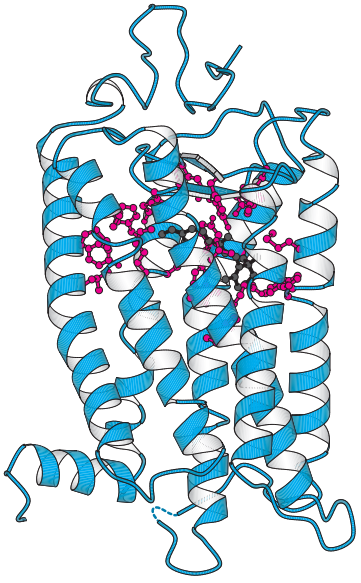
Sensory Systems



Color perception requires specific photoreceptors. The photoreceptor rhodopsin (right), which absorbs light in the process of vision, consists of the protein opsin and a bound vitamin A derivative, retinal. The amino acids (shown in red) that surround the retinal determine the color of light that is most efficiently absorbed. A person lacking a light-absorbing photoreceptor for the color green will see a colorful fruit stand (left) as mostly yellows (middle). [(Left and middle) From L. T. Sharpe et al. In *Color Vision: from Genes to Perception,* K. Gegenfurtner and L. T. Sharpe, Eds. (Cambridge University Press, 1999), pp. 3–51.]

Our senses provide us with means for detecting a diverse set of external signals, often with incredible sensitivity and specificity. For example, when fully adapted to a darkened room, our eyes allow us to sense very low levels of light, down to a limit of less than 10 photons. With more light, we are able to distinguish millions of colors. Through our senses of smell and taste, we are able to detect thousands of chemicals in our environment and sort them into categories: pleasant or unpleasant? healthful or toxic? Finally, we can perceive mechanical stimuli in the air and around us through our senses of hearing and touch.

How do our sensory systems work? How are the initial stimuli detected? How are these initial biochemical events transformed into perceptions and experiences? We have already encountered systems that sense and respond to chemical signals—namely, receptors that bind to growth factors and hormones. Our knowledge of these receptors and their associated signal-transduction pathways provides us with concepts and tools for unraveling some of the workings of sensory systems. For example, 7TM receptors (seven-transmembrane receptors; Section 14.1) play key roles in olfaction, taste, and vision. Ion channels that are sensitive to mechanical stress are essential for hearing and touch.

In this chapter, we focus on the five major sensory systems found in human beings and other mammals: olfaction (the sense of smell—i.e., the detection of small molecules in the air), taste, or gustation (the detection of selected organic compounds and ions by the tongue), vision (the detection

OUTLINE

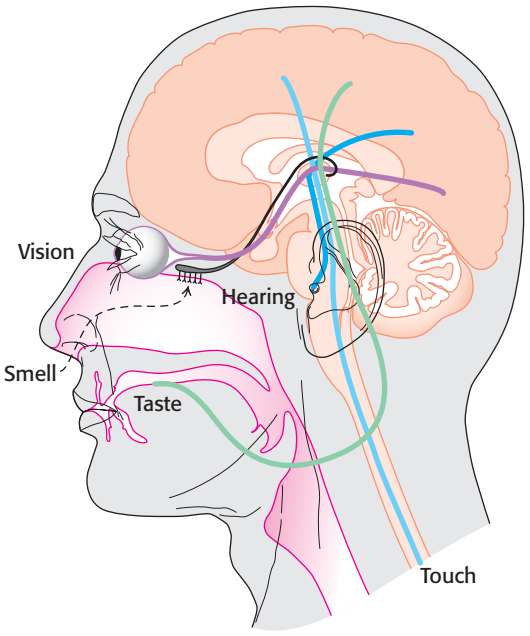
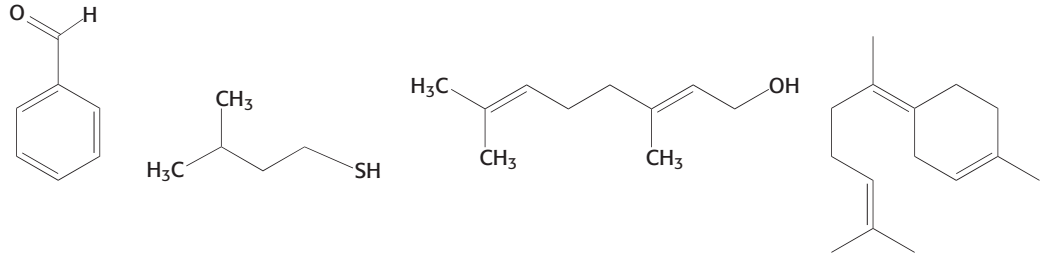


Figure 33.1 Sensory connections to the brain.
Sensory nerves connect sensory organs to the brain and spinal cord.

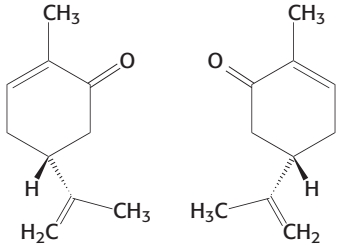
of light), hearing (the detection of sound, or pressure waves in the air), and touch (the detection of changes in pressure, temperature, and other factors by the skin). Each of these primary sensory systems contains specialized sensory neurons that transmit nerve impulses to the central nervous system (Figure 33.1). In the central nervous system, these signals are processed and combined with other information to yield a perception that may trigger a change in behavior. By these means, our senses allow us to detect changes in our environments and to adjust our behavior appropriately.

33.1 A Wide Variety of Organic Compounds Are Detected by Olfaction

Human beings can detect and distinguish thousands of different compounds by smell, often with considerable sensitivity and specificity. Most odorants are small organic compounds with sufficient volatility that they can be carried as vapors into the nose. For example, a major component responsible for the odor of almonds is the simple aromatic compound benzaldehyde, whereas the sulfhydryl compound 3-methylbutane-1-thiol is a major component of the odor of skunks.



Benzaldehyde (Almond) | **3-Methylbutane-1-thiol (Skunk)** | **Geraniol (Rose)** | **Zingiberene (Ginger)**



***R*-Carvone (Spearmint)** | ***S*-Carvone (Caraway)**

What properties of these molecules are responsible for their odors? First, *the shape of the molecule rather than its other physical properties is crucial.* We can most clearly see the importance of shape by comparing molecules such as those responsible for the odors of spearmint and caraway. These compounds are identical in essentially all physical properties such as hydrophobicity because they are exact mirror images of one another. Thus, the odor produced by an odorant depends not on a physical property but on the compound's interaction with a specific binding surface, most likely a protein receptor. Second, some human beings (and other animals) suffer from *specific anosmias;* that is, they are incapable of smelling specific compounds even though their olfactory systems are otherwise normal. Such anosmias are often inherited. These observations suggest that mutations in individual receptor genes lead to the loss of the ability to detect a small subset of compounds.

Olfaction is mediated by an enormous family of seven-transmembrane-helix receptors

Odorants are detected in a specific region of the nose, called the *main olfactory epithelium,* that lies at the top of the nasal cavity (Figure 33.2). Approximately 1 million sensory neurons line the surface of this region. Cilia containing the odorant-binding protein receptors project from these neurons into the mucous lining of the nasal cavity.

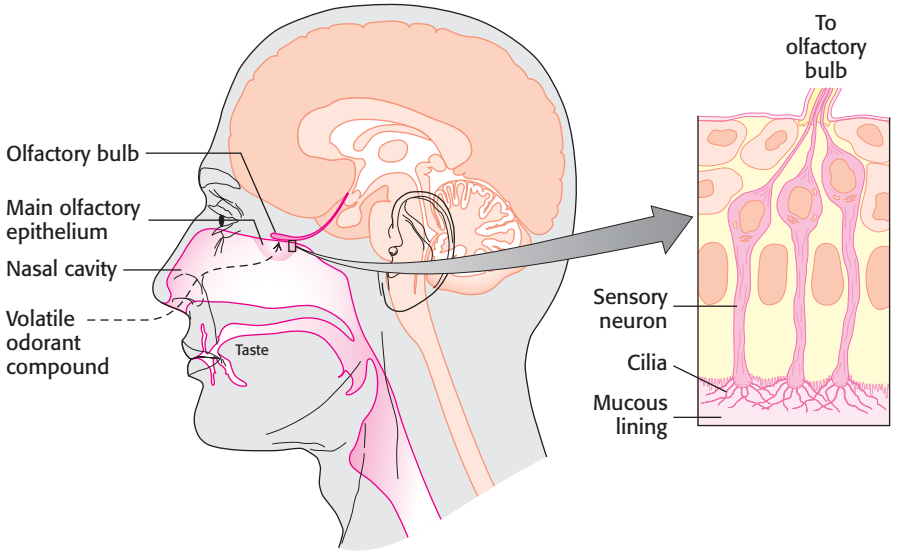


Figure 33.2 The main nasal epithelium. This region of the nose, which lies at the top of the nasal cavity, contains approximately 1 million sensory neurons. Nerve impulses generated by odorant molecules binding to receptors on the cilia travel from the sensory neurons to the olfactory bulb.

Biochemical studies in the late 1980s examined isolated cilia from rat olfactory epithelium that had been treated with odorants. Exposure to the odorants increased the cellular level of cyclic AMP, and this increase was observed only in the presence of GTP. On the basis of what was known about signal-transduction systems, *the participation of cAMP and GTP strongly suggested the involvement of a G protein and, hence, 7TM receptors.* Indeed, Randall Reed purified and cloned a G-protein α subunit, termed $G_{(olf)}$, which is uniquely expressed in olfactory cilia. The involvement of 7TM receptors suggested a strategy for identifying the olfactory receptors themselves. Complementary DNAs were sought that (1) were expressed primarily in the sensory neurons lining the nasal epithelium, (2) encoded members of the 7TM-receptor family, and (3) were present as a large and diverse family to account for the range of odorants. Through the use of these criteria, cDNAs for odorant receptors from rats were identified in 1991 by Richard Axel and Linda Buck.

The odorant receptor (hereafter, OR) family is even larger than expected: *more than 1000 OR genes are present in the mouse and the rat, whereas the human genome encodes approximately 380 ORs.* In addition, the human genome includes approximately 500 OR pseudogenes containing mutations that prevent the generation of a full-length, proper odorant receptor. The OR family is thus one of the largest gene families in human beings. Further analysis of primate OR genes reveals that the fraction of pseudogenes is greater in species more closely related to human beings (Figure 33.3). Thus, we may have a glimpse at the evolutionary loss of acuity in the sense of smell as higher mammals presumably became less dependent on this sense for survival. For rodents that are highly dependent on their sense of smell, essentially all OR genes encode functional proteins.

The OR proteins are typically 20% identical in sequence with the β-adrenergic receptor (Section 14.1) and from 30% to 60% identical with one another. Several specific sequence features are present in most or all OR family members (Figure 33.4). The central region,

Figure 33.3 Evolution of odorant receptors. Odorant receptors appear to have lost function through conversion into pseudogenes in the course of primate evolution. The percentage of OR genes that appear to be functional for each species is given in parentheses.

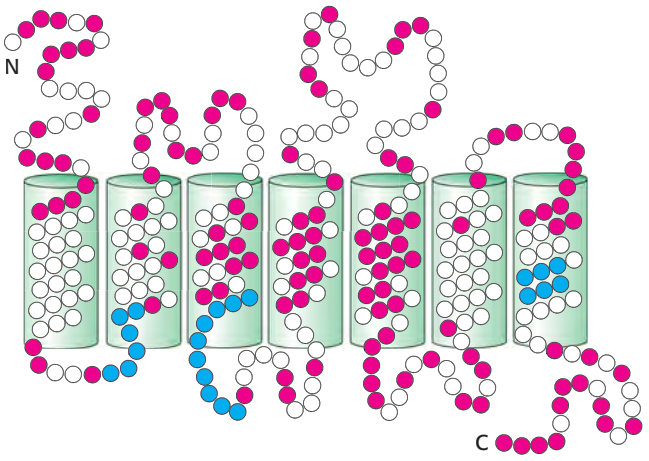


Figure 33.4 Conserved and variant regions in odorant receptors. Odorant receptors are members of the 7TM-receptor family. The green cylinders represent the seven presumed transmembrane helices. Strongly conserved residues characteristic of this protein family are shown in blue, whereas highly variable residues are shown in red.

particularly transmembrane helices 4 and 5, is highly variable, suggesting that this region is the site of odorant binding. That site must be different in odorant receptors that bind distinct odorant molecules.

What is the relation between OR gene expression and the individual neuron? Interestingly, *each olfactory neuron expresses only a single OR gene,* among hundreds available. Apparently, the precise OR gene expressed is determined largely at random. After one OR gene is expressed and a functional OR protein is produced, the expression of all other OR genes is suppressed by a feedback mechanism that remains to be fully elucidated.

The binding of an odorant to an OR on the neuronal surface initiates a signal-transduction cascade that results in an action potential (Figure 33.5). The ligand-bound OR activates $G_{(olf)}$, the specific G protein mentioned earlier. $G_{(olf)}$ is initially in its GDP-bound form. When activated, it releases GDP, binds GTP, and releases its associated βγ subunits. The α subunit then activates a specific adenylate cyclase, increasing the intracellular concentration of cAMP. The rise in the intracellular concentration of cAMP activates a nonspecific cation channel that allows calcium and other cations into the cell. The flow of cations through the channel depolarizes the neuronal membrane and initiates an action potential. This action potential, combined with those from other olfactory neurons, leads to the perception of a specific odor.

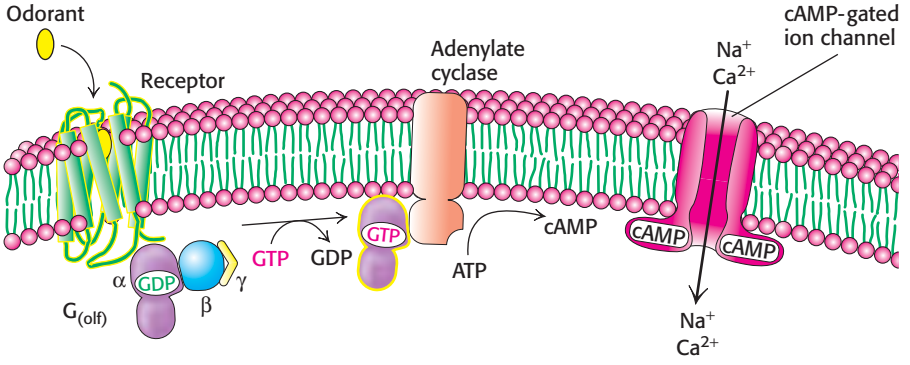


Figure 33.5 The olfactory signal-transduction cascade. The binding of odorant to the olfactory receptor activates a signaling pathway similar to those initiated in response to the binding of some hormones to their receptors. The final result is the opening of cAMP-gated ion channels and the initiation of an action potential.

Odorants are decoded by a combinatorial mechanism

An obvious challenge presented to an investigator by the large size of the OR family is to match each OR with the one or more odorant molecules to which it binds. Exciting progress has been made in this regard. Initially, an OR was matched with odorants by overexpressing a single, specific OR gene in rats. This OR responded to straight-chain aldehydes, most favorably to *n*-octanal and less strongly to *n*-heptanal and *n*-hexanal. More-dramatic progress was made by taking advantage of our knowledge of the OR signal-transduction pathway and the power of the polymerase chain reaction (PCR). A section of nasal epithelium from a mouse was loaded with the calcium-sensitive dye Fura-2 (Section 14.1). The tissue was then treated with different odorants, one at a time, at a specific concentration. If the

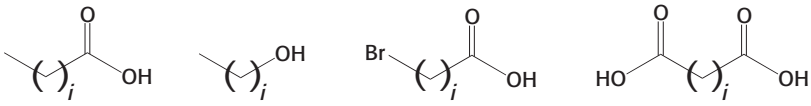


Figure 33.6 Four series of odorants tested for olfactory-receptor activation.

odorant had bound to an OR and activated it, that neuron could be detected under a microscope by the change in fluorescence caused by the influx of calcium that takes place as part of the signal-transduction process. To determine which OR was responsible for the response, cDNA was generated from mRNA that had been isolated from single identified neurons. The cDNA was then subjected to PCR with the use of primers that are effective in amplifying most or all OR genes. The sequence of the PCR product from each neuron was then determined and analyzed.

Using this approach, investigators analyzed the responses of neurons to a series of compounds having varying chain lengths and terminal functional groups (Figure 33.6). The results of these experiments appear surprising at first glance (Figure 33.7). Importantly, there is not a simple 1 : 1 correspondence between odorants and receptors. *Almost every odorant activates a number of receptors* (usually to different extents) and *almost every receptor is activated by more than one odorant*. Note, however, that each odorant activates a unique combination of receptors. In principle, this combinatorial mechanism allows even a small array of receptors to distinguish a vast number of odorants.

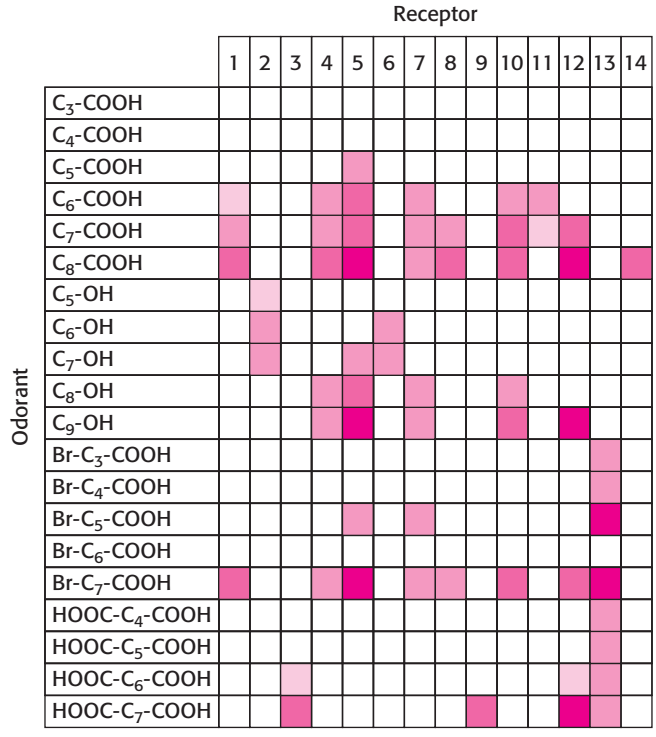


Figure 33.7 Patterns of olfactory-receptor activation. Fourteen different receptors were tested for responsiveness to the compounds shown in Figure 33.6. A colored box indicates that the receptor at the top responded to the compound at the left. Darker colors indicate that the receptor was activated at a lower concentration of odorant.

How is the information about which receptors have been activated transmitted to the brain? Recall that each neuron expresses only one OR and that the pattern of expression appears to be largely random. A substantial clue to the connections between receptors and the brain has been provided by the creation of mice that express a gene for an easily detectable colored marker in conjunction with a specific OR gene. Olfactory neurons that express the OR–marker-protein combination were traced to their destination in the brain, a structure called the olfactory bulb (Figure 33.8). The processes from neurons that express the same OR gene were found to connect to the same location in the olfactory bulb. Moreover, this pattern of neuronal connection was found to be identical in all mice examined. Thus, *neurons that express specific ORs are linked to specific sites in the brain*. This property creates a spatial map of odorant-responsive neuronal activity within the olfactory bulb.

Can such a combinatorial mechanism truly distinguish many different odorants? An electronic "nose" that functions by the same principles pro-

Figure 33.8 Converging olfactory neurons. This section of the nasal cavity is stained to reveal processes from sensory neurons expressing the same olfactory receptor. The processes converge to a single location in the olfactory bulb. [From P. Mombaerts, et al. *Cell* 87:675–689, 1996.]

vides compelling evidence that it can (Figure 33.9). The receptors for the electronic nose are polymers that bind a range of small molecules. Each polymer binds every odorant, but to varying degrees. Importantly, the electrical properties of these polymers change on odorant binding. A set of 32 of these polymer sensors, wired together so that the pattern of responses can be evaluated, is capable of distinguishing individual compounds such as *n*-pentane and *n*-hexane as well as complex mixtures such as the odors of fresh and spoiled fruit.

Figure 33.9 The Cyranose 320. The electronic nose may find uses in the food industry, animal husbandry, law enforcement, and medicine. [Courtesy of Cyrano Sciences.]

33.2 Taste Is a Combination of Senses That Function by Different Mechanisms

The inability to taste food is a common complaint when nasal congestion reduces the sense of smell. Thus, smell greatly augments our sense of taste (also known as *gustation*), and taste is, in many ways, the sister sense to olfaction. Nevertheless, the two senses differ from each other in several important ways. First, we are able to sense several classes of compounds by taste that we are unable to detect by smell; salt and sugar have very little odor, yet they are primary stimuli of the gustatory system. Second, whereas we are able to discriminate thousands of odorants, discrimination by taste is much more modest. Five primary tastes are perceived: *bitter, sweet, sour, salty,* and *umami* (the taste of glutamate and aspartate from the Japanese word for "deliciousness"). These five tastes serve to classify compounds into potentially nutritive and beneficial (sweet, salty, umami) or potentially harmful or toxic (bitter, sour). Tastants (the molecules sensed by taste) are quite distinct for the different groups (Figure 33.10).

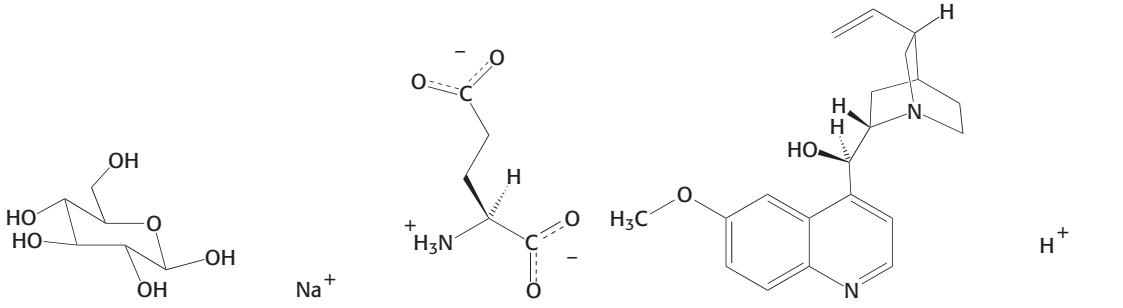
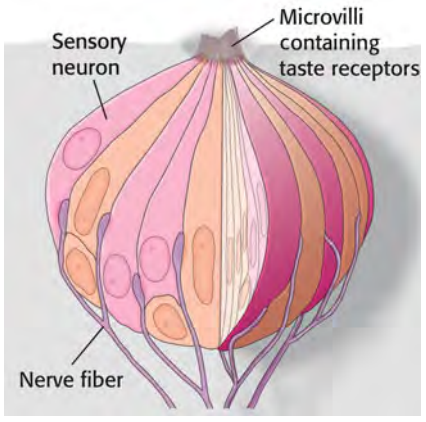


Figure 33.10 Examples of tastant molecules. Tastants fall into five groups: sweet, salty, umami, bitter, and sour.



The simplest tastant, the hydrogen ion, is perceived as sour. Other simple ions, particularly sodium ion, are perceived as salty. The taste called umami is evoked by the amino acids glutamate and aspartate, the former often encountered as the flavor enhancer monosodium glutamate (MSG). In contrast, *tastants perceived as sweet and, particularly, bitter are extremely diverse*. Many bitter compounds are alkaloids or other plant products of which many are toxic. However, they do not have any common structural elements or other common properties. Carbohydrates such as glucose and sucrose are perceived as sweet, as are other compounds including some simple peptide derivatives, such as aspartame, and even some proteins.

Figure 33.11 A taste bud. Each taste bud contains sensory neurons that extend microvilli to the surface of the tongue, where they interact with tastants.

These differences in specificity among the five tastes are due to differences in their underlying biochemical mechanisms. The sense of taste is, in fact, a number of independent senses all utilizing the same organ, the tongue, for their expression.

Tastants are detected by specialized structures called *taste buds,* which contain approximately 150 cells, including sensory neurons (Figure 33.11). Fingerlike projections called *microvilli,* which are rich in taste receptors, project from one end of each sensory neuron to the surface of the tongue. Nerve fibers at the opposite end of each neuron carry electrical impulses to the brain in response to stimulation by tastants. Structures called *taste papillae* contain numerous taste buds.

Sequencing of the human genome led to the discovery of a large family of 7TM bitter receptors

Just as in olfaction, a number of clues pointed to the involvement of G proteins and, hence, 7TM receptors in the detection of bitter and sweet tastes. The evidence included the isolation of a specific G-protein a subunit termed *gustducin,* which is expressed primarily in taste buds (Figure 33.12). How could the 7TM receptors be identified? The ability to detect some compounds depends on specific genetic loci in both human beings and mice. For instance, the ability to taste the bitter compound 6-*n*-propyl-2-thiouracil (PROP) was mapped to a region on human chromosome 5 by comparing DNA markers of persons who vary in sensitivity to this compound.

This observation suggested that this region might encode a 7TM receptor that responded to PROP. Approximately 450 kilobases in this region had been sequenced early in the human genome project. This sequence was searched by computer for potential 7TM-receptor genes, and, indeed, one was detected and named *T2R1*. Additional database searches detected approximately 30 sequences similar to *T2R1* in the human genome. The encoded proteins are between 30 and 70% identical with T2R1 (Figure 33.13).

Are these proteins, in fact, bitter receptors? Several lines of evidence suggest that they are. First, their genes are expressed in taste-sensitive cells—in fact, in many of the same cells that express gustducin. Second, cells that express individual members of this family respond to specific bitter compounds. For example, cells that express a specific mouse receptor (mT2R5) responded when exposed specifically to cycloheximide. Third, mice that had been found unresponsive to cycloheximide were found to have point mutations in the gene encoding mT2R5. Finally, cycloheximide

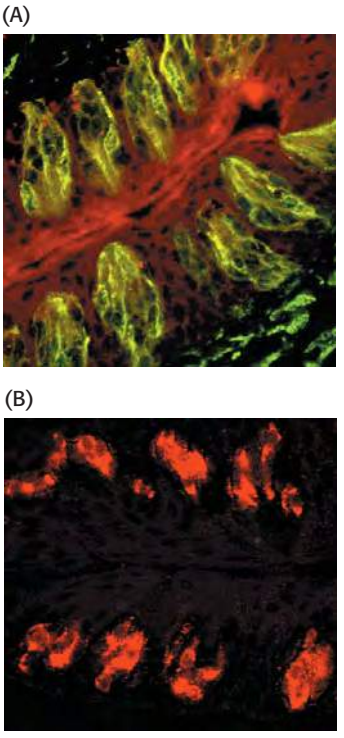
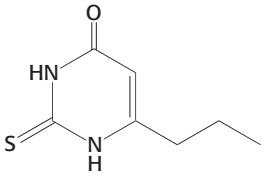


Figure 33.12 Expression of gustducin in the tongue. (A) A section of tongue stained with a fluorescent antibody reveals the position of the taste buds.(B) The same region stained with an antibody directed against gustducin reveals that this G protein is expressed in taste buds. [Courtesy of Dr. Charles S. Zuker.]



6-*n*-Propyl-2-thiouracil (PROP)

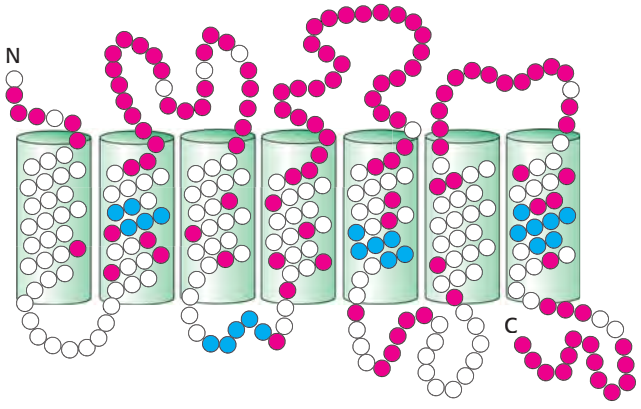


Figure 33.13 Conserved and variant regions in bitter receptors. The bitter receptors are members of the 7TM-receptor family. Strongly conserved residues characteristic of this protein family are shown in blue, and highly variable residues are shown in red.

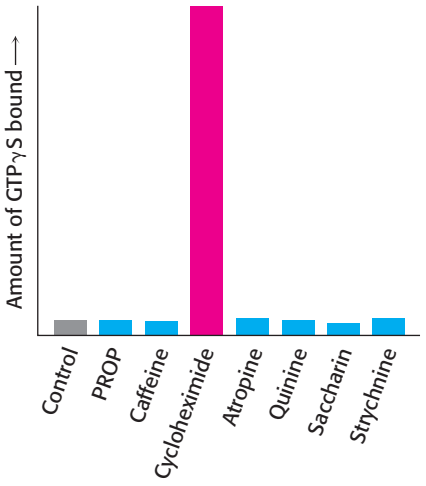


Figure 33.14 Evidence that T2R proteins are bitter taste receptors. Cycloheximide uniquely stimulates the binding of the GTP analog GTPγS to gustducin in the presence of the mT2R protein. [After J. Chandrashekar, et al. *Cell* 100:703–711, 2000.]

specifically stimulates the binding of GTP analogs to gustducin in the presence of the mT2R5 protein (Figure 33.14).

Importantly, each taste-receptor cell expresses many different members of the T2R family. This pattern of expression stands in sharp contrast to the pattern of one receptor type per cell that characterizes the olfactory system (Figure 33.15). The difference in expression patterns accounts for the much greater specificity of our perceptions of odors compared with tastes. *We are able to distinguish among subtly different odors because each odorant stimulates a unique pattern of neurons. In contrast, many tastants stimulate the same neurons*. Thus, we perceive only "bitter" without the ability to discriminate cycloheximide from quinine.

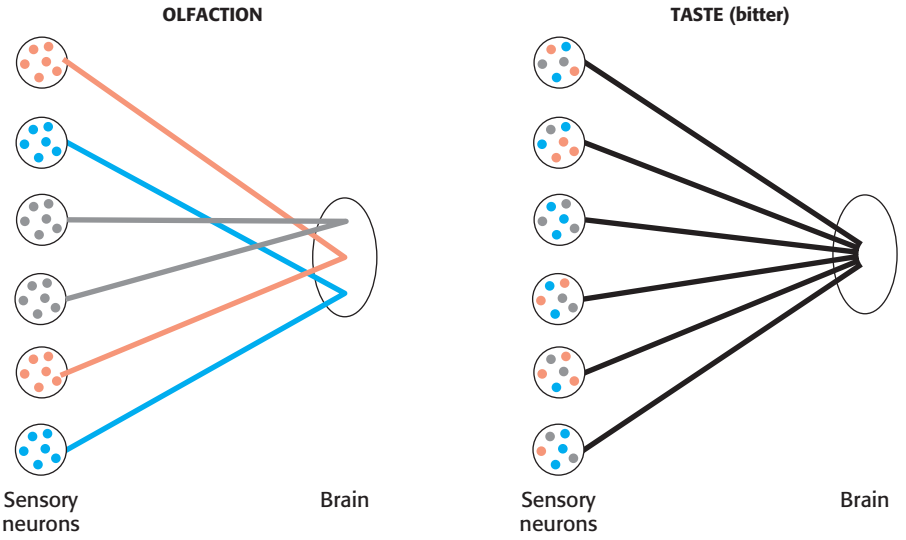


Figure 33.15 Differing gene-expression and connection patterns in olfactory and bitter taste receptors. In olfaction, each neuron expresses a single OR gene, and the neurons expressing the same OR converge to specific sites in the brain, enabling specific perception of different odorants. In gustation, each neuron expresses many bitter receptor genes, and so the identity of the tastant is lost in transmission.

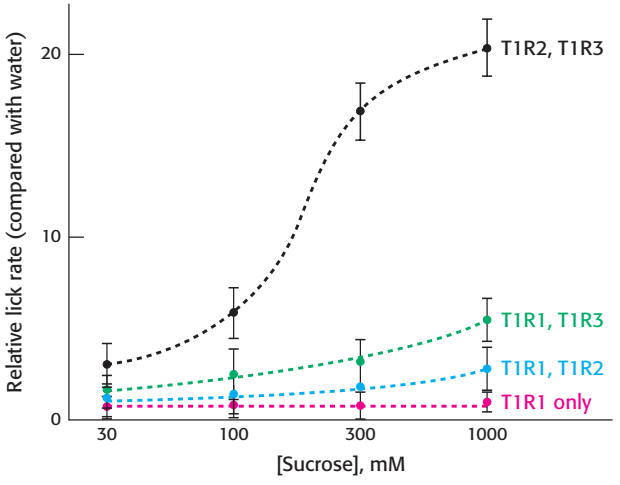


Figure 33.16 Evidence for a heterodimeric sweet receptor. The sensitivity to sweetness of mice with genes for either T1R1, T1R2, T1R3, or both T1R2 and T1R3 was determined by observing the relative rates at which they licked solutions containing various amount of sucrose. These studies revealed that both T1R2 and T1R3 were required for a full response to sucrose. Mice with a disrupted T1R1 gene were indistinguishable from wild-type mice in this assay (not shown). [After G. Q. Zhao, et al. *Cell* 115:255–266, 2003.]

A heterodimeric 7TM receptor responds to sweet compounds

Most sweet compounds are carbohydrates, energy rich and easily digestible. Some noncarbohydrate compounds such as saccharin and aspartame also taste sweet. Members of a second family of 7TM receptors are expressed in taste-receptor cells sensitive to sweetness. The three members of this family, referred to as T1R1, T1R2, and T1R3, are distinguished by their large extracellular domains compared with those of the bitter receptors. Studies in knockout mice have revealed that T1R2 and T1R3 are expressed simultaneously in mice able to taste carbohydrates (Figure 33.16). Thus, T1R2 and T1R3 appear to form a specific heterodimeric receptor responsible for mediating the response to sugars. This heterodimeric receptor also responds to artificial sweeteners and to sweet-tasting proteins and therefore appears to be the receptor responsible for responses to all sweet tastants. Note that T1R2 and T1R3 do respond to sweet tastants individually but only at very high concentrations of tastant.

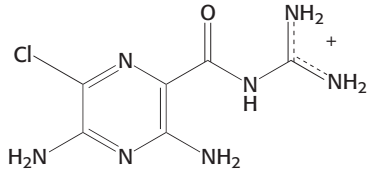
The requirement for an *oligomeric* 7TM receptor for a fully functional response is surprising, considering our previous understanding of 7TM receptors. This discovery has at least two possible explanations. First, the sweet receptor could be a member of a small subset of the 7TM-receptor family that functions well only as oligomers. Alternatively, many 7TM receptors may function as oligomers, but this notion is not clear, because these oligomers contain only one type of 7TM-receptor subunit. Further studies will be required to determine which of these explanations is correct.

Umami, the taste of glutamate and aspartate, is mediated by a heterodimeric receptor related to the sweet receptor

The family of receptors responsible for detecting sweetness is also responsible for detecting amino acids. In human beings, only glutamate and aspartate elicit a taste response. Studies similar to those for the sweet receptor revealed that the umami receptor consists of T1R1 and T1R3. Thus, this receptor has one subunit (T1R3) in common with the sweet receptor but has an additional subunit (T1R1) that does not participate in the sweet response. This observation is supported by the observation that mice in which the gene for T1R1 is disrupted do not respond to aspartate but do respond normally to sweet tastants; mice having disrupted genes for both T1R1 and T1R3 respond poorly to both umami and sweet tastants.

Salty tastes are detected primarily by the passage of sodium ions through channels

Salty tastants are not detected by 7TM receptors. Rather, they are detected directly by their passage through ion channels expressed on the surface of cells in the tongue. Evidence for the role of these ion channels comes from examining known properties of Na^+ channels characterized in other biological contexts. One class of channels, characterized first for its role in salt reabsorption, is thought to be important in the detection of salty tastes because these channels are sensitive to the compound *amiloride,* which mutes the taste of salt and significantly lowers sensory-neuron activation in response to sodium.



Amiloride

An *amiloride-sensitive Na^+ channel* comprises four subunits that may be either identical or distinct but in any case are homologous. An individual subunit ranges in length from 500 to 1000 amino acids and includes two presumed membrane-spanning helices as well as a large extracellular domain in between them (Figure 33.17). The extracellular region includes two (or, sometimes, three) distinct regions rich in cysteine residues (and, presumably, disulfide bonds). A region just ahead of the second membrane-spanning helix appears to form part of the pore in a manner analogous to that of the structurally characterized potassium channel. The members of the amiloride-sensitive Na^+ -channel family are numerous and diverse in their biological roles. We shall encounter them again in the context of the sense of touch.

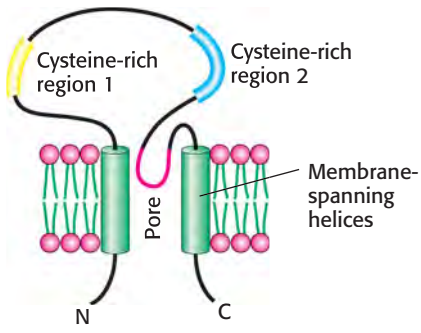


Figure 33.17 Schematic structure of the amiloride-sensitive sodium channel. Only one of the four subunits that constitute the functional channel is illustrated. The amiloride-sensitive sodium channel belongs to a superfamily having common structural features, including two hydrophobic membrane-spanning regions, intracellular amino and carboxyl termini; and a large, extracellular region with conserved cysteine-rich domains.

Sodium ions passing through these channels produce a significant transmembrane current. Amiloride blocks this current, accounting for its effect on taste. However, about 20% of the response to sodium remains even in the presence of amiloride, suggesting that other ion channels also contribute to salt detection.

Sour tastes arise from the effects of hydrogen ions (acids) on channels

Like salty tastes, *sour tastes are detected by direct interactions with ion channels,* but the incoming ions are hydrogen ions (in high concentrations) rather than sodium ions. For example, in the absence of high concentrations of

sodium, hydrogen ion flow can induce substantial transmembrane currents through amiloride-sensitive Na^+ channels. However, hydrogen ions are also sensed by mechanisms other than their direct passage through membranes. Binding by hydrogen ions blocks some potassium ion channels and activates other types of channels. Together, these mechanisms lead to changes in membrane polarization in sensory neurons that produce the sensation of sour taste. We shall consider an additional receptor related to taste, one responsible for the "hot" taste of spicy food, when we examine mechanisms of touch perception.

33.3 Photoreceptor Molecules in the Eye Detect Visible Light

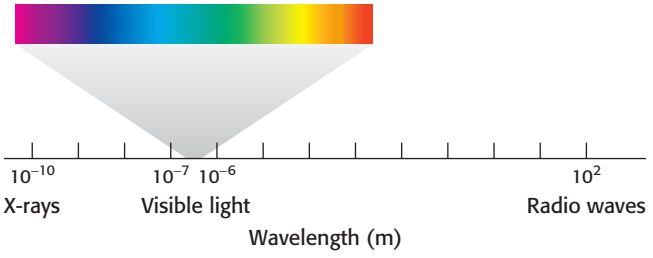
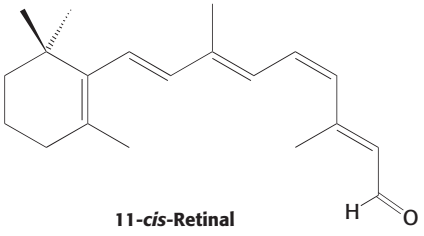


Figure 33.18 The electromagnetic spectrum. Visible light has wavelengths between 300 and 850 nm.

Vision is based on the absorption of light by photoreceptor cells in the eye. These cells are sensitive to light in a narrow region of the electromagnetic spectrum, the region with wavelengths between 300 and 850 nm (Figure 33.18). Vertebrates have two kinds of photoreceptor cells, called *rods* and *cones* because of their distinctive shapes. Cones function in bright light and are responsible for color vision, whereas rods function in dim light but do not perceive color. A human retina contains about 3 million cones and 100 million rods. Remarkably, a rod cell can respond to a single photon, and the brain requires fewer than 10 such responses to register the sensation of a flash of light.

Rhodopsin, a specialized 7TM receptor, absorbs visible light



11-*cis*-Retinal

Rods are slender, elongated structures; the outer segment is specialized for photoreception (Figure 33.19). It contains a stack of about 1000 discs, which are membrane-enclosed sacs densely packed with photoreceptor molecules. The photosensitive molecule is often called a *visual pigment* because it is highly colored owing to its ability to absorb light. The photoreceptor molecule in rods is *rhodopsin* (Section 14.1), which consists of the protein *opsin* linked to *11*-cis-*retinal*, a prosthetic group.

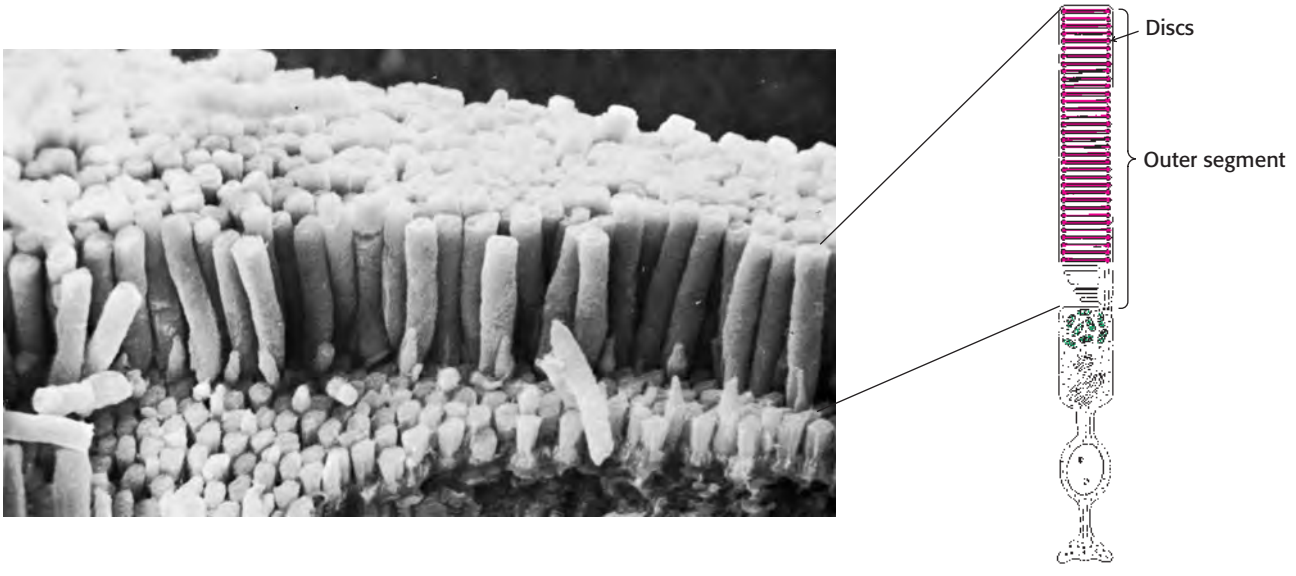


Figure 33.19 The rod cell. (Left) Scanning electron micrograph of retinal rod cells. (Right) Schematic representation of a rod cell. [Photograph courtesy of Dr. Deric Bownds.]

Rhodopsin absorbs light very efficiently in the middle of the visible spectrum, its absorption being centered on 500 nm, which nicely matches the solar output (Figure 33.20). A rhodopsin molecule will absorb a high percentage of the photons of the correct wavelength that strike it, as indicated by the extinction coefficient of 40,000 $M^{-1}cm^{-1}$ at 500 nm. The extinction coefficient for rhodopsin is more than an order of magnitude greater than that for tryptophan, the most efficient absorber in proteins that lack prosthetic groups.

Opsin, the protein component of rhodopsin, is a member of the 7TM-receptor family. Indeed, rhodopsin was the first member of this family to be purified, its gene was the first to be cloned and sequenced, and its three-dimensional structure was the first to be determined. The color of rhodopsin and its responsiveness to light depend on the presence of the light-absorbing group *(chromophore)* 11-*cis*-retinal. This compound is a powerful absorber of light because it is a polyene; its six alternating single and double bonds constitute a long, unsaturated electron network. Recall that alternating single and double bonds account for the chromophoric properties of chlorophyll (Section 19.2). The aldehyde group of 11-*cis*-retinal forms a Schiff base (Figure 33.21) with the ε-amino group of lysine residue 296, which lies in the center of the seventh transmembrane helix. Free retinal absorbs maximally at 370 nm, and its unprotonated Schiff-base adduct absorbs at 380 nm, whereas the protonated Schiff base absorbs at 440 nm or longer wavelengths. Thus, *the 500-nm absorption maximum for rhodopsin strongly suggests that the Schiff base is protonated;* additional interactions with opsin shift the absorption maximum farther toward the red. The positive charge of the protonated Schiff base is compensated by the negative charge of glutamate 113 located in helix 2; the glutamate residue closely approaches the lysine–retinal linkage in the three-dimensional structure of rhodopsin.

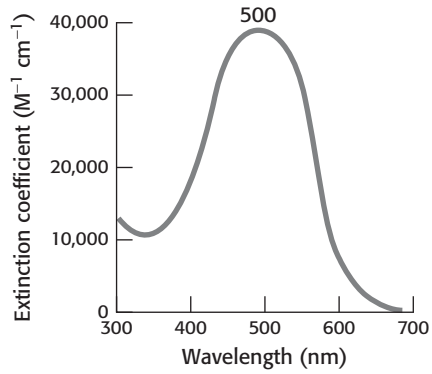
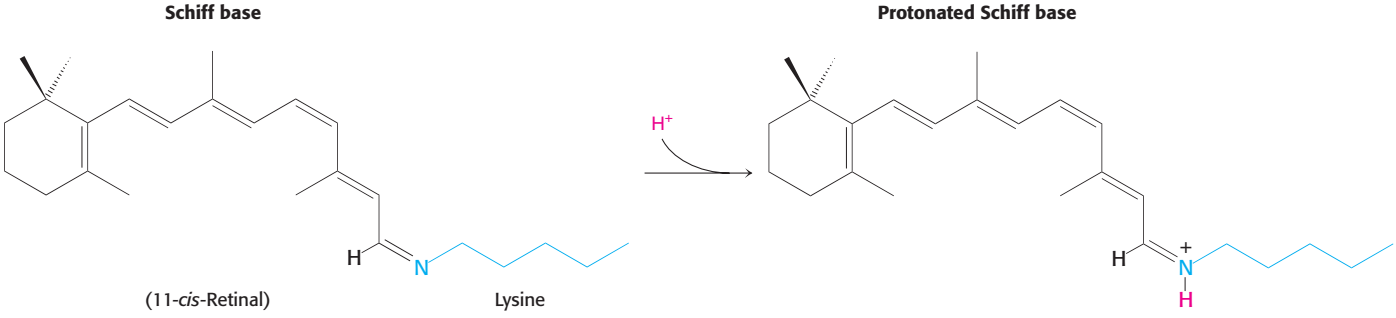


Figure 33.20 Rhodopsin absorption spectrum. Almost all photons with wavelengths near 500 nm that strike a rhodopsin molecule are absorbed.

Figure 33.21 Retinal–lysine linkage. Retinal is linked to lysine 296 in opsin by a Schiff-base linkage. In the resting state of rhodopsin, this Schiff base is protonated.



Light absorption induces a specific isomerization of bound 11-*cis*-retinal

How does the absorption of light by the retinal Schiff base generate a signal? George Wald and his coworkers discovered that *light absorption results in the isomerization of the 11-*cis*-retinal group of rhodopsin to its all-*trans *form* (Figure 33.22). This isomerization causes the Schiff-base nitrogen atom to move approximately 5 Å, assuming that the cyclohexane ring of the retinal group remains fixed. In essence, *the light energy of a photon is converted into atomic motion*. The change in atomic positions, like the binding of a ligand to other 7TM receptors, sets in train a series of events that lead to the closing of ion channels and the generation of a nerve impulse.

The isomerization of the retinal Schiff base takes place within a few picoseconds of a photon being absorbed. The initial product, termed *bathorhodopsin,* contains a strained all-*trans*-retinal group. Within approximately

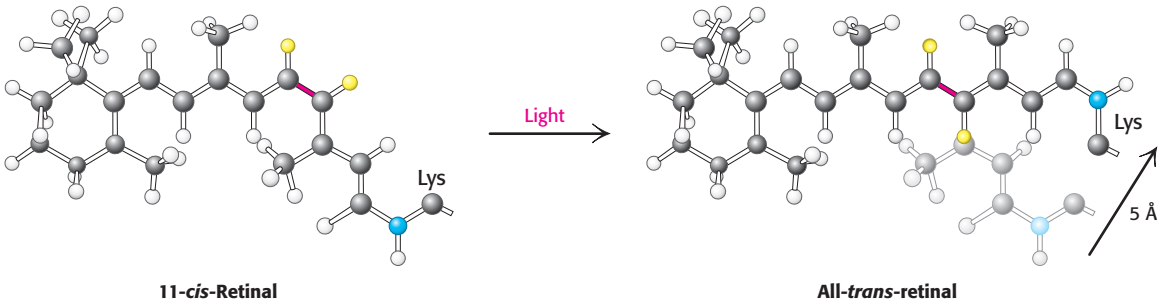


Figure 33.22 Atomic motion in retinal. The Schiff-base nitrogen atom moves 5 Å as a consequence of the light-induced isomerization of 11-*cis*-retinal to all-*trans*-retinal by rotation about the bond shown in red.

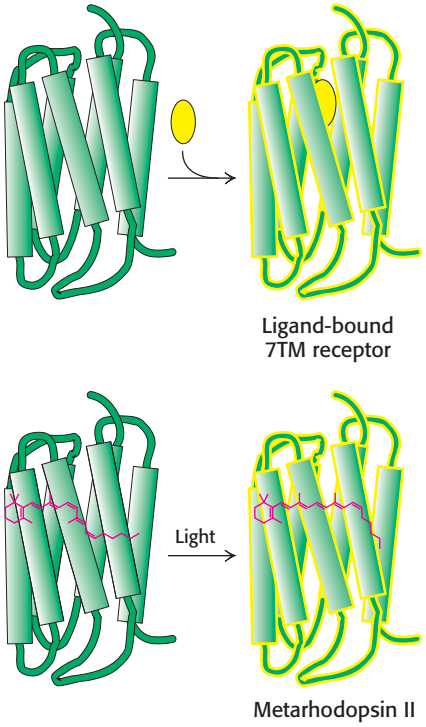


Figure 33.23 Analogous 7TM receptors. The conversion of rhodopsin into metarhodopsin II activates a signal-transduction pathway analogously to the activation induced by the binding of other 7TM receptors to appropriate ligands.

