Chapter 4

The Three-Dimensional Structure of Proteins

1. Structure of proteins

- **Protein molecules adopt a specific three-dimensional conformation in the aqueous solution --- the native fold**
- **This structure is able to fulfill a specific biological function**
- **The native fold has a large number of favorable interactions within the protein**
- **The most important forces stabilizing the specific structures are noncovalent interactions**
- **The three-dimensional structure of a protein is determined by its amino acid sequence**
- **The function of a protein depends on its structure**

Conformation of proteins

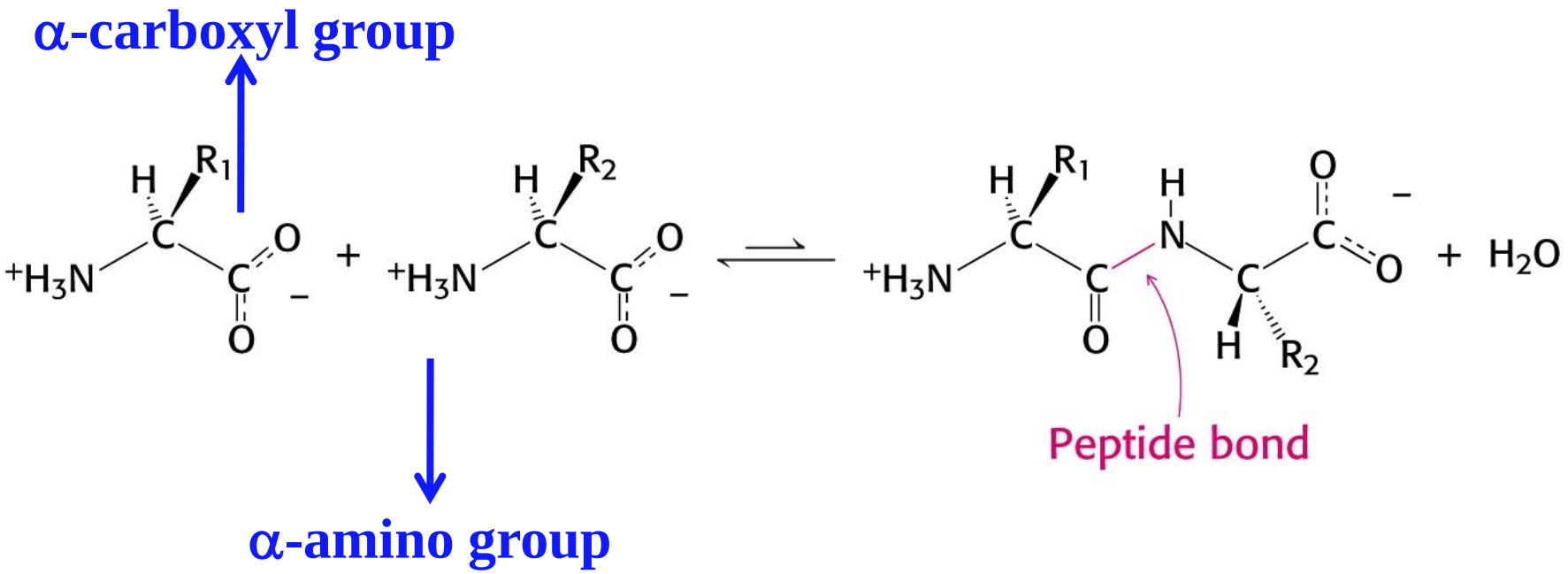
- **Conformation: the spatial arrangement of atoms in a protein**
- **Include any structural state a protein can achieve without breaking covalent bonds**
- **One or a few conformations generally dominate under physiological conditions**
- **A change in conformation could occur by rotation about single bonds**
- **The most stable conformation is the one with the maximum number of weak interactions**

Favorable interactions in proteins

- **Hydrophobic effect**
  - **The association of nonpolar groups with each other in the interior of a protein**
- **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**
- **Van der Waals interactions**
  - **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**

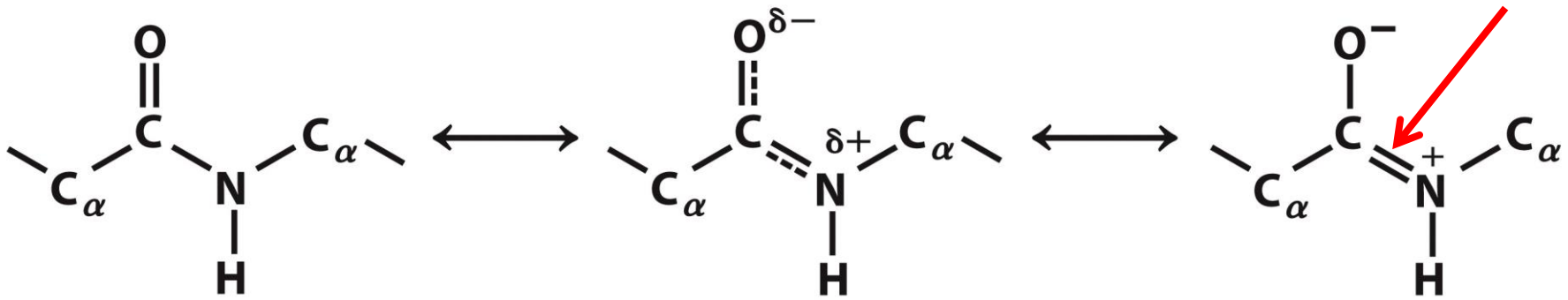
2. Structure of the peptide bond

- **Structure of the protein is partially dictated by the properties of the peptide bond**
- **The peptide bond exhibits a large dipole moment in the favored trans configuration**
- **The peptide bond has a partial double-bond character that keeps the entire six-atom peptide group in a rigid planar configuration**



Peptide bond: linkage between α-carboxyl group of one amino acid to the α-amino group of another amino acid

C_α-C-N-C_α



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

C=N double bond:1.27Å

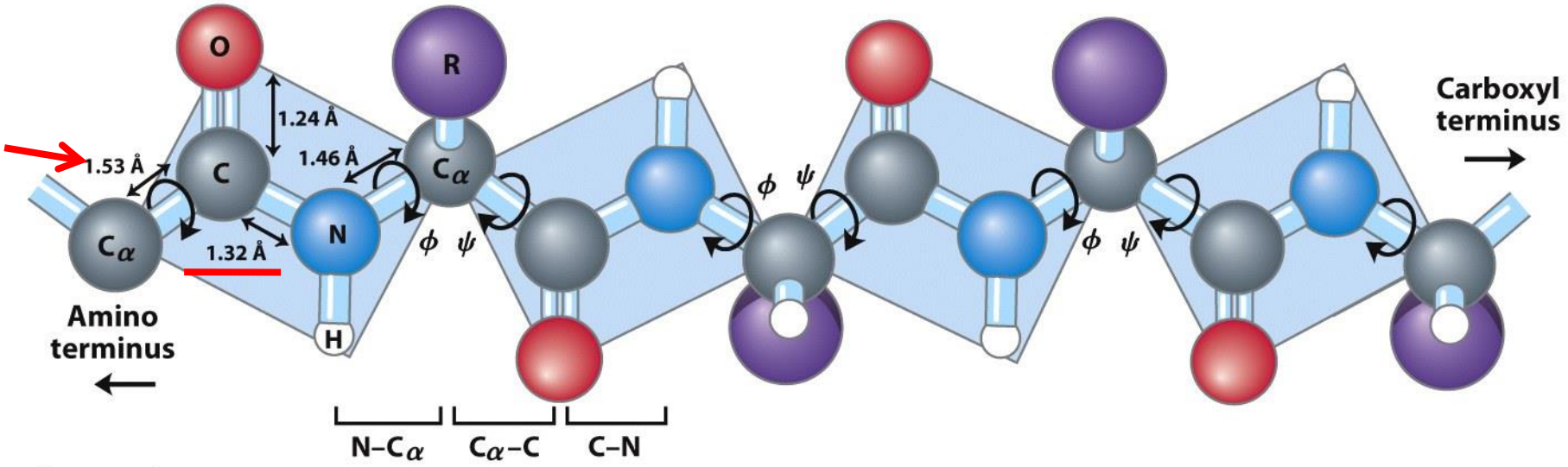
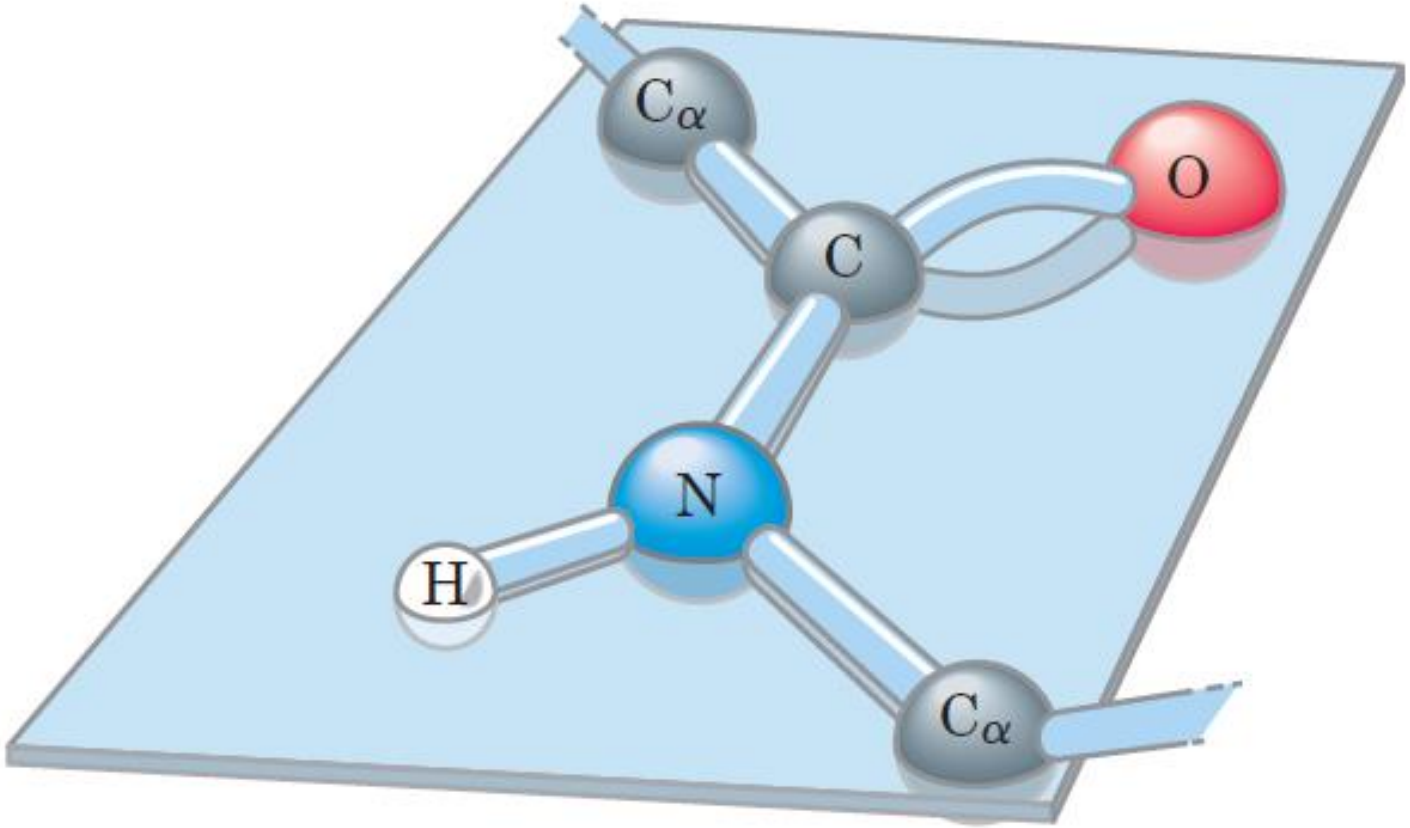


Figure 4-2b
Lehninger Principles of Biochemistry, Fifth Edition

The planar peptide groups

Three bonds separate sequential α carbons in a polypeptide chain. The N—C_α and C_α—C bonds, designated ϕ and ψ, respectively, can rotate. The peptide C—N bond is not free to rotate. Other single bonds in the backbone may also be rotationally hindered, depending on the size and charge of the R groups.



All four atoms attached to the C—N group are located in the same plane

The rigid peptide plane and the partially free rotations

- **Rotation around the peptide bond (C—N) is not permitted**
- **Rotation around bonds connected to the α-carbon is permitted**
- **The freedom of rotation about the two bonds of each amino acid allows proteins to be folded in many different ways**
- **ϕ (phi): the angle of rotation around the α-carbon—amide nitrogen bond (N—C_α)**
- **ψ (psi): the angle of rotation around the α-carbon—carbonyl carbon bond (C_α—C)**

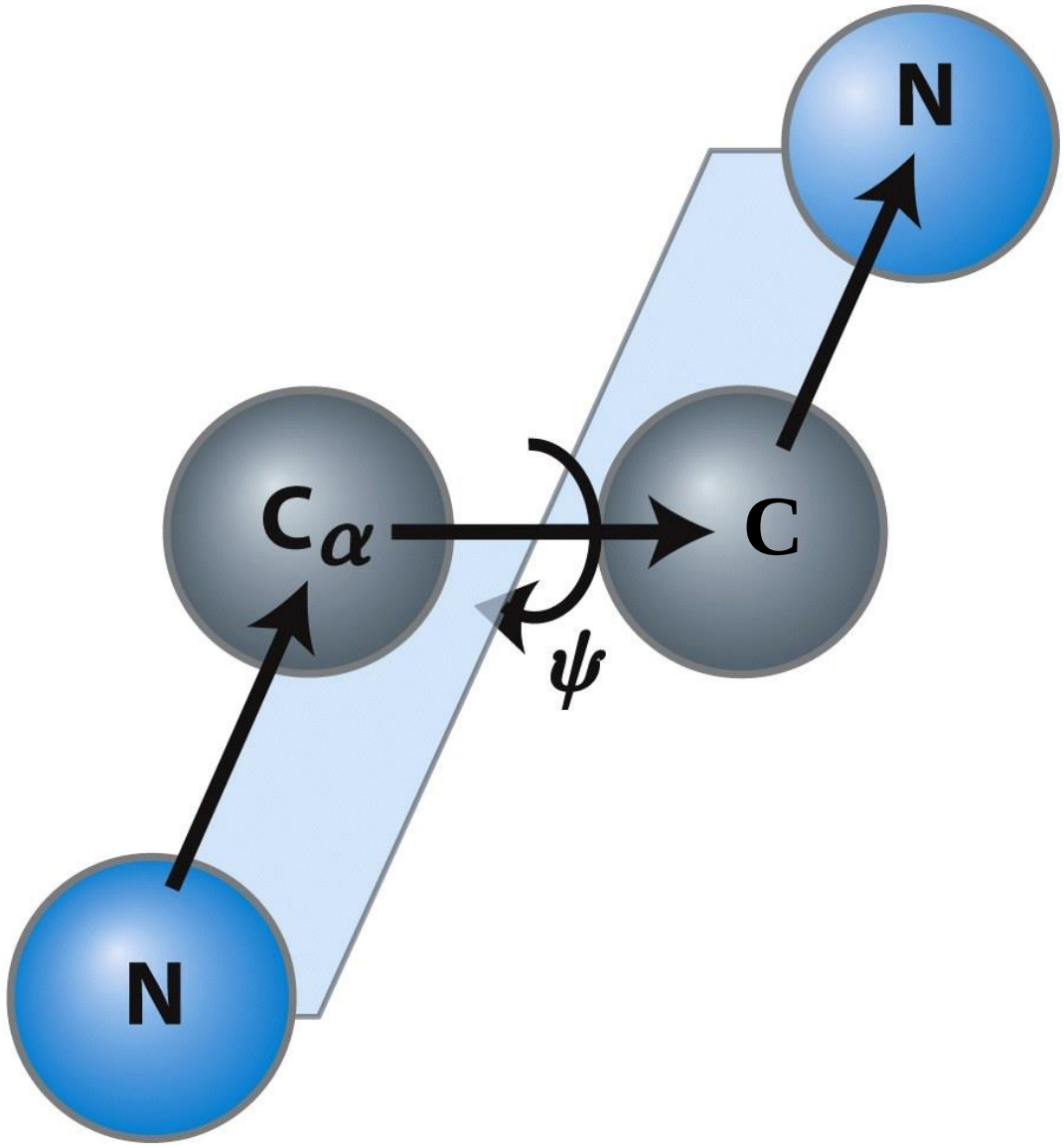


Figure 4-2c
Lehninger Principles of Biochemistry, Fifth Edition

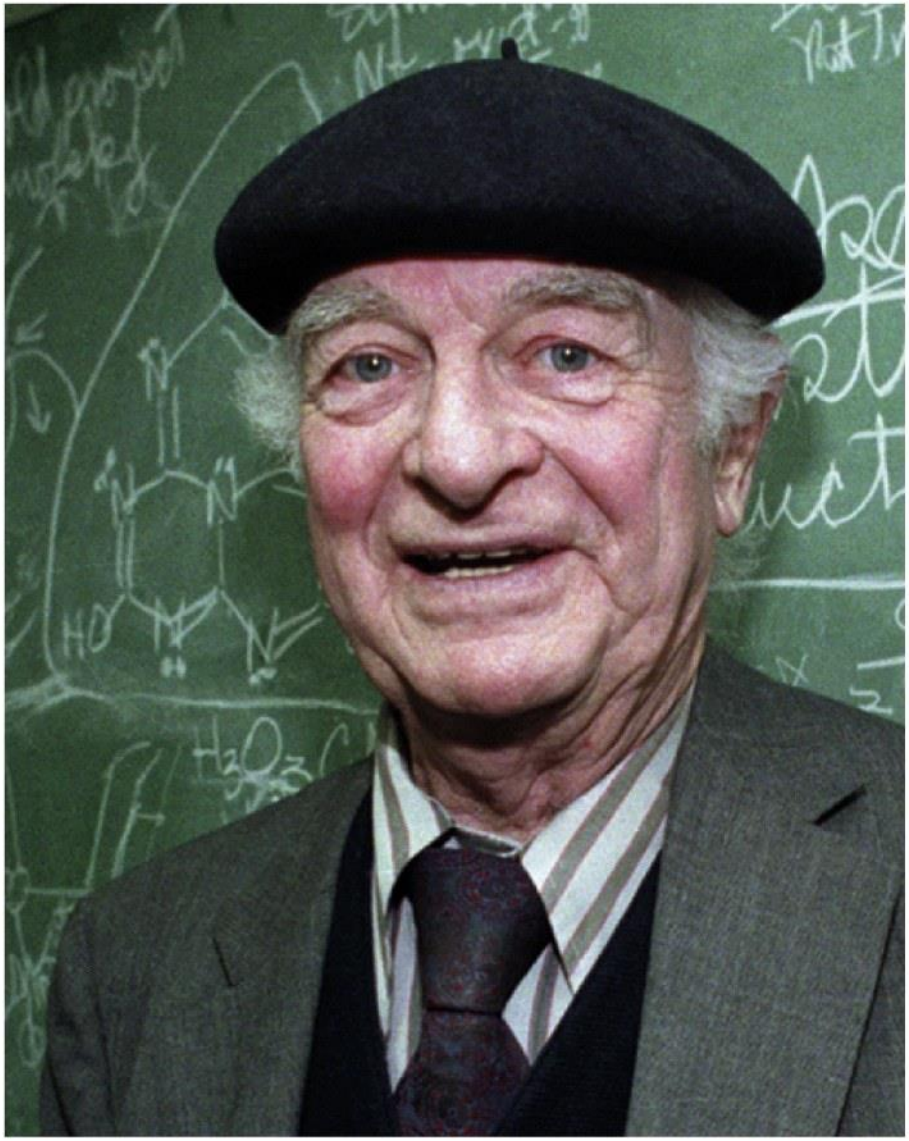
The atoms and planes defining ψ

Distribution of ϕ and ψ dihedral angles

- **In a fully extended polypeptide, both ψ and ϕ are 180°**
- **Some ϕ and ψ combinations are very unfavorable because of steric interference between atoms in the polypeptide backbone and amino acid side chains**
- **Some ϕ and ψ combinations are more favorable because of chance to form favorable H-bonding interactions along the backbone**

3. Protein secondary structure

- **Secondary structure refers to a local spatial arrangement of the polypeptide chain**
- **Two regular arrangements are common (Proposed by Linus Pauling & Robert Corey in 1951)**
- **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**
- **The organization around the peptide bond, paired with the identity of the R groups, determines the secondary structure of the protein.**



Linus Pauling, 1901–1994

Unnumbered 4 p115a

Lehninger Principles of Biochemistry, Fifth Edition

The Nobel Prize in Chemistry 1954

Linus Carl Pauling

USA

California Institute of Technology (Caltech)
Pasadena, CA, USA

"for his research into the nature of the chemical bond and its application to the elucidation of the structure of complex substances"

The Nobel Peace Prize 1962

Linus Carl Pauling

USA

California Institute of Technology (Caltech)
Pasadena, CA, USA

The α helix

- **The backbone is more compact with the ψ dihedral (N–C_α—C–N) in the range (0° < ψ < –70°)**
- **Helical backbone is held together by hydrogen bonds between the nearby backbone amides**
- **Right-handed helix with 3.6 residues (5.4 Å) per turn**
- **Stabilized by the hydrogen bond between the CO and NH groups of the main chain**
- **Peptide bonds are aligned roughly parallel with the helical axis**
- **Side chains point out and are roughly perpendicular with the helical axis**

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*	−60°	−30°
i + 2*	−90°	0°
β Turn type II		
i + 1	−60°	+120°
i + 2	+80°	0°

Note: In real proteins, the dihedral angles often vary somewhat from these idealized values.

*The i+1 and i+2 angles are those for the second and third amino acid residues in the β turn, respectively.

Table 4-1
Lehninger Principles of Biochemistry, Sixth Edition

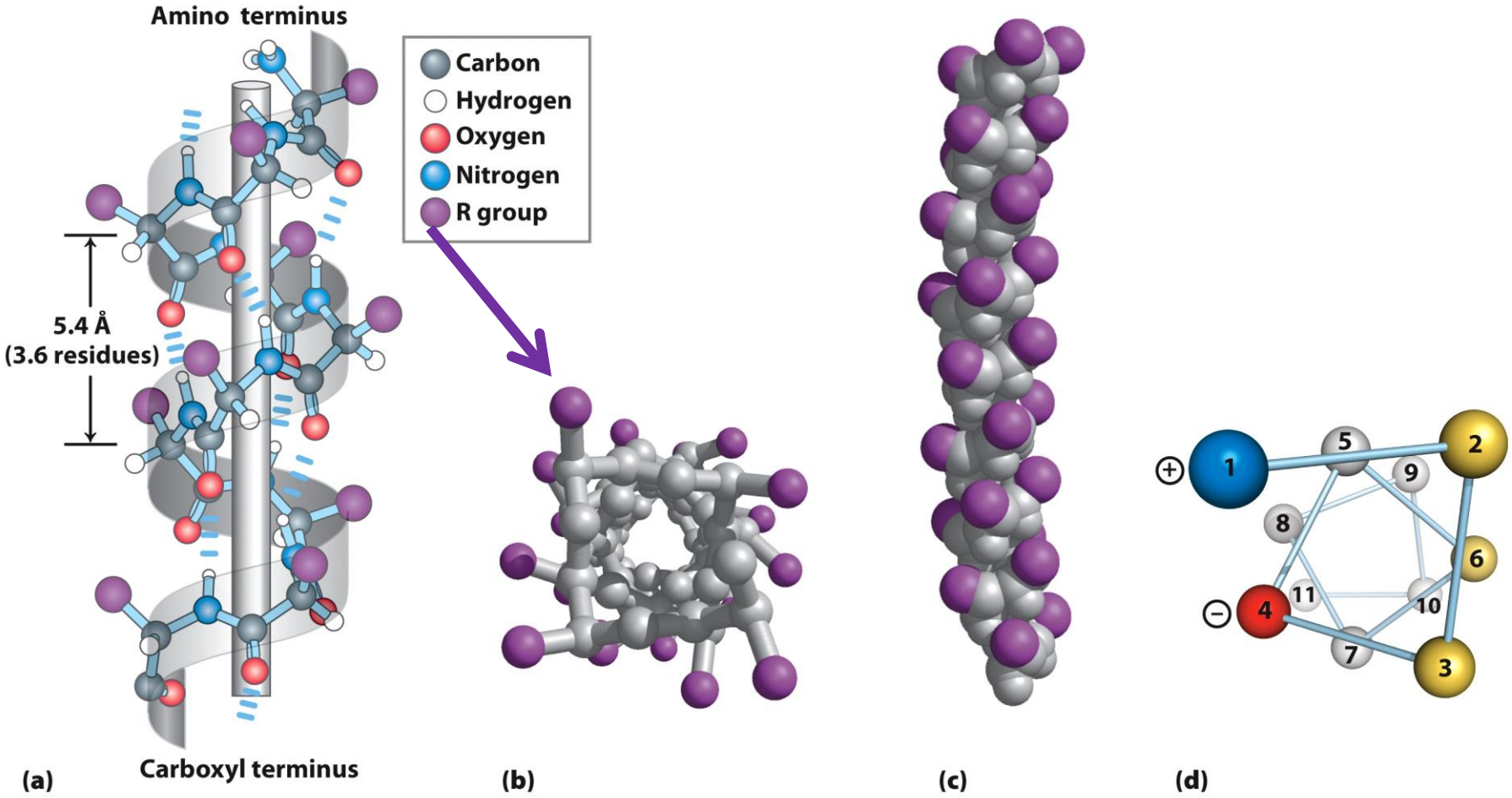
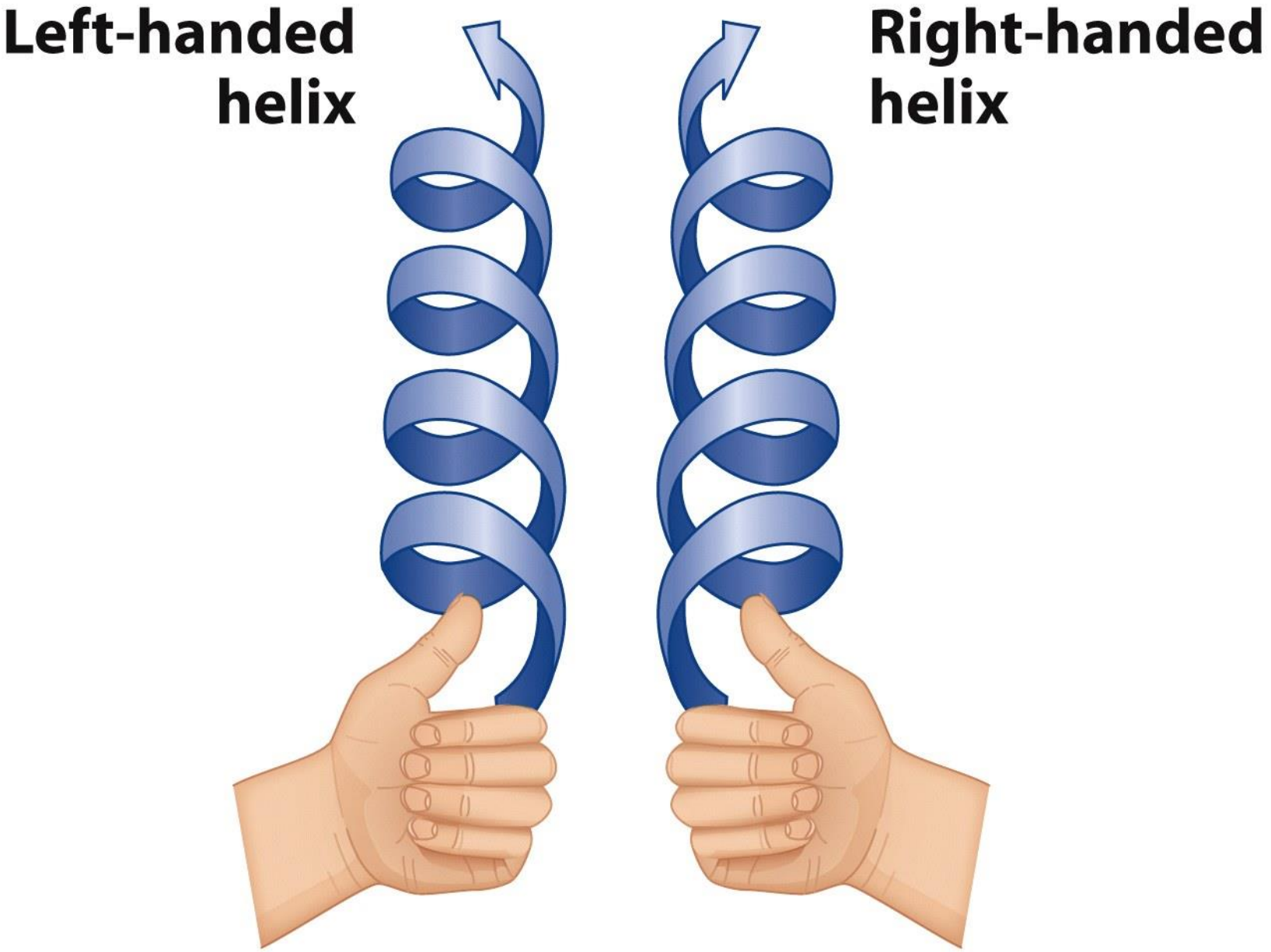


Figure 4-4
Lehninger Principles of Biochemistry, Sixth Edition

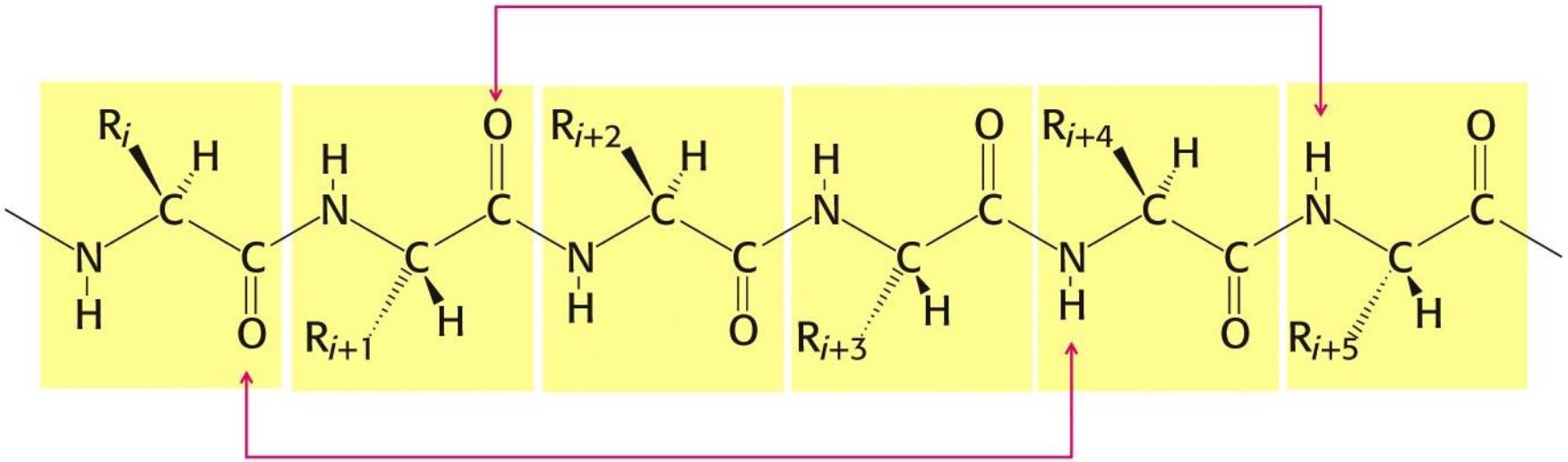
Models of the α helix, showing different aspects of its structure



Box 4-1

Lehninger Principles of Biochemistry, Fifth Edition

Hydrogen bonding in α-helix



In the α helix, the CO group of residue *n* forms a hydrogen bond with the NH group of residue *n* + 4.

