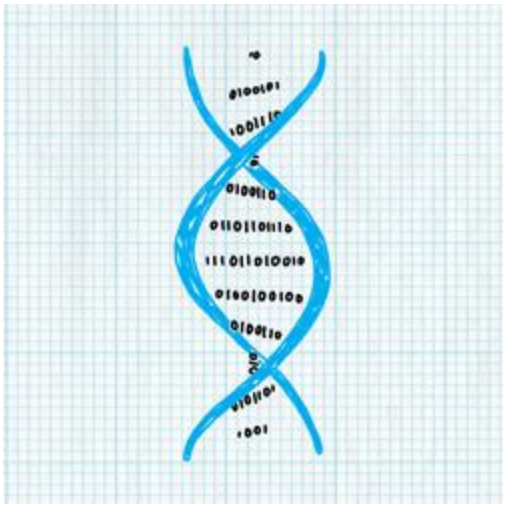
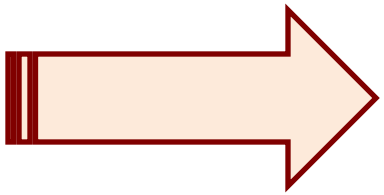
Science of Living System





Nihar Ranjan Jana

School of Bio Science

Email: nihar@iitkgp.ac.in

Tel: 03222-260802

Central Dogma of Life

DNA: Storage Medium

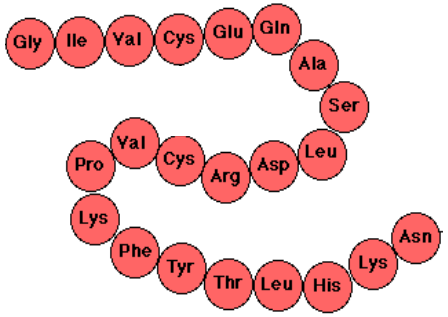
Polymer of nucleotides

CCTGAGCCAACTATTGATGAA

RNA: Transmission Medium

Polymer of nucleotides

CCUGAGCCAACUAUUGAUGAA



Protein: Molecular Machines

Polymer of amino acids

P E P T I D E

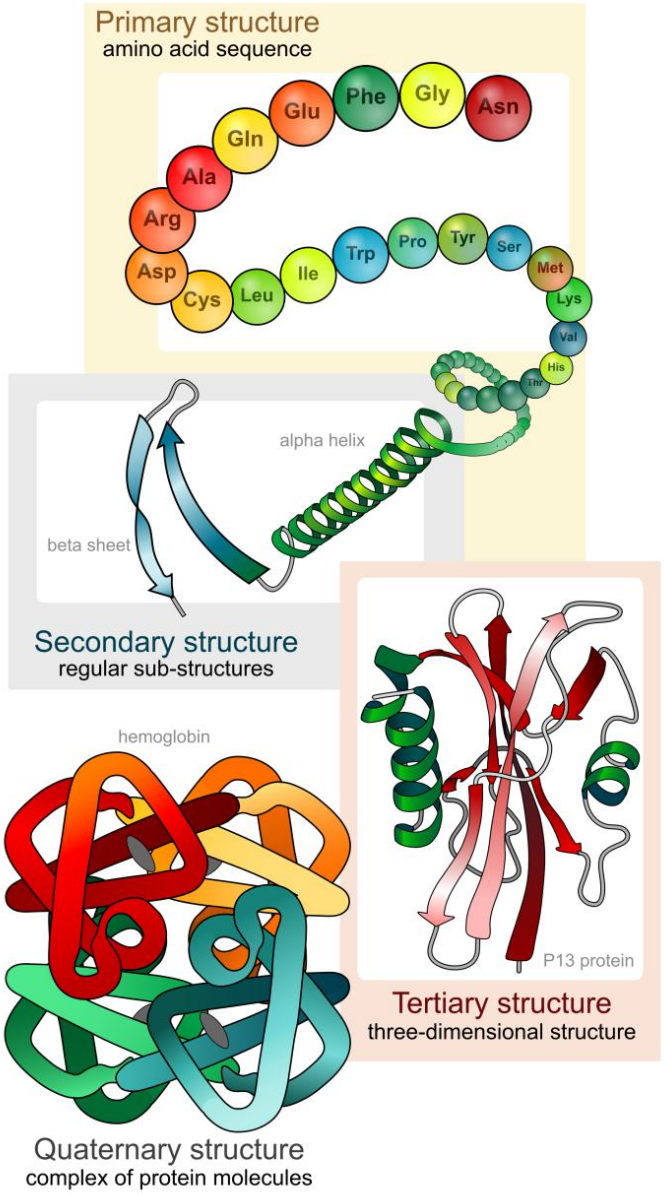
What is a Protein

Proteins are polymers built up from 20 different amino-acids linked by peptide bonds

Ala,Cys,Asp,Glu,Phe,Gly,His,Ile,Lys,Leu
A C D E F G H I K L
Met,Asn,Pro,Gln,Arg,Ser,Thr,Val,Trp,Tyr
M N P Q R S T V W Y

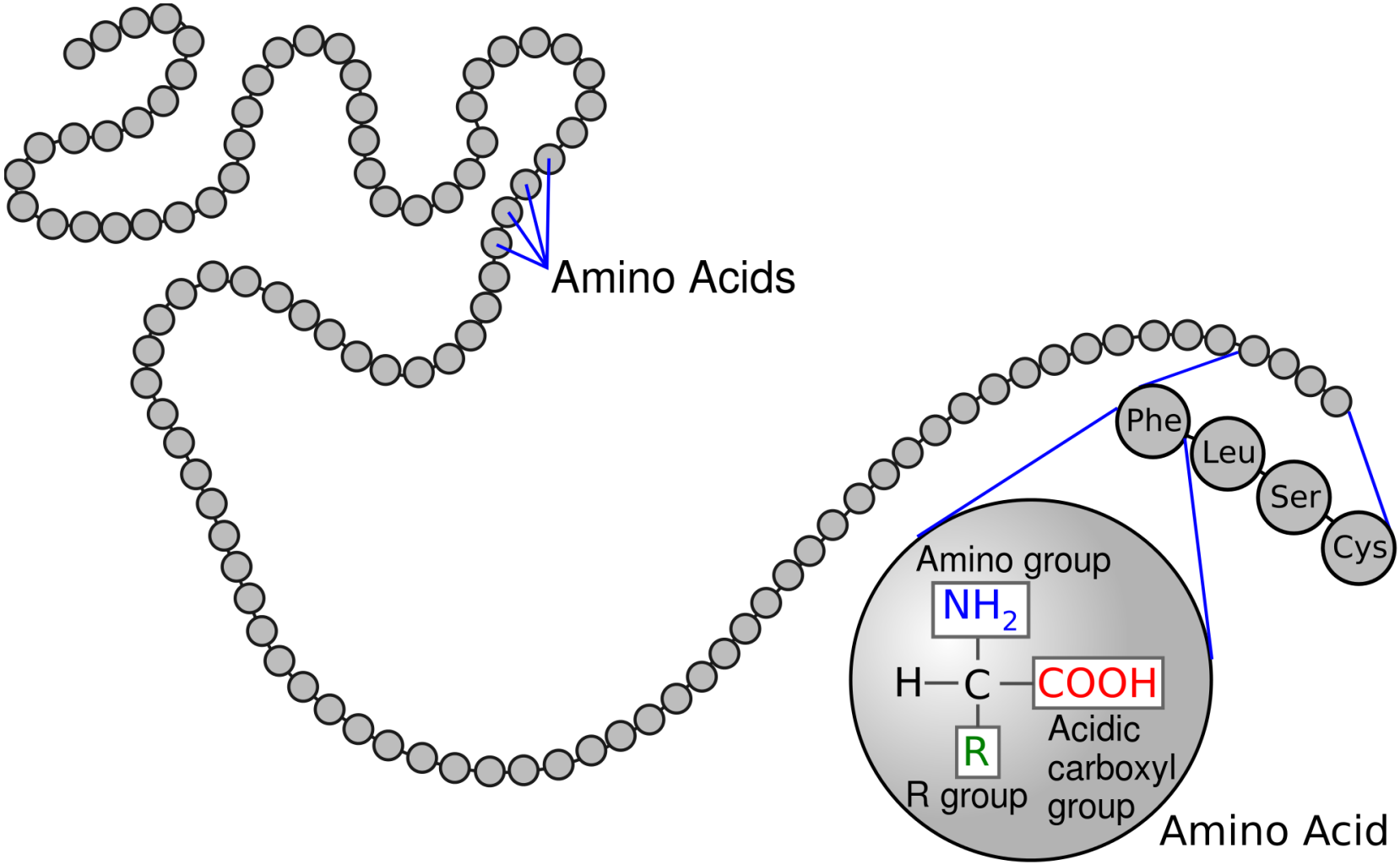
The 3D-structure of a protein is encoded in its amino acid sequence (primary structure)!

Hierarchy of Protein Structure

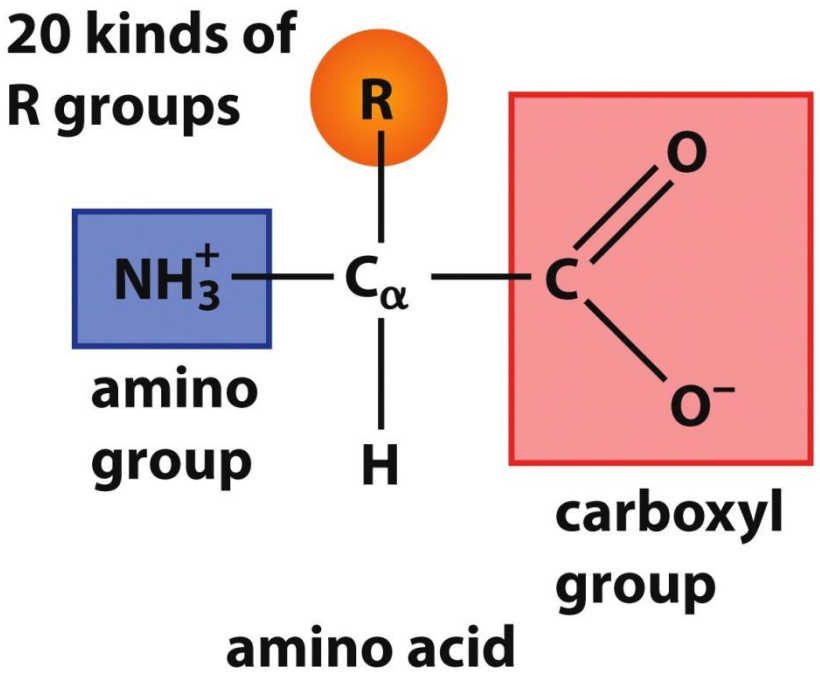


Primary Structure of Proteins

The primary structure of a protein is its amino acid sequence



Amino Acid

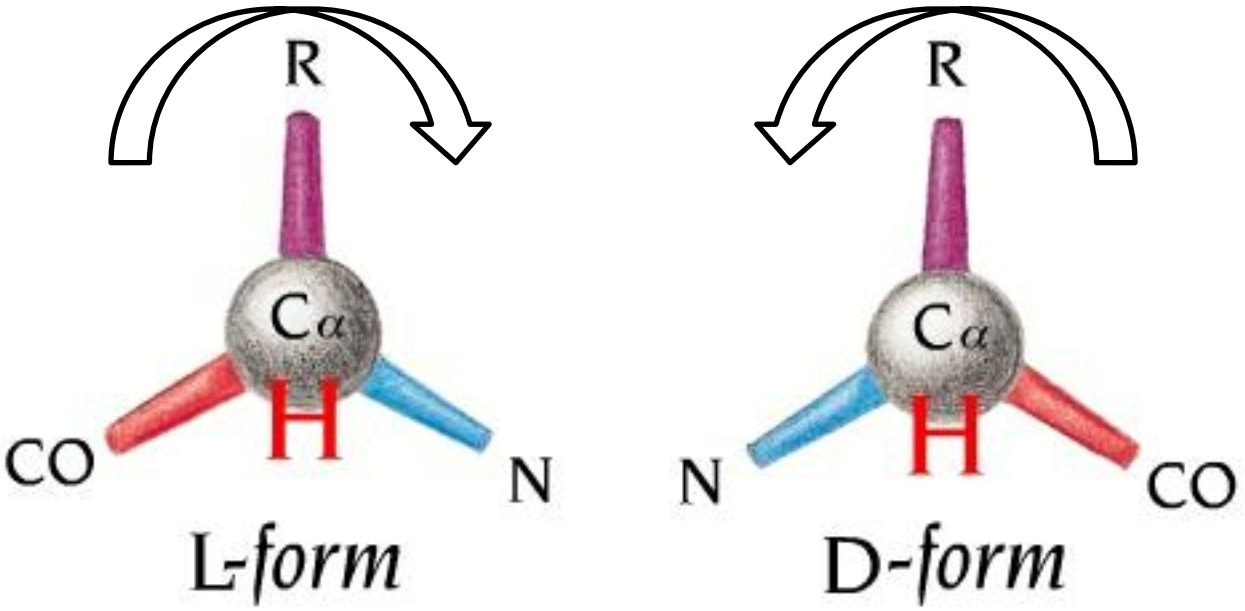


Amino acids have a common chemical structure - A tetrahedral sp^3 carbon (C_α) with four different functional groups:

1. **Amino group**
2. **Carboxyl group**
3. **H-atom**
4. **Side chain (R) with distinct chemical property**

There are 20 common amino acids. The R group (also known as side chain), attached to the α carbon is different in each amino acid

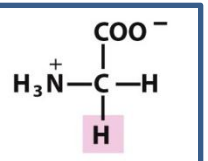
All Amino Acids in Protein Have the "L-form"

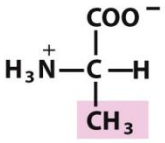


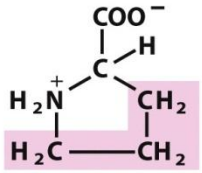
H-atom is coming out of the whiteboard. Looking down the H-C_α bond from the H-atom, the L-form amino acid has CO, R and N going in a clockwise direction. The L-form reads "CORN" in clockwise direction.

