BCHS 6229
Protein Structure and Function

Lecture 2 (October 13, 2011)

Protein Architecture:

Symmetry relationships and protein structure

Primary & Secondary Structure

Motifs & Super-secondary Structure

Tertiary Structure & Fold type

Configuration and Conformation of Molecules

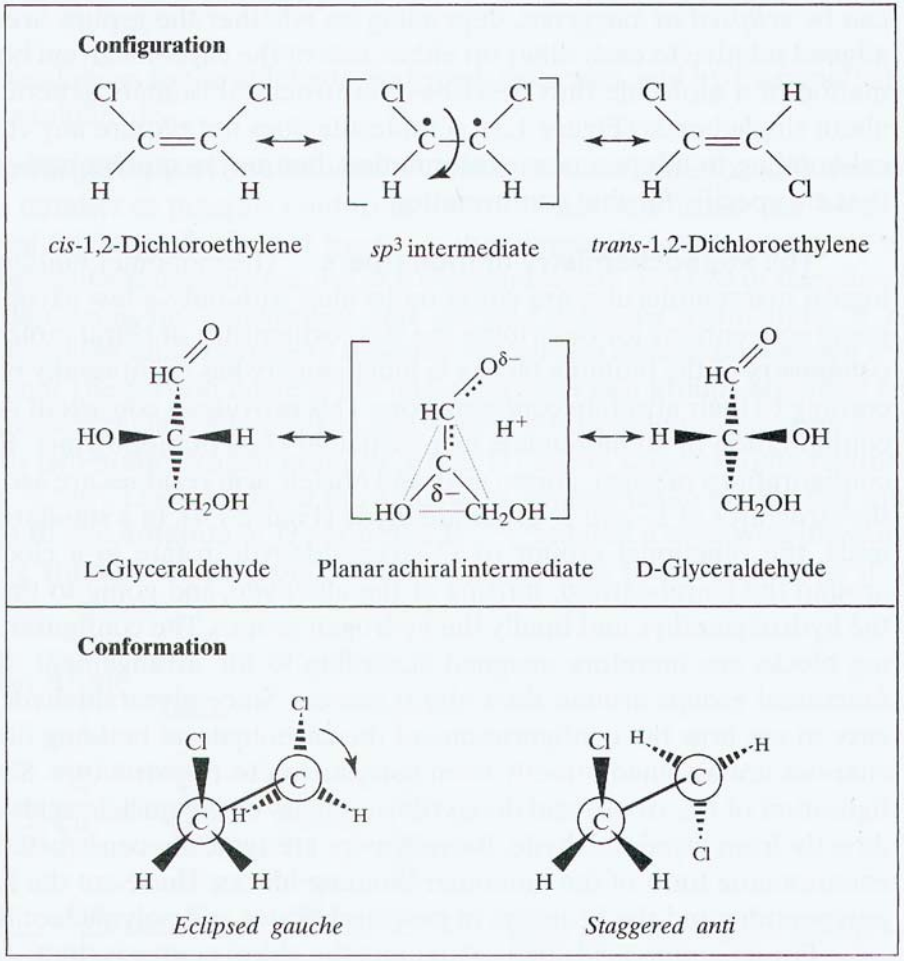


Figure 1.3 Configuration and conformation both describe the geometry of a molecule. The configuration of a molecule can be changed only by breaking and remaking chemical bonds, as in the conversion of a *cis*-double bond to one that is in the *trans*-configuration, or in converting from the L- to the D-stereoisomer of a chiral molecule. Conformations can be changed by simple rotations about a single bond.

Configuration: position of groups

Conformation: spatial arrangement about one or more freely rotating bonds

Conformation of Molecules

is described as a torsion angle about each (*freely*) rotating bond. *Very important in protein structure determination.*

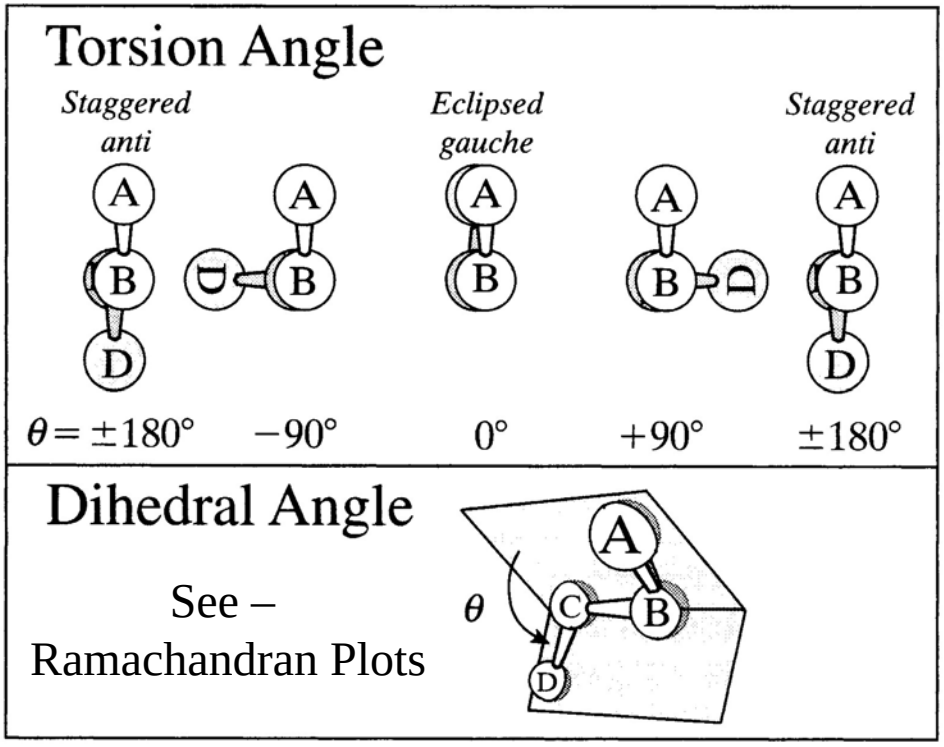


Figure 1.5 Torsion angles and dihedral angles (θ). The rotation around a single bond is described by the torsion angle of the four atoms around the bond (A—B—C—D) and the dihedral angle θ relating the planes defined by atoms A—B—C and by B—C—D.

Three Dimensional Protein Structures

Conformation: Spatial arrangement of atoms that depend on bonds and bond rotations.

Proteins can change conformation, however, most proteins have a stable "native" conformation.

The native protein is folded through weak interactions:

- **a) Hydrophobic interactions**
- **b) Hydrogen-bonds**
- **c) Ionic interactions**
- **d) Van der Waals attractions**

Distance Dependence of Noncovalent Interactions

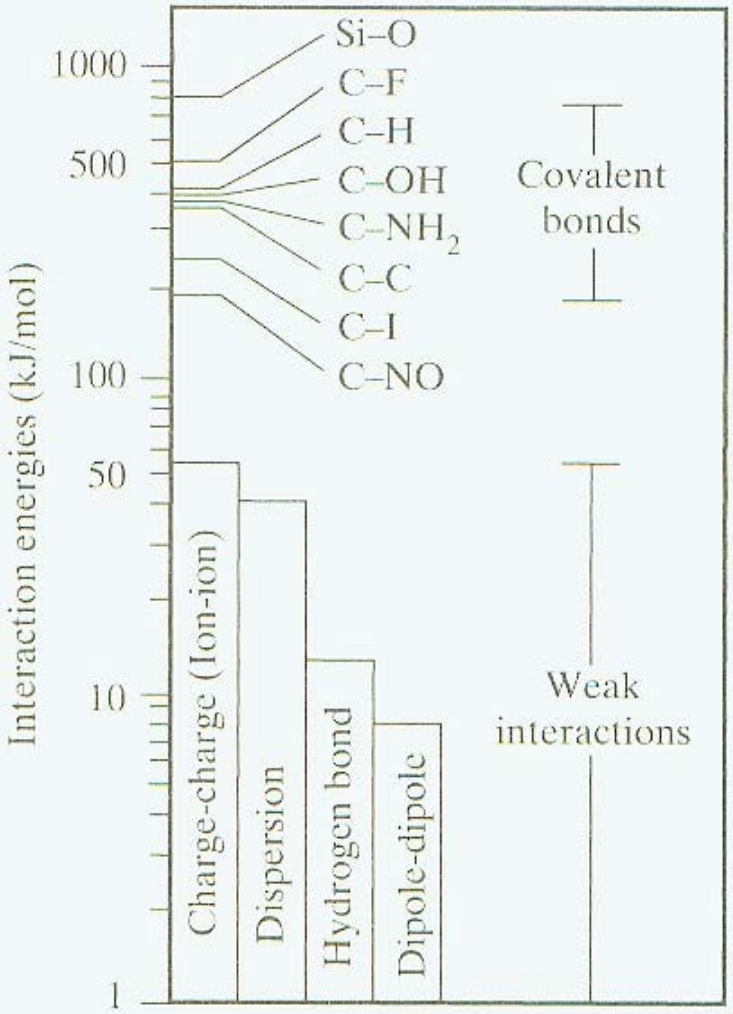
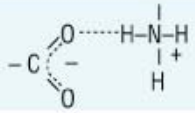
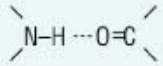
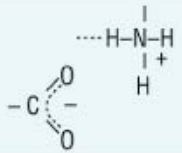
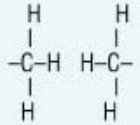


Figure 1.6 Energies of molecular interactions. The interactions that define the structure of a molecule range from the strong interactions of covalent bonds (200 to 800 kJ/mol) to the weak charge-charge (or ion-ion), dipole-dipole, dispersion, and hydrogen-bonding interactions (0 to 60 kJ/mol).

Table 1.1 Relationship of Noncovalent Interactions to the Distance Separating the Interacting Molecules, r

Type of Interaction	Distance Relationship
Charge-charge	$1/r$
Charge-dipole	$1/r^2$
Dipole-dipole	$1/r^3$
Charge-induced dipole	$1/r^4$
Dispersion	$1/r^6$
Repulsion	$1/r^{12}$

Chemical Interactions that Stabilize Polypeptides

Interaction	Example	Distance dependence	Typical distance	Free energy (bond dissociation enthalpies for the covalent bonds)
Covalent bond	$-C_\alpha-C-$	-	1.5 Å	356 kJ/mole (610 kJ/mole for a C=C bond)
Disulfide bond	–Cys–S–S–Cys–	-	2.2 Å	167 kJ/mole
Salt bridge		Donor (here N), and acceptor (here O) atoms <3.5 Å	2.8 Å	12.5–17 kJ/mole; may be as high as 30 kJ/mole for fully or partially buried salt bridges (see text), less if the salt bridge is external
Hydrogen bond		Donor (here N), and acceptor (here O) atoms <3.5 Å	3.0 Å	2–6 kJ/mole in water; 12.5–21 kJ/mole if either donor or acceptor is charged
Long-range electrostatic interaction		Depends on dielectric constant of medium. Screened by water. 1/r dependence	Variable	Depends on distance and environment. Can be very strong in nonpolar region but very weak in water
Van der Waals interaction		Short range. Falls off rapidly beyond 4 Å separation. $1/r^6$ dependence	3.5 Å	4 kJ/mole (4–17 in protein interior) depending on the size of the group (for comparison, the average thermal energy of molecules at room temperature is 2.5 kJ/mole)

Hydrogen bond donors and acceptors (angle-dependence)

Table 1.3 Hydrogen-Bond Donors and Acceptors in Macromolecules

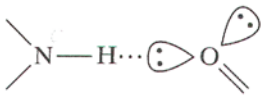
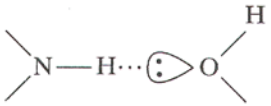
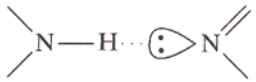
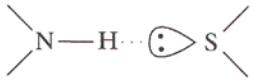
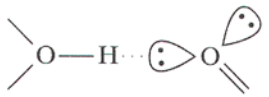
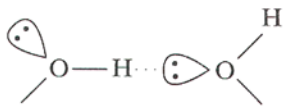
Donor	Acceptor	*r* (nm)
	r	
		0.29
		0.29
		0.31
		0.37
		0.28
		0.28

Table 1.2 Electronegativities of Elements Typically Found in Biological Molecules

Element	Electronegativity
O	3.5
Cl	3.0
N	3.0
S	2.5
C	2.5
P	2.1
H	2.1
Cu^{2+}	1.9
Fe^{2+}	1.8
Co^{2+}	1.8
Mg^{2+}	1.2
Ca^{2+}	1.0
Na^{+}	0.9
K^{+}	0.8

Higher values indicate a higher electron affinity.

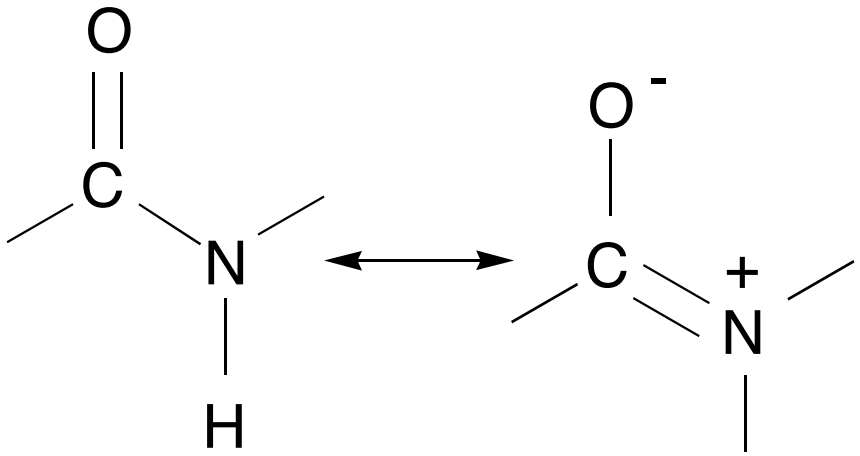
Key Concepts:

- Noncovalent bonds play important roles in determining the physical and chemical properties of water. They also have a significant effect on the structure and function of biomolecules.

- H-bonding is responsible for water's high freezing and boiling points. Because water has a high heat capacity, it can absorb and release heat slowly. Water plays an important role in regulating heat in living organisms.

The Peptide bond

In 1930s-1940s Linus Pauling and Robert Corey determined the structure of the peptide bond by X-ray.

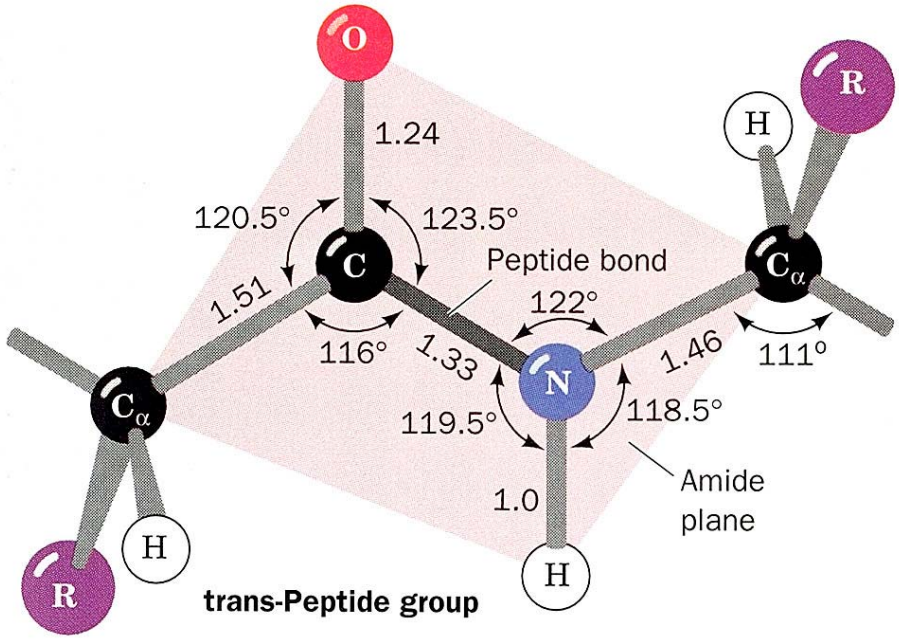


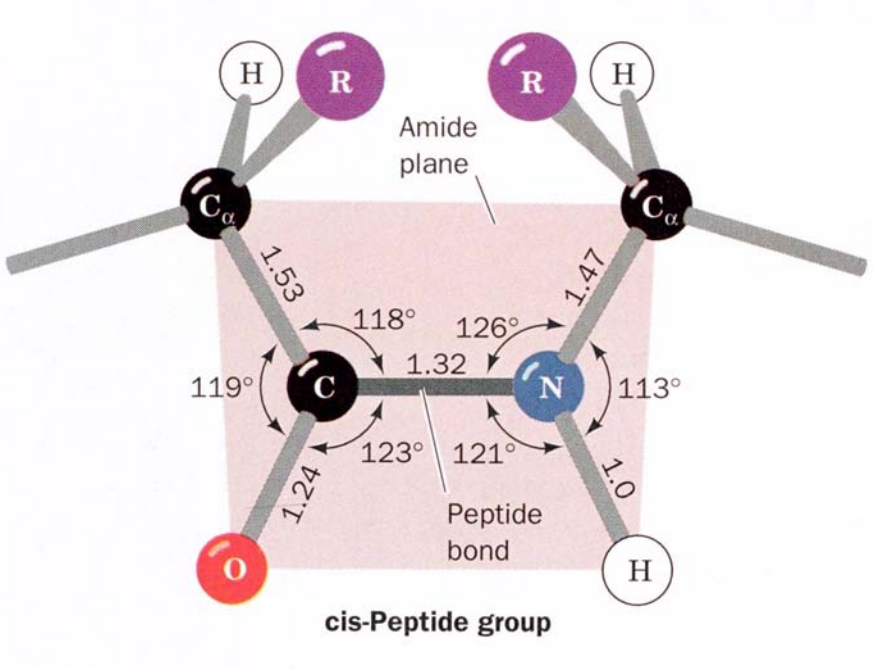
40% double bond character. The amide bond or peptide bond C-N bond is 0.13Å shorter than C_α-N bond.
C=O bond is 0.02 Å longer than that of ketones and aldehydes.
Planar conformation maximizes π-bonding overlap.
Resonance gives 85 kJ/mol stability when bond is planar!!

Peptide bonds are planar

Resonance energy depends on dihedral/torsional angle (Cα-C-N-Cα).
For peptides, this is the angle between the Cα-C and N-Cα bonds.
For a *trans* peptide bond, the dihedral angle is 180° by definition.
In a *cis* peptide bond, the dihedral angle is 0° by definition.
Most peptide bonds are trans, 10% that follow proline may be cis.

Note: differences between bond angles and bond lengths comparing *cis* and *trans* forms of a generic dipeptide.





Peptide Torsion angles

Torsion angles determine flexibility of backbone structure

Rotation or dihedral angles

Cα-N	**ϕ**	**phi**
Cα-C	**ψ**	**psi**

When a peptide chain is fully extended the angles are defined as 180° or -180° (these are the same).
At 180°, one gets a staggered conformation - (all trans) i.e. ethane.

