

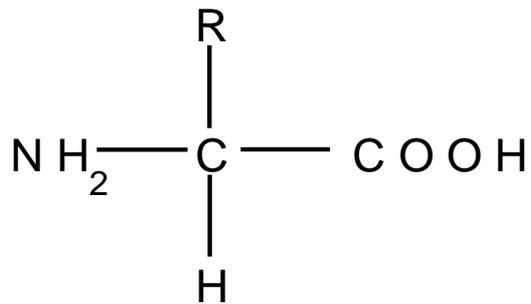
Proteine

Svolgono importanti funzioni in base alle quali sono **classificate come**:

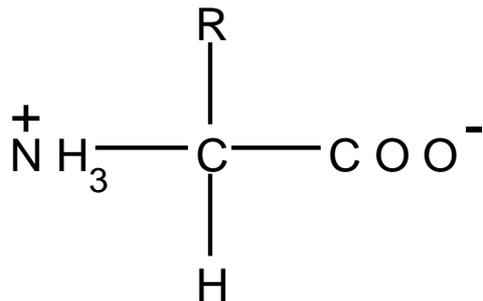
- ▶ **Proteine catalitiche** (enzimi)
- ▶ **Proteine di trasporto** (emoglobina, siero albumina) La mioglobina lega l'ossigeno nelle cellule del muscolo scheletrico e cardiaco; l'emoglobina lega e trasporta ossigeno e anidride carbonica nei globuli rossi del sangue.
- ▶ **Proteine di riserva** (caseina, ovalbumina)
- ▶ **Proteine strutturali** (collagene, elastina, cheratina, tubulina, actina). Sostengono e mantengono la forma della cellula e quindi dei tessuti.
- ▶ **Proteine di difesa** (immunoglobuline o anticorpi, fibrinogeno, trombina). Gli anticorpi difendono i vertebrati dalle infezioni batteriche o virali.
- ▶ **Proteine di regolazione** (ormoni)
- ▶ **Proteine contrattili** (actina, miosina). Possono compiere lavoro meccanico come il movimento dei flagelli, la separazione dei cromosomi durante la mitosi e la contrazione del muscolo.

Proteine

Le proteine sono **copolimeri di condensazione** di α -amminoacidi



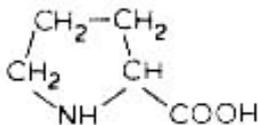
Struttura generale di un
amminoacido



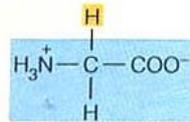
Zwitterion

I venti AMMINOACIDI PRINCIPALI che compongono le proteine

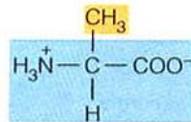
Common name	Adjectival name ^a	3-Letter code	1-Letter code	Side chain R	Remarks
Glycine	glycyl	Gly	G	-H	no sidechain except for hydrogen atom
Alanine	alanyl	Ala	A	-CH ₃	methyl sidechain
Valine	valyl	Val	V	$ \begin{array}{c} \text{CH}_3 \\ \diagup \\ -\text{CH} \\ \diagdown \\ \text{CH}_3 \end{array} $	aliphatic, branched at β -carbon
Threonine	threonyl	Thr	T	$ \begin{array}{c} \text{OH} \\ \diagup \\ -\text{C} \cdots \text{CH}_3 \\ \diagdown \\ \text{H} \end{array} $	has hydroxyl group branched at β -carbon
Aspartate (aspartic acid)	aspartyl	Asp	D	-CH ₂ -COO ⁻	acidic sidechain
Glutamate (glutamic acid)	glutamyl	Glu	E	-(CH ₂) ₂ -COO ⁻	acidic sidechain
Asparagine	asparaginyl	Asn	N	-CH ₂ -CO-NH ₂	corresponds to aspartate with amidated sidechain
Glutamine	glutaminyl	Gln	Q	-(CH ₂) ₂ -CO-NH ₂	corresponds to glutamate with amidated sidechain
Lysine	lysyl	Lys	K	-(CH ₂) ₄ -NH ₃ ⁺	basic sidechain
Histidine	histidyl	His	H	$ \begin{array}{c} -\text{CH}_2 \\ \\ \text{HN} \cdots \text{NH} \\ \diagup \quad \diagdown \\ \text{H} \quad \text{H} \end{array} $	basic and cyclic sidechain

Arginine	arginyl	Arg	R	$-(\text{CH}_2)_3-\text{NH}-\text{C}(\text{NH}_2)_2^+$	basic guanidinyll group sidechain
Phenylalanine	phenylalanyl	Phe	F	$-\text{CH}_2-\text{C}_6\text{H}_5$	aromatic ring in side-chain
Tyrosine	tyrosyl	Tyr	Y	$-\text{CH}_2-\text{C}_6\text{H}_4-\text{OH}$	aromatic ring with phenoxy (OH) group on side-chain
Tryptophan	tryptoph(an)yl	Trp	W	$-\text{CH}_2-\text{Indole}$	has double aromatic and heterocyclic ring in side-chain
Cysteine	cysteinyll	Cys	C	$-\text{CH}_2-\text{SH}$	has sulphydryll group which may be oxidized to form S-S bridge between residues then referred to as 'cystine'
Methionine	methionyl	Met	M	$-(\text{CH}_2)_2-\text{S}-\text{CH}_3$	has sulphur atom, but not sulphydryll group
Proline	prolyll	Pro	P	Structure of whole molecule: 	technically an 'imino acid' sidechain covalently bonded to backbone nitrogen

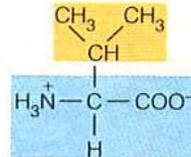
AMINOACIDI ALIFATICI



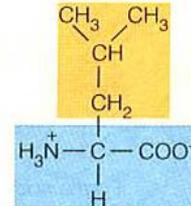
Glicina (Gly) G



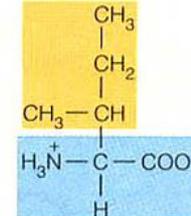
Alanina (Ala) A



Valina (Val) V

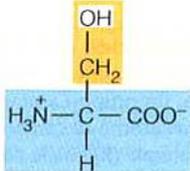


Leucina (Leu) L

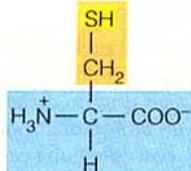


Isoleucina (Ile) I

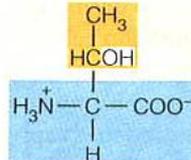
AMINOACIDI CON CATENE LATERALI CONTENENTI ZOLFO O GRUPPI OSSIDRILICI



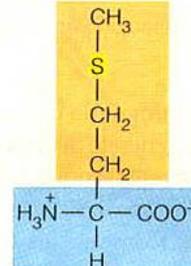
Serina (Ser) S



Cisteina (Cys) C

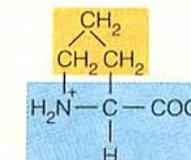


Treonina (Thr) T



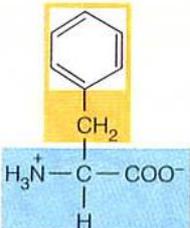
Metionina (Met) M

AMINOACIDO CICLICO

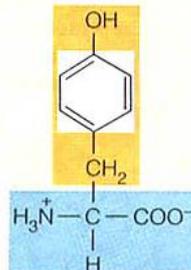


Prolina (Pro) P

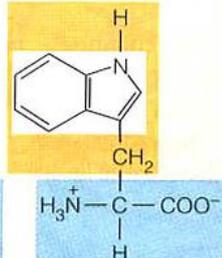
AMINOACIDI AROMATICI



Fenilalanina (Phe) F

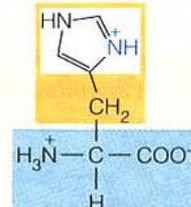


Tirosina (Tyr) Y

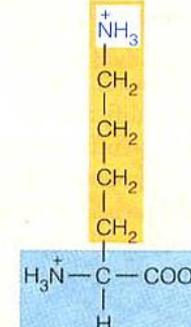


Triptofano (Trp) W

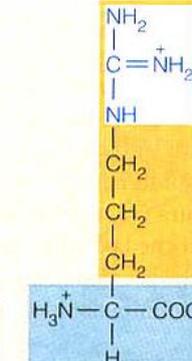
AMINOACIDI BASICI



Istidina (His) H

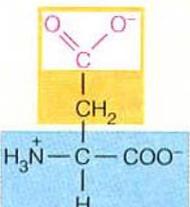


Lisina (Lys) K

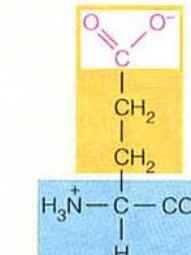


Arginina (Arg) R

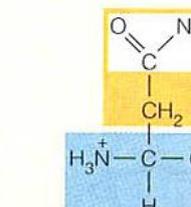
AMINOACIDI ACIDI E LORO AMIDI



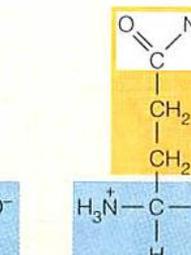
Acido aspartico (Asp) D



Acido glutammico (Glu) E



Asparagina (Asn) N



Glutammina (Gln) Q

Gli aminoacidi incorporati nelle proteine. I 20 α -aminoacidi che sono incorporati nelle proteine sono qui classificati secondo i criteri discussi nel testo. Sotto ciascun aminoacido sono riportati il nome, l'abbreviazione secondo il codice a tre lettere e l'abbreviazione secondo il codice a una lettera.

