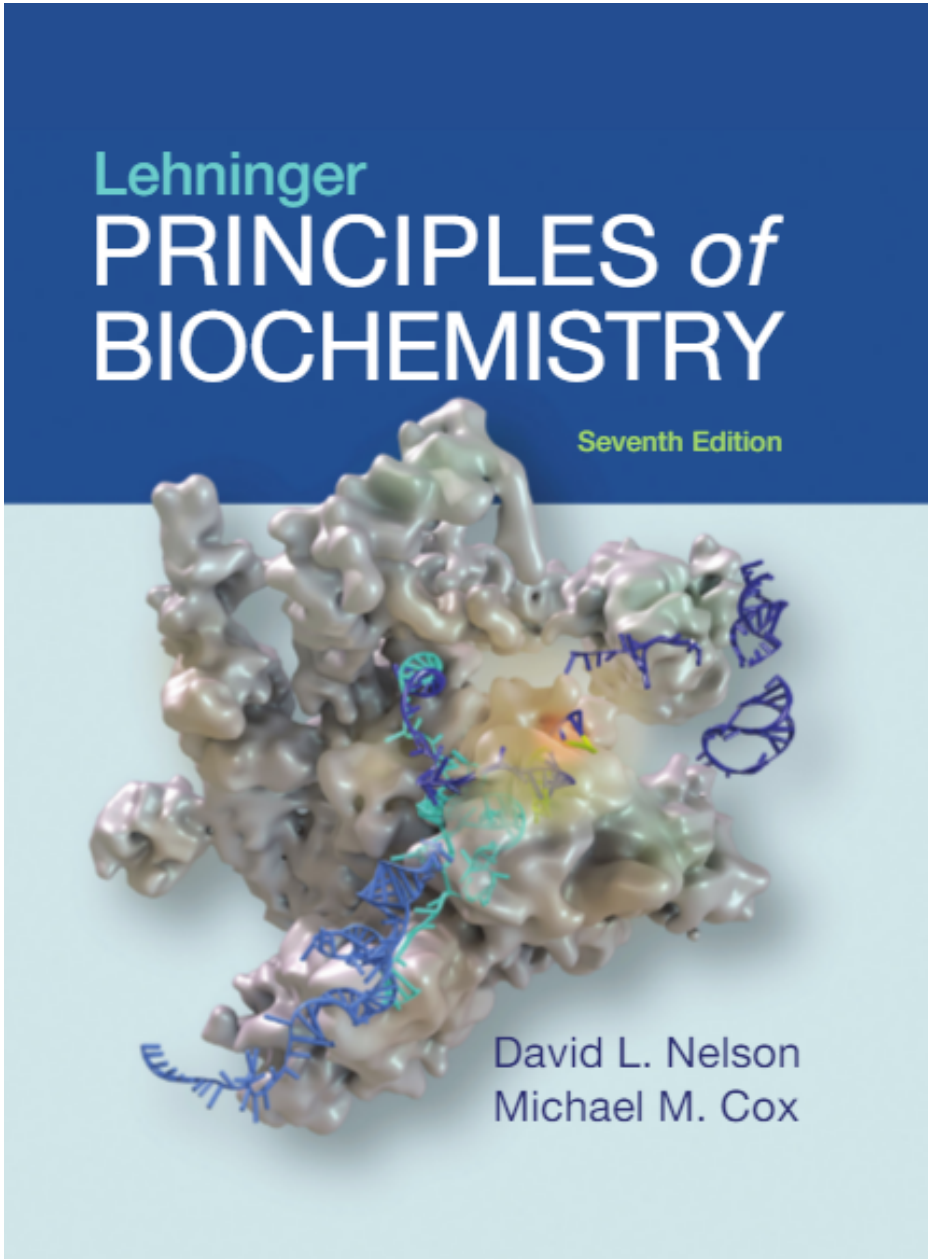
4 | The Three-Dimensional Structure of Proteins

© 2017 W. H. Freeman and Company



CHAPTER 4
The Three-Dimensional Structure of Proteins

Learning goals:

- Structure and properties of the peptide bond
- Structural hierarchy in proteins
- Structure and function of fibrous proteins
- Structure analysis of globular proteins
- Protein folding and denaturation

Structure of Proteins - General Aspects

1. aa sequence determines 3D structure

2. Protein function depends on its structure

3. An isolated protein usually exists in one or few stable structures

4. Noncovalent interactions are the most important forces stabilizing protein structure

5. There are common structural patterns in protein architecture

Structure of Proteins

- Unlike most organic polymers, protein molecules adopt a specific three-dimensional conformation.

✓ ***Conformation:*** *spatial arrangement of atoms in a protein (any structure that can exist without breaking covalent bonds)*

- This structure is able to fulfill a specific biological function
- This structure is called the native fold

✓ ***Native proteins:*** *proteins in any of their functional, folded conformations*

- The native fold has a large number of favorable interactions within the protein
- There is a cost in conformational entropy of folding the protein into one specific native fold

Favorable Interactions in Proteins

- **Hydrophobic effect**
  - Release of water molecules from the structured solvation layer around the molecule as protein folds increases the net entropy
- **Hydrogen bonds**
  - Interaction of N-H and C=O of the peptide bond leads to local regular structures such as α-helices and β-sheets
- **London dispersion**
  - Medium-range weak attraction between all atoms contributes significantly to the stability in the interior of the protein
- **Electrostatic interactions**
  - Long-range strong interactions between permanently charged groups
  - Salt-bridges, esp. buried in the hydrophobic environment strongly stabilize the protein

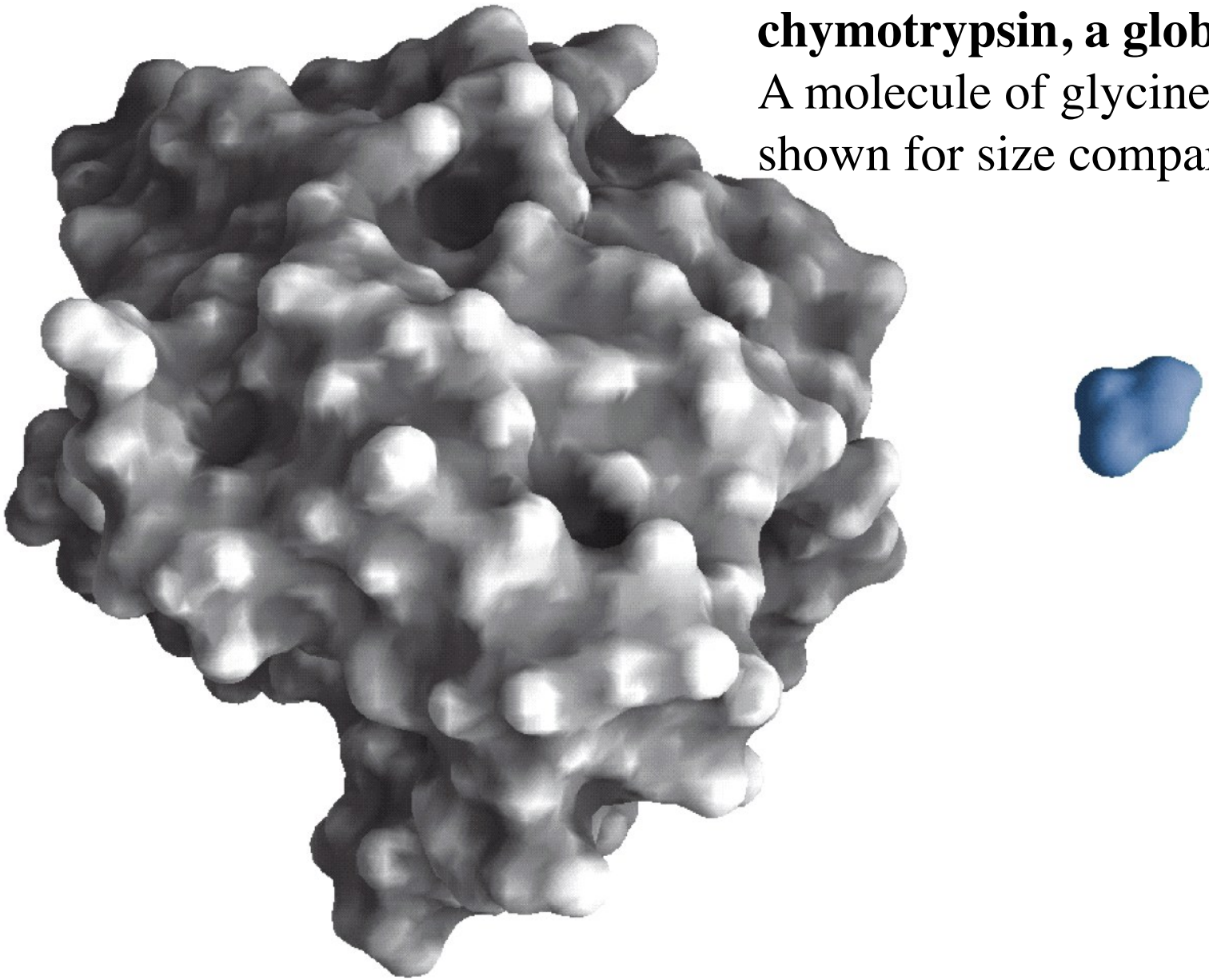
Weak Interactions Stabilize Conformation

- ***Stability:*** tendency to keep native conformation Release of water molecules from the structured solvation layer around the molecule as protein folds increases the net entropy
  - To break 1 covalent bond → 200 – 460 kJ/mol
  - To break a weak interaction (H-bond; hydrophobic interaction, ionic bond, etc.) → 4 – 30 kJ/mol
- **Weak interactions are responsible for native structure**
  - Breaking individual covalent bonds that contribute to native structure of proteins (S-S bonds) is more difficult than breaking 1 weak interaction
  - BUT because there are so many weak interactions → they are responsible for keeping native structure
- **Protein conformation with the lowest free energy (most stable) is the one with MAXIMUM number of weak interactions**

Weak Interactions Stabilize Conformation

- **When H_2O surrounds a hydrophobic molecule, a solvation layer of H_2O molecules is created**
  - increased order of H_2O → ↓ entropy
  - BUT when non polar groups cluster together → ↓ solvation layer → ↑ entropy → ↓ ΔG → more stable
- **Two rules:**

1) Hydrophobic residues are buried inside proteins
2) The number of H-bonds within a protein is maximized



Structure of the enzyme chymotrypsin, a globular protein. A molecule of glycine (blue) is shown for size comparison

Figure 4-1
Lehninger Principles of Biochemistry, Fifth Edition
© 2008 W. H. Freeman and Company

4 Levels of Protein Structure

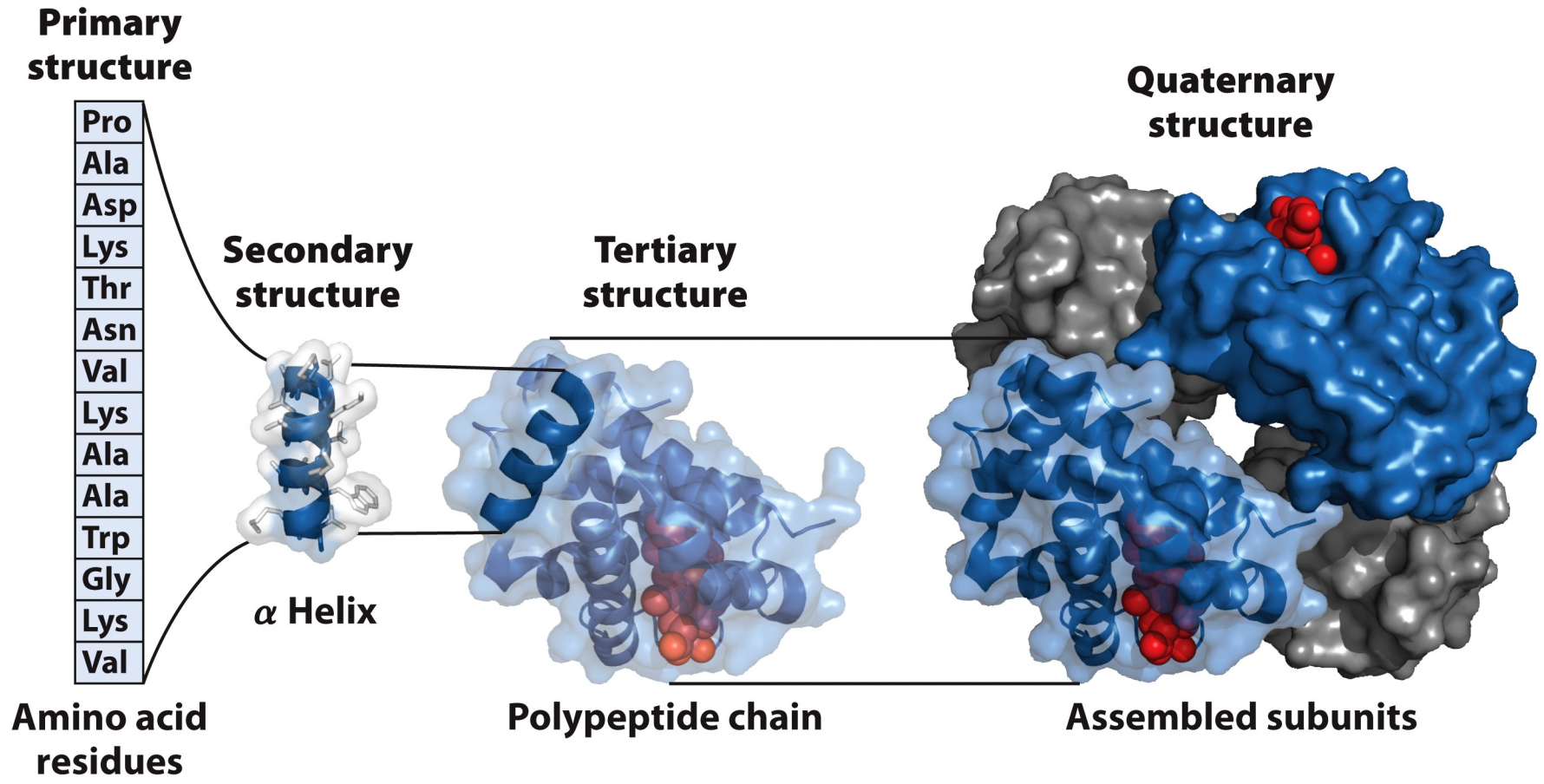
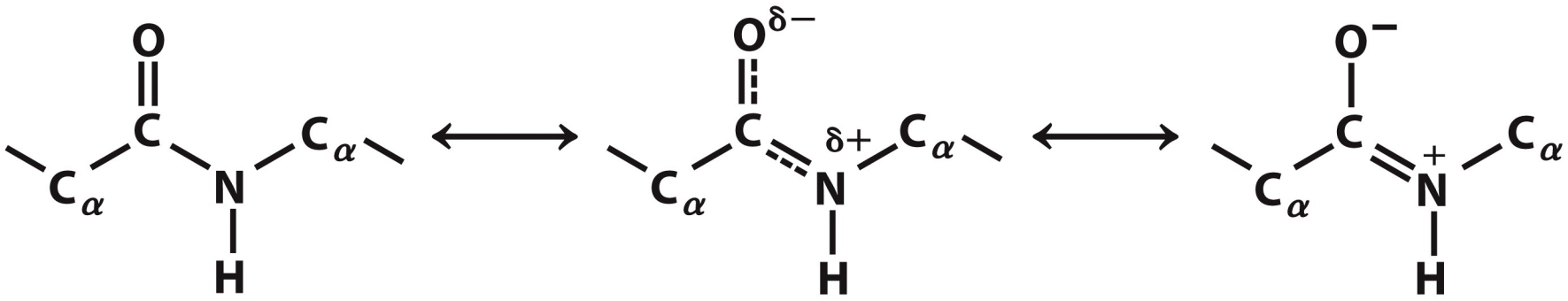


Figure 3-23
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Primary Structure: The Peptide Bond

- Structure of the protein is partially dictated by the properties of the peptide bond
- The peptide bond is a **resonance hybrid** of two structures
- The resonance (electron sharing) causes the peptide bonds
  - to be **less reactive** compared to esters, for example
  - C-N bond in a peptide bond is **shorter** than C-N bond is a simple amine
  - to be quite **rigid** and nearly **planar**
  - to have a **large dipole moment** in the favored **trans** configuration

Resonance in the Peptide Bond



The carbonyl oxygen has a partial negative charge and the amide nitrogen a partial positive charge, setting up a small electric dipole. Virtually all peptide bonds in proteins occur in this trans configuration; an exception is noted in Figure 4–7b.

