

Figure 4.25 Adjacent antiparallel β strands are often joined by hairpin loops. This diagram shows the frequency with which loops of different lengths are found linking two β strands. Such loops are usually short. (Adapted from B.L. Sibanda and J.M. Thornton, *Nature* 316: 170–174, 1985. With permission from Macmillan Publishers Ltd.)

of the turn. The common occurrence of these two specific types of turn structures is a consequence of the fact that the polypeptide backbone is not arbitrarily flexible. The conformation of each residue in the turn has to be in one of the allowed regions of the Ramachandran diagram while allowing the direction of the chain to bend back on itself.

4.13 α helices and β strands are often amphipathic

The most common location for an α helix in a protein structure is along the outside of the protein, with one side of the helix facing the solution and the other side facing the hydrophobic interior of the protein. Therefore, with 3.6 residues per turn,

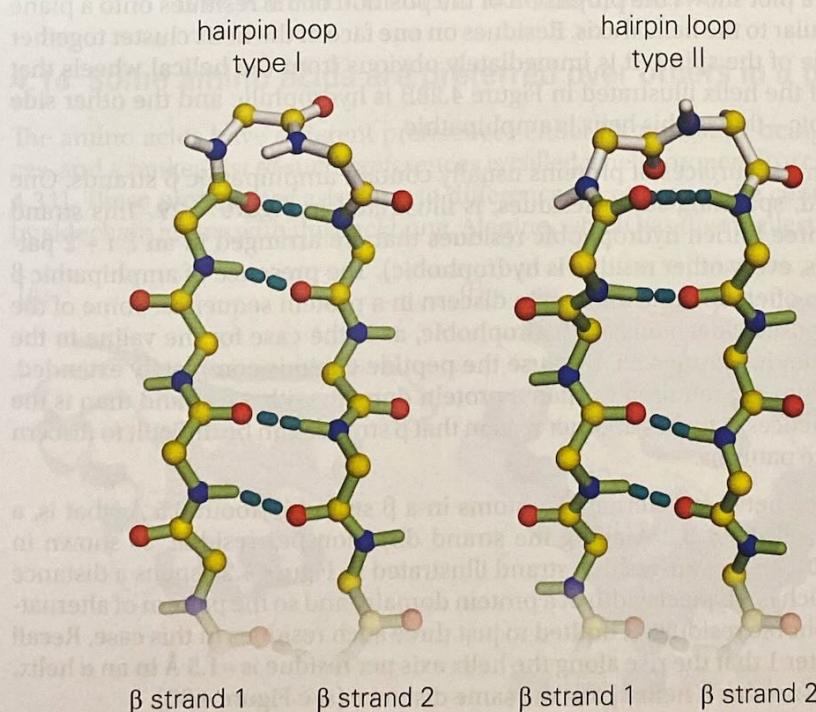
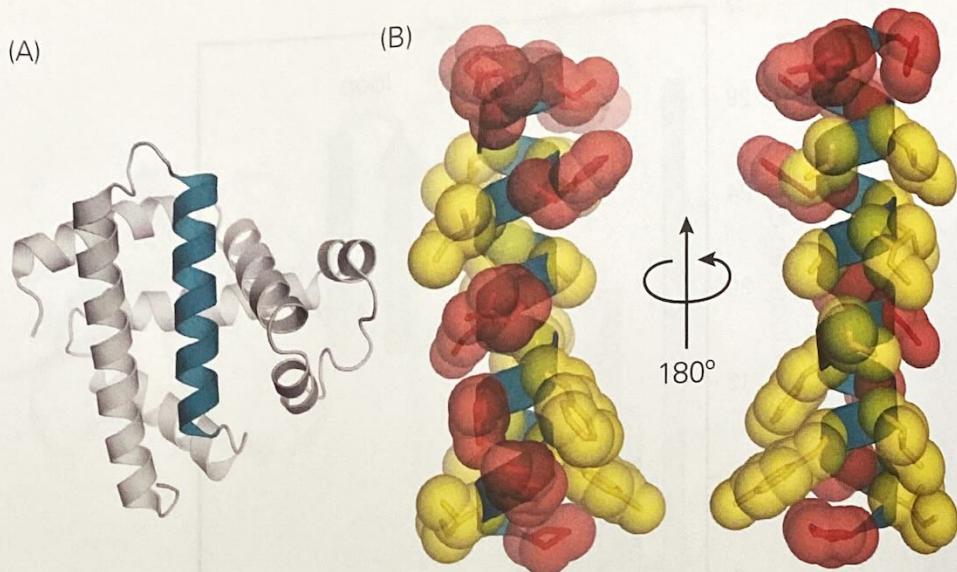


Figure 4.26 Two kinds of β hairpin loops. The diagram shows the two most frequently occurring kinds of two-residue hairpin loops: the type I turn (left) and the type II turn (right). Bonds within the hairpin loop are shown in white. (Adapted from C. Brändén and J. Tooze, *Introduction to Protein Structure*, 2nd ed. New York: Garland Science, 1999.)

Figure 4.27 An amphipathic α helix.

The structure of myoglobin is shown in (A), with one α helix colored blue. (B) This α helix is shown in an expanded view. The view on the left is in the same orientation as in (A). The view on the right is rotated by 180° with respect to the first one, revealing the face of the helix that packs against the rest of the protein. Hydrophobic sidechains are colored yellow and polar and charged sidechains are colored red.



there is a tendency for sidechains to change from hydrophobic to hydrophilic with a periodicity of three to four residues. α helices that have one hydrophobic face and one hydrophilic face are known as **amphipathic helices** (Figure 4.27).

Not all α helices in globular proteins are amphipathic, because residues that face the solution can sometimes be hydrophobic and, furthermore, α helices can be either completely buried within the protein or completely exposed. Figure 4.28 shows the structures of α helices in three different proteins, with their amino acid sequences. One of the helices is totally buried, and most of the residues in this helix are hydrophobic (Figure 4.28A). A partially exposed α helix is shown in Figure 4.28B. The hydrophobic residues in this helix are arranged in an $i, i + 3$, or $i, i + 4$ pattern, where i refers to the position of one residue in the sequence. Figure 4.28C shows the structure of a completely exposed α helix, and most of the residues in this helix are polar. Completely buried or completely exposed α helices are relatively rare in proteins.

Amphipathic α helices and β sheets

Secondary structural elements that have distinctive faces, one hydrophobic and one hydrophilic, are referred to as amphipathic. The hydrophobic faces of amphipathic sheets and helices can pack against each other to form a hydrophobic core, leaving the hydrophilic faces to interact favorably with water. This is a central principle in the architecture of most globular proteins.

A convenient way to illustrate the amino acid sequences in helices is the helical wheel or spiral. Because one turn in an α helix spans 3.6 residues, each residue can be plotted every $360/3.6 = 100^\circ$ around a circle or a spiral, as shown in Figure 4.28. Such a plot shows the projection of the position of the residues onto a plane perpendicular to the helical axis. Residues on one face of the helix cluster together on one side of the spiral. It is immediately obvious from the helical wheels that one side of the helix illustrated in Figure 4.28B is hydrophilic and the other side hydrophobic—that is, this helix is amphipathic.

β sheets on the surfaces of proteins usually contain amphipathic β strands. One such strand, spanning seven residues, is illustrated in Figure 4.29. This strand contains three buried hydrophobic residues that are arranged in an $i, i + 2$ pattern (that is, every other residue is hydrophobic). The presence of amphipathic β strands can often be quite difficult to discern in a protein sequence. Some of the surface exposed sidechains are hydrophobic, as is the case for the valine in the strand shown in Figure 4.29. Because the peptide chain is completely extended, fewer residues are required to span a protein domain with a β strand than is the case for α helices, which is another reason that β strands can be difficult to discern in sequence patterns.

The distance between alternate C_α atoms in a β strand is about 6.5 \AA (that is, a translation of about 3.2 \AA along the strand direction per residue, as shown in Figure 4.30). The seven-residue strand illustrated in Figure 4.29 spans a distance of 20 \AA , which is a typical width of a protein domain, and so the pattern of alternating hydrophobic residues is limited to just three such residues in this case. Recall from Chapter 1 that the rise along the helix axis per residue is $\sim 1.5 \text{ \AA}$ in an α helix, and so a 14-residue α helix spans the same distance (see Figure 4.30).

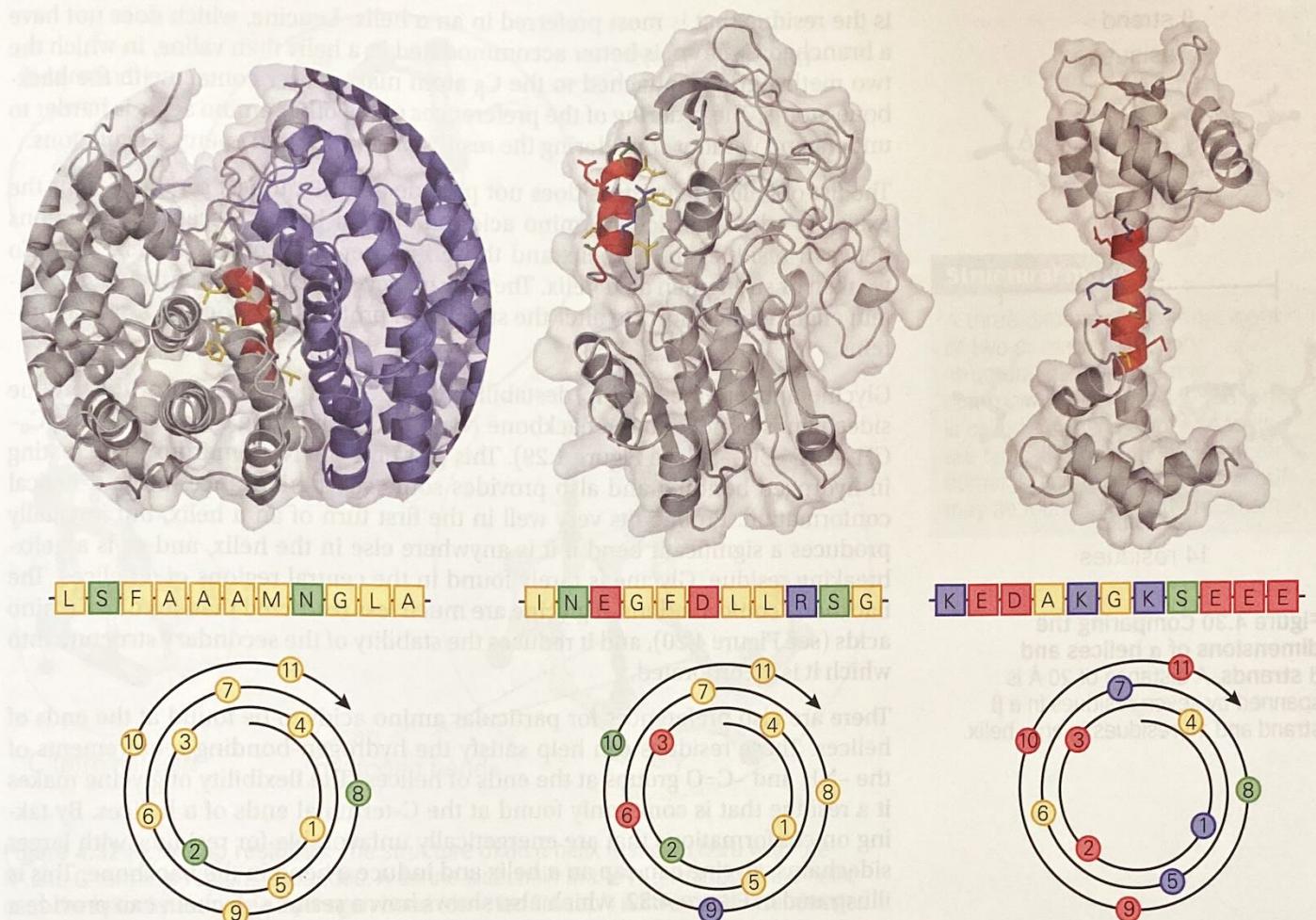


Figure 4.28 Examples of α helices in three different proteins. (A) A completely buried α helix (red) from the enzyme citrate synthase. (B) A partially exposed helix from the enzyme alcohol dehydrogenase. (C) A completely exposed helix from troponin-C. The amino acid sequences of the helices are shown below the structures, with charged residues in red and blue, polar residues in green, and hydrophobic residues in yellow. The sequences of the α helices are also shown on helical wheels or spirals, in which amino acid residues are plotted every 100° around the spiral. (Adapted from C. Brändén and J. Tooze, Introduction to Protein Structure, 2nd ed. New York: Garland Science, 1999; PDB codes: A, 1CTS; B, 1A71; and C, 5TNC.)

