December 30, 2024 Draft

Physical Principles in Biology

Biology 3550/3551

Spring 2025

Chapter 6:
Formation of Biomolecular Structure

David P. Goldenberg
University of Utah
goldenberg@biology.utah.edu

Chapter 6

Formation of Biomolecular Structures

Now, we are going to shift direction a bit and start discussing the question of how biological structures form on the molecular and cellular level. The thermodynamic principles that we have been discussing create a framework for this subject. At a first glance, it seems that biology somehow violates or circumvents the second law of thermodynamics, since highly ordered structures seem to form spontaneously. Our goal is to understand how that can happen within the constraints of the thermodynamic laws.

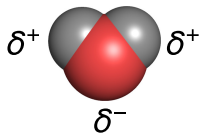
6.1 Water, Ionization and the Hydrophobic Effect

A key factor in biological assemblies is water and its special properties. We tend to take water for granted, since it is the liquid that we know the best, but it is actually a very special liquid.

The unique properties of water become obvious when we try to mix it with other liquids, especially oils. Everyone knows about this experiment: oil and water don't mix. This is, in fact, the major driving force for the formation of biological structures at the molecular level. But, *why* don't they mix? What is so fundamentally different about the two kinds of liquids?

I. Hydrogen bonding

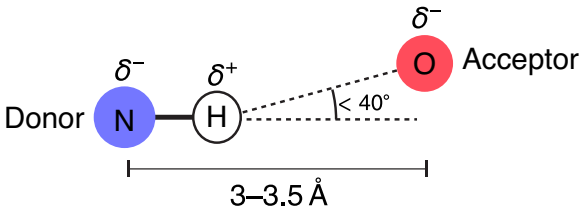
The key property of water is something that we briefly discussed earlier, the unequal sharing of electrons in the chemical bonds between oxygen and hydrogen:



This results in a partial positive charge on the hydrogen atoms and a partial negative charge on the oxygen atoms, which lead to a strong tendency of the molecules to interact. This is an example of a more general phenomenon, the hydrogen bond.

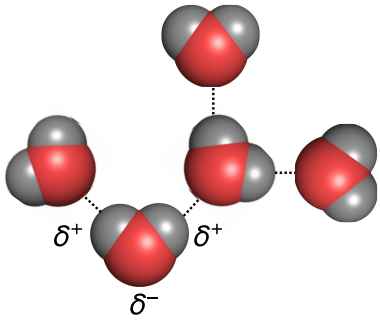
Hydrogen bonds form when a hydrogen atom is covalently bound to an electronegative atom (most often oxygen or nitrogen) and a second electronegative atom is in the vicinity. To a first approximation this is an interaction between charges, but there is also a small degree of covalent bonding, that is electron sharing, involved.

To form a hydrogen bond, the two molecules (or groups within a single molecule) must be arranged so that the two electronegative atoms are about 3–3.5 Å apart, as shown below:



The electronegative atom that is covalently bound to the hydrogen atom is referred to as the hydrogen-bond donor and the other electronegative atom is called the acceptor. In order to form a stable interaction, the three atoms must be approximately collinear with the angle indicated in the drawing no more than about 40°.

The oxygen atom of a water molecule can accept two hydrogen bonds and can act as a donor for two, thus enabling a water molecule to form up to 4 hydrogen bonds.

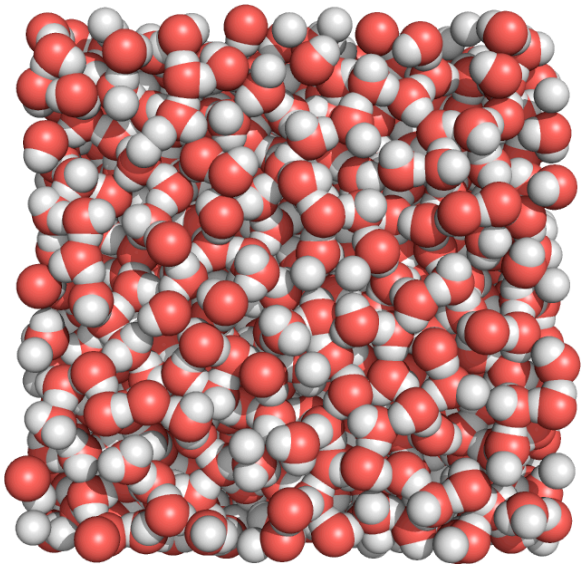


When water freezes it forms a lattice in which each molecule forms all four of the possible hydrogen bonds. The geometry is very similar to the lattice in a diamond, except that the bonds are much weaker.

In liquid water, each molecule forms, on average, three hydrogen bonds at any instant. Thus, only 1/4 of the hydrogen bonds break when ice melts. This is a major reason that the boiling temperature of liquid water is relatively high for a molecule of its size. (In general, the boiling points of liquids increase with the size of the molecules because they can form more extensive van der Waals interactions.)

The hydrogen bonds in liquid water rapidly break and reform constantly. At any instant, there is an extensive network of hydrogen bonds that can be traced from one side of a beaker to another (in principle), but this network is constantly being rearranged.

The picture below is a "frame" from a simulation of 1,000 water molecules, provided by Prof. Valeria Molinero of the University of Utah Chemistry Department:



II. Ionization

Another consequence of the uneven sharing of electrons in water is that the covalent hydrogen-oxygen bonds break rather easily, generating H^+ and OH^- ions. This is a reversible and very rapid process:

$H_2O \rightleftharpoons H^+ + OH^-$

A major reason that this (and similar) reactions occur to a significant degree in water is that the ionic species can interact favorably with the other water molecules. In fact, the H^+ ions are not really free. Instead, they interact with groups of water molecules through hydrogen bonds. The hydrogen ion in solution is often represented as H_3O^+ to indicate that the ion is in close association with water molecules, but this, too, is a simplification. In non-polar solvents, there is essentially no tendency for molecules to ionize, because the solvent does not interact favorably with charged species.

Like any other reversible chemical reaction, there is an equilibrium constant for the dissociation of water:

$$K = \frac{[H^+][OH^-]}{[H_2O]} = 1.8 \times 10^{-16}\,\mathrm{M}$$

Because only a tiny fraction of the water ionizes, the concentration of neutral water is essentially constant, and the usual representation of the equilibrium constant ignores the water:

$$K_{\mathrm{wat}} = [H^+][OH^-] = 10^{-14}\,\mathrm{M}^2$$

It is common to write the concentrations of the H^+ and OH^- ions in a logarithmic form:

$$\mathrm{pH} = -\log[H^+]$$

$$\mathrm{pOH} = -\log[OH^-]$$

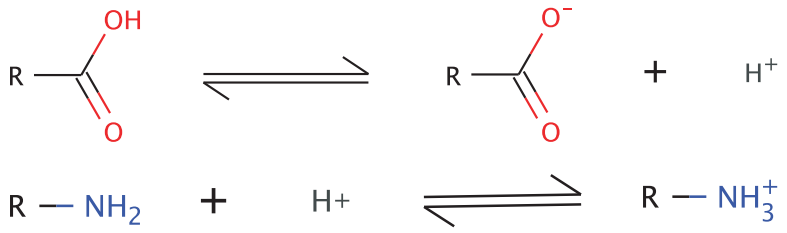
The equilibrium expression can then be written as:

$$\mathrm{pH} + \mathrm{pOH} = 14$$

This equation implies that if, for instance, we add H^+ ions to water (by adding an acid), the concentration of OH^- ions will go down. The reason for this is that some of the added H^+ ions combine with OH^- ions to form water. In general, pH is used much more commonly than pOH, but they both convey the same information, the balance between H^+ and OH^- ions.

The reason that chemists and biochemists give so much attention to pH is that other molecules can release or bind hydrogen ions as well, and the pH determines the balance of charged species.

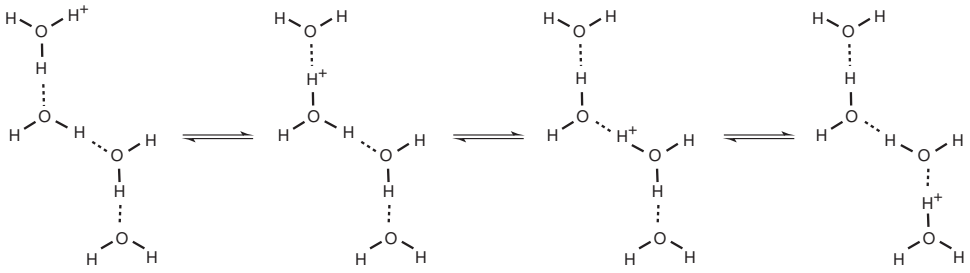
Two important examples of ionizing functional groups in organic molecules are the carboxyl groups and amino groups:



Groups that ionize do so for same basic reason as water does, an uneven distribution of electrons in covalent bonds. Different functional groups have different tendencies to release or take up H^+ ions. Ions are shuttled among different molecules in solution, including the water molecules. These reactions are typically very fast, on the order of microseconds. The exact balance between different charged forms of a molecule depends on the total concentration of free H^+ ions. Because a change in ionization results in a change in the electrical charge of a molecule, its chemical, structural and functional properties can be very sensitive to pH.