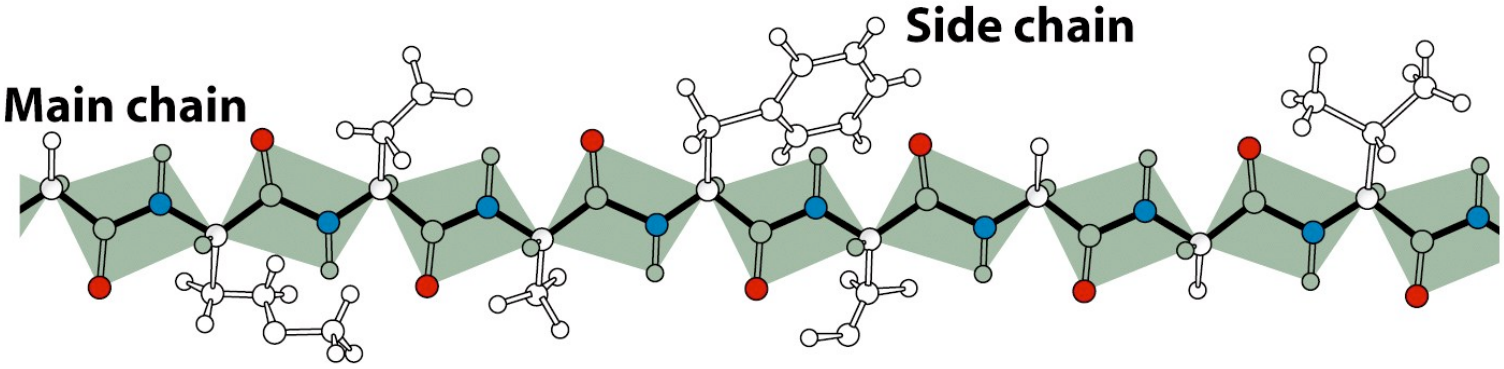
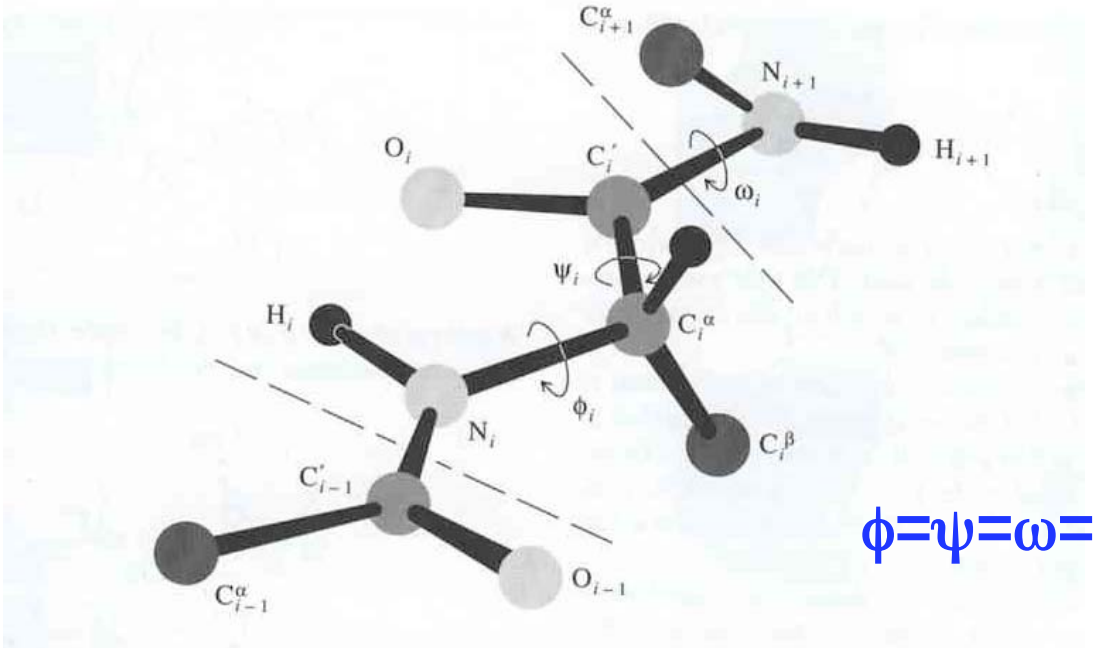


Figure 6-3 Fundamentals of Biochemistry, 2/e

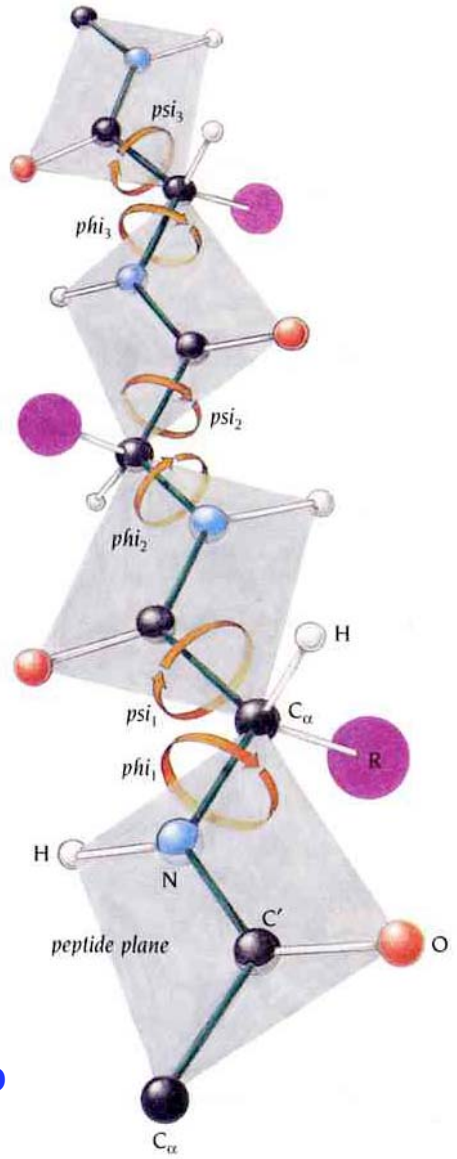
Note: alternating C=O pointing in opposite directions.

When viewed down the Cα-N axis, rotation to the right or clock wise increases the angle of rotation.

Must start with the fully extended form which is defined as 180° or -180°



$\phi=\psi=\omega=180^{o}$



A key to nomenclature for the atoms of the polypeptide chain and backbone dihedral angles ϕ, ψ, and ω.

If all ϕ + ψ angles are defined then the backbone structure of a protein will be known!!

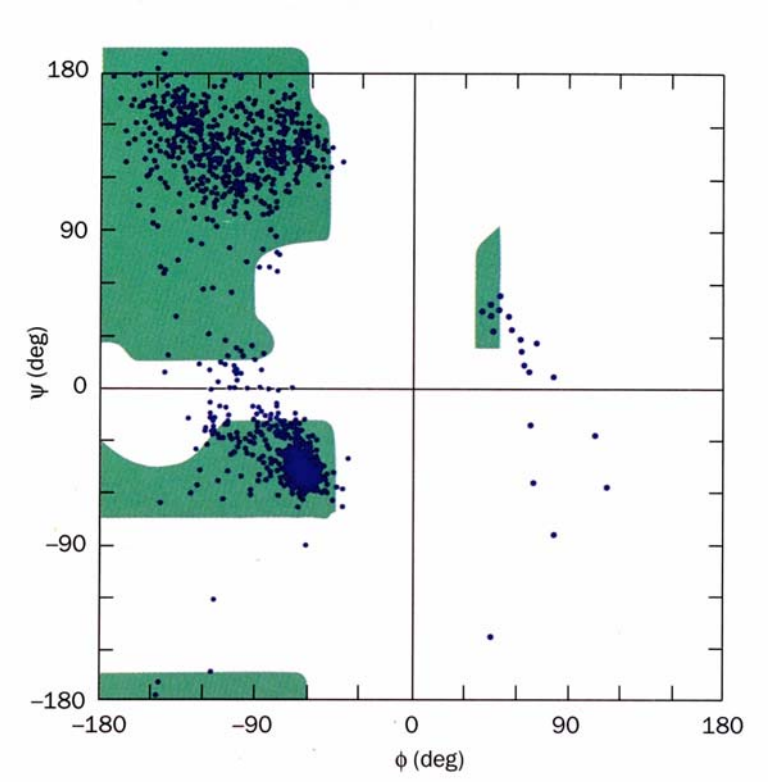
These angles allow a method to describe the protein's structure and all backbone atoms can be placed in a 3D-grid with an X, Y, Z coordinates.

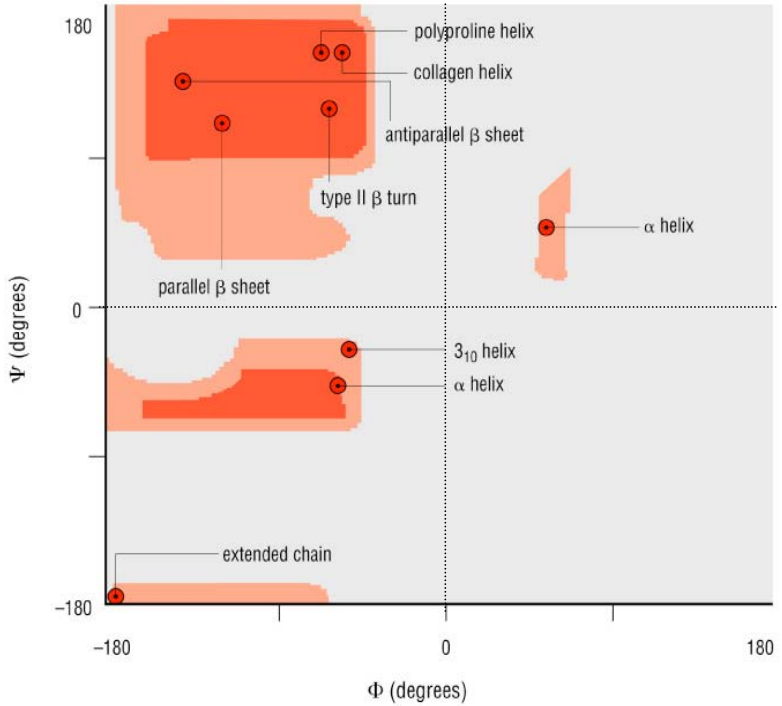
The physical size of atoms and groups of atoms limits the possible ϕ and ψ torsion angles that the backbone of a polypeptide chain can adopt without causing protruding groups like the carbonyl and side chains to bump into each other. Allowed values ▸ Ramachandran plot

Ramachandran diagram

If you plot ψ on the Y-axis and ϕ on the X-axis, you will plot all possible combinations of ϕ, ψ.

You must know the different regions of the Ramachandran diagram.





Protein Structure Facts

- The interior of proteins is hydrophobic.
- The α-helix/β-sheets are important elements of secondary structure.
- The α-helix has a dipole moment.
- α-helices can be amphipathic.
- Some amino acids are preferred in α-helices.
- β-sheets usually have their β strands parallel or anti-parallel.
- Amphipathic β-sheets are found on the surface of proteins.
- Topology diagram are useful for classification of protein structures.
- Large polypeptide chains fold into several domains.

Symmetry Relationships of Molecules

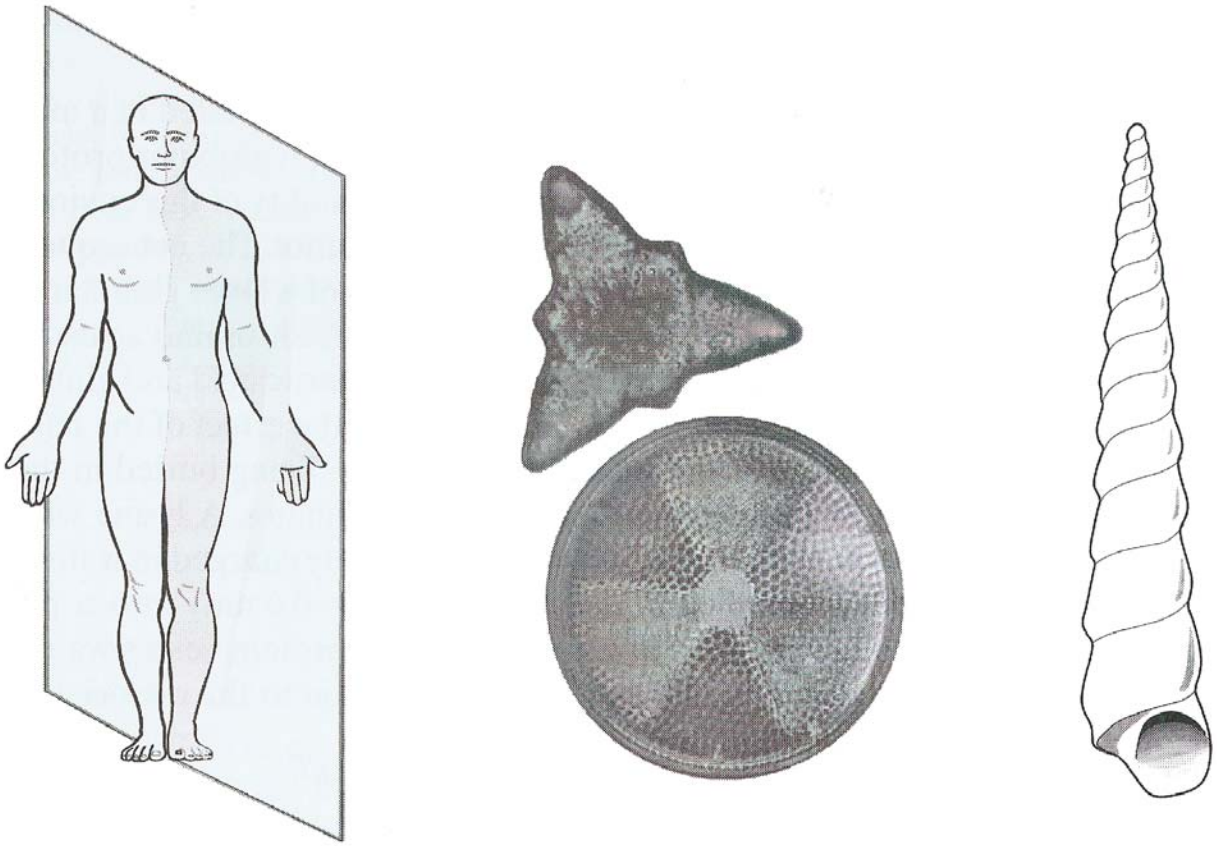


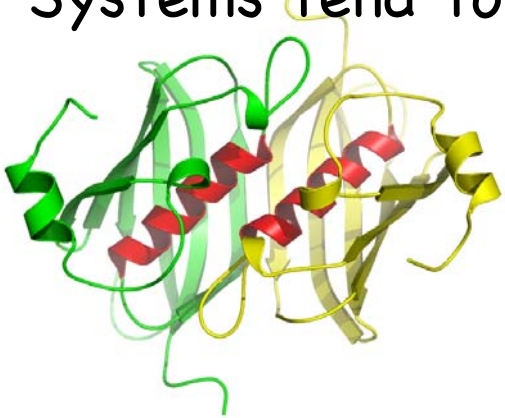
Figure 1.13 Examples of mirror, rotational, and screw symmetry. The human body shows mirror symmetry through a plane, diatoms show rotational symmetry about an axis, and a spiral shell shows screw symmetry about an axis.

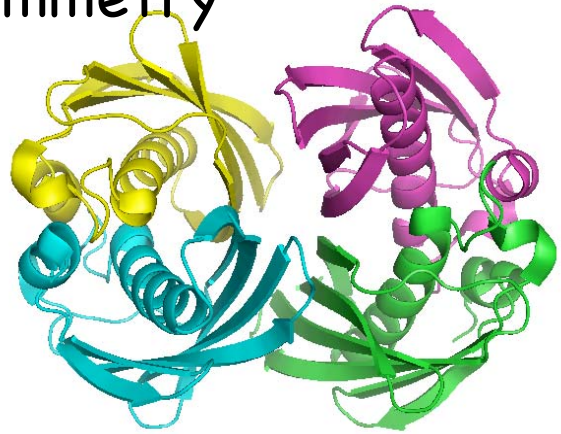
(Fig from van Holde, Johnson & Ho, Principles of Physical Biochem, 2005)

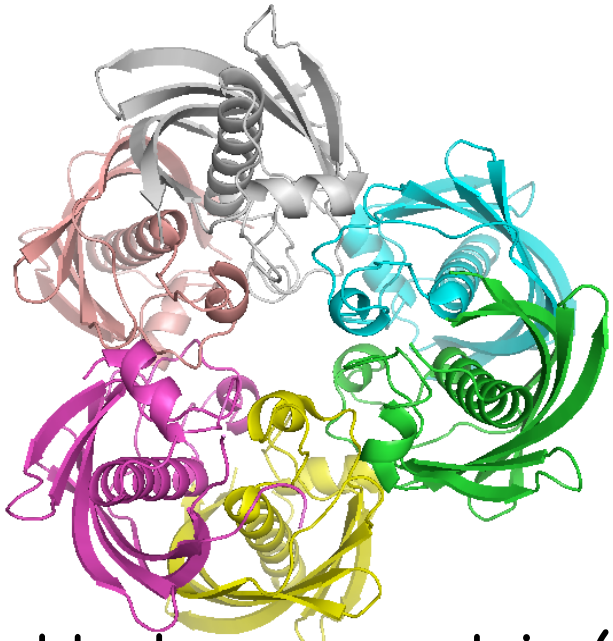
Exact correspondence of form and constituent configuration on opposite sides of a dividing line, mirror plane, about the center of an axis

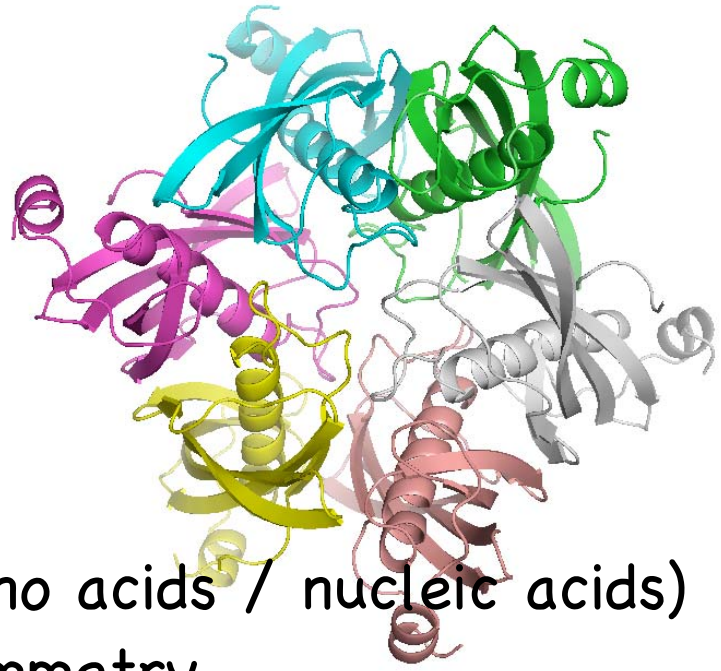
Monomer Building Blocks are Asymmetric

- Biological Systems tend to have symmetry



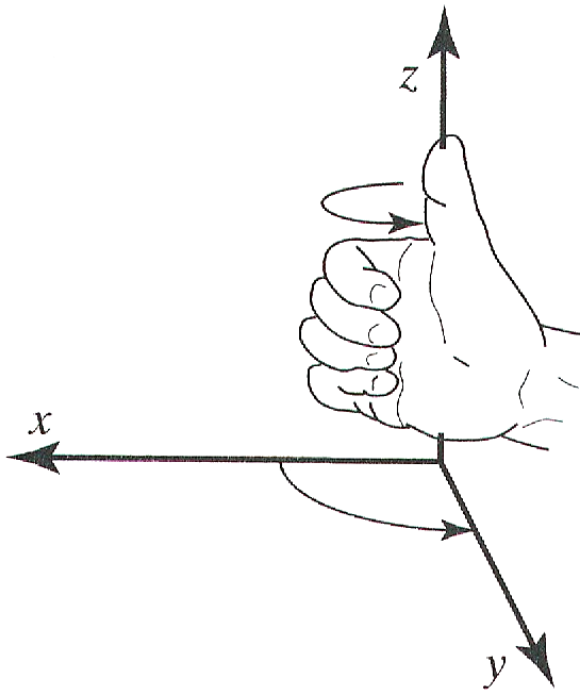






- Building blocks asymmetric (amino acids / nucleic acids)
- Pseudosymmetry – apparent symmetry

Conventions



- Right-handed Cartesian Coordinates and right-handed rotations

Relationship of motif (*m*) and related motif (*m'*) defined as:

$$\hat{O}(m) = m'$$

where (*m*) is a unique object and the symmetry Operator is ($\hat{O}$)

Mirror Symmetry

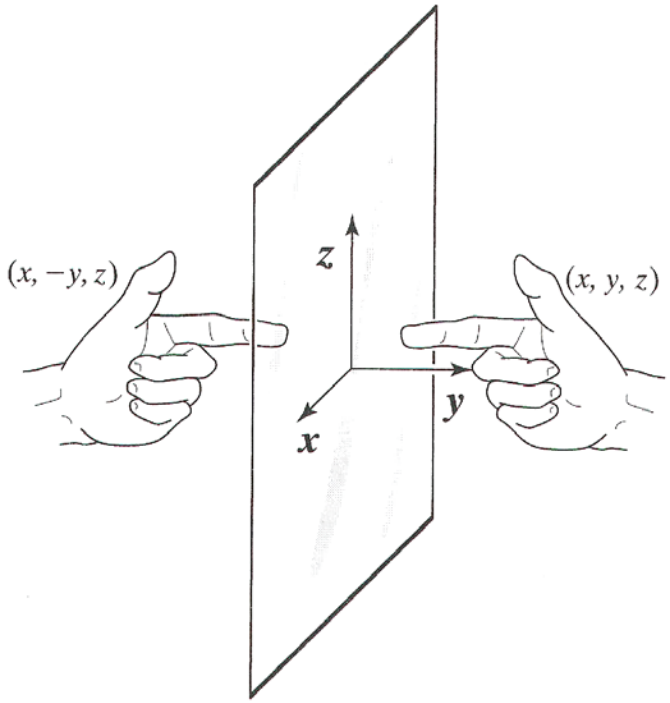


Figure 1.15 Mirror symmetry of left and right hands. The left and right hands are related by mirror symmetry through a plane. In this axis system, the two hands are related by an inversion of the y coordinate through the xz plane.