4.14 Some amino acids are preferred over others in α helices

The amino acids have different preferences either for or against being in α helices, and a ranked list of such preferences is called a helix propensity scale (Figure 4.31). These preferences arise due to differences in energy of the contacts made by sidechain atoms with the backbone. Alanine, which has the smallest sidechain,

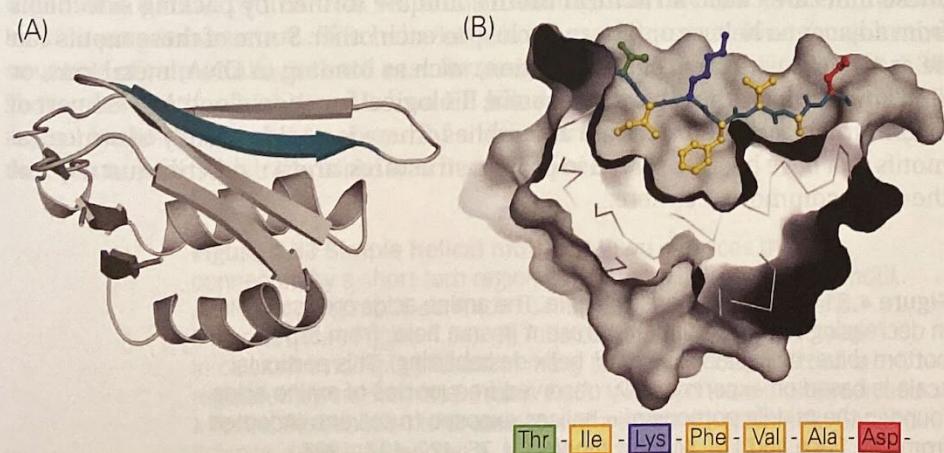


Figure 4.29 An amphipathic β strand. (A) A protein structure is shown in which one β strand is colored teal. (B) A cross section through the protein, showing packing of the same β strand against the rest of the protein. The amino acid sequence of the strand is shown below. Notice that the buried hydrophobic residues (isoleucine, phenylalanine, and alanine) occur in an $i, i + 2$ pattern. (PDB code: 1PLQ.)

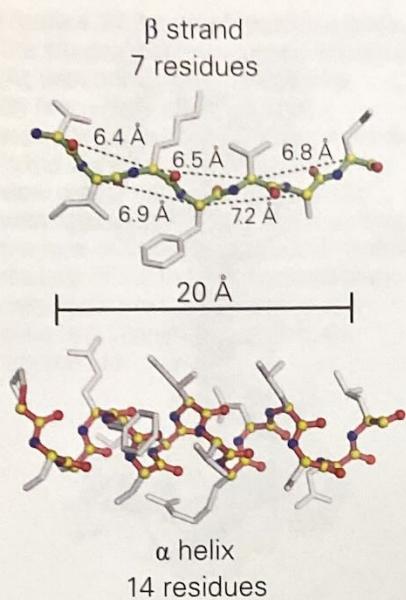


Figure 4.30 Comparing the dimensions of α helices and β strands. A distance of 20 Å is spanned by seven residues in a β strand and 14 residues in an α helix.

is the residue that is most preferred in an α helix. Leucine, which does not have a branched C_β atom, is better accommodated in a helix than valine, in which the two methyl groups attached to the C_β atom make closer contact with the backbone atoms. The ordering of the preferences of the other amino acids is harder to understand without considering the results of complicated energy calculations.

The list of helix preferences does not provide a hard and fast set of rules for the extent to which particular amino acids stabilize α helices because interactions between residues in the helix and those in other parts of the protein can also weaken or strengthen an α helix. They do, however, provide a useful guide in protein engineering efforts to alter the stability of proteins or to design artificial proteins.

Glycine and proline tend to destabilize α helices. The last atom of the proline sidechain is bonded to the backbone N atom, thus forming a ring structure, C_α -CH₂-CH₂-CH₂-N (see Figure 1.29). This prevents the N atom from participating in hydrogen bonding and also provides some steric hindrance to the α -helical conformation. Proline fits very well in the first turn of an α helix, but it usually produces a significant bend if it is anywhere else in the helix, and so is a helix-breaking residue. Glycine is rarely found in the central regions of α helices. The backbone conformations of glycine are much less restricted than for other amino acids (see Figure 4.20), and it reduces the stability of the secondary structure into which it is incorporated.

There are also preferences for particular amino acids to be found at the ends of helices. These residues can help satisfy the hydrogen-bonding requirements of the -NH and -C=O groups at the ends of helices. The flexibility of glycine makes it a residue that is commonly found at the C-terminal ends of α helices. By taking on conformations that are energetically unfavorable for residues with larger sidechains, glycine can cap an α helix and induce a bend in the backbone. This is illustrated in Figure 4.32, which also shows how a serine sidechain can provide a cap for the N-terminal end of a helix.

The propensity of different amino acids to be found in β sheets is less clearly delineated than for α helices, because local interactions between neighboring sidechains play a stronger role in the stability of β -sheet structure.

C. STRUCTURAL MOTIFS AND DOMAINS IN SOLUBLE PROTEINS

4.15 Secondary structure elements are connected to form simple motifs

Simple combinations of a few secondary structure elements with a specific geometric arrangement have been found to occur frequently in protein structures. These units are called **structural motifs**, and are formed by packing sidechains from adjacent α helices or β strands close to each other. Some of these motifs can be associated with a particular function, such as binding to DNA, metal ions, or small molecules; others have no specific biological function alone, but are part of larger structural and functional assemblies. There is a wide variety of structural motifs that have been observed in protein structures, and we describe just a few of the more common ones here.

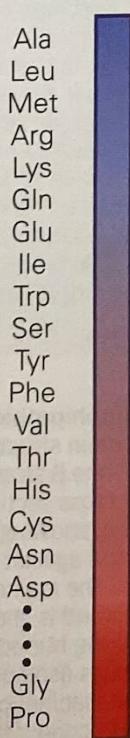


Figure 4.31 A helix propensity scale. The amino acids are listed in decreasing likelihood of being present in an α helix, from top to bottom (blue: helix stabilizing; red: helix destabilizing). This particular scale is based on experimentally observed frequencies of amino acids found in the middle portions of α helices exposed to solvent. (Adapted from C.N. Pace and J.M. Scholtz, *Biophys. J.* 75: 422–427, 1998.)

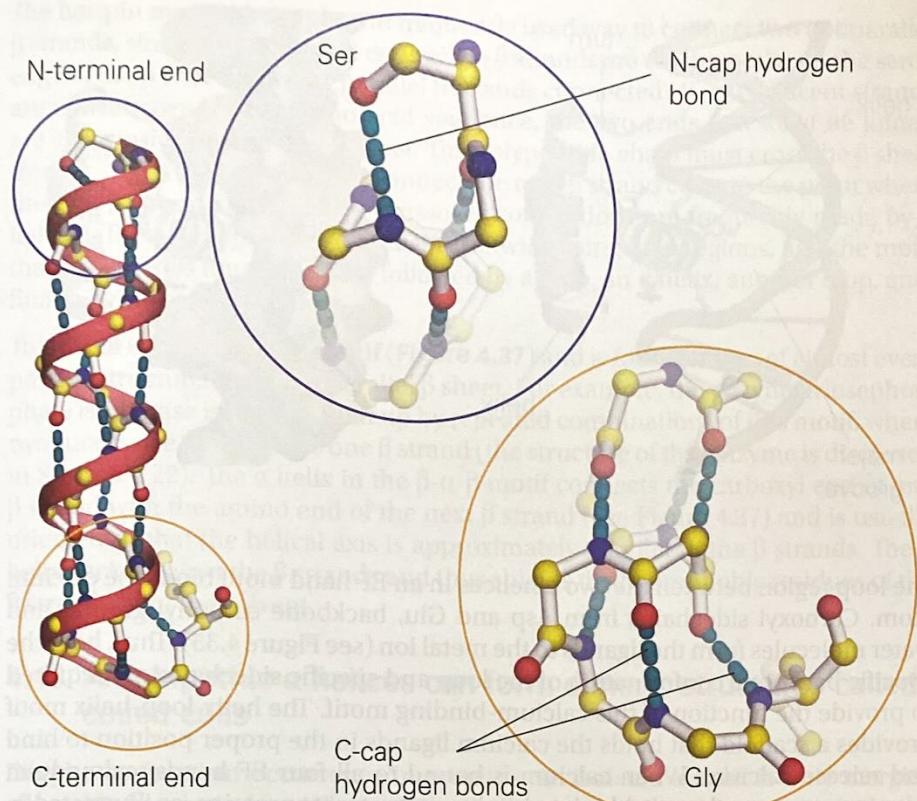


Figure 4.32 Helix cap residues. The structure of an α helix is shown here with the N- and C-terminal regions expanded. A serine sidechain at the N-terminus of the helix caps the helix by forming a hydrogen bond (denoted “N-cap”) with a backbone nitrogen atom. A glycine residue at the other end of the helix causes a bend in the backbone, which allows the backbone nitrogen atoms of the glycine and the next residue to form hydrogen bonds (denoted “C-cap”) with the backbone carbonyl groups of preceding residues. (PDB code: 1TIM.)

A particularly simple motif consists of two α helices joined by a short turn region, called a **helix-turn-helix motif** (Figure 4.33A). Helix-turn-helix motifs are commonly found in proteins that recognize specific sequences in DNA. One of the helices is inserted into the major groove of DNA, where sidechains emanating from the helix make sequence-specific contacts with the bases in the major groove (Figure 4.34).

A similar motif is the **helix-loop-helix motif** (see Figure 4.33B), in which the connection between the two helices is longer, as in the protein calmodulin. Calmodulin is a small protein that responds to changes in calcium levels in the cell by changing its structure. The helix-loop-helix motif appears four times in the structure of calmodulin, in each case forming a calcium-binding site. Figure 4.35 shows this motif, which is called an **EF hand** because the helices of this motif resemble the thumb and forefinger of a right hand, and were labeled E and F in the structure for which the motif was first described.

Figure 4.33 Simple helical motifs. (A) Two α helices that are connected by a short turn region constitute a helix-turn-helix motif. A DNA binding motif, common to many transcription factors, is illustrated (see Figure 4.34). (B) A helix-loop-helix motif involved in calcium binding is present in many proteins whose function is regulated by calcium (see Figure 4.35). (Adapted from C. Brändén and J. Tooze, *Introduction to Protein Structure*, 2nd ed. New York: Garland Science, 1999; PDB codes: A, 1LL1 and B, 1CM1.)

