

Structural hierarchy of proteins

László Smeller
Semmelweis University
Dept. Biophysics and Radiation Biology

Classification of proteins

According to shape and solubility

- **Fibrous proteins (E.g.: collagen...)**
These proteins have a rod like structure. They are not soluble in water.
- **Globular proteins (E.g.: hemoglobin, myoglobin...)**
These proteins more or less spherical in nature. Due to their distribution of amino acids (hydrophobic inside, hydrophilic outside) they are very soluble in aqueous solution.
- **Membrane proteins (E.g.: rhodopsin...)**
These are protein which are in association with lipid membranes. Those membrane proteins that are embedded in the lipid bilayer have extensive hydrophobic amino acids that interact with the non-polar environment of the bilayer interior. Membrane proteins are not soluble in aqueous solution. Rhodopsin is an integral membrane protein and is embedded in the bilayer.

Classification of proteins

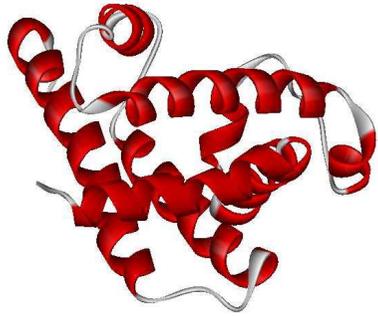
According to the biological function

- Enzymes (E.g.: trypsin, cytochrome-c...)
- Transport proteins (E.g.: hemoglobin...)
- Chaperone proteins
- Toxins (E.g.: ricin, snake venom...)
- Hormones (E.g.: insulin, growth hormone...)
- Contractile proteins (E.g.: myosin, actin...)
- Structure proteins (E.g.: collagen, elastin...)
- Storage proteins (E.g.: ovalbumin, casein, ferritin...)
- Others (E.g.: histone proteins...)

Classification of proteins

According to secondary structure

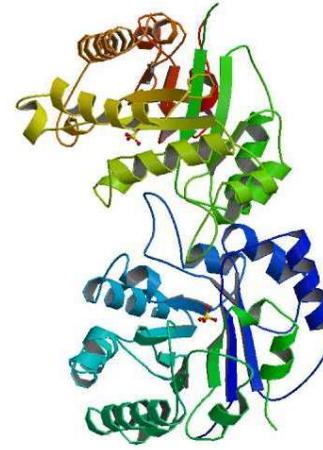
- **All Alpha helical (E.g.: myoglobin...)**
Proteins that contain only (in some exceptional instances there may be isolated beta-sheets) alpha helical secondary structure.
- **Alpha/beta (E.g.: Triose phosphate-isomerase...)**
Proteins that contain alternating alpha-helical and beta-sheet secondary structure elements. The alternating helical and sheet segments may form a closed barrel like structure.
- **Alpha+beta (E.g.: ribonuclease...)**
In these proteins the alpha helical and beta sheet regions occur in independent regions of the molecule.
- **Beta sheet (E.g.: tenascin...)**
Protein that contain only (in some exceptional instances there may be isolated alpha helices) beta-sheet secondary structure.



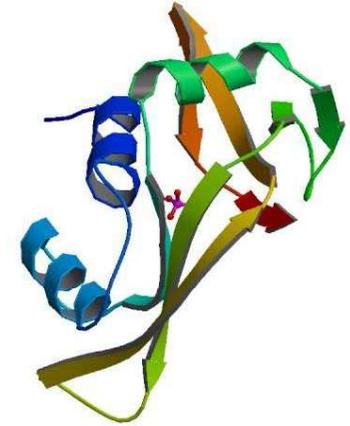
Myoglobin
All alpha



Tenascin
Beta



Alpha/beta structure (Triose phosphate isomerase)



Alpha+beta (ribonuclease)

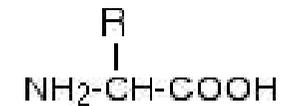
Structural hierarchy

- Primary structure
- Secondary structure
- Tertiary structure
- Quaternary structure

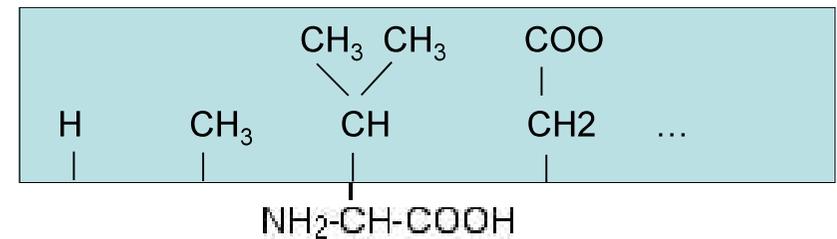
Supramolecular organizations

Amino acids:
the building blocks of proteins

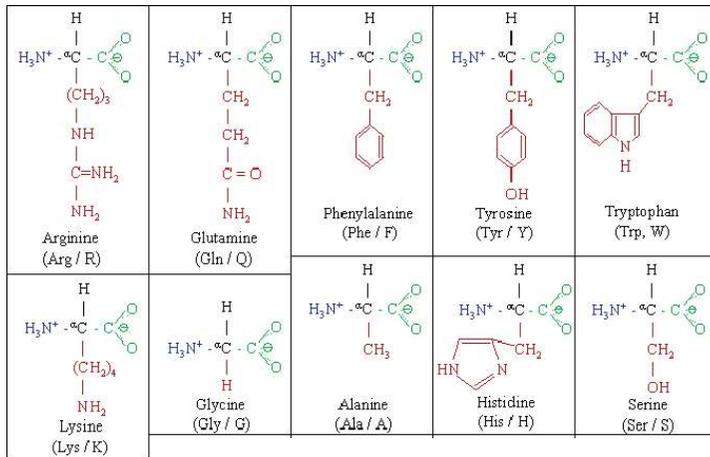
General structure
of amino acids:



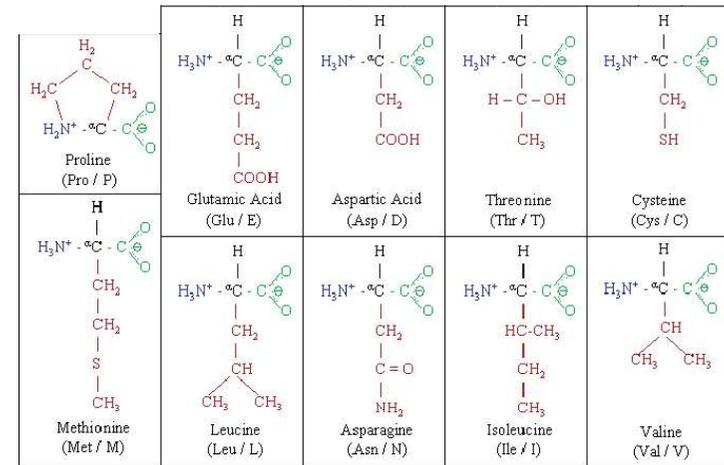
Structural variability:



