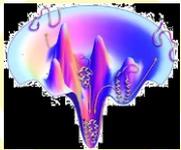


Modulo A1 Biochimica II



Modulo A1 Biochimica II

Ripiegamento e struttura delle proteine

A. Tossi atossi@units.it

2023-24

Metodi per studiare la struttura di proteine

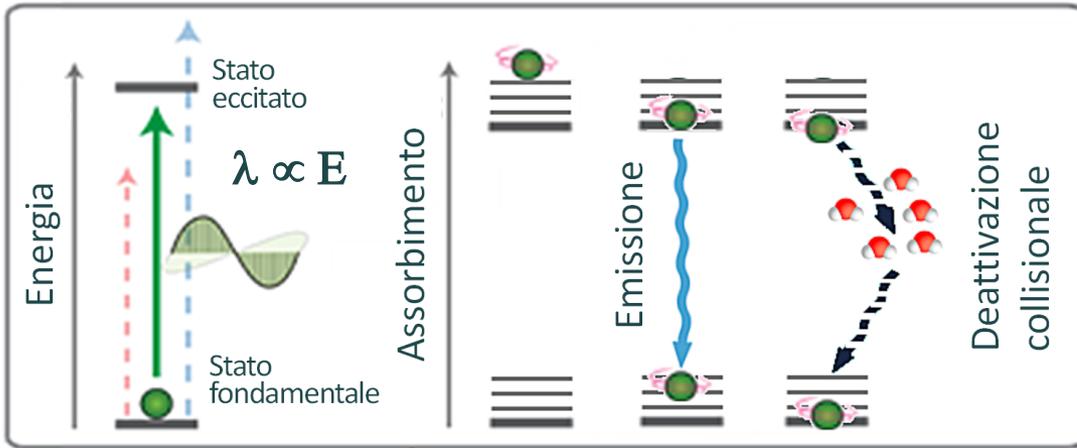
- ▶ Diversi metodi spettrometrici, spettrofotometrici, microscopici, chimici e computazionali permettono di sondare i vari livelli strutturali delle proteine

Livello di struttura	Metodo
Struttura primaria	Degradazione di Edman; Spettrometria di massa (MS; <i>peptide mapping, de novo sequencing</i>)
Struttura secondaria	Dicroismo circolare (CD) Spettrofotometria IR (FTIR) <i>In silico</i> prediction (Alpha-Fold)
Folding (ripiegamento)	CD Spettrofluorimetria MS (<i>hydrogen/deuterium exchange</i>) Rapid mixing
Struttura terziaria	Cristallografia Raggi-X, NMR, Criomicroscopia elettronica MS (<i>hydrogen/deuterium exchange</i>) <i>In silico</i> prediction (Alpha-Fold)
Struttura quaternaria	Cristallografia Raggi-X, Criomicroscopia elettronica MS (<i>hydrogen/deuterium exchange</i>) <i>In silico</i> prediction (Alpha-Fold)

METODI SPETTROSCOPICI: ASSORBIMENTO E FLUORESCENZA

Assorbimento della luce

Diagramma di livelli di energia



Legge di Lambert Beer

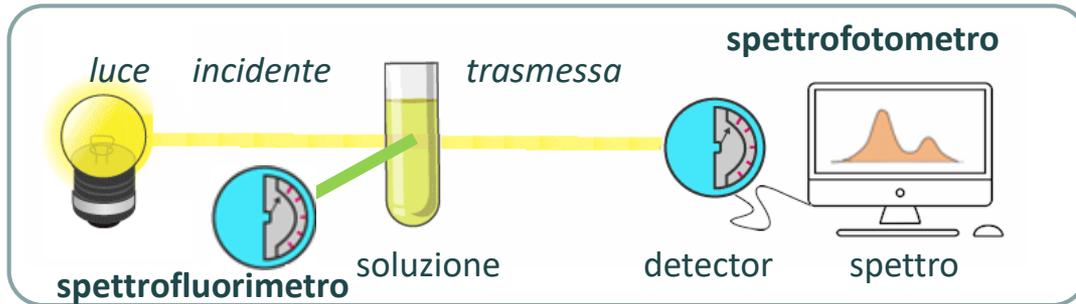
$$A = \epsilon l c$$

A = assorbanza

ϵ_λ = coefficiente d'estinzione

l = cammino ottico

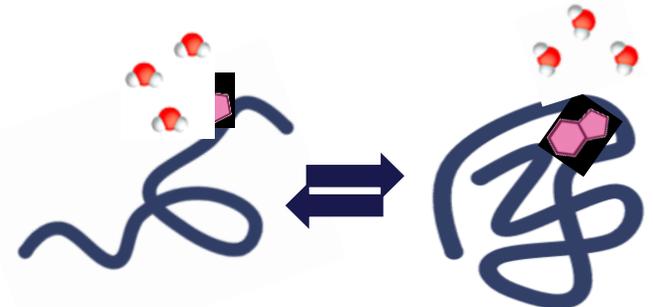
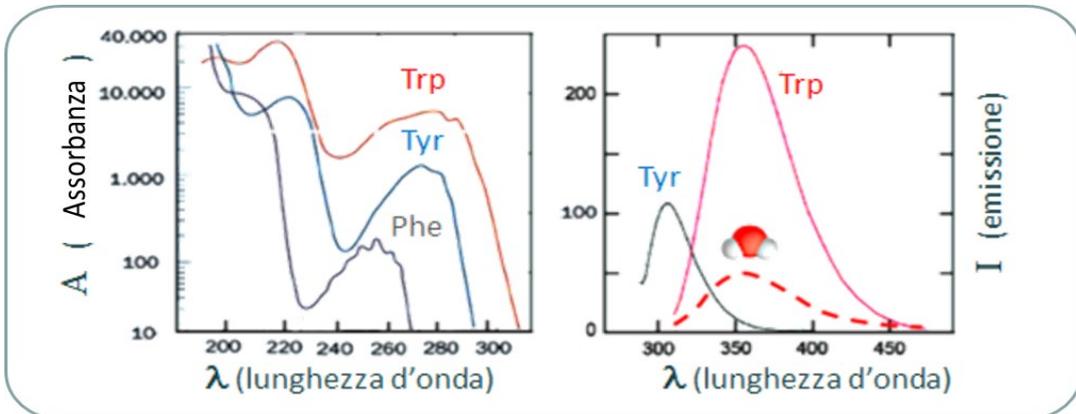
c = concentrazione



Spettroscopia:

1) Assorbanza → concentrazione

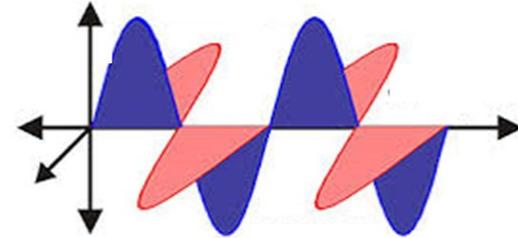
2) Emissione → ambiente



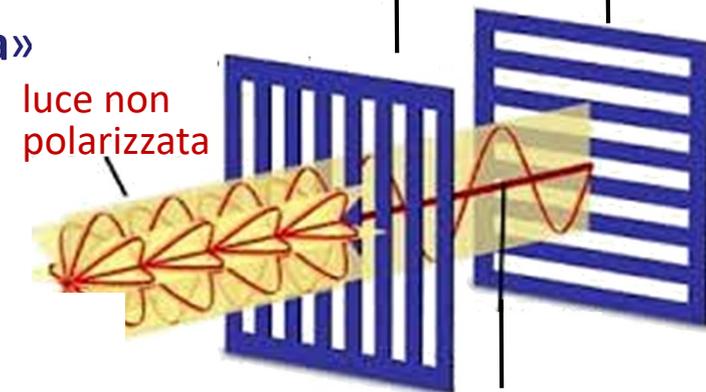
DETERMINAZIONE DELLA STRUTTURA 2^a: CD

► 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**

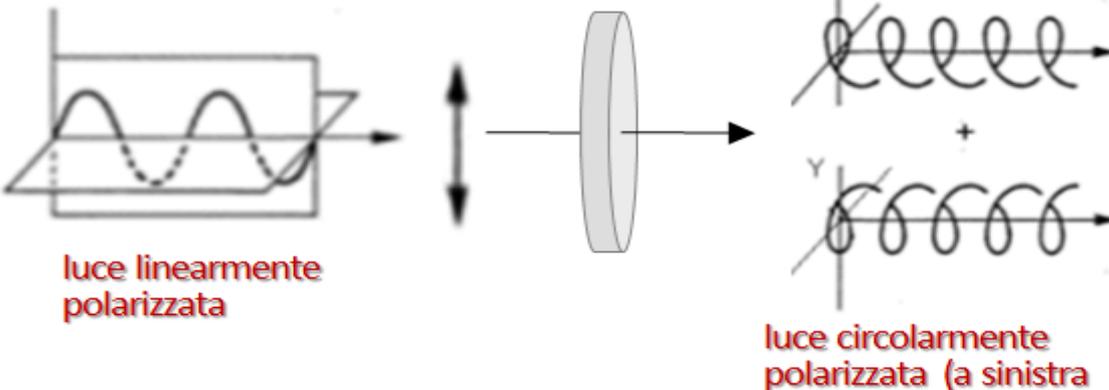


polarizzatore verticale e orizzontale



luce non polarizzata

luce linearmente polarizzata

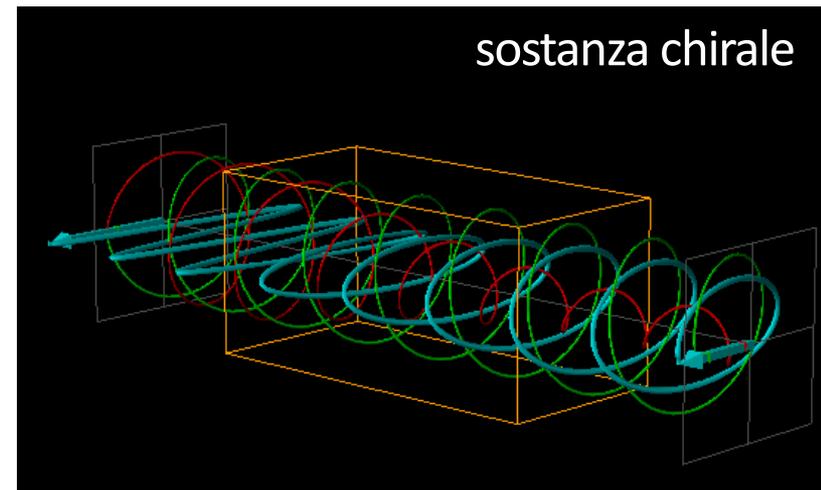
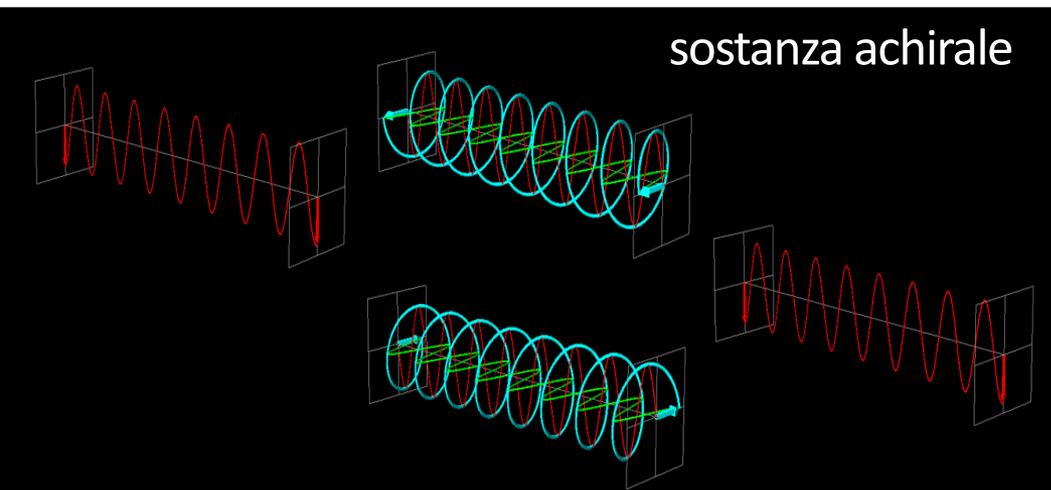
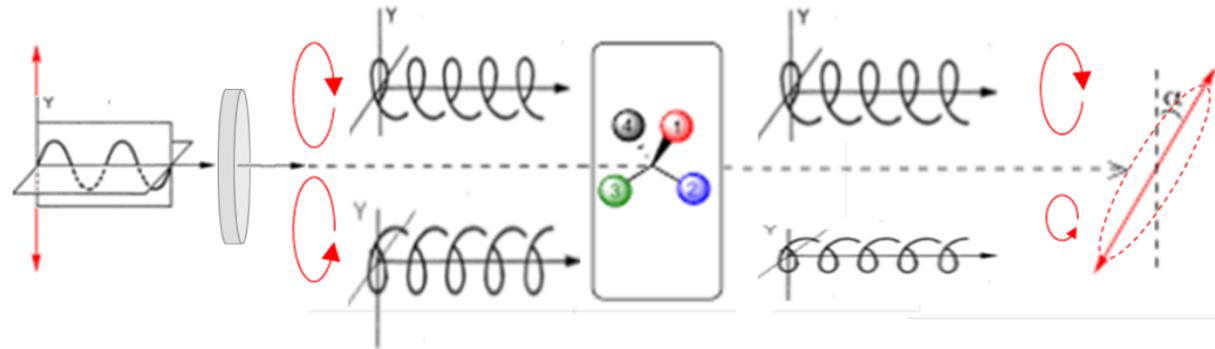