Figure 4-2a
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

The Rigid Peptide Plane and the Partially Free Rotations

- Rotation around the peptide bond is ***not permitted*** due to resonance structure.
- Rotation around bonds connected to the α carbon is permitted.
  - ϕ (phi): angle around the α carbon—amide nitrogen bond
  - ψ (psi): angle around the α carbon—carbonyl carbon bond
- In a fully extended polypeptide, both ψ and ϕ are 180°

The organization around the peptide bond, paired with the identity of the R groups, determines the secondary structure of the protein.

The polypeptide is made up of a series of planes linked at α carbons

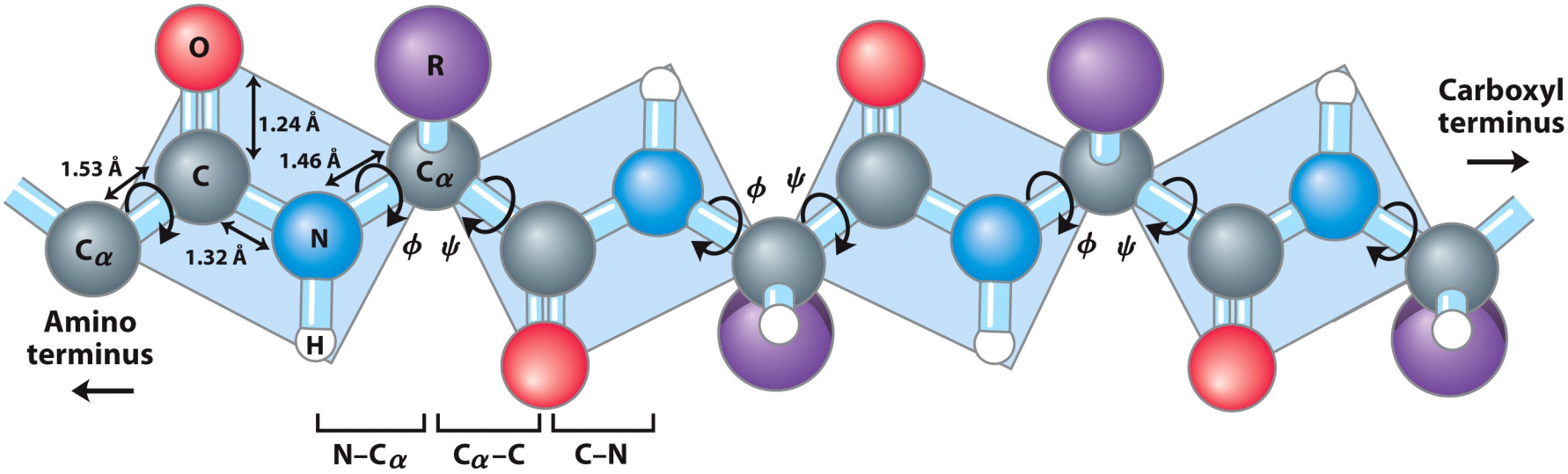


Figure 4-2b
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

- The 6 atoms of the peptide group are in 1 plane
- O atom of carbonyl and H atom of amide N are *trans*
- Due to partial double bond, peptide bond cannot rotate freely → limiting the range of conformation of a polypeptide

Distribution of ϕ and ψ Dihedral Angles

- Some ϕ and ψ combinations are very unfavorable because of **steric crowding** of backbone atoms with other atoms in the backbone or side chains
- Some ϕ and ψ combinations are more favorable because of chance to form favorable H-bonding interactions along the backbone

Secondary Structures

- 2^o structure refers to a local spatial arrangement of the polypeptide backbone (depends on common regular folding patterns of polypeptide backbone)
- Two regular arrangements are common:
- The **α helix**
  - stabilized by hydrogen bonds between nearby residues
- The **β sheet**
  - stabilized by hydrogen bonds between adjacent segments that may not be nearby
- Irregular arrangement of the polypeptide chain is called the random coil

TABLE 4.1 **Idealized ϕ and ψ Angles for Common Secondary Structures in Proteins**

Structure	ϕ	ψ
α Helix	−57°	−47°
β Conformation		
Antiparallel	−139°	+135°
Parallel	−119°	+113°
Collagen triple helix	−51°	+153°
β Turn type I		
i + 1[a]	−60°	−30°
i + 2[a]	−90°	0°
β Turn type II		
i + 1	−60°	+120°
i + 2	+80°	0°

The α Helix

- The **simplest arrangement** a polypeptide can assume (with its rigid peptide bonds and other freely-rotating single bonds) is a helical structure
  - In general, ~ ¼ of all aa residues in polypeptides are found in α helices
  - E.g. in α-keratins, α helix is a predominant structure

The α Helix

- Helical backbone is held together by hydrogen bonds between the backbone amides of an *n* and *n* + 3 amino acids.
- It is a right-handed helix with 3.6 residues (5.4 Å) per turn.
- Peptide bonds are aligned roughly parallel with the helical axis.
- Side chains point out and are roughly perpendicular with the helical axis.

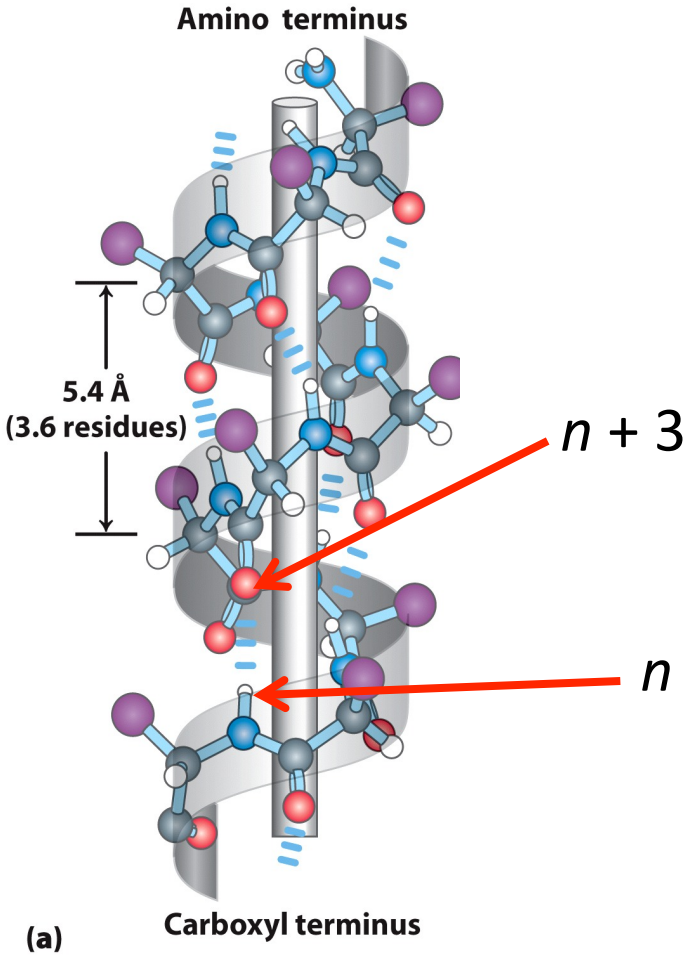
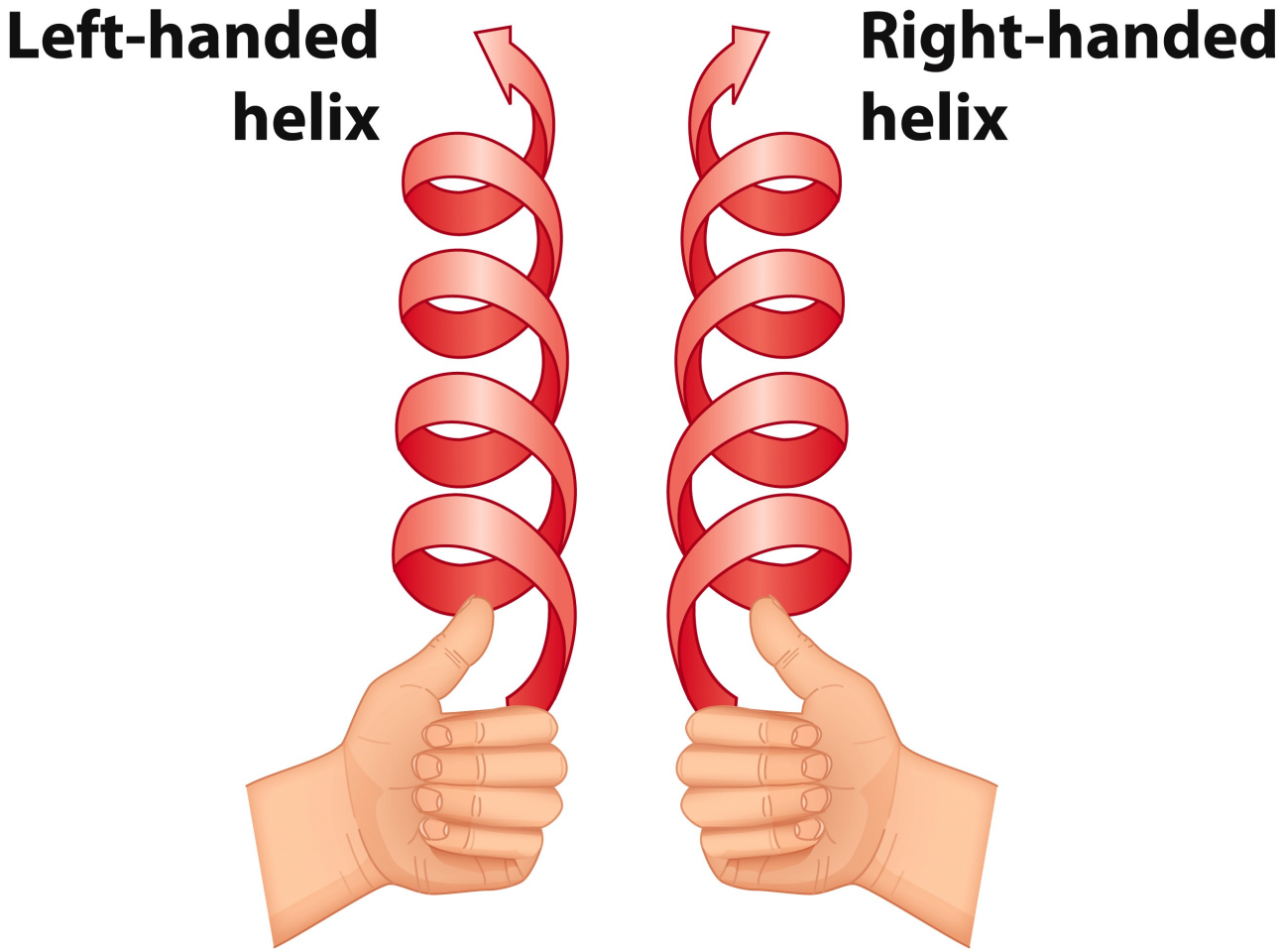


Figure 4-4
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

What is a right-handed helix?



Box 4-1
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

The α Helix: Top View

- The inner diameter of the helix (no side chains) is about 4–5 Å
  - Too small for anything to fit "inside"
- The outer diameter of the helix (with side chains) is 10–12 Å
  - Fits well into the **major groove of dsDNA**
- Residues 1 and 8 align nicely on top of each other
  - What kind of sequence gives an α helix with one hydrophobic face?

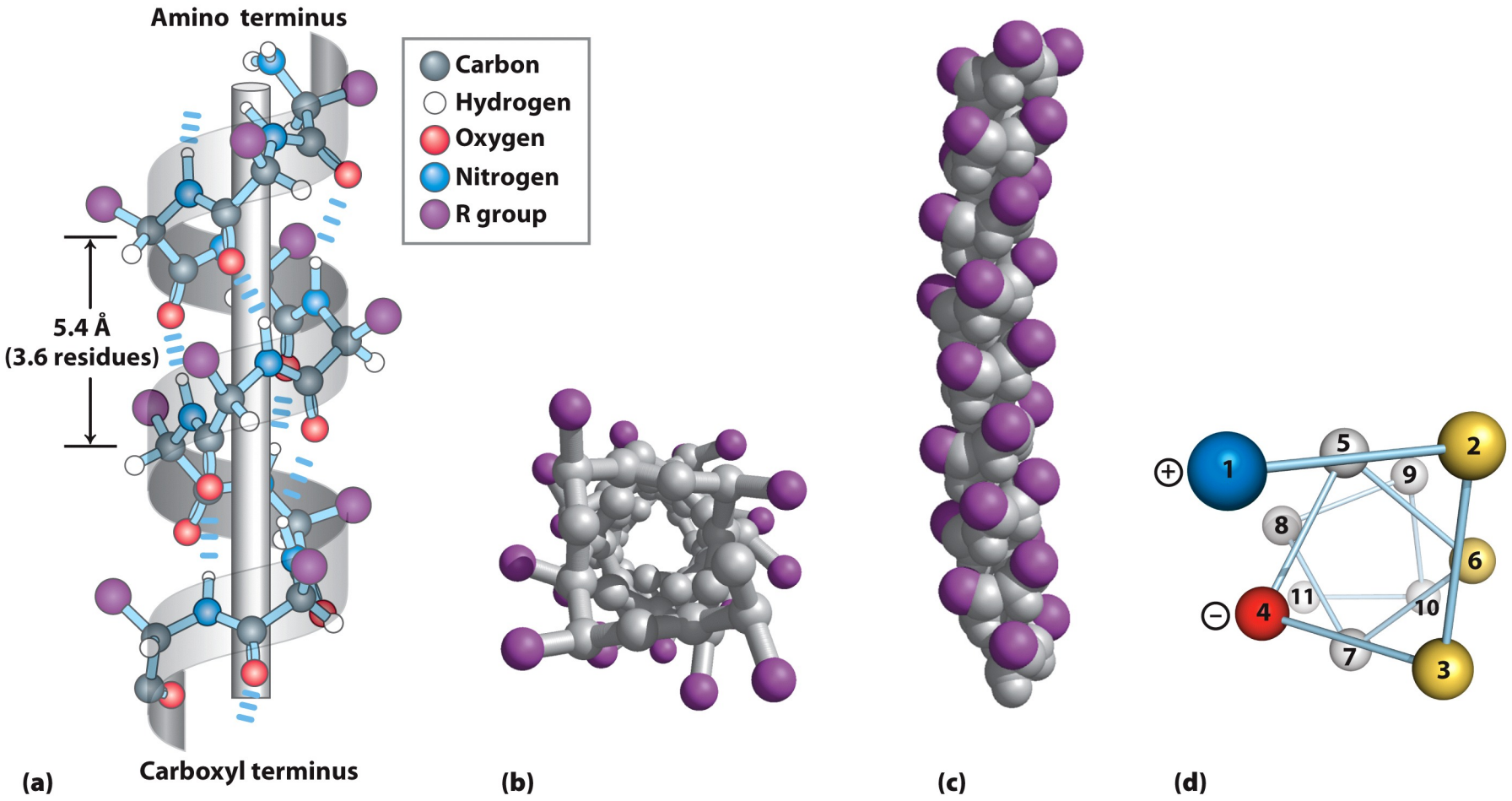


Figure 4-4
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Sequence affects helix stability

- Not all polypeptide sequences adopt α-helical structures
- Small hydrophobic residues such as Ala and Leu are strong helix formers
- Pro acts as a helix breaker because the rotation around the N-C_α bond is impossible
- Gly acts as a helix breaker because the tiny R-group supports other conformations
- Attractive or repulsive interactions between side chains 3–4 amino acids apart will affect formation

Constraints to helix stability

1. Electrostatic repulsion (or attraction) between successive charged aa

- Interactions between aa side chains can stabilize or destabilize a helix:

- E.g. long block of Glu (E) residues → no helix at pH 7 because of –ve charge repulsion

- Same for +vely charged aa

Constraints to helix stability

2. Bulkiness of adjacent R groups

- Near each other, these aa and other bulky aa will not fit well due to steric hindrance

- Small hydrophobic residues such as Ala and Leu are strong helix formers

Constraints to helix stability

3. Interactions between R groups 3 or 4 residues apart

- Due to the twist of α helix → critical interaction between aa side chain and side chain of aa 3 or 4 residues apart in both directions

- Often, +vely charged aa are 3 or 4 aa away from –ve charge → forming ion pairs

- Two hydrophobic side chains 3 or 4 aa apart → hydrophobic interactions

Constraints to helix stability

4. The presence of Pro or Gly

- Pro: rarely found in α helix

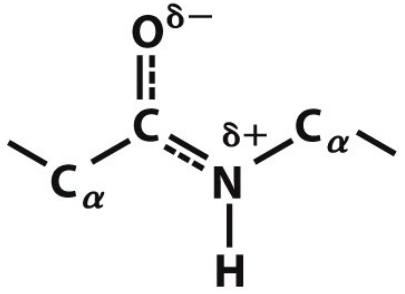
  1) rotation about N-Cα is not possible because N is part of a rigid ring

  2) N has no H to participate in H-bonds

- Gly: infrequently found in α helix
- it has more conformational flexibility than other aa
- makes a different coiled structures than α helix

Constraints to helix stability

5. Dipoles



- A small electric dipole exists in each peptide bond
  - Carbonyl O negative
  - Amide H positive
- Dipoles are connected by H-bonds → large net macroscopic dipole moment extending along helix (increases with helix length)
- The 4 aa at the ends of the helix do not participate fully in H-bonds → on N-terminus partial +ve charge and on C-terminus partial –ve charge
- Therefore, normally –vely charge aa are found near N-termini and vice versa (a +ve aa near N-terminus is destabilizing).

