PROTEINS:
THREE-DIMENSIONAL STRUCTURE AND FUNCTION

Course: Biochemistry I (BIOC 230)
Textbook:
Principles of Biochemistry, 5th Ed., by L. A. Moran and others. 2014, Pearson. . Chapter 4

Terminology

- **Protein:** a chain of amino acids joined by peptide bonds in a specific sequence
- **Conformation** – spatial arrangement of atoms in a protein, that depends on the rotation of a bond or bonds (?)
- **Native conformation** – conformation of functional protein

From DNA to Protein

- In some species, the size and sequence of every polypeptide can be determined from the sequence of the **genome**
- **Genomics:** the study of the structure of whole genome
- In E. coli, there are 4000 different polypeptides with an average size of about 300 aa
- In fruit fly (Drosophila melanogaster) there are 14,000 different polypeptides
- Humans and mammals there are 30,000 different polypeptides
- **Proteomics**: the science that studies large sets of proteins,
- **Proteome**: all proteins produced by a cell

Methods for determining protein structure

- Primary structure or amino acid sequence of a polypeptide is determined using chemical methods such as Edman degradation or indirectly from the sequence of a gene
- Three-dimensional structure of protein:
  - X-ray crystallography
  - Nuclear magnetic resonance (NMR)

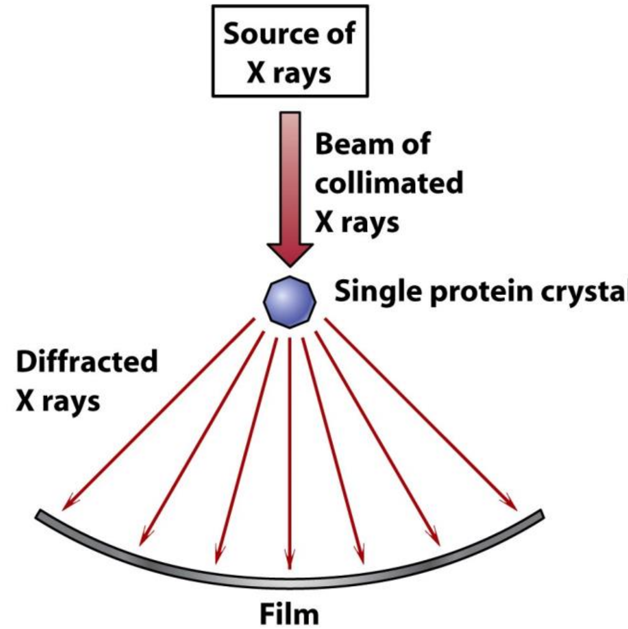


Figure 4-2a Principles of Biochemistry, 4/e

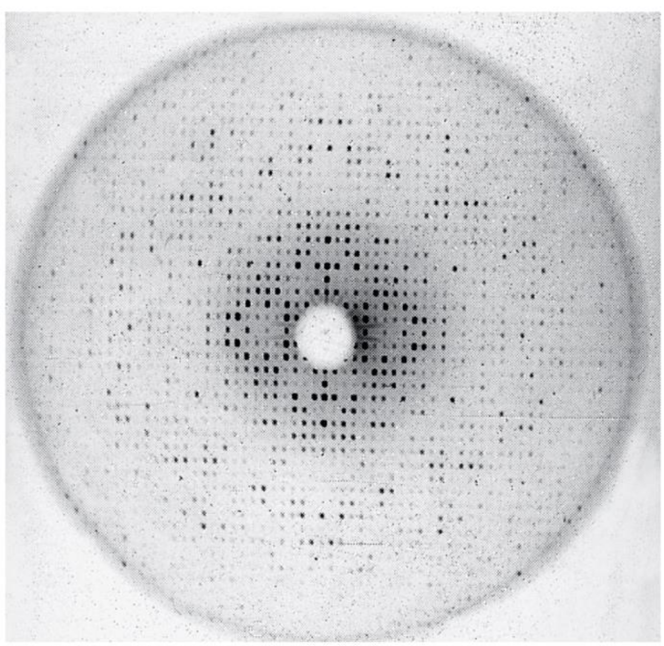
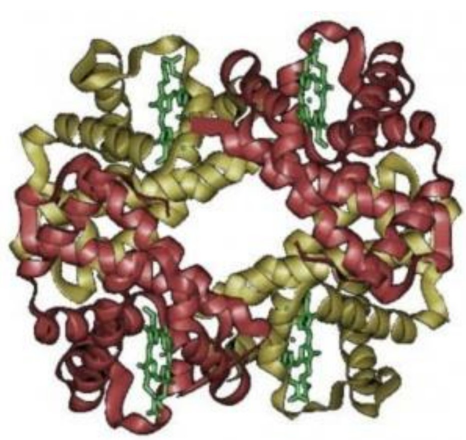
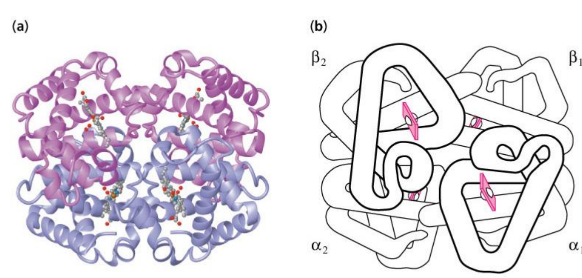


Figure 4-2b Principles of Biochemistry, 4/e



Protein *Classification*

- **Simple** – composed only of amino acid residues
- **Conjugated** – contain prosthetic groups
  (metal ions, co-factors, lipids, carbohydrates)
  Example: Hemoglobin – conjugated to Heme



Protein Classification

- One polypeptide chain: monomeric protein
- More than one - multimeric protein
- Homomultimer - one kind of chain
- Heteromultimer - two or more different chains
- E.g., Hemoglobin is a heterotetramer. It has two alpha chains and two beta chains: $\alpha_2\beta_2$

Protein Classification

Fibrous

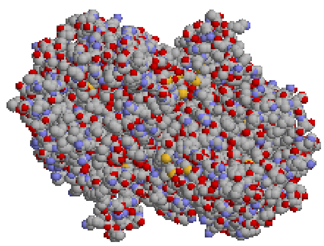
1) Polypeptides arranged in long strands or sheets
2) Water insoluble (lots of hydrophobic aa's)
3) Strong but flexible
4) Structural (e.G., Keratin, collagen)

