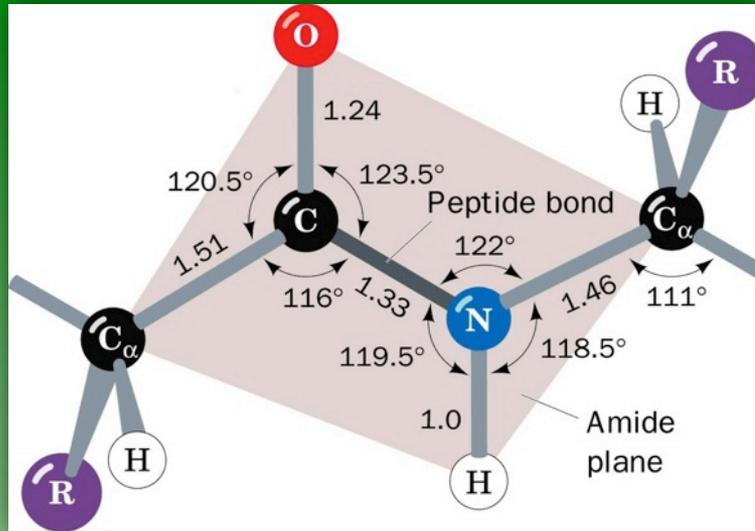


Protein Secondary, Tertiary & Quaternary Structure

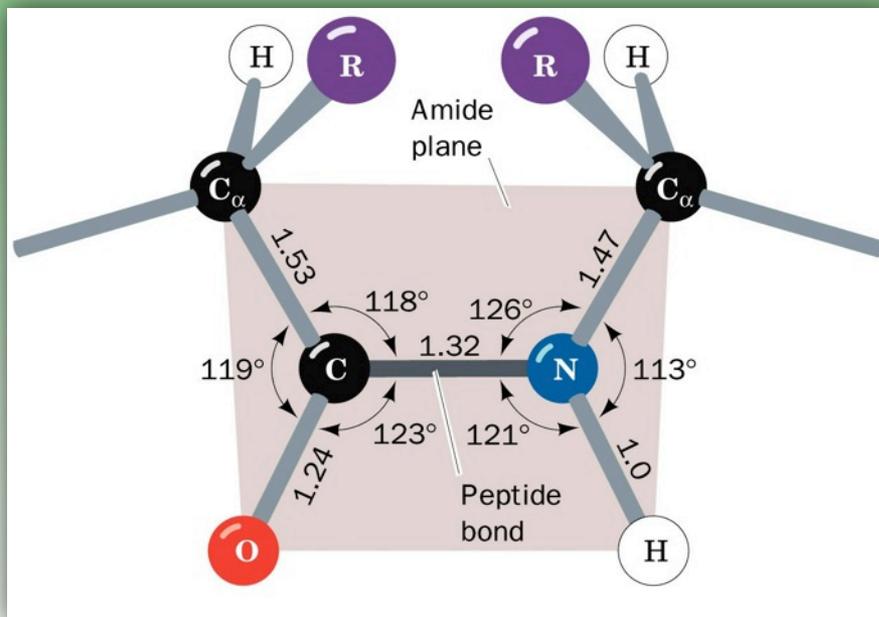
CHEM 420 – Principles of Biochemistry
Instructor – Anthony S. Serianni

Chapter 8: Voet/Voet, *Biochemistry*, 2011
Fall 2015

September 16 & 18



trans peptide
(amide) configuration



cis peptide (amide)
configuration: less stable
for most α -amino acids

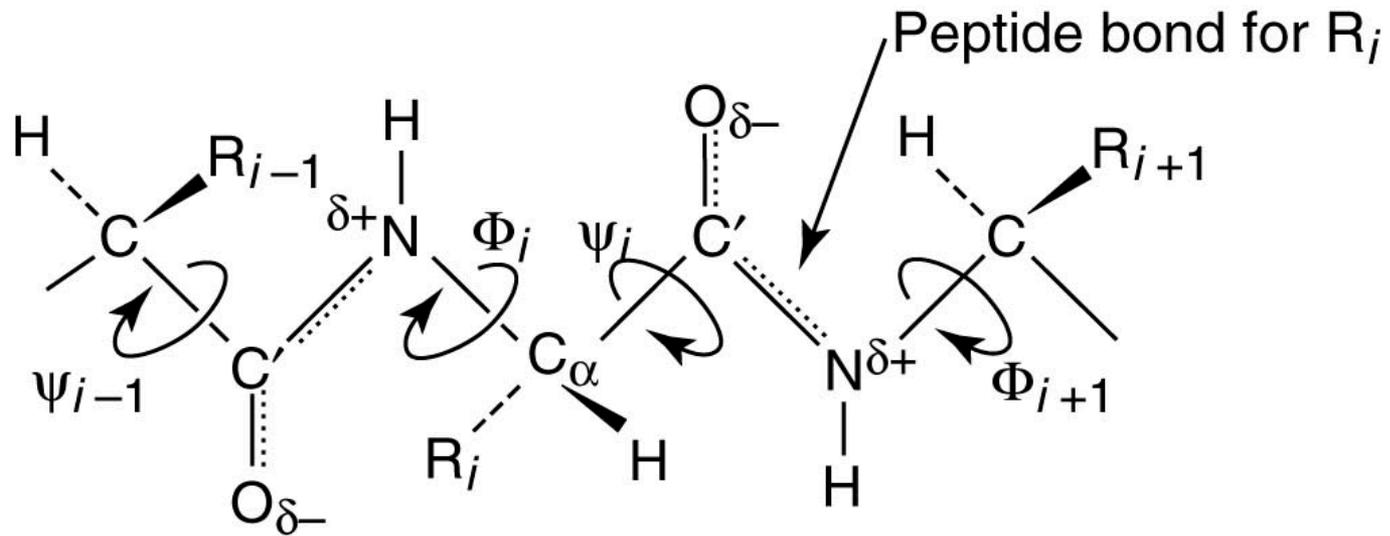
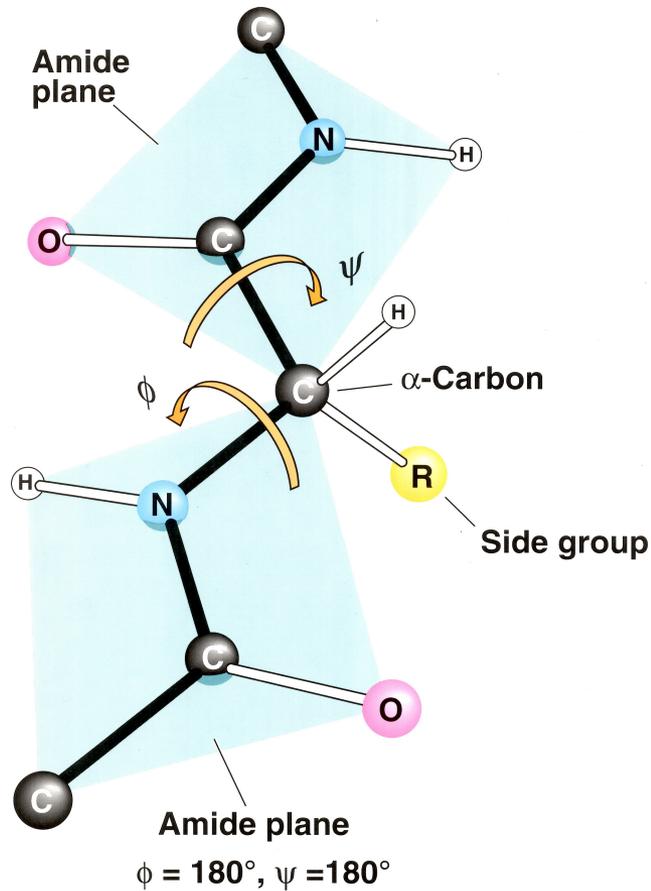


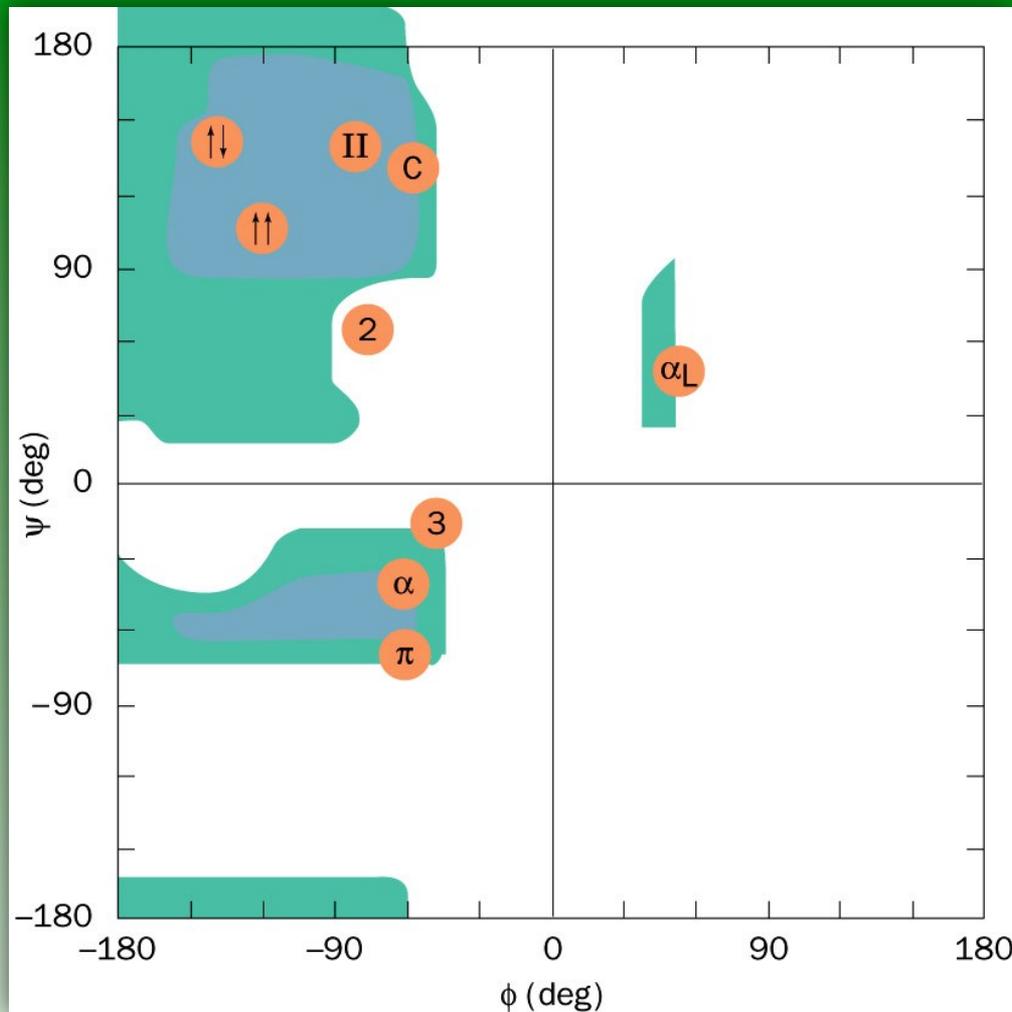
Figure 3.11. Amino acid residue within a polypeptide chain.

Textbook of Biochemistry With Clinical Correlations, Sixth Edition, Edited by Thomas M. Devlin. Copyright © 2006 John Wiley & Sons, Inc.

Three distinct bond types along the backbone of a protein:
 the “rigid” peptide bond and the rotatable
 phi (ϕ) and psi (ψ) bonds involving C_α.

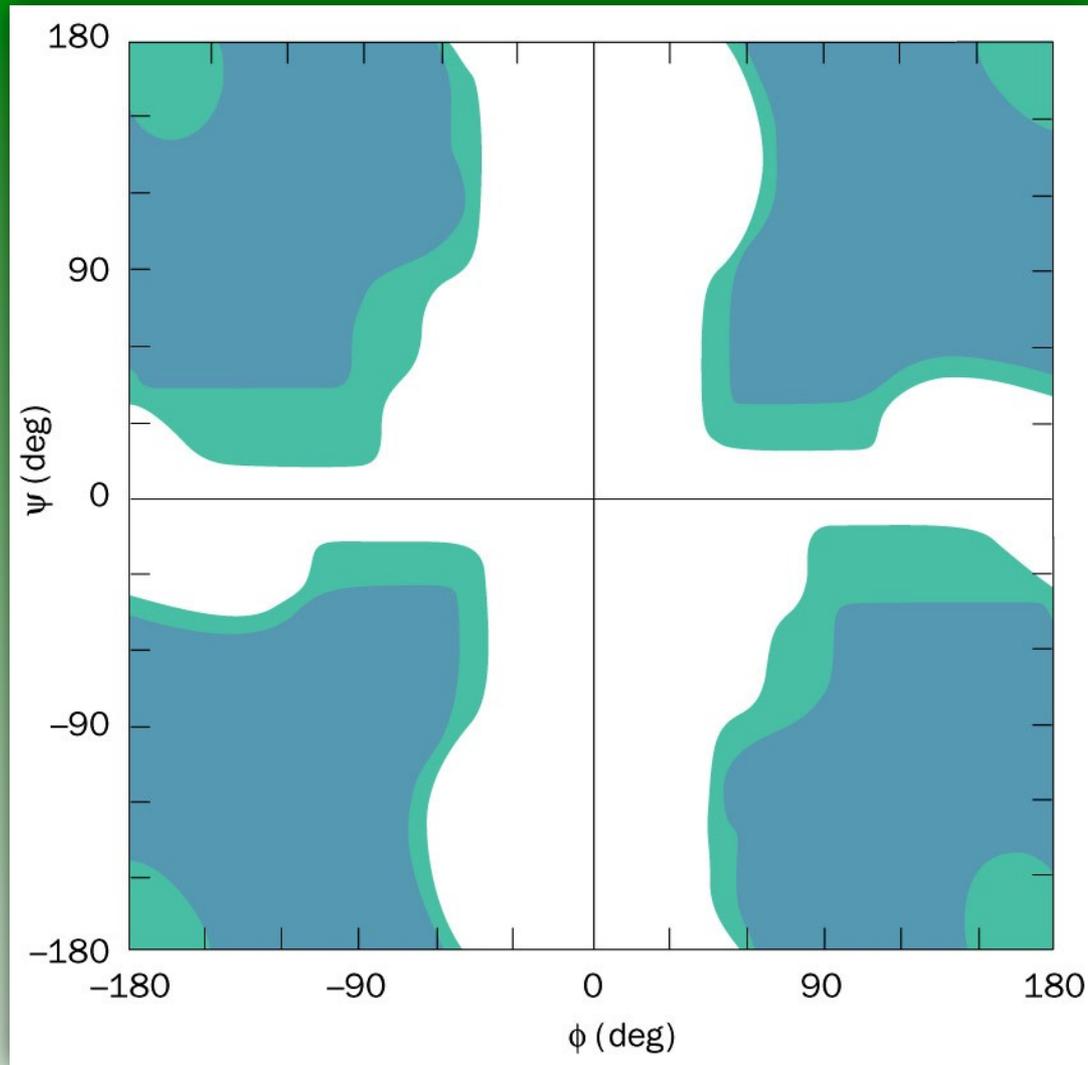


Definitions of
 ϕ and ψ along the
 backbone of a protein



Predicting
protein
secondary
structure

Ramachandran ϕ/ψ plot for proteins **calculated** from analyses of van der Waals radii in proteins



Ramachandran diagram of Gly residues in a polypeptide chain; normally allowed areas are shown in blue.

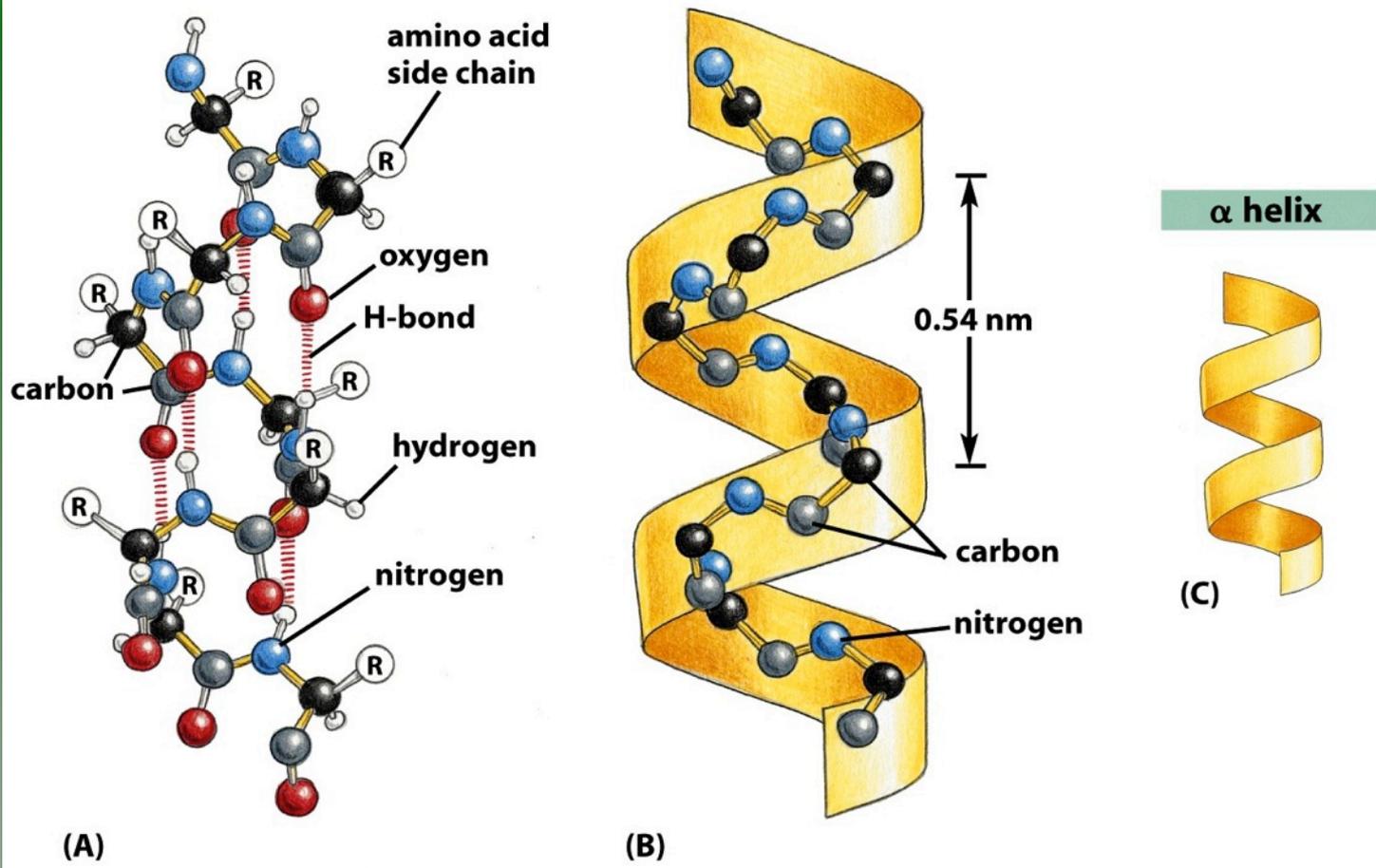


Figure 3-7a-c Molecular Biology of the Cell 5/e (© Garland Science 2008)

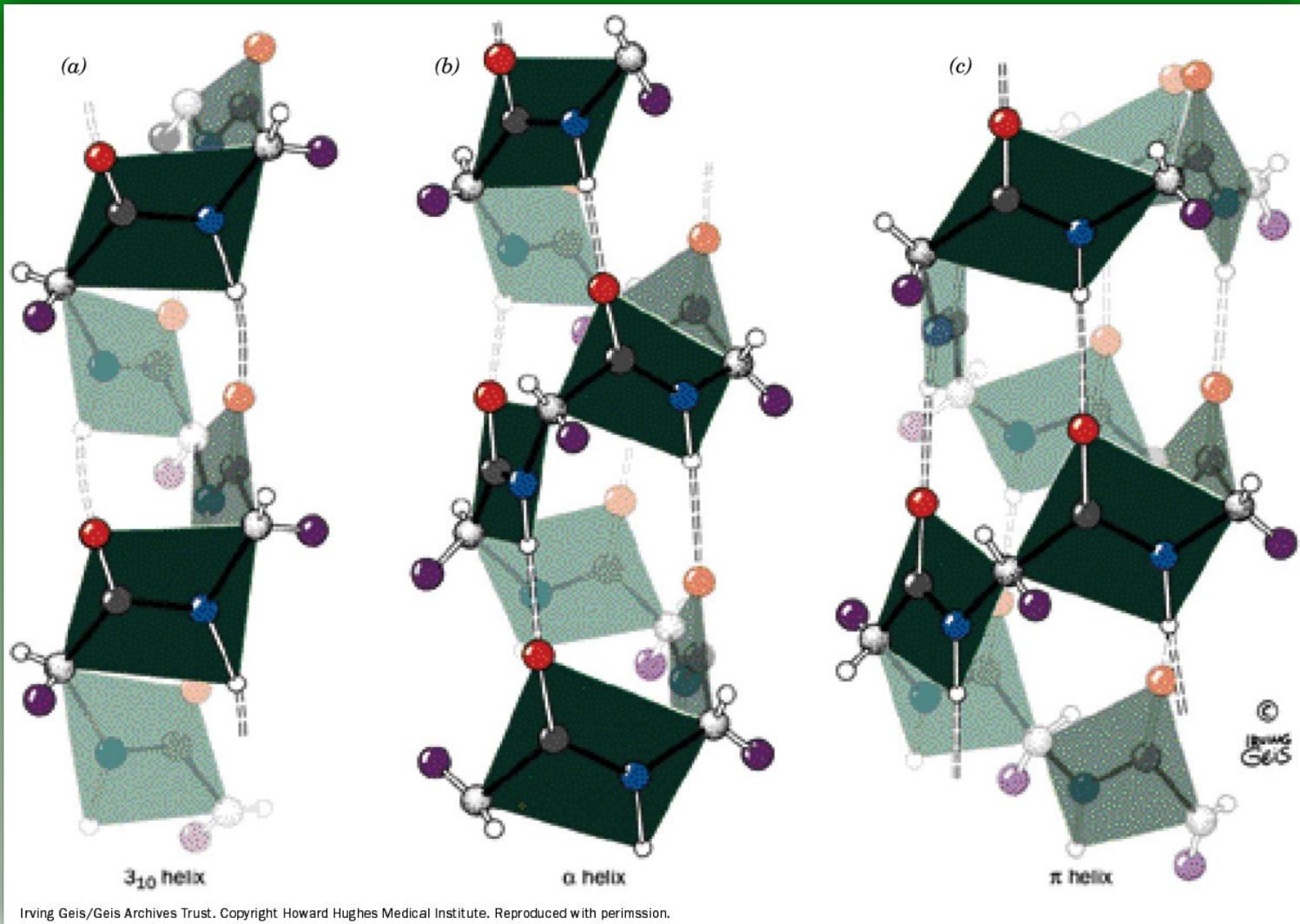
Different representations of the α -helix (3.6₁₃)
secondary structure

Molecular parameters associated with the major protein secondary structures

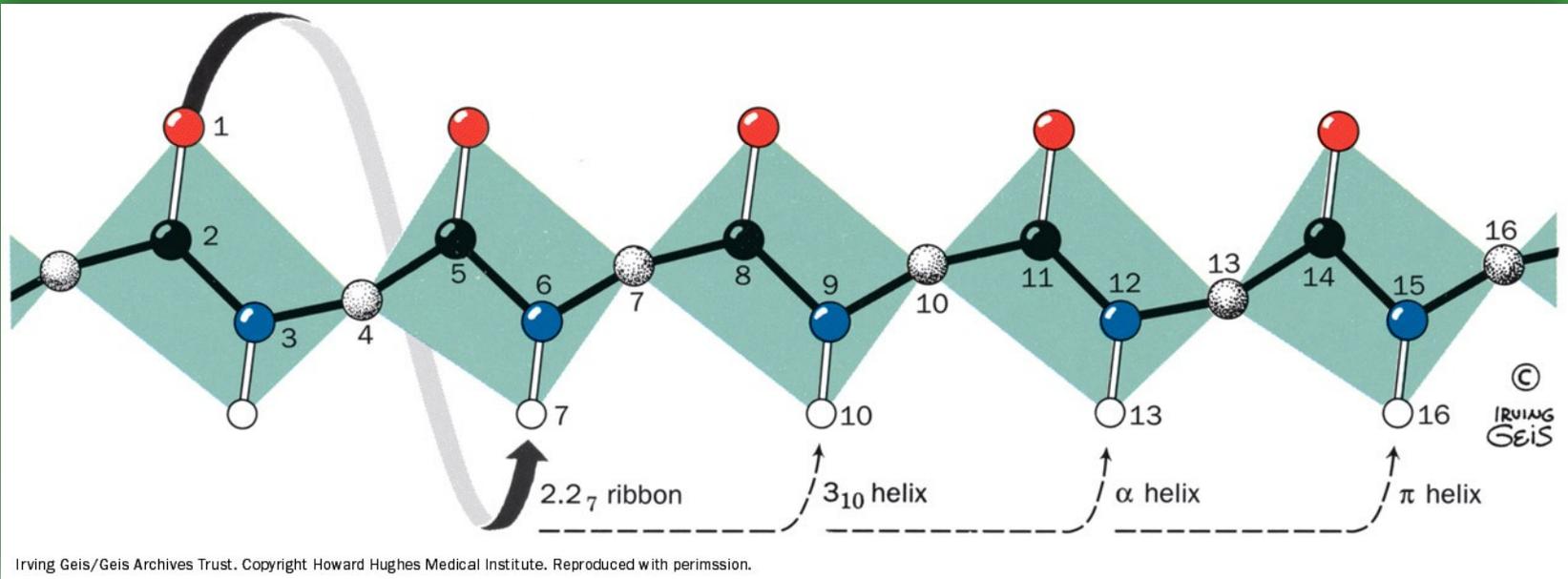
Structure Type	Residues/ Turn	Rise (nm)	Number of Atoms in H-Bonded Ring	ϕ (°)	ψ (°)
Antiparallel β sheet	2.0	0.34	— ^a	-139	+135
Parallel β sheet	2.0	0.32	— ^a	-119	+113
3_{10} helix	3.0	0.20	10	-49	-26
α helix (3.6 ₁₃)	3.6	0.15	13	-57	-47
π helix (4.4 ₁₆) ^b	4.4	0.12	16	-57	-70

^aBonding is between polypeptide chains.

