

## Water Structure

TANFORD, C. (1980) *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*, John Wiley & Sons, New York. The classic text on hydration and hydrophobicity.

## Protein Structure

KABSCH, W. and C. SANDER (1983) "Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometric features," *Biopolymers*, **22**, 2577-2637. This paper describes the hydrogen bonding and geometric parameters for assigning protein secondary structures in proteins.

RAMACHANDRAN, G. N. and V. SASISEKHARAN (1968) "Conformation of polypeptides and proteins," *Adv. Protein Chem.*, **23**, 283-438.

SASISEKHAREN, V. (1962) "Stereochemical criteria for polypeptide and protein structures," in *Collagen*, ed RAMANATHAN, N., Madras, India: John Wiley & Sons, pp. 39-78. This landmark paper (as well as Ramachandran and Sasisekharan's listed above) puts forth the concept that protein conformations can be best described by the main chain torsion angles,  $\phi$  and  $\psi$  of the polypeptide chain.

SCHULZ, G. E. and R. H. SCHIRMER (1979) *Principles of Protein Structure*, Springer-Verlag, New York. A detailed discussion of all aspects of protein structure.

## Nucleic Acid Structure

SAENGER, WOLFRAM (1979) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York. A detailed discussion of all aspects of nucleic acid structure.

WATSON, J. D. and F. H. C. CRICK (1953) "A structure for deoxyribose nucleic acid," *Nature*, **171**, 737-738. This landmark paper describes the anti-parallel double-helical model for B-DNA.

The single crystal structures of A-DNA, B-DNA, and Z-DNA were first published in the following papers:

SHAKED, Z., D. RABINOVICH, W. B. T. CRUISE, E. EGERT, O. KENNARD, S. G. DALA, S. A. SALISBURY, and M. A. VISWAMITRA (1981) "Crystalline A-DNA: The x-ray analysis of the fragment d(G-G-T-A-T-A-C-C)," *Proc. Royal Soc. London, Ser. B*, **231**, 479-487.

WANG, A. H.-J., G. J. QUIGLEY, F. J. KOPAK, J. L. CRAWFORD, J. H. VAN BOOM, G. VAN DER MAREL, and A. RICH (1979) "Molecular structure of a left-handed double-helical DNA fragment at atomic resolution," *Nature*, **282**, 680-686.

WING, R., H. DREW, T. TAKANO, C. BROKA, S. TANAKA, K. ITAKURA, and R. E. DICKERSON (1980) "Crystal structure analysis of a complete turn of DNA," *Nature*, **287**, 755-758.

# 2

## Thermodynamics and Biochemistry

Thermodynamics, with its emphasis on heat engines and abstract energy concepts, has often seemed irrelevant to biochemists. Indeed, a conventional introduction to the subject is almost certain to convince the student that much of thermodynamics is sheer sophistry and unrelated to the real business of biochemistry, which is discovering how molecules make organisms work.

But an understanding of some of the ideas of thermodynamics is important to biochemistry. In the first place, the very abstractness of the science gives it power in dealing with poorly defined systems. For example, we can use the temperature dependence of the equilibrium constant for a protein denaturation reaction to measure the enthalpy change without knowing what the protein molecule looks like or even its exact composition. And the magnitude and the sign of that change tell us something more about protein molecules. Again, modern biochemists continually use techniques that depend on thermodynamic principles. A scientist may measure the molecular weight of a macromolecule or study its self-association by osmotic pressure measurements. All that is observed is a pressure difference, but the observer *knows* that this difference can be quantitatively interpreted to yield an average molecular weight. To use these physical techniques intelligently, we must understand something of their bases—this is what a good deal of this book is about.