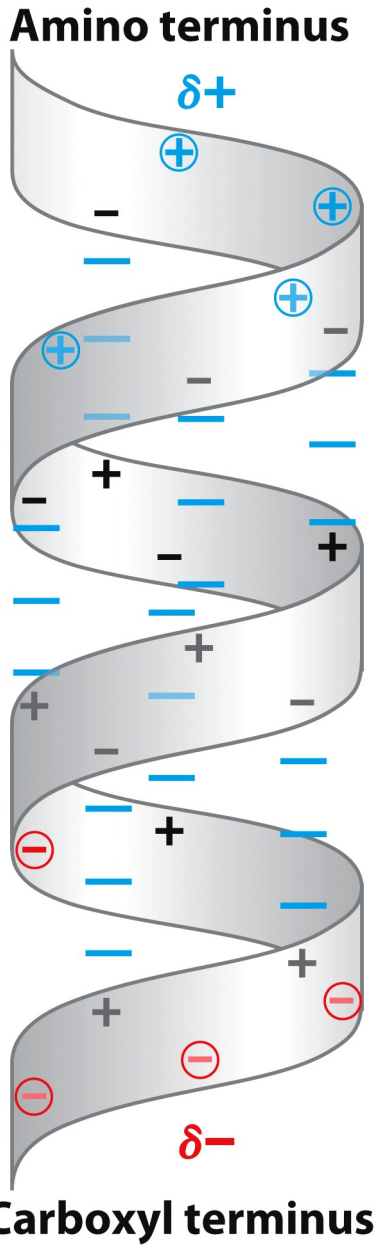


Figure 4-5
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

β Sheets

- The planarity of the peptide bond and tetrahedral geometry of the α-carbon create a pleated sheet-like structure
- Sheet-like arrangement of backbone is held together by hydrogen bonds between the backbone amides in different strands
- Side chains protrude from the sheet alternating in up and down direction

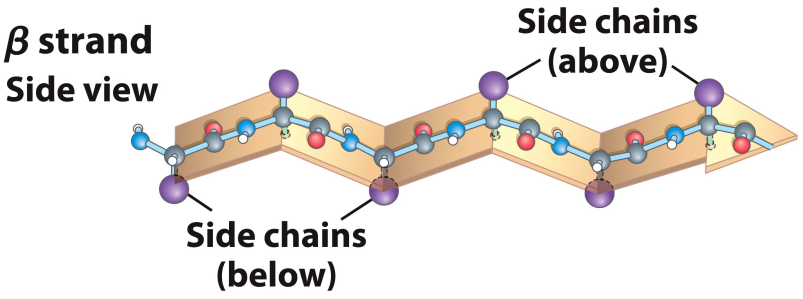


Figure 4-6a
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

β Sheets

- Backbone is extended into a zigzag
- Zigzag polypeptide chains are arranged side-by-side forming the β sheet
- H-bonds form between adjacent segments
- Segments are usually near each other (sometimes can be very far or even between different polypeptides)
- The R groups of adjacent aa come out opposite direction
- Size limits: nearby R groups must be small
- E.g. β-keratins (silk fibroin and fibroin of spider webs) have high content of Gly and Ala (smallest R groups)

Parallel and Antiparallel β Sheets

- Multi β-strand interactions are called **sheets**.
- Sheets are held together by the hydrogen bonding of amide and carbonyl groups of the peptide bond from opposite strands.
- Two major orientations of β sheets are determined by the directionality of the strands within:
  - **Parallel sheets** have strands that are oriented in the same direction.
  - **Antiparallel sheets** have strands that are oriented in opposite directions.

In parallel β sheets, the H-bonded strands run in the same direction.

- Hydrogen bonds between strands are bent (weaker).

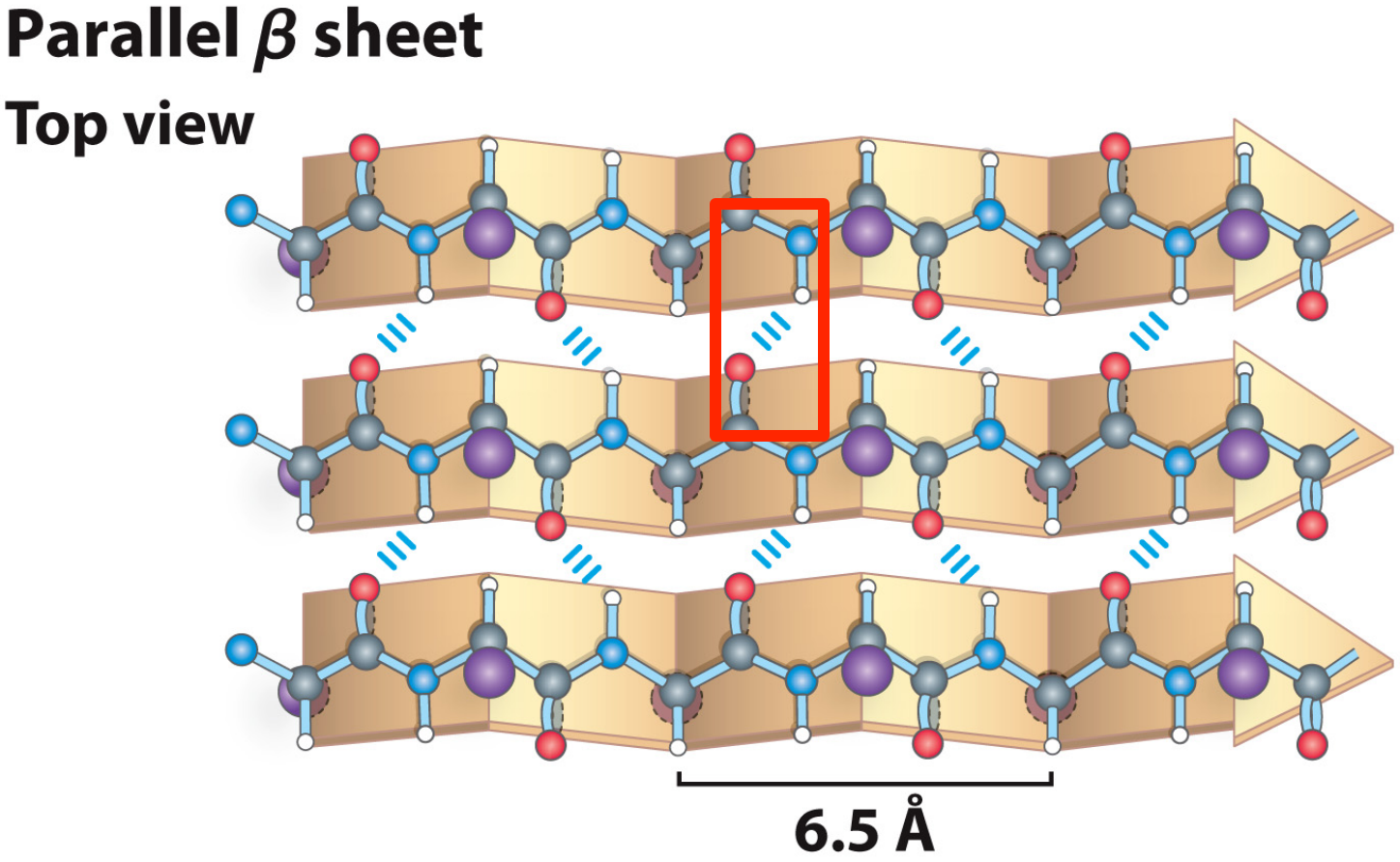


Figure 4-6c
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

In antiparallel β sheets, the H-bonded strands run in opposite directions.

- Hydrogen bonds between strands are linear (stronger).

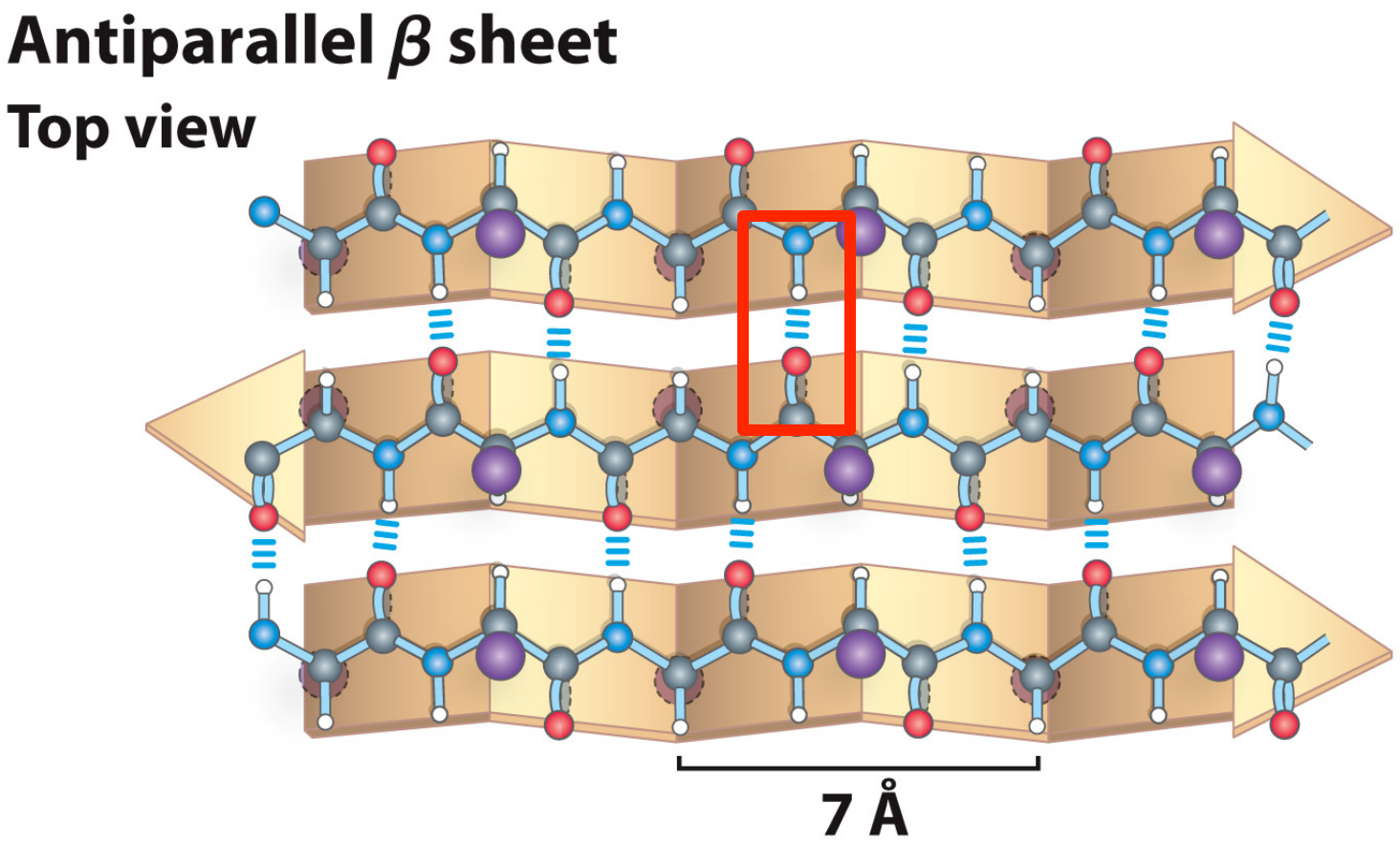


Figure 4-6b
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

β Turns

- Globular proteins have compact folded structures
- ~ $^1/_3$ aa are turns or loops (polypeptide reverses direction)
- β turns occur frequently whenever strands in β sheets change the direction
- Connecting elements linking successive runs of α helix or β sheet
- The 180° turn is accomplished over four amino acids
- The turn is stabilized by a hydrogen bond from a carbonyl oxygen to amide proton three residues down the sequence (aa1 and aa4)
- Proline in position 2 or glycine in position 3 are common in β turns

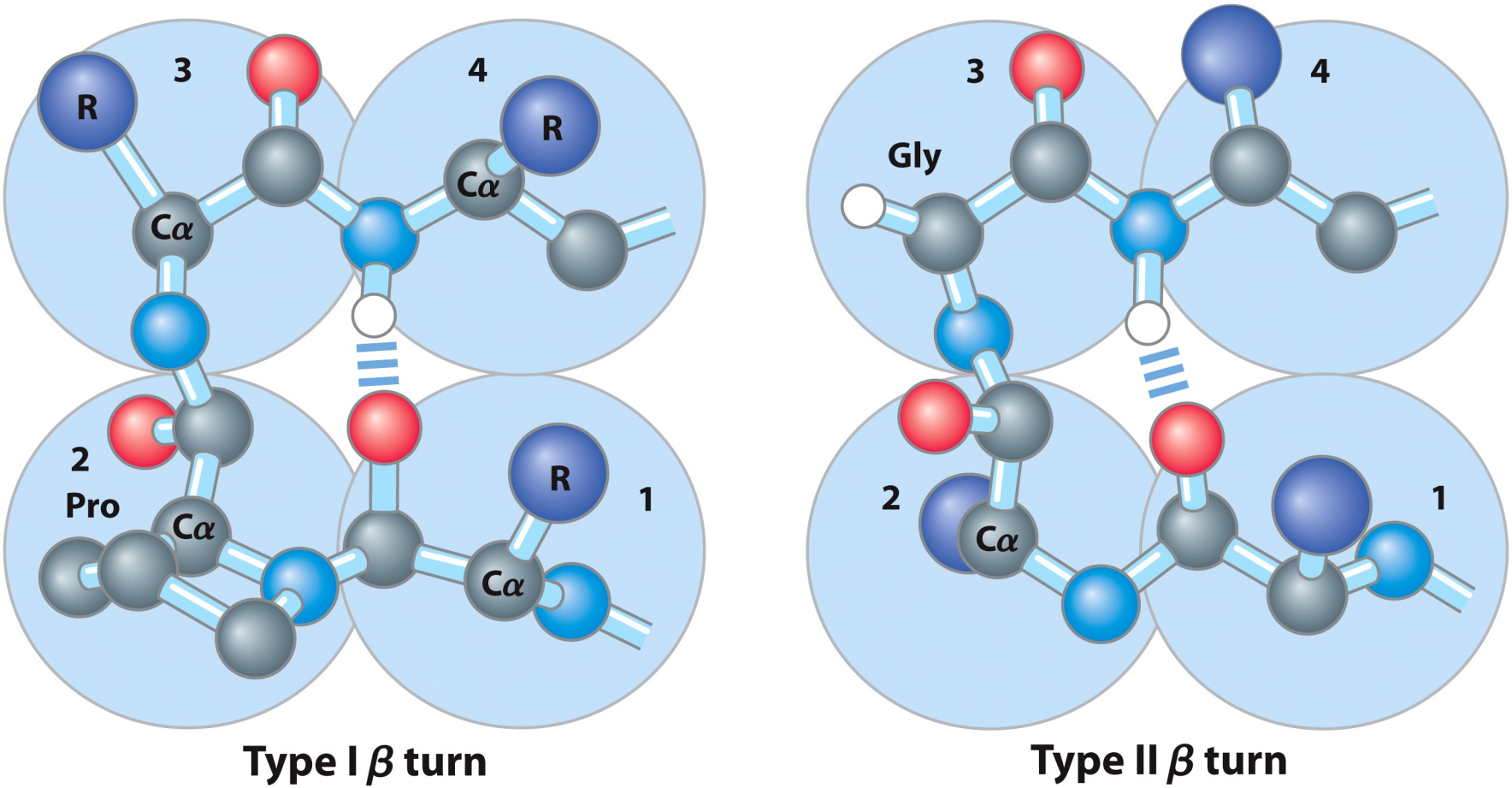
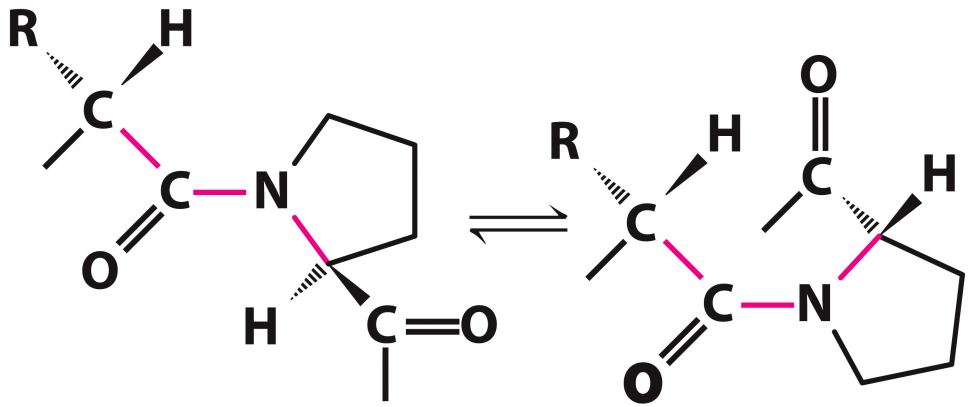


