

Biotecnologie applicate alla progettazione e sviluppo di molecole biologicamente attive

A.A. 2010-2011

Modulo di Biologia Strutturale

Struttura delle Proteine

Marco Nardini

Dipartimento di Scienze Biomolecolari e Biotecnologie

Università di Milano

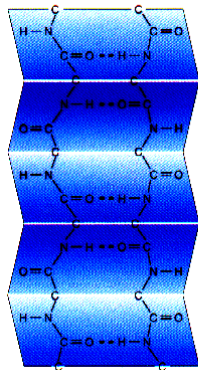
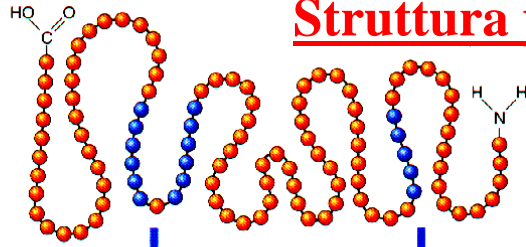
Proteine: la classe più importante dei biopolimeri

- le proteine sono le *macromolecole effettrici* dei sistemi biologici.

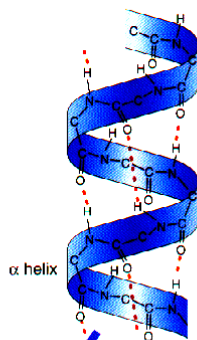
Tra le proteine esistono:

- Proteine Strutturali (collagene, cheratine)
- Proteine Contrattili (actina, miosina, tubuline)
- Proteine del Sistema Immunitario (immunoglobuline)
- Trasportatori: (emoglobina)
- Canali
- Recettori
- Alcuni Ormoni
- Enzimi

Struttura primaria



β pleated sheet



α helix

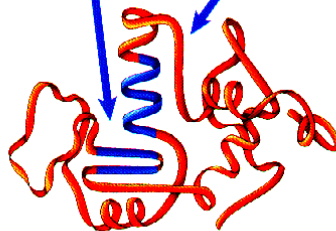


α -elica

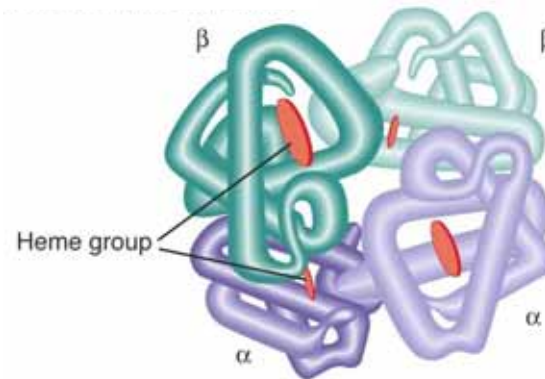


filamenti β e foglietto β

Struttura secondaria



Struttura terziaria



emoglobina

Struttura quaternaria

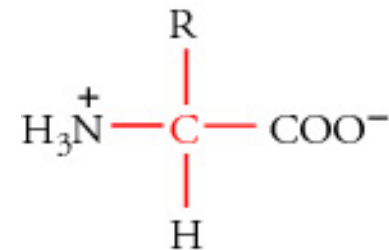
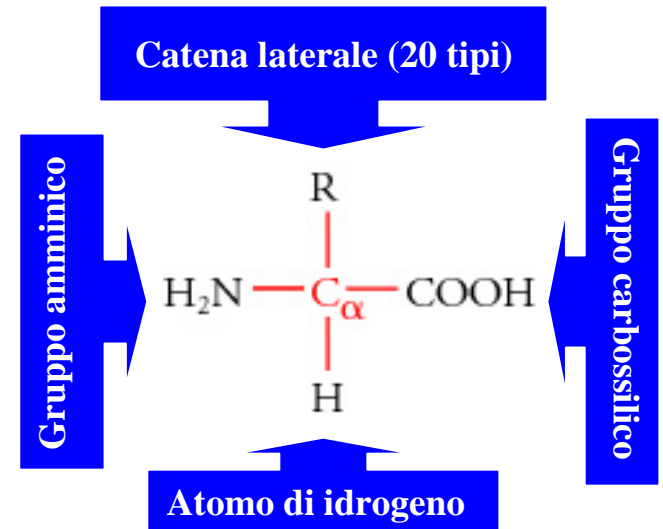
**Legame Peptidico
e
Struttura Primaria**

Amminoacidi

Le **proteine** sono polimeri lineari senza ramificazioni costituiti da **20** unità base, gli **amminoacidi**.

In condizione di pH fisiologico sia il gruppo acido carbossilico sia il gruppo amminico sono completamente ionizzati.

Molecole che possiedono gruppi carichi di opposta polarità si dicono **zwitterioni**.

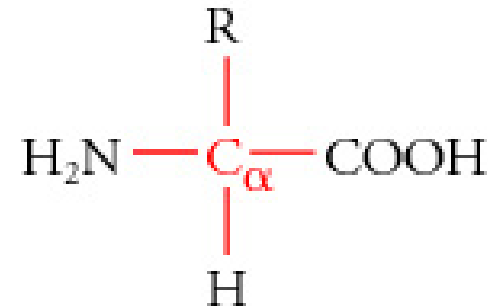


Forma zwitterionica

20 catene laterali diverse per gli amminoacidi: classificazione e caratteristiche

Le **20 diverse** catene laterali (**gruppo R**) che costituiscono gli amminoacidi si differenziano considerevolmente per dimensioni, volume e per le loro caratteristiche fisico chimiche, quali:

- **polarità**
- **acidità e basicità**
- **aromaticità**
- **flessibilità conformazionale**
- **reattività chimica**
- **tendenza a formare legami idrogeno**



Queste diverse caratteristiche sono le maggiori responsabili della grande varietà di proprietà delle proteine.

Gli amminoacidi sono generalmente classificati a seconda della **polarità** delle loro **catene laterali**

Infatti, il ripiegamento della catena polipeptidica nella sua conformazione nativa (**fold**ing) è dovuto principalmente alla **tendenza che hanno le catene laterali idrofobiche a sfuggire il contatto con il solvente e le catene laterali idrofiliche ad essere esposte all'acqua.**

Si possono quindi distinguere **3 gruppi** di amminoacidi:

- gruppo **R non polare**
- gruppo **R polare non carico**
- gruppo **R polare carico**

Per convenzione, i nomi degli amminoacidi sono abbreviati con un **codice a tre lettere** e con uno a **una lettera**.

Gruppo R non polare

Glicina	Gly	G
Alanina	Ala	A
Valina	Val	V
Leucina	Leu	L
Isoleucina	Ile	I
Metionina	Met	M
Prolina	Pro	P
Fenilalanina	Phe	F
Triptofano	Trp	W

Gruppo R polare non carico

Serina	Ser	S
Treonina	Thr	T
Asparagina	Asn	N
Glutammina	Gln	Q
Tirosina	Tyr	Y
Cisteina	Cys	C

Gruppo R polare carico

Lisina	Lys	K
Arginina	Arg	R
Istidina	His	H
Acido Aspartico	Asp	D
Acido Glutammico	Glu	E

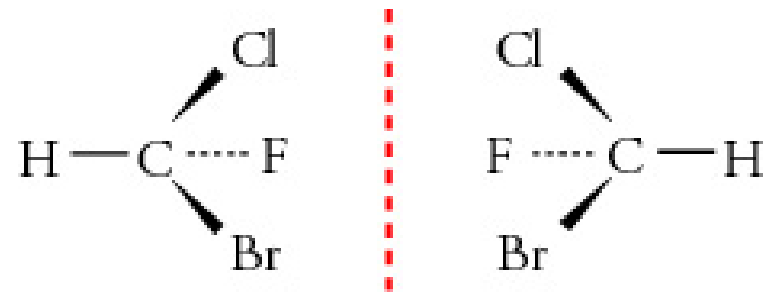
Amminoacidi

Tutti gli amminoacidi (ad eccezione della glicina) derivati da proteine sono **otticamente attivi**, cioè ruotano il piano di polarizzazione della luce.

Le molecole otticamente attive sono **asimmetriche** in modo da non essere sovrapponibili con le loro immagini speculari (**enantiomeri**).

Ciò è caratteristico di molecole contenenti atomi di carbonio tetraedrici con 4 diversi sostituenti.

L'atomo centrale si dice "centro chirale"

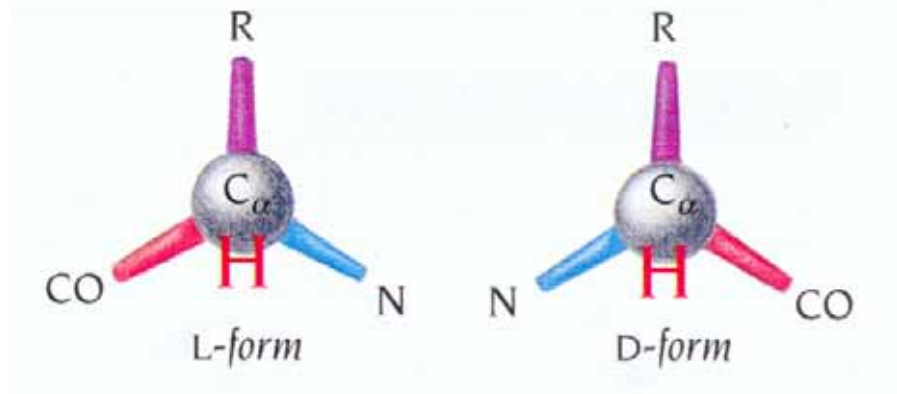


Piano dello specchio

Amminoacidi

Il centro chirale degli amminoacidi è l'atomo C_{α} .

Tutti gli amminoacidi derivati da proteine hanno una configurazione stereochimica di tipo L.



La configurazione assoluta degli amminoacidi L può essere facilmente ricordata:

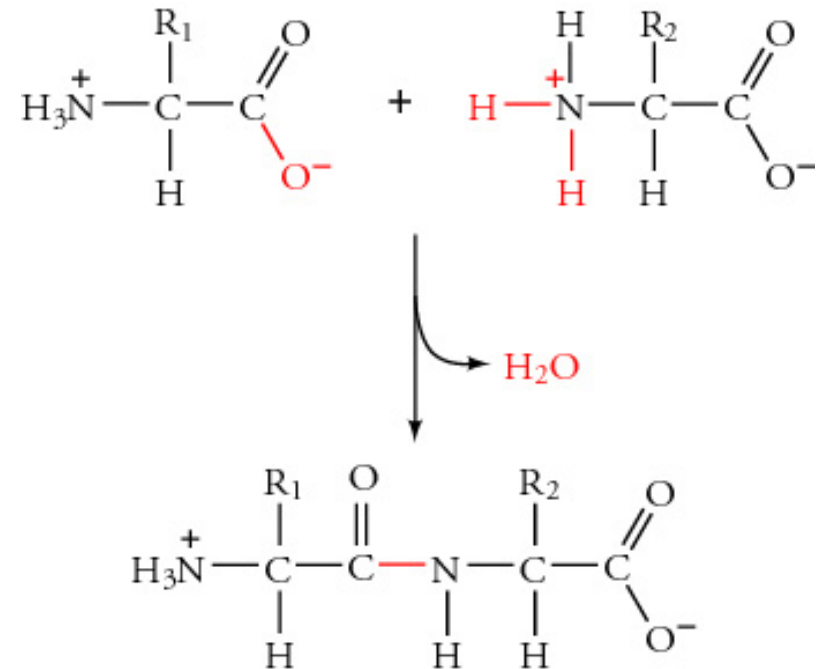
- guardando verso il C_{α} dal suo atomo H, gli altri atomi sostituenti formano, se letti in senso orario, la parola **CORN**.

Legame peptidico

Gli amminoacidi polimerizzano durante la sintesi delle proteine mediante la formazione di legami peptidici.

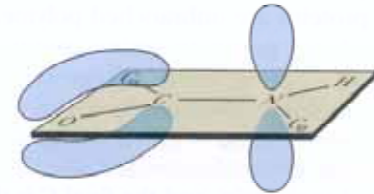
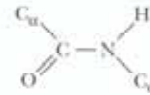
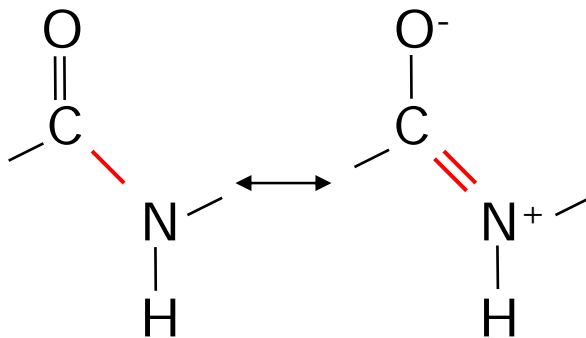
Il **legame peptidico C–N** si ha quando il gruppo carbossilico di un peptide condensa con il gruppo amminico del peptide successivo mediante l'eliminazione di una molecola d'acqua.

Nella cellula, la sintesi dei legami peptidici è controllata enzimaticamente nel ribosoma ed è guidata dal mRNA stampo

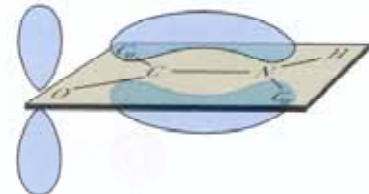
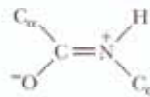


Legame peptidico

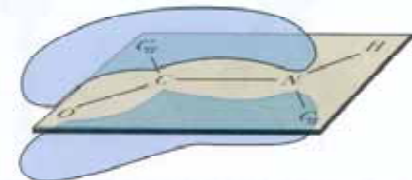
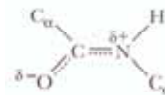
Il gruppo peptidico ha una struttura **rigida** e **planare**, dovuta al parziale (~40 %) carattere di doppio legame del legame peptidico.



A pure double bond between C and O would permit free rotation around the C—N bond.



The other extreme would prohibit C—N bond rotation but would place too great a charge on O and N.

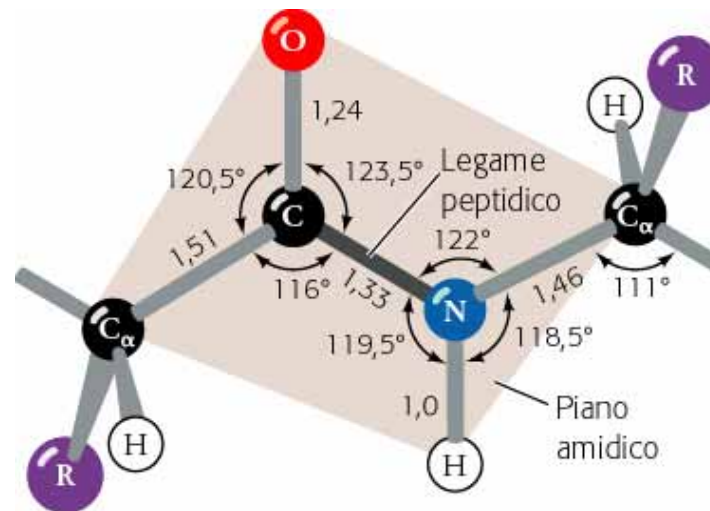


The true electron density is intermediate. The barrier to C—N bond rotation of about 88 kJ/mol is enough to keep the amide group planar.

Legame peptidico

Il legame peptidico C–N è 0.13 Å più corto del legame singolo N–C_α e 0.08 Å più lungo di un doppio legame C=N.

Il legame peptidico, quindi, presenta per il 60 % una natura di legame singolo e per il 40 % una natura di legame doppio.



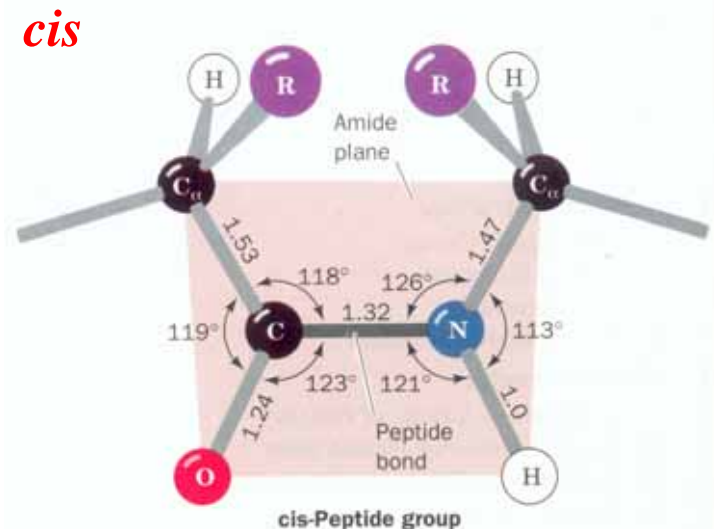
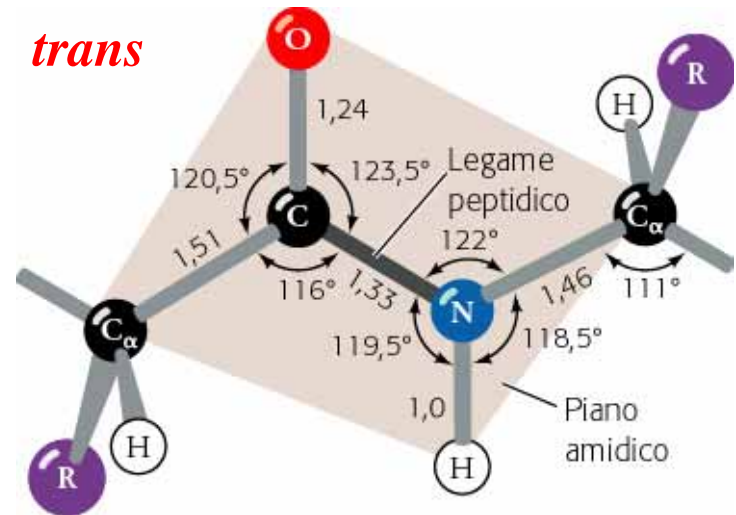
Legame peptidico

Generalmente il gruppo peptidico assume la conformazione *trans*

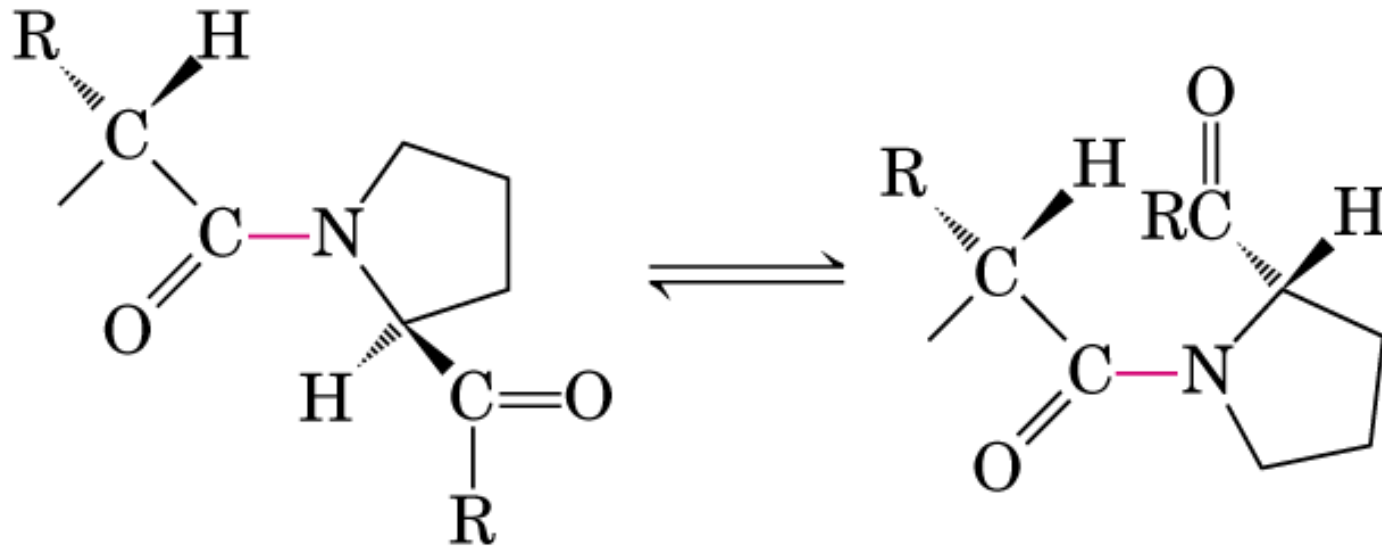
- atomi $(C_{\alpha})_n$ e $(C_{\alpha})_{n+1}$ opposti rispetto al legame peptidico C–N.

In alcuni casi il gruppo peptidico può assumere la conformazione *cis*

~8 kJ·mol⁻¹ meno stabile della conformazione *trans*. Problemi sterici, perché distanza C_{α} - C_{α} = 2.8 Å).



Legame peptidico



trans

cis

Proline isomers

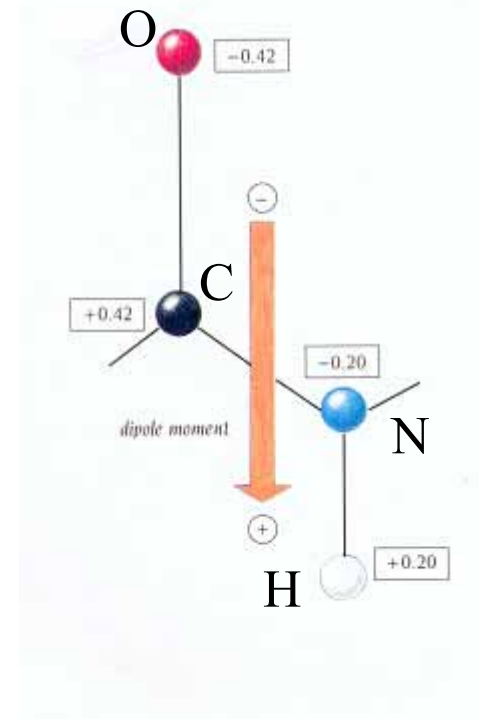
Nel caso in cui il residuo che fornisce il gruppo amminico di un legame peptidico sia una **Pro**, la stabilità reciproca delle due conformazioni del gruppo peptidico è di 4:1.

Legame peptidico

Il gruppo peptidico ha un **momento di dipolo**.

Gli atomi O e N sono rispettivamente più elettronegativi di C ed H. La conseguente delocalizzazione di carica porta alla formazione di due dipoli (CO ed NH) con analoga direzione e verso nel gruppo peptidico.

Il momento di dipolo risultante è di circa **3.5 Debye**.



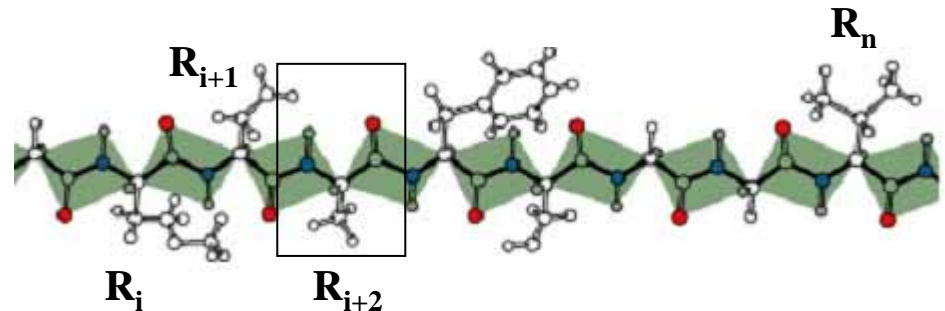
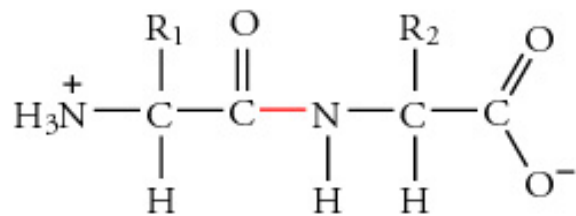
Catena Principale

I polipeptidi sono catene lineari in cui ciascun aminoacido è legato al suo prossimo in modo “*coda-testa*”, senza formare diramazioni

La catena si estende dal suo termine amminico (**N-terminale**) al suo termine carbossilico (**C-terminale**)

La formazione in successione di legami peptidici genera la cosiddetta “*catena principale*” o “*scheletro polipeptidico*” da cui protrudono le catene laterali dei diversi aminoacidi.

