7 Structure–Function Relationships in Proteins

Diseases can be caused by changes in protein structure that affect the protein's ability to bind other molecules and carry out its function. They also can be caused by conformational changes in proteins that affect their solubility and degradability. In amyloidosis AL, immunoglobulin chains form an insoluble protein aggregate called amyloid in organs and tissues. Alzheimer's disease and familial amyloid polyneuropathy are neurodegenerative diseases characterized by the deposition of amyloid. Prion diseases result from misfolding and aggregation of a normal cellular protein. Even in sickle cell anemia, the mutation in hemoglobin principally affects the quarternary structure of hemoglobin and its solubility, and not its ability to bind oxygen.

A multitude of different proteins can be formed from only 20 common amino acids because these amino acids can be linked together in an enormous variety of sequences determined by the genetic code. The sequence of amino acids, its primary structure, determines the way a protein folds into a unique three-dimensional structure, which is its native conformation. Once folded, the three-dimensional structure of a protein forms binding sites for other molecules, thereby dictating the function of the protein in the body. In addition to creating binding sites, a protein must fold in such a way that it is flexible, stable, able to function in the correct site in the cell, and capable of being degraded by cellular enzymes.

Levels of protein structure. *Protein structure is described in terms of four different levels: primary, secondary, tertiary, and quarternary (Fig. 7.1). The* ***primary structure*** *of a protein is the linear sequence of amino acids in the polypeptide chain.* ***Secondary structure*** *consists of local regions of polypeptide chains formed into structures generally stabilized by hydrogen bonds, such as the regular structures called α-helices and β-sheets. The* ***rigidity of the peptide backbone*** *determines the types of secondary structure that can occur. The* ***tertiary structure*** *involves folding of the secondary structural elements into an overall three-dimensional conformation. In* ***globular proteins*** *such as* ***myoglobin****, the tertiary structure generally forms a densely packed* ***hydrophobic core*** *with polar amino acid side chains on the outside. Some proteins exhibit* ***quaternary structure,*** *the combination of two or more* ***subunits,*** *each composed of a polypeptide chain.*

Domains and folds. *The tertiary structure of a globular protein is made up of structural* ***domains,*** *regions of structure that are recognized as separate and linked to other domains in a simple way. Within a domain, a combination of secondary*

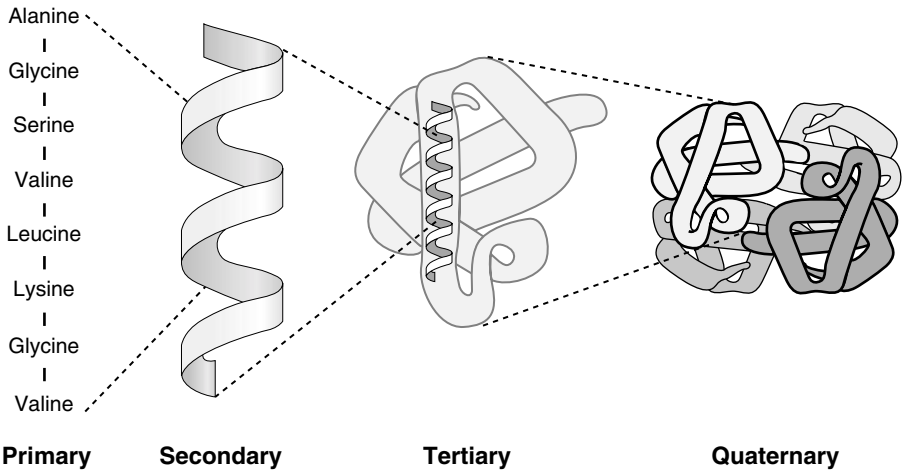


Fig. 7.1. Levels of structure in a protein.

structural elements forms a fold, such as the ***nucleotide binding fold,*** *or an* ***actin fold.*** *Folds are defined by their similarity in a number of different proteins.*

Quarternary structure. *Assembly of globular polypeptide subunits into a multisubunit complex can provide the opportunity for* ***cooperative binding*** *of ligands (e.g., O_2 binding to hemoglobin), form* ***binding sites*** *for complex molecules (e.g., antigen binding to immunoglobulin), and increase* ***stability*** *of the protein. The polypeptide chains of* ***fibrous proteins*** *such as* ***collagen*** *are aligned along an axis, have repeating elements, and are extensively linked to each other through hydrogen bonds.*

Ligand binding. *Proteins form binding sites for a specific molecule, called a* ***ligand*** *(e.g., ATP or O_2) or for another protein. The affinity of a binding site for its ligand is quantitatively characterized by an* ***association*** *or* ***affinity constant, K_a, (or its dissociation constant, K_d).***

Folding of proteins. *The primary structure of a protein dictates the way that it folds into its tertiary structure, which is a* ***stable conformation*** *that is identical to the shape of other molecules of the same protein (that is, its* ***native conformation.)*** ***Chaperonins*** *act as templates to overcome the kinetic barrier to reaching a stable conformation.* ***Prion proteins*** *cause neurodegenerative diseases by acting as a template for misfolding. Heat, acid, and other agents cause proteins to* ***denature,*** *that is, to unfold or refold and lose their native three-dimensional conformation.*

THE WAITING ROOM

Will Sichel, who has sickle cell anemia, was readmitted to the hospital with symptoms indicating that he was experiencing another sickle cell crisis (see Chapter 6).

Anne Jeina is a 54-year-old woman who arrived in the hospital 4 days ago, approximately 5 hours after she began to feel chest pain (see Chapter 6). In the emergency room, the physician had drawn blood for the measurement of myoglobin, CK-MB (creatine kinase, muscle-brain fraction) and cTN-T (cardiac troponin T subunit). The results from these tests had supported the diagnosis of an acute MI (myocardial infarction), and Mrs. Jeina was hospitalized.

Amy Lloyd is a 62-year-old woman who presents with weakness, fatigue, an enlarged tongue (macroglossia), and edema. She has signs and symptoms of cardiac failure, including electrocardiographic abnormalities. Initial laboratory studies showed a serum creatinine of 1.9 mg/dL (reference range [females] = 0.5–1.1) indicating mild renal failure. A urinalysis indicated the presence of a moderate proteinuria and numerous white blood cells in the urinary sediment. She was subsequently diagnosed with amyloidosis/AL secondary to a plasma cell dyscrasia.

Amyloidosis is a term encompassing many diseases that share a common feature—the extracellular deposition of pathologic insoluble fibrillar proteins called amyloid in organs and tissues. In **Amy Lloyd**'s disease, amyloidosis/AL, the amyloid is derived from immunoglobulin light chains (AL = **a**myloidosis, **l**ight-chain related).

Di Abietes returned to her physician's office for a routine visit to monitor her treatment (see Chapters 4, 5, and 6.) Her physician drew blood for an HbA_{1c} (pronounced hemoglobin A-1- c) determination. The laboratory reported a value of 8.5%, compared with to a normal reference range of 5.8 to 7.2%.

I. GENERAL CHARACTERISTICS OF THREE-DIMENSIONAL STRUCTURE

The overall conformation of a protein, the particular position of the amino acid side chains in three-dimensional space, gives a protein its function.

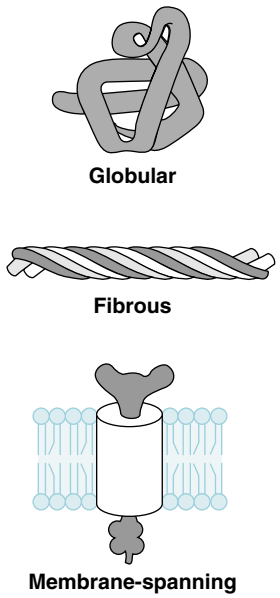


Fig. 7.2. General shapes of proteins.

Requirements of a Protein Structure
Function
Binding specificity
Flexibility
Solubility or lipophilicity
Stability
Degradability

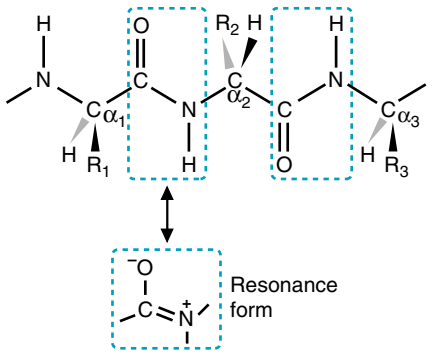


Fig. 7.3. The peptide backbone. Because of the resonance nature of the peptide bond, the C and N of the peptide bonds form a series of rigid planes. Rotation within allowed torsion angles can occur around the bonds attached to the α-carbon. The side chains are *trans* to each other, and alternate above and below the peptide chain. The actual peptide bond is a hybrid between the resonance forms shown, resulting in a partial negative charge on the carbonyl oxygen, a partial positive charge on the nitrogen, and partial double bond character for the peptide bond itself.

A. Descriptions of Protein Structure

Proteins are generally grouped into major structural classifications: globular proteins, fibrous proteins, transmembrane proteins, and DNA-binding proteins (Fig. 7.2). Globular proteins are usually soluble in aqueous medium and resemble irregular balls. The fibrous proteins are geometrically linear, arranged around a single axis, and have a repeating unit structure. Another general classification, transmembrane proteins, consists of proteins that have one or more regions aligned to cross the lipid membrane. DNA-binding proteins, usually classified separately, are considered in Chapter 16.

The structure of these proteins is often described according to levels called primary, secondary, tertiary, and quaternary structure (see Fig. 7.1). The primary structure is the linear sequence of amino acid residues joined through peptide bonds to form a polypeptide chain. The secondary structure refers to recurring structures (such as the regular structure of the α-helix) that form in short localized regions of the polypeptide chain. The overall three-dimensional conformation of a protein is its tertiary structure. The quaternary structure is the association of polypeptide subunits in a geometrically specific manner.

B. Requirements of the Three-Dimensional Structure

The overall three-dimensional structure of a protein must meet certain requirements to enable the protein to function in the cell or extracellular medium of the body. The first requirement is the creation of a binding site that is specific for just one molecule, or a group of molecules with similar structural properties. The specific binding sites of a protein usually define its role. The three-dimensional structure also must exhibit the degrees of flexibility and rigidity appropriate to its function. Some rigidity is essential for the creation of binding sites and for a stable structure (i.e., a protein that just flopped all over the place could not accomplish anything). However, flexibility and mobility in structure enables the protein to fold as it is synthesized, and to adapt as it binds other proteins and small molecules. The three-dimensional structure must have an external surface appropriate for its environment (e.g., cytoplasmic proteins need to keep polar amino acids on the surface to remain soluble in an aqueous environment). In addition, the conformation must also be stable, with little tendency to undergo refolding into a form that cannot fulfill its function or that precipitates in the cell. Finally, the protein must have a structure that can de degraded when it is damaged or no longer needed in the cell. Almost every region in the sequence of amino acids, the primary structure, participates in fulfilling one or more of these requirements through the chemical properties of the peptide bonds and the individual amino acid side chains.

II. THE THREE-DIMENSIONAL STRUCTURE OF THE PEPTIDE BACKBONE

The amino acids in a polypeptide chain are sequentially joined by peptide bonds between the carboxyl group of one amino acid and amide group of the next amino acid in the sequence (Fig. 7.3). Usually the peptide bond assumes a *trans* configuration in which successive α-carbons and their R groups are located on opposite sides of the peptide bond.

