-	-	-
13 (reg)	oI	182	83.8	124.0	230.5	90.0	90.0	90.0	-	-	-
14 (reg)	mC	182	261.7	83.8	124.0	90.0	118.2	90.0	-	-	-
15 (reg)	mP	200	83.8	124.0	124.0	90.0	90.0	90.0	-	-	-

Lattices:

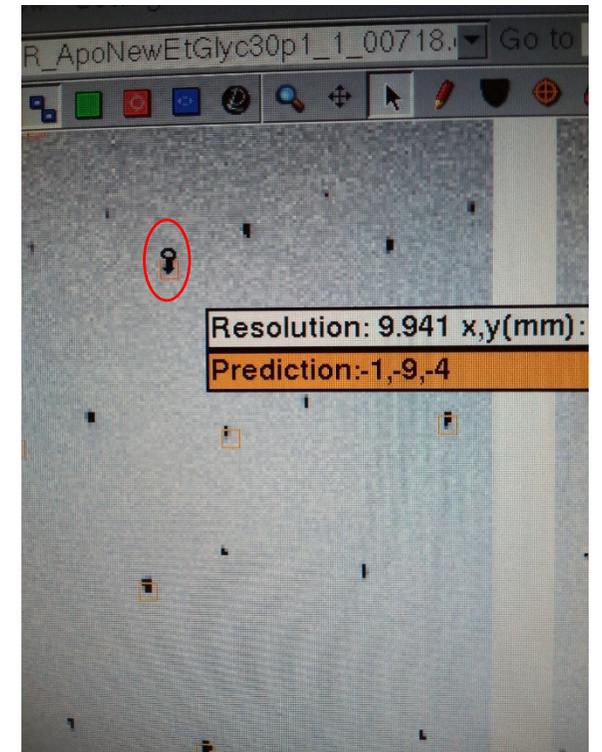
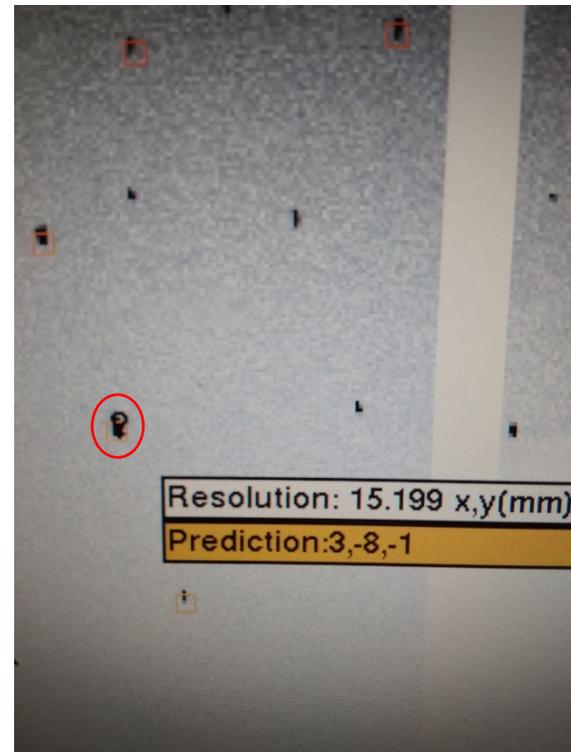
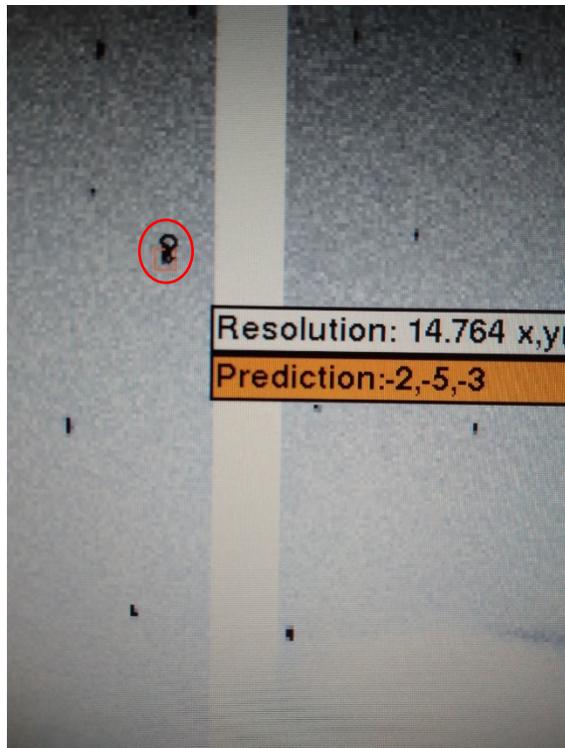
Spacegroup: P3

Mosaicity:

Green warnings 0

# Indicizzazione

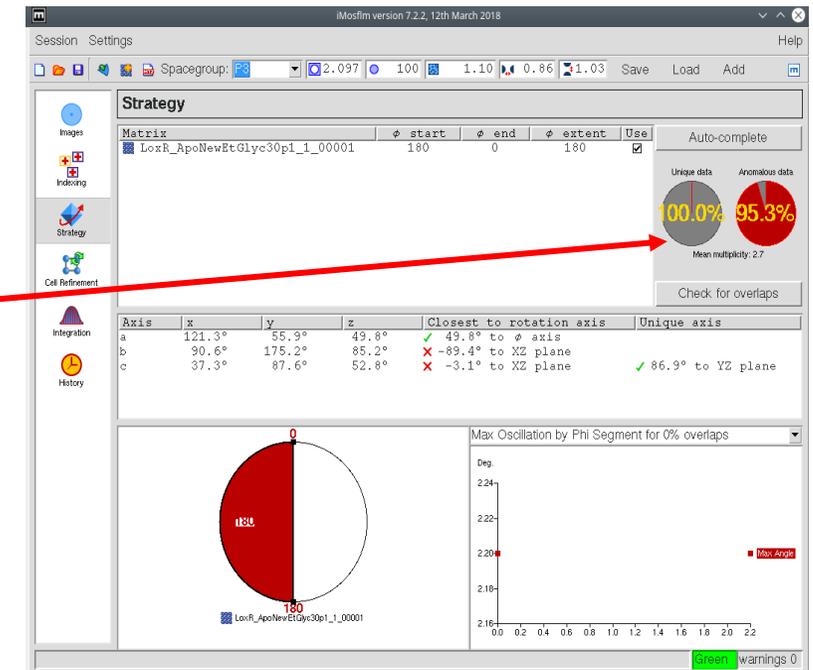
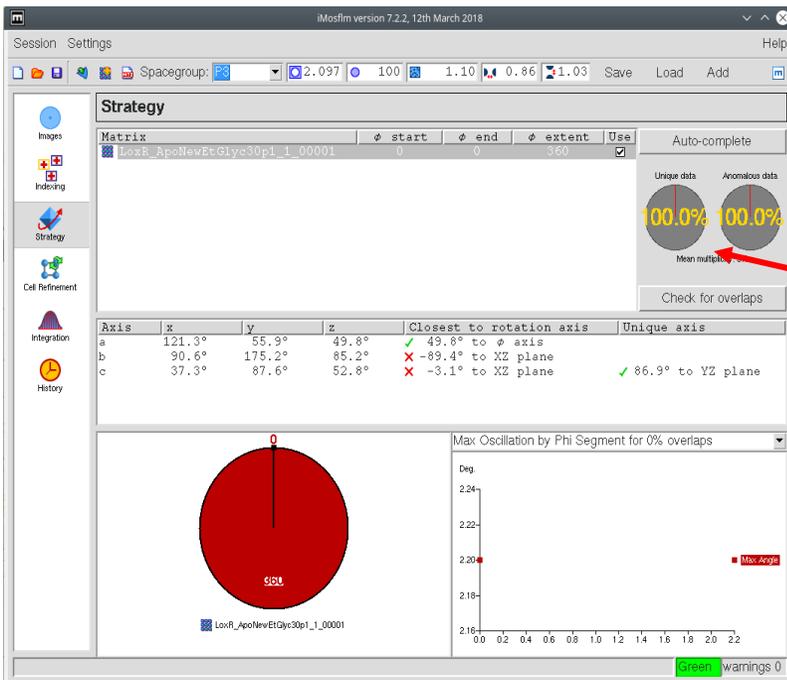
Una volta indicizzato il mio pattern di diffrazione, posso stabilire una relazione certa tra un certo spot osservato nel pattern di diffrazione e gli indici di Miller (hkl) del piano di Bragg corrispondente.



# Strategia

L'indicizzazione del mio pattern di diffrazione mi dice quale è il sistema cristallino e anche **come è orientato il mio cristallo rispetto al sistema di riferimento del laboratorio**.