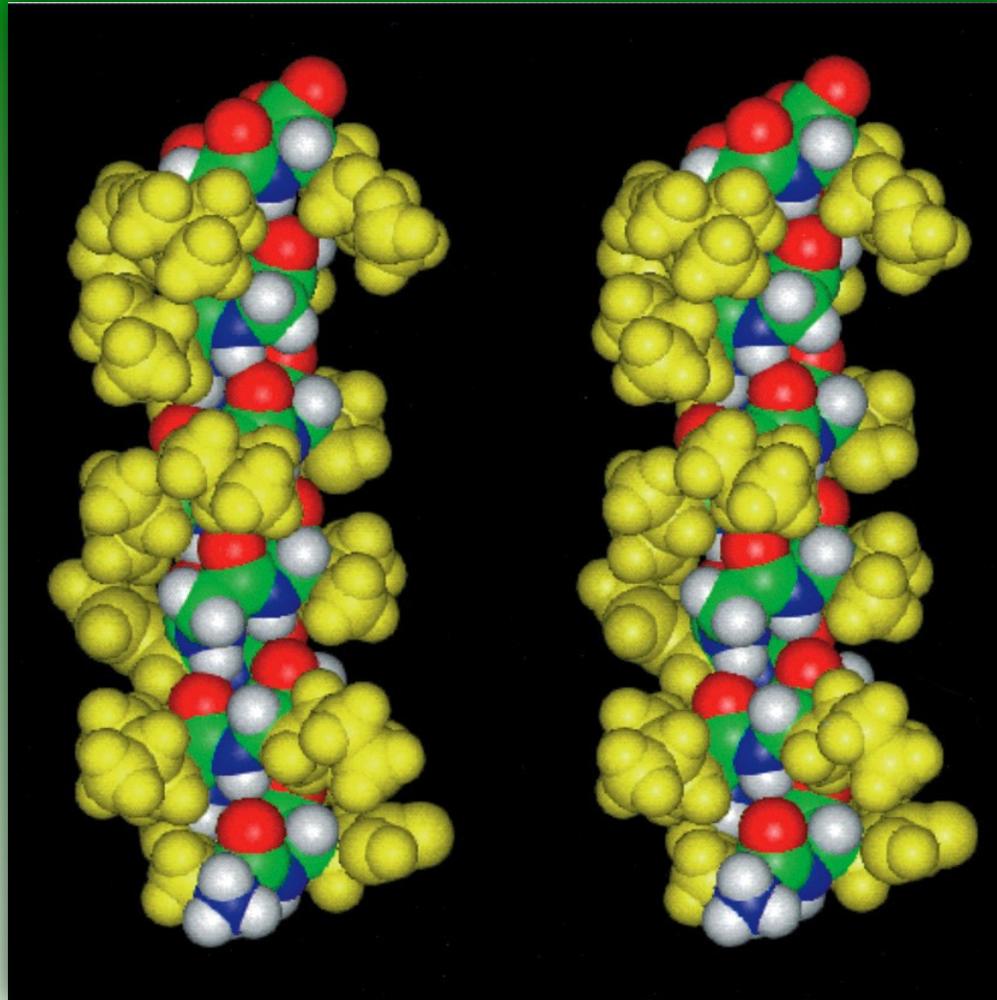
^bSterically permitted but not observed in protein.



Structural comparison of 3_{10} , 3.6_{13} (α) and 4.4_{16} (π) helices

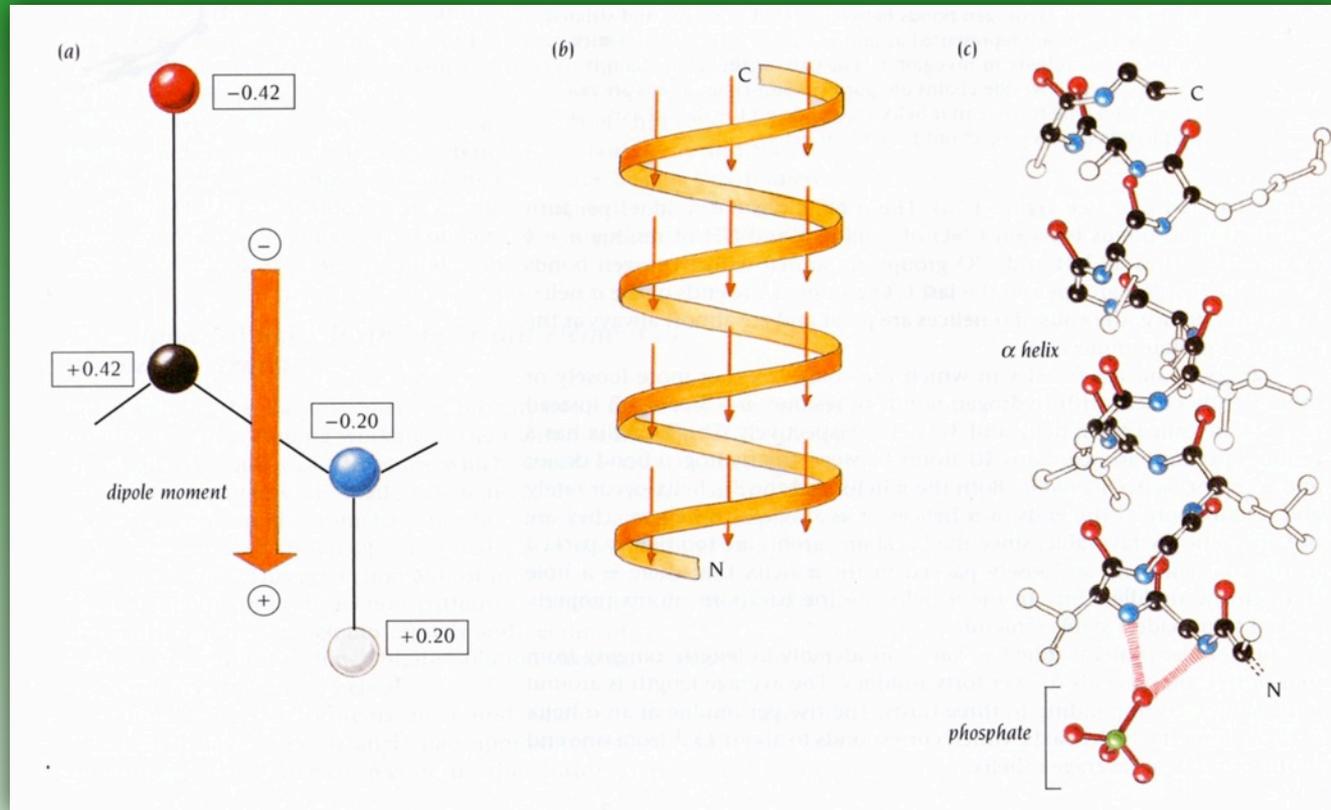


The hydrogen bonding patterns of several polypeptide helices

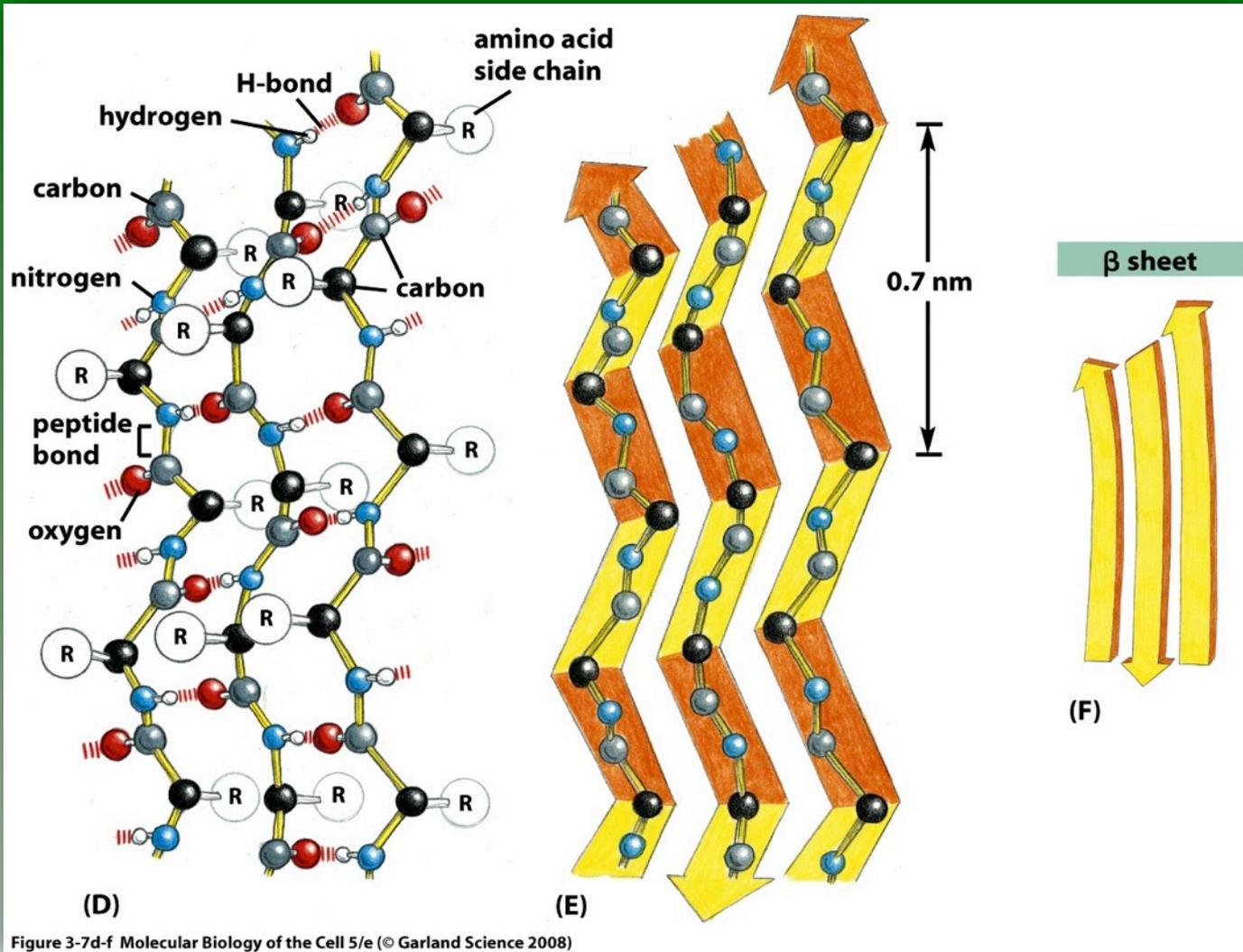


Stereo space-filling representation of an α -helical segment of sperm whale myoglobin (E-helix) determined by X-ray single crystal structure analysis: R-groups shown in yellow

The α -helix has a dipole moment

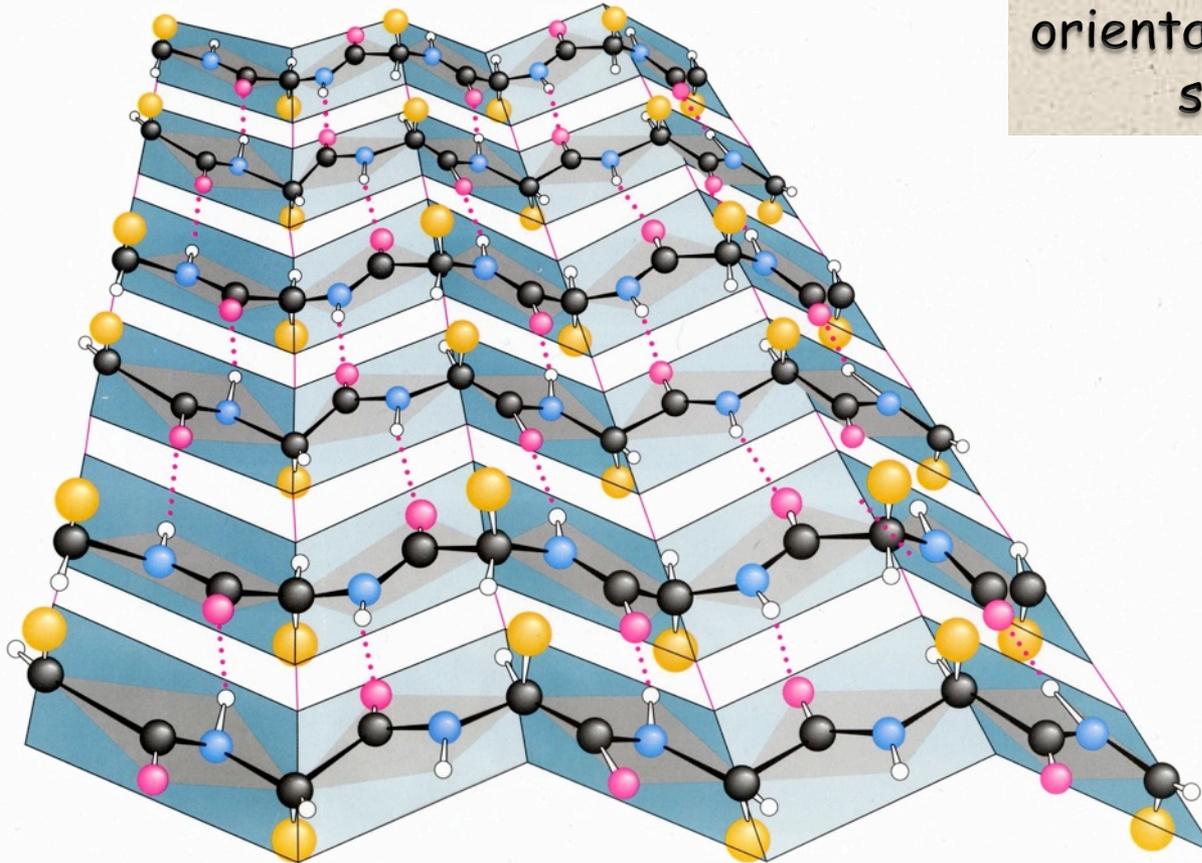


(a) The dipole of a peptide bond showing approximate fractional charges. **(b)** Individual peptide dipoles of the helix are aligned parallel to the helix axis, creating an overall dipole moment, positive at the amino end and negative at the carboxyl end. **(c)** Phosphate H-bonded to the NH end of the helix - binding is facilitated by the helix dipole.

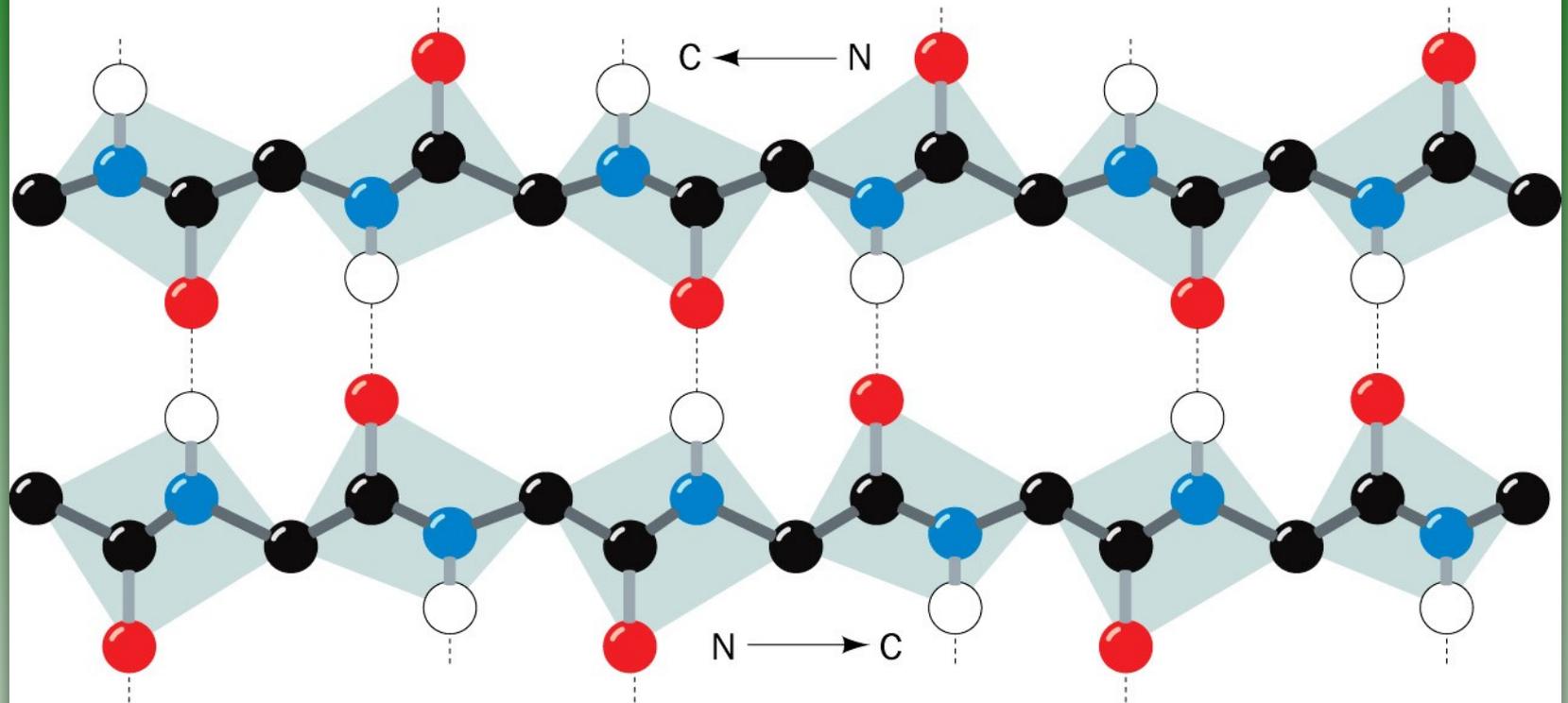


Different representations of the β (pleated) sheet secondary structure

Another view of the side-by-side arrangement of β -sheet secondary structure (antiparallel). Note the up/down orientation of the R-groups shown in yellow.

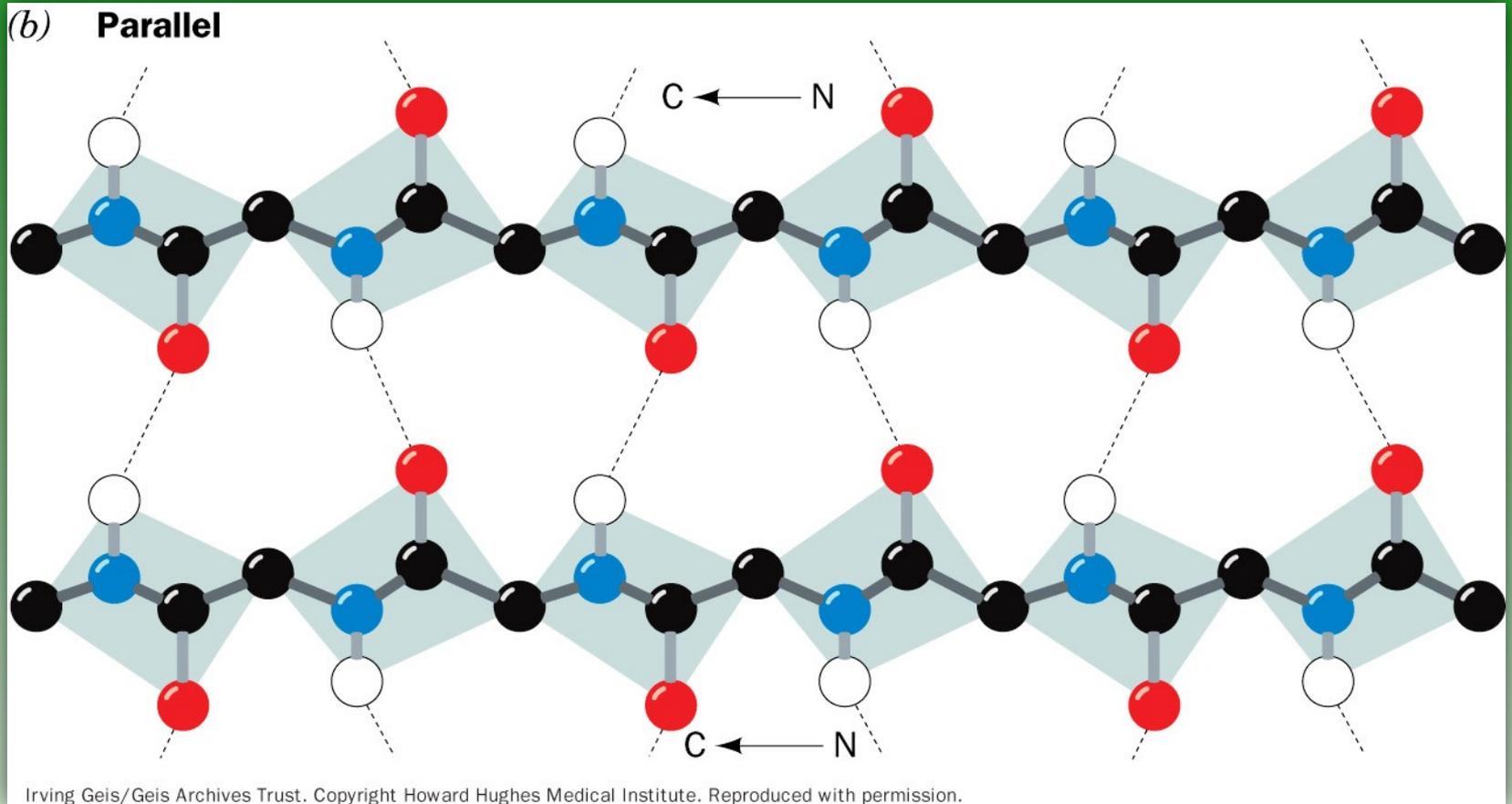


(a) **Antiparallel**

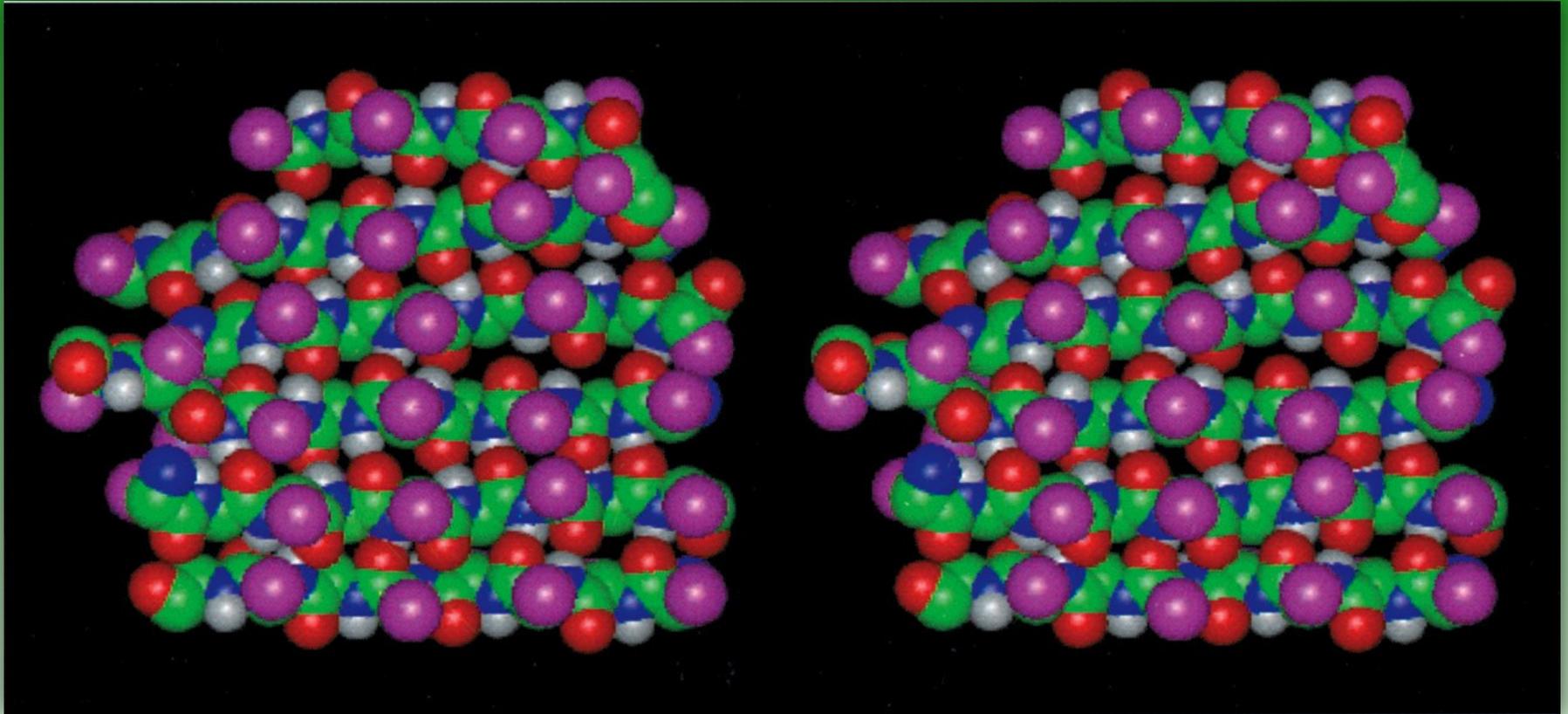


Irving Geis/Geis Archives Trust. Copyright Howard Hughes Medical Institute. Reproduced with permission.