The polypeptide backbone can only bend in a very restricted way. The peptide bond itself is a hybrid of two resonance structures, one of which has double bond character, so that the carboxyl and amide groups that form the bond must, therefore, remain planar (see Fig. 7.3.). As a consequence, the peptide backbone consists of a sequence of rigid planes formed by the peptide groups (see Fig. 7.3). However, rotation within certain allowed angles (torsion angles) can occur around the bond between the α-carbon and the α-amino group and around the bond between the α-carbon and the carbonyl group. This rotation is subject to steric constraints that maximize the

distance between atoms in the different amino acid side chains and forbid torsion (rotation) angles that place the side chain atoms too close to each other. These folding constraints, which depend on the specific amino acids present, limit the secondary and tertiary structures that can be formed from the polypeptide chain.

III. SECONDARY STRUCTURE

Regions within polypeptide chains form recurring, localized structures known as secondary structures. The two regular secondary structures called the α-helix and the α-sheet contain repeating elements formed by hydrogen bonding between atoms of the peptide bonds. Other regions of the polypeptide chain form nonregular nonrepetitive secondary structures, such as loops and coils.

A. The α-Helix

The α-helix is a common secondary structural element of globular proteins, membrane-spanning domains, and DNA-binding proteins. It has a stable rigid conformation that maximizes hydrogen bonding while staying within the allowed rotation angles of the polypeptide backbone. The peptide backbone of the α-helix is formed by strong hydrogen bonds between each carbonyl oxygen atom and the amide hydrogen (N-H) of an amino acid residue located four residues further down the chain (Fig. 7.4). Thus, each peptide bond is connected by hydrogen bonds to the peptide bond four amino acid residues ahead of it and four amino acid residues behind it in the amino acid sequence. The core of the helix is tightly packed, thereby maximizing association energies between atoms. The *trans* side chains of the amino acids project backward and outward from the helix, thereby avoiding steric hindrance with the polypeptide backbone and with each other (Fig. 7.5).

B. β-Sheets

β-Sheets are a second type of regular secondary structure that maximizes hydrogen bonding between the peptide backbones while maintaining the allowed torsion angles. In β-sheets, the hydrogen bonding usually occurs between regions of separate neighboring polypeptide strands aligned parallel to each other (Fig. 7.6). Thus, the carbonyl oxygen of one peptide bond is hydrogen-bonded to the nitrogen of a peptide bond on an adjacent strand. (This pattern contrasts with the α-helix in which the peptide backbone hydrogen bonds are within the same strand.) Optimal hydrogen bonding occurs when the sheet is bent (pleated) to form β-pleated sheets.

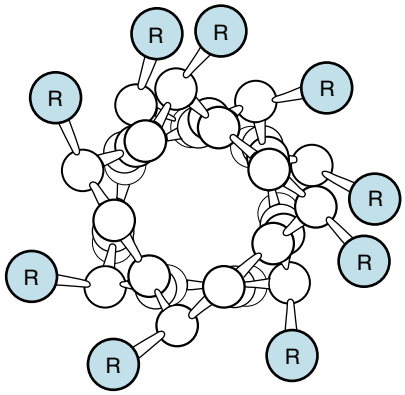


Fig. 7.5. A view down the axis of an α-helix. The side chains (R) jut out from the helix. Steric hindrance occurs if they come within their van der Waals radii of each other, and a stable helix cannot form.

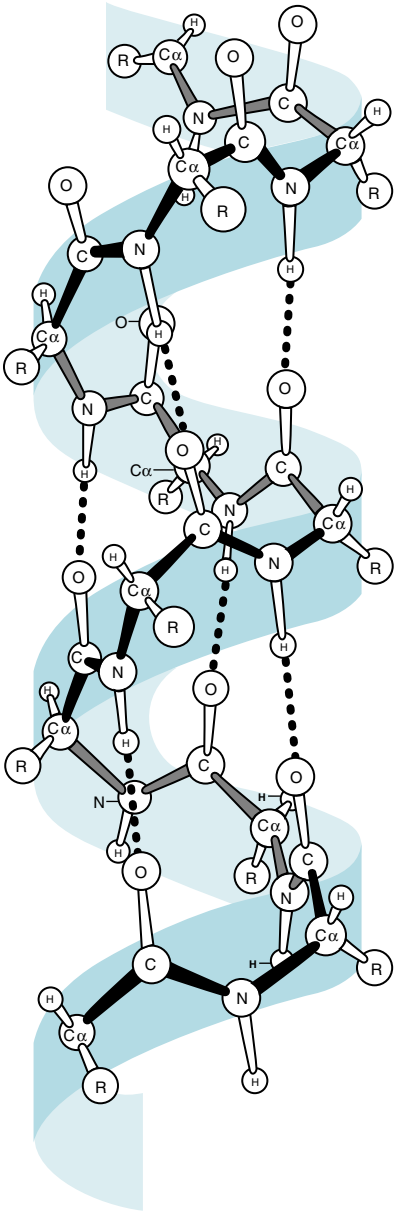


Fig. 7.4. The α-helix. Each oxygen of a carbonyl group of a peptide bond forms a hydrogen bond with the hydrogen atom attached to a nitrogen atom in a peptide bond four amino acids further along the chain. The result is a highly compact and rigid structure.

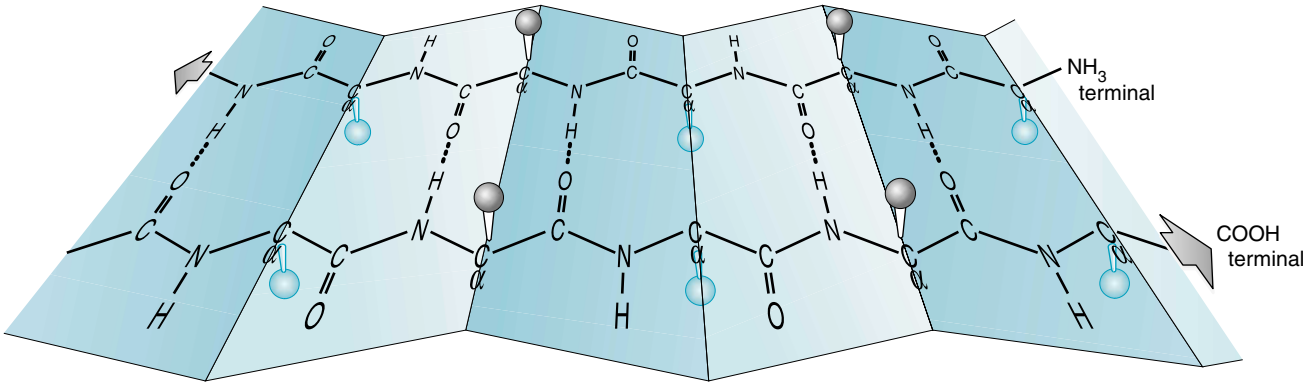


Fig. 7.6. A β-pleated sheet. In this case, the chains are oriented in opposite directions (antiparallel). The large arrows show the direction of the carboxy terminal. The amino acid side chains (R) in one strand are *trans* to each other, and alternate above and below the plane of the sheet, which can have a hydrophobic face and a polar face that engages in hydrogen bonding.

A renal biopsy used in the diagnosis of **Amy Lloyd**'s disease showed amorphous deposits in the glomeruli. When stained with Congo red dye, these deposits appeared red with ordinary light microscopy and exhibited apple green fluorescence when viewed in polarized light. This staining is characteristic of the amyloid fibril structure, which is composed of repeated β-sheets aligned orthogonally (perpendicular) to the axis of the fiber.

A number of diseases involve deposition of a characteristic amyloid fiber. However, in each of these diseases, the amyloid is derived from a different protein that has changed its conformation (three-dimensional structure) to that of the amyloid repeated β-sheet structure. Once amyloid deposition begins, it seems to proceed rapidly, as if the fibril itself were promoting formation and deposition of more fibrils (a phenomenon called "seeding"). The different clinical presentations in each of these diseases results from differences in the function of the native protein and the site of amyloid deposition.

The β-pleated sheet is described as parallel if the polypeptide strands run in the same direction (as defined by their amino and carboxy terminals) and anti-parallel if they run in opposite directions. Antiparallel strands are often the same polypeptide chain folded back on itself, with simple hairpin turns or long runs of polypeptide chain connecting the strands. The amino acid side chains of each polypeptide strand alternate between extending above and below the plane of the β-sheet (see Fig. 7.6). Parallel sheets tend to have hydrophobic residues on both sides of the sheets; antiparallel sheets usually have a hydrophobic side and a hydrophilic side. Frequently, sheets twist in one direction.

C. Nonrepetitive Secondary Structures

α-Helices and β-pleated sheets are patterns of regular structure with a repeating element, the turn of a helix or a pleat. In contrast, bends, loops, and turns are nonregular secondary structures that do not have a repeating element. They are characterized by an abrupt change of direction and are often found on the protein surface. For example, β-turns are short regions usually involving four successive amino acid residues. They often connect strands of antiparallel β-sheets (Fig. 7.7.). The surface of large globular proteins usually has at least one omega loop, a structure with a neck like the capital Greek letter omega, Ω.

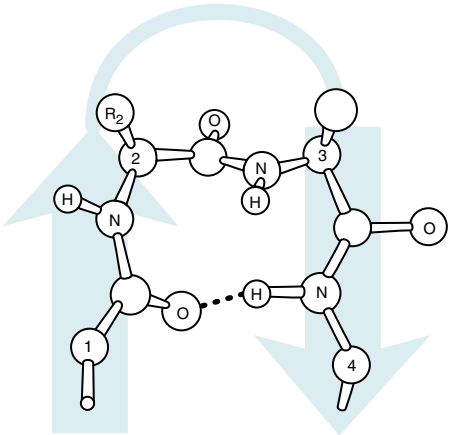


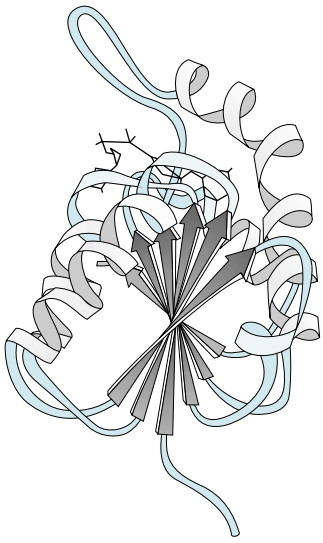
Fig. 7.7. β-turn. The four amino acid residues that form the β-turn (also called a hairpin loop) are held together by hydrogen bonds, which make this an extremely stable structure.

D. Patterns of Secondary Structure

Figure 7.8 is a three-dimensional drawing of a globular domain in the soluble enzyme lactate dehydrogenase (LDH). It illustrates the combination of secondary structural elements to form patterns. This LDH domain is typical of globular proteins, which average approximately 31% α-helical structure and approximately 28% β-pleated sheets (with a wide range of variation.) The helices of globular domains have an average span of approximately 12 residues, corresponding to approximately three to four helical turns, although many are much longer. The β-sheets, represented in diagrams by an arrow for each strand, are an average of six residues long and six strands wide (2–15 strands). Like the β-sheet in the lactate dehydrogenase domain, they generally twist to the right, rather than lie flat (see Fig. 7.8). Most globular domains, such as this LDH domain, also contain motifs. Motifs are relatively small arrangements of secondary structure recognized in many different proteins. For example, certain of the β-strands are connected with α-helices to form the βα βα β structural motif.

The remaining polypeptide segments connecting the helices and β-sheets are said to have a coil or loop conformation (see Fig. 7.8). Although some of the connecting segments recognized in many proteins have been given names (like the Ω-loops), other segments such as those in this LDH domain appear disordered or irregular. These nonregular regions, generally called coils, should never be referred to as "random coils." They are neither truly disordered nor random; they are stabilized through specific hydrogen bonds dictated by the primary sequence of the protein and do not vary from one molecule of the protein to another of the same protein.

