Chapter 2

The First Law of Thermodynamics

A. Introduction

To gain a good understanding of the laws of thermodynamics, it will help to develop an appreciation of the meaning of the words *law* and *thermodynamics*. Let's take a moment to think about these words before launching into a detailed discussion of how we might unpack the content of how the laws can be formulated. We are aided in this quest by the nature of science itself, which unlike ordinary prose and poetry aims to give words a more or less precise definition.

We are familiar with the concept of law from our everyday experience. Laws are rules that we are not supposed to break; they exist to protect someone's interests, possibly our own, and there may be a penalty to pay if the one who breaks a law gets caught. Such are civil and criminal laws. Physical laws are similar but different. They are similar in that they regulate something, namely how matter behaves under given circumstances. They are different in that violations are not known to have occurred, and they describe what is considered to be a basic property of nature. If a violation of a physical law should ever seem to have occurred, you will think first that the experiment has gone wrong at some stage, and second that maybe the "law" isn't a law after all.

Here's an example. Galileo,[1] like Copernicus,[2] believed that the orbits of the known planets were circles; the circle being the shaper of perfection and perfection being of the heavens. This view was inherited from Aristotle. Galileo also thought that the motion of

[1] Galileo Galilei, Italian astronomer and physicist, lived 1564–1642. His model of the Earth, the Moon, the Sun and planets was based on that of Copernicus, who had proposed a Sun-centered planetary system in his *De Revolutionibus Orbium Coelestium* (1543). Galileo is widely considered the father of modern science, because he emphasized the role of observations and experimentation in the discovery of new aspects of nature.

[2] Nicolaus Copernicus (1473–1543) held an ecclesiastical position in a church in Poland and was fascinated by astronomy.

celestial objects like planets was qualitatively different from the motion of terrestrial objects like cannonballs and feathers. But in fact, the orbits of planets are ellipses, not circles,[3] and the mechanical laws of planetary motion are fundamentally the same as those of a missile flying through the air on a battlefield, an object rolling down an inclined plane, and an apple falling to the ground in an orchard.[4] The point is not that Galileo was poor at science: his contributions to science have played an extremely important role in its development. Rather, the point is that what was considered a "law" was later shown not to be a law. (We can also see that at times a great does not get it quite right, in the best cases not through an inherent unwillingness to give all due consideration to available evidence, but because the evidence needed to change a perspective simply did not exist and was not yet sufficiently compelling.) There are many related examples one could cite from the history of science. It is the nature of human awareness of the physical world to develop in this way. It borders on the inhumane to assess the scientific ability of people who lived in a previous age by the standards and knowledge of today.

Whereas a human can break a law intentionally or unwittingly, a basic assumption of science is that a particle cannot break a law of physics. Particle motion is *governed* by the laws of physics (even if we don't know what those laws are). An important fact for us is that no violation of a law of thermodynamics is known to have occurred in nearly two hundred years of research in this area. Because of this many scientists, for example, Einstein, consider the laws of thermodynamics to be the laws of physics least likely to be overturned or superseded by further research. The laws of thermodynamics are generally described as the most general concepts of all of modern science. It behoves the biologist to be familiar with the basic principles of thermodynamics because they are of such basic importance. In view of all this, we might begin to suspect that the concepts we shall discuss are very deep and that considerable study and thought will be the price to pay for mastery of them. Thus has it ever been with basic things.

Energy has been around, well, since "the beginning," but the word *thermodynamics* was not coined until 1840, from the Greek roots *therme*, heat, and *dynamis*, power. The same roots appear in *thermometer* (a device to measure temperature, or heat) and *dynamite* (a powerful explosive). We can guess, then, that thermodynamics

[3] This was demonstrated by the German astronomer Johannes Kepler (1571–1630). In fact, though, the orbit of Earth is remarkably close to circular.

[4] As shown by the English mathematician, natural philosopher, and alchemist Isaac Newton (1642–1727). Sir Isaac is perhaps the greatest scientist of all time. His voluminous writings show that he was apparently as interested in theology and alchemy as in mathematics and natural philosophy, i.e. science. Thomas Jefferson, principal author of the Declaration of Independence and third president of the USA, owned a copy of one of Newton's lesser known works, *Observations upon the Prophecies of Daniel*.

will have to do with heat energy and power or movement. In fact, this branch of physics is concerned with energy storage, transformation, and dissipation. Thermodynamics aims to describe and relate – in relatively simple mathematical terms – the physical properties of systems of energy and matter. Thermodynamics has very much to do with molecular motion.

You might not think so, but you will certainly know something about thermodynamics. If not from having studied physics before starting university, then from having seen what happens when a pan of water is heated on the stove! At first, when the temperature of still water is about 25 °C, nothing seems to be happening; the eye does not detect any motion. But when heat is applied, motion becomes more significant and indeed readily apparent, so that by the time the boiling point is reached the water is moving about rather violently! So you do know something about thermodynamics, even if you don't normally think about it in the framework of today's physics, and a lot was known about thermodynamics well before the word was invented. There is not space to say much about the history of thermodynamics here, but it is worth mentioning that the principles of this science grew out of practical attempts in the nineteenth century to understand how to make a steam engine work as efficiently as possible and why heat is generated when one drills the bore of cannon, not academic speculation on universal law. This suggests that there may be value in avoiding being too prescriptive about how scientific knowledge should develop.

Like Kepler's laws of planetary motion and Newton's laws of mechanics, there are three laws of thermodynamics (plus one). There is a good deal about the first two of them here and in Chapter 3; they form the core of classical thermodynamics. Discussion of the First and Second Laws also provides the necessary context for introducing concepts that underlie the concept of free energy, a useful tool in the biological sciences (Chapters 4 and 5). The Third Law of Thermodynamics is of less immediate importance to biologists, but we'll touch on it at the end of Chapter 3, showing how it raises some very interesting questions about the nature of living organisms. For the purposes of our present subject, the chief practical value of studying the laws of thermodynamics is that they provide insight into how biological systems work and a framework for designing experiments, testing hypotheses, and explaining results.

We're ready for the First Law of Thermodynamics. But before investigating it, let's take one minute to go over the so-called Zeroth Law. The function of the Zeroth Law is to *justify* the concept of temperature and the use of thermometers (two things most of us are accustomed to take for granted!), and it is included here to provide a broader conceptual foundation to our subject. The form of the Zeroth Law is identical to that of a famous logical argument known at least as early as the ancient Greeks. It goes like this: if $\alpha = \beta$ (one premise), and $\beta = \gamma$ (another premise), then $\gamma = \alpha$

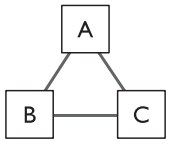


Fig. 2.1 The Zeroth Law of Thermodynamics. If three systems, A, B and C, are in physical contact, at equilibrium all three will have the same temperature. The concept of equilibrium is discussed in depth in Chapter 4.

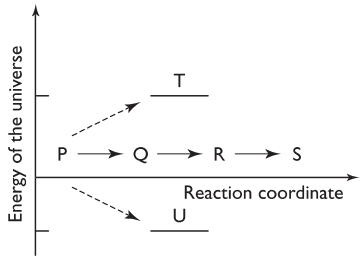


Fig. 2.2 The First Law of Thermodynamics. The total energy of the universe is constant, no matter what changes occur within. This principle also applies to an isolated system. Moreover, it is no less applicable to an open system or a closed system, as long as a complete account can be made of energy exchanged with the surroundings.

(conclusion). The Zeroth Law is built on this *syllogism*, or logical argument consisting of three propositions. It involves the concept of thermal equilibrium, that two objects A and B are in contact and at the same temperature.[5] The Zeroth Law states that if A is in thermal equilibrium with B, and B is in equilibrium with object C, then C is also in thermal equilibrium with A (Fig. 2.1). Simple! In Chapter 1 we touched on how temperature is a measure of the average speed of molecules in a gas. And now that we have the Zeroth Law, we are free to use the concept of temperature as much as we like.

The First Law is a conservation law: energy can be changed from one form to another, but in all its transformations energy is neither created nor destroyed (Fig. 2.2). There is a close resemblance to the conservation of matter, according to which the total amount of matter in a chemical reaction is a constant. The First Law of Thermodynamics is empirical in nature; it cannot be derived from more basic principles. Unlike the Pythagorean theorem,[6] for example, which can be derived from the most basic principles of Euclidean geometry,[7] there is no *mathematical* proof that the First Law of Thermodynamics is right. So then why should you believe it? *Sed solum ego ipse dixi*? Some might question an appeal to "authority" in scientific circles. We accept the First Law on a number of different bases, a most important and necessary one being that it is based on the experience of many, many researchers. The First Law has been tested many times, and as far as anyone knows, it has not been violated even once. It works. It's simple. It makes sense. That alone does not *prove* that the First Law is *true*, but it does at least give a good reason for thinking that it is probably a pretty good description of nature. So we believe in the First Law of the thermodynamics.

Despite its lack of a rigorous mathematical foundation, the First Law is the basis of *all* quantitative accounts of energy, regardless of form. The First Law makes energy the most scientific important concept in physics. And to the extent that physics is the basis of all of science and engineering, energy is the most important scientific concept in these technical areas. We saw in the previous chapter, the energy of a system can be converted from one form to another and distributed in a myriad of ways. And now we assume that energy is

[5] That thermal equilibrium is characterized by the equality of a single parameter (temperature) for all systems was first stated by Joseph Black (1728–1799), a Scottish chemist and physician.

[6] Named after Pythagoras (*c*. 580–500 BC), a mathematically inclined pre-Socratic religious philosopher. The Pythagorean theorem is $a^2 = b^2 + c^2$, where a, b and c are the lengths of the sides of a right triangle. Clay tablets unearthed in present-day Iraq prove that various combinations of integer which satisfy the algebraic equation were known a millennium before Pythagoras was born, historically on the island of Samos, very close to present-day Turkey. A more intuitive, geometrical proof of the theorem requires no knowledge of algebra. The theorem boasts more different proofs than any other theorem of mathematics – literally hundreds of different ones.

[7] The Greek mathematician Euclid lived *c*. 300 BC. His *Elements of Geometry* was the standard work on the subject until other types of geometry were invented in the nineteenth century.

not created or destroyed. The energy of a system plus surroundings is constant in time. For example, you can turn the chemical energy of an aphrodisiac into heat by a series of bodily actions that are better left to the imagination than described in writing or expressed in mathematical formulae, but the amazing thing is that throughout all the underlying changes, the total energy remains the same.

Box 2.1. Thermogenic oscillations keep the home fire burning

Homeotherms like mammals display adaptive thermogenesis. Detected exposure to cold by the brain leads to the activation of efferent pathways which control energy dissipation through the sympathetic nervous system. Mammals maintain body temperature by constant metabolism in cells throughout the body and by circulation of the blood. Plants, by contrast, lack a nervous system and are generally regarded as poikilotherms; most plants are unable to control their own temperature. Some cold-tolerant plants can, however, acclimatize to reduced temperature. So-called thermogenic plants have metabolic pathways that increase the temperature of a particular organ or tissue in response to cold. This group of plants includes the lotus (*Nelumbo nucifera*), a sacred symbol of perpetual regeneration in various religions. The lotus maintains its receptacle temperature between 30 °C and 36 °C during the 2–4 day sequence of anthesis, during which the flower bud opens and is fully functional. A perhaps more prosaic thermogenic plant is the skunk cabbage (*Symplocarpus foetidus*), a species of arum lily whose Japanese name, *Zazen-sou*, means Zen meditation plant, and whose English name comes from its bad smell and cabbage-like leaves. Skunk cabbage blooms in the early spring and maintains the temperature of the female spadix, the spike-like flowering stalk, at close to 20 °C for about a week, even when the ambient air temperature drops below the freezing point of water. Recently, researchers in Japan have shown that the thermogenic oscillations of the skunk cabbage are induced by a change in spadix temperature, through chemical reactions in the cells' mitochondria. The oscillations have a period of around 60 min, and the threshold is less than 0.9 °C, the greatest precision known among plants. There is also some evidence that the thermoregulation process is chaotic. In a model of the oscillatory temperature-sensing ability of skunk cabbage, the temperature of the spadix is maintained at a certain level where heat production and loss, due to radiation, evaporation, conduction, and convection, are balanced. An as-yet unidentified thermal sensor detects changes in temperature, and if the change exceeds the threshold for over about 30 min, thermogenesis is modified. The temperature oscillator in skunk cabbage appears to be distinct from known circadian rhythms and other kinds of biological rhythms.

B. Internal energy

To see more clearly how the First Law operates, we need to add internal energy and work to our conceptual toolkit. As with heat, both internal energy and work are measured in units of joules (or

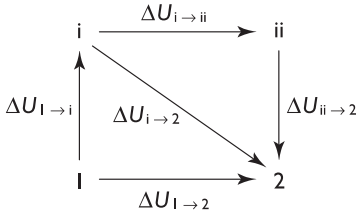


Fig. 2.3 Thermodynamic cycle. The difference in internal energy between state 2 and state I is $\Delta U_{I \to 2}$. Because U is a state function, the energy difference is independent of path; the internal energy of a system is determined by the specific physical properties of the state of the system and not on how the state was prepared. It follows that if the system begins in state I and is brought back to this state, $\Delta U = 0$. In symbols, $\sum_{loop} \Delta X = 0$. This holds not just for the internal energy but for any state function X.

calories). But not all heat is work, and internal energy will be heat or work only under certain circumstances. Say again? The internal energy is the energy *within* the system, U. For our purposes U will represent only those kinds of energy that can be modified by a chemical process – translational, vibrational, rotational, bonding, and non-bonding energies. A particle in a system may translate from A to B, a bond may rotate and vibrate, a bond may break and reform, and particles may interact non-covalently, for instance, by electrostatics. We'll leave out nuclear energy, even though it is always present and important in energy production and fusion reactions in the Sun. Nuclear energy simply does not play much of a role in the typical biochemical reaction, and when we think about a particular biochemical reaction, we take the atoms involved as given and we do not consider their history. So we can leave nuclear energy out of any calculation of the internal energy because the nuclear energy does not change in the typical biochemical reaction. Keeping track of quantities that do change is complicated enough!

The *internal* energy defines the energy of a substance in the absence of *external* effects, for instance, those owing to capillarity, electric fields, and magnetic fields. But U is an *extensive* property of a substance, meaning that its value depends on the size of the sample. For instance, the internal energy of 2 g of fat is twice as great as the internal energy of 1 g of fat under the same conditions. An intensive property; by contrast, for example, the concentration of sodium in a solution of sodium bicarbonate, is independent of the amount of the sample. U is a special kind of thermodynamic quantity called a *state function*. This means that U can be expressed in a certain mathematical form, and that the value of U depends only on the *current* state of the system (e.g. temperature, pressure and number of particles) and not at all on how the particles of the system came to be arranged in a particular way. An example will help to illustrate the point. The internal energy of an aqueous buffer solution depends only on its current state and not on whether it was made directly by mixing some chemicals with water or was prepared from a concentrated stock solution that had been frozen at $-20\,^{\circ}C$ for however long. Other general properties of state functions will be introduced as we go along.

The internal energy of a system cannot be measured *directly*; it is calculated from other measured properties. Moreover, it is not U that is measured but a change in U. But this presents no problems because normally we are interested in changes in U and not U itself. When a process brings about a change in a system from state 1 to state 2, the internal energy changes from U_1 to U_2, and the difference $\Delta U = U_2 - U_1$. For example, when salt is dissolved in water a large amount of heat is released, and solvation of the ions can be measured as a change in temperature. State 1 is the crystalline form of the salt and pure water, and state 2 is the salt when it is completely dissociated into ions and solvated. It does not matter whether we think of dissolution occurring in several steps (e.g. separation of ions in vacuum followed by solvation) or all in one go (Fig. 2.3); the

computed energy difference between states 1 and 2 is the same. This implies that ΔU for a complete cycle, say, a change from state 1 to state 2 and back again, will be 0, regardless of the *path* of the process – the succession of states through which the system passes. Many experiments corroborate the rule, and no exception is known. This is the experimental basis on which U is considered a state function. All state functions are path-independent.

The path-independence of U has the ring of the First Law about it. In fact, changes in U are what the First Law is about! In the money analogy of Chapter 1, the total amount at the end of the day did not depend at all on whether payment was made in coins and banknotes, nor on the order in which the customers made their purchases (and definitely not on the identity of the customers); it depended only on which purchases were made on a particular day. There are many, many ways in which money could change hands and still compute to a net change of Δm. The situation with ΔU is clearly very similar. Let's now set internal energy aside for a moment and have a look at work.

C. Work

Work, w, is a key physical quantity in thermodynamics; we had better know something about it. Adequate treatment of work requires the concept of force, F, so let's look at F and then w. A force is any influence that can cause an object to be accelerated, and it is usually measured in newtons, N ($1\ \mathrm{N} = 1\ \mathrm{kg\text{-}\ m\text{-}\ s^{-2}}$). There are as many kinds of force as there are kinds of energy. A familiar example of a force is the gravitational force of attraction of Earth for some object, e.g. the book you are reading. The force of gravity on an object is proportional to the quantity of matter (its mass), and for objects near the surface of Earth, the acceleration due to gravity, g, $9.8\ \mathrm{m\text{-}\ s^{-2}}$. These concepts are combined in Newton's famous second law, $F = ma$, where m is the mass of the object and a is the acceleration. When the force is the gravitational force, $a = g$ and $F = mg$. The gravitational force on a 70 kg man, i.e. his weight, is $70\ \mathrm{kg} \times 9.8\ \mathrm{m\text{-}\ s^{-2}} = 690\ \mathrm{N}$. The ratio of weight ($F$) to mass ($m$), namely, the acceleration due to gravity (g), is the same for a large object and a small object. Similarly, the ratio of circumference to diameter is the same for a large circle and a small circle. Time to get to work!

Work is similar to heat. When heat is added to a system, the internal energy changes (it increases). When work is done on a system, for example by compressing the volume of a gas, the internal energy changes (it increases). Both heat and work are forms of energy transfer *across* the boundary of a system; q and w are "boundary phenomena" (Chapter 1). You may have noticed that, unlike the internal energy, both heat and work are represented by lower case symbols. This is because U is a state function, but neither q nor w is a state function. Instead, q and w are *path* functions. In practical terms

this means that both *q* and *w* are *transient* quantities, unlike *U*, which is stored in a system. So heat and work are similar. But they are also different, and that's why different names are needed. They differ in that work is the *equivalent* of a force (e.g. gravity) acting through the displacement of an object, while heat is the transfer of energy owing to a temperature difference. Work involves the *non-random* movement of particles, heat the *random* movement of particles.

There are many different kinds of work. Not many different possible jobs for readers of this book, but different kinds of *w*. Here are a few of the latter sort: lifting of a weight against the force of gravity, expansion of a gas against an external pressure, movement of a charge against an electric field gradient (voltage), rotation of a shaft driving a fan, driving of an electric generator, expansion of a liquid film against its surface tension. In each case a force acts through a displacement, resulting in work being done by or on the system. The system does work when it pushes on the surroundings, and work is done on the system when the surroundings push on the system. When a system consisting of gas in a syringe is compressed by pushing on the plunger, the surroundings, which may include your hand and arm, do work on the system, and the internal energy of the gas increases.

The technical definition of *work* is similar to the one we are familiar with from everyday life. If someone "works hard," they put a lot of effort into accomplishing a task or reaching a goal. Similarly, in physics work is done when an object is moved against an opposing force. For example, when a crate is lifted vertically against the opposing force of gravity, the atoms of the box are involved in an *organized* transfer of energy; all the atoms of the box move together in the same direction. The heavier the crate, the more work done. But there are also differences in meaning, so we need to clarify what is meant by *work*. Although it may be income-earning work (non-technical meaning) for a porter to tote baggage at some distance above the ground, if the distance above ground is fixed no work (technical meaning) is done. This is because the suitcase is not displaced against the opposing force of gravity, though it is maintained at a fixed height against the pull of gravity. We have assumed that the suitcase alone is the system. If the person carrying the suitcase is made part of the system – and we are free to define the system however we like – then a sort of work, "physiological" work, is done in holding the suitcase at a fixed distance off the floor. Physiological work is done to maintain the muscle tension needed to hold the suitcase in place. This work results from transformation of chemical energy into mechanical energy in striated muscle (red meat) – a process that involves a lot of nerve impulses, the energy molecule ATP (Chapter 5), and polymerized actin and myosin (Chapter 8).

Another example will help to illustrate how work is done in coordinated movement against an opposing force. Pumping air into a bicycle tire is easy if the tire is flat; there is not much stuff in the tube to resist the flow of air in. Once the tube starts to fill, however,

and the pressure begins to build, it gets harder and harder to force air in. Here air in the pump is being moved in an organized way against the opposing force of the compressed air in the tire. During inflation, the tire expands somewhat but not much. This is because tires are made to adopt a certain shape, one well-suited to their function. The volume of the tire is approximately constant. Why does the pressure increase as you pump more air in? More and more molecules get stuffed into more or less the same volume.

Similarly, you may have been to a nightclub or concert, or been wearing a headphone, and found the music painfully loud. This is not merely a psychological effect: it comes from huge waves of air pressure smashing into your eardrum! Even in the middle of a grassy plain on a still day, where there is not a sound to be heard anywhere, the eardrum is in contact with the air. Nitrogen, oxygen, and carbon dioxide bombard it continually. We usually take no notice of this, because the same types of molecules are bombarding the eardrum from the opposite side with *equal* force. We detect a disturbance of the air as sound only when there is a pressure *difference* across the eardrum (a type of membranous boundary) and the difference is large enough. When a sound is so loud that it makes your ears hurt, it's because the pressure on the outside of your eardrum is far greater than the pressure on the inside, and to save your hearing it would be advisable to plug your ears or leave!

When particles of a system have reached *thermal equilibrium*, all the particles will have the same (average) thermal energy. In a gas, as we said above, some particles will be moving faster than others, but we can think of the like particles as having the same average speed. By the Zeroth Law, the temperature of each object in the system will be the same at equilibrium. What we're getting at here is that the *thermal energy* of a collection of particles is proportional to T, the *absolute* temperature. Suppose our system is a gas at thermal equilibrium. The gas particles move about freely within the system in all directions; particle movement is *disorganized*, or *random*. From time to time (very often!) a gas particle will collide with a wall of the container, the system boundary. The impact of such collisions will give rise to a pressure exerted by the particles on the contained. If we keep the volume constant but increase the number of particles in the container, the number of collisions between particles and boundary rises and the pressure is increased. If the volume and number of particles is constant but heat is added, the speed of the particles goes up and so does the temperature. Faster particles strike the walls of the system more often, increasing pressure. This descriptive view of temperature and pressure fits the simple mathematical relationship called the ideal gas law: $pV = nRT$, the pressure times the volume equals the number of moles of gas times the gas constant times the absolute temperature. This law has a long empirical foundation, and it can be derived from more basic principles. The ideal gas law is too simplistic to provide an accurate description of most real gases, but like many of the examples of elementary physics, it is none the less useful for

making sensible qualitative predictions about the behavior of matter under most of the conditions that most of us are likely to care about. We can leave the ideal gas law for now, but there will be reasons to pay it another visit later on.

What if we have a mixture of two different kinds of gas particles, a "light" gas like hydrogen and a "heavy" one like nitrogen? At thermal equilibrium, all the particles will have the same thermal energy. But will the helium molecule and nitrogen molecules bombard the walls of the container with equal force? No! Why not? Their masses are different. From physics the energy of motion of a particle, its *kinetic* energy (K.E.), is proportional to its mass times its velocity *squared*: K.E. $\propto mv^2$. (The direction of motion is not important here, so we can think of velocity as speed.) K.E. is a *non-linear* function of v. If the velocity of an object doubles, say, from 1 to 2 m s^{-1}, its K.E. quadruples. We can make the proportionality into an equality by including a multiplicative factor, in this case 1/2, but it's not needed for the present discussion. The *momentum* of a particle $\mathbf{p} = mv$. It is a *linear* function of v. The momentum of an automobile traveling at velocity v_1 is clearly much greater than that of a bicycle traveling at v_1. Historically, it took a good long while for physicists and philosophers to clarify the difference between K.E. and momentum and the relation of these quantities to energy and force, but we now think this is understood pretty well. By simple substitution K.E. $\propto \mathbf{p}^2/m$. Momentum matters to our present study because a change in momentum per unit time is proportional to a pressure. A change in momentum per unit time is a *force*, and a force per unit area is a *pressure*. In symbols, $\Delta\mathbf{p}/\Delta t = F = p/A$. At thermal equilibrium, the (average) kinetic energy of a particle is equal to the (average) thermal energy, so $\mathbf{p}^2/m \propto T$ at equilibrium. Solving this relationship for $\mathbf{p}$ gives $\mathbf{p} \propto (Tm)^{1/2}$. Thus, in our mixture of gases a "heavy" nitrogen molecule will have a greater average momentum than a "light" hydrogen molecule. Such ideas underlie all of physical biochemistry, and it will help us to keep them running in the background of our thinking as we make our way through the many avenues of our subject.

Before concluding this section, we want to introduce the concept of a heat engine. We'll give the heat engine the once-over-lightly treatment here and come back to it with vigor in Chapter 3. As shown in Fig. 2.4A, a heat engine does work by transferring heat from a source (e.g. a radiator) to a sink (e.g. a cold room). Only some of the heat transferred can be used to do work, because there is a fundamental limit on engine efficiency. (This limitation is a statement of the Second Law of Thermodynamics, as we shall see in Chapter 3.) The heat energy that is *not* used to do work enters the heat sink as randomized molecular motion. Work is energy transfer by ordered motion, heat is energy transfer by random motion. Note here how the First Law applies: the energy lost by the heat source (ΔU) is either converted into work (w) or transferred to the heat sink (q), and w and q must sum to ΔU. Figure 2.4B shows a diagram of a cell, say, an epithelial cell; it could just as well represent a tissue, organ or entire organism. The

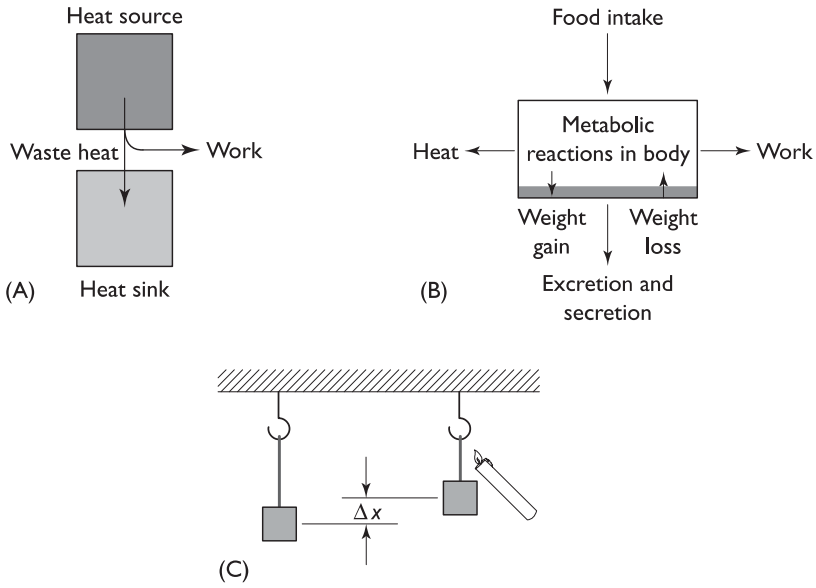


Fig. 2.4 Heat transfer. (A) Heat is transferred from a source (a warm body) to a sink (a cold body). Some of this heat can be used to do work, but certainly not all of it (Chapter 3). (B) Schematic representation of energy transformations within the body. The energy "input" is food. There are several "outputs." Aside from heat and work, which are mentioned explicitly in the First Law, there is excretion and change in body weight. In general, (food intake) – (waste excreted) = (change in body weight) + (heat) + (work), according to the First Law of Thermodynamics. (C) A very simple heat engine. The rubber band contracts when heated, lifting a weight. Because the weight is translated against the force of gravity, work is done. Thus, some portion of the heat is turned into work. The efficiency of this engine is low! Panel (C) is based on Fig. 44–1 of Feynman *et al.* (1963).

system of inputs and outputs resembles the situation in panel (A), but in panel (B) *everything is at the same temperature*. An organism is an isothermal system. Figure 2.4C shows how the heat energy of a candle can be used to do work. A rubber band dangles from a horizontal support, and attached to the rubber band is a weight (a mass accelerating under the force of gravity). When heat from the candle is absorbed by the molecules of the rubber band, the rubber band contracts. The attached weight is translated a distance Δx against the opposing force of gravity, and work w is done. Some of the heat of the candle will of course be lost to the surrounding air (this heat engine is rather inefficient), and only if adequate care is taken will the rubber not melt before our eyes, leaving no engine at all! Bearing all this in mind, let's take a closer look at how the First Law works.

D. The First Law in operation

By convention, the internal energy of a system will *increase* either by transferring heat *to* it or by doing work *on* it (Table 2.1). Knowing this, we can express the First Law of Thermodynamics as follows:

$$\Delta U = q + w. \tag{2.1}$$

Note that, in keeping with our earlier comments on measuring energy, the First Law defines only changes in ΔU. The conceptual background to Eqn. (2.1) was formulated in 1847 by the eminent German physicist and physiologist Hermann Ludwig Ferdinand von Helmholtz (1821–1894).[8] The idea of energy conservation had been

[8] See *Ueber die Erhalting der Kraft* (Berlin: Reimer, 1847). Helmholtz was son of a teacher of philosophy and literature at the Potsdam Gymnasium, a top-grade secondary school; his mother was descended from William Penn, a Quaker who

Table 2.1. *Sign conventions for heat and work*

Heat is transferred to the system	$q > 0$
Heat is transferred to the surroundings	$q < 0$
The system expands against an external pressure	$w < 0$
The system is compressed because of an external pressure	$w > 0$

proposed in 1842 by the German physiologist Julius Robert von Mayer (1814–1878).[9] It is interesting that a physiologist played such an important role in establishing one of the most important concepts of thermodynamics. When a system does work on its surroundings, w makes a negative contribution to ΔU because the system loses energy. Similarly, if heat is lost from the system, q makes a negative contribution to ΔU. In other words, ΔU measures the *net* amount of energy change in the system; it is the difference between the energy gained from the surroundings and the energy lost to the surroundings.

Let's look at some examples of Eqn. (2.1) in action. James Prescott Joule (1818–1889), son of a brewer in Manchester, England, is famous for his studies on the conservation of thermal energy understood as the mechanical equivalent of heat (1843). Perhaps the best-known experiment Joule did was to monitor the temperature of a vat of water during stirring. In this experiment, increases in water

was expelled from Oxford University for refusing to conform to Anglicanism and who later founded the American Commonwealth of Pennsylvania. Helmholtz's earned degree was in medicine; he would eventually receive an honorary doctorate in physics. Music and painting played a large part in his science. The present author, who was a Ph.D. student at Johns Hopkins University and a post-doctoral research fellow at University of Oxford, is connected to Helmholtz in various ways. Helmholtz's student Otto Richard Lummer (1860–1925) was awarded the doctorate in physics at the Humboldt University of Berlin. And Lummer was thesis adviser of George Ernest Gibson, a Scot, at University of Breslau, who was the adviser of Henry Eyring (see Chapter 8) in the Department of Chemistry at University of California at Berkeley, who was adviser of *I* in the Department of Chemistry at University of Utah, who was thesis adviser of *J* in the Department of Chemistry at University of Minnesota, who was thesis adviser of *K* in the Department of Pharmacology at University of Virginia, who was thesis adviser of the author. Johns Hopkins University was founded in 1876 by a wealthy Quaker merchant on the German model, which emphasized specialized training and research. The person who advised Helmholtz's medical thesis was Johannes Peter Müller, son of a shoemaker. Müller's academic degree was in medicine. His *Handbuch der Physiologie des Menschen für Vorlesungen* was recognized throughout the world, and it established a positive interchange between physiology and hospital practice in Germany, stimulated further basic research, and became a starting point for the mechanistic concept of life processes.

[9] The conservation of *mechanical* energy (kinetic energy + potential energy = constant) had in fact been proposed much earlier, by the German philosopher and mathematician Gottfried Wilhelm Leibniz (1646–1716), son of a professor of moral philosophy, and was an accepted principle of mechanics. Mayer's more general statement stemmed, curiously, from an analysis of the color of blood – a key means of distributing food energy throughout the body.

temperature represent positive increments in q, the heat transferred to the system. A motor turns a wheel in contact with water. The system is the water plus the wheel. As the wheel turns, mechanical energy is converted into increased motion of the water, and as we have seen, the motion of water is related to its temperature. Individual water molecules collide with the wheel and with each other. Vigorous and protracted stirring could eventually bring the vat of water to the boil. The *system* does no work; it does not expand against an opposing pressure or anything like that, so $w = 0$ and $\Delta U = q$. As a second example, suppose we have a motor. Suppose it has been determined that the motor generates 30 kJ of mechanical work per second, and that 9 kJ is lost to the surroundings as heat in the same amount of time. The change in internal energy of the motor per second is $-9\,\text{kJ} - 30\,\text{kJ} = -39\,\text{kJ}$. The energy produced by the motor is negative because work is done by the system on the surroundings and heat is lost to the surroundings. OK, but we also want to see how these ideas can be applied to something biological.

Above we saw that no work is done in holding up a heavy crate when the person is excluded from the system. And even when the person is included, no physical work is done, just physiological work. And if the energy expended in holding up the crate is not replenished by food, there will be a net decrease in the internal energy of the body. Our bodies do work even when we're sleeping! When you touch a metallic door handle on a wintry day, unless you have gloves on you can feel the heat being transferred from your flesh to the metal, and it might hurt! This heat transfer makes a negative contribution to the internal energy of the body, and the energy used to keep the body warm comes from food. When you walk up a flight of stairs, you do work against the force of gravity. If there are many steps to climb, as for instance on the way up to the cupola of Santa Maria del Fiore in Florence, or to Lady Liberty's torch in the Big Apple, by the time you've reached the top you may well be out of breath and dripping with perspiration. But if you're captivated by the view, you will not mind too much! Not only will you have lost energy to move your weight against Earth's gravitational pull on your body, you will be losing a lot of heat energy to the surroundings to keep your body cool. In any case, the energy used to climb stairs and the energy lost as heat comes from food.

We can be more quantitative about work with relatively little additional effort. From physics, the work done when an object is displaced a distance Δx ($x_{\text{final}} - x_{\text{initial}}$, where x refers to position) against an *opposing* force (hence the minus sign) of constant magnitude F is calculated as

$$w = -F\Delta x. \tag{2.2}$$

We see that work is the product of an "intensity factor" that is independent of the size of the system (the force) and a "capacity factor" (the change in the position of the object on which the force acts). For instance, the work done against gravity by a 50 kg woman in climbing

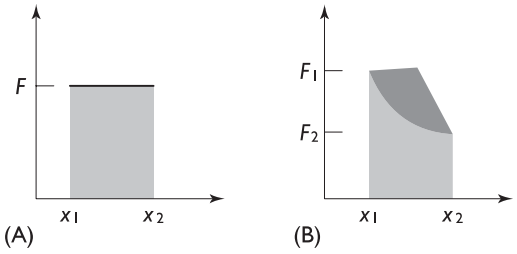


Fig. 2.5 Graphical representation of work: $|w| = F\Delta x$, where $|w|$ means "the absolute magnitude of w." For example, $|-3| = 3$. (A) For a constant force, the magnitude of work is the area $F \times \Delta x$. (B) If the force is variable, $|w|$ can no longer be calculated simply as $F \times \Delta x$. The figure illustrates why w cannot be considered a state function: its value depends on the path. See Fig. 2.3. The graphical method of calculating the work done by a system is said to have been introduced by James Watt.

to a point on a ladder 4 m above the ground is $-(50\,\text{kg} \times 9.8\,\text{m s}^{-2}) \times 4\,\text{m} = -1960\,\text{kg m}^2\,\text{s}^{-2} = -1.96\,\text{kJ}$. (Note: $1\,\text{J} = 1\,\text{kg m}^2\,\text{s}^{-2}$. Oddly, the dimensions of energy are [mass][length]^2[time]$^{-2}$.) The minus sign indicates that energy has been expended by the system, in this case, the woman. Diagrams help visualize the situation. The work done in Eqn. (2.2) can be represented graphically as an *area*, as shown in Fig. 2.5A. Figure 2.5B shows that the work done during a process depends on the path, because the shaded area need not be the same for all processes. This is a way of depicting that w is a path function; its magnitude depends on the path.

When a piston in a cylinder moves against a *constant* external pressure p_{ex} (as for instance in Fig. 2.6), the work done is

$$w = -p_{ex}\Delta V, \tag{2.3}$$

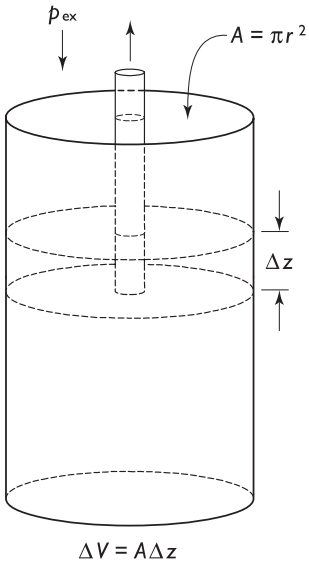


Fig. 2.6 Gas-filled cylinder. There is a constant external pressure, p_{ex}. For a change in position of the piston, Δz, there is a corresponding change in volume, ΔV. The work done *by* the system is $w = -p_{ex}\Delta V$. If ΔV is negative, if the gas in the cylinder is compressed, there is a positive contribution to ΔU.

where ΔV represents the change in the volume of the *system*; and $p_{ex} = nRT/V$, to the extent that whatever is pushing on the system can be modeled as an ideal gas. This type of work is called pV-work. Again, the work done is the product of an "intensity factor" (p_{ex}) and a "capacity factor" (ΔV). If the volume *of the system* increases, $\Delta V > 0$; the energy for expansion against an opposing force comes from the system itself, so the work done is negative. If there is no external pressure (if the surroundings are vacuum), then $p_{ex} = 0$; there is no force to oppose expansion of the system, and no work is done as V increases. Both p and V are state variables: they specify the state of the system.

E. Enthalpy

Another thermodynamic state function we need to know about is the enthalpy, H. It is covered in this book for several reasons, the most important one being that H is a component of the Gibbs free energy (Chapters 4 and 5). The term *enthalpy* is from the Greek *enthalpein*, to warm in, and it was coined around 1850 by the German physicist Rudolf Julius Emanuel Clausius (1822–1888), son of a pastor and schoolmaster. The enthalpy is the heat absorbed by a system at constant pressure (subscript "p"). Let's suppose we are working under constant pressure. Rewriting the First Law in terms of q, we have

$$q_p = \Delta U - w. \tag{2.4}$$

When the pressure is constant and the system expands from state 1 to state 2, the system does work on the surroundings. If the only type of work is pV-work, Eqn. (2.4) becomes

$$q_p = U_2 - U_1 + p\ (V_2 - V_1). \tag{2.5}$$

We can rearrange Eqn. (2.5) as

$$q_p = (U_2 + pV_2) - (U_1 + pV_1) = \Delta U + p\Delta V. \tag{2.6}$$

The complicated quantity $\Delta U + p\Delta V$ is equivalent to the heat exchanged at constant pressure. The right-hand side of Eqn. (2.6) is a state function, called the enthalpy, H:

$$H = U + pV. \tag{2.7}$$

Equation (2.7) can seem confusing and abstract; we need a way of making things fit better with what we have said already and our everyday experience of the world. We said above that $w = -p_{ex}\Delta V$ is a *path* function. But how can a state function plus a path function equal a state function? It is precisely because, although the product of p and V is a path function, p and V themselves specify the state of the system, and like U, they are independent of how that state was reached. H is therefore a state function, and it is no less or more of a state function than U is. And the development leading up to Eqn. (2.7), e.g. Eqn. (2.6), says that if the pressure is constant, the amount of heat exchanged during a reaction is independent of whether the product is formed directly or indirectly, in one or in a series of steps.[10] This statement, which is based on the results of experiments, says that state functions and state function differences (e.g. ΔU or ΔH) are additive (compare Fig. 2.3).

Looked at another way, the enthalpy can be thought of as the amount of energy *in* a thermodynamic system for *transfer* between itself and the environment. For example, in the calorimetry experiments in Chapter 1, the change in enthalpy was (very close to) the heat of oxidation, which was the energy transferred from the oxidized compounds to the calorimeter. When a system changes *phase*, for example, when a quantity of liquid water becomes solid, the change in enthalpy of the system is the "latent heat" of fusion, the heat given off to the environment in the freezing process. And in a temperature change, for example, the cooling off of a food item when placed in a fridge, the change in the enthalpy per unit temperature reflects a property of the material.

Let's see what happens when the enthalpy varies by a small but measurable amount. From Eqn. (2.7) we have

$$\Delta H = \Delta(U + pV) = \Delta U + \Delta(pV) = \Delta U + p\Delta V + V\Delta p. \tag{2.8}$$

[10] This is known as the "law of constant summation" of the Swiss–Russian chemist Germain Henri Hess (1802–1850), a physician and chemist. Hess's Law restates the First Law, though historically the former preceded the latter.

Note that p and V are assumed to vary independently. If the external pressure is constant, $\Delta p = 0$ and the last term vanishes. Substituting in Eqn. (2.1) and requiring pV-work only gives

$$\Delta H = q_p - p\Delta V + p\Delta V. \tag{2.9}$$

The last two terms on the right-hand side cancel, and we are left with

$$\Delta H = q_p. \tag{2.10}$$

Just as we said, the heat transferred to a system at constant pressure measures the change in the enthalpy of the system. Why the emphasis on heat transfer at constant pressure in a book on biological thermodynamics? Most of the reactions biochemists study are carried out at constant pressure (usually 1 atm), and as we shall see in Chapter 4, H is a component of a state function known as the Gibbs free energy, G, which predicts the direction of spontaneous change for a process at constant pressure and temperature, the biological scientist's favorite experimental constraints.

How can we *understand* the difference between ΔH and ΔU? Equations presented above make them out to be quite different, but the discussion about them sounds quite similar. In fact, the difference between ΔH and ΔU is often small enough to be neglected, but not always. If a reaction occurs in solution, and gas is neither produced nor consumed, $\Delta V \approx 0$. Under such circumstances $\Delta U \approx q_p$, as we can see from Eqn. (2.5), and so $\Delta U \approx \Delta H$. An example will help to illustrate that as a general rule it is a mistake not to take account of differences when there are reasons to suspect they might be significant. From Eqn. (2.7),

$$\Delta H = \Delta U + \Delta(pV). \tag{2.11}$$

The ideal gas law is $pV = nRT$, so assuming that the gas involved in our experiment can be modeled as an ideal gas, Eqn. (2.11) can be written as

$$\Delta H = \Delta U + \Delta(nRT). \tag{2.12}$$

If we now require constant temperature, $\Delta(nRT) = RT(\Delta n)$, where Δn represents the change in the number of moles of gas in the reaction; R, the universal gas constant, is 8.3145 J K^{-1} mol^{-1} in SI units (1.9872 cal K^{-1} mol^{-1} is also still in use); and T is the absolute temperature.

To illustrate, let's express the combustion of ethanol as:

$$C_2H_5OH(1) + 3O_2(g) \rightarrow 2CO_2(g) + 3H_2O(1). \tag{2.13}$$

From a bomb calorimetry experiment at 298 K and constant volume, 1368 kJ mol^{-1} of heat are released in the reaction. Now, $\Delta n = 2 - 3 = -1$. Therefore, by Eqn. (2.12), $\Delta H(298\,K) = \Delta U(298\,K) - RT = -1\,368\,000\,J\,mol^{-1} - 2480\,J\,mol^{-1} = -1370\,kJ\,mol^{-1}$. This is a small difference between ΔH and ΔU – less than 1%, i.e. ΔH is approximately equal to ΔU for beer, wine, and other such beverages – but it is a difference. We learn from this example that, although the

oxidation heats of Chapter 1 are changes in internal energy, they are very close to the corresponding changes in enthalpy. Check whether the combustion data you use for calculations do or do not take the pV term into account! A process for which the change in enthalpy is negative is called *exothermic*, as heat is let out of the system into the surroundings; a process for which the change in enthalpy is positive is called *endothermic*, as heat is let into the system from the surroundings.

Combustion of food in a bomb calorimeter tells us more than just how much heat is produced when food is completely burnt to a crisp. Indeed, tables of oxidation would be of little use to nutritionists if the numbers did not say something about the energetics of metabolism. Such tables are useful to the physicist, the biochemist, and the nutritionist because the laws of physics are assumed to be independent of time and location. In other words, the enthalpy of oxidation of glucose is not one value in a bomb calorimeter and some other value in the striated muscle connecting your big toe to the rest of your body. By Hess's Law, this enthalpy equivalence holds despite the fact glucose oxidation occurs in the body by a large number of sequential steps involving a large number of chemical intermediates. This discussion suggests that we can use machines like calorimeters to investigate the thermodynamic properties of the body and the molecules the body is made of. It also suggests that our bodies themselves are very much like machines.

Finally, suppose we have a system of pure water. We know from careful measurements that when ice melts at $+0.1\,^{\circ}\mathrm{C}$, barely above the melting temperature, $\Delta H = 1437.2\,\mathrm{cal\ mol^{-1}}$. When melting occurs at $-0.1\,^{\circ}\mathrm{C}$, just below the freezing point, $\Delta H = 1435.4\,\mathrm{cal\ mol^{-1}}$. The difference in enthalpy differences, $\Delta(\Delta H)$, is $1.8\,\mathrm{cal\ mol^{-1}}$. This is only about *half* the enthalpy change we would expect on changing the temperature of water by $0.2\,^{\circ}\mathrm{C}$ *in the absence of melting*. (See Section H below.) The difference arises from the change of phase that occurs between the initial state and final state, the melting of the solid into a liquid. It is necessary to account for the heat effect of any changes in the state of matter (solid, liquid or gas) when calculating ΔH.

F. Standard state

Changes in enthalpy (and other state functions) in tables of thermodynamic data are generally given for processes occurring under a standard set of conditions. The *standard state* is usually defined as one mole of a pure substance at 298.15 K (25.00 °C) and 1 bar ($1\ \mathrm{bar} = 10^5\,\mathrm{Pa} = 0.986\ 932\,\mathrm{atm}$). An example is the standard enthalpy change accompanying the conversion of pure solid water to pure liquid water at the melting temperature and a pressure of 1 bar:

$$\mathrm{H_2O(s)} \rightarrow \mathrm{H_2O(l)} \quad \Delta H^{\ominus}(273\,\mathrm{K}) = 6.01\ \mathrm{kJ\ mol^{-1}}. \tag{2.14}$$

Note that this enthalpy change is *positive*: heat must be added to ice at 0 °C in order to melt it. The standard enthalpy change used by the biochemist, ΔH°, is the change in enthalpy for a process in which the initial and final states of *one mole* of a substance in pure form are in their standard state: 25 °C and 1 atm pressure. The difference in enthalpy from the difference between 1 bar and 1 atm is almost always small enough to be neglected in biochemical reactions. But one should nevertheless be aware of the different ways in which thermodynamic data of use to the biochemist are presented in tables and be on the lookout for situations where the small differences cannot be neglected.

G. Some examples from biochemistry

Equation (2.10) is useful to the biochemist. As we have seen, it helps to make sense of the oxidation heats measured by bomb calorimetry. It can also be used in the study of protein stability, an important subject for several reasons. One is that about half of the dry mass of the human body is protein, and knowing how a polypeptide folds up into its native state would be of tremendous value in making good use of all the protein-encoding DNA sequence data that has been revealed by the Human Genome Project.

How can Eqn. (2.10) be used to study the thermodynamic properties of proteins (or of nucleic acids)? The native state of a protein is like an organic crystal. It is fairly rigid, and held together by a large number of different kinds of "weak" non-covalent interactions, including hydrogen bonds (largely electrostatic in character), van der Waals interactions[11] and "salt bridges" (electrostatic attractions between ionized amino acid side chains) (Tables 2.2 and 2.3). A native protein is "folded." In the "core" of a folded protein, apolar amino acid side chains interdigitate and are tightly packed, forming rigid and specific contacts. The rigidity of a folded protein is important to its biological function and, in favorable circumstances, permits determination of its structure at atomic resolution. This is not to say that a folded protein exhibits no fluctuations of structure or rotations of bonds. Native protein structure certainly does fluctuate, as we know for example by nuclear magnetic resonance studies, and such fluctuations can be important in the binding of small compounds to macromolecules (Chapter 7) and to enzyme function (Chapter 8). But the folded state of a typical protein is nevertheless quite rigid. In contrast, the unfolded state of a protein is more flexible and fluid-like. Bonds in amino acid side chains rotate relatively freely in an unfolded protein, and in the

[11] Named after the Dutch physicist Johannes Diderik van der Waals (1837–1923). Van der Waals was awarded the Nobel Prize in Physics in 1911.

Table 2.2. *Energetics of non-covalent interactions between molecules*

Type of interaction	Equation	Approximate magnitude (kcal mol^{-1})
Ion–ion	$E = q_1 q_2 / Dr$	14
Ion–dipole	$E = q\mu\theta / Dr^2$	−2 to +2
Dipole–dipole	$E = \mu_1 \mu_2 \theta' / Dr^3$	−0.5 to +0.5
Ion–induced dipole	$E = q^2\alpha / 2Dr^2 r^4$	0.06
Dispersion	$E = 3h\nu\alpha^2 / 4r^6$	0 to 10

[a] Charge q_1 interacts with charge q_2 at a distance r in medium of dielectric D.
[b] Charge q interacts with dipole μ at a distance r from the dipole in medium of dielectric D. θ and θ' are functions of the orientation of the dipoles.
[c] Dipole μ_1 interacts with dipole μ_2 at an angle q relative to the axis of dipole μ_2 and a distance r from the dipole in medium of dielectric D.
[d] Charge q interacts with molecule of polarizability at α distance r from the dipole in medium of dielectric D.
[e] Charge fluctuations of frequency ν occur in mutually polarizable molecules of polarizability α separated by a distance r.
The data are from Table 1.1 of van Holde (1985).

Table 2.3. *Characteristics of hydrogen bonds of biological importance*

Bond type	Mean bond distance (nm)	Bond energy (kJ mol^{-1})
O–H···O	0.270	−22
O–H···O^-	0.263	−15
O–H···N	0.288	−15 to −20
N^+–H···O	0.293	−25 to −30
N–H···O	0.304	−15 to −25
N–H···N	0.310	−17
HS–H···SH_2	—	−7

The data are from Watson (1965).

ideal case all amino acid side chains are completely exposed to solvent (Table 2.4).

The non-covalent interactions that stabilize folded protein structure (or double-stranded DNA or folded RNA structure) can be "broken" in a number of ways. One is by adding heat. If all the non-covalent bonds break simultaneously, in an all-or-none fashion ("cooperative" unfolding), then there are in essence just two states of the protein: the folded state and the unfolded state. The *transition* from the folded state to the unfolded state is like melting. So inducing the unfolding of protein by heat or some other means is something like melting a solid. This is true even if one is working not with a mass of freeze-dried protein but with folded proteins dissolved in aqueous solution. The *cooperativity* of the transition, the all-or-none character of going from being folded to being unfolded,

Table 2.4. *Principal features of protein structure*

Folded (native) state	Unfolded (denatured) state
Highly ordered polypeptide chain	Highly disordered chain – "random coil"
Intact elements of secondary structure, held together by hydrogen bonds	No secondary structure
Intact tertiary structure contacts, as in an organic crystal, held together by van der Waals interactions	No tertiary structure
Limited rotation of bonds in the protein core	Free rotation of bonds throughout polypeptide chain
Desolvated side chains in protein core	Solvated side chains
Compact volume	Greatly expanded volume

results from the concurrent breaking of a large number of weak interactions. In water, these interactions are hydrogen bonds; in proteins, they are the several kinds mentioned above. The melting of pure water or any other pure solid is a cooperative phenomenon. That is, melting occurs at a single or over a very narrow range of temperatures, not over a broad range. The same is true of cooperative protein unfolding or the melting of DNA.

