

Lezione 9

Folding

Il folding è il processo di ripiegamento delle proteine.

Sappiamo che esiste una sequenzialità molto precisa che determina le proprietà della proteina. Alla base sta il DNA codificante per la proteina. La sequenza determina la struttura della proteina. La strutturazione della proteina nella sua forma tridimensionale (stato nativo) determina la sua funzione.

Per comprendere bene come avviene il folding dobbiamo capire quali sono le forze che intervengono nel far sì che la struttura nativa abbia una stabilità sufficiente a poter permanere in soluzione ed esplicare la sua funzione.

Interessante è la comprensione delle regole che governano il ripiegamento della proteina. Il folding è strettamente correlato alla sequenza, tanto che si parla di un secondo codice (oltre al codice genetico che associa la sequenza di DNA alla sequenza amminoacidica) che è l'insieme di interazioni e regole che partendo da una certa sequenza portano ad una certa struttura.

Si sa molto sul secondo codice, ma ancora non è del tutto compreso. Questo è un problema molto complesso, data la grande quantità di interazioni che vanno tenute in considerazione nel ripiegamento della proteina.

È comunque importante capire cosa porta una certa sequenza a ripiegarsi in un certo modo, sia per predizioni delle strutture 3D delle proteine, che soprattutto per il design *de novo* delle proteine.

Un altro argomento che tratteremo è la cinetica del folding.

Il ripiegamento comporta aspetti termodinamici ed aspetti cinetici, correlati al percorso che la catena polipeptidica fa nel corso del suo ripiegamento.

Ci occuperemo sia dell'aggregazione tra proteine che della formazione di corpi di inclusione.

Ci chiediamo anche se una proteina deve anche avere una struttura tridimensionale specifica per avere una funzione specifica.

Si è visto che in realtà ci sono proteine che non hanno una struttura tridimensionale definita, che chiamiamo *proteine intrinsecamente disordinate*.

Queste proteine non si riuscivano a cristallizzare perché avevano una struttura talmente dinamica da rendere impossibile la formazione dei cristalli. Queste proteine possono arrivare al 30% del contenuto della cellula, e la loro funzione biologica è strettamente correlata al fatto di non avere una struttura prefissata, ed al fatto che queste riescono ad adattarsi all'interazione con altre proteine.

Queste proteine sono particolarmente rappresentate tra quelle che intervengono nella trasduzione del segnale. Ci sono anche proteine intrinsecamente disordinate tra i virus, che ne sfruttano la versatilità ottimizzando così la quantità di materiale genetico trasportato.

Cosa serve conoscere?

- Proprietà degli aminoacidi e struttura di questi
- Legami deboli rilevanti per le strutture proteiche
- Strutture secondarie
- Basi di termodinamica

Proteine

Le proteine hanno una struttura primaria, che è la catena lineare di aminoacidi codificati dalla sequenza di DNA. La struttura primaria si può organizzare in modo locale creando strutture secondarie. Le strutture secondarie interagiscono tra loro anche a lungo raggio, compattandosi e dando luogo alla struttura terziaria.

Più strutture terziarie, quindi più polipeptidi, possono interagire associandosi in una struttura quaternaria.

Legami nelle proteine

Le proteine sono polimeri di aminoacidi legati tra loro da un legame covalente, chiamato *legame peptidico*.

L'unico altro legame covalente che troviamo nelle proteine è il *ponte disolfuro*, che si forma tra due residui di *cisteina*, che possono anche essere lontani tra loro nella struttura primaria, ma spazialmente vicini nella struttura terziaria.

Tutti gli altri legami sono legami deboli, importantissimi sia per la struttura secondaria che per la funzione della proteina.

I legami deboli permettono di legare ligandi, ad esempio ioni metallici, e permettono interazioni locali tra le catene laterali degli aminoacidi.

Nel ripiegamento è importante sia l'interazione tra gli aminoacidi che compongono la proteina sia l'interazione tra aminoacidi e solvente.

Legame a idrogeno

Nelle proteine sono molto importanti i legami idrogeno, legami deboli con una energia di legame che va dai 2 ai 6 kJ/mole, con una lunghezza di legame che varia tra i 2,5 e i 3,5 Å.