Figure 4-7
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Proline Isomers

- Most peptide bonds not involving proline are in the trans configuration (>99.95%)
- For peptide bonds involving proline, about 6% are in the cis configuration. Most of this 6% involve β-turns
- Proline isomerization is catalyzed by **proline isomerases**

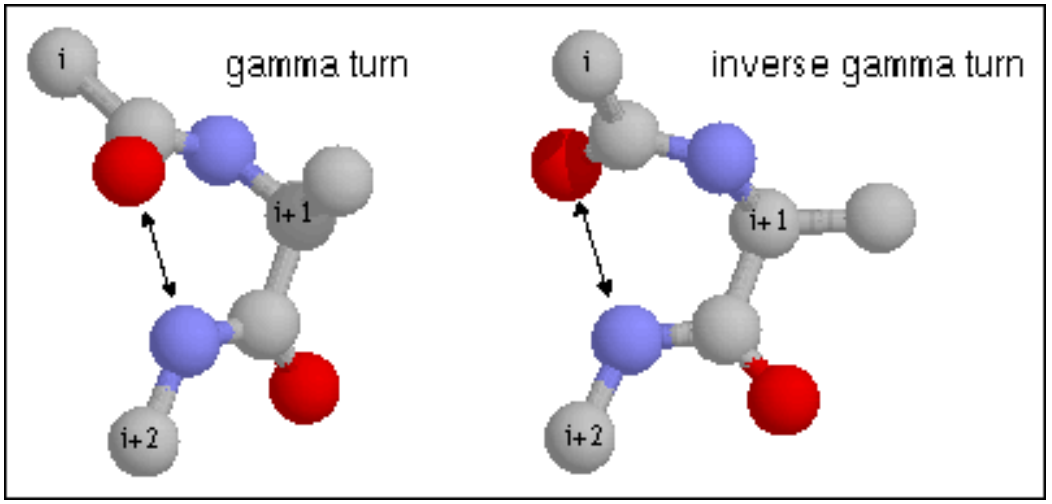


Proline isomers

Figure 4-8
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

γ turns

- Less common
- 3 aa turn
- H-bonding between 1st and 3rd aa



Circular Dichroism (CD) Analysis

- A technique used to determine the secondary structure fold of purified proteins
- CD measures the molar absorption difference $\Delta\varepsilon$ of left- and right-circularly polarized light: $\Delta\varepsilon = \varepsilon_L - \varepsilon_R$
- Chromophores (mainly aromatic aa) in the chiral environment produce characteristic signals
- CD signals from peptide bonds depend on the chain conformation
- Characteristic spectra for different secondary structures
- Combined spectra give the percentage of α to β in a protein

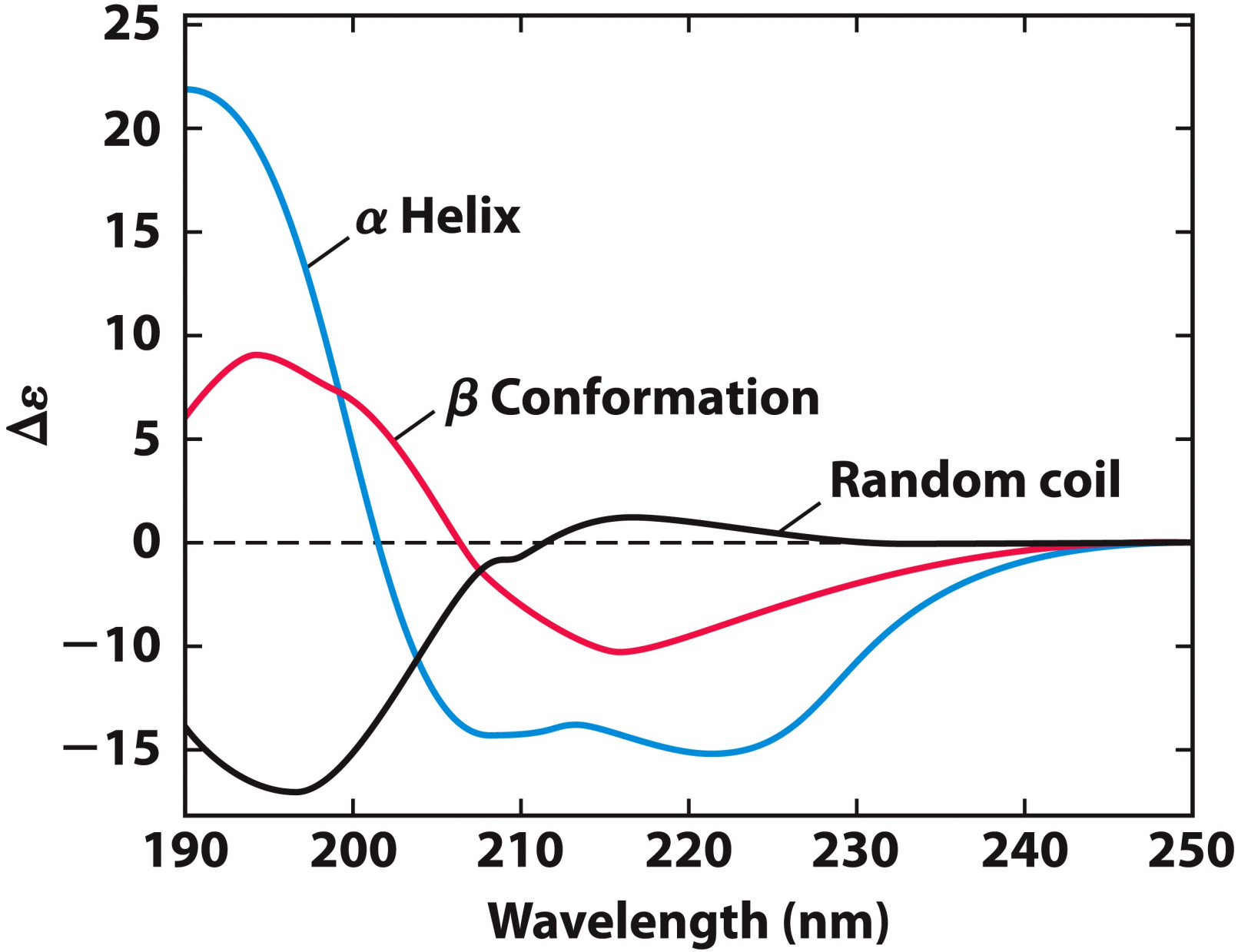


Figure 4-10
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Protein Tertiary Structure

- Tertiary structure refers to the overall spatial arrangement of atoms in a protein

- Stabilized by numerous weak interactions between amino acid side chains.
  - Largely hydrophobic and polar interactions
  - Can be stabilized by disulfide bonds

- Interacting amino acids are not necessarily next to each other in the primary sequence.

- Two major classes
  - **Fibrous** and **globular**

Protein Tertiary Structure

- **Fibrous** proteins – polypeptide chain arranged in long strands or sheets
  – consist largely of 1 type of 2^o structure
  – needed for structure and shape of cells
  – water insoluble (hydrophobic)

- **Globular** proteins – polypeptide chain folded into spherical shape
  – many types of 2^o structures
  – enzymes and regulators
  – largely water soluble (hydrophilic)

Fibrous Proteins: From Structure to Function

Function	Structure	Example
Tough, rigid, hard (nails, horns)	Cross-linked α-helixes Rigid linker (S—S)	α-keratin
Tensile strength, non-stretching (tendons, cartilage)	Cross-linked triple-helixes Flexible linker (Lys-HyLys)	Collagen
Soft, flexible non-stretchy (egg sac, nest, web)	Non-covalently held β-sheets van der Waals interaction	Silk fibroin

Fibrous Proteins

- Have properties that give them strength and flexibility
- All fibrous proteins are water insoluble (high concentration of hydrophobic aa inside and on protein surface)
- **α-keratin:** ~ all dry weight of hair, wool, nail, horns, hooves, etc.
- Part of a family called intermediate filaments (IF)
- Coiled coil, parallel (N-termini at same side)
- Supertwists are left-handed
- At the surface of a helix where supertwists exist there are many hydrophobic aa

Structure of α-Keratin in Hair

- **Simple 3º structure:** dominated by a-helical 2º structure with its helical axis twisted in a left-handed helix
- **Complex 4º structure:** many coiled coils assembled into large supramolecular complexes

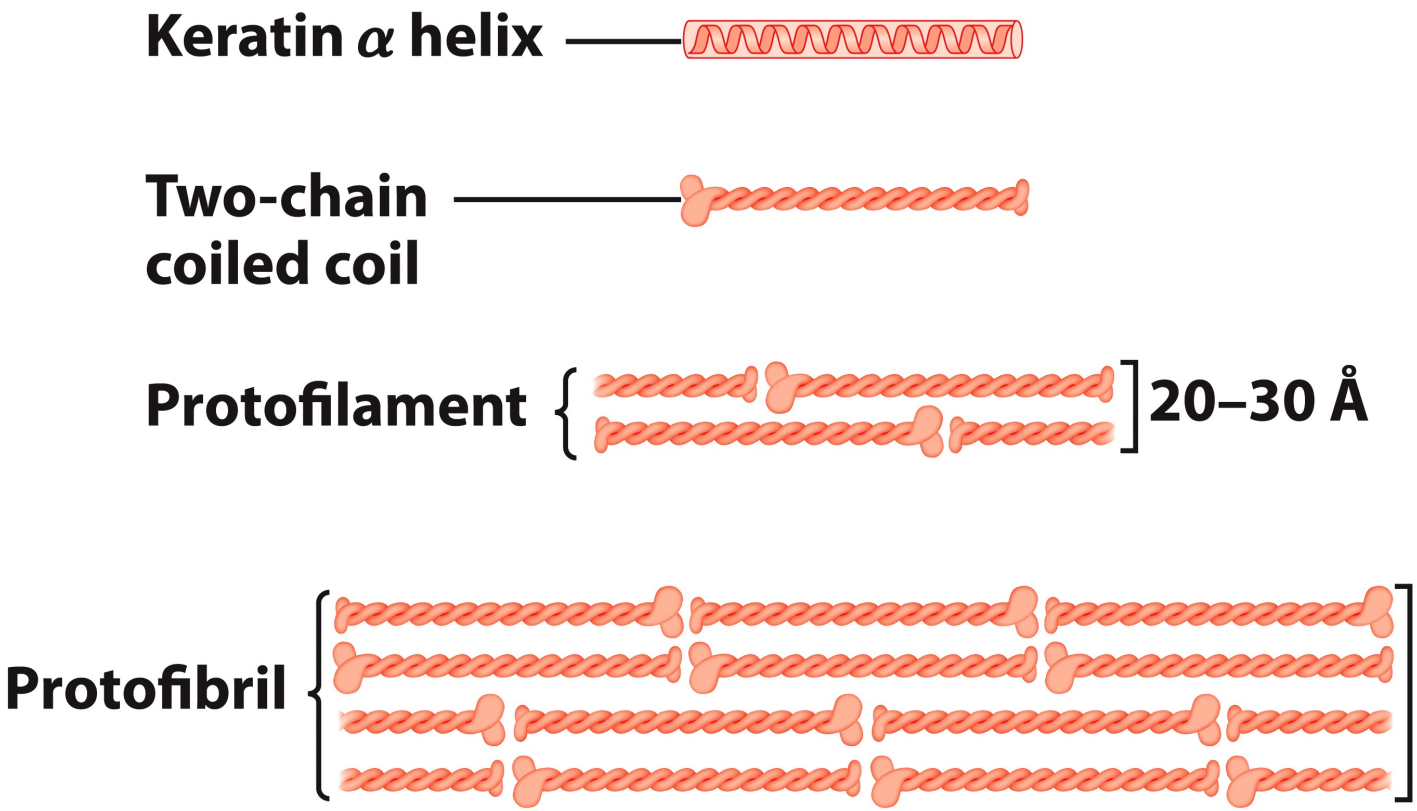


Figure 4-11a
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

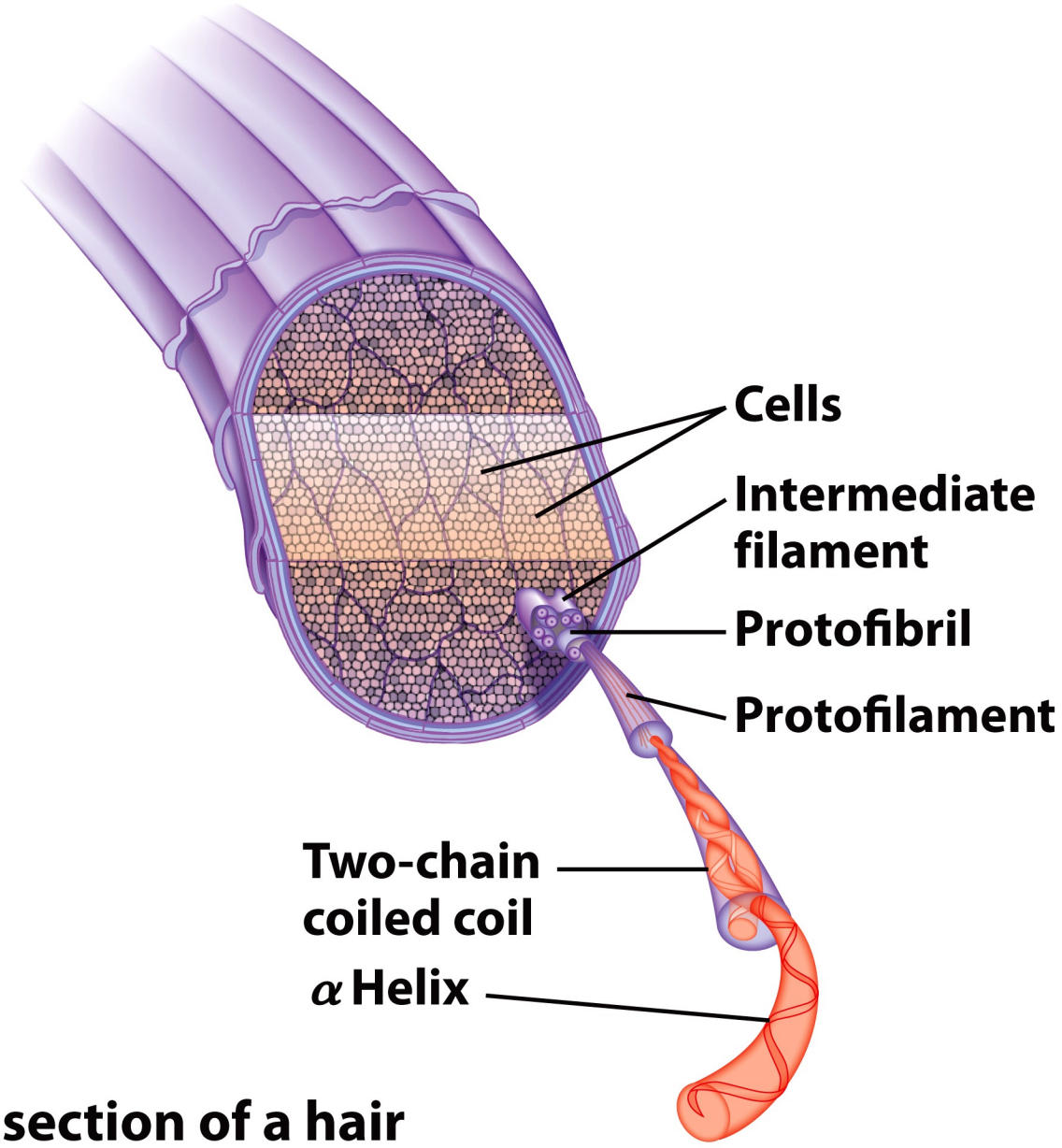
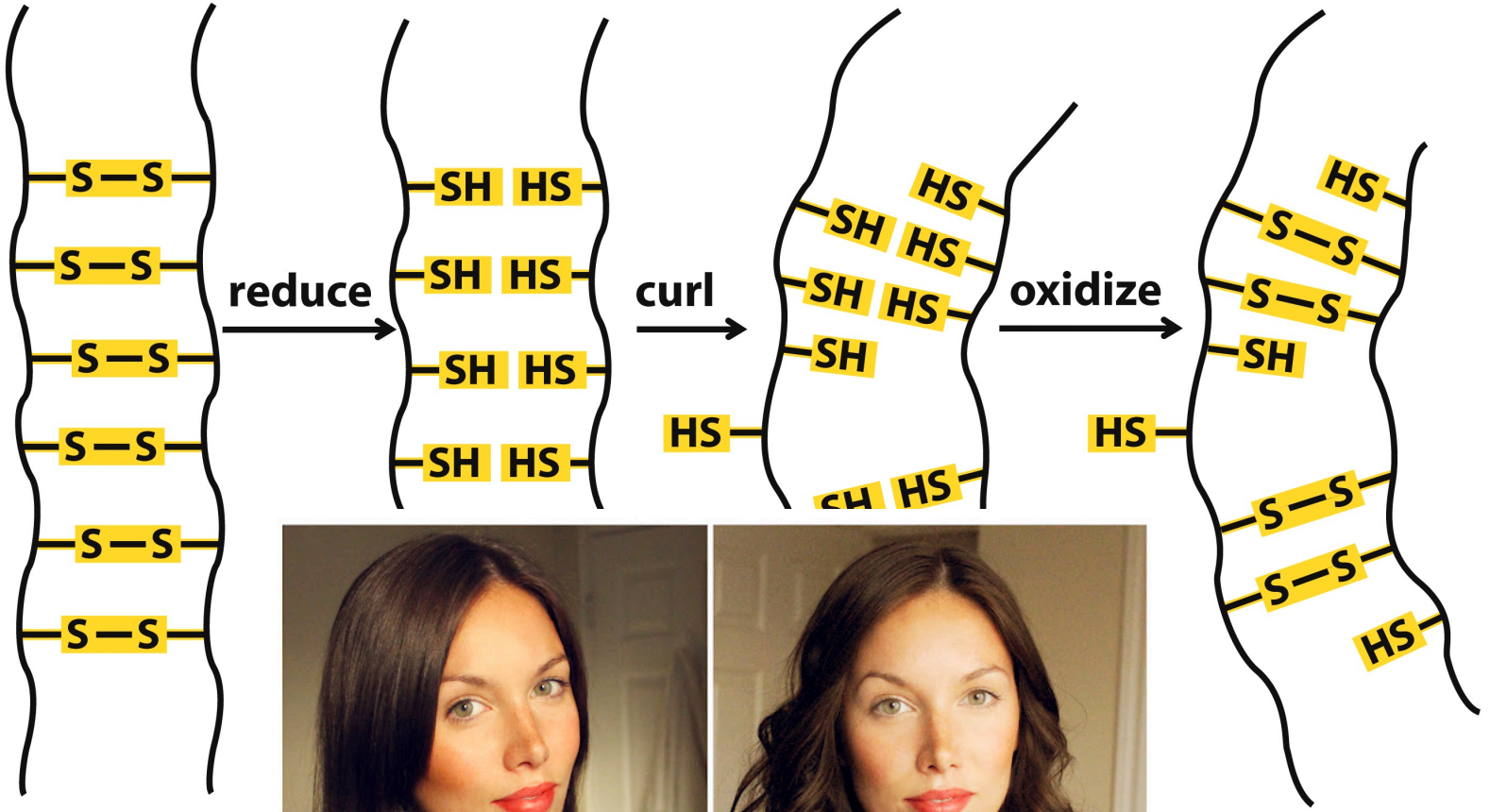


