Lezione 4

Andrea Ilari

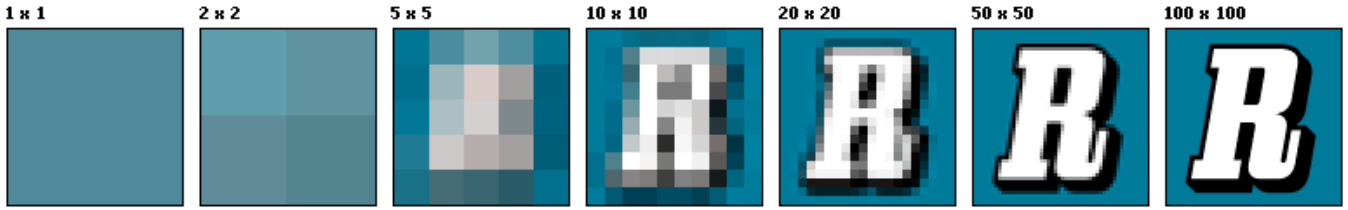
La CRISTALLOGRAFIA AI RAGGI X

I cristalli e la cristallizzazione

Cristallografia a raggi-X Parte 1



RISOLUZIONE



Il concetto di risoluzione è legato al campionamento dei dati

1-2 Å

10-20 Å

ALTA

Metodi indiretti

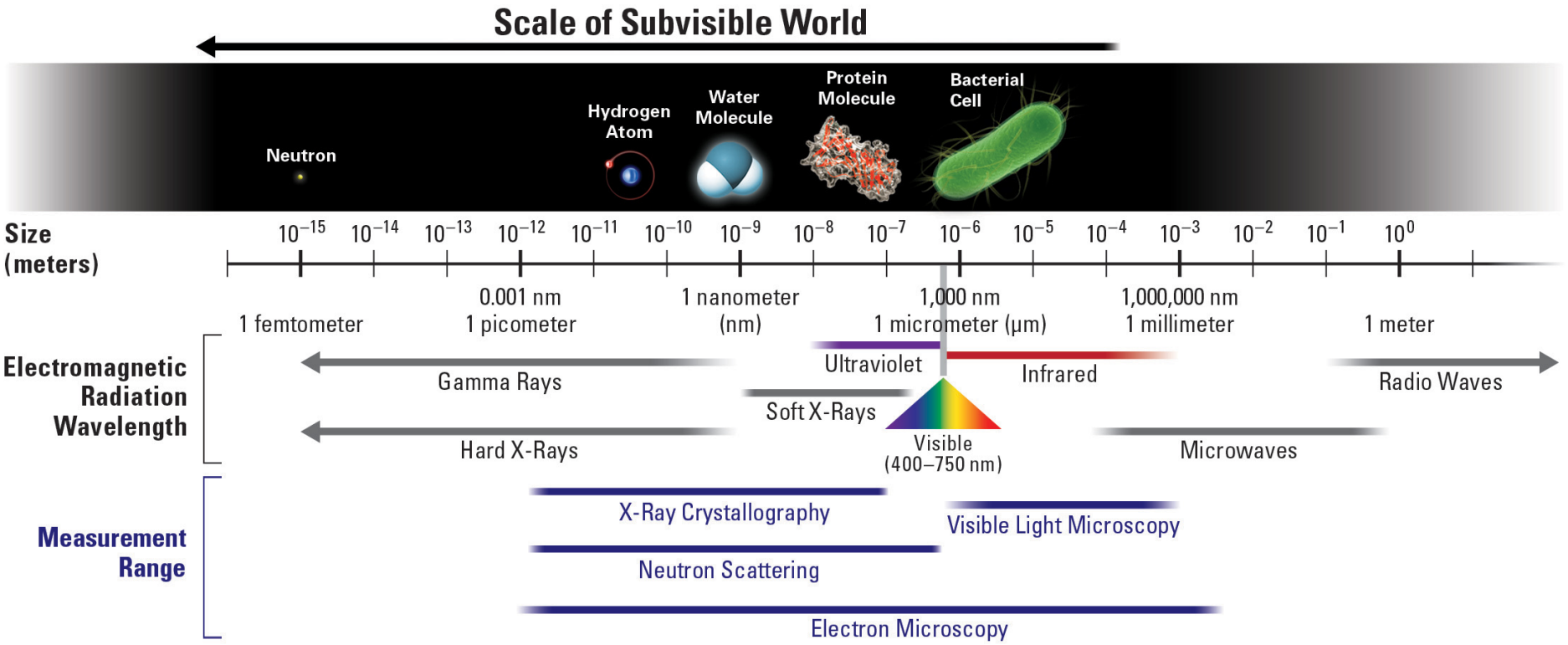
- **Cristallografia a raggi-X**
- Risonanza Magnetica Nucleare – NMR
- Cristallografia elettronica
- Cryo microscopia Elettronica - Cryo-EM

MEDIA

Metodi diretti

- Microscopia Elettronica - EM

Perché i Raggi-X ?



Per poter visualizzare due oggetti come entità separate, la lunghezza d'onda della radiazione elettromagnetica deve essere dello stesso ordine di grandezza della distanza tra gli oggetti:

distanza interatomica ≈1 Å λ raggi-X= 0.1÷2Å

⇒ con i raggiX "vedo" gli atomi

Perchè CRISTALLOGRAFIA e non MICROSCOPIA

NON ESISTONO LENTI PER FOCALIZZARE I RAGGI X

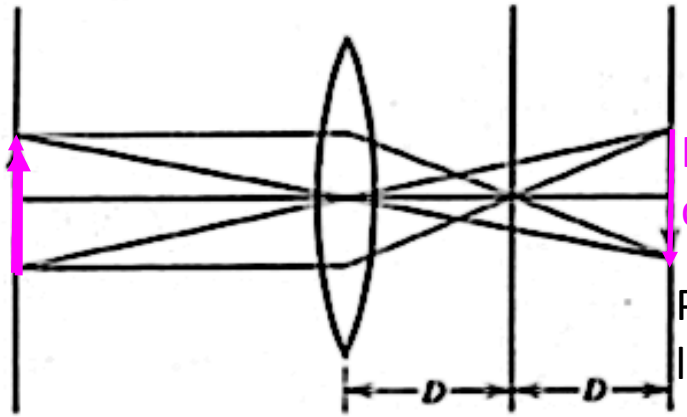


Immagine dell'oggetto

Piano del-l'immagine

L'immagine dell'oggetto e l'immagine di diffrazione sono in relazione tra di loro tramite una **trasformata di Fourier (TF)**

Ora, separiamo la lente in due metà:

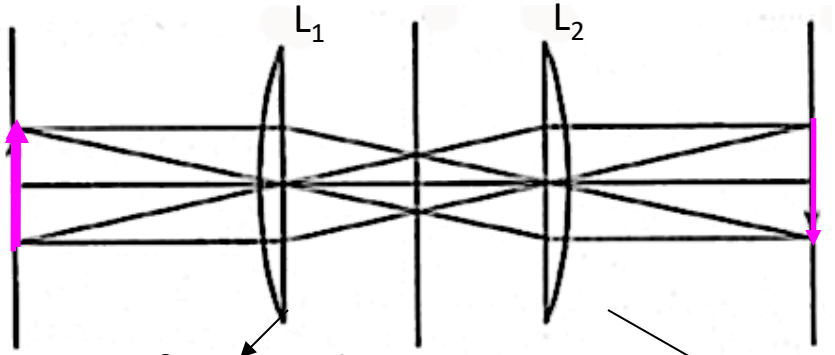


Immagine dell'oggetto (diffrazione di una diffrazione)

Trasformata di Fourier "fisica"

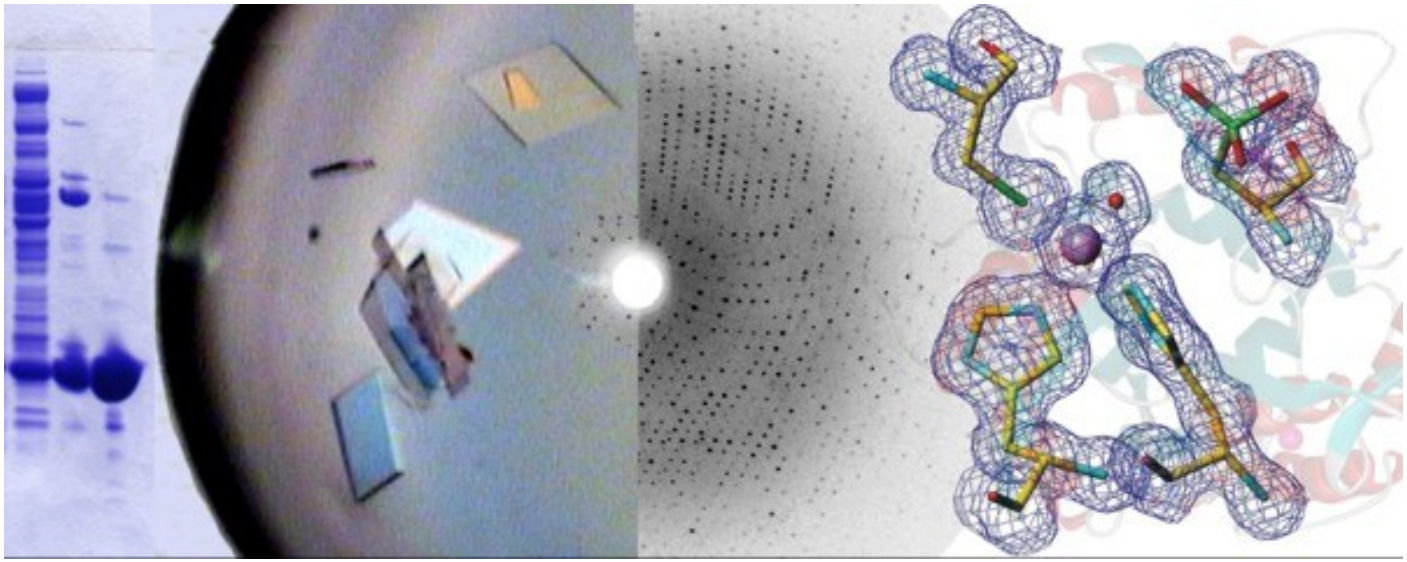
Trasformata di Fourier "matematica"

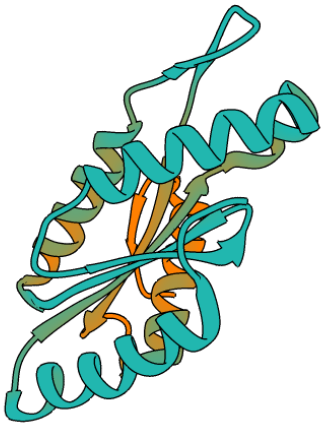
Piano focale, contenente l'immagine di diffrazione

L'immagine dell'oggetto è una TF della diffrazione
La diffrazione è una TF dell'oggetto

Opening the box

Non esistono lenti in grado di focalizzare i raggi-X ⇒ la matematica ci viene in aiuto





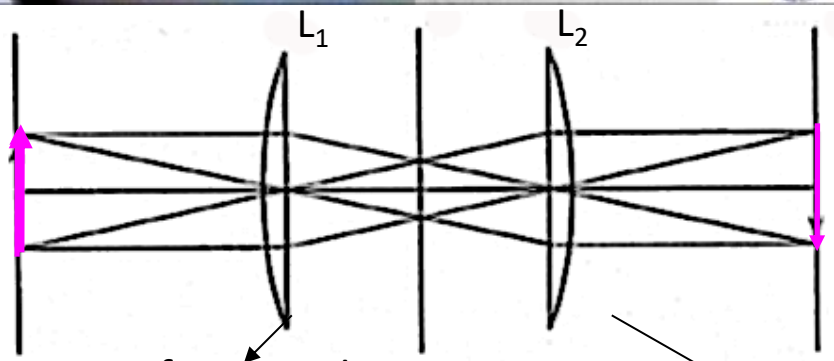


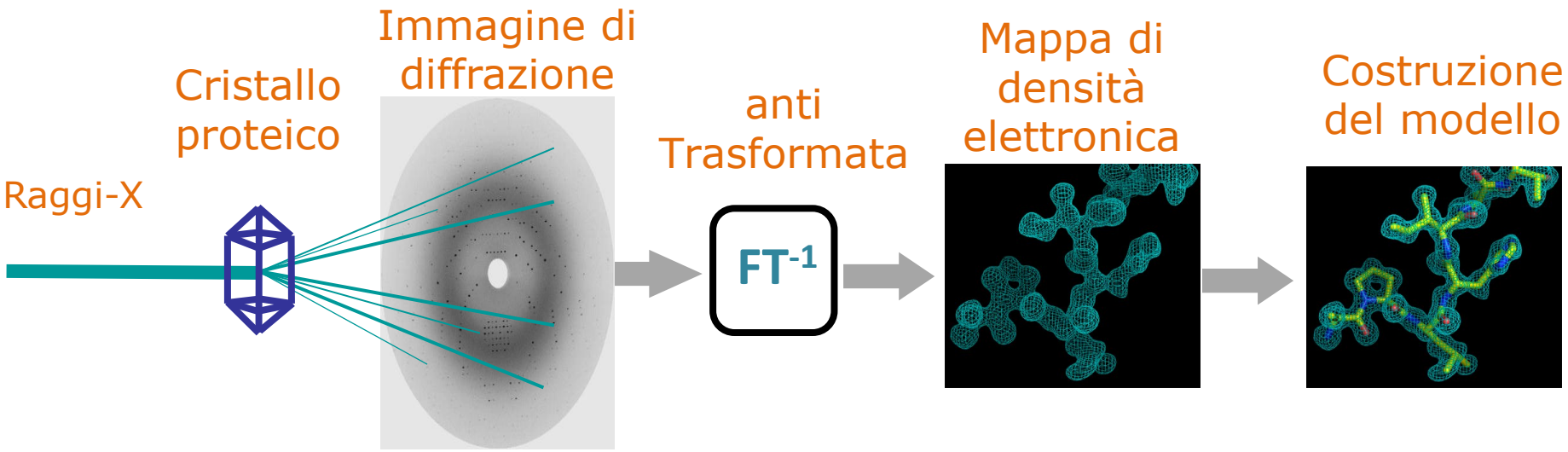
Immagine dell'oggetto (diffrazione di una diffrazione)

Trasformata di Fourier "fisica"

Piano focale, contenente l'immagine di diffrazione

Trasformata di Fourier "matematica"

L'esperimento di diffrazione



Expression
Purification
Crystallization



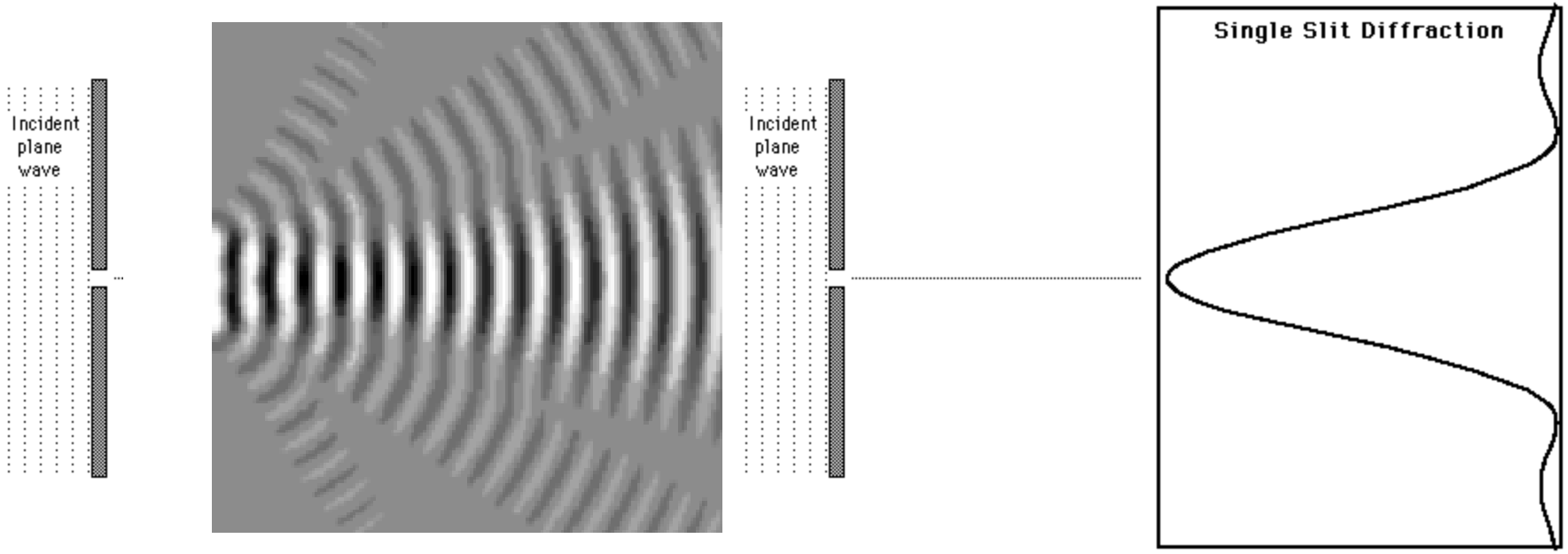
X-ray diffraction (data collection)

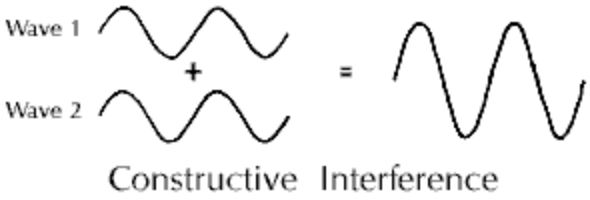


Phasing and model building

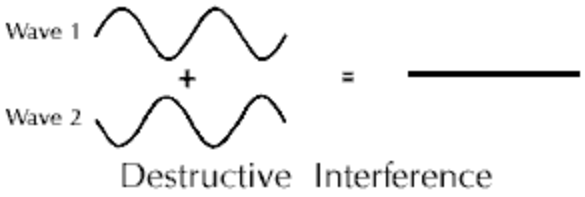


Diffrazione di un'onda elettromagnetica





Se le onde diffratte sono in fase avremo
Un'interferenza costruttiva (le ampiezze si sommano)



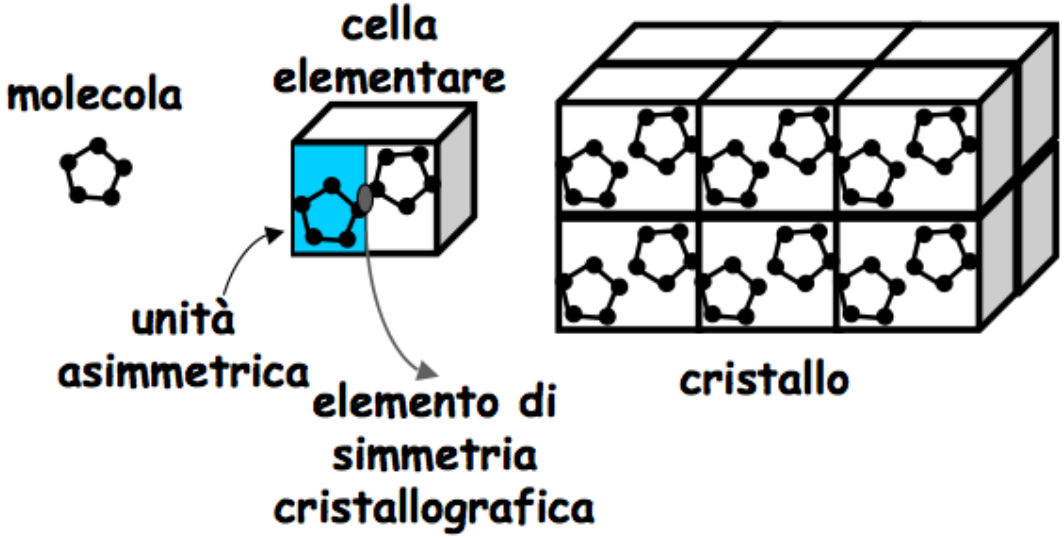
Se le onde diffratte non sono in fase avremo
un'interferenza distruttiva (le ampiezze si sottraggono)

Perchè un cristallo?

Un problema fondamentale è che l' intensità dell'onda diffusa da una singola molecola è troppo debole per dare informazioni utilizzabili,

⇒impossibile ricostruire la molecola

Utilizzando un cristallo il segnale viene amplificato.

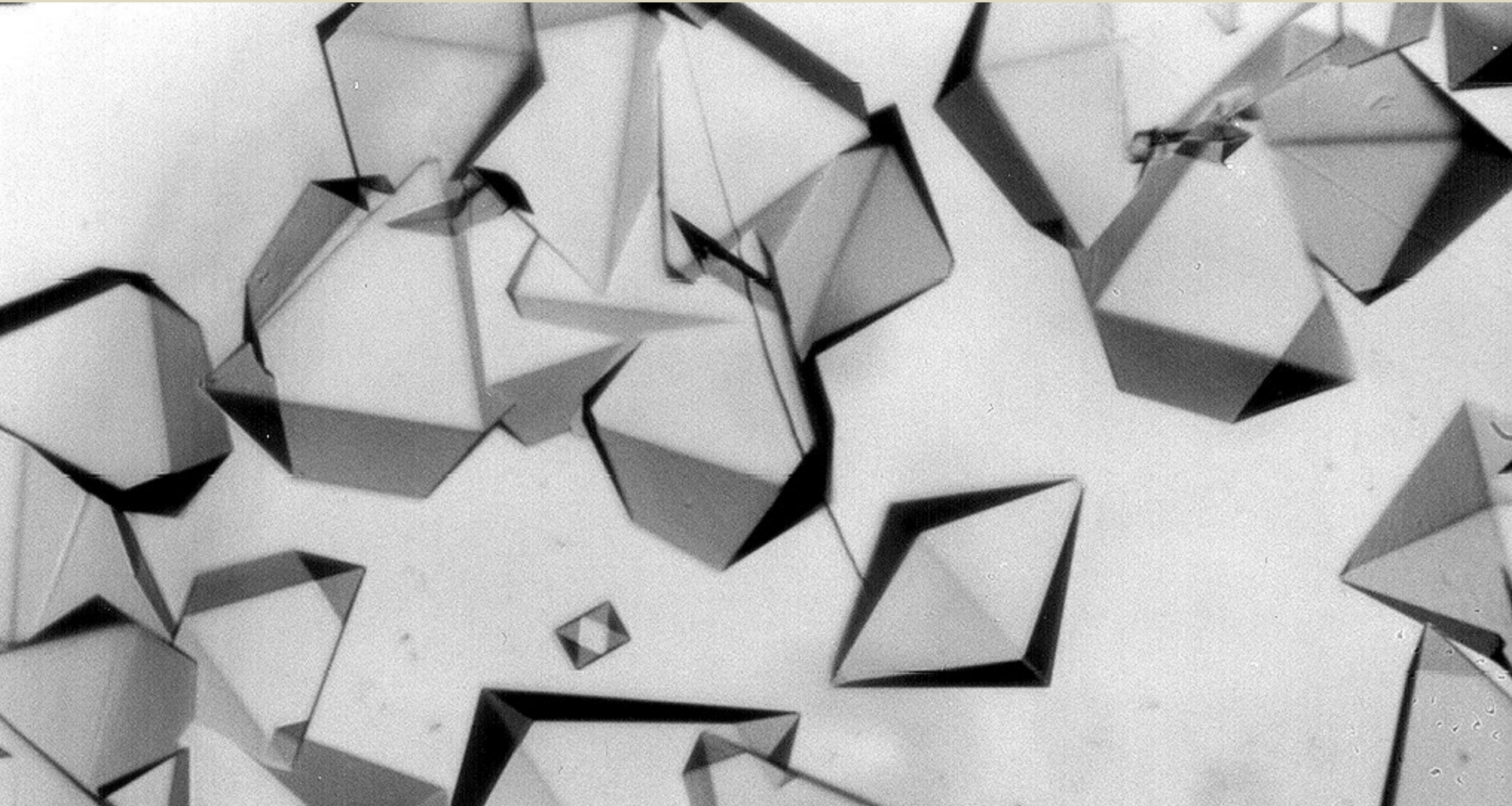


In un cristallo, miliardi di molecole sono disposte in modo ripetuto e ordinato nelle tre dimensioni dello spazio

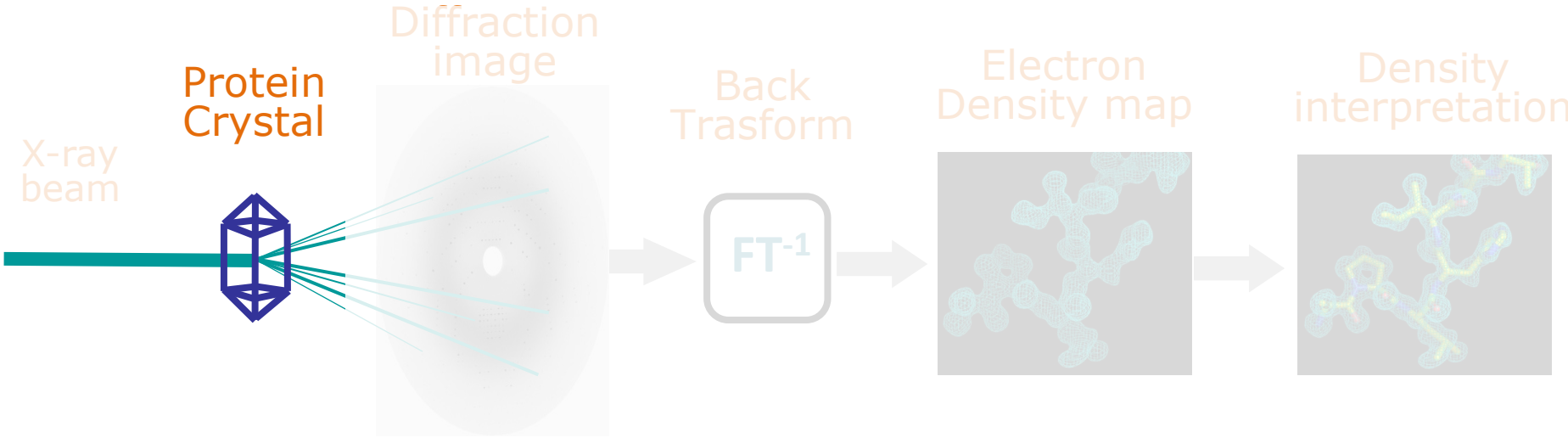
.