Amminoacidi con catena alifatica

Gly, Ala, Leu, Val e Ile sono amminoacidi idrofobici spesso localizzati all'interno della molecola proteica, dove risultano isolati dall'acqua. La glicina è l'amminoacido più piccolo e non è chirale: può alloggiare in piccole cavità che non possono accogliere altri amminoacidi. La **Pro** ha una catena laterale che si ripiega a legare il gruppo amminico creando un ciclo. Di conseguenza la prolina contiene un gruppo amminico secondario e non primario. L'anello pirrolidinico eterociclico della prolina limita la geometria dei polipeptidi introducendo a volte bruschi cambiamenti nella direzione della catena polipeptidica.

Amminoacidi con catene laterali contenenti atomi di zolfo o gruppi ossidrilici

Ad eccezione di **Met** che risulta particolarmente idrofobica, gli amminoacidi **Ser, Cys** e **Thr** sono più idrofilici degli amminoacidi con catena alifatica. La **Met** svolge un ruolo particolare nella sintesi proteica in quanto costituisce quasi sempre il primo amminoacido della catena.

L'ossidazione di due residui di **Cys** può portare alla formazione di un ponte a disolfuro nella cistina. L'ossidazione avviene più facilmente a pH leggermente alcalini.

Amminoacidi aromatici

Phe è tra gli amminoacidi più idrofobici mentre in **Tyr** e **Trp** il carattere idrofobico è moderato da gruppi polari. Il gruppo alcolico della tirosina è ionizzabile ma nelle normali condizioni fisiologiche mantiene l'idrogeno. Il triptofano contiene un gruppo indolico biciclico. Gli amminoacidi aromatici assorbono la luce UV a circa 280 nm (Phe a 260 nm) in quanto contengono elettroni π delocalizzati.

Amminoacidi basici

Lys, **Arg** e **His** sono residui basici fortemente polari e perciò si trovano sulla superficie delle proteine dove possono essere idratati dall'ambiente acquoso. Le loro catene sono cariche positivamente in condizioni fisiologiche. In particolare **His** incorporata in proteine raggiunge pKa di circa 7. Poichè His può scambiare protoni a valori di pH prossimi al pH fisiologico, è spesso coinvolta in catalisi enzimatiche in cui si ha trasferimento di protoni. L'istidina contiene un anello imidazolico.

Amminoacidi acidi e loro ammidi

Asp e **Glu** sono carichi negativamente a pH fisiologico. **Asp**, **Glu**, **Asn** e **Gln** sono molto idrofilici e tendono a disporsi sulla superficie delle proteine per interagire favorevolmente con l'acqua.

Otto amminoacidi naturali vengono definiti **essenziali** in quanto devono essere assunti con la dieta visto che il nostro organismo non è in grado di sintetizzarli. Gli amminoacidi essenziali sono: **Valina, Leucina, Isoleucina, Fenilalanina, Treonina, Lisina, Triptofano e Metionina.**

Negli organismi viventi sono presenti **più di 200 diversi amminoacidi.**

Oltre ai 20 standard, tutte le specie contengono L-aa precursori degli amminoacidi comuni o intermedi in altre vie metaboliche. Ad es. la S-adenosilmetionina è un donatore di gruppi metilici in molte vie biochimiche.

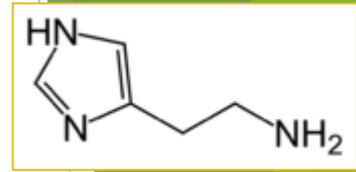
Batteri e funghi sintetizzano aa **D** per le pareti batteriche. Nel cervello dei mammiferi il glutammato viene convertito in γ -amminobutirrato (GABA), un neurotrasmettitore. L'istamina viene ottenuta nei mammiferi dall'istidina e controlla la contrazione di alcuni vasi sanguigni e la produzione di HCl nello stomaco.

La tirosina è il precursore nel surrene della noradrenalina che può trasformarsi poi in adrenalina. La tirosina è anche il precursore degli ormoni tiroidei (tiroxina e triiodotironina).

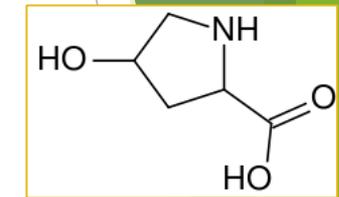
Alcuni aa vengono **modificati chimicamente dopo essere stati incorporati nella catena polipeptidica.** Per es. alcuni residui di Pro vengono convertiti in idrossiprolina nel collagene. Alcuni residui possono venire legati a catene carboidratiche (glicosilazione). Molte proteine vengono fosforilate per aggiunta di gruppi fosforici alle catene laterali di Ser Thr o Tyr. Anche la formazione di gruppi disolfuro avviene a catena polipeptidica assemblata.

Alcuni aminoacidi biologicamente importanti non riscontrati nelle proteine

Nome	Formula	Fonte biochimica e funzione
β -Alanina	$\text{H}_3\text{N}^+ - \text{CH}_2 - \text{CH}_2 - \text{COO}^-$	Presente nell'acido pantotenico (una vitamina) e in alcuni importanti peptidi naturali
D-Alanina	$\begin{array}{c} \text{COO}^- \\ \\ \text{H} - \text{C} - \text{NH}_3^+ \\ \\ \text{CH}_3 \end{array}$	Nei polipeptidi di alcune pareti cellulari batteriche
Acido γ -amino-butirrico	$\text{H}_3\text{N}^+ - \text{CH}_2 - \text{CH}_2 - \text{CH}_2 - \text{COO}^-$	Nel cervello, in altri tessuti animali; ha la funzione di neurotrasmettitore
Acido D-glutammico	$\begin{array}{c} \text{COO}^- \\ \\ \text{H} - \text{C} - \text{NH}_3^+ \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 - \text{COO}^- \end{array}$	Nei polipeptidi di alcune pareti cellulari batteriche
L-Omoserina	$\begin{array}{c} \text{COO}^- \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{H} \\ \\ \text{CH}_2 - \text{CH}_2\text{OH} \end{array}$	In molti tessuti; è un intermedio del metabolismo degli aminoacidi
L-Ornitina	$\begin{array}{c} \text{COO}^- \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{H} \\ \\ \text{CH}_2 - \text{CH}_2 - \text{CH}_2\text{NH}_3^+ \end{array}$	In molti tessuti; è un intermedio nella sintesi dell'arginina
Sarcosina	$\begin{array}{c} \text{CH}_3 - \text{N} - \text{CH}_2 - \text{COO}^- \\ \\ \text{H} \end{array}$	In molti tessuti; è un intermedio nella sintesi degli aminoacidi
L-Tiroxina	$\begin{array}{c} \text{COO}^- \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{H} \\ \\ \text{CH}_2 - \text{C}_6\text{H}_3(\text{I})_2 - \text{O} - \text{C}_6\text{H}_3(\text{I})_2 - \text{OH} \end{array}$	Nella ghiandola tiroide; è un ormone tiroideo (I = iodio)