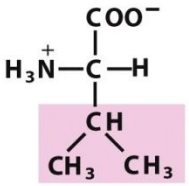
The 20 Common Amino Acids of Proteins

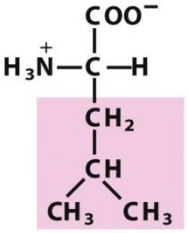
Nonpolar, aliphatic R groups

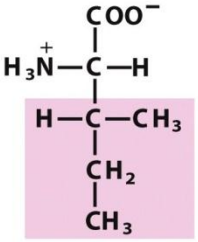

Glycine

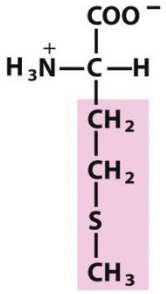

Alanine


Proline

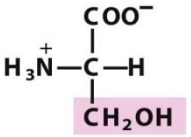

Valine

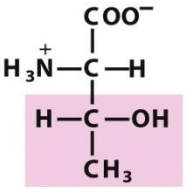

Leucine

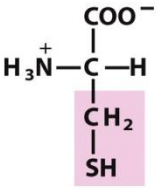

Isoleucine

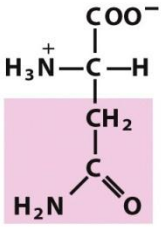

Methionine

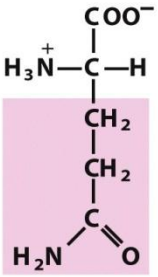
Polar, uncharged R groups


Serine

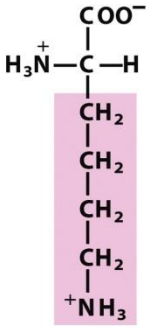

Threonine

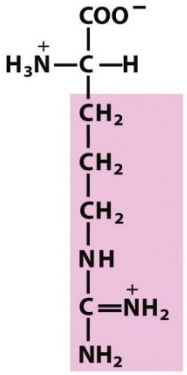

Cysteine

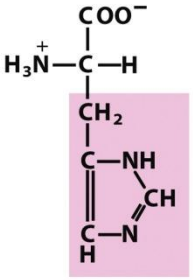

Asparagine


Glutamine

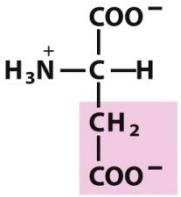
Positively charged R groups

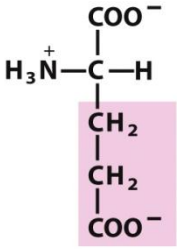

Lysine


Arginine

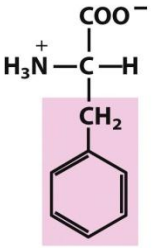

Histidine

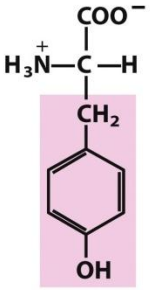
Negatively charged R groups

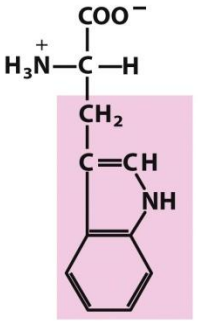

Aspartate


Glutamate

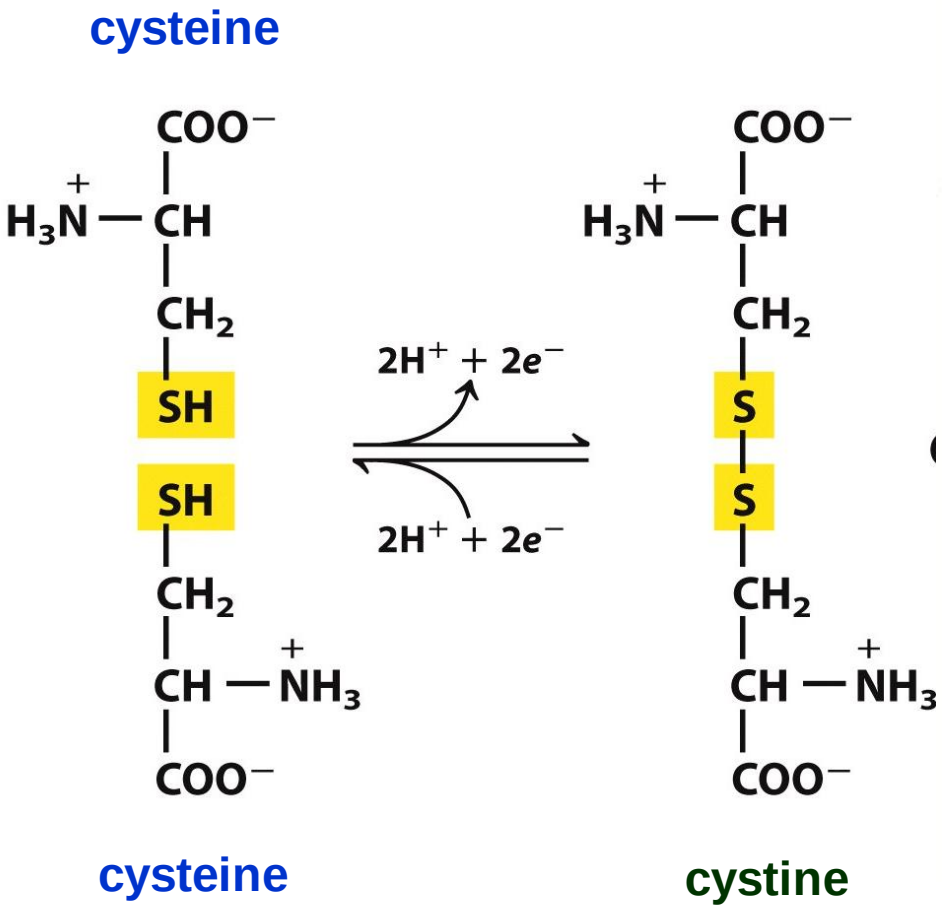
Aromatic R groups

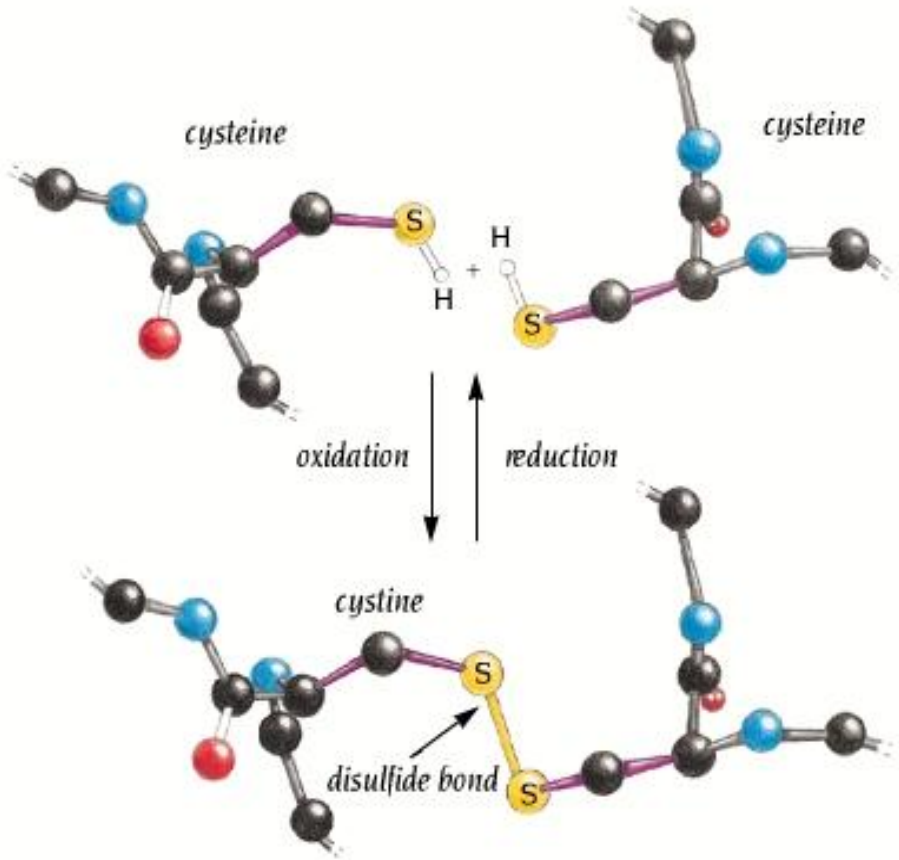

Phenylalanine


Tyrosine


Tryptophan

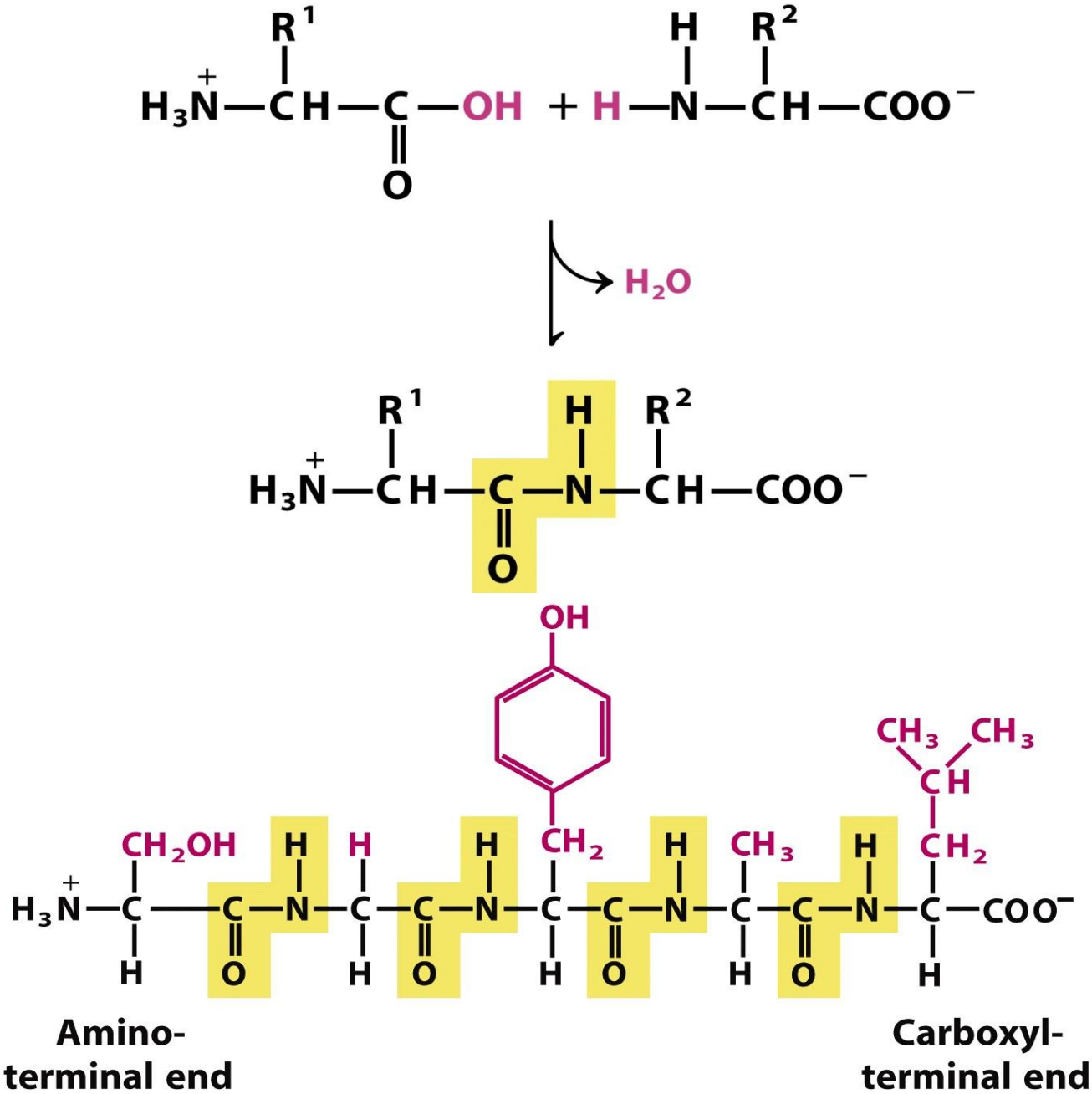
Amino Acid with Special Characteristics



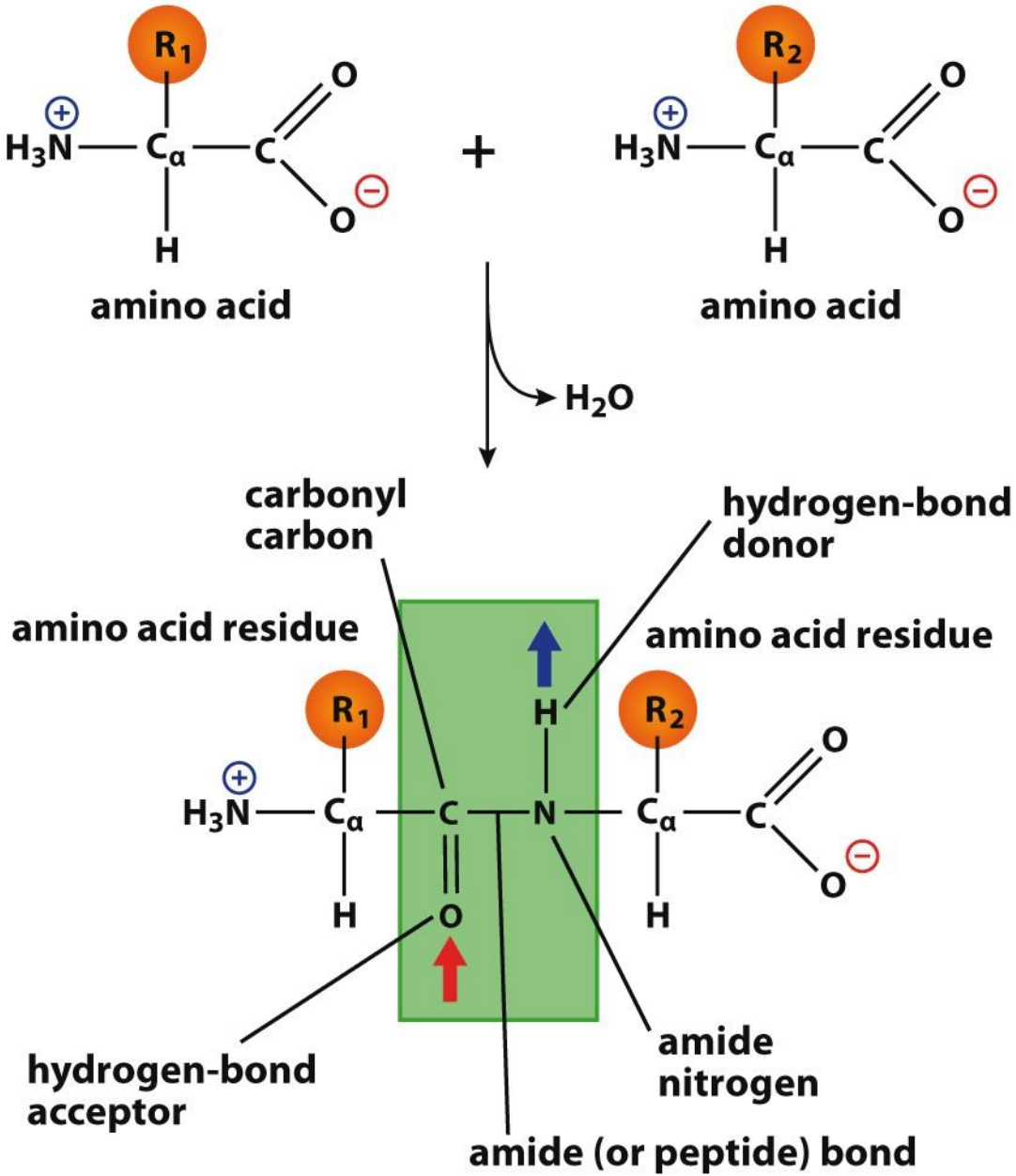


Reversible formation of a disulfide bond by the oxidation of two molecules of cysteine. Disulfide bonds between Cys residues stabilize the structures of many proteins