Simultaneous equations:

$$\begin{aligned} a_1x + b_1y + c_1z &= x' \\ a_2x + b_2y + c_2z &= y' \\ a_3x + b_3y + c_3z &= z' \end{aligned}$$

Matrix form

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \times \begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} x' \\ y' \\ z' \end{vmatrix}$$

Where the matrix dot product is

$$\begin{aligned} 1 \times x + 0 \times y + 0 \times z &= x' \\ 0 \times x - 1 \times y + 0 \times z &= y' \\ 0 \times x + 0 \times y + 1 \times z &= z' \end{aligned}$$

Matrix form of $\hat{\imath}$

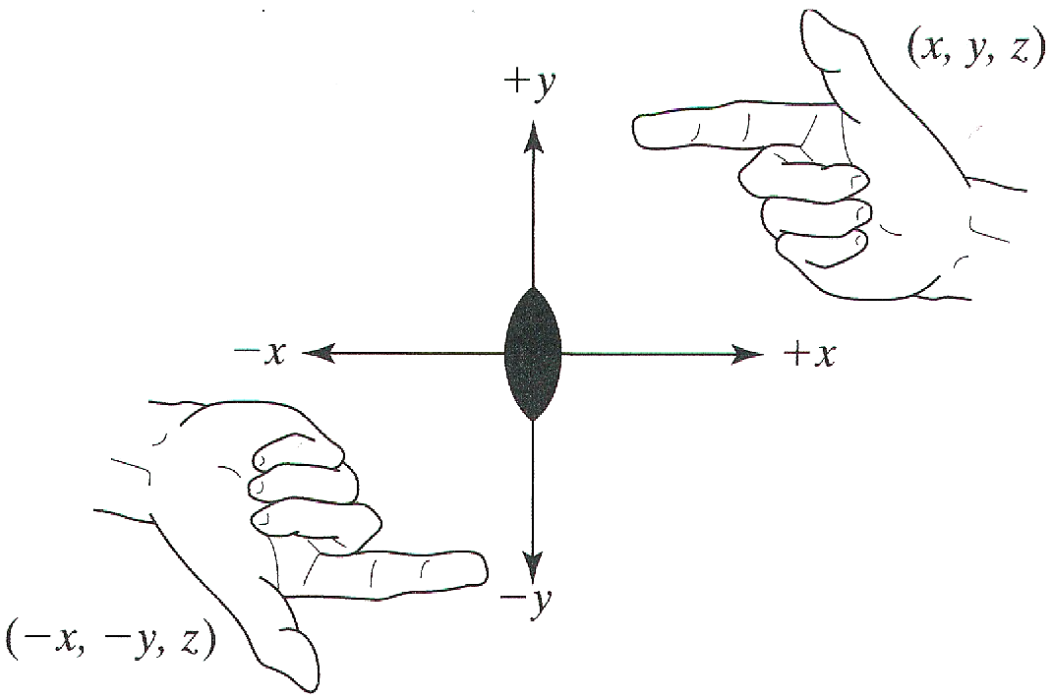
$$\hat{\imath} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

Point Group related through symmetry operator $\hat{\imath}$:

$$\hat{\imath}(x,y,z) = (x',y',z') = (x,-y,z)$$

Rotational Symmetry

Symmetry around a point or axis



Point Group related through two fold rotational axis C_2 through operator *ĉ:*

ĉ(x,y,z) = (x',y',z') = (-x,-y,z)

Matrix form of *ĉ*

$$\hat{c} = \begin{vmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

General operator for rotation about the z-axis by a rotation angle *Θ*

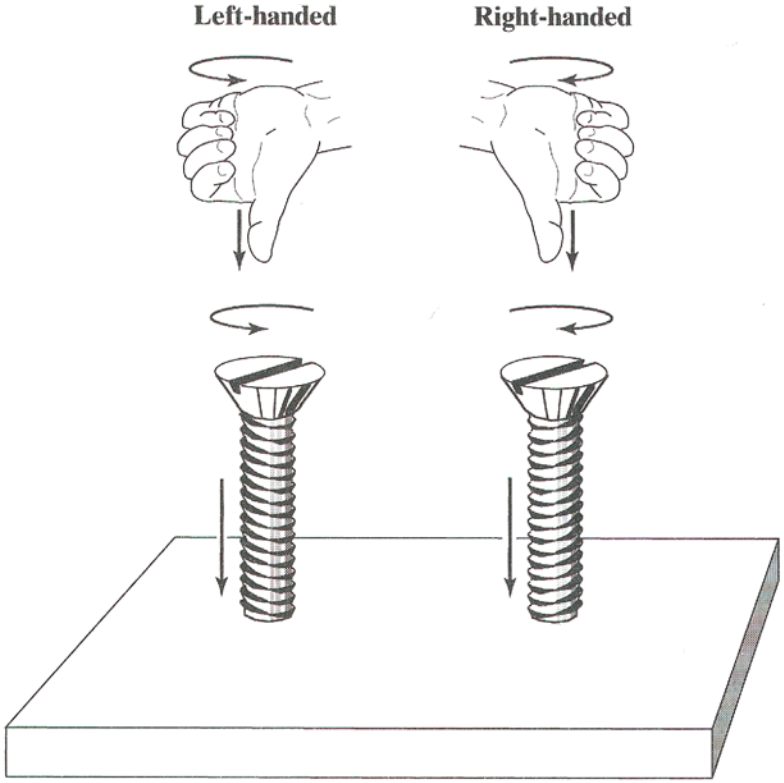
$$\begin{vmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

Solve for angle *Θ = 180°*

$$\begin{vmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} \times \begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} x' \\ y' \\ z' \end{vmatrix}$$

Screw Symmetry

Defined by rotational and translational motion as well as handedness.



Combine rotational $\hat{c}$ and translational T elements:

$$\hat{c}(x,y,z) + T = (x',y',z')$$

where the translation results from a θ of 360°

Translational Operator T acts without changing orientation:

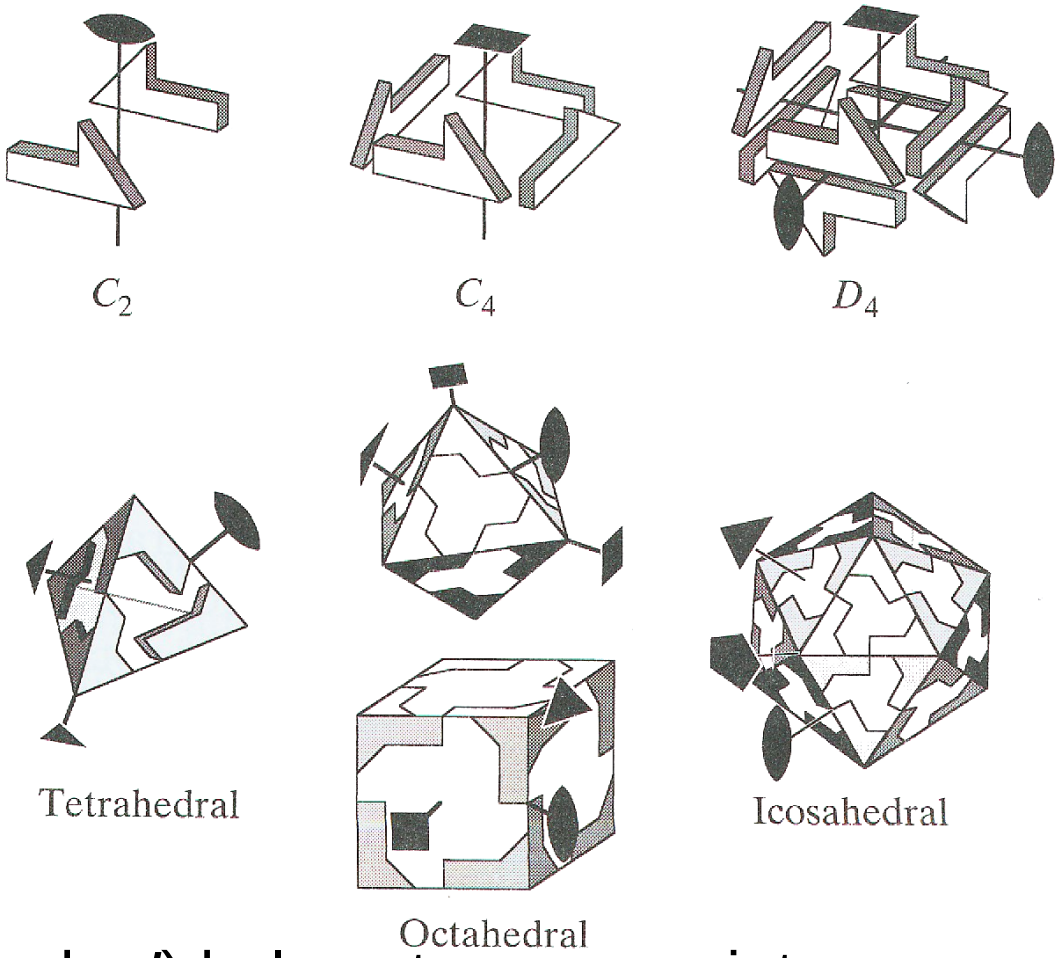
$$(x,y,z) + T = (x + Tx,\ y + Ty,\ z + Tz)$$

Where Tx, Ty, and Tz are the x, y, and z components of the translational operator

Multiple Symmetry Relationships and Point Groups

Rotational Symmetry Groups

(C) Central Point
(D) Dihedral
(T) Tetragonal
(O) Octahedral
(I) Icosahedral



Related motifs (e.g. *m* and *m'*) belong to same point group

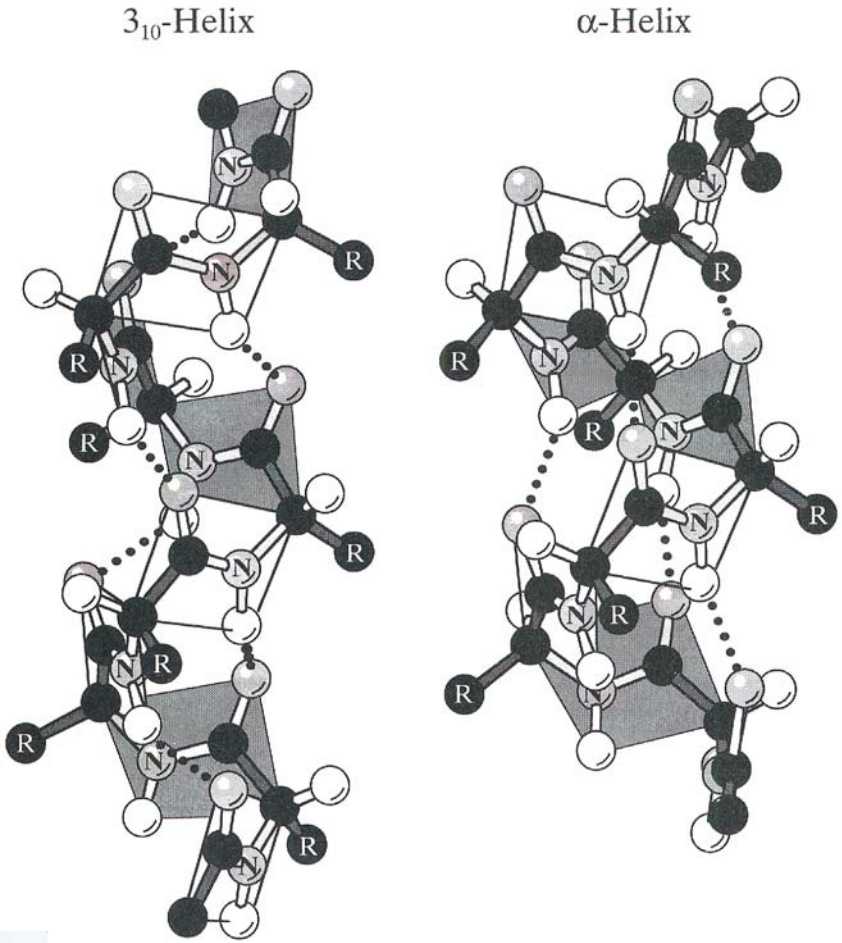
Symmetry Symbols

Table 1.4 Symbols for Symmetry

Symbol	Symmetry	Motif
	C_2 (two-fold)	Monomer
	2_1 (two-fold screw)	Monomer
	C_3 (three-fold)	Monomer
	3_1 (right-handed three-fold screw)	Monomer
	3_2 (left-handed three-fold screw)	Monomer
	C_4 (four-fold)	Monomer
	4_1 (right-handed four-fold screw)	Monomer
	4_2 (four-fold screw)	Dimer
	4_3 (left-handed four-fold screw)	Monomer
	C_6 (six-fold)	Monomer
	6_1 (right-handed six-fold screw)	Monomer
	6_2 (right-handed six-fold screw)	Dimer
	6_3 (six-fold screw)	Trimer
	6_4 (left-handed six-fold screw)	Dimer
	6_5 (left-handed six-fold screw)	Monomer

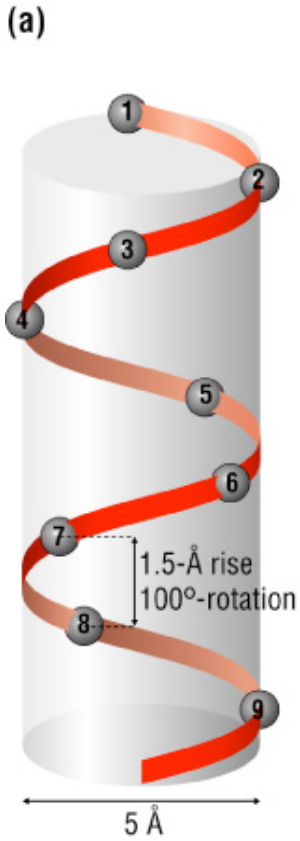
α-helix - regular secondary structure

- The most favorable ϕ and ψ angles with little steric hindrance (Typically ϕ = -60 º and ψ = -50 º).
- Forms repeated hydrogen-bonds.
- N = 3.6 residues per turn
- P = 5.4 Å (What is the d for an a-helix?) d=p/n=5.4Å/3.6=1.5
- The C=O of the n^{th} residue points towards the N-H of the $(N+4)^{th}$ residue.

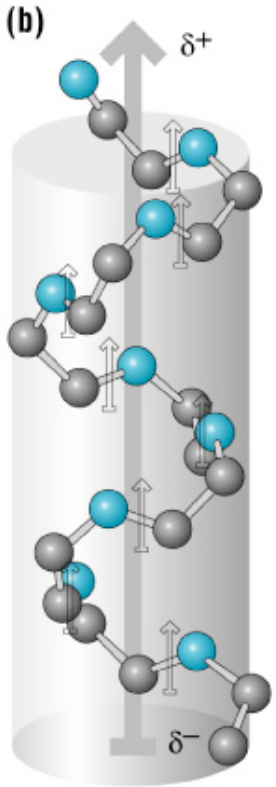


$$\begin{vmatrix} -0.174 & -0.985 & 0 \\ 0.985 & -0.174 & 0 \\ 0 & 0 & 1 \end{vmatrix} \times \begin{vmatrix} x \\ y \\ z \end{vmatrix}_i + \begin{vmatrix} 0 \\ 0 \\ 0.15\ \text{nm} \end{vmatrix} = \begin{vmatrix} x \\ y \\ z \end{vmatrix}_{i+1}$$

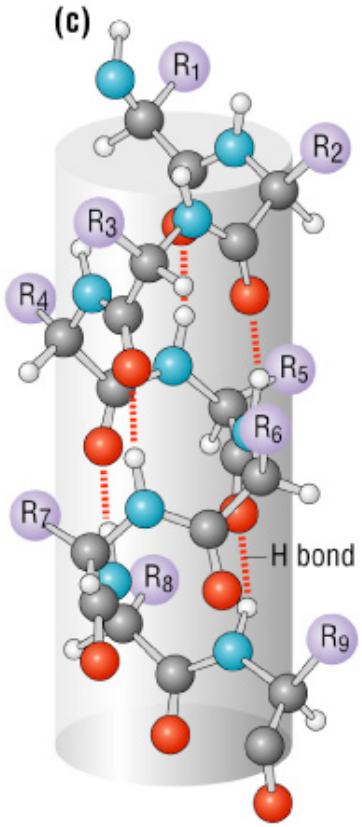
α-helix formed by 1-5 (n+4th), N-H·····O H-bond



Cα

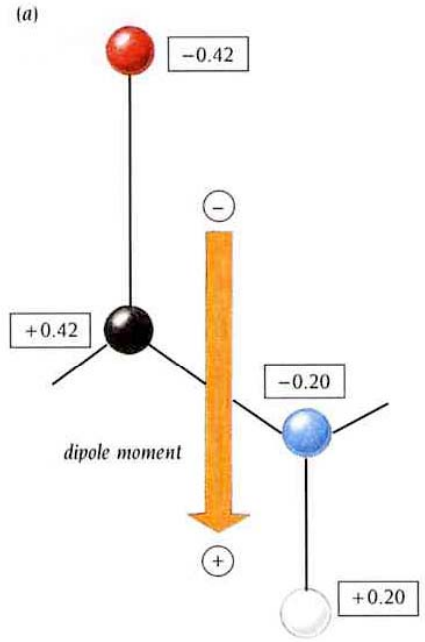


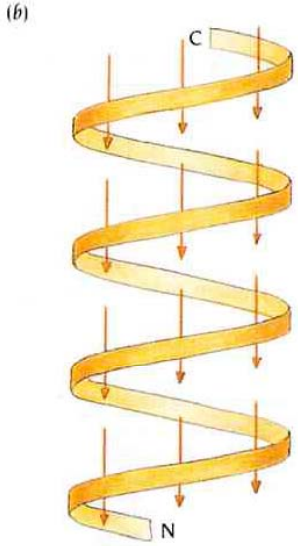
Backbone fold

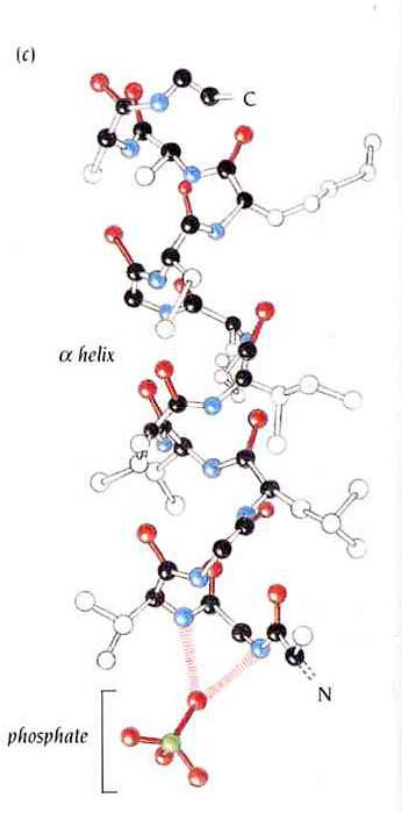


Full structure

The α-helix has a dipole moment







Dipole of a peptide unit

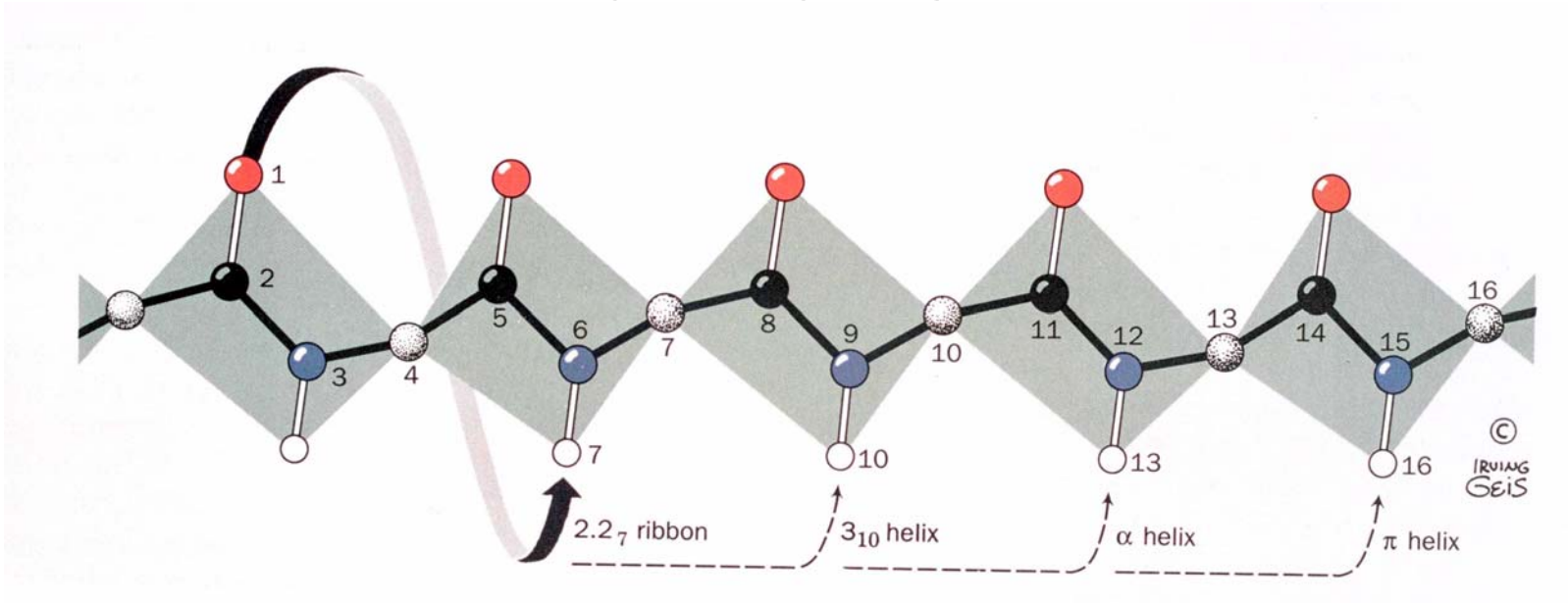
Macro-dipole
along with helix axis

Negatively charged groups frequently bind to the amino ends of α helices.

The N_m nomenclature for helices

N = the number of repeating units per turn.

M = the number of atoms that complete the cyclic system that is enclosed by the hydrogen bond.



