

BREVE INTRODUZIONE SULLE PROTEINE

La parola **PROTEINA** viene dal Greco "prota" che significa "di primaria importanza. Questo termine fu introdotto per la prima volta da Jons Jakob Berzelius nel 1838 per grandi molecole organiche.

Le proteine sono macromolecole biologiche con un peso molecolare di alcuni milioni di Dalton (1 Dalton = 1 u.m.a. = $1,66 \cdot 10^{-27}$ Kg).

Le proteine sono costituite da ammino acidi legati in catene lineari tramite un legame peptidico. Il legame peptidico si forma tra il gruppo carbossilico di un ammino acido ed il gruppo amminico di quello successivo, tramite una reazione di condensazione.

Negli esseri viventi esistono 20 ammino acidi e il numero di proteine prodotte dalle diverse combinazioni è enorme. Per esempio per una sequenza di 10 a.a è possibile avere 20^{10} sequenze diverse, che è approssimativamente uguale a 10^{13} .

Negli esseri viventi le proteine presiedono a tutte le funzioni e ogni proteina o gruppo di proteine sono responsabili di una specifica funzione.

Perché è importante studiare la struttura 3D delle proteine?

La conoscenza accurata della struttura è un requisito fondamentale per lo studio della funzione di proteine e per gli studi di drug design nello sviluppo di nuovi farmaci.

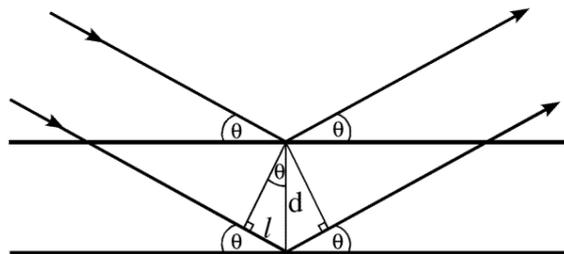
Alcune date importanti.....(cenni storici)

- Nel **1926** James B. Sumner dimostrò che gli enzimi potevano essere isolati, purificati e cristallizzati.
- Nel **1955** Sir Frederick Sanger determinò la sequenza amminoacidica completa di una proteina, l'insulina.
- Nel **1958** Max Perutz and Sir John Cowdery Kendrew risolvettero la struttura tridimensionale dell'emoglobina e della mioglobina, rispettivamente. Queste strutture vennero determinate tramite la diffrazione di raggi X.

Che cos'è la cristallografia di raggi X?

E' una tecnica sperimentale che si basa sullo diffrazione dei raggi X da parte dei cristalli.

I raggi X hanno una lunghezza d'onda nell'ordine degli Ångström ($\sim 10^{-8}$ cm) e vengono deviati dalla nuvola elettronica di un atomo di dimensioni simili.

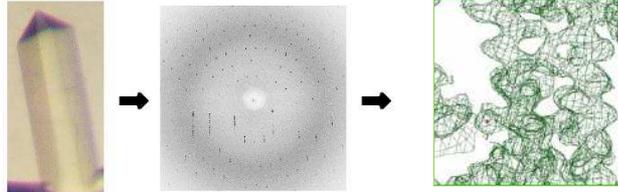


$$n\lambda = 2 d \sin\theta$$

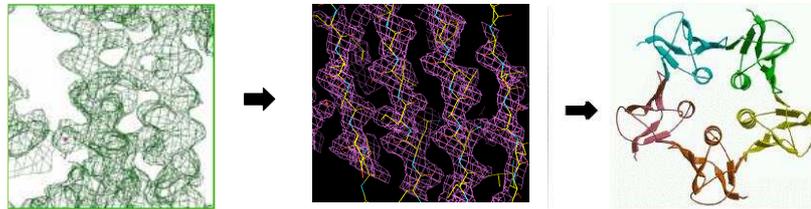
Legge di Bragg

Come si determina una struttura?

Sulla base dell'immagine di diffrazione ottenuta quindi dalla diffusione dei raggi X da parte della nuvola elettronica dell'assemblamento periodico delle molecole o degli atomi nel cristallo, si può ricostruire la densità elettronica.



Nella densità elettronica sperimentale si costruisce un modello della proteina, che viene progressivamente affinato per ottenere un'accurata struttura molecolare .



Vantaggi della cristallografia di raggi X ?

Le principali tecniche tecniche di indagine strutturale per proteine:

- Raggi X
- NMR (nuclear magnetic resonance)
- Microscopia Elettronica in condizioni criogeniche

La **cristallografia di raggi X**:

- 1- non ha limitazioni di peso molecolare al contrario di NMR
- 2- permette di raggiungere una maggiore risoluzione rispetto a EM

Svantaggi della cristallografia di raggi X ?

- Per effettuare un esperimento di diffrazione è necessario produrre dei cristalli, che siano anche di buona qualità per avere una struttura con elevata accuratezza.

- Da un singolo esperimento di diffrazione si ha un numero limitato di informazioni sul comportamento dinamico della molecola in soluzione, anche se nel cristallo la proteina mantiene una certa mobilità dovuto ad un alto contenuto di solvente (dal 20% fino al 70%) e la sua attività.

