



Amino Acids, Polypeptides and Proteins

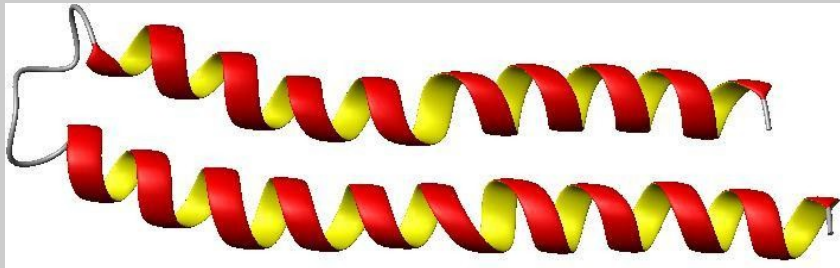
Homayoun Valafar

Department of Computer Science and Engineering,
USC



Protein

- Proteins are functional elements of a cell
- Proteins are made of 20 “amino acid” subunits
- Proteins fold to create a their own characteristic fold (three dimensional shape)

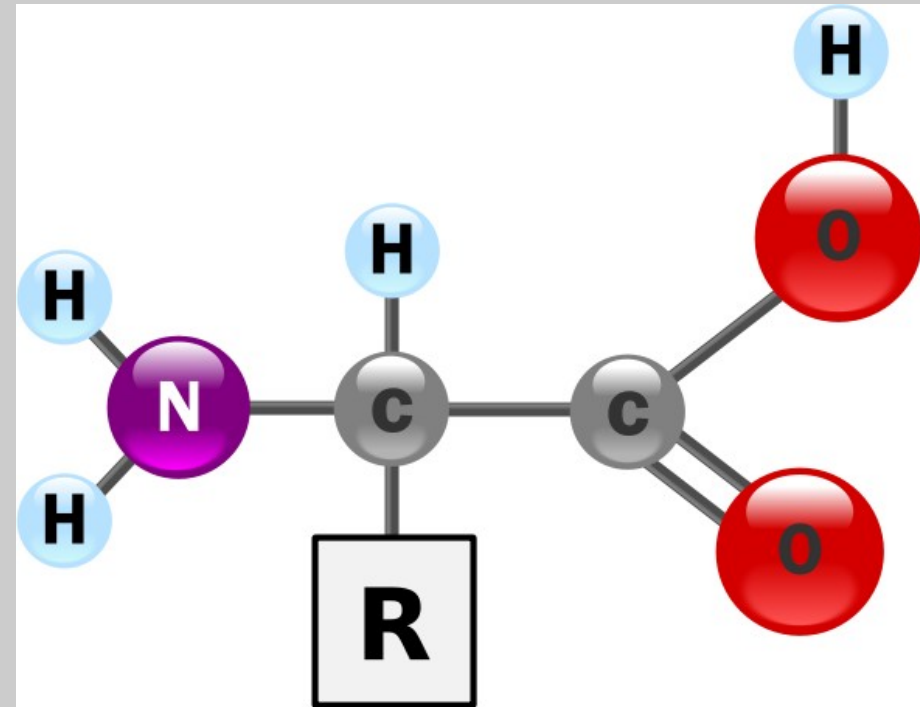


$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ (\text{CH}_2)_3 \\ \\ \text{NH} \\ \\ \text{C}=\text{NH}_2 \\ \\ \text{NH}_2 \end{array}$ <p>Arginine (Arg / R)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{C}=\text{O} \\ \\ \text{NH}_2 \end{array}$ <p>Glutamine (Gln / Q)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH}_2 \\ \\ \text{C}_6\text{H}_5 \end{array}$ <p>Phenylalanine (Phe / F)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH}_2 \\ \\ \text{C}_6\text{H}_4 \\ \\ \text{OH} \end{array}$ <p>Tyrosine (Tyr / Y)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH}_2 \\ \\ \text{C}_8\text{H}_6\text{N} \\ \\ \text{H} \end{array}$ <p>Tryptophan (Trp / W)</p>
$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ (\text{CH}_2)_4 \\ \\ \text{NH}_2 \end{array}$ <p>Lysine (Lys / L)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{H} \end{array}$ <p>Glycine (Gly / G)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH}_3 \end{array}$ <p>Alanine (Ala / A)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH}_2 \\ \\ \text{C}_4\text{H}_3\text{N} \end{array}$ <p>Histidine (His / H)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH}_2 \\ \\ \text{OH} \end{array}$ <p>Serine (Ser / S)</p>