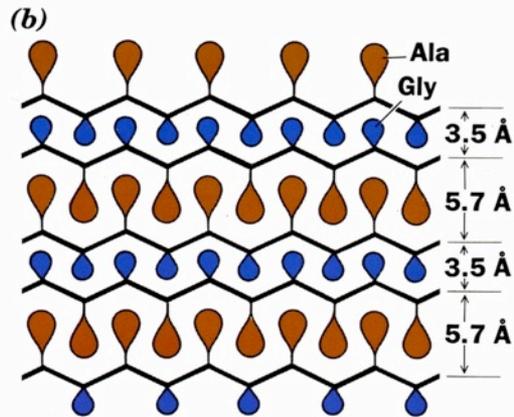
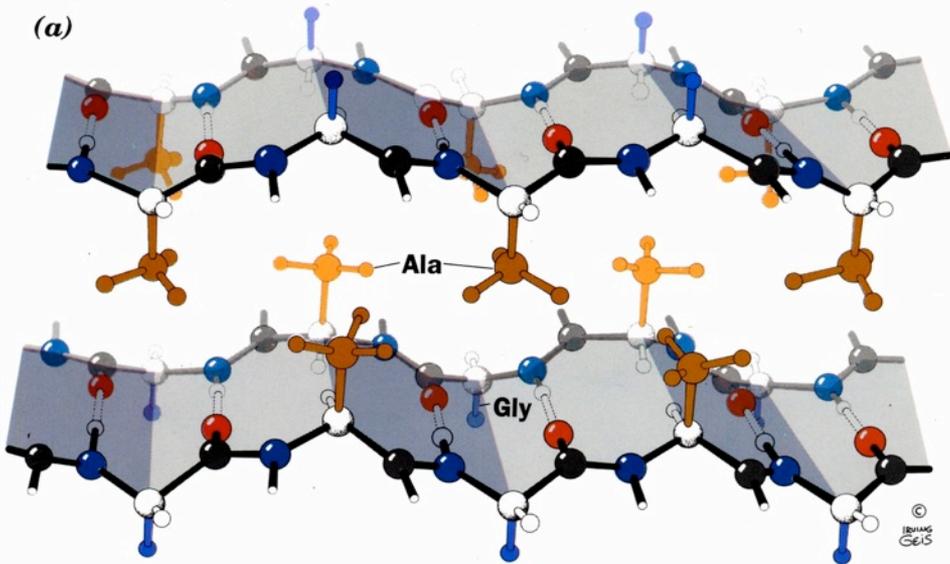
β pleated sheet: antiparallel orientation



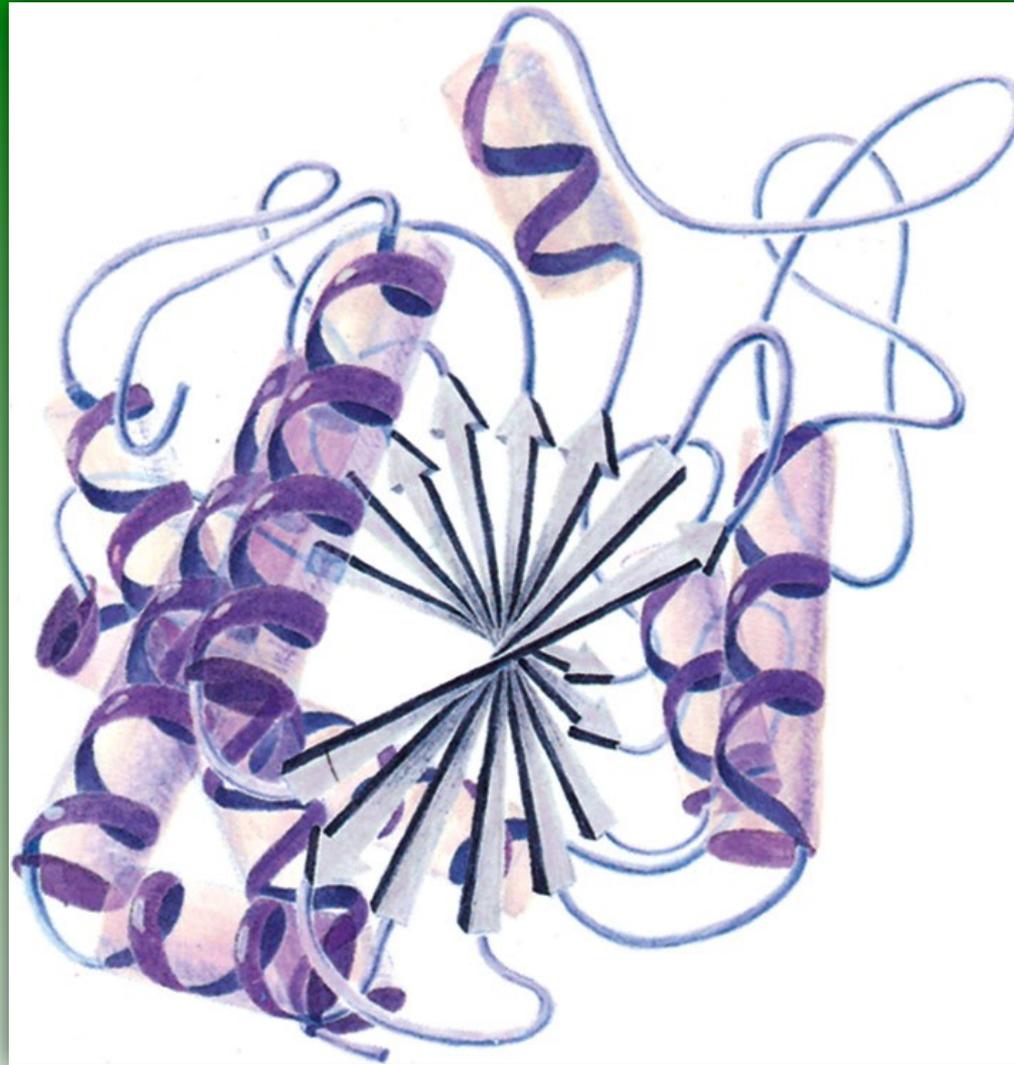
β pleated sheet: parallel orientation



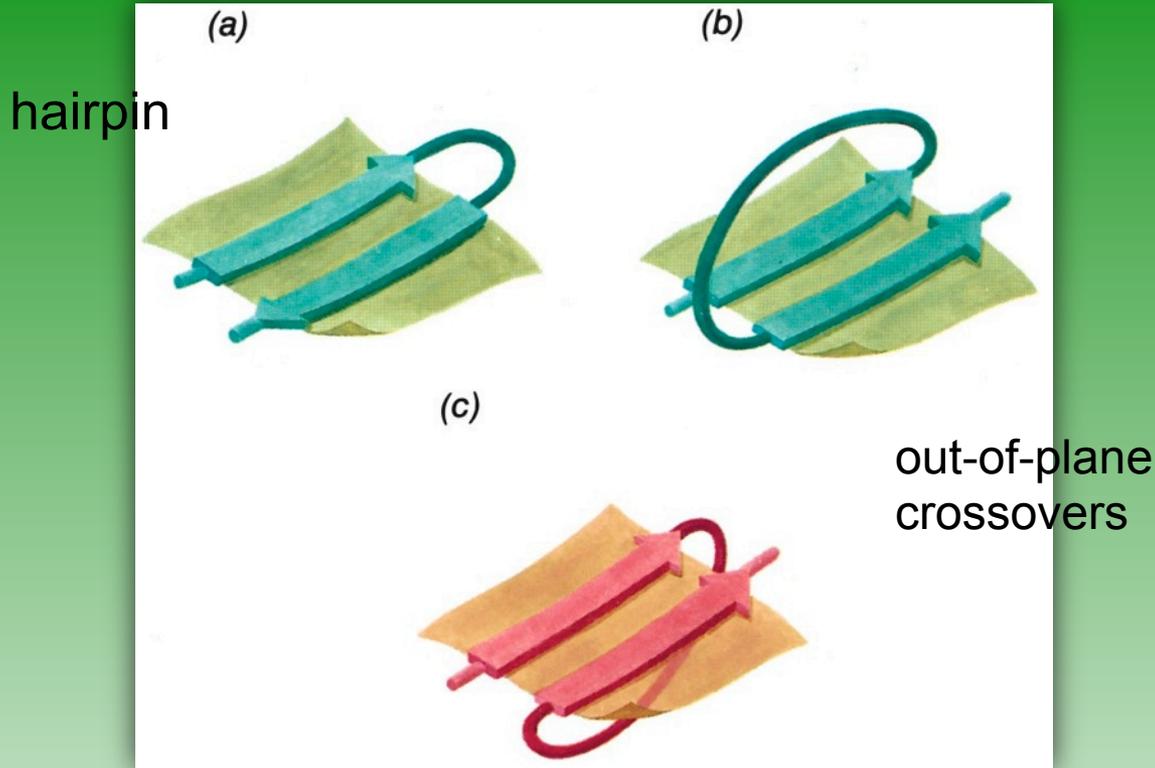
Stereo space-filling representation of the 6-stranded antiparallel β pleated sheet in jack bean concanavalin A as determined by crystal X-ray analysis; β structure in a globular protein



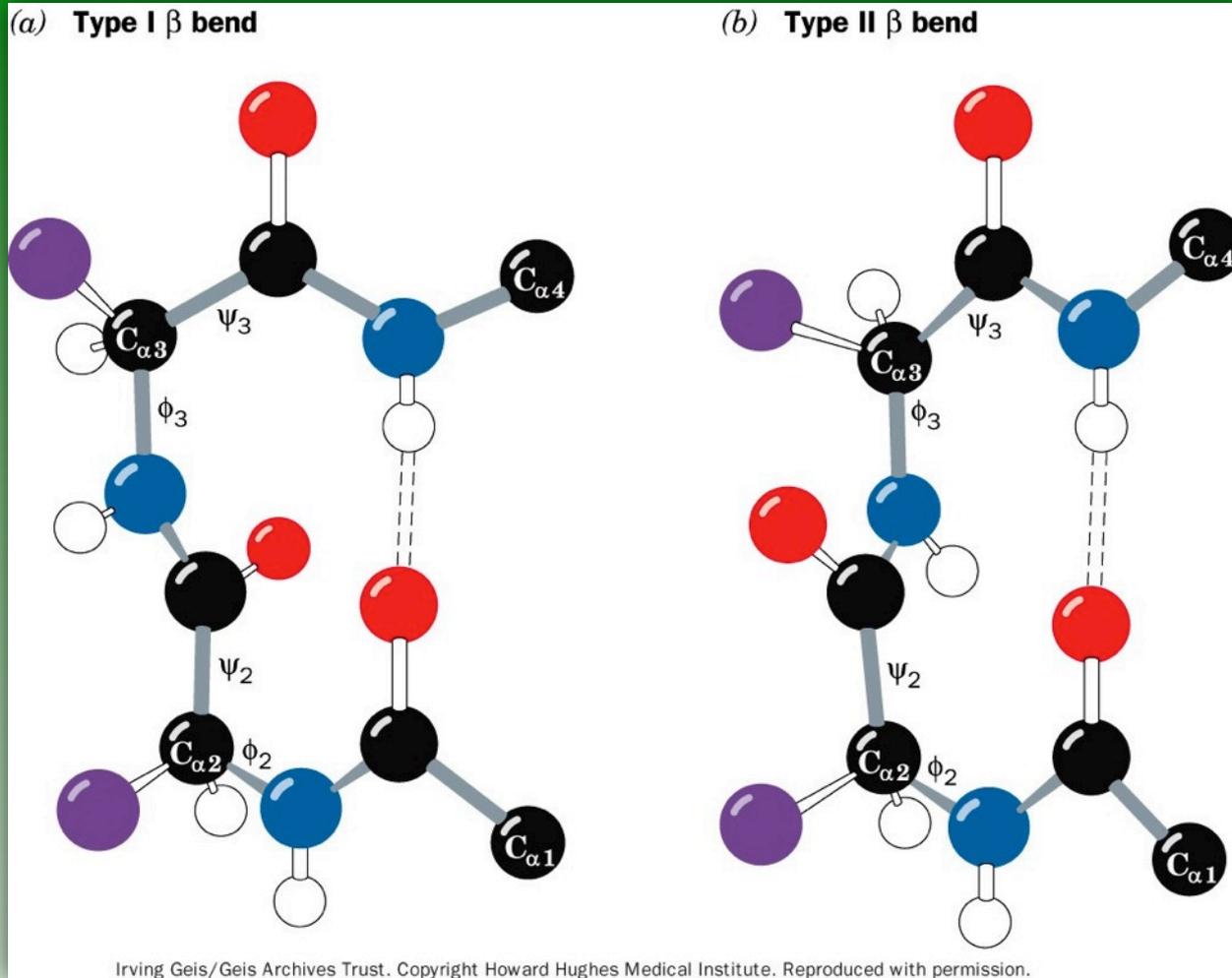
Stacking of side-by-side β -sheet structures showing “registry” of their R-groups. Spatial complementarity produces protein strength (illustrated for the silk protein)



Polypeptide chain folding in a globular protein illustrating the right-handed twist of β sheets: **bovine carboxypeptidase A**



Connections between adjacent polypeptide segments in β -pleated sheets: hairpins are sufficient for antiparallel β -sheet formation; crossovers are minimally required for parallel β -sheet formation



Reverse turns in polypeptide chains: two (2) residues per turn for a β -bend, stabilized by a single H-bond

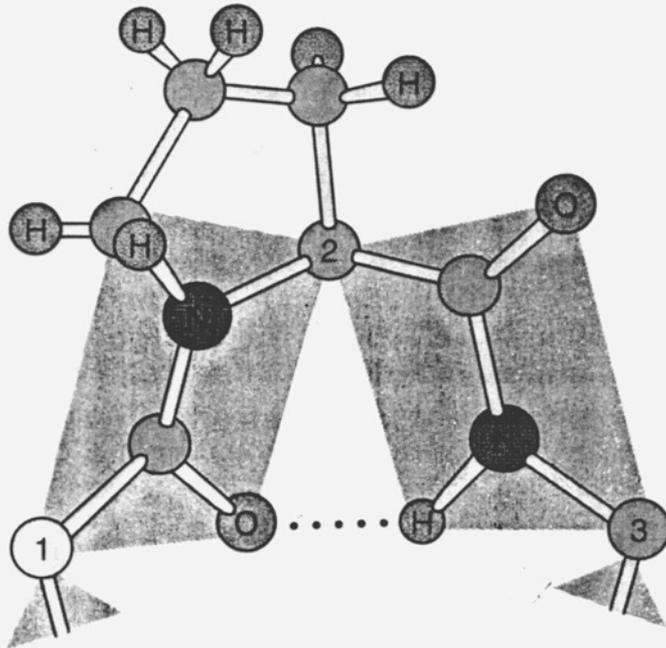
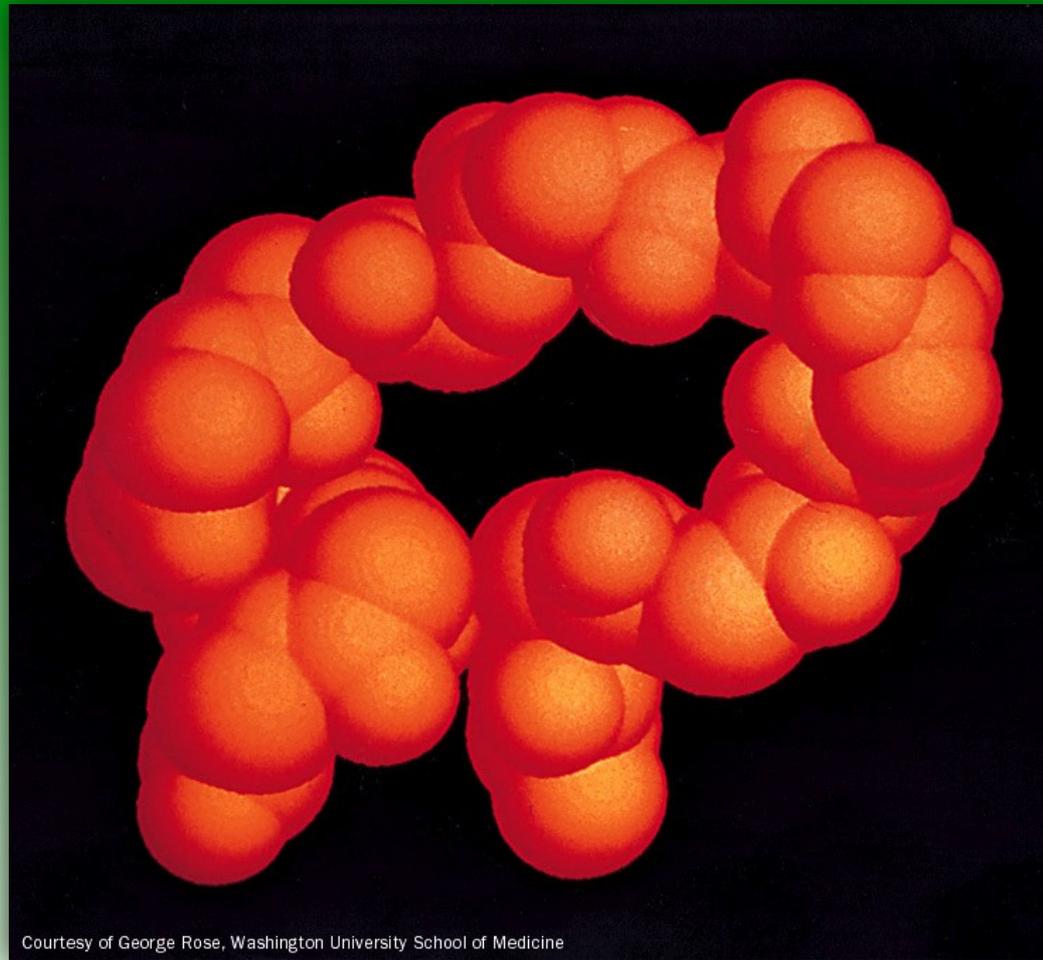


FIGURE 6.19

A γ turn. Only one residue is out of the hydrogen bonding sequence. In this case it is a proline, which cannot make such a bond in any case.

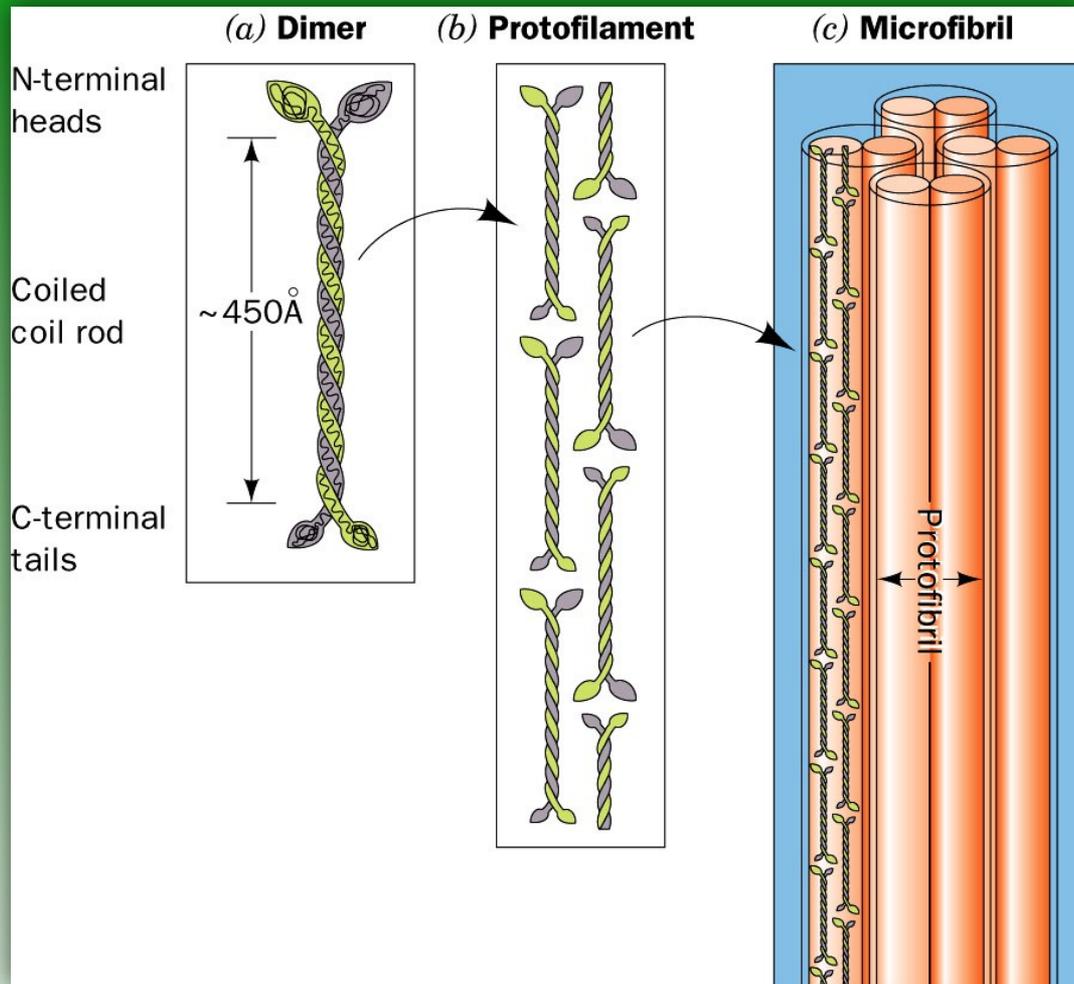
Gamma (γ) turns (tighter)



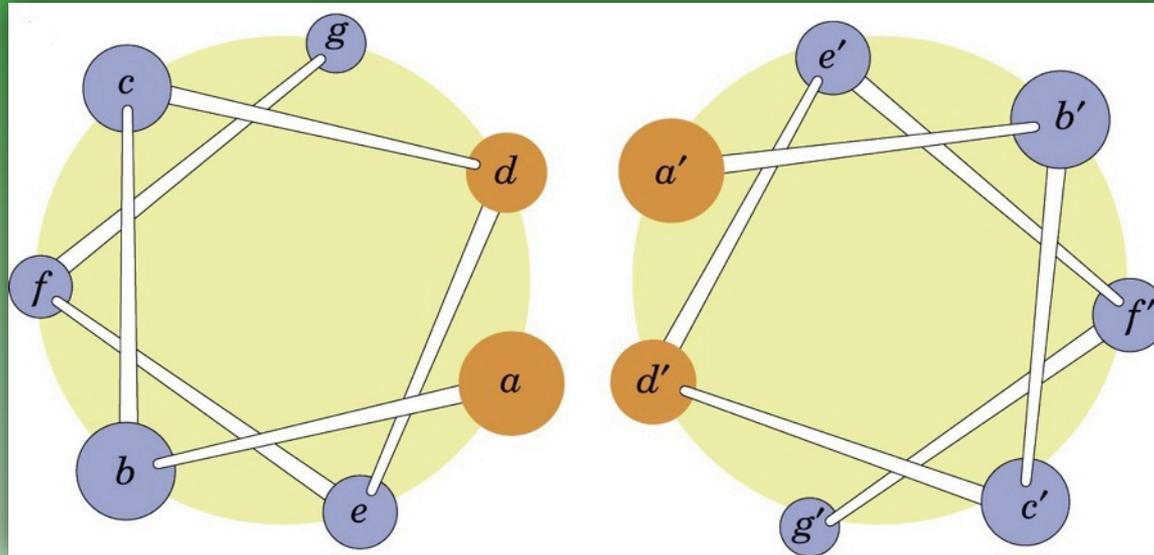
Another backbone bending motif: Space-filling representation of an Ω (omega) loop comprising residues 40 to 54 of cytochrome *c*

Fibrous proteins

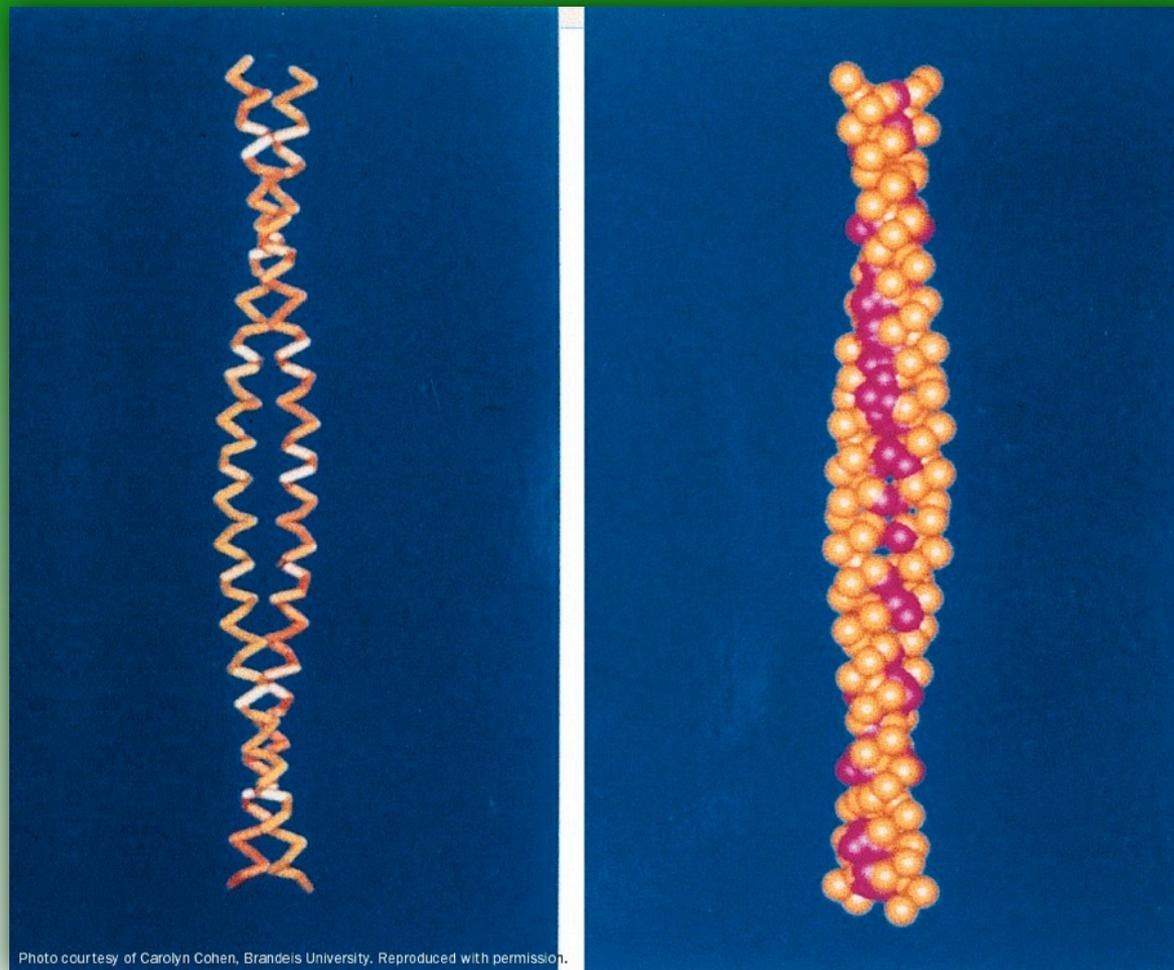
- α -keratins
- collagen



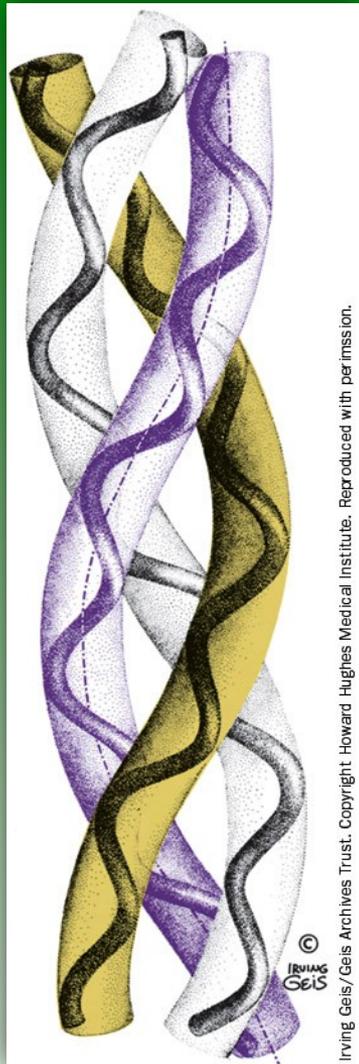
A fibrous protein: The structural organization of α -keratin
 What stabilizes the formation of the coiled coil?



