



6

Protein Structure & Folding

Goal

To understand protein folding as a chemical equilibrium.

Objectives

After this chapter, you should be able to:

- describe the four levels of protein structure and the thermodynamic forces that stabilize them.
- explain how entropy (S) and enthalpy (H) contribute to Gibbs free energy.
- use the equation $\Delta G = \Delta H - T\Delta S$ to determine the dependence of the favorability of a reaction on temperature.
- explain the hydrophobic effect and its role in protein folding.

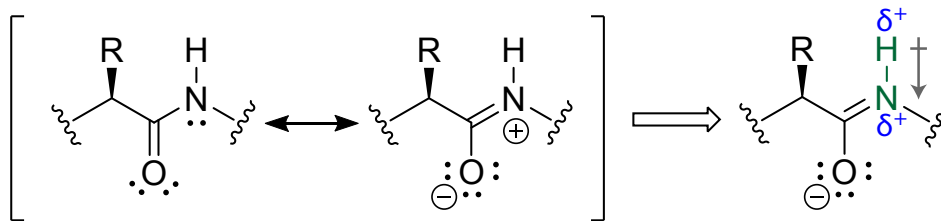
In the last chapter we learned that proteins are composed of amino acids linked together by peptide bonds. We also learned that the twenty amino acids display a wide range of chemical properties. In this chapter we will see that how a protein folds is determined by its amino acid sequence and that the three-dimensional shape of a folded protein determines its function by the way it positions these amino acids. Finally, we will see that proteins fold because doing so minimizes Gibbs free energy and that this minimization involves both making the most favorable bonds and maximizing disorder.

Proteins exhibit four levels of structure

The structure of proteins can be broken down into four levels of organization. The first is **primary structure**, the linear sequence of amino acids in the polypeptide chain. By convention, the primary sequence is written in order from the amino acid at the N-terminus (by convention usually on the left) to the amino acid at the C-terminus. The second level of protein structure, **secondary structure**, is the local conformation adopted by stretches of contiguous amino acids. Two common types of secondary structures in proteins are **alpha (α) helices** and **beta (β) sheets**. The third level is **tertiary structure**, which is the three-dimensional folded architecture of a polypeptide chain. Tertiary structure results from the interaction of regions of secondary structure with each other in precisely defined ways to form a distinct shape. Finally, **quaternary structure** is the interaction of individual folded polypeptide chains to form a higher order complex. Not all proteins exhibit quaternary structure, but many do as in

Figure 1 Amide hydrogen atoms carry a partial positive charge

The nitrogen of the peptide bond carries a substantial amount of positive charge due to resonance. Since the N-H bond (green) is polar and electrons in the bond are more strongly attracted to nitrogen than hydrogen (indicated with gray arrow). This particular bond is even more polar than a typical N-H bond because the excess positive charge carried by nitrogen increases its tendency to attract electrons. As a consequence, the partial positive charge carried by the amide hydrogen atom is larger in magnitude than the charge carried by the hydrogen in a typical N-H bond.



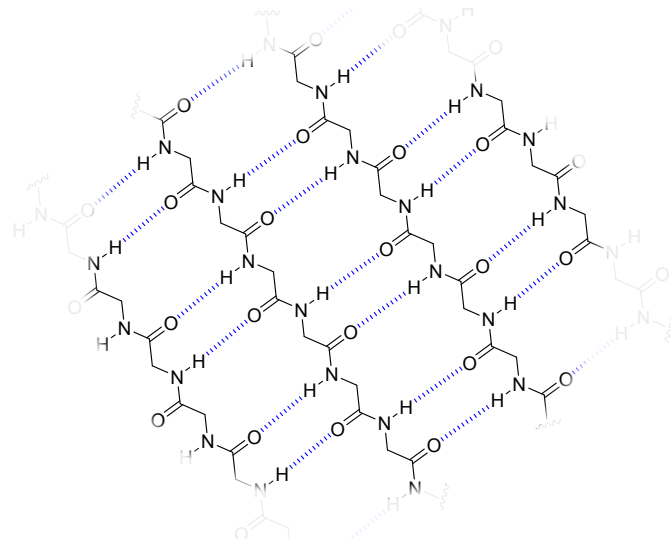
the case of hemoglobin, which is a tetramer of four globin polypeptides (Chapter 5). Some higher order complexes consist of multiple copies of the same protein whereas others are heterocomplexes consisting of two or more kinds of proteins. (Hemoglobin is a heterocomplex of two kinds of globin polypeptides.) Describing protein structure in four levels should not be taken to indicate sequential steps in protein folding. For example, the formation of secondary structure does not necessarily precede the formation of tertiary structure; instead, these two levels of structure tend to form concurrently, with the tertiary structure influencing the formation of secondary structural elements and vice versa.

Secondary structure results from hydrogen bonding between peptide bonds

As you will recall, peptide bonds are polar, owing to a resonance form that imparts a partial positive charge on the hydrogen atom of the N-H bond (Figure 1) and a partial negative charge on the carbonyl oxygen. Because these atoms carry charge, they have a tendency to form hydrogen bonds with oppositely charged atoms. In many cases, these atoms simply form hydrogen bonds with water, a process that is thermodynamically favorable. These atoms can also form hydrogen bonds with each other, which is even more energetically favorable. The reason for this is that the atoms in peptide bonds carry more charge than the hydrogen and oxygen atoms in water molecules and interactions that involve more charge are stronger. Thus, a hydrogen bond between two peptide bonds is stronger than a hydrogen bond between a peptide bond and water. When peptide bonds form hydrogen bonds with each other, the hydrogen atom attached to the amide nitrogen is donated to the carbonyl oxygen, which acts as the hydrogen bond acceptor. As we learned in the Chapter 2, hydrogen bonds are directional (exhibiting a bond angle of 180°); therefore, the peptide backbone must adopt specific conformations in order for atoms in peptide bonds to hydrogen bond with each other. Hydrogen bonding between atoms in peptide bonds is a common theme in protein structure and forms the basis for all secondary structure.

The great 20th century chemist Linus Pauling made fundamental contributions to our understanding of the structure and function of biomolecules, including enzymes. In the early 1950's, before entire protein structures had been solved by X-ray crystallography (a technique that will be introduced in subsequent chapters), Pauling was studying the crystal

Figure 2 The crystal structure of a synthetic glycine polymer suggested that folded proteins feature hydrogen bonds between backbone amides



structure of simple, synthetic polypeptides. X-ray experiments had shown that many proteins gave a periodic diffraction pattern, suggesting a regular feature in their structures, but the best crystallographers of the day, including William Lawrence Bragg, were unable to explain the pattern. Pauling thought that simple polymers might provide insight into how peptides fold. Shown in Figure 2 is one of Pauling's synthetic polypeptides. Pauling noticed in crystal structures of this simple polymer that the peptide bond was flat and that the dipoles of the carbonyls and amides in adjacent strands were aligned. Based on his observations about the synthetic polymer, he predicted that proteins could similarly fold into structures in which peptide bond dipoles were aligned.

In Pauling's seminal work, individual peptide chains were interacting intermolecularly to form extended sheets of polyglycine. As we now discuss, identical interactions also occur within a single polypeptide chain of a protein. Because of the directional requirement and the limited conformational freedom of the peptide backbone, only two secondary structures are permitted that allow for a repeating pattern of hydrogen bonds between backbone atoms within the polypeptide chain of a protein. These two forms are the beta (β) sheet and the alpha (α) helix, which are highly favored in protein structure because they maximize intramolecular hydrogen bonding.

β -strands are discontinuous stretches of amino acids that interact with each other via hydrogen bonds to form β -sheets

A β -sheet is composed of stretches of amino acids known as **β -strands** that are embedded within a single polypeptide chain. β -strands associate intramolecularly by aligning the dipoles of their peptide backbones such that complementary charges interact via hydrogen bonds. To do this, the strands in these β -sheets assume an extended, zigzag conformation in which carbonyls and amides of successive residues point in opposite directions (Figures 3 and 4). In this extended conformation, there is no opportunity for hydrogen bonding between backbone atoms within a single β -strand. Instead, hydrogen bonds form between the backbone atoms of separate β -strands. As few as two strands can interact in this way. When