Figure 4-11b
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Chemistry of Permanent Waving



Box 4-2
Lehninger Principles of Biochemistr
© 2013 W. H. Freeman and Compa

Structure of Collagen

- Collagen is an important constituent of connective tissue: tendons, cartilage, bones, cornea of the eye
- Each collagen chain is a long Gly- and Pro-rich left-handed helix
- Collagen helix is different from α helix:

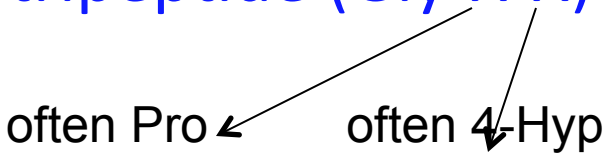
    1) it is left handed

    2) it has 3 aa per turn

- Three collagen chains intertwine into a right-handed superhelical triple helix
- The triple helix has higher tensile strength than a steel wire of equal cross section

Structure of Collagen

- Typical collagen has ~35% Gly, 11% Ala and 21% Pro and 4-Hyp
  - aa sequence is generally a tripeptide (Gly-X-X)



- Many triple-helices assemble into a collagen fibril

- Collagenase is an enzyme that breaks down collagen

Fibrous Proteins

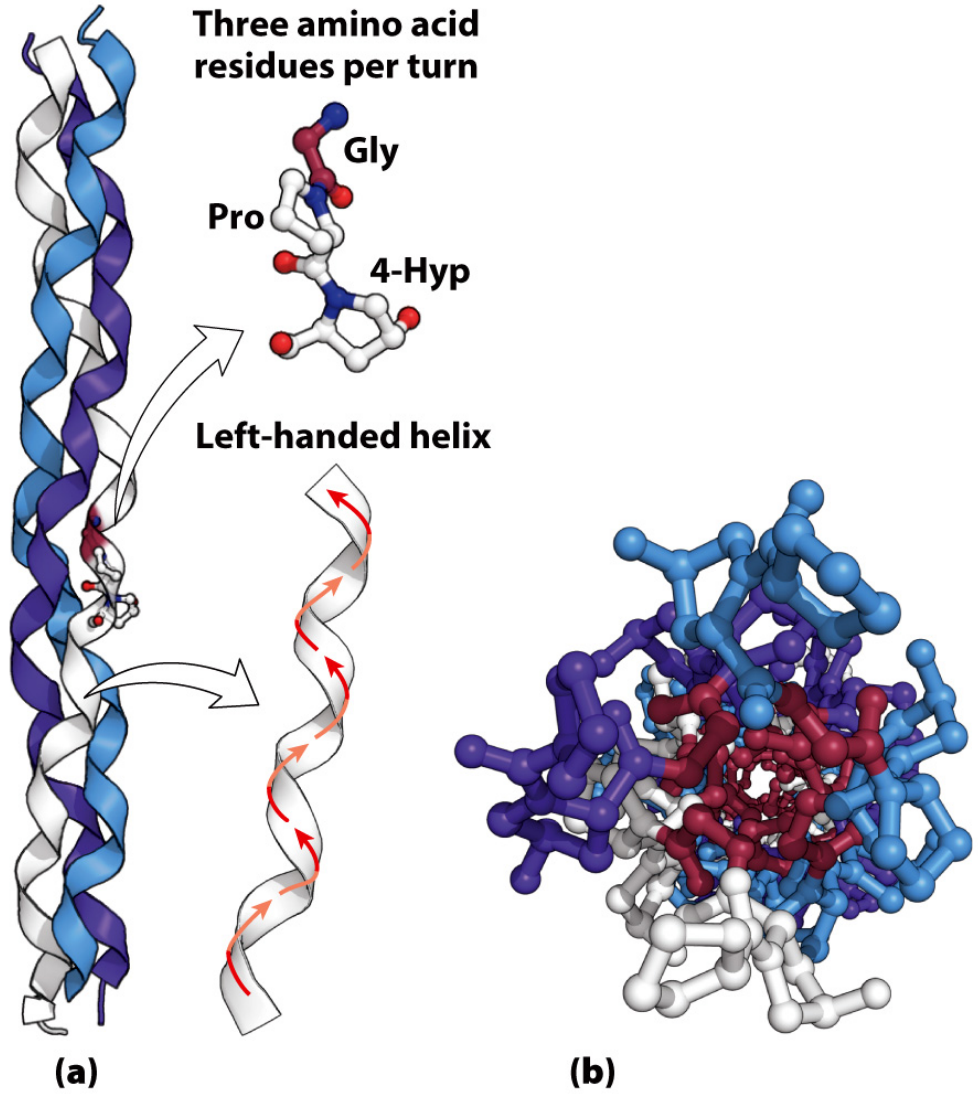


Figure 4-12
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

Collagen Fibrils

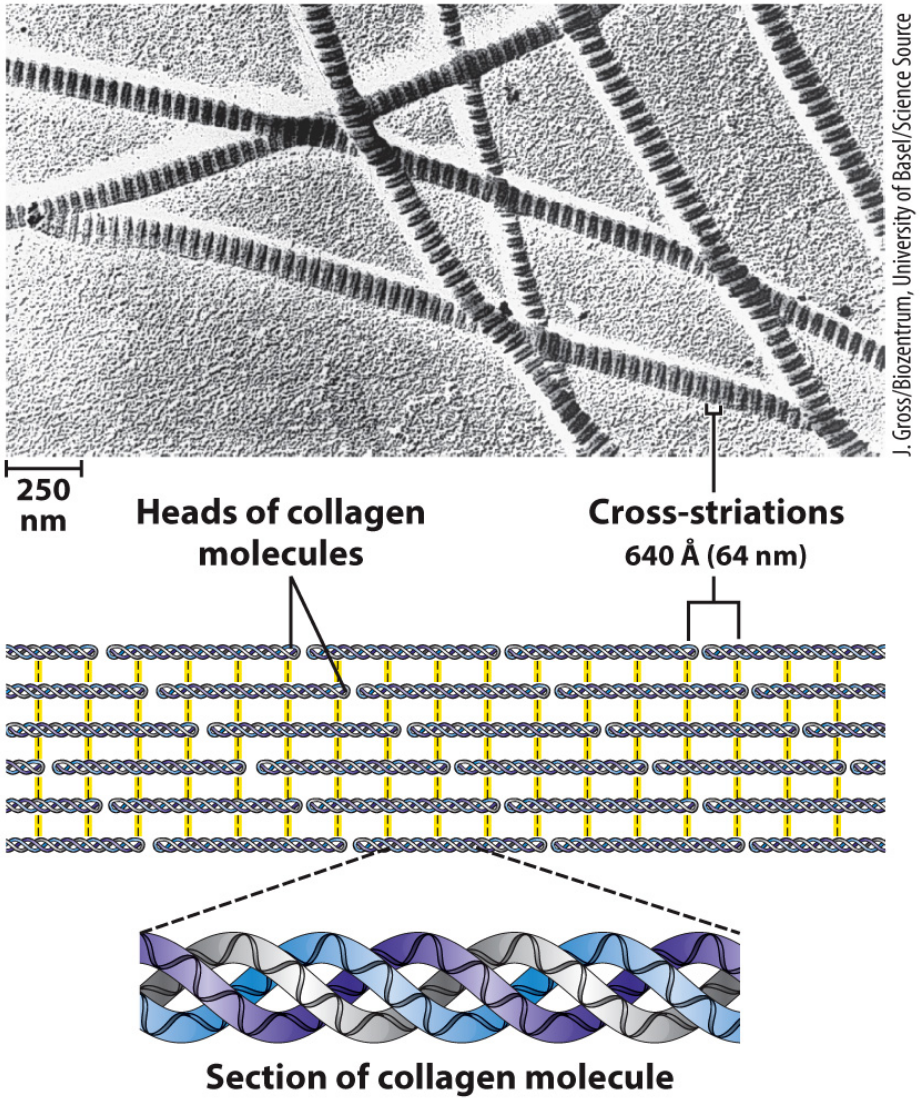


Figure 4-13
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

- Collagen superstructures are formed by cross-linking of collagen triple-helices to form collagen fibrils.
- Crosslinks are covalent bonds between Lys or HyLys, or His amino acid residues.

Silk Fibroin

- Fibroin is the main protein in silk from moths and spiders
- Antiparallel β sheet structure
- Small side chains (Ala and Gly) allow the close packing of sheets
- Silk does not stretch (β conformation is highly extended)
- Flexible because it is held together by weak interactions
- Structure is stabilized by
  - hydrogen bonding within sheets
  - London dispersion interactions between sheets

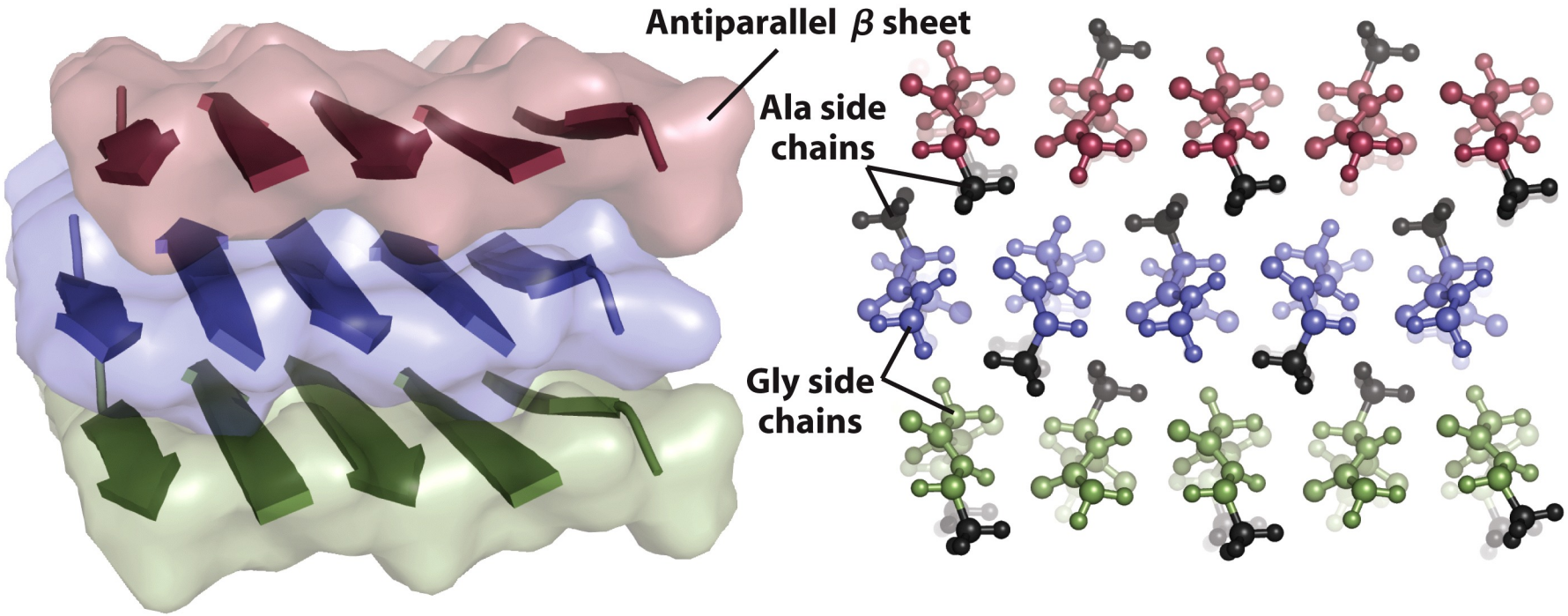


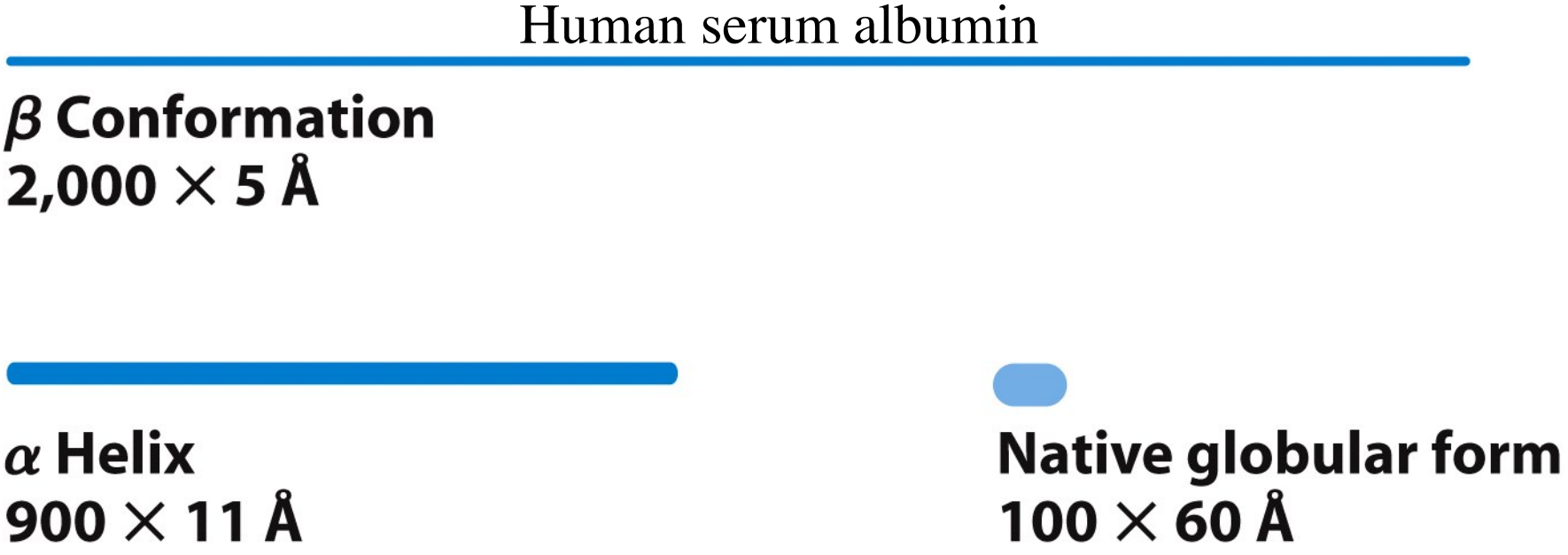
Figure 4-14a
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Spider Silk

