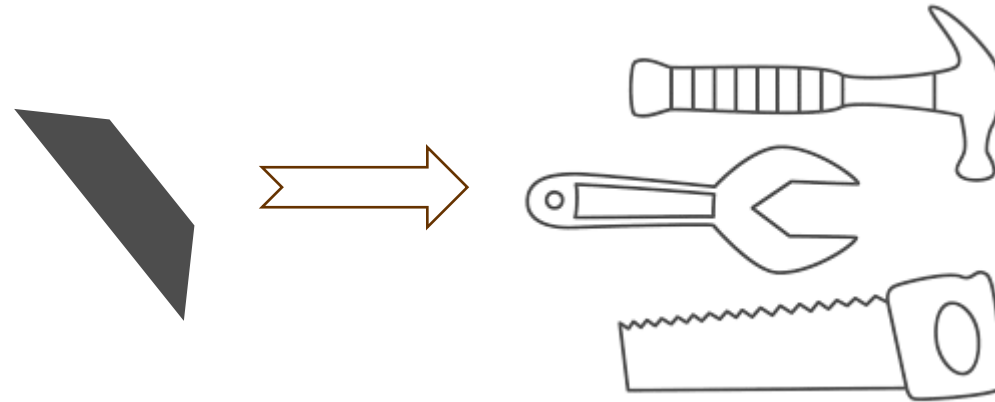


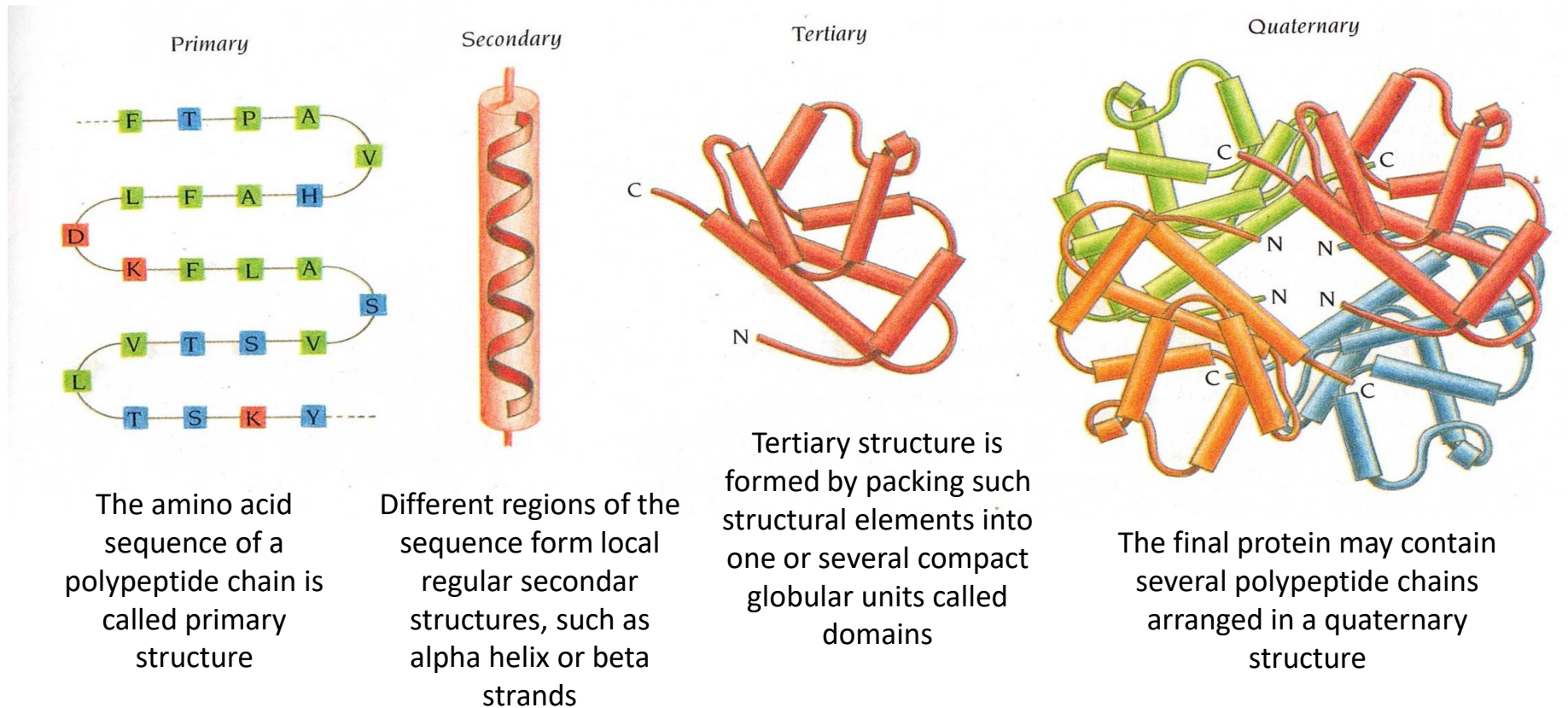
LS2101-Biochemistry

Structure function correlation of Proteins:

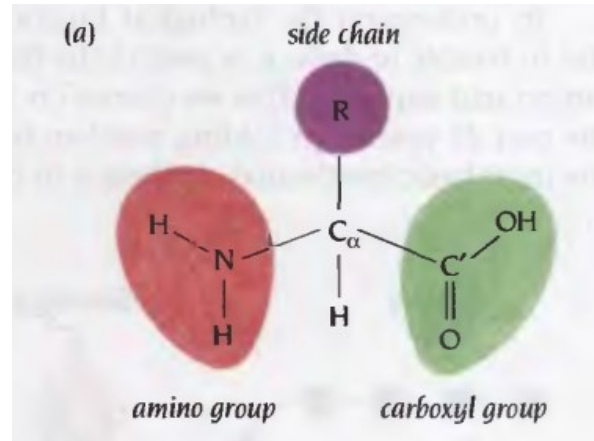


- Function of each protein is absolutely determined by the its 3D structure

- Proteins have evolved in nature to perform specific functions
- Functional properties of protein depend upon three-dimensional structure
- Functionally active structure of protein arises because of particular sequence of amino acids in polypeptide chains fold to generate compact domains with specific three dimensional structure

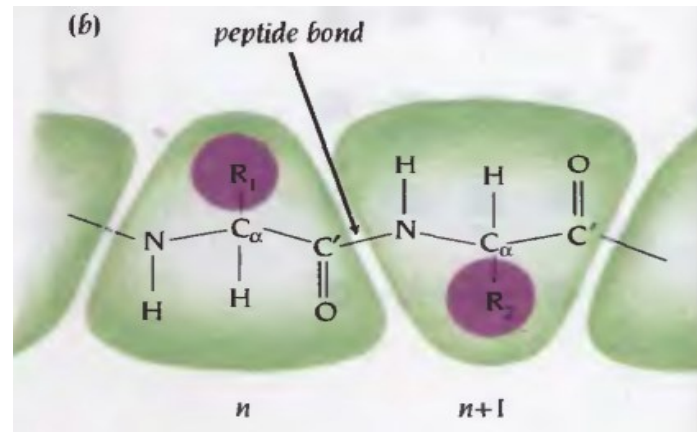


Protein structure



Amino acid – building block
...all amino acids are chiral except glycine
L-form

Amino acids are joined by peptide bond in protein



Repeating unit: amino acid residue

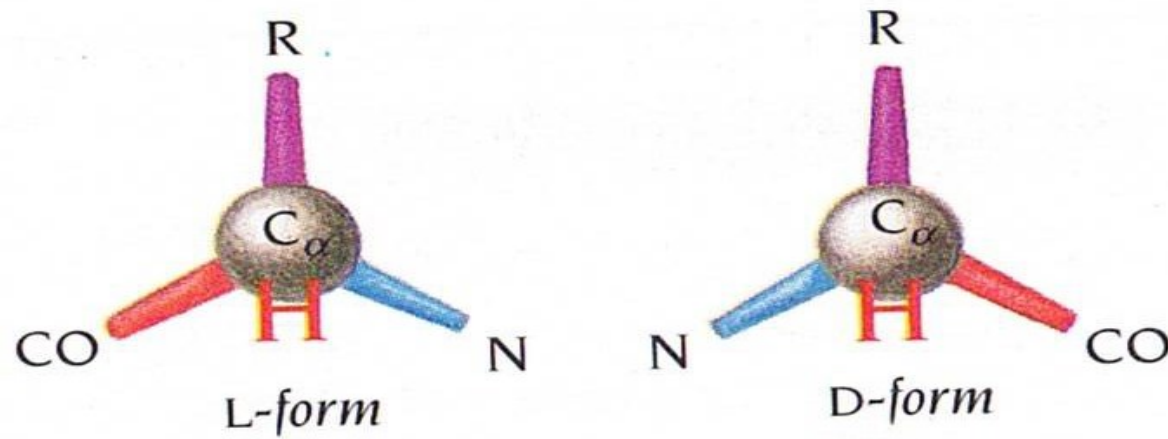


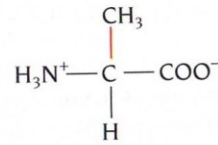
Figure 1.3 The “handedness” of amino acids. Looking down the H– C_α bond from the hydrogen atom, the L-form has CO, R, and N substituents from C_α going in a clockwise direction. There is a mnemonic to remember this; for the L-form the groups read CORN in clockwise direction.

(d) Glycine

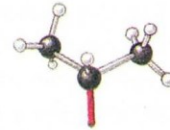


G Gly, Glycine

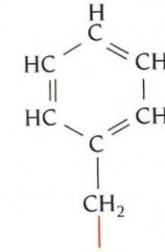
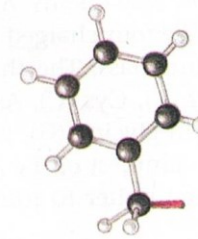
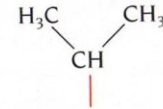
(a) Hydrophobic amino acids



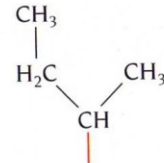
A Ala, Alanine



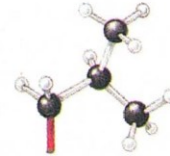
V Val, Valine



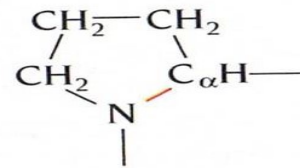
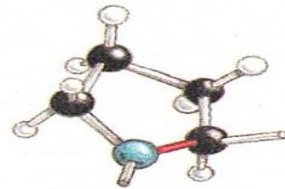
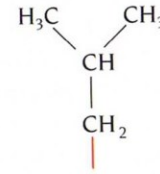
F Phe, Phenylalanine



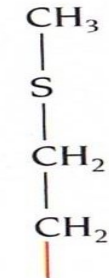
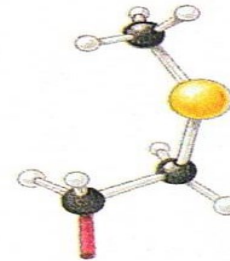
I Ile, Isoleucine



L Leu, Leucine

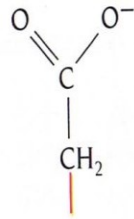


P Pro, Proline

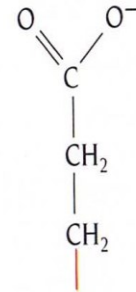
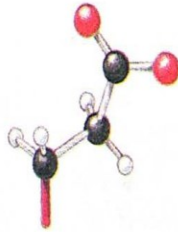


M Met, Methionine

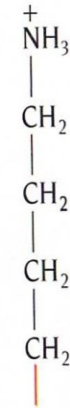
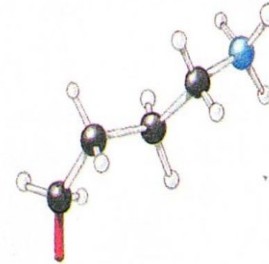
(b) Charged amino acids



