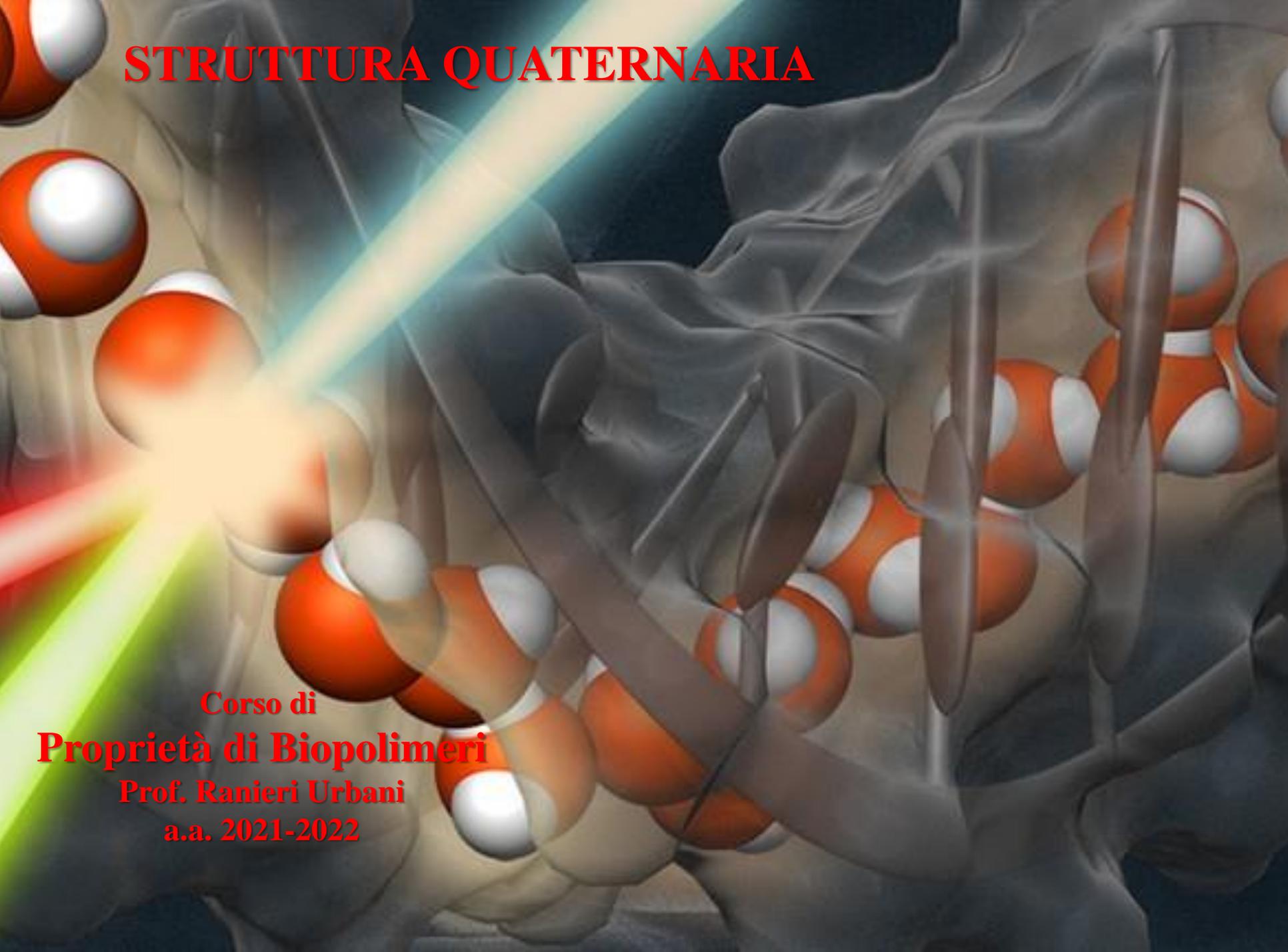


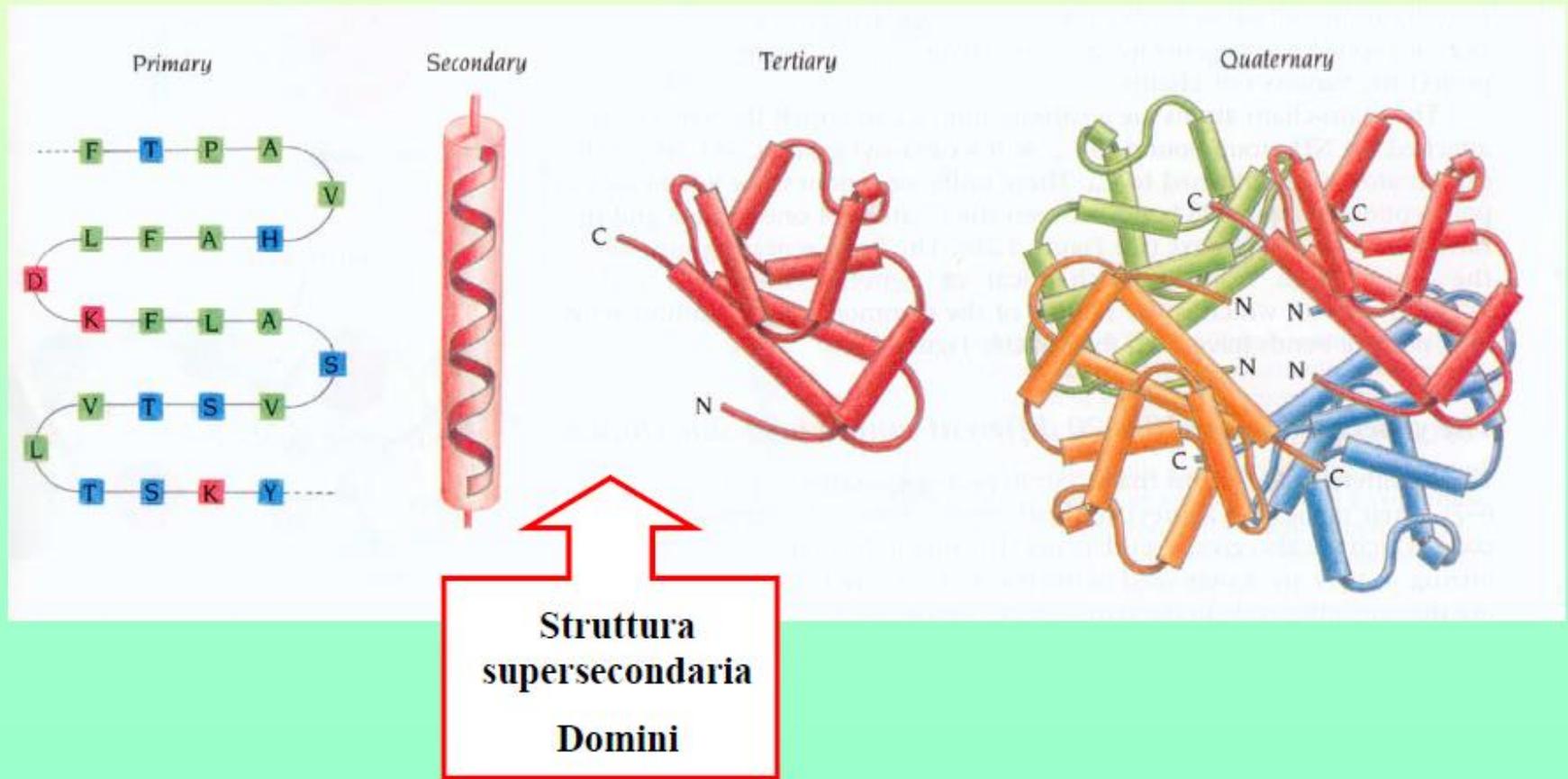
# STRUTTURA QUATERNARIA

A 3D molecular model of a protein quaternary structure. The protein is shown as a grey, semi-transparent surface with a complex, multi-subunit arrangement. Several subunits are visible, each containing a heme-like prosthetic group represented by a red and white sphere. A bright, multi-colored beam of light (yellow, green, blue) passes through the center of the structure, highlighting its internal architecture.

Corso di  
**Proprietà di Biopolimeri**  
Prof. Ranieri Urbani  
a.a. 2021-2022

# La struttura quaternaria

L'ultimo livello nella gerarchia strutturale delle proteine è rappresentato dalla **struttura quaternaria**.



# **STRUTTURA QUATERNARIA**

E' data dall'insieme delle interazioni tra singole catene biopolimeriche (subunità)

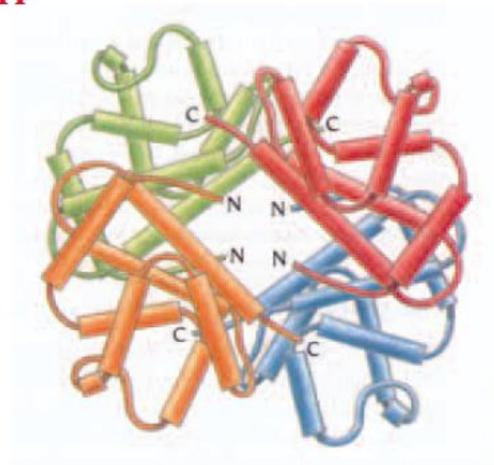
E' stabilizzata da legami non covalenti.

## 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 quaternaria**

- 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

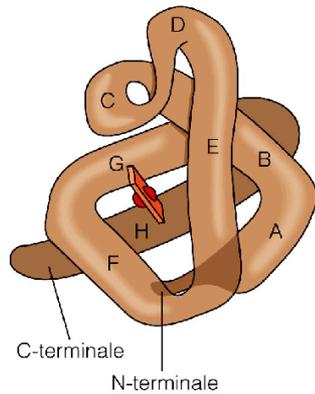


# Caratteristiche principali

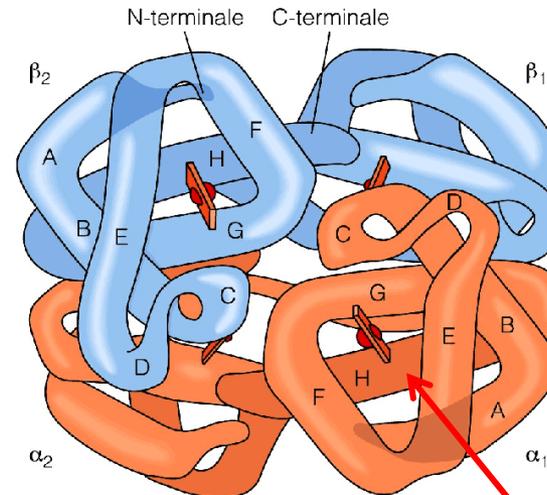
- Proteine formate da più catene polipeptidiche
- Stechiometria definita
- Legami ionici ed idrofobici
- Gli oligomeri sono più stabili delle subunità dissociate
- Siti attivi si possono formare tra le catene
- Il legame di ligandi può cambiare la struttura.