The nonregular coils, loops, and other segments are usually more flexible than the relatively rigid helices and β-pleated sheets. They often form hinge regions that allow segments of the polypeptide chain to move as a compound binds or to move as the protein folds around another molecule.



Lactate dehydrogenase domain 1

Fig. 7.8. Ribbon drawing showing the arrangement of secondary structures into a three-dimensional pattern in domain 1 of lactate dehydrogenase. The individual polypeptide strands in the six-stranded β-sheet are shown with arrows. Different strands are connected by helices and by nonrepetitive structures (turns, coils and loops), shown in blue. This domain is the nucleotide binding fold. NAD^+ is bound to a site created by the helices (upper left of figure.) (Modified from Richardson JS. Adv Protein Chem. The anatomy and taxonomy of protein structure 1981;34:167).

Although it is usually assumed that proteins can have truly disordered regions, the more that is learned about protein structure, the less disordered these regions seem. Even regions that look truly disordered may form a specific binding site for another molecule (e.g., the intracellular C-terminal of transmembrane proteins). Eventually all regions of all proteins may be classified as a particular named pattern.

Flexibility is one of the most important features of protein structure. Although every portion of every amino acid in a protein is engaged in bonding to the rest of the protein, to water, or to a ligand, proteins do not have rigid structures. The three-dimensional structure is flexible and dynamic, with rapidly fluctuating movement in the exact position of amino acid side chains and domains. These fluctuations are like wiggling or shaking that can occur without unfolding. They allow ions and water to diffuse through the structure and provide alternate conformations for ligand binding.

IV. TERTIARY STRUCTURE

The tertiary structure of a protein is the folding pattern of the secondary structural elements into a three-dimensional conformation, as shown for the LDH domain in Figure 7.8. As illustrated with examples below, this three-dimensional structure is designed to serve all aspects of the protein's function. It creates specific and flexible binding sites for ligands (the compound that binds), illustrated with actin and myoglobin. The tertiary structure also maintains residues on the surface appropriate for the protein's cellular location, polar residues for cytosolic proteins, and hydrophobic residues for transmembrane proteins (illustrated with the β_2-adrenergic receptor).

A. Domains in the Tertiary Structure

The tertiary structure of large complex proteins is often described in terms of physically independent regions called structural domains. You can usually identify domains from visual examination of a three-dimensional figure of a protein, such as the three-dimensional figure of G-actin shown in Fig. 7.9. Each domain is formed from a continuous sequence of amino acids in the polypeptide chain that are folded into a three-dimensional structure independently of the rest of the protein, and two domains are connected through a simpler structure like a loop (e.g., the hinge region of Fig. 7.9). The structural features of each domain can be discussed independently of another domain in the same protein, and the structural features of one domain may not match that of other domains in the same protein.

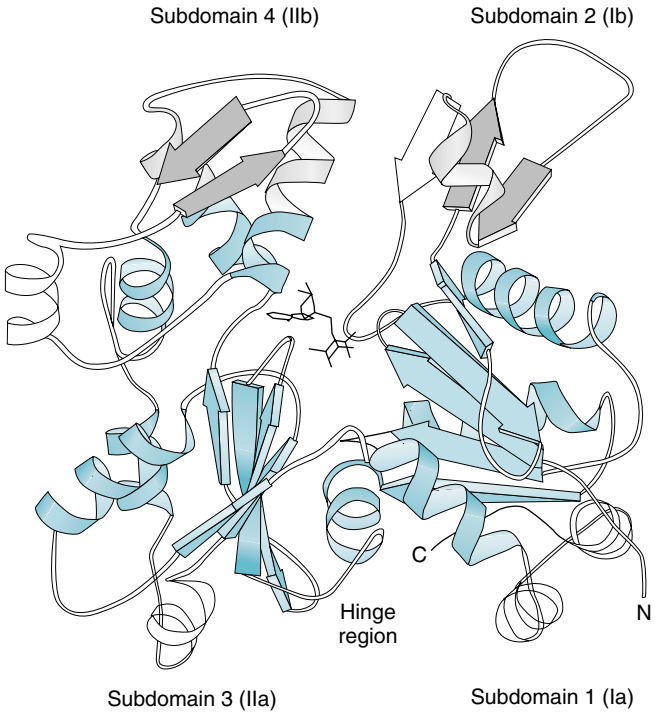


Fig. 7.9. G-Actin. ATP binds in the center of the cleft. The two domains that form the cleft are further subdivided into subdomains 1–4. The overall structure is found in many ATP-binding proteins and is called the actin fold. The conformations of the regions shown in blue are nearly superimposable among the proteins that contain the actin fold. (From Kabsch W, Holmes KC. The actin fold. FASEB J 1995:9:167–174.)

B. Folds in Globular Proteins

Folds are relatively large patterns of three-dimensional structure that have been recognized in many proteins, including proteins from different branches of the phylogenetic tree. A characteristic activity is associated with each fold, such as ATP binding and hydrolysis (the actin fold) or NAD^+ binding (the nucleotide-binding fold). These three examples of fold patterns are discussed below.

More than 1000 folds have now been recognized, and it is predicted that there are only a few thousand different folds for all the proteins that have ever existed. One of the values of classifying proteins into fold patterns is that a specific function, such as ATP binding and hydrolysis, is associated with each fold. Currently, we know the primary structure of many proteins involved in inherited diseases without knowing the function of the protein. However, the function of a domain in a protein can be predicted if its amino acid sequence can be matched to that of other proteins in a fold family with a known function. Classifying proteins into fold families also may be useful for drug design. The structure of a bacterial protein targeted by a drug can be compared with that of human proteins in the same fold family so that an antibacterial drug can be designed that more specifically targets the bacterial protein.

1. THE ACTIN FOLD

In the three-dimensional drawing of G-actin shown in Figure 7.9, all four subdomains contribute to a folding pattern called the actin fold, named for the first protein in which it was described. ATP is bound into the middle of the cleft of the actin fold by amino acid residues contributed by domains on both sides; thus, ATP binding promotes a conformational change that closes the cleft. Once bound, ATP is cleaved to ADP and phosphate.

The actin fold is found in proteins as diverse as actin, which polymerizes to form the cytoskeleton, heat shock protein 70 (hsp70), which uses ATP energy in changing the conformation of other proteins and hexokinase, which catalyzes phosphorylation of glucose (see Chapter 8 for further discussion of hexokinase.) Although these proteins have very little sequence identity, three-dimensional drawings of their actin folds are almost superimposable. The amount of sequence identity they do have is consistent with their membership in the same fold family and establishes that they are all homologs of the same ancestral protein. In all of these proteins, ATP binding results in large conformational changes that contribute to the function of the protein.

G-actin subunits polymerize to form F-actin, which forms the cytoskeleton (see Chapter 10). The conformational change caused by ATP binding to the G-actin subunits promotes their addition to the growing ends of the F-actin polymers. Hydrolysis of bound ATP to ADP by the G-actin subunit promotes dissociation.

2. THE NUCLEOTIDE BINDING FOLD

A fold also can be formed by one domain. In the example of secondary structures provided by lactate dehydrogenase (see Fig. 7.8), domain 1 alone forms the nucleotide binding fold. This fold is a binding site for NAD^+ or, in other proteins, molecules with a generally similar structure (e.g., riboflavin). However, many proteins that bind NAD^+ or $NADP^+$ contain a very different fold from a separate fold family. These two different NAD^+ binding folds arise from different ancestral lines and have different structures, but have similar properties and function. They are believed to be the product of convergent evolution.

C. The Solubility of Globular Proteins in an Aqueous Environment

Most globular proteins are soluble in the cell. In general, the core of a globular domain has a high content of amino acids with nonpolar side chains (val, leu, ile, met, and phe), out of contact with the aqueous medium. This hydrophobic core is densely packed to maximize attractive van der Waals forces, which exert themselves over short distances. The charged polar amino acid side chains (arg, his, lys, asp, and glu) are generally located on the surface of the protein, where they form ion pairs (salt bridges) or are in contact with aqueous solvent. Charged side chains often bind inorganic ions (e.g., K^+, PO_4^{3-} or Cl^-) to decrease repulsion between like charges. When charged amino acids are located on the interior, they are generally involved in forming specific binding sites. The polar uncharged amino acid side chains of ser, thr, asn, gln, tyr, and trp are also usually found on the surface of the protein, but may occur in the interior, hydrogen bonded to other side chains. Although cystine disulfide bonds (the bond formed by two cysteine sulfhydryl groups) are sometimes involved in the formation of tertiary structure, they are generally not needed.

Many features of the adrenaline (β2-adrenergic) receptor are typical of hormone receptors. Each extracellular loop is actually a structural domain, but several loops together form the binding site for the hormone. The binding site is sometimes referred to as a binding domain (a functional domain), even though it is not formed from a continuous segment of the polypeptide chain. The amino terminus (residues 1–34) extends out of the membrane and has branched high mannose oligosaccharides linked through N-glycosidic bonds to the amide of asparagine. It is anchored in the lipid plasma membrane by a palmitoyl group (shown as a squlggle in Fig. 7.10) that forms a thioester with the SH residue of a cysteine. The COOH terminus, which extends into the cytoplasm, has a number of serine and threonine phosphorylation sites (shown as blue circles in Fig. 7.10).

D. Tertiary Structure of Transmembrane Proteins

Transmembrane proteins, such as the β_2-adrenergic receptor, contain membrane-spanning domains and intracellular and extracellular domains on either side of the membrane (Fig. 7.10). Many ion channel proteins, transport proteins, neurotransmitter receptors, and hormone receptors contain similar membrane-spanning segments that are α-helices with hydrophobic residues exposed to the lipid bilayer. These rigid helices are connected by loops containing hydrophilic amino acid side chains that extend into the aqueous medium on both sides of the membrane. In the β_2-adrenergic receptor, the helices clump together so that the extracellular loops form a surface that acts as a binding site for the hormone adrenaline (epinephrine), our fight or flight hormone. The binding site is sometimes referred to as a binding domain (a functional domain), even though it is not formed from a continuous segment of the polypeptide chain. Once adrenaline binds to the receptor, a conformational change in the arrangement of rigid helical structures is transmitted to the intracellular domains that form a binding site for another signaling protein, a heterotrimeric G protein (a guanosine triphosphate [GTP]-binding protein composed of three different subunits). Thus, receptors require both rigidity and flexibility to transmit signals across the cell membrane.