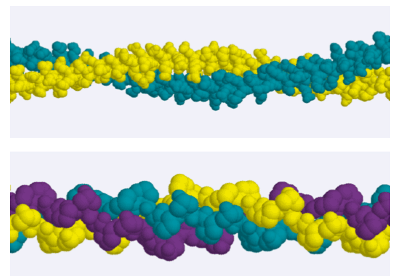
Globular

1) Polypeptide chains folded into spherical or globular form
2) Water soluble
3) Contain several types of secondary structure
4) Diverse functions (enzymes, regulatory proteins)

catalase

keratin



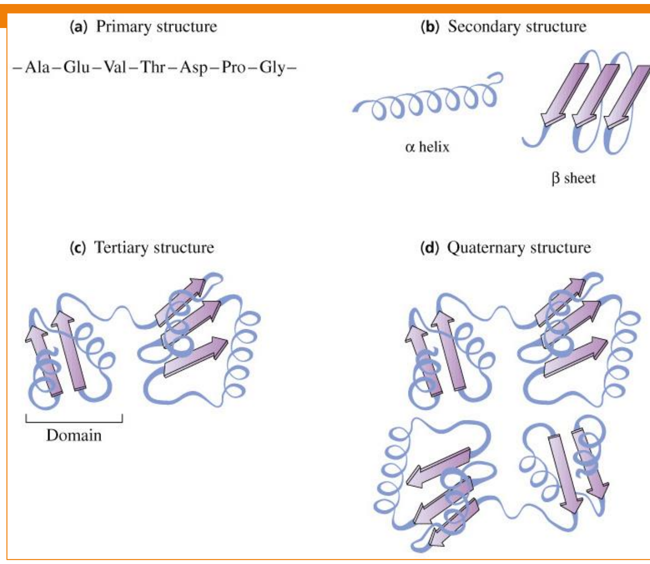


collagen

Protein Function

- Catalysis – enzymes
- Structural – keratin
- Transport – hemoglobin
- Trans-membrane transport – Na^+/K^+ ATPases
- Toxins – rattle snake venom, ricin
- Contractile function – actin, myosin
- Hormones – insulin
- Storage Proteins – seeds and eggs
- Defensive proteins – antibodies

4 Levels of Protein Structure



Non-covalent forces important in determining protein structure

- van der Waals: 0.4 - 4 kJ/mol
- hydrogen bonds: 12-30 kJ/mol
- ionic bonds: 20 kJ/mol
- hydrophobic interactions: <40 kJ/mol

Questions!

1. Which of the following level of protein structure determines all other levels?
   a. Primary structure
   b. Secondary structure
   c. Tertiary structure
   d. Quaternary structure
2. The peptide bond that links amino acids in a polypeptide is formed between
   a. Alpha-carboxyl group of first amino acid and alpha-amino group of the next amino acid
   b. Alpha-carboxyl group of first amino acid and R-group of the next amino acid

1° Structure Determines 2°, 3°, 4° Structure

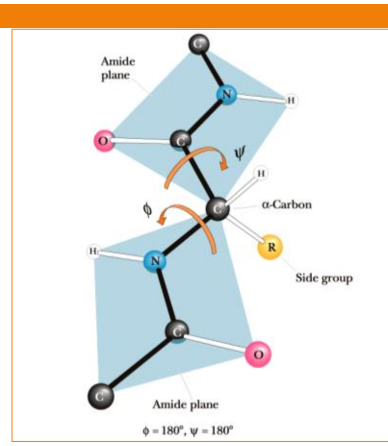
- Sickle Cell Anemia – single amino acid change in hemoglobin related to disease
- Osteoarthritis – single amino acid change in collagen protein causes joint damage

Classes of 2° Structure

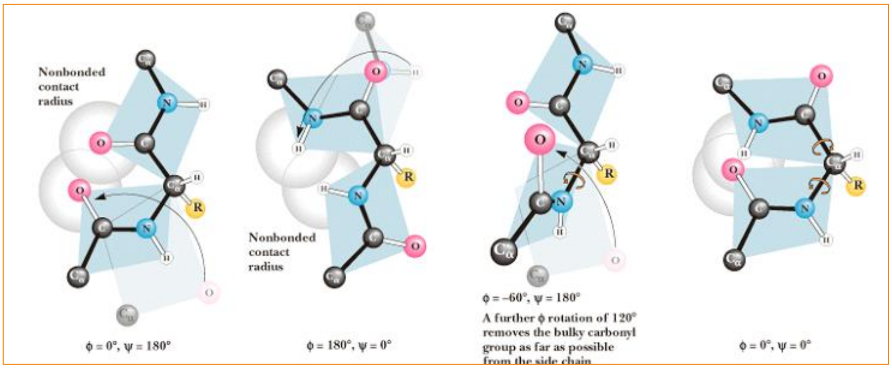
- Alpha helix
- B-sheet
- Loops and turns

2° Structure Related to Peptide Backbone

- Double bond nature of peptide bond cause planar geometry
- Free rotation at N - αC and αC-carbonyl C bonds
- Angle about the C(alpha)-N bond is denoted phi (ϕ)
- Angle about the C(alpha)-C bond is denoted psi (ψ)
- The entire path of the peptide backbone is known if all phi and psi angles are specified



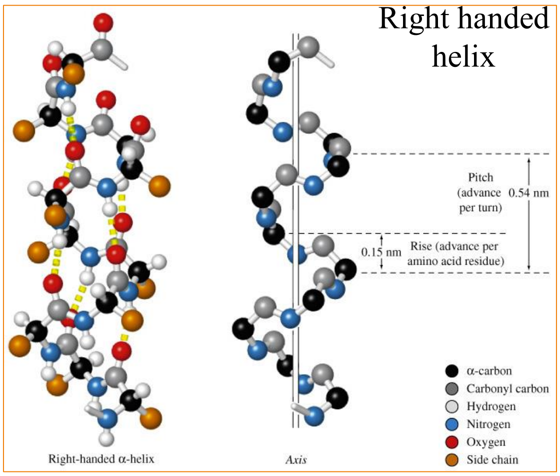
Not all ϕ/ψ angles are possible



Alpha-Helix

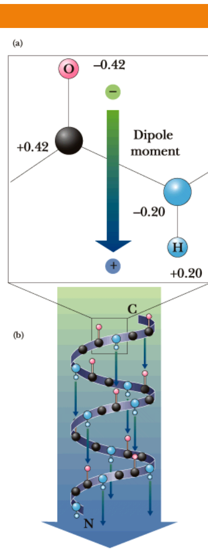
- First proposed by Linus Pauling and Robert Corey in 1951
- Identified in keratin by Max Perutz
- A ubiquitous component of proteins
- Stabilized by H-bonds

Alpha-Helix



- Residues per turn: 3.6
- Rise per residue: 1.5 Angstroms
- Rise per turn (pitch): 3.6 x 1.5A = 5.4 Angstroms
- amino hydrogen H-bonds with carbonyl oxygen located 4 AA's away forms 13 atom loop

Alpha-Helix



All H-bonds in the alpha-helix are oriented in the same direction giving the helix a dipole with the N-terminus being positive and the C-terminus being negative.