## Proteine coniugate: emoproteine

### Mioglobina



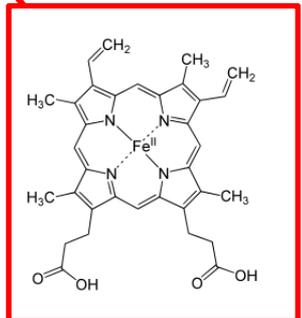
### Emoglobina



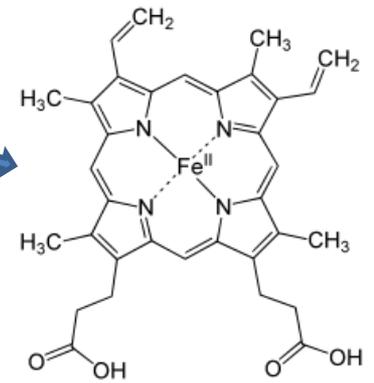
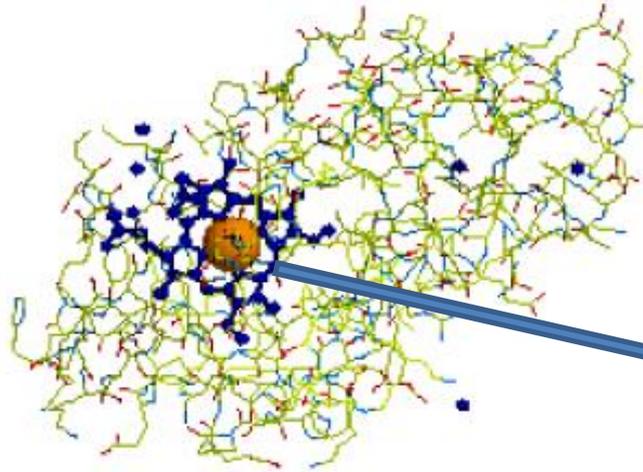
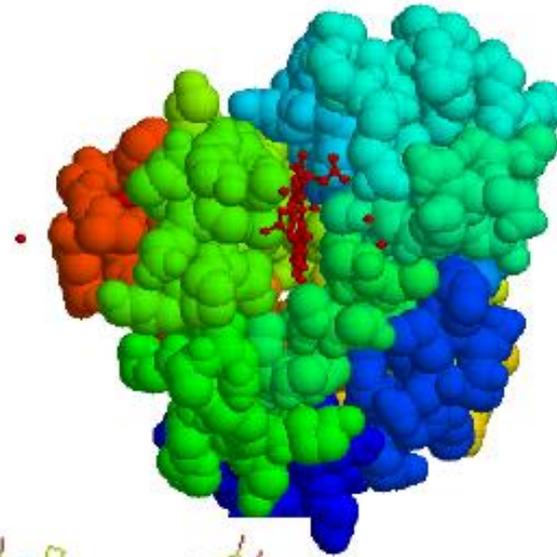
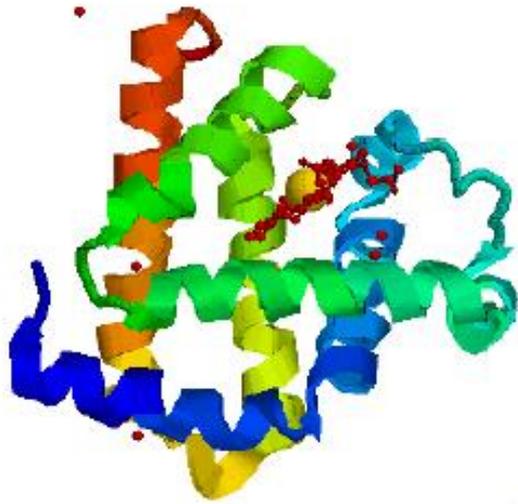
**N.B. Motivo strutturale comune :  $\alpha$ -elica**

La mioglobina è costituita da un'unica catena polipeptidica (struttura III)

L'emoglobina è un tetramero composto da due catene di tipo  $\alpha$  e due catene di tipo  $\beta$  (struttura IV)



# Mioglobina

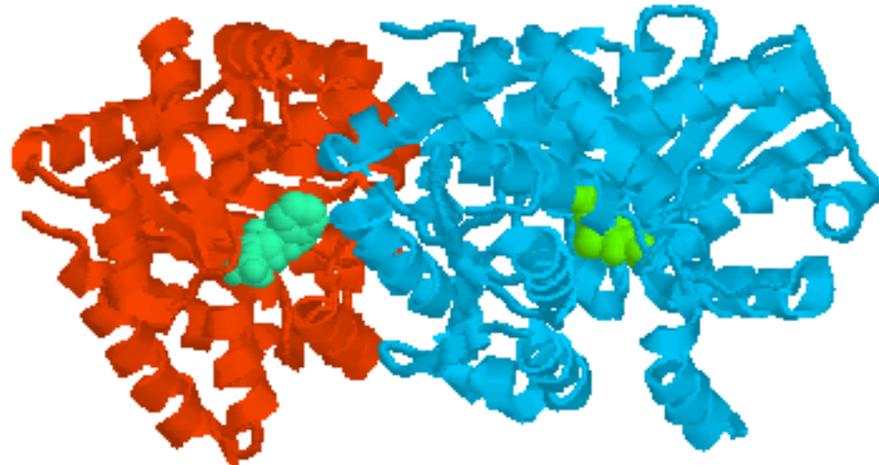


# Strutture etero-multimeriche



- Può succedere che due catene possono unirsi per formare una struttura "secondaria" di dimensioni maggiori
- Questo è il caso della lectina (2LTN) dove il  $\beta$ -sheet è fatto di filamenti provenienti da differenti catene polipetidiche.

## Connessione covalente



- *E. coli* produce un enzima che catalizza sia l'isomerizzazione del fosfo-ribosil antranilato E la sintesi di indolo-gliceril fosfato, (biosintesi del triptofano)
- È formato da due  $\alpha\beta$  barrels molto simili, ognuno dei quali ha una propria attività enzimatica.

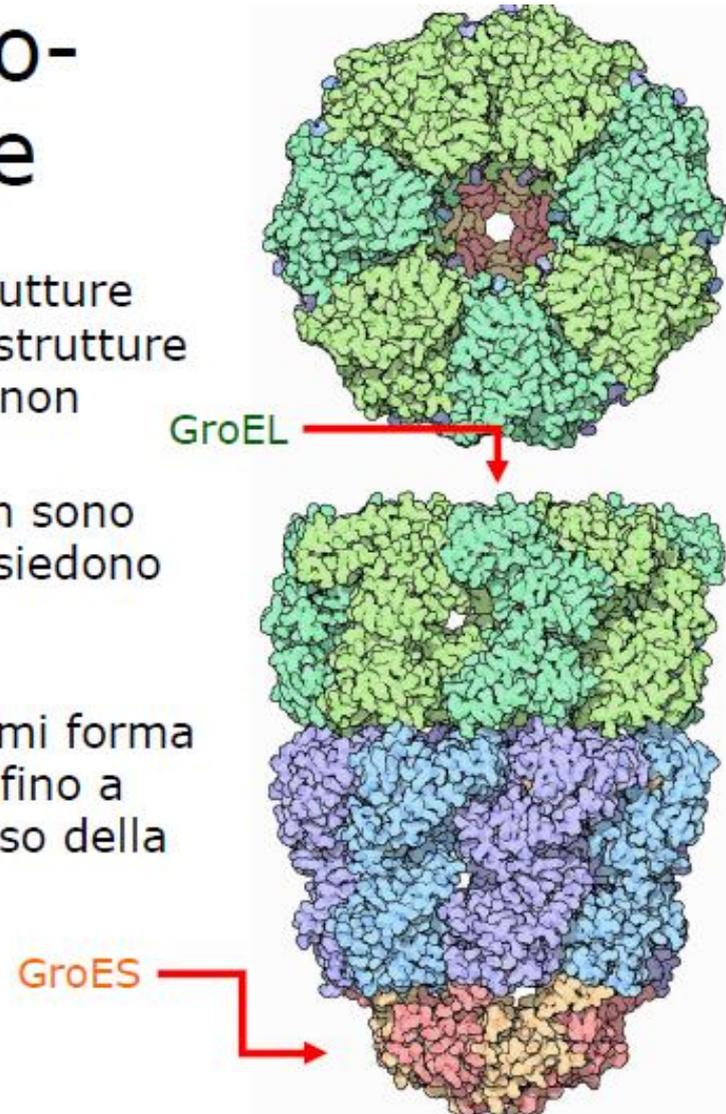
# Strutture etero-multimeriche

- Differenti strutture si aggregano per formare una unica struttura.



# Strutture omo-multimeriche

- È molto comune trovare strutture multimeriche con la stesse strutture terziarie associate in modo non covalente.
- Spesso queste strutture non sono simmetriche, in genere possiedono una simmetria rotazionale attraverso uno più assi.
- La maggior parte degli enzimi forma dimeri, trimeri, tetrameri ... fino a tetradecameri (come nel caso della chaperonina GroEL).

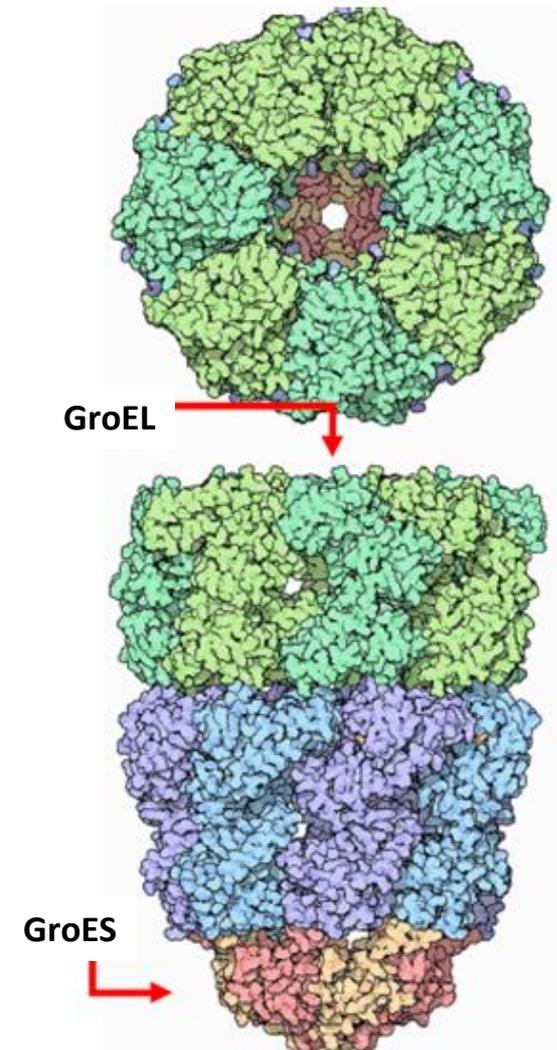


## chaperonina GroEL)

Gli **chaperones molecolari** comprendono diverse famiglie di proteine altamente conservate; molti di questi sono anche **Heat Shock Proteins** (Hsp).

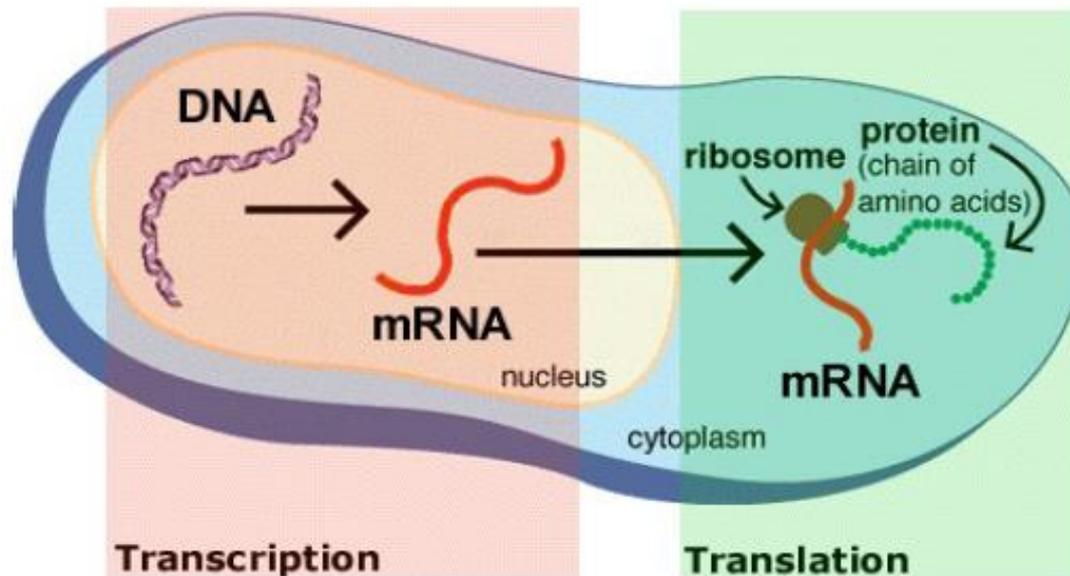
Il loro ruolo è legato al mantenimento della struttura delle proteine secondo quattro aspetti principali:

- a) assicurano il raggiungimento e il mantenimento del corretto stato conformazionale delle catene polipeptidiche appena sintetizzate;
- b) dirigono l'assemblaggio di complessi multienzimatici;
- c) partecipano al mantenimento o alla creazione di uno stato di parziale denaturazione delle proteine, favorendone così il trasporto attraverso le membrane dei mitocondri o dei plastidi;
- d) stabilizzano le proteine danneggiate formatesi a seguito di stress chimici o fisici facilitandone la rinaturazione e/o la degradazione.



## Il ripiegamento corretto è indispensabile per il funzionamento delle proteine

La biosintesi proteica rappresenta l'ultima tappa del percorso molecolare seguito dall'informazione genetica che procede dal DNA, al RNA, alle proteine.