The two-stranded coiled coil: view down the coil axis showing the interactions between the non-polar edges of the α -helices. The α -helices have a repeating heptameric sequence $a-b-c-d-e-f-g$ in which residues a and d are predominantly non-polar.



The two-stranded coiled coil: side view in which the polypeptide back bone is represented by skeletal (*left*) and space-filling (*right*) forms.



The right-handed triple helix of collagen

Another fibrous protein

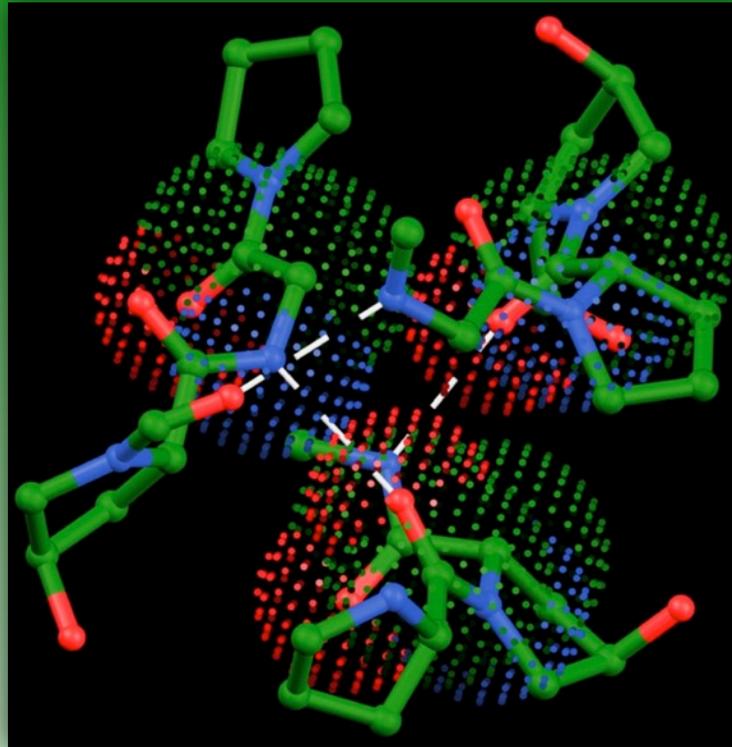
Collagen is an extracellular protein organized into insoluble fibers having great strength; **a major component of connective tissue.**

Its amino acid composition is distinctive:
~33% Gly and 15-30% Pro and 4-hydroxyproline (Hyp). 3-Hydroxyproline and 5-hydroxylysine (Hyl) are also present.

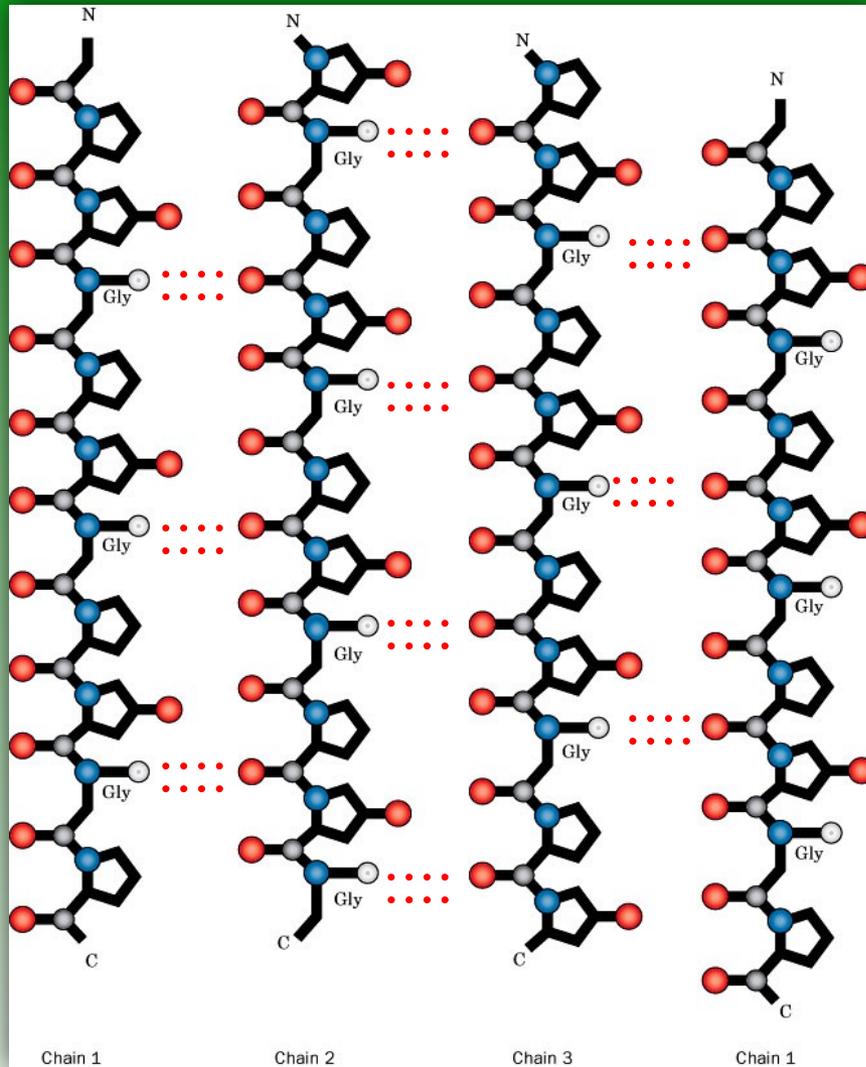
Collagen has a triple helical structure.
The individual polypeptide chains (polyproline-like helices - left-handed) are parallel and wound around each other with a right-handed rope-like twist to form the triple helical structure.

958	Gly	Pro	Arg	Gly	Pro	Hyp	Gly	Ser	Ala
967	Gly	Ser	Hyp	Gly	Lys	Asp	Gly	Leu	Asn
976	Gly	Leu	Hyp	Gly	Pro	Ile	Gly	Hyp*	Hyp
985	Gly	Pro	Arg	Gly	Arg	Thr	Gly	Asp	Ala
994	Gly	Pro	Ala	Gly	Pro	Hyp	Gly	Pro	Hyp
1003	Gly	Pro	Hyp	Gly	Pro	Hyp	Gly	Pro	Pro

The amino acid sequence at the C-terminal end of the triple helical region of the bovine $\alpha 1(I)$ collagen chain. Note repeating triplets Gly-X-Y where X is often Pro and Y is Hyp.



View along helix axis showing **inter-chain H-bonding** that stabilizes the triple helix structure (Gly N atoms and Pro O atoms in adjacent chains).

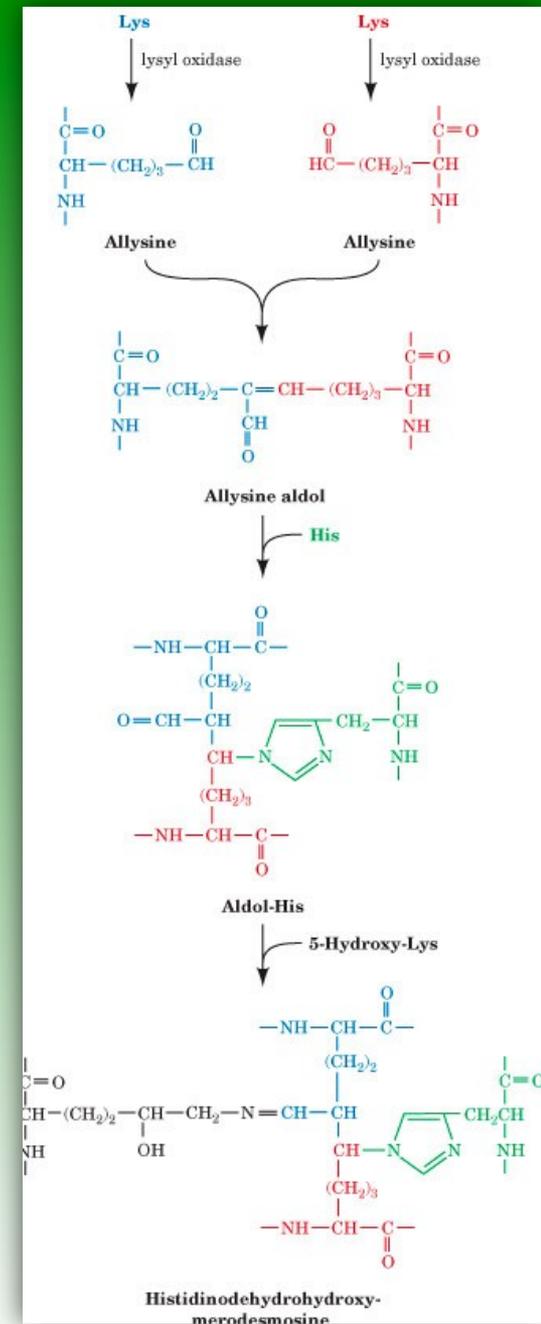


A schematic diagram showing **inter-chain H-bonding** in the Gly-containing regions of the triple helix of collagen.

Collagen is also **glycosylated** at Hyl residues with a Glc-Gal disaccharide (a post-translational modification / *O*-glycosylation).

Collagen (tropocollagen units) is organized into fibrils; the fibrils are **covalently crosslinked**.

A biosynthetic pathway for cross-linking **Lys**, **Hyl**, and **His** side-chains in collagen; crosslinking is enzyme-catalyzed.



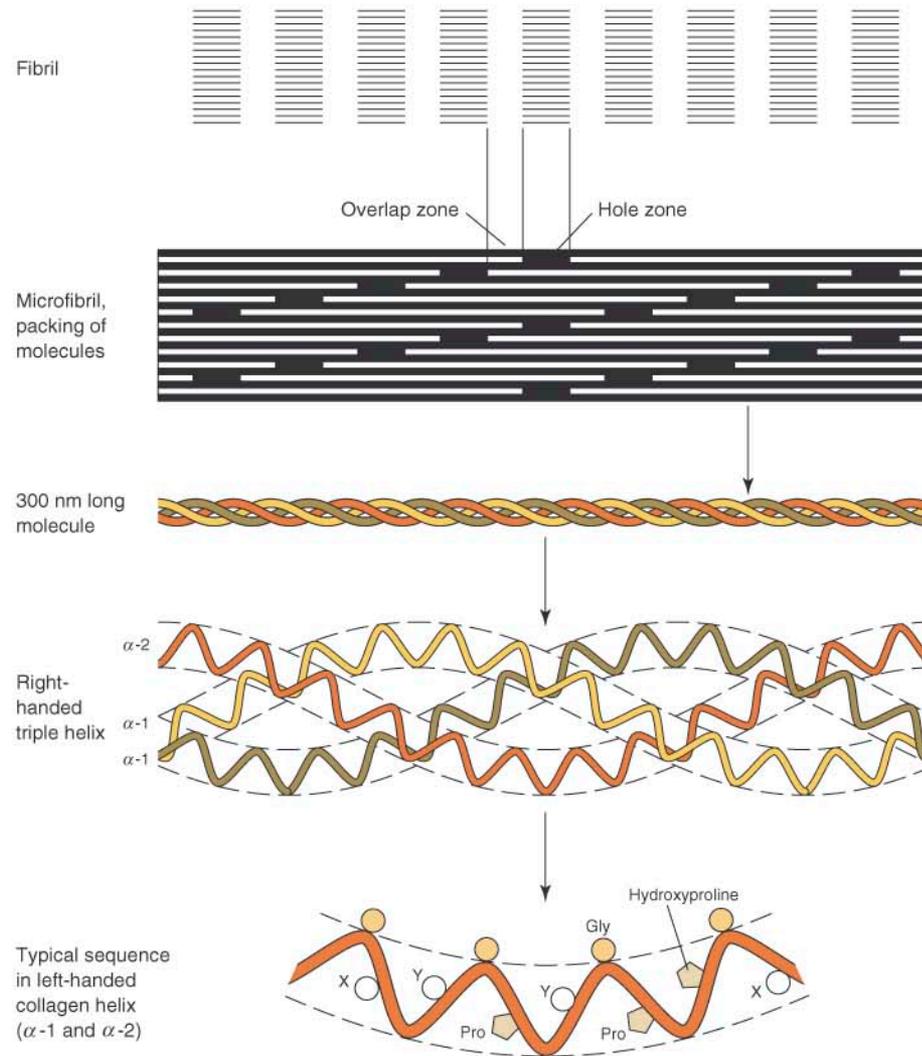
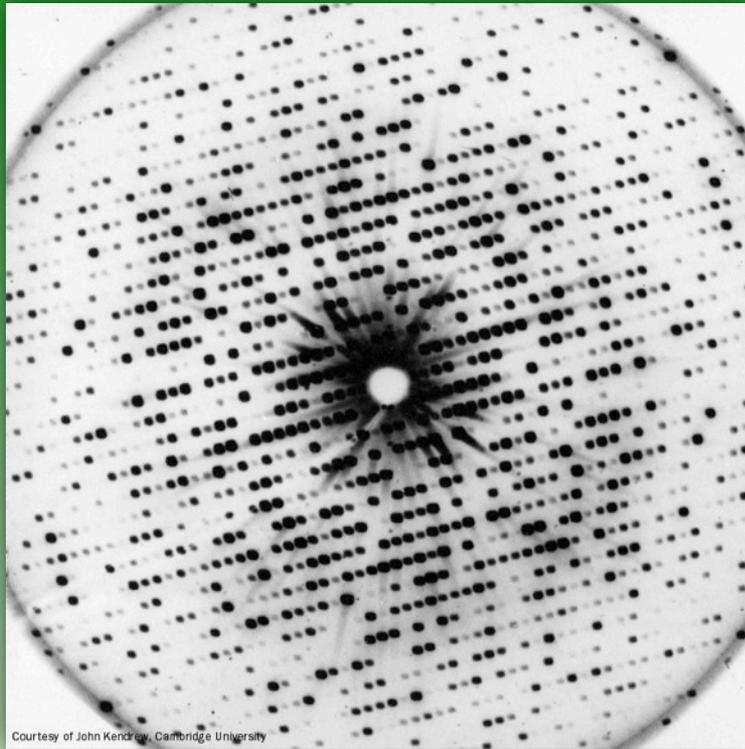


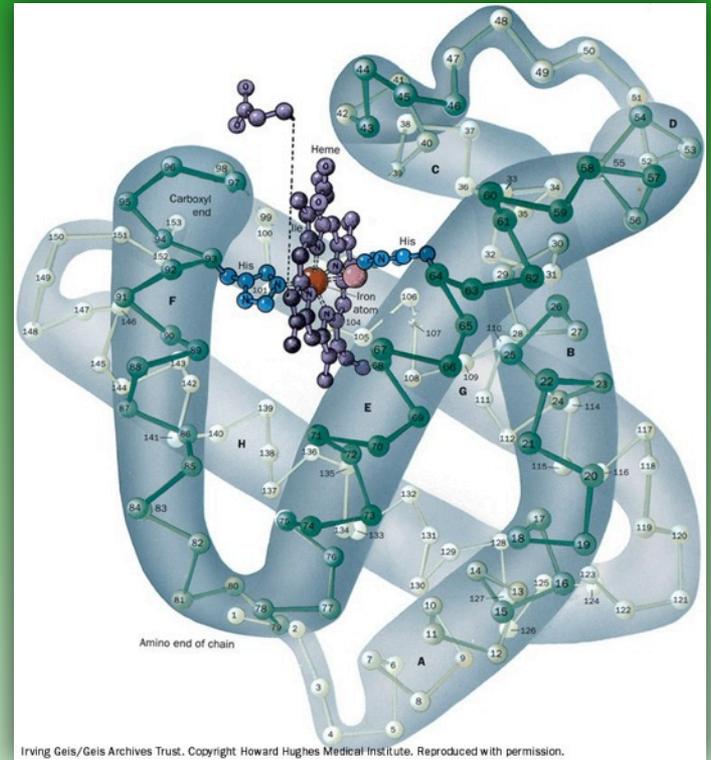
Figure 6.21. Collagen structure, illustrating (bottom to top) the regularity of primary sequence in a left-handed polyproline type II helix; the right-handed triple helix; the 300-nm molecule; and the organization of molecules in a typical fibril, within which collagen molecules are cross-linked.

Determination of protein 3D structure (globular):
Single crystal X-ray crystallography and NMR



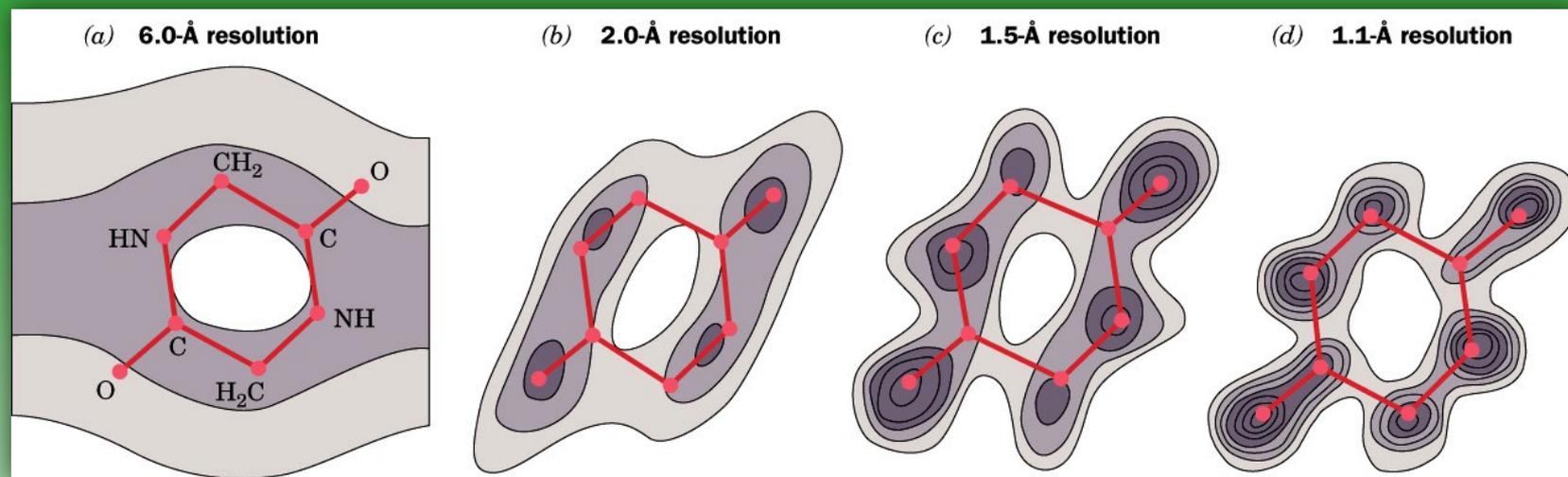
Courtesy of John Kendrew, Cambridge University

X-Ray diffraction photograph of a single crystal of sperm whale myoglobin

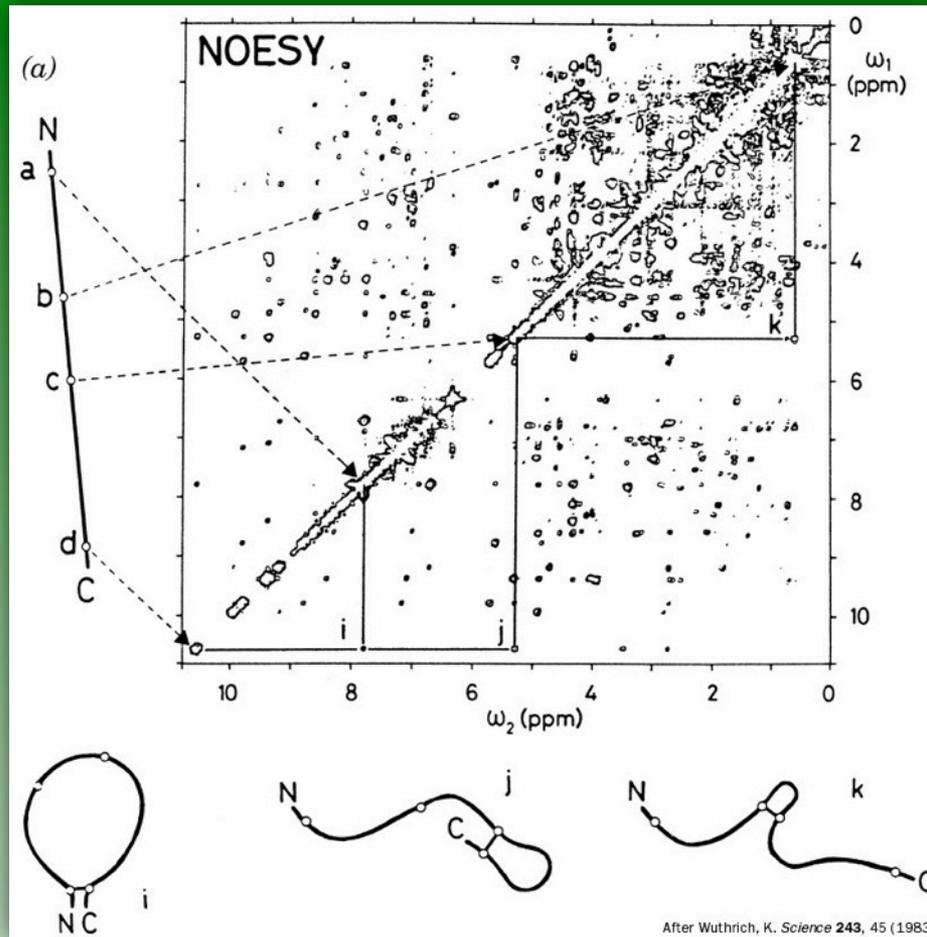


Irving Geis/Geis Archives Trust. Copyright Howard Hughes Medical Institute. Reproduced with permission.