As discussed in Chapter 6, transmembrane proteins usually have a number of posttranslational modifications that provide additional chemical groups to fulfill requirements of the three-dimensional structure. The amino terminus (residues 1–34) extends out of the membrane and has branched high mannose oligosaccharides linked through *N* -glycosidic bonds to the amide of asparagine (see Fig. 7.10). It is anchored in the lipid plasma membrane by a palmitoyl group that forms a thioester with the SH residue of a cysteine. The COOH terminus, which extends

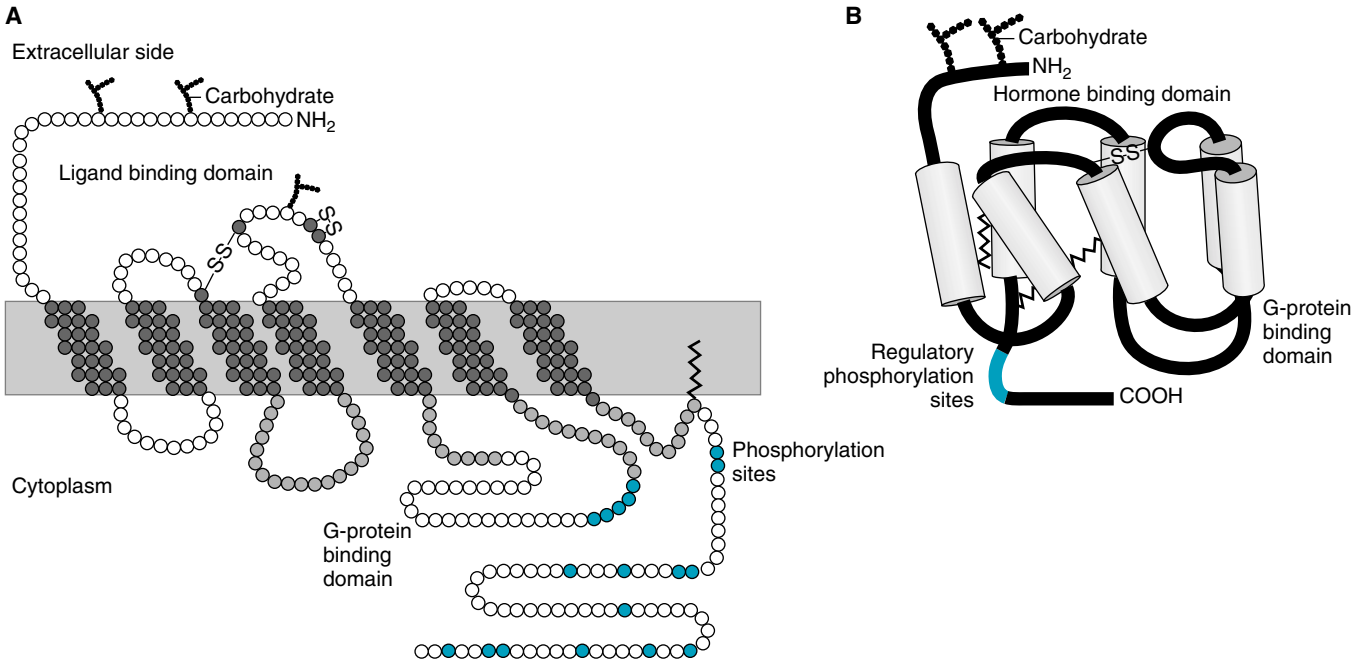


Fig. 7.10. β_2-Adrenergic receptor. The receptor has seven α-helix domains that span the membrane and is therefore a member of the heptahelical class of receptors. **A.** The transmembrane domains are drawn in an extended form. **B.** The seven transmembrane helices (shown as tubes) form a cylindrical structure. Loops connecting helices form the hormone binding site on the external side of the plasma membrane, and a binding site for a G-protein is on the intracellular side. The protein also contains oligosaccharide chains, palmityl groups, and phosphorylation sites (shown in blue; see Chapter 6).

into the cytoplasm, has a number of serine and threonine phosphorylation sites (shown as blue circles) that regulate receptor activity.

V. QUATERNARY STRUCTURE

The quaternary structure of a protein refers to the association of individual polypeptide chain subunits in a geometrically and stoichiometrically specific manner. Many proteins function in the cell as dimers, tetramers, or oligomers, proteins in which two, four, or more subunits, respectively, have combined to make one functional protein. The subunits of a particular protein always combine in the same number and in the same way, because the binding between the subunits is dictated by the tertiary structure, which is dictated by the primary structure, which is determined by the genetic code.

A number of different terms are used to describe subunit structure. The prefixes "homo" or "hetero" are used to describe identical or different subunits, respectively, of 2, 3, or 4 subunit proteins (e.g., heterotrimeric G proteins have three different subunits). A protomer is a unit structure composed of nonidentical subunits. In contrast, F-actin is an oligomer, a multisubunit protein composed of identical G-actin subunits. "Multimer" is sometimes used as a more generic term to designate a complex with many subunits of more than one type.

The contact regions between the subunits of globular proteins resemble the interior of a single subunit protein; they contain closely packed nonpolar side chains, hydrogen bonds involving the polypeptide backbones and their side chains, and occasional ionic bonds or salt bridges. The subunits of globular proteins are very rarely held together by interchain disulfide bonds, and never by other covalent bonds. In contrast, fibrous and other structural proteins may be extensively linked to other proteins through covalent bonds.

Assembly into a multisubunit structure increases the stability of a protein. The increase in size increases the number of possible interactions between amino acid residues and therefore makes it more difficult for a protein to unfold and refold. As a result, many soluble proteins are composed of two or four identical or nearly identical subunits with an average size of approximately 200 amino acids.

A multisubunit structure has many advantages besides increased stability. It may enable the protein to exhibit cooperativity between subunits in binding ligands (illustrated later with hemoglobin) or to form binding sites with a high affinity for large molecules (illustrated with antigen binding to the immunoglobulin molecule. IgG). An additional advantage of a multisubunit structure is that the different subunits can have different activities and cooperate in a common function. Examples of enzymes that have regulatory subunits or exist as multiprotein complexes are provided in Chapter 9.

Insulin is composed of two nonidentical polypeptide chains attached to each other through disulfide bonds between the chains (see Chapter 6, Fig. 6.13). The subunits of globular proteins are generally not held together by disulfide bonds, but regions of the same chain may be connected by disulfide bonds that form as the chain folds. Insulin actually fits this generalization because it is synthesized as a single polypeptide chain, which forms the disulfide bonds. Subsequently, a proteolytic enzyme in secretory vesicles clips the polypeptide chain into two nonidentical subunits. Generally, each subunit of most protomers and oligomers is synthesized as a separate polypeptide chain. In fibrous proteins, which have a regular, sometimes repeating sequence of amino acids, interchain binding serves different functions. In collagen, for example, extensive interchain binding provides great tensile strength. The structure of collagen is discussed in Chapter 49 on connective tissue.

Creatine phosphokinase (CK), one of the proteins measured to follow **Ann Jeina**'s myocardial infarction (see Chapter 6) is present in cells as dimers (two subunits). The dimers may be homodimers (two identical subunits of either the M [muscle] isozyme or the B [brain] isozyme), or heterodimers (MB) The MB isozyme is produced only by the heart and readily released from injured cardiomyocytes into the blood (see Chapter 6).

Serum levels of lactate dehydrogenase (LDH) were formerly used to diagnose an acute MI. LDH is present in cells as a tetramer of four identical, or nearly identical, subunits. Each subunit is a separate polypeptide chain with a molecular weight of 35 kD (approximately 35,000 g/mole). These subunits are present as two isoforms, the H isoform (for heart) and the M isoform (for skeletal muscle). Although the heart produces principally the H_4 form (four H subunits combined into a tetramer) and skeletal muscles produce principally the M_4 isoform, the heart, skeletal muscle, and other tissues produce several intermediate combinations (e.g., H_3M, H_2M_2). These tetrameric isoforms all have similar activities, but the individual monomeric subunits are inactive. Measurements of LDH isozymes in the serum are no longer used for diagnosis of a recent MI because the enzyme is large, released slowly, and the isozyme pattern is not as specific for the heart as is CK.

Equation 7.1. The association constant, K_a for a binding site on a protein.

Consider a reaction in which a ligand (L) binds to a protein (P) to form a ligand–protein complex (LP) with a rate constant of k_1. LP dissociates with a rate constant of k_2.:

$$L + P \underset{k_2}{\overset{k_1}{\rightleftharpoons}} LP$$

then,

$$K_{eq} = \frac{k_1}{k_2} = \frac{[LP]}{[L][P]} = \frac{K_a}{K_d} = 1$$

The equilibrium constant, K_{eq}, is equal to the association constant (K_a) or $1/K_d$, the dissociation constant. Unless otherwise given, the concentrations of L, P, and LP are expressed as mol/L, and K_a has the units of $(mol/L)^{-1}$.

VI. QUANTITATION OF LIGAND BINDING

In the examples of tertiary structure discussed above, the folding of a protein created a three-dimensional binding site for a ligand (NAD^+ for the lactate dehydrogenase domain 1, ATP for hexokinase, or adrenaline for the β_2 adrenergic receptor). The binding affinity of a protein for a ligand is quantitatively described by its association constant, K_a, which is the equilibrium constant for the binding reaction of a ligand (L) with a protein (P) (Eq. 7.1).

K_a is equal to the rate constant (k_1) for association of the ligand with its binding site divided by the rate constant (k_2) for dissociation of the ligand–protein complex (LP). K_d, the dissociation constant for ligand–protein binding, is the reciprocal of K_a. The tighter the binding of the ligand to the protein, the higher is the K_a and the lower is the K_d. The K_a is useful for comparing proteins produced by different alleles, or for describing the affinity of a receptor for different drugs.

VII. STUCTURE–FUNCTION RELATIONSHIPS IN MYOGLOBIN AND HEMOGLOBIN

Myoglobin and hemoglobin are two oxygen-binding proteins with a very similar primary structure (Fig. 7.11). However, myoglobin is a globular protein composed of a single polypeptide chain that has one O_2 binding site. Hemoglobin is a tetramer composed of two different types of subunits (2α and 2β polypeptide chains, referred to as two αβ protomers). Each subunit has a strong sequence homology to myoglobin and contains an O_2 binding site. A comparison between myoglobin and hemoglobin illustrates some of the advantages of a multisubunit quaternary structure.

The tetrameric structure of hemoglobin facilitates saturation with O_2 in the lungs and release of O_2 as it travels through the capillary beds (Fig. 7.12). When the amount of oxygen bound to myoglobin or hemoglobin is plotted against the partial pressure of oxygen (pO_2), a hyperbolic curve is obtained for myoglobin, whereas that for hemoglobin is sigmoidal. These curves show that when the pO_2 is high, as in the lungs, both myoglobin and hemoglobin are saturated with oxygen. However,

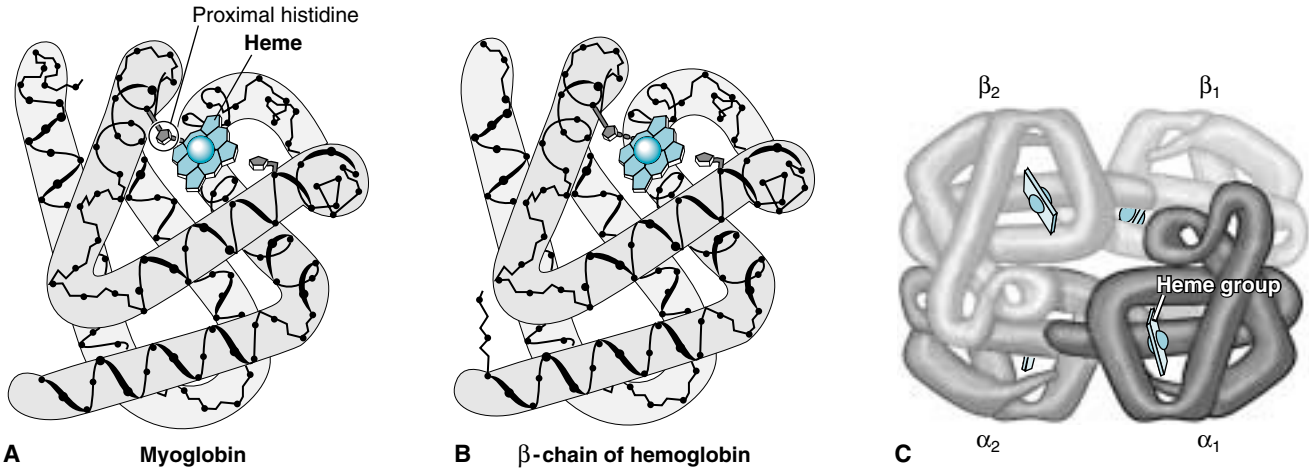


Fig. 7.11. Myoglobin and hemoglobin. Myoglobin consists of a single polypeptide chain, which is similar in structure to the α and β subunits of hemoglobin. In all of the subunits, heme is tightly bound in a hydrophobic binding pocket. The proximal histidine extends down from a helix to bind to the Fe atom. The oxygen binds between the distal histidine and the heme. Panel C displays the quaternary structure of hemoglobin (From Frescht A. Structure and Mechanism in Protein Science. New York: WH Freeman and Company, 1999. Used with permission.)

