

Modulo 3

Lo studio delle proteine

2022-23

Come isolare e purificare le proteine

► Fattori determinanti - QUALITÀ, QUANTITÀ, COSTO

- la rispettiva rilevanza dipende dal contesto: ricerca di base, R&D o produzione
- le proteine si possono ottenere o dal sito di produzione endogena (es. una cellula) o mediante tecniche ricombinanti.
- devono poi essere separate da tutti gli altri componenti cellulari

► La lisi cellulare libera le proteine

Metodi fisici

- omogenizzazione meccanica
- ultrasonicazione
- estrusione (pressa di French)
- cavitazione (shock osmotico)

Metodi chimici/biochimici

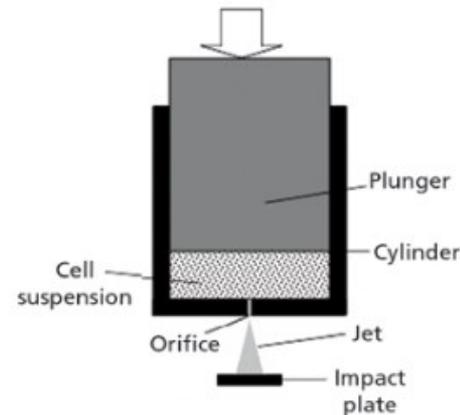
- solventi (es. toluene)
- detergenti (es. triton-X)
- enzimi (es. lipasi, lisozima)



omogenizzatore meccanico



sonicatore



pressa di French



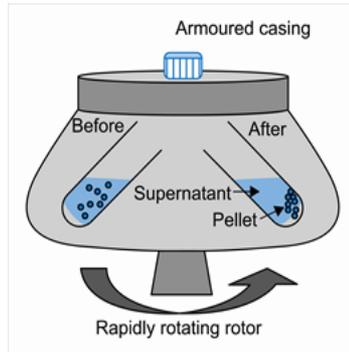
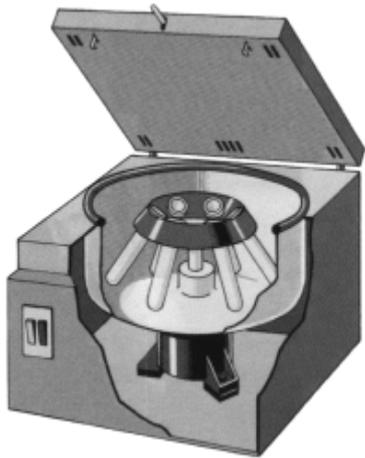
bomba osmotica

Raccolta del materiale proteico

► La lisi cellulare libera tutto il materiale presente nel citosol

- si liberano proteine, piccole biomolecole, sali, polisaccaridi, acidi nucleici, ecc.
- se si desidera isolare proteine da **altri compartimenti** (es. di membrana, dal ER o mitocondri, è necessario seguire **protocolli specifici** (es. centrifugazione differenziale a diverse velocità)
- è necessario **inattivare enzimi degradativi** (es. lavorando a freddo e aggiungendo alla sospensione inibitori per gli enzimi proteolitici)
- si procede quindi a **separare le proteine** nella sospensione da detriti cellulari mediante, per esempio:

centrifugazione



dialisi

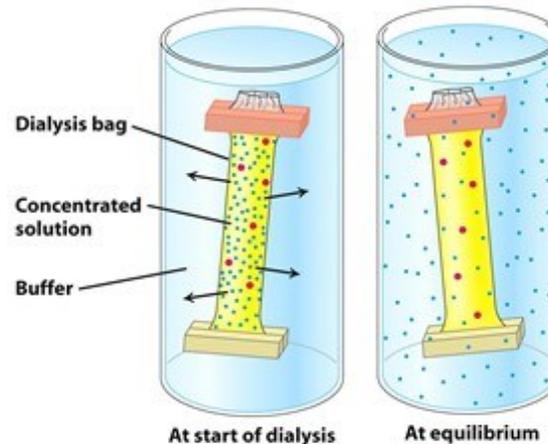


Figure 3.2
Biochemistry, Seventh Edition
© 2012 W. H. Freeman and Company

ultrafiltrazione

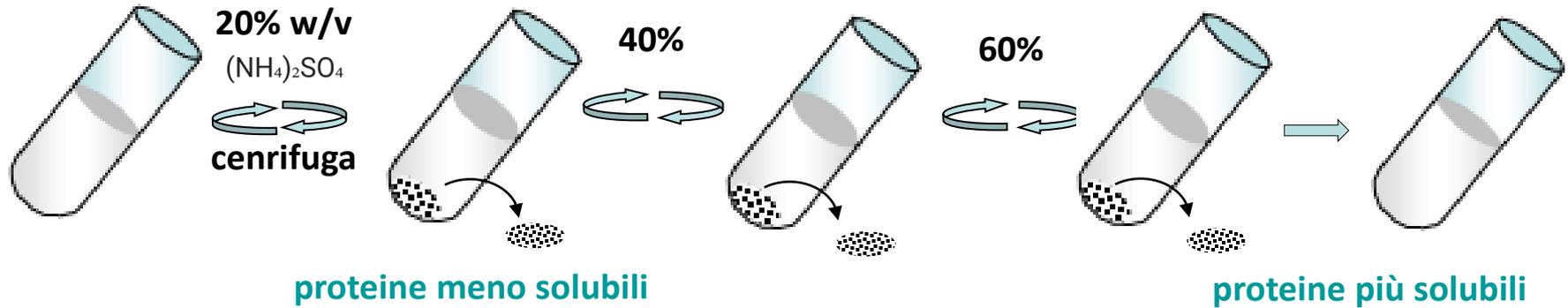


- si eliminano altri tipi di macromolecole (es. acidi nucleici) con enzimi specifici che li degradano (nucleasi)

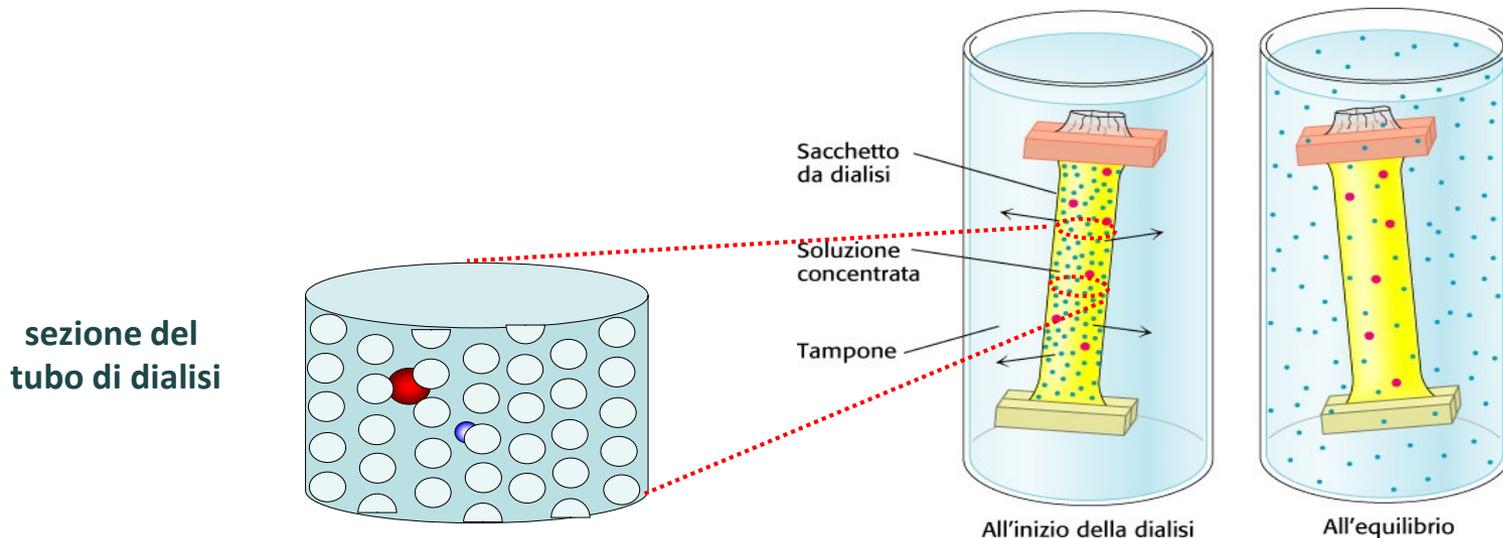
1° passaggio di purificazione: salting out

► Il metodo del **SALTING OUT** separa le proteine in base alla loro idrofilicità

- aggiunta di solfato di ammonio, $(\text{NH}_4)_2\text{SO}_4$, un sale altamente igroscopico



- utilizzo della dialisi per rimuovere l'eccesso di $(\text{NH}_4)_2\text{SO}_4$ e altre specie di piccole dimensioni



GEL ELETTROFORESI

► Una molecola carica in soluzione si muove in un campo elettrico

- la velocità di movimento dipende da vari fattori

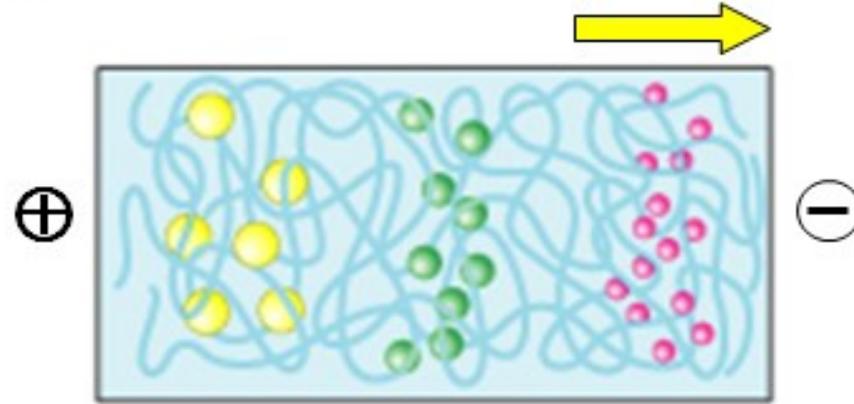
$$v = \frac{V \cdot q}{f \cdot d}$$

voltaggio
carica
coefficiente di frizione

distanza
elettrodi

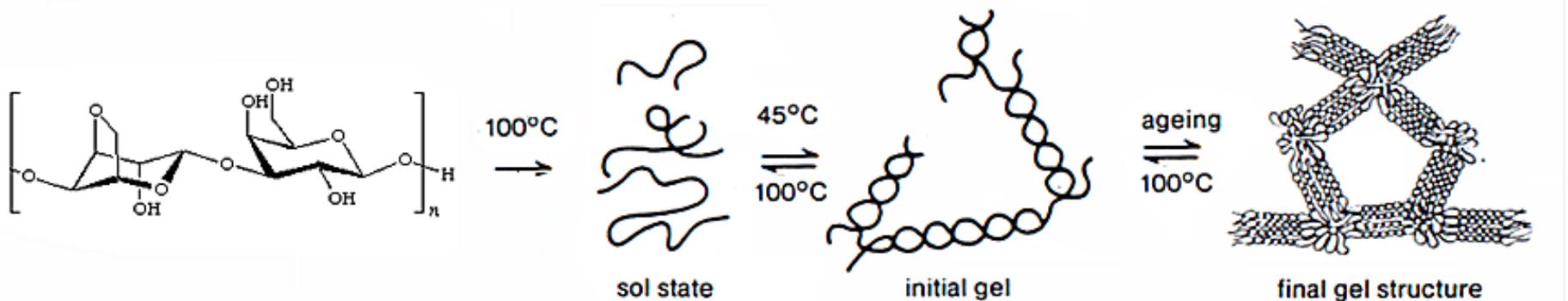
$$f \propto r \cdot \eta$$

grandezza
viscosità
ambiente



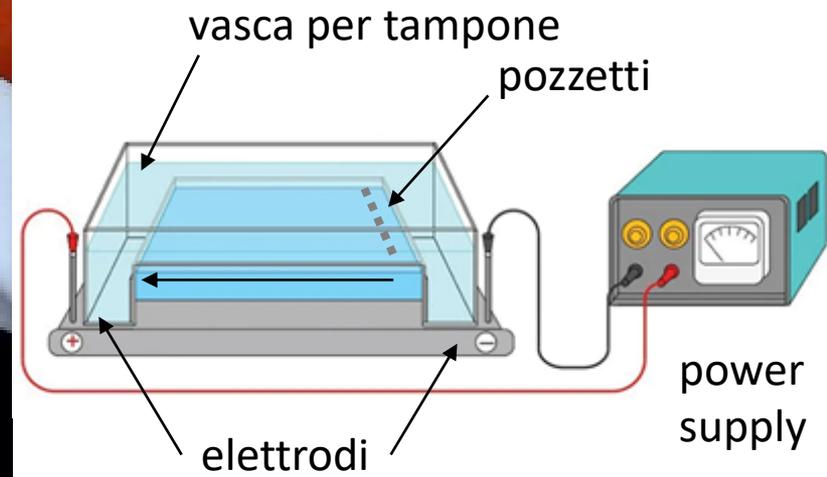
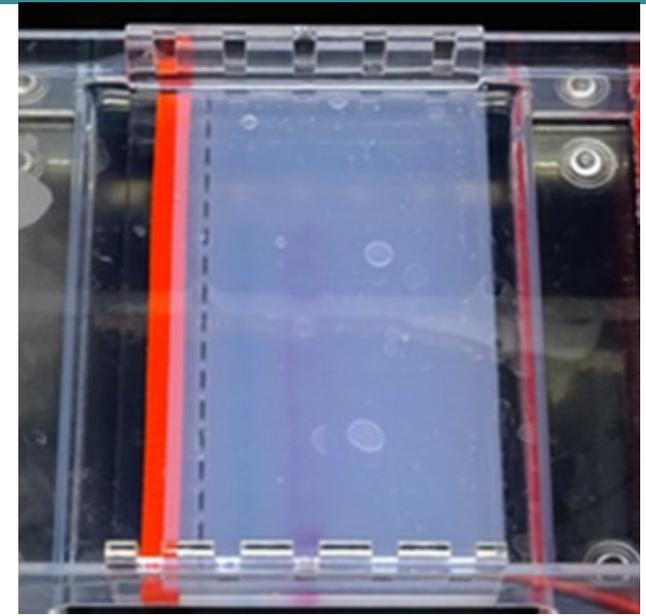
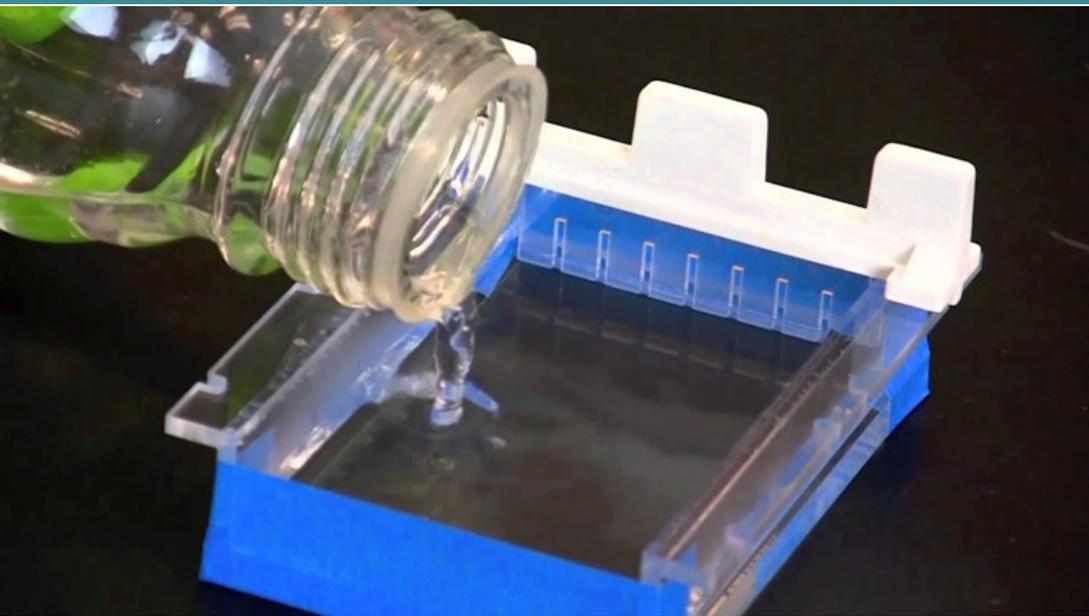
► Elettroforesi su gel di agarosio

- utilizza gel di agarosio, polimero di $[(1 \rightarrow 6) \alpha\text{-D-galattosio}-(1 \rightarrow 4)\beta\text{-3,6-anidrogalattosio}]_n$
- metodo utilizzato per separare DNA a doppio filamento (ds)



- maggiore la concentrazione d'agarosio, più fitta è la trama del gel

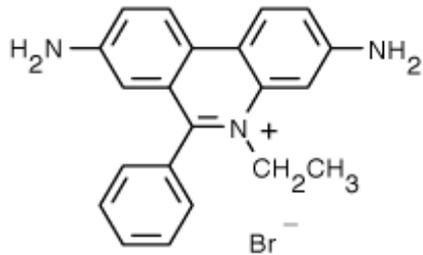
Gel elettroforesi all'agarosio



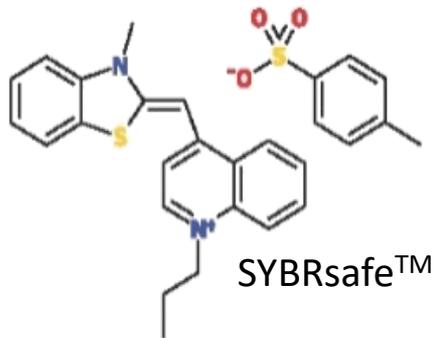
Gel elettroforesi all'agarosio: visualizzazione dei frammenti