$\begin{array}{c} \text{H}_2 \\ \\ \text{C} \\ / \quad \backslash \\ \text{H}_2\text{C} \quad \text{CH}_2 \\ \quad \quad \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \end{array}$ <p>Proline (Pro / P)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{COOH} \end{array}$ <p>Glutamic Acid (Glu / E)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH}_2 \\ \\ \text{COOH} \end{array}$ <p>Aspartic Acid (Asp / D)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{H} - \text{C} - \text{OH} \\ \\ \text{CH}_3 \end{array}$ <p>Threonine (Thr / T)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH}_2 \\ \\ \text{SH} \end{array}$ <p>Cysteine (Cys / C)</p>
$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{S} \\ \\ \text{CH}_3 \end{array}$ <p>Methionine (Met / M)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH}_2 \\ \\ \text{CH} \\ / \quad \backslash \\ \text{CH}_3 \quad \text{CH}_3 \end{array}$ <p>Leucine (Leu / L)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH}_2 \\ \\ \text{C}=\text{O} \\ \\ \text{NH}_2 \end{array}$ <p>Asparagine (Asn / N)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{HC} - \text{CH}_3 \\ \\ \text{CH}_2 \\ \\ \text{CH}_3 \end{array}$ <p>Isoleucine (Ile / I)</p>	$\begin{array}{c} \text{H} \\ \\ \text{H}_3\text{N}^+ - \text{C} - \text{C} - \text{O}^- \\ \\ \text{CH} \\ / \quad \backslash \\ \text{CH}_3 \quad \text{CH}_3 \end{array}$ <p>Valine (Val / V)</p>



Amino Acids

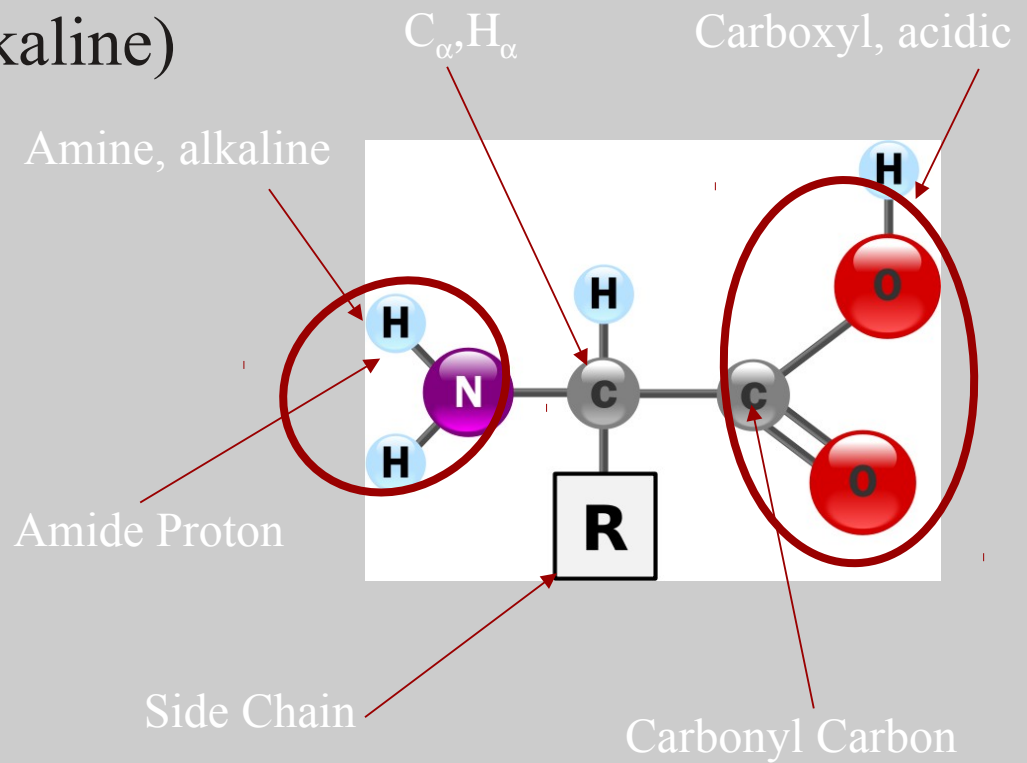
- 20 amino acids
- General structure of an A.A.
- Differences due to side chain R
- Can be categorized as acidic/alkaline, hydrophobic/hydrophilic or positively/negatively charged