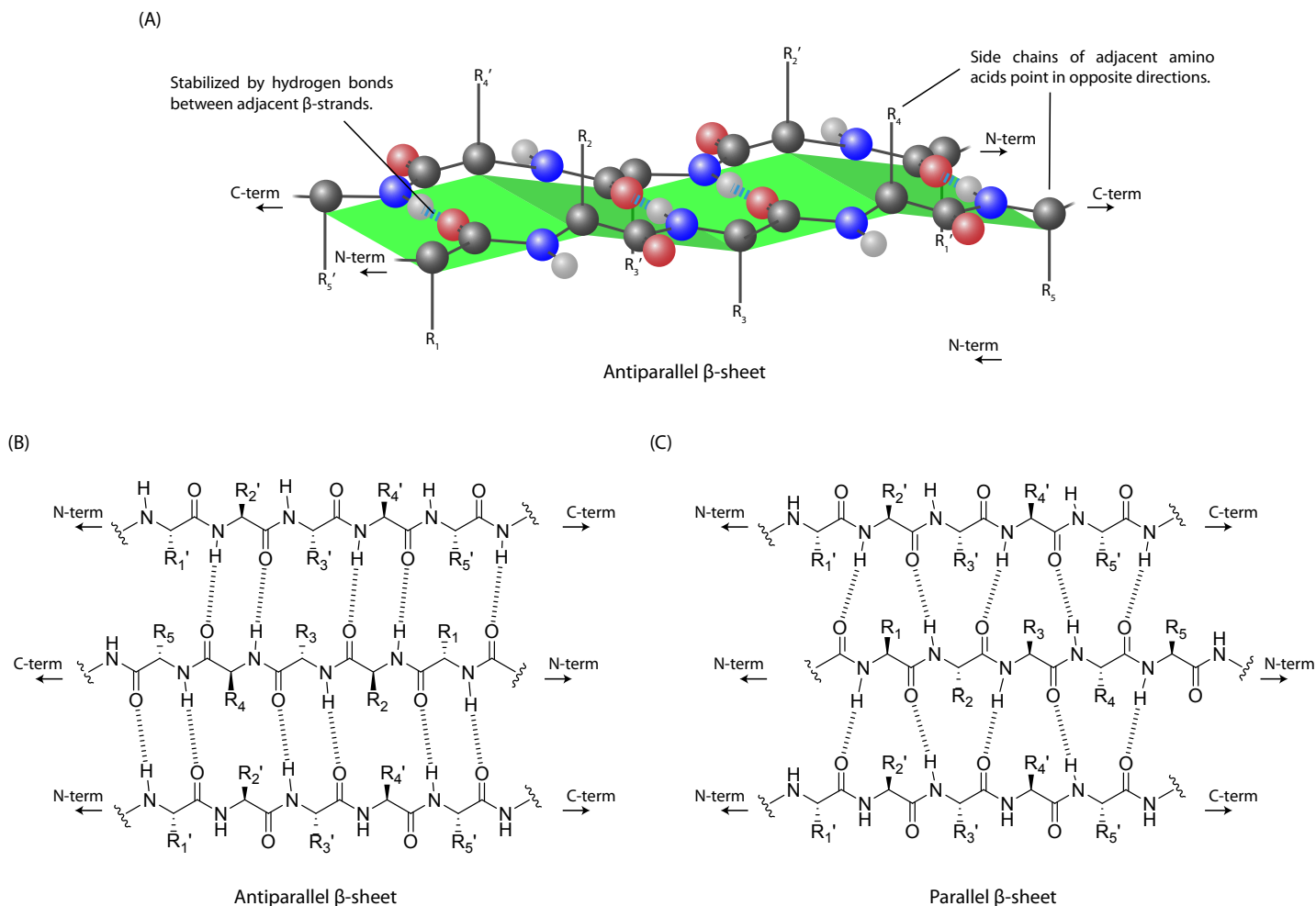


Figure 3 Peptide bonds in a β -sheet point in alternate directions

(A) Line drawing of an antiparallel β -sheet. (B) Line drawing of a parallel β -sheet. (C) Cartoon diagram of a generic antiparallel β -sheet. The green planes highlight the pleated shape of the β -sheet. Hydrogen bonds between backbone atoms are indicated with dashed lines (light blue). Carbon, gray; oxygen, red; nitrogen, blue; hydrogen, white.

multiple strands from different regions of a polypeptide interact in this way a relatively flat, sheet-like surface is created, the β -sheet. Notice that the side chains in β -sheets point outward from each side of the surface with the side chains of adjacent amino acids alternately displayed above and below the sheet. For example, if one side chain in a β -sheet points to the bottom of a β -sheet, the following residue in the primary sequence will have a side chain that points to the top of the β -sheet.

The orientation of β -sheets can be parallel or antiparallel. In parallel sheets, the β -strands are aligned such that their N-termini (and likewise C-termini) all point in the same direction (Figure 3B). In antiparallel sheets, the N-termini and C-termini of adjacent strands are orientated oppositely (Figure 3A). There are subtle variations in hydrogen bonding patterns between adjacent strands in parallel and antiparallel β -sheets but the basic features of cross strand interactions are maintained. An example of a two-stranded β -sheet in which two strands are linked by a short loop of amino acids that allows the two strands to fold back on themselves is shown in Figure 4. Notice that this results in an anti-parallel orientation.

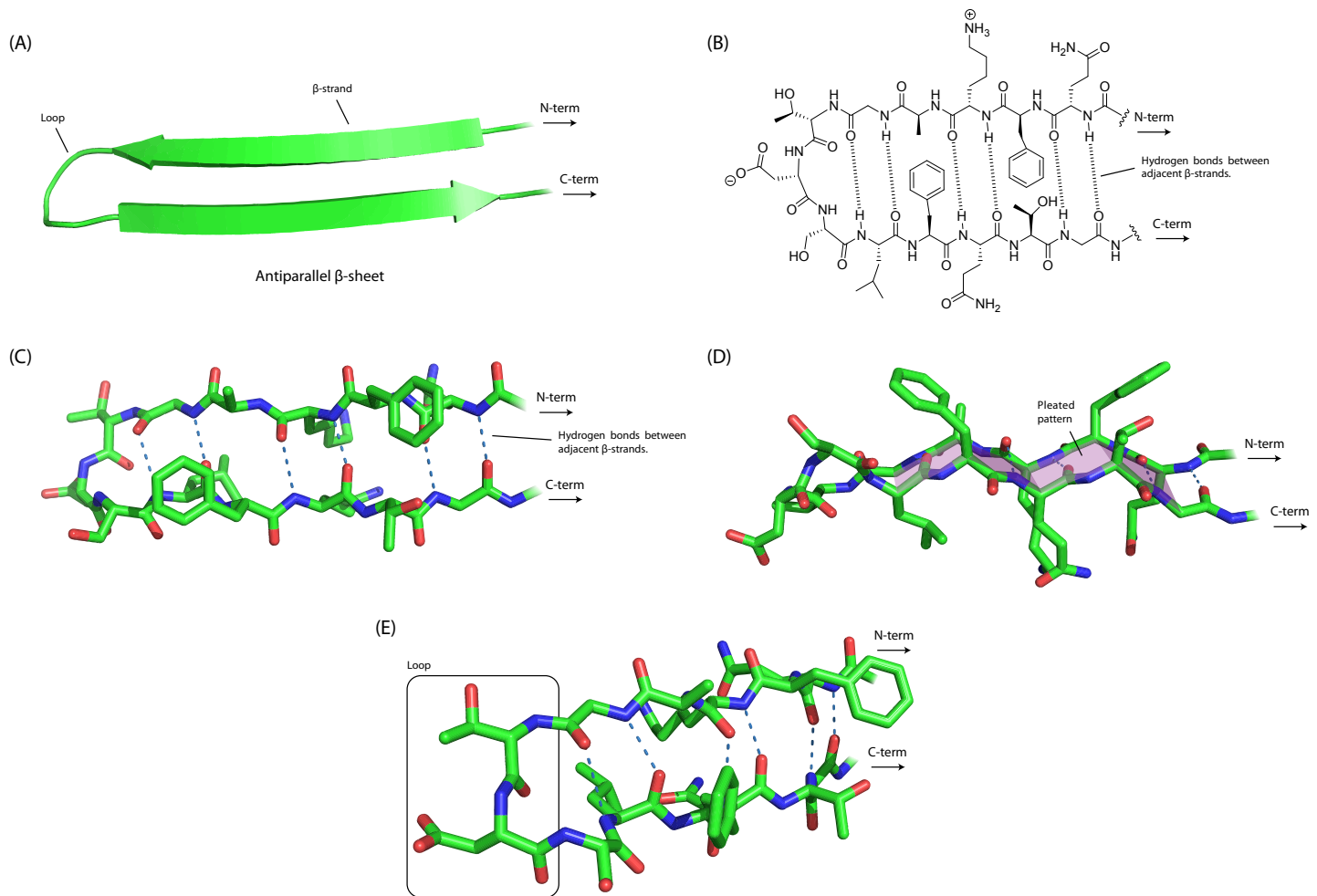


Figure 4 An example of a β -sheet formed from two β -strands connected by a loop

(A) Cartoon representation of an excerpt from a larger protein. In this example, two β -strands are connected by a short loop to form an antiparallel β -sheet. Arrows representing β -strands point toward C-terminus by convention. (B) Line drawing of the chemical structure of the polypeptide from (A). (C) Stick representation of the polypeptide from (A) and (B) that highlights the hydrogen bonds that stabilize the β -sheet (shown in blue). Carbon, green; oxygen, red; nitrogen, blue. (D) Alternate view of (C) that highlights the pleated pattern of the β -sheet (highlighted in purple). (E) Alternate view of (C) and (D) that highlights the location of the loop that connects the two β -strands.

Some β -strands can assemble into very extensive sheets as new strands are added to each side of the sheet. This can produce insoluble aggregates, or clumps of protein. In Alzheimer's disease a protein in the brain is cleaved, and one of the resulting fragments forms a highly extensive β -sheet that aggregates to form plaques whose presence is correlated with tissue death.