A number of experimental studies have been carried out to measure the energy required to break a hydrogen bond at room temperature. This is pertinent not only to the unfolding of proteins but also to the "melting" of double-stranded DNA, which is held together by hydrogen bonds. Estimates of the bond energy vary, but a reasonable and generally agreed rough figure is 1 kcal mol^{-1}. Individual hydrogen bonds are weak; collections can be quite strong.

In terms of Eqn. (2.10), the enthalpy of the folded state of a protein is H°_F, the enthalpy of the unfolded state is H°_U, and the difference, $H^\circ_U - H^\circ_F$, is the enthalpy of denaturation or unfolding, ΔH°_d. In this case the folded state of the protein is the *reference state*, as the enthalpy of the unfolded state is measured with respect to it. What is this enthalpy difference? As discussed above, the enthalpy change for a process is equal to the heat absorbed by the system at constant pressure, and the rigid folded state of a protein can be pictured as a solid, and the flexible unfolded state as a liquid. So the enthalpy difference between the unfolded and folded states of a protein is the amount of heat needed to unfold the protein. As we shall see, the magnitude of that heat depends on the temperature.

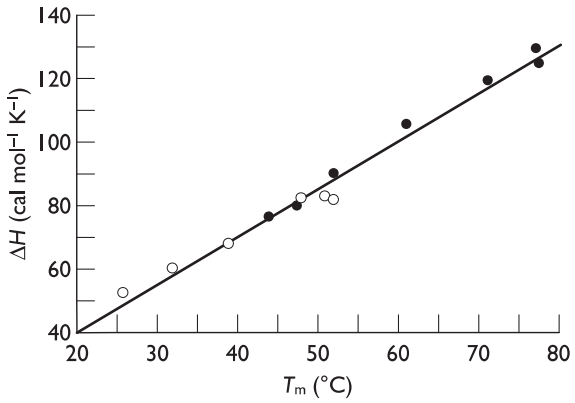


Fig. 2.7 Enthalpy of unfolding of hen egg white lysozyme as a function of transition temperature. Filled symbols: intact lysozyme. Open symbols: lysozyme in which one of the four native disulfide bonds has been removed. When folded, 3-SS lysozyme closely resembles the native state of intact lysozyme. Change in transition temperature was induced by a change of pH. Note that ΔH is approximately linear in T_m. The data are from Cooper *et al.* (1991).

The temperature at which a protein unfolds (or double-stranded DNA melts) is called the melting temperature, T_m. This temperature depends not only on the number and type of non-covalent bonds in the folded state but also on the pH and other solution conditions. T_m also depends on the pressure, but most biological science experiments are done at 1 atm pressure. In the case of proteins, changing the pH of solution changes the net charge on the protein surface. This can have a marked impact on T_m and ΔH°_d, as shown in Fig. 2.7 for the example of hen egg white lysozyme, a well-studied small globular protein. The figure also illustrates that the slope of ΔH° against T_m for this protein is more or less constant throughout the pH range shown.

Above we saw how a bomb calorimeter can be used to obtain thermodynamic information. Here we introduce isothermal titration calorimetry (ITC)[12] and explain how it can be used to measure the enthalpy of a biochemical process (Fig. 2.8). By Eqn. (2.10) the heat absorbed at constant pressure measures the enthalpy change. Suppose, for example, we are interested in the energetics of the binding of the F_c portion of immunoglobulin G (IgG), important in humoral immunity and biotechnology, to soluble protein A, a bacterial protein. We need not be concerned at the moment just which part of IgG the F_c portion of is: we just need to know that antibody molecules can be dissected into components and that the F_c portion is one of them. The thermodynamic states of interest here are the unbound state, where protein A is free in solution, and the bound state, where protein A is physically associated with F_c. The heat exchanged at constant pressure upon injection of protein A into a calorimetric cell containing the antibody can thus be used to determine ΔH°_b, the enthalpy of binding (under standard state conditions). The heat of injection will change as the number of vacant binding sites decreases.

[12] The isothermal titration calorimeter was first described in 1922 by Théophile de Donder, founder of the Brussels School of thermodynamics.

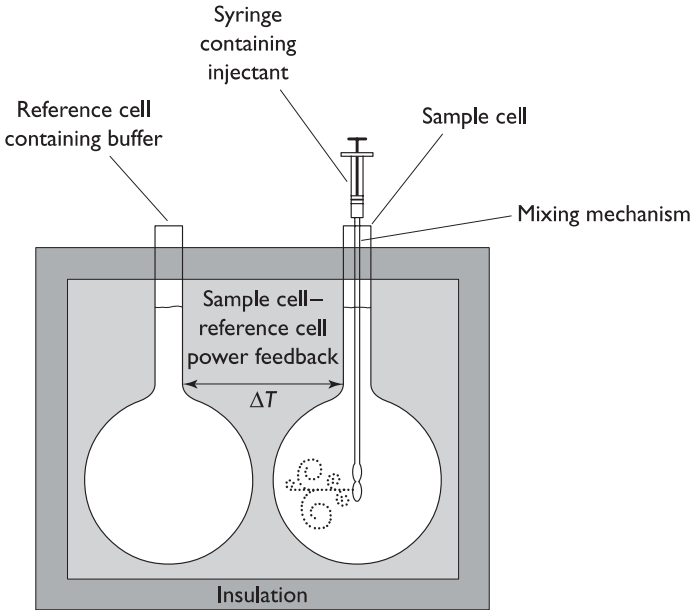


Fig. 2.8 Isothermal titration calorimeter. The temperature is constant. There are two identical chambers, the sample cell and the reference cell. In most cases, the sample cell will contain a macromolecule, and the syringe/stirrer is used to inject a ligand into the sample cell. The syringe is usually coupled to an injector system under software control and rotated at a constant speed. The reference cell is filled with buffer; no reaction occurs there. ΔT measures the temperature difference between cells, which are surrounded by insulation to minimize heat exchange with the surroundings. Electronic (power feedback) circuitry minimizes ΔT on a continuous basis. If injection of ligand results in binding, there will ordinarily be a change in the temperature of the sample. The sign of the change will depend on whether the reaction is exothermic or endothermic. An experiment consists of equal-volume injections from the syringe into the sample cell.

What if we're interested in the energetics of an enzyme binding to its substrate? This can be measured if a suitable substrate analog can be found or the enzyme can be modified. For instance, ITC has been used to measure the enthalpy of binding of a small compound called 2′-cytidine monophoshate (2′CMP) to ribonuclease A, which hydrolyzes RNA to its component nucleotides. 2′CMP binds to and inhibits the enzyme. If the enzyme of interest is, say, a protein phosphatase with a nucleophilic cysteine in the active site, mutation of the Cys to Ser or Asn will abolish catalytic activity, as in the N-terminal domain of the cytoskeleton-associated protein tensin, and the energetics of binding can be studied. A deeper understanding of binding will be sought in Chapters 5 and 7.

If you've spent any time in a biochemistry lab, you may have experienced the large heat given off by a salt solution as the salt dissolves. There are several contributions to the effect, but the main one is the enthalpy of hydration. This is the enthalpy change that occurs when an ion in vacuum is dropped into a sea of pure water. Water molecules form what is called a hydration shell around the ion, the number depending on the radius of the ion and its charge. Calorimetry can be used to measure the hydration enthalpy of biologically important ions. Values are given in Table 2.5. Why is this important? In one example, some of the water molecules hydrating an ion must be stripped away before the ion can pass through a selective ion channel in the plasma membrane, and this requires an input of energy. Complete dehydration of the ion would require a very large input of energy, so it is easy to imagine that a few water molecules remain associated with an ion as it passes through a pore. Ion channels that are specific for the passage of certain types of ion are part of the molecular machinery underlying the transmission of nerve impulses.

Table 2.5. *Standard ion hydration enthalpies*

H^+	−1090	Mg^{2+}	−1920
Li^+	−520	Ca^{2+}	−1650
Na^+	−405	Ba^{2+}	−1360
K^+	−321	Fe^{2+}	−1950
—	—	Zn^{2+}	−2050
NH^{4+}	−301	Fe^{3+}	−4430

The data refer to $X^+(g) \rightarrow X^+(aq)$ at 1 bar and are from Table 2.6c in Atkins (1998). 1 bar = 10^5 Pa = 10^5 N m^{-2} = 0.987 atm. 1 Pa = 1 pascal. Blaise Pascal (1623–1662) was a French scientist and religious philosopher.

H. Heat capacity

Above we noted that the heat taken up or released per unit change in temperature from a material at constant pressure is a property of that material. The name of this property is the heat capacity at constant pressure, C_p.[13] This quantity is readily measured, and it can be used to calculate changes in the enthalpy. The heat capacity per unit mass of Coke, for example, which is mostly water, differs from the heat capacity per unit mass of an egg, which contains a large amount of protein; the amount of heat energy that must be extracted from 1 g of each in order to lower the temperature by 1 degree K will not be the same in the two cases. The heat capacity tells us how much energy is in the system for transfer between itself and the environment *per degree*.

It is evident from Fig. 2.7 that the enthalpy of a protein rises as its temperature is increased. This is true of substances in general. The numerical relationship between H and T, however, depends on the conditions. We concern ourselves here with constant pressure only. The slope of a graph of H versus T at constant pressure is the heat capacity at constant pressure.

We are all familiar with heat capacity, even in the absence of a formal introduction. Returning to our water-in-a-saucepan example, if the water is heated at a high enough rate, it will eventually boil. The amount of heat that must be added to increase the temperature of 1 g of a substance by 1 degree K is the *specific heat capacity*. At 1 atm pressure, the heat capacity of liquid water varies only very slightly with temperature over the range 0–100 °C. This comes from the bonding structure of water (Fig. 2.9). Just as the extent of motion changes substantially when a quantity of water freezes or vaporizes, the heat capacity of water depends substantially on its state. This is true of substances in general. But the number of hydrogen bonds

[13] The heat capacity per unit mass of material, or specific heat, was first described in detail by Joseph Black.

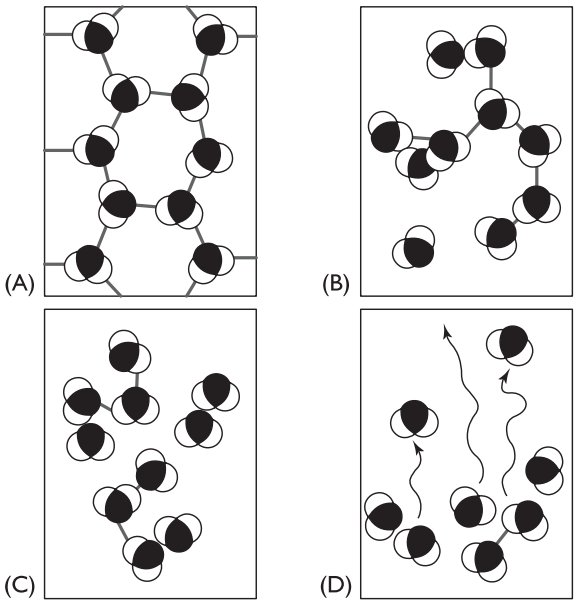


Fig. 2.9 Schematic diagram of the structure of water under different conditions: (A) solid state, (B) melting point, (C) liquid state, and (D) boiling point. Oxygen is shown in black, hydrogen in white. The black bars represent hydrogen bonds. Hydrogen bonds are relatively persistent in the solid state. The number of bonds decreases at the melting point, as molecules move out of the lattice. In the liquid state, hydrogen bonds are present, but they are formed only transiently. Boiling water has such a high thermal energy that persistent hydrogen bonds are rare. As the temperature increases, there are increases in translational, vibrational and rotational energy. The change in translational energy is not very difficult to detect; for example, when water is brought to the boil on the stove. Increases in vibrational and rotational motion of water cannot be seen with the naked eye. Based on Fig. 3.3 in Voet and Voet (1995).

formed by an individual water molecule is roughly constant throughout the temperature range 0–100 °C at 1 atm pressure. A quantity of water vapor can be much hotter than 100 °C, and water vapor must lose a good deal more energy than liquid water to fall in temperature by 1 degree K. This makes steam hazardous for people, but it reduces the time to cook the claw muscle of Chesapeake Bay blue crabs.

With regard to purified proteins in solution, just as the folded and unfolded states have different enthalpies, they also have different heat capacities. The heat capacity of the folded state is $C_{p,F}$, while that of the unfolded state is $C_{p,U}$. The heat capacity difference between states at constant pressure is $C_{p,U} - C_{p,F} = \Delta C_{p,d}$. In principle, $C_{p,F}$, $C_{p,U}$ and therefore $\Delta C_{p,d}$ are temperature-dependent. In practice, however, the variation with temperature can be and often is sufficiently small to be ignored. An example of a case where $\Delta C_{p,d}$ is only slightly dependent on temperature is shown in Fig. 2.7. That $\Delta C_{p,d}$ is positive is related to the increase in hydrophobic surface that is in contact with the solvent. The side chains of the hydrophobic core are largely sequestered from the solvent in the folded state.

Now we are in a position to write a general expression for the enthalpy of a substance as a function of temperature. It is

$$H(T_2) = H(T_1) + C_p(T_2 - T_1), \tag{2.15}$$

where T_1 is the temperature of the system in state 1 and $H(T_2)$ is the enthalpy of the system in state 2. Another way of writing Eqn. (2.15) is

$$\Delta H = C_p \Delta T, \tag{2.16}$$

where $\Delta H = H(T_2) - H(T_1)$ and $\Delta T = T_2 - T_1$. Note that ΔH would have the same magnitude but the opposite sign if the labels attached to the

states were reversed; the enthalpy is a state function. From a mathematical point of view Eqn. (2.16), which can be written $C_p = \Delta H/\Delta T$, tells us that the constant pressure heat capacity can be obtained from a plot of H versus T in the interval ΔT. When C_p is constant throughout the temperature range, H versus T will be constant. As we have said, C_p is effectively constant over small temperature ranges for many materials in the absence of a change of phase. But the unfolding of a protein can be described as a phase change (melting of a solid), and we should therefore expect that there will be a difference in heat capacity between the folded and unfolded states. The corresponding expression to Eqn. (2.8) for the enthalpy difference between the unfolded and folded states of a protein is

$$\Delta H_d^\circ(T_2) = \Delta H_d^\circ(T_1) + \Delta C_{p,d}(T_2 - T_1), \tag{2.17}$$

where the heat capacity change is independent of temperature. Equations (2.15) and (2.17) apply to many different situations (not just protein folding/unfolding!) and are known as Kirchhoff's enthalpy law, after the German physicist Gustav Robert Kirchhoff (1824–1887).

One way of determining $\Delta C_{p,d}$ for protein unfolding is to denature the protein under different conditions. A common method is to measure ΔH_d° and T_m for different values of pH, as shown in Fig. 2.7. This can be done with a technique called differential scanning calorimetry (DSC), which measures the heat absorbed as a function of temperature (Fig. 2.10). The experiment is repeated at a variety of pH values to generate a curve like that shown in Fig. 2.7. As we shall see in Chapter 5, the relatively large $\Delta C_{p,d}$ of protein unfolding has a big impact on how much work must be done to unfold a protein, and how this amount of work depends on temperature.

Box 2.2. A micromachined nanocalorimeter for life sciences research and diagnostics

Receptors are membrane-embedded protein molecules that recognize and respond to the body's own chemical messengers, for example, hormones or neurotransmitters. In pharmacology, receptor affinity and efficacy together determine the potency of a drug. Differences in efficacy make a drug an agonist or antagonist. A drug of sufficient efficacy and affinity for a receptor to affect cell function is an agonist. A drug that binds the receptor but does not elicit a response is an antagonist. Pre-clinical screening of agonists and antagonists and assessment of the toxicity of novel lead compounds is generally done with specialized cell-based assays. Often, a specific cell line is required. An understanding of the nature of the molecules and cellular pathways involved is generally needed to interpret the results of such assays, and therefore a substantial investment of time and money. Close monitoring of cell temperature could provide a means of detecting changes in cell metabolism that are not stimulus-specific, enabling simpler, less expensive and more general cell-based screening than with specialized cell-based assays. Researchers in the Department of Electronic and Electrical Engineering at University of Glasgow have developed a

Box 2.2. Cont.

micromachined nanocalorimeter which functions as a biosensor. A small number of living cells are present in a sub-nanoliter chamber. The small size of the chamber could be useful for rapid screening of small samples. The sensor comprises a 10-junction gold and nickel thermopile on a silicon chip. A thermopile is a number of thermocouples, 10 in this case, connected end on end, and a thermocouple is simply a temperature-measuring device consisting of two wires of different metals fused at each end. A temperature difference between the metals results in a difference in an electrical potential, which can be calibrated to a temperature. The nanocalorimeter of the Glasgow group can detect a mere 13 nW of power generated by the cells on exposure to a chemical stimulus, the temperature resolution is 0.125 mK, the heat capacity is 1.2 nJ mK^{-1}, and the response time is 12 ms. Primary cell lines or tissue biopsies can be analyzed.

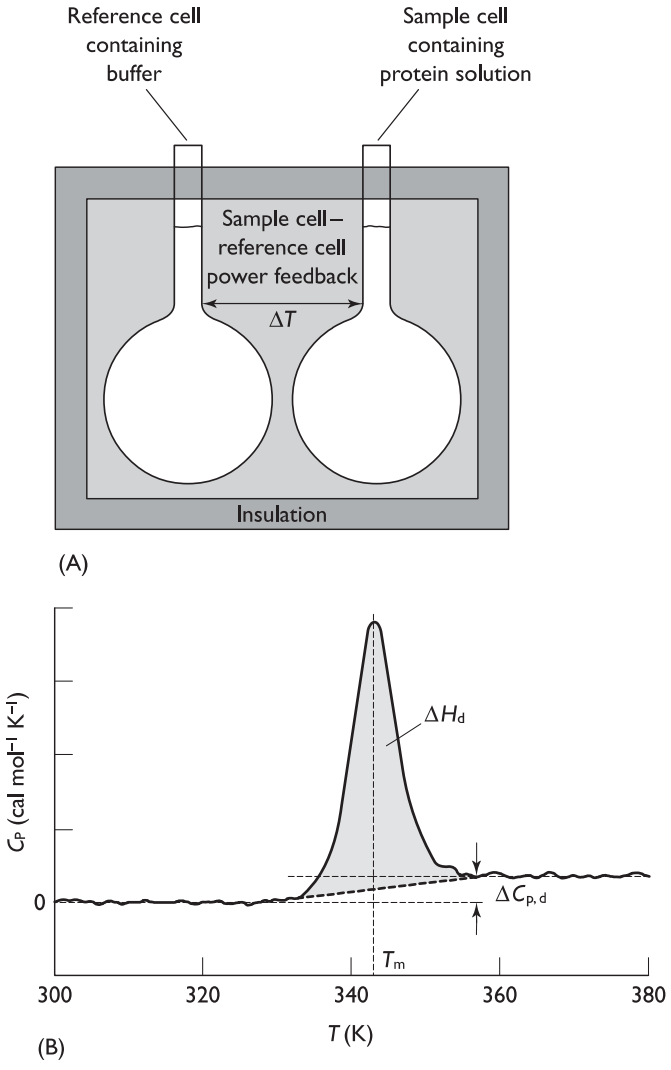


Fig. 2.10 Differential scanning calorimetry. (A) Schematic diagram of the instrument. In this case the reference cell contains buffer only, and the sample cell contains the macromolecule dissolved in buffer. Both cells are heated very slowly (e.g. 1 °C min^{-1}) in order to maintain equilibrium, and feedback electronic circuitry is used to add heat so that $\Delta T \approx 0$ throughout the experiment. Other types of DSC have been used for other purposes in biophysics, for example, to investigate the physiological limits of the freeze tolerance and freeze-avoidance strategies taken by different insect species to survive subzero temperatures. (B) Data. The heat added to keep $\Delta T \approx 0$ can be plotted as a function of temperature. The endothermic peak corresponds to heat absorbed, for example, on protein denaturation. The peak maximum corresponds roughly to the transition temperature, or melting temperature. The area below the peak is $\Delta H_d(T_m)$. The heat capacity of the unfolded state of a protein minus the heat capacity of the folded state is $\Delta C_{p,d}$. There is more about DSC in Chapter 5.

I. Energy conservation in the living organism

The First Law tells us that, if a system does work, w makes a negative contribution to ΔU; the system loses energy. This implies that not even the most sophisticated known "machine" – the human body, as far as we know – can do work without an energy source. And no matter how much the urge to eat might conflict with other ways we might rather spend time, there is no getting around having to eat – relatively often. But this does not necessarily mean that the First Law applies to living organisms.

In Chapter 1 we noted that calorimetry experiments on whole organisms were carried out as early as 1781 by Lavoisier and Laplace. They measured the heat given off by animals (and other objects) as the amount of water produced by melting ice, relative to a control in which no animal was present. The greater the volume of water at the end of the experiment, the greater the amount of heat given off during the experiment. Lavoisier and Laplace also collected and measured the gaseous "waste" from the animals used in their experiments. The quantity of heat and carbon dioxide produced by a guinea pig was compared with what was found for the combustion of carbon. Lavoisier later used the data from such experiments to establish that the combustion of food in animals leads to the production of heat, CO_2, and H_2O. About a century later, in 1904, a German physiologist named Max Rubner (1854–1932) reported on similar experiments with dogs. Rubner's work was effectively the final word on whether thermochemistry applied to physiology. For he was able to show that the heat production of a dog can be accounted for by the carbon and hydrogen balance of its respiration and the heats of combustion of fat, protein, and excrement. And on that cheerful note, we bring the text of this chapter to a close.

J. References and further reading

Atkins, P.W. (1998). *Physical Chemistry*, 6th edn, ch. 2. Oxford: Oxford University Press.

Atkinson, D.E. (1977). *Cellular Energy Metabolism and Its Regulation*. New York: Academic Press.

Bergethon, P.R. (1998). *The Physical Basis of Biochemistry: the Foundations of Molecular Biophysics*, ch. 11. New York: Springer-Verlag.

Blandamer, M.J., Cullis, P.M. & Engberts, J.B.F.N. (1995). Differential scanning and titration calorimetric studies of macromolecules in aqueous solution. *Journal of Thermal Analysis*, **45**, 599–613.

Burton, R.F. (1998). *Biology by Numbers: an Encouragement to Quantitative Thinking*, ch. 3.1–3.3. Cambridge: Cambridge University Press.

Cahan, D. (1993). *Letters of Hermann von Helmholtz to His Parents: The Medical Education of a German Scientist*, 1837–1846. Stuttgart: Franz Steiner Verlag.

Cahan, D. (2004). *An Institute for an Empire: The Physikalish–Technische Reichsanstalt*, 1871–1918. Cambridge: Cambridge University Press.

Christensen, H. N. & Cellarius, R. A. (1972). *Introduction to Bioenergetics: Thermodynamics for the Biologist: A Learning Program for Students of the Biological and Medical Sciences.* Philadelphia: W. B. Saunders.

Cooper, A., Eyles, S. J., Radford, S. E. & Dobson, C. M. (1991). Thermodynamic consequences of the removal of a disulfide bridge from hen lysozyme. *Journal of Molecular Biology,* **225**, 939–43.

Creighton, T. E. (1993). *Proteins: Structure and Molecular Properties*, 2nd edn, ch. 4.4.3. New York: W. H. Freeman.

Darveau, C.-A., Suarez, R. K., Andrews, R. D. & Hochachka, P. W. (2002). Allometric cascade as a unifying principle of body mass effects on metabolism. *Nature*, **417**, 166–70.

Dunham, W. (1994). *The Mathematical Universe: An Alphabetical Journey through the Great Proofs, Problems, and Personalities*, ch. H. New York: John Wiley & Sons.

Encyclopdia Britannica CD 98, "Drug," "Enthalpy," "Heat Capacity," "Heat of Reaction," "Heat Transfer," "Infection," "Internal Energy," "Kinetic Theory of Gases," "Latent Heat," "Maxwell–Boltzmann Distribution Law," "Principles of Thermodynamics," "Specific Heat," "Temperature," "Thermal Energy," "Thermocouple," and "Work."

Feynman, R. P., Leighton, R. B. & Sands, M. (1963). *Lectures on Physics*, vol. I, cc. 14–1, 44–1, 45–1 & 45–2. Reading, Massachusetts: Addison-Wesley.

Fruton, J. S. (1999). *Proteins, Enzymes, Genes: the Interplay of Chemistry and Biology*. New Haven: Yale University Press.

Ganong, W.F. (1989). *Review of Medical Physiology*, 13th edn, ch.17. New York: McGraw-Hill/Appleton & Lange.

Gillispie, C. C. (ed.) (1970). *Dictionary of Scientific Biography*. New York: Charles Scribner.

Gislason, E. A. & Craig, N. C. (1987). General definitions of work and heat in thermodynamic processes, *Journal of Chemical Education*, **64**, 660–8.

Harold, F. M. (1986). *The Vital Force: a Study of Bioenergetics*, ch. 1. New York: W. H. Freeman.

Haynie, D. T. (1993). *The Structural Thermodynamics of Protein Folding*, ch. 2. Ph.D. thesis, The Johns Hopkins University.

Haynie, D. T. & Ponting, C. P. (1996). The N-terminal domains of tensin and auxilin are phosphatase homologues. *Protein Science*, **5**, 2543–646.

Hewitt, P. G. (2006) *Conceptual Physics*, 10th edn. ch. 4, San Francisco: Pearson Addison-Wesley.

Ito, K. & Ito, T. (2005). Nonlinear dynamics of homeothermic temperature control in skunk cabbage, *Symplocarpus foetidus, Physical Review E*, **72**, 051909, 6 pages.

Ito, K., Ito, T., Onda, Y. & Uemura, M. (2004). Temperature-triggered periodical thermogenic oscillations in skunk cabbage (*Symplocarpus foetidus*); *Plant and Cell Physiology*, **45**, 257–64.

Johannessen, E. A., Weaver, J. M. R., Bourova, L., Svoboda, P., Cobbold, P. H. & Cooper, J. M. (2002). Micromachined nanocalorimetric sensor for ultra-low volume cell based assays. *Analytical Chemistry*, **74**, 2190–7.

Jones, C. W. (1976). *Biological Energy Conservation*. London: Chapman & Hall.

Katchalsky, A. & Curran, P. F. (1967). *Nonequilibrium Thermodynamics in Biophysics*, ch. 1. Cambridge, Massachusetts: Harvard University Press.

Klotz, I. M. (1986). *Introduction to Biomolecular Energetics*, ch. 1. Orlando: Academic Press.

Kondepudi, D. & Prigogine, I. (1998). *Modern Thermodynamics: from Heat Engines to Dissipative Structures*, ch. 2. Chichester: John Wiley.
Lawrence, C., Roger, A. & Compton, R. (1996). *Foundations of Physical Chemistry*. Oxford: Oxford University Press.
Lazarides, T., Archontis, G. & Karplus, M. (1995). Enthalpic contribution to protein stability: insights from atom-based calculations and statistical mechanics. *Advances in Protein Chemistry*, **47**, 231–306.
McNab, B.K. (2003). Metabolism: ecology shapes bird bioenergetics. *Nature*, **426**, 620–1.
Morowitz, H.J. (1978). *Foundations of Bioenergetics*, ch. 3. New York: Academic.
Microsoft Encarta 96 Encyclopedia, "Thermodynamics."
Pearson, H. (2004). Low-carb diets get thermodynamic defence. *Nature News*, 16 August.
Peusner, L. (1974). *Concepts in Bioenergetics*, ch. 2. Englewood Cliffs: Prentice-Hall.
Polanyi, M. (1946). *Science, Faith and Society*, ch. 1. Chicago: University of Chicago Press.
Price, G. (1998). *Thermodynamics of Chemical Processes*, cc. 1 & 2. Oxford: Oxford University Press.
Roberts, T.J., Marsh, R.L., Weyland, P.G. & Taylor, C.R. (1997). Muscular force in running turkeys: the economy of minimizing work, *Science*, **275**, 1113–15.
Schneck, D.J. (2006a). What is this thing called "me"? Part 4: The *buffered*, isothermal, living engine. *American Laboratory News*, January, 4–8.
Schneck, D.J. (2006b). What is this thing called "me"? Part 5 : The stationary, buffered, isothermal living engine. *American Laboratory*, May, 4–10.
Schneck D.J. (2006c). What is this thing called "me"? Part 6: The controlled, stationary, buffered, isothermal living engine. *American Laboratory*, November/December, 6–12.
Smith, C.A. & Wood, E.J. (1991). *Energy in Biological Systems*, cc. 1.2 & 1.3. London: Chapman & Hall.
Treptow, R.S. (1995). Bond energies and enthalpies: an often neglected difference. *Journal of Chemical Education*, **72**, 497–9.
van Holde, K.E. (1985). *Physical Biochemistry*, 2nd edn, ch. 1.1. Englewood Cliffs: Prentice-Hall.
Voet, D. & Voet, J.G. (1995). *Biochemistry*, 2nd edn, cc. 2-2, 3, 4, 11-2, 15-4–15-6, 16, 18-1, 19-1, 28-3 & 34-4B. New York: Wiley.
Watson, J.D. (1965). *The Molecular Biology of the Gene*. New York: Benjamin.
Westerterp, K.R. (2001). Pattern and intensity of physical activity. *Nature*, **410**, 539.
Williams, T.I. (ed.) (1969). *A Biographical Dictionary of Scientists*. London: Adam & Charles Black.
Zachariassen, K.E. (1985). Physiology of cold tolerance in insects. *Physiological Review*, **65**, 799–832.

K. Exercises

1. Invent three syllogisms. See Aristotle's *Topics* for ways of making use of syllogisms in formulating arguments.

2. Give the etymologies of *kinetic* and *potential*.

3. Give an example of a law of biology. What makes it a law?

4. Equation (2.1) involves a difference in internal energy. Differences in energy are much easier to measure than absolute magnitudes. Explain.

5. Figure 2.3 shows a thermodynamic cycle. The state function shown is U, though in principle a cycle of this sort could be given for any state function. Suppose that each arrow represents an experimental process, and that each internal energy represents an experimentally determined quantity. Give representative values for each energy change so that the condition $\sum_{\text{loop}} X = 0$ is satisfied.

6. The "Δ" in Eqn. (2.1) represents, effectively, a *measurable* change. What does this mean? Strictly speaking, the "Δ" should be used with state functions only; it should not be used to represent changes in q or w. Given this, and referring to Fig. 2.5, suggest a definition of *path function*. Does it follow that q (or w) can *never* be considered a state function? Why or why not?

7. Show that the right-hand sides of Eqns. (2.2) and (2.3) have the same dimensions.

8. We used Eqn. (2.2) to show that -1.96 kJ of work is done against gravity as a 50 kg woman climbs 4 m. Let the system be the woman. Evaluate ΔU? Explain how energy is conserved.

9. How many joules are expended by a 70 kg man climbing up 6 m of stairway? Does this quantity represent a maximum or minimum energy expenditure? Why? How much work is done if the climbing takes place on the surface of the moon? (Assume that the acceleration due to gravity on the moon's surface is 1.6 m s^{-2}.)

10. How many meters of stairway could a 70 kg man climb if all the energy available in metabolizing an 11 g spoonful of sugar to carbon dioxide and water could be converted to work?

11. A cylinder of compressed gas has a cross-sectional area of 50 cm^2. How much work is done by the system as the gas expands, moving the piston 15 cm against an external pressure of 121 kPa?

12. Indicate whether the temperature increases, decreases or remains the same in the following four situations: an endothermic/exothermic process in an adiabatic/non-adiabatic system. An adiabatic process is one in which no heat is exchanged with the surroundings.

13. A mathematical statement of the First Law of Thermodynamics is $\Delta U = q + w$. This holds for all processes. Assume that the only type of work done is pV-work. Show that $\Delta U = +w$ for an *adiabatic* process. Show that $\Delta U = 0$ for a process in an isolated system. Show that $\Delta U = q$ for a process that occurs at constant volume. Show that $\Delta H = 0$ for an adiabatic process at constant pressure.

14. When glucose is burned completely to carbon dioxide and water,

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O,$$

673 kcal are given off per mole of glucose oxidized at 25 °C. What is ΔU at this temperature? Why? What is ΔH at this temperature? Why? Suppose that glucose is fed to a culture of bacteria, and 400 kcal mol^{-1} of glucose is given off while the growing bacteria converted the glucose to CO_2 and H_2O. Why there is a discrepancy between the oxidation heats?

15. Conservation of energy is said to be implicit in the symmetrical relation of the laws of physics to time. Explain.

16. A person weighing 60 kg drinks 0.25 kg of water. The latter has a temperature of 62 °C. Assume that body tissues have a specific heat capacity of 0.8 kcal kg^{-1} K^{-1}. The specific heat of water is 1.0 kcal kg^{-1} K^{-1}. By how many degrees will the hot drink raise the person's body temperature from 37 °C? Explain how arriving at the answer involves the First Law of Thermodynamics.

17. Prove that Eqn. (2.14) follows from Eqn. (2.13).

18. Non-polar moieties in proteins make a positive contribution to $\Delta C_{p,d}$. This is known from measurements of the change in heat capacity of water on dissolution of non-polar compounds, e.g. cyclohexane. Is this true for polar moieties as well? What is the sign of their contribution to $\Delta C_{p,d}$? Explain your reasoning.

19. Early protein microcalorimetry studies were done by Peter Privalov, a Soviet biophysicist who emigrated to the United States in the early 1990s. One of the most thorough of all microcalorimetric studies of a protein is Privalov and Pfeil's work on hen egg white lysozyme, published in 1976. According to this work and later studies, $\Delta C_{p,d} = 1.5$ kcal mol^{-1} K^{-1}, and at pH 4.75, $\Delta H_d(25\,°C) = 52$ kcal mol^{-1}. Calculate the enthalpy difference between the unfolded and folded states of lysozyme at (a) 78 °C, the transition temperature, and (b) −10 °C. What is the physical meaning of ΔH in part (b)?

20. You have been asked to investigate the thermodynamic properties of a newly identified small globular protein by differential scanning calorimetry. The following results were obtained.

pH	T_m (°C)	$\Delta H_d(T_m)$ (kJ mol^{-1})
2.0	68.9	238
2.5	76.1	259
3.0	83.2	279
3.5	89.4	297
4.0	92.0	305
4.5	92.9	307
5.0	93.2	308
5.5	91.3	303
6.0	88.9	296
6.5	85.9	287
7.0	82.0	276
7.5	79.4	268
8.0	77.8	264

Plot $\Delta H_d(T_m)$ v. T_m. Describe the curve and rationalize its shape. Now plot $\Delta H_d(T_m)$ v. pH. What is happening?

21. ITC can be used to measure the enthalpy of protonation of amino acid side chains. Suppose three peptides were separately dissolved in weak phosphate buffer at pH 8 and injected into weak phosphate buffer at pH 2.5. There is a change in side chain ionization in going from one pH to the other. The peptides and the measured heats of reaction were Gly–Asp–Gly (-7.2 ± 0.8 μcal), Gly–Glu–Gly (-5.4 ± 0.8 μcal) and Gly–His–Gly (-5.5 ± 1.0 μcal). The data represent an average of 10 experimental data points, heat of injection minus background signal (injection of the pH 8 buffer into the pH 2 buffer in the absence of peptide). Gly = glycine, Asp = aspartate, Glu = glutamate, His = histidine. The peptide concentrations for the experiments were 0.64 mM, 0.57 mM and 0.080 mM, respectively. At pH 8, the side chains are approximately completely deprotonated, while at pH 2 they are approximately completely protonated. These solutions were injected into a sample cell in 10 μl aliquots. What is the physical basis of the background signal? What are the approximate protonation enthalpies of the Asp, Glu and His side chains? Suggest why tripeptides were used for these experiments rather than free amino acids. Would pentapeptides be any better? What could be done to account for the possible contribution of the terminal amino or carboxyl group?

22. Table C in Appendix C gives enthalpies of protonation for a number of popular biochemical buffers. Which five of these are likely to be best for thermodynamic measurements? Why?

23. The conditions of the standard state are chosen arbitrarily. What additional condition(s) might a biochemist add to those given in the text? Why?

24. Explain in structural and thermodynamic terms how the unfolding of a protein is like the melting of an organic crystal.

25. A protein called α-lactalbumin is a close homolog of hen egg white lysozyme. Unlike lysozyme, α-lactalbumin binds Ca^{2+} with high affinity. The measured enthalpy of binding, however, is much smaller in magnitude than the enthalpy of hydration. Explain.

26. Design a series of experiments to test whether the First Law of Thermodynamics applies to all living organisms.

27. Figure 2.7 shows that the enthalpy change on protein folding is large and positive. Suggest what gives rise to this.

28. Matter can neither be created nor destroyed, merely interconverted between forms. Discuss the statement in terms of the First Law of Thermodynamics.

29. Living organisms excrete the excess nitrogen from the metabolism of amino acids in one of the following ways: ammonia, urea, or uric acid. Urea is synthesized in the liver by enzymes of the urea cycle, excreted into the bloodstream, and accumulated by the kidneys for excretion in urine. The urea cycle – the first known metabolic cycle – was elucidated in outline by Hans Krebs and Kurt Henseleit in 1932. As we shall see in Chapter 5, urea is a strong chemical denaturant that is used to study the structural stability of proteins. Solid urea combusts to liquid water and gaseous carbon dioxide and nitrogen according to the following reaction scheme:

$$CO(NH_2)_{2(8)}(s) + 1.5O_2(g) \rightarrow CO_2(g) + N_2(g) + 2H_2O(l).$$

According to bomb calorimetry measurements, at 25 °C this reaction results in the release of 152.3 kcal mol^{-1}. Calculate ΔH for this reaction.

30. Giant sequoias, an indigenous species of California, are among the tallest trees on Earth. Some individuals live to be 3500 years old. Water entering at the roots must be transported up some 300 m of xylem in order to nourish cells at the top of the tree. Calculate the work done against gravity in transporting a single water molecule this distance.

31. Suggest three proofs that heat is not a fluid is the sense that liquid water is a fluid.

Chapter 3

The Second Law of Thermodynamics

A. Introduction

We have seen that a given amount of energy can be distributed in many different ways – something like how a certain volume of fluid can adopt many different shapes and adapt itself to its container. In this chapter we turn the spotlight on a thermodynamic function that enables us to *measure* how "widely" a quantity of energy is distributed.

The First Law of Thermodynamics relates heat, work and internal energy, and it tells us that energy is neither created nor destroyed in all its changes of form; the total energy of a reaction, and indeed of the universe, is *constant*. The First Law tells us with breathtaking generality that a boundary on the possible is a basic characteristic of our universe. It is not hard to see, though, that the First Law does not tell us some things we would like to know. For instance, if we put a "hot" system into contact with a "cold" one and allow them to come to thermal equilibrium, we find that the final temperature of the two objects, which will persist indefinitely if the combined system is isolated, is at some intermediate value. The value of ΔU for this reaction, however, which obviously proceeds spontaneously, is 0. Similarly, if we add a "concentrated" solution of substance A to a "dilute" solution of substance A, we find that the final concentration of the combined system, which will persist indefinitely if the system is isolated, is between the initial concentrations (Fig. 3.1). Again, $\Delta U = 0$, this time for spontaneous mixing. We see a similarity in behavior between heat energy and matter, and there is a correlation between $\Delta U = 0$ and spontaneity of reaction. Wait a minute. Won't $\Delta U = 0$ for a system that undergoes no change at all? In general the magnitude or sign of ΔU does not indicate the direction of spontaneous change!

Could we get our two objects in thermal equilibrium to return from being two warm objects to one that's hot and one that's cold? Could we get the solution at the intermediate concentration to spontaneously unmix and return to the concentrated solution and the dilute one? No! At least not spontaneously. For in both cases

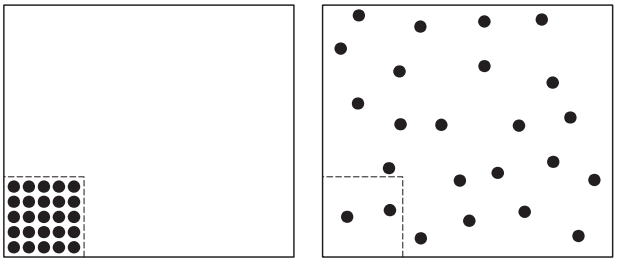


Fig. 3.1 Distribution of substance A before and after mixing. Panel (A) shows the situation before mixing, panel (B) when mixing is complete. Let substance A be perfume molecules on a heavily scented person who enters a small room, say, an elevator (lift). The molecules are initially very close together; the concentration of perfume is high. In such a high dose, it might be hard to distinguish perfume from insect repellent! Visual inspection of the wearer might not help much. After a while, maybe by the time you've reached the 57th storey, the perfume molecules are spread approximately randomly throughout the accessible space; the perfume is much more diffuse than at first; the concentration is uniform. This process is irreversible! The driving force for change is a movement of the system toward the most probable distribution of perfume molecules. Entropy, the key thermodynamic function of this chapter, measures the change in distribution. Further drops of understanding can be squeezed out of this figure. Imagine that the dots correspond not to molecules but to heat energy. Let the region where all the heat is collected in panel (A) be one object, and let the rest of the area be another object. The small object is hot, the big one is cold. Panel (B) shows the situation some time later, when the two objects have reached thermal equilibrium. (For example, 500 mL cans of Guinness Draught say that the contents should be chilled at 4 °C for 3 h, giving the gas plenty of time to dissolve in the liquid. The head is not quite so creamy on less cooling.) The heat energy has been redistributed throughout the matter, and both objects are at the same temperature. The driving force for change is a movement of the system toward the most probable distribution of heat energy. The comparison of perfume and heat suggests something exceptionally similar about matter and energy. Indeed, this relationship lies at the heart of $E = mc^2$.

something has been lost, *something* has changed – and the change is *irreversible*. The First Law, useful as it is, does not provide even the slightest clue about what that *something* is. Nor does it answer any of the following important questions: In which direction will a reaction proceed spontaneously? Is there a limit to how much work can be obtained from a reaction? If so, what is it? Why do highly purified enzymes degrade even when stored in the cold?

To be able to do more than just wave hands in positing answers to these queries, we must turn from the First to the Second Law of Thermodynamics. Like the First Law, the Second is an empirical result and a window on the relationship of heat to work. In still qualitative terms, the Second Law provides a way of describing the conversion of heat into work. It gives a precise definition of a thermodynamic state function called the entropy, and the sign of this function (plus or minus, not Leo or Libra!) tells us whether a process will occur spontaneously or not. This is something that no ΔU or w or q alone can do. Ice changes spontaneously into liquid water at 1 °C and 1 atm, despite the increase in translational motion (K.E.) of the molecules.

Our approach in this chapter is mainly that of "classical" thermodynamics. But we should be careful not to prejudge the discussion and think that its age makes it somehow less illuminating or useful. People who think that way are all too common, and few of them are good at thermodynamics. In learning a foreign language, one starts with relatively simple prose and not with poetry of the highest art! Moreover, this chapter is not an end in itself: it is a piece of the foundation for what comes next. In Chapters 4 and 5 we'll see the role played by the entropy in the Gibbs free energy, the biochemist's favorite thermodynamic function. And in Chapter 6 we'll turn our attention to statistical thermodynamics, and that will enable us to see how the statistical behavior of particles underlies the classical concept of entropy and other thermodynamic functions we'll have met by then.

We're all familiar with becoming aware of the cologne, aftershave, frou-frou juice, what-have-you someone nearby is wearing, or maybe the malodorous molecules underarm bacteria are pumping into the environment. The particles waft along an air current to the olfactory apparatus, gain entry, bind to receptors embedded in cell membranes, and thereby cause signals to be induced in brain-bound neurons. In some cases, for instance, when you walk by someone who is wearing a scent, it's your motion relative to that person that explains the phenomenon. But the relative motion of one person's neck to another's nose is not the essential ingredient of the present point. For on entering a room in which a heavily scented person has been present for a short while, you can smell the perfume immediately, even if that person's no longer present. What in the Milky Way does this have to do with thermodynamics?

The sweet-smelling volatile molecules in perfume may consist of various organic chemical components such as aromatic aldehydes. When heated by the body to 37 ^{o}C and exposed to the atmosphere, these molecules quickly become airborne. Then convection currents resulting from differences of temperature in patches of air play a role in spreading the cologne about. Finally, and most important here, the aromatic amines are bombarded constantly by the random motion of gaseous nitrogen and oxygen, and this moves them around a good deal – even in the absence of convection currents – by a process called *diffusion*. After some time, the concentration of perfume molecules will be approximately *uniform* throughout the room. Amazing! And no machine was required to achieve this end. Experience tells us that such behavior is the *only* thing that will happen (as long as there is no open window or something like that). That is, leaving a perfume-filled room and returning to find that all the aromatic aldehydes had, like some Arabian genie, somehow gone back into the bottle, seems extremely highly improbable. It *is* extremely highly improbable! As we shall see, diffusion is "governed" by the Second Law of Thermodynamics.

The Second Law is about the tendency of *particles* to go from being concentrated to being spread out in space – spontaneously. It is also about the tendency of *energy* to go from being "concentrated" to being

"spread out" – spontaneously. Consider a mass in a gravitational field, for example, a football that has been kicked high above the pitch. The driving force for change of the airborne ball is motion towards the most probable state, the state of lowest potential energy, the state of lowest energy of *position*. In this example, the state of lowest potential energy is just the ball at rest somewhere on the football field. The tendency of concentrated particles to become more uniformly dispersed is a reflection of the tendency of (chemical) energy to disperse itself into its most probable distribution, the state of lowest potential energy. We see this tendency in action on adding some cold cream to a cup of tea or coffee. Initially, the cream is seen to form distinct swirls, but before long the color and temperature of the liquid become more uniform. The Second Law, which helps to describe this process, is marvelously general: it applies not just to the mixing of cream and coffee but also (and equally well) to the spontaneous dissipation of aromatic aldehyde molecules from an open scent bottle, the spontaneous cooling of a hot saucepan when removed from the stove, the spontaneous folding of a polypeptide into a protein, the spontaneous movement of ions down their concentration gradient when a membrane channel opens.

Some students find the Second Law hard to grasp. One reason is that there are numerous formulations of the Law, and it's not always readily apparent that they are equivalent. It can be instructive to look at matter through the lenses of different observers, so to speak, as each sheds new light on the topic and helps to understand it in a different way. One of the earliest formulations of the Second Law from the historical point of view is of particular interest here because it helps to reveal the practical nature of the human activity out of which the science of thermodynamics developed. It is *impossible* for a system to turn a given amount of heat into an equivalent amount of work. In other words, if we put some quantity of heat q into the system, whatever work w is done by the system will be such that $w < q$. This comes to us from the work of the visionary French military engineer Nicolas Léonard Sadi Carnot (1796–1832). Publication of Carnot's *Réflexions sur la puissance motrice du feu et es machines propre à développer cette puissance*[1] (at the age of 28!) outlined a theory of the steam engine and inaugurated the science of thermodynamics. We shall encounter M. Carnot again shortly, but in a different chamber of the mansion of biological thermodynamics.

B. Entropy

The foregoing discussion brings us to the thermodynamic state function S, entropy (Greek, *en*, in + *trope*, transforming; also coined by Clausius). Being a state function, the entropy change for a process is independent of path, regardless of whether the change is

[1] *Report on the driving force of heat and the proper machines to develop this power.*

reversible or irreversible. The entropy is an index of the tendency of a system to undergo spontaneous change; it is a measure of the state of differentiation or distribution of the energy of the system. The entropy is the key to understanding energy transformation. As such, the entropy enables us to rationalize why solutes diffuse from a concentrated solution to a dilute one without exception, why smoke leaves a burning log and never returns, why wind-up clocks always run down, why magnets demagnetize spontaneously, why heat always flows from a hot body to a cold one. All this might suggest that entropy is something physical, as indeed many people have believed. It is important to realize, however, that entropy is not so much a "thing" as a highly useful mathematical object that provides insight to the nature of change in the material world.

As we have said, the entropy is a measure of the *order* of a system. For now, let a non-technical definition of *order* suffice; in Chapter 6, a more precise definition will be given. Entropy is less a "thing" than a way of describing how particles are *arranged* in space (e.g. perfume in the bottle or distributed throughout the room) and how particle arrangement changes as a system is subjected to changes of temperature, pressure, number of particles, volume, etc. The tendency toward a condition of *no further change* that we have seen in the examples above is a general property of thermodynamic systems. In fact, it is so fundamental to all of physics that most scientists consider the Second Law of Thermodynamics the most universal "governor" of natural activity that has ever been revealed by scientific study. Entropy measures how close a system is to the state corresponding to no further change, or equilibrium.

Let's avoid drifting into the ethereal world of abstractions and come back to Earth with another illustration, one that is closely related to the phenomenon of diffusion. Suppose we have two glass bulbs of equal volume connected by a stopcock, as shown in Fig. 3.2. Initially, the stopcock is closed. A gas occupies one bulb only; the other is evacuated. When the stopcock is opened, the gas molecules stampede into the evacuated bulb. There is a net flow of molecules into the formerly empty bulb until the concentration of molecules is identical (on the average) throughout the accessible volume. Such expansion of a gas is accompanied by an *irreversible* increase in entropy. The process is irreversible because a substantial amount of *work* would have to be done to herd all the gas molecules back into one bulb. The state in which all the molecules are distributed at random throughout the volume of the two bulbs – state 2 at equilibrium – is *less ordered* than the state in which all the molecules were randomly distributed throughout the volume of one bulb – state 1 at equilibrium. In this context, *equilibrium* just means a state of no further (net) change. Similarly, perfume is more ordered when it's still in the bottle than when it's been applied, volatilized by strenuous activity and increased body temperature, and scattered to the four corners of a room.

In Fig. 2.4C we saw how heat can be used to make a rubber band contract and lift a weight, to do work. Now let's see what insights

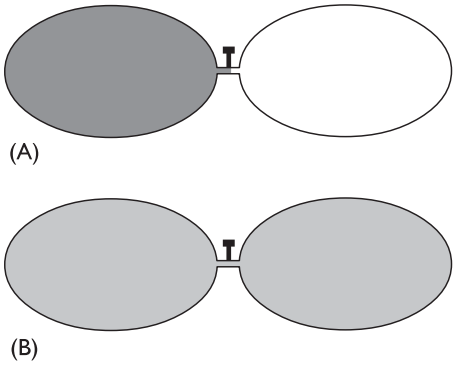


Fig. 3.2 Irreversible expansion of a gas. In panel (A), when the stopcock is closed, the gas is concentrated and confined to the left bulb. The right bulb is evacuated. When the stopcock is opened, as shown in panel (B), the gas flows rapidly from the bulb on the left to the bulb on the right. After a sufficiently long time, the condition known as equilibrium is reached; the concentration of gas is the same in both bulbs and the net flow of molecules between the bulbs is zero. This is just a more idealized view of the heavily scented person we met in the elevator.

can be gained by shrinking in size by several orders of magnitude and examining the mechanics of our rubber band machine. Rubber consists of long, chainlike molecules (Fig. 3.3). Stretching makes them align and become more orderly, decreasing the number of different spatial arrangements. Like forcing all the gas molecules back into one bulb, stretching a rubber band requires *work* to be done *on* the system, so $w > 0$. Your fingers, hands and arms together do the stretching; the energy comes from the cleavage of chemical bonds in your muscles. The contracting muscle is then used to stretch the rubber beyond its relaxed, equilibrium position, and there is a change in mechanical energy.