L'unità base ripetuta lungo la catena principale è (**NH–C_αH–C=O**), che deriva dalle parti comuni degli aminoacidi dopo che si è formato il legame peptidico.



Le infinite possibili strutture primarie

Per ciascun amminoacido costituente la catena polipeptidica si presentano **20** diverse **possibilità di catena laterale** (o **residui**), per cui è facile immaginare l'enorme numero di diverse catene polipeptidiche che possono essere costituite

dipeptide \Rightarrow si avranno $20^2 = 400$ possibili dipeptidi diversi

tripeptide \Rightarrow si avranno $20^3 = 8000$ possibili tripeptidi diversi

Nel caso delle proteine, una piccola proteina è costituita da una singola catena polipeptidica di circa **100 residui**

$\Rightarrow 20^{100} = 1.27 \times 10^{130}$ possibili catene polipeptidiche diverse!

Gli organismi sulla terra sintetizzano un gran numero di proteine, con caratteristiche fisico-chimiche differenti, che derivano dalle diverse proprietà dei 20 amminoacidi 'standard' e da come questi si combinano nella catena polipeptidica.

mioglobina

Numero	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
Uomo	G	L	S	D	G	E	W	Q	L	V	L	N	V	W	G			
Capodoglio	V	L	S	E	G	E	W	Q	L	V	L	H	V	W	A			
Numero	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30			
Uomo	K	V	E	A	D	I	P	G	H	G	Q	E	V	L	I			
Capodoglio	K	V	E	A	D	V	A	G	H	G	Q	D	I	L	I			
Numero	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45			
Uomo	R	L	F	K	G	H	P	E	T	L	E	K	F	D	K			
Capodoglio	R	L	F	K	S	H	P	E	T	L	E	K	F	D	R			
Numero	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60			
Uomo	F	K	H	L	K	S	E	D	E	M	K	A	S	E	D			
Capodoglio	F	K	H	L	K	T	E	A	E	M	K	A	S	E	D			
Numero	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75			
Uomo	L	K	K	H	G	A	T	V	L	T	A	L	G	G	I			
Capodoglio	L	K	K	H	G	V	T	V	L	T	A	L	G	A	I			
Numero	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90			
Uomo	L	K	K	K	G	H	H	E	A	E	I	K	P	L	A			
Capodoglio	L	K	K	K	G	H	H	E	A	E	L	K	P	L	A			
Numero	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105			
Uomo	Q	S	H	A	T	K	H	K	I	P	V	K	Y	L	E			
Capodoglio	Q	S	H	A	T	K	H	K	I	P	I	K	Y	L	E			
Numero	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120			
Uomo	F	I	S	E	C	I	I	Q	V	L	Q	S	K	H	P			
Capodoglio	F	I	S	E	A	I	I	H	V	L	H	S	R	H	P			
Numero	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135			
Uomo	G	D	F	G	A	D	A	Q	G	A	M	N	K	A	L			
Capodoglio	G	N	F	G	A	D	A	Q	G	A	M	N	K	A	L			
Numero	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153
Uomo	E	L	F	R	K	D	M	A	S	N	Y	K	E	L	G	F	Q	G
Capodoglio	E	L	F	R	K	D	I	A	A	K	Y	K	E	L	G	Y	Q	G

- ogni proteina di ogni organismo è dotata di una **sequenza definita**
- la sequenza di una proteina costituisce la sua **struttura primaria**
- tutte le molecole di una data proteina di un certo organismo hanno la stessa struttura primaria
- proteine corrispondenti (**omologhe**) di organismi appartenenti a specie diverse presentano strutture primarie simili (diverse solo parzialmente)

Legenda:

- Aminoacidi identici
- Sostituzioni conservative
- Sostituzioni non conservative

Escherichia coli: circa 3000 proteine diverse

Homo sapiens: 50,000 –100,000 proteine diverse

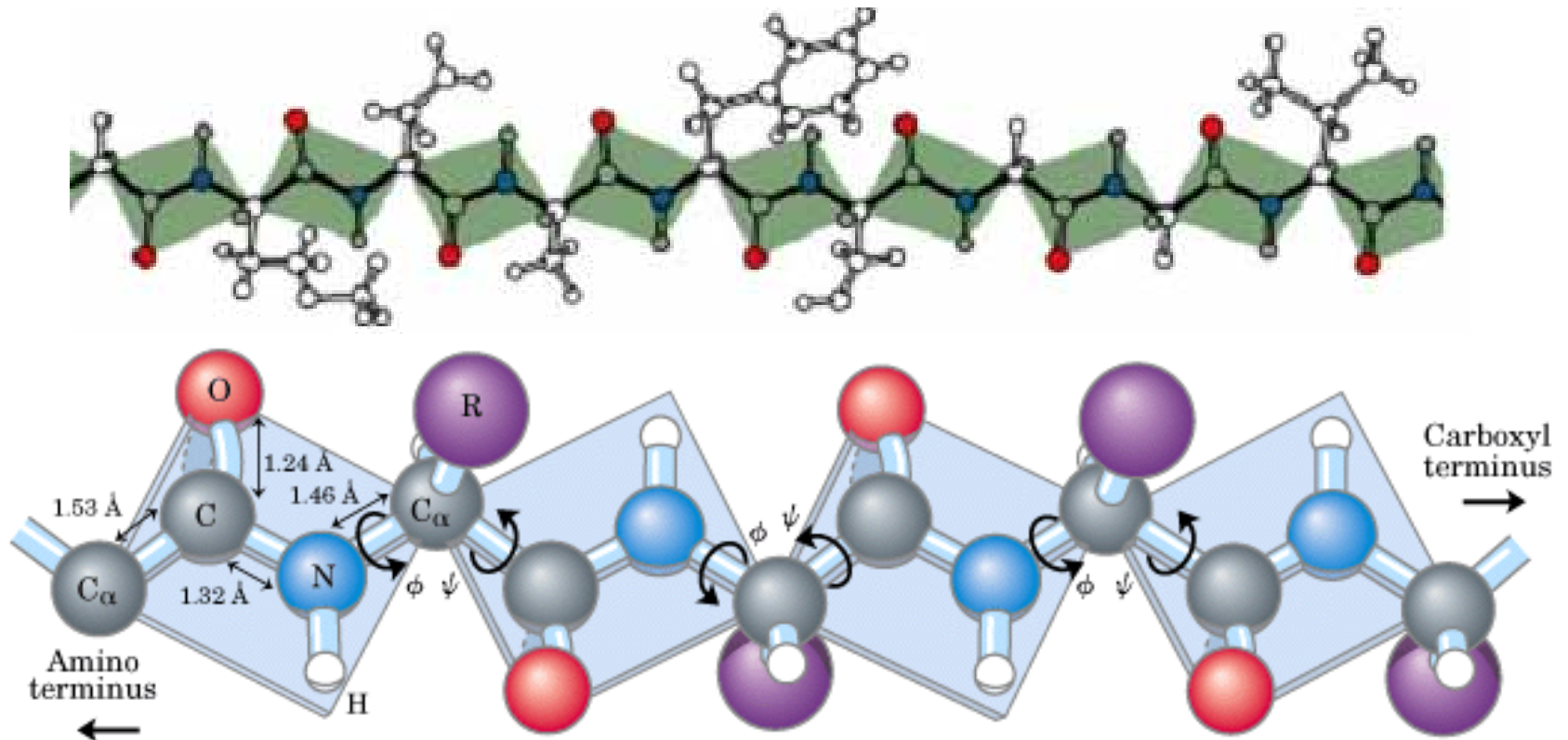
Proteine diverse con **struttura primaria simile** hanno una **struttura tridimensionale simile** e, molto spesso, proprietà e funzioni simili.

Alterazioni della struttura primaria di una proteina (sostituzione anche di un solo residuo amminoacidico) ne possono **alterare drasticamente la funzionalità**

Diagramma di Ramachandran

Catena polipeptidica

La formazione in successione di legami peptidici genera la cosiddetta “*catena principale*” o “*scheletro polipeptidico*” da cui protrudono le catene laterali dei diversi amminoacidi.



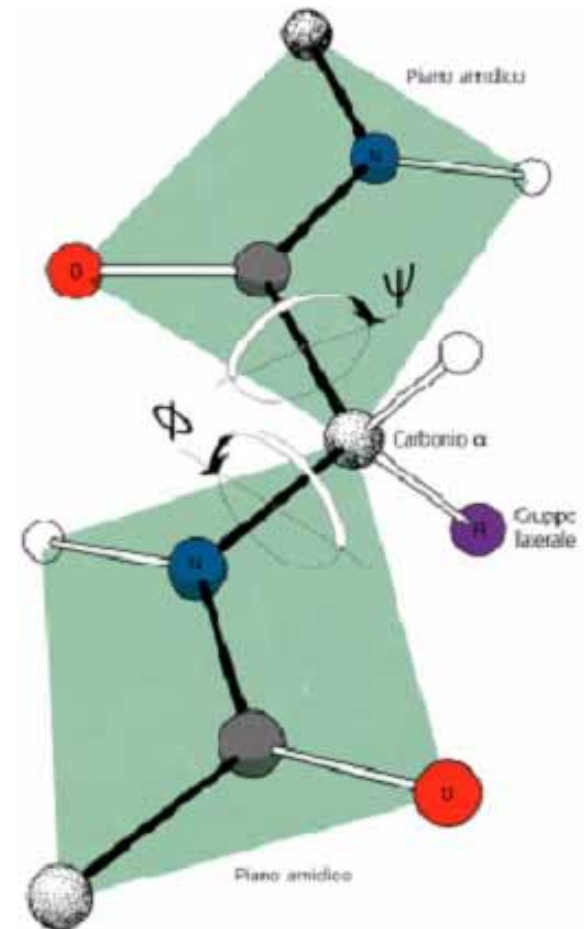
Angoli diedri ϕ e ψ

Proprietà fondamentale del **legame peptidico** è di essere **rigido e planare**.

I soli **gradi di libertà** dell'unità peptidica rigida sono:

- rotazione intorno al legame N–C_α (**angolo ϕ**)
- rotazione intorno al legame C_α–C (**angolo ψ**)

A ciascun amminoacido viene associata una coppia di angoli diedri (ϕ , ψ), che determina in modo univoco la conformazione della catena principale.

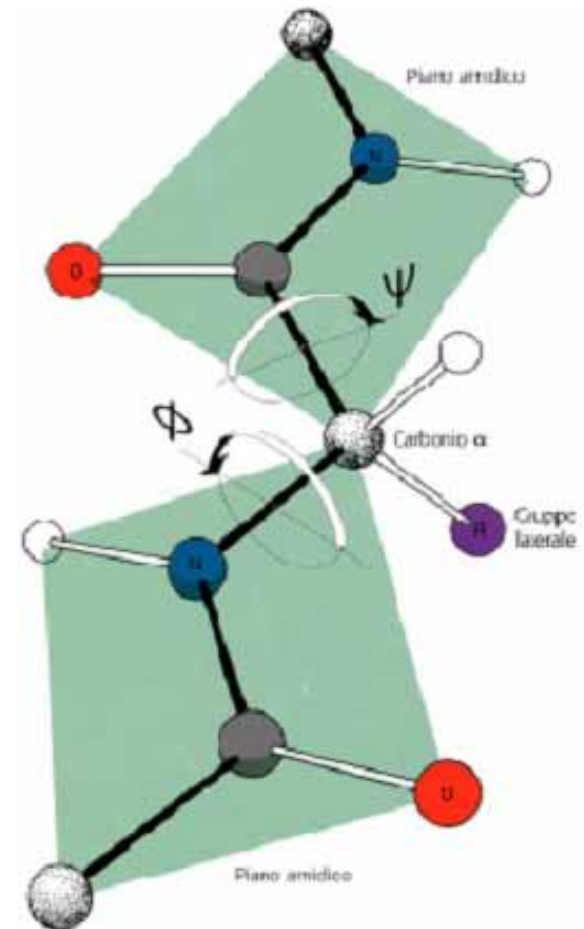


Angoli diedri ϕ e ψ

Per convenzione, agli angoli diedri ϕ e ψ viene assegnato valore di 0° quando la catena polipeptidica è nella sua conformazione planare completamente estesa, e aumentano quando si ruota in senso orario (quando visti dal C_α).

Esistono **vincoli sterici** sugli angoli di torsione ϕ e ψ di una catena polipeptidica, che ne limitano la variabilità conformazionale.

La struttura elettronica dei legami singoli C–C e N–C farebbe pensare ad una completa libertà di rotazione intorno ad essi.



Angoli diedri ϕ e ψ

Molte combinazioni di angoli (ϕ , ψ) per gli amminoacidi **non sono permesse**, a causa delle collisioni steriche fra atomi della catena principale (e/o delle catene laterali).

I valori di (ϕ , ψ) stericamente permessi si possono determinare calcolando le distanze fra gli atomi di un tripeptide in corrispondenza di tutti i valori (ϕ , ψ) per l'unità peptidica centrale.

Le conformazioni stericamente proibite sono quelle per cui la distanza interatomica di un'interazione non covalente è inferiore alla corrispondente somma di raggi di van der Waals.

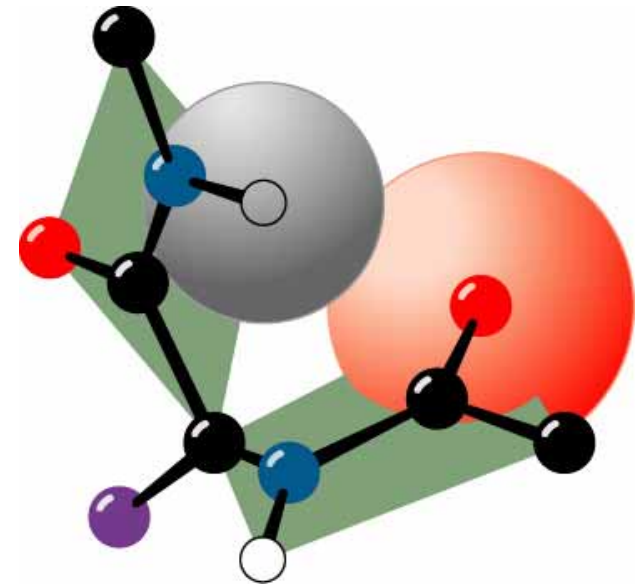


Diagramma (“Plot”) di Ramachandran

Le coppie (ϕ , ψ) permesse per gli amminoacidi sono riportate in un diagramma chiamato **plot di Ramachandran**, dal nome del fisico indiano G.N. Ramachandran che per primo (fine anni '60) calcolò le regioni stericamente permesse.

Ramachandran e colleghi usarono il modello che schematizza gli atomi come sfere rigide e fissa le geometrie dei legami.

Misure più precise del plot di Ramachandran risultano dal calcolo delle energie relative a ciascuna conformazione, permettendo una certa flessibilità per i contatti di van der Waals e per gli angoli di legame, e calcolando i contributi di tutte le interazioni favorevoli e non.

Diagramma di Ramachandran

Diagramma di Ramachandran relativo ad un tripeptide costituito da Ala.

Qualunque scelta accettabile si faccia del raggio di van der Waals, si ottengono comunque solo **3 regioni del plot di Ramachandran** fisicamente **accessibili** alla catena polipeptidica.

Le zone in azzurro sono quelle largamente permesse, mentre le zone in verde indicano le conformazioni, comunque possibili, che hanno distanze di van der Waals al limite inferiore dell'accettabilità.

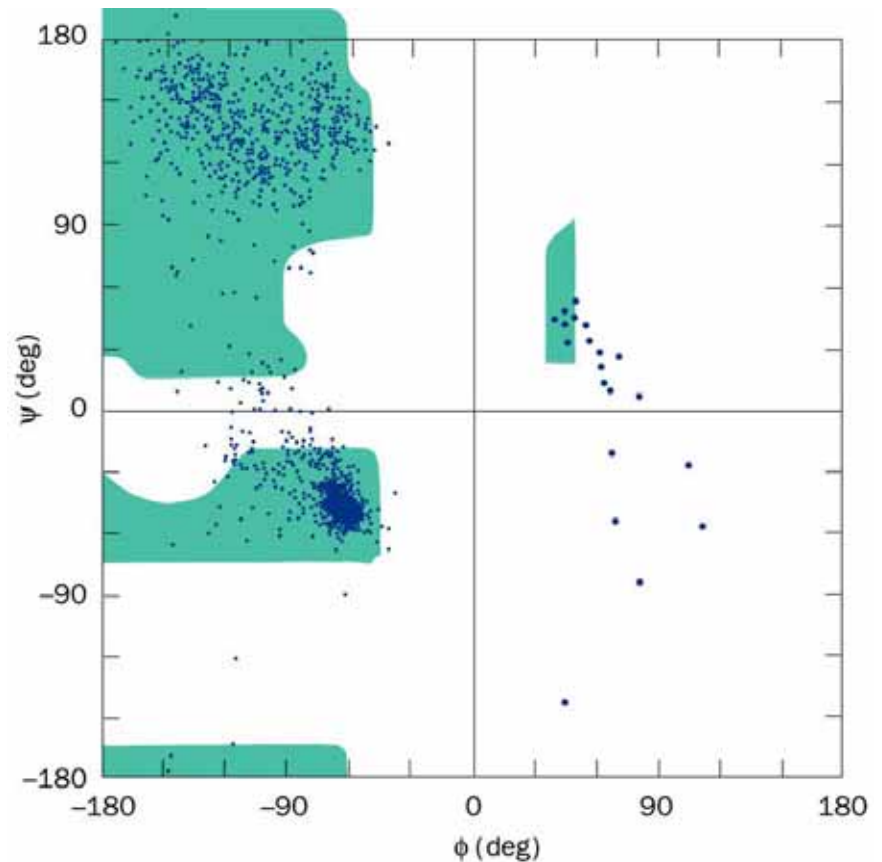


Diagramma di Ramachandran

Solo il **25-30 %** del diagramma è occupato da zone stericamente permesse.

Analoghi risultati valgono anche per i plot di Ramachandran calcolati per gli altri amminoacidi (escluse Gly e Pro).

La delimitazione delle zone permesse nel plot di Ramachandran dipende dalla scelta del raggio di van der Waals usato per calcolarle.

TABLE 7-1. VAN DER WAALS DISTANCES FOR INTERATOMIC CONTACTS

Contact Type	Normally Allowed (Å)	Outer Limit (Å)
H ··· H	2.0	1.9
H ··· O	2.4	2.2
H ··· N	2.4	2.2
H ··· C	2.4	2.2
O ··· O	2.7	2.6
O ··· N	2.7	2.6
O ··· C	2.8	2.7
N ··· N	2.7	2.6
N ··· C	2.9	2.8
C ··· C	3.0	2.9
C ··· CH ₂	3.2	3.0
CH ₂ ··· CH ₂	3.2	3.0

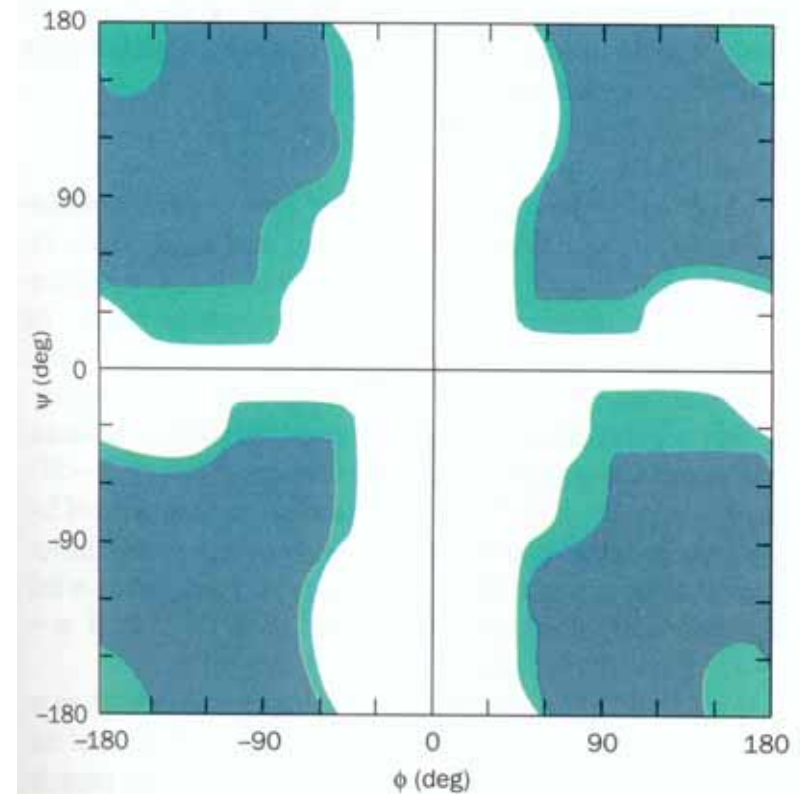
Source: Ramachandra, G.N. and Sasisekharan, V., *Adv. Protein Chem.* 23, 326 (1968).

Gly e Pro

Un caso a parte è rappresentato dalla **glicina**: le zone permesse occupano fino al 45% del plot di Ramachandran.

Il plot di Ramachandran per la glicina è centrosimmetrico, riflettendo il fatto che la glicina è l'unico amminoacido a non essere asimmetrico.

La glicina gioca un ruolo strutturale molto importante, permettendo conformazioni inusuali alla catena principale delle proteine.



Gly e Pro

Una situazione opposta a quella della glicina è rappresentata dall'amminoacido **prolina**, la cui catena laterale ciclica limita i valori di ϕ intorno a -60° ($\pm 25^\circ$).

Ciò rende la prolina l'amminoacido con maggiori restrizioni conformazionali.

