15. Sequence Determination of the Brain Peptide Leucine Enkephalin A group of peptides that influence nerve transmission in certain parts of the brain has been isolated from normal brain tissue. These peptides are known as opioids, because they bind to specific receptors that also bind opiate drugs, such as morphine and naloxone. Opioids thus mimic some of the properties of opiates. Some researchers consider these peptides to be the brain's own pain killers. Using the information below, determine the amino acid sequence of the opioid leucine enkephalin. Explain how your structure is consistent with each piece of information.

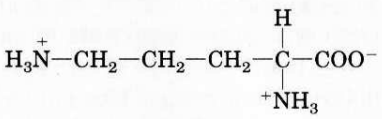
(a) Complete hydrolysis by 6 M HCl at 110 °C followed by amino acid analysis indicated the presence of Gly, Leu, Phe, and Tyr, in a 2:1:1:1 molar ratio.

(b) Treatment of the peptide with 1-fluoro-2,4-dinitrobenzene followed by complete hydrolysis and chromatography indicated the presence of the 2,4-dinitrophenyl derivative of tyrosine. No free tyrosine could be found.

(c) Complete digestion of the peptide with pepsin followed by chromatography yielded a dipeptide containing Phe and Leu, plus a tripeptide containing Tyr and Gly in a 1:2 ratio.

16. Structure of a Peptide Antibiotic from *Bacillus brevis* Extracts from the bacterium *Bacillus brevis* contain a peptide with antibiotic properties. This peptide forms complexes with metal ions and apparently disrupts ion transport across the cell membranes of other bacterial species, killing them. The structure of the peptide has been determined from the following observations.

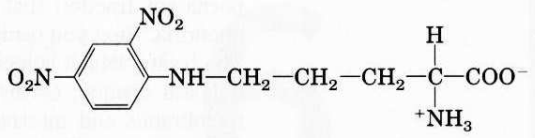
(a) Complete acid hydrolysis of the peptide followed by amino acid analysis yielded equimolar amounts of Leu, Orn, Phe, Pro, and Val. Orn is ornithine, an amino acid not present in proteins but present in some peptides. It has the structure



(b) The molecular weight of the peptide was estimated as about 1,200.

(c) The peptide failed to undergo hydrolysis when treated with the enzyme carboxypeptidase. This enzyme catalyzes the hydrolysis of the carboxyl-terminal residue of a polypeptide unless the residue is Pro or does not contain a free carboxyl group for some reason.

(d) Treatment of the intact peptide with 1-fluoro-2,4-dinitrobenzene, followed by complete hydrolysis and chromatography, yielded only free amino acids and the following derivative:



(Hint: Note that the 2,4-dinitrophenyl derivative involves the amino group of a side chain rather than the α-amino group.)

(e) Partial hydrolysis of the peptide followed by chromatographic separation and sequence analysis yielded the following di- and tripeptides (the amino-terminal amino acid is always at the left):

Leu–Phe Phe–Pro Orn–Leu Val–Orn
Val–Orn–Leu Phe–Pro–Val Pro–Val–Orn

Given the above information, deduce the amino acid sequence of the peptide antibiotic. Show your reasoning. When you have arrived at a structure, demonstrate that it is consistent with *each* experimental observation.

17. Efficiency in Peptide Sequencing A peptide with the primary structure Lys–Arg–Pro–Leu–Ile–Asp–Gly–Ala is sequenced by the Edman procedure. If each Edman cycle were 96% efficient, what percentage of the amino acids liberated in the fourth cycle would be leucine? Do the calculation a second time, but assume a 99% efficiency for each cycle.

18. Biochemistry Protocols: Your First Protein Purification As the newest and least experienced student in a biochemistry research lab, your first few weeks are spent washing glassware and labeling test tubes. You then graduate to making buffers and stock solutions for use in various laboratory procedures. Finally, you are given responsibility for purifying a protein. It is a citric acid cycle enzyme, citrate synthase, located in the mitochondrial matrix. Following a protocol for the purification, you proceed through the steps below. As you work, a more experienced student questions you about the rationale for each procedure. Supply the answers. (Hint: See Chapter 2 for information on separation of organelles from cells, and Chapter 4 for information about osmolarity).

(a) You pick up 20 kg of beef hearts from a nearby slaughterhouse. You transport the hearts on ice, and perform each step of the purification in a walk-in cold room or on ice. You homogenize the beef heart tissue in a high-speed blender in a medium containing ~0.2 M sucrose, buffered to a pH of 7.2. *Why do you use beef heart tissue, and in such large quantity? What is the purpose of keeping the tissue cold and suspending it in 0.2 M sucrose, at pH 7.2? What happens to the tissue when it is homogenized?*

chapter

7

Protein Function

Knowing the three-dimensional structure of a protein is an important part of understanding how the protein functions. However, the structure shown in two dimensions on a page is deceptively static. Proteins are dynamic molecules whose functions almost invariably depend on interactions with other molecules, and these interactions are affected in physiologically important ways by sometimes subtle, sometimes striking changes in protein conformation.

In this chapter, we explore how proteins interact with other molecules and how their interactions are related to dynamic protein structure. The importance of molecular interactions to a protein's function can hardly be overemphasized. In Chapter 6, we saw that the function of fibrous proteins as structural elements of cells and tissues depends on stable, long-term quaternary interactions between identical polypeptide chains. As we will see in this chapter, the functions of many other proteins involve interactions with a variety of different molecules. Most of these interactions are fleeting, though they may be the basis of complex physiological processes such as oxygen transport, immune function, and muscle contraction, the topics we examine in detail in this chapter. The proteins that carry out these processes illustrate the following key principles of protein function, some of which will be familiar from the previous chapter:

The functions of many proteins involve the reversible binding of other molecules. A molecule bound reversibly by a protein is called a **ligand.** A ligand may be any kind of molecule, including another protein. The transient nature of protein-ligand interactions is critical to life, allowing an organism to respond rapidly and reversibly to changing environmental and metabolic circumstances.

A ligand binds at a site on the protein called the **binding site,** which is complementary to the ligand in size, shape, charge, and hydrophobic or hydrophilic character. Furthermore, the interaction is specific: the protein can discriminate among the thousands of different molecules in its environment and selectively bind only one or a few. A given protein may have separate binding sites for several different ligands. These specific molecular interactions are crucial in maintaining the high degree of order in a living system. (This discussion excludes the binding of water, which may interact weakly and nonspecifically with many parts of a protein. In Chapter 8, we consider water as a specific ligand for many enzymes.)

Proteins are flexible. Changes in conformation may be subtle, reflecting molecular vibrations and small movements of amino acid residues

throughout the protein. A protein flexing in this way is sometimes said to "breathe." Changes in conformation may also be quite dramatic, with major segments of the protein structure moving as much as several nanometers. Specific conformational changes are frequently essential to a protein's function.

The binding of a protein and ligand is often coupled to a conformational change in the protein that makes the binding site more complementary to the ligand, permitting tighter binding. The structural adaptation that occurs between protein and ligand is called **induced fit.**

In a multisubunit protein, a conformational change in one subunit often affects the conformation of other subunits.

Interactions between ligands and proteins may be regulated, usually through specific interactions with one or more additional ligands. These other ligands may cause conformational changes in the protein that affect the binding of the first ligand.

Enzymes represent a special case of protein function. Enzymes bind and chemically transform other molecules—they catalyze reactions. The molecules acted upon by enzymes are called reaction **substrates** rather than ligands, and the ligand-binding site is called the **catalytic site** or **active site.** In this chapter we emphasize the noncatalytic functions of proteins. In Chapter 8 we consider catalysis by enzymes, a central topic in biochemistry. You will see that the themes of this chapter—binding, specificity, and conformational change—are continued in the next chapter, with the added element of proteins acting as reactants in chemical transformations.

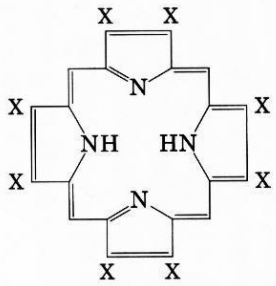
Reversible Binding of a Protein to a Ligand: Oxygen-Binding Proteins

Myoglobin and hemoglobin may be the most-studied and best-understood proteins. They were the first proteins for which three-dimensional structures were determined, and our current understanding of myoglobin and hemoglobin is garnered from the work of thousands of biochemists over several decades. Most important, they illustrate almost every aspect of that most central of biochemical processes: the reversible binding of a ligand to a protein. This classic model of protein function will tell us a great deal about how proteins work.

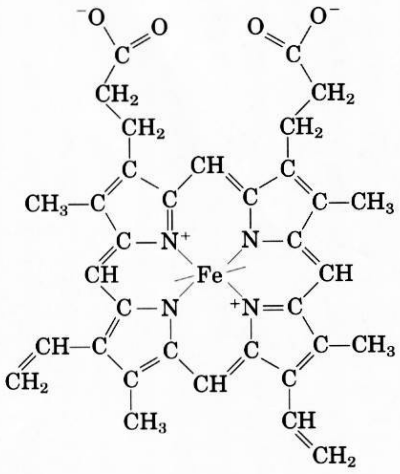
Oxygen Can Be Bound to a Heme Prosthetic Group

Oxygen is poorly soluble in aqueous solutions (see Table 4–3) and cannot be carried to tissues in sufficient quantity if it is simply dissolved in blood serum. Diffusion of oxygen through tissues is also ineffective over distances greater than a few millimeters. The evolution of larger, multicellular animals depended on the evolution of proteins that could transport and store oxygen. However, none of the amino acid side chains in proteins is suited for the reversible binding of oxygen molecules. This role is filled by certain transition metals, among them iron and copper, that have a strong tendency to bind oxygen. Multicellular organisms exploit the properties of metals, most commonly iron, for oxygen transport. However, free iron promotes the formation of highly reactive oxygen species such as hydroxyl radicals that can damage DNA and other macromolecules. Iron used in cells is therefore bound in forms that sequester it and/or make it less reactive. In multicellular organisms—especially those in which iron, in its oxygen-carrying capacity, must be transported over large distances—iron is often incorporated into a protein-bound prosthetic group called **heme.** (A prosthetic group is a compound permanently associated with a protein that contributes to the protein's function.)

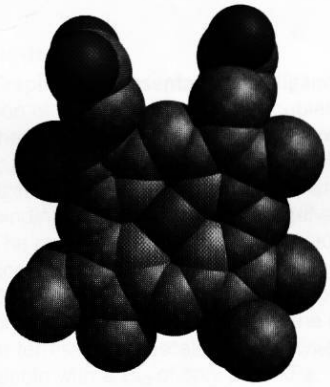
figure 7–1
Heme. The heme group is present in myoglobin, hemoglobin, and many other proteins, designated heme proteins. Heme consists of a complex organic ring structure, protoporphyrin IX, to which is bound an iron atom in its ferrous (Fe^{2+}) state. Porphyrins, of which protoporphyrin IX is only one example, consist of four pyrrole rings linked by methene bridges **(a),** with substitutions at one or more of the positions denoted X. Two representations of heme are shown in **(b)** and **(c).** The iron atom of heme has six coordination bonds: four in the plane of, and bonded to, the flat porphyrin ring system, and two perpendicular to it **(d).**

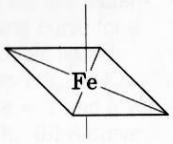


(a)



(b)


(c)



(d)

Heme (or haem) consists of a complex organic ring structure, **protoporphyrin,** to which is bound a single iron atom in its ferrous (Fe^{2+}) state (Fig. 7–1). The iron atom has six coordination bonds, four to nitrogen atoms that are part of the flat **porphyrin ring** system and two perpendicular to the porphyrin. The coordinated nitrogen atoms (which have an electron-donating character) help prevent conversion of the heme iron to the ferric (Fe^{3+}) state. Iron in the Fe^{2+} state binds oxygen reversibly; in the Fe^{3+} state it does not bind oxygen. Heme is found in a number of oxygen-transporting proteins, as well as in some proteins, such as the cytochromes, that participate in oxidation-reduction (electron transfer) reactions (Chapter 19).

In free heme molecules, reaction of oxygen at one of the two "open" coordination bonds of iron (perpendicular to the plane of the porphyrin molecule, above and below) can result in irreversible conversion of Fe^{2+} to Fe^{3+}. In heme-containing proteins, this reaction is prevented by sequestering the heme deep within a protein structure where access to the two open coordination bonds is restricted. One of these two coordination bonds is occupied by a side-chain nitrogen of a His residue. The other is the binding site for molecular oxygen (O_2) (Fig. 7–2). When oxygen binds, the electronic properties of heme iron change; this accounts for the change in color from the dark purple of oxygen-depleted venous blood to the bright red of oxygen-rich arterial blood. Some small molecules, such as carbon monoxide (CO) and nitric oxide (NO), coordinate to heme iron with greater affinity than does O_2. When a molecule of CO is bound to heme, O_2 is excluded, which is why CO is highly toxic to aerobic organisms. By surrounding and sequestering heme, oxygen-binding proteins regulate the access of CO and other small molecules to heme iron.