1 ms, this intermediate is converted through several additional intermediates into *metarhodopsin II*. In metarhodopsin II, the Schiff base is deprotonated and the opsin protein has undergone significant reorganization.

Metarhodopsin II (also referred to as R*) is analogous to the ligand-bound state of 7TM receptors such as the β_2-adrenergic receptor (Section 14.1) and the odorant and tastant receptors discussed previously (Figure 33.23). Like these receptors, this form of rhodopsin activates a heterotrimeric G protein that propagates the signal. The G protein associated with rhodopsin is called *transducin*. Metarhodopsin II triggers the exchange of GDP for GTP by the α subunit of transducin (Figure 33.24). On the binding of GTP, the βγ subunits of transducin are released and the α subunit switches on a *cGMP phosphodiesterase* by binding to an inhibitory subunit and removing it. The activated phosphodiesterase is a potent enzyme that rapidly hydrolyzes cGMP to GMP. The reduction in cGMP concentration causes *cGMP-gated ion channels* to close, leading to the hyperpolarization of the membrane and neuronal signaling. *At each step in this process, the initial signal—the absorption of a single photon—is amplified so that it leads to sufficient membrane hyperpolarization to result in signaling.*

Light-induced lowering of the calcium level coordinates recovery

As we have seen, the visual system responds to changes in light and color within a few milliseconds, quickly enough that we are able to perceive continuous motion at nearly 1000 frames per second. To achieve a rapid response, the signal must also be terminated rapidly and the system must be returned to its initial state. First, activated rhodopsin must be blocked from continuing to activate transducin. *Rhodopsin kinase* catalyzes the phosphorylation of the carboxyl terminus of R* at multiple serine and threonine residues.

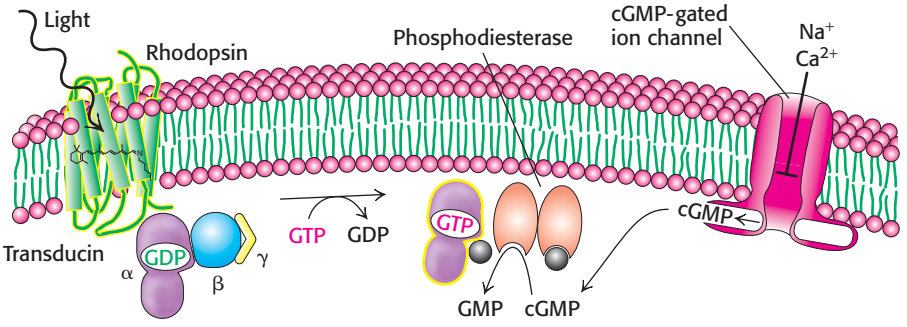


Figure 33.24 Visual signal transduction. The light-induced activation of rhodopsin leads to the hydrolysis of cGMP, which in turn leads to ion-channel closing and the initiation of an action potential.

Arrestin, an inhibitory protein then binds phosphorylated R* and prevents additional interaction with transducin.

Second, the α subunit of transducin must be returned to its inactive state to prevent further signaling. Like other G proteins, the α subunit possesses built-in GTPase activity that hydrolyzes bound GTP to GDP. Hydrolysis takes place in less than a second when transducin is bound to the phosphodiesterase. The GDP form of transducin then leaves the phosphodiesterase and reassociates with the βγ subunits, and the phosphodiesterase returns to its inactive state. Third, the level of cGMP must be raised to reopen the cGMP-gated ion channels. *The action of guanylate cyclase accomplishes this third step by synthesizing cGMP from GTP.*

Calcium ion plays an essential role in controlling guanylate cyclase because it markedly inhibits the activity of the enzyme. In the dark, Ca^{2+} as well as Na^+ enter the rod outer segment through the cGMP-gated channels. Calcium ion influx is balanced by its efflux through an exchanger, a transport system that uses the thermodynamically favorable flow of four Na^+ ions into the cell and one K^+ ion out of the cell to extrude one Ca^{2+} ion. After illumination, the entry of Ca^{2+} through the cGMP-gated channels stops, but its export through the exchanger continues. Thus, the cytoplasmic Ca^{2+} level drops from 500 nM to 50 nM after illumination. This drop markedly stimulates guanylate cyclase, rapidly restoring the concentration of cGMP to reopen the cGMP-gated channels.

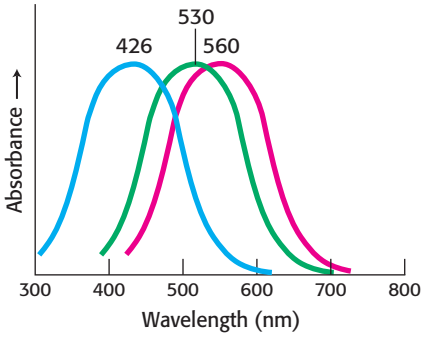
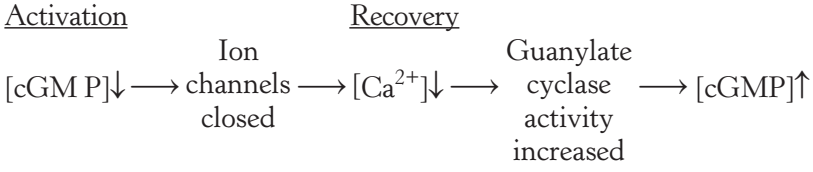


Figure 33.25 Cone-pigment absorption spectra. The absorption spectra of the cone visual pigment responsible for color vision.



By controlling the rate of cGMP synthesis, Ca^{2+} levels govern the speed with which the system is restored to its initial state.

Color vision is mediated by three cone receptors that are homologs of rhodopsin

Cone cells, like rod cells, contain visual pigments. Like rhodopsin, these photoreceptor proteins are members of the 7TM-receptor family and use 11-*cis*-retinal as their chromophore. In human cone cells, there are three distinct photoreceptor proteins with absorption maxima at 426, 530, and ~560 nm (Figure 33.25). *These absorbances correspond to (in fact, define) the blue, green, and red regions of the spectrum.* Recall that the absorption maximum for rhodopsin is 500 nm.

The amino acid sequences of the cone photoreceptors have been compared with one another and with rhodopsin. The result is striking. Each of the cone photoreceptors is approximately 40% identical in sequence with rhodopsin. Similarly, the blue photoreceptor is 40% identical with each of the green and red photoreceptors. The green and red photoreceptors, however, are >95% identical with each other, differing in only 15 of 364 positions (Figure 33.26).

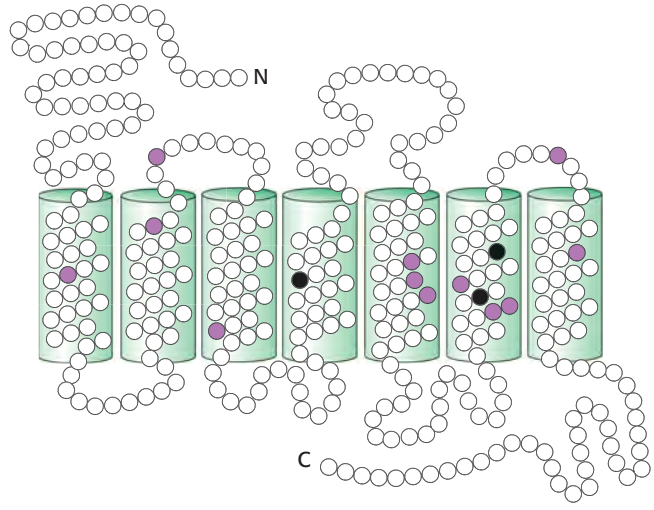


Figure 33.26 Comparison of the amino acid sequences of the green and red photoreceptors. Open circles correspond to identical residues, whereas colored circles mark residues that are different. The differences in the three black positions are responsible for most of the difference in their absorption spectra.

These observations are sources of insight into photoreceptor evolution. First, the green and red photore-

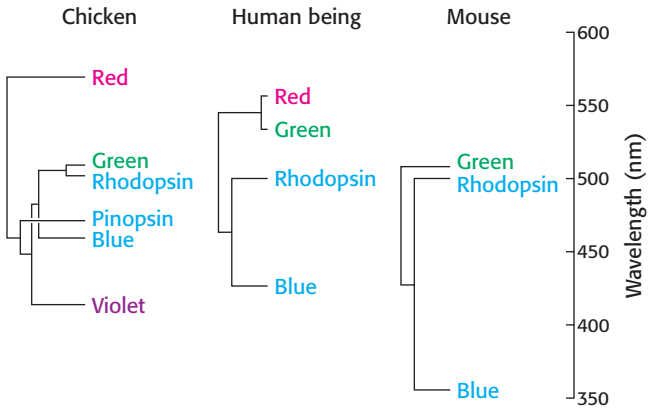


Figure 33.27 Evolutionary relationships among visual pigments. Visual pigments have evolved by gene duplication along different branches of the animal evolutionary tree. The branch lengths of the "trees" correspond to the percentage of amino acid divergence. [After J. Nathans. *Neuron* 24:299–312, 1999; by permission of Cell Press.]

ceptors are clearly products of a recent evolutionary event (Figure 33.27). The green and red pigments appear to have diverged in the primate lineage approximately 35 million years ago. Mammals, such as dogs and mice, that diverged from primates earlier have only two cone photoreceptors, blue and green. They are not sensitive to light as far toward the infrared region as we are, and they do not discriminate colors as well. In contrast, birds such as chickens have a total of six pigments: rhodopsin, four cone pigments, and a pineal visual pigment called *pinopsin*. Birds have highly acute color perception.

Second, the high level of similarity between the green and the red pigments has made the identification of the specific amino acid residues responsible for spectral tuning possible. Three residues (at positions 180, 277, and 285) are responsible for most of the difference between the green and the red pigments. In the green pigment, these residues are alanine, phenylalanine, and alanine, respectively; in the red pigment, they are serine, tyrosine, and threonine. A hydroxyl group has been added to each amino acid in the red pigment. The hydroxyl groups can interact with the photoexcited state of retinal and lower its energy, leading to a shift toward the lower-energy (red) region of the spectrum.

Rearrangements in the genes for the green and red pigments lead to "color blindness"

Homologous recombination
The exchange of DNA segments with substantial sequence similarity at equivalent positions between chromosomes.

The genes for the green and red pigments lie adjacent to each other on the human X chromosome. These genes are more than 98% identical in nucleotide sequence, including introns and untranslated regions as well as the protein-coding region. Regions with such high similarity are very susceptible to unequal homologous recombination.

Recombination can take place either between or within transcribed regions of the gene (Figure 33.28). If recombination takes place between transcribed regions, the product chromosomes will differ in the number of pigment genes that they carry. One chromosome will lose a gene and thus may lack the gene for, say, the green pigment; the other chromosome will gain a gene. Consistent with this scenario, approximately 2% of human X chromosomes carry only a single color-pigment gene, approximately 20% carry two, 50% carry three, 20% carry four, and 5% carry five or more. A person lacking the gene for the green pigment will have trouble distinguishing red and green color, characteristic of the most common form of color blindness. Approximately 5% of males have this form of color blindness. Recombination can also take place within the transcription units,

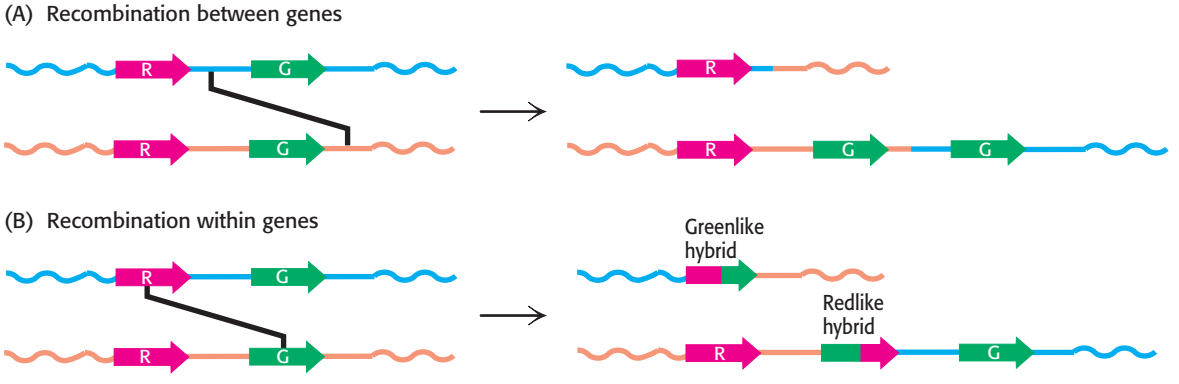


Figure 33.28 Recombination pathways leading to color blindness. Rearrangements in the course of DNA replication may lead to (A) the loss of visual pigment genes or (B) the formation of hybrid pigment genes that encode photoreceptors with anomolous absorption spectra. Because the amino acids most important for determining absorption spectra are in the carboxyl-terminal half of each photoreceptor protein, the part of the gene that encodes this region most strongly affects the absorption characteristics of hybrid receptors. [After J. Nathans. *Neuron* 24:299–312, 1999; by permission of Cell Press.]

resulting in genes that encode hybrids of the green and red photoreceptors. The absorption maximum of such a hybrid lies between that of the red and green pigments. A person with such hybrid genes who also lacks either a functional red- or a functional green-pigment gene does not discriminate color well.

33.4 Hearing Depends on the Speedy Detection of Mechanical Stimuli

Hearing and touch are based on the detection of mechanical stimuli. Although the proteins of these senses have not been as well characterized as those of the senses already discussed, anatomical, physiological, and biophysical studies have elucidated the fundamental processes. *A major clue to the mechanism of hearing is its speed*. We hear frequencies ranging from 200 to 20,000 Hz (cycles per second), corresponding to times of 5 to 0.05 ms. Furthermore, our ability to locate sound sources, one of the most important functions of hearing, depends on the ability to detect the time delay between the arrival of a sound at one ear and its arrival at the other. Given the separation of our ears and the speed of sound, we must be able to accurately sense time differences of 0.7 ms. In fact, human beings can locate sound sources associated with temporal delays as short as 0.02 ms. This high time resolution implies that hearing must employ direct transduction mechanisms that do not depend on second messengers. Recall that, in vision, for which speed also is important, the signal-transduction processes take place in milliseconds.

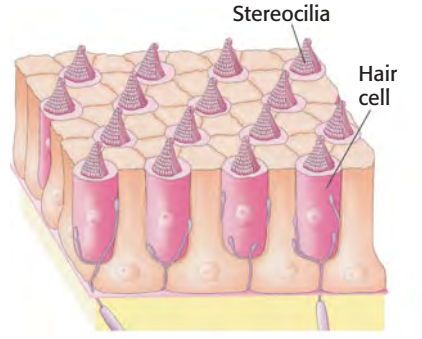


Figure 33.29 Hair cells, the sensory neurons crucial for hearing. These specialized neurons are capped with hairlike projections called stereocilia that are responsible for detecting very subtle vibrations. [After A. J. Hudspeth. *Nature* 341:397–404, 1989.]

Hair cells use a connected bundle of stereocilia to detect tiny motions

Sound waves are detected inside the cochlea of the inner ear. The *cochlea* is a fluid-filled, membranous sac that is coiled like a snail shell. The primary detection is accomplished by specialized neurons inside the cochlea called *hair cells* (Figure 33.29). Each cochlea contains approximately 16,000 hair cells, and each hair cell contains a hexagonally shaped bundle of 20 to 300 hairlike projections called *stereocilia* (Figure 33.30). These stereocilia are graded in length across the bundle. Mechanical deflection of the hair bundle, as takes place when a sound wave arrives at the ear, creates a change in the membrane potential of the hair cell.

Micromanipulation experiments have directly probed the connection between mechanical stimulation and membrane potential. Displacement toward the direction of the tallest part of the hair bundle results in the depolarization of the hair cell, whereas displacement in the opposite direction results in the hyperpolarization (Figure 33.31). Motion perpendicular to the hair-length gradient does not produce any change in resting potential.

Figure 33.30 An electron micrograph of a hair bundle. [Courtesy of Dr. A. Jacobs and Dr. A.J. Hudspeth.]

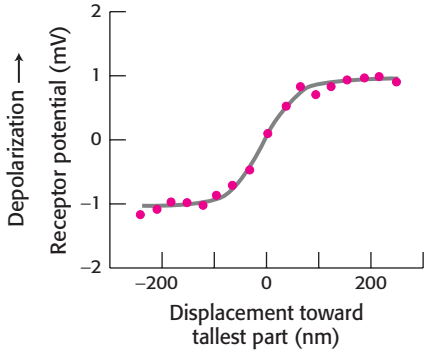


Figure 33.31 Micromanipulation of a hair cell. Movement toward the tallest part of the bundle depolarizes the cell as measured by the microelectrode. Movement toward the shortest part hyperpolarizes the cell. Lateral movement has no effect. [After A.J. Hudspeth. *Nature* 341:397–404, 1989.]

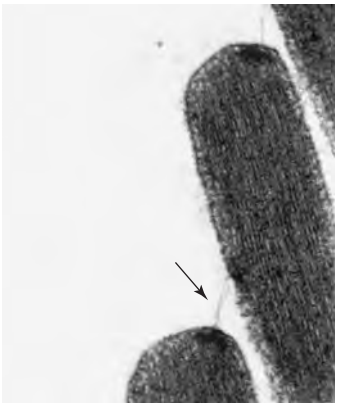


Figure 33.32 Electron micrograph of tip links. The tip link between two hair fibers is marked by an arrow. [Courtesy of Dr. A. Jacobs and Dr. A.J. Hudspeth.]

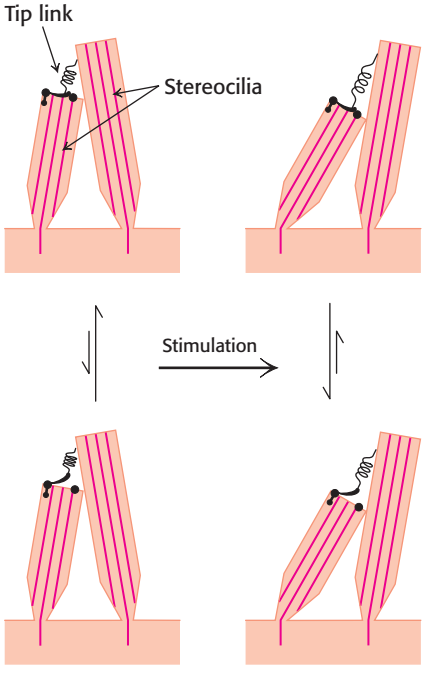


Figure 33.33 Model for hair-cell transduction. When the hair bundle is tipped toward the tallest part, the tip link pulls on an ion channel and opens it. Movement in the opposite direction relaxes the tension in the tip link, increasing the probability that any open channels will close. [After A.J. Hudspeth. *Nature* 341:397–404, 1989.]

Remarkably, *displacement of the hair bundle by as little as 3 Å (0.3 nm) results in a measurable (and functionally important) change in membrane potential.* This motion of 0.003 degree corresponds to a 1-inch movement of the top of the Empire State Building.

How does the motion of the hair bundle create a change in membrane potential? The rapid response, within microseconds, suggests that the movement of the hair bundle acts on ion channels directly. An important observation is that adjacent stereocilia are linked by individual filaments called *tip links* (Figure 33.32).

The presence of these tip links suggests a simple mechanical model for transduction by hair cells (Figure 33.33). The tip links are coupled to ion channels in the membranes of the stereocilia that are gated by mechanical stress. In the absence of a stimulus, approximately 15% of these channels are open. When the hair bundle is displaced toward its tallest part, the stereocilia slide across one another and the tension on the tip links increases, causing additional channels to open. The flow of ions through the newly opened channels depolarizes the membrane. Conversely, if the displacement is in the opposite direction, the tension on the tip links decreases, the open channels close, and the membrane hyperpolarizes. *Thus, the mechanical motion of the hair bundle is directly converted into current flow across the hair-cell membrane.*

Mechanosensory channels have been identified in *Drosophila* and vertebrates

The search for ion channels that respond to mechanical impulses has been pursued in a variety of organisms. *Drosophila* have sensory bristles used for detecting small air currents. These bristles respond to mechanical displacement in ways similar to those of hair cells; displacement of a bristle in one direction leads to substantial transmembrane current. Strains of mutant fruit flies that show uncoordinated motion and clumsiness have been examined for their electrophysiological responses to displacement of the sensory bristles. In one set of strains, transmembrane currents were dramatically reduced. The mutated gene in these strains was found to encode a protein of 1619 amino acids, called NompC for *no mechanoreceptor potential*.

The carboxyl-terminal 469 amino acids of NompC resemble a class of ion-channel proteins called TRP (*transient receptor potential*) channels. This region includes six putative transmembrane helices with a pore-like region between the fifth and the sixth helices. The amino-terminal 1150 amino acids consist almost exclusively of 29 *ankyrin repeats* (Figure 33.34).

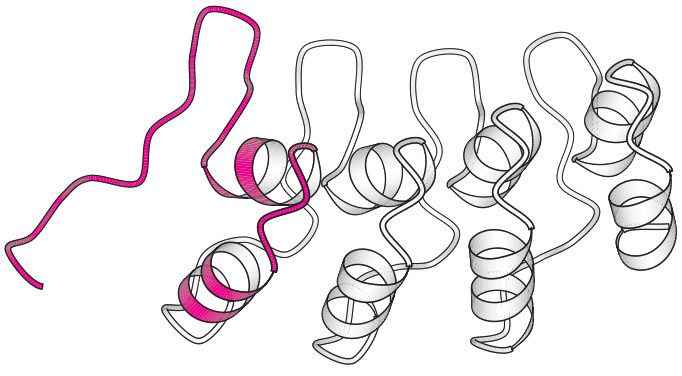


Figure 33.34 Ankyrin repeat structure. One ankyrin domain is shown in red in this series of four ankyrin repeats. *Notice* the hairpin loop followed by a helix-turn-helix motif in the red-colored ankyrin unit. Ankyrin domains interact with other proteins, primarily through their loops. [Drawn from 1AWC.pdb.]

Ankyrin repeats are structural motifs consisting of a hairpin loop followed by a helix-turn-helix. Importantly, in other proteins, regions with tandem arrays of these motifs mediate protein–protein interactions, suggesting that these arrays couple the motions of other proteins to the activity of the NompC channel.

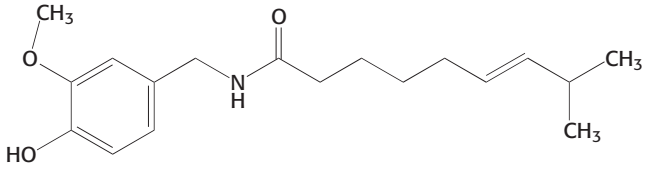
A candidate for at least one component of the mechanosensory channel taking part in hearing has been identified. The protein, TRPA1, is a member of the TRP channel family. The sequence of TRPA1 also includes 17 ankyrin repeats. TRPA1 is expressed in hair cells, particularly near their tips, and is likely to be at least one component of the mechanosensory channel that is central to hearing. Further studies are under way to confirm and extend this exciting discovery.

33.5 Touch Includes the Sensing of Pressure, Temperature, and Other Factors

Like taste, touch is a combination of sensory systems that are expressed in a common organ—in this case, the skin. The detection of pressure and the detection of temperature are two key components. Amiloride-sensitive Na^+ channels, homologous to those of taste, appear to play a role. Other systems are responsible for detecting painful stimuli such as high temperature, acid, or certain specific chemicals. Although our understanding of this sensory system is not as advanced as that of the other sensory systems, recent work has revealed a fascinating relation between pain and taste sensation, a relation well known to anyone who has eaten "spicy" food.

Studies of capsaicin reveal a receptor for sensing high temperatures and other painful stimuli

Our sense of touch is intimately connected with the sensation of pain. Specialized neurons, termed *nociceptors,* transmit signals from skin to pain-processing centers in the spinal cord and brain in response to the onset of tissue damage. What is the molecular basis for the sensation of pain? An intriguing clue came from the realization that *capsaicin,* the chemical responsible for the "hot" taste of spicy food, activates nociceptors.



Capsaicin

Early research suggested that capsaicin would act by opening ion channels that are expressed in nociceptors. Thus, a cell that expresses the *capsaicin receptor* should take up calcium on treatment with the molecule. This insight led to the isolation of the capsaicin receptor with the use of cDNA from cells expressing this receptor. Such cells had been detected by their fluorescence when loaded with the calcium-sensitive compound Fura-2 and then treated with capsaicin or related molecules. Cells expressing the capsaicin receptor, which is called VR1 (for *v*anilloid *r*eceptor 1), respond to capsaicin below a concentration of 1 μM. The deduced 838-residue sequence of VR1 revealed it to be a member of the TRP channel family (Figure 33.35). The amino-terminal region of VR1 includes three ankyrin repeats.

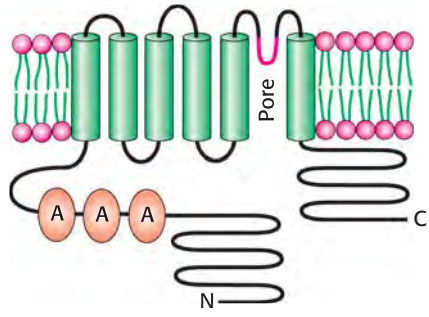


Figure 33.35 The membrane topology deduced for VR1, the capsaicin receptor. The proposed site of the membrane pore is indicated in red, and the three ankyrin (A) repeats are shown in orange. The active receptor comprises four of these subunits. [After M. J. Caterina, et al. *Nature* 389:816–824, 1997.]

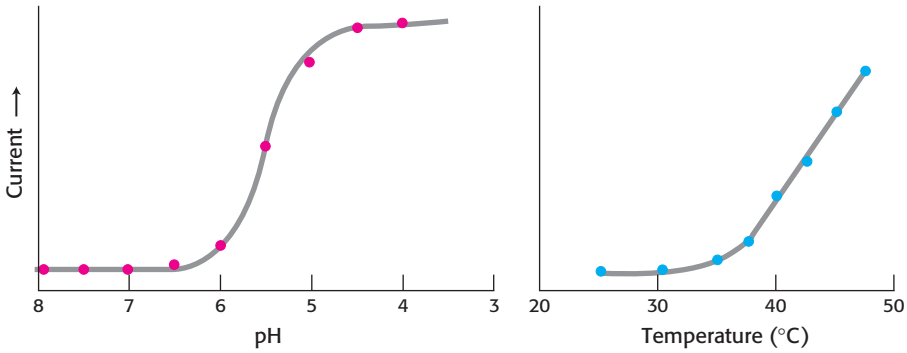


Figure 33.36 Response of the capsaicin receptor to pH and temperature. The ability of this receptor to respond to acid and to increased temperature helps detect potentially noxious situations. [After M. Tominaga, et al. *Neuron* 21:531–543, 1998.]

Currents through VR1 are also induced by temperatures above 40°C and by exposure to dilute acid, with a midpoint for activation at pH 5.4 (Figure 33.36). Temperatures and acidity in these ranges are associated with infection and cell injury. The responses to capsaicin, temperature, and acidity are not independent. The response to heat is greater at lower pH, for example. Thus, *VR1 acts to integrate several noxious stimuli*. We feel these responses as pain and act to prevent the potentially destructive conditions that cause the unpleasant sensation. Mice that do not express VR1 suggest that this is the case; such mice do not mind food containing high concentrations of capsaicin and are, indeed, less responsive than control mice to normally noxious heat. Plants such as chili peppers presumably gained the ability to synthesize capsaicin and other "hot" compounds to protect themselves from being consumed by mammals. Birds, which play the beneficial role of spreading pepper seeds into new territory, do not appear to respond to capsaicin.

Because of its ability to simulate VR1, capsaicin is used in pain management for arthritis, neuralgia, and other neuropathies. How can a compound that induces pain assist in its alleviation? Chronic exposure to capsaicin overstimulates pain-transmitting neurons, leading to their desensitization.

More sensory systems remain to be studied

There may exist other subtle senses that are able to detect environmental signals that then influence our behavior. The biochemical basis of these senses is now under investigation. One such sense is our ability to respond, often without our awareness, to chemical signals called pheromones, released by other persons. Another is our sense of time, manifested in our daily (circadian) rhythms of activity and restfulness. Daily changes in light exposure strongly influence these rhythms. The foundations for these senses have been uncovered in other organisms; future studies should reveal to what extent these mechanisms apply to human beings as well.

Summary

Smell, taste, vision, hearing, and touch are based on signal-transduction pathways activated by signals from the environment. These sensory systems function similarly to the signal-transduction pathways for many hormones. These intercellular signaling pathways appear to

have been appropriated and modified to process environmental information.

33.1 A Wide Variety of Organic Compounds Are Detected by Olfaction

The sense of smell, or olfaction, is remarkable in its specificity; it can, for example, discern stereoisomers of small organic compounds as distinct aromas. The 7TM receptors that detect these odorants operate in conjunction with $G_{(olf)}$, a G protein that activates a cAMP cascade resulting in the opening of an ion channel and the generation of a nerve impulse. An outstanding feature of the olfactory system is its ability to detect a vast array of odorants. Each olfactory neuron expresses only one type of receptor and connects to a particular region of the olfactory bulb. Odors are decoded by a combinatorial mechanism: each odorant activates a number of receptors, each to a different extent, and most receptors are activated by more than one odorant.

33.2 Taste Is a Combination of Senses That Function by Different Mechanisms

We can detect only five tastes: bitter, sweet, salt, sour, and umami. The transduction pathways that detect taste are, however, diverse. Bitter, sweet, and umami tastants are experienced through 7TM receptors acting through a special G protein called gustducin. Salty and sour tastants act directly through membrane channels. Salty tastants are detected by passage though Na^+ channels, whereas sour taste results from the effects of hydrogen ions on a number of types of channels. The end point is the same in all cases—membrane polarization that results in the transmission of a nerve impulse.

33.3 Photoreceptor Molecules in the Eye Detect Visible Light

Vision is perhaps the best understood of the senses. Two classes of photoreceptor cells exist: cones, which respond to bright lights and colors, and rods, which respond only to dim light. The photoreceptor in rods is rhodopsin, a 7TM receptor that is a complex of the protein opsin and the chromophore 11-*cis*-retinal. The absorption of light by 11-*cis*-retinal changes its structure into that of all-*trans*-retinal, setting in motion a signal-transduction pathway that leads to the breakdown of cGMP, to membrane hyperpolarization, and to a subsequent nerve impulse. Color vision is mediated by three distinct 7TM photoreceptors that employ 11-*cis*-retinal as a chromophore and absorb light in the blue, green, and red parts of the spectrum.

33.4 Hearing Depends on the Speedy Detection of Mechanical Stimuli

The immediate receptors for hearing are found in the hair cells of the cochleae, which contain bundles of stereocilia. When the stereocilia move in response to sound waves, cation channels will open or close, depending on the direction of movement. The mechanical motion of the cilia is converted into current flow and then into a nerve impulse.

33.5 Touch Includes the Sensing of Pressure, Temperature, and Other Factors

Touch, detected by the skin, senses pressure, temperature, and pain. Specialized nerve cells called nociceptors transmit signals that are interpreted in the brain as pain. A receptor responsible for the perception of pain has been isolated on the basis of its ability to bind capsaicin, the molecule responsible for the hot taste of spicy food. The capsaicin receptor, also called VR1, functions as a cation channel that initiates a nerve impulse.

Key Terms

main olfactory epithelium (p. 958)
$G_{(olf)}$ (p. 959)
gustducin (p. 963)
amiloride-sensitive Na^+ channel (p. 965)
rod (p. 966)
cone (p. 966)
rhodopsin (p. 966)
opsin (p. 966)
retinal (p. 966)
chromophore (p. 967)
transducin (p. 968)
cGMP phosphodiesterase (p. 968)
cGMP-gated ion channel (p. 968)
rhodopsin kinase (p. 968)
arrestin (p. 969)
guanylate cyclase (p. 969)
hair cell (p. 971)
stereocilium (p. 971)
tip link (p. 972)
nociceptor (p. 973)
capsaicin receptor (VR1) (p. 973)

Problems

1. *Olfaction in worms*. Unlike the olfactory neurons in the mammalian systems discussed herein, olfactory neurons in the nematode *C. elegans* express multiple olfactory receptors. In particular, one neuron (called AWA) expresses receptors for compounds to which the nematode is attracted, whereas a different neuron (called AWB) expresses receptors for compounds that the nematode avoids. Suppose that a transgenic nematode is generated such that one of the receptors for an attractant is expressed in AWB rather than AWA. What behavior would you expect in the presence of the corresponding attractant?

2. *Odorant matching*. A mixture of two of the compounds illustrated in Figure 33.6 is applied to a section of olfactory epithelium. Only receptors 3, 5, 9, 12, and 13 are activated, according to Figure 33.7. Identify the likely compounds in the mixture.

3. *Timing*. Compare the aspects of taste (bitter, sweet, salty, sour) in regard to their potential for rapid time resolution.

4. *Two ears*. Our ability to determine the direction from which a sound is coming is partly based on the difference in time at which our two ears detect the sound. Given the speed of sound (350 m s^{-1}) and the separation between our ears (0.15 m), what difference is expected in the times at which a sound arrives at our two ears? How does this difference compare with the time resolution of the human hearing system? Would a sensory system that utilized 7TM receptors and G proteins be capable of adequate time resolution?

5. *It's better to be bitter*. Some nontoxic plants taste very bitter to us. Suggest one or more explanations.

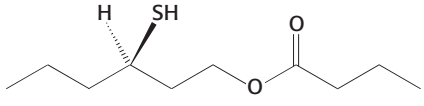
6. *Of mice and men*. In human beings, the umami taste is triggered only by glutamate and aspartate. In contrast, mice respond to many more amino acids. Design an experiment to test which of the subunits (T1R1 or T1R3) determines the specificity of this response. Assume that all desired mouse strains can be readily produced.

7. *Color discrimination*. Some women who have had male children who display color blindness have been demonstrated to distinguish between colors that are not readily distinguished by other people. Propose an explanation.

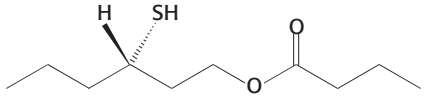
8. *Combinatorial power*. How many different odorants could a person distinguish if each odorant is bound by a single type of receptor and each receptor binds only a single odorant? How many if each odorant binds to two different odorant receptors? Three odorant receptors?

9. *Light action*. Describe the effect of light absorption on 11-*cis*-retinal bound within rhodopsin..

10. *Mirror, mirror*. Compounds A and B typically evoke different odors, with compound A described as smelling like tropical fruit and compound B smelling oniony.



A



B

Explain how two such similar compounds can smell so different.

11. *Channeling*. Provide an example of an ion channel in vision, taste, and hearing.

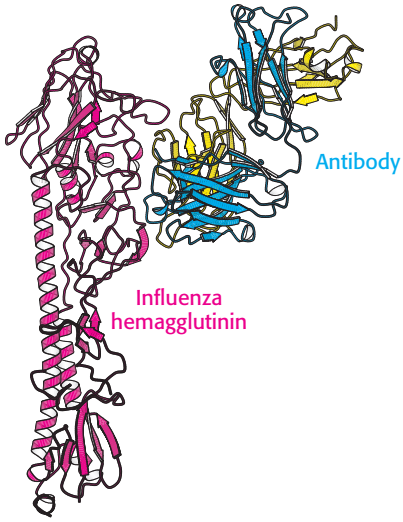
Chapter Integration Problem

12. *Energy and information*. The transmission of sensory information requires the input of free energy. For each sensory system (olfaction, gustation, vision, hearing, and touch), identify mechanisms for the input of free energy that allow the transmission of sensory information.

Mechanism Problem

13. *Schiff-base formation*. Propose a mechanism for the reaction between opsin and 11-*cis*-retinal.

The Immune System



Just as medieval defenders used their weapons and the castle walls to defend their city, the immune system constantly battles against foreign invaders such as viruses, bacteria, and parasites to defend the organism. Antibody molecules provide a key element in the immune system's defensive arsenal. For example, specific antibodies can bind to molecules on the surfaces of viruses and prevent the viruses from infecting cells. Above right, an antibody binds to one subunit on hemagglutinin from the surface of influenza virus. [(Left) The Granger Collection.]

We are constantly exposed to an incredible diversity of bacteria, viruses, and parasites, many of which would flourish in our cells or extracellular fluids were it not for our immune system. How does the immune system protect us? The human body has two lines of defense: an *innate immune system* that responds rapidly to features present in many pathogens, and an *adaptive immune system* that responds to specific features present only in a given pathogen. *Both the innate and the adaptive immune systems first identify features on disease-causing organisms and then work to eliminate or neutralize those organisms.* While a thorough description of the immune system is certainly beyond the scope of this book, this chapter will focus on how biochemical concepts such as protein structure, receptor-ligand interactions, and signal transduction are applied to the identification of pathogens.

The immune system must meet two tremendous challenges in the identification of pathogens: (1) to produce a system of receptors diverse enough to recognize a wide array of potential pathogens and (2) to distinguish invaders and their disease-causing products from the organism's own products (i.e., self- versus nonself-recognition). To meet these challenges, the innate immune system evolved the ability to recognize structural

OUTLINE

elements, such as specific glycolipids or forms of nucleic acid, that are well conserved in pathogens but absent in the host organism. The repertoire of such elements is limited, however, and so some pathogens have strategies to escape detection. The adaptive immune system has the remarkable ability to produce more than 10^8 distinct proteins, called antibodies, that can recognize different foreign molecules, and more than 10^{12} receptors on immune cells, called T-cell receptors (TCRs), each of which presents a different surface with the potential to specifically bind a structure from a foreign organism. In producing this vast range of defensive molecules, however, the adaptive immune system has the potential to create antibodies and T-cells that recognize and attack cells or molecules normally present in our bodies—a situation that can result in autoimmune diseases.

This chapter will examine these challenges, focusing first on the structures of proteins that recognize foreign organisms and then on the mechanisms for protecting us from a specific pathogen once it has been recognized. The chapter will closely examine the modular construction of the proteins of the immune system—identifying structural motifs and considering how spectacular diversity can arise from modular construction.

Innate immunity is an evolutionarily ancient defense system

Innate immunity is an evolutionarily ancient defense system found, at least in some form, in all multicellular plants and animals. The innate immune system represents the first line of defense against foreign pathogens, relying on common features of invading organisms to identify and eliminate these threats. Components of the innate immune system include the epithelial lining that surrounds host cells and the specialized cells, called *phagocytes*, that can ingest and destroy pathogens without the aid of the adaptive immune system.

The innate immune system also includes a family of receptors that can recognize specific features present in most pathogens and yet not respond to materials normally present in the host. The best-understood of these receptors are the *Toll-like receptors* (TLRs). The name "toll-like" is derived from a receptor known as Toll encoded in the *Drosophila* genome; Toll was first identified in a screen for genes important for *Drosophila* development and was subsequently shown to play a key role in the innate immune system later in development. The TLRs have a common structure (Figure 34.1). Each receptor consists of a large extracellular domain built primarily from repeated amino acid sequences termed leucine-rich repeats (LRRs). Each LRR typically contains 20–30 residues, including 6 that are usually leucine. The human TLRs have from 18 to 27 LRR repeats that are followed by a sequence forming a single transmembrane helix and an intracellular signaling domain. This signaling domain is not a protein kinase but acts as a docking site for other proteins. Most TLRs are expressed in the cell membrane for the detection of extracellular pathogens such as fungi and bacteria. Other TLRs are located in the membranes of internal compartments for the detection of intracellular pathogens such as viruses and some bacteria.

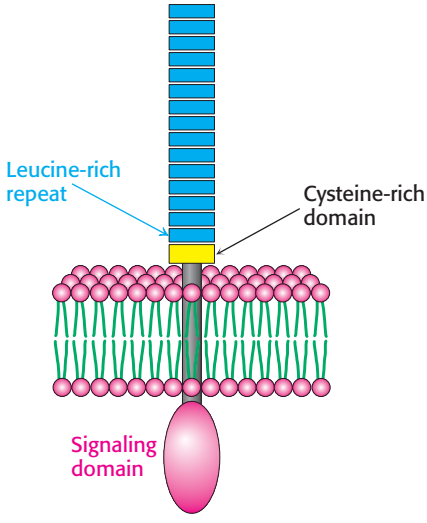


Figure 34.1 Toll-like receptor. Each receptor is made up of a set of 18 or more leucine-rich repeat sequences, followed by a cysteine-rich domain, a single transmembrane helix, and an intracellular domain that functions in signal transduction.

Each TLR targets a specific molecular characteristic, often called a *pathogen-associated molecular pattern* (PAMP), found primarily on invading organisms. Typically, a PAMP is a critical component of the pathogen's function: mutations in these targets cannot easily block recognition by the TLR without compromising the activity of the pathogen. One particularly important PAMP is lipopolysaccharide (LPS), also referred to as *endotoxin*, a specific class of glycolipid found in the cell walls of Gram-negative bacteria such as *E. coli*. LPS is recognized by TLR-4. The response of the innate immune system to LPS can be easily demonstrated. Injection of less than

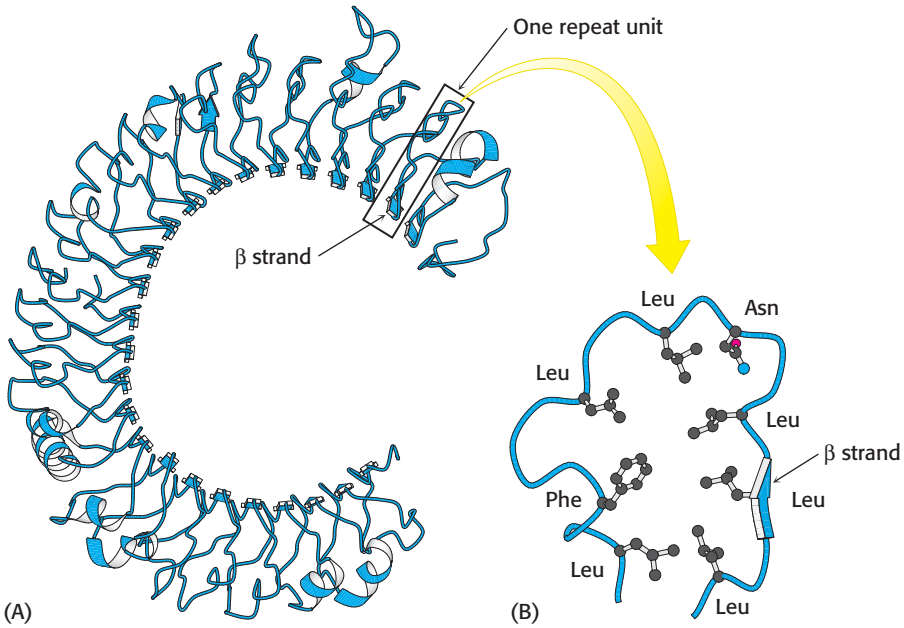


Figure 34.2 PAMP-recognition unit of the Toll-like receptor. (A) The structure of the leucine-rich repeat (LRR) domain from human TLR-3. *Notice* that the LRR units come together to form a central parallel β sheet that curls to form a concave structure. (B) The structure of a single LRR showing the positions of the residues that are generally approximately conserved. *Notice* that the leucine residues come together to form a hydrophobic core with the single β strand along on one side. [Drawn from 1ZIW.pdb].

1 mg of LPS into a human being produces a fever and other signs of inflammation even though no living organisms are introduced.

How do TLRs recognize PAMPs? The leucine-rich repeat domain from human TLR-3 has a remarkable structure (Figure 34.2). Each of its LRR units contributes a single β strand to a large parallel β sheet that lines the inside of a concave structure. This hooklike structure immediately suggests a model for how TLRs bind PAMPs—namely, that the PAMP lies on the inside of the "hook." This model is likely accurate for some TLRs. However, for other TLRs, the PAMP-binding site appears to lie on one side of the structure, and the central hole is blocked by host carbohydrates linked to the structure.

Regardless of the details of the interaction, PAMP binding appears to lead to the formation of a specific dimer of the TLR. The cytoplasmic side of this dimer is a signaling domain that initiates the signal-transduction pathway. Because the TLRs and other components of the innate immune system are always expressed, they provide the host organism with a rapid response to resist attack by pathogens. However, a number of pathogens have evolved the ability to escape detection by the innate immune system. For protection against such pathogens, the host relies on the adaptive immune system, which is able to target specific pathogens, even those that it has never encountered in the course of evolution.

The adaptive immune system responds by using the principles of evolution

The adaptive immune system comprises two parallel but interrelated systems: humoral and cellular immune responses. In the *humoral immune response,* soluble proteins called *antibodies (immunoglobulins)* function as recognition elements that bind to foreign molecules and serve as markers signaling foreign invasion. Antibodies are secreted by *plasma cells,* which are derived from *B lymphocytes (B cells)* (Figure 34.3). A foreign macromolecule that binds selectively to an antibody is called an *antigen.* In a physiological context, if the binding of the foreign molecule stimulates an immune response, that molecule is called an *immunogen*. The specific affinity of an antibody is not for the entire macromolecular antigen but for a particular

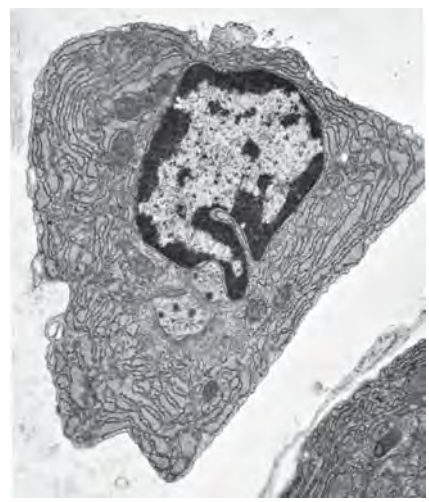


Figure 34.3 Immunoglobulin production. An electron micrograph of a plasma cell shows the highly developed rough endoplasmic reticulum necessary for antibody secretion. [Courtesy of Lynne Mercer.]

site on the antigen called the *epitope* or *antigenic determinant*. Each B cell produces just one type of antibody that can recognize a single epitope.

In the *cellular immune response*, cells called *cytotoxic T lymphocytes* (also commonly called *killer T cells*) destroy cells that have been invaded by a pathogen. Because intracellular pathogens do not leave markings on the exteriors of infected cells, vertebrates have evolved a mechanism to mark the exterior of cells with a sample of the interior contents, both self and foreign. Some of the internal proteins are broken into peptides, which are then bound to a complex of integral membrane proteins encoded by the *major histocompatbility complex* (MHC). T cells continually scan the bound peptides to find and kill cells that display foreign motifs on their surfaces. Another class of T cells called *helper T lymphocytes* contributes to both the humoral and the cellular immune responses by stimulating the differentiation and proliferation of appropriate B cells and cytotoxic T cells. The cellular immune response is mediated by specific receptors that are expressed on the surfaces of the T cells.

The remarkable ability of the immune system to adapt to an essentially limitless set of potential pathogens requires a powerful system for transforming the immune cells and molecules present in our systems in response to the presence of pathogens. *This adaptive system operates through the principles of evolution, including reproduction with variation followed by selection of the most well suited members of a population.*

If the human genome contains, by the latest estimates, only 25,000 genes, how can the immune system generate more than 10^8 different antibody proteins and 10^{12} T-cell receptors? The answer is found in a novel mechanism for generating a highly diverse set of genes from a limited set of genetic building blocks. Linking different sets of DNA regions in a combinatorial manner produces many distinct protein-encoding genes that are not present in the genome. A rigorous selection process then leaves for proliferation only cells that synthesize proteins determined to be useful in the immune response. The subsequent reproduction of these cells without additional recombination serves to enrich the cell population with members expressing particular protein species.

Critical to the development of the immune response is the selection process, which determines which immune cells will reproduce. The process comprises several stages. In the early stages of the development of an immune response, cells expressing molecules that bind tightly to self-molecules are destroyed or silenced, whereas cells expressing molecules that do not bind strongly to self-molecules and that have the potential for binding strongly to foreign molecules are preserved. The appearance of an immunogenic invader at a later time will stimulate cells expressing antibodies or T-cell receptors that bind specifically to elements of that pathogen to reproduce—in evolutionary terms, such cells are selected for. Thus, the immune response is based on the selection of cells expressing molecules that are specifically effective against a particular invader; the response evolves from a population with wide-ranging specificities to a more-focused collection of cells and molecules that are well suited to defend the host when confronted with that particular challenge.

Not only are antibodies and T-cell receptors a result of genetic diversity and recombination, but antibodies have highly diverse structures as well. Antibodies require many different structural solutions for binding many different antigens, each of which has a different form. T-cell receptors, in contrast, are not structurally diverse, because they have coevolved with the MHC. The docking mode of a T-cell receptor to the peptide bound to MHC is similar for all structures. As a consequence of this coevolution, every T-cell receptor has an inherent reactivity with every MHC. The

coevolution ensures that all T-cell receptors can scan all peptide–MHC complexes on all tissues. The genetic diversity of the 10^{12} different T-cell receptors is concentrated in a highly diverse set of residues in the center of the MHC groove. This localized diversity allows the T-cell receptor to recognize the many different foreign peptides bound to the MHC. T-cell receptors must survey many different MHC–peptide complexes with rapid turnover. Therefore, the binding affinities between T-cell receptors and the MHC are weaker than those between antibody and antigen.

34.1 Antibodies Possess Distinct Antigen-Binding and Effector Units

Antibodies are central molecular players in the immune response. In Chapter 3, we discussed the application of antibodies as tools to study proteins; let us now consider the native structures and functions of these remarkable molecules. In 1959, Rodney Porter showed that *immunoglobulin G* (IgG), the major antibody in serum, can be cleaved into three 50-kd fragments by the limited proteolytic action of papain. Two of these fragments bind antigen. They are called F_{ab} (*F* stands for *f*ragment, *ab* for *a*ntigen *b*inding). The other fragment, called F_c because it crystallizes readily, does not bind antigen, but it has other important biological activities, including the mediation of responses termed *effector functions*. These functions include the initiation of the *complement cascade*, a process that leads to the lysis of target cells. Although such effector functions are crucial to the functioning of the immune system, they will not be considered further here.