Per cui un cristallo funziona come un **amplificatore** del segnale

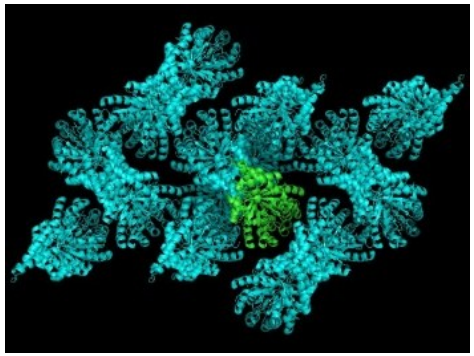
CRISTALLIZZAZIONE

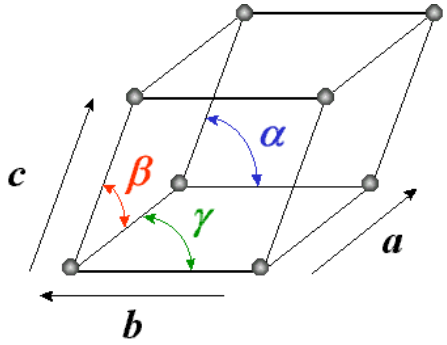


Perché servono CRISTALLI



CRYSTALS: a solid material whose constituents are arranged in a **highly ordered** structure, forming a crystal lattice that extends in all directions



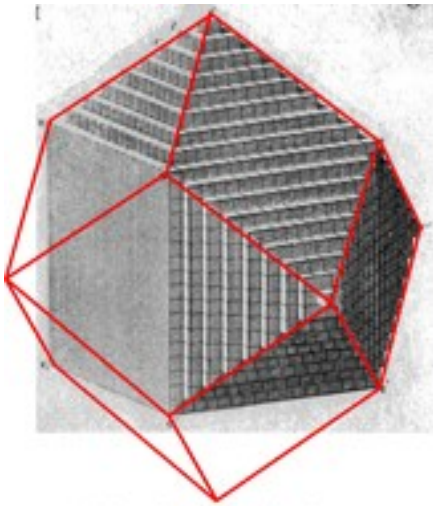


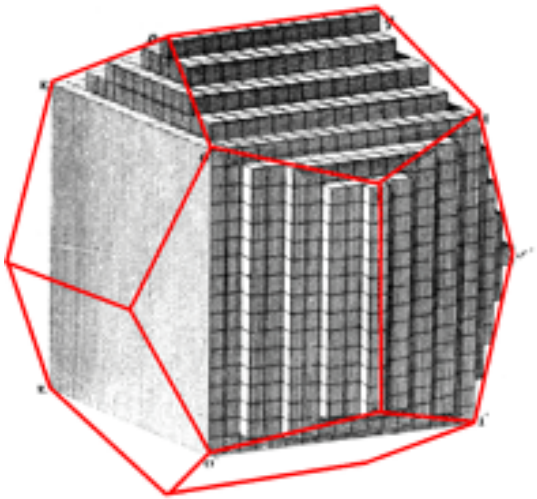
Unit Cell

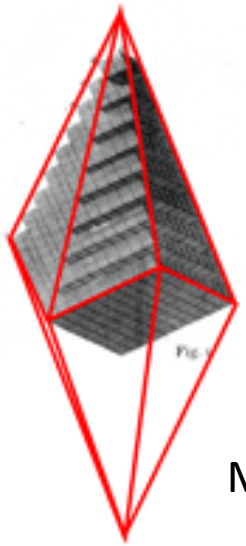
SIGNAL AMPLIFICATION

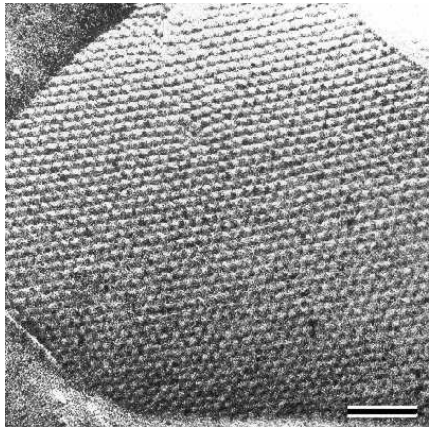
Cristallo ⇒ amplificatore del segnale
⇒ buone immagini di diffrazione (↑ segnale/rumore)

Modello di cristallo Haüy


Rhomb-dodecahedron


Pentagon-dodecahedron

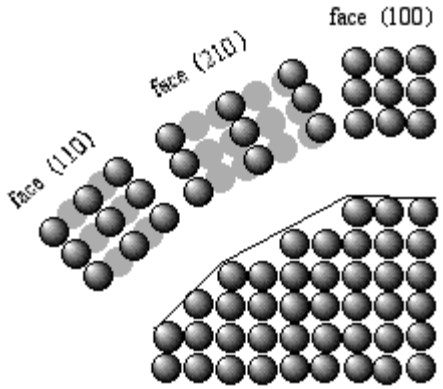

Scalenohedron


Microscopia elettronica

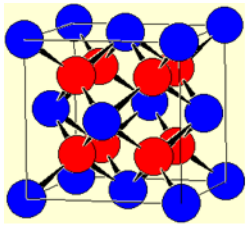
1801 *Traité de Minéralogie*

- La morfologia del cristallo deriva dalla ripetizione regolare di un parallelepipedo chiamato

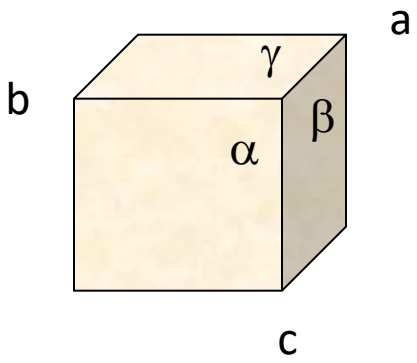
cella elementare

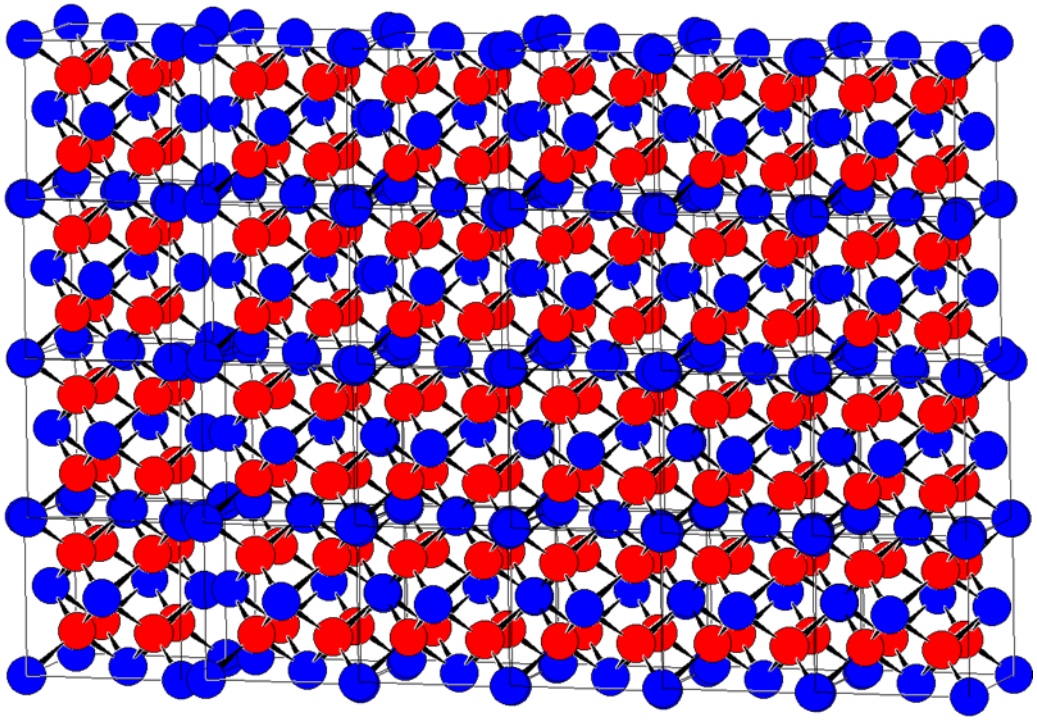


Cristalli e Simmetria



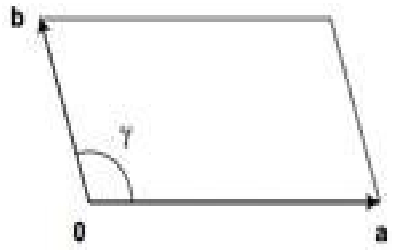
Cella elementare

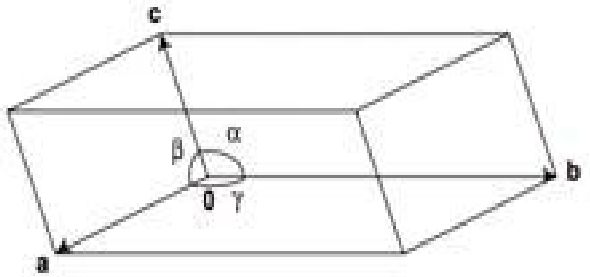




In un solido cristallino (cristallo), qualunque sia la sua natura: inorganico, organico o proteico, la materia deve soddisfare delle molto precise. La materia si dispone in modo omogeneo e periodico In un cristallo, gli atomi, ioni o molecole, chimicamente ben definiti, si dispongono secondo simmetrie ben precise

L’unità ripetitiva Cella elementare

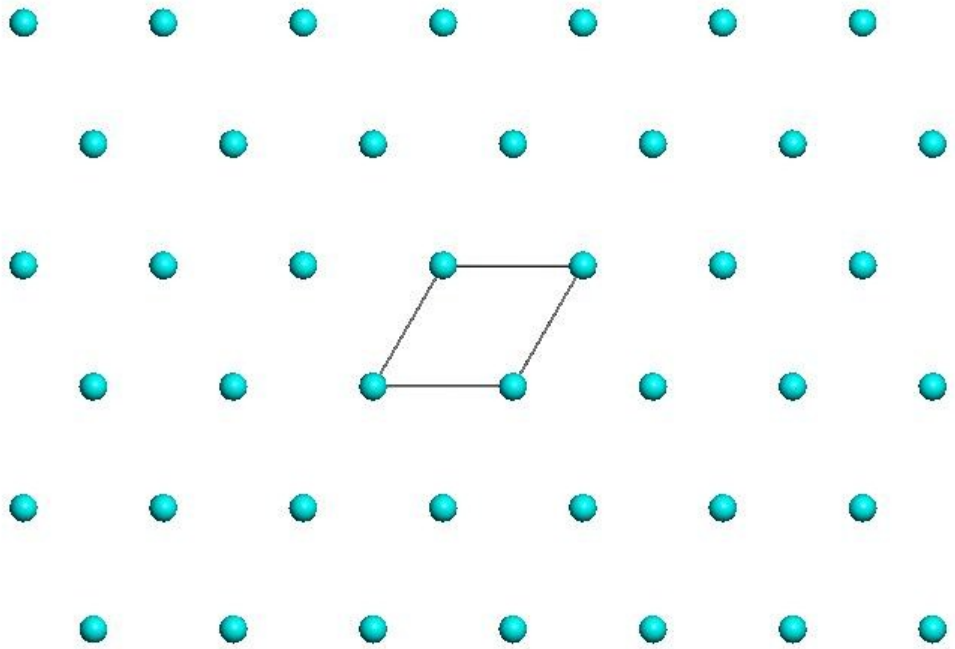




Parametri
di cella
a, b, c, α, β, γ

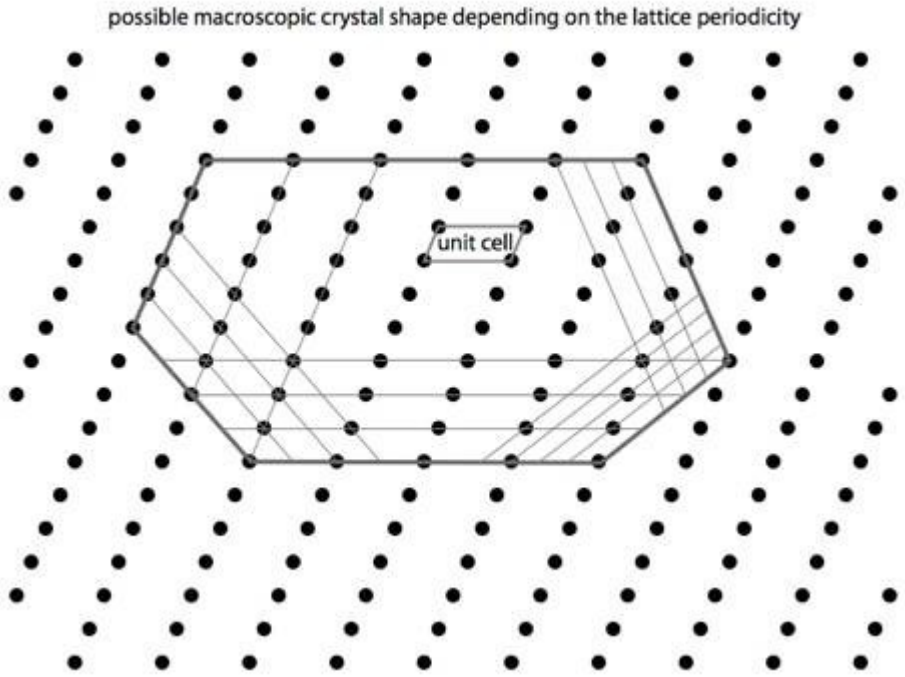
Reticolo cristallino

Nodi reticolari

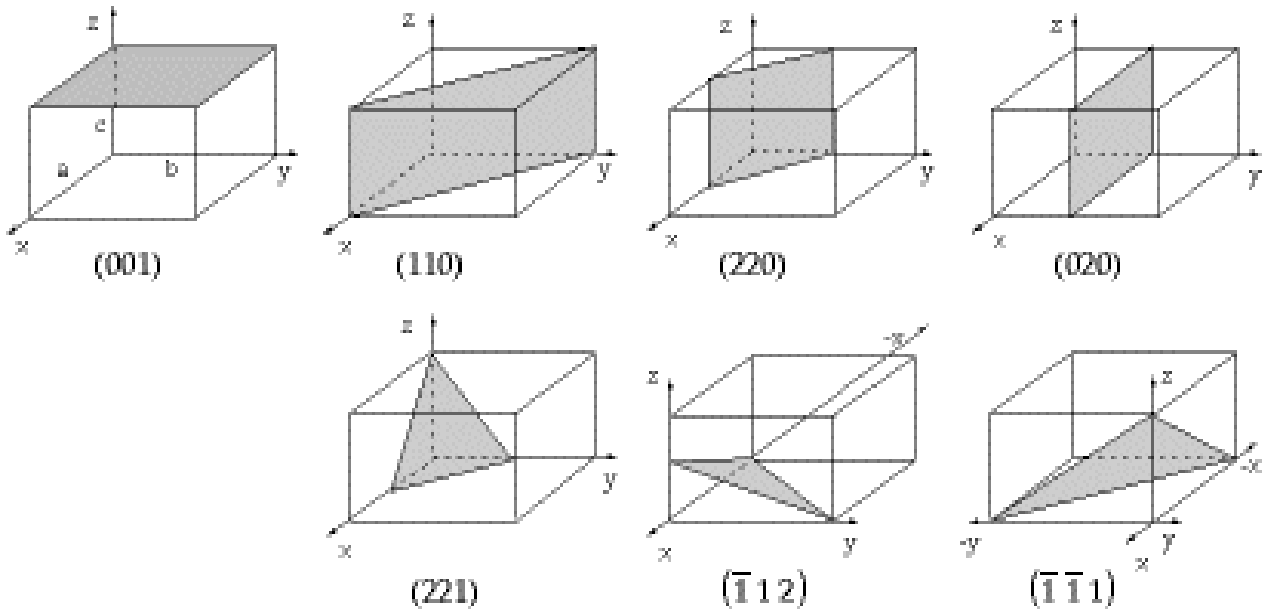


Legge di Bravais

Le facce dei cristalli sono parallele a piani reticolari con alta densità di nodi



I piani reticolari vengono individuati da una terna di numeri interi

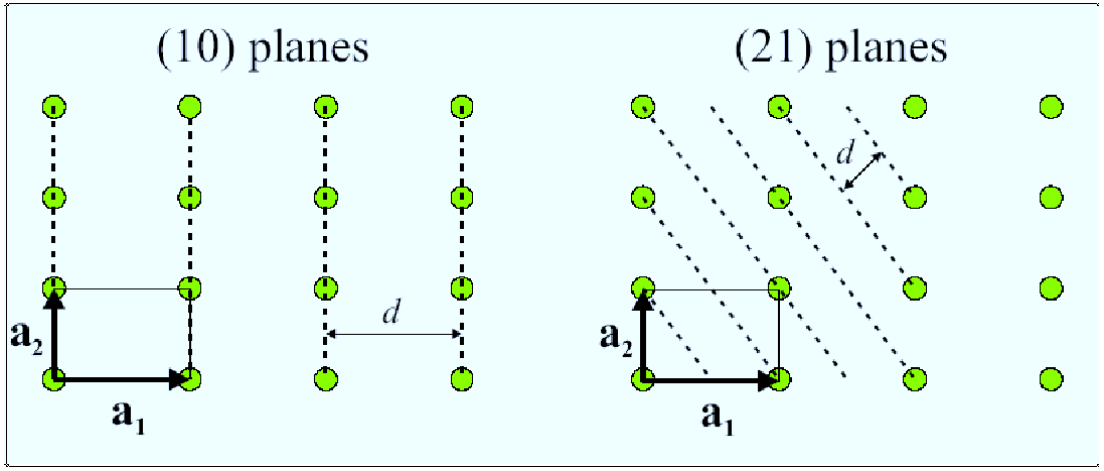


Indici di Miller

Il piano con indici di Miller *hkl* intercetta gli assi nei punti: ***a***/*h* ***b***/*k* ***c***/*l*

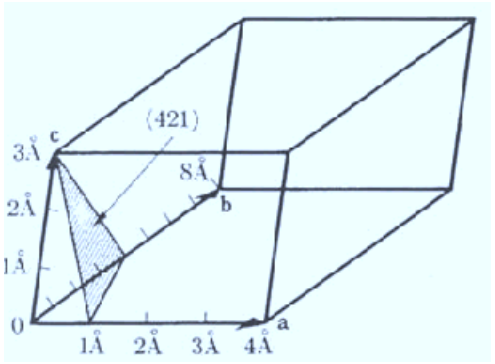
Piani reticolari

Un piano cristallografico passa almeno per tre nodi reticolari
Per ogni nodo passa un piano di una data famiglia di piani cristallografici



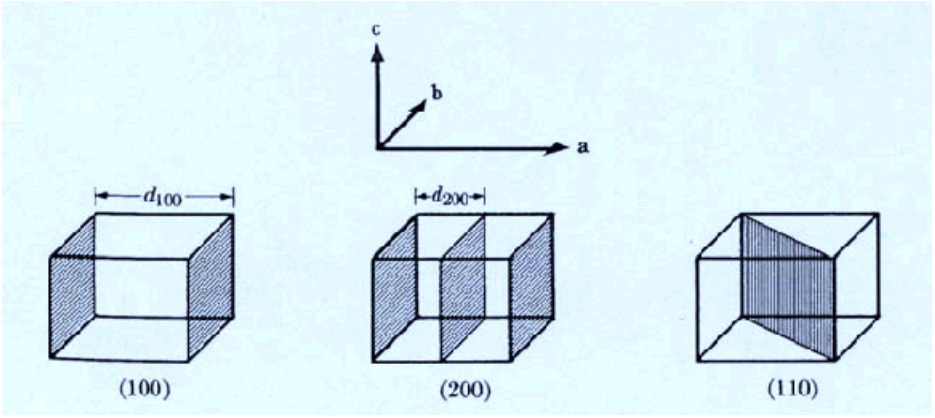
Le proprietà del reticolo implicano che in una famiglia di piani, i piani sono tra loro equispaziati da una grandezza costante indicata solitamente con la lettera *d*.

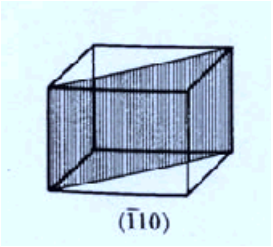
Fissata una certa origine esisterà un piano della famiglia di piani con indici di Miller (hkl) a una distanza d_{hkl} dall'origine che intersecherà gli assi della cella cristallografica nelle coordinate *a*/h, *b*/k e *c*/l.



Famiglie di piani

Nel caso particolare di una famiglia di piani parallela ad un asse, questo piano cristallografico avrà 0 come indice di Miller in quanto l'intercetta rispetto a questo asse è ∞.