Box 3.1. The celebrated jumping insect of *Cantabrigiensis*

Insecta is the largest class of the phylum Arthropoda, which is itself the largest of the various phyla of the kingdom Animalia. Of all the animal species so far described by science, five of every six are insects. Beasts like bedbugs, beetles, and butterflies are a few of the more familiar members of the class. About 10^6 different insect species are known in total, and it is estimated that about as many more have yet to be described. Insects are poikilotherms: body temperature follows that of the surrounding environment within the range of tolerance for a species. Most of our knowledge of genetics has come not from humans but from experimental studies on the insect *Drosophila melanogaster*, one of about 10^3 different species of vinegar fly (more commonly but also misleadingly called fruit flies). Insects have jointed legs, and a segmented body covered by a hard exoskeleton (external skeleton) which is composed in part of the protein chitin. The exoskeleton serves as the point of attachment for muscles, which consist

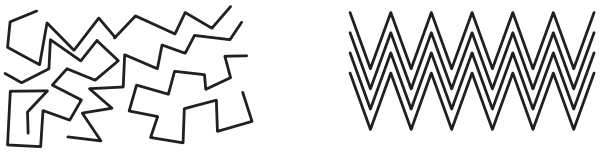


Fig. 3.3 Schematic diagram of a rubber band in the unstretched (equilibrium) state (left) and stretched state (right). In the unstretched state, the molecules are highly disordered. When the band is stretched, the molecules form a relatively orderly array. The entropy of the molecules is lower in the stretched state than in the unstretched state.

Box 3.1. Cont.

largely of the proteins actin and myosin. Locomotion is effected by muscles acting on the exoskeleton. In leaping insects (fleas, grasshoppers, locusts and the like) the force of muscle contraction compresses a "pad" of resilin. This protein is a member of a family of elastic proteins that includes elastin, gluten and spider silks. Resilin is also found in the sound-producing organs of cicadas. This long-chain molecule consists of many copies of a short elastic repeat sequence of amino acid residues joined in tandem. Chains are crosslinked between tyrosine residues, providing low stiffness, high strain, efficient energy storage, and a very high fatigue lifetime. The elasticity of resilin is thought to be the result of the extremely dynamic nature of amorphous hydrophobic regions which form a kinetically free, random-network polymer. At the microscopic level, rubbery materials like resilin become more ordered when extended, decreasing entropy, and they uncoil in elastic recovery. The energy stored in an elastic material can be transformed under stress into other forms of energy, for example, the K.E. of a projectile from a catapult. In the case of a cricket, the energy stored in coiled resilin is released by way of a catch mechanism that allows force generated before the jump to release rapidly; the projectile is the entire body of the bug. The current champion jumper of all species is the insect *Philaenus spumarius*, the froghopper (spittle bug), according to a recent report by a researcher at Cambridge University. The body of a froghopper accelerates at 2800–4000 m s^{-2} in a jump. The average jump requires about 50 μJ, which translates into a power output of 36 W g^{-1}. The force exerted in an average jump is 34 mN, about the same as the gravitational pull of the Earth on one-third of a gram of water at the planet's surface.

But how is work done by the *contraction* of rubber? Just as heating decreases the order of perfume molecules by bringing about a sort of thermodynamic diaspora, the addition of heat decreases the order of the rubber molecules and the rubber band contracts to its equilibrium position. Similar to muscle, the contraction of a rubber band permits work to be done. In this case, however, work is not done *on* the system, it is done *by* the system. We can *calculate* the magnitude of this work knowing gravitational acceleration, the mass of the object lifted by the rubber band,[2] and the distance the mass is moved against the force of gravity (Δx in Fig. 2.4C).

Stretching increases the order of the long, stringy molecules. In the taut state, the molecules make fewer random collisions with each other; the entropy of the system is reduced relative to the relaxed state. When the rubber band relaxes, the orderliness of the molecules is lost and the entropy increases. The closer the molecules are to being randomly ordered, the greater the entropy of the system. This study of rubber bands has helped to show how changes in entropy relate to transfers of heat, and heat is a form of energy related to molecular motion, as discussed in Chapter 2.

[2] By Newton's Second Law $\Delta \mathbf{p}/\Delta t = F = ma = mg$, where m is mass, a is acceleration, and g is the gravitational acceleration.

A related illustration. When a scented candle burns, the thermal energy of aromatic molecules embedded in the wax goes up. The greater the thermal energy of one of these molecules, the more easily it can overcome the attractive forces of other molecules in the melted wax, break forth from the surface, escape into the air, and after some time go spelunking in the cavernous recesses of your nasal passage. Blimey! Can this camp example bespeak aught of biology? The point is that the addition of heat increases the entropy of the system, the "breadth" of the distribution of the energy of the system. And in fact, both lords and ladies of the plant *Arum maculatum* attract pollinating insects to their flowers by heating and vaporizing aromatic molecules. At certain times of the day, the temperature of the part of the flower called the appendix increases rapidly, resulting in the evaporation of volatile compounds produced within. Insects "smell" the aromatic compounds, find them stimulating, and move up the concentration gradient until reaching the flower. Then, movement within the flower as the bug gathers up its prize leads to pollination.

Let's have another quick look at Fig. 2.4. Panel (A) shows heat being transferred from a hot system to a cold one. This is similar to Panel (C), where heat is being transferred from the candle to the rubber band. Back in Panel (A), kinetic energy is being shared as heat by means of multiple collisions between the molecules of the heat source and the molecules of the heat sink. A warm object is in contact with a cool one, and the warm molecules are banging on the cool ones harder than the cool ones are banging on the warm ones (or each other!). Heat transfer from the warm object to the cool one enables work to be done, as in the rubber band heat engine, and this situation continues as long as there is a temperature *difference* between the two systems.

When a thing is cold it moves relatively little. K.E. is proportional to thermal energy is proportional to absolute temperature. So addition of heat to something cold can disorder that thing more than the addition of the same amount of heat to the same thing at a higher temperature. On this basis we can guess that if a given amount of heat q is transferred to a system, it will increase the randomness of the system by an amount that is inversely proportional to the absolute temperature. Because we have described entropy as a measure of the breadth of a distribution of energy, we should expect $\Delta S \propto q/T$, though at this point we don't know if it's the first power of the temperature that's needed or something else.

Now let's become definite about the relationship between heat and entropy. The entropy change, ΔS, on heat transfer at *absolute* temperature T, is *defined* as

$$\Delta S \geq q/T, \tag{3.1}$$

where, by the convention we've adopted, $q > 0$ if heat is added *to* the system. Equation (3.1) is a fruit of labor on quantifying the maximum amount of work that can be obtained from an ideal reversible

engine. The equality holds only under an exceptional but none the less important constraint: when heat transfer is carried out *very* slowly and any change in the system is reversible, *i.e.*, when both the system *and* its surroundings can be returned to their *original* states. A reversible process is one that occurs through a succession of equilibrium or near-equilibrium states. The inequality corresponds to any other kind of process. No wonder we said that the expansion of a gas into vacuum is irreversible! Equation (3.1) tells us that although heat is a path function, the heat exchanged in a reversible *and* isothermal process is independent of path. So q can be independent of path but generally is not. S is always independent of path.

Equation (3.1) says that when a quantity of heat is transferred from a hot system to a cold one, $\Delta S_{\text{hot}} < 0$, $\Delta S_{\text{cold}} > 0$, $|\Delta S_{\text{hot}}| < |S_{\text{cold}}|$, and $\Delta S_{\text{total}} = \Delta S_{\text{hot}} + S_{\text{cold}} > 0$. Regardless of the magnitudes of q, T_{hot}, and T_{cold}, the *total* entropy change *must* be greater than zero. To make the discussion more concrete, let's look at an example. Suppose we wish to calculate the entropy change in the surroundings when 1.00 mol of $H_2O(l)$ is formed from H_2 and O_2 at 1 bar and 298 K. This is of interest here because liquid water is the only known matrix in which life occurs. We require the reaction to occur slowly. A table of standard thermodynamic quantities tells us that $\Delta H = -286$ kJ; the reaction is exothermic. Heat is transferred from the system to the surroundings, and $q_{\text{sur}} = +286$ kJ. Substituting this into Eqn. (3.1) and solving for ΔS_{sur} gives 286 kJ/298 K $= +959$ J K^{-1}. The entropy of the surroundings increases as heat is transferred *to* it. To put the numbers into perspective, 1 mol of water has a mass of 18 g and a volume of 18 ml, about the same as a large sip. Formation of a mole of water from hydrogen and oxygen at room temperature and ambient pressure increases the entropy of the universe by $\sim$1000 J K^{-1}.

C. Heat engines

This section describes how living organisms do *not* behave. You might guess therefore that this will be the usual sort of apparently useless, academic exercise that one is required to do to earn a degree. But in fact in trying to understand what something is, it can be instructive to seek to know *why* it is not what it is not. That's the spirit, anyhow, in which we discuss heat engines.

Let's suppose, as Carnot did, that heat q is transferred from a heat source to a heat sink (Fig. 2.4A). How much of this heat is available to do work? How much work can be done? No more than meets the requirement of the Second Law; the work done is usually much less than the Second Law allows! Let's calculate the limit. We set

$$\Delta S_{\text{hot}} + S_{\text{cold}} = 0. \tag{3.2}$$

Plugging in Eqn. (3.1), and calling the cold sink the place where the waste heat goes, we have

$$-q_{\text{transferred}}/T_{\text{hot}} + q_{\text{waste}}/T_{\text{cold}} = 0. \tag{3.3}$$

Rearranging,

$$q_{\text{waste}} = q_{\text{transferred}} T_{\text{cold}}/T_{\text{hot}}. \tag{3.4}$$

Here, q_{waste} is the *minimum* amount of heat transferred to sink; q_{waste} *cannot* be used to do work. In designing a heat engine, we would want to make q_{waste} as small as possible, making T_{cold} as small as possible and T_{hot} as large as possible, and remembering that energy must be "consumed" to make $T_{\text{cold}} < T_{\text{surroundings}}$ or $T_{\text{hot}} > T_{\text{surroundings}}$. The *maximum* work one can do is to use all the heat that remains, and that is $q_{\text{transferred}}$ less q_{waste}:

$$w_{\max} = q_{\text{transferred}} - q_{\text{transferred}} T_{\text{cold}}/T_{\text{hot}} = q_{\text{transferred}}(1 - T_{\text{cold}}/T_{\text{hot}}). \tag{3.5}$$

A simple numerical example is the following. If 30 J is transferred from a heat source at 300 K to a heat sink at 200 K (cold!), the maximum work that can be done is $30\ \text{J} \times [1 - (200\ \text{K} / 300\ \text{K})] = 10\ \text{J}$. The efficiency of this process $= w_{\max}/q_{\text{transferred}} = 10/30 = 33\%$. We can see now that *an engine in which all the heat is converted to mechanical work cannot exist* – a suggestion that seems suspiciously like a limitation on what is possible in our universe. Which brings us to another way of formulating the Second Law, one due to Carnot himself: *heat of itself cannot pass from a colder body to a hotter one*; work is required. What does this have to do with biology? It helps us to realize that cells cannot do work by heat transfer because they are essentially isothermal systems (Fig. 2.4B). This applies not only to terrestrial mammals like armadillos, but also to the hyperthermophilic bacteria that live on the ocean floor in thermal vents, and presumably to any living thing anywhere in the universe.

We have seen that the transfer of heat can be used to do work but that the process generates waste heat, q_{waste}. By Eqn. (3.1), $S_{\text{irreversible}}$, the *minimum* irreversible entropy produced by heat transfer, is

$$S_{\text{irreversible}} = q_{\text{waste}}/T_{\text{cold}}. \tag{3.6}$$

Does this mean that an irreversible entropy increase must be written off as a pure loss? No! To see how a loss might be a gain of a sort, let's pay another visit to our friend the rubber band. When stretched, its entropy is low; the long rubber molecules are ordered. The release of tension results in decreased ordering of the molecules, so $\Delta S > 0$. We should therefore expect $q > 0$ on release of tension, and this is easily verified by experiment (try it!). The heat q_{waste} is lost to the surroundings. This heat is not completely useless, however, because the contraction of the rubber could be used to do something constructive, for example, lift a weight (Fig. 2.4C). An irreversible increase in entropy can be used to do work.

Having covered the necessary background, let's look at a biological example of an irreversible increase in entropy being used to do

work. Grasshoppers (and other hopping insects) store elastic energy in the compressed form of a protein called resilin, from *resilient*. This is something like a compressed spring, for instance, the spring in a loaded jack-in-the-box. When the insect leaps, elastic energy is released and the resilin becomes less ordered. ΔS for this process is large and positive. This form of energy release is just about as fast as the transmission of a nerve impulse and much faster than a typical metabolic reaction, enabling the grasshopper to make tracks if it senses danger from a predator. You will certainly know something about this if you have ever tried to catch one of these crispy critters while it was, so to speak, in the pink.

Now, before we wander too far away from equations, we ask: which is greater in magnitude, $q_{\text{transferred}}$ or q_{waste}? Or, suppose we have a process that can be carried out either reversibly or irreversibly. For which process will q be larger? Combining Eqn. (3.1) with the First Law gives

$$\Delta U \leq T\Delta S + w, \tag{3.7}$$

which, upon rearrangement, becomes

$$w \geq \Delta U - T\Delta S. \tag{3.8}$$

The most negative value of w that this expression can yield, and therefore the greatest amount of work that can be done *by* the system, is

$$w_{\text{max}} = \Delta U - T\Delta S. \tag{3.9}$$

That is, the work done is maximal when the process is carried out reversibly. (Note that if $w_{\text{max}} = 0$ and $\Delta U = 0$, then $\Delta S = 0$ at any T.) By the First Law,

$$\Delta U_{\text{rev}} = q_{\text{rev}} + w_{\text{rev}} \tag{3.10}$$

for a reversible process, and

$$\Delta U_{\text{irrev}} = q_{\text{irrev}} + w_{\text{irrev}} \tag{3.11}$$

for an irreversible one. But if the starting and ending points are the same, then $\Delta U_{\text{rev}} = \Delta U_{\text{irrev}} = \Delta U$. And if work is done by the system on the surroundings, then the sign of w is negative, and Eqns. (3.10) and (3.11) are, respectively,

$$\Delta U = q_{\text{rev}} - w_{\text{rev}} \tag{3.12}$$

and

$$\Delta U = q_{\text{irrev}} - w_{\text{irrev}}. \tag{3.13}$$

Combining these equations, which we can do because the change of state is identical in the two cases, gives

$$\Delta U = q_{\text{rev}} - w_{\text{rev}} = q_{\text{irrev}} - w_{\text{irrev}} \tag{3.14}$$

or, upon rearrangement,

$$q_{\text{rev}} - q_{\text{irrev}} = w_{\text{rev}} - w_{\text{irrev}}. \tag{3.15}$$

Above we found that $w_{rev} \geq w_{irrev}$, which means that both sides of Eqn. (3.15) must be positive. This implies that

$$q_{rev} \geq q_{irrev}. \tag{3.16}$$

And so we have answered the question we set ourselves. What does Eqn. (3.16) mean? In an endothermic process, the heat extracted from the surroundings will be greatest when the process is reversible. In an exothermic process, the heat released to the surroundings will be smallest when the process is reversible. So, living organisms would release the smallest possible amount of energy as heat if the processes going on inside them were reversible. And we do release a lot of heat to the surroundings. But many of the processes going on inside you are irreversible! To keep from cooling down to the temperature of the surroundings, we must consume energy. And it comes from food.

D. Entropy of the universe

As we have seen, the total entropy of an isolated system increases in the course of a spontaneous change. Put another way, the Second Law says that no natural process can occur unless it is accompanied by an increase in the entropy of the universe (Fig. 3.4 and Table 3.1). The upshot is that every process that occurs in nature is ultimately irreversible and unidirectional, the direction being dictated by the requirement of an overall increase in entropy. This can be symbolized in a compact mathematical form as $\Delta S_{total} = \Delta S_{hot} + \Delta S_{cold} > 0$. Rewriting this in a more general way, we have

$$\Delta S_{system} + \Delta S_{surroundings} = \Delta S_{universe} > 0. \tag{3.17}$$

In order for a physical change to occur spontaneously, the entropy of the universe *must* increase. Leafing back to a preceding page of this chapter, we see that ΔS in Eqn. (3.1) is we see that $\Delta S_{universe}$, and in the previous two sections the entire universe consisted of just a heat source and a heat sink, one being the system and the other being the surroundings.

It is important to realize that Eqn. (3.17) does *not* say that entropically "unfavorable" reactions (ones for which the entropy change is negative) *cannot* occur. Such reactions can and do occur, albeit *not spontaneously*. When an entropically unfavorable process is made to occur, the *overall* change in the entropy of the universe *will* be greater than zero, by the Second Law of Thermodynamics. Will this be true for a rather complicated and organized thing like an amoeba, an ant, or an aardvark? Yes! And what if we want to *measure* the entropy production of an organism? Can we do that and decide whether the change occurred spontaneously? No! Because the sign of ΔS for a system indicates whether a reaction will proceed spontaneously *only* if the system is isolated from its surroundings or both the entropy

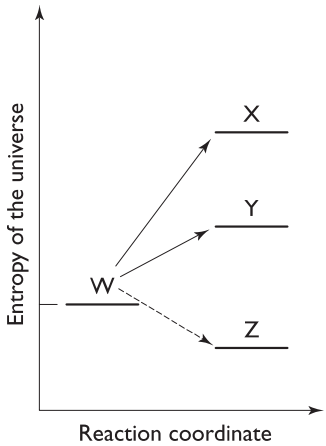


Fig. 3.4 The Second Law of Thermodynamics. No process will occur spontaneously unless it is accompanied by an increase in the entropy of the universe. This applies to an isolated system, a closed system, and an open system.

Table 3.1. *Comparison of the "orderliness" of different types of energy*

Form of energy	Entropy per unit energy
Nuclear reactions	10^{-6}
Internal heat of stars	10^{-3}
Sunlight	1
Chemical reactions	1–10
Terrestrial waste heat	10–100

Note how the entropy of a given amount of energy increases as it is transformed from a nuclear reaction to the heat given off by biological organisms on the surface of Earth.

change of the system *and* the surroundings have been measured. These conditions are not easily met. As we have seen, $q_p = T\Delta S$ *only* if a reaction is reversible. The decomposition of an organism into its various molecular components is irreversible! Also, we must consider how one might go about measuring the entropy change of the surroundings. There is nothing to say such measurements are impossible as a matter of principle, but in many cases they cannot be made in practice! The next chapter will show us a way of circumventing this obstacle, enabling us to determine whether the change in a system has occurred spontaneously. We'll do that by employing a less general but in many respects more useful index of reaction spontaneity, the Gibbs free energy.

E. Isothermal systems

Now we wish to look a little more closely at the specialized situation of constant temperature. Isothermal conditions are of great importance to the biochemist, because the temperature of a living organism is more or less uniform throughout its body, and it is practical for bench-top experiments to be done at constant temperature. In making a biochemical measurement, say of enzyme activity, one would not want to have to report that the temperature at which the reaction took place fluctuated over a range of a dozen degrees during the experiment.

Human beings and many other organisms can tolerate a change in body temperature of no more than a few degrees. We have terribly sophisticated negative feedback systems for thermoregulation and a circulatory system for distributing heat energy; when these systems are functioning normally they keep the entire body at about the same temperature. Now, consider a cell deep within your body to be a thermodynamic system. And suppose this cell undergoes a reversible change at constant pressure that transfers heat q to the surrounding cells of the tissue. Then

$$\Delta S_{\text{surroundings}} = -q/T, \tag{3.18}$$

where $T_{surroundings}$ is so close to T_{system} that both temperatures are T. Because $\Delta T \approx 0$, the pV-work that can be done from the heat transfer is practically negligible, and $q = \Delta H$. Substituting into Eqns. (3.17) and (3.18) gives

$$\Delta S_{system} - \Delta H/T > 0, \tag{3.19}$$

which, after rearrangement, can be written as

$$\Delta H - T\Delta S_{system} < 0. \tag{3.20}$$

The quantity $H - TS$ is a thermodynamic state function called the Gibbs free energy. ΔH measures the heat exchanged at constant pressure, as discussed in Chapter 2, and $T\Delta S$ can be thought of "isothermally unavailable" energy, as transfer of this energy out of the system would result in a temperature decrease, and we have said that the temperature is constant. The Gibbs energy enables prediction of the direction of spontaneous change of a system under the constraints of constant temperature and constant pressure.

We can approach the idea of a portion of heat released by a reaction being unavailable for work from another direction. First, let's see if we can convince ourselves (again) that some of the available energy cannot be used to do work. Suppose we have a system which undergoes a process that results in a decrease in the entropy of the system, for example, liquid water allowed to come to equilibrium in the solid state at -1 °C. If all the heat released by this process, which is about 6 kcal mol^{-1}, were exported to the surroundings as *work*, there would be *no* increase in the entropy of the surroundings, and the overall entropy for the process would be negative. But this contradicts Eqn. (3.17), the requirement of the Second Law that any real process will result in an increase in the entropy of the universe. So, at least some of the heat generated by the process must be unavailable to do work. But how much?

We can find the answer by continuing with a qualitative treatment. Let's suppose that the reaction reduces the entropy of the system by some amount, ΔS. In order for the reaction to be spontaneous, the entropy of the surroundings must increase by at least as much. By Eqn. (3.6), this is q_{waste}/T. Solving for the heat supplied, $q_{waste} = T\Delta S$, where ΔS is the *decrease* in the entropy of the system (q_{waste} is negative). This is the energy that *is not* available to do work. The energy that *is* available to do work is the difference between the total energy and the energy that is not available to do work. And this energy is the Gibbs free energy. Note the resemblance of this qualitative example to Eqn. (3.20). The Gibbs free energy is so important to the biological sciences that the next two chapters will be devoted to it.

Above we combined the First and Second Laws to arrive at the Gibbs free energy. Now we wish to combine these laws again, but in a somewhat different way. The result will be different, and it will provide new insight into the expansion of a gas into vacuum

discussed above and prepare the way for a discussion of osmosis in Chapter 5. Moreover, it will provide a good starting point for our coverage of statistical thermodynamics (Chapter 6). Let's require the system to be adiabatic; energy can neither enter nor leave, $\Delta U = 0$. By the First Law, $q = -w$. If we require pV-type work only, then $q = p\Delta V$; and if the system is an ideal gas, then $q = nRT\Delta V/V$. Suppose now that the process we'd like to carry out is reversible. Then, by the Second Law, $q = T\Delta S$. Combining the First and Second Laws gives

$$T\Delta S = nRT\Delta V/V. \tag{3.21}$$

Cancelling the Ts leaves

$$\Delta S = nR\Delta V/V. \tag{3.22}$$

This is the increment in entropy when an ideal gas is allowed to expand by a small but measurable volume, ΔV. If the expansion is carried out reversibly over a large change in volume, the total change in entropy is the sum of all the small changes, and the result, which can be found with a little calculus, is

$$\Delta S = nR\ln(V_f/V_i), \tag{3.23}$$

where V_f is the final volume and V_i is the initial volume. The entropy is a state function, so as long as the initial and final states are whatever they are, the entropy difference between states will be independent of how the change occurs. The entropy change on opening the stopcock of an adiabatic version of the system shown in Fig. 3.2 can be found by Eqn. (3.23). It is $\Delta S = nR\ln(2V_i/V_i) = nR\ln 2$. The entropy change of the universe for irreversible change must be greater than the entropy change of the universe of a reversible change.

F. Protein denaturation

Let's see how Eqn. (3.18) (or Eqn. (3.1)) can be used to describe the reversible isothermal entropy change of any biochemical system we like. Here, we apply it to protein denaturation, but it describes equally well the "melting" of DNA, the dissociation of the double-stranded helix into two single strands. In Chapter 4 we shall see how it is yet even more general.

Rewriting Eqn. (3.18) with symbols introduced in the previous chapter, we have

$$\Delta S_d = \Delta H_d/T_m. \tag{3.24}$$

The minus sign has vanished because heat is being transferred *to* the system and we are describing the entropy change of the protein system, not the surroundings. Suppose T_m is 342 K and we wish to

know ΔS_d at 25 °C. What can we do? If the heat transfer is carried out reversibly, then, by Eqns. (3.1) and (3.18), $\Delta S = -q/T$. The heat transferred, q, will increase the enthalpy of the system according to Eqn. (2.16), if the system is at constant pressure. Combining these equations gives

$$\Delta S = C_p \Delta T/T. \tag{3.25}$$

If we now sum up all these small contributions to find the entropy change over a measurable range of temperatures (T_1 – T_2) and use a wee bit of mathematical wizardry (the same sort used in the previous section), the result is

$$\Delta S(T_2) = \Delta S(T_1) + C_p \ln(T_2/T_1), \tag{3.26}$$

where $\Delta S(T_i)$ is the change in entropy evaluated at temperature T_i, not the mathematical product of ΔS and T_i, and we have assumed that C_p is constant throughout the temperature range. It can be shown that Eqn. (3.26) becomes

$$\Delta S(T_2) = \Delta S(T_1) + \Delta C_p \ln(T_2/T_1), \tag{3.27}$$

if state 1 and state 2 differ in heat capacity and ΔC_p is constant throughout the relevant temperature range. As an example, suppose that $\Delta S = 354$ cal mol^{-1} K^{-1} at 80 °C and $\Delta C_p = 1500$ cal mol^{-1} K^{-1}. Then $\Delta S(25\,°C) = 354$ cal mol^{-1} K^{-1} + (1500 cal mol^{-1} K^{-1}) × ln(298.16 K/353.16 K) = 100 cal mol^{-1} K^{-1}.

There is another way of thinking about proteins and entropy, one that does not involve a large change in heat capacity. As the pH decreases, acidic amino acid side chains become protonated. From an exercise in Chapter 2, the enthalpy change of amino acid side chain protonation is about 1 kcal mol^{-1}. This is so small as to be negligible in comparison with the enthalpy change of unfolding the protein in the absence of protonation effects. Changes in pH can nevertheless have a dramatic effect on protein stability; indeed, we have already seen how lowering the pH reduces the transition temperature of hen egg white lysozyme. It follows that protonation of Glu and Asp is mainly an *entropic* effect, with regard to the binding of the proton to the amino acid side chain *and* the effect on protein unfolding. The T_m of the protein decreases upon reduction of pH because the entropy difference between the folded and unfolded states decreases at a faster rate than the enthalpy difference, making $\Delta H_d/\Delta S_d$ progressively smaller. The ionization of food molecules in the low-pH environment of the gut denatures proteins, facilitating their breakdown into short peptides by digestive proteases.

How do we *interpret* the pH denaturation of proteins, and explain it on a more detailed level? As the pH goes *down* there is a change in the ionization state of the *acidic* side chains. This results in a net

increase in the surface charge of the protein. So at low pH, the positive charges repel each other by electrostatic interactions more than at high pH, destabilizing the folded conformation. The situation can be represented as follows:

$$P_{folded} + nH^+ \rightarrow P^* \rightarrow P_{unfolded}, \tag{3.28}$$

where P* is an unstable folded conformation. The effect of charge on protein stability was first described mathematically in 1924 by the Danish physical biochemist Kaj Ulrik Linderstrøm-Lang (1896–1959). The earliest known experiments on the use of a strong acid to denature proteins were done about a century earlier, by a Dutch chemist named Gerardus Johannes Mulder (1802–1880), who is also said to have stimulated the eminent Swedish chemist Jöns Jacob Berzelius (1779–1848) to coin the word *protein* (Greek, of the highest importance). Linderstrøm-Lang's mentor at Carlsberg Laboratory, Copenhagen, Denmark, was the Danish biochemist Søren Peter Laurtiz Sørensen (1868–1939), who developed the now universally adopted pH scale for measuring the acidity of an aqueous solution and ushered in the modern era of protein chemistry. We shall meet Linderstrøm-Lang again in Chapters 6 and 8.

G. The Third Law and biology

Yet another way of stating the Second Law brings us to the Third Law. *Any system not at absolute zero has some minimum amount of energy that is a necessary property of that system at that temperature*. This energy, of magnitude *TS*, is the "isothermally unavailable" energy from above (Section E). Now, the Third Law of Thermodynamics states that the entropy of a *perfect* crystal is *zero* when the absolute temperature is *zero* (0 K = −273 °C = cold!). A perfect crystal is like the ideal diamond in which each atom is at its proper place in an orderly array. The reason why we care about the Third Law is that it implies that the rapid and complex changes exhibited by living organisms, for instance, in a eukaryotic cell migrating during embryogenesis, can only occur far from thermodynamic equilibrium.

There is a substantial scientific literature on the freezing of *living* organisms. Some creatures have been stuck in a rather sophisticated sort of meat locker, taken to an extremely low temperature (~4 K), and allowed to thaw again. It is remarkable that some relatively "simple" organisms, e.g. bacteria, some types of microscopic animals, and plant seeds, return to room temperature from the deadly cold and function normally. Even some nematodes, which are comparatively complex organisms, having on the order of 10^3 cells at adulthood, are known to be able to withstand this process (depending on how it is carried out).

As discussed in Chapter 2, temperature measures the average kinetic energy of a collection of molecules. So, when the temperature is made to approach 0 K, all molecular motion ceases (excluding that

required by the Heisenberg uncertainty principle[3]). Thus, close to absolute zero, the only "memory" a biological system has of its life-before-deep-freeze is the *information* contained in the structure and arrangement of its macromolecules. When the organism thaws out, no new information is added; in fact, information is removed, because heating is a disordering process, as we have seen throughout this chapter.

The foregoing discussion would suggest that "all" one would have to do to "create" a cell would be to construct a being of the appropriate configuration of atoms. If the configuration (structure) were "the right type," the cell would function "on its own." From this point of view, it seems that a cell of an organism or indeed an entire organism might not be qualitatively different from some other collection of organic molecules. But on another view, a cell is simultaneously the most highly organized and the most complex collection of organic molecules of its size that we can imagine! Will it be possible to make "artificial" cells? What will be their properties?

H. Irreversibility and life

In the first several chapters of this book we have looked at ways in which living organisms can be thought of as machines. (Note: this does not mean that living organisms *are* machines!) Back in Chapter 1, for example, we discussed energy "consumption" as though biological machines really do "consume" energy. We've covered the First and Second Laws in some degree of depth, and we want to take a more critical look at energy consumption and relate it to life. If food is potential energy and we consume food, we consume energy; a sound argument if we agree on the meaning of *consume*. There is, however, another way of looking at energy consumption, and that is what we want to do now.

We do of course consume food, but that should not be taken to mean that we *consume* energy. For all a living organism or any type of system whatsoever can do is *transform* energy from one form to another; the total energy of the universe remains the same throughout such transformations by the First Law. The amount of energy returned to the environment by an organism, for instance as excretory products or heat, is equivalent in magnitude to the energy taken in, assuming no change in the weight. In this sense, living things do not consume energy at all; energy simply *flows through* them.

Just as important, the energy an organism returns to the environment *must* be less *useful* than the energy it "consumed."

[3] Named after the German mathematical physicist and philosopher Werner Karl Heisenberg (1901–1976), son of a professor of ancient history. Heisenberg was awarded the Nobel Prize in Physics in 1932.

Sure, excretory products make great fertilizer, but there are several good reasons why certain animals would not want to feed on them! As we have seen, any real process *must* increase the entropy of the universe; any change in the universe must result in an overall decrease in order. And as we shall see in the next chapter, *biologically useful* energy, or free energy, is the energy that can be used to do work under isothermal conditions. It is the sort of energy that humans, sheep, goats, even sea slugs, need to live. As the example of the heat engine has shown, heat transfer cannot be used to perform a substantial amount of work in biological systems, because all parts of a cell and its surroundings are effectively at the same temperature (and pressure). We have thus eliminated a major class of ways in which cells could conceivably do work, at least within the constraints of the physical properties of our universe. And we have narrowed the path to understanding. Whew! In playing the game of twenty questions, the ideal strategy is to pose a query whose yes-or-no reply will eliminate the largest number of possible answers and enable you to close in on the right one. The direction of spontaneous change in an isothermal system from a non-equilibrium to an equilibrium state is determined by the requirement that the extent of change be a maximum at every point on the reaction pathway. The suggestion that, say, uniformly dispersed smoke particles could somehow move spontaneously from all corners of a room back into a burning cigar seems absurd – except in a videotape run backwards.

There are various causes of the irreversibility of real-world processes. These include friction between two objects during relative motion, unrestrained expansion of a gas or liquid without production of work, the mixing of different substances that would require the input of work to separate them – all common phenomena. Because all atoms interact with each other, even noble gases, it would appear that there must be at least a small amount of irreversibility in any actual process. The inexorable increase in the entropy of the universe resembles, broadly speaking, the unidirectional flow of time. For as far as anyone knows, time moves in one direction only: forward! Why should this be so remarkable?

Time moves forward, and the past is, well, the past. This apparent conjunction of an interpretation of a scientific theory (the Second Law of Thermodynamics) and our ordinary (psychological?) perception of time is all the more intriguing because *all* organisms come into being and pass out of being *in time* and *all* the fundamental laws of physics are *time-reversible*.[4] Newton's laws of motion work equally well in either direction of time; they are time symmetrical. Maxwell's equations of electromagnetism work equally

[4] The decay of kaons and other sub-nuclear particles violates time symmetry; these particles appear to possess an intrinsic "sense" of past-future. See Christenson *et al.* (1964).

well forwards and backwards.[5] The time-dependent Schrödinger[6] equation of quantum theory is equally happy whether time is positive or negative. Einstein's theory of relativity works just as well in either direction of time. The time-reversibility or time-symmetry of laws of physics is related to energy conservation (the First Law).

The widely accepted mathematical formulations of physical law help us to rationalize many aspects of the nature of the universe and, moreover, provide tools for the creation of technology. Because of this, we cannot but be convinced that physics gives us at least an approximately right sense of the nature of reality. Nevertheless, and regardless of one's familiarity with physics, time marches on. The only law of physics that jibes with this aspect of our everyday experience of the world is the Second Law of Thermodynamics. This is all the more noteworthy here in that life on Earth has grown increasingly complex since the advent of the first cell; humans, composed as they are of billions of cells, are a good deal more complex than one-celled beasts like bacteria! We'll come back to this point in Chapter 9.

One can think about the irreversibility of chemical processes and life on different levels. Just as the increase in complexity of life forms on Earth is irreversible, in that it cannot be undone (though we could, it seems, blast ourselves to oblivion by means of well-placed and sufficiently large nuclear bombs): so at certain points in the development of an organism "commitment" occurs. For instance, in higher eukaryotes, once embryonic cells have "differentiated" into mesoderm or ectoderm, they ordinarily do not and in many cases apparently cannot become endoderm. If you have had the extreme misfortune of losing a limb, you will be acutely aware of the fact that a new one won't grow in to take its place. Some researchers think that biological ageing can be described in terms of the Second Law. On this view, what we call ageing is the process whereby a biological system moves from a point far from equilibrium toward equilibrium, a state of no further change. Another way of stating this is that order is a basic property of a living organism, and disorder, a dead one. There is a great deal that could be said on this topic, and it is a pity that there is not enough space to do more with it here.

The concepts of entropy and irreversibility (and energy conservation) have had a huge impact on humankind's view of the universe. Indeed, the concept of entropy has thrown into high relief philosophies of progress and development. "How is it possible to understand life

[5] James Clerk Maxwell, a Scot, lived 1831–1879. He is regarded as the nineteenth-century scientist who had the greatest influence on twentieth-century physics and is ranked with Isaac Newton and Albert Einstein for the fundamental nature of his contributions. He did important work in thermodynamics and the kinetic theory of gases.

[6] The Austrian physicist Erwin Schrödinger (1887–1961) was awarded the Nobel Prize in Physics in 1933. His little book *What is Life?* had a significant impact on the early development of molecular biology.

when the entire world is ordered by a law such as the second principle of thermodynamics, which points to death and annihilation?"[7] It is hard to see how a definitive answer can be given. Again, this topic deserves far more attention than it can be given here.

Finally, we wish to touch on the origin of irreversibility in many-body systems like large collections of small interacting particles. The thermodynamic description of such systems is so useful precisely because, in the usual case, there is no detailed knowledge or *control* over the (microscopic) variables of position and momentum for each individual particle. If such control were possible, the dynamics of many-body systems would presumably be reversible. When the number of microscopic variables is large, the state of maximum entropy is *overwhelmingly* probable, and the only lack of *certainty* that the entropy is maximal is the requirement that statistical "fluctuations" be allowed to occur. Under given constraints, the maximum entropy (equilibrium) state is the macroscopic state that can be formed in the greatest number of microscopic ways. More on this in Chapter 6.

I. References and further reading

Allen, A. (1983). Entropy still in a state of disorder. *Trends in Biochemical Sciences*, **8**, 81–2.

Atkins, P. W. (1994). *The Second Law: Energy, Chaos, and Form*. New York: Scientific American.

Atkinson, D. E. (1977). *Cellular Energy Metabolism and Its Regulation*. New York: Academic Press.

Barón, M. (1989). With Clausius from energy to entropy. *Journal of Chemical Education*, **66**, 1001–4.

Bennet, C. H. (1987). Demons, engines and the Second Law. *Scientific American*, **257**, no. 5, 108–12.

Bent, H. A. (1965). *The Second Law*. New York: Oxford University Press.

Bergethon, P. R. (1998). *The Physical Basis of Biochemistry: the Foundations of Molecular Biophysics*, cc. 12.1–12.2. New York: Springer-Verlag.

Burrows, M. (2003). Froghopper insects leap to new heights. *Nature*, **424**, 509.

Cano, R. J. & Borucki, M. K. (1995). Revival and identification of bacterial spores in 25- to 40-million-year old Dominican amber. *Science*, **268**, 1060–4.

Christensen, H. N. & Cellarius, R. A. (1972). *Introduction to Bioenergetics: Thermodynamics for the Biologist: A Learning Program for Students of the Biological and Medical Sciences.* Philadelphia: W. B. Saunders.

Christerson, J. H., Cronin, J. W., Fitch, V. L. & Turlay, R. (1964). Evidence for the 2π decay of the K_2^0 meson. *Physical Review Letters*, **13**, 138–40.

Clugston, M. J. (1990). A mathematical verification of the second law of thermodynamics from the entropy of mixing. *Journal of Chemical Education*, **67**, 203–5.

Craig, N. C. (1988). Entropy analyses of four familiar processes. *Journal of Chemical Education*, **65**, 760–4.

[7] Léon Brillouin, *Life, Thermodynamics, and Cybernetics*.

Cropper, W. H. (1988). Walther Nernst and the last law. *Journal of Chemical Education*, **64**, 3–8.

Davies, P. W. C. (1995). *About Time: Einstein's Unfinished Revolution*. London: Penguin.

Djurdjevic, P. & Gutman, I. (1988). A simple method for showing that entropy is a function of state. *Journal of Chemical Education*, **65**, 399.

Einstein, A. (1956). *Investigations on the Theory of Brownian Movement*. New York: Dover.

Elvin, C. M., Carr, A. G., Huson, M. G., Maxwell, J. M., Pearson, R. D., Vuocolo, T., Liyou, N. E., Wong, D. C. C., Merritt, D. J. & Dixon, N. E. (2005). Synthesis and properties of crosslinked recombinant pro-resilin. *Nature*, **437**, 999–1002.

Encyclopdia Britannica CD 98, "Absolute Zero," "Elasticity," "Entropy," "Insects," "Metabolism," "Onsager, Lars," "pH," "Reversibility," and "Principles of Thermodynamics."

Entropy: An International and Interdisciplinary Journal of Entropy and Information Studies. See *http://www.mdpi.org/entropy/*.

Fenn, J. B. (1982). *Engines, Energy and Entropy*. New York: W. H. Freeman.

Feynman, R. P., Leighton R. B. & Sands, M. (1963). *Lectures on Physics*, Vol. I, cc. 44–2–44–6. Reading, Massachusetts: Addison-Wesley.

Fruton, J. S. (1999). *Proteins, Enzymes, Genes: the Interplay of Chemistry and Biology*. New Haven: Yale University Press.

Gardner, M. (1967). Can time go backward? *Scientific American*, **216**, no. 1, 98–108.

Gillispie, C. C. (ed.) (1970). *Dictionary of Scientific Biography*. New York: Charles Scribner.

Gutfreund, H. (1951). The nature of entropy and its role in biochemical processes. *Advances in Enzymology*, **11**, 1–33.

Hale, F. J. (1993). Heat engines and refrigerators. In *Encyclopedia of Applied Physics*, vol. 7, ed. G. L. Trigg, pp. 383–403. New York: VCH.

Harold, F. M. (1986). *The Vital Force: a Study of Bioenergetics*, ch. 1. New York: W. H. Freeman.

Haynie, D. T. (1993). *The Structural Thermodynamics of Protein Folding*, ch. 3. Ph.D. thesis, The Johns Hopkins University.

Hollinger, H. B. & Zenzen, M. J. (1991). Thermodynamic irreversibility: 1. What is it? *Journal of Chemical Education*, **68**, 31–4.

Holter, H. & Møller, K. M. (ed.) (1976). *The Carlsberg Laboratory 1876–1976*. Copenhagen: Rhodos.

Katchalsky, A. & Curran, P. F. (1967). *Nonequilibrium Thermodynamics in Biophysics*, ch. 2. Cambridge, Massachusetts: Harvard University Press.

Klotz, I. M. (1986). *Introduction to Biomolecular Energetics*, ch. 2. Orlando: Academic Press.

Kondepudi, D. & Prigogine, I. (1998). *Modern Thermodynamics: from Heat Engines to Dissipative Structures*, cc. 3 & 4. Chichester: John Wiley.

Maddox, J. (1993). When entropy does not seem extensive. *Nature*, **365**, 103.

Microsoft Encarta 96 Encyclopedia, "Thermodynamics."

Morowitz, H. J. (1970). *Entropy for Biologists*. New York: Academic Press.

Morowitz, H. J. (1978). *Foundations of Bioenergetics*, cc. 4 & 5. New York: Academic Press.

Ochs, R. S. (1996). Thermodynamics and spontaneity. *Journal of Chemical Education*, **73**, 952–4.

Parker, F. L. & Krenkel, P. A. (1970) *Physical and Engineering Aspects of Thermal Pollution.* Cleveland, Ohio: CRC Press.

Penrose, O. (2005). An asymmetric world. *Nature*, **438**, 919.
Peusner, L. (1974). *Concepts in Bioenergetics*, cc. 3 & 10. Englewood Cliffs: Prentice-Hall.
Planck, M. (1991). The Second Law of Thermodynamics. In *The World Treasury of Physics, Astronomy, and Mathematics*, ed. T. Ferris, pp. 163–9. Boston: Little Brown.
Price, G. (1998). *Thermodynamics of Chemical Processes*, ch. 3. Oxford: Oxford University Press.
Prigogine, I. (1967). *Introduction to Thermodynamics of Irreversible Processes*. New York: John Wiley.
Schrödinger, E. (1945). *What is Life? The Physical Aspect of the Living Cell*. Cambridge: Cambridge University Press.
Seidman, K. & Michalik, T. R. (1991). The efficiency of reversible heat engines: the possible misinterpretation of a corollary to Carnot's theorem. *Journal of Chemical Education*, **68**, 208–10.
Skoultchi, A. I. & Morowitz, H. J. (1964). Information storage and survival of biological systems at temperatures near absolute zero. *Yale Journal of Biology and Medicine*, **37**, 158.
Smith, C. A. & Wood, E. J. (1991). *Energy in Biological Systems*, cc. 1.2, 1.3 & 2.3. London: Chapman & Hall.
Williams, T. I. (ed.) (1969). *A Biographical Dictionary of Scientists*. London: Adam & Charles Black.
Wrigglesworth, J. (1997). *Energy and Life*, cc. 1.3–1.4.1. London: Taylor & Francis.

J. Exercises

1. Is the word *entropy* a misnomer? Why or why not?

2. State whether the following phrases pertain to (A) the First Law of Thermodynamics, (B) the Second Law, (C) both the First and Second Law, or (D) neither of the Laws.

(1) Is concerned with the transfer of heat and the performance of work.
(2) Is sufficient to describe energy transfer in purely mechanical terms in the absence of heat transfer.
(3) Indicates whether a process will proceed quickly or slowly.
(4) Predicts the direction of a reaction.
(5) Is a statement of the conservation of energy.
(6) Says that the capacity to do work decreases as the organization of a system becomes more uniform.
(7) Is a statement of the conservation of matter.
(8) Says that a quantity of heat cannot be converted into an equivalent amount of work.
(9) Says that the capacity to do work decreases as objects come to the same temperature.

3. Examine Eqn. (3.1). What happens to ΔS as $T \rightarrow 0$? In order to ensure that this equation remains physically meaningful as

$T \to 0$, what must happen to ΔS? The answer to this question is a statement of the Third Law of Thermodynamics.

4. Consider a heat engine. Suppose 45 J is transferred from a heat source at 375 K to a heat sink at 25 °C. Calculate the maximum work that can be done and the efficiency of the process.

5. We said that heat engines do not tell us very much about how living organisms work. Show that if the human body depended on *thermal* energy to do work, it would cook before it could demonstrate its efficiency as a heat engine. Assume that the "engine" has an efficiency of 20%.

6. One calorie (1 cal) is produced for every 4.1840 J (Joules) of work done. If 1 cal of heat is available, can 4.1840 J of work be accomplished with it? Why or why not?

7. In Chapter 2 we learned about thermal equilibrium. In the approach to thermal equilibrium when two objects of differing initial temperature are brought into contact, although no energy is lost (by the First Law of Thermodynamics), *something* certainly is lost. What is it?

8. Entropy change of protein unfolding. Suppose that $\Delta H_d(25\ ^\circ C) = 10$ kcal mol^{-1}, $T_m = 68\ ^\circ C$ and $\Delta C_p = 1{,}650$ cal mol^{-1} K^{-1}. Calculate $\Delta S_d(T_m)$, $\Delta S_d(37\ ^\circ C)$ and $\Delta S_d(15\ ^\circ C)$. At what temperature is $\Delta S_d = 0$? Give the thermodynamic significance of ΔS_d in molecular terms at each temperature.

9. Recall Exercise 19 from Chapter 2. Use the same data to evaluate $\Delta S_d(T_m)$ at each pH value. Rationalize the entropy values.

10. For irreversible pathways, q/T is generally dependent on the path. How can one discover the entropy change between two states? Knowing that $q_{reversible} > q_{irreversible}$ (Eqn. (3.16)), use the First Law to write down a similar inequality for $w_{reversible}$ and $w_{irreversible}$.

11. Explain in thermodynamic terms why water freezes.

12. Suppose you have a cyclic process, as shown in Fig. 2.3. The entropy change for the system must be 0. Is there any inconsistency with the Second Law of Thermodynamics? Explain.

13. In his book *What is Life?*, Erwin Schrödinger says "an organism feeds with negative entropy." What does he mean? (Hint: consider an organism that is able to maintain its body temperature and weight in an isolated system.)

14. Consider a gas, a liquid and a crystal at the same temperature. Which system has the lowest entropy? Why?

15. Can a machine exist in which energy is continually drawn from a cold environment to do work in a hot environment at no cost? Explain.

16. There are a number of different causes of undernutrition. Some of these are: failure of the food supply; loss of appetite; fasting and anorexia nervosa; persistent vomiting or inability to swallow; incomplete absorption, comprising a group of diseases in which digestion and intestinal absorption are impaired and there is excess loss of nutrients in the feces; increased basal metabolic rate, as in prolonged fever, overactivity of the thyroid gland, or some cancers; and loss of calories from the body; *e.g.* glucose in the urine in diabetes. Rationalize each type of undernutrition in terms of the First and Second Laws of thermodynamics.

17. The macroscopic process of diffusion can be identified with microscopic Brownian motion,[8] which subjects molecules to repeated collisions with the atoms of their environment and results in their random rotation and translation. Some people say that the time-asymmetry in the inevitable increase of randomness of the universe is not strictly true as Brownian motion may contravene it. What is your view? Support it with well-reasoned arguments.

18. Consensus is a weak but nonetheless important criterion of truth, particularly in the scientific community. Doig and Williams[9] claim that disulfide bonds make a substantial contribution to the enthalpy change of protein unfolding. Their view is rejected by most researchers who study protein thermodynamics. In the light of the results of the study by Cooper *et al.*[10], and considering the structure of a disulfide bond, rationalize the long-standing view of the scientific community to the thermodynamic role of disulfide bonds in proteins.

19. The Gibbs paradox: consider two gas bulbs separated by a stopcock. The stopcock is closed. Both bulbs are filled with the same inert gas at the same concentration. What is the change in entropy when the stopcock is opened?

20. Is it possible for heat to be taken in to a system and converted into work with no other change in the system or surroundings? Explain.

21. Organisms are highly ordered, and they continually create highly ordered structures in cells from less-ordered nutrient

[8] Brownian motion is named after the Scottish botanist Robert Brown, who was the first to observe it, in 1827.

[9] See Doig, A. J. & Williams, D. H. (1991). Is the hydrophobic effect stabilizing or destabilizing in proteins – the contribution of disulfide bonds to protein stability. *J. Mol. Biol*, **217**, 389–98.

[10] See Cooper, A., Eyles, S. J., Radford, S. E. & Dobson, C. M. (1992). Thermodynamic consequences of the removal of a disulfide bridge from hen lysozyme, *J. Mol. Biol*, **225**, 939–43.

molecules. Does this mean that organisms violate the Second Law of Thermodynamics? Explain.

22. The process whereby the Earth was formed and living organisms grew increasingly complex with time is "essentially irreversible," says Thomas Huxley. It "gives rise to an increase of variety and an increasingly high level of organization." Thus, this process appears not to square with the Second Law of thermodynamics. Explain.

23. It would appear that all living organisms on Earth are, essentially, isothermal systems. Relatively few organisms live where the surroundings are at a higher temperature than they are. Rationalize this observation in thermodynamic terms.

24. Tube worms thrive at black smokers at the bottom of the ocean. These invertebrates live as long as 250 years, longer than any other known spineless animal. Tubeworms have no mouth, stomach, intestine, or way to eliminate waste. The part of the worm that produces new tube material and helps to anchor the worm in its protective tube, a chitin proteoglycan/protein complex, is often planted deep within the crevices of a black smoker. The soft, bright-red structure (made so by hemoglobin) at the other end of the worm serves the same purpose as a mouth and can be extended or retracted into the surrounding water. Giant tubeworms are over 1 m long, and they have to cope with a dramatic temperature gradient across their length. The temperature at a worm's plume is about 2 °C, just above the freezing point of pure water at 1 atm, while that at its base is about 30 °C! Can tube worms be modeled as isothermal systems? Why or why not?

25. Individual model hydrogen bond donors and acceptors do not often form hydrogen bonds in aqueous solution. Why not?

26. You may have noted that Carnot's formulation of the Second Law of Thermodynamics involves a very bold and unusually strong word: "impossible." Is this always true? Why or why not?

27. The contraction of rubber is largely an entropic phenomenon. What are the sources of the enthalpic component?

28. Recall the example used to illustrate the entropy change in the surroundings when a mole of liquid water is formed from molecular hydrogen and molecular oxygen at 298 K. Use the data given in the text to calculate the entropy change per water molecule formed.

29. Protonation of the side chains of Glu and Asp is mainly an entropic effect. Why is this not true of His as well?

30. Show that when a system gains heat reversibly from surroundings held at constant temperature, there is no change in entropy.

31. "The entropy change during an irreversible process is higher than the entropy change during a reversible process." Is the statement true? Under what conditions?

32. What bearing does the Second Law have on pollution? (See pp. 241, 247 of Peusner (1974).)

33. Discuss Fig. 1.5 in terms of the concepts of Chapter 3.

Chapter 4

Gibbs free energy – theory

A. Introduction

This chapter discusses a thermodynamic relationship that provides a basis for explaining spontaneous chemical reactivity, chemical equilibrium, and the phase behavior of chemical compounds. The relationship involves a thermodynamic state function that enables prediction of the direction of a chemical reaction *at constant temperature and pressure*. The constraints of fixed T and p might seem annoyingly restrictive, because they are less general than the requirements of the Second Law, but in fact the gains made on imposing the constraints will outweigh the losses. How is that? One reason is at any given time an individual organism is practically at uniform pressure and temperature (but be sure to see the Exercises at the end of the chapter). Another is that constant temperature and pressure are the very conditions under which nearly all bench-top biochemistry experiments are done. Yet another is that, although the total entropy of the universe must increase in order for a process to be spontaneous, evaluation of the *total* entropy change *requires* measurement of both the entropy change of the system *and* the entropy change of the surroundings. Whereas ΔS_{system} can often be found without too much difficulty, albeit only indirectly, $\Delta S_{\text{surroundings}}$ can be *hard* to measure! How could one measure the entropy change of the rest of the universe? The subject of the present chapter provides a way around the difficulty.