III. Dynamics of hydrogen ion diffusion.

Water, and molecules in it, form a highly dynamic solution, with the charges of molecules rapidly changing. Rates of exchange of H^+ ions from one molecule to another occur with times on the order of 1 μs, or less, allowing electric charge to be displaced in water very rapidly through relay processes. A mechanism for the rapid diffusion of hydrogen ions in water was was proposed by Theodor Grotthuss in 1806 and is illustrated below:



If you examine this diagram closely, you will see that the hydrogen ion that starts on the top-most water molecule doesn't really change position, but the electric charge moves through the rearrangement of hydrogen bonds. The basic idea of this mechanism is still thought to be correct, but details are still being studied and debated. This type of mechanism may also be important for the transport of hydrogen ions through some membrane channels.

IV. The hydrophobic effect

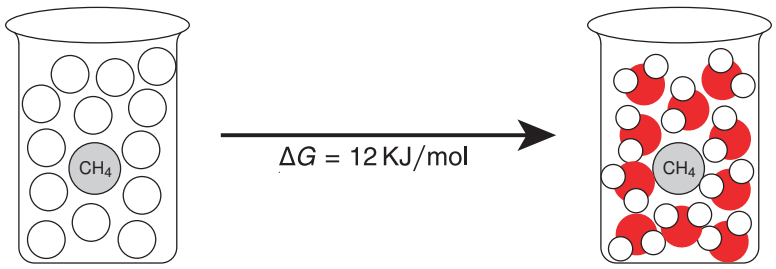
As mentioned earlier, one of the most important features of water is the fact that it doesn't mix well with non-polar molecules. This is a major driving force for the assembly of biological structures, because it leads to structures in which non-polar parts of molecules are sequestered away from water.

Why are non-polar molecules not very soluble in water? This is often referred to as the "hydrophobic effect", but, as we will see, non-polar molecules aren't really afraid of water.

Remember, water molecules *love* to form hydrogen bonds. What happens if a non-polar molecule does try to enter water? One thing it *doesn't* do is form hydrogen bonds with the water!

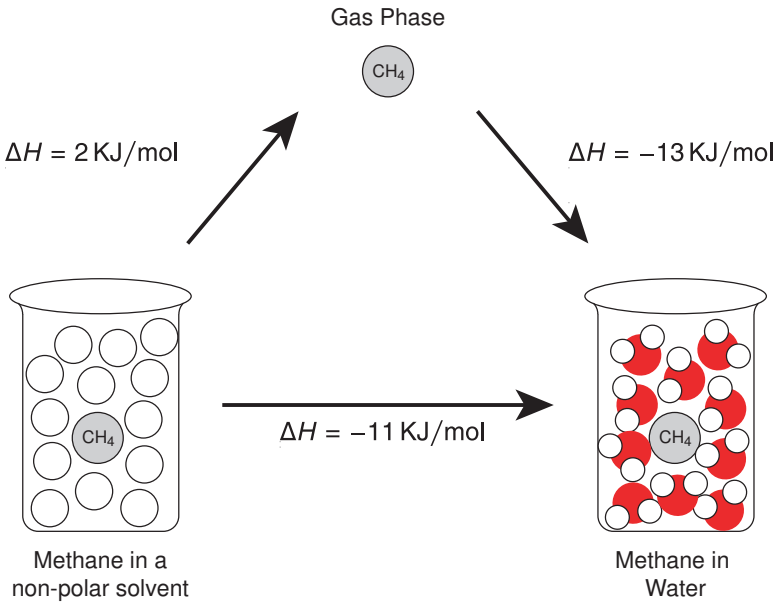
Does the non-polar molecule cause hydrogen bonds to break? It seems plausible. How can we find out?

Most of what we know about this phenomenon (or think we know) comes from thermodynamic measurements. The hydrophobic effect can be quantified by measuring the thermodynamics of transferring a non-polar molecule from a non-polar solvent to water. For instance, a molecule of methane from octanol to water:



In practice, the free-energy change is measured by determining the solubility of the molecule in each of the two solvents. As expected, the free energy of the process is positive, which simply means that oil and water don't mix.

We can also measure the enthalpy change for this process:



What is surprising is that ΔH is actually negative, which means that heat is released. This implies that, on average, there are more or stronger bonds in the aqueous solution with the non-polar molecule dissolved than when the non-polar molecule is in a non-polar solvent.

It's also possible to measure the enthalpy change for transfer of the non-polar molecule from each of the solutions to the gas phase. There is actually only a small positive ΔH for moving from the non-polar liquid to the gas phase, but a large negative ΔH for moving from the gas phase to water. Water and non-polar molecules actually interact quite strongly.

If ΔG is positive and ΔH is negative, ΔS must be negative for the transfer of the non-polar molecule to water. Somehow or other, the molecules become more ordered when a non-polar molecule is dissolved in water. Favorable processes for which ΔH is less than zero are often said to be *entropically driven*.

These observations lead to a model in which the water molecules rearrange themselves in some way around the non-polar molecule so that they lose entropy but actually form more or stronger hydrogen bonds. This seems rather counter intuitive, since it would seem easy for the water to just give up a few hydrogen bonds to accommodate the non-polar molecule. None the less, it seems that giving up some entropy is least bad way for water molecules to live with a non-polar molecule in their midst.

Further support for this model comes from another thermodynamic parameter, the heat capacity change at constant pressure, ΔC_{p}. This parameter is the derivative of ΔH with respect to temperature:

$$\Delta C_{\mathrm{p}} = \frac{d\Delta H}{dT}$$

ΔC_{p} can be measured by measuring ΔH as a function of temperature. For the transfer processes we are discussing, this gives a linear, or very nearly linear, plot and the slope is ΔC_{p}.

We can think about heat capacity as the amount of heat that is required to raise the temperature of a substance by 1°C. For either side of the reaction, we can write:

$$dH = C_p dT$$

Different substances have different heat capacities, because they have different ways of absorbing heat, including modes of motion and potentially breaking bonds.

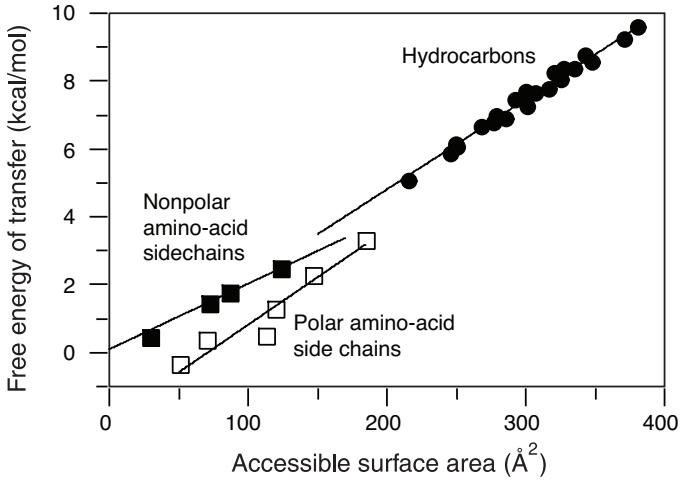
The heat capacity change for the transfer of a non-polar molecule to water is quite large and positive. This means that it takes more heat to raise the temperature of water when a non-polar molecule is present. This is consistent with the idea that the water forms more or stronger hydrogen bonds when the non-polar molecule is present, and these bonds break or weaken when the temperature is increased.

The positive heat capacity change also means that ΔH becomes less negative as the temperature increases. At higher temperatures, ΔH becomes positive, suggesting that introducing a non-polar molecule in water *does* lead to a net loss of hydrogen bonds.

This model is sometimes referred to as an "iceberg" model, but no one *really* understands it on a detailed structural basis.

"Hydrophobic effect" is a bad name, but it's somewhat better than "hydrophobic bonds" or "hydrophobic interactions", which are also used. The non-polar molecule doesn't fear water; it actually likes it pretty well! It's the water that has problems with its guest.