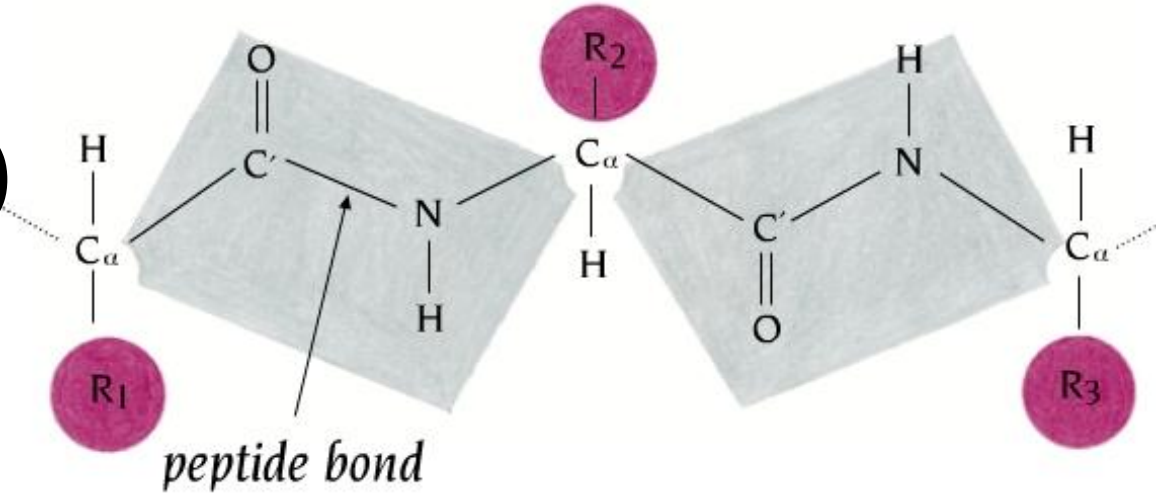
Formation of the Peptide Bond

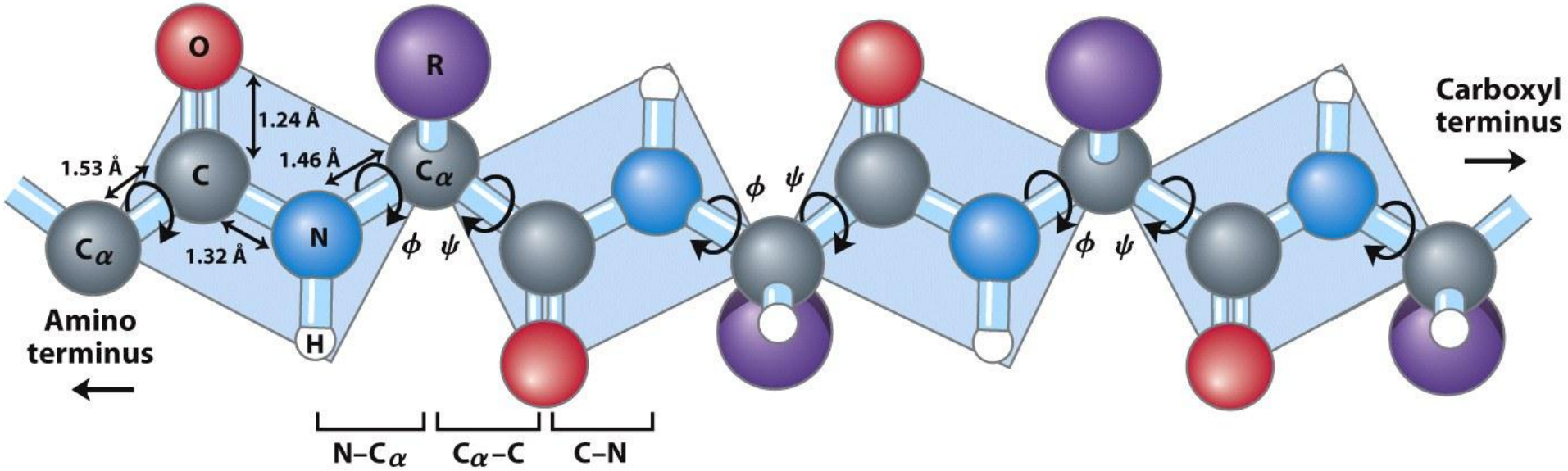


Formation of the Peptide Bond



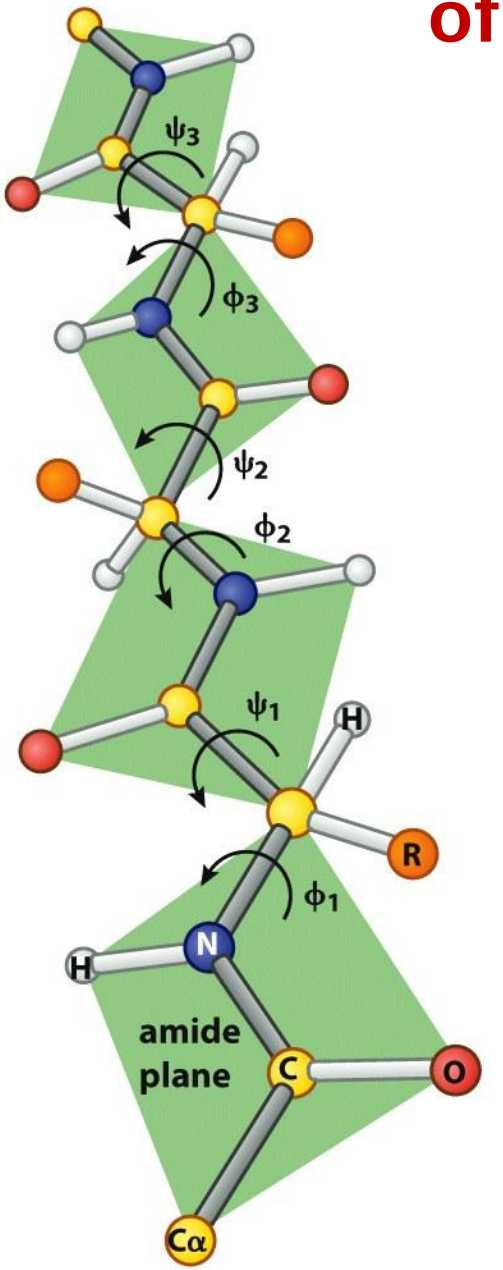
Torsion angles: Φ (phi) and Ψ (psi)





The N—C_α and C_α—C bonds can rotate, and designated as Φ and Ψ angles respectively. The peptide C—N bond is not free to rotate. Other single bonds in the backbone may also be rotationally hindered, depending on the size and charge of the R groups

φ and ψ Torsion Angles are the Only Degrees of Freedom for the Backbone



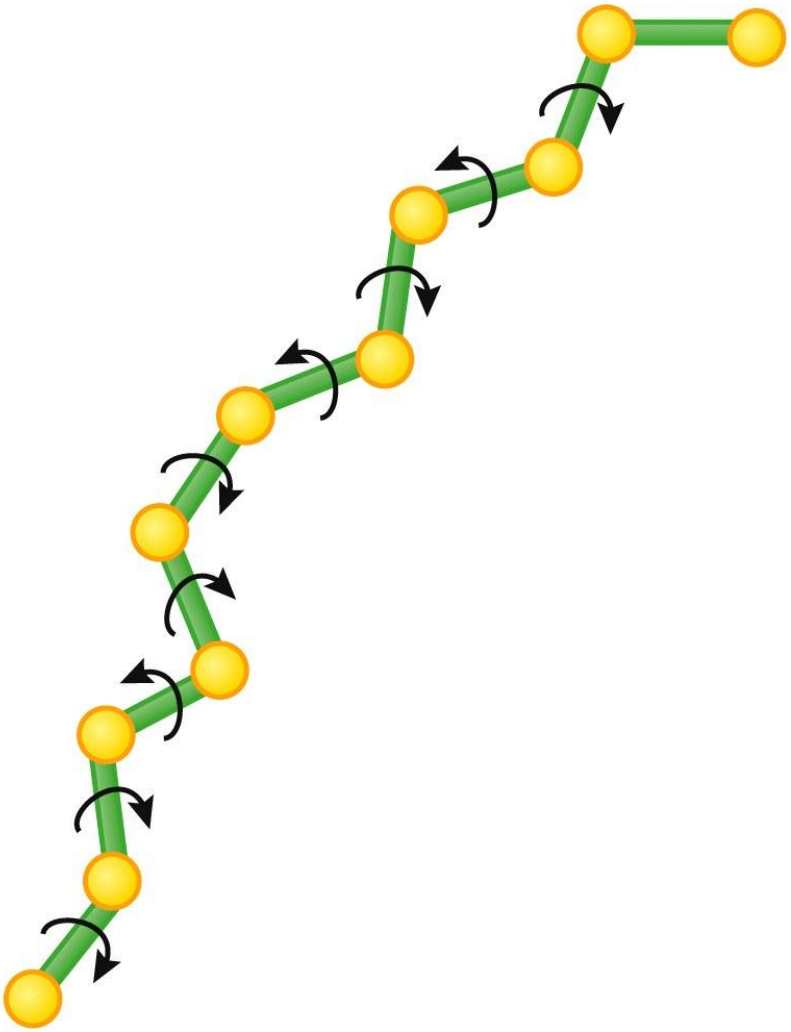
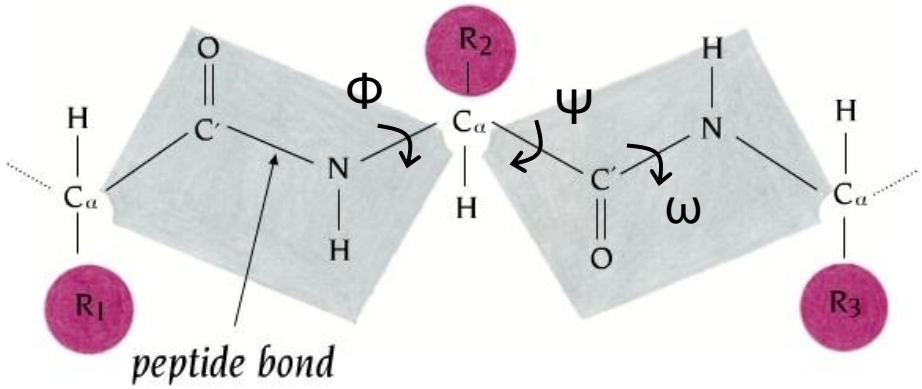
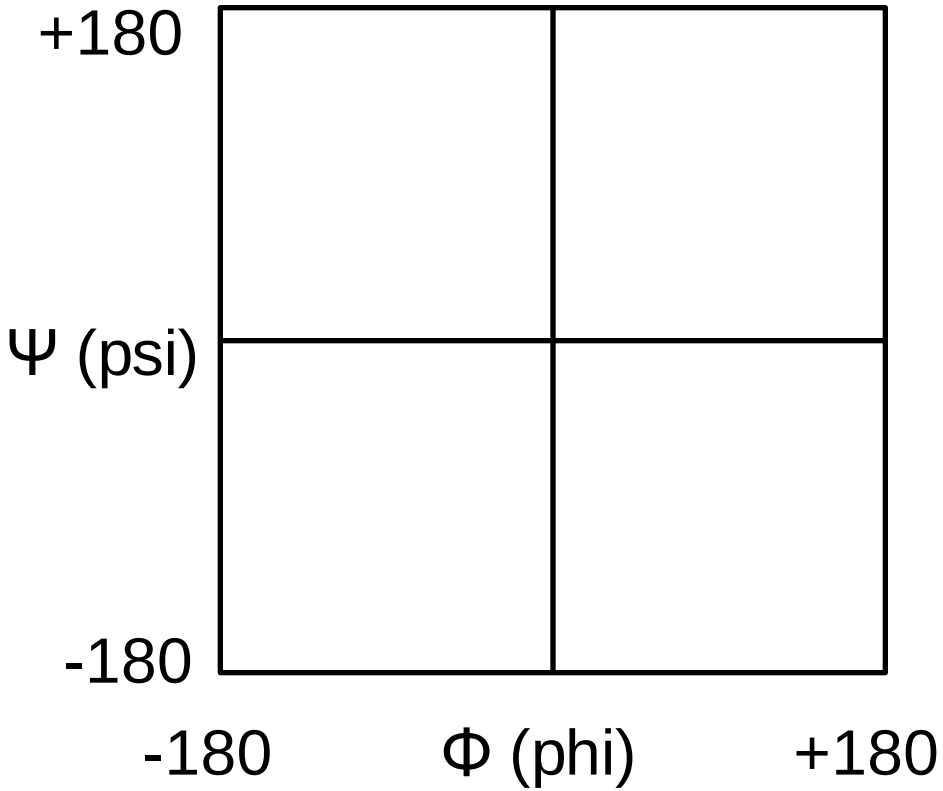


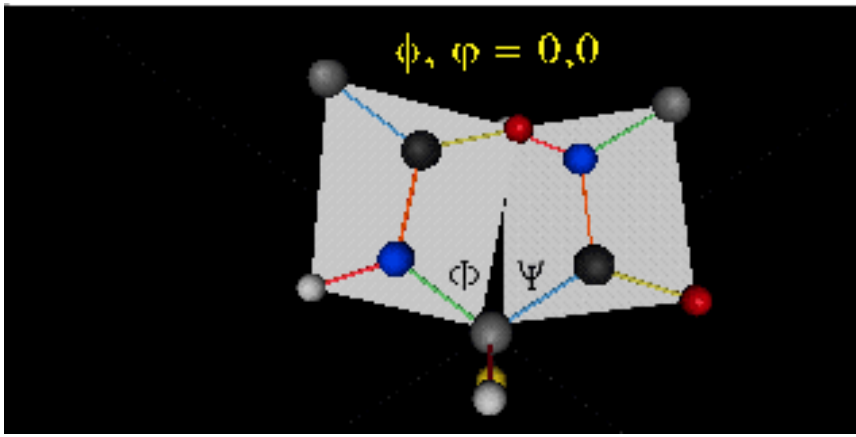
Figure 4.11b The Molecules of Life (© Garland Science 2013)

Ramachandran Plot: The φ - ψ Space



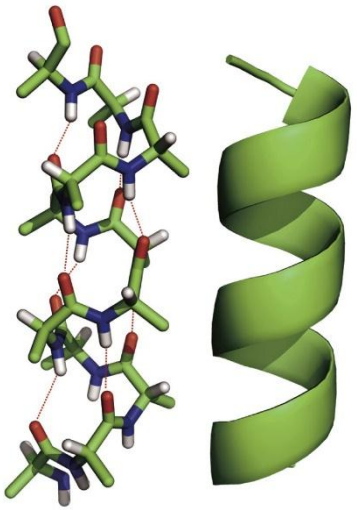
G. N. Ramachandran



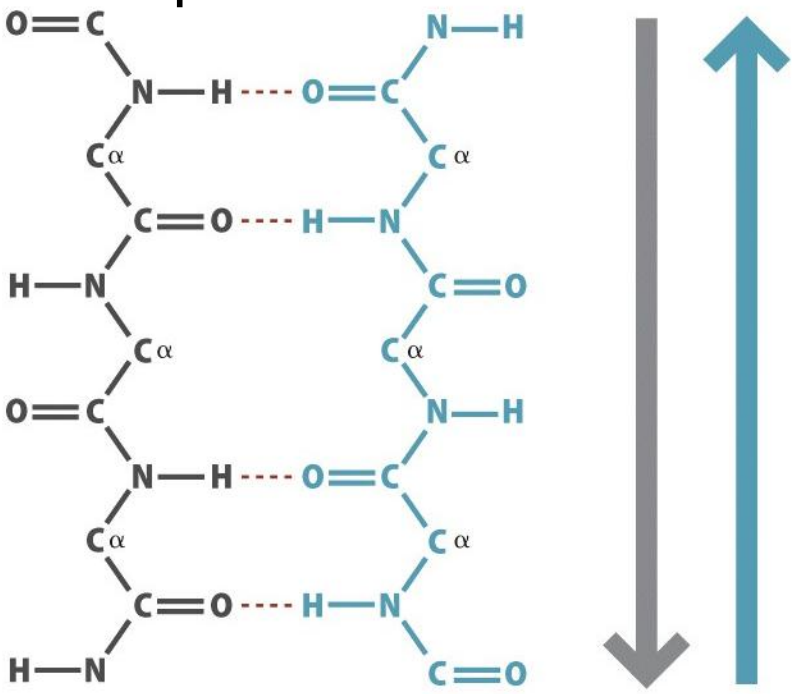


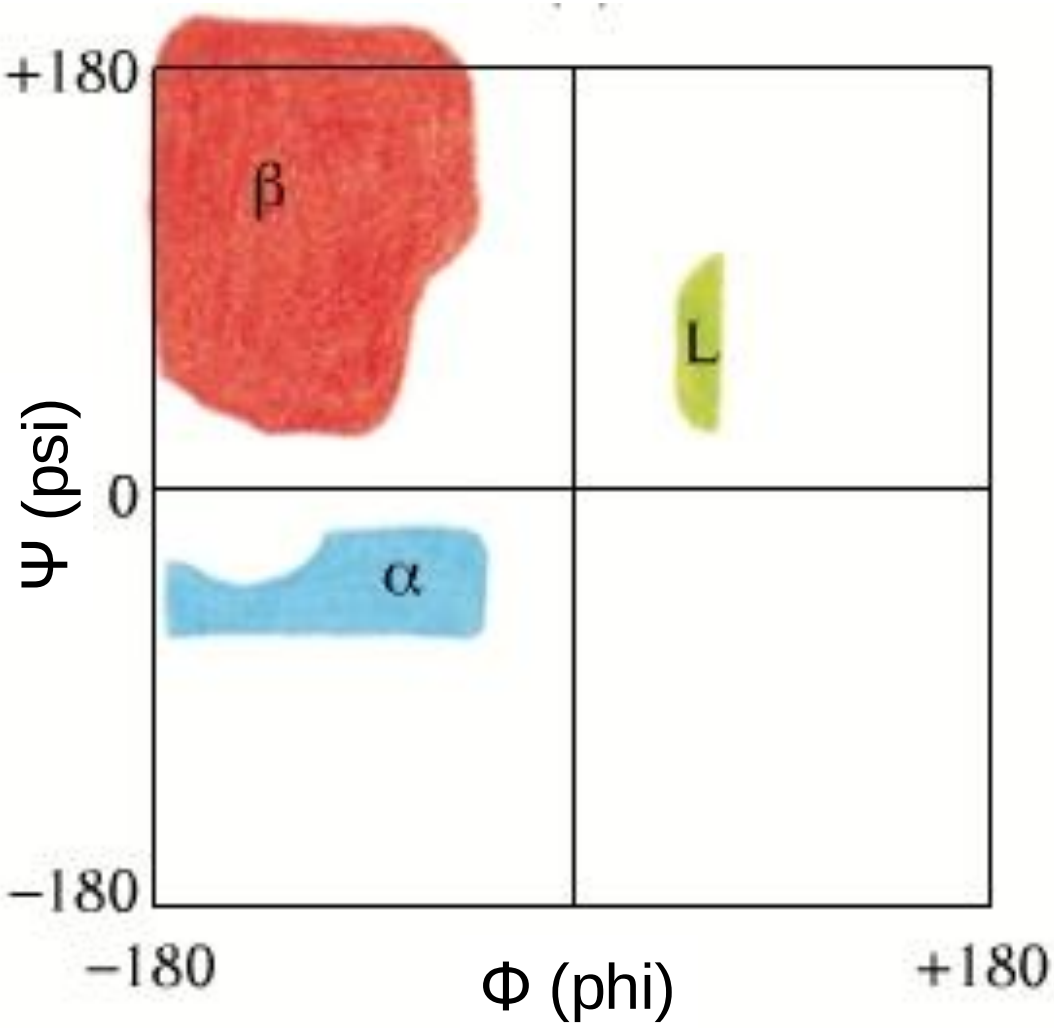
Ramachandran Plot

α-Helix



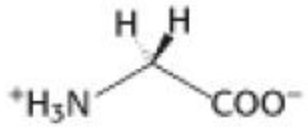
β-Strand

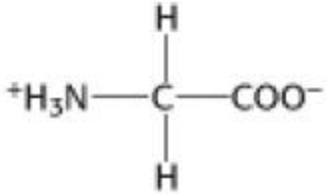




Glycine Residues can Adopt Many Different Conformations

Glycine
(Gly, G)