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**

TABLE 4–1 Propensity of Amino Acids to Take Up an α-Helical Conformation

Amino acid	$\Delta\Delta G°$ (kJ/mol)*	Amino acid	$\Delta\Delta G°$ (kJ/mol)*
Ala	0	Leu	0.79
Arg	0.3	Lys	0.63
Asn	3	Met	0.88
Asp	2.5	Phe	2.0
Cys	3	Pro	>4
Gln	1.3	Ser	2.2
Glu	1.4	Thr	2.4
Gly	4.6	Tyr	2.0
His	2.6	Trp	2.0
Ile	1.4	Val	2.1

Sources: Data (except proline) from Bryson, J.W., Betz, S.F., Lu, H.S., Suich, D.J., Zhou, H.X., O'Neil, K.T., & DeGrado, W.F. (1995) Protein design: a hierarchic approach. *Science* **270**, 935. Proline data from Myers, J.K., Pace, C.N., & Scholtz, J.M. (1997) Helix propensities are identical in proteins and peptides. *Biochemistry* **36**, 10,926.

*$\Delta\Delta G°$ is the difference in free-energy change, relative to that for alanine, required for the amino acid residue to take up the α-helical conformation. Larger numbers reflect greater difficulty taking up the α-helical structure. Data are a composite derived from multiple experiments and experimental systems.

Table 4-1

Lehninger Principles of Biochemistry, Fifth Edition

The helix dipole

- **The peptide bond has a strong dipole moment**
  - **Carbonyl O negative**
  - **Amide H positive**
- **All peptide bonds in the α helix have a similar orientation**
- **The α helix has a large macroscopic dipole moment**
- **Negatively charged residues often occur near the amino terminus of the helical segment to interact with the positive charge of the helix dipole**

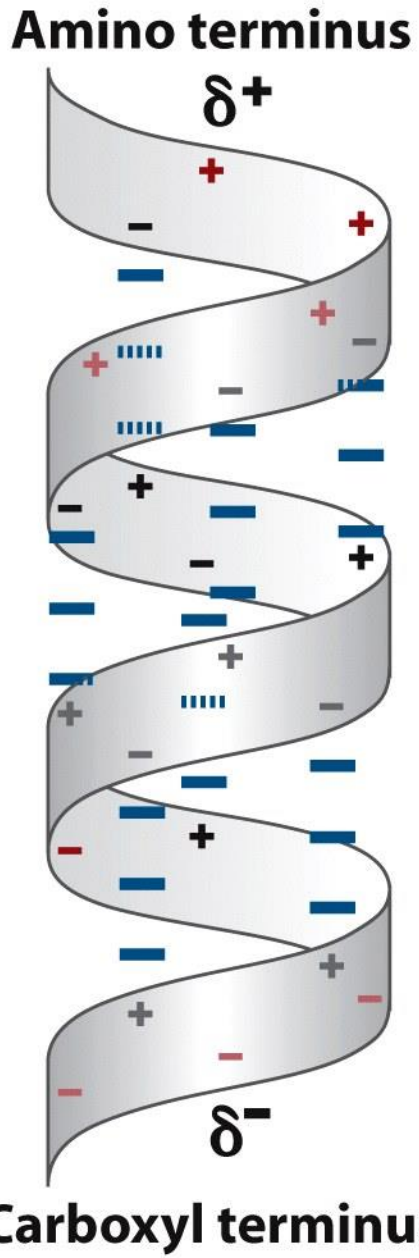
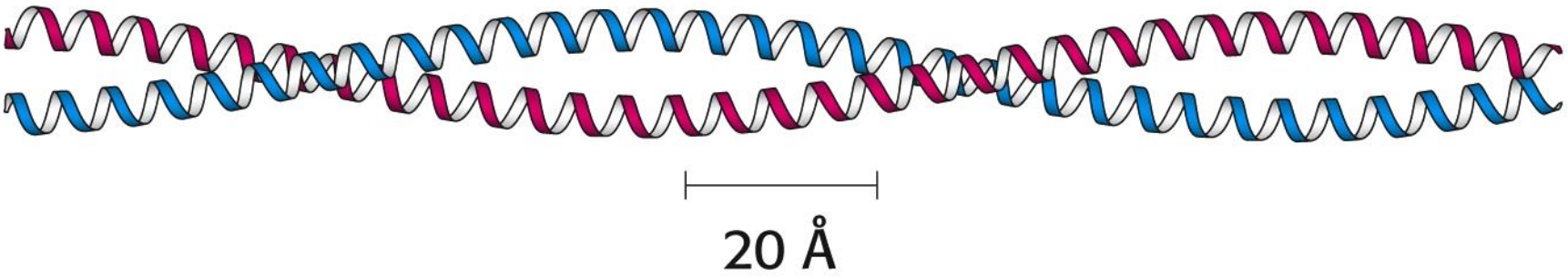


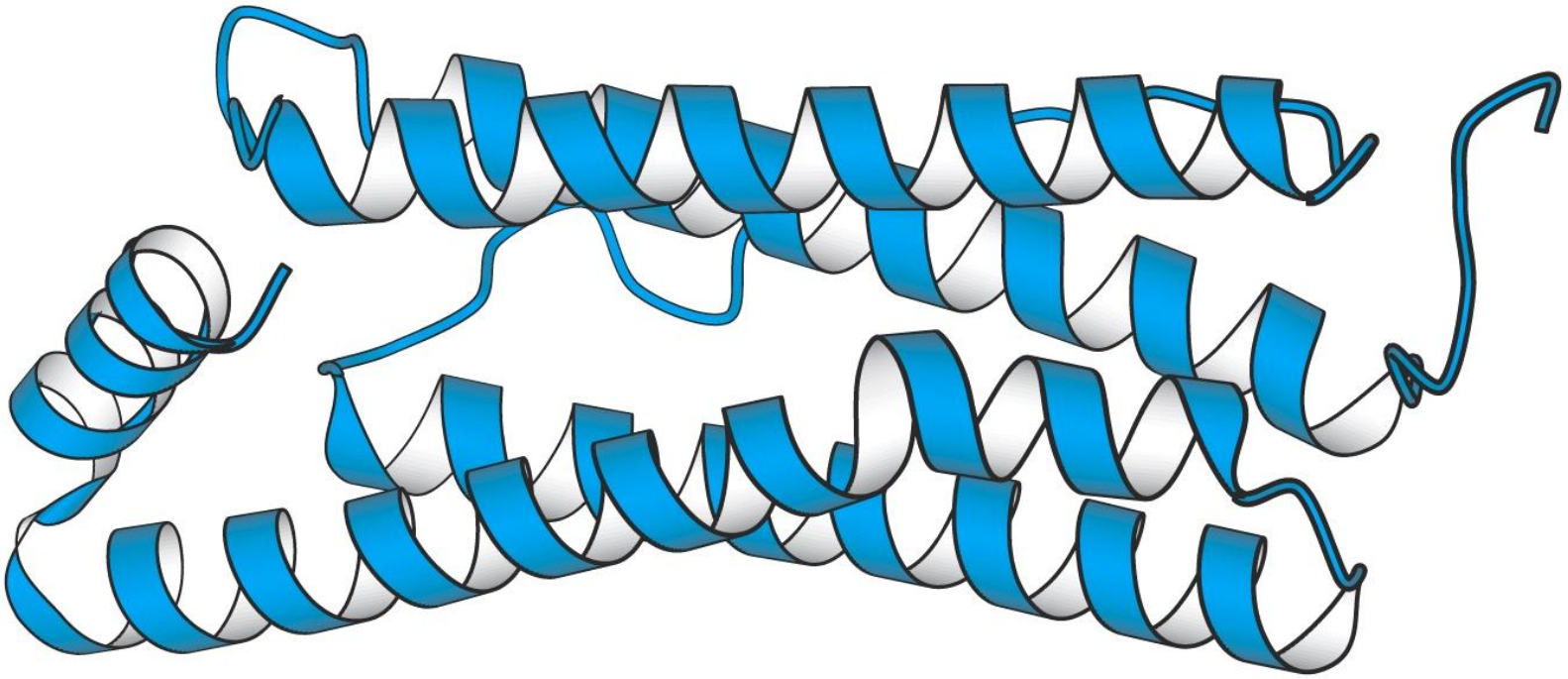
Figure 4-5
Lehninger Principles of Biochemistry, Fifth Edition

Helix dipole

α-helical coiled coil



The two helices wind around one another to form a superhelix. Such structures are found in many proteins including keratin in hair, quills, and horns.

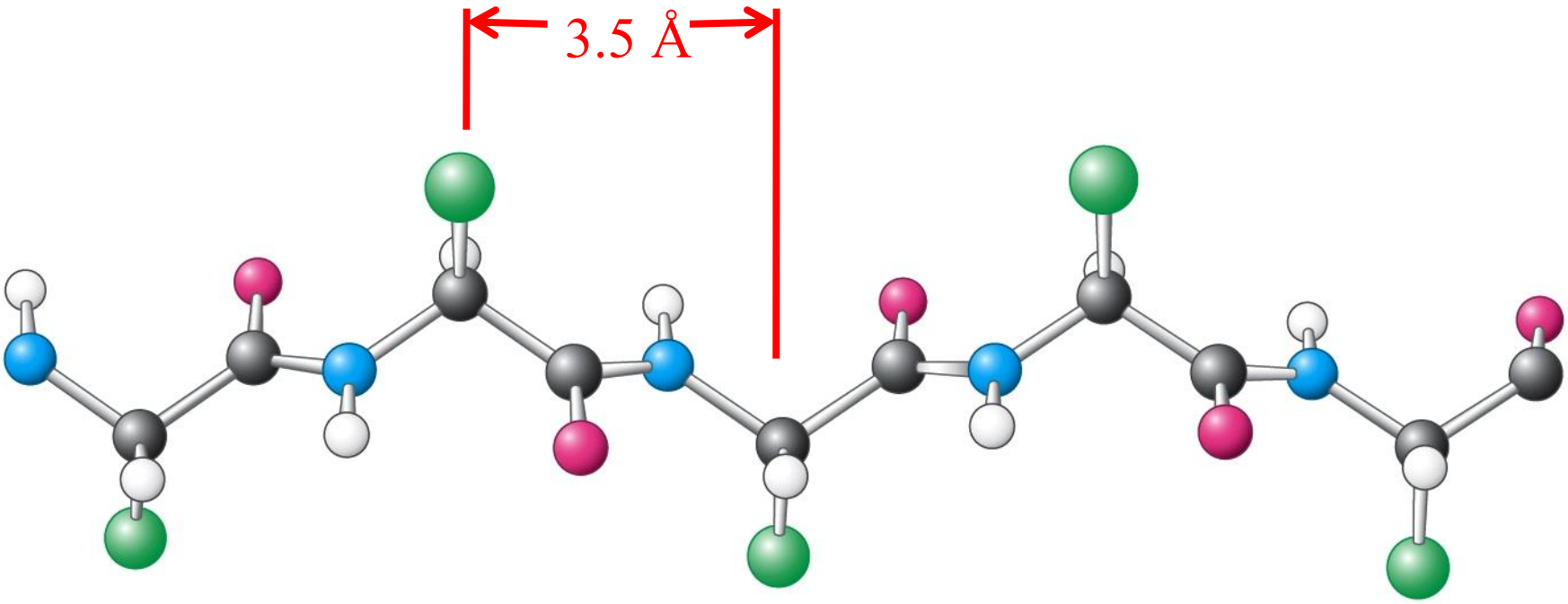


Ferritin, an iron-storage protein, is built from a bundle of α helices.

β Sheet

- **The backbone is more extended with the ψ dihedral (N–C_α—C–N) in the range (90° < ψ < 180°)**
- **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 more distal backbone amides**
- **Side chains protrude from the sheet alternating in up and down direction**

β Sheet



The side chains (green) are alternately above and below the plane of the strand.

Parallel and antiparallel β Sheets

- Parallel or antiparallel orientation of two chains within a sheet are possible
- In parallel β sheets the H-bonded strands run in the same direction
- In antiparallel β sheets the H-bonded strands run in opposite directions
- Which one is more stable???

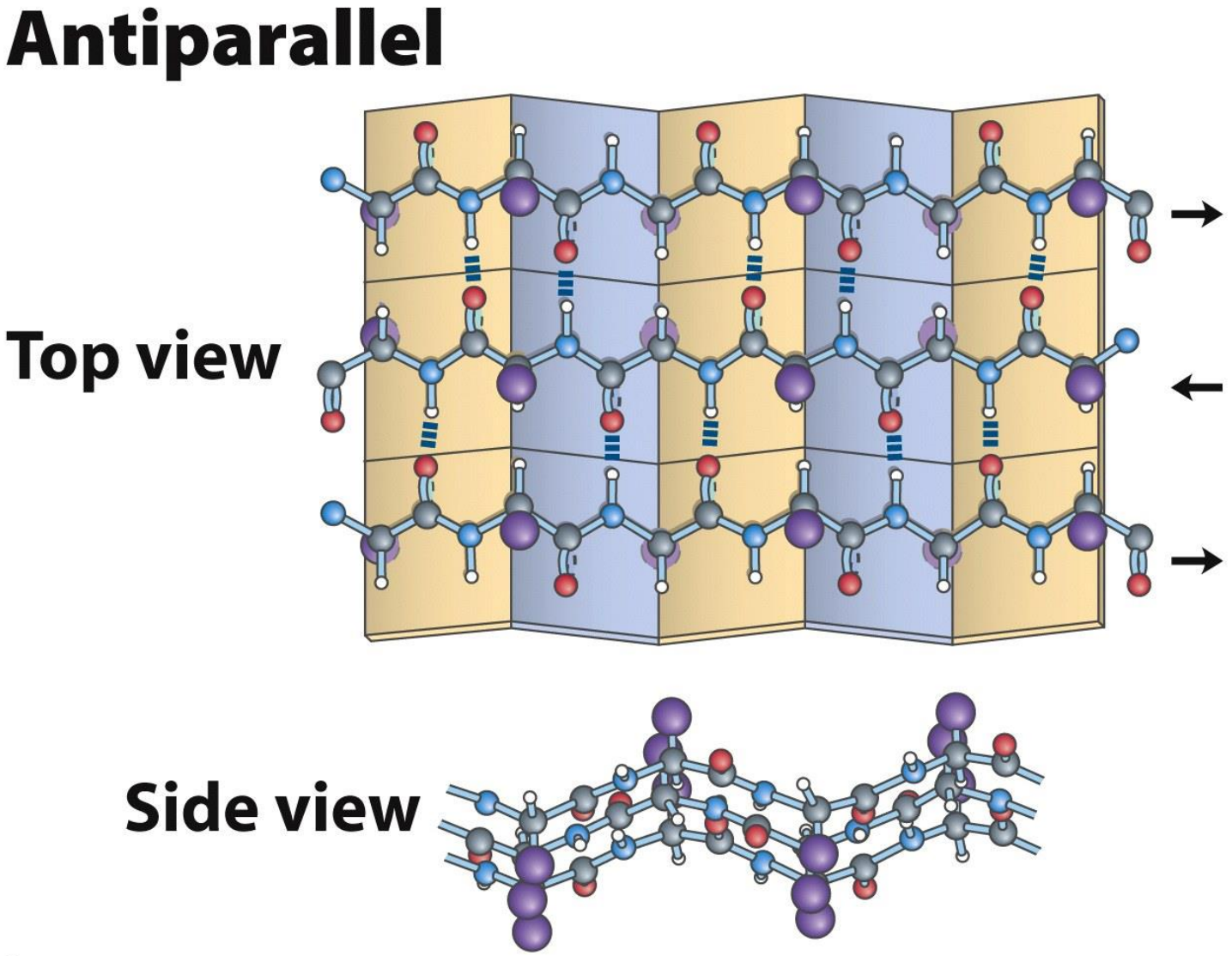


Figure 4-6a
Lehninger Principles of Biochemistry, Fifth Edition

The β conformation of polypeptide chains

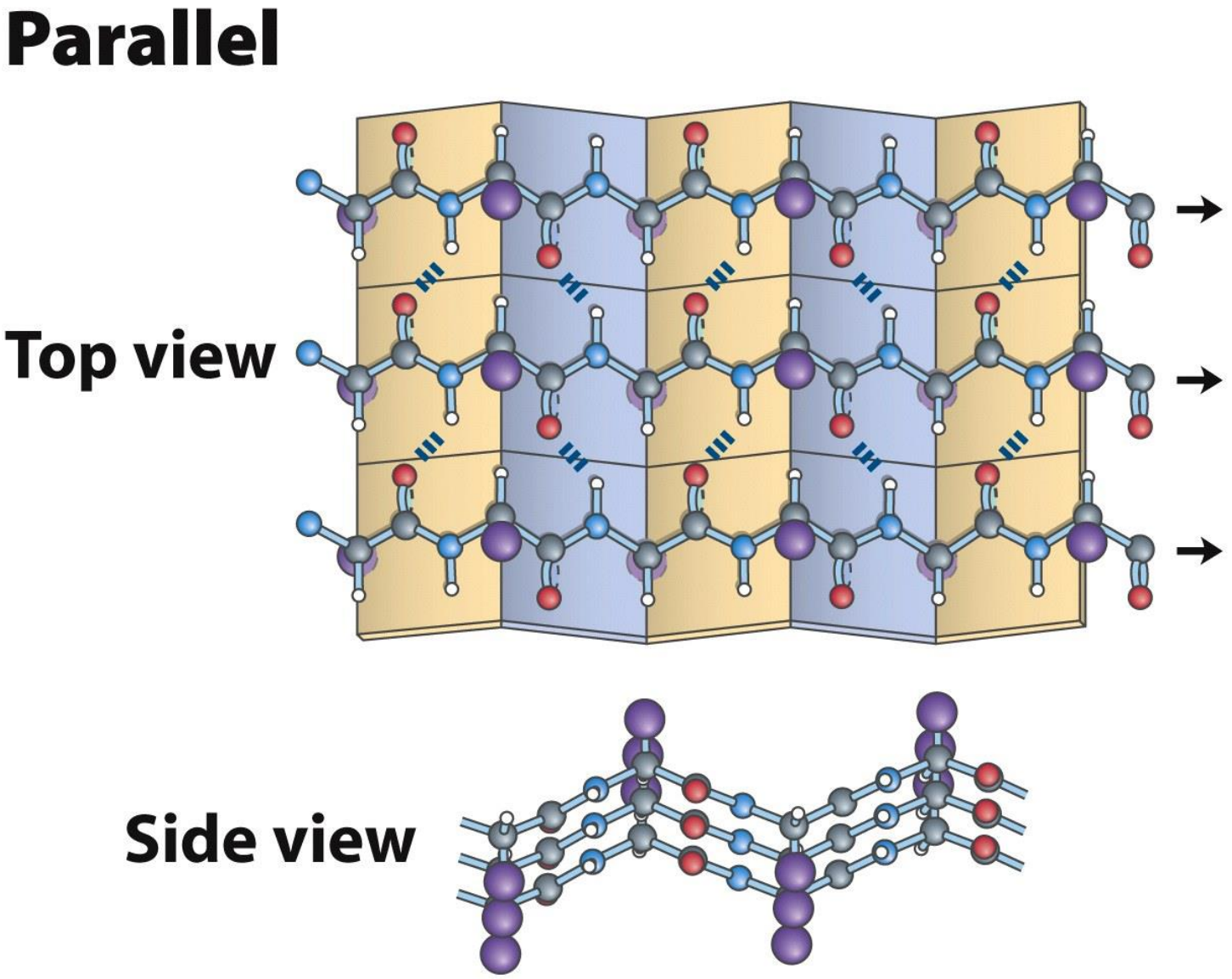


Figure 4-6b
Lehninger Principles of Biochemistry, Fifth Edition

The β conformation of polypeptide chains

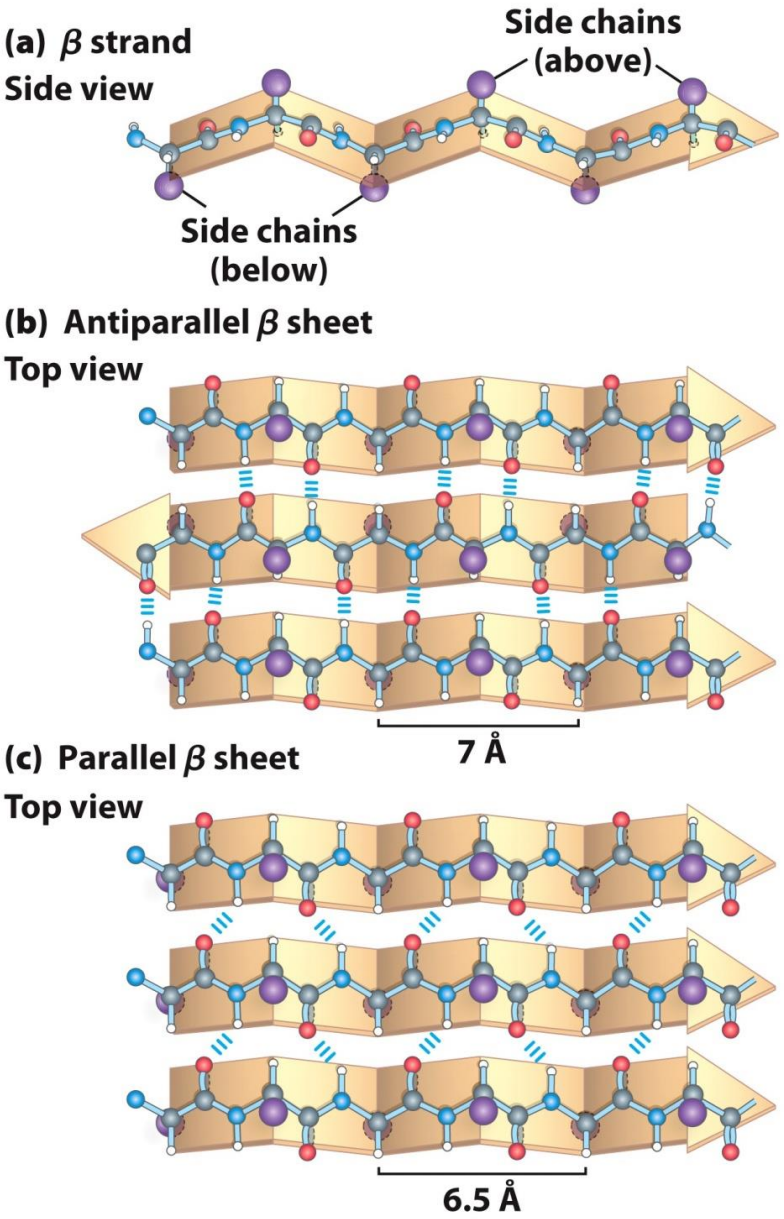
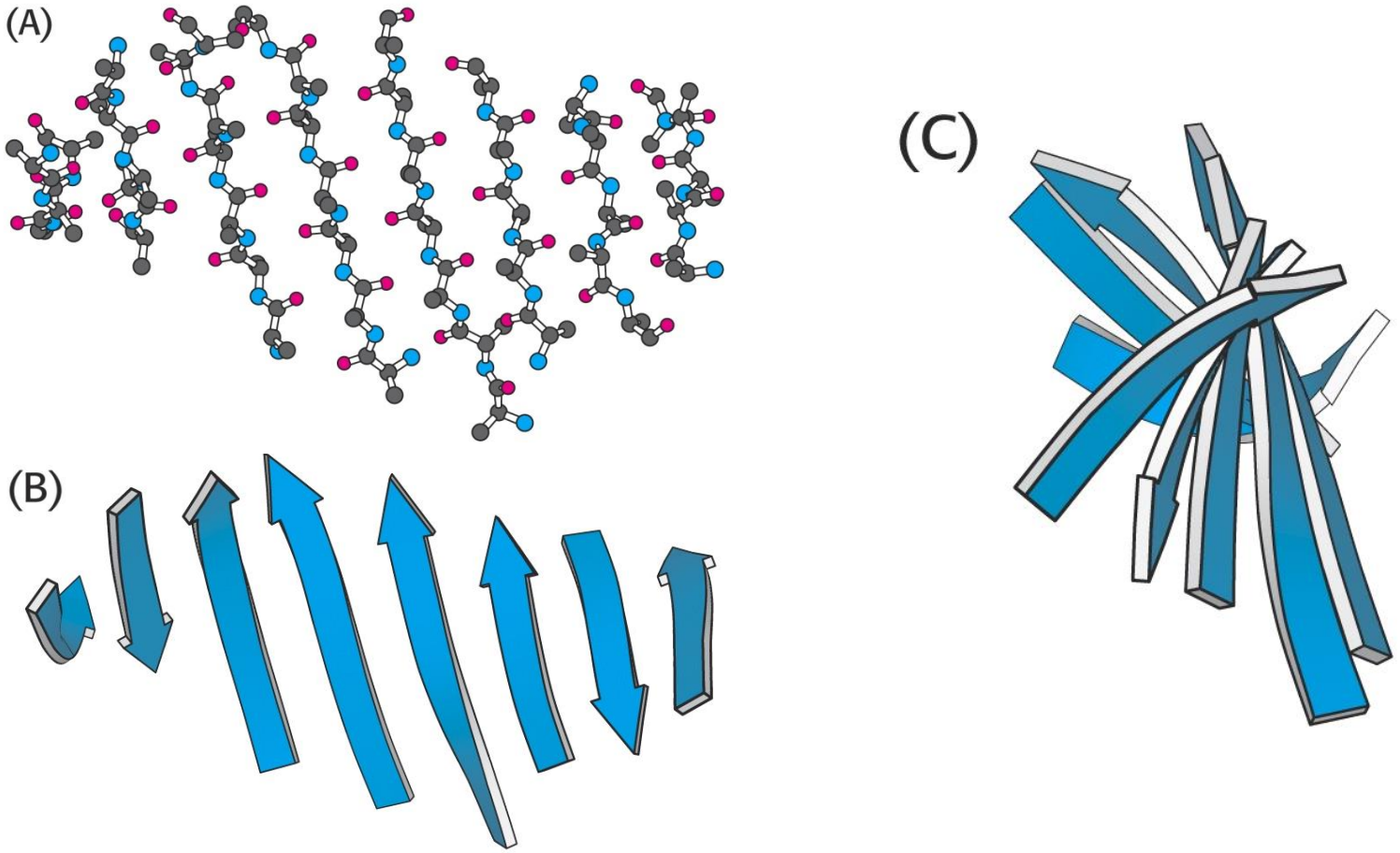
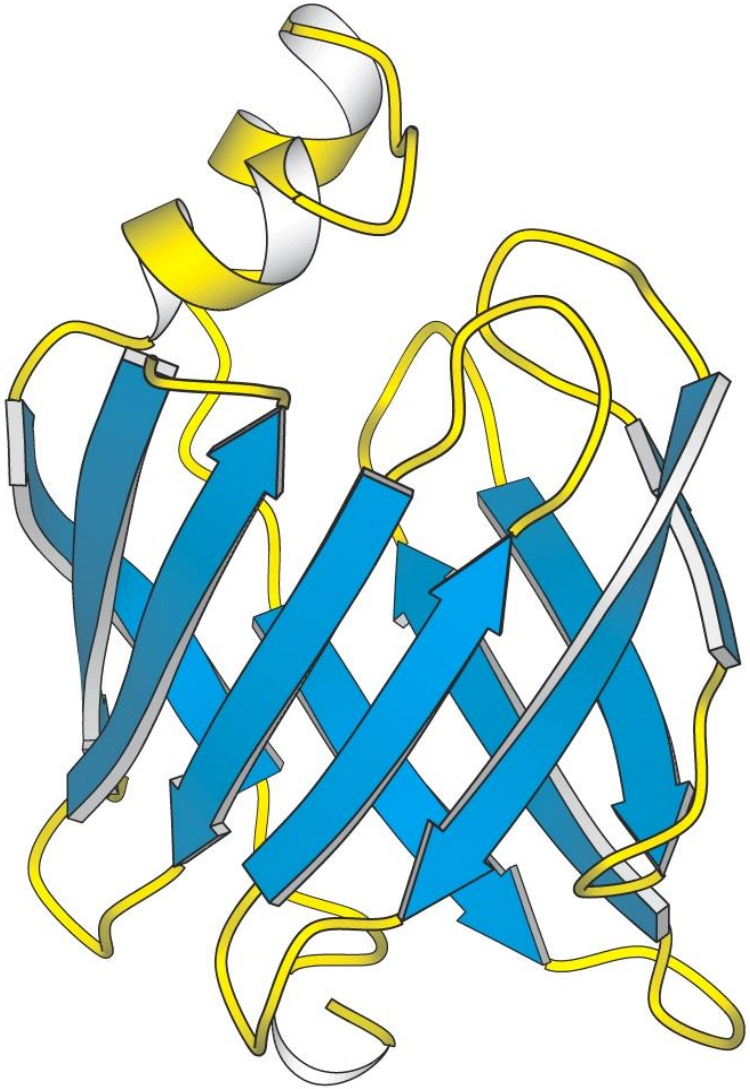


Figure 4-6
Lehninger Principles of Biochemistry, Sixth Edition

The β conformation of polypeptide chains



A twisted β sheet. (A) A ball-and-stick model. (B) A schematic model. (C) The schematic view rotated by 90 degrees to illustrate the twist more clearly.