Structural motif

A three-dimensional arrangement of two or more secondary structural elements that is commonly found in many proteins is called a structural motif. Motifs are typically components of larger domains, and more than one motif may be found in a protein domain.

Figure 4.36 The Greek key motif.
(A) An antiparallel β -sheet with the strands arranged in a Greek key motif.
(B) A topology diagram for the Greek key motif. (PDB code: 4J3NC)

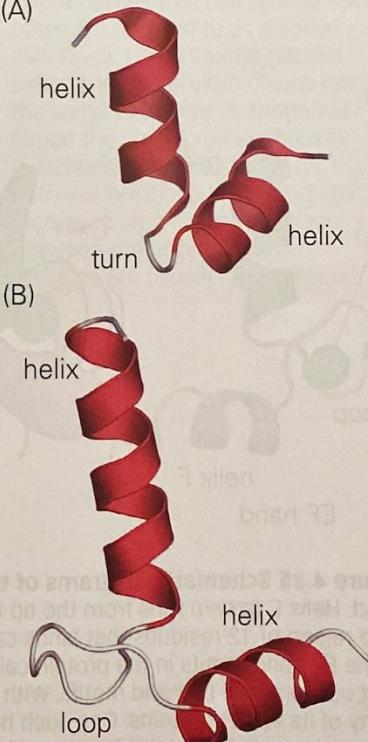
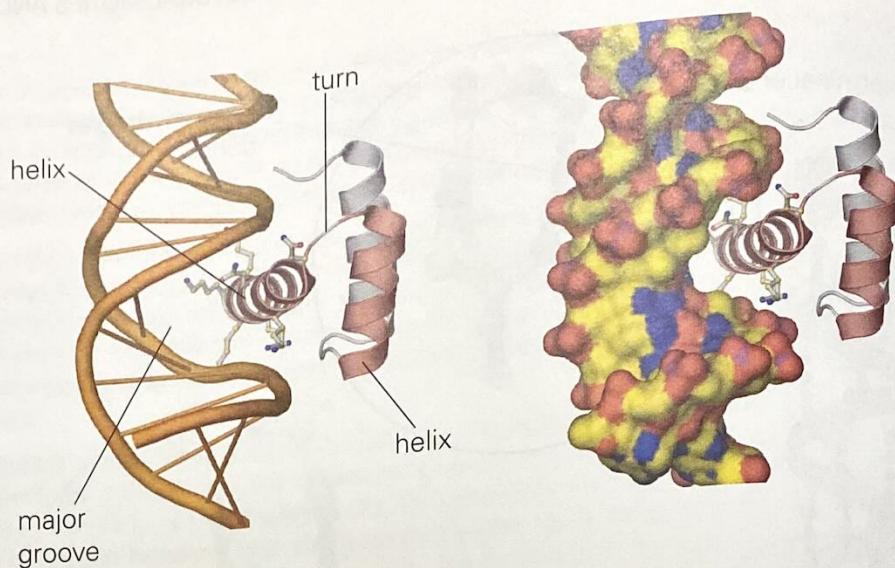


Figure 4.34 Interaction between DNA and the helix-turn-helix motif of a DNA binding protein. (PDB code: 1DU0.)



The loop region between the two α helices in an EF hand motif binds the calcium atom. Carboxyl sidechains from Asp and Glu, backbone carbonyl groups, and water molecules form the ligands to the metal ion (see Figure 4.35). Thus, both the specific backbone conformation of the loop and specific sidechains are required to provide the function of this calcium-binding motif. The helix-loop-helix motif provides a scaffold that holds the calcium ligands in the proper position to bind and release calcium. When calcium is bound to all four EF hands, calmodulin adopts a structure that enables it to bind to many target proteins, as illustrated in Figure 4.35. The ability to bind to these proteins is usually lost when Ca^{2+} dissociates from calmodulin.

Antiparallel β strands are connected together by turns or hairpins (see Figure 4.26). For example, four adjacent antiparallel β strands are frequently arranged in a pattern similar to the repeating unit of an ornamental pattern, or fret, used in ancient Greece, which is now called a Greek key. In proteins, the motif is therefore called a **Greek key motif**, and Figure 4.36 shows an example of such a motif. The Greek key motif is not associated with any specialized function, but it occurs frequently in protein structures.

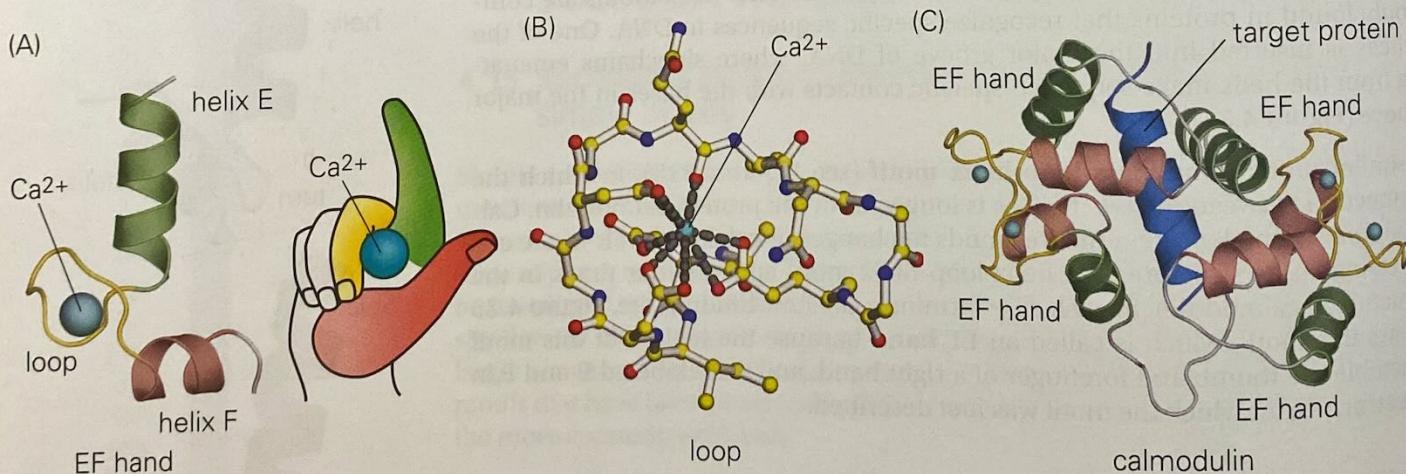


Figure 4.35 Schematic diagrams of the calcium-binding EF hand motif. (A) The calcium-binding motif is symbolized by a right hand. Helix E (green) runs from the tip to the base of the forefinger (also green). The flexed middle finger (yellow) corresponds to the loop region of 12 residues that binds calcium (blue). Helix F (red) runs to the end of the thumb. (B) The calcium ion is bound to one of the EF hand motifs in the protein calmodulin through several oxygen atoms (red) of the protein. (C) The structure of calmodulin is built up from four EF hand motifs. With four Ca^{2+} ions bound to it, calmodulin adopts a structure that enables it to bind to helices in many of its target proteins. One such helix is shown in blue. (Adapted from C. Brändén and J. Tooze, *Introduction to Protein Structure*, 2nd ed. New York: Garland Science, 1999; PDB code: 1CM1.)

The hairpin motif is a simple and frequently used way to connect two antiparallel β strands, since the connected ends of the β strands are close together at the same edge of the β sheet. How are parallel β strands connected? If two adjacent strands are consecutive in the amino acid sequence, the two ends that must be joined are at opposite edges of the β sheet. The polypeptide chain must cross the β sheet from one edge to the other and connect the next β strand close to the point where the first β strand started. Such crossover connections are frequently made by α helices. The polypeptide chain must turn twice using loop regions, and the motif that is formed is thus a β strand followed by a loop, an α helix, another loop, and, finally, the second β strand.

This motif is called a β - α - β motif (Figure 4.37) and is found as part of almost every protein structure that has a parallel β sheet. For example, the enzyme triosephosphate isomerase is entirely built up by repeated combinations of this motif, where two successive motifs share one β strand (the structure of this enzyme is discussed in Section 4.22). The α helix in the β - α - β motif connects the carboxyl end of one β strand with the amino end of the next β strand (see Figure 4.37) and is usually oriented so that the helical axis is approximately parallel to the β strands. The α helix packs against the β strands and thus shields the hydrophobic residues of the β strands from the solvent.

4.16 Amphipathic α helices can form dimeric structures called coiled coils

Despite its frequent occurrence in proteins, an isolated α helix is only marginally stable in solution (see Figure 4.9). α helices are stabilized in proteins by being packed together through hydrophobic sidechains. The simplest way to achieve such stabilization is to pack two α helices together. As early as 1953, Francis Crick reasoned that the sidechain interactions are optimized if the two α helices are not straight rods but are wound around each other in a supercoil, a so-called **coiled-coil** arrangement (Figure 4.38). Coiled coils are the basis for the structures of many fibrous proteins, and coiled coils in fibers can extend over many hundreds of amino acid residues to produce long, flexible structures that contribute to the strength and flexibility of the fibers (see Figure 4.38A). Much shorter coiled coils are used in a variety of proteins to promote formation of homo- and heterodimers (see Figure 4.38B).

The coiled-coil arrangement of α helices is called a **supercoil**, or **superhelix**. These terms reflect the fact that the α helix is itself a coil, and the two α helices coil around each other (see Figure 4.38). Coiled coils can either be parallel, as is the

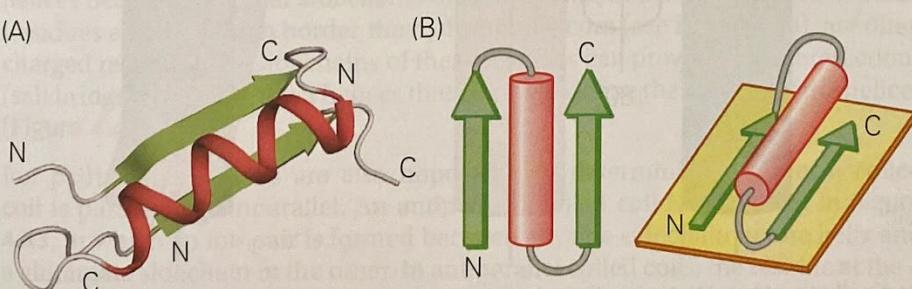


Figure 4.37 β - α - β motifs. (A) Two adjacent parallel β strands are usually connected by an α helix from the C-terminus of strand 1 to the N-terminus of strand 2. Most protein structures that contain parallel β sheets are built up from combinations of such β - α - β motifs. (B) Topological diagrams of the β - α - β motif. (Adapted from J.S. Richardson, *Adv. Protein Chem.* 34: 167–339, 1981.)

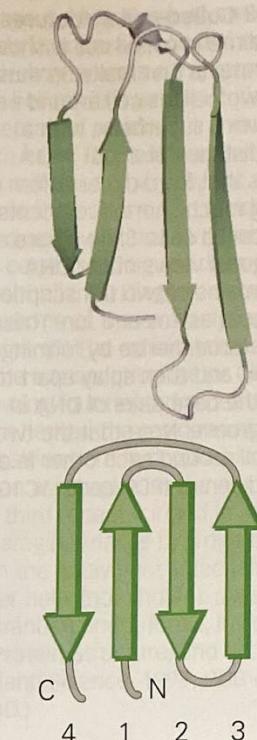


Figure 4.36 The Greek key motif.