How do these fragments relate to the three-dimensional structure of whole IgG molecules? Immunoglobulin G consists of two kinds of polypeptide chains, a 25-kd *light* (L) *chain* and a 50-kd *heavy* (H) *chain* (Figure 34.4). The subunit composition is L_2H_2. Each L chain is linked to an H chain by a disulfide bond, and the H chains are linked to each other by at least one disulfide bond. Examination of the amino acid sequences and three-dimensional structures of IgG molecules reveals that each L chain comprises two homologous domains, termed *immunoglobulin domains*, to be described in detail in Section 34.2. Each H chain has four immunoglobulin domains. Overall, the molecule adopts a conformation that resembles the letter Y, in which the stem, corresponding to the F_c fragment obtained by cleavage with papain, consists of the two carboxyl-terminal immunoglobulin domains of each H chain and in which the two arms of the Y, corresponding to the two F_{ab} fragments, are formed by the two amino-terminal domains of each H chain and

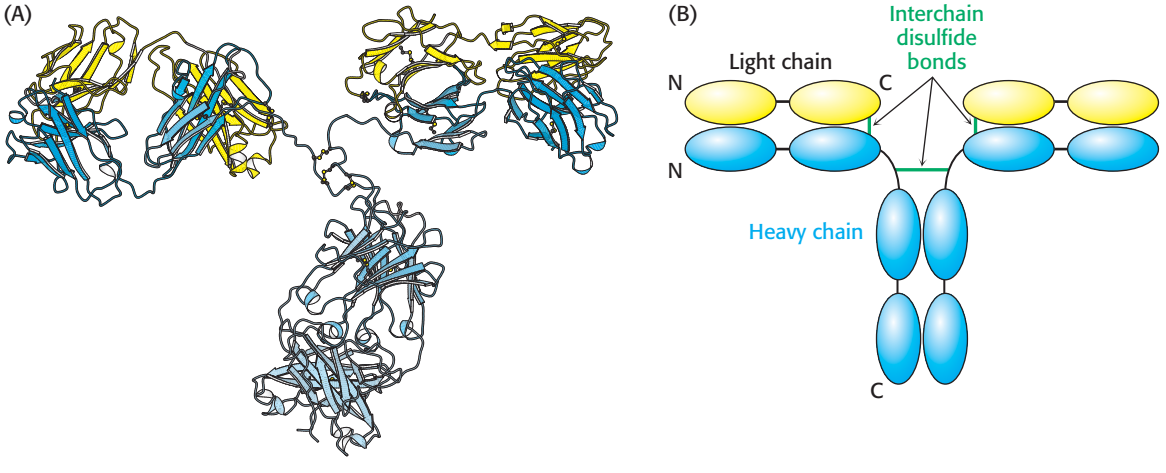


Figure 34.4 Immunoglobulin G structure. (A) The three-dimensional structure of an IgG molecule showing the light chains in yellow and the heavy chains in blue. (B) A schematic view of an IgG molecule indicating the positions of the interchain disulfide bonds. Abbreviations: N, amino terminus; C, carboxyl terminus. [Drawn from 1IGT.pdb.]

Table 34.1 Properties of immunoglobulin classes

Class	Serum concentration (mg ml^{-1})	Mass (kd)	Light chains	Heavy chains	Chain structure
IgG	12	150	κ or λ	γ	$\kappa_2\gamma_2$ or $\lambda_2\gamma_2$
IgA	3	180–500	κ or λ	α	$(\kappa_2\alpha_2)_n$ or $(\lambda_2\alpha_2)_n$
IgM	1	950	κ or λ	μ	$(\kappa_2\mu_2)_5$ or $(\kappa_2\mu_2)_5$
IgD	0.1	175	κ or λ	δ	$\kappa_2\delta_2$ or $\lambda_2\delta_2$
IgE	0.001	200	κ or λ	ε	$\kappa_2\varepsilon_2$ or $\lambda_2\varepsilon_2$

Note: n = 1, 2, or 3. IgM and oligomers of IgA also contain J chains that connect immunoglobulin molecules. IgA in secretions has an additional component.

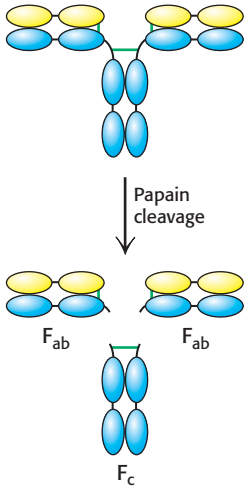


Figure 34.5 Immunoglobulin G cleavage. Treatment of intact IgG molecules with the protease papain results in the formation of three large fragments: two F_{ab} fragments that retain antigen-binding capability and one F_c fragment that does not.

the two amino-terminal domains of each L chain. The linkers between the stem and the two arms consist of extended polypeptide regions within the H chains and are quite flexible.

Papain cleaves the H chains on the carboxyl-terminal side of the disulfide bond that links each L and H chain (Figure 34.5). Thus, each F_{ab} consists of an entire L chain and the amino-terminal half of an H chain, whereas F_c consists of the carboxyl-terminal halves of both H chains. Each F_{ab} contains a single antigen-binding site. Because an intact IgG molecule contains two F_{ab} components and therefore has two binding sites, it can cross-link multiple antigens (Figure 34.6). Furthermore, the F_c and the two F_{ab} units of the intact IgG are joined by flexible polypeptide regions that allow facile variation in the angle between the F_{ab} units through a wide range (Figure 34.7). This kind of mobility, called *segmental flexibility,* can enhance the formation of an antibody–antigen complex by enabling both combining sites on an antibody to bind an antigen that possesses multiple binding sites, such as a viral coat composed of repeating identical monomers or a bacterial cell surface. The combining sites at the tips of the F_{ab} units simply move to match the distance between specific determinants on the antigen.

Immunoglobulin G is the antibody present in highest concentration in the serum, but other classes of immunoglobulin also are present (Table 34.1). Each class includes an L chain (either κ or λ) and a distinct H chain (Figure 34.8). The heavy chains in IgG are called γ chains, whereas those in immunoglobulins A, M, D, and E are called α, μ, δ, and ε, respectively. *Immunoglobulin M* (IgM) is the first class of antibody to appear in the serum after exposure to an antigen. The presence of 10 combining sites enables IgM to bind especially tightly to antigens containing multiple identical epitopes. The strength of an interaction comprising multiple independent binding interactions between partners is termed *avidity* rather than affinity, which denotes the binding strength of a single combining site.

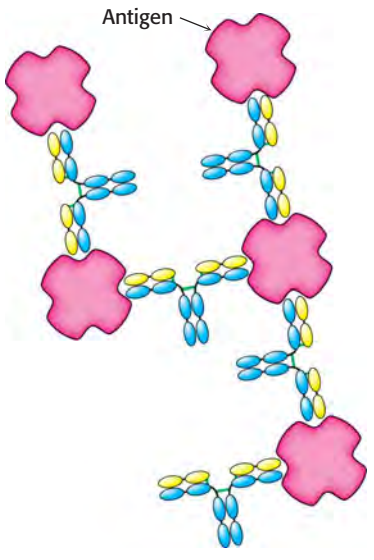


Figure 34.6 Antigen cross-linking. Because IgG molecules include two antigen-binding sites, antibodies can cross-link multivalent antigens such as viral surfaces.

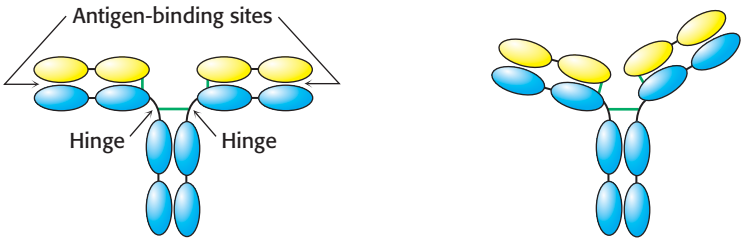


Figure 34.7 Segmental flexibility. The linkages between the F_{ab} and the F_c regions of an IgG molecule are flexible, allowing the two antigen-binding sites to adopt a range of orientations with respect to each other. This flexibility allows effective interactions with a multivalent antigen without requiring that the epitopes on the target be a precise distance apart.

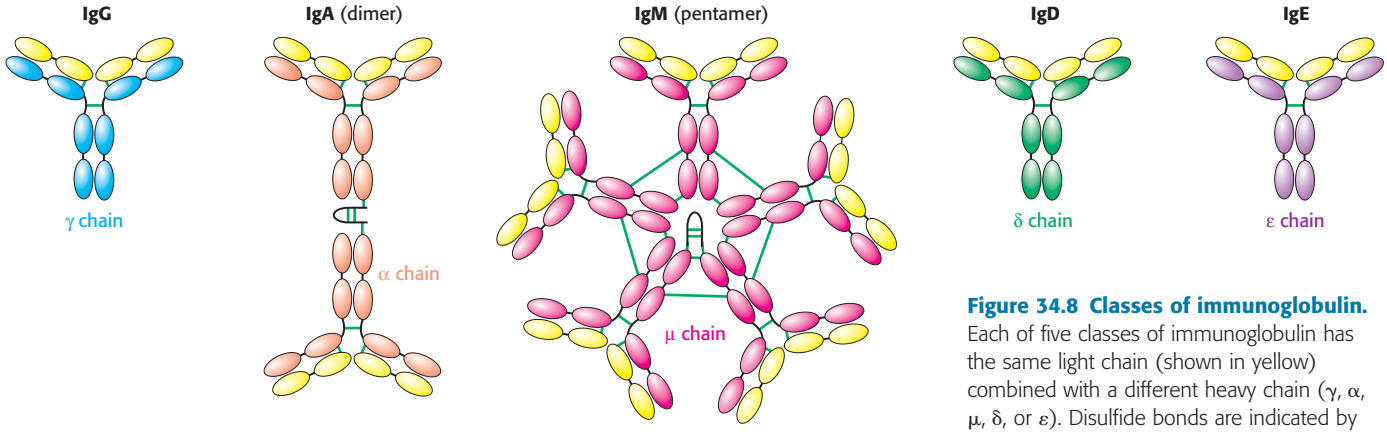


Figure 34.8 Classes of immunoglobulin. Each of five classes of immunoglobulin has the same light chain (shown in yellow) combined with a different heavy chain (γ, α, μ, δ, or ε). Disulfide bonds are indicated by green lines. The IgA dimer and the IgM pentamer have a small polypeptide chain in addition to the light and heavy chains.

Immunoglobulin A (IgA) is the major class of antibody in external secretions, such as saliva, tears, bronchial mucus, and intestinal mucus. Thus, IgA serves as a first line of defense against bacterial and viral antigens. The role of *immunoglobulin D* (IgD) is not yet known. *Immunoglobulin E* (IgE) is important in conferring protection against parasites, but IgE also participates in allergic reactions. IgE–antigen complexes form cross-links with receptors on the surfaces of mast cells to trigger a cascade that leads to the release of granules containing pharmacologically active molecules. Histamine, one of the agents released, induces smooth-muscle contraction and stimulates the secretion of mucus.

34.2 Antibodies Bind Specific Molecules Through Hypervariable Loops

A comparison of the amino acid sequences of different IgG antibodies from human beings or mice shows that the carboxyl-terminal half of the L chains and the carboxyl-terminal three-quarters of the H chains are very similar in all of the antibodies. Importantly, the amino-terminal domain of each chain is more variable, including three stretches of approximately 7 to 12 amino acids within each chain that are hypervariable, as shown for the H chain in Figure 34.9. The amino-terminal immunglobulin domain of each chain is thus referred to as the *variable region,* whereas the remaining immunoglobulin

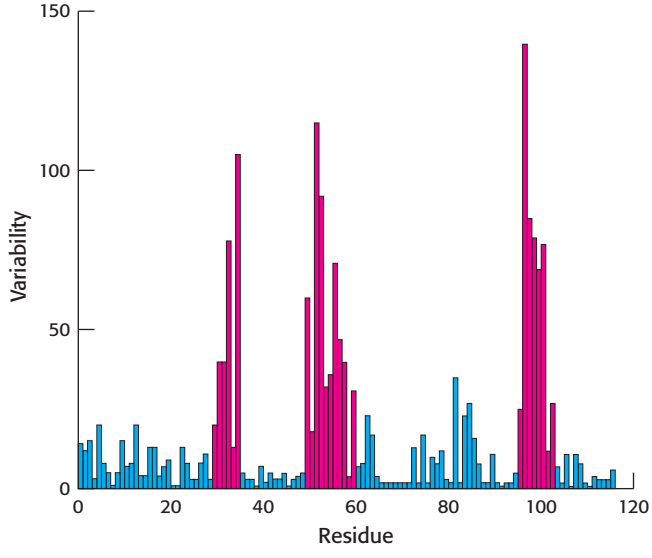


Figure 34.9 Immunoglobulin sequence diversity. A plot of sequence variability as a function of position along the sequence of the amino-terminal immunoglobulin domain of the H chain of human IgG molecules. Three regions (in red) show remarkably high levels of variability. These hypervariable regions correspond to three loops in the immunoglobulin domain structure. [After R. A. Goldsby, T. J. Kindt, and B. A. Osborne, *Kuby Immunology,* 4th ed. (W. H. Freeman and Company, 2000), p. 91.]

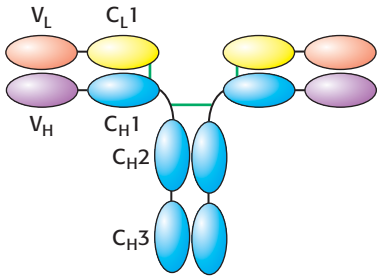


Figure 34.10 Variable and constant regions. Each L and H chain includes one immunoglobulin domain at its amino terminus that is quite variable from one antibody to another. These domains are referred to as V_L and V_H. The remaining domains are more constant from one antibody to another and are referred to as constant domains (C_L1, C_H1, C_H2, and C_H3).

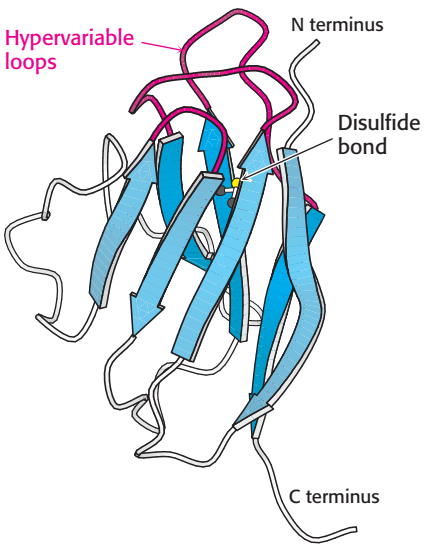


Figure 34.11 Immunoglobulin fold. An immunoglobulin domain consists of a pair of β sheets linked by a disulfide bond and hydrophobic interactions. *Notice* that three hypervariable loops lie at one end of the structure. [Drawn from 1DQJ.pbd.]

domains are much more similar in all antibodies and are referred to as *constant regions* (Figure 34.10).

The immunoglobulin fold consists of a beta-sandwich framework with hypervariable loops

An IgG molecule consists of a total of 12 immunoglobulin domains. These domains have many sequence features in common and adopt a common structure, the *immunoglobulin fold* (Figure 34.11). Remarkably, this same structural domain is found in many other proteins that play key roles in the immune system and in nonimmune functions.

The immunoglobulin fold consists of a pair of β sheets, each built of antiparallel β strands, that surround a central hydrophobic core. A single disulfide bond bridges the two sheets. Two aspects of this structure are particularly important for its function. First, three loops present at one end of the structure form a potential binding surface. These loops contain the hypervariable sequences present in antibodies and in T-cell receptors. Variation of the amino acid sequences of these loops provides the major mechanism for the generation of the vastly diverse set of antibodies and T-cell receptors expressed by the immune system. These loops are referred to as *hypervariable loops* or *complementarity-determining regions* (CDRs). Second, the amino terminus and the carboxyl terminus are at opposite ends of the structure, which allows structural domains to be strung together to form chains, as in the L and H chains of antibodies. Such chains are present in several other key molecules in the immune system.

The immunoglobulin fold is one of the most prevalent domains encoded by the human genome: more than 750 genes encode proteins with at least one immunoglobulin fold recognizable at the level of amino acid sequence. Such domains are also common in other multicellular animals such as flies and nematodes. However, from inspection of amino acid sequence alone, immunoglobulin-fold domains do not appear to be present in yeast or plants, although these organisms possess other structurally similar domains, including the key photosynthetic electron-transport protein plastocyanin in plants (Section 19.3). Thus, the immunoglobulin-fold family appears to have expanded greatly along evolutionary branches leading to animals—particularly, vertebrates.

X-ray analyses have revealed how antibodies bind antigens

For each class of antibody, the variable domains at the amino-terminal ends of the L and H chains (designated V_L and V_H) come together to form a binding surface. The positions of the complementarity-determining regions are striking. These hypervariable sequences, present in three loops of each domain, come together so that all six loops form a single surface at the end of each arm (Figure 34.12). Because virtually any V_L can pair with any V_H, *a very large number of different binding sites can be constructed by their combinatorial association*.

The results of x-ray crystallographic studies of several hundred large and small antigens bound to F_{ab} molecules have been sources of much insight into the structural basis of antibody specificity. The binding of antigens to antibodies is governed by the same principles that govern the binding of substrates to enzymes. The interaction between complementary shapes results in numerous contacts between amino acids at the binding surfaces of both molecules. Many hydrogen bonds, electrostatic interactions, and van der Waals interactions, reinforced by hydrophobic interactions, combine to give specific and strong binding.

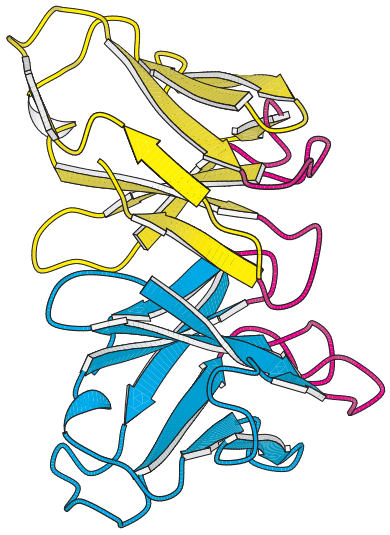


Figure 34.12 Variable domains. A side view of the variable domains of the L chain (yellow) and the H chain (blue); the complementarity-determining regions (CDRs) are shown in red. *Notice* that the six CDRs come together to form a binding surface. The specificity of the surface is determined by the sequences and structures of the CDRs. [Drawn from 1DQJ.pdb.]

A few aspects of antibody binding merit specific attention, inasmuch as they relate directly to the structure of immunoglobulins. The binding site on the antibody incorporates some or all of the CDRs in the variable domains of the antibody. Small molecules are likely to make contact with fewer CDRs, with perhaps 15 residues of the antibody participating in the binding interaction. Macromolecules often make more extensive contact, sometimes interacting with all six CDRs and 20 or more residues of the antibody. Small molecules often bind in a cleft of the antigen-binding region. Macromolecules, such as globular proteins, tend to interact across larger, fairly flat apposed surfaces bearing complementary protrusions and depressions.

A well-studied case of small-molecule binding is seen in an example of phosphorylcholine bound to F_{ab}. Crystallographic analysis revealed phosphorylcholine bound to a cavity lined by residues from five CDRs—two from the L chain and three from the H chain (Figure 34.13). The positively charged trimethylammonium group of phosphorylcholine is buried inside the wedge-shaped cavity, where it interacts electrostatically with two negatively charged residues, a glutamate and an aspartate. The negatively charged phosphoryl group of phosphorylcholine binds to the positively charged guanidinium group of an arginine residue at the mouth of the crevice and is hydrogen bonded to the side chain of a nearby tyrosine residue. Numerous van der Waals interactions, such as those made by a tryptophan side chain, also stabilize this complex.

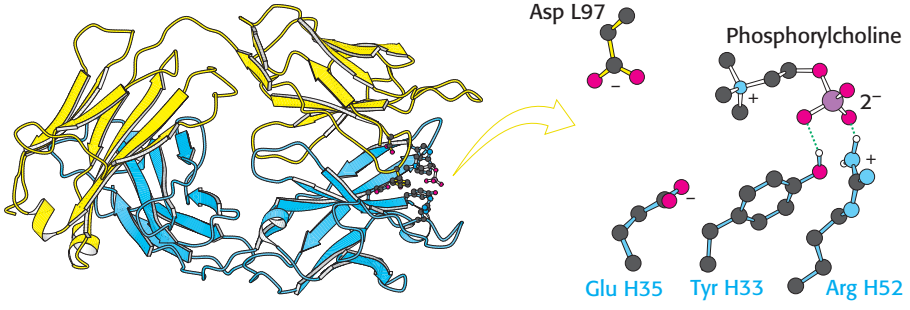


Figure 34.13 Binding of a small antigen. The structure of a complex between an F_{ab} fragment of an antibody and its target—in this case, phosphorylcholine. *Notice* that residues from the antibody interact with phosphorylcholine through hydrogen bonding and electrostatic interactions [Drawn from 2MCP.pdb.]

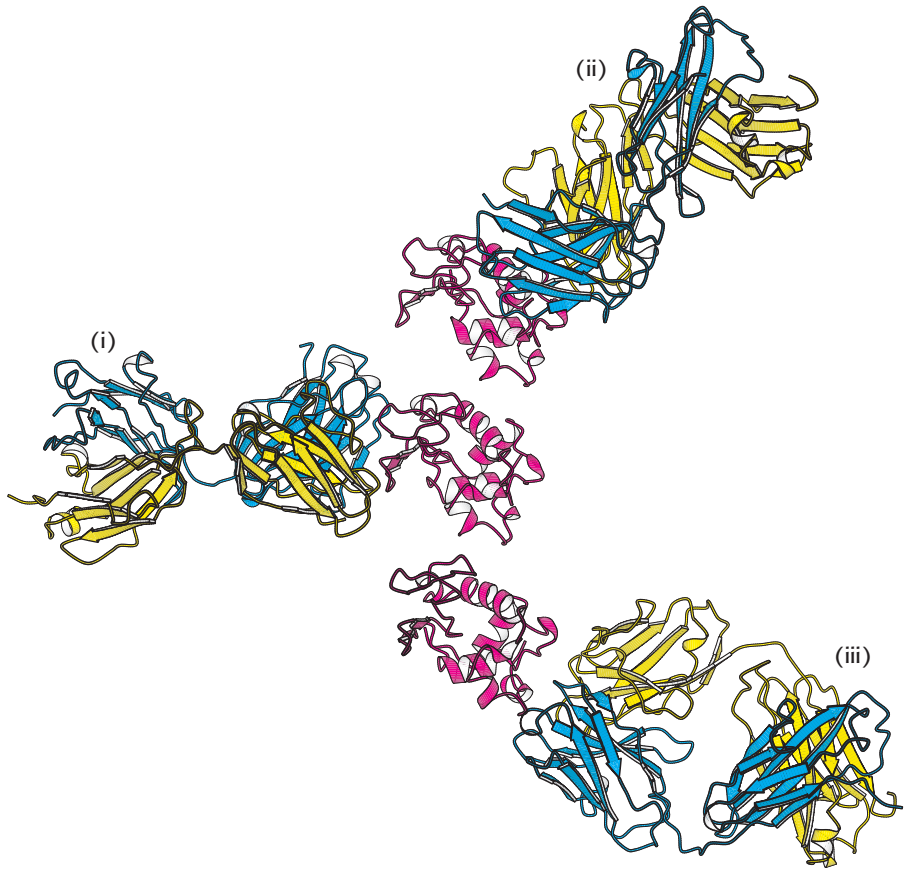


Figure 34.14 Antibodies against lysozyme. The structures of three complexes (i, ii, iii) between F_{ab} fragments (blue and yellow) and hen egg-white lysozyme (red) shown with lysozyme in the same orientation in each case. *Notice* that the three antibodies recognize completely different epitopes on the lysozyme molecule. [Drawn from 3HFL, 1DQJ, and 1FDL.pdb.]

Residues from five CDRs participate in the binding of phosphorylcholine to human F_{ab}. This binding does not significantly change the structure of the antibody, yet induced fit plays a role in the formation of many antibody–antigen complexes. A malleable binding site can accommodate many more kinds of ligands than can a rigid one. Thus, induced fit increases the repertoire of antibody specificities.

Large antigens bind antibodies with numerous interactions

How do large antigens interact with antibodies? A large collection of antibodies against hen egg-white lysozyme has been structurally characterized in great detail (Figure 34.14). Each different antibody binds to a distinct surface of lysozyme. Let us examine the interactions in one of these complexes (complex ii in Figure 34.14) in detail. This antibody binds two polypeptide segments of lysozyme that are widely separated in the primary structure (Figure 34.15).

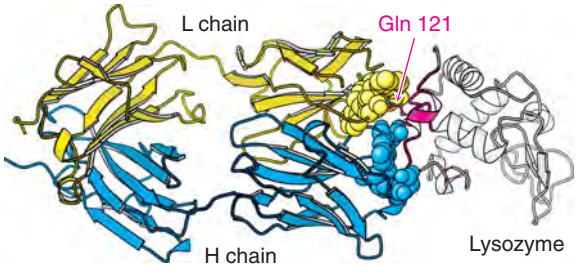


Figure 34.15 Antibody–protein interactions. *Notice* that a single residue of lysozyme, glutamine 121, penetrates more deeply into the antibody combining site. [Drawn from 1FDL.pdb.]

All six CDRs of the antibody make contact with this epitope. The region of contact is quite extensive (about 30 × 20 Å). The apposed surfaces are rather flat. The only exception is the side chain of glutamine 121 of lysozyme, which penetrates deeply into the antibody's binding site, where it forms a hydrogen bond with a main-chain carbonyl oxygen atom and is surrounded by three aromatic side chains. The formation of 12 hydrogen bonds and numerous van der Waals interactions contributes to the high affinity (K_d = 20 nM) of this antibody–antigen interaction. Examination of the F_{ab} molecule without bound protein reveals that the structures of the V_L and V_H domains change little on binding, although they slide 1 Å apart to allow more intimate contact with lysozyme.

34.3 Diversity Is Generated by Gene Rearrangements

A mammal such as a mouse or a human being can synthesize large amounts of specific antibody against virtually any foreign determinant within a matter of days after exposure. We have seen that antibody specificity is determined by the amino acid sequences of the variable regions of both light and heavy chains, which brings us to the key question: How are different variable-region sequences generated?

The discovery of distinct variable and constant regions in the L and H chains raised the possibility that the genes that encode immunoglobulins have an unusual architecture that facilitates the generation of a diverse set of polypeptide products. In 1965, William Dreyer and Claude Bennett proposed that multiple *V (variable) genes* are separate from a single *C (constant) gene* in embryonic (germ-line) DNA. According to their model, one of these V genes becomes joined to the C gene in the course of differentiation of the antibody-producing cell. A critical test of this novel hypothesis had to await the isolation of pure immunoglobulin mRNA and the development of techniques for analyzing mammalian genomes. Twenty years later, Susumu Tonegawa found that V and C genes are indeed far apart in embryonic DNA but are closely associated in the DNA of antibody-producing cells. Thus, immunoglobulin genes are rearranged in the differentiation of lymphocytes.

J (joining) genes and D (diversity) genes increase antibody diversity

Sequencing studies carried out by Susumu Tonegawa, Philip Leder, and Leroy Hood revealed that V genes in embryonic cells do not encode the entire variable region of L and H chains. Consider, for example, the region that encodes the κ light-chain family. A tandem array of 40 segments, each of which encodes approximately the first 97 residues of the variable domain of the L chain, is present on human chromosome 2 (Figure 34.16).

Figure 34.16 The κ light-chain locus. This part of human chromosome 2 includes an array of 40 segments that encode the variable (V) region (approximately residues 1–97) of the light chain, an array of 5 segments that encode the joining (J) region (residues 98–110), and a single region that encodes the constant (C) region.

However, the variable region of the L chain contains 110 residues. Where is the DNA that encodes the last 13 residues of the variable domain? For L chains in undifferentiated cells, this stretch of DNA is located in an unexpected place: near the C gene. It is called the *J gene* because it joins the V and C gene segments in a differentiated cell. In fact, a tandem array of five J gene segments is located near the C gene in embryonic cells. In the differentiation of an antibody-producing cell, a V gene segment becomes spliced to a J gene segment to form a complete gene for the variable region (Figure 34.17). RNA splicing generates an mRNA molecule for the

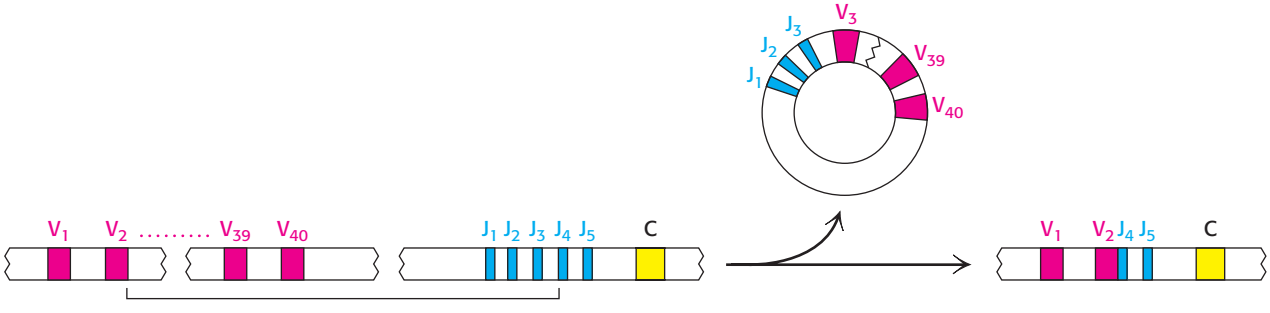


Figure 34.17 VJ recombination. A single V gene (in this case, V_2) is linked to a J gene (here, J_4) to form an intact VJ region. The intervening DNA is released in a circular form. Because the V and J regions are selected at random and the joint between them is not always in exactly the same place, many VJ combinations can be generated by this process.

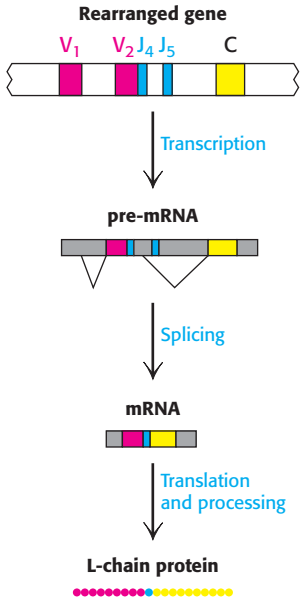


Figure 34.18 Light-chain expression. The light-chain protein is expressed by the transcription of the rearranged gene to produce a pre-RNA molecule with the VJ and C regions separated. RNA splicing removes the intervening sequences to produce an mRNA molecule with the VJ and C regions linked. Translation of the mRNA and processing of the initial protein product produce the light chain.

complete L chain by linking the coding regions for the rearranged VJ unit with that for the C unit (Figure 34.18).

J gene segments are important contributors to antibody diversity because they encode part of the last hypervariable segment (CDR3). In forming a continuous variable-region gene, any of the 40 V gene segments can become linked to any of 5 J gene segments. Thus, somatic recombination of these gene segments amplifies the diversity already present in the germ line. The linkage between V and J is not precisely controlled. Recombination between these genes can take place at one of several bases near the codon for residue 95, generating additional diversity. A similar array of V and J gene segments encoding the λ light chain is present on human chromosome 22. This region includes 30 V_λ gene segments and four J_λ gene segments. In addition, this region includes four distinct C genes, in contrast with the single C gene in the κ locus.

In human beings, the genes encoding the heavy chain are present on chromosome 14. Remarkably, the variable domain of heavy chains is assembled from *three* rather than two segments. In addition to V_H gene segments that encode residues 1 through 94 and J_H gene segments that encode residues 98 through 113, this chromosomal region includes a distinct set of segments that encode residues 95 through 97 (Figure 34.19). These gene segments are called D (for *d*iversity). Some 27 D gene segments lie between 51 V_H and 6 J_H gene segments. The recombination process first joins a D gene segment to a J_H gene segment; a V_H gene segment is then joined to DJ_H. A greater variety of antigen-binding patches and clefts can be formed by the H chain than by the L chain because the H chain is encoded by three rather than two gene segments. Moreover, CDR3 of the H chain is diversified by the action of terminal deoxyribonucleotidyl transferase, a special DNA polymerase that requires no template. This enzyme inserts extra nucleotides between V_H and D. The *V(D)J recombination* of both the L and the H chains is executed by specific enzymes present in immune cells. These proteins, called *RAG-1* and *RAG-2*, recognize specific DNA sequences called *recombination signal sequences* (RSSs) adjacent to the V, D, and J gene segments and facilitate the cleavage and religation of the DNA segments.

More than 10^8 antibodies can be formed by combinatorial association and somatic mutation

Let us recapitulate the sources of antibody diversity. The germ line contains a rather large repertoire of variable-region genes. For κ light chains, there are about 40 V-segment genes and 5 J-segment genes. Hence, a total of $40 \times 5 = 200$ kinds of complete V_κ genes can be formed by the combinations of V and J. A similar analysis suggests that at least 120 different λ light chains can be generated. A larger number of heavy-chain genes can be formed because of the role of the D segments. For 51 V, 27 D, and 6 J gene segments, the number of complete V_H genes that can be formed is 8262. The

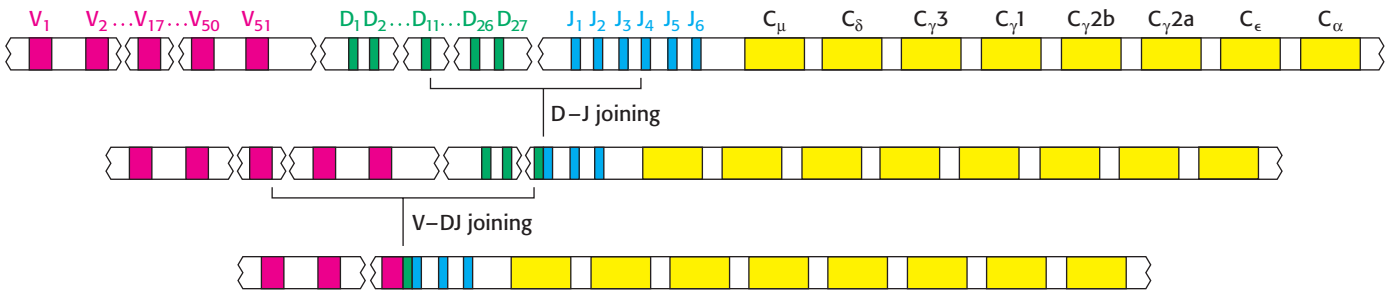


Figure 34.19 V(D)J recombination. The heavy-chain locus includes an array of 51 V segments, 27 D segments, and 6 J segments. Gene rearrangement begins with D–J joining, followed by further rearrangement to link the V segment to the DJ segment.

association of 320 kinds of L chains with 8262 kinds of H chains would yield 2.6×10^6 different antibodies. Variability in the exact points of segment joining and other mechanisms increases this value by at least two orders of magnitude.

Even more diversity is introduced into antibody chains by *somatic mutation*—that is, by the introduction of mutations into the recombined genes. In fact, a 1000-fold increase in binding affinity is seen in the course of a typical humoral immune response, arising from somatic mutation, a process called *affinity maturation*. The generation of an expanded repertoire leads to the selection of antibodies that more precisely fit the antigen. Thus, nature draws on each of three sources of diversity—a germ-line repertoire, somatic recombination, and somatic mutation—to form the rich variety of antibodies that protect an organism from foreign incursions.

The oligomerization of antibodies expressed on the surfaces of immature B cells triggers antibody secretion

The processes heretofore described generate a highly diverse set of antibody molecules—a key first step in the generation of an immune response. The next stage is the selection of a particular set of antibodies directed against a specific invader. How is this selection accomplished? Each immature B cell, produced in the bone marrow, expresses a specific monomeric form of IgM attached to its surface (Figure 34.20). Each cell expresses approximately 10^5 IgM molecules, but *all of these molecules are identical in amino acid sequence and, hence, in antigen-binding specificity*. Thus, the selection of a particular immature B cell for growth will lead to the amplification of an antibody with a unique specificity. The selection process begins with the binding of an antigen to the membrane-bound antibody.

Associated with each membrane-linked IgM molecule are two molecules of a heterodimeric membrane protein called Ig-α–Ig-β (see Figure 34.20). Examination of the amino acid sequences of Ig-α and Ig-β is highly instructive. The amino terminus of each protein lies outside the cell and corresponds to a single immunoglobulin fold, and the carboxyl terminus, which lies inside the cell, includes a sequence of 18 amino acids called an *immunoreceptor tyrosine-based activation motif* (ITAM; see Figure 34.20). As its name suggests, each ITAM includes key tyrosine residues, which are subject to phosphorylation by particular protein kinases present in immune-system cells.

A fundamental observation with regard to the mechanism by which the binding of antigen to membrane-bound antibody triggers the subsequent steps of the immune response is that *oligomerization or clustering of the antibody molecules is required* (Figure 34.21). The requirement for oligomerization is reminiscent of the dimerization of receptors triggered by epidermal growth factor and insulin encountered in Section 14.3; indeed, the associated signaling mechanisms appear to be quite similar. The oligomerization of the membrane-bound antibodies results in the phosphorylation of the tyrosine residues within the ITAMs by protein tyrosine kinases including Lyn, a homolog of Src (Section 14.5). The phosphorylated ITAMs serve as docking sites for a protein kinase termed spleen tyrosine *k*inase (Syk), which has two SH2 domains that interact with the pair of phosphorylated tyrosine residues in each

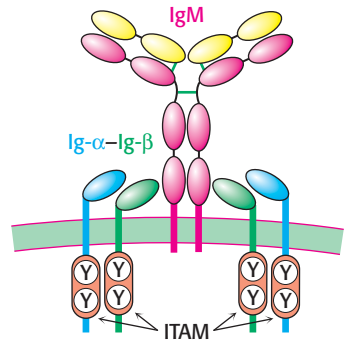


Figure 34.20 B-cell receptor. This complex consists of a membrane-bound IgM molecule noncovalently bound to two Ig-α–Ig-β heterodimers. The intracellular domains of each of the Ig-α and Ig-β chains include an immunoreceptor tyrosine-based activation motif (ITAM).

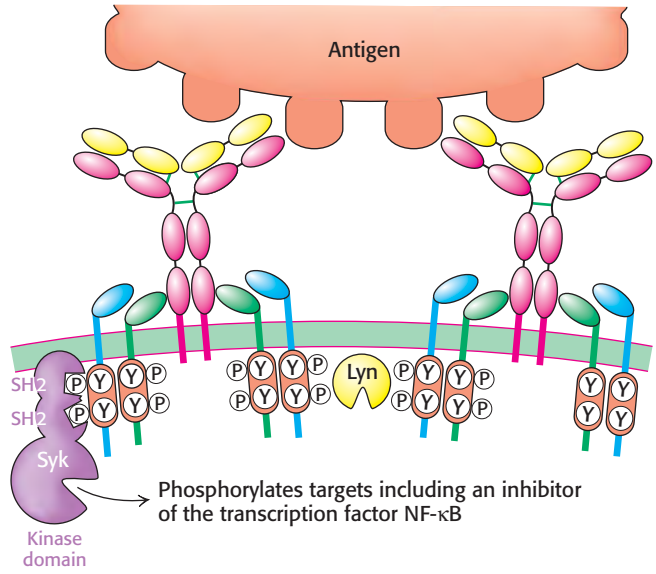
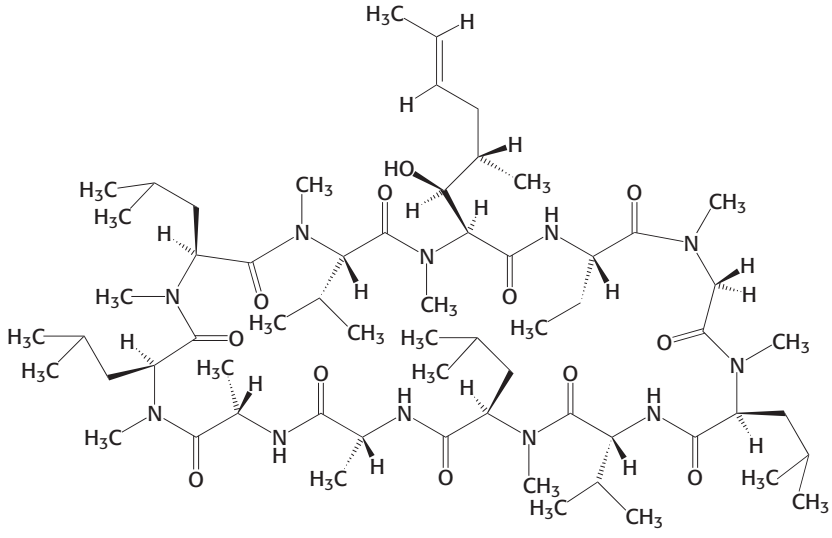


Figure 34.21 B-cell activation. The binding of a multivalent antigen such as a bacterial or viral surface links membrane-bound IgM molecules. This oligomerization triggers the phosphorylation of tyrosine residues in the ITAM sequences by protein tyrosine kinases such as Lyn. After phosphorylation, the ITAMs serve as docking sites for Syk, a protein kinase that phosphorylates a number of targets, including transcription factors.

ITAM. Syk, when activated by phosphorylation, proceeds to phosphorylate other signal-transduction proteins including an inhibitory subunit of a transcription factor called NF-κB and an isoform of phospholipase C. The signaling processes continue downstream to activate gene expression, leading to the stimulation of cell growth and initiating further B-cell differentiation.

Drugs that modulate the immune system have served as sources of insight into immune-system signaling pathways. For example, *cyclosporin,* a powerful suppressor of the immune system, acts by blocking a phosphatase called *calcineurin,* which normally activates a transcription factor called NF-AT by dephosphorylating it.



Cyclosporin A

The potent inhibition of the immune system, or *immunosuppression*, induced by cyclosporin reveals how crucial the activity of NF-AT is to the development of an immune response. Without such drugs, organ transplantation would be extremely difficult because transplanted tissue expresses a wide range of foreign antigens, causing the host immune system to reject the new tissue.

The role of oligomerization in the B-cell signaling pathway is illuminated when we consider the nature of many antigens presented by pathogens. The surfaces of many viruses, bacteria, and parasites are characterized by arrays of identical membrane proteins or membrane-linked carbohydrates. Thus, most pathogens present multiple binding surfaces that will naturally cause membrane-associated antibodies to oligomerize as they bind adjacent epitopes. In addition, the mechanism accounts for the observation that most small molecules do not induce an immune response; however, coupling multiple copies of a small molecule to a large oligomeric protein such as keyhole limpet hemocyanin (KLH), which has a molecular mass of close to 1 million daltons or more, promotes antibody oligomerization and, hence, the production of antibodies against the small-molecule epitope. The large protein is called the *carrier* of the attached chemical group, which is called a *haptenic determinant*. The small foreign molecule by itself is called a *hapten.* Antibodies elicited by attached haptens will bind unattached haptens as well.

Different classes of antibodies are formed by the hopping of V_H genes

The development of an effective antibody-based immune response depends on the secretion into the blood of antibodies that have appropriate effector

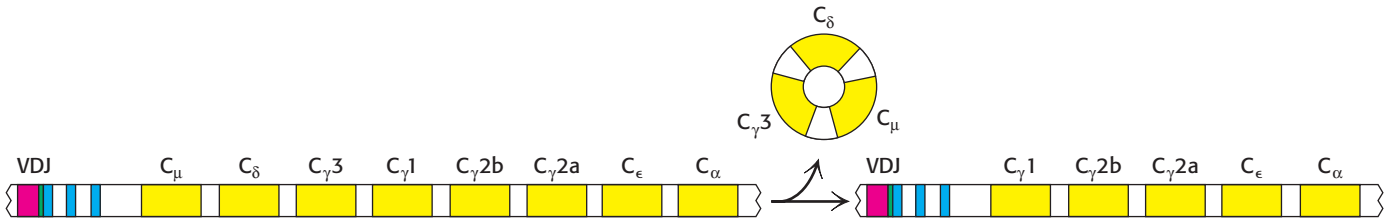


Figure 34.22 Class switching. Further rearrangement of the heavy-chain locus results in the generation of genes for antibody classes other than IgM. In the case shown, rearrangement places the VDJ region next to the Cγ1 region, resulting in the production of IgG1. Note that no further rearrangement of the VDJ region takes place, and so the specificity of the antibody is not affected.

functions. At the beginning of this response, an alternative mRNA-splicing pathway is activated so that the production of membrane-linked IgM is supplanted by the synthesis of secreted IgM. As noted in Section 33.1, secreted IgM is pentameric and has a high avidity for antigens containing multiple identical epitopes. Later, the antibody-producing cell makes either IgG, IgA, IgD, or IgE of the same specificity as that of the intially secreted IgM. In this switch, the light chain and the variable region of the heavy chain are unchanged. Only the constant region of the heavy chain changes. This step in the differentiation of an antibody-producing cell is called *class switching* (Figure 34.22). In undifferentiated cells, the genes for the constant region of each class of heavy chain, called C_μ, C_δ, C_γ, C_ε, and C_α, are next to one another. There are eight in all, including four genes for the constant regions of γ chains. A complete gene for the heavy chains of IgM antibody is formed by the translocation of a V_H gene segment to a DJ_H gene segment.

How are other heavy chains formed? Class switching is mediated by a gene-rearrangement process that moves a VDJ gene from a site near one C gene to a site near another C gene. Importantly, *the antigen-binding specificity is conserved in class switching because the entire V_HDJ_H gene is translocated in an intact form*. For example, the antigen-combining specificity of IgA produced by a particular cell is the same as that of IgM synthesized at an earlier stage of its development. The biological significance of C_H switching is that a whole recognition domain (the variable domain) is shifted from the early constant region (C_μ) to one of several other constant regions that mediate different effector functions.

34.4 Major-Histocompatibility-Complex Proteins Present Peptide Antigens on Cell Surfaces for Recognition by T-Cell Receptors

Soluble antibodies are highly effective against extracellular pathogens, but they confer little protection against microorganisms that are predominantly intracellular, such as some viruses and mycobacteria (which cause tuberculosis and leprosy). These pathogens are shielded from antibodies by the host-cell membrane (Figure 34.23). A different and more subtle strategy, *cell-mediated immunity,* evolved to cope with intracellular viral pathogens. *T cells* continually scan the surfaces of all cells and kill those that exhibit foreign markings. The task is not simple; intracellular microorganisms are not so obliging as to intentionally leave telltale traces on the surface of their host. Quite the contrary, successful pathogens are masters of the art of camouflage. Vertebrates have evolved an ingenious mechanism—cut and display—to reveal the presence of stealthy intruders. Nearly all vertebrate cells exhibit on their surfaces a sample of peptides derived from the digestion of proteins in their

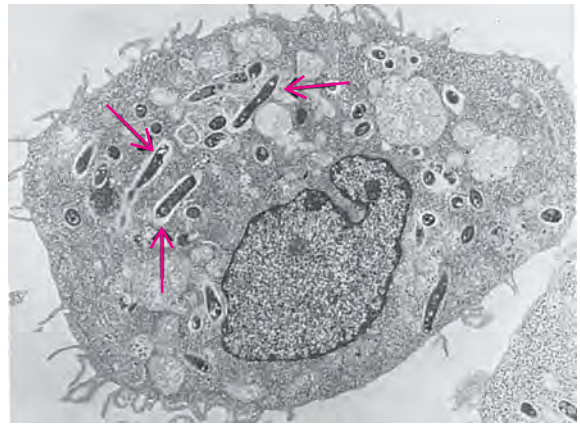


Figure 34.23 Intracellular pathogen. An electron micrograph showing mycobacteria (arrows) inside an infected macrophage. [Courtesy of Dr. Stanley Falkow.]

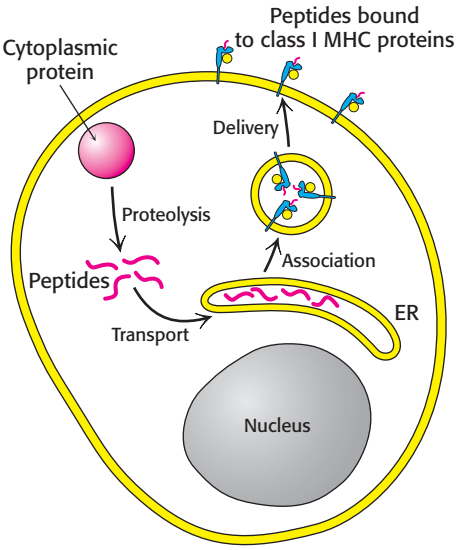


Figure 34.24 Presentation of peptides from cytoplasmic proteins. Class I MHC proteins on the surfaces of most cells display peptides that are derived from cytoplasmic proteins by proteolysis.

cytoplasm. These peptides are displayed by integral membrane proteins that are encoded by the *major histocompatibility complex* (MHC). Specifically, peptides derived from cytoplasmic proteins are bound to and displayed by *class I MHC proteins*. The dendritic cells of the innate immune system that subject pathogens to phagocytosis migrate to lymphatic tissue, where they use an MHC-like mechanism to present foreign peptides or lipid components to T cells—thus linking the innate and adaptive immune responses to pathogens.

How are these peptides generated and delivered to the plasma membrane? The process starts in the cytoplasm with the degradation of proteins—self-proteins as well as those of pathogens (Figure 34.24). Digestion is carried out by proteasomes (Section 23.2). The resulting peptide fragments are transported from the cytoplasm into the lumen of the endoplasmic reticulum by the *TAP protein* (for *t*ransporter associated with *a*ntigen *p*rocessing), a member of the ABC transporter family of ATP-driven pumps (Section 13.2). In the ER, peptides combine with nascent class I MHC proteins; these complexes are then targeted to the plasma membrane.

MHC proteins embedded in the plasma membrane tenaciously grip their bound peptides so that they can be touched and scrutinized by T-cell receptors on the surface of a killer cell. Foreign peptides bound to class I MHC proteins signal that a cell is infected and mark it for destruction by cytotoxic T cells. An assembly consisting of the foreign peptide—MHC complex, the T-cell receptor, and numerous accessory proteins triggers a cascade that induces apoptosis in the infected cell. Strictly speaking, infected cells are not killed but, instead, are triggered to commit suicide to aid the organism.