Gli intervalli dipendono dal fatto che i legami ad idrogeno non sono tutti uguali tra loro dal punto di vista dell'energia e della geometria.

Cambierà quindi la forza di legame in base alla geometria del legame e al tipo di atomi coinvolti.

Nelle proteine troveremo sempre intervenire l'ossigeno (O) e l'azoto (N). I casi in cui il legame idrogeno risulta importante sono la formazione delle strutture secondarie, nei legami nel solvente e nella formazione della struttura nativa. Il legame idrogeno è un legame debole ma nelle proteine ce ne sono molti.

Vedremo che nel processo di folding si romperanno i legami idrogeno col solvente e si formeranno legami idrogeno intramolecolari.

Interazioni elettrostatiche

Sono molto importanti, specialmente nell'interazione tra aminoacidi carichi in particolari condizioni di pH, che potrebbero interagire tra loro se vicini nello spazio (ad esempio Lisina ed Acido Aspartico), formando ponti salini.

Forze di van der Waals

Sono forze che riguardano strutture non polari in cui ci può essere la formazione di dipoli temporanei (CH₃ di valina e leucina).

L'interazione di van der Waals dipende da quanto sono vicine le parti che devono interagire. C'è una distanza a cui questa interazione è relativamente forte; se le parti si avvicinano oltre questa soglia l'interazione diventa di tipo repulsivo.

Queste interazioni sono importanti perché coinvolgono *aminoacidi non polari*, spesso correlati a gruppi metilici di aminoacidi come *valina* e *leucina*. Questi sono aminoacidi idrofobici, che spesso si trovano nella parte interna della proteina foldata. Sono legami importanti per il mantenimento della struttura nativa, e si formano nelle fasi finali del folding, quando la proteina è già in gran parte ripiegata.

Impacchettamento della proteina

La formazione della struttura secondaria dipende dal formarsi di legami ad idrogeno. Spesso la proteina contiene zone strutturate in strutture secondarie, alternate a zone poco strutturate come i loop.

I residui polari possono fare varie interazioni, interagendo tra loro e con il solvente. Troveremo molti residui polari sulla superficie esterna della proteina, e questi tenderanno a fare interazioni col solvente. Altri residui polari si trovano all'interno della proteina, e tendono a fare legami tra loro. I residui idrofobici tendono ad essere all'interno della proteina. L'interno della proteina è sempre più ricco di aminoacidi idrofobici rispetto alla superficie.

All'interno della proteina c'è una condizione di abbondante impacchettamento delle catene laterali. Si stima che circa il 70-77% del volume di una struttura proteica globulare sia occupato, mentre la restante parte è formata da cavità. Le cavità sono importantissime per la funzione, perché un enzima deve avere un certo grado di flessibilità. Piccoli cambiamenti conformazionali nella struttura della proteina sono importanti per la funzione della proteina, così come lo sono i loop, per la loro mancanza di strutturazione.

Sono proprio i loop le regioni che più si prestano a modifiche per l'evoluzione della proteina.

Definizioni

- *Stato nativo*: generalmente è la conformazione che manifesta l'attività biologica.

Sappiamo comunque che esistono proteine disordinate, nelle quali non ci sarà una conformazione ben definita.

- *Dominio*: è una unità strutturale della proteina con una precisa organizzazione. La singola catena polipeptidica si organizza in domini strutturali, che possono anche essere diversi a seconda della regione della proteina.

La proteina nel suo ambiente fisiologico non è rigida, ma assume un insieme di stati conformazionali, che possono essere evidenziati mediante NMR oppure con tecniche bioinformatiche di Dinamica Molecolare.

- *Stato denaturato*: si parla in questo caso spesso di *random coil*, che è uno stato in cui la proteina è quasi del tutto distesa e quasi libera di ruotare attorno ai suoi legami. Nella realtà questa condizione non si presenta, perché le proteine tendono sempre ad organizzarsi anche localmente.

Lo stato denaturato (*U*) è lo stato completamente denaturato. È una conformazione della proteina che otteniamo solamente trattandola con un agente denaturante.