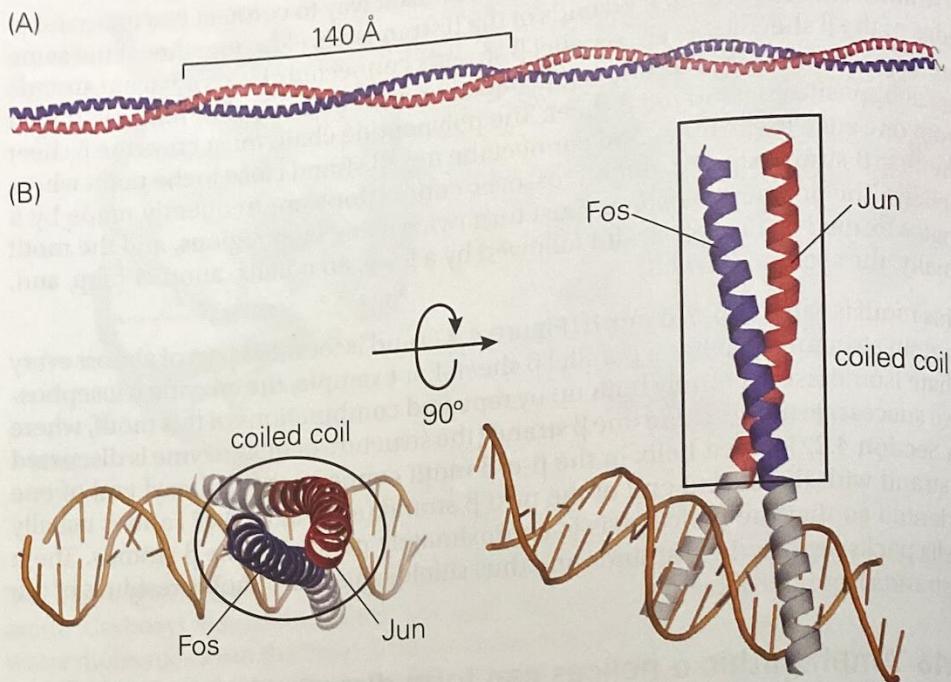
(A) An antiparallel β sheet with the strands arranged in a Greek key motif. (B) A topology diagram for the Greek key motif. (PDB code: A, 1SNC.)

Coiled coil

Two α helices that coil around each other are referred to as a coiled coil. These can either be parallel (when the two protein chains run in the same direction) or antiparallel (when the chains run in opposite directions). Residues at the i , $i + 3$, and $i + 7$ positions in a coiled coil face each other at the interface between the helices. These residues are usually hydrophobic.

Figure 4.38 Coiled-coil structures.

(A) A long dimeric coiled coil is shown, found in proteins that make up muscle fiber. The two helices coil around each other to form a superhelix, with a repeating distance of about 140 Å. (B) Proteins that form dimers often do so by using much shorter segments that form coiled coils. Shown here are two orthogonal views of the DNA-binding domains of two transcription factors known as Fos and Jun. These proteins heterodimerize by forming a coiled coil, and then splay apart to recognize the base pairs of DNA in the major groove. Note that the two proteins coil around each other in a left-handed sense. (PDB code: 1C1G and 1A02.)



case for the structure shown in Figure 4.39, with the chains running in the same direction, or antiparallel (see below). Parallel coiled coils are more common than antiparallel ones.

Supercoil

When the axis of a helix is coiled rather than straight, the resulting structure is called a supercoil or a superhelix.

For a coiled-coil structure to be stable, hydrophobic residues have to be brought into register at the interface between the helices. In order to achieve this, a coiled coil must have a left-handed superhelical structure. To see why this is so, consider the straight α helix illustrated in Figure 4.39A. If one residue is aligned with a line parallel to the helix axis, the residue two turns up from it will be to the *left* of this line. This is because the position exactly two turns up from a residue at position i will be $3.6 \times 2 = 7.2$ residues along the helix backbone—that is, ahead of the position of the residue at the $i + 7$ position. If two α helices were to pack together, as shown in Figure 4.39B and C, the hydrophobic residues would not be in register unless the helices were supercoiled in a left-handed sense.

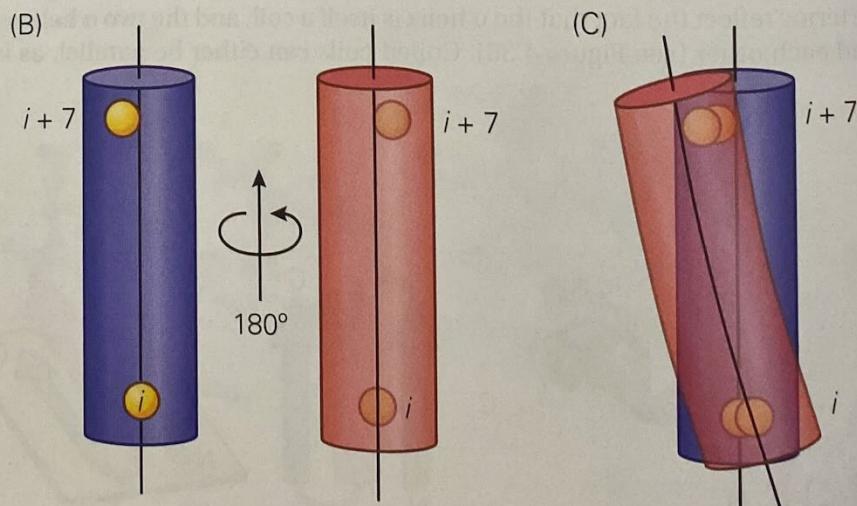
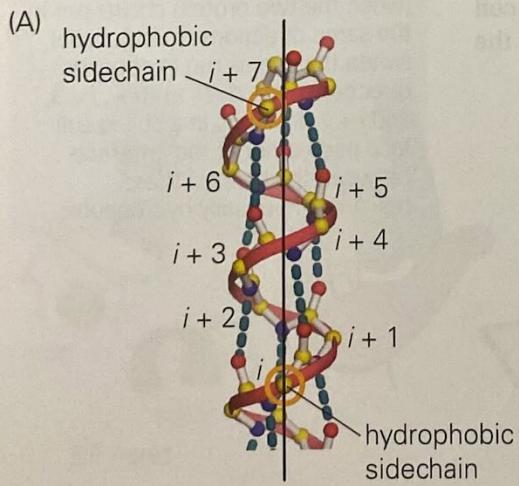
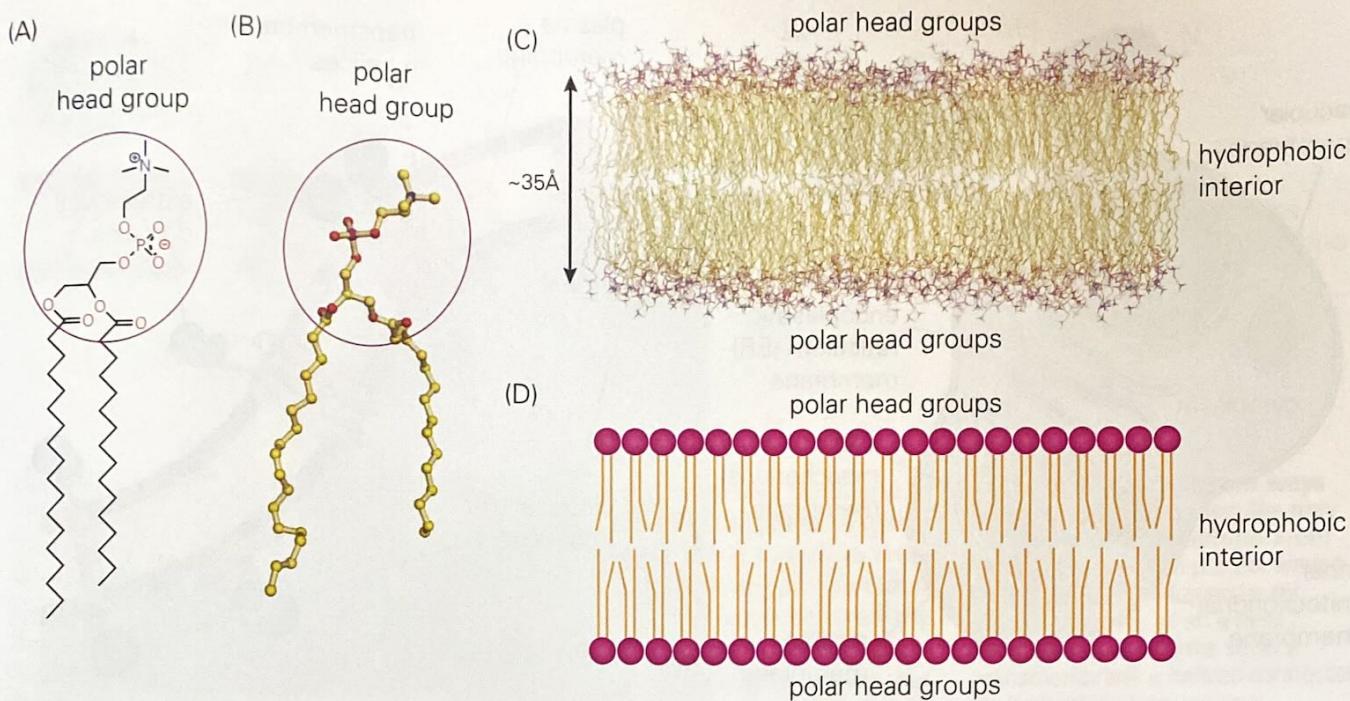


Figure 4.39 Schematic diagram of the coiled-coil structure. (A) An α helix is depicted with the C_α atoms of the residues at the i and the $i + 7$ positions circled. A vertical line is drawn parallel to the helix axis and running through the residue at the i^{th} position. Note that the residue at the $i + 7$ position is to the *left* of this line. (B) Schematic representation of two α helices, with one rotated by 180° with respect to the other. Hydrophobic residues at i and $i + 7$ are shown in yellow. (C) Packing of the two helices depicted in (C). In order for the hydrophobic residues at i and $i + 7$ to be aligned, the red helix is shown curling to the left. In real coiled coils, such as the ones shown in Figure 4.38, both helices are curled in a left-handed sense and so the extent of curling for each helix is less than that shown here for the red helix.



oriented in such a way that the carboxyl edges of both β sheets point toward the active site. Notice, in Figure 4.63, that one of the domains is inserted as a unit into the connection between a strand and a helix in the other. Loop regions adjacent to the switch points (see Figure 4.53) of both domains participate in forming the active site. In enzymatic reactions where two different substrates participate, they could be bound to different domains and brought together for catalytic reactions by the orientation of these domains. In other proteins, the two domains bind different regions of the same ligand. The bacterial arabinose-binding protein is an example of this second case.

D. STRUCTURAL PRINCIPLES OF MEMBRANE PROTEINS

4.31 Lipid bilayers form barriers that are nearly impermeable to polar molecules

Cells and the organelles within them are bounded by membranes, which are extremely thin ($\sim 35 \text{ \AA}$) films of lipids (see Chapter 3). The principal lipid components of membranes are **phospholipids**, in which a phosphorylated polar or charged scaffold, known as the **head group**, is attached to two hydrocarbon chains. Cell membranes are composed of two sheets of phospholipids that are packed against each other to form a **lipid bilayer**, in which the polar head groups are on the outside and the hydrophobic chains are on the inside (Figure 4.64).

A biological membrane functions as a permeability barrier that impedes the movement of polar or charged molecules across the membrane (Figure 4.65). The interior of the bilayer is hydrophobic, with no hydrogen-bonding capability. If a polar molecule were to enter the lipid bilayer, it would lose hydrogen bonds with water without gaining compensatory hydrogen bonds from the lipid. The loss of hydrogen bonding imposes a large energy penalty that opposes the movement of polar molecules into the bilayer.