Peptides bonds in α -helices point in the same direction

Another secondary structure that maximizes hydrogen bond formation within a single polypeptide chain is the α -helix. In α -helical regions of secondary structure, the polypeptide chain is organized in a right-handed corkscrew with all of the carbonyl oxygen atoms along the helix pointing towards the C-terminus and all of the amide hydrogen atoms pointing towards the N-terminus (Figure 5). This arrangement of carbonyls and amide hydrogens with respect to the axis of the helix is a consequence of the angles that the bonds in the peptide backbone can assume. In an α -helix, hydrogen bonds can form between amino acids in one turn of a helix and

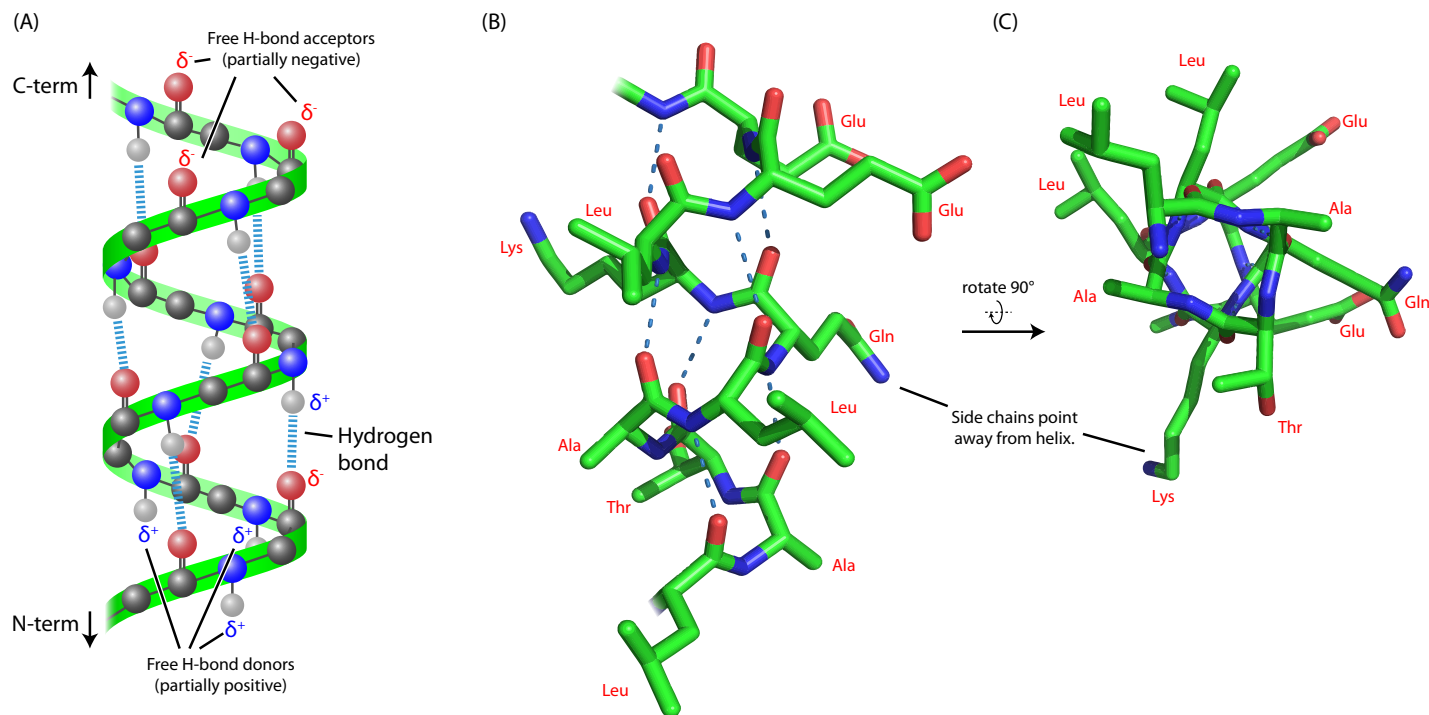


Figure 5 Peptide bonds in an α -helix all point in the same direction

(A) Cartoon diagram of a generic α -helix. Side chains are omitted from this cartoon. Hydrogen bonds between backbone atoms are indicated with dashed lines (light blue). Carbon, gray; oxygen, red; nitrogen, blue; hydrogen, white. (B) Stick representation of an α -helix with the sequence "NH₂-LATALQKLEE-COOH." Hydrogen bonds between backbone atoms are indicated with dashed lines (light blue). Carbon, green; oxygen, red; nitrogen, blue. (C) Rotated view of (B). Notice that side chains radiate outward, away from the center of the helix.

amino acids in the succeeding turn. The carbonyl oxygen of a particular residue in the helix is perfectly aligned with the amide hydrogen of the residue four amino acids away, allowing these two atoms to form a strong hydrogen bond. Measuring from the carbonyl carbon at one end of a turn of the helix and the amide nitrogen at the other end, each full turn of the helix (360°) is 3.6 amino acids in length. The rise of the α -helix per amino acid is 1.5 Å, meaning that the length of a helix can be easily estimated from the number of amino acids it contains. For example, a twenty amino acid long helix would be approximately 30 Å long. A key feature of α -helices is that the side chains are displayed pointing outward around the axis of the helix to avoid steric clashes with each other. This means that interactions between α -helices are mediated by favorable contacts between side chains.

Within a stretch of α -helix, all of the backbone amides and carbonyls make hydrogen bonds to each other except for the carbonyl oxygen atoms at the C-terminal end of the helix and the amide hydrogen atoms at the N-terminal end of the helix. Note that the first three amide hydrogen atoms and the last three carbonyl oxygen atoms in the α -helix (going from N- to C-terminus) are unpaired. Because there are unsatisfied hydrogen donors and acceptors at the amino and carboxy termini, respectively, and because all of the dipoles in the helix point in the same direction, there is a significant partial positive charge at the N-terminal end of the α -helix and a corresponding partial negative charge at the C-terminal end. Because a large amount of charge is separated across a large distance, α -helices are said

to contain **macroipoles**. The charges associated with these macroipoles can allow the two ends of the helix to participate in attractive interactions with oppositely charged moieties.

Finally, to appreciate that the α -helix is right-handed, in your mind's eye superimpose your right hand over the helix in Figure 5a with your thumb pointing up. Your fingers will wrap around the helix with the helix rising in the direction of your thumb. Conversely, if you do this with your left hand, the helix will be descending away from the direction that your thumb is pointing. We will use this mnemonic, also known as the physicist's right-hand rule, in Chapter 7 to see that the DNA is a right-handed double helix.

Proteins can contain different types of secondary structure

Most proteins contain a mixture of α -helices and β -sheets, but some are composed exclusively of just one type of secondary structure (Figure 6). Segments of the peptide backbone that do not form secondary structure are called **loops**. Loops are linkers that connect regions of secondary structure. The cartoon representation of a protein's structure represents α -helices and β -sheets as coils and flat arrows, respectively. Each coil and arrow represents a contiguous stretch of amino acids. Loops are shown in cartoon representations as lines connecting regions of secondary structure. It should be noted that proline is disfavored in secondary structures because its amide cannot participate in hydrogen bonding (Figure 7). Instead proline is often found in loops because it can assume the *cis* configuration, which forces the chain to bend back on itself. Because loops are often flexible regions of proteins, glycine is also often found in loops.

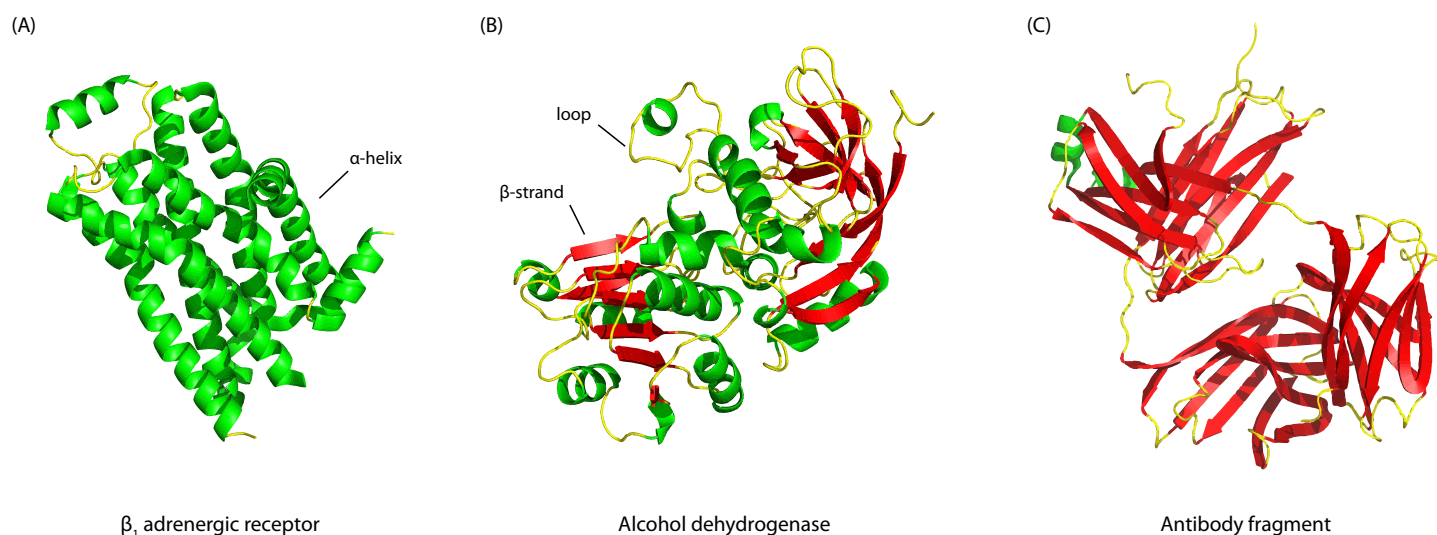
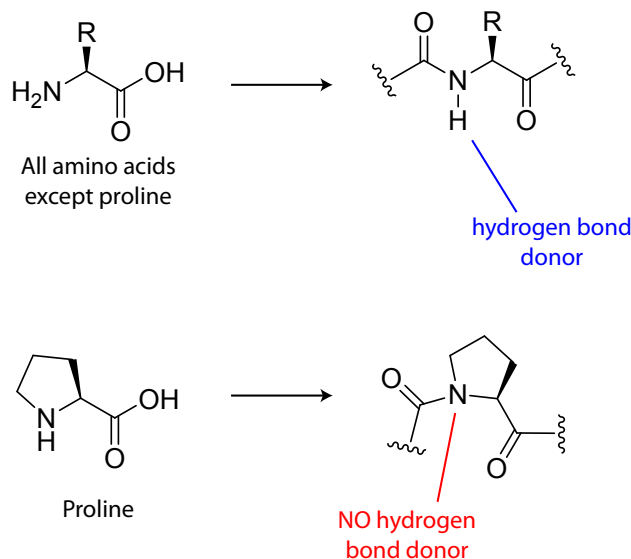