Flowchart della determinazione strutturale di una proteina



Processo di cristallizzazione

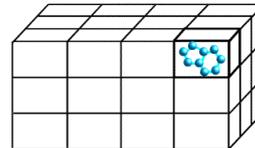
molecule



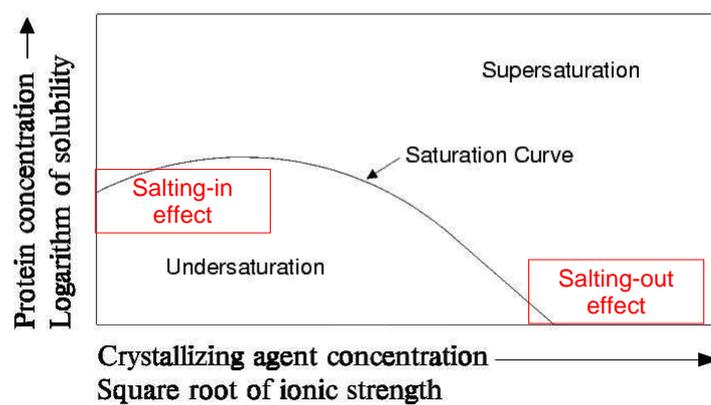
unit cell



crystal



Curva di solubilità di una generica proteina



Funzione e natura del precipitante

Il **precipitante** è una sostanza che ha il ruolo di favorire l'aggregazione delle macromolecole quindi il passaggio dalla soluzione allo stato solido. L'interazione e aggregazione viene favorita catturando le molecole di H_2O di solvatazione delle molecole proteiche.

I precipitanti più utilizzati appartengono a 3 categorie:

- **Sali**: gli ioni derivanti dalla dissociazione dei sali in soluzione competono con le proteine per le molecole di H_2O e risultano più efficienti perché la carica è concentrata in dimensioni più piccole. A bassa concentrazione di sale (bassa forza ionica) si ha una diminuzione di solubilità della proteina fenomeno detto **salting-in**. Ad alta concentrazione di sale (alta forza ionica) si ha diminuzione di solubilità della proteina fenomeno **salting-out**.
- **Polimeri**: abbassano la concentrazione di proteina perché catturano le molecole di H_2O nelle maglie delle catene polimeriche disposte in maniera random
- **Solventi organici**: disidratano le macromolecole biologiche, catturando le molecole di H_2O di solvatazione.

Barriera energetica nel processo di cristallizzazione

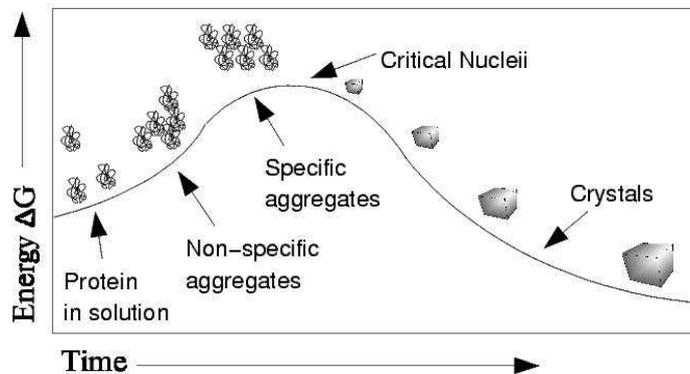
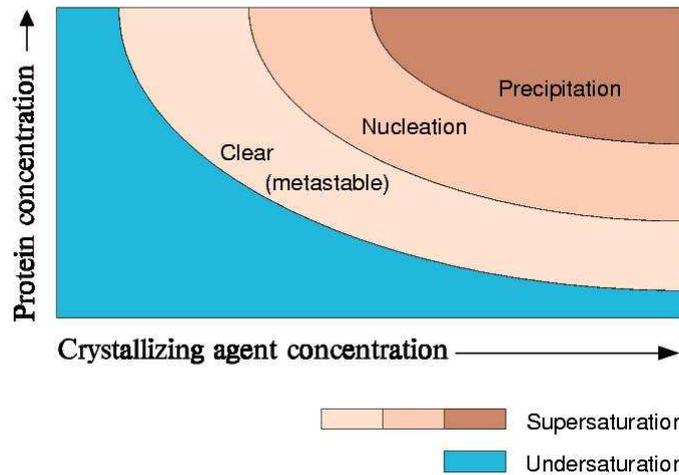


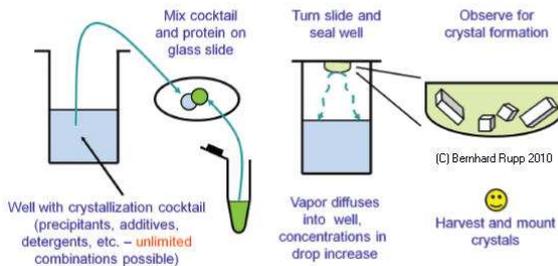
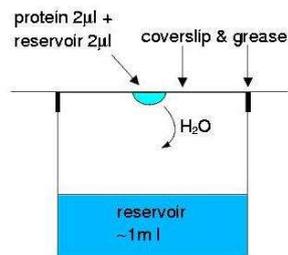
Diagramma di fase per una generica proteina



Tecniche di cristallizzazione

- Diffusione di vapore con goccia pendente

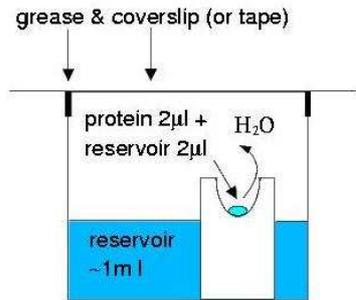
Piccoli volumi di soluzione proteica e di precipitante vengono mescolati insieme in una goccia ed equilibrati con un volume più grande di soluzione riserva contenente il precipitante, tampone eventuale additivi.



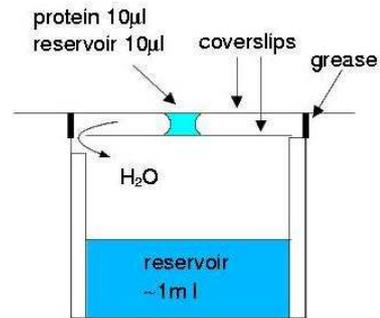
Per differenza di concentrazione del precipitante tra la goccia e la riserva si crea una diffusione di vapore che determina una diminuzione di volume della goccia con conseguente aumento della concentrazione della proteina.