Istamina

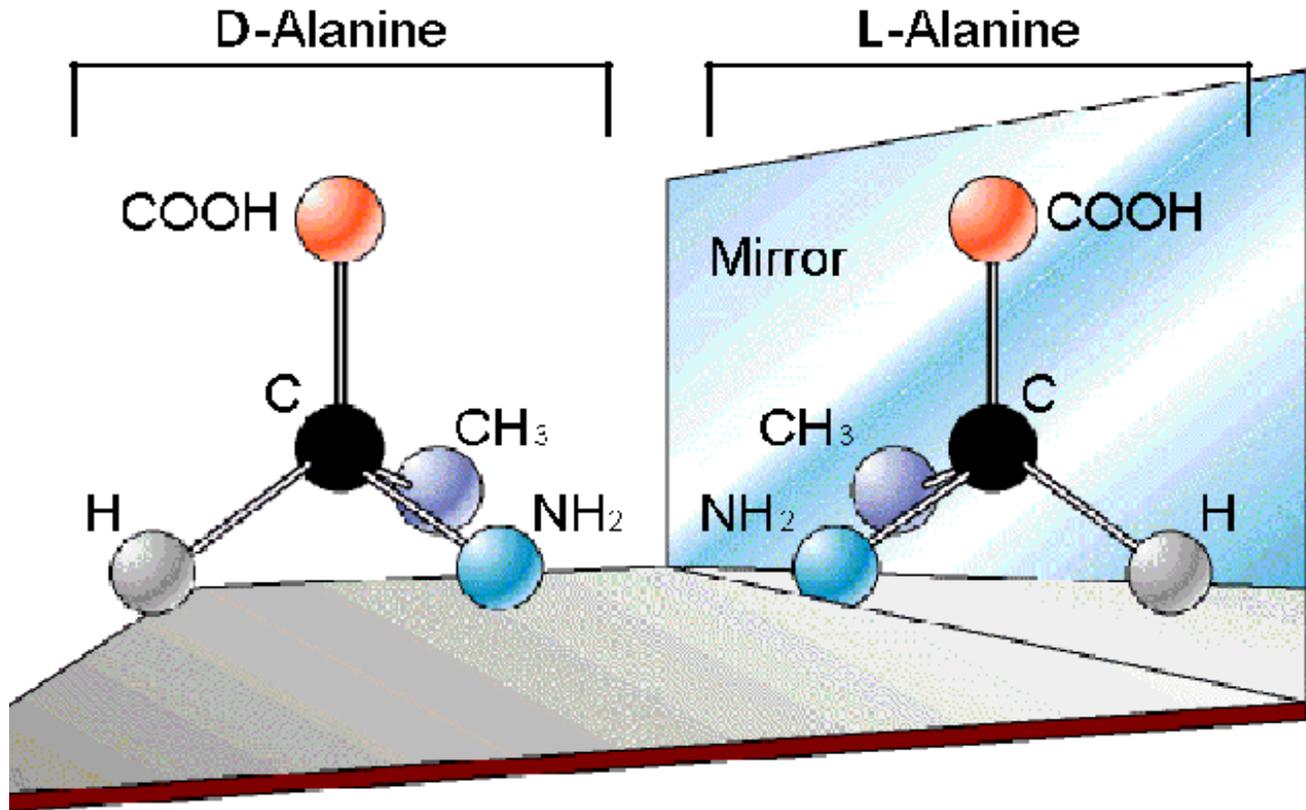


Idrossiprolina

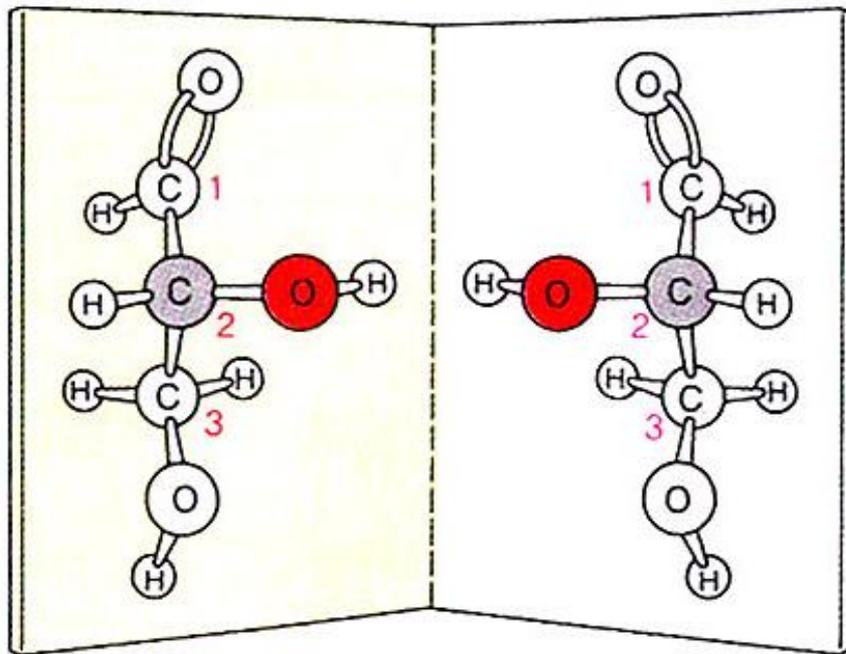
pKa di gruppi dissociabili delle proteine

<i>Tipo di gruppo</i>	<i>intervallo di pKa</i>
Carbossilico	3.5 - 4.0
Carbossilico della catena laterale di Asp e Glu	4.0 - 4.8
Imidazolico (His)	6.5 - 7.4
Cisteinico (SH)	8.5 - 9.0
Fenolico (Tyr)	9.5 - 10.5
Alfa-amminico	8.0 - 9.0
Amminico della catena laterale di Lys	9.8 - 10.4
Guanidinico (Arg)	12

STEREOISOMERIA DEGLI AMMINOACIDI



(Karp 1997)



D-Gliceraldeide

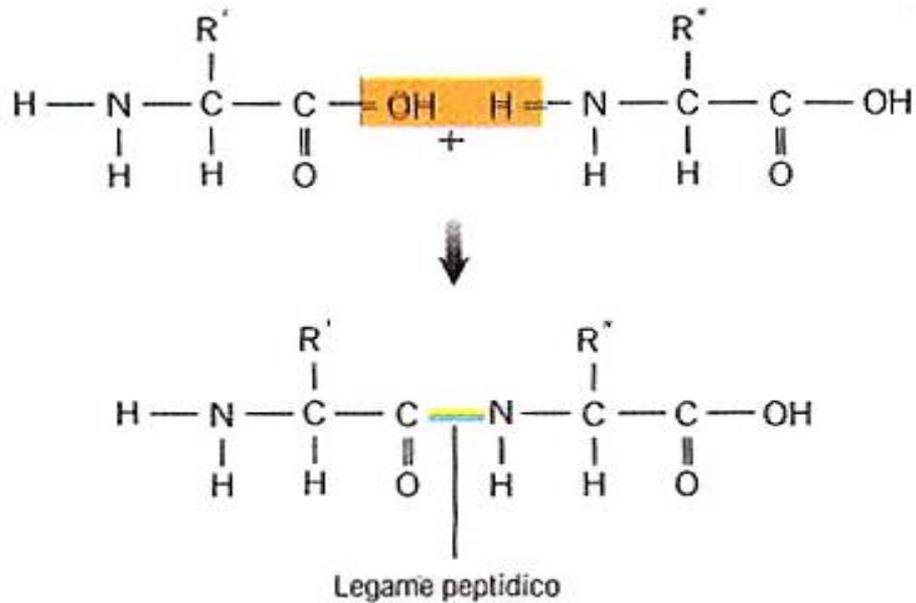
L-Gliceraldeide

Gli enantiomeri della gliceraldeide. La configurazione dei gruppi intorno all'atomo di carbonio 2 chirale (raffigurato in grigio scuro) permette di distinguere la D-gliceraldeide dalla L-gliceraldeide. Le due molecole sono immagini speculari e non possono essere sovrapposte l'una all'altra.

Numero piccolo di residui amminoacidici → oligopeptidi

Numero elevato di residui amminoacidici → polipeptidi o peptidi

LA FORMAZIONE DEL LEGAME PEPTIDICO (AMMIDICO) AVVIENE PER CONDENSAZIONE (-H₂O) DI DUE AMMINOACIDI



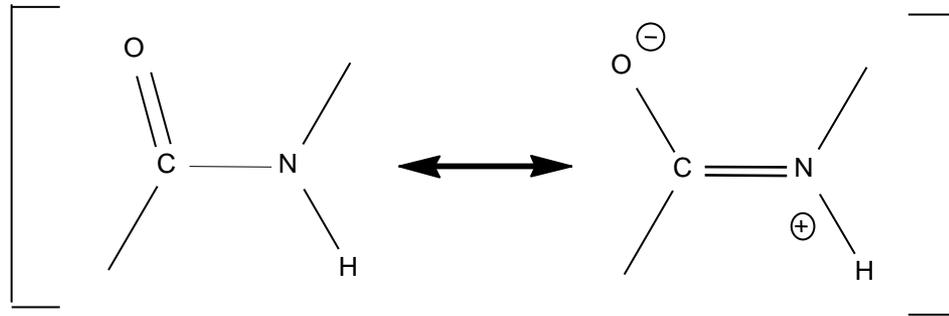
L'analisi mediante cristallografia ai raggi X rivela che nel legame peptidico:

-la lunghezza del legame C-N è più corta del previsto, intermedia tra quella di un legame singolo e quella di un legame doppio;

-Il legame C-O è più lungo rispetto a quello di un legame C=O

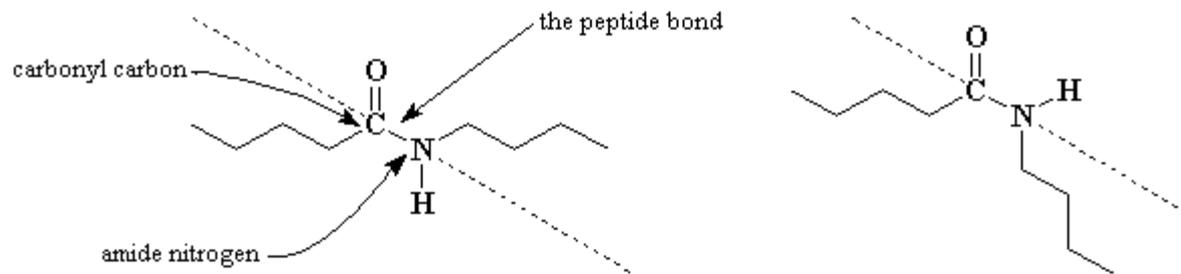
Tali misure evidenziano un **carattere parziale di doppio legame** per il legame C-N come ben descritto dall'ibrido di risonanza.