Anatomy of Amino Acids

- Much simpler structure compared to nucleotides
- Side chain R
- Carboxylic group (acidic)
- Amide/amine group (alkaline)
- Zwitterion
- C_{α} and H_{α}
- Amide proton
- Carbonyl carbon





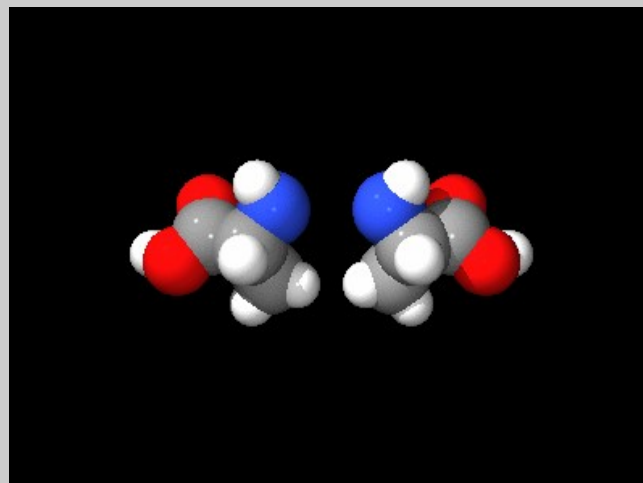
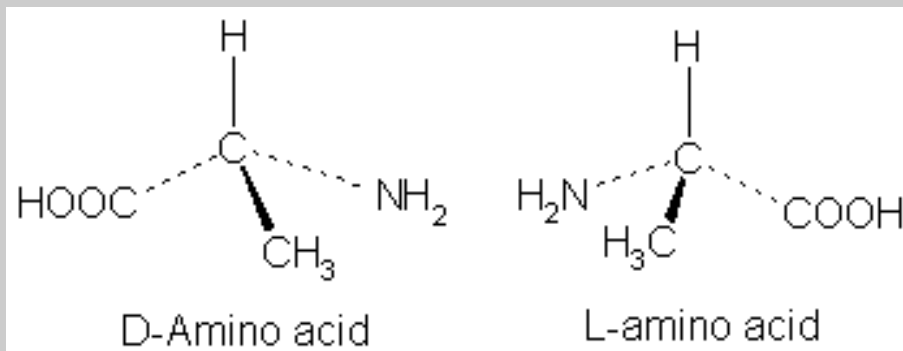
Chemical/Physical Properties of A.A.

- Hydrophobic versus hydrophilic
- Isoelectric point (pI)
- Acidic versus alkaline (basic)
- Mass
- Amino acid full names, three letter and single letter abbreviations
 - Alanine, ALA, A
- Etc.



Stereo Chemistry of Amino Acids

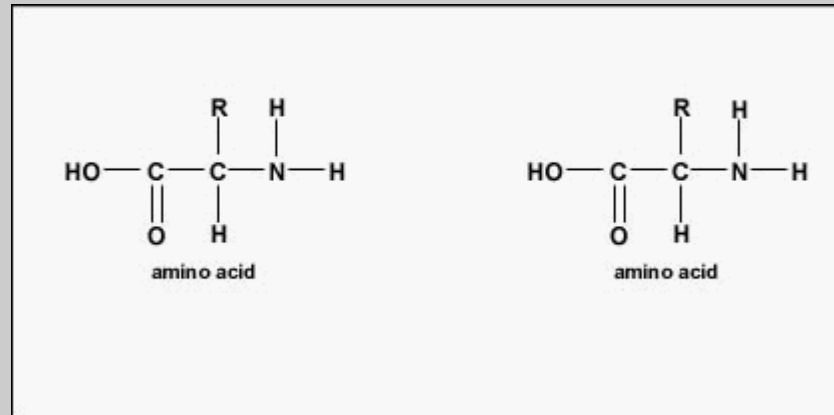
- Amino acids are three dimensional entities
- Carbon forms 4 bonds
- In SP^3 hybridization, these four bonds form a tetrahedron
- Stereoisomer: two identical molecules that can not be superimposed
- D and L are two different enantiomers
- A mixture of equal amounts of both enantiomers is said to be a racemic mixture
- Vast majority of a.a. found in proteins are L a.a.





Peptide (Polymerization of A.A.)

- The family of molecules formed from the linking of various amino acids.
- Two amino acids link through an amide bond or peptide bond.