Figure 6 Examples of secondary structure elements in proteins

(A) The β_1 adrenergic receptor is an example of a protein that is made entirely of α -helices. (B) Alcohol dehydrogenase is an example of a protein that is a mixture of both α -helices and β -sheets. (C) Antibodies are examples of proteins that almost exclusively contain β -sheets. In each structure loops are shown in yellow, α -helices are shown in green, and β -sheets are shown in red.

Figure 7 Proline disrupts secondary structure

All amino acids except proline have a hydrogen attached to their amide nitrogen when they are part of a peptide (top). Proline lacks this hydrogen (bottom). This missing hydrogen atom is needed as a donor to form the hydrogen bonds that stabilize secondary structure. As a result, proline disrupts secondary structure because it prevents hydrogen bonding between backbone atoms.



Intra- and intermolecular interactions help to stabilize tertiary and quaternary structure

Whereas hydrogen bonding interactions between backbone atoms stabilize secondary structure, ionic interactions, hydrogen bonds, and van der Waals interactions help stabilize tertiary and quaternary structure. These interactions occur between functional groups in different portions of the same molecule in tertiary structure and between molecules in quaternary structure. Whereas secondary structure is largely stabilized by interactions between backbone atoms, tertiary and quaternary structure is principally stabilized by interactions between side chains. Side chains are chemically heterogeneous; consequently, these interactions can involve ions, permanent dipoles, and induced dipoles (Figure 8).

Ionic interactions take place between amino acid side chains that are ionized at physiological pH, meaning that they carry full positive or negative

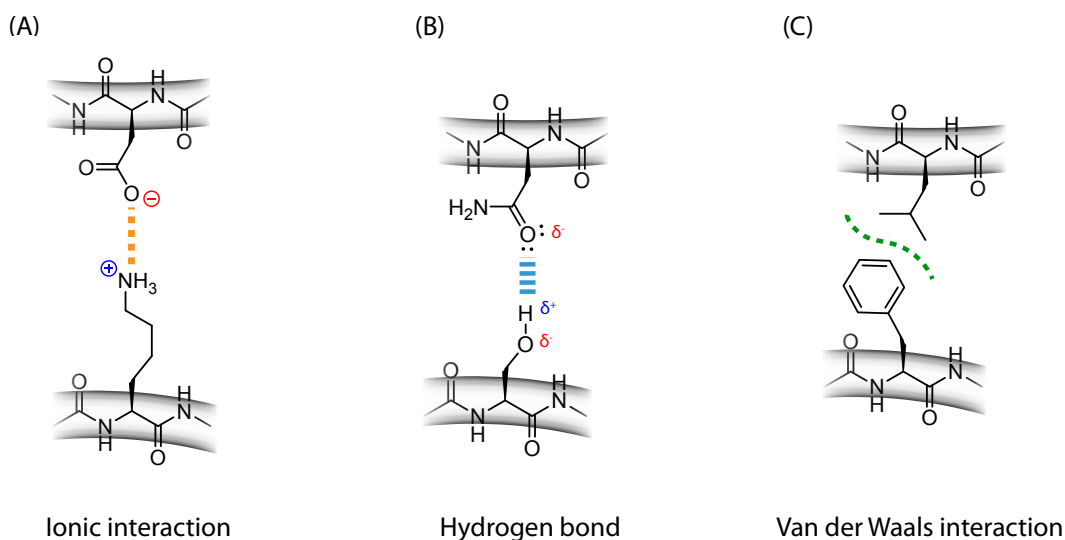


Figure 8 Interactions between side chains help stabilize tertiary and quaternary structure

Interactions between amino acid side chains help to stabilize the folded structures of proteins. These interactions can include (A) ionic interactions, (B) hydrogen bonds, and (C) van der Waals interactions.

charges. Ionic side chains of opposite charge (e.g., lysine and aspartate) can form favorable ionic interactions that contribute to the thermodynamic favorability of protein folding. Ionic interactions that form in proteins are called **salt bridges**. As discussed in Chapter 2, the attractive force between two opposite charges is inversely dependent on the polarity of the environment around the two charges. For example, the attractive interaction between two charges buried in the nonpolar interior of a protein is approximately twenty times stronger than the attractive interaction between two charges in water. The dipoles of the water molecules surrounding each charge align such that they partially attenuate the ions' charges, decreasing the attraction between them. In effect, the formation of a salt bridge is in competition with interactions with water molecules in an aqueous environment. Conversely, because there is no water on the interior of a protein, and because most of the buried side chains are nonpolar, the interactions between side chains of opposite charge are highly favorable in an anhydrous environment.

Hydrogen bonds also contribute to the stability of protein structure. Because hydrogen bonds are directional, and because unfavorable alignment of dipoles is repulsive, hydrogen bonds between side chains play key roles in determining the unique structures that different proteins form.

Protein structure is also stabilized by van der Waals interactions between nonpolar surfaces. As we have seen, nonpolar bonds can be transiently polarized by the random motion of electrons or by the presence of a nearby charge, and the resulting instantaneous dipoles can attract one another. Such interactions between nonpolar surfaces are weak, but the strength of van der Waals interactions scales with the surface area along which they occur. Proteins tend to densely pack numerous hydrophobic side chains into the protein interior, resulting in a significant amount of surface area over which van der Waals interactions occur, contributing to the stability of tertiary and quaternary structure.

The hydrophobic effect is a principal force stabilizing tertiary and quaternary structures

In addition to ionic, hydrogen bonding, and van der Waals interactions, an important driving force for protein folding is the **hydrophobic effect**. This phenomenon is the tendency for nonpolar surfaces to interact with each other rather than with water. The hydrophobic effect leads to the burial of nonpolar side chains in the interior of proteins, which in turn collapses polypeptide chains into compact, globular structures. This collapse *does not*, however, happen because van der Waals interactions in the interior are favorable (they are but only weakly so). Rather, the principal driving force for burying nonpolar side chains comes from minimizing their interactions with water molecules. The hydrogen bonds formed between water molecules are far stronger and hence highly favored over the dipole-induced dipole interactions between water molecules and nonpolar side chains. As a consequence, water molecules tend to become highly ordered around nonpolar surfaces (much like the ordering of water molecules in ice), maximizing hydrogen bonding with each other and minimizing interactions with the nonpolar side chains. This ordering of water molecules

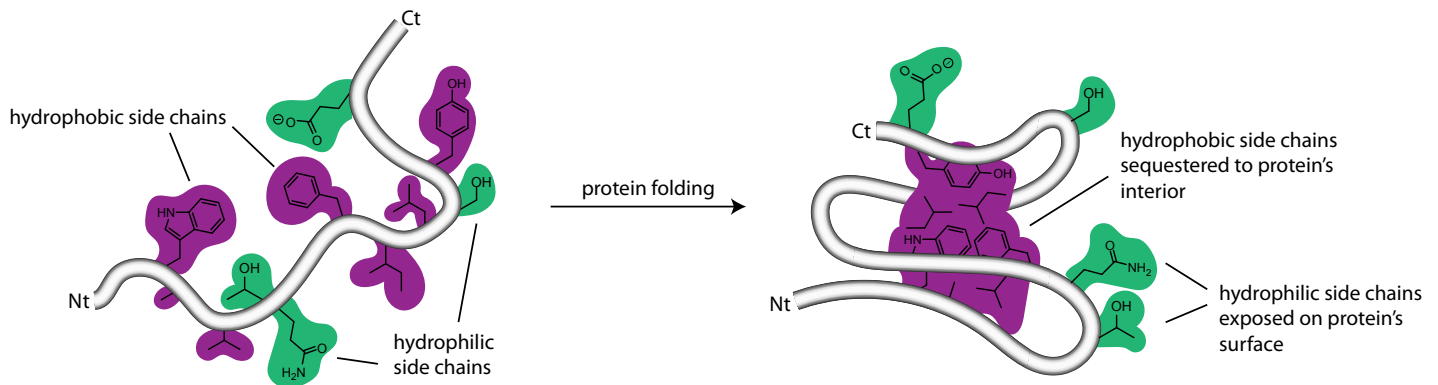


Figure 9 The arrangement of hydrophobic and hydrophilic side chains in a protein determines its folded structure

Because of the hydrophobic effect, hydrophobic side chains (purple) are sequestered to the protein's interior when it folds. Hydrophilic side chains (green) are usually left exposed on the protein's surface.

on hydrophobic surfaces increases the degree of order in the system. As we will come to presently, disorder (**entropy**) is thermodynamically favored over order. Thus, burying nonpolar surfaces in the interior of a protein is favored because it *maximizes disorder*. That is, minimizing the nonpolar surface area (by collapsing the polypeptide chain into a globule) minimizes the number of water molecules that become ordered into ice-like structures. Minimizing the ordering of water molecules is a principal driving force in the folding of proteins. As we shall see in Chapter 7, the hydrophobic effect is also a principal driving force for stabilizing the wrapping of two polynucleotide strands around each other in the DNA double helix.