Utilizzo del dicroismo circolare

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

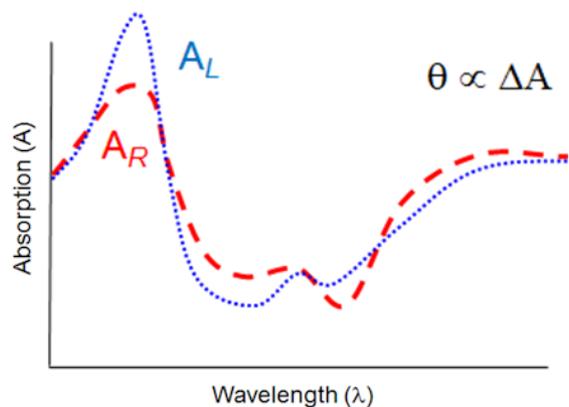
- $A = \varepsilon c l$ il coefficiente d'estinzione (ε) è diverso per le due componenti
- luce linearmente polarizzata che transita per una soluzione di sostanza chirale viene **ellitticamente polarizzata**. Il grado di ellitticità (θ) può essere misurato

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

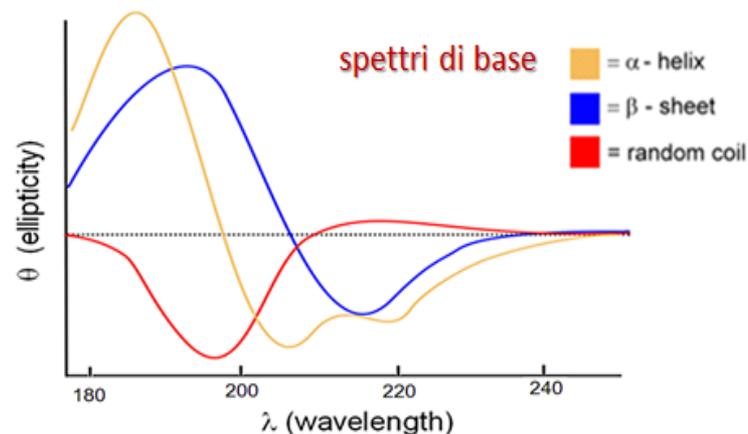


Utilizzo del dicroismo circolare (cont)

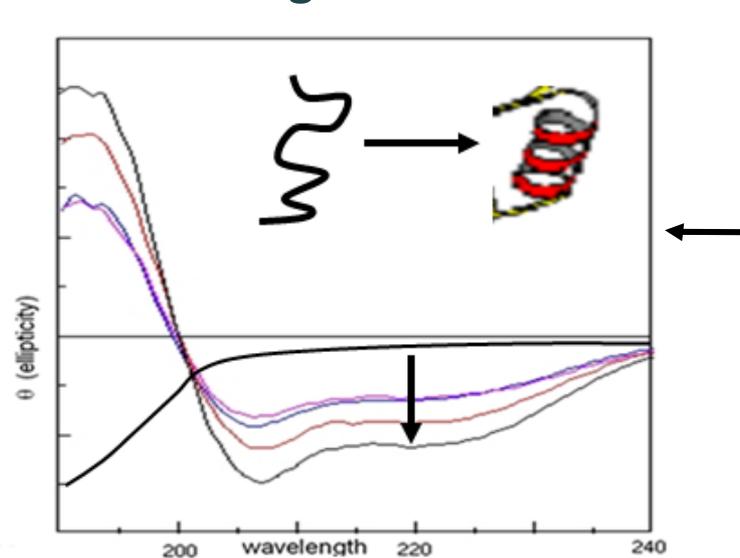
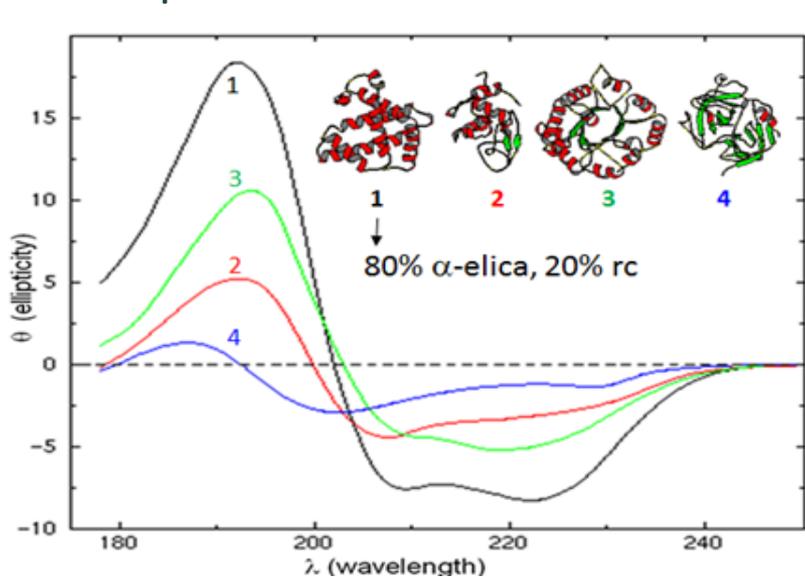
- Il θ può essere misurato e fornire uno spettro di «dicroismo circolare»
- nelle catene proteiche ci sono molti centri chirali posti in segmenti con **strutture 2^e a loro volta asimmetriche** – è questo che **determina lo forma dello spettro**



$$\epsilon_L^\lambda - \epsilon_R^\lambda$$
$$\theta \cong 3300 \Delta \epsilon$$

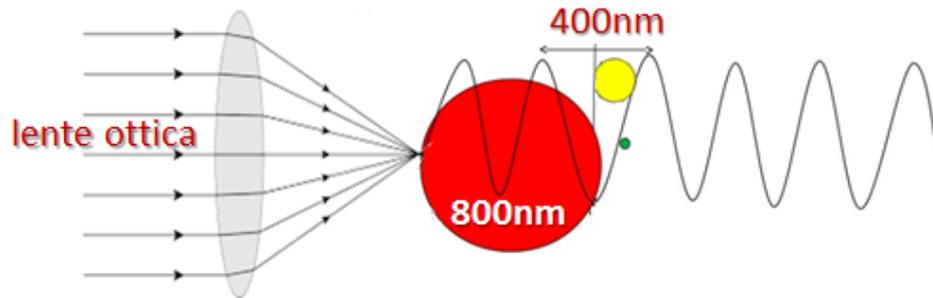


- serve per determinare il **contenuto conformazionale** e il **grado di strutturazione**

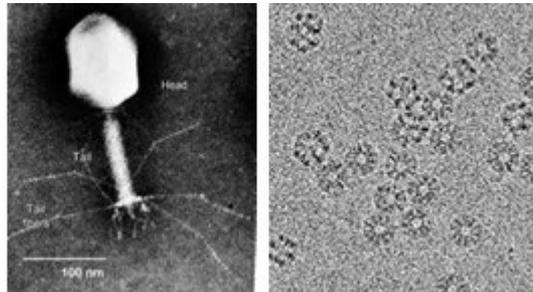


DETERMINAZIONE DELLE STRUTTURE 3^a e 4^a

- ▶ 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:

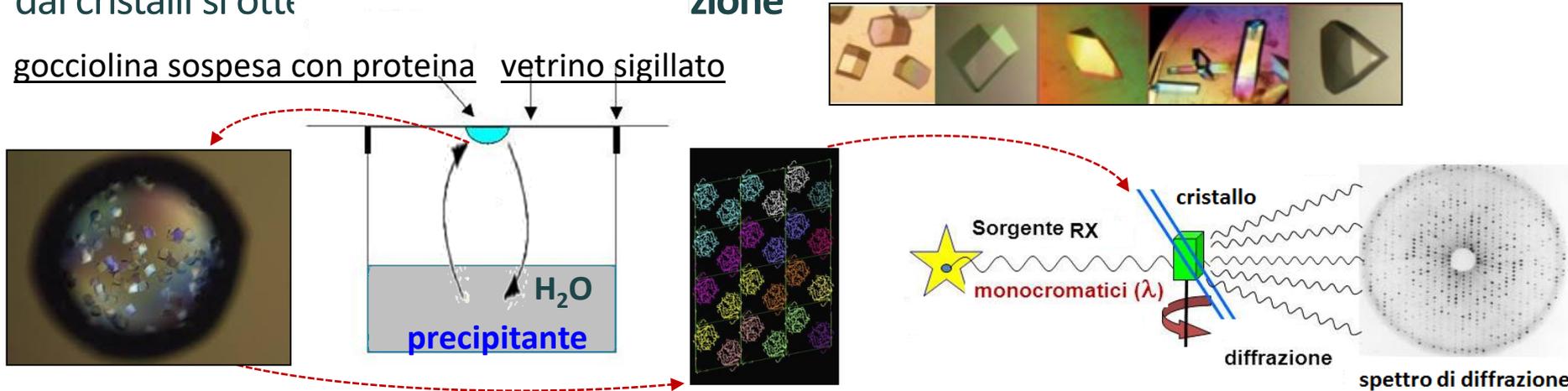
- spettroscopia a **risonanza magnetica nucleare** bidimensionale (2D-NMR)
- **microscopia crioelettronica (Cryo-EM)**