Lo stato denaturato (*D*) è uno stato più fisiologico, in cui la proteina è denaturata, ma in condizione di folding più permissive. È una condizione di denaturazione meno estrema rispetto allo stato *U*.

La conformazione nativa è data da un insieme ristretto di conformazioni in cui la proteina è strutturata ed attiva.

La conformazione denaturata *D* è un insieme molto grande di conformazioni in cui la proteina è poco strutturata e non attiva, simili tra loro per energia, e interconvertibili tra loro.

Processi alla base del folding

L'importanza del solvente

Nell'acqua liquida abbiamo associazioni transienti tra le molecole d'acqua legate tra loro da legami ad idrogeno; queste associazioni vengono chiamate *flickering cluster*.

Quando parliamo del ripiegamento delle proteine dobbiamo sempre tenere presente non solo le caratteristiche del polipeptide, ma anche quelle del solvente in cui questo è disciolto.

Quando noi mettiamo in soluzione una molecola, nel nostro caso un polipeptide, induciamo degli effetti sulla struttura del solvente. Questi dipenderanno dalle caratteristiche del soluto, tra cui la sua idrofobicità ed idrofilicità.

Quando introduciamo una molecola idrofobica, non polare, che quindi non tende a formare legami ad idrogeno col solvente, questa rimarrà esclusa dalla soluzione. La sua presenza in ogni caso indurrà le molecole del solvente ad assumere una maggiore strutturazione. Le molecole d'acqua quindi tendono ad organizzarsi, facendo legami tra loro in modo da escludere la molecola non polare dalla soluzione.

Molecole non polari in acqua

Prendendo ad esempio una molecola apolare come la molecola di metano, questa si comporterà in maniera diversa in solvente polare (come l'acqua) ed in un solvente apolare.

Sappiamo che il metano è scarsamente solubile in acqua, quindi la dissoluzione del metano in acqua sarà sfavorita dal punto di vista termodinamico, secondo la reazione:

$$\Delta G^0 = -R T \ln(K_{eq})$$

dove K_{eq} dipende dalla solubilità nel solvente polare e nel solvente apolare.

Su tutta questa parte ho dei dubbi nella parte dei calcoli, perché dalla formula:

$$\Delta G^0 = \Delta H^0 - T \Delta S^0$$

non risulta questo valore con i dati forniti dalla prof in videolezione folding_2, da min. 4:10.

Il ΔG della reazione è sfavorito da un punto di vista termodinamico. ΔG del metano in acqua è:

$$\Delta G^0 = + 2,8 \text{ Kcal/mol.}$$

$$\Delta H^0 = -2,7 \text{ Kcal/mol}$$

$$\Delta S^0 = -18 \text{ Kcal/mol.}$$

Se facciamo i calcoli:

$$-2700 - ((32+273) * -18000) = 5487300$$

Se invece sostituiamo -18 Kcal/mol con -18 cal/mol:

$$-2700 - ((32+273) * -18) = 2790$$

che è coerente con un $\Delta G^0 = +2,8 \text{ Kcal/mol}$

Il problema è quindi probabilmente un K di troppo nel ΔS^0 della reazione.

Fine Dubbi

Quello che succede è che l'arrivo della molecola apolare costringe le molecole del solvente a fare un network di interazione reciproca. Questo vuol dire liberare energia ed infatti:

$$\Delta H^0 = -2,7 \text{ Kcal/mol.}$$

Questo tenderebbe a spingere la reazione verso la solubilità.

Il fattore entropico risente molto di questa condizione, perché la riorganizzazione delle molecole d'acqua causa un abbassamento dell'entropia. Il fattore entropico quindi prevale (ΔS negativo), ed il risultato è un ΔG positivo, quindi un aumento dell'energia libera della soluzione. La reazione è così sfavorita.

Effetto idrofobico

Quando la proteina è nel suo stato disteso gli aminoacidi polari faranno legami idrogeno con il solvente, mentre gli aminoacidi non polari resteranno esclusi dal solvente e ne causeranno una riorganizzazione locale.

Quando la proteina si ripiega tende a portare le zone idrofobiche verso l'interno della struttura, lasciando le zone idrofiliche esposte al solvente.