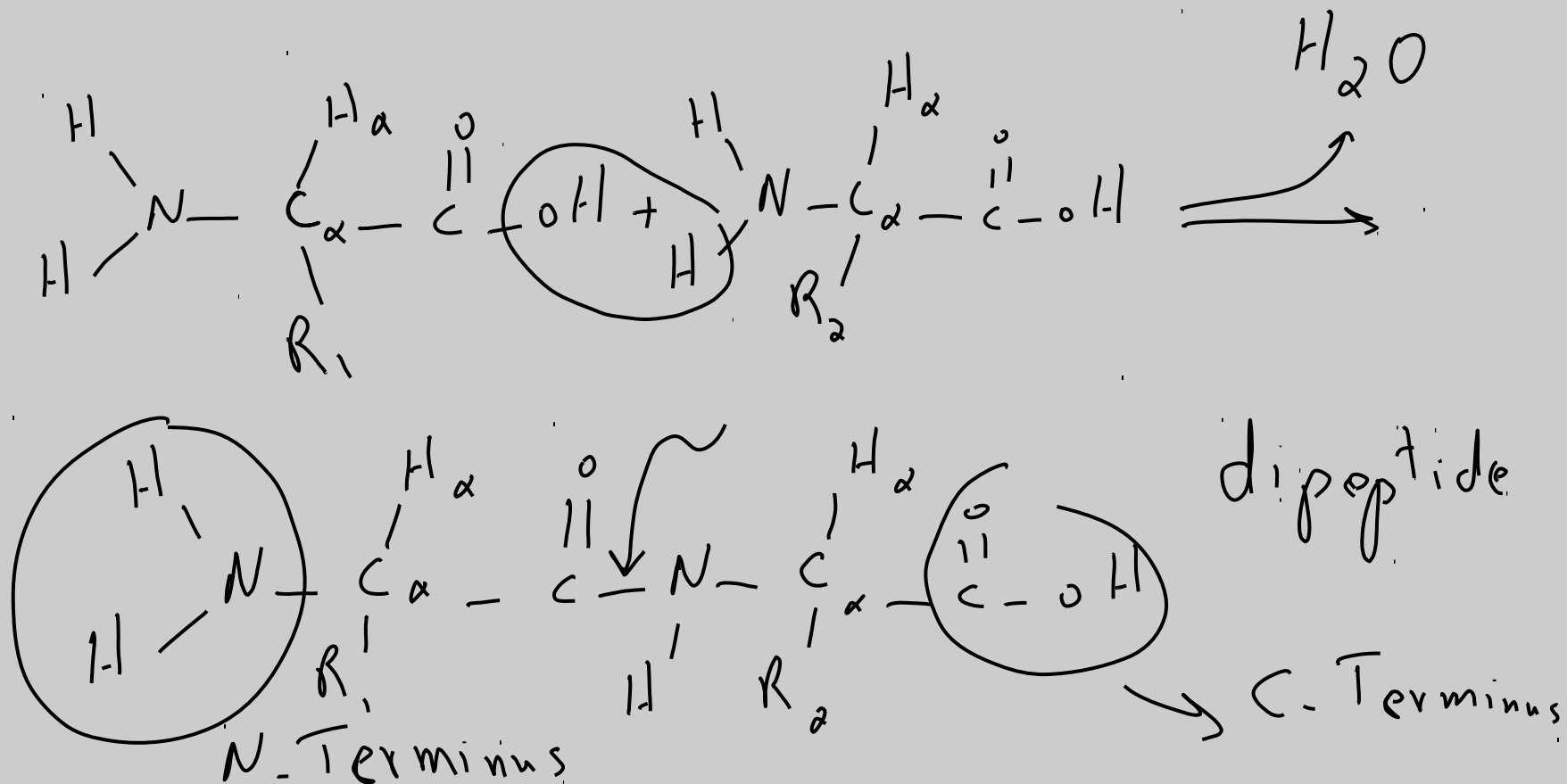


- Two linked amino acids form a peptide plane (more later).



Polymerization of A.A.