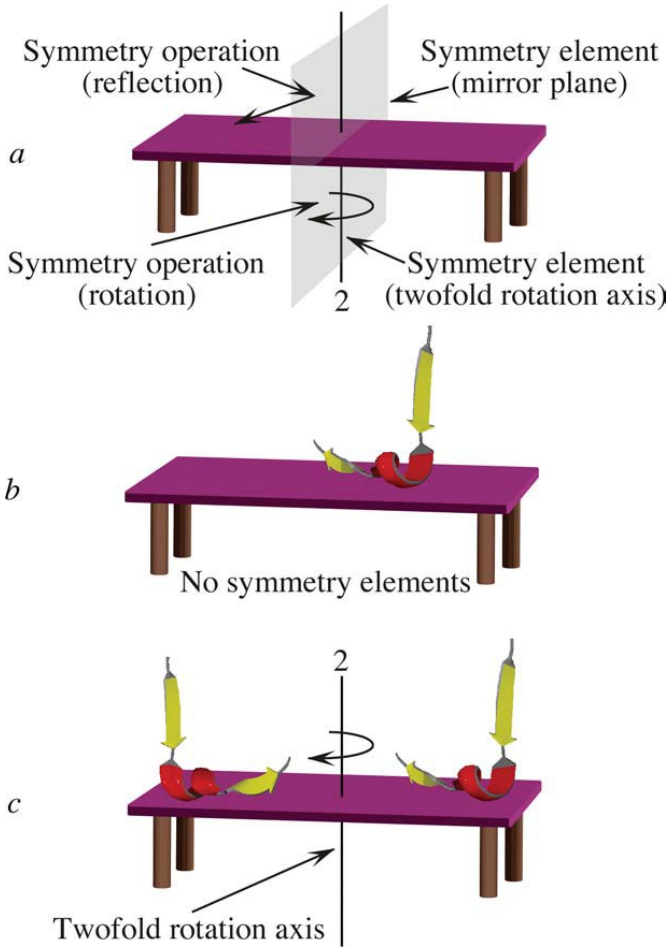


h=1; k=0; l=0
h=1 significa che i piani intercettano l'asse x nel punto a/1 e i suoi multipli
k=0 significa che i piani intercettano l'asse y nel punto b/0, quindi i piani sono paralleli all'asse y
c=0 significa che i piani intercettano l'asse y nel punto c/0, quindi i piani sono paralleli all'asse z

I piani relazionati alle facce di un cristallo sono piani reticolari in senso stretto (cioè i piani che passano per i nodi reticolari) ed hanno come indici di Miller dei numeri interi più piccoli possibili tra loro

Operazioni di simmetria

Una cella contiene più di una molecola proteica correlata con le altre molecole della cella
Da operazioni di simmetria. Nelle celle proteiche non ci sono «Mirror plane» perché le proteine
Sono oggetti chirali e asimmetrici



I principali elementi di simmetria
Sono:

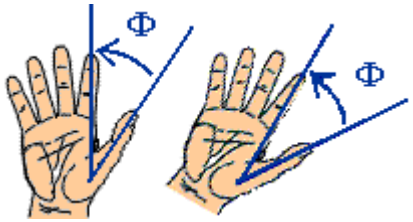
a) La traslazione: lo spostamento
Di a/n, b/n o c/n con n intero

b) La rotazione indicata con n con n intero
2 indica l'asse binario
3 indica l'asse ternario
4 indica l'asse quaternario
6 indica l'asse senario

c) Rototraslazione indicata con n/m che indica una rotazione ennesima seguita da una translazione pari a n/m

Congruenza

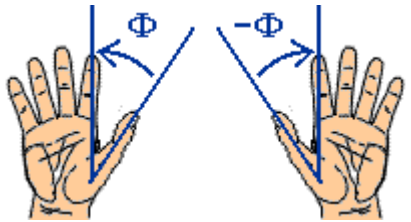
Diretta



- Traslazione
- Rotazione
- Roto-traslazione

compatibili
con tutte le molecole anche molecole chirali enantiopure

Inversa



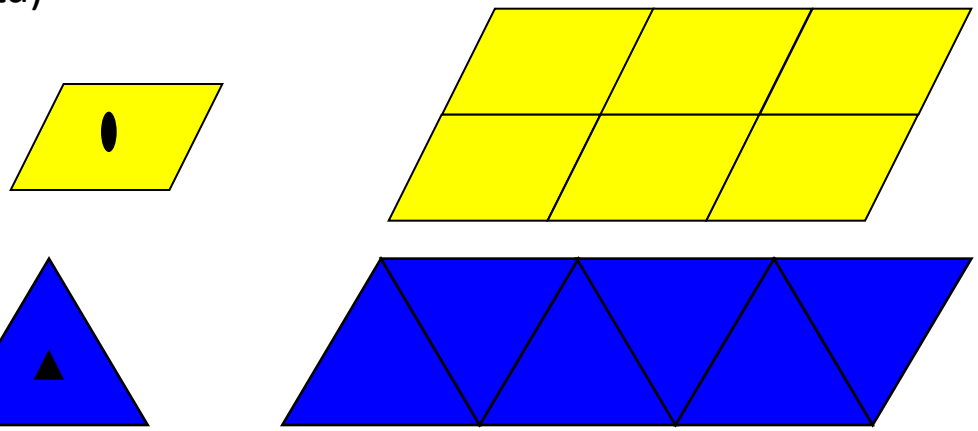
- Inversione
- Roto-inversione
- Roto-inversione-traslazione

incompatibili
con molecole chirali enantiopure
compatibili
con miscele racemiche e molecole achirali

Assi di rotazione compatibili con il reticolo cristallino

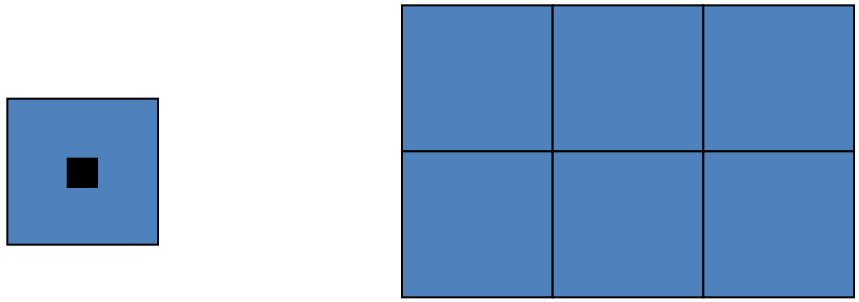
Asse di ordine 1 rotazione di 360° (identità)

Asse di ordine 2 (180°)

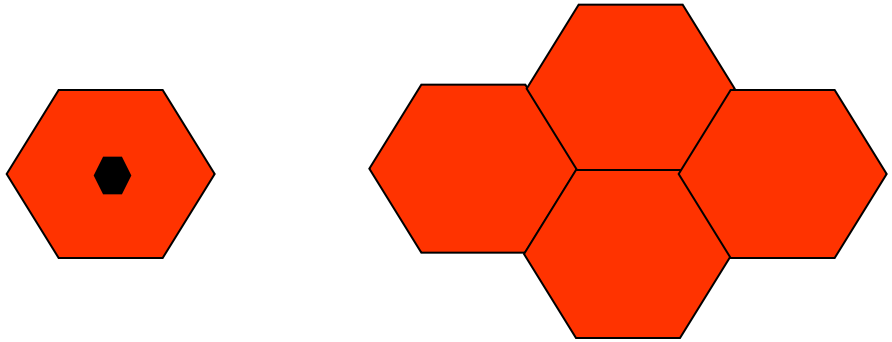


Asse di ordine 3 (120°)

Asse di ordine 4 (90°)

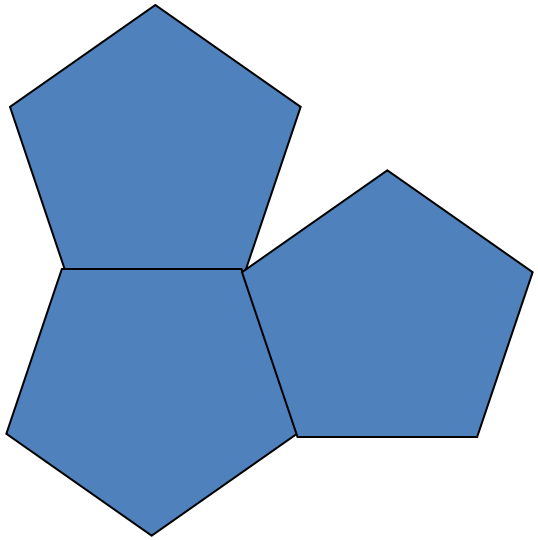


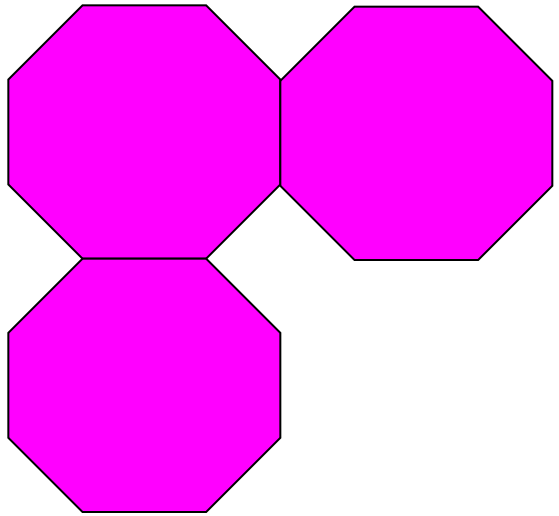
Asse di ordine 6 (60°)



Assi di rotazione non compatibili con il reticolo cristallino

Assi di ordine 5 di ordine 7 e superiori

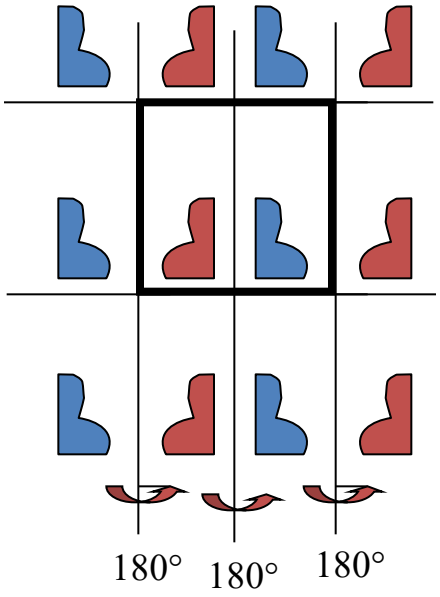




Assi di roto-traslazione compatibili con il reticolo cristallino

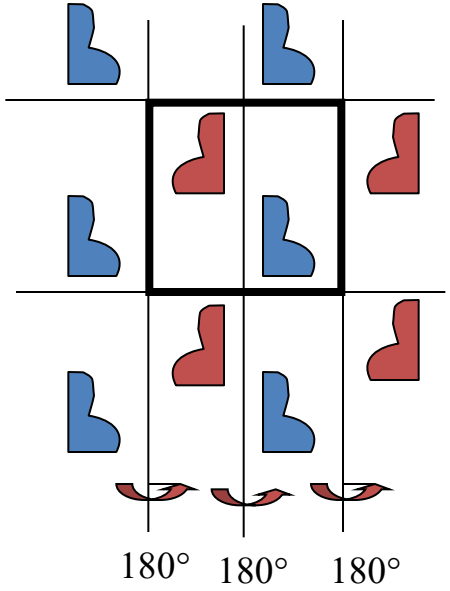
Asse 2





Asse screw 2_1





Asse 3



Assi screw 3_1 3_2



Asse 4



Assi screw 4_1 4_2 4_3



Asse 6



Assi screw 6_1 6_2 6_3 6_4 6_5

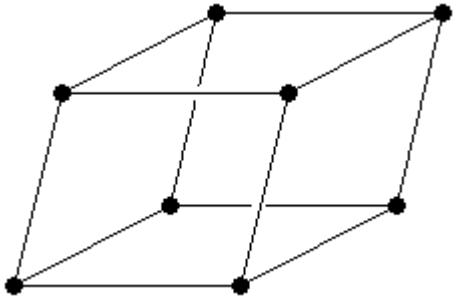


7 sistemi cristallini

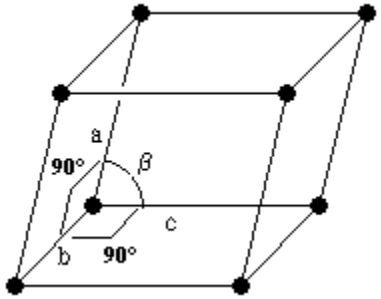
Simmetria minima	SISTEMA	vincoli sulla cella
nessuna	Triclino	nessuno
asse di ordine 2	Monoclino	$\alpha=\gamma=90°$
3 assi di ordine 2 perpendicolari	Ortorombico	$\alpha=\beta=\gamma=90°$
asse di ordine 4	Tetragonale	a=b $\alpha=\beta=\gamma=90°$
asse di ordine 3	Esagonale Trigonale	a=b $\gamma=120°$ $\alpha=\beta=90°$
asse di ordine 6	Esagonale	a=b $\gamma=120°$ $\alpha=\beta=90°$
4 assi di ordine 3 lungo le diagonali	Cubico	a=b=c $\alpha=\beta=\gamma=90°$

14 reticoli di Bravais

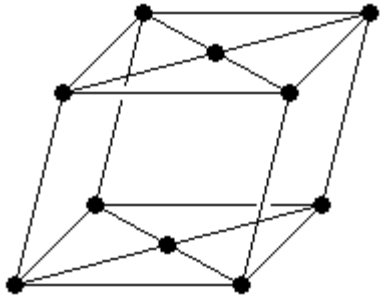
*a***P**



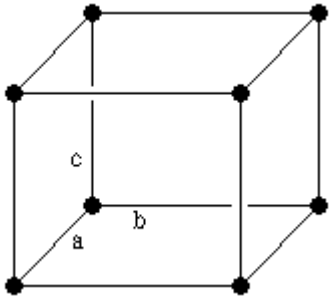
*m***P**



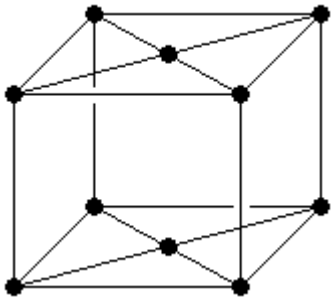
*m***C**



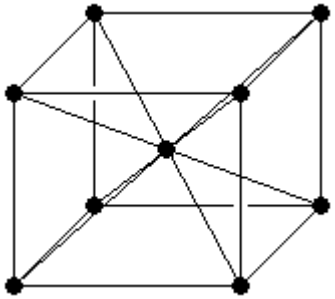
*o***P**



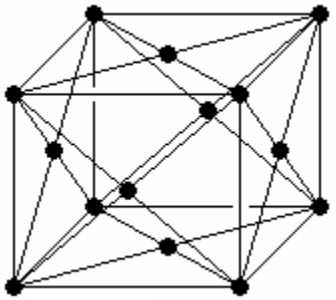
*o***C**



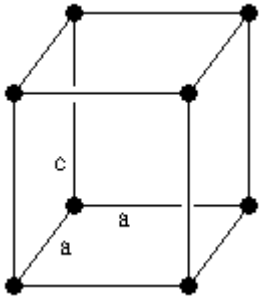
*o***I**



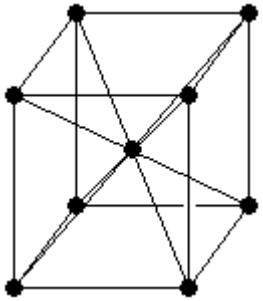
*o***F**



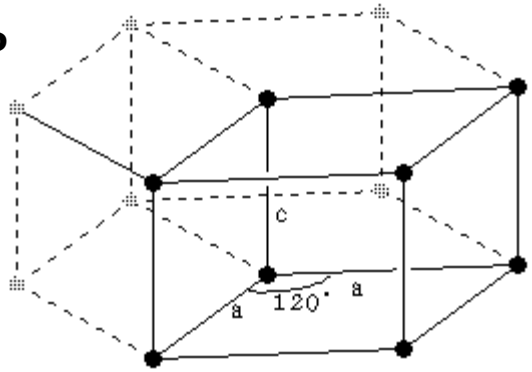
*t***P**



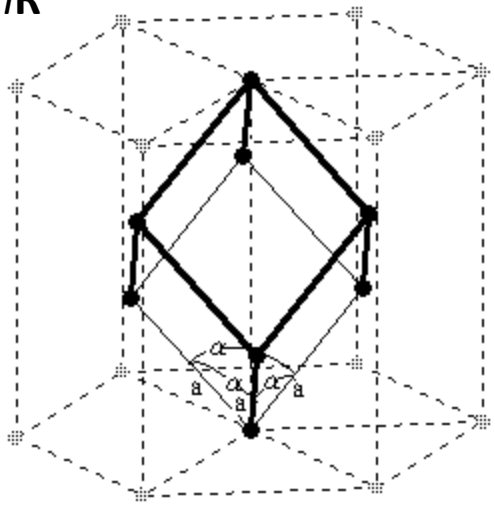
*t***I**



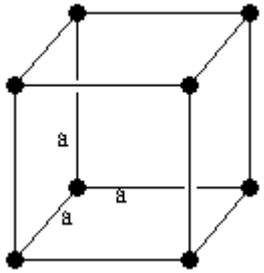
*h***P**



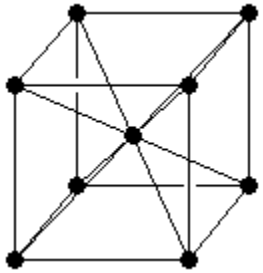
*h***R**



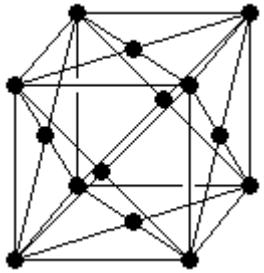
*c***P**



*c***I**



*c***F**



32 classi cristalline (gruppi puntuali cristallografici)

Vengono generate da tutte le possibili combinazione dei semplici elementi di simmetria puntuale (senza traslazione) compatibili con il reticolo

11 Classi di simmetria Laue (gruppi puntuali centrosimmetrici)

Triclino 1, **$\bar{1}$**

Monoclino 2, m, **2/m**

Ortorombico 222, mm2, **mmm**

Tetragonale 4, $\bar{4}$, **4/m**, 422, 4mm, $\bar{4}$2m, **4/mm$\bar{m}$**

Trigonale 3, **$\bar{3}$**, 32, $\bar{3}$m, **$\bar{3}$m**

Esagonale 6, $\bar{6}$, **6/m**, 622, 6mm, $\bar{6}$2m, **6/mm$\bar{m}$**

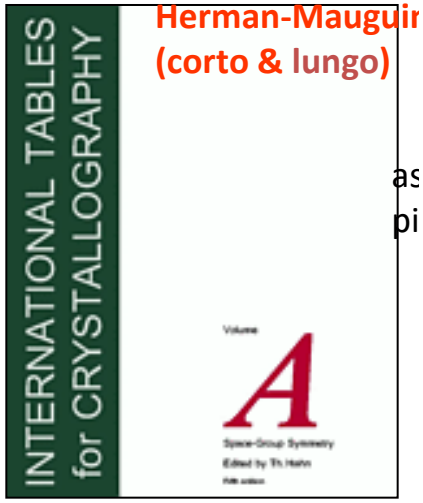
Cubico 23, **m$\bar{3}$**, 432, $\bar{4}$3m, **m$\bar{3}$m**

230 gruppi spaziali

Generati dalla combinazione di tutti gli elementi di simmetria compatibili con il reticolo, comprese le operazioni con traslazione (screw e glide) con i 14 reticoli di Bravais

65 gruppi spaziali compatibili con le molecole biologiche enantiomorfe in quanto non contengono elementi di rotoinversione

Tavole internazionali per la cristallografia



Simbolo gruppo s

Herman-Mauguin
(corto & lungo)

as

pi

CONTINUED No. 14 $P\,2_1/c$

Generators selected (1); $t(1,0,0)$; $t(0,1,0)$; $t(0,0,1)$; (2); (3)

Positions

Multiplicity, Wyckoff letter, Site symmetry			Coordinates				Reflection conditions
							General:
4	e	1	(1) x,y,z	(2) $\bar{x},y+\frac{1}{2},\bar{z}+\frac{1}{2}$	(3) $\bar{x},\bar{y},\bar{z}$	(4) $x,\bar{y}+\frac{1}{2},z+\frac{1}{2}$	$h0l$: $l=2n$ $0k0$: $k=2n$ $00l$: $l=2n$
							Special: as above, plus
2	d	$\bar{1}$	$\frac{1}{2},0,\frac{1}{2}$	$\frac{1}{2},\frac{1}{2},0$			hkl : $k+l=2n$
2	c	$\bar{1}$	$0,0,\frac{1}{2}$	$0,\frac{1}{2},0$			hkl : $k+l=2n$
2	b	$\bar{1}$	$\frac{1}{2},0,0$	$\frac{1}{2},\frac{1}{2},\frac{1}{2}$			hkl : $k+l=2n$
2	a	$\bar{1}$	$0,0,0$	$0,\frac{1}{2},\frac{1}{2}$			hkl : $k+l=2n$

Coordinate frazionarie

Posizioni generali equivalenti

Siti Wyckoff

Symmetry of special projections

Along [001] $p\,2g\,m$
$\boldsymbol{a}'=\boldsymbol{a}_p$ $\boldsymbol{b}'=\boldsymbol{b}$
Origin at $0,0,z$

Along [100] $p\,2g\,g$
$\boldsymbol{a}'=\boldsymbol{b}$ $\boldsymbol{b}'=\boldsymbol{c}_p$
Origin at $x,0,0$

Along [010] $p\,2$
$\boldsymbol{a}'=\frac{1}{2}\boldsymbol{c}$ $\boldsymbol{b}'=\boldsymbol{a}$
Origin at $0,y,0$

Maximal non-isomorphic subgroups

I [2]$P\,1\,2_1\,1$ $(P\,2_1)$ 1; 2
[2]$P\,\bar{1}$ 1; 3
[2]$P\,1\,c\,1$ $(P\,c)$ 1; 4

Subgruppi e supergruppi

IIa none

IIb none

Maximal isomorphic subgroups of lowest index

IIc [3]$P\,1\,2_1/c\,1$ $(\boldsymbol{b}'=3\boldsymbol{b})$ $(P\,2_1/c)$; [2]$P\,1\,2_1/c\,1$ $(\boldsymbol{a}'=2\boldsymbol{a}$ or $\boldsymbol{a}'=2\boldsymbol{a},\boldsymbol{c}'=2\boldsymbol{a}+\boldsymbol{c})$ $(P\,2_1/c)$

Minimal non-isomorphic supergroups

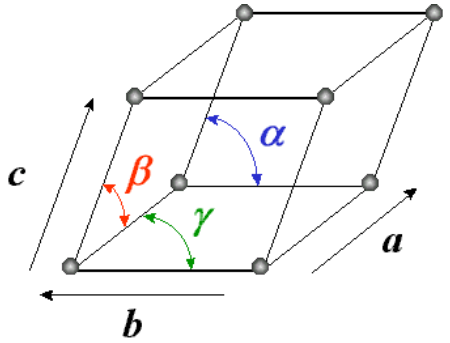
I [2]$P\,n\,n\,a$; [2]$P\,m\,n\,a$; [2]$P\,c\,c\,a$; [2]$P\,b\,a\,m$; [2]$P\,c\,c\,n$; [2]$P\,b\,c\,m$; [2]$P\,n\,n\,m$; [2]$P\,b\,c\,n$; [2]$P\,b\,c\,a$; [2]$P\,n\,m\,a$; [2]$C\,m\,c\,a$

II [2]$C\,1\,2/c\,1$ $(C\,2/c)$; [2]$A\,1\,2/m\,1$ $(C\,2/m)$; [2]$I\,1\,2/c\,1$ $(C\,2/c)$; [2]$P\,1\,2_1/m\,1$ $(2\boldsymbol{c}'=\boldsymbol{c})$ $(P\,2_1/m)$; [2]$P\,1\,2/c\,1$ $(2\boldsymbol{b}'=\boldsymbol{b})$ $(P\,2/c)$