► I frammenti di dsDNA sono visualizzati utilizzando una sonda fluorescente

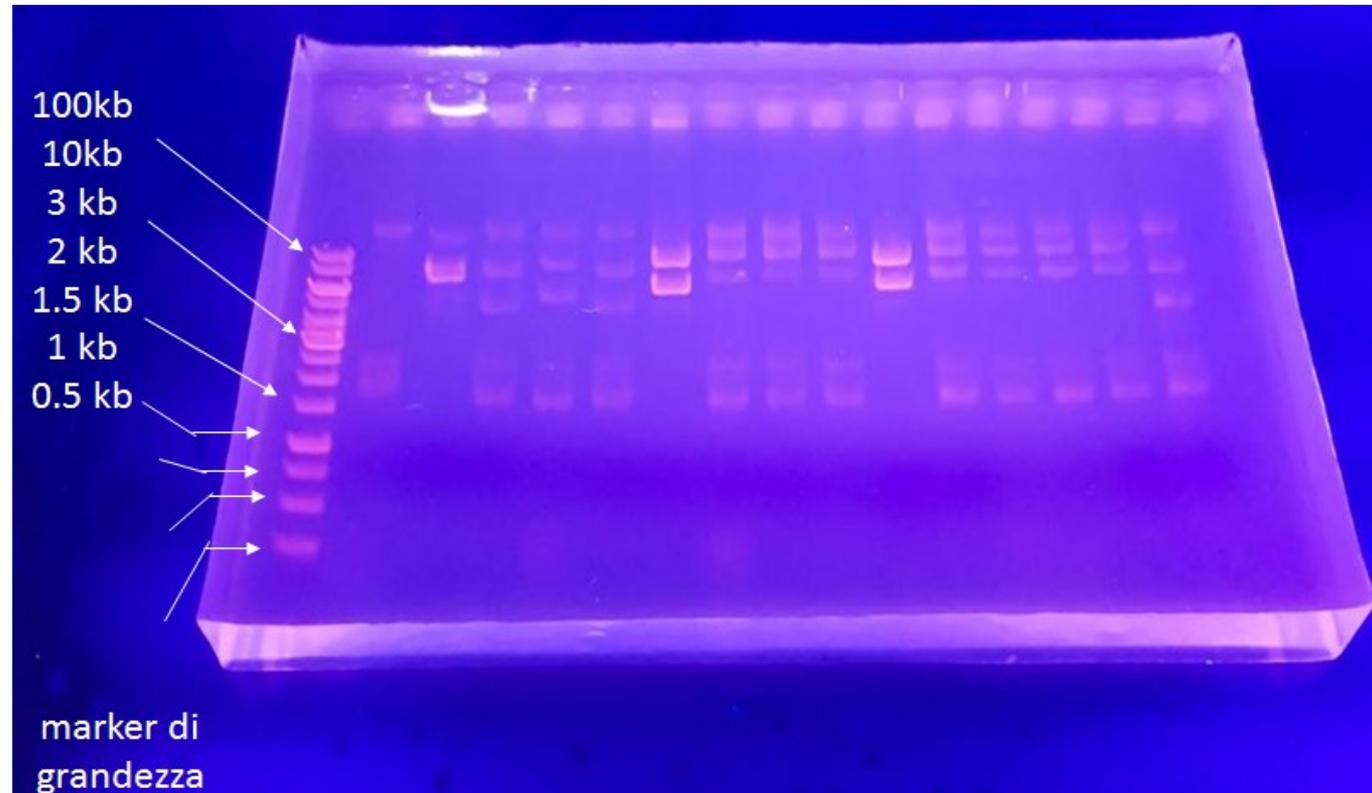
- originariamente si utilizzava Etidio Bromuro (ER), che **intercala** in dsDNA e diventa **fluorescente**
- i frammenti di DNA viaggiano con **velocità proporzionale alla grandezza** (corti = veloci)
- per il DNA la **grandezza** si misura in kilobasi (kb), determinata confrontando la **posizione** della banda con quella di **marker** a lunghezza nota (kit commerciali)
- **ER è tossico e teratogeno**, è stato rimpiazzato da altre sonde (es. SYBRsafe)



etidio bromuro



SYBRsafe™

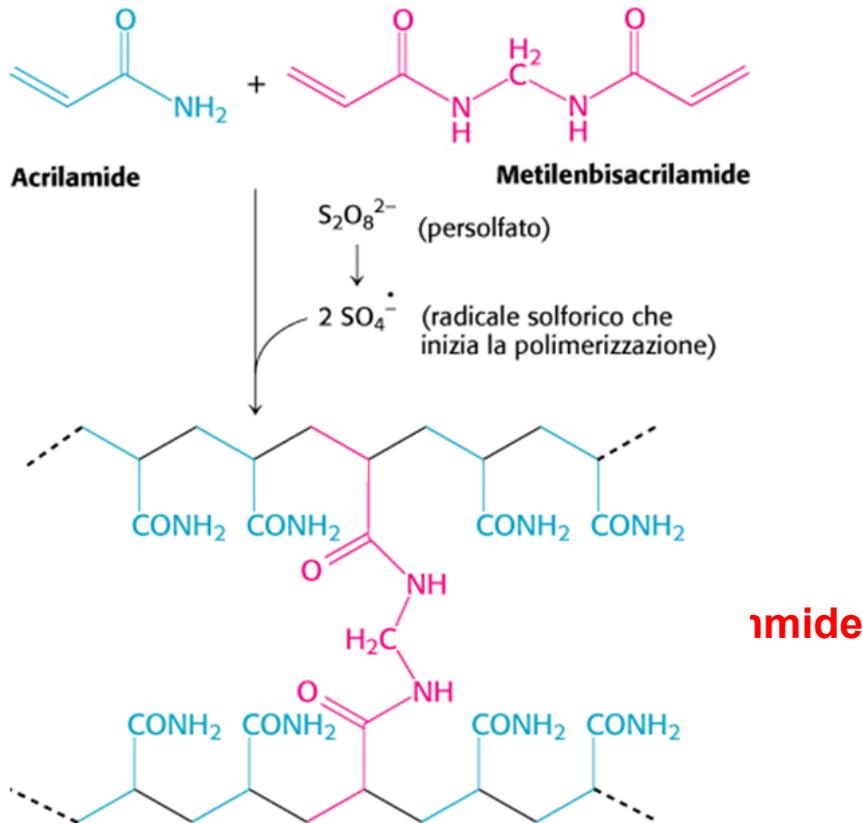


marker di grandezza

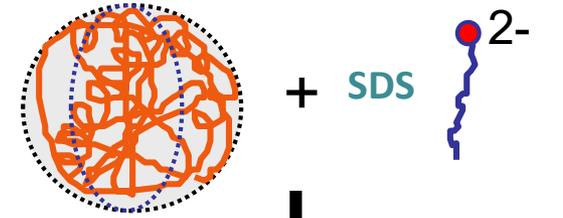
Elettroforesi al poliacrilammide (PAGE e SDS-PAGE)

► PAGE è l'acronimo di PolyAcrilamide Gel Electrophoresis

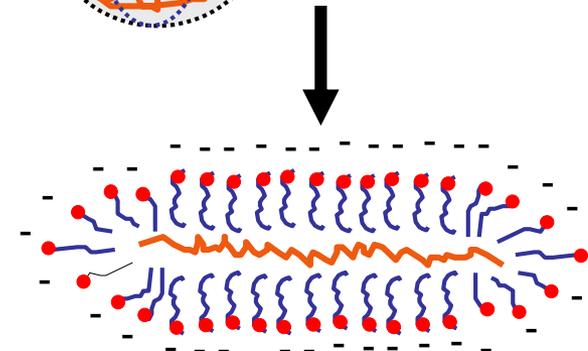
- questo metodo è utilizzato per la **separazione di oligonucleotidi a singolo filamento (ss)**
- La variante SDS-PAGE tiene conto che a differenza degli acidi nucleici, la **carica e forma** delle proteine **non è correlata alla lunghezza** della catena
- Utilizza il **Sodio Dodecil Solfato** per denaturarle ed inglobarle, omogeneizzando la carica alla lunghezza
- permette la visualizzazione dei frammenti con coloranti o mediante marker radioattivi



proteina nella forma nativa

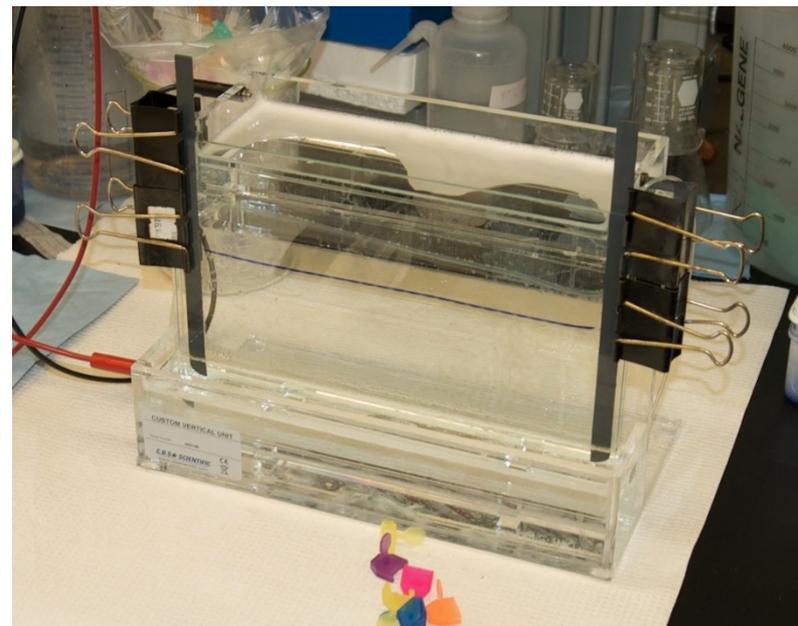
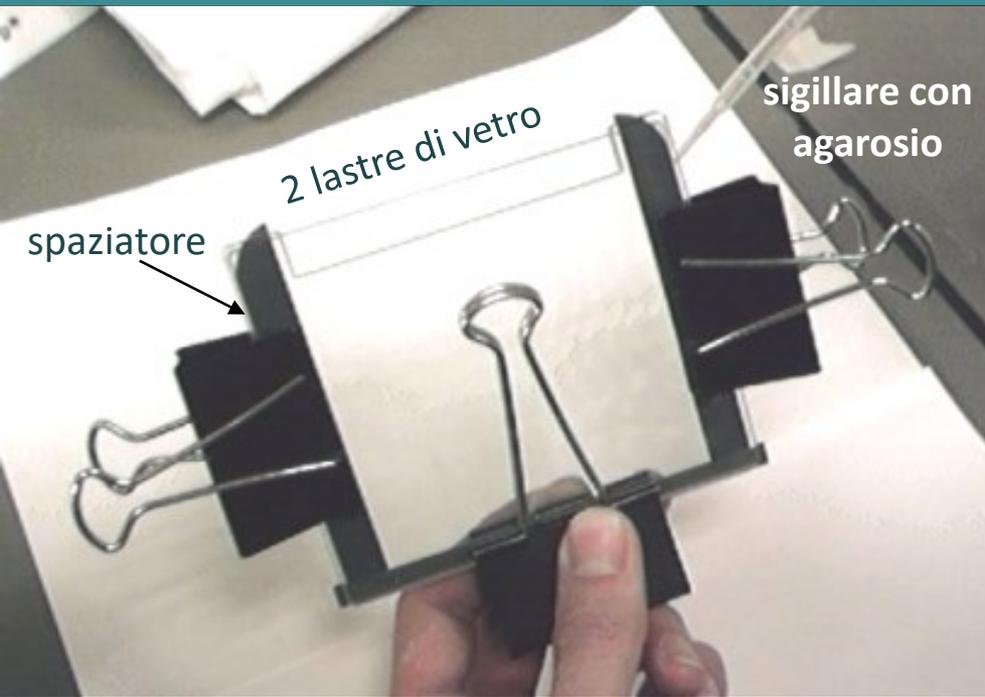


+ SDS

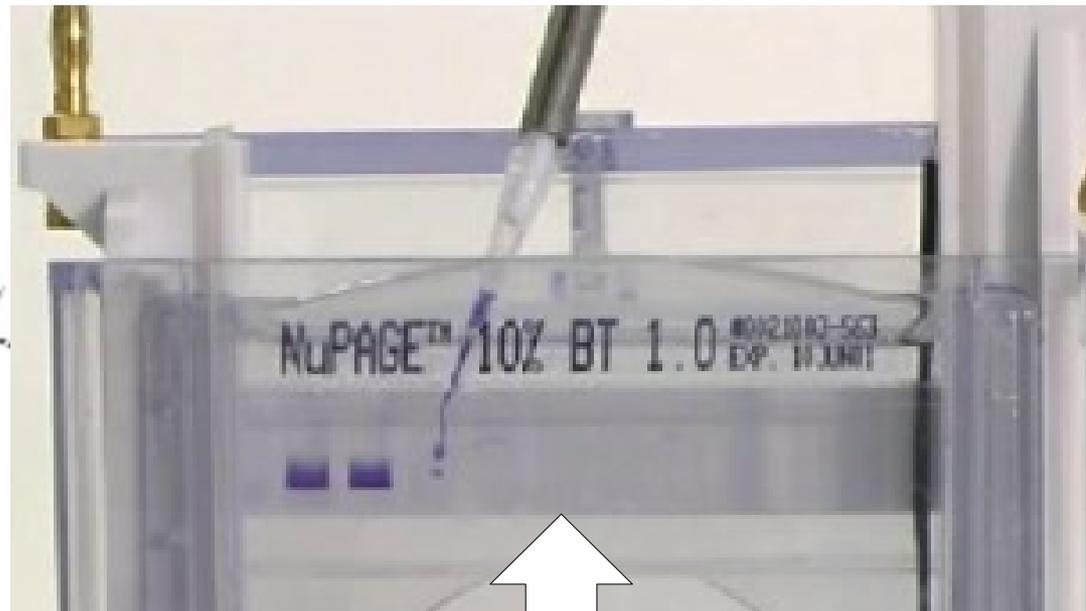
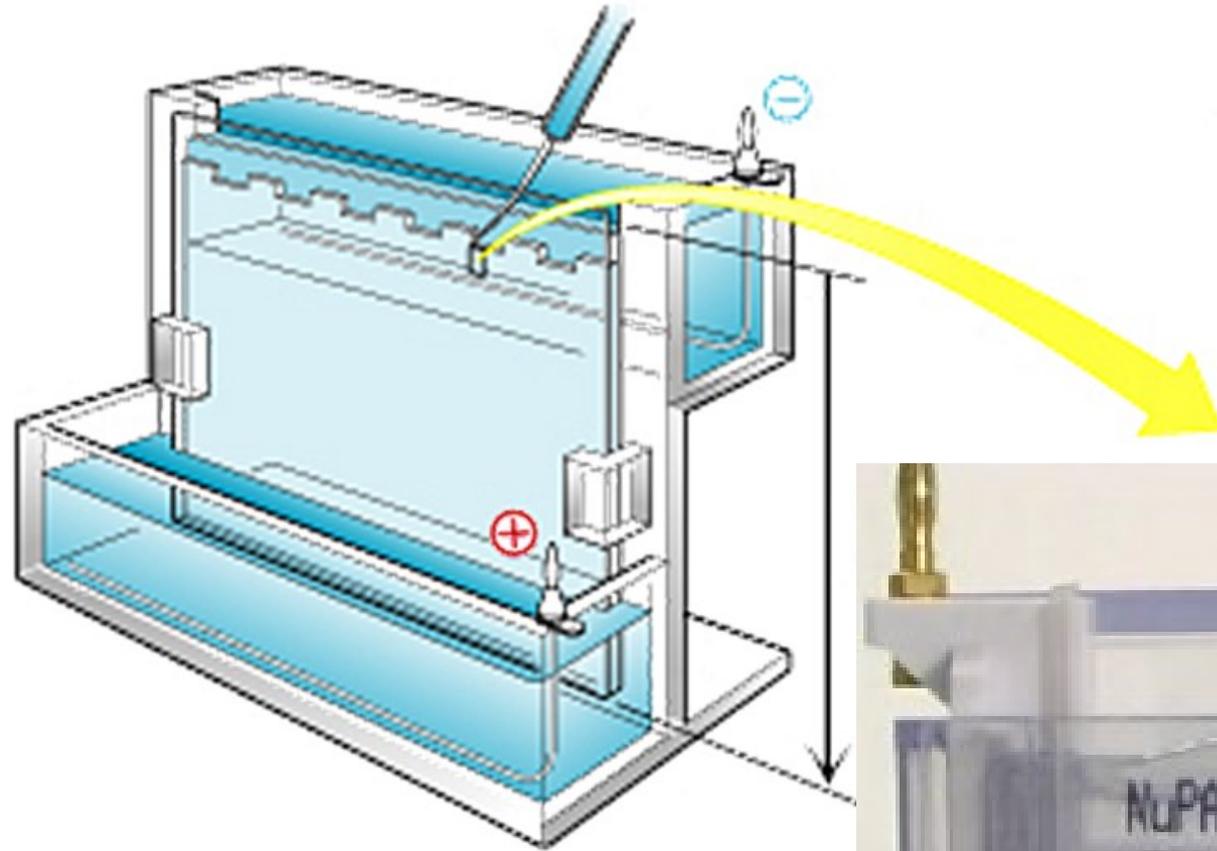


proteina denaturata incapsulata da molecole di SDS

SDS-PAGE – preparazione del gel



SDS-PAGE – caricamento dei campioni



Loading buffer:
Proteine da separare + blu di bromofenolo
(colorante) + glicerolo (10%, addensante)

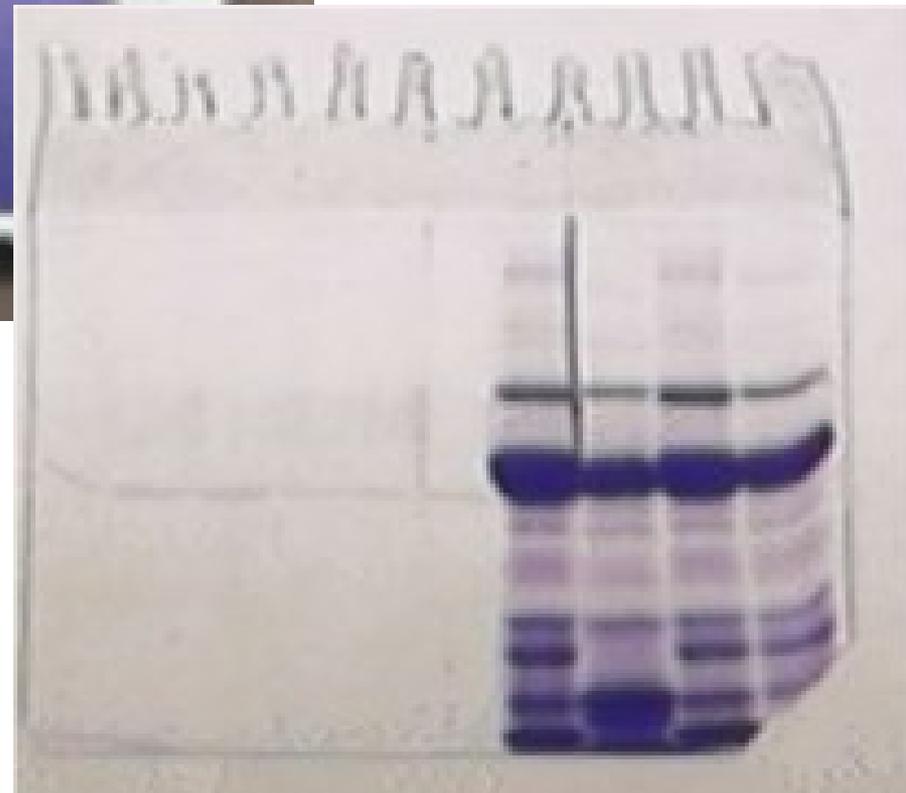
SDS-PAGE – sviluppo del gel



- 1) colorazione delle proteine con colorante
- 2) rimozione del colorante dal gel