Il prodotto finale della biosintesi proteica non è una proteina perfettamente funzionante, infatti, per essere definita tale, deve anzitutto ripiegarsi, deve andare incontro al fenomeno del “folding” (ripiegamento).

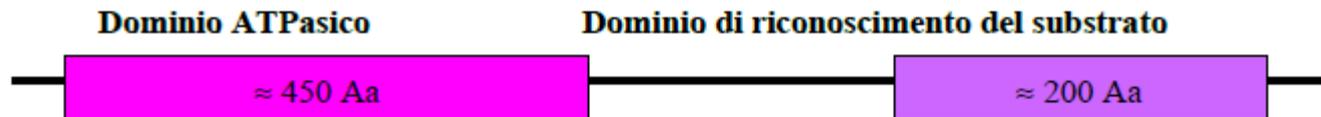
In realtà le Hsp 70 appartengono ad una sottoclasse di chaperonine che fanno capo ad una classe più vasta di proteine prodotte dalla cellula, non solo, in risposta all'aumento della temperatura, ma a svariati stimoli o agenti tossici quali:

**Fattori ambientali:** shock termico, metalli pesanti, agenti chemioterapici

**Stati patologici:** infezioni virali, infiammazioni, febbre, neoplasie.

**Fattori cellulari normali:** ciclo di divisione cellulare, fattori di crescita

Tra i vari chaperon molecolari esiste una forte omologia di sequenza, com'è emerso dal confronto della sequenza amminoacidica di varie proteine da stress. Esse possiedono all'estremità N-terminale una sequenza di circa 450 Aa detta dominio ATPasico, ed una sequenza di circa 200 Aa all'estremità C-terminale detta dominio di riconoscimento del substrato.



Tramite gli amminoacidi che formano il dominio di riconoscimento, i chaperon molecolari riconoscono e legano le proteine da assistere, mentre tramite il dominio ATPasico riconoscono e legano l'ATP da utilizzare come fonte d'energia per compiere la loro azione ausiliaria nel ripiegamento corretto delle proteine mal ripiegate o denaturate.

## Funzione generale dei chaperon molecolari

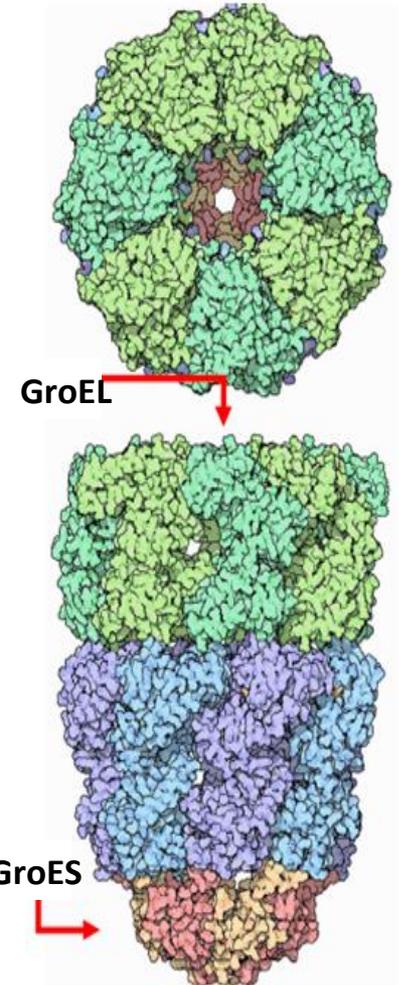
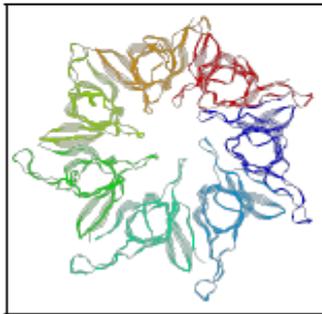
I chaperon molecolari svolgono la loro azione legandosi a residui idrofobici esposti della proteina target. La loro azione, in questo modo, è duplice:

1. legandosi ai residui idrofobici, impediscono che questi possano interagire con altri residui idrofobici di proteine vicine nel tentativo di proteggersi dal contatto con l'acqua, e quindi impediscono la formazione degli aggregati insolubili.
2. impediscono che i residui idrofobici della stessa proteina si aggregino prematuramente tra loro originando una conformazione errata.

## Complesso GroEL/GroES

GroEL è costituita da 14 subunità che si riuniscono in maniera tale da formare due gruppi, di 7 subunità ciascuno. Questi due gruppi sono sovrapposti e formano due alloggiamenti uno superiore ed uno inferiore.

GroES è invece costituita da 7 subunità unite assieme a formare una sorta di unico disco.

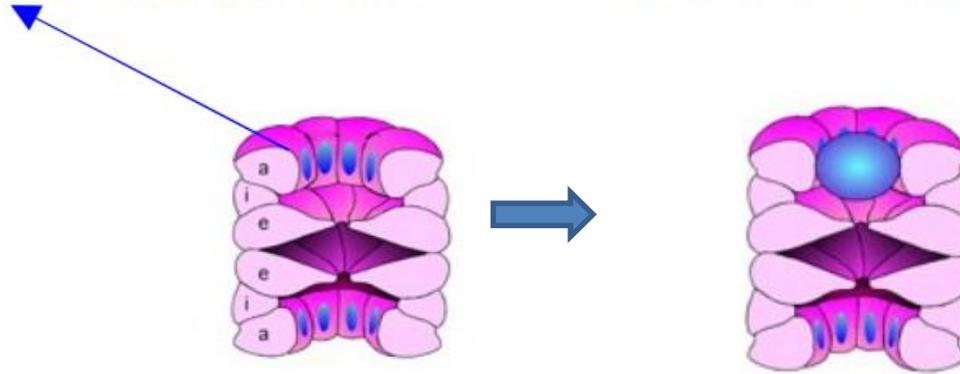


Il meccanismo d'azione del complesso GroEL/GroES, è paragonabile a quello di una gabbia che imprigiona la proteina non ancora ripiegata e che si apre liberando le proteine solo a ripiegamento avvenuto correttamente.

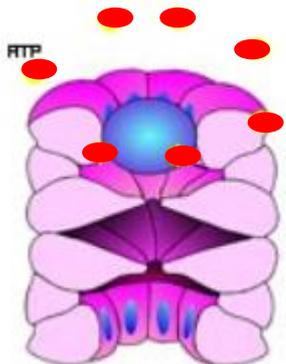
Per questa ragione si dice che GroEL/GroES funga da gabbia: **LA GABBIA DI ANFISEN**

## LA GABBIA DI ANFISEN

In condizioni di riposo la superficie della cavità interna di tali alloggiamenti espone verso il centro del canale: **residui amminoacidici idrofobici** (quelli che nella figura sottostante sono rappresentati in blu elettrico).

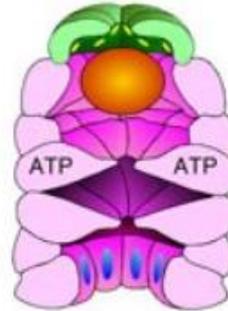


La proteina non ripiegata e che pertanto presenta i suoi amminoacidi idrofobici non ancora racchiusi all'interno della sua struttura, potrà allora interagire con gli amminoacidi idrofobici dell'alloggiamento superiore di GroEL.



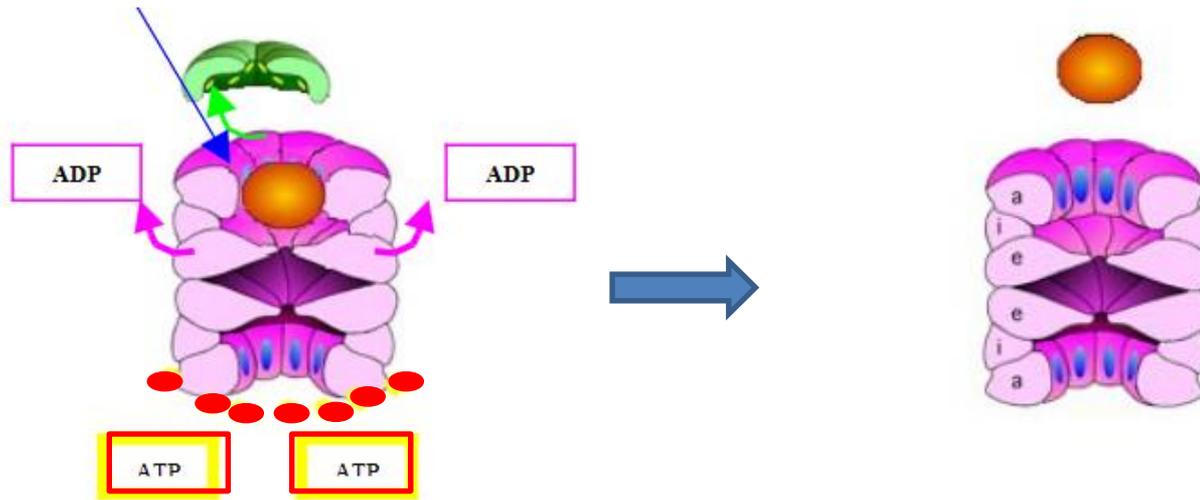
Una volta che il polipeptide entra nella cavità superiore, intervengono 7 molecole di ATP che si legano in prossimità dei domini ATPasici di ciascuna delle 7 subunità costituenti l'alloggiamento superiore.

Assieme all'ATP, entra in gioco anche GroES che si dispone al di sopra dell'alloggiamento in cui è contenuta la proteina, come fosse un tappo.....**La proteina a questo punto è proprio in gabbia!!!**



L'idrolisi dell'ATP a ADP da parte di ciascuna delle 7 sub unità fa sì che in esse si verifichi una modificazione conformazionale tale che, le subunità non esporranno più, verso il centro della cavità, i loro amminoacidi idrofobici, bensì i loro amminoacidi idrofilici (come si vede dalla figura, infatti,

A questo punto intervengono altre 7 molecole di ATP che si legano, stavolta, in prossimità dei siti ATPasici delle 7 subunità costituenti l'alloggiamento inferiore; in questo modo saranno allontanati sia GroES (aprendo la gabbia) sia l'ADP (derivante dall'idrolisi dell'ATP che si era precedentemente legato alle sub unità superiori). A questo punto, tali subunità esportano nuovamente verso la cavità interna dell'alloggiamento, gli Aa idrofobici



La proteina correttamente ripiegata può così uscire dalla trappola in cui era imprigionata, infatti, essa non può più interagire con la superficie della cavità che è tornata idrofobica.

Il **dogma di Anfinsen**, conosciuto anche come ipotesi termodinamica di Anfinsen, è un postulato della biologia molecolare espresso dal premio Nobel per la chimica Christian B. Anfinsen.

Il dogma afferma che, almeno per le piccole proteine globulari, la struttura nativa è determinata **solamente** dalla sequenza di amminoacidi che costituiscono la proteina.

Ciò equivale a dire che, nelle condizioni ambientali (temperatura, concentrazione del solvente e composizione, ecc.) alle quali avviene il ripiegamento proteico, la struttura nativa corrisponde a un unico minimo di energia libera, stabile e cineticamente accessibile.

## dogma di Anfinsen

**Unicità:** richiede che la sequenza non possieda qualche altra configurazione dotata di energia libera comparabile. Quindi il minimo di energia libera deve essere univoco.

**Stabilità:** piccoli cambiamenti nell'ambiente circostante non possono produrre cambiamenti nella configurazione a energia minima.

**Accessibilità cinetica:** significa che il percorso nella superficie di energia libera dallo stato denaturato a quello con riavvolgimento proteico deve essere tale che il ripiegamento della catena non implichi cambiamenti altamente complessi nella forma

Come la proteina raggiunga questa struttura rientra nel campo di studio del ripiegamento proteico, che si basa un dogma correlato chiamato **paradosso di Levinthal**.