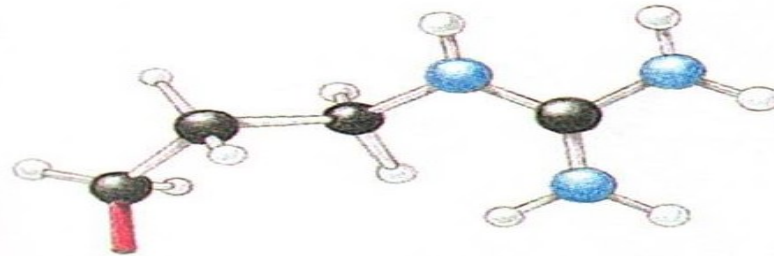
D Asp, Aspartic acid



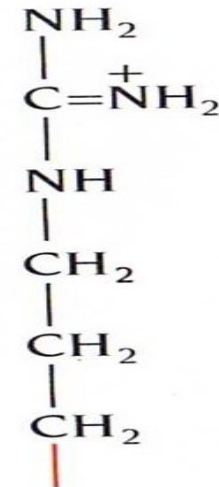
E Glu, Glutamic acid



K Lys, Lysine



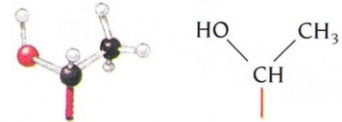
R Arg, Arginine



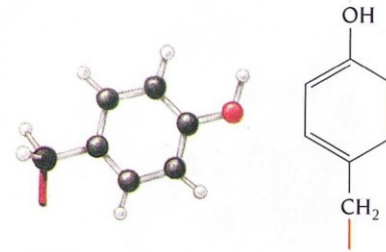
(c) Polar amino acids



S Ser, Serine



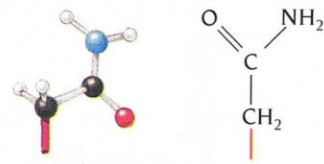
T Thr, Threonine



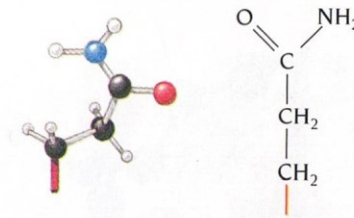
Y Tyr, Tyrosine



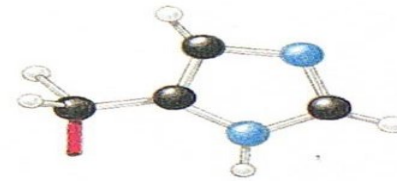
C Cys, Cysteine



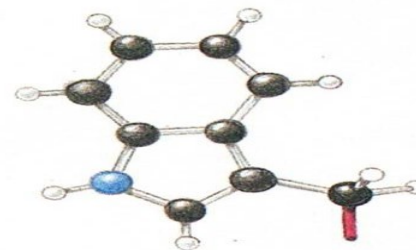
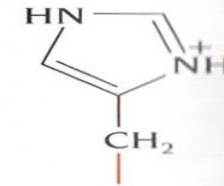
N Asn, Asparagine



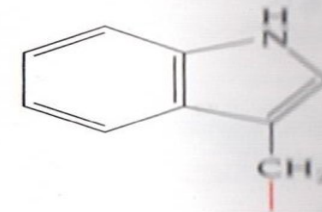
Q Gln, Glutamine



H His, Histidine



W Trp, Tryptophan



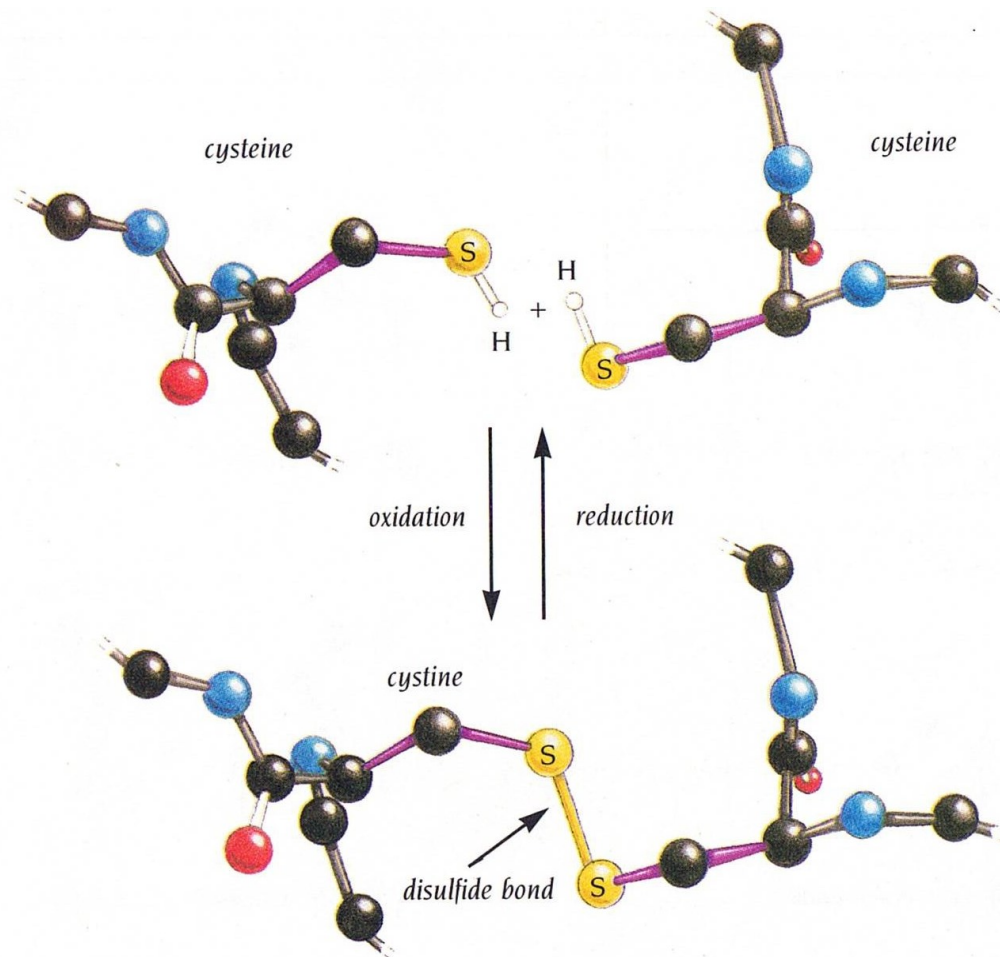
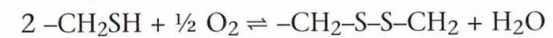


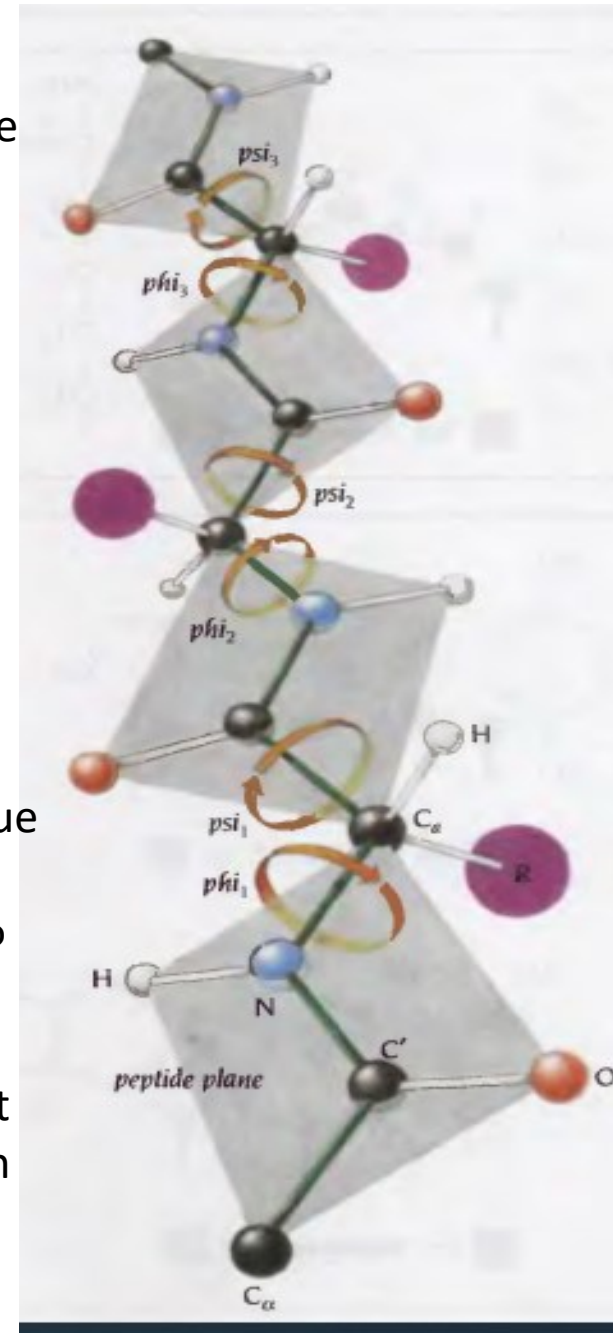
Figure 1.4 The disulfide is usually the end product of air oxidation according to the following schematic reaction scheme:



Disulfide bonds form between the side chains of two cysteine residues. Two SH groups from cysteine residues, which may be in different parts of the amino acid sequence but adjacent in the three-dimensional structure, are oxidized to form one S-S (disulfide) group.

➤ Different way to visualise the polypeptide chain:

- Peptide units do not have the side chains
- Peptide units are rigid
- Each unit can rotate around $C_\alpha-C'$ and $N-C_\alpha$ bond
- Angle of rotation around $N-C_\alpha$: phi (ϕ)
- Angle of rotation around $C_\alpha-C'$: psi (ψ)
- Each amino acid residue is associated with two conformational angle: ϕ and ψ
- These are the only degrees of freedom that define the conformation of whole main chain



First X-ray crystallographic structure of Protein in 1958: Myoglobin



➤ Sausage shaped region of Myoglobin represents α helices

John Kendrew discovered crystal structure of Myoglobin in 1958.....“Perhaps the most remarkable features of the molecule are its complexity and its lack of symmetry it is more complicated than has been predicted by any theory of protein structure”

=> Such structural irregularity is required for proteins to fulfill their diverse function