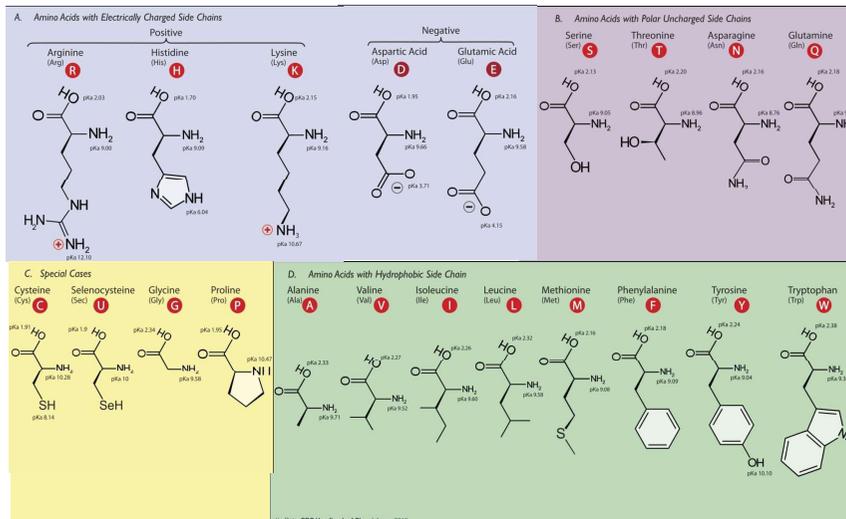
Amino acids of the proteins



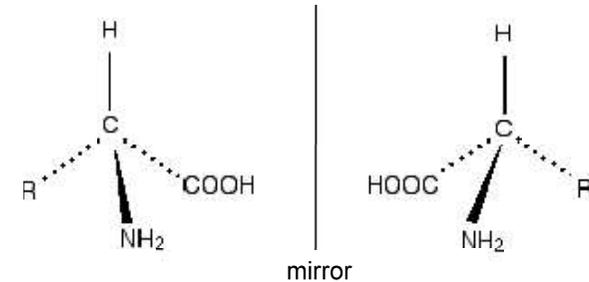
Amino acids of the proteins



Properties of amino acids



Chirality



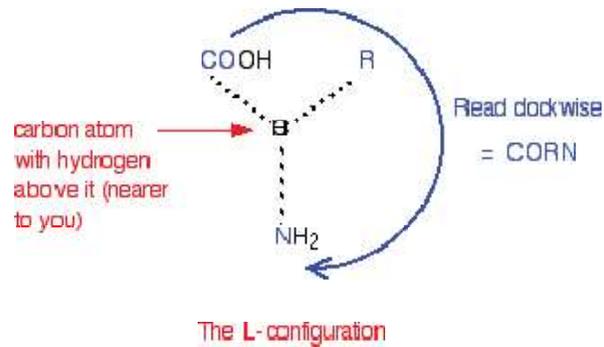
Chirality center:
a carbon atom, connected to
four different atoms.

→ Optical activity
(rotation of the plane of polarization
of linearly polarized light)

Hand:



D and L enantiomers

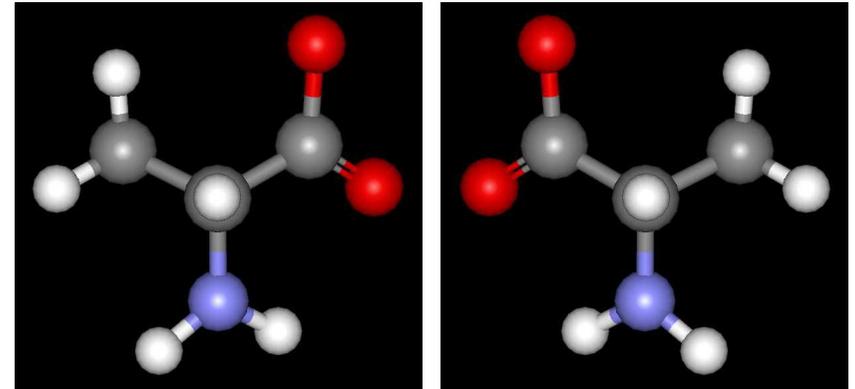


In living systems: L type !

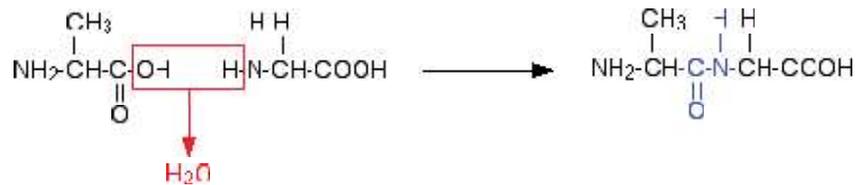
No direct connection between the L-D enantiomer forms and the direction of the rotation of the polarization plane.

E.g.: (+)alanine (-)cysteine (-)tyrosine (+)valine

D and L enantiomers of alanine



Connection of the amino acids: the peptide bond



Peptide 2.. c.a 20 amino acids

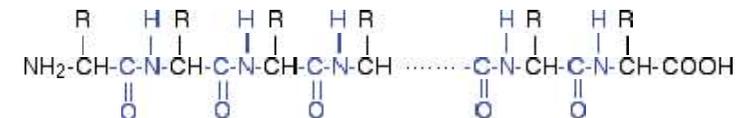
Protein: more than 20 amino acids

The primary structure

Primary structure:

the amino acid sequence in the polypeptide chain

Which direction? N terminal -> C-terminal



E.g.: (myoglobin, 1YMB)

GLY LEU SER ASP GLY GLU TRP GLN GLN VAL LEU ASN VAL...