Tecniche di cristallizzazione

**- Diffusione di vapore
- con goccia seduta**



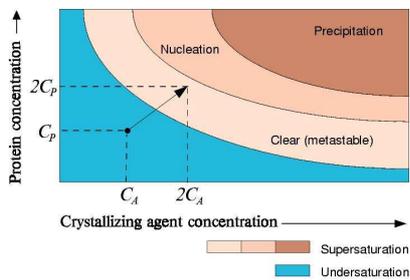
**- Diffusione di vapore
con goccia a sandwich**



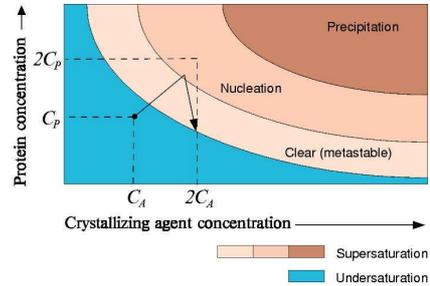
Tecniche di cristallizzazione

Che cosa avviene durante un esperimento di diffusione di vapore?

- Nessun cristallo

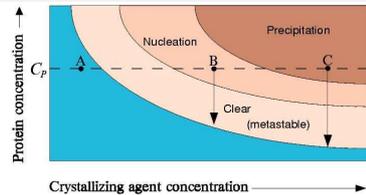
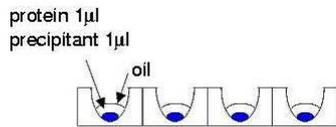


- Crescita cristalli

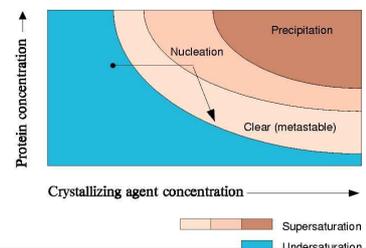
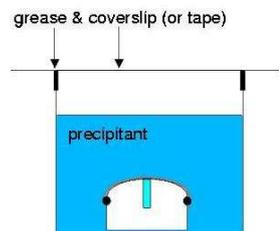


Tecniche di cristallizzazione

- **Batch crystallization:** la proteina ed il precipitante vengono mescolati insieme in olio per evitare evaporazione



- **Dialisi:** la soluzione proteica viene equilibrata con un volume maggiore di precipitante attraverso una membrana.



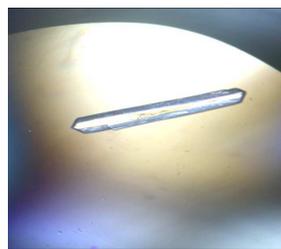
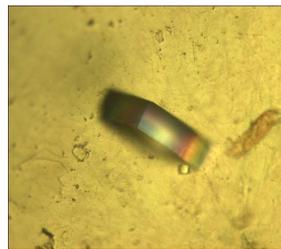
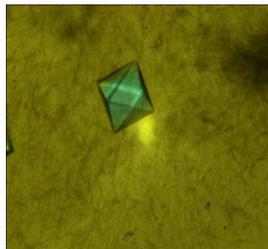
Parametri che influenzano il processo di cristallizzazione

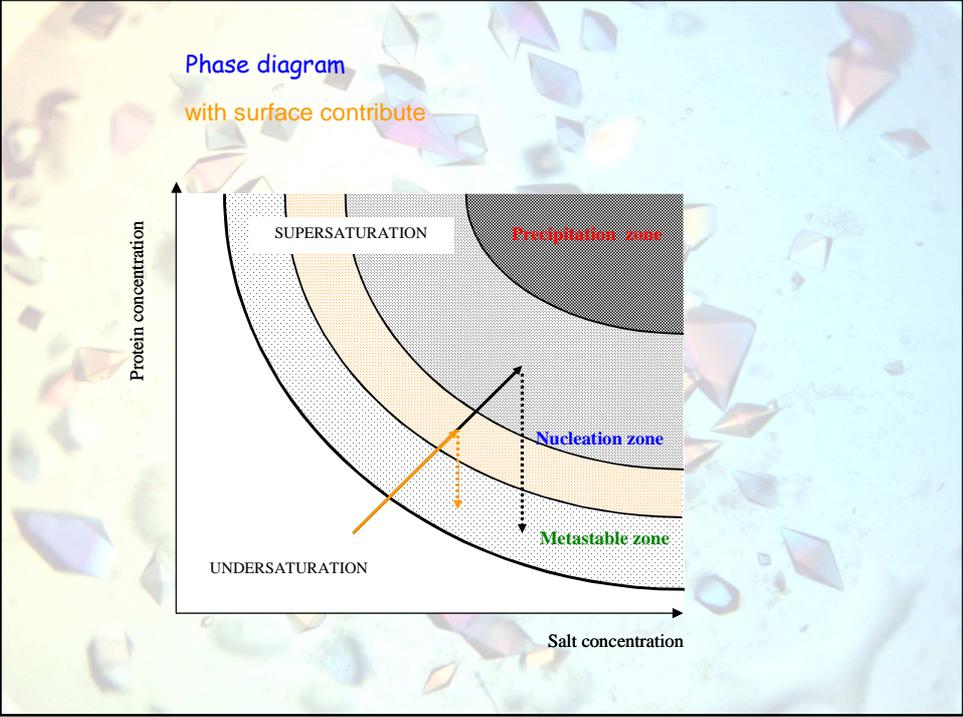
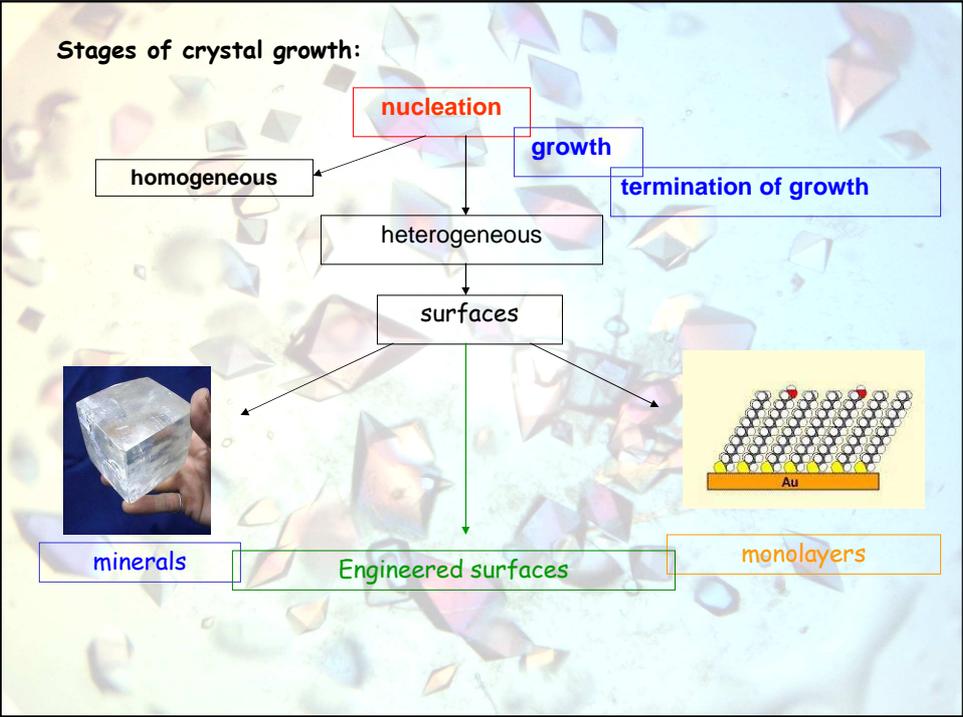
- 1- pH and buffer
- 2- Ionic strength
- 3- Temperature and temperature fluctuations
- 4- Concentration and nature of precipitant
- 5- Concentration of macromolecule
- 6- Purity of macromolecules
- 7- Additives, effectors and ligands
- 8- Organism source of macromolecule
- 9- Substrates, coenzymes, inhibitors
- 10- Reducing or oxidizing environment
- 11- Metal and other specific ions
- 12- Equilibration and growth rate
- 13- Surfactants or detergents
- 14- Gravity convection and sedimentation
- 15- Vibrations and sound
- 16- Volume of crystallization sample
- 17- Presence of amorphous or particulate sample
- 18- Surfaces of crystallization vessels
- 19- Proteolysis
- 20- Contamination by microbes
- 21- Pressure
- 22- Electric and magnetic fields
- 23- Handling by investigator and cleanliness
- 24- Viscosity of mother liquor
- 25- Heterogeneous or epitaxial nucleating agents