From **Protein Structure and Function** by Gregory A Petsko and Dagmar Ringe

Average Conformational Parameters of Helical Elements

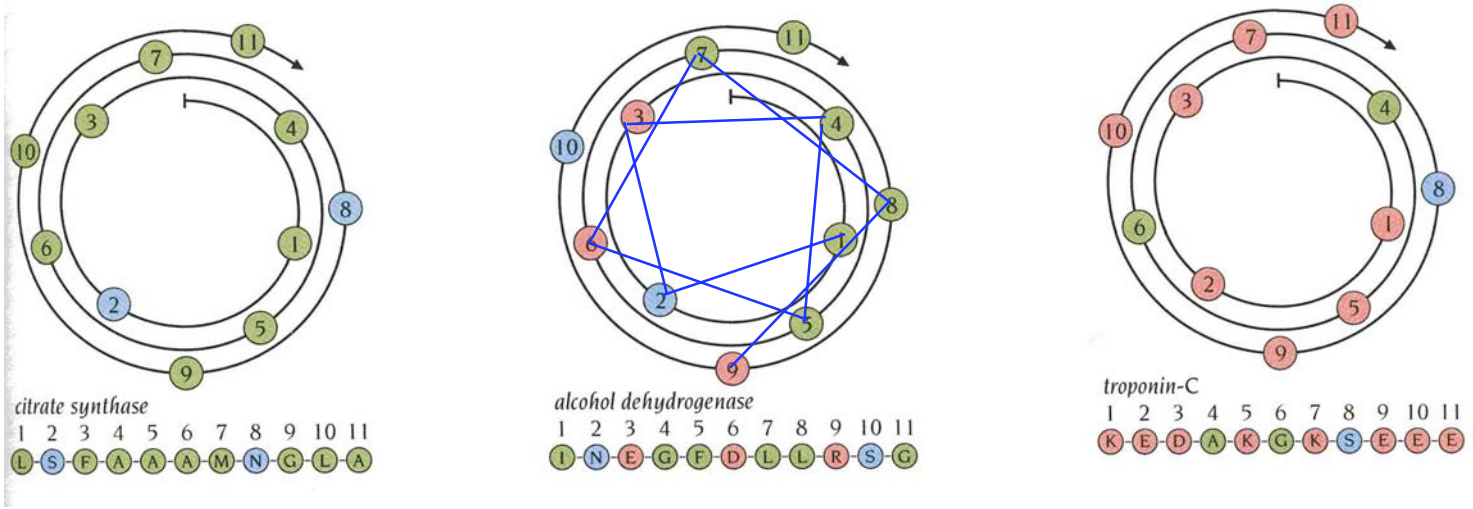
Conformation	Phi	Psi	Omega	Residues per turn	Translation per residue
Alpha helix	–57	–47	180	3.6	1.5
3-10 helix	–49	–26	180	3.0	2.0
Pi-helix	57	–70	180	4.4	1.15
Polyproline I	–83	+158	0	3.33	1.9
Polyproline II	–78	+149	180	3.0	3.12
Polyproline III	–80	+150	180	3.0	3.1

The helical wheel or Spiral

Table 2.1 Amino acid sequences of three α helices

1. - Leu - Ser - Phe - Ala - Ala - Ala - Met - Asn - Gly - Leu - Ala -
2. - Ile - Asn - Glu - Gly - Phe - Asp - Leu - Leu - Arg - Ser - Gly -
3. - Lys - Glu - Asp - Ala - Lys - Gly - Lys - Ser - Glu - Glu - Glu -

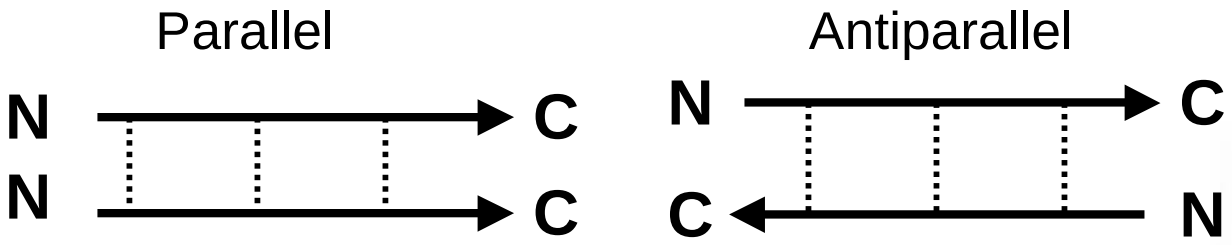
The first sequence is from the enzyme citrate synthase, residues 260–270, which form a buried helix; the second sequence is from the enzyme alcohol dehydrogenase, residues 355–365, which form a partially exposed helix; and the third sequence is from troponin-C, residues 87–97, which form a completely exposed helix. Charged residues are colored red, polar residues are blue, and hydrophobic residues are green.



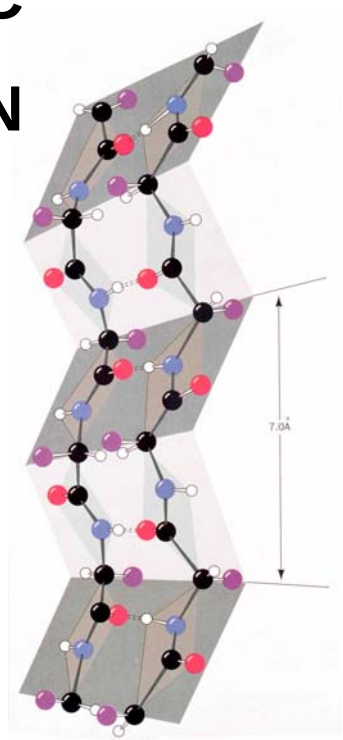
aa residues plotted every 100° round spiral.

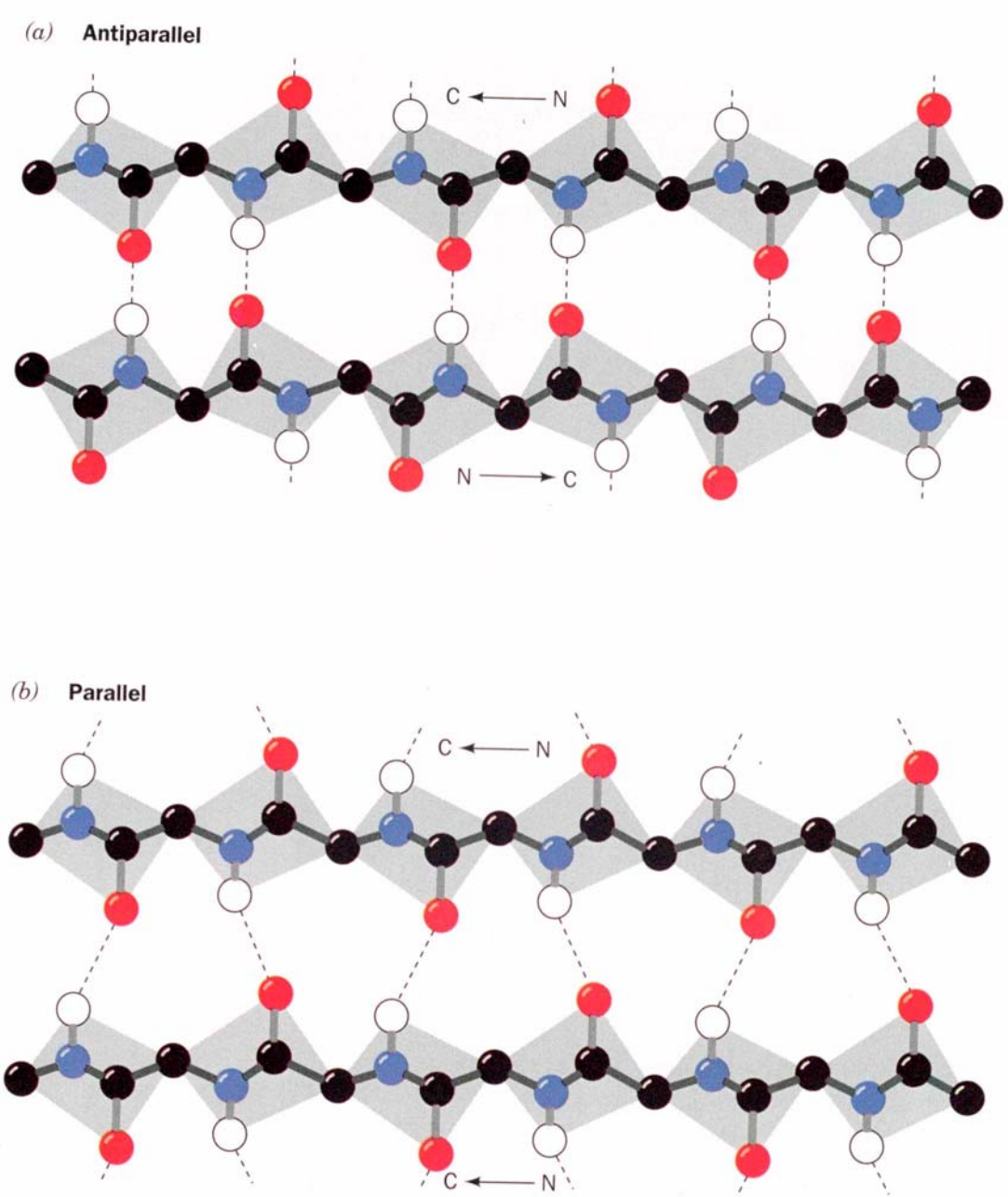
beta (β) sheet - regular secondary structure

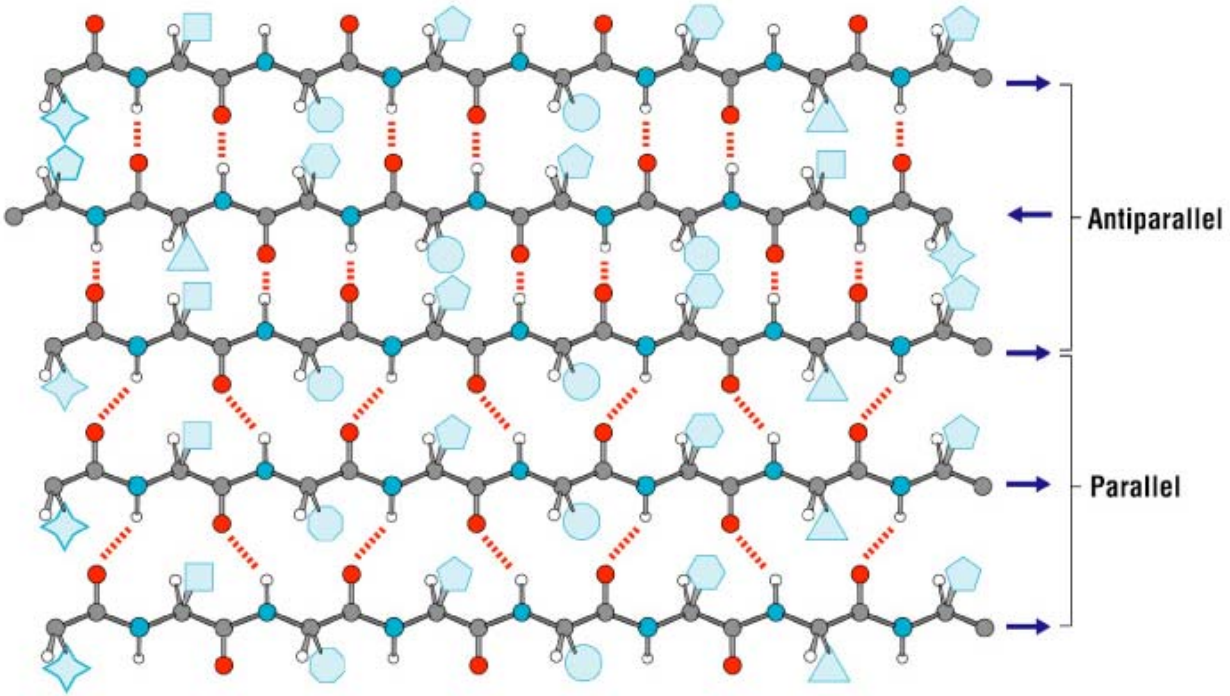
- Hydrogen-bonding between adjacent peptide chains.
- Almost fully extended but have a buckle or a pleat.
  Much like a Ruffles potato chip
- Two types:

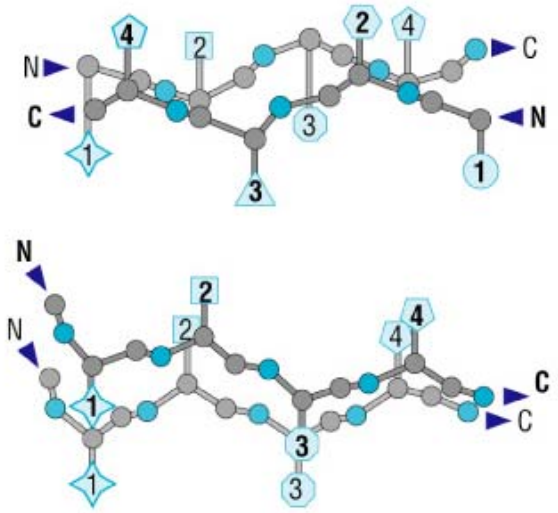


- 7.0 Å between pleats on the sheet
- Widely found pleated sheets exhibit a right-handed twist, seen in many globular proteins.




(a) Antiparallel
C ← N
N → C
(b) Parallel
C ← N
C ← N



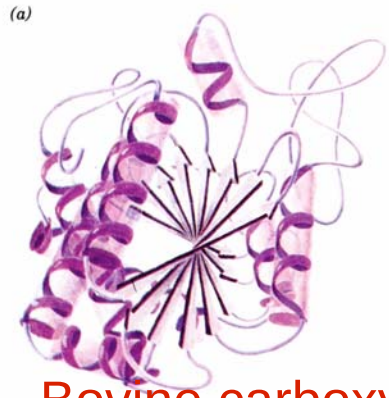


Mixed β sheets

edge-on views

β sheets are twisted

two proteins exhibiting a twisting β sheet



Bovine carboxypeptidase

Triose phosphate isomerase

- The twist is due to chiral L-amino acids in the extended plane.
- This chirality gives the twist and distorts H-bonding.
- A little tug of war exists between conformational energies of the side chain and maximal H-bonding.
- These structures are not “static” but breathe and vibrate with a change in structure due to external circumstances.

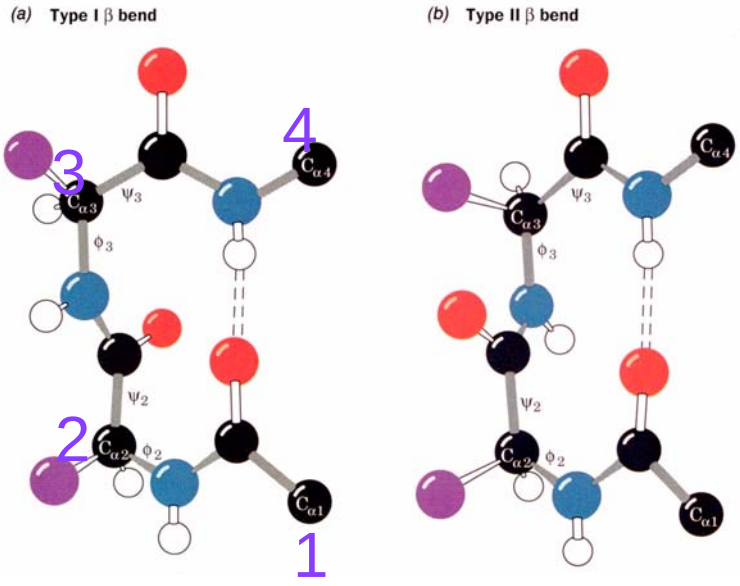
Non-repetitive regions

Turns - coils or loops

50% of structure of globular proteins are not repeating structures

β-bends

type I and type II: hairpin turn between anti-parallel sheets



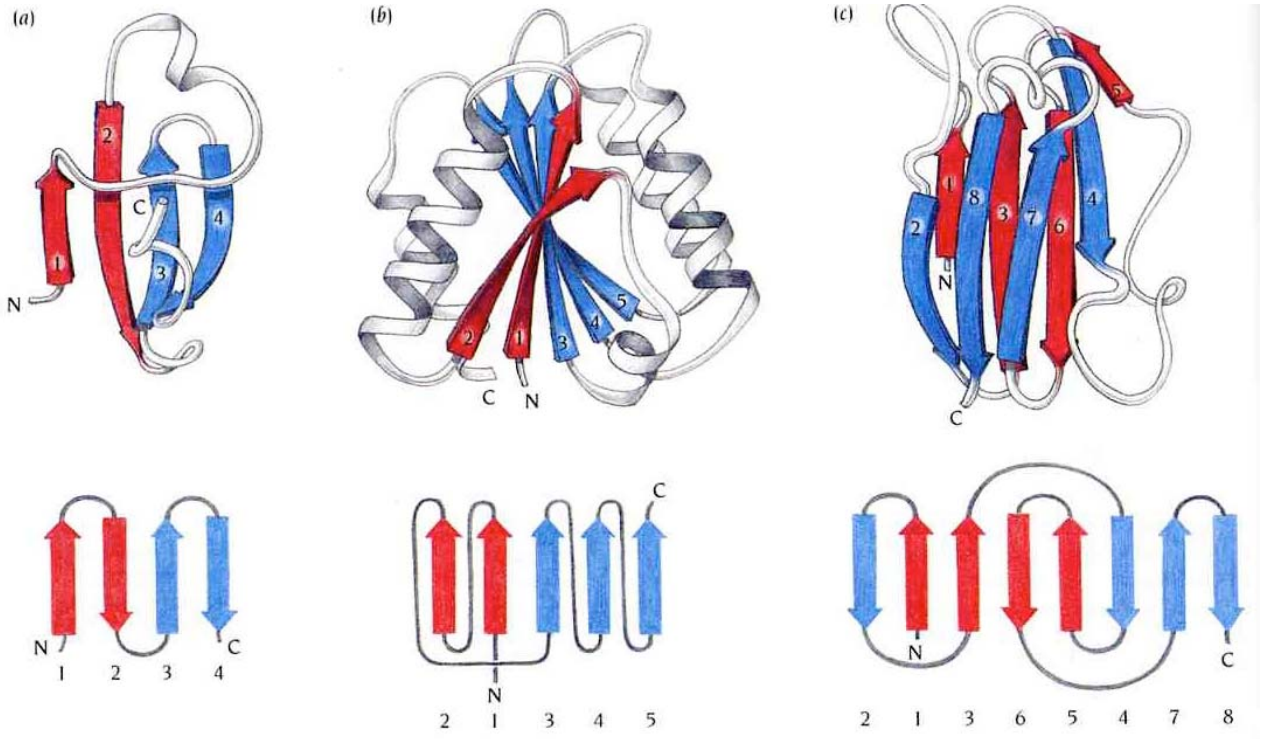
Type I $\phi_2 = -60^o, \psi_2 = -30^o$

$\phi_3 = -90^o, \psi_3 = 0^o$

Type II $\phi_2 = -60^o, \psi_2 = 120^o$

$\phi_3 = 90^o, \psi_3 = 0^o$

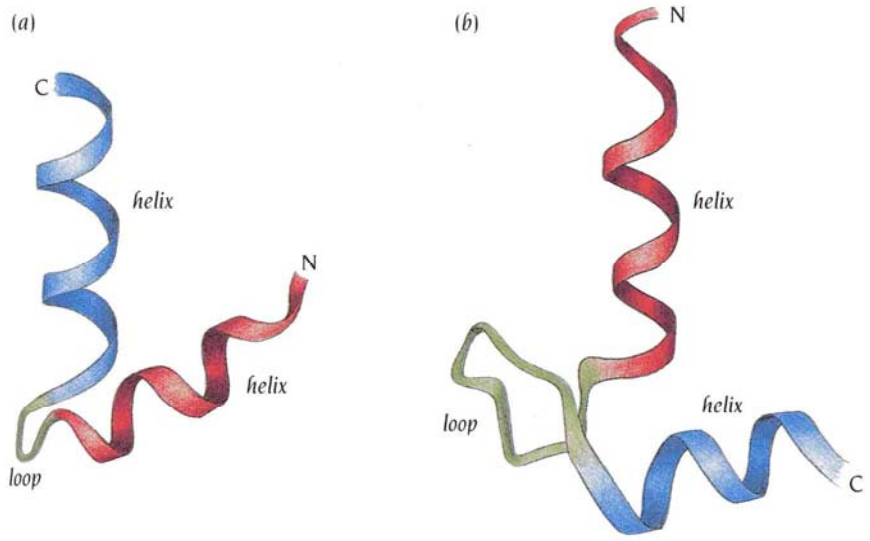
Topology diagrams are useful for classification of protein structure



Adapted from Branden & Tooze

β sheets by arrows (direction and connectivity)