- Used for webs, egg sacks, and wrapping the prey
- Extremely strong material
  - stronger than steel
  - can stretch a lot before breaking
- A composite material
  - crystalline parts (fibroin rich)
  - rubber-like stretchy parts

Globular Proteins

- Different segments of polypeptide chain (or multiple polypeptide chains) fold back on each other → compact form
- Provides structural diversity to carry out a wide range of biological functions



Myoglobin

- Small (M_r 16700)
- O_2–binding protein of muscle cells
- Storage of O_2 and rapid supply to contracting muscles
- Has a heme group
- Abundant in muscles of diving mammals
- 8 straight α helices interrupted by bends (some of which are β turns)
- Longest helix – 23 aa; shortest helix – 7 aa
- 70% of myoglobin is α helix
- Most hydrophobic R groups are inside the protein

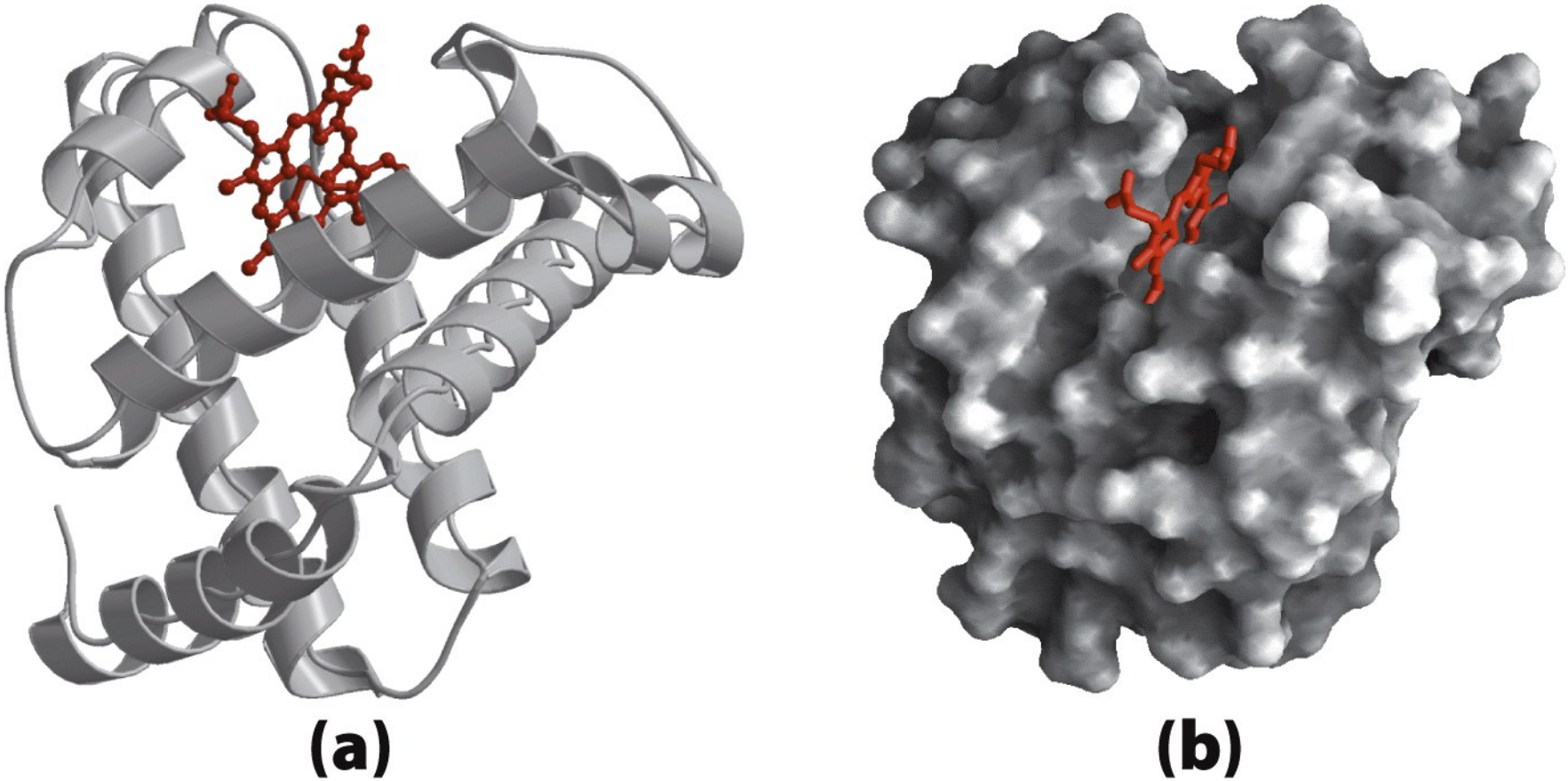
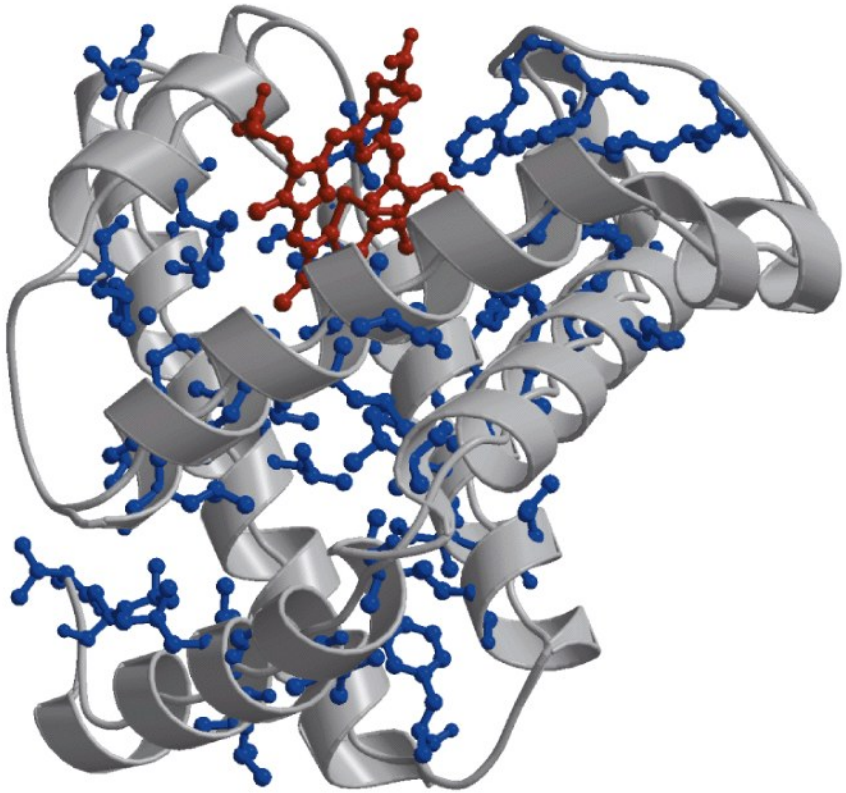
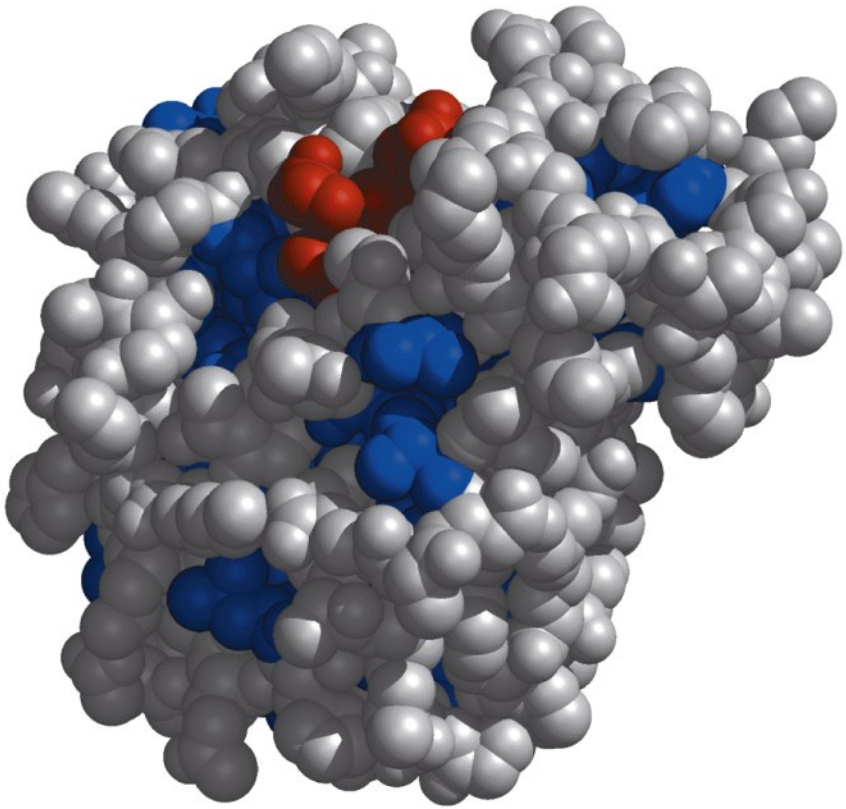


Figure 4-15ab
Lehninger Principles of Biochemistry, Fifth Edition
© 2008 W. H. Freeman and Company



(c)



(d)

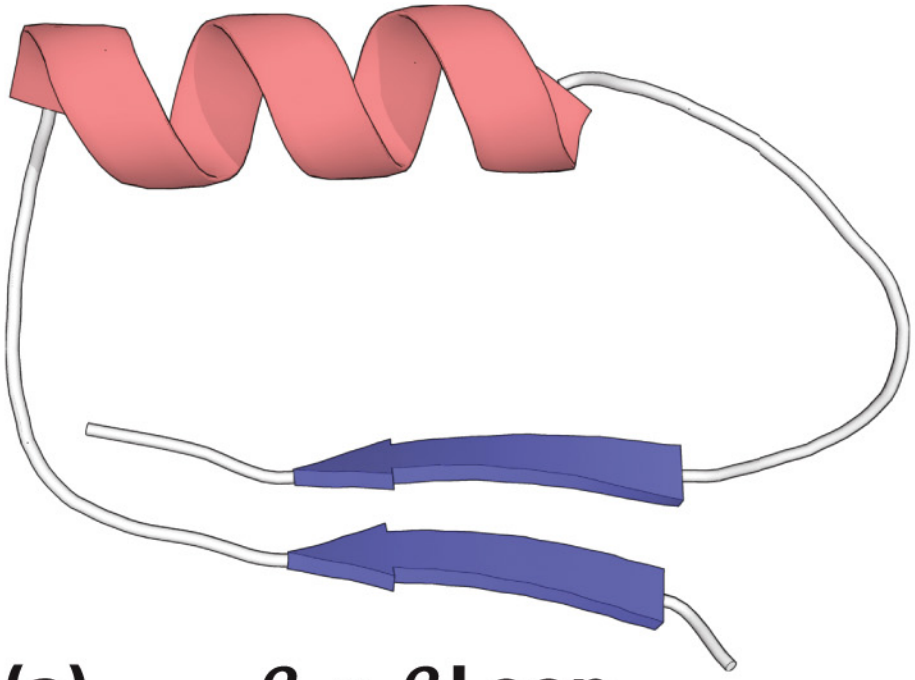
Figure 4-15cd
Lehninger Principles of Biochemistry, Fifth Edition
© 2008 W. H. Freeman and Company

The hydrophobic residues are shown in blue; most are buried in the interior of the protein and thus not visible

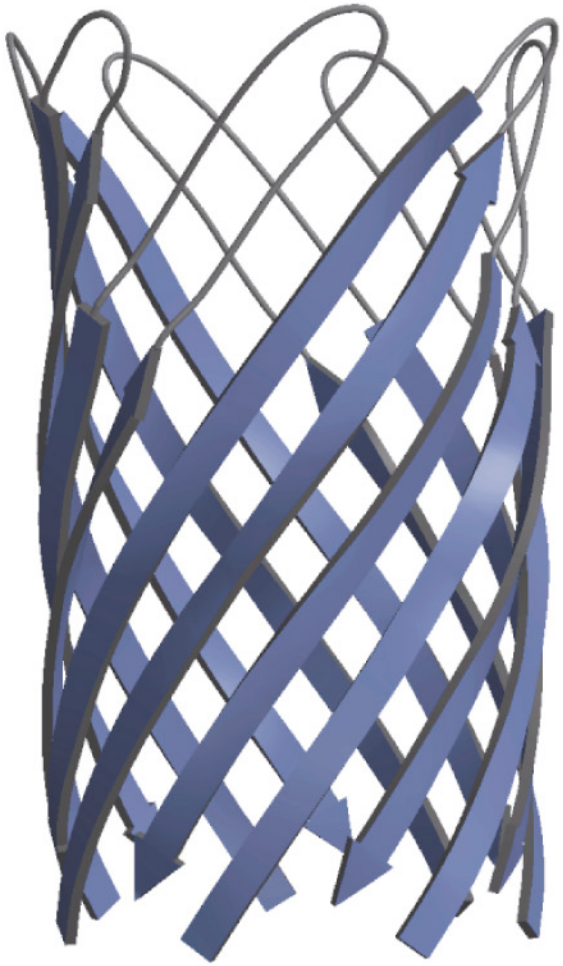
Motifs (folds or supersecondary structures)

- Specific arrangement of several secondary structure elements
  - All alpha-helix
  - All beta-sheet
  - Both
- Motifs can be found as reoccurring structures in numerous proteins
- Proteins are made of different motifs folded together

Motifs (folds)



(a) β-α-β Loop



(b) β Barrel

Figure 4-18
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

- **Complex motifs:**

- A series of β−α−β loops can form a barrel, called α/β **barrel**
- All connections are right handed