Xtal screening 1 e 2

Tube #	SALT	BUFFER	Precipitant
1	0.02M Calcium Chloride	0.1M Na Acetate pH 4.6	30% w/v 2-methyl-2,4-pentanediol
2	None	None	0.4M K ₂ Na Tartrate tetrahydrate
3	None	None	0.4M Ammonium dihydrogen phosphate
4	None	0.1M Tris-HCl pH 8.5	2.0M Ammonium Sulfate
5	0.2M tri-sodium citrate	0.1M Na HEPES pH 7.5	30% w/v 2-methyl-2,4-pentanediol
6	0.2M Magnesium chloride	0.1M Tris-HCl pH 8.5	30% w/v PEG 4000
7	None	0.1M Na Cacodylate pH 6.5	1.4M Sodium acetate trihydrate
8	0.2M tri-sodium citrate	0.1M Na Cacodylate pH 6.5	30% v/v 2-propanol
9	0.2M Ammonium acetate	0.1M Na Citrate pH 5.6	30% w/v PEG 4000
10	0.2M Ammonium acetate	0.1M Na Acetate pH 4.6	30% w/v PEG 4000
11	None	0.1M Na Citrate pH 5.6	1.0M Ammonium dihydrogen phosphate
12	0.2M Magnesium chloride	0.1M Na HEPES pH 7.5	30% v/v 2-propanol
13	0.2M tri-sodium citrate	0.1M Tris-HCl pH 8.5	30% v/v PEG 400
14	0.2M Calcium Chloride	0.1M Na HEPES pH 7.5	20% v/v PEG 400
15	0.2M Ammonium acetate	0.1M Na Cacodylate pH 6.5	30% w/v PEG 8000
16	None	0.1M Na HEPES pH 7.5	1.5M Lithium sulfate monohydrate
17	0.2M Lithium sulfate	0.1M Tris-HCl pH 8.5	30% w/v PEG 4000
18	0.2M Magnesium acetate	0.1M Na Cacodylate pH 6.5	20% w/v PEG 8000
19	0.2M Ammonium acetate	0.1M Tris-HCl pH 8.5	30% v/v 2-propanol
20	0.2M Ammonium sulfate	0.1M Na Acetate pH 4.6	25% w/v PEG 4000
21	0.2M Magnesium acetate	0.1M Na Cacodylate pH 6.5	30% w/v 2-methyl-2,4-pentanediol
22	0.2M Sodium acetate	0.1M Tris-HCl pH 8.5	30% w/v PEG 4000
23	0.2M Magnesium chloride	0.1M Na HEPES pH 7.5	30% v/v PEG 400
24	0.2M Calcium Chloride	0.1M Na Acetate pH 4.6	20% v/v 2-propanol
25	None	0.1M Imidazole pH 6.5	1.0M Sodium acetate trihydrate
26	0.2M Ammonium acetate	0.1M Na Citrate pH 5.6	30% v/v 2-methyl-2,4-pentanediol
27	0.2M tri-sodium citrate	0.1M Na HEPES pH 7.5	20% v/v 2-propanol
28	0.2M Sodium acetate	0.1M Na Cacodylate pH 6.5	30% w/v PEG 8000
29	None	0.1M Na HEPES pH 7.5	0.8M K ₂ Na Tartrate
30	0.2M Ammonium acetate	None	30% w/v PEG 8000
31	0.2M Ammonium acetate	None	30% w/v PEG 4000
32	None	None	2.0M Ammonium Sulfate
33	None	None	4.0M Sodium formate
34	None	0.1M Na Acetate pH 4.6	2.0M Sodium formate
35	None	0.1M Na HEPES pH 7.5	1.6M K ₂ Na Tartrate
36	None	0.1M Tris-HCl pH 8.5	8% w/v PEG 8000
37	None	0.1M Na Acetate pH 4.6	8% w/v PEG 4000
38	None	0.1M Na HEPES pH 7.5	1.8M Sodium citrate dihydrate
39	None	0.1M Na HEPES pH 7.5	2% v/v PEG 400 & 2.0M Ammonium Sulfate
40	None	0.1M Na Citrate pH 5.6	20% v/v 2-propanol & 20% w/v PEG 4000
41	None	0.1M Na HEPES pH 7.5	10% v/v 2-propanol & 20% w/v PEG 4000
42	0.05 K-phosphate monobasic	None	20% w/v PEG 8000
43	None	None	30% w/v PEG 1500
44	None	None	0.2M Magnesium formate
45	0.2M Zinc acetate	0.1M Na Cacodylate pH 6.5	18% w/v PEG 8000
46	0.2M Calcium acetate	0.1M Na Cacodylate pH 6.5	18% w/v PEG 8000
47	None	0.1M Na Acetate pH 4.6	2.0M Ammonium Sulfate
48	None	0.1M Tris-HCl pH 8.5	2.0M Ammonium dihydrogen phosphate
49	1.0M Lithium Sulfate	None	2% w/v PEG 8000
50	0.5M Lithium Sulfate	None	15% w/v PEG 8000