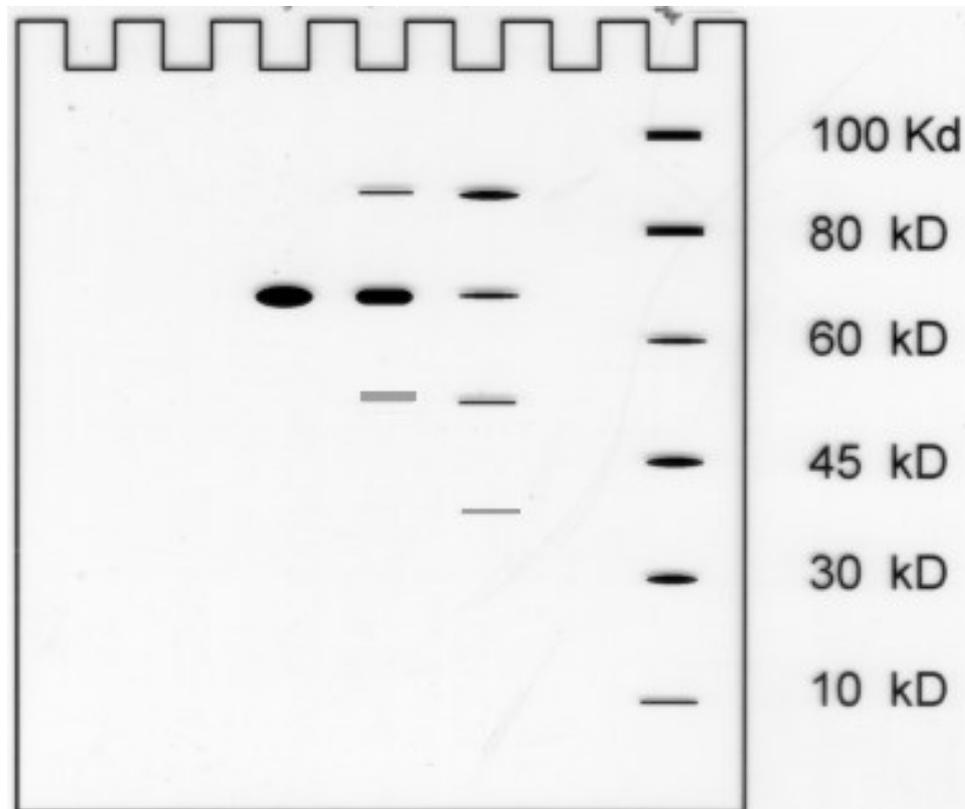
Coloranti:

- colorante coumassie blu
- colorante argentario
- colorante fluorescente



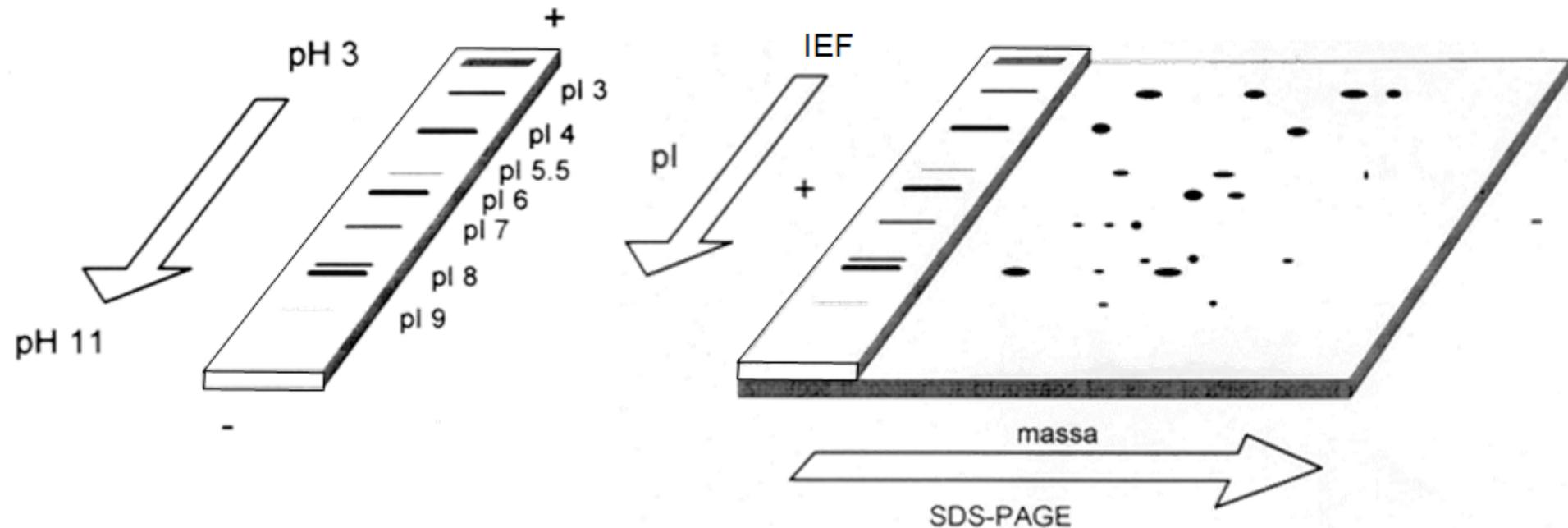
Gel Elettroforesi: a cosa serve?

- nell'elettroforesi **la mobilità** nel gel è generalmente **proporzionale alla grandezza**
- l'utilizzo di marker a grandezza nota **permette di stimare la massa** delle proteine che formano una banda (**approssimativamente**)
- Il gel fornisce indicazioni
 - 1) su quante diverse proteine sono presenti (**purezza**)
 - 2) sulla loro **abbondanza relativa** (intensità della banda)
- permette di **identificare la frazione** dove la proteina è più **pura e/o abbondante**



ELETTROFORESI BIDIMENSIONALE – focusing isoelettrico (IEF)

- ▶ **le proteine sono molecole anfoteriche (hanno residui sia acidi che basici)**
 - per ogni proteina c'è un valore di pH al quale il numero di cariche + equivale quelle -
 - questo è il **punto isoelettrico (pI)** al quale la carica netta della catena = 0
 - IEF utilizza una striscia di gel preparato con **molecole anfoteriche**, che formano un **gradiente di pH**; le proteine si fermano al pH che corrisponde al loro pI
 - la striscia di gel è posta in seguito su un gel SDS-PAGE, si ottiene così un **gel bidimensionale**
 - IEF associato a spettrometria di massa (permette di identificare le proteine in ogni banda in base al preciso peso molecolare) è di **fondamentale importanza per la proteomica**

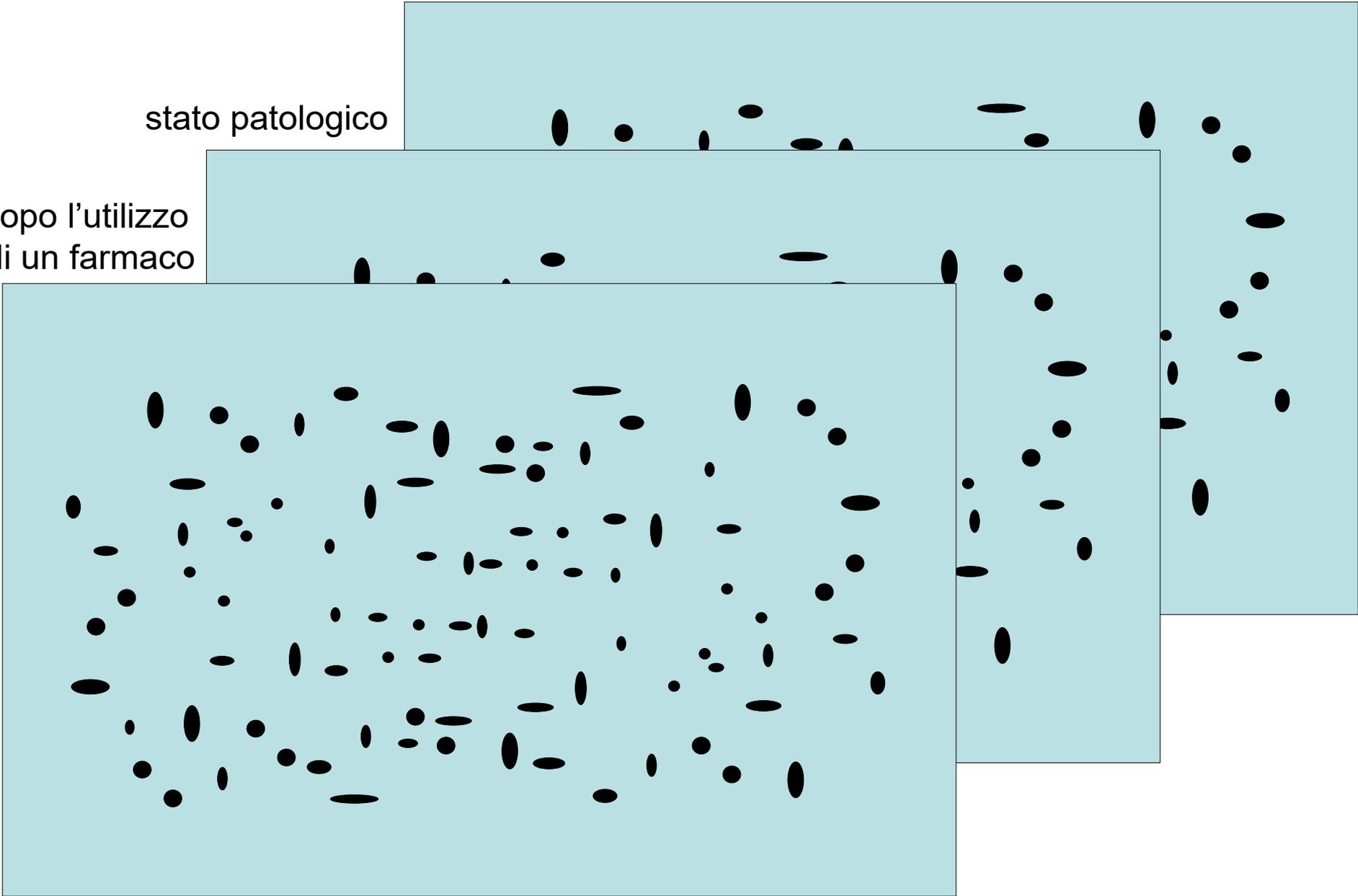


Elettroforesi bidimensionale – stato metabolico delle cellule

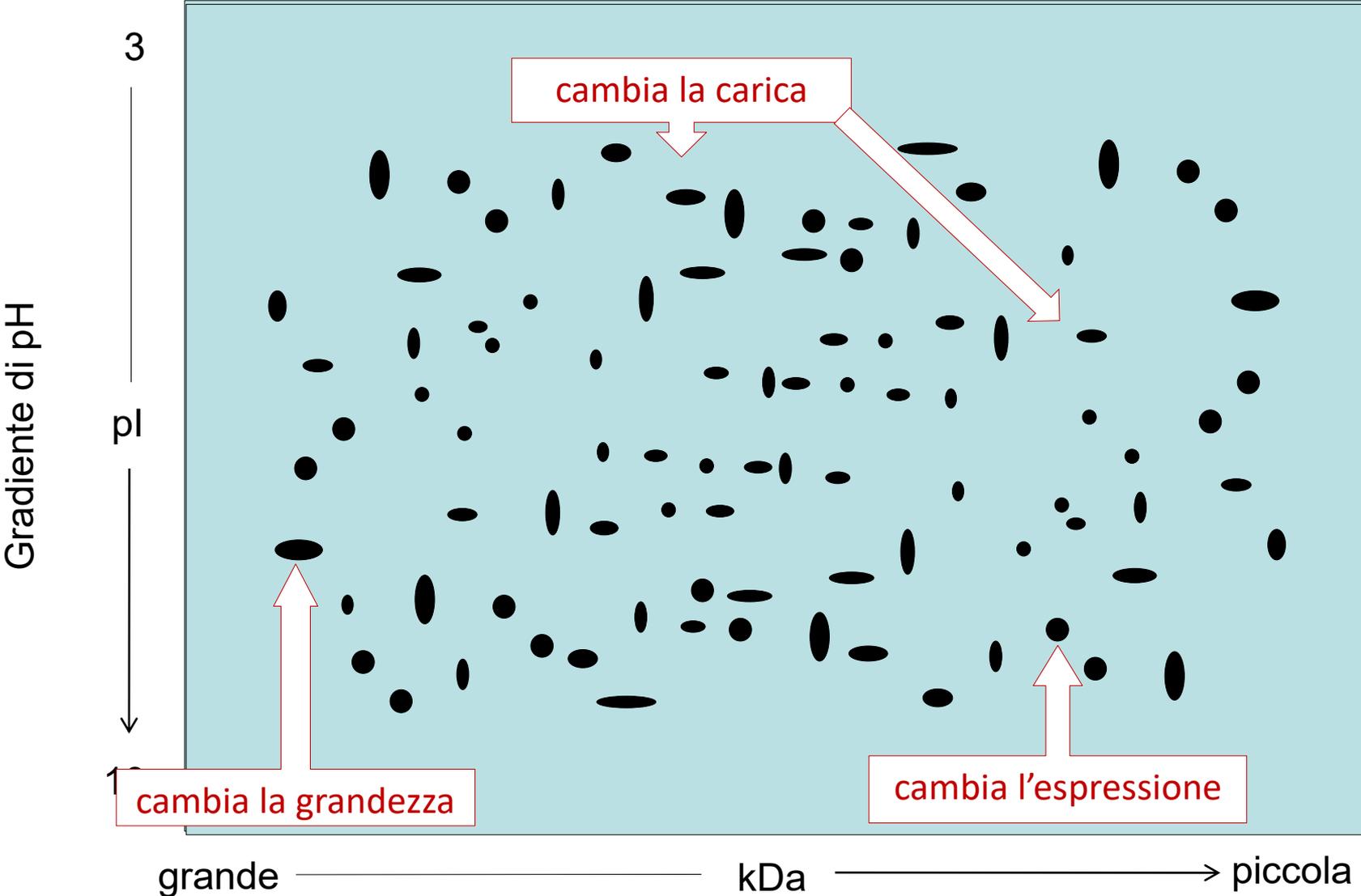
stato fisiologico

stato patologico

dopo l'utilizzo
di un farmaco

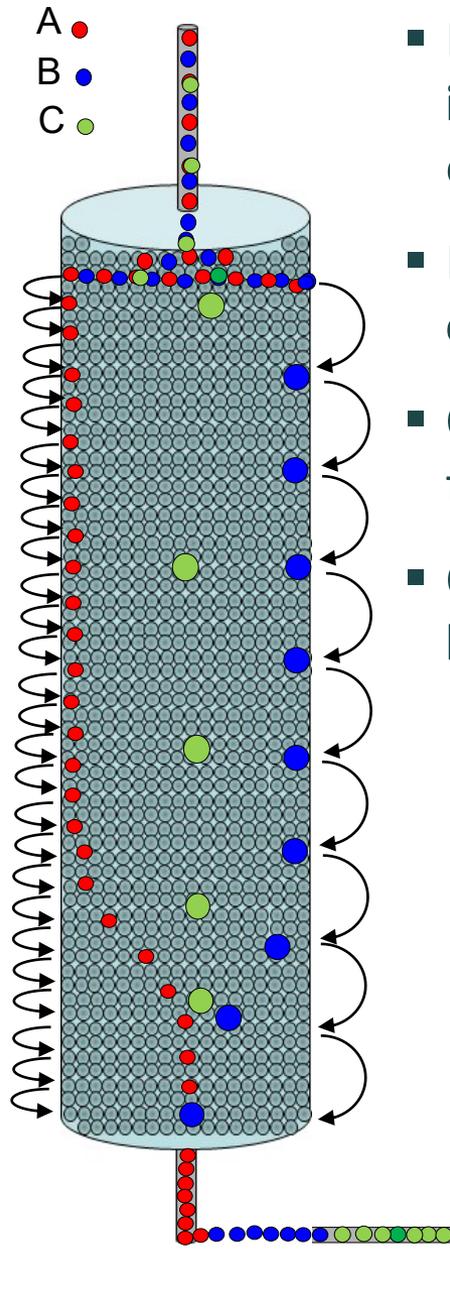


Elettroforesi bidimensionale – stato metabolico delle cellule

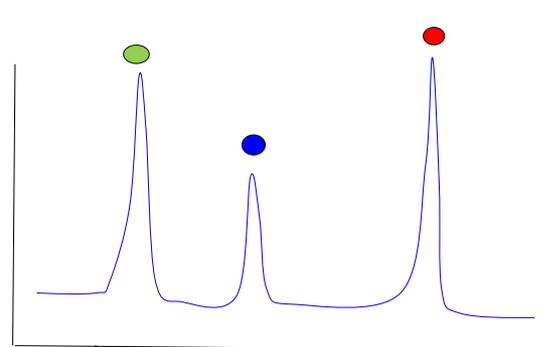


SDS - PAGE

METODI CROMATOGRAFICI



- La separazione di macromolecole in una miscela dipende dall' **interazione differenziale con la matrice** presente nella colonna cromatografica (nota come la **fase stazionaria**).
- Le proteine sono disciolte nella **fase mobile** (un solvente noto come **eluente**) che transita nella colonna ad una velocità definita (ml/min)
- Quelle che non interagiscono con la fase stazionaria escono con il fronte dell'eluente.
- Quelle che interagiscono sono rallentate. **Più forte l'interazione, più lento il tempo di transito.**



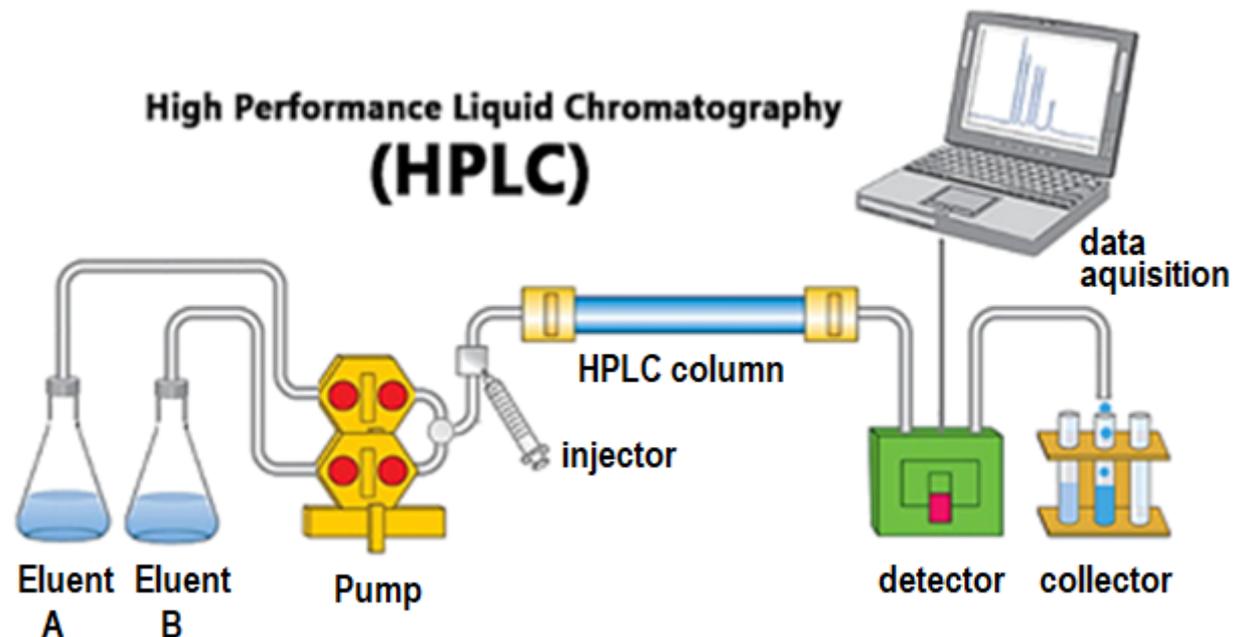
Le proteine si separano in base alle loro **caratteristiche chimico-fisiche**:

- 1) **Idrofobicità/polarità** della superficie
- 2) **Carica**
- 3) **Grandezza**
- 4) **Interazioni specifiche** con ligandi

Rilevatore
(es. spettrofotometro)

HPLC

- ▶ La separazione cromatografica delle macromolecole biologiche in generale e proteine in particolare richiede l'uso di HPLC (High Performance Liquid Cromatography)
- questa tecnica utilizza colonne cromatografiche impaccate con **matrici formate da particelle molto piccole** per **aumentare la risoluzione** della separazione
- questo rallenta il transito dell'eluente; è un problema perché questo riduce la risoluzione dovuto al movimento browniano (diffusione) delle molecole da separare
- per **aumentare la velocità di eluizione** serve quindi pompare l'eluente nella colonna con **pressioni elevate** (inizialmente HPLC = High Pressure Liquid Chromatography)
- Richiede colonne resistenti (spesso in metallo), e strumentazione particolare



CROMATOGRAFIA A FASE INVERSA

► La cromatografia a fase inversa è associata a HPLC → RP-HPLC (reversed phase)

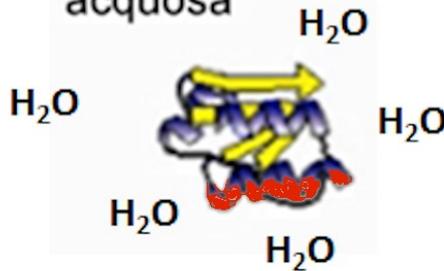
- fase stazionaria organica (**idrocarburo**); fase mobile acqua → organica

proteine legano alla fase stazionaria mediante **interazioni idrofobiche**
no in base ad un' **aumentata *organicità*** della fase mobile

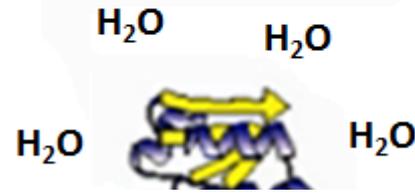
C18

C4

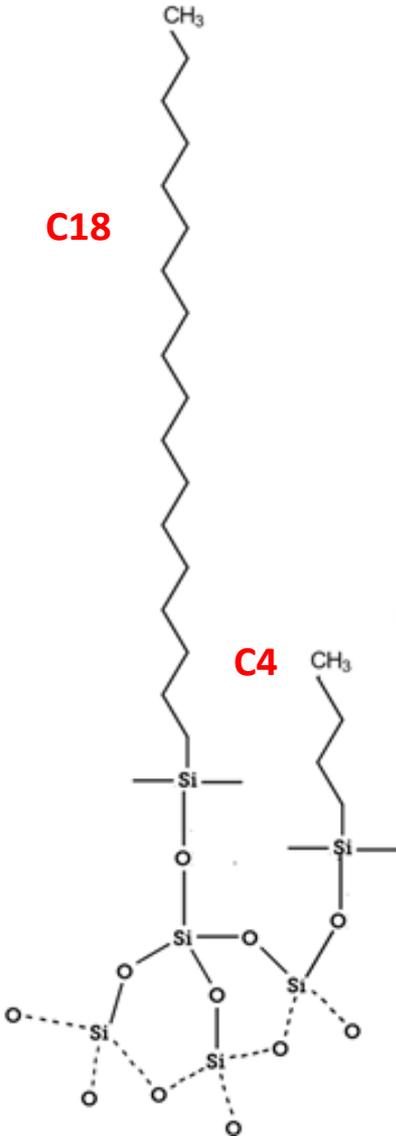
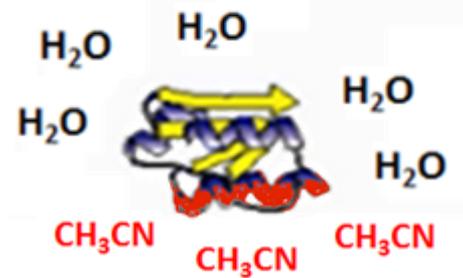
Il polipeptide
entra nella
colonna
disciolto nella
fase mobile
acquosa



Adsorbe nella superficie
organica stazionaria
composta da catene
alchiliche C4-C18



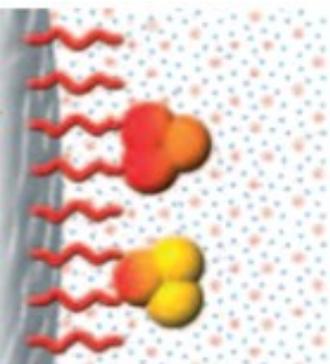
Desorbe dalla
superficie RP quando la
concentrazione di
solvente organico
nella fase acquosa
raggiunge un valore
critico



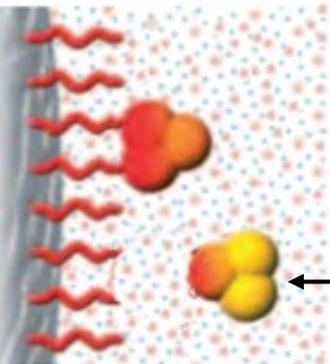
RP-HPLC

Reverse phase

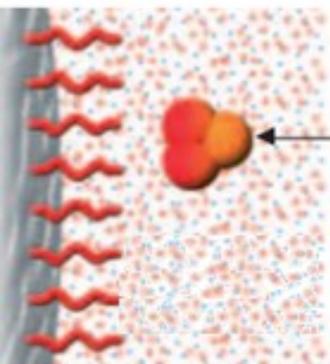
- RP-HPLC utilizza un gradiente di eluizione $H_2O \rightarrow Org$ (es. CH_3CN)



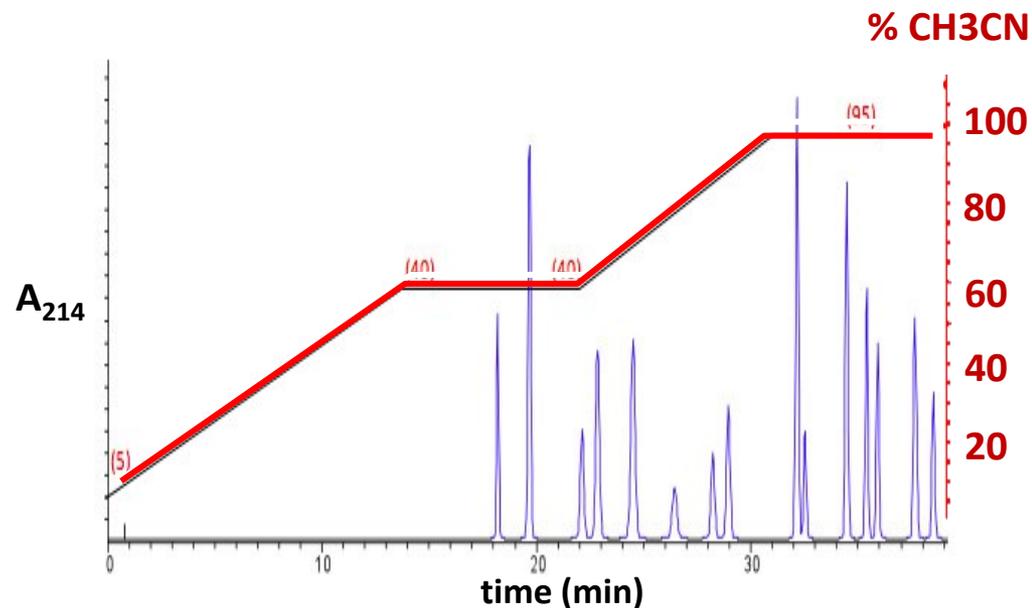
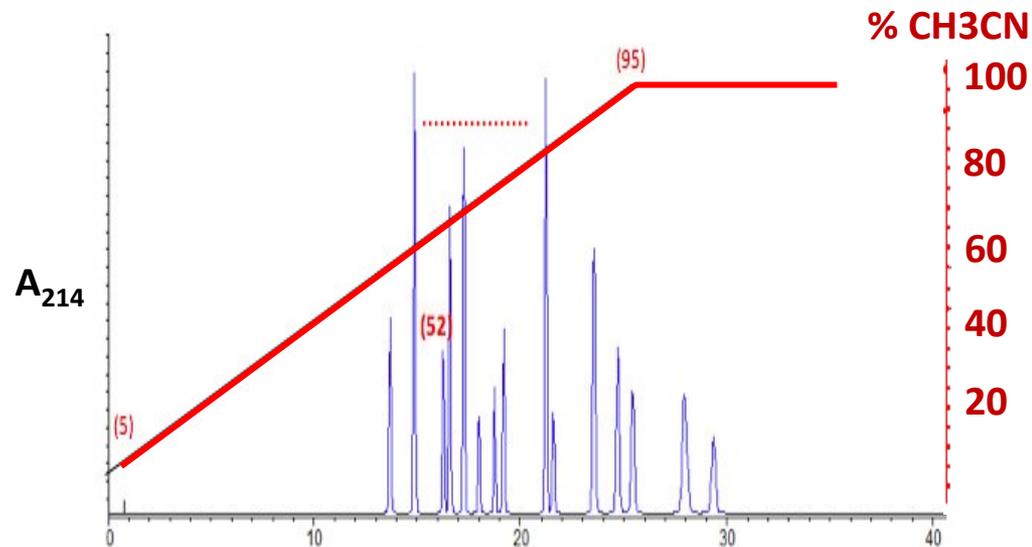
95% H_2O , 0.05% TFA



75% H_2O , 0.05% TFA
25% CH_3CN , 0.05% TFA
meno idrofobica



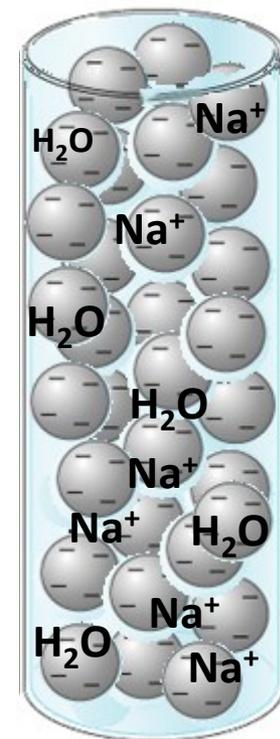
More hydrophobic
più idrofobica
50% H_2O , 0.05% TFA
50% CH_3CN , 0.05% TFA



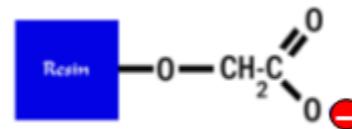
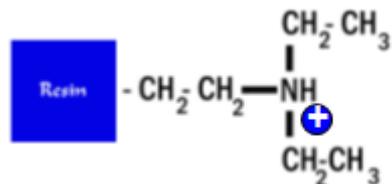
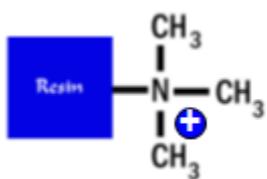
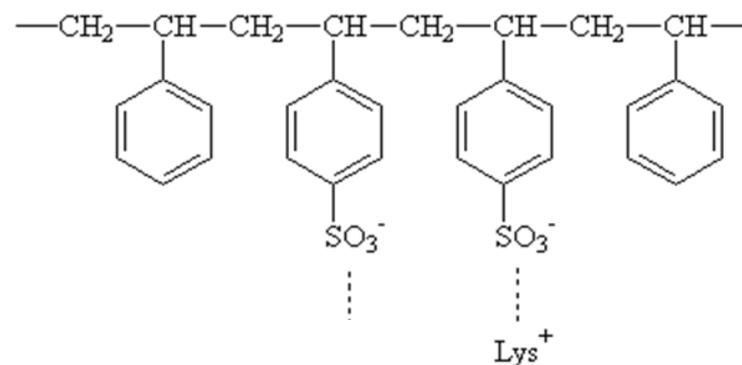
CROMATOGRAFIA A SCAMBIO IONICO (ION EXCHANGE - IEX)

► Fase stazionaria (polistirene o polisaccaride) modificata con gruppi cationici (+) o anionici (-)

- peptidi/proteine legano alla fase stazionaria mediante interazioni elettrostatiche.
- la fase mobile è acquosa con concentrazione crescente di sali o pH
- i polipeptidi si staccano in base alla crescente forza ionica o pH della fase mobile
- gruppi ionici comunemente utilizzati per modificare la fase stazionaria sono:

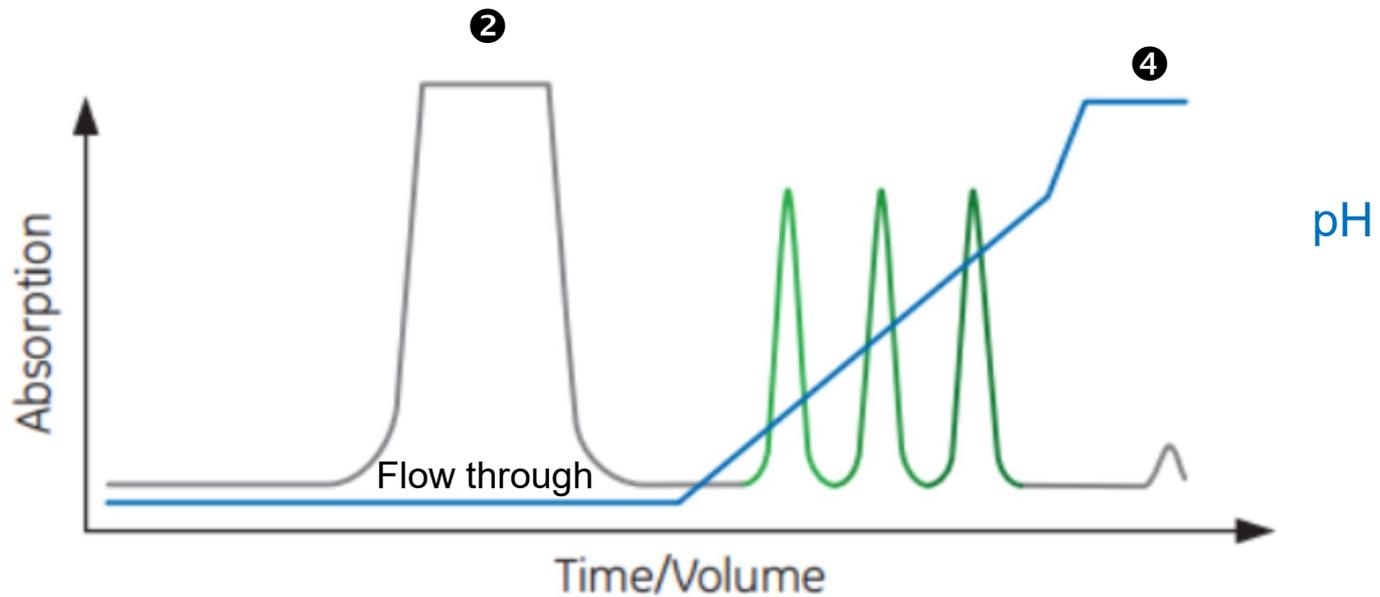
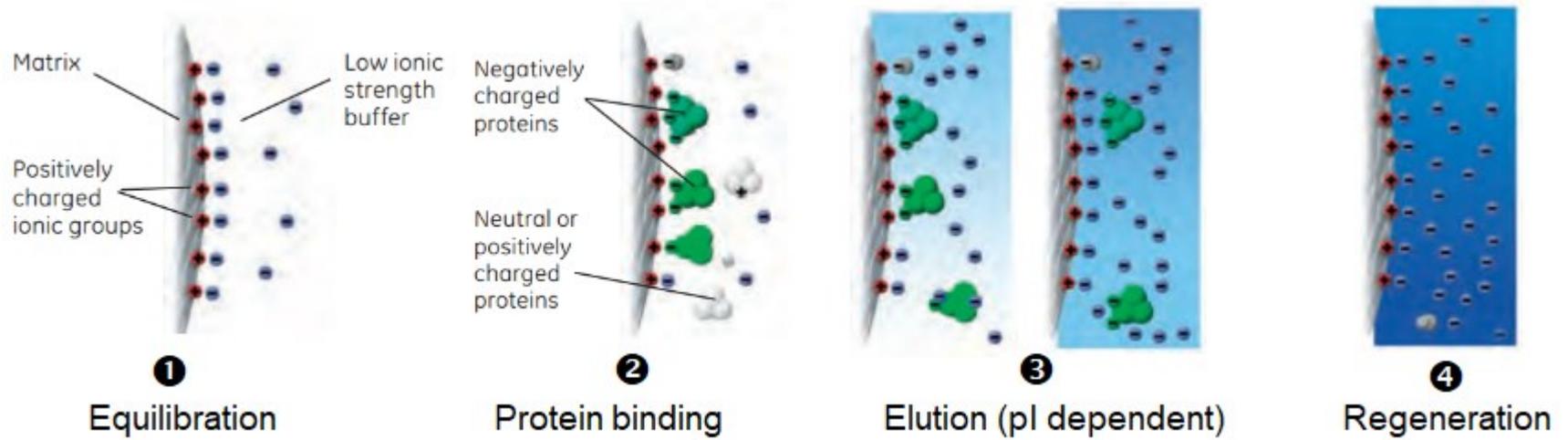


- gruppi carbossimetile **(anionico debole)**
- gruppi solfato **(anionico forte)**
- gruppi dietilammino etile **(cationico debole)**
- gruppi amminoetile 4° **(cationico forte)**