A particularly clear example of the inadequacy of ΔS_{system} to predict the direction of spontaneous change is given by the behavior of water at its freezing point. Table 4.1 shows the thermodynamic properties of water for the liquid $\rightarrow$ solid phase transition. The decrease in internal energy (which is practically identical to the enthalpy as long as the number of moles of gas doesn't change; see Chapter 2) would suggest that water freezes spontaneously in the range 263–283 K. Going on internal energy alone, spontaneous freezing would seem even more probable at $+10\ ^{\circ}$C than at $-10\ ^{\circ}$C,

Table 4.1. *Thermodynamics of the liquid → solid transition of water at 1 atm pressure*

Temperature (°C)	ΔU (J mol^{-1})	ΔH (J mol^{-1})	ΔS (J mol^{-1} K^{-1})	$-T\Delta S$ (J mol^{-1})	ΔG (J mol^{-1})
−10	−5619	−5619	−21	5406	−213
0	−6008	−6008	−22	6008	0
−10	−6397	−6397	−23	6623	−226

Table 4.2. *Sign of ΔG and direction of change*

Sign of ΔG	Direction of change
$\Delta G > 0$	The forward reaction is energetically unfavorable, the reverse reaction proceeds spontaneously
$\Delta G = 0$	The system is at equilibrium, there is no further change
$\Delta G < 0$	The forward reaction is energetically favorable, the forward reaction proceeds spontaneously

because ΔU for this system becomes increasingly negative, meaning that the internal energy of the system decreases, with increasing temperature. The entropy change too is negative at all three temperatures, consistent with the solid state being more ordered than the liquid one. So the sign and magnitude of the entropy of the *system* does not predict the direction of spontaneous change (unless the system is isolated).

In contrast to ΔU (which is hardly different from ΔH in this case) and ΔS, the last column, ΔG, matches what we know about the physical chemistry of water: below 0 °C, it freezes spontaneously ($\Delta G < 0$), at 0 °C solid water and liquid water coexist ($\Delta G = 0$), and above 0 °C, ice is unstable ($\Delta G > 0$). *ΔG is negative for what we know is a spontaneous process, and it is positive for the reverse process* (Table 4.2). As we know from experience, a stretched rubber band will contract when released. What is the sign of ΔG for this process? Negative! ΔS is large and positive, making $-T\Delta S$ negative, and ΔH is negative. More generally, *$\Delta G < 0$ is a basis for explaining chemical reactivity, equilibrium, and phase behavior*. Providing a good understanding of the Gibbs free energy and how the biochemist can use it is one of the most important purposes of this book.

The thermodynamic state function of chief interest in this chapter is G, the Gibbs free energy. This quantity is an eponym of Josiah Willard Gibbs (1839–1903),[1] the American theoretical physicist and

[1] Gibbs was the fourth child and only son of Josiah Willard Gibbs, Sr, professor of sacred literature at Yale University. In 1863 the younger Gibbs became the first person to receive the doctorate in engineering in the USA. Gibbs never married, lived with his sister, and spent all of his life in New Haven, Connecticut, apart from

chemist who was the first to describe it. Like its cousins U and H, which we met in Chapter 2, G is measured in units of joules. If all these state functions have the same units, what distinguishes G from U and H? What *sort* of energy is the Gibbs free energy?

Free energy is energy that is available in a form that can be used to do work. You should not find this statement terribly surprising. Remember from the last chapter that *some* energy is *not* free to do work; heat transfer *always* generates waste heat, and waste heat *cannot* be used to do work. The Gibbs free energy measures the maximum amount of work that can be done by a process going from non-equilibrium to equilibrium (at constant temperature and pressure). The First and Second Laws place boundaries on what is possible in extremely general terms. The Gibbs free energy tells us how much work can be done by a system under the constraints of the First Law, the Second Law, constant temperature, and constant pressure. But there is no law of free energy!

Like H and U, G is defined for *macroscopic* systems, ones that involve a very large number of particles. This implies that, although measurement of ΔG does tell us how much work must be done to convert one state of a system into the other, or how much work could be done by a process, it does not explain *why* that much work should done. Nevertheless, thermodynamics is often all the more useful for its very power in dealing with systems described in qualitative of terms. For instance, a microcalorimeter enables the enthalpy change of a process, say, the unfolding of a protein, to be measured directly and with extraordinary accuracy, regardless of how little or much one knows about the structure of the protein. And a bomb calorimeter can be used to measure the heat of combustion of a beignet from Café du Monde in New Orleans or a *crème brûlé* from the Latin Quarter in Paris without knowing the slightest thing about ingredients, shape, structure of the molecules involved, or the nature of the interactions in either case. Indeed, one need not have heard of beignets to record good combustion data! A biophysicist can measure the heat effects of pure protein or DNA molecules whose three-dimensional structures are not known in detail. When a molecular interpretation of thermodynamic quantities is needed, one turns to a branch of physical chemistry called statistical mechanics (Chapter 6).

G is a thermodynamic *potential function*. As such, it is analogous to the gravitational potential function of classical mechanics, which describes how the gravitational energy of an object varies with position in a gravitational field. If you take a coin out of your pocket and let it go, it will change position spontaneously and rapidly! The coin falls because its gravitational potential energy is greater in the air than on

three years in Europe shortly after the American Civil War. He was appointed professor of mathematical physics at Yale University in 1871. Gibbs is arguably the most famous American-born scientist to date, owing partly to the promotion of his work by Maxwell.

the ground, and air gives relatively little resistance to a change of position. The coin moves down its gradient of gravitational potential. A potential function, like G, permits prediction of whether a system will change or stay the same under given conditions.

Of course any change in a system will happen at some rate – some extent of change per unit time. As we shall see, though, you don't need to know the rate of a reaction in order to say whether it will occur spontaneously. And the rate of a reaction is no predictor of the energy difference between reactants and products. But let's keep things simple and focus on energy differences here and tackle reaction rates in Chapter 8.

Finally, the mathematics of this chapter will be a little rougher ride than earlier on, so fasten your seatbelt. Don't worry, though, because we'll find that we can make the complete journey without anything more perplexing than algebraic equations and the odd logarithm. The ideas themselves will always be considered more important than the specific mathematical tricks by which the results are obtained. This is not to imply that the mathematics is unimportant; it's to put our priorities in proper order. Mathematics may be Queen of the Sciences, but Applied Mathematics is definitely Handmaid of the Sciences. Those who are prepared for a more demanding mental journey and wish to explore an all-terrain-vehicle treatment of the subject might consider consulting the more advanced references at the end of the chapter.

B. Equilibrium

We turn now to the very important concept of chemical equilibrium. It was first proposed by the Norwegian chemists Cato Maximilian Guldberg (1836–1902) and Peter Waage (1833–1900) in the 1860s, in the form of the law of *mass action*: when a system is at equilibrium, an increase (decrease) in the amount of reactants (products) will result in an increase (decrease) in the amount of products (reactants). An equilibrium system responds to change by minimizing changes in the relative amounts of reactants and products. For example, suppose we have a solution of our favorite protein molecule. At equilibrium some molecules will be in the folded state (reactant), some in the unfolded state (product). Now add a dash of protease. In the usual case, unfolded proteins are much more susceptible to proteolytic attack than folded proteins. If a proteolyzed protein is unable to refold, proteolysis will change the balance of folded proteins and unfolded ones. By the law of mass action, the response of the system to the decrease in the number unfolded proteins will be for folded proteins to unfold, in order to minimize the change in the relative amounts of reactants and products. Equilibrium is such an important concept that the Swedish chemist Svante August Arrhenius (1859–1927) called it "the central problem of physical chemistry." That was in 1911, eight years after

he had won the Nobel Prize in Chemistry. There will be more on Arrhenius in Chapters 8 and 9.

A detailed discussion of equilibrium was avoided in previous chapters for the simple reason that neither the First Law nor the Second Law depends on it. Another reason is that no living organism functions and practically no real process occurs under such conditions! But this hardly implies that thinking about systems at equilibrium is unimportant. That's because abstractions and idealizations play an extremely important role in scientific study, serving as models of reality or simple generalizations of otherwise complex phenomena. A basic idea of thermodynamics is that *any* physical system will *inevitably* and *spontaneously* approach a stable condition called equilibrium. This concept is bound up in the Second Law; it resembles the proposed relationship between entropy and ageing that we touched at the close of the previous chapter.

A system will exhibit a net change in time if it is *not* at equilibrium, even if the rate of change is imperceptible. An example of a system that is not at equilibrium is a pet parrot. It's prudent to keep it caged to prevent it becoming predator Pussy's supper, but unless Polly's fed, its protected perch won't lengthen its lifespan. This is because the metabolic reactions of the body require a continual input of chemical energy, and when the energy requirements are not met, the body winds down and dies. All living organisms are highly non-equilibrium systems. All have a tendency to decay. The tendency cannot be overcome without energy. Input panels (A) and (B) of Fig. 4.1 are clearly not at equilibrium; the fluid height is not level. There will be a net flow of liquid from one side to the other. The flow rate in one direction will be greater than the flow rate in the opposite direction. The system is at equilibrium in panel (C). There is no net flow. The flow rates in opposite directions are equal.

A system will not show net change if it is *at equilibrium* and *left unperturbed*. For instance, a plugged test-tube filled with a

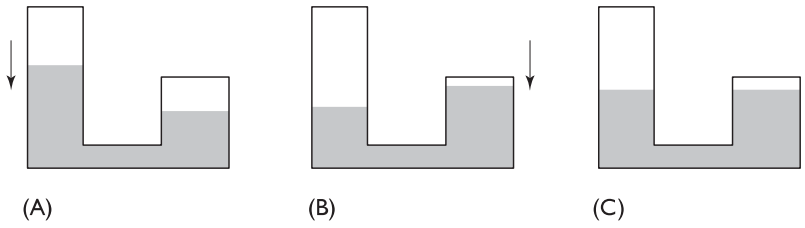


Fig. 4.1 Movement of a liquid system toward equilibrium. In panels (A) and (B) the system is not at equilibrium; we know from experience that change will occur. The driving force for change in this case is a difference in hydrostatic pressure, which is related to a difference in gravitational potential energy. The pressure difference is proportional to the difference in the height of the fluid in the two arms of the vessel. Water flows downhill! The rate of flow in one direction is greater than the rate of flow in the opposite direction. The system will continue to change until the fluid level is the same on both sides of the vessel. In panel (C), the flow rate is the same in both directions; the system is at equilibrium; no further change occurs.

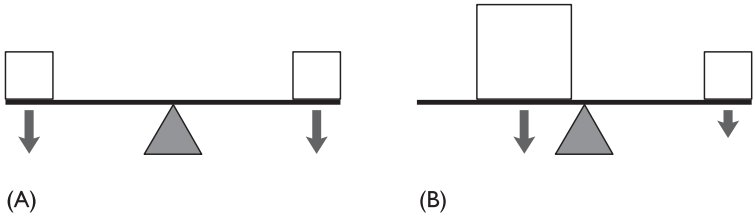


Fig. 4.2 Equilibrium. We know from experience that if the weights are the same, as in panel (A), the distance of each weight from the fulcrum must be the same. But if the weights are different, as in panel (B), the distance from the fulcrum cannot be the same. The lighter weight must be farther from the fulcrum than the heavier weight. By Newton's Second Law, at equilibrium the clockwise torque equals the counterclockwise torque, where torque = mass × gravitational acceleration × distance from fulcrum.

biochemical buffer and kept at constant temperature will not change (barring bacterial contamination and chemical degradation, of course). The system is in a rather stable equilibrium. A less stable equilibrium might be a heavy snow lying peacefully on a mountainside. Demonstration of the tenuous nature of this equilibrium could be regrettable for anyone or anything nearby: for a slight disturbance could turn a huge pile of resting snow into a raging avalanche! Another example of an unstable equilibrium is a balanced seesaw. The weights on either side of the fulcrum need not be equal, but if they are unequal, adjustments must be made in their distances from the fulcrum to achieve a balance, as in Newton's laws of mechanics (Fig. 4.2). What might be called a semi-stable equilibrium is one in which the energy of the system is at a minimum but not at the lowest possible minimum. For example, suppose a rock is resting at the bottom of a ravine between two hills. The rock will not leap up spontaneously and move elsewhere! If given a hard enough kick, though, it might roll up the side of the ravine, reach the crest of a hill, and tumble down the hill into the valley below. Various types of equilibrium are summarized in Fig. 4.3. The equilibrium state is one in which no further macroscopic change takes place because all forces acting on the system are balanced.

Let's look at the relationship of diffusion to equilibrium. Diffusion, a type of transport process, is important in many different biochemical reactions, for example, the chemical transmission of nerve impulses across synapses. Figure 4.4 shows one of many ways in which diffusion can be illustrated. We have two miscible liquids, and for the sake of argument we suppose that one can be layered aside the other, giving a sharp boundary, until we say "go." The top solution is colorless, like water, the bottom one opaque, like India ink. The densities are assumed to be about the same. Just after mixing has begun, which occurs spontaneously, there is a large concentration gradient across the boundary. But we know that this situation is unstable and will not persist. The result after a long time is that the dark liquid will be distributed uniformly throughout the combined volume.

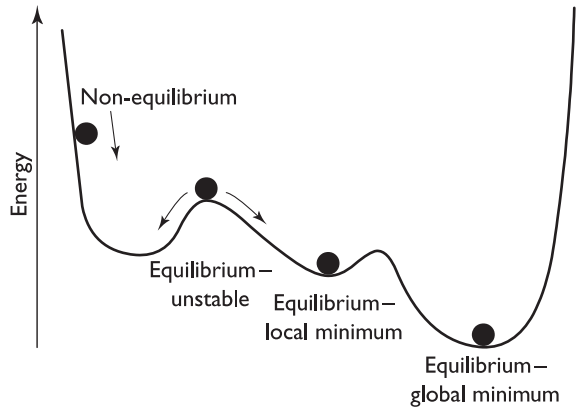


Fig. 4.3 Different types of equilibrium. The non-equilibrium situation will change immediately. The unstable equilibrium, like the snow at rest on a mountainside, will not change without the addition of energy. But the amount of energy needed to get over the barrier is very small. There are two main types of energy minimum: local minimum and global minimum. A local minimum has the lowest energy in the vicinity. The global minimum has the lowest free energy of all. An ongoing debate in protein folding research is whether the folded state of a protein corresponds to a local or a global free energy minimum. In some case, the functional state of a protein or enzyme might be a kinetically trapped conformation.

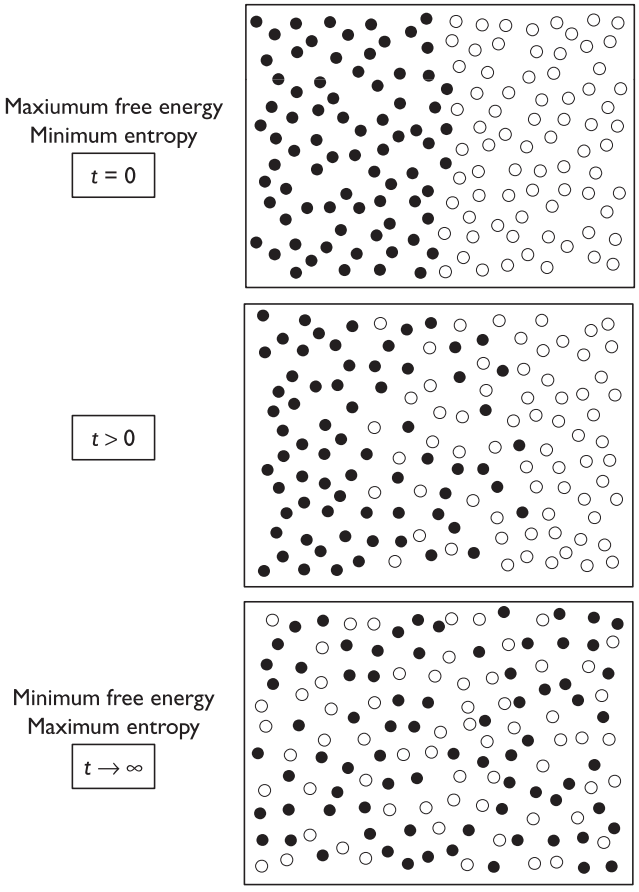


Fig. 4.4 Mixing. At time $t = 0$, when the partition is removed, the liquids are completely separate. The liquids are miscible so mixing occurs. After a sufficiently long time, the liquids are completely mixed. The unmixed state has a maximum of free energy for the system and a minimum of entropy. The completely mixed state has a minimum of free energy and a maximum of entropy. Based on Fig. 1.5 of van Holde (1985).

Figure 4.4A corresponds to the minimum entropy and maximum free energy of the system, while Fig. 4.4C, representing equilibrium, corresponds to the maximum entropy and minimum free energy of the system. Equilibrium will be achieved only if the temperature of the system is uniform throughout. For when the temperature is not uniform, convection currents and differences in particle concentration will be present, as for example in the swirls one sees shortly after pouring milk into a cup of coffee or tea. The convection currents and inhomogeneities will eventually go away, and when they have vanished, then the system will be at equilibrium. If we have two systems of identical molecules A and B, and the systems are in thermal equilibrium, then the distribution of the kinetic energy of the molecules of one system is identical to that of the other system. Some of the molecules will be moving very rapidly, others not so fast, but the distributions will be the same (Fig. 1.9). At equilibrium, the kinetic energy distribution of the molecules is one of maximum probability; the entropy, both of the individual systems and in combination, is a maximum; and the free energy, both of the individual systems and in combination, is a minimum.

An exercise from the previous chapter asked you to show that under the three constraints of *pV*-work only, constant internal energy, and reversibility, $\Delta S = 0$ at any temperature. We will now build on this foundation to develop the concept of equilibrium. Recall that when we required all change to a system to be reversible, we were saying that the system had to be at (or very close to) equilibrium throughout the process. In this context we interpret $\Delta S = 0$ to mean that the entropy must be at an extreme value when equilibrium is reached. By the Second Law, we know that the extremum is a maximum and not a minimum – the entropy of the universe always increases. In any non-equilibrium state of an isolated system, $S < S_{\max}$, and the system will change *spontaneously* until $S = S_{\max}$. Once equilibrium has been reached, there is no further increase in the entropy, $\Delta S = 0$. Somehow or other, a system "knows" when to stop "generating" entropy. This provides yet another way of stating the Second Law: an *isolated* system will change spontaneously until a maximum state of disorder is obtained.

Consider the reaction $\mathrm{Y} \Leftrightarrow \mathrm{Z}$. How far will the reaction go? Until equilibrium is reached, until there is no longer a tendency for a (macroscopic) change to occur spontaneously. At equilibrium, the average concentrations of Y and Z are constant in time; there is no further *macroscopic* change to the system. But this does not mean that the particles are no longer moving! For as long as the temperature is above 0 K, all the particles will be in motion, the amount depending on the temperature and whether the substance is in the solid, liquid or gas phase. If follows that at equilibrium Y and Z can and will interconvert, even if this happens only very slowly. Chemical equilibrium is a *dynamic equilibrium*. But the *concentrations* of Y and Z will not change – on the average.

If a process is made to occur through a succession of near-equilibrium states, the process must be slow, allowing the system to come to equilibrium after each small change. Real processes necessarily occur at a finite rate, so the best one can achieve in practice is a near-reversible process. Many biological macromolecules, for example small proteins, exhibit highly reversible order–disorder transitions ($> 95\%$) on thermal or chemical denaturation. It would appear in such cases that all the information required for the protein to fold to its native state is present in the matter it's made of, i.e. in the amino acid sequence. We'll return to protein folding in later chapters.

There are a number of other important features of the equilibrium state. One is that *for a system truly to be at equilibrium, it must be closed*. For example, if an unopened Pepsi can has reached thermal equilibrium, the contents will be at equilibrium. The amount of carbonic acid present and the amount of gaseous CO_2 will be constant, even if individual CO_2 molecules are constantly escaping from the liquid and returning thither. On opening the can, there will be a very rapid change in pressure and a jump to a non-equilibrium state. And a net loss of CO_2 will ensue for three main reasons. One, the system will no longer be closed, and gaseous CO_2 will escape immediately. Two, the decrease in abundance of CO_2 gas near the liquid–gas interface will promote the loss of CO_2 from the liquid. And three, if the can is held in your hand and not kept on ice, its contents will begin warming up, and this will drive off CO_2 because the solubility of the gas varies inversely with the temperature. We infer from all this that for a system to remain at equilibrium, variables such as T, p, V, and pH must be constant. For if any of them should change, or if the concentration of any component of the system should change, a non-equilibrium state would result, and the system as a whole would continue to change until equilibrium was reached.

C. Reversible processes

Now it's time for a mathematical statement of the Gibbs free energy:

$$G = H - TS. \tag{4.1}$$

We see that G is a sort of combination of the First and Second Laws, as it involves both enthalpy and entropy. We must bear in mind, though, that the temperature and pressure are constant for G to predict the direction of spontaneous change of a system.

What can be done with Eqn. (4.1)? For an incremental measurable change in G,

$$\Delta G = \Delta H - T\Delta S - S\Delta T. \tag{4.2}$$

If T is constant, the last term on the right-hand side vanishes, leaving $\Delta G = \Delta H - T\Delta S$. This tells us that the gain in *useful* work

from an isothermal system must be *less* than the gain in energy or enthalpy (ΔH). The difference is measured by the product of gain in entropy (ΔS) and the temperature at which the reaction occurs. $T\Delta S$ is "isothermally unavailable energy." ΔG is also the *minimum* work required to take a system from one equilibrium state to another.

Let's require that pV-work only be done in a reversible system. Then, because $\Delta U = T\Delta S - p\Delta V$, $H = U + pV$, and $\Delta H = \Delta U + p\Delta V + V\Delta p$, substitution into Eqn. (4.2) gives $\Delta G = [(T\Delta S - p\Delta V) + p\Delta V + V\Delta p] - T\Delta S - S\Delta T$, which simplifies to

$$\Delta G = V\Delta p - S\Delta T. \tag{4.3}$$

If we further require p and T to be constant, then $\Delta G = 0$. Just as with ΔS, we interpret this to mean that a reversible system has a maximal or minimal value of G when T and p are constant and the system is at equilibrium. In this case, and opposite to the entropy, the extremum is a minimum, just as with gravitational potential. In other words, the magnitude of ΔG measures the extent of displacement of the system from equilibrium, and $\Delta G = 0$ for a system at equilibrium.

Now we are positioned to see how the Gibbs free energy can be of great utility in predicting the direction of a process in a *closed* biochemical system. Have another look at Table 4.1, particularly the ΔG column. When ΔG is *positive* (if the energy change is *endergonic*), the process will *not* occur *spontaneously*. This is because the final state of the process has a higher free energy than the initial state, and this can be achieved only at the expense of the energy of the surroundings. When ΔG is *negative* for a process (if the energy change is *exergonic*), the reaction proceeds *spontaneously* in the direction of equilibrium, and when equilibrium is reached no further change will occur. *For any real process to occur spontaneously at constant temperature and pressure, the Gibbs free energy change must be negative*. Equation (4.1) shows that the lower the enthalpy (energy), the lower G, and the higher the entropy, the lower G. This tells us that spontaneity of a reaction is favored by a reduction of enthalpy (exothermic reactions) and by an increase of entropy (heat-releasing leading to increased disorder or increased energy that cannot be used to do work).

It must be emphasized that while the magnitude of ΔG tells us the size of the driving force in a spontaneous reaction, ΔG says nothing at all about the time required for the reaction to occur. Real physical, chemical, and biological processes occur at a finite rate, and all real chemical reactions are, to some extent, irreversible. Nevertheless, the basic principles of thermodynamics hold. Reversibility can be approached in the real world of biochemistry by having the process take place in a controlled manner, and that is what the biochemist often aims to do by studying reactions *in vitro*. The rate of a reaction will be developed in Chapter 8.

D. Phase transitions

We have already encountered phase transitions in two different contexts in this book: in likening the thermal denaturation of a protein molecule to the melting of an organic crystal, and in describing the physical properties of water. A *phase* is a system or part of a system that is homogeneous and has definite boundaries. A phase need not be a chemically pure substance. A phase transition is ordinarily caused by heat uptake or release, and when a phase change does occur it is at a definite temperature and involves a definite amount of heat. Phase changes are associated with a variety of fascinating and general aspects of biological thermodynamics. Let's take a few minutes to know more about them!

Phase transitions might look simple "from the outside;" in fact they are actually rather complex. No one really knows how to describe a phase change on the level of individual particles, though quantum mechanics must come into play at some level. Nevertheless, it is possible to give a description of a phase change on the macroscopic level in terms of classical thermodynamics. Some energetic quantities, for instance enthalpy and entropy, exhibit a *discontinuous* change at a phase boundary. What this means is that the enthalpy curve with temperature has the shape of a "step"; there is an abrupt change as the solid becomes a liquid or the liquid a solid, not a smooth change. Same for entropy. (See Fig. 4.5.) The amount of heat exchanged on melting or boiling is a heat capacity change, the latent heat of melting or vaporization, respectively. There is a relatively large change in heat capacity over an extremely small temperature range on a change of phase. Transitions of this type are "all-or-none" transitions; the material is completely in one phase on one side of the phase boundary and completely in another phase on the other side of the phase boundary. The two phases can coexist at the transition temperature. Such transitions are known as *first-order* phase transitions. These transitions resemble the "catastrophes" of an area of mathematics imaginatively called catastrophe theory. In many cases protein folding/unfolding closely resembles a first-order phase transition, for example, hen egg white lysozyme in Chapter 2. In more complicated situations, for example, gradual unfolding with heat increase, such a description is clearly inadequate (see Chapter 6).

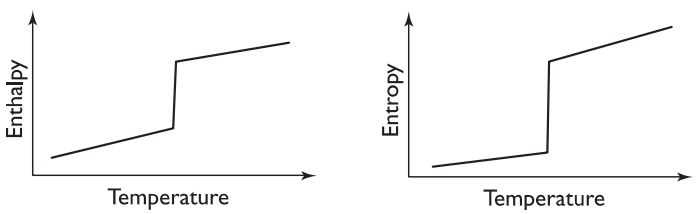


Fig. 4.5 First-order phase transition. The graphs show the behavior of the enthalpy and entropy functions versus temperature. Both of these thermodynamic quantities are *discontinuous* at the transition temperature, as is C_P.

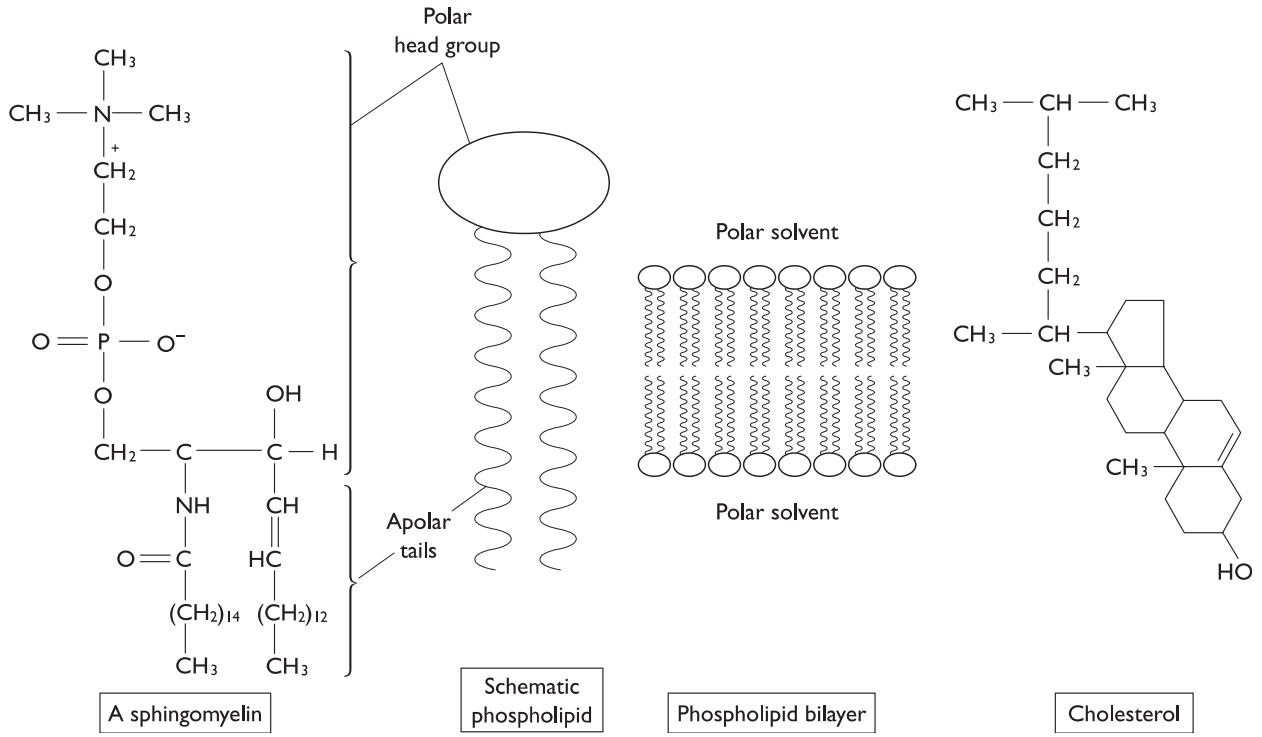


Fig. 4.6 Lipid bilayers. Lipids, or phospholipids, have two main regions, a polar "head" and aliphatic "tails." The head group is in contact with the solvent in a lipid bilayer, as shown. Sphingomyelins are the most common of the sphingolipids, which together form a major component of biological membranes. Sphingomyelins are just one type of phospholipid. The myelin sheath that surrounds and electrically insulates many nerve cell axons is rich in sphingomyelin. The most abundant steroid in animals is cholesterol. The metabolic precursor of steroid hormones, cholesterol is a major component of animal plasma membranes. In animals or plants, biological membranes are highly heterogeneous. They include not just several kinds of lipids and cholesterol, but membrane-spanning and membrane-associated proteins as well. At physiological temperatures, membranes are gel-like and allow lateral diffusion of their components.

The phase transition in water from the solid state to the liquid state is a first-order phase transition.

The liquid–solid phase boundary of water plays a key role in life on Earth – in more than one way. For instance, when the temperature drops, water begins to freeze on the surface of a pond, not at the bottom, and ice remains up top. Water on the surface loses its heat to the surroundings, the temperature of which can dip well below 0 °C. The density of solid water is lower than that of the liquid. Water is a peculiar substance! The physical properties of water also play a critical role in determining the level of the oceans and shaping the world's weather, determining what fraction of the oceans' water is liquid and how much is in the polar icecaps. In short, water determines the character of the biosphere.

Less well-known and obvious than the phase changes of water and proteins, perhaps, are those of lipids. These mainly water-insoluble molecules can undergo changes in state just as other compounds do. We say "mainly water-insoluble" because lipids are made of two parts, a small water-soluble "head" and a long water-insoluble "tail" (Fig. 4.6). In bilayers, which are surfaces of two layers of lipids in

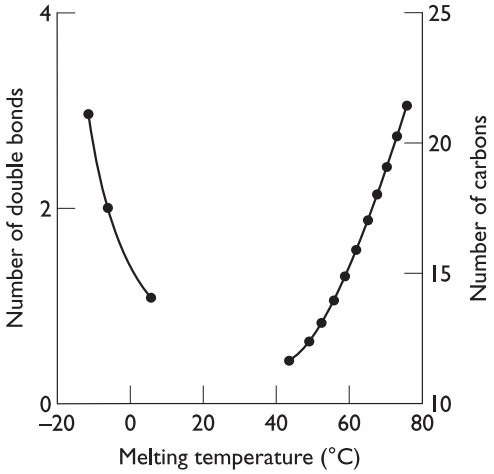


Fig. 4.7 Membrane melting temperature. In general, the melting temperature of a phospholipid bilayer decreases with increasing heterogeneity. An exception to the rule is cholesterol, which increases the melting temperature by increasing the rigidity of the bilayer. Increasing the number of double bonds in the aliphatic tail in lipids decreases the melting temperature by decreasing the ability of the molecules to pack against each other. A double bond introduces a kink in the tail. Increasing the length of the tail increases the melting temperature, because the aliphatic portions of lipids can interact favorably with each other by means of van der Waals forces. Redrawn from Fig. 7 of Bergethon (1998).

which the tails face each other, lipids can be in the liquid crystalline or gel state. This is an intermediate level of organization between the solid state, which is rigid, and the liquid state, which is fluid. The gel state is the one in which lipids are found in the membranes of the cells of living organisms. In pure lipid bilayers, there is a definite melting temperature between the solid state and the gel state, just as there is for water at the liquid–gas phase boundary and for a protein like lysozyme between its folded and unfolded states. The solid–gel transition of pure lipid bilayers is highly cooperative, also similar to the behavior of water and some proteins. Decreasing the purity of a bilayer, for instance by introducing a second type of lipid that is miscible with the first or by adding transmembrane proteins, the cooperativity of the solid–gel transition shows a corresponding decrease. Biological membranes in the living organism are highly heterogeneous: they are made of lipids in the fluid–lamellar phase, proteins, and some carbohydrate. Such membranes therefore do not exhibit very cooperative transitions, and melting occurs over a range of temperatures, usually 10–40 °C.

Other physical properties of lipids influence bilayer fluidity. One is length of the non-polar hydrocarbon tails (Fig. 4.7). The longer the chain, the higher the transition temperature. This is because the hydrophobic stuff of one lipid can interact fairly strongly with the hydrophobic stuff of another lipid. Another is the degree of saturation of the carbon–carbon bonds in the tails. Unsaturated bonds (double bonds) introduce kinks into the chain, making it more difficult for lipids in a bilayer to form an orderly array. Variation in the number and location of double bonds ensures that biological membranes do not become rigid. A contributor to membrane rigidity is lipid-soluble cholesterol, which decreases membrane fluidity by disrupting orderly interactions between fatty acid tails. Cholesterol itself is a rigid molecule (Fig. 4.6). The melting temperature also depends on solute and counter ion concentration! The physical properties of lipids, which are evidently complex, are of

vital important to the cell. Biological membranes permit membrane-bound proteins some degree of lateral movement, enable the cell to change shape and migrate, and make tissue formed by millions of cells relatively soft to the touch. Would a baby so enjoy nourishing itself at its mother's breast if her nipples were no more flexible than fingernails?

E. Chemical potential

Life as we know it could not exist without water. Small wonder that Thales of Miletus (fl. *c.* 500 BCE) considered water the ultimate substance, or *Urstoff*, the stuff of which all things are made! All physiological biochemical reactions take place in a largely aqueous environment, from enzyme catalysis to the folding of proteins to the binding of proteins to DNA to the assembly of large macromolecular complexes. It is probable that life on Earth began in a primordial sea of salt water (if it did not arrive from space). We had therefore better devote some energy to learning the basics of the thermodynamics of solutions. The path we shall take towards greater understanding will involve scaling a few jagged mathematical formulas, but the view we shall have at the end of the journey will make the effort worthwhile.

If you have doubted the usefulness of mathematics in previous study of biochemistry, it might help to bear in mind that mathematics is to biochemistry as a protocol is to producing a biochemical result. Mathematics is handmaiden, not master. Nevertheless, protocols themselves can be extremely useful, particularly when they are rather general in scope. One of the most highly cited scientific journal articles of all time is one of the most boring biochemistry articles ever published. This is because it has to do with a technique for separating proteins on the basis of size and says nothing specific about biology. But the paper in question is very important because the method it outlines (polyacrylamide gel electrophoresis) can be used in a very broad range of situations. In this sense, the protocol is even more important than any particular result it might be used to produce. In the same way, getting to know something of the mathematical background to a formula can be worth the time and effort, because the sort of thinking involved is in many cases of general utility.

Being quantitative about free energy changes is a matter of both being careful in making measurements and being clear about the conventions one has adopted. Teaching the first of these is beyond the scope of this book! To do the latter, we need to return to a topic introduced in Chapter 2: the standard state. Previously, we *defined* the standard enthalpy change, ΔH°, as the change in enthalpy for a process in which the initial and final states of one mole of a pure substance are at 298 K and 1 atm. Now we wish to define the *standard free energy change*, ΔG°. Here, the superscript indicates *unit activity* at standard temperature (298.15 K; three significant digits are usually

enough) and pressure (1 atm). The *activity* of a substance, a concept introduced by the American Gilbert Newton Lewis (1875–1946), is its concentration after correcting for non-ideal behavior, its effective concentration, its tendency to function as a reactant in a given chemical environment. There are many sources of non-ideality, an important one being the ability of a substance to interact with itself.

Ideal behavior of solute A is approached only in the limit of infinite dilution. That is, as $[A] \to 0$, $\gamma_A \to 1$. In the simplest case, the activity of substance A, a_A, is defined as

$$a_A = \gamma_A [A], \tag{4.4}$$

where γ_A is the *activity coefficient* of A on the *molarity* scale. When a different concentration scale is used, say the molality scale, a different activity coefficient is needed. The concept of activity is basically the same in both cases. According to Eqn. (4.4), $0 < a_A < [A]$ because $0 < \gamma_A < 1$. Activity is a *dimensionless* quantity; the units of the *molar* activity coefficient are $1\ mol^{-1}$.

Defining ΔG° at unit activity, while conceptually simple, is problematic for the biochemist. This is because free energy change depend on the concentrations of reactants and products, and the products and reactants are practically never maintained at molar concentrations throughout a reaction! Moreover, most reactions of interest do not occur at standard temperature. Furthermore, biochemistry presents many cases where the solvent itself is part of a reaction of interest. We need a way to take all these considerations into account when discussing free energy change.

The relationship between the concentration of a substance A and its free energy is *defined* as

$$\mu_A - \mu_A^\circ = RT \ln a_A, \tag{4.5}$$

where μ_A is the *partial molar free energy*, or *chemical potential*, of A, and μ_A° is the standard state chemical potential of A. The partial molar free energy of A is, in essence, just $\Delta G_A / \Delta n_A$, or how the free energy of A changes when the number of molecules of A in the system changes by one (Fig. 4.8). The chemical potential of A is a function of its chemical potential in the standard state and its concentration. Equation (4.5) could include a volume term and an electrical term (there are numerous other kinds of work, see Chapter 2), but let's assume for the moment that the system does not expand against a constant pressure and that no charged particles are moving in an electric field. It is appropriate to call μ the chemical potential because at constant T and p, G is a function of chemical composition alone.

Equation (4.5) tells us that when $a_A = 1$, $\mu_A - \mu_A^\circ = 0$. That is $\mu_A - \mu_A^\circ$ measures the chemical potential of A relative to the standard state conditions; the activity of a substance is 1 in the standard state. The chemical potential also depends on temperature as shown, and the gas constant puts things on a per-mole basis. Equation (4.5) also says that the chemical potential of a solvent

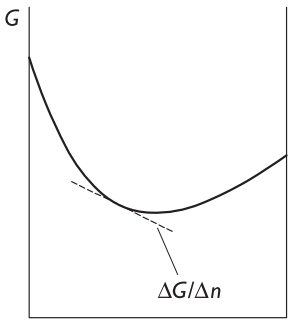


Fig. 4.8 Thermodynamic potential and solute concentration. The Gibbs free energy of a solute varies with concentration. The chemical potential measures the rate of change of G with n, or the slope of the curve at a given value of n ($\Delta G / \Delta n$). Note that G can decrease or increase on increases in concentration.

decreases as solute is added. The activity of a substance is always highest when that substance is pure. The reduction in chemical potential on mixing occurs even if the solution is ideal (enthalpy of mixing of zero),[2] as in the case where the solute is "inert" and does not interact with the solvent at all. This tells us that the decrease in chemical potential on mixing is fundamentally an entropic effect, even if for any real solvent and solute the change in chemical potential will contain an enthalpic component stemming from interparticle interactions.

Before getting too far down this road, let's look at an example that is aimed at helping to clarify the difference between ΔG° and ΔG (and thus $\Delta\mu^\circ$ and $\Delta\mu$). Suppose you're doing a study of the binding of a peptide hormone to a receptor situated in the plasma membrane of a cell. The cheapest way to obtain large quantities of pure hormone might be to synthesize it chemically and purify it using a type of liquid chromatography – quite possibly no problem at all if the peptide does not have to be glycosylated. Suppose that the sequence of the peptide you're studying is X–X–X–X–X–X–X–X–Ala–Gly, and because solid-phase chemical synthesis of peptides is done from the C-terminus to the N-terminus (the reverse of how it happens during translation of mRNA on a ribosome), the first coupling is of Ala to Gly attached to a resin. You'd like to know something about the energetics of formation of the peptide bond between these two amino acids. But when two free amino acids join to form a peptide bond, a water molecule is produced; the reaction is a type of dehydration synthesis. When the reaction occurs *in aqueous solution* under *standard state conditions* – that is, all products and reactants are at a concentration of 1 M *except water* (we'll say why later on) – the free energy change, ΔG°, is 4130 cal mol^{-1} – far from equilibrium! The driving force for change is in the direction of the reactants. When Ala and Gly are at 0.1 M, and Ala–Gly is at 12.5 μM, the reactants and products are no longer in their standard states, and the free energy difference is not ΔG° but ΔG. On doing an experiment to determine the energetics of peptide bond formation, one finds that the reaction is at equilibrium and that no change in the concentrations of reactants takes place unless the system is perturbed. In other words, $\Delta G = 0$. If the concentrations are again changed, so that Ala and Gly are at 1 M, as in the standard state reaction, but Ala–Gly is at 0.1 mM, you will find that the reaction proceeds in the direction of the products and $\Delta G = -1350$ cal mol^{-1}. ΔG on its own measures how far away a reaction is from equilibrium, and $\Delta G - \Delta G^\circ$ measures

[2] The given definition of an ideal solution is less general than it could be. We could say instead that an ideal solution is one for which the enthalpy of mixing is zero *at all temperatures*, implying that the solute always interacts with the solvent as it interacts with itself. An even more technical definition is that an ideal solution is a homogeneous mixture of substances whose physical properties are linearly proportional to the properties of the pure components, which holds in fact for many dilute solutions. This is known as Raoult's Law, after François-Marie Raoult (1830–1901), a French chemist.

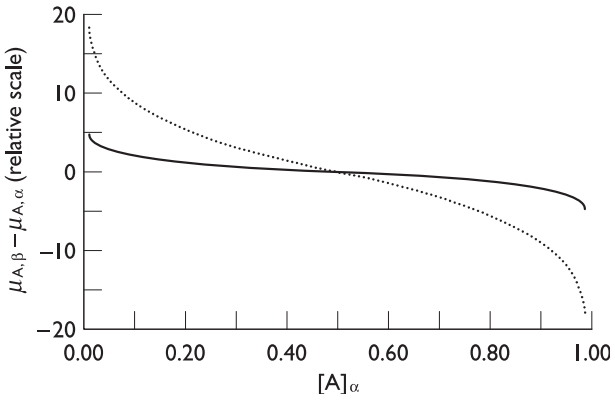


Fig. 4.9 Chemical potential difference as a function of concentration. The figure illustrates the behavior of the left-hand side of Eqn. (4.7) when the total concentration of A is fixed; $[A]_\alpha + [A]_\beta =$ constant. The dashed line represents a change in temperature.

how much the conditions differ from the standard state conditions. Yippee!

Let's have the discussion become a little more difficult. Suppose we have a two-component solution made of solvent and solute. The solute is some general substance A that is soluble in the solvent. The solvent could be water and the solute a metabolite. Assuming ideal behavior, $a_A = [A]$ and Eqn. (4.5) becomes

$$\mu_A - \mu_A^\circ = RT \ln[A]. \tag{4.6}$$

Let's now construct a *notional* partition between two regions of the solution and require that the system *not* be at equilibrium. Our mental boundary could be a liquid–gas interface or a membrane permeable to A. Substance A in any case can move across the boundary and back again by *random* motion. Calling the two regions of the solution α and β, the difference in chemical potential between them is

$$\Delta\mu_A = \mu_{A,\beta} - \mu_{A,\alpha} = RT \ln([A]_\beta / [A]_\alpha). \tag{4.7}$$

The standard state terms have vanished because the standard state free energy of A is the same in both regions. Intuitively, we would expect $\Delta\mu_A < 0$ when $[A]_\alpha > [A]_\beta$. When $[A]_\alpha > [A]_\beta$, the argument of the logarithm $([A]_\beta/[A]_\alpha)$ is less than one, and because $\ln x < 0$ for $x < 1$, our expectation is met (Fig. 4.9). When $\Delta\mu_A$ is negative, the solute particle will move spontaneously down its concentration gradient from α to β. If the concentration is greater in region α than region β, $\Delta\mu_A > 0$, and A will move spontaneously from β to α.

So far, so good. Backing up a couple of steps and combining Eqns. (4.4) and (4.5), we have

$$\mu_1 = \mu_1^\circ + RT \ln \gamma_1[1] \tag{4.8}$$

for component 1, the solvent. We can rewrite Eqn. (4.8) as

$$\mu_1 = \mu_1^\circ + RT \ln f_1 X_1, \tag{4.9}$$

where f_1 is the activity coefficient of component 1 on the *mole fraction*

scale. The mole fraction X_i is the *number of molecules* of i (i.e. n_i) expressed as a fraction of the total *number of molecules* in the system, n. In other words, $X_i = n_i/n$. We are still dealing with a two component system, so $X_1 = n_1/(n_1 + n_2) = 1 - X_2$, where X_2 is the mole fraction of solute A; the mole fractions of the individual components of solution must sum to 1. Writing $RT \ln f_1 X_1$ as $RT \ln f_1 + RT \ln X_1$ (which we can do because $\ln ab = \ln a + \ln b$), using $\ln(1+x) = x - x^2/2 + x^3/3 \ldots$ (a relationship from mathematics which is valid for $-1 < x < 1$), and rearranging terms, Eqn. (4.7) becomes

$$\mu_1 - \mu_1^\circ = RT(-X_2 + \ldots) + RT \ln f_1. \tag{4.10}$$

We can simplify this beastly equation in two ways if the solution is dilute ($n_2 << n_1$). One is that the "higher order" terms in X_2 (namely, the square of X_2, the cube of X_2, etc.) are small because n_2 and therefore X_2 is small. So we consciously neglect them! The other assumption is that

$$X_2 \approx C_2 V_1^\circ / M_2, \tag{4.11}$$

where C_2 is the concentration of the solute in *molal* units (g l^{-1}), M_2 is the molecular weight of the solute, and V_1° is the molar volume of pure solvent. Equation (4.11) comes from $X_2 = n_2/(n_1 + n_2) \approx n_2/n_1$ when $n_2 << n_1$. Because $n_1 V_1^\circ = V$, the total volume of *solvent*, $X_2 \approx n_2 V_1^\circ / V = M_2 n_2 V_1^\circ / M_2 V = (M_2 n_2/V)(V_1^\circ/M_2) = C_2 V_1^\circ / M_2$. The dimensions of Eqn. (4.11) are (g l^{-1}) × (l mol^{-1}) / (g mol^{-1}) = 1 (i.e., no units). That is, X_2 is dimensionless, as it must be! Substitution of Eqn. (4.11) into Eqn. (4.10) gives

$$\Delta\mu_1 = \mu_1 - \mu_1^\circ \approx -RTC_2 V_1^\circ / M_2 + RT \ln f_1 \tag{4.12}$$

for a dilute solution of component 2, the solute. Assuming that the second term on the right-hand side is small enough to be neglected (because $f_1 \approx 1$ and $\ln 1 \approx 0$), as is often approximately the case in biochemical experiments, we have a relatively simple and mathematical expression and can see that the chemical potential of a single-component solution relative to the pure solvent ($\Delta\mu_1$) is (approximately) directly proportional to the concentration of the solute and inversely proportional to solute mass. The higher the concentration of a solute, the higher its chemical potential. *A substance always diffuses from a region of higher concentration to a region of lower concentration.* The greater the mass of a particle, the lower its chemical potential. If we start to think about real biological macromolecules like proteins or DNA, things become more complex, because such molecules are almost always charged. Let's wait until Chapter 5 to see how to take charge into account.

F. Effect of solutes on boiling points and freezing points

Our comments on phase transitions have suggested that we should be aware of them in biochemistry. And the previous section

showed that the chemical potential of a substance depends on its concentration. Now we want to combine these areas of knowledge to obtain an expression for the change in boiling or freezing point of a solvent which results from the addition of a solute. We know that salt is used to keep roads clear of ice in the winter, and that salt is added to water to cook pasta, but why either is done may have been a mystery to you before now. Let's slog through a few more formulas and then look at examples of how to apply the gained knowledge to biochemistry.

From Eqn. (4.9) we have

$$\mu_1 - \mu_1^\circ = RT \ln f_1 X_1. \tag{4.13}$$

This can also be expressed as

$$(\mu_1 - \mu_1^\circ)/RT = \Delta G_{1,\mathrm{m}}/RT = (\Delta H_{1,\mathrm{m}} - T\Delta S_{1,\mathrm{m}})/RT = \ln f_1 X_1, \tag{4.14}$$

where, for instance, $\Delta G_{1,\mathrm{m}} = \Delta G_1/\Delta n_1$, and ΔG_1 represents the difference in the Gibbs free energy of component 1 between the one phase and the other. Because ΔH_m *and* ΔS_m *are relatively insensitive to temperature over short temperature ranges*, $\Delta H_\mathrm{m} = \Delta H_\mathrm{tr}$ and $\Delta S_\mathrm{m} = \Delta H_\mathrm{tr}/T_\mathrm{tr}$, where 'tr' stands for 'transition.' Substitution of these relations into Eqn. (4.14) gives

$$\Delta H_\mathrm{tr}/R \times (1/T - 1/T_\mathrm{tr}) = \ln f_1 X_1 \tag{4.15}$$

The difference $T - T_\mathrm{tr} = \Delta T$ is very small, so $1/T - 1/T_\mathrm{tr} \approx -\Delta T/T_\mathrm{tr}^2$. When this and Eqn. (4.10) are substituted in, we obtain

$$-\Delta H_\mathrm{tr}/R \times \Delta T/T_\mathrm{tr}^2 = -(X_2 + \cdots) + \ln f_1. \tag{4.16}$$

If the concentration of the solute, component 2, is small, then $\ln f_1 \approx 0$ and Eqn. (4.16) equation simplifies to

$$\Delta T \approx RX_2 T_\mathrm{tr}^2/\Delta H_\mathrm{tr}. \tag{4.17}$$

That's it! It is easy to see that ΔT varies proportionately with X_2. The greater the mole fraction of solute, the more the temperature of the phase transition will differ from that of the pure solvent. Because the effect depends on the mole fraction of the solute but not the solute's identity, it is called a colligative property, one that depends "on the collection" of solute molecules.

Are there practical applications in biological science of this hard-won knowledge of physical chemistry? Yes! In both cases, the addition of the "solute" changes the temperature of the phase change, in the one case lowering the freezing point and in the other raising the boiling point. There are also applications in the biochemistry lab. For instance, glycerol is often used to reduce the freezing point of aqueous protein solutions. The restriction enzymes so important for molecular biology and biotechnology are stored in *c.* 50% glycerol, which lowers the freezing point of water to below $-20\,^\circ$C! The enzymes are more thermostable at this temperature

than at 4 °C, and this preserves them for longer than if stored in the fridge by reducing the rate of spontaneous inactivation. Maintaining enzymes in the freezer comes with a cost – that of keeping the icebox running – so this is not a case of getting something for nothing! Then there are proteins that either raise or lower the freezing point of water with biological effect. The ones that lower it are known as anti-freeze proteins. Such proteins bind to and arrest the growth of ice crystals in the fish and thereby prevent it from freezing. By contrast, some bacterial proteins are known to *increase* the probability that supercooled water[3] will freeze.