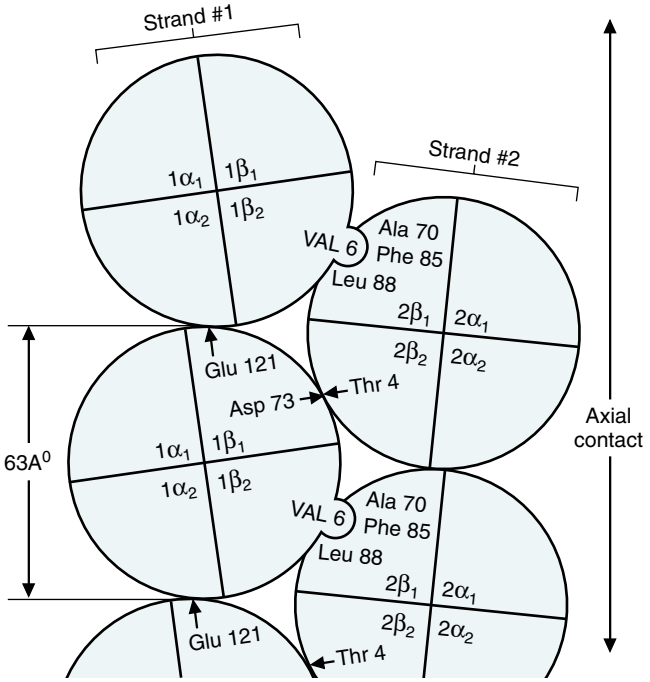
at the lower levels of pO_2 in oxygen-using tissues, hemoglobin cannot bind oxygen as well as myoglobin (i.e., its percent saturation is much lower). Myoglobin, which is present in heart and skeletal muscle, can bind the O_2 released by hemoglobin, which it stores to meet the demands of contraction. As O_2 is used in the muscle cell for generation of ATP during contraction, it is released from myoglobin and picked up by cytochrome oxidase, a heme-containing enzyme in the electron transport chain that has an even higher affinity for oxygen than myoglobin.

A. Oxygen Binding and Heme

The tertiary structure of myoglobin consists of eight α-helices connected by short coils, a structure that is known as the globin fold (see Fig. 7.11). This structure is unusual for a globular protein in that it has no β-sheets. The helices create a hydrophobic O_2 binding pocket containing tightly bound heme with an iron atom (Fe^{2+}) in its center.

Heme consists of a planar porphyrin ring composed of four pyrrole rings that lie with their nitrogen atoms in the center, binding an Fe^{2+} atom (Fig. 7.13).

Sickle cell anemia is really a disease caused by the wrong quarternary structure. The painful vasooclusive crises experienced by **Will Sichel** are caused by the polymerization of HbS molecules into long fibers that distort the shape of the red blood cells into sickle cells. The substitution of a hydrophobic valine for a glutamate in the β_2 chain creates a knob on the surface of deoxygenated hemoglobin that fits into a hydrophobic binding pocket on the β_1 subunit of a different hemoglobin molecule. A third hemoglobin molecule, which binds to the first and second hemoglobin molecules through aligned polar interactions, binds a fourth hemoglobin molecule through its valine knob. Thus the polymerization continues until long fibers are formed.



Polymerization of the hemoglobin molecules is highly dependent on the concentration of HbS and is promoted by the conformation of the deoxygenated molecules. At 100% oxygen saturation, even high concentrations of HbS will not polymerize. A red blood cell spends the longest amount of time at the lower oxygen concentrations of the venous capillary bed, where polymerization is most likely initiated.

Myoglobin is readily released from skeletal muscle or cardiac tissue when the cell is damaged. It has a small molecular weight, 17,000 kDa, and is not complexed to other proteins in the cell. (Da is the abbreviation for Dalton, which is a unit of mass approximately equal to one H atom. Thus, a molecular weight of 17,000 kDa is equal to approximately 17,000 g/mole.) Large injuries to skeletal muscle that result from physical crushing or lack of ATP production result in cellular swelling and the release of myoglobin and other proteins into the blood. Myoglobin passes into the urine and turns the urine red because the heme (which is red) remains covalently attached to the protein. During an acute MI, myoglobin is one of the first proteins released into the blood from damaged cardiac tissue. (However, the amount released is not high enough to cause myoglobinuria.) Laboratory measurements of serum myoglobin are used for early diagnosis in patients such as **Ann Jeina.** Because myoglobin is not present in skeletal muscle and the heart as tissue-specific isozymes, and the amount released from the heart is much smaller than the amount that can be released from a large skeletal muscle injury, myoglobin measurements are not specific for an MI.

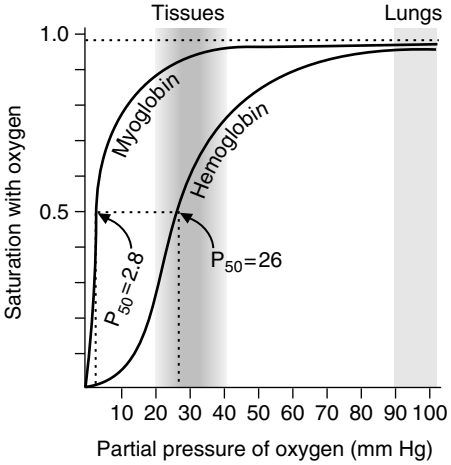


Fig. 7.12. Oxygen saturation curves for myoglobin and hemoglobin. Note that the curve for myoglobin is hyperbolic, whereas that for hemoglobin is sigmoidal. The effect of the tetrameric structure is to inhibit O_2 binding at low O_2 concentrations. P_{50} is the partial pressure of O_2 (pO_2) at which the protein is half-saturated with O_2. P_{50} for myoglobin is 2.8, and that for hemoglobin is 26.

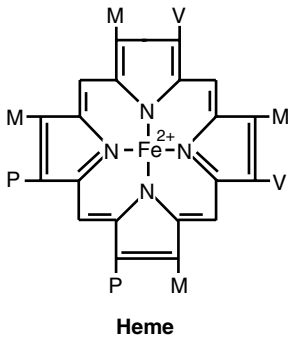


Fig. 7.13. Heme. The Fe is bound to four nitrogen atoms in the center of the heme porphyrin ring. Methyl (M, CH_3), vinyl (V, -CH = CH_2), and propionate (P, $CH_2CH_3COO^-$) side chains extend out from the four pyrrole rings that constitute the porphyrin ring.

Negatively charged propionate groups on the porphyrin ring interact with arginine and histidine side chains from the hemoglobin, and the hydrophobic methyl and vinyl groups that extend out from the porphyrin ring interact with hydrophobic amino acid side chains from hemoglobin. All together, there are approximately 16 different interactions between myoglobin amino acids and different groups in the porphyrin ring.

Organic ligands that are tightly bound, such as the heme of myoglobin, are called prosthetic groups. A protein with its attached prosthetic group is called a holoprotein; without the prosthetic group, it is called an apoprotein. The tightly bound prosthetic group is an intrinsic part of the protein and does not dissociate until the protein is degraded.

Within the binding pocket of myoglobin, O_2 binds directly to the Fe^{2+} atom on one side of the planar porphyrin ring (Fig. 7.14). The Fe^{2+} atom is able to chelate six different ligands; four of the ligand positions are in a plane and taken by the central nitrogens in the planar porphyrin ring. Two ligand positions are perpendicular to this plane. One of these positions is taken by the nitrogen atom on a histidine, called the proximal histidine, which extends down from a myoglobin helix. The other position is taken by O_2.

The proximal histidine of myoglobin and hemoglobin is sterically repelled by the heme porphyrin ring. Thus, when the histidine binds to the Fe^{2+} in the middle of the

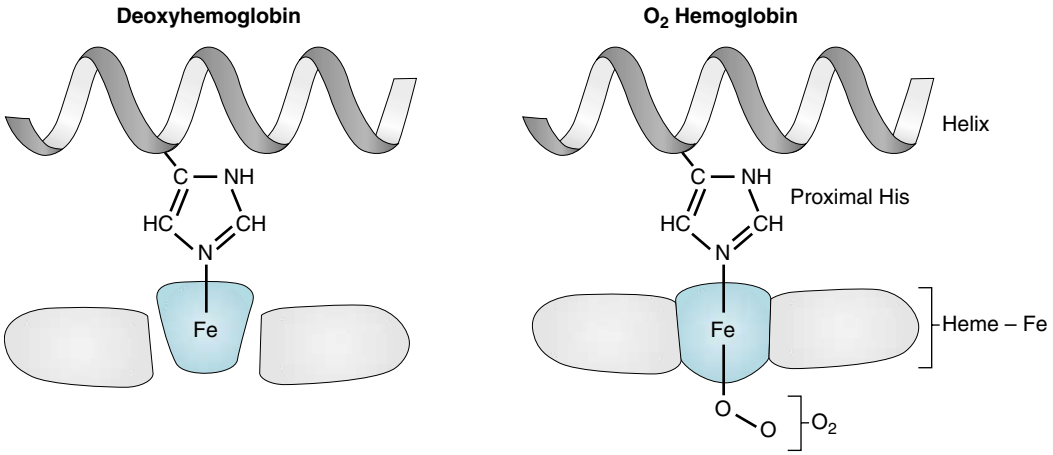


Fig. 7.14. Oxygen binding to the Fe^{2+} of heme in myoglobin. A histidine residue called the proximal histidine binds to the Fe^{2+} on one side of the porphyrin ring; O_2 binds to Fe $^{2+}$ on the other side. O_2 binding causes a conformational change that pulls the Fe^{2+} back into the plane of the ring. As the proximal histidine moves, it moves the helix that contains it.

ring, it pulls the Fe^{2+} above the plane of the ring. When oxygen binds on the other side of the ring, it pulls the Fe^{2+} back into the plane of the ring. The pull of O_2 binding moves the proximal histidine toward the porphyrin ring, which moves the helix containing the proximal histidine. This conformational change has no effect on the function of myoglobin. However, in hemoglobin, the movement of one helix leads to the movement of other helices in that subunit, including one in a corner of the subunit that is in contact with a different subunit through salt bridges. The loss of these salt bridges then induces conformational changes in all other subunits, and all four subunits may change in a concerted manner from their original conformation to a new conformation.