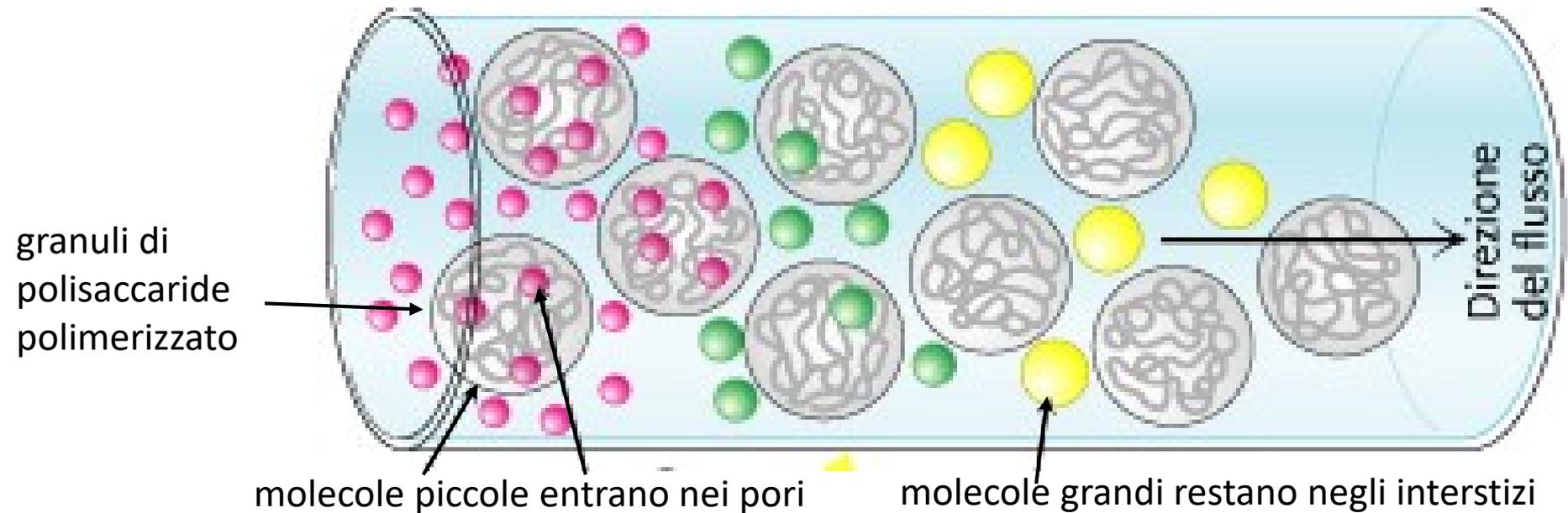
resina PS-SO₃⁻

IEX-HPLC



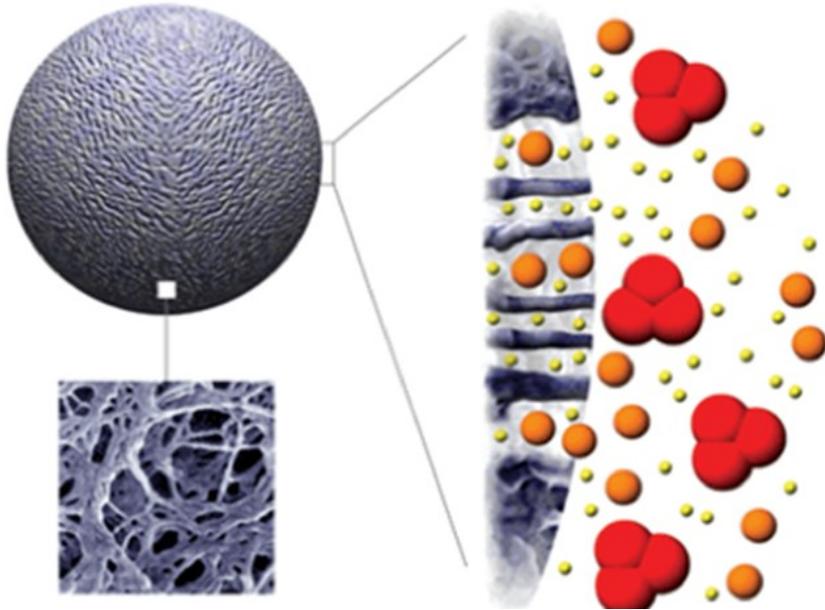
CROMATOGRAFIA A GEL PERMEAZIONE (GPC)

- ▶ Un tipo di cromatografia ad esclusione (*Size Exclusion Chromatography* - SEC) nota anche come *molecular sieve chromatography* (setaccio molecolare)
 - la fase stazionaria è un gel di polisaccaride supportato da una matrice solida
 - le proteine si separano in base alla loro grandezza.
 - quelle più piccole entrano nei pori del gel e sono rallentate
 - quelle più grandi sono escluse dai pori e passano per gli interstizi fra le particelle della fase stazionaria, eluendo più velocemente.



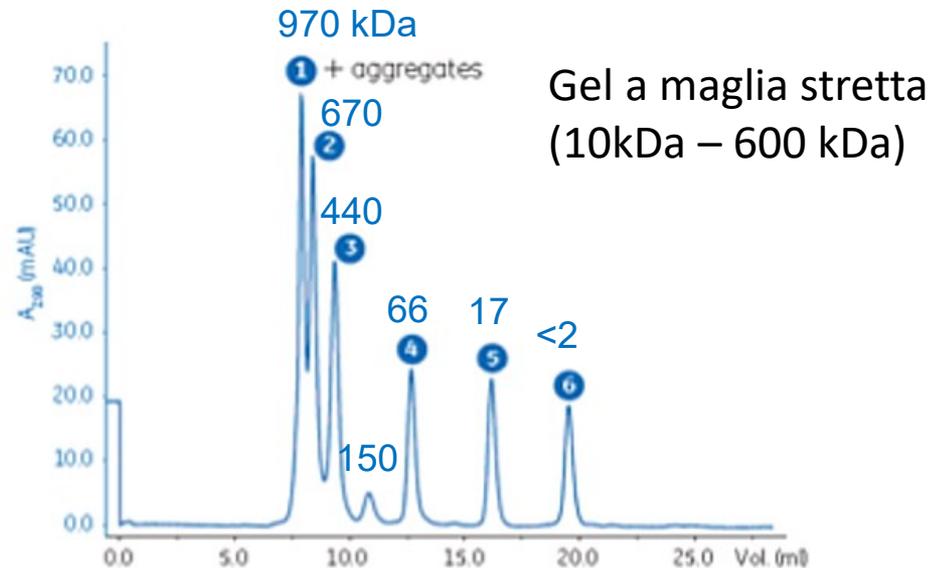
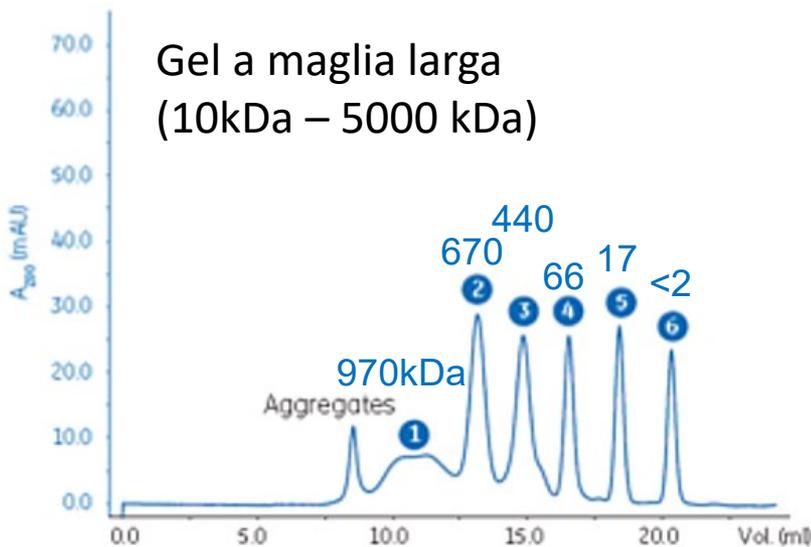
GPC-HPLC

► Questa cromatografia è eseguita in eluizione isocratica (tampone acquoso)



- nolecole grandi non riescono ad entrare nei pori delle particelle cromatografiche
- roteina separata che entra nei pori maggiori ed è rallentata
- piccole proteine e sali entrano in tutti i pori in profondità e sono molto rallentati

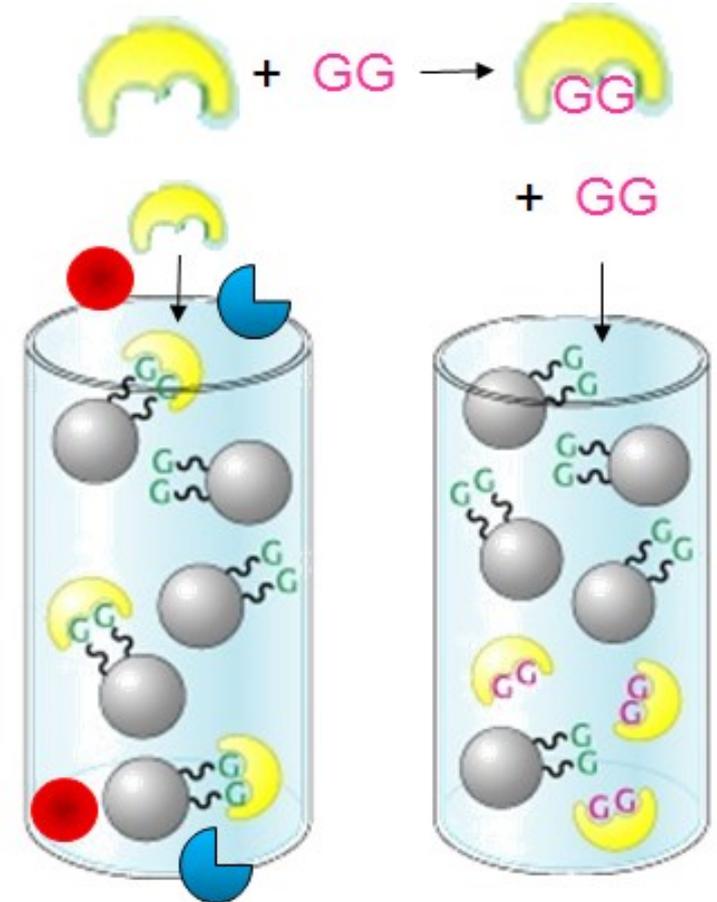
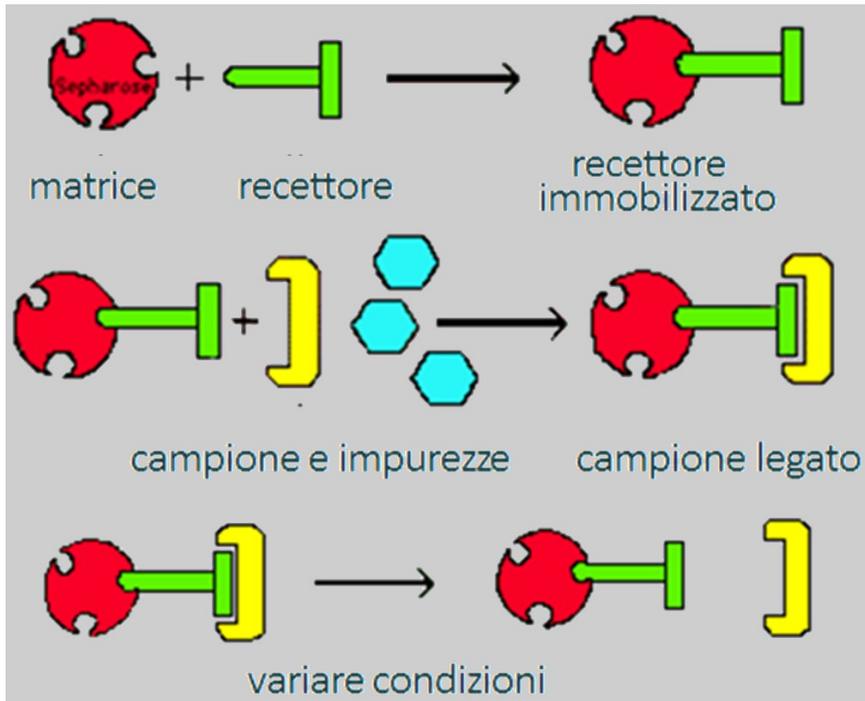
Una buona risoluzione richiede la scelta della maglia del gel più adatta alla proteina da purificare



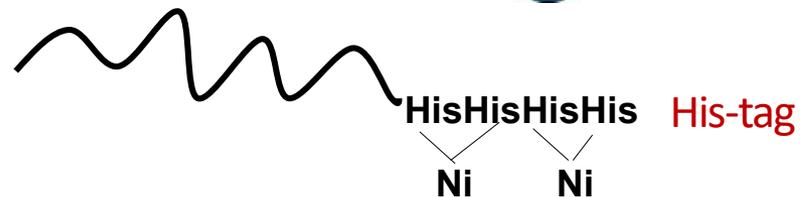
CROMATOGRAFIA AD AFFINITÀ (X-AC)

► Separazione in base all' interazione specifica della specie da separare con una specie 'recettore' (o *ligando*) immobilizzata alla fase stazionaria

- interazioni proteina/ligando; anticorpo/antigene; chelante/metallo, oligonucleotide/complementare

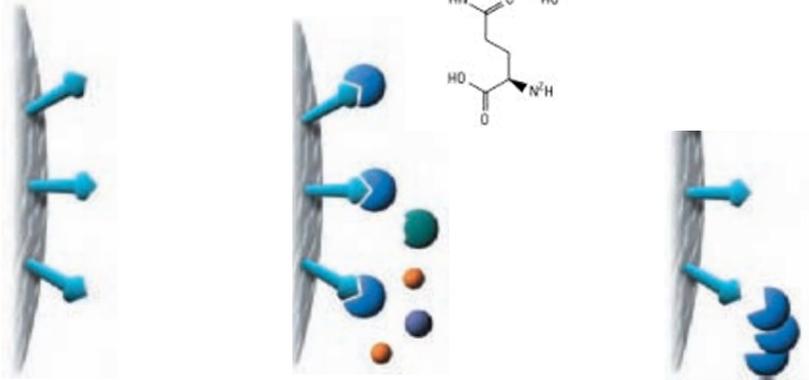
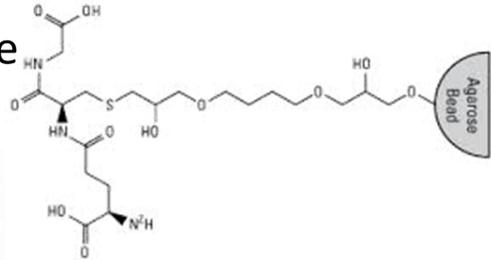


- Es. IMAC – Immobilized Metal AC



AC-HPLC

glutathione

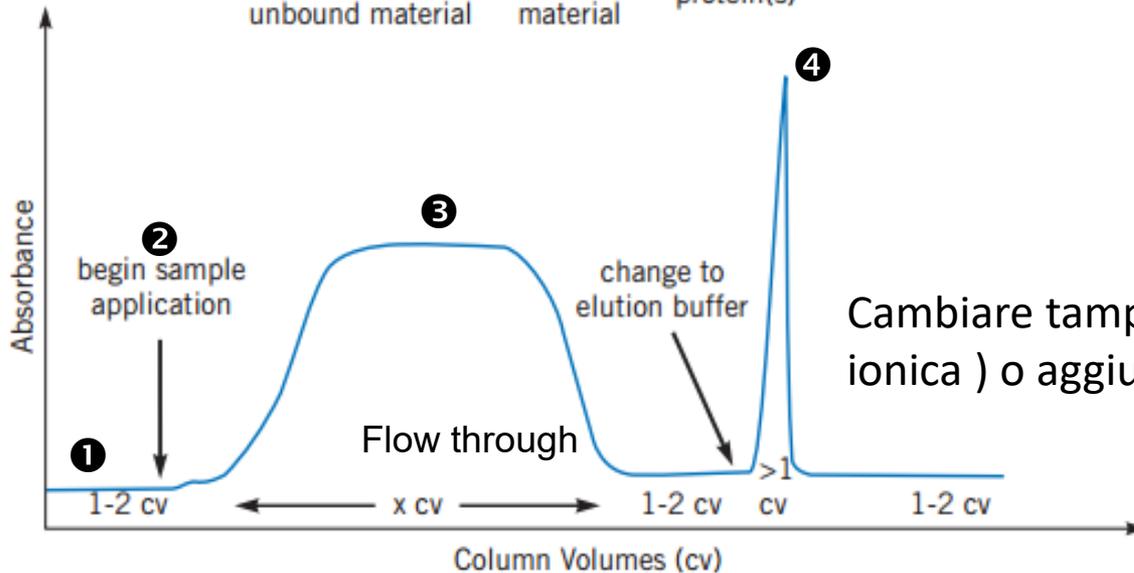
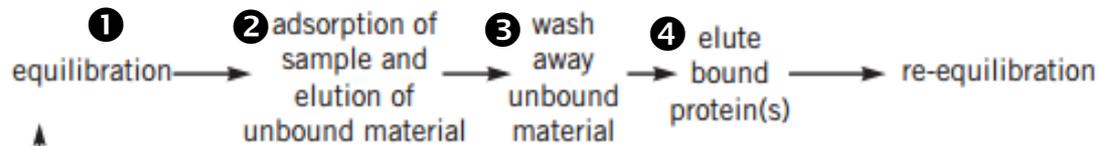


RECETTORE

- ligando \leftrightarrow
- anticorpo \leftrightarrow
- acido nucl. \rightarrow
- ormone \leftrightarrow
- metallo \rightarrow

SPECIE DA SEPARARE

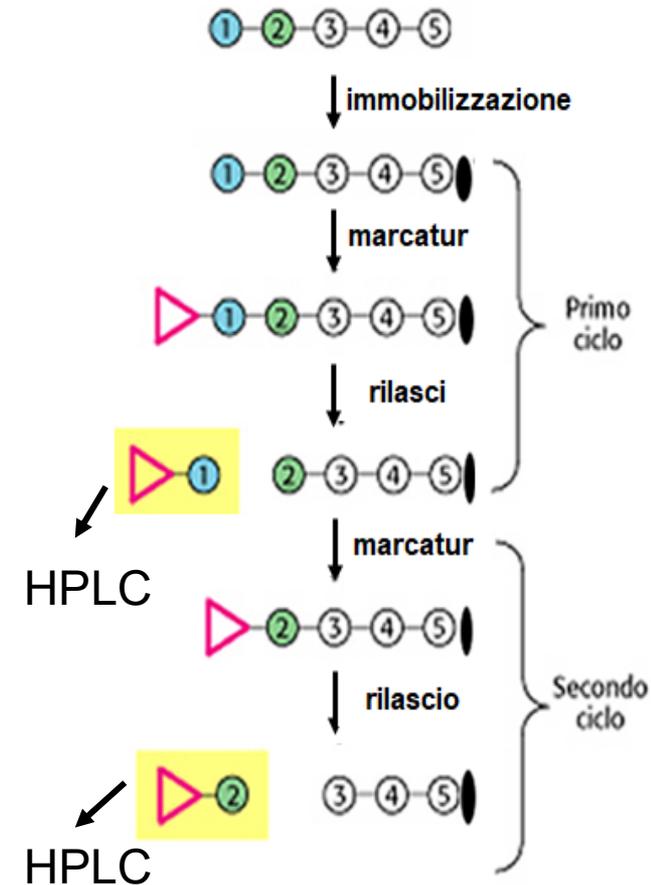
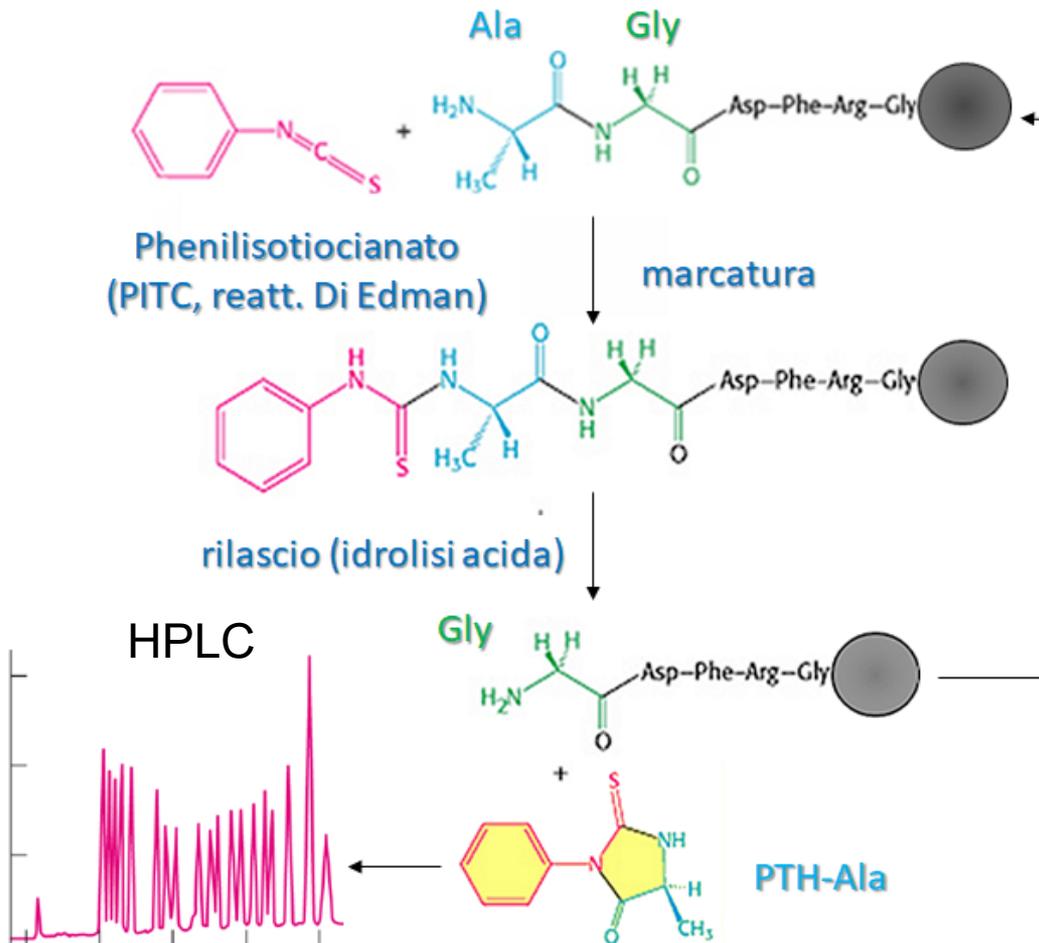
- \leftrightarrow proteina
- \leftrightarrow antigene
- \rightarrow filamento complementare
- \leftrightarrow recettore
- \rightarrow proteina con His-tag