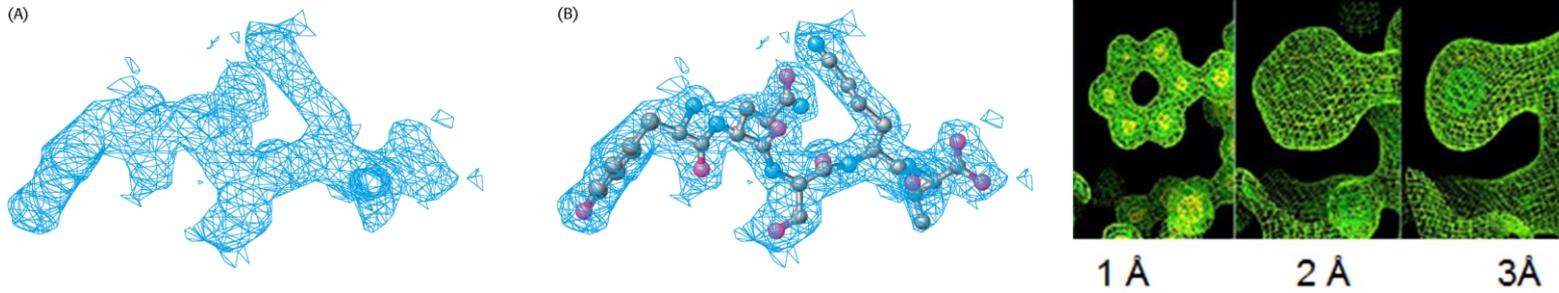
► Cristalli di proteine

- 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

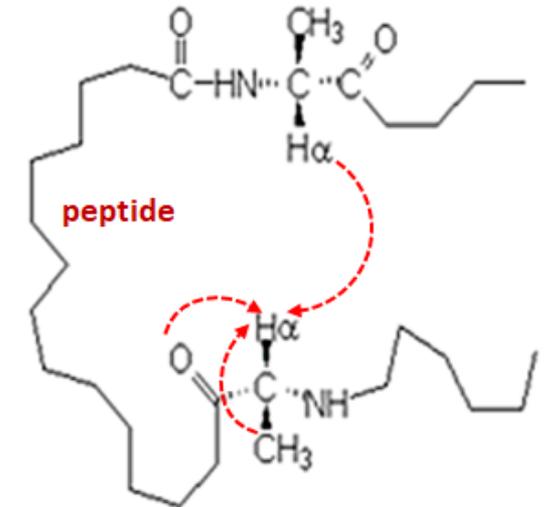
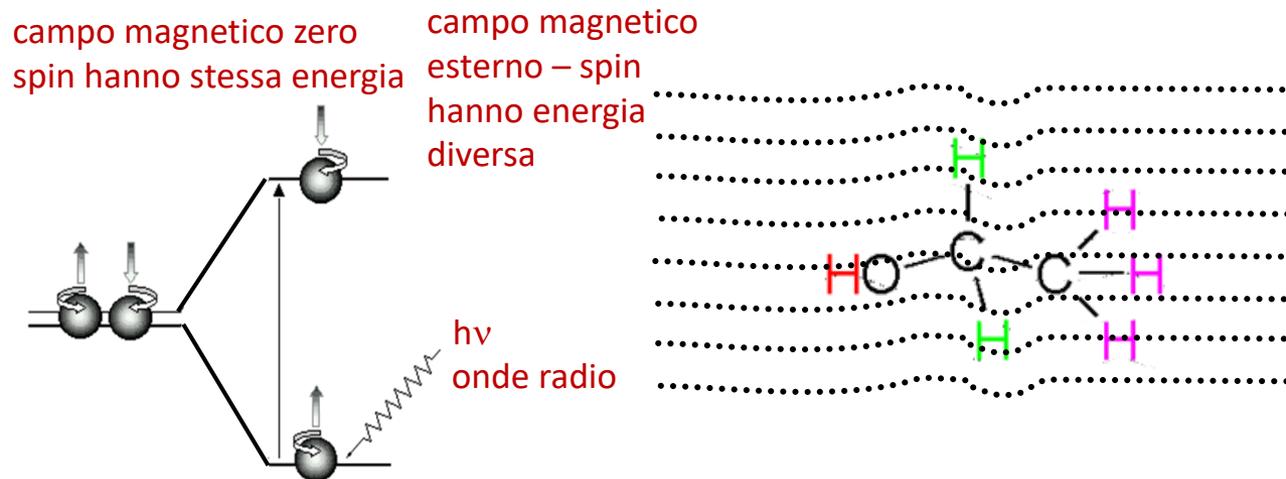


- la densità elettronica attorno agli atomi diffrange i raggi X; dalla mappa di diffrazione raccolta, si ottengono informazioni sulla densità elettronica e quindi sulla posizione degli atomi.
- migliore la qualità del cristallo e maggiore intensità e coerenza della luce, più elevata è la risoluzione



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

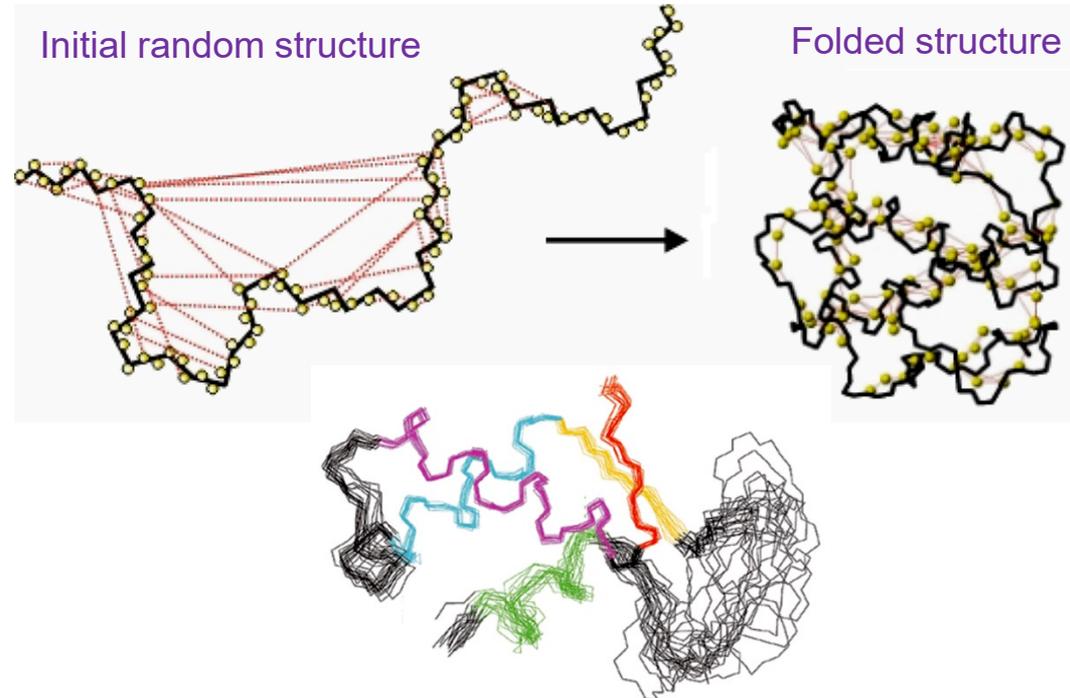
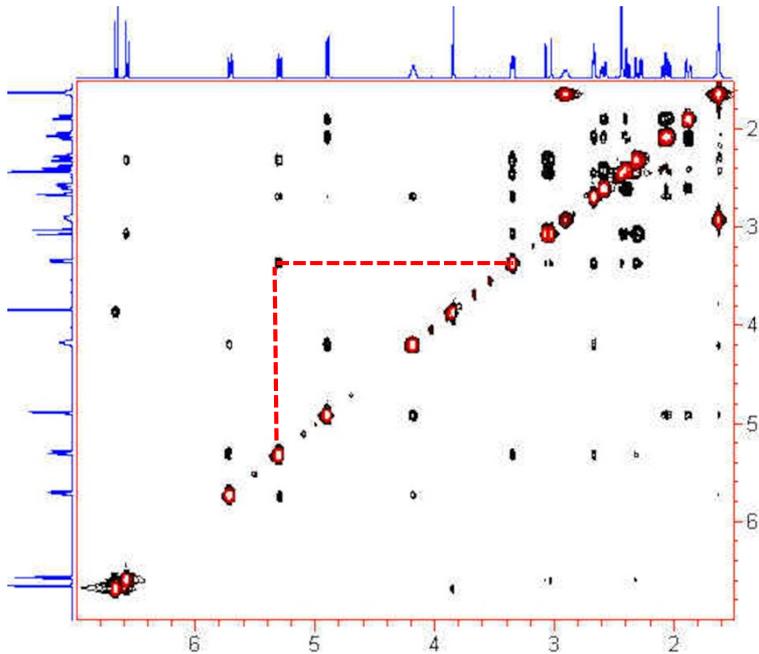
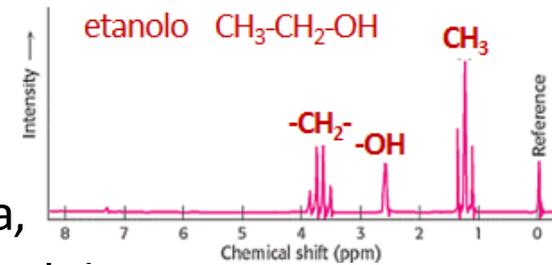
- si basa sulla transizione fra stati del *momento magnetico nucleare di spin* per alcuni isotopi di atomi presenti nelle biomolecole (^1H , ^{13}C , ^{15}N , ^{31}P)
- immersi in un forte campo magnetico esterno, gli stati energetici degli spin si separano
- quello meno energetico è favorito, ma assorbendo 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.)

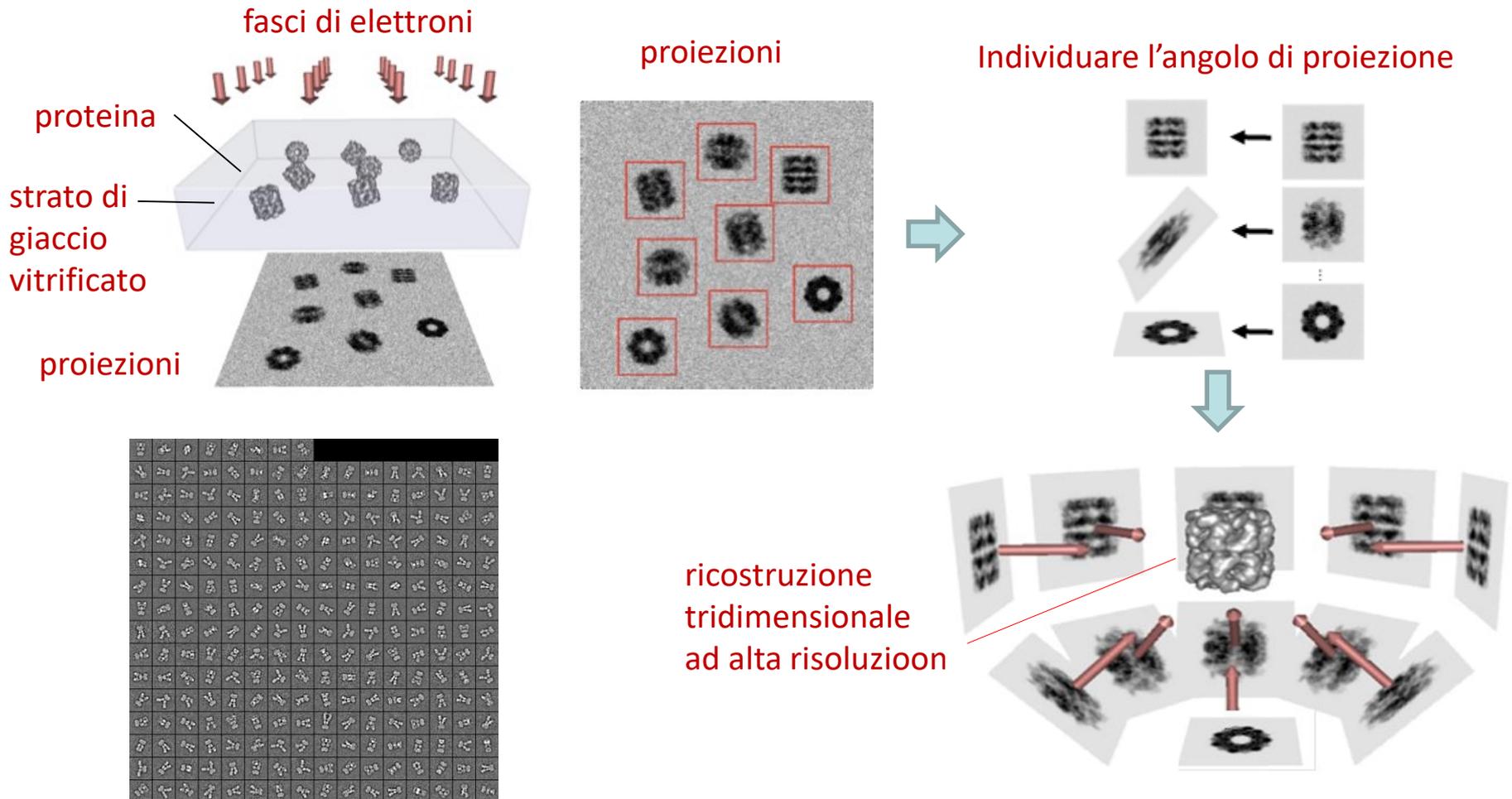
► $^1\text{H-NMR}$ permette la risoluzione di strutture relativamente semplici (piccole biomolecole)

