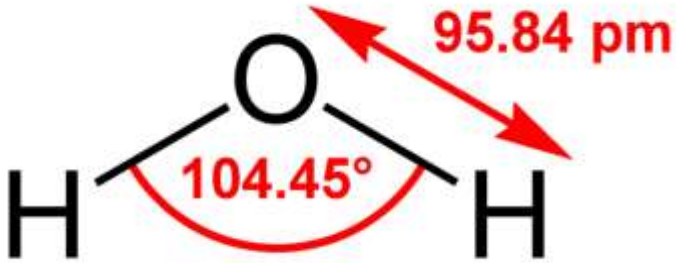
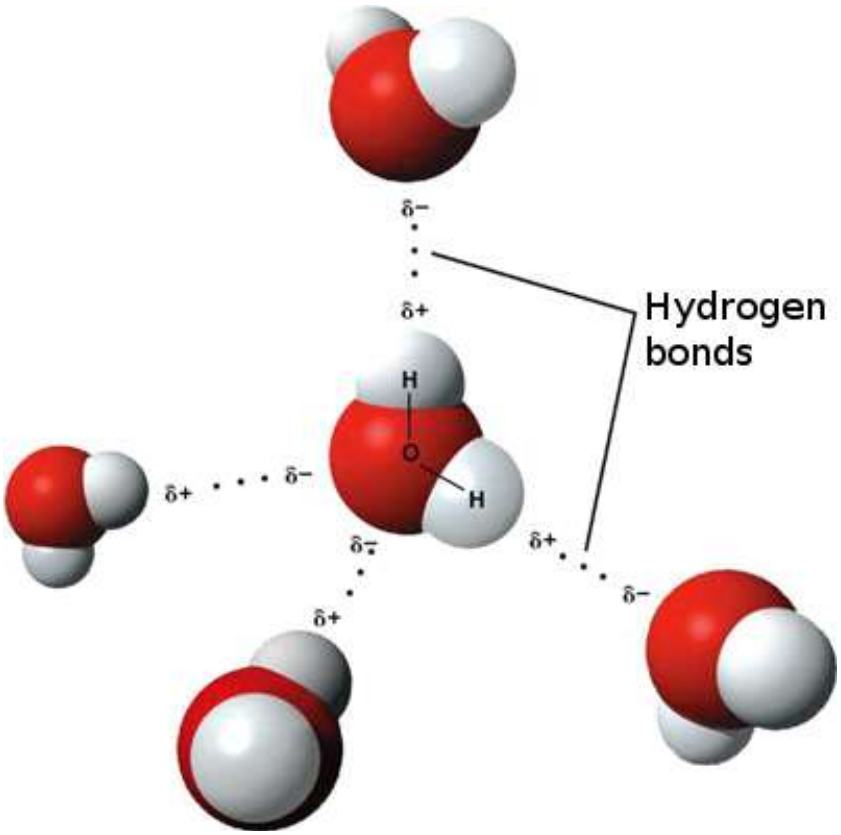
Water

In this supplement the central focus is the molecule water, H_2O and its importance in biological self-assembly. Water's physical properties will be discussed [Wiki]. These are the basis for understanding the remarkable process of self-assembly that is driven by the translational entropy of water molecules. Self-assembly applies to protein aggregations and enzyme complexes as well as to membranes. In addition, water plays a major role in understanding the general mechanism for all synthetic processes in which polymers are produced. In fact an incredible number of biological synthetic reactions are designed around achieving *dehydrations*, the elimination of water. The mechanism of dehydration is explained in the short cut [Polymer biosynthesis] as the means by which cells drive the transition from monomers to oligomers and polymers. The details of these chemical syntheses for proteins and fatty acids are found there.

The chemical formula for water is H_2O. It has a molecular weight (MW) of 18 d (dalton; 1 dalton = 1.67×10^{-24} g). Its structure is given in figure below.



The van der Waals radii for oxygen and hydrogen are 0.14 nm and 0.12 nm respectively. Nevertheless, the O-H bond length is only 0.0958 nm, not 0.26 nm. The O-H bond has 33% ionic character because O has much greater electronegativity than does H. This means that water molecules are dipoles and their dipole moment is 1.85 D (Debye; 1 D = 10^{-18} esu-cm). The Debye unit is roughly the product of the elementary charge, 4.8×10^{-10} esu (electrostatic unit), and the characteristic atomic length scale, 10^{-8} cm = 0.1 nm. In liquid water there is a great deal of intermolecular *hydrogen bonding*. Two water molecules may share a hydrogen proton in a hydrogen bond. The hydrogen bond length is 0.18 nm, about twice the O-H covalent bond length. The proton can shift over from one water molecule to the other, thereby forming OH^- and H_3O^+. This is shown in the figure below.