DETERMINAZIONE DELLA STRUTTURA 1^a: Metodo di Edman

► Degradazione (sequenziamento) di Edman (metodo archaico)

- rilascio sequenziale di singoli AA marcati (derivatizzati) con una sonda colorata all'N-terminale di una proteina immobilizzata su un supporto solido
- rilevazione dei singoli AA mediante RP-HPLC e nuovo ciclo di degradazione



STRUTTURA PRIMARIA: Metodi alternativi (moderni)

► La sequenza amminoacidica di una proteina può essere dedotta da quella del gene che la codifica

- questo metodo è **più facile ed efficace per sequenze lunghe**
- è **più semplice sequenziare DNA** che proteine, la **sequenza è spesso già in banche dati**
- utilizzando la nota tabella di codoni, e poi semplice ottenere la sequenza proteica

Sequenza del DNA	GGG	TTC	TTG	GGA	GCA	GCA	GGA	AGC	ACT	ATG	GGC	GCA
Sequenza degli amminoacidi	Gly	Phe	Leu	Gly	Ala	Ala	Gly	Ser	Thr	Met	Gly	Ala

- La degradazione di Edman (arriva a ~ 50 residui) può essere comunque necessaria per isolare il gene che codifica una proteina o identificarlo dalle banche dati

► La sequenza di proteine può essere determinata mediante spettrometria di massa (MS)

- Si usa un metodo noto come mass fingerprinting, frammentando la proteina e confrontando la massa dei frammenti con quelli in una banca dati.

FRAMMENTAZIONE DELLE PROTEINE

- ▶ Per determinarne la sequenza è utile poter frammentare le proteine
 - la scissione della catena in frammenti è normalmente operata da **proteasi**
 - i vari **frammenti sono separati** usando RP-HPLC
 - i **frammenti sono poi sequenziati** usando il metodo di Edman o MS fingerprinting
 - Le sequenze dei frammento sono poi ordinate con il **metodo della sovrapposizione (*overlapping*)** che permette la ricostruzione della sequenza

Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa

sequenza ignota

digestione con tripsina
(taglia dopo Lys o Arg)

digestione con chimotripsina
(taglia dopo aromatici)

sequenziamento

Ala-Ala-Trp-Gly-Lys
Thr-Phe-Val-Lys

Val-Lys-Ala-Ala-Trp
Gly-Lys **Thr-Phe**

sovrapposizione

Thr-Phe-Val-Lys•Ala-Ala-Trp-Gly-Lys

Thr-Phe•Val-Lys-Ala-Ala-Trp•Gly-Lys

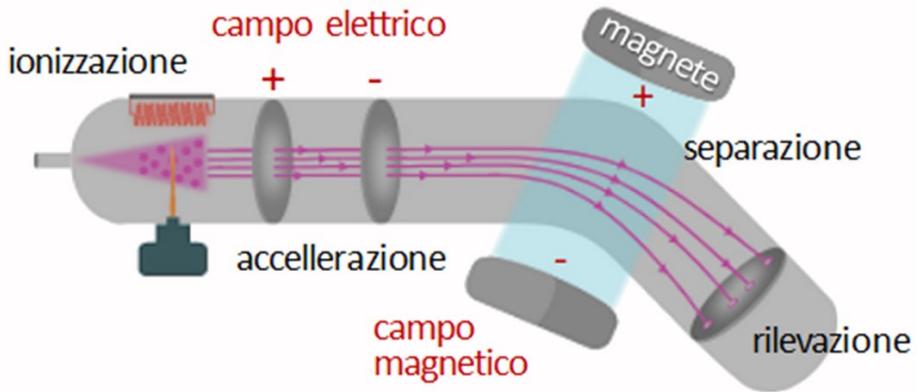
DETERMINAZIONE DELLA GRANDEZZA: Spettrometria di Massa

► Per ottenere un spettro di massa di una molecola serve:

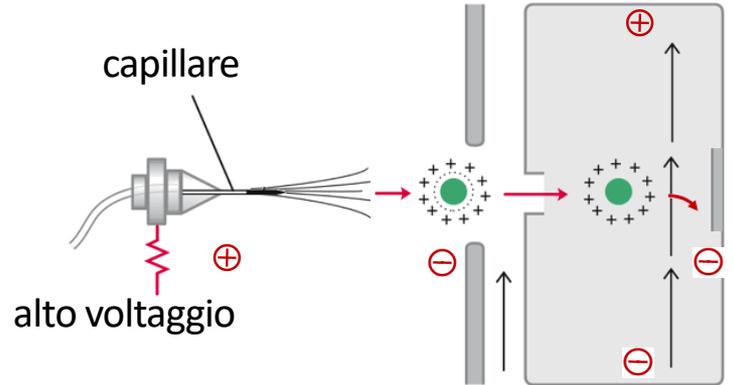
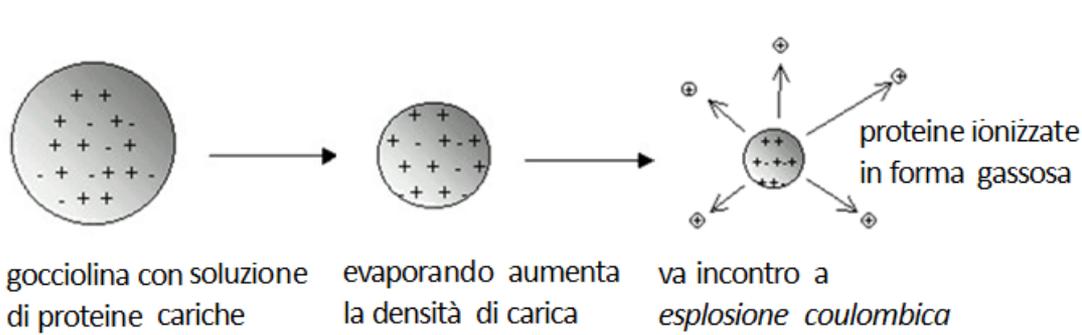


- la proteina viene accelerata e separata in base al rapporto massa/carica (m/z)

massa ↑ accelerazione ↓ deflessione ↓
carica ↑ accelerazione ↑ deflessione ↑



► ElectroSpray Ionisation (ESI-MS)



Spettrometria di massa MALDI-TOF

► Matrix Assisted Laser Desorption Ionization – Time Of Flight (MALDI-TOF)

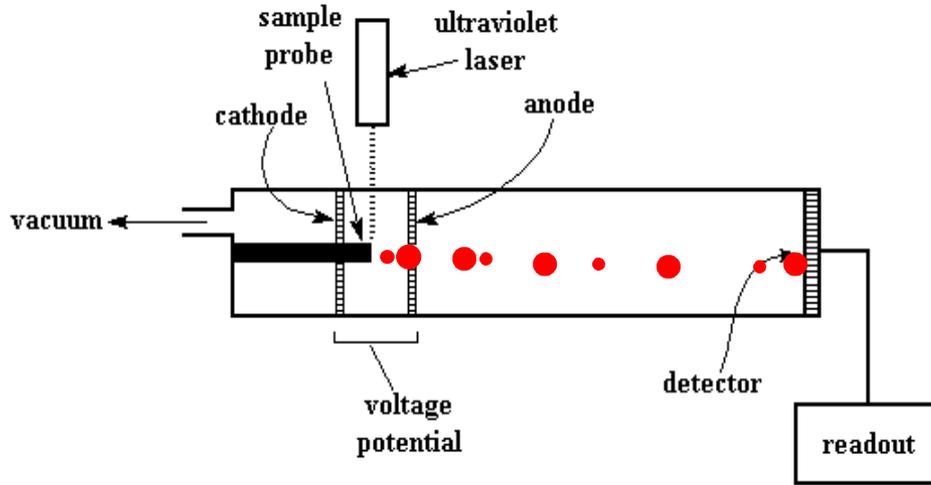
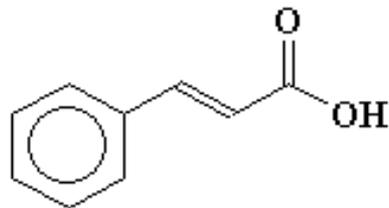


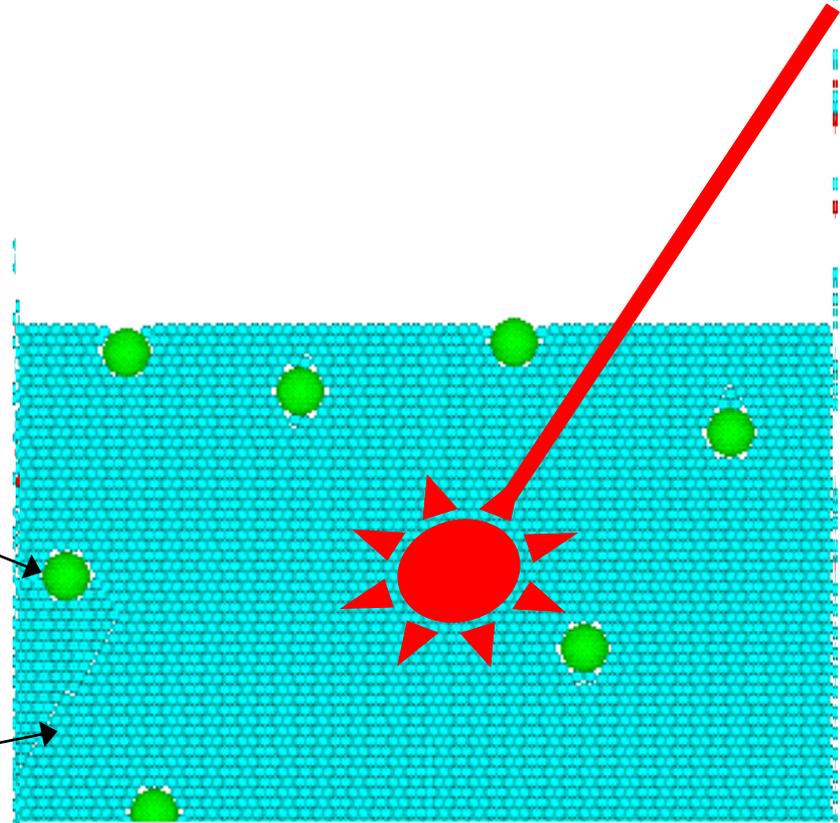
Diagramma semplificato dell'apparecchiatura MALDI



acido trans-cinnamico
acido

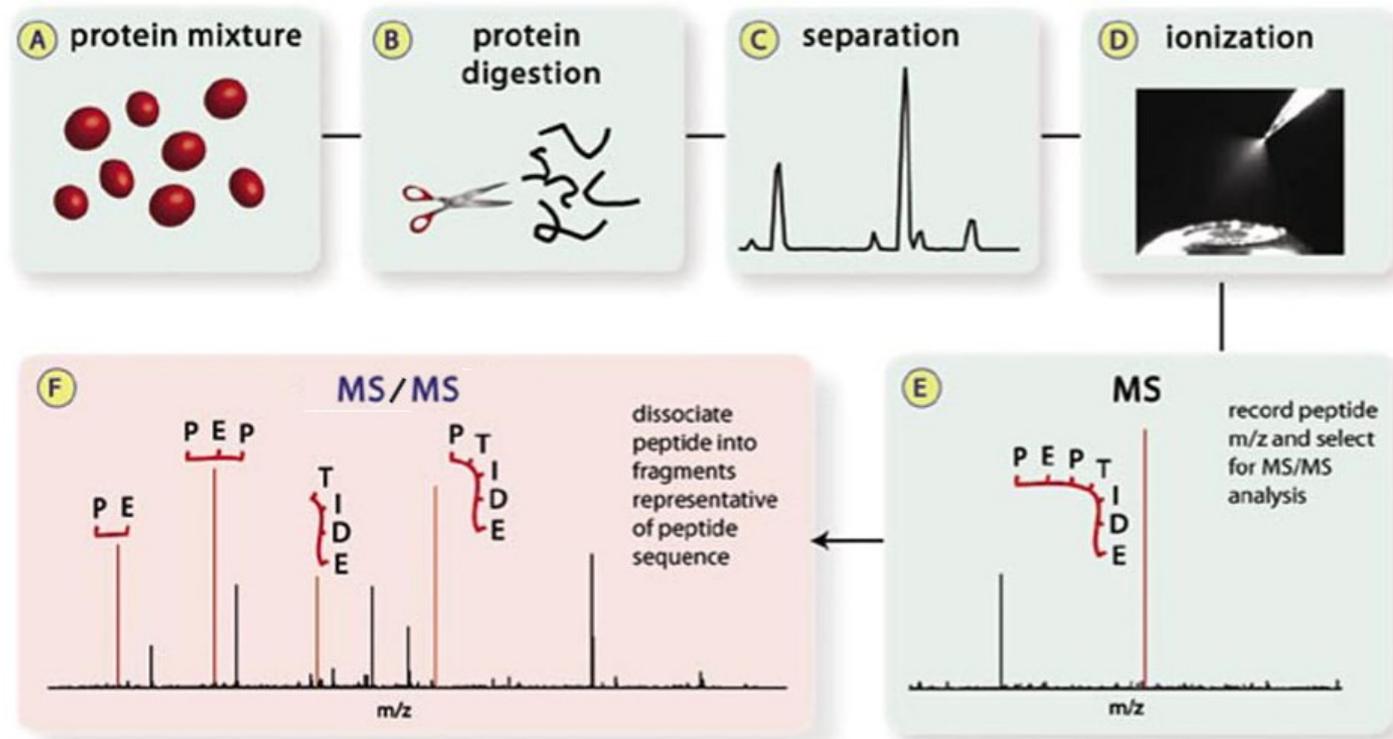
proteine

biochimica



Utilizzi della spettrometria di massa

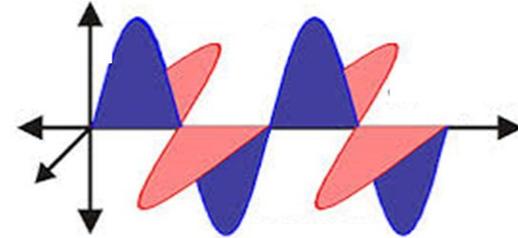
- ▶ La spettrometria di massa (MS) è un tecnica fondamentale per la biochimica
 - è spesso associata con la cromatografia in fase liquida (LC-MS), permette l'identificazione univoca delle proteine in un campione (esempio in un gel bidimensionale)
 - permette di identificare mutazioni, o modifiche post-traduzionali (PTM) nelle proteine (es. fosforilazione, glicosilazione ecc.).
 - permette il sequenziamento di frammenti di una proteina per poterla identificare (richiede un spettrometro tandem - MS/MS)



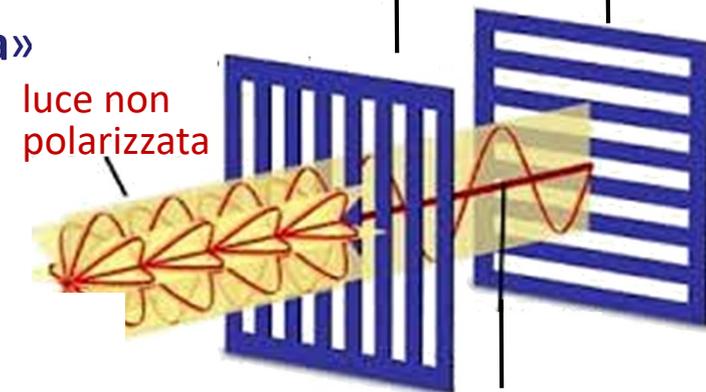
DETERMINAZIONE DELLA STRUTTURA 2^a: CD

► Dicroismo Circolare (CD) è una tecnica spettrofotometrica che permette di determinare il contenuto di elementi di struttura 2^a nelle proteine