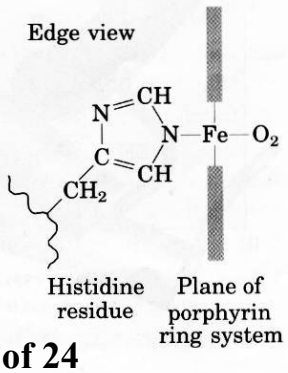


figure 7–2
The heme group viewed from the side. This view shows the two coordination bonds to Fe^{2+} perpendicular to the porphyrin ring system. One of these two bonds is occupied by a His residue, sometimes called the proximal His. The other is the binding site for oxygen. The remaining four coordination bonds are in the plane of, and bonded to, the flat porphyrin ring system.

Myoglobin Has a Single Binding Site for Oxygen

Myoglobin (M_r 16,700; abbreviated Mb) is a relatively simple oxygen-binding protein found in almost all mammals, primarily in muscle tissue. It is particularly abundant in the muscles of diving mammals such as seals and whales that must store enough oxygen for prolonged excursions undersea. Proteins very similar to myoglobin are widely distributed, occurring even in some single-celled organisms. Myoglobin stores oxygen for periods when energy demands are high and facilitates its distribution to oxygen-starved tissues.

Myoglobin is a single polypeptide of 153 amino acid residues with one molecule of heme. It is typical of the family of proteins called **globins,** which have similar primary and tertiary structures. The polypeptide is made up of eight α-helical segments connected by bends (Fig. 7–3). About 78% of the amino acid residues in the protein are found in these α helices.

Any detailed discussion of protein function inevitably involves protein structure. Our treatment of myoglobin will be facilitated by introducing some structural conventions peculiar to globins. As seen in Figure 7–3, the helical segments are labeled A through H. An individual amino acid residue may be designated either by its position in the amino acid sequence or by its location within the sequence of a particular α-helical segment. For example, the His residue coordinated to the heme in myoglobin, His^{93} (the 93rd amino acid residue from the amino-terminal end of the myoglobin polypeptide sequence), is also called His F8 (the 8th residue in α helix F). The bends in the structure are labeled AB, CD, EF, and so forth, reflecting the α-helical segments they connect.

Protein-Ligand Interactions Can Be Described Quantitatively

The function of myoglobin depends on the protein's ability not only to bind oxygen, but also to release it when and where it is needed. Function in biochemistry often revolves around a reversible protein-ligand interaction of this type. A quantitative description of this interaction is therefore a central part of many biochemical investigations.

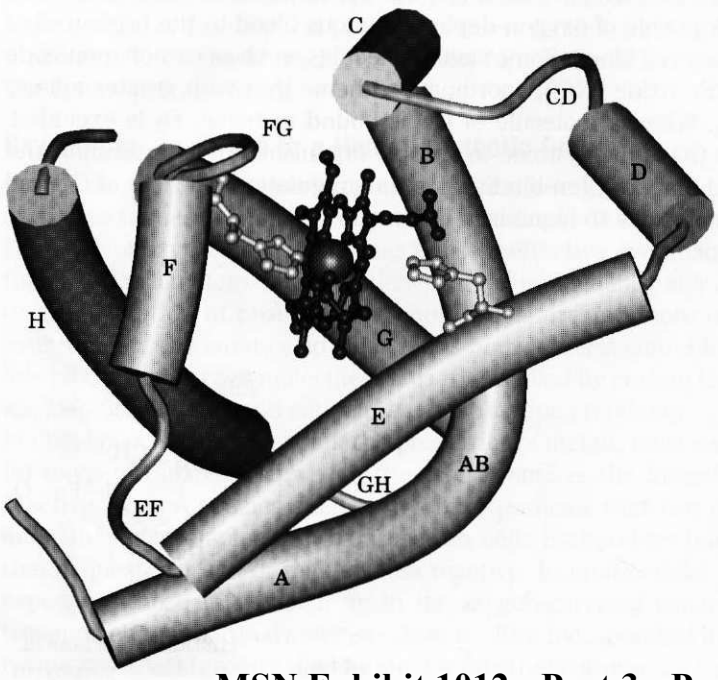


figure 7–3
The structure of myoglobin. The eight α-helical segments (shown here as cylinders) are labeled A through H. Nonhelical residues in the bends that connect them are labeled AB, CD, EF, and so forth, indicating the segments they interconnect. A few bends, including BC and DE, are abrupt and do not contain any residues; these are not normally labeled. (The short segment visible between D and E is an artifact of the computer representation.) The heme is bound in a pocket made up largely of the E and F helices, although amino acid residues from other segments of the protein also participate.

In general, the reversible binding of a protein (P) to a ligand (L) can be described by a simple **equilibrium expression:**

$$P + L \rightleftharpoons PL \tag{7–1}$$

The reaction is characterized by an equilibrium constant, K_a, such that

$$K_a = \frac{[PL]}{[P][L]} \tag{7–2}$$

The term $\boldsymbol{K_a}$ is an **association constant** (not to be confused with the K_a that denotes an acid dissociation constant; see p. 98). The association constant provides a measure of the affinity of the ligand L for the protein. K_a has units of M^{-1}; a higher value of K_a corresponds to a higher affinity of the ligand for the protein. A rearrangement of Equation 7–2 shows that the ratio of bound to free protein is directly proportional to the concentration of free ligand:

$$K_a[L] = \frac{[PL]}{[P]} \tag{7–3}$$

When the concentration of the ligand is much greater than the concentration of ligand-binding sites, the binding of the ligand by the protein does not appreciably change the concentration of free (unbound) ligand—that is, [L] remains constant. This condition is broadly applicable to most ligands that bind to proteins in cells and simplifies our description of the binding equilibrium.

Thus we can consider the binding equilibrium from the standpoint of the fraction, θ (theta), of ligand-binding sites on the protein that are occupied by ligand:

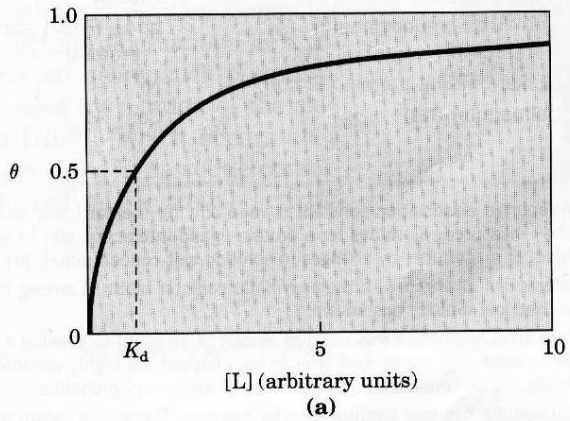
$$\theta = \frac{\text{binding sites occupied}}{\text{total binding sites}} = \frac{[PL]}{[PL] + [P]} \tag{7–4}$$

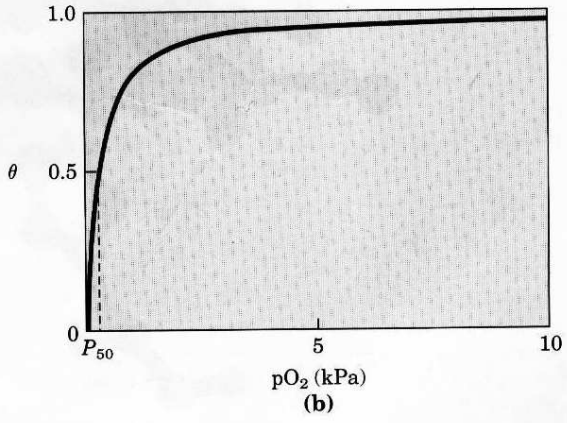
Substituting $K_a[L][P]$ for [PL] (see Eqn 7–3) and rearranging terms gives

$$\theta = \frac{K_a[L][P]}{K_a[L][P] + [P]} = \frac{K_a[L]}{K_a[L] + 1} = \frac{[L]}{[L] + \frac{1}{K_a}} \tag{7–5}$$

The term K_a can be determined from a plot of θ versus the concentration of free ligand, [L] (Fig. 7–4a). Any equation of the form $x = y/(y + z)$ describes a hyperbola, and θ is thus found to be a hyperbolic function of [L]. The fraction of ligand-binding sites occupied approaches saturation asymptotically as [L] increases. The [L] at which half of the available ligand-binding sites are occupied (at $\theta = 0.5$) corresponds to $1/K_a$.

figure 7–4
Graphical representations of ligand binding. The fraction of ligand-binding sites occupied, θ, is plotted against the concentration of free ligand. Both curves are rectangular hyperbolas. **(a)** A hypothetical binding curve for a ligand L. The [L] at which half of the available ligand-binding sites are occupied is equivalent to $1/K_a$, or K_d. The curve has a horizontal asymptote at $\theta = 1$ and a vertical asymptote (not shown) at $[L] = -1/K_a$. **(b)** A curve describing the binding of oxygen to myoglobin. The partial pressure of O_2 in the air above the solution is expressed in terms of kilopascals (kPa). Oxygen binds tightly to myoglobin with a P_{50} of only 0.26 kPa.





It is sometimes intuitively simpler to consider the **dissociation constant, K_d,** which is the reciprocal of K_a ($K_d = 1/K_a$) and is given in units of molar concentration (M). K_d is the equilibrium constant for the release of ligand. The relevant expressions change to

$$K_d = \frac{[P][L]}{[PL]} \tag{7–6}$$

$$[PL] = \frac{[P][L]}{K_d} \tag{7–7}$$

$$\theta = \frac{[L]}{[L] + K_d} \tag{7–8}$$

When [L] is equal to K_d, half of the ligand-binding sites are occupied. When [L] is lower than K_d, little ligand binds to the protein. In order for 90% of the available ligand-binding sites to be occupied, [L] must be nine times greater than K_d. In practice, K_d is used much more often than K_a to express the affinity of a protein for a ligand. Note that a lower value of K_d corresponds to a higher affinity of ligand for the protein. The mathematics can be reduced to simple statements: K_d is the molar concentration of ligand at which half of the available ligand-binding sites are occupied. At this point, the protein is said to have reached half saturation with respect to ligand binding. The more tightly a protein binds a ligand, the lower the concentration of ligand required for half the binding sites to be occupied, and thus the lower the value of K_d. Some representative dissociation constants are given in Table 7–1.

The binding of oxygen to myoglobin follows the patterns discussed above, but because oxygen is a gas, we must make some minor adjustments to the equations. We can simply substitute the concentration of dissolved oxygen for [L] in Equation 7–8 to give

$$\theta = \frac{[O_2]}{[O_2] + K_d} \tag{7–9}$$

As for any ligand, K_d is equal to the $[O_2]$ at which half of the available ligand-binding sites are occupied, or $[O_2]_{0.5}$. Equation 7–9 becomes

$$\theta = \frac{[O_2]}{[O_2] + [O_2]_{0.5}} \tag{7–10}$$

table 7–1

Some Protein Dissociation Constants

Protein	Ligand	K_d (M)*
Avidin (egg white)†	Biotin	1×10^{-15}
Insulin receptor (human)	Insulin	1×10^{-10}
Anti-HIV immunoglobulin (human)‡	gp41 (HIV-1 surface protein)	4×10^{-10}
Nickel-binding protein (*E. coli*)	Ni^{2+}	1×10^{-7}
Calmodulin (rat)§	Ca^{2+}	3×10^{-6}
		2×10^{-5}

*A reported dissociation constant is valid only for the particular solution conditions under which it was measured. K_d values for a protein-ligand interaction can be altered, sometimes by several orders of magnitude, by changes in solution salt concentration, pH, or other variables.

†Interaction of avidin with the enzymatic cofactor biotin is among the strongest noncovalent biochemical interactions known.

‡This immunoglobulin was isolated as part of an effort to develop a vaccine against HIV. Immunoglobulins (described later in the chapter) are highly variable, and the K_d reported here should not be considered characteristic of all immunoglobulins.

§Calmodulin has four binding sites for calcium. The values shown reflect the highest- and lowest-affinity binding sites observed in one set of measurements.

The concentration of a volatile substance in solution, however, is always proportional to its partial pressure in the gas phase above the solution. In experiments using oxygen as a ligand, it is the partial pressure of oxygen, pO_2, that is varied because this is easier to measure than the concentration of dissolved oxygen. If we define the partial pressure of oxygen at $[O_2]_{0.5}$ as P_{50}, substitution in Equation 7–10 gives

$$\theta = \frac{pO_2}{pO_2 + P_{50}} \qquad (7\text{–}11)$$

A binding curve for myoglobin that relates θ to pO_2 is shown in Figure 7–4b.

Protein Structure Affects How Ligands Bind

The binding of a ligand to a protein is rarely as simple as the above equations would suggest. The interaction is greatly affected by protein structure and is often accompanied by conformational changes. For example, the specificity with which heme binds its various ligands is altered when the heme is a component of myoglobin. CO binds to free heme molecules over 20,000 times better than does O_2 (the K_d or P_{50} for CO binding is more than 20,000 times lower than that for O_2) but binds only about 200 times better when the heme is bound in myoglobin. The difference is partly explained by steric hindrance. When O_2 binds to free heme, the axis of the oxygen molecule is positioned at an angle to the Fe—O bond (Fig. 7–5a). In contrast, when CO binds to free heme, the Fe, C, and O atoms lie in a straight line (Fig. 7–5b). In both cases, the binding reflects the geometry of hybrid orbitals in each ligand. In myoglobin, His^{64} (His E7), on the O_2-binding side of the heme, is too far away to coordinate with the heme iron, but it does interact with a ligand bound to heme. This residue, called the *distal His,* does not affect the binding of O_2 (Fig. 7–5c) but may preclude the linear binding of CO, providing one explanation for the diminished binding of CO to heme in myoglobin (and hemoglobin). This effect on CO binding is physiologically important, because CO is a low-level byproduct of cellular metabolism. Other factors, not yet well-defined, also seem to modulate the interaction of heme with CO in these proteins.