The hydrogen bond energy is about 20 kJ/mol = 4.8 kcal/mol = 8 $k_B T_{298K}$. While this is several times the thermal energy at physiological temperatures, it is nothing compared to typical covalent bond energies (83 kcal/mol for a C-C bond).

Erwin Schrodinger deduced from statistical mechanics the formula

$$t = \tau \exp\left[\frac{E}{k_B T}\right]$$

for the expectation time, t, for a thermal collision involving the energy, E, in which τ is the reciprocal of the collision frequency. In liquid water at 25 ^{o}C, τ ~ 10^{-13} sec. Thus, for a hydrogen bond

$$t = 10^{-13}\,\text{sec} \times e^{8} \cong 3 \times 10^{-10}\,\text{sec}$$

This means that thermal agitation destroys hydrogen bonds several times every nanosecond. Thus, the hydrogen bonds are incessantly forming, breaking and reforming. In contrast, the O-H covalent bond energy is 460 kJ/mol (110 kcal/mol = 183 $k_B T_{298K}$). In this case the result for t is

$$t = 10^{-13}\,\text{sec} \times e^{183} \cong 3 \times 10^{66}\,\text{sec}$$

There are 3.16 x 10^7 seconds per year which means that a covalent O-H bond will be destroyed by physiological thermal motions about once every 10^{59} years (the age of the universe is only of order 10^{10} years).

Ice is a tetrahedral arrangement of water molecules held together by hydrogen bonds. The hydrogen protons of one water molecule are attracted to lone pairs of electrons on another water molecule (each completed octet of electrons in the outer shell of O in water is comprised of two shared pairs of electrons in the two O-H bonds and two unshared pairs). This structure makes ice less dense than liquid water. At 0 ^oC,

$$\rho_{H_2O,liq.} = 1.00\ g/ml \text{ and } \rho_{ice} = 0.92\ g/ml$$

where the densities, ρ, are given in g/ml (grams per milliliter). On a nm scale even liquid water has a local ice like structure that is rapidly changing as water molecules reorient at a rate of 10^{12} sec^{-1}. While a water molecule in ice has 4 hydrogen bonded neighbors, in liquid water it has an average of 3.4. Liquid water is highly cohesive.

There are many other hydrogen bonds inside the cell. Amine hydrogens will hydrogen bond with carbonyl oxygens. This is an important contributor to structure in proteins, particularly in α helicies and β sheets. The base pairs of the double helix structure of DNA are held together by hydrogen bonds. Water molecules compete for such hydrogen bonds. Amine hydrogens can hydrogen bond to water oxygens and water hydrogens can hydrogen bond to carbonyl oxygens. In aqueous environment, molecules joined by a single hydrogen bond are highly susceptible to attack by water molecules that will compete for the hydrogen bonds. Since covalent bonds are 10 to 20 times stronger than hydrogen bonds, hydrogen bonds are referred to as weak bonds. This weakness is shared by hydrophobic interactions, van der Waals interactions and most ionic bonds. The strengths of these weak bonds vary from a few k_BT to 10 k_BT for T ~ 25 ^oC.

Self-assembly

Self-assembly is an extremely important generator of structure in nanobiology. It is based on multiple weak bonds and complementarity. Complementarity is a property of two macromolecular surfaces in which they each contain groups that are attrached to complementary groups of the other surface. This complementarity refers to both the nature of the groups and their geometric location. Obviously, positioning is critical since these weak bonds are short ranged (a few Angstroms). In the case of hydrogen bonds and ionic bonds, the interacting groups are of different kind (charges are opposite in ionic bonds and hydrogen bonds may join O and N say). The hydrophobic and van der Waals interactions may involve identical groups but usually involve different groups that merely share the same general character. The virtue of weak bonds is that they can be formed, broken and reformed readily. The virtue of multiple complementary bonds is that while an individual bond may be unstable, several such bonds confer a degree of stability. Nanobiology takes advantage of the plasticity of these multiple weak interactions.