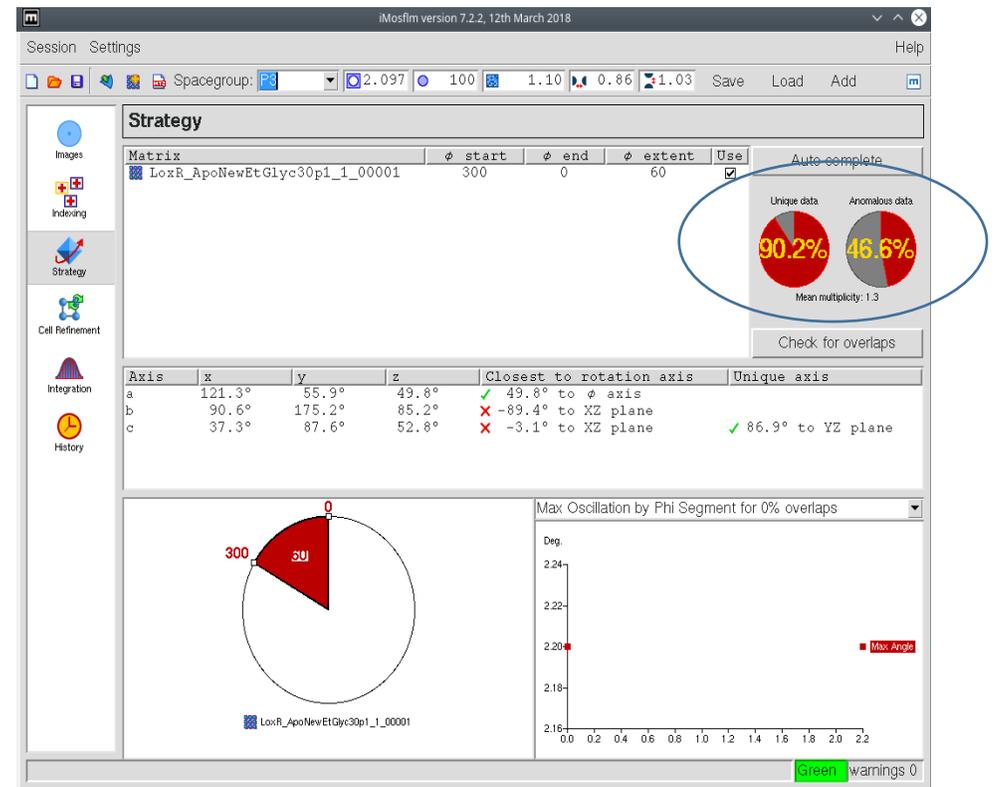
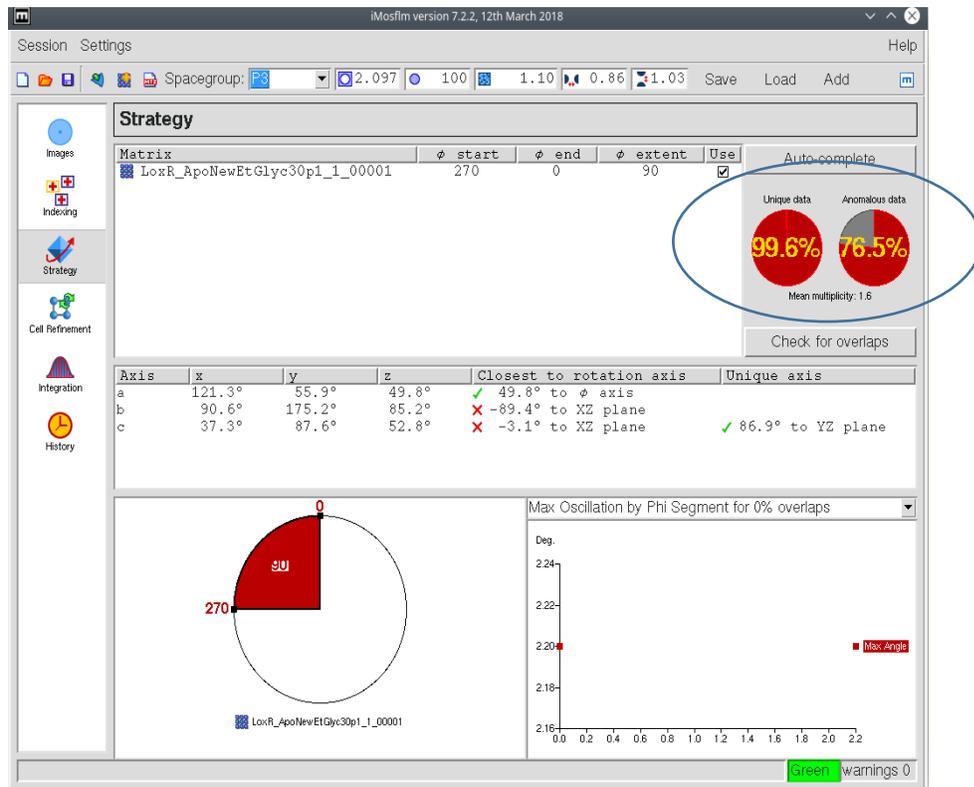
Conoscendo queste informazioni posso calcolare le migliori condizioni di raccolta (strategia della raccolta) al fine di avere un dataset finale di qualità adeguata.



Dimezzando il tempo di acquisizione ottengo comunque una completezza del 100%

# Strategia

Conoscendo l'orientazione del cristallo, posso calcolare come acquisire tutti i dati necessari utilizzando l'angolo totale più piccolo, pur avendo una completezza elevata (> 90%). **In tal modo riduco il tempo di esposizione totale e riduco il danno da radiazione.**



# Integrazione - 1

Una volta stabiliti gli indici di Miller, ovvero i piani di Bragg associati, di ogni singola intensità diffratta, posso **'estrarre' le intensità**, ovvero determinare per ogni diffrazione (spot) in ogni immagine acquisita, l'intensità relativa.

Una volta indicizzato il mio pattern, posso predire dove saranno posizionati le mie intensità e quindi creare una **'maschera'** intorno ad ogni spot da cui estrarre le intensità

The screenshot displays the XDS-INT software interface for data integration. The main window is titled "Integration" and shows the following components:

- Parameter Tables:**

Parameter	Value	Fix
Beam x	228.96	<input type="checkbox"/>
Beam y	215.51	<input type="checkbox"/>
Distance	429.70	<input type="checkbox"/>
Y-scale	1.0000	<input type="checkbox"/>
Tilt	-0.01	<input type="checkbox"/>
Twist	-0.04	<input type="checkbox"/>
Tangential offset	0.000	<input type="checkbox"/>
Radial offset	0.000	<input type="checkbox"/>
RMS residual	0.091	
RMS res. (central)	0.060	
RMS res. (weighted)	0.400	