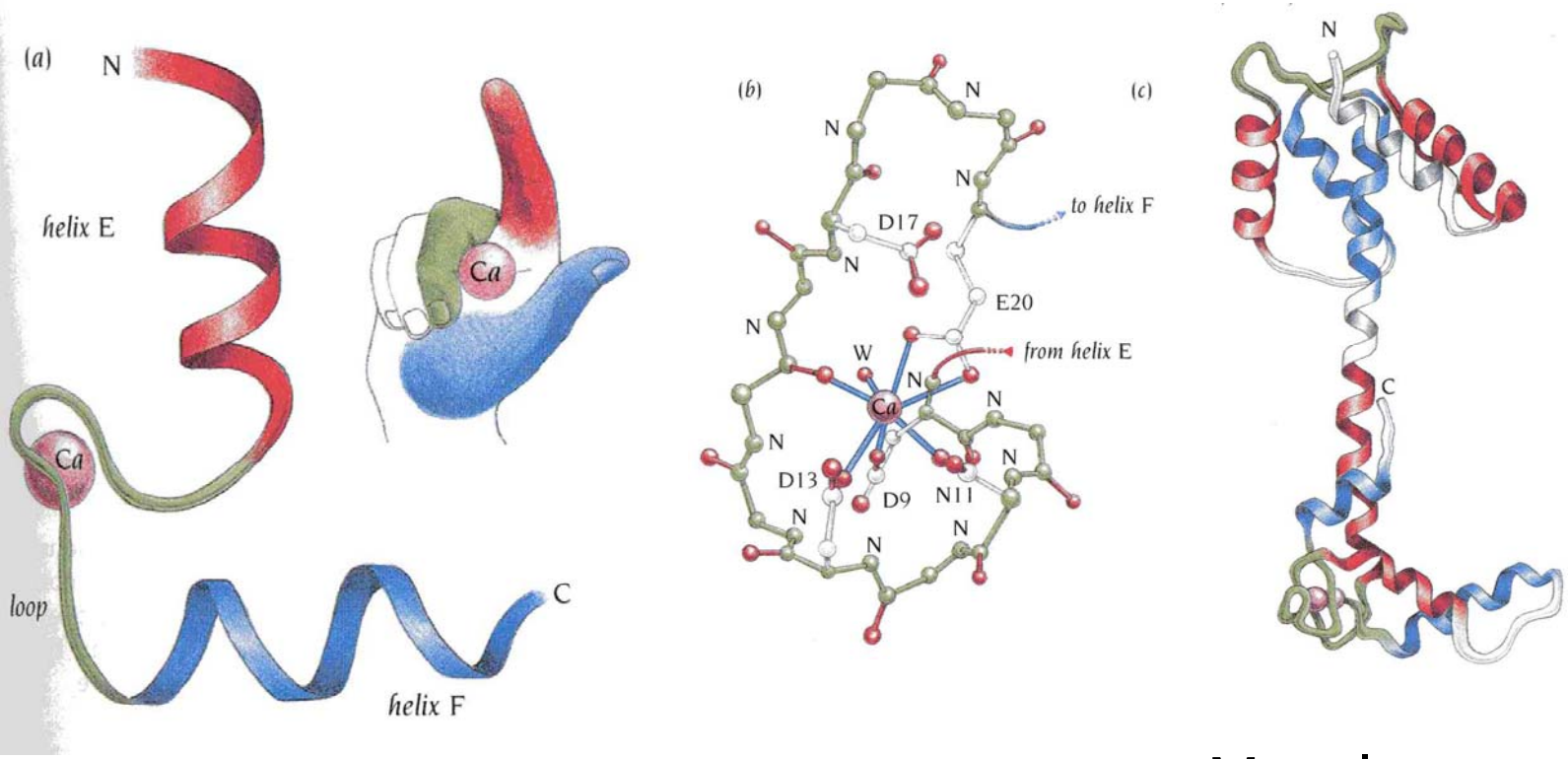
Two α helices connected by a loop in a specific geometric arrangement: helix-turn-helix



Helix-turn-helix motif: DNA-binding motif

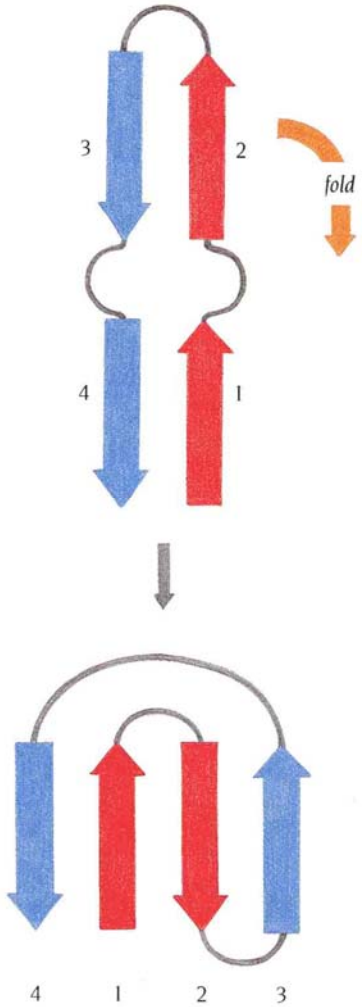
Ca^{2+}-binding motif

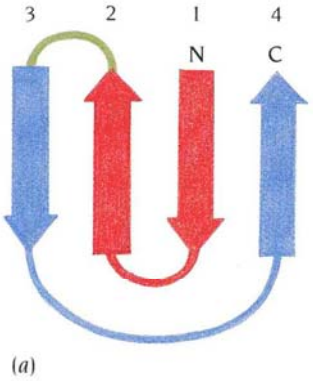
Calcium-binding motif: EF hand

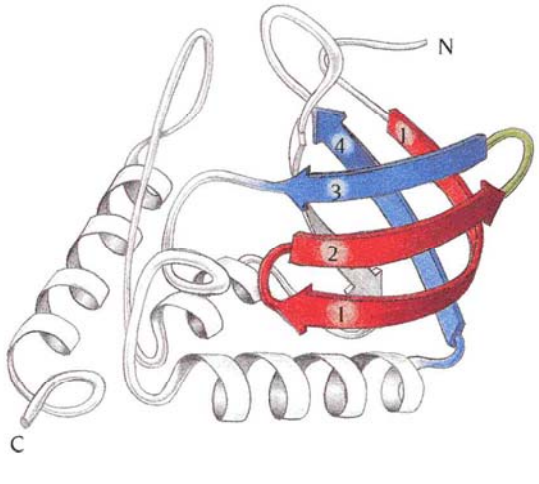


Muscle protein troponin-C

Greek key motif



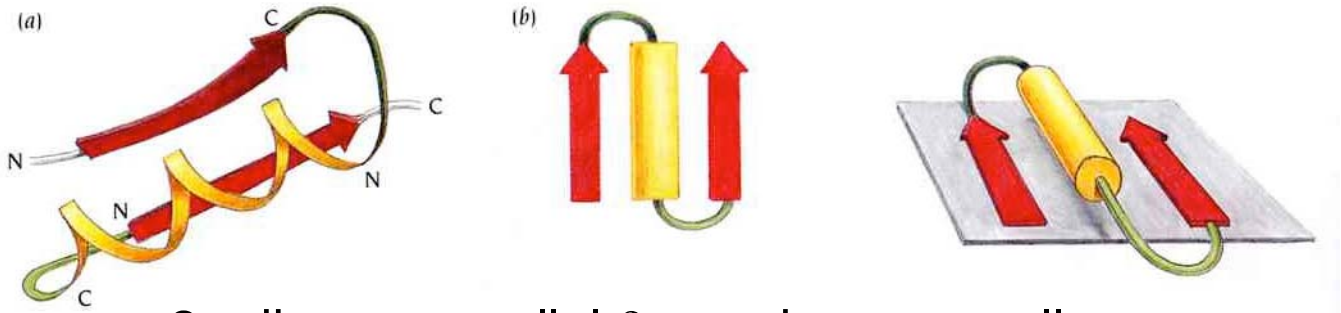




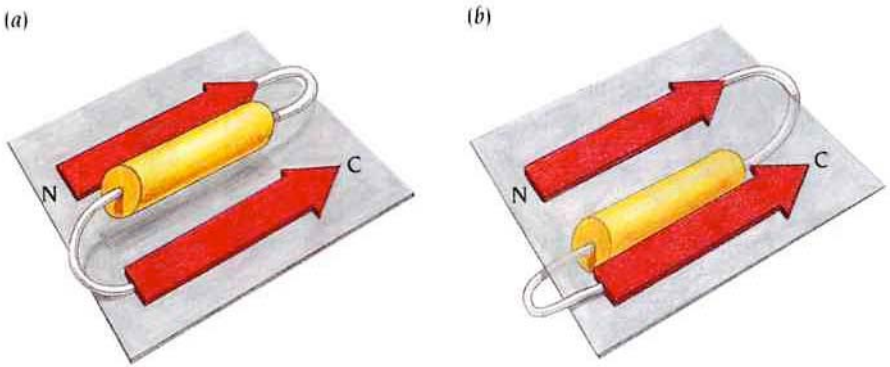
4 adjacent β strands arrangement

Suggested folding pathway

β–α–β motif



2 adjacent parallel β stands are usually connected by an α-helix from C-ter of helix 1 to N-ter of helix 2



Right handed - same as α-helix

In principle, two “hands”

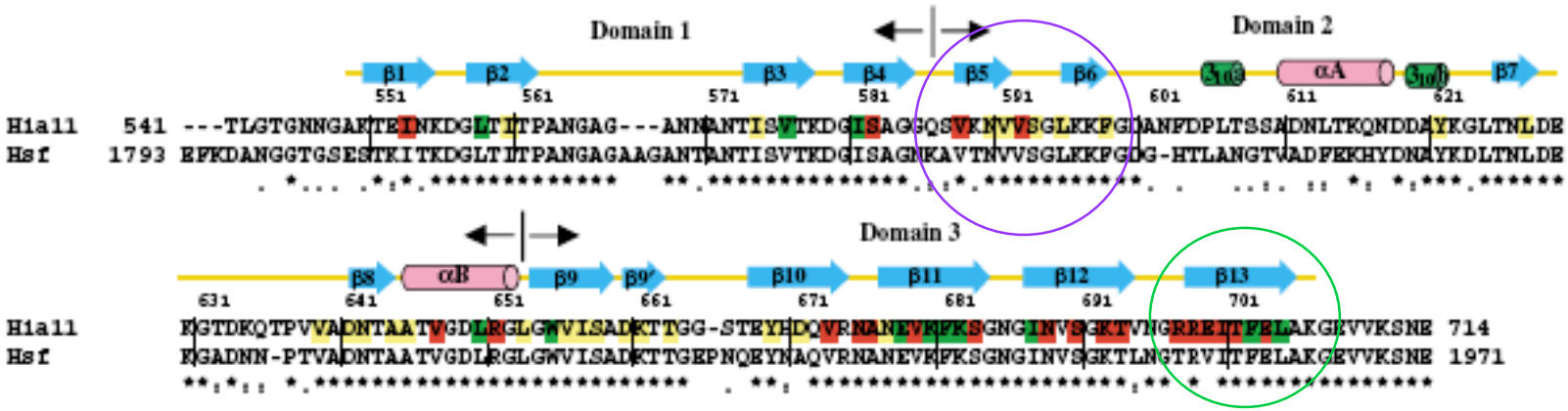
Conformational Preferences of the Amino Acids			
Amino acid	Preference		
	α-helix	β-strand	Reverse turn
Glu	**1.59**	0.52	1.01
Ala	**1.41**	0.72	0.82
Leu	**1.34**	1.22	0.57
Met	**1.30**	1.14	0.52
Gln	**1.27**	0.98	0.84
Lys	**1.23**	0.69	1.07
Arg	**1.21**	0.84	0.90
His	**1.05**	0.80	0.81
Val	0.90	**1.87**	0.41
Ile	1.09	**1.67**	0.47
Tyr	0.74	**1.45**	0.76
Cys	0.66	**1.40**	0.54
Trp	1.02	**1.35**	0.65
Phe	1.16	**1.33**	0.59
Thr	0.76	**1.17**	0.90
Gly	0.43	0.58	**1.77**
Asn	0.76	0.48	**1.34**
Pro	0.34	0.31	**1.32**
Ser	0.57	0.96	**1.22**
Asp	0.99	0.39	**1.24**

From Petsko and Ringe

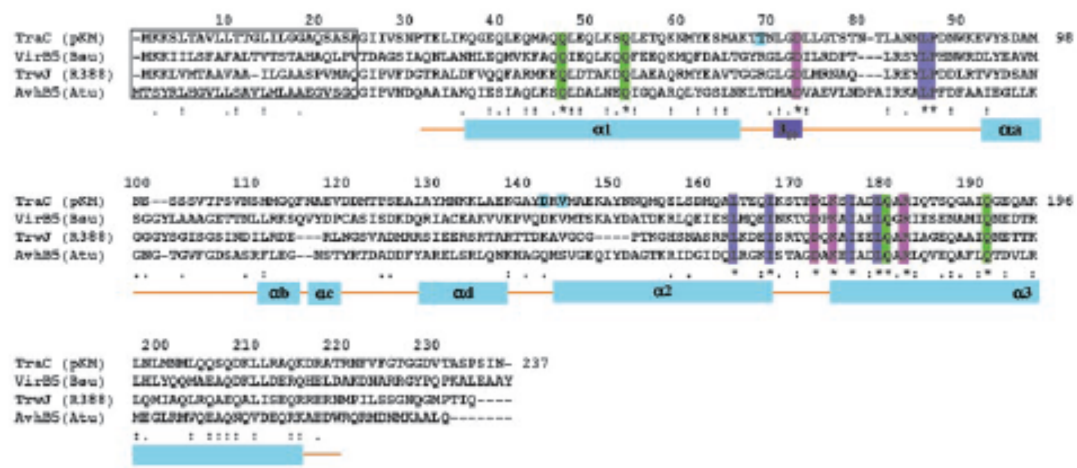
Certain amino acids are more usually found in α helices, others in β sheets

Prediction of Secondary Structure

```
                  10        20        30        40        50        60        70        80        90        100
                   |         |         |         |         |         |         |         |         |         |
UNK_279690 GPLGSTLGTGNNGAKTEINKDGLTITPANGAGANNANTISVTKDGISAGGQSVKNVVSGLKKFGDANFDPLTSSADNLTKQNDDAYKGLTNLDEKGTDKQ
DPM        tccccccttttccccchettccecececcctctcctcceeeeecccctttcccceeeeeeccccctctcccccctcccccttccccccccchttctcccc
HNNC       ccccccccccccccceeecccceeeeccccccccccceeeeeccccccccchhhhhhhhhhhhcccccccccccchhhccchhhhccccccccccccc
MLRC       ccceeeeccccccceeeecccceeeecccccccccceeeeeeccccccchhhhhhhhhhhhcccccccccccchhhhhccchhhchcccccccccc
PHD        ccccceccccccceeeeccccceeecccccccccceeeeeccccecccceeeeeeceeeeeccccccccccccccccccchhhhhhcccccccccc
Predator   ccccccccccccceeeeeccccccccccccccccceeeeeccccccccceeecccccccccccccccccccccccccccccccccccccccccc
SOPM       cceeeeeeccctteeeeecttteeeecccccccccceeeeeetccttcchhhhhhhhhhhcccccccccchhhhhhhhtchhhhhhhhhhctttcccc
Sec.Cons.  ccccc?ccccccccceeeeccccceeeecccccccccceeeeeccccccccch??hhhhhhhhcccccccccccccc???cccccchh?c?ccccccccccc

                  110       120       130       140       150       160       170
                   |         |         |         |         |         |         |
UNK_279690 TPVVADNTAATVGDLRGLGWVISADKTTGGSTEYHDQVRNANEVKFKSGNGINVSGKTVNGRREITFELAKGEVVKSNE
DPM        cceecccchheecceteceeeehccccteteccccceehthchehetttcceteccctetcchheehhhhcchheecccc
HNNC       cceeccccccecccccceeeeeecccccccchhhhhhcccceeeecccceeeccccccchheeehhcccceeeccc
MLRC       ceeeeccccccccchcccceeeeecccccccccchhhhhhcccceeeecccccceeeecccccccceeeeeeecccceeeccc
PHD        ceeeeecccceeeeccceeeeeeeeecccccceeehhhhcccceeeeecccceeeeeeecccceeeeeeeecccceeeeecc
Predator   ccccccccccccccccccccccccccccccccccccccccccccccceeeeecccccccccceeeecchhhhhhhhhccccccc
SOPM       ceeeetcchhhhhhhttttteeeeecccccccchhhhhcchhheeecttcceeeeccceettcceeeeeectteeehtc
Sec.Cons.  c?ee?ccccceccccccceeeeeecccccccchhhhh?ccceeeeecccceee?cce?ccc??eee??ccccceeeccc
```



```
                     10        20        30        40        50        60        70        80        90       100
                      |         |         |         |         |         |         |         |         |         |
UNK_23720 MKKSLTAVLLTTGLILGGAQSASAGIIVSNPTELIKQGEQLEQMAQQLEQLKSQLETQKNMYESMAKTTNLGDLLGTSTNTLANNLPDNWKEVYSDAMNS
DPM       cchhhhhhhheechehcchhhhhhceeecccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhcchcchccccccchhcccccchhhhhhhhhcc
HNNC      cchhhhhhhhhhhhhhccccccceeeeecccchhhhccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhccccchcccchhhhhhccccchhhhhhhcccc
MLRC      cccceeeeeeechheeeccccccccceeeecccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhcccccccccccchhhhcccccchhhhhhhhccc
PHD       ccccceeeehhhcceechhhhcccceeeecchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhcccccccchhhhhhchhhhhhhhhhhhhhcc
Predator  ccchhhhhhhhcccccccccccccceeeecccccccccccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhcccccccccccccccccccccccccccccccc
SOPM      hhhhhhheeeeeeeeetcccccccceeeecccchhecttthhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhcchhhhhhccccchhhhhhhhhcc
Sec.Cons. cc?hhhh??h??c?eecccccccceeeeccc?hhh?hhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhccccc?cccc?hhhhcccccchhhhhhhh?cc

                    110       120       130       140       150       160       170       180       190       200
                      |         |         |         |         |         |         |         |         |         |
UNK_23720 SSSVTPSVNSMMGQFNAEVDDMTPSEAIAYMNKKLAEKGAYDRVMAEKAYNNQMQELSDMQALTEQIKSTPDLKSIADLQARIQTSQGAIQGEQAKLNLM
DPM       cccecccecchhhhhhhhhhhhcchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhhcccchhhhhhhhhhhhhhhchhhhhhhhhhhh
HNNC      cccccccchhhhcccccccccccchhhhhhhhhhhhhccchhhhhhhhhhhhchhhhhhhhhhhhhhhccccchhhhhhhhhhhhcccccchhhhhhhhh
MLRC      ccccchhhhhhhhhhchchccccchhhhhhhhhhhhhccchhhhhhhhhhhhhhhhhhhhhhhhhhhhcccchhhhhhhhhhhhhcccccchhhhhhhhh
PHD       ccccccchhhhhhcccceecccchhhhhhhhhhhhhhccchhhhhhhhhhhhhhhhhhhhhhhhhhhhcccchhhhhhhhhhhhhhhhhhhhhhhhhhhh
Predator  ccceeeccchhhhhhhhhhcccccchhhhhhhhhhhhhccchhhhhhhhhchhhhhhhhhhhhhhhhcccccchhhhhhhhhhhhhhhhhhhhhhhhhhh
SOPM      ccccchhhhhhhhhhhhhhhtccchhhhhhhhhhhhhttchhhhhhhhhhhhhhhhhhhhhhhhhhhcccccccchhhhhhhhhhccttccchhhhhhhhh
Sec.Cons. ccccccchhhhhhhh?hhhccccchhhhhhhhhhhhhccchhhhhhhhhhhhhhhhhhhhhhhhhhhhcccc?hhhhhhhhhhhh?hc???hhhhhhhhh

                    210       220       230
                      |         |         |
UNK_23720 NMLQQSQDKLLRAQKDRATRNFVFGTGGDVTASPSIN
DPM       hhhhhhhhhhhhhhhhhhhhhheetctcccccccccc
HNNC      hhhhhhhhhhhhhhhhhccceeeeecccccccccccc
MLRC      hhhhhhhhhhhhhhhcccceeeeeecccccccccccc
PHD       hhhhhhhhhhhhhhhhhhcceeeeecccceecccccc
Predator  hhhhhhhhhhhhhhhhhhhhheeeccccccccccccc
SOPM      hhhhhhhhhhhhhhcccccceeeeecccceeeccccc
Sec.Cons. hhhhhhhhhhhhhhhhh?cceeeeecccccccccccc
```

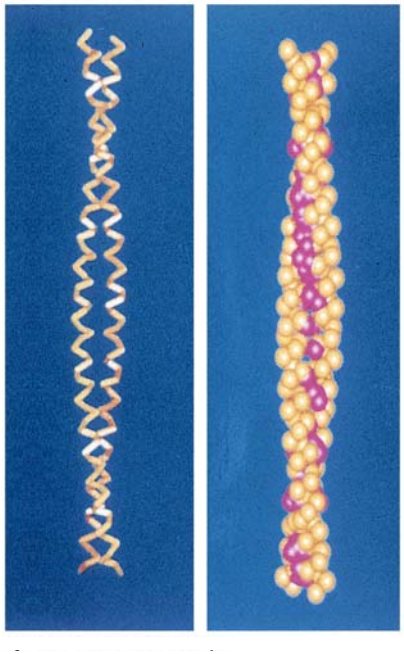


Alpha-domain structures

1. Coiled-coil helices: the basis for some fibrous proteins

Coiled-coils in fibers > many hundreds of amino acids

Shorter coiled-coils: transcription factors



α keratin proteins are helical but spacing differs from a regular α-helix

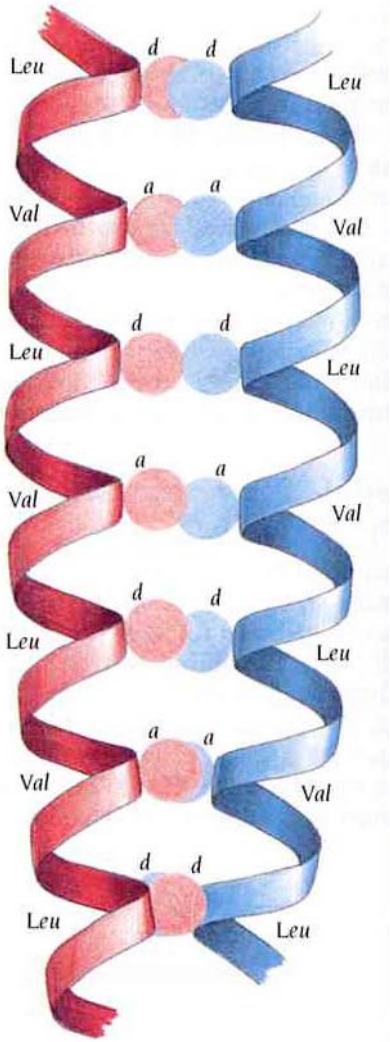
5.1 Å vs. **5.4** Å pitch.