Alcuni cristalli proteici di varie morfologie.....





Flow-chart della determinazione strutturale di una proteina



Raccolta dati

Sincrotroni: sorgenti di radiazioni ad ampio spettro



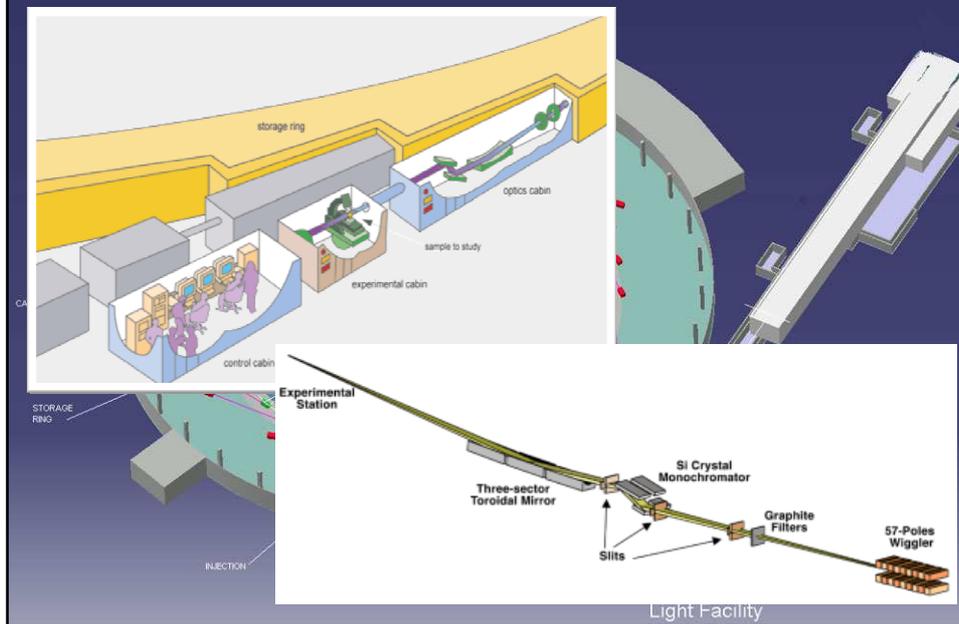
Le sorgenti convenzionali di raggi X che usano la radiazione $\text{CuK}\alpha$ ($\lambda = 1.54056 \text{ \AA}$) sono progressivamente sostituite da sorgenti di luce di sincrotrone (acceleratori di elettroni).



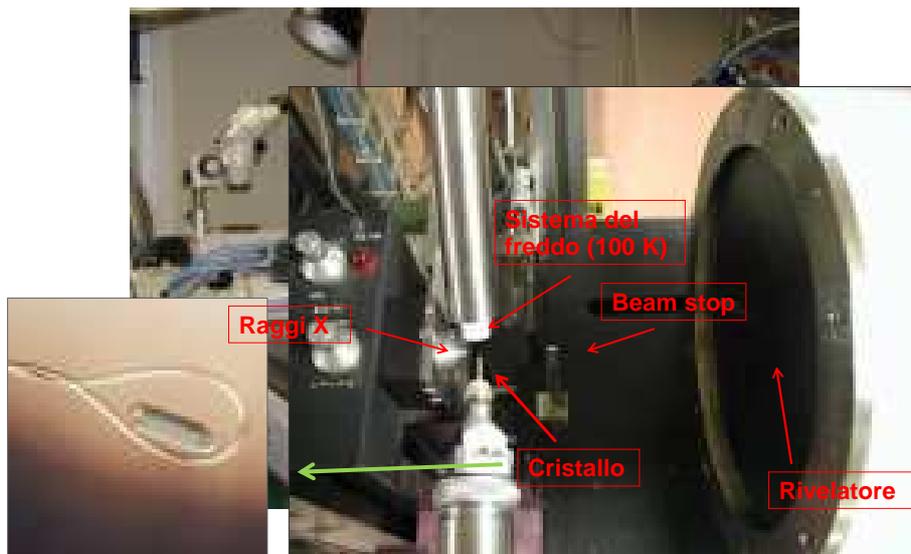
1995: 3rd generation synchrotron. ESRF (European Synchrotron Radiation Facility), Grenoble (France). Diameter 844 m

- Used to solve protein structure
- Large central service units
- Use magnets that bend the electrons
- Wavelength range, typically 0.5-3.0 \AA
- Faster measurement due to up to 1000 fold higher energy

Schema di una sorgente di luce di sincrotrone



Postazione di misura in una linea di diffrazione di raggi X



- La raccolta dei dati viene comunemente eseguita a 100 K

Per diminuire la mobilità dei radicali formati dall'irraggiamento dei raggi X. Questi radicali provocherebbero l'innesco di reazioni che danneggerebbero le molecole stesse.
Danni comunemente osservati sono ad es. la decarbossilazione di glutammati e aspartati.