The structure of a fatty acid-binding protein, which is rich in β sheets.

β Turns

- β-turns occur frequently whenever strands in β sheets change the direction
- 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
- Proline in position 2 (Type I) or glycine in position 3 (Type II) are common in β-turns

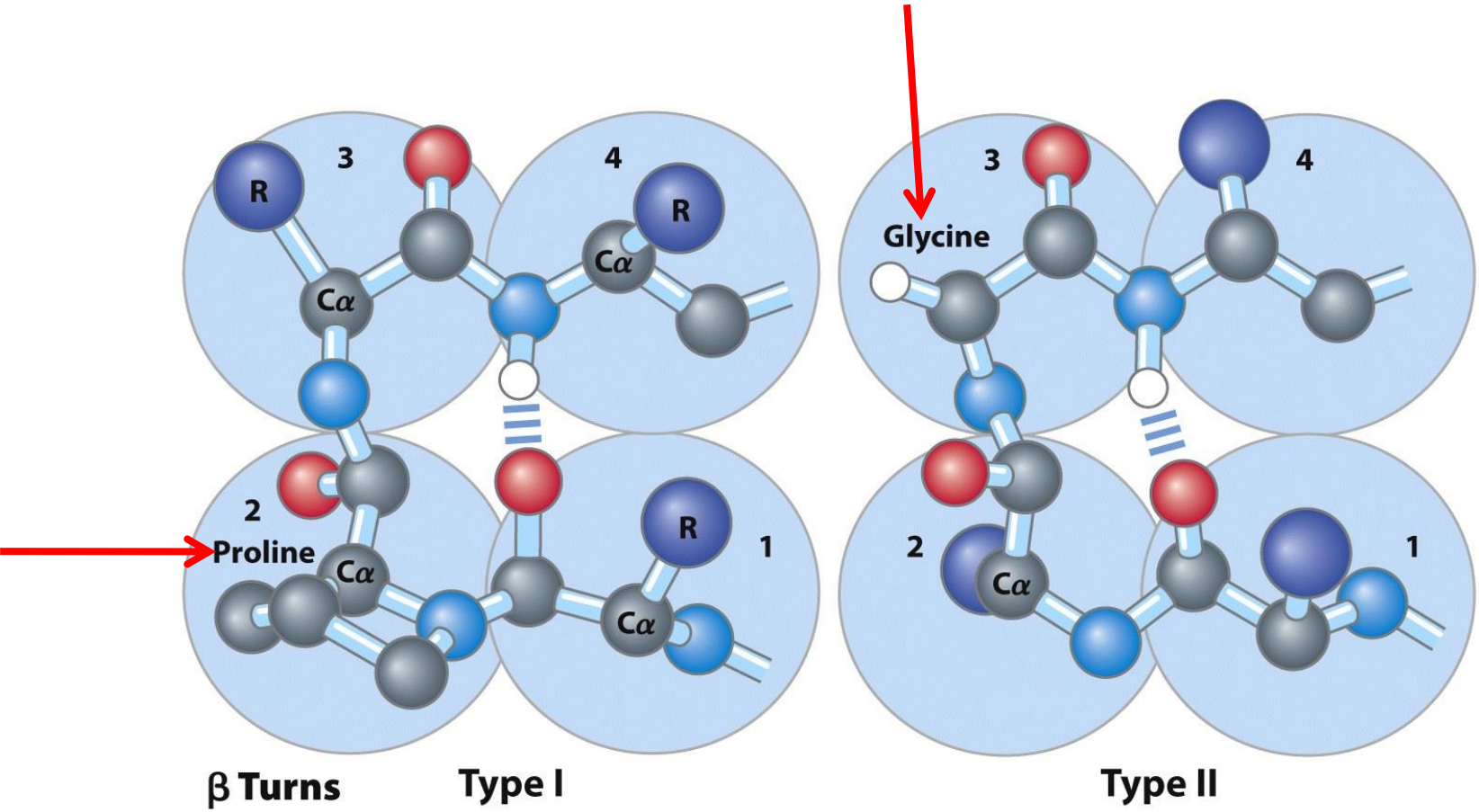


Figure 4-7a
Lehninger Principles of Biochemistry, Fifth Edition

Structures of β turns

Note the hydrogen bond between the peptide groups of the first and fourth residues of the bends

4. Protein tertiary structure

- **Tertiary structure refers to the overall spatial arrangement of all atoms in a polypeptide chain or in a protein**

- **Two major classes of proteins:**

  – fibrous proteins: made from a single type of secondary structure, insoluble in H_2O

  – globular proteins: made from several types of secondary structure in the same polypeptide chain

Structure of α-keratin in hair

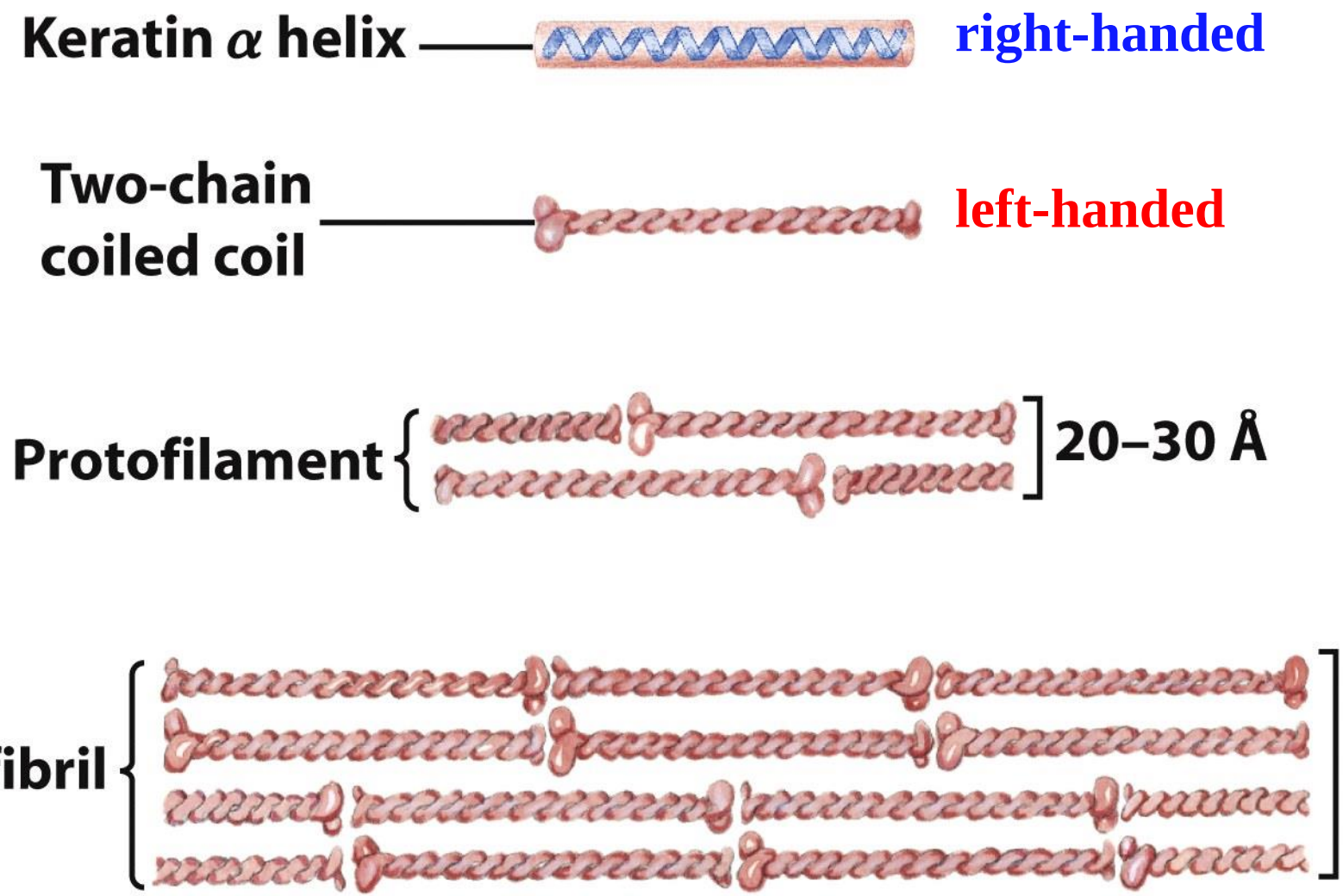


Figure 4-10a
Lehninger Principles of Biochemistry, Fifth Edition

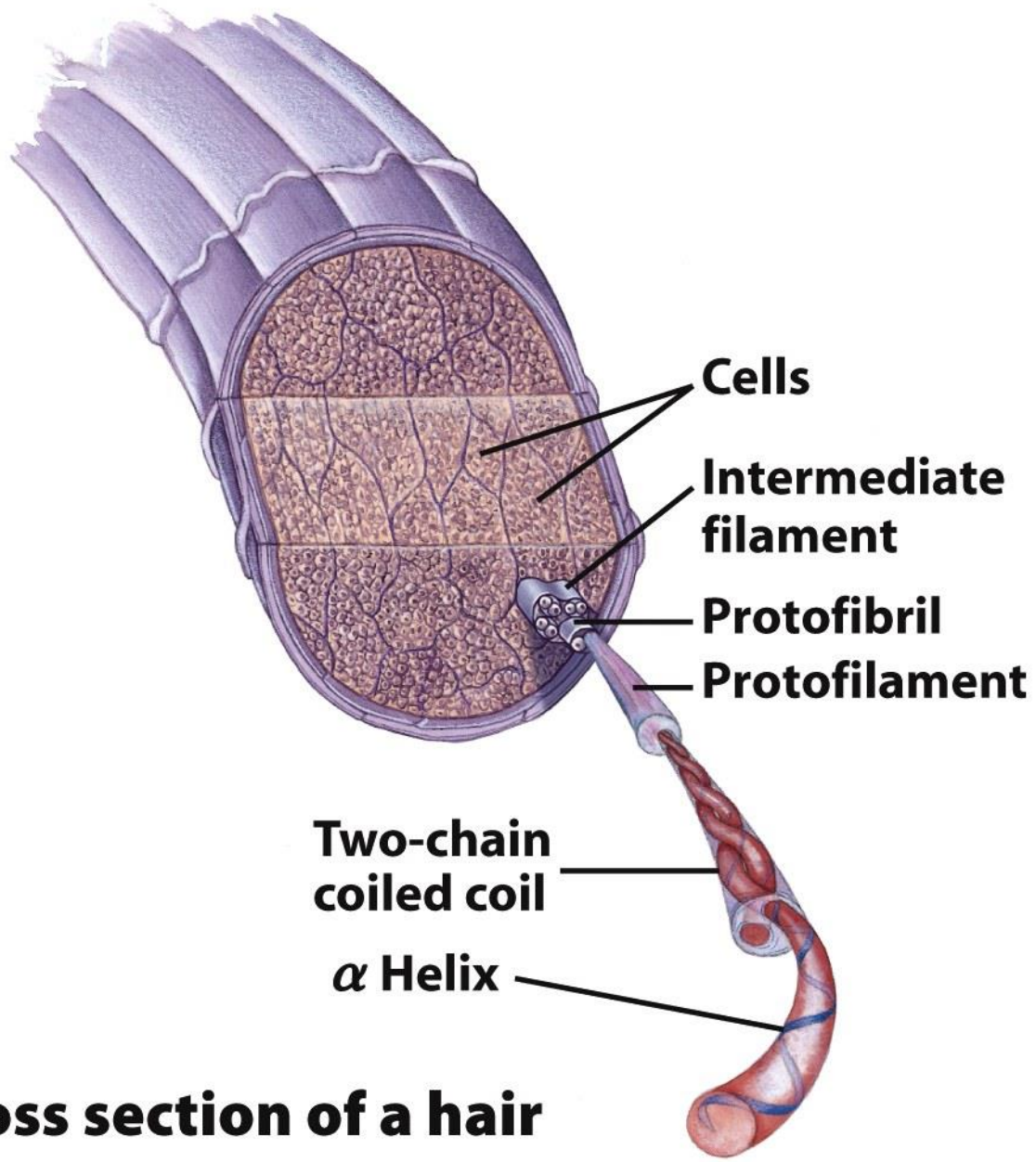
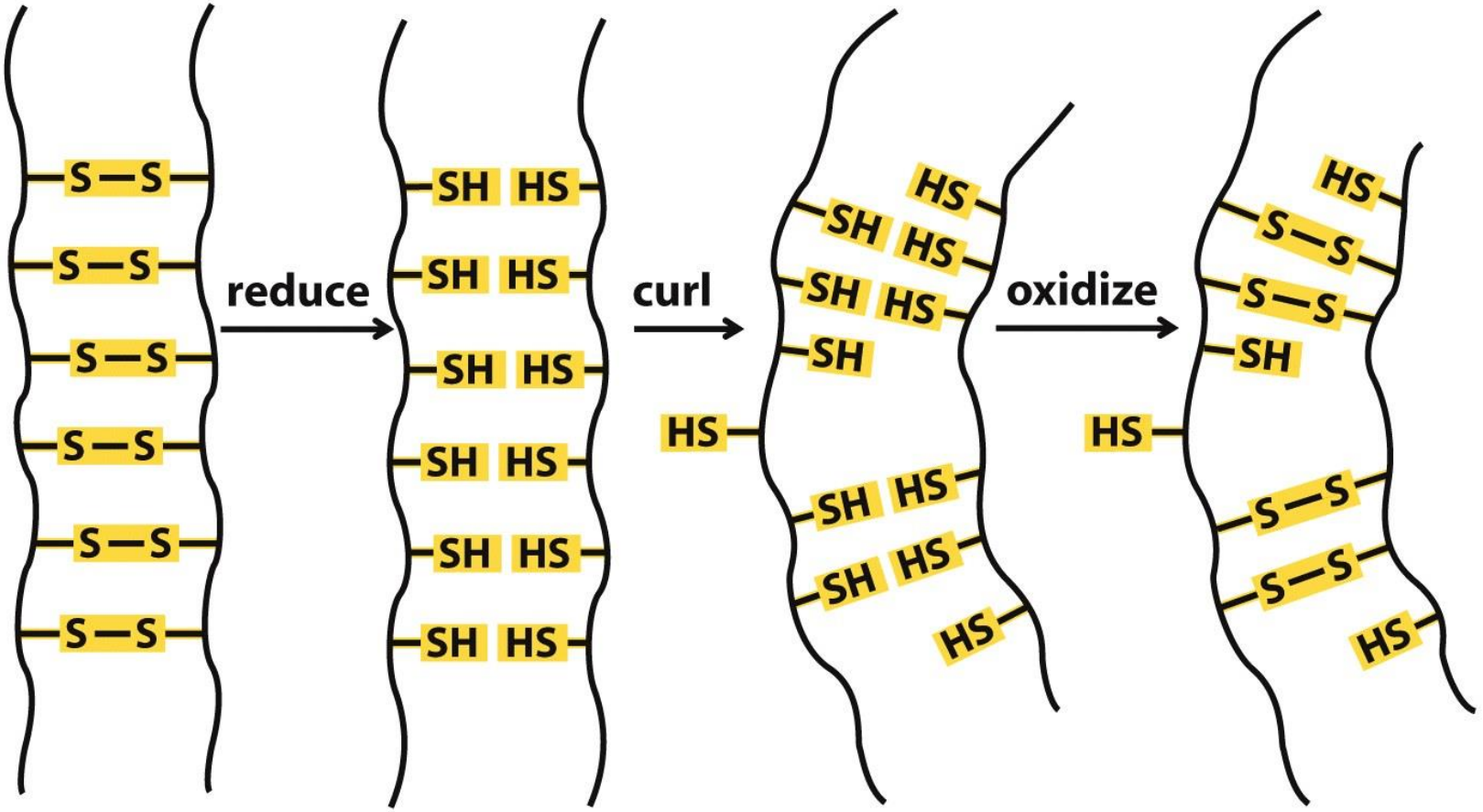


Figure 4-10b
Lehninger Principles of Biochemistry, Fifth Edition

Structure of hair

Chemistry of permanent waving



Box 4-2
Lehninger Principles of Biochemistry, Fifth Edition

The stretchability of α-keratins and their numerous disulfide cross-linkages are the basis of permanent waving

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 with (Gly-Pro-4-Hyp) repeats**
- **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**
- **Many triple-helices assemble into a collagen fibril**

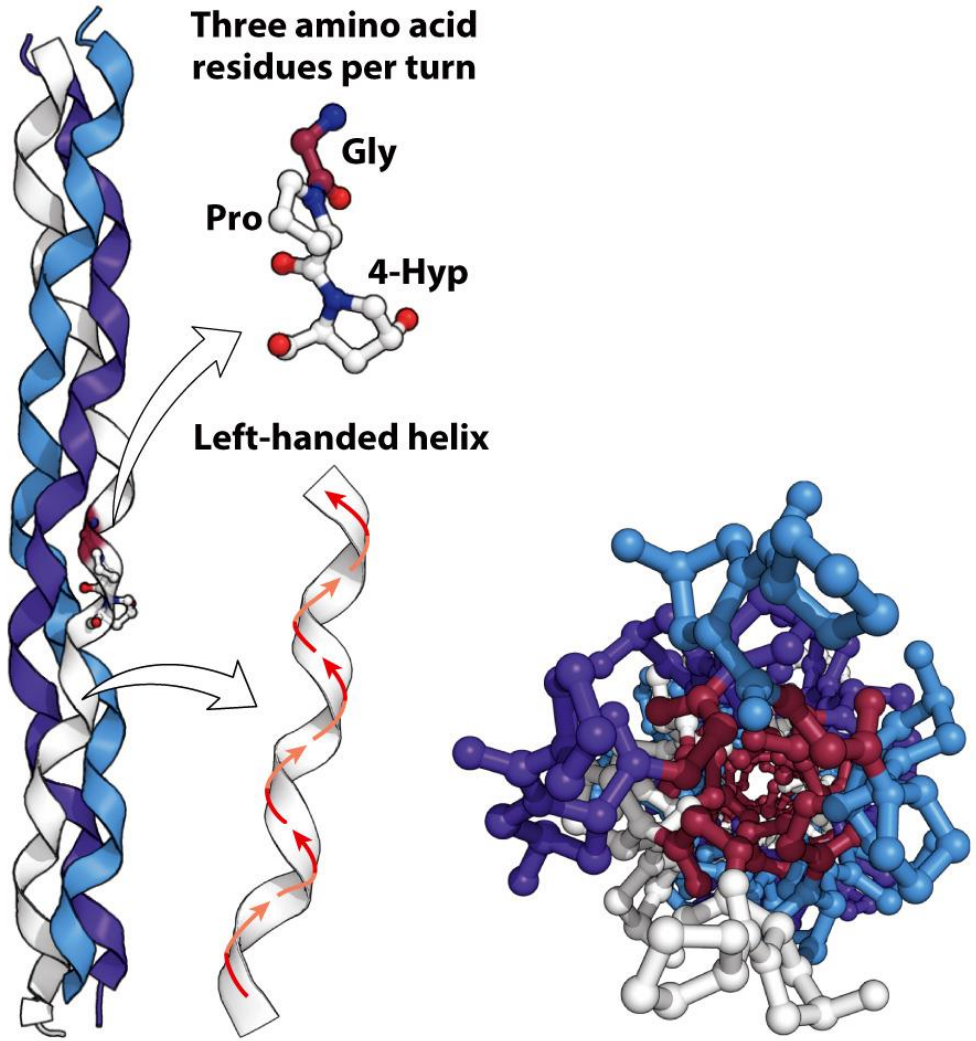


Figure 4-12
Lehninger Principles of Biochemistry, Seventh Edition

Structure of collagen

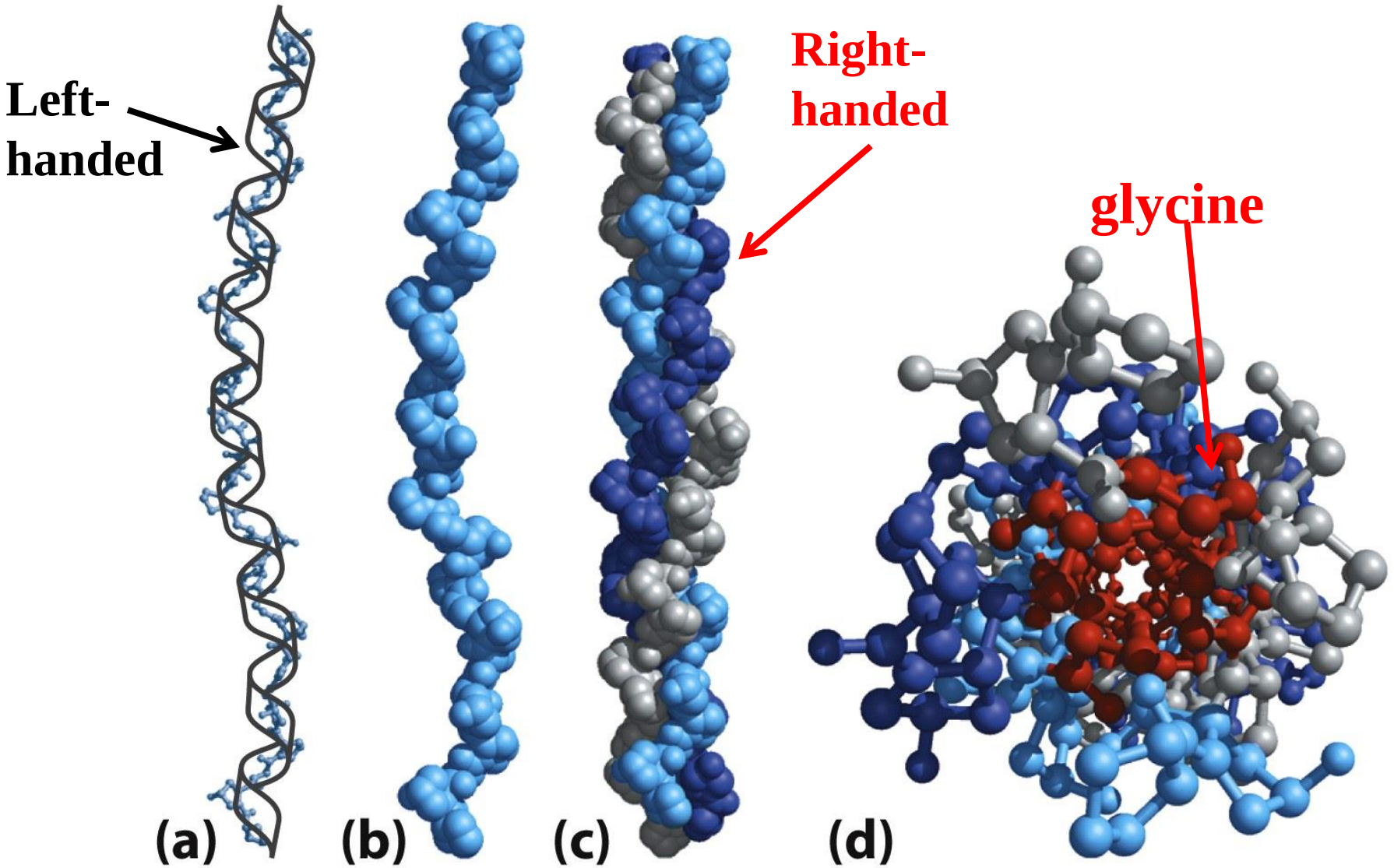


Figure 4-11
Lehninger Principles of Biochemistry, Fifth Edition

Structure of collagen

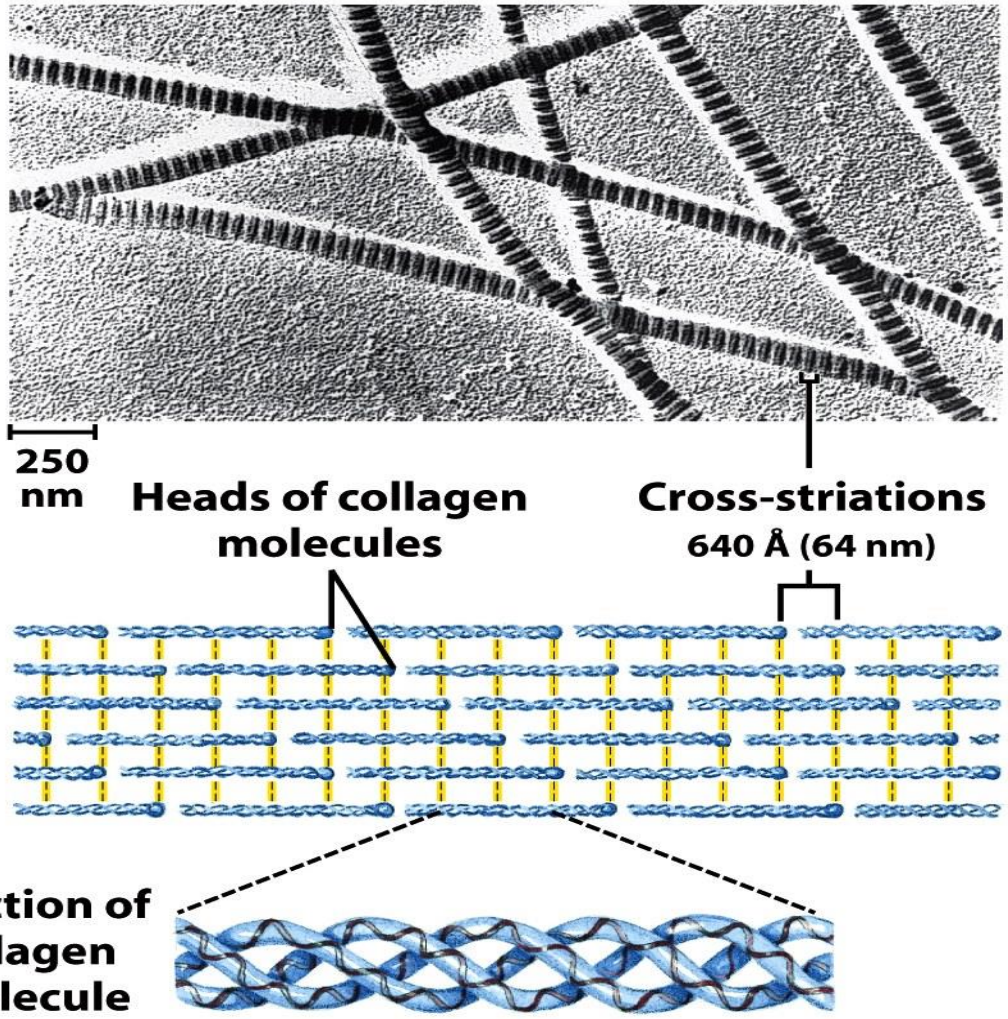


Figure 4-12
Lehninger Principles of Biochemistry, Fifth Edition

Structure of collagen fibrils

4-Hydroxyproline in collagen

- **Forces the proline ring into a favorable pucker**
- **Offer more hydrogen bonds between the three strands of collagen**
- **The post-translational processing is catalyzed by prolyl hydroxylase and requires α-ketoglutarate, molecular oxygen, and ascorbate (vitamin C)**
  (read Box 4-3 to know why you should eat fresh vegetables and fruits)

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**
- **Structure is stabilized by**
  - **hydrogen bonding**
  - **van der Waals interactions**