Parameter	Value	Fix
$\phi$ (X)	-0.04	<input type="checkbox"/>
$\phi$ (Y)	0.03	<input type="checkbox"/>
$\phi$ (Z)	0.02	<input type="checkbox"/>
a	124.20	<input checked="" type="checkbox"/>
b	124.20	<input checked="" type="checkbox"/>
c	83.83	<input checked="" type="checkbox"/>
$\alpha$	90.00	<input checked="" type="checkbox"/>
$\beta$	90.00	<input checked="" type="checkbox"/>
$\gamma$	120.00	<input checked="" type="checkbox"/>
Mosaicity	0.695	<input type="checkbox"/>

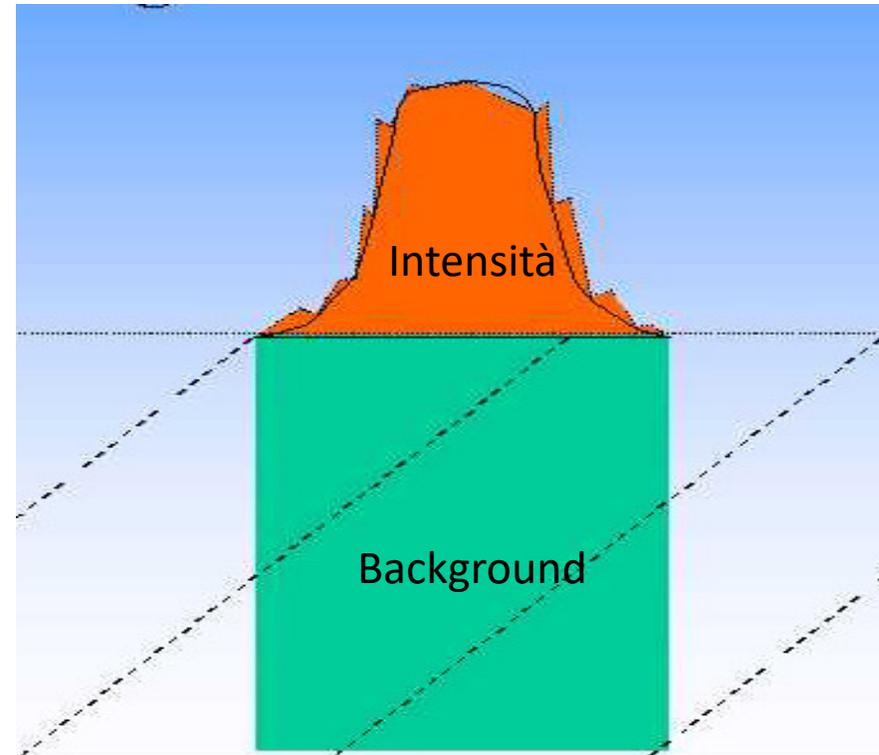
  

Parameter	Full	Partial
$\langle I/\sigma(I) \rangle$ (prf)	0.00	5.30
$\langle I/\sigma(I) \rangle$ (sum)	0.00	5.40
Reflections	0	2750
$\langle I/\sigma(I) \rangle$ HR (prf)	0.00	0.40
$\langle I/\sigma(I) \rangle$ HR (sum)	0.00	0.40
Reflections HR	0	497
Overloads	0	
Bad spots	2	
Spatial overlaps	5	
Lattice overlaps	0	
- Plots:** Three plots showing integration statistics vs. image number (0-20). The top plot shows Tilt (red squares) and Twist (yellow squares) in degrees. The middle plot shows Mosaicity (red circles) and  $\langle I/\sigma(I) \rangle$  (blue squares) in degrees. The bottom plot shows  $\langle I/\sigma(I) \rangle$  (prf) full (red squares) and part (yellow squares) in degrees.
- Image and Block Views:** The top right shows a single diffraction image (Image 12) with a central spot. The middle right shows a grid of images (Block 1) with a central spot. The bottom right shows a histogram of profile fits ( $I/\sigma(I)$ ) vs. Resolution (Å), with full reflections in red and partial reflections in yellow.
- Status Bar:** Shows "Refining image 24" and "Green warnings 0".

# Integrazione - 2

Per ogni singolo spot va determinata una intensità diffratta a cui è sottratto il background.

Inoltre sono stimati parametri come la mosaicità e la divergenza del fascio di raggi-X



# Postrefinement/Scalatura

- tutte le intensità integrate sono **normalizzate** (le singole immagini possono avere avuto esposizioni leggermente diverse) e i riflessi parziali sommati.
- Vengono **applicati fattori di correzione** (polarizzazione...)
- Vengono **calcolate le deviazioni standard** delle intensità e dei fattori di struttura
- **Le intensità legate da relazioni di simmetria devono essere uguali tra loro (*riflessi simmetrici*)**, quindi **queste intensità sono mediate tra di loro**.
- Sono prodotti degli **indicatori di qualità** per il dataset

