Chapter 8

Three-Dimensional Structures of Proteins
Peptide bonds

Partial double bond character
(No rotation around C-N bond, Planar)

Trans-Peptide bond
(a)

Resonance hybrid
(b)

Amide planes within a fully extended tetrapeptide residue
(c)

Amino terminus

Carboxyl terminus
Torsion (Dihedral) angles

\[ \phi \text{ (phi): rotation around } C\alpha-N \]
\[ \psi \text{ (psi) rotation around } C\alpha-C \]
Torsion angles

- When both peptide groups are in the same plane and the polypeptide is stretched out, the torsion angles are defined as 180°
- Looking from the Cα rotating clockwise increases the angle
- The torsion angles describes how the backbone is folded up
- Only a small portion of φ and ψ angle combinations are sterically possible
φ and ψ are 180°

φ and ψ are 0°
Most torsion angle combinations are prohibited.
The Ramachandran diagram

Calculated for poly-L-alanine

Observed for proteins (except for Gly and Pro)
Bonds and Interactions in Proteins

- Hydrogen bonding between peptide groups
- Metal ion coordination
- Hydrophobic interactions among nonpolar side chains
- Ionic linkages

- α-Helical structure
- β-Sheet structure
- Disulfide bond
- Side chain hydrogen bonding

N → R_{Leu} → R_{Val} → S → R_{Ile} → C
Elements of secondary structure

- $\alpha$-Helix
- $\beta$-Sheet
  - Anti-parallel
  - Parallel
- Bend (Reverse turn)
- Loop
Helices are characterized by p, n, and d

- p - pitch (distance the helix rises per turn)
- n - number of repeating units per turn
- d - helical rise per repeating unit \((p/n)\)
Characteristics of $\alpha$-helix

- Right-handed
- $\phi (-57^\circ)$ and $\psi (-47^\circ)$
- $n = 3.6$ residues per turn
- Pitch is 5.4 Å
- Hydrogen bonds between $n$ (carbonyl oxygen) and $n+4$ (amide hydrogen)
- Side chains are facing outwards
Other polypeptide helices

\[ n_m \]

\( n \) = number of residues per turn

\( m \) = number of atoms in the ring that is closed by hydrogen bond
Occasionally observed  
Mostly observed  
Rarely observed
\( \beta \) Sheets

- Fully extended peptide conformation \((\phi = \psi = \pm 180^\circ)\)
- Parallel or antiparallel
- Hydrogen bonding between neighboring polypeptide chains
Anti-parallel $\beta$-sheets
Parallel $\beta$-sheets
Connecting $\beta$-sheets

- Hairpin
- Right-handed crossover
- Left-handed crossover
β Bend (Reverse turn)

Type I and II

To N-terminus to C-terminus
Ω Loop

Courtesy of George Rose, Washington University School of Medicine
Fibrous proteins

- Structural materials (e.g. skin, tendon and bone proteins) – Protective, connective or supportive roles
- Motor proteins (e.g. muscle and ciliary proteins) – Motive functions

<table>
<thead>
<tr>
<th>Structure</th>
<th>Characteristics</th>
<th>Examples of occurrence</th>
</tr>
</thead>
<tbody>
<tr>
<td>α Helix, cross-linked by disulfide bonds</td>
<td>Tough, insoluble protective structures of varying hardness and flexibility</td>
<td>α-Keratin of hair, feathers, and nails</td>
</tr>
<tr>
<td>β Conformation</td>
<td>Soft, flexible filaments</td>
<td>Silk fibroin</td>
</tr>
<tr>
<td>Collagen triple helix</td>
<td>High tensile strength, without stretch</td>
<td>Collagen of tendons, bone matrix</td>
</tr>
</tbody>
</table>
α Keratin – A helix of helices
Two-stranded coiled coil

Nonpolar residues
"Permanent wave"
Collagen – A triple helical cable
Amino acid sequence of the collagen

