

Bioorganic Chemistry

Chem 468

August 30, 2010

Bioorganic Chemistry

Developments in medicinal chemistry, biotechnology and biotherapeutics have blurred the traditional boundaries between chemistry and biology.

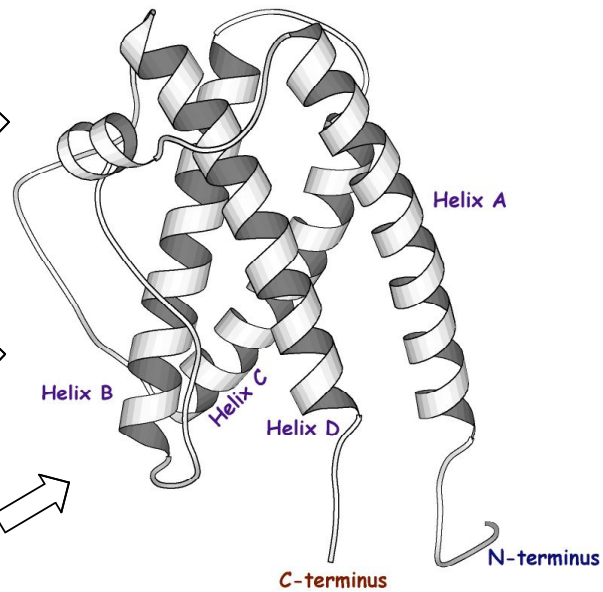
Week	Date	Topic	In Class	Reading
1	08/30/10	Introduction – amino acids and proteins	Lecture	Lecture Notes (supplement with a general biochem. text)
2	09/06/10	Labor Day – No Classes		
3	09/13/10	Peptide chemical synthesis	Lecture	Lecture Notes (supplement with organic text)
4	09/20/10	Peptide Mimetics	Lecture + Paper discussion	Lecture Notes + Papers to be assigned
5	09/27/10	Combinatorial Chemistry	Lecture + Paper discussion	Lecture Notes + Papers to be assigned
6	10/04/10	Test I		Lectures from weeks 1-5
7	10/12/10	Protein/peptide modification	Lecture + Paper discussion	Lecture Notes + Papers to be assigned
8	10/18/10	Enzymes in Chemical Synthesis	Lecture + Paper discussion	Lecture Notes + Papers to be assigned
9	10/25/10	TBA	TBA	TBA
10	11/01/10	Introduction to carbohydrates	Lecture	Lecture Notes (supplement with general biochem text)
11	11/08/10	Test II		Lectures from weeks 7-10
12	11/15/10	Chemical synthesis of oligosaccharides and carbohydrates	Lecture + Paper discussion	Lecture Notes + Papers to be assigned
13	11/22/10	Cyclodextrins & Student Presentations	Lecture + Paper	Lecture Notes + Papers to be assigned
14	11/29/10	Student Presentations		
15	12/06/10	Student Presentations		

Granulocyte colony-stimulating factor (G-CSF)

A protein hormone, stimulating the production and activation of neutrophils (a type of white blood cell).

The natural protein was developed by AMGEN into a therapeutic (Neupogen®) used to reduce the risks of infection in cancer patients undergoing chemotherapy.

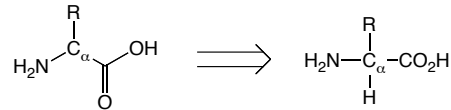
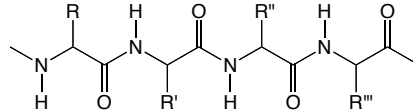
Amgen now produces a chemically-modified G-CSF (Neulasta®) with improved therapeutic properties.



Amino Acids

Amino Acids, Peptides and Proteins

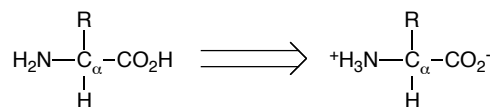
- Proteins and peptides have a covalently linked linear sequence.
- Amino acids are the basic building blocks of peptides and proteins.
- All proteins are assembled from the 20 "standard" amino acids, α -amino acids.
- Amino acid side chains confer distinct chemical properties.
- The observed physical and chemical properties of proteins are direct results of the sequence of these amino acid.



5

General Properties

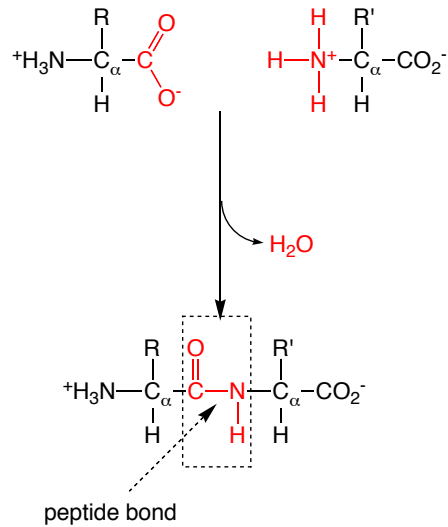
- pK of the α -carboxylic acids are in the range near 2.2.
- pK of the α -amino groups are near 9.4.
- At physiological pH, both the carboxylic acid and the amino groups are ionized, zwitterions.
- Amino acids can serve in acid or base capacities, they are amphoteric.



6

Peptide Bonds

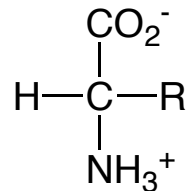
- Peptides and proteins are linear polymers of amino acids linked by "peptide bonds".
- Peptide bonds are amide bonds between amino acids.
- Formation of peptide bond produces a water molecule.
- Dipeptides, tripeptides, oligopeptides and polypeptides... proteins.



7

Classification of Amino Acids

- Classification based on properties and characteristics of the amino acid side chains.
 - Nonpolar
 - Uncharged Polar
 - Charged polar



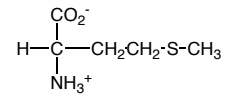
8

Nonpolar Side Chains

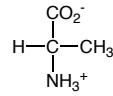
- Nine amino acids with hydrophobic side chains.
- Aliphatic side chains.
- Aromatic side chains.



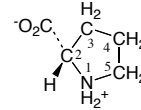
Glycine (Gly, G)



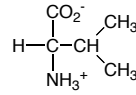
Methionine (Met, M)



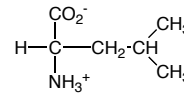
Alanine (Ala, A)



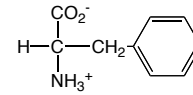
Proline (Pro, P)



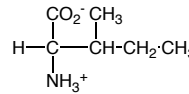
Valine (Val, V)



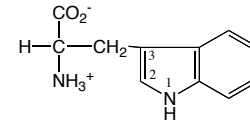
Leucine (Leu, L)



Phenylalanine (Phe, F)



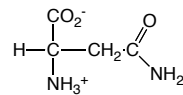
Isoleucine (Ile, I)



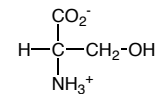
Tryptophan (Trp, W) 9

Uncharged Polar Side Chains

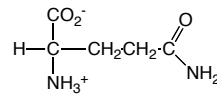
- Six amino acids with uncharged polar side chains.
- Side chain hydroxyl.
- Side chain amide.
- Side chain phenol.
- Side chain thiol.