B. Cooperativity of O_2 Binding in Hemoglobin

The cooperativity in oxygen binding in hemoglobin comes from conformational changes in tertiary structure that takes place when O_2 binds (Fig. 7.15). The conformational change of hemoglobin is usually described as changing from a T (tense) state with low affinity for O_2 to an R (relaxed) state with a high affinity for O_2. Breaking the salt bridges in the contacts between subunits is an energy-requiring

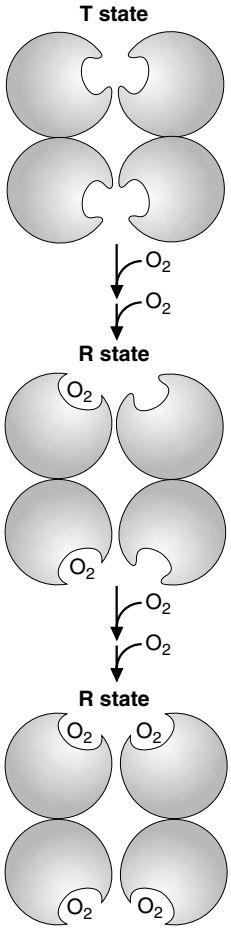


Fig. 7.15. Equilibria for binding of O_2 molecules to hemoglobin according to the concerted model of Monod, Wyman, and Changeux. Hemoglobin exists in two alternate conformations, the T (tense) state with a low affinity for O_2 and the R (relaxed) state with a higher affinity. In the T subunits, the binding sites are hindered, and in the R state the binding sites are open. Each successive addition of O_2 shifts the equilibrium further toward the R state. Because the conformation of all the subunits can change when O_2 binds to one subunit, oxygen binding is said to follow the concerted model. Most of the molecules change to the R state when 2 O_2 molecules have bound.

Amy Lloyd's serum protein electrophoresis indicated the presence of a sharp narrow peak or homogeneous "spike" in the characteristic γ-globulin zone known as an M protein (monoclonal protein) component. A narrow peak of spike in electrophoresis, which separates proteins according to charge distribution of the side chains, suggests an elevation of proteins with a similar or identical structure. Subsequently, it was shown that Amy Lloyd's immunoglobulin M-component was composed of a single homogeneous type of immunoglobulin (just one amino acid sequence in the N-terminal variable region). Thus, the M protein was produced by a single clone of antibody-secreting cells (cells that all arose from proliferation of one cell) in the bone marrow (called plasma cell dyscrasia). In amyloidosis, substitutions of particular amino acids in the light chain variable region are thought to destabilize its native conformation, resulting in fibrillogenesis.

In AL amyloidosis, amyloid is formed from degradation products of the λ or κ light chains that deposit most frequently in the extracellular matrix of the kidney and the heart but also may deposit in the tongue. In other types of amyloidosis, the amyloid arises from other proteins and deposits in a characteristic organ. For example, the amyloid associated with chronic inflammatory conditions, such as tuberculosis or rheumatoid arthritis, is derived from an acute phase serum protein called serum amyloid A that is produced by the liver in response to inflammation. It deposits most frequently in the kidney, and cardiac involvement is rare.

process and, consequently, the binding rate for the first oxygen is very low. When the next oxygen binds, many of the hemoglobin molecules containing one O_2 will already have all four subunits in the R state, and therefore the rate of binding is much higher. With two O_2 molecules bound, an even higher percentage of the hemoglobin molecules will have all four subunits in the R state. This phenomenon, known as positive cooperativity, is responsible for the sigmoidal oxygen saturation curve of hemoglobin (see Fig. 7.12).

VIII. STRUCTURE–FUNCTION RELATIONSHIPS IN IMMUNOGLOBULINS

The immunoglobulins (or antibodies) are one line of defense against invasion of the body by foreign organisms. In this capacity, they function by binding to ligands called antigens on the invading organisms, thereby initiating the process by which these organisms are inactivated or destroyed.

Immunoglobulins all have a similar structure; each antibody molecule contains two identical small polypeptide chains (the light or L chains) and two identical large polypeptide chains (the heavy or H chains) (Fig. 7.16). The chains are joined to each other by disulfide bonds.

The body has five major classes of immunoglobulins. The most abundant immunoglobulins in human blood are the γ-globulins, which belong to the IgG class. The γ-globulins have approximately 220 amino acids in their light chains and 440 in their heavy chains. Like most serum proteins, they have attached oligosaccharides that participate in targeting the protein for clearance from the blood. Both the light and heavy chains consist of domains known as the immunoglobulin fold, which is a collapsed β-barrel made from a number of β-sheets (Fig. 7.17).

Both the light and heavy chains contain regions termed variable (V) and constant (C) regions. The variable regions of the L and H chains (V_L and V_H, respectively) interact to produce a single antigen binding site at each branch of the Y-shaped molecule. Each population (clone) of B cells produces an antibody with a different amino acid composition in the variable region that is complementary to the structure of the antigen eliciting the response. The K_d of antibodies for their specific antigens is extremely small and varies from approximately 10^{-7} to 10^{-11} M. The antigen thus binds very tightly with almost no tendency to dissociate and can be removed from circulation as the antigen–antibody complex is ingested by macrophages. The constant domains that form the Fc part of the antibody are important for binding of the antigen–antibody complex to phagocytic cells for clearance and for other aspects of the immune response.

IX. PROTEIN FOLDING

Although the peptide bonds in a protein are rigid, flexibility around the other bonds in the peptide backbone allow an enormous number of possible conformations for each protein. However, every molecule of the same protein folds into the same stable three-dimensional structure. This shape is known as the native conformation.

The tight binding affinity of immunoglobulins for their specific antigen makes them useful for the measurement of small amounts of other compounds in various radioimmunoassays. The principle of the radioimmunoassay is that the immunoglobulin will specifically bind the compound being measured, and an additional component of the system that is labeled with a radioactive chemical will bind the immunoglobulin. The complex is then separated from the solution and the bound radioactivity measured. The different isozymes of CK and Tn-T used to track **Ann Jeina**'s MI are measured with a type of radioimmunassay by using antibodies specific for each isozyme. This method is much faster than the laborious separation of isozymes by electrophoresis. Radioimmunoassays are also useful for measuring the small amounts of hormones present in the blood for diagnosis of endocrine diseases.

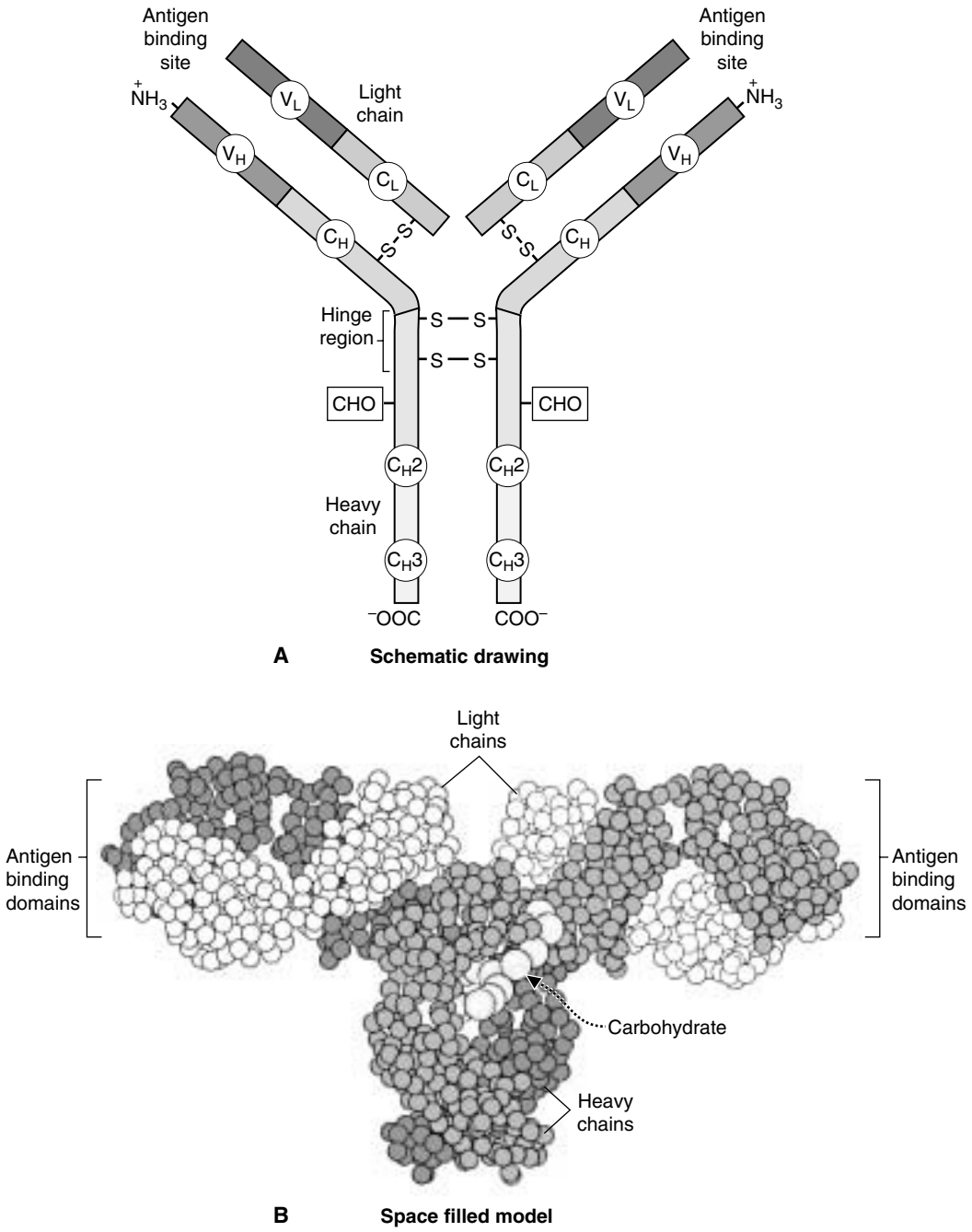


Fig. 7.16. Structure of immunoglobulins. **A.** Each IgG molecule contains two light (L) and two heavy (H) chains joined by disulfide bonds. Each light chain contains two domains, a variable domain (V_L) and a region of constant amino acid sequence (C_L). Each heavy chain has four domains: one variable domain (V_H) and three constant domains (C_H). The conformation of the constant domain contains the β-sheets that are called the immunoglobulin fold. The variable domains are specific for the antigen that is bound, whereas the constant regions are the same for all antibody molecules of a given class. Carbohydrate (CHO) is bound as indicated within the constant region of the heavy chains (CH). The hinge region allows flexibility when the molecule binds antigen. **B.** In the space-filled model, the light chains are light in color and the heavy chains are two different shades of gray. (Modified from Silverton EW, Navia MA, Davies DR, et al. Proc Natl Acad Sci, Three-dimensional structure of an intact human immunoglobulin USA 1977;11:5142)

Very little difference is seen in the energy state of the native conformation and a number of other stable conformations that a protein might assume. It appears that the prion protein, the cause of mad cow disease, is a normal cellular protein that has refolded into a different stable conformation with a lower energy state than its normal functional conformation (discussed under Biochemical Comments, later). If misfolded proteins do not precipitate into aggregates, they can be degraded in the cell by proteolytic reactions, or even refolded.

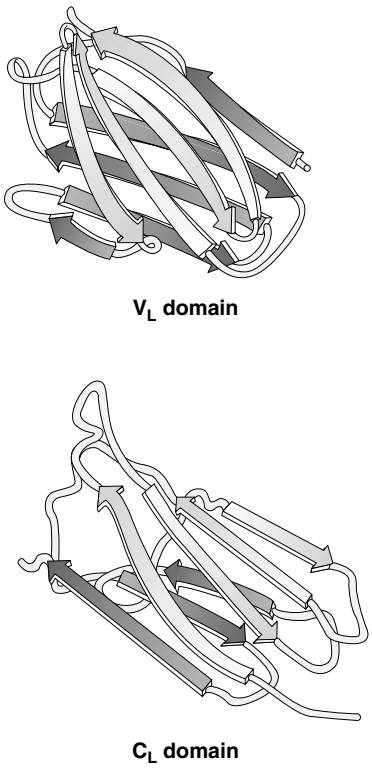


Fig. 7.17. Structure of the V_L and C_L domains of IgG. Layers of antiparallel β-sheets are stacked in these domains, which have been referred to as collapsed β-barrels. The antigen binds between the V_H and V_L immunoglobulin folds, and NOT in the barrel. The C_L domain is also called the immunoglobulin fold. (Top modified from Richardson JS. Adv Protein Chem. The anatomy and taxonomy of protein structure 1981;34:167; bottom reprinted in part with permission from Edmundson AB, et al. Biochemistry 1975;14:3954. © 1975 American Chemical Society.)