In this chapter we shall briefly review some of the ideas of thermodynamics that are of importance to biochemistry and molecular biology. While most readers will have taken an undergraduate course in physical chemistry, it has been our experience that this usually leads, insofar as thermodynamics is concerned, to a fairly clear understanding of the first law and some confusion about the second. Since the aim of this section is the use of thermodynamics rather than contemplation of its abstract beauty, we shall emphasize some molecular interpretations of thermodynamic principles. But it should never be forgotten that thermodynamics does not depend for its rigor on explicit details of molecular behavior. It is, however, sometimes easier to visualize thermodynamics in this way.

## 2.1 HEAT, WORK, AND ENERGY—FIRST LAW OF THERMODYNAMICS

The intention of biochemistry is ultimately to describe certain macroscopic systems involving multitudes of molecules, in terms of individual molecular properties. The fact is, however, that such systems are so complex that a complete description is beyond the capabilities of present-day physical chemistry. On the other hand, a whole field of study of energy relationships in macroscopic systems has been developed that makes no appeal whatsoever to molecular explanations. This discipline, thermodynamics, allows very powerful and exact conclusions to be drawn about such systems. The laws of thermodynamics are quite exact for systems containing many particles and this gives us a clue as to their origin. They are essentially statistical laws.

The situation is in a sense like that confronting an insurance company; the behavior of the individuals making up its list of insureds is complex, and an attempt to trace out all of their interactions and predict the fate of any one of them would be a staggering task. But if the number of individuals is very large, the company can rely with great confidence on statistical laws, which say that so many will perish or become ill in any given period. Similarly, physical scientists can draw from their experiences with macroscopic bodies (large populations of molecules) laws that work very well indeed, even though the laws may leave obscure the mechanism whereby the phenomena are produced. Just as the insurance company with only 10 patrons is in a precarious position (it would not be too unlikely for all 10 people to die next month), so is the chemist who attempts to apply thermodynamics to systems of a few molecules. But 1 mole is  $6 \times 10^{23}$  molecules, a large number indeed. However, a bacterial cell may contain only a small number of some kinds of molecules; this means that some care must be taken when we apply thermodynamic ideas to systems of this kind (but see Hill 1963).

For review, let us define a few fundamental qualities.

**System.** A part of the universe chosen for study. It will have spatial boundaries but may be *open* or *closed* with respect to the transfer of matter. Similarly, it may or may not be thermally insulated from its surroundings. If insulated, it is said to be an *adiabatic* system.

**State of the system.** The thermodynamic state of a system is clearly definable only for systems at equilibrium. In this case, specification of a certain number of variables (two of three variables—temperature, pressure, and volume—plus the masses and identities of all the independent chemical substances in the system) will specify the state of the system. In other words, specification of the state is a recipe that allows us to reproduce the system at any time. It is an observed fact that if the state of a system is specified, its properties are given. The properties of a system are of two kinds. *Extensive* properties, such as volume and energy, require for the definition specification of the thermodynamic state including the amounts of all substances. *Intensive* properties, such as density or viscosity, are fixed by giving less information; only the *relative* amounts of dif-

ferent substances are needed. For example, the density of a 1 M NaCl solution is independent of the size of the sample, though it depends on temperature  $T$ , pressure  $P$ , and the concentration of NaCl.

Thermodynamics is usually concerned with changes between equilibrium states. Such changes may be *reversible* or *irreversible*. If a change is reversible, the path from initial to final state leads through a succession of near-equilibrium states. The system always lies so close to equilibrium that the direction of change can be reversed by an infinitesimal change in the surroundings.

**Heat,  $q$ .** The energy transferred into or out of a system as a consequence of a temperature difference between the system and its surroundings.

**Work,  $w$ .** Any other exchange of energy between a system and its surroundings. It may include such cases as volume change against external pressure, changes in surface area against surface tension, electrical work, and so forth.

**Internal energy,  $E$ .** The energy within the system. In chemistry we usually consider only those kinds of energy that might be modified by chemical processes. Thus the energy involved in holding together the atomic nuclei is generally not counted. The internal energy of a system may then be taken to include the following: translational energy of the molecules, vibrational energy of the molecules, rotational energy of the molecules, the energy involved in chemical bonding, and the energy involved in nonbonding interactions between molecules. Some such interactions are listed in Table 2.1. How these interactions affect the structure and stability of proteins and nucleic acids will be discussed in Chapter 3.