- si basa sull'**assorbimento differenziale di luce polarizzata**
- normalmente la luce non è polarizzata; le onde elettromagnetiche della luce oscillano in tutte le direzioni
- speciali filtri detti «**polarizzatori**» selezionano le componenti producendo luce «**linearmente polarizzata**»
- luce linearmente polarizzata a sua volta è composta da luce **circularmente polarizzata**



polarizzatore verticale e orizzontale

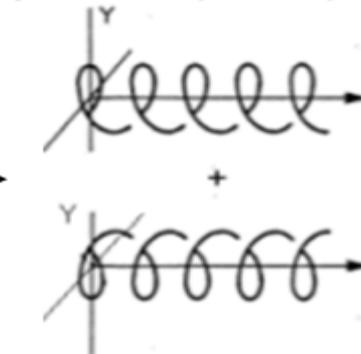


luce non polarizzata

luce linearmente polarizzata

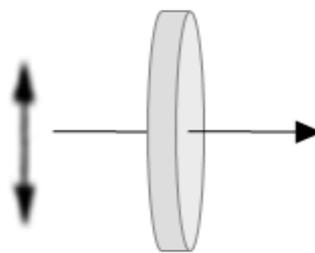
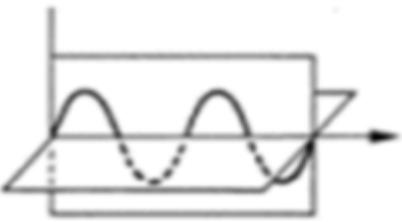


luce circularmente polarizzata (a destra)



luce circularmente polarizzata (a sinistra)

luce linearmente polarizzata

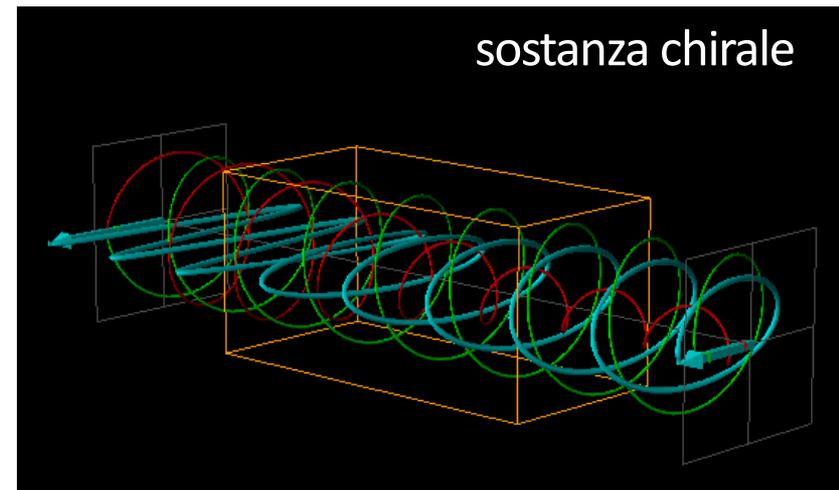
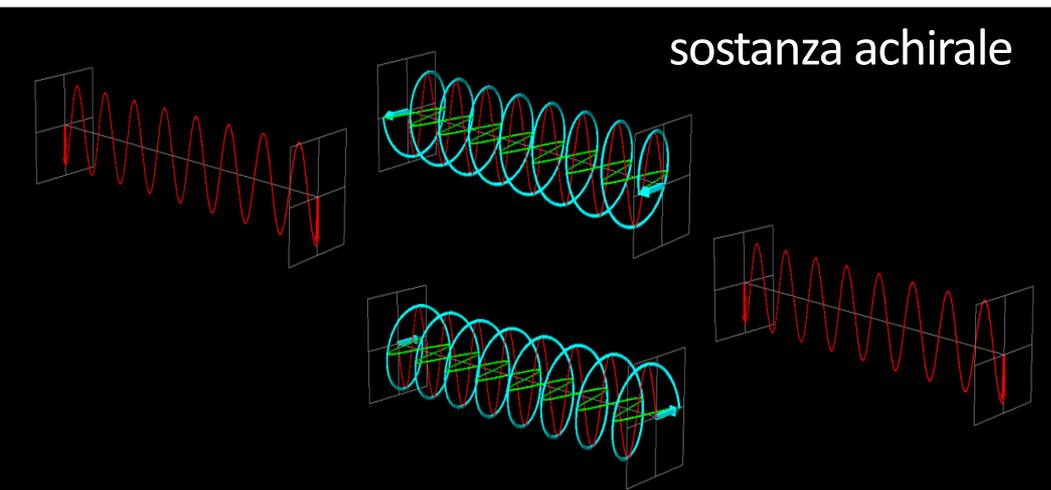
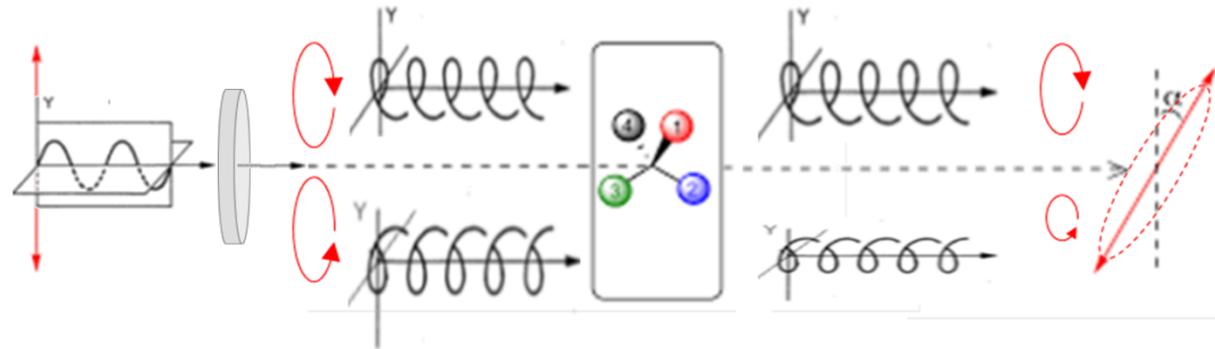


Utilizzo del dicroismo circolare

► Molecole chirali assorbono luce circolarmente polarizzata a destra o sinistra diversamente

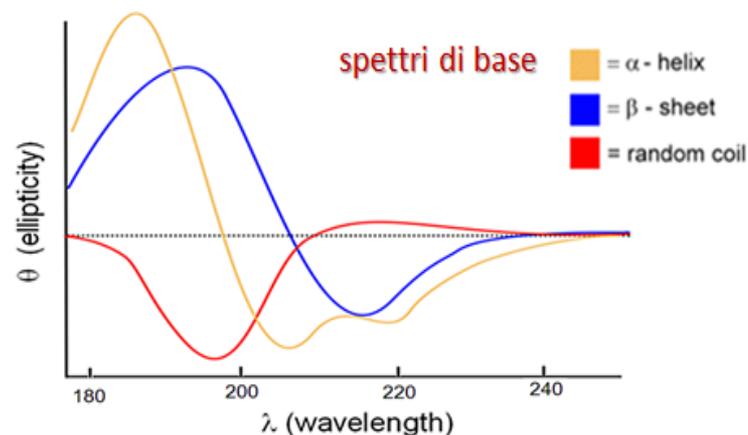
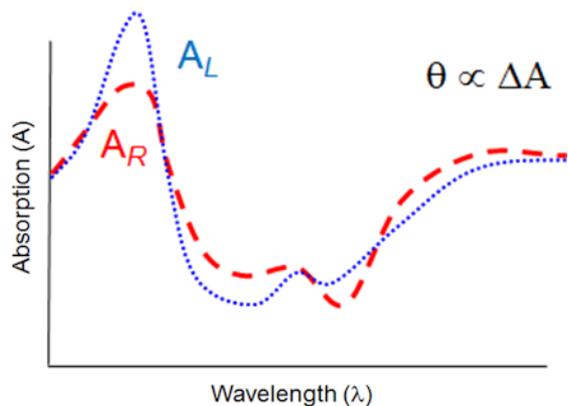
- $A = \varepsilon c l$ il coefficiente d'estinzione (ε) è diverso per le due componenti
- luce linearmente polarizzata che transita per una soluzione di sostanza chirale viene **ellitticamente polarizzata**. Il grado di ellitticità (θ) può essere misurato

$$\begin{aligned}\Delta A &= (\varepsilon_L - \varepsilon_R) \cdot c \cdot l \\ &= \Delta \varepsilon \cdot c \cdot l \\ \theta &\propto \Delta \varepsilon \propto \Delta A\end{aligned}$$

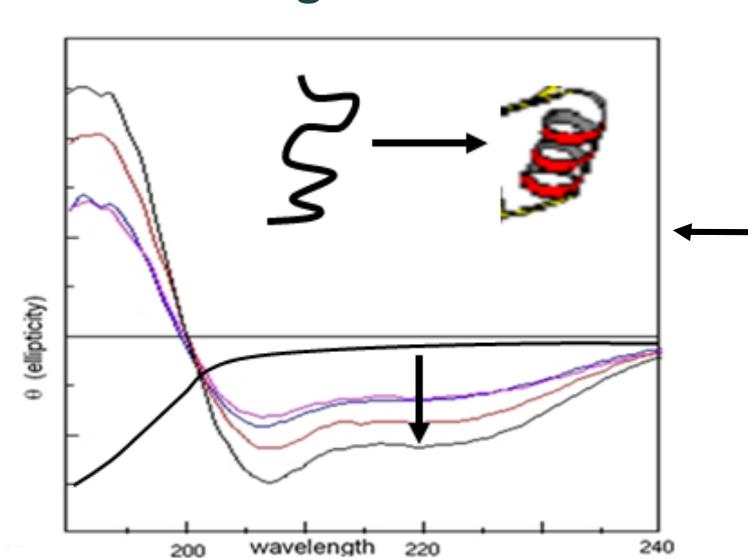
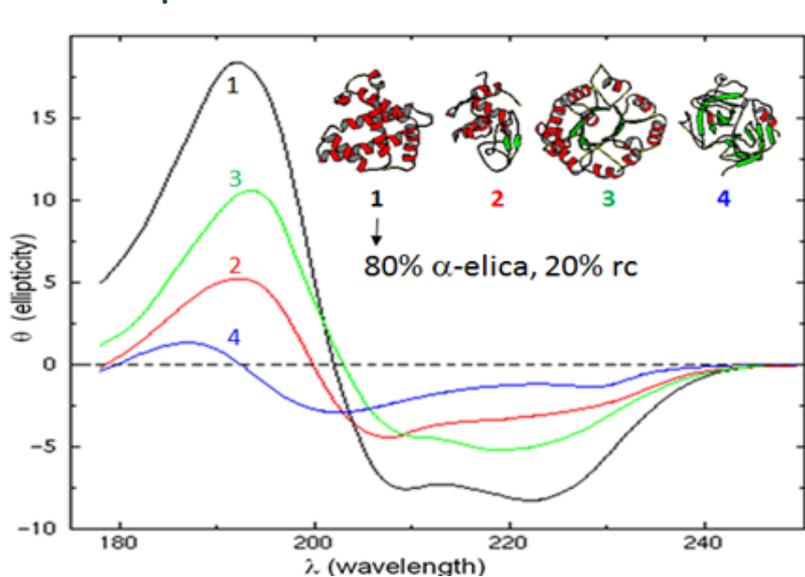


Utilizzo del dicroismo circolare (cont)

- Il grado di ellitticità può essere misurato e fornire uno spettro di «dicroismo circolare»
- nelle catene proteiche ci sono molti centri chirali posti in segmenti con **strutture 2^e a loro volta asimmetriche** – è questo che **determina lo forma dello spettro**

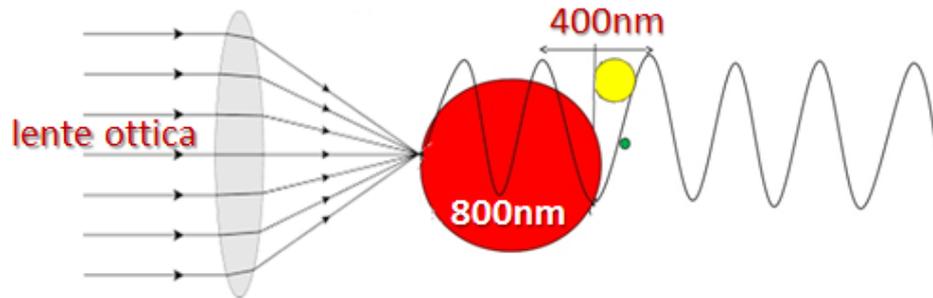


- serve per determinare il **contenuto conformazionale** e il **grado di strutturazione**

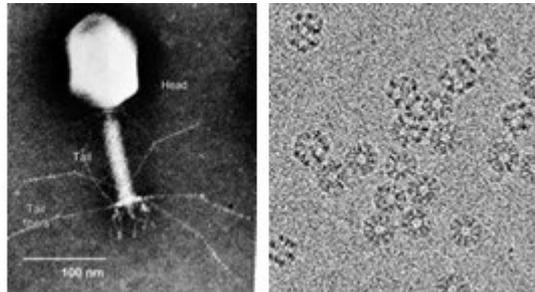


DETERMINAZIONE DELLE STRUTTURE 3^a e 4^a

- ▶ Nella microscopia ottica, Il limite di risoluzione è metà della lunghezza d'onda ($\frac{1}{2}\lambda$)



- ▶ Microscopia elettronica convenzionale il limite è ~ 10 nm

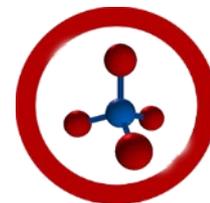


- ▶ La diffrazione dei raggi X ha un limite di ~ 1Å (~ lunghezza legame covalente)

- richiede la diffrazione da parte degli atomi in **molte molecole disposte in modo molto regolare**
- per questa ragione è necessario disporre di **cristalli**

- ▶ Altre tecniche utili per determinare la struttura di proteine sono:

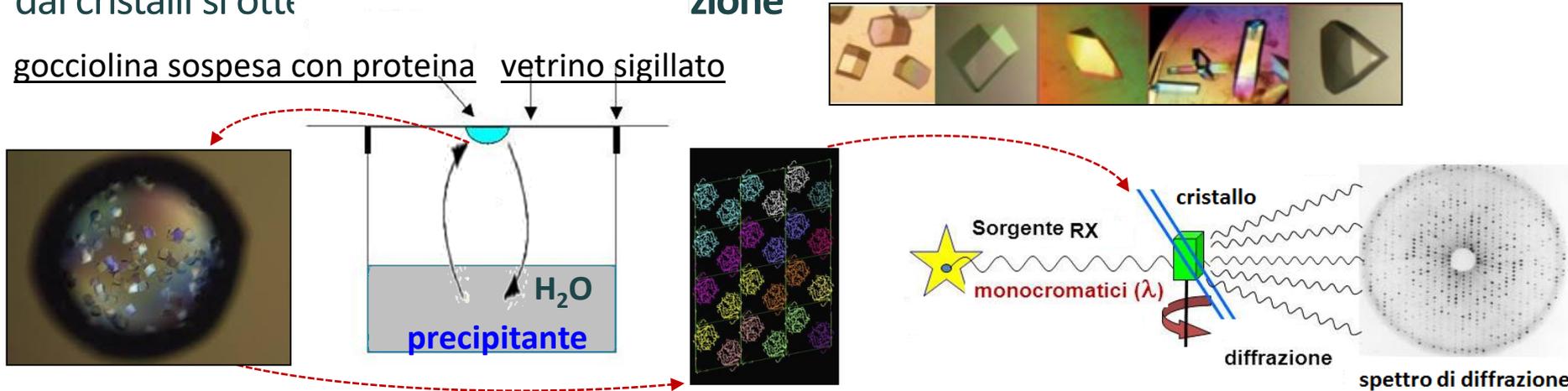
- spettroscopia a **risonanza magnetica nucleare** bidimensionale (2D-NMR)
- **microscopia crioelettronica (Cryo-EM)**



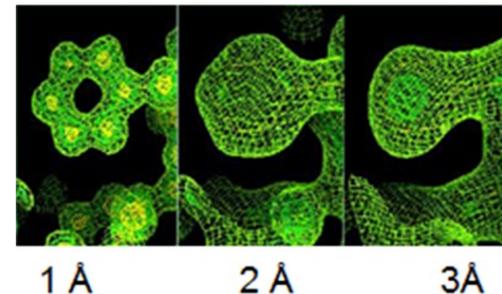
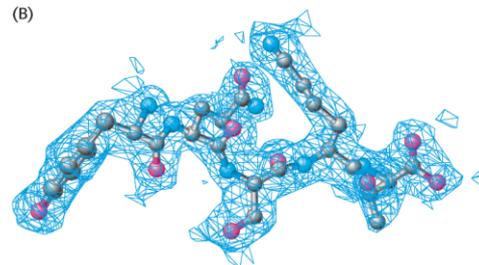
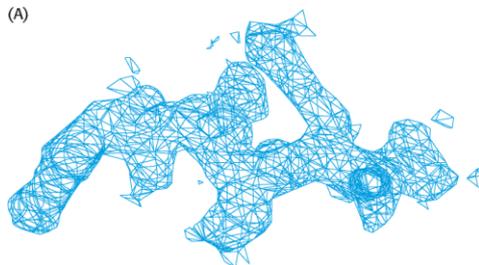
► Cristalli di proteine

- per ottenerli servono diversi mg/ml di proteina pura (spesso ricombinante)
- Il processo di cristallizzazione deve essere molto lento (es. **metodo hanging drop**)
- dai cristalli si ottengono **mappe di diffrazione**

gocciolina sospesa con proteina vetrino sigillato

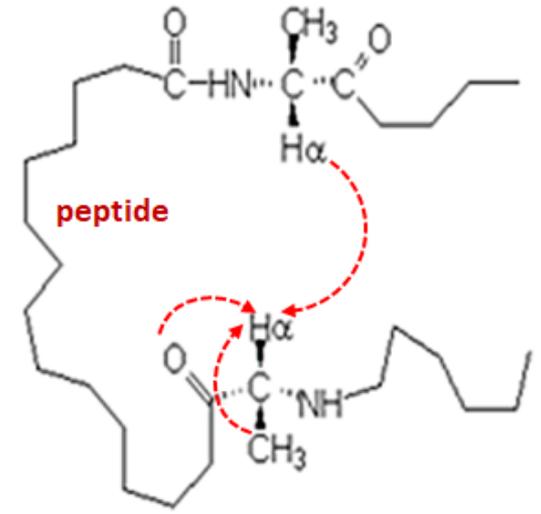
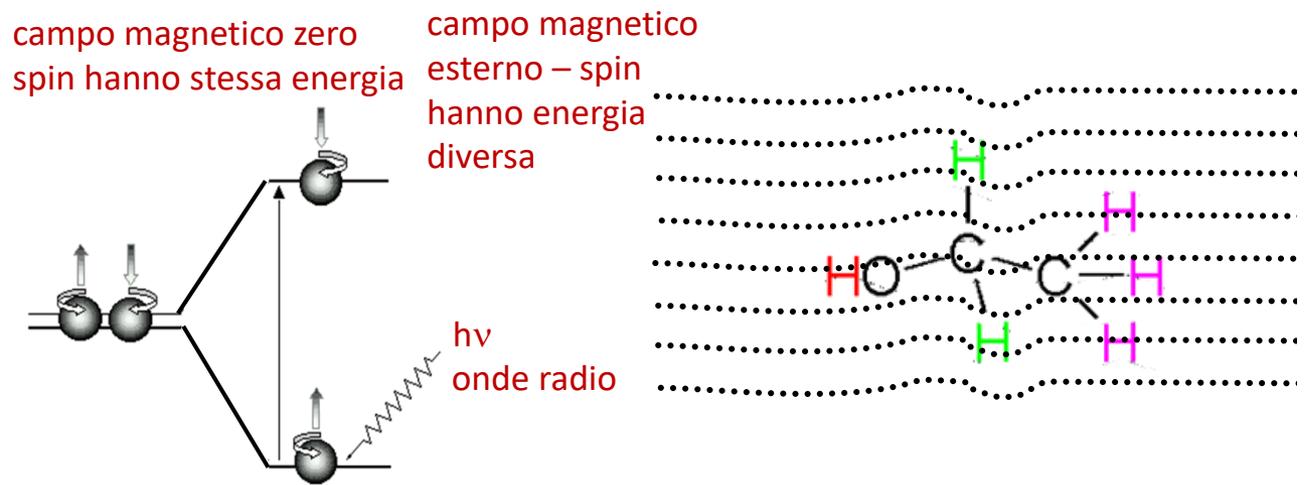


- la densità elettronica attorno agli atomi diffrange i raggi X; dalla mappa di diffrazione raccolta, si ottengono informazioni sulla densità elettronica e quindi sulla posizione degli atomi.
- migliore la qualità del cristallo e maggiore intensità e coerenza della luce, più elevata è la risoluzione