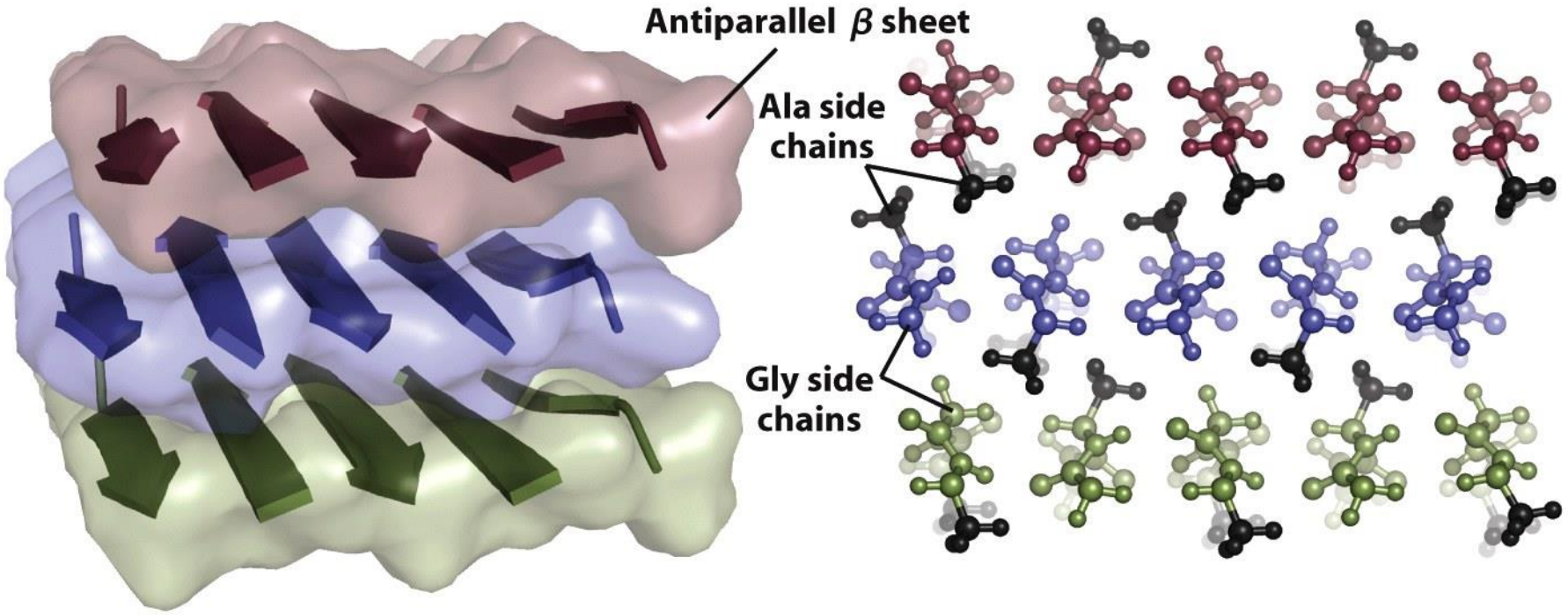


Figure 4-13a
Lehninger Principles of Biochemistry, Fifth Edition

Fibroin consists of layers of antiparallel β sheets rich in Ala and Gly residues

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**

TABLE 4–2 Secondary Structures and Properties of Some Fibrous Proteins *****

Structure	Characteristics	Examples of occurrence
α Helix, cross-linked by disulfide bonds	Tough, insoluble protective structures of varying hardness and flexibility	α-Keratin of hair, feathers, and nails
β Conformation	Soft, flexible filaments	Silk fibroin
Collagen triple helix	High tensile strength, without stretch	Collagen of tendons, bone matrix

Table 4-2
Lehninger Principles of Biochemistry, Fifth Edition

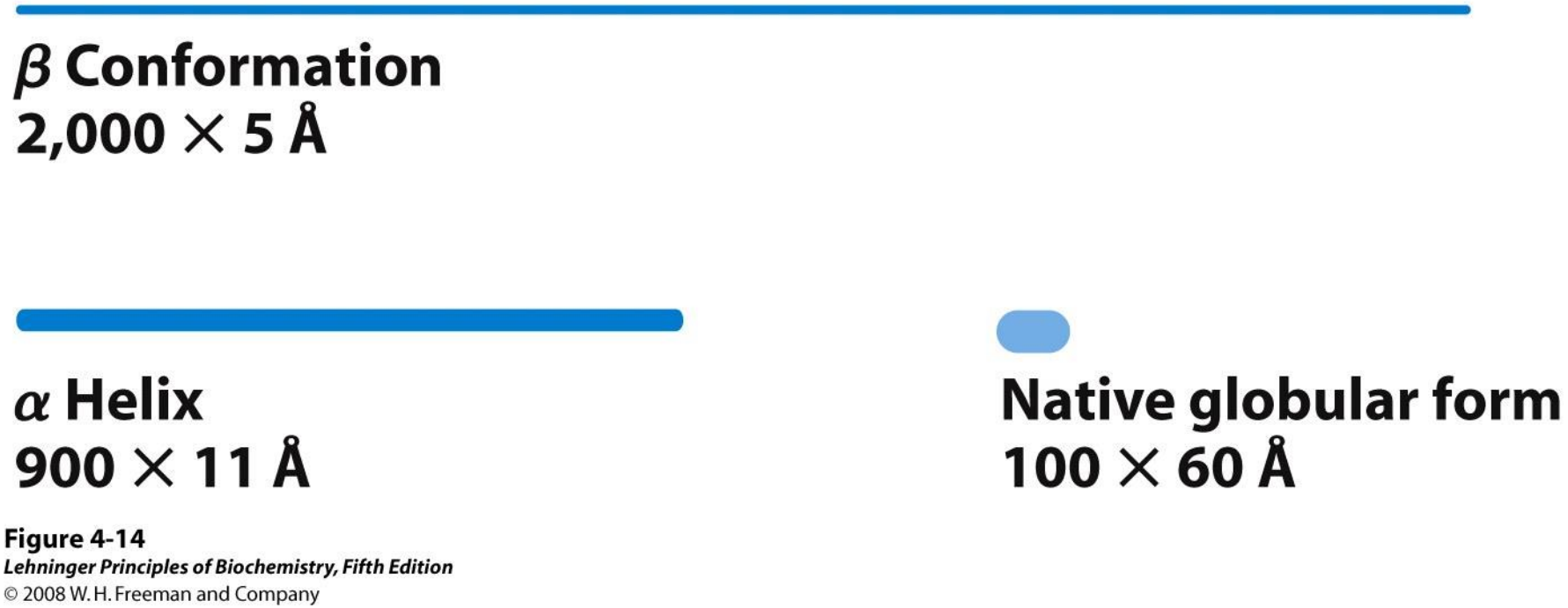


Figure 4-14
Lehninger Principles of Biochemistry, Fifth Edition

Human serum albumin: 585 residues, M_r 64,500

Globular protein structures are compact and varied

Myoglobin

- **The oxygen carrier in muscle**
- **Single polypeptide chain, 153 amino acids, M_r 16,700**
- **Heme: prosthetic group, capacity to bind oxygen**
- **Extremely compact; 70% main chain folded into 8 α helices (right-handed)**

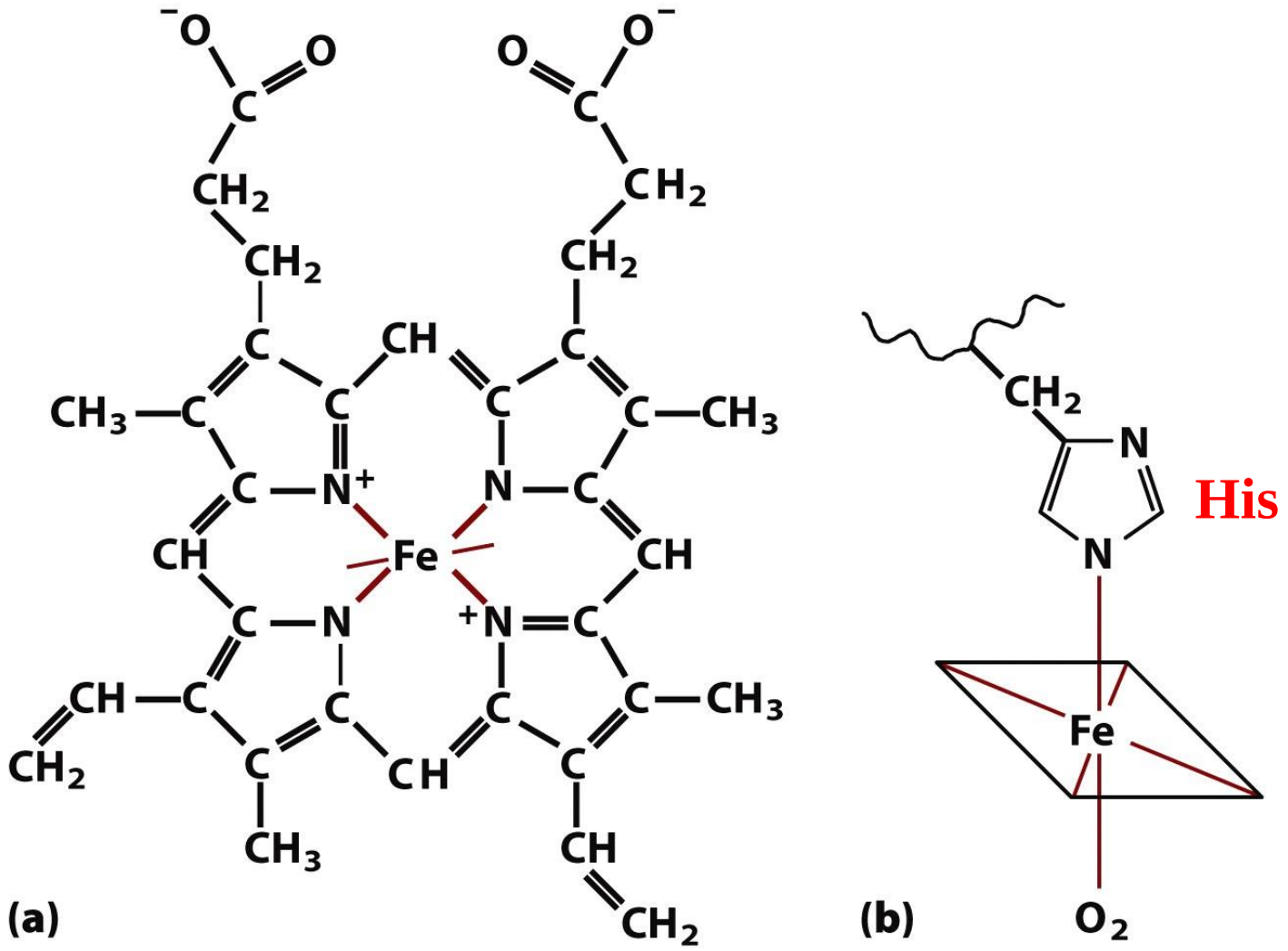


Figure 4-16
Lehninger Principles of Biochemistry, Fifth Edition

The heme group

Tertiary structure of myoglobin

- **Interior consists almost entirely of nonpolar residues**
- **The only polar residues inside are 2 His, which play critical roles in binding iron and oxygen**
- **Charged residues are on the outer surface**
- **Little empty space inside**

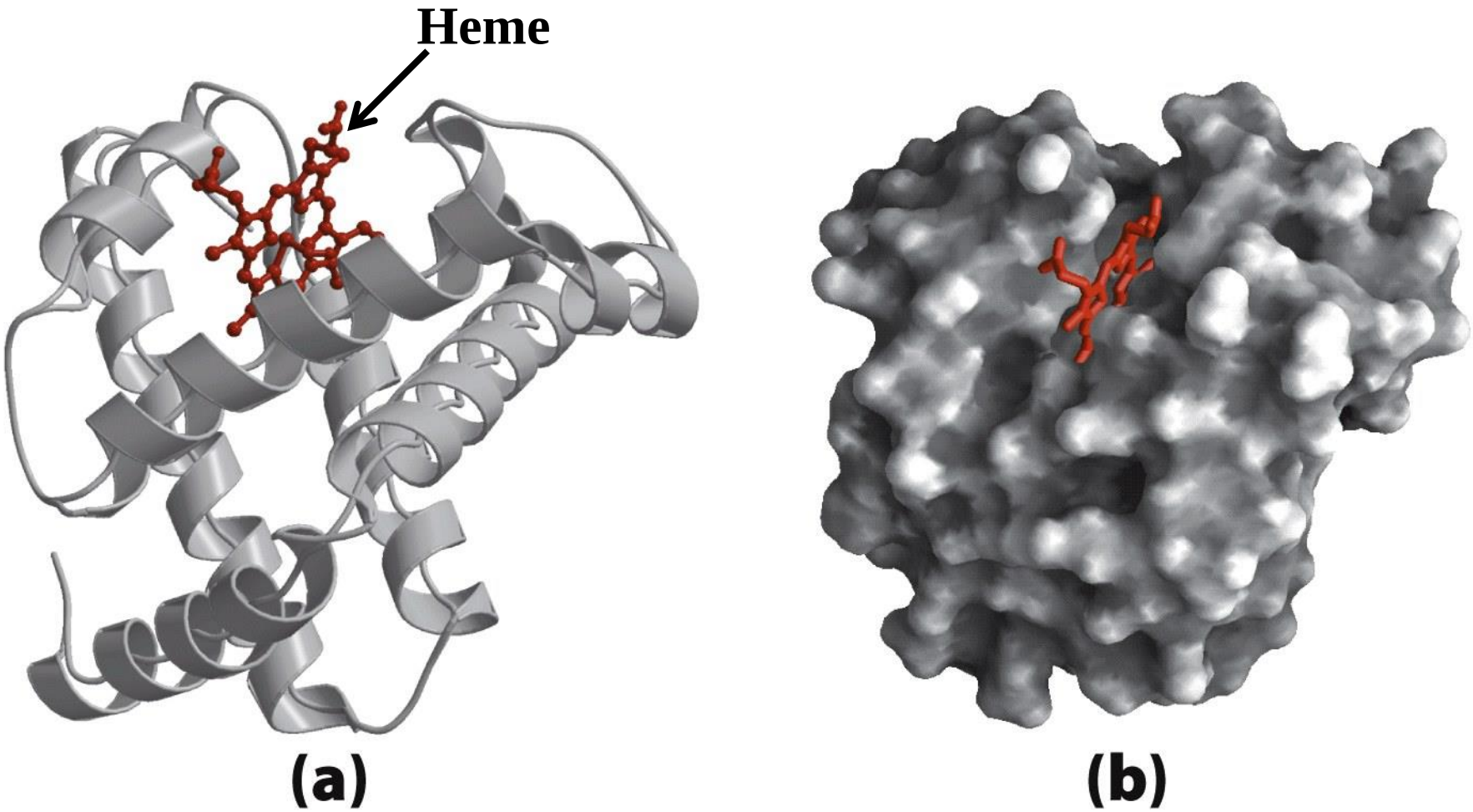


Figure 4-15ab
Lehninger Principles of Biochemistry, Fifth Edition

Tertiary structure of sperm whale myoglobin

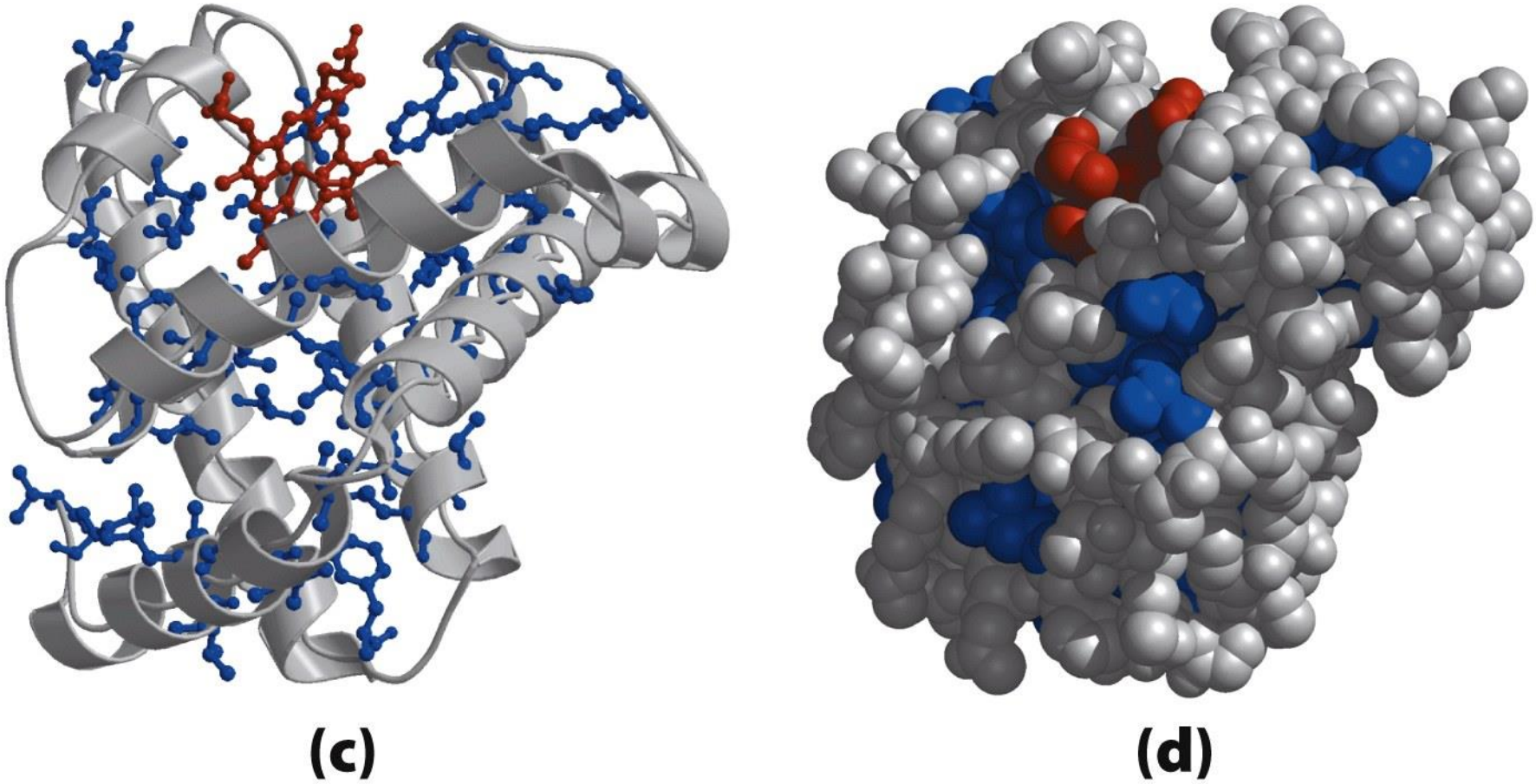
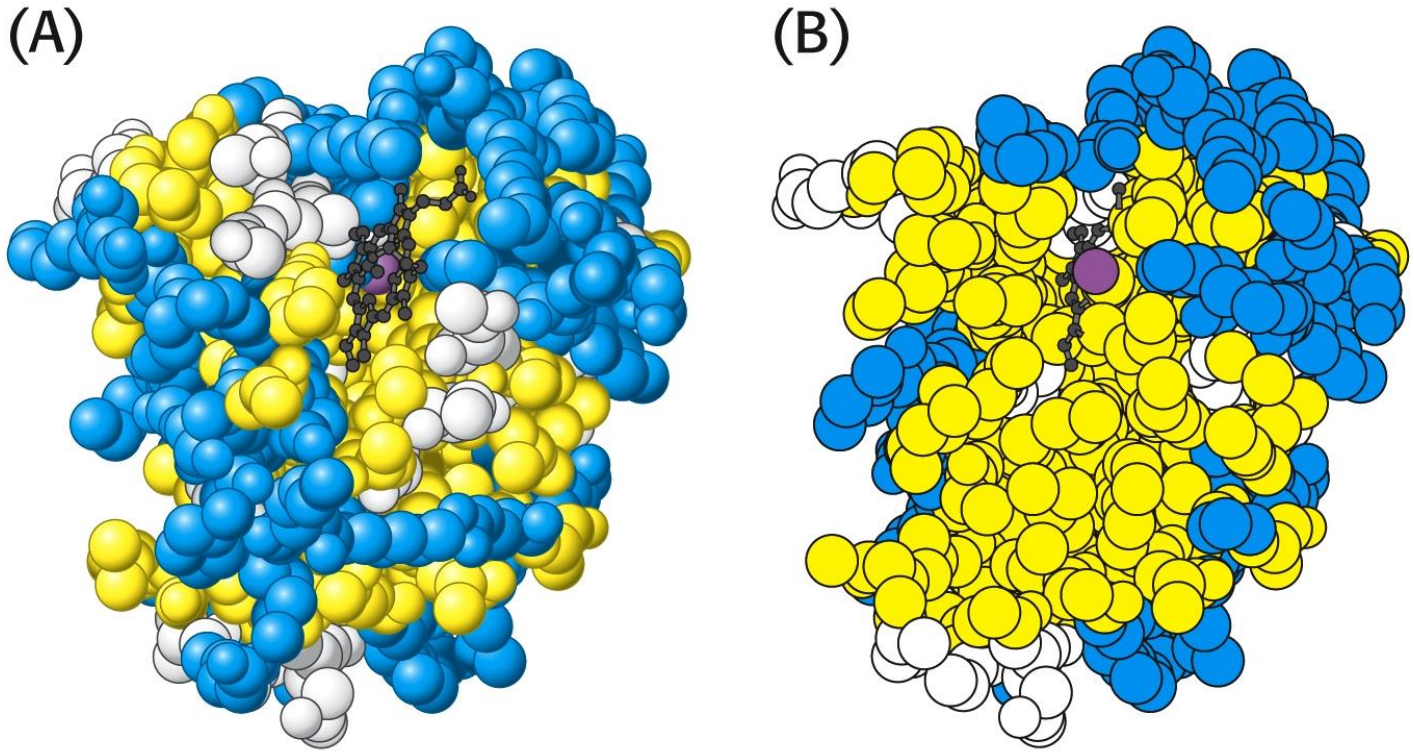


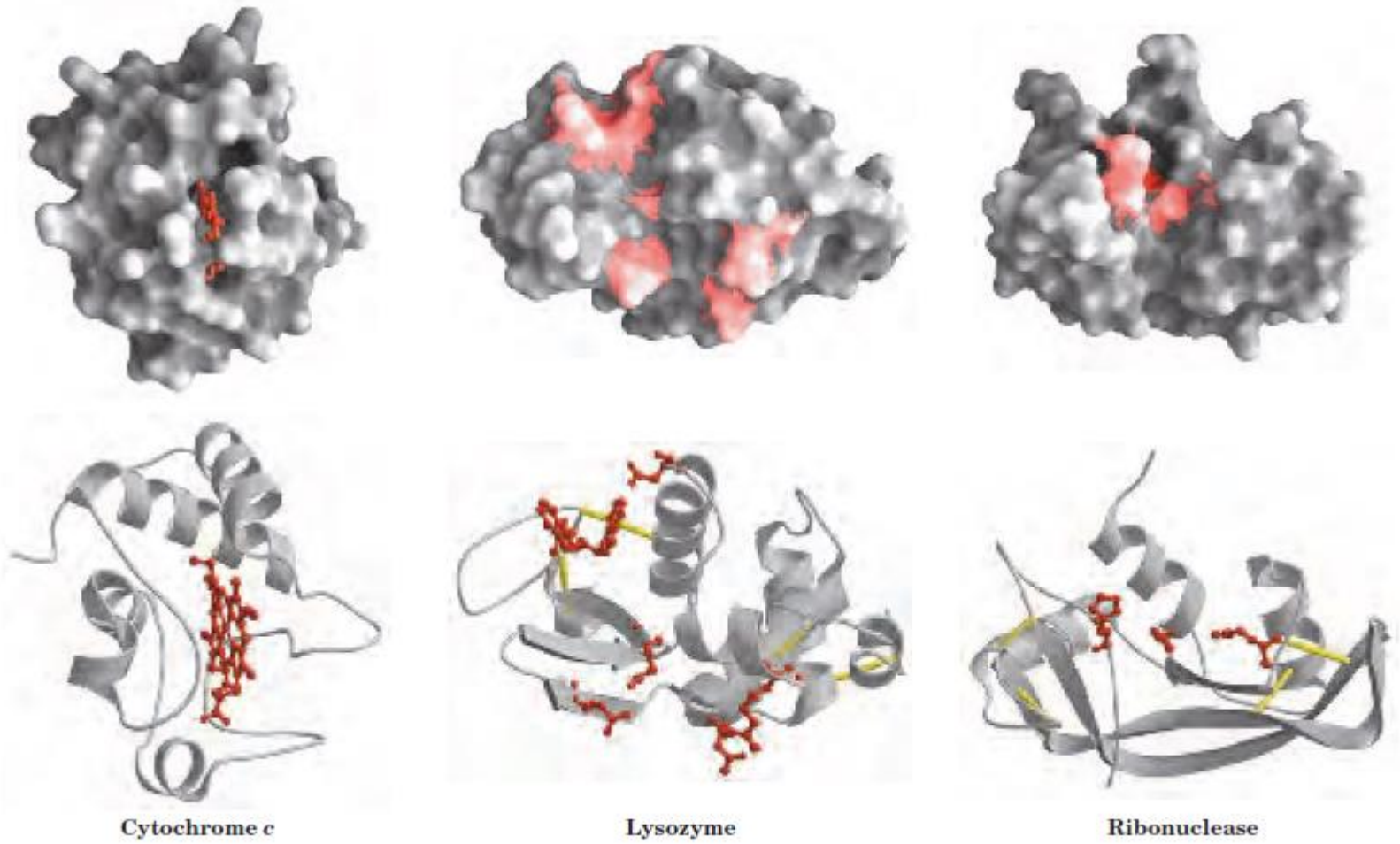
Figure 4-15cd
Lehninger Principles of Biochemistry, Fifth Edition

Tertiary structure of sperm whale myoglobin



Distribution of amino acids in myoglobin

(A) A space-filling model of myoglobin with hydrophobic amino acids shown in yellow, charged amino acids shown in blue, and others shown in white. The surface of the molecule has many charged amino acids, as well as some hydrophobic amino acids. (B) A cross-sectional view shows that hydrophobic amino acids are found on the inside of the structure, whereas the charged amino acids are found on the protein surface.



Three-dimensional structures of some small proteins

TABLE 4–3 Approximate Proportion of α Helix and β Conformation in Some Single-Chain Proteins

Protein (total residues)	Residues (%)*	
	α Helix	β Conformation
Chymotrypsin (247)	14	45
Ribonuclease (124)	26	35
Carboxypeptidase (307)	38	17
Cytochrome *c* (104)	39	0
Lysozyme (129)	40	12
Myoglobin (153)	78	0

Source: Data from Cantor, C.R. & Schimmel, P.R. (1980) *Biophysical Chemistry,* Part I: *The Conformation of Biological Macromolecules,* p. 100, W. H. Freeman and Company, New York.

*Portions of the polypeptide chains not accounted for by α helix or β conformation consist of bends and irregularly coiled or extended stretches. Segments of α helix and β conformation sometimes deviate slightly from their normal dimensions and geometry.

Table 4-3

Lehninger Principles of Biochemistry, Fifth Edition

Motif and domain

- **Motif (also called supersecondary structure, or fold): a recognizable folding pattern involving two or more elements of secondary structure and the connections between them**
- **A motif may be simple or complex, and can represent all or just a small part of a protein**
- **Domain: A distinct region of a protein that can fold stably and independently**
- **Domain may have separate functions**

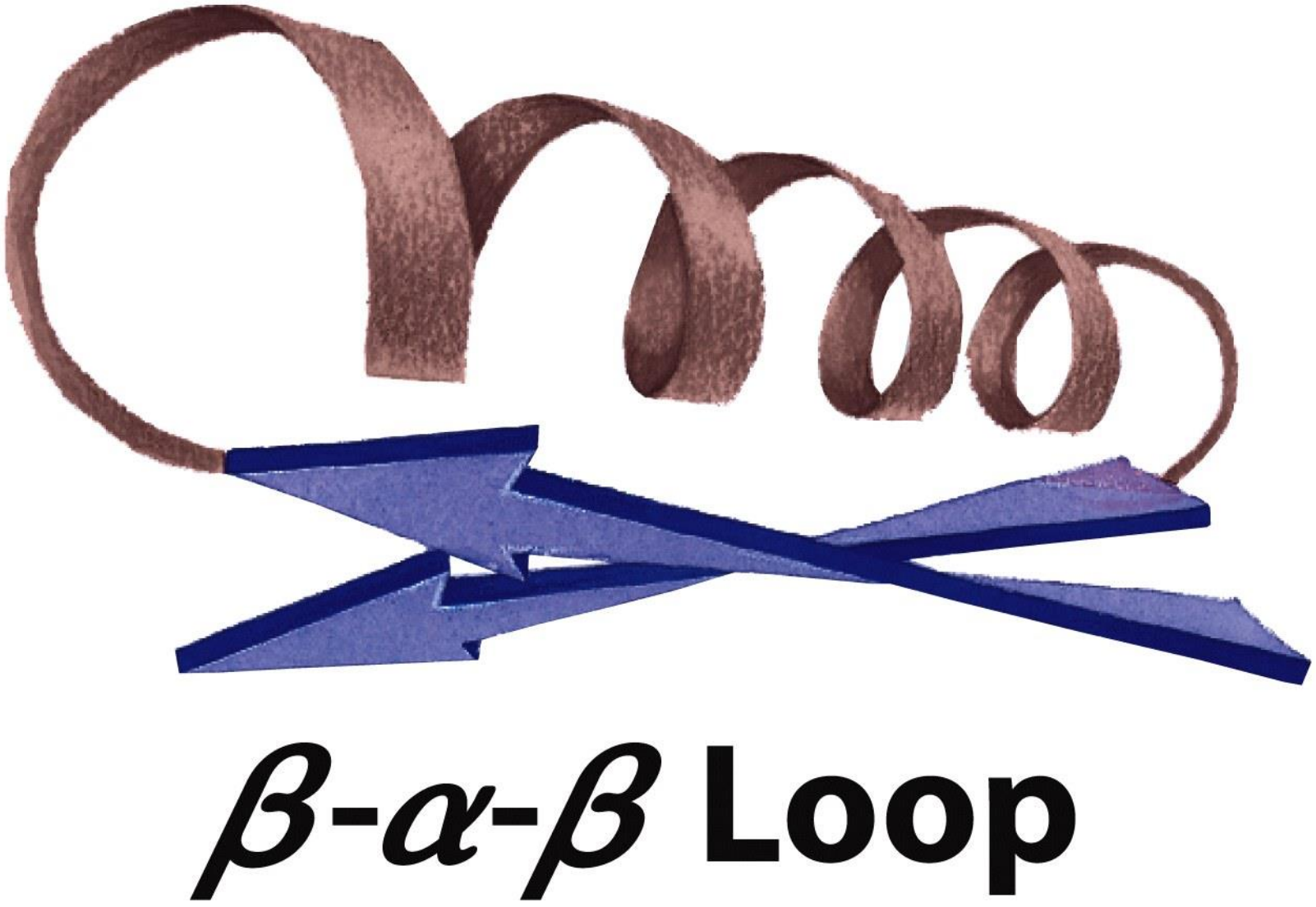


Figure 4-17a
Lehninger Principles of Biochemistry, Fifth Edition

A simple motif, the β-α-β loop

Figure 4-17b
Lehninger Principles of Biochemistry, Fifth Edition

This β barrel is a single domain of α-hemolysin (a toxin that kills a cell by creating a hole in its membrane) from the bacterium *Staphylococcus aureus*.