L'effetto idrofobico è una delle forze che spingono la proteina a raggiungere la sua forma nativa, rimuovendo i residui apolari dal contatto con il solvente, le cui molecole non sono quindi più costrette ad organizzarsi tra loro.

Succede anche che mentre la proteina si ripiega i gruppi polari rompono i legami che prima formavano col solvente.

Succede che questi accettori e donatori di legami ad idrogeno si riorganizzano all'interno della proteina, dando luogo a legami idrogeno all'interno della molecola e partecipando alla formazione di strutture secondarie.

Il fatto che la proteina si ripieghi ha 2 effetti positivi:

- si riduce la superficie idrofobica a contatto con l'acqua
- i gruppi idrofobici possono stabilire interazioni tra loro all'interno della molecola (van der Waals)

Le proteine che contengono molti aminoacidi polari tendono a stare in soluzione in una forma più distesa. Questo è ciò che avviene nelle *proteine intrinsecamente disordinate*.

Strutture secondarie

Sia alfa-eliche che beta-foglietti sono dovuti a legami idrogeno che si instaurano al livello del backbone della proteina.

Le catene laterali degli aminoacidi non partecipano direttamente alla formazione delle strutture secondarie.

Differential Scanning Calorimetry per monitorare la denaturazione delle proteine.

La calorimetria differenziale è un metodo biochimico che permette di misurare come aumenta la temperatura di una certa soluzione man mano che noi forniamo calore.

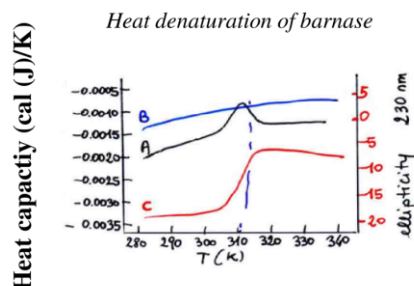
Misurando la differenza tra un campione di solvente ed un campione con la soluzione di solvente e proteina possiamo trarre delle conclusioni sullo stato di ripiegamento della proteina.

La *capacità termica* indica quanta energia devo dare alla soluzione per aumentare la sua temperatura di 1° C o K.

L'esperimento si fa aumentando la temperatura della soluzione e misurando la variazione del parametro della capacità termica. Nel caso in analisi viene usata la proteina modello *barnasi*.

Nella misurazione della capacità termica del solo solvente (acqua in questo caso) si evidenzia una curva con un aumento regolare della capacità termica.

Quando si misura la capacità termica della soluzione, all'inizio la curva è parallela a quella del solvente. Ad un certo punto si evidenzia un picco, successivamente al quale la curva torna parallela a quella del solvente. Il picco nella capacità termica indica che c'è un intervallo di temperatura



B = buffer

C = circular dichroism measure

A = heat capacity

in cui serve fornire più energia per aumentare la temperatura della soluzione. Il picco indica l'aumento della capacità termica dovuto all'energia necessaria a rompere i legami della proteina in soluzione durante il processo di denaturazione. Dopo la denaturazione della proteina la capacità termica torna ad essere quella del solvente.

Un esperimento di dicroismo circolare evidenzia la transizione, che indica il passaggio della proteina dallo stato nativo allo stato denaturato **D**.

Nel punto in cui la proteina denatura (*midpoint*) la capacità termica aumenta. Il picco corrisponde alla temperatura di melting ***T_m*** o, più correttamente, *temperatura di midpoint*.

Quando la proteina denatura, gli aminoacidi idrofobici che erano all'interno della struttura vengono esposti al solvente. Questo causa una maggiore organizzazione del solvente, che necessita di maggiore energia per essere alterata. Da questo l'aumento di capacità termica.

Da queste misure di capacità termica si possono ricavare i parametri termodinamici della reazione di transizione da stato nativo a stato denaturato.

Termodinamica del folding

Le proteine possono essere denaturate *in vitro* tramite diversi metodi:

- Calore
- Denaturanti
- pH
- Pressione

Spesso la denaturazione è reversibile, vuol dire che quando tolgo l'elemento che ha perturbato il folding, la proteina riprende la sua forma nativa.

Non sempre il folding è reversibile. A volte la proteina può aggregare con altre proteine, il che ne impedisce il folding.