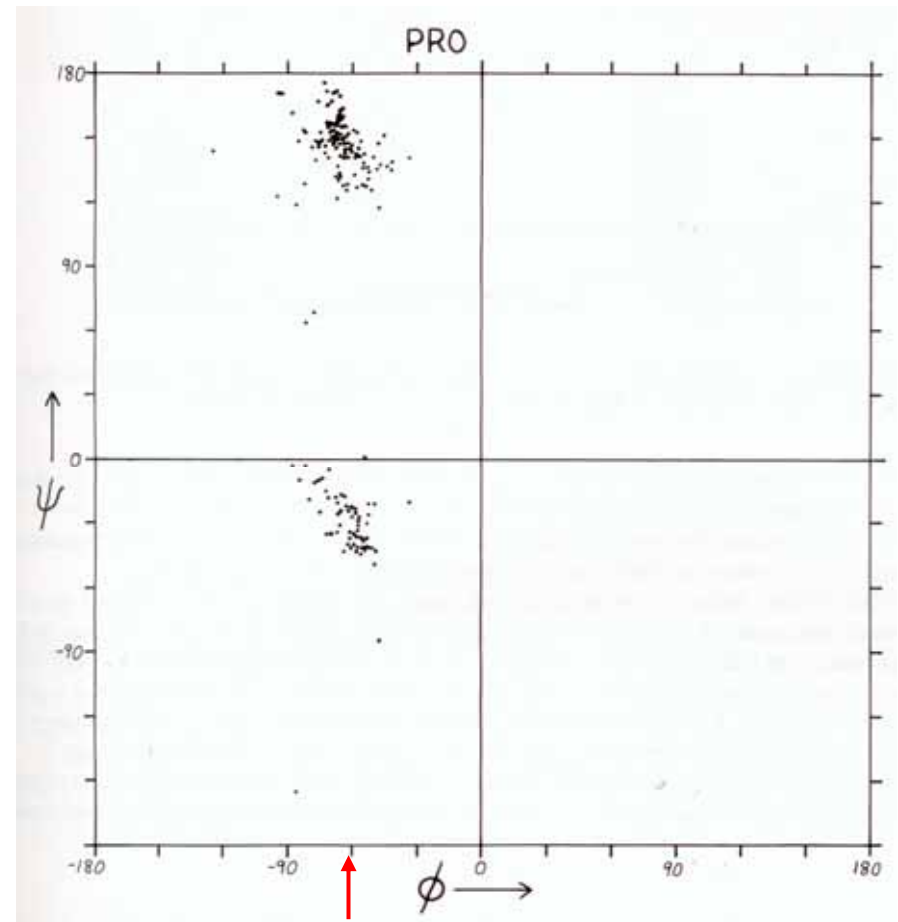
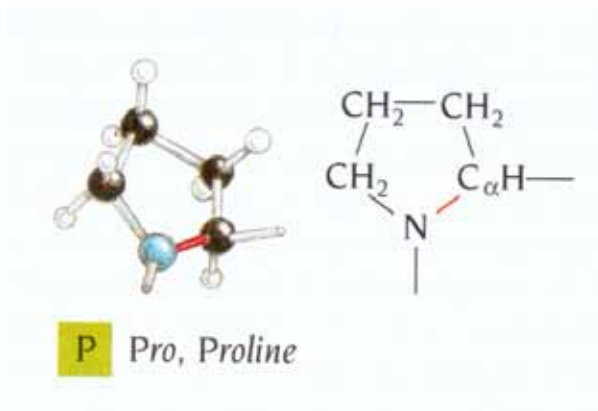
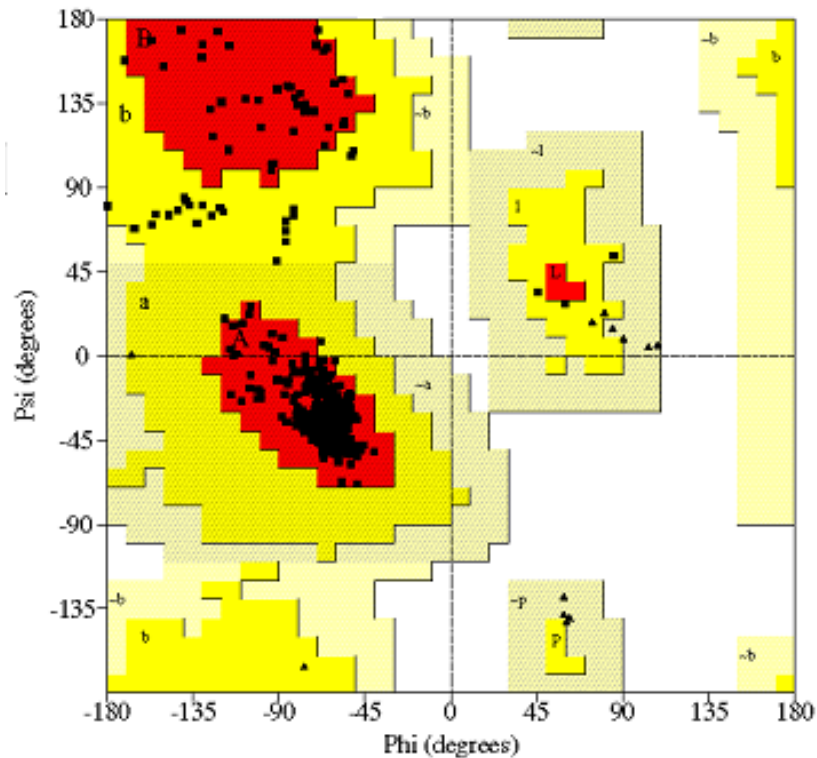


Diagramma (“plot”) di Ramachandran



Plot statistics

Residues in most favoured regions [A,B,L]	469	94.2%
Residues in additional allowed regions [a,b,l,p]	29	5.8%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%

Number of non-glycine and non-proline residues	498	100.0%
Number of end-residues (excl. Gly and Pro)	13	
Number of glycine residues (shown as triangles)	34	
Number of proline residues	35	

Total number of residues	580	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Struttura Secondaria

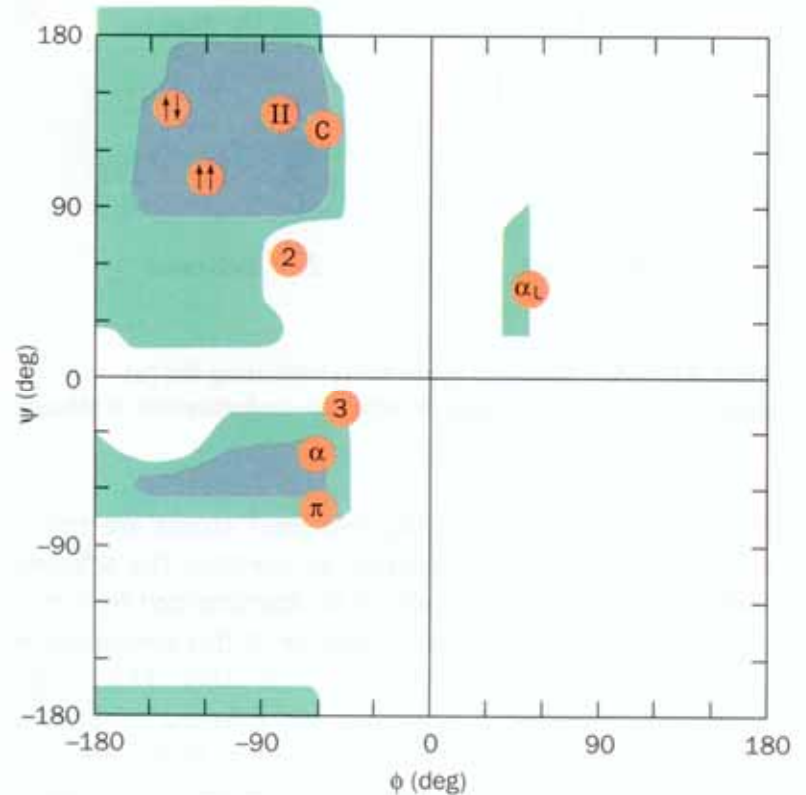
Struttura secondaria

Le regioni del diagramma di Ramachandran sono indicate con il nome della conformazione risultante in un peptide se i corrispondenti angoli (ϕ , ψ) sono ripetuti negli amminoacidi successivi lungo la catena polipeptidica.

α -eliche destrorse

filamenti β antiparalleli e paralleli

α -eliche sinistrorse



Struttura secondaria

Valori di (ϕ , ψ) relativi a ciascuna struttura secondaria

Secondary Structure	ϕ (deg)	ψ (deg)
Right-handed α helix (α)	-57	-47
Parallel β pleated sheet ($\uparrow\uparrow$)	-119	113
Antiparallel β pleated sheet ($\uparrow\downarrow$)	-139	135
Right-handed 3_{10} helix (3)	-49	-26
Right-handed π helix (π)	-57	-70
2.2 ₇ ribbon (2)	-78	59
Left-handed polyglycine II and poly-L-proline II helices (II)	-79	150
Collagen (C)	-51	153
Left-handed α helix (α_L)	57	47

[After Flory, P.J., *Statistical Mechanics of Chain Molecules*, p. 253, Interscience (1969); and IUPAC-IUB Commission on Biochemical Nomenclature, *Biochemistry* **9**, 3475 (1970).]

Struttura secondaria

- le **strutture secondarie** sono disposizioni regolari locali della catena polipeptidica principale, che vengono classificate senza fare riferimento al tipo di catena laterale degli amminoacidi
- la regolarità della conformazione risulta dalla regolarità della struttura atomica della catena polipeptidica, ed è evidenziata dai valori degli **angoli diedri** (ϕ , ψ) di ciascun gruppo peptidico.
- la struttura secondaria descrive la conformazione della catena principale di ciascun residuo amminoacidico di quel tratto rispetto a quelli adiacenti (**conformazione locale**)
- essa è stabilizzata da **legami a H** fra il **gruppo amminico** e il **gruppo carbonilico** di determinati gruppi peptidici.

Struttura Secondaria sperimentalmente osservata:

Si distinguono fondamentalmente **tre tipi di strutture secondarie**:

α elica

foglietto β

reverse turn

- circa il 70-80 % degli amminoacidi delle proteine globulari assume la conformazione regolare tipica di uno dei tre tipi di struttura secondaria

- i segmenti di catena polipeptidica che non sono in α elica, foglietto β o turn, assumono la conformazione chiamata **loop** o '**random coil**', struttura non ripetitiva né regolare, spesso priva di legami idrogeno tra gli amminoacidi che la compongono

α elica

- struttura secondaria più frequentemente adottata dalla catena polipeptidica delle proteine (~32-38% degli amminoacidi delle proteine globulari di struttura 3D nota)

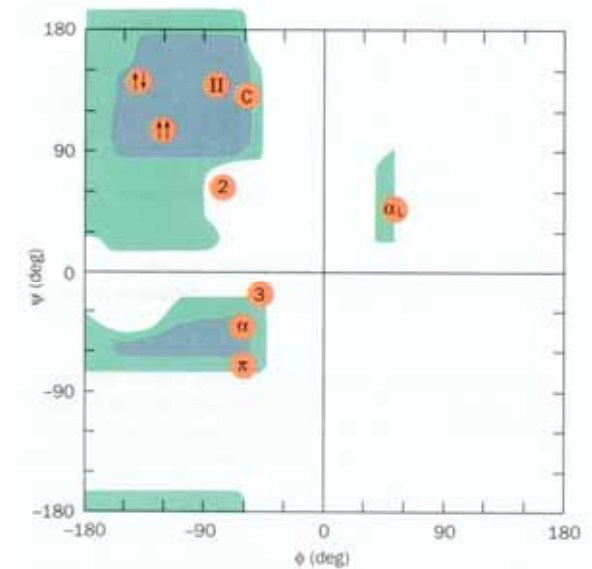
- una α elica destrorsa è caratterizzata da:

$$(\phi, \psi) = (-57^\circ, -47^\circ)$$

$$n = 3.6 \text{ residui per giro}$$

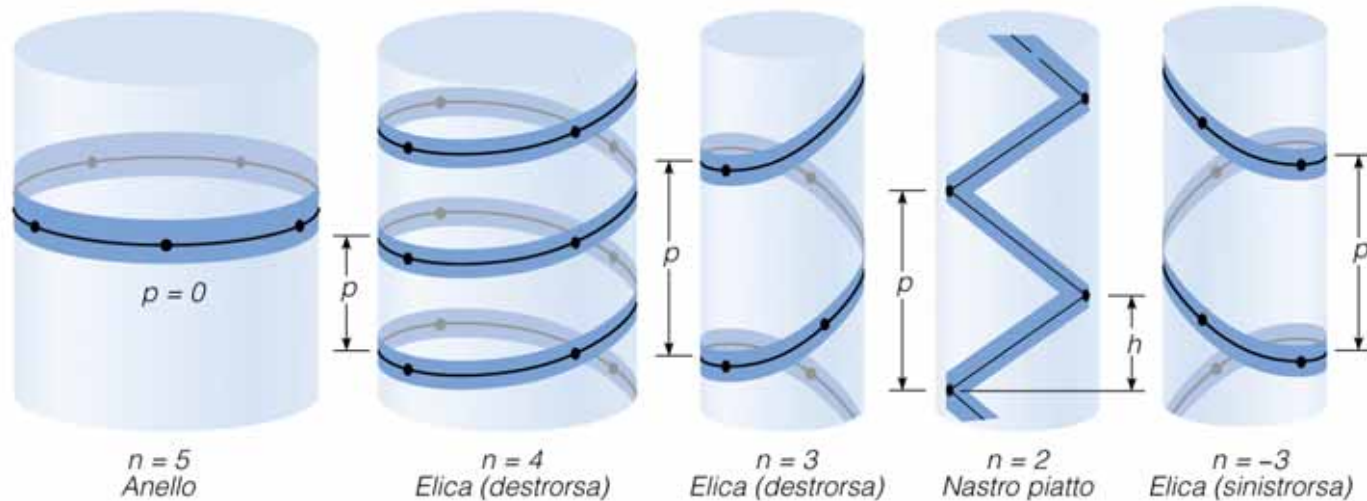
$$p = 5.4 \text{ \AA}$$

$$d = 1.5 \text{ \AA}$$



- le α eliche di tipo sinistrorso molto rare e limitate a pochi amminoacidi e sono caratterizzate da valori $(\phi, \psi) = (+57^\circ, +47^\circ)$ e da stessi valori per n e p rispetto a quelle destrorse (ma n negativo).

Parametri geometrici delle Strutture Secondarie regolari



eliche idealizzate

Parametri per la descrizione di un'elica:

n : numero di residui per giro d'elica

p : passo dell'elica

h : passo per residuo

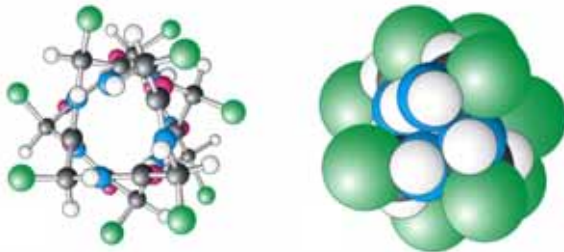
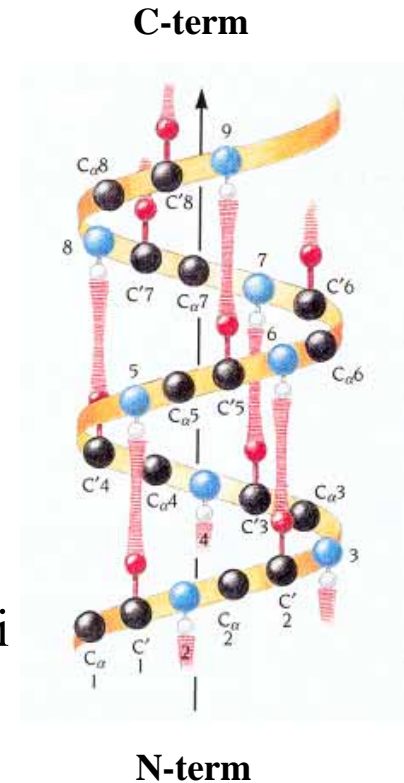
segno di n : verso dell'elica (positivo = destrorso, negativo = sinistrorso)

“regola della mano destra”

$$h = \frac{p}{n}$$

α elica

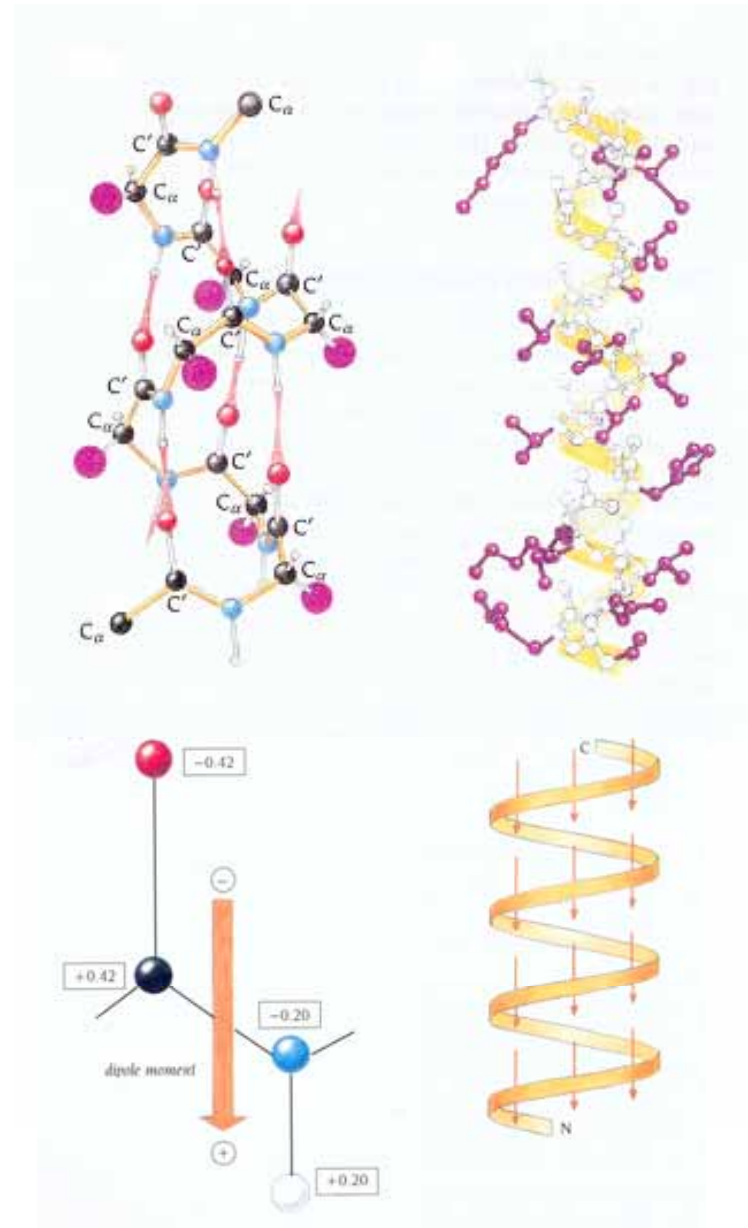
- legami a H fra il gruppo carbonilico dell'amminoacido n e il gruppo amminico dell'amminoacido $n+4$.
(distanza $N\cdots O$ quasi ottimale, di 2.8 Å)
- legami a H unidirezionali paralleli all'asse dell'elica
- tutti i gruppi NH e CO saranno legati da legami a H, eccetto i primi 3 gruppi NH e gli ultimi 3 gruppi CO alle estremità dell'elica
- ottimizzazione interazioni di van der Waals fra gli atomi lungo l'intero asse dell'elica



\Rightarrow struttura fortemente impaccata dell' α elica
(schematizzata come cilindro pieno)

Struttura Secondaria

- catene laterali rivolte verso l'esterno e verso l'estremità N-term dell'elica (struttura ad “*albero di Natale*”)
 - ⇒ minimizzazione interferenze steriche con catena polipeptidica principale e catene laterali
- una **α elica sinistrorsa** si trova raramente nelle proteine (3-5 amminoacidi) perché le catene laterali interferirebbero stericamente con la catena principale
- all' α elica è associato un **macro dipolo** con momento di dipolo totale $\sim n \times 3.5$ Debye (n = numero aa in elica)
 - ⇒ parziale carica positiva e negativa rispettivamente alle estremità N- e C-terminali dell' α elica (attrazione di ligandi di carica opposta)



Propensione ad α -elica

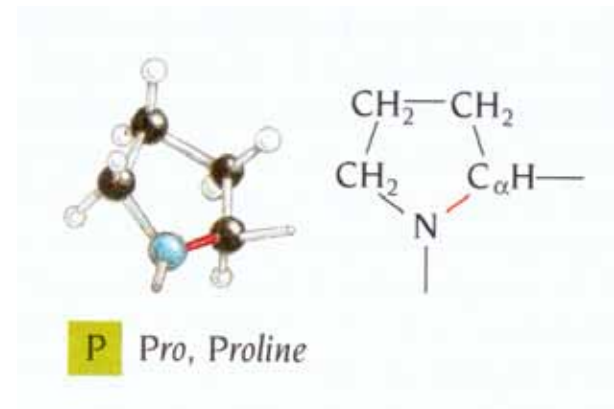
Ciascuno dei 20 amminoacidi mostra una specifica **propensione**, o meno, per assumere conformazione in α elica.

Ala, Glu, Leu, Met: buoni iniziatori di α eliche

Gly, Tyr, Ser: deboli iniziatori di α eliche

Pro: impossibile la formazione del legame a H con il gruppo CO.

La **prolina** si adatta bene nel **primo giro** dell' α elica, mentre se si trova in qualunque altra posizione nell'elica di solito produce una significativa distorsione dell'asse dell'elica dalla linearità (circa 25°).

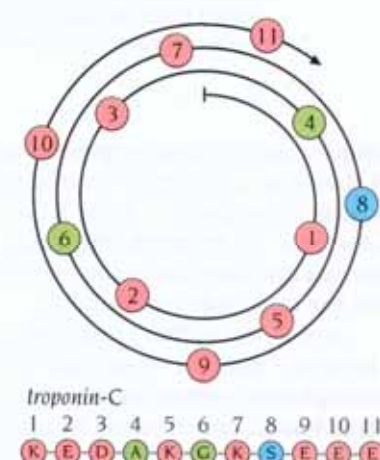
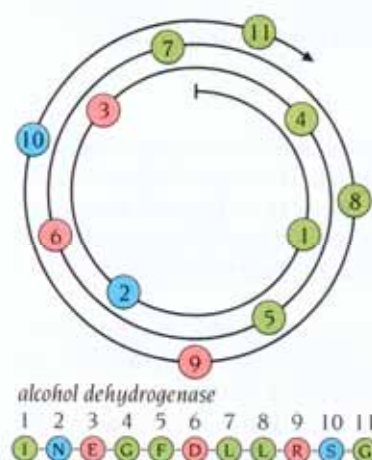
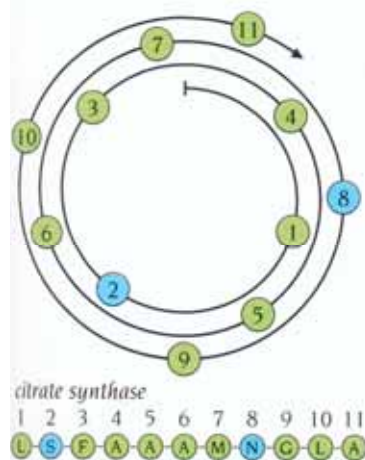


Distribuzione amminoacidi in α -elica

Le α eliche possono essere:

- completamente esposte al solvente
- completamente sepolte nel cuore della proteina
- sulla parte esterna della superficie proteica
(un lato affacciato sul solvente e l'altro all'interno idrofobico della proteina)

⇒ tendenza delle catene laterali a cambiare da idrofobiche a idrofiliche con una periodicità di 3-4 residui.