- 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 struttura solo di piccole proteine nel loro stato dinamico (in soluzione)

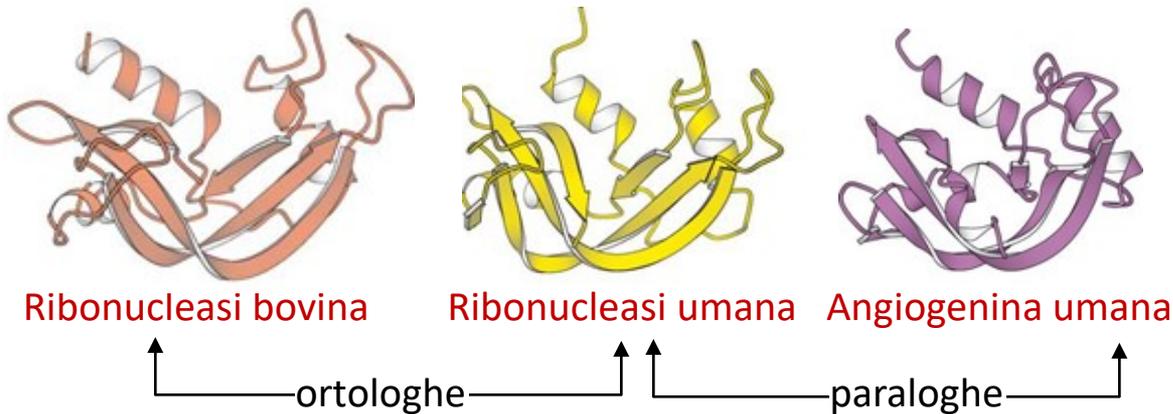
- ▶ La microscopia elettronica risale al 1931 e permette di ottenere una risoluzione di ~ 10 nm
- la cryo-EM è una forma di EM a trasmissione dove singole proteine sono immobilizzate in uno strato di ghiaccio vitrificato (amorfo, che ne preserva la struttura).
- l'immagine ad elevata risoluzione risulta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

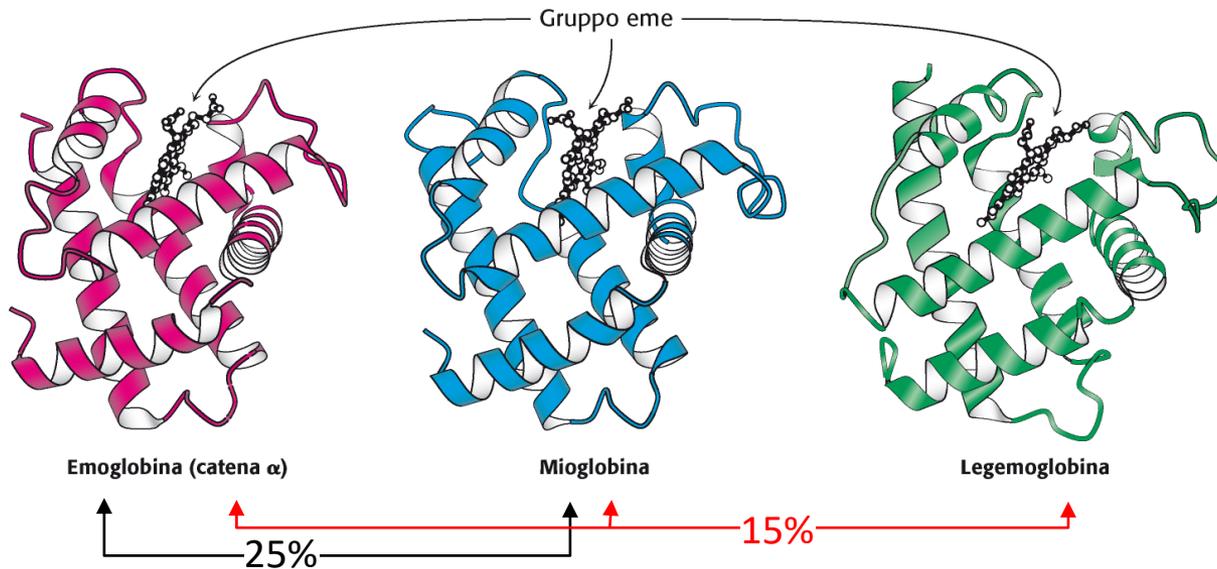
► Relazioni evolutive fra proteine

- 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



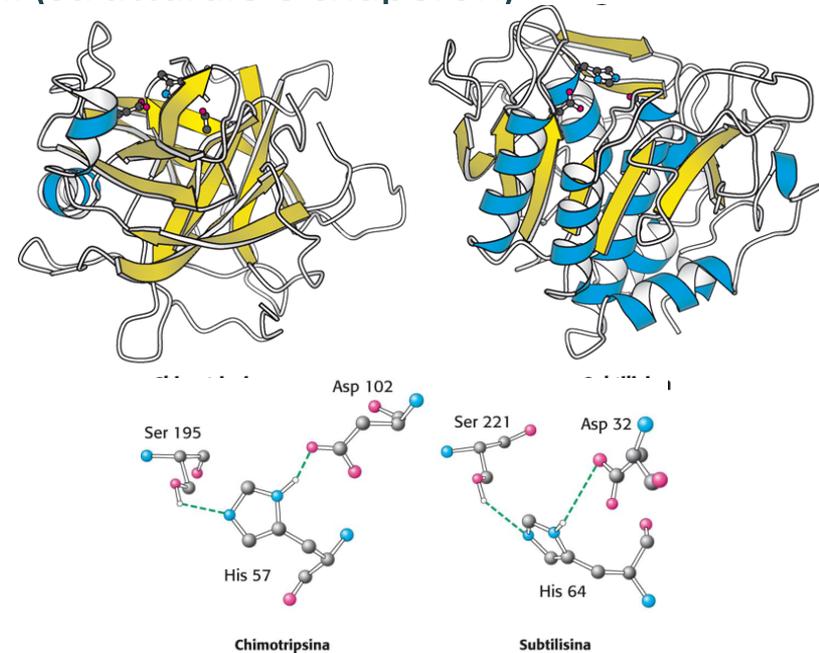
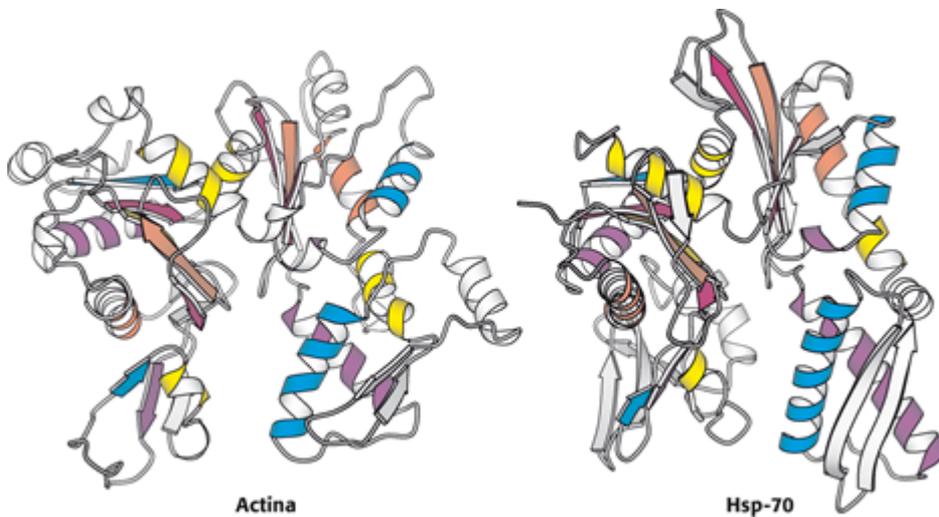
► Esempio:

- 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 la 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 (strutturale e chaperon)



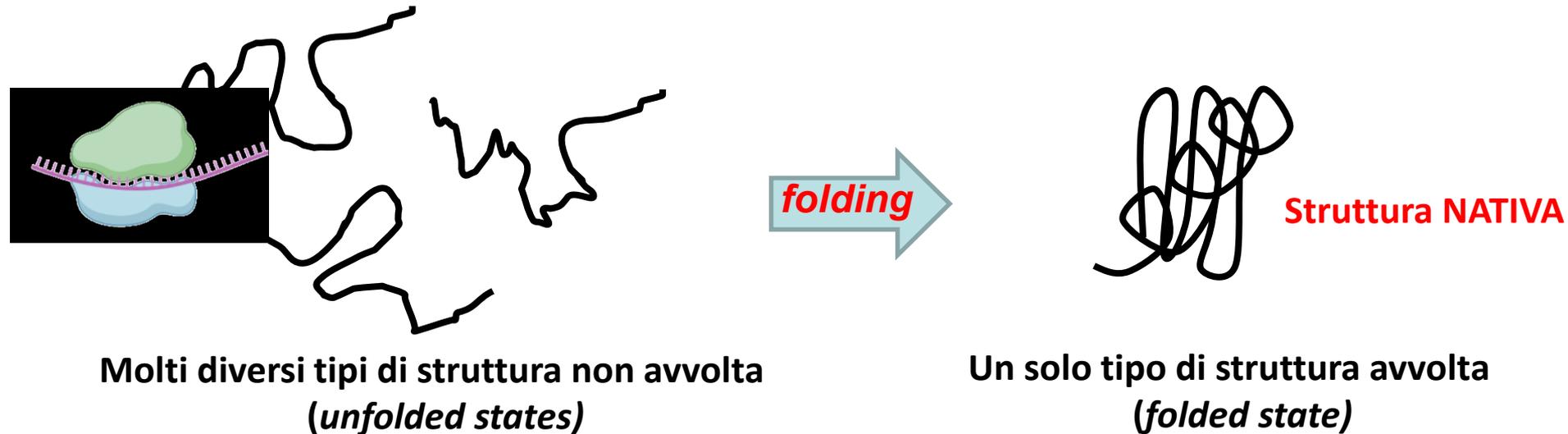
► Evoluzione convergente:

- Ci sono casi di proteine con struttura molto diversa ma funzione molto simile
- la sequenza della serina 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

RIPIEGAMENTO E STRUTTURAZIONE DELLE PROTEINE

► Principi della strutturazione proteica

- la **sequenza determina la struttura**
- la **struttura determina la funzione**
- esiste (generalmente) **solo una struttura nativa**
- la struttura dipende da **interazioni deboli** (non covalenti: elettrostatiche, idrofobiche, VdW, legami-H)
- la struttura finale crea il **massimo numero di legami-H** possibile (che dipendono da una **precisa distanza ed orientamento**)
- la struttura finale è quella con l'**energia interna minore** (è la più stabile)
- strutture native diverse hanno **componenti strutturali comuni** (α -eliche, β -strutture, β o γ -turn ecc.) organizzati in topologie ricorrenti (es. β -forcine, α - β - α , α -loop- α , ecc)



PARADOSSO DI LEVINTHAL

- ▶ **DOGMA CENTRALE:** “ La sequenza determina la struttura “ (Anfinsen)
- consideriamo una catena di 100 residui, se permettiamo 3 conformazioni/residuo (molto conservativo) → **3^{100} possibili strutture**
- se la catena **saggia ogni possibile conformazione** per trovare il minimo energetico, **restando 10^{-13} s** in ciascuna conformazione (tempo medio di rotazione attorno ad un legame) richiederebbe **$3^{100} \times 10^{-13} \text{ s} = 3 \times 10^{88} \text{ s} = 10^{27} \text{ anni}$ (> vita dell'universo)**
- con 10 possibili conformazioni (più realistico) → **10^{77} anni**
- è chiaro che quando le proteine si ripiegano **non saggia ogni possibile conformazione**
- Per evitare di farlo, una caratteristica essenziale del processo di ripiegamento è la **conservazione di elementi parzialmente corretti**
- questo riduce il tempo necessario per il ripiegamento a millisecondi/secondi

CONSERVAZIONE DI ELEMENTI PARZIALMENTE CORRETTI

```
200 ?T(\G{+s x[A.N5~, #ATxSGpn`e00
400 oDr'Jh7s DFR:W4]'u+^v6zpJse0i
600 e2ih'8zs n527x8l8d_ih=Hldseb.
800 S#dh>)/s )tZqC%lP%DK<|!^ase2.
1000 V0th>nLs ut/isjl_kwojjwMasef.
1200 juth+nvs it is(lukh?SCw-ase5.
1400 Iithdn4s it is0l/ks/IxwLase~.
1600 M?thinrs it is lXk?T"_woasel.
1800 MStthinWs it is lwkN7Qkw{asel.
2000 Mhthin`s it is likv,aww_asel.
2200 MMthinns it is lik+5avwlasel.
2400 MethinXs it is likydaqw)asel.
2600 Methin4s it is lik2dasweasel.
2800 MethinHs it is likeQaTweasel.
2883 Methinks it is like a weasel.
```

```
200 )z~hg)W4({cu!kO{d6jS!NlEyUx}p
400 "W hi\kR.<&CfA%4-Y1G!iT$6({|6
600 .L=hinkm4 (uMGP^lAWoE6klwW=yiS
800 AthinkaPa_vYH liR\Hb,Uo4\-"(
1000 OFthinksP)@fZO li8v] /+Eln26B
1200 6ithinksMvt -V likm+gl#K~)BFk
1400 vxthinksaEt Qw like.SlGeutks.
1600 :Othinks<it MC likesN2[eaVe4.
1800 uxthinksqit Or likeQh)weaoeW.
2000 Y/thinks it id like7alwea)e&.
2200 Methinks it iW like a[weaWel.
2400 Methinks it is like a;weasel.
2431 Methinks it is like a weasel.
```

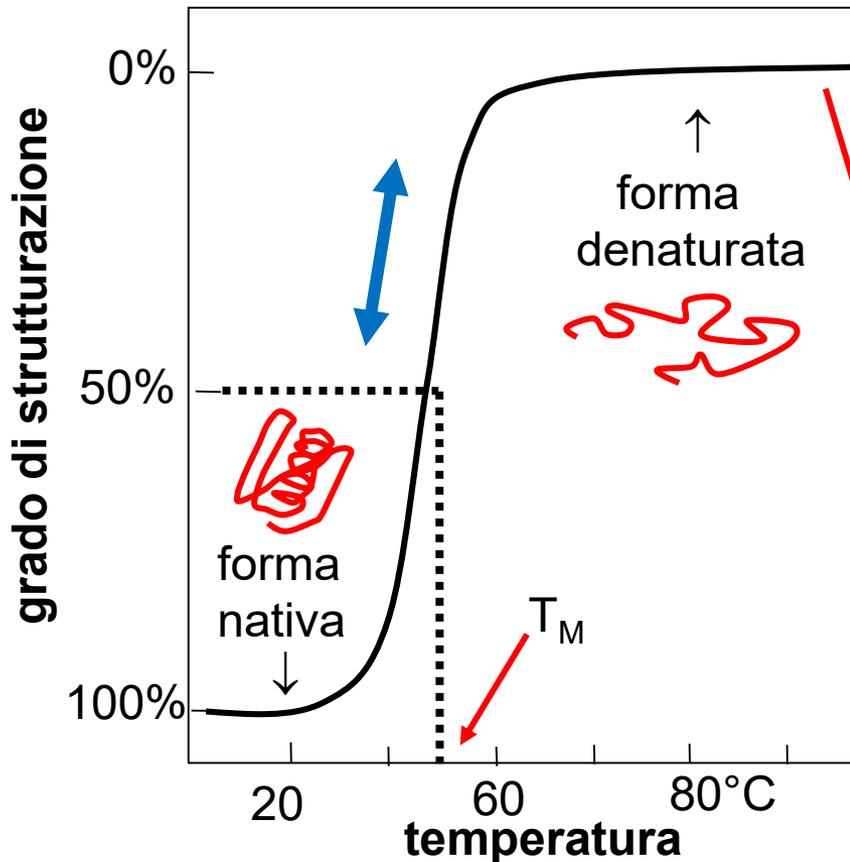
► Durante il ripiegamento della proteina:

- **ogni tappa** nel ripiegamento **porta ad una riduzione di energia**
- questa riduzione è **dell'ordine dell'energia termica** (può essere facilmente invertita)
- la **formazione di un segmento con conformazione corretta** influisce sulla formazione di un **successivo segmento a conformazione corretta**
- il ripiegamento è un **processo altamente cooperativo**

DENATURAZIONE / RINATURAZIONE DELLE PROTEINE

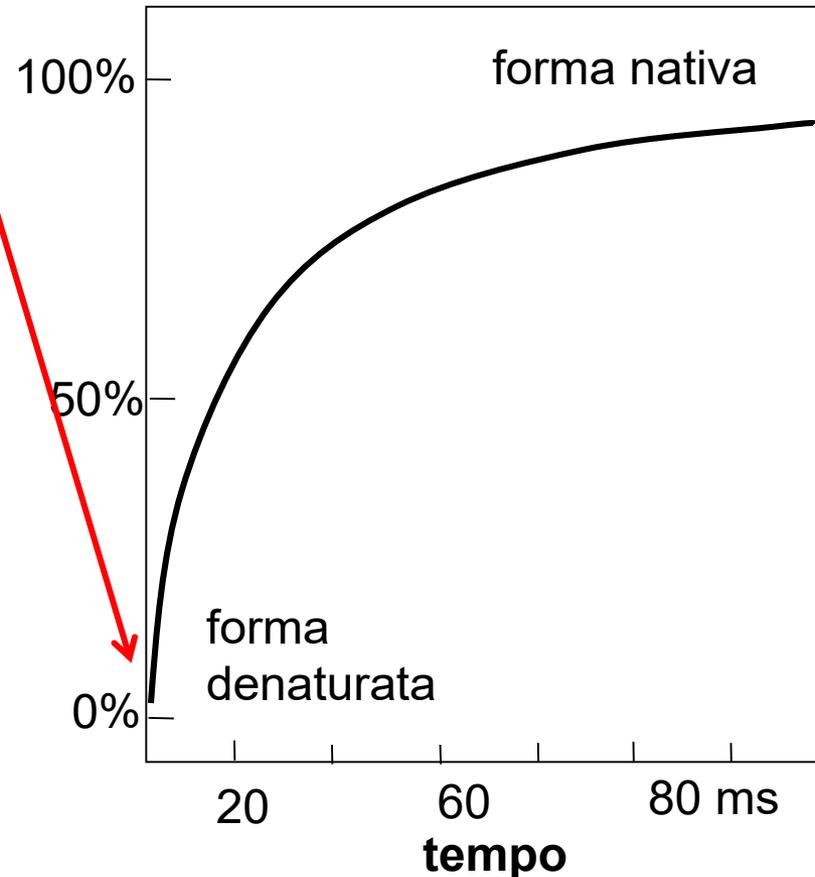
- Ripiegamento e la denaturazione delle proteine sono processi altamente cooperativi e reversibili

Curva di denaturazione



(oppure pH, [agente caotropico], ecc.)

Cinetica di ripiegamento (o di rinaturazione)



Riportare rapidamente in condizioni rinaturanti

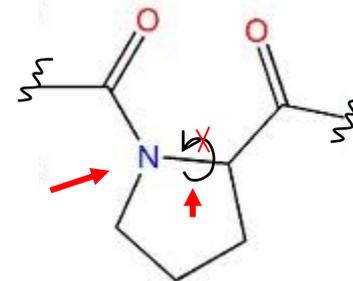
RIPIEGAMENTO: ruolo dei singoli amminoacidi

► I diversi amminoacidi hanno diverse tendenze intrinseche a formare strutture 2^{ie}

- **Ala, Leu, Arg, Gln** → favoriscono la formazione di α -elica
- Altri amminoacidi invece non la favoriscono per ragioni steriche:

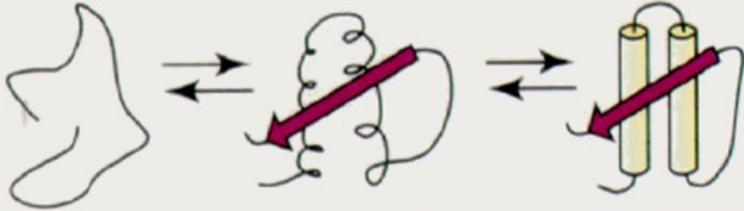


Pro → destabilizza α -elica e β -foglietto per ragioni sia **steriche (angolo torsionale bloccato)** sia per l'incapacità di formare legami-H con il gruppo imminico

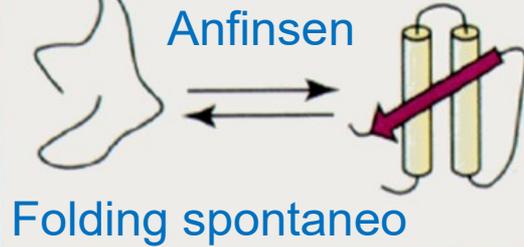
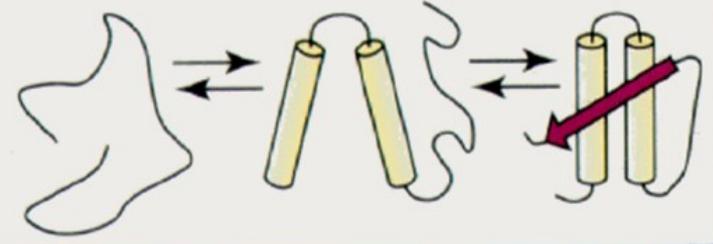