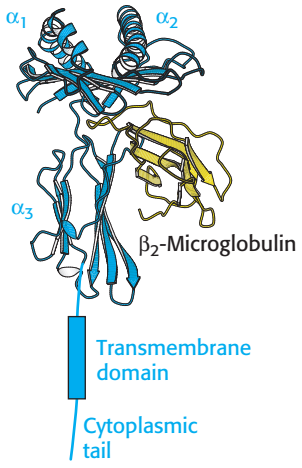


Figure 34.25 Class I MHC protein. A protein of this class consists of two chains. *Notice* that the α chain begins with two domains (α_1, α_2) that include α helices and continues with an immunoglobulin domain (α_3), a transmembrane domain, and a cytoplasmic tail. The second chain, β_2-microglobulin, adopts an immunoglobulin fold. [Drawn from 1HHK.pdb.]

Peptides presented by MHC proteins occupy a deep groove flanked by alpha helices

The three-dimensional structure of a large extracellular fragment of a human MHC class I protein, *human leukocyte antigen A2* (HLA-A2), was solved in 1987 by Don Wiley and Pamela Bjorkman. Class I MHC proteins consist of a 44-kd α chain noncovalently bound to a 12-kd polypeptide called *β_2-microglobulin*. The α chain has three extracellular domains (α_1, α_2, and α_3), a transmembrane segment, and a tail that extends into the cytoplasm (Figure 34.25). The β_2 microglobulin and the α_3 domains have immunoglobulin folds, although the pairing of the two domains differs from that in

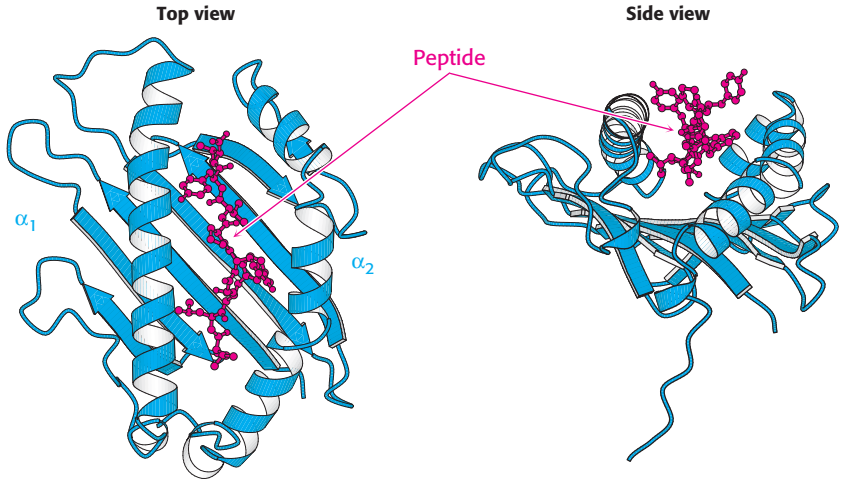
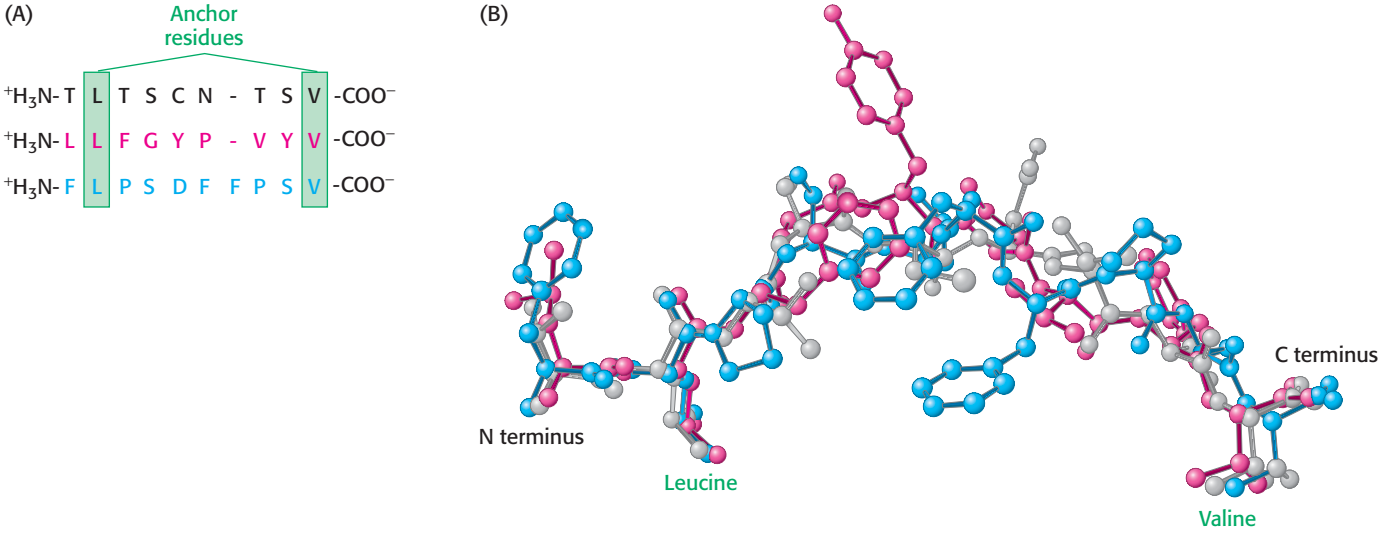


Figure 34.26 Class I MHC peptide-binding site. The α_1 and α_2 domains come together to form a groove in which peptides are displayed. *Notice* that that the peptide is surrounded on three sides by a β sheet and two α helices, but it is accessible from the top of the structure. [Drawn from 1HHK.pdb.]

antibodies. The α_1 and α_2 domains exhibit a novel and remarkable architecture. They associate intimately to form a deep groove that serves as the peptide-binding site (Figure 34.26). The floor of the groove, which is about 25 Å long and 10 Å wide, is formed by eight β strands, four from each domain. A long helix contributed by the α_1 domain forms one side, and a helix contributed by the α_2 domain forms the other side. *This groove is the binding site for the presentation of peptides*.

The groove can be filled by a peptide from 8 to 10 residues long in an extended conformation. As we shall see (p. 998), MHC proteins are remarkably diverse in the human population; each person expresses as many as six distinct class I MHC proteins, and many different forms are present in different people. The first structure determined, HLA-A2, binds peptides that almost always have leucine in the second position and valine in the last position (Figure 34.27). Side chains from the MHC molecule interact with the amino and carboxyl termini and with the side chains in these two key positions. These two residues are often referred to as the *anchor residues*. The other residues are highly variable. Thus, many millions of different

Figure 34.27 Anchor residues. (A) The amino acid sequences of three peptides that bind to the class I MHC protein HLA-A2 are shown. Each of these peptides has leucine in the second position and valine in the carboxyl-terminal position. (B) Comparison of the structures of these peptides reveals that the amino and carboxyl termini, as well as the side chains of the leucine and valine residues, are in essentially the same positions in each peptide, whereas the remainder of the structures are quite different.



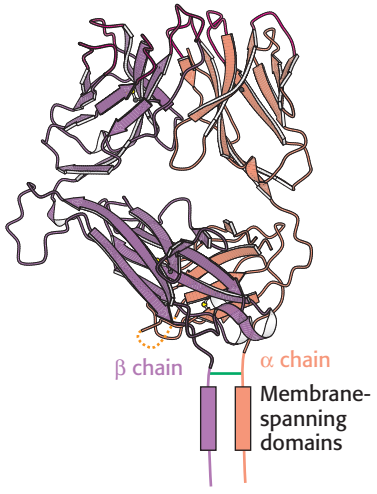


Figure 34.28 T-cell receptor. This protein consists of an α chain and a β chain, linked by a disulfide bond. *Notice* that each chain consists of two immunoglobulin domains on the cell surface, a membrane-spanning domain, and a short cytoplasmic tail. [Drawn from 1BD2.pdb.]

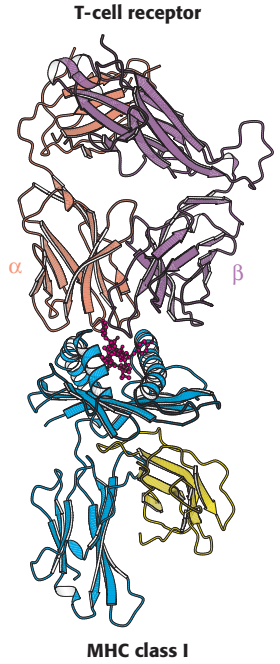


Figure 34.29 T-Cell Receptor–class I MHC complex. The T-cell receptor binds to a class I MHC protein containing a bound peptide. *Notice* that the T-cell receptor contacts both the MHC protein and the peptide. [Drawn from 1BD2.pdb.]

peptides can be presented by this particular class I MHC protein; the identities of only two of the nine residues are crucial for binding. Each class of MHC molecules requires a unique set of anchor residues. Thus, a tremendous range of peptides can be presented by these molecules. Note that *one face of the bound peptide is exposed to solution, where it can be examined by other molecules, particularly T-cell receptors*. An additional remarkable feature of MHC–peptide complexes is their kinetic stability; once bound, a peptide is not released, even after a period of days.

T-cell receptors are antibody-like proteins containing variable and constant regions

We are now ready to consider the receptor that recognizes peptides displayed by MHC proteins on target cells. The *T-cell receptor* consists of a 43-kd α chain joined by a disulfide bond to a 43-kd β chain (Figure 34.28). Each chain spans the plasma membrane and has a short carboxyl-terminal region on the cytoplasmic side. A small proportion of T cells express a receptor consisting of γ and δ chains in place of α and β. The α and β chains of the T-cell receptor, like immunoglobulin L and H chains, consist of *variable* and *constant* regions. Indeed, *these domains of the T-cell receptor are homologous to the V and C domains of immunoglobulins*. Furthermore, hypervariable sequences present in the V regions of the α and β chains of the T-cell receptor form the binding site for the epitope.

The genetic architecture of these proteins is similar to that of immunoglobulins, though the antibody genetic diversity is distributed over all the CDR loops, whereas T-cell-receptor genetic diversity is concentrated in the CDR3 loop that interacts with the peptide bound to the MHC. The variable region of the T-cell receptor α chain is encoded by about 50 V gene segments and 70 J gene segments. The T-cell receptor β chain is encoded by two D gene segments in addition to 57 V gene segments and 13 J gene segments. Again, the diversity of gene segments and the use of slightly imprecise modes of joining them increase the number of distinct proteins formed. *At least 10^{12} different specificities could arise from combinations of this repertoire of genes*. Thus, T-cell receptors, like immunoglobulins, can recognize a very large number of different epitopes. All the receptors on a particular T cell have the same specificity.

How do T cells recognize their targets? The variable regions of the α and β chains of the T-cell receptor form a binding site that recognizes a combined epitope–foreign peptide bound to an MHC protein (Figure 34.29). Neither the foreign peptide alone nor the MHC protein alone forms a complex with the T-cell receptor. Thus, fragments of an intracellular pathogen are presented in a context that allows their detection, leading to the initiation of an appropriate response.

CD8 on cytotoxic T cells acts in concert with T-cell receptors

The T-cell receptor does not act alone in recognizing and mediating the fate of target cells. Cytotoxic T cells also express a protein termed *CD8* on their surfaces that is crucial for the recognition of the class I MHC–peptide complex. The abbreviation *CD* stands for *cluster of differentiation*, referring to a cell-surface marker that is used to identify a lineage or stage of differentiation. Antibodies specific for particular CD proteins have been invaluable in following the development of leukocytes and in discovering new interactions between specific cell types.

Each chain in the CD8 dimer contains a domain that resembles an immunoglobulin variable domain (Figure 34.30). CD8 interacts primarily with the constant α_3 domain of class I MHC proteins. This interaction

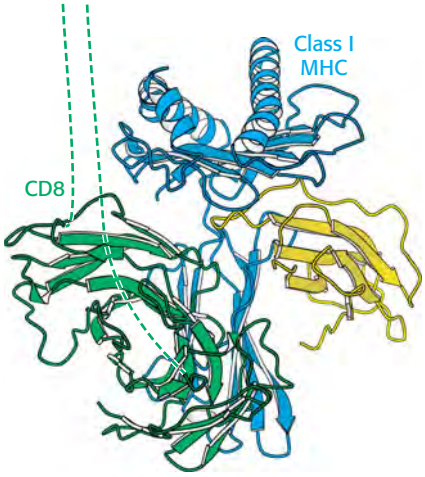


Figure 34.30 The coreceptor CD8. This dimeric protein extends from the surface of a cytotoxic T cell and binds to class I MHC molecules that are expressed on the surface of the cell that is bound to the T cell. The dashed lines represent extended polypeptide chains that link the immunoglobulin domains of CD8 to the membrane. *Notice* that the coreceptor interacts primarily with the constant α_3 domain of the class I MHC domain. [Drawn from 1AKJ.pdb.]

further stabilizes the interactions between the T cell and its target. The cytoplasmic tail of CD8 contains a docking site for Lck, a cytoplasmic tyrosine kinase akin to Src. The T-cell receptor itself is associated with six polypeptides that form the CD3 complex (Figure 34.31). The γ, δ, and ε chains of CD3 are homologous to Ig-α and Ig-β associated with the B-cell receptor (see Figure 34.20); each chain consists of an extracellular immunoglobulin domain and an intracellular ITAM region. These chains associate into CD3-γε and CD3- δε heterodimers. An additional component, the CD3-ζ chain, has only a small extracellular domain and a larger intracellular domain containing three ITAM sequences.

On the basis of these components, a model for T-cell activation can be envisaged that is closely parallel to the pathway for B-cell activation (Section 34.3; Figure 34.32). The binding of the T-cell receptor to the class I MHC–peptide complex and the concomitant binding of CD8 from the T-cell to the MHC molecule link the kinase Lck to the ITAM substrates of the components of the CD3 complex. The phosphorylation of the tyrosine residues in the ITAM sequences generates docking sites for a protein kinase called ZAP-70 (for 70-kd *z*eta-*a*ssociated *p*rotein) that is homologous to Syk in B cells. Docked by its two SH2 domains, ZAP-70 phosphorylates

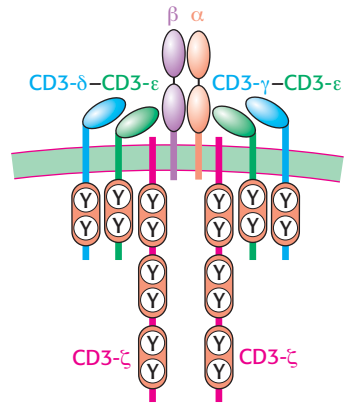


Figure 34.31 T-cell receptor complex. The T-cell receptor is associated with six CD3 molecules: a CD3-γ–CD3-ε heterodimer, a CD3-δ–CD3-ε heterodimer, and two chains of CD3-ζ. Single ITAM sequences are present in the cytoplasmic domains of CD3-γ, CD3-δ, and CD3-ε, whereas three such sequences are found in each CD3-ζ chain.

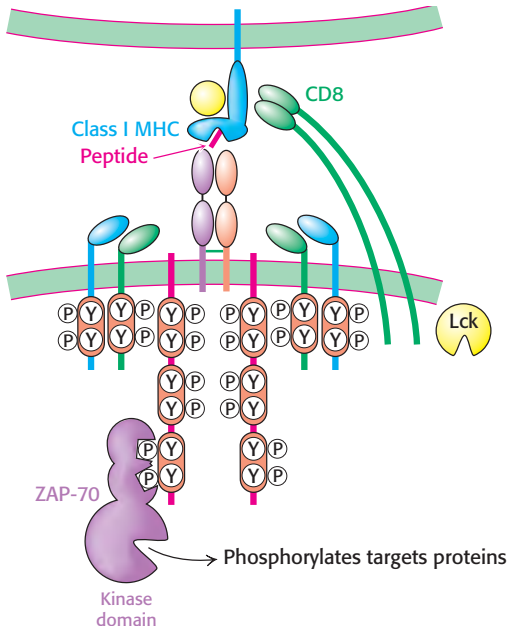


Figure 34.32 T-cell activation. The interaction between the T-cell receptor and a class I MHC–peptide complex results in the binding of CD8 to the MHC protein, the recruitment of the protein tyrosine kinase Lck, and the phosphorylation of tyrosine residues in the ITAM sequences of the CD3 chains. After phosphorylation, the ITAM regions serve as docking sites for the protein kinase ZAP-70, which phosphorylates protein targets to transmit the signal.

downstream targets in the signaling cascade. Additional molecules, including a membrane-bound protein phosphatase called CD45 and a cell-surface protein called CD28, play ancillary roles in this process.

T-cell activation has two important consequences. First, the activation of cytotoxic T cells results in the secretion of several proteins, including *perforin* and *granzymes*. Perforin is a 70-kd protein that destabilizes the plasma membrane of the target cell, enabling the entry of *granzymes* into the cytoplasm of the target cell. Granzymes are serine proteases (Section 9.1) that initiate the pathway of apoptosis, leading to the death of the target cell and the fragmentation of its DNA, including any viral DNA that may be present. Second, after it has stimulated its target cell to commit suicide, the activated T cell disengages and is stimulated to reproduce. Thus, additional T cells that express the same T-cell receptor are generated to continue the battle against the invader after these T cells have been identified as a suitable weapon.

Helper T cells stimulate cells that display foreign peptides bound to class II MHC proteins

Not all T cells are cytotoxic. *Helper T cells, a different class, stimulate the proliferation of specific B lymphocytes and cytotoxic T cells and thereby serve as partners in determining the immune responses that are produced*. The importance of helper T cells is graphically revealed by the devastation wrought by AIDS, a condition that destroys these cells. Helper T cells, like cytotoxic T cells, detect foreign peptides that are presented on cell surfaces by MHC proteins. However, the source of the peptides, the MHC proteins that bind them, and the transport pathway are different.

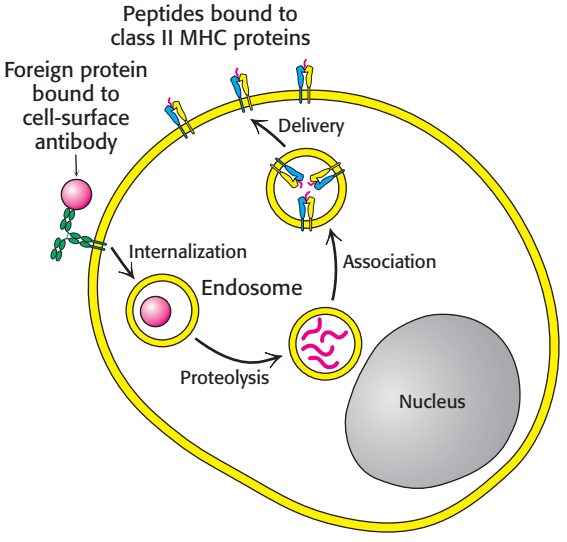


Figure 34.33 Presentation of peptides from internalized proteins. Antigen-presenting cells bind and internalize foreign proteins and display peptides that are formed from the digestion of these proteins in class II MHC proteins.

Helper T cells recognize peptides bound to MHC molecules referred to as class II. Their helping action is focused on B cells, macrophages, and dendritic cells. *Class II MHC proteins are expressed only by these antigen-presenting cells*, unlike class I MHC proteins, which are expressed on nearly all cells. The peptides presented by class II MHC proteins do not come from the cytoplasm. Rather, *they arise from the degradation of proteins that have been internalized by endocytosis*. Consider, for example, a virus particle that is captured by membrane-bound immunoglobulins on the surface of a B cell (Figure 34.33). This complex is delivered to an endosome, a membrane-enclosed acidic compartment, where it is digested. The resulting peptides become associated with class II MHC proteins, which move to the cell surface. Peptides from the cytoplasm cannot reach class II proteins, whereas peptides from endosomal compartments cannot reach class I proteins. This segregation of displayed peptides is biologically critical. The association of a foreign peptide with a class II MHC protein signals that a cell has *encountered* a pathogen and serves as a call for *help*. In contrast, association with a class I MHC protein signals that a cell has *succumbed* to a pathogen and is a call for *destruction*.

Helper T cells rely on the T-cell receptor and CD4 to recognize foreign peptides on antigen-presenting cells

The overall structure of a class II MHC molecule is remarkably similar to that of a class I molecule. Class II molecules consist of a 33-kd α chain and a noncovalently bound 30-kd β chain (Figure 34.34). Each contains two extracellular domains, a transmembrane segment, and a short cytoplasmic tail. The peptide-binding site is formed by the α_1 and β_1 domains, each of

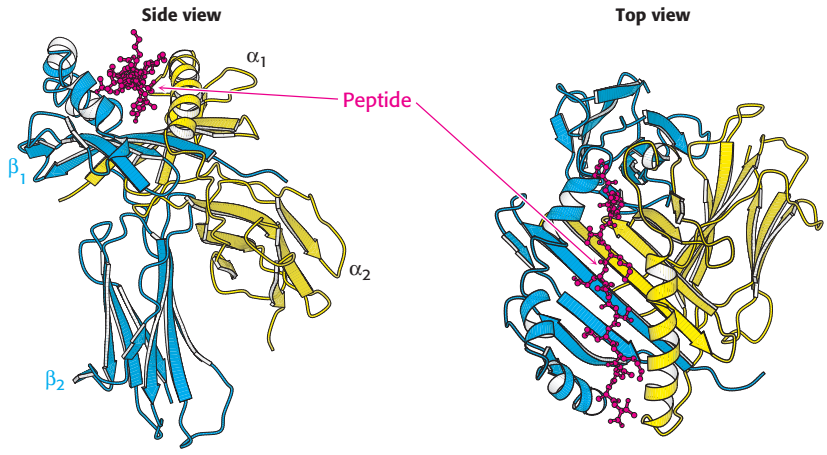


Figure 34.34 Class II MHC protein. A class II MHC protein consists of homologous α and β chains, each of which has an amino-terminal domain that constitutes half of the peptide-binding structure, as well as a carboxyl-terminal immunoglobulin domain. *Notice* the trough-like peptide-binding site, which is similar to that in class I MHC proteins except that it is open at both ends, allowing class II MHC proteins to bind longer peptides than those bound by class I. [Drawn from 1DLH.pdb.]

which contributes a long helix and part of a β sheet. Thus, the same structural elements are present in class I and class II MHC molecules, but they are combined into polypeptide chains in different ways. The peptide-binding site of a class II molecule is open at both ends, and so this groove can accommodate longer peptides than can be bound by class I molecules; typically, peptides between 13 and 18 residues long are bound. The peptide-binding specificity of each class II molecule depends on binding pockets that recognize particular amino acids, also known as anchor residues, in specific positions along the sequence.

Helper T cells express T-cell receptors that are produced from the same genes as those on cytotoxic T cells. These T-cell receptors interact with class II MHC molecules in a manner that is analogous to T-cell-receptor interaction with class I MHC molecules. Nonetheless, helper T cells and cytotoxic T cells are distinguished by other proteins that they express on their surfaces. In particular, helper T cells express a protein called CD4 instead of expressing CD8. *CD4* consists of four immunoglobulin domains that extend from the T-cell surface, as well as a small cytoplasmic region (Figure 34.35). The amino-terminal immunoglobulin domains of CD4 interact with the base on the class II MHC molecule. Thus, helper T cells bind cells expressing class II MHC specifically because of the interactions with CD4 (Figure 34.36).

Figure 34.35 Coreceptor CD4. This protein comprises four tandem immunoglobulin domains that extend from the surface of a helper T cell. [Drawn from 1WIO.pdb.]

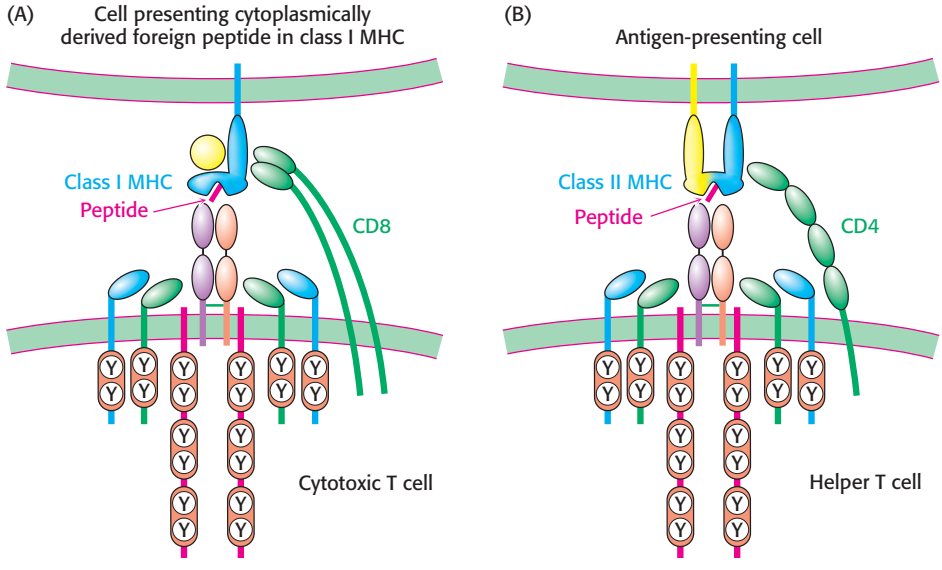


Figure 34.36 Variations on a theme. (A) Cytotoxic T cells recognize foreign peptides presented in class I MHC proteins with the aid of the coreceptor CD8. (B) Helper T cells recognize peptides presented in class II MHC proteins by specialized antigen-presenting cells with the aid of the coreceptor CD4.

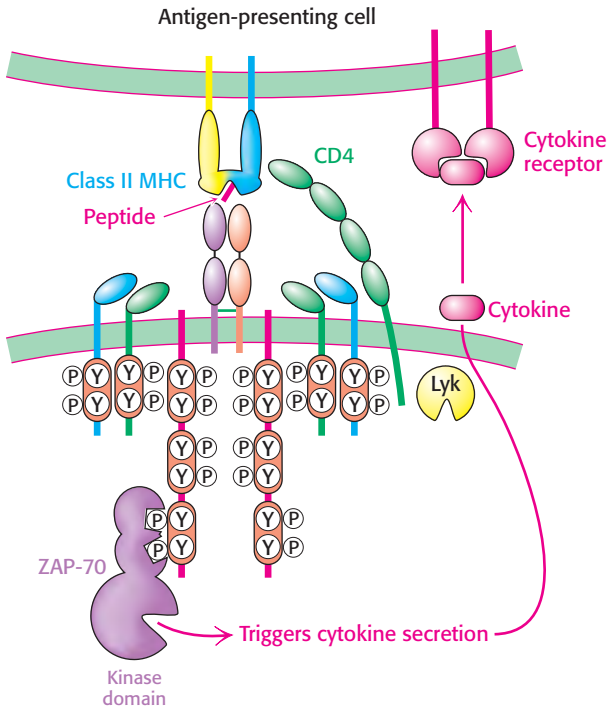


Figure 34.37 Helper-T-cell action. The engagement of the T-cell receptor in helper T cells results in the secretion of cytokines. These cytokines bind to cytokine receptors expressed on the surface of the antigen-presenting cell, stimulating cell growth, differentiation, and, in regard to a B cell, antibody secretion.

When a helper T cell binds to an antigen-presenting cell expressing an appropriate class II MHC–peptide complex, signaling pathways analogous to those in cytotoxic T cells are initiated by the action of the kinase Lck on ITAMs in the CD3 molecules associated with the T-cell receptor. However, rather than triggering events leading to the death of the attached cell, *these signaling pathways result in the secretion of cytokines from the helper cell.* Cytokines are a family of molecules that include, among others, interleukin-2 and interferon-γ. Cytokines bind to specific receptors on the antigen-presenting cell and stimulate growth, differentiation, and, in regard to plasma cells, which are derived from B cells, antibody secretion (Figure 34.37). Thus, the internalization and presentation of parts of a foreign pathogen help to generate a local environment in which cells taking part in the defense against this pathogen can flourish through the action of helper T cells.

MHC proteins are highly diverse

MHC class I and II proteins, the presenters of peptides to T cells, were discovered because of their role in *transplantation rejection.* A tissue transplanted from one person to another or from one mouse to another is usually rejected by the immune system. In contrast, tissues transplanted from one identical twin to another or between mice of an inbred strain are accepted. Genetic analyses revealed that rejection occurs when tissues are transplanted between individual organisms having different genes in the major histocompatibility complex, a cluster of more than 75 genes playing key roles in immunity. The 3500-kb span of the MHC is nearly the length of the entire *E. coli* chromosome. The MHC encodes class I proteins (presenters to cytotoxic T cells) and class II proteins (presenters to helper T cells), as well as class III proteins (components of the complement cascade) and many other proteins that play key roles in immunity.

Human beings express six different class I genes (three from each parent) and six different class II genes. The three loci for class I genes are called HLA-A, -B, and -C; those for class II genes are called HLA-DP, -DQ, and -DR. These loci are *highly polymorphic:* many alleles of each are present in

the population. For example, more than 50 each of HLA-A, -B, and -C alleles are known; the numbers discovered increase each year. Hence, the likelihood that two unrelated persons have identical class I and II proteins is very small ($<10^{-4}$), accounting for transplantation rejection unless the genotypes of donor and acceptor are closely matched in advance.

Differences between class I proteins are located mainly in the α_1 and α_2 domains, which form the peptide-binding site (Figure 34.38). The α_3 domain, which interacts with a constant β_2-microglobulin, is largely conserved. Similarly, the differences between class II proteins cluster near the peptide-binding groove. Why are MHC proteins so highly variable? *Their diversity makes the presentation of a very wide range of peptides to T cells possible*. A *particular* class I or class II molecule may not be able to bind any of the peptide fragments of a viral protein. The likelihood of a fit is markedly increased by having several kinds (usually six) of each class of presenters in each individual organism. If all members of a species had identical class I or class II molecules, the population would be much more vulnerable to devastation by a pathogen that had mutated and thereby evaded presentation. The evolution of the diverse human MHC repertoire has been driven by the selection for individual members of the species who resist infections to which other members of the population may be susceptible.

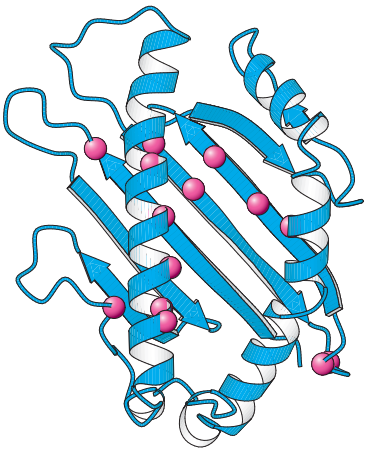


Figure 34.38 Polymorphism in class I MHC proteins. *Notice* that the positions of sites with a high degree of polymorphism in the human population are displayed as red spheres on the structure of the amino-terminal part of a class I MHC protein. [Drawn from 1HHK.pdb.]

Human immunodeficiency viruses subvert the immune system by destroying helper T cells

In 1981, the first cases of a new disease now called *acquired immune deficiency syndrome* (AIDS) were recognized. The victims died of rare infections because their immune systems were crippled. The cause was identified 2 years later by Luc Montagnier and coworkers. AIDS is produced by *human immunodeficiency virus* (HIV), of which two major classes are known: HIV-1 and the much less common HIV-2. Like other *retroviruses*, HIV contains a single-stranded RNA genome that is replicated through a double-stranded DNA intermediate. This viral DNA becomes integrated into the genome of the host cell. In fact, viral genes are transcribed only when they are integrated into the host DNA.

The HIV virion is enveloped by a lipid-bilayer membrane containing two glycoproteins: gp41 spans the membrane and is associated with gp120, which is located on the external face (Figure 34.39). The core of the virus contains two copies of the RNA genome and associated transfer RNAs, as well as several molecules of reverse transcriptase. They are surrounded by many copies of two proteins called p18 and p24. *The host cell for HIV is the helper T cell*. The gp120 molecules on the membrane of HIV bind to CD4 molecules on the surface of the helper T cell (Figure 34.40). This interaction allows the associated viral gp41 to insert its amino-terminal head into the host-cell membrane. The viral membrane and the helper-T-cell membrane fuse, and the viral core is released directly into the cytoplasm. Infection by

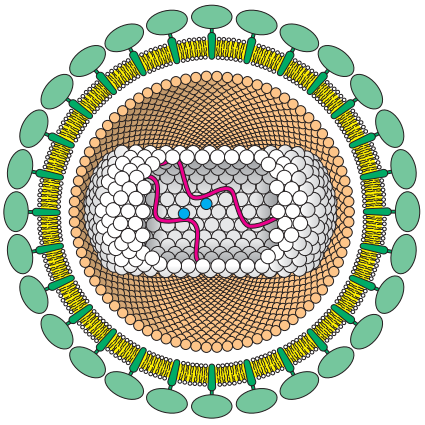


Figure 34.39 Human immunodeficiency virus. A schematic representation of HIV reveals its proteins and nucleic acid components. The membrane-envelope glycoproteins gp41 and gp120 are shown in dark and light green. The viral RNA is shown in red, and molecules of reverse transcriptase are shown in blue. [After R. C. Gallo. The AIDS virus. Copyright © 1987 by Scientific American, Inc. All rights reserved.]

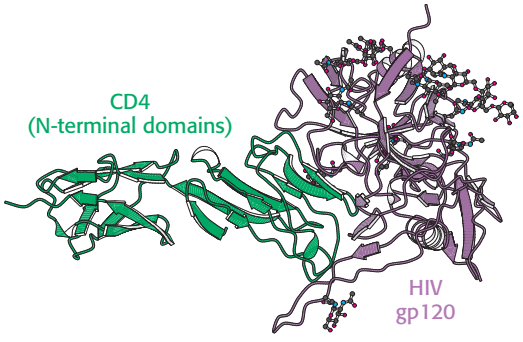


Figure 34.40 HIV receptor. A complex between a modified form of the envelope glycoprotein gp120 from HIV and a peptide corresponding to the two amino-terminal domains from the helper-T-cell protein CD4 reveals how viral infection of helper T cells is initiated. [Drawn from 1GC1.pdb.]

HIV leads to the destruction of helper T cells because the permeability of the host plasma membrane is markedly increased by the insertion of viral glycoproteins and the budding of virus particles. The influx of ions and water disrupts the ionic balance, causing osmotic lysis.

34.5 The Immune System Contributes to the Prevention and the Development of Human Diseases

The primary function of the immune system is to protect the host from invasion by foreign organisms. But how does the immune system prevent itself from mounting attacks against the host organism? In other words, how does the immune system distinguish between self and nonself? Clearly, proteins from the organism itself do not bear some special tag identifying them. Instead, selection processes early in the developmental pathways for immune cells kill or suppress those immune cells that react strongly with self-antigens. The evolutionary paradigm still applies; immune cells that recognize self-antigens are generated, but selective mechanisms eliminate such cells in the course of development.

T cells are subjected to positive and negative selection in the thymus

T cells derive their name from the location of their production—the thymus, a small organ situated just above the heart. Examination of the developmental pathways leading to the production of mature cytotoxic and helper T cells reveals the selection mechanisms that are crucial for distinguishing self from nonself. These selection criteria are quite stringent; approximately 98% of the thymocytes, the precursors of T cells, die before the completion of the maturation process.

Thymocytes produced in the bone marrow do not express the T-cell receptor complex, CD4, or CD8. On relocation to the thymus and rearrangement of the T-cell-receptor genes, the immature thymocyte expresses all of these molecules. These cells are first subjected to *positive selection* (Figure 34.41). Cells for which the T-cell receptor can bind with reasonable affinity to either class I or class II MHC molecules survive this selection; those for which the T-cell receptor does not participate in such an interaction undergo apoptosis and die. *The role of the positive selection step is to prevent the production of T cells that will not bind to any MHC complex present, regardless of the peptide bound.*

The cell population that survives positive selection is subjected to a second step, *negative selection.* Here, T cells that bind with high affinity to MHC complexes bound to self-peptides expressed on the surfaces of

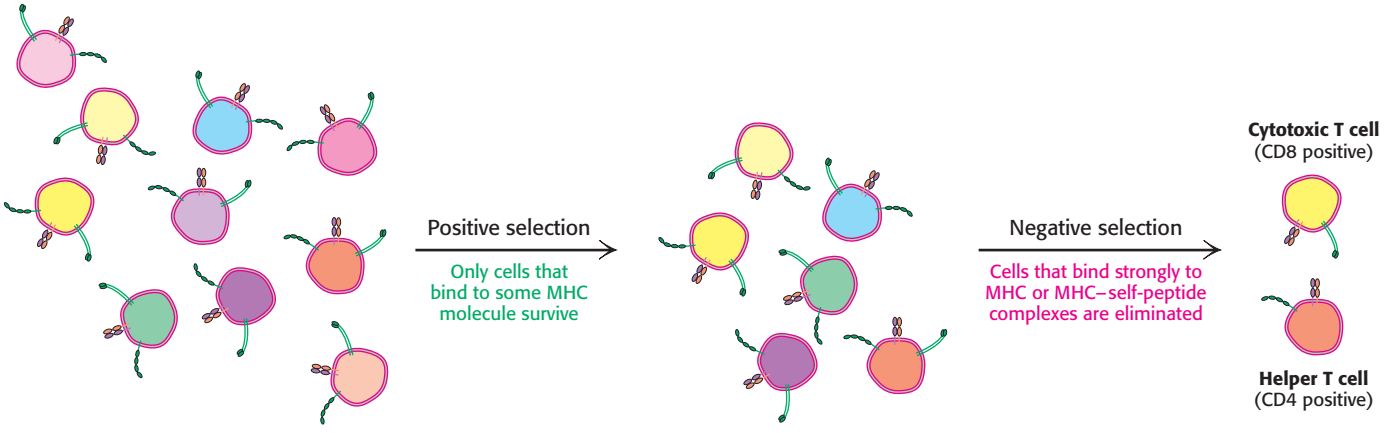


Figure 34.41 T-cell selection. A population of thymocytes is subjected first to positive selection to remove cells that express T-cell receptors that will not bind to MHC proteins expressed by the individual organism. The surviving cells are then subjected to negative selection to remove cells that bind strongly to MHC complexes bound to self-peptides.

antigen-presenting cells in the thymus undergo apoptosis or are otherwise suppressed. Those that do not bind too avidly to any such MHC complex complete development and become mature cytotoxic T cells (which express only CD8) or helper T cells (which express only CD4). The negative selection step leads to *self-tolerance;* cells that bind an MHC–self-peptide complex are removed from the T-cell population. Similar mechanisms apply to developing B cells, suppressing B cells that express antibodies that interact strongly with self-antigens.

Autoimmune diseases result from the generation of immune responses against self-antigens

Although thymic selection is remarkably efficient in suppressing the immune response to self-antigens, failures do occur. Such failures result in *autoimmune diseases.* These diseases include common illnesses such as insulin-dependent diabetes mellitus, multiple sclerosis, and rheumatoid arthritis. In these illnesses, immune responses against self-antigens result in damage to selective tissues that express the antigen (Figure 34.42).

In many cases, the cause of the generation of self-reactive antibodies or T cells is unclear. However, in other cases, infectious organisms such as bacteria or viruses may play a role. Infection leads to the generation of antibodies and T cells that react with many different epitopes from the infectious organism. If one of these antigens closely resembles a self-antigen, an autoimmune response can result. For example, *Streptococcus* infections sometimes lead to rheumatic fever owing to the production of antibodies to streptococcal antigens that cross-react with exposed epitopes in heart muscle.

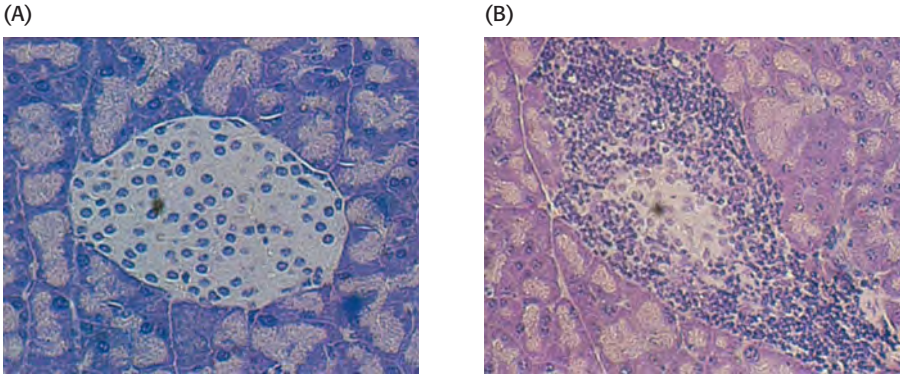


Figure 34.42 Consequences of autoimmunity. Photomicrographs of an islet of Langerhans (A) in the pancreas of a normal mouse and (B) in the pancreas of a mouse with an immune response against pancreatic β cells, which results in a disease resembling insulin-dependent diabetes mellitus in human beings. *Notice* that the relatively pale cellular area in the center of part A is populated with the dark nuclei of inflammatory cells in part B. [From M. A. Atkinson and N. K. Maclaren. What causes diabetes? Copyright © 1990 by Scientific American, Inc. All rights reserved.]

The immune system plays a role in cancer prevention

The development of immune responses against proteins encoded by our own genomes can be beneficial in some circumstances. Cancer cells have undergone significant changes that often result in the expression of proteins that are not normally expressed. For example, the mutation of genes can generate proteins that do not correspond in amino acid sequence to any normal protein. Such proteins may be recognized as foreign, and an immune response will be generated specifically against the cancer cell. Alternatively, cancer cells often produce proteins that are expressed during

embryonic development but are not expressed or are expressed at very low levels after birth. For example, a membrane glycoprotein called *carcinoembryonic antigen* (CEA) appears in the gasterointestinal cells of developing fetuses but is not normally expressed at significant levels after birth. More than 50% of patients with colorectal cancer have elevated serum levels of CEA. Immune cells recognizing epitopes from such proteins will not be subjected to negative selection and, hence, will be present in the adult immune repertoire. These cells may play a cancer surveillance role, killing cells that overexpress antigens such as CEA and preventing genetically damaged cells from developing into tumors.

Vaccines are a powerful means to prevent and eradicate disease

The discovery and utilization of *vaccines*, biological preparations that stimulate immunity to a particular pathogen or disease, represent a significant milestone in modern medical history. The pioneering work of Edward Jenner and Louis Pasteur in the eighteenth and nineteenth centuries revealed that inoculation with inactivated forms of a pathogen could protect against subsequent infection with the active pathogen. Treatment of a person with a vaccine, or *immunization*, leads to the development of *immunological memory*, the ability of the immune system to respond more rapidly and effectively to pathogens that have been encountered previously. At the cellular level, memory arises because positively selected B cells and T cells that have a high affinity for their target persist within the organism as *memory B cells* and *memory T cells*.

In order for a vaccine to be effective, it must generate a sustained protective response appropriate for the targeted pathogen. For example, pathogens such as *Clostridium tetani*, the causative agent of tetanus, cause disease by the release of destructive proteins called *toxins*. Hence, an effective tetanus vaccine should generate a strong B-cell-mediated response. In contrast, a vaccine designed to prevent disease caused by an intracellular pathogen should generate a strong T-cell-mediated response. The active component of a vaccine that is responsible for stimulating this immune response can take many forms:

Killed or *inactivated vaccines* contain pathogens that have been rendered harmless by treatment with chemicals or high heat.

Live attenuated vaccines contain live pathogens that have accumulated mutations so that they are no longer virulent to human cells. These vaccines are most commonly generated by repeated infection of the pathogen in cultured cells until its virulence is lost.

Subunit vaccines contain a purified protein component of the pathogen. Such proteins can be either isolated from infected material (such as blood from chronically infected patients) or generated by recombinant methods.

Toxoid vaccines are used against pathogens that employ an extracellular toxin to cause disease. These vaccines contain a form of the toxin that has been inactivated by treatment with chemicals or high heat.

The effectiveness of immunization efforts has made the global eradication of certain diseases possible. In 1966, 20 million cases of smallpox were reported worldwide. By 1980, coordinated vaccination efforts in regions of endemic disease had enabled the World Health Assembly to declare that smallpox had been eradicated. Nevertheless, despite the success of vaccines in the prevention of many devastating diseases, several pathogens have posed a significant challenge to vaccine development. For example, the

development of an effective HIV vaccine has been complicated by the antigenic diversity of HIV strains. Because its mechanism for replication is prone to error, a population of HIV presents an ever-changing array of coat proteins. Indeed, the mutation rate of HIV is more than 65 times higher than that of influenza virus. Nevertheless, the identification of protective antibodies in individuals resistant to HIV and advancements in vaccine design suggest that the development of an effective HIV vaccine remains a real possibility.

Summary

Two lines of defense against pathogens are the innate immune system and the adaptive immune system. The innate immune system targets features present on many different pathogens but misses those pathogens lacking the targeted features. The adaptive immune system is both more specific and wide reaching. To respond effectively to a vast array of pathogens, this type of immune system must be tremendously adaptable. Adaptation by the adaptive immune system follows the principles of evolution: an enormously diverse set of potentially useful proteins is generated; these proteins are then subjected to intense selection so that only cells that express useful proteins flourish and continue development, until an effective immune response to a specific invader is generated.

34.1 Antibodies Possess Distinct Antigen-Binding and Effector Units

The major immunoglobulin in the serum is immunoglobulin G. An IgG protein is a heterotetramer with two heavy chains and two light chains. Treatment of IgG molecules with proteases such as papain produces three fragments: two F_{ab} fragments that retain antigen-binding activity and an F_c fragment that retains the ability to activate effector functions such as the initiation of the complement cascade. The F_{ab} fragments include the L chain and the amino-terminal half of the H chain; the F_c domain is a dimer consisting of the carboxyl-terminal halves of two H chains. Five different classes of antibody—IgG, IgM, IgA, IgD, and IgE—differ in their heavy chains and, hence, in their effector functions.

34.2 Antibodies Bind Specific Molecules Through Hypervariable Loops

One particular protein fold is found in many of the key proteins of the immune system. The immunoglobulin fold consists of a pair of β sheets that pack against each other, linked by a single disulfide bond. Loops projecting from one end of the structure form a binding surface that can be varied by changing the amino acid sequences within the loops. Domains with immunoglobulin folds are linked to form antibodies and other classes of proteins in the immune system, including T-cell receptors. Two chains come together to form the binding surface of an antibody. Three loops from each domain, the complementarity-determining regions, form an essentially continuous surface that can vary tremendously in shape, charge, and other characteristics to allow particular antibodies to bind to molecules ranging from small molecules to large protein surfaces.

34.3 Diversity Is Generated by Gene Rearrangements

The tremendous diversity of the amino acid sequences of antibodies is generated by segmental rearrangements of genes. For antibody κ light chains, 1 of 40 variable regions is linked to 1 of 5 joining regions. The combined VJ unit is then linked to the constant region. Thousands of

different genes can be generated in this manner. Similar arrays are rearranged to form the genes for the heavy chains, but an additional region called the diversity region lies between the V and the J regions. The combination of L and H chains, each obtained through such rearranged genes, can produce more than 10^8 distinct antibodies. Different classes of antibodies are also generated by gene rearrangements that lead to class switching. Oligomerization of membrane-bound antibody molecules initiates a signal-transduction cascade inside B cells. Key steps in this signaling process include the phosphorylation of specific tyrosine residues in sequences termed immunoreceptor tyrosine-based activation motifs, present in proteins that associate with the membrane-bound antibodies.

34.4 Major-Histocompatibility-Complex Proteins Present Peptide Antigens on Cell Surfaces for Recognition by T-Cell Receptors

Intracellular pathogens such as viruses and mycobacteria cannot be easily detected. Intracellular proteins are constantly being cut into small peptides by proteasomes and displayed in class I major-histocompatibility-complex proteins on cell surfaces. Such peptides lie in a groove defined by two helices in the class I MHC proteins. The combination of MHC protein and peptide can be bound by an appropriate T-cell receptor. T-cell receptors resemble the antigen-binding domains of antibodies in structure, and diversity in T-cell-receptor sequence is generated by V(D)J gene rearrangements. The T-cell receptor recognizes features of both the peptide and the MHC molecule that presents it. Cytotoxic T cells initiate apoptosis in cells to which they bind through interactions between T-cell receptors and class I MHC–peptide complexes aided by interactions with the coreceptor molecule CD8. Helper T cells recognize peptides presented in class II MHC proteins, a distinct type of MHC protein expressed only on antigen-presenting cells, such as B cells and macrophages. Helper T cells express the coreceptor CD4 rather than CD8. CD4 interacts with class II MHC proteins present on antigen-presenting cells. Signaling pathways, analogous to those in B cells, are initiated by interactions between MHC–peptide complexes and T-cell receptors and the CD8 and CD4 coreceptors. Human immunodeficiency virus damages the immune system by infecting cells that express CD4, such as helper T cells.

34.5 The Immune System Contributes to the Prevention and the Development of Human Diseases

In principle, the immune system is capable of generating antibodies and T-cell receptors that bind to self-molecules—that is, molecules that are normally present in a healthy and uninfected individual organism. Selection mechanisms prevent such self-directed molecules from being expressed at high levels. The selection process includes both positive selection, to enrich the population of cells that express molecules that have the potential to bind foreign antigens in an appropriate context, and negative selection, which eliminates cells that express molecules with too high an affinity for self-antigens. Autoimmune diseases such as insulin-dependent diabetes mellitus can result from the amplification of a response against a self-antigen. Vaccines stimulate immunological memory so as to prevent, and even eradicate, disease.