Gly-X-Y

X= Pro etc
Y= Hyp etc

<table>
<thead>
<tr>
<th>Gly</th>
<th>Pro</th>
<th>Arg</th>
<th>Gly</th>
<th>Pro</th>
<th>Hyp</th>
<th>Gly</th>
<th>Ser</th>
<th>Ala</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gly</td>
<td>Ser</td>
<td>Hyp</td>
<td>Gly</td>
<td>Lys</td>
<td>Asp</td>
<td>Gly</td>
<td>Leu</td>
<td>Asn</td>
</tr>
<tr>
<td>Gly</td>
<td>Leu</td>
<td>Hyp</td>
<td>Gly</td>
<td>Pro</td>
<td>Ile</td>
<td>Gly</td>
<td>Hyp</td>
<td>Hyp</td>
</tr>
<tr>
<td>Gly</td>
<td>Pro</td>
<td>Arg</td>
<td>Gly</td>
<td>Arg</td>
<td>Thr</td>
<td>Gly</td>
<td>Asp</td>
<td>Ala</td>
</tr>
<tr>
<td>Gly</td>
<td>Pro</td>
<td>Ala</td>
<td>Gly</td>
<td>Pro</td>
<td>Hyp</td>
<td>Gly</td>
<td>Pro</td>
<td>Hyp</td>
</tr>
<tr>
<td>Gly</td>
<td>Pro</td>
<td>Hyp</td>
<td>Gly</td>
<td>Pro</td>
<td>Hyp</td>
<td>Gly</td>
<td>Pro</td>
<td>Pro</td>
</tr>
</tbody>
</table>

4-Hydroxyproply residue (Hyp)

3-Hydroxyproply residue

5-Hydroxylysyl residue (Hyl)
Interchain hydrogen bonds
Packing of collagen molecules

Collagen molecule

Packing of molecules

Hole zone 0.6D

Overlap zone 0.4D

Courtesy of Karl A. Plew, Collagen Corporation
Covalent cross-linking of collagens

Lys

lysyl oxidase

\[ \text{C} = \text{O} \]
\[ \text{CH} - (\text{CH}_2)_3 - \text{CH} \]
\[ \text{NH} \]

Allysine

lysyl oxidase

\[ \text{O} \]
\[ \text{HC} - (\text{CH}_2)_3 - \text{CH} \]
\[ \text{NH} \]

Allysine

\[ \text{C} = \text{O} \]
\[ \text{CH} - (\text{CH}_2)_2 - \text{C} = \text{CH} - (\text{CH}_2)_3 - \text{CH} \]
\[ \text{NH} \]
\[ \text{O} \]

Allysine aldol

Aldol-His

\[ \text{O} \]
\[ \text{OH} \]
\[ \text{NH} \]
\[ \text{CH} - (\text{CH}_2)_2 - \text{CH} - (\text{CH}_2)_3 - \text{NH} \]

His

Histidinodehydrohydroxymerodesmosine
Silk Fibroin

- Produced by insects and spiders
- Packing of $\beta$ sheets by interlocking arrangement of side chains (Gly and Ala)
Fibroin is a stack of $\beta$ sheets
Globular proteins

- Enzymes
- Transport proteins
- Receptors
Single crystal X-ray diffraction of proteins
Electron density map
Resolution of the electron density maps

(a) 6.0-Å resolution
(b) 2.0-Å resolution
(c) 1.5-Å resolution
(d) 1.1-Å resolution
Most crystalline proteins maintain their native conformations

- Hydrated (40~60% water in a typical protein crystal)
- X-ray crystal structure $\approx$ NMR structure
- Many enzymes are catalytically active in the crystalline state
Protein structure determination by NMR

Useful for studying protein folding and dynamics
Structural representations of proteins

Stick

Cartoon

Cα backbone
Tertiary structure of proteins

- $\alpha$ helices and/or $\beta$ sheets
- Side chain location varies with polarity
  - Nonpolar residues – mainly interior of proteins (Hydrophobic interactions)
  - Charged polar residues – mainly surface of proteins
  - Uncharged polar residues – both surface and interior (hydrogen bonding interactions)
- Cores maintain relaxed conformations
- Domains
- Structural motifs (Supersecondary structures)
Localization of side chains

- Red: Hydrophobic side chains
- Green: Hydrophilic side chains
Protein domains and families

- Larger proteins folds into two or more globular clusters known as Domains
- Domains in a protein are structurally independent units
- Family of proteins - proteins with similar folding pattern (look similar) and sequence similarities
Domain has a separate function

- Catalytic domain
- Binding domain
  - Small molecules
  - DNA or RNA
  - Proteins
- Transmembrane domain
Structural motifs (Supersecondary structure)

- $\beta\alpha\beta$ motif
- $\beta$ haipin motif
- $\alpha\alpha$ motif
- Greek key motif ($\beta\beta\beta\beta$ motif)
α Domains

*E. coli* cytochrome *b*562

Human growth hormone
β Domains

Immunoglobulin (Ig) fold

Retinol binding protein (Up-and-down β barrel)
Two other types of β barrels

Greek key motif

γ-B crystallin

Jelly roll barrel (Swiss roll barrel)
α/β Barrel (TIM barrel)

Triose phosphate isomerase (TIM)
Open $\beta$ sheets

Carboxypeptidase A

Glyceraldehyde-3-phosphate dehydrogenase

Adenylate kinase
Rossman fold
(Dinucleotide-binding fold)