► **NMR rivela la struttura di molecole in soluzione ad elevata concentrazione**

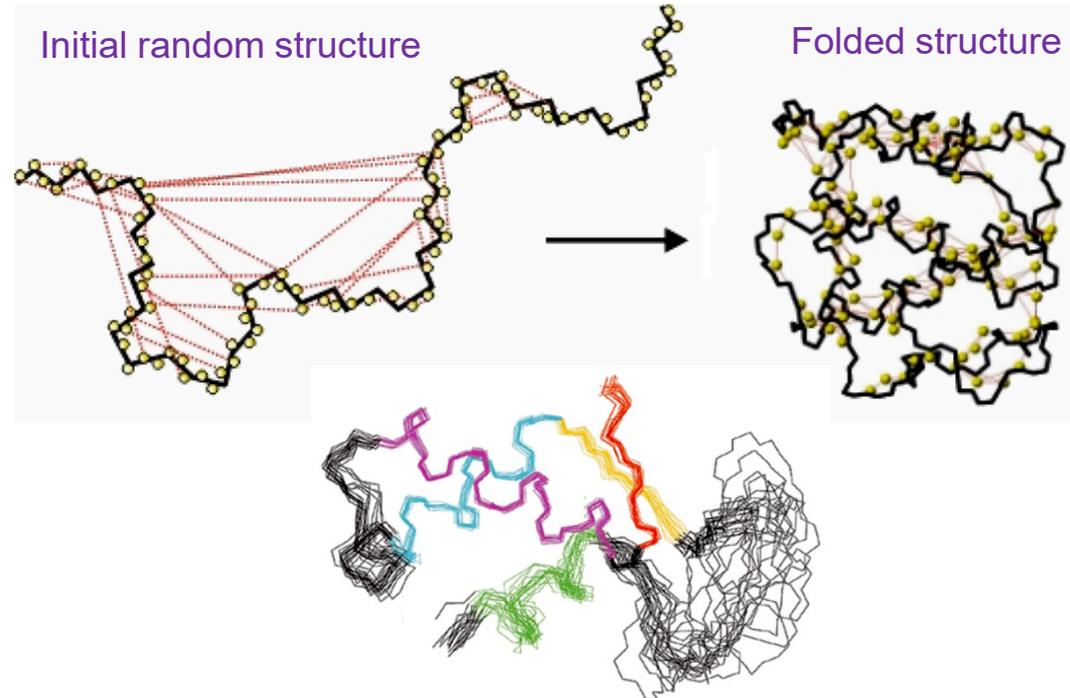
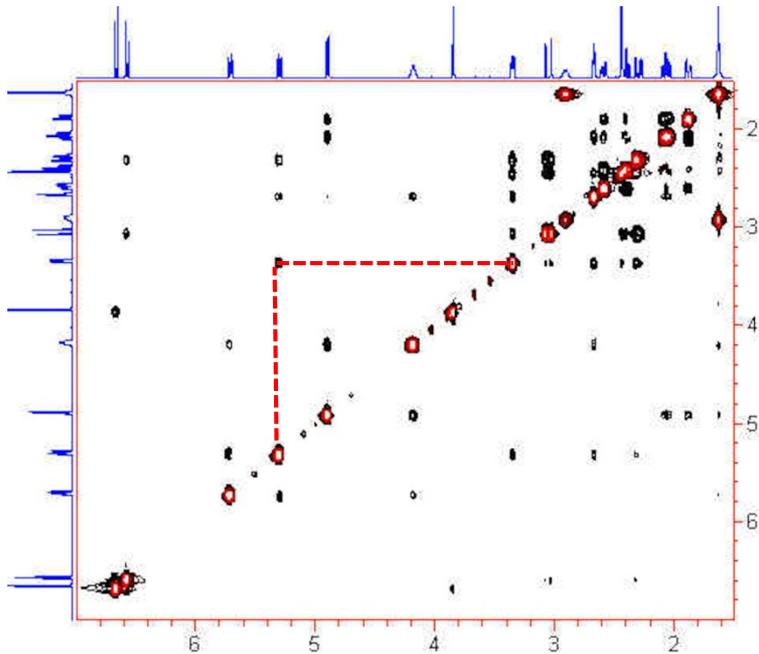
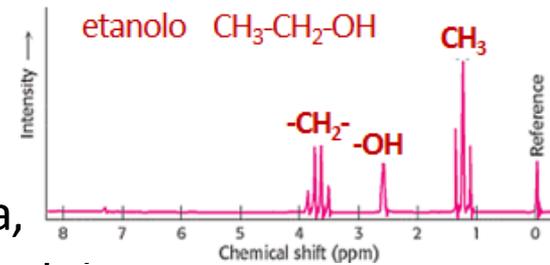
- si basa sulla transizione fra stati del *momento magnetico nucleare di spin* per alcuni isotopi di atomi presenti nelle biomolecole (^1H , ^{13}C , ^{15}N , ^{31}P)
- immersi in un forte campo magnetico esterno, gli stati energetici degli spin si separano
- quello meno energetico è favorito, ma assorbendo radiazione alla frequenze delle onde radio si ha la transizione al livello energetico più elevato
- oltre al campo magnetico esterno, gli stati energetici risentono dall'ambiente chimico attorno ad ogni nucleo
- Il campo magnetico è modificato dai nuclei vicini (connessi da legami o vicini nello spazio)



Spettrometria 2D-NMR (cont.)

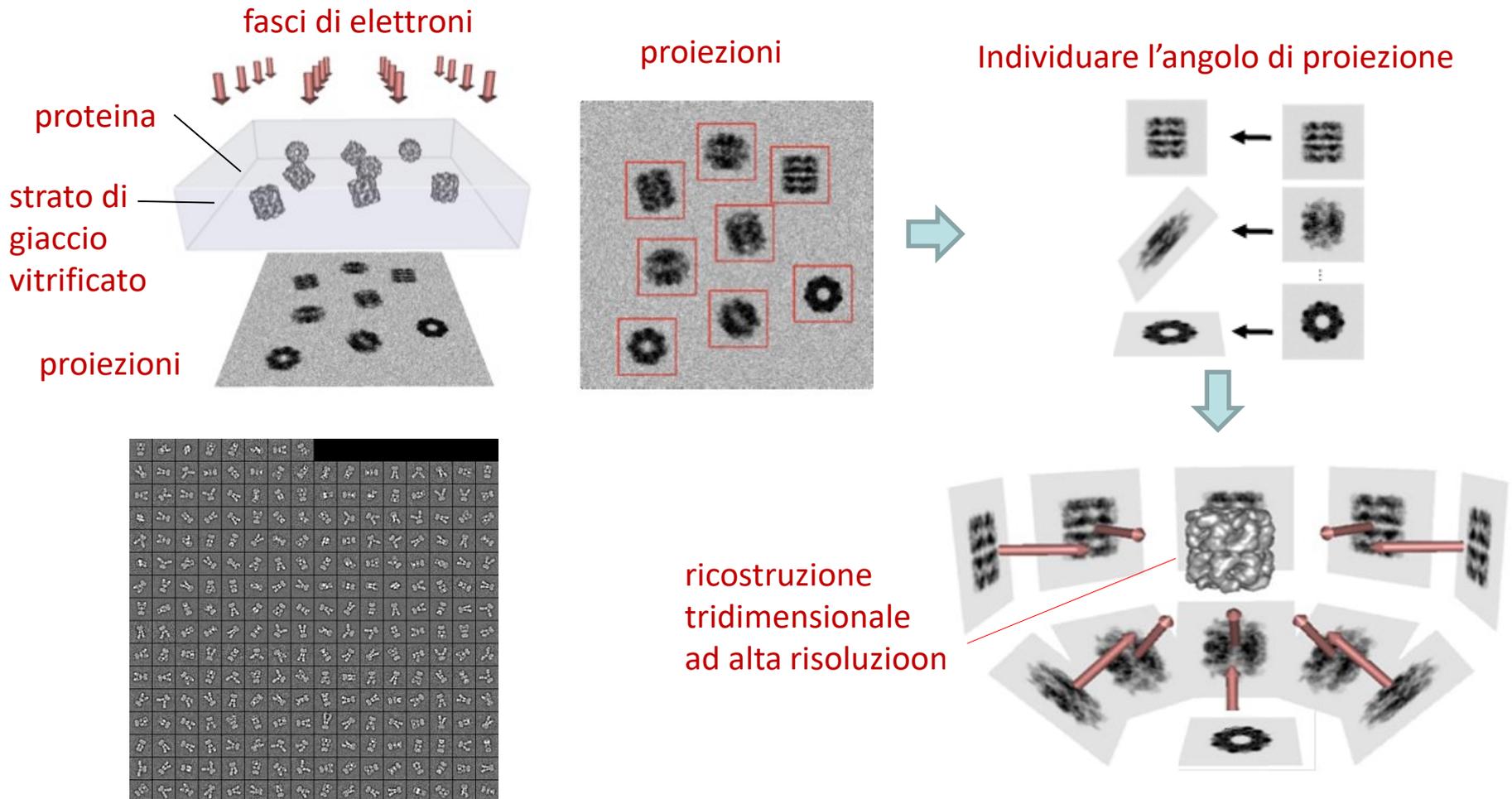
► $^1\text{H-NMR}$ permette la risoluzione di strutture relativamente semplici (piccole biomolecole)

- NMR convenzionale tiene conto solo di effetti trans-legame e non è adatto a macromolecole. Per questo serve il 2D-NMR che tiene conto anche di interazioni vicine nello spazio ($< 5\text{\AA}$)
- le distanze interatomiche rilevate, la conoscenza della sequenza, e dei vincoli conformazionali permette di derivare la struttura proteica



- 2D-NMR generalmente rivela le strutture solo di piccole proteine nel loro stato dinamico (in soluzione)

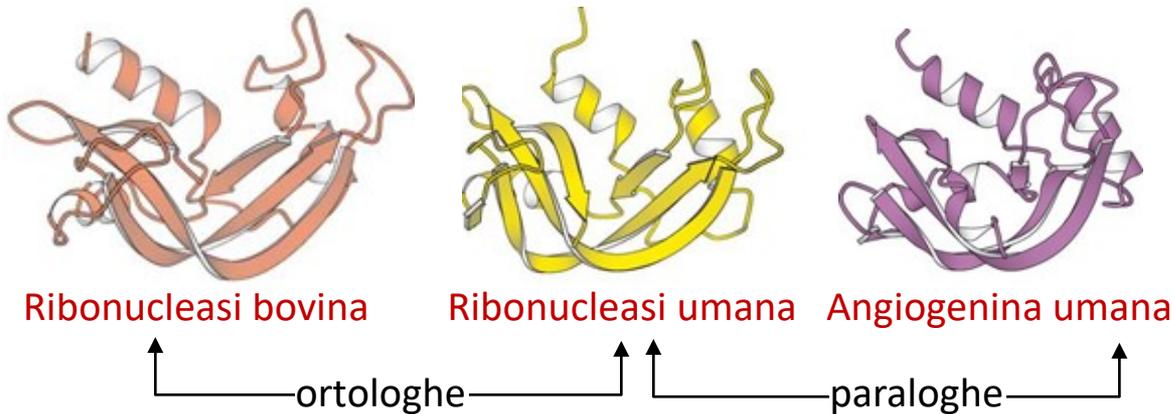
- ▶ La microscopia elettronica risale al 1931 e permette di ottenere una risoluzione di ~10 nm
- la cryo-EM è una forma di EM a trasmissione dove singole proteine sono immobilizzate in uno strato di ghiaccio vitrificato (amorfo, che ne preserva la struttura).
- l'immagine ad elevata risoluzione risulta dalla ricostruzione computazionale di molte immagini prese con diversi orientamenti



Studio dell'evoluzione molecolare delle proteine

► La sequenza amminoacidica di una proteina può fornire importanti informazioni

- il confronto con proteine note fornisce informazioni sull'origine, e possibile struttura e funzione; è importante perché è molto più facile conoscere la struttura 1^a che quella 3^o
- nell'**evoluzione molecolare** si cerca di comprendere come le proteine si sono trasformate strutturalmente e funzionalmente nel tempo



- le ribonucleasi e l'angiogenina hanno strutture simili ma solo il 35% degli AA della sequenza in comune.
- derivano da un precursore comune?
- hanno mantenuto la stessa funzione?

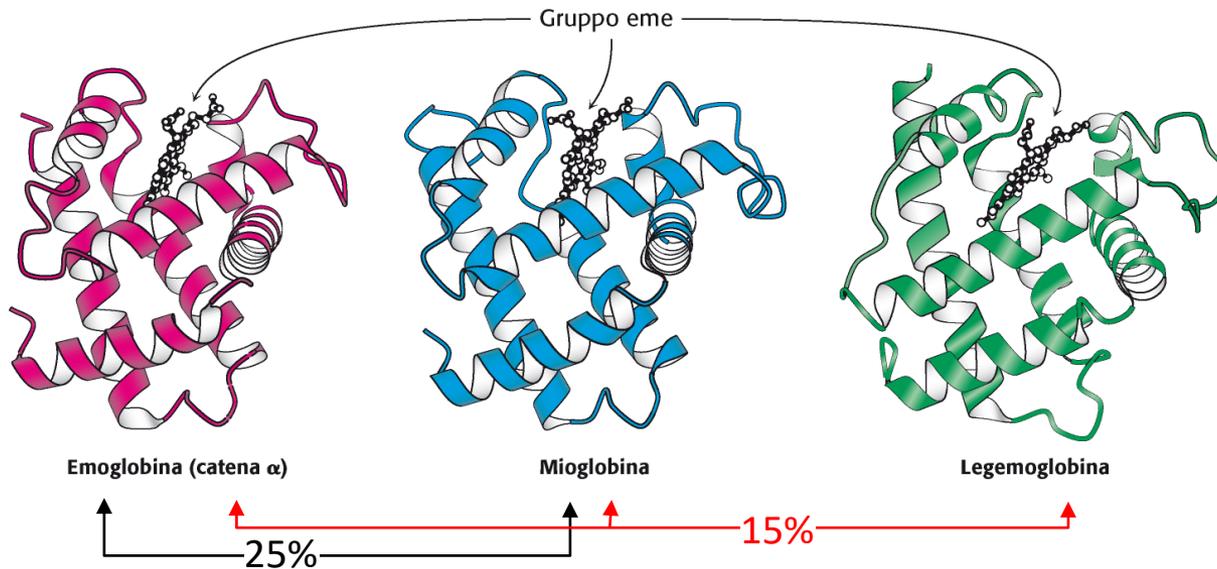
► Relazioni evolutive fra proteine

- due proteine sono **omologhe** se derivano da un progenitore ancestrale comune; questo si manifesta con una **significativa** somiglianza nella sequenza e struttura
- possono essere **ortologhe**: codificate da geni equivalenti in specie diverse
- paraloghe**: codificate da geni che si sono duplicati in una specie e che spesso hanno funzioni diverse

Rapporti evolutivi fra proteine

► Struttura tridimensionale e rapporti evolutivi tra proteine

- la *sequenza determina la struttura*, ma è la *struttura che determina la funzione*
- nell'evoluzione molecolare, la **struttura 3^a è più conservata che quella 1^a**
- La selezione naturale può agire in modo che la sequenza vari notevolmente senza alterare marcatamente la struttura e di conseguenza neanche la funzione



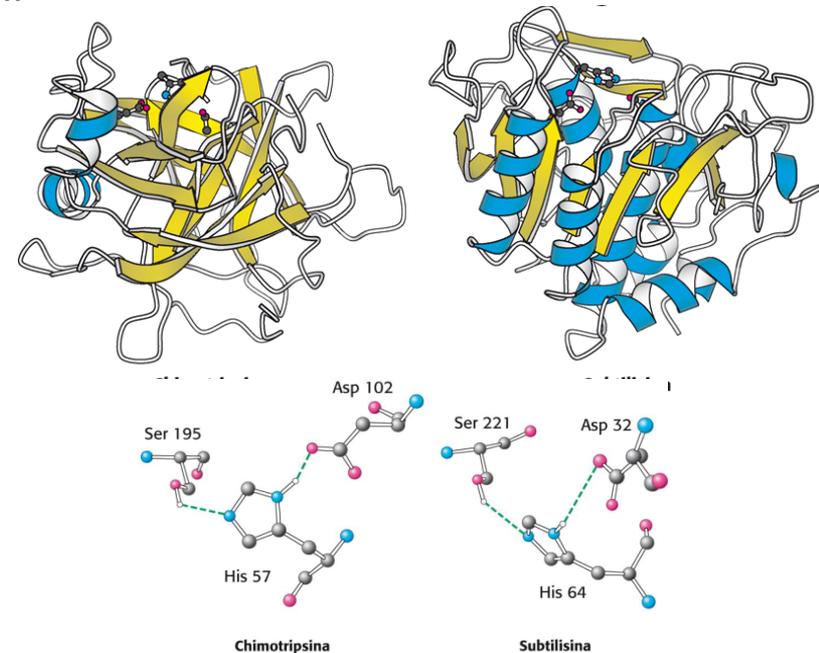
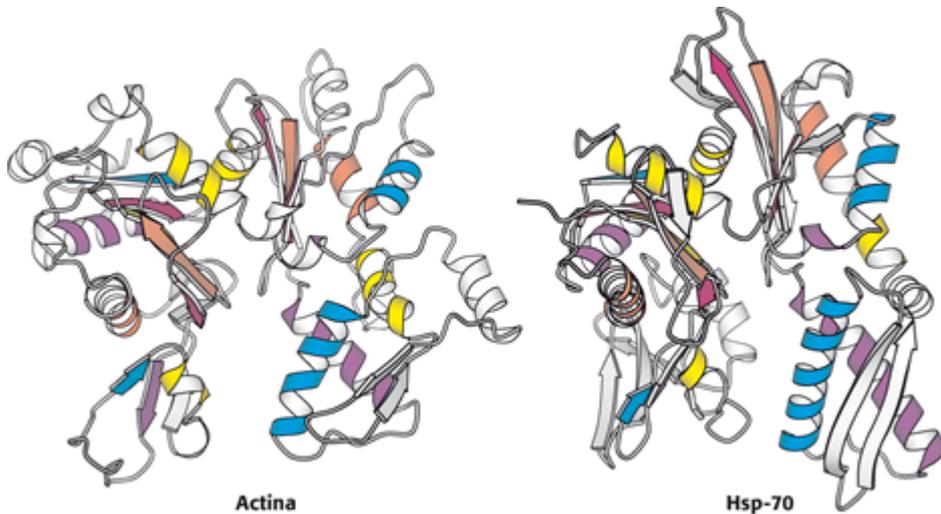
► Esempio:

- la sequenza della mioglobina e dell'emoglobina sono molto diverse da quelle della leghemoglobina delle leguminose (solo il 15% di AA in comune)
- le loro strutture terziarie sono simili e anche la loro funzione, che è quella di legare l'ossigeno.

Rapporti evolutivi fra proteine (cont.)

► Evoluzione divergente

- la struttura può rimanere simile anche in proteine con funzioni diverse
- actina e HSP-70 hanno strutture abbastanza simili anche se le sequenze hanno solo ~16% di amminoacidi in comune, e presumibilmente hanno geni paraloghi
- hanno un precursore comune ma diverse funzioni



► Evoluzione convergente:

- Ci sono casi di proteine con struttura molto diversa ma funzione molto simile
- la sequenza della proteasi chimotripsina (animale) non è omologa a quella della subtilisina (batterica) ma i residui del sito catalitico e la funzione sono simili
- l'evoluzione ha trovato due volte, indipendentemente, la stessa soluzione funzionale