## Paradosso di Levinthal

molte proteine si ripiegano nella loro conformazione nativa in meno di pochi secondi

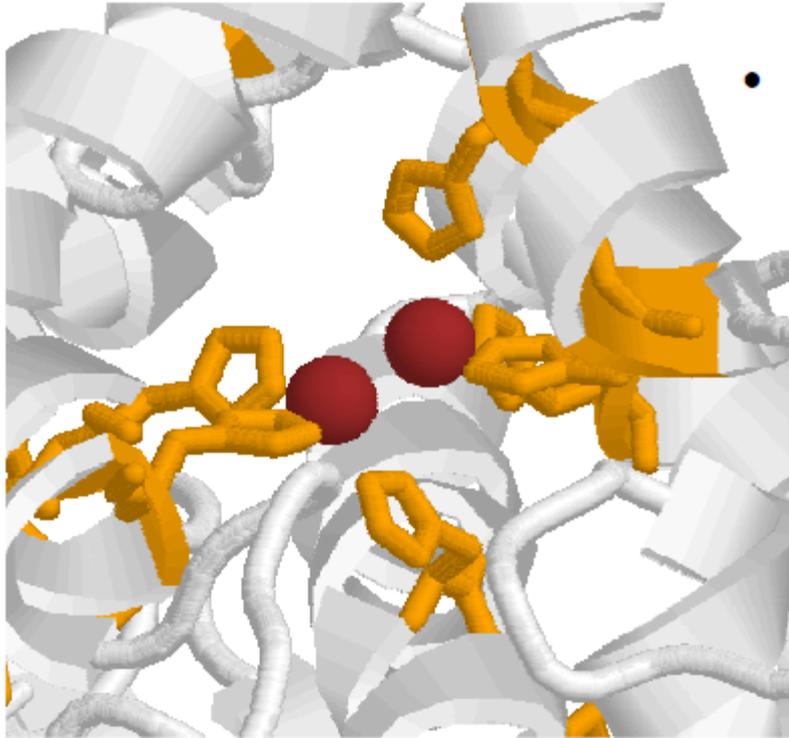
- ⇒ deve esistere un **percorso** abbastanza **preciso e diretto di ripiegamento**
- ⇒ durante il corretto avvolgimento, la stabilità della struttura proteica aumenta rapidamente (e l'energia libera diminuisce rapidamente) rendendo il **processo sostanzialmente irreversibile**

# Gruppi prostetici

- Metalli
  - Legati direttamente ( $\text{Fe}^{++}$ ,  $\text{Fe}^{+++}$ ,  $\text{Zn}^{++}$ ,  $\text{Ca}^{++}$ ,  $\text{Co}^{++}$ ,  $\text{Cu}^{++}$ ,  $\text{Mg}^{++}$ ,  $\text{Mn}^{++}$  ...)
  - Attraverso un gruppo eme ( $\text{Fe}^{++}$ ,  $\text{Fe}^{+++}$ ,  $\text{Mg}^{++}$ ...)
  - Come complessi con lo zolfo ( $\text{Fe}^{++}$ ,  $\text{Fe}^{+++}$ )
- Biotina
- Acido lipoico
- Retinale
- Piridossale
- ...

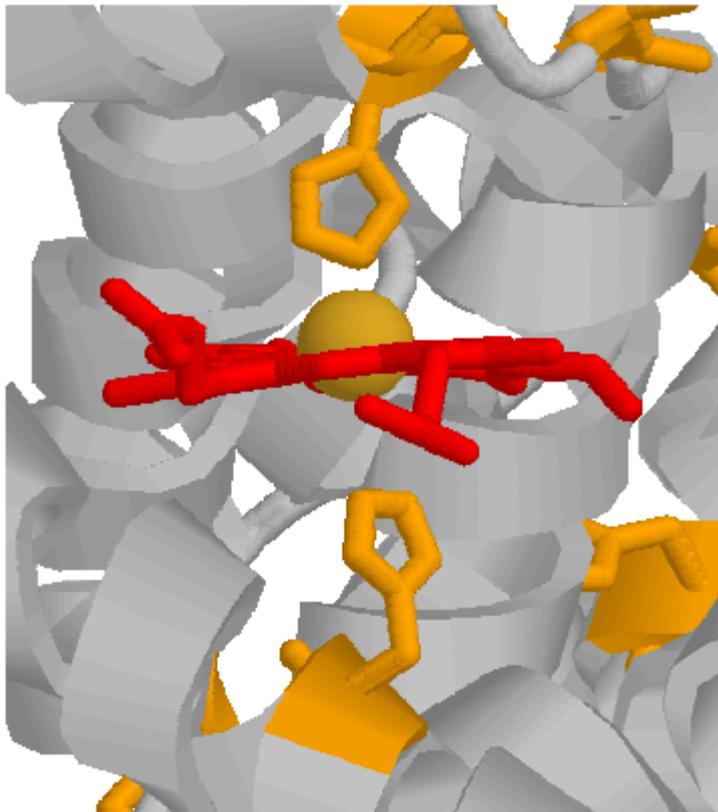


# Legati direttamente



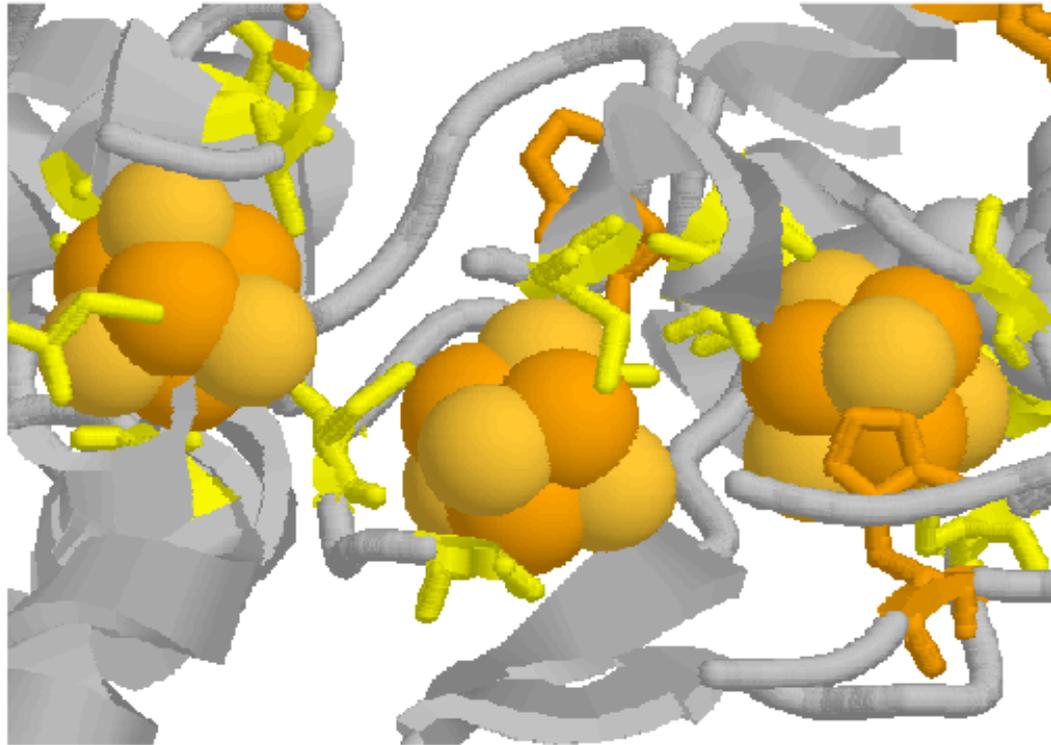
- Cu<sup>2+</sup> nell'emocianina

# Legati attraverso un gruppo eme



- Fe<sup>2+</sup> nell'emoglobina

# Come complessi FeS

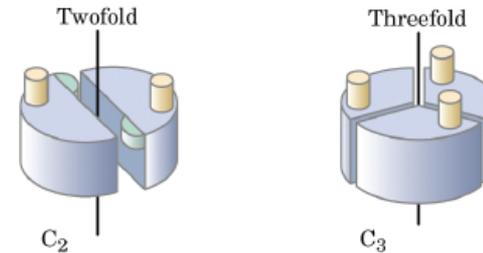


- Clusters  $\text{Fe}^{++}\text{-S}$  nella citocromo ossidasi

## Struttura quaternaria

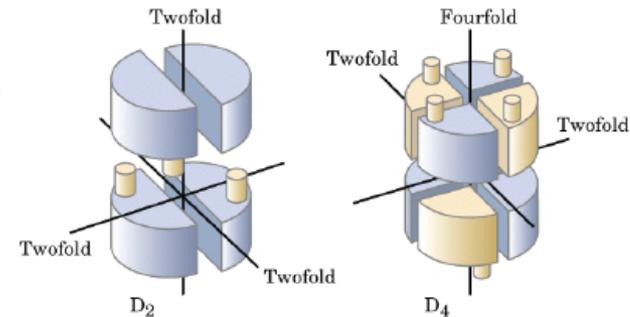
le subunità si dispongono in modo simmetrico, cioè occupano posizioni geometricamente equivalenti secondo **operazioni di simmetria rotazionale**:

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



Two types of cyclic symmetry

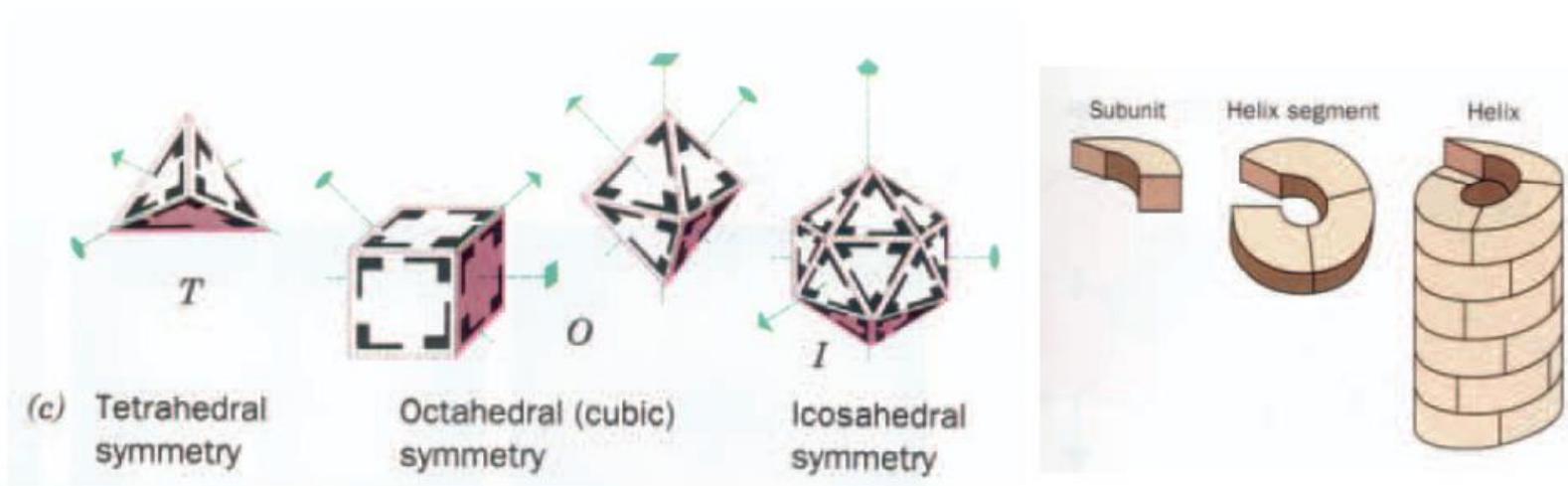
- **simmetria diedra**, caratterizzata dalla composizione di un asse di rotazione di ordine  $n$  con un asse di rotazione di ordine 2 (che si intersecano perpendicolarmente)



