BCMB 3100: Partial notes
Chapter 4 (Part 1)

- Diversity of proteins
- 3D structure of proteins
- Fibrous vs globular proteins
- Conformation vs configuration
- 1°, 2°, 3° and 4° structure
- Peptide groups in polypeptide
- $\phi$ vs $\Psi$ angles, Ramachandran plot
- Xray crystallography & NMR
- $\alpha$ helix vs $\beta$-sheet

Diversity of proteins

- ______________ - study of large sets of proteins, such as the entire complement of proteins produced by a cell
- $E. \ coli$ has about _________ different polypeptides (average size 300 amino acids, M, 33,000)
- Fruit fly ($Drosophila melanogaster$) about 16,000, humans, other mammals about 40,000 different polypeptides

$E. \ coli$ proteins on 2D gel electrophoresis

See Page 74/76

$E. \ coli$ expresses ~4000 proteins

3D STRUCTURE OF PROTEINS

Two classes of proteins:

- ______________: water insoluble, static, "tough", extended, provide mechanical support ($\alpha$-keratin, collagen)
- ______________: compact, "spherical", usually: hydrophobic interior & hydrophilic exterior enzymes
The biological activity of a protein depends on its **conformation**

**Conformation**: spatial arrangement of substituent groups that are free to assume different positions in space, without breaking any bonds, because of the freedom of bond rotation

The number of potential conformations of a protein is **infinite**. Under physiological conditions the protein assumes a single stable shape: **native conformation**

**Native conformation**: a spatial arrangement of atoms that can not be changed without breaking covalent bonds

---

**Levels of Protein Structure**

___________: the covalent backbone of a polymer

___________: the residue-by-residue conformation of the backbone of a polymer

___________: the 3D conformation of a polymer in its native folded state

___________: the 3D structure of a multisubunit, particularly the manner in which the subunits fit together

**Supersecondary structure**: clusters of secondary structure (e.g. $\beta\alpha\beta$)

**Domain**: a distinct structural unit of a polypeptide; domains have separate functions and may fold as independent, compact units
BCMB 3100: Partial notes for Chapter 4

- Diversity of proteins
- 3D structure of proteins
- Fibrous vs globular proteins
- Conformation vs configuration
- 1°, 2°, 3° and 4° structure
- Peptide groups in polypeptide
- \( \phi \) vs \( \Psi \) angles, Ramachandran plot
- Xray crystallography & NMR
- \( \alpha \) helix vs \( \beta \)-sheet

---

**Conformation of the Peptide Group:** The peptide group consists of 6 atoms

**Resonance structure of the peptide bond**

(a) Peptide bond shown as a C-N single bond
(b) Peptide bond shown as a double bond
(c) Actual structure is a hybrid of the two resonance forms. Electrons are delocalized over three atoms: O, C, N

---

**Planar peptide groups in a polypeptide chain**

- Rotation around C-N bond is restricted due to the double-bond nature of the resonance hybrid form
- Peptide groups (blue planes) are therefore planar

---

**See Fig. 4.8 Trans and cis conformations of a peptide group**

Nearly all peptide groups in proteins are in the trans conformation

- _____ conformation is less favorable than trans due to steric interference of \( \alpha \)-carbon side chains
- _____ conformation is established protein during synthesis
BCMB 3100: Partial notes for Chapter 4

- Diversity of proteins
- 3D structure of proteins
- Fibrous vs globular proteins
- Conformation vs configuration
- 1°, 2°, 3° and 4° structure
- Peptide groups in polypeptide
- \( \phi \) vs \( \Psi \) angles, Ramachandran plot
- X-ray crystallography & NMR
- \( \alpha \) helix vs \( \beta \)-sheet

---

Rotation of atoms in a peptide group

- Peptide group has a repeating \( \text{N-C}_\alpha \text{-C} \) backbone
- Rotation about both the \( \text{N-C}_\alpha \) (\( \phi \)) and \( \text{C}_\alpha \text{-C} \) (\( \Psi \)) bonds is possible
- Rotation of the \( \text{N-C}_\alpha \) bond in proline is restricted because of the pyrrolidine ring structure

---

See Pgs. 49 Rotation around the \( \text{N-C}_\alpha \) and \( \text{C}_\alpha \text{-C} \) bonds that link peptide groups