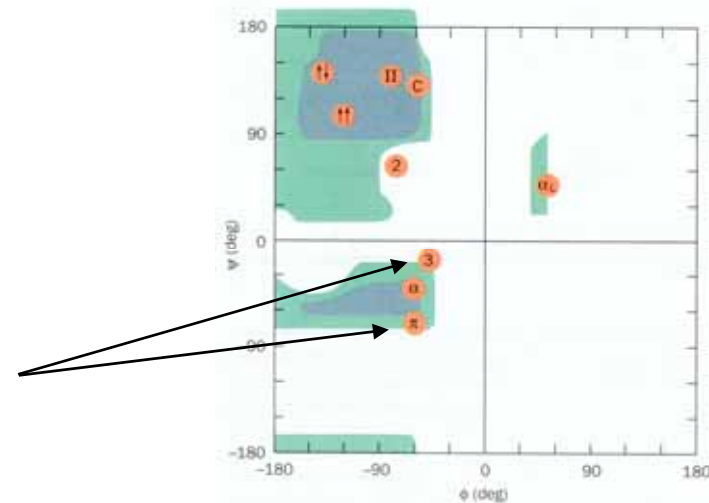
Altri tipi di elica

Variazioni dell' α elica, in cui la catena polipeptidica è avvolta più o meno strettamente, con legami a H fra coppie di amminoacidi ($n \rightarrow n+3$) o ($n \rightarrow n+5$), sono chiamate **elica 3_{10}** ed **elica π**

Elica 3_{10} : $(\phi, \psi) = (-49^\circ, -26^\circ)$

Elica π : $(\phi, \psi) = (-57^\circ, -70^\circ)$

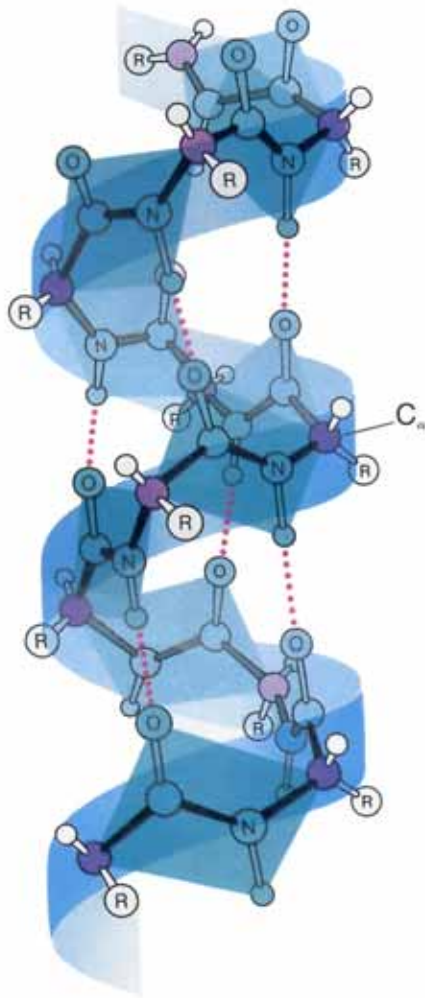
- entrambe sono ai limiti della regione permessa per le strutture elicoidali destrorse nel plot di Ramachandran:



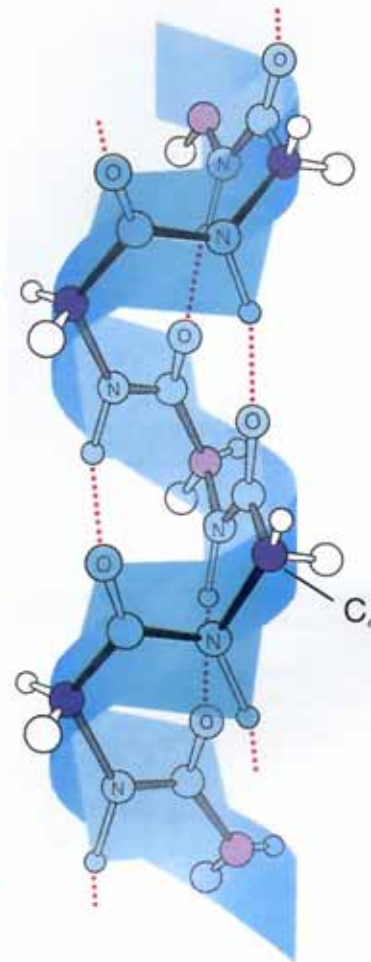
- le elica 3_{10} e π ricorrono **poco frequentemente** nelle proteine

- le eliche 3_{10} possono essere presenti alle **estremità delle α eliche**, come segmenti molto corti (3-4 amminoacidi)

Altri tipi di elica

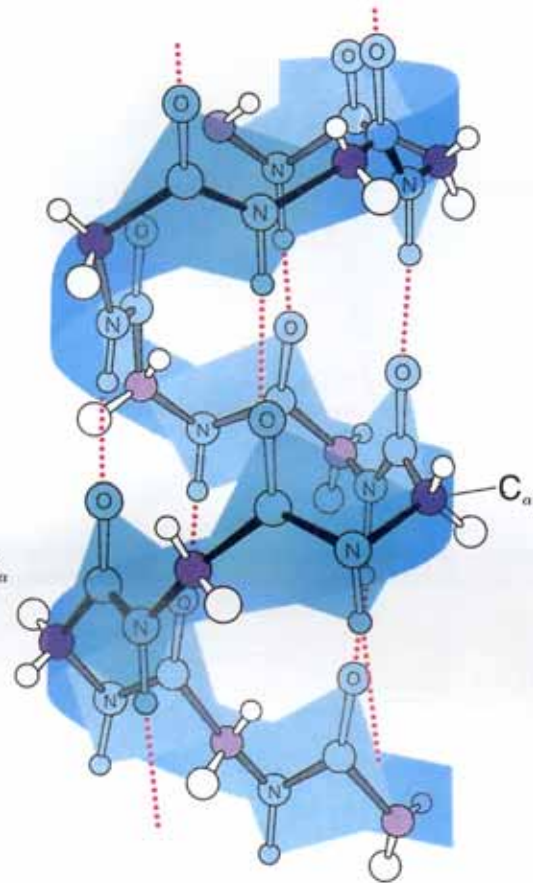


(a) α elica



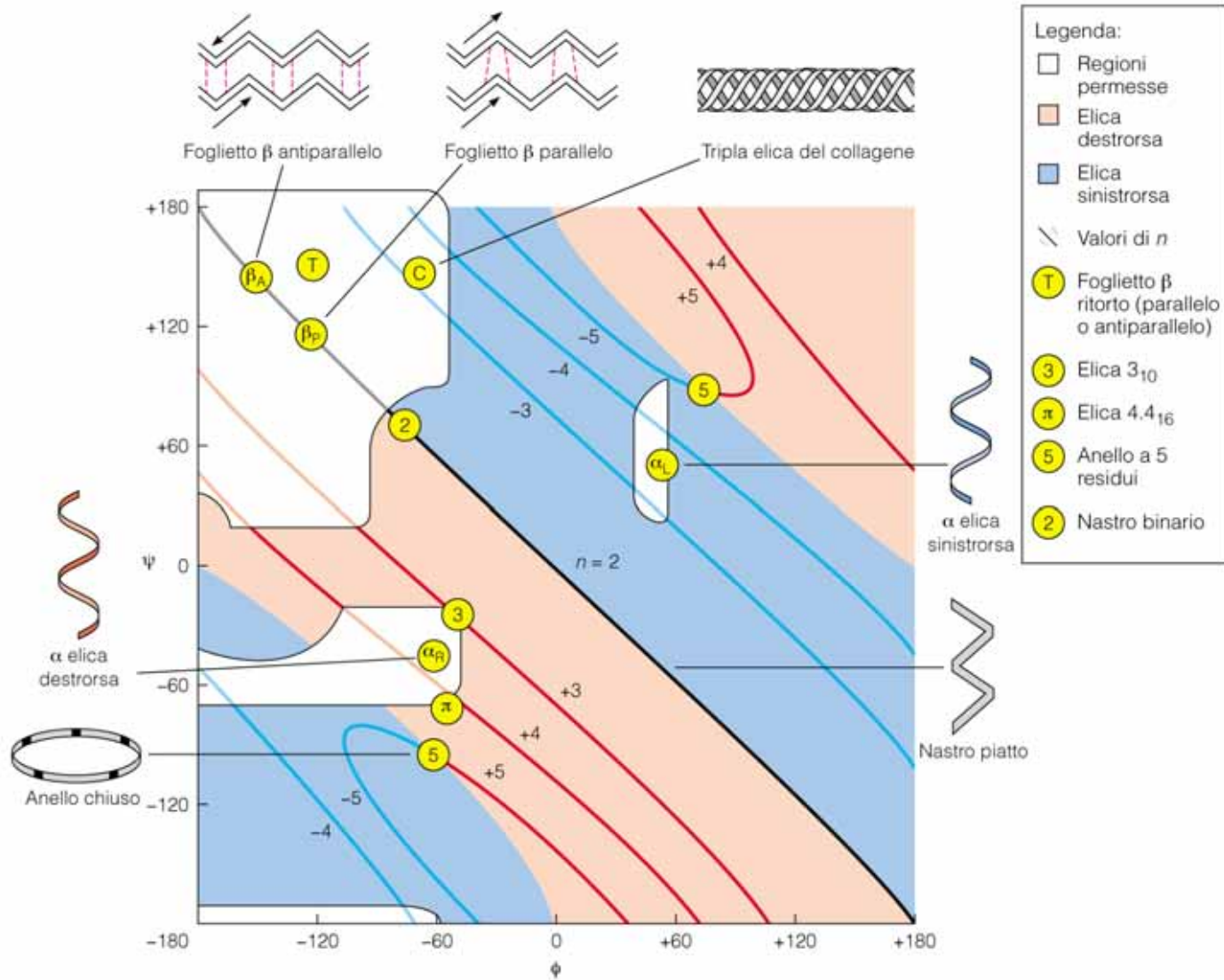
(a) Elica 3_{10}

(b) Foglietto β



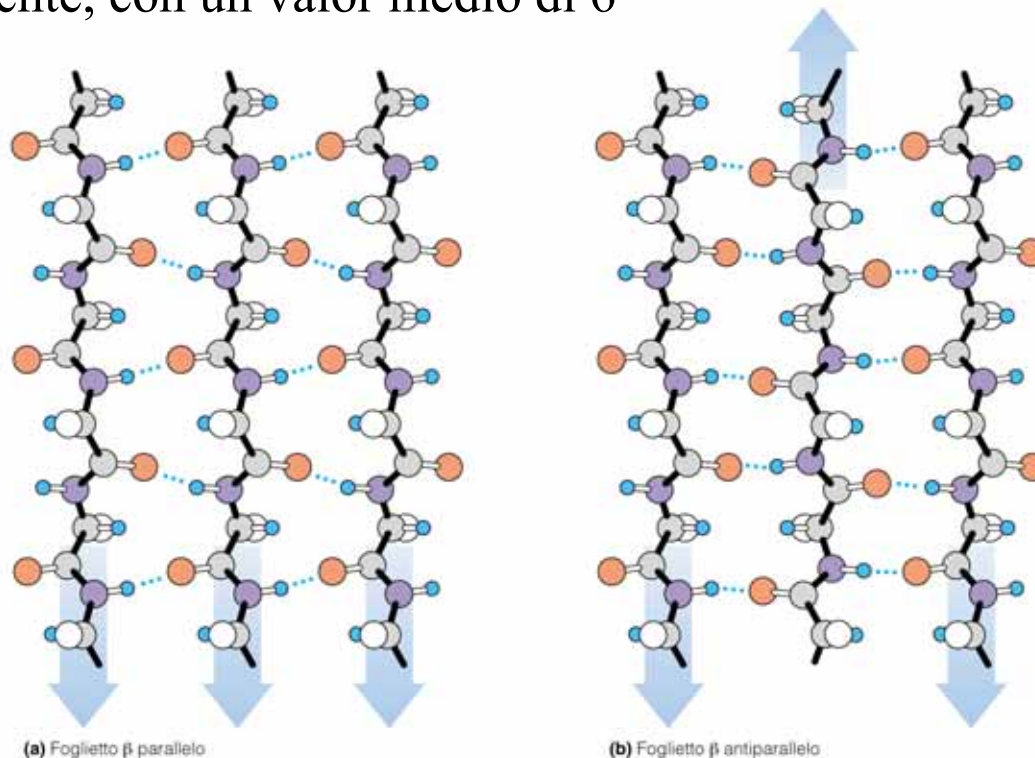
(b) Elica π

Struttura Secondaria



Foglietto- β

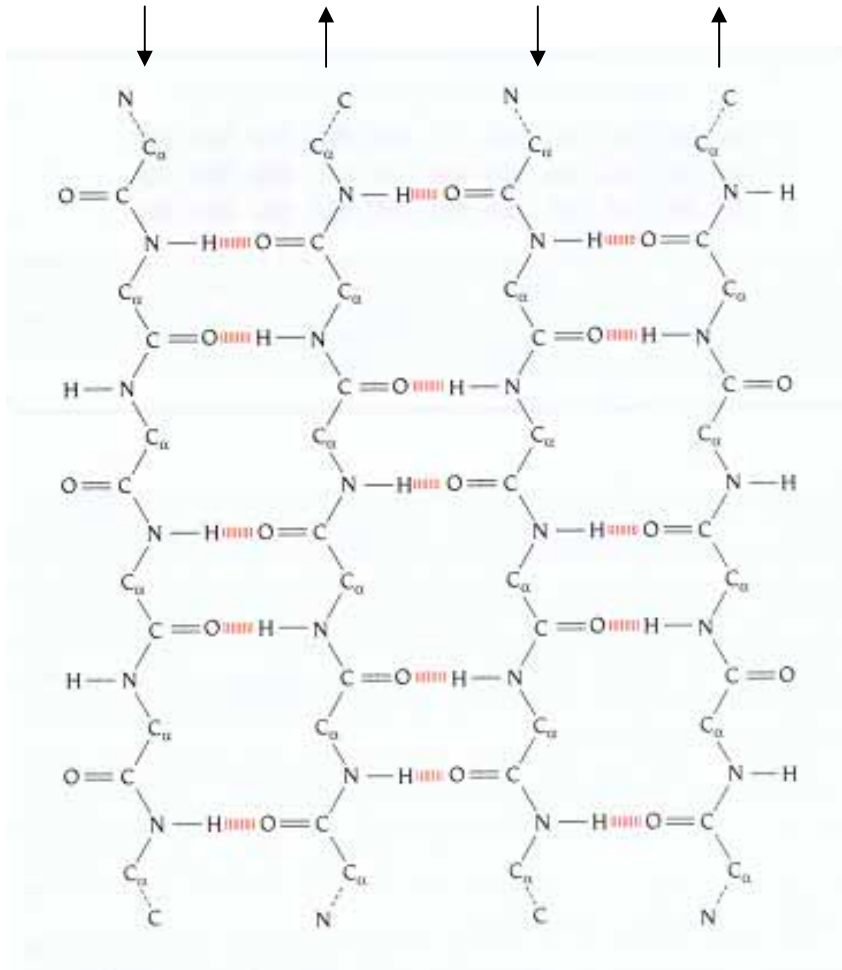
I foglietti β nelle proteine globulari sono costituiti da 2 a 15 filamenti β abbinati lateralmente, con un valor medio di 6



I filamenti β sono allineati uno vicino all'altro, in modo tale che si possano formare **legami a H** tra i **gruppi CO** di un filamento e i **gruppi NH** del filamento β adiacente e viceversa

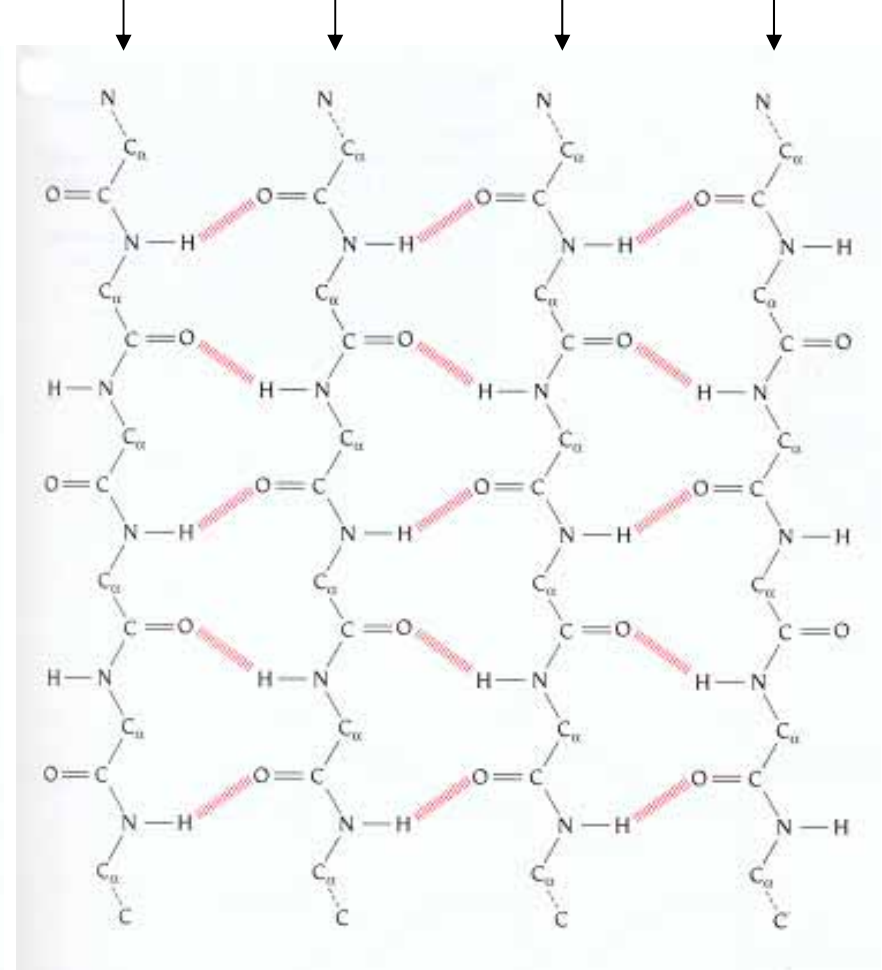
Foglietto- β

foglietto- β antiparallelo



$(\phi, \psi) = (-139^\circ, 135^\circ)$

foglietto- β parallelo



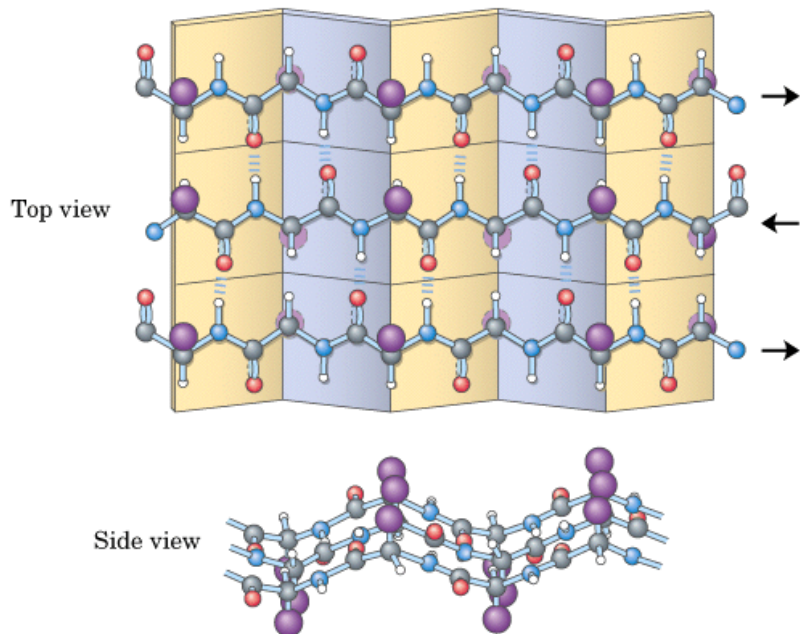
$(\phi, \psi) = (-119^\circ, 113^\circ)$

Foglietto- β

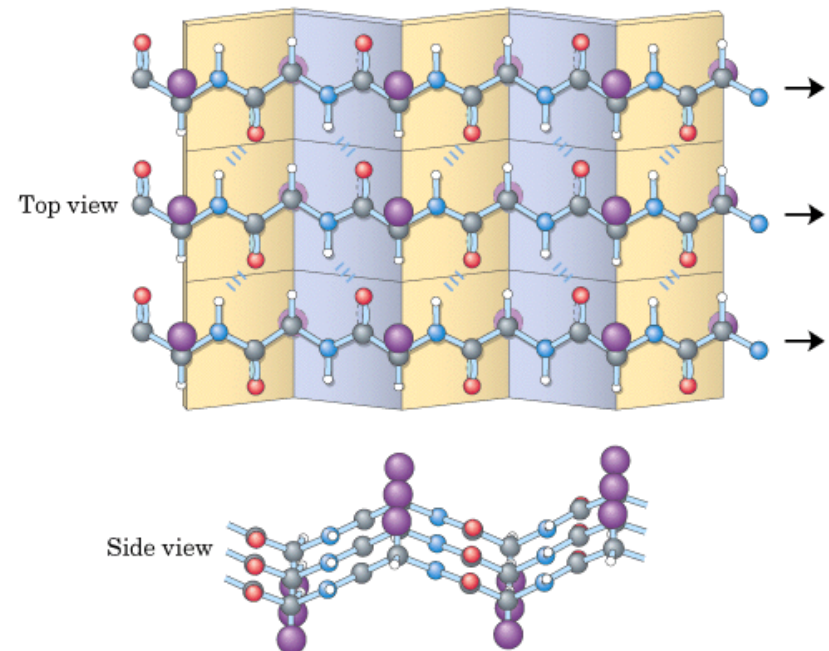
I foglietti β formati da un certo numero di filamenti β sono ‘**pieghettati**’ (**pleated**)

Le **catene laterali** degli puntano **alternativamente sopra e sotto il foglietto β** . Spesso un lato del foglietto β può presentare tutte catene laterali polari e l'altro tutte non polari.

(a) Antiparallele

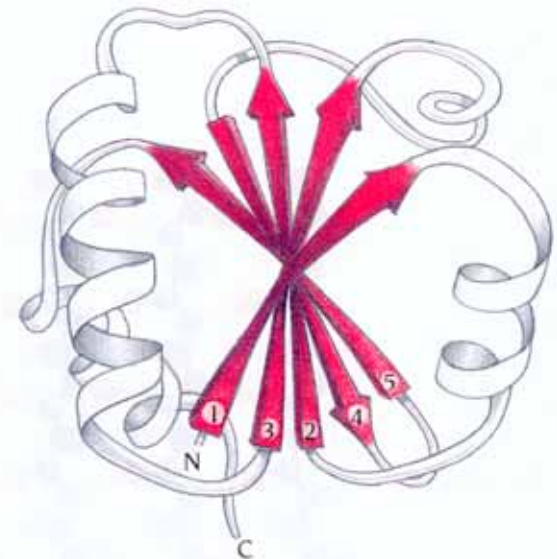


(b) Parallele



Foglietto- β

- i foglietti β antiparalleli sono più stabili di quelli paralleli: i legami-H hanno nei primi una geometria più favorevole
- i foglietti β non hanno momento dipolare
- tutti i foglietti β osservati (paralleli, antiparalleli e misti) presentano una **torsione (twist) destrorsa** (fino a 30°) di ciascun filamento β che li compone
 - \Rightarrow compromesso tra l'ottimizzazione dell'energia conformazionale delle catene polipeptidiche costituenti i filamenti β e il mantenimento della geometria dei legami idrogeno



Reverse turn

Le proteine globulari hanno una forma compatta, dovuta a numerose **inversioni** della direzione della catena polipeptidica che le compone

Molte di queste inversioni sono dovute alla presenza di un comune elemento strutturale, chiamato **reverse turn**

Gli amminoacidi coinvolti nella formazione dei reverse turn si trovano generalmente sulla superficie delle proteine e hanno natura polare e/o carica

Esistono vari tipi di reverse turn, a seconda del numero di amminoacidi che li costituiscono e degli elementi di struttura secondaria che collegano:

β -turn

γ -turn

Ω -loop

Random coil (loop)

Le regioni della catena polipeptidica che non assumono alcun tipo di struttura secondaria

- struttura irregolare nel senso che le coppie (ϕ , ψ) degli amminoacidi che li compongono assumono valori tra i più svariati e non costanti
- generalmente flessibili e possono adottare diverse conformazioni: per questo motivo molto spesso sono difficili da rilevare sperimentalmente

Per **esempio**: estremità N- e C-terminali, zone particolarmente ricche di amminoacidi carichi (Lys in particolare)

- oltre ad avere la funzione di collegare elementi di struttura secondaria, possono anche partecipare alla formazione di siti di legame e di siti attivi negli enzimi, per cui possono essere **disordinati** in assenza della molecola specifica e **ordinati** in presenza della molecola che legano

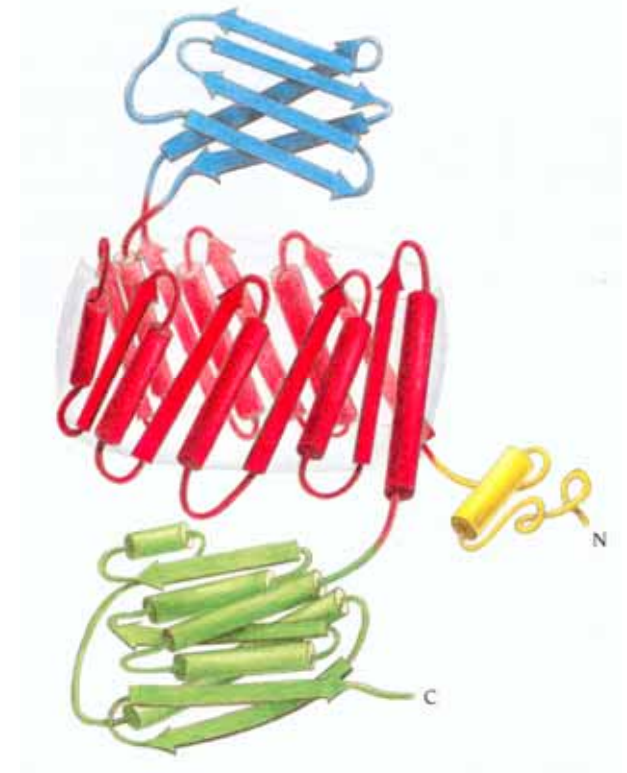
Struttura Terziaria

Struttura terziaria

L'arrangiamento spaziale degli amminoacidi di una singola catena polipeptidica a formare la sua struttura 3D viene chiamata **struttura terziaria** (o **struttura tridimensionale**).