A. Primary Structure Determines Folding

The primary structure of a protein determines its three-dimensional conformation. More specifically, the sequence of amino acid side chains dictates the fold pattern of the three-dimensional structure and the assembly of subunits into quaternary structure. Under certain conditions, denatured proteins can refold into their native conformation, regaining their original function. Proteins can be denatured with organic solvents such as urea that disrupt hydrophobic interactions and convert the protein to a soluble random coil. Many simple single-subunit proteins like ribonuclease that are denatured in this way spontaneously refold into their native conformation if carefully brought back to physiologic conditions. Even complex multi-subunit proteins containing bound cofactors can sometimes spontaneously renature under the right conditions. Thus, the primary structure essentially specifies the folding pattern.

In the cell, not all proteins fold into their native conformation on their own. As the protein folds and refolds while it is searching for its native low energy state, it passes through many high-energy conformations that slow the process (called kinetic barriers). These kinetic barriers can be overcome by heat shock proteins, which use energy provided by ATP hydrolysis to assist in the folding process (Fig. 7.18). Heat shock proteins were named for the fact that their synthesis in bacteria increased when the temperature was suddenly raised. They are

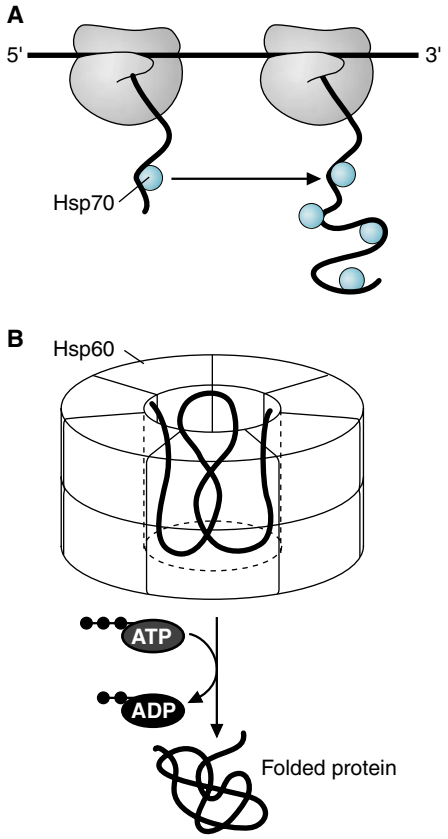


Fig. 7.18. Role of heat shock proteins in folding. A. The Hsp70 family of proteins prevent folding of the nascent chain and promote unfolding. The ATPase domain of the protein has the actin fold. B. The Hsp60 class of protein has a barrel shape into which the protein fits. It acts as a template, binding and rebinding portions of the unfolded protein until folding is completed. It hydrolyzes many ATP bonds to provide energy for the process.

present in human cells as different families of proteins with different activities. For example, the hsp70 proteins bind to nascent polypeptide chains as their synthesis is being completed to keep the uncompleted chains from folding prematurely. They also unfold proteins prior to their insertion through the membrane of mitochondria and other organelles. The multi-subunit barrel-shaped hsp60 family of proteins is called chaperonins. The unfolded protein fits into the barrel cavity that excludes water and serves as a template for the folding process. The hydrolysis of several ATP molecules is used to overcome the energy barriers to reaching the native conformation.

A *cis-trans* isomerase and a disulfide isomerase also participate in folding. The *cis-trans* isomerase converts a trans peptide bond preceding a proline into the *cis* conformation, which is well suited for making hairpin turns. The disulfide isomerase breaks and reforms disulfide bonds between the -SH groups of two cysteine residues in transient structures formed during the folding process. After the protein has folded, cysteine-SH groups in close contact in the tertiary structure can react to form the final disulfide bonds.

B. Protein Denaturation

1. DENATURATION THROUGH NONENZYMATIC MODIFICATION OF PROTEINS

Amino acids on proteins can undergo a wide range of chemical modifications that are not catalyzed by enzymes, such as nonenzymatic glycosylation or oxidation. Such modifications usually lead to a loss of function and denaturation of the protein, sometimes to a form that cannot be degraded in the cell. In nonenzymatic glycosylation, glucose that is present in blood, or in interstitial or intracellular fluid, binds to an exposed amino group on a protein (Fig. 7.19). The two-step process forms an irreversibly glycosylated protein. Proteins that turn over very slowly in the body, such as collagen or hemoglobin, exist with a significant fraction present in the glycosylated form. Because the reaction is nonenzymatic, the rate of glycosylation is proportionate to the concentration of glucose present, and individuals with hyperglycemia have much higher levels of glycosylated proteins than individuals with normal blood glucose levels. Collagen and other glycosylated proteins in tissues are further modified by nonenzymatic oxidation and form additional cross-links. The net result is the formation of large protein aggregates referred to as AGEs (advanced glycosylation end-products). AGE is a meaningful acronym because AGEs accumulate with age, even in individuals with normal blood glucose levels.

Di Abietes' physician used her glycosylated hemoglobin levels, specifically the HbA_{1c} fraction, to determine whether she had sustained hyperglycemia over a long period. The rate of irreversible nonenzymatic glycosylation of hemoglobin and other proteins is directly proportional to the glucose concentration to which they are exposed over the last 4 months. The danger of sustained hyperglycemia is that, over time, many proteins become glycosylated and subsequently oxidized, affecting their solubility and ability to function. The glycosylation of collagen in the heart, for example, is believed to result in a cardiomyopathy in patients with chronic uncontrolled diabetes mellitus. In contrast, glycosylation of hemoglobin has little effect on its function.

2. PROTEIN DENATURATION BY TEMPERATURE, pH, AND SOLVENT

Proteins can be denatured by changes of pH, temperature, or solvent that disrupt ionic, hydrogen, and hydrophobic bonds. At a low pH, ionic bonds and hydrogen bonds formed by carboxylate groups would be disrupted; at a very alkaline pH, hydrogen and ionic bonds formed by the basic amino acids would be disrupted. Thus, the pH of the body must be maintained within a range compatible with three-dimensional structure. Temperature increases vibrational and rotational energies in the bonds, thereby affecting the energy balance that goes into making a stable three-dimensional conformation.

Hydrophobic molecules can also denature proteins by disturbing hydrophobic interactions in the protein. For example, long-chain fatty acids can inhibit many enzyme-catalyzed reactions by binding nonspecifically to hydrophobic pockets in proteins and disrupting hydrophobic interactions. Thus, long-chain fatty acids and other highly hydrophobic molecules have their own binding proteins in the cell.

Proteins are denatured in the gastric juice of the stomach, which has a pH of 1 to 2. Although this pH cannot break peptide bonds, disruption of the native conformation makes the protein a better substrate for digestive enzymes.

Thermal denaturation is often illustrated by the process of cooking an egg. In the presence of heat, the protein albumin converts from its native translucent state to a denatured white precipitate. Protein precipitates can sometimes be dissolved by amphipathic agents such as urea, guanidine HCl, or SDS (sodium dodecylsulfate) that form extensive hydrogen bonds and hydrophobic interactions with the protein.

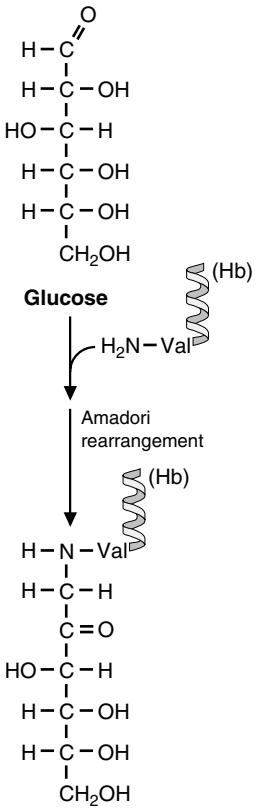


Fig. 7.19. Nonenzymatic glycosylation of hemoglobin. Glucose forms a Schiff base with the N-terminal amino group of the protein, which rearranges to form a stable glycosylated product. Similar nonenzymatc glycosylation reactions occur on other proteins.

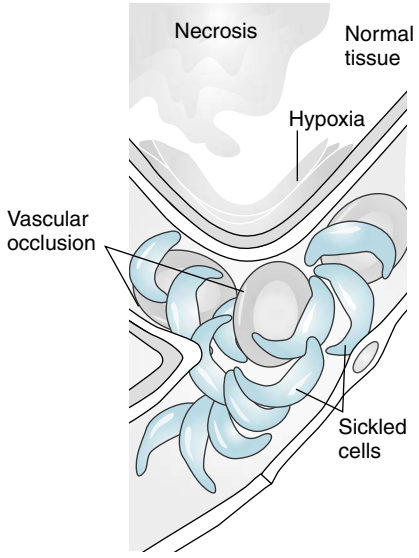


Fig. 7.20 Sickled red blood cells occlude a blood vessel, causing hypoxia (low O_2 in cells) and necrosis (cell death).

CLINICAL COMMENTS

Will Sichel. Will Sichel continues to experience severe low back and lower extremity pain for many hours after admission. The diffuse pains of sickle cell crises are believed to result from occlusion of small vessels in a variety of tissues, thereby depriving cells of oxygen and causing ischemic or anoxic damage to the tissues. In a sickle cell crisis, long hemoglobin polymers form, causing the red blood cells to become distorted and change from a biconcave disc to an irregular shape, such as a sickle (for which the disease was named) or a stellate structure (Fig.7.20). The aggregating Hb polymers damage the cellular membrane and promote aggregation of membrane proteins, increased permeability of the cell, and dehydration. Surface charge and antigens of red blood cells are carried on the transmembrane proteins glycophorin and Band 3 (the erythrocyte anion exchange channel, see Chapter 10). Hemoglobin S binds tightly to the cytoplasmic portion of band 3, contributing to further polymer aggregation and uneven distribution of negative charge on the sickle cell surface. As a result, the affected cells adhere to endothelial cells in capillaries, occluding the vessel and decreasing blood flow to the distal tissues. The subsequent hypoxia in these tissues causes cellular damage and even death.

The sickled cells are sequestered and destroyed mainly by phagocytic cells, particularly those in the spleen. An anemia results as the number of circulating red blood cells decreases and bilirubin levels rise in the blood as hemoglobin is degraded.

After a few days of treatment, **Will Sichel's** crisis was resolved. In the future, should Will suffer a cerebrovascular accident as a consequence of vascular occlusion or have recurrent life-threatening episodes, a course of long-term maintenance blood transfusions to prevent repeated sickle crises may be indicated. Iron chelation would have to accompany such a program to prevent or delay the development of iron overload. Although a few individuals with this disease have survived into the sixth decade, mean survival is probably into the fourth decade. Death usually results from renal failure or cardiopulmonary disease.