Lipid bilayers establish discrete compartments within the cell and prevent the random mixing of the contents of one compartment with those of another. But,

Figure 4.64 A lipid bilayer.

(A) Chemical structure of a phospholipid. (B) Molecular model of a phospholipid. (C) Molecular structure of a lipid bilayer. (D) Schematic diagram of a lipid bilayer.

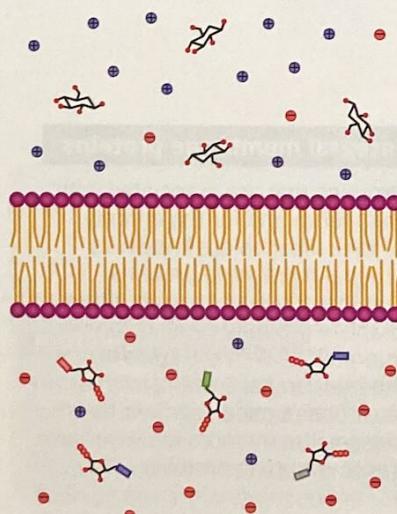


Figure 4.65 Phospholipid bilayers are impermeable to polar molecules. The hydrophobic center of the lipid bilayer cannot form hydrogen bonds, and so polar molecules do not readily pass through membranes.

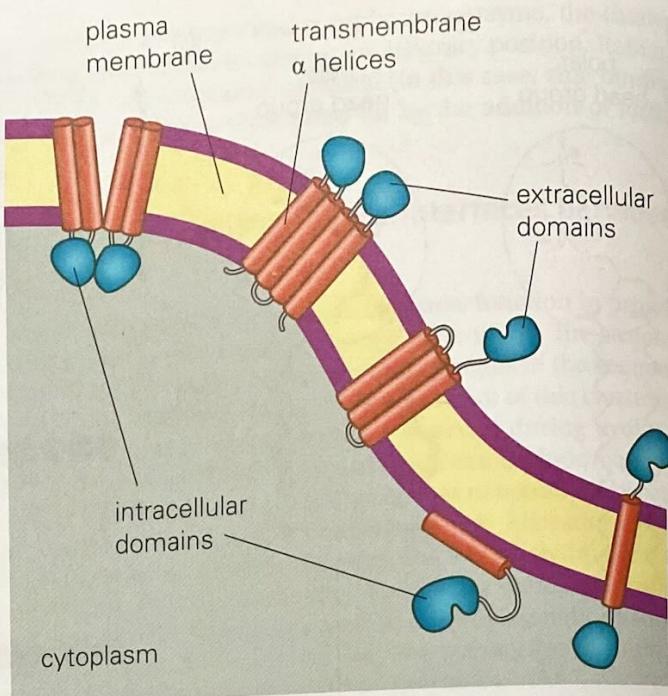
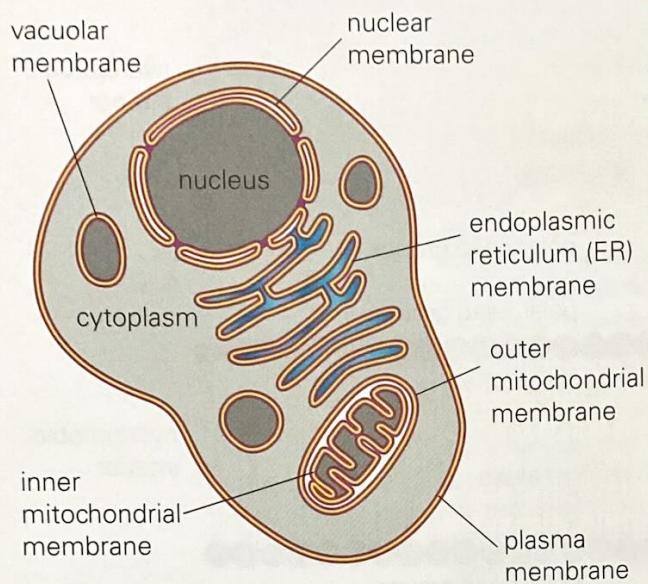


Figure 4.66 Cellular membranes. A schematic diagram of a eukaryotic cell is shown, with various cellular compartments and membranes indicated. The drawing is not to scale, and the width of the membranes with respect to the size of the cell is exaggerated. The exploded view shows a small segment of the plasma membrane along with several membrane proteins.

a living cell depends on the exchange of molecules across these membranes, without which the cell could not maintain its energy level or synthesize its molecular components. Thus, the transport of polar molecules across the membrane is facilitated by proteins embedded in the lipid bilayer.

A eukaryotic cell is illustrated schematically in Figure 4.66, in which various cellular compartments are shown along with the membranes that encapsulate them. The plasma membrane, for example, is the outer membrane of the cell. Other membranes, such as that of the endoplasmic reticulum, serve to separate internal compartments from the cytoplasm. All of these membranes are decorated with a dense array of proteins, known as **integral membrane proteins**, as shown in Figure 4.66.

Integral membrane proteins serve as mediators between the cell and its environment or the interior of an organelle and the cytoplasm. They catalyze the specific transport of metabolites and ions across the membrane barriers. They convert the energy of sunlight into chemical and electrical energy, and they couple the flow of electrons to the synthesis of ATP. Some membrane proteins act as signal receptors and transduce signals across the membrane. The signals can be, for example, neurotransmitters, growth factors, hormones, light, or chemotactic stimuli. Membrane proteins of the plasma membrane are also involved in cell–cell recognition.

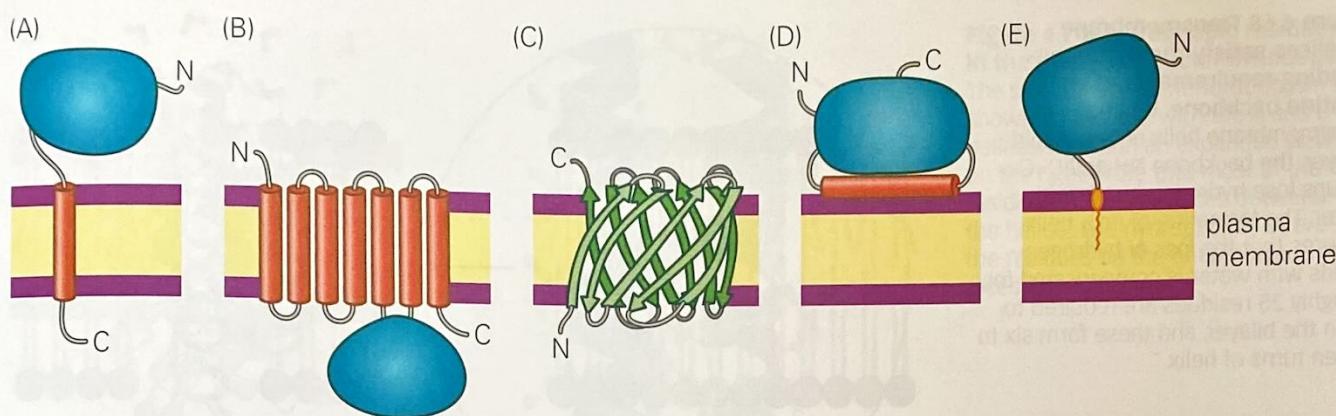
Integral membrane proteins

Proteins that are associated with the membrane for the entirety of their lifespan are known as integral membrane proteins. The three-dimensional structure of these proteins depends on interactions with the lipids in the membrane. Proteins with at least one peptide segment that crosses the membrane bilayer are transmembrane proteins.

4.32 Membrane proteins have distinct regions that interact with the lipid bilayer

Membrane proteins with the simplest architecture have three distinct regions: one segment that crosses the membrane, known as the **transmembrane segment**, and two segments, one on each side of the membrane, that are exposed to water (Figure 4.67A). The amino acid residues within the transmembrane segment are almost exclusively hydrophobic, a point which we discuss in more detail in Section 4.34. The segments outside the membrane have amino acid distributions that are similar to those found in water-soluble proteins, and these segments often form folded structures with a hydrophobic core and a hydrophilic exterior.

The polypeptide chains of other kinds of transmembrane proteins pass through the membrane several times, usually as α helices (Figure 4.67B). Recall from



Chapter 1 that the rise along the helix axis is 1.5 \AA per residue in an α helix, and so ~25 residues (six to seven turns, with 3.6 residues per turn) are required to span a bilayer that is $\sim 35 \text{ \AA}$ thick. Some transmembrane proteins use β strands to cross the membrane (Figure 4.67C). In these cases, the hydrophilic regions on either side of the membrane are the termini of the chain and the loops between the membrane-spanning parts. Some membrane-associated proteins do not traverse the membrane, but are instead attached to one side, either through α helices that lie parallel to the membrane surface (Figure 4.67D) or by fatty acids, covalently linked to the protein, that intercalate in the lipid bilayer of the membrane (Figure 4.67E). These are known as **peripheral membrane proteins**.

4.33 The hydrophobicity of the lipid bilayer requires the formation of regular secondary structure within the membrane

We had noted, in Section 4.4, that when the peptide backbone of a water-soluble protein travels into the hydrophobic core, it is required to form regular secondary structure (either α helices or β sheets; see Figure 4.8). This is a consequence of the hydrogen-bonding requirements of the peptide backbone. Because the sidechains in the hydrophobic core are nonpolar and do not form hydrogen bonds, the peptide backbone must hydrogen bond with itself. α helices and β sheets can be thought of as efficient mechanisms for completely satisfying the hydrogen-bonding requirements of the peptide backbone.

The hydrophobic interior of a lipid bilayer, like the hydrophobic core of a protein, provides no hydrogen-bond donors or acceptors. The energetic penalty for removing a polar group from water, with which it can form strong hydrogen bonds, into the lipid bilayer can be very substantial unless the hydrogen-bonding capacity of the group is satisfied. An α helix is completely self-contained in terms of its backbone hydrogen bonding, and proteins that cross the membrane only once always do so by forming a membrane-spanning α helix (Figure 4.68).

One corollary of the need to satisfy the hydrogen-bonding requirement of the backbone is that membrane-spanning α helices are highly stable on their own, in contrast to isolated α helices in soluble proteins. As discussed in Section 4.4, water forms hydrogen bonds with the peptide backbone when it is unfolded, thereby weakening the α -helical conformation in soluble proteins. The situation within the membrane is quite different because of the absence of water, and isolated membrane-spanning α helices do not unfold easily within the membrane.