Result

Summary data for Project: New Crystal: New Dataset: New

	Overall	InnerShell	OuterShell
Low resolution limit	83.83	83.83	2.38
High resolution limit	2.30	8.91	2.30
Rmerge (within I+/I-)	0.136	0.068	1.186
Rmerge (all I+ and I-)	0.139	0.071	1.221
Rmeas (within I+/I-)	0.144	0.072	1.257
Rmeas (all I+ & I-)	0.143	0.073	1.258
Rpim (within I+/I-)	0.047	0.023	0.414
Rpim (all I+ & I-)	0.034	0.018	0.297
Rmerge in top intensity bin	0.058	-	-
Total number of observations	601991	9954	57730
Total number unique	33476	645	3247
Mean(I)/sd(I)	12.6	29.3	2.6
Mn(I) half-set correlation CC(1/2)	0.999	1.000	0.763
Completeness	100.0	99.7	100.0
Multiplicity	18.0	15.4	17.8
Mean(Chi^2)	0.77	0.70	0.85
Anomalous completeness	100.0	98.9	100.0
Anomalous multiplicity	9.2	9.0	9.0
DelAnom correlation between half-sets	0.174	0.542	-0.010
Mid-Slope of Anom Normal Probability	0.836	-	-

# Indicatori di qualità

Alla fine del processo di 'riduzione dei dati' (Data reduction) sono prodotti degli indicatori di qualità:

$$R_{merge} = \frac{\sum_h \sum_{j=1}^N |I_{h,j} - \bar{I}_h|}{\sum_h \sum_{j=1}^N I_{h,j}}$$

La somma delle differenze tra le singole intensità (h k l), che devono essere uguali per simmetria, e il loro valor medio, diviso per la somma delle intensità.

**Più è piccolo e più la qualità è elevata**

$$\left\langle \frac{I}{\sigma(I)} \right\rangle = \frac{1}{N} \sum_h \frac{|I_h|}{\sigma(I_h)}$$

Indica il rapporto segnale/rumore medio.

Più è alto e meglio è ( $\sigma(I_h)$  è la deviazione standard di  $I_h$ )

**Ridondanza** (Redundancy):

Indica il numero di volte che una data intensità di indici (h, k, l) è stata acquisita in quanto tale, o come intensità diffratta da un piano di Bragg legato da relazione di simmetria.

**Completezza**

Per una **risoluzione** data, indica la percentuale di intensità effettivamente acquisite rispetto al numero massimo teorico (> 90%, più vicino è al 100% meglio è)

# R-merge

L'R-merge è un parametro che indica la qualità dell'insieme complessivo di dati acquisiti (data set)

La simmetria impone che determinate intensità di diffrazione  $I(hkl)$  siano uguali ad altre intensità  $I(h'k'l')$  con indici diversi (**riflessi equivalenti**) ma legati da relazioni ben definite. Inoltre in virtù dell'angolo totale di rotazione, una stessa intensità  $I(hkl)$  può essere stata acquisita più di una volta. Al termine del processo di *data reduction*, tutti queste intensità diffratte acquisite più volte, come tali o simmetrici, vengono confrontate e mediate tra loro, con un valore più accurato rispetto alle misure individuali.

L'R-merge misura la dispersione delle misure individuali rispetto al valor medio dell'intensità e deve essere numericamente piccolo (si esprime spesso in %)

$$R_{merge} = \frac{\sum_h \sum_{j=1}^N |I_{h,j} - \bar{I}_h|}{\sum_h \sum_{j=1}^N I_{h,j}}$$

Qualità dei dati	R-merge %
Eccellente	< 5
Buona	5-10
Accettabile	10-20

R-merge troppo elevato può indicare che la simmetria cristallina applicata è sbagliata

**E' forse il parametro più importante per valutare la qualità del data set**

# R-merge

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE  $\geq -3.0$  AS FUNCTION OF RESOLUTION

RESOLUTION LIMIT	NUMBER OF REFLECTIONS			COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR COMPARED expected	I/SIGMA	R-meas	CC(1/2)	Anomal Corr	SigAno	Nano	
	OBSERVED	UNIQUE	POSSIBLE										
6.83	23492	2440	2449	99.6%	3.4%	3.3%	23492	60.39	3.6%	99.9*	29*	1.097	1076
4.86	44542	4382	4382	100.0%	4.3%	4.1%	44542	47.29	4.5%	99.9*	17*	1.001	2044
3.97	57703	5635	5635	100.0%	4.4%	4.3%	57703	46.09	4.6%	99.9*	1	0.835	2668
3.45	69584	6675	6675	100.0%	6.7%	6.4%	69584	31.48	7.0%	99.9*	-2	0.826	3191
3.08	77663	7550	7550	100.0%	11.9%	11.8%	77663	18.08	12.5%	99.7*	0	0.812	3628
2.82	87215	8336	8336	100.0%	24.9%	26.3%	87215	9.05	26.2%	98.5*	-2	0.757	4022
2.61	93311	9094	9096	100.0%	47.3%	50.7%	93311	4.82	49.8%	94.4*	0	0.734	4398
2.44	102507	9770	9770	100.0%	78.5%	83.9%	102507	2.93	82.6%	85.8*	-1	0.727	4737
2.30	105417	10314	10331	99.8%	126.3%	139.3%	105391	1.73	133.0%	69.7*	0	0.689	4989
total	661434	64196	64224	100.0%	9.1%	9.3%	661408	17.55	9.6%	99.9*	1	0.787	30753

# $I/\sigma(I)$

$$\left\langle \frac{I}{\sigma(I)} \right\rangle = \frac{1}{N} \sum_h^N \frac{|I_h|}{\sigma(I_h)}$$

L'  $I/\sigma(I)$  medio si calcola mediando i rapporti del valor medio di ogni specifica  $I(hkl)$  diviso per la sua deviazione standard stimata