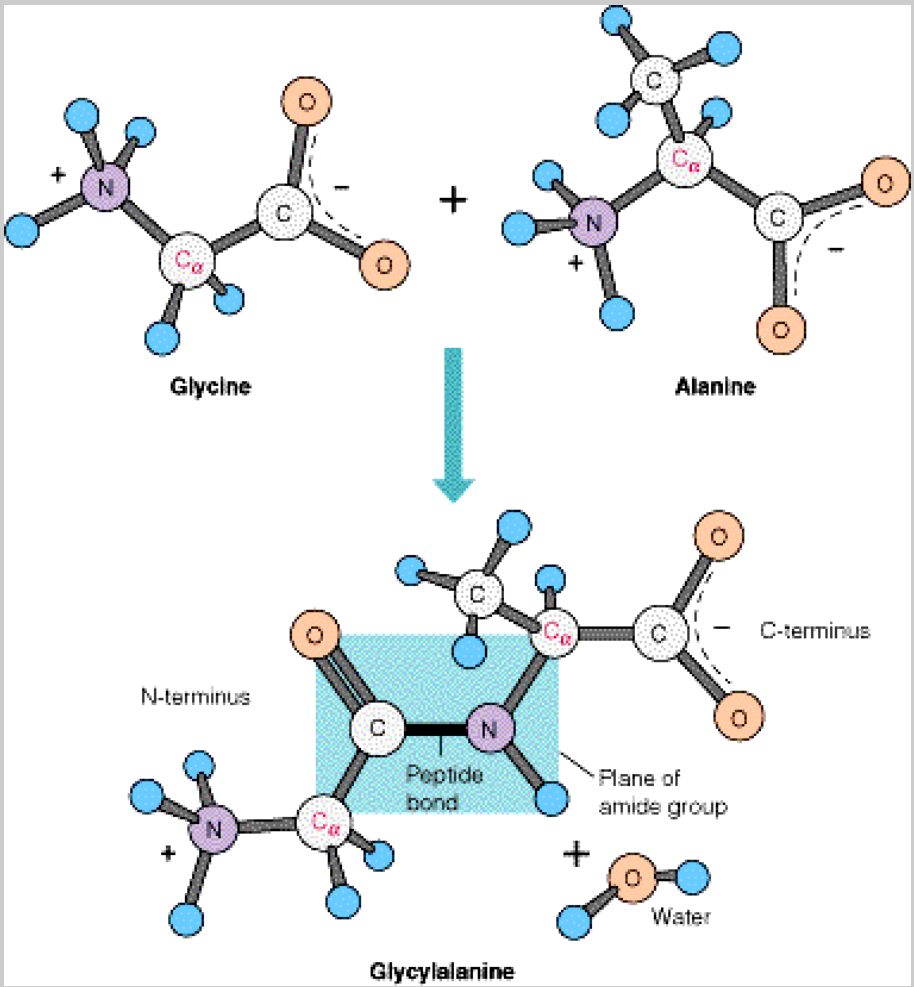
- N-terminus, C-terminus





Polymerization of A.A.

- N-terminus, C-terminus





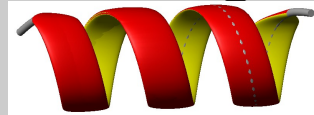
Protein Structure Hierarchy

- Primary sequence (1°):
 - linear order of connected amino acids.

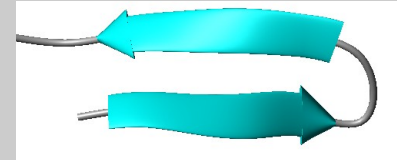
ALA-GLY-LYS-PRO-...

- Secondary Structure (2°):
 - Internal stable segments.

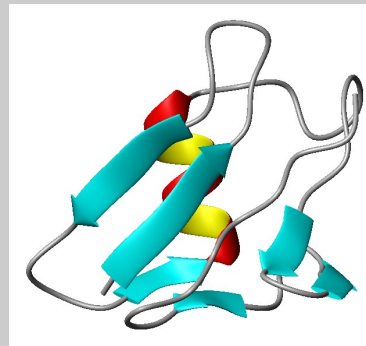
α Helix



β Strand



- Tertiary Structure (3°):





Protein Data Bank (PDB)

- Most prominent database of structures.
- Contains structures for proteins, DNA, RNA, carbohydrates and other biomolecules.
- Data disseminated in two formats: PDB and mmCIF.
- Become familiar with this database!
 - Learn to download files and use them.
 - Learn search and other various tools on this site.
 - Browse links...



Molecular Visualization Tools

- A number of tools available or different computing environments.
 - Rasmol
 - MolMol
 - PyMol
 - VMD
 - Etc.



Why Protein Structure?

- Proteins provide metabolic and mechanical support for biological organisms.
- Structure gives rise to function.
- Structure is necessary (not sufficient) for function.
- Proteins are of special interest due to their therapeutic potential (why not DNA)?