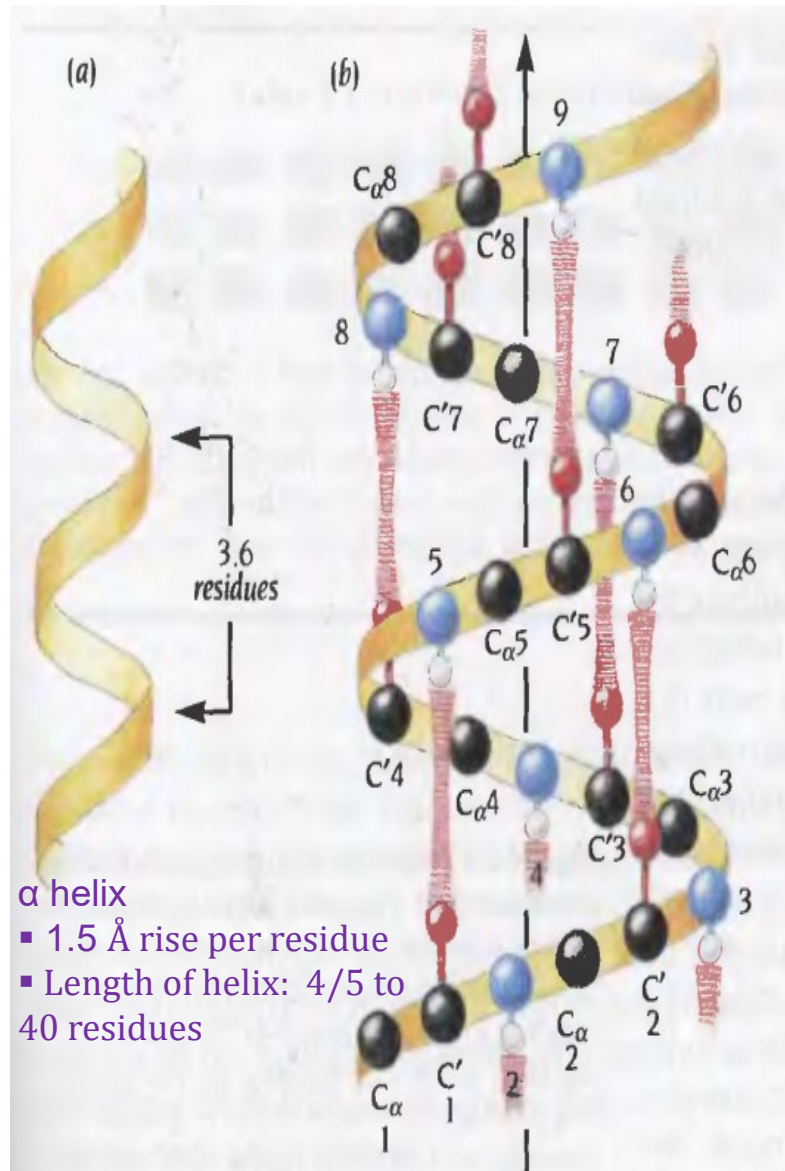
✓ The main driving force for folding water soluble protein molecules is to pack hydrophobic side chains into the interior of the molecule, thus creating hydrophobic core and hydrophilic surface

➤ The main chain is highly polar: one hydrogen bond donor, NH and one hydrogen bond acceptor, C'=O for each peptide unit : Main chain polar groups can be neutralised via H-bonds

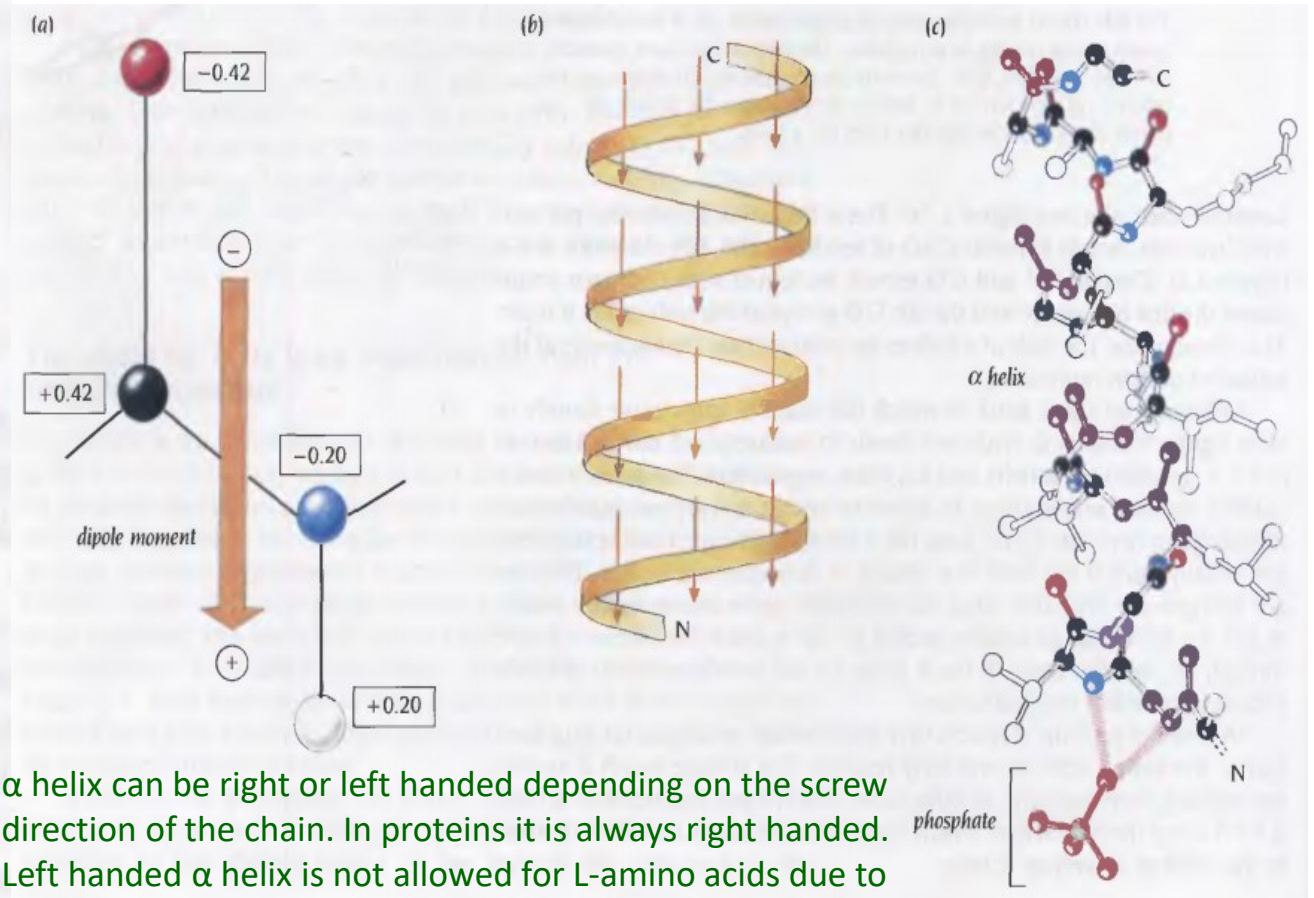
➤ The problem can be solved by the formation of two types of secondary structure : α helices and β sheets \Rightarrow Both types are characterised by hydrogen bonding between main chain NH and C'=O groups

➤ The secondary structure elements, formed in this way and held together by the hydrophobic core, provide a rigid stable frame work

- α helix has 3.6 residue per turn with H-bond between $C'=O$ of residue n and NH of residue $n+4$
- Main chain NH and $C'=O$ groups are joined with H-bond except first NH group at N-terminus and $C'=O$ group at C-terminus \Rightarrow α helix has a dipole moment with magnitude about 0.5–0.7 unit charge at each end of the helix



- -vely charged ligands, e.g., PO_4^{3-} , bind at N-terminus
- +vely charged ligands bind at C-terminus



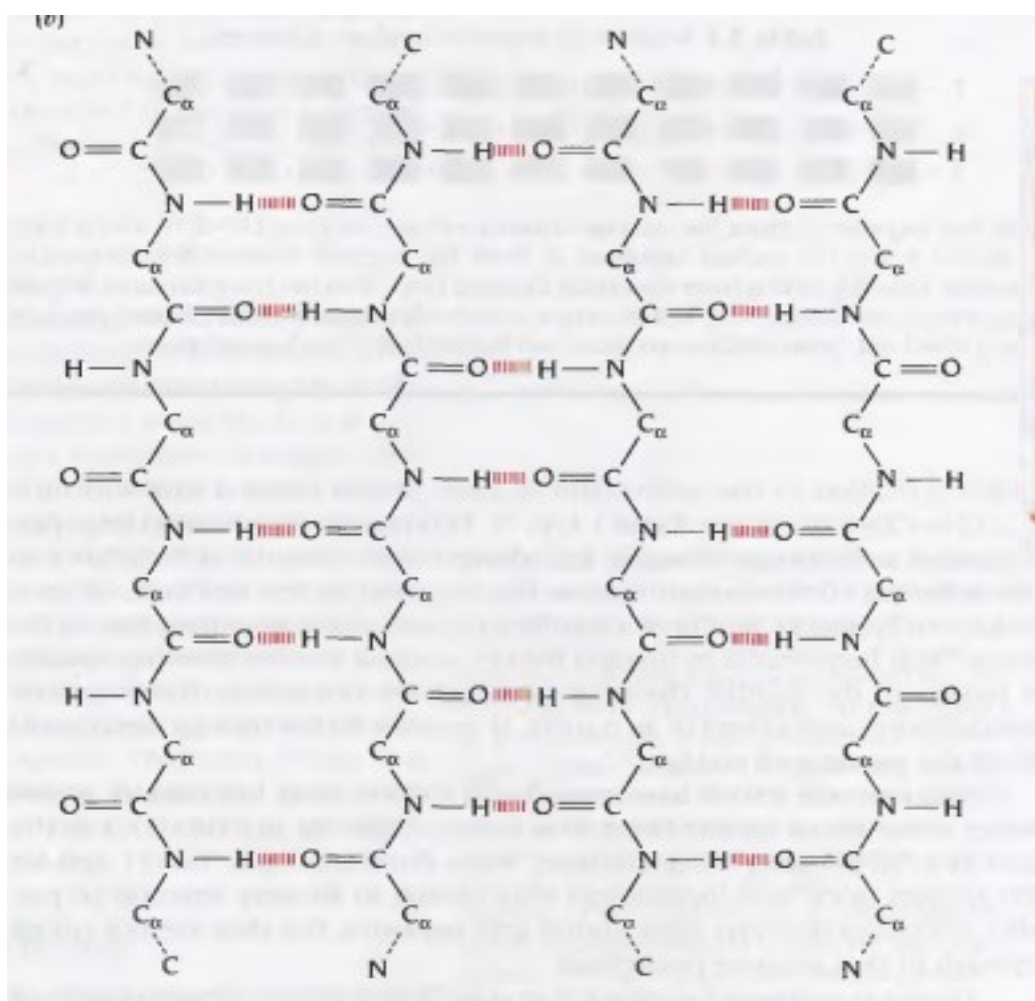
α helix can be right or left handed depending on the screw direction of the chain. In proteins it is always right handed. Left handed α helix is not allowed for L-amino acids due to close approach of the side chains and CO group.

α helices are almost always at the surface of proteins

It also occurs occasionally with 3 to 5 residues

π helix: H-bond to residues $n+5 \Rightarrow$ more loosely coiled than α helix

3_{10} helix: H-bond to residues $n+3 \Rightarrow$ more tightly coiled than α helix