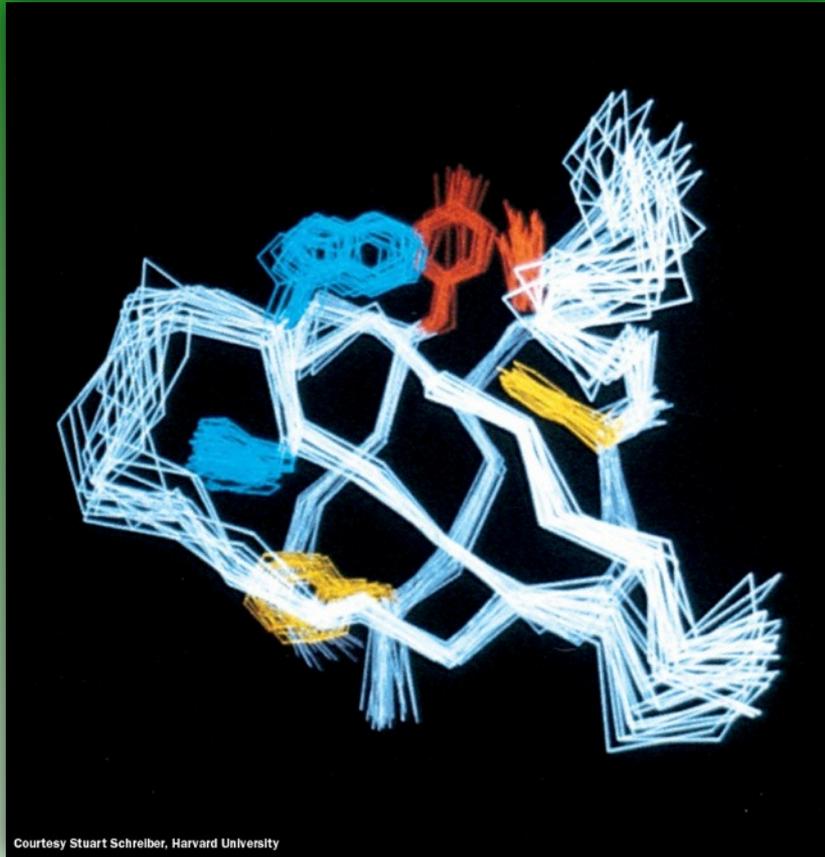
98-residue globular protein



Sections through the electron density map of a small organic molecule, diketopiperazine, calculated at the indicated resolution. As atomic resolution decreases, the ability to measure accurate bond lengths, angles and torsions also decreases (implications for structure/mechanism work, determination of substrate structure/conformation in a co-crystal).



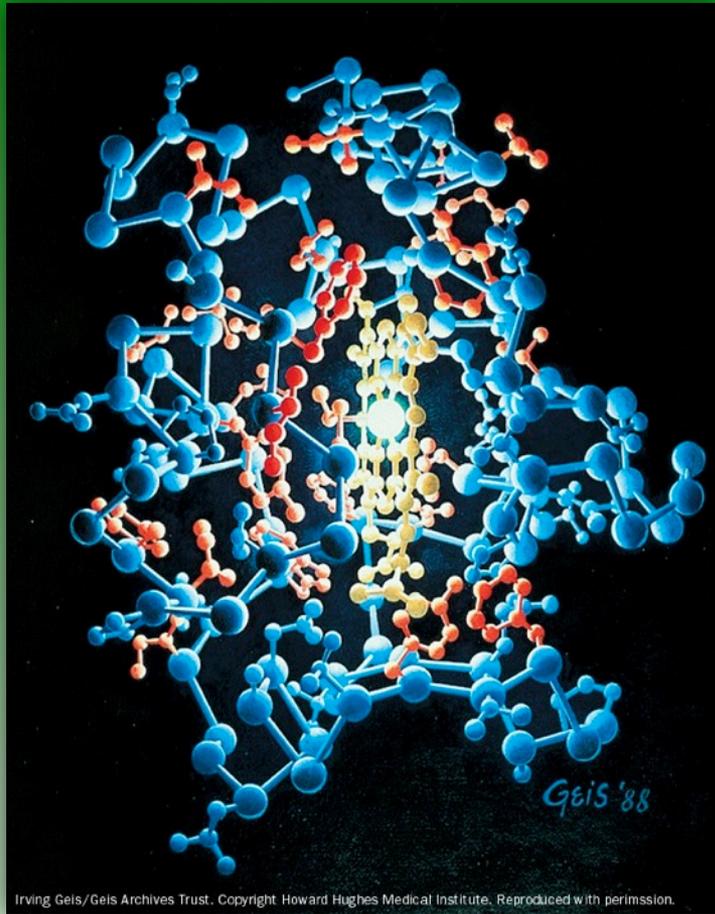
2D ^1H NMR spectra of proteins: a 2D NOESY spectrum of a protein presented as a contour plot with two frequency axes ω_1 and ω_2 .
 NOESY provides information about the relative internuclear distances between specific proton pairs in a molecule



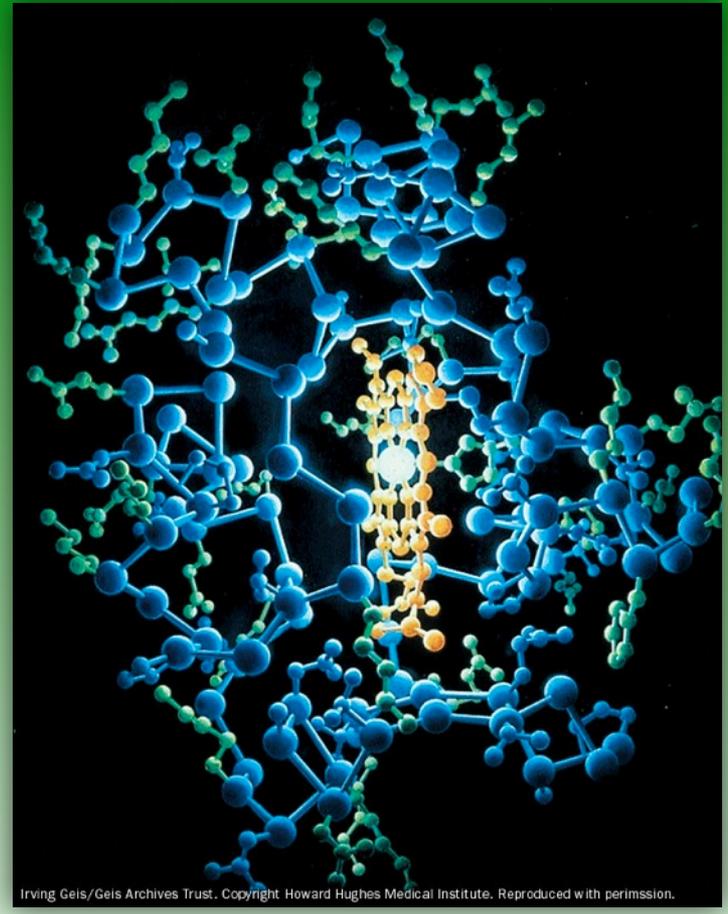
2D NMR structure of a 64-residue polypeptide comprising the Src protein SH3 domain.

NMR data can be collected in 3D and 4D dimensions using isotopically labeled protein. These multidimensional NMR datasets allow signal assignments in, and 3D structure determinations of, larger proteins.

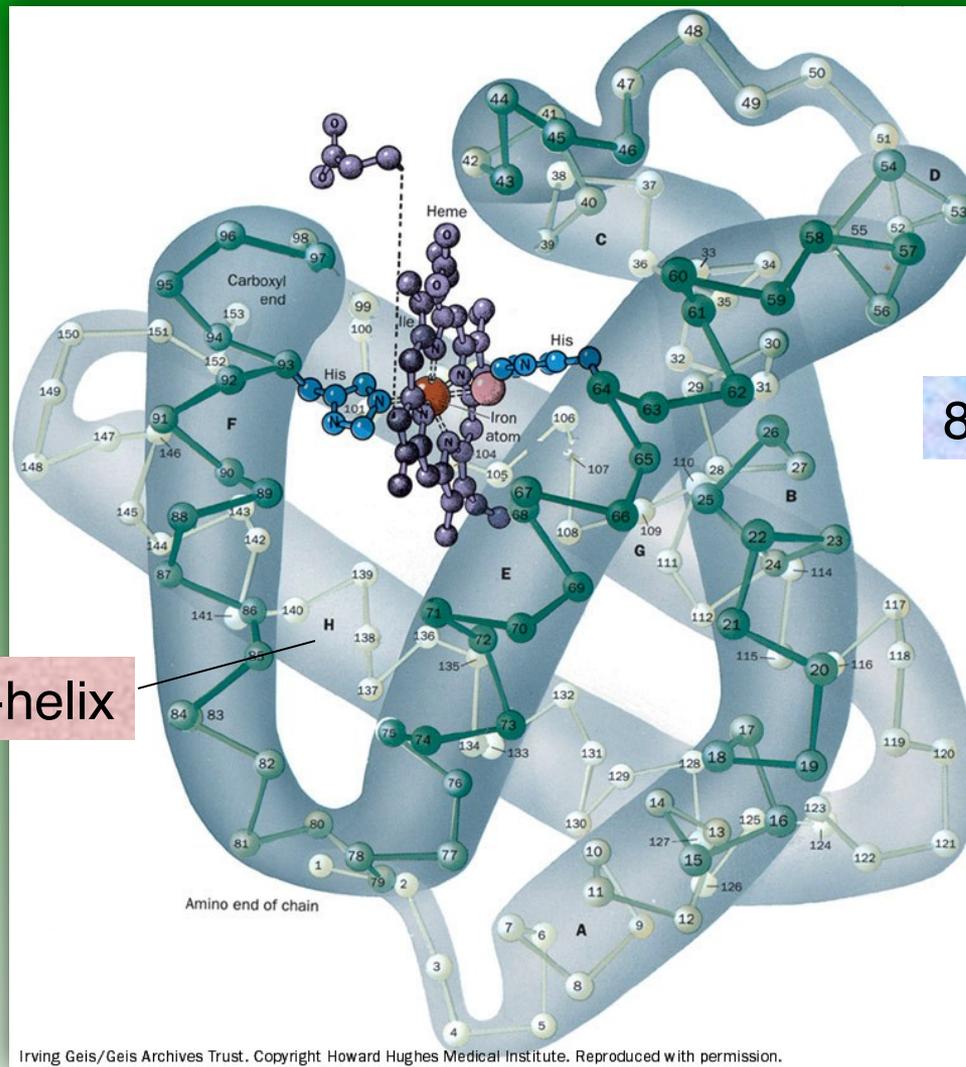
Some generalities about globular protein structure



X-ray structure of horse heart cytochrome c: hydrophobic residues (red) are largely buried



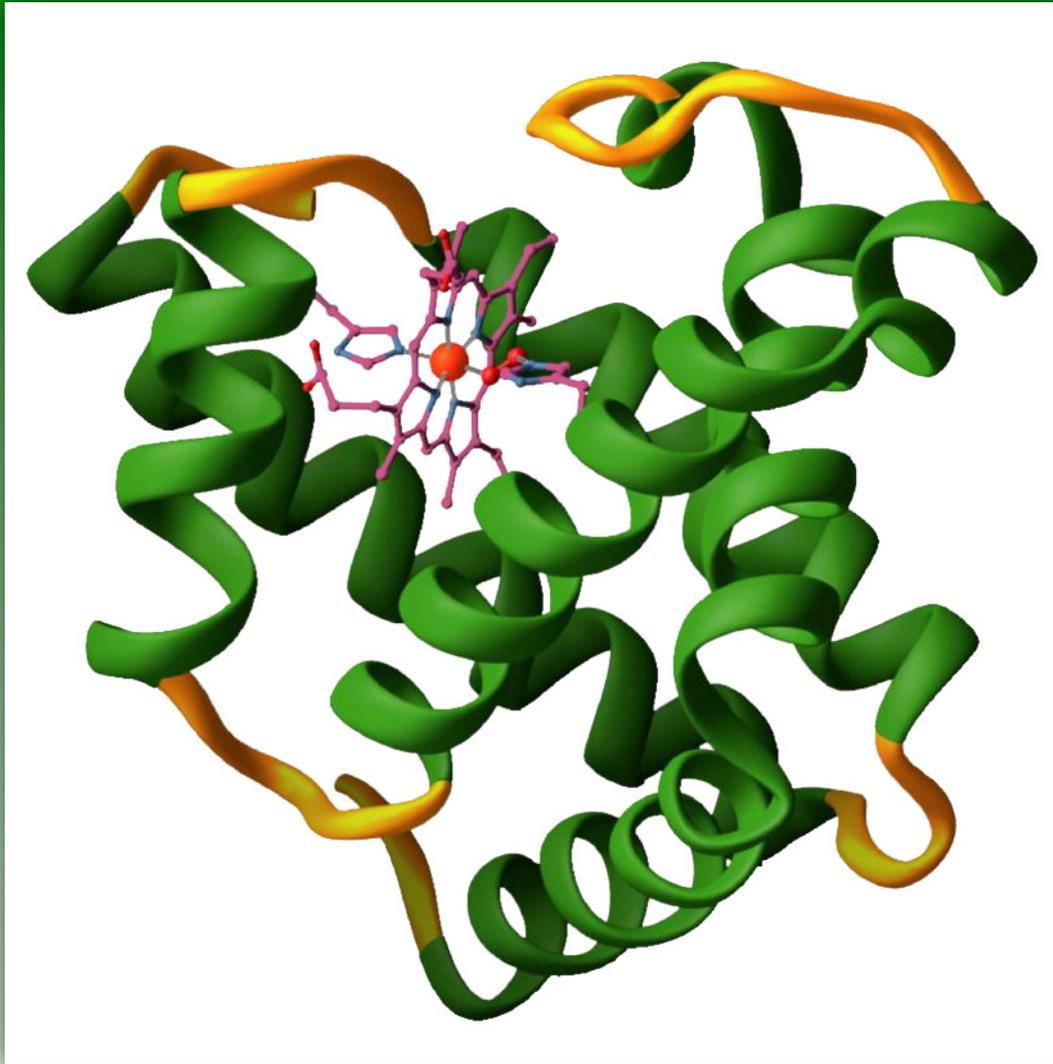
X-ray structure of horse heart cytochrome c: hydrophilic residues (green) are largely solvent exposed (on surface)



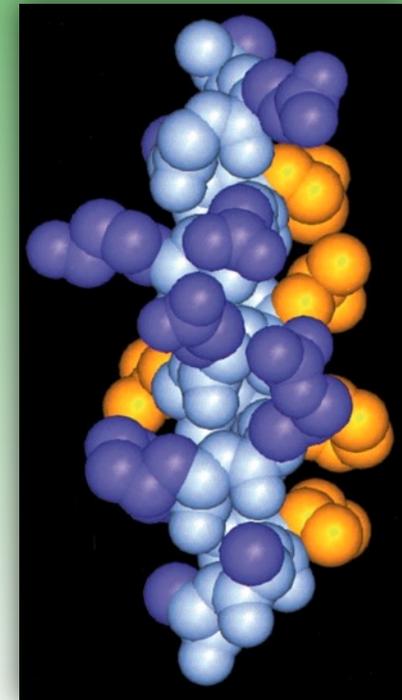
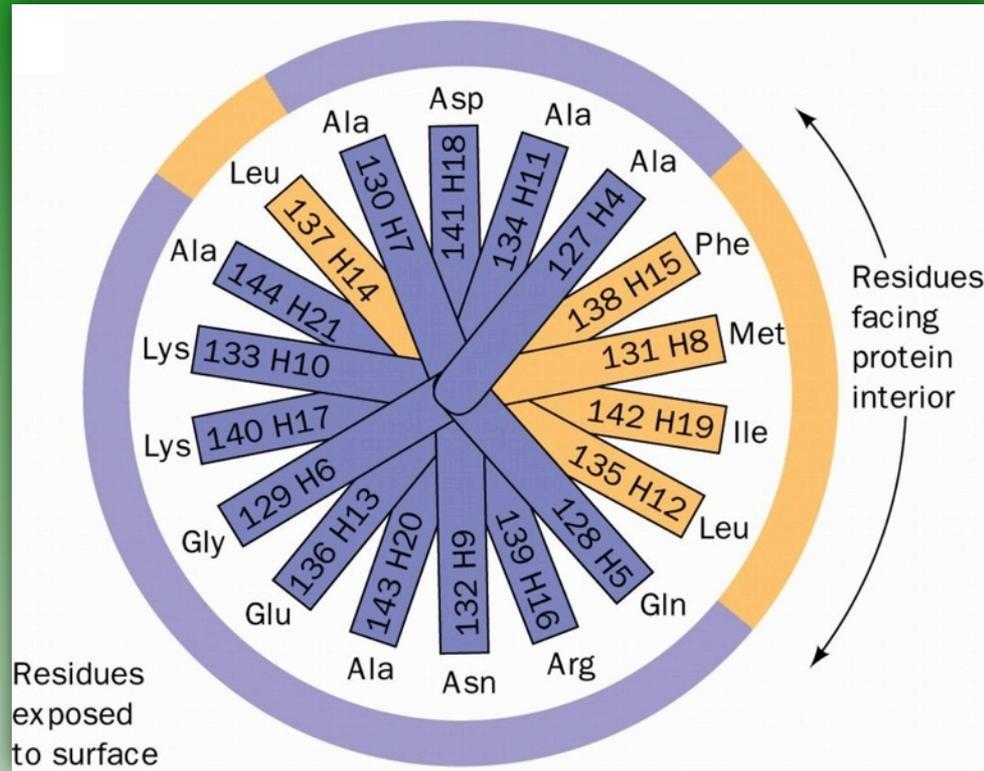
8 helices: A-H

H-helix

X-ray structure of sperm whale myoglobin identifying the H-helix

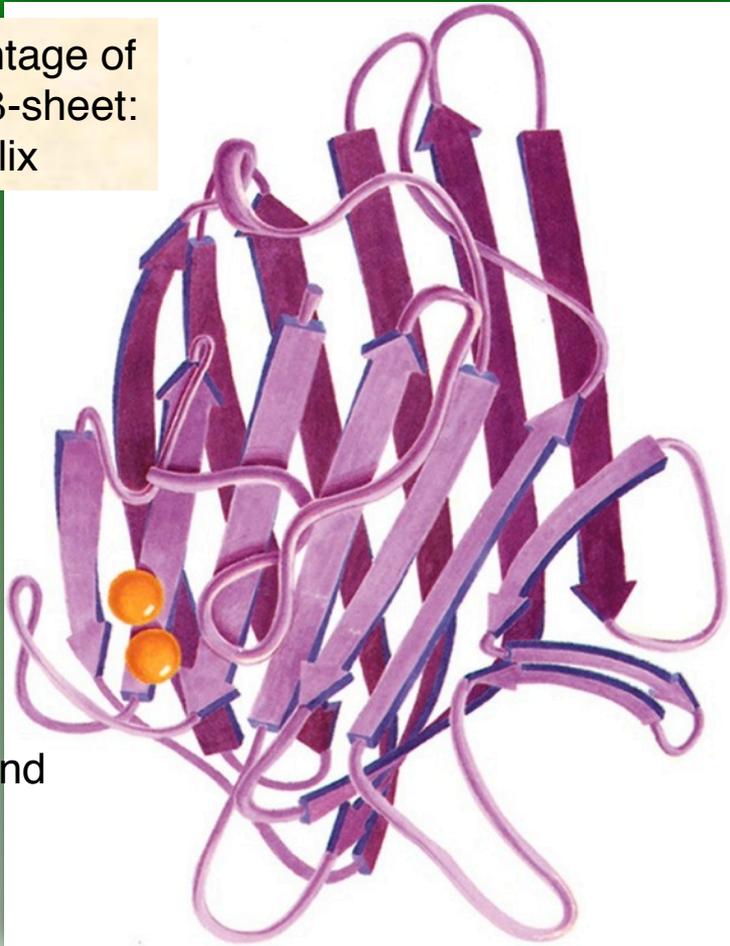


X-ray structure of sperm whale myoglobin: a computer-generated ribbon drawing



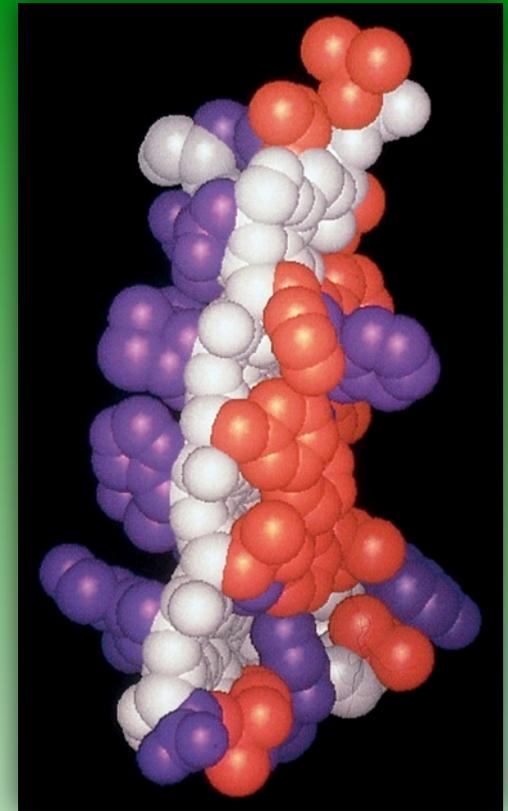
The **H helix** of sperm whale myoglobin. A **helical wheel** representation in which the side chain positions about the α helix are projected down the helix axis onto a plane.