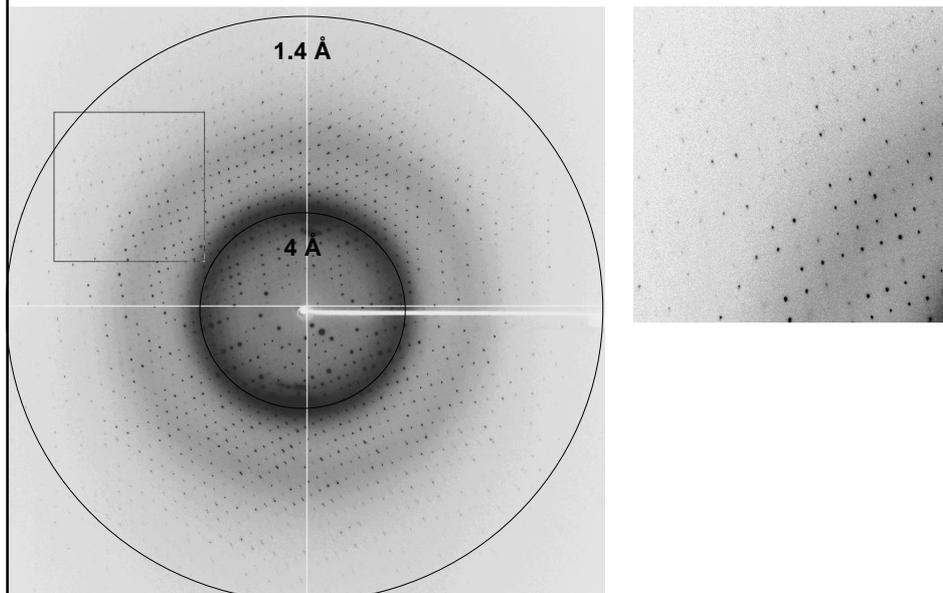
- Come viene trattato e preparato il cristallo prima della raccolta

- Il cristallo deve essere protetto dalla formazione di ghiaccio che per aumento di volume, ne provocherebbe la rottura.

-Il cristallo viene "pescato" ed immerso in una soluzione formata dalla soluzione di cristallizzazione più un *crio-protettore* spesso costituito da glicerolo, PEG a basso peso molecolare (200-400), saccarosio in percentuale dal 10 al 40%.

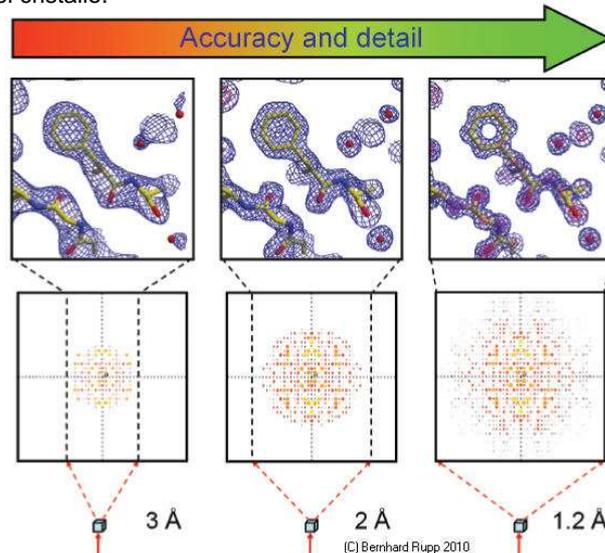
-Il cristallo viene di nuovo pescato e rapidamente posto sotto il fascio di N_2 a 100 K.

Immagine di diffrazione per una macromolecola biologica



La qualità del cristallo e dei dati di diffrazione determina l'accuratezza della struttura 3D

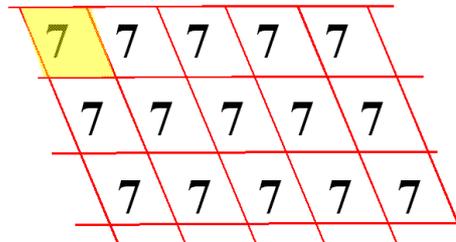
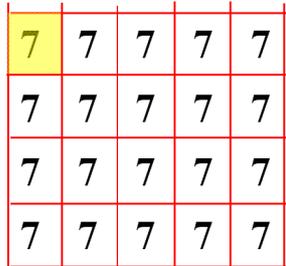
La **risoluzione** (espressa di solito in $\text{\AA} = 0,1 \text{ nm} = 10^{-8} \text{ cm}$) corrisponde al massimo angolo di diffrazione a cui si osservano i riflessi, quindi alla più piccola distanza distinguibile nel cristallo.



Cosa otteniamo dai dati di diffrazione?