The binding of O_2 to the heme in myoglobin also depends on molecular motions, or "breathing," in the protein structure. The heme molecule is deeply buried in the folded polypeptide, with no direct path for oxygen to go from the surrounding solution to the ligand-binding site. If the protein were rigid, O_2 could not enter or leave the heme pocket at a measurable rate. However, rapid molecular flexing of the amino acid side chains produces transient cavities in the protein structure, and O_2 evidently makes its way in and out by moving through these cavities. Computer simulations of rapid structural fluctuations in myoglobin suggest that there are many such pathways. One major route is provided by rotation of the side chain of the distal His (His^{64}), which occurs on a nanosecond (10^{-9} s) time scale. Even subtle conformational changes can be critical for protein activity.

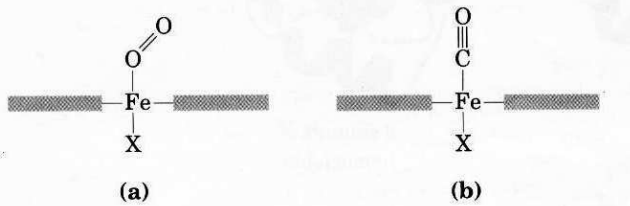
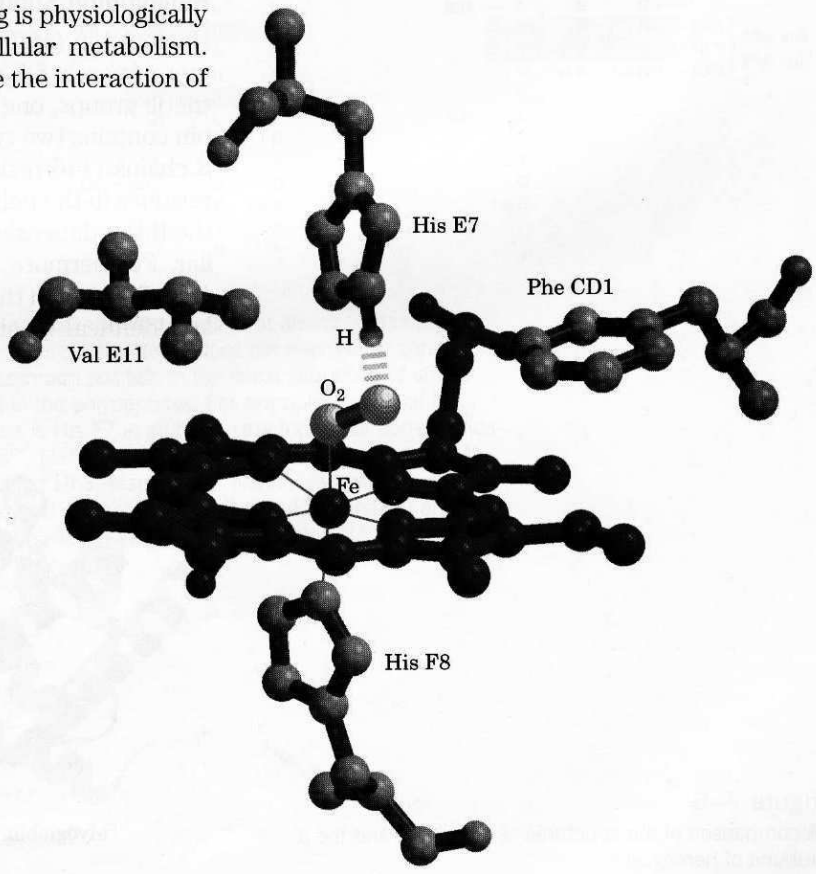


figure 7–5

Steric effects on the binding of ligands to the heme of myoglobin. (a) Oxygen binds to heme with the O_2 axis at an angle, a binding conformation readily accommodated by myoglobin. **(b)** Carbon monoxide binds to free heme with the CO axis perpendicular to the plane of the porphyrin ring. CO binding to the heme in myoglobin is forced to adopt a slight angle because the perpendicular arrangement is sterically blocked by His E7, the distal His. This effect weakens the binding of CO to myoglobin. **(c)** Another view showing the arrangement of key amino acid residues around the heme of myoglobin. The bound O_2 is hydrogen-bonded to the distal His, His E7 (His^{64}), further facilitating the binding of O_2.



Oxygen Is Transported in Blood by Hemoglobin

Nearly all the oxygen carried by whole blood in animals is bound and transported by hemoglobin in erythrocytes (red blood cells). Normal human erythrocytes are small (6 to 9 μm in diameter), biconcave disks. They are formed from precursor stem cells called **hemocytoblasts.** In the maturation process, the stem cell produces daughter cells that form large amounts of hemoglobin and then lose their intracellular organelles—nucleus, mitochondria, and endoplasmic reticulum. Erythrocytes are thus incomplete, vestigial cells, unable to reproduce and, in humans, destined to survive for only about 120 days. Their main function is to carry hemoglobin, which is dissolved in the cytosol at a very high concentration (~34% by weight).

In arterial blood passing from the lungs through the heart to the peripheral tissues, hemoglobin is about 96% saturated with oxygen. In the venous blood returning to the heart, hemoglobin is only about 64% saturated. Thus, each 100 mL of blood passing through a tissue releases about one-third of the oxygen it carries, or 6.5 mL of O_2 gas at atmospheric pressure and body temperature.

Myoglobin, with its hyperbolic binding curve for oxygen (Fig. 7–4b), is relatively insensitive to small changes in the concentration of dissolved oxygen and so functions well as an oxygen-storage protein. Hemoglobin, with its multiple subunits and O_2-binding sites, is better suited to oxygen transport. As we will see, interactions between the subunits of a multimeric protein can permit a highly sensitive response to small changes in ligand concentration. Interactions among the subunits in hemoglobin cause conformational changes that alter the affinity of the protein for oxygen. The modulation of oxygen binding allows the O_2-transport protein to respond to changes in oxygen demand by tissues.

Hemoglobin Subunits Are Structurally Similar to Myoglobin

Hemoglobin (M_r 64,500; abbreviated Hb) is roughly spherical, with a diameter of nearly 5.5 nm. It is a tetrameric protein containing four heme prosthetic groups, one associated with each polypeptide chain. Adult hemoglobin contains two types of globin, two α chains (141 residues each) and two β chains (146 residues each). Although fewer than half of the amino acid residues in the polypeptide sequences of the α and β subunits are identical, the three-dimensional structures of the two types of subunits are very similar. Furthermore, their structures are very similar to that of myoglobin (Fig. 7–6), even though the amino acid sequences of the three polypeptides are identical at only 27 positions (Fig. 7–7). All three polypeptides are

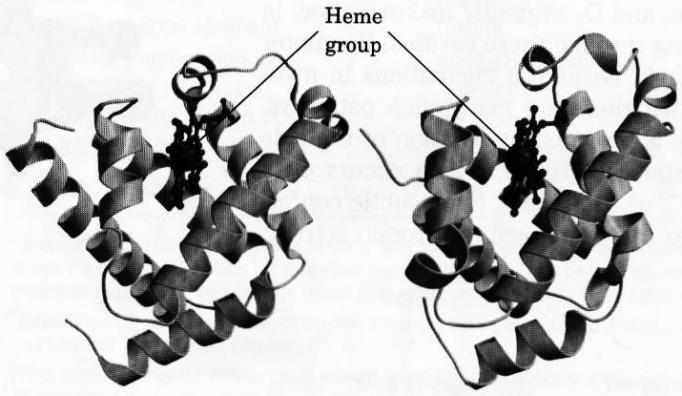


figure 7–6
A comparison of the structures of myoglobin and the β subunit of hemoglobin.

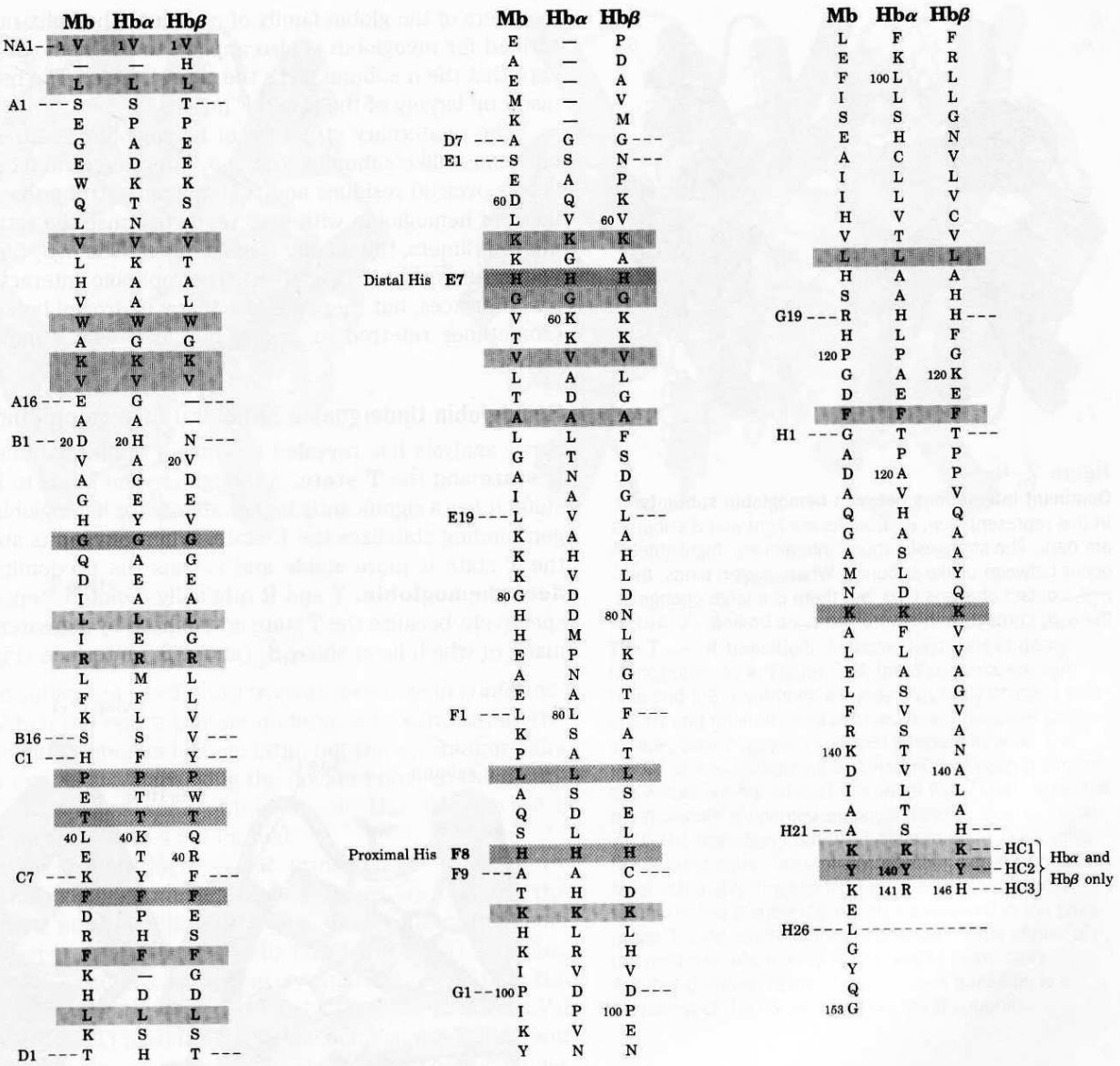


figure 7–7

The amino acid sequences of whale myoglobin and the α and β chains of human hemoglobin. Dashed lines mark helix boundaries. To align the sequences optimally, short breaks must be incorporated into both Hb sequences where a few amino acids are present in the other sequences. With the exception of the missing D helix in Hbα, this alignment permits the use of the helix lettering convention that emphasizes the common positioning of amino acid residues that are identical in all three structures (shaded). Residues shaded in red are conserved in all known globins. Note that a common letter-and-number designation for amino acids in two or three different structures does not necessarily correspond to a common position in the linear sequence of amino acids in the polypeptides. For example, the distal His residue is His E7 in all three structures, but corresponds to His^{64}, His^{58}, and His^{63} in the linear sequences of Mb, Hbα, and Hbβ, respectively. Nonhelical residues at the amino and carboxyl termini, beyond the first (A) and last (H) α-helical segments, are labeled NA and HC, respectively.