Large percentage of anti-parallel β -sheet:
no α -helix



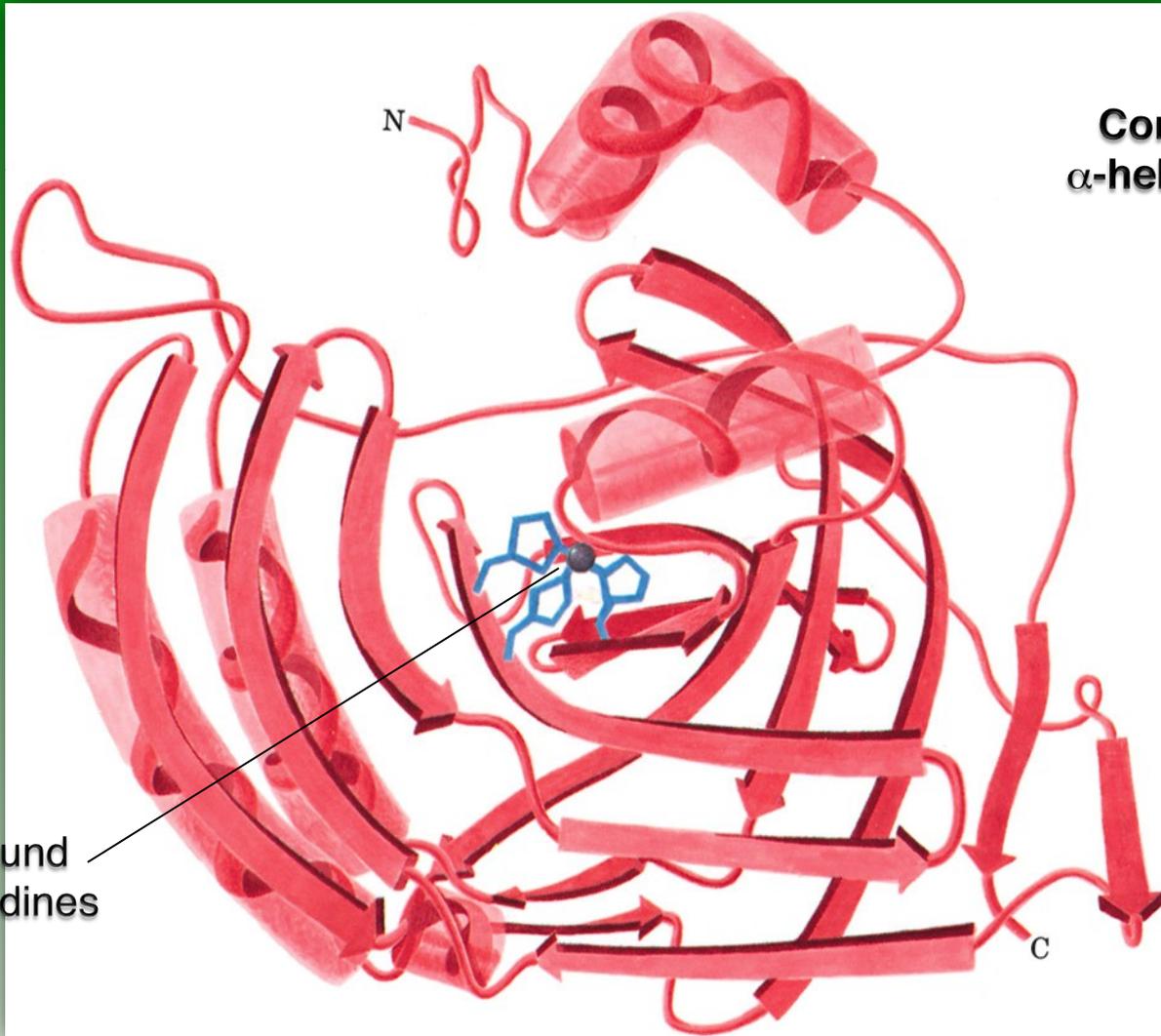
contains bound
metal

The X-ray structure of jack
bean protein concanavalin A



red = nonpolar
purple = polar

A space-filling model of
an antiparallel β sheet
from concanavalin A

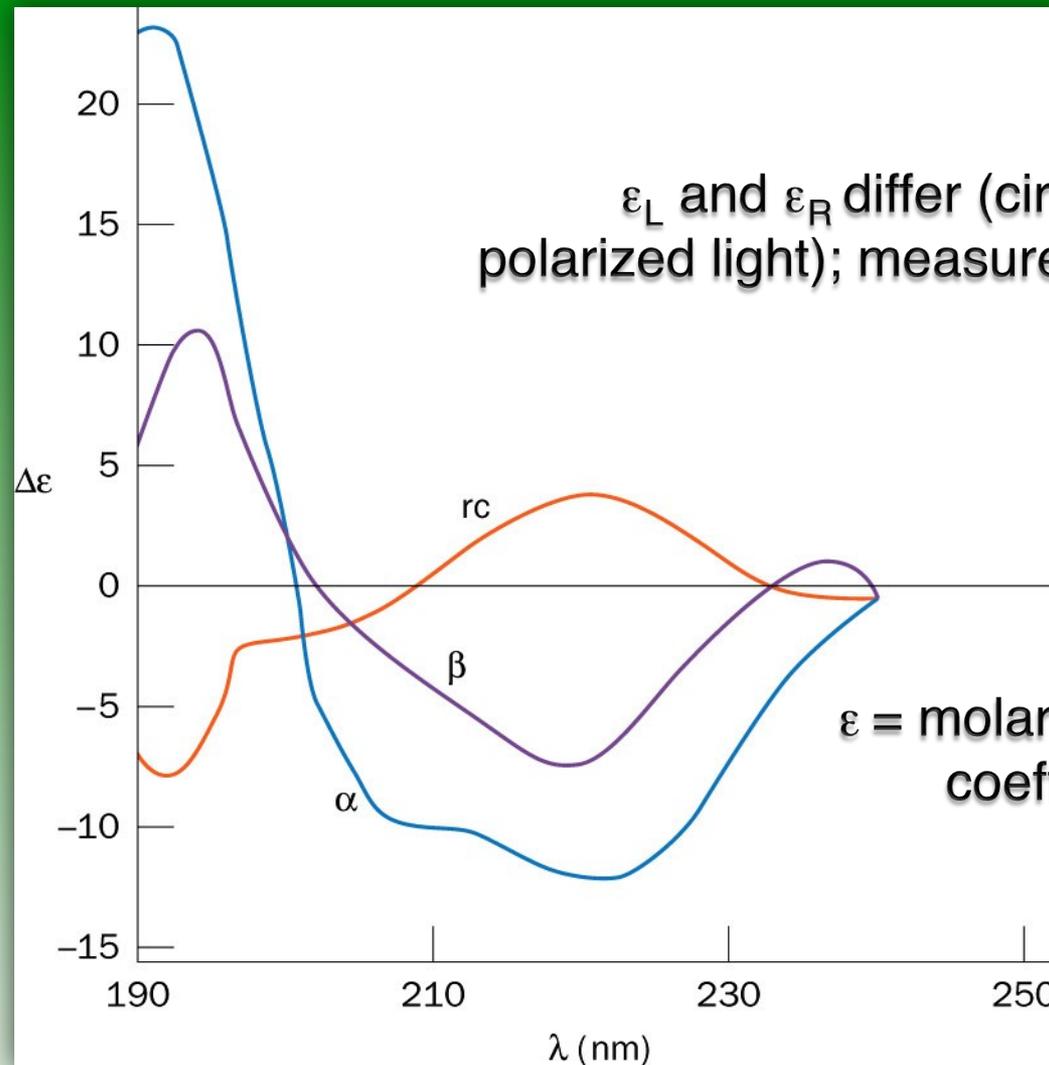


Contains mix of
 α -helix and β -sheet

Zn²⁺ ion bound
by three histidines

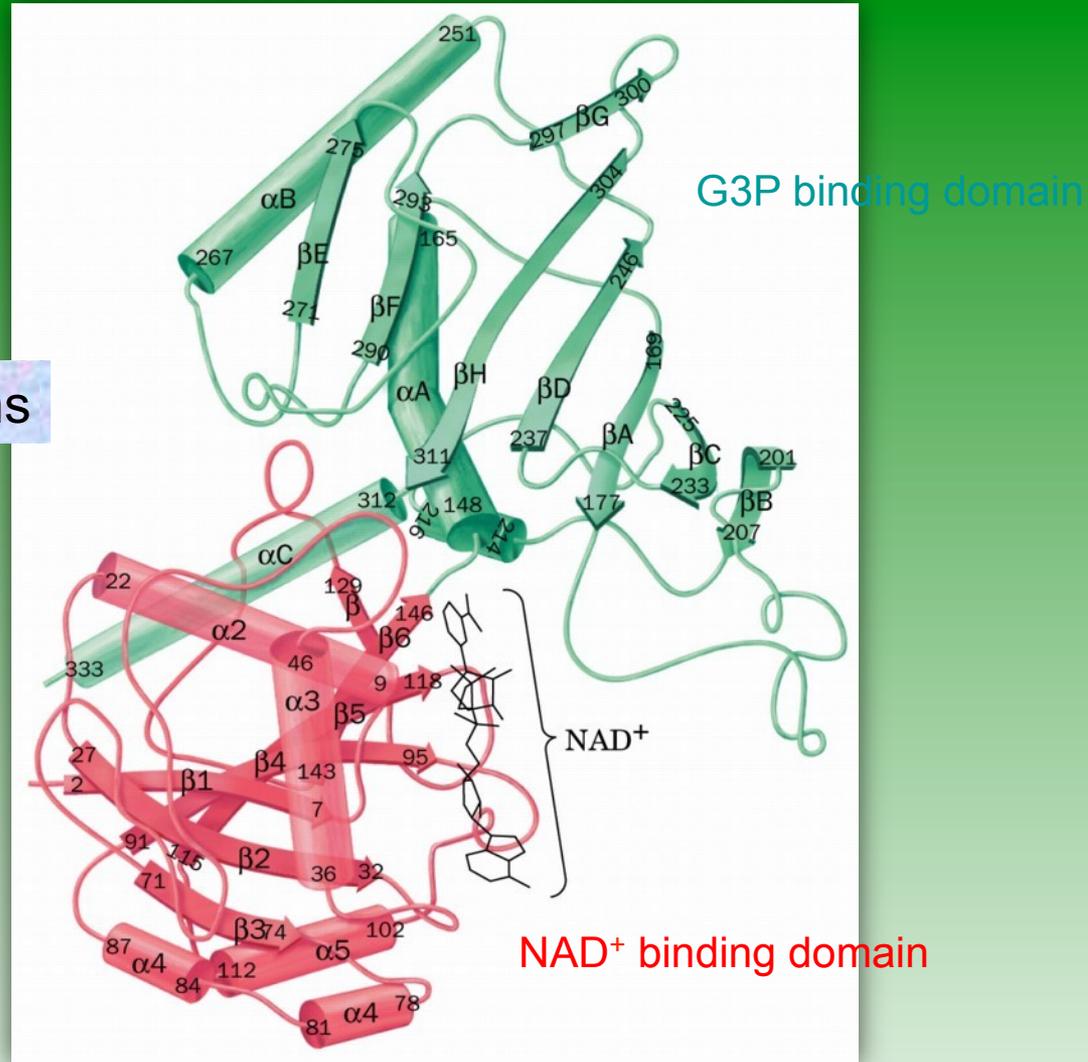
Human carbonic anhydrase

Only chiral molecules give a CD spectrum.



Circular dichroism (CD) spectra of polypeptides:
 α = α -helix; β = β -sheet; rc = random coil

two domains



One subunit of the enzyme, glyceraldehyde-3-phosphate dehydrogenase, from *Bacillus stearothermophilus*

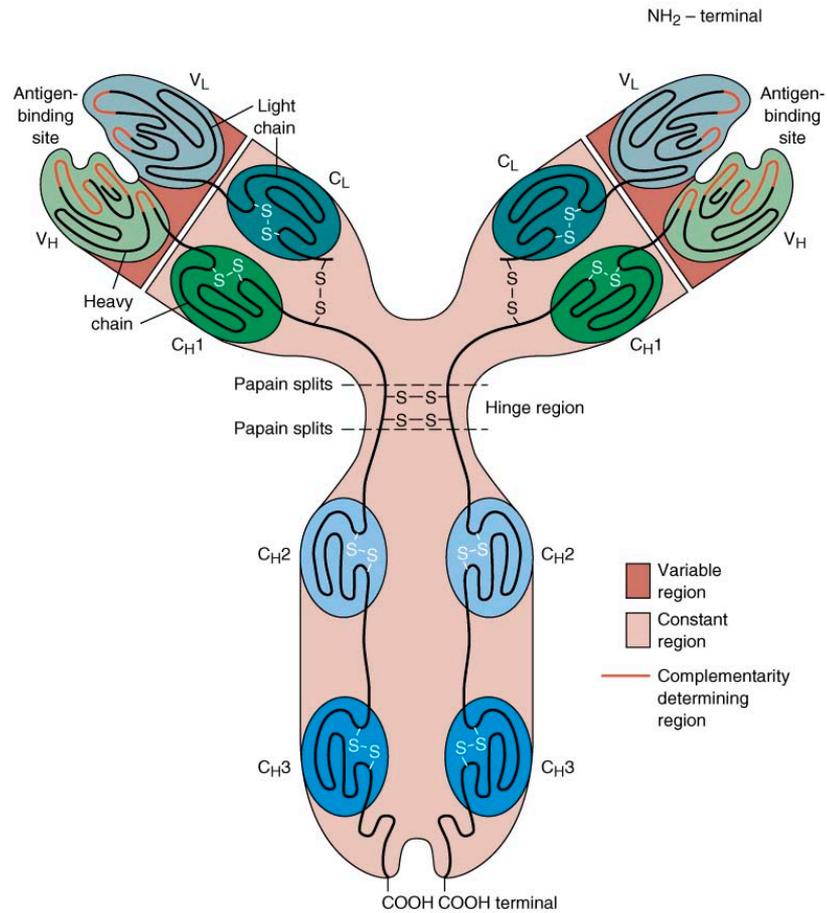
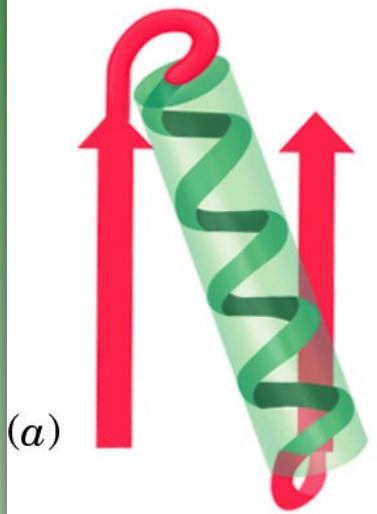


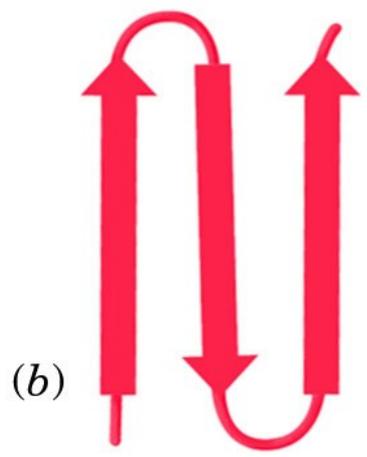
Figure 9.2. Diagrammatic structure of IgG. From Cantor, C. R. and Schimmel, P. R. *Biophysical Chemistry*, Part I, San Francisco: Freeman, 1980. Reprinted with permission of Mr. Irving Geis, New York.

**Schematic diagrams of some supersecondary
structures**

$\beta\alpha\beta$



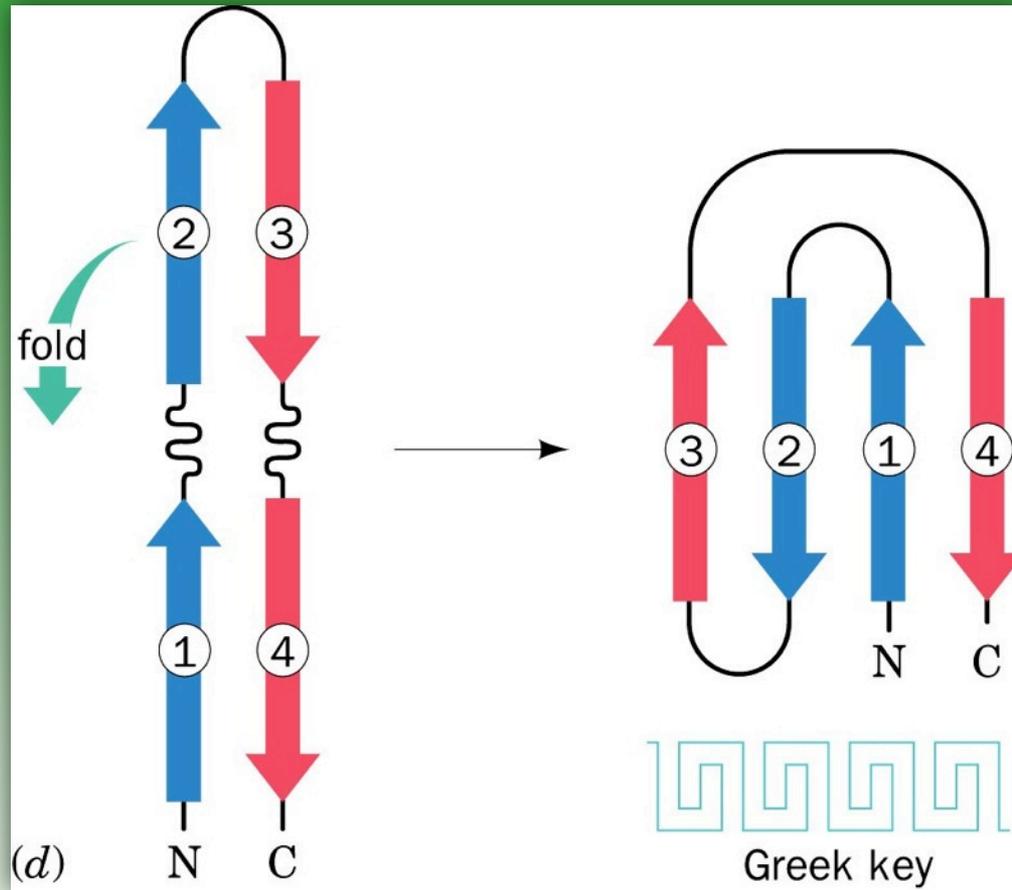
β -hairpin

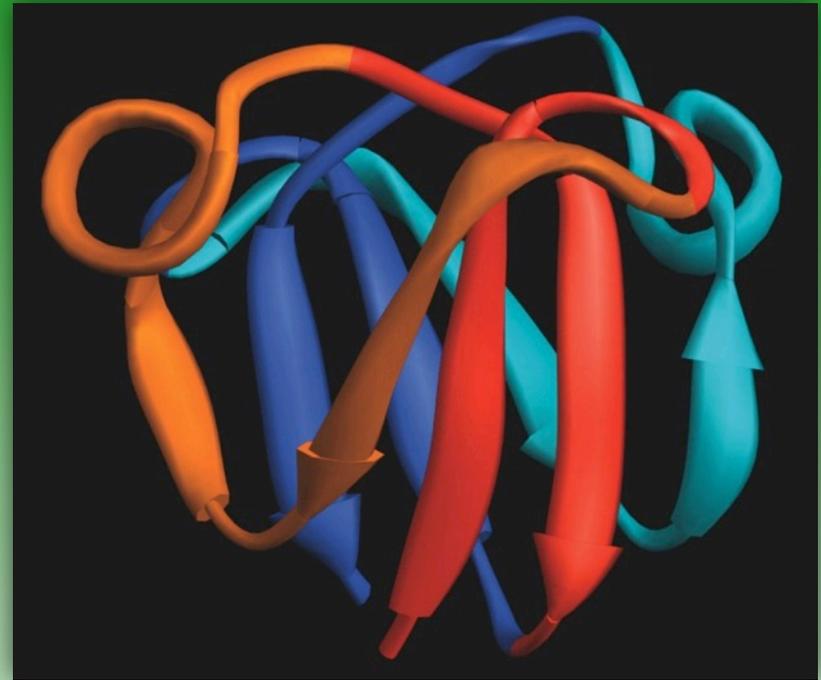
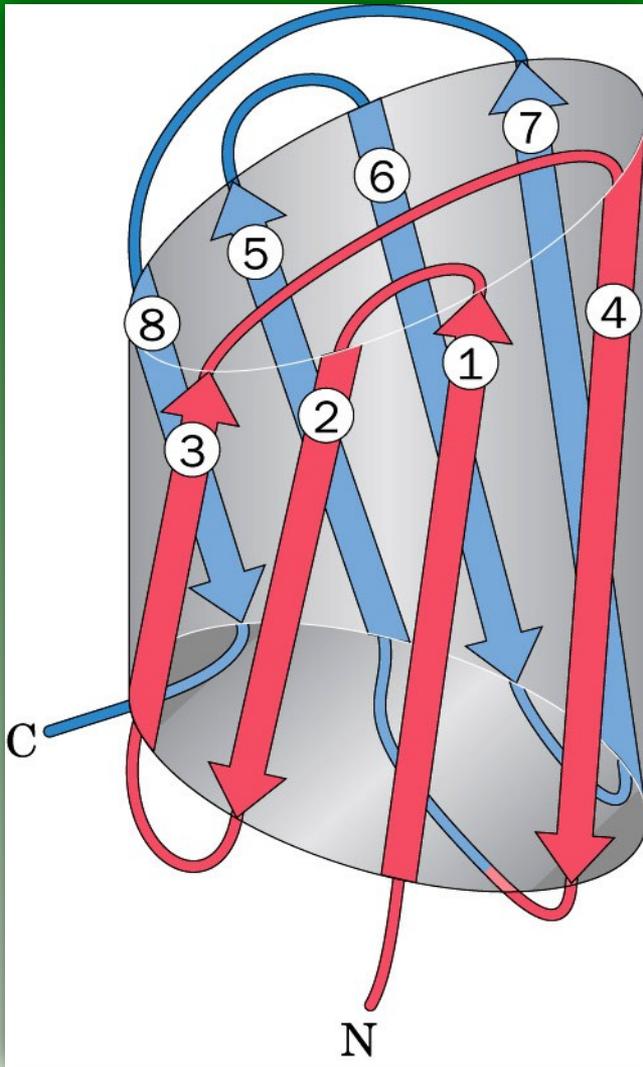


$\alpha\alpha$



Greek key motif





ribbon form

X-ray structure of the C-terminal domain (83 residues) of bovine γ - β crystallin: a topological diagram showing how its two Greek key motifs are arranged in a β barrel.

Protein-DNA binding motifs

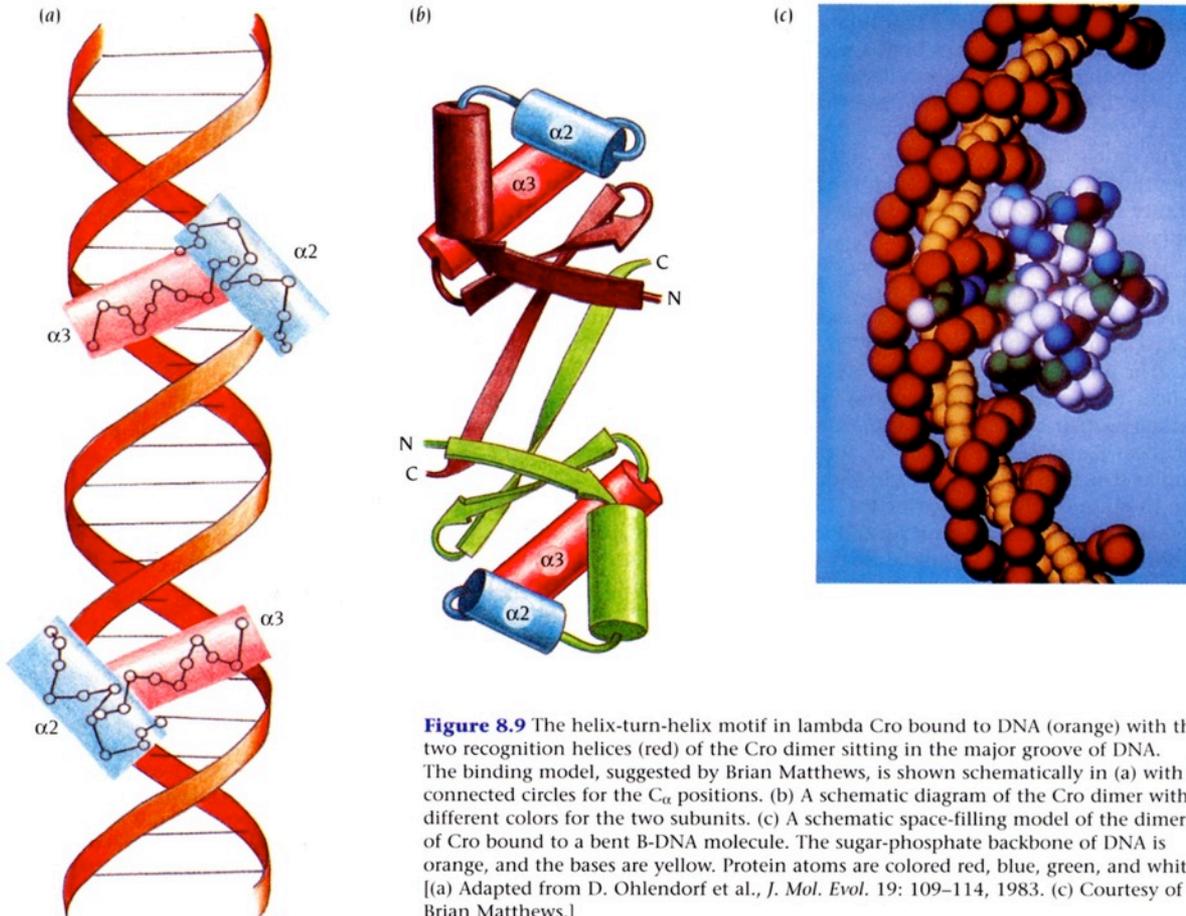
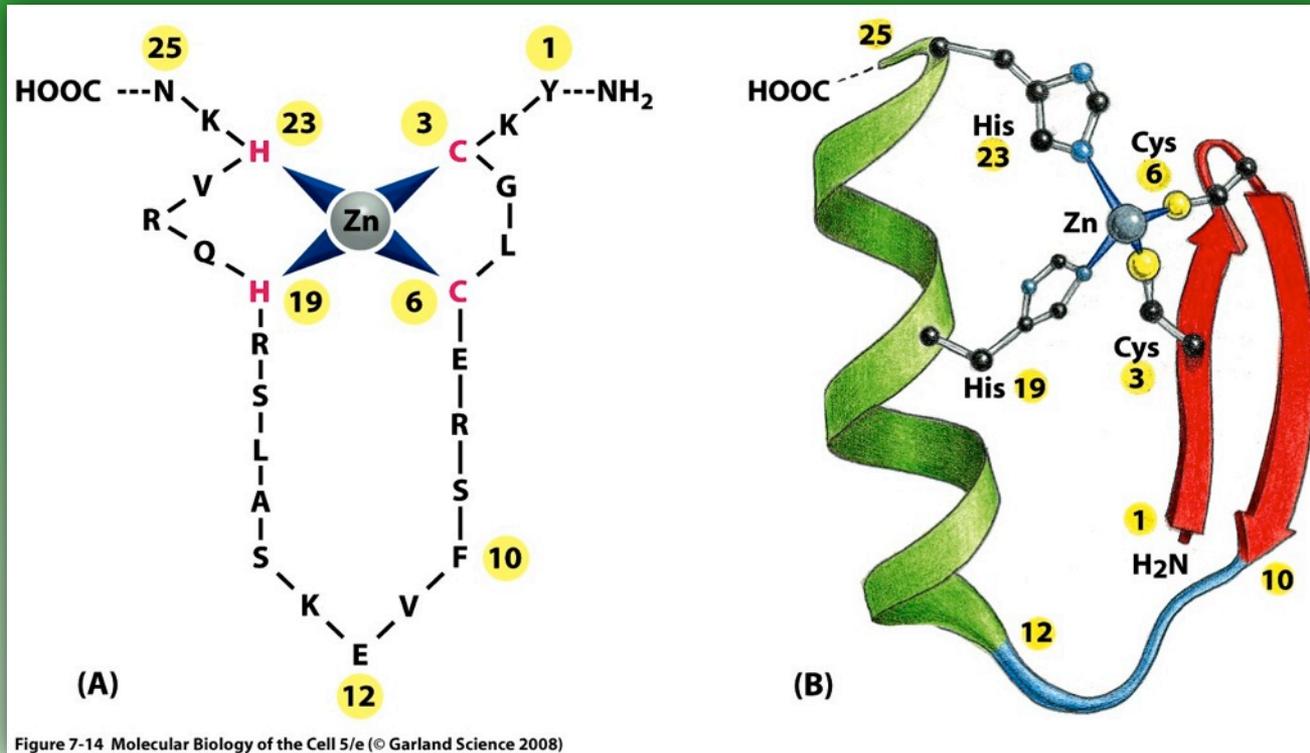


Figure 8.9 The helix-turn-helix motif in lambda Cro bound to DNA (orange) with the two recognition helices (red) of the Cro dimer sitting in the major groove of DNA. The binding model, suggested by Brian Matthews, is shown schematically in (a) with connected circles for the C_{α} positions. (b) A schematic diagram of the Cro dimer with different colors for the two subunits. (c) A schematic space-filling model of the dimer of Cro bound to a bent B-DNA molecule. The sugar-phosphate backbone of DNA is orange, and the bases are yellow. Protein atoms are colored red, blue, green, and white. [(a) Adapted from D. Ohlendorf et al., *J. Mol. Evol.* 19: 109–114, 1983. (c) Courtesy of Brian Matthews.]