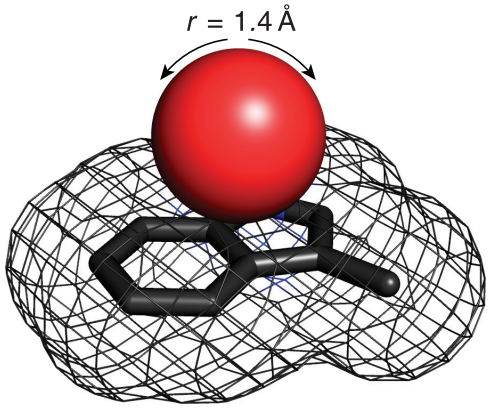
Another important observation about the hydrophobic effect is that the magnitude of the transfer free energy is proportional to the size of the non-polar molecule. This was pointed out in a classic review article by Fred Richards in 1977, which included a graph like the one below[1]:



In this graph, the filled circles represent hydrocarbons, the filled squares represent non-polar amino-acid side chains and the open squares represent polar amino-acid side chains.

[1]Figure adapted from Richards, F. M. (1977). Areas, Volumes, Packing and Protein Structure. *Annu. Rev. Biophys. Bioeng.*, 6, 151–176. http://dx.doi.org/10.1146/annurev.bb.06.060177.001055

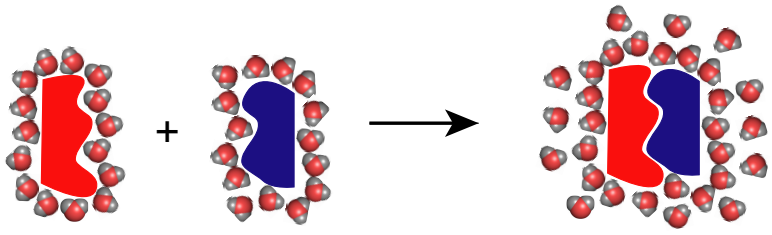
The quantity plotted on the horizontal axis is called the *accessible surface area* (ASA) and has units of $Å^2$. The process for calculating the accessible surface area for a known structure can be visualized as rolling a sphere representing a solvent molecule over the surface of the molecule and recording path followed by the point of the sphere closest to the molecule, as diagrammed in the figure below:



In this illustration, the molecule of interest is the side chain of a tryptophan residue and probe sphere has a radius of 1.4 Å, representing a water molecule. The surface area is represented by the wire cage surrounding the stick representation of the tryptophan side chain. For these calculations, the hydrogen atoms of both the molecule of interest and the water molecule are often ignored because they contribute relatively little to volume and surface area. Other measures of molecular size, such as volume, do not correlate with free energy of transfer nearly as well as surface area does. This is consistent with the idea that it is the interaction of the non-polar molecule with water and the interference with hydrogen bonding that define the hydrophobic effect.

Another important point that is illustrated in the graph on the previous page is that the amino-acid side chains that include polar groups lie on a distinct line from the one for those that contain only carbon and hydrogen atoms. The free energy of transfer for the side chains with polar groups are systematically less positive than the hydrocarbons. This reflects the more favorable interactions of non=polar groups with water.

The relationship between the hydrophobic effect and accessible surface area helps to explain how the hydrophobic effect can favor the assembly of biological structures, as illustrated in the cartoon below:



The cartoon represents two molecules with non-polar surfaces. As discussed above, these surfaces favor the ordering of water molecules in their vicinities, a process that

leads to the reduction of solvent entropy. When the two molecules interact via their complementary surfaces, the water that is closely associated with those surfaces is displaced, leading to an increase in entropy. Because many water molecules may be released by the association of two larger molecules, the increase in the solvent entropy can be significantly greater than the decrease in the associating molecules.

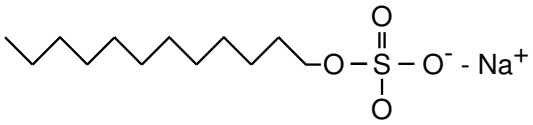
Other molecules or ions that are closely associated with the surfaces of relatively large molecules can also contribute to a favorable association reactions. An important example is the association of nucleic acids with proteins. Because they have a high density of negative charge at their surfaces, nucleic acids in solution are frequently associated with divalent cations, such as Mg^{2+}. The proteins that bind nucleic acids often contain positively charged side chains that interact with the negative charges of the nucleic acids and displace the cations. This results in an increase in entropy of the cations and contributes to a favorable free energy change for interaction between the protein and nucleic acid.

6.2 Lipid Bilayers and Membranes

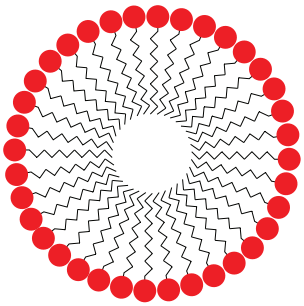
Some of the most important structures in living organisms are the membranes that separate the contents of cells from the extracellular environment and separate different intracellular compartments. The structures and properties of these membranes are largely determined by the hydrophobic effect discussed in the previous section.

I. Amphiphilic molecules, micelles and bilayers

When non-polar molecules exceed their solubility in water they simply form a separate phase, like in salad dressing. But, there are molecules that contain both polar and non-polar parts, and these molecules, called amphiphiles, can form more specific structures. An example of this type of molecule is a detergent. The structure of a typical detergent, sodium dodecyl sulfate (SDS), is shown below:



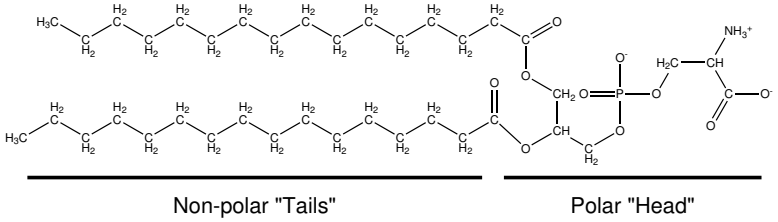
SDS is commonly used in biochemistry experiments, especially electrophoresis, and is also a common ingredient of household cleaning products, such as shampoos. The two parts of the molecule are a hydrocarbon chain (commonly referred to as the tail) with 12 carbon atoms and a sulfate group, a highly charged ion (the polar head). When molecules like this exceed their solubility, they assemble into a structures that sequester the non-polar part away from water while keeping the charged polar group exposed. These structures are called micelles and are roughly spherical, as illustrated in cross-section below:



These structures have a modest degree of specificity. For a given molecule, there will be a characteristic preferred size of micelle that optimizes the packing of the hydrophobic tails, while keeping the polar head-groups solvated with water. Typical micelle diameters lie in the range of about 3 to 50 nm.

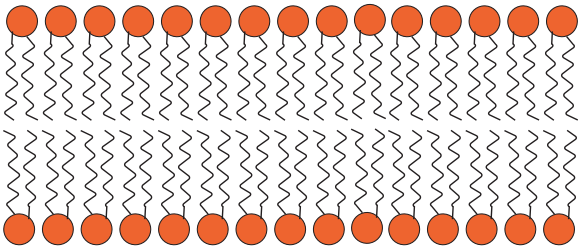
Molecules like this help us clean things up by solubilizing greasy molecules in the middle of the micelles.

The major amphiphilic molecules in biology have a slightly more complicated structure. These molecules typically have two hydrocarbon tails that are linked together by a glycerol molecule, a three carbon sugar, which is also linked to a polar phosphate group:



These molecules are called phospholipids, and there is wide variety in their structures. Different phospholipids have different hydrocarbon tails and different chemical groups attached to the phosphate.

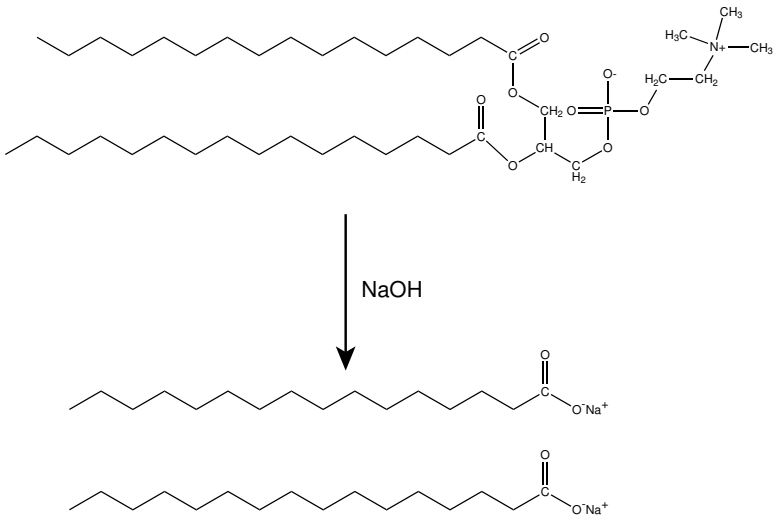
Like detergents, phospholipids form structures in water, with the non-polar groups sequestered and the polar groups interacting with water. But, rather than forming spheres, these molecules form extended (nearly) flat structures with two layers, as diagrammed below:



These structures form the membranes of cells, creating compartments with different chemical compositions.