Box 4.1 Puzzling proteins keep flounders from freezing

Pleuronectes americanus, the winter flounder, flourishes in the freezing waters of polar and subpolar oceans. Plasma proteins in the blood of this fish bind to ice crystals and stop them from growing. Type I anti-freeze protein, discovered in winter flounder 30 years ago, is a 3.3 kDa alanine-rich amphipathic molecule that forms a single α helix and binds to a pyramidal plane of ice. The concentration of this protein in fish blood plasma in winter is about 10–15 mg mℓ^{-1}, giving a non-colligative freezing-point depression, or thermal hysteresis, of 0.7 °C; far short of the −1.9 °C freezing point of sea water. The thermal hysteresis attributable to the colligative effects of other blood solutes adds another 0.8 °C to the total. But the 1.5 °C sum is not quite enough to keep the fish thawed. A piece of the jigsaw was missing. Recently, researchers in Ontario and Newfoundland, Canada, identified a previously unknown 16.7 kDa anti-freeze protein in the fish. This protein provides a thermal hysteresis of 1.1 °C at a concentration of just 0.1 mg mℓ^{-1}, about half the circulating concentration. It would appear that this new protein resolves the question of how the fish survives icy polar waters.

G. Ionic solutions

Ions in solution are called electrolytes. Charged particles get this name from their ability to conduct an electric current. Discussion about them takes up space in the present tome because water is found everywhere in the biological world, life as we know it could not exist without water, and most of the time plenty of ions are present in water. (An expensive water purification system is required to do work on ordinary water and separate ions from the solvent!) Moreover, many representatives of the three major classes of biological *macromolecule* – namely, proteins, nucleic acids, and polysaccharides – are charged at neutral pH, even if their net charge might be zero (though it usually isn't). And many species of lipid are charged. Charge properties help to give biomacromolecules important physical properties that are closely related to their physiological function.

[3] Liquid water cooled below its normal freezing point.

Box 4.2. Surviving a dip in the ice-cold ocean in a Speedo

In December 2005 British swimmer Mr. Lewis Pugh broke two records previously held by US swimmer Ms. Lynne Cox: the most southerly swim in the ocean, and the longest-duration swim near a Pole. For his first feat, Mr. Pugh swam 1 km under conditions that would kill a typical person within minutes. He spent 18 min 10 s in salt water just below 0 °C in the seas off the Antarctic Peninsula (latitude, 65° S). His second feat was a 1 mile swim for 30 min 30 s in mere 2–3 °C water. How was Mr. Pugh able to maintain body temperature and stave off the normally debilitating effects of the body's reaction to cold? In part by habituation, through increasingly frequent exposure to cold. The body can be trained not to shiver, allowing muscles to work more effectively in frigid surroundings. But when the surroundings are so cold, how is it possible to keep the temperature of the body above 35 °C, the cut-off point for hypothermia? Mental imagery plays a key role, stimulating the production of large amounts of heat by "anticipatory thermogenesis." Just as in sprinting and other sports, the release of stress hormones by the brain increases the metabolic rate of the body. Finally, there are the fitness and fatness factors. The latter is related to why women tend to be better than men at outdoor swimming: females tend to have more fat over their leg and arm muscles, which keeps them better insulated. Fitness favors the frantic movement of appendages, muscle activity, generation of heat, endurance, and speed. And men are generally ahead of women in that category. Mind over matter?

Ionic solutions tend to be dominated by electrical forces, which can, be very strong. Referring back to Table 2.2, you will see that the electrostatic energy of two electronic charges can be as great as 14 kcal mol^{-1}, an energy as large or larger than the free energy difference between the folded and unfolded states of a protein at room temperature. Often, though, electrostatic interactions are substantially weaker, as they depend not only on the distance between the interacting charges but also on the dielectric constant of the medium, D. In the middle of a protein, which is something like an oily solid, $D \approx 4$, and the electrostatic interaction is reduced to just 75 % from its vacuum value. In bulk aqueous solution, by contrast, $D \approx 80$! The polarity of water greatly reduces the distance over which the strength of the electric field created by a charge is significant. The strength of charge–charge interactions in water is often reduced even further by the orderly arrangement of a few water molecules in a "solvation shell" around an ion.

EDTA is a cation[4] chelator. It is a useful tool in the biochemist's kit because it can be used in a variety of practical situations. For instance, when preparing dialysis tubing, EDTA is used to "trap" divalent metal ions, "removing" them from solution, limiting their ability to bind to biomacromolecules one might want to prepare by

[4] Positive ions, e.g. Na^+, are cations. Cl^- and other negative ions are anions.

dialysis and proteases that might require them for activity. The activity of such proteases is greatly reduced in the presence of EDTA. ΔH for the binding of Mg^{2+} to EDTA is positive, but because ΔS is very negative, $\Delta G < 0$ and chelation occurs. Anyone who has ever tried to make an aqueous EDTA solution will know very well that the sodium salt does not dissolve very quickly at room temperature. If enough EDTA is present, when the solution comes to equilibrium only some of the salt will be dissolved. It might be possible to get more EDTA into solution by heating, which changes its solubility and the rate of dissolution.

What we want to do now is look at some general ways of thinking about thermodynamic properties of electrolytes. We'll do this by way of the example of EDTA. But let's bear in mind that the example will show how the approach is really much broader in scope. The solubility equilibrium of EDTA can be written as

$$Na_4EDTA(s) \Leftrightarrow 4Na^+ + EDTA^{4-}. \tag{4.18}$$

Based on this equation, at equilibrium,

$$\mu_{NaEDTA} = 4\mu_{Na^+} + \mu_{EDTA^-}. \tag{4.19}$$

The positive and negative ions appear as a pair because it is not possible to make separate measurements of the chemical potentials on the right hand side. To take this doubling effect into account, we define the *mean chemical potential*, $\mu_\pm$, which in this case is

$$\mu_\pm = \frac{4}{5}\mu_{Na^+} + \frac{1}{5}\mu_{EDTA^-}. \tag{4.20}$$

The coefficients account for the stoichiometry of dissociation of EDTA. Equation (4.19) can now be rewritten as

$$\mu_{NaEDTA} = 5\mu_\pm. \tag{4.21}$$

More generally,

$$W \Leftrightarrow \nu_+ A^{z+} + \nu_- B^{z-}. \tag{4.22}$$

In this equation W is a neutral compound, A and B are positive and negative ions with ion numbers z^+ and z^-, and ν_+ and ν_- are stoichiometric coefficients. The mean (not meaner!) chemical potential is

$$\mu_\pm = \frac{(\nu_+ A^{z+} + \nu_- B^{z-})}{V_+ + V_-} = \frac{\mu_{salt}}{\nu_+ + \nu_-}. \tag{4.23}$$

This is just W divided by the sum of the stoichiometric coefficients. Try working through the equations using EDTA as the example.

Substituting Eqn. (4.9) into Eqn. (4.20), we have

$$\mu_\pm = \frac{4}{5}(\mu^\circ_{Na^+} + RT\ln(f_{Na^+}X_{Na^+})) + \frac{1}{5}(\mu^\circ_{EDTA^-} + RT\ln(f_{EDTA^-}X_{EDTA^-})). \tag{4.24}$$

Making use of $x\ln a = \ln a^x$, a handy formula from mathematics, gives

$$= \mu^\circ_\pm + RT\ln\sqrt[5]{(f_{Na^+}X_{Na^+})^4 f_{EDTA^-}X_{EDTA^-}}, \tag{4.25}$$

where the standard state chemical potentials of Na^+ and pure $EDTA^-$ have been combined in the first term on the right-hand side of Eqn. (4.25).

Just as the chemical potentials of the ions cannot be measured separately, neither can one measure the activity coefficients separately. We therefore define a *mean ionic activity coefficient*, which for the present example is

$$f_\pm = {f_{Na^+}}^{4/5}{f_{EDTA^-}}^{1/5}. \tag{4.26}$$

This comes from $(ab)^{1/2}$, the *geometric mean* of a and b. In the more general case, Eqn. (4.26) looks like

$$f_\pm = (f_+^{\nu_+}f_-^{\nu_-})^{1/(\nu_+ + \nu_-)}, \tag{4.27}$$

where f_+ and f_- are the activity coefficients of the positive and negative ions on the mole fraction scale.

Knowing the mean activity coefficient of a salt can be important for interpreting the results of a biochemistry experiment. This will be especially true when the solution conditions are far from ideal, and particularly when the salt concentration is high. The bacteria that live in the Dead Sea, known as halophiles, thrive in a high salt environment. Somehow or other the molecular machinery of these bugs can cope with the high-salt surroundings. A high-salt solution is far from ideal, so the activity coefficients of ions in the surrounding environment of halobacteria deviate substantially from unity. In a similar example, the salt guanidinium chloride (GuHCl) is a strong protein denaturant. Most proteins are unfolded at a concentration of about 6 M GuHCl (a large concentration indeed, but one still many times smaller than the concentration of pure water, which is about ten-fold greater). Like HCl, GuHCl dissociates completely in aqueous solution to guanidinium ion and chloride ion (the solubility limit of GuHCl is well above 5 M at 25 °C). To explain in molecular terms what GuHCl does to protein structure, one needs to know its activity coefficient. We'll learn more about guanidinium chloride-induced unfolding of proteins in Chapter 5.

Now we want to think about electrolytes in a slightly different way. What follows is a simplified version of the theory of strong electrolytes developed by the Netherlander Petrus Josephus Wilhelmus Debye[5] (1884–1966) and the German Erich Hückel (1896–1980), published in 1923. The activity of an ion depends on a quantity known as the ionic strength, I, which is defined as

$$I = \frac{1}{2}\sum_i z_i^2 m_i. \tag{4.28}$$

[5] Debye was awarded the Nobel Prize in Chemistry in 1936.

Here m_i, the *molality*, is defined as X_i/M_s, the ratio of the mole fraction of solute i to the molecular mass of the solvent in *kilograms*, M_s. Note that if the salt you're working with is relatively simple, like NaCl, there is no problem in computing I: NaCl dissociates below its solubility limit, each particle carries just one charge, so a one molal solution of this salt has an ionic strength of $[(1^2\times 1)+(1^2\times 1)]/2=1$. $CaCl_2$ is somewhat more complicated, because the ions involved no longer carry the same charge.

Finally, it can be shown[6] that the activity coefficient of ion i on the molality scale is

$$\log \gamma_i = -Hz_i^2\sqrt{I}, \tag{4.29}$$

where γ_i is the activity coefficient on the *molality* scale. H is a complicated expression that depends the density of the solvent, the absolute temperature, the charge on an electron, the dielectric constant of the solvent ... For ions in water at 25 °C, $H\approx 0.5$. And so there is a way to calculate γ_i. Things are much more complicated when dealing with polyvalent ions like proteins, because the degree of ionization is sensitive to pH. Polyvalent ions and ionization will be treated in greater depth below. Enough about ions for now!

H. Equilibrium constant

We have looked at the concept of equilibrium from various directions throughout the present chapter. But in fact, we have only scratched the surface of what could be known. Many things in life work like that. Here, we approach equilibrium in yet another way, one that is very useful to biological scientists, particularly biochemists.

Given a general reaction

$$a\mathrm{A} + b\mathrm{B} \Leftrightarrow c\mathrm{C} + d\mathrm{D} \tag{4.30}$$

the overall free energy change is

$$\Delta G = c\mu_\mathrm{C} + d\mu_\mathrm{D} - a\mu_\mathrm{A} - b\mu_\mathrm{B}. \tag{4.31}$$

Substituting Eqn. (4.31) into Eqn. (4.5), we have

$$\Delta G = \Delta G^\circ + RT\ln\left(\frac{[\mathrm{C}]^c\,[\mathrm{D}]^d}{[\mathrm{A}]^a\,[\mathrm{B}]^b}\right), \tag{4.32}$$

where $\Delta^\circ_G = c\mu^\circ_C + d\mu^\circ_D - a\mu^\circ_A - b\mu^\circ_B$ and we have not distinguished between activity and concentration, in case the reaction is not carried out at infinite dilution. The quantity $[\mathrm{C}]^c[\mathrm{D}]^d/[\mathrm{A}]^a[\mathrm{B}]^b$, called the *mass action ratio*, is based on Eqn. (4.30). Equation (4.32) indicates that the free energy change of a reaction has two parts: a constant term that depends only on the particular reaction taking place, and a

[6] See, for example, pp. 250–2 of Atkins (1998).

Table 4.3. *Relationship between ΔG° and K_{eq}*

Free energy change	Equilibrium constant
$\Delta G^\circ < 0$	$K_{eq} > 1$
$\Delta G^\circ = 0$	$K_{eq} = 1$
$\Delta G^\circ > 0$	$K_{eq} < 1$

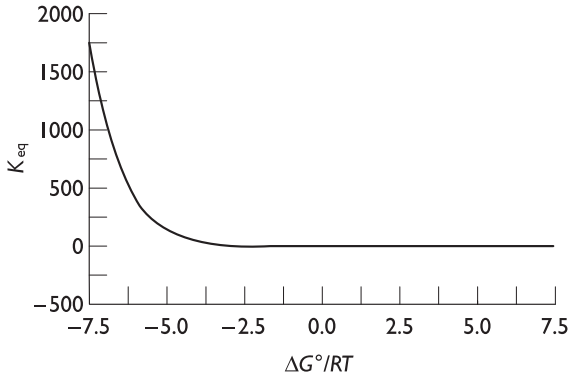


Fig. 4.10 Variation of K_{eq} with ΔG°. Note that K_{eq} is a function of the negative logarithm of ΔG°. When ΔG° is large and negative, K_{eq} is very large. When ΔG° is large and positive, K_{eq} is very small. In most biochemical reactions ΔG° will fall within the range shown here.

variable term that depends on temperature, concentrations of reactants and products, and stoichiometric relationships.

At equilibrium, the forward reaction balances the reverse reaction, and $\Delta G = 0$. So

$$\Delta G^\circ = -RT \ln K_{eq} = -RT \ln \left(\frac{[C]_{eq}^c \, [D]_{eq}^d}{[A]_{eq}^a \, [B]_{eq}^b} \right), \tag{4.33}$$

where the subscript "eq" signifies "equilibrium." The concentrations of reactants and products are the concentrations at equilibrium. The equilibrium constant of the reaction, K_{eq}, is *defined* as

$$K_{eq} = \left(\frac{[C]_{eq}^c \, [D]_{eq}^d}{[A]_{eq}^a \, [B]_{eq}^b} \right). \tag{4.34}$$

Note that K_{eq} will be unitless if $a + b = c + d$. Equation (4.33) indicates that K_{eq} can be calculated from standard state free energies. The form of relationship is illustrated in Fig. 4.10. You can get a feel for magnitudes by substituting in values (see Table 4.3). For instance, when $\Delta G^\circ = 0$, $K_{eq} = 1$. A 10-fold change in K_{eq} at 25 °C corresponds to $\Delta G^\circ = 5.7$ kJ mol^{-1}, an energy difference two to three times greater than thermal energy at the same temperature. Deviations from equilibrium will stimulate a change in the system towards the equilibrium concentrations of reactants and products. This is known as *Le Châtelier's principle*.[7] A reaction will go to completion for $K_{eq} \geq 10^4$.

[7] Named after the French chemist Henri Louis Le Châtelier (1850–1936). The principle was first enunciated in 1884, and it applies equally as well to reversible chemical

Comparison of Eqn. (4.32) with Eqn. (4.34) tells us that it will be hard to change the direction of a reaction with a very large or very small value of K_{eq} by changing the mass action ratio. In Chapter 5 we shall see how this difficulty is relevant in glycolysis, the process of glucose metabolism. Finally, Eqn. (4.34) also has a relationship to reaction rates, which was put forward by van't Hoff. We mention this here as a foretaste of the development in Chapter 8.

Here's a worked example of the relationship between Gibbs free energy change and equilibrium constant. Consider the glycolytic reaction in which a phosphoryl group is moved from carbon 1 of glucose to carbon 6:

$$\text{glucose-1-phosphate} \Leftrightarrow \text{glucose-6-phosphate}. \tag{4.35}$$

The equilibrium constant for this reaction, which is catalyzed by the enzyme phosphoglucomutase, is 19. How big is ΔG° at room temperature? By Eqn. (4.31), $\Delta G^\circ = -(8.314\ \text{J mol}^{-1}\ \text{K}^{-1}) \times 298\ \text{K} \times \ln(19) = -7.3\ \text{kJ mol}^{-1}$. The negative sign indicates that reaction proceeds to the right spontaneously *under standard conditions*. We need not have evaluated the magnitude of ΔG° to know this (Table 4.1), because when $K_{eq} > 1$, $\Delta G^\circ < 0$ and the forward reaction will be spontaneous. If $K_{eq} < 1$, the reaction will be energetically unfavorable. Let's see what happens when we alter the concentrations of reactants and products in Eqn. (4.35) and require the concentrations of reactants and products to be held at 10 mM and 1 mM, respectively. By Eqn. (4.32), $\Delta G = -7.3\ \text{kJ mol}^{-1} + (8.314\ \text{J mol}^{-1}\ \text{K}^{-1}) \times 298\ \text{K} \times \ln(1/10) = -13\ \text{kJ mol}^{-1}$. The reaction is considerably more exergonic than under standard state conditions. This tells us both that we are far from equilibrium, where $\Delta G = 0$, and the reaction will proceed to the right to reach equilibrium. Clearly, the magnitude of ΔG can depend significantly on the concentrations of reactants and products. In fact, the concentrations can be different enough in some cases (when $\Delta G \approx 0$) to reverse the direction of the reaction, a situation that is relatively common in metabolism and occurs in glycolysis. What is the position of the enzyme in this picture? Will it influence the rate of reaction? Yes! Will it change the free energy difference between products and reactants? No! Why not? The Gibbs free energy is a thermodynamic state function. As such, what it measures depends only on the state of the system and not on how the system was prepared.

I. Standard state in biochemistry

Most *in vitro* biochemical experiments are carried out at constant temperature and pressure in dilute *aqueous solution* near *neutral* pH. To be maximally useful, the *biochemist's definition of standard state*

reactions and reversible physical processes. The conclusions reached by Le Châtelier in his 1884 work had been anticipated in part by the American J. Willard Gibbs.

should take all these conditions into account. Doing this will seem to make things more complicated at first, but not ultimately.

We *define* the standard state of water to be that of the pure liquid. This means that the activity of water is *set* to 1, even though its concentration is 55.5 M. One justification for this is that if a reaction is carried out in dilute aqueous solution, the percentage change in concentration of water will ordinarily be negligible. Another is that it is pretty easy to multiply and divide by 1! You might agree that many *in vitro* experiments are carried out under relatively close-to-ideal conditions, but object that such conditions will often differ greatly from those in a cell. Which cell? Which organism? Where? We take 25 °C as the standard temperature because it is convenient for bench-top experiments. It is also close to the temperature of many organisms. The hydrogen ion activity is *defined* as unity at pH 7 (neutral solution), not pH 0 (highly acidic solution where the activity of H_3O^+ is 1).[8] Finally, no account is taken of the various ionization states of molecules that might be present at pH 7. This is particularly relevant to biological macromolecules, especially proteins, which can be ionized in a multitude of different ways. When all the stated conditions are accounted for, the standard state free energy change is symbolized as $\Delta G^{\circ\prime}$. The prime indicates pH 7. See Table 4.4.

To put the foregoing discussion in practice, consider the following chemical reaction, a sort of variation on Eqn. (4.30):

$$\mathrm{A} + \mathrm{B} \Leftrightarrow \mathrm{C} + \mathrm{D} + n\mathrm{H_2O}. \tag{4.36}$$

From Eqn. (4.34),

$$\begin{aligned} \Delta G^\circ &= RT \ln K_{\mathrm{eq}} \\ &= RT \ln([\mathrm{C}]_{\mathrm{eq}}[\mathrm{D}]_{\mathrm{eq}}[\mathrm{H_2O}]^n / [\mathrm{A}]_{\mathrm{eq}}[\mathrm{B}]_{\mathrm{eq}}) \\ &= RT \ln([\mathrm{C}]_{\mathrm{eq}}[\mathrm{D}]_{\mathrm{eq}} / [\mathrm{A}]_{\mathrm{eq}}[\mathrm{B}]_{\mathrm{eq}}) - nRT \ln[\mathrm{H_2O}], \end{aligned} \tag{4.37}$$

which, on invoking the biochemist's conventions, is

$$\Delta G^{\circ\prime} = RT \ln K_{\mathrm{eq}}{}' = RT \ln([\mathrm{C}]_{\mathrm{eq}}[\mathrm{D}]_{\mathrm{eq}} / [\mathrm{A}]_{\mathrm{eq}}[\mathrm{B}]_{\mathrm{eq}}). \tag{4.38}$$

Thus, the relationship between ΔG° and $\Delta G^{\circ\prime}$ is

$$\Delta G^{\circ\prime} = \Delta G^\circ + nRT \ln [\mathrm{H_2O}]. \tag{4.39}$$

What if protons are "consumed" in the reaction, taken up from the solvent by solute molecules? The situation can be modeled as

$$\mathrm{A} + \nu \mathrm{H^+(aq)} \rightarrow \mathrm{P}. \tag{4.40}$$

By Eqn. (4.31),

$$\Delta G = \mu_{\mathrm{p}} - \mu_{\mathrm{A}} - \nu \mu_{\mathrm{H^+}}. \tag{4.41}$$

If A and P are in their standard states, then

$$\Delta G = \mu^\circ_{\mathrm{P}} - \mu^\circ_{\mathrm{A}} - \nu \mu_{\mathrm{H^+}}. \tag{4.42}$$

[8] H_3O^+ symbolizes the hydronium ion, the proton in aqueous solution.

Table 4.4. | *Values of $\Delta G^{\circ\prime}$ for some important biochemical reactions*

Reaction	$\Delta G^{\circ\prime}$ (kcal mol^{-1})
HYDROLYSIS	
Acid anhydrides:	
Acetic anhydride + $H_2O \rightarrow$ 2 acetate	−21.8
$PP_i + H_2O \rightarrow 2P_i$	−8.0
$ATP + H_2O \rightarrow ADP + 2P_i$	−7.3
Esters:	
Ethylacetate + $H_2O \rightarrow$ ethanol + acetate	−4.7
Glucose-6-phosphate + $H_2O \rightarrow$ glucose + P_i	−3.3
Amides:	
Glutamine + $H_2O \rightarrow$ glutamate + NH_4^+	−3.4
Glycylglycine + $H_2O \rightarrow$ 2 glycine (a peptide bond)	−2.2
Glycosides:	
Sucrose + $H_2O \rightarrow$ glucose + fructose	−7.0
Maltose + $H_2O \rightarrow$ 2 glucose	−4.0
ESTERIFICATION	
Glucose + $P_i \rightarrow$ glucose-6-phosphate + H_2O	+3.3
REARRANGEMENT	
Glucose-1-phosphate $\rightarrow$ glucose-6-phosphate	−1.7
Fructose-6-phosphate $\rightarrow$ glucose-6-phosphate	−0.4
Glyceraldehyde-3-phosphate $\rightarrow$ dihydroxyacetone phosphate	−1.8
ELIMINATION	
Malate $\rightarrow$ fumarate + H_2O	+0.75
OXIDATION	
Glucose + $6O_2 \rightarrow 6CO_2 + 6H_2O$	−686
Palmitic acid + $23O_2 \rightarrow 16CO_2 + 16H_2O$	−2338
PHOTOSYNTHESIS	
$6CO_2 + 6H_2O \rightarrow$ six-carbon sugars + $6O_2$	+686

The data are from p. 397 of Lehninger, A. L. (1975) *Biochemistry*, 2nd edn. New York: Worth.

The chemical potential of H^+, from Eqn. (4.5), is

$$\mu_{H^+} = \mu^\circ_{H^+} + RT\ln a_{H^+} = \mu^\circ_{H^+} - 2.303RT(\text{pH}), \tag{4.43}$$

where we have assumed ideal conditions and used the definition of pH ($\log[H^+] \approx -\text{pH}$, see below). Combining Eqns. (4.42) and (4.43) gives

$$\Delta G = \mu^\circ_P = \mu^\circ_A - \nu(\mu^\circ_{H^+} - 2.303RT(\text{pH})) = \Delta G^\circ + \nu 2.303\text{RT}(\text{pH}). \tag{4.44}$$

The biochemist's standard state, however, is defined for pH = 7. So,

$$\Delta G^{\circ\prime} = \Delta G^\circ - \nu 16.121RT \tag{4.45}$$

The free energy difference varies linearly with T and ν, the number of protons transferred in the reaction. Happily, if neither H_2O nor H^+ is involved in the reaction, then $n = 0$ and $\nu = 0$, and $\Delta G^{\circ\prime} = \Delta G^\circ$.

J. Effect of temperature on K_{eq}

The concentrations of reactants and products depend on the physical conditions, so too must the equilibrium constant. We can see that K_{eq} varies with temperature as follows:

$$\ln K_{eq} = -\Delta G^\circ/RT = -(\Delta H^\circ/R)(1/T) + \Delta S^\circ/R. \tag{4.46}$$

(See Fig. 4.11.) As before, the superscript indicates standard state, but no prime appears because we have not specified pH 7. In general, ΔH° and ΔS° depend on T. Often, though, the enthalpy and entropy changes will be relatively weak functions of temperature, and a plot $\ln K_{eq}$ versus $1/T$, called a *van't Hoff graph*, will be approximately linear. The slope and intercept of the line are $-\Delta H^\circ/R$ and $\Delta S^\circ/R$, respectively. The upshot is that if K_{eq} can be measured at several temperatures in the range, ΔH° and ΔS° can be *measured*, albeit indirectly. ΔH° determined from the behavior of K_{eq} is called the *van't Hoff enthalpy*, or ΔH_{vH}.

If ΔH° and ΔS° depend significantly on temperature, as with proteins (because of the large ΔC_p of unfolding), as long as folding/unfolding is cooperative, the phase transition will occur over a relatively narrow temperature range, as in the melting of a pure solid, and ΔH° and ΔS° will be *approximately* constant in the range. A van't Hoff plot can then be used to estimate these thermodynamic functions *at the transition temperature*. Such indirect measurements of protein energetics have been corroborated by direct calorimetric measurements. In scanning microcalorimetry (Fig. 2.10B), the heat measurement is made at constant pressure, so $q = \Delta H = \Delta H_{cal}$, the *calorimetric enthalpy*, the area below the heat absorption peak. If the folding/unfolding reaction is highly cooperative and involves effectively only two states, then the ratio $\Delta H_{vH}/\Delta H_{cal} = 1$. Deviations from such behavior, whether from oligomerization of the folded state or stabilization of partly folded species, have an enthalpy ratio greater than or less than one, respectively. We shall return to this subject in Chapter 5 and discuss it in considerable detail in Chapter 6.

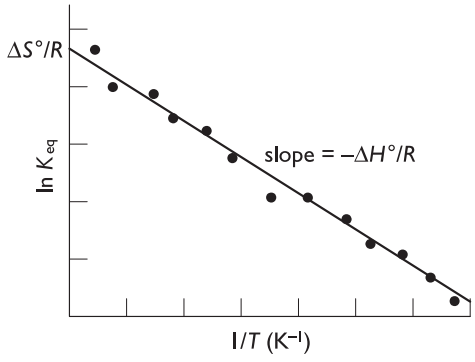


Fig. 4.11 A van't Hoff plot. This approach to data analysis will be useful when only two states are involved in the transition *and* if either of two conditions is met: ΔC_p is of negligible magnitude; or, if ΔC_p is large, the temperature range is small enough that ΔH° and ΔS° can be considered approximately independent of temperature. van't Hoff analysis is often used when data are collected by a technique that does not make a direct measurement of heat (e.g. any type of optical spectroscopy, NMR spectroscopy, viscometry, X-ray scattering, electrophoresis). See Chapter 7 for a discussion of multi-state equilibria.

How does a fluctuation in T translate into a fluctuation in $\ln K_{eq}$? In other words, how does $T \rightarrow (T + \Delta T)$ affect $\ln K_{eq}$? If the change brought about in $\ln K_{eq}$ by the temperature fluctuation is $\Delta \ln K_{eq}$, by Eqn. (4.44) we have

$$\ln K_{eq} + \Delta \ln K_{eq} = -(\Delta H^\circ/R)(1/(T + \Delta T)) + \Delta S^\circ/R. \tag{4.47}$$

To evaluate $1/(T + \Delta T) = (T + \Delta T)^{-1}$, we make use of a famous relationship in mathematics known as the binomial theorem, according to which $(x+y)^n = x^n + nx^{n-1}y + n(n-1)x^{n-2}y^2/2 + \ldots$ for $y^2 < x^2$. Our requirement is $\Delta T < T$, so the theorem can be applied! Making the substitution, we have

$$\begin{aligned} \ln K_{eq} + \ln K_{eq} &= -(\Delta H^\circ/R)[T^{-1} + (-1)T^{-2}\Delta T + (-1)(-2)T^{-3}(\Delta T)^2/2 + \cdots] \\ &\quad + \Delta S^\circ/R \ldots \\ &= -(\Delta H^\circ/R)(1/T) + (\Delta H^\circ/R)(T^{-2})(\Delta T) \\ &\quad - (\Delta H^\circ/R)(T^{-3})(\Delta T)^2/2 + \cdots + \Delta S^\circ/R. \end{aligned} \tag{4.48}$$

The first and last terms on the right-hand side of Eqn. (4.48) sum to $\ln K_{eq}$, so these terms can be dropped from both sides, leaving

$$\Delta \ln K_{eq} = (\Delta H^\circ/R)(T^{-2})(\Delta T) - (\Delta H^\circ/R)(T^{-3})(\Delta T)^2/2 + \cdots \tag{4.49}$$

If ΔT is small, then the second and following terms on the right-hand side will be much smaller than the first one. So we have

$$\Delta \ln K_{eq} \approx (\Delta H^\circ/R)(T^{-2})(\Delta T). \tag{4.50}$$

Does the approximation work? Let's test it! Suppose $\Delta H^\circ = 50$ kcal mol^{-1} at 300 K, a representative value for a small protein. If $\Delta T \sim 1$ K, then by Eqn. (4.50) $\Delta \ln K_{eq} = (50 \text{ kcal mol}^{-1}) \times (300 \text{ K})^{-2} \times (1 \text{ K}) / (1.9872 \text{ cal mol}^{-1} \text{ K}^{-1}) - (50 \text{ kcal mol}^{-1}) \times (300 \text{ K})^{-3} \times (1 \text{ K})^2 / (1.9872 \text{ cal mol}^{-1} \text{ K}^{-1}) = 0.2796 + 0.0009$. That is, in neglecting the second term, we incur an error of less than 1% – smaller than the usual uncertainty in measuring protein concentration! The error of the neglected terms would be greater if all of them were taken into account, but it can be shown that the total error would still be small. If, however, living organisms flourished at a temperature much closer to absolute zero, the relative magnitude of a 1 degree fluctuation would be much larger, the calculation we have just done would not involve division by such a large number, and the terms we have neglected would be much bigger. Life seems to depend on liquid water – the ability of water molecules to move around with a degree of independence – and on fluctuations in temperature and therefore thermal energy to be relatively small. The Antarctic bacteria mentioned in Chapter 1 are hardly "living" when frozen, even if they are not "dead." Finally, Eqn. (4.50) does provide a handy way of calculating the effect of temperature on $\ln K_{eq}$. You may have noticed, though, that we made a key simplifying assumption in the calculation – that the variation of ΔH° and ΔS° with temperature is small!

K. Acids and bases

A key application of the concept of equilibrium in biology concerns acids and bases in solution. According to the *Brønsted–Lowry definitions*, an *acid* is a proton donor and a *base* is a proton acceptor.[9] Biological situations usually concern water in some way or other. A measure of the acidity of water, due to the Danish biochemist Sørenson, is its pH, defined as

$$\mathrm{pH} = -\log a_{\mathrm{H_3O^+}}. \tag{4.51}$$

The pH of a solution determines the extent of proton dissociation from ionizable chemical groups in biological macromolecules and thus can have a profound effect on enzyme activity, protein–protein association, protein–DNA binding, and other types of biochemical reaction. We therefore had better put some effort into knowing this subject!

Suppose we have an acid HA. It participates in the following reaction in water:

$$\mathrm{HA(aq) + H_2O(l) \Leftrightarrow H_3O^+(aq) + A^-(aq)}. \tag{4.52}$$

The *acidity constant* for this reaction is *defined* as

$$K_\mathrm{a} = \left(\frac{a_{\mathrm{H_3O^+}} a_{\mathrm{A^-}}}{a_{\mathrm{HA}} a_{\mathrm{H_2O}}}\right) \approx \left(\frac{a_{\mathrm{H_3O^+}} a_{\mathrm{A^-}}}{a_{\mathrm{HA}}}\right). \tag{4.53}$$

The approximation is justified on the same grounds as before: if our concern is dilute aqueous solutions, the activity of water is close to 1 and essentially unchanging. At low concentrations of HA, the activity of hydronium ions is roughly equal to their molar concentration, and the acidity constant is usually written as

$$K_\mathrm{a} \approx [\mathrm{H_3O^+}][\mathrm{A^-}]/[\mathrm{HA}], \tag{4.54}$$

where the concentrations are in mol ℓ^{-1}. This approximation is valid only when *all* the ions in solution are present in low concentrations. Neglecting the limitations of the approximation can introduce complications. For instance, 1 mM imidazole[10] has an activity coefficient of about 0.95, and ignoring this can affect the equilibrium constant by as much as 10%! Acidity constant values are often tabulated in terms of their negative logarithm, $\mathrm{p}K_\mathrm{a}$:

$$\mathrm{p}K_\mathrm{a} = -\log K_\mathrm{a}. \tag{4.55}$$

The $\mathrm{p}K_\mathrm{a}$ is related to the standard state free energy of ionization as

[9] Johannes Nicolaus Brønsted, a Danish physical chemist, lived 1879–1947. Thomas Martin Lowry was an Englishman. Brønsted and Lowry introduced their definitions simultaneously but independently in 1923.

[10] A common buffer in biochemistry. The structure of imidazole is the same as that of the side chain of histidine. The $\mathrm{p}K_\mathrm{a}$ of imidazole is about 6.5, close to neutral pH.

Table 4.5. | pK_a' values of acidic and basic groups in proteins.

Group	Amino acid residue	pK_a'(25 °C)
α-Carboxyl		3.0–3.2
Carboxyl	Aspartic acid	3.0–4.7
	Glutamic acid	~4.5
Imidazolyl	Histidine	5.6–7.0
α-Amino		7.6–8.4
Sulfhydryl	Cysteine	9.1–10.8
Phenolic hydroxyl	Tyrosine	9.8–10.4
Guanidino	Arginine	11.6–12.6

Note that a range of values is given. The acidity constant will depend in part on the specific electronic environment of the acidic or basic group.

$$pK_a = \Delta G^\circ/(2.303RT). \tag{4.56}$$

Taking the logarithm of both sides of Eqn. (4.54) yields the *Henderson–Hasselbalch equation*:

$$pH = pK_a - \log([HA]/[A^-]). \tag{4.57}$$

The form of this relation is the same as that of Eqn. (4.7). Equation (4.57) tells us that the pK_a of an acid indicates the pH at which half of the protons are dissociated (when $[HA] = [A^-]$, $\log([HA]/[A^-]) = 0$). A buffer exhibits its greatest *buffering capacity* at this pH: a change in $[H_3O^+]$ or $[OH^-]$ has the smallest effect on the solution pH at the pK_a. The A^- ions at this pH can react with the largest amount hydronium ions produced on addition of strong acid, and the HA present can react with strong base added. The pK_a values of acidic and basic groups in proteins are given in Table 4.5. Proteins themselves are crucial for buffering fluids in living organisms. For example, about 80% of the buffering capacity of human blood comes from proteins, principally serum albumin and hemoglobin.

Let's use Eqn. (4.57) to calculate the pH of buffer solution containing 0.200 mol ℓ^{-1} KH_2PO_4 (monobasic potassium phosphate) and 0.100 mol ℓ^{-1} K_2HPO_4 (dibasic potassium phosphate). The equation describing the equilibrium between acid and base is as follows:

$$H_2PO_4^-(aq) + H_2O(1) \Leftrightarrow H_3O^+(ap) + HPO_4^{2-}(aq). \tag{4.58}$$

The pK_a for this reaction, which is based on measurement, can be found in standard tables and is 7.21. Plugging all the numbers into Eqn. (4.57) gives

$$pH = 7.21 - \log(0.2/0.1) = 6.91. \tag{4.59}$$

It is clear why potassium phosphate is a favorite buffer of biochemists.

Despite the neat appearance of the calculation we've just done, the situation with phosphate buffer is rather complex. This is

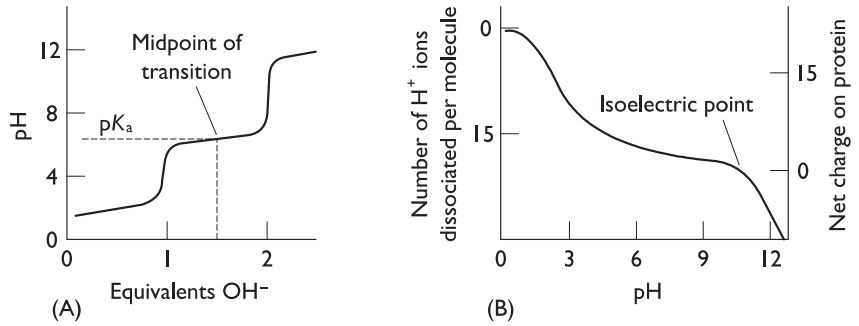


Fig. 4.12 Acid–base titration curves. Panel (A) shows titration of sodium phosphate at 25 °C. Phosphoric acid has three dissociable protons and therefore three pK_a values. The pK_as of phosphate are well-separated on the pH axis. This means that titration of the first site is effectively complete before titration of the second site begins, and so forth. Panel (B) shows the titration of ribonuclease at 25 °C. Ribonuclease is an extremely complicated molecule. Not only are there several different types of titratable group (the different types of ionizable amino acids), but the specific chemical environment of a dissociable proton can result in a substantial shift in its pK_a relative to the value for the free amino acid (in some cases by more than two pH units, or a 100-fold change in H_3O^+ concentration). This makes the titration curve of a protein very complicated. Panel (B) is based on a figure in Tanford & Hauenstein (1956).

because, unlike a simple acid like HCl, phosphoric acid is *polyprotic*. That is, phosphoric acid can donate more than one proton, so it has more than one pK_a. One of the other two pK_as is much higher than pH 7, the other is much lower. Fortunately, neither of the others makes a large contribution to the acid–base equilibrium at pH 7, and we can justify ignoring them in the neutral pH range. But don't neglect other contributions of a polyprotic acid if it has acidity constants that differ by less than about two pH units! Figure 4.12 shows how the net charge varies with pH for two polyprotic acids, phosphoric acid (panel (A)), in which there are three, well-separated ionization constants, and the protein ribonuclease (panel (B)), in which there are numerous ionizable groups of similar value. We'll have a further look at this topic in Chapter 6.

L. Chemical coupling

The Gibbs free energy is a state function. Individual contributions to the overall free energy change are therefore additive (see Fig. 2.3). Well, OK, but what does this say about biology? An endergonic reaction ($\Delta G > 0$) can be "powered" by an exergonic reaction ($\Delta G < 0$) if the two reactions are *chemically* "coupled" and the overall free energy change under the same conditions is negative. This is important! An analog will help to illustrate the situation. An example of *mechanical* coupling is the use of the downhill flow of a stream of water to turn a wheel to drive a system of gears to do something useful, for example, convert mechanical energy into electrical energy as at a hydroelectric dam or grind grain into flour.

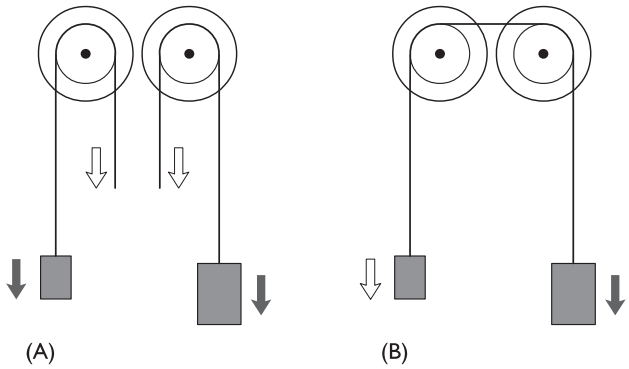


Fig. 4.13 Coupling. The figure shows a mechanical analog of chemical coupling, an extremely important means by which an endergonic reaction can occur. In panel (A), the two weights attached to strings are uncoupled. Both fall in accordance with Newton's law of gravitation. Neither weight does work on the other. In contrast, when the two weights are coupled, as in panel (B), one weight falls and the other one rises. The heavy weight does work on the light one. Note, however, that the net force acting on the "system" in panel (B) is less than that acting on the heavy weight in panel (A). The apparent mass of the heavy weight is the difference in mass between the coupled weights. Based on the figure on p. 167 of Atkins (1994).

Another type of mechanical coupling is depicted in Fig. 4.13. In panel (A), each weight represents an energetically favorable situation; both will fall to the ground spontaneously. The reverse reactions, which involve an increase in gravitational potential energy, will not occur spontaneously! But if the pulleys are coupled, as in panel (B), the more massive weight can be used to do work on the lighter one. Experimental studies have shown that coupling of biochemical reactions is an essential thermodynamic principle for the operation of metabolic pathways, including for instance the citric acid cycle (discussed in Chapter 5).

As an example of chemical coupling, consider the following two-step process:

$$A + B \Leftrightarrow C + D \qquad \Delta G_1, \tag{4.60}$$

$$D + E \Leftrightarrow F + G \qquad \Delta G_2, \tag{4.61}$$

Reaction (4.60) will occur spontaneously if $\Delta G_1 < 0$. But let's suppose $\Delta G_1 > 0$. Let's also suppose that the reaction of Eqn. (4.61) is spontaneous. Then the second reaction can be used to drive the first one if two conditions are met: the reactions involve a common compound (in this case, substance D), and the overall free energy change ($\Delta G_1 + \Delta G_2$) is negative. When both conditions are satisfied, the overall reaction proceeds spontaneously, even if the amount of compound D formed in Eqn. (4.60) is very small. The two reactions are said to be *coupled*, and D is called a *common intermediate* or *energy transducer*.

Let's look at a specific biochemical example. Transport of glucose into the cell is accompanied by phosphorylation. The negative charge on glucose prevents the diffusion of this valuable free energy molecule back out of the cell. The farmer will want to ensure that energy spent on harvesting wheat isn't wasted, so the grain is stored up a in building where the wind can't carry it off. The interior of the plasma membrane is made of hydrocarbon and, like the interior of a protein, is a region with a low dielectric. It is energetically unfavorable for something charged to pass into a membrane from bulk aqueous solution. Glucose is in the cell and phosphorylated. The overall coupled reaction here is one in which

a phosphoryl group is transferred from ATP to glucose, but it should be mentioned that neither of the relevant *half-reactions* (ATP hydrolysis or glucose phosphorylation) *obligatorily* drives the other reaction, as one might require in a more restricted definition of *coupled reaction*.

Let's look at some details of the energetics of glucose phosphorylation:

$$\begin{aligned} &\text{glucose} + P_i \Leftrightarrow \text{glucose-6-phosphate} + H_2O \\ &\Delta G^{\circ\prime} = 13.8\,\text{kJ mol}^{-1}. \end{aligned} \tag{4.62}$$

The blood concentration of glucose, the brain's primary fuel source, is ~5 mM. Given a cellular glucose concentration of about 300 μM, for the reaction to proceed to the right on the basis of concentration differences alone, the concentration of glucose-6-phosphate in the blood would have to be large – over 100 mM! (Can you prove it?) What actually happens in our bodies is that glucose phosphorylation is coupled to ATP hydrolysis. The overall reaction can be written as

$$\text{glucose} + \text{ATP} \Leftrightarrow \text{glucose-6-phosphate} + \text{ADP}. \tag{4.63}$$

This coupled reaction, for which $\Delta G^{\circ\prime} = -17.2$ kJ mol^{-1}, is clearly energetically favorable and will proceed to the right spontaneously – if the right "molecular hardware" is present.

Examination of Eqn. (4.63) can give a good sense of how this reaction coupling works on the molecular level. Minimizing the concentration of P_i would promote the forward reaction (by mass action), so an enzymatic nanomachine we might design to carry out the phosphorylation reaction should avoid the accumulation of P_i. Similarly, minimizing the concentration of H_2O in the vicinity of our molecule-sized "workbench" would minimize the probability of transfer of P_i from ATP to water; we want the phosphate group to be transferred to glucose! Crystallographic studies of the enzyme hexokinase, one of nature's own nanomachines, have revealed much about how the process actually occurs. Binding of glucose induces a conformational change in the enzyme that increases its affinity for ATP 50-fold and excludes water from the catalytic site. The functional groups of the amino acid side chains involved in catalysis move into proper alignment and a phosphoryl group is transferred from ATP to glucose. The low activity (concentration) of water in the active site of the enzyme is crucial for the reaction. Measurements have shown that the conformational change induced in the enzyme upon glucose binding results in the release of about 60 water molecules into bulk solution. This contributes a substantial increase in entropy to the overall energetics of the reaction, offsetting the unfavorable entropy change of bringing glucose and ATP simultaneously into an orientation that permits phosphoryl transfer to the appropriate hydroxyl group on glucose. Amazing!

M. Redox reactions

Utilization of the free energy of glucose and other nutrients consumed by organisms is controlled by means of *oxidation–reduction reactions*, or *redox reactions*. Some of these reactions occur in organelles called mitochondria, the "power-houses" of the cell. Redox reactions are of such great and general importance that much more than one section of one chapter of this book could be allocated to them. The present aim, however, is a brief introduction to the topic and not comprehensive treatment.

In redox reactions, electrons are transferred from one molecule to another. Electron transfer can be accompanied by the transfer of an atom or ion, but our main concern at the moment is electrons and changes in oxidation. The electron donor is called the *reductant* and the acceptor is the *oxidant*. Reductant and oxidant work in pairs, or couples.

Any redox reaction can be expressed as the difference of two reduction *half-reactions*. Conceptual in nature, half-reactions facilitate calculations by showing the electrons explicitly. An example of a half-reaction is the following:

$$\text{oxidized compound}(A_{ox}) + ne^- \rightarrow \text{reduced compound}(A_{red}), \tag{4.64}$$

where n electrons are transferred to the oxidized compound, giving the reduced compound. A more complicated redox reaction is

$$B_{ox} + ne^- + mH^+ \rightarrow B_{red}mH, \tag{4.65}$$

where n and m could be different. Note the proton (H^+) transfer. An example of a full redox reaction, where both redox pairs are given, is

$$A_{red} + B_{ox} \rightarrow A_{ox} + B_{red}. \tag{4.66}$$

A key physiological redox reaction is the reduction of oxygen to water by cytochrome *c*:

$$4\text{cyt}c^{2+} + 4H^+ + O_2 \rightarrow 2H_2O + 4\text{cyt}c^{3+}. \tag{4.67}$$

The transferred electrons are only implicit in this representation. Note that water is produced in this reaction.

The *standard redox potential*, $\Delta V^{\circ\prime}$, represents the strength of a redox pair to exchange electrons, the "electron transfer potential." A redox potential is measured as an *electromotive force* (e.m.f.), or *voltage*,[11] of a half-cell consisting of both members of the redox couple (Fig. 4.14). Just as the Gibbs free energy (a *thermodynamic* potential) must be measured with respect to a chosen reference state, so the redox voltage measurement is made relative to a standard value. In most cases the standard is provided by the

[11] The volt, the SI unit of electrical potential, or voltage, gets its name from Count Alessandro Giuseppe Antonio Anastasio Volta (1745–1827), an Italian physicist.

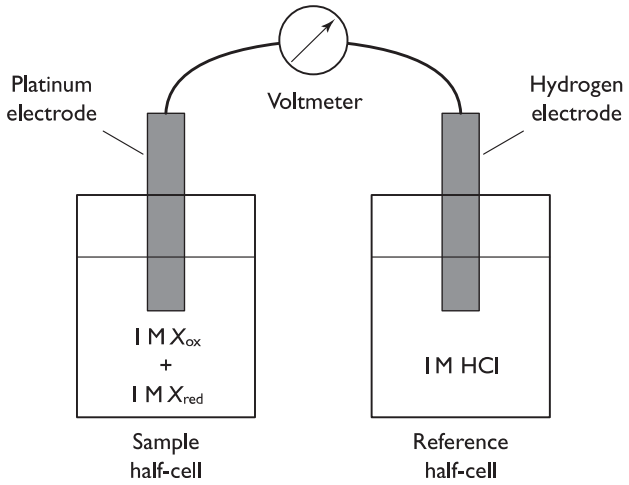


Fig. 4.14 Measurement of the standard redox potential. This device can be used to measure the standard redox potential of a compound. The hydrogen electrode in the reference cell contains 1 atm $H_2(g)$.

hydrogen electrode, the voltage of which is set to 0 V at pH 0. The redox potential is like the Gibbs free energy in that the magnitude depends on the concentrations of the reacting species.

To find the concentration-dependence of the redox potential, we note that work is done when a charge, $\boldsymbol{q}$, is moved through a difference in electrical potential, $\Delta \boldsymbol{V}$:

$$w' = -\boldsymbol{q}\Delta \boldsymbol{V}, \tag{4.68}$$

where the prime indicates *non-pV* work. This is the electrical work we mentioned in Chapter 2. As with *pV*-work, electrical work is the product of an "intensity factor" ($\boldsymbol{q}$) and a "capacity factor" ($\Delta \boldsymbol{V}$). If the amount of charge in Eqn. (4.68) is the same as that on 1 mole of protons, then $\boldsymbol{q} = F$, the Faraday constant (96.494 kJ V^{-1} mol^{-1}),[12] and the relationship is put on a molar basis. Recalling that the free energy change for a process is the maximum work that the system can do, and supposing that the stoichiometry of the reaction involves the transfer of n moles of *electrons*, we have

$$\Delta\mu = -nF\Delta \boldsymbol{V}. \tag{4.69}$$

The greater the number of positive charges transferred up a potential gradient, the greater the work that must be done by the system to bring about the change, and the less spontaneous the change. Rearranging Eqn. (4.69) in terms of $\Delta \boldsymbol{V}$ gives

$$\Delta \boldsymbol{V} = -\frac{\Delta\mu}{nF} \tag{4.70}$$

[12] The Faraday constant represents the electronic charge on 1 mole of electrons and is named after the renowned British physical chemist Michael Faraday (1791–1867), the first person to quantify the relationship between a chemical reaction and electric current. Faraday is a sort of anomaly in the history of science for his era: Most of the significant contributions to science were made by men of wealth or some form of high social status, and Faraday was from a poor family. Nevertheless, Faraday went on to become president of the Royal Institution, and the simple quantitative relationship he found between magnetic flux and induced e.m.f. was adopted by Maxwell as one of four elementary equations of electromagnetism.