Key Terms

innate immune system (p. 977)
adaptive immune system (p. 977)
phagocyte (p. 978)
Toll-like receptor (TLR) (p. 978)
pathogen-associated molecular pattern (PAMP) (p. 978)

endotoxin (p. 978)
humoral immune response (p. 979)
B lymphocyte (B cell) (p. 979)
antigen (p. 979)
antigenic determinant (epitope) (p. 980)
cellular immune response (p. 980)
cytotoxic T lymphocyte (killer T cell) (p. 980)
helper T lymphocyte (p. 980)
immunoglobulin G (IgG) (p. 981)
F_{ab} (p. 981)
F_c (p. 981)
light (L) chain (p. 981)
heavy (H) chain (p. 981)
segmental flexibility (p. 982)
immunoglobulin M (IgM) (p. 982)
immunoglobulin A (IgA) (p. 983)
immunoglobulin D (IgD) (p. 983)
immunoglobulin E (IgE) (p. 983)
variable region (p. 983)
constant region (p. 984)
immunoglobulin fold (p. 984)
hypervariable loop (p. 984)
complementarity-determining region (CDR) (p. 984)
V(D)J recombination (p. 988)
immunoreceptor tyrosine-based activation motif (ITAM) (p. 989)
cyclosporin (p. 990)
hapten (p. 990)
class switching (p. 991)
T cell (p. 991)
major histocompatibility complex (MHC) (p. 992)
class I MHC protein (p. 992)
TAP (transporter associated with antigen processing) protein (p. 992)
human leukocyte antigen (HLA) (p. 992)
β_2-microglobulin (p. 992)
T-cell receptor (p. 994)
CD8 (p. 994)
perforin (p. 996)
granzyme (p. 996)
helper T cell (p. 996)
class II MHC protein (p. 996)
CD4 (p. 997)
human immunodeficiency virus (HIV) (p. 999)
positive selection (p. 1000)
negative selection (p. 1000)
autoimmune disease (p. 1001)
carcinoembryonic antigen (CEA) (p. 1002)
vaccine (p. 1002)
immunological memory (p. 1002)
killed, or inactivated, vaccines (p. 1002)
live attenuated vaccines (p. 1002)
subunit vaccines (p. 1002)
toxoid vaccines (p. 1002)

Problems

1. *First things first.* Distinguish between the innate and adaptive immune systems.

2. *Antibody diversity.* What are the mechanisms used by B cells to generate antibody diversity?

3. *Hang in there.* Explain the difference between *affinity* and *avidity.* For which immunoglobulin class might avidity be particularly important in antigen recognition?

4. *Innate abilities.* A strain of mice has been identified that does not respond to LPS. This lack of response is due to a single amino acid change in the intracellular domain of mouse TLR-4. Propose an explanation for the lack of response.

5. *TLR ligands.* The PAMP recognized by TLR-3 is double-stranded RNA (dsRNA). Against which pathogens would TLR-3 be an effective immune receptor?

6. *Energetics and kinetics.* Suppose that the dissociation constant of an F_{ab}–hapten complex is 3×10^{-7} M at 25°C.

(a) What is the standard free energy of binding?

(b) Immunologists often speak of affinity (K_a), the reciprocal of the dissociation constant, in comparing antibodies. What is the affinity of this F_{ab}?

(c) The rate constant for the release of hapten from the complex is $120\ s^{-1}$. What is the rate constant for association? What does the magnitude of this value imply about the extent of structural change in the antibody on binding hapten?

7. *A brilliant emitter.* Certain naphthalene derivatives, such as the dansyl group, exhibit a weak yellow fluorescence when they are in a highly polar environment (such as water) and an intense blue fluorescence when they are in a markedly nonpolar environment (such as hexane). The binding of ε-dansyl-lysine to specific antibody is accompanied by a marked increase in its fluorescence intensity and a shift in color from yellow to blue. What does this finding reveal about the hapten–antibody complex?

8. *Miniantibody.* The F_{ab} fragment of an antibody molecule has essentially the same affinity for a monovalent hapten as does intact IgG.

(a) What is the smallest unit of an antibody that can retain the specificity and binding affinity of the whole protein?

(b) Design a compact single-chain protein that is likely to specifically bind antigen with high affinity.

9. *Turning on B cells.* B lymphocytes, the precursors of plasma cells, are triggered to proliferate by the binding of multivalent antigens to receptors on their surfaces. The cell-surface receptors are transmembrane immunoglobulins. Univalent antigens, in contrast, do not activate B cells.

(a) What do these findings reveal about the mechanism of B-cell activation?

(b) How might antibodies be used to activate B cells?

10. *An ingenious cloning strategy.* In the cloning of the gene for the α chain of the T-cell receptor, T-cell cDNAs were

hybridized with B-cell mRNAs. What was the purpose of this hybridization step? Can the principle be applied generally?

11. *Pathogen susceptibility.* Patients carrying specific mutations in the gene encoding the TLR-4 protein are susceptible to infections from Gram-negative bacteria. Why are these patients vulnerable to this particular type of pathogen?

12. *Matchmaker, matchmaker.* Why is it important to match HLA alleles between donor and recipient in organ transplantation?

13. *Instruction.* Before the mechanism for generating antibody diversity had been established, a mechanism based on protein folding around an antigen was proposed, primarily by Linus Pauling. In this model, antibodies that had different specificities had the same amino acid sequence but were folded in different ways. Propose a test of this model.

14. *Dealing with nonsense.* Cells, including immune cells, degrade mRNA molecules in which no long open reading frame is present. The process is called nonsense-mediated RNA decay. Suggest a role for this process in immune cells.

15. *Down, but not out.* To understand the genes responsible for growth and infectivity in a disease-causing bacterial strain, you perform chemical mutagenesis on a culture of these bacteria. In the course of your investigation into the properties of the resulting mutants, you identify a set of mutant bacteria that is still viable but their virulence is significantly impaired. How might these mutants be useful for vaccine development?

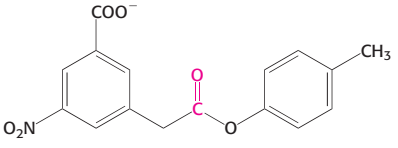
16. *Presentation.* The amino acid sequence of a small protein is

MSRLASKNLIRSDHAGGLLQATYSAVSS-
IKNTMSFGAWSNAALNDSRDA

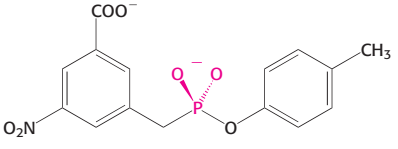
Predict the most likely peptide to be presented by the class I MHC molecule HLA-A2.

Mechanism Problem

17. *Catalytic antibody.* Antibody is generated against a transition state for the hydrolysis of the following ester.



Ester



Transition-state analog

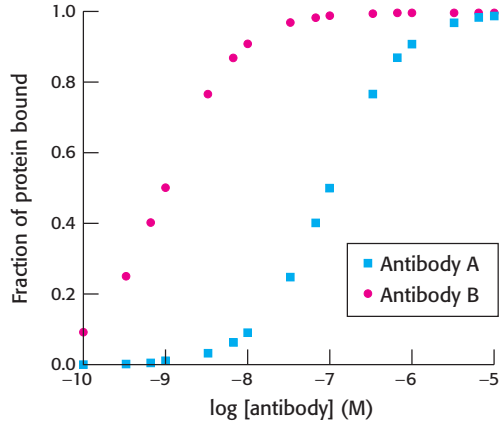
Some of these antibodies catalyze the hydrolysis of the ester. What amino acid residue might you expect to find in the binding site on the antibody?

Chapter Integration Problem

18. *Signaling.* Protein tyrosine phosphatases, such as the molecule CD45 expressed in both B cells and T cells, play important roles in activating such protein tyrosine kinases as Fyn and Lck, which are quite similar to Src. Suggest a mechanism for the activation of such protein kinases by the removal of a phosphoryl group from a phosphotyrosine residue.

Data Interpretation Problem

19. *Affinity maturation.* A mouse is immunized with an oligomeric human protein. Shortly after immunization, a cell line that expresses a single type of antibody molecule (antibody A) is derived. The ability of antibody A to bind the human protein is assayed with the results shown in the graph below.



After repeated immunizations with the same protein, another cell line is derived that expresses a different antibody (antibody B). The results of analyzing the binding of antibody B to the protein also are shown. From these data, estimate

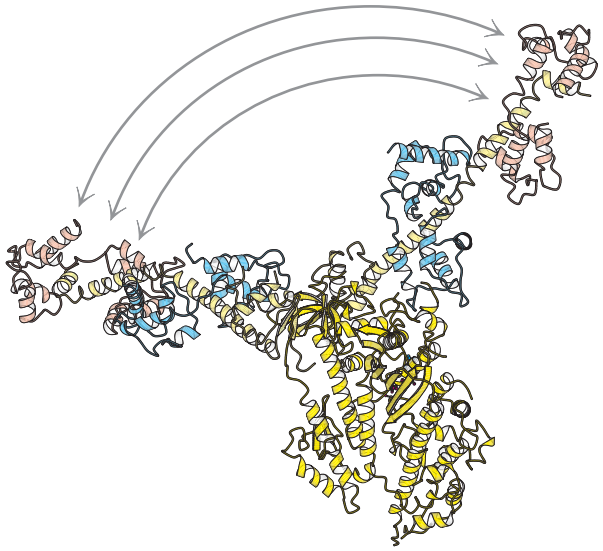
(a) the dissociation constant (K_d) for the complex between the protein and antibody A.

(b) the dissociation constant for the complex between the protein and antibody B.

Comparison of the amino acid sequences of antibody A and antibody B reveals them to be identical except for a single amino acid. What does this finding suggest about the mechanism by which the gene encoding antibody B was generated?

CHAPTER 35

Molecular Motors



The powerful muscles of the horse, like the muscles of all animals, is powered by the molecular-motor protein myosin. A part of myosin moves dramatically (as shown above) in response to ATP binding, hydrolysis, and product release, propelling myosin along an actin filament. This molecular movement is translated into movement of the entire animal, vividly depicted in da Vinci's rearing horse. [(Left) Leonardo da Vinci's study of a rearing horse for the *Battle of Anghiari* (ca. 1504) from The Royal Collection © Her Royal Majesty Queen Elizabeth II.]

Organisms, from human beings to bacteria, move to adapt to changes in their environments, navigating toward food and away from danger. Cells themselves are not static but are bustling assemblies of moving proteins, nucleic acids, and organelles. This motion is enabled by two elements: molecular-motor proteins and complex networks of filamentous proteins termed the *cytoskeleton* (Figure 35.1). The dynamic networks that determine the shape and mobility of calls are among the most active areas of investigation in modern cell biology. Remarkably, the fundamental biochemical mechanisms that produce contractions in our muscles are the same as those that propel organelles along the cytoskeleton. In fact, many of the proteins that play key roles in converting chemical energy into kinetic energy are members of the same protein family, the P-loop NTPases, the hugely important group of proteins that we first examined in Chapter 9. These molecular motors are homologous to proteins that we have encountered in other contexts, including the G proteins in protein synthesis, signaling, and other processes. Once again, we see the economy of evolution in adapting existing proteins to perform new functions.

Molecular motors operate by small increments, converting changes in protein conformation into directed motion. Orderly motion across distances requires a track that steers the motion of the motor assembly. Indeed, we have already encountered a class of molecular motors that utilize mechanisms

OUTLINE

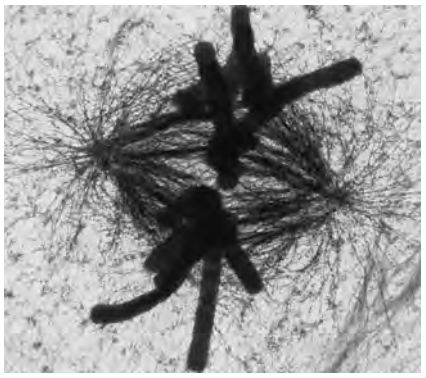


Figure 35.1 Motion within cells. This high-voltage electron micrograph shows the mitotic apparatus in a metaphase mammalian cell. The large cylindrical objects are chromosomes, and the threadlike structures stretched across the center are microtubules, key components of the skeleton. Microtubules serve as tracks for the molecular motors that move chromosomes. Many processes, including chromosome segregation in mitosis, depend on the action of molecular-motor proteins. [Courtesy of Dr. J. R. McIntosh.]

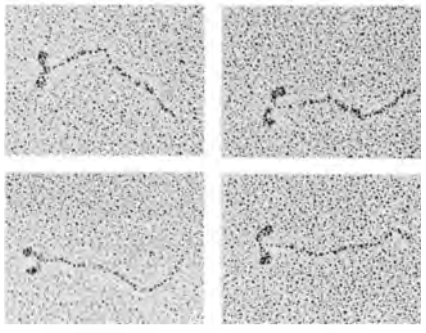


Figure 35.2 Myosin structure at low resolution. Electron micrographs of myosin molecules reveal a two-headed structure with a long, thin tail. [Courtesy of Dr. Paula Flicker, Dr. Theo Walliman, and Dr. Peter Vibert.]

that we will examine here—namely, the helicases that move along DNA during DNA replication (Section 28.1). The proteins on which we will focus in this chapter move along actin and microtubules—protein filaments composed of repeating subunits. The motor proteins cycle between forms having high or low affinity for the filament tracks in response to ATP binding and hydrolysis, enabling a bind, pull, and release mechanism that generates motion.

We will also consider a completely different strategy for generating motion, one used by bacteria such as *E. coli*. A set of flagella act as propellers, rotated by a motor in the bacterial cell membrane. This rotary motor is driven by a proton gradient across the membrane, rather than by ATP hydrolysis. The mechanism for coupling the proton gradient to rotatory motion is analogous to that used by the F_0 subunit of ATP synthase. Thus, both of the major modes for storing biochemical energy—namely, ATP and ion gradients—have been harnessed by evolution to drive organized molecular motion.

35.1 Most Molecular-Motor Proteins Are Members of the P-Loop NTPase Superfamily

Eukaryotic cells contain three major families of motor proteins: myosins, kinesins, and dyneins. Members of each of these classes move along components of the cytoskeleton, but, at first glance, these protein families appear to be quite different from one another. *Myosin*, first characterized on the basis of its role in muscle, moves along filaments of the protein actin. Each molecule of muscle myosin consists of two copies each of a *heavy chain* with a molecular mass of 220 kd, an *essential light chain*, and a *regulatory light chain*. The human genome encodes more than 40 distinct myosins; some function in muscle contraction, and others participate in a variety of other processes. *Kinesins,* which have roles in protein, mRNA, and vesicle transport as well as construction of the mitotic spindle and chromosome segregation, are generally dimers of two polypeptides. The human genome encodes more than 40 kinesins. *Dyneins* power the motion of cilia and flagella, and a general cytoplasmic dynein contributes to a variety of motions in all cells, including vesicle transport and various transport events in mitosis. Dyneins are enormous, with heavy chains of molecular mass greater than 500 kd. The human genome encodes approximately 10 dyneins.

Initially, comparison of the amino acid sequences of myosins, kinesins, and dyneins did not reveal significant relationships between these protein families but, after their three-dimensional structures were determined, members of the myosin and kinesin families were found to have remarkable similarities. In particular, both myosin and kinesin contain P-loop NTPase cores homologous to those found in G proteins. Sequence analysis of the dynein heavy chain reveals it to be a member of the AAA subfamily of P-loop NTPases that we encountered in the context of the 19S proteasome (Section 23.2). Dynein has six sequences encoding such P-loop NTPase domains arrayed along its length, although only four actually appear to bind nucleotides. Thus, we expect similarities in the mechanisms of action, and we can draw on our knowledge of P-loop NTPases in general as we analyze the mechanisms of action of these motor proteins.

Molecular motors are generally oligomeric proteins with an ATPase core and an extended structure

Let us first consider the structure of myosin, which we examined briefly in Chapter 9. The results of electron microscopic studies of skeletal-muscle

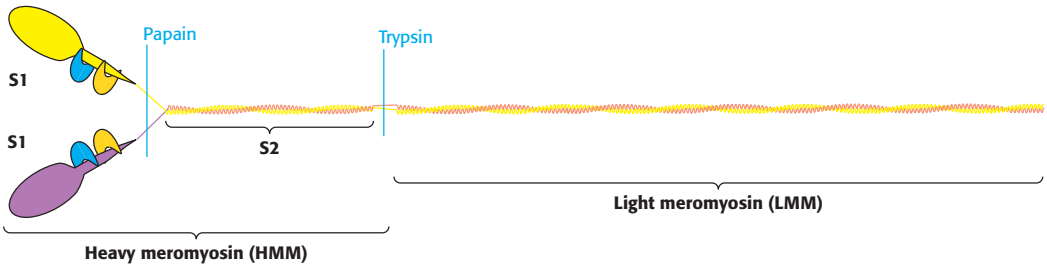


Figure 35.3 Myosin dissection. Treatment of muscle myosin with proteases forms stable fragments, including subfragments S1 and S2 and light meromyosin. Each S1 fragment includes a head (shown in yellow or purple) from the heavy chain and one copy of each light chain (shown in blue and orange).

myosin show it to be a two-headed structure linked to a long stalk (Figure 35.2). The treatment of myosin with trypsin and papain results in the formation of four fragments: two S1 fragments; heavy meromyosin (HMM) which consists of the S1 fragments and an additional region termed S2; and a fragment called light meromyosin (LMM; Figure 35.3). Each *S1 fragment* corresponds to one of the heads of the intact structure and includes 850 amino-terminal amino acids from one of the two heavy chains as well as one copy of each of the light chains. Examination of the structure of an S1 fragment at high resolution reveals the presence of the P-loop NTPase-domain core that is the site of ATP binding and hydrolysis (Figure 35.4). We examined the structure and mechanism of action of this motor domain in Chapter 9.

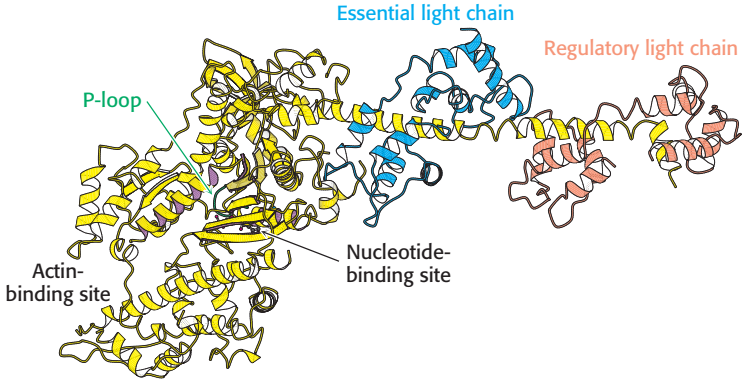


Figure 35.4 Myosin structure at high resolution. The structure of the S1 fragment from muscle myosin reveals the presence of a P-loop NTPase domain (shaded in purple). *Notice* that an α helix extending from this domain is the binding site for the two light chains. [Drawn from 1DFL.pdb.]

Extending away from this structure is a long α helix from the heavy chain. This helix is the binding site for the two light chains. The light chains are members of the EF-hand family, similar to calmodulin (see Figure 14.16), although most of the EF hands in light chains do not bind metal ions (Figure 35.5).

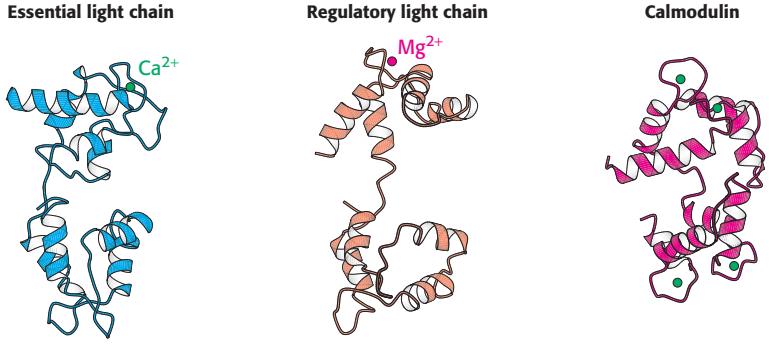


Figure 35.5 Myosin light chains. The structures of the essential and regulatory light chains of muscle myosin are compared with the structure of calmodulin. *Notice* the similarities in the structures that allow each of these homologous proteins to bind an α helix (not shown) by wrapping around it. [Drawn from 1DFL.pdb and 1CM1.pdb.]

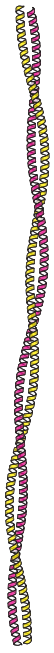


Figure 35.6 Myosin two-stranded coiled coil. The two α helices form left-handed supercoiled structures that spiral around each other. Such structures are stabilized by hydrophobic residues at the contact points between the two helices. [Drawn from 2TMA.pdb.]

Like calmodulin, these proteins wrap around an α helix, serving to thicken and stiffen it. The remaining fragments of myosin—S2 and light meromyosin—are largely α helical, forming two-stranded coiled coils created by the remaining lengths of the two heavy chains wrapping around each other (Figure 35.6). These structures, together extending approximately 1700 Å, link the myosin heads to other structures. In muscle myosin, several LMM domains come together to form higher-order bundles.

Conventional kinesin (kinesin 1), the first kinesin discovered, has several structural features in common with myosin. The dimeric protein has two heads connected by an extended structure (Figure 35.7). The size of the head domain is approximately one-third that of myosin. Determination of the three-dimensional structure of a kinesin fragment revealed that the head domain also is built around a P-loop NTPase core (Figure 35.8). The myosin domain is so much larger than that of kinesin because of two large insertions in the myosin domain that bind to actin filaments. For conventional kinesin, a region of approximately 500 amino acids extends from the head domain. Like the corresponding region in myosin, the extended part of kinesin forms an α-helical coiled coil. Conventional kinesin also has light chains, but, unlike those of myosin, these light chains bind near the carboxyl terminus of the heavy chain and are thought to link the motor to intracellular cargo.

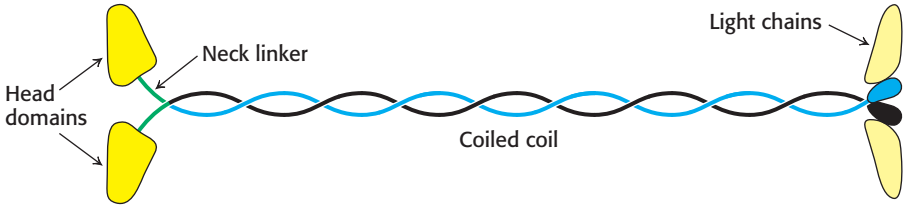


Figure 35.7 Structure of kinesin. The elongated structure has the head domains at one end and the cargo-binding domains at the other, linked by a long coiled-coil region.

Dynein has a somewhat different structure. As noted earlier, the dynein heavy chain includes six regions that are homologous to the AAA subfamily of ATPase domains. Although no crystallographic data are yet available, the results of electron microscopic studies and comparison with known structures of other AAA ATPases have formed the basis for the construction of a model of the dynein structure (Figure 35.9). The head domain is appended to a region of approximately 1300 amino acids that forms an extended structure that links dynein units together to form oligomers and interacts with other proteins.

Although the structures of these three classes of molecular motors have significant differences, some common features emerge. Each structure is dimeric with two head domains, has regions of extended but quite rigid structures, and has regions for interacting with other proteins. As we shall see, these structures are suitable for actions that resemble climbing a rope, hand over hand. The regions for interacting with other proteins represent the grasping hands, the extended structures represent the arms that act as levers to promote larger-scale motion, and the head domains are the engines that provide the necessary mechanical energy.

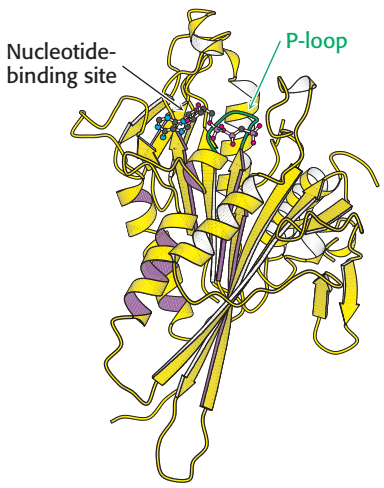


Figure 35.8 Structure of head domain of kinesin at high resolution. *Notice* that the head domain of kinesin has the structure of a P-loop NTPase core (indicated by purple shading). [Drawn from 1I6I.pdb.]

ATP binding and hydrolysis induce changes in the conformation and binding affinity of motor proteins

In Chapter 9, we examined the conformational changes that take place in the myosin ATPase domain from the slime mold *Dictyostelium*. The struc-

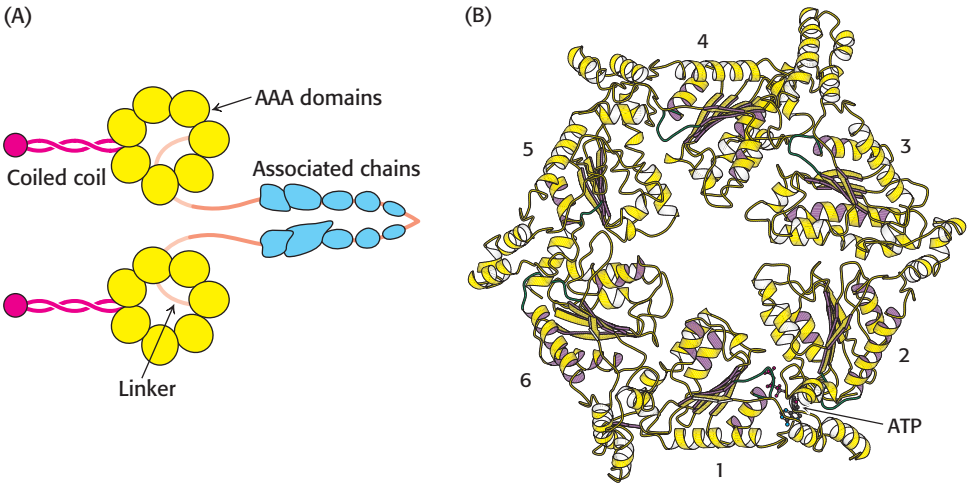


Figure 35.9 Dynein structure. (A) The overall structure of dynein. (B) A model of the motor domain of dynein. *Notice* the six P-loop NTPase domains, some of which bind to and hydrolyze ATP. [Drawn from 1HN5.pdb.]

tures of myosin ATP domains from other sources have been elucidated in a variety of forms as well. The S1 fragment of myosin from scallop muscle provides a striking example of the changes observed (Figure 35.10). The structure of this S1 fragment has been determined in a number of forms including that without bound nucleotide and that bound to a complex formed of ADP and vanadate (VO_4^{3-}), which, as mentioned in Chapter 9, is an analog of the ATP-hydrolysis transition state. The long helix that binds the light chains (hereafter referred to as the *lever arm*) protrudes outward from the head domain. Comparison of the structures reveals that the lever arm has rotated by nearly 90 degrees in the ADP–VO_4^{3-} complex compared with its position in the nucleotide-free form. How does the species in the nucleotide-binding site cause this dramatic transition? Two regions around the nucleotide-binding site (termed *switch I* and *switch II*) tightly conform to the shape of the γ-phosphoryl group of the ATP analog

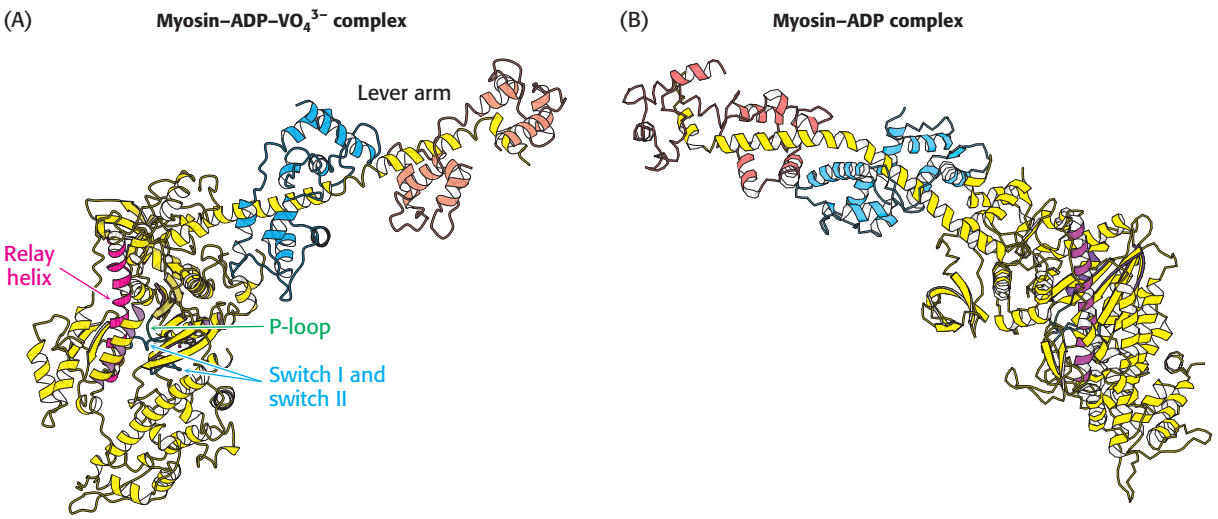


Figure 35.10 Lever-arm motion. Two forms of the S1 fragment of scallop-muscle myosin. *Notice* the dramatic conformational changes when the identity of the bound nucleotide changes from the ADP–VO_4^{3-} complex to nucleotide-free form or vice versa, including a nearly 90-degree reorientation of the lever arm. [Drawn from 1DFL.pdb and 1SR6.pdb.]

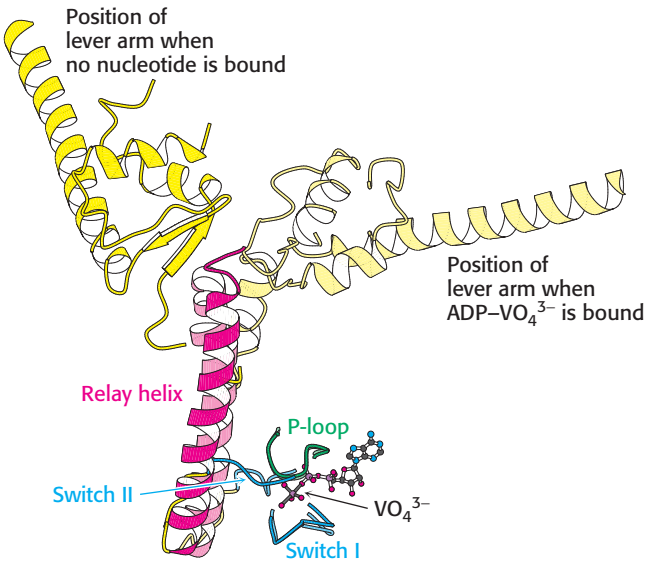


Figure 35.11 Relay helix. A superposition of key elements in two forms of scallop myosin reveals the structural changes that are transmitted by the relay helix from the switch I and switch II loops to the base of the lever arm. *Notice* that the switch I and switch II loops interact with VO_4^{3-} in the position that would be occupied by the γ-phosphoryl group of ATP. The structure of the myosin–ADP–VO_4^{3-} complex is shown in lighter colors. [Drawn from 1DFL.pdb and 1SR6.pdb.]

and adopt a looser conformation when the γ-phosphoryl group is absent (Figure 35.11). This conformational change allows a long α helix (termed the *relay helix*) to adjust its position. The carboxyl-terminal end of the relay helix interacts with structures at the base of the lever arm, and so a change in the position of the relay helix leads to a reorientation of the lever arm.

Analogous conformational changes take place in kinesin. The kinesins also have a relay helix that can adopt different configurations when kinesin binds different nucleotides. Kinesin lacks an α-helical lever arm, however. Instead, a relatively short segment termed the *neck linker* changes conformation in response to nucleotide binding (Figure 35.12). The neck linker binds to the head domain of kinesin when ATP is bound but is released when the nucleotide-binding site is vacant or occupied by ADP.

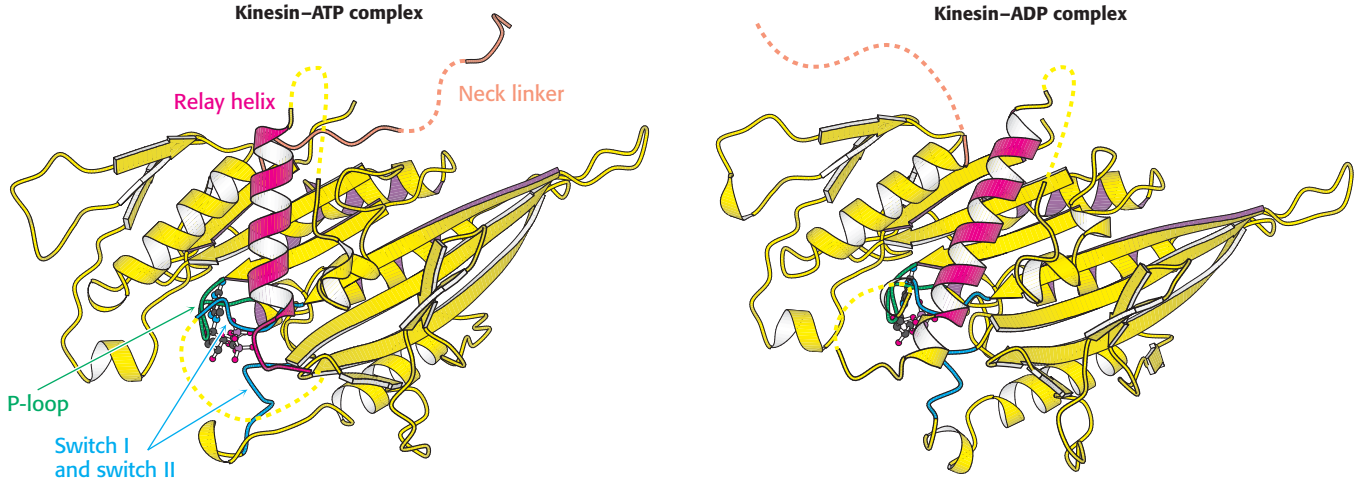
35.2 Myosins Move Along Actin Filaments

Myosins, kinesins, and dyneins move by cycling between states with different affinities for the long, polymeric macromolecules that serve as tracks along which they move. For myosin, the molecular track is a polymeric form of *actin*, a 42-kd protein that is among the most abundant proteins in eukaryotic cells, typically accounting for as much as 10% of the total protein. We begin with a general discussion of the polymeric structure of actin and its assembly. We then examine the interactions between myosin and actin, including both structure and the dynamic interactions between these two proteins. Finally, we turn to the structure of muscle and the roles of myosin and actin in muscle contraction.

Figure 35.12 Neck linker. A comparison of the structures of a kinesin bound to ADP and bound to an ATP analog. *Notice* that the neck linker (orange), which connects the head domain to the remainder of the kinesin molecule, is bound to the head domain in the presence of the ATP analog but is free in the presence of ADP only. [Drawn from 1I6I.pdb and 1I5S.pdb.]

Actin is a polar, self-assembling, dynamic polymer

The structure of the actin monomer has been determined to atomic resolution by x-ray crystallography and has been used to interpret the structure of actin filaments, already somewhat understood through electron microscopy studies at lower resolution. Each actin monomer comprises four domains (Figure 35.13). These domains come together to surround a bound nucleotide,



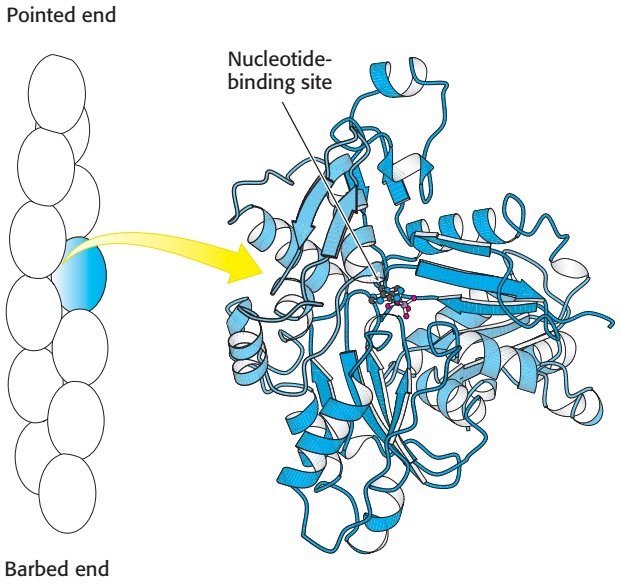


Figure 35.13 Actin structure. (Left) Schematic view of actin monomers (one in blue) of an actin filament. (Right) The domains in the four-domain structure of an actin monomer are identified by different shades of blue. *Notice* the nucleotide-binding site at the center of the structure. [Drawn from 1J6Z.pdb.]

either ATP or ADP. The ATP form can be converted into the ADP form by hydrolysis.

Actin monomers (often called *G-actin* for globular) come together to form actin filaments (often called *F-actin;* see Figure 35.13). F-actin has a helical structure; each monomer is related to the preceding one by a translation of 27.5 Å and a rotation of 166 degrees around the helical axis. Because the rotation is nearly 180 degrees, F-actin resembles a two-stranded cable. Note that each actin monomer is oriented in the same direction along the F-actin filament, and so the structure is polar, with discernibly different ends. One end is called the barbed (plus) end, and the other is called the pointed (minus) end. The names "barbed" and "pointed" refer to the appearance of an actin filament when myosin S1 fragments are bound to it.

How are actin filaments formed? Like many biological structures, actin filaments self-assemble; that is, under appropriate conditions, actin monomers will come together to form well-structured, polar filaments. The aggregation of the first two or three monomers to form a filament is highly unfavorable. Thus, specialized protein complexes, including one called Arp2/3, serve as nuclei for actin assembly in cells. Once such a filament nucleus exists, the addition of subunits is more favorable. Let us consider the polymerization reaction in more detail. We designate an actin filament with n subunits A_n. This filament can bind an additional actin monomer, A, to form A_{n+1}.

$$K_d = \frac{[A_n][A]}{[A_{n+1}]}$$

The dissociation constant, K_d, for this reaction, defines the monomer concentrations at which the polymerization reaction will take place, because the concentration of polymers of length $n + 1$ will be essentially equal to that for polymers of length n. Thus,

$$[A_n] \sim [A_{n+1}] \qquad \text{and} \qquad K_d = \frac{[A_n][A]}{[A_{n+1}]} \sim [A]$$

In other words, the polymerization reaction will proceed until the monomer concentration is reduced to the value of K_d. If the monomer concentration is below the value of K_d, the polymerization reaction will not proceed at all; indeed, existing filaments will depolymerize until the monomer concentration reaches the value of K_d. Because of these phenomena, K_d is referred to as the *critical concentration* for the polymer. Recall that actin contains a nucleotide-binding site that can contain either ATP or ADP. The critical concentration for the actin–ATP complex is approximately 20-fold lower than that for the actin–ADP complex; actin–ATP polymerizes more readily than does actin–ADP.

Actin filaments inside cells are highly dynamic structures that are continually gaining and losing monomers. Nucleation by complexes such as Arp2/3 can initiate the polymerization of actin–ATP. In contrast, the hydrolysis of bound ATP to ADP favors actin *d*epolymerization. This reaction acts as a timer to make actin filaments kinetically unstable. Proteins that bind actin monomers or promote the severing of actin filaments also play roles. Polymerization reactions can exert force, pushing or pulling on cell membranes. *Regulated actin polymerization is central to the changes in cell shape associated with cell motility in amebae as well as in human cells such as macrophages.*

Myosin head domains bind to actin filaments

It has not been possible to determine the in vivo structure of a complex between actin and myosin at sufficiently high resolution to discern molecular details. However, treatment of actin filaments with myosin S1 fragments in the absence of ATP results in a complex referred to as *decorated actin* for which the structure has been determined by cryoelectron microscopy to a resolution of 13 Å. Although a structure at this resolution alone would not be adequate to observe molecular details, superimposition of the high-resolution structures of actin monomers and the myosin S1 fragment on the structure of decorated actin can be a source of insight into the details of its structure (Figure 35.14). The myosin head domain is in a conformation close to that observed for the nucleotide-free form. This structure also reveals the interaction surfaces between myosin and actin. The modeling suggests that the myosin head-domain conformation changes somewhat to increase its interaction with the actin filament. These conformational changes result in a slight opening of the nucleotide-binding site in myosin. This observation has implications for the mechanism by which myosin moves along actin filaments.

Motions of single motor proteins can be directly observed

Now that we understand the conformational changes behind myosin's action, we can explore how myosin "walks" along its actin track. Studies of *single myosin molecules* moving relative to actin filaments have been sources of deep insight into the mechanisms underlying muscle contraction and

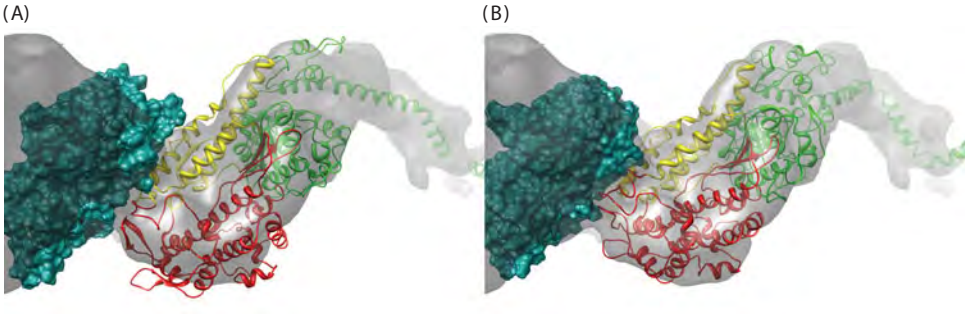


Figure 35.14 The structure of myosin bound to actin. (A) The gray surface represents the structure observed by cryoelectron microscopy, with the green space-filling model representing one actin subunit. The ribbon diagram shows the structure of the S1 fragment of myosin docked into the cryoelectron microscopic structure. *Notice* that some of the myosin structure lies outside the gray surface. (B) The structure after the structure of the myosin S1 fragment has been allowed to adjust to more closely match the structure observed by cryoelectron microscopy. *Notice* that the myosin structure now more closely matches the gray surface. [M. Lorenz and K. C. Holmes. *PNAS* 107:12529–12534, 2010. Copyright 2010 National Academy of Sciences, U.S.A.]

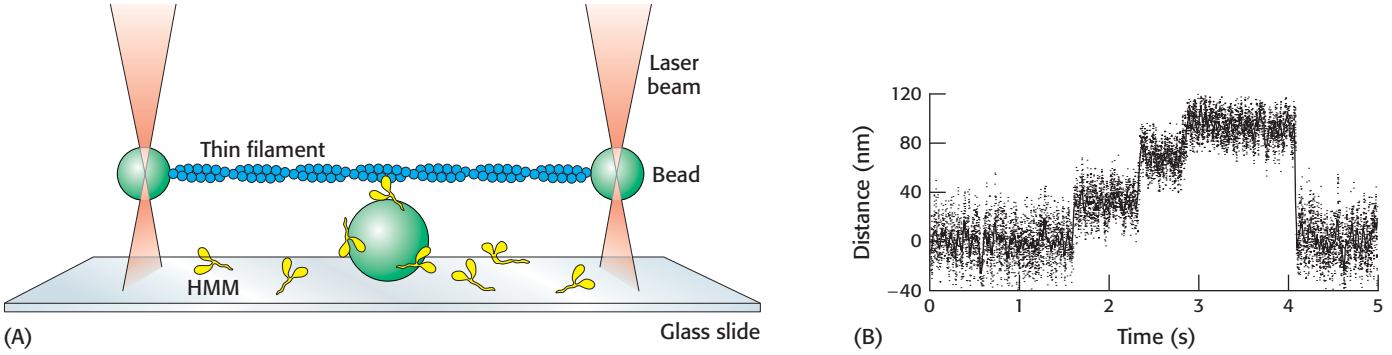


Figure 35.15 Watching a single motor protein in action. (A) An actin filament (blue) is placed above a heavy meromyosin (HMM) fragment (yellow) that projects from a bead on a glass slide. A bead attached to each end of the actin filament is held in an optical trap produced by a focused, intense infrared laser beam (orange). The position of these beads can be measured with nanometer precision. (B) Recording of the displacement of an actin filament due to a myosin derivative attached to a bead, influenced by the addition of ATP. Note the fairly uniform step sizes that are observed. [(A) After J. T. Finer, R. M. Simmons, and J. A. Spudich. *Nature* 368:113–119, 1994. (B) From R. S. Rock, M. Rief, A. D. Metra, and J. A. Spudich. *Methods* 22:378 –381, 2000.]

other complex processes. A powerful tool for these studies, called an *optical trap*, relies on highly focused laser beams (Figure 35.15). Small beads can be caught in these traps and held in place in solution.

The position of the beads can be monitored with nanometer precision. James Spudich and his coworkers designed an experimental arrangement consisting of an actin filament that had a bead attached to each end. Each bead could be caught in an optical trap (one at each end of the filament) and the actin filament could be pulled taut over a microscope slide containing other beads that had been coated with fragments of myosin such as the heavy meromyosin fragment (see Figure 35.15). On the addition of ATP, transient displacements of the actin filament were observed along its long axis. The size of the displacement steps was fairly uniform with an average size of 11 nm (110 Å).

The results of these studies, performed in the presence of varying concentrations of ATP, are interpreted as showing that individual myosin heads bind the actin filament and undergo a conformational change (the *power stroke*) that pulls the actin filament, leading to the displacement of the beads. After a period of time, the myosin head releases the actin, which then snaps back into place.

Phosphate release triggers the myosin power stroke

How does ATP hydrolysis drive the power stroke? A key observation is that the addition of ATP to a complex of myosin and actin results in the dissociation of the complex. Thus, ATP binding and hydrolysis cannot be directly responsible for the power stroke. We can combine this fact with the structural observations described earlier to construct a mechanism for the motion of myosin along actin (Figure 35.16). Let us begin with nucleotide-free myosin bound to actin. The binding of ATP to actin results in the dissociation of myosin from actin. With ATP bound and free of actin, the myosin domain can undergo the conformational change associated with the formation of the transition state for ATP hydrolysis. This conformational change results in the reorientation of the lever arm. In this form, the myosin head can dock onto the actin filament; phosphate is released with concomitant motion of the lever arm. *This conformational change represents the power stroke and moves the body of the myosin molecule relative to the actin filament by approximately 110 Å*. The release of ADP completes the cycle.

Muscle is a complex of myosin and actin

The mechanism of moving a single myosin molecule relative to an actin filament explains how muscles contract. Vertebrate muscle that is under voluntary control, such as the biceps and triceps in your upper arm, has a banded (striated) appearance when examined under a light microscope. It consists of multinucleated cells that are bounded by an electrically excitable

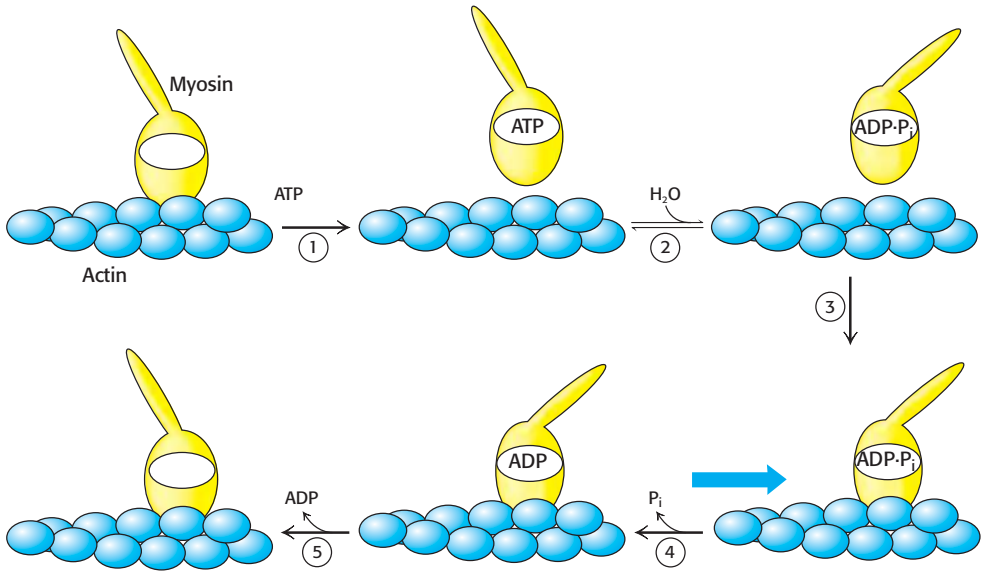
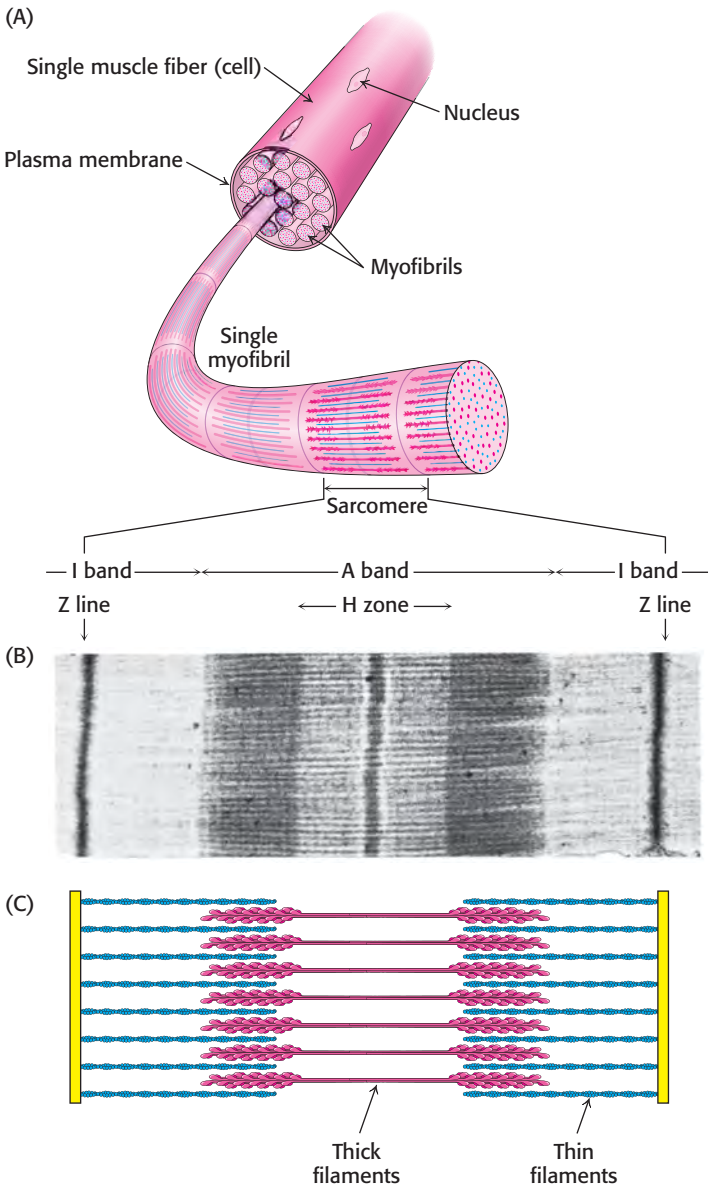


Figure 35.16 Myosin motion along actin. A myosin head (yellow) in the apo form is bound to an actin filament (blue). The binding of ATP (1) results in the release of myosin from actin. The reversible hydrolysis of ATP bound to myosin (2) can result in the reorientation of the lever arm. With ATP hydrolyzed but still bound to actin, myosin can bind actin (3). The release of P_i (4) results in the reorientation of the lever arm and the concomitant motion of actin relative to myosin. The release of ADP (5) completes the cycle.



plasma membrane. A muscle cell contains many parallel *myofibrils*, each about 1 m in diameter. The functional unit, called a *sarcomere*, typically repeats every 2.3 μm (23,000 Å) along the fibril axis in relaxed muscle (Figure 35.17). A dark *A band* and a light *I band* alternate regularly. The central region of the A band, termed the *H zone*, is less dense that the rest of the band. The I band is bisected by a very dense, narrow *Z line*.