The very different structures formed by detergents and phospholipids is largely due to the difference in the shapes of the molecules. Detergents tend to have a wedge shape that leads to sharp curvature, while phospholipids, because they contain two lipid tails, are more rectangular and form flat structures.

Treating phospholipids with strong bases (lye) hydrolyzes the ester bonds between fatty acids and glycerol, resulting in salts of the fatty acids, which are soaps.

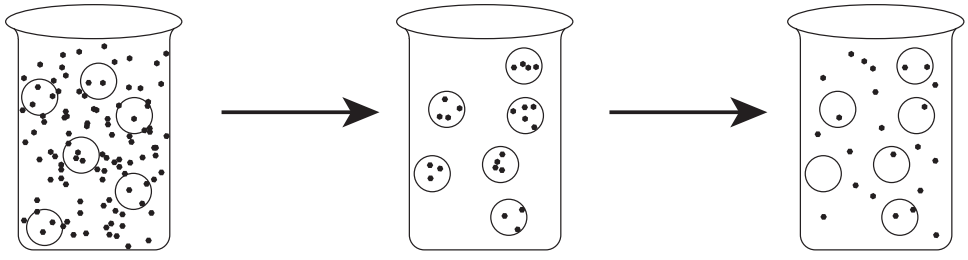


Like the detergents, soaps have a bit of a wedge-like structure, leading them to form micelles, rather than bilayers, so that they act in the same way as synthetic detergents. This reaction is probably one of the earliest examples of practical chemistry, dating back to the Roman Empire.

II. Permeability of bilayers

The major role of lipid bilayers is to enclose cells and form compartments within them with distinct chemical compositions. So, an important property is their permeability to different molecules.

Permeability can be measured by forming vesicles in the presence of specific molecules, separating the vesicles from free molecules, and then measuring the rates at which the molecules diffuse out of the vesicles, as illustrated below:



The rate at which the concentrations equilibrate is determined by Fick's first law:

$$J = -D\frac{dC}{dx}$$

where J is the flux (with units of $\text{mol} \cdot \text{s}^{-1}\text{m}^{-2}$); D is the diffusion coefficient (with units of m^2s^{-1}) and dC/dx is the concentration gradient across the membrane.

In this case, the area is the total area defined by the surface of the vesicle. The concentration gradient is the difference in the concentration divided by the thickness of the bilayer. The diffusion coefficient is a property of the molecule and the bilayer, and will generally be very different (smaller) than the diffusion coefficient of the same molecule in water. A typical vesicle might have a diameter of about 50 nm and typical bilayers have thicknesses of 3–4 nm. Thickness is a bit ambiguous, since it depends on how much of the polar head group is included, and different lipids have different lengths. In fact, the diffusion coefficient is not so easily defined because of the heterogeneous structure that the molecule has to cross.

Instead of using diffusion coefficients, the common practice for describing diffusion across bilayers is to introduce a *permeability coefficient* defined so that:

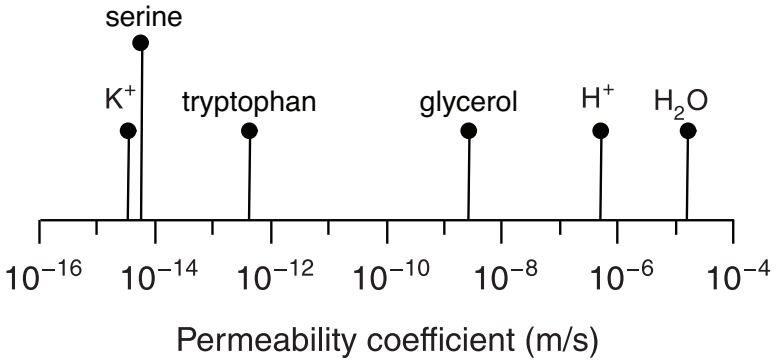
$$J = -D\frac{dC}{dx} = -P\Delta C$$

where ΔC is the difference in concentration across the membrane. In effect, the permeability coefficient combines the diffusion coefficient and the thickness of the membrane, so that:

$$P = \frac{D}{\Delta x}$$

where Δx is the thickness of the membrane, which is about 4 nm. The units for the permeability coefficient are m/s.

Some experimental values for the permeability coefficient are plotted below on a logarithmic scale[2]:



[2]Values in the figure are from: Chakrabarti, A. C. & Deamer, D. W. (1992). Permeability of lipid bilayers to amino acids and phosphate. *Biochim. Biophys. Acta - Biomembranes*, 111, 171–177. https://doi.org/10.1016/0005-2736(92)90308-9
and
Paula, S., Volkov, A. G., Van Hoek, A. N., Haines, T. H. & Deamer, D. W. (1996). Permeation of protons, potassium ions and small polar molecules through phospholipid bilayers as a function of membrane thickness. *Biophys. J.*, 70, 339–348. https://doi.org/10.1016/S0006-3495(96)79575-9

Some important points to note from these values are:

- The permeabilities cover 9 orders of magnitude.
- Ions have extremely low permeabilities, but H^+ is a notable exception.
- Polar molecules also have low permeabilities.
- Water has a quite high permeability.

We can compare the permeability coefficients to the diffusion coefficients for small molecules in water. The two parameters are related to one another according to:

$$D = P\Delta x$$

For the amino acid serine, assuming $\Delta x = 4\,\text{nm}$:

$$\begin{aligned} D &= 5 \times 10^{-15}\,\text{m} \cdot \text{s}^{-1} \times 4 \times 10^{-9}\,\text{m} \\ &= 2 \times 10^{-23}\,\text{m}^2\text{s}^{-1} \end{aligned}$$

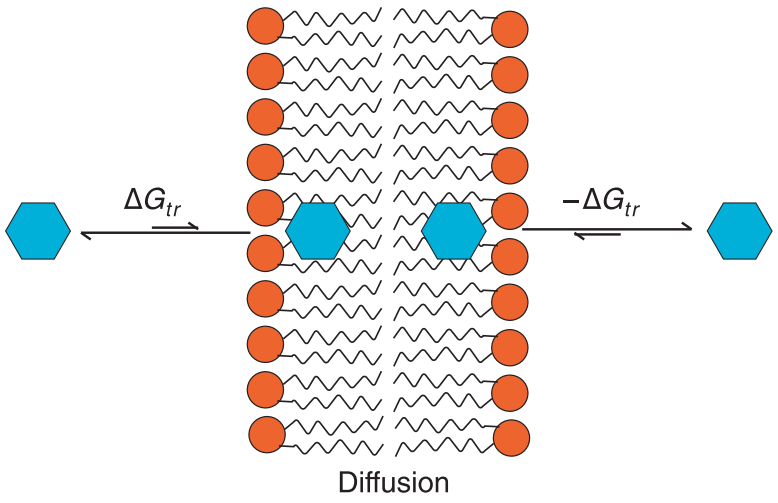
This is *much* less than the diffusion coefficient of a small molecule in water, which is about $10^{-10}\,\text{m}^2\text{s}^{-1}$.

For water:

$$\begin{aligned} D &= 1.6 \times 10^{-5}\,\text{m} \cdot \text{s}^{-1} \times 4 \times 10^{-9}\,\text{m} \\ &= 6.4 \times 10^{-14}\,\text{m}^2\text{s}^{-1} \end{aligned}$$

This is still quite small.