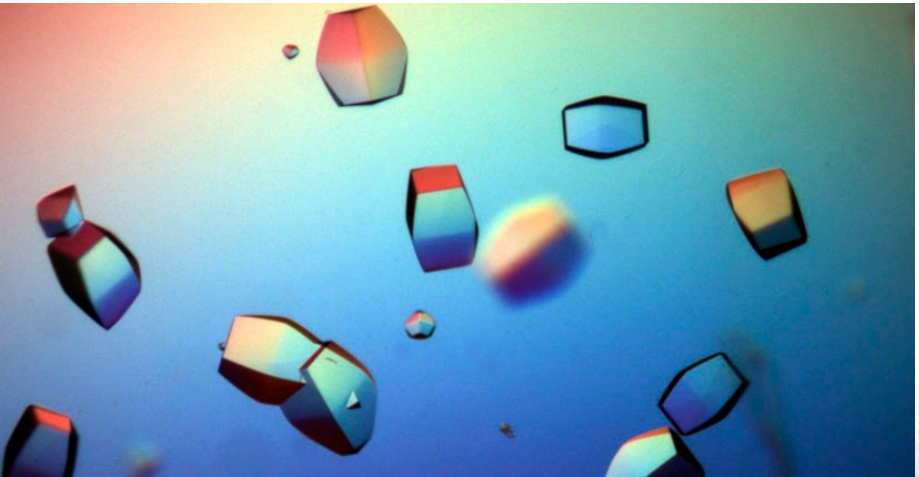
CARATTERISTICHE DEI CRISTALLI DI PROTEINE

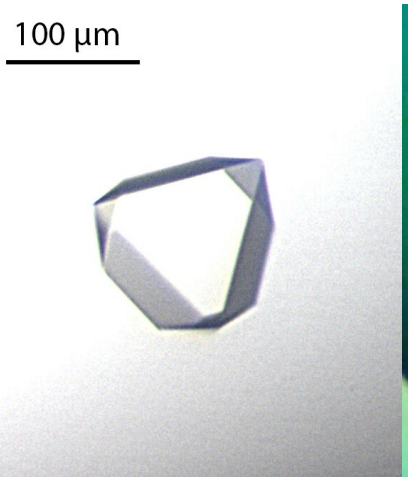


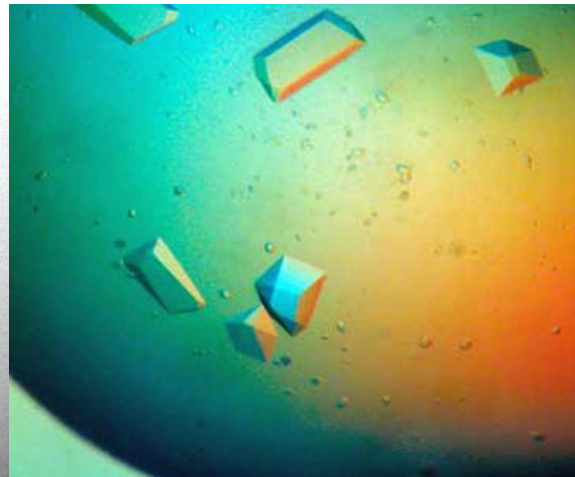
- **Dimensioni ridotte** (40µm-500µm)
- Meccanicamente **fragili**
- Sensibili alla temperatura
- Sensibili al **danno radiologico**
- **Alto contenuto di solvente** (40%-60%)
- Si fratturano se disidratati
- Deboli proprietà ottiche (basso potere diffrangente)



Unit Cell



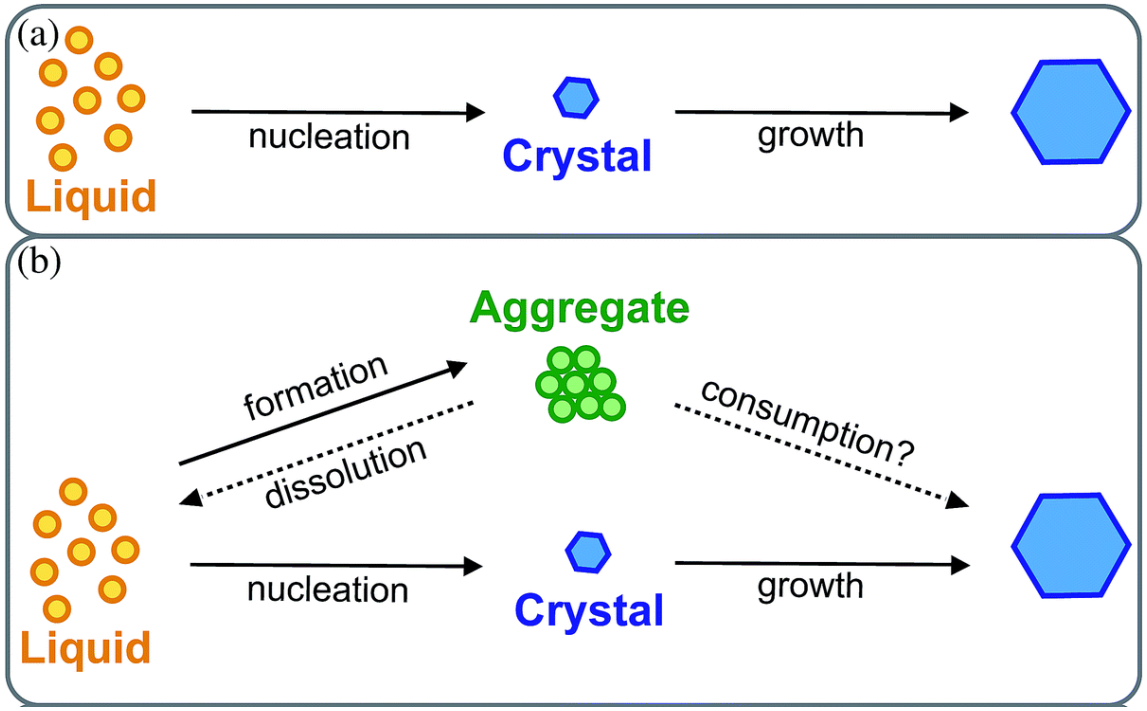




COME SI FORMANO I CRISTALLI



La cristallizzazione è un **passaggio di fase**



La nucleazione che prevede **un'aggregazione lenta e ORDINATA** delle macromolecole in un reticolo cristallino con contemporanea esclusione del solvente

$$\Delta G = \Delta H - T\Delta S$$

La cristallizzazione è **un processo spontaneo**, con un **fattore entropico sfavorevole** e **interazioni molecolari altamente favorite**

PARAMETRI CHE INFLUENZANO LA CRISTALLIZZAZIONE



- **TERMODINAMICA**
- Caratteristiche della macromolecola:
  - **Purezza** (>98%)
  - **Omogeneità** e monodispersione
  - **Folding (forma)**
  - **Distribuzione di carica**
- Parametri chimico-fisici
  - **Stabilità**
  - **pH**
- Parametri biochimici e biofisici
  - **Tag di espressione**
  - **Parti flessibili**
- Parametri biologici
  - **cofattori e/o ligandi**
  - **modificazioni post-traduzionali**

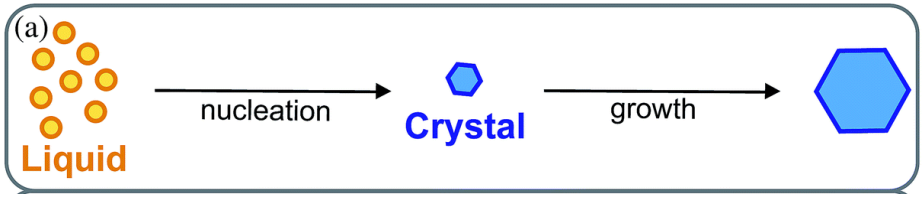
La nucleazione che prevede **un'aggregazione LENTA e ORDINATA** delle macromolecole in un reticolo cristallino con contemporanea esclusione del solvente

$$\Delta G = \Delta H - T\Delta S$$

La cristallizzazione è **un processo spontaneo**, con un **fattore entropico sfavorevole** e **interazioni molecolari altamente favorite**

TERMODINAMICA

$\Delta G = \Delta H - T\Delta S$



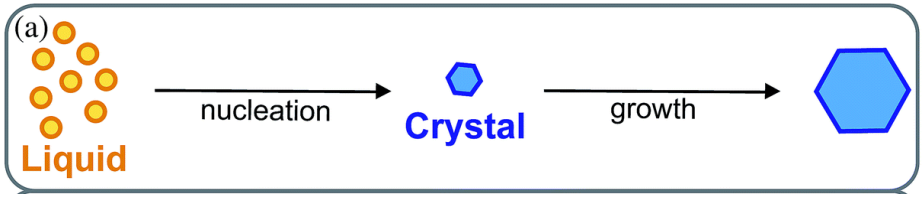


PER INDURRE LA TRANSIZIONE DI FASE
SI USANO **AGENTI PRECIPITANTI**

- **Sali**: disidratazione per competizione con H_2O
  - Es.: $(NH_4)_2SO_4$; $(Li)_2SO_4$; NaCl; etc..
- **Polimeri organici**: segregazione entropica (esclusione di volume)
  - Es.: Poli Etilen Glicol (PEG) con MW range 400 ÷ 20000
- **Solventi organici**: riduzione delle proprietà elettriche
  - Es.: poli-alcooli; Etilenglicole; etc..

TERMODINAMICA

$\Delta G = \Delta H - T\Delta S$



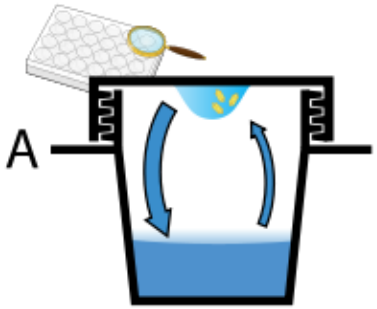


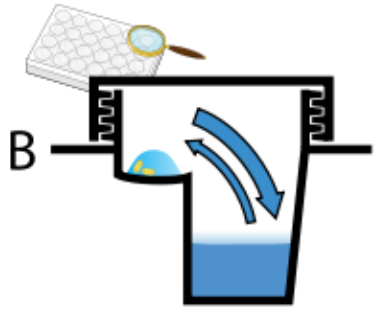
METODI DI CRISTALIZZAZIONE

Si miscelano la:

soluzione proteica + precipitante

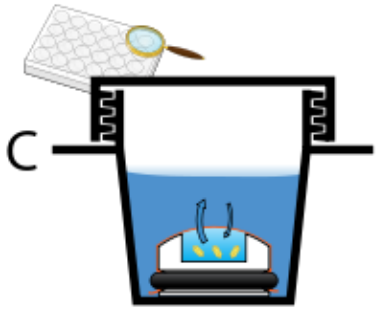
Goccie da 1-2µL circa





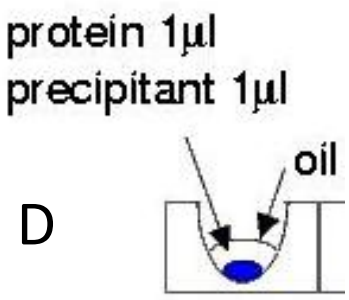
Diffusione in fase di vapore

- Goccia capovolta (hanging drop - A)
- Goccia seduta (sitting drop – B)



Altri metodi di scambio

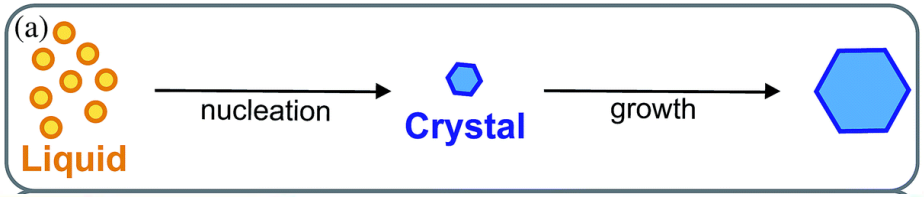
- Micro dialisi (C)
- Batch (D)



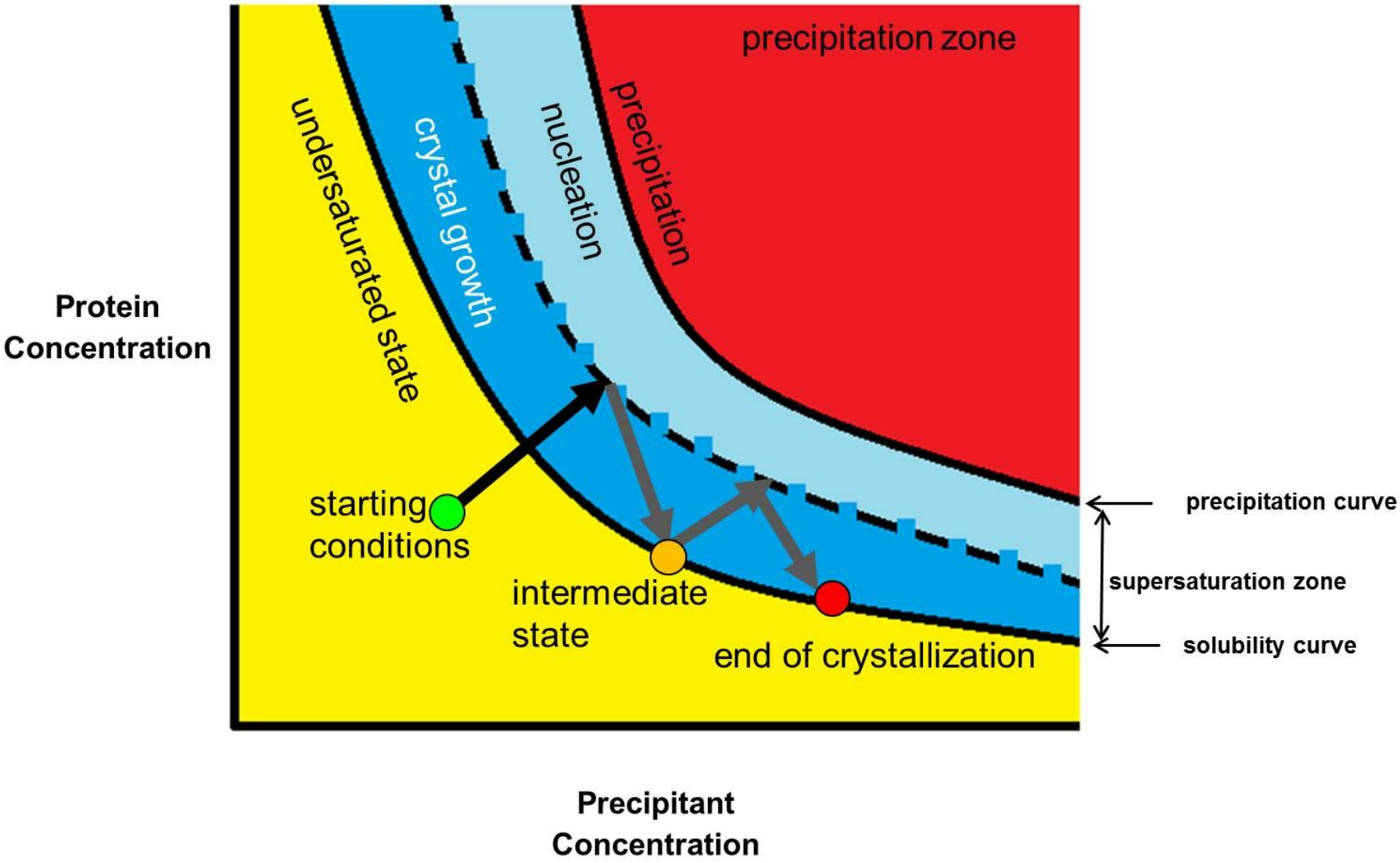
CRISTALLI DI PROTEINE

TERMODINAMICA

$\Delta G = \Delta H - T\Delta S$

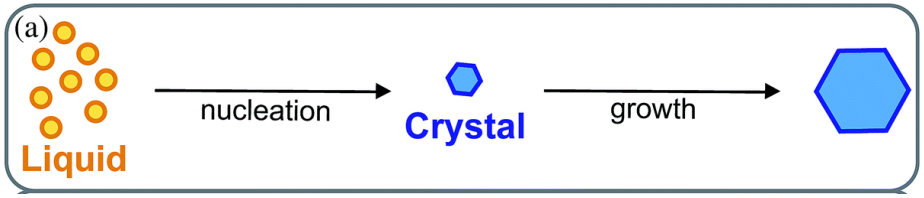






TERMODINAMICA

$\Delta G = \Delta H - T\Delta S$





LA CRISTALLIZZAZIONE È UN PROCESSO
EMPIRICO (TRIAL & ERROR)

Per testare **MOLTE CONDIZIONI DIVERSE**
utilizzo gli **SCREEN di CRISTALLIZZAZIONE**

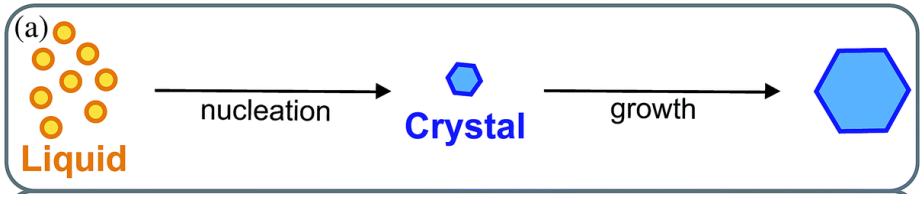
96 diverse soluzione composta da:

1. **Agente precipitante**
2. **Buffer (pH costante)**
3. **+/- Sale**



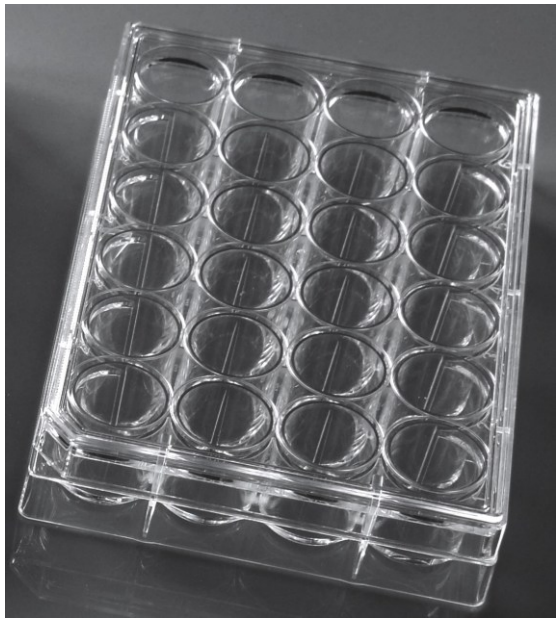
TERMODINAMICA

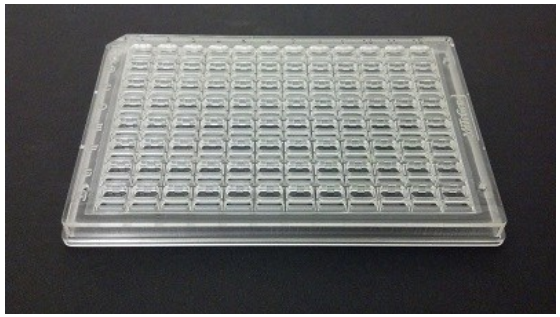
$\Delta G = \Delta H - T\Delta S$





METODI DI CRISTALIZZAZIONE





A MANO

24 condizioni/piastra

Gocce 2-5 µL

ROBOT

96 condizioni/piastra

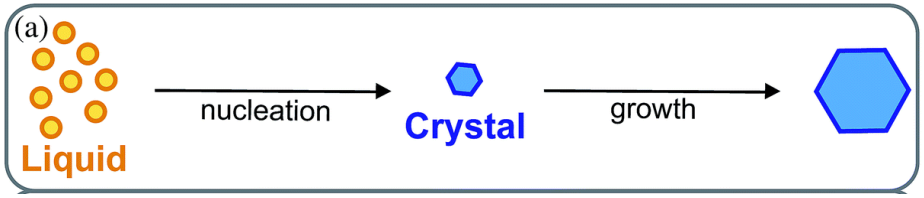
Gocce 0.2-0.8 µL



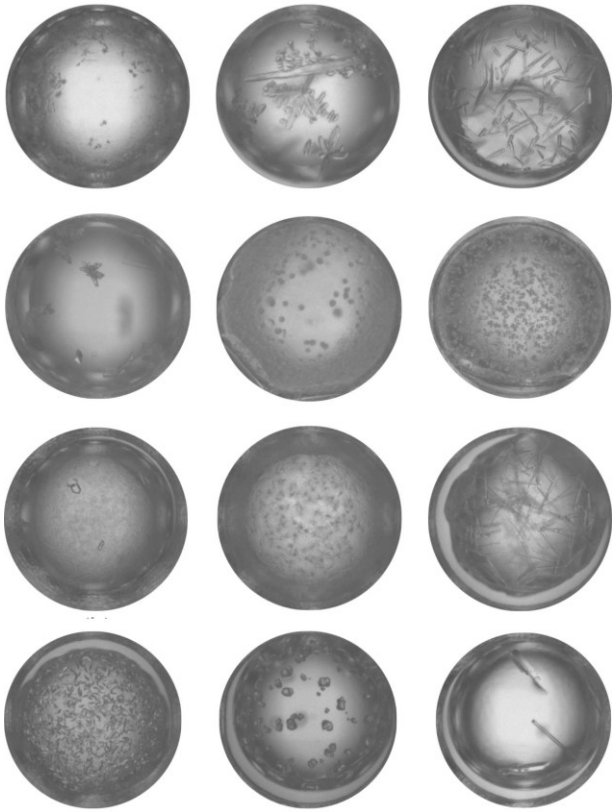
CRISTALLI DI PROTEINE

TERMODINAMICA

$\Delta G = \Delta H - T\Delta S$



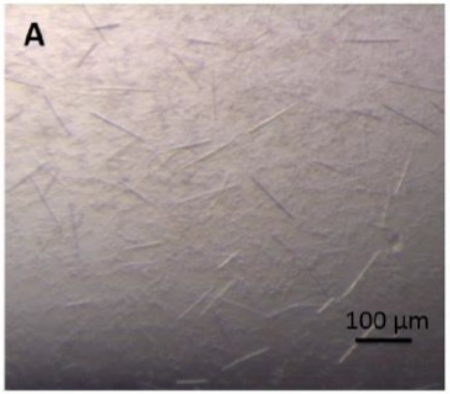


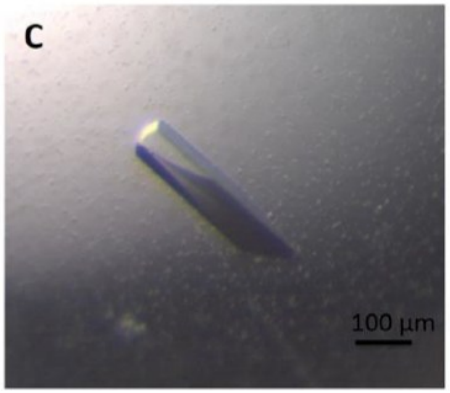


Condizioni di cristallizzazione **iniziali** dallo screen

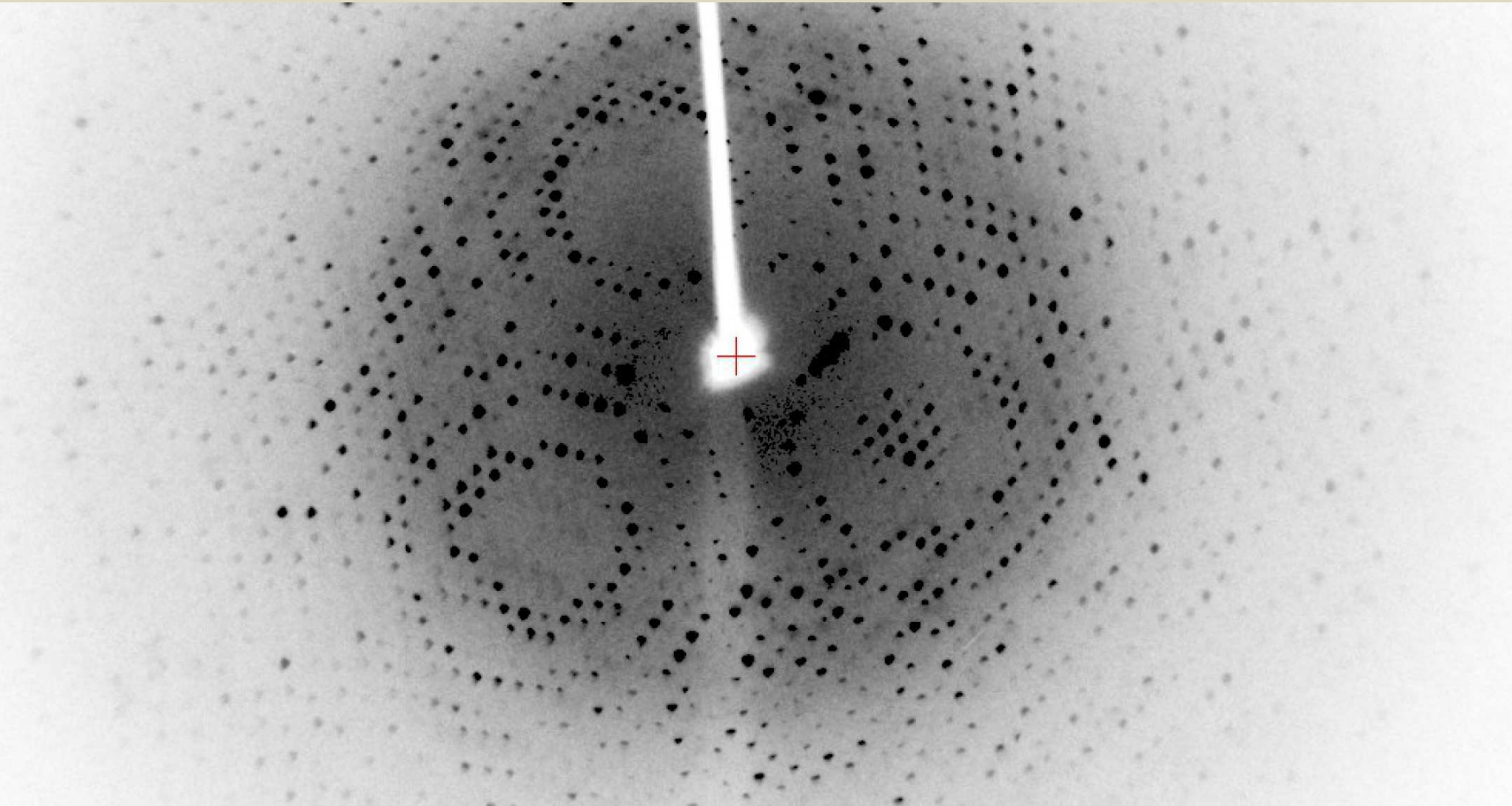
Condizioni di cristallizzazione **OTTIMIZZATE** A mano variando:

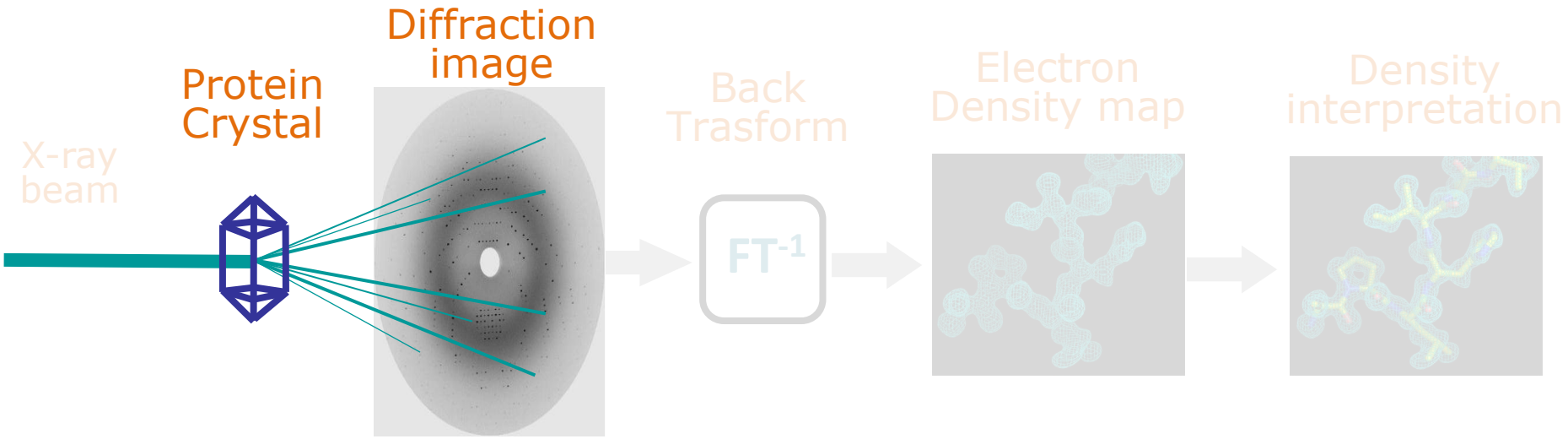
[precipitante]
[proteina]
[sali, additivi]
pH; T





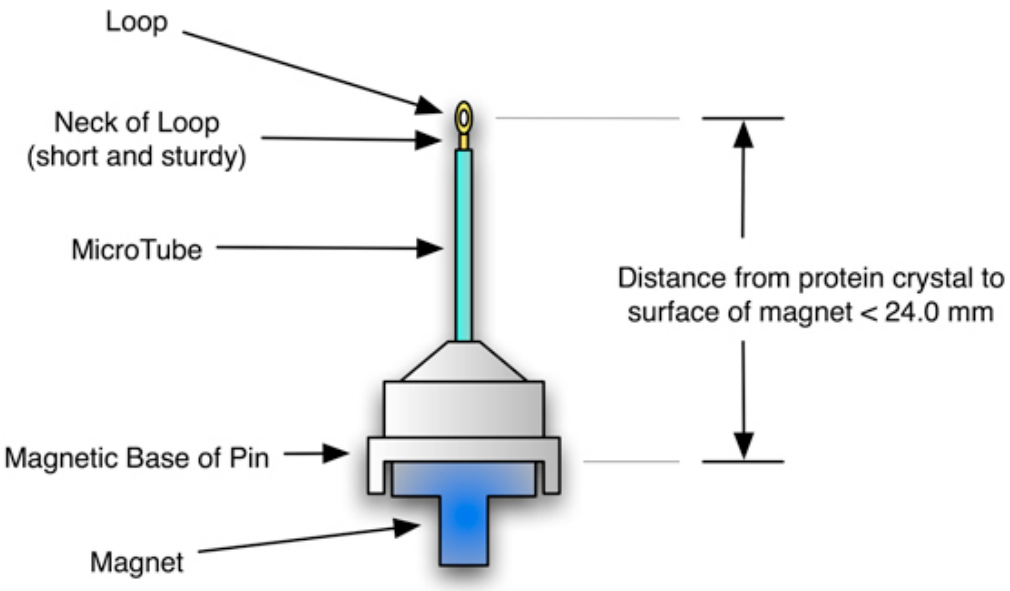
RACCOLTA DATI DI DIFFRAZIONE

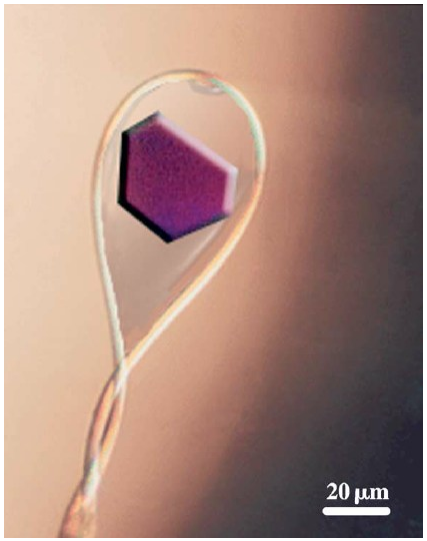




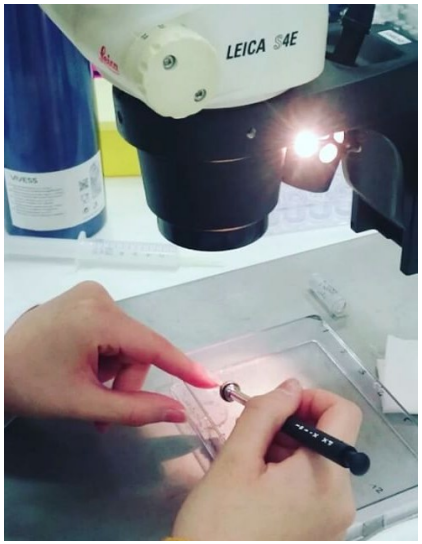
Per minimizzare il **danno radiologico** il cristallo viene **congelato in N_2 liquido (100K) - vitrificazione**

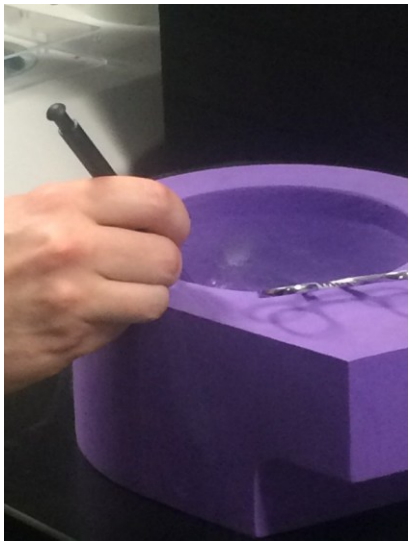
RACCOLTA DATI











RACCOLTA DATI

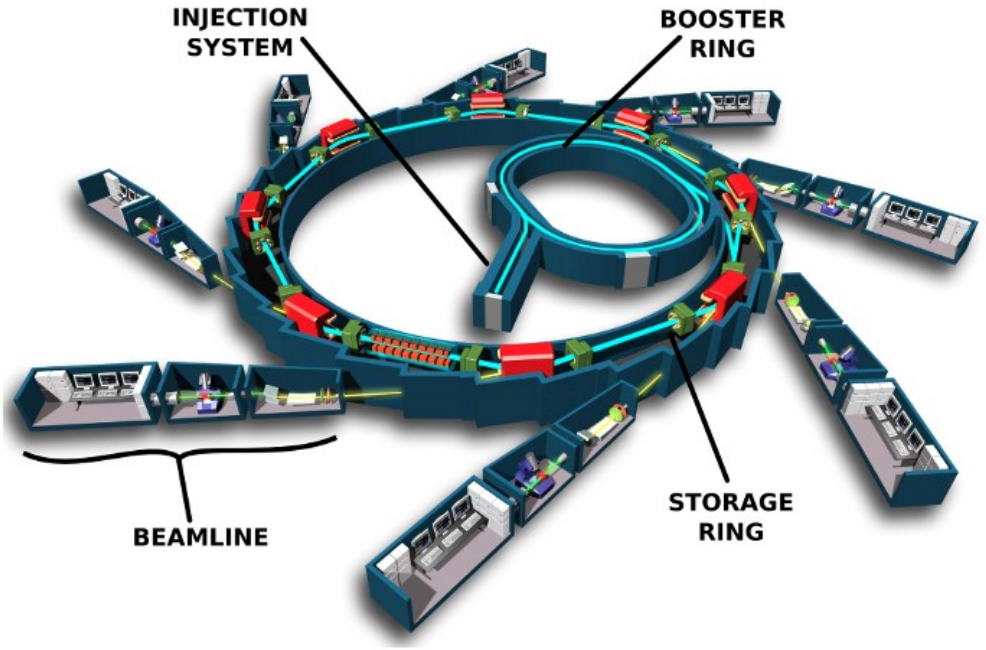




Anodo rotante



SINCROTRONE



RACCOLTA DATI

Lunghezza d'onda fissa
$\lambda(Cu_{k\alpha}) = 1.5486Å$

- bassa del intensità raggio diretto
- tempi di esposizione lunghi
- basso danno radiologico

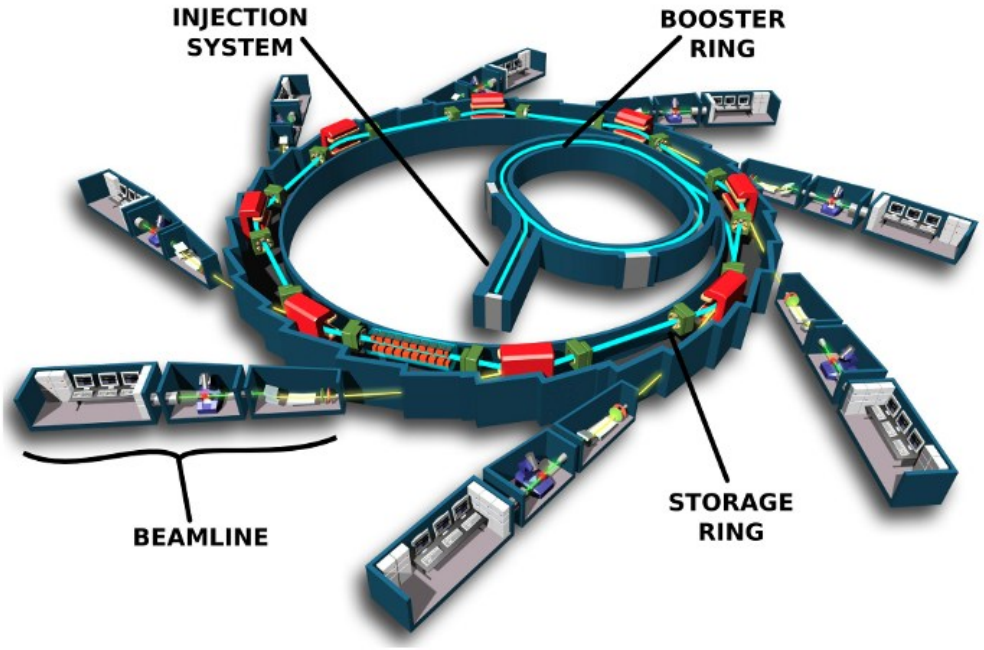
Anodo rotante



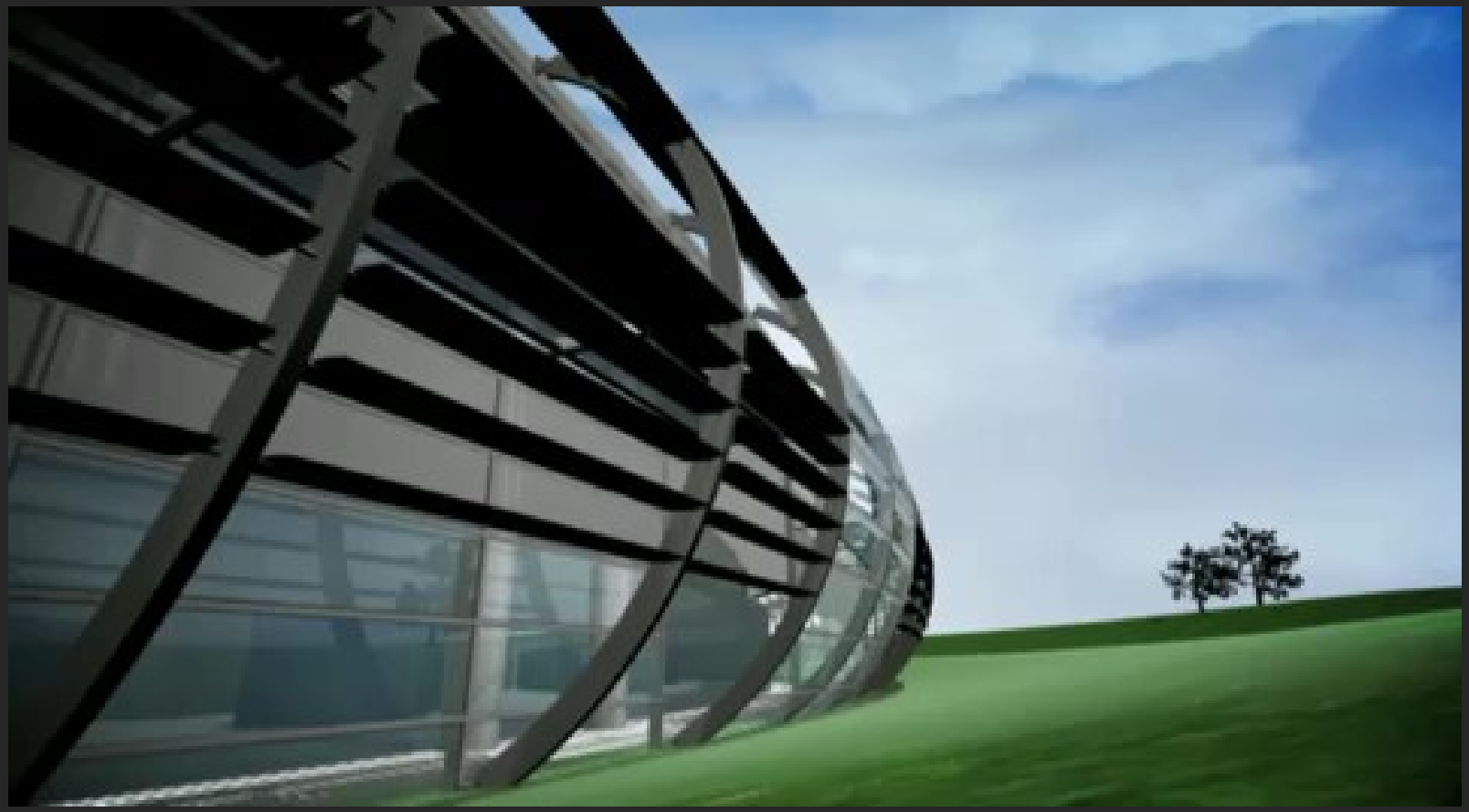
Lunghezza d'onda fissa e/o variabile

- intensità 100-1000 maggiore
- tempi di esposizione brevi
- alto danno radiologico
- λ modulabile

SINCROTRONE

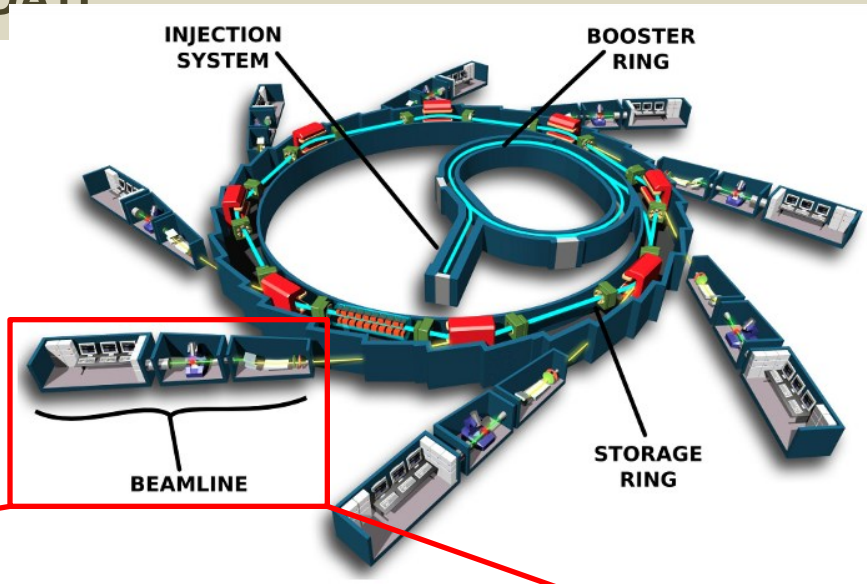


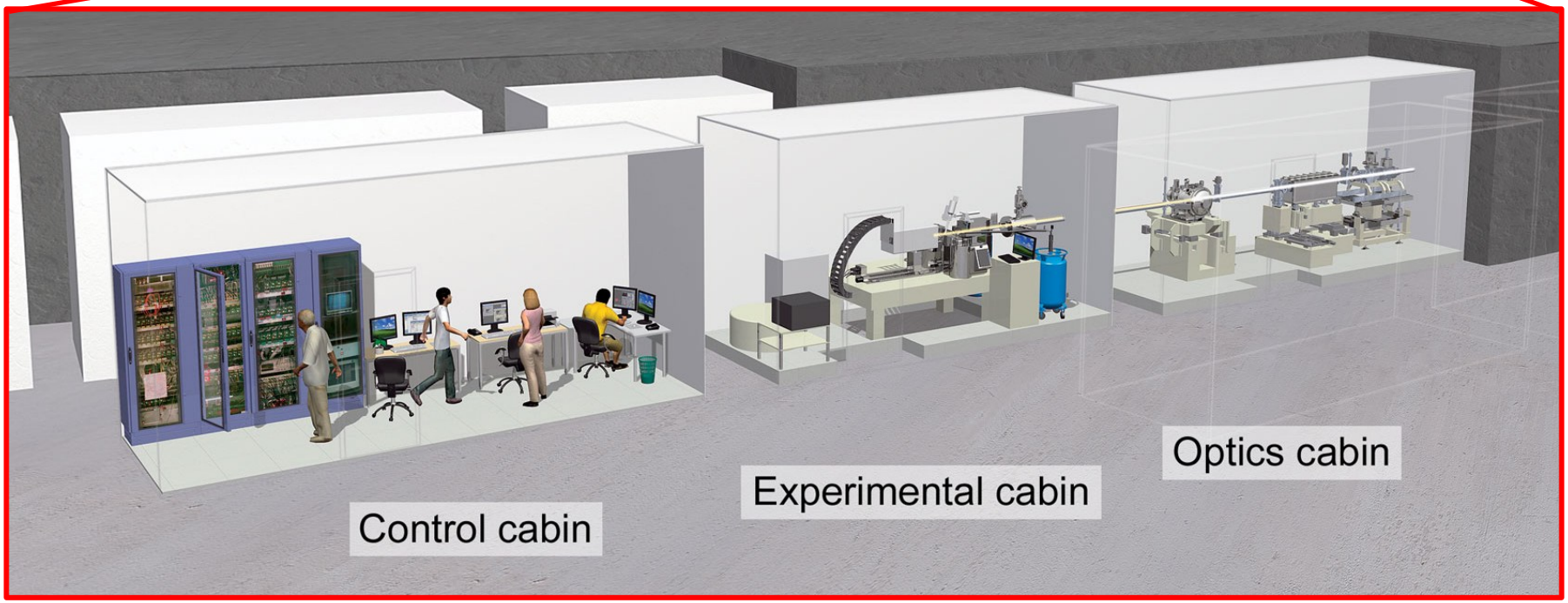
LUCE DI SINCROTRONE



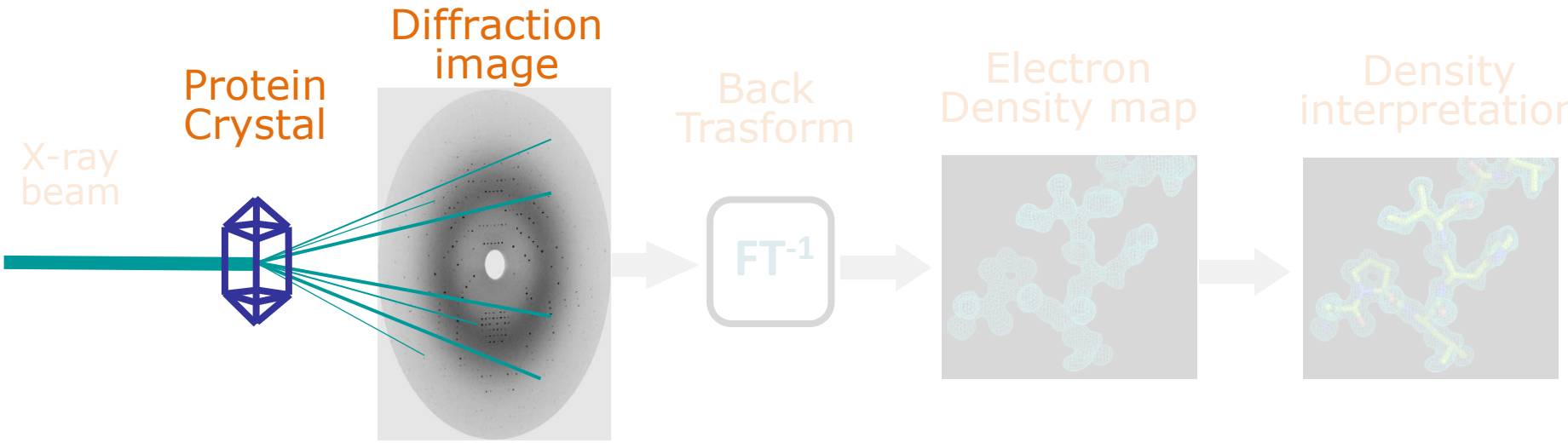
https://www.youtube.com/watch?v=qjob1EOeM5M