Helix dipole

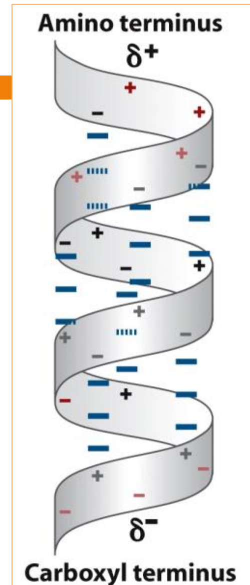
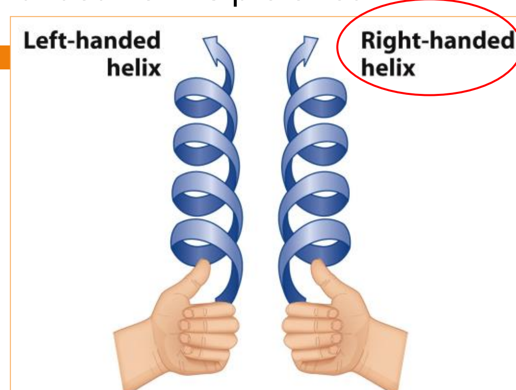


Figure 4-5
Lehninger Principles of Biochemistry, Fifth Edition

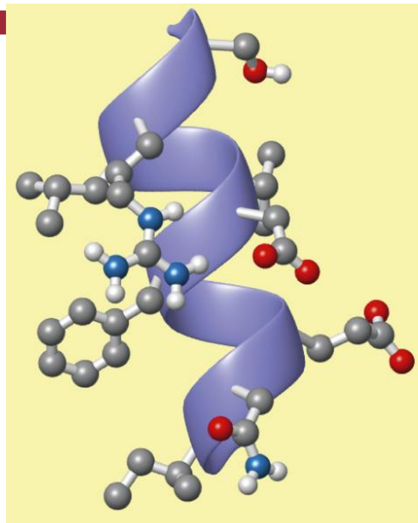
α-helix: Why right-handed helix is preferred?



Box 4-1
Lehninger Principles of Biochemistry, Fifth Edition

In principle, naturally occurring L-amino acids can form either right or left-handed α-helix, but extended left-handed α-helices are theoretically less stable & have not been observed in proteins

Alpha-Helix

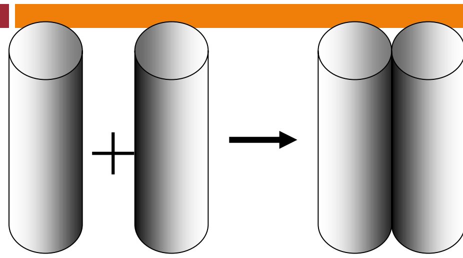


•Side chain groups point outwards from the helix

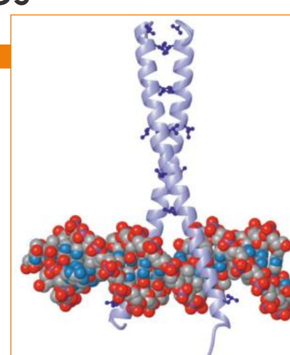
•AA's with bulky side chains less common in alpha-helix

•Glycine and proline destabilizes alpha-helix

Amphipathic Alpha-Helices



Amphipathic helices



Leucine zipper of yeast

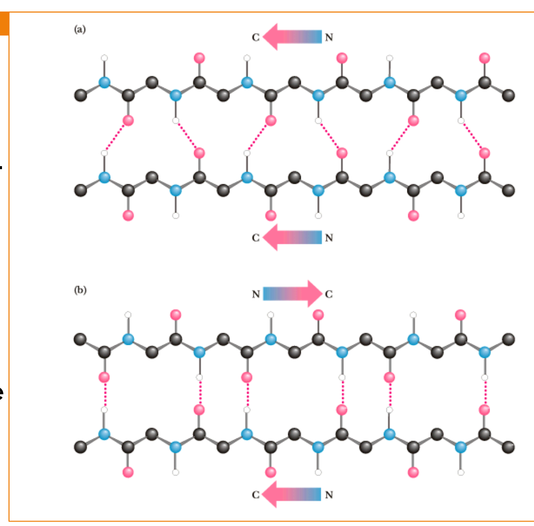
One side of the helix (dark) has mostly hydrophobic AA's
Two amphipathic helices can associate through hydrophobic interactions. Such helices often occur on protein surface

Beta-Strands and Beta-Sheets

- Also first postulated by Pauling and Corey, 1951
- Strands may be parallel or antiparallel
- Rise per residue:
  - 3.47 Angstroms for antiparallel strands
  - 3.25 Angstroms for parallel strands
  - Each strand of a beta sheet may be pictured as a helix with two residues per turn

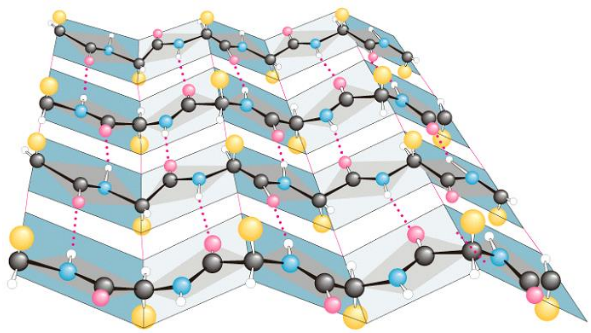
Beta-Sheets

- Beta-sheets formed from multiple side-by-side beta-strands.
- Can be in parallel or anti-parallel configuration
- Anti-parallel beta-sheets more stable
- Typically B-strands twist slightly in a right hand direction or twist clockwise



Beta-Sheets

- Side chains point alternately above and below the plane of the beta-sheet
- A typical B-sheet contains 2 to 15 beta-strands/beta-sheet
- Each strand is made of ~ 6 amino acids on average



Loops and turns

Loops

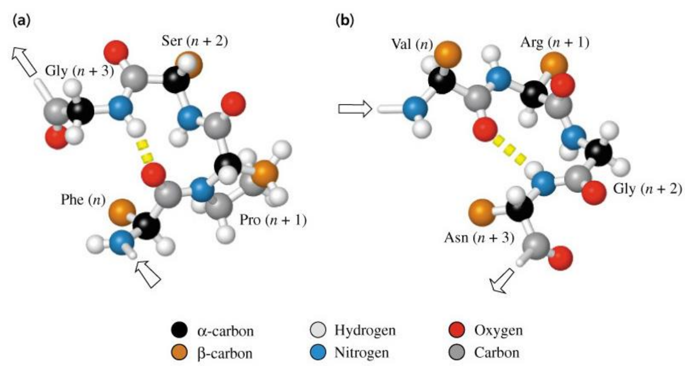
- Loops usually contain hydrophillic residues.
- Found on surfaces of proteins
- Connect alpha-helices and beta-sheets
- ~10% of residues are found in loops

Turns

- Loops with < 5 AA’s are called turns
- Beta-turns are common

Beta-turns

- allows the peptide chain to reverse direction
- carbonyl C of one residue is H-bonded to the amide proton of a residue three residues away
- proline and glycine are prevalent in beta turns