For many molecules, the observed permeability coefficients can be accounted for by a model in which there is an equilibrium between the molecule in water and in the lipid bilayer, coupled to diffusion within the bilayer and then rapid escape from the bilayer back to the water phase:



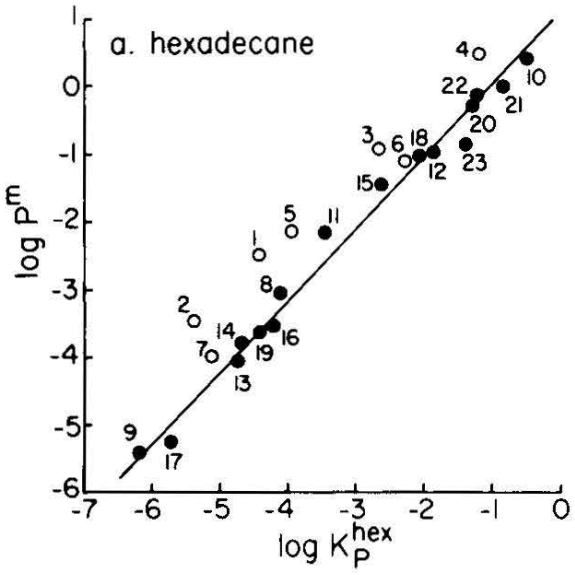
$\Delta G_{\rm tr}$ is the free energy of transfer from water to a non-polar environment and is very unfavorable for a polar molecules . This model, referred to as the *solubility-diffusion model*, can be used to predict the permeabilities of molecules from the values of ΔG_{tr} and the diffusion coefficients in oils. This model assumes that the overall rate of crossing the bilayer is determined by the fraction of molecules that are present in the bilayer, relative to the water phase on each side, and the diffusion coefficient in the non-polar part of the bilayer. The fraction of molecules in the bilayer is equal to the equilibrium constant for transfer from water to the non-polar environment, $K_{\rm tr}$, provided that this equilibrium constant is much less than 1. The effective diffusion coefficient is then given by:

$$D_{\rm eff} = K_{\rm tr} \times D$$

and the permeability coefficient can be expressed as:

$$P = P = \frac{D_{\rm eff}}{\Delta x} K_{\rm tr} D/\Delta x$$

This model predicts that the permeability coefficients for different molecules should be correlated with their relative solubilities in non-polar liquids. This prediction has been borne out for many, but not all, molecules that have been examined, as illustrated in the figure below:[3]



This correlation provides a strong argument that the solubility-diffusion model is a good description for the permeabilities of small molecules, both polar and non-polar, and most small ions. The largest discrepancies appear to be for water and H^+ ions, which have anomolously high permeability coefficients.

[3]Figure from Walter, A. & Gutknecht, J. (1986). Permeability of small nonelectrolytes through lipid bilayer membranes. *J. Membrane Biol.*, 90, 207–217. http://dx.doi.org/10.1007/BF01870127

An alternative model would be that holes transiently form in the bilayer and allow molecules to pass through. However, this model would suggest that the permeability coefficients would be relatively independent of the polarity of the molecules, which is clearly not the case. On the other hand, the permeability data for some species, especially H^+ ions and water, do not fit the solubility-diffusion model, indicating that other factors may play a role. It has been suggested that chains of hydrogen-bonded water molecules may cross the bilayer and allow the net movement of both water molecules and H^+ ions, the latter by a Grotthuss mechanism (page 170). In addition, transient defects may form and contribute to permeability, but this appears to be a relatively small factor for most molecules.

One important application of kind of data is in the design of pharmaceuticals, since drug molecules generally have to cross multiple membranes to reach their targets. It may also be important to keep the drug from crossing other membranes. This can make or break a potential drug.

We can apply the permeability coefficients to estimate the rate of molecules entering or leaving cells by passive diffusion across the bilayer (in the absence of transport by specific membrane proteins). Suppose that we have a cell with a diameter of $20\,\mu$m, and it contains no glucose, but is surrounded by a solution that is 0.1 M in glucose. How rapidly will glucose, with a permeability coefficient of about 5×10^{-10} m/s, enter this cell? From the permeability coefficient and the concentration difference, we can calculate the flux, J, in units of $\text{mol} \cdot \text{s}^{-1}\text{m}^{-2}$:

$$\begin{aligned} J &= P\Delta C \\ &= 5 \times 10^{-10}\,\text{m/s} \times 0.1\,\text{M} \\ &= 5 \times 10^{-11}\,\text{mol} \cdot \text{L}^{-1}\text{m} \cdot \text{s}^{-1} \\ &= 5 \times 10^{-8}\,\text{mol} \cdot \text{s}^{-1}\text{m}^{-2} \end{aligned}$$

Next, we calculate the surface area of the cell:

$$\begin{aligned} A &= 4\pi r^2 \\ &= 4\pi(10^{-5}\,\text{m})^2 \\ &\approx 10^{-9}\,\text{m}^2 \end{aligned}$$

The total flow into the cell is the flux multiplied by the surface area:

$$\begin{aligned} \text{flow} &= 5 \times 10^{-8}\,\text{mol} \cdot \text{s}^{-1}\text{m}^{-2} \times 10^{-9}\,\text{m}^2 \\ &= 5 \times 10^{-17}\,\text{mol} \cdot \text{s}^{-1} \end{aligned}$$

That's not a lot of moles per second, but it is about 30 million molecules per second.

How rapidly would the intracellular concentration change? We need to calculate the volume:

$$\begin{aligned} V &= \frac{4}{3}\pi r^3 \\ &= \frac{4}{3}\pi(10^{-5}\,\mathrm{m})^3 \\ &\approx 4\times 10^{-15}\,\mathrm{m}^3 \times \frac{10^3\,\mathrm{L}}{1\,\mathrm{m}^3} \\ &\approx 4\times 10^{-12}\,\mathrm{L} \end{aligned}$$

So, the rate of change in concentration is:

$$\begin{aligned} \frac{dC}{dT} &= \frac{5\times 10^{-17}\,\mathrm{mol}\cdot\mathrm{s}^{-1}}{4\times 10^{-12}\,\mathrm{L}} \\ &\approx 10^{-5}\,\mathrm{mol}\cdot\mathrm{L}^{-1}\mathrm{s}^{-1} \end{aligned}$$

It will take a long time for the intracellular concentration to equilibrate with the extracellular environment.

III. Primitive membranes

The impermeability of bilayers raises an interesting question with regard to the origins of life, one of the great intellectual challenges. The basic problem is that modern life forms appear to be so perfect and complicated that it is hard to imagine how any part of it could have evolved by itself. Modern organisms have to:

- Collect nutrients
- Convert nutrients into useable forms of energy
- Build complicated macromolecules, including enzymes and genetic material
- Create compartments bounded by membranes
- Reproduce themselves

For a long time, a big issue was whether proteins or nucleic acids came first. Proteins are needed for enzymes, but DNA and RNA are needed to encode proteins. In the 1980s, however, it was discovered that some RNA molecules have catalytic activities. It's now widely believed that the earliest biological macromolecules were RNA molecules that had very limited ability to catalyze their own replication.

Membranes also present a problem. A key event in the evolution of early cells must have been the formation of membranes, so that nutrient molecules could be sequestered and not shared with competing cells or molecules. Modern lipid bilayers are extremely impermeant to polar molecules and ions. Translocation of molecules across membranes depends on protein molecules embedded in the bilayer, which allows for the transport to be controlled. But, how could primordial membranes have worked?

It turns out that some fatty acids can form bilayers that are much more permeant to polar and even charged species. A current idea is that vesicles of this type formed and trapped RNA, or related molecules, that could polymerize. Precursors to polymers could diffuse across the membranes, but when they were incorporated in polymers they were trapped. As molecules inside the vesicles got larger and more numerous, vesicles were forced to expand and eventually divide.

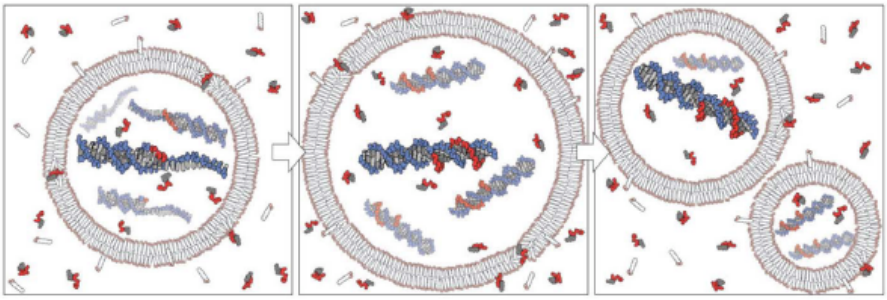


Figure from:

Mansy, S. S., Schrum, J. P., Krishnaurthy, M., Tobé, S., Treco, D. A. & Szostak, J. W. (2008). Template-directed synthesis of a genetic polymer in a model protocell. *Nature*, 454, 122–125.
http://dx.doi.org/10.1038/nature07018

Another, related, reference:

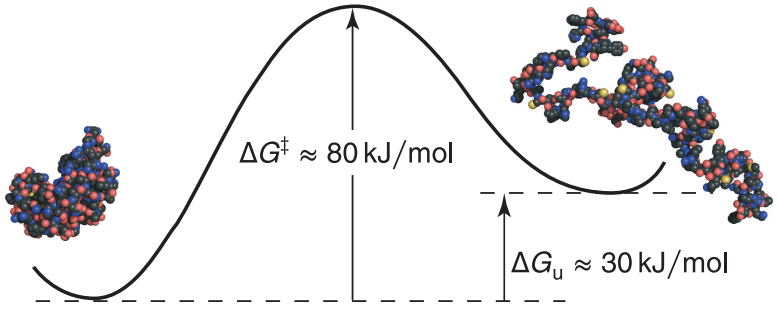
Monnard, P.-A., Luptak, A. & Deamer, D. W. (2007). Models of primitive cellular life: polymerases, and templates in liposomes. *Phil. Trans. Royal. Soc. Lond. B*, 362, 1741–1750.
http://dx.doi.org/10.1098/rstb.2007.2066

6.3 Protein Folding and Unfolding

One of the most important examples of biological structure formation is the folding of polypeptide chains into stable three dimensional structures. This process occurs largely after proteins are synthesized by ribosomes, although the first segments of a protein that are synthesized may begin to fold while the rest of the chain is still being synthesized. For most proteins, the folded conformation is required for function, and this conformation is specified by the amino acid sequence. As a consequence the folding process represents a point in which one-dimensional information encoded in the genome is expressed as a three-dimensional structure. With many proteins, it is possible to unfold the folded, or native, structure by altering the solution conditions and then reform the structure by again changing the conditions. This provides the means of studying the thermodynamics and mechanisms of protein folding. Because of its central importance in biology, this process has been studied extensively, both experimentally and by theoretical and computational methods.