Glycine
(Gly, G)

- Glycine with only a H-atom as side chain can adopt a much wider range of Φ-Ψ conformations than the other residues
- It thus plays a structurally important role; it allows unusual main chain conformations in proteins
- This is the main reasons why a high proportion of Glycine residues are conserved among homologous protein sequences

Central Dogma of Life

DNA: Storage Medium

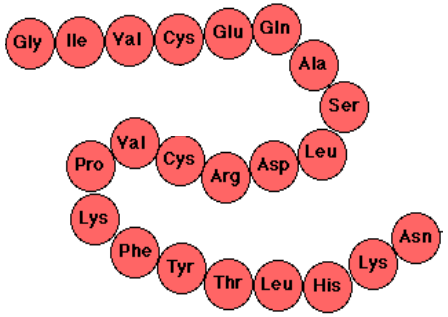
Polymer of nucleotides

CCTGAGCCAACTATTGATGAA

RNA: Transmission Medium

Polymer of nucleotides

CCUGAGCCAACUAUUGAUGAA

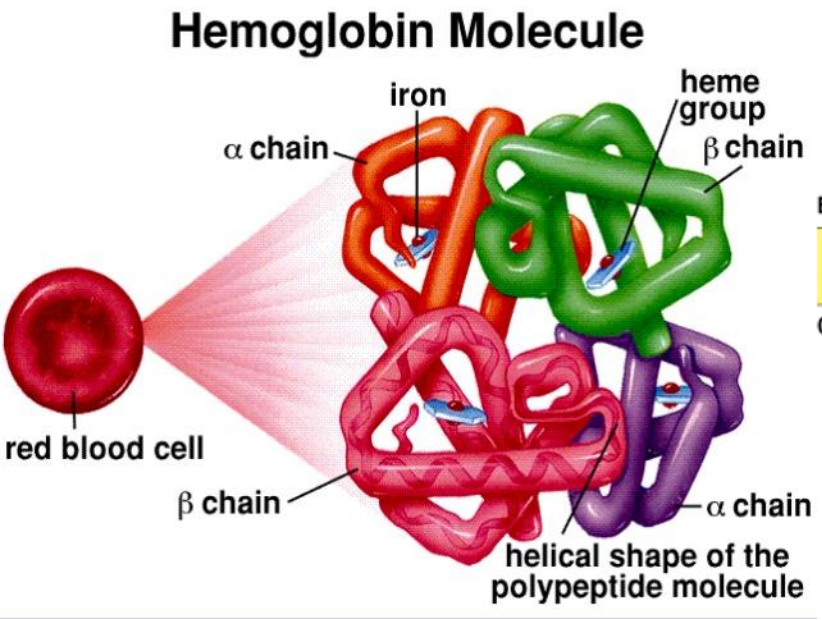


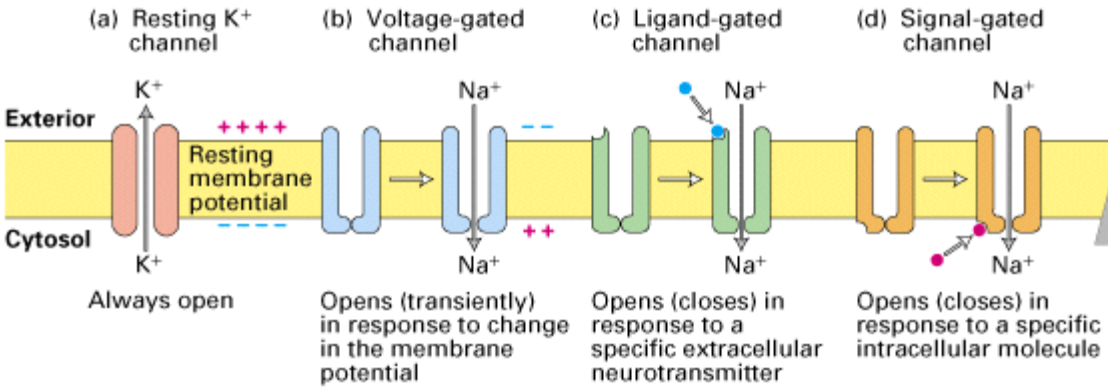
Protein: Molecular Machines

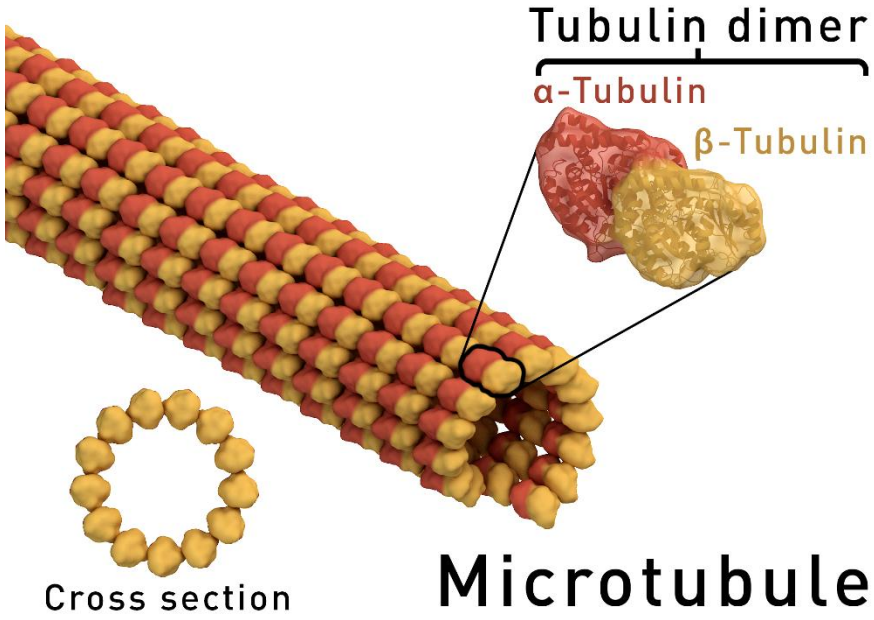
Polymer of amino acids

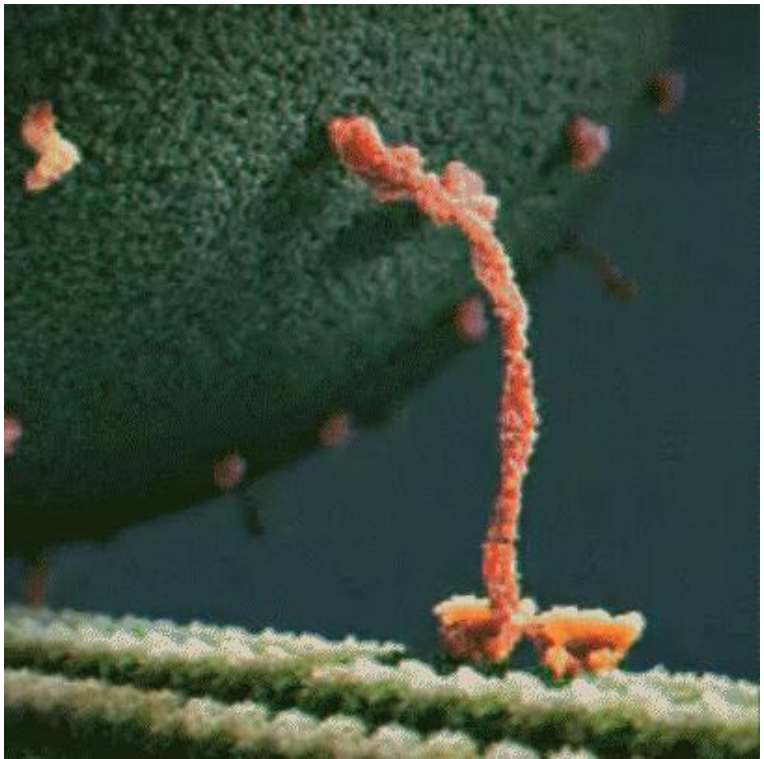
P E P T I D E

Proteins come in various shapes and sizes

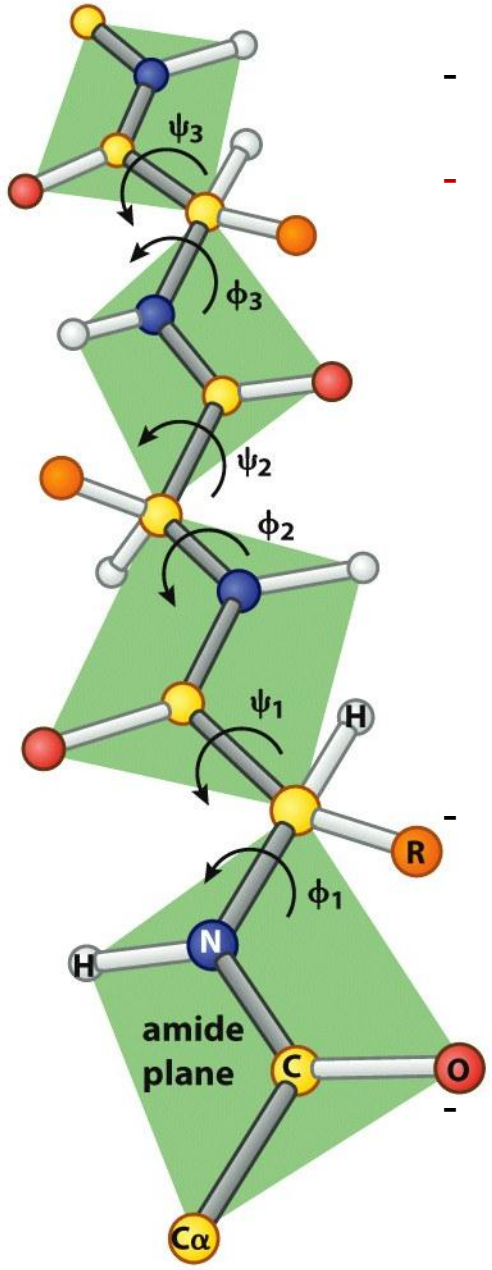








The protein folding problem

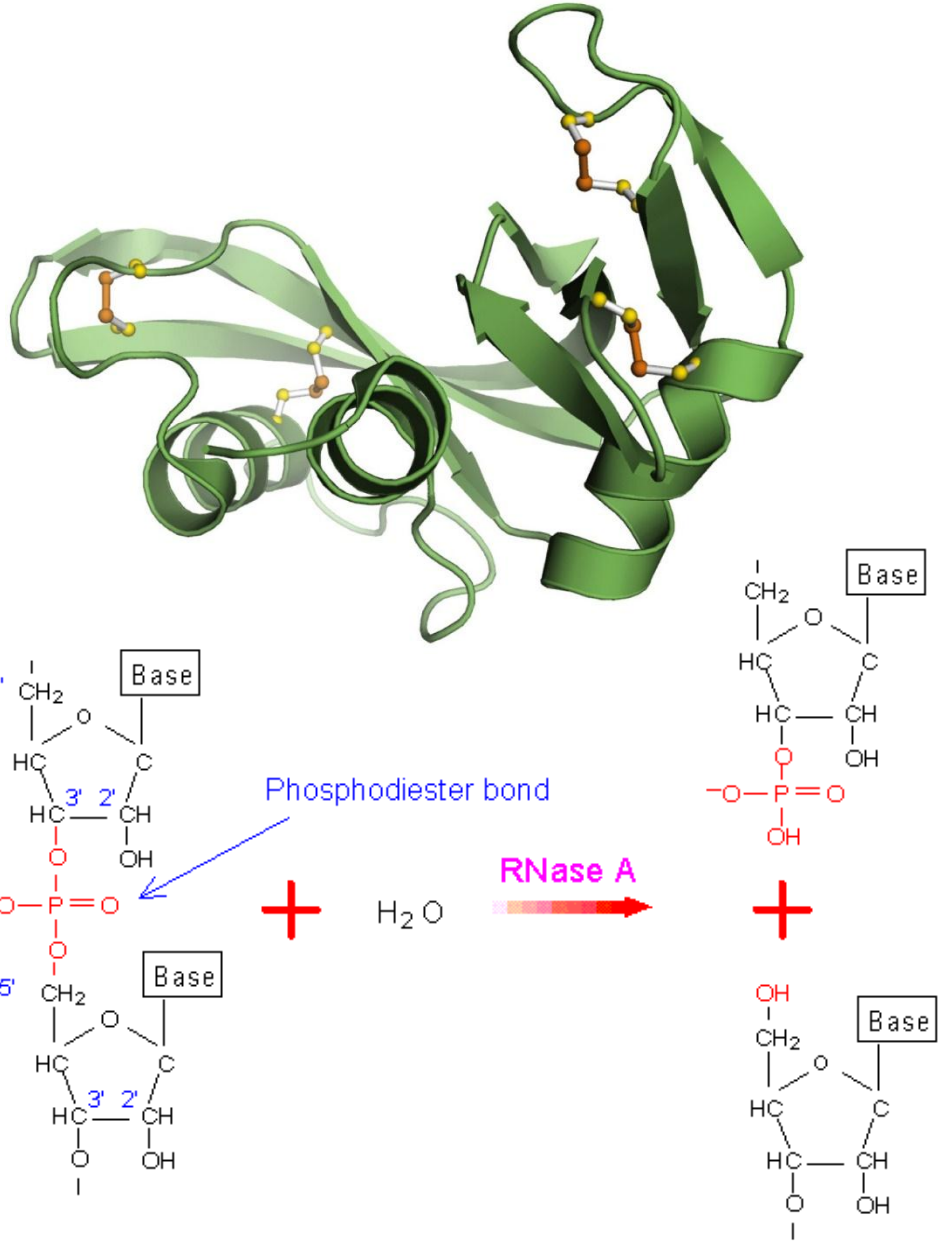


- Consider a small protein with 100 residues.
- Cyrus Levinthal calculated that, if each residue can assume three different conformations, the total number of structures would be $\mathbf{3^{100}}$, which is equal to 5×10^{47}. If it takes 10^{-13} s to convert one structure into another, the total search time would be $5 \times 10^{47} \times 10^{-13}$ s, which is equal to 5×10^{34} s, or $\mathbf{10^{27}}$ **years** i.e. longer than the age of the universe!
- Clearly, it would take much too long for even a small protein to fold properly by randomly trying out all possible conformations.
- The enormous difference between calculated and actual folding times is called ***Levinthal's paradox.***

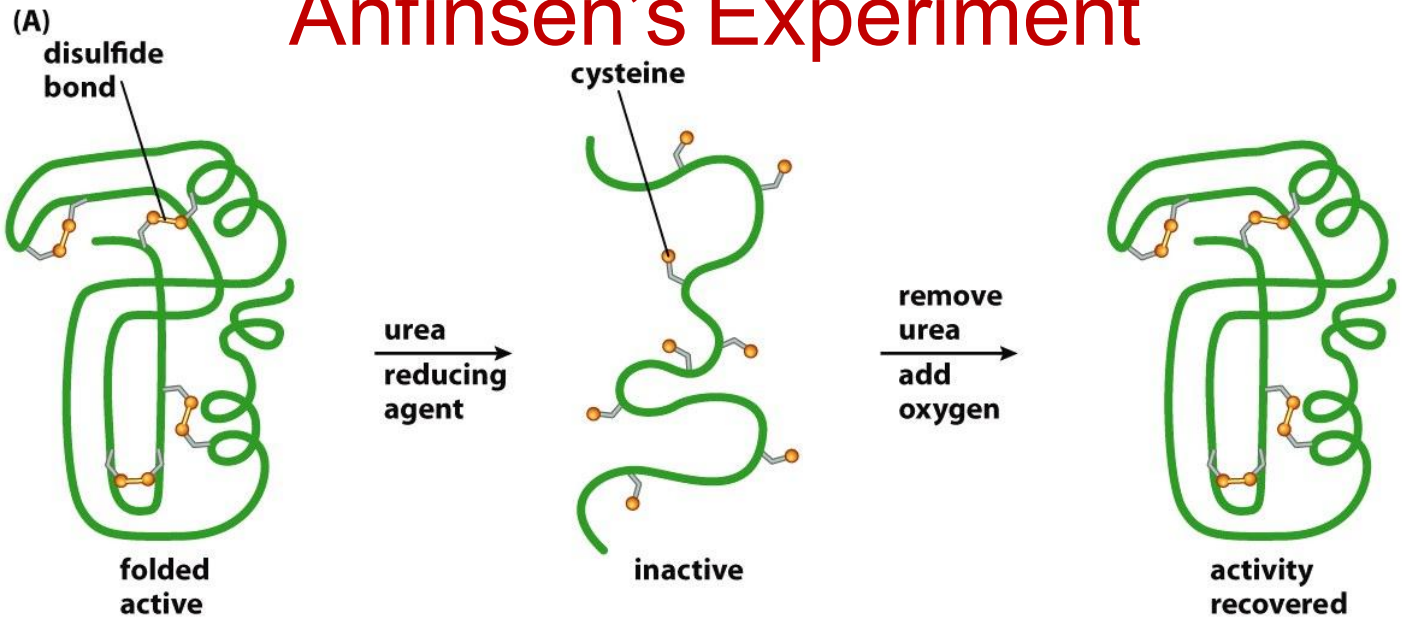
The 3D structure of a protein is encoded in its primary sequence: Anfinsen's Experiment

Thermodynamic hypothesis of Protein Folding: The interactions between the atoms in a protein control the folding of the protein molecule into a well-defined three-dimensional structure.