The internal energy is a function of the state of a system. That is, if the state is specified, the internal energy is fixed at some value regardless of how the system came to be in that state. Since we are usually concerned with energy changes, internal energy is defined with respect to some arbitrarily chosen standard state.

**Enthalpy,  $H = E + PV$ .** The internal energy of a system plus the product of its volume and the external pressure exerted on the system. It is also a function of state.

With these definitions, we state the first law of thermodynamics, an expression of the conservation of energy. For a change in state,

$$\Delta E = q - w \quad (2.1)$$

which takes the convention that heat absorbed by a system and work done by a system are positive quantities. For small changes, we write

$$dE = dq - dw \quad (2.2)$$

The slashes through the differential symbols remind us that whereas  $E$  is a function of state and  $dE$  is independent of the path of the change,  $q$  and  $w$  do depend on the path. The first law is entirely general and does not depend on assumptions of reversibility and the like.

TABLE 2.1 NONCOVALENT INTERACTIONS BETWEEN MOLECULES

Type of Interaction	Equation <sup>a</sup>	Order of Magnitude <sup>b</sup> (kJ/mol)
Ion-ion	$E = \frac{Z_1 Z_2 e^2}{Dr}$	60
Ion-dipole	$E = \frac{Z_1 e \mu_2 \theta}{Dr^2}$	-8 to +8
Dipole-dipole	$E = \frac{\mu_1 \mu_2 \theta'}{Dr^3} - \frac{3(\mu_1 r \theta'')(\mu_2 r \theta'')}{Dr^5}$	-2 to +2
Ion-induced dipole	$E = \frac{Z_1 e^2 \alpha_2}{2D^2 r^4}$	0.2
Dispersion <sup>c</sup>	$E = \frac{3h\nu_0 \alpha^2}{4r^6}$	0 to 40

<sup>a</sup>In these equations  $e$  is the charge of a proton (or the magnitude of the charge of an electron),  $Z$  is the valence charge,  $\mu$  is the dipole moment of a dipole, and  $\alpha$  is the molecular polarizability.  $D$  is the dielectric constant of the medium and  $r$  is the distance between the molecules. The factors  $\theta'$  and  $\theta''$  are functions of the orientations of dipoles. (See Moelwyn-Hughes 1961.)

<sup>b</sup>Calculations were made with the following assumptions: (1) molecules are 0.3 nm apart; (2) all charges are  $4.8 \times 10^{-10}$  esu (electron charge); (3) all dipole moments are 2 debye units; and (4) all polarizabilities are  $2 \times 10^{-24}$  cm<sup>3</sup>. These are typical values for small molecules and ions. The dielectric constant was taken to be 8, a reasonable value for a *molecular* environment; energies would be lower in aqueous solution, where  $D \approx 80\epsilon_0$ . Since the ion-dipole and dipole-dipole interactions depend strongly on dipole orientation, we have given the *extreme* values. For comparison, covalent bond energies range between 120 and 600 = kJ/mol.

<sup>c</sup>Dispersion interactions are between mutually polarizable molecules. The charge fluctuations in the molecules with frequency  $\nu_0$  interact, producing a net interaction. Since this depends so strongly on distance and becomes important only for very close molecules, a range of values is given.