The underlying molecular plan of a sarcomere is revealed by cross sections of a myofibril. These cross sections show the presence of two kinds of interacting protein filaments. The *thick filaments* have diameters of about 15 nm (150 Å) and consist primarily of myosin. The *thin filaments* have diameters of approximately 8 nm (80 Å) and consist of actin as well as *tropomyosin* and the *troponin complex*. Muscle contraction is achieved through the sliding of the thin filaments along the length of the thick filaments, driven by the hydrolysis of ATP (Figure 35.18).

To form the thick filaments, myosin molecules self-assemble into thick bipolar structures, with the myosin heads protruding at both ends of a bare region in the center (Figure 35.19A). Approximately 500 head domains line the surface of each thick filament. Each head-rich region associates with two actin filaments, one on each side of the myosin molecules (Figure 35.19B). The interaction of individual myosin heads with actin units creates the sliding force that gives rise to muscle contraction.

Figure 35.17 Sarcomere. (A) Structure of muscle cell and myofibril containing sarcomeres. (B) Electron micrograph of a longitudinal section of a skeletal-muscle myofibril showing a single sarcomere. (C) Schematic representation of the sarcomere corresponding to the regions in the micrograph. [(B) Courtesy of Dr. Hugh Huxley.]

Tropomyosin and the troponin complex regulate this sliding in response to nerve impulses. Under resting conditions, tropomyosin blocks the intimate interaction between myosin and actin. A nerve impulse leads to an increase in calcium ion concentration within the muscle cell. A component of the troponin complex senses the increase in Ca^{2+} and, in response, relieves the inhibition of myosin–actin interactions by tropomyosin.

How does the myosin reaction cycle apply to muscle contraction? Recall that hundreds of head domains project from the ends of each thick filament. The head domains are paired in myosin dimers, but the two heads within each dimer act independently. Actin filaments associate with each head-rich region, with the barbed ends of actin toward the Z line. In the presence of normal levels of ATP, most of the myosin heads are detached from actin. Each head can independently hydrolyze ATP, bind to actin, release P_i, and undergo its power stroke. Because few other heads are attached, the actin filament is relatively free to slide. Each head cycles approximately five times per second with a movement of 110 Å per cycle. However, when hundreds of heads are interacting with the same actin filament, the overall rate of movement of myosin relative to the actin filament may reach 80,000 Å per second, allowing a sacromere to contract from its fully relaxed to its fully contracted form rapidly. Having many myosin heads briefly and independently attaching and moving an actin filament allows for much greater speed than could be achieved by a single motor protein.

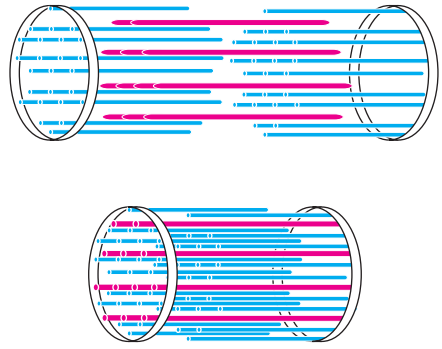


Figure 35.18 Sliding-filament model. Muscle contraction depends on the motion of thin filaments (blue) relative to thick filaments (red). [After H. E. Huxley. The mechanism of muscular contraction. Copyright © 1965 by Scientific American, Inc. All rights reserved.]

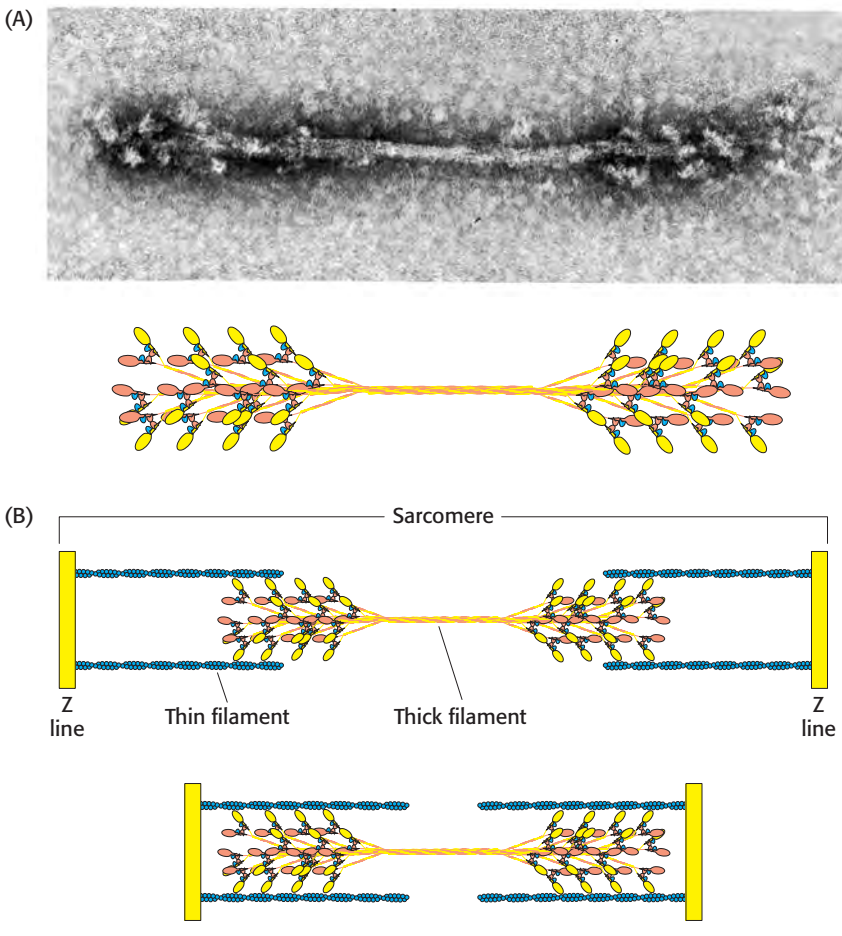


Figure 35.19 Thick filament. (A) An electron micrograph of a reconstituted thick filament reveals the presence of myosin head domains at each end and a relatively narrow central region. A schematic view below shows how myosin molecules come together to form the thick filament. (B) A diagram showing the interaction of thick and thin filaments in skeletal-muscle contraction. [(A, top) Courtesy of Dr. Hugh Huxley.]

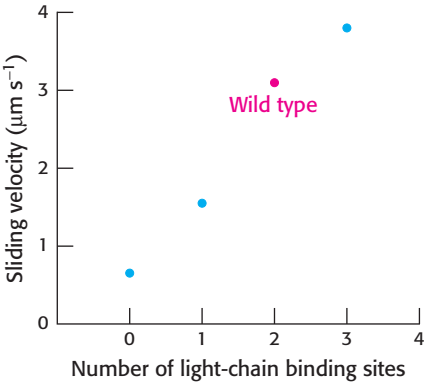


Figure 35.20 Myosin lever-arm length. Examination of the rates of actin movement supported by a set of myosin mutants with different numbers of light-chain binding sites revealed a linear relation; the greater the number of light-chain binding sites (and, hence, the longer the lever arm), the faster the sliding velocity. [After T. Q. P. Uyeda, P. D. Abramson, and J. A. Spudich. *Proc. Natl. Acad. Sci. U. S. A.* 93:4459 –4464, 1996.]

The length of the lever arm determines motor velocity

A key feature of myosin motors is the role of the lever arm as an amplifier. The lever arm amplifies small structural changes at the nucleotide-binding site to achieve the 110-Å movement along the actin filament that takes place in each ATP hydrolysis cycle. A strong prediction of the mechanism proposed for the movement of myosin along actin is that the length traveled per cycle should depend on the length of this lever arm. Thus, the length of the lever arm should influence the overall rate at which actin moves relative to a collection of myosin heads.

This prediction was tested with the use of mutated forms of myosin with lever arms of different lengths. The lever arm in muscle myosin includes binding sites for two light chains (Section 35.1). Thus investigators shortened the lever arm by deleting the sequences that correspond to one or both of these binding sites. They then examined the rates at which actin filaments were transported along collections of these mutated myosins (Figure 35.20). As predicted, the rate decreased as the lever arm was shortened. A mutated form of myosin with an unusually long lever arm was generated by inserting 23 amino acids corresponding to the binding site for an additional regulatory light chain. Remarkably, this form was found to support actin movement that was *faster than that of the wild-type protein*. These results strongly support the proposed role of the lever arm in contributing to myosin motor activity.

35.3 Kinesin and Dynein Move Along Microtubules

In addition to actin, the cytoskeleton includes other components, notably intermediate filaments and microtubules. Microtubules serve as tracks for two classes of motor proteins—namely, kinesins and dyneins. Kinesins moving along microtubules usually carry cargo such as organelles and vesicles from the center of a cell to its periphery. Dyneins are important in sliding microtubules relative to one other during the beating of cilia and flagella on the surfaces of some eukaryotic cells. Additionally, dynein carries cargo from the cell periphery to the cell center.

Some members of the kinesin family are crucial to the transport of organelles and other cargo to nerve endings at the peripheries of neurons. It is not surprising, then, that mutations in these kinesins can lead to nervous system disorders. For example, mutations in a kinesin called KIF1β can lead to the most common peripheral neuropathy (weakness and pain in the hands and feet), Charcot-Marie-Tooth disease, which affects 1 in 2500 people. A glutamine-to-leucine mutation in the P-loop of the motor domain of this kinesin has been found in some affected persons. Knockout mice with a disruption of the orthologous gene have been generated. Mice heterozygous for the disruption show symptoms similar to those observed in human beings; homozygotes die shortly after birth. Mutations in other kinesin genes have been linked to human spastic paraplegia. In these disorders, defects in kinesin-linked transport may impair nerve function directly, and the decrease in the activity of specific neurons may lead to other degenerative processes.

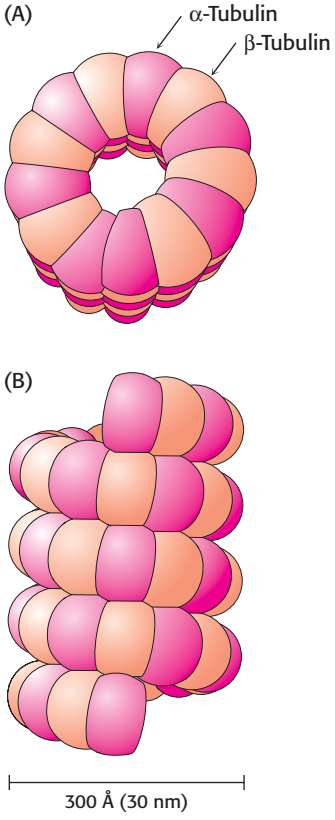


Figure 35.21 Microtubule structure. Schematic views of the helical structure of a microtubule. α-Tubulin is shown in dark red and β-tubulin in light red. (A) Top view. (B) Side view.

Microtubules are hollow cylindrical polymers

Microtubules are a major component of the cytoskeleton. They are built from two kinds of homologous 50-kd subunits, α- and β-tubulin, which assemble in a helical array of alternating tubulin types to form the wall of a hollow cylinder (Figure 35.21). Alternatively, a microtubule can be regarded as 13 protofilaments that run parallel to its long axis. The outer

diameter of a microtubule is 30 nm, much larger than that of actin (8 nm). Like actin, microtubules are polar structures. The minus end of a microtubule is anchored near the center of a cell, whereas the plus end extends toward the cell surface.

Microtubules are also key components of cilia and flagella present on some eukaryotic cells. For example, sperm propel themselves through the motion of flagella containing microtubules. The microtubules present in these structures adopt a common architecture (Figure 35.22). A bundle of microtubules called an *axoneme* is surrounded by a membrane contiguous with the plasma membrane. The axoneme is composed of a peripheral group of nine microtubule pairs surrounding two singlet microtubules. This recurring motif is often called a *9 + 2 array.* Dynein drives the motion of one member of each outer pair relative to the other, causing the overall structure to bend.

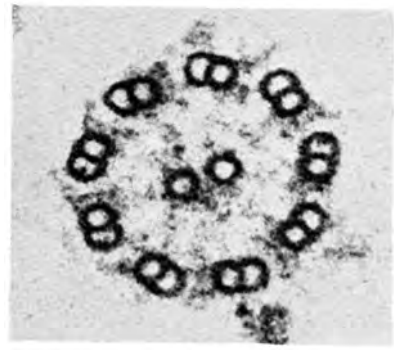


Figure 35.22 Microtubule arrangement. Electron micrograph of a cross section of a flagellar axoneme shows nine microtubule doublets surrounding two singlets. [Courtesy of Dr. Joel Rosenbaum.]

Microtubules are important in determining the shapes of cells and in separating daughter chromosomes in mitosis. They are highly dynamic structures that grow through the addition of α- and β-tubulin to the ends of existing structures. Like actin, *tubulins* bind and hydrolyze nucleoside triphosphates, although, for tubulin, the nucleotide is GTP rather than ATP. The critical concentration for the polymerization of the GTP forms of tubulin is lower than that for the GDP forms. Thus, a newly formed microtubule consists primarily of GTP-tubulins. Through time, the GTP is hydrolyzed to GDP. The GDP-tubulin subunits in the interior length of a microtubule remain stably polymerized, whereas GDP subunits exposed at an end have a strong tendency to dissociate. Marc Kirschner and Tim Mitchison found that some microtubules in a population lengthen while others simultaneously shorten. This property, called *dynamic instability,* arises from random fluctuations in the number of GTP- or GDP-tubulin subunits at the plus end of the polymer. The dynamic character of microtubules is crucial for processes such as mitosis, which require the assembly and disassembly of elaborate microtubule-based structures such as spindle fibers.

The structure of tubulin was determined at high resolution by electron crystallographic methods (Figure 35.23). As expected from their 40% sequence identity, α- and β-tubulin have very similar three-dimensional structures. Further analysis revealed that the tubulins are members of the P-loop NTPase family and contain a nucleotide-binding site adjacent to the P-loop. Tubulins are present only in eukaryotes, although a prokaryotic homolog has been found. Sequence analysis identified a prokaryotic protein called FtsZ (for *f*ilamentous temperature-sensitive mutant *Z*) that is quite similar to the tubulins. The homology was confirmed when the structure was determined by x-ray crystallography. Interestingly, this protein participates in bacterial cell division, forming ring-shaped structures at the constriction that arises when a cell divides. These observations suggest that tubulins may have evolved from an ancient cell-division protein.

The continual lengthening and shortening of microtubules is essential to their role in cell division. *Taxol,* a compound isolated from the bark of the Pacific yew tree, was discovered through its ability to interfere with cell proliferation.

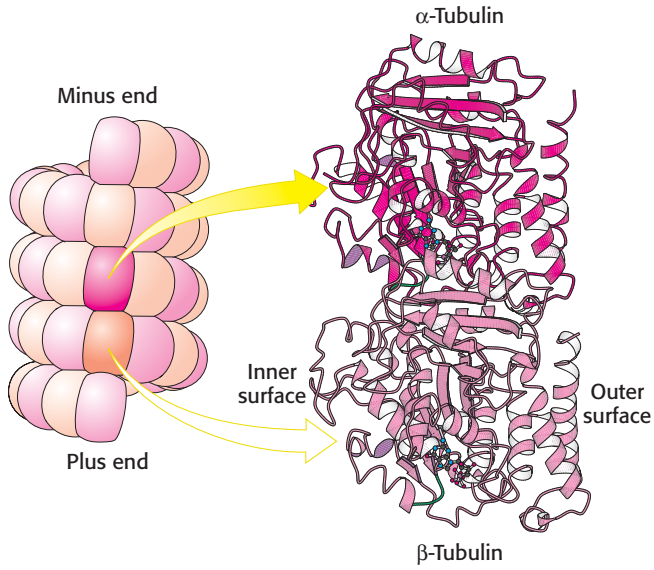
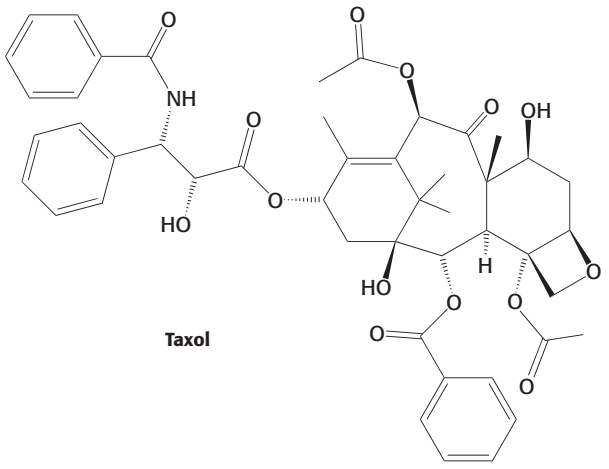


Figure 35.23 Tubulin. Microtubules can be viewed as an assembly of α-tubulin– β-tubulin dimers. The structures of α-tubulin and β-tubulin are quite similar. *Notice* that each includes a P-loop NTPase domain (shown in purple) and a bound guanine nucleotide. [Drawn from 1JFF.pdb.]



Taxol binds to microtubules and stabilizes the polymerized form. Taxol and its derivatives have been developed as anticancer agents because they preferentially affect rapidly dividing cells, such as those in tumors.

Kinesin motion is highly processive

Kinesins are motor proteins that move along microtubules. We have seen that myosin moves along actin filaments by a process in which actin is released in each cycle; a myosin head group acting independently dissociates from actin after every power stroke. In contrast, when a kinesin molecule moves along a microtubule, the two head groups of the kinesin molecule operate in tandem: one binds, and then the next one does. A kinesin molecule may take many steps before both head groups are dissociated at the same time. In other words, the motion of kinesin is highly processive. Single-molecule measurements allow processive motion to be observed (Figure 35.24). A single kinesin molecule will typically take 100 or more steps toward the plus end of a microtubule in a period of seconds before the molecule becomes detached from the microtubule. These measurements also revealed that the average step size is approximately 80 Å, a value that corresponds to the distance between consecutive α- or β-tubulin subunits along each protofilament.

An additional fact is crucial to the development of a mechanism for kinesin motion—namely, that the addition of ATP strongly *increases* the affinity of kinesin for microtubules. This behavior stands in contrast with the behavior of myosin; ATP binding to myosin promotes its *dissociation* from actin. Do these differences imply that kinesin and myosin operate by completely different mechanisms? Indeed not. Kinesin-generated movement

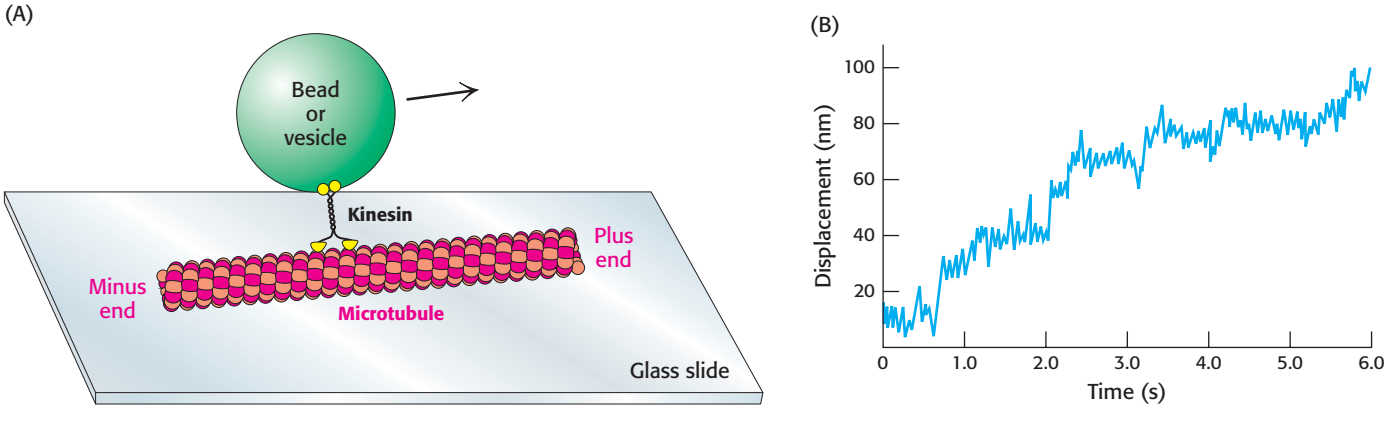


Figure 35.24 Monitoring movements mediated by kinesin. (A) The movement of beads or vesicles, carried by individual kinesin dimers along a microtubule, can be directly observed. (B) A trace shows the displacement of a bead carried by a kinesin molecule. Multiple steps are taken in the 6-s interval. The average step size is about 8 nm (80 Å). [(B) After K. Svoboda et al. *Nature* 365:721–727, 1993.]

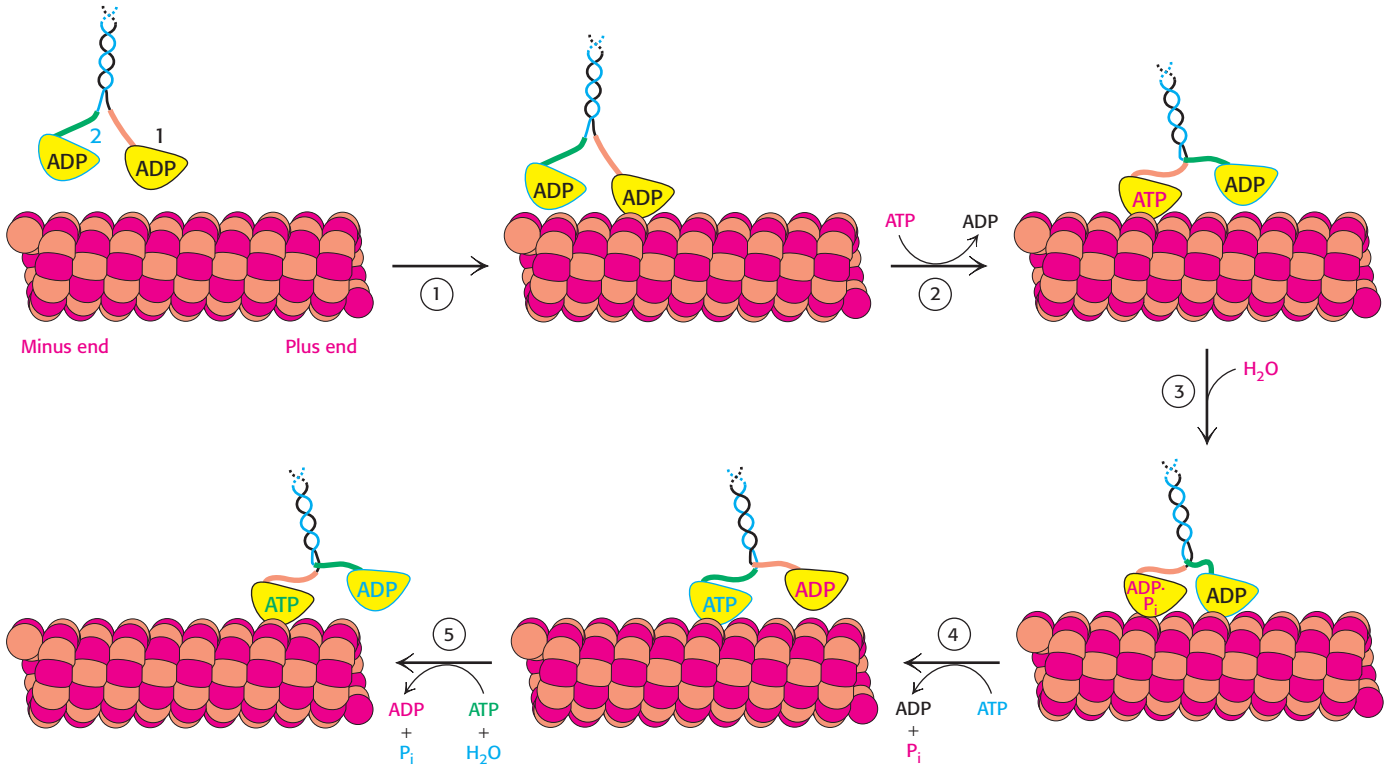


Figure 35.25 Kinesin moving along a microtubule. (1) One head of a two-headed kinesin molecule, initially with both heads in the ADP form, binds to a microtubule. (2) The release of ADP and the binding of ATP results in a conformational change that locks the head to the microtubule and pulls the neck linker (orange) to the head domain, throwing the second domain toward the plus end of the microtubule. (3) ATP undergoes hydrolysis while the second head interacts with the microtubule. (4) The exchange of ATP for ADP in the second head pulls the first head off the microtubule, releasing P_i and moving the first domain along the microtubule. (5) The cycle repeats, moving the kinesin dimer farther down the microtubule.

appears to proceed by a mechanism that is quite similar to that used by myosin (Figure 35.25). Let us begin with a two-headed kinesin molecule in its ADP form, dissociated from a microtubule. Recall that the neck linker binds the head domain when ATP is bound and is released when ADP is bound. The initial interaction of one of the head domains with a tubulin dimer on a microtubule stimulates the release of ADP from this head domain and the subsequent binding of ATP. The binding of ATP triggers a conformational change in the head domain that leads to two important events. First, the affinity of the head domain for the microtubule increases, essentially locking this head domain in place. Second, the neck linker binds to the head domain. This change, transmitted through the coiled-coil domain that connects the two kinesin monomers, repositions the other head domain. In its new position, the second head domain is close to a second tubulin dimer, 80 Å along the microtubule in the direction of the plus end. Meanwhile, the intrinsic ATPase activity of the first head domain hydrolyzes the ATP to ADP and P_i. When the second head domain binds to the microtubule, the first head releases ADP and binds ATP. Again, ATP binding favors a conformational change that pulls the first domain forward. This process can continue for many cycles until, by chance, both head domains are in the ADP form simultaneously and kinesin dissociates from the microtubule. Because of the relative rates of the component reactions, a simultaneous dissociation takes place approximately every 100 cycles. Thus, kinesin "walks" hand over hand along the microtubule for approximately 100 steps before both "hands" simultaneously let go. Although recent studies have led to a structural model for the dynein power stroke, the mechanism by which it works is unknown and an interesting research frontier.

Kinesin hydrolyzes ATP at a rate of approximately 80 molecules per second. Thus, given the step size of 80 Å per molecule of ATP, kinesin moves along a microtubule at a speed of 6400 Å per second. This rate is considerably lower than the maximum rate for myosin, which moves

relative to actin at 80,000 Å per second. Recall, however, that myosin movement depends on the independent action of hundreds of different head domains working along the same actin filament, whereas the movement of kinesin is driven by the processive action of kinesin head groups working in pairs. Muscle myosin evolved to maximize the speed of the motion, whereas kinesin functions to achieve steady, but slower, transport in one direction along a filament.

35.4 A Rotary Motor Drives Bacterial Motion

In 1 s, a motile bacterium can move approximately 25 μm, or about 10 body lengths. A human being sprinting at a proportional rate would complete the 100-meter dash in slightly more than 5 s. The motors that power this impressive motion are strikingly different from the eukaryotic motors that we have seen so far. In the bacterial motor, an element spins around a central axis rather than moving along a polymeric track. The direction of rotation can change rapidly, a feature that is central to chemotaxis, the process by which bacteria swim preferentially toward an increasing concentration of certain useful compounds and away from potentially harmful ones. One type of flagellar motor, powered by a Na^+ gradient, turns at a rate of 200,000 revolutions per minute.

Bacteria swim by rotating their flagella

Bacteria such as *Escherichia coli* and *Salmonella typhimurium* swim by rotating flagella that lie on their surfaces (Figure 35.26). When the flagella rotate in a counterclockwise direction (viewed from outside a bacterium), the separate flagella form a bundle that very efficiently propels the bacterium through solution.

Bacterial flagella are polymers approximately 15 nm in diameter and as much as 15 μm in length, composed of 53-kd subunits of a protein called *flagellin* (Figure 35.27). These subunits associate into a helical structure that has 5.5 subunits per turn, giving the appearance of 11 protofilaments. Each flagellum has a hollow core. Remarkably, flagella form not by growing at the base adjacent to the cell body but, instead, by the addition of new subunits that pass through the hollow core and add to the free end. Each flagellum is intrinsically twisted in a left-handed sense. At its base, each flagellum has a rotory motor.

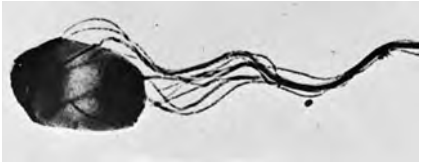


Figure 35.26 Bacterial flagella. Electron micrograph of *Salmonella typhimurium* shows flagella in a bundle. [Courtesy of Dr. Daniel Koshland, Jr.]

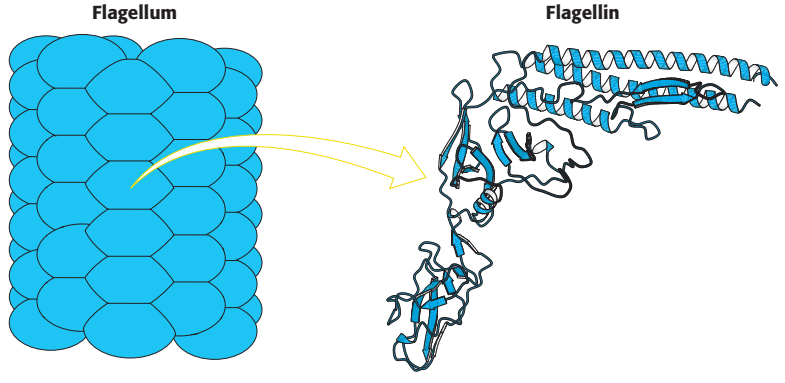


Figure 35.27 Structure of flagellin. A bacterial flagellum is a helical polymer of the protein flagellin. *Notice* that each subunit corresponds to a bent structure with a relatively flat surface facing the hollow core of the flagellum. [Drawn from 1IO1.pdb.]

Proton flow drives bacterial flagellar rotation

Early experiments by Julius Adler demonstrated that ATP is *not* required for flagellar motion. What powers these rotary motors? The necessary free

energy is derived from the proton gradient that exists across the plasma membrane. The flagellar motor is quite complex, containing as many as 40 distinct proteins (Figure 35.28). Five components particularly crucial to motor function have been identified through genetic studies. MotA is a membrane protein that appears to have four transmembrane helices as well as a cytoplasmic domain. MotB is another membrane protein with a single transmembrane helix and a large periplasmic domain. Approximately 11 *MotA–MotB pairs* form a ring around the base of the flagellum. The proteins *FliG*, *FliM*, and *FliN* are part of a disc-like structure called the MS (membrane and supramembrane) ring, with approximately 30 FliG subunits coming together to form the ring. The three-dimensional structure of the carboxyl-terminal half of FliG reveals a wedge-shaped domain with a set of charged amino acids, conserved among many species, lying along the thick edge of the wedge (Figure 35.29).

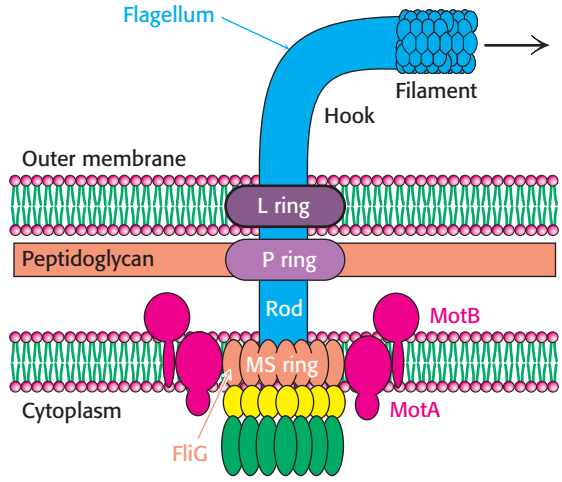


Figure 35.28 Flagellar motor. A schematic view of the flagellar motor, a complex structure containing as many as 40 distinct types of protein. The approximate positions of the proteins MotA and MotB (red), FliG (orange), FliN (yellow), and FliM (green) are shown.

The MotA–MotB pair and FliG combine to create a proton channel that drives the rotation of the flagellum. How can proton flow across a membrane drive mechanical rotation? We have seen such a process earlier in regard to ATP synthase (Section 18.4). Recall that the key to driving the rotation of the γ subunit of ATP synthase is the **a** subunit of the F_0 fragment. This subunit appears to have two half-channels; protons can move across the membrane only by moving into the half-channel from the side of the membrane with the higher local proton concentration, binding to a disc-like structure formed by the **c** subunits, riding on this structure as it rotates to the opening of the other half-channel, and exiting to the side with the lower local proton concentration. Could a similar mechanism apply to flagellar rotation? Indeed, such a mechanism was first proposed by Howard Berg to explain flagellar rotation before the rotary mechanism of ATP synthase was elucidated. Each MotA–MotB pair is conjectured to form a structure that has two half-channels; FliG serves as the rotating proton carrier, perhaps with the participation of some of the charged residues identified in crystallographic studies (Figure 35.30). In this scenario, a proton from the periplasmic space passes into the outer half-channel and is transferred to an FliG subunit. The MS ring rotates, rotating the flagellum with it and allowing the proton to pass into the inner half-channel and into the cell. Ongoing structural and mutagenesis studies are testing and refining this hypothesis.

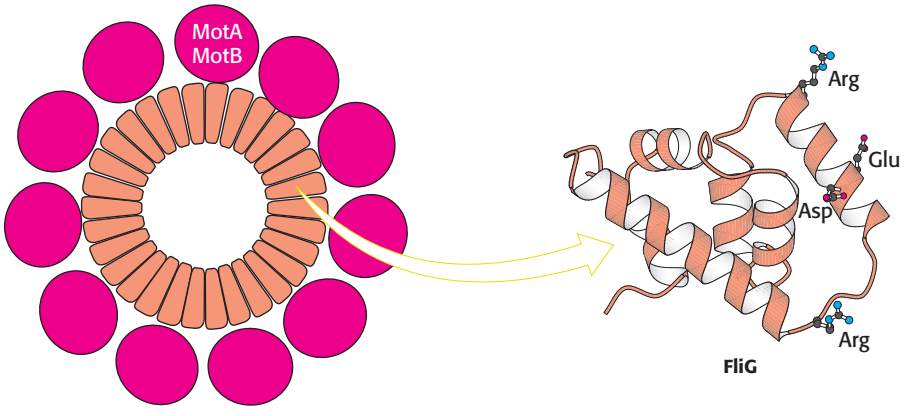


Figure 35.29 Flagellar motor components. Approximately 30 subunits of FliG assemble to form part of the MS ring. The ring is surrounded by approximately 11 structures consisting of MotA and MotB. *Notice* that the carboxyl-terminal domain of FliG includes a ridge lined with charged residues that may participate in proton transport. [Drawn from 1QC7.pdb.]

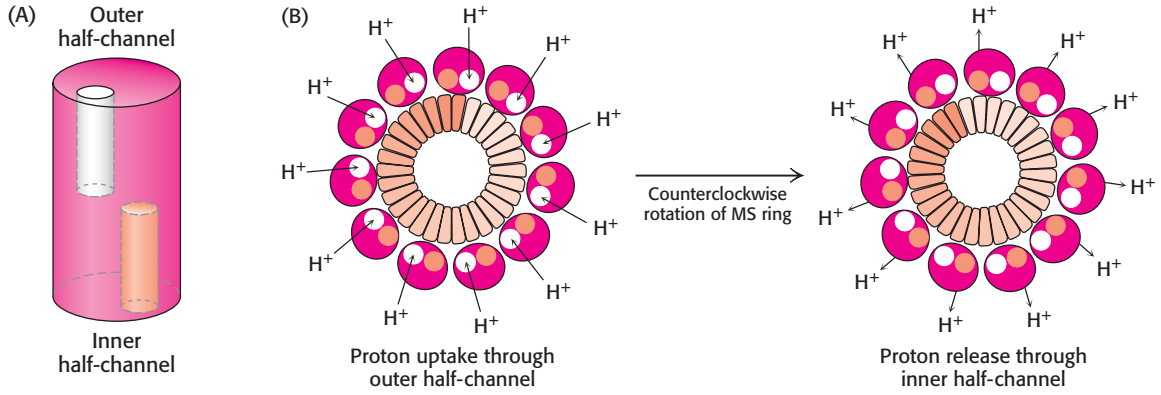


Figure 35.30 Proton-transport-coupled rotation of the flagellum. (A) MotA–MotB may form a structure having two half-channels. (B) One model for the mechanism of coupling rotation to a proton gradient requires protons to be taken up into the outer half-channel and transferred to the MS ring. The MS ring rotates in a counterclockwise direction, and the protons are released into the inner half-channel. The flagellum is linked to the MS ring and so the flagellum rotates as well.

Bacterial chemotaxis depends on reversal of the direction of flagellar rotation

Many species of bacteria respond to changes in their environments by adjusting their swimming behavior. Examination of the paths taken is highly revealing (Figure 35.31). The bacteria swim in one direction for some length of time (typically about a second), tumble briefly, and then set off in a new direction. The tumbling is caused by a brief reversal in the direction of the flagellar motor. When the flagella rotate counterclockwise, the helical filaments form a coherent bundle favored by the intrinsic shape of each filament, and the bacterium swims smoothly. When the rotation reverses, the bundle flies apart because the screw sense of the helical flagella does not match the direction of rotation. Each flagellum then pulls in a different direction and the cell tumbles.

In the presence of a gradient of certain substances such as glucose, bacteria swim preferentially toward the direction of the higher concentration of the substance. Such compounds are referred to as *chemoattractants*. Bacteria also swim preferentially away from potentially harmful compounds such as phenol, a *chemorepellant*. The process of moving in specific directions in response to environmental cues is called *chemotaxis*. In the presence of a gradient of a chemoattractant, bacteria swim for longer periods of time without tumbling when moving toward higher concentrations of the chemoattractant. In contrast, they tumble more frequently when moving toward lower concentrations of the chemoattractant. This behavior is reversed for chemorepellants. The result of these actions is a *biased random walk* that facilitates net motion toward conditions more favorable to the bacterium.

Chemotaxis depends on a signaling pathway that terminates at the flagellar motor. The signaling pathway begins with the binding of molecules to receptors in the plasma membrane (Figure 35.32). In their *unoccupied* forms, these receptors initiate a pathway leading eventually to the phosphorylation of a specific aspartate residue on a soluble protein called *CheY*. In its phosphorylated form, CheY binds to the base on the flagellar motor. When bound to phosphorylated CheY, the flagellar motor rotates in a clockwise rather than a counterclockwise direction, causing tumbling.

The binding of a chemoattractant to a surface receptor blocks the signaling pathway leading to CheY phosphorylation. Phosphorylated CheY spontaneously hydrolyzes and releases its phosphoryl group in a process accelerated by another protein, CheZ. The concentration of phosphorylated CheY drops, and the flagella are less likely to rotate in a clockwise direction. Under these conditions, bacteria swim smoothly without tumbling. Thus, the reversible rotary flagellar motor and a phosphorylation-based signaling pathway work together to generate an effective means for responding to environmental conditions.

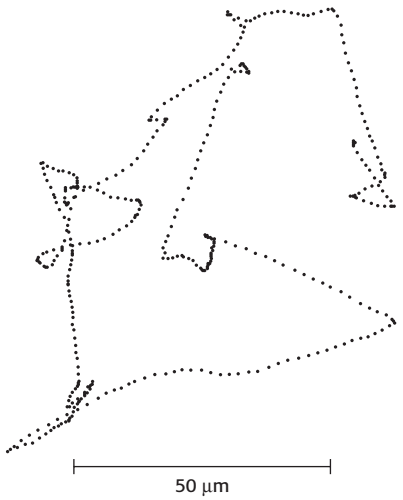


Figure 35.31 Charting a course. This projection of the track of an *E. coli* bacterium was obtained with a microscope that automatically follows bacterial motion in three dimensions. The points show the locations of the bacterium at 80-ms intervals. [After H. C. Berg. *Nature* 254:389–392, 1975.]

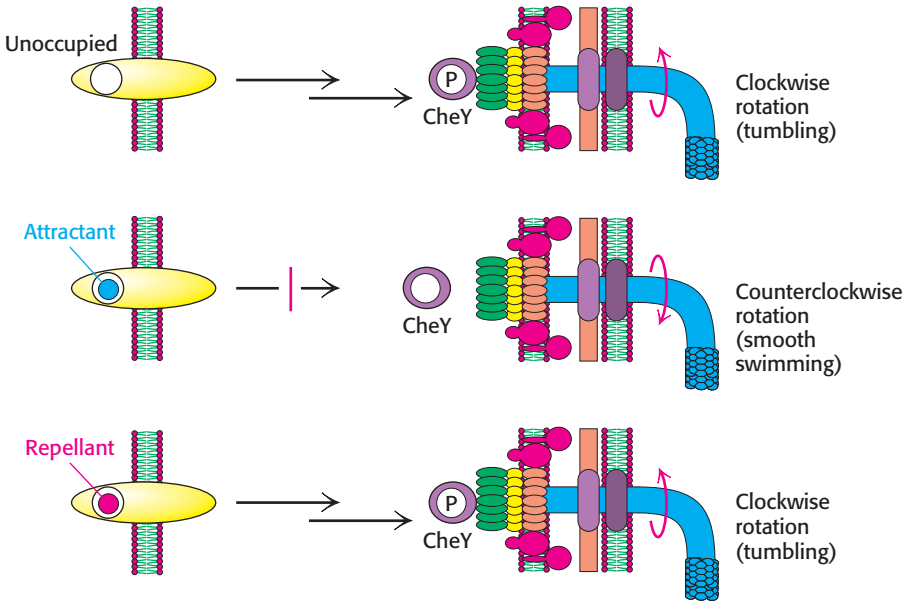


Figure 35.32 Chemotaxis signaling pathway. Receptors in the plasma membrane initiate a signaling pathway leading to the phosphorylation of the CheY protein. Phosphorylated CheY binds to the flagellar motor and favors clockwise rotation. When an attractant binds to the receptor, this pathway is blocked, and counterclockwise flagellar rotation and, hence, smooth swimming result. When a repellant binds, the pathway is stimulated, leading to an increased concentration of phosphorylated CheY and, hence, more-frequent clockwise rotation and tumbling.

Bacteria sense spatial gradients of chemoattractants by measurements separated in time. A bacterium sets off in a random direction and, if the concentration of the chemoattractant has increased after the bacterium has been swimming for a period of time, the likelihood of tumbling decreases and the bacterium continues in roughly the same direction. If the concentration has decreased, the tumbling frequency increases and the bacterium tests other random directions. The success of this mechanism once again reveals the power of evolutionary problem solving: many possible solutions are tried at random, and those that are beneficial are selected and exploited.

Summary

35.1 Most Molecular-Motor Proteins Are Members of the P-Loop NTPase Superfamily

Eukaryotic cells contain three families of molecular-motor proteins: myosins, kinesins, and dyneins. These proteins move along tracks defined by the actin and microtubule cytoskeletons of eukaryotic cells, contributing to cell and organismal movement and to the intracellular transport of proteins, vesicles, and organelles. Despite considerable differences in size and a lack of similarity detectable at the level of amino acid sequence, these proteins are homologous, containing core structures of the P-loop NTPase family. The ability of these core structures to change conformations in response to nucleoside triphosphate binding and hydrolysis is key to molecular-motor function. Motor proteins consist of motor domains attached to extended structures that serve to amplify the conformational changes in the core domains and to link the core domains to one another or to other structures.

35.2 Myosins Move Along Actin Filaments

The motile structure of muscle consists of a complex of myosin and actin, along with accessory proteins. Actin, a highly abundant 42-kd protein, polymerizes to form long filaments. Each actin monomer can bind either ATP or ADP. A myosin motor domain moves along actin filaments in a cyclic manner, beginning with myosin free of bound

nucleotides bound to actin: (1) ATP binds to myosin and the myosin is released from actin; (2) a reversible conformational change associated with the hydrolysis of ATP while still bound to myosin leads to a large motion of a lever arm that extends from the motor domain; (3) myosin with bound ADP and P_i binds to actin; (4) P_i is released from myosin, resulting in resetting the position of the lever arm and moving actin relative to myosin; and (5) the release of ADP returns the motor domain to its initial state. The length of the lever arm determines the size of the step taken along actin in each cycle. The ability to monitor single molecular-motor proteins has provided key tests for hypotheses concerning motor function. Muscle contraction entails the rapid sliding of thin filaments, composed of actin, relative to thick filaments, composed of myosin. The thick filaments consist of multiple myosin molecules bound together. Each myosin molecule has two heads that can bind to actin and move it relative to myosin, driven by the hydrolysis of ATP by myosin. Muscle contraction is regulated by tropomyosin and the troponin complex. These proteins prevent actin and myosin from interacting until an increase in calcium concentration associated with a nerve impulse results in calcium-induced changes in troponin and tropomyosin.

35.3 Kinesin and Dynein Move Along Microtubules

Kinesin and dynein move along microtubules rather than actin. Microtubules are polymeric structures composed of α- and β-tubulin, two very similar guanine-nucleotide-binding proteins. Each microtubule comprises 13 protofilaments with alternating α- and β-tubulin subunits. Kinesins move along microtubules by a mechanism quite similar to that used by myosin to move along actin but with several important differences. First, ATP binding to kinesin favors motor-domain binding rather than dissociation. Second, the power stroke is triggered by the binding of ATP rather than the release of P_i. Finally, kinesin motion is processive. The two heads of a kinesin dimer work together, taking turns binding and releasing the microtubule, and many steps are taken along a microtubule before both heads dissociate. Most kinesins move toward the plus end of microtubules.

35.4 A Rotary Motor Drives Bacterial Motion

Many motile bacteria use rotating flagella to propel themselves. When rotating counterclockwise, multiple flagella on the surface of a bacterium come together to form a bundle that effectively propels the bacterium through solution. A proton gradient across the plasma membrane, rather than ATP hydrolysis, powers the flagellar motor. The mechanism for coupling transmembrane proton transport to macromolecular rotation appears to be similar to that used by ATP synthase. When rotating clockwise, the flagella fly apart and the bacterium tumbles. Bacteria swim preferentially toward chemoattractants in a process called chemotaxis. When bacteria are swimming in the direction of an increasing concentration of a chemoattractant, clockwise flagellar motion predominates and tumbling is suppressed, leading to a biased random walk in the direction of increasing chemoattractant concentration.

Key Terms

cytoskeleton (p. 1007)
myosin (p. 1008)
kinesin (p. 1008)
dynein (p. 1008)
S1 fragment (p. 1009)
conventional kinesin (p. 1010)

lever arm (p. 1011)
relay helix (p. 1012)
neck linker (p. 1012)
actin (p. 1012)
G-actin (p. 1013)
F-actin (p. 1013)
critical concentration (p. 1014)
optical trap (p. 1015)
power stroke (p. 1015)
myofibril (p. 1016)
sarcomere (p. 1016)
tropomyosin (p. 1016)
troponin complex (p. 1016)
microtubule (p. 1018)
tubulin (p. 1019)
dynamic instability (p. 1019)
flagellin (p. 1022)
MotA–MotB pair (p. 1023)
FliG (p. 1023)
chemoattractant (p. 1024)
chemorepellant (p. 1024)
chemotaxis (p. 1024)
CheY (p. 1024)

Problems

1. *Diverse motors.* Skeletal muscle, eukaryotic cilia, and bacterial flagella use different strategies for the conversion of free energy into coherent motion. Compare and contrast these motility systems with respect to (a) the free-energy source and (b) the number of essential components and their identity.

2. *You call that slow?* At maximum speed, a kinesin molecule moves at a rate of 6400 Å per second. Given the dimensions of the motor region of a kinesin dimer of approximately 80 Å, calculate its speed in "body lengths" per second. To what speed does this body-length speed correspond for an automobile 10 feet long?

3. *Heavy lifting.* A single myosin motor domain can generate a force of approximately 4 piconewtons (4 pN). How many times its "body weight" can a myosin motor domain lift? Note that 1 newton = 0.22 pounds (100 g). Assume a molecular mass of 100 kd for the motor domain.

4. *Compare and contrast.* Describe two similarities and two differences between actin filaments and microtubules.

5. *Lighten up.* What is the primary role of the light chains in myosin? In kinesin?

6. *Rigor mortis.* Why does the body stiffen after death?

7. *Now you see it, now you don't.* Under certain stable concentration conditions, actin monomers in their ATP form will polymerize to form filaments that disperse again into free actin monomers over time. Explain.

8. *Helicases as motors.* Helicases can use single-stranded DNA as tracks. Consider a helicase that moves one base in the $3' \longrightarrow 5'$ direction in each cycle. Assuming that the helicase can hydrolyze ATP at a rate of 50 molecules per second in the presence of a single-stranded DNA template, calculate the velocity of the helicase in micrometers per second. How does this velocity compare with that of kinesin?

9. *New moves.* When bacteria such as *E. coli* are starved to a sufficient extent, they become nonmotile. However, when such bacteria are placed in an acidic solution, they resume swimming. Explain.

10. *Going straight.* Suppose you measure the mean distance that an *E. coli* bacterium moves along a straight path before it tumbles over a period of time. Would you expect this distance to change in the presence of a gradient of a chemoattractant? Would it be longer or shorter?