In other words, the protein sequence contains enough information required for the proper folding of the protein into its functional three-dimensional structure.

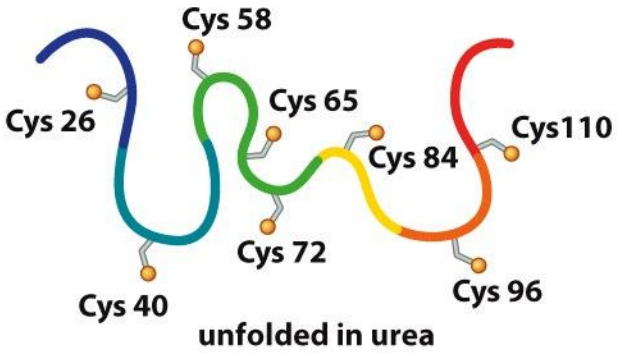


Anfinsen's Experiment



If we understand HOW PROTEINS FOLD, we can predict their structure from sequence! Then we can design proteins with novel functions.

Anfinsen's Experiment



Forces that stabilize a protein structure

- **Hydrophobic effect**
- **Conformational entropy**
- **Electrostatics**
- **Hydrogen bonding**
- **van der Waals interaction**

Most important feature: The interior of proteins is hydrophobic!

The main driving force for folding water soluble globular protein molecules is to pack hydrophobic side chains into the interior of the molecule, thus creating a **HYDROPHOBIC CORE** and a **HYDROPHILLIC SURFACE**.

Problem: **How to create such a hydrophobic core from a linear protein chain** ???

Hydrophobic core formation drives protein folding

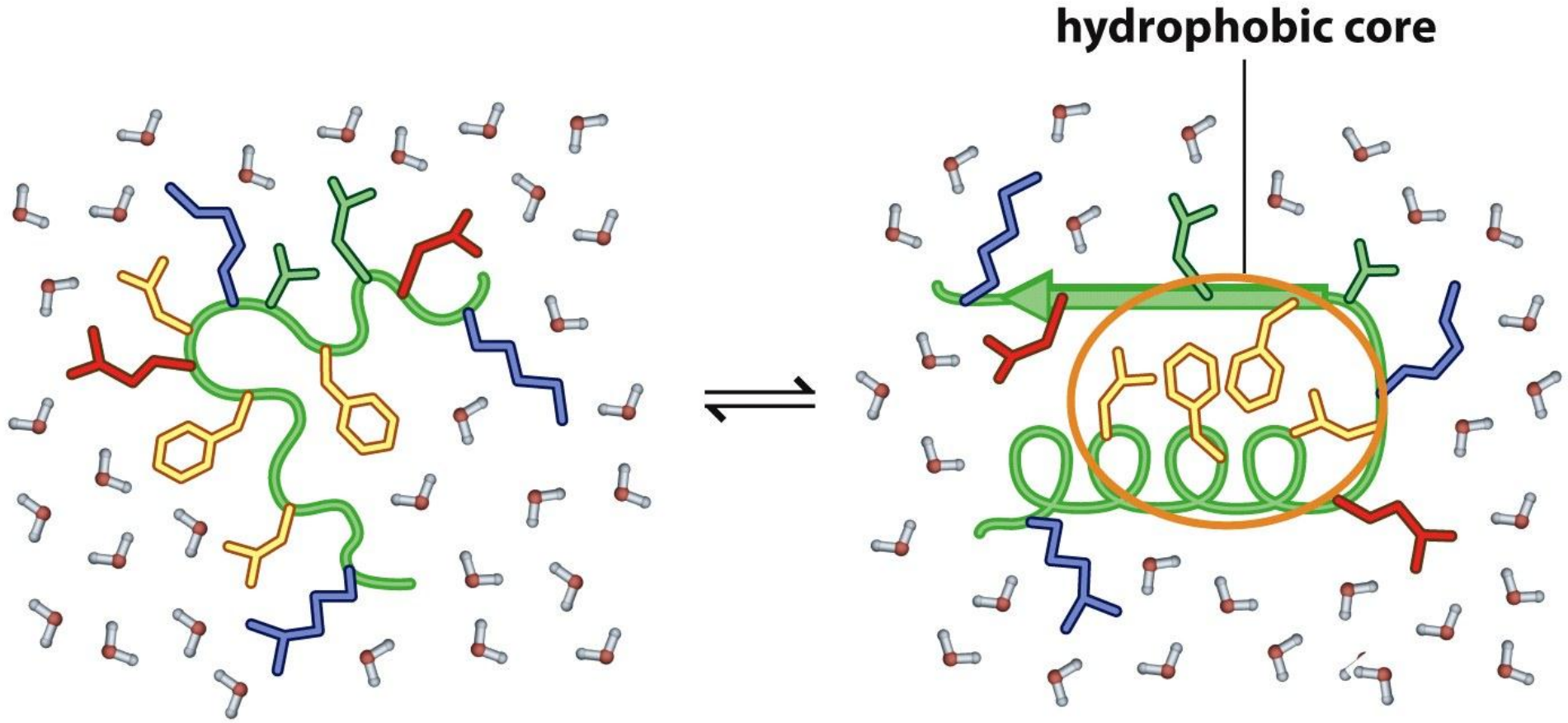


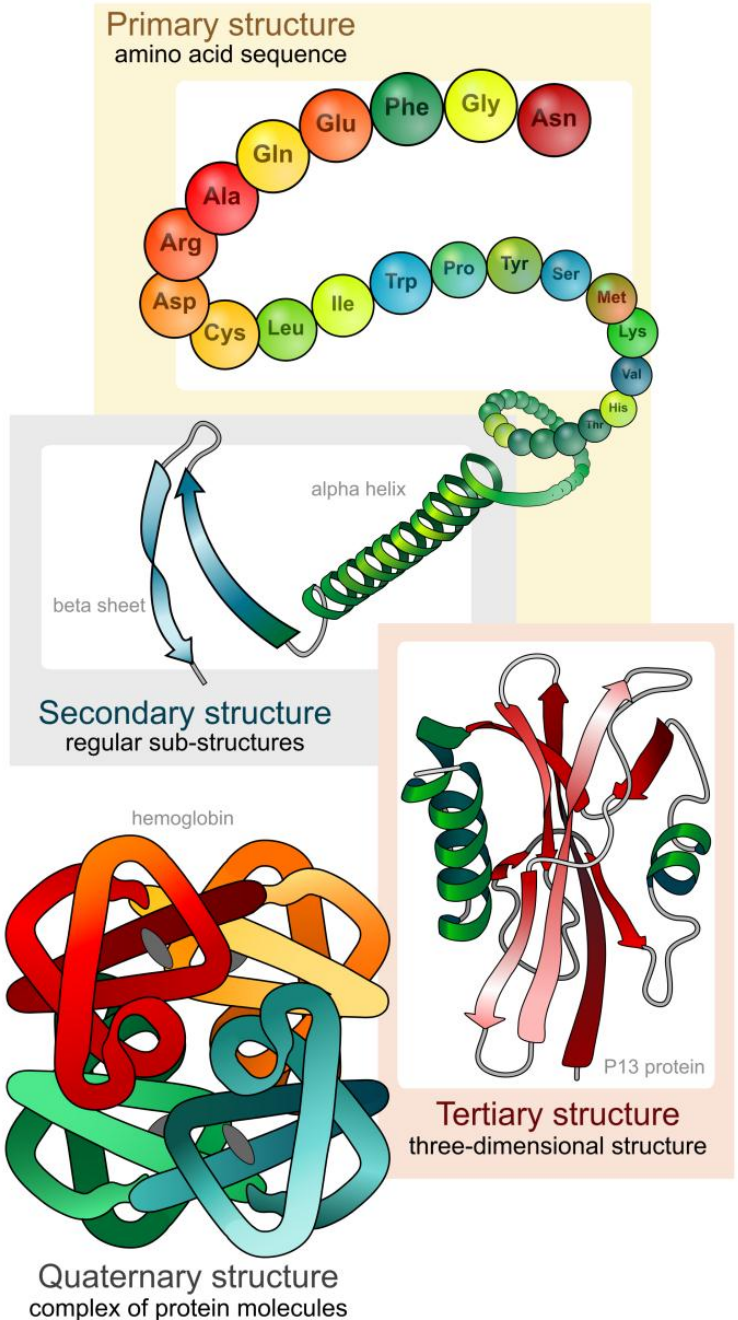
Figure 4.7 The Molecules of Life (© Garland Science 2013)