- E' una proprietà del data set
- **Più è elevata e più è elevata, in linea di principio, la qualità del data set**
- E' un indicatore del rapporto segnale/rumore

$$I/\sigma(I)$$

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE  $\geq -3.0$  AS FUNCTION OF RESOLUTION

RESOLUTION LIMIT	NUMBER OF REFLECTIONS			COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR expected	COMPARED	I/SIGMA	R-meas	CC(1/2)	Anomal Corr	SigAno	Nano
	OBSERVED	UNIQUE	POSSIBLE										
6.83	23492	2440	2449	99.6%	3.4%	3.3%	23492	60.39	3.6%	99.9*	29*	1.097	1076
4.86	44542	4382	4382	100.0%	4.3%	4.1%	44542	47.29	4.5%	99.9*	17*	1.001	2044
3.97	57703	5635	5635	100.0%	4.4%	4.3%	57703	46.09	4.6%	99.9*	1	0.835	2668
3.45	69584	6675	6675	100.0%	6.7%	6.4%	69584	31.48	7.0%	99.9*	-2	0.826	3191
3.08	77663	7550	7550	100.0%	11.9%	11.8%	77663	18.08	12.5%	99.7*	0	0.812	3628
2.82	87215	8336	8336	100.0%	24.9%	26.3%	87215	9.05	26.2%	98.5*	-2	0.757	4022
2.61	93311	9094	9096	100.0%	47.3%	50.7%	93311	4.82	49.8%	94.4*	0	0.734	4398
2.44	102507	9770	9770	100.0%	78.5%	83.9%	102507	2.93	82.6%	85.8*	-1	0.727	4737
2.30	105417	10314	10331	99.8%	126.3%	139.3%	105391	1.73	133.0%	69.7*	0	0.689	4989
total	661434	64196	64224	100.0%	9.1%	9.3%	661408	17.55	9.6%	99.9*	1	0.787	30753

# Ridondanza

La ridondanza è un numero che **indica quante volte, in media, una certa  $I(hkl)$  è stata misurata come tale o come equivalente per simmetria.**

Si calcola come rapporto tra il numero di intensità totali misurate e il numero di **intensità indipendenti**, ottenute al termine della media dei riflessi simmetrici e di quelli acquisiti più volte.

$$\frac{\text{Numero Intensità totali misurate}}{\text{Numero Intensità indipendenti}}$$

Per il teorema del limite centrale, **più è elevato e più il data set sarà accurato** (media di molti valori)

# Ridondanza

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE  $\geq -3.0$  AS FUNCTION OF RESOLUTION

RESOLUTION LIMIT	NUMBER OF REFLECTIONS			COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR expected	COMPARED	I/SIGMA	R-meas	CC(1/2)	Anomal Corr	SigAno	Nano
	OBSERVED	UNIQUE	POSSIBLE										
6.83	23492	2440	2449	99.6%	3.4%	3.3%	23492	60.39	3.6%	99.9*	29*	1.097	1076
4.86	44542	4382	4382	100.0%	4.3%	4.1%	44542	47.29	4.5%	99.9*	17*	1.001	2044
3.97	57703	5635	5635	100.0%	4.4%	4.3%	57703	46.09	4.6%	99.9*	1	0.835	2668
3.45	69584	6675	6675	100.0%	6.7%	6.4%	69584	31.48	7.0%	99.9*	-2	0.826	3191
3.08	77663	7550	7550	100.0%	11.9%	11.8%	77663	18.08	12.5%	99.7*	0	0.812	3628
2.82	87215	8336	8336	100.0%	24.9%	26.3%	87215	9.05	26.2%	98.5*	-2	0.757	4022
2.61	93311	9094	9096	100.0%	47.3%	50.7%	93311	4.82	49.8%	94.4*	0	0.734	4398
2.44	102507	9770	9770	100.0%	78.5%	83.9%	102507	2.93	82.6%	85.8*	-1	0.727	4737
2.30	105417	10314	10331	99.8%	126.3%	139.3%	105391	1.73	133.0%	69.7*	0	0.689	4989
total	661434	64196	64224	100.0%	9.1%	9.3%	661408	17.55	9.6%	99.9*	1	0.787	30753

# Completezza

Indica la **percentuale di Intensità indipendenti presenti nel data set, rispetto a quelli teoricamente possibili per quella data risoluzione.**

**Deve avvicinarsi quanto più al 100%, ma valori sopra il 90% sono buoni**

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

**Un data set 'incompleto' può produrre aberrazioni nella struttura finale**

# Completezza

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE  $\geq -3.0$  AS FUNCTION OF RESOLUTION