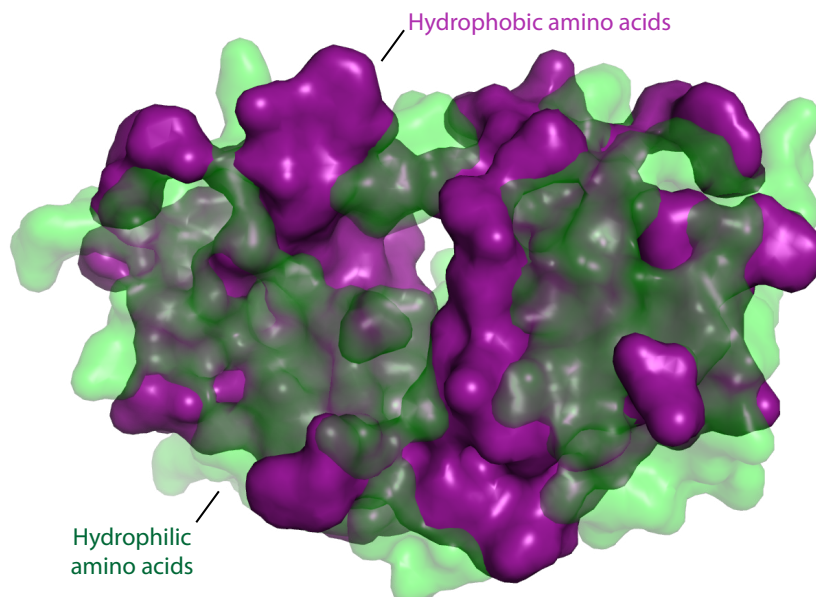
Because of the importance of the hydrophobic effect in protein folding, the distribution of nonpolar and polar amino acids within primary structure plays a critical role in determining a protein's architecture. Nonpolar, hydrophobic side chains, such as those belonging to phenylalanine, leucine, isoleucine, valine, methionine, and tryptophan, tend to cluster in the interior of the folded protein. In contrast, polar side chains, such as those belonging to arginine, glutamine, glutamate, and lysine tend to arrange themselves near the outside of the protein, where they can form hydrogen bonds with water and other polar molecules (Figures 9 and 10). Some ionic and polar amino acid side chains are, nonetheless, located in the protein's interior. When they are located in the interior, they are inevitably in close proximity to other ionic or polar bonds with which they can interact. Interactions between complementary, charged amino acids (e.g., lysine with aspartate) in the interior are particularly strong (due to the absence of competing water molecules) and thus make a disproportionately important energetic contribution to the precise folded conformation of the protein.

The energy released during protein folding is small

Protein folding is in effect a chemical reaction in which the equilibrium is defined by the concentrations of the product (the folded protein) and the reactant (the unfolded protein). Proteins fold because the folded conformation is lower in energy than the unfolded conformation; protein

Figure 10 Polar side chains are generally found on the surfaces of folded proteins, whereas hydrophobic side chains are sequestered to the protein interior

Shown is a surface representation of an example protein in which hydrophobic amino acids are colored in purple and hydrophilic amino acids are colored in translucent green.



folding is thermodynamically favorable. Consequently, K_{eq} for protein folding is greater than one and ΔG_{rxn} is negative. In most cases, however, it does not take much energy to unfold a protein. Some proteins **denature** (that is, they unfold) at just a few degrees above body temperature. In fact, when you get a fever, some proteins actually do denature, causing the body to make special stress response proteins that help denatured proteins refold correctly. Because small amounts of energy are required to unfold a protein, we can deduce that ΔG_{rxn} for protein folding, while negative, is relatively small.

Gibbs free energy is a function of enthalpy (H) and entropy (S)

To understand the energetics of protein folding more deeply, we need to introduce the concepts of **enthalpy (H)** and **entropy (S)**. Enthalpy is a measure of the potential to take up or give off energy in the form of heat, and entropy is a measure of disorder. Both properties are directly related to Gibbs free energy according to the following equation in which T is the temperature in Kelvin:

$$G = H - T \cdot S$$

Enthalpy and entropy, like Gibbs free energy, describe properties of a system at a given state. Like Gibbs free energy, the absolute amount of entropy or enthalpy in a system is not directly measurable; instead, we typically discuss changes in these values as a system moves from one state to another. In Chapter 3 we defined ΔG_{rxn} as the change in Gibbs free energy as a reaction proceeds from 100% reactant to 100% product under standard state conditions. We can similarly define ΔH_{rxn} and ΔS_{rxn} as the changes in enthalpy and entropy, respectively, for a reaction that proceeds from 100% reactant to 100% product under standard state conditions. ΔG_{rxn} , ΔH_{rxn} and ΔS_{rxn} are related in the same way as G, H, and S:

$$\Delta G_{rxn} = \Delta H_{rxn} - T \cdot \Delta S_{rxn}$$

As we shall see this dependence of ΔG_{rxn} on temperature is important because it means that the favorability of some reactions is temperature dependent.

ΔH_{rxn} measures the strengths of bonds formed and broken in a reaction

The change in enthalpy (ΔH_{rxn}) for a reaction equals sum of the bond energies of the bonds broken during a reaction ($\Sigma BE_{\text{broken}}$) minus the sum of the bond energies of the bonds formed during the reaction ($\Sigma BE_{\text{formed}}$):

$$\Delta H_{\text{rxn}} = \Sigma BE_{\text{broken}} - \Sigma BE_{\text{formed}}$$

ΔH_{rxn} is a measure of the heat given off or taken up during a reaction. When a covalent bond is formed, heat is given off because the potential energy of the system decreases. Conversely, when a covalent bond is broken, energy is put into the system. When the sum of the bond energies that are broken is smaller than that for those that are formed, heat is given off and the sign of ΔH_{rxn} is negative. Reactions that give off heat (i.e., ΔH_{rxn} is negative) are **exothermic**. Conversely, reactions that absorb heat (i.e., ΔH_{rxn} is positive) are **endothermic**. Because ΔH_{rxn} contributes directly to ΔG_{rxn} , some reactions are favorable because their ΔH_{rxn} values are large and negative. Keep in mind that temperature and ΔS_{rxn} also affect ΔG_{rxn} ; as a result, not all exothermic reactions are favorable and not all endothermic reactions are unfavorable as we will explain.

Entropy measures the number of ways energy can be distributed in a system (disorder)

Entropy, the other component of Gibbs free energy, is a measure of disorder. Entropy relates to the number of ways energy can be distributed in a system. Entropy is higher where there are more ways to distribute energy. Molecules that move freely and adopt many different conformations have more ways to distribute energy and are higher in entropy than molecules whose movements are constrained. This is because the movements of atoms and bonds within molecules (e.g., rotations, vibrations, etc.) are all associated with specific amounts of energy, and when more movements are possible, there are more ways to distribute energy. The peptide chain of an unfolded protein can adopt millions of conformations, whereas the folded protein adopts a specific conformation. Because of this, folded proteins are more constrained and have lower entropy than unfolded proteins.

ΔS_{rxn} describes the difference in entropy between the products and reactants of a reaction:

$$\Delta S_{\text{rxn}} = S_{\text{products}} - S_{\text{reactants}}$$

ΔS_{rxn} equals the entropy of the final state of the reaction minus the entropy of the initial state; therefore, reactions with positive values of ΔS_{rxn} generate more disorder and reactions with negative values of ΔS_{rxn} increase order. Because systems that are high in entropy (i.e., disorder) are favored, large, positive ΔS_{rxn} values lead to more favorable ΔG_{rxn} values.