⇒ ripiegamento nello spazio dell'intera catena polipeptidica (catena principale e catene laterali)

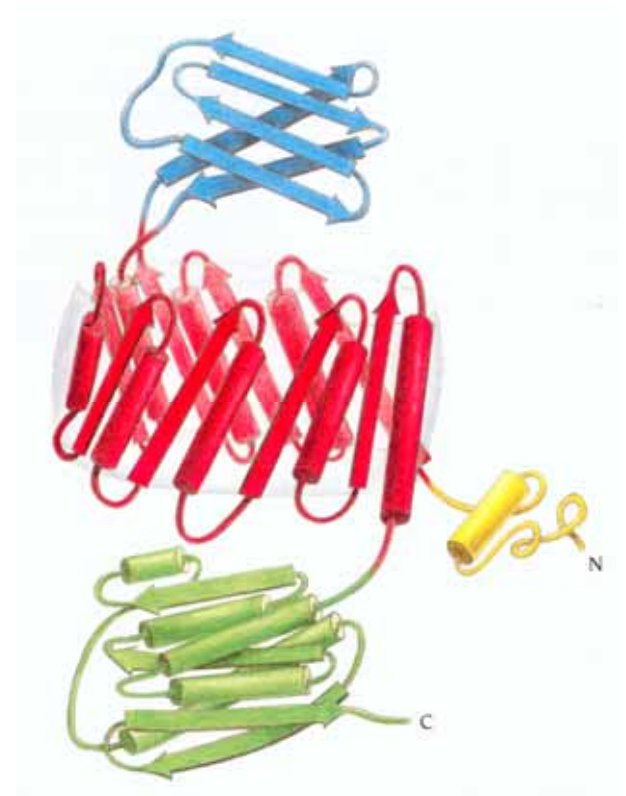
La struttura terziaria può essere vista anche come il modo in cui i diversi elementi di struttura secondaria (eliche, foglietti, ripiegature, anse) si associano tra loro



Architettura della Struttura terziaria

Le strutture terziarie delle proteine sono molto diversificate.
E' tuttavia possibile trovare una logica nella loro architettura:

- alcuni elementi di struttura secondaria possono organizzarsi reciprocamente a dare **motivi strutturali** o **strutture supersecondarie** (gli stessi motivi sono ricorrenti in molte proteine)
- più motivi possono unirsi a e dar luogo ad un **dominio strutturale** (regione strutturalmente – parzialmente – autonoma)
- una molecola proteica può essere costituita da **uno o più domini**

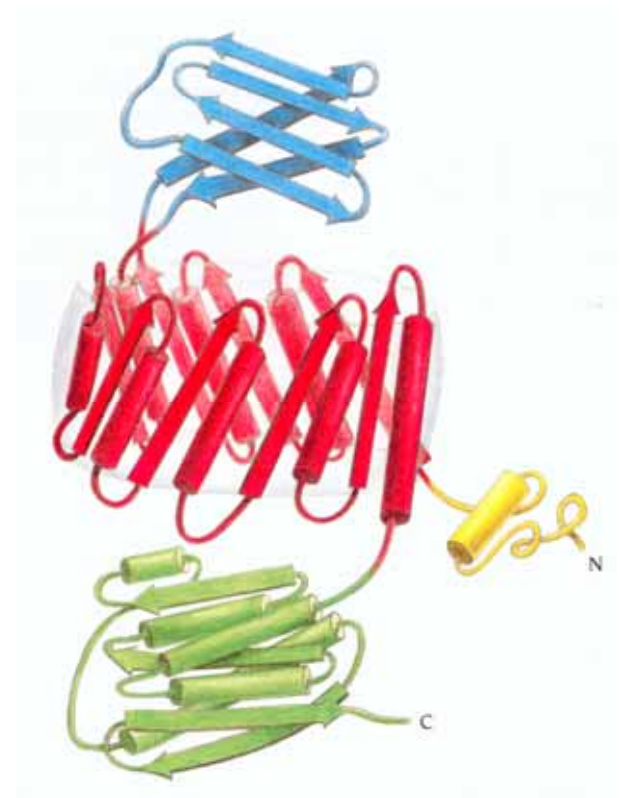


Architettura della Struttura terziaria

L'unità fondamentale della struttura terziaria è il **dominio**

“*struttura terziaria*” viene usato sia per indicare il modo in cui le strutture secondarie si arrangiano a formare i domini, sia il modo in cui una singola catena polipeptidica si ripiega in uno o più domini

- se esiste una significativa omologia di sequenza fra due domini di proteine diverse, questi domini hanno struttura terziaria simile



Domini

I **domini** sono definiti come una catena polipeptidica o parte di essa che si ripiega indipendentemente in una **struttura stabile**, tale da esistere da solo in soluzione acquosa.

- i domini hanno dimensione variabile, di solito < 250 aa
- il 49% di tutti i domini varia tra 51 e 150 aa
- il dominio più grande ha 907 aa
- il numero max di domini in una proteina è 13

I domini sono unità funzionali e spesso a domini diversi di una proteina sono associate funzioni diverse.

I domini hanno cuori idrofobici

- ⇒ minimizzazione numero interazioni sfavorevoli dei gruppi idrofobici con il solvente acquoso
- ⇒ massimizzazione numero interazioni favorevoli di van der Waals dei gruppi idrofobici fra loro

Domini e struttura terziaria

Nei domini e, più in generale, nella struttura terziaria di una proteina gli a.a. sono spazialmente distribuiti a seconda della loro **polarità**

- amminoacidi **non polari**: si trovano nella parte interna della proteina, per evitare contatti con il solvente. Tale distribuzione viene promossa dalle interazioni idrofobiche

- amminoacidi **polari carichi** si trovano sulla superficie della proteina, in contatto con il solvente. La loro presenza all'interno della proteina è dovuta ad una loro specifica funzione chimica, come promuovere la catalisi o partecipare al legame di ioni metallo

- amminoacidi **polari non carichi** di solito si trovano sulla superficie della proteina, ma spesso anche nella parte interna della macromolecola. In questo secondo caso, questi amminoacidi formano legami a H con altri gruppi della proteina, neutralizzando così la polarità di questi gruppi

I domini sono costituiti da diverse **combinazioni di elementi di struttura secondaria** e di **motivi**

Struttura supersecondaria o motivi

- le α eliche e i filamenti β costituenti i motivi sono **adiacenti uno all'altro nella struttura tridimensionale** e connessi da regioni di loop; a loro volta, i motivi adiacenti o formati da regioni consecutive della catena polipeptidica sono vicini nella struttura tridimensionale

- alcuni di questi motivi si possono associare ad **una particolare funzione**, come ad esempio il legame del DNA, mentre altri non hanno una funzione biologica specifica, ma sono semplicemente **parte di organizzazioni strutturali più ampie e complesse**.

I principali motivi individuati nelle proteine sono:

(1) **motivi α** , (2) **motivi β** , (3) **motivi α/β**

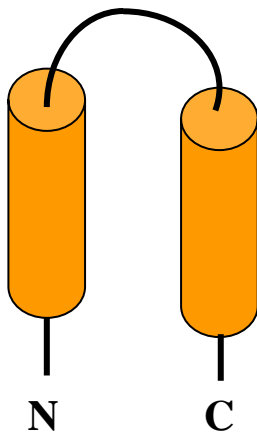
Motivi α

Le α eliche isolate non sono stabili in soluzione

\Rightarrow nelle strutture terziarie delle proteine le α eliche si impaccano in **modo adiacente** una all'altra attraverso interazioni tra le catene laterali idrofobiche.

I principali **motivi α** sono: (1) **α -loop- α** e (2) **EF-hand**

α -loop- α



Il motivo α più semplice consiste di 2 α eliche antiparallele collegate da una regione di loop, chiamato **α hairpin**.

La più breve connessione fra 2 α eliche coinvolge 2 amminoacidi, di cui il secondo è sempre Gly, orientati perpendicolarmente agli assi delle eliche.

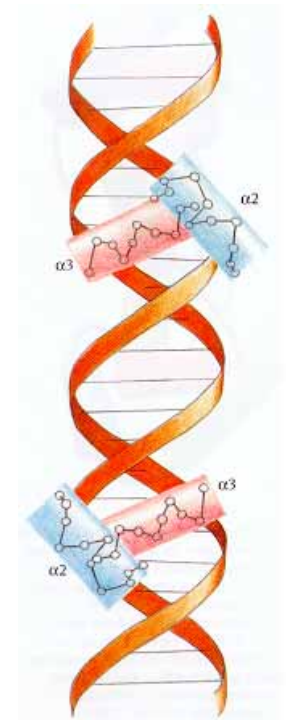
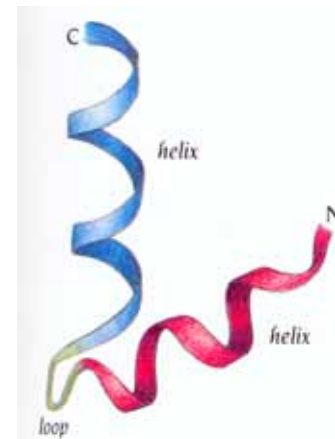
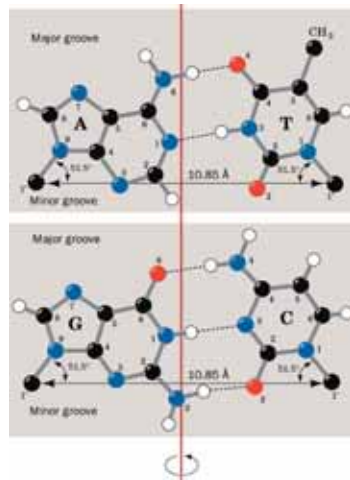
Motivi α

α -loop- α

Un particolare motivo α -loop- α è caratteristico di alcune proteine che riconoscono e legano specifiche zone di DNA.

un'elica si va ad inserire nel solco maggiore del DNA, e riconosce le basi nucleotidiche,

mentre l'altra interagisce con i gruppi fosfato dello scheletro desossiribosio-fosfato.

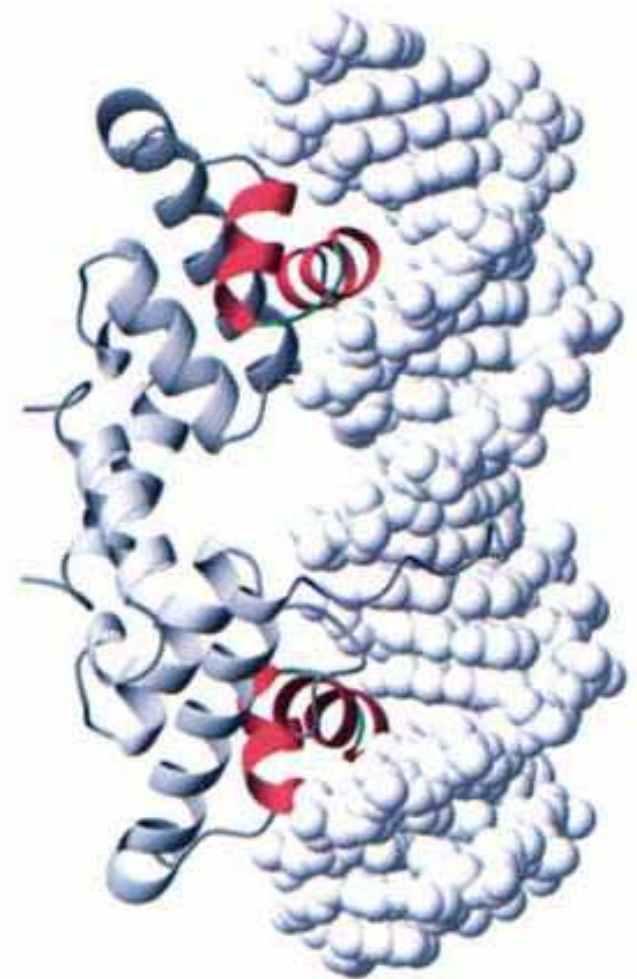


Motivi α

α -loop- α

Esempio: Proteina repressore lambda

Elica-turn-elica È riportato in colore il dominio che lega DNA della proteina repressore lambda, regolatoria di un gene batterico, contenente due motivi elica-turn-elica. Le due eliche più vicine al DNA sono le eliche di lettura o di riconoscimento, che si legano nel solco maggiore e riconoscono sequenze di DNA specifiche per la regolazione di geni. (PDB 1lmb.)



Motivi α

EF-hand

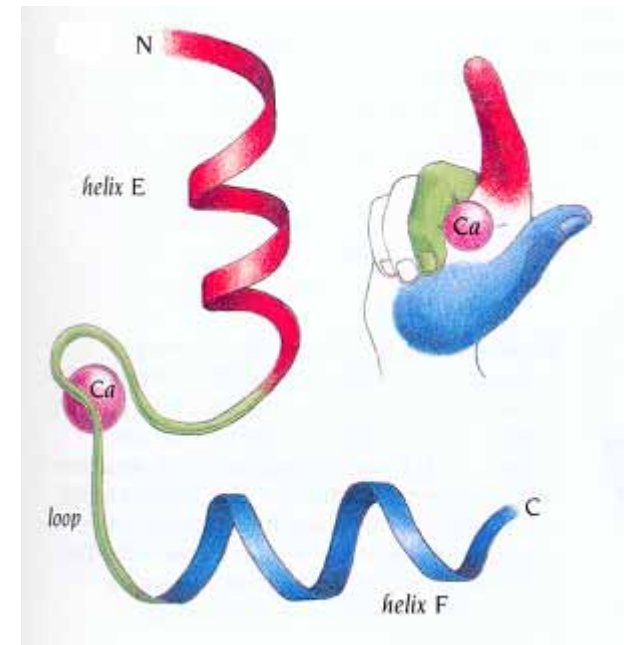
motivo α è specifico per il legame del calcio ed è presente in proteine che legano il calcio quali parvalbumina, calmodulina e troponina C, che regolano l'attività cellulare

Il loop fra le 2 eliche (12 aa) lega l'atomo di calcio

5 aminoacidi legano il calcio e le loro catene laterali devono possedere cariche negative (Asp e Glu);

- il sesto aminoacido del loop deve essere Gly;

- un certo numero di aminoacidi devono essere idrofobici per formare una piccola zona idrofobica fra le 2 α eliche



Motivi β

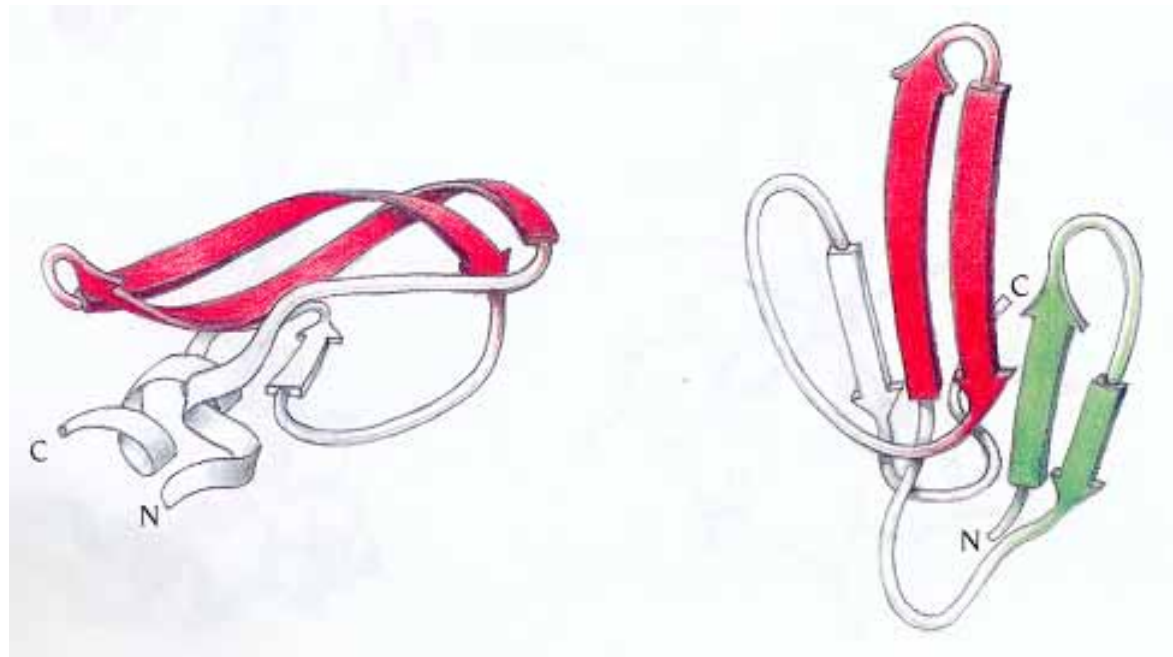
β -hairpin

2 filamenti β antiparalleli adiacenti collegati da un tratto di loop.

ricorre molto frequentemente nelle strutture β antiparallele, come motivo isolato o come parte di un foglietto β più complesso.

La lunghezza del tratto di loop tra i filamenti β è variabile, ma di solito è costituito da 2-5 amminoacidi.

A questo motivo β non è associata nessuna funzione specifica.



Motivi $\alpha\beta$

Crossover connection

La **cross over connection** consiste di di due filamenti β paralleli, un' α elica e due loop (che possono variare notevolmente in lunghezza).

L' α elica si impacca con i 2 filamenti β , riparando dal solvente gli amminoacidi idrofobici dei filamenti β .

La cross over connection può essere considerata come un largo giro di superelica, a partire dal primo filamento β , attraverso la connessione, fino al secondo filamento β .



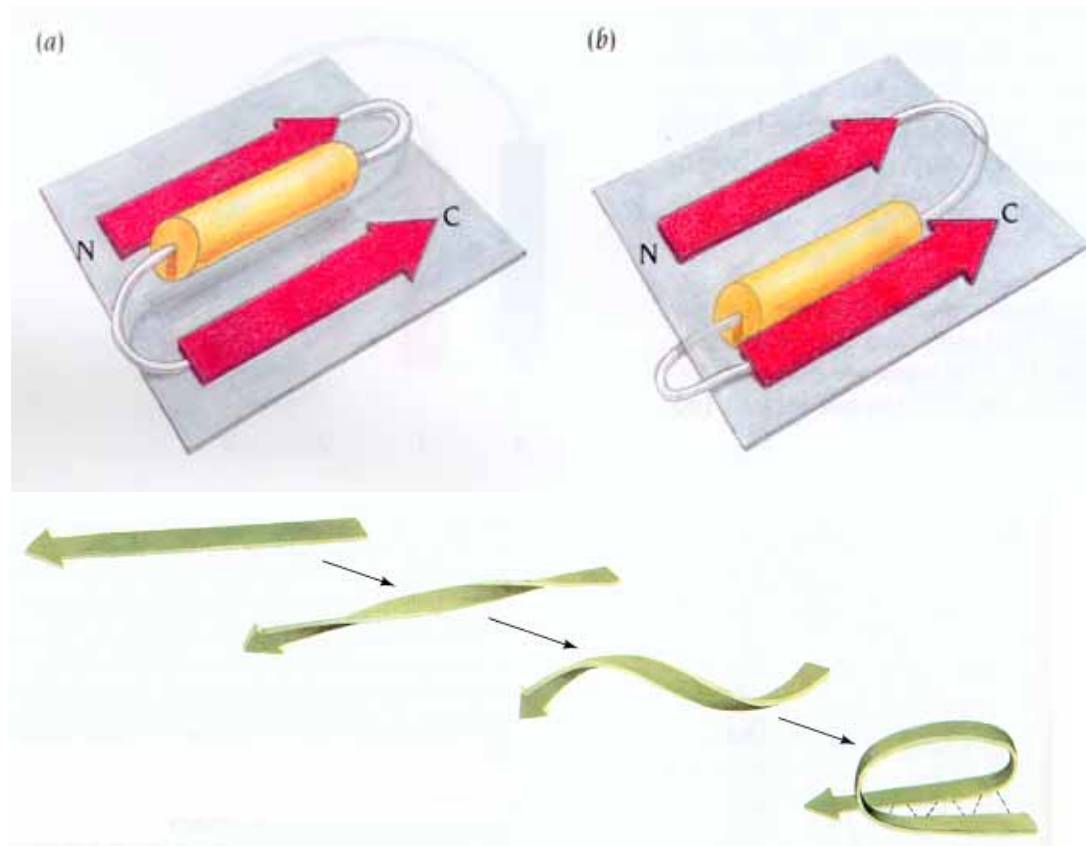
Motivi $\alpha\beta$

Crossover connection

La cross over connection può essere di tipo destrorso (a) o sinistrorso (b).

Quasi tutte le proteine presentano una cross over connection destrorsa.

cross-over connection destrorsa per meglio adattarsi al twist destrorso dei foglietti β .



(oppure: il twist destrorso dei foglietti β favorisce la formazione delle cross-over connection destrorse)

Domini

Michael Levitt and Cyrus Chothia, sulla base di semplici considerazioni sulla connessione dei motivi, hanno classificato i **domini** in **3 gruppi** principali, a seconda delle strutture secondarie e dei motivi coinvolti nella loro formazione:

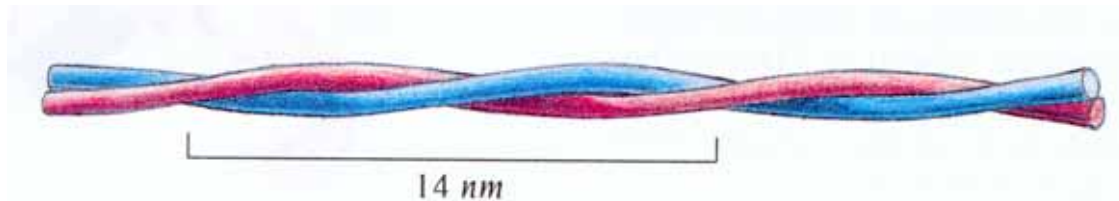
- **domini α** (coiled-coil, fascio a 4 eliche, fold globinico)
- **domini β** (topologia ad up-and-down, chiave greca, jelly roll)
- **domini α/β**
- **altro**

Domini α

Coiled-coil

Il modo più semplice di associazione compatta di α eliche è quello di impaccarsi a coppie.

I coiled coil sono la base di alcune proteine fibrose, quali l' α cheratina e la miosina, e di proteine che legano DNA o RNA.



Superelica sinistrorsa costituita da 2 α eliche destrorse (3.5 residui per giro), in modo tale che le interazioni delle catene laterali fra le 2 eliche si ripetano ogni 7 amminoacidi (2 giri di elica).

Motivo di sequenza (eptapeptide con residui indicati con le lettere *a-g*)

Domini α

Coiled-coil

Gli aminoacidi *a* e *d* sono idrofobici (Val, Leu o Ile) e si impaccano *d-d* e *a-a* ogni 2 giri delle α eliche.

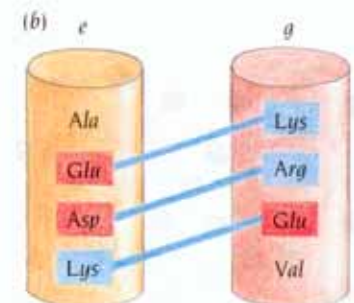
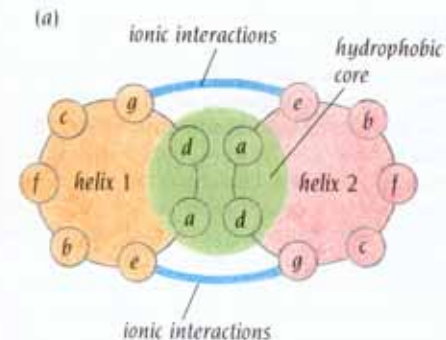
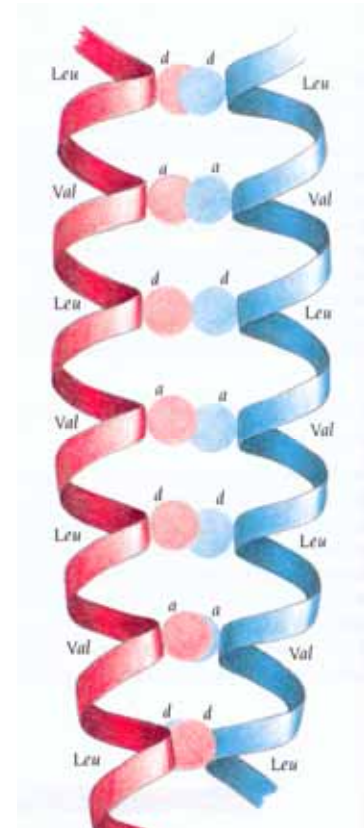
Gli aminoacidi *a* e *d* formano il core idrofobico della superelica coiled coil.