3° & 4° structure and protein folding

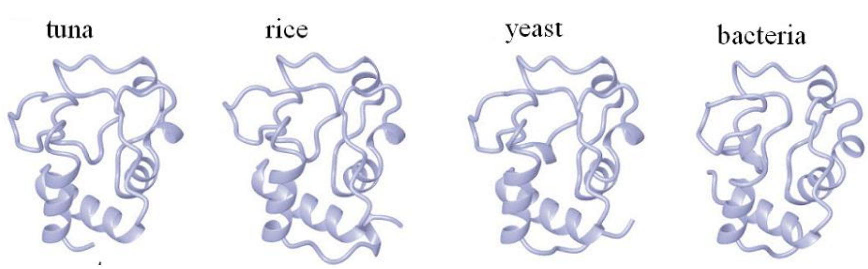
3° Structure

- Third level of protein organization
- Folding of polypeptide chain causes 2° structures to interact.
- AAs that are far apart in 1ry structure are brought together, permitting interactions among their side chains.
- 3° structure is stabilized primarily by non-covalent interactions (mostly hydrophobic interactions) between side chains of aa residues.
- Disulfide bridges are also elements of 3ry structure
- Formation of motifs and domains

Proteins with similar 1° structure also have similar 3° structure

```
tuna     1  GDVAKGKKTFVQKCAQCHTVENGGKHKVGPNLWGLFGRKTGQAEGYSYTDANKSKGIVWN
yeast    1  GSAKKGATLFKTRCLQCHTVEKGGPHKVGPNLHGIFGRHSGQAEGYSYTDANIKKNVWDE
rice     1  GNPKAGEKIFKTKCAQCHTVDKGAGHKQGPNLNGLFGRQSGTTPGYSYSTANKMAVIWEE

tuna    61  ETLMEYLENPKKYIPGTKMIFAGIKKKGERQDLVAYLKSATS
yeast   61  NNMSEYLTNPKKYIPGTKMAFGGLKKEKDRNDLITYLKKACE
rice    61  NTLYDYLLNPKKYIPGTKMVFPGLKKPQERADLISYLKEATS
```



Super-secondary structures

- Super-secondary structures or motifs: are recognizable combinations of α–helices, β–strands and loops that appear in a different number of proteins.
- Sometimes, motifs are associated with a particular function, but similar motifs may have different functions.
-

Common Motifs

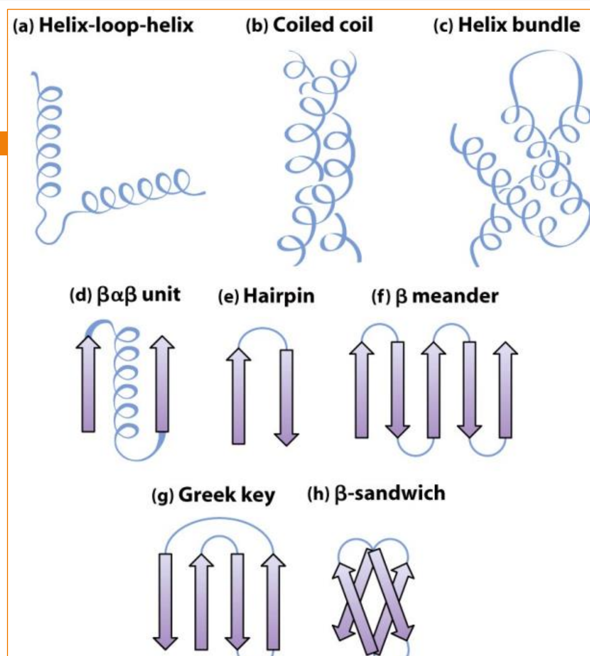


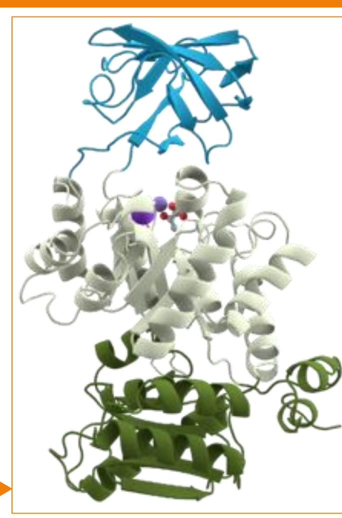
Figure 4-19 Principles of Biochemistry, 4/e

Motifs

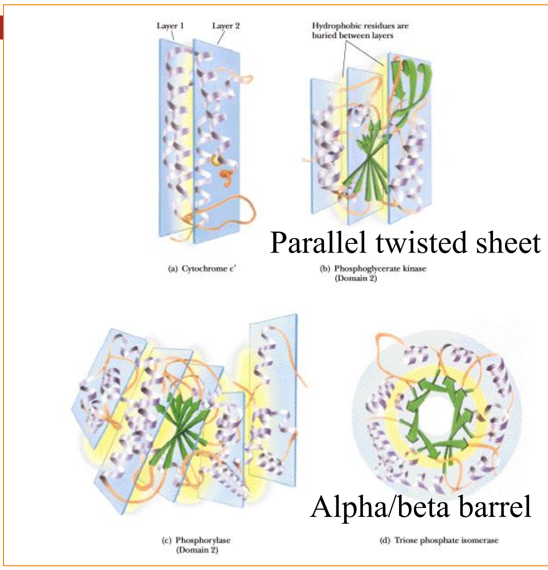
- Helix-loop-helix: occurs in calcium-binding proteins. Glu and Asn forms part of the loop and Ca+2 binding site. In some DNA binding proteins it is called, Helix-turn-helix because the loop forms a reverse turn
- Coiled –coil motif: two amphipathic a-helices that interact through hydrophobic edges, as in leucine zipper
- "βαβ":
- B-meander
- Hairpin
- Greek key
- B-sandwich

Domains

- **Domains:** are independently folded compact units
- Domains may consist of combinations of motifs
- Individual domains have specific function
- Size of domain vary from 25 to 30 aa residues to more than 300 aa residues
- Example: Pyruvate dehydrogenase has 3 domains



Motifs Combine to form Domains

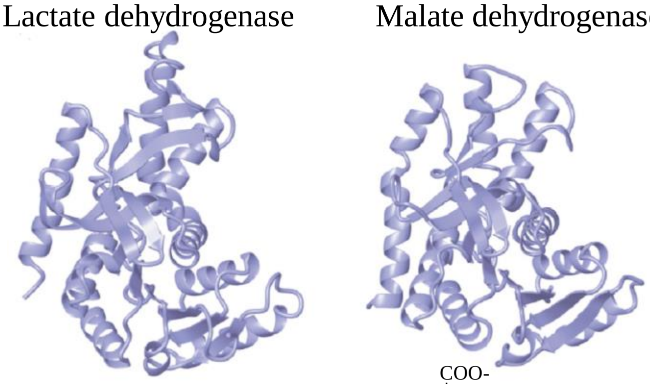


- Domains are independent folding units in a 3° structure of a protein
- Individual domains have specific function

- Hydrophobic interactions are the major driving force in folding domains

Protein family members share common domain structures

Lactate dehydrogenase | Malate dehydrogenase



$$O{=}C(COO^-)(CH_3) + NADH \rightarrow HO{-}CH(COO^-)(CH_3) + NAD^+$$

$$O{=}C(COO^-)(H_2C{-}COO^-) + NADH \rightarrow HO{-}CH(COO^-)(CH_2{-}COO^-) + NAD^+$$