Table 4.6. *Standard redox potentials of some important biochemical substrates*

Electrode equation	n	$\Delta V^{\circ\prime}$ (V)
$\text{Acetate} + 2H^+ + 2e^- \Leftrightarrow \text{acetaldehyde}$	2	−0.58
$2H^+ + 2e^- \Leftrightarrow H_2$	2	−0.42
$NAD^+ + H^+ + 2e^- \Leftrightarrow NADH$	2	−0.32
$\text{Pyruvate} + 2H^+\ 2e^- \Leftrightarrow \text{lactate}$	2	−0.19
Cytochrome c $(Fe^{3+}) + e^- \Leftrightarrow$ cytochrome c (Fe^{2+})	1	+0.22
$\frac{1}{2}O_2 + 2H^+ + 2e^- \Leftrightarrow H_2O$	2	+0.82

$$\Delta \boldsymbol{V} = -\frac{\Delta\mu^{\circ\prime} + RT\ln\frac{[\text{products}]}{[\text{reactants}]}}{nF} = \Delta \boldsymbol{V}^{\circ\prime} - \frac{RT}{nF}\ln\frac{[\text{products}]}{[\text{reactants}]}, \tag{4.71}$$

where the standard state potential, $\Delta\mu^{\circ\prime}/nF$, is written as $\Delta \boldsymbol{V}^{\circ\prime}$. Standard state conditions are pH 7.0, 1 atm, and 25 °C. The right hand side of Eqn. (4.71) is known as the *Nernst equation*.[13] As we shall see below, it has many applications in biological systems. At equilibrium,

$$\Delta \mathrm{V}^{\circ\prime} = -\frac{RT}{nF}\ln K'_{\text{eq}} \tag{4.72}$$

Now we have a way to relate the standard cell e.m.f. to the equilibrium constant. Some standard redox potentials are given in Table 4.6.

Equation (4.65) can be split into two parts,

$$\mathrm{A_{ox}} + n\mathrm{e}^- \rightarrow \mathrm{A_{red}} \tag{4.73}$$

and

$$\mathrm{B_{ox}} + n\mathrm{e}^- \rightarrow \mathrm{B_{red}}. \tag{4.74}$$

The half-cell reduction potentials here are

$$\Delta \boldsymbol{V}_{\mathrm{A}} = \Delta \boldsymbol{V}_{A}^{\circ\prime} - \frac{RT}{nF}\ln\frac{[\mathrm{A_{red}}]}{[\mathrm{A_{ox}^{n+}}]} \tag{4.75}$$

and

$$\Delta \boldsymbol{V}_{\mathrm{B}} = \Delta \boldsymbol{V}_{\mathrm{B}}^{\circ\prime} - \frac{RT}{nF}\ln\frac{[\mathrm{B_{red}}]}{[\mathrm{B_{ox}^{n+}}]}. \tag{4.76}$$

Note that the form of Eqns. (4.75) and (4.76) is exactly the same as that of Eqn. (4.7). The overall redox potential of any two half-reactions is

$$\Delta \boldsymbol{V} = \Delta \boldsymbol{V}_{\mathrm{e^-\,acceptor}} - \Delta \boldsymbol{V}_{\mathrm{e^-\,donor}} \tag{4.77}$$

A biochemical example will help to illustrate how these equations are put into practice. An important reaction of glycolysis is the reduction

[13] Named after Walther Hermann Nernst (1864–1941), the German physicist and chemist. Nernst is connected with the Third Law of Thermodynamics and was awarded the Nobel Prize in Chemistry in 1920.

of pyruvate to lactate by nicotinamide adenine dinucleotide (NADH, reduced form). The reaction is catalyzed with complete stereospecificity by the enzyme lactate dehydrogenase:

$$\text{pyruvate} + \text{NADH} + \text{H}^+ \Leftrightarrow \text{lactate} + \text{NAD}^+. \tag{4.78}$$

The sum of the respective half-cell potentials, $\Delta \boldsymbol{V}^{\circ\prime}$, $+0.32$ V and -0.19 $\boldsymbol{V}$ (Table 4.6), exceeds 0. Therefore, by Eqn. (4.69) $\Delta\mu^{\circ\prime} = -2 \times 96.5$ kJ V^{-1} mol$^{-1} \times 0.13$ V $= -25.1$ kJ mol^{-1}, the forward reaction is the more probable one, and we should expect the spontaneous oxidation of NADH by pyruvate. We note that, although they do not appear explicitly in Eqn. (4.78), two electrons are transferred in this reaction. Nicotinamide adenine dinucleotide phosphate (NADPH, reduced form), a close relative of NADH, plays a key role in the cellular redox reactions that enable the synthesis of compounds that are thermodynamically less stable than glucose, a starting material.

What if we change the concentrations of reactants and products? What effect will this have on the spontaneity of electron transfer? As we have seen, the standard state redox potential of the NAD_{ox}/NAD_{red} couple is -0.32 V. Now we wish to calculate the potential of this half-cell reaction under non-standard state conditions. Let the couple be 75% reduced at $T = 70$ °C. By Eqn. (4.75),

$$\Delta \boldsymbol{V} = -0.32\,\text{V} - \frac{8.314\ \text{J mol}^{-1}\ \text{K}^{-1} \times 343\ \text{K}}{2 \times 96\,494\ \text{J V}^{-1}\,\text{mol}^{-1}} \ln\left(\frac{75}{25}\right) = -0.34\,\text{V} \tag{4.79}$$

There is relatively little deviation from the standard state potential at 75% reduction. At 50% reduction, $\Delta \boldsymbol{V} = \Delta \boldsymbol{V}^{\circ\prime}$. Marked deviations from the standard state potential occur only for extremes of temperature or extremes of the ratio of the concentration of oxidant to concentration of reductant.

Before bringing down the curtain on the text of this chapter, let us return the spotlight to photosynthesis. This biological process makes sugar, a reduced form of carbon, using water, a reducing agent. The overall chemical reaction can be expressed as

$$6\ \text{CO}_2 + 6\ \text{H}_2\text{O} + \text{light} \rightarrow \text{C}_6\text{H}_{12}\text{O}_6 + 6\text{O}_2 + \text{energy}. \tag{4.80}$$

We covered some of the conceptual background to this equation in Chapter 1. Water is a relatively poor reductant, having a reduction potential of +820 mV; energy is required to separate electrons from water. This energy comes from the photons absorbed by chlorophylls *a* and *b* in photosystems I and II. Energy trapped by chlorophyll is transferred to the reaction center, whence electrons are transferred to pheophytin and plastoquinone. The reaction center is regenerated by "replacement" electrons from water, releasing the oxygen animals need to respire and generating a proton gradient across a lipid membrane. A more in-depth look at photosynthesis and glycolysis must wait until Chapter 5.

N. References and further reading

Arrhenius, S. (1912). *Theories of Solution*. New Haven: Yale University Press.

Atkins, P. W. (1994). *The Second Law: Energy, Chaos, and Form*, p. 167. New York: Scientific American.

Atkins, P. W. (1998). *Physical Chemistry*, 6th edn, cc. 10.2 & 23.2. Oxford: Oxford University Press.

Bada, J. L. & Miller, S. L. (1968). Equilibrium constant for the reversible deamination of aspartic acid. *Biochemistry*, **7**, 3403–8.

Banks, B. E. C. & Vernon, C. A. (1978). Biochemical abuse of standard equilibrium constants. *Trends in Biochemical Sciences*, **3**, N156–8.

Bates, R. G. (1973). *Determination of pH; Theory and Practice*. New York: John Wiley.

Ben-Shem, A., Frolow, F. & Nelson, N. (2003). Crystal structure of plant photosystem I. *Nature*, **426**, 630–5.

Bergethon, P. R. (1998). *The Physical Basis of Biochemistry: the Foundations of Molecular Biophysics*, cc. 2.17, 13 & 21.4–21.6. New York: Springer-Verlag.

Blandamer, M. J. (1992). *Chemical Equilibrium in Solution*. Hemel Hempstead: Ellis Horwood/Prentice-Hall.

Bradley, J. (1990). Teaching electrochemistry. *Education in Chemistry*, **27**, 48–50.

Bretscher, M. S. (1985). The molecules of the cell membrane. *Scientific American*, **253**, no. 4, 86–90.

Butcher, J. (2005). Profile: Lewis Gordon Pugh – polar swimmer. *The Lancet*, **366**, S23–4.

Christensen, H. N. & Cellarius, R A. (1972). *Introduction to Bioenergetics: Thermodynamics for the Biologist: A Learning Program for Students of the Biological and Medical Sciences*. Philadelphia: W. B. Saunders.

Creighton, T. E. (1993). *Proteins: Structures and Molecular Properties*, 2nd edn, ch. 7. New York: W. H. Freeman.

Dawes, E. A. (1962). *Quantitative Problems in Biochemistry*, 2nd edn, ch. 1. Edinburgh: E. & S. Livingstone.

Dawson, R. M. C., Elliott, D. C., Elliott, W. H. & Jones, K. M. (1986). *Data for Biochemical Research*, 3rd edn. Oxford: Clarendon Press.

DeKock, R. L. (1996). Tendency of reaction, electrochemistry, and units. *Journal of Chemical Education*, **73**, 955–6.

DeVries, A. L. (1982). Biological antifreeze agents in coldwater. *Comparative Biochemistry and Physiology*, **73A**, 627–40.

Ellis, R. J. (2001). Macromolecular crowding: Obvious but underappreciated. *Trends in Biochemical Sciences*, **26**, 597–604.

Encyclopdia Britannica CD98, "Activity Coefficient," "Arrhenius Theory," "Boiling Point," "Debye–Hückel Equation," "Equilibrium," "Free Energy," "Freezing Point," "Ideal Solution," "Life," "Melting Point," and "Phase."

Feeney, R. E. (1974). A biological antifreeze. *American Scientist*, **62**, 712–19.

Feiner, A. S. & McEvoy, A. J. (1994). The Nernst equation. *Journal of Chemical Education*, **71**, 493–4.

Franzen, H. F. (1988). The freezing point depression law in physical chemistry: is it time for a change? *Journal of Chemical Education*, **65**, 1077–8.

Fruton, J. S. (1999). *Proteins, Enzymes, Genes: the Interplay of Chemistry and Biology*. New Haven: Yale University Press.

Gennix, R. B. (1989). *Biomembranes: Molecular Structure and Function*. New York: Springer-Verlag.
German, B. & Wyman, J. (1937). The titration curves of oxygenated and reduced hemoglobin. *Journal of Biological Chemistry*, **117**, 533–50.
Gillispie, Charles C. (ed.) (1970). *Dictionary of Scientific Biography*. New York: Charles Scribner.
Good, N. E., Winter, W., Connolly, T. N., Izawa, S. & Singh, R. M. M. (1966). Hydrogen ion buffers for biological research. *Biochemistry*, **5**, 467–76.
Green, R. L. & Warren, G. J. (1985). Physical and functional repetition in a bacterial ice nucleation gene. *Nature*, **317**, 645–8.
Gurney, R. W. (1953). *Ionic Processes in Solution*. New York: McGraw-Hill.
Harned, H. S. & Owen, B. B. (1958). *The Physical Chemistry of Electrolytic Solutions*. Princeton: Van Nostrand-Reinhold.
Harold, F. M. (1986). *The Vital Force: a Study of Bioenergetics*, cc. 1 & 2. New York: W. H. Freeman.
Harris, D. A. (1995). *Bioenergetics at a Glance*, ch. 1. Oxford: Blackwell Science.
Haynie, D. T. (1993). *The Structural Thermodynamics of Protein Folding*, ch. 4. Ph.D. thesis, The Johns Hopkins University.
Hodgman, C. D. (ed.) (1957). *C.R.C. Standard Mathematical Tables*, 11th edn. Cleveland, Ohio: Chemical Rubber Company.
Hopkin, M. (2005). Man breaks world records with Antarctic swim. *Nature News*, 20 December.
Howlett, G. J., Blackburn, M. N., Compton, J. G. & Schachman, H. K. (1977). Allosteric regulation of aspartate transcarbamoylase. Analysis of the structural and functional behavior in terms of a two-state model. *Biochemistry*, **126**, 5091–9.
Hubbard, R. (1966). The stereoisomerization of 11-*cis*-retinal. *Journal of Biological Chemistry*, **241**, 1814–18.
Kemp, H. R. (1987). The effect of temperature and pressure on equilibria: a derivation of the van't Hoff rules. *Journal of Chemical Education*, **64**, 482–4.
Klotz, I. M. (1986). *Introduction to Biomolecular Energetics*, cc. 3–7. Orlando: Academic Press.
Kondepudi, D. & Prigogine, I. (1998). *Modern Thermodynamics: from Heat Engines to Dissipative Structures*, cc. 7.5, 8.2 & 8.3. Chichester: John Wiley.
Koryta, J. (1992). *Ions, Electrodes, and Membranes*. New York: John Wiley.
Lodish, H., Baltimore, D., Berk, A., Zipursky, S. L., Matsudaira, P. & Darnell, J. (1995). *Molecular Cell Biology*, 3rd edn, ch. 2. New York: W. H. Freeman.
Marshall, C. B., Fletcher, G. L. & Davies, P. L. (2004) Hyperactive antifreeze protein in a fish, *Nature*, **429**, 153.
McPartland, A. A. & Segal, I. H. (1986). Equilibrium constants, free energy changes and coupled reactions: concepts and misconceptions. *Biochemical Education*, **14**, 137–41.
Millar, D., Millar, I., Millar, J. & Millar, M. (1989). *Chambers Concise Dictionary of Scientists*. Cambridge: Chambers.
Nicholls, D. G. & Ferguson, S. J. (1992). *Bioenergetics 2*. London: Academic Press.
Ochs, R. S. (1996). Thermodynamics and spontaneity. *Journal of Chemical Education*, **73**, 952–4.
Ostro, M. J. (1987). Liposomes. *Scientific American*, **256**(1), 102–11.
Palmer, C. M., Siebke, K. & Yeates, D. K. (2004). Infrared video thermography: a technique for assessing cold adaptation in insects. *BioTechniques*, **37**, 212–17.
Peusner, L. (1974). *Concepts in Bioenergetics*, cc. 3, 5, 6 & 7. Englewood Cliffs: Prentice-Hall.

Price, G. (1998). *Thermodynamics of Chemical Processes*, ch. 4. Oxford: Oxford University Press.
Schrake, A., Ginsburg, A. & Schachman, H.K. (1981). Calorimetric estimate of the enthalpy change for the substrate-promoted conformational transition of aspartate transcarbamoylase from *Escherichia coli*. *Journal of Biological Chemistry*, **256**, 5005–15.
Schultz, S.G. (1980). *Basic Principles of Membrane Transport*. Cambridge: Cambridge University Press.
Segal, I.H. (1976). *Biochemical Calculations: How to Solve Mathematical Problems in General Biochemistry*, 2nd edn, ch. 3. New York: John Wiley.
Smith, C.A. & Wood, E.J. (1991). *Energy in Biological Systems*, cc. 1.3 & 1.4. London: Chapman & Hall.
Snell, F.M., Shulman, S., Spencer, R.P. & Moos, C. (1965). *Biophysical Principles of Structure and Function*. Reading, Massachusetts: Addison-Wesley.
Spencer, J.N. (1992). Competitive and coupled reactions, *Journal of Chemical Education*, **69**, 281–4.
Tanford, C. & Hauenstein, J.D. (1956). Hydrogen ion equilibria of ribonuclease. *Journal of the American Chemical Society*, **78**, 5287–91.
Tombs, M.P. & Peacocke, A.R. (1974). *The Osmotic Pressure of Biological Macromolecules*. Oxford: Clarendon Press.
Tydoki, R.J. (1995). Spontaneity, accessibility, irreversibility, 'useful work': the availability function, the Helmholtz function and the Gibbs function, *Journal of Chemical Education*, **72**, 103–12.
Tydoki, R.J. (1996). The Gibbs function, spontaneity, and walls. *Journal of Chemical Education*, **73**, 398–403.
van Holde, K.E. (1985). *Physical Biochemistry*, 2nd edn, cc. 2.1, 2.3, 2.4 & 3. Englewood Cliffs: Prentice-Hall.
Voet, D. & Voet, J.G. (1995). *Biochemistry*, 2nd edn, cc. 2–2, 3, 4, 11–2, 15–4–15–6, 16, 18–1, 19–1, 28–3 & 34–4B. New York: Wiley.
Williams, T.I. (ed.) (1969). *A Biographical Dictionary of Scientists*. London: Adam & Charles Black.
Wood, S.E. & Battino, R. (1996). The Gibbs function controversy. *Journal of Chemical Education*, **73**, 408–11.
Wrigglesworth, J. (1997). *Energy and Life*, cc3, 5.7.2, 7.1, 7.3 & 7.5.1. London: Taylor & Francis.

O. Exercises

1. State whether the following phrases pertain to (A) the expansion of a gas into a vacuum, (B) two objects coming to thermal equilibrium, (C) both of these processes, or (D) neither of these processes.
(1) Involves a change in enthalpy.
(2) Involves an increase in entropy.
(3) Involves a decrease in Gibbs free energy.
(4) Can be made to proceed in the opposite direction.

2. State whether the following phrases pertain to (A) spontaneity, (B) reversibility, (C) both spontaneity and reversibility, or (D) neither spontaneity nor reversibility.

(1) Established for $\Delta G < 0$ at constant T.
(2) Established for $\Delta S < 0$.
(3) Established for a process in which the work done is a maximum.
(4) Illustrated by the migration of a solute from a region of high concentration to low concentration.
(5) Required for determination of ΔS by the heat transferred.
(6) Implies that the coupling of coupled reaction is very efficient.

3. State whether the following phrases pertain to (A) ΔU, (B) ΔG, (C) both ΔU and ΔG, or (D) neither ΔU nor ΔG.
(1) Does not depend on pathway during a change of state.
(2) Consists of the heat transferred and the work done.
(3) Must be negative if an isothermal and isobaric process is spontaneous.
(4) Measures the degree of disorder of a system.
(5) Is zero at equilibrium for an isothermal and isobaric process.
(6) Used to determine whether one reaction can drive another by coupling.
(7) Includes only energy that can do work, at constant temperature and pressure.

4. State whether the following phrases pertain to (A) ΔG, (B) ΔG°, (C) both ΔG and ΔG°, or (D) neither ΔG nor ΔG°.
(1) Equals $-RT \ln K_{eq}$.
(2) Equals $-nF\Delta V^{\circ\prime}$.
(3) Is zero if the change in the state of the system is spontaneous.
(4) Equals $\Delta H - T\Delta S$ at constant T.
(5) Is given for one mole of the reactants for a given reaction.
(6) Is equal to the sum of the chemical potentials of the products minus the chemical potentials of the reactants, with each chemical potential multiplied by the number of moles involved in the reaction.
(7) Is independent of the concentration of the components of a reaction.

5. State whether the following phrases pertain to (A) a, (B) μ, (C) both a and μ, or (D) neither a nor μ.
(1) Equals the concentration times the activity coefficient.
(2) Needed to calculate ΔG if ΔG° is known under specified conditions for all components of a reaction.
(3) Used to calculate ΔG for a process after multiplication by the number of moles of that component involved in the process.

6. In Chapter 1 we said that all living organisms depend on the Sun in order to meet the energy requirements of life. This is only partially true of the chemosynthetic bacteria that live at the bottom of the ocean. Explain the energy requirements for life in completely general terms. Although it may be that the Sun

played an indispensable role in the formation of life as we know it, is the Sun absolutely necessary for life? Why or why not?

7. What are the units of K_{eq}? Explain.

8. Calculate ΔG°(25 °C) for $K_{eq} = 0.001$, 0.01, 0.1, 1, 10, 100, and 1000.

9. The multi-component enzyme aspartate transcarbamoylase catalyzes the formation of *N*-carbamoylaspartate from carbamoyl phosphate and aspartate. Arthur Pardee has demonstrated that this reaction is the first step unique to the biosynthesis of pyrimidines, including cytosine, thymine and uracil, major components of nucleic acids. Aspartate transcarbamoylase has at least two stable folded conformations, known as R (high substrate affinity) and T (low substrate affinity). Interestingly, the relative stability of the T and R states is affected by the binding of ATP (a purine) to R and CTP (a pyrimidine) to T, a topic covered in Chapter 7. Measurement of the standard state free energy difference between R and T in the absence of ATP and CTP yielded the value 3.3 kcal mol^{-1}. Calorimetric determination of ΔH° for the transition was –6 kcal mol^{-1}. Calculate the standard state entropy change for the T → R transition.

10. When a photon in the visible range is absorbed in the retina by rhodopsin, the photoreceptor in rod cells, 11-*cis*-retinal is converted to the all-*trans* isomer. Light energy is transformed into molecular motion. The efficiency of photons to initiate the reaction is about 20% at 500 nm (57 kcal mol^{-1}). About 50% of the absorbed energy is available for the next signaling step. This process takes about 10 ms. In the absence of light, spontaneous isomerization of 11-*cis*-retinal is *very* slow, on the order of 0.001 yr^{-1}! Experimental studies have shown that the equilibrium energetics of retinal isomerization are $\Delta S^{\circ} = 4.4$ cal $mol^{-1}K^{-1}$ and $\Delta H^{\circ} = 150$ cal mol^{-1}. Calculate the equilibrium constant for the reaction.

11. Which one of the following equations is used to evaluate free energy changes in cells under physiological conditions? What makes it appropriate?
(a) $\Delta G = RT \ln K_{eq}'$.
(b) $\Delta G = \Delta G^{\circ\prime} + RT \ln$[products]/[reactants].
(c) $\Delta G = RT \ln$[products]/[reactants].
(d) $\Delta G = \Delta H - T\Delta S$.
(e) $\Delta G = \Delta G^{\circ\prime} + RT$ [products]/[reactants].

12. The direction of a reaction with a very large or very small value of K_{eq} is difficult, though not impossible, to alter by changing the mass action ratio. Explain.

13. Show that for a reaction at 25 °C which yields 1 mol of H_2O, $\Delta G^{\circ\prime} = \Delta G^{\circ} + 9.96$ kJ mol^{-1}.

14. Calculate K_{eq} for the hydrolysis of the following compounds at neutral pH and 25 °C: phosphoenolpyruvate ($\Delta G^{\circ\prime} = -61.9$ kJ mol^{-1}), pyrophosphate ($\Delta G^{\circ\prime} = -33.5$ kJ mol^{-1}), and glucose-1-phosphate ($\Delta G^{\circ\prime} = -20.9$ kJ mol^{-1}). Assume that the equilibrium constant includes water, accounting for the possibility that the water concentration is relatively low, as in the cell. These compounds are involved in the glycolytic pathway.

15. $\Delta G^{\circ\prime}$ for the conversion of fructose 1,6-bisphosphate (FBP) into glyceraldehyde-3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP) is +22.8 kJ mol^{-1}. This reaction is step four of the glycolytic pathway and is catalyzed by aldolase. In the cell at 37 °C the mass action [DHAP]/[GAP] = 5.5. What is the equilibrium ratio of [FBP]/[GAP] when [GAP] = 2×10^{-5} M? When [GAP] = 1×10^{-3} M?

16. Calculate ΔG when the concentrations of glucose-1-phosphate and glucose-6-phosphate are maintained at 0.01 mM and 1 mM, respectively. Compare the sign of ΔG with what was obtained in the worked example above. Suggest how this might be significant in metabolism.

17. Lactate dehydrogenase (LDH) catalyzes the oxidation of pyruvate to lactate and NADH to NAD^+ in glycolysis, the pathway by which glucose is converted to pyruvate with the generation of 2 mol of ATP mol^{-1} of glucose. The reaction is particularly important during strenuous activity, when the demand for ATP is high and oxygen is depleted. The relevant half-reactions and their standard reduction potentials are given in Table 4.6. Calculate ΔG for the reaction under the following conditions: [lactate]/[pyruvate] = [NAD^+]/[NADH] = 1; [lactate]/[pyruvate] = [NAD^+]/[NADH] = 160; [lactate]/[pyruvate] = [NAD^+]/[NADH] = 1000. What conditions are required for the reaction to spontaneously favor oxidation of NADH? [NAD^+]/[NADH] must be maintained close to 10^3 in order for the free energy change of the glyceraldehyde-3-phosphate reaction to favor glycolysis. This function is performed by LDH under anaerobic conditions. What is the largest [lactate]/[pyruvate] can be in order for the LDH reaction to favor the production of NAD^+ and maintain [NAD^+]/[NADH] = 10^3?

18. The citric acid cycle is the common mode of oxidative degradation in eukaryotes and prokaryotes (Chapter 5). Two components of the citric acid cycle are α-ketoglutarate and isocitrate. Let [NAD_{ox}]/[NAD_{red}] = 8; [α-ketoglutarate] = 0.1 mM; [isocitrate] = 0.02 mM. Assume 25 °C and pH 7.0. Calculate ΔG. Is this reaction a likely site for metabolic control? Explain.

19. Refer to Fig. 4.4. Mixing. In the text we noted that at first entropy is at a minimum and free energy is at a maximum.

Later, ... if the two liquids are the same, what are ΔS and ΔG of mixing?

20. Refer to Fig. 3.2. The stopcock is closed in panel (A). All of the inert gas is in the bulb on the left-hand side. In panel (B), the bulb on the left-hand side has inert gas at concentration x, and the bulb on the right has the same inert gas at concentration x. What are the entropy and the free energy differences between panels (A) and (B)?

21. Rationalize the change of sign in Eqn. (4.32).

22. Cytochromes are redox-active proteins that occur in all organisms except a few types of obligate anaerobes. These proteins contain heme groups, the iron atom of which reversibly alternates between the Fe(II) and Fe(III) oxidation states during electron transport. Consider the reaction

$$\text{cyt}c\ (\text{Fe}^{2+}) + \text{cyt}f\,(\text{Fe}^{3+}) \Leftrightarrow \text{cyt}c\ (\text{Fe}^{3+}) + \text{cyt}f\,(\text{Fe}^{2+})$$

involving cytochromes c and f. If $V^{\circ\prime} = 0.365$ V for electron transfer to cyt$f$ (Fe^{3+}), and $V^{\circ\prime} = 0.254$ V for electron transfer to cyt$c$ (Fe^{3+}), can ferrocytochrome c (2+ oxidation state) reduce ferricytochrome f (3+ oxidation state) spontaneously?

23. Calculate $\Delta \boldsymbol{V}$ in Eqn. 4.71 when the couple is 99% reduced and the temperature is 37 °C.

24. Table 4.1 presents the thermodynamic properties of water. On the basis of these data, rationalize the suitability, or lack thereof, of each thermodynamic function as an index of spontaneous change.

25. Cholesterol increases membrane rigidity. What effect will it have on the character of the lipid bilayer order–disorder transition? Why?

26. Some organisms are able to tolerate a wide range of ambient temperature, for instance bacteria and poikilothermic (cold-blooded) animals such as fish. The membrane viscosity of *E. coli* at its growth temperature is approximately constant over the range $15 - 43$ °C. Knowing aspects of the physical basis of the solid-gel transition in lipid bilayers, suggest how bacteria and fish might cope with changes of temperature.

27. Use your knowledge of the physical properties of lipids to outline several design characteristics of a liposome-based drug delivery system. A liposome is a bilayer structure that is self-enclosing and separates two aqueous phases.

28. The reversible deamination of aspartate yields ammonium and fumarate. Fumarate is a component of the citric acid cycle. Aspartate deamination is catalyzed by the enzyme aspartase.

Experimental studies on the deamination reaction have shown that

$$\log K_{eq} = 8.188 - (2315.5/T) - 0.010\,25T,$$

where T is in degrees kelvin (K). Note that the units of the coefficient of $1/T$ (i.e. 2315.5) must be K, while those of 0.01025 are K^{-1}. Calculate ΔG° at 25°C. Remember that $K = 10^{\log K}$ and $2.303\log x \approx \ln x$. Follow the development leading up to Eqn. (4.50) to show that $\Delta H^\circ = 2.303 \times R \times (2315.5 - 0.010\ 25T^2)$. Calculate ΔH° and ΔS° at 25 °C. From Chapter 2, $\Delta C_p = \Delta(\Delta H^\circ)/\Delta T$. Use this to show that $\Delta C_p = -2.303 \times R \times 0.0205T$. Evaluate ΔC_p at 25 °C.

29. State whether the following phrases pertain to (A) chemical potential of the solute, (B) chemical potential of the solvent, (C) both of these chemical potentials, or (D) neither chemical potential.
(1) Equals $RT \ln a$.
(2) Equals $\mu^\circ + RT \ln a$.
(3) At equilibrium, its value is the same on both sides of a membrane.
(4) Is proportional to the osmotic pressure (see Chapter 5).

30. Calculate the value of x for which the approximation $\ln(1 + x) \approx x$ gives an error of 5%.

31. State whether the following phrases pertain to (A) ΔG, (B) $\Delta \boldsymbol{V}$, (C) both ΔG and $\Delta \boldsymbol{V}$, or (D) neither ΔG nor $\Delta \boldsymbol{V}$.
(1) Indicates whether an oxidation-reduction reaction is spontaneous.
(2) Standard value for a reaction is determined with all components in their standard states.
(3) Is positive for a spontaneous reaction.
(4) Is called the standard electrode reduction potential.
(5) Can be used to calculate the equilibrium constant for a reaction for a known set of concentrations of all components of a reaction at a given temperature.

32. Chemical coupling. The equilibrium constant for $Glu^- + NH_4^+ \Leftrightarrow Gln + H_2O$ is 0.003 15 M^{-1} at pH 7 and 310 K; the reaction lies far to the left. The synthesis of Gln from Glu is made energetically favorable by coupling it to hydrolysis of the terminal phosphodiester bond of ATP. The products of ATP hydrolysis are ADP and P_i. The equilibrium constant for the coupled reaction, which is known from experiments with glutamine synthase, is 1200. Calculate the phosphate bond energy in ATP at pH 7 and 310 K.

33. What is the pH value of 0.001 M HCl solution?

34. Calculate the hydrogen ion concentration of solution of pH 6.0.

35. Calculate the ionic strength of a 0.35 molal aqueous solution of $MnCl_2$. Assume that dissociation of the salt into ions is complete at this concentration.

36. Calculate the ionic strength of 0.01 N acetic acid if the dissociation constant of the acid is 1.8×10^{-5}.

37. Calculate the activity coefficient and activities of the ions in aqueous solution of (a) 5 mM H_2SO_4 and (b) 2 mM NaCl.

38. The following data were obtained by German and Wyman (1937) for horse hemoglobin, an oxygen-binding blood protein, in 0.333 M NaCl.

Deoxygenated hemoglobin		Oxygenated hemoglobin	
Acid (−) or base (+) per gram Hb	pH	Acid (−) or base (+) per gram Hb	pH
−0.514	4.280	−0.514	4.280
−0.452	4.415	−0.453	4.410
−0.419	4.525	−0.420	4.525
−0.390	4.610	−0.392	4.618
−0.323	4.842	−0.324	4.860
−0.258	5.160	−0.259	5.188
−0.224	5.320	−0.225	5.430
−0.172	5.590	−0.173	5.800
−0.130	6.072	−0.130	6.055
−0.064	6.541	−0.063	6.430
0.0	6.910	+0.001	6.795
+0.070	7.295	+0.072	7.130
+0.131	7.660	+0.133	7.510
+0.171	7.860	+0.172	7.725
+0.208	8.140	+0.209	8.043
+0.254	8.545	+0.254	8.450
+0.288	8.910	+0.288	8.890
+0.311	9.130	+0.292	8.990
+0.331	9.350	+0.311	9.130
+0.350	9.410	+0.331	9.355
+0.357	9.465	+0.350	9.410
+0.407	9.800	+0.357	9.480
		+0.407	9.800

Plot the titration data to find which form of hemoglobin is the stronger acid. The stronger an acid, the more readily it gives up protons. We shall study hemoglobin in considerably greater depth in Chapters 5 and 7.

39. The history of science is full of "partly true" ideas pursued with vigor until they no longer became tenable. As we have

seen in Chapter 2, Galileo's assumption about the shapes of planetary orbits, which was based on the speculations of thinkers of classical antiquity, was eventually superseded by the very detailed measurements of the Danish astronomer Tycho Brahe (1546–1601) and analysis of Johannes Kepler. Similarly, Galileo's work on the relative motion of bodies was a great addition to the physics of his day (mostly that of Aristotle), and it prepared the way for Newton; but in the twentieth century, Galilean (Newtonian) relativity is seen to be a limiting case of the more general view proposed by Einstein. In the nineteenth century, research in thermochemistry was motivated in part by the belief that the heat of a reaction measured its "affinity": the greater the energy liberated, the greater the affinity of the reactants for each other. This view became untenable by the discovery of spontaneous endothermic reactions. Explain.

40. ΔG cannot generally be equated with ΔG°. To a very good first approximation ΔH can be equated with ΔH°. Explain.

41. Calculate the percentage non-ionized for an acid with a pK_a of 4 in an environment of pH 1.

42. Tris, a base, is a popular buffer for biochemical research. Its pK_a is strongly dependent on temperature. Would it make a very good buffer for a scanning calorimetry experiment? Why or why not? Assuming that $K_a = 8.3 \times 10^{-9}$ M, calculate the ratio of acid to base at pH 8.0. Let the total concentration of Tris be 150 mM, and divide the stock into to two parts. To one, add 10 mM HCl. How does the pH change? Hint: assume complete dissociation of HCl. To the other, reduce the concentration to 30 mM. How does the pH change?

43. The concentration of creatine in urine is *c.* 40-fold greater than in serum. Calculate the free energy change per molecule required for transfer of creatine from blood to urine at 37°C.

44. Which of the following redox pairs is the strongest reducing agent?

Redox pair	$V^{\circ\prime}$ in volts
Oxidized ferrodoxin/reduced ferrodoxin	−0.43
NADP/NADPH	−0.32
Oxidized glutathione/reduced glutathione	−0.23
Pyruvate/lactate	−0.19
Ubiquinone/hydroquinone	0.10

Chapter 5

Gibbs free energy – applications

A. Introduction

The Gibbs free energy is important in biology research because it enables one to predict the direction of spontaneous change for a system under the constraints of constant temperature and pressure. These constraints generally apply to all living organisms. In the previous chapter we discussed basic properties of the Gibbs free energy, showed how its changes underlie a number of aspects of physical biochemistry, and touched on what the biological scientist might do with such knowledge. Here, we build on the introductory material and explore how it can be applied to a wide variety of topics of interest to the biological scientist. A range of examples illustrate when, where, why, and how the Gibbs free energy is such a useful concept. We shall discuss the energetics of different types of biological structure, including small organic molecules, membranes, nucleic acids, and proteins. This will help to give a deeper sense of the relatedness of some seemingly very different topics one encounters in biological science.

B. Photosynthesis, glycolysis, and the citric acid cycle

This section presents a low-resolution view of the energetics of photosynthesis, glycolysis, and the citric acid cycle. There can be no doubt that the details we omit are important: entire books have been written on each subject! But our aim here is to consider biological energy in a global, qualitative way. We want to try to see "the big picture." So many of the protein, enzyme, chemical intermediate players do not have a speaking part in the present dramatic performance. Such details can be found in any good biochemistry textbook.

Over 99% of the free energy in our biosphere is from the Sun. Green plants, certain unicellular organisms like diatoms,

cyanophytes (blue-green algae), and various kinds of bacteria, collectively known as photoautotrophs, convert the light energy of the Sun and CO_2 into the chemical energy of bonding electrons in sugar molecules. The energy conversion process is called *photosynthesis*. The remaining less than 1% of our biosphere's free energy comes from the oxidation of inorganic matter, mainly hydrogen and sulfur, by microorganisms called chemolithotrophs. Whether photoautotrophs preceded or followed chemolithotrophs in the flowering of life on Earth is an intriguing open question (see Chapter 9).

The overall chemical reaction of photosynthesis is:

$$CO_2 + H_2O + \text{light} \rightarrow (CH_2O) + O_2. \quad (5.1)$$

CO_2 and H_2O are reduced to sugar and oxygen in this redox reaction. The process carried out in photosynthetic protists and cyanophytes resembles that in green plants, while compounds other than water serve as a reactant in photosynthetic bacteria and oxygen is not produced. All photosynthetic organisms[1] contain the light-absorbing pigment chlorophyll (Fig. 1.3). This molecule plays a key role in the transformation of light energy to chemical compounds. Chlorophyll, like the heme group (see below) of the vertebrate oxygen transport protein hemoglobin and the heme group of the electron transport protein cytochrome *c*, is derived from protoporphyrin IX, a complex ring structure synthesized from glycine and acetate (Fig. 5.1).

Figure 5.2 depicts the energetics of photosynthesis in schematic form. Absorption of photons ($h\nu$) results in the ejection of electrons from P680, the reaction center chlorophyll of photosystem II.[2] Each electron passes through a chain of electron carriers to plastoquinone, giving plastoquinol. By means of a series of redox reactions, the electrons are delivered to plastocyanin, which regenerates photooxidized P700, the reaction center chlorophyll of photosystem I. The electron ejected from P700 then passes through a chain of electron carriers to the oxidized form of nicotinamide adenine dinucleotide phosphate ($NADP^+$), an intracellular electron carrier. Photosynthetic electron transport drives the formation of a proton (pH) gradient, a difference in the concentration of protons on opposite sides of a membrane (in plants, the thylakoid membrane in chloroplasts). Movement of protons from a region of high chemical potential to low chemical potential powers the synthesis of ATP in manner that closely resembles oxidative phosphorylation, the endergonic synthesis of ATP from ADP and P_i in mitochondria in animal cells (see below). Plants also use light energy to make cellulose and other sugar molecules.

Glucose is the six-carbon sugar that is quantitatively the most important source of energy for cellular processes in all known

[1] Excluding halobacteria but including all other types of photosynthetic prokaryotes. Halobacteria thrive in the high salt environment of the Dead Sea.

[2] So called because 680 nm is the wavelength of the absorption maximum of the reaction center chlorophyll.

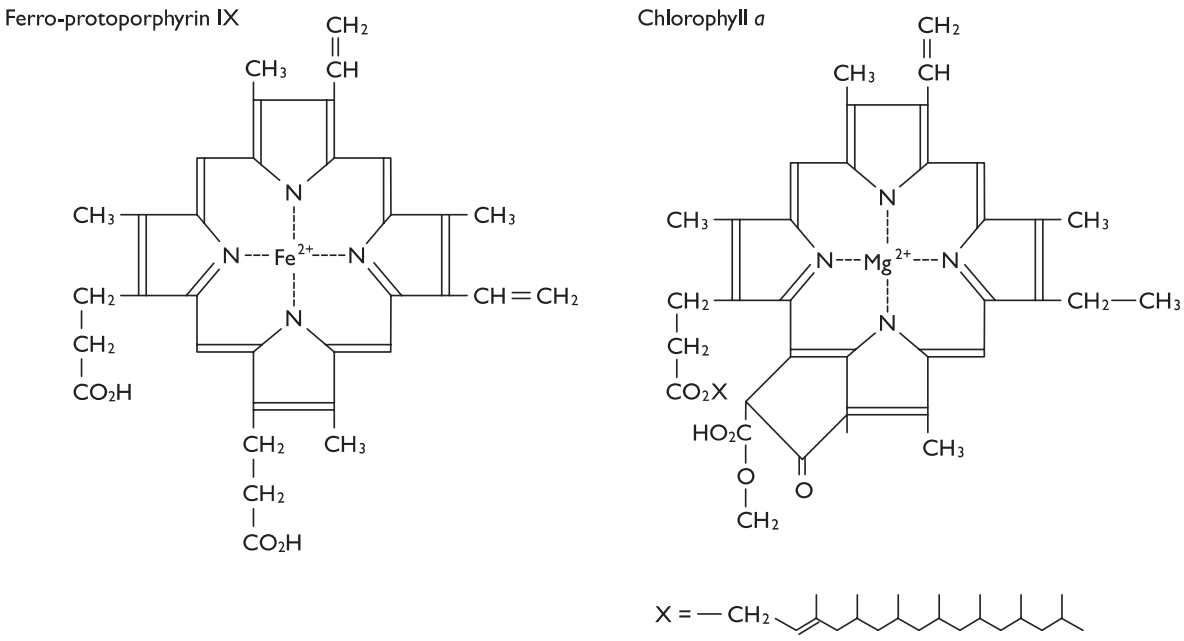


Fig. 5.1 Molecular formulas of the heterocyclic ring systems of ferro-protoporphyrin IX (heme) and chlorophyll a. The bound iron of heme is normally in the Fe(II) (ferrous, 2+) oxidation state regardless of whether oxygen is bound. The structure of chlorophyll b is nearly identical to that of chlorophyll a: a formyl group is found in place of a methyl group. Bacteriochlorophylls a and b, which are important in photon capture in photosynthetic bacteria, are very similar in structure to chlorophyll a. The long aliphatic tail of chlorophyll increases its solubility in a nonpolar environment. Note that in both heme and chlorophyll a divalent cation is bound. It is a remarkable indication of the unity of all known living things that such similar ring structures should play important roles in biological energetics in organisms as different as bacteria and humans.

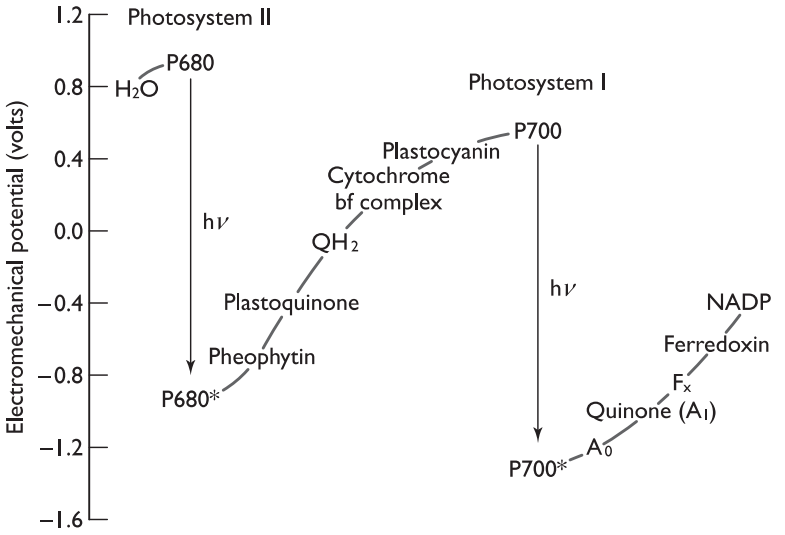


Fig. 5.2 Schematic diagram ("Z-scheme") of the energetics of electron transport in photosynthesis. The electrochemical potential (free energy) is measured in volts. Electrons tend to flow spontaneously from a state of higher to lower free energy. In terms of electrochemical potential, electrons migrate spontaneously from a more negative to a more positive reduction potential. PSII is coupled to PSI via the quinone Q and plastocyanin.

organisms. *Glycolysis*, the metabolism[3] of glucose, is a sequence of biochemical reactions by which one molecule of glucose is oxidized to two molecules of pyruvate, a three-carbon molecule (Fig. 5.3).

[3] Greek, *metabolikon*, disposed to cause or suffer change; coined by the German biologist Theodor Schwann (1810–1882).

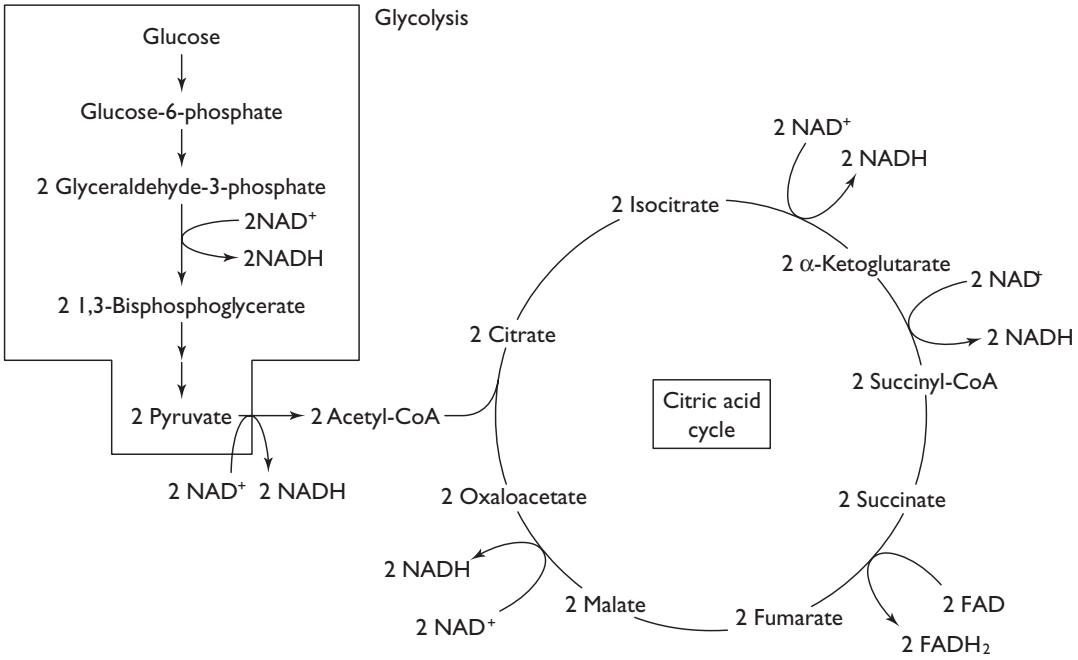


Fig. 5.3 Schematic diagram of glycolysis and the citric acid cycle. The figure shows the points at which the electron carriers NAD^+ and FAD are reduced by electron transfer to form NADH and $FADH_2$.

Pyruvate is then converted by a series of reactions to carbon dioxide and water. In combination with other aspects of oxidative carbohydrate metabolism, glycolysis is essentially the reverse process of photosynthesis. The *overall* chemical reaction for glucose metabolism is

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O. \qquad (5.2)$$

Compare Eqn. (5.2) with Eqn. (5.1). The free energy change for the *complete* redox reaction is $\Delta G^{\circ\prime} = -2823$ kJ mol^{-1}, and 24 electrons are transferred in the process. The standard state free energy change ($\Delta G^{\circ\prime}$) for glycolysis alone is -43.4 kJ mol^{-1}, while the physiological free energy change (ΔG) for glycolysis, which includes the synthesis of 2 moles of ATP, is -74 kJ mol^{-1}. Figure 5.4 depicts the physiological energetics of glycolysis in schematic form.

Glycolysis is similar in all organisms. Once a glucose molecule has entered the cell, it is immediately phosphorylated at the expense of one molecule of ATP. (It is interesting that ATP is expended in a process which, as we shall see, leads to ATP production.) Glucose phosphorylation is an essentially irreversible reaction because the free energy change of removal of the phosphoryl group from ATP is large and negative. Phosphorylation ensures that once it has entered the cell, the chemical energy of glucose is trapped there. The fate of pyruvate depends on the organism, tissue and conditions. In stressed, oxygen-depleted skeletal muscle, for instance, pyruvate is converted to lactate (the conjugate base of lactic acid) and one molecule of ATP is produced. Fermentation of yeast, a process integral to making bread, beer, and wine, involves the

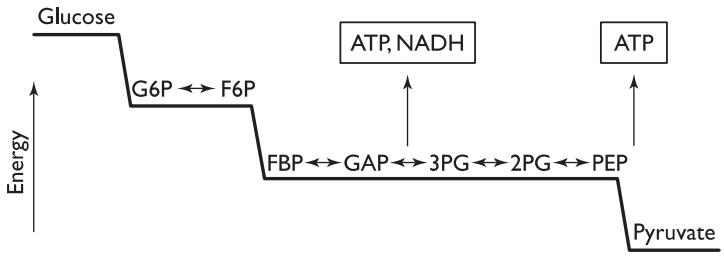
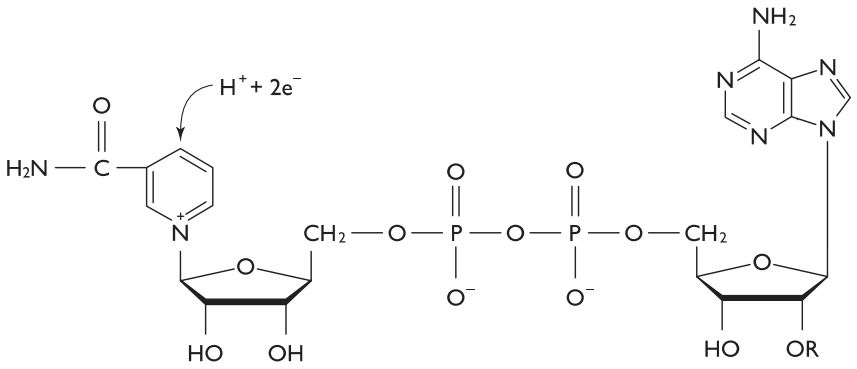


Fig. 5.4 Control by phosphofructokinase (PFK) of the flux of glycolysis breakdown products. Abbreviations: G6P, glucose-6-phosphate; F6P, fructose-6-phosphate; FBP, fructose bisphosphate; GAP, glyceraldehyde-3-phosphate; 3PG, 3-phosphoglycerate; 2PG, 2-phosphoglycerate; PEP, phosphoenolpyruvate. The physiological free energy changes (in kJ mol^{-1}) are: −27.2, − 1.4, −25.9, −5.9, +3.3, −0.6, −2.4, −13.9 (from Table 16–1 in Voet and Voet (1995)). There are three irreversible steps in the metabolism of glucose to pyruvate. These occur between glucose and G6P, F6P and FBP, and PEP and pyruvate. The irreversibility of these reactions is extremely important for cellular function. For only at an irreversible step of a process can control be exerted; irreversibility permits regulation of the speed of the reaction. The most important regulatory enzyme of glycolysis is PFK. This allosteric enzyme has four subunits and is controlled by several activators and inhibitors (see Chapters 7 and 8). PFK catalyzes the conversion of F6P to FBP. Because regulation of a pathway at a particular point affects all reactions that occur downstream, PFK controls the flux of glycolysis. Based on Fig. 1.3 of Harris (1995).

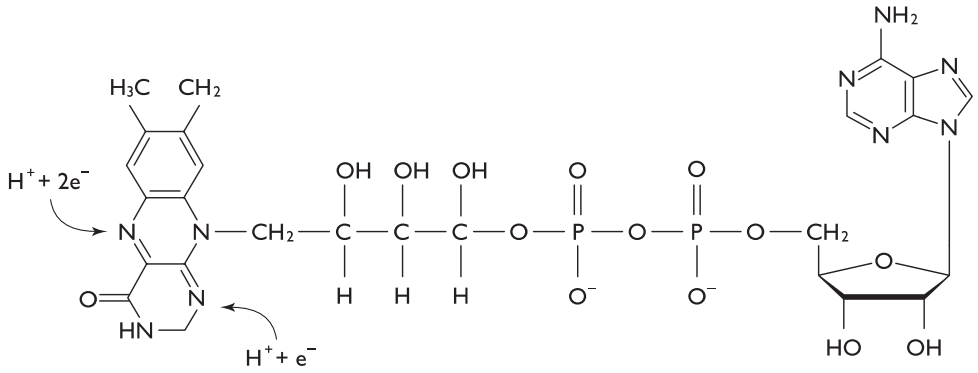
conversion of pyruvate to ethanol and CO_2. In the presence of oxygen, the three carbons of pyruvate are completely oxidized to CO_2.

In Chapter 4 we saw how an extreme value of K_{eq} corresponds to a mass action ratio that is difficult to shift by changes in the concentrations of reactants or products alone. Nevertheless, the thermodynamic unfavorability of a process can be overcome by the cell's maintaining concentrations that promote the reaction. One such reaction occurs in glycolysis. Fructose-1,6-bisphosphate (FBP) is cleaved by aldolase into two triose phosphates, dihydroxyacetone phosphate and glyceraldehyde phosphate (GAP). (Note that both trioses are phosphorylated, preventing escape from the cell!) Cleavage of the C–C bond is highly endergonic; ΔG° is large and positive. In order for the reaction to occur, $\ln([GAP]^2/[FBP])$ must be negative; the mass action ratio must be much less than 1. This step of glycolysis occurs only because the cellular concentrations of the products are kept below 1 μM; the mass action ratio is less than 1 for concentrations of FBP greater than 1 pM! There is a sense in which the cell is a sort of finely tuned machine.