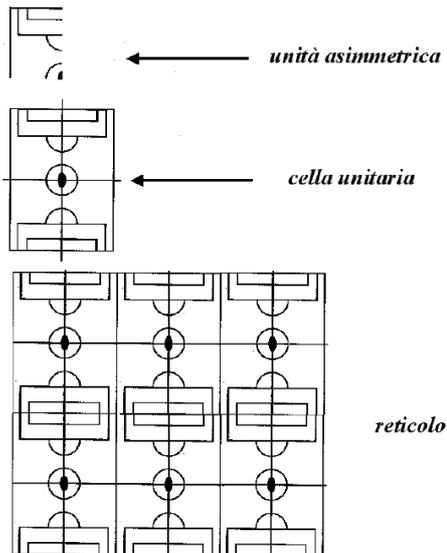
- Indici dei riflessi (h,k,l)
- Cella unitaria
- Gruppo spaziale
- Intensità dei riflessi $I(h, k, l)$

La **cella unitaria** è l'unità minima che si ripete traslazionalmente in un reticolo cristallino



La simmetria nella cella unitaria è definita rispetto a precisi punti ed operazioni. L'insieme delle operazioni di simmetria appartenenti ad una cella è un *gruppo puntuale*. Sono possibili al massimo 32 gruppi puntuali. La combinazione dei gruppi puntuali con le operazioni traslazionali sia semplici che complesse dà i *gruppi spaziali*, cioè tutte le possibili disposizioni in uno spazio a tre dimensioni di oggetti tridimensionali. I gruppi spaziali sono 230.

L'**unità asimmetrica** è l'unità minima che è in grado di generare l'intero reticolo (tramite le operazioni di simmetria e le traslazioni)



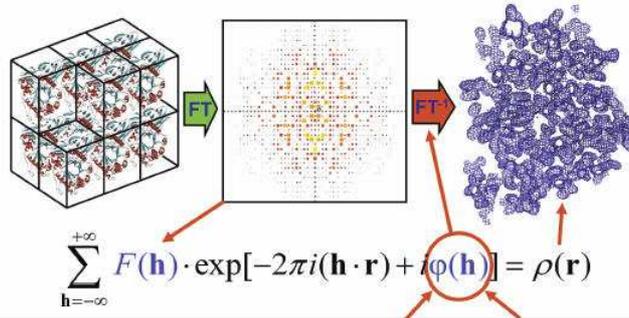
Che cos'è il "problema della fase" in cristallografia?

In un **esperimento di diffrazione** gli atomi disposti in posizioni ordinate diffrangono i raggi X in direzioni ben definite. Registrando l'immagine di diffrazione tramite un rivelatore possiamo misurare l'intensità $I(h,k,l)$ di ciascun riflesso.

Per ricostruire la densità elettronica, quindi la posizione degli atomi e la forma della molecola che diffrangono i raggi X, è necessario applicare all'immagine di diffrazione l'inverso della trasformata di Fourier.

Per ogni riflesso è necessario conoscere 2 componenti:

- $F(h)$ **ampiezza**, è misurabile ed è proporzionale al quadrato dell'intensità
- $\phi(h)$ fase, la sua determinazione necessita ulteriori esperimenti



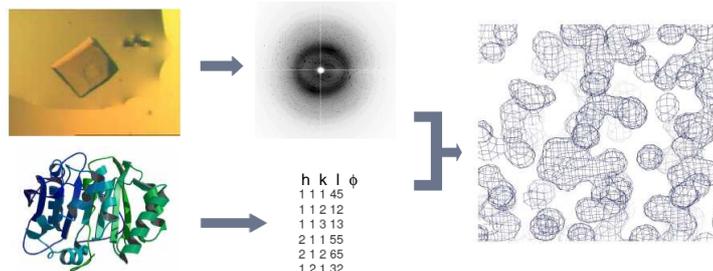
METODO DELLA SOSTITUZIONE MOLECOLARE (Molecular replacement)

Può essere utilizzato quando si dispone di un buon modello per una frazione ragionevole della molecola.
Il livello di similitudine tra due proteine è correlato al grado di identità di sequenza.

➡ **Come regola generale il MR funziona se l'identità di sequenza tra la proteina da determinare e il modello scelto è almeno del 40%**

Per eseguire MR bisogna posizionare la struttura modello nella cella unitaria determinata sperimentalmente. Per orientare correttamente la molecola è necessario determinare 3 angoli di rotazione e 3 valori per la traslazione.

Se nell'unità asimmetrica c'è una sola molecola, il MR è un problema a 6 dimensioni, il quale può essere diviso in due problemi a 3 dimensioni.



Metodi comuni per ottenere fasi sperimentali iniziali

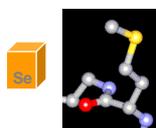


Multiple Isomorphous Replacement (MIR)

- È necessario raccogliere i dati da un cristallo nativo e da due cristalli immersi in una soluzione di un atomo pesante (Hg, Pt, Au, Cd). Limite importante: i 3 cristalli devono essere isomorfi.
- È un metodo molto buono e potente per ottenere fasi sperimentali anche a bassa risoluzione. Si può eseguire con sorgenti convenzionali ($\text{CuK}\alpha$)

Single Anomalous Diffraction (SAD)

- Il cristallo deve contenere un elemento adatto che dia scattering anomalo. Il metodo è un buon metodo per ottenere fasi ma deve essere combinato con altre tecniche per migliorare la densità elettronica ottenuta. Molto comune utilizzare lo zolfo nativo come elemento di scattering anomalo.