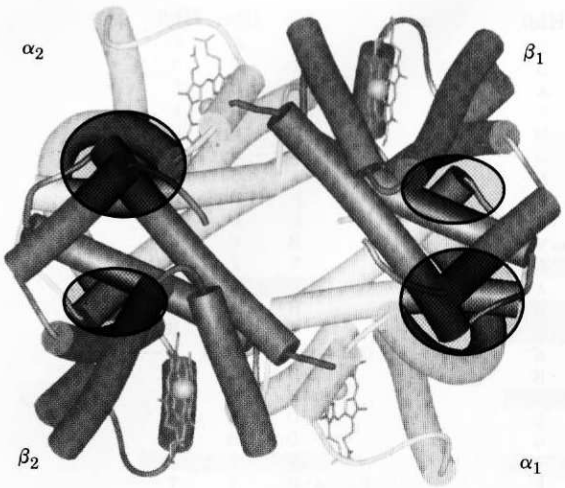


figure 7–8
Dominant interactions between hemoglobin subunits. In this representation, α subunits are light and β subunits are dark. The strongest subunit interactions, highlighted, occur between unlike subunits. When oxygen binds, the $\alpha_1\beta_1$ contact changes little, but there is a large change at the $\alpha_1\beta_2$ contact, with several ion pairs broken.

members of the globin family of proteins. The helix-naming convention described for myoglobin is also applied to the hemoglobin polypeptides, except that the α subunit lacks the short D helix. The heme-binding pocket is made up largely of the E and F helices.

The quaternary structure of hemoglobin features strong interactions between unlike subunits. The $\alpha_1\beta_1$ interface (and its $\alpha_2\beta_2$ counterpart) involves over 30 residues and is sufficiently strong that although mild treatment of hemoglobin with urea tends to cause the tetramer to disassemble into $\alpha\beta$ dimers, the dimers remain intact. The $\alpha_1\beta_2$ (and $\alpha_2\beta_1$) interface involves 19 residues (Fig. 7–8). Hydrophobic interactions predominate at the interfaces, but there are also many hydrogen bonds and a few ion pairs (sometimes referred to as salt bridges), whose importance is discussed below.

Hemoglobin Undergoes a Structural Change on Binding Oxygen

X-ray analysis has revealed two major conformations of hemoglobin: the **R state** and the **T state.** Although oxygen binds to hemoglobin in either state, it has a significantly higher affinity for hemoglobin in the R state. Oxygen binding stabilizes the R state. When oxygen is absent experimentally, the T state is more stable and is thus the predominant conformation of **deoxyhemoglobin.** T and R originally denoted "tense" and "relaxed," respectively, because the T state is stabilized by a greater number of ion pairs, many of which lie at the $\alpha_1\beta_2$ (and $\alpha_2\beta_1$) interface (Fig. 7–9). The binding

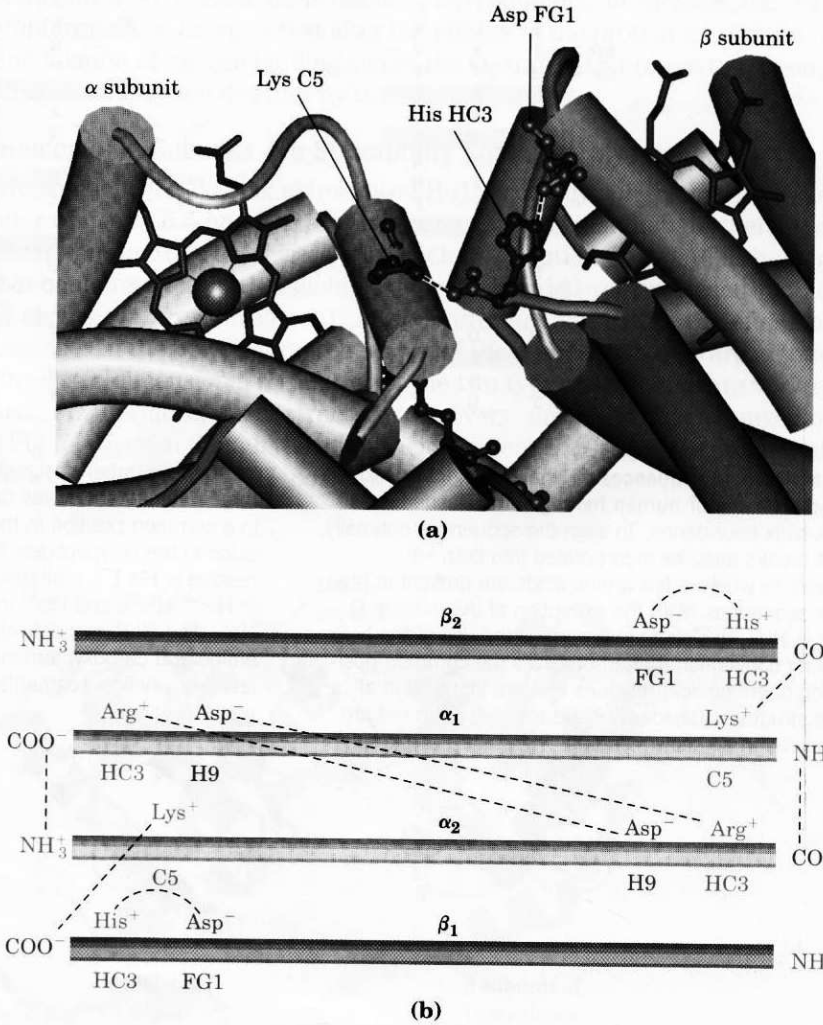


figure 7–9
Some ion pairs that stabilize the T state of deoxyhemoglobin. (a) A close-up view of a portion of a deoxyhemoglobin molecule in the T state. Interactions between the ion pairs His HC3 and Asp FG1 of the β subunit (blue) and between Lys C5 of the α subunit (gray) and the α-carboxyl group of His HC3 of the β subunit are shown with dashed lines. (Recall that HC3 is the carboxyl-terminal residue of the β subunit.) **(b)** The interactions between these ion pairs and others not shown in **(a)** are schematized in this representation of the extended polypeptide chains of hemoglobin.

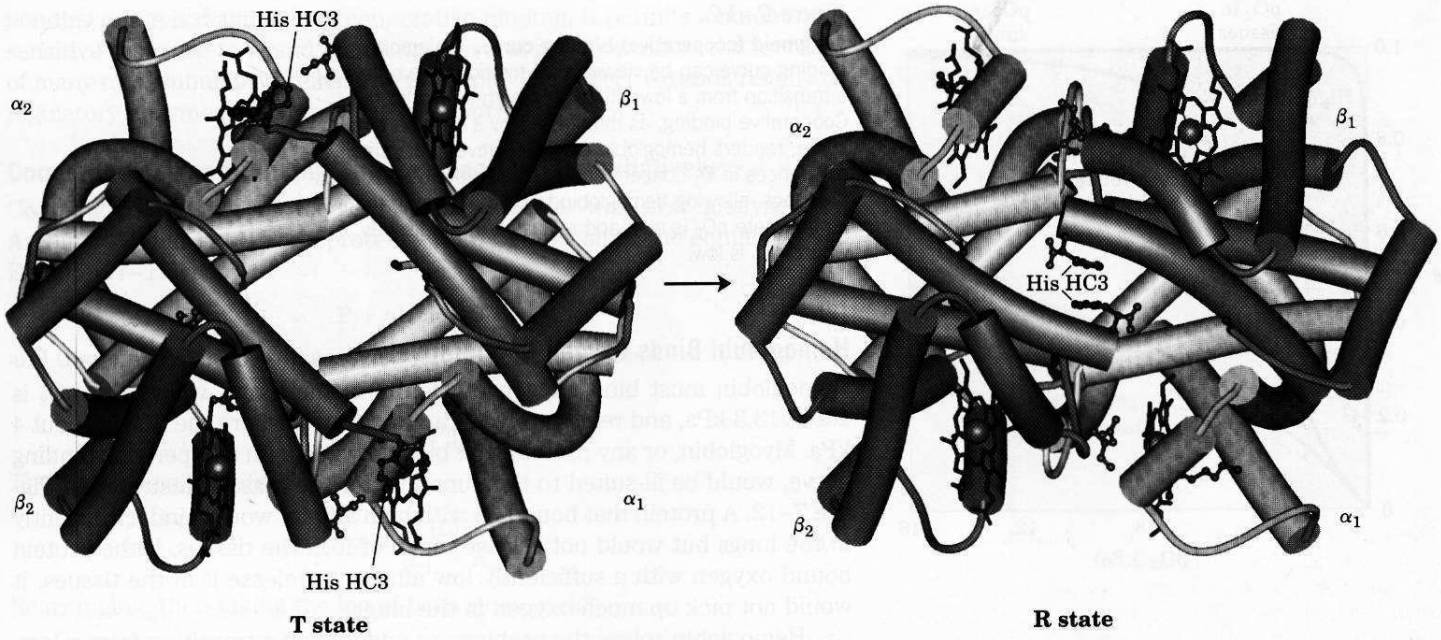


figure 7–10
The T ⟶ R transition. In these depictions of deoxyhemoglobin, as in Figure 7–9, the β subunits are light blue and the α subunits are gray. Positively charged side chains and chain termini involved in ion pairs are shown in blue, their negatively charged partners in pink. The Lys C5 of each α subunit and Asp FG1 of each β subunit are visible but not labeled (compare Fig. 7–9a). Note that the molecule is oriented slightly differently than in Figure 7–9. The transition from the T state to the R state shifts the subunit pairs substantially, affecting certain ion pairs. Most noticeably, the His HC3 residues at the carboxyl termini of the β subunits, which are involved in ion pairs in the T state, rotate in the R state toward the center of the molecule where they are no longer in ion pairs. Another dramatic result of the T ⟶ R transition is a narrowing of the pocket between the β subunits.

of O_2 to a hemoglobin subunit in the T state triggers a change in conformation to the R state. When the entire protein undergoes this transition, the structures of the individual subunits change little, but the $\alpha\beta$ subunit pairs slide past each other and rotate, narrowing the pocket between the β subunits (Fig. 7–10). In this process, some of the ion pairs that stabilize the T state are broken and some new ones are formed.

Max Perutz proposed that the T ⟶ R transition is triggered by changes in the positions of key amino acid side chains surrounding the heme. In the T state, the porphyrin is slightly puckered, causing the heme iron to protrude somewhat on the proximal His (His F8) side. The binding of O_2 causes the heme to assume a more planar conformation, shifting the position of the proximal His and the attached F helix (Fig. 7–11). Also, a Val residue in the E helix (Val E11) partially blocks the heme in the T state and must swing out of the way for oxygen to bind (Fig. 7–10). These changes lead to adjustments in the ion pairs at the $\alpha_1\beta_2$ interface.

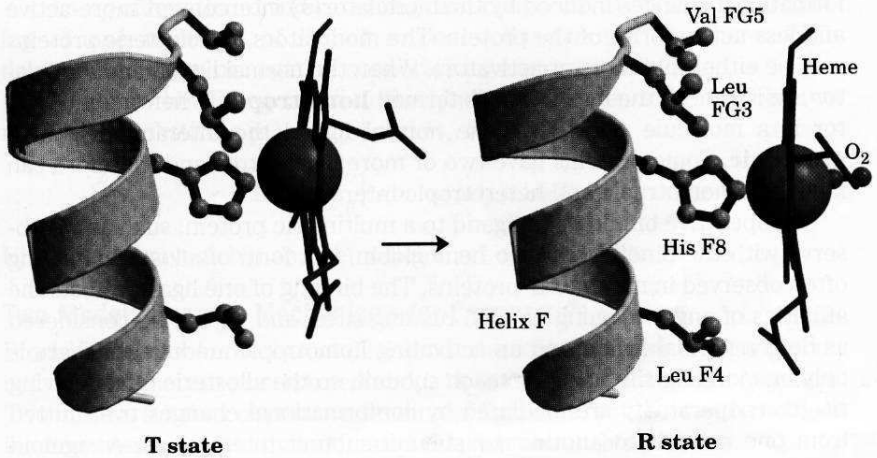


figure 7–11
Changes in conformation near heme on O_2 binding. The shift in the position of the F helix when heme binds O_2 is one of the adjustments that is believed to trigger the T ⟶ R transition.

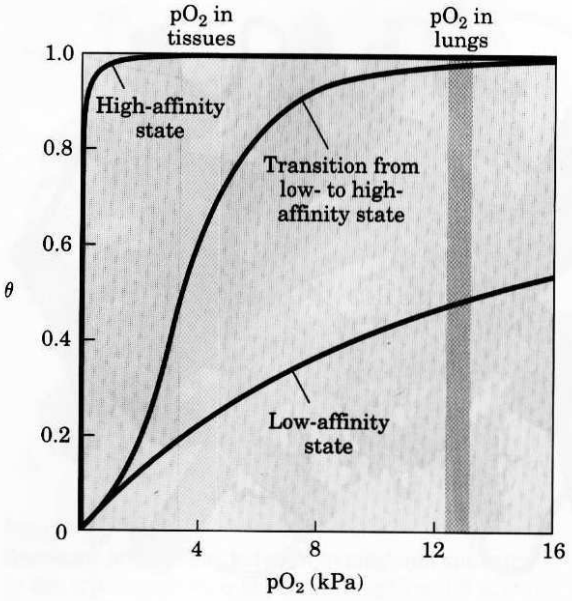


figure 7–12
A sigmoid (cooperative) binding curve. A sigmoid binding curve can be viewed as a hybrid curve reflecting a transition from a low-affinity to a high-affinity state. Cooperative binding, as manifested by a sigmoid binding curve, renders hemoglobin more sensitive to the small differences in O_2 concentration between the tissues and the lungs, allowing hemoglobin to bind oxygen in the lungs where pO_2 is high and release it in the tissues where pO_2 is low.