11. *Hauling a load.* Consider the action of a single kinesin molecule in moving a vesicle along a microtubule track. The force required to drag a spherical particle of radius a at a velocity v in a medium having a viscosity η *is*

$$F = 6\pi\eta a v$$

Suppose that a 2-μm-diameter bead is carried at a velocity of 0.6 μm s^{-1} in an aqueous medium (η = 0.01 poise = 0.01 g cm^{-1} s^{-1}).

(a) What is the magnitude of the force exerted by the kinesin molecule? Express the value in dynes (1 dyne = 1 g cm s^{-2}).

(b) How much work is performed in 1 s? Express the value in ergs (1 erg = 1 dyne cm).

(c) A kinesin motor hydrolyzes approximately 80 molecules of ATP per second. What is the energy associated with the hydrolysis of this much ATP in ergs? Compare this value with the actual work performed.

12. *Unusual strides.* A publication describes a kinesin molecule that is claimed to move along microtubules with a step size of 6 nm. You are skeptical. Why?

13. *The sound of one hand clapping.* KIF1A is a motor protein that moves toward the plus end of microtubules *as a monomer.* KIF1A has only a single motor domain. What additional structural elements would you expect to find in the KIF1A structure?

14. *Building blocks.* Actin filaments, microtubules, and bacterial flagella are all built from small subunits. Describe three advantages of assembling long filamentous structures from subunits rather than from single, long proteins.

Mechanism Problem

15. *Backward rotation.* On the basis of the proposed structure in Figure 35.30 for the bacterial flagellar motor, suggest

a pathway for transmembrane proton flow when the flagellar motor is rotating clockwise rather than counterclockwise.

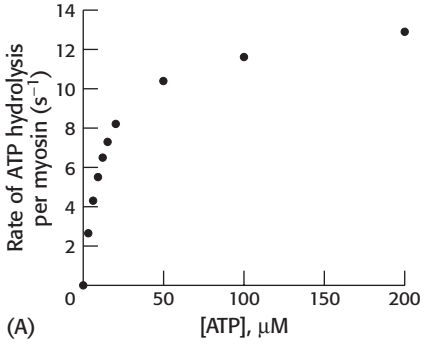
Chapter Integration Problem

16. *Smooth muscle.* Smooth muscle, in contrast with skeletal muscle, is not regulated by a tropomyosin–troponin mechanism. Instead, vertebrate smooth-muscle contraction is controlled by the degree of phosphorylation of its light chains. Phosphorylation induces contraction, and dephosphorylation leads to relaxation. Like that of skeletal muscle, smooth-muscle contraction is triggered by an increase in the cytoplasmic calcium ion level. Propose a mechanism for this action of calcium ion on the basis of your knowledge of other signal-transduction processes.

Data Interpretation Problem

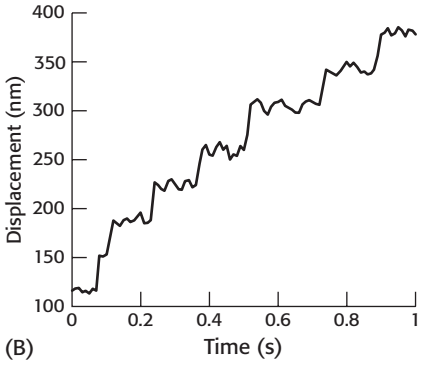
17. *Myosin V.* An abundant myosin-family member, myosin V is isolated from brain tissue. This myosin has a number of unusual properties. First, on the basis of its amino acid sequence, each heavy chain has six tandem binding sites for calmodulin-like light chains. Second, it forms dimers but not higher-order oligomers. Finally, unlike almost all other myosin-family members, myosin V is highly processive.

The rate of ATP hydrolysis by myosin has been examined as a function of ATP concentration, as shown in graph A.



(a) Estimate the values of k_{cat} and K_M for ATP.

With the use of optical-trap measurements, the motion of single myosin V dimers could be followed, as shown in graph B.



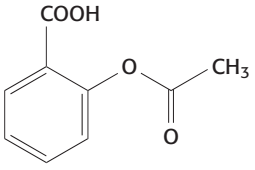
[Based on M. Rief et al. *Proc. Natl. Acad. Sci. U. S. A.* 97:9482–9486, 2000.]

(b) Estimate the step size for myosin V.

The rate of ADP release from myosin V is found to be approximately 13 molecules per second.

(c) Combine the observations about the amino acid sequence of myosin, the observed step size, and the kinetics results to propose a mechanism for the processive motion of myosin V.

Drug Development



Many drugs are based on natural products. Aspirin (above) is a chemical derivative of a compound isolated from willow bark (near left). Extracts of willow bark had long been known to have medicinal properties. The active compound was isolated, modified, and, beginning in 1899, packaged for consumers (far left). [(Far left) Used with permission of Bayer Corporation. (Near left) Image Ideas/ Picture Quest.]

The development of drugs is among the most important interfaces between biochemistry and medicine. In most cases, drugs act by binding to specific receptors or enzymes and inhibiting or otherwise modulating their activities. Thus, knowledge of these molecules and the pathways in which they participate is crucial to drug development. An effective drug is much more than a potent modulator of its target, however. Drugs must be readily administered to patients, preferably as small tablets taken orally, and must survive within the body long enough to reach their targets. Furthermore, to prevent unwanted physiological effects, drugs must not modulate the properties of biomolecules other than the target molecules. These requirements tremendously limit the number of compounds that have the potential to be clinically useful drugs.

Drugs have been discovered by two fundamentally opposite approaches (Figure 36.1). The first approach identifies a substance that has a desirable physiological consequence when administered to a human being, to an appropriate animal, or to cells. Such substances can be discovered by serendipity, by the fractionation of plants or other materials known to have medicinal properties, or by screening natural products or other "libraries" of compounds. In this approach, a biological effect is known before the

OUTLINE

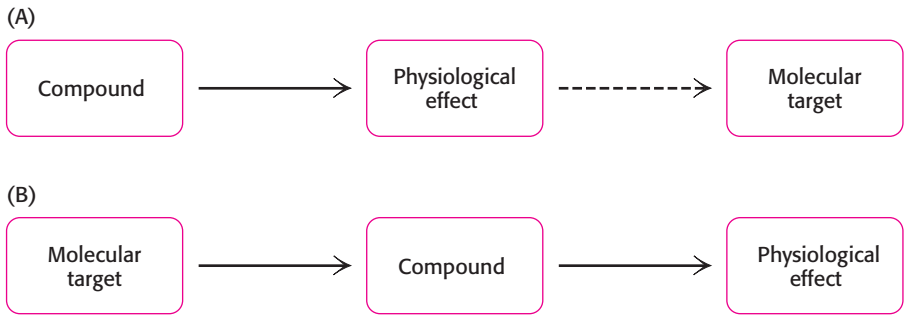


Figure 36.1 Two paths to drug discovery. (A) A compound is discovered to have a desirable physiological effect. The molecular target can be identified in a separate step as needed. (B) A molecular target is selected first. Drug candidates that bind to the target are identified and then examined for their physiological effects.

molecular target is identified. The mode of action of the substance is only later identified after substantial additional work. The second approach begins with a known molecular target. Compounds are sought, either by screening or by designing molecules with desired properties, that bind to the target molecule and modulate its properties. After such compounds are available, scientists can explore their effects on appropriate cells or organisms. Many unexpected results may be encountered in this process as the complexity of biological systems reveals itself.

In this chapter, we explore the science of pharmacology. We examine a number of case histories that illustrate drug development—including many of its concepts, methods, and challenges. We then see how the concepts and tools from genomics are influencing approaches to drug development. The chapter ends with a summary of the stages along the way to developing a drug.

Pharmacology
The science that deals with the discovery, chemistry, composition, identification, biological and physiological effects, uses, and manufacture of drugs.

36.1 The Development of Drugs Presents Huge Challenges

Many compounds have significant effects when taken into the body, but only a very small fraction of them have the potential to be useful drugs. A foreign compound, not adapted to its role in the cell through long evolution, must have a range of special properties to function effectively without causing serious harm. We next consider some of the challenges faced by drug developers.

Drug candidates must be potent modulators of their targets

Most drugs bind to specific proteins, usually receptors or enzymes, within the body. To be effective, a drug needs to bind a sufficient number of its target proteins when taken at a reasonable dose. One factor in determining drug effectiveness is the strength of the interaction between the drug and its target. A molecule that binds to some target molecule is often referred to as a *ligand*. A ligand-binding curve is shown in Figure 36.2. Ligand molecules occupy progressively more target binding sites as ligand concentration increases until essentially all of the available sites are occupied. The tendency of a ligand to bind to its target is measured by the *dissociation constant*, K_d, defined by the expression

$$K_d = [R][L]/[RL]$$

where [R] is the concentration of the free receptor, [L] is the concentration of the free ligand, and [RL] is the concentration of the receptor–ligand complex. The dissociation constant is a measure of the strength of the

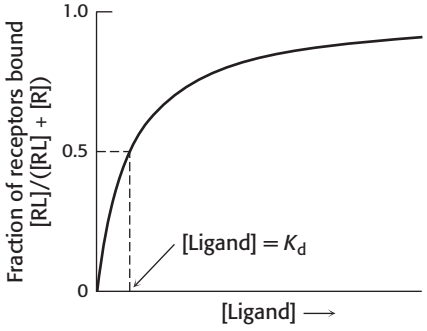


Figure 36.2 Ligand binding. The titration of a receptor, R, with a ligand, L, results in the formation of the complex RL. In uncomplicated cases, the binding reaction follows a simple saturation curve. Half of the receptors are bound to ligand when the ligand concentration equals the dissociation constant, K_d, for the RL complex.

interaction between the drug candidate and the target; the lower the value, the stronger the interaction. The concentration of free ligand at which one-half of the binding sites are occupied equals the dissociation constant, as long as the concentration of binding sites is substantially less than the dissociation constant.

In many cases, biological assays in the context of living cells or tissues (rather than direct enzyme or binding assays) are used to examine the potency of drug candidates. For example, the fraction of bacteria killed by a drug might indicate the potency of a potential antibiotic. In these cases, values such as EC_{50} are used. EC_{50} is the concentration of the drug candidate required to elicit 50% of the maximal biological response (Figure 36.3). Similarly, EC_{90} is the concentration required to achieve 90% of the maximal response. In the example of an antibiotic, EC_{90} would be the concentration required to kill 90% of bacteria exposed to the drug. For drug candidates that are inhibitors, the corresponding terms IC_{50} and IC_{90} are often used to describe the concentrations of the inhibitor required to reduce a response to 50% or 90% of its value in the absence of inhibitor, respectively.

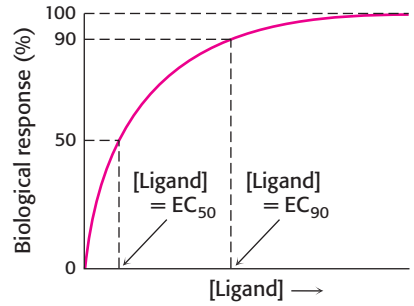


Figure 36.3 Effective concentrations. The concentration of a ligand required to elicit a biological response can be quantified in terms of EC_{50}, the concentration required to give 50% of the maximum response, and EC_{90}, the concentration required to give 90% of the maximum response.

Values such as the IC_{50} and EC_{50} are measures of the potency of a drug candidate in modulating the activity of the desired biological target. To prevent unwanted effects, often called *side effects*, ideal drug candidates should not bind biomolecules other than the target to any significant extent. Developing such a drug can be quite challenging, particularly if the drug target is a member of a large family of evolutionarily related proteins. The degree of specificity can be described in terms of the ratio of the K_d values for the binding of the drug candidate to any other molecules to the K_d value for the binding of the drug candidate to the desired target.

Many complicating factors are present under physiological conditions. Many drug targets also bind ligands normally present in tissues; often, the drug and these ligands will compete for binding sites on the target. We encountered this situation when we considered competitive inhibitors in Chapter 8. Suppose that the drug target is an enzyme and the drug candidate is a competitive inhibitor. The concentration of the drug candidate necessary to inhibit the enzyme will depend on the physiological concentration of the enzyme's normal substrate (Figure 36.4). Biochemists Yung-Chi Cheng and William Prusoff described the relation between the IC_{50} of an enzyme inhibitor and its *inhibition constant* K_i (analogous to the dissociation constant, K_d, of a ligand):

$$IC_{50} = K_i\,(1 + [S]/K_M)$$

This relation, referred to as the *Cheng–Prusoff equation*, demonstrates that the IC_{50} of a competitive inhibitor will depend on the concentration and the Michaelis constant (K_M) for the substrate S. The higher the concentration of the natural substrate, the higher the concentration of drug needed to inhibit the enzyme to a given extent.

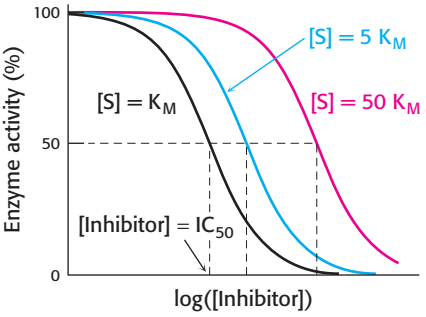


Figure 36.4 Inhibitors compete with substrates for enzyme active sites. The measured IC_{50} of a competitive inhibitor to its target enzyme depends on the concentration of substrate present.

Drugs must have suitable properties to reach their targets

Thus far, we have focused on the ability of molecules to act on specific target molecules. However, an effective drug must also have other characteristics. It must be easily administered and must reach its target at sufficient concentration to be effective. A drug molecule encounters a variety of obstacles on its way to its target, related to its absorption, distribution, metabolism, and excretion after it has entered the body. These processes are interrelated, as summarized in Figure 36.5. Taken together, a drug's ease of absorption, distribution, metabolism, and excretion are often referred to as *ADME* (pronounced "add-me") properties.

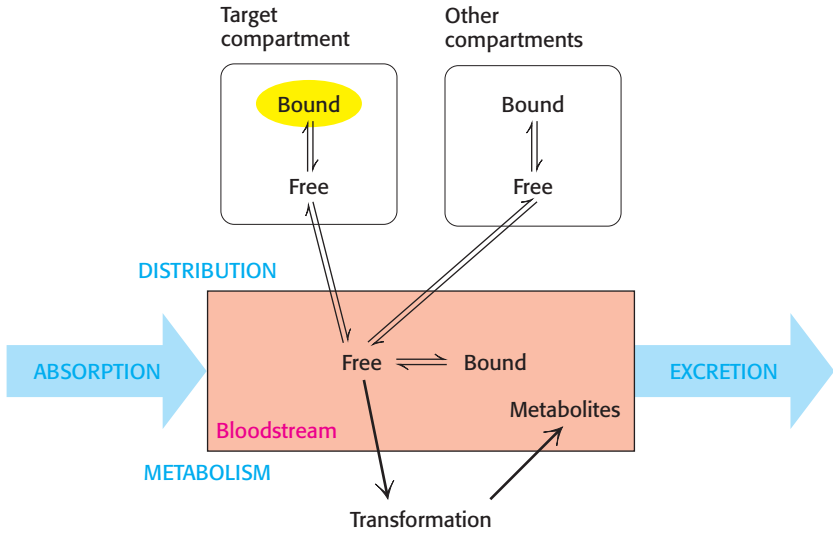
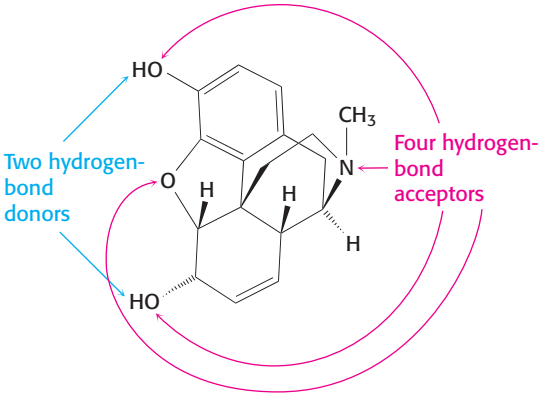


Figure 36.5 Absorption, distribution, metabolism, and excretion (ADME). The concentration of a compound at its target site (yellow) is affected by the extents and rates of absorption, distribution, metabolism, and excretion.

Administration and absorption. Ideally, a drug can be taken orally as a small tablet. An orally administered active compound must be able to survive the acidic conditions in the gut and then be absorbed through the intestinal epithelium. Thus, the compound must be able to pass through cell membranes at a significant rate. Larger molecules such as proteins cannot be administered orally, because they often cannot survive the acidic conditions in the stomach and, if they do, are not readily absorbed. Even many small molecules are not absorbed well; they may be too polar and not pass through cell membranes readily, for example. The ability to be absorbed is often quantified in terms of the *oral bioavailability*. This quantity is defined as the ratio of the peak concentration of a compound given orally to the peak concentration of the same dose injected directly into the bloodstream. Bioavailability can vary considerably from species to species, and so results from animal studies may be difficult to apply to human beings. Despite this variability, some useful generalizations have been made. One powerful set consists of *Lipinski's rules*.

Lipinski's rules tell us that poor absorption is likely when

1. the molecular weight is greater than 500.
2. the number of hydrogen-bond donors is greater than 5.
3. the number of hydrogen-bond acceptors is greater than 10.
4. the partition coefficient [measured as log(P)] is greater than 5.



Morphine ($C_{17}H_{19}O_3N$)

Molecular weight = 285

log(P) = 1.27

Figure 36.6 Lipinski's rules applied to morphine. Morphine satisfies all of Lipinski's rules and has an oral bioavailability in human beings of 33%.

The partition coefficient is a way to measure the tendency of a molecule to dissolve in membranes, which correlates with its ability to dissolve in organic solvents. It is determined by allowing a compound to equilibrate between water and an organic phase, *n*-octanol. The log(P) value is defined as $\log_{10}$ of the ratio of the concentration of a compound in *n*-octanol to the concentration of the compound in water. For example, if the concentration of the compound in the *n*-octanol phase is 100 times that in the aqueous phase, then log(P) is 2. Although the ability of a drug to partition in organic solvents is ideal, because it implies that the compound can penetrate membranes, a log(P) value that is too high suggests that the molecule may be poorly soluble in an aqueous environment.

Morphine, for example, satisfies all of Lipinski's rules and has moderate bioavailability (Figure 36.6). A drug that violates one or

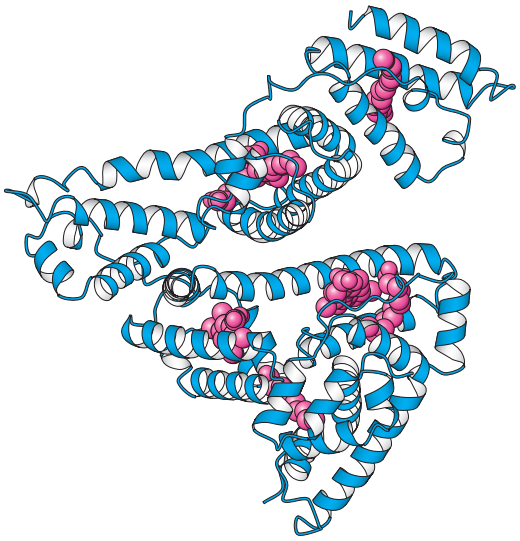


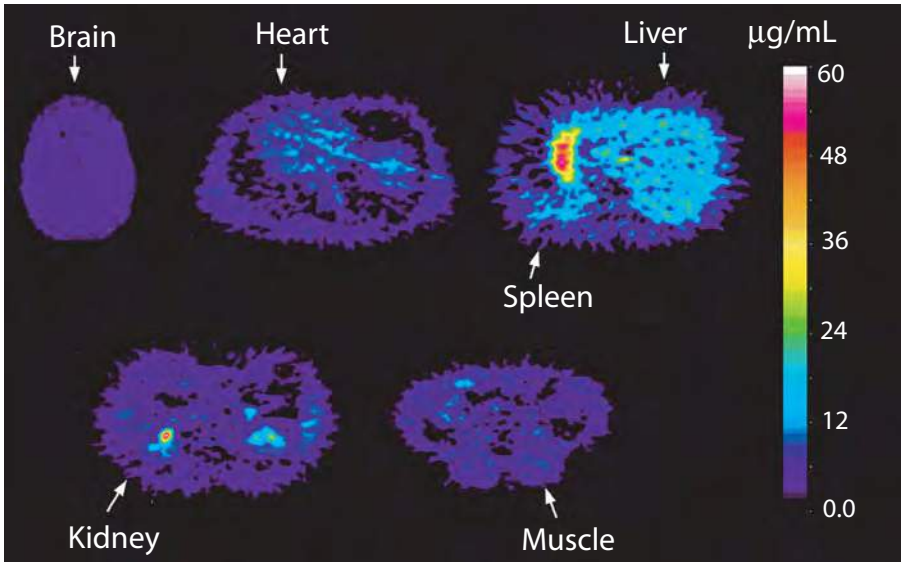
Figure 36.7 Structure of the drug carrier human serum albumin. Seven hydrophobic molecules (in red) are shown bound to the molecule. [Drawn from 1BKE.pdb.]

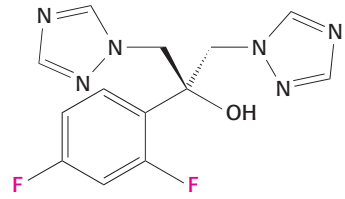
more of these rules may still have satisfactory bioavailability. Nonetheless, these rules serve as guiding principles for evaluating new drug candidates.

Distribution. Compounds taken up by intestinal epithelial cells can pass into the bloodstream. However, hydrophobic compounds and many others do not freely dissolve in the bloodstream. These compounds bind to proteins, such as albumin (Figure 36.7), that are abundant in the blood serum and by this means are carried everywhere that the bloodstream goes.

When a compound has reached the bloodstream, it is distributed to different fluids and tissues, which are often referred to as *compartments.* Some compounds are highly concentrated in their target compartments, either by binding to the target molecules themselves or by other mechanisms. Other compounds are distributed more widely (Figure 36.8). An effective drug will reach the target compartment in sufficient quantity; the concentration of the compound in the target compartment is reduced whenever the compound is distributed into other compartments.

Some target compartments are particularly hard to reach. Many compounds are excluded from the central nervous system by the *blood–brain barrier,* the tight junctions between endothelial cells that line blood vessels within the brain and spinal cord.





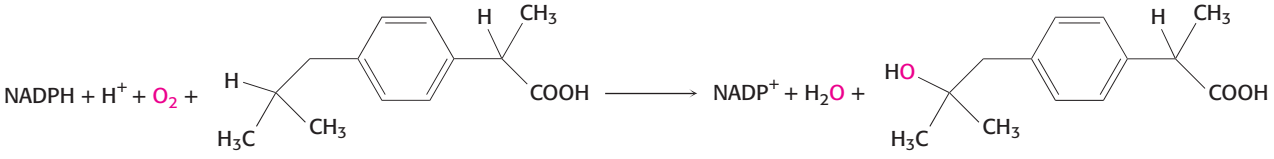
Fluconazole

Figure 36.8 Distribution of the drug fluconazole. After having been taken in, compounds distribute themselves to various organs within the body. The distribution of the antifungal agent fluconazole has been monitored through the use of positron emission tomography (PET) scanning. These images were taken of a healthy human volunteer 90 minutes after injection of a dose of 5 mg kg^{-1} of fluconazole containing trace amounts of fluconazole labeled with the positron-emitting isotope ^{18}F. [From A. J. Fischman et al., *Antimicrob. Agents Chemother.* 37:1270–1277, 1993.]

Metabolism and excretion. A final challenge to a potential drug molecule is to evade the body's defenses against foreign compounds. Many such compounds (called *xenobiotic compounds*) are released from the body in the urine or stool, often after having been metabolized to aid in excretion. This *drug metabolism* poses a considerable threat to drug effectiveness because the concentration of the desired compound decreases as it is metabolized. Thus, a rapidly metabolized compound must be administered more frequently or at higher doses.

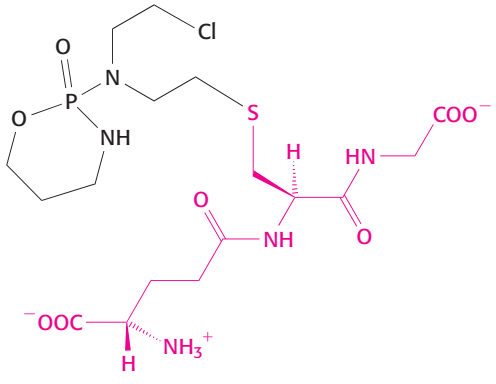
Two of the most common pathways in xenobiotic metabolism are *oxidation* and *conjugation*. Oxidation reactions can aid excretion in at least two ways: by increasing water solubility, and thus ease of transport, and by introducing functional groups that participate in subsequent metabolic steps. These reactions are often promoted by cytochrome P450 enzymes in the liver (Section 26.4). The human genome encodes more than 50 different P450 isozymes, many of which participate in xenobiotic metabolism. A typical reaction catalyzed by a P450 isozyme is the hydroxylation of ibuprofen (Figure 36.9).

Figure 36.9 P450 conversion of ibuprofen. Cytochrome P450 isozymes, primarily in the liver, catalyze xenobiotic metabolic reactions such as hydroxylation. The reaction introduces an oxygen atom derived from molecular oxygen.

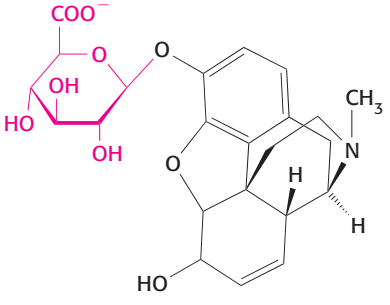


Ibuprofen

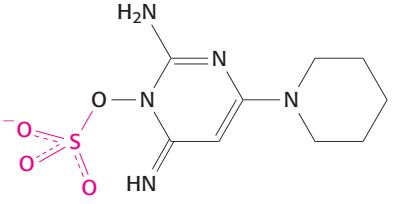
Conjugation is the addition of particular groups to the xenobiotic compound. Common groups added are glutathione (Section 20.5), glucuronic acid, and sulfate (Figure 36.10). These additions often increase water solubility and provide labels that can be recognized to target excretion. Examples of conjugation include the addition of glutathione to the anticancer drug cyclophosphamide, the addition of glucuronidate to the analgesic morphine, and the addition of a sulfate group to the hair-growth stimulator minoxidil.



Cyclophosphamide-glutathione conjugate



Morphine glucuronidate



Minoxidil sulfate

Interestingly, the sulfation of minoxidil produces a compound that is more active in stimulating hair growth than is the unmodified compound. Thus, the metabolic products of a drug, though usually less active than the drug, can sometimes be more active.

Note that an oxidation reaction often precedes conjugation because the oxidation reaction can generate hydroxyl and other groups to which groups

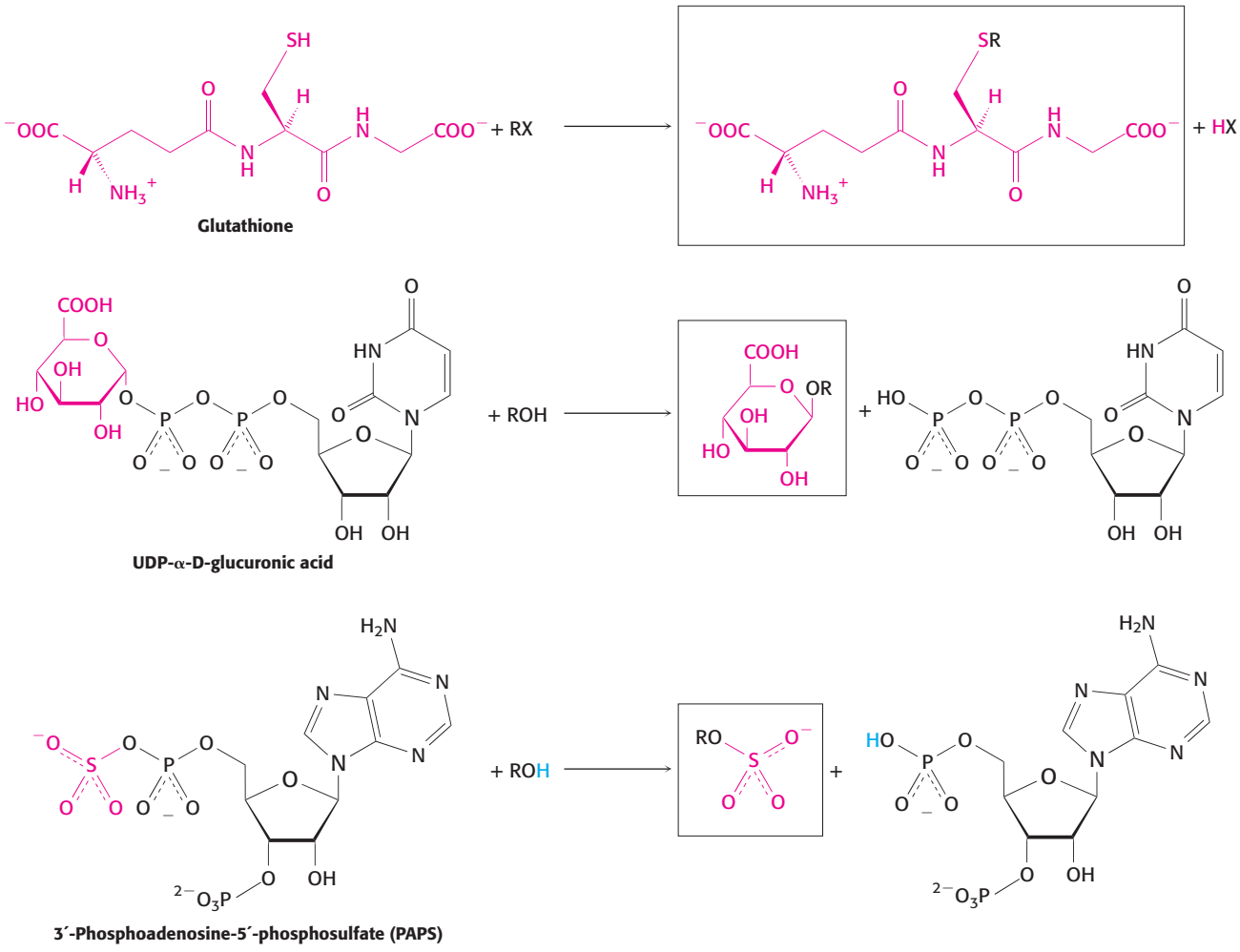


Figure 36.10 Conjugation reactions. Compounds that have appropriate groups are often modified by conjugation reactions. Such reactions include the addition of glutathione (top), glucuronic acid (middle), or sulfate (bottom). The conjugated product is shown boxed.

such as glucuronic acid can be added. The oxidation reactions of xenobiotic compounds are often referred to as *phase I transformations,* and the conjugation reactions are referred to as *phase II transformations.* These reactions take place primarily in the liver. Because blood flows from the intestine directly to the liver through the portal vein, xenobiotic metabolism often alters drug compounds before they ever reach full circulation. This *first-pass metabolism* can substantially limit the availability of compounds taken orally.

After compounds have entered the bloodstream, they can be removed from circulation and excreted from the body by two primary pathways. First, they can be absorbed through the kidneys and excreted in the urine. In this process, the blood passes through *glomeruli,* networks of fine capillaries in the kidney that act as filters. Compounds with molecular weights less than approximately 60,000 pass though the glomeruli. Many of the water molecules, glucose molecules, nucleotides, and other low-molecular-weight compounds that pass through the glomeruli are reabsorbed into the bloodstream, either by transporters that have broad specificities or by the passive transfer of hydrophobic molecules through membranes. Drugs and metabolites that pass through the first filtration step and are not reabsorbed are excreted.

Second, compounds can be actively transported into bile, a process that takes place in the liver. After concentration in the gall bladder, bile flows into the intestine. In the intestine, the drugs and metabolites can be excreted through the stool, reabsorbed into the bloodstream, or further degraded by digestive enzymes. Sometimes, compounds are recycled from

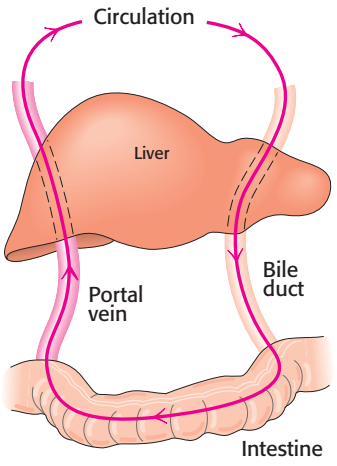


Figure 36.11 Enterohepatic cycling. Some drugs can move from the blood circulation to the liver, into the bile, into the intestine, to the liver, and back into circulation. This cycling decreases the rate of drug excretion.

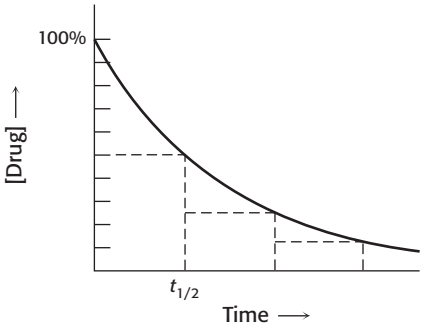


Figure 36.12 Half-life of drug excretion. In the case shown, the concentration of a drug in the bloodstream decreases to one-half of its value in a period of time, $t_{1/2}$, referred to as its half-life.

the bloodstream into the intestine and back into the bloodstream, a process referred to as *enterohepatic cycling* (Figure 36.11). This process can significantly decrease the rate of excretion of some compounds because they escape from an excretory pathway and reenter the circulation.

The kinetics of compound excretion is often complex. In some cases, a fixed percentage of the remaining compound is excreted over a given period of time (Figure 36.12). This pattern of excretion results in exponential loss of the compound from the bloodstream that can be characterized by a half-life ($t_{1/2}$). The half-life is the fixed period of time required to eliminate 50% of the remaining compound. It is a measure of how long an effective concentration of the compound remains in the system after administration. As such, the half-life is a major factor in determining how often a drug must be taken. A drug with a long half-life might need to be taken only once per day, whereas a drug with a short half-life might need to be taken three or four times per day.

Toxicity can limit drug effectiveness

An effective drug must not be so toxic that it seriously harms the person who takes it. A drug may be toxic for any of several reasons. First, it may modulate the target molecule itself *too* effectively. For example, the presence of too much of the anticoagulant drug coumadin can result in dangerous, uncontrolled bleeding and death. Second, the compound may modulate the properties of proteins that are distinct from, but related to, the target molecule itself. Compounds that are directed to one member of a family of enzymes or receptors often bind to other family members. For example, an antiviral drug directed against viral proteases may be toxic if it also inhibits proteases normally present in the body such as those that regulate blood pressure.

A compound may also be toxic if it modulates the activity of a protein unrelated to its intended target. For example, many compounds block ion channels such as the potassium channel hERG (p. 393), causing potentially life threatening disturbances of the heartbeat. To prevent cardiac side effects, many compounds are screened for their ability to block such channels.

Finally, even if a compound is not itself toxic, its metabolic by-products may be. Phase I metabolic processes can generate damaging reactive groups in products. An important example is liver toxicity observed with large doses of the common pain reliever acetaminophen (Figure 36.13). A particular cytochrome P450 isozyme oxidizes acetaminophen to *N*-acetyl-*p*-benzoquinone imine. The resulting compound is conjugated to glutathione. With large doses, however, the liver concentration of glutathione drops dramatically, and the liver is no longer able to protect itself from this reactive compound and others. Initial symptoms of excessive acetaminophen include nausea and vomiting. Within 24 to 48 hours, symptoms of liver failure may appear. Acetaminophen poisoning accounts for about 35% of cases of severe liver failure in the United States. A liver transplant is often the only effective treatment.

The toxicity of a drug candidate can be described in terms of the *therapeutic index*. This measure of toxicity is determined through animal tests, usually with mice or rats. The therapeutic index is defined as the ratio of the dose of a compound that is required to kill one-half of the animals (referred to as the LD_{50} for "lethal dose") to a comparable measure of the effective dose, usually the EC_{50}. Thus, if the therapeutic index is 1000, then lethality is significant only when 1000 times the effective dose is administered. Analogous indices can provide measures of toxicity less severe than lethality.

Many compounds have favorable properties in vitro, yet fail when administered to a living organism because of difficulties with ADME and

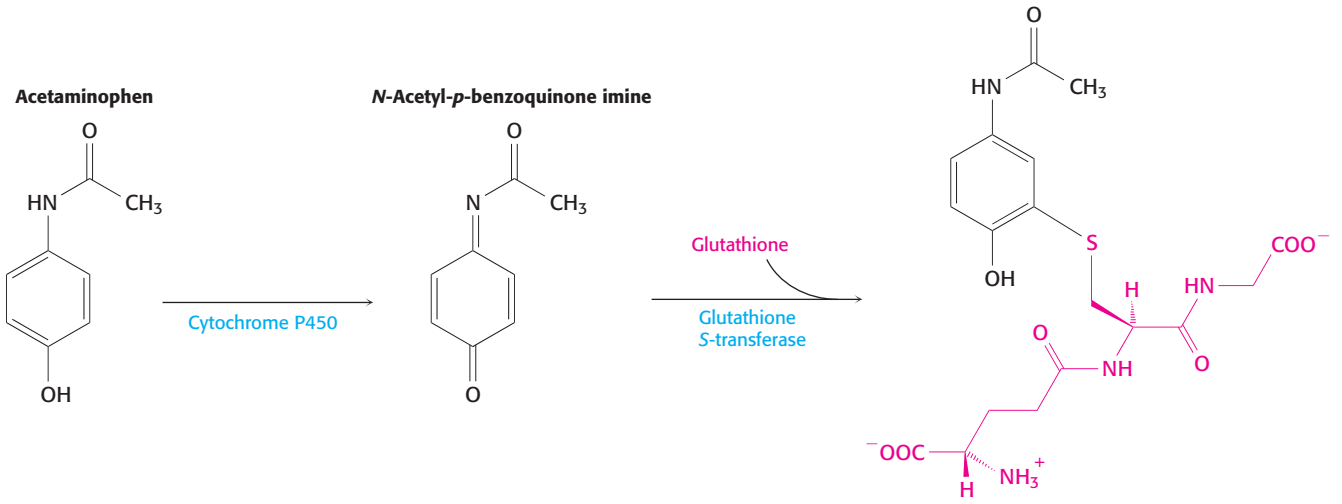


Figure 36.13 Acetaminophen toxicity. A minor metabolic product of acetaminophen is *N*-acetyl-*p*-benzoquinone imine. This metabolite is conjugated to glutathione. Large doses of acetaminophen can deplete liver glutathione stores.

toxicity. Expensive and time-consuming animal studies are required to verify that a drug candidate is not toxic, yet differences between animal species in their response can confound decisions about moving forward with a compound toward human studies. One hope is that, with more understanding of the biochemistry of these processes, scientists can develop computer-based models to replace or augment animal tests. Such models would need to accurately predict the fate of a compound inside a living organism from its molecular structure or other properties that are easily measured in the laboratory without the use of animals.

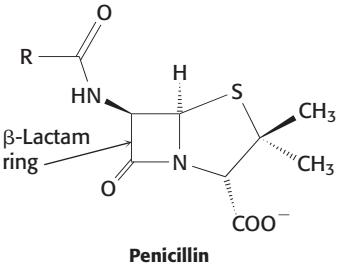
36.2 Drug Candidates Can Be Discovered by Serendipity, Screening, or Design

Traditionally, many drugs were discovered by serendipity, or chance observation. More recently, drugs have been discovered by screening collections of natural products or other compounds for compounds that have desired medicinal properties. Alternatively, scientists have designed specific drug candidates by using their knowledge about a preselected molecular target. We will examine several examples of each of these pathways to reveal common principles.

Serendipitous observations can drive drug development

Perhaps the most well known observation in the history of drug development is Alexander Fleming's chance observation in 1928 that colonies of the bacterium *Staphylococcus aureus* died when they were adjacent to colonies of the mold *Penicillium notatum*. Spores of the mold had landed accidentally on plates growing the bacteria. Fleming soon realized that the mold produced a substance that could kill disease-causing bacteria. This discovery led to a fundamentally new approach to the treatment of bacterial infections. Howard Florey and Ernest Chain developed a powdered form of the substance, termed penicillin, that became a widely used antibiotic in the 1940s.

The structure of this antibiotic was elucidated in 1945. The most notable feature of this structure is the four-membered β-lactam ring. This unusual feature is key to the antibacterial function of penicillin, as noted earlier (Section 8.5).



Penicillin

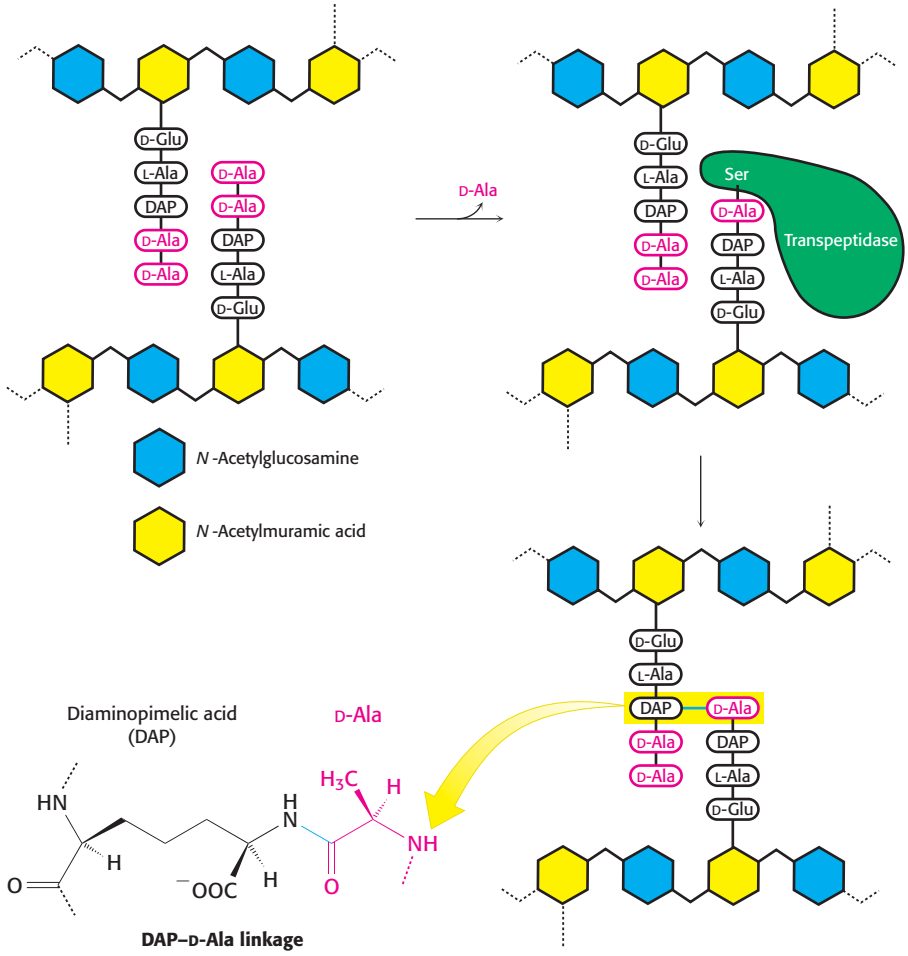
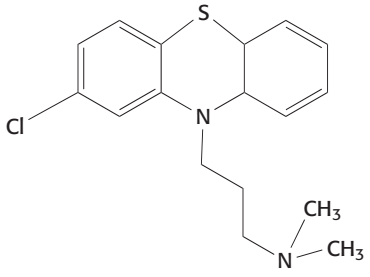
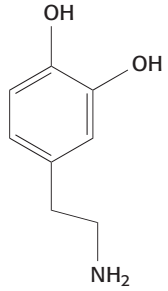


Figure 36.14 Mechanism of cell-wall biosynthesis disrupted by penicillin. A transpeptidase enzyme catalyzes the formation of cross-links between peptidoglycan groups. In the case shown, the transpeptidase catalyzes the linkage of D-alanine at the end of one peptide chain to the amino acid diaminopimelic acid (DAP) on another peptide chain. The diaminopimelic acid linkage (bottom left) is found in Gram-negative bacteria such as *E. coli.* Linkages of glycine-rich peptides are found in Gram-positive bacteria. Penicillin inhibits the action of the transpeptidase; so bacteria exposed to the drug have weak cell walls that are susceptible to lysis.



Chlorpromazine



Dopamine

Three steps were crucial to fully capitalize on Fleming's discovery. First, an industrial process was developed for the production of penicillin from *Penicillium* mold on a large scale. Second, penicillin and penicillin derivatives were chemically synthesized. The availability of synthetic penicillin derivatives opened the way for scientists to explore the relations between structure and function. Many such penicillin derivatives have found widespread use in medicine. Finally, in 1965, Jack Strominger and James Park independently determined that penicillin exerts its antibiotic activity by blocking a critical transpeptidase reaction in bacterial cell-wall biosynthesis (Figure 36.14), as introduced in Section 8.5.

Many other drugs have been discovered by serendipitous observations. The antineuroleptic drug chlorpromazine (Thorazine) was discovered in the course of investigations directed toward the treatment of shock in surgical patients. In 1952, French surgeon Henri Laborit noticed that, after taking the compound, his patients were remarkably calm. This observation suggested that chlorpromazine could benefit psychiatric patients, and, indeed, the drug has been used for many years to treat patients with schizophrenia and other disorders. The drug does have significant side effects, and its use has been largely superceded by more recently developed drugs.

Chlorpromazine acts by binding to receptors for the neurotransmitter dopamine and blocking them (Figure 36.15). Dopamine D2 receptors are the targets of many other psychoactive drugs. In the search for drugs with more-limited side effects, studies are undertaken to correlate drug effects with biochemical parameters such as dissociation constants and binding and release rate constants.

A more-recent example of a drug discovered by chance observation is sildenafil (Viagra). This compound was developed as an inhibitor of phosphodiesterase 5, an enzyme that catalyzes the hydrolysis of cGMP to GMP (Figure 36.16). The compound was intended as a treatment for hypertension and angina because cGMP plays a central role in the relaxation of smooth-muscle cells in blood vessels (Figure 36.17). The inhibition of phosphodiesterase 5 was expected to increase the concentration of cGMP by blocking the pathway for its degradation. In the course of early clinical trials in Wales, some men reported unusual penile erections. Whether this chance observation by a few men was due to the compound or to other effects was unclear. However, the observation made some biochemical sense because smooth-muscle relaxation due to increased cGMP levels had been discovered to play a role in penile erection. Subsequent clinical trials directed toward the evaluation of sildenafil for erectile dysfunction were successful. This account testifies to the importance of collecting comprehensive information from clinical-trial participants. In this case, incidental observations led to a new treatment for erectile dysfunction and a multibillion-dollar-per-year drug market.

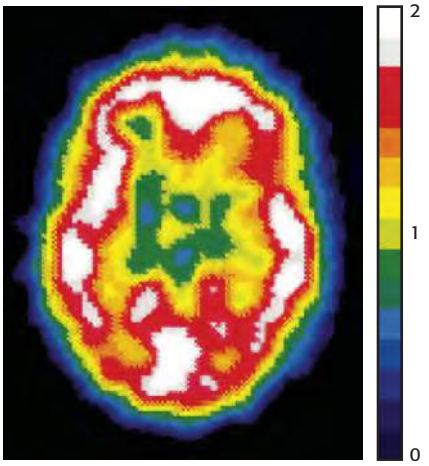


Figure 36.15 Chlorpromazine targets. This positron emission tomographic image shows the distribution of dopamine D2 receptors in the brain. The colors on the bar at the right represent the relative distribution of the receptor: white regions have the highest concentration; dark blue regions have no receptors. These sites are blocked by treatment with chlorpromazine. [From C. Trichard et al. *Am. J. Psychiatry* 155:505–508, 1998; reprinted with permission conveyed through Copyright Clearance Center, Inc.]

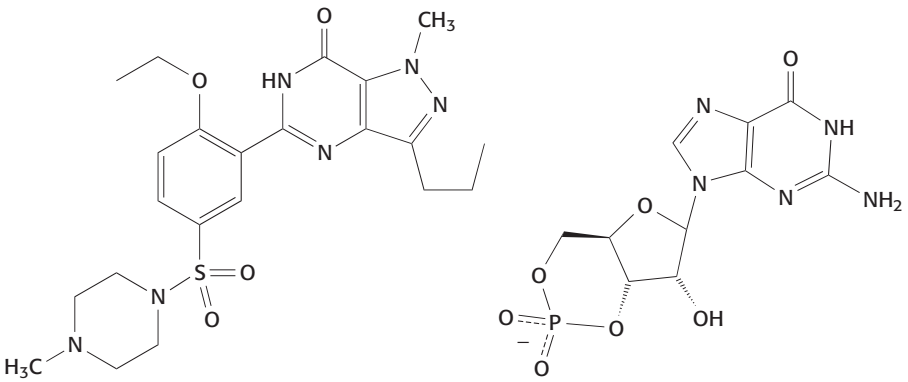


Figure 36.16 Sildenafil, a mimic of cGMP. Sildenafil was designed to resemble cGMP, the substrate of phosphodiesterase 5.

Screening libraries of compounds can yield drugs or drug leads

No drug is as widely used as aspirin. Observers at least as far back as Hippocrates (~400 B.C.) have noted the use of extracts from the bark and leaves of the willow tree for pain relief. In 1829, a mixture called *salicin* was isolated from willow bark. Subsequent analysis identified salicylic acid as the active component of this mixture. Salicylic acid was formerly used to treat pain, but this compound often irritated the stomach. Several investigators attempted to find a means to neutralize salicylic acid. Felix Hoffmann, a

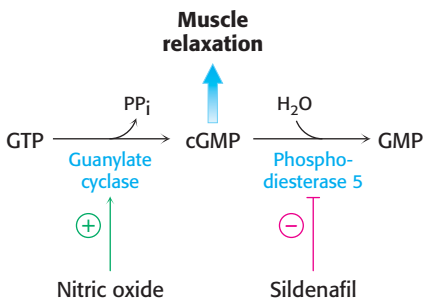
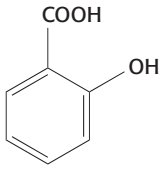
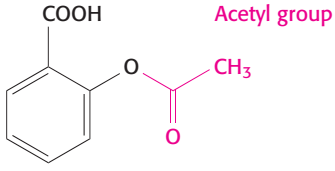


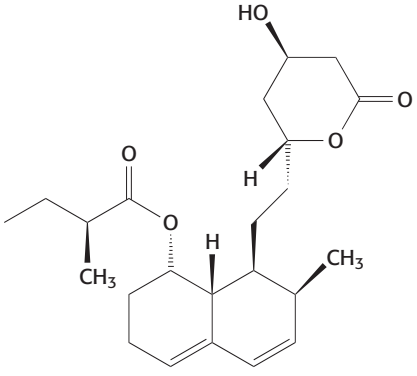
Figure 36.17 Muscle-relaxation pathway. Increases in nitric oxide levels stimulate guanylate cyclase, which produces cGMP. The increased cGMP concentration promotes smooth-muscle relaxation. Phosphodiesterase 5 hydrolyzes cGMP, which lowers the cGMP concentration. The inhibition of phosphodiesterase 5 by sildenafil maintains elevated levels of cGMP.