Figure 11 The entropy of universe always increases

The second law of thermodynamics states that the total entropy of the universe (S_{total}) always increases. We can, however, define a “system” that contains only the matter we wish to study. This could be a chemical reaction in a flask, or a cell, or whatever we are interested in observing. The entropy of that system (S_{system}) can decrease, but only if the entropy somewhere else in the universe ($S_{\text{surroundings}}$) increases.

$$\underbrace{\Delta S_{\text{total}}}_{\Delta S_{\text{total}} \text{ is always positive.}} = \underbrace{\Delta S_{\text{system}} + \Delta S_{\text{surroundings}}}_{\Delta S_{\text{system}} \text{ can be negative, but } \Delta S_{\text{surroundings}} \text{ must be positive and large enough to make } \Delta S_{\text{total}} \text{ positive.}}$$

The entropy of the universe is always increasing, but the entropy of a given system can decrease

The field of thermodynamics is based upon the Laws of Thermodynamics, which are a set of fundamental principles that inform our understanding of the universe. The First Law of Thermodynamics states that energy cannot be created or destroyed, but only converted between forms. The Second Law holds that the entropy of the universe is always increasing. That is, *every* chemical process and transfer of energy results in an increase in the total entropy of the universe. It is possible to observe chemical reactions that are associated with a decrease in entropy ($\Delta S_{\text{rxn}} < 0$). For this to occur, however, there *must* be an increase in entropy elsewhere in the universe. In general, we define a system in which the reaction is taking place, such as a cell or a flask, as distinct from its surroundings. Within that system, entropy can either increase or decrease, but there must always be a corresponding change in the entropy of the surroundings such that the total entropy of the system and surroundings is increasing (Figure 11). Imagine, for example, a reaction that reduces entropy but gives off heat ($\Delta S_{\text{rxn}} < 0$ and $\Delta H_{\text{rxn}} < 0$). Although the entropy of the system decreases as the reaction proceeds, the heat it releases increases the entropy of the surroundings.

The favorability of certain reactions is dependent on temperature

The signs (positive or negative) of ΔH_{rxn} and ΔS_{rxn} tell us how changes in temperature will affect ΔG_{rxn} , K_{eq} , and the favorability of a chemical reaction. Figure 12 lists the four possible permutations for the signs of ΔH_{rxn} and ΔS_{rxn} . The effect of temperature on ΔG_{rxn} is different in each of these cases. For example, for a reaction in which the change in enthalpy is negative and the change in entropy is positive, ΔG_{rxn} is always a negative value. This is because the “ $-T \cdot \Delta S_{\text{rxn}}$ ” term is negative if ΔS_{rxn} is positive, and no matter the magnitude of ΔS_{rxn} , the sum of ΔH_{rxn} and $-T \cdot \Delta S_{\text{rxn}}$ is always negative. On the other hand, if ΔH_{rxn} is negative and ΔS_{rxn} is also negative, then ΔG_{rxn} can be positive or negative depending on the temperature. In

Figure 12 The signs of ΔH_{rxn} and ΔS_{rxn} determine how the sign of ΔG_{rxn} is affected by temperature

ΔH_{rxn}	ΔS_{rxn}	ΔG_{rxn}
Positive	Positive	Negative at high temperatures
Negative	Negative	Negative at low temperatures
Negative	Positive	Always negative
Positive	Negative	Always positive

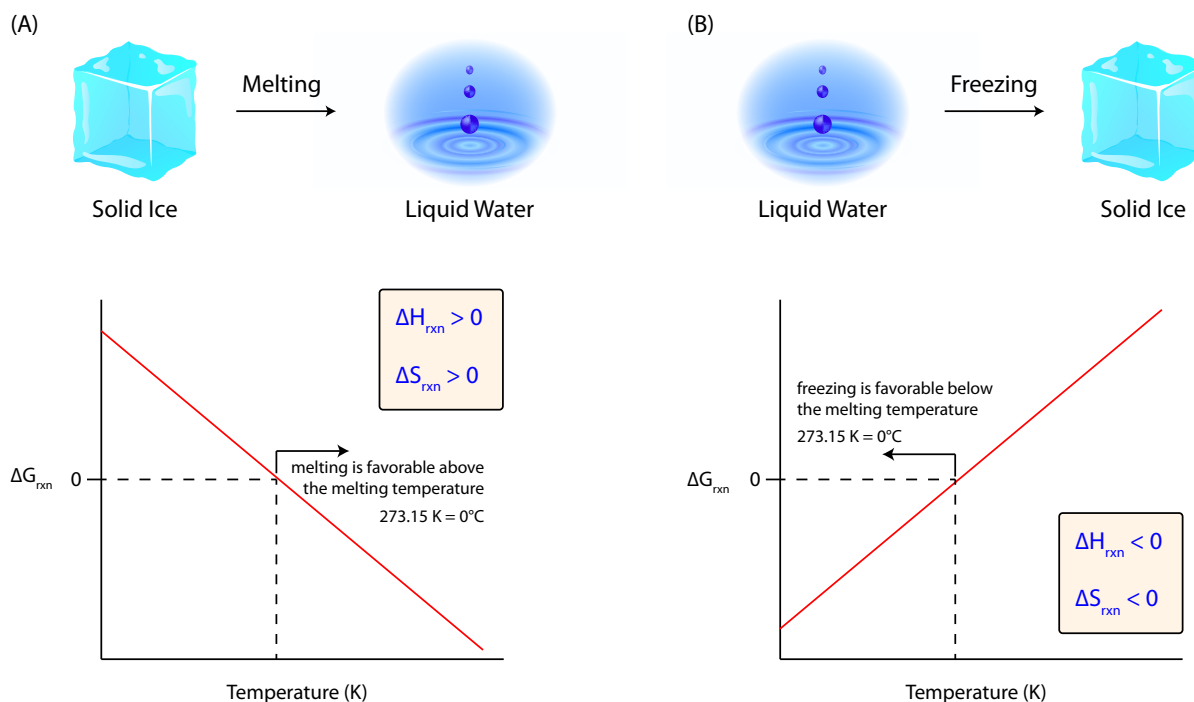


Figure 13 The temperature dependence of melting and freezing are explained by the signs of ΔH_{rxn} and ΔS_{rxn} for each process

(A) When solid ice melts into liquid water, the sign of ΔG_{rxn} is negative at high temperatures, corresponding to ΔH_{rxn} and ΔS_{rxn} values that are both positive. (B) When liquid water freezes into solid ice, the sign of ΔG_{rxn} is negative at low temperatures, corresponding to ΔH_{rxn} and ΔS_{rxn} values that are both negative.

sum, two of the four cases are independent of temperature (one case is always favorable and the other is never favorable), and the other two are dependent on temperature (in one case the reaction becomes favorable as the temperature increases and in the other the reaction becomes favorable as the temperature decreases).

Consider the example of ice melting, which becomes a favorable process as the temperature rises above freezing. That means that ΔG_{rxn} for ice melting must be negative at temperatures above 0°C . From this information alone, we can conclude that the signs of ΔH_{rxn} and ΔS_{rxn} for ice melting must both be positive because the process only becomes favorable as the temperature increases (first case in Figure 12 and see Figure 13).

Why are ΔH_{rxn} and ΔS_{rxn} for ice melting positive? When you hold an ice cube, your hand becomes cold as the ice begins to melt, and so it should make sense that ice melting is an endothermic process. Heat is withdrawn from your hand to break the hydrogen bonds that hold the water molecules in ice together, and as a result, your hand becomes cold. There are hydrogen bonds in liquid water as well, but they are not as strong as the hydrogen bonds in ice (because the optimal bond angle for hydrogen bonds is better preserved in ice than in water), and because the melting of ice requires the net breaking of bonds, ΔH_{rxn} is positive. Therefore, it requires the input of more energy to break the more favorable hydrogen bonds in ice than the amount of energy that is released by the formation of bonds in water. Because the entropy of liquid water is higher than the entropy of solid ice, ΔS_{rxn} for ice melting is positive. This should make sense because the molecules in liquid water are far less ordered than the molecules in ice. Ice

is a crystalline solid in which each water molecule is confined to a defined position in the crystal lattice. In the liquid phase, however, water molecules are constantly moving around.

The reverse process, the freezing of water, occurs spontaneously as the temperature decrease below 0°C. So the signs of ΔH_{rxn} and ΔS_{rxn} for water freezing must both be negative (second case in Figure 12 and see Figure 13). Even though ΔS_{rxn} is negative, the Second Law of Thermodynamics is not violated, as the heat released by the reaction causes molecules in the surroundings to move around more, increasing their entropy. Although the entropy of the system consisting of the ice/water reaction decreases, the entropy of the system plus its surroundings increases; therefore, the total entropy of the universe increases, consistent with the Second Law of Thermodynamics.

The basis for the hydrophobic effect is an increase in the entropy of water

Let us return to the hydrophobic effect using what we have learned about entropy and how it contributes to Gibbs free energy to understand better how it contributes to the thermodynamic favorability of protein folding. As we discussed above, water molecules that interact with hydrophobic surfaces form ordered, ice-like structures around the hydrophobic surface. These ordered water molecules have less entropy than free water molecules that make up the bulk of the solution. When hydrophobic surfaces coalesce during protein folding and exclude water, ordered water molecules are released into the bulk solution. The water molecules that are freed have more entropy than they did when they formed ordered structures around the hydrophobic surface (Figure 14). The increase in the entropy of water is the driving force of the hydrophobic effect ($\Delta S_{\text{water}} > 0$). Note that as a protein folds, the polypeptide chain loses entropy as it adopts a single conformation but this decrease in entropy is often more than offset by the concomitant gain in entropy as the folded structure buries its hydrophobic side chains.