The Protein Folding Game - Foldit

https://fold.it/portal/

Motifs of Protein Structure

Protein Molecules are Organized in a Structural Hierarchy



Primary

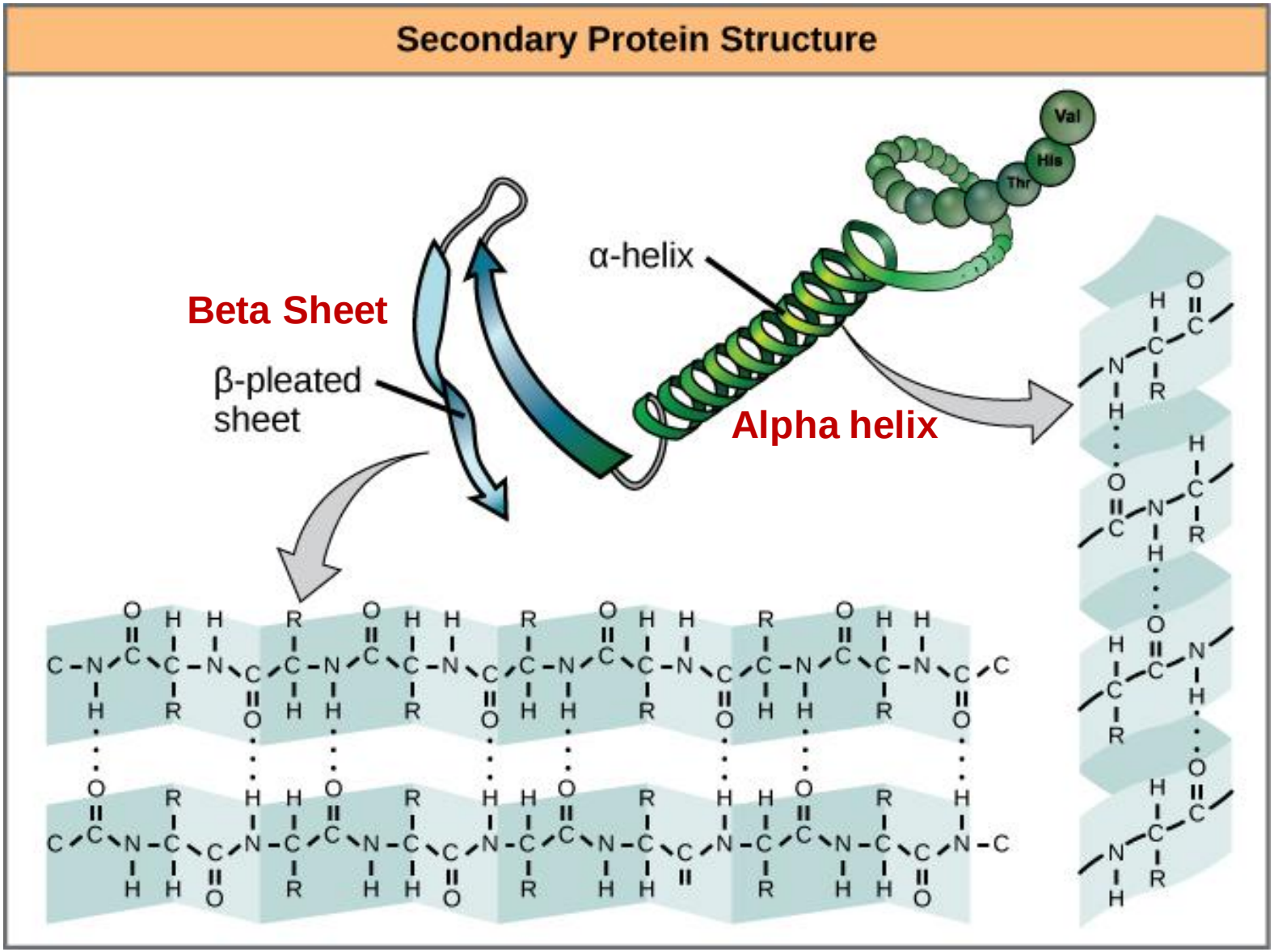
Secondary

Tertiary

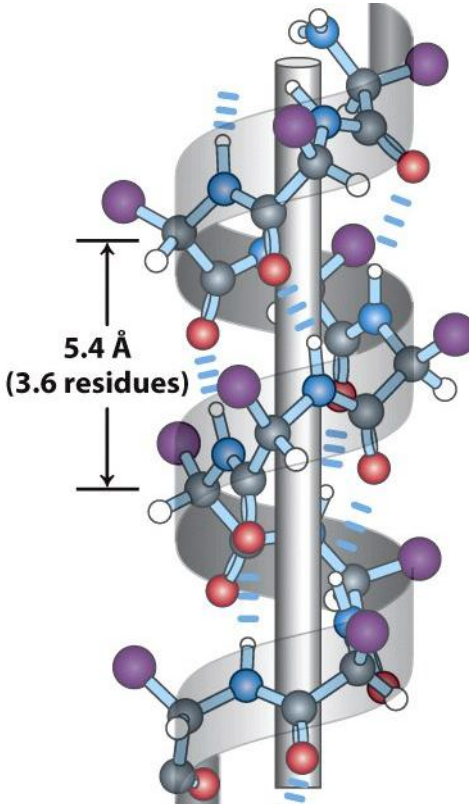
Quaternary

Secondary Protein Structure

Characterized by main chain NH and CO groups participating in H-bonds

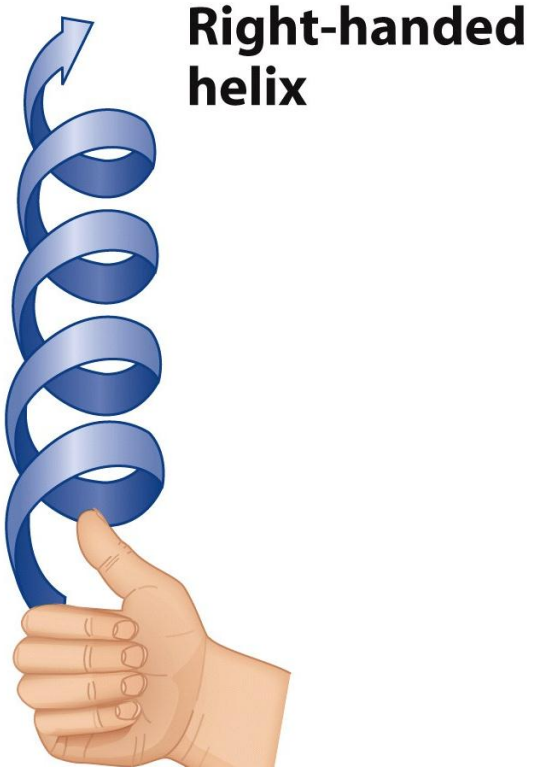


Alpha Helix

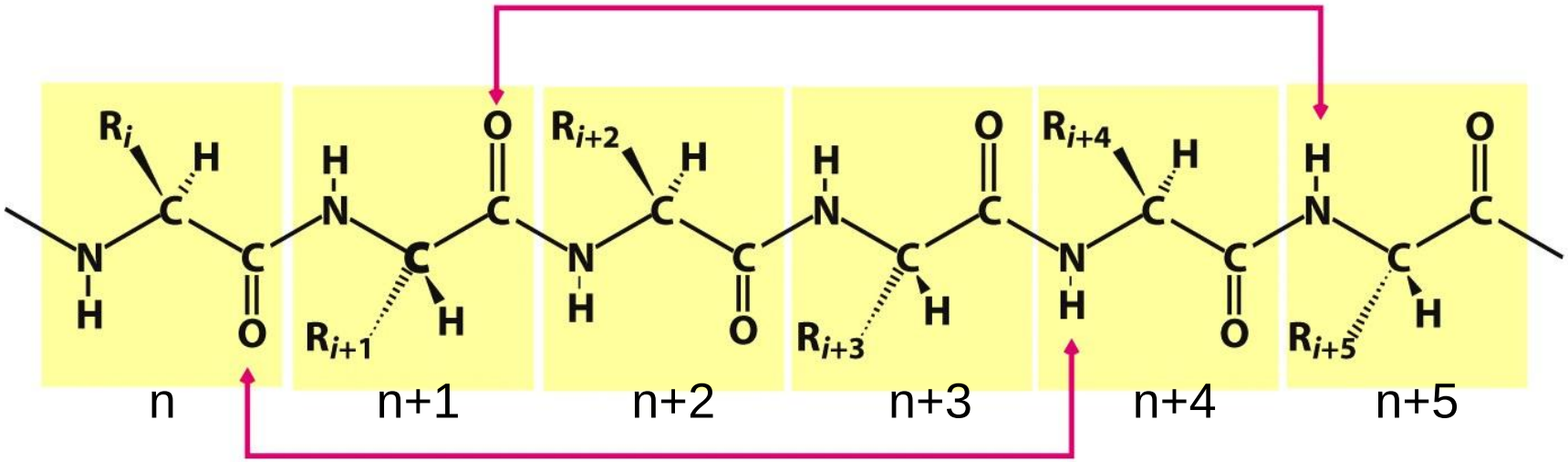


Every 3.6 residues make one turn

The distance (pitch of helix) between two turns is 5.4 Å



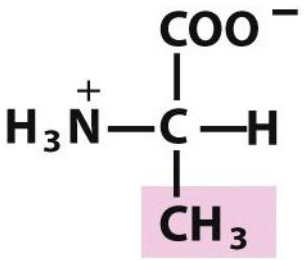
The C=O of residue ‘n’ is hydrogen bonded to N-H of residue ‘n+4’



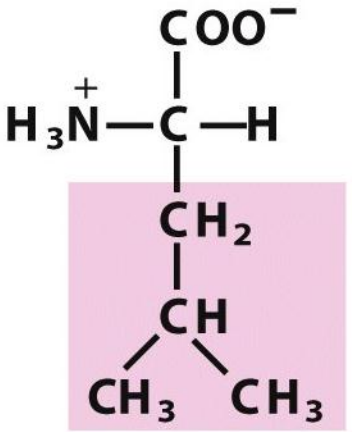
Some Amino Acids are Preferred in α-Helices

Good helix formers:

Ala , Glu, Leu , Met



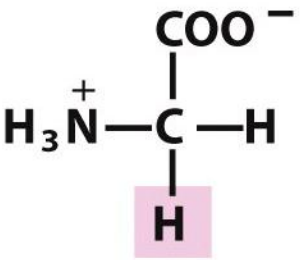
Alanine



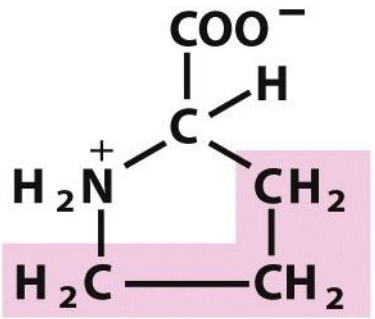
Leucine

Less Preferred:

Pro, Gly, Tyr, Ser



Glycine



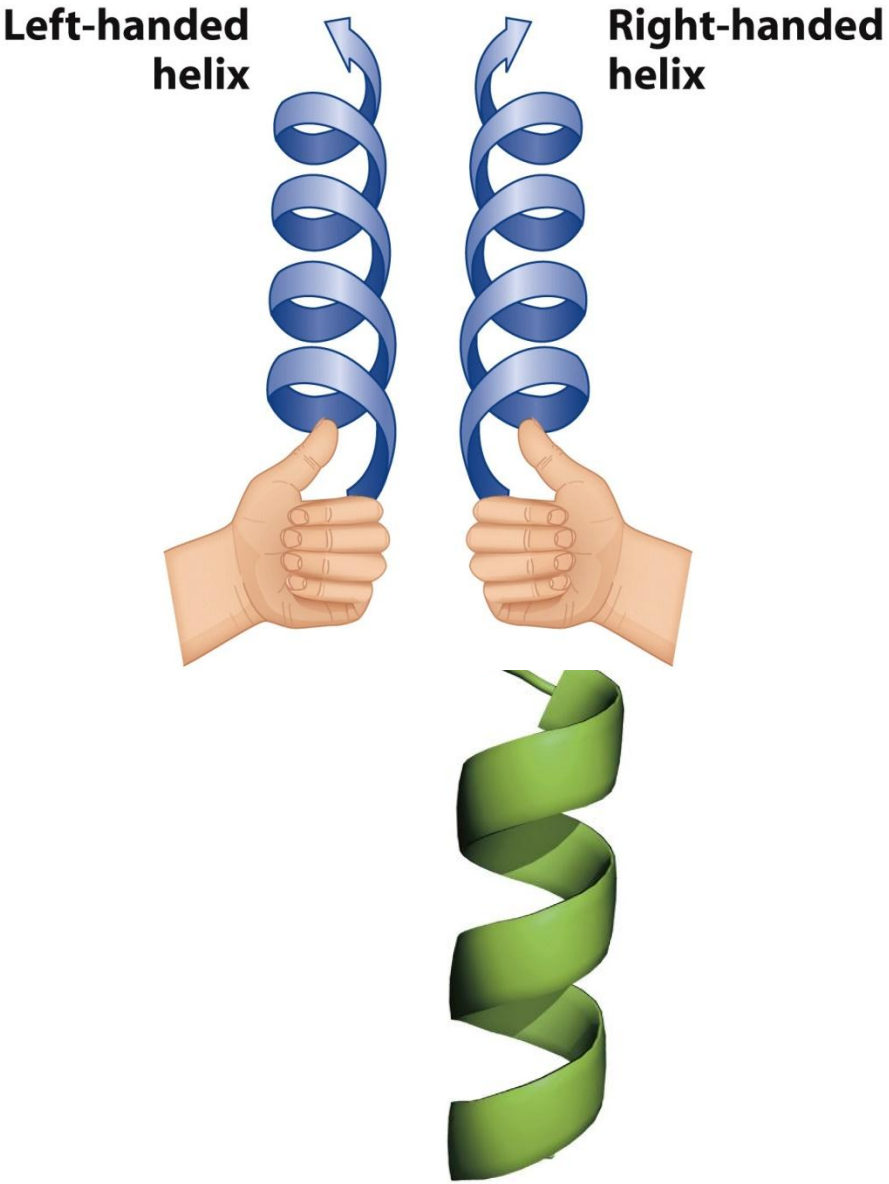
Proline

Alpha Helix: Right-handed or Left-handed?

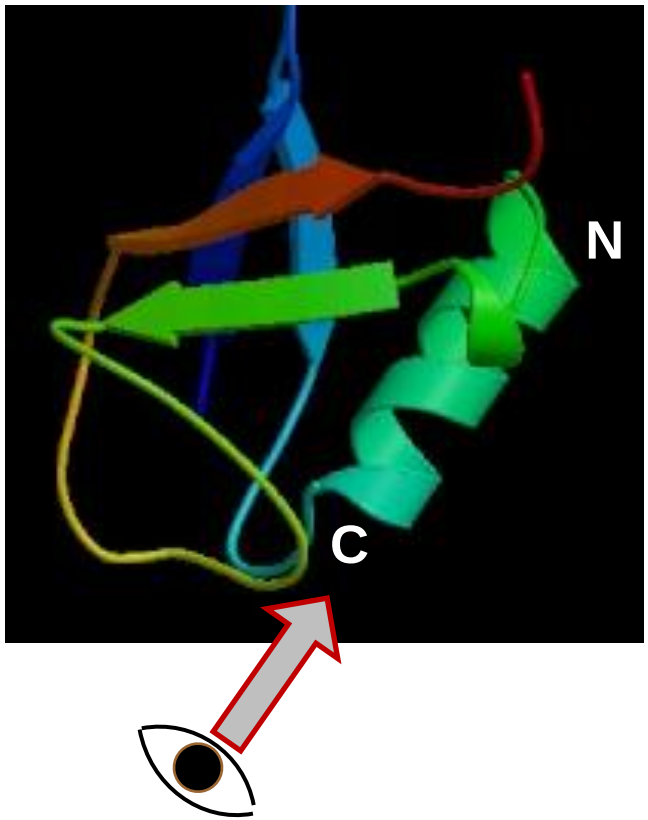
Alpha helix can be – Right-handed or Left handed

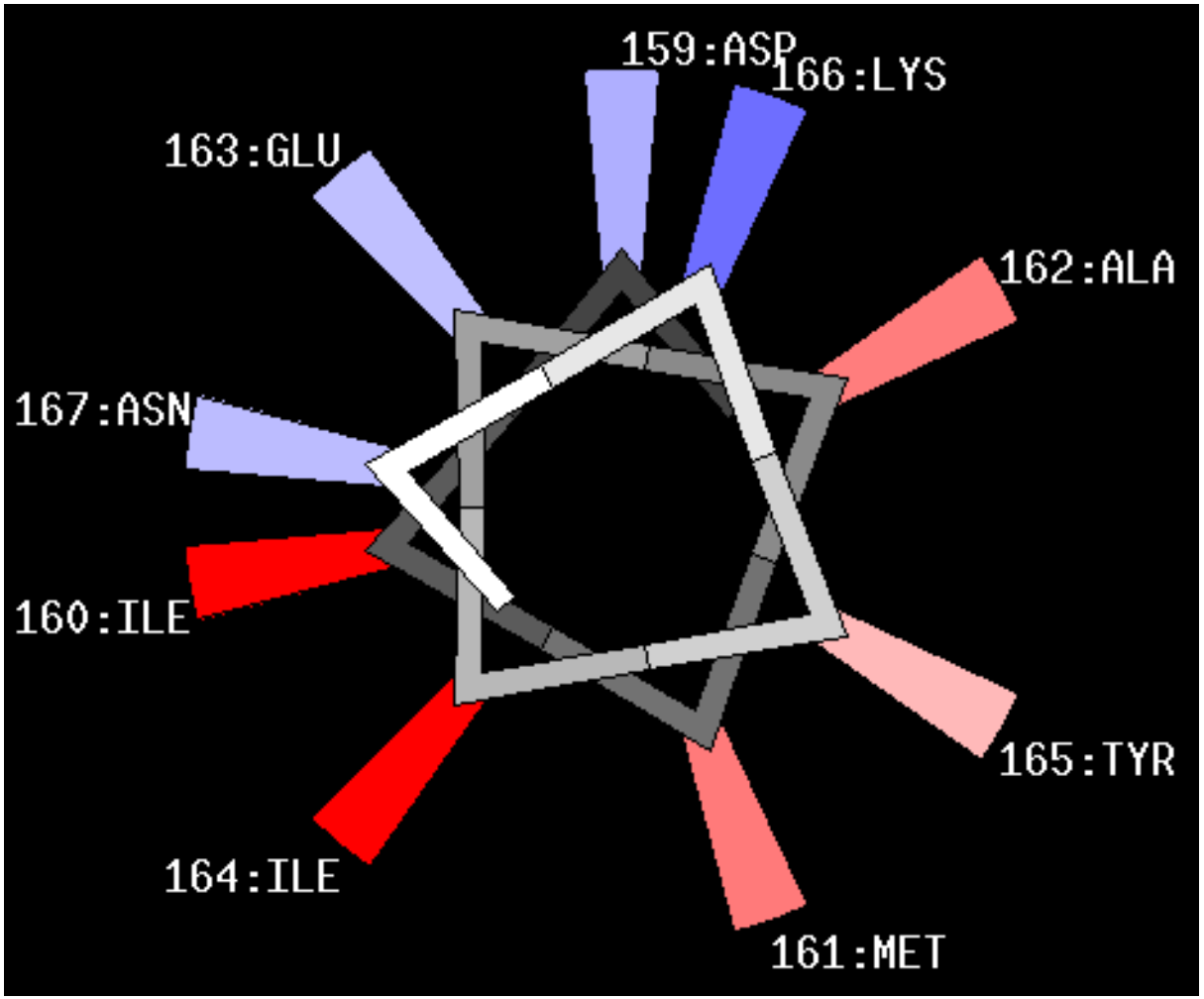
BUT, left handed helix is not possible for L-amino acids due to close approach of the side chains and CO group.

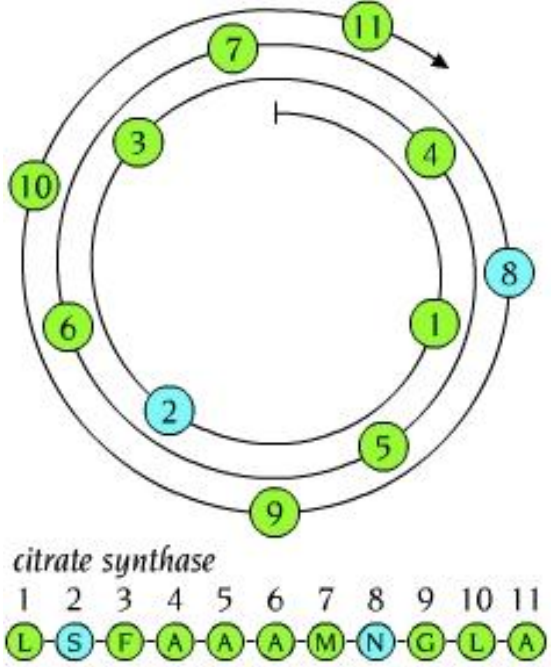
Right handed – most commonly observed in proteins.