... ALA LYS TYR LYS GLU LEU GLY PHE GLN GLY

Example: Myoglobin

Primary structure with 3-letter code (153 aa.):

```

GLY LEU SER ASP GLY GLU TRP GLN GLN VAL LEU ASN VAL
TRP GLY LYS VAL GLU ALA ASP ILE ALA GLY HIS GLY GLN
GLU VAL LEU ILE ARG LEU PHE THR GLY HIS PRO GLU THR
LEU GLU LYS PHE ASP LYS PHE LYS HIS LEU LYS THR GLU
ALA GLU MET LYS ALA SER GLU ASP LEU LYS LYS HIS GLY
THR VAL VAL LEU THR ALA LEU GLY GLY ILE LEU LYS LYS
LYS GLY HIS HIS GLU ALA GLU LEU LYS PRO LEU ALA GLN
SER HIS ALA THR LYS HIS LYS ILE PRO ILE LYS TYR LEU
GLU PHE ILE SER ASP ALA ILE ILE HIS VAL LEU HIS SER
LYS HIS PRO GLY ASP PHE GLY ALA ASP ALA GLN GLY ALA
MET THR LYS ALA LEU GLU LEU PHE ARG ASN ASP ILE ALA
ALA LYS TYR LYS GLU LEU GLY PHE GLN GLY
    
```

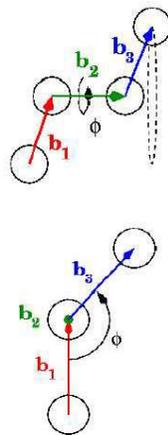
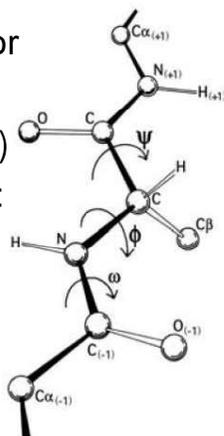
Rotation around the chemical bonds of the backbone

3 backbone bonds for each AA

1 rigid (delocalization)

2 rotation possible:

Φ , Ψ dihedral angles



2N rotational degree of freedom

Example: Myoglobin

Primary structure with one letter code (153 aa.):

```
>1YMB:A|PDBID|CHAIN|SEQUENCE
```

```

GLSDGEWQQVNLVWGKVEADIAGHGQEVLRIRLFTGHPETLEKFDKFKHLKTEAE
MKASEDLKKHGTVVLTALGGILKKKGHHEAELKPLAQSHATKHKIPIKYLEFIS
DAIIHVLHSHKHPGDFGADAQGAMTKALELFRNDIAAKYKELGFQG
    
```

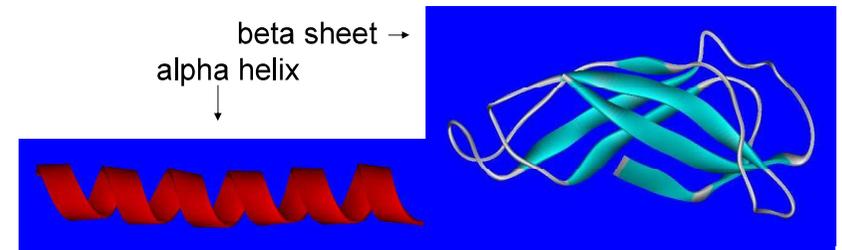
(FASTA format)

Secondary structure

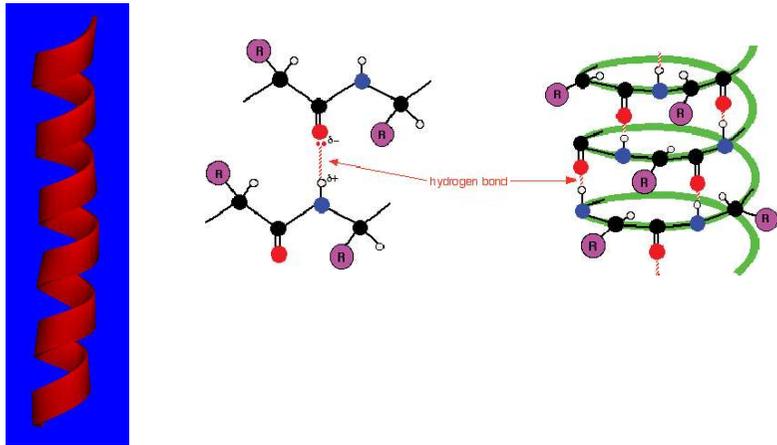
Regular local structures of linear segments of polypeptide chains.

They are stabilized by H-bonds

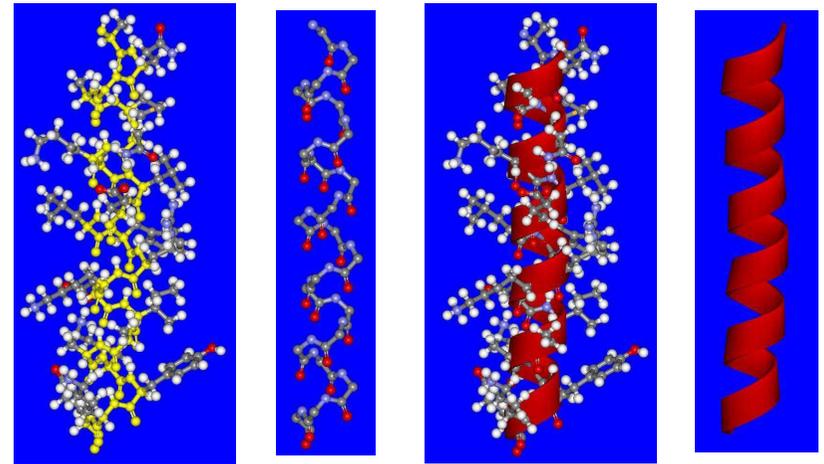
Typical forms:



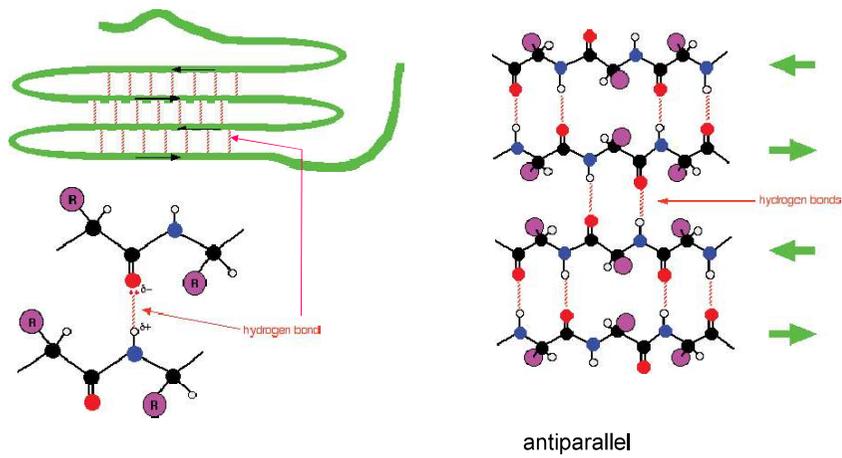
Alpha helix



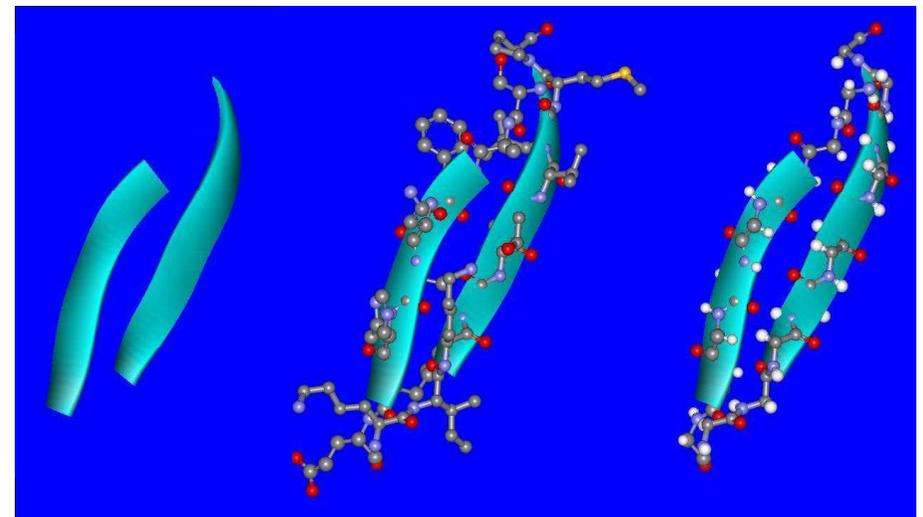
Alpha helix



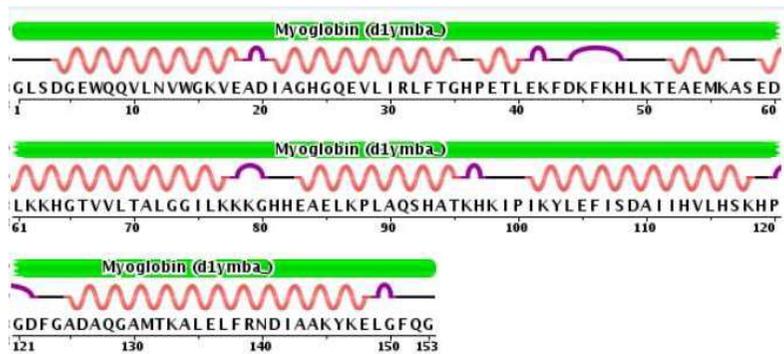
Beta sheet



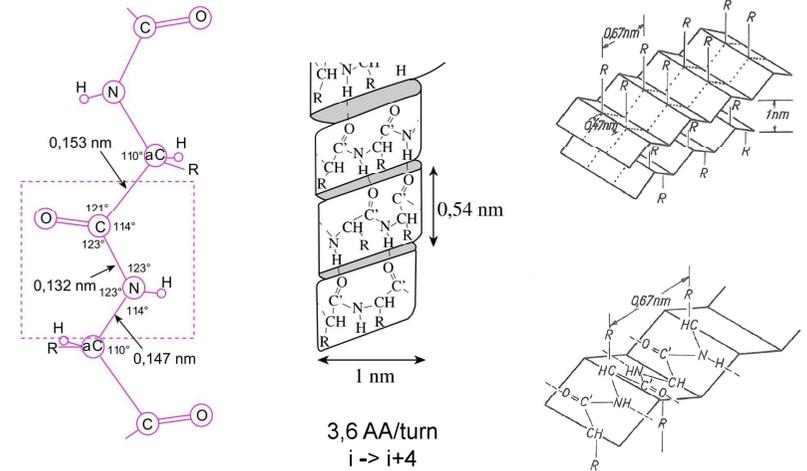
Beta sheet



One dimensional representation of the secondary structure



Size



Stabilization by H-bonds

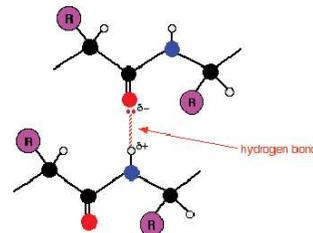
12-30 kJ/mol

Cf: Covalent bond: 200 kJ/mol

van der Waals: 1-2 kJ/mol

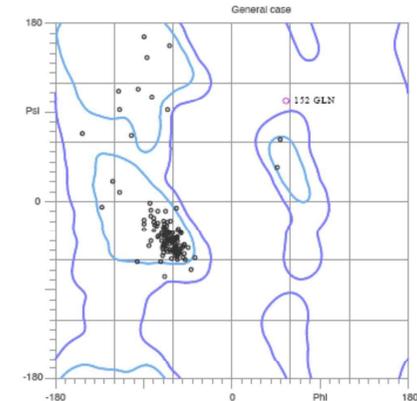
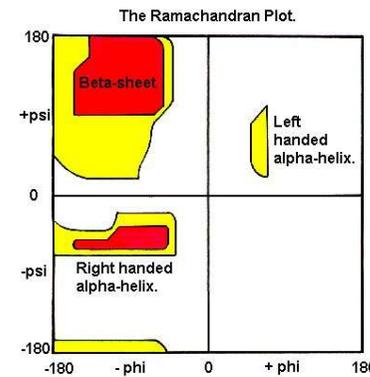
thermal energy (RT):

2.5 kJ/mol (T=300K)



Boltzmann factor: $e^{-\frac{\Delta E}{RT}} = 0.000335 = \frac{1}{2981} \approx \frac{1}{3000}$
 ($\Delta E=20\text{kJ/mol}$)

Ramachandran plot



1YMB

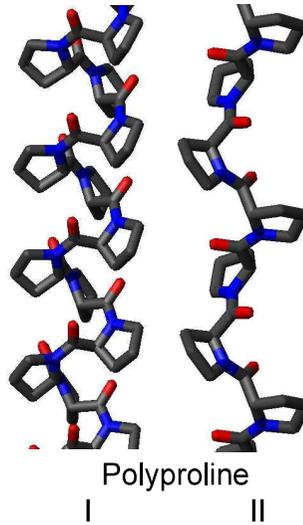
Special helices

3_{10} -helix* $i \rightarrow i+3$ (10 atom)