If the only kind of work done involves change of volume of the system against an external pressure ( $pdV$  work),

$$dE = \delta q - PdV \quad (2.3)$$

Similarly, we write for the change in the enthalpy of a system the general expression

$$\begin{aligned} dH = d(E + PV) &= dE + PdV + VdP \\ &= \delta q - \delta w + PdV + VdP \end{aligned} \quad (2.4)$$

For systems doing only  $P dV$ -type work,  $\delta w = PdV$ , and

$$dH = \delta q + VdP \quad (2.5)$$

Equations 2.3 and 2.5 point up the meaning of  $dE$  and  $dH$  in terms of measurable quantities. For changes at constant volume

$$\begin{aligned} dE &= dq \\ \Delta E &= q_v \quad \text{for a finite change of state} \end{aligned} \quad (2.6)$$

whereas, for processes occurring at constant pressure, Eq. 2.5 gives

$$\begin{aligned} dH &= dq \\ \Delta H &= q_p \quad \text{for a finite change of state} \end{aligned} \quad (2.7)$$

That is, the heat absorbed by a process at constant volume measures  $\Delta E$ , and the heat absorbed by a process at constant pressure measures  $\Delta H$ . These quantities of heat will in general differ, because in a change at constant pressure some energy exchange will be involved in the work done in the change of volume of the system.

The thermochemistry of biological systems is almost always concerned with  $\Delta H$ , since most natural biochemical processes occur under conditions more nearly approaching constant pressure than constant volume. However, since most such processes occur in liquids or solids rather than in gases, the volume changes are small. To a good approximation, we can often neglect the difference between  $\Delta H$  and  $\Delta E$  in biochemistry and simply talk about the energy change accompanying a given reaction.

Figure 2.1 summarizes the relationships among the quantities  $q$ ,  $w$ ,  $\Delta E$ , and  $\Delta H$ . Note that we begin with the perfectly general first law and specialize to particular kinds of processes by adding more and more restrictions.

## 2.2 MOLECULAR INTERPRETATION OF THERMODYNAMIC QUANTITIES

We have seen that from the first law, together with the assertion that the internal energy is a function of the state of a system, powerful and general conclusions can be drawn. Although these have not required molecular models to attest to their validity, the student should keep in mind that quantities such as the internal energy and the energy changes in chemical reactions are ultimately expressible in terms of the behavior of atoms and molecules. It will be worth our while to explore the point in more detail.

Suppose that we ask the following question: If we put energy into a system to give an increase in the internal energy, where has the energy gone? Surely some has appeared as increased kinetic energy; but if the molecules are complex, some must be stored in rotational and vibrational energy and in intermolecular interactions, and perhaps some is accounted for by excited electronic states of a few molecules.\* Therefore, the question is really one of how the energy is *distributed*.

For thermodynamic properties, we are talking about large numbers of molecules, generally in or near states of equilibrium. The first means that a *statistical* point of view may be taken; we need not follow the behavior of any one molecule. The second implies that we should look for the *most probable* distribution of energy, for we would not expect an equilibrium state to be an improbable one. Although any system might, by momentary fluctuations, occasionally distribute its energy in some improbable way (like having almost all of the energy in a few molecules), the relative occurrence of such extreme fluctuations becomes vanishingly small as the number of molecules becomes very large.