SINCROTRONE – Beamline

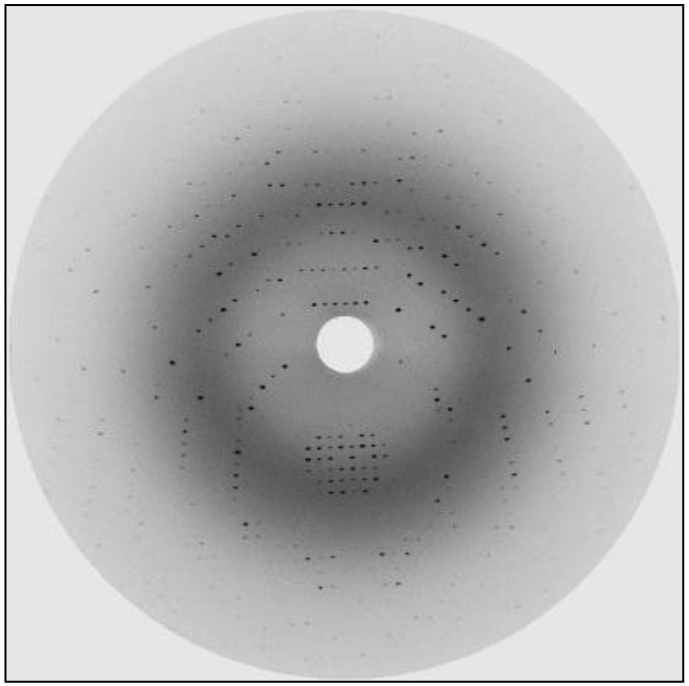




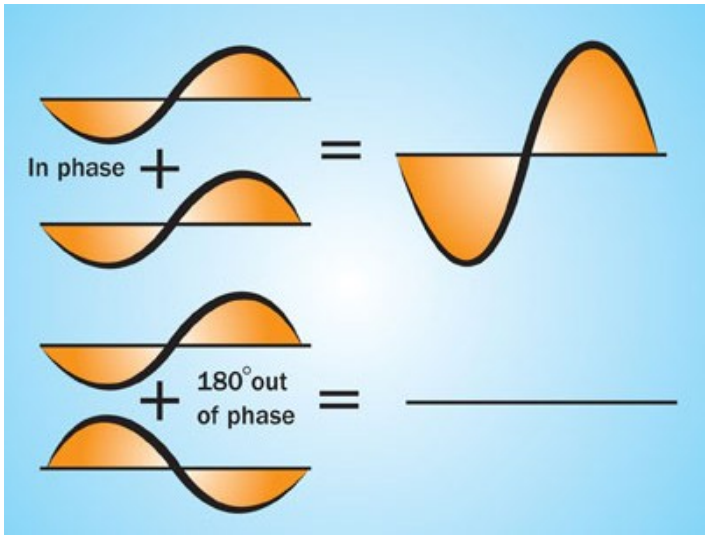
RACCOLTA DATI

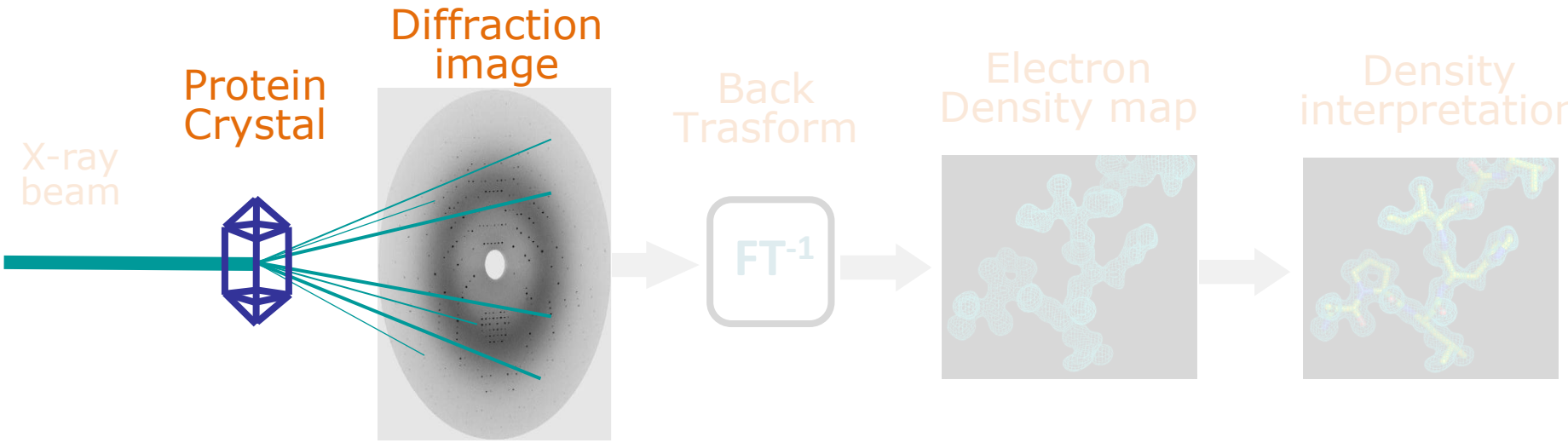


Perché vediamo i puntini (SPOT)?

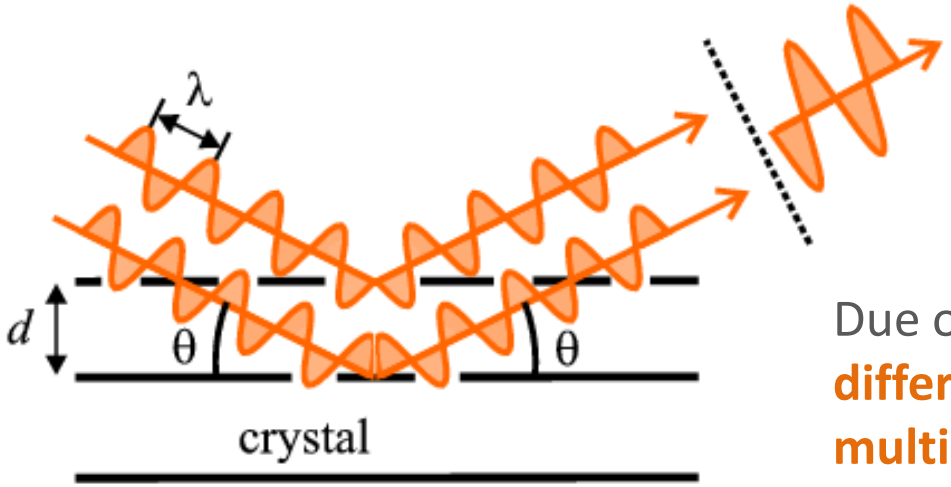


l'immagine di diffrazione è la **SOMMA DI TUTTI I FOTONI DIFFRATTI.** Solo le onde **IN FASE** si **SOMMERANNO** ed emergeranno dal cristallo



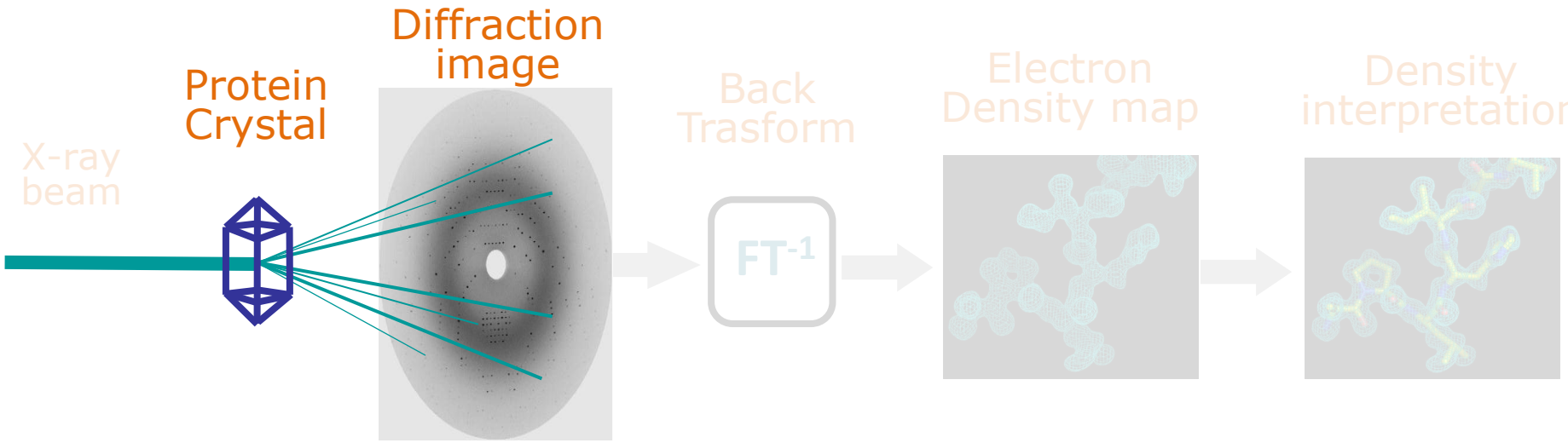


Perché vediamo i puntini (SPOT)?

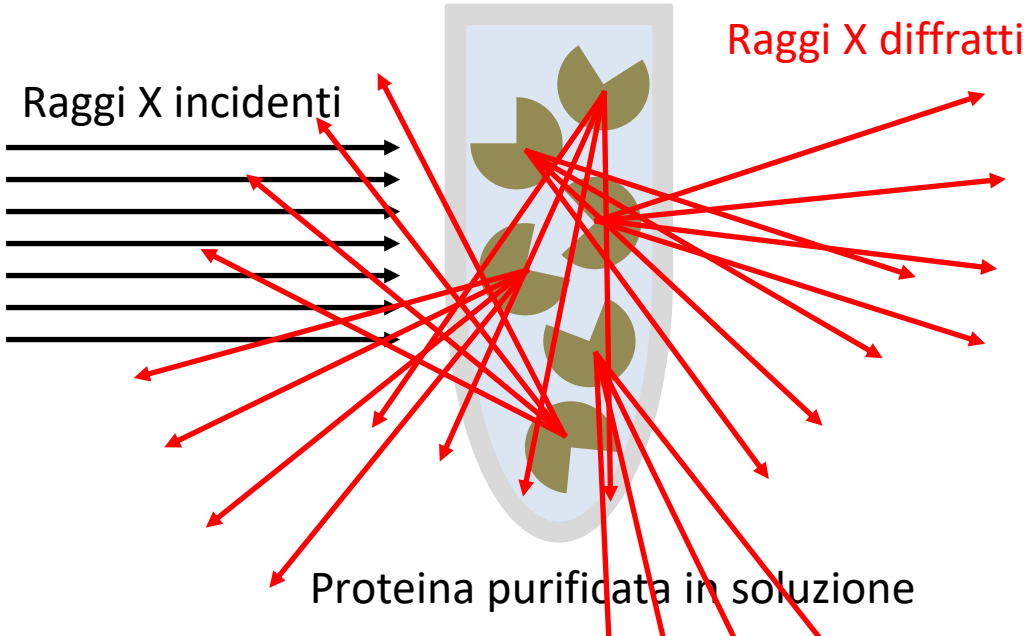


SPOT

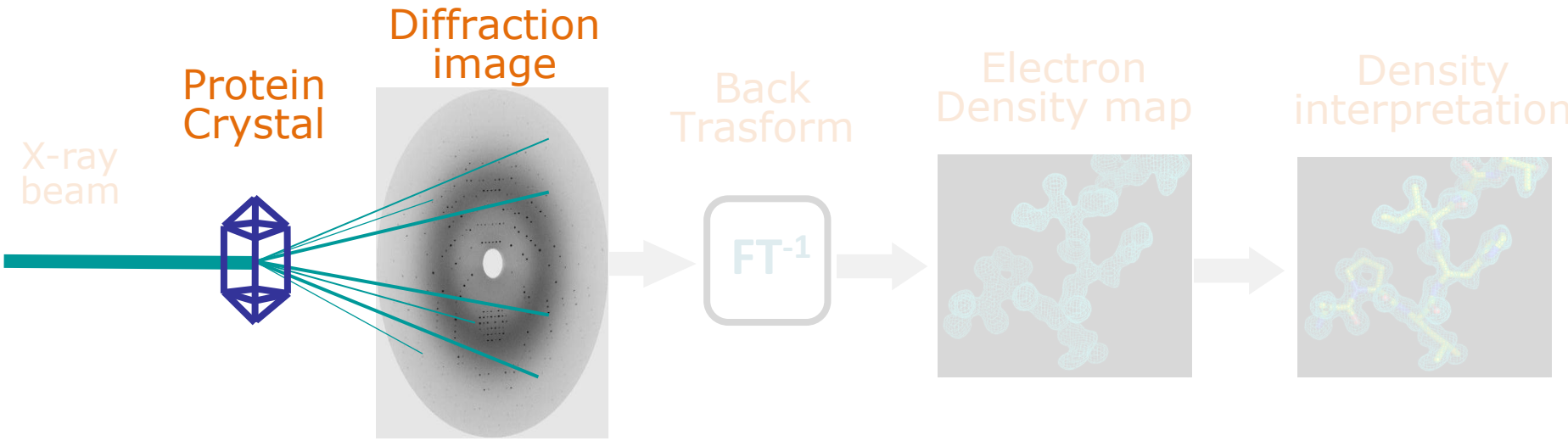
Due onde sono in fase quando **la differenza di cammino ottico è un multiplo della lunghezza d'onda**



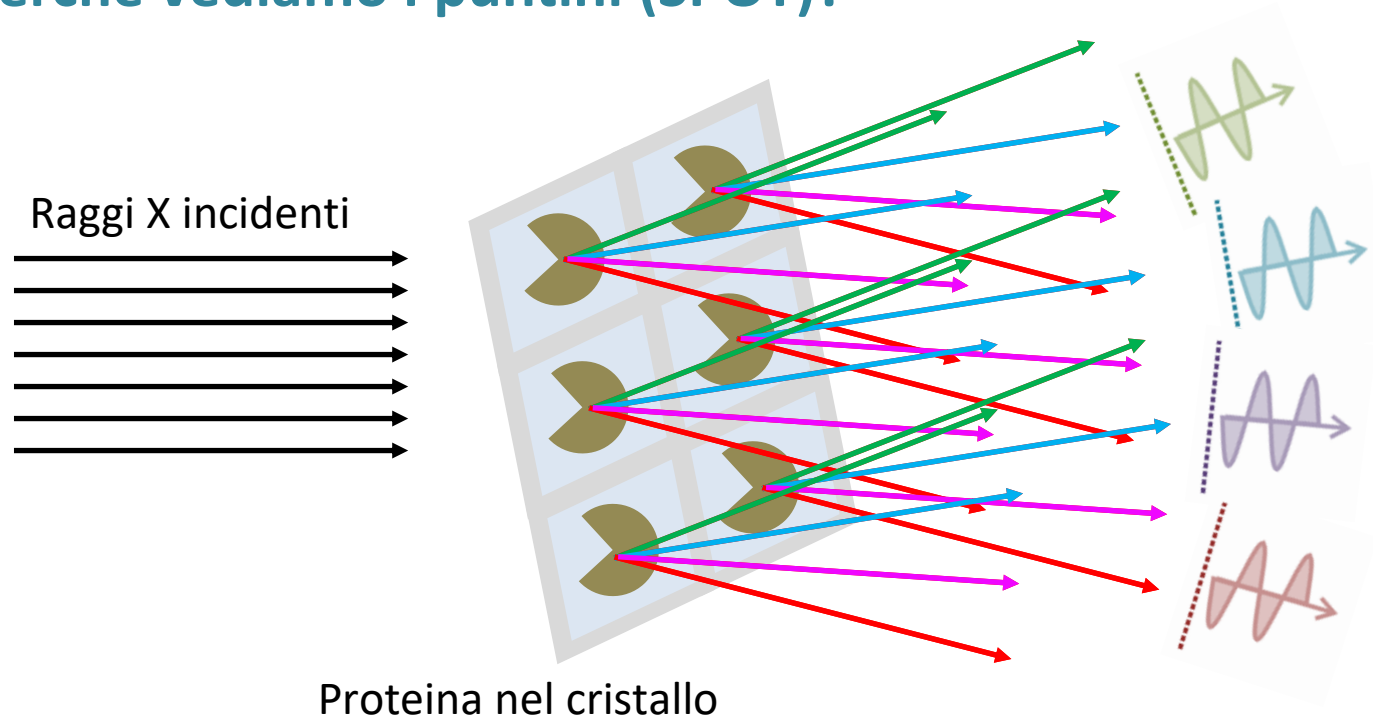
Perché vediamo i puntini (SPOT)?



In soluzione ogni proteina diffrange i raggi X, ma **le direzioni dei raggi diffratti** sono casuali e quindi **la probabilità che un numero sufficiente di essi sia in fase** per essere rilevato **è praticamente nulla**



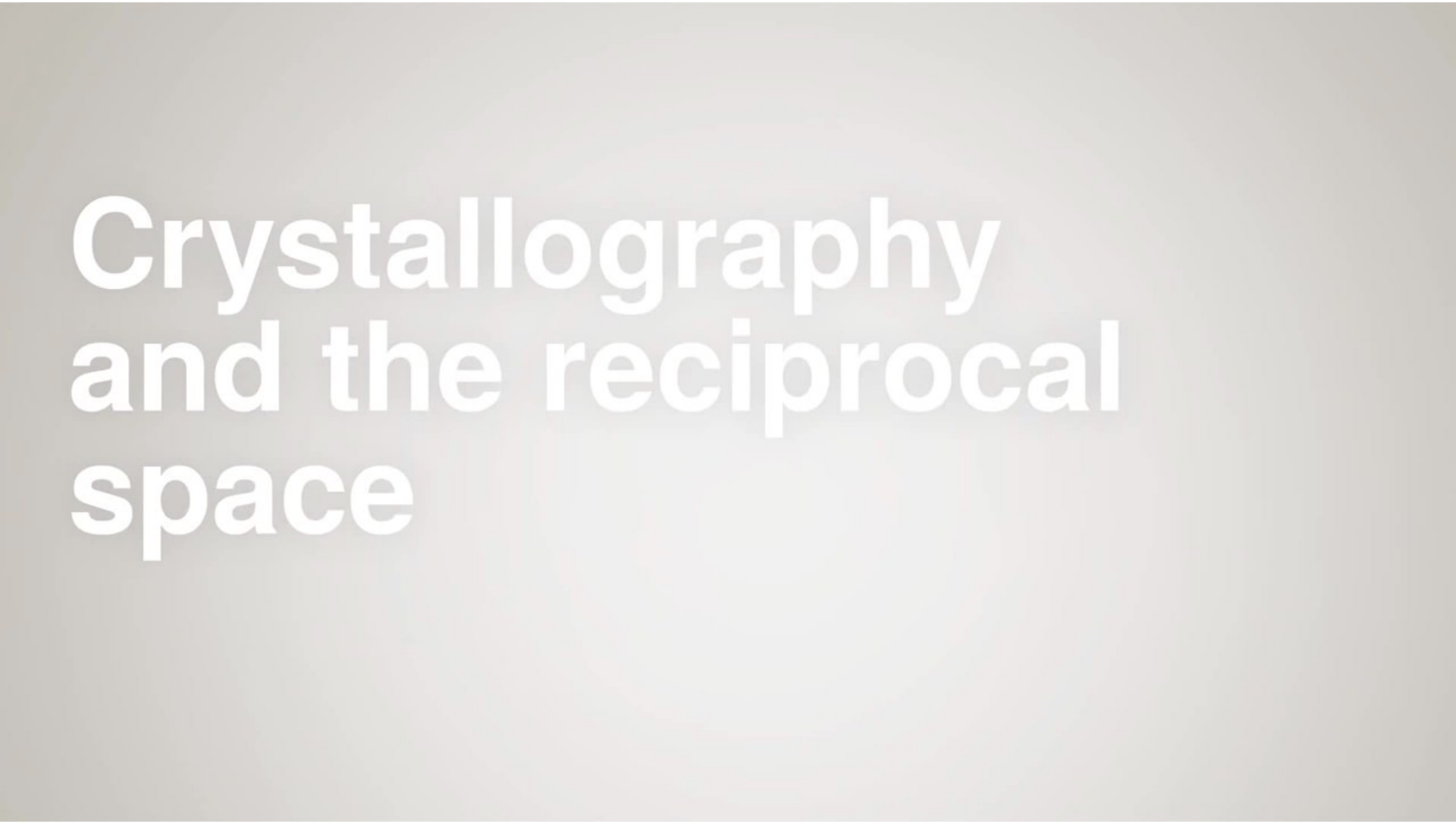
Perché vediamo i puntini (SPOT)?