Il folding di una proteina globulare può essere descritto dalla equazione:

$$\Delta G^0 = \Delta H^0 - T \Delta S^0$$

Dove il ΔS^0 ed il ΔH^0 sono negativi.

Dato che il ripiegamento delle proteine è favorito bisogna capire come variano il ΔS ed il ΔH in modo da dare un ΔG negativo.

Contributo entalpico e contributo entropico ΔH e ΔS sono il risultato di tanti elementi che si sommano.

Entropia

$$S = K * \ln W$$

dove K è la costante di Boltzman = (R/N) J/K

W è il numero di stati conformazionali

R è la costante universale dei gas

N è il numero di Avogadro

La proteina denaturata ha un alto numero di stati conformazionali. Quando la proteina si ripiega gli stati conformazionali diminuiscono, così con l'entropia S (entropia conformazionale). *Questo è un fattore **destabilizzante**.*

Quando la proteina si ripiega gli aminoacidi idrofobici non sono più a contatto con l'acqua, quindi l'acqua non deve più organizzarsi e la sua entropia aumenta. *Questo è un fattore **stabilizzante**.*

Resta il fatto che siccome ΔS diminuisce nel processo di folding, e davanti a ΔS c'è un segno -, il contributo entropico comunque è positivo. Per avere un ΔG negativo dobbiamo contare sul bilanciamento del fattore entalpico ΔH .

Entalpia

Fattori stabilizzanti $\Delta H < 0$

Contributi stabilizzanti sono tutti quelli in cui $\Delta H < 0$, perché dobbiamo avere un ΔG negativo. Tipicamente ciò a cui contribuisce il fattore entalpico è la formazione di legami. I legami contribuiscono con la loro energia.

Forze di Van der Waals

Quando la proteina è distesa, le forze di van der Waals non sono rilevanti. Quando la proteina è ripiegata, aminoacidi vicini possono interagire e dar luogo a legami di van der Waals.

Ponti ad idrogeno

Quando la proteina è denaturata forma legami idrogeno con l'acqua, ma quando si ripiega questi legami idrogeno verranno rotti, formandone altri all'interno della sua struttura.

La formazione di ponti ad idrogeno è importante nel folding, perché in una proteina nella sua conformazione nativa i legami ad idrogeno hanno una energia maggiore, perché distanza e geometria di legame sono ottimali.

Ponti salini

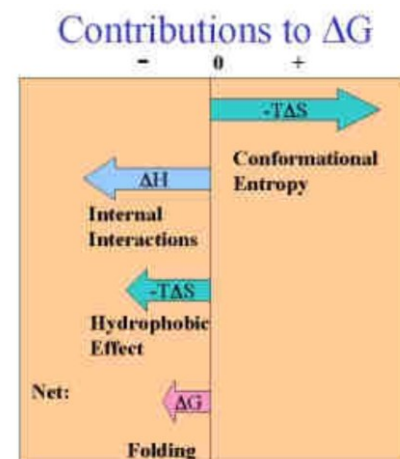
I ponti salini sono interazioni tra cariche negative e positive che stabilizzano la struttura nativa. Si formano spesso reti di ponti salini di stabilizzazione.

Fattori destabilizzanti $\Delta H > 0$

Quando la proteina si ripiega ci sono zone che sono distorte dal punto di vista conformazionale, che creano tensioni nella struttura che tendono a destabilizzare la proteina.

Si parla sempre di legami deboli, perché nel corso del folding il legame peptidico non viene intaccato. Questa grande quantità di interazioni deboli causa comunque una ampia variazione di energia totale nei due processi.

La differenza nell'energia libera tra proteina nello stato nativo rispetto allo stato denaturato è solitamente piccola, nell'ordine di 20 KJ/mol. Sembra in contraddizione con quello detto precedentemente, ma le energie vanno in direzioni opposte, quindi il risultato è una piccola variazione del ΔG . Le proteine quindi sono stabili solo quel tanto che basta a mantenere la struttura.