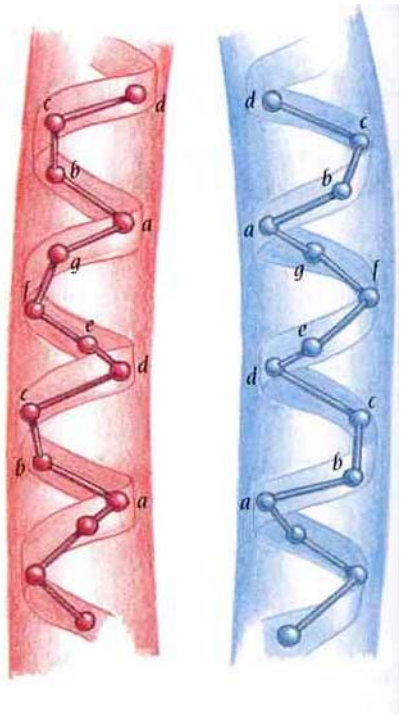
This change in pitch forms closely associated pairs of helices.

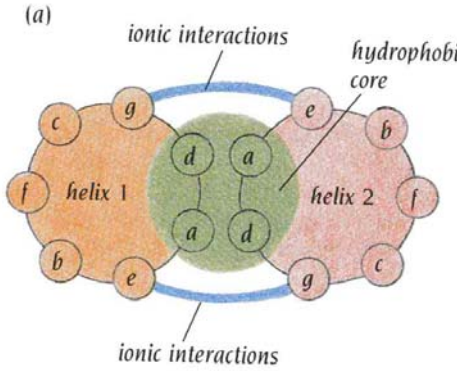
Left-handed coil

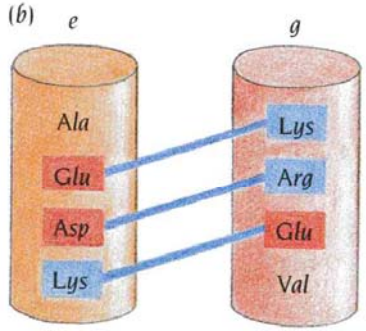
GNC4 showing a heptad repeat of Leu residues (Leu zipper)

	a	b	c	d	e	f	g	
NH_2 -	Met	- Lys	- Gln	- **Leu**	- Glu	- Asp	- Lys	-
	Val	- Glu	- Glu	- **Leu**	- Leu	- Ser	- Lys	-
	Asn	- Tyr	- His	- **Leu**	- Glu	- Asn	- Glu	-
	Val	- Ala	- Arg	- **Leu**	- Lys	- Lys	- Leu	- COOH



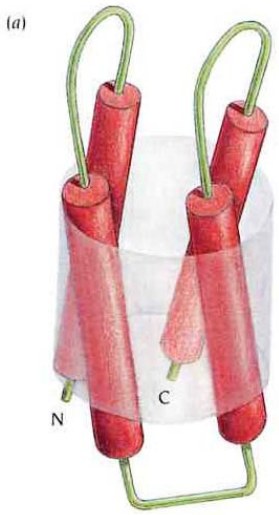


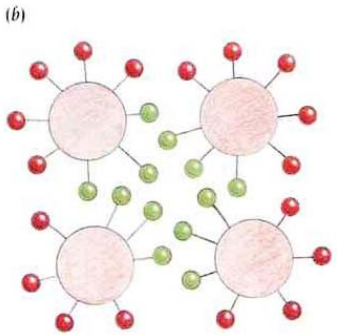


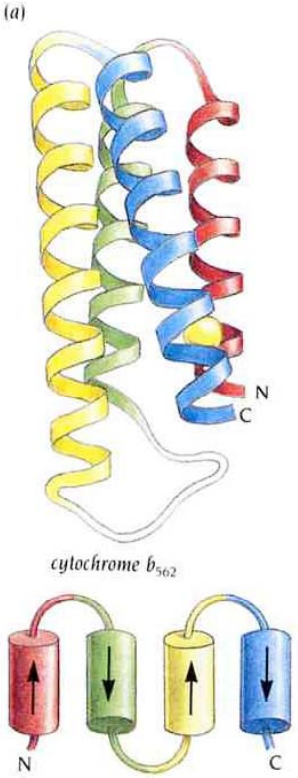


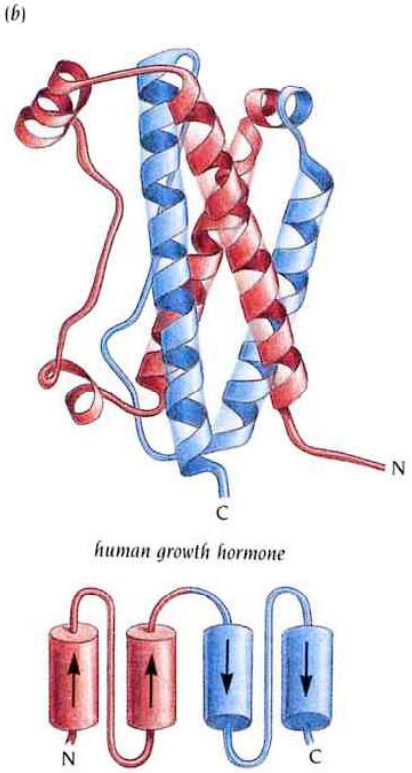
Helical repeat: 3.5 residues

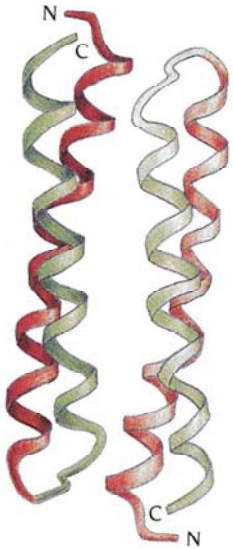
2. Four helix bundle





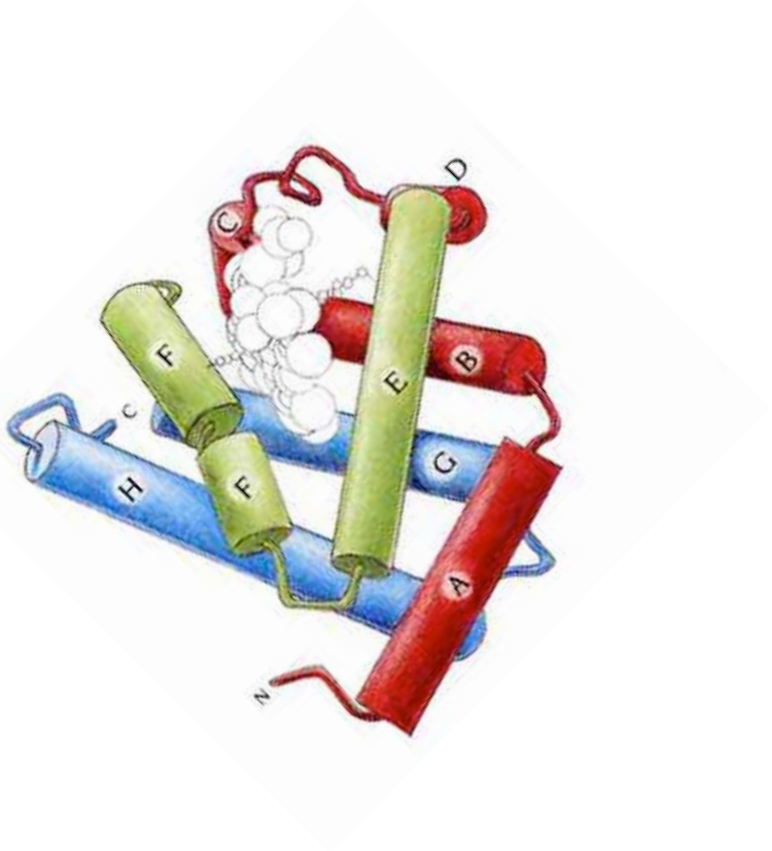






Dimeric Rop

3. The globin fold (myoglobin, hemoglobin, phycocyanins)



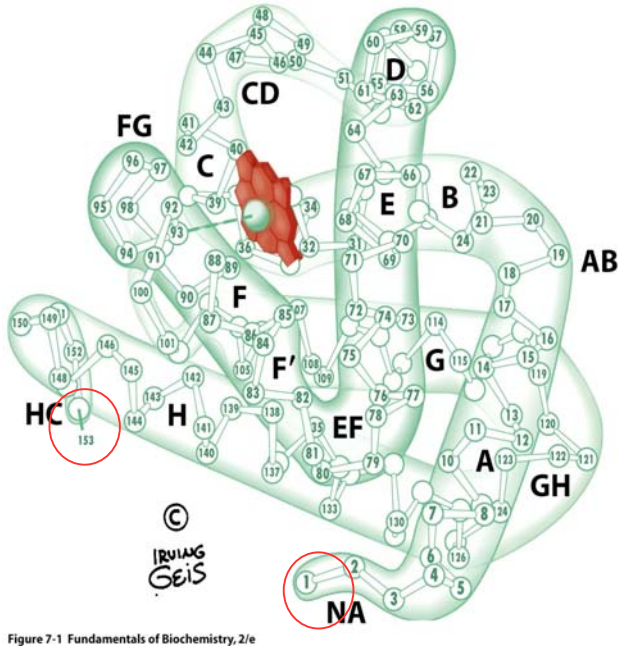
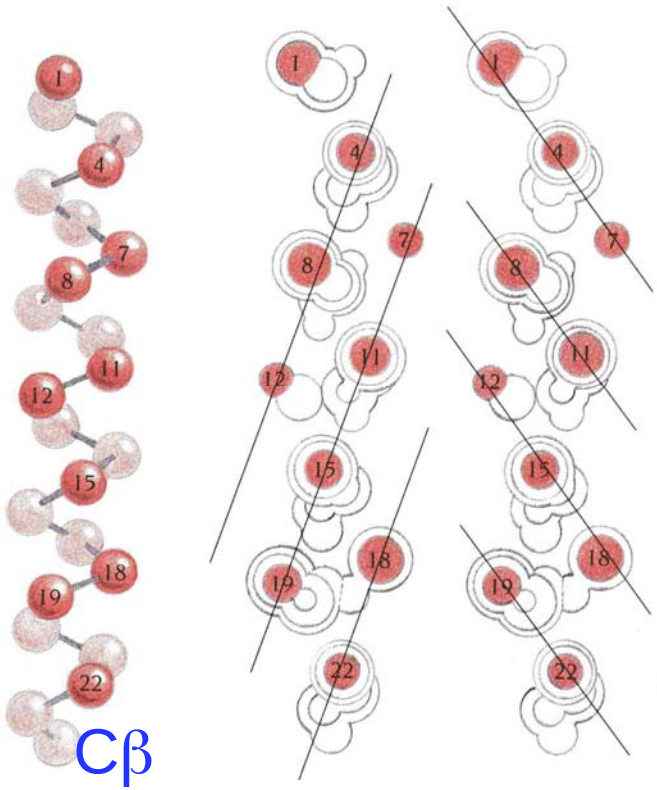


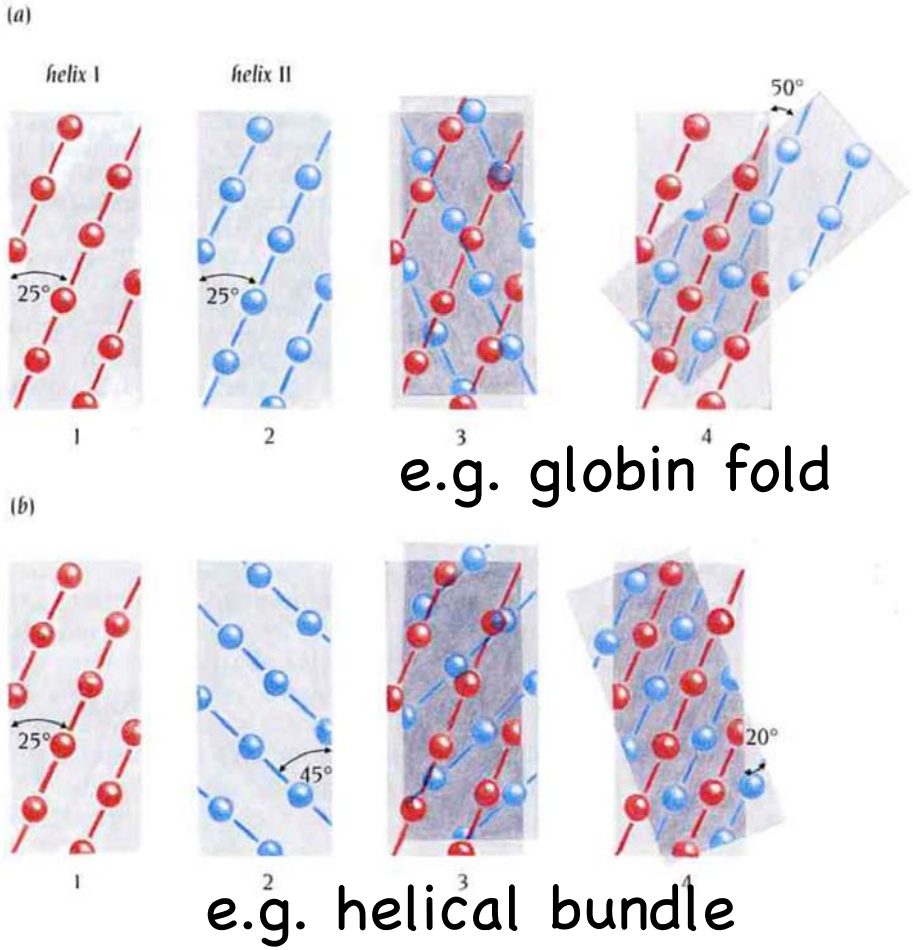
Figure 7-1 Fundamentals of Biochemistry, 2/e

Some features of general structural interests of globins

Geometric considerations determine α-helix packing



Side chains on the surface of an α helix form ridges



Fitting the ridges of side chains

- Coiled-coil α-helical structures are found both in fibrous proteins and as a part of smaller domains in many globular proteins.
- Two common motifs for alpha domains are the four helix bundle and the globin fold.
- Rules have been derived that explain the different geometrical arrangements of α helices observed in α-domain structures.
- The globin fold has been used to study evolutionary constraints for maintaining structure and function.

Alpha/Beta structures

Most frequent of the domain structures

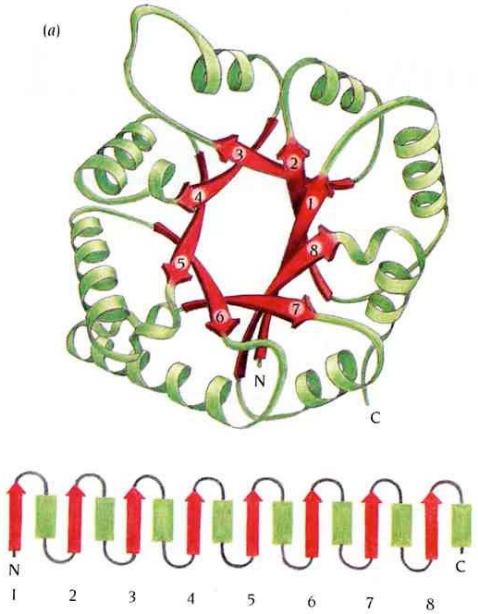
Parallel β strands are arranged in barrels or sheets

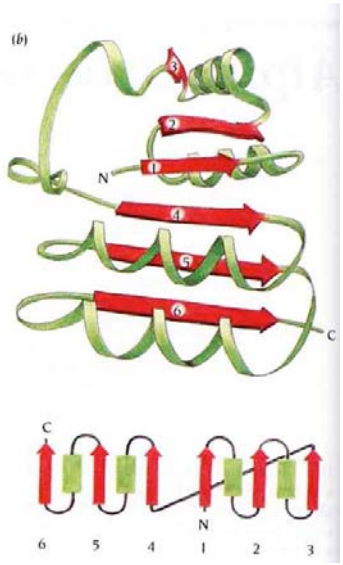
3 main classes of β–α–β motifs:

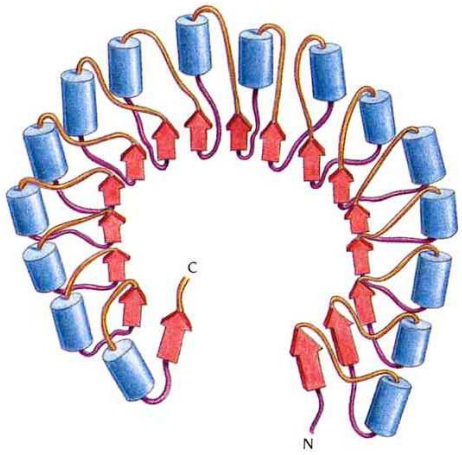
TIM barrel (triosephosphate isomerase)

Rossman fold (lactate dehydrogenase)

Horseshoe fold - Leucine-rich motifs



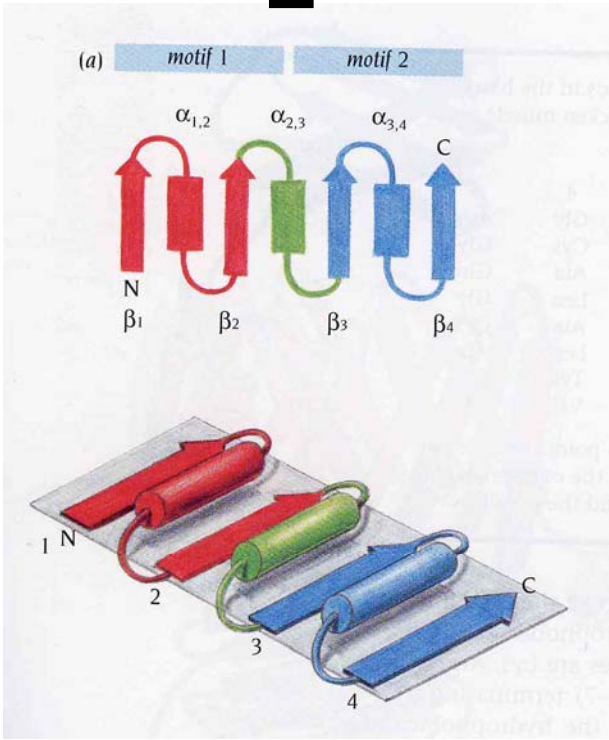




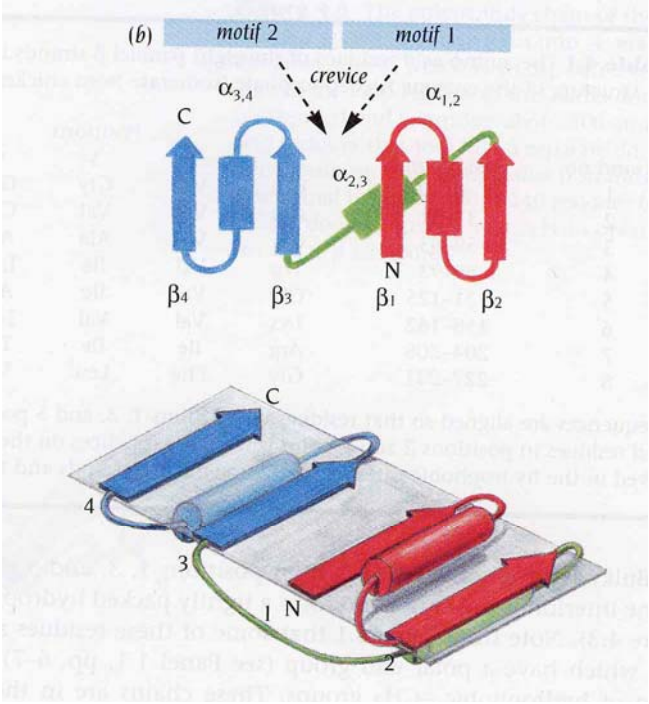
A β–α–β motif is a right-handed structure.

Two such motifs can be joined into a 4-stranded parallel β sheet in two different ways.