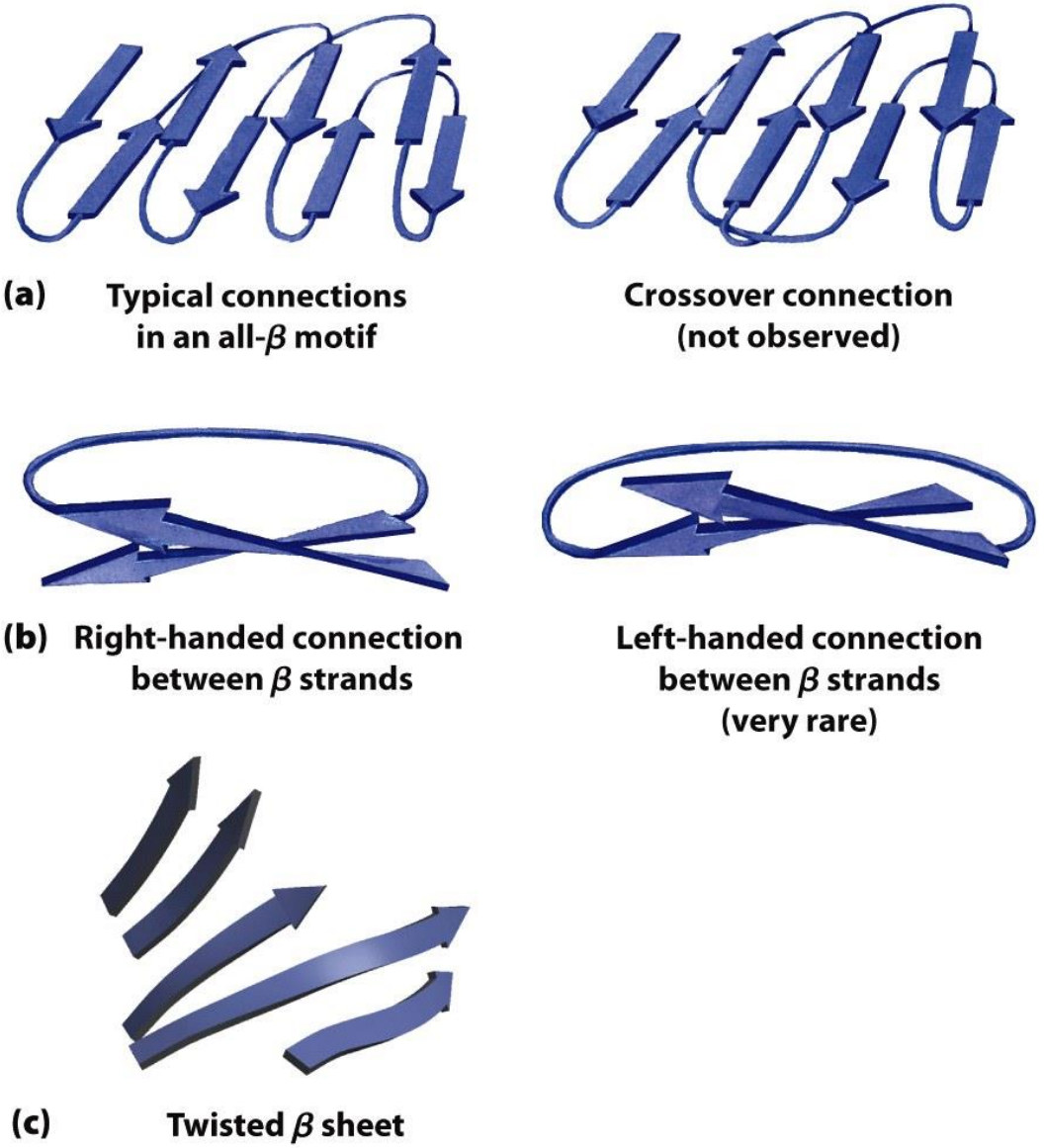


Figure 4-19
Lehninger Principles of Biochemistry, Fifth Edition

Stable folding patterns in proteins

Constructing large motifs from smaller ones. The α/β barrel is a commonly occurring motif constructed from repetitions of the α-β-α loop motif. This α/β barrel is a domain of pyruvate kinase.

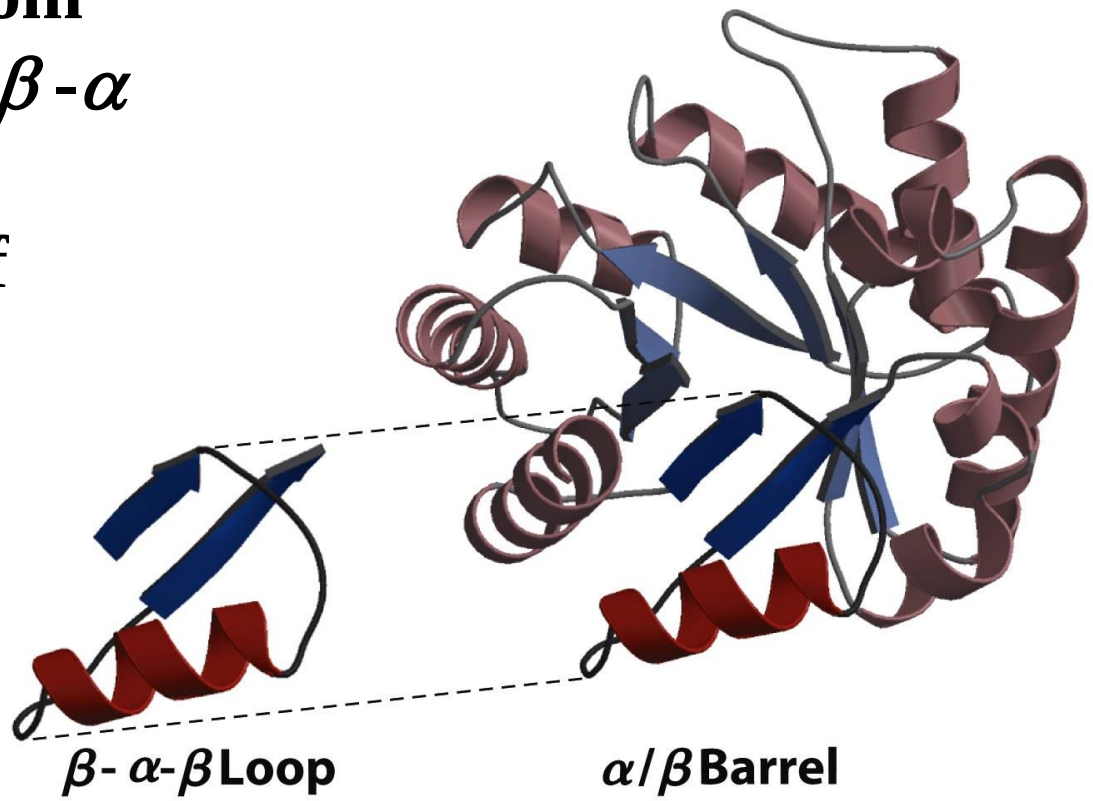


Figure 4-20
Lehninger Principles of Biochemistry, Fifth Edition

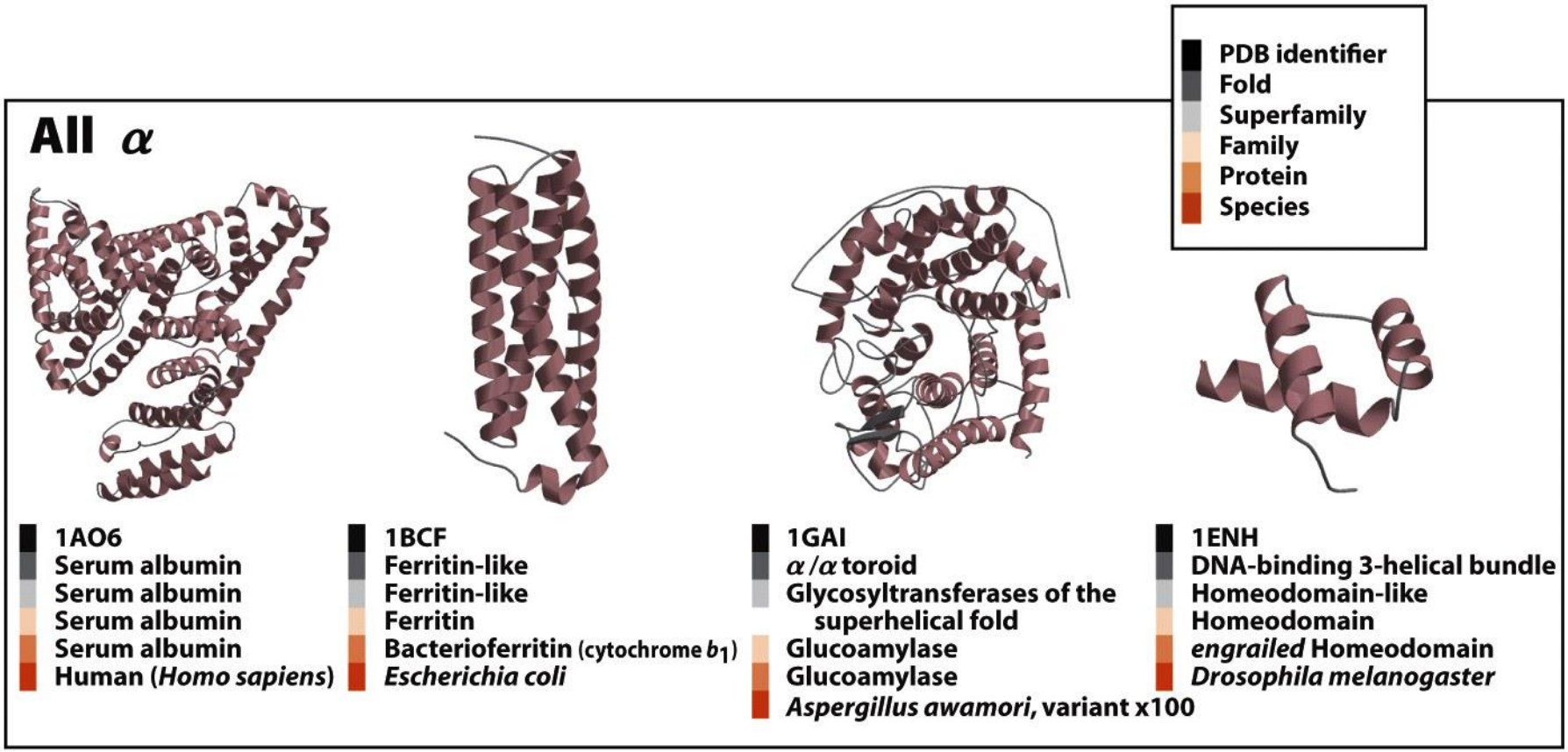


Figure 4-21 part 1

Lehninger Principles of Biochemistry, Fifth Edition

Organization of proteins based on motif

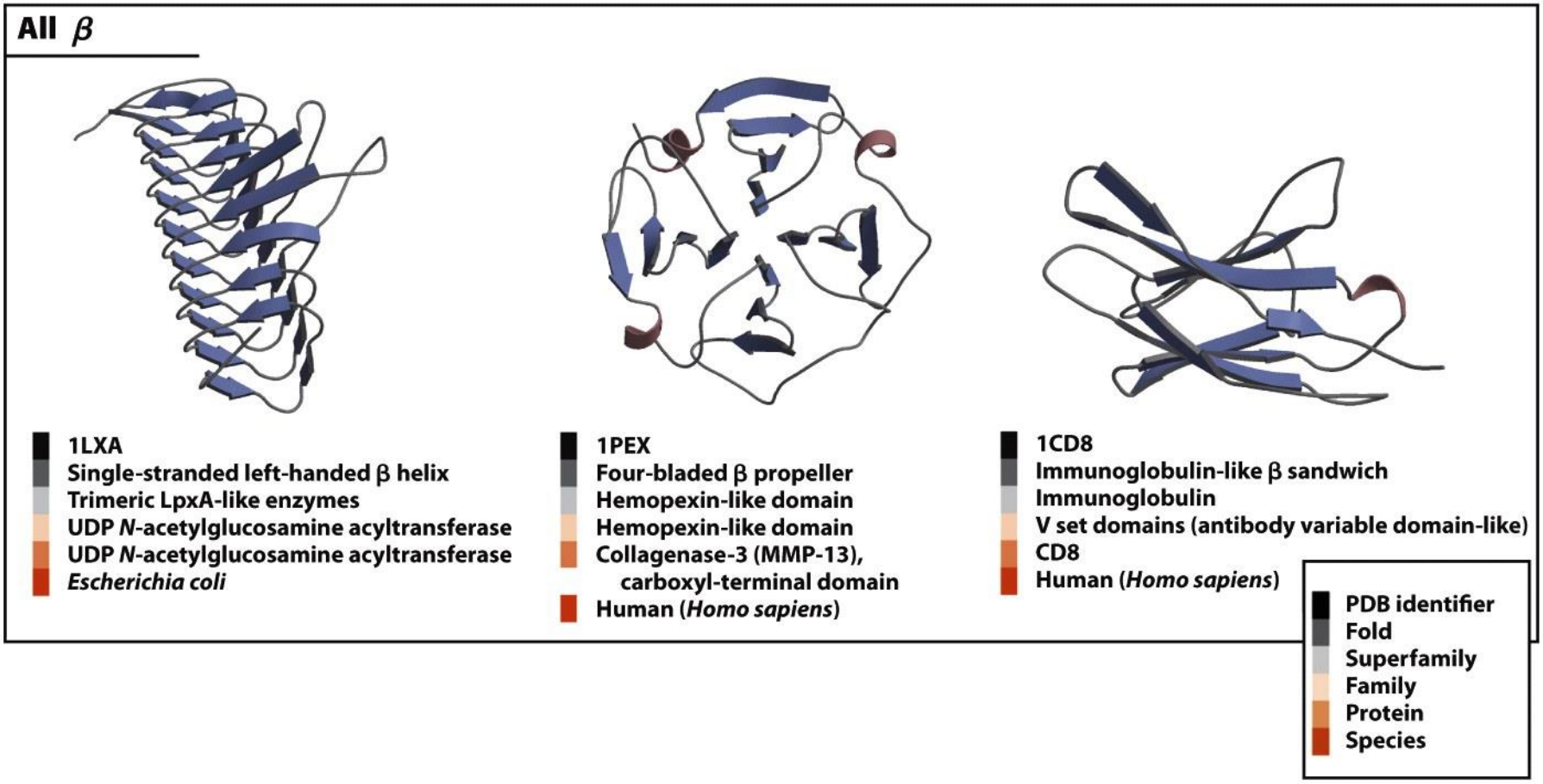


Figure 4-21 part 2
Lehninger Principles of Biochemistry, Fifth Edition

Organization of proteins based on motif

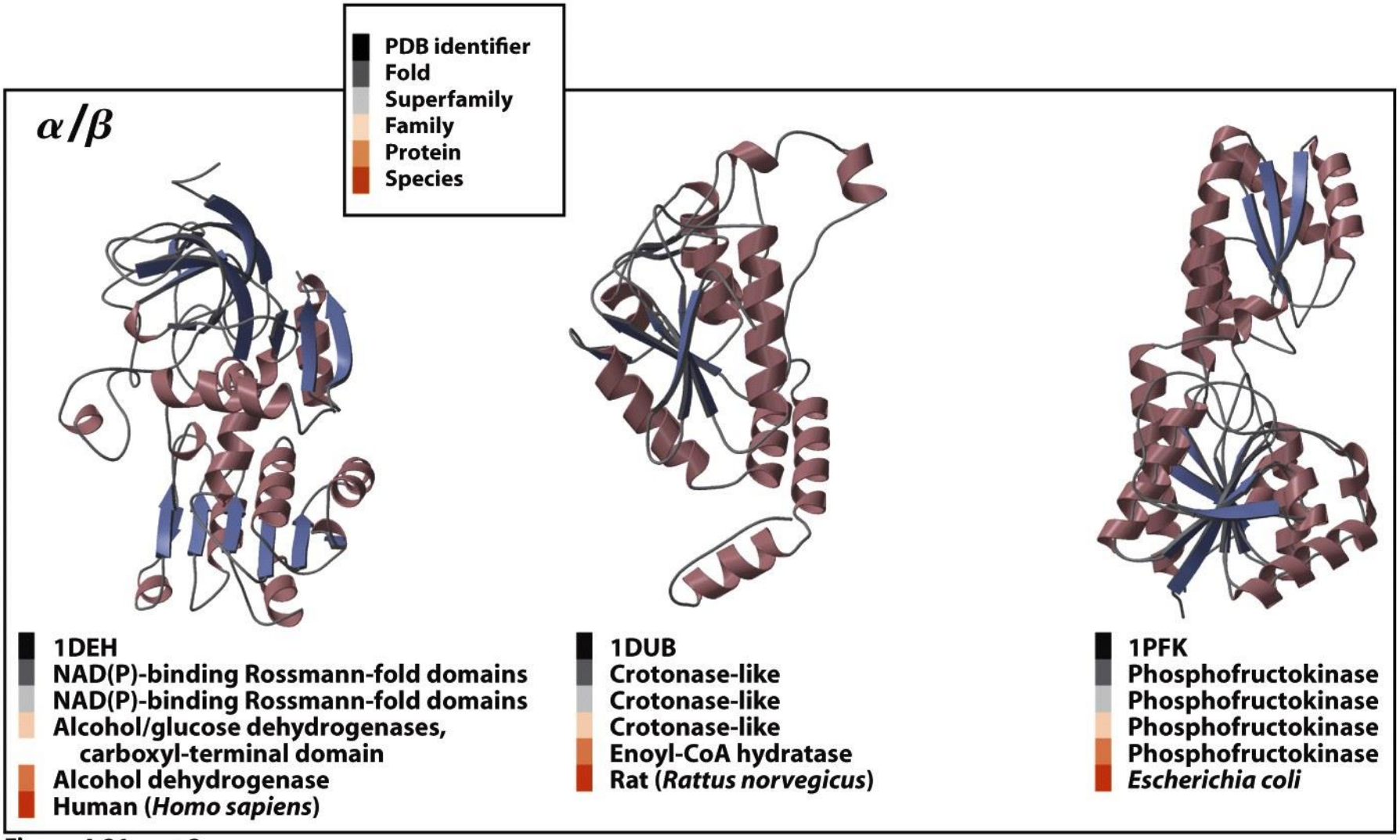


Figure 4-21 part 3

Lehninger Principles of Biochemistry, Fifth Edition

Organization of proteins based on motif

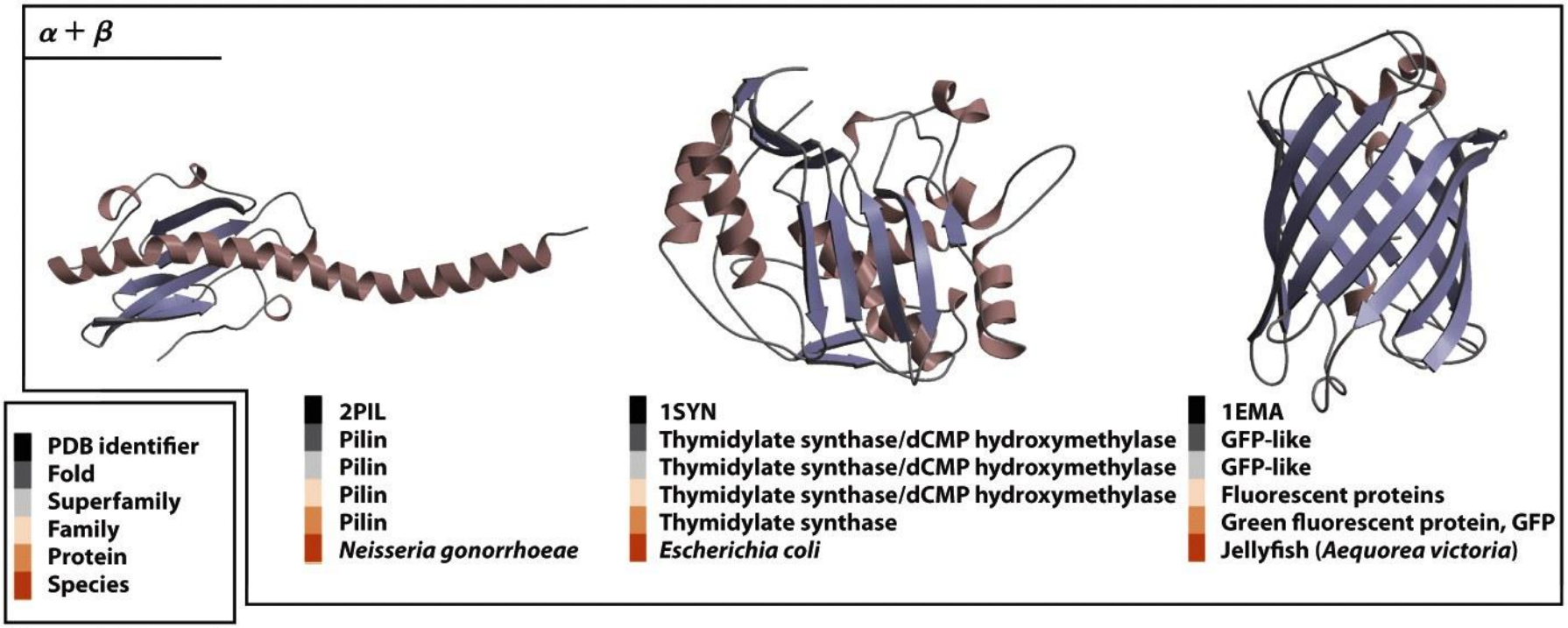


Figure 4-21 part 4

Lehninger Principles of Biochemistry, Fifth Edition

Organization of proteins based on motif

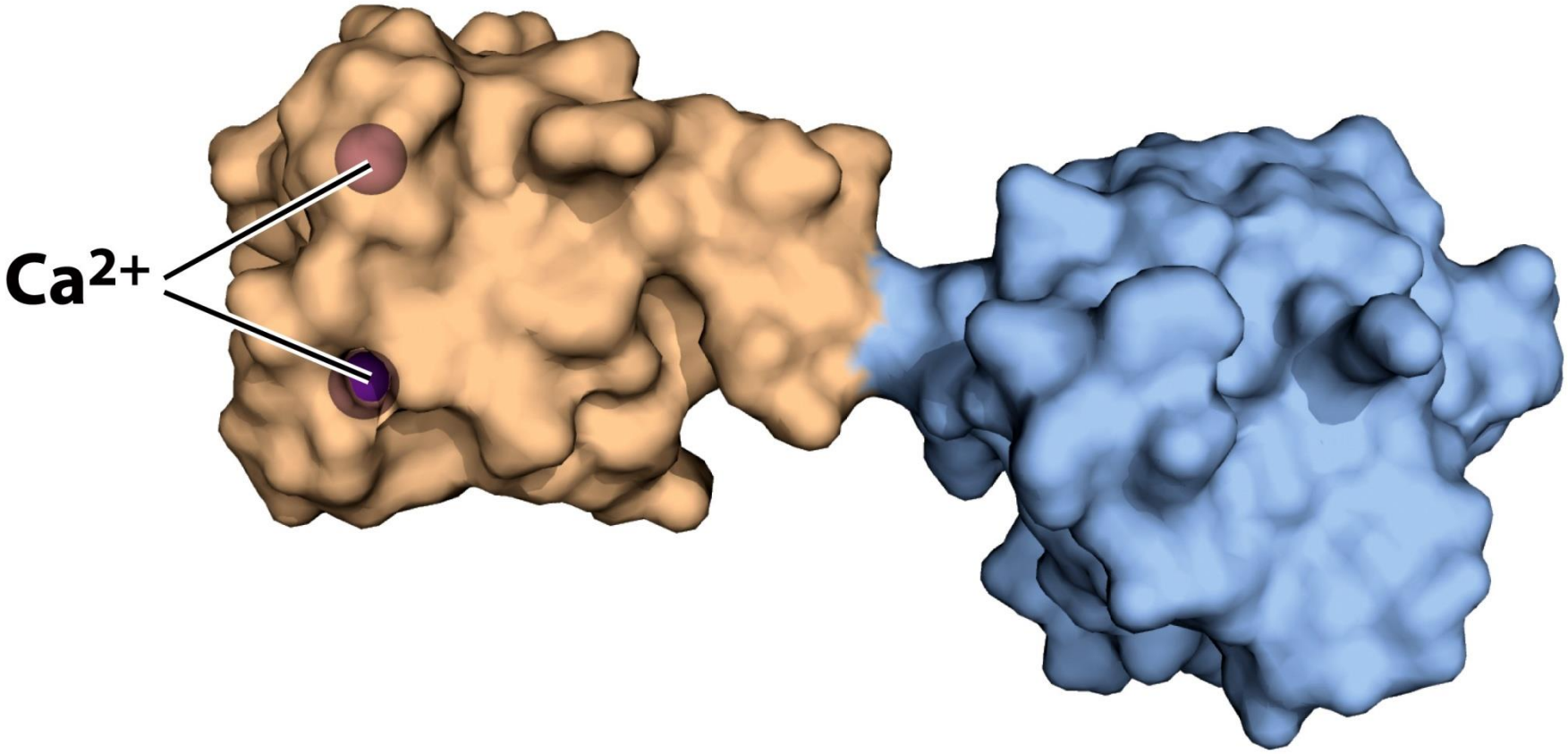
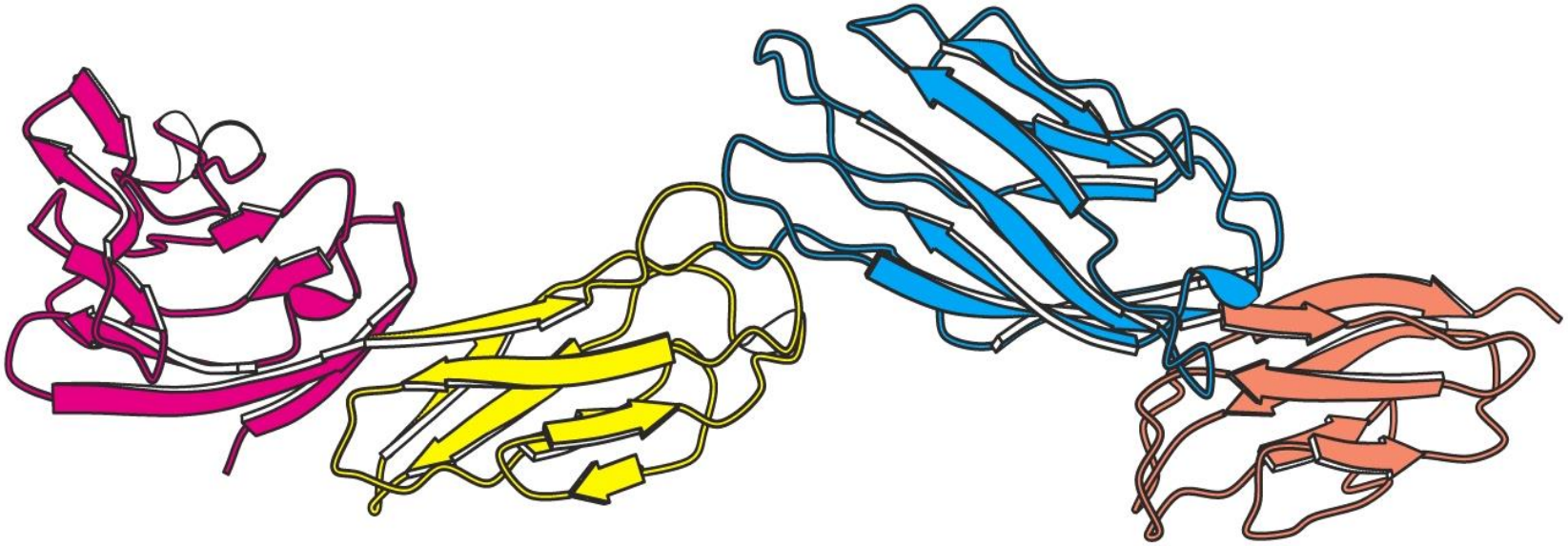


Figure 4-19
Lehninger Principles of Biochemistry, Sixth Edition

Structural domains in the polypeptide troponin C

This calcium-binding protein associated with muscle has two separate calcium-binding domains.



The cell-surface protein CD4 consists of four similar domains.