I. Native and unfolded protein states

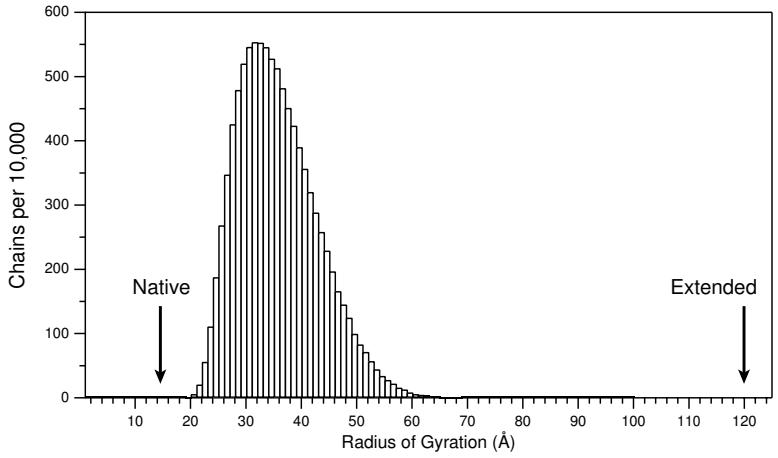
Although there is tremendous variety in the size and architecture of protein molecules, the great majority of the experimental work that has been done to date has focused on relatively small proteins, containing about 50–200 amino-acid residues, and our discussion will be largely limited to this class of proteins. Proteins of this size typically fold into a single compact structure, termed a domain, whereas larger proteins are often formed of multiple domains. One important characteristic feature of single-domain proteins is that the unfolding and folding processes is highly cooperative, meaning that partially folded structures are significantly less stable than either the native state or completely disordered molecules. In addition, the kinetics of folding and unfolding can often be described by models involving a single major transition state or energy barrier, as diagrammed below:



The structures in the diagram represent ribonuclease A (RNAse A), which contains 124 amino-acid residues, in its folded conformation (on the left) and a disordered conformation. As indicated in the diagram, the native state is more stable than the unfolded by a few tens of kJ/mol. The activation free energy indicated in the diagram, 80 kJ/mol corresponds to a rate constant for folding of about $0.1\,\mathrm{s}^{-1}$, or a half time of about 6 s. This value falls within the quite wide range of observed folding times, from minutes to microseconds.

Whereas the folded conformation is a relatively unique structure, the unfolded state is a broad distribution of rapidly interconverting conformations, contrary to the impression that may be conveyed by the figure above. The conformation shown in the diagram is one of many generated in a computational simulation designed to explore the properties of unfolded proteins[4] This simulation generated approximately 200,000 conformations, and the figure below shows their distribution with respect to overall dimension:

[4]Goldenberg, D. P. (2003). Computational simulation of the statistical properties of unfolded proteins. *J. Mol. Biol.*, 326, 1615–1633. http://dx.doi.org/10.1016/S0022-2836(03)00033-0



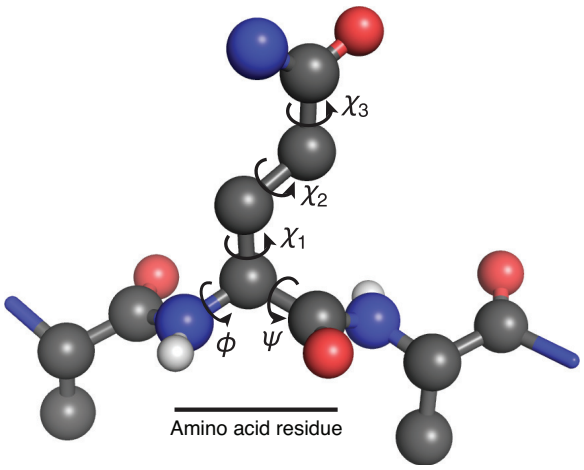
The radius of gyration, as plotted on the horizontal axis, of a polypeptide chain is the root-mean-square distance between the center of mass and each non-hydrogen atom in the molecule[5]. It is thus a useful measure of the overall size of a molecule in a particular conformation. For reference, the figure also indicates the radii of gyration of native RNAse A and a fully extended conformation. Although the distribution includes conformations that are nearly as compact as the folded state, there are none that approach the radius of gyration of a fully extended chain. One consequence of the very broad distribution of unfolded states is that it has much more entropy than the native conformation, as discussed further below.

II. Entropy of the unfolded state

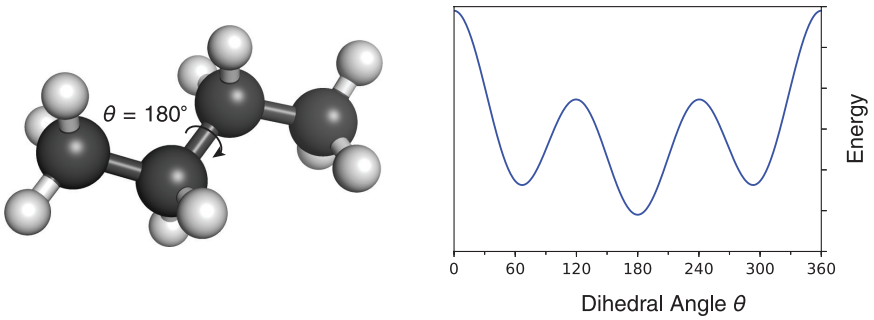
The large difference in entropy between the native and unfolded states of a protein is expected to disfavor the native state. It is thus useful to try to estimate the magnitude of this difference. If we can somehow count the number of microstates making up the two states, we can use the Boltzmann equation to calculate the entropy difference. To do this rigorously would require consideration of all of the possible alternate conformations in the two states, which is a formidable challenge. We can, however, make some simplifying assumptions that allow for a very rough approximation. The key assumptions are that the folded protein has a single, unique conformation and that the individual amino-acid residues are able to take on multiple, independent conformations in the unfolded state.

The conformations accessible to an individual reside can be defined by the dihedral angles that describe the rotations about single covalent bonds, as illustrated below for a glutamine residue:

[5]This definition is somewhat simplified from the more general definition of the radius of gyration, but it is a good approximation for an object in which all of the elements have approximately the same mass.



For each residue (other than prolines), there two rotatable bonds in the polypeptide backbone, labeled ϕ and ψ, and additional rotatable bonds in the side chain, which are labeled χ_1, χ_2 and so on, depending on the particular residue type. These bonds can undergo rotation, to varying degrees, in both the folded and unfolded states. However, these rotations are associated with changes in the potential energy of the molecule, due to steric and other interactions among the atoms. These effect create a pattern of valleys and peaks in the energy as a bond is rotated, as illustrated in the diagram below for the simple case of a propane molecule:



As the central bond is rotated, the lowest energy is found when the methyl groups on the two sides of the bond are pointed in opposite directions, and the highest energy is observed when the methyl groups are on the same side of the central bond. There are two other minima, where the bond is rotated by 120° from the lowest energy position. Bonds in different contexts have different rotational energy profiles, but the patterns are similar.

Within a folded protein, most of the rotatable bonds are restricted a single minimum, but the dihedral angles fluctuate about those minima. In the unfolded state, the dihedral angles can sample all of the minima and fluctuate within these minima. Thus, we can think of the reduction in conformational entropy as a reduction in the number of accessible minima, and we can define the microstates in term of distinct conformations with the dihedrals in specified minima. This approximation is referred to as the *rotational isomeric state* model.