The *citric acid cycle* (Fig. 5.3) is the terminal stage of the chemical processes by which the major portion of carbohydrates, fatty acids, and amino acids are converted into a form of chemical energy that is more useful to the cell. The cycle is the common mode of oxidative degradation in cells in animals, plants, microorganisms, and fungi; it is a main feature of cellular chemistry that is shared by all known forms of life. One complete cycle yields two molecules of carbon dioxide, one molecule of ATP, and numerous biosynthetic precursors. The cycle is entered twice in the oxidation of a single glucose molecule (one glucose gives two pyruvates), producing six molecules of nicotinamide adenine dinucleotide (NADH) and two molecules of flavin adenine dinucleotide ($FADH_2$) per glucose molecule by way of redox reactions (Fig. 5.5). The electron carriers NADH and $FADH_2$, which are synthesized from vitamin precursors, are of great importance to ATP production in oxidative phosphorylation (see below). The citric acid cycle was first proposed in 1937 by Sir Hans Adolf Krebs (1900–1981), a biochemist who emigrated from Germany to England in 1933. Krebs shared the 1953 Nobel Prize in Medicine or Physiology with the American Fritz Albert Lipmann (1899–1986).



(A) Nicotinamide adenine dinucleotide (oxidized).



(B) Flavin adenine dinucleotide (oxidized).

Fig. 5.5 Electron carriers in metabolism. NAD is a major soluble redox intermediate in metabolism. It is closely related to NADP, another redox intermediate. NAD and NADP differ in that the latter is phosphorylated on the adenylate ribose (R = phosphate in NADP, R = H in NAD). NADH shuttles electrons to electron transfer chains, NADPH provides electrons for biosynthesis. Neither NADH nor NADPH can form a stable one-electron intermediate, whereas FAD, a protein-bound cofactor, can form a one-electron semiquinone. Both NAD^+ and FAD comprise ADP and are synthesized from ATP (see Fig. 5.7). Energy and matter, matter and energy, energy and matter . . .

C. Oxidative phosphorylation and ATP hydrolysis

The NADH and $FADH_2$ molecules generated by the citric acid cycle play a central role in *oxidative phosphorylation*, the complex process whereby ADP and inorganic phosphate are combined to form ATP. From a quantitative point of view, oxidative phosphorylation is the most important means by which a cell generates ATP: complete metabolism of 1 mole of glucose by the citric acid cycle yields a maximum of 38 moles of ATP (2 from glycolysis, 2 from the citric acid cycle, and 34 from reoxidation of NADH and $FADH_2$). ATP is the most commonly utilized form of energy in a cell (Chapter 1).

The term *bioenergetics* usually refers to the way the in which cells generate energy from foodstuffs. The main concept of bioenergetics

is chemiosmotic theory, which states that energy stored as a proton gradient across a biological membrane (the so-called *proton motive force*) is converted to useful chemical energy in the form of ATP. One of the key contributors to the understanding of biological energy transfer has been the British biological chemist Peter Dennis Mitchell (1920–1992), who was awarded the Nobel Prize in Chemistry for his work in 1978.

The proton motive force is built up across the inner membrane of mitochondria in animals, the inner membrane of chloroplasts in plants, and the plasma membrane of aerobic bacteria (Fig. 5.6). Energy released from electron-transfer events in membrane-bound proteins is harnessed to generate the gradient. The chain of electron transfer reactions in mitrochondria terminates in the reduction of oxygen to water and the otherwise thermodynamically unfavorable pumping of protons across the membrane against the concentration gradient. The movement of protons down their gradient through the enzyme ATP synthase, the most complex structure in the inner mitochondrial membrane, results in the synthesis of ATP from ADP and inorganic phosphate. The difference in proton concentration across the membrane can be measured as a difference in pH. The role of mitochondria in coupling the phosphorylation of ADP to the electron transfer from reduced NAD to oxygen was shown by Albert Lester Lehninger (1917–1986) and associates at Johns Hopkins. Lest anyone think that the underlying electron-transfer reactions are unimportant to the end result, the mechanism underlying the toxicity of the highly poisonous cyanide ion involves binding to and inhibition of the cytochrome *a*-cytochrome a_3 complex (cytochrome oxidase) in mitochondria, and the poison sodium azide, which is added to protein solutions to inhibit the growth of bacteria, inhibits cytochrome *c* oxidase and thereby ATP synthase.

ATP, once synthesized, is put to use by the cell in many ways. For example, the free energy change of ATP hydrolysis is employed to power a tremendous variety of otherwise thermodynamically unfavorable biochemical reactions. In essence what ATP does in this context is provide free energy on the loss of its terminal phosphate group by hydrolysis of the *phosphoanhydride bond* (Fig. 5.7). Chemical coupling of ATP hydrolysis (Chapter 4) then "energizes" metabolic reactions which on their own cannot occur spontaneously. ATP is a common intermediate of energy transfer during anabolism, cellular processes by which energy is used to synthesize complex molecules from simpler ones.

In certain specialized cells or tissues, the chemical energy of ATP is used to do other kinds of chemical work, for example, the mechanical work of muscle contraction and cell movement (Chapter 8). ATP is required for osmotic work, the transport of ions other than H_3O^+ or metabolites through a membrane against a concentration gradient (below). ATP is also a major energy source in the synthesis of macromolecules from monomers, e.g. polysaccharides from individual sugar molecules and polypeptides from amino acids

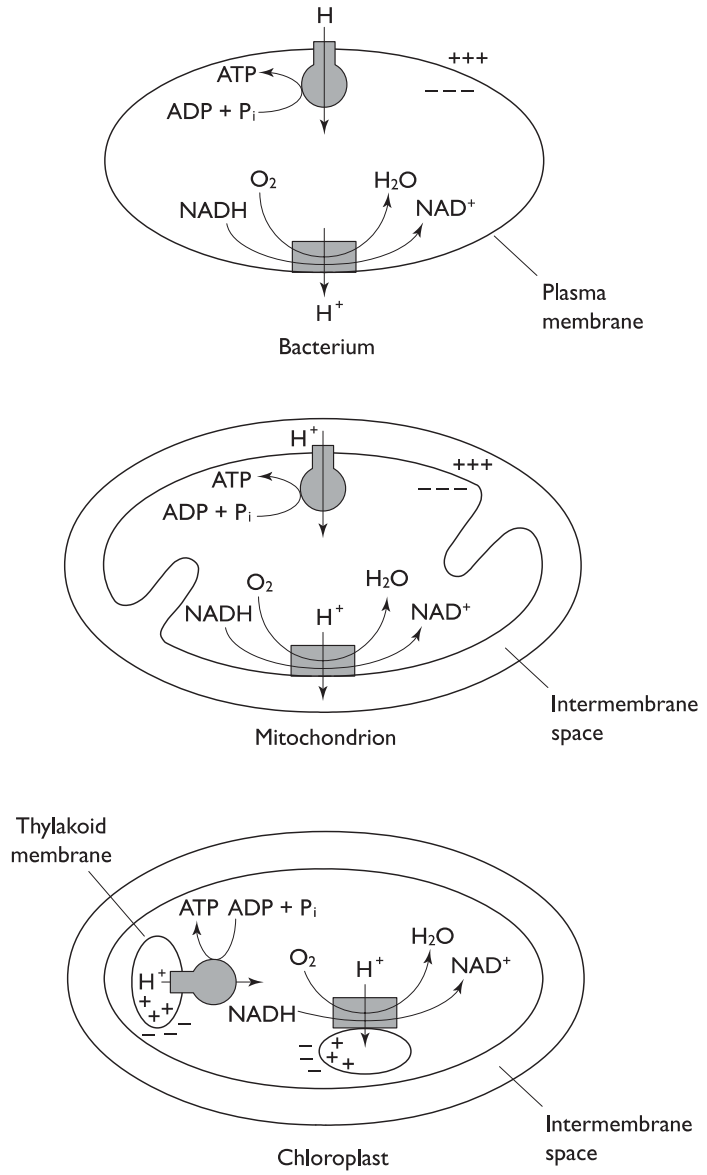


Fig. 5.6 Proton movement in bacteria, mitochondria, and chloroplasts. Note the similarities and differences in membrane orientation and direction of proton movement. In bacteria, mitochondria, and chloroplasts, the protein complex in which ATP is synthesized in situated on the cytosolic face of the membrane. Electron transport results in translocation of protons from the cytosolic side to the exoplasmic side of the membrane, creating a pH gradient. This is used to generate ATP as protons move down the pH gradient into cytoplasmic side. The similarities in ATP generation in bacteria, mitochondria, and chloroplasts point to the profound unity of all known living organisms. Adapted from Fig. 17–14 of Lodish *et al.* (1995).

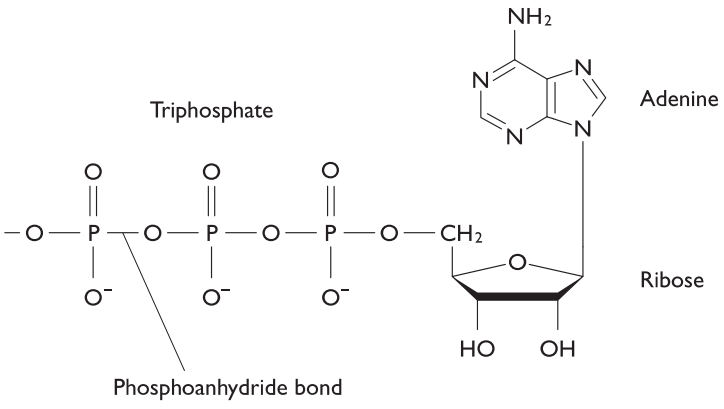


Fig. 5.7 The structure of adenosine triphosphate. There are three main components: adenine, a base found in RNA and DNA; ribose, a sugar; and triphosphate. In most biochemical reactions in which it is involved, ATP is hydrolyzed to ADP and inorganic phosphate. The bond broken in this reaction is a phosphoanhydride bond. The pK_as of the dissociable protons are different (see Chapter 4).

Table 5.1. *ATP requirements of macromolecule formation*

Macromolecule	Subunit type	ATP expenditure per monomer added (mol mol^{-1})
Polysaccharide	Sugar	2
Protein	Amino acid	4
Lipid	CH_2 unit from acetic acid	1
DNA/RNA polymerization	Nucleotide	2

(Table 5.1). In respect of all this, ATP is known as the "universal biochemical energy currency" (Chapter 1). We can see that there are many possible ways in which the free energy of a single glucose molecule can be distributed throughout a cell!

The vital importance of ATP in metabolism was first recognized by Fritz Lipmann and Herman Kalckar in 1941. Over 60 years on, the role of ATP in the cell is no less important than at any other point in the Earth's history (it seems). So we had better know something about it! The hydrolysis of ATP to ADP and P_i can be symbolized as

$$ATP + H_2O \Leftrightarrow ADP + P_i + H^+. \tag{5.3}$$

Using Eqns. (4.32) and (4.38), the free energy change for this reaction can be expressed as

$$\Delta G = \Delta G^{\circ\prime} + RT\ln[ADP][P_i]/[ATP]. \tag{5.4}$$

To keep things simple, we assume ideal behavior. Note that $[H^+]$ and $[H_2O]$, which are practically independent of the concentrations of the other species, are not included explicitly in Eqn. (5.4) (refer to the previous chapter if you are not sure why!). $\Delta G^{\circ\prime}$ for Eqn. (5.4) is about -7 kcal mol^{-1}. Does this hold for the cell, where conditions are of course very different from the standard state? Assuming that the cellular concentration of each species is 10 mM (a *very rough* estimate), Eqn. (5.4) says that $\Delta G = -7$ kcal $mol^{-1} + [1.987$ cal mol^{-1} $K^{-1} \times 298$ K $\times \ln(0.010)] = -7$ kcal $mol^{-1} - 2.7$ kcal $mol^{-1} \approx -10$ kcal $mol^{-1} \approx -42$ kJ mol^{-1}.[4] That's a 40% increase in the driving force for hydrolysis over standard state conditions! In other words, the equilibrium in Eqn. (5.3) makes a big shift towards the products when the solution is dilute. And according to the Second Law, if ATP hydrolysis releases about 10 kcal mol^{-1} at cellular concentrations, *at least* that much energy must have been consumed to synthesize ATP in the first place! Where does the energy come from?

The foregoing discussion increases our sense of awe of how the world is put together, but it also teaches some practical lessons.

[4] In skeletal muscle, [ATP] is $\sim 50 \times$[AMP] and $\sim 10 \times$[ADP]. Using these values, ΔG is even more exergonic, possibly as large as -60 kcal mol^{-1}.

Hydrolysis of ATP is clearly spontaneous in aqueous solution, and the reaction occurs relatively rapidly at 25 °C. (*In vitro*, the half-life of ATP is on the order of days at this temperature, and in the cell, where it is needed for metabolism, it is less than 1 s.) If the ratio of the *in vitro* to *in vivo* half-life were not large, ATP would be less a useful energy storage molecule than we know it to be. The hydrolysis rate of ATP and its dependence on concentration in the laboratory require that ATP-containing buffers be made up fresh and stored cold. For the same reason solutions of the free nucleotides used in the polymerase chain reaction (PCR, see below) are usually stored frozen at −20°C and thawed immediately before use.

Measurement of the enthalpy change of ATP hydrolysis shows that $\Delta H^\circ = -4$ kcal mol^{-1}. That is, hydrolysis of one mole of ATP at 25 °C results in about 4 kcal being transferred to the solution in the form of heat and about 3 kcal remaining with ADP and P_i in the form of increased random motion. We can combine our knowledge of the free energy and enthalpy changes to calculate the entropy change of ATP hydrolysis. Solving Eqn. (4.2) for ΔS when $\Delta T = 0$, we have $\Delta S^\circ = (\Delta H^\circ - \Delta G^\circ)/T$. At 310 K, $\Delta S^\circ = (-4 \text{ kcal mol}^{-1} - (-7 \text{ kcal mol}^{-1}))/(310 \text{ K}) = 10 \text{ cal mol}^{-1} \text{ K}^{-1}$. This is *roughly* the amount of entropy your body generates every time an ATP molecule is hydrolyzed. So, no matter how much you might feel your eyes glazing over at the sight of more Δs and cals, and no matter how much you might feel that biological thermodynamics is catalyzing the transition of your brain from a normal to a vegetative state, because you're simply alive you're doing a very fine job indeed of degrading the useful energy of the universe!

A couple of other pertinent points can be made here. One is that three of the four phosphate hydroxyl groups of ATP have pK_a values around 1.5. These are effectively completely ionized at neutral pH. In contrast, the fourth one has a pK_a of 6.5. This suggests that the net charge on any given ATP molecule might have a large impact on its cellular function. A second point is that the free energy difference between ATP and ADP + P_i is *not* the same as that between the plus-phosphate and minus-phosphate forms of other biomolecules. Glucose-6-phosphate, for instance, an important molecule in glycolysis, transfers its phosphate group to water with a standard state free energy change of about −3 kcal mol^{-1}. This is a substantially smaller energy change than for hydrolysis of ATP. The driving force for the chemical transfer of a phosphoryl group is known as *phosphoryl group-transfer potential*. ATP has the higher phosphoryl group-transfer potential of the two molecules. One might wonder whether ATP has the *highest* standard free energy of hydrolysis of all naturally occurring phosphates? No! ATP occupies a position about midway between extremes in tables of the standard free energy of hydrolysis of phosphate compounds (Table 5.2). ATP's being small and in the middle of the phosphate energy scale is likely an important determinant of its role in the cell.

Table 5.2. *Standard free energy changes of hydrolysis of some phosphorylated compounds*

Compound	$\Delta G^{\circ\prime}$ (kJ mol^{-1})
Glucose-1-phosphate	−20.9
Glucose-6-phosphate	−13.8
Fructose-6-phosphate	−13.8
ATP → ADP + P_i	**−30.5**
ATP → AMP + P_i	−32.5
Phosphocreatine	−43.1
Phosphoenolpyruvate	−61.9

Data are from Jencks, W. P., in Fasman, G. D. (ed.) (1976) *Handbook of Biochemistry and Molecular Biology*, 3rd edn, *Physical and Chemical Data, Vol. I*, pp. 296–304. Boca Raton: CRC Press.

Now let's look at a few other aspects of the cellular role of ATP: activity of glycogen synthase, synthesis of cyclic AMP, binding of ATP to hemoglobin, and inhibition of thermogenin in heat generation. Glycogen is a polymeric form of glucose that can be readily metabolized in times of need. Synthesis of glycogen involves the transfer of the glycosyl unit of uridine diphosphate glucose (UDPG) to an existing carbohydrate chain. UDPG is synthesized from glucose-6-phosphate and uridine triphosphate (UTP), a molecule involved in the synthesis of mRNA. Note the close "coupling" between energy storage and metabolism and information storage and expression. Marvellous efficiency! Replenishment of UTP occurs by means of a phosphoryl transfer reaction mediated by nucleotide diphosphate kinase. This enzyme catalyzes the transfer of a phosphoryl group from ATP to UDP, yielding ADP and UTP. Then replenishment of ATP occurs by means of a phosphoryl reaction mediated by ATP synthase and a proton gradient, and replenishment of the proton gradient occurs by means of oxidation of glucose . . .

ATP is a precursor in the synthesis of 3′,5′-cyclic AMP (cAMP), an important intracellular signaling molecule known as a *second messenger* (Fig. 5.8).[5] The concentration of cAMP in the cell increases or decreases in response to the tight and specific binding of an extracellular molecule to a cell-surface receptor. For instance, [cAMP] goes up when a specific odorant receptor on a cell in the olfactory epithelium binds an odorant molecule, for instance, one of the aromatic ketones or amines mentioned in Chapter 3. Binding induces a conformational change in the receptor, and an intracellular protein that interacts with the cytoplasmic part of the receptor then activates adenylyl cyclase, the membrane-bound enzyme responsible for synthesis of cAMP from ATP. Once made, cAMP then moves throughout the cytoplasm, interacting with a wide range

[5] This term was introduced in 1964 by Earl Sutherland (1915–1974), an American, the discoverer of cAMP.

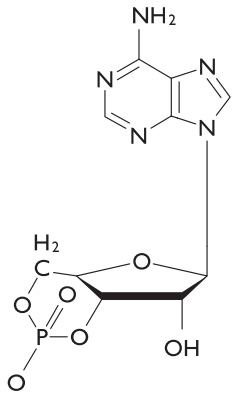


Fig. 5.8 Cyclic AMP. This molecule, which is synthesized from ATP, plays a key role in a variety of cellular processes. Principal among these is the control of glycogen metabolism in muscle. Glycogen is the highly branched high molecular mass glucose polysaccharide that higher animals synthesize to protect themselves from potential fuel shortage. The corresponding polymer in plants is starch (Fig. 1.1). Glycogen synthesis involves glycogen synthase. This enzyme catalyzes the transfer of the glucosyl unit of UDP-glucose (itself synthesized from glucose-1-phosphate and UTP, one of ATP's brother molecules) to glycogen. In glycogen breakdown, the enzyme glycogen phosphorylase cleaves the glycosidic bond linking glucose monomers by the substitution of a phosphoryl group. The products are a slightly smaller glycogen molecule and one molecule of glucose-1-phosphate (G1P), which is converted to glucose-6-phosphate by phosphoglucomutase. The nucleotides of information storage in genetic material play an important role in energy storage and utilization in all known living organisms. cAMP activates a protein kinase which activates phosphorylase kinase which, through phosphorylation, activates glycogen phosphorylase and inactivates glycogen synthase. The cellular concentration of cAMP is increased by adenylate cyclase, which is activated by the binding of glucagon or epinephrine to its receptor in the plasma membrane. When the hormone insulin binds to its receptor, glycogen phosphorylase is inactivated and glycogen synthase is activated.

of proteins. In this way, cAMP "mediates" the response of the cell to the ligand, be it an odorant molecule, hormone, or neurotransmitter. Again, there is a connection between energy and information, in that the small energy molecule ATP is involved in the communication throughout the cell of a signal received at the cell membrane. Later in this chapter we shall look at an example of the mechanics of binding interactions, and the subject will be covered in considerable detail in Chapter 7.

The physiological role of ATP does not always involve hydrolysis or chemical conversion into an electron carrier or second messenger. In fish and most amphibians, ATP binds tightly to deoxygenated hemoglobin but only weakly to oxygenated hemoglobin. The protein hemoglobin plays a crucial role in respiration by transporting oxygen to cells for oxidative phosphorylation. Binding to ATP regulates the function of hemoglobin by reducing its affinity for oxygen (see below and Chapter 7).

Box 5.1 Cool mice live longer

Obesity results when energy intake exceeds the energy expenditure. Experimental studies have shown that calorie restriction reduces core body temperature in rodents and primates. But is a lower core body temperature a simple consequence of calorie restriction, or is lower body temperature itself beneficial for health? Separate studies have found that lowering the core body temperature of poikilotherms like fish slows aging and prolongs life. But is this true of homeotherms like humans? To investigate the matter, Bruno Conti of the Scripps Research Institute in La Jolla, California and his colleagues created an engineered strain of laboratory mice. These animals have an overheated hypothalamus, the preoptic area of which is the brain's central thermostat. Heating up the hypothalamus dispatches "Chill out!" signals to the rest of the body and thus decreases the core temperature. Hypothalamic overheating in the engineered mice was achieved by over-expressing uncoupling protein 2 (UCP2). Found in the inner membrane of mitochondria, the powerhouses of the cell, UCP2 uncouples electron transport from ATP production and thus dissipates as heat the energy stored in the proton gradient across the mitochondrial membrane. UCP2 over-expression dissipated more proton gradient energy as heat than in normal mice, elevated the

Box 5.1. Cont.

temperature of the hypothalamus, and dropped the core body temperature 0.3–0.5 °C. Experimental female mice lived an average of 20% longer than controls; experimental males, 12% longer. Sex hormones might influence the rate at which some mice attain a lower core temperature. The experimental animals also appeared to show greater metabolic efficiency than controls, suggesting that fewer calories were needed to live. A take-home message for all you Ponce de Leon wannabes out there: the fountain of youth is within you.

Above we saw how the proton gradient in mitochondria can be coupled to the membrane protein ATP synthase and used to synthesize ATP. In brown adipose tissue, which contains large amounts of triacylglycerols (fatty acid triesters of glycerol, or fats, see Table 1.2) and many mitochondria, the proton gradient can be uncoupled from ATP synthesis by means of a channel protein called thermogenin. Dissipation of the proton gradient in the absence of ATP generation means that brown adipose tissue acts as a "built-in heating pad." Thermogenin is particularly plentiful in cold-adapted animals. The activity of thermogenin is under hormonal control. The adrenal hormone norepinephrine binds to its receptor and activates adenylate cyclase, which makes cAMP, which activates a kinase that phosphorylates a lipase, which hydrolyzes triacylglycerols to free fatty acids. When the concentration of free fatty acids is sufficiently high, thermogenesis is activated, and thermogenin changes the permeability of the inner mitochondrial membrane to protons and allows them back into the mitochondrial matrix without ATP production. Proton flow under the control of thermogenin is inhibited by ATP, GTP, and the diphosphate forms of these nucleotides.

D. Substrate cycling

The reaction catalyzed by the glycolytic enzyme phosphofructokinase is highly exergonic. Under physiological conditions,

$$\text{fructose-6-phosphate} + \text{ATP} \rightarrow \text{fructose-1,6-bisphosphate} + \text{ADP}, \quad (5.5)$$

with $\Delta G = -25.9$ kJ mol^{-1}. This reaction is so favorable that it is essentially irreversible. But the reverse reaction can occur! It just won't do so on its own. In fact, the enzyme fructose-1,6-bisphosphatase is present in many mammalian tissues, and it catalyzes the removal of a phosphate group from fructose-1,6-bisphosphate as follows:

$$\text{fructose-6-bisphosphate} + \text{H}_2\text{O} \rightarrow \text{fructose-6-phosphate} + \text{P}_\text{i}. \quad (5.6)$$

This also occurs spontaneously because $\Delta G = -8.6$ kJ mol^{-1}, but the reverse reaction is more probable than in Eqn. (5.5). The net reaction

is simply ATP hydrolysis, and $\Delta G = -34.5$ kJ mol^{-1}. Note that, although the overall free energy change is negative, this coupled reaction is less favorable than transfer of the terminal phosphoryl group of ATP to water. The opposing reactions of Eqns. (5.5) and (5.6) are called a *substrate cycle*.

Substrate cycles might seem to serve no useful purpose, since all they do is consume energy. But nature is a subtle lover, and she is more apt to reveal her charms to persistent humility than audacious presumption. The reverse reaction, far from being futile, constitutes a means of regulating the generation of product by the forward reaction, because enzyme activity itself is regulated. In cases where a substrate cycle is operative, metabolic flux is not simply a matter of the activity of an enzyme, but the combined activity of the enzymes working in opposite directions. There is thus exquisite regulation of a metabolic pathway, adjusting as needed to the cell's metabolic needs. The price paid for such control is the energy lost in the combined forward and reverse reactions.

Substrate cycles also function to produce heat, helping to maintain an organism's temperature.[6] So, although there is a high cost to control, it is clear that organisms make highly efficient use of the resource. It is also possible that heat production is important for controlling the rate of enzyme activity by controlling the temperature of the enzymatic reaction (Chapter 8). In bumblebees,[7] the presence of fructose-1,6-bisphosphatase in flight muscle is thought to enable these insects to fly at temperatures as low as 10 °C: honeybees, which do not have fructose-1,6-bisphosphatase, cannot fly when it's cold. Substrate cycling probably plays a key role in maintaining body heat in many animals, including humans. It is stimulated by thyroid hormones, which are activated upon exposure of the organism to cold. It's time for a change of topic.

E. Osmosis

We covered the concept of chemical potential in Chapter 4. Let's use it to develop a topic of general importance in biochemistry: osmosis (Greek, push). When mineral ions and charged molecules are absorbed by the small intestine, water follows *by osmosis*. We treat this subject in a fair amount depth for two reasons: *osmotic work* underlies many physiological functions – nerve conduction, secretion of hydrochloric acid in the stomach, and removal of water from the kidneys – and the subject involves a number of key subtleties of thermodynamics. Before looking at the mathematics of osmosis, let's first think about the situation qualitatively. This way, we'll be more certain to have a general sense of the matter before facing a

[6] Some organisms obtain a significant amount of heat from their surroundings, for instance basking lizards and snakes, which absorb heat from the Sun.

[7] Previously called humble-bees in the UK.

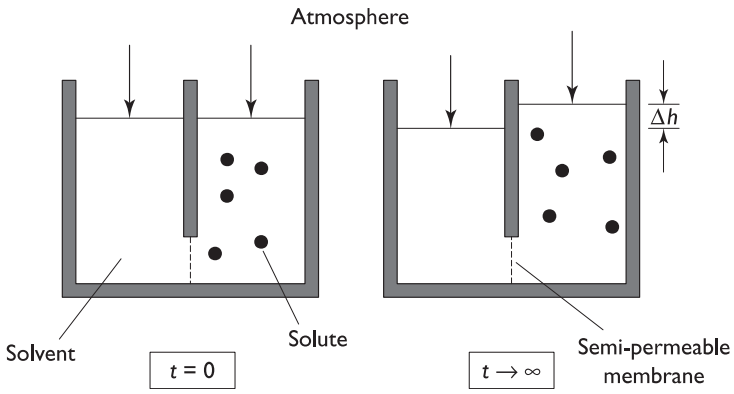


Fig. 5.9 A simple osmometer. A solute can move freely in a fraction of the total volume of solvent. The solution is separated from pure solvent by a membrane that is permeable to the solvent but not the solute. There is a net flow of solvent from the pure solvent to the solution, resulting in the development of a head of pressure. This pressure is the osmotic pressure, $\pi = \rho g \Delta h$, where ρ is density of the solvent, g is gravitational acceleration, and Δh is the difference in fluid levels. As described by van't Hoff, $\pi = CV_o RT/m$, where C is the mass of solute in the volume of solvent, V_o is the partial molar volume of the solvent, and m is the molecular mass of the membrane-impermeant solute. Note that π is an approximately linear function of C under some conditions. Osmotic pressure data can thus be used to measure the molecular mass of an osmotic particle.

page filled with equations. That *osmosis* is a pleasant-sounding word might help to move the discussion along ...

Osmosis is an equilibrium phenomenon that involves a semi-permeable membrane (not necessarily a biological membrane). *Semi-permeable* in this context means that there are pores in the membrane that allow small molecules like solvents, salts, and metabolites to pass through but prevent the passage of macromolecules like DNA, polysaccharides, and proteins. Biological membranes are semi-permeable: large solute molecules are *impermeant*. Like freezing point depression and boiling point elevation, osmosis is a colligative property.

Suppose we have an osmometer, also called a U-tube, with arms separated by a semi-permeable membrane (Fig. 5.9). Let the temperature be constant. If no solute is present the height of the solvent is the same on both sides, because the pressure of the external environment is the same on both sides. The situation changes on introduction of an impermeant solute to one side. Let the solute be a largish protein, say hemoglobin, and let it be freeze-dried before being added to the solvent. Freeze-dried protein occupies a relatively small volume. Initially, the height of the fluid is the same on both sides of the osmometer, just as when no solute was present. But whereas before the solute occupied a small volume on the bench-top, now it is able to move freely throughout one side of the osmometer. There has been a large increase in the entropy of the solute! (If you are not sure why, see the discussion on perfume in Chapter 3.) We require that the solute particles be free to roam about the entire volume on their side of the membrane, but that they not be able pass through the membrane. And just as a confined gas pushes against the walls of its container (Chapter 2), the solution pushes against the atmosphere and against the walls of the osmometer. What happens? There is a net transfer of solvent from the side where no solute is present to the other side. This decreases the volume of pure solvent and increases the volume of solution. How can we explain what has happened?

Addition of solute to solvent reduces the chemical potential of the solvent (Chapter 4). This creates a difference in the chemical potential of the solvent between the pure side and the impure side. The difference in chemical potential is thermodynamically

unstable; change must occur. The impure side has a lower solvent chemical potential, so water moves down its concentration gradient until equilibrium is reached. From an entropic point of view, the flow of water into the side with solute increases the entropy of the solute, making the situation more stable. How is entropy increased? Neglecting interactions between the solute and solvent, the greater the volume of solvent present, the larger the volume in which the solute can distribute itself. There is a resemblance to the ideal gas expansion discussed in Chapter 3 (see Eqn. (3.23)). In the context of the perfume example, if it is applied in the bathroom, the perfume molecules become distributed throughout the bathroom, but when the door is opened, the molecules begin to spread into the corridor. At equilibrium, the perfume molecules will occupy the bathroom and corridor, i.e. the entire accessible volume. The concentration is reduced in the process, entropy is increased, and more work would have to be done to gather all the molecules back together into the same place.

The flow of water from one side of the U-tube to the other must result in a change in the height of the water on the two sides. It becomes lower on the side of the pure solvent and higher on the side of the impure solvent. After enough time, the system comes to equilibrium, and the driving force for water to move through the membrane from the pure solvent to the solution will be equal in magnitude to the hydrostatic pressure arising from the difference in height of the water in the two arms ($p_{\text{hydrostatic}} = \rho g \Delta h$, where ρ is the density of the solution). The hydrostatic pressure is the same as the *additional* pressure one would have to apply to the side of the U-tube with solute in order to equalize the height on the two sides of the membrane. This pressure is called the *osmotic pressure*, and it was first studied in the 1870s by the German botanist and chemist Wilhelm Friedrich Philipp Pfeffer (1845–1920), the son of an apothecary.

Now let's take a more mathematical approach to osmosis. This way of thinking about the subject is not necessarily superior to the qualitative approach just because it involves more equations, but it will provide additional insight to our subject, and that's what we want. Mathematical or computational modeling of the physical world makes sense as long as it leads to insights that can be tested experimentally. In our approach to modeling, the system is regarded as consisting of two *phases*, x and y. In x, the impermeant molecules (component 2) are dissolved in the solvent (component 1). In y, only solvent molecules are present. Considering the *solvent* alone, the requirement for equilibrium between the two phases is

$$\Delta G = \mu_1^x \Delta n_1^x + \mu_1^y \Delta n_1^y = 0, \tag{5.7}$$

where Δn_1 stands for an incremental change in the number of moles of solvent. (See Eqn. (4.5).) Because $\Delta n_1^x = -\Delta n_1^y$ (because the gain of solvent molecules in one phase must come at the expense of the same number of molecules from the other phase),

$$\mu_1^x = \mu_1^y. \tag{5.8}$$

The ledger balances. But wait! Something funny's going on. For regardless of the amount of solvent transferred through the membrane, we can't avoid the requirement that $\mu_1 - \mu_1^o = \Delta\mu < 0$ (see Eqn. (4.10)). That is, the chemical potential of the solvent plus solute *must* be lower than that of the pure solvent. Nevertheless, Eqn. (5.8) does say that the chemical potentials of the solvent in the two phases must be equivalent. Where did we err?

We didn't! We conclude that there is a contradiction, that biological thermodynamics is illogical and therefore a waste of time, that it was a mistake to study biology or in any case to do it in a place where biological thermodynamics forms part of the curriculum, and that our best option would be to make our way to the college bar and drink away our misery. Right? No way! Things are just starting to get interesting! Let's see if we can't crack this nut now, and think about a celebratory night out later on. But what can we do?

Above we showed that $\Delta G = V\Delta p - S\Delta T$. Under isothermal conditions, $\Delta T = 0$ and the free energy change is proportional to Δp. To make the resulting expression tell us what will happen when the number of solvent molecules is changed, we divide both sides by Δn_1. This gives

$$\Delta G/\Delta n_1 = \mu_1 = V_{m,1}\Delta p, \tag{5.9}$$

where $V_{m,1}$ is the molar volume of component 1. We have found the "missing" term from our expression of the chemical potential earlier in this section! Taking into account the development leading up to Eqn. (4.12), where we saw how the chemical potential of a solvent changes when a solute is added, and adding in Eqn. (5.9), we have

$$\mu_1 - \mu_1^\circ \approx -RTC_2V_1^\circ/M_2 + RT\ln f_1 + V_1^\circ\pi, \tag{5.10}$$

where the pressure difference has been symbolized as π (this has nothing to do with the ratio of the circumference of a circle to its diameter; it is the symbol that is traditionally used to designate the osmotic pressure; π starts with the same sound as *pressure*). Adding in the extra term (Eqn. (5.9)) might strike you as a rather arbitrary way of doing math – one that would be fine for getting a correct result on an exam but maybe dubiously valid. But we need to remember that in order for the equilibrium condition to be met, we must have a balance of forces, and we can write down an equation – an expression of balance – only if we take into account *everything* that's relevant. There was indeed a contradiction earlier because we had assumed that the system was at equilibrium when in fact we hadn't taken the pressure term into account. Note that in Eqn. (5.10) we have assumed $V_{m,1} \approx V_1^\circ$, the molar volume of pure solvent, which is valid for dilute solutions. The pressure difference π is the pressure that must be applied to the solute side of the U-tube to make the fluid height the same on both sides.

We can simplify Eqn. (5.10) a bit. If the solution is ideal, $f_1 \approx 1$ and $RT\ \ln f_1 \approx 0$. At equilibrium, $\mu_1 - \mu_1^o = 0$. It follows that

$$\pi = RTC_2/M_2. \tag{5.11}$$

This is the van't Hoff law of osmotic pressure for ideal dilute solutions, named in honor of the scientist who gave Pfeffer's work a mathematical foundation.[8] Equation (5.11) can be used to measure the mass of an impermeant solute particle (though there are easier and more accurate ways to do it). Note how Eqn. (5.11) looks like Eqn. (4.12). You may already have noticed how closely Eqn. (5.11) resembles the ideal gas law ($pV = nRT$ or $p = nRT/V = CRT$, where n is number of particles and C is concentration). C_2, the concentration of solute, is the mass of solute particles added to a known volume of pure solvent. What van't Hoff found was that the measured osmotic pressure was basically the pressure of n solute particles moving around in volume V, the volume of the solvent through which the solute particles are free to move!

The degree to which Eqn. (5.11) matches experimental results varies with concentration and solute (Fig. 5.10). There are several different ways of trying to cope with the situation, but our concern will be with just one of them here. Time is spent on it at all because it's a generally useful method. We express the thermodynamic observable quantity (here, π) as a series of increasing powers of an independent variable, (here, C) and check that the dominant term is the same as we found before (Eqn. (5.10)) when the independent variable takes on an extreme value (low concentration limit, as we assumed above):

$$\pi = \frac{C_2 RT}{M_2}(1 + B_1(T)C_2 + B_2(T)C_2^2 + \ldots). \tag{5.12}$$

The $B_i(T)$ terms are *constant* coefficients whose values are solute- and temperature-dependent and must be determined *empirically*. If C_2 is small, only the first term makes a significant contribution to π (convince yourself of this!), just as in Eqn. (5.10). If only the first two

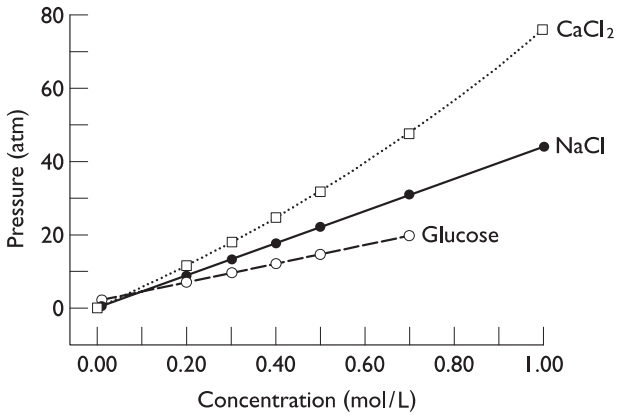


Fig. 5.10 Osmotic pressure measurements. Osmotic pressure increases with concentration of solute, as predicted by the van't Hoff law. The pressure at a given concentration of solute depends significantly on the solute. If the solute is a salt, dissociation in aqueous solution will result in a greater number of particles than calculated from the molecular mass of the salt. The van't Hoff law is exact for an ideal solution. At high solute concentrations, non-linear behavior can be detected. Such behavior can be accounted for by higher order terms in C. The data are from Table 6–5 of Peusner (1974).

[8] The Dutch physical chemist Jacobus Henricus van't Hoff (1852–1911) was the recipient of the Nobel Prize in Chemistry in 1901, the first year in which the prestigious awards were made.

terms make a significant contribution to π, a plot of π/C_2 will be linear in C_2 with slope $B_1(T)RT/M_2$ and intercept RT/M_2. This permits indirect measurement of M_2 and $B_1(T)$. Equation (5.12) can readily be generalized to include contributions from different species of osmotic particle:

$$\pi_{\text{total}} = \pi_1 + \pi_2 + \cdots + \pi_n = \Sigma\pi_i. \tag{5.13}$$

If a solute species is present on both sides of the membrane, and if this solute cannot pass through the membrane, it will make a contribution to the total osmotic pressure, but only if there is a concentration difference. In such cases, π_i is proportional not to C_i, as in Eqn. (5.12), but to ΔC_i, the concentration difference across the membrane.

Now let's leave the wispy world of mathematics and migrate over to a more material physical biochemistry. Osmosis can be a very strong effect. At 25°C, a 1 M solution of glucose, a relatively small "osmolyte," gives a pressure more than 25 times greater than that of the atmosphere; 1 M solutions of salts give even larger osmotic pressures (see Fig. 5.10), though the ions are smaller than glucose, even when hydration is taken into account. Osmotic forces are important in biology because they play a key role in membrane transport of water, in all kinds of situations. For example, red blood cells are full of impermeant solute particles, mainly hemoglobin; red cells have been called "bags of hemoglobin." M_2 is large, about 68 000, and C_2 is high, about 0.3 M or higher. When placed in pure water, there is initially a very large π across the membrane – about 8 atm or greater, the approximate pressure when scuba diving at a depth of 70 m! The plasma membrane cannot withstand an osmotic pressure of this magnitude and breaks, spewing its hemoglobin into the surrounding medium. Blood banks limit damage to red cells after separating them from plasma by centrifugation by resuspending the cells in a sucrose solution (sucrose is membrane impermeant) of approximately the same solute concentration as blood plasma (an *isotonic* solution).

Note the difference between the red blood cell in a solution of low solute concentration (hypotonic solution) and impermeant particles in an osmometer. In the osmometer, there is a real pressure difference. The presence of the impermeant particles results in the formation of a pressure head, and the solution pushes down harder than it would if the solute were not there. But what if the osmometer were configured as a capillary tube oriented horizontally with a semi-permeable membrane in the middle? Water would move through the membrane as before, and one could think of the effect as arising from the pressure of n solute particles confined to a volume V of solvent. The thought experiment suggests that red blood cells burst in hypotonic solution because the hemoglobin molecules inside the cell bang on the cell membrane much harder than the water molecules bang on the membrane from the outside. Is that right?

If a solution is at thermal equilibrium, then *all* the particles in the system have the same average thermal energy, irrespective of size. Big molecules like hemoglobin are relatively slow, little molecules like water are relatively fast (Chapter 1). But these molecules do not have the same *momentum* (Chapter 2). From physics, the K.E. of a particle is $\frac{1}{2}mv^2 = \mathbf{P}^2/m$, where m is the mass, v is the velocity and $\mathbf{P} = mv$ is the momentum. Thermal energy is proportional to T, and at thermal equilibrium the K.E. of a particle is equal to its thermal energy. So, $\mathbf{P}$ $(mT)^{\frac{1}{2}}$. In other words, the more massive the particle, the greater its momentum. Note that $\mathbf{P}$ has nothing to do with particle *volume*. So in a hypotonic solution, where there is nothing outside the cell that cannot get inside, hemoglobin molecules do bang into the membrane from all directions a lot harder than water molecules bang on the membrane from all directions. And the concentration of water molecules on the inside certainly is lower than the concentration on the outside, but the concentration of hemoglobin outside is 0. Therefore, it might be supposed, the banging of hemoglobin molecules on the cell membrane causes the cell to burst. But wait, is not water also rushing into the cell? Water is indeed pouring in, at a very high rate, seeking to destroy the gradient. The driving force for this flood, moreover, is very large. It is also driven almost entirely by an increase in entropy. Water could enter the cell before the surroundings become hypotonic, but the rate in and rate out were the same. Now, with a net flow of water inwards, the red blood cell swells, creating a larger and larger volume in which the massive hemoglobin molecules can diffuse, allowing them to display their Second Law tendency to disperse. It has been said that thermodynamics is a difficult subject because there are so many different ways of thinking about the same thing. Osmosis is a good example of the truth of that statement.

The actual situation with hemoglobin is more complicated than we've made it sound. This is because hemoglobin does not just float around in a sea of water in a red blood cell; the protein interacts with the solvent. When particle interactions are taken into account, the volume of the particle does matter, because the more space it takes up, the more surface it will expose to the solvent. This is the source of the higher-order terms in Eqn. (5.12).

In contrast to red cells, some bacteria do not burst when placed in a hypotonic solution. This is because these organisms (as well as plant cells and fungi) can withstand high osmotic pressure gradients by means of a rigid cell wall. When certain bacteria come into contact with lysozyme, an enzyme we have encountered already several times in this book, however, the bugs can be made to spill their guts. The biochemical activity of lysozyme is to cleave certain glycosidic bonds in the polysaccharides that give the bacterial cell wall its strength, weakening the cell wall. It is a rather good thing that our bodies station lysozyme molecules at common points of entry of foreign microbes, for example, the mucosal membrane in

the nasal passage. People with a mutant lysozyme gene have a tougher time than most in fighting off infection, and they tend to die relatively young. One might surmise that early death results from too little lysozyme being available to make the cell walls of bacteria more susceptible to osmotic stress. But the actual situation is more complex than that. That's because the immune system has a role to play in fighting off infection, few pathogenic bacteria are susceptible to lysozyme alone, and the mutant lysozyme proteins, which are less active than the wild-type enzyme, are also less thermostable than the wild-type enzyme and give rise to amyloid fibril formation. Not only is the amyloidogenic lysozyme less active and therefore less able to fight off infection, there is a net incorporation of the protein molecules into rigid fibril structures where they have effectively no enzymatic activity all. To make things worse, the body has a hard time ridding itself of the fibrils, and their continued increase in size can be pathological (Chapter 8).

F. Dialysis

This section is a close relative of the previous one. There are two basic forms of dialysis in biochemistry: non-equilibrium dialysis and equilibrium dialysis. We look at both here; the physics is basically the same in both cases. Dialysis is useful to the biochemist because it can be used to separate molecules according to size. It does this by means of a semi-permeable membrane, like the membrane in the section on osmosis (Fig. 5.11). Many semi-permeable membranes used for dialysis are made of cellophane (cellulose acetate).

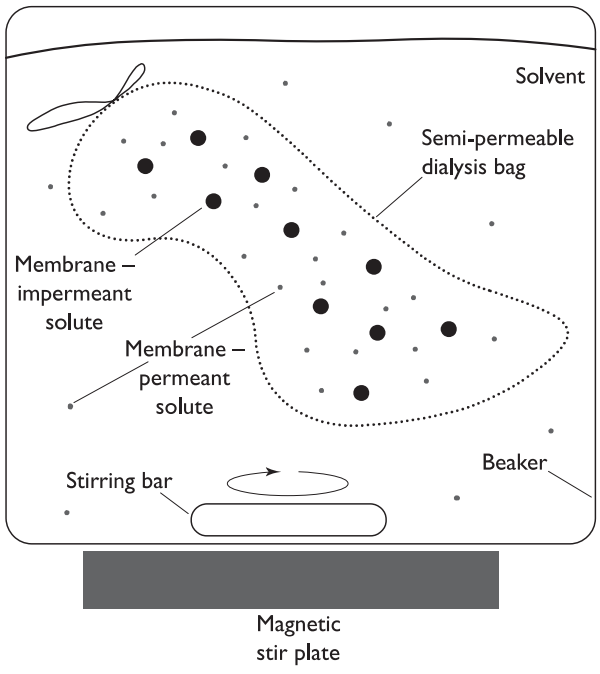


Fig. 5.11 Dialysis. A dialysis bag containing a membrane-impermeant solute is submerged in solvent in a beaker. Membrane-permeant solute appears on both sides of the membrane. The dialysis bag-solvent system is not at equilibrium. At equilibrium, the concentration of membrane-permeant solute will be the same on both sides of the membrane. A magnetic stir plate and stirring bar are used to accelerate the approach to equilibrium; the flow rate of membrane-permeant solute out of the dialysis bag is related to the concentration gradient of that solute.

Non-equilibrium dialysis is the use of a semi-permeable membrane to change the composition of the solution in which macromolecules are dissolved. For instance, one means of purifying recombinant proteins from *E. coli* host cells is to lyse the cells in 8 M urea, a small organic compound. Urea at a concentration of 8 M or higher will denature most proteins at room temperature. Once the recombinant proteins have been separated from the bacterial ones (often by a type of affinity chromatography), the recombinant protein solution is transferred to a dialysis bag. Then the bag is sealed and placed in a large volume of buffer containing no urea. When equilibrium is reached, several hours later, the concentration of urea inside the dialysis bag has decreased and the concentration outside has increased, and the concentration of urea is about the same on both sides of the membrane. What drives the change?

Initially, the concentration of solute particles is much higher within the dialysis bag than in the solution outside; some osmosis occurs. The chemical potential of urea is very much higher in the bag and very low outside, at least initially; there will be a net migration of urea out of the bag until the concentration gradient has vanished. The continuous change in concentration of urea both inside the bag and outside until equilibrium is reached and gives rise to a continuous change in osmotic pressure. The osmotic effect can be substantial when working with a high urea concentration, leading to a substantial increase in the volume of material inside the bag during dialysis. So, to avoid possible rupture of tubing and loss of a precious sample, leave room in the bag for an influx of water!

Is any work done in the net migration of urea out of the bag? No! Despite similarities to osmosis, the situation here is qualitatively different. True, urea moves down its concentration gradient, but there is no corresponding development of a pressure head as in osmosis. In other words, nothing retards the dilution of urea, just like nothing opposes the expansion of gas into vacuum, so no pV-work is done. The experiment can be carried out in isolation, so $q = 0$. Then by the First Law, $\Delta U = 0$. If the pressure is constant, then $\Delta H = 0$. And if the temperature is constant as well, G is a thermodynamic potential function. But the process we've outlined is irreversible! If we carry out the process very slowly, though, having the system pass through a series of near equilibrium states, we can evaluate ΔG. Recall that G is a state function, so its value depends only on the initial and final states of the system, and not on whether the process was reversible or irreversible. The expansion of the urea, just like the expansion of gas into vacuum, is spontaneous, so $\Delta G < 0$. This can only be true if $\Delta S > 0$, because $\Delta H = 0$. No violation the Second Law, no problems.

Another concern of non-equilibrium dialysis is charge. Typically, the macromolecule in the dialysis bag will be ionized, and this will affect osmosis by interactions with water. The solution being dialyzed out of the bag or into it will usually be a buffer of some sort, containing both charged and uncharged solute particles, and the

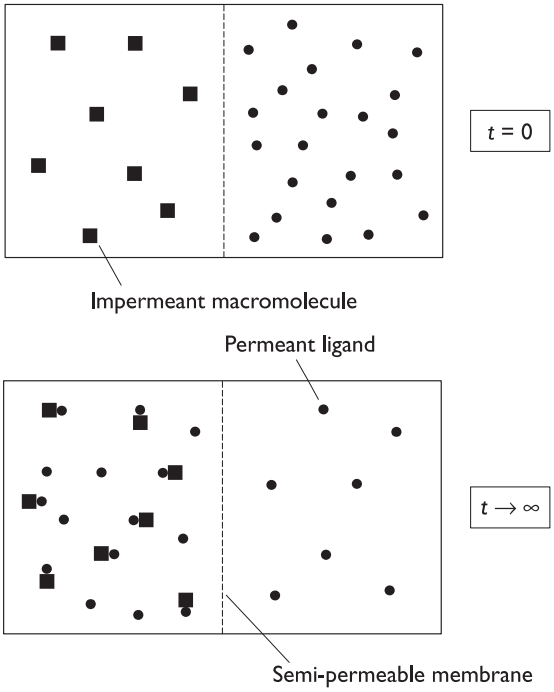


Fig. 5.12 Equilibrium dialysis. At the beginning of the experiment ($t = 0$), the membrane-impermeant macromolecule and membrane-permeant ligand are on opposite sides of a semi-permeable dialysis membrane. The two-chambered system is not at equilibrium. After a long time ($t \rightarrow \infty$), the concentration of *free* ligand is approximately the same on both sides of the membrane, in accordance with the Second Law of Thermodynamics. The number of ligand molecules is not the same on both sides of the membrane, however, as some ligands are bound to the membrane-impermeant macromolecules. The bound ligand molecules are nevertheless in equilibrium with the free ones. Measurement of the concentration of free ligand at equilibrium and the total concentration of ligand determines the amount of bound ligand at equilibrium.

ratio and relative abundance of these will have an impact on the migration of water through the membrane.

And equilibrium dialysis? In some respects it's rather similar to non-equilibrium dialysis. In others, it has a more specific meaning than *dialysis* and therefore deserves to be treated somewhat separately. Suppose you are interested in the binding of a macromolecule to a membrane-permeant ligand. This presents an opportunity for quantitative analysis of the binding interaction. To see how, suppose we have a two-chambered device like that shown in Fig. 5.12. In the left side, you introduce a known amount of macromolecule in your favorite buffer, and on the right side, a known amount of ligand dissolved in the same buffer. The ligand will diffuse in solution, and the net effect will be movement down its concentration gradient, through the membrane. By mass action the ligand will bind to the macromolecule. After a sufficiently long time, the two chambers will be at equilibrium; the concentration of free ligand will be the same on both sides of the membrane. The amount of ligand on the side of the macromolecule, however, will be higher by an amount depending on the strength of interaction between macromolecule and ligand. You can then use a suitable assay to measure the amount of ligand on both sides of the membrane, and the difference will be the amount bound to the macromolecule. You then compare the concentration of "bound" ligand to the concentration of macromolecule and to the concentration of "free" ligand, and use the results to calculate the binding constant and the number of ligand molecules bound per macromolecule. This is an important topic. See Chapter 7.

G. Donnan equilibrium

In our discussion of dialysis we barely mentioned charge effects. Here, we'll see just how much more complicated things are when charge is taken into account more formally. We need to engage with the subject with this added degree of complexity, because all major classes of biological macromolecule – proteins, nucleic acids, and some polysaccharides – are charged. Moreover, in the living organism these molecules are found not in pure water but in a saline solution.