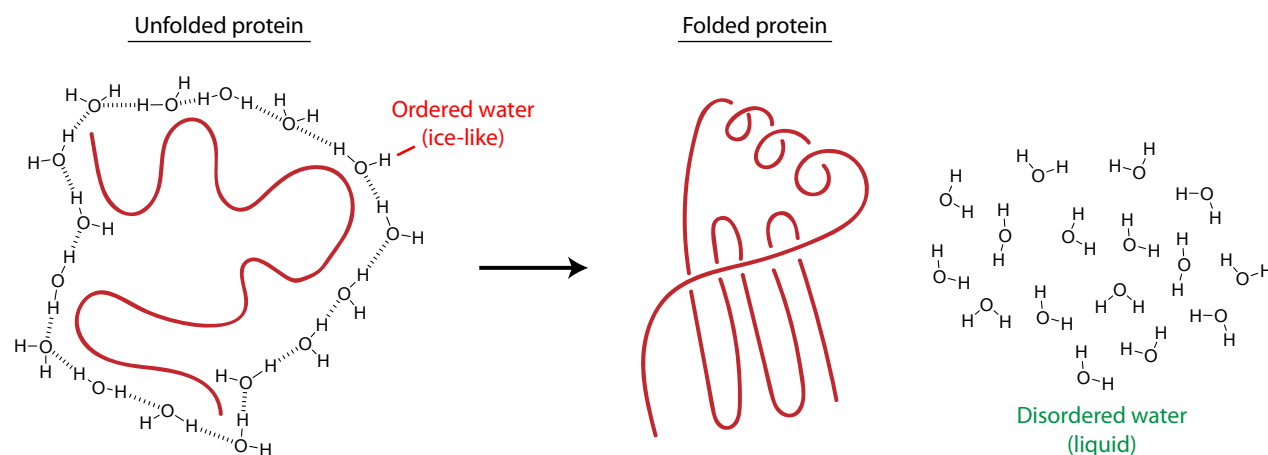


Figure 14 The hydrophobic effect contributes to the thermodynamic favorability of protein folding

When water interacts with hydrophobic side chains in unfolded proteins, it forms ordered, ice-like structures that are low in entropy. When these hydrophobic surfaces coalesce and exclude water (as the protein folds), the ordered water molecules become disordered and return to the bulk solvent. The resulting increase in the water molecules' entropy is the thermodynamic driving force behind the hydrophobic effect.

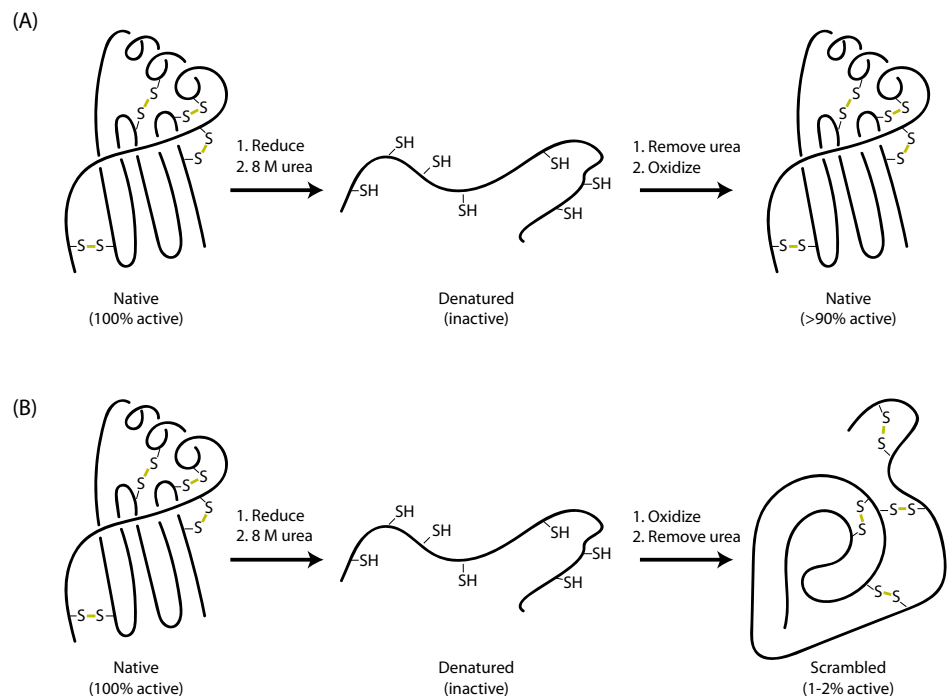
Anfinsen's experiment shows that primary structure determines protein conformation

The Nobel-winning scientist Christian Anfinsen proposed several decades ago that the three-dimensional structure of a protein is a direct consequence of its primary structure. That is, proteins find the folded conformation that is lowest in energy, meaning the conformation in which favorable interactions are maximized and unfavorable interactions are minimized. To test this hypothesis, Anfinsen denatured (unfolded) a protein and then asked if it would fold back into its native conformation in a test tube. He used the enzyme ribonuclease A, which cleaves RNA. This protein is relatively small, and because it is an enzyme, its folded state could be assessed by measuring its enzymatic activity.

Anfinsen took two samples of ribonuclease A and added both reducing agent and 8 M urea ($\text{CO}(\text{NH}_2)_2$) to each of them. The reducing agent was added in order to break the four disulfide bonds that stabilize the structure of ribonuclease A. Urea causes proteins to unfold. At high concentrations, urea alters the extensive hydrogen bonding network between water molecules (urea contains both hydrogen bond donors and acceptors) and thus interferes with the forces that promote protein folding (in effect, it makes the hydrophobic effect less important in protein folding). Anfinsen then took one of the samples and used dialysis to remove the urea. Next, Anfinsen treated this sample with oxidizing agent to promote disulfide bond formation. The protein was then tested for enzymatic activity, and it was found to have 90% of the activity of the original sample, implying that most of the protein molecules had refolded into their correct, or native, shape (Figure 15A). In a parallel experiment, Anfinsen took the other sample and, after it had been treated with reducing agent and urea, oxidized it first to promote disulfide bond formation and then removed the urea. In other words, he reversed the order of the final two steps (Figure 15B). This sample was only 1-2% as active as the original, implying that the

Figure 15 The entropy of universe always increases

(A) Ribonuclease A was denatured and its disulfide bonds were reduced. The denaturant was removed and the protein was allowed to refold before the disulfide bonds were restored with an oxidizing agent. The resulting protein regained its enzymatic activity, demonstrating that it had returned to its proper folded state. (B) In a parallel experiment, the protein's disulfide bonds were restored by adding an oxidizing agent *before* removing the denaturant. This resulting in scrambled disulfide bonds that prevented the protein from refolding into its native conformation. Consistent with expectations, the scrambled control protein had regained little of its enzymatic activity.



procedure of oxidizing while the protein was still denatured was ineffective in restoring enzymatic activity. By oxidizing the protein before removing urea, Anfinsen “trapped” conformations containing incorrect disulfide bonds. These bonds then prevented the protein from adopting its native structure again.

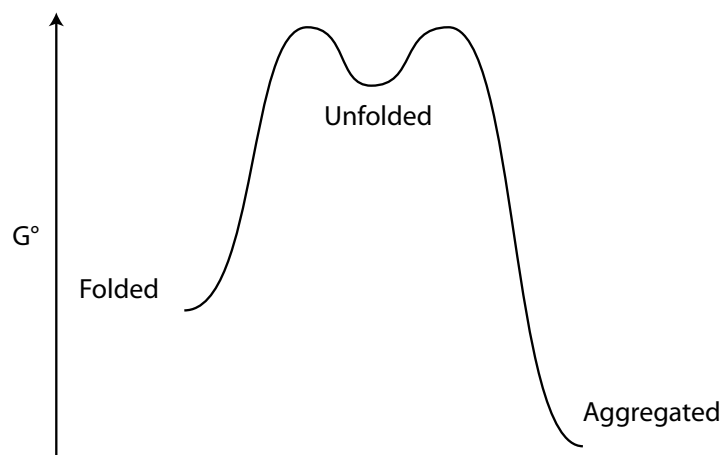
Anfinsen’s experiment showed that the native structure of ribonuclease A will form following denaturation provided that premature oxidation is prevented. Therefore, the protein is intrinsically capable of finding its lowest energy conformation. Anfinsen concluded that the information required to fold a protein into its native, lowest energy conformation is entirely contained within its sequence of amino acids.