RESOLUTION LIMIT	NUMBER OF REFLECTIONS OBSERVED	UNIQUE	POSSIBLE	COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR expected	COMPARED	I/SIGMA	R-meas	CC(1/2)	Anomal Corr	SigAno	Nano
6.83	23492	2440	2449	99.6%	3.4%	3.3%	23492	60.39	3.6%	99.9*	29*	1.097	1076
4.86	44542	4382	4382	100.0%	4.3%	4.1%	44542	47.29	4.5%	99.9*	17*	1.001	2044
3.97	57703	5635	5635	100.0%	4.4%	4.3%	57703	46.09	4.6%	99.9*	1	0.835	2668
3.45	69584	6675	6675	100.0%	6.7%	6.4%	69584	31.48	7.0%	99.9*	-2	0.826	3191
3.08	77663	7550	7550	100.0%	11.9%	11.8%	77663	18.08	12.5%	99.7*	0	0.812	3628
2.82	87215	8336	8336	100.0%	24.9%	26.3%	87215	9.05	26.2%	98.5*	-2	0.757	4022
2.61	93311	9094	9096	100.0%	47.3%	50.7%	93311	4.82	49.8%	94.4*	0	0.734	4398
2.44	102507	9770	9770	100.0%	78.5%	83.9%	102507	2.93	82.6%	85.8*	-1	0.727	4737
2.30	105417	10314	10331	99.8%	126.3%	139.3%	105391	1.73	133.0%	69.7*	0	0.689	4989
total	661434	64196	64224	100.0%	9.1%	9.3%	661408	17.55	9.6%	99.9*	1	0.787	30753

# Determinazione della risoluzione massima - 1

La risoluzione massima della struttura non è qualcosa di univocamente determinato, in generale si possono seguire criteri più o meno restrittivi.

In generale **si divide la risoluzione in tanti intervalli di  $d_{hkl}$  decrescente** e si valuta quale è il  $d_{hkl}$  più piccolo a cui gli indicatori di qualità sono ancora accettabili.

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE >= -3.0 AS FUNCTION OF RESOLUTION													
RESOLUTION LIMIT	NUMBER OF REFLECTIONS		COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR COMPARED expected	I/SIGMA	R-meas	CC(1/2)	Anomal Corr	SigAno	Nano		
	OBSERVED	UNIQUE POSSIBLE											
6.83	23492	2440	2449	99.6%	3.4%	3.3%	23492	60.39	3.6%	99.9*	29*	1.097	1076
4.86	44542	4382	4382	100.0%	4.3%	4.1%	44542	47.29	4.5%	99.9*	17*	1.001	2044
3.97	57703	5635	5635	100.0%	4.4%	4.3%	57703	46.09	4.6%	99.9*	1	0.835	2668
3.45	69584	6675	6675	100.0%	6.7%	6.4%	69584	31.48	7.0%	99.9*	-2	0.826	3191
3.08	77663	7550	7550	100.0%	11.9%	11.8%	77663	18.08	12.5%	99.7*	0	0.812	3628
2.82	87215	8336	8336	100.0%	24.9%	26.3%	87215	9.05	26.2%	98.5*	-2	0.757	4022
2.61	93311	9094	9096	100.0%	47.3%	50.7%	93311	4.82	49.8%	94.4*	0	0.734	4398
2.44	102507	9770	9770	100.0%	78.5%	83.9%	102507	2.93	82.6%	85.8*	-1	0.727	4737
2.30	105417	10314	10331	99.8%	126.3%	139.3%	105391	1.73	133.0%	69.7*	0	0.689	4989
total	661434	64196	64224	100.0%	9.1%	9.3%	661408	17.55	9.6%	99.9*	1	0.787	30753

# Determinazione della risoluzione - 2

Anche in questo caso ci sono programmi che possono suggerire la risoluzione in modo automatico.

Un approccio *classico* e conservativo prevede per l'ultimo intervallo:

- R-merge < 100
- $\langle I/\sigma(I) \rangle \sim 2$
- Completezza > 90%

```
Estimates of resolution limits: overall
  from half-dataset correlation CC(1/2) > 0.30: limit = 2.30A == maximum resolution
  from Mn(I/sd) > 1.50: limit = 2.30A == maximum resolution
  from Mn(I/sd) > 2.00: limit = 2.30A == maximum resolution

Estimates of resolution limits in reciprocal lattice directions:
  Along h k plane
    from half-dataset correlation CC(1/2) > 0.30: limit = 2.30A == maximum resolution
    from Mn(I/sd) > 1.50: limit = 2.30A == maximum resolution
  Along l axis
    from half-dataset correlation CC(1/2) > 0.30: limit = 2.30A == maximum resolution
    from Mn(I/sd) > 1.50: limit = 2.38A

Anisotropic deltaB (i.e. range of principal components), A^2: 6.40

Average unit cell: 124.20 124.20 83.83 90.00 90.00 120.00
Space group: P 31 2 1
Average mosaicity: 0.67
```

# Scelta del Gruppo Spaziale - 1

La scelta della cella effettuato comparando gli indicatori di qualità, specialmente  $R_{\text{merge}}$  e  $\langle I/\sigma(I) \rangle$ , per le varie possibilità.

La scelta migliore sarà ovviamente quella che da indicatori di qualità migliori.

La scelta del reticolo di Bravais e del gruppo spaziale viene effettuato osservando se esistono delle classi di intensità diffratte con valori di  $I/\sigma(I) \sim 0$  in modo sistematico, in virtù di

- Cella non primitiva (Facce centrate, Corpo centrato...)
- Esistenza di elementi di simmetria con traslazione (come  $2_1$ ,  $3_1$ ,  $3_2$ ...)

Queste classi di intensità diffratte sono **le estinzioni sistematiche**

# Determinazione del Gruppo Spaziale

Esistono programmi automatici che fanno questo tipo di analisi in modo sofisticato.

Alcune ambiguità non possono però essere risolte a questo livello.