Hemoglobin Binds Oxygen Cooperatively

Hemoglobin must bind oxygen efficiently in the lungs, where the pO_2 is about 13.3 kPa, and release oxygen in the tissues, where the pO_2 is about 4 kPa. Myoglobin, or any protein that binds oxygen with a hyperbolic binding curve, would be ill-suited to this function, for the reason illustrated in Figure 7–12. A protein that bound O_2 with high affinity would bind it efficiently in the lungs but would not release much of it in the tissues. If the protein bound oxygen with a sufficiently low affinity to release it in the tissues, it would not pick up much oxygen in the lungs.

Hemoglobin solves the problem by undergoing a transition from a low-affinity state (the T state) to a high-affinity state (the R state) as more O_2 molecules are bound. As a result, hemoglobin has a hybrid S-shaped, or sigmoid, binding curve for oxygen (Fig. 7–12). A single-subunit protein with a single ligand-binding site cannot produce a sigmoid binding curve—even if binding elicits a conformational change—because each molecule of ligand binds independently and cannot affect the binding of another molecule. In contrast, O_2 binding to individual subunits of hemoglobin can alter the affinity for O_2 in adjacent subunits. The first molecule of O_2 that interacts with deoxyhemoglobin binds weakly, because it binds to a subunit in the T state. Its binding, however, leads to conformational changes that are communicated to adjacent subunits, making it easier for additional molecules of O_2 to bind. In effect, the T $\longrightarrow$ R transition occurs more readily in the second subunit once O_2 is bound to the first subunit. The last (fourth) O_2 molecule binds to a heme in a subunit that is already in the R state, and hence it binds with much higher affinity than the first molecule.

An **allosteric protein** is one in which the binding of a ligand to one site affects the binding properties of another site on the same protein. The term allosteric derives from the Greek *allos,* "other," and *stereos,* "solid" or "shape." Allosteric proteins are those having "other shapes" or conformations induced by the binding of ligands referred to as modulators. The conformational changes induced by the modulator(s) interconvert more-active and less-active forms of the protein. The modulators for allosteric proteins may be either inhibitors or activators. When the normal ligand and modulator are identical, the interaction is termed **homotropic.** When the modulator is a molecule other than the normal ligand the interaction is **heterotropic.** Some proteins have two or more modulators and therefore can have both homotropic and heterotropic interactions.

Cooperative binding of a ligand to a multimeric protein, such as we observe with the binding of O_2 to hemoglobin, is a form of allosteric binding often observed in multimeric proteins. The binding of one ligand affects the affinities of any remaining unfilled binding sites, and O_2 can be considered as both a normal ligand and an activating homotropic modulator. There is only one binding site for O_2 on each subunit, so the allosteric effects giving rise to cooperativity are mediated by conformational changes transmitted from one subunit to another by subunit-subunit interactions. A sigmoid

bonding curve is diagnostic of cooperative binding. It permits a much more sensitive response to ligand concentration and is important to the function of many multisubunit proteins. The principle of allostery extends readily to regulatory enzymes, as we will see in Chapter 8.

Cooperative Ligand Binding Can Be Described Quantitatively

Cooperative binding of oxygen by hemoglobin was first analyzed by Archibald Hill in 1910. For a protein with n binding sites, the equilibrium of Equation 7–1 becomes

$$P + nL \rightleftharpoons PL_n \tag{7–12}$$

and the expression for the association constant becomes

$$K_a = \frac{[PL_n]}{[P][L]^n} \tag{7–13}$$

The expression for θ (see Eqn 7–8) is

$$\theta = \frac{[L]^n}{[L]^n + K_d} \tag{7–14}$$

Rearranging, then taking the log of both sides, yields

$$\frac{\theta}{1-\theta} = \frac{[L]^n}{K_d} \tag{7–15}$$

$$\log\left(\frac{\theta}{1-\theta}\right) = n \log [L] - \log K_d \tag{7–16}$$

Equation 7–16 is the **Hill equation,** and a plot of log $[\theta/(1-\theta)]$ versus log [L] is called a **Hill plot.** Based on the equation, the Hill plot should have a slope of n. However, the experimentally determined slope actually reflects not the number of binding sites, but the degree of interaction between them. The slope of a Hill plot is therefore denoted n_H, the **Hill coefficient,** which is a measure of the degree of cooperativity. If n_H equals 1, ligand binding is not cooperative, a situation that can arise even in a multisubunit protein if the subunits do not communicate. An n_H of greater than 1 indicates positive cooperativity in ligand binding. This is the situation observed in hemoglobin, in which the binding of one molecule of ligand facilitates the binding of others. The theoretical upper limit for n_H is reached when $n_H = n$. In this case the binding would be completely cooperative: all binding sites on the protein would bind ligand simultaneously, and no protein molecules partially saturated with ligand would be present under any conditions. This limit is never reached in practice, and the measured value of n_H is always less than the actual number of ligand-binding sites in the protein.

An n_H of less than 1 indicates negative cooperativity, in which the binding of one molecule of ligand *impedes* the binding of others. Well-documented cases of negative cooperativity are rare.

To adapt the Hill equation to the binding of oxygen to hemoglobin we must again substitute pO_2 for [L] and P_{50} for K_d:

$$\log\left(\frac{\theta}{1-\theta}\right) = n \log pO_2 - \log P_{50} \tag{7–17}$$

Hill plots for myoglobin and hemoglobin are given in Figure 7–13.

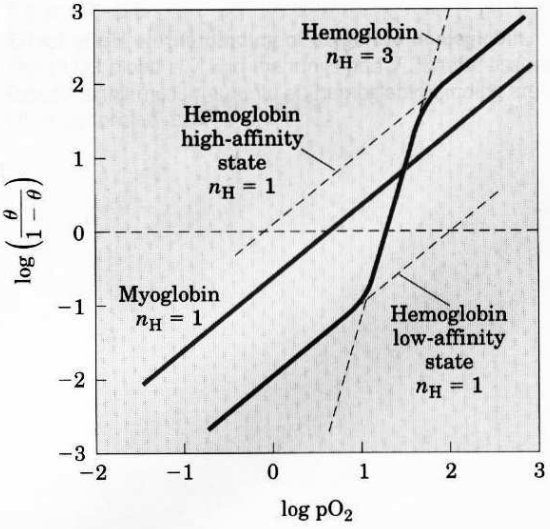


figure 7–13

Hill plots for the binding of oxygen to myoglobin and hemoglobin. When $n_H = 1$, there is no evident cooperativity. The maximum degree of cooperativity observed for hemoglobin corresponds approximately to $n_H = 3$. Note that while this indicates a high level of cooperativity, n_H is less than n, the number of O_2-binding sites in hemoglobin. This is normal for a protein that exhibits allosteric binding behavior.

Two Models Suggest Mechanisms for Cooperative Binding

Biochemists now know a great deal about the T and R states of hemoglobin, but much remains to be learned about how the T ⟶ R transition occurs. Two models for the cooperative binding of ligands to proteins with multiple binding sites have greatly influenced thinking about this problem.

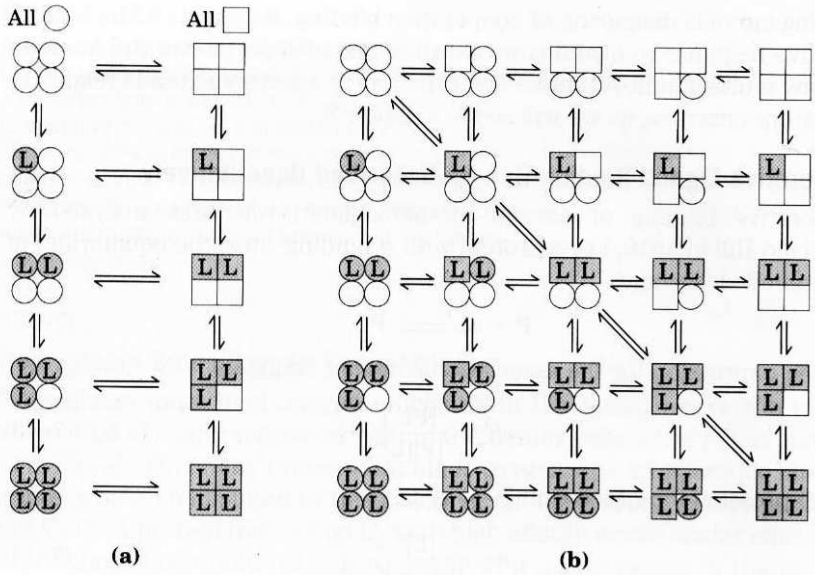


figure 7–14
Two general models for the interconversion of inactive and active forms of cooperative ligand-binding proteins. Although the models may be applied to any protein—including any enzyme (Chapter 8)—that exhibits cooperative binding, four subunits are shown because the model was originally proposed for hemoglobin. In the concerted, or all-or-none, model **(a)** all the subunits are postulated to be in the same conformation, either all ○ (low affinity or inactive) or all □ (high affinity or active). Depending on the equilibrium, K_1, between ○ and □ forms, the binding of one or more ligand molecules (L) will pull the equilibrium toward the □ form. Subunits with bound L are shaded. In the sequential model **(b)** each individual subunit can be in either the ○ or □ form. A very large number of conformations is thus possible.

The first model was proposed by Jacques Monod, Jeffries Wyman, and Jean-Pierre Changeux in 1965, and is called the **MWC model** or the **concerted model** (Fig. 7–14a). The concerted model assumes that the subunits of a cooperatively binding protein are functionally identical, that each subunit can exist in (at least) two conformations, and that all subunits undergo the transition from one conformation to the other simultaneously. In this model, no protein has individual subunits in different conformations. The two conformations are in equilibrium. The ligand can bind to either conformation, but binds each with different affinity. Successive binding of ligand molecules to the low-affinity conformation (which is more stable in the absence of ligand) makes a transition to the high-affinity conformation more likely.

In the second model, the **sequential model** (Fig. 7–14b), proposed in 1966 by Daniel Koshland and colleagues, ligand binding can induce a change of conformation in an individual subunit. A conformational change in one subunit makes a similar change in an adjacent subunit, as well as the binding of a second ligand molecule, more likely. There are more potential intermediate states in this model than in the concerted model. The two models are not mutually exclusive; the concerted model may be viewed as the "all-or-none" limiting case of the sequential model. In Chapter 8 we will use these models when we investigate allosteric enzymes.

Hemoglobin Also Transports H^+ and CO_2

In addition to carrying nearly all the oxygen required by cells from the lungs to the tissues, hemoglobin carries two end products of cellular respiration—H^+ and CO_2—from the tissues to the lungs and the kidneys, where they are excreted. The CO_2, produced by oxidation of organic fuels in mitochondria, is hydrated to form bicarbonate:

$$CO_2 + H_2O \rightleftharpoons H^+ + HCO_3^-$$

This reaction is catalyzed by **carbonic anhydrase,** an enzyme particularly abundant in erythrocytes. Carbon dioxide is not very soluble in aqueous solution, and bubbles of CO_2 would form in the tissues and blood if it were not converted to bicarbonate. As you can see from the equation, the hydration of CO_2 results in an increase in the H^+ concentration (a decrease in pH) in

the tissues. The binding of oxygen by hemoglobin is profoundly influenced by pH and CO_2 concentration, so the interconversion of CO_2 and bicarbonate is of great importance to the regulation of oxygen binding and release in the blood.

Hemoglobin transports about 20% of the total H^+ and CO_2 formed in the tissues to the lungs and the kidneys. The binding of H^+ and CO_2 is inversely related to the binding of oxygen. At the relatively low pH and high CO_2 concentration of peripheral tissues, the affinity of hemoglobin for oxygen decreases as H^+ and CO_2 are bound, and O_2 is released to the tissues. Conversely, in the capillaries of the lung, as CO_2 is excreted and the blood pH consequently rises, the affinity of hemoglobin for oxygen increases and the protein binds more O_2 for transport to the peripheral tissues. This effect of pH and CO_2 concentration on the binding and release of oxygen by hemoglobin is called the **Bohr effect,** after Christian Bohr, the Danish physiologist (and father of physicist Niels Bohr) who discovered it in 1904.

The binding equilibrium for hemoglobin and one molecule of oxygen can be designated by the reaction

$$Hb + O_2 \rightleftharpoons HbO_2$$

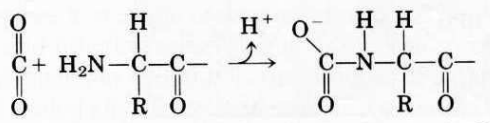
but this is not a complete statement. To account for the effect of H^+ concentration on this binding equilibrium, we rewrite the reaction as

$$HHb^+ + O_2 \rightleftharpoons HbO_2 + H^+$$

where HHb^+ denotes a protonated form of hemoglobin. This equation tells us that the O_2-saturation curve of hemoglobin is influenced by the H^+ concentration (Fig. 7–15). Both O_2 and H^+ are bound by hemoglobin, but with inverse affinity. When the oxygen concentration is high, as in the lungs, hemoglobin binds O_2 and releases protons. When the oxygen concentration is low, as in the peripheral tissues, H^+ is bound and O_2 is released.