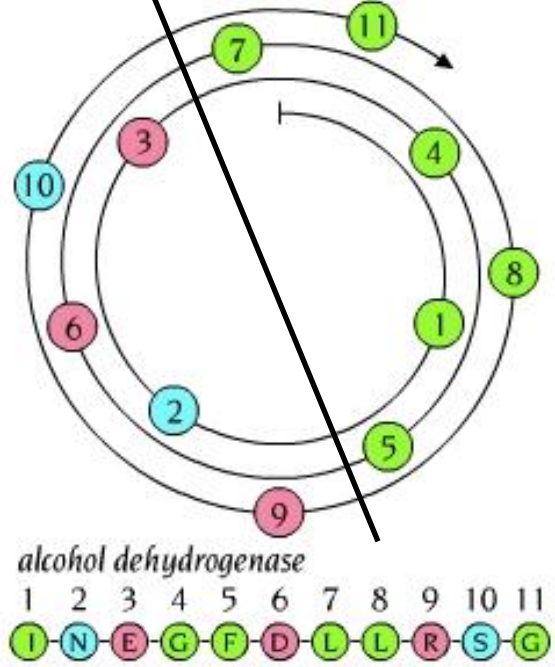
Helical Wheel Plot







Totally buried

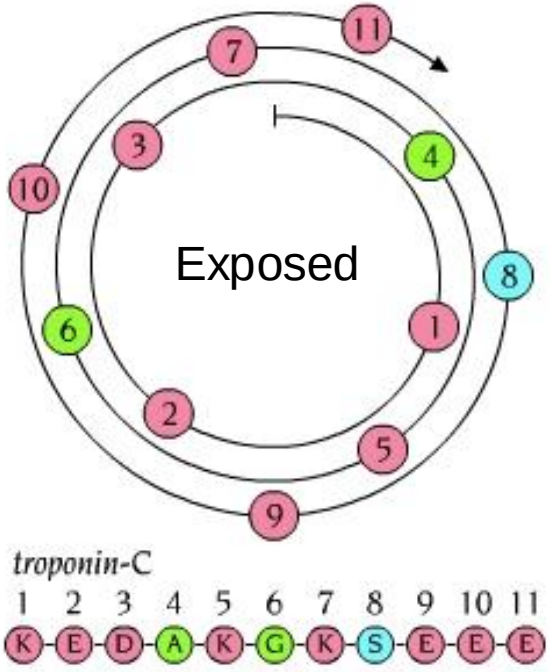


Partially buried

Hydrophobic

Hydrophilic

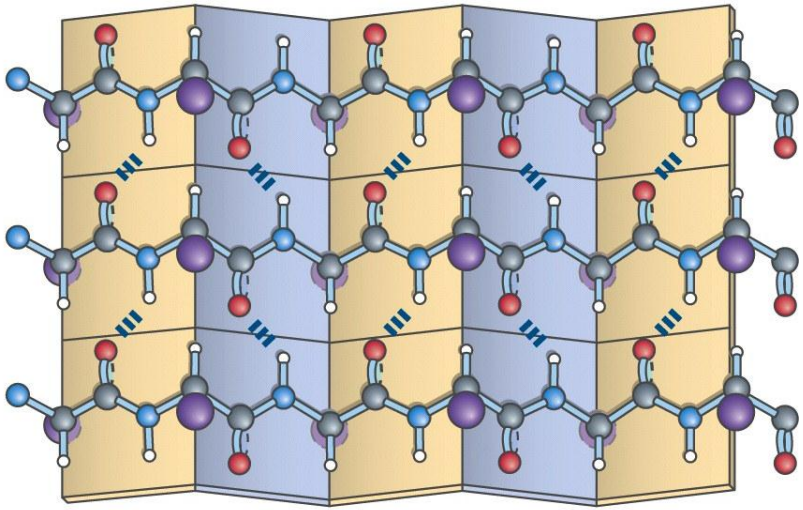
Charged



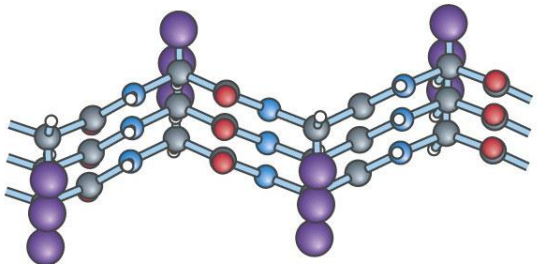
Helical Wheel: Each residue can be plotted every 360/3.6=100° around a circle or spiral

β-sheet
(Number of β-Strands are Involved)

Top view



Side view



β-sheet from several regions of the chain; Each β-strand, typically 5-10 residues long

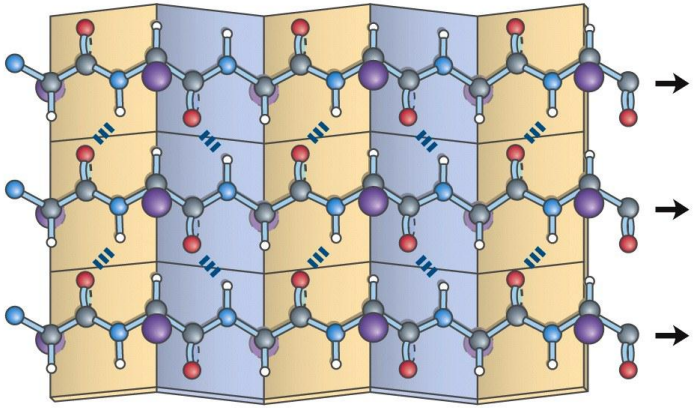
H-bonds are perpendicular to strands

α-helix: from one continuous region

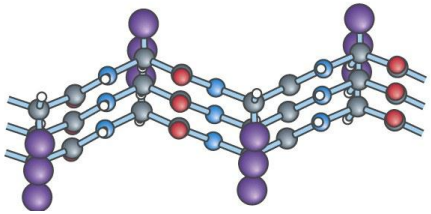
Parallel and Antiparallel β-sheet

Parallel

Top view

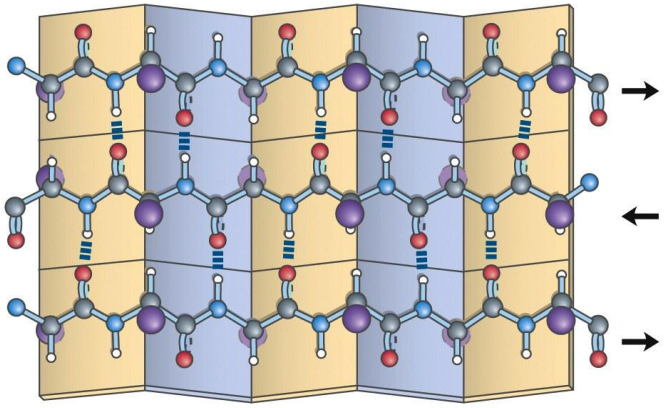


Side view

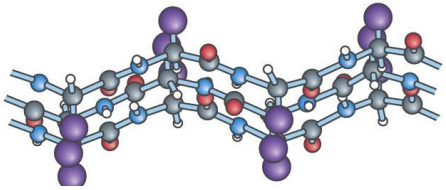


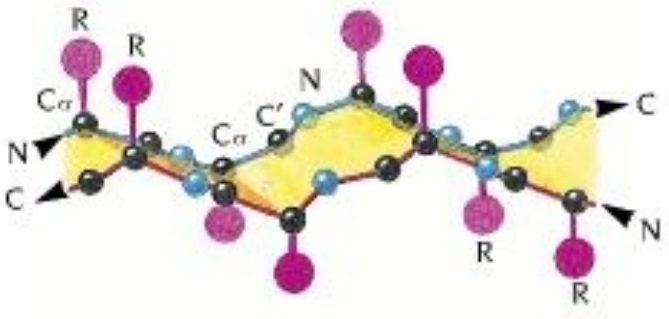
Antiparallel

Top view



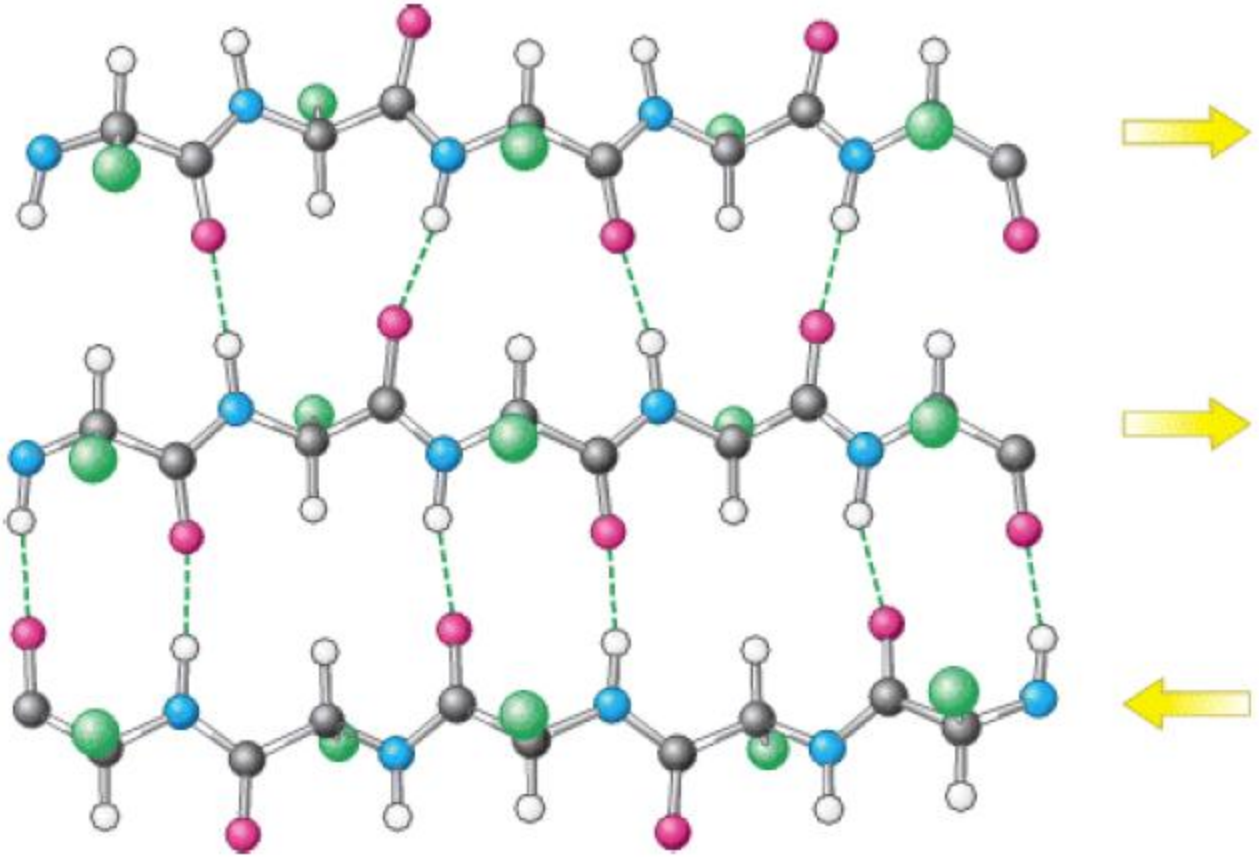
Side view





β-pleated sheet: 'pleated' because side chains point up and down alternatively

Mixed β-sheet



Hair Keratin

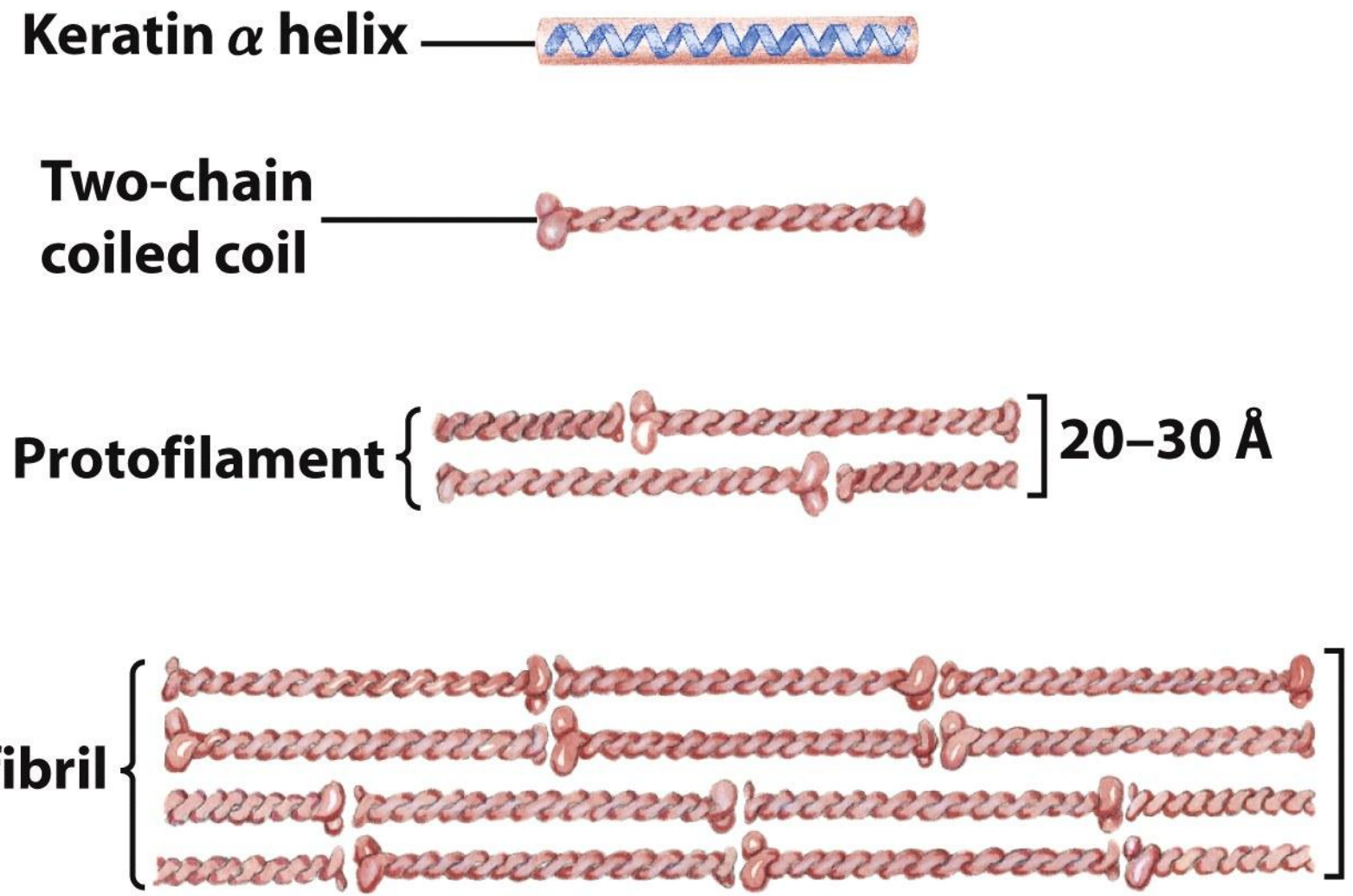
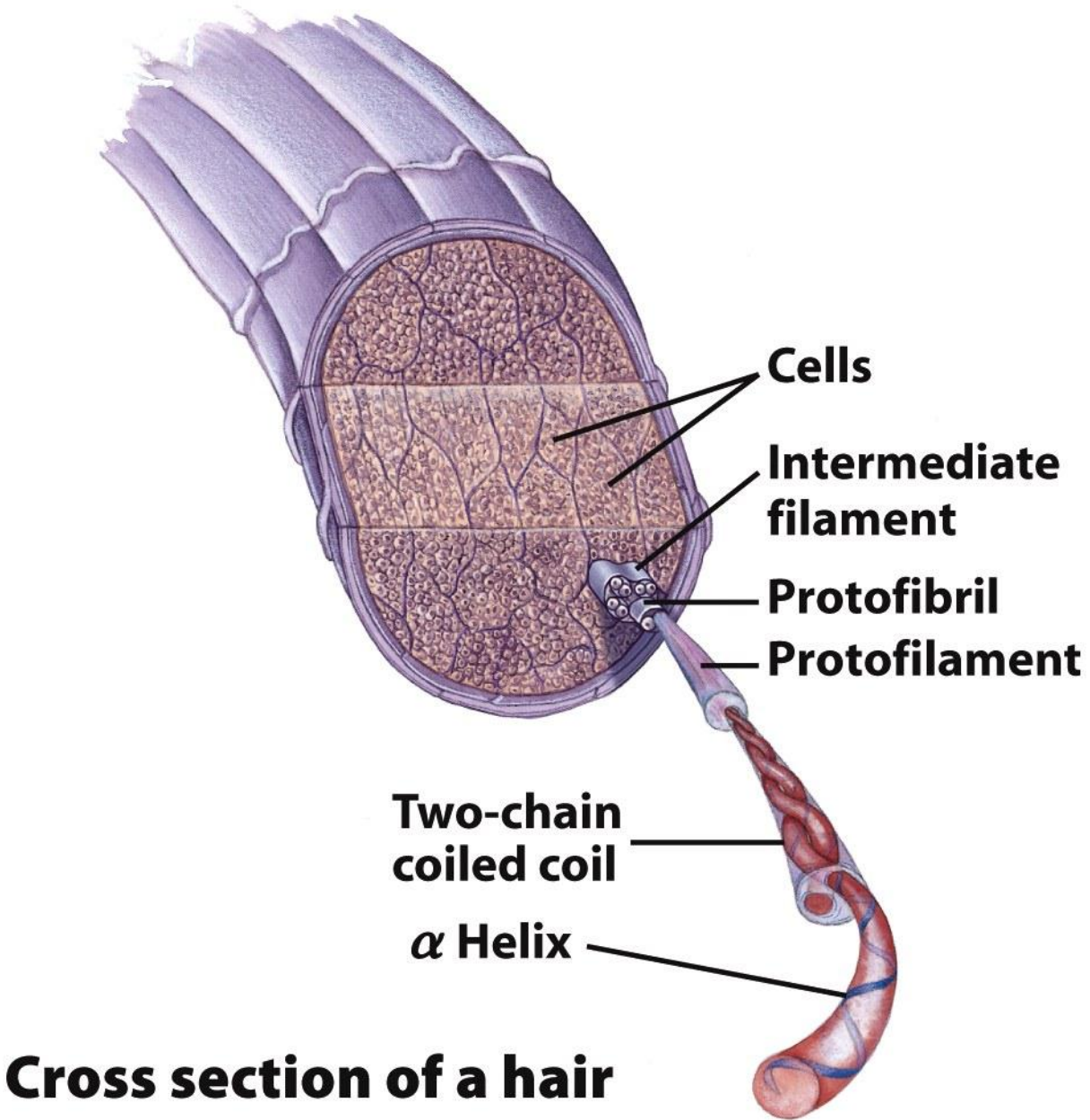
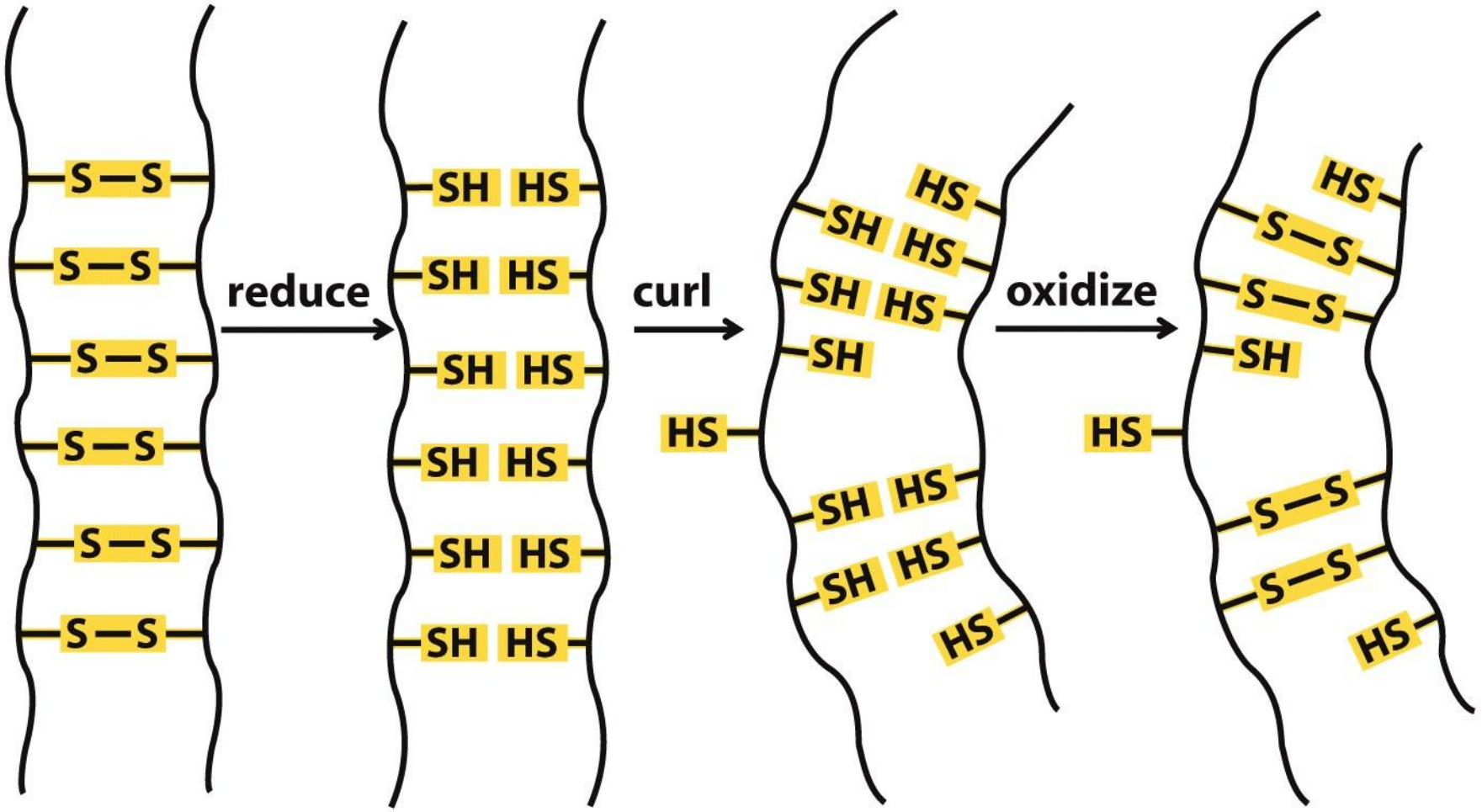