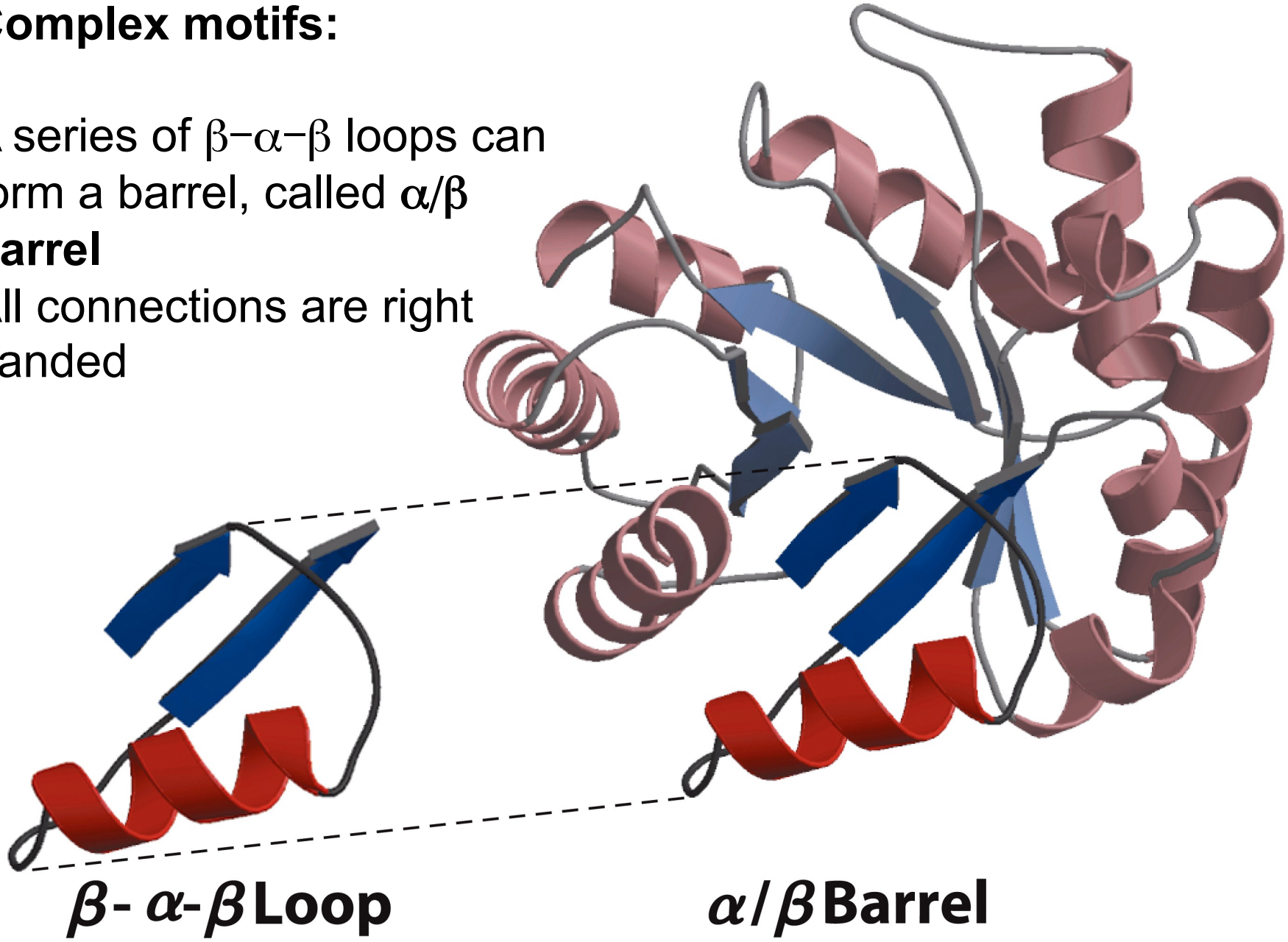


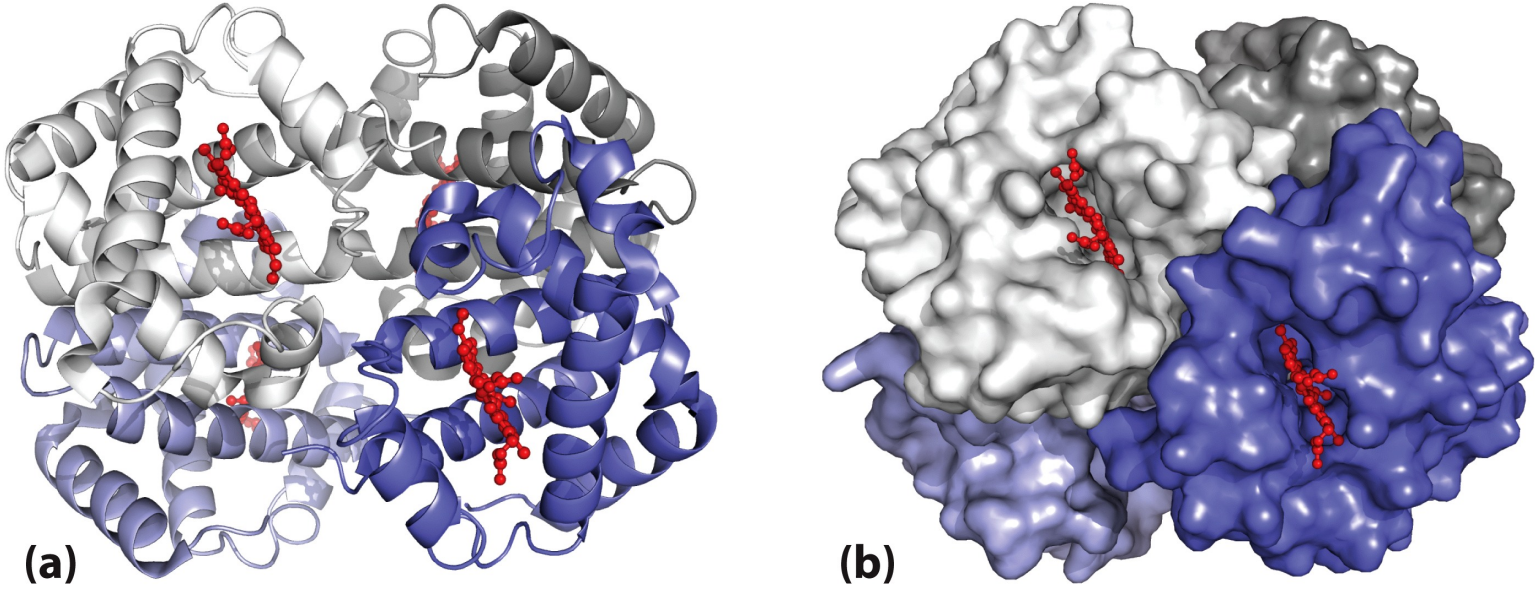
Figure 4-21
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Intrinsically Disordered Proteins

- Contain protein segments that lack definable structure
- Composed of amino acids whose higher concentration forces less-defined structure
  - Lys, Arg, Glu, and Pro
- Disordered regions can conform to many different proteins, facilitating interaction with numerous different partner proteins.

Protein Quaternary Structure

- Quaternary structure is formed by the assembly of individual polypeptides into a larger functional cluster
- It describes the arrangement of protein subunits in 3D complexes
- ***Only*** in multisubunit proteins



Tetramer or dimer of αβ protomers

Proteostasis

Maintenance of cellular protein activity is accomplished by the coordination of many different pathways.

3 processes contribute to proteostasis:

1) protein synthesis on ribosomes
2) protein folding (involving complexes called chaperones)
3) the sequestration and degradation of proteins that are irreversibly unfolded

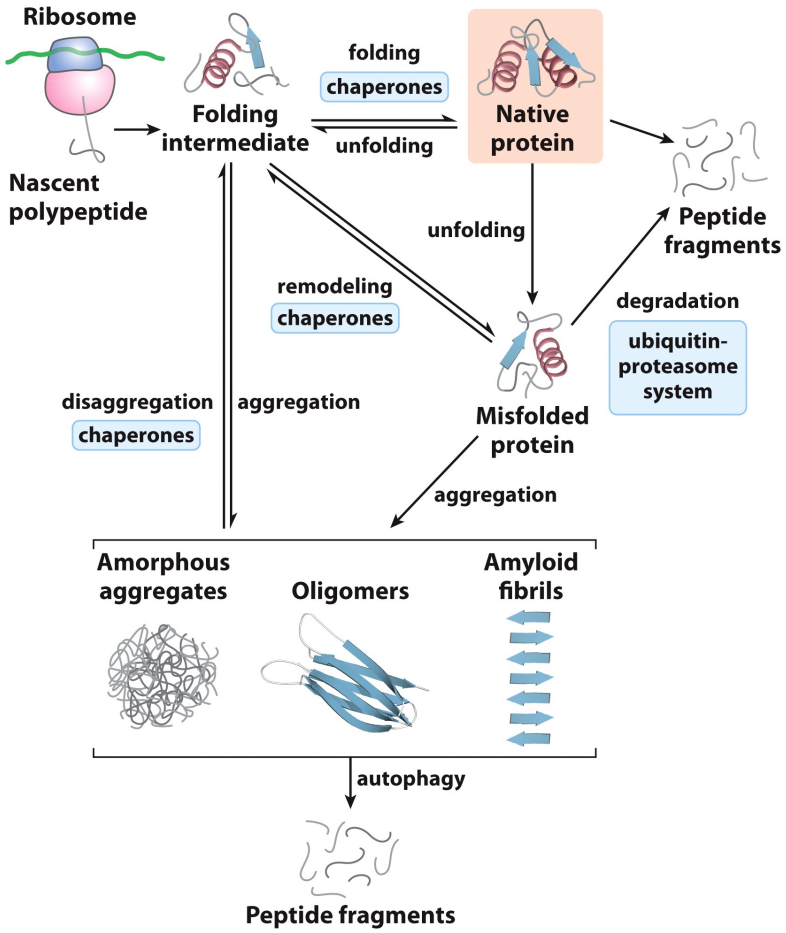


Figure 4-25
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Protein Stability and Folding

- A protein's function depends on its 3D-structure
- Denaturation: the loss of native structure sufficient to cause loss of function
- Does not mean that protein is always unfolded
- Under most conditions, denatured proteins exist in partially folded states

- **Loss of structure → Loss of function**
  - Causes of denaturation:

1) **Heat**: affecting weak interactions (mainly H-bonds)
   - gradual increase in temp, conformation remains intact until an abrupt change occurs at a narrow temp range
   - due to cooperativity (loss of structure in one part of the protein destabilizes other parts)

2) **Extremes of pH**: change net charge on protein → electrostatic repulsion and disruption of some H-bonds

3) **Organic solvents, detergents**: no covalent bonds in the protein are broken. Disruption of hydrophobic interactions

4) **Chaotropic agents (urea and guanidine HCl):** disruption of the arrangement of water molecules that solvate the proteins

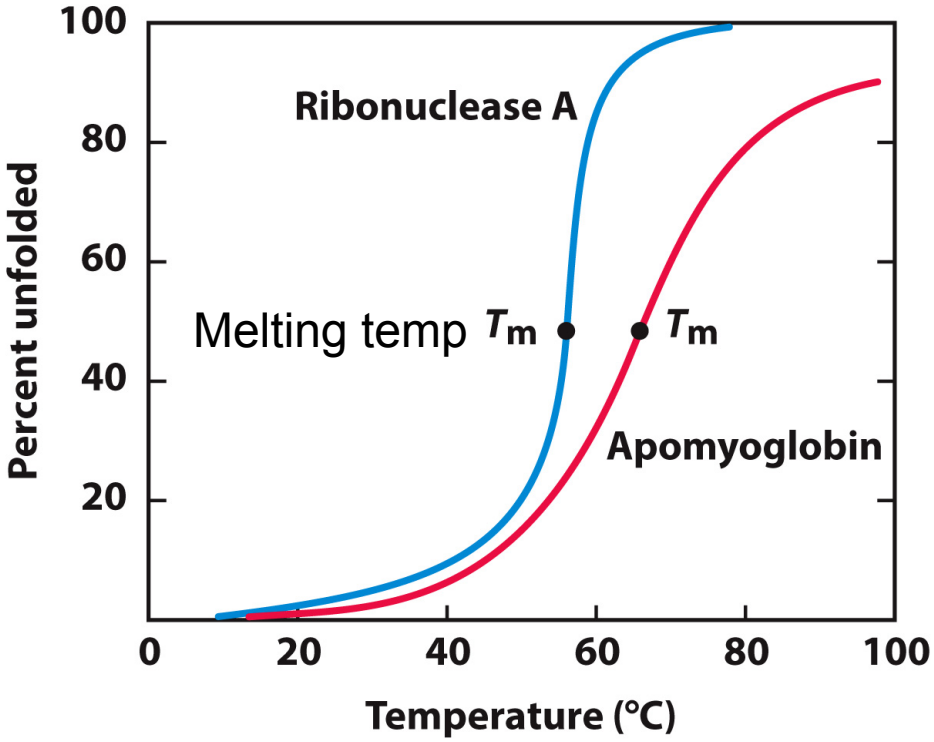


Figure 4-26a
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

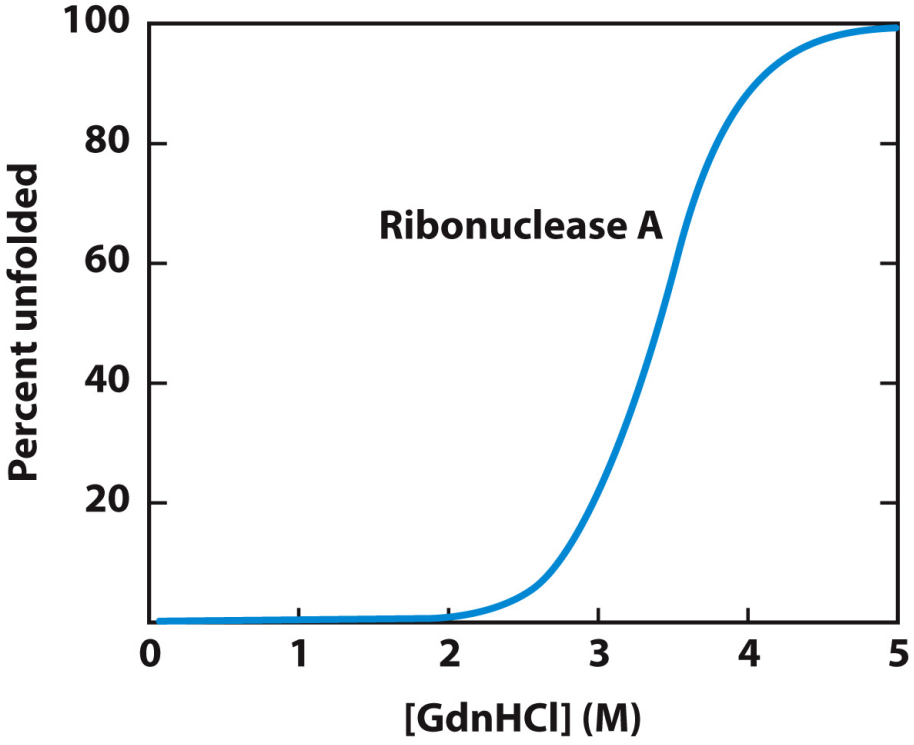


Figure 4-26b
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

Ribonuclease Refolding Experiment

- **Renaturation:** some denatured proteins can get back to their native conformation and biological activity if returned to conditions where native conformation is stable

- E.g. **Ribonuclease A** is a small protein that contains 8 cysteines linked via 4 disulfide bonds. It can be *fully denatured* by urea and a reducing agent (to break its 4 disulfide bonds)
- When urea and 2-mercaptoethanol are removed, the protein *spontaneously* refolds, and the correct disulfide bonds are reformed
- *The sequence alone determines the native conformation*

Quite "simple" experiment, but so important it earned Chris Anfinsen the 1972 Chemistry Nobel Prize

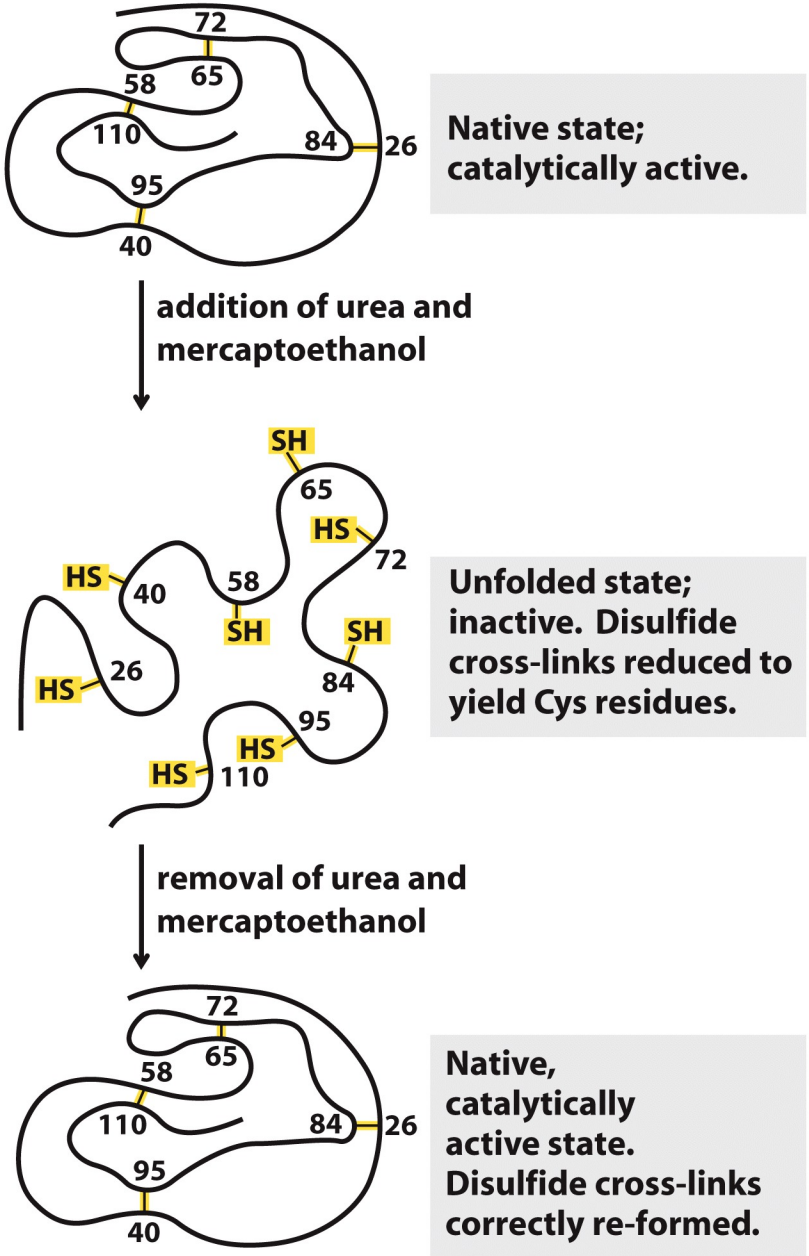


Figure 4-27
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Protein misfolding is the basis of numerous human diseases