4º Structure

- Quaternary structure describes the organization of subunits in a protein with multiple subunits (oligomeric protein)
- Can have homo-multimers or hetero-multimers
- Subunits are held together by weak interactions, primarily by hydrophobic interactions

4º Structure

- Determine molecular weight of native protein by gel permeation chromatography
- Determine molecular weight of individual subunits by SDS-PAGE
- Can use the information to determine subunit composition

If…….
Native protein – 160,000 daltons and
α-Subunit – 50,000 daltons, β-Subunit – 30,000 daltons

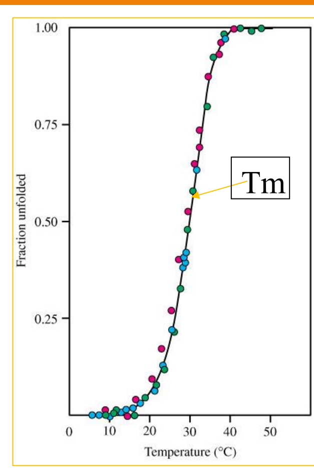
Then……
Protein can have $\alpha_2\beta_2$ structure

4° Structure

- Subunits held together by non-covalent interactions
- Oligomeric protein is more stable than disassociated subunits
- Active site often made up of AA residues from different subunits
- 4° and 3° structure is often affected by ligand (substrate or inhibitor) binding. Important in enzyme regulation

Protein denaturation

- Denaturation – disruption of native conformation
- Heat commonly used to denature proteins
- Tm = temperature where 50% folded/50% unfolded. You can calculate Tm from curve?
- Under physiological conditions, most proteins are stable up to 50-60°C
- Typical Tm = 40-60°C
- Tm depends on pH and ionic strength
- Tm for thermophiles >100°C (E.g., Taq DNA polymerase)
- Chemical denaturants Chaotrophic agents = Urea, guanidinium salts, KCN detergents = SDS



Protein Folding

- Ribonuclease A (RNase A) will refold to native structure spontaneously (1 minute)
- $>10^{50}$ possible conformations
- If 10^{-13} sec per conformation would take 10^{30} years to sample enough to determine structure
- How do proteins fold so quickly?

Protein Folding

- Structures of globular proteins are not static
- Proteins "breathing" between different conformations
- Proteins fold towards lowest energy conformation
- Multiple paths to lowest energy form
- All folding paths funnel towards lowest energy form
- Local low energy minimum can slow progress towards lowest energy form

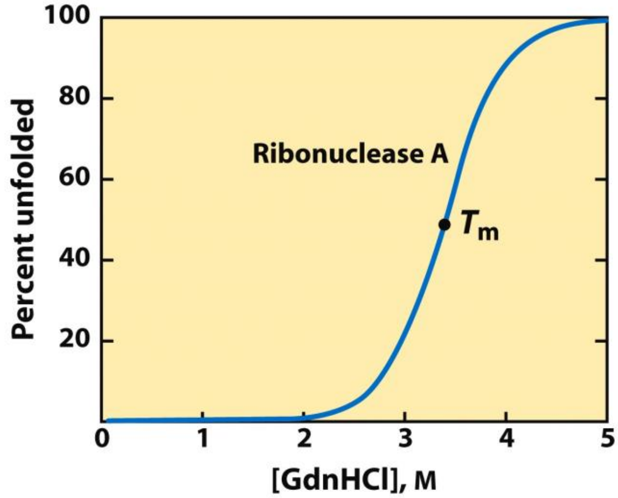


Figure 4-25b
Lehninger Principles of Biochemistry, Fifth Edition

Denaturation by Guanidine hypochloride

Protein **renaturation**

- Protein renaturation: certain globular proteins when denatured can regain their native structure and biological activity
- Example: Ribonuclease experiment *(see next figure)*

Renaturation of unfolded, denatured Ribonuclease

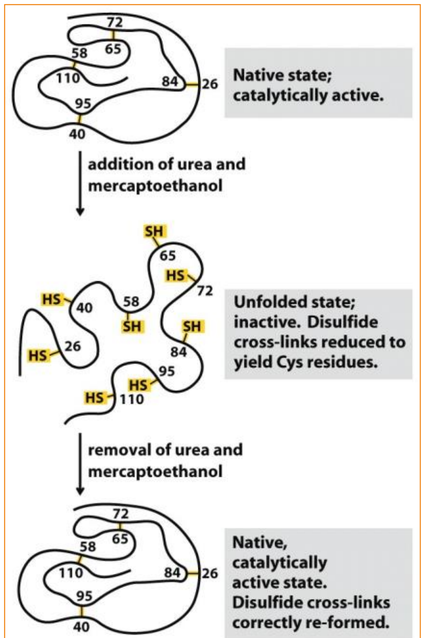


Figure 4-26
Lehninger Principles of Biochemistry, Fifth Edition

Ribonuclease Denaturation/Renaturation experiment

- Conclusions from the Ribonuclease experiment
- ➔ AA sequence contains all information required to fold the p.p chain into its native, 3D structure.

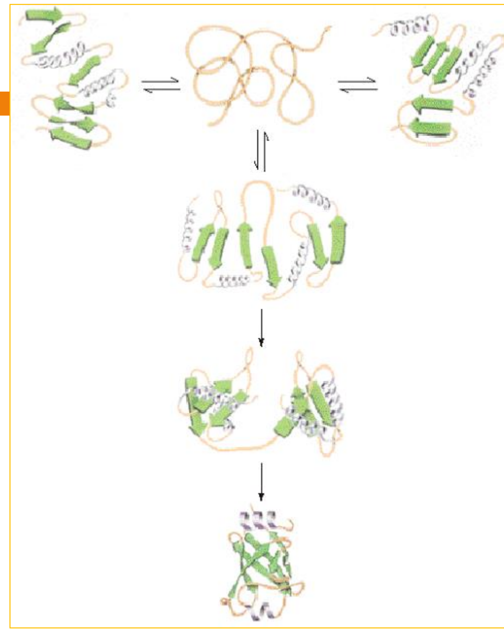
- This is true for a minority of proteins, which is small and inherently stable.
- Even though all proteins have the potential to fold into their native structure, many require some assistance

Pathway of Protein Folding

1) Nucleation of folding - Rapid and reversible formation of local 2° structures form

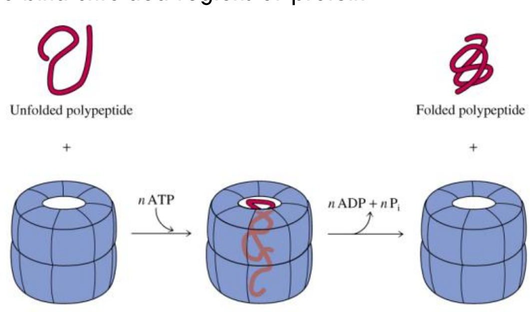
2) Formation of domains (Molten Globular intermediates) through aggregation of local 2° structures