MODELLI PER IL RIPIEGAMENTO

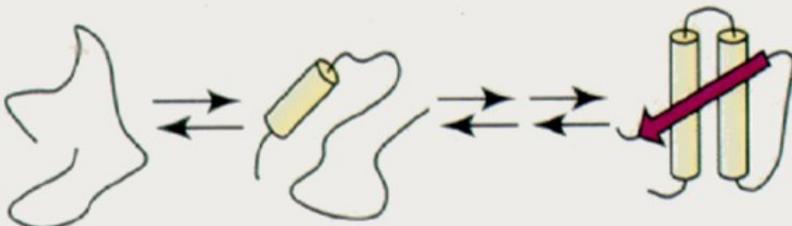
Collasso idrofobico
(globulo fuso)



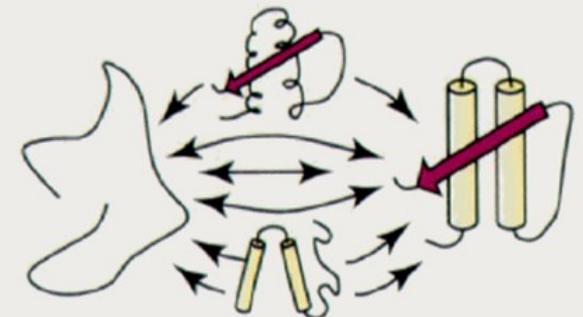
Modello gerarchico



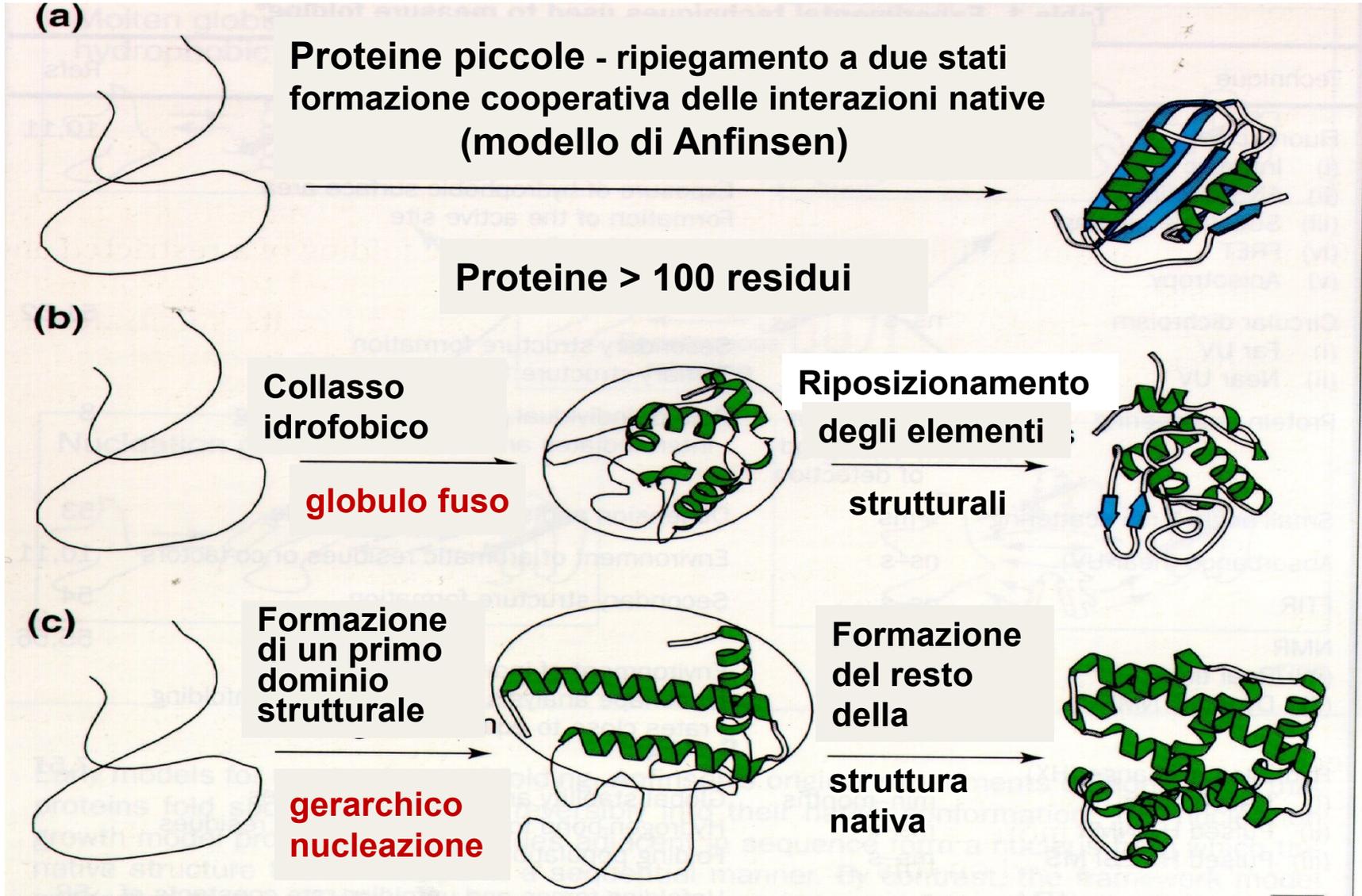
Nucleazione



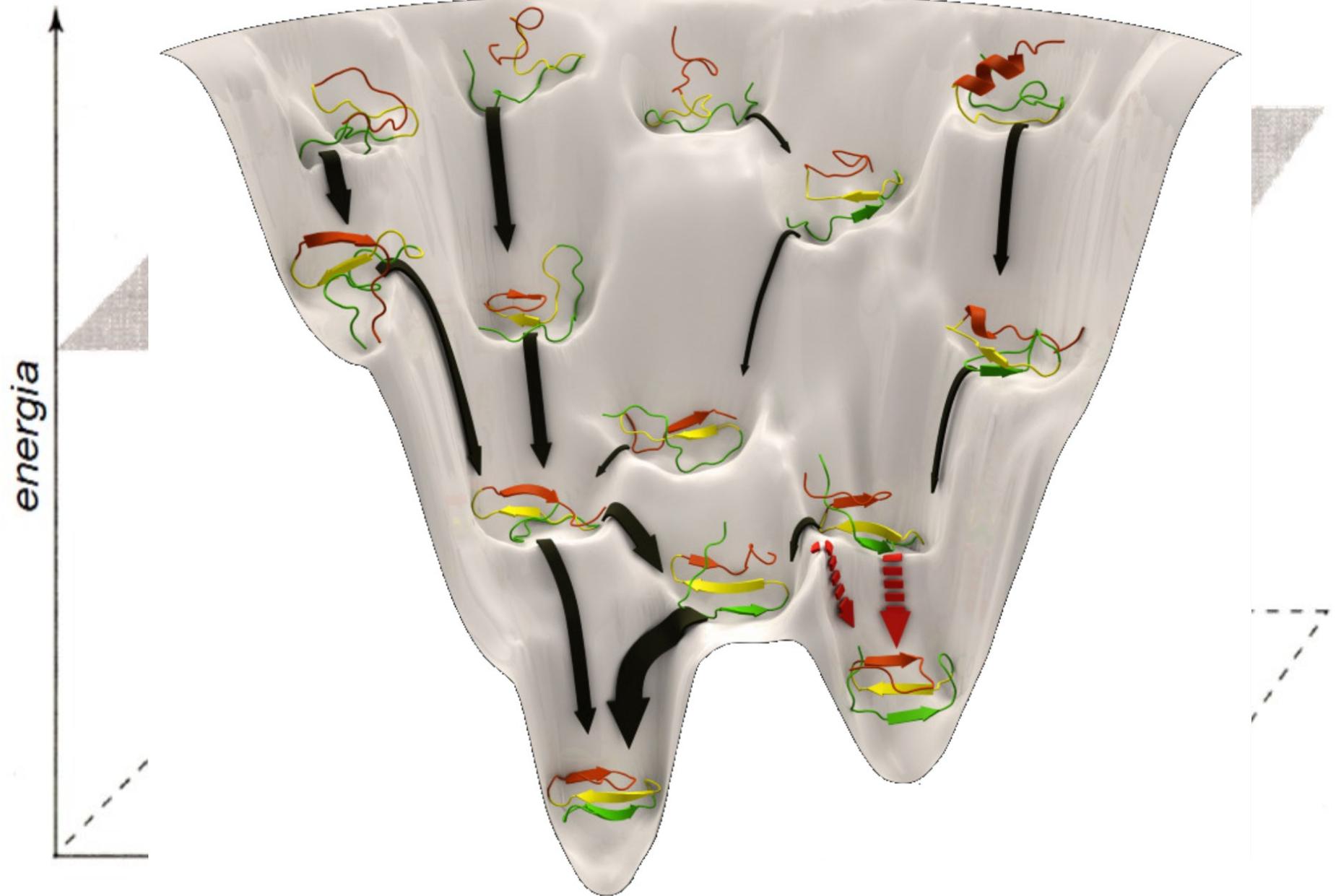
Modello multimodale



PROCESSI DI RIPIEGAMENTO

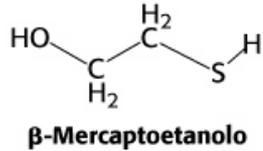
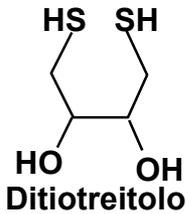


PANORAMA ENERGETICO PER IL RPIEGAMENTO

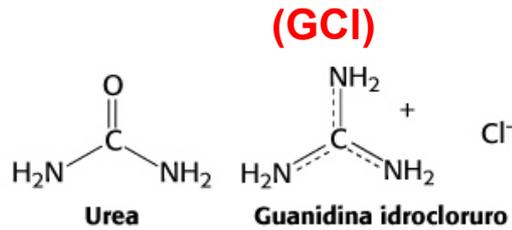


TECNICHE PER STUDIARE IL FOLDING

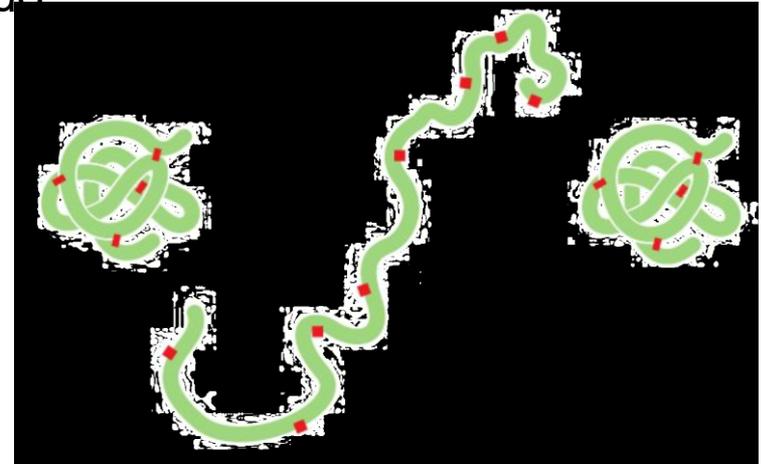
E' necessario partire da uno stato denaturato usando
AGENTI DENATURANTI