**A structure-
function
correlation:
Role of α -helices
in the binding of
Cro dimer
To DNA**

**β -Structure
at a protein dimer
interface:
The stretch of
sequence
at the C-terminus
of each
Cro monomer
participates
in a β -sheet
secondary
structure, which
promotes
dimerization**



A zinc finger protein: Cys-Cys-His-His family

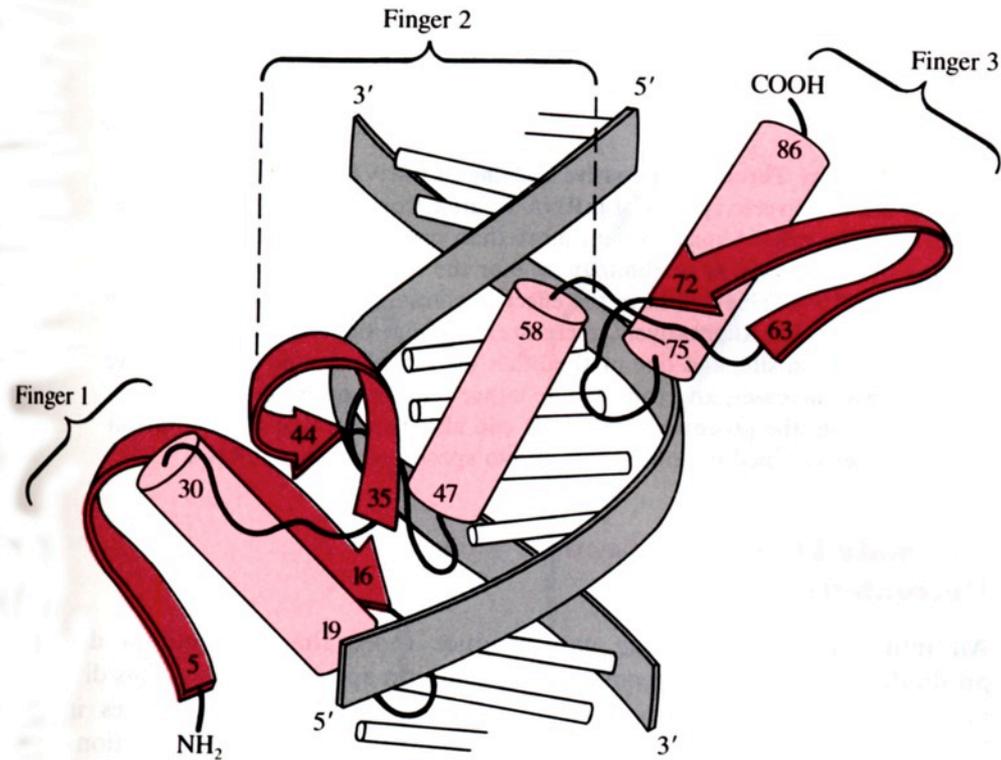


Figure 13-17

A Graphic Illustration of the X-ray Crystal Structure of the Zinc Finger Protein Zif268 Bound to DNA. Each zinc finger consists of an α -helix, shown as a cylinder, and a β -strand, shown as an arrow ribbon, held together through coordination to zinc. [After N. P. Pavletich and C. Pabo, *Science* (Washington, D.C.) 252 (1991): 809–817, Fig. 2.]



A **leucine zipper**
dimer bound to DNA

Protein fold classification

The columns of the table are based on domain architectures as defined by the CATH hierarchical classification. Each cell provides information on an interesting fold group within that architecture and highlights a particular structural domain from that group. The first row of each column typically contains the most basic fold group for that architecture followed by fold groups with more complex structures. The population given as a percentage for each architecture is calculated from the 527 genomes present in Gene3D version 6.0.

Known functions have been automatically assigned to one of eight categories in the Gene Ontology (GO) molecular function classification (see legend). These categories are represented as a coloured octagon around the structure and are based on a classification scheme devised by Christos Ouzounis. For each fold group, the GO categories are those identified for all structures within that fold group, excluding electronically inferred annotations, as well as all annotated sequence homologues to those structures (at 60% sequence identity, 80% overlap of the larger domain) in Gene3D. Functions are assigned based on the whole structure to which the domain belongs and may therefore not always represent a specific functional attribute of that domain.

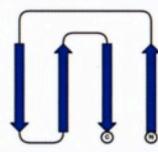
A white octagon tile means that no proteins in that fold group have that function. The incremental filling of the tile (by 1/4, 1/2, 3/4, 1/1) indicates the presence of the respective functions in the fold group and their relative importance (i.e. up to 25 / 50 / 75 / 100% of all proteins in that fold group have that function). For the fibrous proteins the functional mapping is simply that of the particular structure shown and its 60% sequence identity homologues. A completely blank octagon reflects the fact that currently no function can be automatically mapped to that fold, but not necessarily that no known function exists.



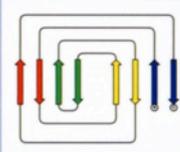
Basic topologies of secondary structure



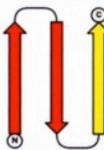
β -Hairpin



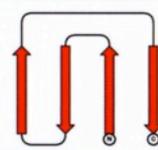
N-type Greek key



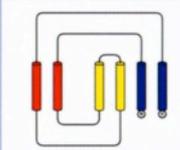
Jelly roll



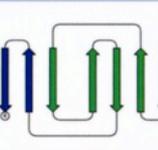
β -Meander



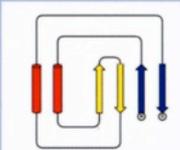
C-type Greek key



α Solenoid



Ig domain

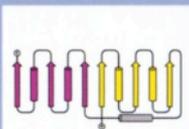


$\alpha\beta$ -Plait

The β -hairpin and the β -meander are examples of the simplest type of up-and-down antiparallel β -sheet topologies. The latter has an additional strand (yellow) with respect to the former. They can be observed in innumerable examples of β -structures shown in the main table.

The **Ig fold** is one of the most well-known of all protein structures. An Ig constant domain, as shown here, has an **N-type Greek key** embedded within it (shown in green).

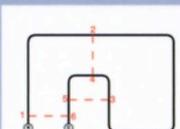
The **jelly roll**, **α -solenoid** and **$\alpha\beta$ -plait** topologies are surprisingly similar despite their different secondary structures. All are based on a giant hairpin which is subsequently rolled up to form a compact domain in which the secondary structures form two layers (the β -sheets in the case of the jelly roll). The latter is probably the most well-known of the three folds and takes its name from the topology diagram of its β -strands, which resembles a slice of a jelly roll (Swiss roll).



Rossmann fold

The **Rossmann fold** was first described by Michael Rossmann as a motif observed in lactate dehydrogenase. It is common in α/β proteins with open β -sheets and particularly

in nucleotide binding proteins. It is composed of two topologically identical and pseudo-symmetrically related substructures, which are shown in different colours. The term Rossmann fold is also occasionally used to refer to such substructures. Several examples can be seen in the main table.



Kringle

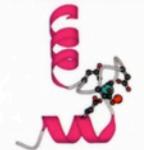
The **kringle** takes its name from the topology imposed by its three disulphide bridges, which in two dimensions resembles a Danish pastry of the same name. It is a common modular element in proteins of the coagulation pathways.

Important structural motifs



HTH motif

The **helix-turn-helix** and **EF-hand** motifs are both characterized by two orthogonal α -helices. The former is a specific example of an α/α corner and is found in DNA binding proteins, where the second (recognition) helix inserts into the major groove. The EF-hand is observed in Ca^{2+} binding proteins, where the Ca^{2+} is bound by a loop between the two helices.



EF hand

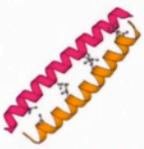


Greek key

The **Greek key** and the interdigitated **β -arches** are two of the most commonly observed motifs (or sub-structures) in β -proteins. The former is predominantly observed at the edges of antiparallel β -sheets where the motif is often divided between two such sheets. The latter has been described as the most common sub-structure observed in β -sandwiches.



2 β -Arches



Leucine zipper

Both the **leucine zipper** and the **helix-loop-helix** are dimerization motifs. In the former case this occurs via the formation of a classical left-handed coiled coil, with leucines at every 7th position (the d position of the coiled coil). In the case of the HLH, the two helices of the motif come together with those of the second monomer to form a 4-helix bundle.

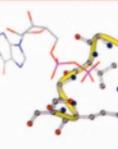


Zinc finger

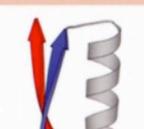
Zinc fingers are metal binding motifs involved in DNA recognition. They differ in their Zn^{2+} ligands, 3D structures and DNA binding modes. The example shown is a 'classical' Zinc finger involving two His and two Cys ligands. The **P-loop** is a glycine rich motif involved in nucleotide binding, where it interacts directly with the α and β phosphate moieties.



HLH motif



P-loop

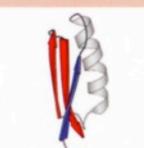


Right handed $\beta\alpha\beta$



right-handed $\beta\beta\beta$ motif
left-handed

Most connections between parallel β -strands are right-handed, but exceptions are to be seen in the left-handed β -helices of the main table. The **$\beta\alpha\beta\beta$ motif** includes an additional intervening antiparallel strand.



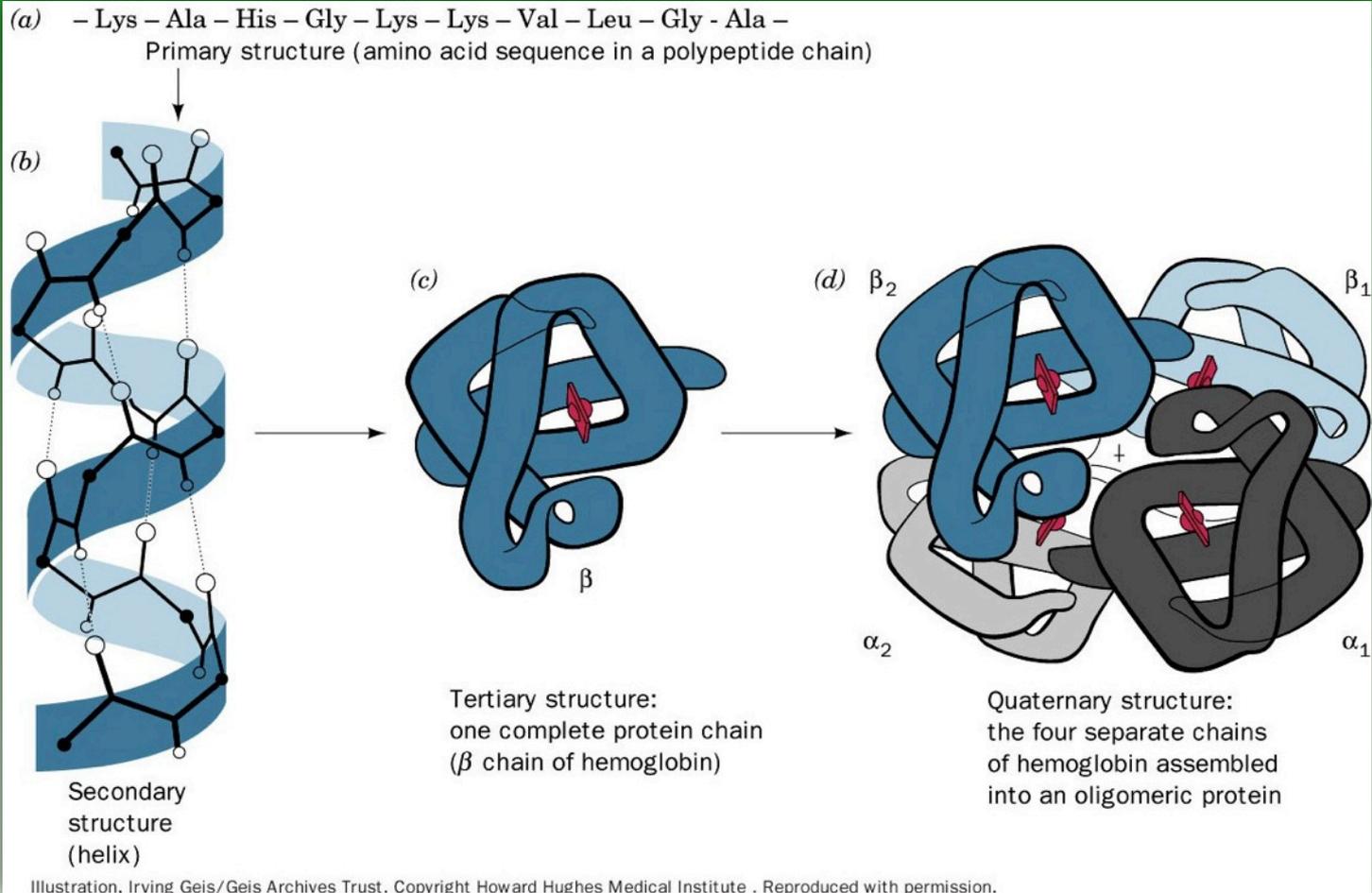
$\beta\alpha\beta\beta$ motif



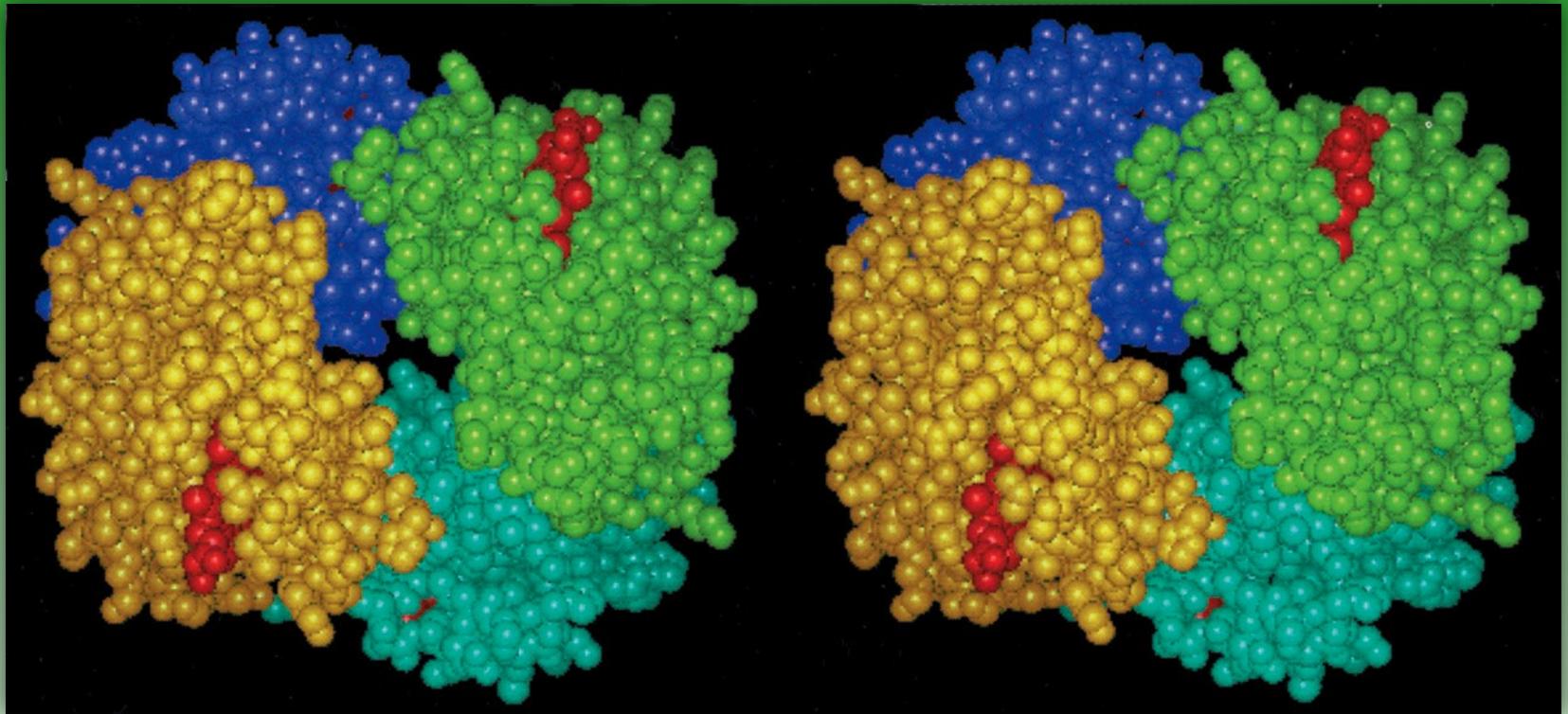
Leucine rich repeat

The **Leucine rich repeat (LRR)** is characterized by a sequence motif which typically contains 6 leucines. They form a structural motif of a β -strand, α -helix and connecting loop. Several examples of proteins, containing different numbers of repeats, can be seen in the $\alpha\beta$ -horseshoes of the main table.

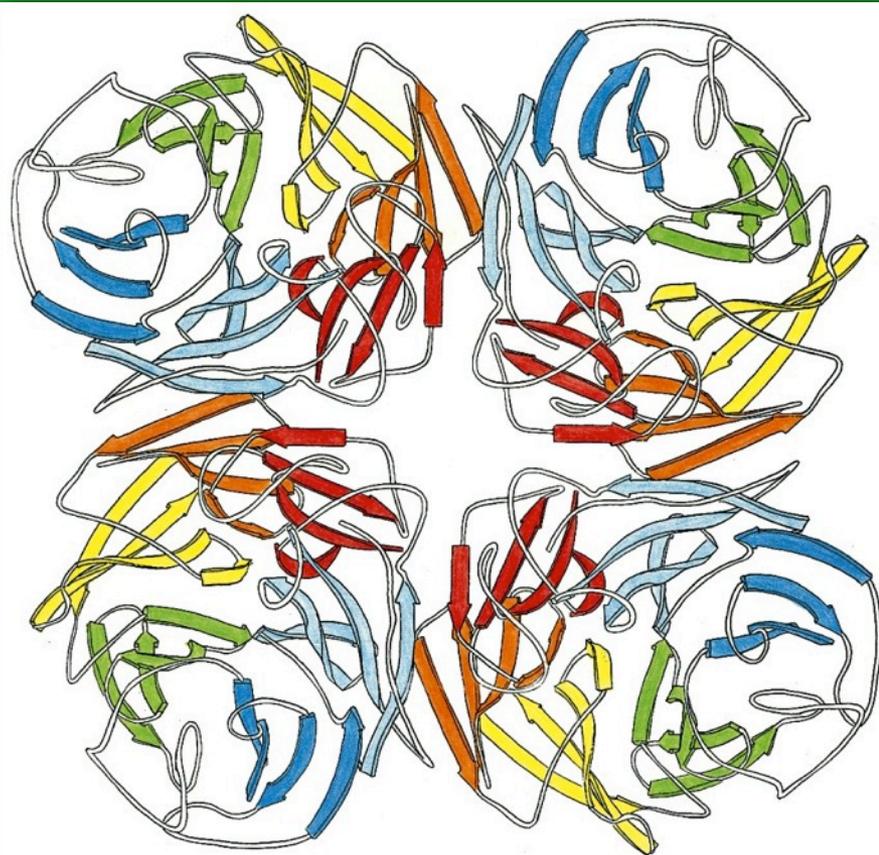
Multi-subunit (oligomeric) proteins
4° structure (quaternary structure)



The structural hierarchy in proteins: 4^o structure

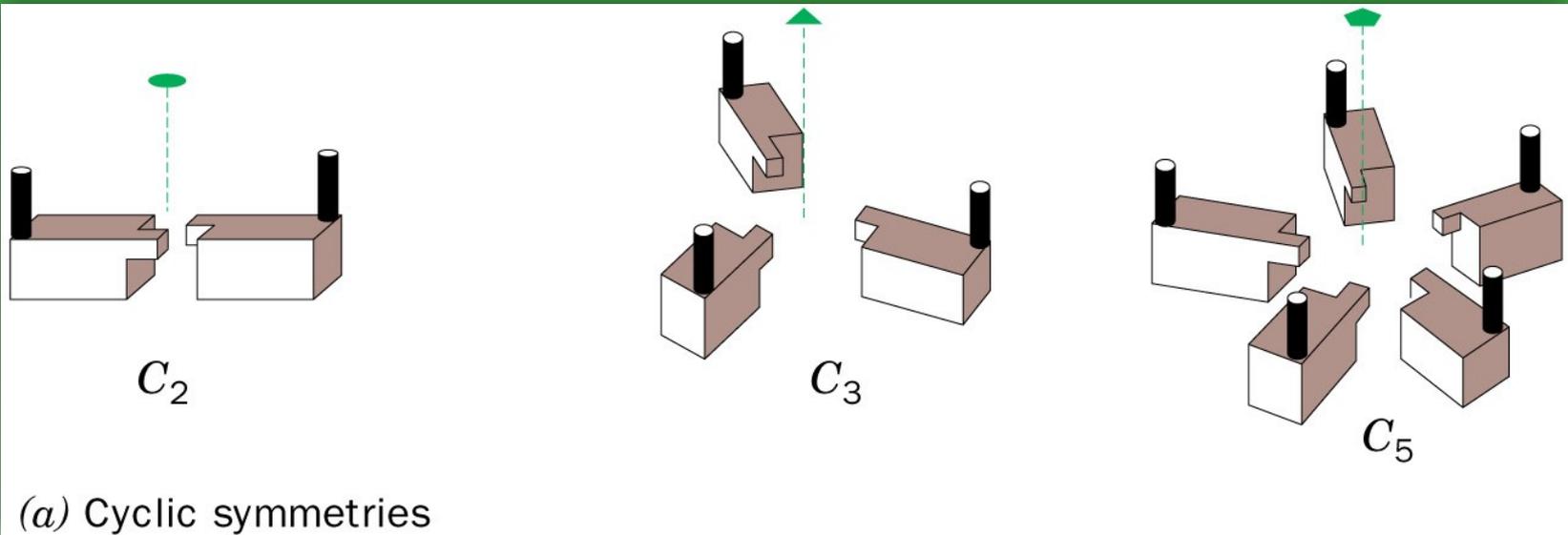


The quaternary structure of hemoglobin (tetramer; two α and two β polypeptides; four O_2 binding sites/hemoglobin tetramer); a binding protein (no catalytic activity)

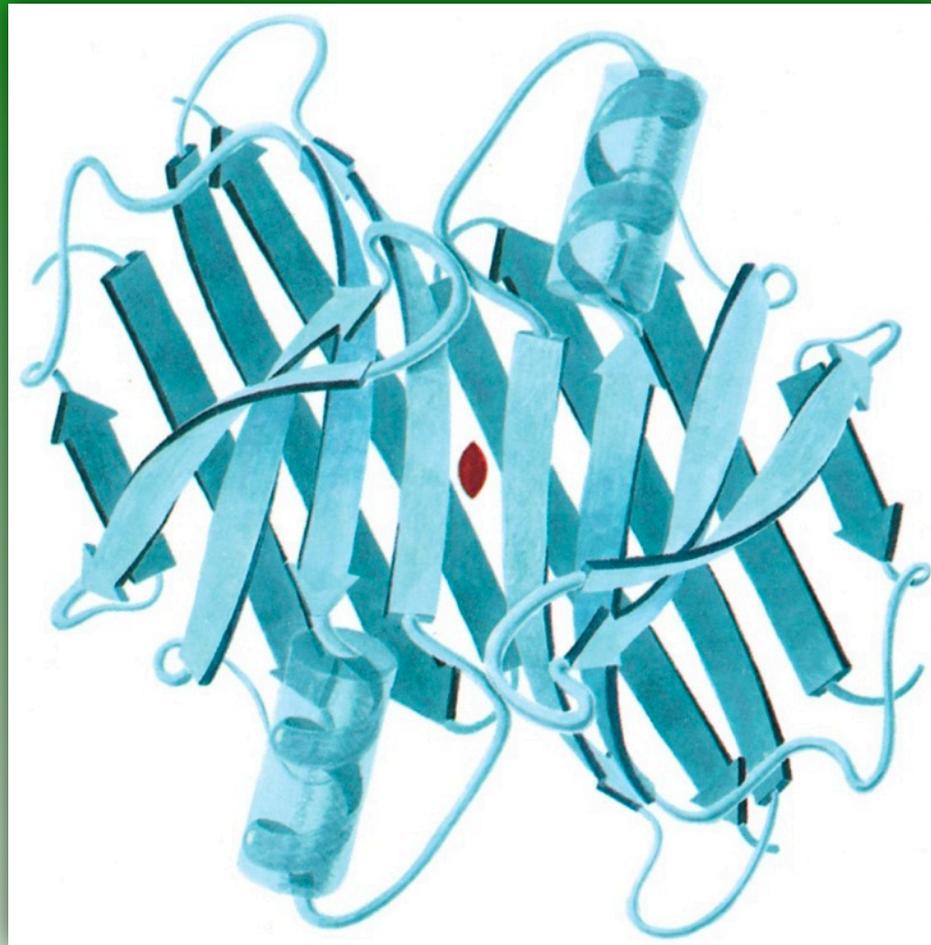


tetramer of neuraminidase protein

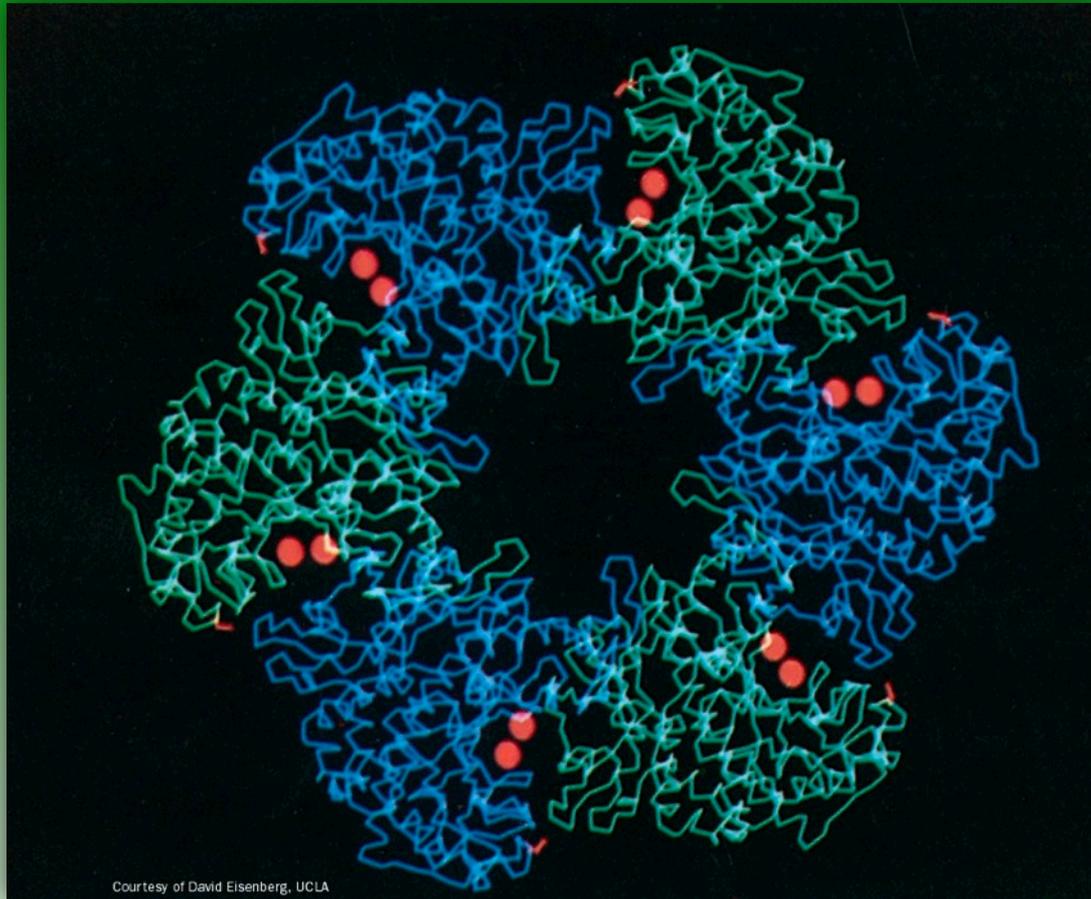
Neuraminidase:
A tetrameric enzyme
comprised of four
identical subunits