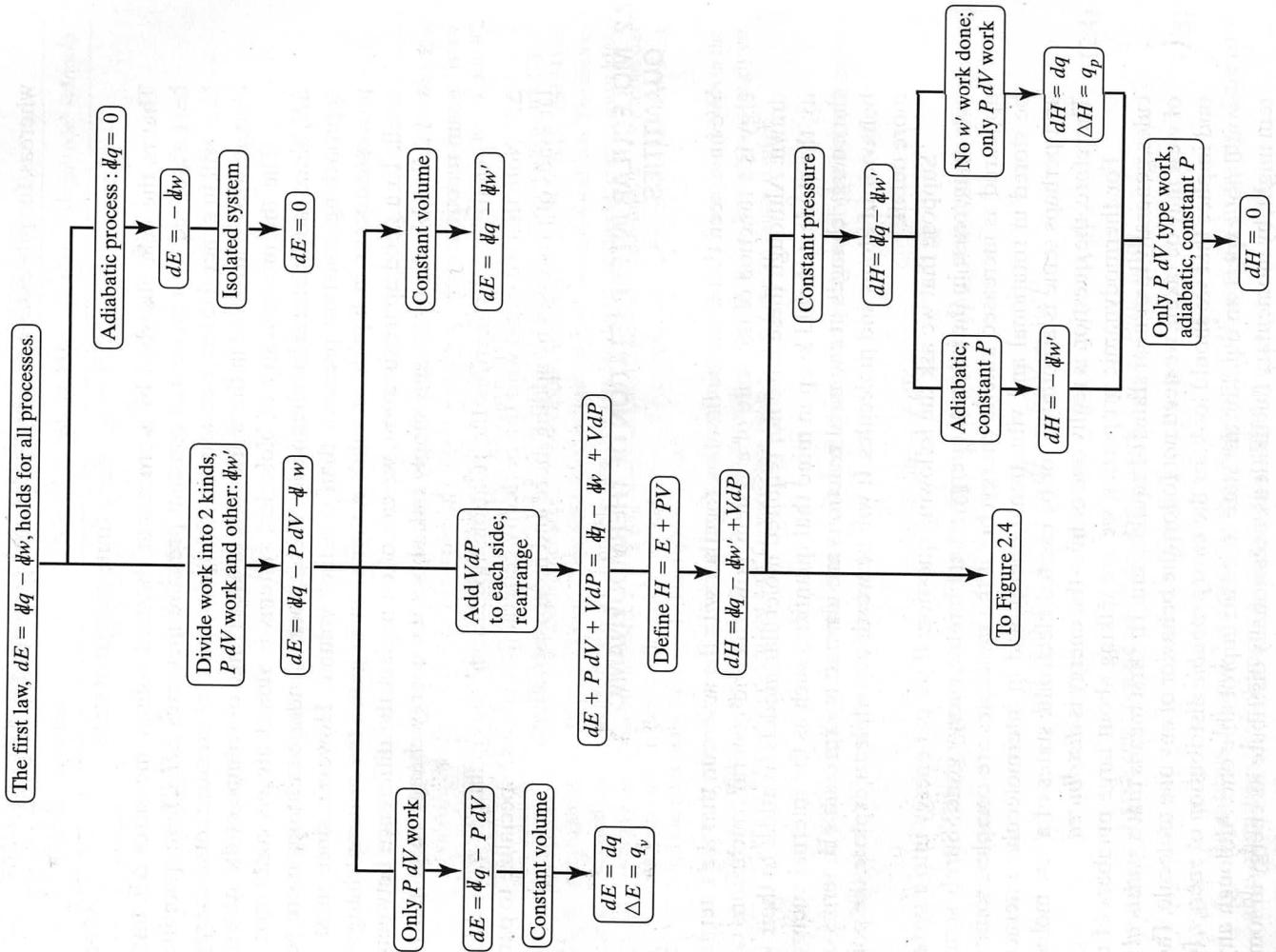


Figure 2.1 Consequences of the first law and the definition of  $\Delta H$ . The other kind of work  $dw'$  may be identified with electrical work, work done in expanding a surface and so forth.

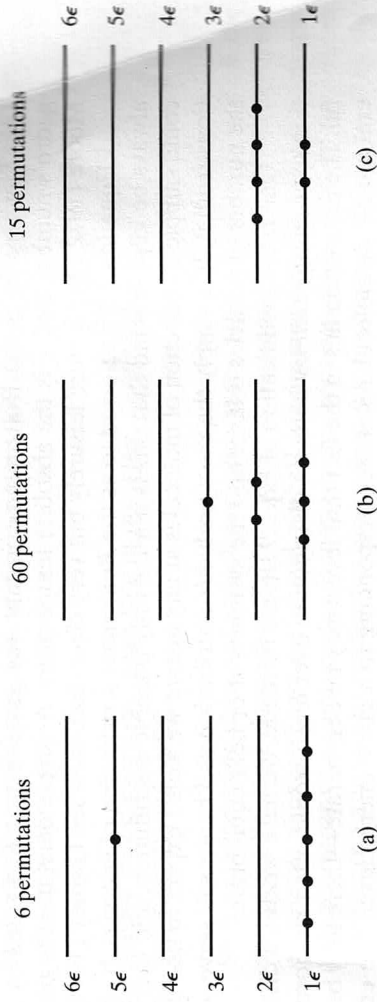


Figure 2.2 Some distributions of particles over energy states, subject to the constraints that  $N = 6$  and  $E = 10\epsilon$ . The numbers are calculated from Eq. 2.8 (remember  $0! = 1$ ). The Boltzmann distribution most closely resembles (b). Of course, there are not enough particles for it to hold accurately in this simple case.