3) Domain conformations adjust to form native protein



Chaperonins

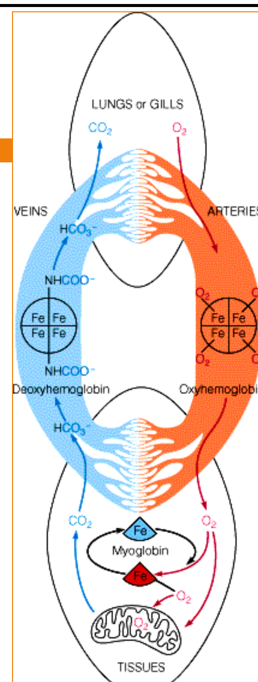
- Protein complexes that promote protein folding
- Chaperonins don't determine native structure
- Prevent misfolding and aggregation of protein
- Sequesters unfolded protein from other proteins
- Require ATP for protein binding, after ATP hydrolysis native protein released
- Thought to bind unfolded regions of protein



Protein 3-D Structure/ Function

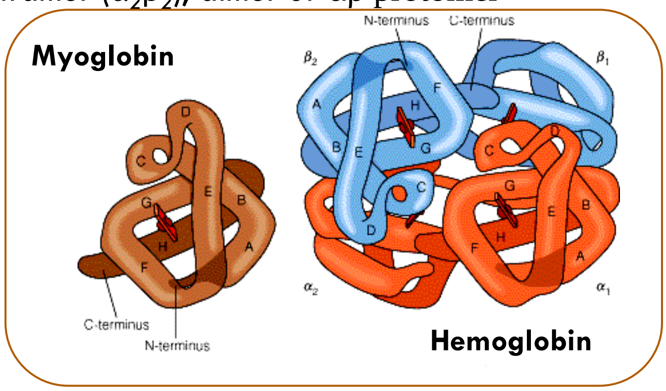
Myoglobin/ Hemoglobin

- First protein structures determined
- Oxygen carriers
- Reversible binding of O2 is called Oxygenation
- Hemoglobin transport O_2 from lungs to tissues
- Myoglobin: O_2 storage protein, supply O2 to muscle tissues in reptiles, birds and mammals



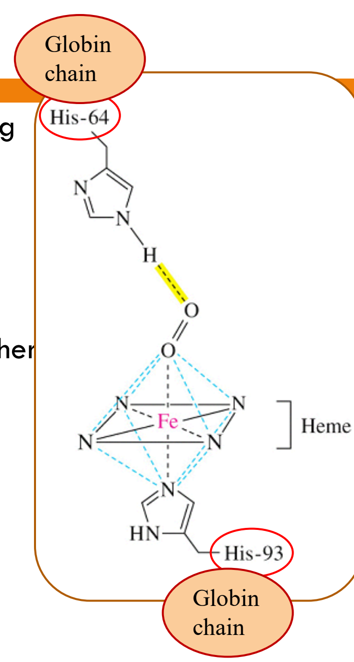
Mb and Hb subunits structurally similar

- 8 alpha-helices
- Contain heme group
- Mb monomeric protein
- Hb heterotetramer ($\alpha_2\beta_2$), dimer of $\alpha\beta$ protomer



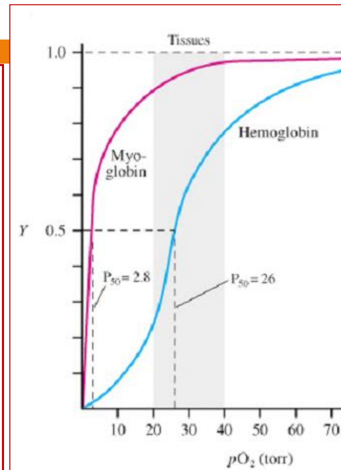
Heme group

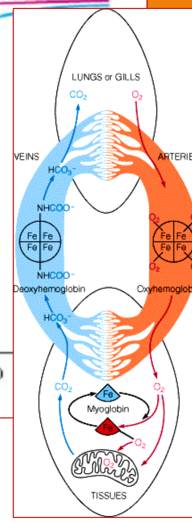
- Heme = Fe^{++} bound to tertapyrrole ring (protoporphyrin IX complex)
- Heme non-covalently bound to globin proteins through His residue
- O_2 binds non-covalently to heme Fe^{++}, stabilized through H-bonding with another His residue
- Heme group in hydrophobic crevice of globin protein



Oxygen Binding Curves

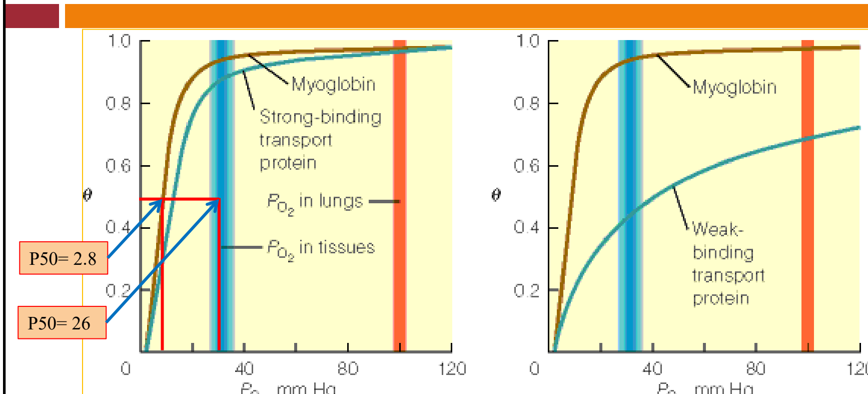
- Mb has hyberbolic O_2 binding curve
- Mb binds O_2 tightly. Releases at very low $_pO_2$
- Hb has sigmoidal O_2 binding curve
- Hb high affinity for O_2 at high $_pO_2$ (lungs)
- Hb low affinity for O_2 at low $_pO_2$ (tissues)





Partial pressure at P50 is a measure of the affinity of the protein for O2

Oxygen Binding Curve



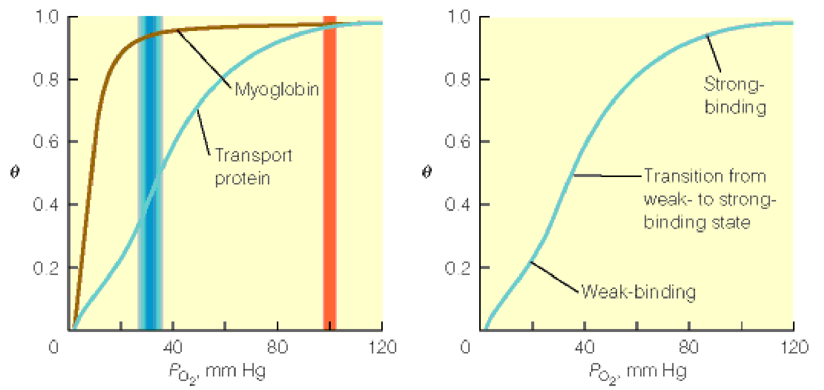
(a) Transport protein efficient in binding but inefficient in unloading (hyperbolic binding curves)

(b) Transport protein efficient in unloading but inefficient in binding (hyperbolic binding curves)

- Mb is half saturated (P_{50}) at PO_2 of 2.8 torr, while that of Hb is 26 torr.