π -helix $i \rightarrow i+5^*$

Polyproline I helix cis

Polyproline II helix** trans



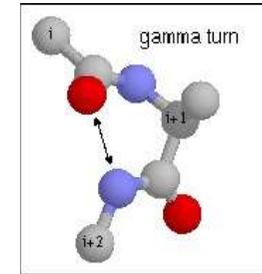
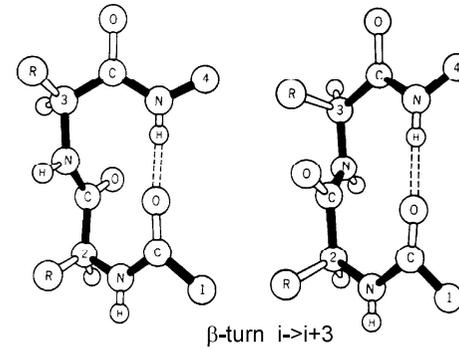
* α -helix: $i \rightarrow i+4$ $3,6_{16}$ helix
** in water

Other nonhelical structures

Loops and turns

(loop)

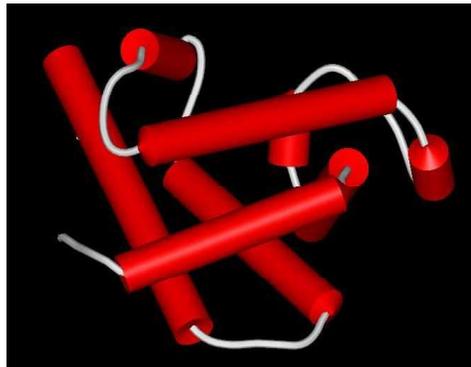
(turn)



γ -turn $i \rightarrow i+2$

Tertiary structure

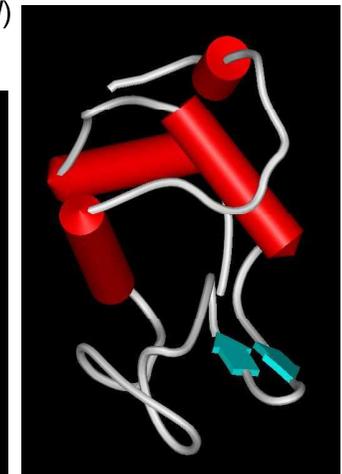
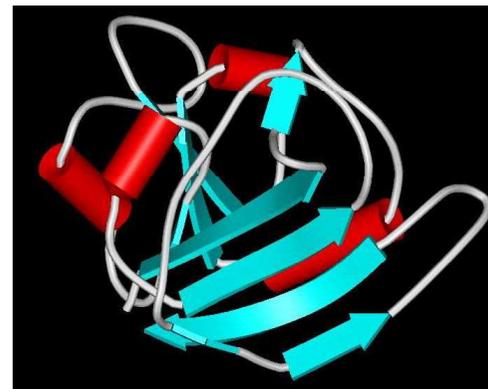
Overall topology of the folded polypeptide chain
(Organization of the secondary structure elements)



Myoglobin

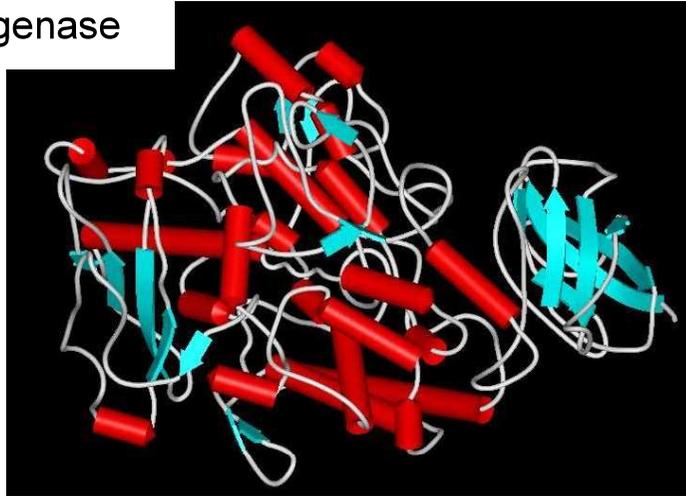
Examples

Lysozyme (HEW)
Dihydrofolate reductase

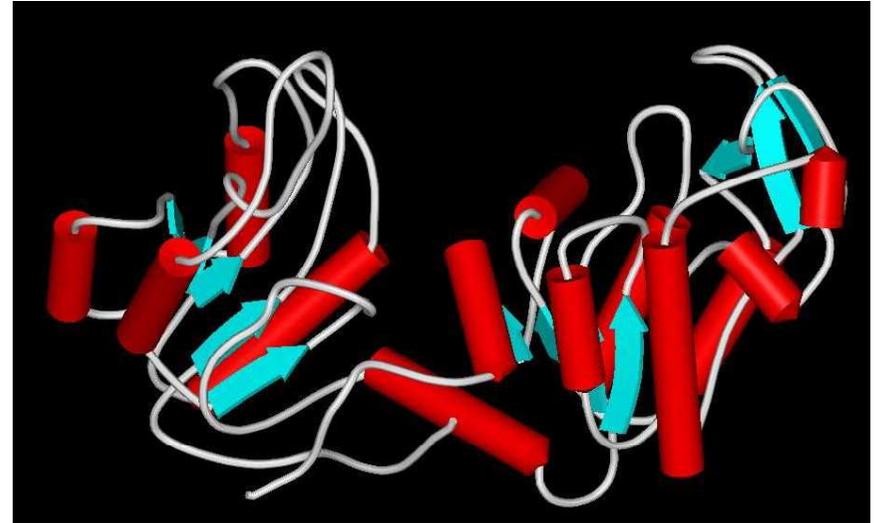


Examples

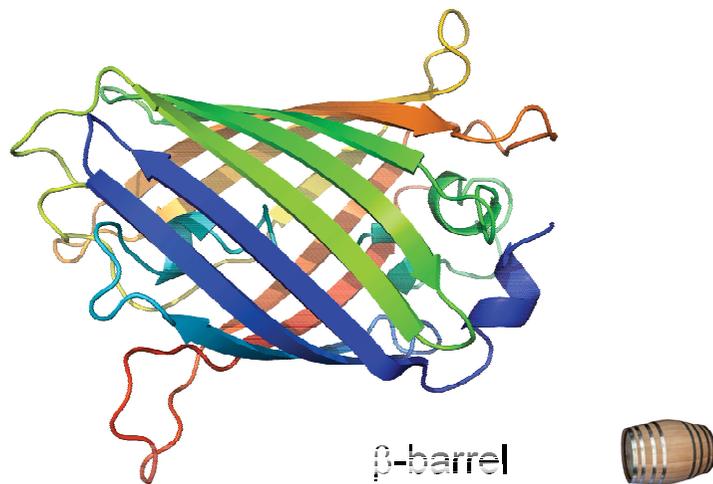
Lipoxygenase



Examples: Phosphoglycerate-kinase (PGK)