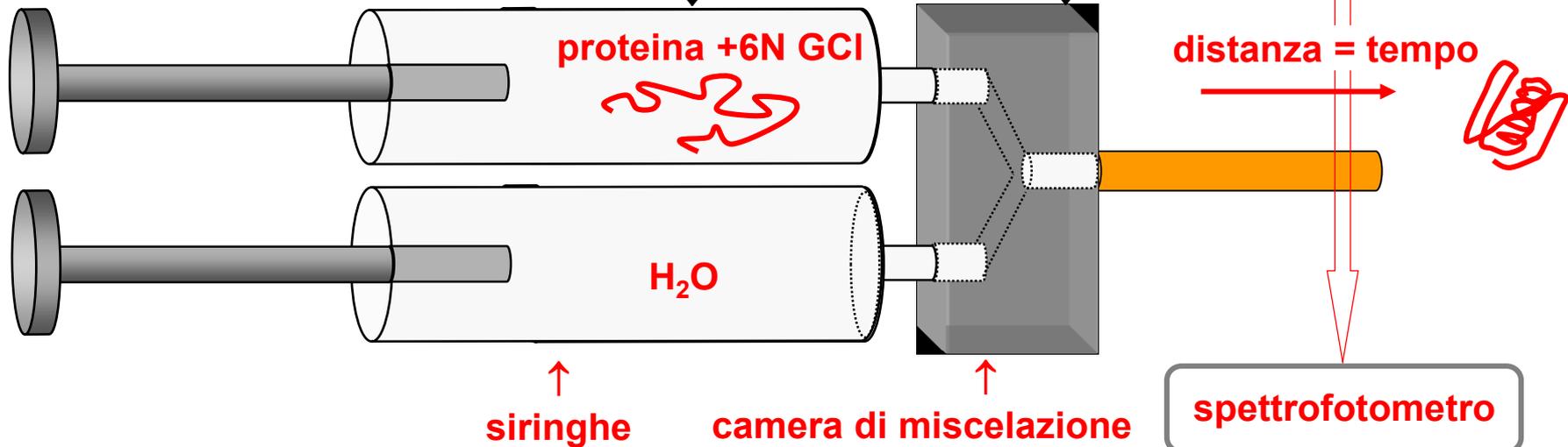
rompere legami S-S



rompere legami-H



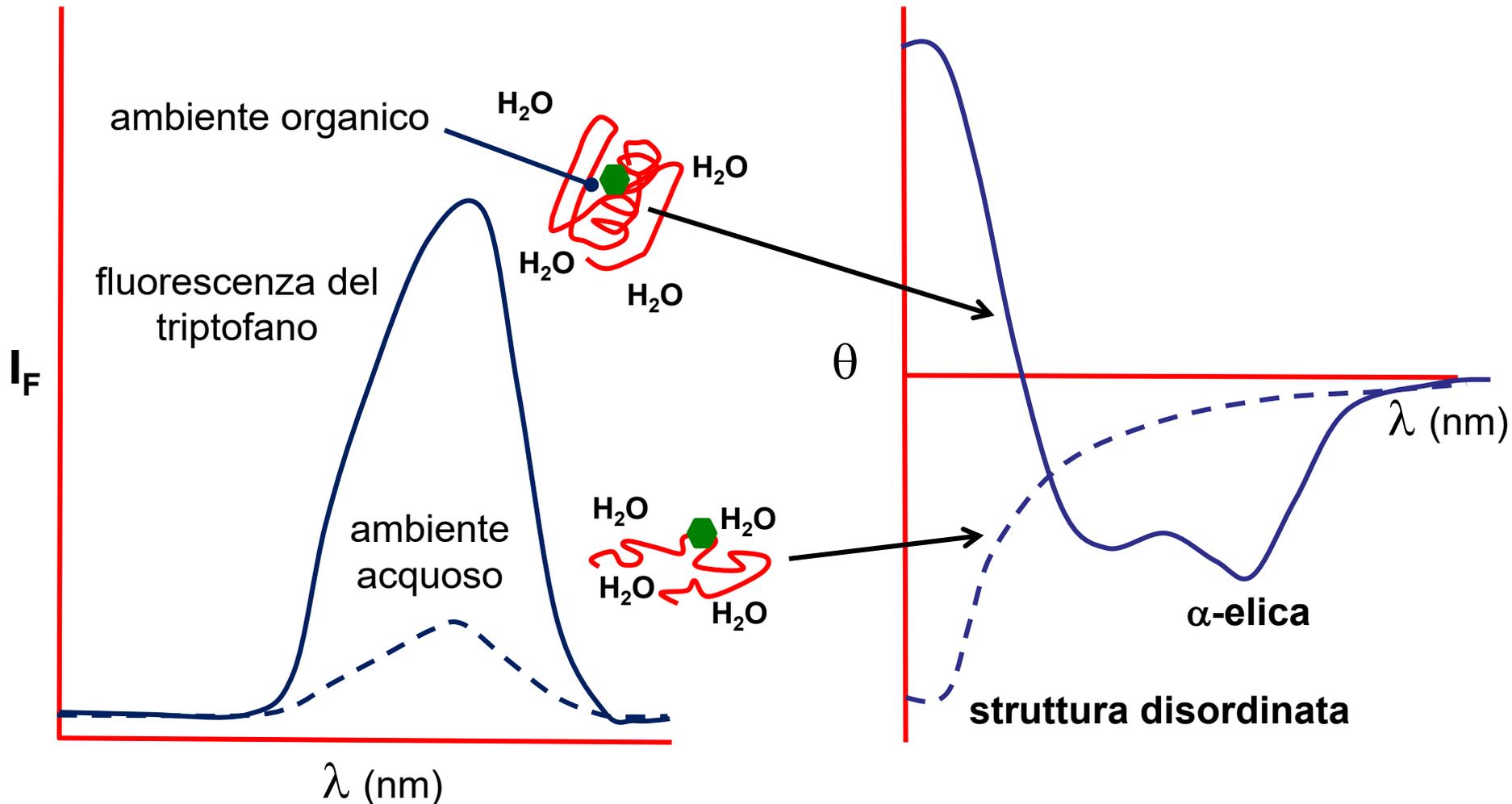
Apparato STOP-FLOW



TECNICHE PER STUDIARE IL FOLDING (cont.)

► Monitoraggio sperimentale del ripiegamento

- Il **dicroismo circolare** e la **spettroscopia a fluorescenza** sono tecniche utili per seguire i processi di ripiegamento:



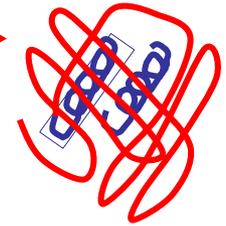
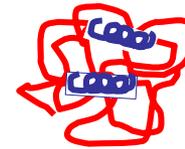
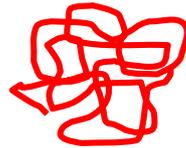
TECNICHE PER STUDIARE IL FOLDING (cont.)

random coil

molten globule

2^y structure

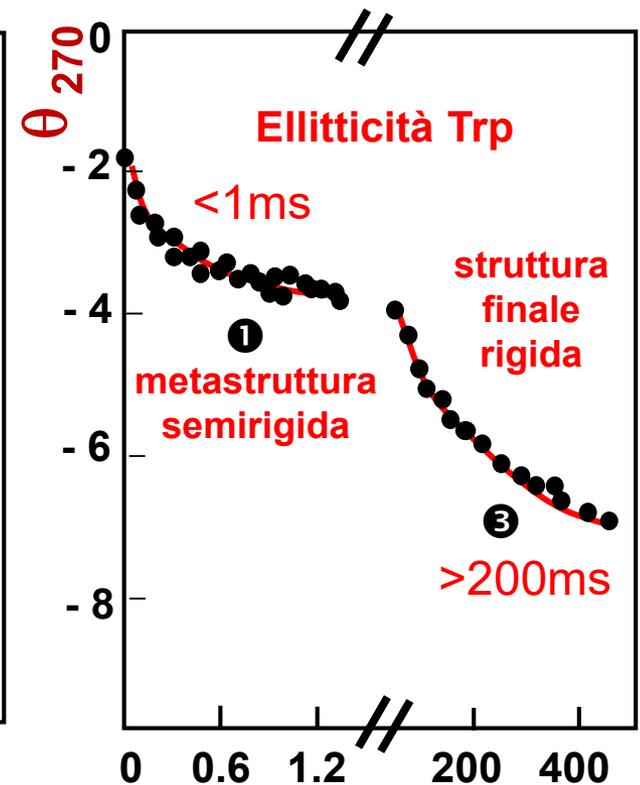
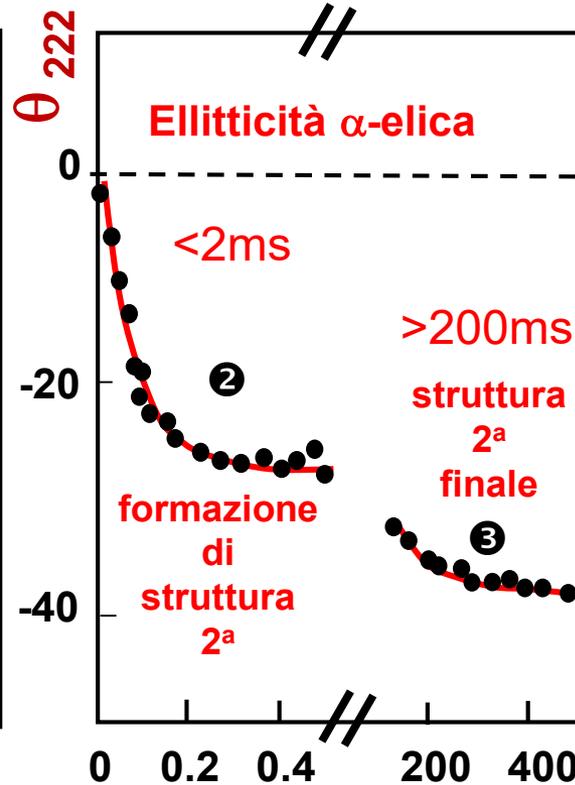
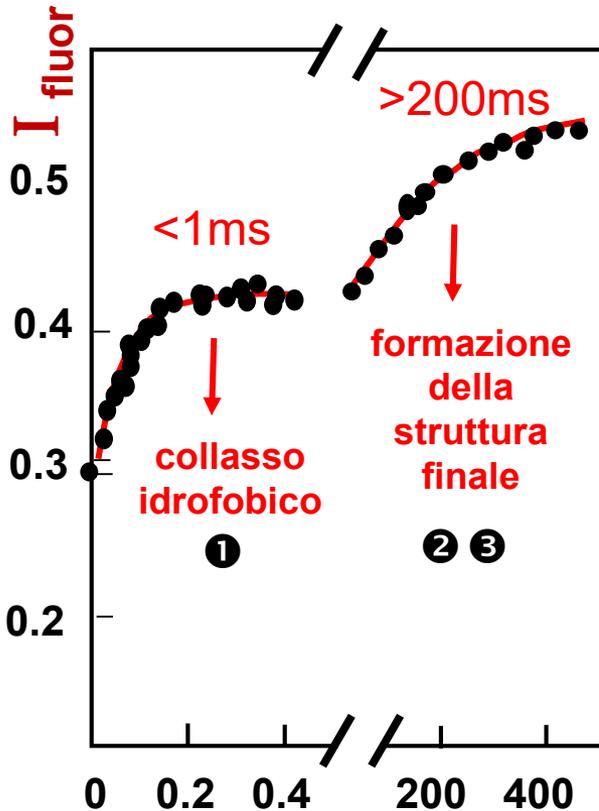
3^y structure



fluorescenza (Trp)

dicroismo circolare catena

dicroismo circolare Trp



tempo di ripiegamento (millisec)