IL LEGAME PEPTIDICO HA NATURA INTERMEDIA TRA LEGAME SEMPLICE E LEGAME DOPPIO

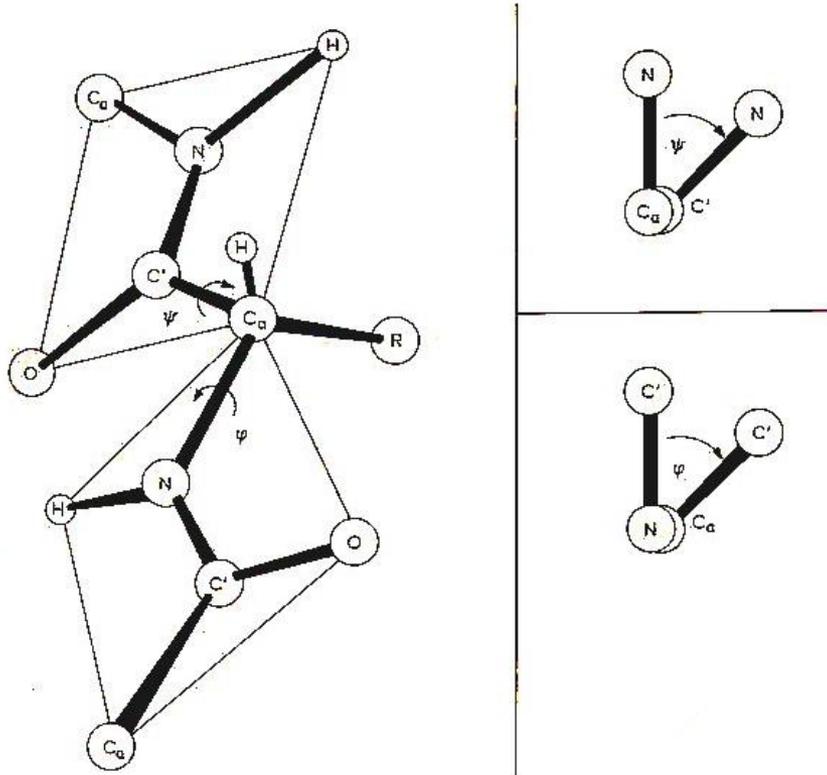


La configurazione del doppio legame può essere *cis* o *trans* e si forma durante la sintesi della proteina.

La forma *cis* è meno favorita della *trans* a causa dell'ingombro sterico delle catene laterali degli atomi in alfa. Eccezione: per la Pro la forma *cis* è solo leggermente superiore per energia alla forma *trans*. Le peptidil prolil *cis/trans* isomerasi possono catalizzare la conversione.

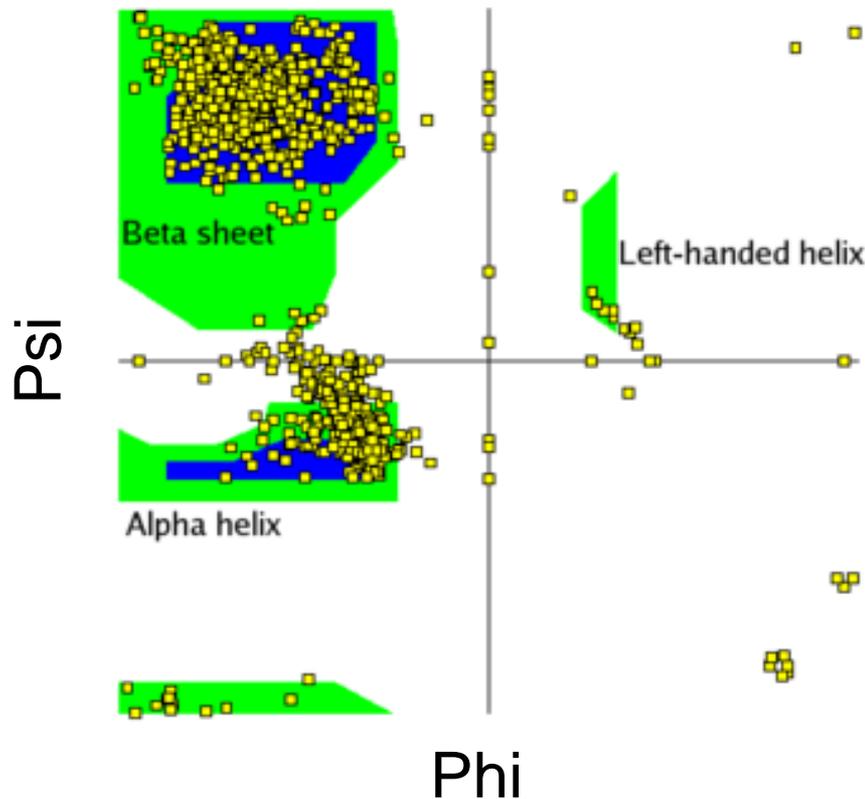


Il parziale carattere di doppio legame del gruppo peptidico costringe sullo stesso piano 6 atomi, ma la rotazione attorno ai legami N-C_{alfa} e C_{carbonilico}-C_{alfa} risulta permessa



Φ (Phi) angolo di rotazione attorno al legame N-C_α
Ψ (Psi) angolo di rotazione attorno al legame C_α-CO

La **rotazione di phi e psi** è però limitata dall'ingombro sterico degli atomi della catena principale e dalle catene laterali dei residui adiacenti. Gli angoli possibili sono indicati nelle regioni ombreggiate delle **mappe steriche di Ramachandran**.



Alcuni residui hanno mappe molto ristrette come la prolina, altri zone più ampie come la glicina.

STRUTTURA PRIMARIA

Specifica sequenza dei residui amminoacidici caratteristici di ciascun peptide o proteina

Esempio:

Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly

Sequenza primaria dell'ossitocina

Le proteine evolvono per cambiamenti conservativi e non.



AGGS → Amminoacido C-terminale



Amminoacido N-terminale

(posto per convenzione a sinistra)

Composizione:

- Analisi amminoacidica (idrolisi, proteasi, BrCN)
- Sequenziamento

STRUTTURA SECONDARIA

Struttura corrispondente ad un determinato insieme di valori degli angoli di torsione φ e ψ

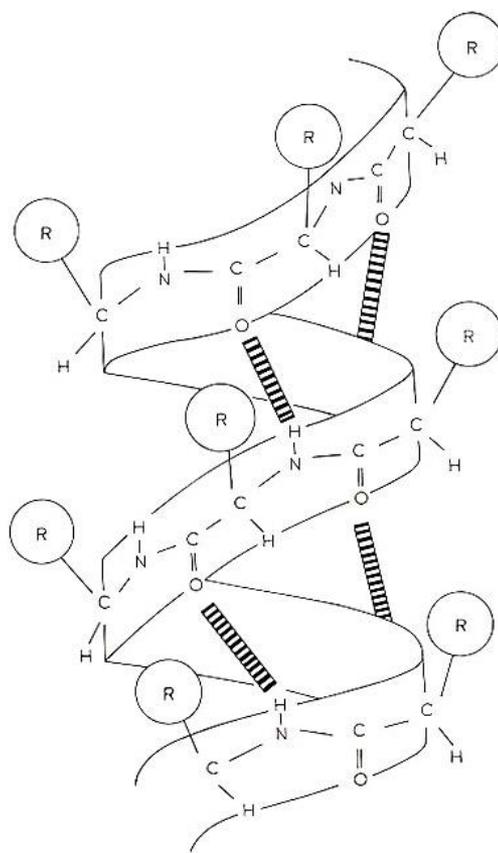


Orientamento relativo e regolare dei diversi segmenti di una proteina

La conoscenza sulla struttura secondaria delle proteine deriva in gran parte dagli studi di Linus Pauling degli anni '30. Secondo i **postulati di Pauling** qualsiasi struttura secondaria deve possedere i seguenti requisiti:

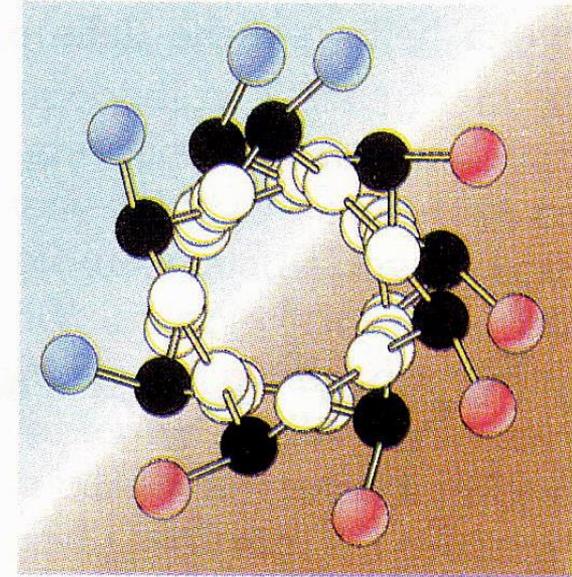
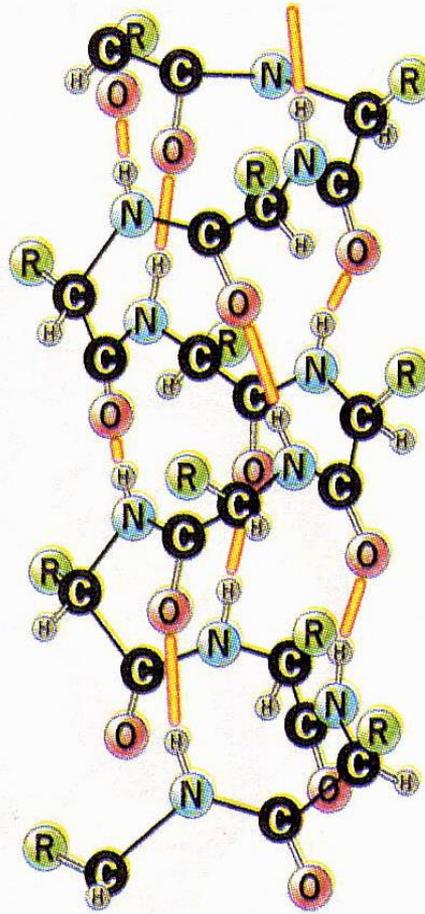
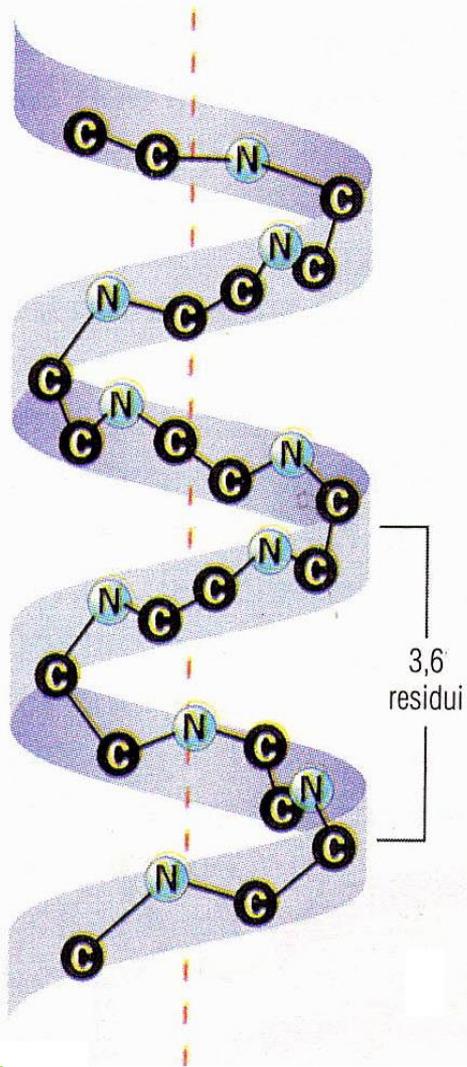
1. Le lunghezze e gli angoli di legame devono essere quelli riscontrati tramite diffrazione dei raggi X;
2. Due atomi non devono avvicinarsi più di quanto sia loro consentito dai rispettivi raggi di Van der Waals;
3. Il gruppo ammidico deve rimanere planare e in configurazione trans;
4. Dovrebbero essere presenti alcuni tipi di legame non covalente per stabilizzare i ripiegamenti regolari. Le configurazioni favorite sono quelle che permettono la formazione del maggior numero di legami a ponte di idrogeno

α -elica



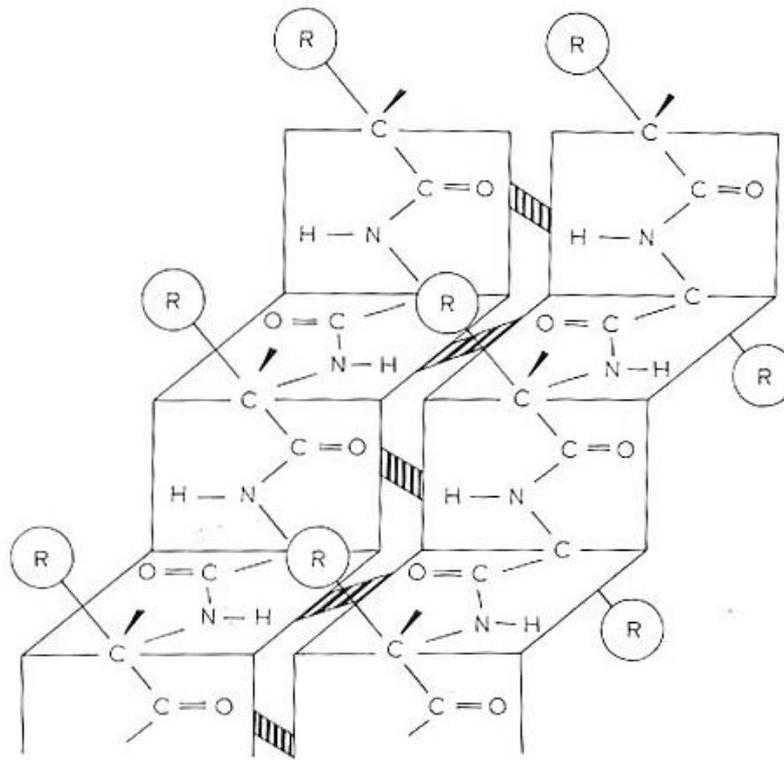
La struttura α -elicoidale solitamente è **destrogiro**. Possiede **3.6 residui per giro**. Il gruppo carbonilico del primo residuo lega con ponte ad H il gruppo amminico del quarto residuo successivo chiudendo un'ansa di **13 atomi**. I legami ad H sono circa paralleli all'asse. $\Phi = -57^\circ$ $\Psi = -47^\circ$

L'effetto cumulativo di molti legami a H stabilizza la conformazione. Nelle regioni proteiche interne idrofobiche le strutture elicoidali sono particolarmente stabili perché l'acqua non può competere nella formazione dei legami a ponte d'idrogeno. Ala si trova spesso in strutture α -elicoidali mentre Gly e Pro destabilizzano l'elica. La sua lunghezza varia da 4 a 20 residui ma mediamente ne comprende 12. Molte α -eliche hanno **natura anfipatica**. Più strutture elicoidali anfipatiche possono unirsi tramite le facce idrofobiche (strutture coiled coil; leucin zipper)

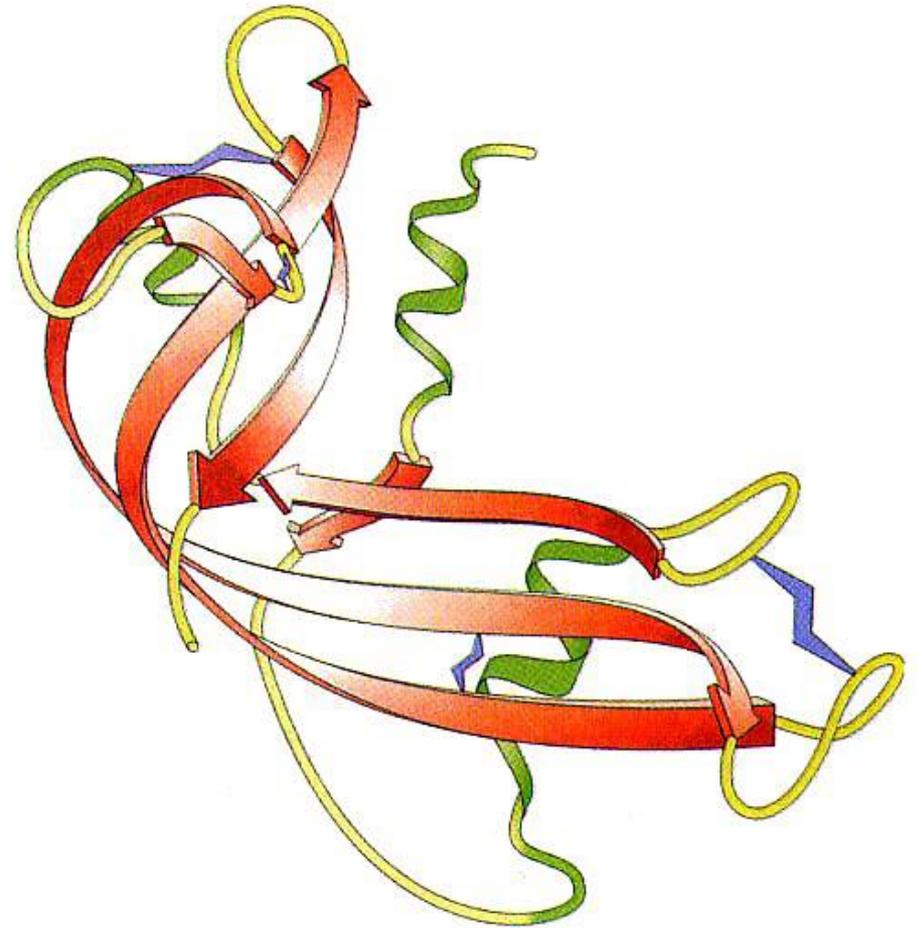
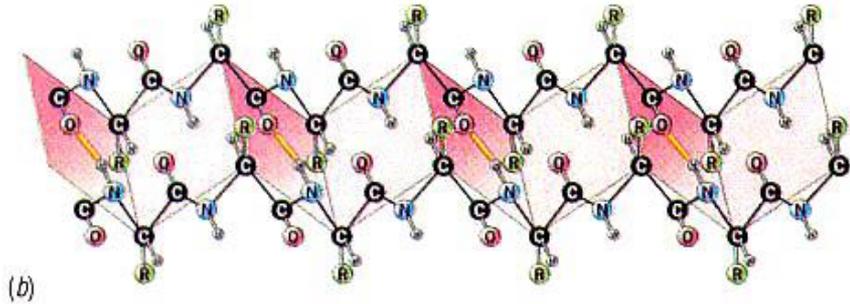
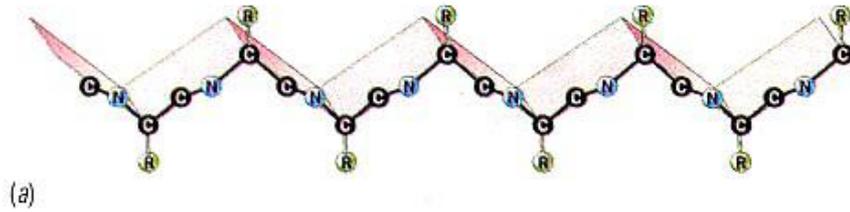


Alcune proteine contengono tratti di eliche 3_{10} .