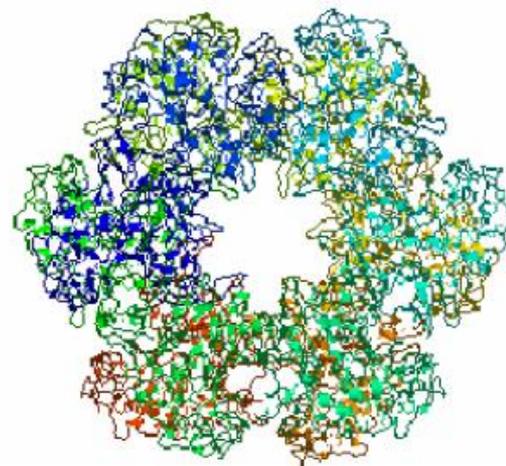
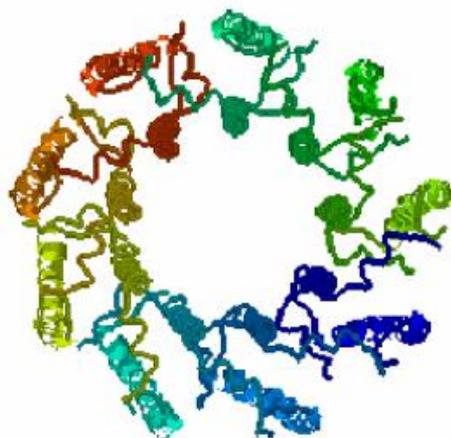
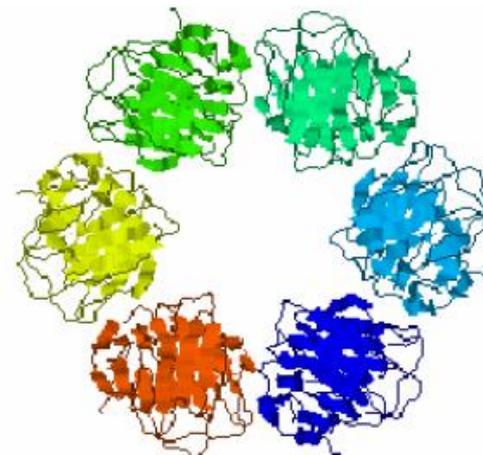
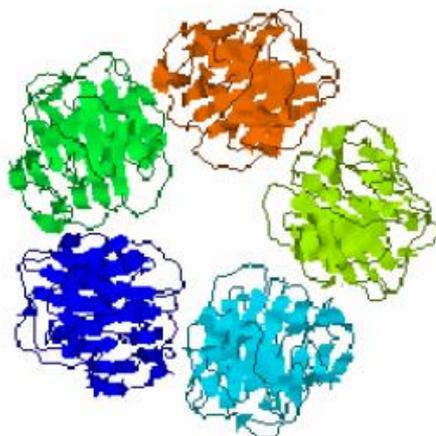
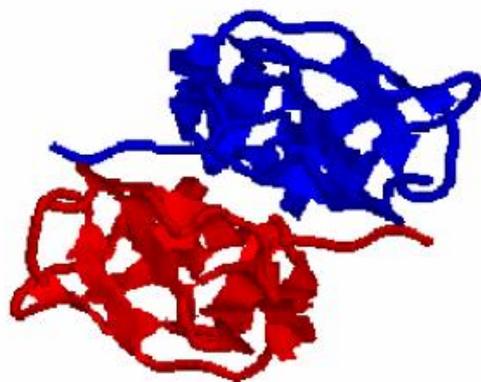
Two types of dihedral symmetry

## Struttura quaternaria

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



## Strutture quaternarie e simmetrie



# I virus

I virus sono parassiti costituiti da molecole di acido nucleico (DNA o RNA, a doppia o singola elica) circondate da un rivestimento protettivo costituito da proteine, chiamato **capside**. Alcuni virus hanno anche una membrana a doppio strato lipidico, chiamato **pericapside** o **envelope**, che riveste il capsid.

Un acido nucleico non potrebbe mai codificare una singola molecola proteica tanto grande da racchiuderlo. Il capsid dei virus, quindi, è costituito da innumerevoli copie di una o alcune catene polipeptidiche.

I virus hanno forme e dimensioni diverse: in particolare, il capsid può essere sferico o a bastoncino.

La conoscenza in dettaglio (struttura tridimensionale a livello atomico) del capsid dei virus potrebbe essere importante per **progettare farmaci antivirali**.

Per esempio, l'entrata dei virus nelle cellule dei mammiferi dipende dal riconoscimento specifico, da parte del capsid virale, di proteine o carboidrati (**recettori**) che si trovano sulla superficie delle cellule bersaglio. Questa interazione gioca un ruolo cruciale e composti che interferiscono con questa interazione dovrebbero bloccare l'infezione.

I siti di legame virale sulla superficie delle cellule ospiti sono anche i **target degli anticorpi**, per cui il sistema immunitario dell'ospite ha con tutta probabilità giocato un ruolo importante nell'evoluzione dei virus animali. La conoscenza, quindi, delle regioni di superficie di questi virus dovrebbe aiutare a capire meglio gli adattamenti strutturali che i virus hanno evoluto per eludere le difese dell'ospite.

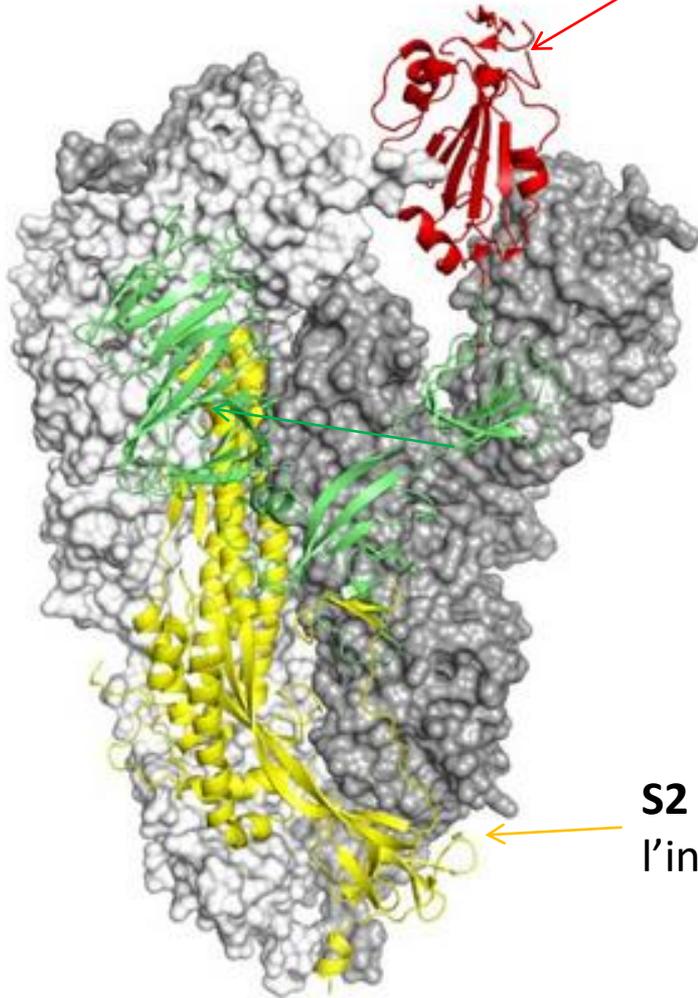
# Cos'è e come funziona la proteina Spike dei Coronavirus

Una tra le proteine bersaglio del virus più interessanti a questo scopo è la **proteina Spike (S)**. Questa proteina decora la superficie del virus formando delle protuberanze caratteristiche (**facendolo sembrare una corona – da cui il nome “Coronavirus”**). La proteina S si suddivide in due parti:

- **S1**, che contiene una regione che serve a legarsi alla cellula bersaglio aderendo al recettore ACE2;
- **S2**, che in una seconda fase consente l'ingresso del virus nella cellula.

Quindi, una molecola che fosse capace di **impedire l'interazione tra la proteina Spike e il recettore ACE2** sarebbe potenzialmente in grado di prevenire l'infezione da coronavirus e, di conseguenza, la malattia.

la porzione che serve ad interagire con il recettore ACE2 per infettare le cellule



La catena **S1** è colorata in verde

**S2** la parte della molecola necessaria per l'ingresso del virus nella cellula

## I virus

Le proteine che costituiscono il capsido dei virus si dispongono in modo simmetrico o quasi-simmetrico.

Sono note due possibili disposizioni (assemblaggi di natura quaternaria, quindi):

- virus con **capside elicoidale**, in cui le subunità proteiche si associano a formare un bastoncino elicoidale;
- virus con **capside sferico**, in cui le subunità proteiche si aggregano a formare un poliedro.

Considerato l'elevato numero di subunità proteiche (fino ad alcune centinaia) coinvolte nella formazione del capsido dei virus, essi rappresentano

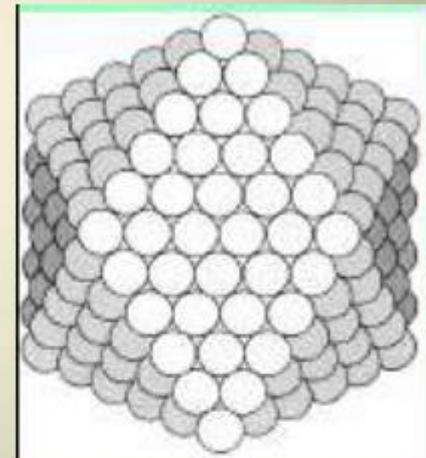
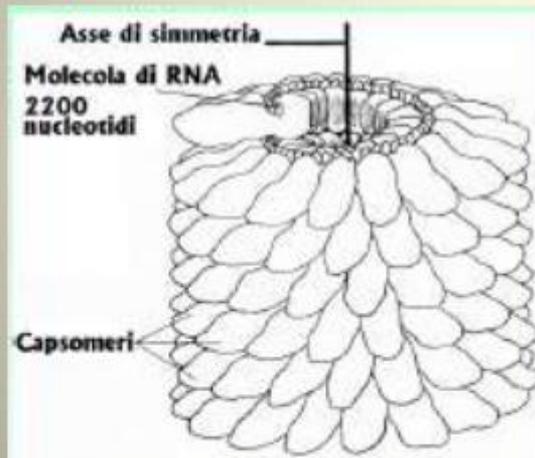
**l'esempio più complesso di organizzazione delle proteine in struttura quaternaria.**

# SIMMETRIA CAPSIDICA

Le proteine destinate a formare il capside sono capaci di autocombinarsi (**assemblaggio**) seguendo due tipi di schemi fondamentali, definiti:

- simmetria **ELICOIDALE**

- simmetria cubica (detta anche **ICOSAEDRICA**)



La simmetria virale è dovuta all'interazione genoma-proteine.

## ESEMPIO:

### The Contribution of Tobacco Mosaic Virus to Biology and Medicine

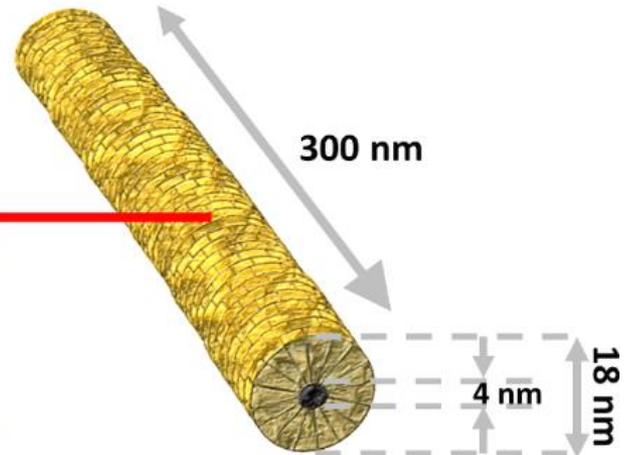
Uninfected leaf



TMV-infected leaf



TMV particle



Colpisce:

- barbabietola da zucchero
- cetriolo
- mais
- patata
- pomodoro
- pisello
- tabacco

È un virus altamente infettivo che si trasmette per contatto senza il concorso di vettori.

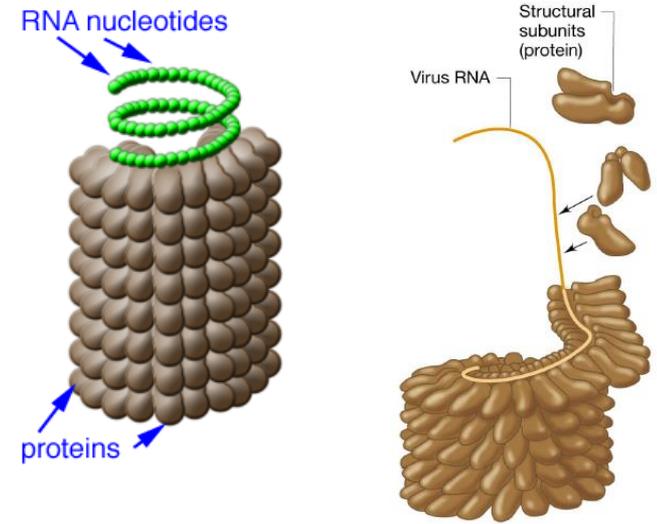
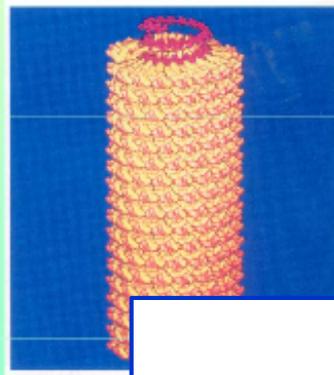
## Virus con capside elicoidale

Un esempio di virus con capside elicoidale è dato dal **virus del mosaico del tabacco (TMV)**.