---

(steric contour diagram of allowed values of rotation permitted about the \( \text{N-C}_\alpha \) (\( \phi \)) and \( \text{C}_\alpha \text{-C} \) (\( \Psi \)) bonds.)
See Fig. 4.10 (a) Ramachandran Plot

(a) Solid lines: range of permissible $\phi$ and $\psi$ values
Dashed lines: outer limits for an alanine residue
Blue dots: values for known conformations

(b) Ramachandran Plot

• Observed $\phi$ and $\psi$ values in known structures. Crosses are typical values for a single protein.
• $\alpha$-Helix residues (red), $\beta$-Strand residues (blue), others (green).

BCMB 3100: Partial notes for Chapter 4

• Diversity of proteins
• 3D structure of proteins
• Fibrous vs globular proteins
• Conformation vs configuration
• $1^\circ$, $2^\circ$, $3^\circ$ and $4^\circ$ structure
• Peptide groups in polypeptide
• $\phi$ vs $\Psi$ angles, Ramachandran plot
• X-ray crystallography & NMR
• $\alpha$ helix vs $\beta$-sheet

---

Technique that reveals 3D position of atoms in a protein.

1. Need crystallized protein (often produced in presence of high salt)

2. Source of ______

X-ray strikes protein & in part is scattered. Scattering is detected by X-ray film or solid state electronic detector
3. Electrons in protein scatter X-rays (amplitude of wave is proportional to number of electrons)

4. Scattered waves recombine:
   in phase waves $\rightarrow$ complement
   out-of-phase waves $\rightarrow$ cancel

5. Pattern of recombinant waves (angles and intensities) depend upon atomic structure.

6. Pattern converted to atomic image by computer analysis (Fourier transformation).

---

**Ribonuclease A**

(a) Space-filling model
   (bound substrate analog black)
(b) Cartoon ribbon model
   (shows secondary structure)
(c) Substrate-binding site view

---

**Methods for Determining Protein Structure**

- X-ray crystallography is used to determine the three-dimensional conformation of proteins

---

**NMR determination of protein structure**

- a method used to determine structure of small proteins in solution

Ribonuclease A determined by NMR
   (polypeptide chain backbone)

Protein structures are deposited in the databases:
   e.g., Protein Data Bank (PDB)
BCMB 3100: Partial notes for Chapter 4

- Diversity of proteins
- 3D structure of proteins
- Fibrous vs globular proteins
- Conformation vs configuration
- 1°, 2°, 3° and 4° structure
- Peptide groups in polypeptide
- $\phi$ vs $\Psi$ angles, Ramachandran plot
- Xray crystallography & NMR
- $\alpha$ helix vs $\beta$-sheet

Secondary structure in proteins

Linus Pauling & Robert Corey (1950s)
proposed ______ and ______ as types of structures in proteins

right-handed

carbonyl O of aa residue n is H-bonded to $\alpha$-amino N of residue n+4

rise = 0.15nm, pitch = 0.54, 3.6 residue per turn

Fibrous proteins (e.g. keratin) may be largely $\alpha$-helix

Globular proteins vary greatly in $\alpha$-helix content: average $\alpha$ helix content of 26%

All side chains point outward from cylinder of helix

 Ala often found in $\alpha$ helix; Pro & Gly usually not present in helix but may be at ends

See Fig. 4.11
The $\alpha$-helix
See Fig. 4.11 Stereo view of right-handed $\alpha$ helix

- All side chains project outward from helix axis

Helix in horse liver alcohol dehydrogenase

(a) Amino acid sequence, (b) Helical wheel diagram

Horse liver alcohol dehydrogenase

- Amphipathic $\alpha$ helix (blue ribbon)
- Hydrophobic residues (blue) directed inward, hydrophilic (red) outward

Fig. 4.14 Leucine zipper of yeast

(See Fig. 4.21)

- DNA binding region consists of two amphipathic $\alpha$ helices, one from each of two protein subunits.

This transcription factor has amphipathic helices known as a leucine zipper. Leucine zippers have only a loose coiled structure compared to other coiled-coil amphipathic helices.
β sheet: extended polypeptide strands (β strands) stabilized by H-bonds between carbonyl O and amide H. Each amino acid accounts for 0.32-0.34 nm of length of polypeptide chain.