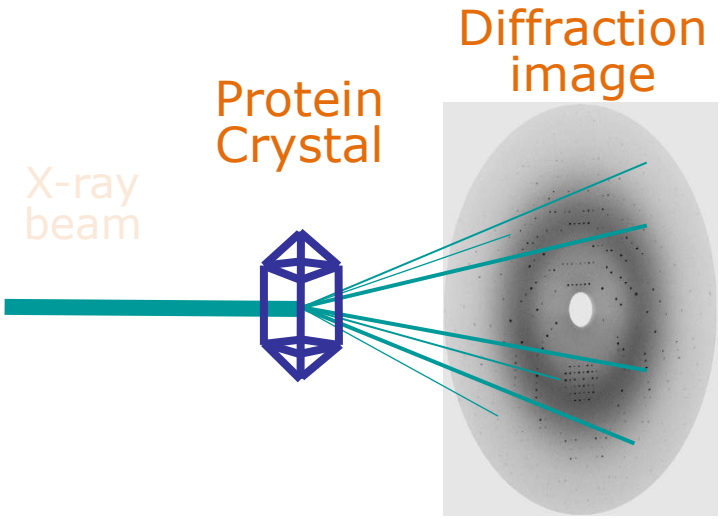
Nel cristallo le singole proteine sono disposte in maniera ordinata quindi **c'è una elevata probabilità che IN ALCUNE DIREZIONI i raggi X diffratti siano in fase e producano gli spot**

RACCOLTA DATI – il reticolo reciproco

La Trasformata di Fourier di un reticolo è sempre un reticolo le cui dimensioni sono **inversamente proporzionali** a quelle del reticolo reale

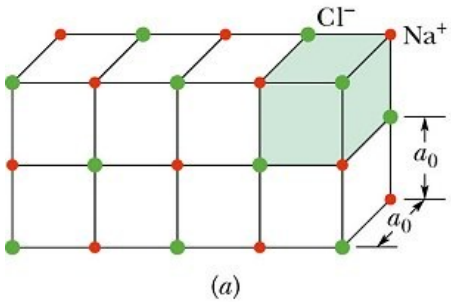


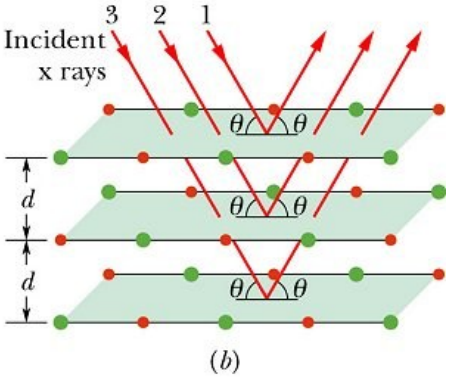
https://www.youtube.com/watch?v=DFFU39A3fPY



Ogni SPOT è **COME SE** fosse **RIFLESSO** da una specifica FAMIGLIA DI PIANI CRISTALLOGRAFICI

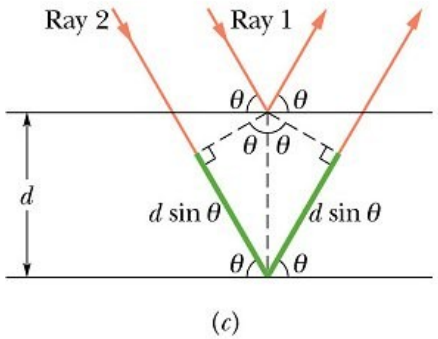
SPOT = RIFLESSIONI

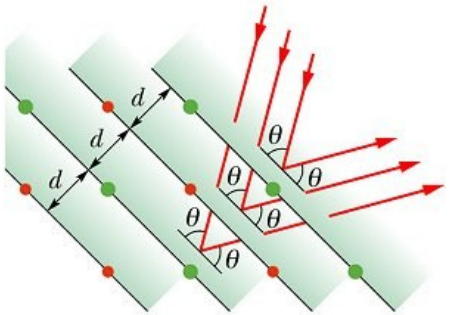


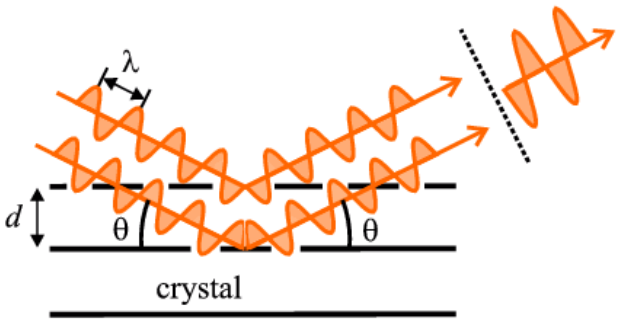


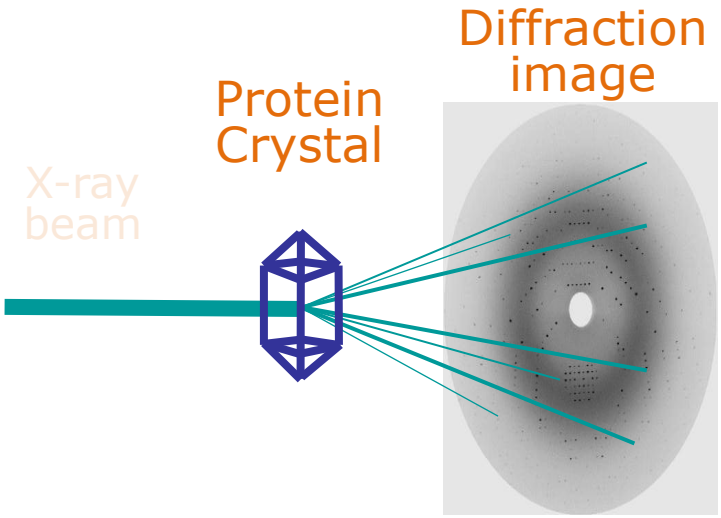
Bragg's law

$$2d \sin\theta = n\lambda$$



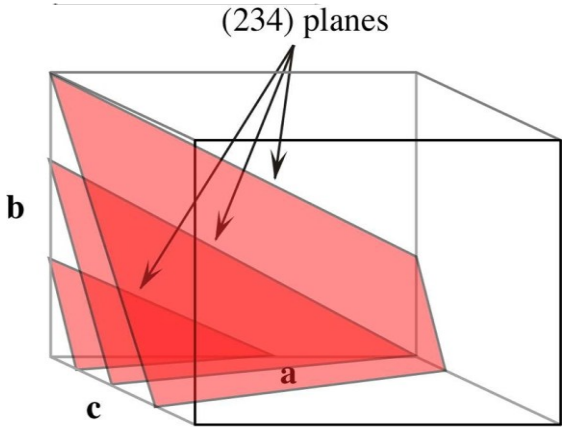




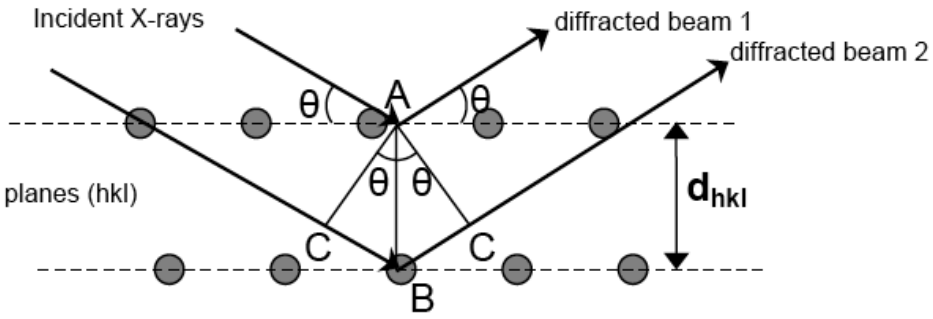


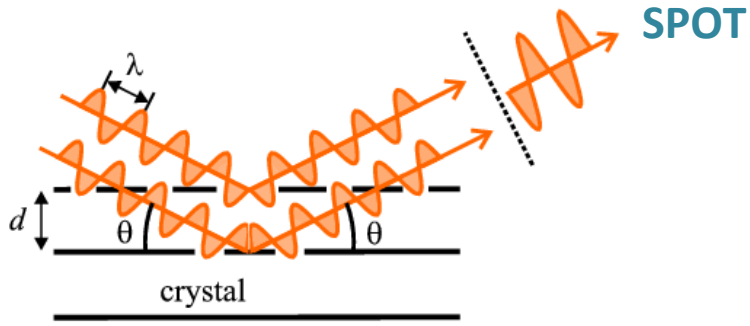
Ogni **SPOT** del reticolo reciproco è **COME SE** fosse originato dal **RIFLESSO** dei raggi-X su una specifica **FAMIGLIA DI PIANI CRISTALLOGRAFICI** Del reticolo reale

SPOT = RIFLESSIONI

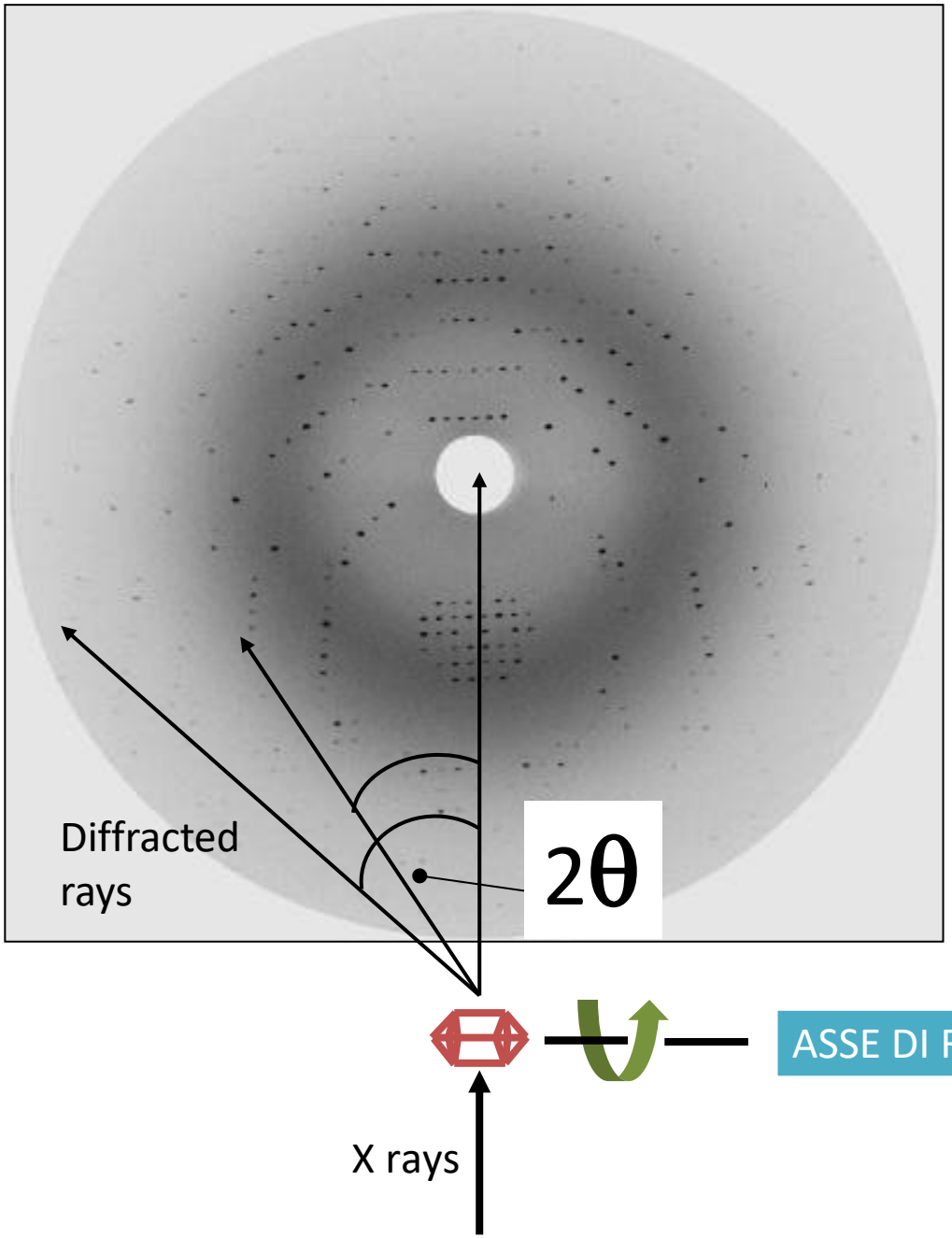


- **Ad ogni riflessione corrisponde una famiglia di piani**
- **Ad ogni famiglia di piani assegno degli indici: *h k l***

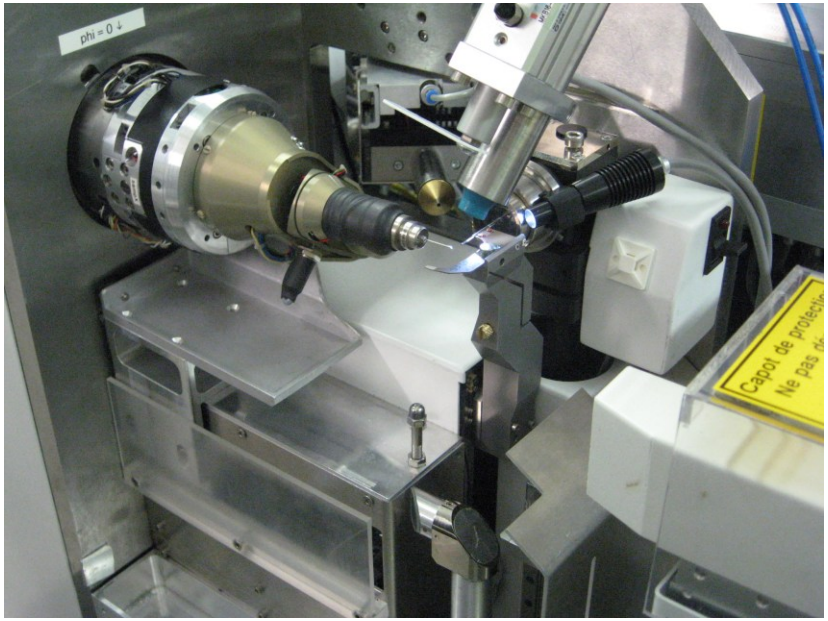




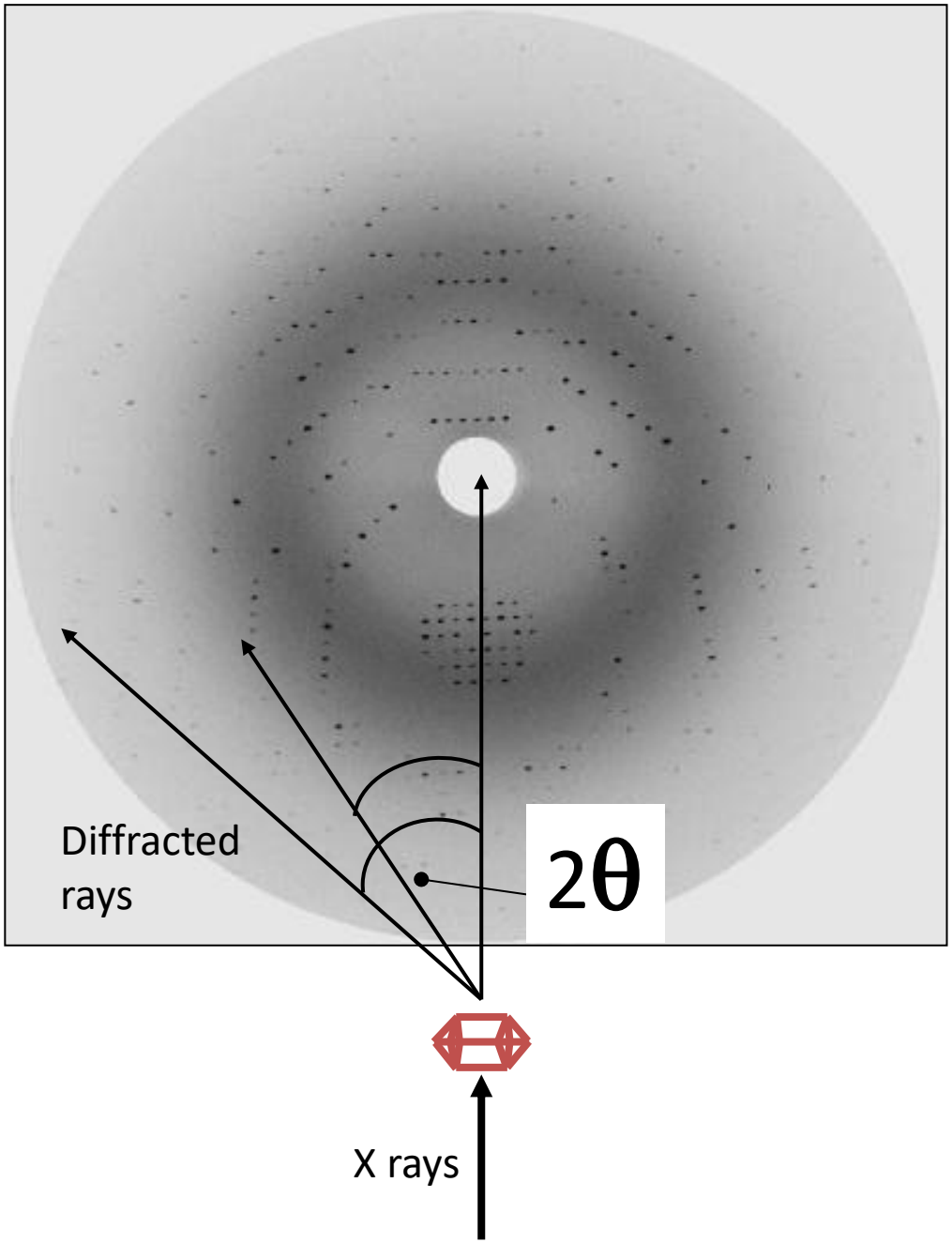
RACCOLTA DATI



PER OTTENERE UNA
RACCOLTA DATI COMPLETA
(TUTTE LE POSSIBILI RIFLESSIONI)
RUOTO IL CRISTALLO
E RACCOLGO PIÙ IMMAGINI



ASSE DI ROTAZIONE DEL CRISTALLO

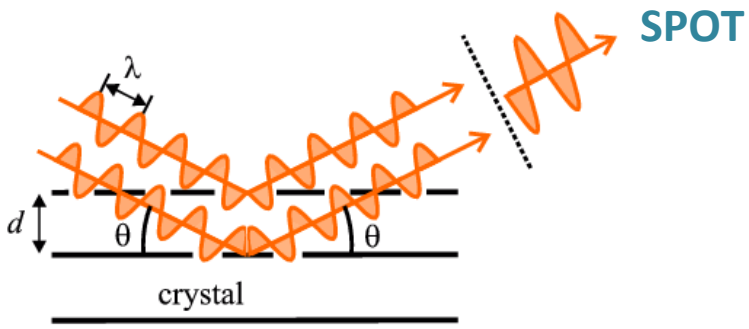


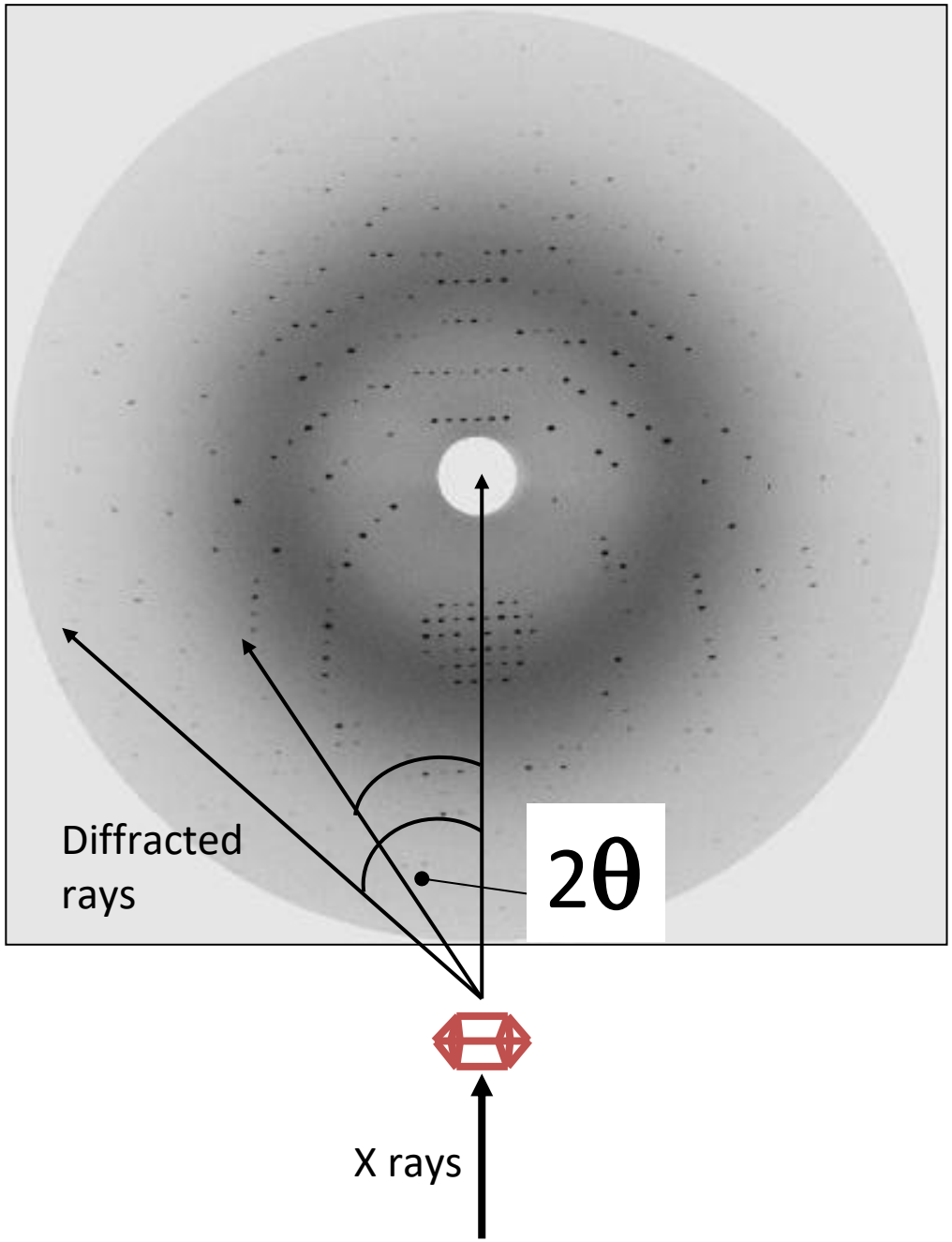
Bragg's law

$2d \sin\theta = n\lambda$

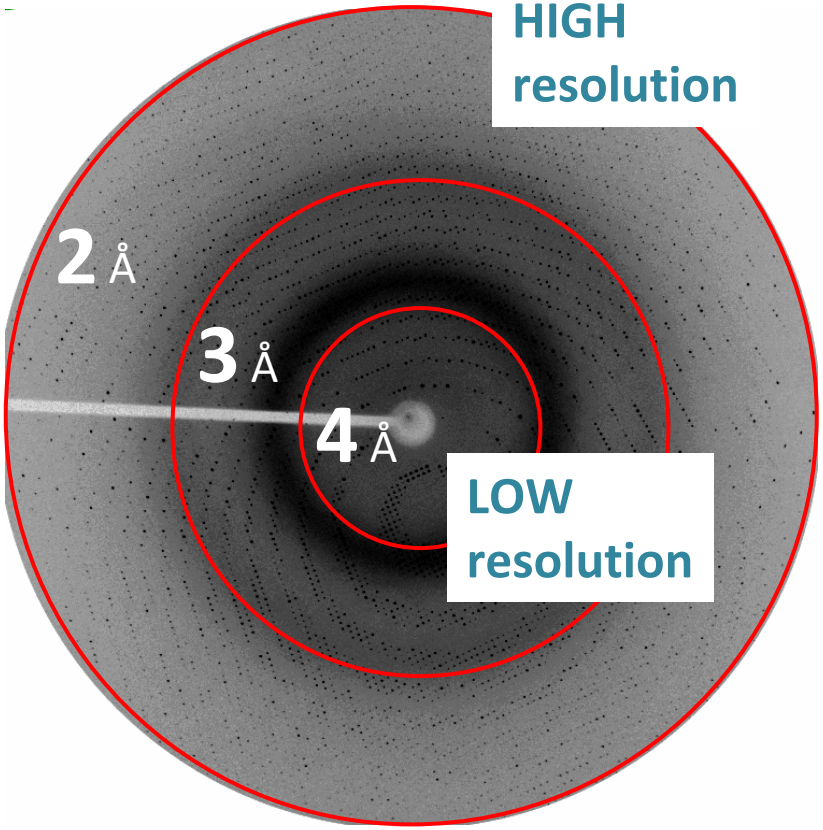
$d = n\lambda / 2\sin\theta$

d_{min} = RESOLUTION [Å]