Examples: GFP



Stabilization of the tertiary structure

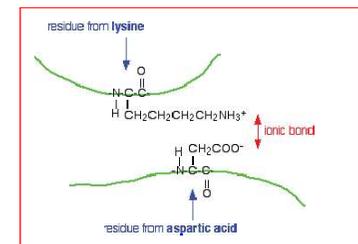
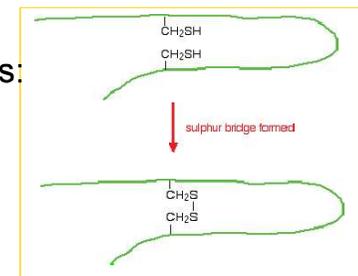
Between the side chains:

disulfide bond

ionic bonds

H-bond

Van der Waals int.



Stabilization of the tertiary structure

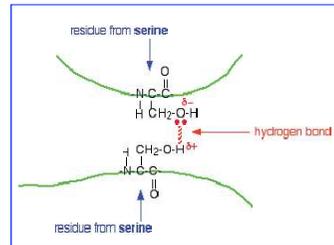
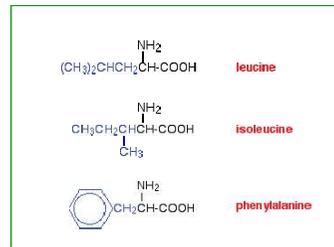
Between the side chains:

disulfide bond

ionic bonds

H-bond

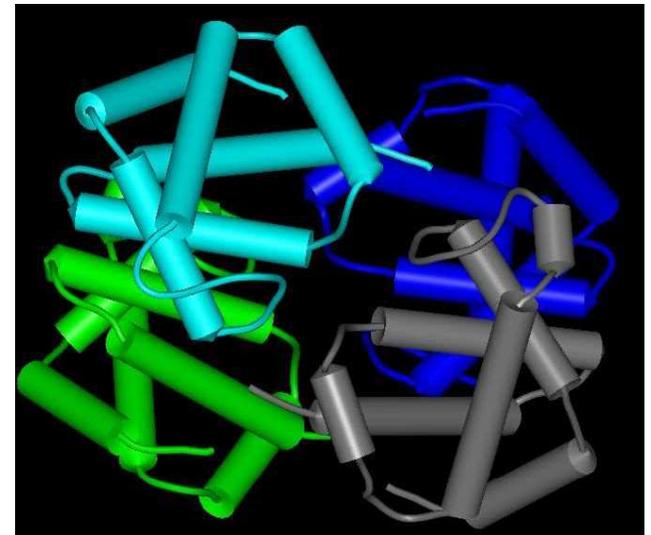
Van der Waals int.



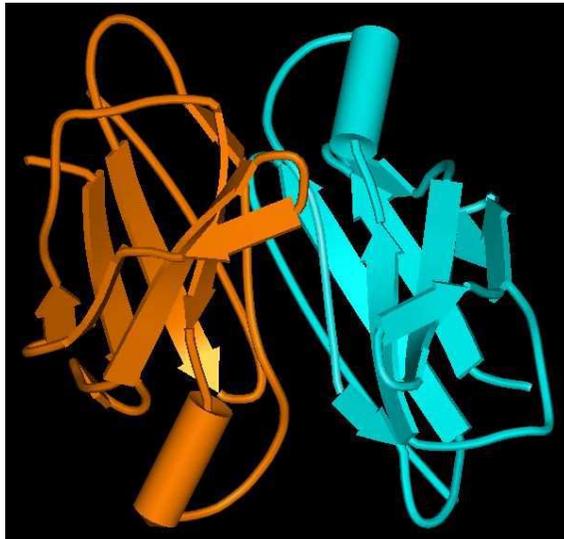
Quaternary structure

only for
proteins with
more than
one
polypeptide
chains.

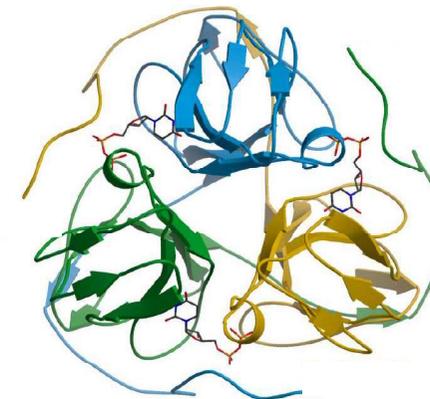
E.g.:
Hemoglobin
tetramer



Further examples: Transthyretin



Further examples: DUTPase

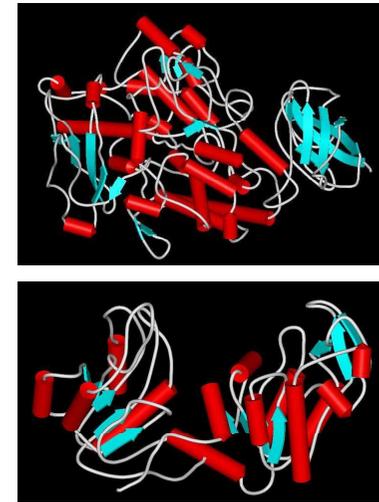


Important further aspects of the protein structure

- Domain
- Prosthetic group
- Posttranslational modifications
- Active site
- Pocket

Domain

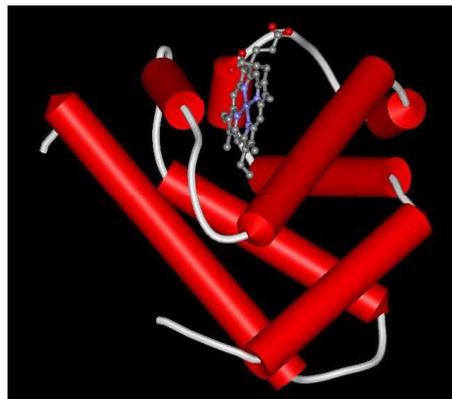
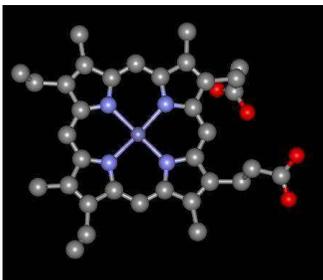
Part of the protein, which can fold into an ordered structure. Its structure is stable, it can function without the presence of the rest of the protein. The different domains of a protein may have different functions: e.g.: ATP binding domain, etc.