β strands may be parallel or antiparallel
side chains point alternatively above and below plane of strand
β sheets contain 2-15 strands with average of 6 aa residues per strand
β strand content of globular proteins is variable: Average of 19% β structure

Fig 4.15 β-Sheets (a) parallel, (b) antiparallel
(See Figures 4.16 & 4.17)

Stereo view of antiparallel β sheet
• Side chains (front β strand) alternate sides
• β-Strands twist in a right-handed direction

Interactions of β sheets
• β Sheet side chains project alternately above and below the plane of the β strands
• One surface of a β sheet may consist of hydrophobic side chains that can interact with other hydrophobic residues in protein interior
• __________ helices have hydrophobic side chains projecting outward that can interact with hydrophobic faces of β sheets or other helices
Structure of PHL P2 protein
(a) Blue/purple antiparallel β-sheets within a protein
(b) Stereo view of the β sandwich. Polar residues (red), hydrophobic residues (blue)

Example of amphipathic interaction in a β sheet

Loops and Turns (1)
- Loops and turns connect α helices and β strands and allow a peptide chain to fold back on itself to make a compact structure
- ________ - often contain hydrophilic residues and are found on protein surfaces
- ________ - loops containing 5 residues or less
- ________________ - connect different antiparallel β strands (also called hairpin loop)

Loops and Turns (2)
- ________: 4 aa, stabilized by H-bonds between α-carbonyl O of residue n and α-NH of residue (n+3); Pro is often 2nd residue
- __________: (also called glycine turn, β bend) like Type I Turn but 3rd residue is Gly

Reverse turns
See Fig. 4.20
(a) Type I, and (b) Type II
**Common motifs**

**Motifs - recurring protein structures**

- Helix-loop-helix
- Coiled-coil
- Helix bundle
- Hairpin
- Greek key
- β-Sandwich

**Supersecondary Structures (Motifs), Domains, Folds, Quaternary Structure**

- Anfinsen’s Experiment: denaturation, reduction & refolding
- Protein folding, Chaperones
- Collagen, Myoglobin & Hemoglobin

**Pyruvate Kinase**

- Main polypeptide chain folds into three distinct domains

- Independently folded, compact, distinct structural unit in proteins
- ~25 to ~300 amino acid residues
- Connected to each other by loops, bound by weak interactions between side chains
- May have separate functions
- Illustrate evolutionary conservation of protein structure
**Cytochrome c**

- Conservation of cyt c structure

(a) Tuna (+heme)
(b) Tuna
(c) Rice
(d) Yeast
(e) Bacteria

See Last slides of Chapter 3 Lecture for Phylogenetic Tree

**Structural similarity of** (a) lactate dehydrogenase, (b) malate dehydrogenase

(only 23% identical amino acids!!)

(a) *B. stereothermophilus*, (b) *E. coli*

**Four categories of protein domains**

1. All α - almost entirely α helices and loops
2. All β - only β sheets and non-repetitive structures that link the β strands
3. Mixed α/β - alternate regions of α helix and β strand (e.g. αβα motif)
4. α + β - local clusters of α helices and β sheet in separate, contiguous regions of the polypeptide chain

**Folds**

- A “fold” is a combination of secondary structures that form the core of a domain
- Some domains have simple folds, others have more complex folds
Common domain folds

Examples of tertiary structure

Tertiary Protein Structures (cont)

Domain Structure and Function

- A ________ may have a particular function
- ________ between 2 domains provide crevices, grooves, and pockets on the surface of a protein for binding or catalytic sites
- In multifunctional enzymes, each catalytic activity can be on one of several domains
**Quaternary Structure**

- **organization of subunits** in a protein with multiple subunits (an “oligomer”)
- Subunits (may be identical ($\alpha\alpha$) or different ($\alpha\beta$)) have a defined stoichiometry and arrangement
- Subunits are held together by many weak, noncovalent interactions (hydrophobic, electrostatic)

---

**Quaternary structure of multidomain proteins**

(a) Chicken trerose phosphate isomerase

(b) HIV-1 aspartic protease

Subunits = $\alpha_2$

---

(continued)

(c) Staphylococcus potassium channel prot

Subunits = $\alpha_4$

(d) Bacillus sp. K12 epsilon protein

Subunits = $\alpha_3$

---

(contin.)