1 Atmosphere =760 torr. PO2 in lungs is ~100 torr and in tissues is ~20-40 torr.

Oxygen Binding Curve

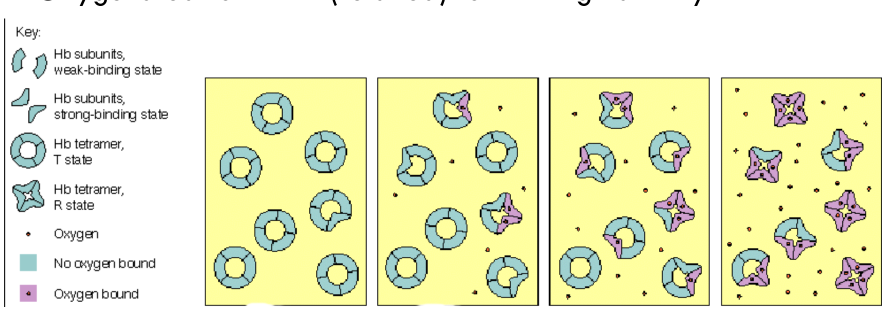


(c) Transport protein efficient in both binding and unloading, because it has a sigmoidal binding curve.

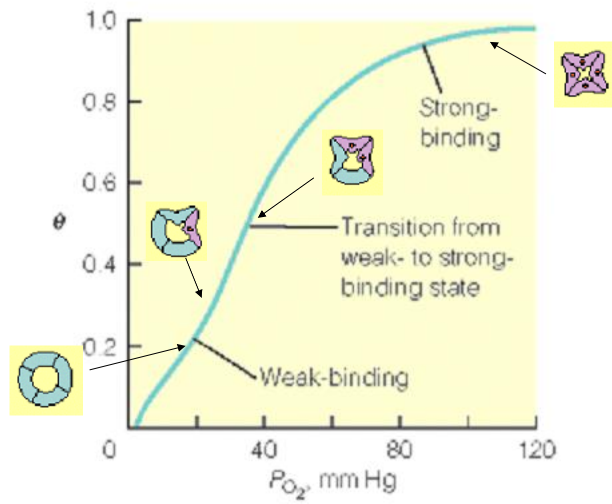
(d) Switch from weak– to strong–binding state explains the sigmoidal curve

O_2 Binding to Hb shows positive cooperativity

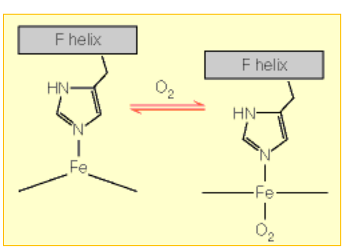
- Hb binds four O_2 molecules
- O_2 affinity increases as each O_2 molecule binds
- Increased affinity due to conformation change
- Deoxygenated form = T (tense) form = low affinity
- Oxygenated form = R (relaxed) form = high affinity



O_2 Binding to Hb shows positive co-operativity



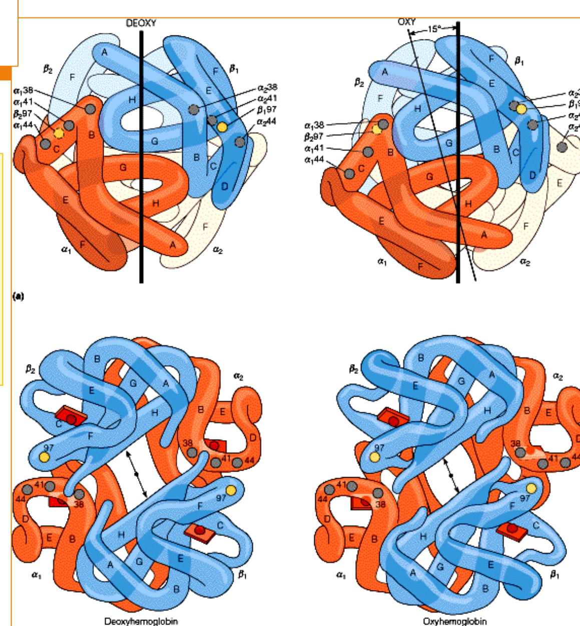
O_2 Binding induces conformation change



Heme moves 0.34 nm

Exposing crystal of deoxy-form to air cause crystal to crack

T-conformation R-conformation



Allosteric Interactions

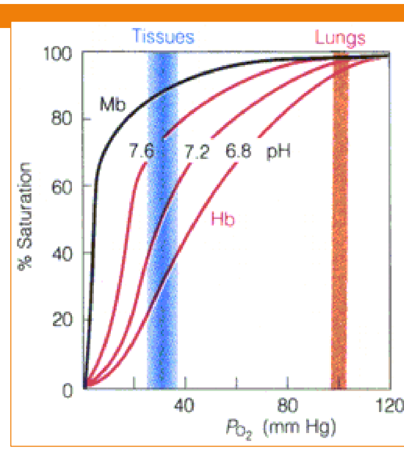
- Allosteric interaction occur when specific molecules bind a protein and modulates activity
- Allosteric modulators or allosteric effectors
- Bind reversibly to site separate from functional binding or active site
- Modulation of activity occurs through change in protein conformation
- 2,3 bisphosphoglycerate (BPG), CO_2 and protons are allosteric effectors of Hb binding of O_2
- The binding of 2,3BPG to Hb raises its P50 to 26 torr, much higher than the P50 for O2 binding to purified Hb in aqueous solution (12 torr).

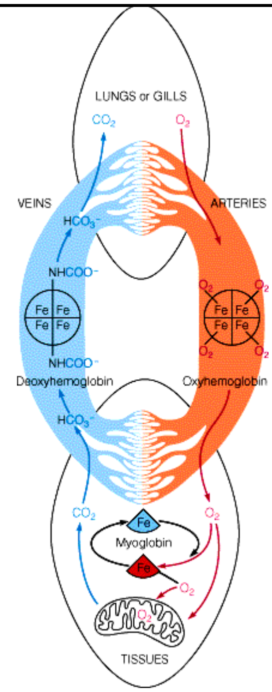
Bohr Effect

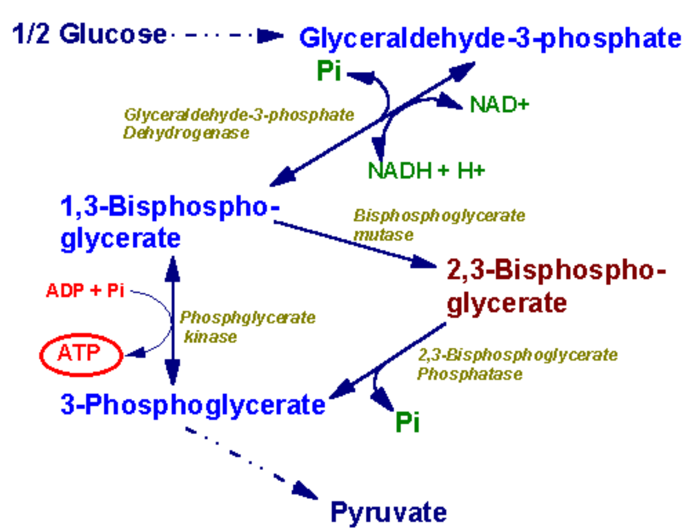
- Increased CO_2 leads to decreased pH

  $CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$
- At decreased pH several key AA's protonated, causes Hb to take on T-conformation (low affinity)
- In R-form same AA's deprotonated, form charge interactions with positive groups, stabilize R-conformation (High affinity)
- HCO_3- combines with N-terminal alpha-amino group to form carbamate group.