TMV ha una forma a bastoncino, è lungo circa **3000 Å**, ha un **diametro di 180 Å** e ha una massa di **40 milioni di dalton**.

Il suo capside è costituito da circa **2130 copie identiche di una proteina (158 aminoacidi)**, disposte a formare un'elica destrorsa cava (diametro cavità circa **40 Å**), caratterizzata da  $n = 16.3$  subunità per giro,  $p = 23$  Å.

TMV contiene una molecola di RNA a singola elica che si avvolge nella cavità interna seguendo i giri dell'elica di proteine, in modo tale che **3 nucleotidi interagiscono con una subunità proteica**.



Il capside è costituito dalla associazione di subunità proteiche in protomeri ripetuti (capsomeri)

5 subunità (pentameri) o da 6 subunità (esameri)

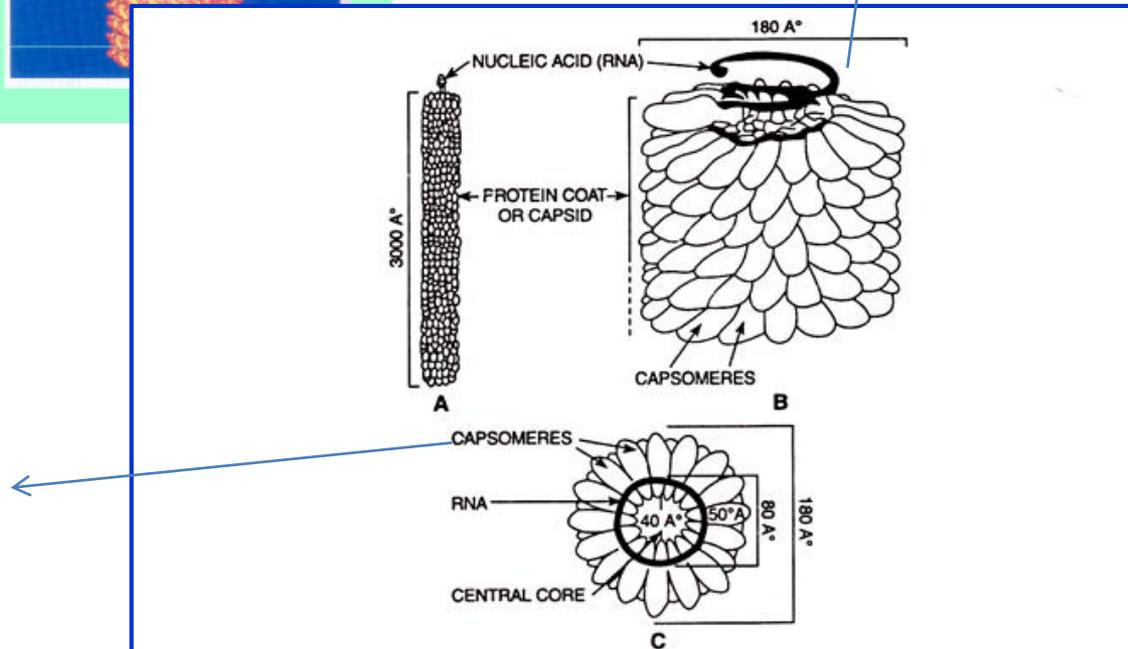
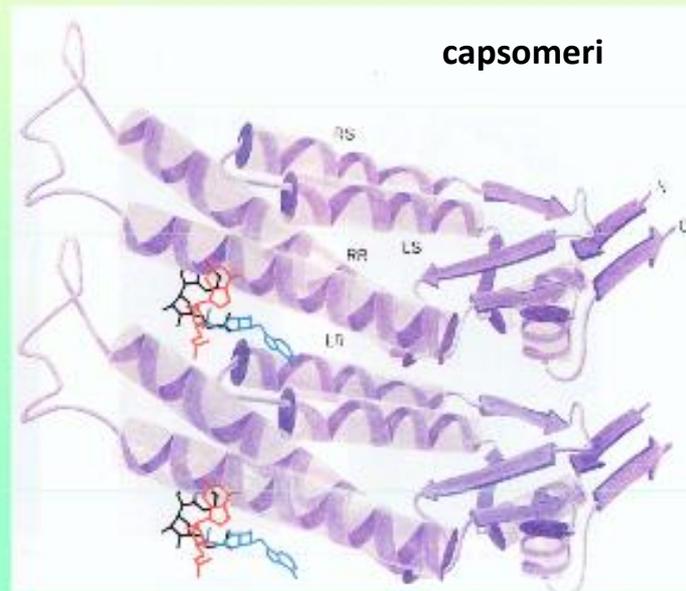


FIG. 13.20. Tobacco mosaic virus (TMV). A. surface view; B. an enlarged portion showing RNA-capsomere arrangement; C. view in section.

## Virus con capside elicoidale

L'analisi cristallografica a raggi X ha mostrato che ciascuna subunità del capside del virus MTV è costituita da un **bundle di 4 eliche**, disposto in modo radiale rispetto all'asse dell'elica del virus.

asse elica virus



## Virus con capside sferico

Un altro modo in cui subunità equivalenti si dispongono a costituire il capside dei virus è quello di formare un poliedro sferico, secondo due principi base:

- **specificità**: le subunità devono riconoscersi con precisione per formare la giusta interfaccia di interazioni non-covalenti, visto che i virus si assemblano spontaneamente dai loro componenti individuali;

- **economia genetica**: il capside è costituito da molte copie di alcuni tipi di subunità.

Questi due principi implicano la simmetria: pattern di legame ripetuti e specifici di unità costituenti identiche portano ad una struttura finale simmetrica.

## Virus con capside sferico

Come disporre in modo simmetrico oggetti identici in modo tale da costruire una sfera?

Esiste solo un numero limitato di modi per farlo e fra questi la **simmetria icosaedrica** è la più alta possibile, permettendo al maggior numero di oggetti identici di formare una sfera.

Lo studio del capsido dei virus sferici con il microscopio elettronico e con la cristallografia a raggi X ha indicato che la simmetria icosaedrica sta alla base della loro architettura.

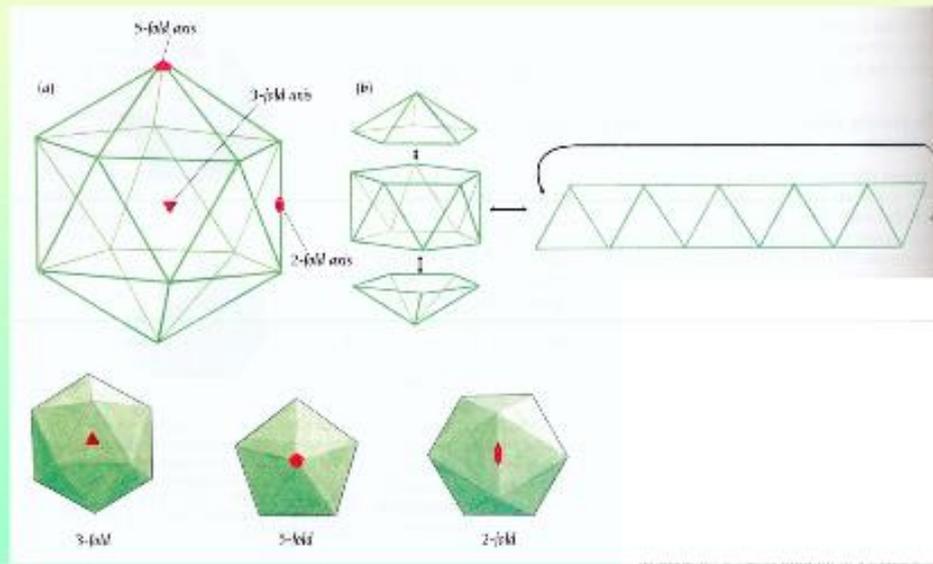
## Virus con capside sferico

L'**icosaedro** è un oggetto approssimativamente sferico costituito da 20 triangoli equilateri identici. La sua simmetria è descritta da tipi diversi di rotazioni, che lo portano a coincidere con se stesso.

**12 vertici**  
(asse di rotazione di ordine 5)

**20 facce**  
(asse di rotazione di ordine 3)

**30 spigoli**  
(asse di rotazione di ordine 2)



Kin exercises/Viral-coat

## Virus con capside sferico

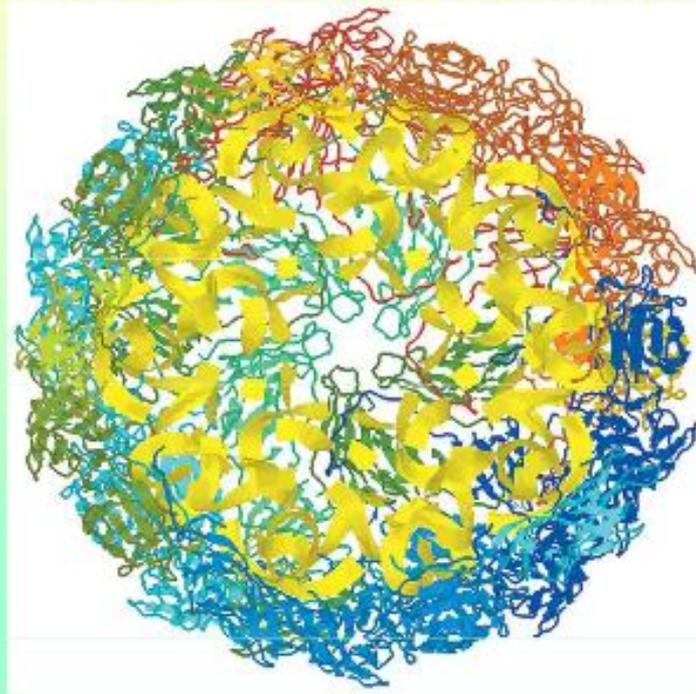
L'icosaedro presenta  $12 \times 5 = 60$  (o  $20 \times 3$ , o  $30 \times 2$ ) posizioni equivalenti.

Ogni oggetto che presenta simmetria è costituito da parti più piccole, identiche e messe in relazione una con l'altra da operazioni di simmetria; tali parti vengono chiamate "unità connesse da simmetria" o "**unità asimmetriche**".

Le subunità proteiche che costituiscono il capside dei virus sferici sono unità asimmetriche: il numero minimo di subunità proteiche necessarie a formare il capside con simmetria icosaedrica è quindi uguale alle posizioni equivalenti dell'icosaedro.

Le proprietà di simmetria dell'icosaedro non si limitano alla superficie, ma si estendono all'intero volume. Un'unità asimmetrica è quindi una parte di questo volume: è un cuneo che parte dalla superficie, fino al centro dell'icosaedro.

X-Ray structure of satellite tobacco mosaic virus (STMV). The virion is shown in cutaway view along its icosahedral (T-1) 5-fold axis.



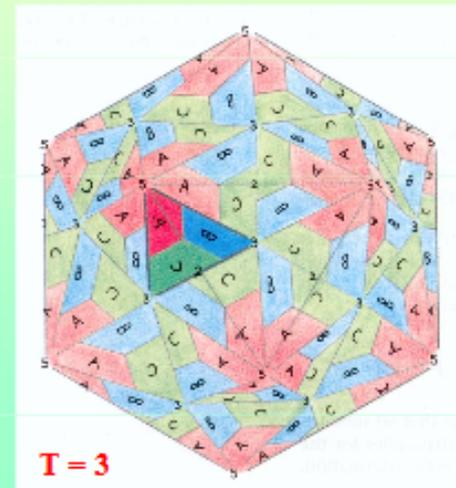
## Virus con capside sferico: $T = 3$

Nel caso di  $T = 3$  (180 subunità), ciascuna unità asimmetrica contiene 3 subunità proteiche (A, B e C) con un diverso environment.

Le subunità **A interagiscono intorno all'asse di rotazione di ordine 5**, mentre le subunità **B e C si alternano intorno all'asse di rotazione di ordine 3**.

Ci sono, quindi, 6 subunità (3 B e 3 C) disposte in modo pseudo-simmetrico intorno all'asse di rotazione di ordine 3, che diventa quindi anche un asse di pseudo-simmetria di ordine 6.