Anne Jeina. Mrs. Jeina's diagnosis of an acute myocardial infarction (MI) was based partly on measurements of CK-MB, myoglobin, and cTN-T (the cardiac isozyme of troponin-T, a subunit of the regulatory protein troponin). Early diagnosis is critical for a decision on the type of therapeutic intervention to be used. Of these proteins, myoglobin appears in the blood most rapidly. However, its levels are relatively nonspecific for cardiac injury because the amino acid sequences of cardiac and skeletal muscle myoglobins are identical. Myoglobin measurements do have a very high negative predictive value within the 2- to 6-hour period after the onset of symptoms (i.e., if myoglobin is not elevated, a myocardial infarction did not occur). In contrast, serum cardiac troponin-T is a relatively late, but highly specific, marker of myocardial injury. It is typically detected in an acute MI within 3 to 5 hours after onset of symptoms, is positive in most cases within 8 hours, and approaches 100% sensitivity at 10 to 12 hours. It remains elevated for 5 to 10 days.

Mrs. Jeina stayed in the hospital until she had been free of chest pain for 5 days. She was discharged on a low-fat diet and was asked to participate in the hospital patient exercise program for patients recovering from a recent heart attack. She was scheduled for regular examinations by her physician.

Troponin is a heterotrimeric protein involved in the regulation of striated and cardiac muscle contraction. Most troponin in the cell is bound to the actin–tropomyosin complex in the muscle fibril. The three subunits of troponin consist of troponin-C, troponin-T, and troponin-I, each with a specific function in the regulatory process. Troponin-T and troponin-I exist as different isoforms in cardiac and skeletal muscle (sequences with a different amino acid composition), thus allowing the development of specific antibodies against each form. As a consequence, either cardiac troponin-T or cardiac troponin-I may be rapidly measured in blood samples by immunoassay with a good degree of specificity.

Amy Lloyd. Amy Lloyd has AL amyloidosis, which is characterized by deposition of amyloid fibers derived principally from the variable region of λ or κ immunoglobulin light chains. Increased amounts of the fragments of the light chains called Bence-Jones proteins appeared in her urine. Fibril deposition in the extracellular matrix of her kidney glomeruli has resulted in mild renal failure. Deposition of amyloid in the extracellular matrix of her heart muscle resulted in the cardiac arrhythmia seen on an electrocardiogram. In addition to other signs of right-sided heart failure, she had peripheral edema. The loss of weight may have been caused by infiltrations of amyloid in the gastrointestinal tract or by constipation and diarrhea resulting from involvement of the autonomic nervous system. Treatment may be directed against the plasma cell proliferation, or against the symptomatic results of organ dysfunction.

During Amy Lloyd's evaluation, she developed a cardiac arrhythmia that was refractory to treatment. The extensive amyloid deposits in her heart had disrupted conduction of electrical impulses in the heart muscle, ultimately resulting in cardiac arrest. On autopsy, amyloid deposits were found within the heart, tongue, liver, adipose tissue, and every organ examined except the central nervous system, which had been protected by the blood-brain barrier.

Dianne Abietes. Di Abietes' HbA_{1c} of 8.5% was just above the normal range (5.8–7.2% of total hemoglobin), and her physician decided not to alter her insulin treatment plan. Glycosylation is a nonenzymatic reaction that occurs with a rate directly proportionate to the concentration of glucose in the blood. In the normal range of blood glucose concentrations (approximately 80–140 mg/dL, depending on time after a meal), 6 to 7% of the hemoglobin is glycosylated to form $HbA_{1c.}$ Hemoglobin turns over in the blood as red blood cells are phagocytosed and their hemoglobin degraded and new red blood cells are derived from retic-

Four minor components of adult hemoglobin (HbA) result from posttranslational, nonenzymatic glycosylation of different amino acid residues (HbA_{1a1}, HbA_{1a2}, HbA_{1b1}, and HbA_{1c}). In HbA_{1c}, the fraction that is usually measured, the glycosylation occurs on an N-terminal valine.

Prion diseases are categorized as transmissable spongiform encephalopathies, which are neurodegenerative diseases characterized by spongiform degeneration and astrocytic gliosis in the central nervous system. Frequently, protein aggregates and amyloid plaques are seen. These aggregates are resistant to proteolytic degradation.

Familial prion diseases are caused by point mutations in the gene encoding the Pr protein (point mutations are changes in one base in the DNA nucleotide sequence). The diseases have a variety of names related to the different mutations and the clinical syndrome (e.g., Gertsmann-Straussler-Scheinker disease and familial Creutzfeld-Jakob disease). Familial Creutzfeld-Jakob disease (fCJD) arises from an inherited mutation and has an autosomal dominant pedigree. It typically presents in the fourth decade of life. The mutation lowers the activation energy for refolding and the prion proteins fold into the PrP^{Sc} conformation more readily. It is estimated that the rate of generating prion disease by refolding of PrP^{c} in the normal cell is approximately 3,000 to 4,000 years. Lowering of the activation energy for refolding by mutation presumably decreases this time to the observed 30- to 40-year prodromal period. Sporadic Creutzfeld-Jakob disease may arise from somatic cell mutation or rare spontaneous refoldings that initiate a cascade of refolding into the PrP^{Sc} conformation. The sporadic form of the disease accounts for 85% of all cases of CJD.

ulocytes. The average lifespan of a red blood cell is 120 days. Thus, the extent of hemoglobin glycosylation is a direct reflection of the average serum glucose concentration to which the cell has been exposed over its 120-day lifespan. **Di Abietes'** elevated Hb_{A1c} indicates that her blood glucose levels have been elevated on an average over the preceding 6 weeks to 4 months. An increase of Di's insulin dosage would decrease her hyperglycemia but increase her risk of hypoglycemic events.

BIOCHEMICAL COMMENTS

Protein Misfolding and Prions. Prion proteins are believed to cause a neurodegenerative disease by acting as a template to misfold other cellular prion proteins into a form that cannot be degraded. The word *prion* stands for proteinaceous infectious agent. The prion diseases can be acquired either through infection (mad-cow disease), or from sporadic or inherited mutations (e.g., Creutzfeldt-Jakob disease). Although the infectious prion diseases represent a small proportion of human cases, their link to mad cow disease in the United Kingdom (new variant CJD), to growth hormone inoculations in the United States and France (iatrogenic or "doctor-induced" CJD), and to ritualistic cannibalism in the Fore tribespeople (Kuru) have received the most publicity.

The prion protein is normally found in the brain and is encoded by a gene that is a normal component of the human genome. The disease-causing form of the prion protein has the same amino acid composition but is folded into a different conformation that aggregates into multimeric protein complexes resistant to proteolytic degradation (Fig. 7.21). The normal conformation of the prion protein has been designated PrP^{c} and the disease-causing form PrP^{Sc} (sc for the prion disease scrapies in sheep). Although PrP^{Sc} and PrP^{c} have the same amino acid composition, the PrP^{Sc} conformer is substantially enriched in β-sheet structure compared with the normal PrP^{c} conformer, which has little or no β-sheet structure and is approximately 40% α-helix. This difference favors the aggregation of PrP^{Sc} into multimeric complexes. These two conformations presumably have similar energy levels. Fortunately, spontaneous refolding of PrP proteins into the PrP^{Sc} conformation is prevented by a large activation energy barrier that makes this conversion extremely slow. Thus, very few molecules of PrP^{Sc} are normally formed during a lifetime.

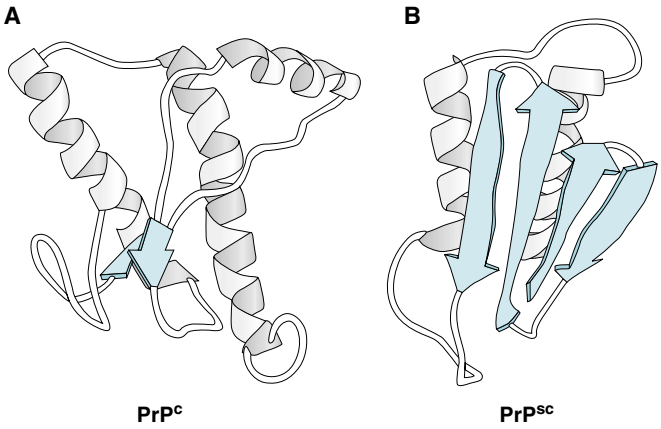


Fig. 7.21. The conformation of PrP^{c} and PrP^{sc}. The prion proteins have two domains, an N-terminal region that binds four Cu^{2+} per chain, and a C-terminal region. In PrP^{c}, the C-terminal regions contain three substantial helices and two 3-residue β strands joined by two to three hydrogen bonds (approximately 40% a-helix and almost no β-sheet structure). It exists as a monomer. In PrP^{sc}, the C-terminal region is folded into an extensive β-sheet. The overall structure is approximately 40 to 50% β-sheet, 20 to 30% α-helices. This conformation promotes aggregation.

The infectious disease occurs with the ingestion of PrP^{Sc} dimers in which the prion protein is already folded into the high β structure. These PrP^{Sc} proteins are thought to act as a template to lower the activation energy barrier for the conformational change, causing native proteins to refold into the PrP^{Sc} conformation much more rapidly (much like the role of chaperonins). The refolding initiates a cascade as each new PrP^{Sc} formed acts as a template for the refolding of other molecules. As the number of PrP^{Sc} molecules increase in the cell, they aggregate into a multimeric assembly that is resistant to proteolytic digestion.

Suggested References

Khan MF, Falk RH. Amyloidosis. Postgrad Med J 2001;77:686–693.

Horwich AL, Weissman JS. Deadly conformations-protein misfolding in prion diseases. Cell 1997;89:499–510.

Raj DSC, Choudhury D, Welbourne TC, Levi M. Advanced glycation end products: a nephrologist's perspective. Am J Kidney Dis 2000;35:365–380.

REVIEW QUESTIONS—CHAPTER 7

1. Which of the following characterize α-helix regions of proteins?

   (A) They all have the same primary structure.
   (B) They are formed principally by hydrogen bonds between a carbonyl oxygen atom in one peptide bond and the amide hydrogen from a different peptide bond.
   (C) They are formed principally by hydrogen bonds between a carbonyl atom in one peptide bond and the hydrogen atoms on the side chain of another amino acid.
   (D) They are formed by hydrogen bonding between two adjacent amino acids in the primary sequence.
   (E) They require a high content of proline and glycine.

2. Which of the following is a characteristic of globular proteins?

   (A) Hydrophilic amino acids tend to be on the inside.
   (B) Hydrophobic amino acids tend to be on the outside.
   (C) Tertiary structure is formed by hydrophobic and electrostatic interactions between amino acids, and by hydrogen bonds between amino acids and between amino acids and water.
   (D) Secondary structures are formed principally by hydrophobic interactions between amino acids.
   (E) Covalent disulfide bonds are necessary to hold the protein in a rigid conformation.

3. A protein has one transmembrane domain composed entirely of α-helical secondary structure. Which of the following amino acids would you expect to find in the transmembrane domain?

   (A) Proline
   (B) Glutamate
   (C) Lysine
   (D) Leucine
   (E) Arginine

4. Autopsies of patients with Alzheimer's disease show protein aggregates called neurofibrillary tangles and neuritic plaques in various regions of the brain. These plaques exhibit the characteristic staining of amyloid. Which of the following structural features is the most likely characteristic of at least one protein in these plaques?

   (A) A high content of β-pleated sheet structure
   (B) A high content of α-helical structure
   (C) A high content of random coils
   (D) Disulfide bond crosslinks between polypeptide chains
   (E) A low-energy native conformation

5. While studying a novel pathway in a remote species of bacteria, you discover a new globular protein that phosphorylates a substrate, using ATP as the phosphate donor. This protein most likely contains which of the following structures?

   (A) An actin fold
   (B) An immunoglobulin fold
   (C) A nucleotide binding fold
   (D) A globin fold
   (E) A β-barrel