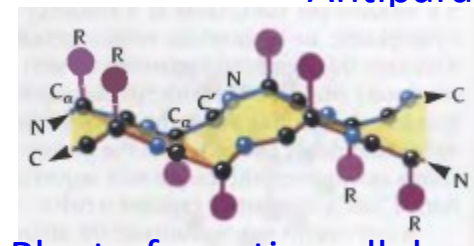


Antiparallel β sheet



β strand

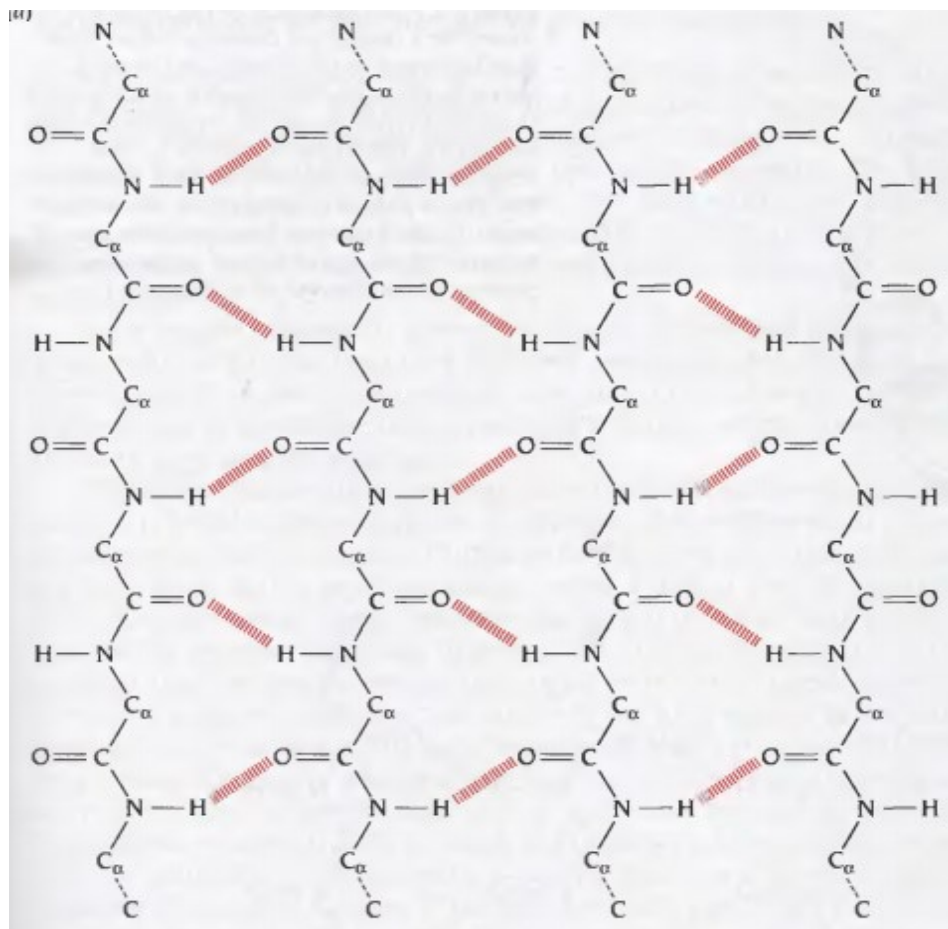
- β sheet is built up from a combination of several regions of the polypeptide chain (in contrast to α helix, built up from one continuous region)
- β strands are usually 5 to 10 residue long
- β strands are adjacent in the β sheets where H-bonds form between $C'=O$ group of one β strand and NH group of adjacent β strand



Pleat of a antiparallel β sheet

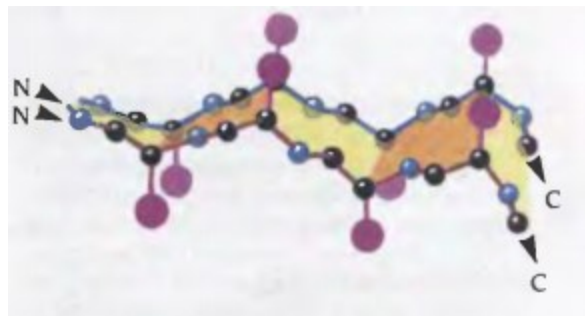
β sheets formed from several β strands are pleated with C_{α} atoms successively a little above and below the plane of β sheet

- Antiparallel β sheet: Amino acids in the successive strands run in alternating directions, amino terminal to carboxy terminal followed by carboxy terminal to amino terminal and so on
- Antiparallel β sheet has narrowly spaced H-bond pairs that alternate with widely spaced pairs
- Except two flanking β strands of β sheet, all possible main chain H-bonds are formed across constituent β strands

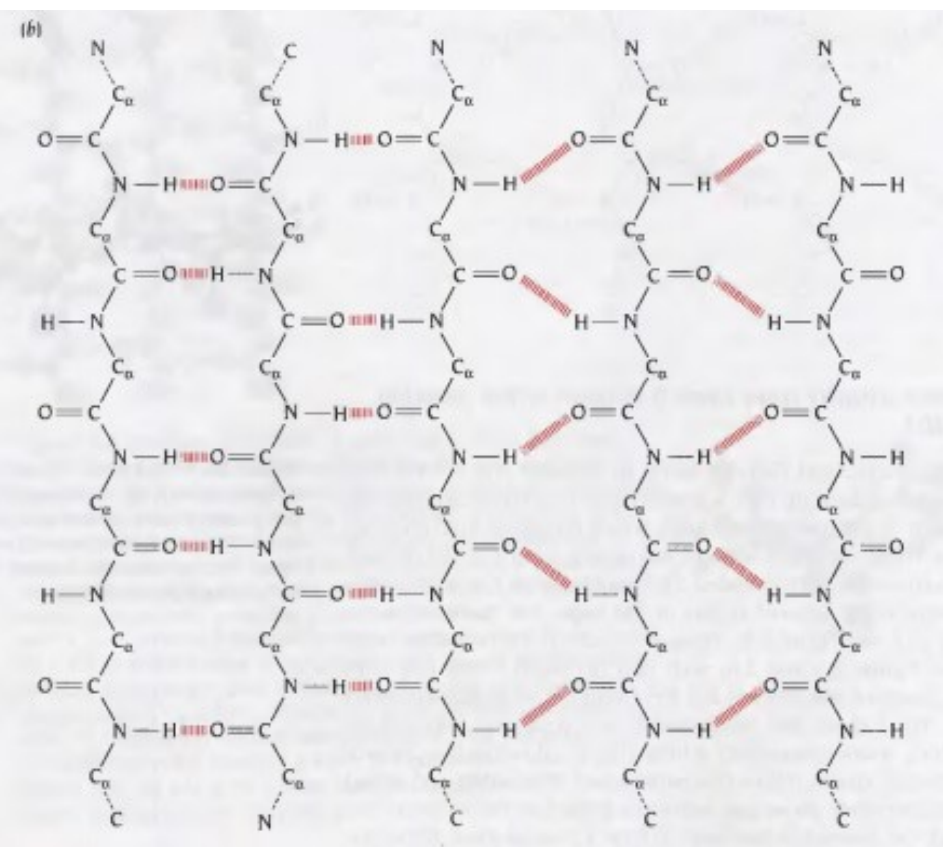


Parallel β sheet

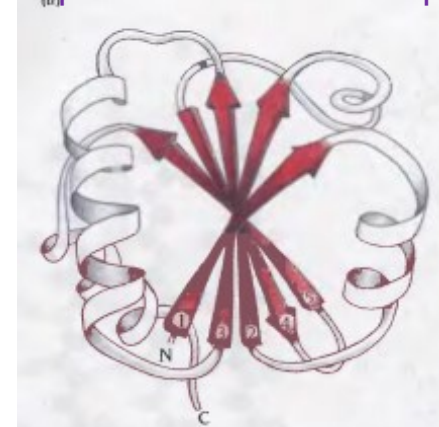
- Amino acids in all strands run in same directions
- Evenly spaced H-bonds that bridge β strands at an angle



Pleat of a parallel β sheet

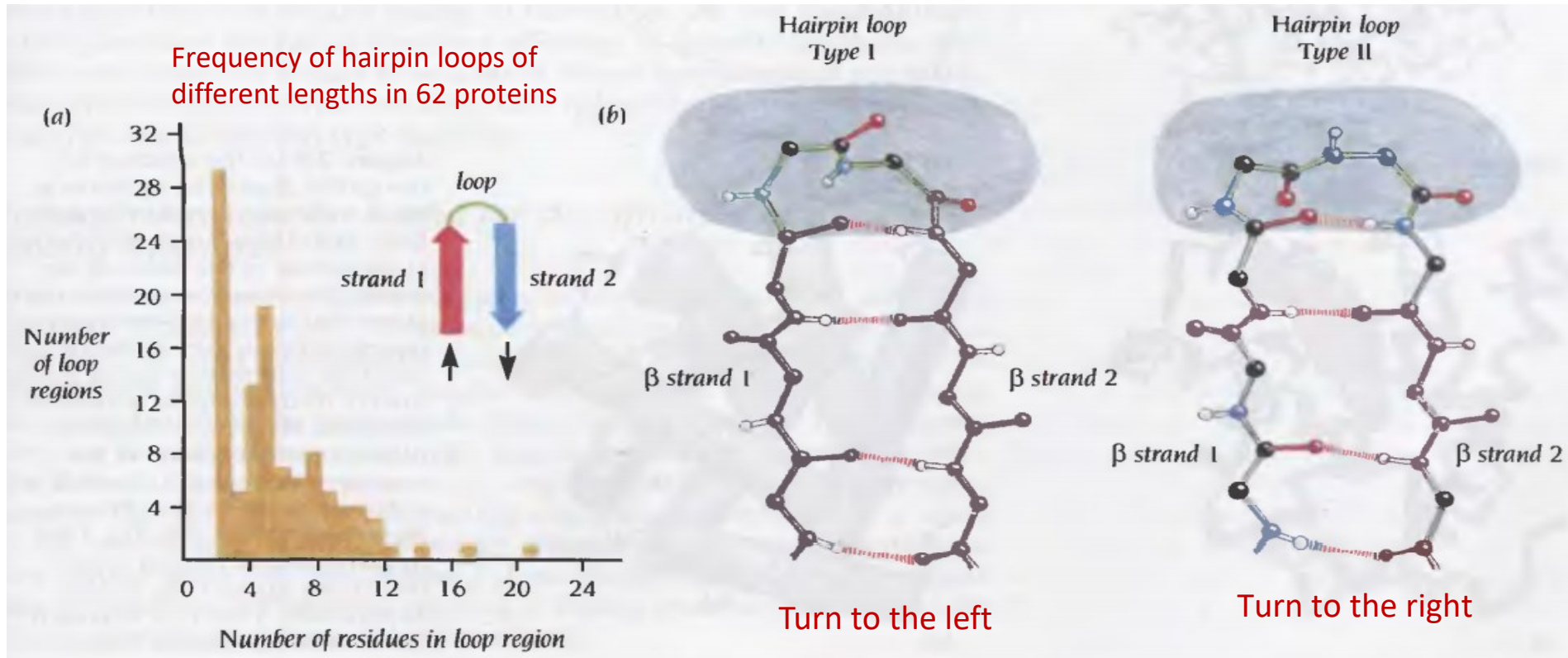


β strands can combine into mixed β sheets with some β strand pairs parallel and some antiparallel



Protein thioredoxin from *E.coli*

In protein structure the secondary structural elements α helices and β strands are connected by loop regions of various lengths and irregular shapes. The loops are at the surface of the molecule and are rich in polar hydrophilic residues. The main chain C'=O and NH groups are exposed to the solvents and form H-bonds with water.

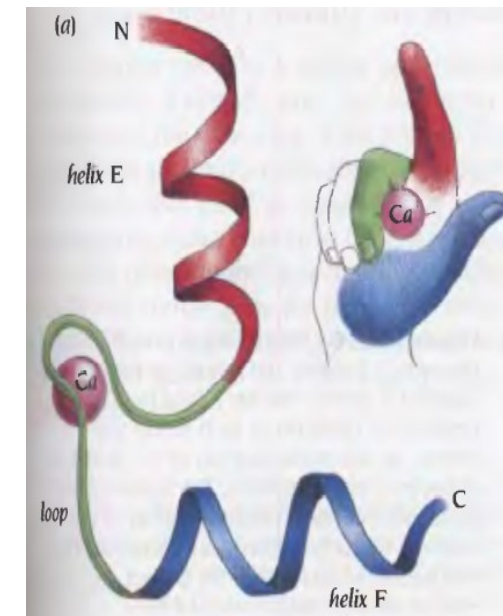


Loops that connect two adjacent β strands are called hairpin loops or reverse turns.

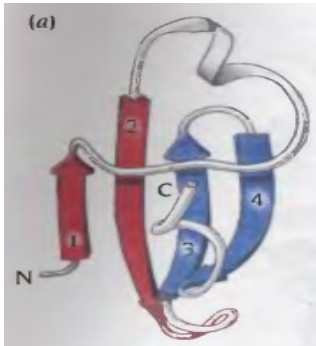
- Long loop regions are flexible and can frequently adopt several different conformation
- Loop regions participates in forming binding sites and enzyme active sites
- Long loops are in many cases susceptible to proteolytic degradation. However it can be protected by binding metal ions, such as calcium ion

Motif: Combination of a few secondary structural elements with a specific geometric arrangement in protein structure is defined as Motif or super-secondary structure. It might have specific function.