Protein aggregation can be more favorable than folding at high concentration

Despite the example of ribonuclease A, many denatured proteins fail, in practice, to fold back into their native state on their own. A well-known example is egg white. Egg white principally consists of the protein albumin in an aqueous solution. When egg white is heated, the proteins unfold and clump together in a process called **aggregation**. When the egg white is removed from the heat, the albumin molecules remain clumped together; they do not refold as they were in the original egg white. The reason is that the proteins in the egg white are highly concentrated. The same is true in cells. A typical cell is densely packed and has a protein concentration of about 200-300 g/l, which is the equivalent of dissolving a half pound of salt in a liter of water. Whereas salt is extremely soluble, proteins are not. When proteins in a concentrated solution are denatured, they interact with neighboring protein molecules instead of folding back into the correct shape. For example, instead of burying hydrophobic residues in the protein interior, a denatured protein will aggregate with a neighboring protein by coalescing its hydrophobic side chains with those of its neighbor. Such interactions would liberate ordered water just as effectively as protein folding (i.e., ΔS_{water} is equal for both aggregation and folding) whereas not requiring the protein to become as ordered (i.e., $\Delta S_{\text{protein}}$ for aggregation is less negative than $\Delta S_{\text{protein}}$ for folding). Although protein folding is a favorable process, protein aggregation is likely a *more favorable* process (Figure 16). Anfinsen’s classic experiment circumvented this problem by

Figure 16 Protein aggregates are thermodynamically more favorable than folded proteins

Unfolded proteins tend to aggregate in concentrated solutions because aggregation is more thermodynamically favorable than folding. Folded proteins do not tend to aggregate because they require an input of energy to reach the unfolded conformation before they can aggregate.



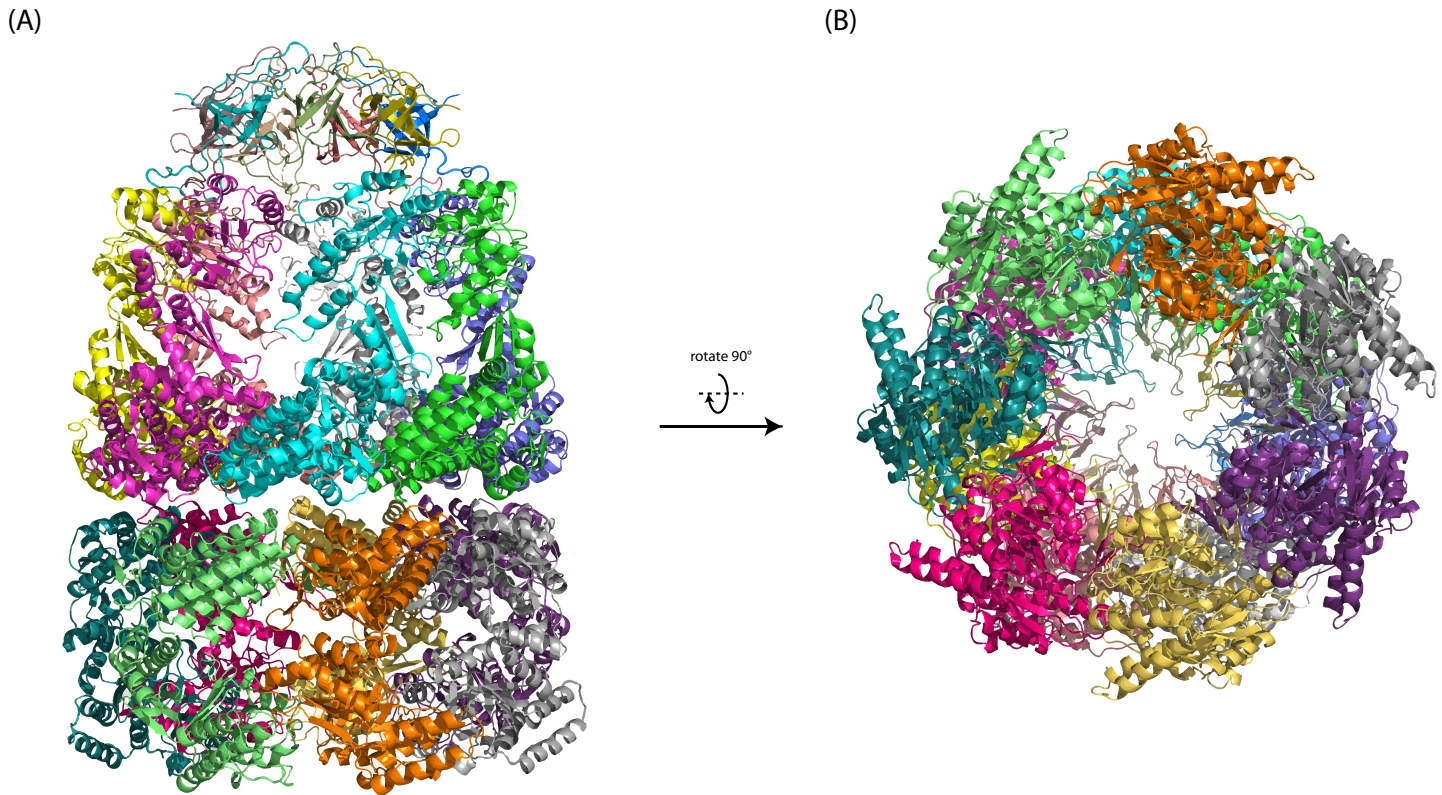


Figure 17 Chaperone proteins help other proteins fold in the densely packed environment of the cell

(A) Cartoon representation of a chaperone. Individual polypeptide subunits of the chaperone are uniquely colored. (B) Alternate view of (A). The open cavity that is evident in (B) prevents aggregation by providing a space in which proteins can fold in an isolated environment that is free of other polypeptides.

using extremely dilute solutions of ribonuclease A. This gave the protein time to fold before it encountered other protein molecules with which it could aggregate.

You might imagine that it would be hard for a protein in the densely packed environment of the cell to fold correctly. In many cases this does present a challenge. Cells deal with this issue by means of **chaperones**, which are proteins that help other proteins fold. As an example, Figure 17 shows a chaperone that contains a cavity that is just big enough to hold an individual protein. This cavity allows the protein to fold in isolation, mimicking the “high dilution” of Anfinsen’s ribonuclease A refolding experiment.

Protein aggregation can be more favorable than folding at high concentration

Proteins are organized in four levels of structure. The primary structure of a protein is the linear sequence of amino acids in the polypeptide chain. The secondary structure is the local structure of contiguous stretches of amino acids that form via hydrogen bonding between peptide bonds. Tertiary structure is the three-dimensional, folded architecture that results from interactions between various units of secondary structure. Many proteins also exhibit quaternary structure in which two or more polypeptide chains interact to form a higher order complex.

Proteins exhibit two kinds of secondary structure, the β -sheet and the α -helix. In β -sheets the peptide backbone adopts an extended conformation in which both amide and carbonyl groups point in alternating directions. Each segment of the peptide backbone with this conformation is known as a β -strand. A β -sheet is formed when multiple β -strands lie side-by-side and form hydrogen bonds between their backbone atoms. In α -helices all of the peptide bond N-H groups are pointed in the same direction and all of the carbonyl C=O groups are pointed in the same direction (opposite to the N-H groups). The peptide backbone in an α -helix forms a right-handed coil in which the carbonyl oxygen of each amino acid forms a hydrogen bond to the amide hydrogen of the amino acid four residues closer to the C-terminus.

Tertiary and quaternary structures are stabilized by intra- and intermolecular interactions between side chains and the hydrophobic effect. Intra- and intermolecular interactions between side chains are diverse in nature, owing to the chemical heterogeneity of the amino acid side chains. They include ionic interactions, hydrogen bonds, and van der Waals' interactions. The hydrophobic effect occurs when hydrophobic side chains are sequestered away from water and buried in the interior of the protein. The hydrophobic effect is distinct from van der Waals' interactions. An ordered shell of water forms over hydrophobic surfaces that are exposed to water. When hydrophobic surfaces, such as amino acid side chains, are sequestered away from water, the outer shell of water is released, and the previously ordered water molecules are released into the bulk solvent where they move more freely. Because the released water molecules are more disordered than the ordered water molecules, the hydrophobic effect causes an increase in entropy that is responsible for the thermodynamic favorability of the hydrophobic effect.

Gibbs free energy consists of enthalpic (H) and entropic (S) components. Like Gibbs free energy, enthalpy and entropy are values that describe the state of a system. The change in enthalpy for a process (ΔH_{rxn}) describes the strength of the bonds broken and formed during the process. When more energy is released from the forming of bonds than absorbed by the breaking of bonds, ΔH_{rxn} is negative. When ΔH_{rxn} is negative, the process releases heat and is exothermic. When ΔH_{rxn} is positive, the process absorbs heat and is endothermic. Negative values of ΔH_{rxn} contribute favorably to ΔG_{rxn} . Entropy is a measure of disorder, or more specifically, the number of ways in which energy can be distributed in a system. Entropy is highest when the atoms in molecules are free to vibrate, rotate, and translate back and forth. The change in entropy for a process (ΔS_{rxn}) describes the difference in entropy between the products and reactants. Positive values of ΔS_{rxn} contribute favorably to ΔG_{rxn} and indicate an increase in disorder as the reaction proceeds. The contribution of entropy to ΔG_{rxn} depends on temperature as indicated by the equation:

$$\Delta G_{\text{rxn}} = \Delta H_{\text{rxn}} - T \cdot \Delta S_{\text{rxn}}$$

This, in turn, means that the favorability of some reactions is dependent on temperature.

Finally, Anfinsen's experiment showed that the primary structure of a protein determines its three-dimensional, folded structure. Proteins often exist in dense solutions, such as in an egg or a cell, in which they can aggregate, or clump together, if they become unfolded. This is because aggregation can be thermodynamically favored over folding.