5. Protein quaternary structure

- **Quaternary structure is formed by spontaneous assembly of individual polypeptides into a larger functional cluster**
- **Quaternary structure refers to the spatial arrangement of subunits and the nature of their interactions**
- **In proteins containing more than one polypeptide chains, each polypeptide chain is called a subunit.**

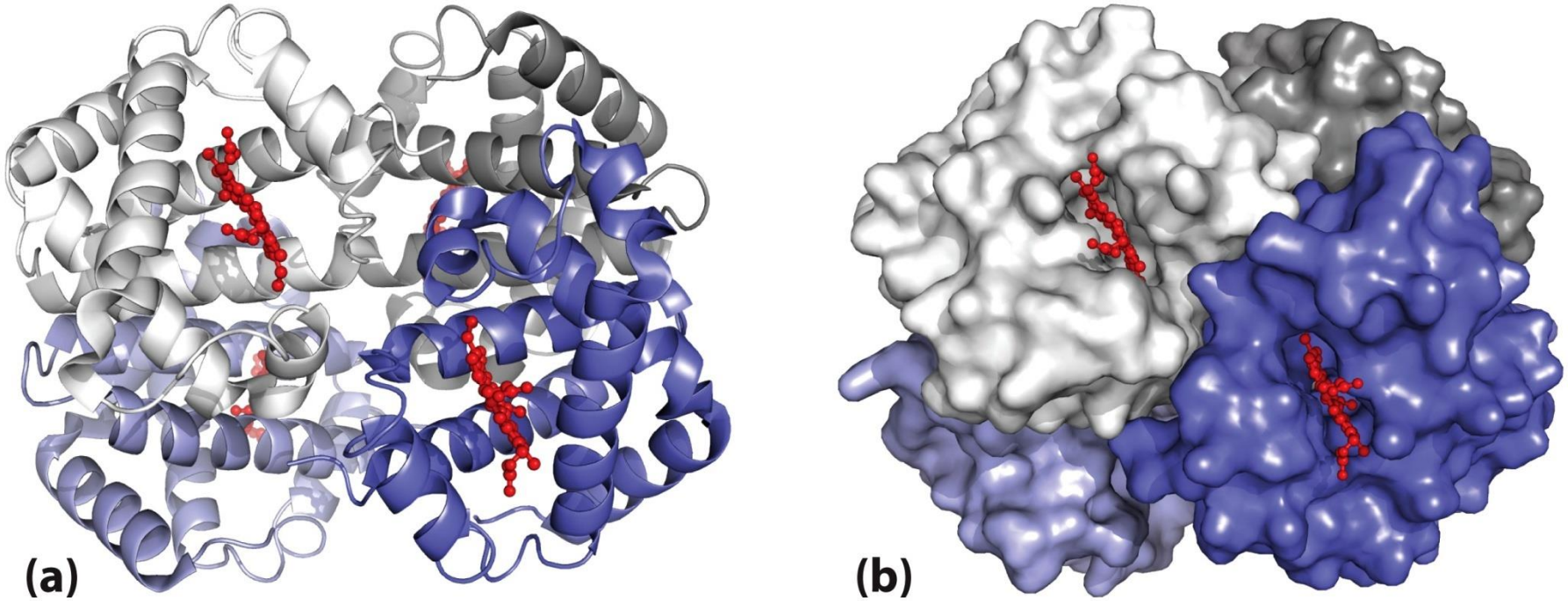
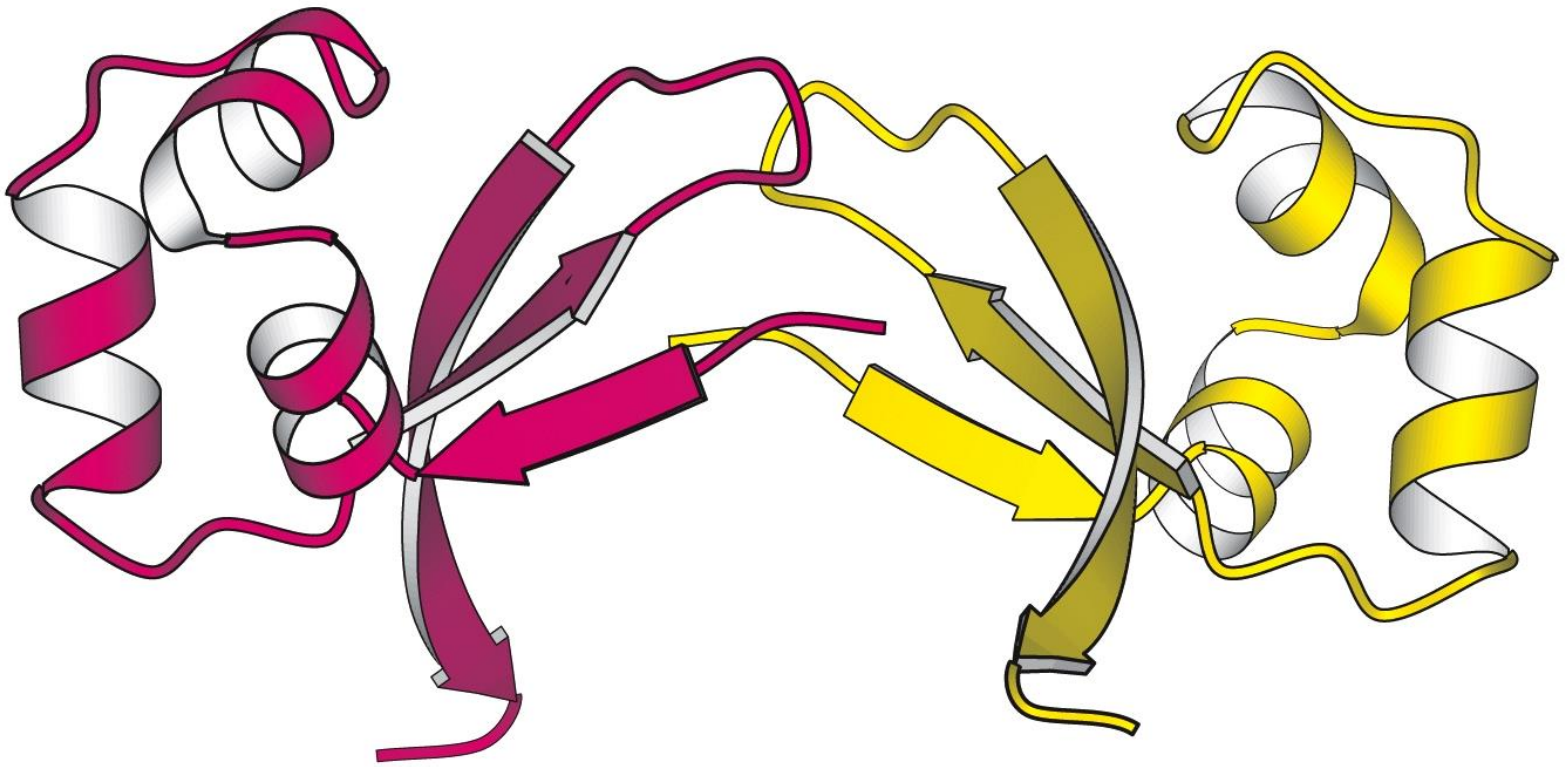


Figure 4-23
Lehninger Principles of Biochemistry, Sixth Edition

Quaternary structure of deoxyhemoglobin



The Cro protein of bacteriophage A is a dimer of identical subunits.

Protein structure

- **Primary structure: the amino acid sequence**
- **Secondary structure: the spatial arrangement of amino acid residues that are nearby in the sequence (α helix, β sheet)**
- **Tertiary structure: the spatial arrangement of amino acid residues that are far apart in the sequence; the pattern of disulfide bonds**
- **Quaternary structure: the spatial arrangement of subunits and the nature of their interactions**

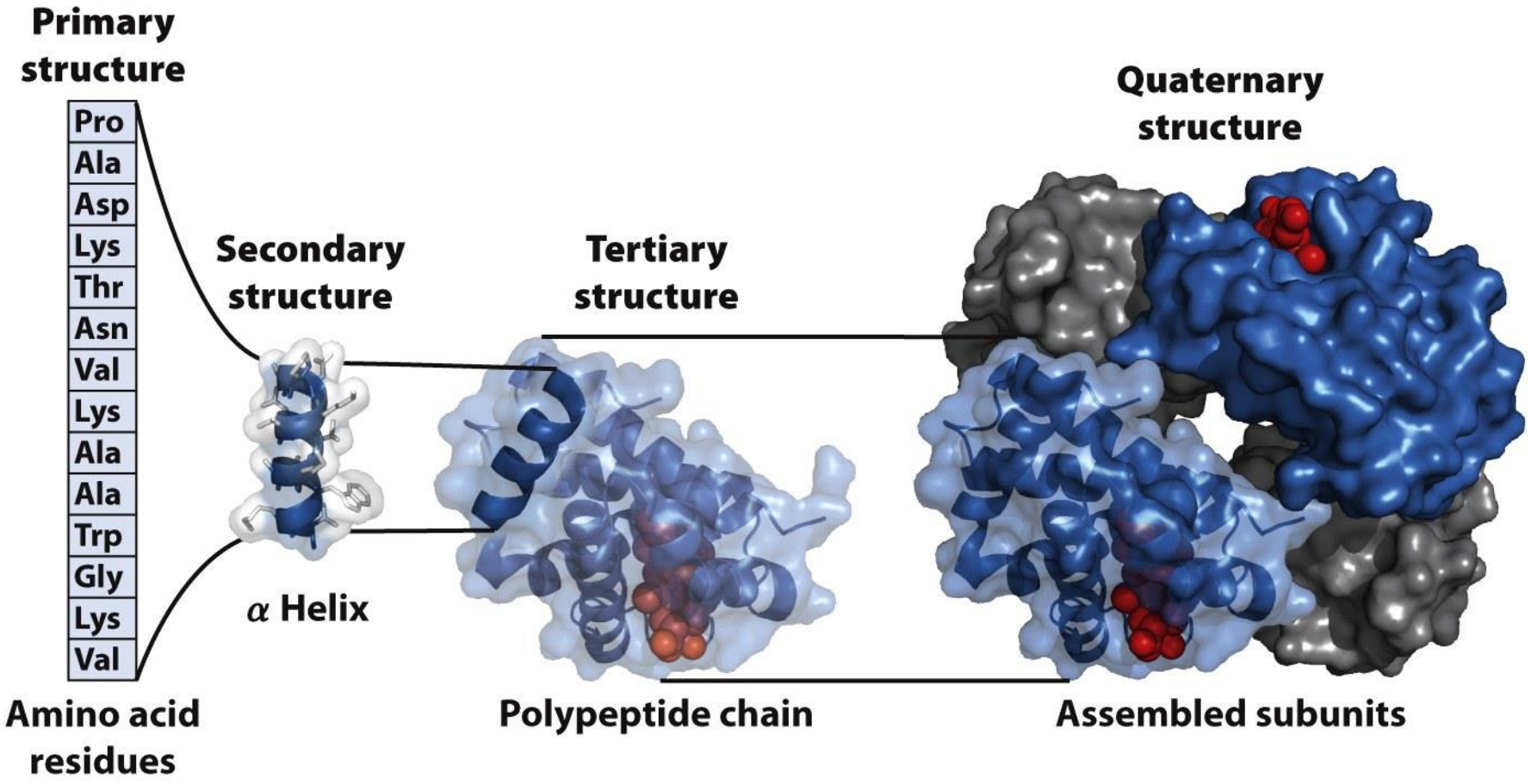


Figure 3-23
Lehninger Principles of Biochemistry, Fifth Edition

Levels of structure in proteins

6. Protein denaturation and folding

- **A protein's function depends on its three-dimensional structure**
- **Loss of structural integrity with accompanying loss of activity is called denaturation**
- **Proteins can be denatured by**
  - **heat**
  - **pH extremes**
  - **organic solvents: alcohol, acetone**
  - **denaturing reagents: urea, guanidine hydrochloride**

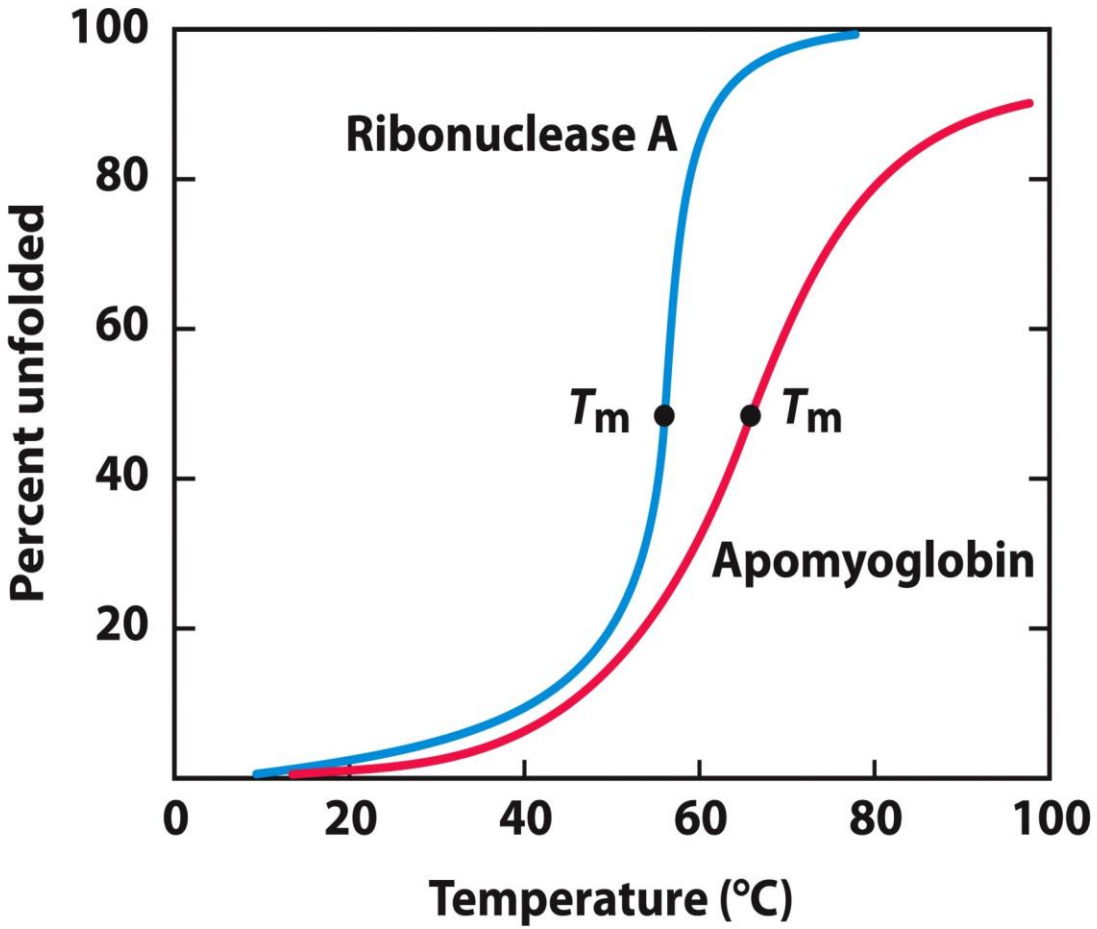


Figure 4-26a
Lehninger Principles of Biochemistry, Sixth Edition

The transition from the folded to the unfolded state is abrupt, suggesting cooperativity in the unfolding process. The midpoint of the temperature range over which denaturation occurs is called the melting temperature, or T_m.

Thermal denaturation

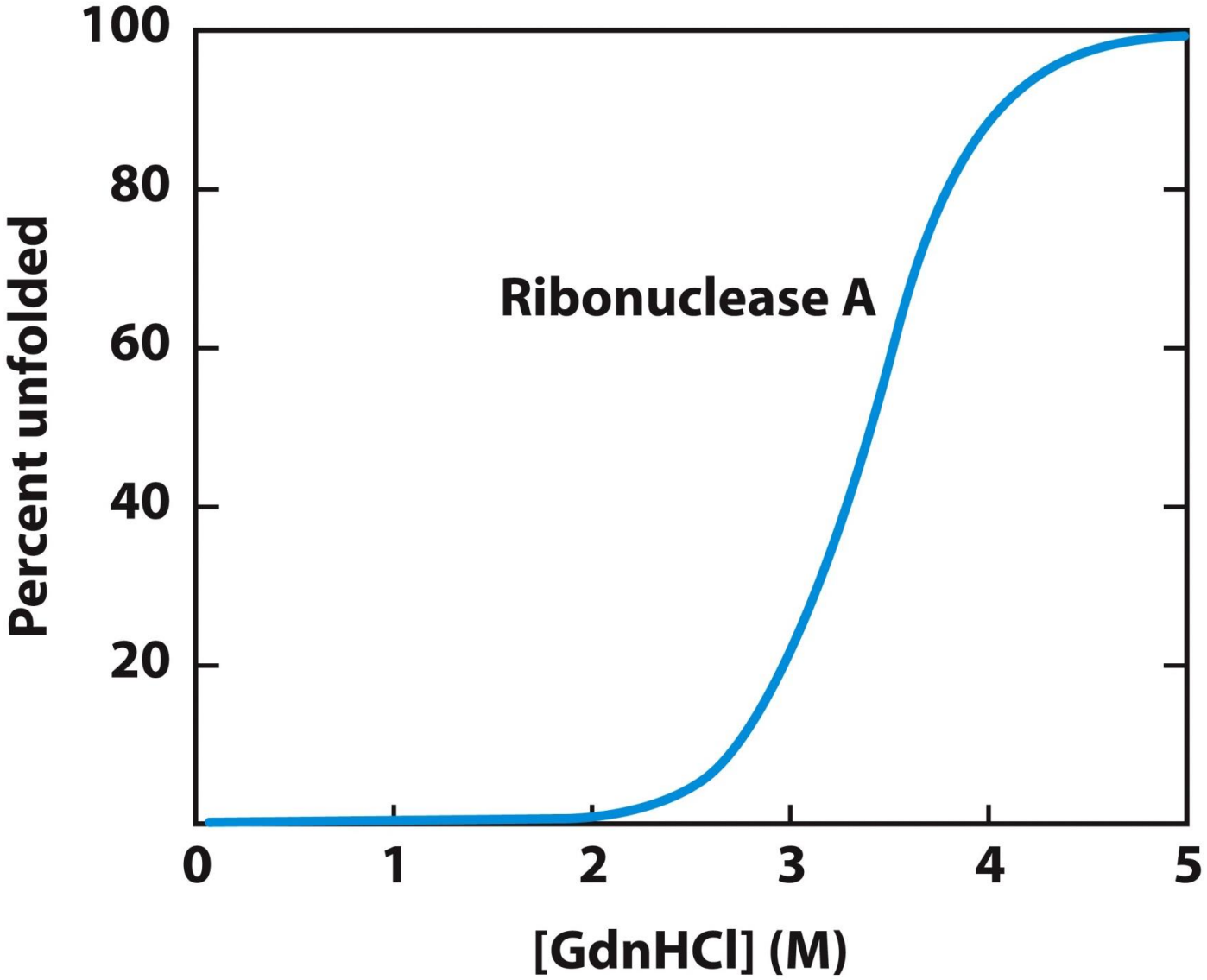
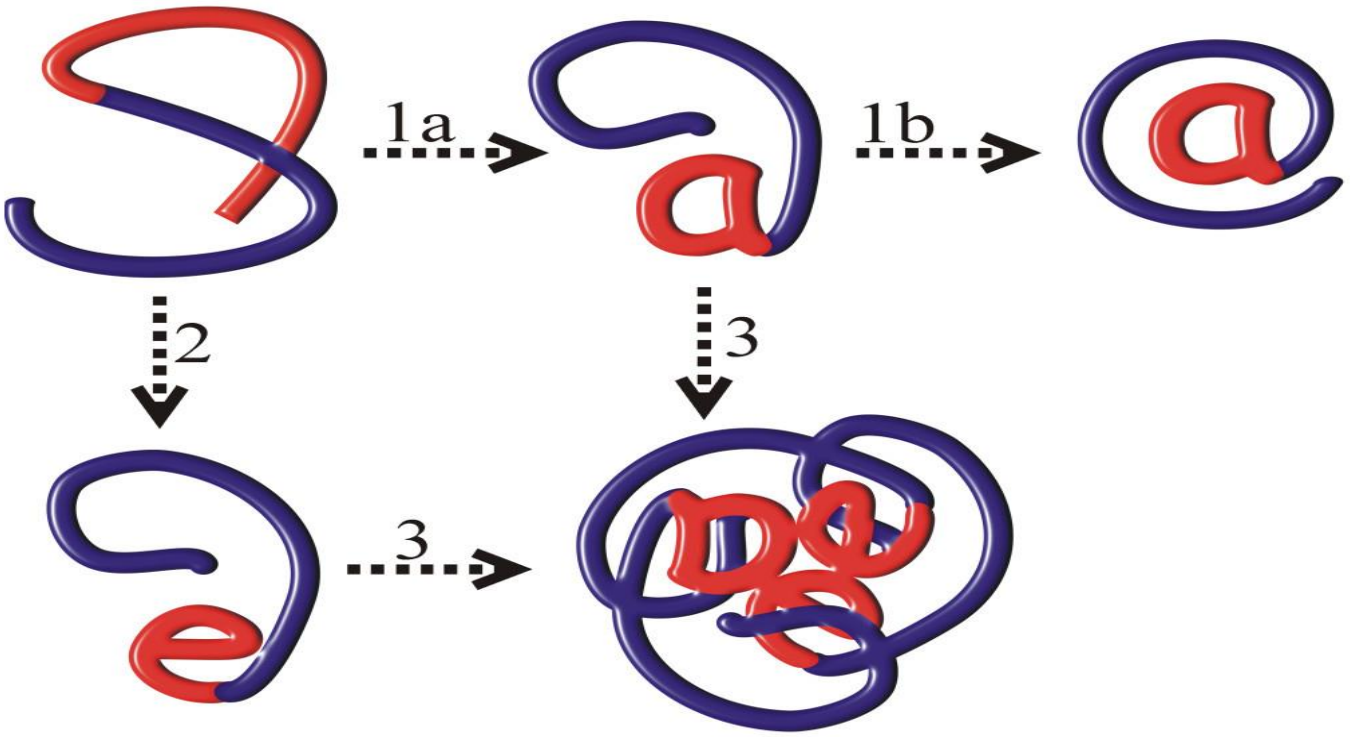


Figure 4-26b
Lehninger Principles of Biochemistry, Sixth Edition

Denaturation by guanidine hydrochloride

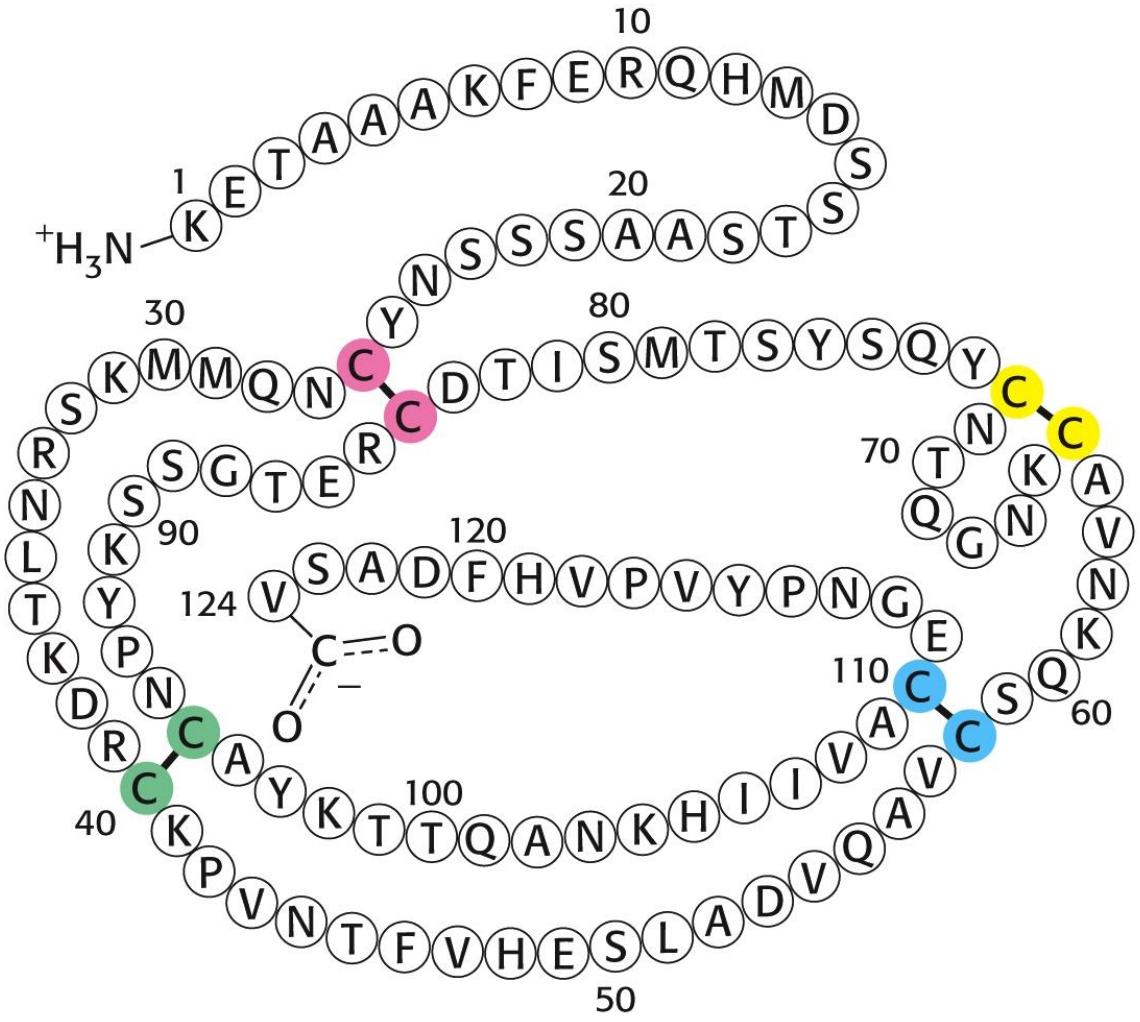
Protein Folding

1D Structure → **3D Structure**



Ribonuclease refolding experiment

- **Ribonuclease is a small protein that contains 8 cysteine linked via four disulfide bonds**
- **Urea in the presence of β-mercaptoethanol fully denatures ribonuclease**
- **When urea and β-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 that Chris Anfinsen won the Nobel Prize in Chemistry in 1972**

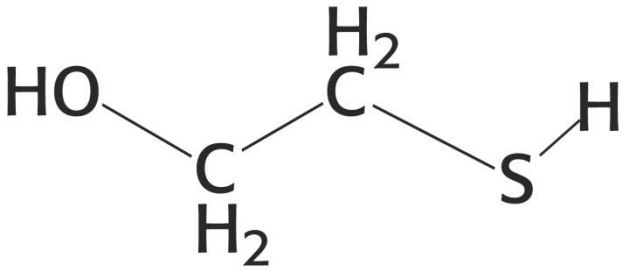


Amino acid sequence of bovine ribonuclease

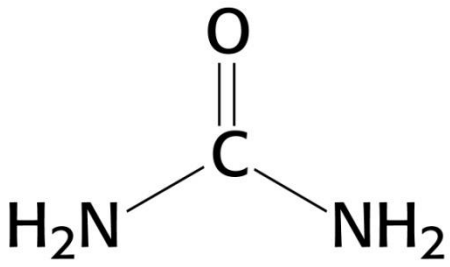
(single chain; 124 amino acids; 4 disulfide bonds)

What conditions are required to restore the structure when the 3D structure of ribonuclease is disrupted?

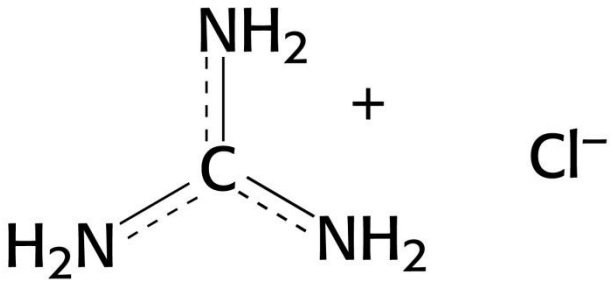
Step 1. Disrupt the 3D structure of ribonuclease by β-mercaptoethanol & urea



β-Mercaptoethanol



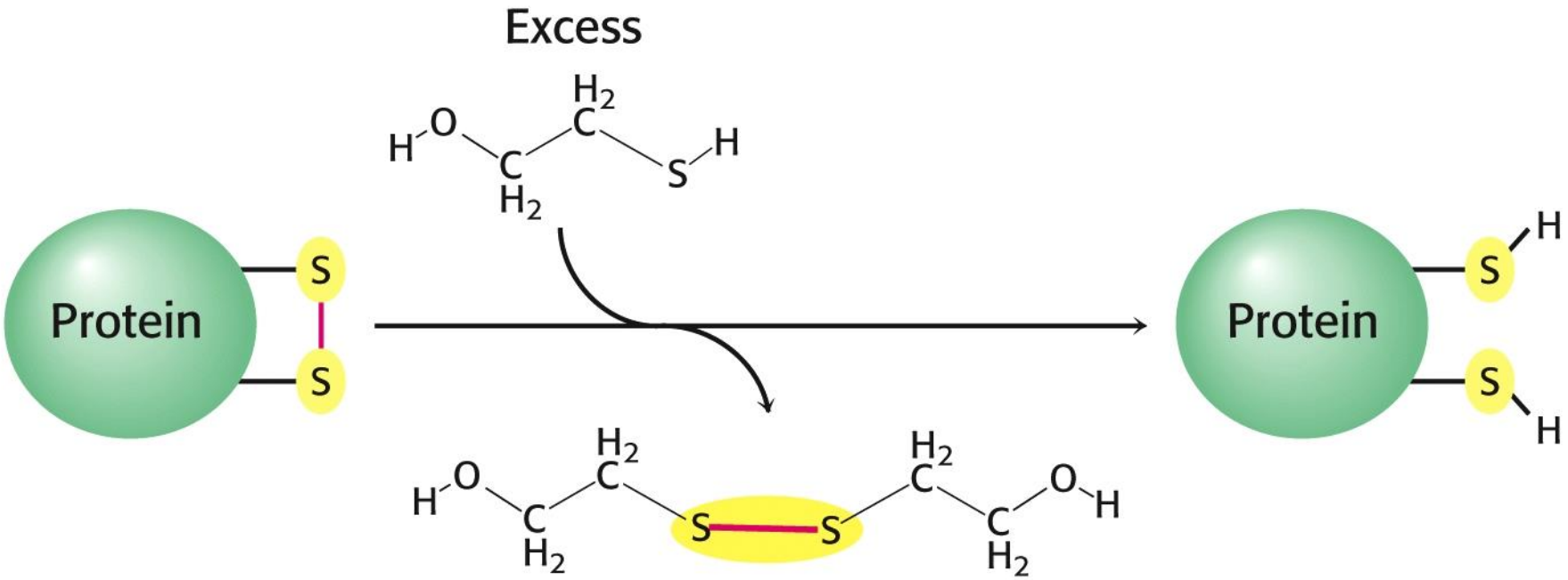
Urea



Guanidinium chloride

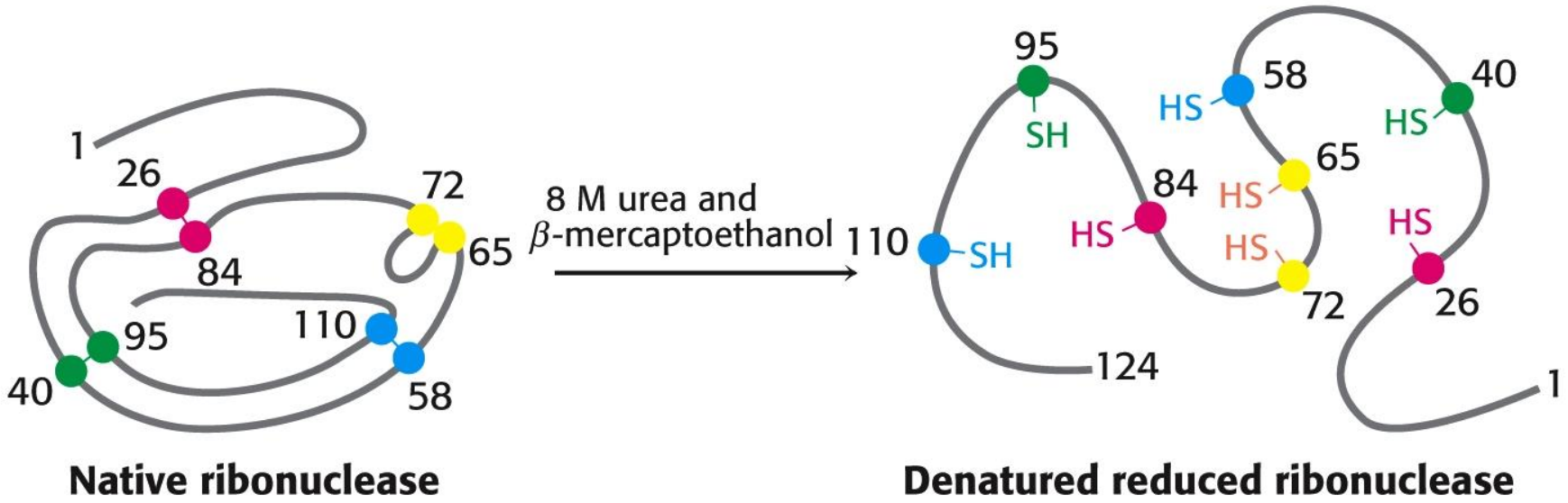
Urea or guanidinium chloride: disrupt the noncovalent bonds

β-mercaptoethanol (BME): cleave the disulfide bonds reversibly



Role of β-mercaptoethanol in reducing disulfide bonds

As the disulfides are reduced, the β-mercaptoethanol is oxidized and forms dimers.