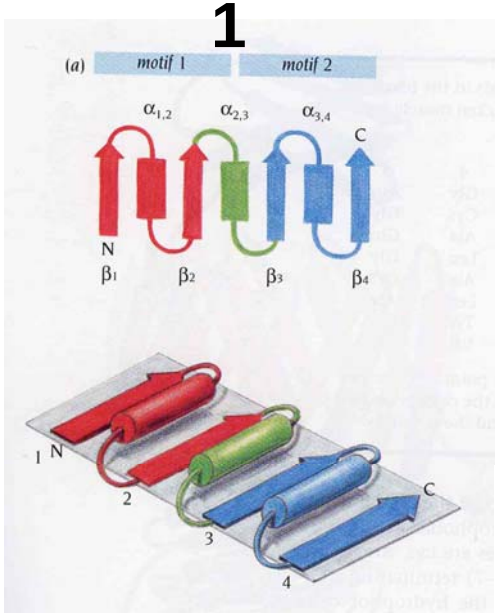
1



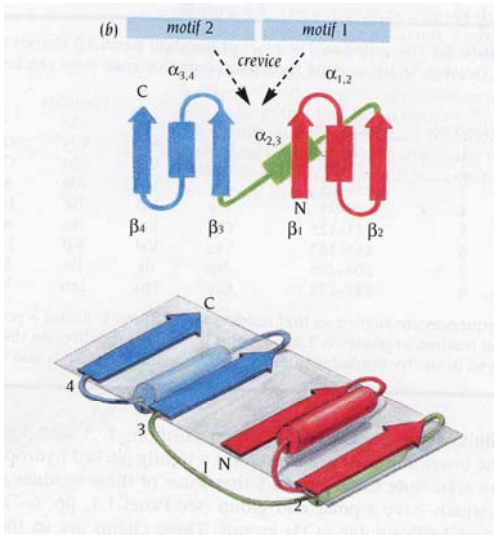
2

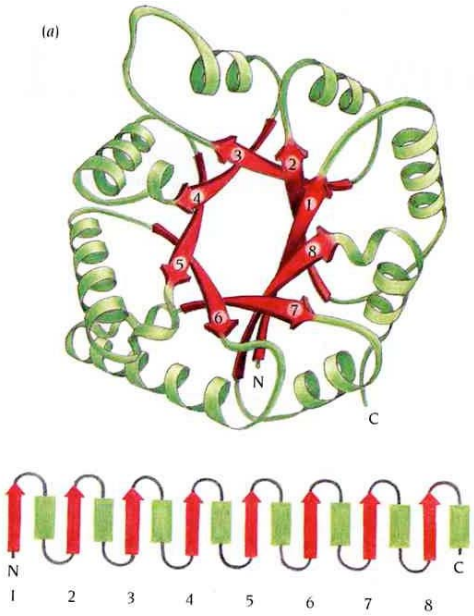


1


(a)
motif 1
motif 2
α1,2
α2,3
α3,4
C
N
β1
β2
β3
β4
1
N
2
3
4
C

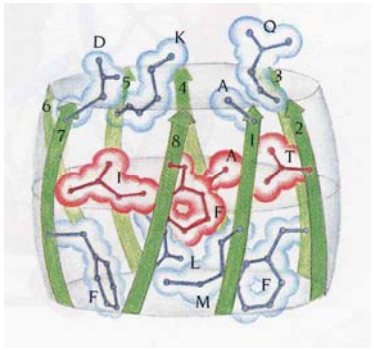
2

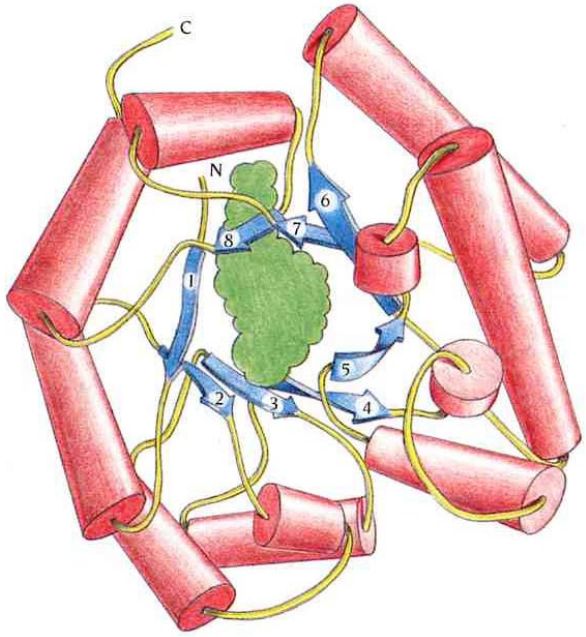

(b)
motif 2
motif 1
crevice
α3,4
α1,2
α2,3
C
N
β4
β3
β1
β2
C
4
3
1
N
2


(a)
3
2
4
1
5
8
6
7
N
C
N
C
1
2
3
4
5
6
7
8


(b)
3
2
N
1
4
5
6
C
C
N
6
5
4
1
2
3

α/β barrels occur in many different enzymes.





In most α/β barrel structures the 8 β strands of the barrel enclose a tightly packed hydrophobic core formed entirely by the side chains from the β strands.

Requirement of bulky hydrophobic residues!

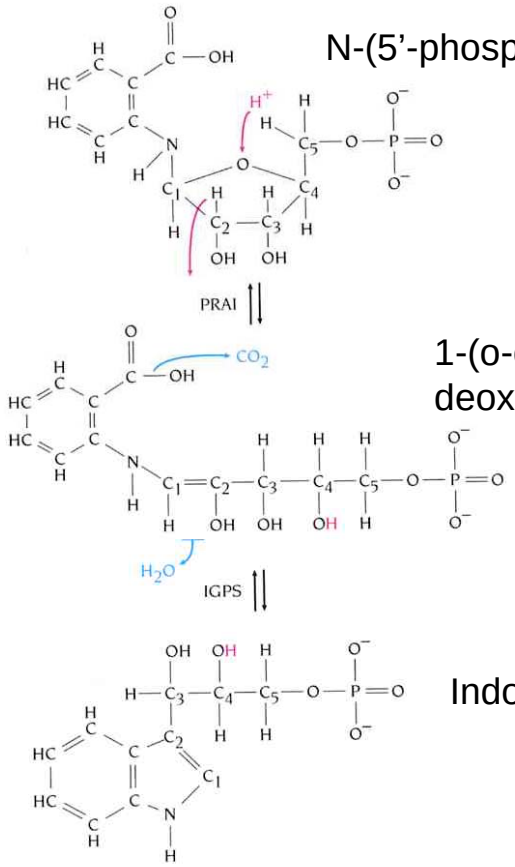
Exception:

methylmalonyl-CoA mutase

The inside of the barrel is filled with the small hydrophilic side chains (T, S)

Double barrels have occurred by gene fusion.

PRA-isomerase (PRAI): IGP-synthase
B. Subtilis (2 separate), *E. coli* (Bi), *Neurospora crassa* (Tri)

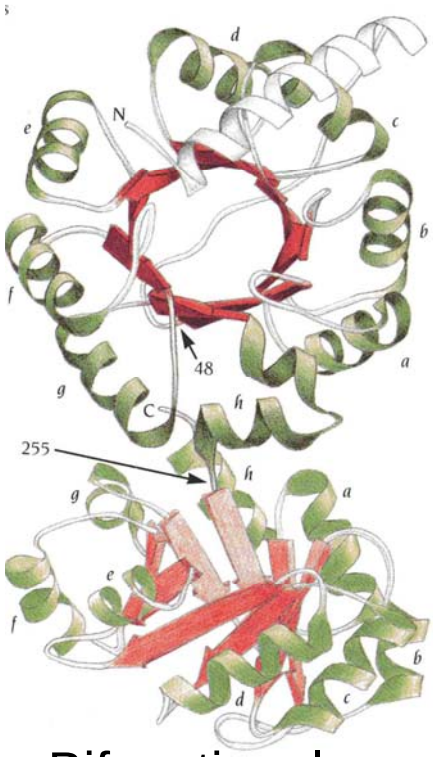


N-(5'-phosphoribosyl) anthranilate (PRA)

1-(o-carboxyphenylamino)-1-deoxyribulose 5-phosphate

Indole-3-glycerol phosphate (IGP)

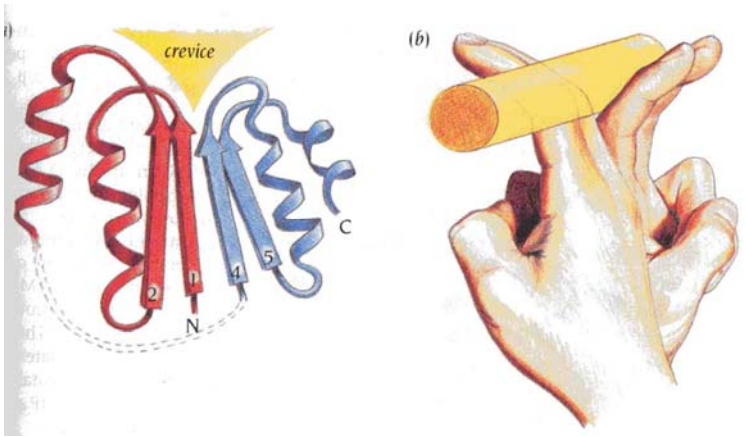
Biosynthesis of Trp pathway



Bifunctional enzyme
Active site: opposite

Structure ⇔ Function

- The active site is formed by loops at one end of the α/β barrel.
- α/β twisted open-sheet structures contain α helices on both sides of the β sheet.
- The active site in open twisted α/β domains is in a crevice outside the C-ends of the β-strands.
- Open β sheet structures have a variety of topologies.



The positions of active sites can be predicted in α/β structures

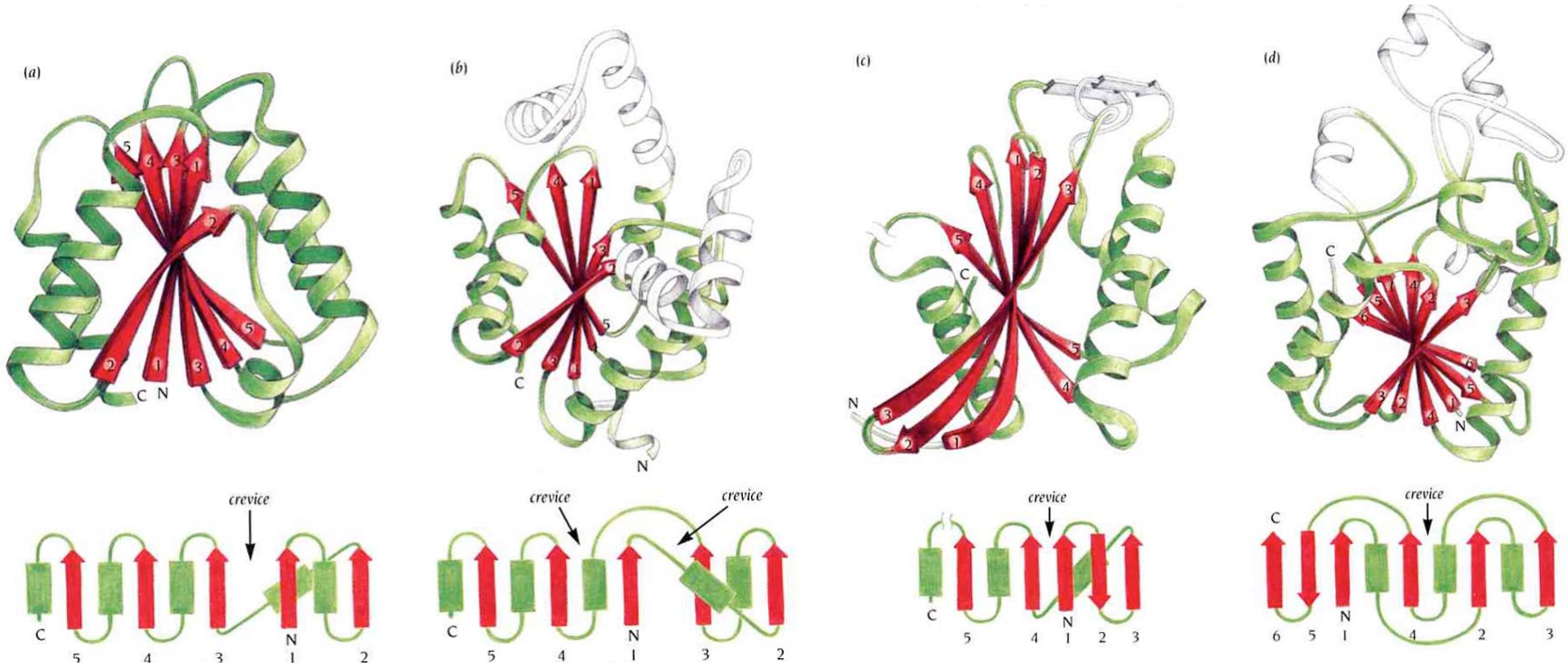
The α/β barrels have the active site at the same position with respect to their common structure in spite of having different functions as well as the primary sequences.

The open α/β sheet structures vary considerably in size, number of β strands, and strand order.

Independent of these variations, they all have their active sites at the carboxy edge of the β strands, and these active sites are lined by the loop regions that connect the β strands with the α helices.

The positions of these regions can be predicted from topology diagrams.

Examples of open twisted α/β structures

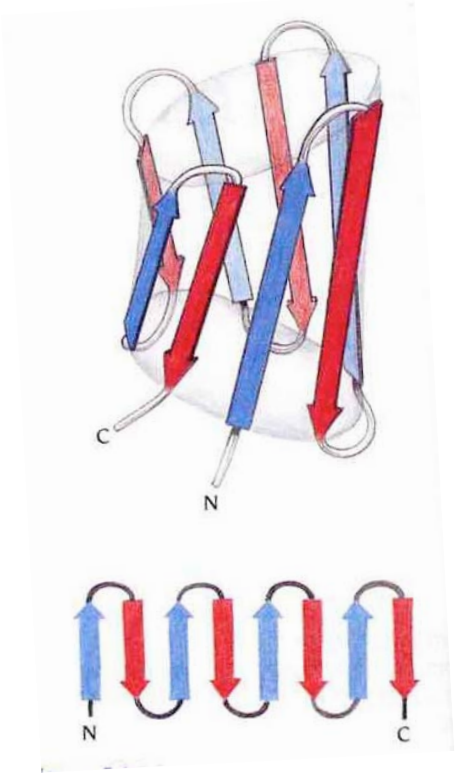


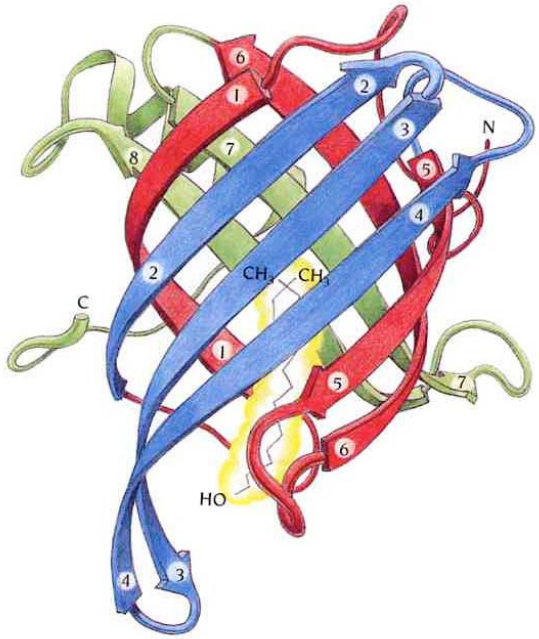
Flavodox adenylate kinase hexokinase glycerate mutase

Beta structures

Antiparallel β structures: the 2nd large group of the protein domain structures. Functionally, the most diverse!

1. Up-and-down β barrels

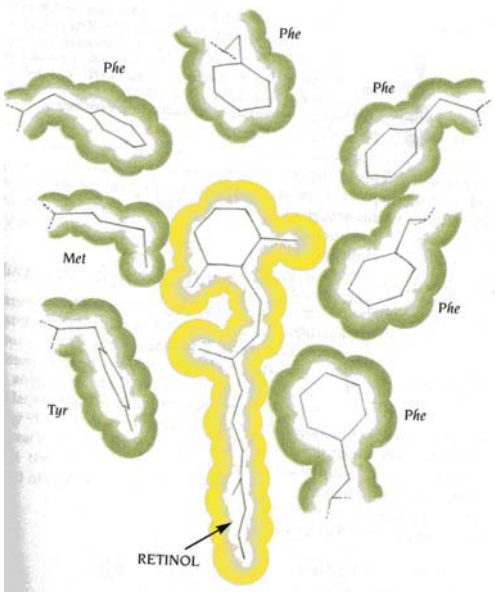


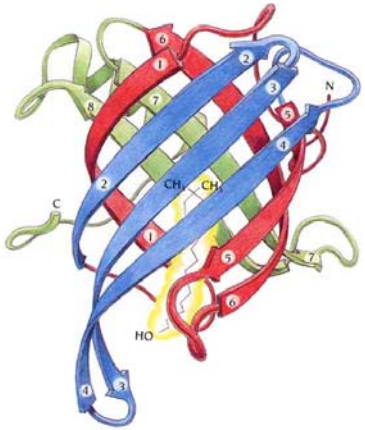


Retinol-binding protein

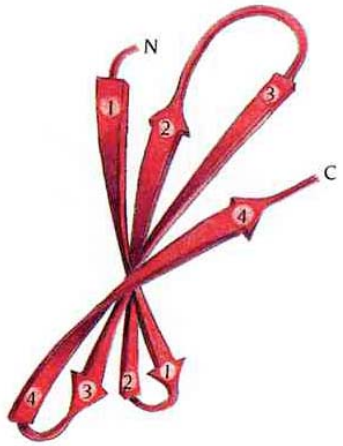
Amino acid sequence reflects β structure

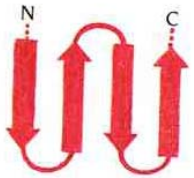
strand no.	residue no.	amino acid sequence
2	41–48	– Ile – Val – Ala – Glu – Phe – Ser – Val – Asp –
3	53–60	– Met – Ser – Ala – Thr – Ala – Lys – Gly – Arg –
4	71–78	– Ala – Asp – Met – Val – Gly – Thr – Phe – Thr –

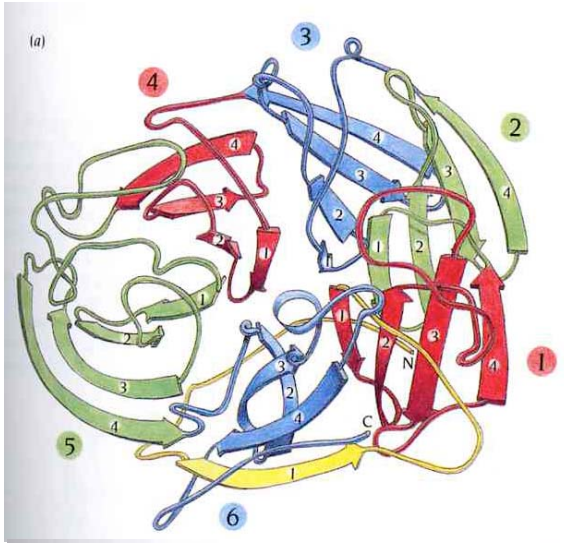


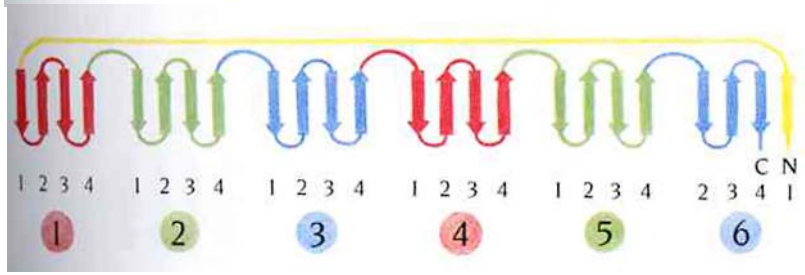


2. Up-and-down β sheets



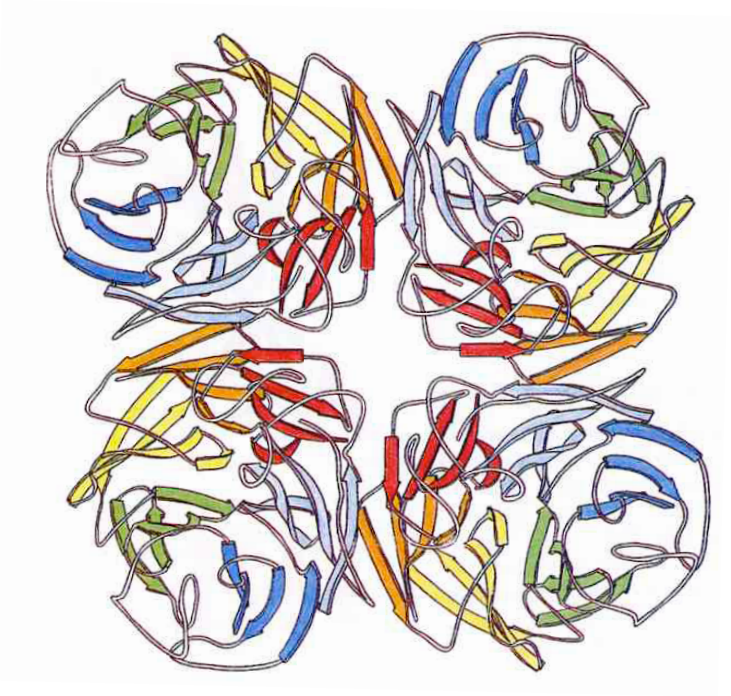


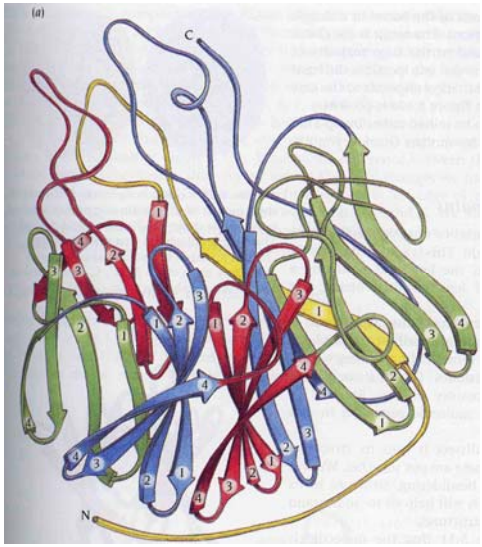


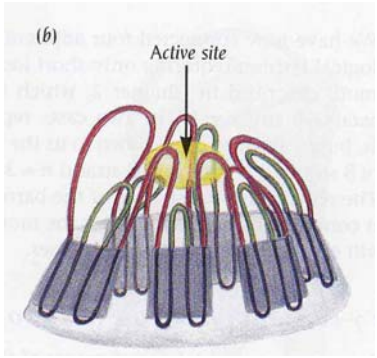


Neuraminidase headpiece (Influenza virus)