For the native protein we assume that there is a single microstate, recognizing that this microstate includes all of the fluctuations about the dihedral minima, as well as vibrational motions. For the unfolded state, we assume that the number of microstates of each residue is approximately 10-times the number in the native state. Therefore the ratio of the number of microstates for a single residue in the two states is:

$$\frac{\Omega_{\mathrm{U}}}{\Omega_{\mathrm{N}}} = 10$$

For two residues, if the accessible conformations are independent, the ratio is 10^2; for three residues, the ratio is 10^3, and so on. For an n-residue protein, therefore, the ratio is:

$$\frac{\Omega_{\mathrm{U}}}{\Omega_{\mathrm{N}}} = 10^n$$

The entropy change is then:

$$\Delta S_{\mathrm{conf}} = k \ln \frac{\Omega_{\mathrm{U}}}{\Omega_{\mathrm{N}}} = k \ln 10^n$$

This quantity is designated ΔS_{conf} to emphasize that it reflects only the change in polypeptide conformation for unfolding and that there are other contributions to the overall entropy change for the process. For a protein containing 100 amino-acid residues, the conformational entropy change is:

$$\Delta S_{\mathrm{conf}} = k \ln 10^{100} = 3.3 \times 10^{-21}\,\mathrm{J/K}$$

On a molar basis:

$$\Delta S_{\mathrm{conf}} = R \ln 10^{100} = 1900\,\mathrm{J/(K \cdot mol)}$$

and the free energy contribution to unfolding at 300 K is:

$$-T\Delta S_{\mathrm{conf}} = -570\,\mathrm{kJ/mol}$$

Note that this is a large factor favoring unfolding. It is approximately ten-fold greater than the net free energy change that favors folding for a typical protein of this size.

One important, and unrealistic, assumption that went into this calculation is that the conformations of individual amino-acid residues are independent of one another in an unfolded protein. Manipulating a physical model of a peptide containing only a few residues will demonstrate that many combinations of dihedral angles will lead to steric clashes among the atoms. Suppose that we assume that only one in 10^{10} (one in 10 billion) of the conformations that we assumed in our calculations is actually possible. For the 100-residue protein, the ratio of microstates in the unfolded and native states is then reduced to:

$$\frac{\Omega_{\mathrm{U}}}{\Omega_{\mathrm{N}}} = 10^{100} \div 10^{10} = 10^{90}$$

The corresponding molar entropy and free energy changes are:

$$\Delta S_{\text{conf}} = R \ln 10^{90} = 1700\,\text{J/mo}$$
$$-T\Delta S_{\text{conf}} = 520\,\text{kJ/mol}$$

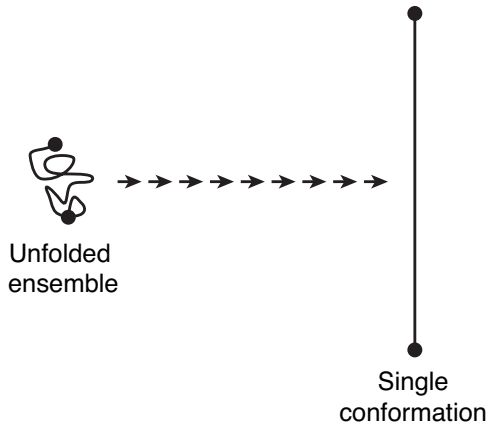
Thus, even reducing the initial estimate of the number of possible conformations by a factor of 10 billion does not alter the general conclusion from this calculation.

It is also possible to estimate the value of ΔS_{conf} from experimental measurements. Recall, from Chapter 6, the general definition of the entropy change (for the system) for a process at constant temperature:

$$\Delta S_{\text{sys}} = \frac{q_{\text{rev}}}{T}$$

where q_{rev} is the heat absorbed by the system during the reversible change from one state to another. If the process is carried out at constant temperature, and there are no changes in the potential energies of the molecules making up the system, $\Delta E = 0$ and $q_{\text{rev}} = -w_{\text{rev}}$, where w_{rev} is the work done on the system during the process. If we could somehow measure the work required to convert the very broad ensemble of unfolded conformations (very slowly) to a single conformation, we could determine the associated entropy change.

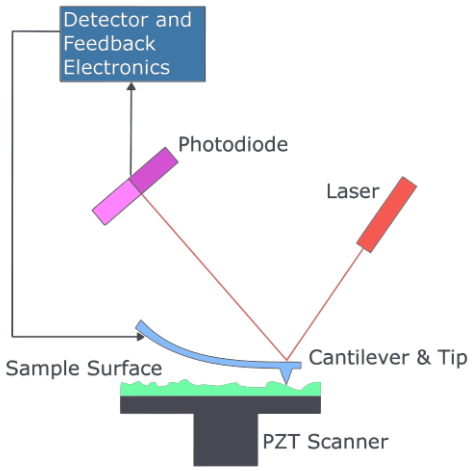
It might seem that the process to examine in this way would be the conversion of the unfolded state to the native state. But, this is problematic because folding is associated with the formation and rearrangement of numerous non-covalent interactions that alter the potential energy of the molecule, so that ΔE is not equal to zero. On the other hand, we can consider the conversion of the unfolded ensemble to a fully extended conformation, as diagrammed below:



Because the intramolecular interactions in the folded state are largely broken in the unfolded state, there should not be significant changes in the interactions of the polypeptide chain with either itself or the solvent when the molecules is stretched to its maximum extension. Thus, in terms of conformational entropy, the fully extended chain is equivalent to the fully folded chain! A caveat to this assumption is that the side

chains of many residues will be more restricted in the native state than in the fully extended conformation. Thus, the decrease in conformational entropy for stretching the unfolded ensemble to a fully extended conformation is expected to be somewhat less than that for folding.

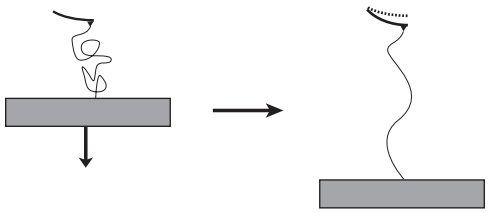
To actually measure the work for stretching a single protein requires very sensitive instrumentation, and two types of instrument have been employed in this kind of experiment, optical tweezers and atomic-force microscopes (AFM). Here we will focus on the AFM, a version of which is diagrammed below[6]:



This instrument was invented primarily for recording very high resolution images of surfaces. The essential elements are a very fine probe, with a tip diameter of a few nm or less, and a sample stage that can be moved with nanometer precision. The probe is attached to a flexible cantilever and brought into contact with the sample surface mounted on the stage. As the surface is moved in two directions, the probe tip follows the surface, and the cantilever bends slightly to accommodate this motion. The cantilever contains a reflective surface, and the light from a stationary laser is reflected from it. As the sample is scanned, the direction of the reflected light changes in response to the changes in vertical position of the tip. The fluctuations are recorded and converted into a record of the height of the surface as a function of position, thus generating an image. The tip can also be used to manipulate or move individual molecules, and even atoms. The key to this precision is the ability to move the stage in tiny, reproducible steps. This is made possible by the use of piezoelectric crystals that undergo small size changes in response to electrical voltage changes.

An AFM can also be used to measure forces generated by the motion of the stage. For this purpose, the cantilever, which acts as a spring, is calibrated so that the displacement of the tip can be converted into a force value. An arrangement for stretching a polypeptide chain is shown in the diagram below:

[6]Figure from https://en.wikipedia.org/wiki/Atomic-force_microscopy



Using some combination of genetic engineering and protein chemistry, one end of the chain is attached to the probe tip and the other to the movable stage. The stage is then moved downward, which reduces the entropy of the chain and creates a downward force on the tip. In response, the cantilever bends and, in doing so, exerts an opposite force, which increases as the stretching proceeds. By carrying out this process very slowly, to approach the ideal of reversibility, and continuously recording the force of the cantilever, the total work can be determined as:

$$w = \int F dx$$

Actually carrying out this kind of experiment, and properly analyzing the data is very challenging. None the less, the measurement has been performed for a variety of proteins and the quantitative results are quite consistent with the estimates of ΔS_{conf} based on counting rotational isomers[7].

To summarize, these experiments and calculations based on the rotational isomeric state model indicate that a reasonable estimate for the change in conformational entropy for a polypeptide of n residues is on the order of:

$$\Delta S_{\text{conf}} = k \ln 10^n$$

For a 100-residue protein, ΔS_{conf} is calculated from this relationship to be $1900\,\text{J}/(\text{K}\cdot\text{mol}$. This factor favors unfolding, by $570\,\text{kJ/mol}$ at $300\,\text{K}$. In the next section, we consider the factors that overcome this entropy penalty to make the folded structures of proteins stable under physiological conditions.