- Defects in protein folding cause many diseases:

❑ **Cystic fibrosis** (CF التليف الكيسي) is caused by a mutation in the gene for the protein cystic fibrosis transmembrane conductance regulator (CFTR), a Cl^- channel. A mutation in this protein causes misfolding

❑ **Mad cow disease** caused by a misfolded protein (prion)

❑ Amyloid fibrils accumulate in **Alzheimer's disease** and many other **neurodegenerative diseases**. They arise from misfolded normal proteins and polypeptides

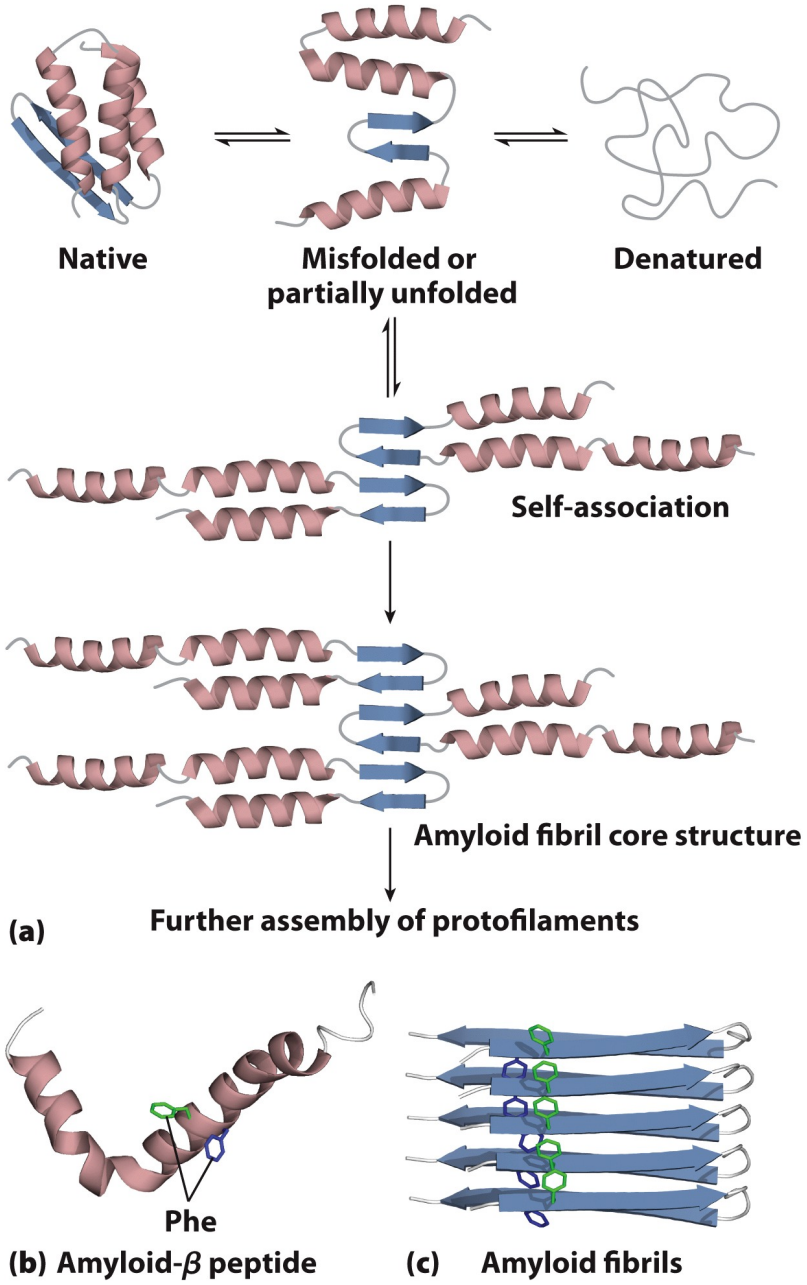


Figure 4-32
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Assisted Folding of Proteins

- Not all proteins fold spontaneously
- Molecular chaperones are proteins that help proteins fold by interacting with partially folded polypeptides and providing microenvironments for proper folding
- Two classes:

1. **Hsp70** – heat shock proteins of MW ~70 kDa
   - bind to unfolded polypeptides rich in hydrophobic aa
   - need another Hsp40
   - In bacteria, homologs are DnaK and DnaJ
   - prevent the aggregation of unfolded proteins
2. **Chaperonins** – (GroEL/GroES in bacteria)
   - barrel and lid, provide good microenvironment for proteins to fold properly

Chaperones prevent misfolding

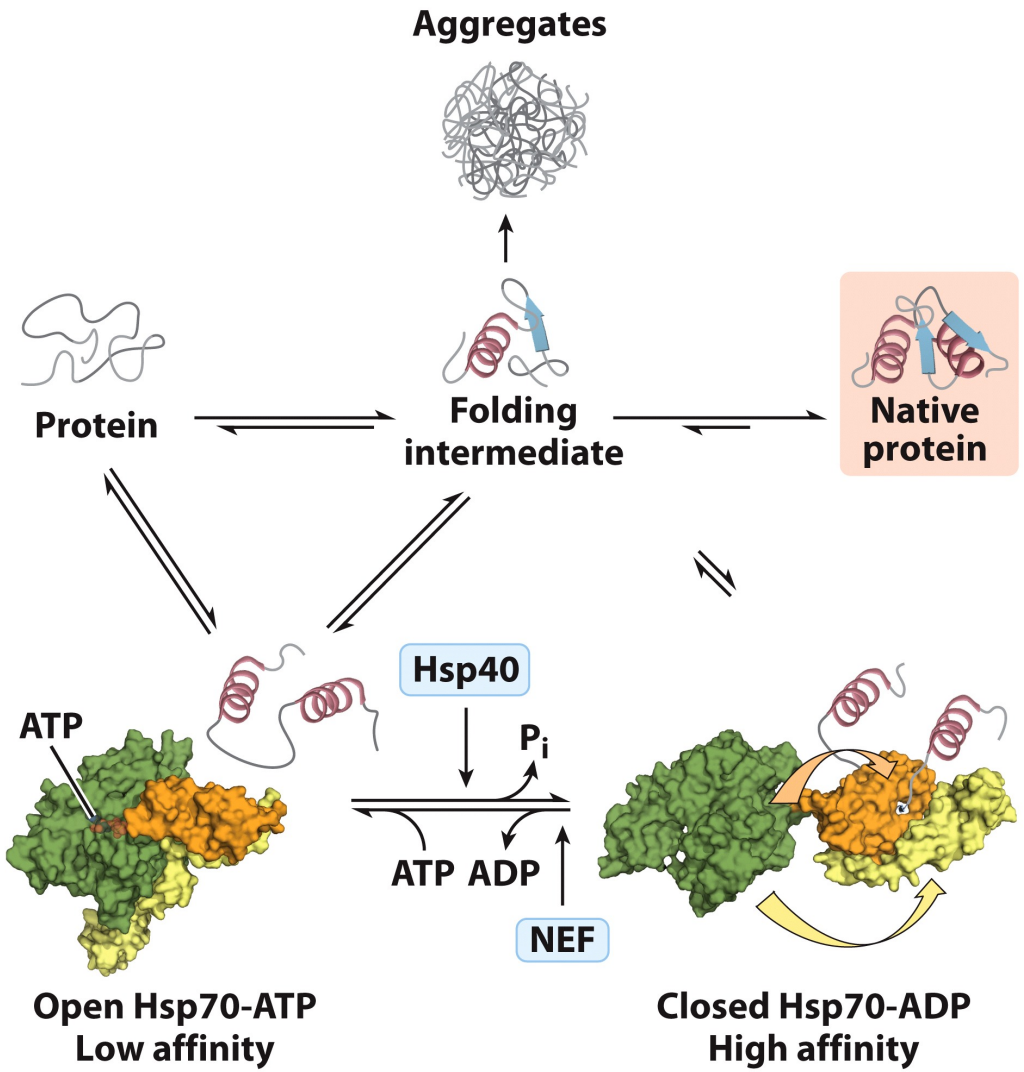


Figure 4-30
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Chaperonins facilitate folding

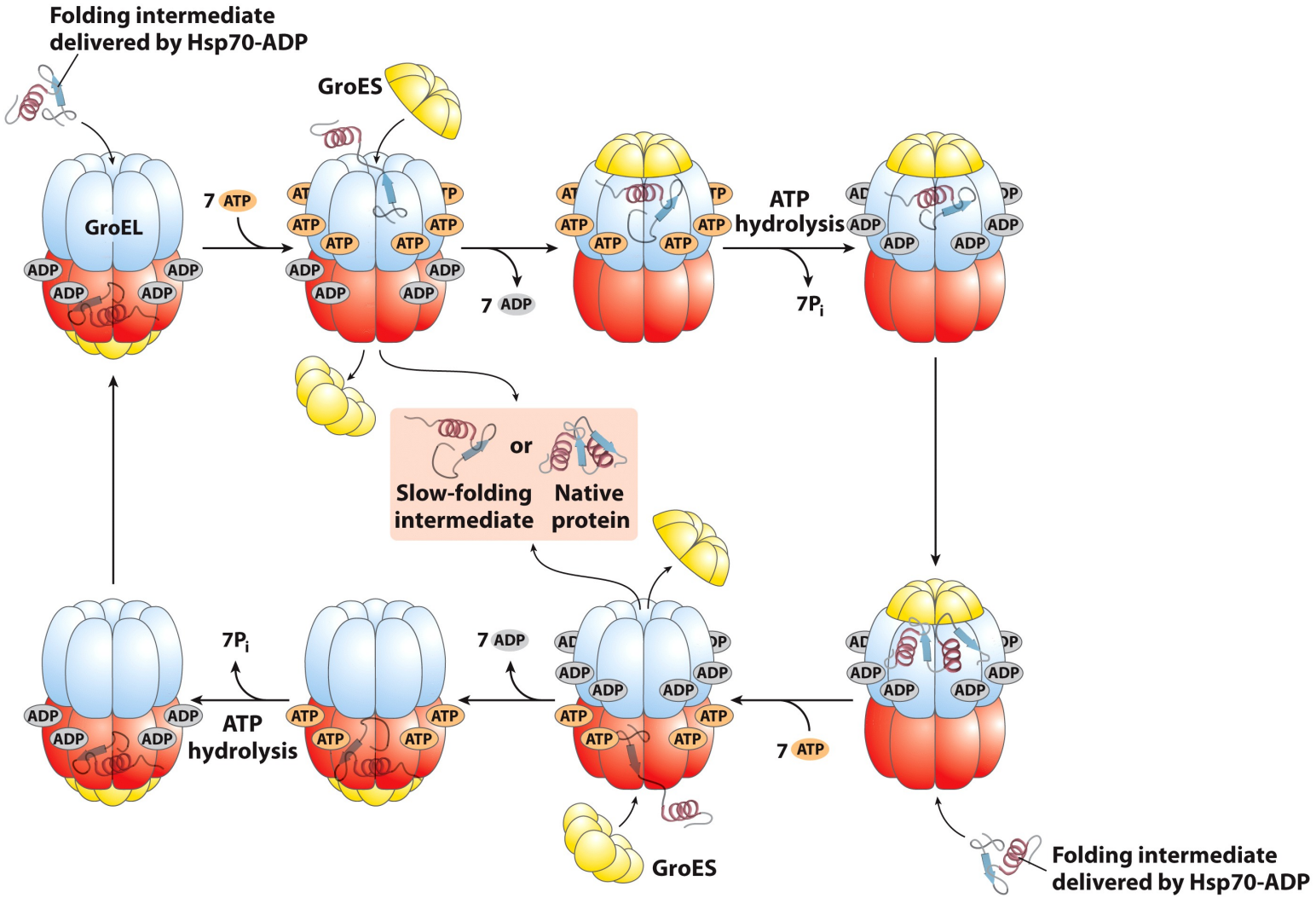


Figure 4-31a
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

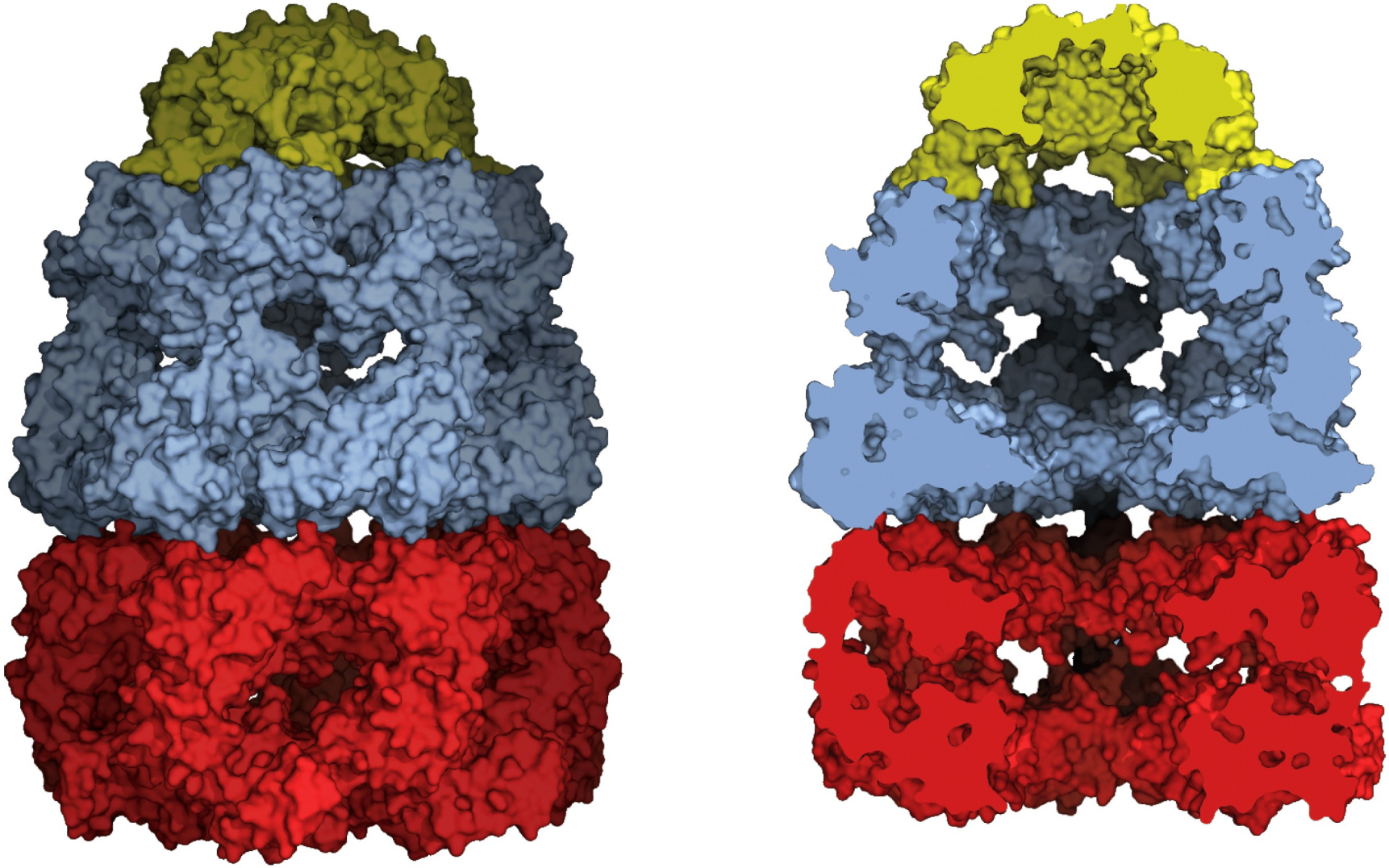


Figure 4-31b
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Chapter 4: Summary

In this chapter, we learned about:

- the two most important secondary structures
  - α helices
  - β sheets
- how properties and function of fibrous proteins are related
- how to determine three-dimensional structures of proteins
- one of the largest unsolved puzzles in modern biochemistry: how proteins fold