La disposizione delle subunità intorno all'asse di simmetria di ordine 5 e a quello di pseudo-simmetria di ordine 6 dovrebbe essere molto simile, con piccole alterazioni delle modalità di impaccamento.



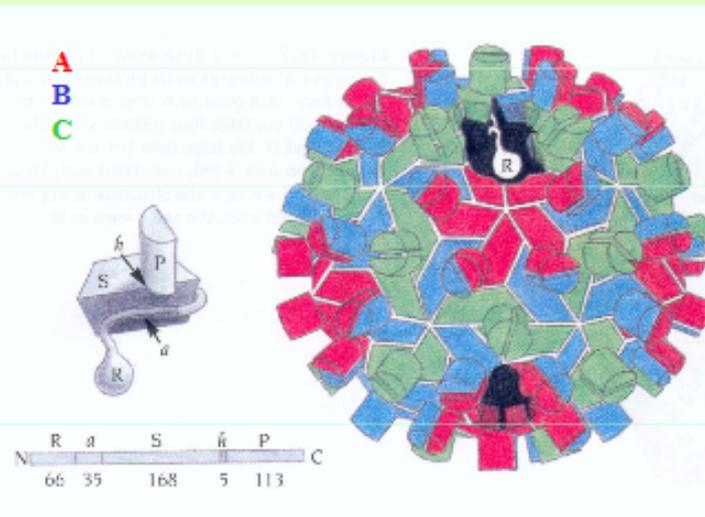
Numeri di triangolazione T

## Virus con capside sferico: $T = 3$

Un **esempio** di capside di virus sferico con  $T = 3$  è dato dal virus del pomodoro (tomato bushy stunt virus). Questo virus contiene 180 catene polipeptidiche chimicamente identiche (386 aminoacidi) per un diametro totale di 330 Å.

Ciascuna subunità è costituita da un **dominio R**, una regione di loop (a) che connette R con il **dominio S**, a formare il guscio virale, e da un **dominio P** che si proietta fuori dalla superficie. I domini P ed S sono collegati da una cerniera (h).

Le catene polipeptidiche identiche hanno diverse strutture tridimensionali quando formano le 3 subunità A, B e C.



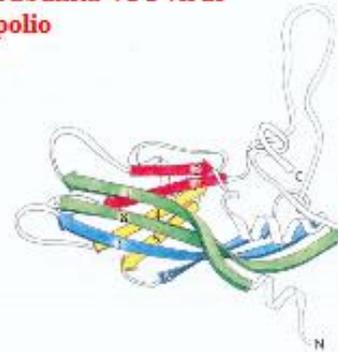
## Virus con capside sferico

In generale si osserva che i loop ad una estremità del barile sono sempre corti, mentre quelli all'estremità opposta sono molto lunghi e di dimensioni variabili. Tali strutture proteiche si possono schematizzare con un cuneo.

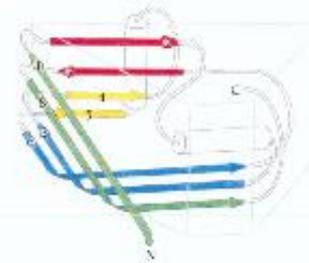
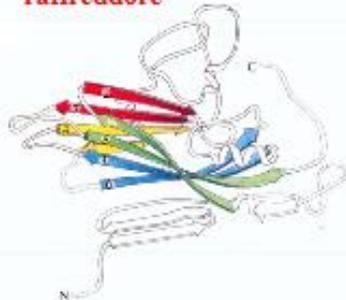
**Virus satellite della necrosi del tabacco**



**Subunità VP1 virus polio**



**Subunità VP2 virus raffreddore**



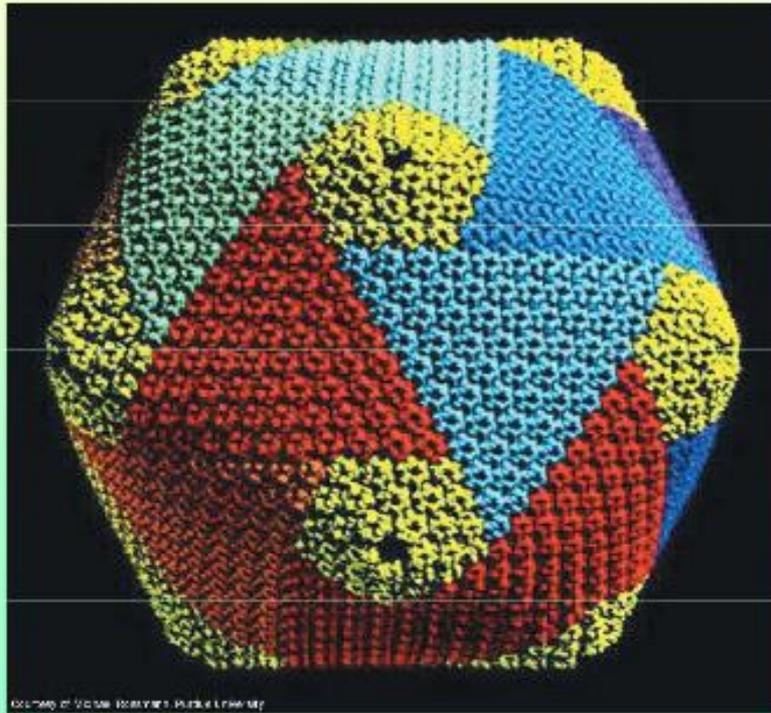
## Virus con capside sferico: $T = 4$

Nel caso di  $T = 4$  (240 subunità), ciascuna unità asimmetrica contiene 4 subunità proteiche (A, B, C e D) con un diverso environment.

Le subunità A interagiscono intorno all'asse di rotazione di ordine 5, le subunità D intorno all'asse di rotazione di ordine 3, mentre le subunità B e C sono disposte in modo tale che 2 copie di ciascuna interagiscono intorno all'asse di ordine 2, oltre a 2 subunità D.



Structure of the PBCV-1 capsid. A quasi-atomic model based on fitting the X-ray structure of Vp54 to the cryo-EM-based image of the capsid.



# **Struttura tridimensionale delle macromolecole biologiche: Risorse Bioinformatiche**

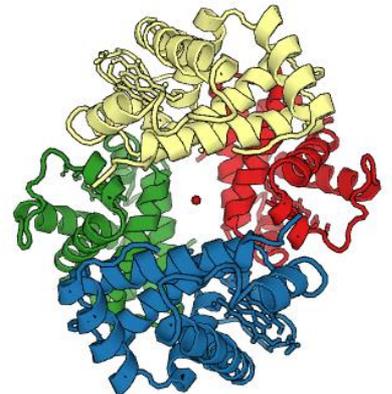


Protein Data Bank:

<http://www.rcsb.org/pdb/home/home.do>

Since 1971, the Protein Data Bank archive (PDB) has served as the single repository of information about the 3D structures of proteins, nucleic acids, and complex assemblies.

The Worldwide PDB (wwPDB) organization manages the PDB archive and ensures that the PDB is freely and publicly available to the global communi



## Banche dati di Proteine: il Protein Data Bank

- PDB E' la principale risorsa per le **strutture di macromolecole**: Proteine, Supercomplessi, Acidi nucleici
- Le strutture sono determinate mediante cristallografia e NMR
- Oltre ai file di struttura (.PDB) sono presenti informazioni sulle sequenze e molti strumenti per:

Analisi di struttura, Visualizzazione di ligandi,  
Determinazione delle similitudini

# Accesso a PDB

- RCSB PDB ([www.rcsb.org/pdb/home/home.do](http://www.rcsb.org/pdb/home/home.do))
- PDBe (<http://www.ebi.ac.uk/pdb/e/>)
- PDBj ([www.pdbj.org/](http://www.pdbj.org/))
- BMRB (Biological Magnetic Resonance Data Bank, [www.bmrw.wisc.edu](http://www.bmrw.wisc.edu))

4 siti membri: i dati sono gli stessi, cambia l'interfaccia e i tool



# Esempio di ricerca: 1A3I

Descrizione

Citazione  
bibliografica di  
riferimento

Descrizione  
molecolare

The screenshot displays the RCSB PDB entry page for 1A3I. The main title is "X-RAY CRYSTALLOGRAPHIC DETERMINATION OF A COLLAGEN-LIKE PEPTIDE WITH THE REPEATING SEQUENCE (PRO-PRO-GLY)". The ID "1A3I" is prominently displayed in a red circle. The primary citation is "Kramer, R.Z., Vitagliano, L.P., Bella, J.P., Boreio, R.P., Mazzarella, L.P., Brodsky, B.P., Zagan, A.P., Berman, H.M.P. (1998) J Mol Biol. 280: 623-638". The molecular description table is as follows:

Molecular Description			
Classification:	Extracellular Matrix		
Structure Weight:	1933.14		
Molecule:	COLLAGEN-LIKE PEPTIDE		
Polymer: 1	Type: protein	Length:	9
Chains:	A		
Molecule:	COLLAGEN-LIKE PEPTIDE		
Polymer: 2	Type: protein	Length:	6
Chains:	B, C		

The 3D structure is shown as a ribbon model in the "Biological Assembly" section.

ID della  
proteina

Struttura 3D

# Il file .PDB: formato testuale per la descrizione di strutture 3D di macromolecole biologiche

Contiene la descrizione e l'annotazione di strutture di proteine e acidi nucleici tra cui: ***coordinate atomiche, rotameri di catene laterali*** osservati, ***assegnazione a particolari strutture secondarie***, e ***connettività atomica***. Altre molecole come acqua, ioni, acidi nucleici, ligandi e così via possono essere descritti nel formato pdb.

# Il formato PDB

	Classificazione	Data sottomissione	PDB ID
	EXTRACELLULAR MATRIX	22-JAN-98	1A3I
TITOLO	X-RAY CRYSTALLOGRAPHIC DETERMINATION OF A COLLAGEN-LIKE		
TITOLO	2 PEPTIDE WITH THE REPEATING SEQUENCE (PRO-PRO-GLY)		
...			
Tipo di dato sperimentale	EXPDTA X-RAY DIFFRACTION		
Autori	AUTHOR R.Z.KRAMER,L.VITAGLIANO,J.BE		
Autori	AUTHOR 2 B.BRODSKY,A.ZAGARI,H.M.BERM		
...			
Commenti più dettagliati	REMARK 350 BIOMOLECULE: 1		
Commenti più dettagliati	REMARK 350 APPLY THE FOLLOWING TO CHAI		
Commenti più dettagliati	REMARK 350 BIOMT1	1	1.000000 0.0
Commenti più dettagliati	REMARK 350 BIOMT2	1	0.000000 1.000000 0.000000 0.000000
...			
Sequenza delle catene peptidiche	SEQRES	1 A	9 PRO PRO GLY PRO PRO GLY PRO PRO GLY
Sequenza delle catene peptidiche	SEQRES	1 B	6 PRO PRO GLY PRO PRO GLY
Sequenza delle catene peptidiche	SEQRES	1 C	6 PRO PRO GLY PRO PRO GLY
...			
Coordinate degli atomi che fanno parte della proteina	ATOM	1 N	PRO A 1 8.316 21.206 21.530 1.00 17.44 N
Coordinate degli atomi che fanno parte della proteina	ATOM	2 CA	PRO A 1 7.608 20.729 20.336 1.00 17.44 C
Coordinate degli atomi che fanno parte della proteina	ATOM	3 C	PRO A 1 8.487 20.707 19.092 1.00 17.44 C
Coordinate degli atomi che fanno parte della proteina	ATOM	4 O	PRO A 1 9.466 21.457 19.005 1.00 17.44 O
Coordinate degli atomi che fanno parte della proteina	ATOM	5 CB	PRO A 1 6.460 21.723 20.211 1.00 22.26 C
...			
HETATM: atomi che non fanno parte della proteina (piccole molecole o ioni)	HETATM	130 C	ACY 401 3.682 22.541 11.236 1.00 21.19 C
HETATM: atomi che non fanno parte della proteina (piccole molecole o ioni)	HETATM	131 O	ACY 401 2.807 23.097 10.553 1.00 21.19 O
HETATM: atomi che non fanno parte della proteina (piccole molecole o ioni)	HETATM	132 OXT	ACY 401 4.306 23.101 12.291 1.00 21.19 O
...			