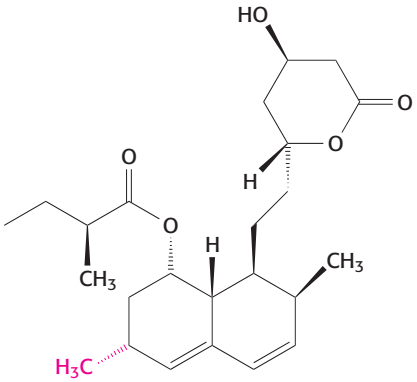
Salicylic acid



Aspirin (acetylsalicylic acid)



Compactin



Lovastatin

chemist working at the German company Bayer, developed a less-irritating derivative by treating salicylic acid with a base and acetyl chloride. This derivative, acetylsalicylic acid, was named *aspirin* from "a" for acetyl chloride, "spir" for *Spiraea ulmaria* (meadowsweet, a flowering plant that also contains salicylic acid), and "in" (a common ending for drugs). Each year, approximately 35,000 tons of aspirin are taken worldwide, nearly the weight of the *Titanic*.

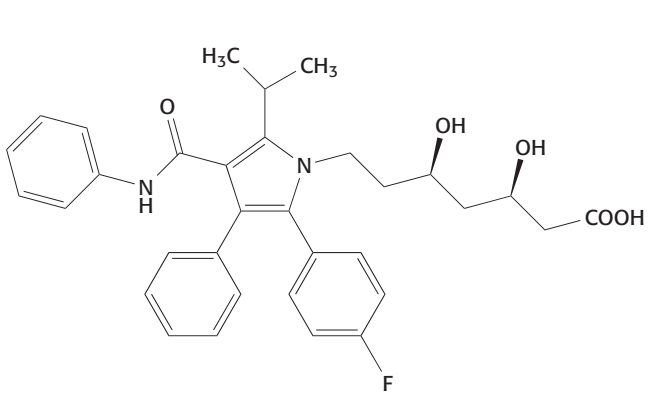
As discussed in Chapter 12, the acetyl group in aspirin is transferred to the side chain of a serine residue that lies along the path to the active site of the cyclooxygenase component of prostaglandin H_2 synthase (see Figure 12.25). In this position, the acetyl group blocks access to the active site. Thus, even though aspirin binds in the same pocket on the enzyme as salicylic acid, the acetyl group of aspirin dramatically increases its effectiveness as a drug. The account illustrates the value of screening extracts from plants and other materials that are believed to have medicinal properties for active compounds. The large number of herbal and folk medicines are a treasure trove of new drug leads.

More than 100 years ago, a fatty, yellowish material was discovered on the arterial walls of patients who had died of vascular disease. The presence of the material was termed *atheroma* from the Greek word for porridge. This material proved to be cholesterol. The Framingham heart study, initiated in 1948, documented a correlation between high blood-cholesterol levels and high mortality rates from heart disease. This observation led to the notion that blocking cholesterol synthesis might lower blood-cholesterol levels and, in turn, lower the risk of heart disease. Drug developers had to abandon an initial attempt at blocking the cholesterol-synthesis pathway at a late step because cataracts and other side effects developed, caused by the accumulation of the insoluble substrate for the inhibited enzyme. Investigators eventually identified a more-favorable target—namely, the enzyme HMG-CoA reductase (Section 26.2). This enzyme acts on a substrate, HMG-CoA (3-hydroxy-3-methylglutaryl coenzyme A), that can be used by other pathways and is water soluble.

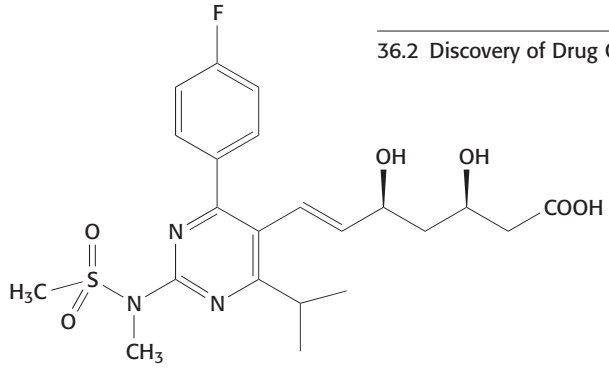
A promising natural product, compactin, was discovered in a screen of compounds from a fermentation broth from *Penicillium citrinum* in a search for antibacterial agents. In some, but not all, animal studies, compactin was found to inhibit HMG-CoA reductase and to lower serum-cholesterol levels. In 1982, a new HMG-CoA reductase inhibitor was discovered in a fermentation broth from *Aspergillus cereus*. This compound, now called lovastatin, was found to be structurally very similar to compactin, bearing one additional methyl group.

In clinical trials, lovastatin significantly reduced serum-cholesterol levels with few side effects. Most side effects could be prevented by treatment with mevalonate (the product of HMG-CoA reductase), indicating that the side effects were likely due to the highly effective blocking of HMG-CoA reductase. One notable side effect is muscle pain or weakness (termed *myopathy*), although its cause remains to be fully established. After many studies, the Food and Drug Administration (FDA) approved lovastatin for treating high serum-cholesterol levels.

A structurally related HMG-CoA reductase inhibitor was later shown to cause a statistically significant decrease in deaths due to coronary heart disease. This result validated the benefits of lowering serum-cholesterol levels. Further mechanistic analysis revealed that the HMG-CoA reductase inhibitor acts not only by lowering the rate of cholesterol biosynthesis, but also by inducing the expression of the low-density-lipoprotein (LDL) receptor (Section 26.3). Cells with such receptors remove LDL particles from the bloodstream, and so these particles cannot contribute to atheroma.



Atorvastatin



Rosuvastatin

Figure 36.18 Synthetic statins. Atorvastatin (Lipitor) and rosuvastatin (Crestor) are completely synthetic drugs that inhibit HMG-CoA reductase.

Lovastatin and its relatives are natural products or they are compounds readily derived from natural products. The next step was the development of totally synthetic molecules that are more-potent inhibitors of HMG-CoA reductase (Figure 36.18). These compounds are effective at lower dose levels, reducing side effects.

The original HMG-CoA reductase inhibitors or their precursors were found by screening libraries of natural products. More recently, drug developers have tried screening large libraries of both natural products and purely synthetic compounds prepared in the course of many drug-development programs. Under favorable circumstances, hundreds of thousands or even millions of compounds can be tested in this process, termed *high-throughput screening*. Compounds in these libraries can be synthesized one at a time for testing. An alternative approach is to synthesize a large number of structurally related compounds that differ from one another at only one or a few positions all at once. This approach is often termed *combinatorial chemistry*. Here, compounds are synthesized with the use of the same chemical reactions but a variable set of reactants. Suppose that a molecular scaffold is constructed with two reactive sites and that 20 reactants can be used in the first site and 40 reactants can be used in the second site. A total of $20 \times 40 = 800$ possible compounds can be produced.

A key method in combinatorial chemistry is *split-pool synthesis* (Figure 36.19). The method depends on solid-phase synthetic methods, first developed for the synthesis of peptides (Section 3.5). Compounds are synthesized on small beads. Beads containing an appropriate starting *scaffold* are produced and divided (split) into n sets, with n corresponding to the number of building blocks to be used at one site. Reactions adding the reactants at the first site are run, and the beads are isolated by filtration. The n sets of beads are then combined (pooled), mixed, and split again into m sets, with m corresponding to the number of reactants to be used at the second site. Reactions adding these m reactants are run, and the beads are again isolated. The important result is that each bead contains only one compound, even though the entire library of beads contains many. Furthermore, although only $n + m$ reactions were run, $n \times m$ compounds are produced. With the preceding values for n and m, $20 + 40 = 60$ reactions produce $20 \times 40 = 800$ compounds. In some cases, assays can be performed directly with the compounds still attached to the bead to find compounds with desired properties (Figure 36.20). Alternatively, each bead can be isolated and the compound can be cleaved from the bead to produce free compounds for analysis. After an interesting compound has been identified, analytical methods of various types must be used to identify which of the $n \times m$ compounds is present.

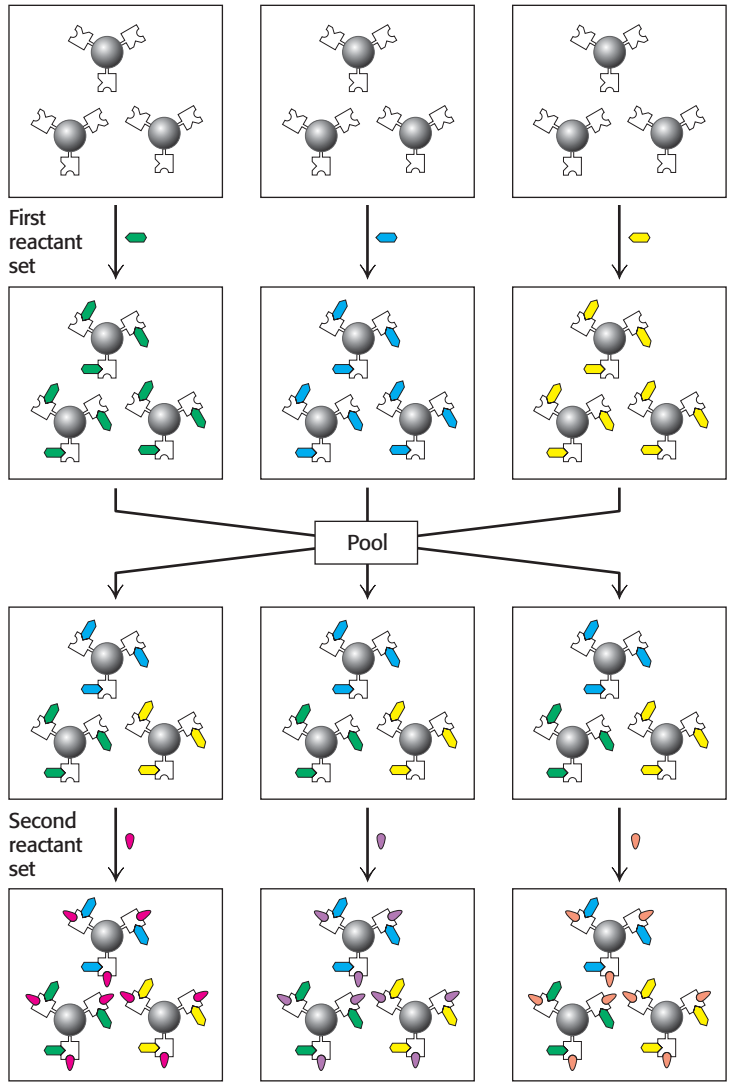


Figure 36.19 Split-pool synthesis. Reactions are performed on beads. Each of the reactions with the first set of reactants is performed on a separate set of beads. The beads are then pooled, mixed, and split into sets. The second set of reactants is then added. Many different compounds will be produced, but all of the compounds on a single bead will be identical.

Note that the "universe" of druglike compounds is vast. More than an estimated 10^{40} compounds are possible with molecular weights less than 750. Thus, even with "large" libraries of millions of compounds, only a tiny fraction of the chemical possibilities are present for study.

Drugs can be designed on the basis of three-dimensional structural information about their targets

Many drugs bind to their targets in a manner reminiscent of Emil Fischer's lock and key (see Figure 8.8). Therefore, we should be able to design a key, given enough knowledge about the shape and chemical composition of the lock. In the idealized case, we would like to design a small molecule that is complementary in shape and electronic structure to a target protein so that it binds effectively to the targeted site. Despite our ability to determine three-dimensional structures rapidly, the achievement of this goal remains in the future. Designing stable compounds from scratch that have the correct shape and other properties to fit precisely into a binding site is difficult because predicting the structure that will best fit into a binding site is difficult. Prediction of binding affinity requires a detailed understanding of the interactions between a compound and its binding partner *and* of the interactions between the compound and the solvent when the compound is free in solution.

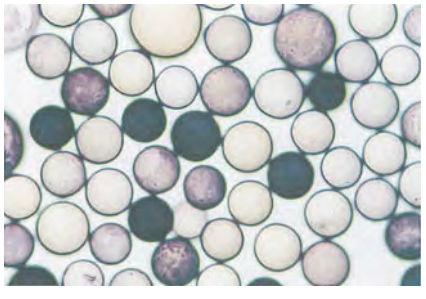


Figure 36.20 Screening a library of synthesized carbohydrates. A small combinatorial library of carbohydrates synthesized on the surface of 130-μm beads is screened for carbohydrates that are bound tightly by a lectin from peanuts. Beads that have such carbohydrates are darkly stained through the action of an enzyme linked to the lectin. [From R. Liang et al., *Proc. Natl. Acad. Sci. U. S. A.* 94:10554–10559, 1997; © 2004 National Academy of Sciences, U. S. A.]

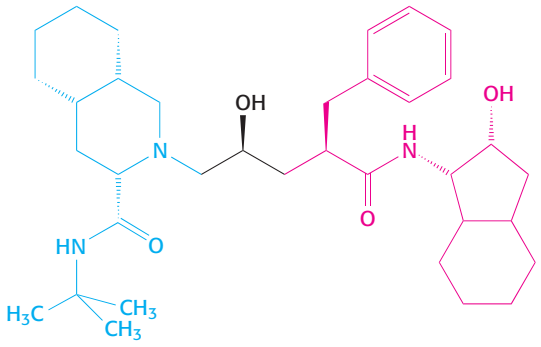
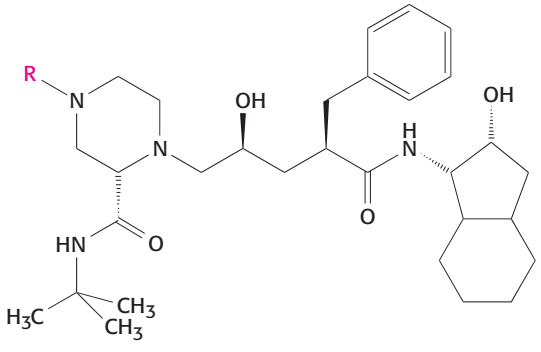


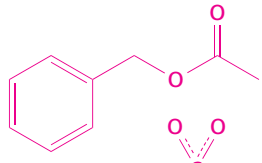
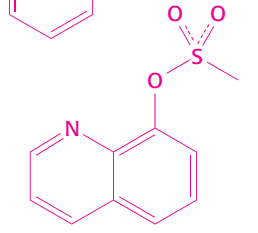
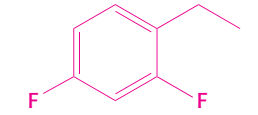
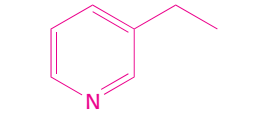
Figure 36.21 Initial design of an HIV protease inhibitor. This compound was designed by combining part of one compound with good inhibition activity but poor solubility (shown in red) with part of another compound with better solubility (shown in blue).

Nonetheless, *structure-based drug design* has proved to be a powerful tool in drug development. Among its most prominent successes has been the development of drugs that inhibit the protease from the HIV virus. Consider the development of the protease inhibitor indinavir (Crixivan; p. 264). Two sets of promising inhibitors that had high potency but poor solubility and bioavailability were discovered. X-ray crystallographic analysis and molecular-modeling findings suggested that a hybrid molecule might have both high potency and improved bioavailability (Figure 36.21). The synthesized hybrid compound did show improvements but required further optimization. The structural data suggested one point where modifications could be tolerated. A series of compounds were produced and examined (Figure 36.22). The most-active compound showed poor bioavailability, but one of the other compounds showed good bioavailability and acceptable activity. The maximum serum concentration available through oral administration was significantly higher than the levels required to suppress replication of the virus. This drug, as well as other protease inhibitors developed at about the same time, has been used in combination with other drugs to treat AIDS with much more encouraging results than had been obtained previously (Figure 36.23).

Aspirin targets the cyclooxygenase site in prostaglandin H_2 synthase, as discussed earlier. Animal studies suggested that mammals contain not one but two distinct cyclooxygenase enzymes, both of which are targeted by

Figure 36.22 Compound optimization. Four compounds are evaluated for characteristics including the IC_{50}, log(*P*), and c_{max} (the maximal concentration of compound present) measured in the serum of dogs. The compound shown at the bottom has the weakest inhibitory power (measured by IC_{50}) but by far the best bioavailability (measured by c_{max}). This compound was selected for further development, leading to the drug indinavir (Crixivan).



R =	IC_{50} (nmol)	log(*P*)	c_{max} (μM)
	0.4	4.67	< 0.1
	0.01	3.70	< 0.1
	0.3	3.69	0.7
	0.6	2.92	11

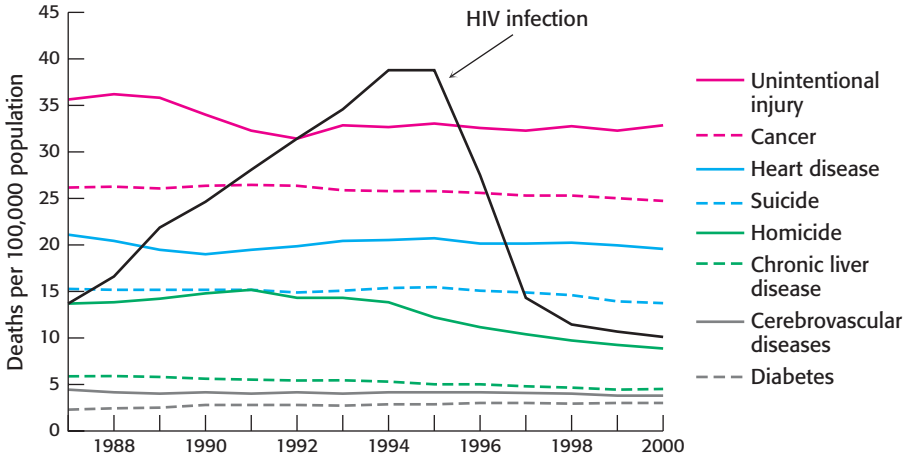
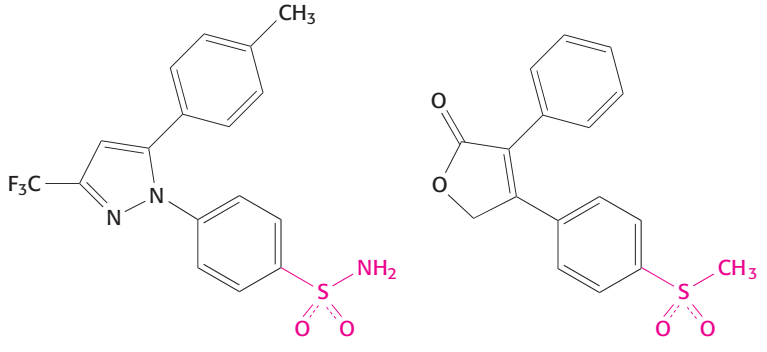


Figure 36.23 The effect of anti-HIV drug development. Death rates from HIV infection (AIDS) reveal the tremendous effect of HIV protease inhibitors and their use in combination with inhibitors of HIV reverse transcriptase. The death rates in this graph are from the leading causes of death among persons 24 to 44 years old in the United States. [From Centers for Disease Control.]

aspirin. The more recently discovered enzyme, cyclooxygenase 2 (COX2), is expressed primarily as part of the inflammatory response, whereas cyclooxygenase 1 (COX1) is expressed more generally. These observations suggested that a cyclooxygenase inhibitor that was specific for COX2 might be able to reduce inflammation in conditions such as arthritis without producing the gastric and other side effects associated with aspirin.

The amino acid sequences of COX1 and COX2 were deduced from cDNA cloning studies. These sequences are more than 60% identical, clearly indicating that the enzymes have the same overall structure. Nevertheless, there are some differences in the residues around the aspirin-binding site. X-ray crystallography revealed that an extension of the binding pocket was present in COX2, but absent in COX1. This structural difference suggested a strategy for constructing COX2-specific inhibitors—namely, to synthesize compounds that had a protuberance that would fit into the pocket in the COX2 enzyme. Such compounds were designed and synthesized and then further refined to produce effective drugs familiar as Celebrex and Vioxx (Figure 36.24). Vioxx was subsequently withdrawn from the market because some people experienced adverse effects. These effects appear to be due to the inhibition of COX2, the intended target. Thus, although the development of these drugs is a triumph for structure-



Celecoxib (Celebrex) **Rofecoxib (Vioxx)**

Figure 36.24 COX2-specific inhibitors. These compounds have protuberances (shown in red) that fit into a pocket in the COX2 isozyme but sterically clash with the COX1 isozyme.

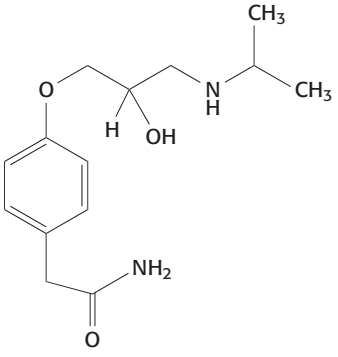
based drug design, these outcomes highlight the fact that the inhibition of important enzymes can lead to complex physiological responses.

36.3 Analyses of Genomes Hold Great Promise for Drug Discovery

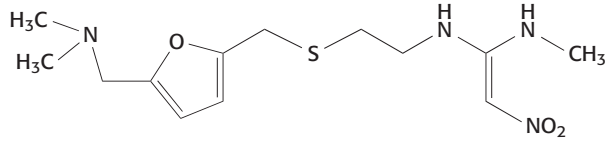
The completion of the sequencing of the human and other genomes is a potentially powerful driving force for the development of new drugs. Genomic sequencing and analysis projects have vastly increased our knowledge of the proteins encoded by the human genome. This new source of knowledge may greatly accelerate early stages of the drug-development process or even allow drugs to be tailored to the individual patient.

Potential targets can be identified in the human proteome

The human genome encodes approximately 25,000 proteins, not counting the variation produced by alternative mRNA splicing and posttranslational modifications. Many of these proteins are potential drug targets, in particular those that are enzymes or receptors and have significant biological effects when activated or inhibited. Several large protein families are particularly rich sources of targets. For example, the human genome includes genes for more than 500 protein kinases that can be recognized by comparing the deduced amino acid sequences. Many of these kinases are known to play a role in the progression of a variety of diseases. For example, Bcr-Abl kinase, a dysregulated kinase formed from a specific chromosomal defect, is known to contribute to certain leukemias and is the target of the drug imatinib mesylate (Gleevec; Section 14.5). Some of the other protein kinases undoubtedly play central roles in particular cancers as well. Similarly, the human genome encodes approximately 800 7TM receptors (Section 14.1) of which approximately 350 are odorant receptors. Many of the remaining 7TM receptors are potential drug targets. Some of them are already targets for drugs, such as the β-blocker atenolol, which targets the β-adrenergic receptor, and the antiulcer medication ranitidine (Zantac). The latter compound is an antagonist of the histamine H_2 receptor, a 7TM receptor that participates in the control of gastric acid secretion.



Atenolol



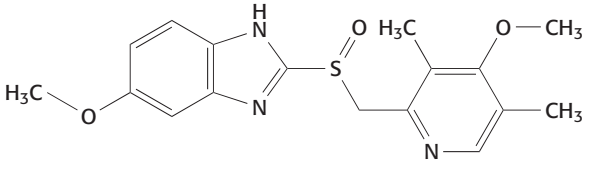
Ranitidine

Novel proteins that are not part of large families already supplying drug targets can be more readily identified through the use of genomic information. There are a number of ways to identify proteins that could serve as targets of drug-development programs. One way is to look for changes in expression patterns, protein localization, or posttranslational modifications in cells from disease-afflicted organisms. Another is to perform studies of tissues or cell types in which particular genes are expressed. Analysis of the

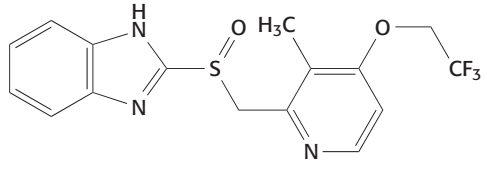
human genome should increase the number of actively pursued drug targets by a factor of an estimated two or more.

Animal models can be developed to test the validity of potential drug targets

The genomes of a number of model organisms have now been sequenced. The most important of these genomes for drug development is that of the mouse. Remarkably, the mouse and human genomes are approximately 85% identical in sequence, and more than 98% of all human genes have recognizable mouse counterparts. Mouse studies provide drug developers with a powerful tool—the ability to disrupt ("knock out") specific genes in the mouse (p. 164). If disruption of a gene has a desirable effect, then the product of this gene is a promising drug target. The utility of this approach has been demonstrated retrospectively. For example, disruption of the gene for the α subunit of the H^+–K^+ ATPase, the key protein for secreting acid into the stomach, produces mice with less acid in their stomachs. The stomach pH of such mice is 6.9 in circumstances that produce a stomach pH of 3.2 in their wild-type counterparts. This protein is the target of the drugs omeprazole (Prilosec) and lansoprazole (Prevacid and Takepron), used for treating gastroesophageal reflux disease.



Omeprazole



Lansoprazole

Several large-scale efforts are underway to generate hundreds or thousands of mouse strains, each having a different gene disrupted. The phenotypes of these mice are a good indication of whether the protein encoded by a disrupted gene is a promising drug target. This approach allows drug developers to evaluate potential targets without any preconceived notions regarding physiological function.

Potential targets can be identified in the genomes of pathogens

Human proteins are not the only important drug targets. Drugs such as penicillin and HIV protease inhibitors act by targeting proteins within a pathogen. The genomes of hundreds of pathogens have now been sequenced, and these genome sequences can be mined for potential targets.

New antibiotics are needed to combat bacteria that are resistant to many existing antibiotics. One approach seeks proteins essential for cell survival that are conserved in a wide range of bacteria. Drugs that inactivate such proteins are expected to be broad-spectrum antibiotics, useful for treating infections from any of a range of different bacteria. One such protein is peptide deformylase, the enzyme that removes formyl groups present at the amino termini of bacterial proteins immediately after translation (see Figure 30.19).

Alternatively, a drug may be needed against a specific pathogen. A recent example of such a pathogen is the organism responsible for severe acute respiratory syndrome (SARS). Within one month of the recognition of this emerging disease, investigators had isolated the virus that causes the syndrome, and, within weeks, its 29,751-base genome had been completely sequenced. This sequence revealed the presence of a gene encoding a viral protease, known to be essential for viral replication from studies of other members of

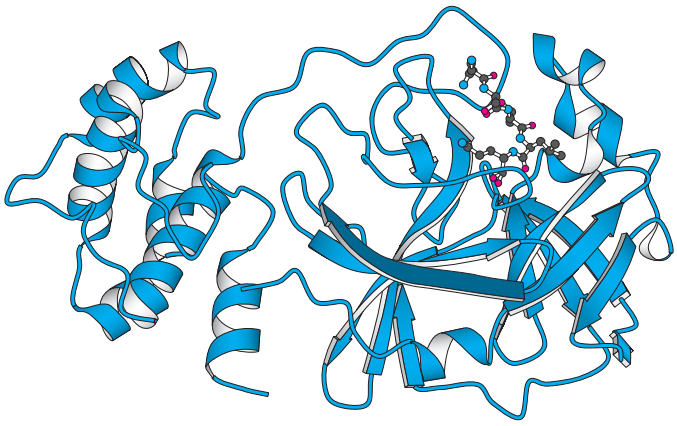


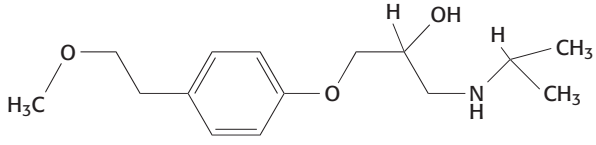
Figure 36.25 Emerging drug target. The structure of a protease from the coronavirus that causes SARS (severe acute respiratory syndrome) is shown bound to an inhibitor. This structure was determined less than a year after the identification of the virus. [Drawn from 1P9S.bdb.]

the coronavirus family to which the SARS virus belongs. Drug developers are already at work seeking specific inhibitors of this protease (Figure 36.25).

Genetic differences influence individual responses to drugs

Many drugs are not effective in everyone, often because of genetic differences among people. Nonresponding persons may have slight differences in either a drug's target molecule or the proteins taking part in drug transport and metabolism. The goal of the emerging fields of pharmacogenetics and pharmacogenomics is to design drugs that either act more consistently from person to person or are tailored to individual persons with particular genotypes.

Drugs such as metoprolol that target the β_1-adrenergic receptor are popular treatments for hypertension. These drugs are often referred to as "beta-blockers."



Metoprolol

But some people do not respond well. Two variants of the gene coding for the β_1-adrenergic receptor are common in the U.S. population. The most common allele has serine in position 49 and arginine in position 389. In some persons, however, glycine replaces one or the other of these residues. In studies, participants with two copies of the most common allele responded well to metoprolol: their daytime diastolic blood pressure was reduced by 14.7 ± 2.9 mm Hg on average. In contrast, participants with one variant allele showed a smaller reduction in blood pressure, and the drug had no significant effect on participants with two variant alleles (Figure 36.26). These observations suggest the potential utility of genotyping individual patients at these positions. One could then predict whether treatment with metoprolol or other beta-blockers is likely to be effective.

Given the importance of ADME and toxicity properties in determining drug efficacy, it is not surprising that variations in proteins participating in drug transport and metabolism can alter a drug's effectiveness. An important example is the use of thiopurine drugs such as 6-thioguanine,

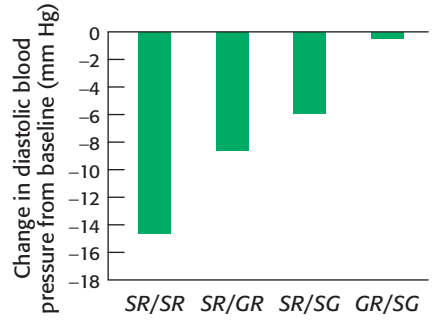
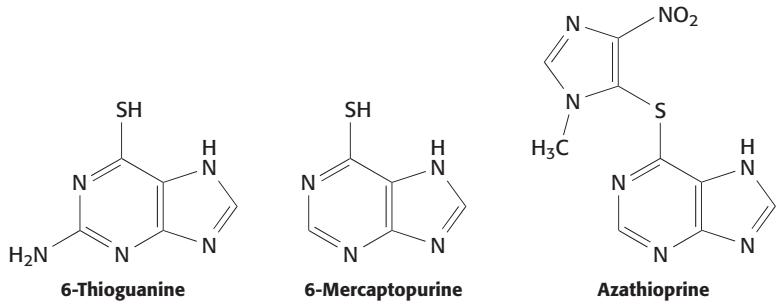
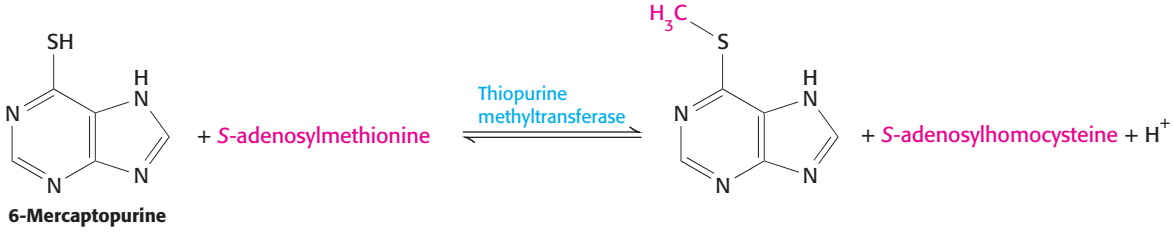


Figure 36.26 Phenotype–genotype correlation. Average changes in diastolic blood pressure on treatment with metoprolol. Persons with two copies of the most common allele ($S_{49}R_{389}$) showed significant decreases in blood pressure. Those with one variant allele (*GR* or *SG*) showed more modest decreases, and those with two variant alleles (*GR/SG*) showed no decrease. [From J. A. Johnson et al., *Clin. Pharmacol. Ther.* 74:44–52, 2003.]

6-mercaptopurine, and azothioprine to treat diseases including leukemia, immune disorders, and inflammatory bowel disease.



A minority of patients who are treated with these drugs show signs of toxicity at doses that are well tolerated by most patients. These differences between patients are due to rare variations in the gene encoding the xenobiotic-metabolizing enzyme thiopurine methyltransferase, which adds a methyl group to sulfur atoms.



The variant enzyme is less stable. Patients with these variant enzymes can build up toxic levels of the drugs if appropriate care is not taken. Thus, genetic variability in an enzyme participating in drug metabolism plays a large role in determining the variation in the tolerance of different persons to particular drug levels. Many other drug-metabolism enzymes and drug-transport proteins have been implicated in controlling individual reactions to specific drugs. The identification of the genetic factors will allow a deeper understanding of why some drugs work well in some persons but poorly in others. In the future, doctors may examine a patient's genes to help plan drug-therapy programs.

36.4 The Development of Drugs Proceeds Through Several Stages

In the United States, the FDA requires that drug candidates be demonstrated to be effective and safe before they may be used in human beings on a large scale. This requirement is particularly true for drug candidates that are to be taken by people who are relatively healthy. More side effects are acceptable for drug candidates intended to treat significantly ill patients such as those with serious forms of cancer, where there are clear, unfavorable consequences for not having an effective treatment.

Clinical trials are time consuming and expensive

Clinical trials test the effectiveness and potential side effects of a candidate drug before it is approved by the FDA for general use. These trials proceed in at least three phases (Figure 36.27). In phase 1, a small number (usually

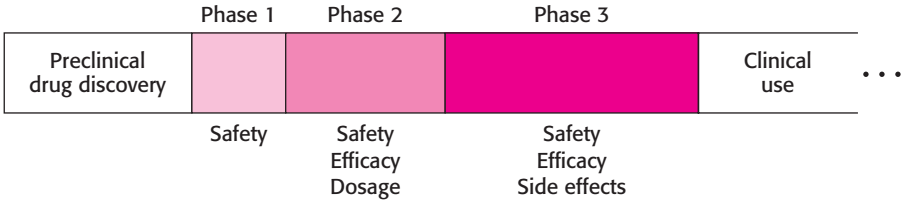


Figure 36.27 Clinical-trial phases. Clinical trials proceed in phases examining safety and efficacy in increasingly large groups.

from 10 to 100) of healthy volunteers take the drug for an initial study of safety. These volunteers are given a range of doses and are monitored for signs of toxicity. The efficacy of the drug candidate is not specifically evaluated.

In phase 2, the efficacy of the drug candidate is tested in a small number of persons who might benefit from the drug. Further data regarding the drug's safety are obtained. Such trials are often controlled and double-blinded. In a controlled study, subjects are divided randomly into two groups. Subjects in the treatment group are given the treatment under investigation. Subjects in the control group are given either a placebo—that is, a treatment such as sugar pills known to not have intrinsic value—or the best standard treatment available, if withholding treatment altogether would be unethical. In a double-blinded study, neither the subjects nor the researchers know which subjects are in the treatment group and which are in the control group. A double-blinded study prevents bias in the course of the trial. When the trial has been completed, the assignments of the subjects into treatment and control groups are unsealed and the results for the two groups are compared. A variety of doses are often investigated in phase 2 trials to determine which doses appear to be free of serious side effects and which doses appear to be effective.

The power of the placebo effect—that is, the tendency to perceive improvement in a subject who believes that he or she is receiving a potentially beneficial treatment—should not be underestimated. In a study of arthroscopic surgical treatment for knee pain, for example, subjects who were led to believe that they had received surgery through the use of videotapes and other means showed the same level of improvement, on average, as subjects who were actually operated on.

In phase 3, similar studies are performed on a larger population. This phase is intended to more firmly establish the efficacy of the drug candidate and to detect side effects that may develop in a small percentage of the subjects who receive treatment. Thousands of subjects may participate in a typical phase 3 study.

Clinical trials can be extremely costly. Hundreds or thousands of patients must be recruited and monitored for the duration of the trial. Many physicians, nurses, clinical pharmacologists, statisticians, and others participate in the design and execution of the trial. Costs can run from tens of millions to hundreds of millions of dollars. Extensive records must be kept, including documentation of any adverse reactions. These data are compiled and submitted to the FDA. The full cost of developing a drug is currently estimated to be from $400 million to $800 million.

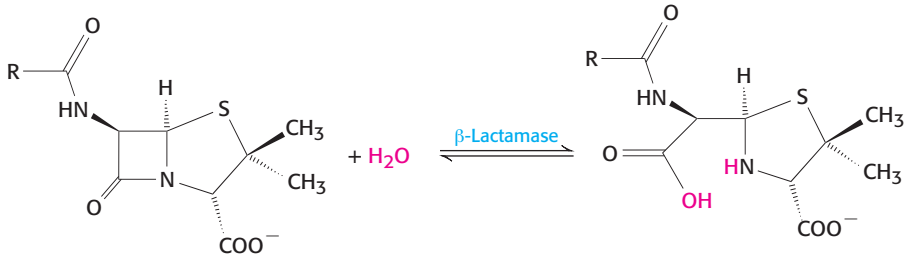
Even after a drug has been approved and is in use, difficulties can arise. As mentioned earlier, rofecoxib (Vioxx), for example, was withdrawn from the market after significant cardiac side effects were detected in additional clinical trials. Such events highlight the necessity for users of any drug to balance beneficial effects against potential risks.

The evolution of drug resistance can limit the utility of drugs for infectious agents and cancer

Many drugs are used for long periods of time without any loss of effectiveness. However, in some cases, particularly for the treatment of cancer or infectious diseases, drug treatments that were initially effective become less effective. In other words, the disease becomes resistant to the drug therapy. Why does this resistance develop? Infectious diseases and cancer have a common feature—namely, that an affected person contains many cells (or viruses) that can mutate and reproduce. These conditions are necessary for evolution to take place. Thus, an individual microorganism or cancer cell may by chance have a genetic variation that makes it more suitable for growth and reproduction in the presence of the drug than is the population of microorganisms or cancer cells at large. These microorganisms or cells are more fit than others in their population, and they will tend to take over the population. As the selective pressure due to the drug is continually applied, the population of microorganisms or cancer cells will tend to become more and more resistant to the presence of the drug. Note that resistance can develop by a number of mechanisms.

The HIV protease inhibitors discussed earlier provide an important example of the evolution of drug resistance. Retroviruses are very well suited to this sort of evolution because reverse transcriptase carries out replication without a proofreading mechanism. In a genome of approximately 9750 bases, each possible single point mutation is estimated to appear in a virus particle more than 1000 times per day in each infected person. Many multiple mutations also occur. Most of these mutations either have no effect or are detrimental to the virus. However, a few of the mutant virus particles encode proteases that are less susceptible to inhibition by the drug. In the presence of an HIV protease inhibitor, these virus particles will tend to replicate more effectively than does the population at large. With the passage of time, the less-susceptible viruses will come to dominate the population and the virus population will become resistant to the drug.

Pathogens may become resistant to antibiotics by completely different mechanisms. Some pathogens contain enzymes that inactivate or degrade specific antibiotics. For example, many organisms are resistant to β-lactams such as penicillin because they contain β-lactamase enzymes. These enzymes hydrolyze the β-lactam ring and render the drugs inactive.



Penicillin

Many of these enzymes are encoded in plasmids, small circular pieces of DNA often carried by bacteria. Many plasmids are readily transferred from one bacterial cell to another, transmitting the capability for antibiotic resistance. Plasmid transfer thus contributes to the spread of antibiotic resistance, a major health-care challenge. On the other hand, plasmids have been harnessed for use in recombinant DNA methods (Section 5.2).

Drug resistance commonly emerges in the course of cancer treatment. Cancer cells are characterized by their ability to grow rapidly without the

constraints that apply to normal cells. Many drugs used for cancer chemotherapy inhibit processes that are necessary for this rapid cell growth. However, individual cancer cells may accumulate genetic changes that mitigate the effects of such drugs. These altered cancer cells will tend to grow more rapidly than others and will become dominant within the cancer-cell population. This ability of cancer cells to mutate quickly has been a challenge to one of the major breakthroughs in cancer treatment: the development of inhibitors for proteins specific to cancer cells present in certain leukemias (Section 14.5). For example, tumors become undetectable in patients treated with imatinib mesylate, which is directed against the Bcr-Abl protein kinase. Unfortunately, the tumors of many of the patients treated with imatinib mesylate recur after a period of years. In many of these cases, mutations have altered the Bcr-Abl protein so that it is no longer inhibited by the concentrations of imatinib mesylate used in therapy.

Cancer patients often take multiple drugs concurrently in the course of chemotherapy and, in many cases, cancer cells become simultaneously resistant to many or all of them. This multiple-drug resistance can be due to the proliferation of cancer cells that overexpress a number of ABC transporter proteins that pump drugs out of the cell (Section 13.2). Thus, cancer cells can evolve drug resistance by overexpressing normal human proteins or by modifying proteins responsible for the cancer phenotype.

Summary

36.1 The Development of Drugs Presents Huge Challenges

Most drugs act by binding to enzymes or receptors and modulating their activities. To be effective, drugs must bind to these targets with high affinity and specificity. However, even most compounds with the desired affinity and specificity do not make suitable drugs. Most compounds are poorly absorbed or rapidly excreted from the body or they are modified by metabolic pathways that target foreign compounds. Consequently, when taken orally, these compounds do not reach their targets at appropriate concentrations for a sufficient period of time. A drug's properties related to its absorption, distribution, metabolism, and excretion are called ADME properties. Oral bioavailability is a measure of a drug's ability to be absorbed; it is the ratio of the peak concentration of a compound given orally to the peak concentration of the same dose directly injected. The structure of a compound can affect its bioavailability in complicated ways, but generalizations called Lipinski's rules provide useful guidelines. Drug-metabolism pathways include oxidation by cytochrome P450 enzymes (phase I metabolism) and conjugation to glutathione, glucuronic acid, and sulfate (phase II metabolism). A compound may also not be a useful drug because it is toxic, either because it modulates the target molecule too effectively or because it also binds to proteins other than the target. The liver and kidneys play central roles in drug metabolism and excretion.

36.2 Drug Candidates Can Be Discovered by Serendipity, Screening, or Design

Many drugs have been discovered by serendipity—that is, by chance observation. The antibiotic penicillin is produced by a mold that accidentally contaminated a culture dish, killing nearby bacteria. Drugs such as chlorpromazine and sildenafil were discovered to have beneficial effects on human physiology that were completely different from those expected. The cholesterol-lowering statin drugs were developed after large collections of compounds were screened for potentially

interesting activities. Combinatorial chemistry methods have been developed to generate large collections of chemically related yet diverse compounds for screening. In some cases, the three-dimensional structure of a drug target is available and can be used to aid the design of potent and specific inhibitors. Examples of drugs designed in this manner are the HIV protease inhibitors indinavir and cyclooxygenase 2 inhibitors such as celecoxib.

36.3 Analyses of Genomes Hold Great Promise for Drug Discovery

The human genome encodes approximately 25,000 proteins, and many more if derivatives due to alternative mRNA splicing and post-translational modification are included. The genome sequences can be examined for potential drug targets. Large families of proteins known to participate in key physiological processes such as the protein kinases and 7TM receptors have each yielded several targets for which drugs have been developed. The genomes of model organisms also are useful for drug-development studies. Strains of mice with particular genes disrupted have been useful in validating certain drug targets. The genomes of bacteria, viruses, and parasites encode many potential drug targets that can be exploited owing to their important functions and their differences from human proteins, minimizing the potential for side effects. Genetic differences between individuals can be examined and correlated with differences in responses to drugs, potentially aiding both clinical treatments and drug development.

36.4 The Development of Drugs Proceeds Through Several Stages

Before compounds can be given to human beings as drugs, they must be extensively tested for safety and efficacy. Clinical trials are performed in stages: first testing safety, then safety and efficacy in a small population, and finally safety and efficacy in a larger population to detect rarer adverse effects. Largely due to the expenses associated with clinical trials, the cost of developing a new drug has been estimated to be as much as $800 million. Even when a drug has been approved for use, complications can arise. With regard to infectious diseases and cancer, patients often develop resistance to a drug after it has been administered for a period of time because variants of the disease agent that are less susceptible to the drug arise and replicate, even when the drug is present.

Key Terms

ligand (p. 1030)
dissociation constant (K_d) (p. 1030)
side effect (p. 1031)
inhibition constant (K_i) (p. 1031)
Cheng–Prusoff equation (p. 1031)
ADME (p. 1031)
oral bioavailability (p. 1032)
Lipinski's rules (p. 1032)
compartment (p. 1033)
blood–brain barrier (p. 1033)
xenobiotic compounds (p. 1034)
drug metabolism (p. 1034)
oxidation (p. 1034)
conjugation (p. 1034)
phase I transformation (p. 1035)
phase II transformation (p. 1035)
first-pass metabolism (p. 1035)
glomerulus (p. 1035)
enterohepatic cycling (p. 1036)
therapeutic index (p. 1036)
atheroma (p. 1040)
myopathy (p. 1040)
high-throughput screening (p. 1041)
combinatorial chemistry (p. 1041)
split-pool synthesis (p. 1041)
structure-based drug design (p. 1043)

Problems

1. *Routes to discovery.* For each of the following drugs, indicate whether the physiological effects of the drug were known before or after the target was identified.

(a) Penicillin
(b) Sildenafil (Viagra)
(c) Rofecoxib (Vioxx)
(d) Atorvastatin (Lipitor)
(e) Aspirin
(f) Indinavir (Crixivan)

2. *Lipinski's rules.* Which of the following compounds satisfy all of Lipinski's rules? [Log(*P*) values are given in parentheses.]

(a) Atenolol (0.23)
(b) Sildenafil (3.18)
(c) Indinavir (2.78)

3. *Calculating log tables.* Considerable effort has been expended to develop computer programs that can estimate log(*P*) values entirely on the basis of chemical structure. Why would such programs be useful?

4. *An ounce of prevention.* Legislation has been proposed that would require that *N*-acetylcysteine be added to acetaminophen tablets. Speculate about the role of this additive.

5. *Clinical-trial design.* Distinguish between phase 1 and phase 2 clinical trials in regard to number of persons enrolled, the state of health of the subjects, and the goals of the study.

6. *Drug interactions.* As noted in this chapter, coumadin can be a very dangerous drug because too much can cause uncontrolled bleeding. Persons taking coumadin must be careful about taking other drugs, particularly those that bind to albumin. Propose a mechanism for this drug–drug interaction.

7. *A bad combination.* Explain why drugs that inhibit P450 enzymes may be particularly dangerous when used in combination with other medications.

8. *Mechanistically speaking.* Name one advantage of a noncompetitive inhibitor as a potential drug compared with a competitive inhibitor.

9. *A helping hand.* You have developed a drug that is capable of inhibiting the ABC transporter MDR. Suggest a possible application for this drug in cancer chemotherapy.

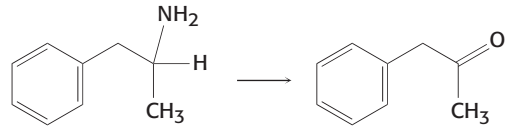
10. *Find the target.* Trypanosomes are unicellular parasites that cause sleeping sickness. During one stage of their life cycle, these organisms live in the bloodstream and derive all of their energy from glycolysis, which takes place in a specialized organelle called a glycosome inside the parasite. Propose potential targets for treating sleeping sickness. What are some potential difficulties with your approach?

11. *Knowledge is power.* How might genomic information be helpful for the effective use of imatinib mesylate (Gleevec) in cancer chemotherapy?

12. *Multiple targets, same goal.* Sildenafil induces its physiological effects by increasing the intracellular concentrations of cGMP, leading to muscle relaxation. On the basis of the scheme shown in Figure 36.17, identify another approach for increasing cGMP levels with a small molecule.

Mechanism Problem

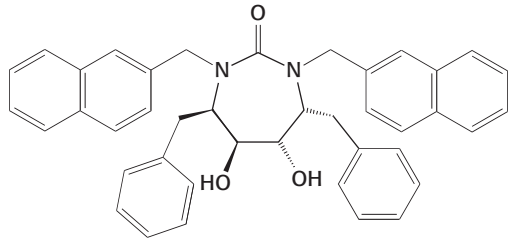
13. *Variations on a theme.* The metabolism of amphetamine by cytochrome P450 enzymes results in the conversion shown here. Propose a mechanism and indicate any additional products.



Amphetamine

Data Interpretation Problem

14. *HIV protease inhibitor design.* Compound A is one of a series that were designed to be potent inhibitors of HIV protease.



Compound A

Compound A was tested by using two assays: (1) direct inhibition of HIV protease in vitro and (2) inhibition of viral RNA production in HIV-infected cells, a measure of viral replication. The results of these assays are shown below. The HIV protease activity is measured with a substrate peptide present at a concentration equal to its K_M value.

Compound A (nM)	HIV protease activity (arbitrary units)
0	11.2
0.2	9.9
0.4	7.4
0.6	5.6
0.8	4.8
1	4.0
2	2.2
10	0.9
100	0.2

Compound A (nM)	Viral RNA production (arbitrary units)
0	760
1.0	740
2.0	380
3.0	280
4.0	180
5.0	100
10	30
50	20

Estimate the values for the K_I of compound A in the protease-activity assay and for its IC_{50} in the viral-RNA-production assay.

Treating rats with the relatively high oral dose of 20 mg kg^{-1} results in a maximum concentration of compound A of 0.4 μM. On the basis of this value, do you expect compound A to be effective in preventing HIV replication when taken orally?