Oxygen and H^+ are not bound at the same sites in hemoglobin. Oxygen binds to the iron atoms of the hemes, whereas H^+ binds to any of several amino acid residues in the protein. A major contribution to the Bohr effect is made by His^{146} (His HC3) of the β subunits. When protonated, this residue forms one of the ion pairs—to Asp^{94} (Asp FG1)—that helps stabilize deoxyhemoglobin in the T state (Fig. 7–9). The ion pair stabilizes the protonated form of His HC3, giving this residue an abnormally high pK_a in the T state. The pK_a falls to its normal value of 6.0 in the R state because the ion pair cannot form, and this residue is largely unprotonated in oxyhemoglobin at pH 7.6, the blood pH in the lungs. As the concentration of H^+ rises, protonation of His HC3 promotes release of oxygen by favoring a transition to the T state. Protonation of the amino-terminal residues of the α subunits, certain other His residues, and perhaps other groups has a similar effect.

Thus we see that the four polypeptide chains of hemoglobin communicate with each other not only about O_2 binding to their heme groups, but also about H^+ binding to specific amino acid residues. And there is still more to the story. Hemoglobin also binds CO_2, again in a manner inversely related to the binding of oxygen. Carbon dioxide binds as a carbamate group to the α-amino group at the amino-terminal end of each globin chain, forming carbaminohemoglobin:



Amino-terminal residue → Carbamino-terminal residue

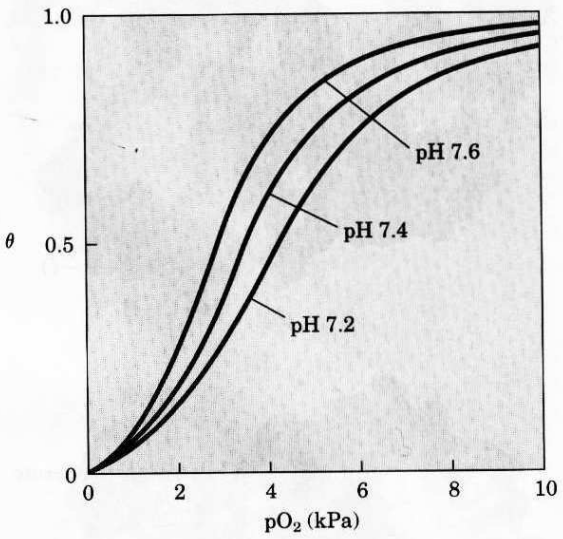


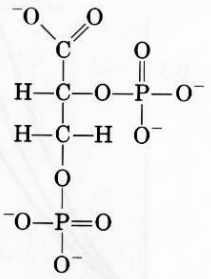
figure 7–15
Effect of pH on the binding of oxygen to hemoglobin. The pH of blood is 7.6 in the lungs and 7.2 in the tissues. Experimental measurements on hemoglobin binding are often performed at pH 7.4.

This reaction produces H^+, contributing to the Bohr effect. The bound carbamates also form additional salt bridges (not shown in Fig. 7–9) that help to stabilize the T state and promote the release of oxygen.

When the concentration of carbon dioxide is high, as in peripheral tissues, some CO_2 binds to hemoglobin and the affinity for O_2 decreases, causing its release. Conversely, when hemoglobin reaches the lungs, the high oxygen concentration promotes binding of O_2 and release of CO_2. It is the capacity to communicate ligand-binding information from one polypeptide subunit to the others that makes the hemoglobin molecule so beautifully adapted to integrating the transport of O_2, CO_2, and H^+ by erythrocytes.

Oxygen Binding to Hemoglobin Is Regulated by 2,3-Bisphosphoglycerate

The interaction of **2,3-bisphosphoglycerate** (BPG) with hemoglobin provides an example of heterotropic allosteric modulation. BPG is present in relatively high concentrations in erythrocytes. When hemoglobin is isolated, it contains substantial amounts of bound BPG, which can be difficult to remove completely. In fact, the O_2-binding curves for hemoglobin that we have examined to this point were obtained in the presence of bound BPG. 2,3-Bisphosphoglycerate is known to greatly reduce the affinity of hemoglobin for oxygen—there is an inverse relationship between the binding of O_2 and the binding of BPG. We can therefore describe another binding process for hemoglobin:



2,3-Bisphosphoglycerate

$$HbBPG + O_2 \rightleftharpoons HbO_2 + BPG$$

BPG binds at a site distant from the oxygen-binding site and regulates the O_2-binding affinity of hemoglobin in relation to the pO_2 in the lungs. BPG plays an important role in the physiological adaptation to the lower pO_2 available at high altitudes. For a healthy human strolling by the ocean, the binding of O_2 to hemoglobin is regulated such that the amount of O_2 delivered to the tissues is equivalent to nearly 40% of the maximum that could be carried by the blood (Fig. 7–16). If the same person is quickly transported to a mountainside at an altitude of 4,500 meters, where the pO_2 is considerably lower, the delivery of O_2 to the tissues is reduced. However,

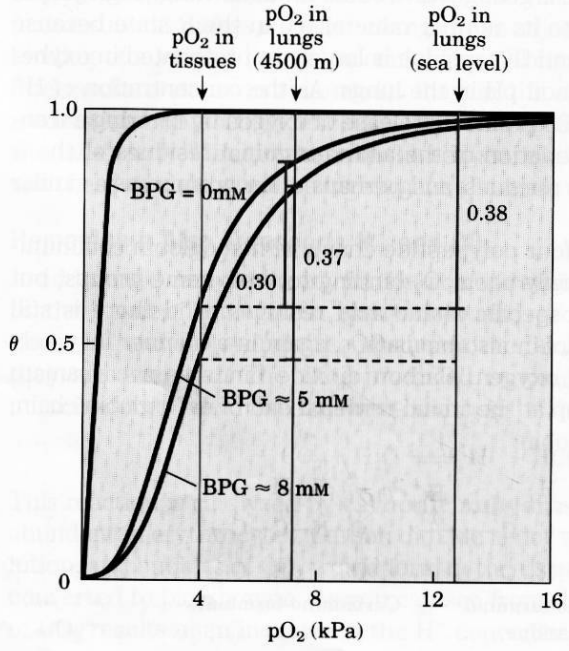


figure 7–16
Effect of BPG on the binding of oxygen to hemoglobin. The BPG concentration in normal human blood is about 5 mM at sea level and about 8 mM at high altitudes. Note that hemoglobin binds to oxygen quite tightly when BPG is entirely absent, and the binding curve appears to be hyperbolic. In reality, the measured Hill coefficient for O_2-binding cooperativity decreases only slightly (from 3 to about 2.5) when BPG is removed from hemoglobin, but the rising part of the sigmoid curve is confined to a very small region close to the origin. At sea level, hemoglobin is nearly saturated with O_2 in the lungs, but only 60% saturated in the tissues, so that the amount of oxygen released in the tissues is close to 40% of the maximum that can be carried in the blood. At high altitudes, O_2 delivery declines by about one-fourth, to 30% of maximum. An increase in BPG concentration, however, decreases the affinity of hemoglobin for O_2 so that nearly 40% of what can be carried is again delivered to the tissues.

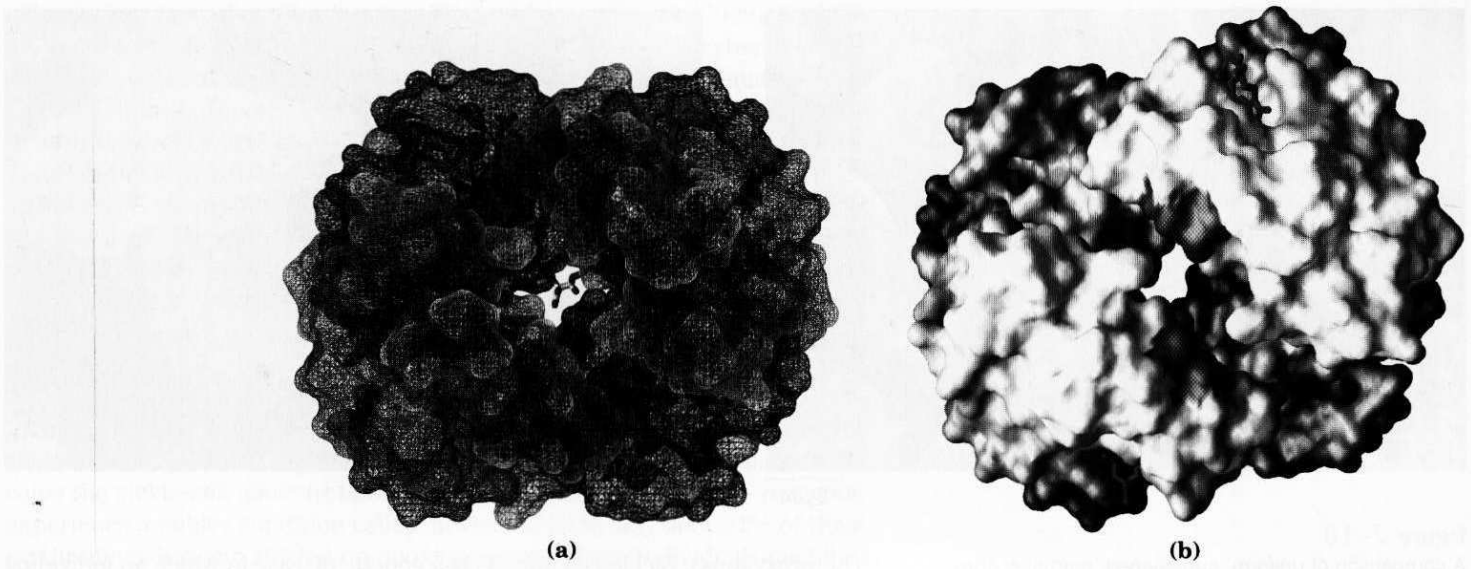


figure 7–17
Binding of BPG to deoxyhemoglobin. (a) BPG binding stabilizes the T state of deoxyhemoglobin, shown here as a mesh surface image. **(b)** The negative charges of BPG interact with several positively charged groups (shown in blue in this GRASP surface image) that surround the pocket between the β subunits in the T state. **(c)** The binding pocket for BPG disappears on oxygenation, following transition to the R state. (Compare **(b)** and **(c)** with Fig. 7–10.)

after just a few hours at the higher altitude, the BPG concentration in the blood has begun to rise, leading to a decrease in the affinity of hemoglobin for oxygen. This adjustment in the BPG level has only a small effect on the binding of O_2 in the lungs but a considerable effect on the release of O_2 in the tissues. As a result, the delivery of oxygen to the tissues is restored to nearly 40% of that which can be transported by the blood. The situation is reversed when the person returns to sea level. The BPG concentration in erythrocytes also increases in people suffering from **hypoxia,** lowered oxygenation of peripheral tissues due to inadequate function of the lungs or circulatory system.

BPG binds to hemoglobin in the cavity between the β subunits in the T state (Fig. 7–17). This cavity is lined with positively charged amino acid residues that interact with the negatively charged groups of BPG. Unlike O_2, only one molecule of BPG is bound to each hemoglobin tetramer. BPG lowers hemoglobin's affinity for oxygen by stabilizing the T state. The transition to the R state narrows the binding pocket for BPG, precluding BPG binding. In the absence of BPG, hemoglobin is converted to the R state more easily.

Regulation of oxygen binding to hemoglobin by BPG has an important role in fetal development. Because a fetus must extract oxygen from its mother's blood, fetal hemoglobin must have greater affinity than the maternal hemoglobin for O_2. In fetuses, γ subunits are synthesized rather than β subunits, and $\alpha_2\gamma_2$ hemoglobin is formed. This tetramer has a much lower affinity for BPG than normal adult hemoglobin, and a correspondingly higher affinity for O_2.