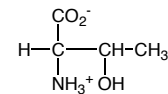
Asparagine (Asn, N)



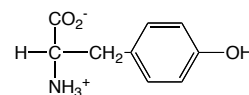
Serine (Ser, S)



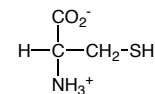
Glutamine (Gln, Q)



Threonine (Thr, T)



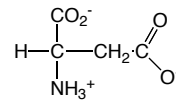
Tyrosine (Tyr, Y)



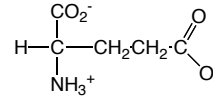
Cysteine (Cys, C)

Charged Polar Side Chains

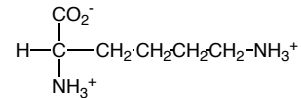
- Five amino acids have charged polar side chains.
- Side chain carboxylic acid groups.
- Side chain amines and guanidino group.
- Sidechain imidazole group.



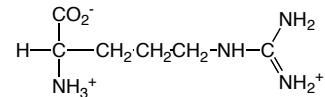
Aspartic Acid (Asp, D)



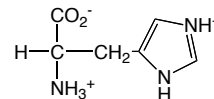
Glutamic Acid (glu, E)



Lysine (Lys, K)



Arginine (Arg, R)



Histidine (His, H)

11

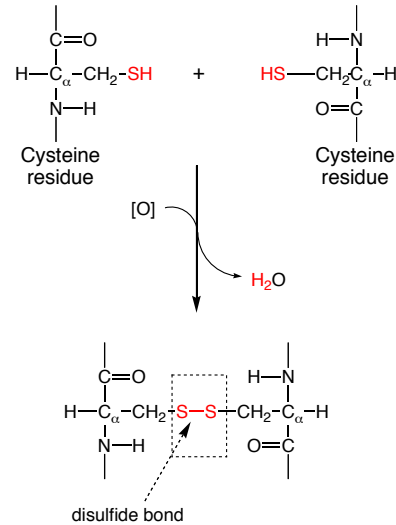
Hydropathy of Amino Acid Side Chains

Amino Acid	Hydropathy
Ile	4.5
Val	4.2
Leu	3.8
Phe	2.8
Cys	2.5
Met	1.9
Ala	1.8
Gly	-0.4
Thr	-0.7
Ser	-0.8
Trp	-0.9
Tyr	-1.3
Pro	-1.6
His	-3.2
Glu	-3.5
Gln	-3.5
Asp	-3.5
Asn	-3.5
Lys	-3.9
Arg	-4.5

12

Disulfide Bonds

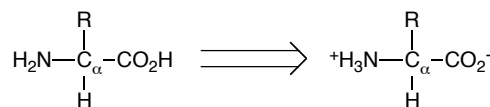
- Cysteine (Cys, C) residues are unique.
- They can form disulfide bonds between two Cys residues.
- Disulfide bonds can link separate polypeptide chains.
- ... or they can link cysteine residues within the same polypeptide chain.



13

General Properties

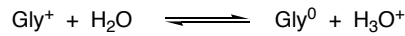
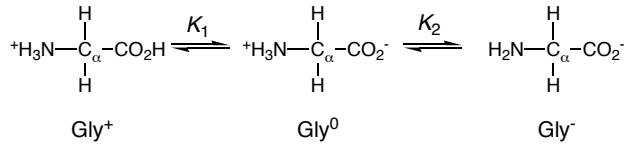
- pK of the α -carboxylic acids are in the range near 2.2.
- pK of the α -amino groups are near 9.4.
- At physiological pH, both the carboxylic acid and the amino groups are ionized, zwitterions.
- Amino acids can serve in acid or base capacities, they are amphoteric.



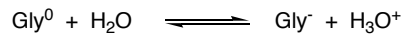
14

Acid-Base Properties

- α -amino acids have 2 (or 3) ionizable groups.
- pK of α -carboxylic acid is influenced by α -NH₃⁺. (pK of acetic acid = 4.76)
- pK of α -NH₃⁺ is influenced by α -carboxylate. (pK of Gly-methyl ester = 7.75)
- Side chain functional groups see similar but weaker influence.
- Titration curves of proteins and peptides tend to be complicated, and rarely reflect individual pK values.



$$K_1 = \frac{[\text{Gly}^0][\text{H}_3\text{O}^+]}{[\text{Gly}^+]} \quad \text{p}K_1 = \sim 2.0-2.4$$



$$K_2 = \frac{[\text{Gly}^-][\text{H}_3\text{O}^+]}{[\text{Gly}^0]} \quad \text{p}K_2 = \sim 9.0-9.8$$

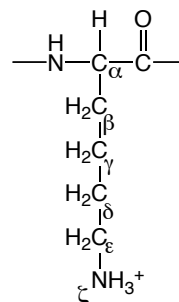
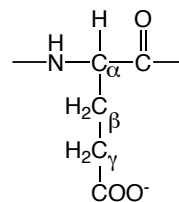
$$\text{pH} = \text{p}K + \log\left(\frac{[\text{A}^-]}{[\text{HA}]}\right)$$

$$\text{pI} = 1/2(\text{p}K_1 + \text{p}K_2)$$

15

Amino Acid Nomenclature

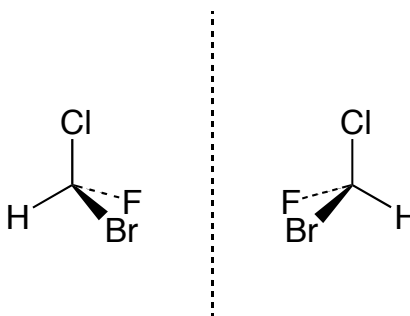
- In peptides amino acid residues are named by replacing -ine with -yl.
- Peptides are written from amine terminus (*N*-terminus) and carboxyl terminus (*C*-terminus).
- Glx reflects uncertainty between Glu or Gln.
- Asx reflects uncertainty between Asp or Asn.
- Position of nonhydrogen side chain atoms indicated by greek alphabet (α , β , γ , δ , ϵ , ζ , ...)



16

Amino Acids: Optical Activity

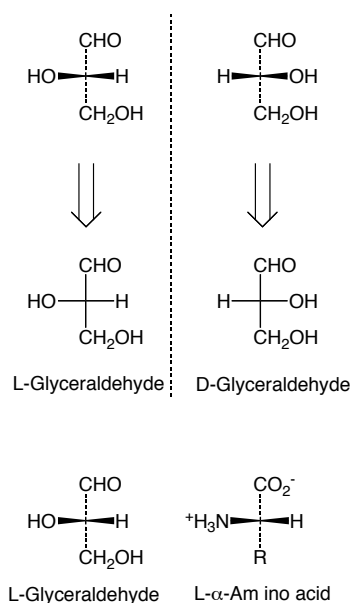
- All amino acids other than glycine are optically active.
- They demonstrate an asymmetry such that their mirror images are not superimposable.
- Asymmetric centers \leftrightarrow chiral centers.
- Enantiomers are molecules that are nonsuperimposable mirror images of each other.