III. Protein-stabilizing factors

The thermodynamics of unfolding have been studied experimentally for a large number of single-domain proteins, most with chain lengths ranging from about 50 to 200 amino acid residues. Although the values of the parameters vary substantially, the values below, for hen egg-white lysozyme at 25°C[8], are typical for a protein in this size class:

[7]Thompson, J. B., Hansma, H. G., Hansma, P. K. & Plaxco, K. W. (2002). The backbone conformational entropy of protein folding: Experimental measures from atomic force microscopy. *J. Mol. Biol.*, 322, 645–652. http://dx.doi.org/10.1016/S0022-2836(02)00801-X

[8]Baldwin, R. L. (1986). Temperature dependence of the hydrophobic interaction in protein folding. *Proc. Natl. Acad. Sci., USA*, 83, 8069–8072. http://dx.doi.org/10.1073/pnas.83.21.8069

ΔG_{u}	$60.7\,\mathrm{kJ/mol}$
ΔH_{u}	$236\,\mathrm{kJ/mol}$
ΔS_{u}	$586\,\mathrm{J/(K \cdot mol)}$
-TΔS_{u}	$175\,\mathrm{kJ/mol}$

Hen lysozyme contains 129 amino-acid residues, leading to a predicted value of ΔS_{conf} of about $2500\,\mathrm{J/(K \cdot mol)}$, based on the assumptions introduced in the previous sections. Note that this value of ΔS_{conf} is approximately four-fold greater than the observed entropy change for unfolding. Further, the magnitude of $-T\Delta S_{\mathrm{conf}}$, $-740\,\mathrm{kJ/mol}$, is about ten-fold greater than that of ΔG_{u}, but of the opposite sign, representing a large factor favoring unfolding. We thus need to account for two apparent discrepancies associated with the large calculated value for ΔS_{conf}:

- The conformational entropy change for unfolding is much larger than the observed entropy change for the overall unfolding process.
- The free energy change associated with the conformational entropy change is far larger than the free energy change for unfolding and greatly favors the unfolded protein under conditions where the folded state is stable.

There must be at least one other factor that contributes a large negative change in entropy upon unfolding and a positive contribution to the free energy change for unfolding. The most likely explanation is the transfer of non-polar parts of the protein, which are buried in the native state, to the water solvent in the unfolded state. Recall from the discussion of the hydrophobic effect (pages 171–175) that the transfer of non-polar molecules from a non-polar liquid to water is associated with both a positive free energy change and a negative entropy change. Although it is not completely understood, it is generally believed that the decrease in entropy is due to an increase in order of water molecules directly surrounding the non-polar atoms, sometimes referred to as an iceberg effect.

As discussed earlier, the magnitude of the positive free energy change for transfer of a non-polar molecule to water is closely correlated with the accessible surface area (ASA) of the molecule (pages 173–174). From transfer measurements of numerous molecules and careful analysis of the data, simple relationships have been derived between surface areas of non-polar parts the molecules (A_{np}) and the thermodynamic parameters for transfer. The values vary quite strongly with temperature, and the expressions below are for 298 K.

$$\begin{aligned}
\Delta H_{\mathrm{tr}} &= A_{\mathrm{np}} \times 7\,\mathrm{J/mol} \\
\Delta S_{\mathrm{tr}} &= -A_{\mathrm{np}} \times 0.3\,\mathrm{J/(mol \cdot K)} \\
-T\Delta S_{\mathrm{tr}} &= -A_{\mathrm{np}} \times 90\,\mathrm{J/mol} \\
\Delta G_{\mathrm{tr}} &= A_{\mathrm{np}} \times 97\,\mathrm{J/mol}
\end{aligned}$$

As discussed earlier, the overwhelmingly predominant component of the unfavorable transfer free energy change is entropic. Polar surface area also influences the transfer thermodynamics, but this effect is much smaller than the influence of the non-polar surface area and will be ignored here.

From the structure of a folded protein, it is relatively straight forward to calculate the accessible surface area, and to distinguish between the polar and non-polar components of that surface. For the unfolded ensemble, some kind of model must be used because of the broad range of conformations. Calculations for hen lysozyme lead to the values listed below for the accessible surface areas of the native and unfolded states:

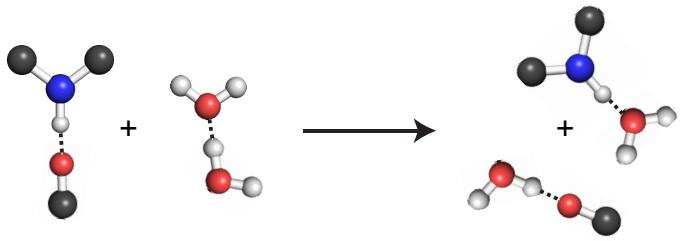
	Native (Å^2)	Unfolded (Å^2)	ΔASA(Å^2)
Total	6,670	15,800	9,130
Non-polar	3,400	9,700	6,300
Polar	3,300	6,100	2,800

Note that both polar and non-polar groups are accessible to solvent in the native state, contrary to the common perception that non-polar groups are almost entirely buried in folded proteins. None the less, non-polar groups are disproportionately buried in the native state and the non-polar accessible surface area increases greatly upon unfolding. Using these values and the expressions relating non-polar surface area to the thermodynamic parameters, we can estimate the contributions of the hydrophobic effect to the overall unfolding thermodynamics. The table below includes these estimates for hen lysozyme in an overall balance sheet.

	ΔH kJ/mol	ΔS J/(mol $\cdot$ K)	ΔG kJ/mol
Conformational entropy		2,500	-740
Hydrophobic effect	44	-1,900	610
Other	192	-14	190
Overall	236	586	61

The row labeled "Other" in the table above represents the additional contributions to the enthalpy, entropy and free-energy changes that must be added to those from conformational entropy and the hydrophobic effect in order to match the observed values for unfolding of hen lysozyme. Note that the hydrophobic effect balances out about 75% of the favorable conformational entropy change for unfolding, leading to a residual (600 J/(mol $\cdot$ K)) that very closely matches the overall entropy change for unfolding. On the other hand, there is an additional contribution of about 200 kJ/mol to ΔH_{u} that is not yet accounted for.

Qualitatively, the positive enthalpy change for unfolding indicates that attractive interactions are broken during unfolding. These interactions likely include hydrogen bonds and van der Waals interactions, both of which are apparent in the folded structures of proteins. Estimating the energetic contributions of individual interactions of this type is quite difficult, however, because interactions in the native state that are disrupted upon unfolding are likely compensated for by new interactions between the protein and the solvent. For instance, the folded proteins contain a large number of hydrogen bonds between the amide nitrogen atoms and carbonyl oxygen atoms of different residues, as illustrated in the right-hand side of the figure below:



At the same time, the water molecules surrounding the protein are extensively hydrogen bonded to each other. When the protein unfolds and the intramolecular hydrogen bonds are broken, the protein nitrogen and oxygen atoms can readily form new hydrogen bonds with water molecules. Breaking a hydrogen bond of this type, without replacing it, requires an energy input of about 50 kJ/mol. However, estimating the net energetic effect of breaking two hydrogen bonds (one in the protein and one between water molecules, as in the illustration above) is very difficult and has been the subject of controversy for nearly sixty years. The net energy difference is likely to be far smaller than the 50 kJ/mol for breaking a hydrogen in isolation, but the number of hydrogen bonds in a folded protein is quite large, about 100 in a protein the size of hen lysozyme. Thus the total contribution to hydrogen bonds could be quite significant.

One important source of information about the contributions of individual interactions to protein stability is experiments in which specific amino-acid residues have been modified by genetic engineering and the effects of these changes on unfolding thermodynamics have been been measured. For instance, a serine residue, in which the side-chain hydroxyl group forms a hydrogen bond with another protein group, can be changed to an alanine residue to eliminate the hydrogen bond. A large number of experiments of this type have been performed over the last few decades, and one of the major observations has been that the effects of this kind of change can vary greatly from protein to protein and among different sites within the same protein. Thus, the contributions of different types of interactions appear to be highly context dependent. For amino-acid replacements that remove hydrogen bonding groups, there is typically a reduction in ΔG_{u} in the range of 5–10 kJ/mol per hydrogen bond. When multiplied by the number of hydrogen bonds in a folded protein, these effects could readily account for a large fraction of the enthalpy change observed for unfolding.

Estimates of the kind discussed here provide a reasonably satisfying accounting for the observed thermodynamics of protein folding, with the predominant contributions

assigned to conformational entropy and the hydrophobic effect. The balancing of the conformational entropy change and the entropy change attributed to the hydrophobic effect seems particularly close. Some caution in interpreting these estimates is called for, however, since they are based on assumptions with significant uncertainties. Some of the greatest uncertainties concern the properties of the unfolded state, which influences both the estimate of ΔS_{conf} and the net contribution of the hydrophobic effect. Unfortunately, the very nature of unfolded states makes them much more difficult than folded proteins to characterize experimentally. This, among other areas of protein conformation and dynamics, continues to be an important area of research.