► Le isomerasi velocizzano il ripiegamento

- sono enzimi che catalizzano **conversioni stereoisomeriche** a livello di singoli residui (attenzioni ! **NON** dirigono il folding)

PPIasi (Peptidilprolil isomerasi oppure *cis/trans isomerasi)*

- proteine **ubiquitarie e abbondanti**, con diversi ruoli.
- catalizzano la conversione fra le **configurazioni *cis* e *trans*** nei residui di prolina

PDIasi - (Proteindisolfuro isomerasi)

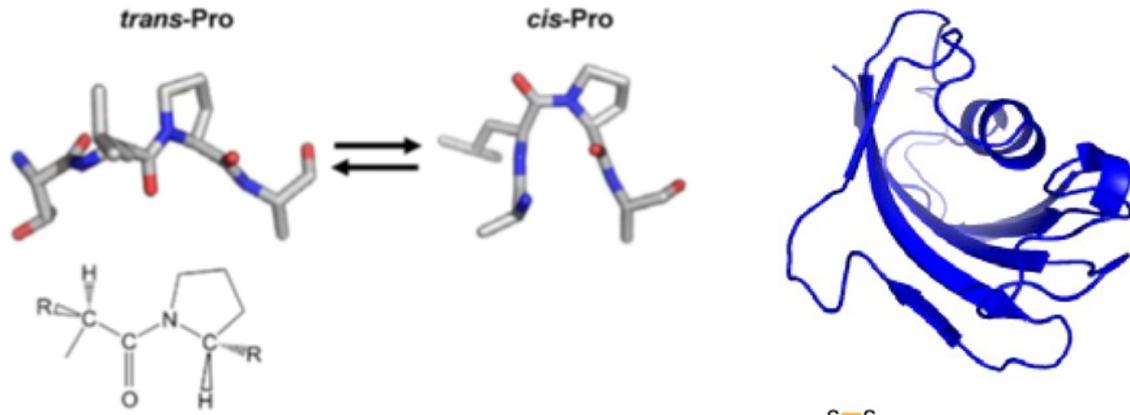
- **ponti disolfuro incorretti** sono una causa principale di **ripiegamento incorretto**
- le *PDIasi* catalizzano il **riarrangiamento di ponti disolfuro** all'interno di proteine.
- contengono **residui di cisteina reattivi** che possono sostituirsi temporaneamente ad una delle due cisteine in un ponte disolfuro, permettendo all'altra di formare un legame con un partner diverso.

► Le “chaperone” molecolari proteggono da interazioni dannose

- *In vitro* (es. negli esperimenti di Anfinsen con *RNAsi*), esiste una sola popolazione di polipeptidi identici in fase di ripiegamento (**condizioni lontane da quelle naturali**).
- *In vivo*, i polipeptidi si ripiegano in un ambiente **molto affollato**, con altre popolazioni in fase di ripiegamento. La possibilità di **interazioni illecite**, che portano ad aggregazione e precipitazione, sono elevate

PEPTIDIL PROLIL ISOMERASI e PROTEINA DISOLFURO ISOMERASI

PPlasi

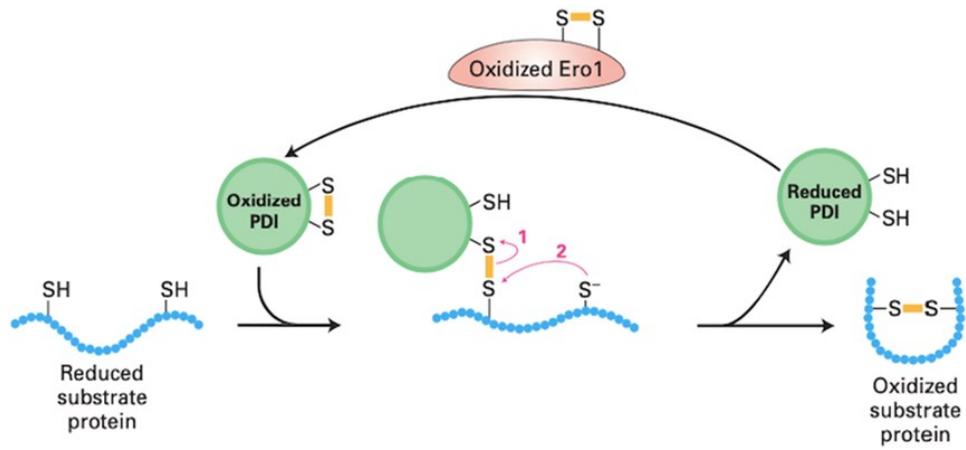


PDIasi

possono agire sia come

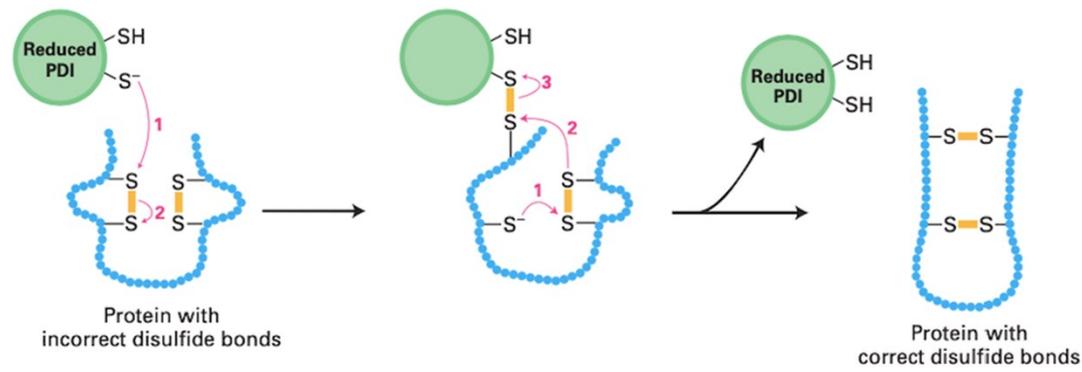
Ossidasi

per proteine nascenti



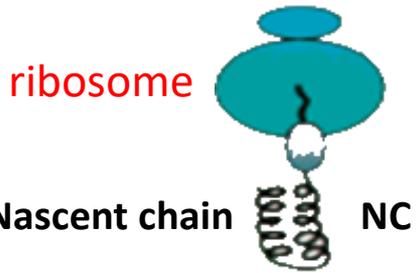
Isomerasi

per proteine misfolded



RIPIEGAMENTO ASSISTITO: Chaperone molecolari e chaperonine

Small proteins

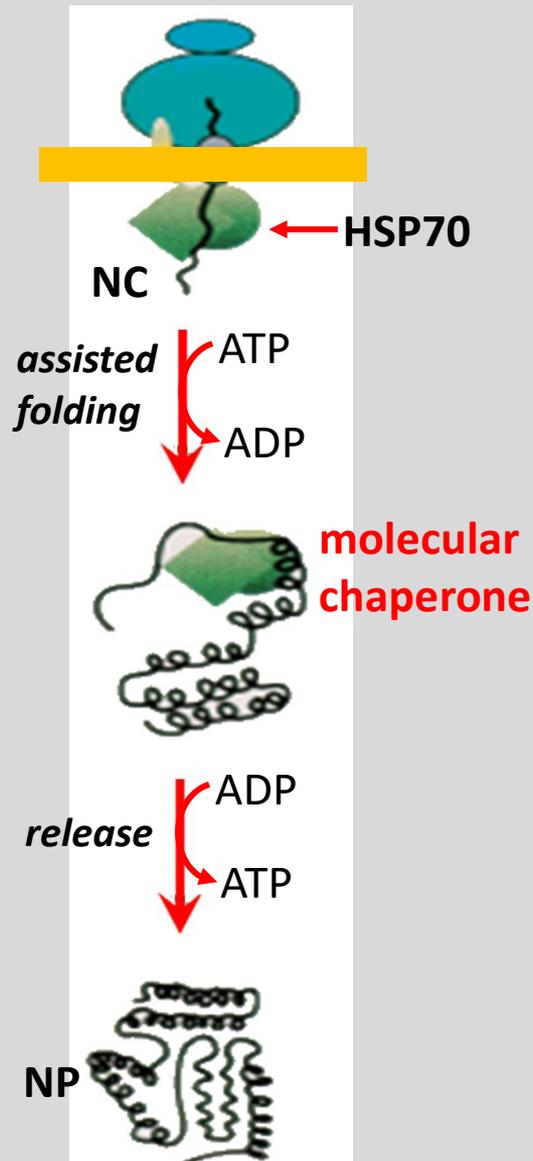


unassisted folding

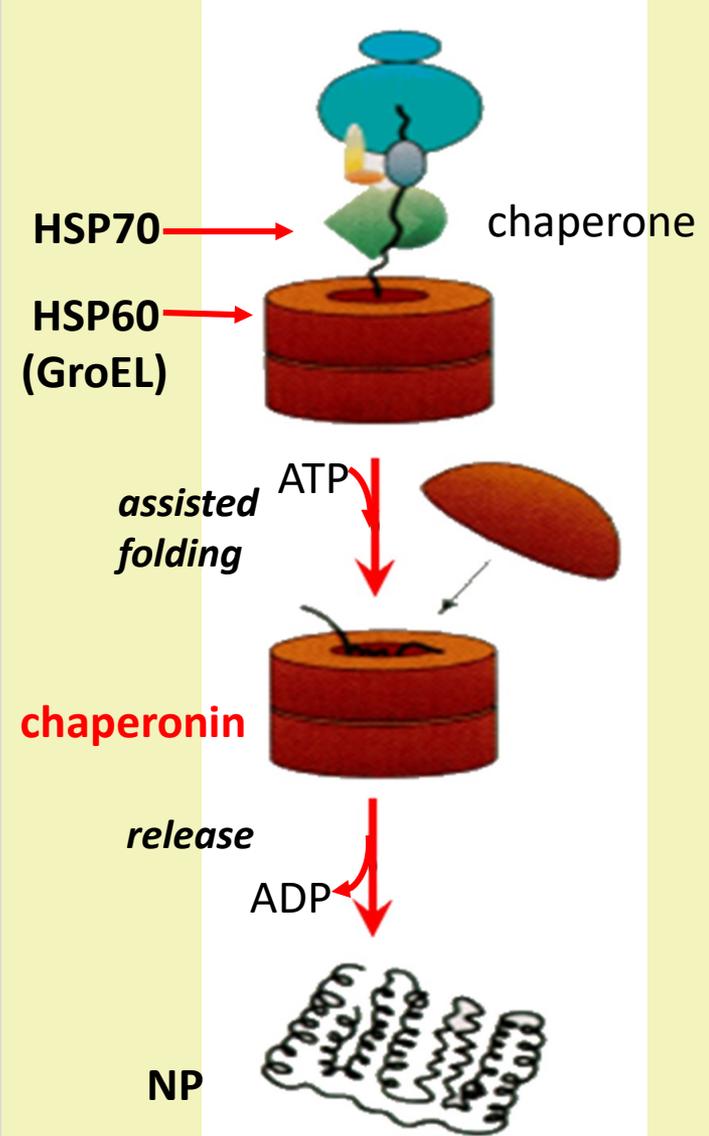


Native (folded) protein NP

ER



Cytoplasm & mitochondria (bacteria)



Meccanismo d'azione e ruoli delle CHAPERONINE