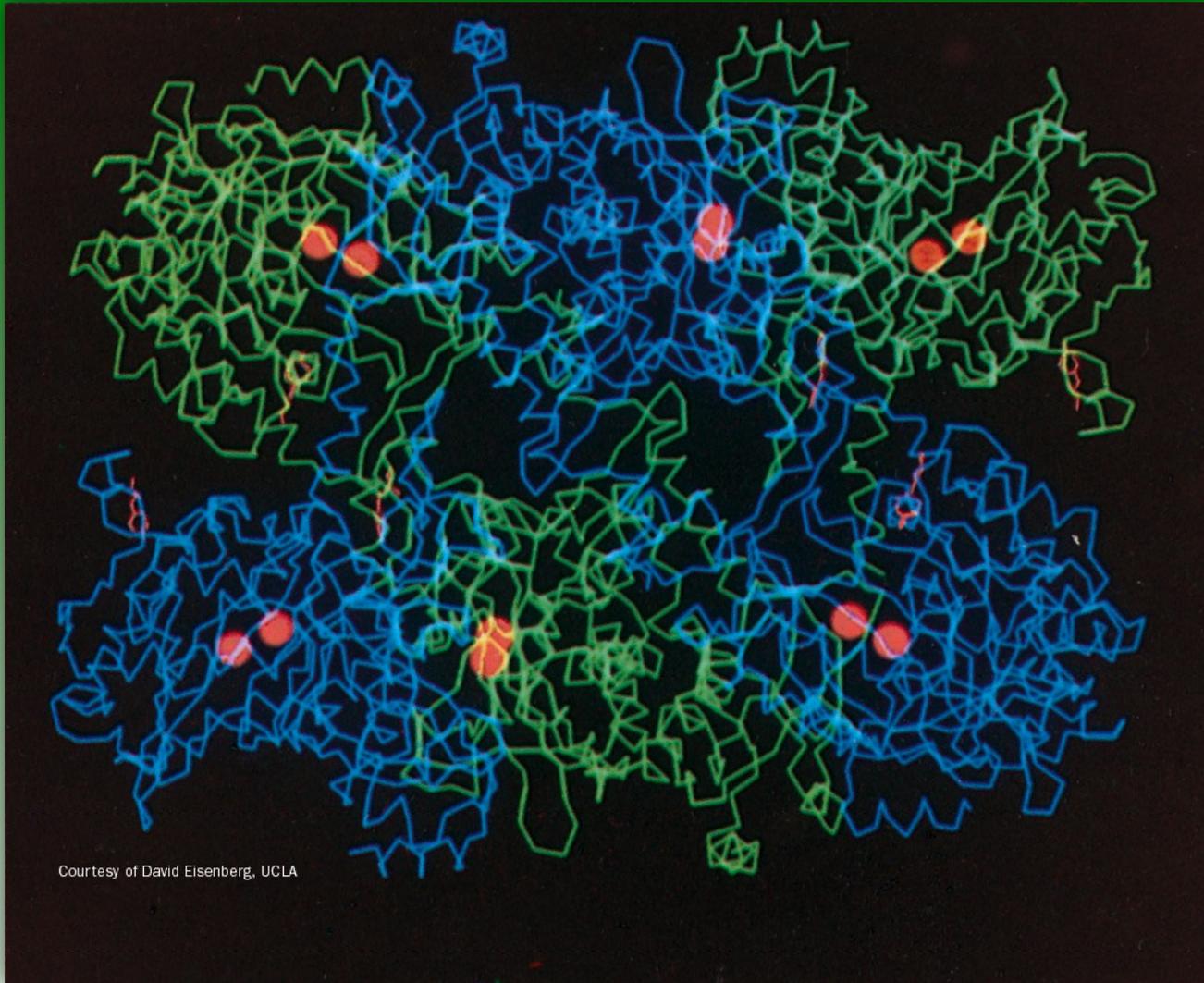
Some possible symmetries of proteins with identical protomers. Assemblies with symmetries C_2 , C_3 , and C_5 .



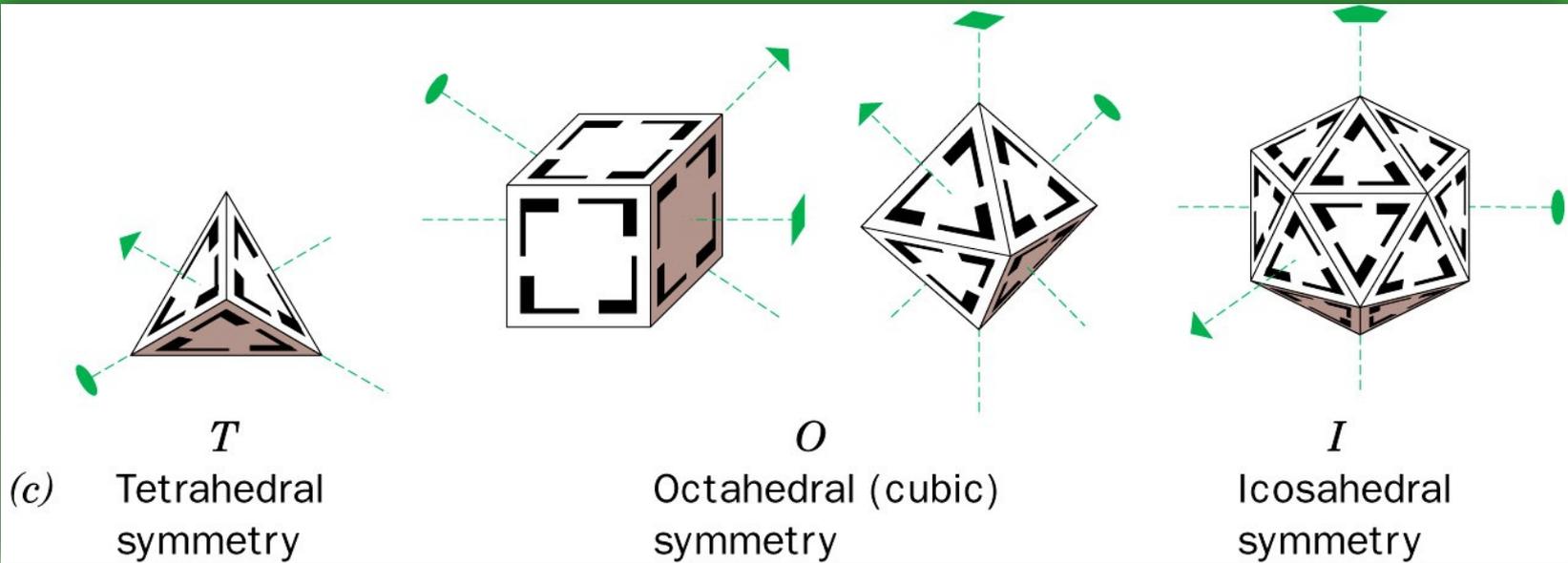
A **dimer** of transthyretin as viewed down its two-fold axis (*red symbol*).



X-ray structure of glutamine synthetase from *Salmonella typhimurium* - view down 6-fold symmetry axis
(contains 12 subunits)



X-ray structure of glutamine synthetase from *Salmonella typhimurium* - view down one of the 2-fold symmetry axes



Some possible symmetries of proteins with identical protomers. Assemblies with *T*, *O*, and *I* symmetries.

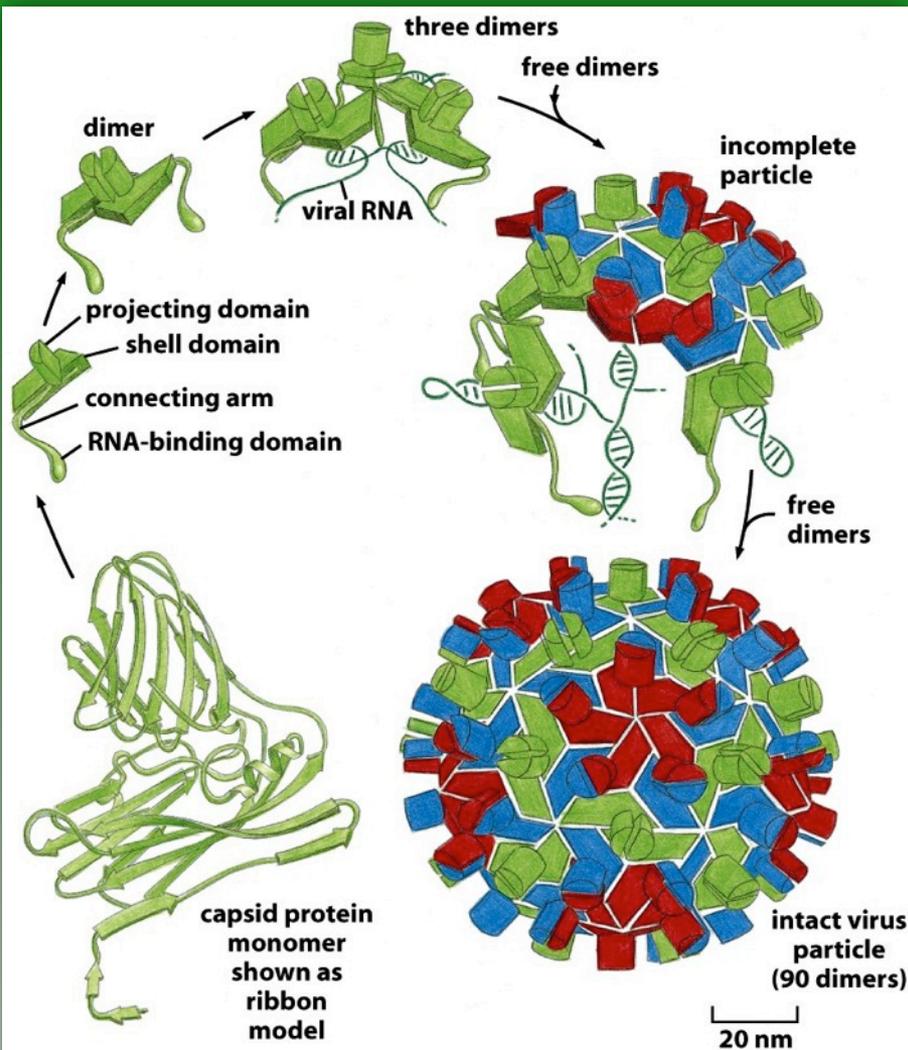


Figure 3-31 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Assembly of a spherical virus (capsid formation): the tomato bushy stunt virus (TBSV).

33 nm in diameter

180 copies of a 386 aa capsid protein subunit

RNA genome of 4500 nucleotides

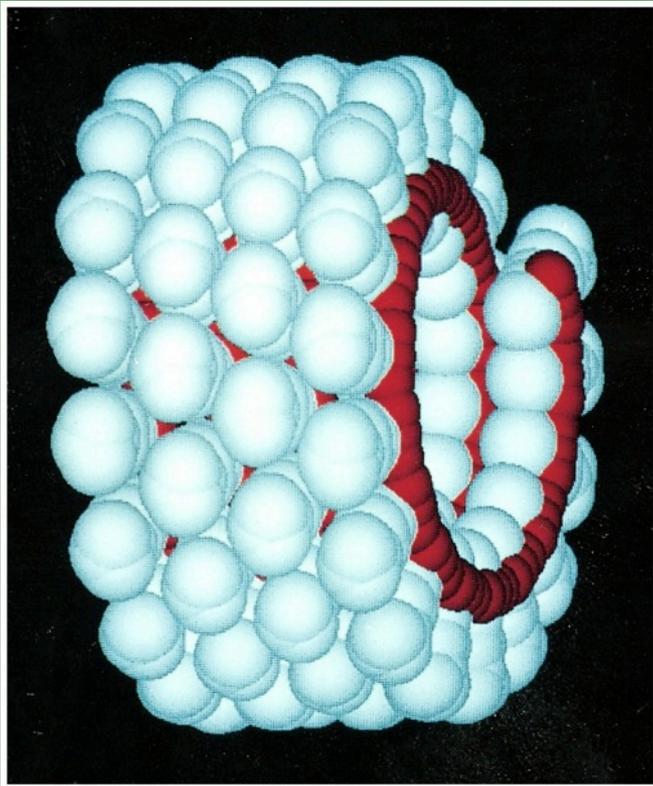


Figure 3-32b Molecular Biology of the Cell 5/e (© Garland Science 2008)

A structural representation of the tobacco mosaic virus (TMV):
A single long RNA molecule (6395 nucleotides) is enclosed in a cylindrical protein (helical coat) composed of 2130 identical protein subunits each containing 158 amino acids.

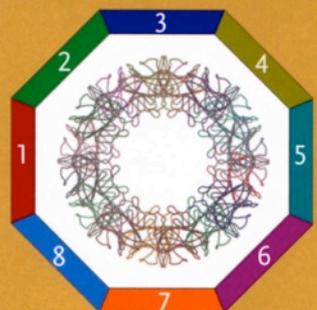
Legend

PDB code for reference structure

Number of subunits

1qtj

16



D₈

822

SAP from horseshoe crab

▲ Protein name ▲
 Point group symmetry (Schoenflies nomenclature) Point group symmetry (International nomenclature)

- 1 Multiple sites, cross-linking, membrane association
- 2 Cooperativity/allostherism
- 3 Cavities, channels and pores
- 4 Functional (active) site formation
- 5 Size and stability
- 6 Economy of genetic material
- 7 "Rulers" (exact separation between binding sites)
- 8 Multiple functions (in hetero-oligomers)

Yellow background indicates dihedral symmetry

Gray background indicates cubic/icosahedral symmetry

WILEY-VCH

Oligomeric Proteins

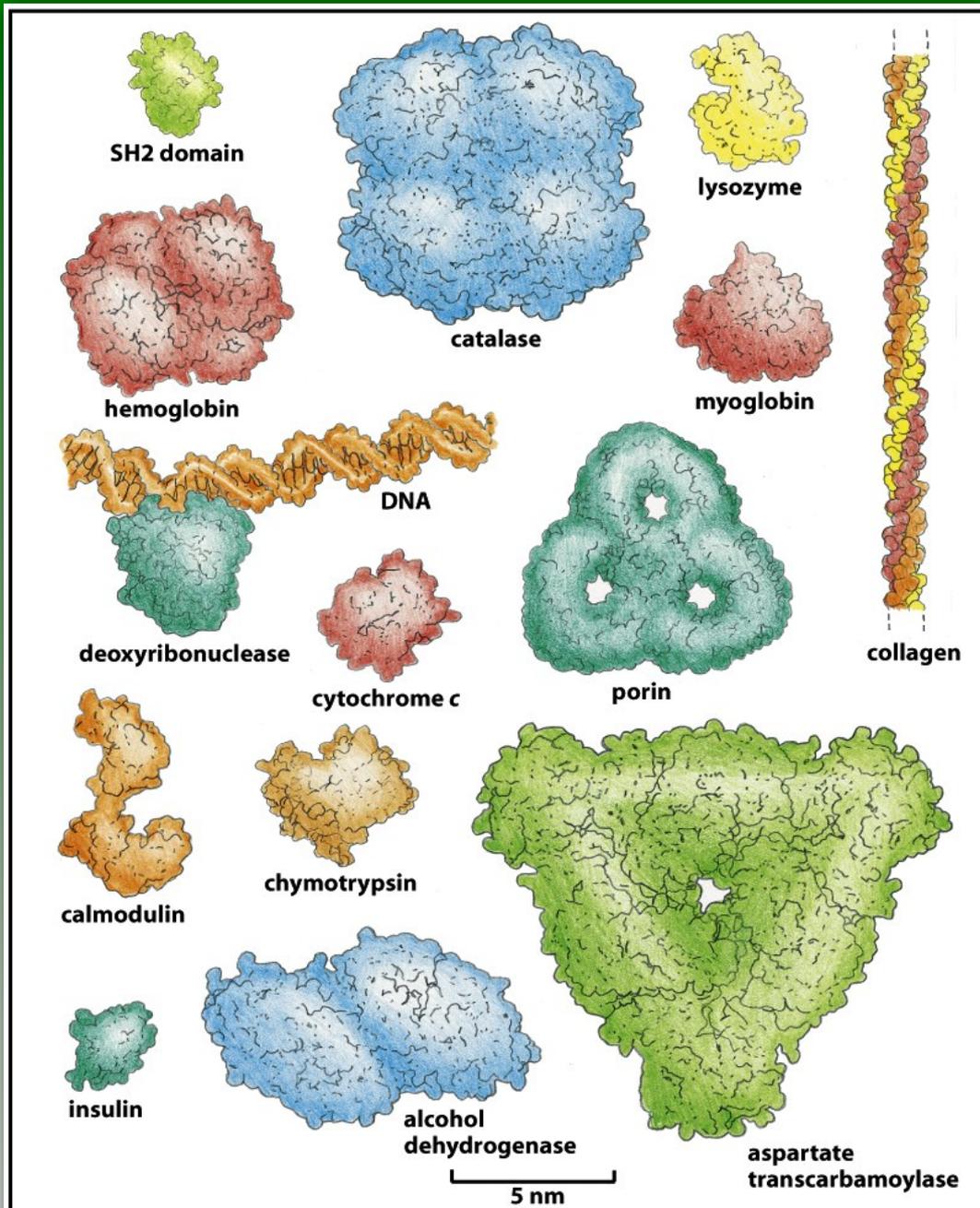
Highest Order Rotation Axis

1	2	3	4	5	6	7	> 7
<p>1a4d 2</p> <p>C₂ HLA class II 1</p>	<p>1wrp 2</p> <p>C₂ Trp repressor 2</p>	<p>1rtm 3</p> <p>C₃ C-type mannose binding protein 3</p>	<p>1bl8 4</p> <p>C₄ Potassium channel 4</p>	<p>1lts 5</p> <p>C₅ Heat labile enterotoxin (B subunit) 5</p>	<p>1do0 6</p> <p>C₆ HslU ATPase 6</p>	<p>18f 7</p> <p>C₇ SmAP 7</p>	<p>1lgh 16</p> <p>C₁₆ LHC II Rhodospirillum molischianum 8</p>
<p>3hvt 2</p> <p>C₂ HIV reverse transcriptase 1</p>	<p>2pol 2</p> <p>C₂ Bacterial polymerase III beta subunit (E. coli) 2</p>	<p>1cd5 6</p> <p>D₆ Glucosamine 6-P deaminase 32</p>	<p>1cuk 4</p> <p>C₄ RUVA (DNA recombination protein) 4</p>	<p>1msl 5</p> <p>C₅ Mechanosensitive channel 5</p>	<p>1g8y 6</p> <p>C₆ Helicase RepA of plasmid RSF 1010 6</p>	<p>7ahl 7</p> <p>C₇ Alpha-hemolysin 7</p>	<p>1nkz 18</p> <p>C₁₈ LHC II Rhodospseudomonas acidophila 9</p>
<p>1hzh 4</p> <p>C₂ IgG 1</p>	<p>3phv 2</p> <p>C₂ HIV protease 2</p>	<p>1raa 12</p> <p>D₁₂ Aspartate transcarbamoylase 32</p>	<p>1dhn 8</p> <p>D₈ 7,8-Dihydroneopterin aldolase 422</p>	<p>1gtp 10</p> <p>D₁₀ GTP cyclohydrolase I 52</p>	<p>1y12 6</p> <p>C₆ Protein secretion apparatus (HCP1) 6</p>	<p>1grl 14</p> <p>D₁₄ GroEL 72</p>	<p>1wap 11</p> <p>C₁₁ TRP RNA binding attenuation protein (TRAP) 11</p>
<p>2aal 2</p> <p>C₂ Ricin 1</p>	<p>4hhb 4</p> <p>C₂ Human haemoglobin 2</p>	<p>1dps 12</p> <p>T₁₂ DNA binding protein Dps 23</p>	<p>1a6d 16</p> <p>D₁₆ Thermosome 422</p>	<p>1rvw 60</p> <p>I₆₀ Icosahedral lumazine synthase (B. subtilis) 532</p>	<p>1f52 12</p> <p>D₁₂ Glutamine synthase 622</p>	<p>1pma 28</p> <p>D₂₈ Proteasome 72</p>	<p>1qtj 16</p> <p>D₁₆ SAP from horseshoe crab 822</p>
<p>4pfk 4</p> <p>D₄ Phosphofructokinase 222</p>	<p>3pcg 24</p> <p>T₂₄ Protocatechuate 3,4-dioxygenase 23</p>	<p>2fha 24</p> <p>O₂₄ Ferritin 432</p>	<p>1stm 60</p> <p>I₆₀ Satellite panicum mosaic virus 532</p>	<p>1g3k 24</p> <p>D₂₄ HslUV ATP dependent protease 622</p>			

ISBN 978-3-527-31963-3

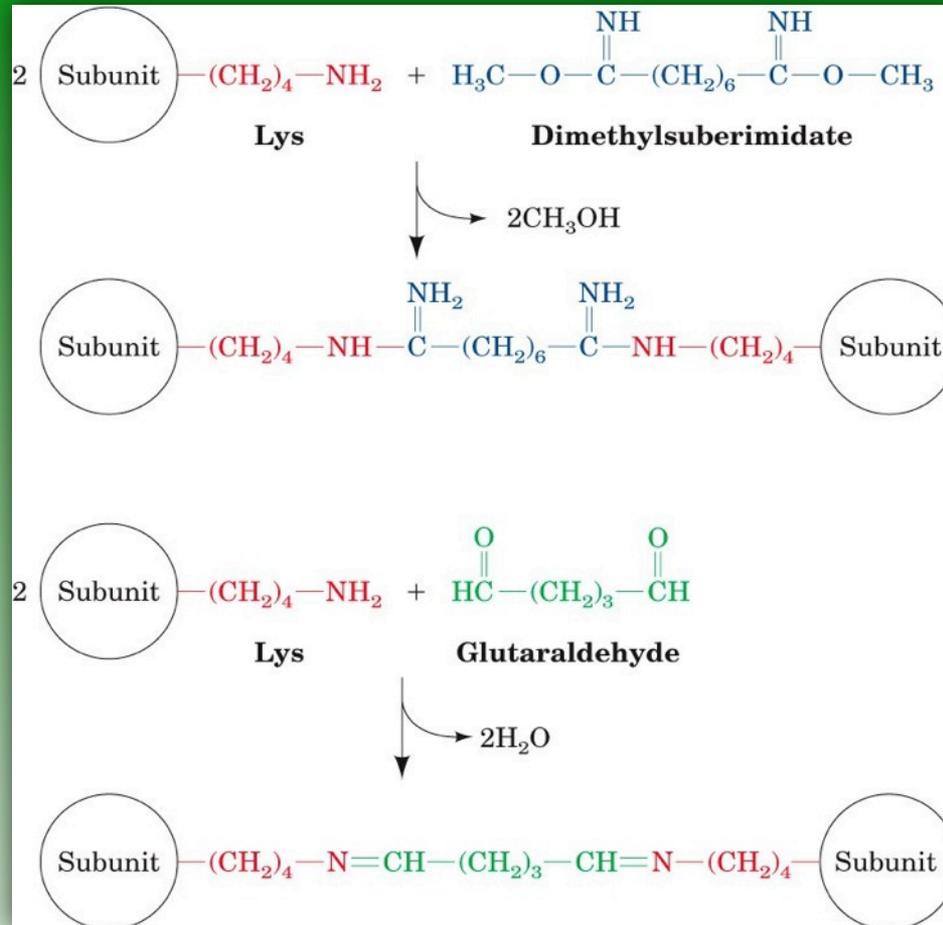


9 783527 319633



A collection of proteins shown on the same scale

Figure 3-23 Molecular Biology of the Cell 5/e (© Garland Science 2008)



Characterization of oligomeric proteins: Some chemical cross-linking agents (bifunctional reagents) used to stabilize protein quaternary structure

Protein-protein interactions (binding): Three common ways in which two proteins can bind to each other.