Reduction and denaturation of ribonuclease

How would you confirm that the 3D structure of ribonuclease is disrupted after β-mercaptoethanol & urea treatment?

- **Random-coil conformation, evidenced by physical properties such as viscosity and optical activity**
- **Loss of enzymatic activity**

Step 2. Restore the 3D structure of ribonuclease by dialysis (a method that slowly removes urea and β-mercaptoethanol)

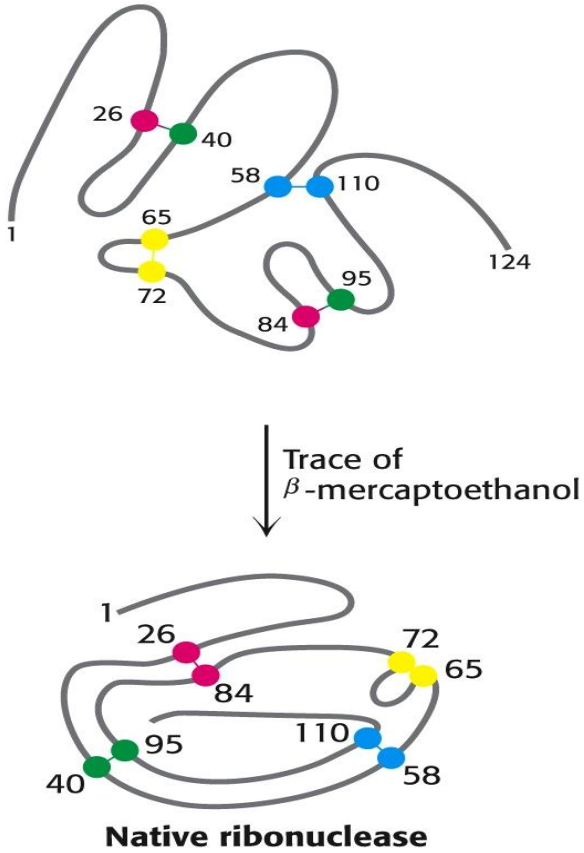
Findings: The denatured ribonuclease slowly regained its enzymatic activity; measured physical and chemical properties of the enzyme were virtually identical with those of the native enzyme. The enzyme spontaneously refolded into a catalytically active form.

Conclusion: The information needed to specify the catalytically active structure of ribonuclease is contained in its amino acid sequence.

Findings: **When reduced ribonuclease was reoxidized while it was still in 8 M urea and the preparation was then dialyzed to remove the urea, ribonuclease had only 1% of the enzymatic activity of the native protein.**

Reason: wrong disulfide bonds were formed in urea.

Conclusion: Weak bonding interactions are required for correct positioning of disulfide bonds and restoration of the native conformation.



Reestablishing correct disulfide pairing

Native ribonuclease can be reformed from scrambled ribonuclease in the presence of a trace of β-mercaptoethanol.

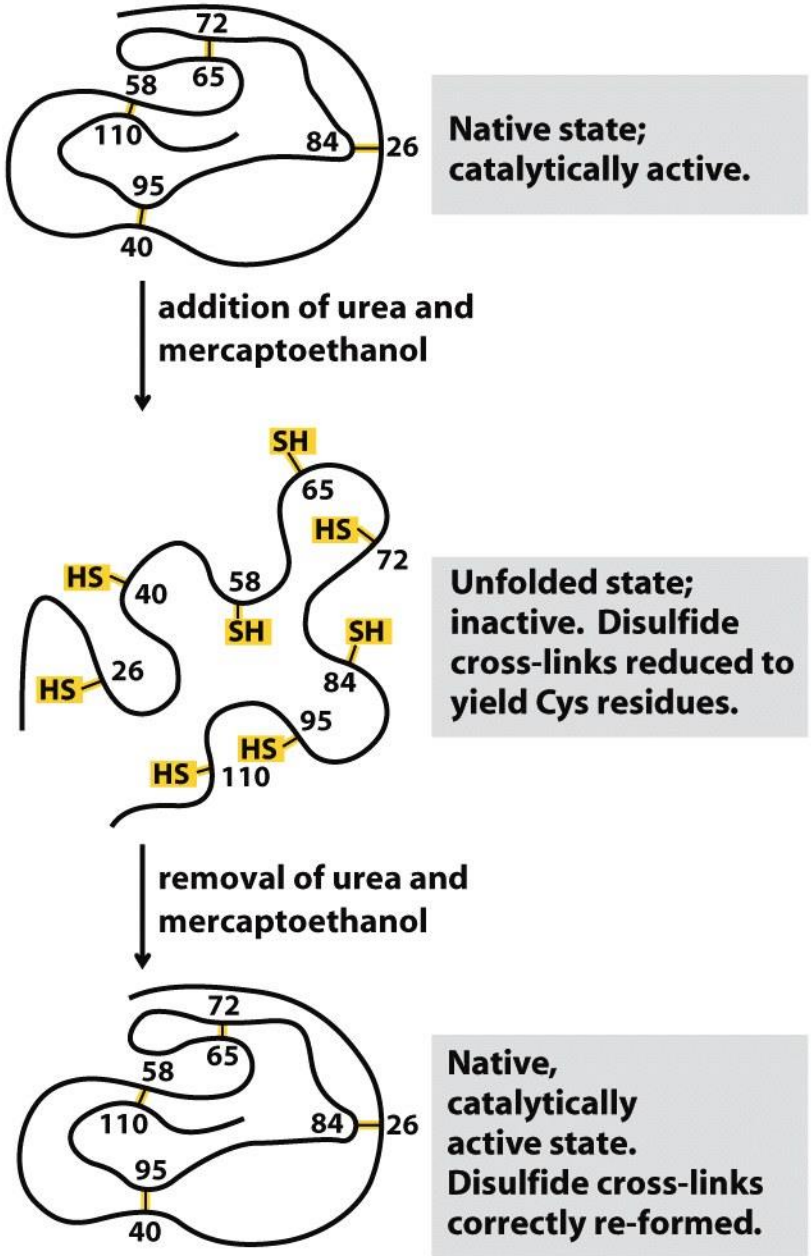


Figure 4-26
Lehninger Principles of Biochemistry, Fifth Edition

Renaturation of unfolded, denatured ribonuclease

The Nobel Prize in Chemistry 1972

Christian B. Anfinsen

USA

National Institutes of Health
Bethesda, MD, USA

"for his work on ribonuclease, especially concerning the connection between the amino acid sequence and the biologically active conformation"

http://nobelprize.org/nobel_prizes/chemistry/laureates/1972/

Proceedings of the

NATIONAL ACADEMY OF SCIENCES

Volume 47 · Number 9 · September 15, 1961

THE KINETICS OF FORMATION OF NATIVE RIBONUCLEASE DURING OXIDATION OF THE REDUCED POLYPEPTIDE CHAIN

By C. B. Anfinsen, E. Haber,* M. Sela,† and F. H. White, Jr.

LABORATORY OF CELLULAR PHYSIOLOGY AND METABOLISM, NATIONAL HEART INSTITUTE, NATIONAL INSTITUTES OF HEALTH

Communicated by John T. Edsall, July 31, 1961

Abstract

Bovine pancreatic ribonuclease is completely reduced by treatment with mercaptoethanol in 8 *M* urea to yield a randomly coiled polypeptide chain containing eight cysteine residues.[1–3] Under optimal conditions of polypeptide concentration and pH, essentially complete reformation of the disulfide bonds of the native enzyme occurs in the presence of molecular oxygen.[2, 3] From chemical and physical studies of the reformed enzyme, it may be concluded that the information for the correct pairing of half-cystine residues in disulfide linkage, and for the assumption of the native secondary and tertiary structures, is contained in the amino acid sequence itself.

Preliminary to studies on the interactions involved in the refolding process, and to establish the order of chemical events during the formation of active protein, we have followed the rates of disappearance of sulfhydryl groups, and of the appearance of the spectral properties characteristic of the native enzyme and its active derivatives. The appearance of increased positive optical rotation associated with secondary structure was also studied. The results rule out the sequential formation of one active molecule after another. They suggest as a major possibility that some disulfide bonds formed during the early stages of oxidation are not identical with those of the native protein but undergo rearrangement to yield the native configuration.

PNAS 47, 1309-1314 (1961)

20 July 1973, Volume 181, Number 4096 SCIENCE

Principles that Govern the Folding of Protein Chains

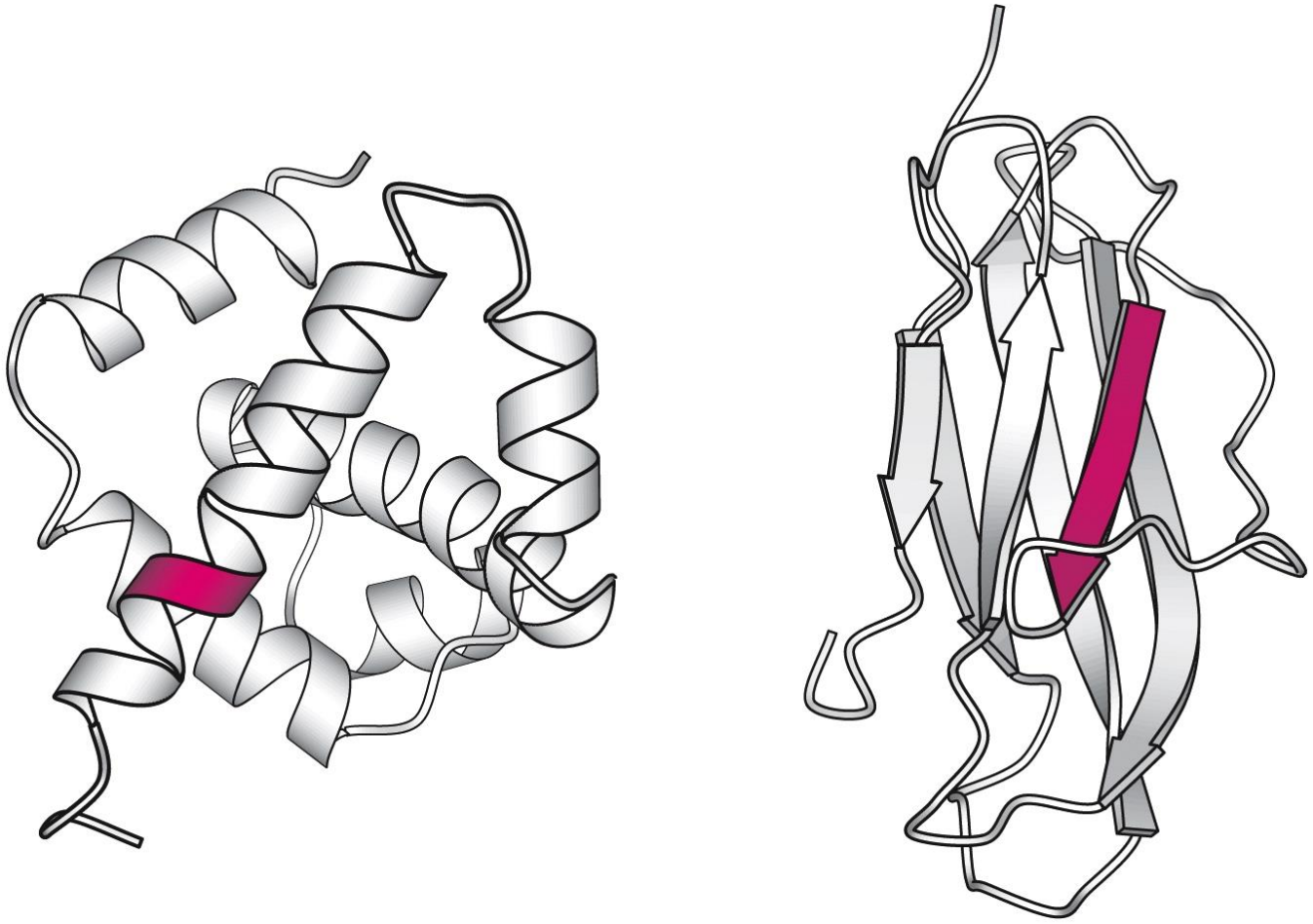
Christian B. Anfinsen

exists under conditions similar to those for which it was selected—the so-called physiological state.

After several years of study on the ribonuclease molecule it became clear to us, and to many others in the field of protein conformation, that proteins devoid of restrictive disulfide bonds or other covalent cross-linkages would make more convenient models for the study of the thermodynamic and kinetic aspects of the nucleation, and subsequent pathways, of polypeptide chain folding. Much of what I will review deals with studies on the flexible and

Prediction of secondary structure remains very difficult

- **Some amino acid sequences do not uniquely determine secondary structure**
- **Many sequences can adopt alternative conformations in different proteins.**
- **Tertiary interactions may be decisive in specifying the secondary structure of some segments.**



Alternative conformations of a peptide sequence

The sequence VDLLKN shown in red is in a α-helix in one protein (left) and a β strand in another protein (right).

How can proteins fold so fast?

- **Proteins fold to the lowest-energy fold in the microsecond to second time scales. How can they find the right fold so fast?**
- **It is mathematically impossible for protein folding to occur by randomly trying every conformation until the lowest energy one is found (Levinthal's paradox)**
- **Search for the minimum is not random because the direction toward the native structure is thermodynamically most favorable**

Levinthal's Paradox
(1968)

Apomyoglobin:

153 amino acid residues

↓

$3^{153} = 10^{73}$ conformations

↓

$10^{73} \times 10^{-13}$ s $= 10^{60}$ s $= 3 \times 10^{52}$ years

Origin of the universe: $(15 \pm 4) \times 10^{9}$ years

Measured rates:

$$U \xrightarrow{< 5\ \text{ms}} I \xrightarrow{\sim 2\ \text{s}} N$$

Possible mechanisms of protein folding

- **Random search for the energetically most favorable conformation: for a protein of 100 residues, this would take 1.6×10^{27} years (Assume each residue can assume 3 conformations; it takes 10^{-13} seconds to convert from one conformation to another)**
- **Partly correct intermediates are retained: 'Cumulative selection'**
- **Protein is folded by progressive stabilization of intermediates rather than by random search**

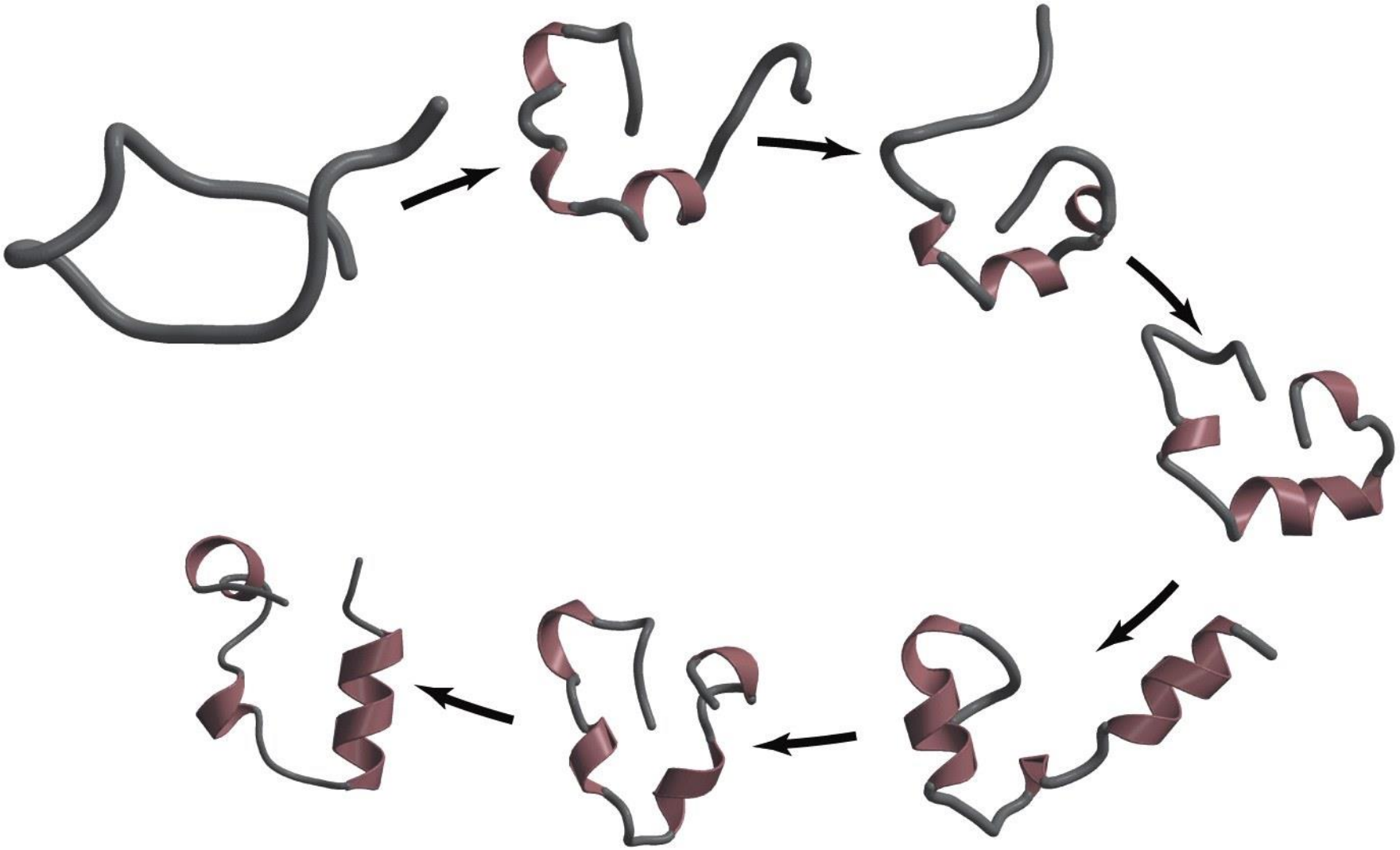


Figure 4-27
Lehninger Principles of Biochemistry, Fifth Edition

A simulated folding pathway

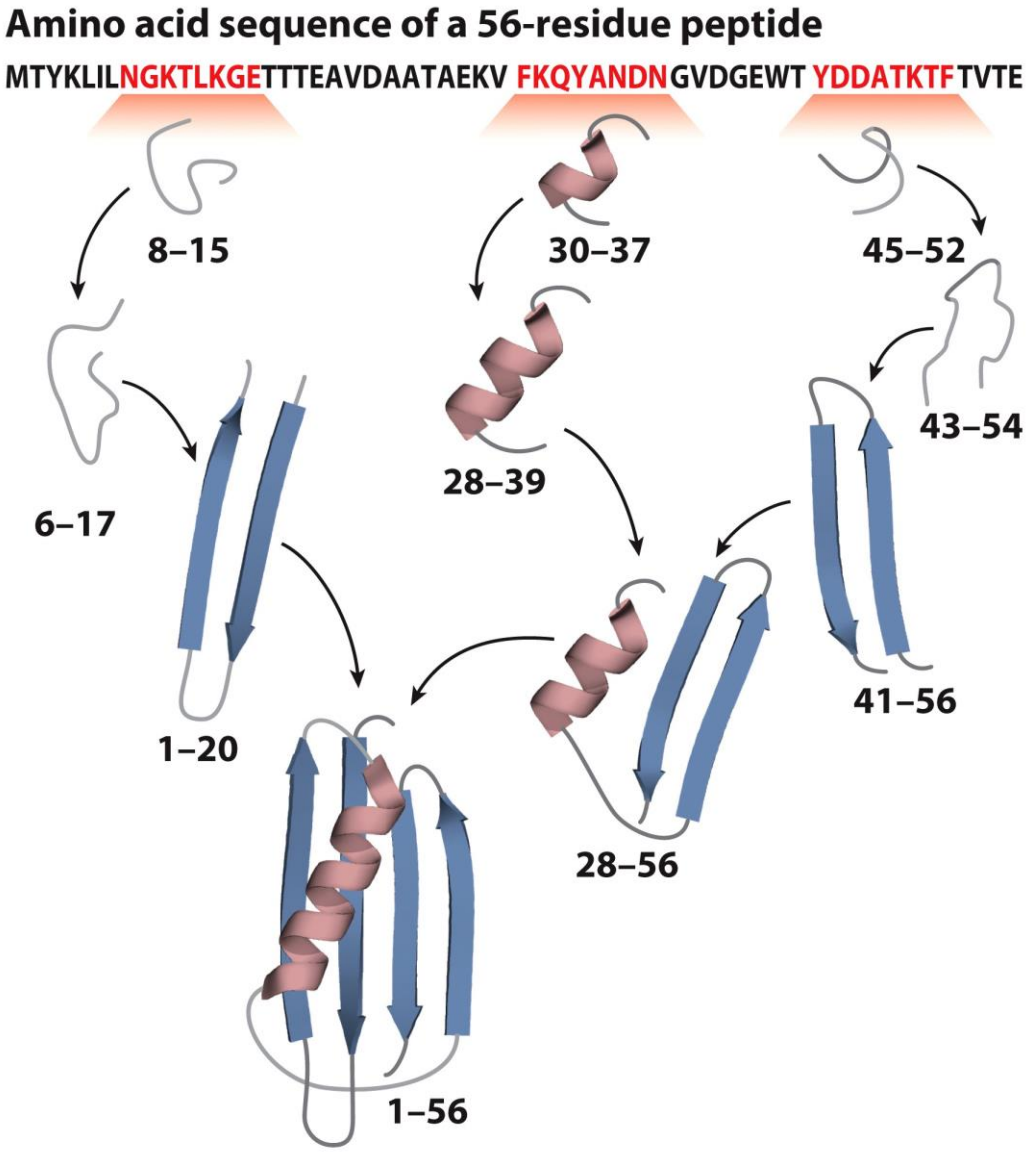


Figure 4-28
Lehninger Principles of Biochemistry, Sixth Edition

A protein-folding pathway as defined for a small protein

Protein folding is a highly cooperative process

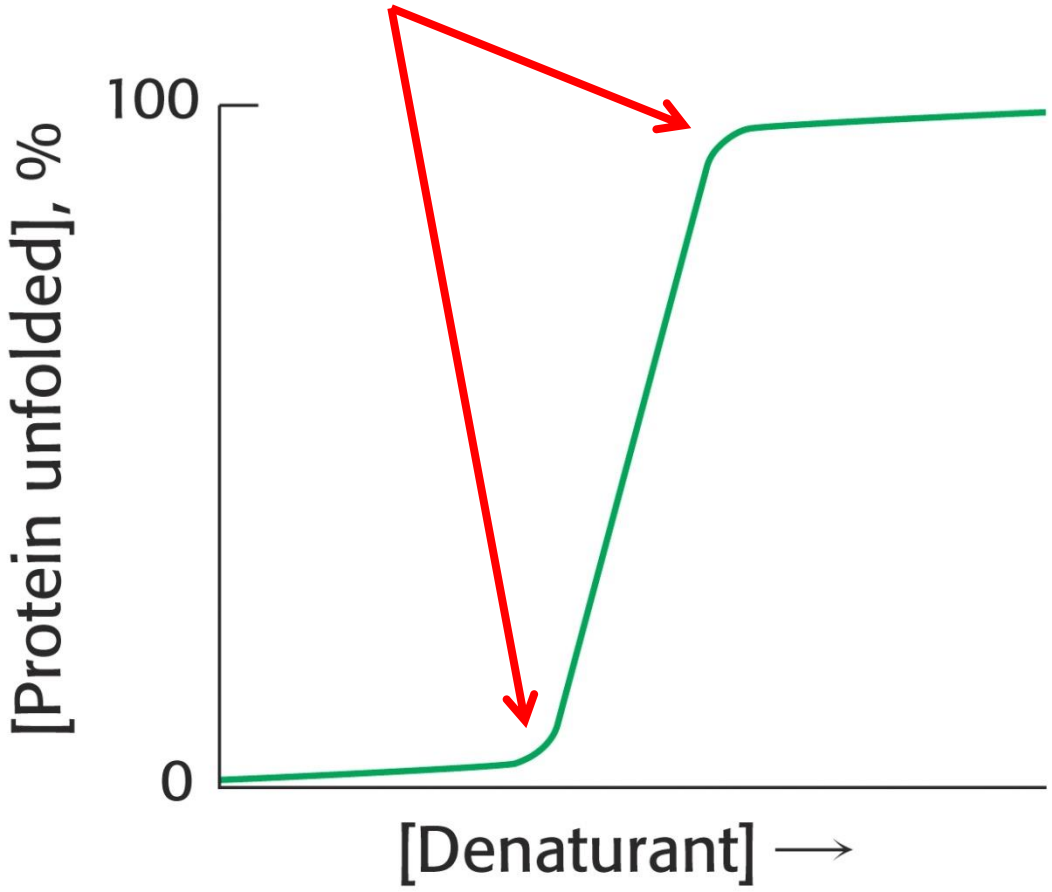
- **As the concentration of denaturant increases, the sharp transition from the folded form to the unfolded form suggests that only these two conformational states are present to any significant extent. (Two States: folded vs unfolded)**

- **Protein folding and unfolding is thus largely an 'all or none' process that results from a cooperative transition.**

Protein folding process

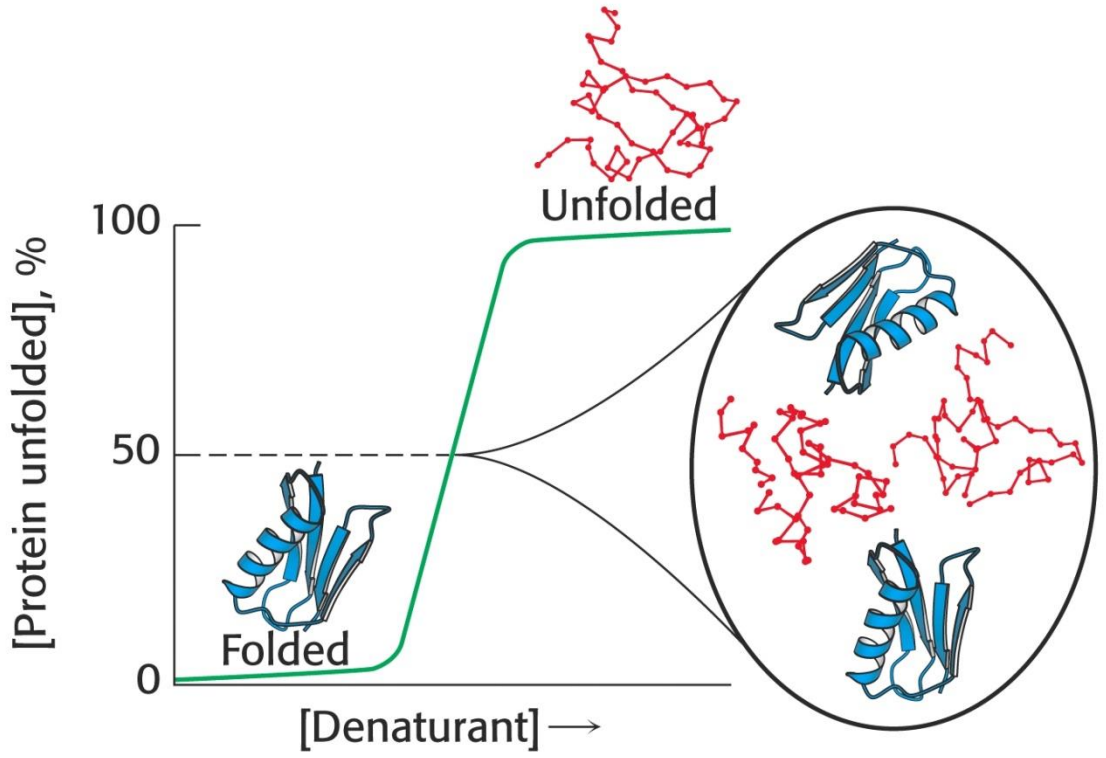
$$U \rightarrow I_1 \rightarrow I_2 \rightarrow \cdots\cdots \rightarrow N$$

Note the sharp transition



Transition from folded to unfolded state

Most proteins show a sharp transition from the folded to unfolded form on treatment with increasing concentrations of denaturants.