Six-blade propeller

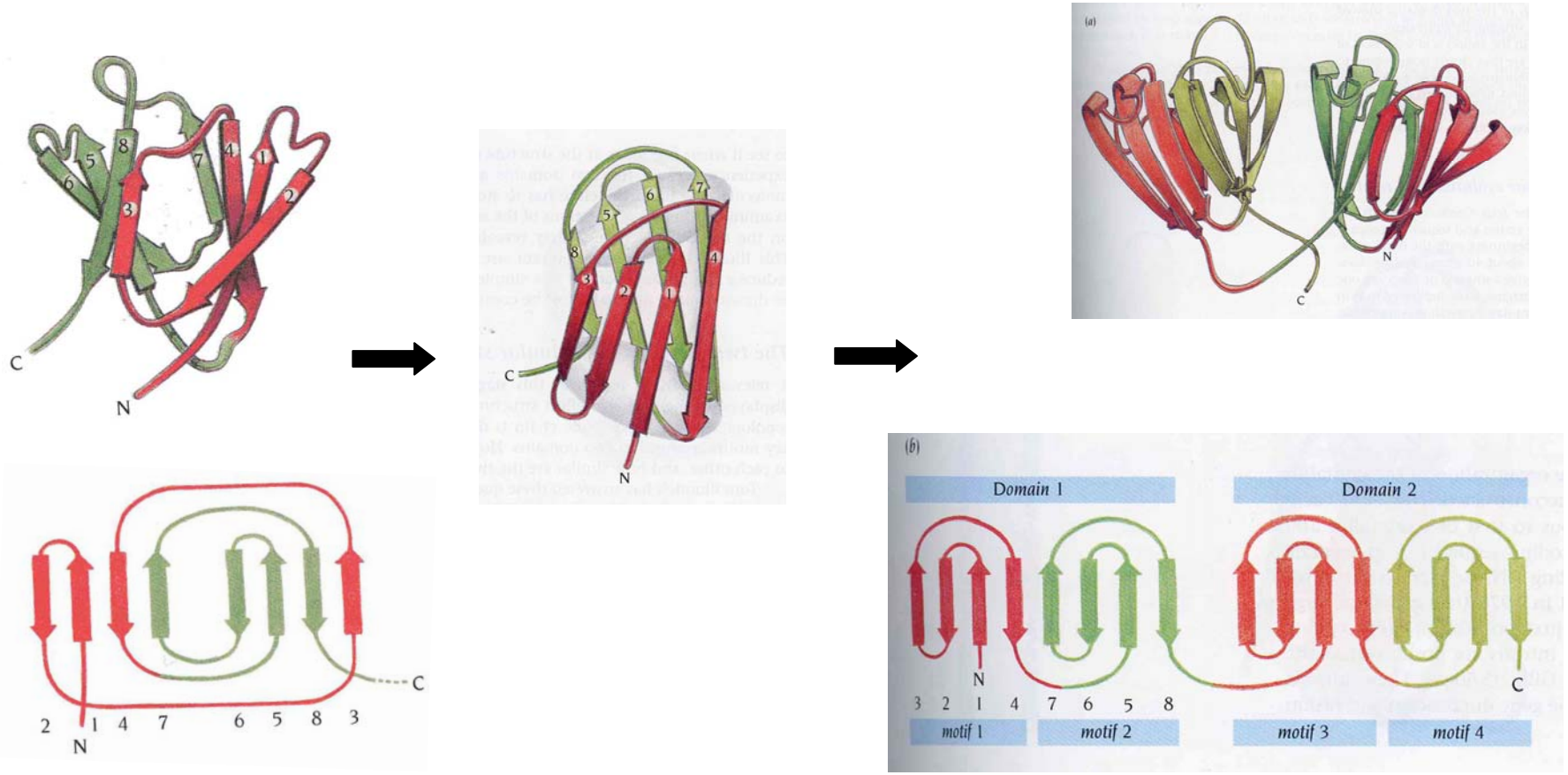




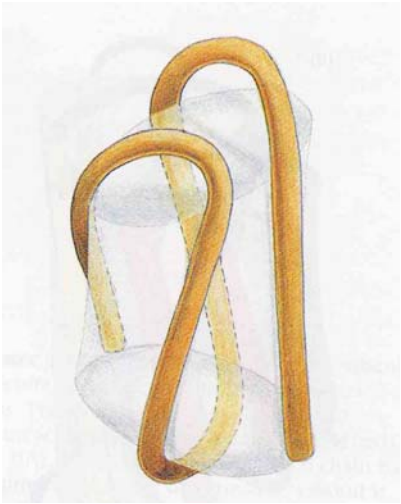


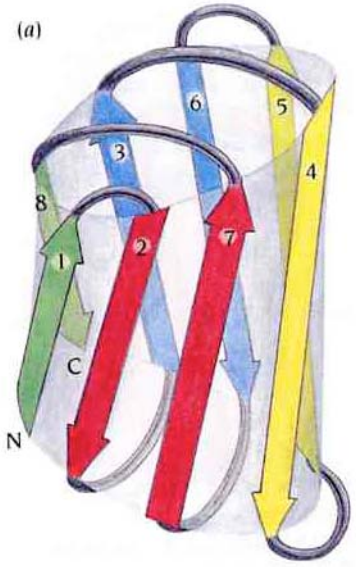
Active site: one side of the propeller

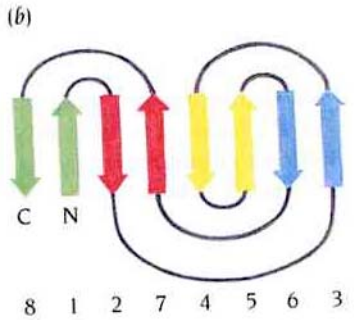
Two Greek key motifs form the domain



Two Greek key motifs can form jelly roll barrels

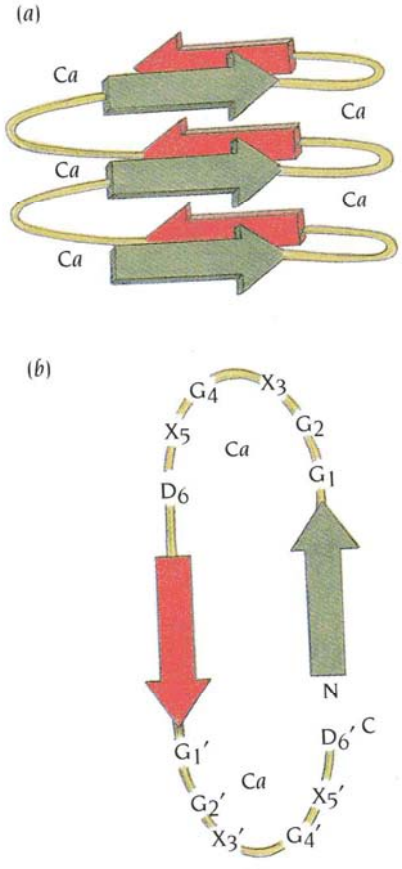


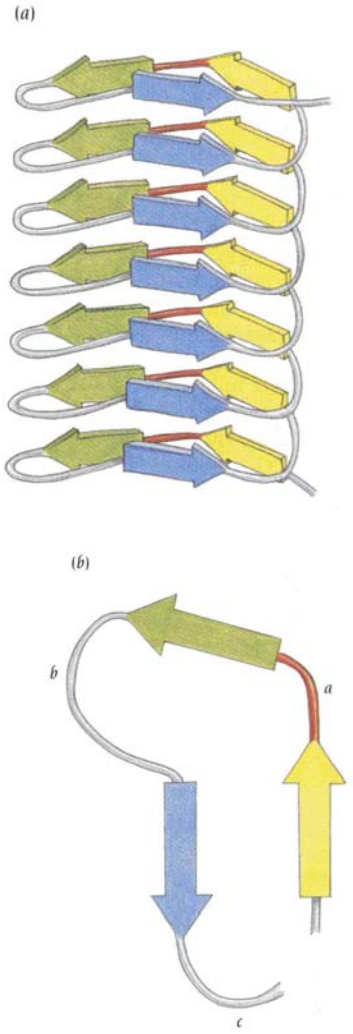




Basic pattern: a piece of string wrapped around a barrel

Parallel β-helix domains

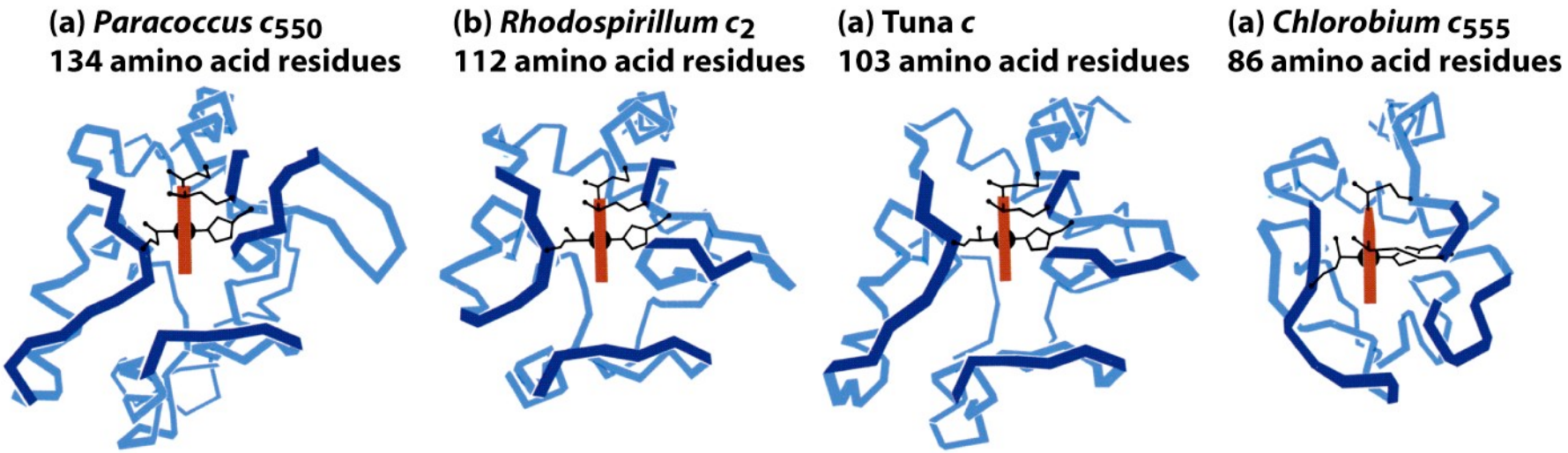




Two-sheet β helix
18 residues: β-loop-β-loop strutcure

Three-sheet β helix

Protein families

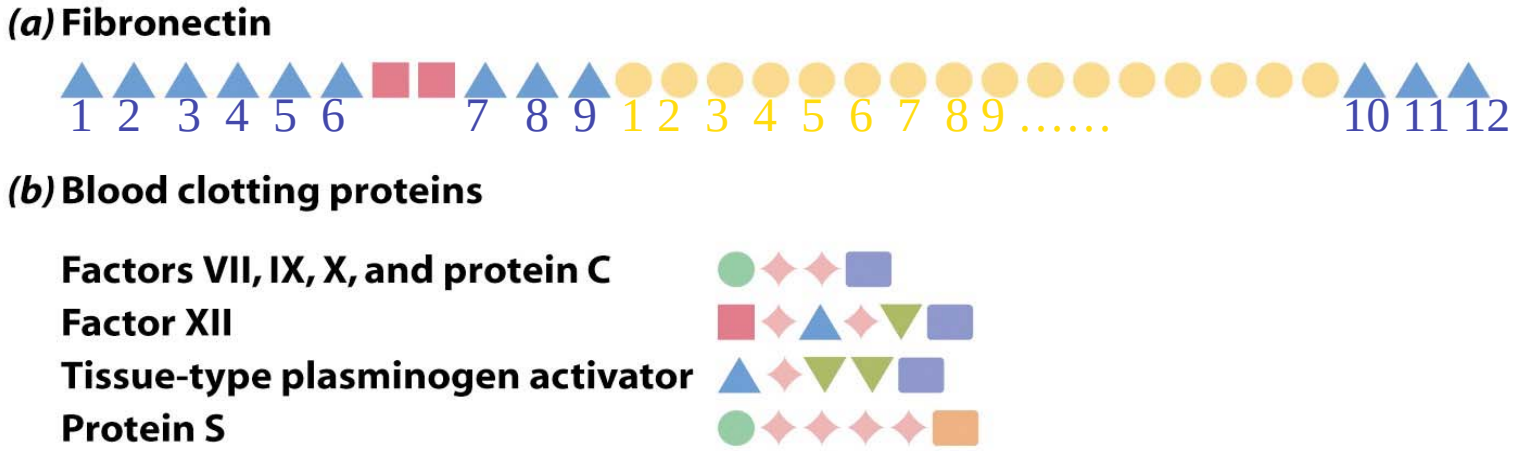


c-type cytochromes from different species exhibit only low degrees of sequence similarity to each other and to eukaryotic cytochrome c.
Yet, their 3D structures are similar, particularly in polypeptide chain folding and side chain packing in the protein interior (hydrophobic core).

It appears that the essential structural and functional elements of proteins, rather than their amino acid residues, are conserved during evolution.

Large polypeptide chains fold into several domains

Modular Construction of some proteins



Modules (sequence motifs):

~ 40 -100 residues

FIBRONECTIN TYPE I MOTIF

Motif = C 6-8X Y 5X (WYF) 10-12X C X C X G 4-7X C

Fn	#1	P G C Y D N - - G K H Y Q I N Q Q W E R T Y - L G - N V L V C T C Y G G S R G - F N C E S K P E E A
	2	E T C F D K Y T G N T Y R V G D T Y E R P K D S - - M I W D C T C I G A G R G R I S C T - I A
	3	N R C H E G - - G Q S Y K I G D T W R R P H E T G G Y M L E C V C L G N G K G E W T C K P I
	4	E K C F D H A A G T S Y V V G E T W E K P Y - Q G W M M V D C T C L G E G S G R I T C T S R
	5	N R C N D Q D T R T S Y R I G D T W S K K D N R G N - L L Q C I C T G N G R G E W K C E R H T S V Q
		G H C V T D S - G V V Y S V G M Q W L K T Q - - G N K Q M L C T C L G N G - - - V S C Q E
	7	F I C T T N F - G V M Y R I G D Q W D K Q H D M G - H M M R C T C V G N G R G E W T C Y A Y S Q I R
	8	D Q C I V D - - D I T Y N V N D T F H K R H E E G - H M L N C T C F G Q G R G R W K C D P V
	9	D Q C Q D S E T G T F Y Q I G D S W E K Y V H - G V R - Y Q C Y C Y G R G I G E W H C Q P
	10	D S C F D P Y T V S H Y A V G D E W E R M S E S G F K - L L C Q C L G F G S G H F R C D S S
	11	R W C H D N - - G V N Y K I G E K W D R Q G E N G - Q M M S C T C L G N G K G E F K C D P H E
	12	A T C Y D D - - G K T Y H V G E Q W Q K E Y L - G - A I C S C T C Q G G Q R G W R C D N C R
tPA		V I C R D E K T Q M I Y Q Q H Q S W L R P V L R S N R V E Y C W C - - N S - G R A Q C H S V
Factor XII		C F E P Q L L R F F H K N E I W Y R T - - E Q A A V A R C Q C K G - - - P D A H C

Type I

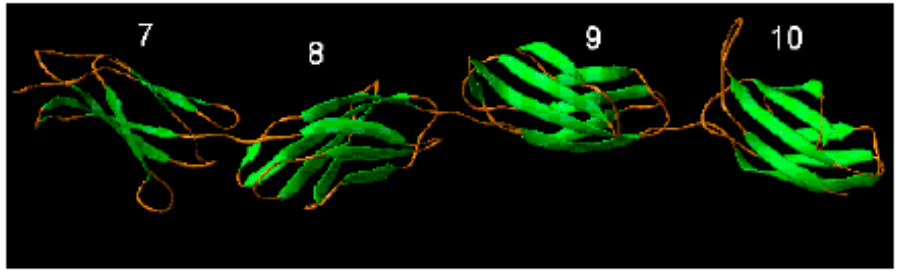
44-48 a.a.
2 disulfides
also found in:
Coag. Factor XII
tissue Pmg Activator

Type II

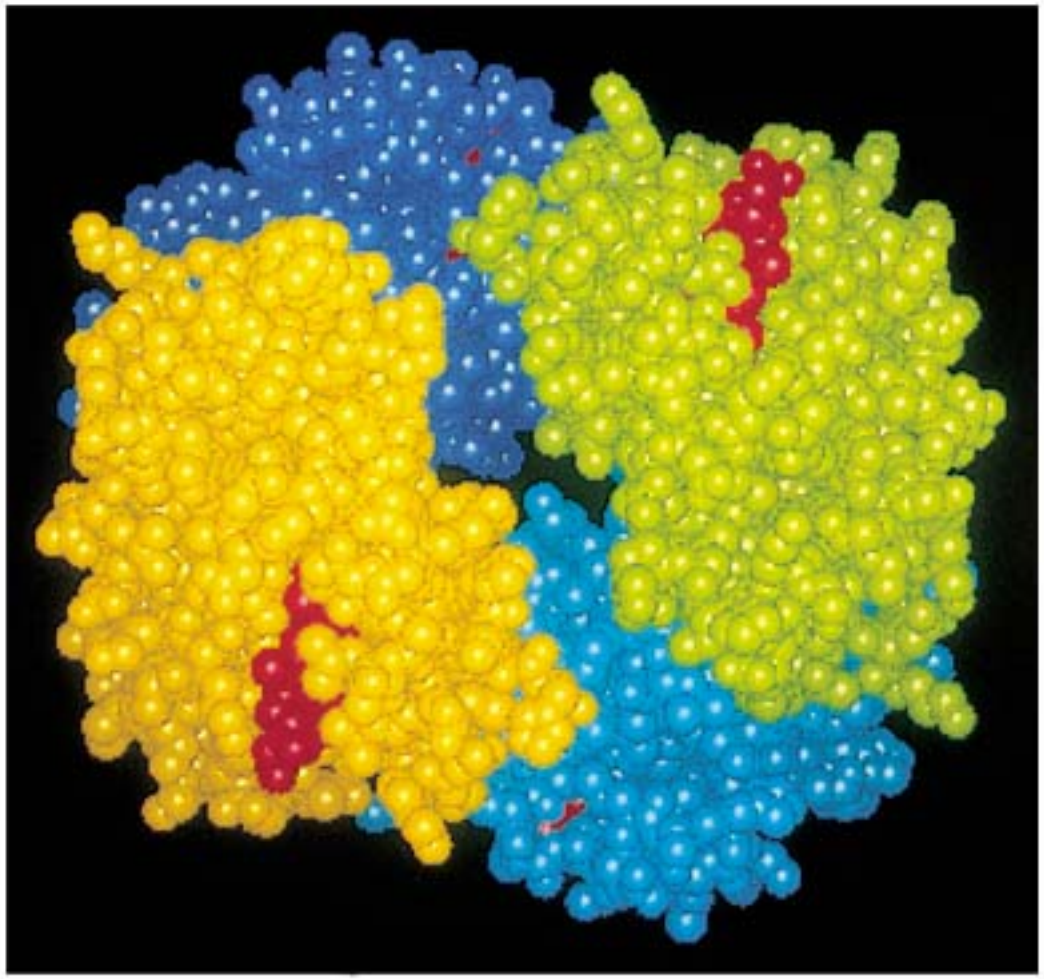
~ 60 a.a.
2 disulfides
also found in:
bovine Seminal Plasma Proteins
MMPs
Coag Factor XII
mannose-6-PO_4 receptors
and others

Type III

87-96 a.a.
no disulfides
also found in:
2% of animal proteins



Quaternary protein structure



4^o structure is the relative placement of different poly-peptide segments

Hemoglobin is shown to the left (α_1-yellow, α_2-green, β_1-cyan, β_2-blue), heme groups are in red - bind O_2

Subunits usually associate noncovalently