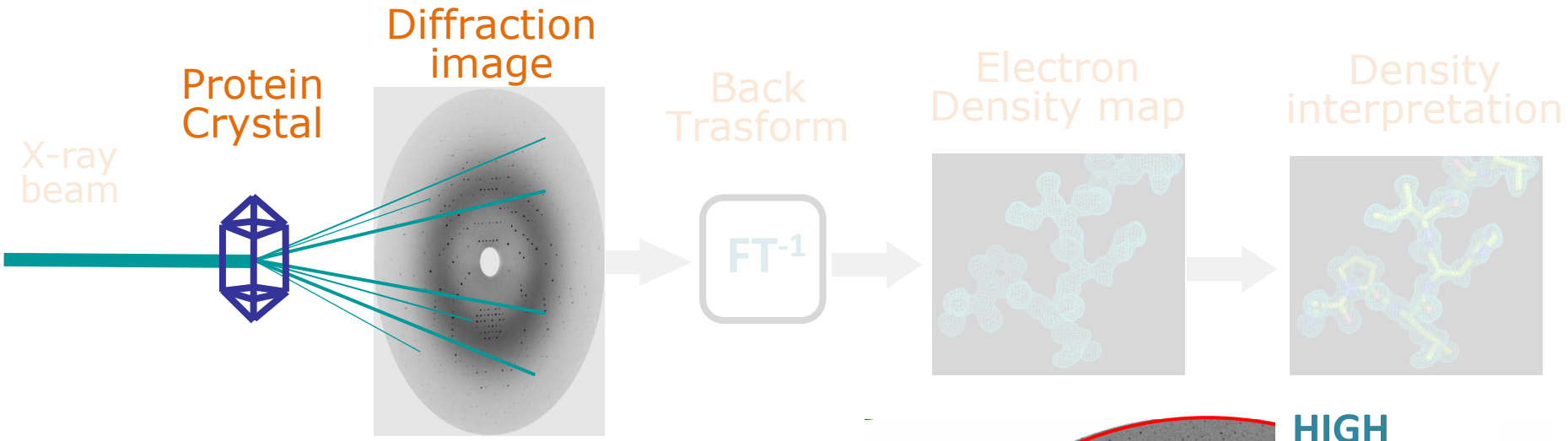




d_{min}= RESOLUTION [Å]

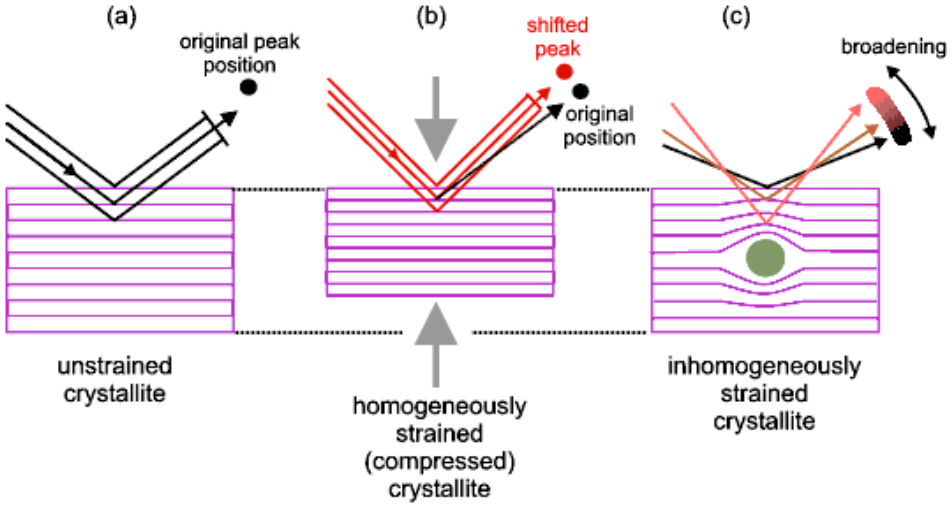


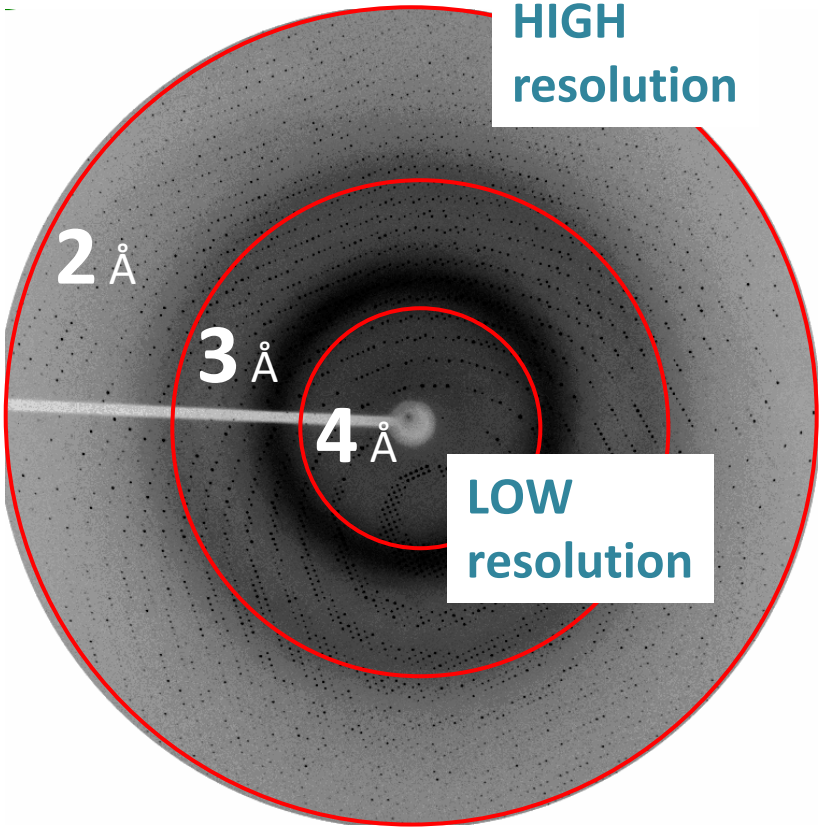
RISOLUZIONE DEI DATI



Resolution depends on the CRYSTAL

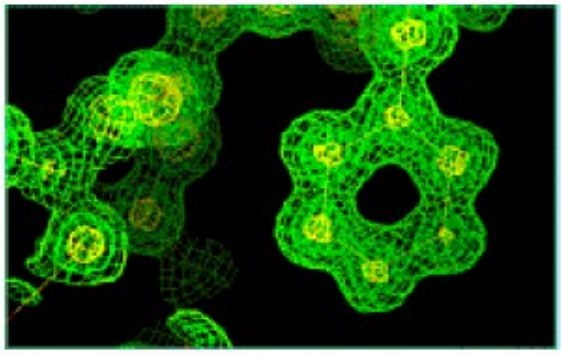
- **Quality (order)**
- **Dimension**



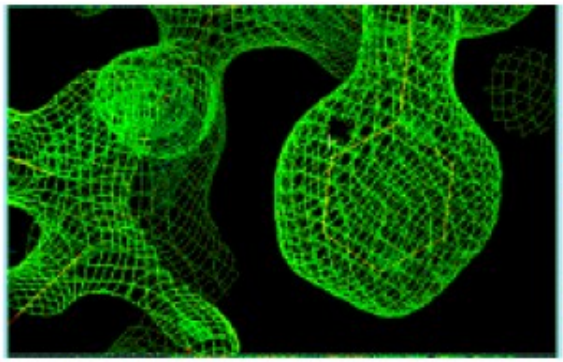


ALTA RISOLUZIONE → MAPPE MIGLIORI

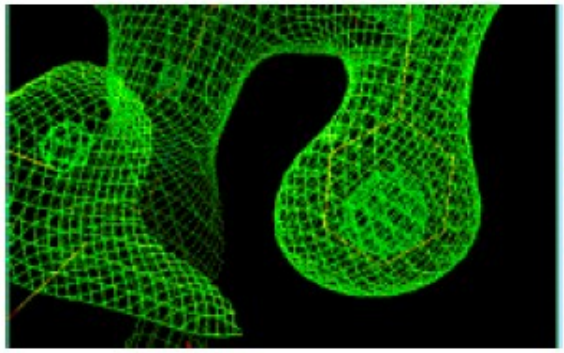
PROTEIN FEATURES	RES. (Å)
Alpha-helix	9
Beta-sheet	4
‘random’ main-chain	3.7
Aromatic side-chains	3.5
Interpretable conformation of side-chains	2.9
Ordered water molecules	2.7
Correct stereochemistry at Cbeta of isoleucine	2.2
Individual atoms	1.5



1.2 Å



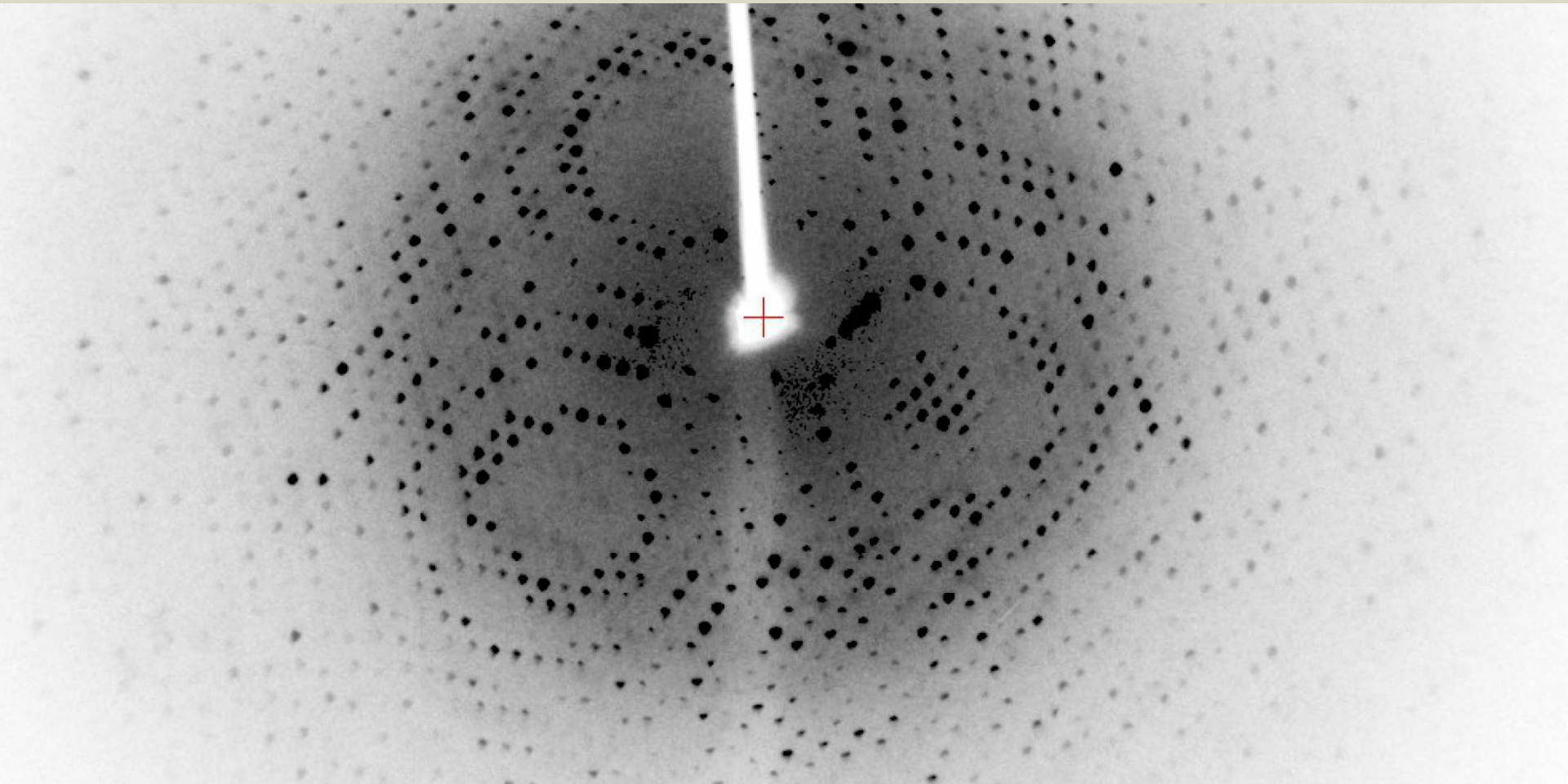
2.0 Å



3.0 Å

ANALISI DEI DATI

Indicizzazione – Integrazione - Scalaggio



Con la RACCOLTA DATI i programmi di analisi fanno 3 cose:

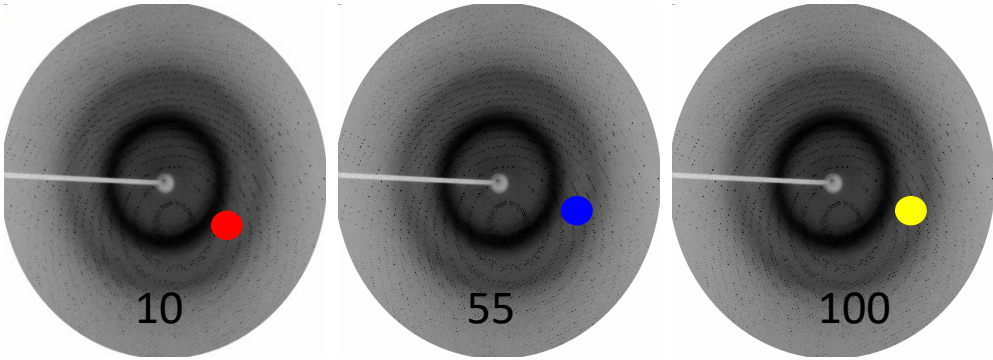
INDICIZZAZIONE

Analizzo le prime immagini di diffrazione:

- misurano **le posizioni** dei picchi

⇓

- ricavano **il gruppo spaziale**



INTEGRAZIONE

Analizzo tutte le immagini di diffrazione:

- misurano le **Intensità di ogni picco (spot)**

$I_1(234)$ $I_2(234)$ $I_3(234)$

SCALAGGIO

Unisco tutti i dati ottenuti:

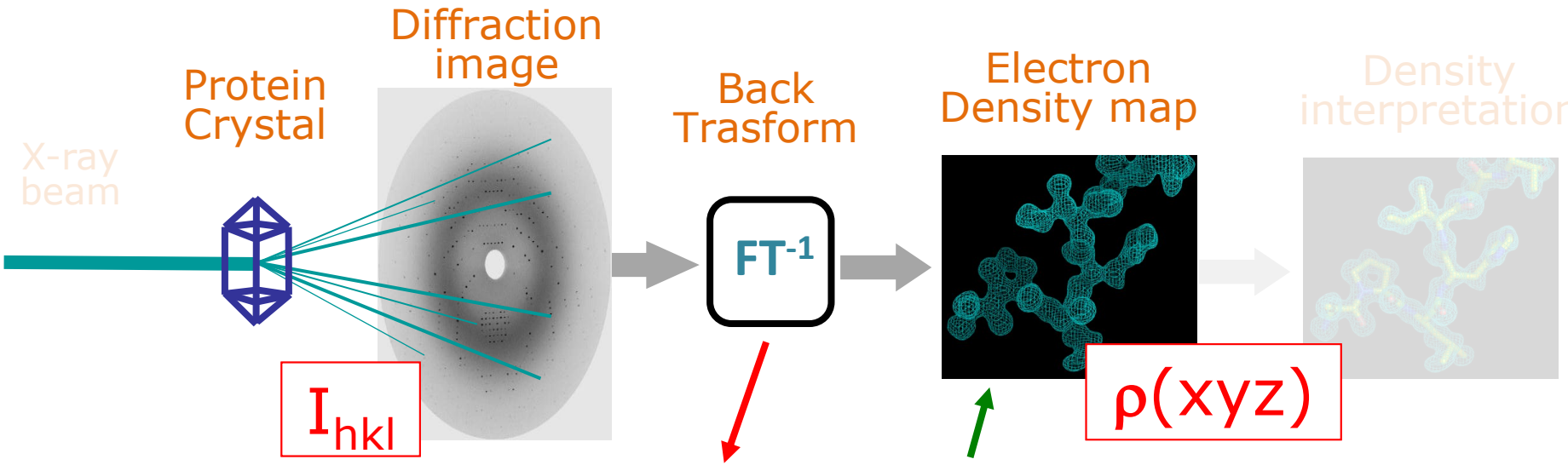
- scrivono in un **output le Intensità relative ad ogni posizione nello spazio reciproco (spot) a cui assegnano gli indici *h, k, l***

$$R_{merge} = \frac{\sum_{hkl}\sum_{i=1}^{n}\left|I(hkl) - I_i(hkl)\right|}{\sum_{hkl} N \times I(hkl)}$$

- **R merge**
- **mosaicità** (disordine intrinseco)
- **completezza**, ridondanza
- ***I/σ(I)*** (segnale/rumore)

STATISTICHE
(indicano la bontà del dato sperimentale)

L'ESPERIMENTO DI DIFFRAZIONE



$$\rho(xyz) = 1/V \sum_{hkl} |F_{hkl}| \exp(i\alpha_{hkl}) \exp(-2\pi i\, hx + ky + lz)$$

α_{hkl} is lost in the experiment - the phase problem!

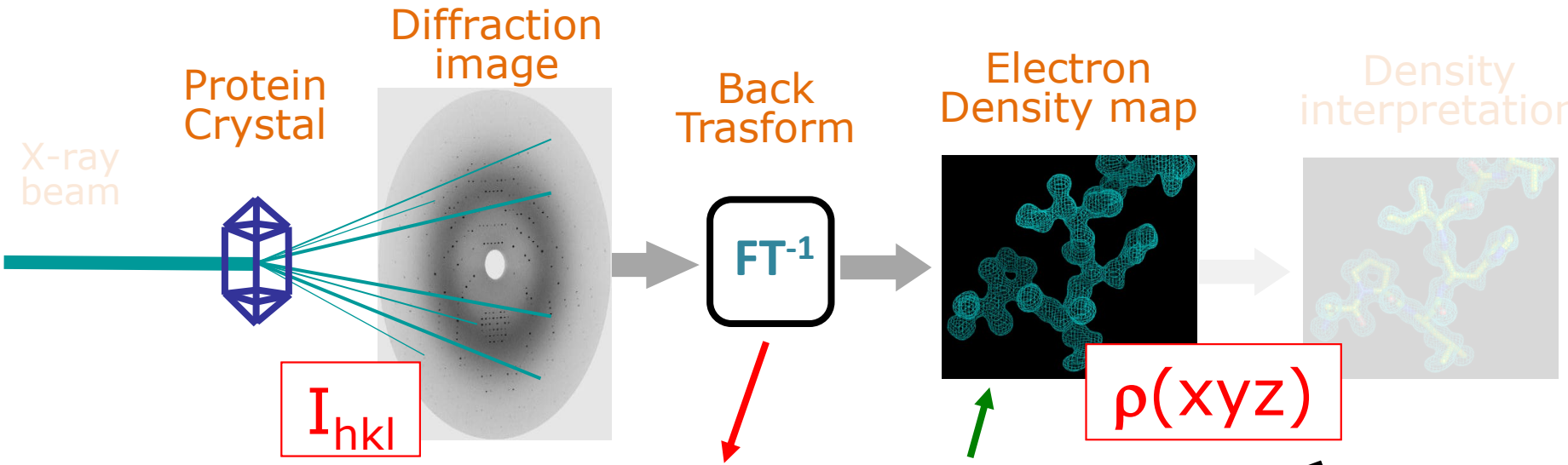
PERSA

structure factor amplitude $|F_{hkl}| \propto (I_{hkl})^{1/2}$

OTTENUTI DAI DATI DI DIFFRAZIONE



L'ESPERIMENTO DI DIFFRAZIONE



$$\rho(xyz) = 1/V \sum_{hkl} |F_{hkl}| \exp(i\alpha_{hkl}) ex$$

...in the experiment -
the phase problem!

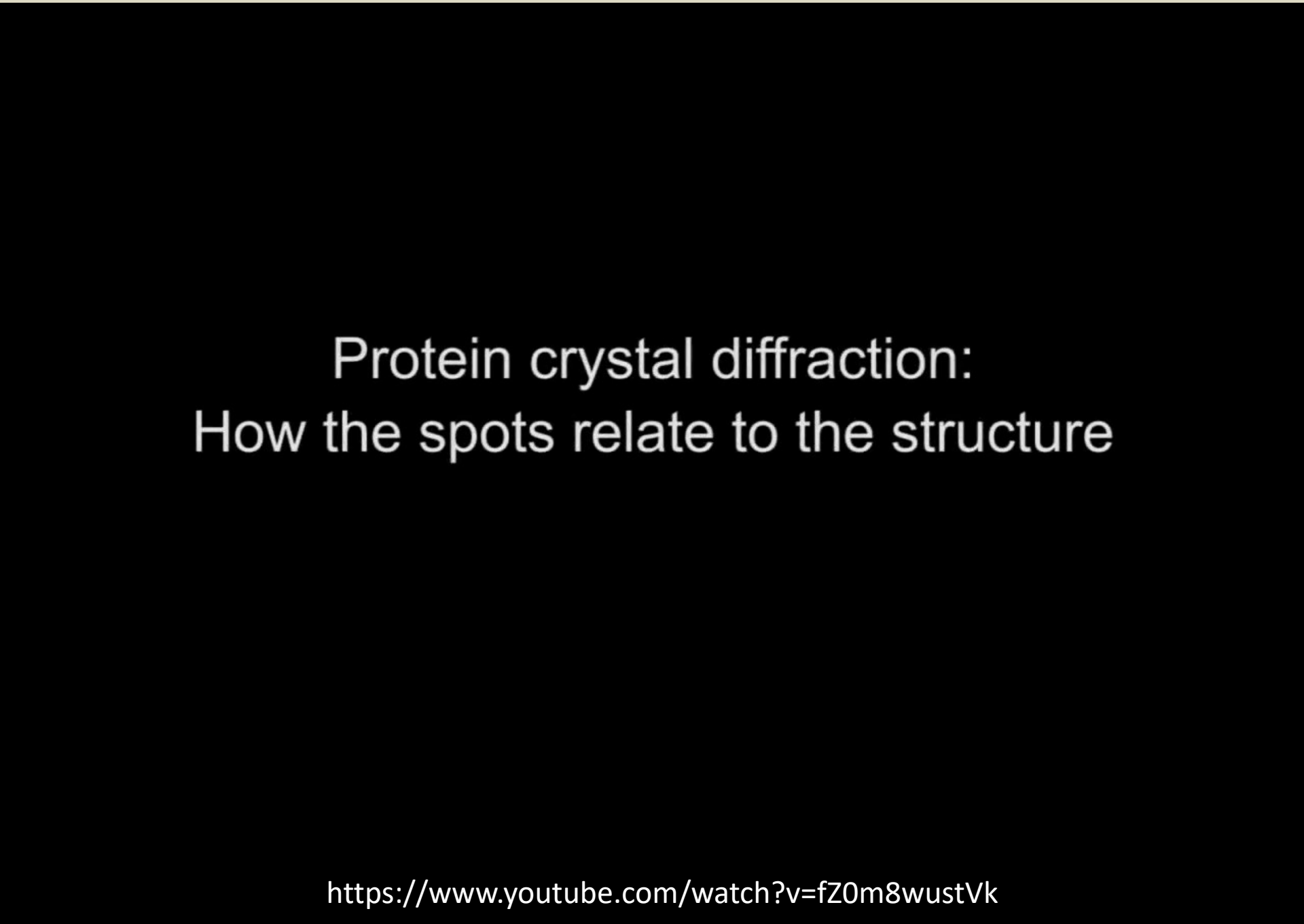
str...

$|F_{hkl}| \propto (I_{hkl})^{1/2}$

PROBLEMA DELLA FASE



L’ESPERIMENTO DI DIFFRAZIONE



https://www.youtube.com/watch?v=fZ0m8wustVk

L’ESPERIMENTO DI DIFFRAZIONE



https://www.youtube.com/watch?v=gLsC4wlrR2A

SERIAL CRYSTALLOGRAPHY

https://www.youtube.com/watch?v=PVlOb6F-MDA

VANTAGGI



- **Risoluzione atomica**
- Non ci sono limitazioni nelle dimensioni delle macromolecole

SVANTAGGI



- **Cristalli singoli**
- Altamente ordinati
- ragionevolmente grandi (30-100 µm)