17

Fischer Convention

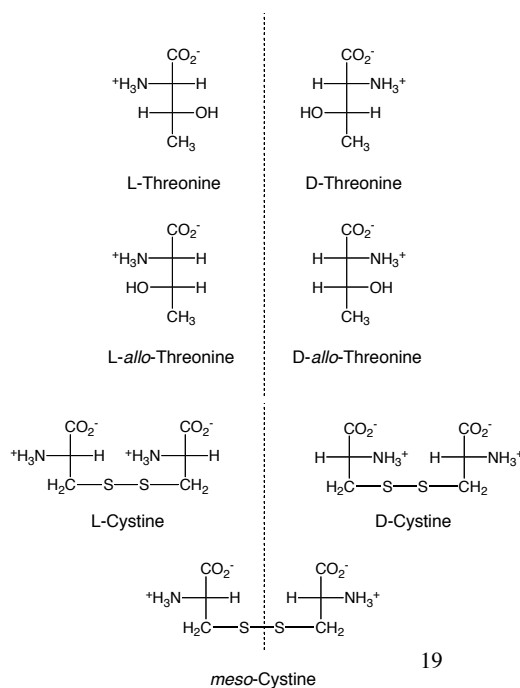
- Configuration of groups around a chiral center can be related to glyceraldehyde.
- Designate:
 - (+) isomer \rightarrow D-glyceraldehyde.
 - (-) isomer \rightarrow L-glyceraldehyde.
- All α -amino acids from proteins have the L stereochemical configuration (glycine being achiral, is an exception.).



18

Diastereomers

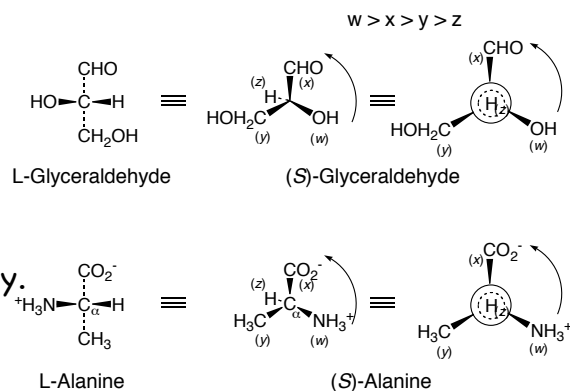
- Molecules with multiple (2 or more) chiral/asymmetric centers have 2^n isomers.
- Diastereomers are stereoisomers that differ by at least one but not all asymmetric centers.
- Diastereomers of L-amino acids are referred to as the allo forms.
- The D-allo and L-allo diastereomers are enantiomers of each other.
- Diastereomers are physically and chemically distinct.
- Meso- refers to molecules with asymmetric centers that have an internal plane of reflection. (such molecules are optically inactive.)



19

Cahn-Ingold-Prelog System

- The Fischer nomenclature system can be ambiguous and awkward.
- The Cahn-Ingold-Prelog system provides a nomenclature that unambiguously indicates the absolute configuration of the asymmetric center.
- Substituents on the chiral carbon atom are ranked in order of priority.
- View chiral center with lowest priority group directed back... into the page.
- If priority of substituent groups decreases going clockwise around the asymmetric center, then designate it *R*.
- If counterclockwise, designated *S*.



Order of priority functional groups common in biomolecules

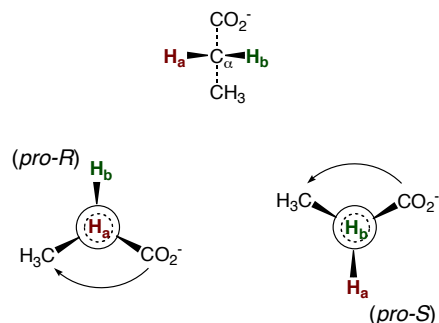
$$\left\{ \begin{array}{l} \text{SH} > \text{OH} > \text{NH}_2 > \text{COOH} > \text{CHO} \\ > \text{CH}_2\text{OH} > \text{C}_6\text{H}_5 > \text{CH}_3 > {}^2\text{H} > {}^1\text{H} \end{array} \right.$$

20

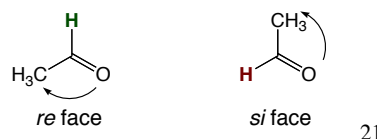
Prochiral Centers

- Two chemically identical substituents on a tetrahedral carbon may be geometrically distinct.
- Two such atoms are referred to as being prochiral.
- Designated as *pro-R* or *pro-S* based on same criteria as *R* and *S*.
- Property of prochirality also applicable to planar carbon centers.
- Faces are designated as *re* face or *si* face.

Tetrahedral centers



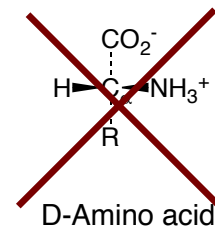
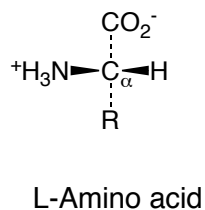
Trigonal planar centers



21

Chirality and Biochemistry

- Chemical synthesis of chiral molecules usually results in racemic mixtures.
- In order to chemically or physically differentiate between enantiomers, the process must involve an asymmetric element or influence. (i.e. chiral reagents or chiral chromatography columns).
- The **biosynthesis** of compounds containing asymmetric centers almost always yields enantiopure products.



22

Covalent Structures of Proteins

Proteins

- Proteins serve in a broad range of capacities.
- They are often referred to as the “building blocks” of life.
- They can vary enormously in size and composition.
- They demonstrate unbelievable functional diversity.
- Their biological function is a direct result of their unique and complex structures.

Protein Functions

- Enzymes
- Hormones
- Antibodies
- Transporters
- Muscle
- Feathers
- Spider Webs
- Rhinoceros Horn
- Poisons
- Antibiotics
- Structural

25

Levels of Organization

- **Primary structure (1° structure):** the amino acid sequence of polypeptide chain.
- **Secondary structure (2° structure):** local spatial organization and arrangement of the peptide backbone. Generally refers to easily localized structural elements (i.e. helices and sheets).
- **Tertiary structure (3° structure):** the comprehensive three-dimensional structure of a protein.
- **Quaternary structure (4° structure):** assembly (through noncovalent interactions) of a larger protein structure from 2 or more polypeptide chains (subunits), and the organization of these subunits.