To see the principles involved, let us take a very simple system, a collection of particles that might be thought of, for example, as atoms in a gas or as protein molecules in a solution. Each of these entities is assumed to have a set of energy states available to it, as shown in Figure 2.2. The energy states available to a particle are not to be confused with the *thermodynamic states* of a system of many particles. Rather, they are the quantized states of energy accessible to any particle under the constraints to which the whole collection of particles is subject. Suppose that we have six particles and a total energy of  $10\epsilon$ , where  $\epsilon$  is some unit of energy. Some distributions are shown in Figure 2.2, each of which satisfies the total energy requirement. Now let us say that for any particle, any state is equally probable. This simply means that there is nothing to prejudice a particle to pick a given state. Then the most probable distribution will be the one that corresponds to the largest number of ways of arranging particles over the states. If we label the particles, we see that there are only six ways of making state (a) and many more ways of making either state (b) or (c). The number of ways of arranging  $N$  particles,  $n_1$  in one group,  $n_2$  in another, and so forth, is

$$W = \frac{N!}{n_1!n_2!n_3!\dots n_i!} \quad (2.8)$$

Since the most probable distribution is the one that corresponds to the largest number of arrangements of particles over energy states, the problem of finding that distribution is a problem of maximizing the number  $W$ , subject to the restrictions that  $N$  and total energy  $E$  are constants. Using standard mathematical techniques for handling such problems, the result for a large number of particles is found to be

$$n_i = n_1 e^{-\alpha(\epsilon_i - \epsilon_1)} \quad (2.9)$$

where  $n_i$  is the number of particles in state  $i$ ,  $n_j$  is the number in the lowest state, and  $\epsilon_i$  and  $\epsilon_1$  are the energies of states  $i$  and 1, respectively. The constant  $\alpha$  turns out to be

$1/k_B T$ , where  $k_B$  is the Boltzmann constant, the gas constant  $R$  divided by Avogadro's number, and  $T$  is the absolute temperature. A simple derivation is given in Moore (1972). For a more leisurely but very clear discussion see Gurney (1949).

Equation 2.9 is referred to as the *Boltzmann distribution* of energies. It should always be kept in mind that this is not the only possible distribution, and that if we could sample a collection of molecules at any instant we would expect to find deviations from it. It is simply the *most probable* distribution and hence will serve well if the number of particles is large and the system is at or near equilibrium.

One more modification of Eq. 2.9 should be made. We have written a distribution over energy states, whereas a distribution over levels would often be more useful. The distinction lies in the fact that levels may be degenerate—there may be several atomic or molecular states corresponding to a given energy level. The energy levels of the hydrogen atom will serve as one example and the possible different conformational states corresponding to a given energy for a random-coil polymer as another. (See Chapters 4 and 8.) If each level contains  $g_i$  states (that is, if the degeneracy is some integer  $g_i$ ), levels should be weighted by this factor. Then

$$n_i = \frac{g_i}{g_1} n_1 e^{-(\epsilon_i - \epsilon_1)/k_B T} \quad (2.10)$$

where  $n_i$  and  $n_1$  now refer to the number of particles in energy levels  $i$  and 1, respectively.