β -sheet



Le proteine raramente contengono filamenti β isolati perché questa struttura di per se non è significativamente più stabile di altre conformazioni. I filamenti possono essere **paralleli o antiparalleli**. Nei filamenti antiparalleli i legami a H sono circa perpendicolari rispetto alla catena. Le catene laterali sono proiettate alternativamente sopra e sotto il piano. Il lato idrofobico tende a interagire con l'interno idrofobico della proteina. I foglietti β interni hanno entrambi i lati idrofobici



Il più elegante utilizzo della struttura β sheet è costituito dalla **seta** e dalla **fibra filata dai ragni**. La fibroina della seta contiene lunghe sequenze di foglietti β antiparallelo date dalla ripetizione di sequenze

$-\text{Gly-Ala-Gly-Ala-Gly-Ser-Gly-Ala-Ala-Gly}-(\text{Ser-Gly-Ala-Gly-Ala-Gly})_8-$

Il risultato è una fibra forte, inestensibile e molto flessibile.

Parametri di strutture secondarie

- **Passo** (p) = distanza tra due spire successive;
- **Altezza** (h) = proiezione della distanza testa-coda di un residuo sull'asse;
- **n** = numero di residui per giro;
- Numero di atomi per ansa chiusa da legame a ponte di H

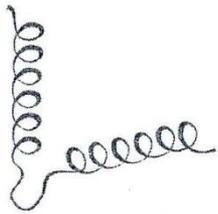
Tipo di struttura	Residui per giro	Passo per res. (nm)	Atomi per ansa chiusa da leg.H	Φ	Ψ
B Sheet antiparallelo	2	0.34		-139	+135
B Sheet parallelo	2	0.32		-119	+113
Elica 310	3	0.20	10	-49	-26
Elica alfa	3.6	0.15	13	-57	-47
Elica pigreco	4.4	0.12	16	-57	-70

Esistono anche **strutture non periodiche** tipo anse (**loop**) o curve (**turn**) che permettono alla catena di ripiegarsi su se stessa.

Un terzo circa dei residui di una proteina si trovano in queste strutture. I β -turn (4 residui, legame a H tra CO del primo e NH del quarto) solitamente connettono tratti di β -sheet

Strutture Supersecondarie o Motivi

(a) Elica-ansa-elica
(*helix-loop-helix*)



(b) Avvolgimento avvolto
(*coiled coil*)



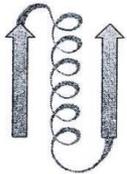
(c) Fascio di eliche
(*helix bundle*)



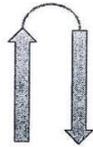
◀ **Figura 4.19**

Motivi comuni. Nelle proteine ripiegate le α -eliche e i filamenti β sono generalmente connessi da anse e curve per formare strutture supersecondarie, mostrate qui come rappresentazioni bidimensionali. Le frecce indicano la direzione da N- a C-terminale della catena peptidica.

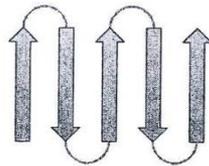
(d) Unità $\beta\alpha\beta$
($\beta\alpha\beta$ unit)



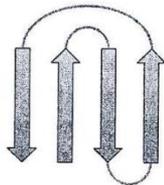
(e) Forcina
(*hairpin*)



(f) Serpentina β
(β meander)



(g) Chiave greca
(*greek key*)



(h) β -sandwich



Sono **combinazioni di α -eliche, filamenti β e anse** che si ritrovano in svariate proteine. Uno dei motivi più ricorrenti è *helix-loop-helix* ricco di Asp e Glu, proprio di differenti proteine che legano il calcio. Il motivo *helix-turn-helix* ricorre in proteine che legano il DNA. Il motivo *coiled coil* è composto di due eliche anfipatiche parallele che interagiscono attraverso le loro estremità idrofobiche (*leucine zipper*). Nel motivo *helix bundle* le α -eliche sono in orientazione opposta.

STRUTTURA TERZIARIA

Particolare avvolgimento della proteina dal quale deriva la sua forma tridimensionale

La struttura terziaria può essere caratterizzata da uno o più domini.

Un dominio è una regione di struttura terziaria compatta e ripiegata localmente. I domini sono interconnessi dalla catena polipeptidica che decorre lungo l'intera molecola.

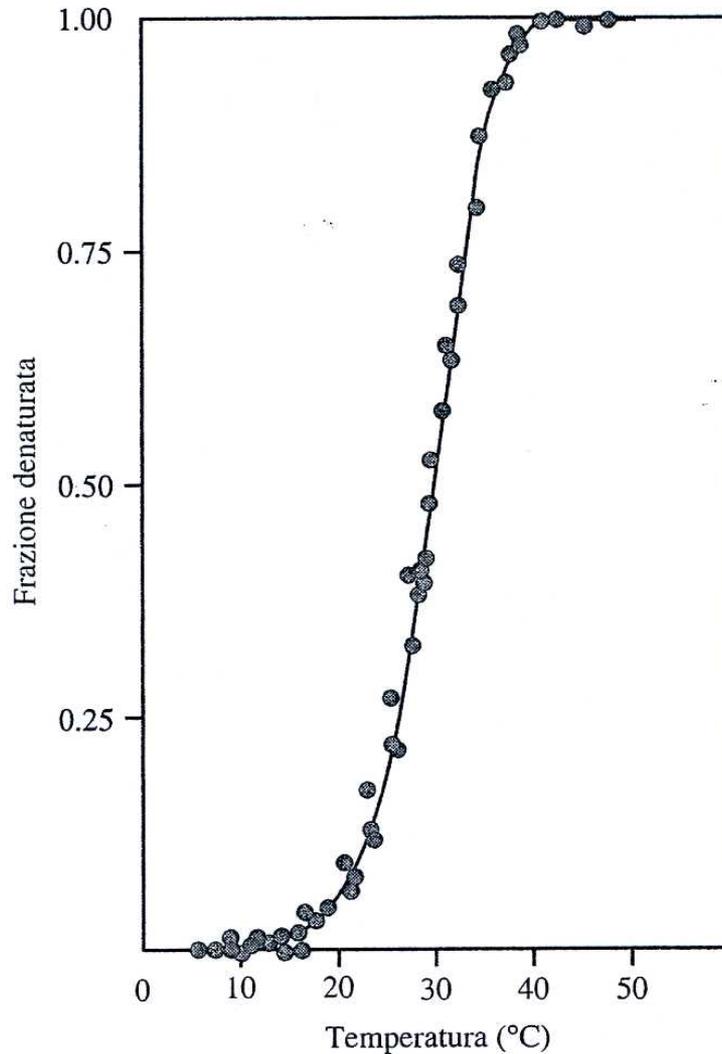
Spesso domini differenti svolgono **funzioni differenziate**.

Un dato dominio può essere riscontrato in molte proteine distinte.

Le informazioni per il corretto ripiegamento che determina la struttura tridimensionale è contenuta nella sequenza primaria. La struttura nativa può essere distrutta (**denaturazione**) da condizioni ambientali (temperatura elevata, pH estremamente acido o alcalino, solventi organici, urea). La catena assume un avvolgimento casuale e perde l'attività biologica. Alcune proteine, come la ribonucleasi, se riportate in condizioni fisiologiche ripristinano la struttura nativa recuperando l'attività biologica (processo reversibile).

E' quindi possibile che una catena polipeptidica appena sintetizzata in condizioni fisiologiche si avvolga e sia in grado di esplicare la sua funzione biologica anche se il processo *in vivo* a volte richiede aiuti per ovviare a folding errati o all'aggregazione. La quantità di energia sufficiente alla denaturazione è quella necessaria per scindere 3 legami a H. La **temperatura di fusione T_m** corrisponde alla temperatura del punto intermedio della transizione da forma nativa a forma denaturata. Le proteine sono mediamente stabili fino a 50-60°C.

Denaturazione



Denaturazione al calore della ribonucleasi A. Una soluzione di ribonucleasi A in 0.02 M KCl a pH 2.1 veniva scaldata. La denaturazione era monitorata dalla variazione di assorbanza nell'ultravioletto (blu), viscosità (rosso) e rotazione ottica (verde). [Adattato da Ginsburg, A. e Carroll, W.R. (1965). Some specific ion effects on the conformation and thermal stability of ribonuclease. *Biochemistry* 4:2159-2174.]

La perdita della conformazione nativa è un **processo cooperativo**: la destabilizzazione di poche interazioni deboli conduce alla perdita del folding

Domini

I domini possono essere composti da pochi aa (25-30 aa) fino a più di 300 aa.