Prosthetic group

A non-protein chemical compound that is required for an enzyme's activity. They are bound strongly to the protein.

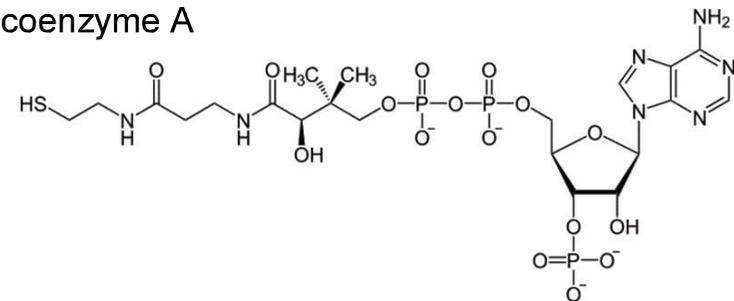
E.g.: hem group



Coenzymes

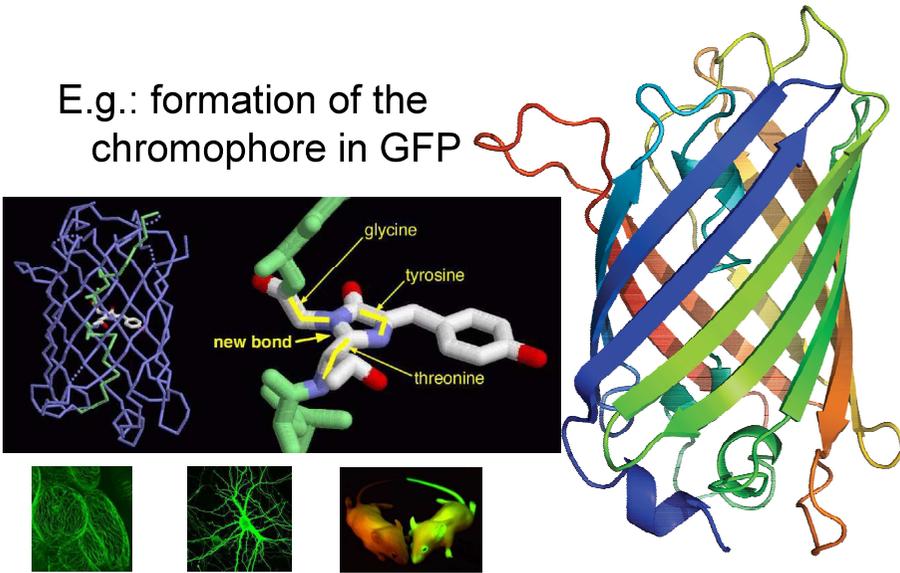
A cofactor is a non-protein chemical compound that is required for an enzyme's activity. They bind weakly and reversibly.

PI: coenzyme A



Posttranslational modifications

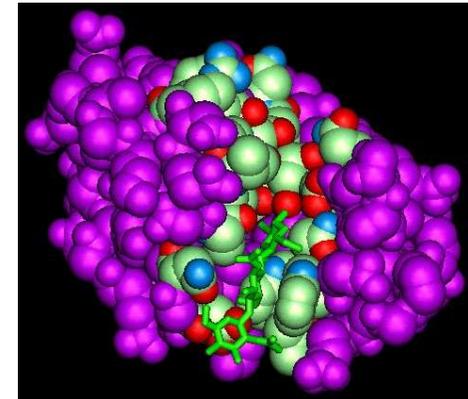
E.g.: formation of the chromophore in GFP



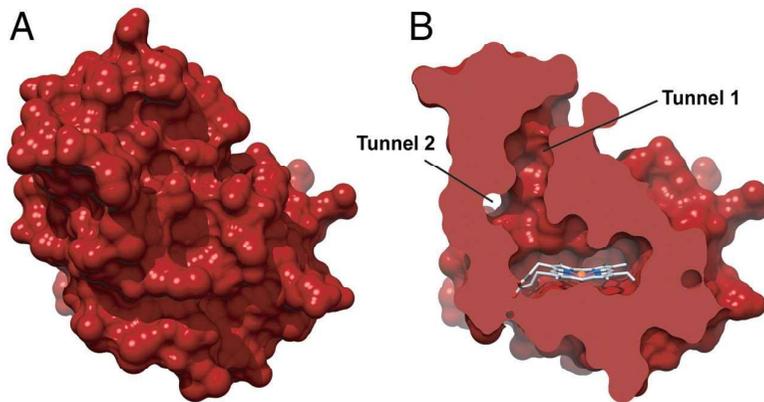
Active site

Active site is the region of an enzyme where substrate molecules bind and undergo a chemical reaction

Binding site
Catalytic site



Hem pocket



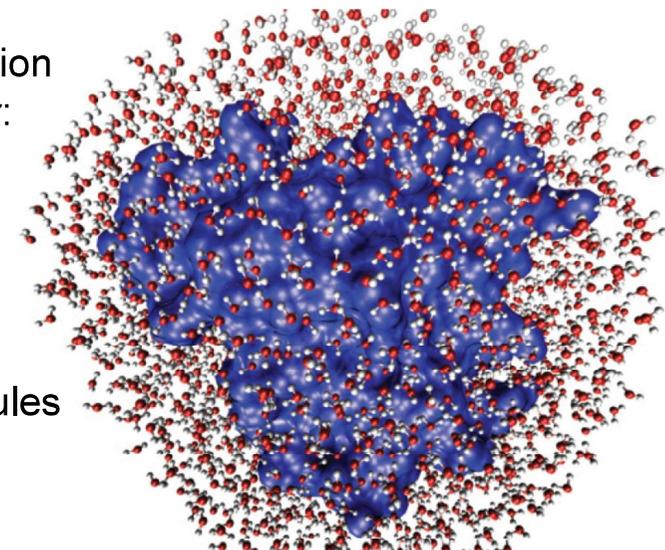
hem nitric oxide/oxygen binding (H-NOX) domain