Sickle-Cell Anemia Is a Molecular Disease of Hemoglobin

The great importance of the amino acid sequence in determining the secondary, tertiary, and quaternary structures of globular proteins, and thus their biological functions, is strikingly demonstrated by the hereditary human disease sickle-cell anemia. More than 300 genetic variants of hemoglobin are known to occur in the human population. Most of these variations consist of differences in a single amino acid residue. The effects on hemoglobin structure and function are often minor but can sometimes be extraordinary. Each hemoglobin variation is the product of an altered gene. The variant genes are called alleles. Because humans generally have two copies of each gene, an individual may have two copies of one allele (thus being homozygous for that gene) or one copy of each of two different alleles

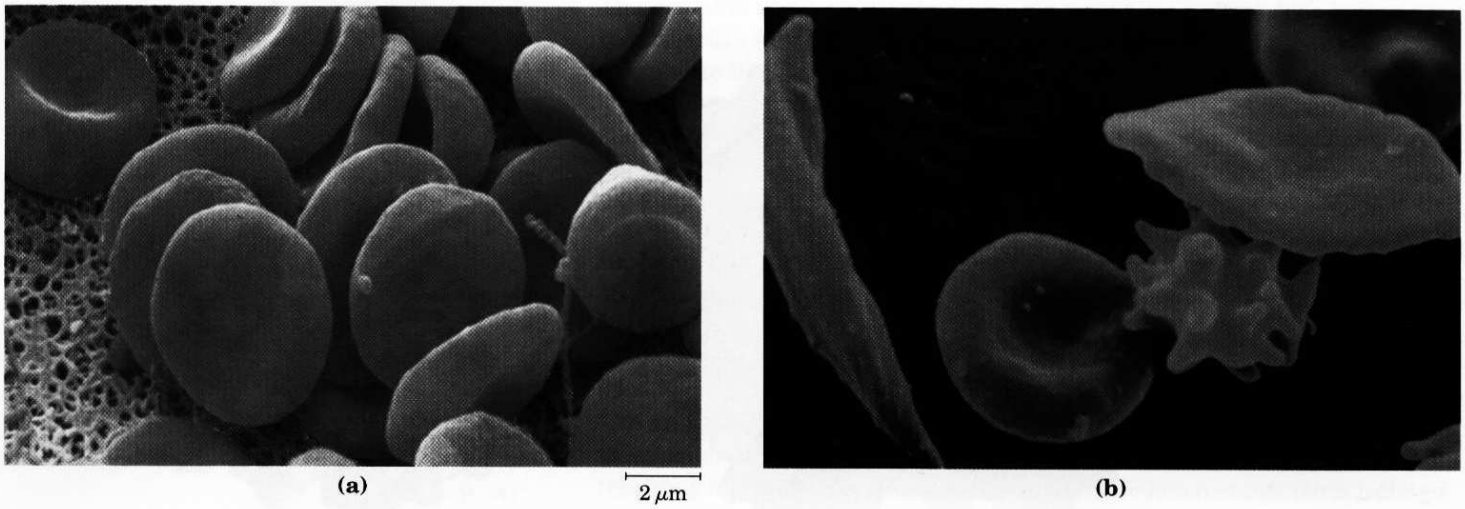


figure 7–18
A comparison of uniform, cup-shaped, normal erythrocytes **(a)** with the variably shaped erythrocytes seen in sickle-cell anemia **(b).** These cells range from normal to spiny or sickle-shaped.

(heterozygous). Sickle-cell anemia is a genetic disease in which an individual has inherited the allele for sickle-cell hemoglobin from both parents. The erythrocytes of these individuals are fewer and also abnormal. In addition to an unusually large number of immature cells, the blood contains many long, thin, crescent-shaped erythrocytes that look like the blade of a sickle (Fig. 7–18). When hemoglobin from sickle cells (called hemoglobin S) is deoxygenated, it becomes insoluble and forms polymers that aggregate into tubular fibers (Fig. 7–19). Normal hemoglobin (hemoglobin A) remains sol-

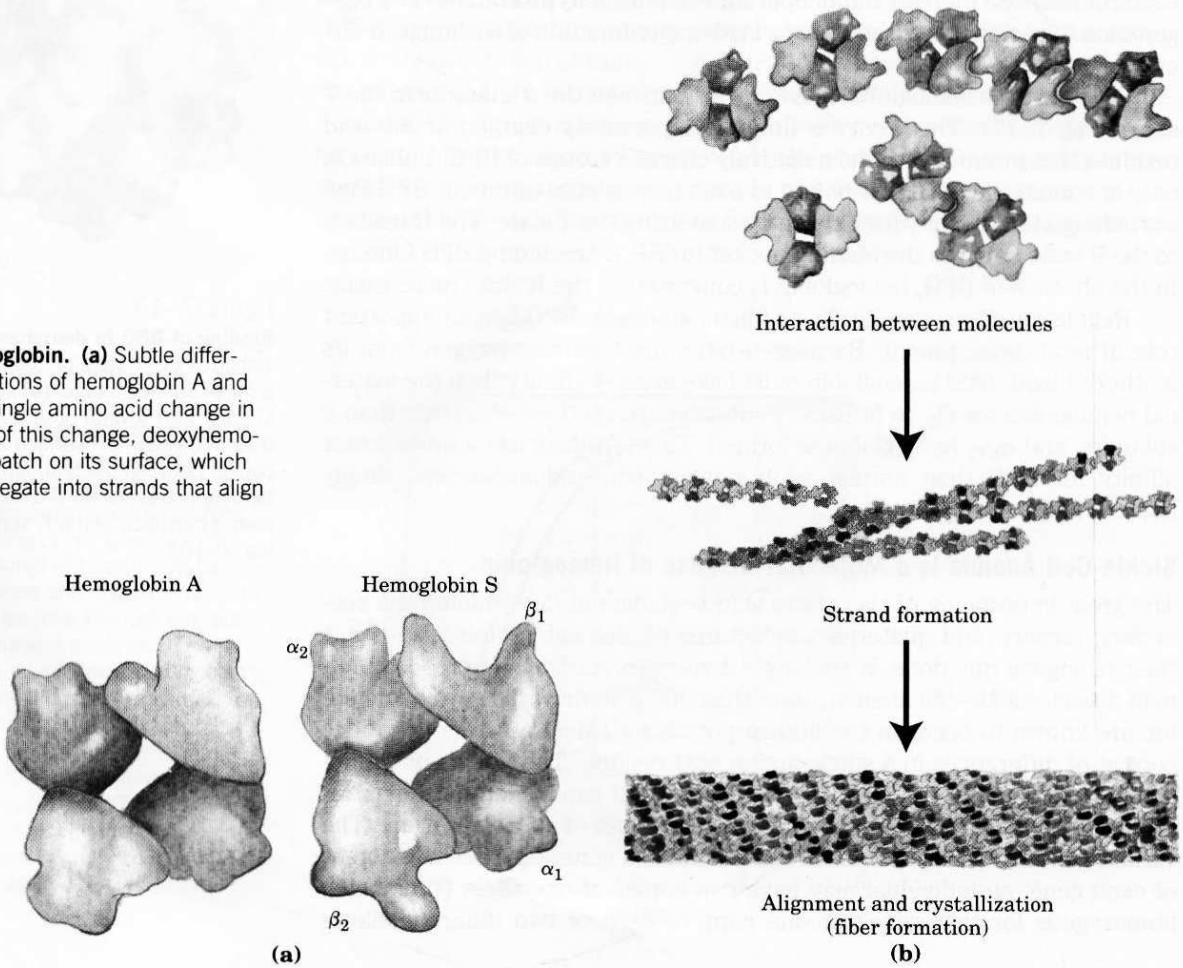


figure 7–19
Normal and sickle-cell hemoglobin. (a) Subtle differences between the conformations of hemoglobin A and hemoglobin S result from a single amino acid change in the β chains. **(b)** As a result of this change, deoxyhemoglobin S has a hydrophobic patch on its surface, which causes the molecules to aggregate into strands that align into insoluble fibers.

uble on deoxygenation. The insoluble fibers of deoxygenated hemoglobin S are responsible for the deformed sickle shape of the erythrocytes, and the proportion of sickled cells increases greatly as blood is deoxygenated.

The altered properties of hemoglobin S result from a single amino acid substitution, a Val instead of a Glu residue at position 6 in the two β chains. The R group of valine has no electric charge, whereas glutamate has a negative charge at pH 7.4. Hemoglobin S therefore has two fewer negative charges than hemoglobin A, one for each of the two β chains. Replacement of the Glu residue by Val creates a "sticky" hydrophobic contact point at position 6 of the β chain, which is on the outer surface of the molecule. These sticky spots cause deoxyhemoglobin S molecules to associate abnormally with each other, forming the long, fibrous aggregates characteristic of this disorder.

Sickle-cell anemia occurs in individuals homozygous for the sickle-cell allele of the gene encoding the β subunit of hemoglobin. Individuals who receive the sickle-cell allele from only one parent and are thus heterozygous experience a milder condition called sickle-cell trait; only about 1% of their erythrocytes become sickled on deoxygenation. These individuals may live completely normal lives if they avoid vigorous exercise or other stresses on the circulatory system.

People with sickle-cell anemia suffer from repeated crises brought on by physical exertion. They become weak, dizzy, and short of breath, and they also experience heart murmurs and an increased pulse rate. The hemoglobin content of their blood is only about half the normal value of 15 to 16 g/100 mL because sickled cells are very fragile and rupture easily; this results in anemia ("lack of blood"). An even more serious consequence is that capillaries become blocked by the long, abnormally shaped cells, causing severe pain and interfering with normal organ function—a major factor in the early death of many people with the disease.

Without medical treatment, people with sickle-cell anemia usually die in childhood. Nevertheless, the sickle-cell allele is surprisingly common in certain parts of Africa. Investigation into the persistence of an allele that is so obviously deleterious in homozygous individuals led to the finding that the allele confers a small but significant resistance to lethal forms of malaria in heterozygous individuals. Natural selection has resulted in an allele population that balances the deleterious effects of the homozygous condition against the resistance to malaria afforded by the heterozygous condition.

Complementary Interactions between Proteins and Ligands: The Immune System and Immunoglobulins

Our discussion of oxygen-binding proteins showed how the conformations of these proteins affect and are affected by the binding of small ligands (O_2 or CO) to the heme group. However, most protein-ligand interactions do not involve a prosthetic group. Instead, the binding site for a ligand is more often like the hemoglobin binding site for BPG—a cleft in the protein lined with amino acid residues, arranged to render the binding interaction highly specific. Effective discrimination between ligands is the norm at binding sites, even when the ligands have only minor structural differences.

All vertebrates have an immune system capable of distinguishing molecular "self" from "nonself" and then destroying those entities identified as nonself. In this way, the immune system eliminates viruses, bacteria, and other pathogens and molecules that may pose a threat to the organism. On a physiological level, the response of the immune system to an invader is an intricate and coordinated set of interactions among many classes of proteins,

molecules, and cell types. However, at the level of individual proteins, the immune response demonstrates how an acutely sensitive and specific biochemical system is built upon the reversible binding of ligands to proteins.

The Immune Response Features a Specialized Array of Cells and Proteins

Immunity is brought about by a variety of **leukocytes** (white blood cells), including **macrophages** and **lymphocytes,** all arising from undifferentiated stem cells in the bone marrow. Leukocytes can leave the bloodstream and patrol the tissues, each cell producing one or more proteins capable of recognizing and binding to molecules that might signal an infection.

The immune response consists of two complementary systems, the humoral and cellular immune systems. The **humoral immune system** (Latin *humor,* "fluid") is directed at bacterial infections and extracellular viruses (those found in the body fluids), but can also respond to individual proteins introduced into the organism. The **cellular immune system** destroys host cells infected by viruses and also destroys some parasites and foreign tissues.

The proteins at the heart of the humoral immune response are soluble proteins called **antibodies** or **immunoglobulins,** often abbreviated Ig. Immunoglobulins bind bacteria, viruses, or large molecules identified as foreign and target them for destruction. Making up 20% of blood protein, the immunoglobulins are produced by **B lymphocytes** or **B cells,** so named because they complete their development in the *b*one marrow.

The agents at the heart of the cellular immune response are a class of **T lymphocytes** or **T cells** (so called because the latter stages of their development occur in the *t*hymus) known as **cytotoxic T cells (T_C cells,** also called killer T cells). Recognition of infected cells or parasites involves proteins called **T-cell receptors** on the surface of T_C cells. Recall from Chapter 2 (p. 30) that receptors are proteins, usually found on the outer surface of cells and extending through the plasma membrane; they recognize and bind extracellular ligands, triggering changes inside the cell.

In addition to cytotoxic T cells, there are **helper T cells (T_H cells),** whose function it is to produce soluble signaling proteins called cytokines, which include the interleukins. T_H cells interact with macrophages. Table 7–2 summarizes the functions of the various leukocytes of the immune system.

table 7–2

Some Types of Leukocytes Associated with the Immune System

Cell type	Function
Macrophages	Ingest large particles and cells by phagocytosis
B lymphocytes (B cells)	Produce and secrete antibodies
T lymphocytes (T cells)	
Cytotoxic (killer) T cells (T_C)	Interact with infected host cells through receptors on T-cell surface
Helper T cells (T_H)	Interact with macrophages and secrete cytokines (interleukins) that stimulate T_C, T_H, and B cells to proliferate.

Each recognition protein of the immune system, either an antibody produced by a B cell or a receptor on the surface of a T cell, specifically binds some particular chemical structure, distinguishing it from virtually all others. Humans are capable of producing over 10^8 different antibodies with distinct binding specificities. This extraordinary diversity makes it likely that any chemical structure on the surface of a virus or invading cell will be recognized and bound by one or more antibodies. Antibody diversity is derived from random reassembly of a set of immunoglobulin gene segments via genetic recombination mechanisms that are discussed in Chapter 25.