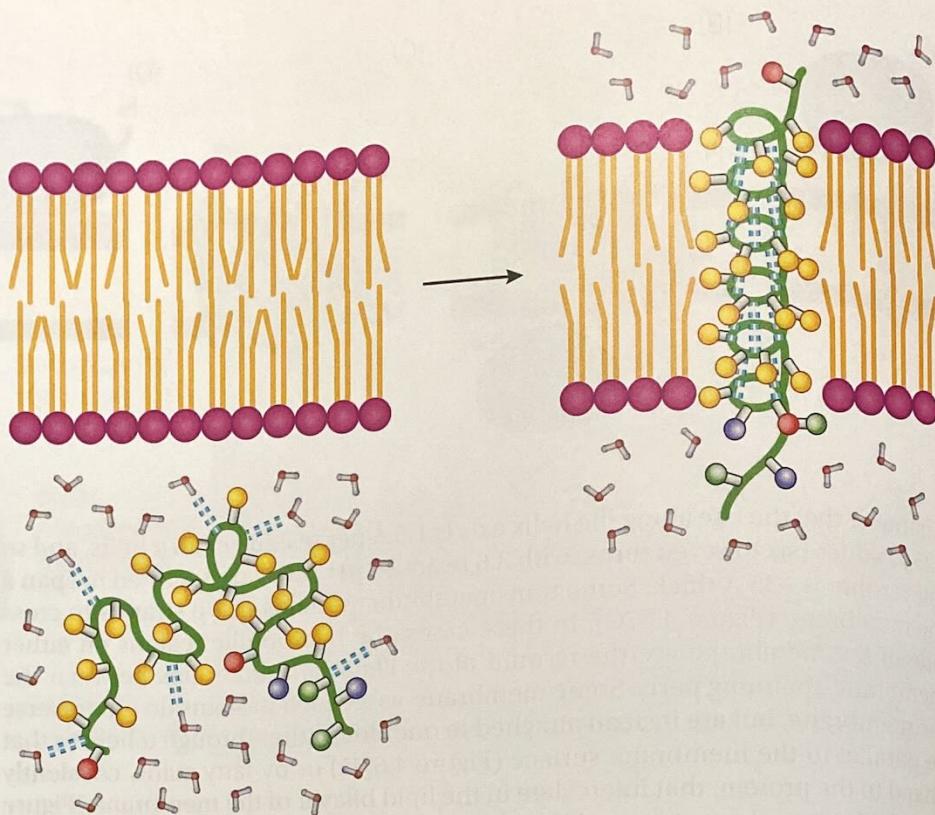
All of the membrane-spanning segments of proteins with multiple transmembrane segments form α helices (Figure 4.69A) or, in rarer cases, β sheets (Figure 4.69B). Proteins that use β sheets to span the membrane must ensure that there are no loose edges to the sheet with uncompensated hydrogen bonds. β sheets that span the membrane are always seen to form closed barrels (see Figure 4.69B).

Figure 4.67 Five different ways in which protein molecules may be bound to a membrane. From left to right are (A) a protein whose polypeptide chain traverses the membrane once as an α helix, (B) a protein that forms several transmembrane α helices connected by hydrophilic loop regions, (C) a protein with several β strands that forms a channel through the membrane, (D) a peripheral membrane protein that is anchored to the membrane by one α helix that is parallel to the plane of the membrane and interacts with it, and (E) a peripheral membrane protein that is covalently linked to a lipid. (Adapted from C. Brändén and J. Tooze, Introduction to Protein Structure, 2nd ed. New York: Garland Science, 1999.)

Peripheral membrane proteins

Proteins that are tightly associated with membranes, but do not traverse the membrane, are known as peripheral membrane proteins. Such proteins may be bound to the membrane by noncovalent interactions or covalently attached to a membrane lipid.

Figure 4.68 Transmembrane α helices satisfy the hydrogen-bonding requirements of the peptide backbone. When a transmembrane helix enters a lipid bilayer, the backbone NH and C=O groups lose hydrogen bonds with water. The formation of an α helix ensures that the loss of hydrogen bonds with water is compensated for. Roughly 25 residues are required to span the bilayer, and these form six to seven turns of helix.



Bacteriorhodopsin

Bacteriorhodopsin was the first integral membrane protein to have its three-dimensional structure determined. Bacteriorhodopsin is a proton pump found in certain “light-harvesting” bacteria, and it couples light energy to the generation of a proton gradient across the cell membrane. Bacteriorhodopsin has seven transmembrane helices, and its structure is reminiscent of that of a large family of transmembrane proteins in eukaryotic cells known as G-protein-coupled receptors (GPCRs). GPCRs transduce signals across the membrane. Rhodopsin, for example, is a GPCR that converts the detection of light photons into a neuronal signal.

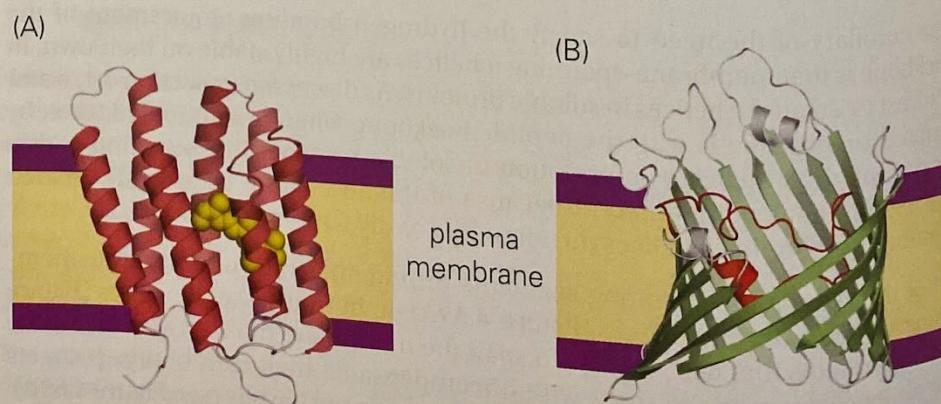
4.34 The more polar sidechains are rarely found within membrane-spanning α helices, except when they are required for specific functions

As we shall see in Section 4.35, the amino acids can be ranked in terms of their hydrophobicity or polarity. Membrane proteins very rarely insert the more polar or charged sidechains into the hydrophobic environment of the membrane. When a polar sidechain is inserted into the membrane, it usually has an interaction partner in the same protein or in an interacting protein or small molecule.

One might imagine that membrane proteins might be constructed in an “inside-out” fashion, with hydrophobic residues on the outer membrane-interacting face and polar sidechains in the interior. This is not the case, however, for α -helical membrane proteins, as illustrated in Figure 4.70 for two of the helices in **bacteriorhodopsin**, an integral membrane protein with seven transmembrane helices (some β barrel proteins provide an exception, as discussed in Section 4.37). The interactions between transmembrane helices within the same protein rely mainly on hydrophobic residues.

One reason for the general exclusion of polar residues at these interfaces is that hydrogen-bonding interactions, in contrast to hydrophobic interactions, require

Figure 4.69 α helices and β sheets in membrane proteins. (A) Structure of bacteriorhodopsin, which consists of seven transmembrane helices. The protein contains a light-sensitive chromophore known as retinal, shown in yellow. (B) Structure of a porin, in which 16 β strands form an antiparallel β barrel that traverses the membrane. A long loop between two β strands (red) constricts the channel of the barrel. The porin protein is part of a trimer, and the two other proteins in the trimer are not shown. The β strands in the front of the molecule are shorter because they pack against other molecules in the trimer rather than the membrane. (PDB codes: A, 1C8S and B, 2POR.)



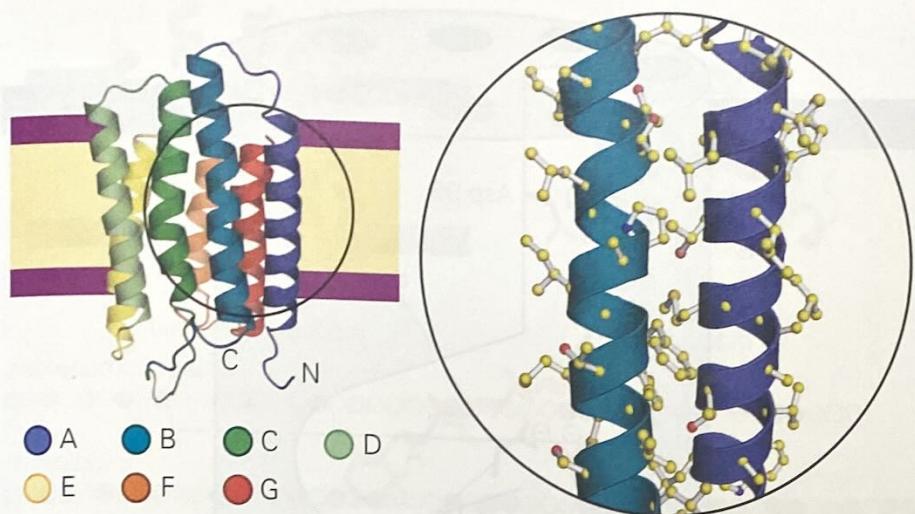


Figure 4.70 Hydrophobic residues in the helices of bacteriorhodopsin. The structure of bacteriorhodopsin is shown on the left, with the A and B helices shown in an expanded view on the right. Note that most of the residues at the interfaces between the helices are hydrophobic, as are the residues on the outer surface that faces the membrane.

the interacting groups to be in quite specific geometries. The proper coordination of one polar residue by another in a hydrophobic environment (where water cannot help) may be difficult to achieve.

Another reason for the dearth of interacting polar sidechains within the membrane could arise from the sensitivity to mutation of interactions between polar sidechains. The accidental loss of one polar residue in such a circumstance is likely to destabilize the protein, because it would leave its hydrogen-bonding partner unsatisfied. Most single-base mutations in codons for hydrophobic residues result in a codon for another hydrophobic residue (see Figure 1.48), making hydrophobic interactions less sensitive to the effects of mutations.

Some polar residues, particularly charged ones, are critical for the proper functioning of the protein and are found within the membrane-spanning segments. Figure 4.71 shows the distribution of amino acid residues in bacteriorhodopsin. Several

Figure 4.71 Amino acid sequence of bacteriorhodopsin. The seven transmembrane helices of bacteriorhodopsin are shown as cylinders. Positively and negatively charged residues are indicated in blue and red, respectively. Notice that there are several of them within the membrane-spanning regions of the helices. These residues have important functional roles, which accounts for their presence within the membrane. (Adapted with permission from D.M. Engelman et al., and B.A. Wallace, *Proc. Natl. Acad. Sci. USA* 77: 2023–2027, 1980.)