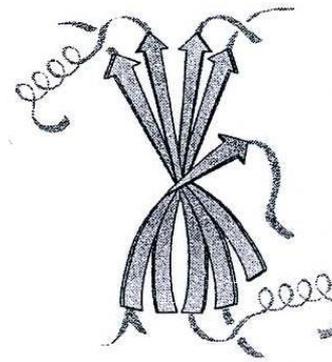
I domini sono solitamente **interconnessi da anse** ma sono spesso legati tra di loro da interazioni deboli tra le catene laterali dei residui. Alcuni domini sono presenti in molte proteine diverse mentre altri sono unici.

In genere le proteine possono essere raggruppate in **famiglie** a seconda della somiglianza dei loro domini e della loro struttura primaria.

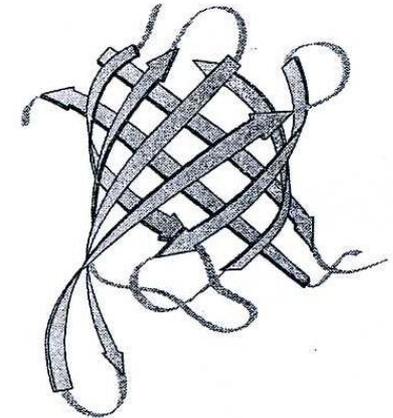
I domini possono essere classificati come solo **alfa**, solo **beta** e **alfa/beta**.

La relazione tra struttura e funzione di un dominio è complessa. Spesso un singolo dominio ha una funzione particolare tipo legare piccole molecole. Negli enzimi multifunzionali ogni attività catalitica può essere associata a uno dei molti domini presenti. Tuttavia in molti casi il legame di piccole molecole e la formazione del sito attivo di un enzima ha luogo all'interfaccia tra due domini. Le interfacce costituiscono crepe, solchi o tasche accessibili dalla superficie proteica. Molti siti di legame degli enzimi sono posizionati verso l'interno della proteina e sono relativamente liberi dall'acqua: quando i substrati si legano il loro inserimento è così preciso che alcune delle poche molecole d'acqua presenti vengono scalzate.

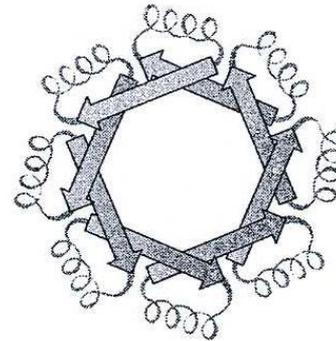
(a) Foglietti paralleli avvolti



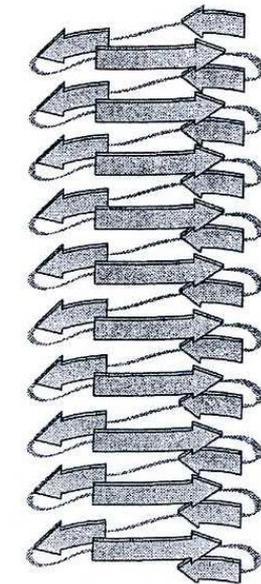
(b) Botte β



(c) Botte α/β



(d) Elica β



Folding

Le proteine ripiegate occupano uno **stato a bassa energia** che rende la struttura nativa molto **più stabile** delle strutture alternative. Quando una proteina si ripiega, le prime interazioni innescano le seguenti (**cooperatività del ripiegamento**). Quando una proteina inizia a ripiegarsi adotta energie sempre più basse e inizia a cadere in un **imbuto energetico**. Nella sua conformazione nativa, infondo all'imbuto, la proteina è molto meno sensibile alla degradazione rispetto ad una catena estesa conquistando un'emivita di molte generazioni cellulari se non di un decennio. Il folding impiega mediamente meno di un secondo. Le interazioni non covalenti sono deboli se prese singolarmente ma complessivamente conferiscono stabilità al folding nativo e nello stesso tempo flessibilità (piccoli cambiamenti conformazionali sono permessi).

Le proteine semplici come la ribonucleasi A possono ripiegarsi spontaneamente nella struttura nativa in una provetta da laboratorio senza che venga fornita energia o aiuto. Anche proteine più grandi possono raggiungere senza problemi la struttura ad energia più bassa ovvero il loro folding nativo. Tuttavia alcune grandi proteine possono restare intrappolate in un avvallamento energetico locale (che si trova tra l'inizio e il termine dell'imbuto). La presenza di **conformazioni metastabili** non corrette nel caso migliore rallenta la velocità del folding, ma in alcuni casi determina l'aggregazione e la precipitazione degli intermedi del ripiegamento. La velocità del ripiegamento corretto è aumentata da un gruppo di proteine ubiquitarie chiamate **chaperoni molecolari** (heat shock proteins)

Termodinamica del ripiegamento

Il ripiegamento di una proteina è un processo termodinamicamente favorito in condizioni fisiologiche ($\Delta G < 0$)

Il processo avviene con una diminuzione del grado di disordine dunque risulta entropicamente sfavorito ($\Delta S < 0$).

È quindi un processo entalpicamente favorito ($\Delta H < 0$). Il contributo principale ad un ΔH negativo è dato dalle interazioni energeticamente favorevoli tra i gruppi funzionali all'interno della molecola ripiegata:

➤ Interazioni tra gruppi carichi (ponti salini)

➤ Legami a ponti di idrogeno intramolecolari tra gruppi accettori e/o donatori quali: gruppi ossidrilici (Thr, Ser), gruppi amminici e ossigeni di gruppi carbonilici (Asn, Gln), azoti di His

➤ Interazioni di Van der Waals

➤ Effetto idrofobico

I composti idrofobici a contatto con l'acqua inducono la formazione da parte delle molecole d'acqua di una struttura a gabbia. Ciò porta ad un sistema più ordinato e quindi termodinamicamente sfavorito. Quando in una proteina globulare i residui idrofobici vengono nascosti all'interno della struttura ciò comporta il rilascio delle molecole che formavano la gabbia e quindi un aumento dell'entropia del sistema.

Residui differenti contribuiscono in modo differente all'effetto idrofobico. Studi di trasferimento di amminoacidi da acqua a solvente idrofobico hanno permesso la formulazione di scale di idrofobicità.

Ponti a disolfuro

Due esempi di scala di idrofobicità

Amino-acido	Scala di Engelman, Steitz e Goldman ^a	Scala di Kyte e Doolittle ^b
Phe	3.7	2.8
Met	3.4	1.9
Ile	3.1	4.5
Leu	2.8	3.8
Val	2.6	4.2
Cys	2.0	2.5
Trp	1.9	-0.9
Ala	1.6	1.8
Thr	1.2	-0.7
Gly	1.0	-0.4
Ser	0.6	-0.8
Pro	-0.2	-1.6
Tyr	-0.7	-1.3
His	-3.0	-3.2
Gln	-4.1	-3.5
Asn	-4.8	-3.5
Glu	-8.2	-3.5
Lys	-8.8	-3.9
Asp	-9.2	-3.5
Arg	-12.3	-4.5

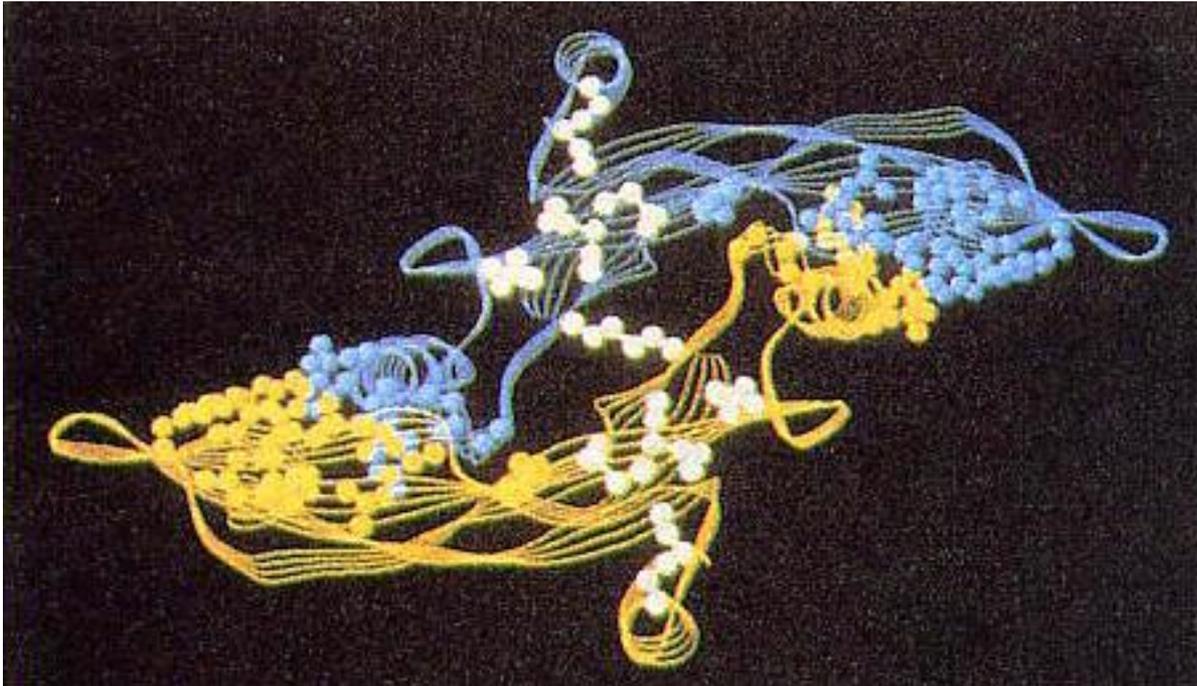
^aDati da D.M. Engelman, T.A. Steitz, and A. Goldman, *Annu. Rev. Biophys. & Biophys. Chem.* (1986) 15:321-353.