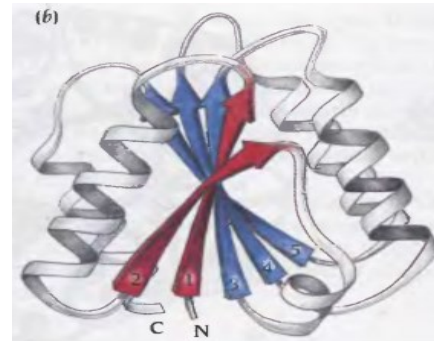
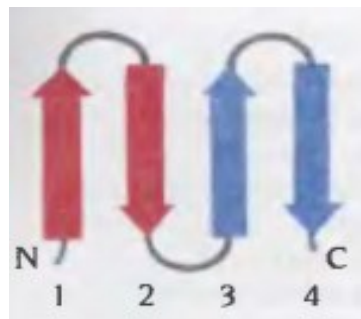
- Two α helices joined by a loop region (helix-loop-helix motif).
- Several different calcium binding protein has this motif, such as parvalbumin, protein troponin-C
- This motif comprises two right handed helices, fifth (E) and sixth (F) helices in the protein structure and a loop of 12 residues.
- Loop region binds calcium ion and the motif provides a scaffold that holds calcium ligands to bind and release calcium during muscle relaxation.



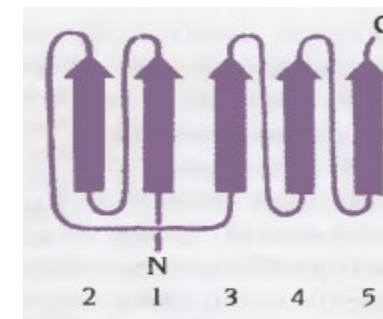
Topology diagram of protein structure: It is the simplified schematic representation of the secondary structure elements, especially β sheet including the number of strands, their relative directions (parallel or antiparallel), and the way strands are connected along polypeptide chain (strand order). However the twist of β sheet is not represented in the topology diagram.



Antiparallel β strands in aspartate transcarbamoylase



Parallel β strands in redox protein flavodoxin

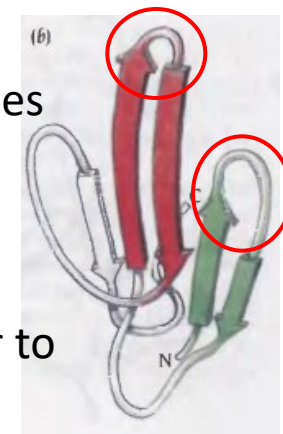


Hairpin β motif:

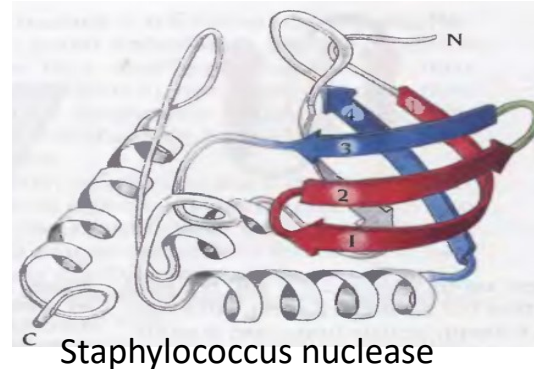
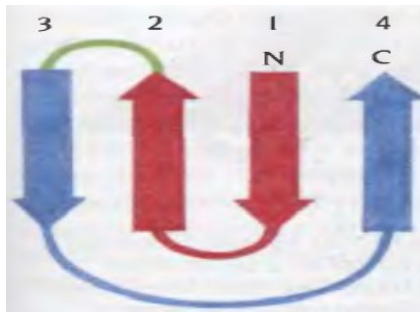
- Two adjacent antiparallel β strands joined by a loop generally from 2 to 5 residues
- ✓ It is present in most antiparallel β structure such as snake venom erabutoxin which binds to and inhibits the acetylcholine receptor in nerve cells

Greek key motif:

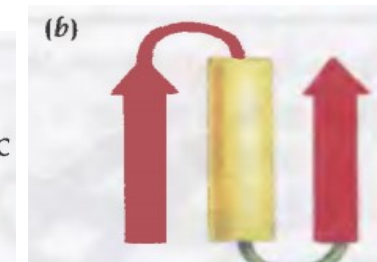
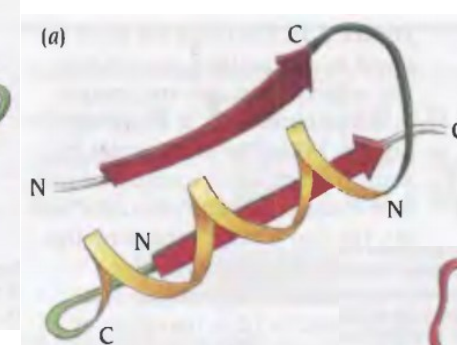
- Four adjacent antiparallel β strands are frequently arranged in a pattern similar to the repeating unit of an ornamental pattern, used in ancient Greece. In protein structure the motif is called Greek key.
- ✓ Staphylococcus nuclease, an enzyme that degrades DNA, has Greek key motif



Erabutoxin

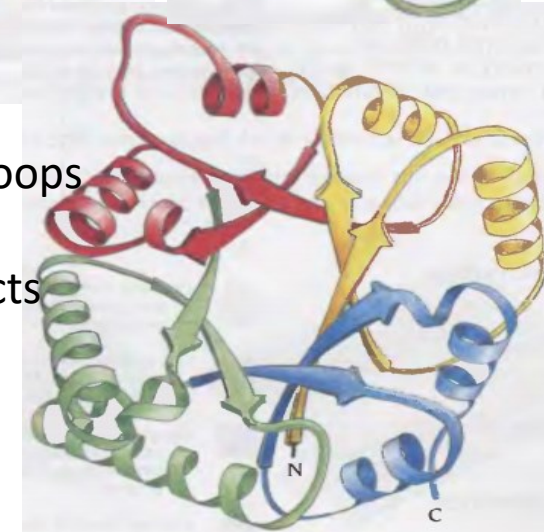


Staphylococcus nuclease



β - α - β motif: It consists of two parallel β strands, an α helix and two loops

- Two adjacent parallel β strands that are consecutive in amino acid sequence are joined through an α helix. α helix in β - α - β motif connects carboxy end of one β strands with amino end of the next β strand.
- The polypeptide chain turns twice using loop regions and it varies from one to two residues to hundred.
- This motif is found out in almost every protein structure having parallel β sheet such as triosephosphate isomerase



Triosephosphate isomerase

Structural hierarchy of protein molecule:

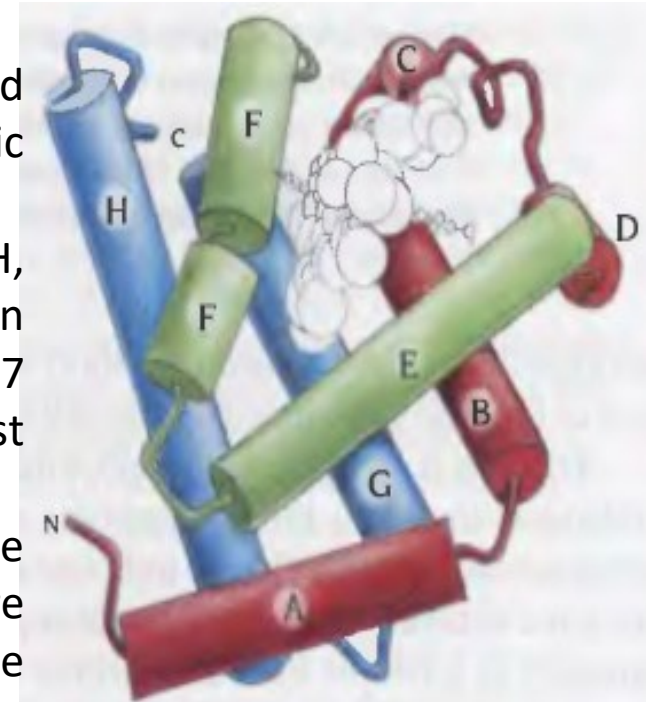
- Primary structure – Amino acid sequence along a polypeptide chain; Two proteins with significant similarity in the primary structure are said to be homologous
- Secondary structure – Secondary structure occurs mainly as α helices and β sheets. Secondary structure elements arrange themselves in simple motif
- Tertiary structure – Several motifs combine to form a compact globular structure called domain which is the fundamental unit of the tertiary structure. Tertiary structure is formed by the packing of domains.
- Quaternary structure – Protein molecules with only one chain are called monomeric and proteins with multiple polypeptide chains are called multimeric proteins. Quaternary structure of a protein consists of several polypeptide chains that are associated into a multimeric molecule.