Schrodinger's persepctive is helpful here. Above it was shown that for a weak bond of strength ~ 8 k_BT, the duration time of the bond is about 3 x 10^{-10} sec. What happens when two or three such bonds are involved in a complementary arrangement? The exponential factor creates a

dramatic result (below, $\tau = 10^{-13}$ sec and n is the number of weak bonds, each of strength ~ 8 k_BT):

$$n = 2 : t = \tau e^{16} = 8.9 \times 10^{-7} \text{ sec}$$
$$n = 3 : t = \tau e^{24} = 2.6 \times 10^{-3} \text{ sec}$$
$$n = 4 : t = \tau e^{32} = 7.9 \text{ sec}$$

Thus, three or four weak bonds in complementary positions are enough to increase stability for lengths of time to be important biologically. As a complex grows by virtue of more and more weak bonds, the structure becomes very stable, so much so that special mechanisms may be needed to break it down again. For this purpose, special catalysts may be required as well as the input of energy. For example, in the case of *actin*, an integral component of muscle fibers, the assembly and disassembly of the actin filaments requires ATP.

Translational entropy

What drives the process of self-assembly? As will now be shown, the expression, *self-assembly*, is a bit of a misnomer. Consider

$$A + B \leftrightarrow AB$$

which denotes the binding of A to B by several complementary weak bonds (imagine that A and B are polypeptides). It would appear that this process decreases translational entropy in going to the right. The translational entropy dominates entropy considerations in self-assembly. The entropy, S, may be expressed as

$$S = NS^0 + \frac{5}{2} Nk_B + Nk_B \ln\left[\left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} \frac{V}{N}\right]$$

in which h is Planck's constant, m is the mass of the molecule N is the number of molecules of a given kind, V is the containing volume and S^0 is the internal entropy of the molecule that includes vibrational, rotational and electronic contributions. The second term contains a portion of the translational and rotational contributions. The last term is the volume, V, translational entropy term because it comes from the freedom of the molecules to move about, i.e. *translate*, in the available volume. In fact, the argument of the natural logarithm in this term is dimensionless because it is the ratio of the available volume per molecule, V/N, and the *thermal DeBroglie volume*, V_T given by

$$V_T = \left(\frac{h^2}{2\pi m k_B T}\right)^{3/2}$$

For liquid water, V/N is 0.03 nm^3 and V_T is 1.4 x 10^{-5} nm^3. The ratio is roughly 2,000. Thus the translational entropy term contributes about 7.6 Nk_B for water. Water is 55 molar, whereas a

protein might be at no more than 10^{-3} molar, making V/N 55 x 10^3 times bigger for a protein molecule. Even a small protein has a molecular weight of at least 65,000, not quite 4,000 times that of water. This makes V_T smaller by a factor of about 250,000. Together, these two effects create a volume ratio 10^{10} times bigger than the ratio of 2,000 for water. This make the translational entropy factor about 30 Nk_B. Notice how the logarithm reduces the effect of this great increase in volume ratio. Nevertheless, the translational entropy factor is clearly the main contribution for a protein molecule.

Assume that initially $N_A^0 = N_B^0 = N^0$ and $N_{AB}^0 = 0$. After binding, assume that the process has gone to completion and that $N_A = N_B = 0$ and $N_{AB} = N^0$. Therefore

$$\Delta S = S_{final} - S_{initial} = N^0 S_{AB}^0 + \frac{5}{2} N^0 k_B + N^0 k_B \ln\left[\left(\frac{2\pi m_{AB} k_B T}{h^2}\right)^{3/2} \frac{V}{N^0}\right]$$
$$- N^0 S_A^0 - \frac{5}{2} N^0 k_B - N^0 k_B \ln\left[\left(\frac{2\pi m_A k_B T}{h^2}\right)^{3/2} \frac{V}{N^0}\right]$$
$$- N^0 S_B^0 - \frac{5}{2} N^0 k_B - N^0 k_B \ln\left[\left(\frac{2\pi m_B k_B T}{h^2}\right)^{3/2} \frac{V}{N^0}\right]$$
$$= -N^0\left(-S_{AB}^0 + S_A^0 + S_B^0\right) - \frac{5}{2} N^0 k_B - N^0 k_B \ln\left[\left(\frac{2\pi \frac{m_A m_B}{m_{AB}} k_B T}{h^2}\right)^{3/2} \frac{V}{N^0}\right]$$