Gli aminoacidi *e* e *g* sono carichi (Glu, Asp, Arg, Lys) e le loro catene laterali formano interazioni fra le 2 α eliche (ponti salini), definendo così la relativa orientazione e allineamento delle 2 catene polipeptidiche.

Esempio

Leu zipper

(proteina GCN4 che lega DNA)
un eptapeptide con Leu in posizione *d*.



Domini α

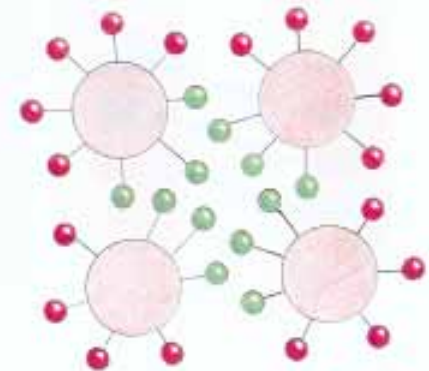
Fascio a 4 eliche (α helical bundle)

4 α eliche disposte con assi reciprocamente quasi paralleli.

Le catene laterali idrofobiche degli amminoacidi di ciascuna α elica sono orientate verso l'interno del fascio mentre le catene laterali idrofiliche degli amminoacidi sono rivolte verso l'esterno.

Le catene laterali idrofobiche rivolte verso l'interno dell' α helical bundle sono così strettamente impaccate che non c'è spazio per molecole d'acqua.

Funzioni biochimiche diverse

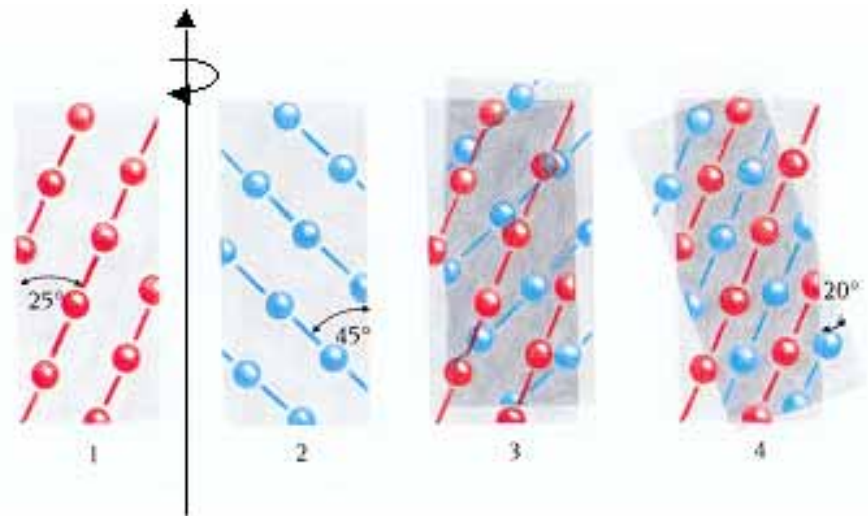


Domini α

Fascio a 4 eliche (α helical bundle)

Le α eliche che costituiscono l' α helical bundle si impaccano con la modalità '**cresta-solco**' (catene laterali di amminoacidi separati da 3-4 residui).

massima corrispondenza fra le creste e i solchi delle 2 eliche in seguito ad un'inclinazione di circa 20° (45° - 25°) fra le 2 eliche.



Domini α

Fold Globinico

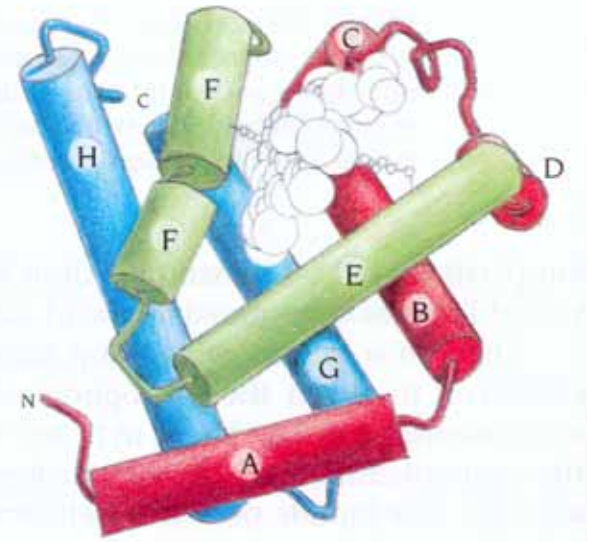
8 α eliche, indicate con le lettere A-H, collegate da loop piuttosto corti, in modo da disporsi a formare la tasca idrofobica in grado di legare gruppi organici o organo-metalli (gruppo eme nella mioglobina e emoglobina)

caratteristico di emoglobine, mioglobine, citocromi, ficocianine
associati a ciascuna α elica si annullano a vicenda.

lunghezza α eliche molto variabile
(da 7 aa per l'elica C a 28 aa per l'elica H)

in forme troncate alcune eliche possono essere
assenti (C e D)

eliche disposte in direzioni diverse, in modo tale che
eliche adiacenti in sequenza non lo sono nella
struttura, con l'eccezione delle eliche G e H.



Domini β

Nei domini β i filamenti β sono disposti in modo tale da formare 2 foglietti β impaccati uno con l'altro (detti “**sandwich**” se aperti, “**barili**” se chiusi), a formare un core idrofobico.

i **domini β** più frequentemente osservati hanno le seguenti 3 topologie:

- **up and down**

- **chiave greca**

- **jelly roll**

Un discorso a parte meritano i domini a **β elica**

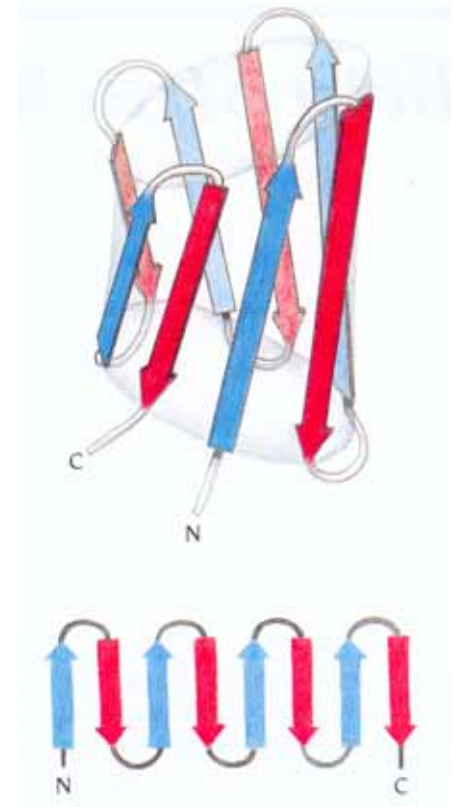
Domini β

topologia ad up-and-down

filamenti β adiacenti sono collegati da regioni di loop (β -hairpin).

Spesso l'ultimo filamento β è collegato al primo da legami a H, a formare un barile (**β barrel**) di 8 filamenti β , (Es: proteine che legano il retinolo o acidi grassi)

In sequenza si alternano catene laterali idrofobiche e polari, in modo tale da formare il core idrofobico all'interno del barile e da interagire esternamente con il solvente.



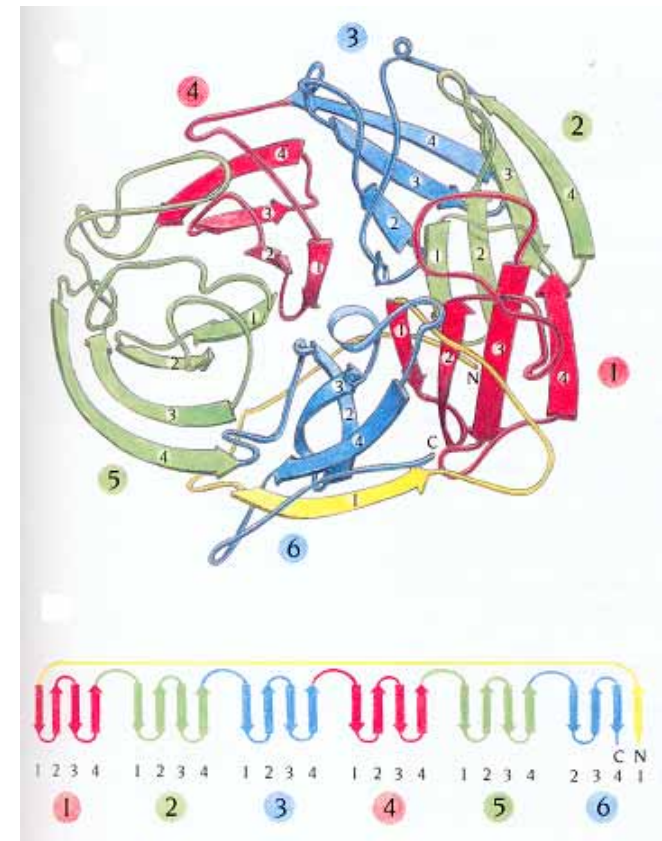
Domini β

topologia ad up-and-down

Es: **neuraminidasi dal virus dell'influenza**

In questo caso i filamenti β non formano un barile ma 6 foglietti β che si organizzano in un β **propeller** (pala di elica)

I loop di ciascun foglietto β che collegano i filamenti β 2 e 3 si trovano tutti dallo stesso lato del β propeller e formano il sito attivo della proteina.



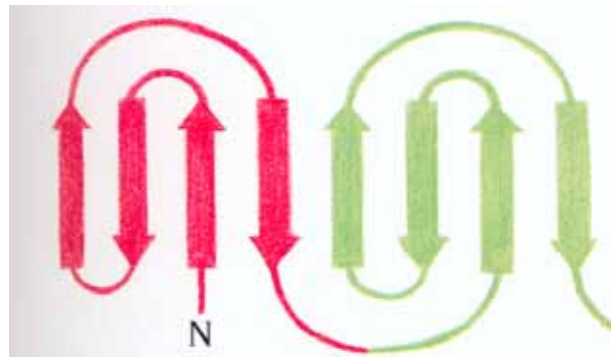
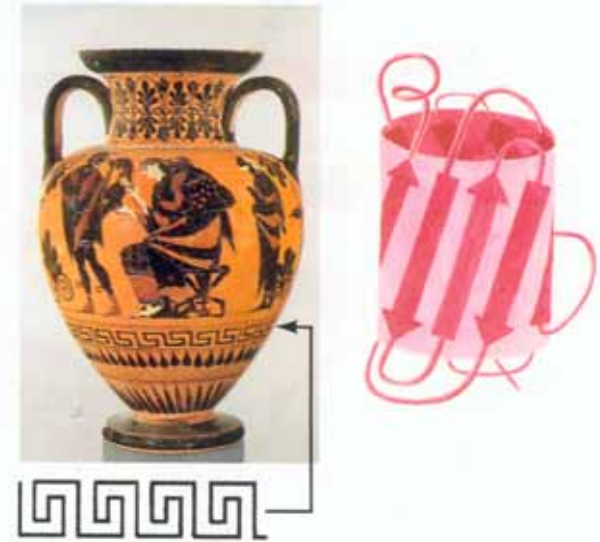
Domini β

topologia a chiave greca

Motivo base costituito da 4 filamenti β : 3 presentano topologia up and down e sono collegati da β hairpin, seguiti da una connessione più lunga al quarto filamento β , che è adiacente al primo.

Esempi: superossido dismutasi a Cu,Zn, immunoglobuline, proteina γ cristallino.

nelle strutture proteiche si è trovata finora sempre la topologia destrorsa



Domini β

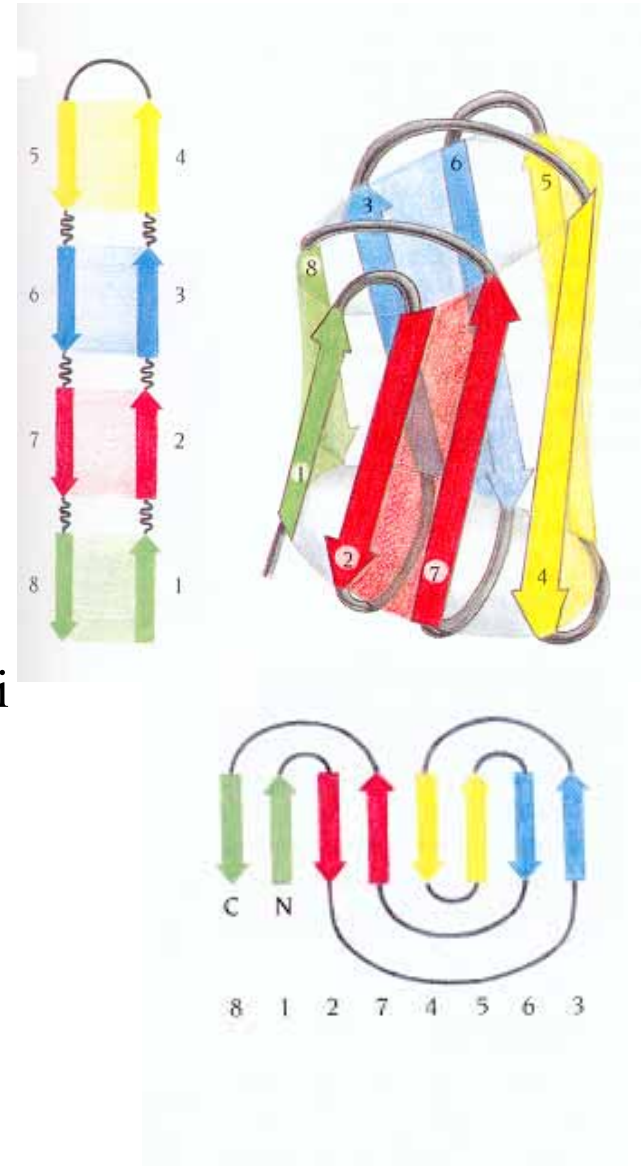
topologia a jelly roll

le coppie di filamenti β (1-8, 2-7, 3-6 e 4-5) sono disposte in modo tale che il filamento β 1 sia adiacente al 2, il 7 al 4, il 5 al 6 e il 3 all'8.

2 connessioni attraversano l'estremità superiore del barile e 2 attraversano quella inferiore.

In più ci sono 2 connessioni tra filamenti β adiacenti nell'estremità superiore del barile e 1 all'estremità inferiore.

Esempi: proteine costituenti il capsid dei virus sferici, emmaglutinina.



Domini β

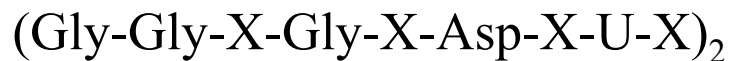
β eliche

- a 2 foglietti β

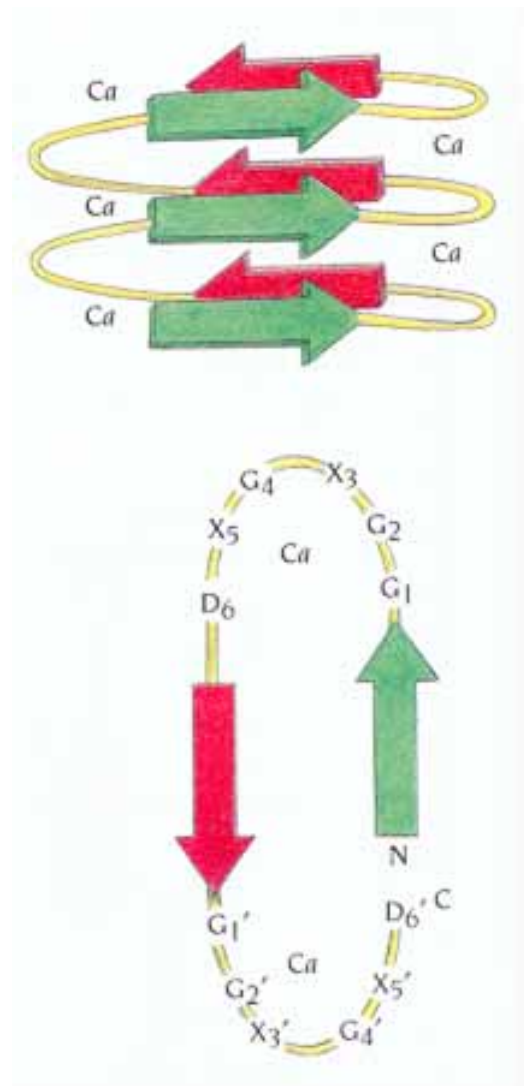
ciascun giro di elica comprende 2 filamenti β e 2 regioni di loop

l'unità strutturale si ripete 3 volte a formare una struttura elicoidale destrorsa costituita da 2 foglietti β paralleli costituiti ciascuno da 3 filamenti β , con un core idrofobico in mezzo.

l'unità strutturale è costituita da 18 aminoacidi, 3 in ciascun filamento β e 6 in ciascun loop. Questa unità base è caratterizzata dalla sequenza consenso:



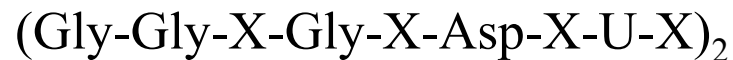
[X: qualsiasi amminoacido; U: idrofobico, spesso Leu]



Domini β

β eliche

- a 2 foglietti β

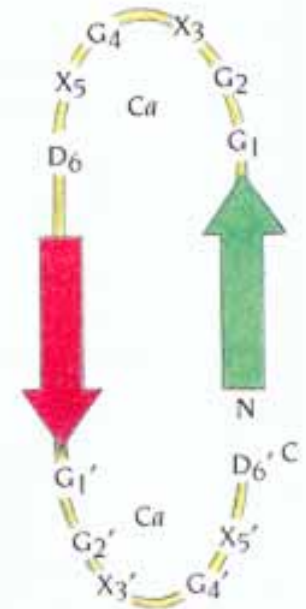
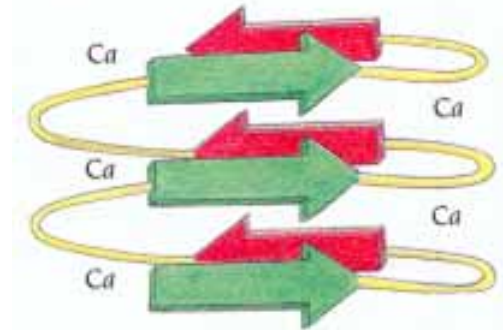


i primi 6 amminoacidi formano un loop e gli altri 3 un filamento β , con la catena laterale di U coinvolta nell'impaccamento idrofobico dei 2 foglietti β

i loop sono stabilizzati da ioni calcio che si legano ad Asp

la sequenza consenso può essere utilizzata per cercare possibili strutture β elica a 2 foglietti β analizzando database di sequenze

Esempi: proteinasi extracellulari batteriche



Domini β

β eliche

- a 3 foglietti β

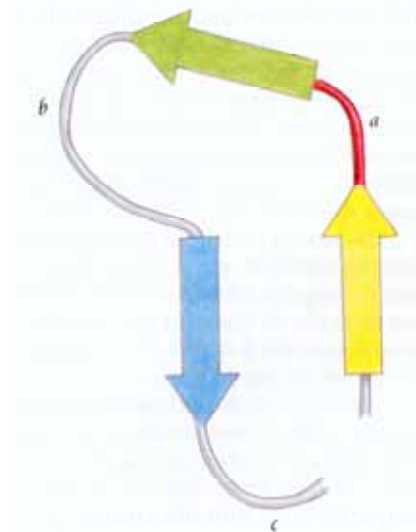
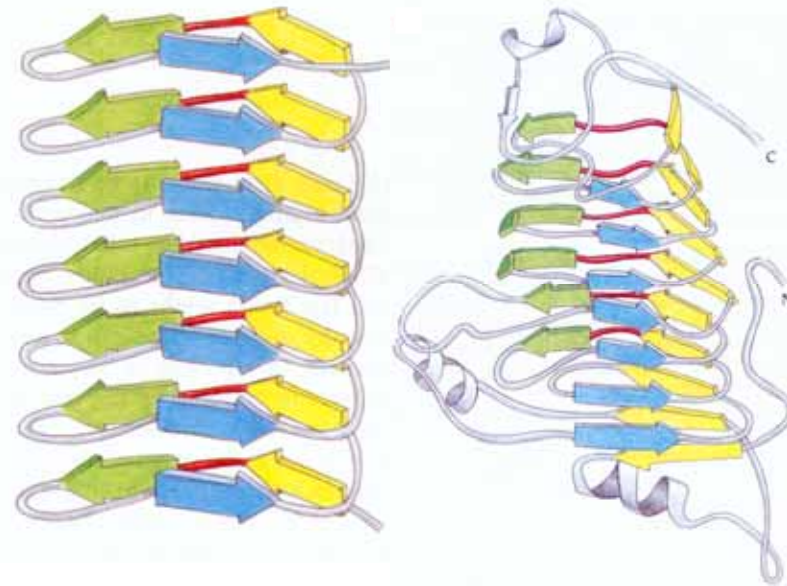
ogni giro di elica è formata da 3 filamenti β (3-5 amminoacidi) collegati da 3 loop

3 foglietti β paralleli, che formano i lati di un prisma (non a sezione triangolare)

il loop (*a*) è corto e quasi sempre formato solo da 2 aa (angolo di circa 120° fra i 2 filamenti β che unisce)

i loop lunghi (variabili) protrudono dai foglietti β e probabilmente formano il sito attivo sulla superficie esterna della proteina \Rightarrow **nessuna sequenza consenso**

Esempi: pectato liasi, proteina tailspike del batteriofago P22



Domini $\alpha\beta$

La **cross over connection** è l'unità costitutiva su cui si basa la topologia di 3 tipi di domini α/β osservati nelle proteine:

- **α/β barrel**
- **motivi ricchi di Leu (fold a ferro di cavallo)**
- **α/β open sheet**

Domini $\alpha\beta$

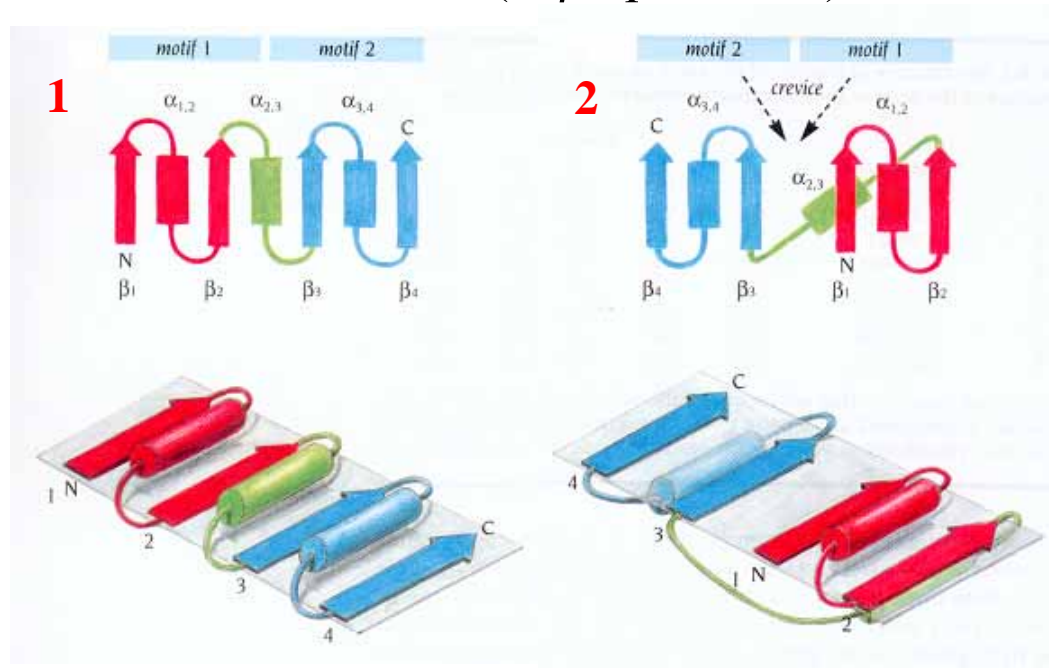
2 cross over connection possono essere collegate in 2 modi diversi (tramite un' α elica), per formare un foglietto β parallelo di 4 filamenti β :