(e) Human epinephrine guaifenesin phosphorylase transaminase

Subunits = $\alpha_3\beta_2$

(f) Rhodopseudomonas photosynthetic reaction center

Subunits = $\alpha_3\beta_2\gamma$

---
Protein Denaturation and Renaturation

- partial or complete unfolding of native conformation
- causes loss of biological activity
- caused by heat, extreme pH, detergents, chaotropic agents; due to disruption of non-covalent interactions
- some proteins can be refolded or renatured

("chaos-promoting"): chemicals that denature proteins (e.g. urea, guanidinium chloride)

Chaotropic agents do NOT cleave covalent bonds but disrupt 2\(^{\circ}\), 3\(^{\circ}\) and 4\(^{\circ}\) structure

Urea and guanidinium chloride
Two chaotropic agents

Disrupt hydrophobic interactions in interior of protein by disordering water molecular adjacent to the protein.

Chaotropic agents increase entropy of system by interfering with hydrogen bonds, Van der Waals forces, and hydrophobic effects.

- Heat denaturation of ribonuclease A
- Unfolding monitored by changes in ultraviolet (blue), viscosity (red), optical rotation (green)

\(T_m\) (melting temperature) = temperature when 50% of particular protein is denatured
oxidation
R-SH + HS-R → R-S-S-R

reduction

_______: the loss of electrons from a compound

_______: the gain of electrons by a compound or ion

Common reducing agents: β mercaptoethanol; dithiothreitol

Disulfide bridges in bovine ribonuclease A

Ribonuclease:
124 aa,
mainly β sheet,
4 disulfide bonds

(a) Location of disulfide bridges
(b) Stereo view of Cys-26 and Cys-84

showed that the information necessary to specify 3D structure of ribonuclease came from the amino acid sequence

The amino acid sequence specifies 3D structure!!

The native form of a protein (e.g. ribonuclease) appears to be the thermodynamically most stable structure

Denaturation and renaturation of ribonuclease A

see Fig. 4.33
Protein Folding and stabilization (1)

Cooperativity of folding: formation of one part of structure (e.g. initial aa interactions) leads to formation of remaining structure.

__________________ is MAJOR driving force in protein folding

Protein Folding and stabilization (2)

Protein folding → nonpolar side chains inside, most polar chains outside, those polar chains inside H-bond with each other → 2° structure

H-bonds and van der Waals forces stabilize globular protein folding

Covalent cross links (R-S-S-R) and ionic interactions may stabilize some globular proteins

CURRENT THEORY OF PROTEIN FOLDING

1. not random
2. cooperative and sequential
3. dependent on 1° structure
4. for some proteins 1° structure alone determines folding
5. some proteins require help to fold (e.g. enzymes or chaperones)
6. extremely rapid, native conformation is generally reached < 1 second
7. NOTE: most protein have single native 3D shape; but some exceptions

* ___________: proteins that bind newly synthesized polypeptides & assist in proper folding
* ___________ increase rate of correct folding and prevent the formation of incorrectly folded intermediates
* ___________ bind unassembled protein subunits, prevent incorrect aggregation before assembly into multisubunit protein
* Most chaperones are heat shock proteins

• ___________: proteins that bind newly synthesized polypeptides & assist in proper folding
• ___________ increase rate of correct folding and prevent the formation of incorrectly folded intermediates
• ___________ bind unassembled protein subunits, prevent incorrect aggregation before assembly into multisubunit protein
• Most chaperones are heat shock proteins
**E. coli chaperonin**

(a) (b) Core consists of 2 identical rings (7 GroE subunits in each ring)
(c) Protein folding takes place inside the central cavity

---

**Chaperonin-assisted protein folding**

- Hydrolysis of several ATP molecules is required

---

**COLLAGEN TRIPLE HELIX (1)**

- most abundant vertebrate protein (25-35% total protein); fibrous; found in bone, tendons, skin, blood vessels, cartilage; gelatin, glue
- Collagen is an aggregate of 3 left-handed helices coiled into a right-handed super helix!
- Has high percentage of Pro, Gly, Hyp (4-hydroxyproline),
- Also contains Hyl (hydroxylysine)