α domain structure: Globular protein – Myoglobin, Hemoglobin

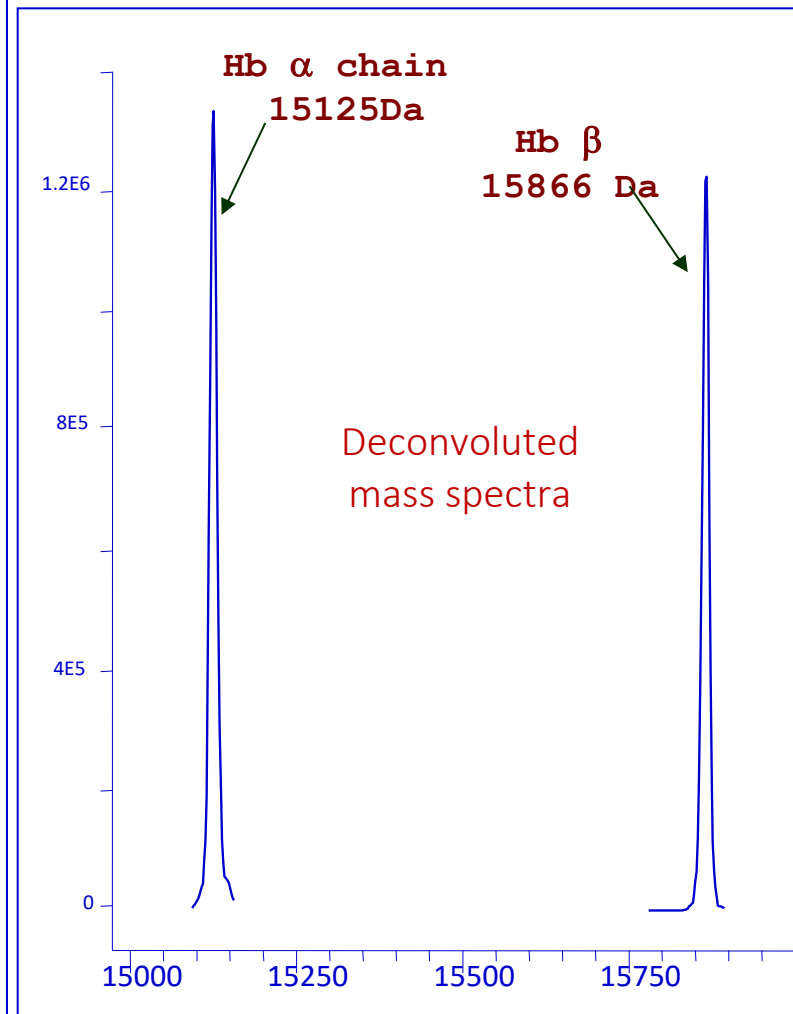
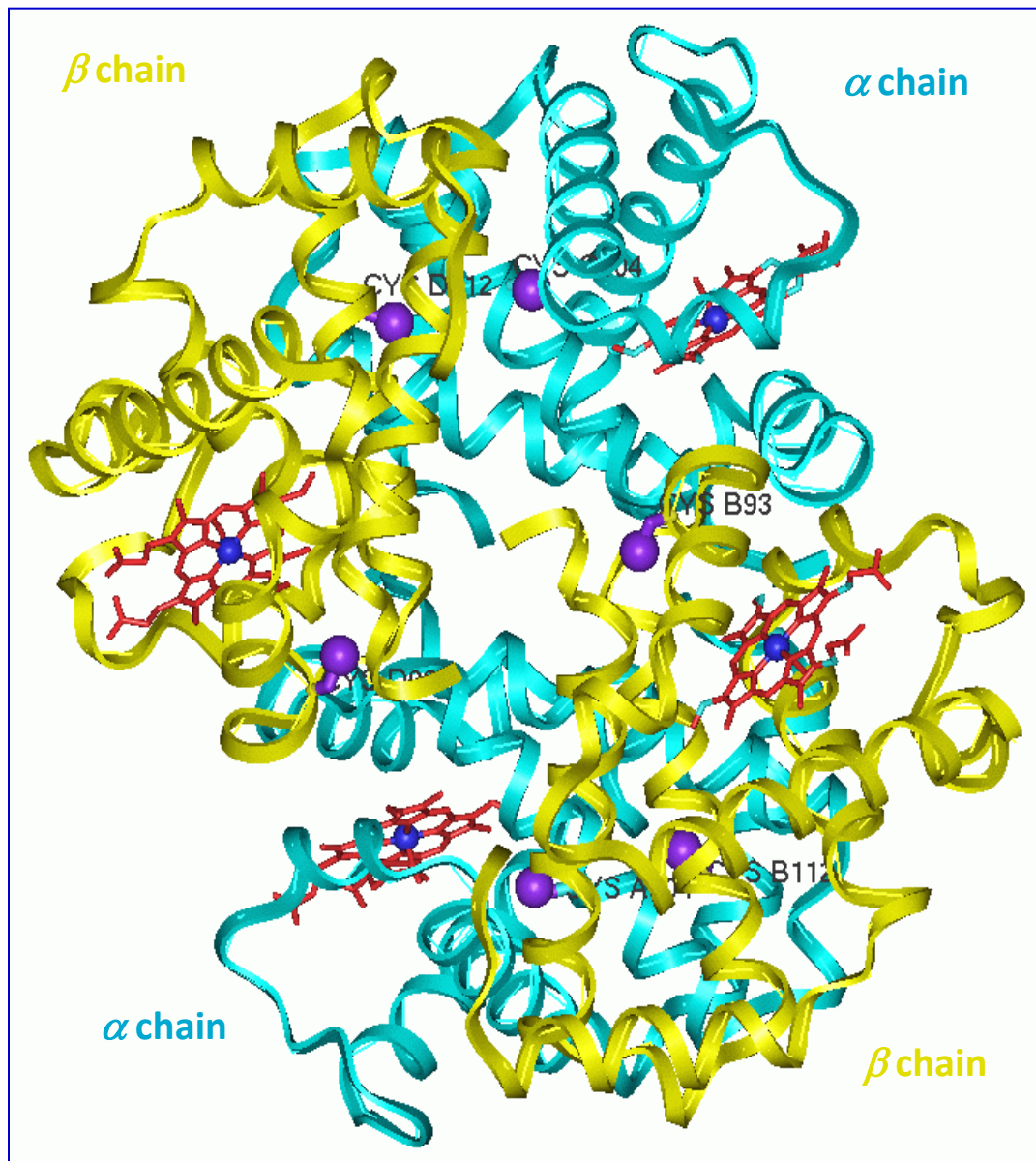
➤ Short α helices are connected by loop regions and packed together to produce a hydrophobic core, while hydrophilic residues on the surface make the protein soluble in water.

- Globin fold: A bundle of eight α helices, usually labeled as A-H, arrange in way that a pocket, an active site is formed, where in myoglobin and hemoglobin a heme group is bound. There are 7 residues in the shortest helix 'C' and 28 residues in the longest helix 'H' of myoglobin.

➤ In membrane bound proteins, the region inside the membranes are frequently α helices whose surfaces are covered by hydrophobic side chains suitable for the hydrophobic environment inside the membranes.



Normal Human Hemoglobin ($\alpha_2\beta_2$)



Structural hierarchy of protein molecule:

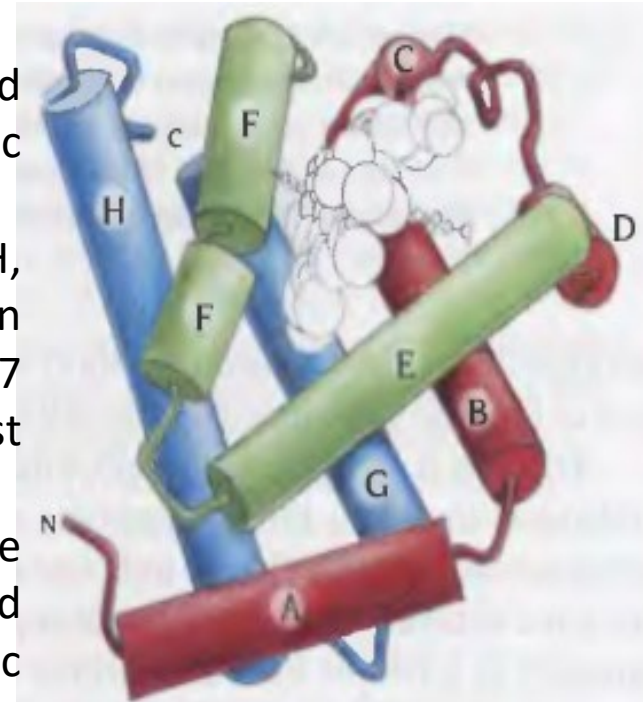
- Primary structure – Amino acid sequence along a polypeptide chain; Two proteins with significant similarity in the primary structure are said to be homologous
- Secondary structure – Secondary structure occurs mainly as α helices and β sheets. Secondary structure elements arrange themselves in simple motif
- Tertiary structure – Several motifs combine to form a compact globular structure called domain which is the fundamental unit of the tertiary structure. Tertiary structure is formed by the packing of domains.
- Quaternary structure – Protein molecules with only one chain are called monomeric and proteins with multiple polypeptide chains are called multimeric proteins. Quaternary structure of a protein consists of several polypeptide chains that are associated into a multimeric molecule.

α domain structure: Globular protein – Myoglobin, Hemoglobin

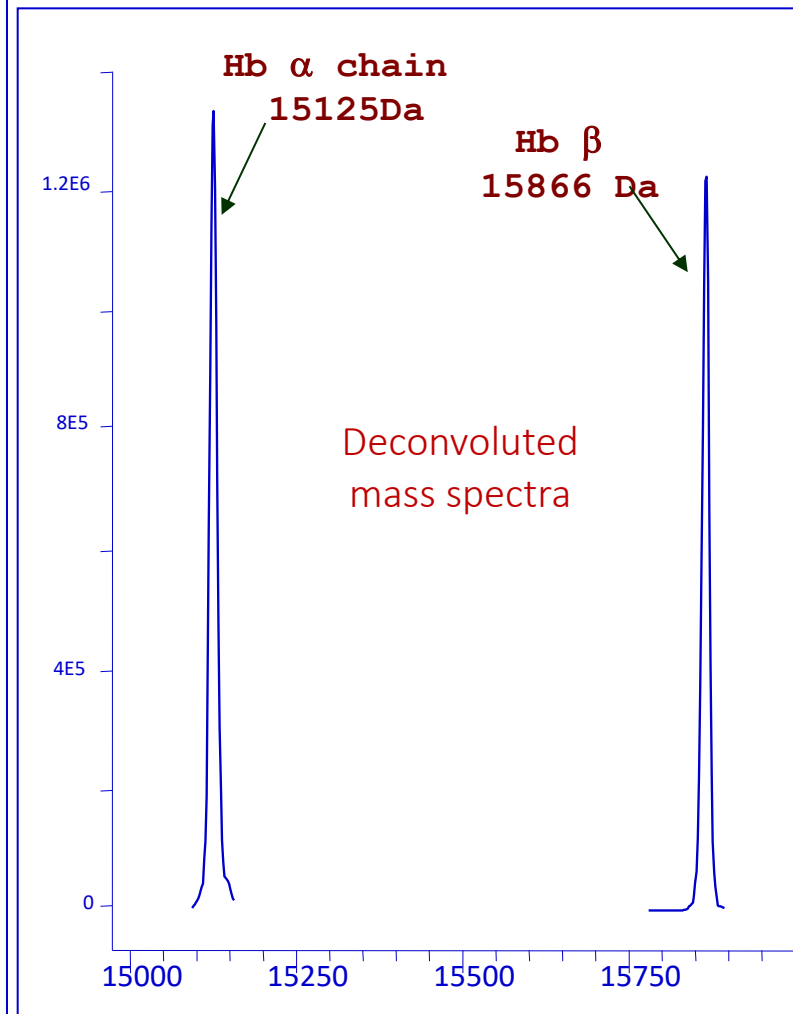
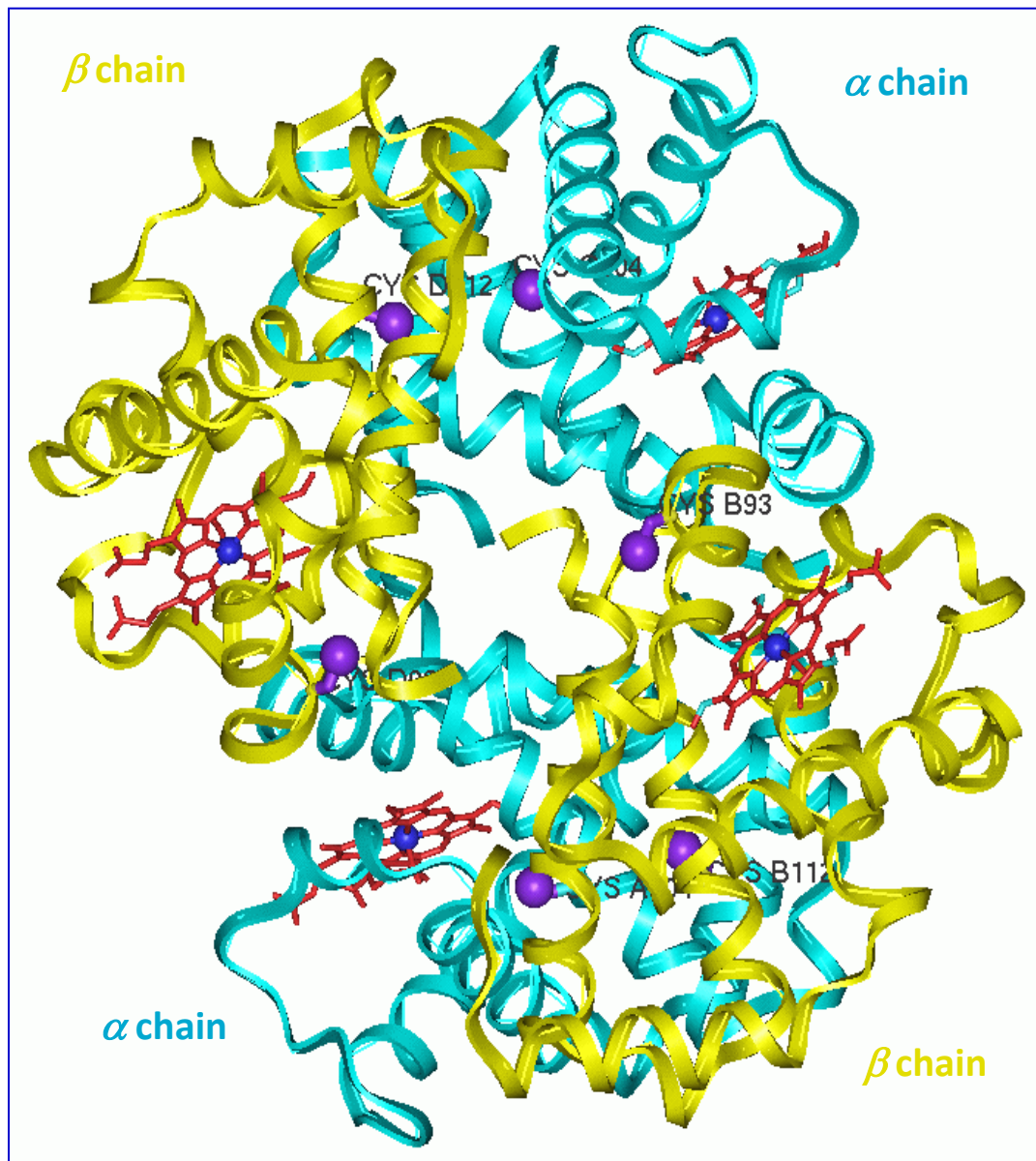
➤ Short α helices are connected by loop regions and packed together to produce a hydrophobic core, while hydrophilic residues on the surface make the protein soluble in water.