**Solo il gruppo spaziale corretto permette di ottenere la struttura cristallografica corretta**  
[in questo caso è P 32 2 1]

```
Run of POINTLESS on 17/11/2019 at 12:23:25

Result
Best Solution:    space group P 31 2 1

Reindex operator: [h,k,l]
Laue group probability: 0.998
Systematic absence probability: 0.972
Total probability: 0.970
Space group confidence: 0.956
Laue group confidence 0.998

WARNING: You will have to resolve the enantiomorphic ambiguity later

Unit cell: 124.20 124.20 83.83 90.00 90.00 120.00
83.83 to 2.76 - Resolution range used for Laue group search
83.83 to 2.30 - Resolution range in file, used for systematic absence

Number of batches in file: 720

The data do not appear to be twinned, from the L-test
```

# Software per il post-refinement

Anche il post-refinement e l'assegnazione del gruppo spaziale sono eseguiti in modo automatico da software specializzati.

L'elevato numero di diffrazioni da valutare (decine di migliaia) richiede l'utilizzo di metodi automatici

- **XDS/SCALA**
- **HKL3000**
- **POINTLESS/AIMLESS (CCP4)**
- **XPREP**

# Il file di dati

h k l       $I_{hkl}$        $\sigma(I_{hkl})$

LIST OF REFLECTIONS

0	0	3	2773.00	83.61	2773.00	83.61	2773.00	83.61
			52.62	0.79	52.62	0.79	52.62	0.79
			1.00					
0	0	6	592.70	18.14	592.70	18.14	592.70	18.14
			24.33	0.37	24.33	0.37	24.33	0.37
			1.00					
0	0	9	822.10	20.64	822.10	20.64	822.10	20.64
			28.66	0.36	28.66	0.36	28.66	0.36
			1.00					

$F_{hkl}$        $\sigma(F_{hkl})$

In fine... Un file ascii

Col num	Sort order	Min	Max	Num Missing	% complete	Mean	Mean abs.	Resolution Low	Resolution High	Type	Column Label
1	ASC	0	48	0	100.00	23.1	23.1	45.30	2.20	H	H
2	NONE	0	28	0	100.00	8.2	8.2	45.30	2.20	H	K
3	NONE	-38	38	0	100.00	0.4	14.4	45.30	2.20	H	L
4	NONE	-15.9	17620.0	0	100.00	171.05	171.30	45.30	2.20	J	IMEAN
5	NONE	0.3	321.2	0	100.00	4.28	4.28	45.30	2.20	Q	SIGMEAN
6	NONE	-16.8	17620.0	27	99.93	171.09	171.44	45.30	2.20	K	I(+)
7	NONE	0.3	327.6	27	99.93	5.87	5.87	45.30	2.20	M	SIGI(+)
8	NONE	-16.9	17620.0	26	99.93	171.25	171.58	45.30	2.20	K	I(-)
9	NONE	0.3	321.2	26	99.93	5.83	5.83	45.30	2.20	M	SIGI(-)
10	NONE	0.5	132.6	0	100.00	8.99	8.99	45.30	2.20	F	FP
11	NONE	0.0	1.6	0	100.00	0.33	0.33	45.30	2.20	Q	SIGFP
12	NONE	0.5	132.6	27	99.93	8.98	8.98	45.30	2.20	G	F(+)
13	NONE	0.1	1.6	27	99.93	0.46	0.46	45.30	2.20	L	SIGF(+)
14	NONE	0.5	132.6	26	99.93	9.00	9.00	45.30	2.20	G	F(-)
15	NONE	0.1	1.6	26	99.93	0.45	0.45	45.30	2.20	L	SIGF(-)
16	NONE	0.0	1.0	0	100.00	0.95	0.95	45.30	2.20	I	FreeRflag

No. of reflections used in FILE STATISTICS 38097

LIST OF REFLECTIONS

0	0	3	2773.00	83.61	2773.00	83.61	2773.00	83.61
			52.62	0.79	52.62	0.79	52.62	0.79
			1.00					
0	0	6	592.70	18.14	592.70	18.14	592.70	18.14
			24.33	0.37	24.33	0.37	24.33	0.37
			1.00					
0	0	9	822.10	20.64	822.10	20.64	822.10	20.64
			28.66	0.36	28.66	0.36	28.66	0.36
			1.00					
0	0	12	13.48	1.68	13.48	1.68	13.48	1.68
			3.65	0.23	3.65	0.23	3.65	0.23
			1.00					
0	0	15	15.44	2.06	15.44	2.06	15.44	2.06
			3.90	0.27	3.90	0.27	3.90	0.27
			1.00					
0	0	21	1141.00	25.09	1141.00	25.09	1141.00	25.09
			33.76	0.37	33.76	0.37	33.76	0.37
			1.00					
0	0	24	210.30	6.25	210.30	6.25	210.30	6.25
			14.49	0.22	14.49	0.22	14.49	0.22
			1.00					
0	0	27	405.80	10.11	405.80	10.11	405.80	10.11
			20.13	0.25	20.13	0.25	20.13	0.25
			1.00					
0	0	30	343.60	8.90	343.60	8.90	343.60	8.90
			18.51	0.24	18.51	0.24	18.51	0.24
			1.00					
0	0	33	-0.23	3.19	-0.23	3.19	-0.23	3.19
			0.99	0.65	0.99	0.65	0.99	0.65
			1.00					