^bDati da J. Kyte, and R.F. Doolittle, *J. Mol. Biol.* (1982) 157:105-132.

STRUTTURA QUATERNARIA

Molte proteine sono di **natura multimerica** ovvero sono costituite di più subunità (catene diverse) aggregate tra di loro mediante **legami non covalenti**. Tale struttura complessa viene definita struttura quaternaria. Le subunità possono essere identiche o differenti. Se sono identiche si parla di dimeri o tetrameri.

Un metodo per classificare le subunità è quello di utilizzare lettere greche per identificare le subunità e numeri a pedice per indicare il numero delle subunità. Es $\alpha_2\beta\gamma$. Le interazioni che legano le subunità sono solitamente deboli con prevalenza di interazioni idrofobiche.



Collagene

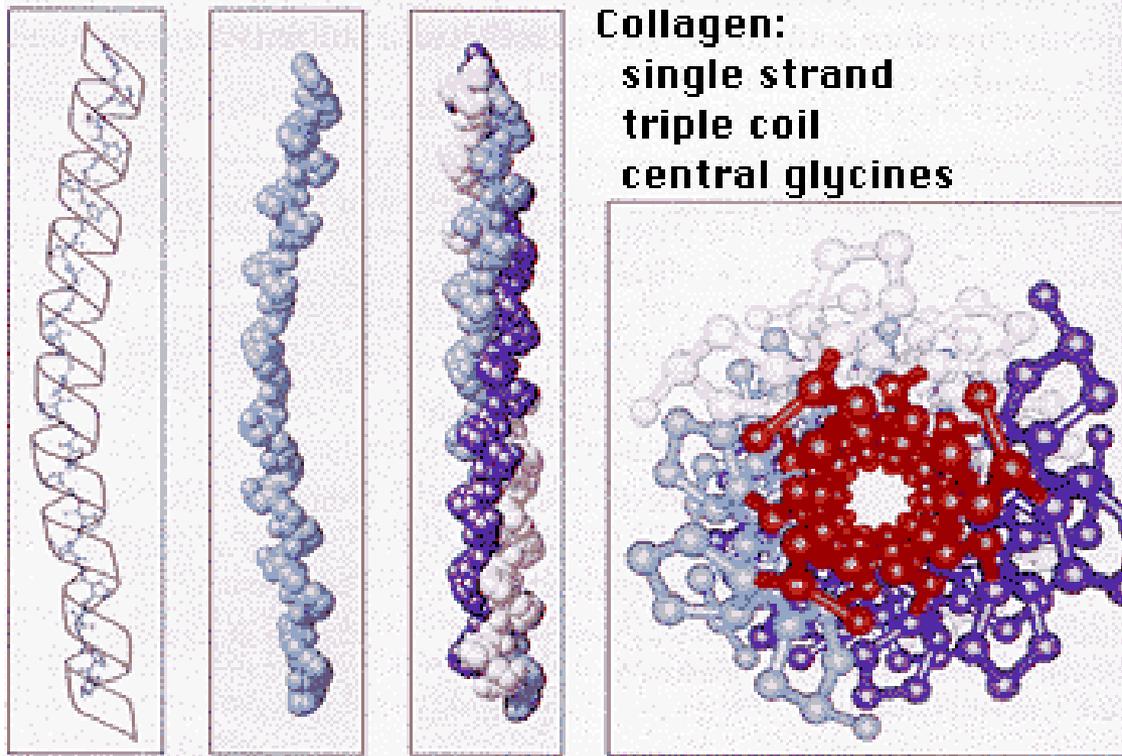
E' il maggior componente del tessuto connettivo dei vertebrati: costituisce dal 25 al 35% delle proteine totali dei mammiferi. Le molecole di collagene hanno forme e funzioni notevolmente diverse. Ad esempio il collagene dei tendini forma fibre dure simili a corde con una straordinaria resistenza alla tensione; nella pelle il collagene forma fibre intrecciate in modo lasso capaci di estendersi in tutte le direzioni.

I tipi di collagene

<i>Tipo</i>	<i>Catene</i>	<i>Struttura</i>	<i>Localizzazione</i>
I	$\alpha 1(I), \alpha 2(I)$	Fibrillare	Cute, tendini, osso, ecc.
II	$\alpha 1(II)$	Fibrillare	Cartilagine, corpo vitreo
III	$\alpha 1(III)$	Fibrillare	Cute, muscoli, ecc.
IV	$\alpha 1(IV), \alpha 2(IV)$	Non fibrillare	Tutte le membrane basali
V	$\alpha 1(V), \alpha 2(V), \alpha 3(V)$	Fibrillare	Molti tessuti interstiziali
VI	$\alpha 1(VI), \alpha 2(VI), \alpha 3(VI)$	Microfibrille	Molti tessuti interstiziali
VII	$\alpha 1(VII)$	Non fibrillare	Fibrille ancoranti
VIII	$\alpha 1(VIII)$?	Alcune cellule endoteliali
IX	$\alpha 1(IX), \alpha 2(IX), \alpha 3(IX)$?	Cartilagine
X	$\alpha 1(X)$?	Cartilagine ipertrofica e mineralizzata
XI	$\alpha 1(XI), \alpha 2(XI), \alpha 3(XI)$	Fibrillare	Cartilagine
XII	$\alpha 1(XII)$?	Cute, tendini

Collagene

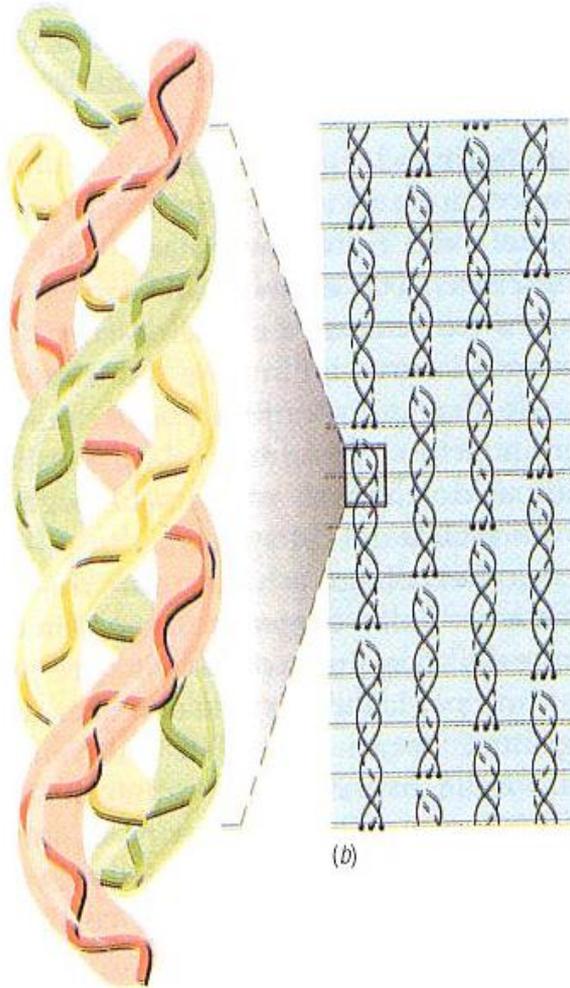
L'unità base della fibra di collagene è il **tropocollagene**, **tripla elica destrorsa di tre catene peptidiche** ciascuna lunga 1000 residui. Le singole catene sono **eliche sinistrorse con 3.3 residui per giro** (eliche più allungate rispetto all' α -elica). Nella super elica un residuo su tre deve essere interno (Gly). Sequenza ripetitiva Gly-Pro-Pro o Gly-Pro-Hyp. L'elica tripla è stabilizzata da **legami a H intercatena** (ponte a H tra NH di Gly e C=O di X della catena adiacente in Gly-X-Y); le singole eliche non hanno legami intracatena. Il collagene contiene anche idrossilisina, sito di aggancio per carboidrati.



Collagene

Le reazioni di **idrossilazione di Pro e Lys** sono catalizzate da enzimi e richiedono vitamina C (acido ascorbico). Le persone affette da scorbuto presentano lesioni alla pelle, fragilità dei vasi sanguigni, perdita dei denti e gengive sanguinanti.

Il tropocollagene si aggrega in modo sfalsato a formare fibre forti e insolubili. La forza del collagene deriva da legami trasversali: i gruppi delle catene laterali di Lys e idrossilisina sono convertiti enzimaticamente in gruppi aldeidici; questi formano basi di Schiff con altri residui di Lys fornendo legami crociati.



Mioglobina complessata con **Eme**.

La mioglobina è una proteina monomerica. L'**emoglobina** è una proteina tetramerica eteromeric ($\alpha_2\beta_2$). Il colore rosso associato alle forme ossigenate della mioglobina o dell'emoglobina è dovuto all'eme. L'eme è il gruppo prostetico di queste proteine ovvero una molecola organica essenziale alla funzione della proteina. L'eme è un anello tetrapirrolico (protoporfirina) che complessa il ferro.