26

Sanger Method

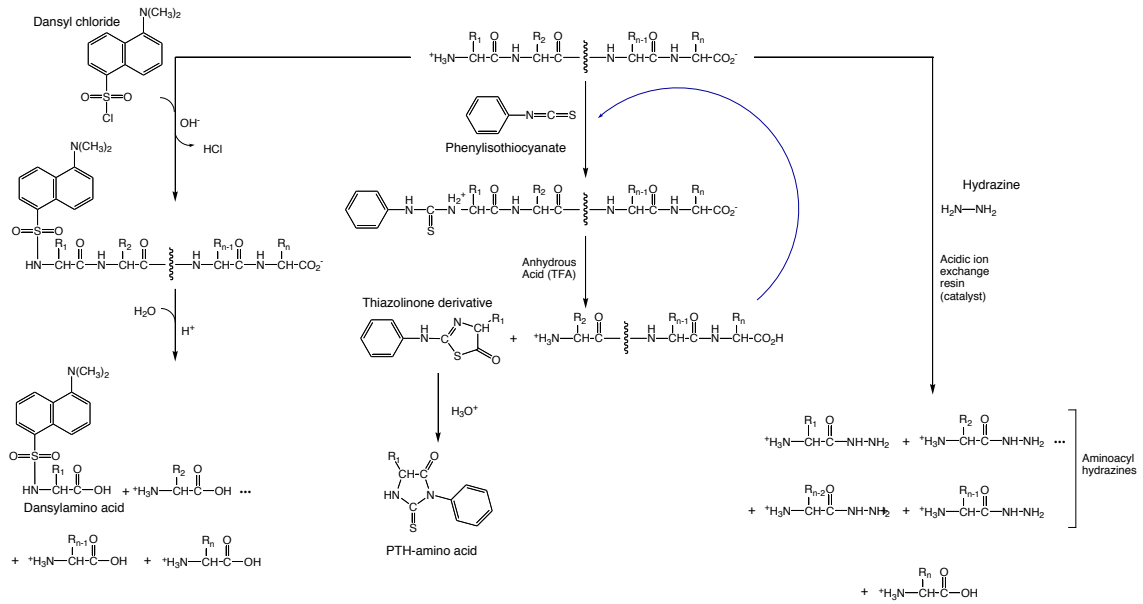
- The basic chemical method for protein sequence determination was developed by Sanger and involves three basic steps:
 1. Prepare the protein for sequencing
 - Determine the number of subunits
 - Cleave disulfide bonds
 - Purify individual subunits
 - Determine the amino acid compositions
 2. Sequence the polypeptide chains
 - Fragment the subunits
 - Separate and purify the fragments
 - Determine the amino acid sequence of each fragment
 - Repeat the above steps with a different fragmentation process
 3. Organize the completed structure
 - Use overlapping peptides to align the fragments
 - Elucidate the positions of disulfide bonds

27

End Group Analysis

- N-terminus identification:
 - 1-Dimethyl-amino-naphthalene-5-sulfonyl chloride (dansyl chloride) reacts with primary amines. Requires acid hydrolysis of peptide. Adducts detectable by fluorescence.
 - Phenylisothiocyanate (PITC, Edman's reagent) reacts with N-terminal amine to form phenylthiocarbamyl (PTC) adducts. PTC adduct can be cleaved from peptide by forming the thiazolinine using anhydrous strong acid.
- C-terminus identification:
 - No chemical method comparable to Edman degradation.
 - Exopeptidases-->carboxypeptidases can be used. Generally demonstrate amino acid specificity. Usually only useful for identifying the first few amino acid residues.
 - Hydrazinolysis: the peptide is treated with hydrazine and heated at 90°C. Treatment with mildly acidic ion exchange resin results in cleavage of peptide bonds producing aminoacyl hydrazides of all the amino acid residues except the C-terminal amino acid.

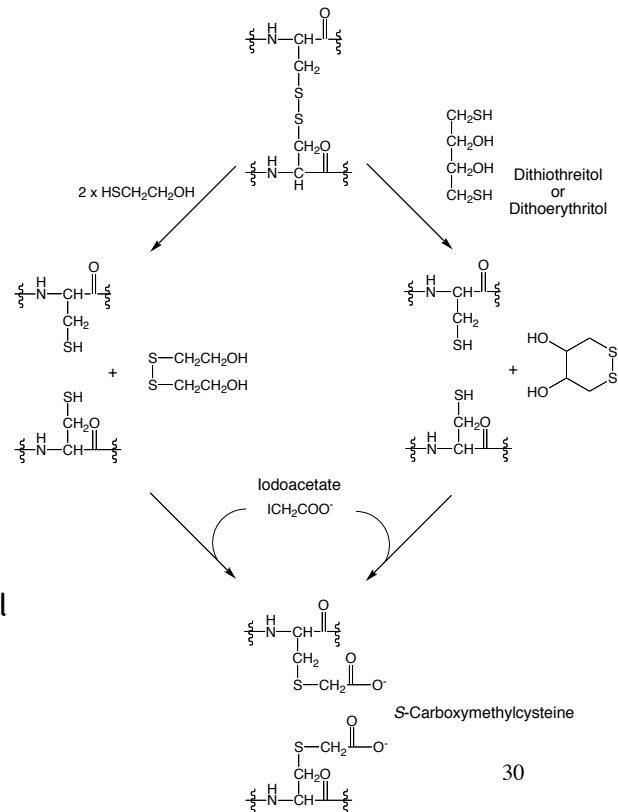
28



29

Reduction of Disulfide Bonds

- Required for separation of disulfide linked polypeptide strands.
- Prevent disulfide stabilized structures from interfering with efficient proteolysis.
- Disulfides usually cleaved through reduction.
 - 2-mercaptoethanol
 - Dithioerythritol/Dithiothreitol
- Cap free thiols with alkylating agent
 - Iodoacetic acid



30

Amino Acid Composition

- Amino acid composition is determined through the complete hydrolysis of the peptide and quantitative analysis of the liberated amino acids.
- Usually achieved through acid hydrolysis using 6M HCl heated at 100–120°C.
 - Involves long hydrolysis times (10–100h)
 - Ser, Thr and Tyr are partially degraded, and the side chain amides of Gln and Asn are hydrolyzed in the process.
 - The process destroys Trp residues.
- Base-catalyzed hydrolysis involves the use of 2–4M NaOH and heating to 100°C.
 - Cys, Ser, Thr and Arg residues decompose under these conditions.
 - Results in partial deamination and racemization of amino acids.
- Complete enzymatic proteolysis requires the use of combinations of proteases/peptidases.
 - Usually employed to quantify Trp, Asn and Gln.

31

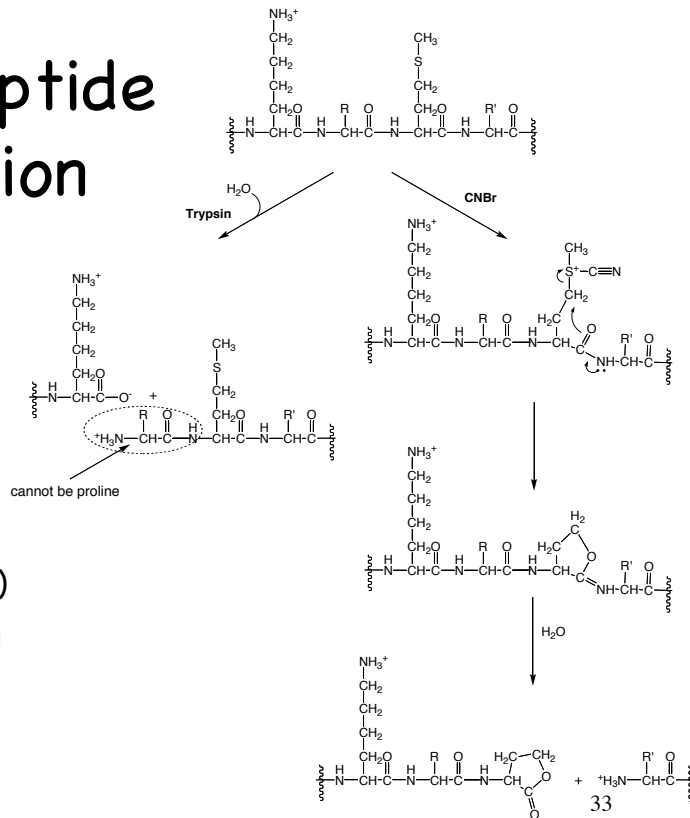
Amino Acid Composition II

- Leu, Ala, Gly, Ser, Val and Glu are the most abundant in proteins (>6%)
- His, Met, Cys and Trp are the least abundant (<3%)
- Ratio of amino acids with polar to nonpolar side chains is >1.
- This tends to decrease with protein size. Reflects increases in volume relative to surface area.
- Folded proteins generally have a hydrophobic “core” and a hydrophilic exterior.