  $\text{--}N_3H^+ + HCO_3^- \leftrightarrow \text{--}NHCOO^-$
- Carbamation stabilizes T-conformation

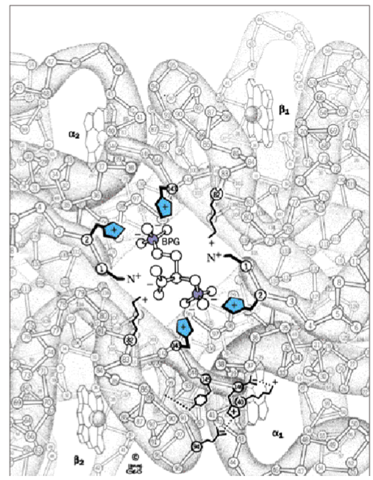



LUNGS or GILLS
CO_2
O_2
VEINS
ARTERIES
HCO_3^-
$NHCOO^-$
Fe Fe
Fe Fe
O_2
$NHCOO^-$
Deoxyhemoglobin
Oxyhemoglobin
HCO_3^-
CO_2
O_2
Fe
Myoglobin
Fe
O_2
TISSUES


1/2 Glucose
Glyceraldehyde-3-phosphate
Pi
NAD+
Glyceraldehyde-3-phosphate
Dehydrogenase
NADH + H+
1,3-Bisphospho-
glycerate
Bisphosphoglycerate
mutase
2,3-Bisphospho-
glycerate
ADP + Pi
Phosphglycerate
kinase
ATP
2,3-Bisphosphoglycerate
Phosphatase
3-Phosphoglycerate
Pi
Pyruvate

Bisphosphoglycerate (BPG)

- BPG involved in adaptation to high altitude
- Binding of BPG to Hb causes low O2 affinity
- BPG binds in the cavity between beta-Hb subunits
- Stabilizes T-conformation
- Fetal Hb (HbF, $\alpha_2\gamma_2$) has low affinity for BPG, allows fetus to compete for O_2 with mother's Hb (HbA, $\alpha_2\beta_2$) in placenta.



Fetal Hb (HbF) versus adult Hb (HbA)

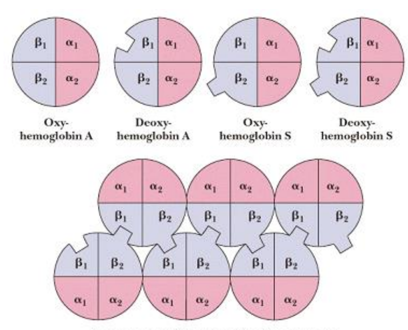
- HbF has higher affinity for O2 than HbA
- It lacks two positively charged amino acids that take part in binding 23BPG, i.e., His-143 of each B-globin chain is replaced by Serine. Thus 2,3BPG binds less tightly to HbF than HbA
- The P50 for HbF is 18 torr compared to 26 torr for HbA. Thus at PO2 of 20-40 torr in tissues HbF has higher affinity for O2. The difference in affinity allows efficient transfer of O2 from maternal blood to the fetus.

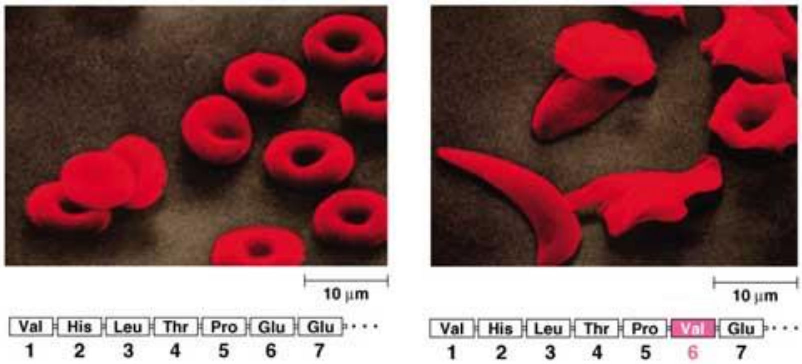
Class activity!

- If HbF and HbA have the same affinity toward oxygen, how this will affect the growth of fetus in utero?

Mutations in α- or β-globin genes can cause disease state

- Sickle cell anemia – E6 to V6
- Causes V6 to bind to hydrophobic pocket in deoxy-Hb
- Polymerizes to form long filaments
- Cause sickling of cells
- Sickle cell trait offers advantage against malaria
- Fragile sickle cells can not support malaria parasite



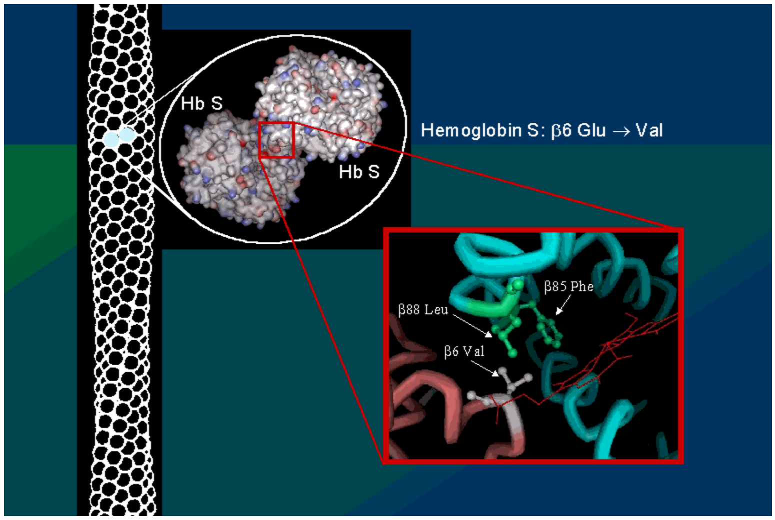


(a) Normal red blood cells and the primary structure of normal hemoglobin

(b) Sickled red blood cells and the primary structure of sickle-cell hemoglobin

Sickle cell (HbS) vs. HbC diseases

- **HbS:** GAG>GTG, Glu>Val;
- **HbC:** GAG>AAG, Glu>Lys



Class activity!

- Can you predict the phenotype or severity of phenotype from the amino acid change?
- How sickle cell disease gives an advantage for patients against Malaria?

End of chapter 4