- Globin fold: A bundle of eight α helices, usually labeled as A-H, arrange in way that a pocket, an active site is formed, where in myoglobin and hemoglobin a heme group is bound. There are 7 residues in the shortest helix 'C' and 28 residues in the longest helix 'H' of myoglobin.

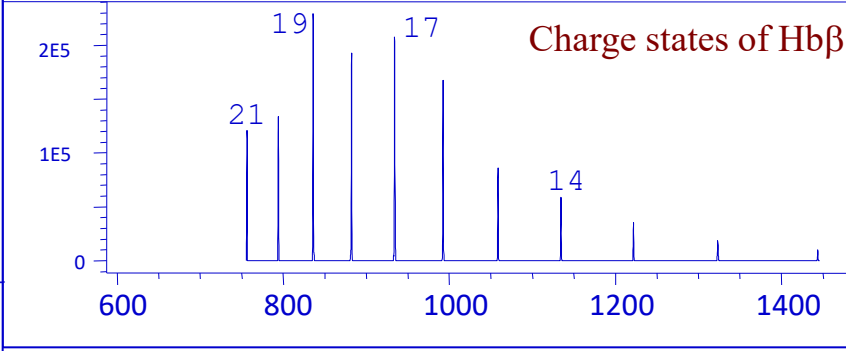
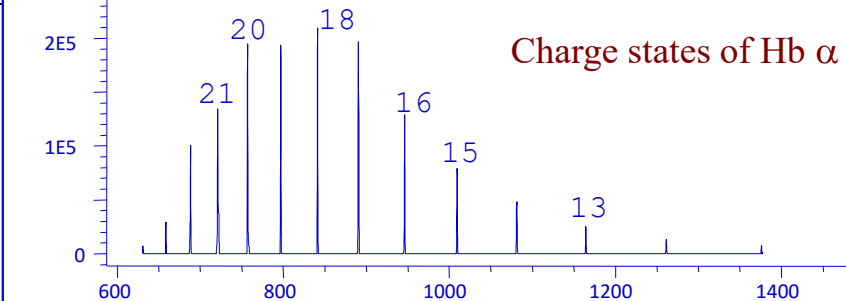
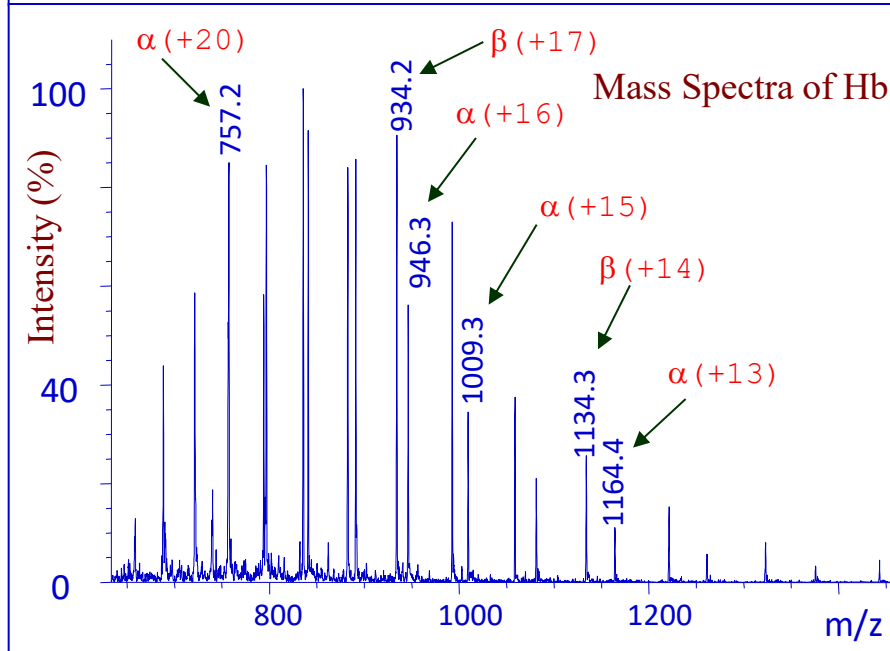
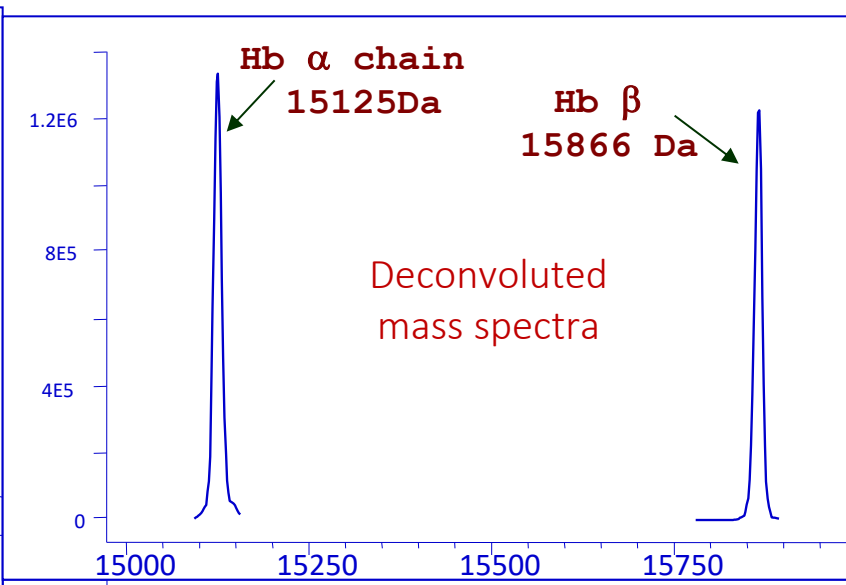
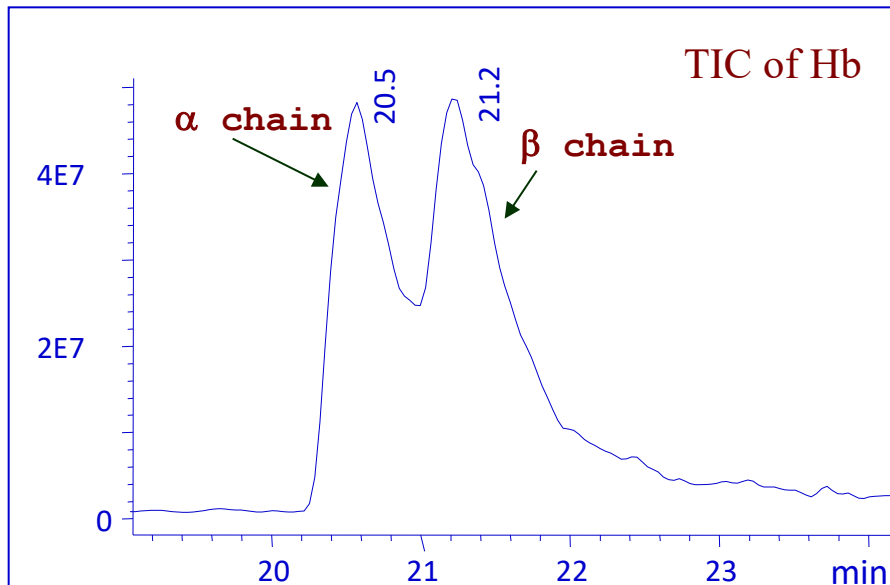
➤ In membrane bound proteins, the region inside the membranes are frequently α helices whose surfaces are covered by hydrophobic side chains suitable for the hydrophobic environment inside the membranes.



Normal Human Hemoglobin ($\alpha_2\beta_2$)

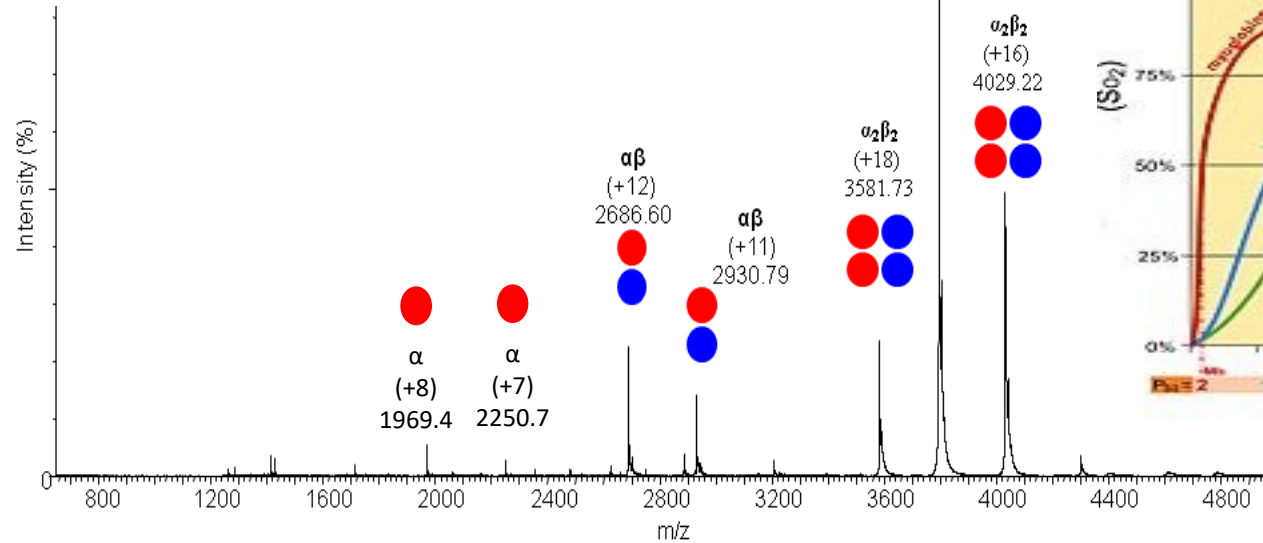
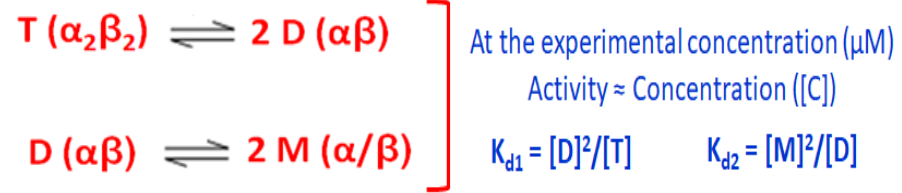


LC/ESI-MS of Normal Human Hemoglobin



Measurement of association/dissociation equilibrium constants:

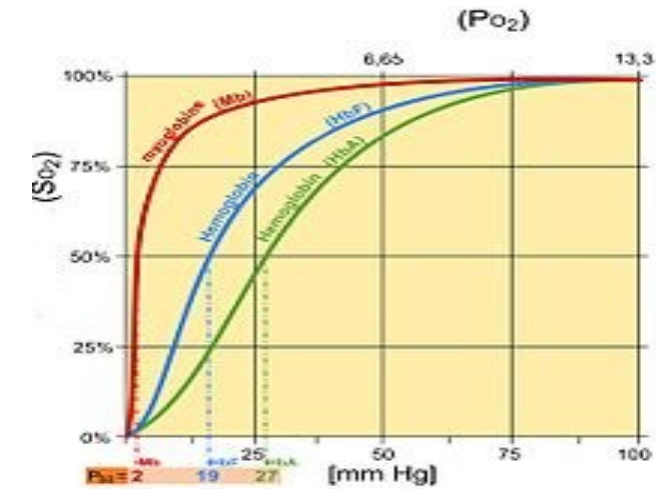
Binding affinities of globin subunits in tetrameric hemoglobin –



Human hemoglobin

● α globin chain

● β globin chain

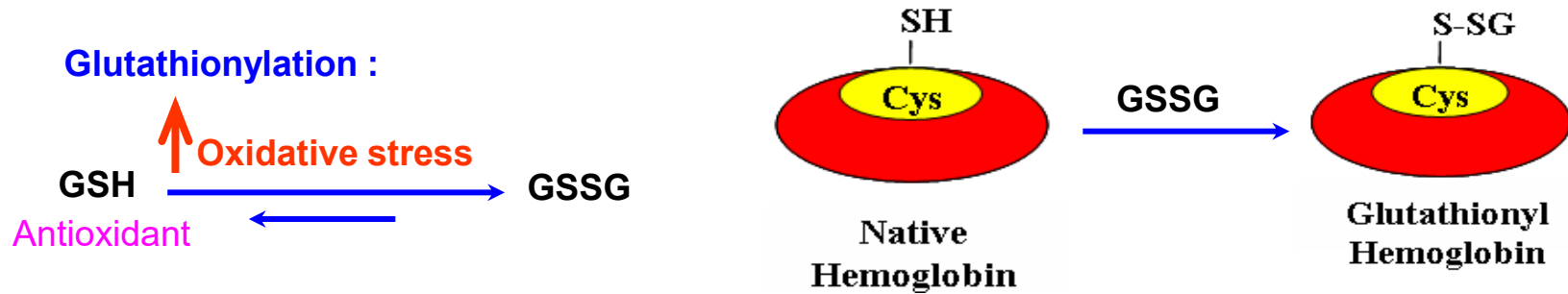


HbA	Replicates			Average \pm
	1	2	3	SD
K_{d1}	1.62	1.85	1.97	1.81 \pm 0.18
K_{d2}	0.11	0.15	0.17	0.14 \pm 0.03

(25 μM Human hemoglobin in 10mM NH_4Ac ; pH 7.4; Temp 310 K)

Journal of Structural Biology, 2019, <https://doi.org/10.1016/j.jsb.2019.09.003>

Posttranslation modification of Human Hemoglobin



Glutathionyl Hemoglobin as a marker of oxidative stress (quantified using *LC/ESI-MS platform*):

Chronic renal failure *Clin Biochem*, 2007, 40, 986-94

Iron Deficiency Anemia *Acta Haematologica*, 2012;127, 26-30

Major Depressive Disorder *Indian Journal of Medical Research*, 2019, 149, 497-502

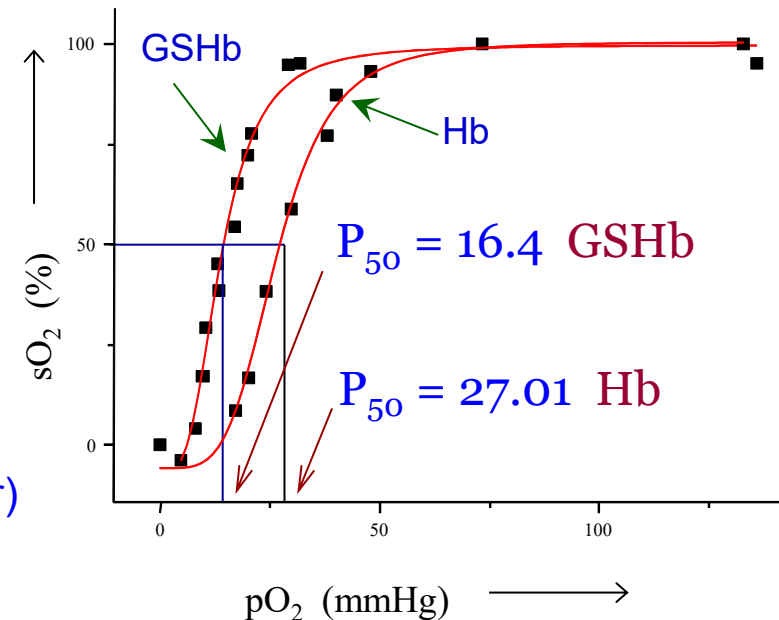
Multiple Sclerosis, Diabetes, Cardio Vascular Disease

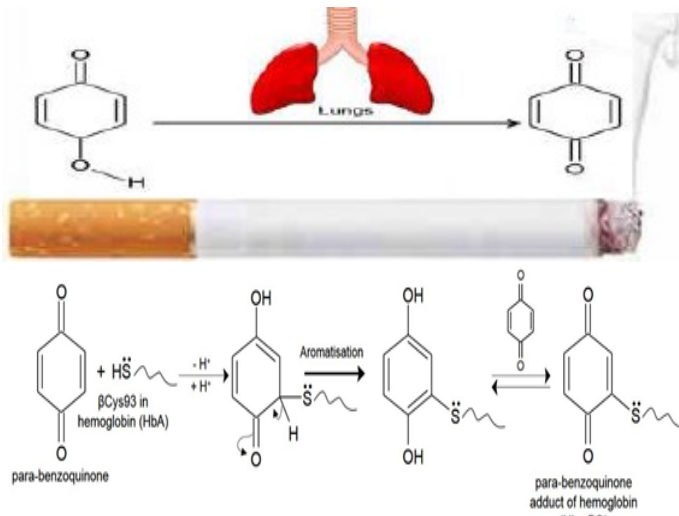
- α globin chain - Cys104
- β globin chain - β Cys93 and β Cys112
- Glutathionylation occurs at β Cys93

Clin Biochem, 2007, 40, 986-94

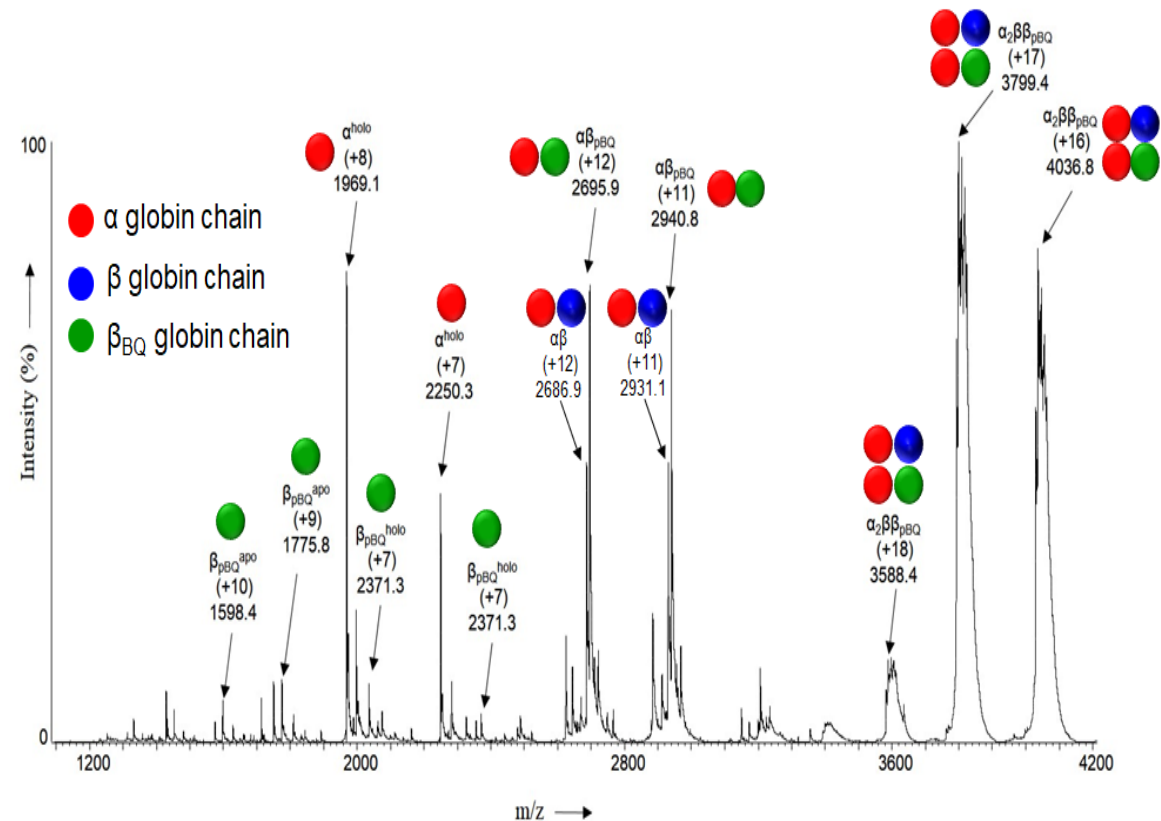
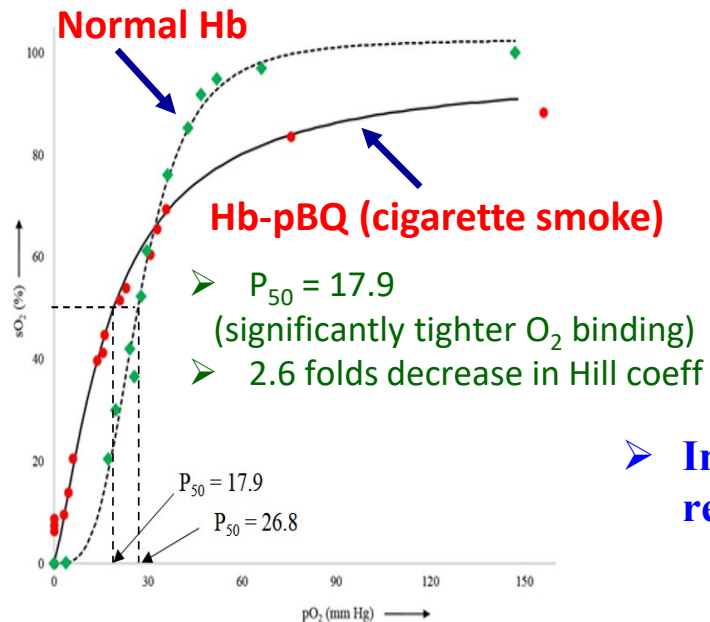
$P_{50} = 27.01$ (Normal Human Hemoglobin)

$P_{50} = 16.4$ (Glutathionyl Hb binds to O_2 six fold tighter)





- pBQ combines with β Cys93
 - HbpBQ increases with no. of cigarette smoked per day
- Toxicology Report, 2016, 3, 295



Molecule	K_{d1}	K_{d2}
HbA	1.81 ± 0.18	0.14 ± 0.03
Hb-pBQ	9.03 ± 1.1	9.0 ± 1.9

- In smoker's blood, conjugation of pBQ with hemoglobin results in reduced O_2 supply to the tissues \Rightarrow **HYPOXIA !!!**
- J Am Soc Mass Spectrom, 2018, 29, 2048-2058*