32

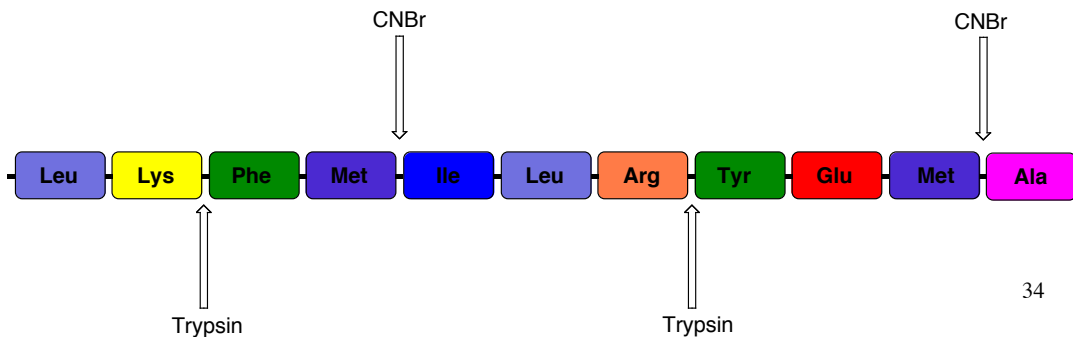
Controlled Peptide Fragmentation

- Endopeptidases:
 - Trypsin--basic residues (Lys, Arg)
 - Other endopeptidases show broader specificities.
- Chemical hydrolysis
 - Cyanogen bromide (Met)
- Peptide fragments can usually be resolved by reversed-phase chromatography -> HPLC.



Sequence Determination

- Sequencing of peptides usually done through repeated cycles of Edman degradation.
- Peptide is usually adsorbed to solid support -- PVDF or glass filter paper impregnated with polybrene.
- Limited to ~40-60 N-terminal amino acid residues.
- Sequences of larger polypeptides can be determined using overlapping peptide fragments.



Characterization by Mass Spectrometry

- Mass spectrometry (MS) is used to accurately measure the mass to charge (m/z) ratio for ions (in the gas phase).
 - **Electrospray ionization (ESI)**: Yields macromolecular ions (+0.5 to +2 per kD). Protonation of basic groups [(M+nH)ⁿ⁺ ions].
 - **Matrix-assisted laser desorption/ionization (MALDI)**: utilizes intense short laser pulses at λ absorbed by matrix material. Intact macromolecules ejected into gas phase usually with charge of +1 (+2 or +3 not uncommon). Can characterize proteins ≥ 300 kD.
 - **Fast atom bombardment (FAB)**: Macromolecules are ejected from low-volatility solvent using a low energy beam of Ar or Xe atoms or Cs⁺ ions. Macromolecules are ejected with charge of +1. Practical limit of ~ 7 kD.
- Can measure m/z with accuracy of $>0.01\%$

35

Peptide Sequencing by MS

- MS can be used to sequence short peptides (<25 residues).
- Uses a tandem mass spectrometer (MS/MS).
- Chemically inert atoms are used to fragment selected peptide ions.
- Masses of fragments can then be determined.
- Analysis of the mass spectra of families of fragments make it possible to assign amino acid sequence of fragments and the full length polypeptide.
 - Cannot differentiate between Ile and Leu.
 - Similarly Gln and Lys can be difficult to differentiate.

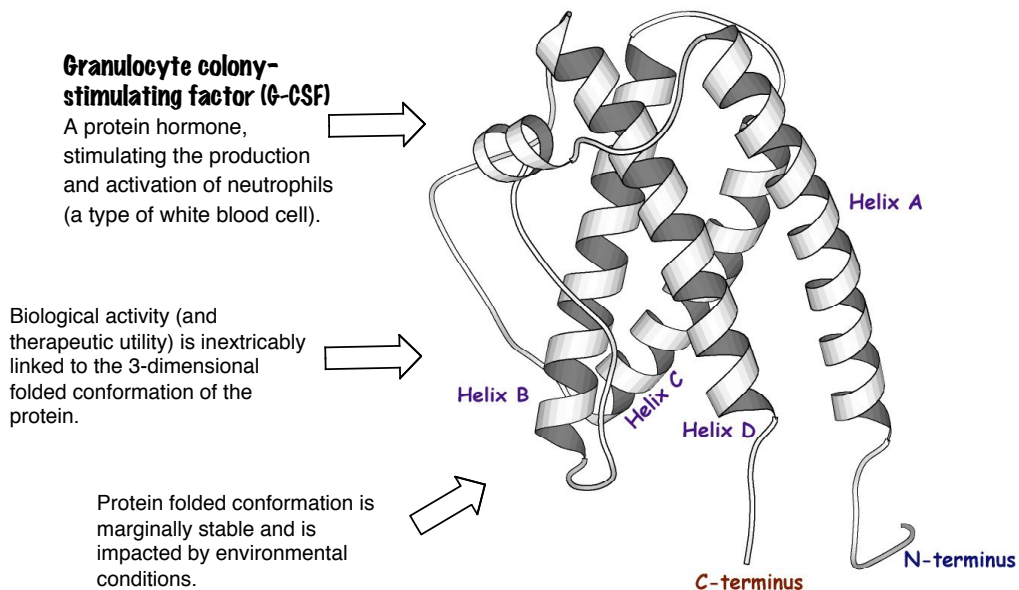
36

Peptide Mapping

- Used to facilitate the determination of proteins similar/related to proteins with known sequence.
- Begins with controlled fragmentation of the known and unknown proteins.
- Comparison of fragments by PAGE or HPLC should indicated fragments in the unknown protein that differ from that of the known protein.
- The variant fragments can be isolated and sequenced.

37

Three Dimensional Structures of Proteins



39

Protein 3-D Structure

- The 3-D structure of a protein is determined by the amino acid sequence.
- A protein has a unique, or nearly unique, structure.
- Protein function depends on structure.
- Although multiple structures are theoretically possible, proteins adopt unique structures.
- Noncovalent interactions stabilize protein structure.
- Protein structure is dynamic in nature.
- Structures of proteins can be elucidated by means of crystallography or by NMR.

40

Levels of Organization

- **Primary structure (1° structure):** the amino acid sequence of polypeptide chain.
- **Secondary structure (2° structure):** local spatial organization and arrangement of the peptide backbone. Generally refers to easily localized structural elements (i.e. helices and sheets).
- **Tertiary structure (3° structure):** the comprehensive three-dimensional structure of a protein.
- **Quaternary structure (4° structure):** assembly (through noncovalent interactions) of a larger protein structure from 2 or more polypeptide chains (subunits), and the organization of these subunits.