Some properties of the interactions between antibodies or T-cell receptors and the molecules they bind are unique to the immune system, and a specialized lexicon is used to describe them. Any molecule or pathogen capable of eliciting an immune response is called an **antigen.** An antigen may be a virus, a bacterial cell wall, or an individual protein or other macromolecule. A complex antigen may be bound by a number of different antibodies. An individual antibody or T-cell receptor binds only a particular molecular structure within the antigen, called its **antigenic determinant** or **epitope.**

It would be unproductive for the immune system to respond to small molecules that are common intermediates and products of cellular metabolism. Molecules of M_r <5,000 are generally not antigenic. However, small molecules can be covalently attached to large proteins in the laboratory, and in this

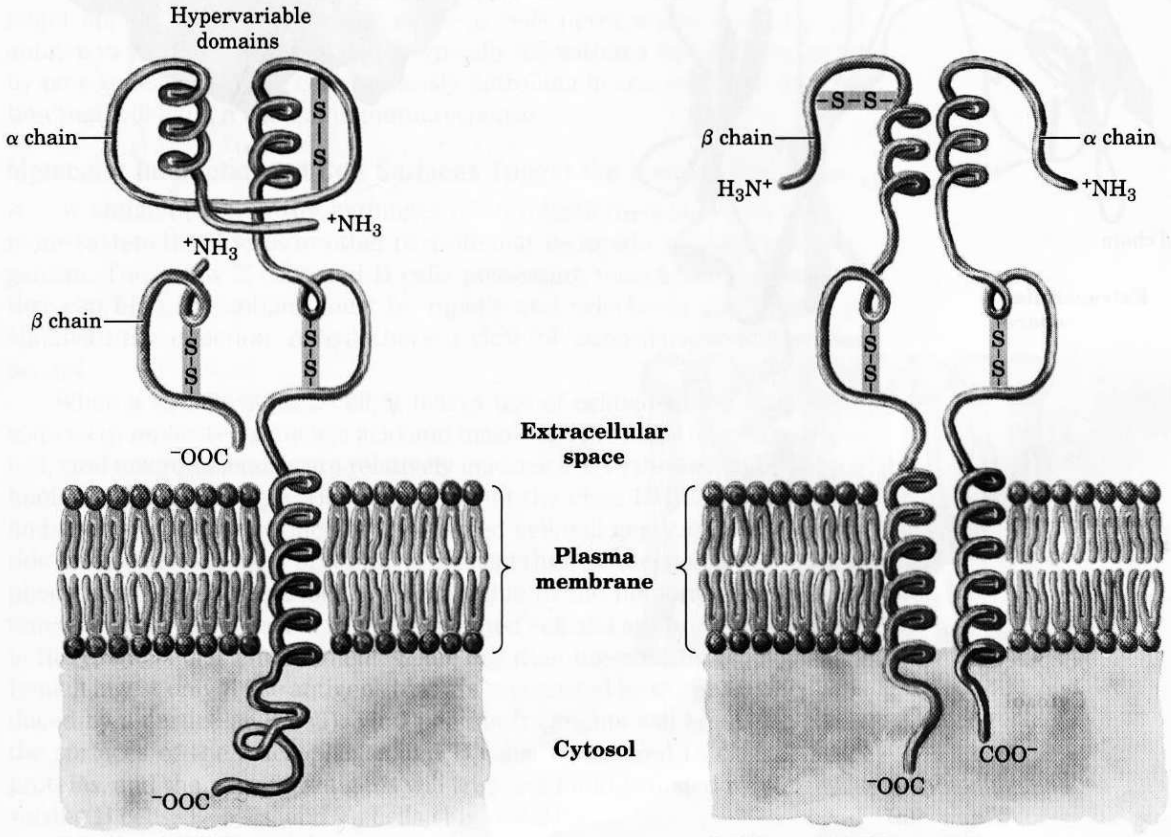
form they may elicit an immune response. These small molecules are called **haptens.** The antibodies produced in response to protein-linked haptens will then bind to the same small molecules when they are free. Such antibodies are sometimes used in the development of analytical tests described later in this chapter or as catalytic antibodies (described in Box 8–3).

The interactions of antibody and antigen are much better understood than are the binding properties of T-cell receptors. However, before focusing on antibodies, we need to look at the humoral and cellular immune systems in more detail to put the fundamental biochemical interactions into their proper context.

Self Is Distinguished from Nonself by the Display of Peptides on Cell Surfaces

The immune system must identify and destroy pathogens, but it must also recognize and *not* destroy the normal proteins and cells of the host organism—the "self." Detection of protein antigens in the host is mediated by **MHC (major histocompatibility complex) proteins.** MHC proteins bind peptide fragments of proteins digested in the cell and present them on the outside surface of the cell. These peptides normally come from the digestion of typical cellular proteins, but during a viral infection viral proteins are also digested and presented by MHC proteins. Peptide fragments from foreign proteins that are displayed by MHC proteins are the antigens the immune system recognizes as nonself. T-cell receptors bind these fragments and launch the subsequent steps of the immune response. There are two classes of MHC proteins (Fig. 7–20), which differ in their distribution among cell types and in the source of digested proteins whose peptides they display.

figure 7–20
MHC proteins These proteins consist of α and β chains. In class I MHC proteins **(a),** the small β chain is invariant but the amino acid sequence of the α chain exhibits a high degree of variability, localized in specific domains of the protein that appear on the outside of the cell. Each human produces up to six different α chains for class 1 MHC proteins. In class II MHC proteins **(b),** both the α and β chains have regions of relatively high variability near their amino-terminal ends.



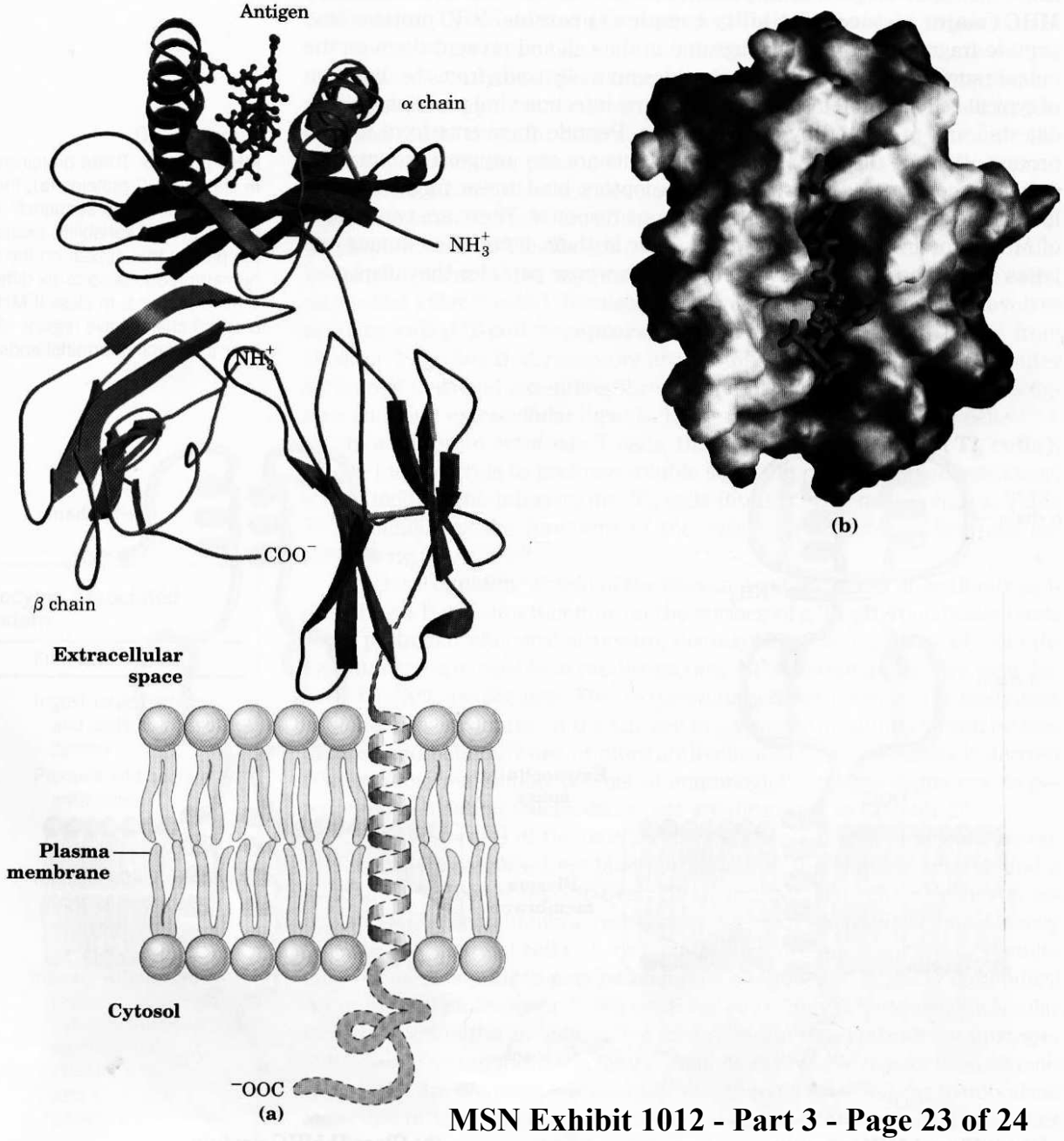
(a) Class I MHC protein **(b) Class II MHC protein**

Class I MHC proteins (Fig. 7–21) are found on the surface of virtually all vertebrate cells. There are countless variants in the human population, placing them among the most polymorphic of proteins. Because individuals produce up to six class I MHC protein variants, any two individuals are unlikely to have the same set. Class I MHC proteins bind and display peptides derived from the proteolytic degradation and turnover of proteins that occurs randomly within the cell. These complexes of peptides and class I MHC proteins are the recognition targets of the T-cell receptors of the T_C cells in the cellular immune system. The general pattern of immune system recognition was first described by Rolf Zinkernagel and Peter Doherty in 1974.

Each T_C cell has many copies of only one T-cell receptor that is specific for a particular class I MHC protein–peptide complex. To avoid creating a

figure 7–21

Structure of a human class I MHC protein. (a) This image is derived in part from the determined structure of the extracellular portion of the protein. The α chain of MHC is shown in gray; the small β chain is blue; the disulfide bonds are yellow. A bound ligand, a peptide derived from HIV, is shown in red. **(b)** Top view showing a surface contour image of the site where peptides are bound and displayed. The HIV peptide (red) occupies the site. This part of the class I MHC protein interacts with T-cell receptors.



legion of T_C cells that would set upon and destroy normal cells, the maturation of T_C cells in the thymus includes a stringent selection process that eliminates more than 95% of the developing T_C cells, including those that might recognize and bind class I MHC proteins displaying peptides from cellular proteins of the organism itself. The T_C cells that survive and mature are those with T-cell receptors that do not bind to the organism's own proteins. The result is a population of cells that bind foreign peptides bound to class I MHC proteins of the host cell. These binding interactions lead to the destruction of parasites and virus-infected cells. When an organ is transplanted, its foreign class I MHC proteins are also bound by T_C cells, leading to tissue rejection.

Class II MHC proteins occur on the surfaces of a few types of specialized cells that take up foreign antigens, including macrophages and B lymphocytes. Like class I MHC proteins, the class II proteins are highly polymorphic, with many variants in the human population. Each human is capable of producing up to 12 variants, and thus it is unlikely that any two individuals have an identical set of variants. The class II MHC proteins bind and display peptides derived not from cellular proteins but from external proteins ingested by the cells. The resulting class II MHC protein–peptide complexes are the binding targets of the T-cell receptors of the various helper T cells. T_H cells, like T_C cells, undergo a stringent selection process in the thymus, eliminating those that recognize the individual's own cellular proteins.

Despite the elimination of most T_C and T_H cells during the selection process in the thymus, a very large number survive, and these provide the immune response. Each survivor has a single type of T-cell receptor that can bind to one particular chemical structure. The T cells patrolling the bloodstream and the tissues carry millions of different binding specificities in the T-cell receptors. Within the highly varied T-cell population there is almost always a contingent of cells that can specifically bind any antigen that might appear. The vast majority of these cells never encounter a foreign antigen to which they can bind and typically die within a few days, replaced by new generations of T cells endlessly patrolling in search of the interaction that will launch the full immune response.

Molecular Interactions at Cell Surfaces Trigger the Immune Response

A new antigen is often the harbinger of an infection—a signal to the immune system that a virus or other parasite may be rapidly growing in the organism. Those few T cells and B cells possessing receptors or antibodies that can bind the antigen must be rapidly and selectively propagated to eliminate the infection. A hypothetical viral infection illustrates how this occurs.

When a virus invades a cell, it makes use of cellular functions and resources to replicate its nucleic acid and make viral proteins. Once inside the cell, viral macromolecules are relatively inaccessible to the antibodies of the humoral immune system. However, some of the class I MHC proteins that find their way to the surface of an infected cell will generally display peptide fragments from viral proteins, which can then be recognized by T_C lymphocytes. Mature viruses become vulnerable to the humoral immune system when they are released from the infected cell and are present for a time in the extracellular environment. Some are then ingested by macrophages (which ingest only those antigens that are recognized by the antibodies produced by a particular B cell). Viral peptide fragments will be displayed on the surfaces of the macrophages and B cells, complexed to class II MHC proteins, and the peptide antigens will trigger a multi-pronged response involving B cells, T_C cells, and T_H cells (Fig. 7–22).