1) α eliche stanno tutte da un lato del foglietto β parallelo

filamenti β nell'ordine 1 2 3 4 (barili α/β e del fold a ferro di cavallo)

2) α eliche stanno sui lati opposti del foglietto β parallelo

filamenti β sono nell'ordine 4 3 1 2 (α/β open sheet)

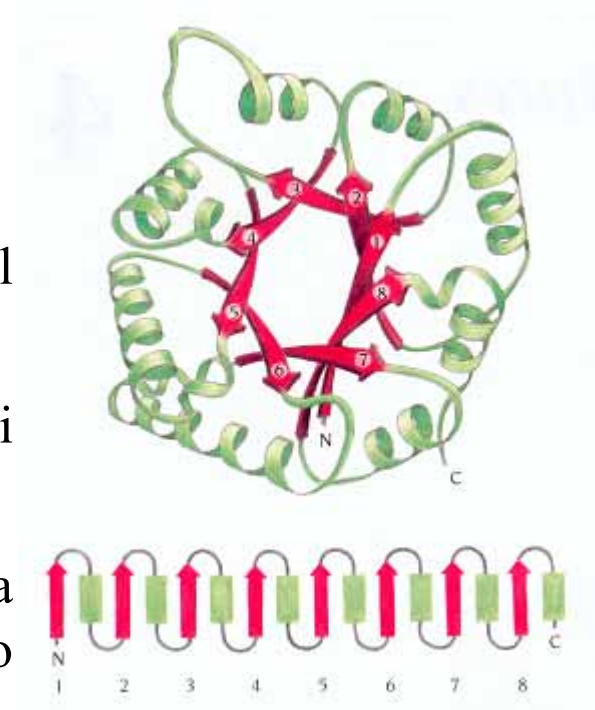


Domini $\alpha\beta$

barili α/β

8 filamenti β paralleli, posizionati in modo tale che il filamento β_8 sia adiacente e faccia legami a H con β_1
dominio che coinvolge almeno 200 aa ed è tipico di molti enzimi

chiamato anche **TIM barrel** dalla struttura dell'enzima trioso fosfato isomerasi, dove fu osservato per la prima volta



Esempi: enzimi coinvolti in isomerizzazione di piccole molecole di zucchero, trasferimento di gruppi fosfato, degradazione di zuccheri polimerici

Domini $\alpha\beta$

barili α/β

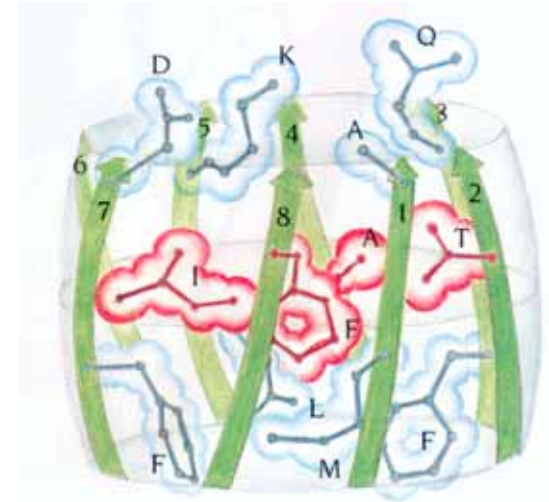
catene laterali idrofobiche delle α eliche si impaccano con le catene laterali idrofobiche dei filamenti β

le altre catene laterali degli aa (in posizione 1, 3, 5, ...) dei filamenti β sono rivolte verso l'interno del barile e formano un core idrofobico strettamente impaccato

Arg, Lys e Gln alle 2 estremità del barile, in modo da poter interagire con il solvente esterno

sito attivo sempre in una tasca costituita dagli 8 loop che connettono l'estremità C-terminale dei filamenti β con l'estremità N-terminale delle α eliche

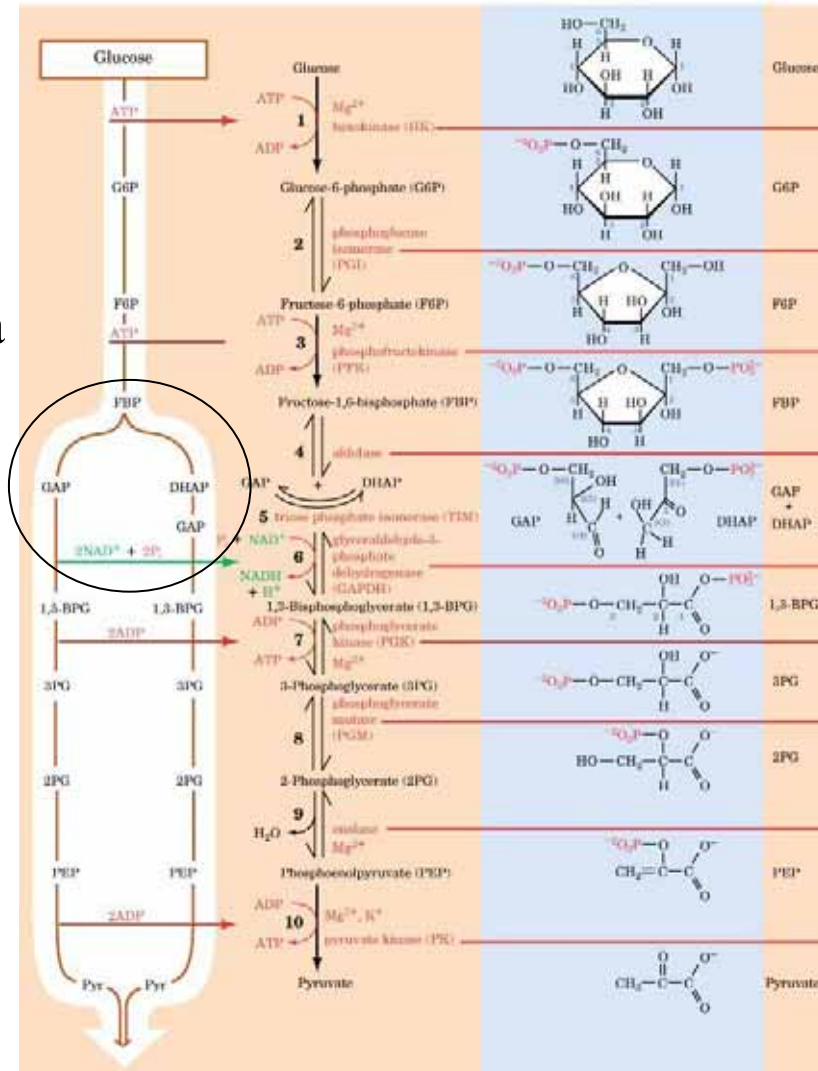
gli aa che partecipano al legame del substrato e alla catalisi appartengono a queste regioni di loop



Trioso fosfato isomerasi (TIM)

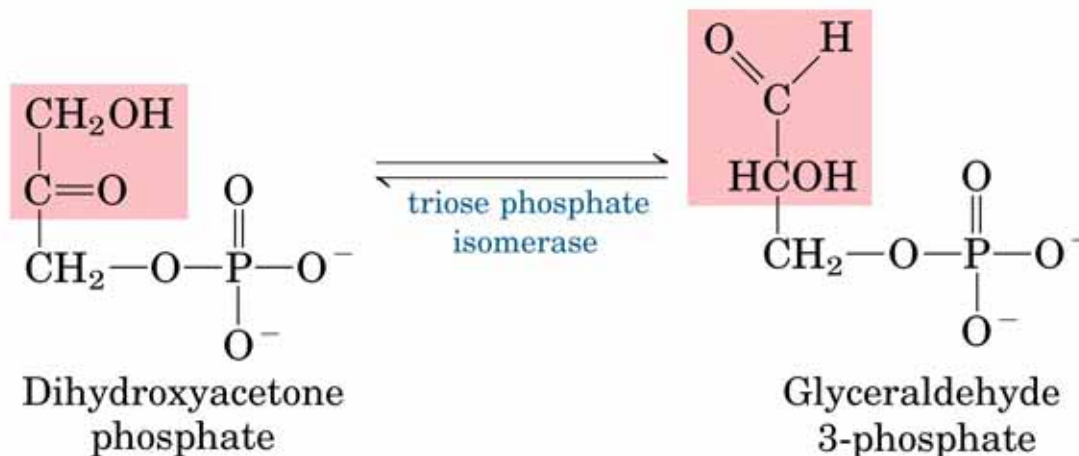
GAP e DHAP (isomeri chetoso-aldosio) prodotti della scissione aldolica. Solo GAP prosegue lungo la via glicolitica

TIM permette ad ambedue i prodotti della reazione aldolasica di procedere nella via glicolitica



Trioso fosfato isomerasi (TIM)

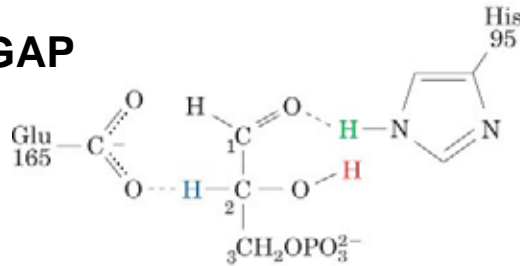
reazione: isomerizzazione con un **intermedio enediolo** (o enediolato) che permette l'interconversione della **gliceraldeide-3-fosfato (GAP)** e del **diidrossiacetone fosfato (DHAP)**



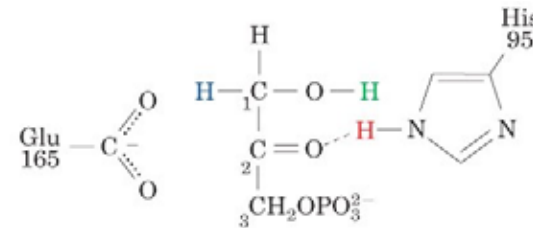
$$\Delta G'^{\circ} = 7.5 \text{ kJ/mol}$$

Trioso fosfato isomerasi (TIM)

conversione di **GAP**
nell'intermedio
enediolo



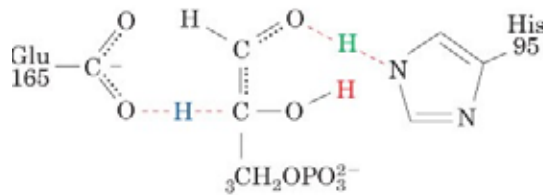
GAP·TIM Michaelis complex



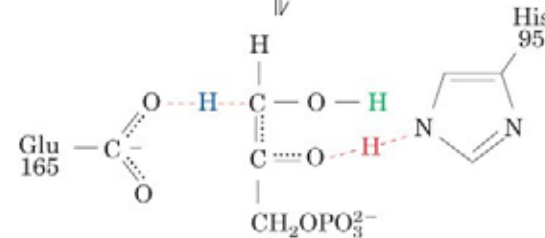
DHAP·TIM Michaelis complex

Glu165 base generale
che sottrae un protone
dal C2 di GAP

His95 acido generale che
protona l'O carbonilico di GAP

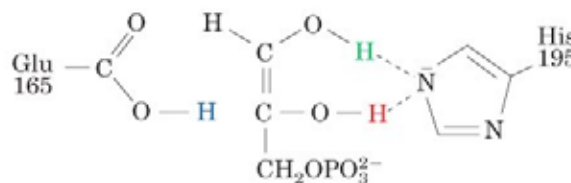


Transition state



Transition state

Glu165 acido generale
che protona il C1



Enediol (or enediolate) intermediate

conversione dell'intermedio
enediolo in DHAP

Trioso fosfato isomerasi (TIM)

- quando il substrato si lega a TIM una ansa di 10 residui (**lid**) si richiude sul substrato

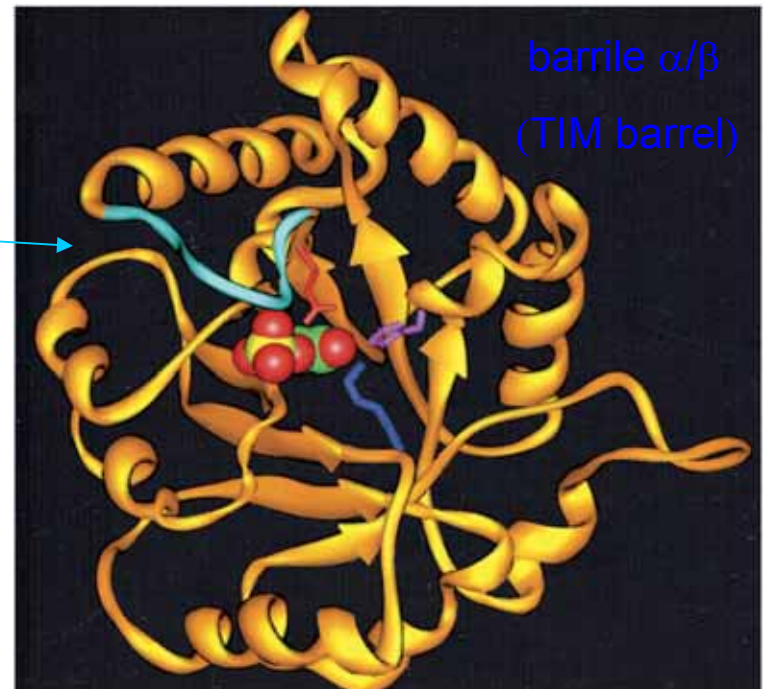
- la chiusura del lid:

(1) stabilizza di preferenza lo stato di transizione della reazione

(2) impedisce la degradazione dell'intermedio enediolo

- in soluzione l'intermedio enediolo si decompone rapidamente in metilgliossale (tossico) e P_i

- altri TIM barrel: aldolasi, enolasi, piruvato chinasi

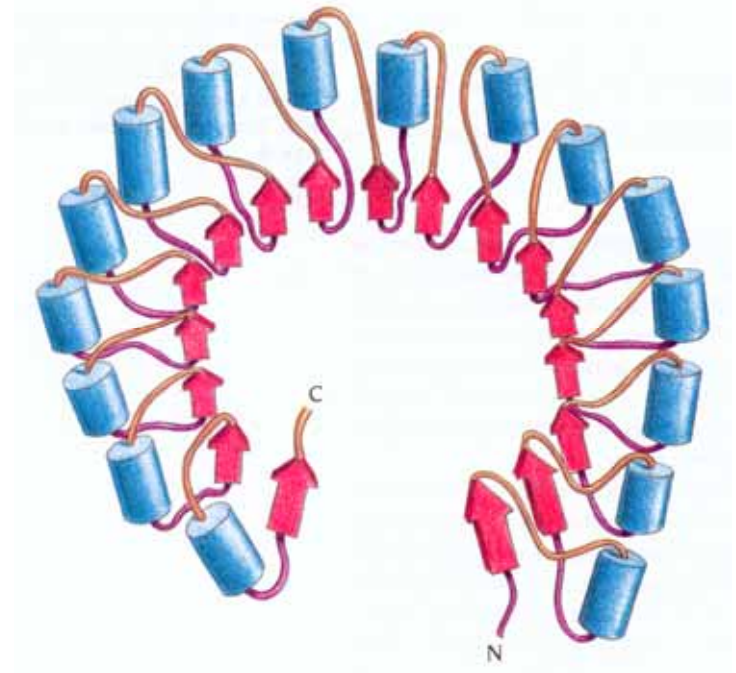


Domini $\alpha\beta$

Fold ferro di cavallo

I filamenti β formano una struttura aperta ricurva, che ricorda un **ferro di cavallo**, con le α eliche sul lato esterno e il foglietto β che forma la parte interna del ferro di cavallo.

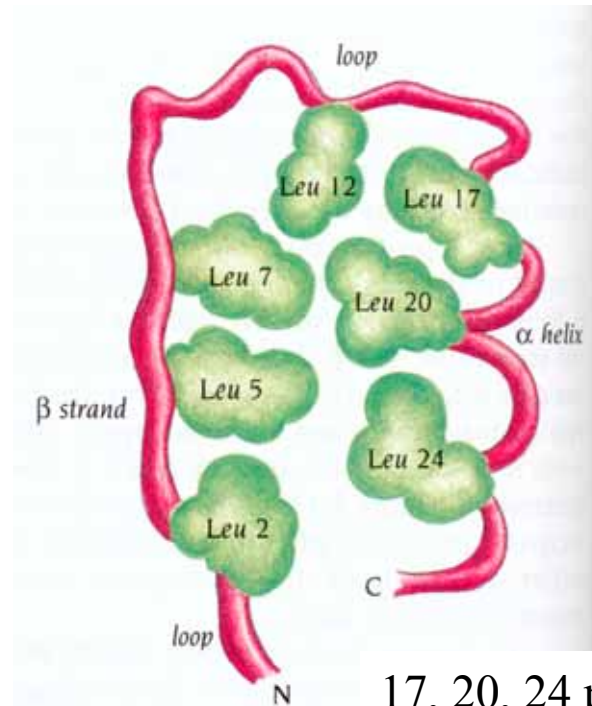
Un lato del foglietto β si interfaccia con le α eliche, formando un core idrofobico, mentre l'altro lato del foglietto β è esposto al solvente, caratteristica che gli altri motivi α/β non hanno.



Domini $\alpha\beta$

Fold ferro di cavallo

Domini ricchi di Leu (ripetizioni di sequenze omologhe di circa 20-30 amminoacidi) sono stati identificati da studi di sequenze in oltre 60 diverse proteine, tra cui recettori, fattori di virulenza batterica, molecole di adesione cellulare, molecole coinvolte nella riparazione del DNA e nello splicing dell'RNA.



17, 20, 24 possono essere \neq da Leu ma comunque idrofobici

(tipo A) NH₂ -X-L-E-X-L-X-L-X-X-C-X-L-T-X-X-X-C-X-X-L-X-X-a-L-X-X-X-X-

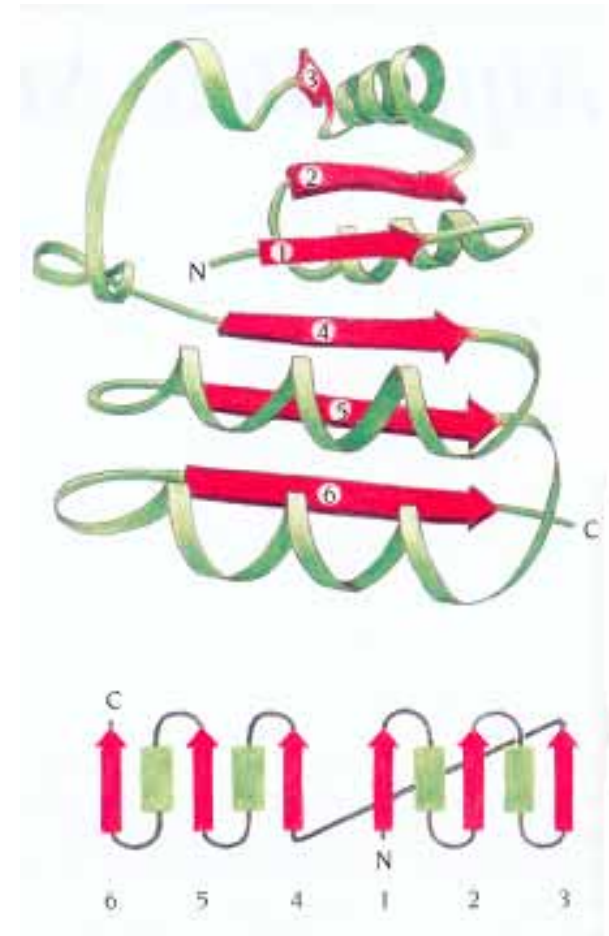
(tipo B) NH₂ -X-L-R-E-L-X-L-X-X-N-X-L-G-D-X-G-a-X-X-L-X-X-X-L-X-X-P-X-X-

(a = amminoacido alifatico)

Domini $\alpha\beta$

α/β open sheet

- le α eliche si trovano su entrambi i lati del foglietto β parallelo. Ciascun filamento β fornisce le catene laterali degli amminoacidi idrofobici per impaccarsi con le α eliche in 2 regioni idrofobiche, una per ciascun lato del foglietto β .
- ci sono sempre 2 filamenti β (1 e 4 in figura) nella parte interna del foglietto β le cui connessioni al filamento β vicino sono sui lati opposti del foglietto β .
- uno dei loop uscenti da uno dei filamenti β va sopra il foglietto β , mentre l'altro loop va sotto.



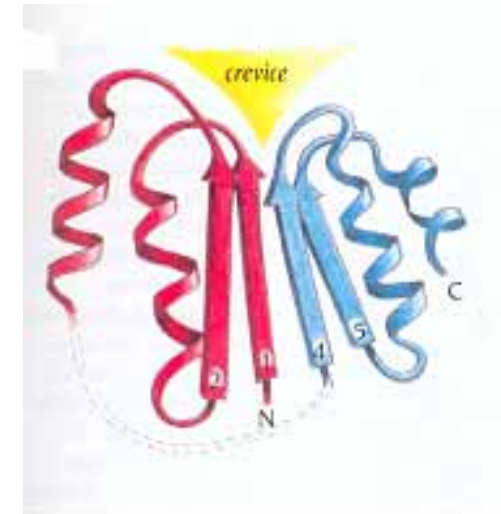
Domini $\alpha\beta$

α/β open sheet

Il punto in cui si ha l'inversione del lato del foglietto β che viene coperto dalle α eliche è chiamato **switch point**.

La zona dello switch point, all'estremità C-terminale del foglietto β , definisce una tasca, in cui si trovano sempre **i siti di legame** di questa classe di proteine.

Negli α/β open sheet non ci sono restrizioni geometriche sul numero dei filamenti β coinvolti: il numero varia da 4 a 10.



Domini $\alpha\beta$

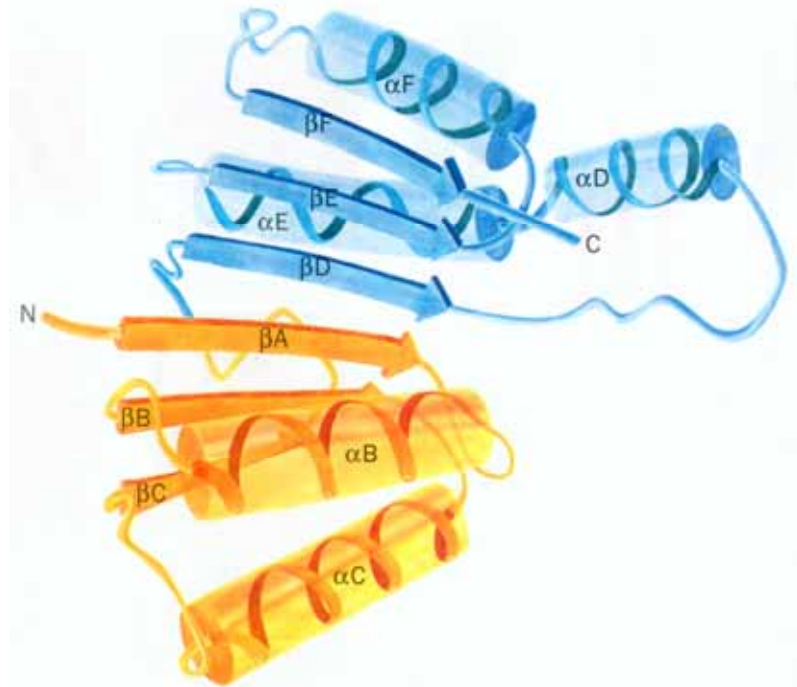
α/β open sheet

Molti enzimi, e in particolare alcune deidrogenasi, hanno il dominio che lega il NAD costituito da due unità $\beta\alpha\beta\alpha\beta$ a formare il **dinucleotide-binding fold** o **Rossmann fold**

Il Rossmann fold è caratterizzato dalla presenza della sequenza consenso:

GXGXXG(17X)D

coinvolta nel legame del dinucleotide.