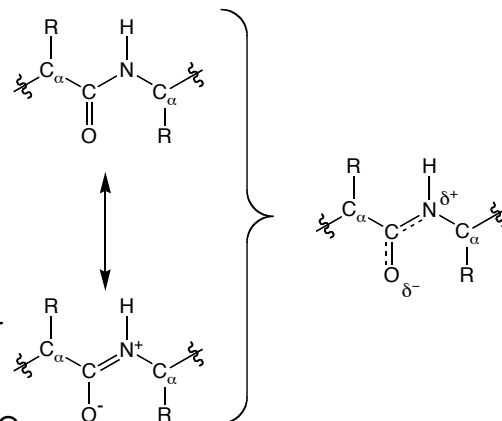
41

Secondary Structure and the Peptide Backbone

- Secondary structure (2° structure): the local conformation of the peptide backbone.
- Regular backbone folding patterns:

- Helices
- Pleated sheets
- Turns

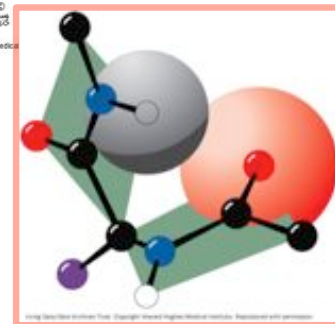
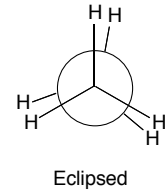
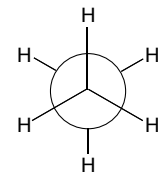
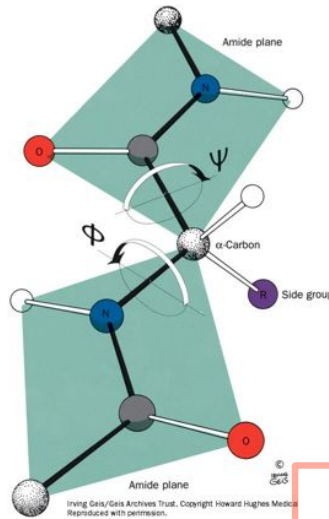
- Peptide group has rigid planar structure.
- Peptide bond has ~40% double bond character due to resonance.
- Peptide bonds are usually in trans conformation.



42

Peptide Conformation and Torsion Angles

- Peptide backbone is a linked sequence of nearly planar peptide groups.
- $\phi = C_{\alpha}-N$
- $\psi = C_{\alpha}-C_{\text{carbonyl}}$
- Steric constraints associated with ϕ and ψ angles.
- Staggered vs eclipsed conformations in ethane.
 - Eclipsed conformation is $12\text{kJ}\cdot\text{mol}^{-1}$ less stable than is the staggered conformation (an energy barrier).
 - Substituents other than H result in even greater steric interference to free rotation.
- Some conformations may become sterically forbidden.



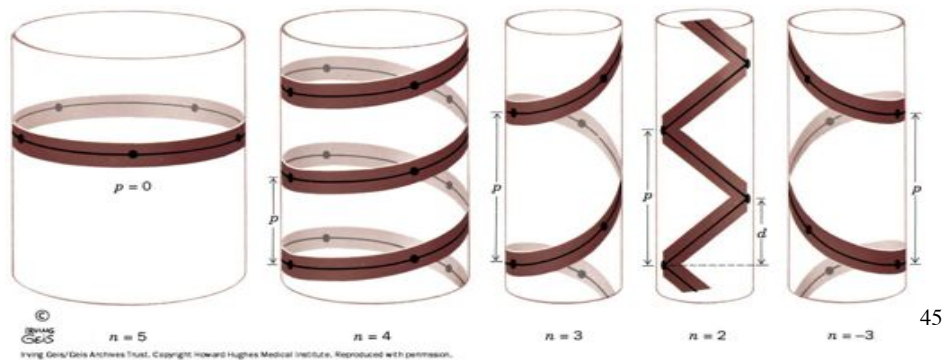
Secondary Structure

- There are three common secondary structural elements:
 - Helices (α -helix)
 - Sheets (β -sheet)
 - Turns (β -turn)
 - Loops (Ω -loop)

Helical Structures

- Helices result when a polypeptide chain is twisted the same extent about each of its C_α 's (ϕ and ψ).
- A helix can be characterized by the number of residues, n , required to make a complete turn and by its pitch, which is the distance a helix rises with each turn.
- Helices have an inherent chirality- can be either right- or left-handed.
- Backbone H-bonding contributes to stability of the helices and other 2° structure elements.

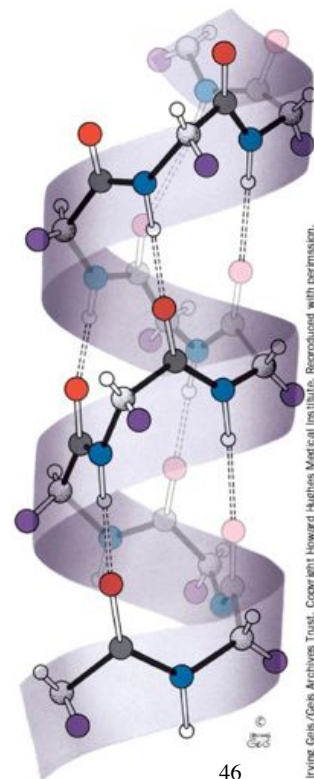
Figure 8-10 Examples of helices.



45

The α -helix

- Right handed helix combining allowed conformational angles and favorable backbone H-bonding.
 - $\phi = -57^\circ$
 - $\psi = -47^\circ$
 - $n = 3.6$ residues/turn
 - pitch = 5.4 \AA
- H-bonds between N-H group (donor) of n th residue and the C=O group (acceptor) of the $n-4$ th residue.
- Core of the helix is tightly packed-van der Waals contacts.
- The side chains are directed outward and backward.
- Present in both fibrous and globular proteins.



46

Sequence and Helix Stability

- The stability of the α -helix is affected by the amino acid sequence.
 - A **stretch or clusters of charged residues** can be destabilizing.
 - Amino acid residues with **bulky side chains** can be destabilizing.
 - Appropriately placed charged residues can form **stabilizing ion pairs**.
 - Appropriately placed **aromatic residues** can have favorable hydrophobic interactions.
 - **Proline** and **Glycine** are not common in α -helices.

47

Other Helical Conformations

- **Less prevalent than the α -helix.**
- Described using **n_m notation**:
 - n = number of residues per turn.
 - m = number of atoms (including H) in the ring formed by the backbone H-bond.
 - An α -helix would be a 3.6_{13} helix.
- The **3_{10} helix**: right handed helix with pitch = 6.0 Å (torsion angles slightly in forbidden range). Usually observed only for short segments.
- The **π -helix** (4.4_{16} helix): the wide-flat conformation results in the an axial hole. A mildly forbidden conformation. Only found in short segments (a few residues), within larger helices.
- **Polyproline(helices)**: left handed helix, with 3 residues per turn and a pitch of 9.4 Å.