Differenze nella sequenza influenzano la stabilità delle proteine

La stabilità è definita come capacità di mantenere la struttura nativa. La variabilità di sequenza è spesso tollerata e non sempre associata a cambiamenti nella stabilità. Sappiamo anche, dagli esperimenti di mutagenesi e patologie ereditarie, che a volte può bastare una sola sostituzione amminoacidica per alterare fortemente la stabilità di un polipeptide.

Queste due osservazioni ci danno indicazioni relative sulla stabilità, ma si può dedurre che nel corso dell'evoluzione la pressione selettiva ha selezionato un livello di stabilità adeguato, non eccessivo né ridotto, ma quando basta per avere una attività alle condizioni di operazione.

La stabilità minima indica flessibilità, e la flessibilità è spesso alla base della funzione enzimatica.

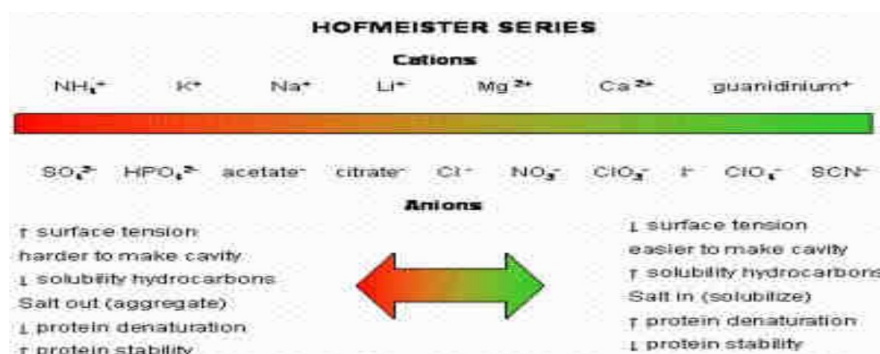
Le condizioni intracellulari favoriscono la forma nativa delle proteine. Nella cellula c'è un'alta concentrazione di molecole e ioni, il cosiddetto crowding, che agevola il ripiegamento e la stabilità.

Denaturazione

Possiamo denaturare una proteina scaldandola. L'aumento di temperatura aumenta l'energia cinetica degli atomi in soluzione, alcune zone della proteina cominceranno quindi a disorganizzarsi, dando il via alla denaturazione.

Possono anche essere usati dei denaturanti chimici.

La serie di Hofmeister classifica vari ioni in base alla loro capacità di precipitare e denaturare proteine.



Precipitazione

Se aggiungiamo certi sali ad una soluzione proteica le proteine tendono a precipitare.

Se si aggiunge solfato di ammonio le proteine precipitano; lo ione solfato è un buon precipitante.

La precipitazione dipende dal fatto che gli ioni rimuovono l'alone di idratazione che circonda il polipeptide; rimuovendo le molecole d'acqua dall'intorno del peptide questo tende a precipitare.

Denaturazione

La capacità di precipitare e la capacità di denaturare vanno in senso inverso. Lo ione solfato SO_4^{2-} è un buon precipitante, ma non è in grado di denaturare una proteina. L'isotiocianato SCN^- invece denaturerà bene la proteina ma non la farà precipitare.

Come denaturante chimico useremo spesso il *guanidinio cloruro* o l'*urea*.

I denaturanti agiscono sulla struttura della proteina e sulla struttura del solvente. Influenzano l'effetto idrofobico.

L'azione di una sostanza denaturante è quella di facilitare l'interazione del peptide con il solvente, favorendo la conformazione distesa.

La stessa cosa fa il detergente ionico SDS. Questa è una molecola con una porzione polare ed una apolare. Le parti non polari della proteina interagiranno con le regioni non polari del detergente, mentre le regioni polari manterranno questo complesso in soluzione.

La rimozione dei denaturanti dalla soluzione avviene dializzando la soluzione e mettendo la proteina in condizioni nuovamente favorevoli al folding.

Possiamo denaturare la proteina anche per mezzo di variazioni del pH.

Alcuni aminoacidi hanno catene laterali ionizzabili, che si caricano con variazioni di pH, in base alla loro pK_a .

Ad esempio l'acido glutammico a pH7 è deprotonato (COO^-) mentre a pH 4,5 è protonato al 50% e muove l'equilibrio verso la denaturazione.