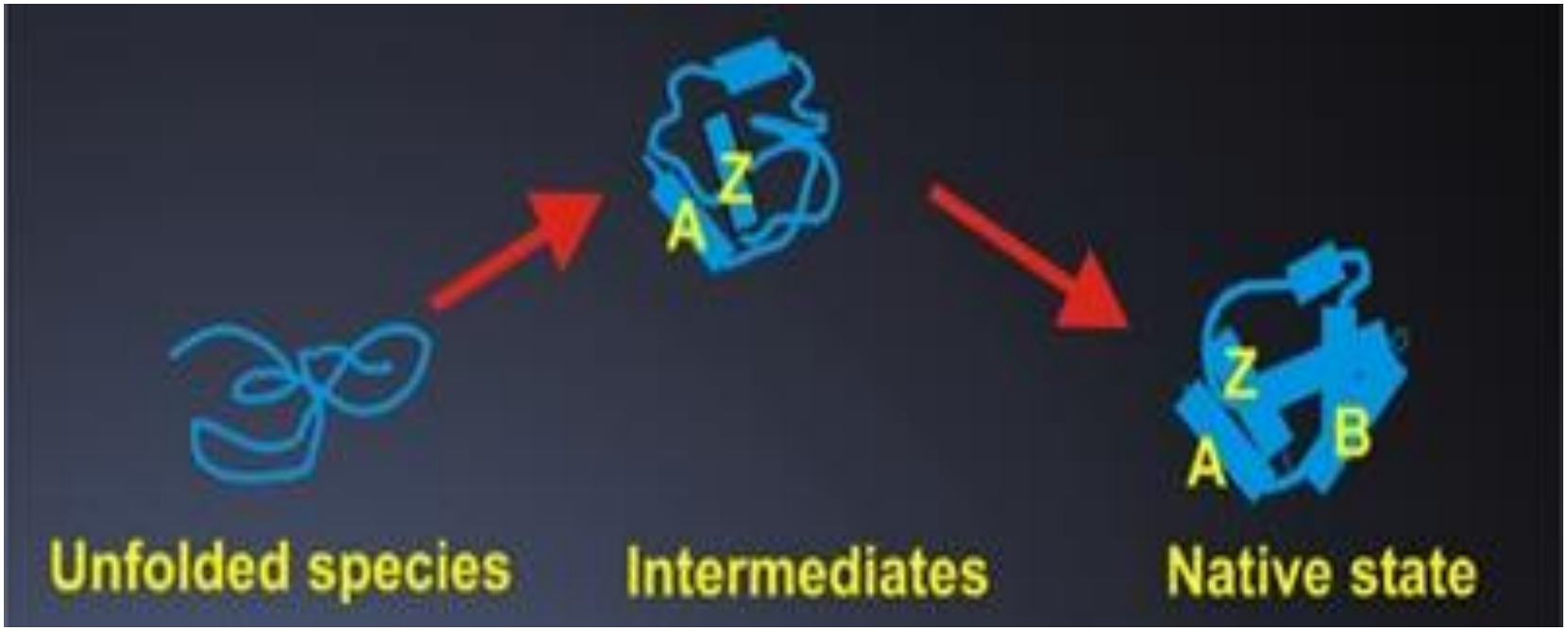


50/50 of fully folded and fully unfolded molecules

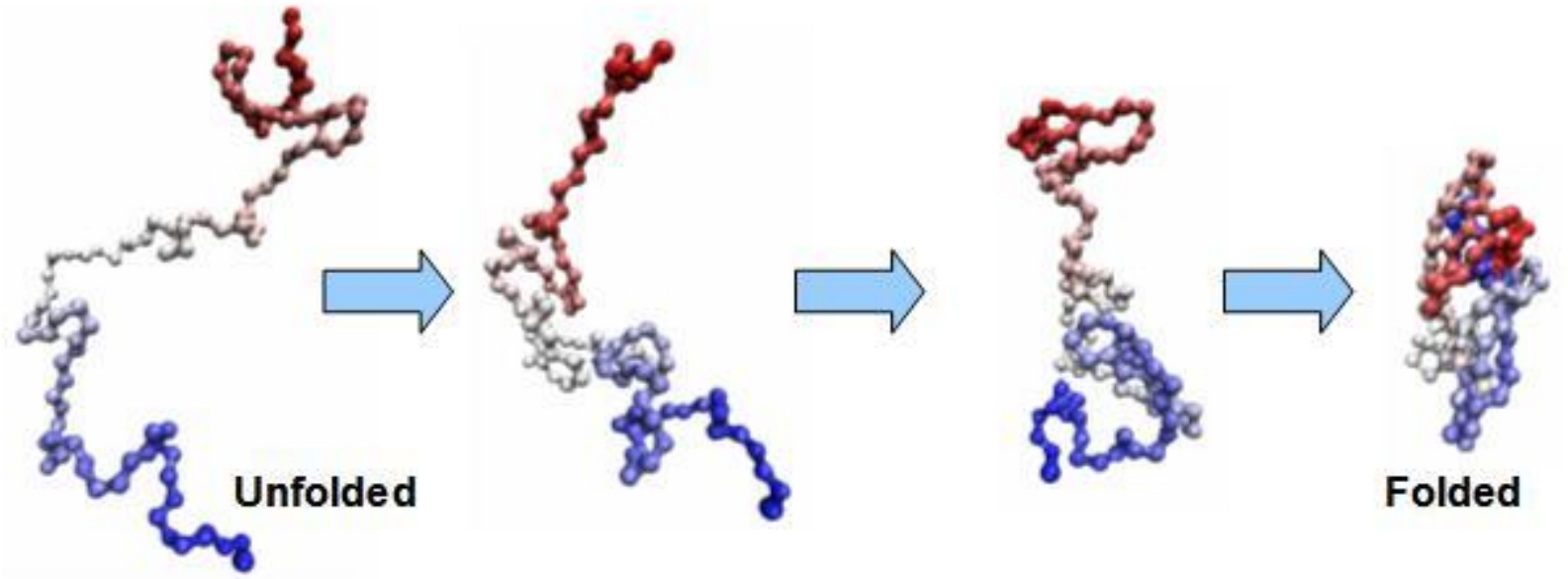
Components of a partially denatured protein solution

In a half-unfolded protein solution, half the molecules are fully folded and half are fully unfolded.

The essence of protein folding is the retention of partly correct intermediates



Protein folding is a highly cooperative process



Regions of secondary structure may form followed by folding into motifs and domains. Large ensembles of folding intermediates are rapidly brought to a single native conformation.

Molecular chaperones

- **Molecular chaperones: proteins that interact with partially folded or improperly folded polypeptides, facilitating correct folding pathways or providing microenvironments in which folding can occur**
- **Hsp70: binds to regions of unfolded polypeptides that are rich in hydrophobic residues, preventing inappropriate aggregation**
- **Hsp70 works with Hsp40, using energy from ATP hydrolysis**
- **DnaK and DnaJ: *E. Coli* homolog of Hsp70 and Hsp40**

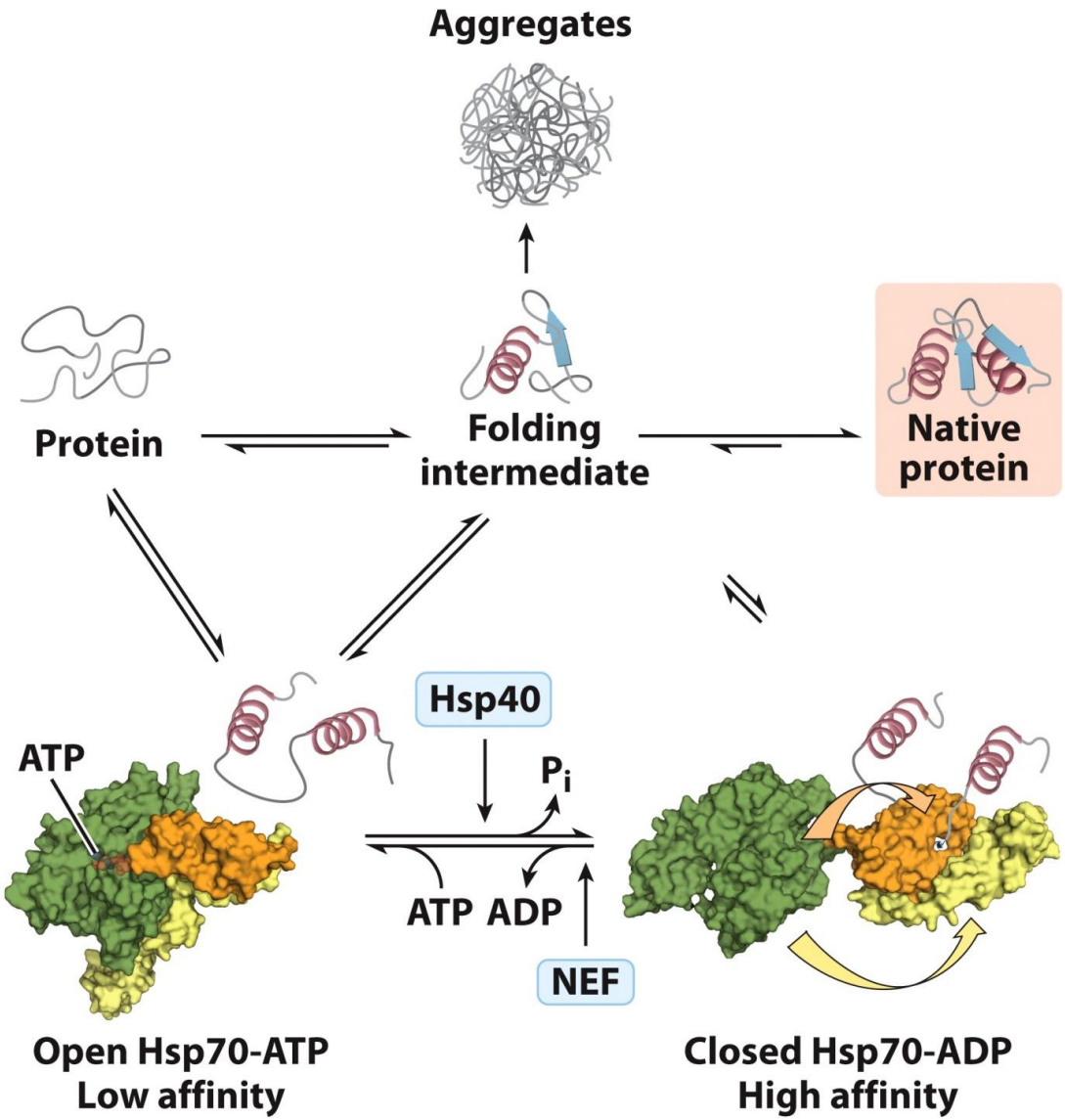


Figure 4-30
Lehninger Principles of Biochemistry, Sixth Edition

Chaperones in protein folding

Molecular chaperones

- **Chaperonin: protein complexes required for the folding of some cellular proteins that do not fold spontaneously**
- **GroEL/GroES: for folding under normal conditions in *E. Coli*, using energy from ATP hydrolysis**
- **Two enzymes required in the folding pathway:**

  PDI: protein disulfide isomerase

  PPI: peptide prolyl cis-trans isomerase

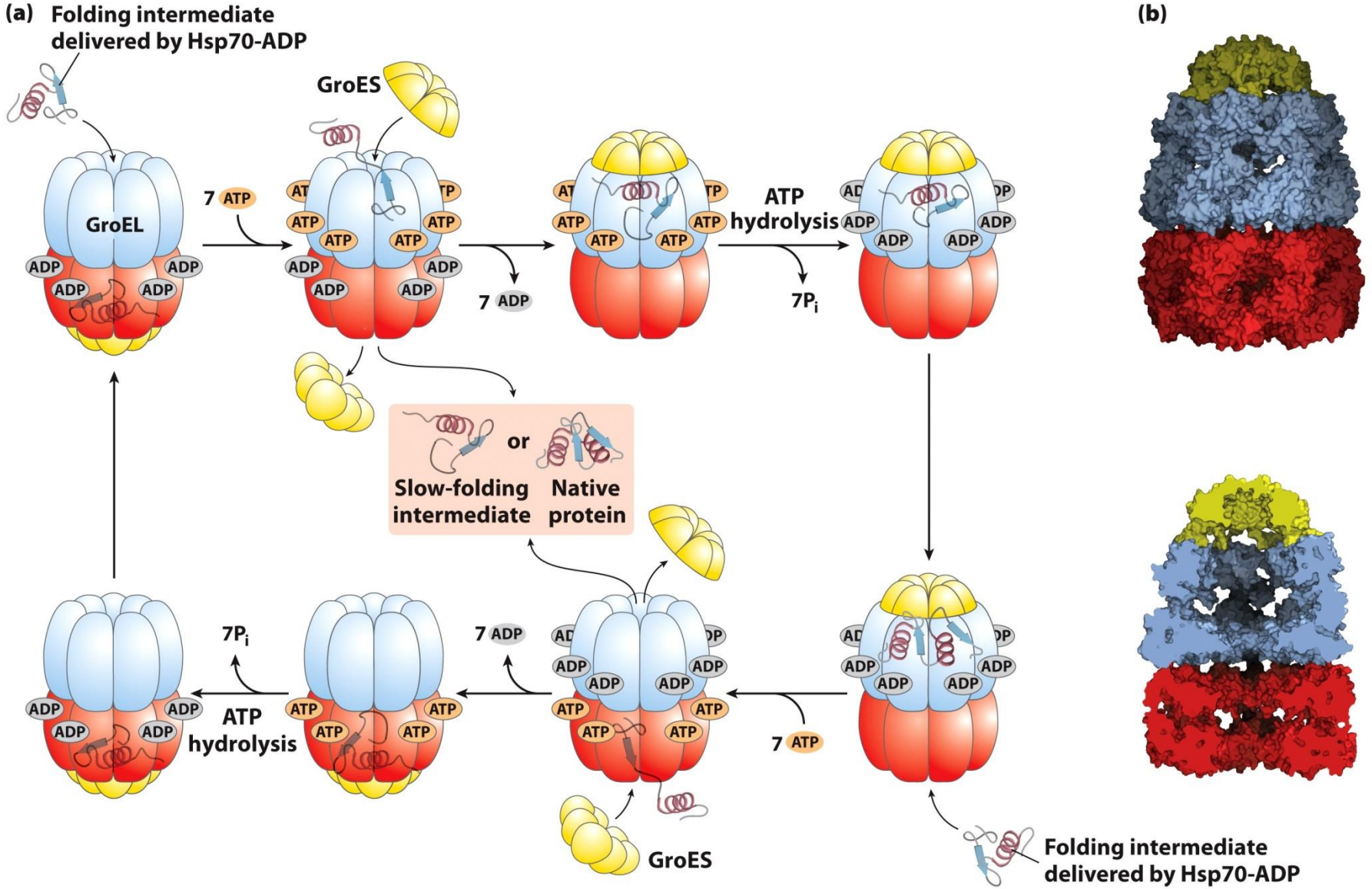


Figure 4-31
Lehninger Principles of Biochemistry, Sixth Edition

Chaperonins in protein folding

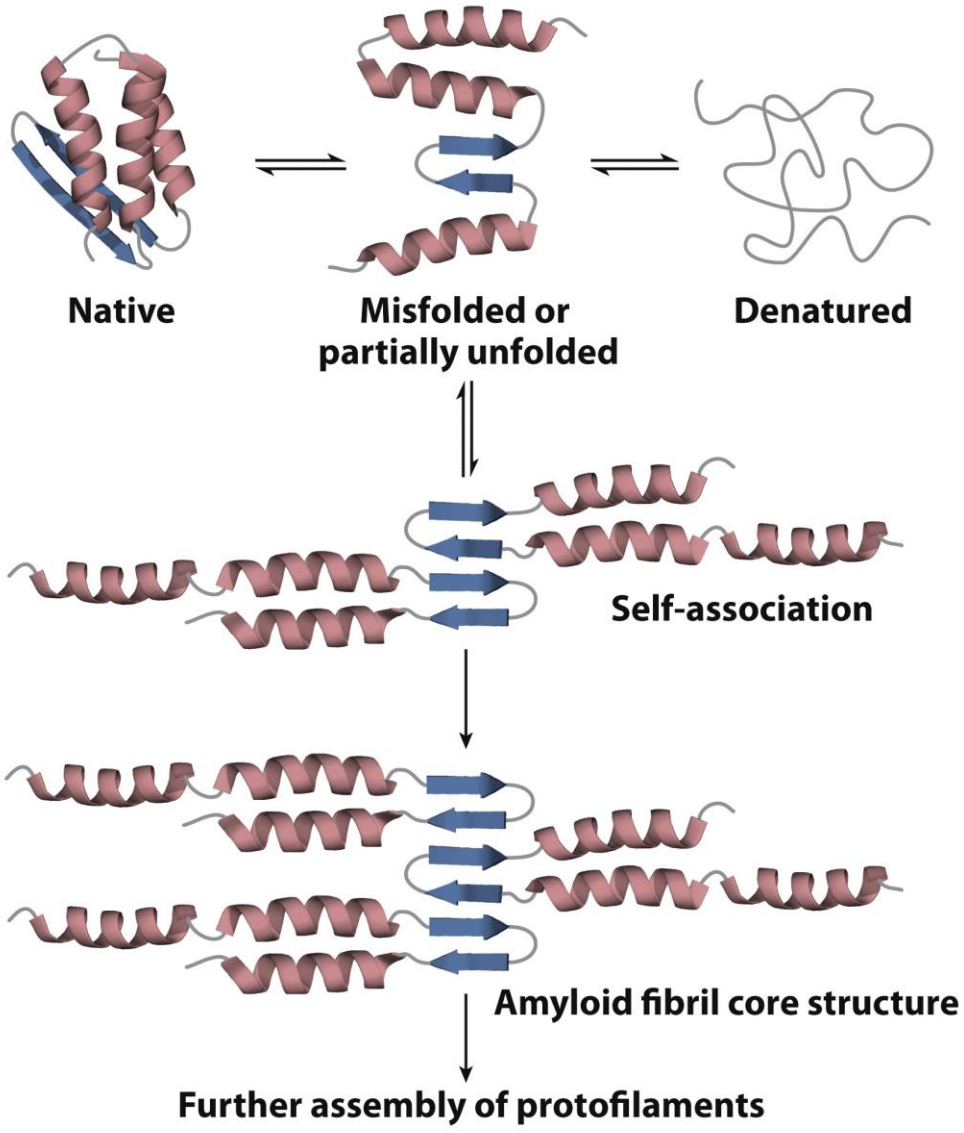


Figure 4-32a
Lehninger Principles of Biochemistry, Sixth Edition

Protein misfolding is the molecular basis of a wide region of human diseases.

Formation of disease-causing amyloid fibrils

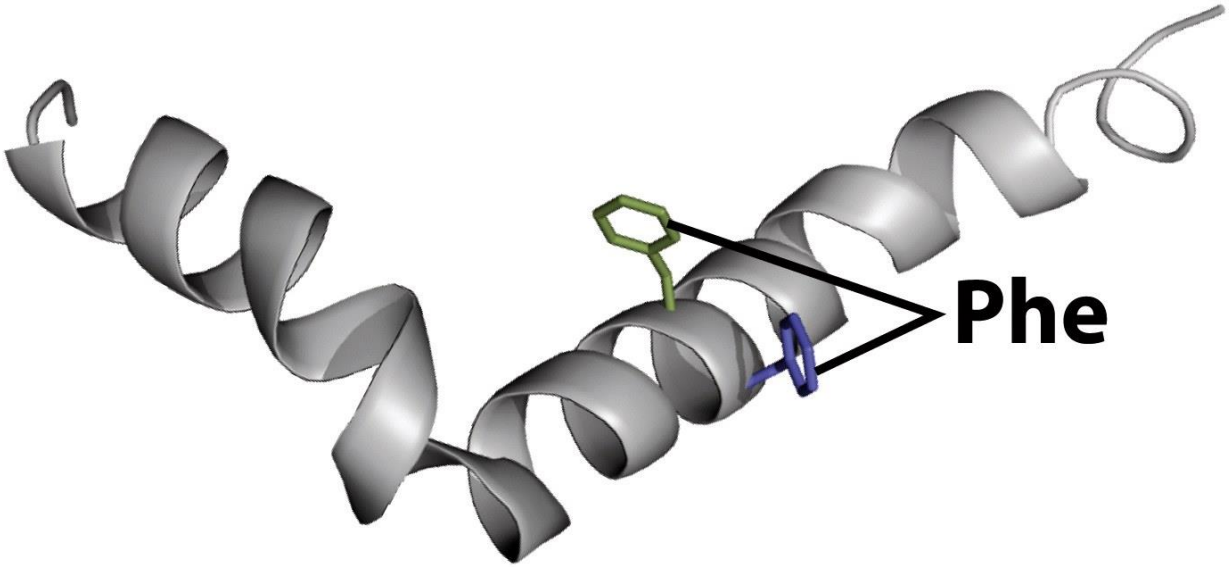


Figure 4-31b
Lehninger Principles of Biochemistry, Fifth Edition

The amyloid-β peptide, which plays a major role in Alzheimer's disease, is derived from a larger transmembrane protein called amyloid- β precursor protein or APP. This protein is found in most human tissues. When it is part of the larger protein, the peptide is composed of two α-helical segments spanning the membrane. When the external and internal domains (each of which have independent functions) are cleaved off by dedicated proteases, the remaining and relatively unstable amyloid-β peptide leaves the membrane and loses its α-helical structure.

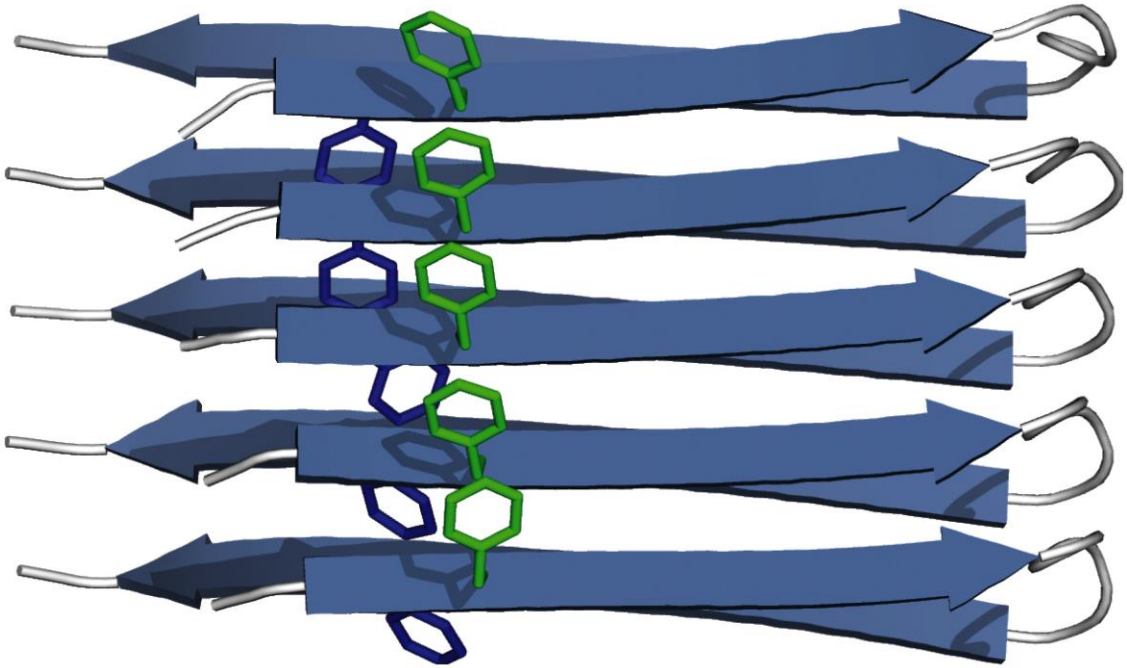


Figure 4-32c
Lehninger Principles of Biochemistry, Sixth Edition

Amyloid fibrils contribute to the characteristic plaques on the exterior of nervous tissue in people with Alzheimer's disease. Amyloid is rich in β-sheet structure, with the β strands arranged perpendicular to the axis of the amyloid fibril.

Determining protein structure by x-ray crystallography

Steps needed:

- **Purify the protein**
- **Crystallize the protein**
- **Collect diffraction data**
- **Calculate electron density**
- **Fit residues into density**

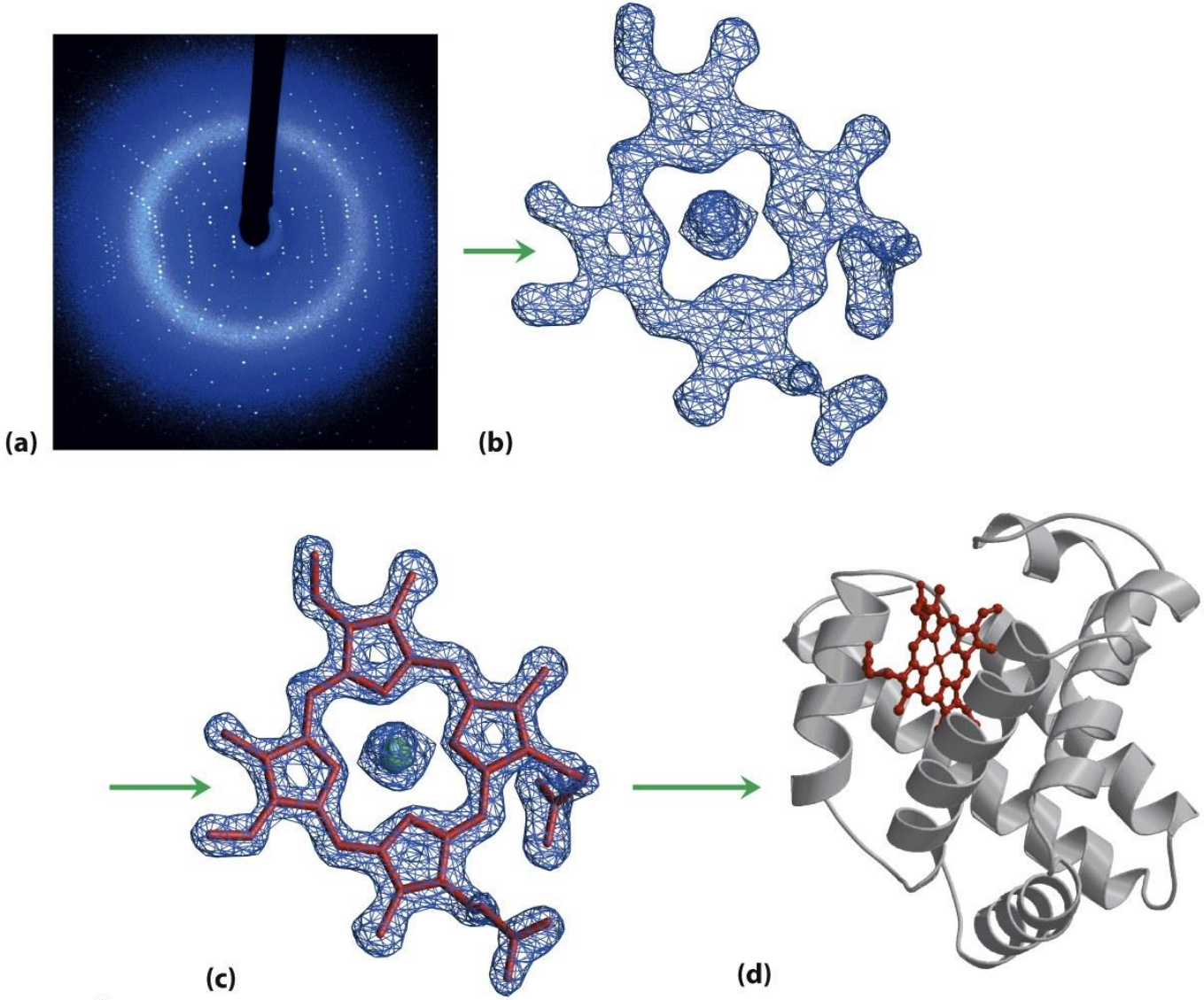
Pros:

- **No size limits**
- **Well-established**

Cons:

- **Difficult for membrane proteins**
- **Cannot see hydrogens**

(read Box 4-5 to get more information)



Box 4-5 figure 1
Lehninger Principles of Biochemistry, Fifth Edition

Steps in determining the structure of sperm whale myoglobin by x-ray crystallography

Determining protein structure by NMR

Steps needed

- **purify the protein**
- **dissolve the protein**
- **collect NMR data**
- **assign NMR signals**
- **calculate the structure**

Pros

- **no need to crystallize the protein**
- **can see many hydrogens**

Cons

- **difficult for insoluble proteins**
- **works best with small proteins**

(read Box 4-5 to get more information)

Protein Database

- **PDB (Protein Data Bank): www.rcsb.org**
- **SCOP (Structural Classification of Protein): http://scop.mrc-lmb.cam.ac.uk/scop**
- **Prosite: http://expasy.org/prosite**
- **PIR (Protein Information Resource): pir.georgetown.edu**
- **NCBI: ncbi.nlm.nih.gov**

Keywords

- **peptide bond**
- **α-helix**
- **β-sheet**
- **protein structure**

Words of the week

- **primary**
- **secondary**
- **tertiary**
- **quaternary**

Textbook-reading of the week

- **Summary 4.1**
- **Summary 4.2**
- **Summary 4.3**
- **Summary 4.4**
- **Protein secondary structure (4.2)**

Summary

- **Structure of the protein is partially dictated by the properties of the peptide bond, which has a partial double-bond character.**
- **The most common regular secondary structures are α-helix, β-sheet and β turns.**
- **Tertiary structure is the complete three-dimensional structure of a polypeptide chain, which can be destroyed by denaturation.**
- **The three-dimensional structure of a protein is determined by its amino acid sequence.**
- **For many proteins, folding is facilitated by Hsp70 chaperones and by chaperonins.**