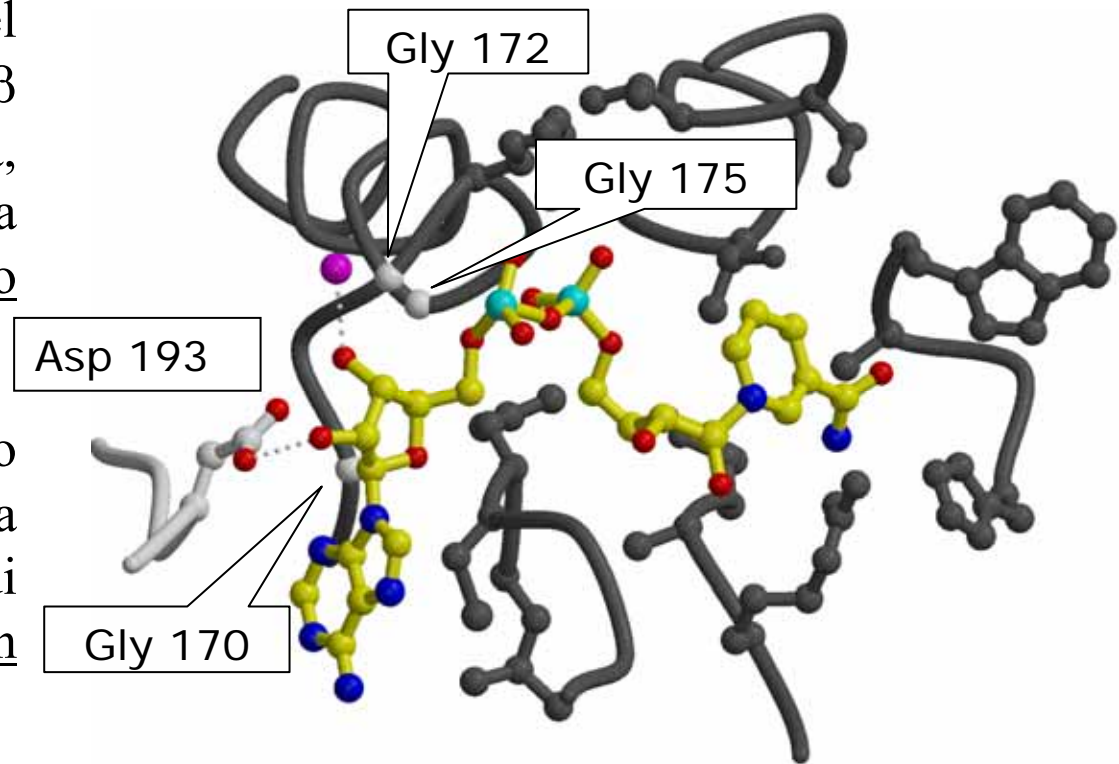


Domini $\alpha\beta$

α/β open sheet

I residui **GXGXXG**, posti nel loop fra il primo filamento β e la susseguente α elica, determinano l'interazione tra la proteina ed il gruppo pirofosfato.

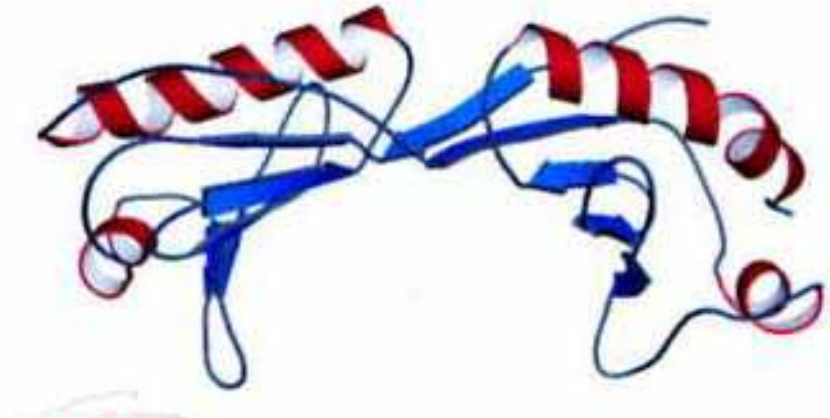
Alla fine del secondo filamento β parallelo la catena laterale carbossilica di un **Asp** o **Glu** interagisce con il diolo del gruppo ribosio.



Domini $\alpha+\beta$

Domini che contengono sia foglietti β che α -eliche, ma essi sono separati

- non esistono particolari principi riguardo le loro organizzazione strutturale
- gruppi di α -eliche interagenti
- foglietti β antiparalleli o misti



Esempio: Sella $\alpha+\beta$

proteina che lega il TATA box
(sito di inizio trascrizione genica negli eucarioti)

- il foglietto β si lega nel solco minore del DNA e lo curva

Domini irregolari con legami crociati

Piccoli domini che ricorrono in proteine extracellulari ed intracellulari monodominio

Domini troppo piccoli per avere:

- 1) un esteso core idrofobico
- 2) numerose interazioni di legame a H dovute a struttura secondaria

⇒ problema di stabilità strutturale

La stabilità strutturale viene fornita dalla presenza di legami crociati fra parti diverse del dominio attraverso legami covalenti

- 1) domini extracellulari: presenza di **ponti disolfuro** (varie coppie Cys-Cys)
- 2) domini intracellulari: presenza di **ioni** (di solito Zinco o Ferro) che formano legami di coordinazione con catene laterali nucleofiliche

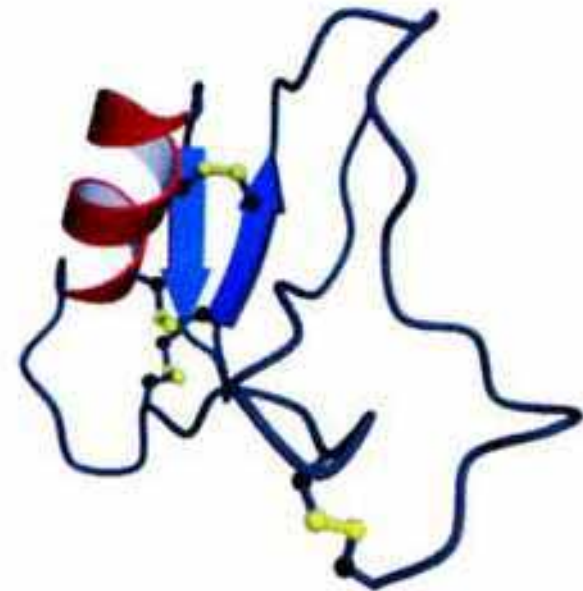
Domini irregolari con legami crociati

1) proteine ricche di **ponti disolfuro**

- spesso tossine che inibiscono proteine cellulari essenziali (veleno cobra, tossina dello scorpione, proteine tossiche dei serpenti marini, inibitori di proteasi secreti, fattore Ra5 dell'allergia al polline)
- proteine molto stabili alla digestione proteolitica ed alla denaturazione per calore

Es: tossina dello scorpione

- 4 ponti disolfuro
- struttura secondaria ridotta al minimo

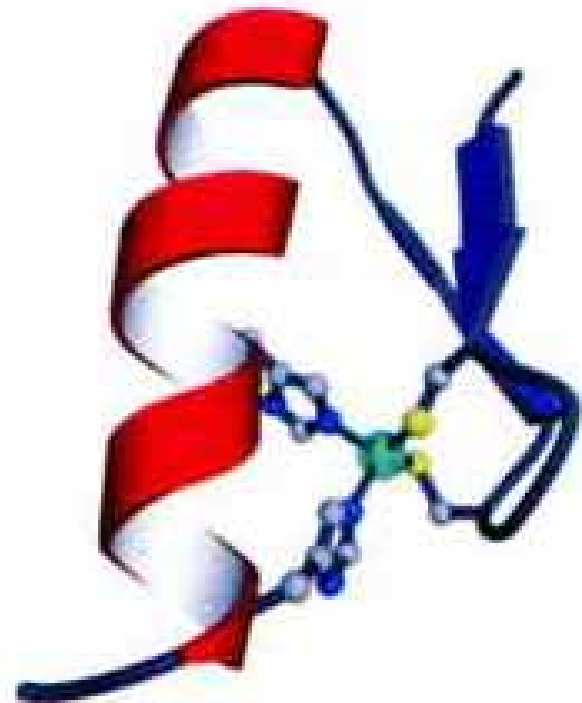


Domini irregolari con legami crociati

2) domini stabilizzati da **ioni**

Zinc finger

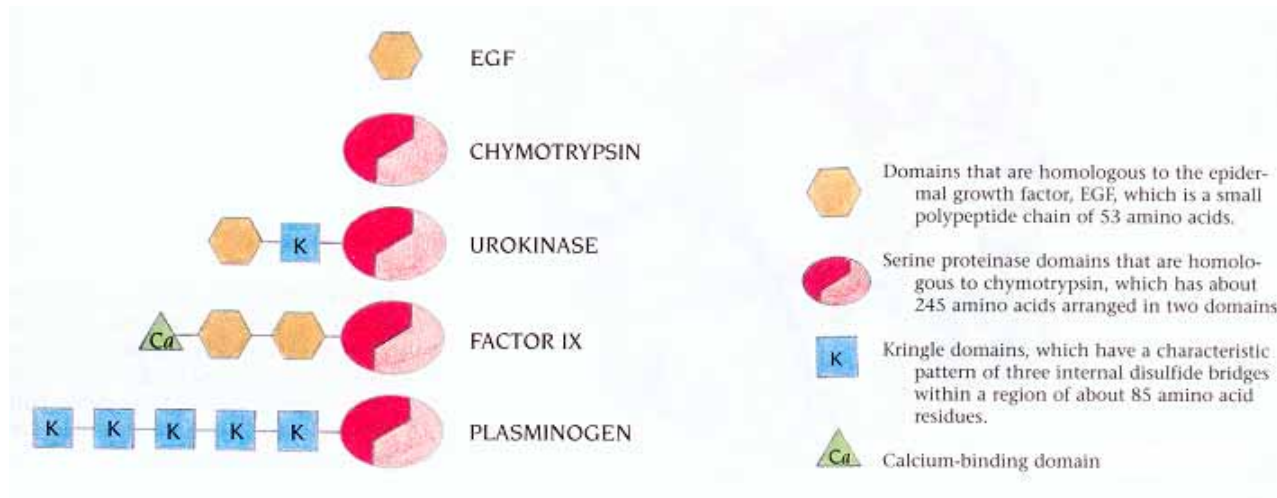
- ione zinco stabilizzato da 2 Cys e 2 His
- dominio unfoldato in assenza dello ione
- il dominio è di solito parte di un fattore di trascrizione più grande



Domini e struttura terziaria

Il numero di combinazioni dei motivi a formare i domini è limitato e alcune combinazioni sono strutturalmente favorite rispetto ad altre.

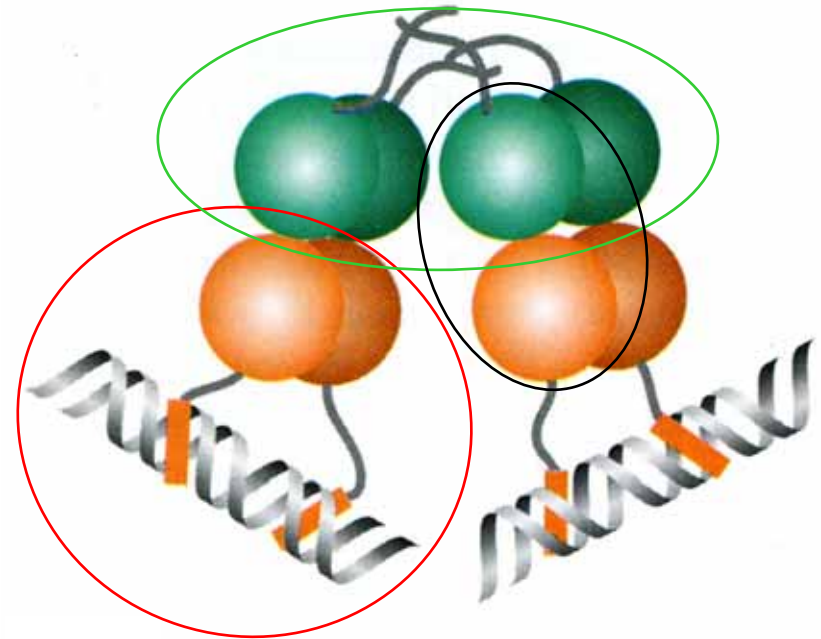
Spesso, domini simili ricorrono in proteine diverse che hanno funzioni diverse e sequenza amminoacidica completamente diversa.



Domini

Esempio:

- uno dei due domini (arancione) può formare dimeri se espresso da solo e legare il DNA con affinità simile alla proteina intera
- l'altro dominio (verde) tetramerizza anche se espresso da solo, ma non lega il DNA



Disegno schematico del tetramero del Lac repressor che lega DNA Ciascun monomero del Lac repressor è costituito da un dominio di tetramerizzazione (in verde) e da un dominio di legame del DNA (in arancione).

Fold (ripiegamenti) proteici

In Natura il **numero di fold proteici** sembra essere **limitato**

Spesso le proteine hanno strutture terziarie simili ad altre già note e/o sono composte da domini che hanno strutture terziarie simili ad altri già noti, sebbene la loro composizione nella proteina possa essere nuova

Quasi tutti i fold delle strutture terziarie scoperti finora compaiono in molte proteine diverse

⇒ le varie combinazioni di un numero limitato di fold proteici genera la diversità proteica riscontrata negli organismi viventi

La natura modulare della struttura delle proteine permette inserzioni e delezioni nella sequenza

⇒ all'interno di un dominio (o di una famiglia proteica) inserzioni/delezioni avvengono in corrispondenza di loop lasciando inalterati gli elementi di struttura secondaria che formano l'intelaiatura del dominio

Fold proteico e funzione

Le proteine possono essere raggruppate a seconda del tipo di fold dei suoi domini

⇒ spesso (**ma non sempre**) la funzione può essere riconosciuta da somiglianze strutturali e di sequenza

Esempio:

- proteine contenenti un dominio chinasi sono quasi sempre chinasi
- proteine contenenti un dominio alfa/beta idrolasi sono quasi sempre idrolasi

L'accoppiamento struttura complessiva-funzione in alcuni casi può essere debole

Esempio:

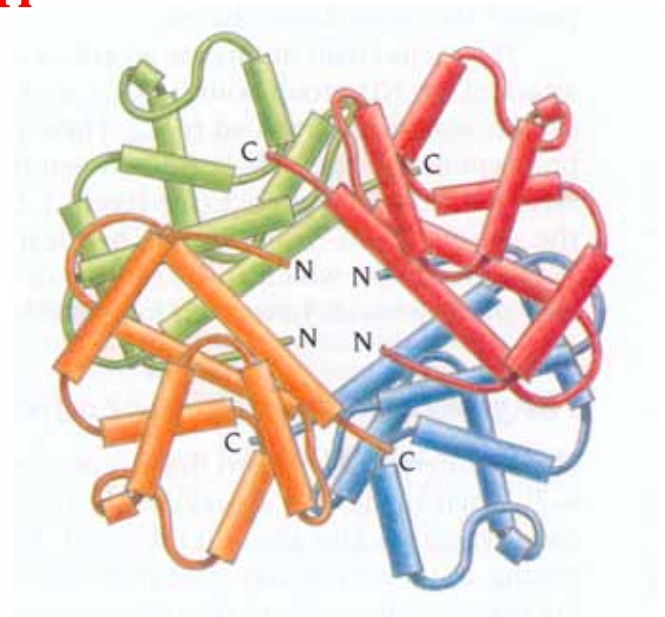
- dominio TIM barrel: comune a molti enzimi con funzioni biochimiche molto diversificate

Struttura quaternaria

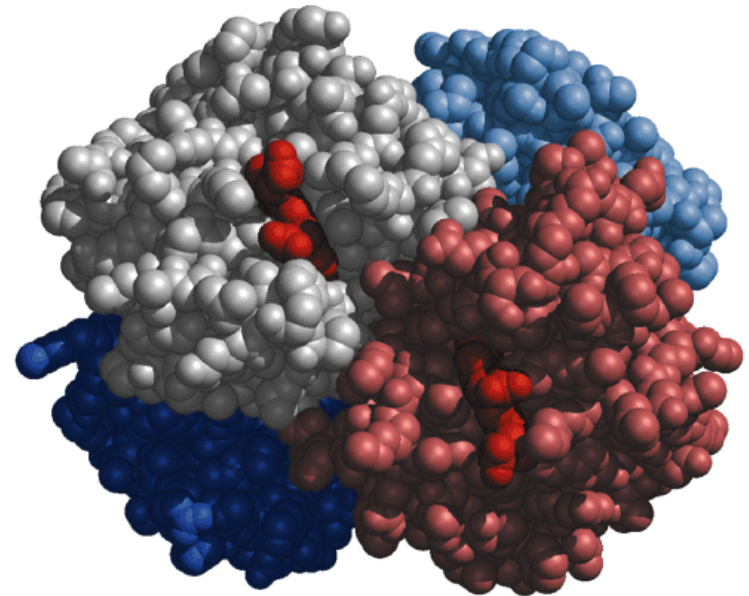
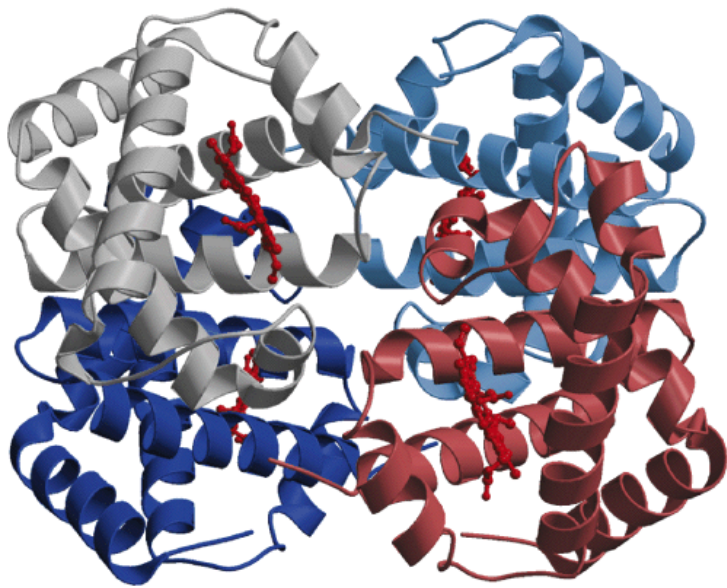
Le proteine che sono costituite da una sola catena polipeptidica sono chiamate **monomeriche**

proteine costituite da un certo numero di catene polipeptidiche identiche, chiamate **subunità**, che si associano in modo specifico a formare una molecola **multimerica** ⇒ **struttura quaternari**

- le subunità possono funzionare in modo indipendente oppure in modo cooperativo
- altre proteine sono costituite da catene polipeptidiche, e quindi da subunità, diverse, ciascuna con una diversa funzione



Struttura quaternaria



Emoglobina: tetramero $\alpha_2\beta_2$

I tipi di legame che determinano la struttura quaternaria sono gli stessi alla base di quella terziaria

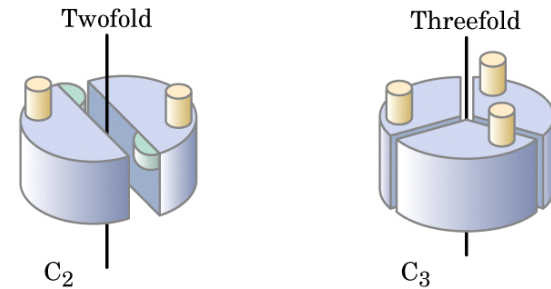
Struttura quaternaria

- eventuali difetti possono essere riparati sostituendo la subunità difettosa.
- il sito di formazione della subunità può essere diverso da quello di assemblaggio della struttura finale e la sola informazione genetica necessaria è specificare le diverse subunità che si devono assemblare.
- nel caso degli enzimi, aumentarne le dimensioni attraverso l'associazione di subunità identiche è più efficiente che aumentare la lunghezza della catena polipeptidica, poiché ogni subunità ha un sito attivo.
- la costituzione a subunità fornisce le basi strutturali per la regolazione della loro attività.

Struttura quaternaria

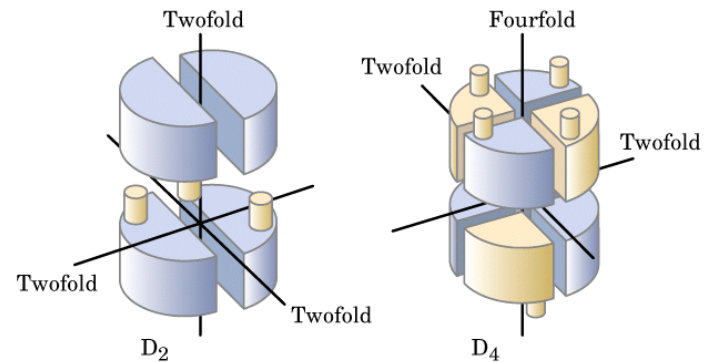
le subunità si dispongono in modo simmetrico, cioè occupano posizioni geometricamente equivalenti secondo **operazioni di simmetria rotazionale**:

- **simmetria ciclica**, in cui le subunità sono correlate da un singolo asse di rotazione



Two types of cyclic symmetry

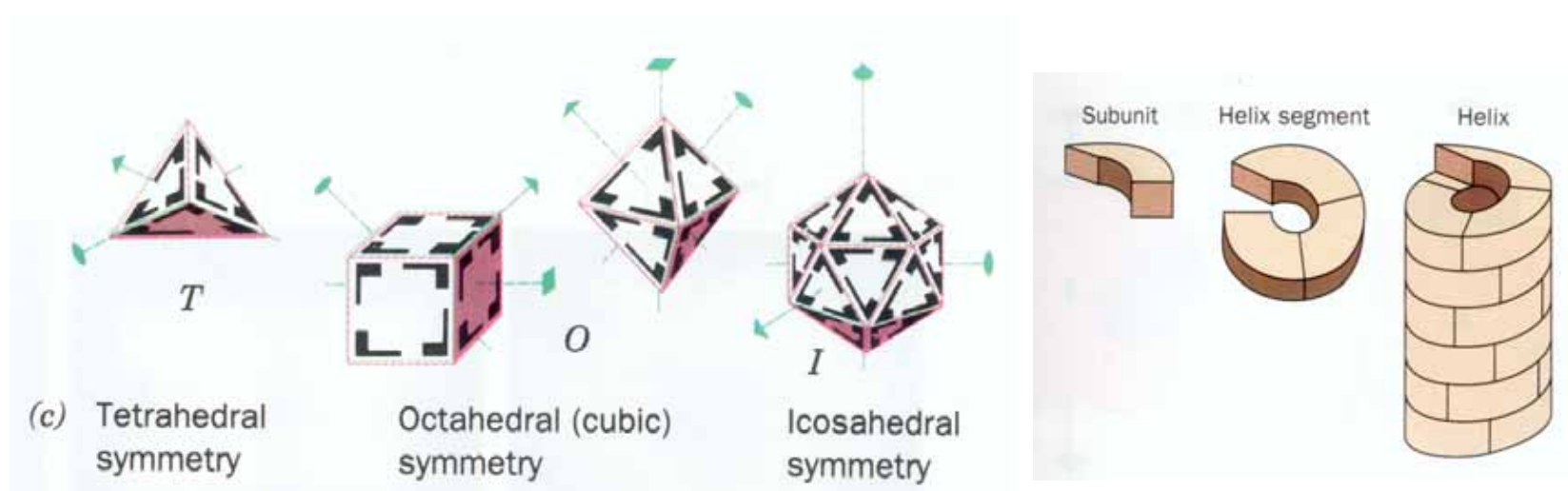
- **simmetria diedra**, caratterizzata dalla composizione di un asse di rotazione di ordine n con un asse di rotazione di ordine 2 (che si intersecano perpendicolarmente)



Two types of dihedral symmetry

Struttura quaternaria

- **altre simmetrie** possibili sono quella tetraedrica, ottaedrica e icosaedrica o quasi-simmetrie cilindriche



Simmetria icosaedrica

Virus a capside sferico

12 vertici (asse di rotazione di ordine 5)

20 facce (asse di rotazione di ordine 3)

30 spigoli (asse di rotazione di ordine 2)

L'icosaedro presenta $12 \times 5 = 60$ (o 20×3 , o 30×2) posizioni equivalenti dette **unità asimmetriche**. Ciascuna unità asimmetrica può contenere una o più catene polipeptidiche.

tutte le proteine costituenti il capside dei virus sferici assumono la struttura β -antiparallela con topologia **jelly roll**