Figure 4-10a
Lehninger Principles of Biochemistry, Fifth Edition

Hair Keratin

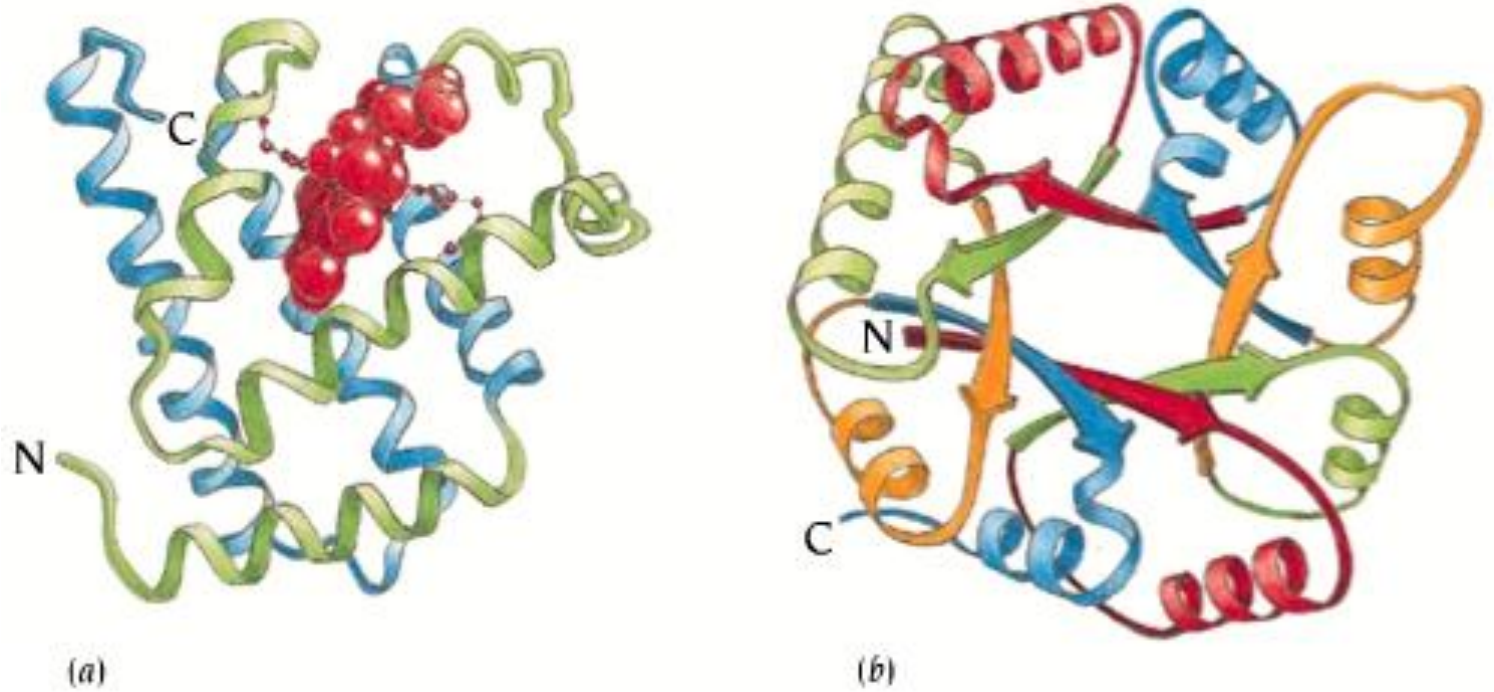


Chemistry of Straight and Curl Hair



Polypeptide Chains Fold into Several Domains

- Fundamental unit of tertiary structure – DOMAIN
- Domain: polypeptide chain or a part of polypeptide chain that can independently fold into a stable tertiary structure
- Domains are also units of function



Tertiary structure refers to the spatial arrangement of amino acid residues that are far apart in the sequence and to the pattern of disulfide bonds.

Quaternary Structure

Proteins containing more than one polypeptide chain exhibit a fourth level of structural organization. Each polypeptide chain in such a protein is called a subunit. Quaternary structure refers to the spatial arrangement of subunits and the nature of their interactions.

The simplest quaternary structure is a dimer, consisting of two identical subunits.

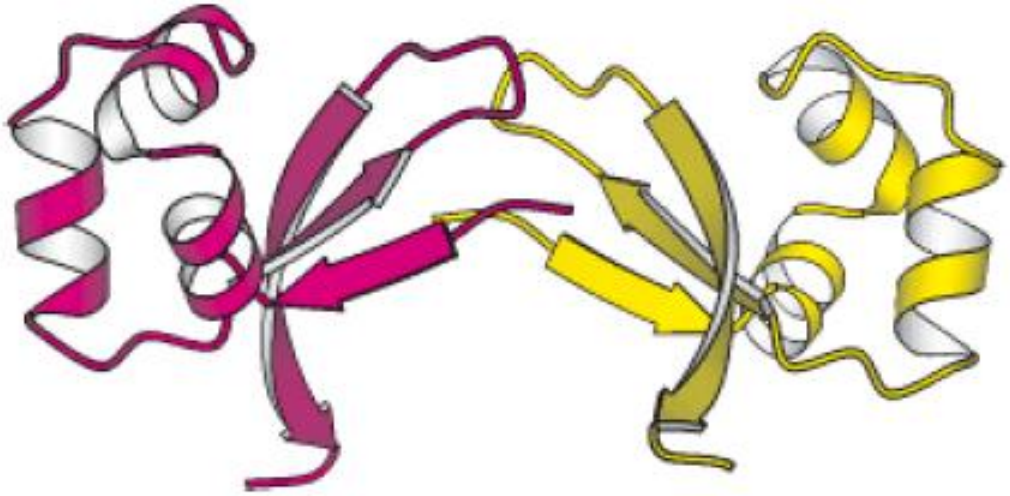
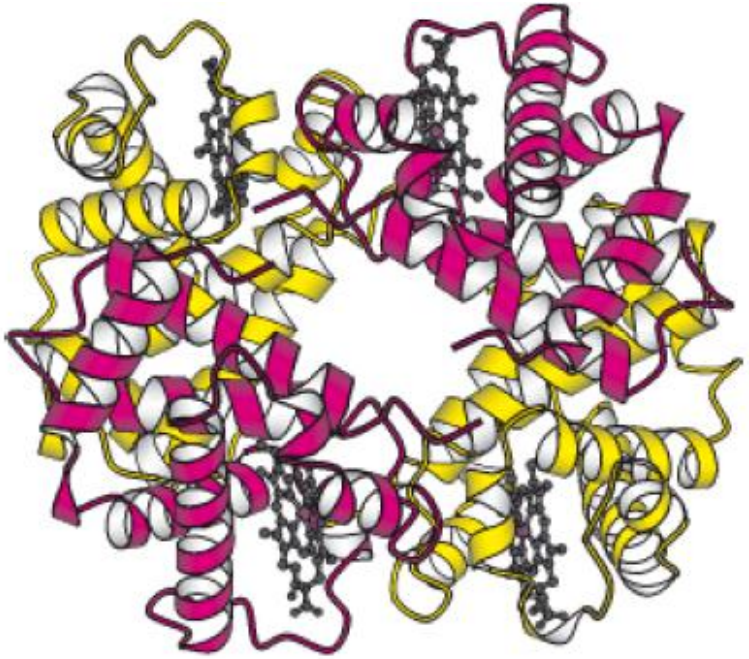
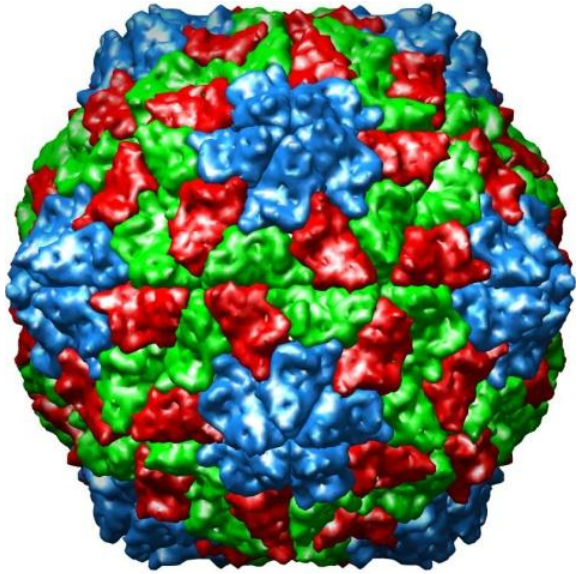


Figure 3.48. Quaternary Structure. The Cro protein of bacteriophage λ is a dimer of identical subunits.

Quaternary Structure (higher order)



The α2β2 tetramer of human haemoglobin. The structure of the two identical α subunits (red) is similar to but not identical with that of the two β subunits (yellow).



Complex Quaternary Structure. The coat of rhinovirus comprises 60 copies of each subunits

Solving Protein Structures

Only 2 kinds of techniques allow one to get atomic resolution pictures of macromolecules

- Structure ⟷ Function
- Structure ⟷ Mechanism
- Structure ⟷ Origins/Evolution
- Structure-based Drug Design
- Solving the Protein Folding Problem

QHTAWCLTSEQHTAAVIWDCETPGKQNGAYQEDCA
HHHHHHCCEEEEEEEEEEECCHHHHHHHCCCCCCC

H: α-helix

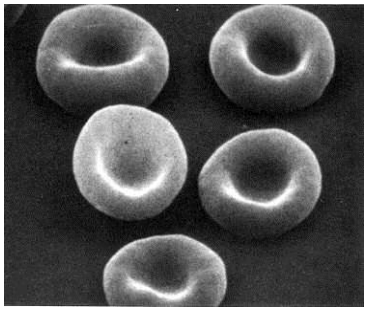
E: β-strand

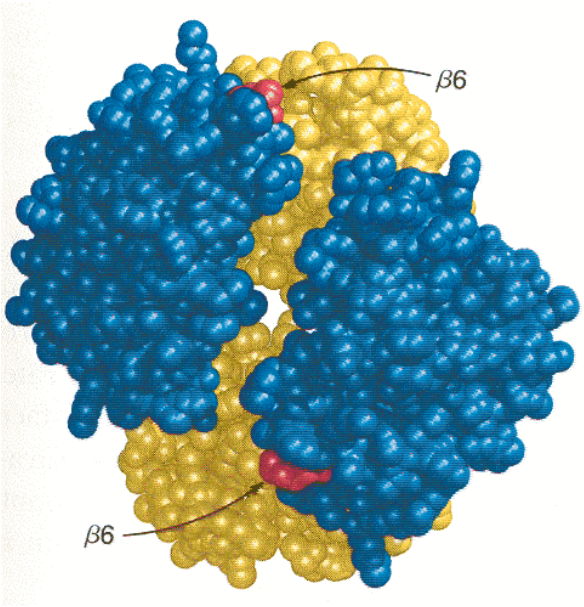
C: unstructured

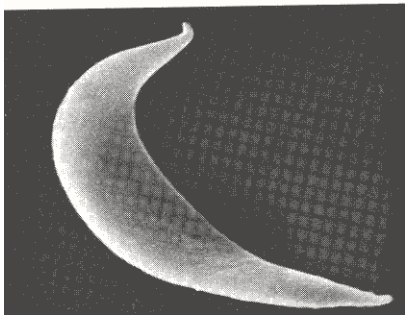
Importance of Protein Structure

Hemoglobin A: Val-His-Leu-Thr-Pro-Glu-Glu-Lys-

Hemoglobin S: Val-His-Leu-Thr-Pro-Val-Glu-Lys-







"sticky patch" causes hemoglobin S to agglutinate (stick together) and form fibers which deform the red blood cell

https://www.youtube.com/watch?v=qBRFIMcxZNM

Books Followed:

Biochemistry (Lubert Stryer)

Principles of Biochemistry (Nelson and Cox)