Lactate dehydrogenase (LDH)

2x βαβαβ motifs → Dinucleotide binding

[Image showing the structure of LDH with βαβαβ motifs and dinucleotide binding]
Doubly wound sheets
Structural bioinformatics

- Structural data bases
- Molecular visualization
- Structural classification and comparison
**Structural Databases**

**Protein Data Bank (PDB):**
http://www.rcsb.org/pdb/

**Nucleic Acid Databank:**
http://ndbserver.rutgers.edu/NDB/ndb.html

**Molecular Modeling Database (MMDB):**

**PQS Protein Quaternary Structure Query Form at the EBI:**
http://pqs.ebi.ac.uk/
Molecular Graphics Programs/Plug-ins

Chime:
http://www.mdli.com/cgi/dynamic/welcome.html

Cn3D:

MAGE:
http://kinemage.biochem.duke.edu/

Protein Explorer:
http://www.umass.edu/microbio/chime/explorer/index.htm

RasMol:
http://www.bernstein-plus-sons.com/software/rasmol/ and
http://www.umass.edu/microbio/rasmol/index.html

Virtual Reality Modeling Language (VRML):
Requires a VRML plug-in, available
through http://www.web3d.org/vrml/vrml.htm
Structural Classification Algorithms

CATH (class, architecture, topology and homologous superfamily):
http://www.biochem.ucl.ac.uk/bsm/cath_new/index.html

CE (combinatorial extension of optimal pathway):
http://cl.sdsc.edu/

FSSP (fold classification based on structure–structure alignment of proteins):
http://www2.ebi.ac.uk/dali/fssp/

SCOP (structural classification of proteins):
http://scop.mrc-lmb.cam.ac.uk/scop/

VAST (vector alignment search tool):
Multisubunit proteins

- Identical or non-identical subunits
- Identical subunits are called protomers
- Subunits are usually non-covalently associated
- Interfaces often contain tightly packed non-polar residues
Why multisubunits?

- Different sites for subunit
- Defective subunits can be replaced
- Less genetic information
- It is easier to maintain smaller genes
- Allosteric effect
Quaternary structure

Hemoglobin
Glutamine synthetase
Helical quaternary structure

Subunit

Helix segment

Helix
Protein stability

- The hydrophobic effect
  - Main driving force for protein native structure
- Van der Waals Interactions
  - Important roles in protein stability
- Electrostatic and hydrogen bonding Interactions
  - Some role in the stability of protein
- Disulfide bonds and metal-ligand interactions
  - Critical role in the stability of some proteins
# Hydrophathic index of amino acids

<table>
<thead>
<tr>
<th>Side Chain</th>
<th>Hydropathy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ile</td>
<td>4.5</td>
</tr>
<tr>
<td>Val</td>
<td>4.2</td>
</tr>
<tr>
<td>Leu</td>
<td>3.8</td>
</tr>
<tr>
<td>Phe</td>
<td>2.8</td>
</tr>
<tr>
<td>Cys</td>
<td>2.5</td>
</tr>
<tr>
<td>Met</td>
<td>1.9</td>
</tr>
<tr>
<td>Ala</td>
<td>1.8</td>
</tr>
<tr>
<td>Gly</td>
<td>0.4</td>
</tr>
<tr>
<td>Thr</td>
<td>0.7</td>
</tr>
<tr>
<td>Ser</td>
<td>0.8</td>
</tr>
<tr>
<td>Trp</td>
<td>0.9</td>
</tr>
<tr>
<td>Tyr</td>
<td>1.3</td>
</tr>
<tr>
<td>Pro</td>
<td>1.6</td>
</tr>
<tr>
<td>His</td>
<td>3.2</td>
</tr>
<tr>
<td>Glu</td>
<td>3.5</td>
</tr>
<tr>
<td>Gln</td>
<td>3.5</td>
</tr>
<tr>
<td>Asp</td>
<td>3.5</td>
</tr>
<tr>
<td>Asn</td>
<td>3.5</td>
</tr>
<tr>
<td>Lys</td>
<td>3.9</td>
</tr>
<tr>
<td>Arg</td>
<td>4.5</td>
</tr>
</tbody>
</table>

Hydropathic index plot
Protein denaturation

• Temperature
• pH
• Detergent
• Chaotrophic agents (Guanidium ion, urea etc)
Thermostable proteins

• Little overall difference in secondary and tertiary structure

• Extensive network of salt bridges on the surface

• Combination of small effects
  – Larger hydrophobic core
  – Larger interface between domains and/or subunits
  – More tightly packed core