Structure elucidated by G. N. Ramachandran

---

**Fig 4.24 Stereo view of human Type III collagen triple helix**
Hyp & Hyl are formed posttranslationally

- Hyp allows more [_________] H-bonding & stabilization of collagen triple helix
- Pro & Hyp prevent formation of \( \alpha \) helices & make collagen fibers rigid
- Strengthened by inter- and intra-covalent crosslinks between allysine-Lys (allysine)
- Vitamin C (ascorbate) deficiency \( \rightarrow \) decrease in Hyp & Hyl formation \( \rightarrow \) [_________] (weakness in blood vessels & skin)

**Collagen triple helix (3)**

- Multiple repeats of \(-\text{Gly-X-Y-}\); X often proline, Y often 4-hydroxyproline
- Glycine residues located along central axis of a triple helix (other residues cannot fit)
- For each \(-\text{Gly-X-Y-}\) triplet, one interchain H bond forms between amide H of Gly in one chain and \( -\text{C=O} \) of residue X in an adjacent chain
- No intrachain H bonds exist in the collagen helix

**Interchain H bonding in collagen**

- Amide H of Gly in one chain is H-bonded to C=O in another chain
Some proteins require **cofactors** or **prosthetic groups** at their active sites for activity

____________: small, usually nonprotein molecule required for enzyme activity

____________: a metal ion or other non-amino acid molecule tightly bound to a protein and essential for its activity

____________: active protein with all its cofactors

____________: protein without its cofactors

---

**Myoglobin (Mb)** [See Chapter 9 !]

Binds $O_2$
Stores & transports $O_2$ in muscle
**Prosthetic group:** heme = Fe-protoporphyrin IX
Single polypeptide (153 aa) + heme
Globular protein with 8 $\alpha$-helices
Oxygenated myoglobin = oxymyoglobin
All polar residues (except 2 His’s) are located on protein surface

---

**Hemoglobin (Hb)**

Binds $O_2$
Transports $O_2$ in vertebrate blood
**Prosthetic group:** heme = Fe-protoporphyrin IX
Tetramer: $\alpha_2\beta_2$; $\alpha$ has 141 aa; $\beta$ has 146 aa; acts as dimer of $\alpha\beta$
$3^\circ$ structure of each subunit similar to myoglobin

---

**Heme Fe(II)-protoporphyrin IX**

- Porphyrin ring provides four of the six ligands surrounding iron atom

The ferrous iron is held in place by binding of nitrogens of the 4 pyrrole rings

See Pg. 143 / 151
The structure of globins has been conserved in many species. Amino acid changes (substitutions) may or may not affect protein conformation. The following changes do NOT significantly effect conformation (i.e. Val → Ile):

- Val

The following change DO effect conformation (Glu → Val):

- Glu

---

**Sperm whale oxymyoglobin**

- Oxygen (red)
- His-93 and His-64 (green)

---

**Hemoglobin tetramer**

See Fig. 4.30  See Fig. 9.6

(a) Human oxyhemoglobin  (b) Tetramer schematic

- α-Globin (blue)
- β-Globin (purple)
- Myoglobin (green)

---

**Tertiary structure of myoglobin, α-globin and β-globin**

- α-Globin (blue)
- β-Globin (purple)
- Myoglobin (green)
**Oxygen-binding site of whale oxymyoglobin**

- Octahedral geometry of coordination complex (six ligands around iron)
- His-93 (proximal histidine) liganded to Fe
- His-64 (distal histidine)

*See Fig. 9.4*

**Oxygen-binding curves**

(a) Comparison of $O_2$-binding to Mb and Hb

Hyperbolic curve (single equilibrium constant)

$$Mb + O_2 \Leftrightarrow MbO_2$$

Fractional Saturation = $Y$

$$Y = \frac{[MbO_2]}{[MbO_2] + [Mb]}$$

$p_{50}$ = partial pressure at half saturation

*See Figures 9.1 & 9.2*

(b) Oxygen-binding curves for observed and for high and low affinity forms of hemoglobin

(b) Binding of the R (high-affinity) and T (low affinity) forms of Hb

**Conformational changes in a hemoglobin chain induced by oxygenation**

- Oxygen binding to Fe pulls the His toward ring plane
- Helix with His shifts position, disrupting some ion pairs between subunits (blue to red position)

*See Fig. 9.7*
Enhanced activity resulting from cooperation between subunits of an allosteric protein.