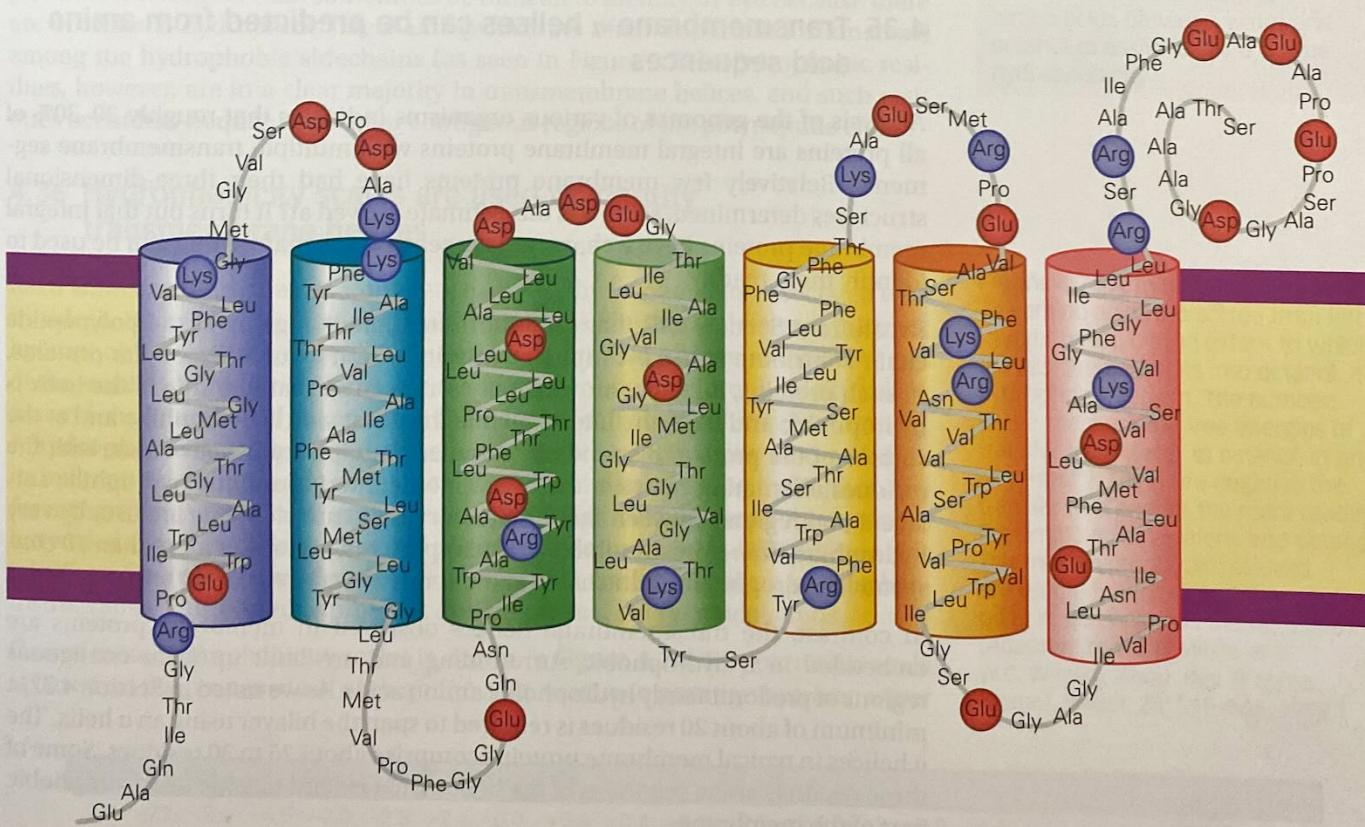
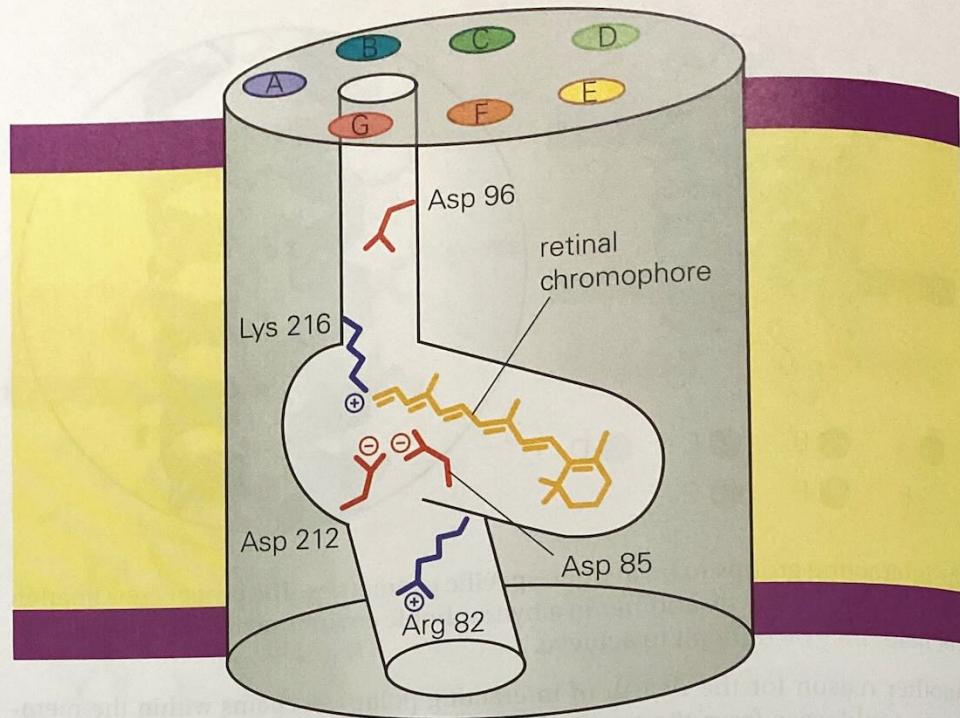


Figure 4.72 Charged residues in bacteriorhodopsin. A proton conducting conduit within the protein is indicated. Charged residues line this pathway and interact with the retinal chromophore. (Adapted from C. Brändén and J. Tooze, Introduction to Protein Structure, 2nd ed. New York: Garland Science, 1999, and R. Henderson et al., and K.H. Downing, *J. Mol. Biol.* 213: 899–929, 1990. With permission from Elsevier.)



charged residues are located within the transmembrane segments. Bacteriorhodopsin contains an internal network of residues that is part of the mechanism by which it functions as a light-driven proton pump, and the retinal chromophore is part of this conduit. Most of the charged residues either line this conduit, where they facilitate the movement of protons, or interact with retinal (Figure 4.72). We discuss the mechanism of bacteriorhodopsin in more detail in Sections 4.38–4.40, where the important role played by some of these charged residues is explained.

4.35 Transmembrane α helices can be predicted from amino acid sequences

Analysis of the genomes of various organisms indicates that roughly 20–30% of all proteins are integral membrane proteins with multiple transmembrane segments. Relatively few membrane proteins have had their three-dimensional structures determined, so how is this estimate arrived at? It turns out that integral membrane proteins have a characteristic sequence signature that can be used to identify them readily.

Recall from Section 4.13 that only short contiguous regions of the polypeptide chain contribute to the hydrophobic interior of water-soluble globular proteins. In such proteins, α helices are generally arranged so that one side of the helix is hydrophobic and faces the interior, while the other side is hydrophilic and at the surface of the protein. β strands in globular proteins are usually short, with the residues alternating between the hydrophobic interior and the hydrophilic surface. Loop regions between these secondary structure elements are usually very hydrophilic. Therefore, in soluble globular proteins, regions of more than 10 consecutive hydrophobic amino acids in the sequence are very rarely encountered.

In contrast, the transmembrane helices observed in membrane proteins are embedded in a hydrophobic surrounding and are built up from contiguous regions of predominantly hydrophobic amino acids. As we noted in Section 4.32, a minimum of about 20 residues is required to span the bilayer using an α helix. The α helices in typical membrane proteins comprise about 25 to 30 residues. Some of these residues are in segments of the helices that extend outside the hydrophobic

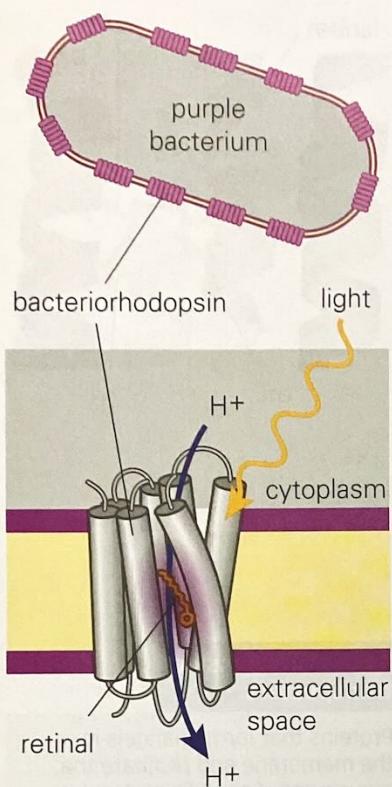


Figure 4.80 Bacteriorhodopsin is a light-driven proton pump. The membrane of the purple bacterium, *Halobacterium halobium*, is densely packed with molecules of bacteriorhodopsin. Light causes isomerization of the retinal chromophore bound to bacteriorhodopsin, which is coupled to the movement of protons out of the cell.

4.40 Bacteriorhodopsin uses light energy to pump protons across the membrane

Particularly important among active transporters are those that are involved in generating a chemical gradient that can be coupled to the production of ATP. Here we provide a brief description of how bacteriorhodopsin utilizes light energy to generate a gradient in proton concentration across the membrane (Figure 4.80).

Bacteriorhodopsin is found in the membrane of *Halobacterium halobium*, a bacterium that gives the salt flats in which it lives a characteristic purple color. Bacteriorhodopsin binds a photosensitive pigment known as retinal (Figure 4.81), which is closely related to the pigment that is used to capture light in our eyes. Bacteriorhodopsin uses the energy of light to pump protons across the membrane and is part of a particularly simple biological system for the conversion of light energy to chemical energy.

When retinal bound to bacteriorhodopsin absorbs a photon, it undergoes an isomerization that involves rotation about a double bond. Excitation by light switches retinal from a conformation in which all of the double bonds are in the *trans* conformation (referred to as all-*trans* retinal) to a conformation where the bond between carbon atom 13 and carbon atom 14 is in a *cis* conformation (13-*cis* retinal; see Figure 4.81). This light-driven conformational change in the retinal molecule is coupled to the movement of protons from the cytosol to the extracellular space, creating a proton gradient. This gradient is used to generate ATP and to transport ions and molecules across the membrane.

4.41 A hydrogen-bonded chain of water molecules can serve as a proton conducting “wire”

The mechanism for proton pumping by bacteriorhodopsin relies on the inability of charged species, such as protons, to transit the hydrophobic interior of the lipid bilayer with an appreciable rate. By spanning the bilayer, proteins such as bacteriorhodopsin can couple conformational changes, such as those induced by light, to the movement of protons through the protein. These mechanisms rely ultimately on the ability of the sidechains of glutamate, aspartate, or histidine to pick up or lose a proton by interconverting between neutral and charged forms (see Chapter 10 for a detailed discussion of the titration properties of these sidechains).

To understand how bacteriorhodopsin works, let us first consider a hypothetical protein that allows protons to move through it from one side of the membrane to the other (Figure 4.82). The protein contains a channel that is filled with water molecules, and two aspartic acid sidechains, labeled A and B, are located at each end of the channel. A chain of three water molecules connects the two aspartic acid sidechains.