Suppose that A and B are proteins with MW~30,000 d each. The mass ratio in the last line is 2.5 x 10^{-20} g. Assume that the initial concentration of A and B is 1 mM. Therefore, V/N^0 = 1.67 x 10^{-18} cm^3. Thus, the volume ratio in the last line's logarithm term is 2.9 x 10^{12}. The natural logarithm of this is 28.7. The internal entropy term difference is approximately zero but at least $S_A^0 + S_B^0 - S_{AB}^0 > 0$. The overall entropy change is quite negative, mostly from the translational entropy change. This appears to make sense since before binding species A and species B are free to move about independently of each other but after binding they move about together. This clearly reduces the motional freedom and decreases the translational entropy. How can the entropy spontaneously decrease? Doesn't this violate the second law of thermodynamics?

Since this process is at constant temperature, T, and pressure, P, the governing thermodynamic potential is the Gibbs free energy instead of the entropy. The condition for a spontaneous process is $\Delta G < 0$. The relationship to the entropy change is given by

$$\Delta G = \Delta U + P\Delta V - T\Delta S$$

in which ΔU is the change in internal energy and ΔV is the change in volume. Since the volume is virtually unchanged (there might be slight change in the molecular volumes), the $P\Delta V$ term can be ignored*. If

$$\Delta U < 0 \text{ and } |\Delta U| > -T\Delta S$$

then $\Delta G < 0$. Apparently the binding of A to B implies that $\Delta U < 0$. This process will go to completion as assumed only if this decrease in binding internal energy exceeds the entropy term.

The preceeding arguement has left out the involvement of water molecules! First of all, the complementary groups on A and B that lead to multiple weak bonds when binding occurs are initially interacting with water molecules that form weak bonds with these groups instead. These associations with water molecules look like little dynamic ice clusters around the binding groups. These water molecules must be displaced when A and B bind through their complementary weak bonds. Thus, it takes energy to displace the water molecules before internal energy is released by A and B binding. These two energies are both caused by weak bonds of comparable magnitude so that the net effect is that $\Delta U \sim 0$ afterall and is not negative, provided the contribution from water is included. But this displaced water also contributes to the entropy change. This contribution is an increase in entropy because each displaced water molecule is now free to move about independently of A or B or AB. Earlier it was argued that the translational contribution for a molecule of water is of order 7.6 k_B. The AB binding translational entropy decreased by 28.7 k_B. If 4 water molecules are released, this entropy decrease for AB will be compensated by the entropy increase for water. In fact, typically there will be at least 2 complementary binding sites on A and on B for the formation of AB and each will initially involve at least 1 water molecule apiece for a total of 4 bound water molecules. This is enough. However, generally there are at least 3 complementary sites and each involves more than 1 water molecule. This means that the entropy contribution from water not only compensates for the decrease in translational entropy associated with the binding of A to B but also compensates for the other small entropy factors as well.

Self-assembly is *entropy driven* by the translational entropy of released water molecules. One doesn't see the water molecules in electron micrographs or by other means that visualize the protein binding, only the proteins are seen, so it *appears* that $\Delta S < 0$. This is the apparent paradox of self-assembly. It is not *self*-assembly but rather an assembly process driven by something else, the translational entropy increase of water molecules. Self-ssembly is important in the formation of enzyme complexes, membranes and protein associations with and in membranes.

* A look at the difference in the volume per water molecule between ice and liquid can be used to assess the importance of the $P\Delta V$ term. From the densities given above for liquid water and ice the volumes per molecule are respectively

$$V_{liq.} = 3 \times 10^{-26} \frac{liter}{molecule} \quad and \quad V_{ice} = \frac{3}{.92} \times 10^{-26} \frac{liter}{molecule}$$

Therefore $P\Delta V$, where $\Delta V = V_{ice} - V_{liq.} = 0.26 \times 10^{-26} \frac{liter}{molecule}$, is given for one atmosphere of pressure by

$$P\Delta V = 0.26 \times 10^{-24} J/molecule$$

because one liter-atmosphere is 101 J. This is only $2.6 \times 10^{-18}\ erg$ per molecule, much less than the energy scale set by translational entropy that is order $k_B T = 4 \times 10^{-14}\ erg$.