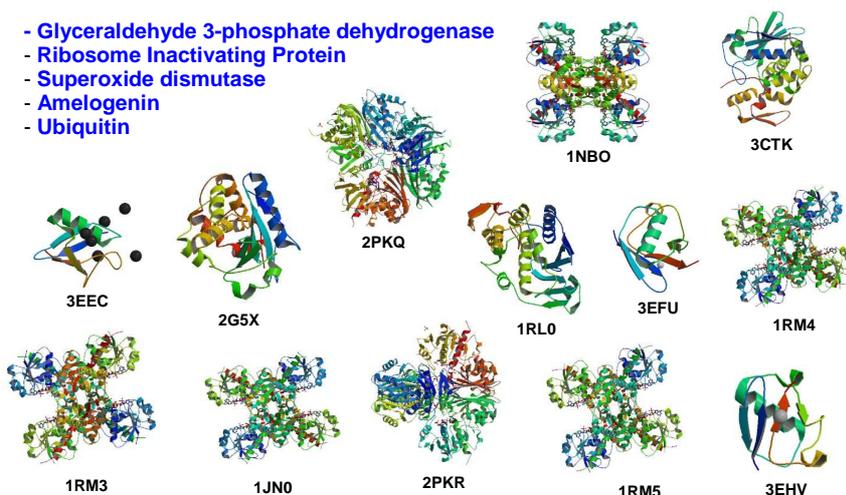
Multiple Anomalous Dispersion (MAD)

- Deve essere presente un elemento di scattering anomalo e devono essere raccolti vari set di dati a diverse lunghezze d'onda. È necessario utilizzare il sincrotrone come sorgente e dati con un basso errore. Si ottengono fasi molto buone utilizzando cristalli con Se-Met.



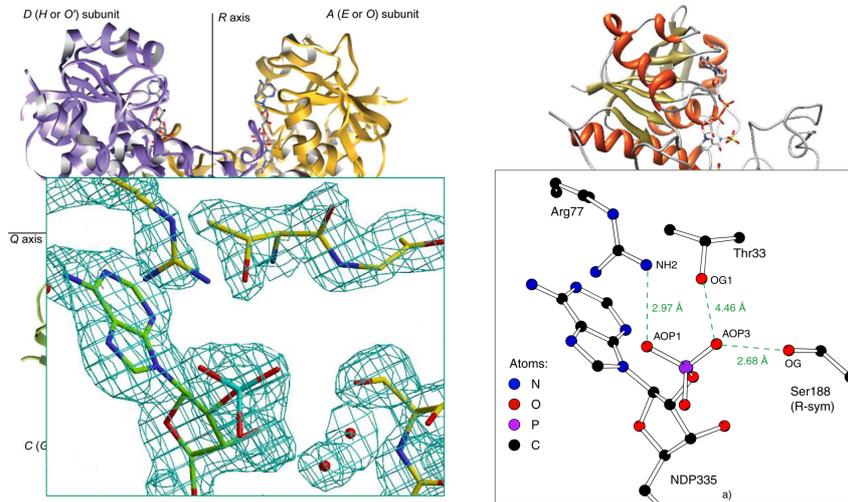
Crystal Structures Solved at the Laboratory of Environmental and Biological Structural Chemistry

- Glyceraldehyde 3-phosphate dehydrogenase
- Ribosome Inactivating Protein
- Superoxide dismutase
- Amelogenin
- Ubiquitin

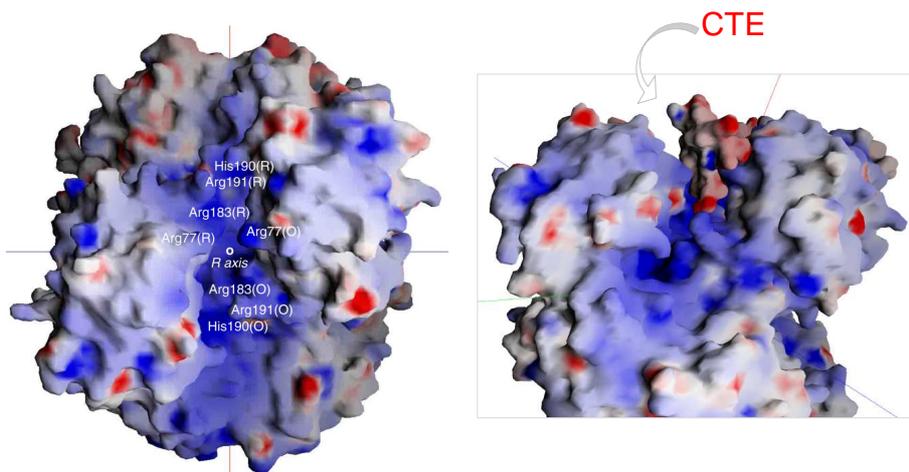


Le strutture proteiche risolte vengono depositate nel PDB (Protein Data Bank)

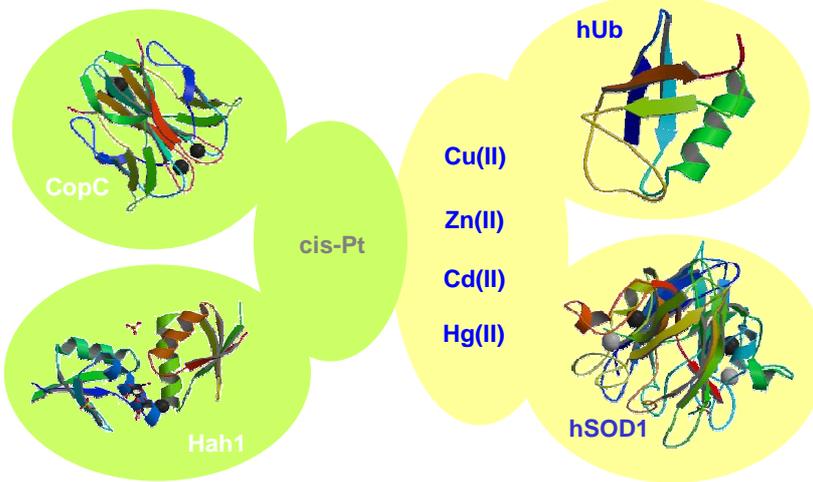
Gliceraldeide-3-fosfato deidrogenasi (GAPDH) fotosintetica



Potenziale elettrostatico superficiale della GAPDH fotosintetica



Interazioni di ioni metallici con proteine



Ubiquitina umana: interazione con i metalli