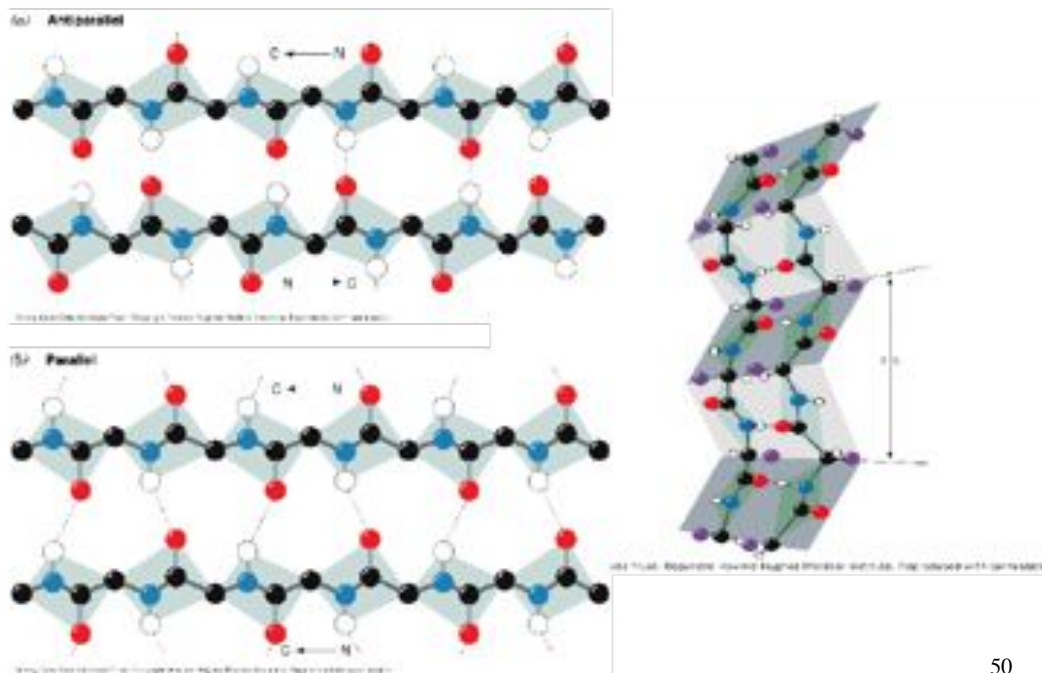
48

β Structures

- The **β -sheet** is characterized by peptide chains in extended conformations with a repeating pattern of ϕ and ψ angles.
- H-bonding between adjacent peptide strands (β -strands).
- Two varieties of β -sheet:
 - Antiparallel β pleated sheet -> H-bonded peptide strands run in opposite directions.
 - Parallel β pleated sheet -> H-bonded peptide strands run in the same direction.
 - (term “pleated” refers to the fact that β -sheets have a rippled or pleated appearance when viewed from edge.)
 - Mixed parallel-antiparallel β -sheets are common.
- In β -sheets, the side chains of successive amino acids are extend to opposite sides of the sheet.

49

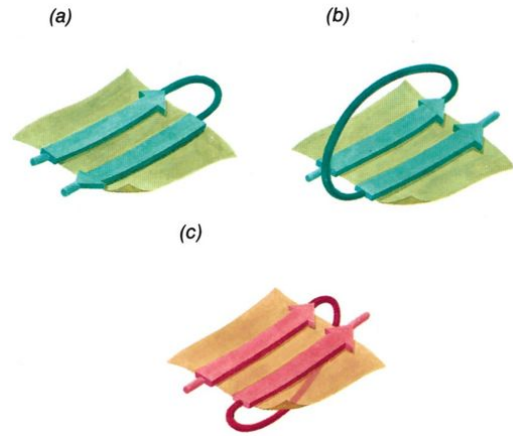
β -sheet Configurations



50

β-sheet Topology

- The β pleated sheets in globular proteins exhibit a pronounced right-handed twist and often form the central cores of proteins.
- The connections between strands may be a simple hairpin turn or may be a more complex crossover either above or below the plane of the sheet.
- Connectivity of strands comprising a β-sheet can be very complex.
- Antiparallel strands may be linked by a simple hairpin turn.
- Parallel strands usually linked by a longer crossover connection that does not lie in the plane of the sheet.



51

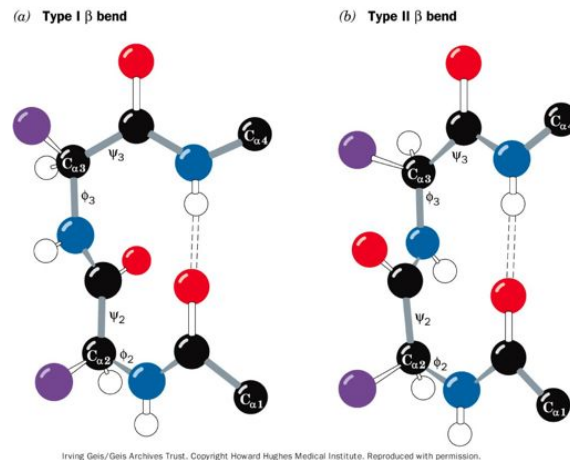
Structural Propensities

Amino Acid	P_{α}	Helix Classification	P_{β}	Sheet Classification
Ala	1.42	H_{α}	0.83	i_{β}
Arg	0.98	i_{α}	0.93	i_{β}
Asn	0.67	b_{α}	0.89	i_{β}
Asp	1.01	I_{α}	0.54	B_{β}
Cys	0.70	i_{α}	1.19	h_{β}
Gln	1.11	h_{α}	1.10	h_{β}
Glu	1.51	H_{α}	0.37	B_{β}
Gly	0.57	B_{α}	0.75	b_{β}
His	1.00	I_{α}	0.87	h_{β}
Ile	1.08	h_{α}	1.60	H_{β}
Leu	1.21	H_{α}	1.30	h_{β}
Lys	1.16	h_{α}	0.74	b_{β}
Met	1.45	H_{α}	1.05	h_{β}
Phe	1.13	h_{α}	1.38	h_{β}
Pro	0.57	B_{α}	0.55	B_{β}
Ser	0.77	i_{α}	0.75	b_{β}
Thr	0.83	i_{α}	1.19	h_{β}
Trp	1.08	h_{α}	1.37	h_{β}
Tyr	0.69	b_{α}	1.47	H_{β}
Val	1.06	h_{α}	1.70	H_{β}

52

Turns and Loops

- These are nonrepetitive structures.
- Reverse turns (β -bends, β -turns).
 - Usually involve 4 successive amino acids.
 - Two classes: Type I and Type II.
- The Ω loops:
 - Usually 6-16 residues.
 - End to end distance of $<10\text{\AA}$.
- Link successive helix or sheet structures



53

Tertiary Structures

- The tertiary structure describes the long-range interactions of amino acids.
- There are two major protein classes:
 - Fibrous
 - Globular
- Proteins can be subdivided into four classes:
 - All α -helix
 - All β -sheet
 - Mixed α/β proteins
 - Mixed $\alpha + \beta$ proteins

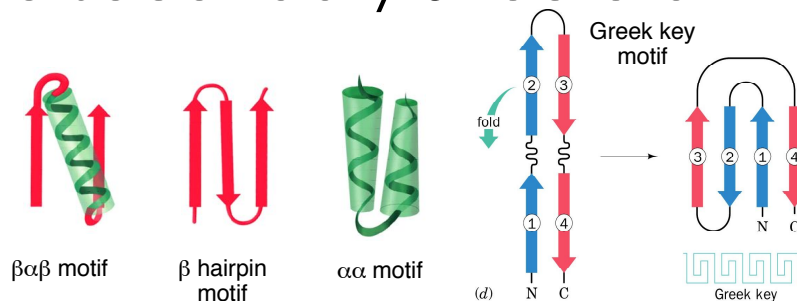
54

Supersecondary Structure

- Secondary structures are often grouped together into supersecondary structures or motifs that are used by many globular proteins.
- Groups of structural motifs combine to form the fold, 3° structure, of a protein.
- There are a finite number of unique folds. The number of unique folds is estimated to be ~1000. To date, only ~600 have been observed in nature.