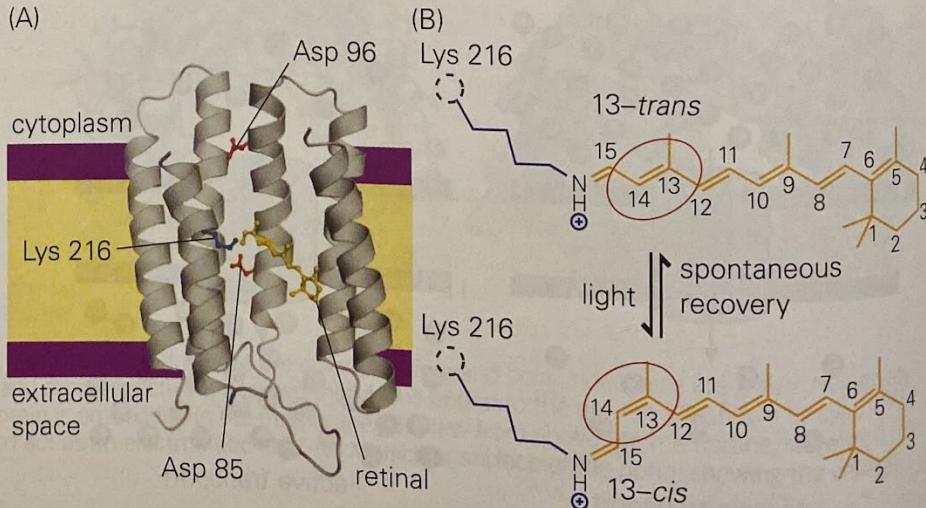
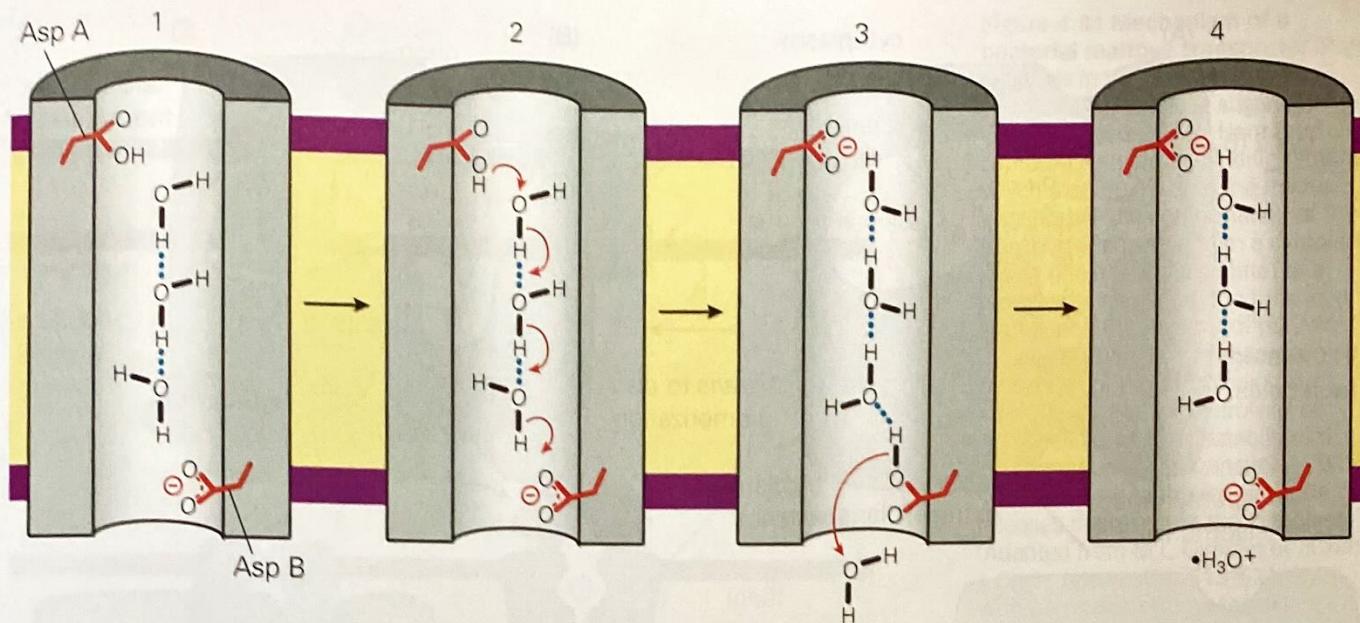


Figure 4.81 The photoisomerization of retinal. (A) The structure of bacteriorhodopsin, with one of the helices removed to show retinal bound in the interior. The retinal molecule forms a Schiff base with Lys 216 (blue). Two aspartic acid residues that are important for the proton-pumping process are shown in red. (B) The chemical structure of retinal in the Schiff base form, showing the light-induced *trans*-*cis* isomerization about the bond between carbon atoms 13 and 14.



To start with, aspartic acid A is protonated. It then releases a proton and becomes negatively charged. The released proton is picked up by the water molecule located next to the aspartic acid sidechain, which in turn donates a proton to the next water molecule, and so on down the chain. The last water molecule in the chain is located next to aspartic acid B, which is negatively charged to begin with, but becomes neutral when it picks up a proton from the last water molecule. Finally, aspartic acid B releases a proton to a water molecule outside the membrane, converting it to a hydronium ion (H_3O^+). The net effect of this series of events is to move a proton from one side of the membrane to the other, using the water network as a proton transfer “wire.”

4.42 Conformational changes in retinal impose directionality to proton flow in bacteriorhodopsin

Like the hypothetical protein depicted in Figure 4.82, bacteriorhodopsin has two critical aspartic acid residues, one closer to the cytoplasm (Asp 96) and the other closer to the extracellular space (Asp 85; see Figure 4.81). The proton transfer pathway for the hypothetical protein, drawn in Figure 4.82, has no directionality to it. Proton transfer could just as well run in the reverse direction, moving a proton from aspartic acid B to aspartic acid A. If such a protein were inserted into a membrane, it would simply cause protons to leak from one side of the membrane to the other, with the direction of net flow determined by which side has a higher proton concentration. The trick to bacteriorhodopsin is that the retinal chromophore is inserted between the two aspartic acid residues, as shown in Figure 4.81. The conformation of the retinal controls the direction of proton transfer and imposes directionality on the process regardless of the concentration of protons on the two sides of the membrane.

Retinal is bound in a pocket of bacteriorhodopsin about equidistant from the two sides of the membrane (see Figure 4.81). The pigment forms a Schiff base with a lysine residue, Lys 216; in other words, it is covalently linked to the nitrogen atom of the lysine sidechain that is protonated and therefore has a positive charge (see Figure 4.81). In the *trans* state of retinal, before light is absorbed, Asp 85 is close to the positive charge of the Schiff base (Figure 4.83A). The structural change of the retinal molecule due to the *trans*-to-*cis* photoisomerization causes the Schiff base to change its position relative to Asp 85, which induces transfer of the Schiff-base proton to the aspartate group (Figure 4.83B). Once the Schiff base-Asp 85 ion pair is converted to a neutral pair by this proton transfer, the protein undergoes a

Figure 4.82 Proton conduction by water molecules through a hypothetical protein. The protein contains a channel in which water molecules are bound. There are two aspartic acid sidechains, labeled A and B, at either end of the channel. In the process shown here, aspartic acid A starts off protonated and ends up transferring a proton to aspartic acid B and, ultimately, to the other side of the membrane. Each red arrow indicates the transfer of a proton from one oxygen atom to an adjacent one, and should not be confused with electron transfer. The proton transfer involves a domino-like transfer of protons along the water network. The mechanism, as shown here, has no intrinsic directionality to it and could just as well run in the reverse direction.

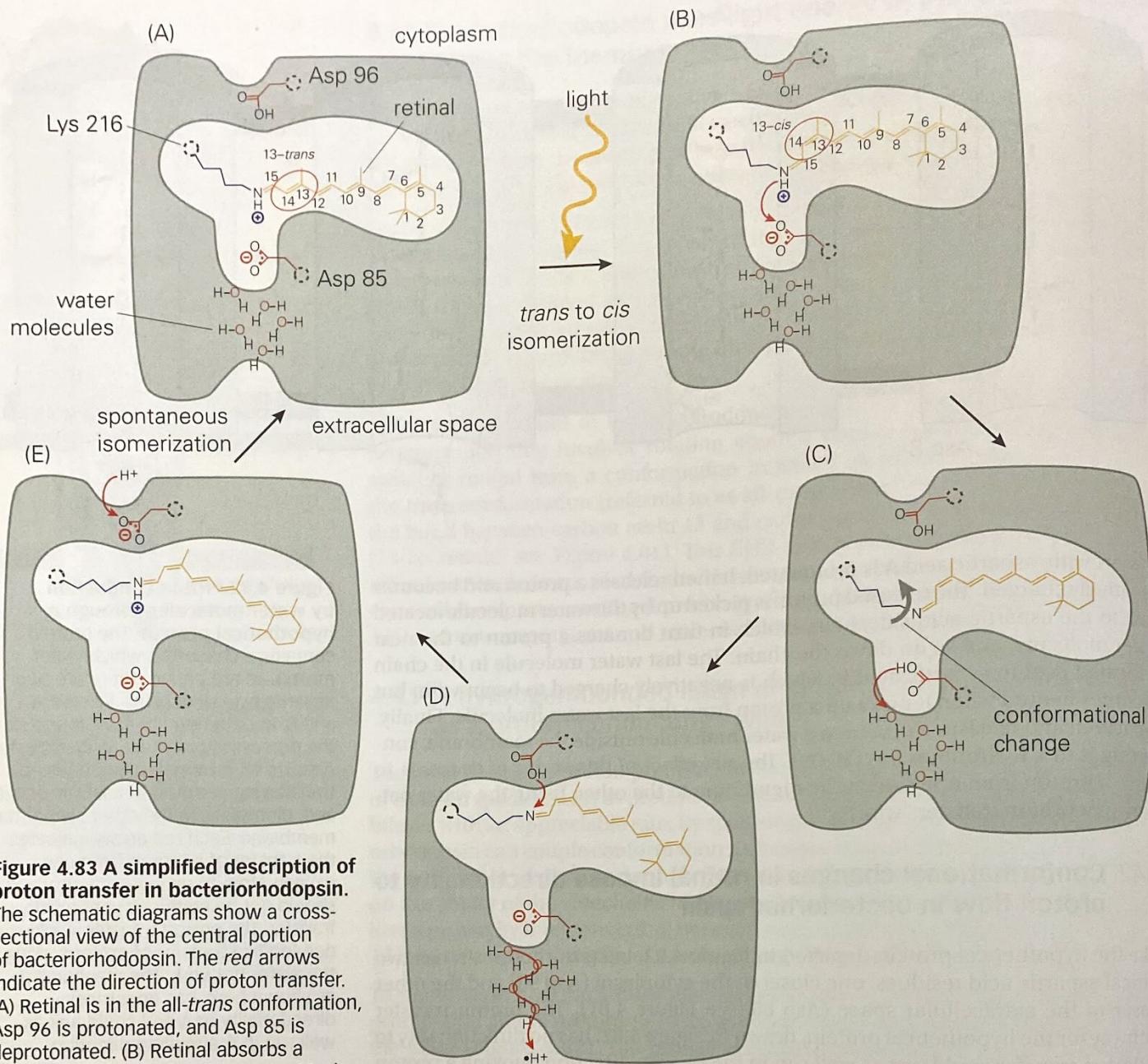


Figure 4.83 A simplified description of proton transfer in bacteriorhodopsin.

The schematic diagrams show a cross-sectional view of the central portion of bacteriorhodopsin. The red arrows indicate the direction of proton transfer. (A) Retinal is in the all-trans conformation, Asp 96 is protonated, and Asp 85 is deprotonated. (B) Retinal absorbs a photon and goes from all-trans to 13-cis. This moves the protonated nitrogen of the Schiff base close to Asp 85, which picks up the proton. (C) Asp 85 donates a proton to a water molecule that is part of a hydrogen-bonded network, leading eventually to the ejection of a proton at the extracellular side. (D) Strain in the retinal molecule causes another conformational change, which moves the unprotonated Schiff-base nitrogen closer to the protonated Asp 96 sidechain. This allows Asp 96 to donate a proton to the Schiff-base nitrogen. (E) Asp 96 picks up a proton from the cytoplasm. The 13-cis conformation of retinal eventually relaxes to the lower-energy all-trans conformation, resetting the system for another cycle. (Adapted from C. Brändén and J. Tooze, Introduction to Protein Structure, 2nd ed. New York: Garland Science, 1999.)

conformational change that involves a reorganization of some of the transmembrane helices that bind retinal, with the consequence that the Schiff base is moved away from Asp 85 and towards Asp 96 (Figure 4.83C). Asp 85 then delivers a proton through a water network to the extracellular space (Figure 4.83D) and Asp 96 re-protonates the Schiff base (Figure 4.83E). Retinal subsequently relaxes back to the *trans* state and the protein is ready for another cycle of photoisomerization-induced proton transfer.

The essential aspect of the bacteriorhodopsin mechanism is that light causes a chemical change at the active site that alters the conformation of the protein, which in turn drives protons from the cytosolic side of the membrane to the extracellular side. The critical role played by the transmembrane protein is to provide a conduit across the hydrophobic span of the membrane for the transit of charged species. This general principle underlies the mechanism of the much more complicated membrane protein assemblies, known as photosynthetic reaction centers (see Figure 4.75), that couples light to chemical energy in plants and photosynthetic bacteria.