Descrive come calcolare le coordinate del multimerico a partire da quelle esplicitate per la singola unità

# Il formato PDB: ATOM

	Numero dell'atomo	Identità dell'atomo		Catena	Coordinate			Temperature factor		
					X	Y	Z			
ATOM	2	CA	GLY	A	672	54.168	8.340	69.707	1.00104.94	C
ATOM	3	C	GLY	A	672	52.692	8.194	69.380	1.00105.46	C
ATOM	4	O	GLY	A	672	51.877	9.045	69.750	1.00108.67	O
ATOM	5	N	GLU	A	673	52.359	7.101	68.691	1.00102.41	N
ATOM	6	CA	GLU	A	673	50.994	6.785	68.274	1.00 89.17	C
ATOM	7	C	GLU	A	673	50.624	5.325	68.585	1.00 81.77	C
ATOM	8	O	GLU	A	673	51.438	4.411	68.405	1.00 81.88	O
ATOM	9	CB	GLU	A	673	50.850	7.050	66.777	1.00 96.53	C
ATOM	10	CG	GLU	A	673	50.252	8.399	66.438	1.00 99.19	C
ATOM	11	CD	GLU	A	673	48.788	8.486	66.827	1.00115.45	C
ATOM	12	OE1	GLU	A	673	48.062	7.477	66.681	1.00116.71	O
ATOM	13	OE2	GLU	A	673	48.356	9.561	67.286	1.00113.58	O
ATOM	14	N	ALA	A	674	49.387	5.109	69.023	1.00 67.27	N
ATOM	15	CA	ALA	A	674	48.912	3.768	69.370	1.00 63.11	C
ATOM	16	C	ALA	A	674	48.702	2.826	68.174	1.00 58.54	C
ATOM	17	O	ALA	A	674	48.064	3.183	67.186	1.00 62.02	O
ATOM	18	CB	ALA	A	674	47.616	3.866	70.189	1.00 47.04	C
ATOM	19	N	PRO	A	675	49.260	1.612	68.240	1.00 55.66	N
ATOM	20	CA	PRO	A	675	49.087	0.665	67.134	1.00 52.95	C
ATOM	21	C	PRO	A	675	47.629	0.261	66.997	1.00 48.19	C

- Occupancy**: indica la frazione di molecole per cui l'atomo si trova in questa posizione. Di solito = 1, sarà inferiore ad 1 in caso di conformazioni multiple
- Temperature factor**: indica la quantità di "sfumatura" della densità elettronica dovuta alla vibrazione degli atomi o a differenze tra le molecole nel cristallo.

## Banche dati – Proteine: Pfam e Prosite

- Utili per studiare e catalogare le strutture proteiche.
- Le *similitudini* (domini, fold, ponti disolfuro, ecc) possono essere usate *per inferire la funzione*.
- Utile tra proteine simili e omologhe in specie diverse



- Suddivide le proteine e ne descrive le caratteristiche in *famiglie* in base a metodi statistici:
  - Allineamenti
  - HMM
- Individua, *data una sequenza*, le *possibili famiglie* di appartenenza.
- Determina possibili *caratteristiche funzionali; domini; cofattori; siti attivi; aminoacidi strutturalmente importanti; livello di conservazione*

Developers: Join the RCSB PDB Team

[Explore Open Positions](#)

Welcome

Deposit

Search

Visualize

Analyze

Download

Learn

## A Structural View of Biology

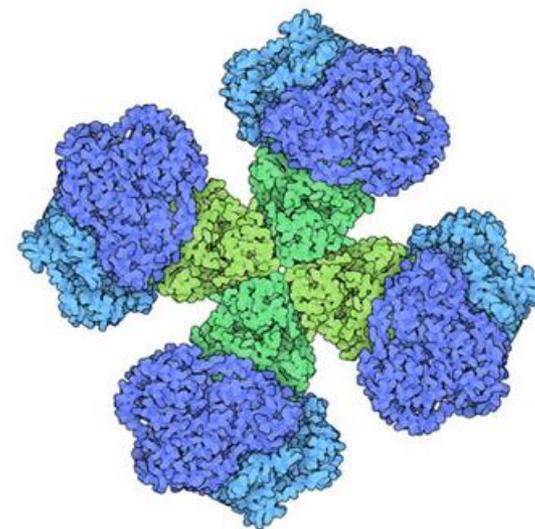
This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

As a member of the wwPDB, the RCSB PDB curates and annotates PDB data.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.



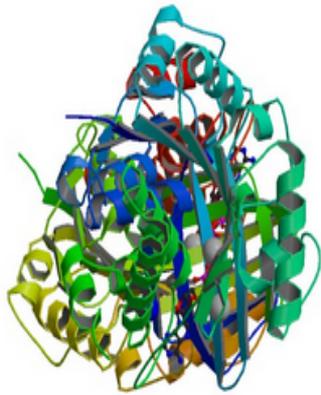
## November Molecule of the Month



Acetohydroxyacid Synthase

## Latest Entries

As of [Tuesday Mar 12 2019](#)



6ILT

[PDB Entry](#)

Structure of Arabidopsis thaliana Ribokinase complexed with ATP and Magnesium ion



## Features & Highlights



Mandatory PDBx/mmCIF format files submission for MX depositions

Submission of PDBx/mmCIF format files for crystallographic depositions to the PDB will be mandatory from July 1<sup>st</sup> 2019 onward. PDB format files will no longer be accepted for deposition of structures solved by MX techniques.



Join Our Team as a Biocurator

Curate, validate, and standardize macromolecular structures from the PDB community at Rutgers, The State University of New Jersey.

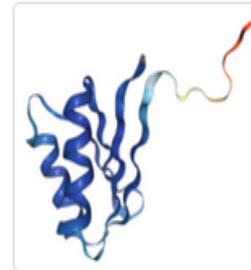


New Architecture and Services Enable Faster Access to More Information

Explore the improved display of PDB Statistics, structure funding information, and 3D views of ligands and electron density.

## News

[Publications](#) ▾



Visualize Structure Quality Metrics in 3D

Use NGL to map wwPDB Validation Report information onto 3D structure, including Random Coil Index data.

» [03/12/2019](#)

[New Video: Penicillin and Antibiotic Resistance](#)

» [03/05/2019](#)

[Superbugs! How Bacteria Evolve Resistance to Antibiotics](#) » [02/26/2019](#)

[Mandatory PDBx/mmCIF format files submission for MX depositions](#) » [02/19/2019](#)

[Join Our Team as a Biocurator](#) » [02/12/2019](#)

[New Online Curriculum: The PDB Pipeline & Data Archiving](#) » [02/05/2019](#)

**PDB at a Glance**

**46473** Distinct Protein Sequences

**41876** Structures of Human Sequences

**10840** Nucleic Acid Containing Structures

[More Statistics](#)

Search by PDB ID, author, macromolecule, sequence, or ligands

Go

Advanced Search | Browse by Annotations

Welcome

Deposit

Search

Visualize

Analyze

Download

Learn

Search Options

Drill Down by Categories

Advanced Search

Sequences

Ligands

Drugs & Drug Targets

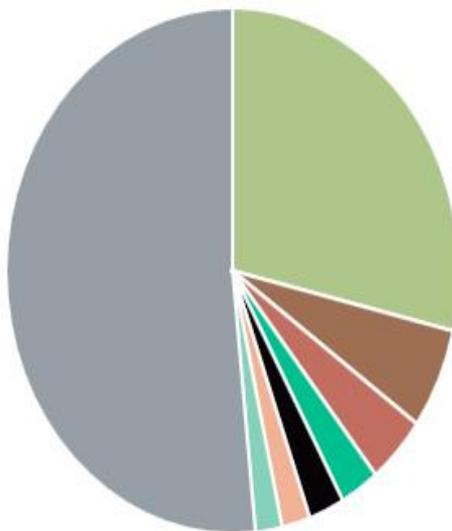
Unreleased & New Entries

Browse by Annotation

PDB Statistics

Explore the PDB Archive

Organism



- Homo sapiens (43274)
- Escherichia coli (9378)
- Mus musculus (6394)
- Saccharomyces cerevisiae (4203)
- synthetic construct (3851)
- Rattus norvegicus (3043)
- Bos taurus (2874)
- Other (77867)

Search by PDB ID, author, macromolecule, sequence, or ligands

Go

[Advanced Search](#) | [Browse by Annotations](#)

Welcome

Deposit

Search

Visualize

Analyze

Download

Learn

Sequence & Structure Alignment

Protein Symmetry

Structure Quality

Map Genomic Position to Protein

EPPIC Biological Assemblies

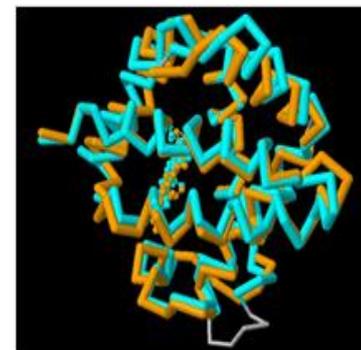
PDB Statistics

Third Party Tools

## Sequence & Structure Alignment

RCSB PDB's Comparison Tool calculates pairwise sequence (blast2seq, Needleman-Wunsch, and Smith-Waterman) and structure alignments (FATCAT, CE, Mammoth, TM-Align, TopMatch).

Comparisons can be made for any protein in the PDB archive and for [customized](#) or [local files](#) not in the PDB. Special features include support for both rigid-body and flexible alignments and detection of circular permutations.



Enter First PDB ID

Enter Second PDB ID



Select Associated Chain ID

Select Associated Chain ID

...

...

- Select Comparison Method -

Align

More Options

43,274 Structures

17533 Citations

14020 Ligands

## Search Parameter:

Refine Search

Save Search to MyPDB

TAXONOMY is just Homo sapiens (human)

## Refinements



Currently showing 1 - 25 of 43274 Page: 1 of 1731

[← Previous](#)[Next →](#)

Displaying 25 Results

## View:

Detailed

## Reports:

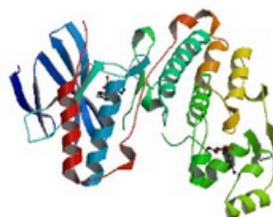
Select a Report

## Sort:

↓ Release Date: Newest to Oldest

[Download Files](#)

## 50MG

[Download File](#)[View File](#)[3D View](#)

## p38alpha in complex with pyrazolobenzothiazine inhibitor COXP4M12

[Bartolini, D.](#), [Buhmann, M.](#), [Barreca, M.L.](#), [Manfroni, G.](#), [Cecchetti, V.](#), [Rauh, D.](#), [Galli, F.](#)

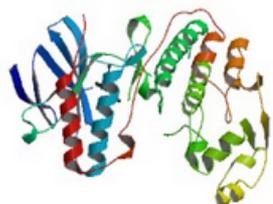
(2019) Biochem Biophys Res Commun

**Released:** 3/13/2019**Method:** X-ray Diffraction**Resolution:** 2.0 Å**Residue Count:** 360**Macromolecule:**

Mitogen-activated protein kinase 14 (protein)

**Unique Ligands:** 9Y8, BOG

## 50MH

[Download File](#)[View File](#)[3D View](#)

## p38alpha in complex with pyrazolobenzothiazine inhibitor COXH11

[Bartolini, D.](#), [Buhmann, M.](#), [Barreca, M.L.](#), [Manfroni, G.](#), [Cecchetti, V.](#), [Rauh, D.](#), [Galli, F.](#)

(2019) Biochem Biophys Res Commun

**Released:** 3/13/2019**Method:** X-ray Diffraction**Resolution:** 2.5 Å**Residue Count:** 360**Macromolecule:**

Mitogen-activated protein kinase 14 (protein)

**Unique Ligands:** 9Y5

## ORGANISM

Homo sapiens only (37491)

Homo sapiens/synthetic co ... (990)

Homo sapiens/Mus musculus (880)

Homo sapiens/Human immuno ... (393)

Homo sapiens/Hirudo medic ... (243)

Homo sapiens/Escherichia ... (228)

Homo sapiens/Rattus norve ... (209)

Other (2840)

## UNIPROT MOLECULE NAME

Beta-2-microglobulin (798)

Carbonic anhydrase 2 (742)

Prothrombin (399)

Cyclin-dependent kinase 2 (391)

Beta-secretase 1 (374)

DNA polymerase beta (345)

DNA cross-link repair 1A ... (312)

Refine Query

## TAXONOMY

Eukaryota only (40132)

Eukaryota/Viruses (1220)

Eukaryota/Other (1175)

Eukaryota/Bacteria (650)

Eukaryota/Archaea (38)