Winter M B et al. PNAS 2011;108:E881-E889

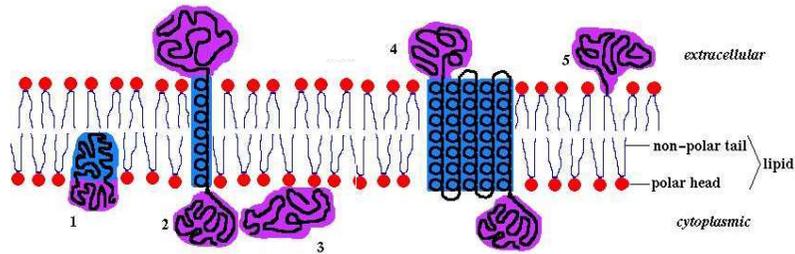
Role of the water

Hydration layer:

2-3
Water
molecules



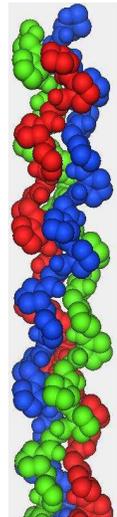
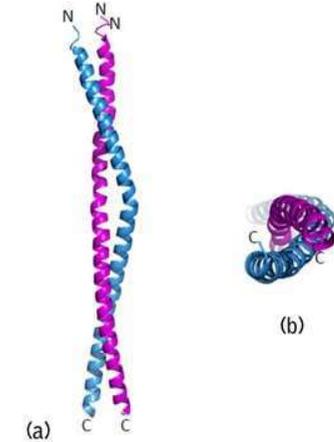
Membrane proteins



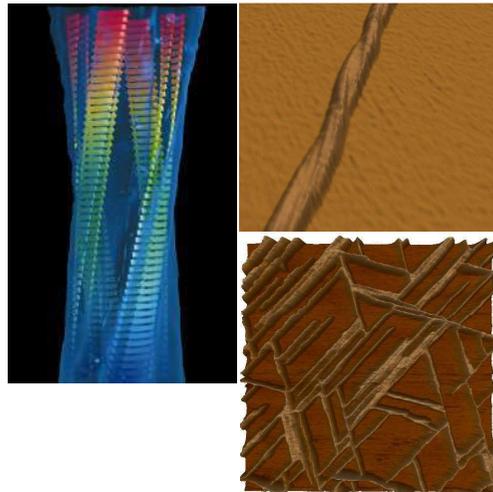
Domains with **hydrophobic** surfaces
 hydrophilic

Supramolecular organizations

- Coiled coil
- Collagen
- Fibrillar structures



Collagen



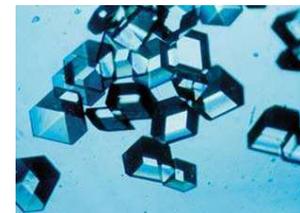
Fibrillar organization

Some methods for determination of the 3D structure of proteins

X-ray crystallography

NMR

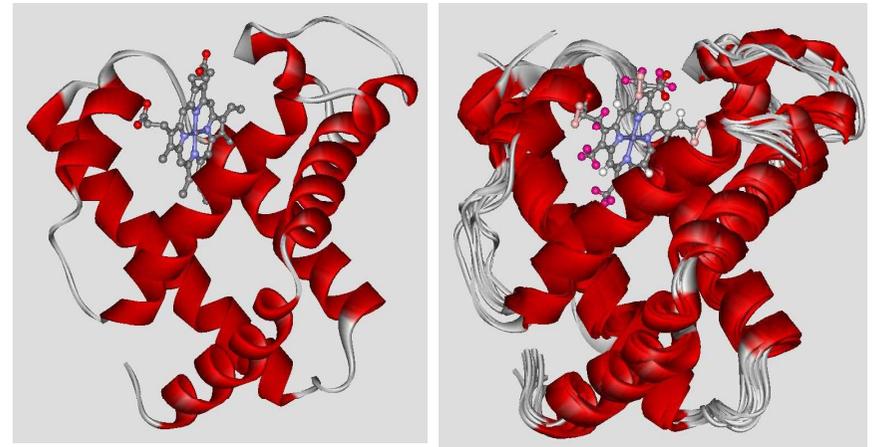
Prediction (homology modelling)



Spectroscopic methods sensitive to the changes of the protein structure

- Circular dichroism (CD)
- Infrared spectroscopy (IR, FTIR)
- Luminescence spectroscopy
- UV absorption spectroscopy
- ...

Crystallography <-> NMR



myoglobin

Protein databases

- PDB
 - Protein Data Bank
 - 3D Structures (c.a.150 000) from
 - X-ray and
 - NMR experiments
- Swiss-prot
 - Protein sequences
 - Proteomics software
 - Structure prediction (homology modeling)
 - Calculation/Estimation of the chemical parameters (e.g. isoelectric point...)
 - Comparison of the sequences...

PDB database: proteins 3D structures based on x-ray crystallography and NMR spectroscopy

RCSB PDB - Mozilla Firefox

26 Feb 2019: 149174

https://www.rcsb.org/

RCSB PDB

149174 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education

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

Go

Advanced Search | Browse by Annotations

Welcome

Deposit

Search

Visualize

Analyze

A Structural View of Biology

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

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

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

Job Opportunity with RCSB PDB at Rutgers

February Molecule of the Month

References

<http://www.molecularmodels.ca/molecule/modelfiles/jb16alan.html>
<http://www.chemguide.co.uk/organicprops/aminoacids/background.html#top>
<http://www.chemguide.co.uk/organicprops/aminoacids/proteinstruct.html>
http://www.enzim.hu/~vertessy/kovari_phd.pdf
<http://www.pdb.org/>
T. E Chreighton: Proteins, Freeman and Company, New York
Orvosi Biofizika Szerk: Damjanovich, Fidy, Szöllősi
Tarján Imre: A biofizika alapjai
Elődi Pál: Biokémia
<http://www.pnas.org/content/108/43/E881.full>
<http://imtech.res.in/raghava/>
<http://www.cryst.bbk.ac.uk/PPS2/course/>
<http://www.cryst.bbk.ac.uk/PPS95/course/>
http://mkk.szie.hu/dep/aet/tanweb/Fogalomtar/index_fogalom.htm
<http://www.med.upenn.edu/shorterlab/research.html>