55

Supersecondary Structural Motifs



- The **$\beta\alpha\beta$ motif** is the most common motif in proteins. An α -helical crossover connects two consecutive parallel strands of a β sheet.
- The **β hairpin** is another common motif. Here two sequential antiparallel strands of a β sheet are connected by a tight reverse turn.
- The **$\alpha\alpha$ motif** involves two successive α helices packing against each other with their axes incline so as to allow efficient hydrophobic packing. The helices are aligned antiparallel to each other.
- In the **Greek key motif**, a β hairpin folds in upon itself so as to form a 4-stranded antiparallel β sheet. Most common motif used to achieve this.

56

Fibrous vs. Globular Proteins

Fibrous Proteins

- Highly elongated molecules.
- Secondary structure is dominant structural motif.
- Relatively simple.
- Generally serve in structural capacities (protective, connective, supportive.)
- Structures not known in great detail.

Globular Proteins

- Large and highly diverse grouping of proteins.
- The 3° structure of globular proteins arises from the folding and organization of 2° structural elements.
- Serve in a variety of capacities.
- Despite vast structural diversity, globular proteins share some common structural features.

57

Globular Proteins

58

Spatial Distribution of Amino Acids

- Amino acid residues with nonpolar side chains (Val, Leu, Ile, Met and Phe) usually lie in the interior of a protein.
- Amino acid residues with charged polar side chains (Arg, His, Lys, Asp and Glu) usually lie at the protein surface.
- Amino acid residues with noncharged polar side chains (Ser, Thr, Asn, Gln, Tyr and Trp) usually lie on the surface of proteins, but frequently are in the interior. (When in the interior, they are usually involved in hydrogen bonding).
- Nearly all buried hydrogen bond donors form H-bonds with buried hydrogen bond acceptors.

59

Globular Protein Interiors

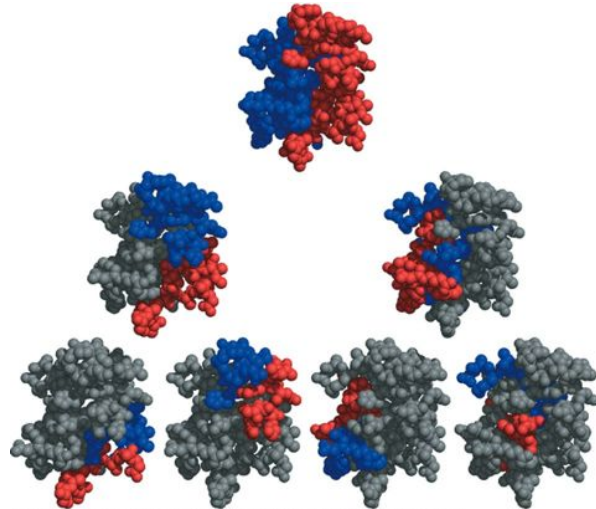
- Globular proteins are relatively compact.
- The interior region is tightly packed, with a packing density of ~ 0.75 .
- This packing density is comparable to that associated with molecular crystals of small organic molecules.
- Despite the tight packing, interior side chains adopt extended/low energy/relaxed conformations.

60

Hierarchal Organization

- **Proteins are hierarchically organized.**

- **Domains:** large protein subunits consisting of contiguous, compact and physically separable segments.
- **Subdomains** and **sub-subdomains:** smaller discreet structural subunits (combine to form domains and subdomains respectively).

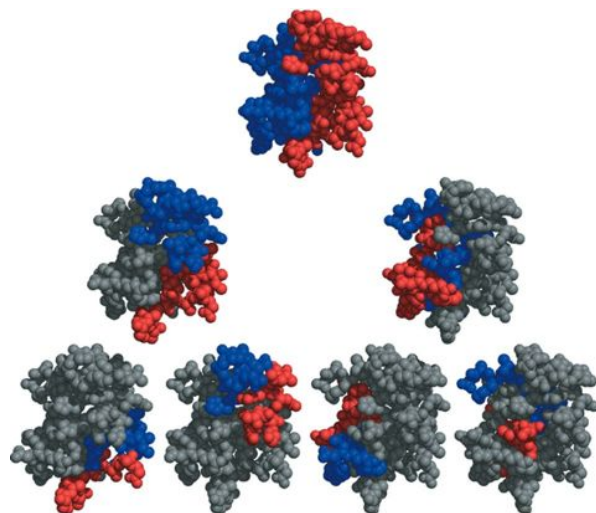


Courtesy of George Rose, The Johns Hopkins University of Medicine, and Robert Baldwin, Stanford University School of Medicine

61

Hierarchal Organization

- Consistent with the observation that most hydrogen bonding in proteins occurs locally.
- They are structurally independent and have characteristics of small globular proteins.
- Domains often have specific functions, such as binding of small molecules. These binding sites are often located near the domain clefts.



Courtesy of George Rose, The Johns Hopkins University of Medicine, and Robert Baldwin, Stanford University School of Medicine

62

Quaternary Structure

- Many proteins consist of more than one peptide chain; they consist of multiple subunits.
- A protein's quaternary structure consists of the spatial arrangement of these subunits.
- The subunit construction of many enzymes provides the structural basis for the regulation of their activities.
- Protomers in the majority of oligomeric proteins are arranged symmetrically.

63

Protein Stability

- Native proteins are only marginally stable under physiological conditions.
- The free energy of denaturation is only ~ 0.4 kJ/mol of amino acid, meaning a 100 residue protein is stable by ~ 40 kJ/mol. (The energy required to break a hydrogen bond is ~ 20 kJ/mol.)
- Protein structure and stability arises through a delicate balance of stabilizing and destabilizing forces.

64

Forces Stabilizing Macromolecular Structure

- Non-covalent interactions are key biological forces.
 - Electrostatic Forces
 - Ionic interactions (2.5Å, 20KJ•mol⁻¹)
 - Dipole Interactions
 - Van der Waals interactions (3-6Å, 0.4-4.0KJ•mol⁻¹)
 - Hydrogen Bonds (3Å, 12-30KJ•mol⁻¹)
 - Hydrophobic interactions (NA, <40KJ•mol⁻¹)
- These forces are transient in nature.
- Several factors influence the strength of these interactions.
- Individually all are weak (C-C bond -> 348KJ•mol⁻¹), together can be very strong.