A protein having multiple active sites as well as distinct regulatory sites that control the flow of biochemicals through a metabolic pathway.

Hemoglobin is an allosteric protein. Regulatory protein whose activity is modulated by noncovalent binding of a specific metabolite at a site other than the active site.

Small molecules that bind to allosteric proteins and regulate their activity. Allosteric regulation is caused by small changes in native conformation of a protein.

Active shape = R (relaxed)  Inactive shape = T (taut)

Allosteric inhibitor: R → T

Allosteric activator

2,3-bisphospho-D-glycerate (2,3BPG) is an allosteric effector of hemoglobin. It lowers the affinity of deoxyhemoglobin for O₂ (raises P50). [Know physiological significance for Exam]

Bohr effect: the increase in P50 of hemoglobin caused by a lowered pH due to an increase in CO₂
Binding of 2,3BPG to deoxyhemoglobin

- Charges on 2,3BPG pair with (+) charges lining the central cavity, stabilizing the DeoxyHb form
- α-Subunits pink, β-subunits blue, heme groups red

Bohr effect

- Lowering the pH decreases the affinity of Hb for oxygen

Antibodies Bind Specific Antigens

- Vertebrate immune systems synthesize protein antibodies (immunoglobulins) to eliminate bacteria, viruses, other foreign substances
- Antibodies specifically recognize and bind antigens
- Antibodies are synthesized by lymphocytes (white blood cells)
Light Chain contains 2 domains; Heavy Chain contains 4 domains
Each domain: 110 aa in common motif known as immunoglobulin fold

Review of Globular Protein 3D Structure
Most globular proteins have compact globular shape due to many reversible turns in direction combined with the α helix and/or β structure. Usually, hydrophobic aa residues are in the interior and hydrophilic aa residues on the exterior of the protein.

The loops & turns contain nonrepetitive regions of 2° structure.
loops: range from ~ 2-16 residues, many hydrophilic residues found at surface of protein (can H-bond with water)
turn: loops having only a few residues (<6)

Average structure of Globular Protein
α helix: 26%; β structure: 19%; turns: 15%; simple loops 21%; complex loops: 10%
Four categories of protein domains

1. **All α** - domains almost entirely α helices and loops
2. **All β** - domains contain only β sheets and non-repetitive structures that link the β strands
3. **Mixed α/β** - supersecondary structures where regions of α helix and β strand alternate (e.g. αβα motif)
4. **α + β** - local clusters of α helices and β sheet in separate, contiguous regions of the polypeptide chain

Protein Denaturation and Renaturation

- **Denaturation** - disruption of native conformation of a protein, with loss of biological activity
- Most denatured proteins adopt a random-coil conformation
- Proteins denatured by heating or chemicals
- Some proteins can be refolded or renatured

Protein Folding and stabilization (1)

**Cooperativity of folding**: formation of one part of structure (e.g. initial aa interactions) leads to formation of remaining structure.

**Hydrophobic effect** is MAJOR driving force in protein folding

Folded proteins occupy a low-energy well that makes the native structure most stable

Many proteins can fold spontaneously to this low-energy conformation

Proteins are thought to fold “cooperatively” ... the first few interactions assist subsequent alignment and folding

Protein folding (more detail)

- extremely rapid, native conformation is generally reached < 1 second
- During folding the polypeptide collapses in upon itself due to the hydrophobic effect
- An intermediate “molten globule” forms with elements of secondary structure
- The backbone is rearranged to achieve a stable native conformation
• **Chaperones**: proteins that bind newly synthesized polypeptides & assist in proper folding

• **Chaperones** increase rate of correct folding and prevent the formation of incorrectly folded intermediates

• Chaperones bind to unassembled protein subunits to prevent incorrect aggregation before they are assembled into a multisubunit protein

• Most chaperones are heat shock proteins (synthesized as temperature increases)

---

Two conformations of hemoglobin: T and R

• **Active (R state)** and **inactive (T state)** forms are in rapid equilibrium in allosteric proteins

• Binding of **substrates** and allosteric **activators** stabilize the R state and shift the equilibrium in the **R direction**

• Allosteric **inhibitors** stabilize the T state and shift the equilibrium in the **T direction**