Equation 2.10 states that if the degeneracies of all states are equal, the lowest states will be the most populated at any temperature. At  $T = 0$ ,  $n_i = 0$  for  $i > 1$ , which means that all particles will be in the lowest level, while as  $T \rightarrow \infty$  the distribution tends to become more and more uniform. At high temperatures, no level is favored over any other, except for the factor of degeneracy. Another useful form of Eq. 2.10 involves  $N$ , the total number of particles, instead of  $n_1$ . If we recognize that  $N = \sum_i n_i$  (the sum being taken over all levels), then

$$\begin{aligned} N &= \frac{n_1}{g_1} \sum_i g_i e^{-(\epsilon_i - \epsilon_1)/k_B T} \\ \frac{n_i}{N} &= \frac{n_1}{g_1} \frac{g_i e^{-(\epsilon_i - \epsilon_1)/k_B T}}{\sum_i g_i e^{-(\epsilon_i - \epsilon_1)/k_B T}} \end{aligned} \quad (2.11)$$

or

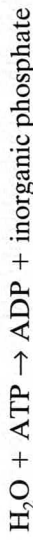
The sum in Eq. 2.11 is frequently encountered in statistical mechanics. It is sometimes called (for obvious reasons) the *sum over states* and more often (for less obvious reasons) the *molecular partition function*.

Since Eq. 2.11 gives the fraction of molecules with energy  $\epsilon_i$ , it is very useful for calculating average quantities. We shall make use of this idea in subsequent chapters.

For example, in Chapter 4 we shall calculate the average structure of helical molecules in just this way.

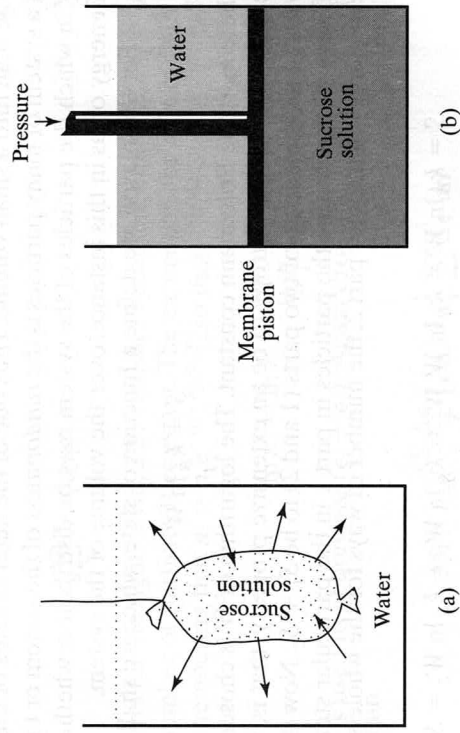
## 2.3 ENTROPY, FREE ENERGY, AND EQUILIBRIUM—SECOND LAW OF THERMODYNAMICS

So far in discussing chemical and physical processes, we have concentrated on the energetics. We have shown that the first law, a restatement of the conservation of energy, leads to exceedingly useful and general conclusions about the energy changes that accompany these processes. But one factor has been pointedly omitted—there has been no attempt to predict the *direction* in which changes will occur. Thus the first law allows us to discuss the heat transfer and work accompanying a chemical reaction, such as the hydrolysis of adenosine triphosphate (ATP) to yield adenosine diphosphate (ADP),



but gives no indication as to whether or not ATP will spontaneously hydrolyze in aqueous solution. Intuitively, we would expect that under the given conditions some particular equilibrium will exist between  $\text{H}_2\text{O}$ , ATP, ADP, and inorganic phosphate, but there is no way that the first law can tell us where that equilibrium lies.

As another example, consider the dialysis experiment shown in Figure 2.3. A solution of sucrose has been placed in a dialysis bag immersed in a container of



**Figure 2.3** (a) A dialysis experiment in which sucrose will diffuse out of a bag and water into the bag until equilibrium is attained. This process is irreversible. No work is done. (b) A way of doing the same experiment reversibly. The membrane piston is impermeable to sucrose and permeable to water. If the pressure on the piston is gradually reduced, the same final state (uniform mixing) approached irreversibly in (a) will be approached reversibly. With this arrangement, work will be done.