Suppose we have a polyelectrolyte like DNA, and let it be dissolved in a solution containing a simple salt, say NaCl. Suppose further that there are two phases to our system, just as in our discussion of osmosis. Now, though, one phase consists of water, Na^+ and Cl^- (phase α), and the other consists of water, Na^+, Cl^- and DNA (phase β). The phases are separated by a semi-permeable membrane, and DNA alone is impermeant. At equilibrium, the concentration of ions will not be the same on the two sides of the membrane except in the limit that $[DNA] \to 0$. Why not? Because DNA is anionic, so we should expect the concentration of sodium to be higher on the side of the membrane with DNA than on the other side. In symbols, $[Na^{+\beta}] > [Na^{+\alpha}]$. Let's see if we can obtain a quantitative expression for the concentration of ions.

At equilibrium, even though the concentrations aren't equal, we must have

$$\mu^{\alpha}_{NaCl} = \mu^{\beta}_{NaCl}. \tag{5.14}$$

Let's keep things simple and assume that the solution is ideal. The chemical potential of the salt is

$$\mu_{NaCl} = \mu^{\circ}_{NaCl} + RT \ln[Na^+][Cl^-]. \tag{5.15}$$

At equilibrium, the standard state chemical potential must be the same in both phases, so

$$[Na^{+\alpha}][Cl^{-\alpha}] = [Na^{+\beta}][Cl^{-\beta}]. \tag{5.16}$$

And the net charge of each phase must be equal to zero, a condition known as *electroneutrality*, which is expressed mathematically as

$$[Cl^{-\alpha}] = [Na^{+\alpha}] \tag{5.17}$$

$$z[DNA^{\beta}] + [Cl^{-\beta}] = [Na^{+\beta}], \tag{5.18}$$

where z is the number of negative charges on the DNA. With a bit of algebra, these equations can be combined to give

$$[Na^{+\beta}] = [Na^{+\alpha}]\left(1 + \frac{z[DNA^{\beta}]}{[Cl^{-\beta}]}\right)^{1/2} \tag{5.19}$$

$$[Cl^{-\beta}] = [Cl^{-\alpha}]\left(1 - \frac{z[DNA^{\beta}]}{[Na^{+\beta}]}\right)^{1/2}. \tag{5.20}$$

As expected, $[Na^{+\beta}] > [Na^{+\alpha}]$, neutralizing the charge on DNA in phase β. Similarly, $[Cl^{-\beta}] > [Cl^{-\alpha}]$, though because of the minus sign in Eqn. (5.20) the difference between phases is not as great as for the DNA counterions (Na). In such situations, the observed osmotic pressure gradient is produced by both the impermeant macromolecule and the asymmetric distribution of small ions. This effect, called the *Donnan equilibrium*, was first described in 1911 by the physical chemist Frederick George Donnan (1870–1956), son of a Belfast merchant. The effect pertains not only to membrane equilibria but to any situation in which there is a tendency to produce a separation of ionic species. The asymmetric distribution of ions arises from the requirement of electroneutrality, and its magnitude decreases with increasing salt concentration and decreasing macromolecule concentration, as can be seen from Eqns. (5.19) and (5.20).

The Donnan effect is even more complicated for proteins than DNA. This is because the net charge on a protein, a sort of weak polyion, is highly dependent on pH, whereas DNA, a sort of strong polyion, has a net charge that varies relatively little with pH. The greater the net charge on a macromolecule, the greater the Donnan effect. For proteins, the Donnan effect is minimized at the isoelectric point, where the net charge on the molecule is zero. There are no conceptual difficulties here, but you might find it tricky to work with proteins at their isoelectric point in the laboratory. Protein solubility tends to be very low at the isoelectric point! A physiological example of where the Donnan effect is relevant is in the red blood cell (RBC). The effect is caused mainly by the huge concentration of hemoglobin inside the cell and the inability of hemoglobin to penetrate the membrane under isotonic conditions. Other ions present, for instance sodium and potassium, do not contribute to the Donnan effect because they are generally impermeant and their effects counterbalance ($[K^+]_{plasma} \approx [Na^+]_{cell}$ and $([K^+]_{cell} \approx [Na^+]_{plasma}$). Chloride, bicarbonate, and hydroxyl ions, by contrast, can cross the membrane, and they contribute to the Donnan equilibrium. Experimental studies have shown that the cell-plasma ratios of these ions are 0.60, 0.685, and 0.63, respectively. The marked deviations from 0.5 arise from the confinement of hemoglobin within the cell. This has an impact on the pH of the blood since both bicarbonate and hydroxyl are bases.

H. Membrane transport

There is metabolic activity within cells, and an individual cell in a higher eukaryote is separated from its surroundings by its plasma membrane. The membrane enclosing the cell is about 10 nm thick. It comprises two layers of phospholipids, with the charged groups on the outside. The interior of a membrane is “oily” and thus generally impermeable to ions and polar compounds. Some charged

substances can pass through membranes, but most only by means of transport proteins embedded in the lipid bilayer.

Membrane transport is said to be *passive* if a solute moves down its concentration gradient, and *active* if it moves against it. An example of active transport in cells is the movement of Na^+ and K^+ across the cell membrane of red blood cells, nerves and muscle cells – against the concentration gradient. The concentration of K^+ in muscle is about 124 mM, some 60-fold greater than in serum. With Na^+ it's the other way around, the concentration being about 4 mM in muscle cells and 140 mM in serum. These ions will of course tend to move down their concentration gradients to minimize free energy. But the gradients, which are important to cell function, are maintained by a membrane-spanning enzyme called Na^+/K^+-transporting adenosine triphosphatase. In order to maintain the gradient, the cell must pay a cost – that of kicking out the unwanted ions that have come in and that of recovering the wanted ones that have left. Gradient maintenance requires moving ions from a region of low concentration to a region of high concentration and therefore the expenditure of energy. The Na^+/K^+-transporter acts as an ion pump and is powered by ATP hydrolysis. Another example of active transport is the secretion of HCl into the gut of mammals by parietal cells – the home of many mitochondria!

A numerical example will help to motivate the discussion that follows. The change in chemical potential of glucose when it is transported down a 1000-fold glucose concentration gradient at 37°C is given by Eqn. (4.32):

$$\Delta\mu = 8.314\ \mathrm{J\,mol^{-1}K^{-1}} \times 310\ \mathrm{K} \times \ln(1/1000) = -17.8\ \mathrm{kJ\,mol^{-1}}. \tag{5.21}$$

That is, if the glucose concentration in the blood is high, as after a meal, and the concentration in cells is low, the sugar molecules enter cells spontaneously. As we have seen, once the sugar gets in, it is "tagged" with a charged phosphoryl group, preventing its escape through the hydrocarbon membrane. And if the concentration of chloride in the blood is about 100 mM, whereas that in the urine is about 160 mM, work must be done to pump chloride out of the blood and into the urine. You can easily calculate the work done by the kidneys in this process: $\Delta\mu = 1.9872\ \mathrm{cal\,mol^{-1}K^{-1}} \times 310\ \mathrm{K} \times \ln(160/100) = 290\ \mathrm{cal\ mol^{-1}}$. Moreover, you can estimate the number of chloride ions transported per ATP molecule hydrolyzed: free energy change of ATP hydrolysis in the cell/energy required to transport $Cl^- = 10\,000\ \mathrm{cal\,mol^{-1}}/290\ \mathrm{cal\,mol^{-1}} \approx 34$. We have ignored charge effects in the second calculation, but the magnitude should be about right.

We know from our study of the Donnan equilibrium that if the solute particle is charged, as in the case of N^+ or K^+, the situation is more subtle. Equation (4.32) does still apply, but we also need to take into account the work done as the charged particle moves through the

electrical potential across the membrane, $\Delta \boldsymbol{V}$. The magnitude of $\Delta \boldsymbol{V}$ is 10−200 mV, depending on the cell type – giving an electric field strength of as much as 200 000 V cm^{-1} across a membrane about 100 Å thick. This is only one order of magnitude smaller than the field where dielectric breakdown of air occurs and lightning strikes! Across every membrane of every cell in your body! In Chapter 4 we saw that the magnitude of the free energy change for electrical work is $\Delta\mu = nF\Delta \boldsymbol{V}$ when the ionic charge is n. Adding this term to Eqn. (4.32), we obtain

$$\Delta\mu = RT \ln[I]_{\mathrm{i}}/[I]_{\mathrm{o}} + nF\Delta \boldsymbol{V}, \tag{5.22}$$

where I represents an ionic solute and $\Delta \boldsymbol{V} = \boldsymbol{V}_{\mathrm{i}} - \boldsymbol{V}_{\mathrm{o}}$. The reference state must be the same for both terms of the right-hand side of this equation; in this case it is the extracellular matrix. When there is no driving force to move an ion from one side of the membrane to the other, $\Delta G = 0$ and

$$nF\Delta \boldsymbol{V} = -RT \ln[I]_i/[I]_o \tag{5.23}$$

which, on rearrangement, becomes

$$\Delta \boldsymbol{V} = -\frac{RT}{nF} \ln \frac{[I]_{\mathrm{i}}}{[I]_{\mathrm{o}}}. \tag{5.24}$$

We can use Eqn. (5.24) and the measured potential across the membrane to determine the ratio of the concentrations of the ionic solute. Let's assume, for instance, that we are working with a monovalent cation ($n = +1$) at 300 K, and let $\Delta \boldsymbol{V} = 120$ mV. Solving for $[I]_{\mathrm{i}}/[I]_{\mathrm{o}}$, we have

$$\begin{aligned} [I]_{\mathrm{i}}/[I]_{\mathrm{o}} &= \exp(-nF\Delta \mathbf{V}/RT) \\ &= \exp[(-96.5\ \mathrm{kJ\,V^{-1}mol^{-1}} \times 0.12\ \mathrm{V})/(8.314\,\mathrm{J\,mol^{-1}\,K^{-1}} \times 300\,\mathrm{K})] \\ &= 0.01. \end{aligned} \tag{5.25}$$

$[I]_{\mathrm{o}}$ is 100 times greater than $[I]_{\mathrm{i}}$.

At a number of points in this book nerve impulses have cropped up, for example, in the context of olfaction, perfume, and the Second Law. Now we wish to expand on the underlying mechanisms, albeit in highly qualitative terms. The aim here is to show how the development of this section fits within the broader picture of how animals work. Neurons, like other cell types, have ion-specific "pumps" situated in the plasma membrane. These protein machines use the energy of ATP hydrolysis to generate ionic gradients across the membrane in a way that resembles how electron transport proteins use the energy of glucose metabolism to generate a proton gradient (below). When at rest, a neuron is not very permeable to Na^+ (which is concentrated outside the cell and dilute inside) or K^+ (concentrated inside, dilute outside). There is a voltage on the order of 60 mV across the "resting" membrane. Stimulation of a nerve

cell results in a "depolarization" of the membrane. In the process, voltage-sensitive channel proteins that are selective for specific ions are "activated" by the decrease in voltage, allowing Na^+ ions in and K^+ ions out. The combined effect of gain and loss of membrane permeability to these ions is a millisecond time scale spike in membrane potential, known as an *action potential* – the potential across the membrane is reflecting the action of nerve impulse transmission. Depolarization of one part of the membrane by an action potential triggers depolarization of the adjacent part of the membrane, thereby propagating the action potential down the axon of neuron. Nerve impulses travel in one direction only because a certain amount of time is required for the recently depolarized part of the cell to regenerate its ion gradient.

A protein that has been mentioned at several points above is ATP synthase, the most famous of all transport proteins. ATPase is a type of molecular motor that plays a vital role in bioenergetics. Equation (5.22) can be used to describe the energetics of the "energy-transducing" membranes involved in ATP synthesis. In this case, the membrane of interest is the inner membrane of mitochondria and the ion is hydronium. The term $\ln[I]_i/[I]_o$ becomes $\ln [H^+]_i/[H^+]_o$, which can be rewritten as $2.3\Delta pH$ ($pH = -\log[H^+]$, and $\ln x \approx 2.3 \log x$). Substituting into Eqn. (5.22) and converting from units of energy to volts gives the *proton motive force* of chemiosmotic theory:

$$\Delta\mu_{H^+} = 2.3RT(pH_o - pH_i) + nF\Delta \boldsymbol{V}. \tag{5.26}$$

The measured membrane potential across the inner membrane of a liver mitochondrion is about -170 mV ($\boldsymbol{V}_i - \boldsymbol{V}_o$), and the pH of its matrix is about 0.75 units *higher* than that of its intermembrane space. Thus,

$$\begin{aligned}\Delta\mu = {} & [2.3 \times 8.314\,\mathrm{J\,mol^{-1}\,K^{-1}} \times 298\,\mathrm{K} \times (-0.75)] \\ & + [1 \times 96\,500\,\mathrm{J\,V^{-1}\,mol^{-1}} \times (-0.17\,\mathrm{V})],\end{aligned} \tag{5.27}$$

the sum total being about -21 kJ mol^{-1} for transport of a proton *into* the matrix.

The basic ideas discussed here apply not only to the synthesis of ATP from ADP and P_i and the accumulation of ions across a membrane, but also to a broad range of transport processes occurring across plasma membranes and neuronal synaptic vesicles. Before concluding this section, let's take the opportunity to see how ATP synthesis is a matter of energy coupling on a grand scale. This will help us to see how things tie together, how marvelously integrated the various aspects of the living cell are.

As we have seen, glucose oxidation in aerobic organisms is coupled to the reduction of oxygen to water. Electron transport proteins play a key role in the process. The overall redox reaction, which is energetically favorable, is used to pump protons *against* their concentration gradient to the opposite side of the membrane. In other words, the pH of solution on one side of the membrane is

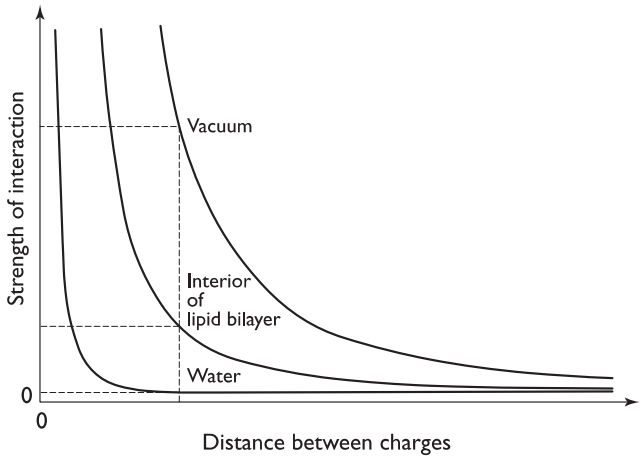


Fig. 5.13 Dependence of electrostatic energy on distance between charges and medium. The energy is inversely proportional to distance, so attraction or repulsion is greatest when the charges are close to each other. The energy also depends substantially on the stuff between the charges, varying inversely with the dielectric constant of the medium. The dielectric constant is constant for a given temperature and pressure, and it must be determined empirically. The interaction between charges is greatest in vacuum. In water, where the dielectric constant is very large, charges must be very close for the interaction between them to be significant. Charge–charge interactions are relatively large in the core of a protein or in the plasma membrane, because the dielectric constant of hydrocarbon is much lower than that of water.

different from that on the other side. And the voltage difference across the membrane, which is only about 10 nm thick, is about 200 mV, so the electric field strength in the middle of the membrane is *huge*! Protons migrate down their concentration gradient through a protein channel in the lipid membrane. Protons don't pass straight through the membrane because it is made of lipids. The channel is lined with polar chemical groups, making proton passage energetically favorable (Fig. 5.13). An amazing thing about this protein channel is that the energy change of proton translocation is coupled to an energetically unfavorable process – ATP synthesis. This is not a trivial chemical coupling; remember, the free energy change on hydrolyzing ATP to ADP is about 10 kcal mol^{-1} at cellular concentrations. For all practical purposes, hydrolysis of ATP is *irreversible*! The point of this discussion is that the cell (in fact, the mitochondria here, but the principles are the same for entire cells) must do work to generate the proton gradient. But there is a sort of purpose to the work. No wonder we need to eat from time to time! The cell is a sort of machine, and a very efficient one with regard to use of energy resources!

I. Enzyme–substrate interaction

In Chapter 2 we touched on the biological function of ribonuclease A (RNase A), a digestive enzyme that hydrolyzes RNA to its component nucleotides. We said that an inhibitor of the enzyme, 2′-cyclic monophosphate, can be used to study the enthalpy of nucleotide binding to RNase A. One aim of Chapter 5 is to illustrate the general utility of Eqn. (4.2). In the present section we focus on how Eqn. (4.2) applies to studies of the energetics of binding of small compounds to protein or DNA, taking RNase A as our example.

Figure 5.14 shows an outline of the reaction catalyzed by RNase A. The scheme is based on the isolation of 2′,3′-cyclic nucleotides from RNase A digests of RNA. There are four types of 2′,3′-cyclic

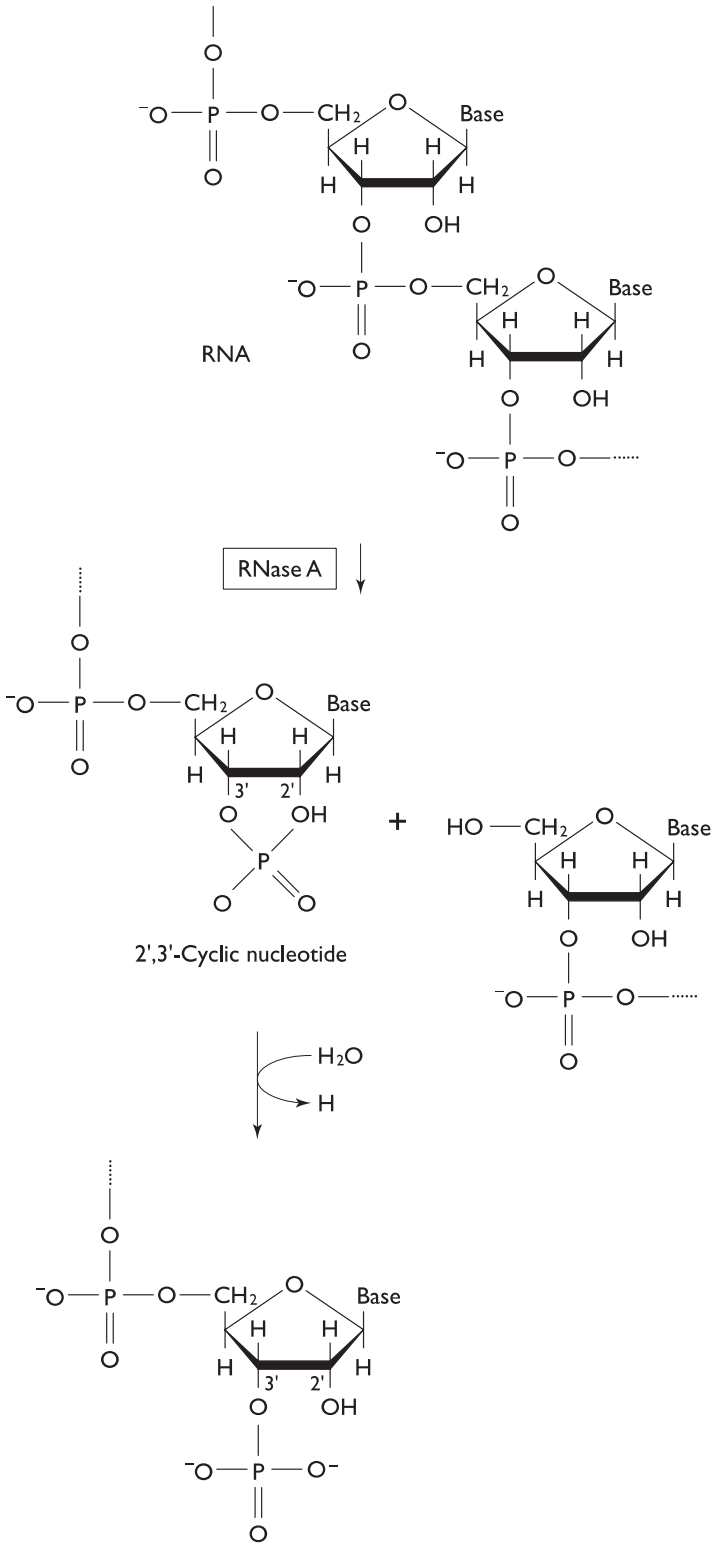


Fig. 5.14 Mechanism of RNase A activity. Bovine pancreatic ribonuclease A is an example of enzyme-mediated acid–base catalysis. The enzyme hydrolyzes RNA to its component nucleotides. The reaction scheme is based on the experimental finding that 2′,3′-cyclic nucleotides are present in RNase digests of RNA. RNase is inhibited by 2′-CMP. This binding interaction has been studied in considerable depth. See Chapter 8 for further information.

nucleotides. RNase A hydrolysis of one of them, cytidine 2′,3′-cyclic phosphate, has been studied extensively. The reaction is

$$\text{cytidine } 2',3'\text{-cyclic phosphate(aq)} + H_2O(l) \rightarrow \text{cytidine } 3'\text{-cyclic phosphate(aq)} \quad (5.28)$$

How might one determine ΔG° for this reaction?

From Eqn. (4.32),

$$\Delta G = \Delta H - T\Delta S. \quad (5.29)$$

If the products and reactants are in the standard state, the thermodynamic relationship is

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ. \quad (5.30)$$

We'll need to know ΔH° and ΔS° to calculate the energy. The enthalpy change of the reaction, which can be estimated by calorimetry, is -2.8 kcal mol^{-1}. But what is ΔS°? One approach would be to make use of the fact that S is a state function and combine measurements that, when summed, give ΔS° for Eqn. (5.28). The reaction scheme might look like this:

$$\text{cytidine } 2',3'\text{-cyclic phosphate(aq)} + H_2O(l) \rightarrow \text{cytidine } 2',3'\text{-cyclic phosphate(s)} + H_2O(l) \quad (5.31)$$

$$\text{cytidine } 2',3'\text{-cyclic phosphate(s)} + H_2O(l) \rightarrow \text{cytidine } 3'\text{-cyclic phosphate(s)} \quad (5.32)$$

$$\text{cytidine } 3'\text{- cyclic phosphate(s)} \rightarrow \text{cytidine } 3'\text{-cyclic phosphate(aq)}. \quad (5.33)$$

Equation (5.31) represents the dissolution of cytidine 2′,3′-cyclic phosphate, Eqn. 5.32 the conversion of cytidine 2′,3′-cyclic phosphate to cytidine 3′-cyclic phosphate in the solid state, and Eqn. (5.33) the dissolution of cytidine 3′-cyclic phosphate. The sum of these reactions is Eqn. (5.28), the conversion of cytidine 2′,3′-cyclic phosphate to cytidine 3′-cyclic phosphate in aqueous solution. If the entropy changes of these several reactions can be measured, ΔS° can be calculated for Eqn. (5.28). And combining ΔS° for Eqn. (5.28) with ΔH° for the overall reaction will give ΔG° for the overall reaction.

The entropy changes for Eqns. (5.31)–(5.33) have in fact been determined experimentally at 25°C. The values are: +8.22 cal mol^{-1} K^{-1}, −9.9 cal mol^{-1} K^{-1}, and +8.28 cal mol^{-1} K^{-1}, respectively. The overall ΔS° for these reactions is just the sum of the individual contributions, 6.6 cal mol^{-1} K^{-1}. Combining this entropy change with the calorimetric enthalpy change gives -2800 cal mol^{-1} $-$ 298 K $\times$ 6.6 cal mol^{-1} K^{-1} $= -$ 4800 cal mol^{-1}. That ΔG° is negative suggests that cytidine 2′,3′-cyclic phosphate will hydrolyze *spontaneously* in aqueous solution, and this is confirmed by experiment. One could test the role that RNase A might play in this process by studying the effect of the wild-type enzyme and point mutants on the rate of reaction. See Chapter 8.

J. Molecular pharmacology

This is an important topic. The equations presented here are more general than the section title may suggest, as they can be applied not only to the interactions of drugs with membrane-bound receptor proteins but also to proteins that bind DNA, small molecules or ions. Binding interactions play a role in regulating enzyme activity and biosynthetic pathways, oxygen transport and regulation of blood pH, and many (!) other physiological processes. But for now, let's think of binding in the context of a single ligand–receptor interaction in pharmacology. This will help to motivate the mathematical development. A more thorough treatment of binding will be given in Chapter 7.

Equation (4.32) can be used to describe a chemical reaction in terms of reactants and products. But it could just as well represent the free energy difference between the "bound" and "free" states of a ligand, a small molecule or an ion. Under appropriate conditions, a ligand will interact with a macromolecule at a binding site. In some cases binding is highly specific; in other cases, not. In either case, ΔG° represents the driving force for binding under standard state conditions.

Here, we'll represent the binding reaction as

$$\mathrm{R} + \mathrm{L} \Leftrightarrow \mathrm{R} \bullet \mathrm{L}, \tag{5.34}$$

where R is the receptor, L signifies *free* ligand molecules, and R • L is the receptor–ligand complex. It is assumed that there is only one binding site for L per receptor molecule. The *association* constant is *defined* as

$$K_a = [\mathrm{R} \bullet \mathrm{L}]/([\mathrm{R}][\mathrm{L}]) \tag{5.35}$$

and the *dissociation* constant is

$$\mathrm{K_d} = \mathrm{K_a^{-1}} = [\mathrm{R}][\mathrm{L}]/[\mathrm{R}\bullet\mathrm{L}] = ([\mathrm{R}]_\mathrm{T} - [\mathrm{R}\bullet\mathrm{L}])[\mathrm{L}]/[\mathrm{R}\bullet\mathrm{L}], \tag{5.36}$$

where $[\mathrm{R}]_\mathrm{T} = [\mathrm{R} \bullet \mathrm{L}] + [\mathrm{R}]$ is the total receptor concentration. The fractional occupancy of ligand-binding sites, F_b, is

$$F_b = [\mathrm{R} \bullet \mathrm{L}]/[\mathrm{R}]_\mathrm{T} = [\mathrm{L}]/(\mathrm{K_d} + [\mathrm{L}]). \tag{5.37}$$

A plot of F_b against [L] is shown in Fig. 5.15A. The shape of the curve is a *rectangular hyperbola*. Equation (5.37) indicates that K_d corresponds to the concentration of L at which the occupancy of binding sites is half-maximal. Many physiological dissociation constants are on the order of μM–nM. A nM binding constant is considered "tight binding." When Eqn. (5.37) is plotted as percentage response against dose (for example, mg of drug per kg of body weight), it is called a *dose–response curve*. The dose is often plotted on a logarithmic scale, giving the curve a sigmoidal appearance (Fig. 5.15B), but the underlying relationship between dose and response is the same in both cases.

We can rearrange Eqn. (5.36) to obtain

$$[\mathrm{R} \bullet \mathrm{L}]/[\mathrm{L}] = ([\mathrm{R}]_\mathrm{T} - [\mathrm{R} \bullet \mathrm{L}])/K_d. \tag{5.38}$$

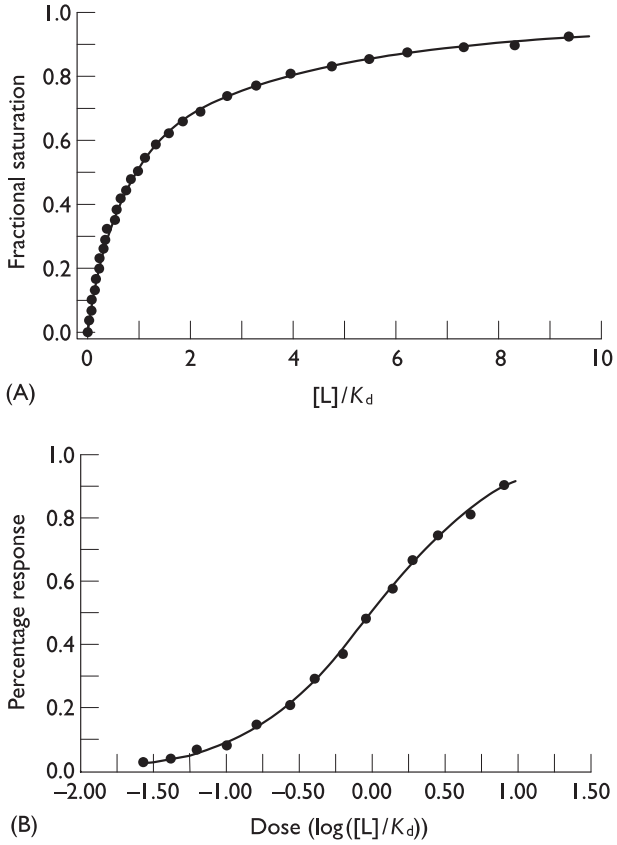


Fig. 5.15 Binding. In panel (A) the circles are experimental data points, the solid line is a theoretical description of binding. There is one ligand-binding site per macromolecule. In such cases, the mathematical relationship between the bound and free ligand concentrations is a rectangular hyperbola. Note that although half-saturation occurs when $[L] = K_d = 1/K_a$, $[L] = 9K_d$ gives only 0.9 saturation and $[L] = 99K_d$ but 0.99 saturation. In other words, most of the information about binding is in the free ligand concentration range $0–2K_d$. Experiments should be designed accordingly. Panel (B) shows a dose–response curve.

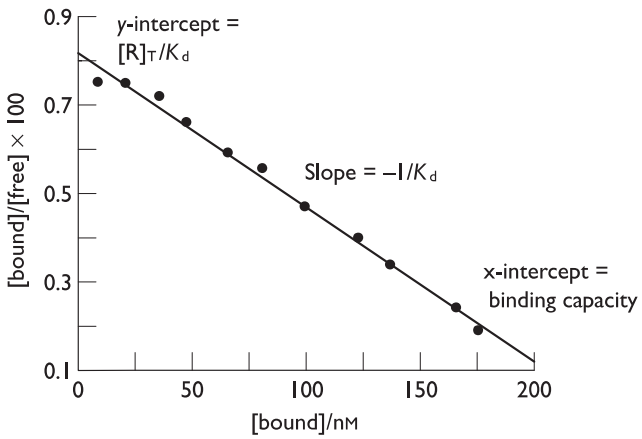


Fig. 5.16 Scatchard plot. The concentration of bound ligand divided by the concentration of free ligand is plotted against the concentration of bound ligand (nM). When binding data are presented in this way, the slope measures the negative inverse of the dissociation constant ($-1/K_d = -K_a$). The vertical axis-intercept is $[R]_T/K_d$, and the horizontal axis-intercept is the binding capacity (the concentration of binding sites).

In this representation [R•L]/[L], the concentration of bound ligand divided by the concentration of free ligand, is a *linear* function of [R•L]. The slope of the curve is $-1/K_d$ (see Fig. 5.16). The axis intercepts themselves represent interesting quantities: the intercept on the vertical axis is $[R_T]/K_d$, and the intercept on the horizontal axis is the "*binding capacity*," the "concentration" of ligand binding sites. A plot of bound/free ligand *versus* bound ligand is called a *Scatchard plot*, after the American physical chemist George Scatchard (1892–1973). Radioactive methods are one way that biological scientists measure the amounts of bound and free ligand.

Experiments can be done to determine the dissociation constant of other ligands that can compete for the same binding site as L. For instance, suppose you wish to test the effectiveness of a number of candidate drugs to compete directly with a physiological ligand L for a specific binding site on R. Let the candidate competitors be I_1, I_2, $I_3 \ldots$ According to this model,

$$K_{d,I_i} = [R][I_i]/[R \bullet I_i] \tag{5.39}$$

for a general inhibitor compound, I_i. It can be shown that in the presence of an inhibitor, the receptor–ligand complex, [R•L], is

$$[R \bullet L] = \frac{[R]_T[L]}{K_{d,L}\left(1 + \frac{[I_i]}{K_{d,I_i}}\right) + [L]}. \tag{5.40}$$

The relative affinity of a ligand in the presence of an inhibitor can be found by dividing Eqn. (5.40) by Eqn. (5.38). This gives

$$\frac{[R \bullet L]_{I_i}}{[R \bullet L]_0} = \frac{K_{d,L} + [L]}{K_{d,L}\left(1 + \frac{[I_i]}{K_{d,I_i}}\right) + [L]}. \tag{5.41}$$

Equation (5.41) is zero for all concentrations of I_i when there is no inhibition (compound I_i has no effect), and it is 1 at 100% inhibition. The concentration of competitor I_i that gives 50% inhibition is designated $[I_{i,50}]$. At this concentration,

$$K_{Ii} = \frac{[I_{i,50}]}{1 + \frac{[L]}{K_{d,L}}}. \tag{5.42}$$

Figure 5.17 shows the percentage inhibition for a number of different inhibitors. Note that the shape of the curves resembles that in Fig. 5.15a.

The above mathematical equations apply not only to natural ligands like the "fight-or-flight" hormone epinephrine and to competitive inhibitors like the "β-blocker" propranolol, which vies with epinephrine for binding sites on β-adrenergic receptors,[9] but also to noxious chemical substances like botulinum toxin. Moreover, the equations also apply to ligands of DNA, for example, repressor proteins that physically block the enzymatic transcription of mRNA by binding to an operator site, and to protein–protein interactions related to signal transduction. An example of ligand binding in a signal transduction cascade is the direct association of the SH2 domain[10] of the protein Grb2 to a specific phosphorylated tyrosine residue on a growth factor receptor (Fig. 5.18).

Phosphotyrosine-mediated binding is of particular interest in biological science for several reasons. One, it involves phosphorylated tyrosine, and the phosphoryl group is acquired via catalysis by a kinase from ATP, the energy molecule. Phosphorylation and

[9] The effects of β-blockers were first described by Sir James W. Black (1924–), a Scot. Sir James was awarded the Nobel Prize in Medicine or Physiology in 1988.

[10] SH2, Src homology 2.

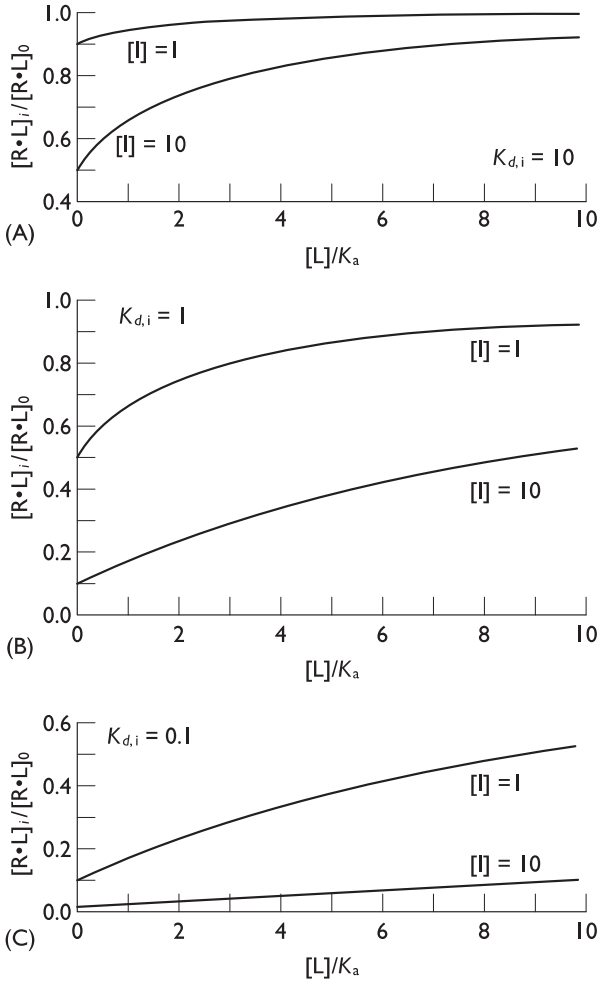


Fig. 5.17 Effect of inhibitor on ligand binding. When the concentration of inhibitor i is low and the inhibitor dissociation constant is high, as in panel (A), $[R\bullet L]_i/[R\bullet L]_0$ is nearly 1 even at low concentrations of ligand. Competition between ligand and inhibitor is more evident when the inhibitor concentration is increased by a factor of 10. In panel (B), the dissociation constant of the inhibitor is 10 times smaller than in panel (A). Note the marked impact this has on $[R\bullet L]_i/[R\bullet L]_0$. The effect of decreasing the dissociation constant by yet another factor of 10 is shown in panel (C). This is the sort of study a pharmaceutical company might do to characterize the properties of inhibitors that could be used as drugs. ITC can be used to screen different compounds. Analysis of such compounds will include not only *in vitro* binding experiments (high-affinity specific binding) but also assessment of side effects (low-affinity non-specific or high-affinity unwanted binding).

dephosphorylation of tyrosine is a type of dynamic molecular switch that regulates cellular activity by controlling which proteins can interact with each other. Phosphorylation also places severe restrictions on the relative orientation of interacting proteins. An important class of phosphotyrosine-mediated interactions is typified by phospholipase $C_{\gamma 1}$ (PLC), an enzyme that interacts with phosphorylated growth factor receptors by means of its two SH2 domains and is involved in lipid metabolism. Binding of Grb2 to a receptor resembles that of PLC, but Grb2 has no catalytic activity; Grb2 is a sort of "adaptor" protein. Two, there are several different types of phosphotyrosine recognition module, and they are found in many different proteins. Two of the best-known phosphotyrosine binding modules are the SH2 domain and the PTB (phosphotyrosine binding) domain. In some cases, both types are found in the same protein, for example, Shc[11] and tensin. Three, the breadth of the range of possible interactions of a given type of module is greatly increased by subtle differences in structure. As a general rule, the amino acid side

[11] Shc, Src homolog, collagen homolog.

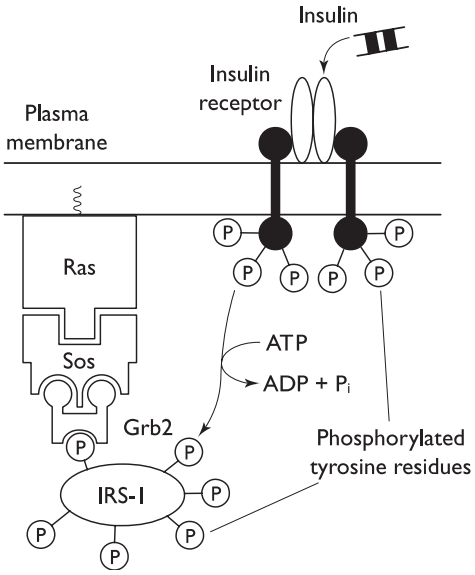


Fig. 5.18 Protein–protein interactions and phosphoryl transfer in signal transduction. The extracellular concentration of the hormone insulin, a peptide signal, is communicated across the plasma membrane by means of dimeric insulin-specific transmembrane receptor molecules. The binding of insulin to its receptor results in receptor autophosphorylation, the catalysis by one receptor molecule of the transfer of a phosphoryl group from ATP to a tyrosine side chain of the other receptor molecule. Phosphorylation of tyrosine acts as a molecular switch in the recruitment of proteins that recognize specific phosphorylated tyrosine residues. One consequence of the chain of events elicited by insulin binding is the phosphorylation of insulin-receptor substrate-1 (IRS-1). Again, ATP is involved. Once phosphorylated, IRS-1 can interact directly with the proteins Grb2, Sos, and Ras. The last of these plays a very important role in cellular signal transduction. The key point here is that several of the protein–protein interactions involved in this and many other signaling cascades are mediated by phosphorylated tyrosine, and phosphorylation of tyrosine depends on ATP. Signal transduction is a form of biological communication and information processing. We shall return to this point in Chapter 9. The figure is based on Fig. 20–48 of Lodish *et al.* (1995).

chains that interact directly with the ligand are conserved from module to module, but side chains nearby are specific to the protein in which the module occurs. Such peculiarities underlie the specificity of the interactions of an otherwise general protein structure. The general and specific interactions combine to give the overall binding free energy. We can see here a close connection between the existence and transmission of biological information and energy.

Finally, it should be mentioned that binding is a far more complicated phenomenon that we have made it seem. For instance, if a macromolecule can interact with more than one type of ligand at different binding sites, there is the possibility that one kind of metabolite can "sense" the concentration of another, despite the absence of a direct interaction between the metabolites. This aspect of the function of biological macromolecules, known as *allostery*, will be developed along with other aspects of binding in Chapter 7.

K. Hemoglobin

Most known organisms require oxygen for life; the only known exceptions are some types of bacteria and archaea. Reduction of molecular oxygen to water plays a key role in the generation of ATP. In every cell in an aerobic organism, oxygen is used as fuel in the combustion of glucose and production of ATP and carbon dioxide. Oxygen and glucose must be delivered to every cell in the body; carbon dioxide, a waste product, must be removed from every cell. Vertebrates carry out this food and gas transport by means of blood or blood-like fluid that moves through a closed system of tubes called the vasculature (Chapter 1). The vasculature makes contact with the lungs, gills or skin on the one hand, and the peripheral tissues on the other hand.

Molecular oxygen is transported throughout the blood by an allosteric transport protein called hemoglobin. In view of this, hemoglobin has been called the "molecular lung." Vertebrate hemoglobin is a tetrameric protein, $\alpha_2\beta_2$; it can be thought of as a dimer of $\alpha\beta$ heterodimers. In invertebrates, hemoglobins range from one to 144 subunits! Each subunit consists of a polypeptide chain called globin and a protoheme IX, a planar complex of an iron and ion protoporphyrin IX (Fig. 5.1). Iron plays a role in the coordination of bound dioxygen. The ability of hemoglobin to bind oxygen depends not only on the structure of the protein and oxygen but also on the partial pressure of oxygen.[12] In hemoglobin, the extent of oxygen loading into binding sites influences the affinity of the other binding sites to bind oxygen. The specific character of the amino acid chains near the protein–heme interface is essential for oxygen binding, as shown by amino acid replacement studies. Mutations in the region of the oxygen binding site can alter affinity for oxygen by over 30 000-fold! In the present section we introduce a number of aspects of hemoglobin thermodynamics. The treatment will be brief. A more in-depth look at oxygen binding is reserved for Chapter 7.

Now, if tetrameric hemoglobin is thermodynamically stable under normal physiological conditions, the tetramer must represent a minimum of free energy; the tetrameric state must be a lower free energy state than the other possible combinations of subunits, for example $\alpha\beta$ dimers. (There are other possibilities, for example, a kinetically trapped tetrameric state, but let's ignore them for now.) A number of natural variants of human hemoglobin are known. One of these is the famous sickle-cell variant. As shown in Table 5.3, the free energy difference between the tetrameric and dimeric states of

[12] The partial pressure of a gas is just the contribution that it makes to the overall gas pressure. By Dalton's Law, which is named after the John Dalton of the atomic hypothesis, the total pressure is just the sum of the partial pressures of the gases present. For example, if the air pressure is 1 atm, the partial pressures of nitrogen, oxygen, and carbon dioxide sum to 1 atm.

Table 5.3. *Thermodynamics of hemoglobin dissociation*

Hemoglobin	Substitution in mutant	ΔG° (kcal mol^{-1} of hemoglobin)
normal	—	8.2
"Kansas"	102β, Asn → Thr	5.1
"Georgia"	95α, Pro → Leu	3.6

The data are from Chapter 4 of Klotz.

hemoglobin can depend substantially on the primary structure, the sequence of amino acid residues. The free energy difference between normal hemoglobin and hemoglobin Kansas is "only" 3.1 kcal mol^{-1}, but the equilibrium constant differs by nearly 200-fold at 25 °C! Considerably less work must be done to dissociate tetrameric hemoglobin Kansas than wild-type hemoglobin into αβ dimers under the same conditions. For comparison, it is well known that inhalation of too much carbon monoxide will normally be fatal, even if exposure lasts just a few minutes. The spectrum of pathological effects of CO poisoning includes damage to the peripheral nervous system, brain damage, cell death in the heart, cell death in other muscles, and pathological accumulation of fluid in the lungs. All this results from the binding of CO to hemoglobin with an affinity constant "only" about 240 times greater than that of oxygen! Hemoglobin Georgia is even less stable than the Kansas variant, so its behavior as an oxygen carrier is very noticeably altered relative to the normal protein.

We have assumed that the tetrameric state of hemoglobin represents a lower free energy state than the dimer. And oxygen associates with hemoglobin. Binding occurs because the bound state is more thermodynamically favorable (has a lower Gibbs free energy) than the unbound state. Let's consider the oxygenation of hemoglobin in solution. For the moment, we'll take a rather simplistic view and assume that hemoglobin has just one binding site, or, more accurately, that each subunit binds O_2 with the same affinity. The reaction can be written as

$$\mathrm{Hb(aq)} + \mathrm{O_2(g)} \;\Leftrightarrow\; \mathrm{HbO_2(aq)}. \tag{5.43}$$

From experiments it is known that $K = 85.5\ \mathrm{atm}^{-1}$ for the reaction as written. At 19 °C, $\Delta G = -2580$ cal mol^{-1}. What is the free energy change when the partial pressure of oxygen is 0.2 atm and oxygen is dissolved in solution with an activity of 1 (as in the standard state)?

The free energy difference between $p = 1$ atm and $p = 0.2$ atm is found using Eqn. (4.5):

$$\begin{aligned}\Delta G &= G(\mathrm{O_2}, 0.2\ \mathrm{atm}) - G^\circ(\mathrm{O_2}, 0.2\ \mathrm{atm}) - (G(\mathrm{O_2}, 1\ \mathrm{atm}) - G^\circ(\mathrm{O_2}, 1\ \mathrm{atm})) \\ &= RT\ln(\mathrm{O_2}, 0.2\ \mathrm{atm}) - RT\ln(\mathrm{O_2}, 1\ \mathrm{atm}) \\ &= RT\ln(0.2/1) \\ &= -930\ \mathrm{cal\ mol^{-1}}.\end{aligned} \tag{5.44}$$

That ΔG is negative is just what we should expect, since a substance will always move spontaneously from a region of higher concentration to a region of lower concentration.

At equilibrium, $\Delta G = 0$ between the oxygen vapor and the dissolved oxygen. To calculate the free energy difference between the concentration of dissolved oxygen in the standard state ($a = 1$) and the concentration at saturation, which is substantially lower, we need to account for the solubility of diatomic oxygen in water. This is 0.000 23 molal (kg l^{-1}) at 19°C. Thus, $\Delta G = RT\ln(1/0.000\ 23) = 4\,860$ cal mol^{-1}. Because the Gibbs free energy is a state function, the *net* free energy change on going from oxygen gas at 1 atm to dissolved oxygen at unit activity is just the sum of the individual contributions, or -930 cal $mol^{-1} + 0 + 4860$ cal $mol^{-1} = 3930$ cal mol^{-1}. The free energy change of the reverse reaction is, of course, -3930 cal mol^{-1}.

Now, the two reactions we're interested in are:

$$\mathrm{Hb(aq) + O_2(g) \Leftrightarrow HbO_2(aq)} \tag{5.45}$$

$$\mathrm{O_2(aq) \Leftrightarrow O_2(g)}, \tag{5.46}$$

which, when summed, give

$$\mathrm{Hb(aq) + O_2(aq) \Leftrightarrow HbO_2(aq)}. \tag{5.47}$$

ΔG for the overall reaction is -2580 cal mol^{-1} -3930 cal $mol^{-1} = -6510$ cal mol^{-1}. We can see that the driving force for oxygen association with hemoglobin is greater when the oxygen is solvated than when it is not solvated.

L. Enzyme-linked immunosorbent assay (ELISA)

Antibodies are protective proteins produced by the immune system in response to the presence of a foreign substance, called an antigen. Antibody recognition of an antigen is mainly a matter of shape complementarity and charge interactions in the antigen-binding site. The shape of the binding site must be a close match to a part of the surface of an antigen for specific binding to occur. Binding can be very tight indeed, with $K_{eq} \sim 10^9$ M^{-1} or greater, and highly specific. The following discussion, though it centers on ELISA, applies to a broad range of immuno-techniques, including for instance western blotting.

ELISA is a useful method for detecting small amounts of specific proteins and other biological substances ("antigens") in laboratory and clinical applications. For instance, it is used to detect the placental hormone chorionic gonadotropin in a commonly available pregnancy test. The assay is so useful because it is very general, antibody binding is very specific (K_{eq} is large), and the sensitivity of the binding "signal" can be increased in various ways, for instance, a covalent link between the antibodies used for detection and an enzyme (often horseradish peroxidase).

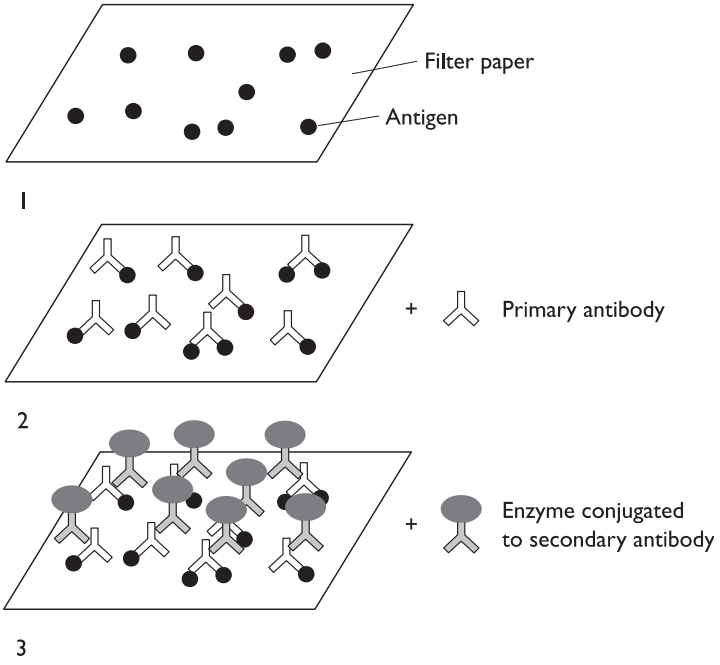


Fig. 5.19 ELISA. This very useful laboratory technique consists of three basic steps. First, the protein antigen is adhered to a solid support, often a nitrocellulose filter. This partially denatures the antigen. Next, (primary) antibodies are allowed to bind to the antigen. Finally, (secondary) antibodies that recognize the first antibody bind are allowed to bind the primary antibodies. Attached to each secondary antibody is an enzyme that is used to catalyze a reaction that facilitates detection. One assumes that detection of the enzyme linked to the secondary antibody implies detection of the antigen. This is often the case because antibody recognition of antigen is highly specific. But because nitrocellulose is very sticky, the milk protein casein is often used to bind sites not occupied by antigen in order to reduce the background signal arising from the non-specific adherence of primary antibodies. After step two, non-specifically adhered primary antibodies are rinsed off with buffer. Specifically bound antibodies are not lost in the rinsing procedure because the rate of dissociation of the antibody from the antigen is very low (Chapter 8).

The ELISA protocol involves adsorbing an "antigen" of interest to an "inert" solid support (usually a type of filter paper, Fig. 5.19). The binding of antigen to the solid support can be very strong indeed, though binding is usually relatively non-specific. The binding process usually results in partial denaturation of a protein antigen. After adsorption, the sample is screened with an antibody preparation (usually a rabbit antiserum) and "rinsed" to remove non-specifically bound antibody ($K_{eq} < 10^4$). The resulting protein–antibody complex on the solid support is reacted with an antibody-specific antibody to which the enzyme used for the detection assay is attached. This second antibody is often from goat. Why does the rinse step does not ruin the experiment?

As we shall see in Chapter 8, $K_{eq} = k_f/k_r$, where k represents reaction rate and "f" and "r" stand for "forward" and "reverse," respectively. When binding is specific, $k_f >> k_r$; the "on rate" (binding) is much greater than the "off rate" (release). So, even during rinsing tightly bound antibodies stay put, despite the requirement for mass action to release antigen to solution, where the antibody concentration is low. To put things into perspective, for $K_{eq} \sim 10^9$, the free energy change on binding is about -50 kJ mol^{-1}! We can get a sense of how big this free energy change is by considering the energy required to raise a 100 g apple a distance of 1 m. It is easy to show that this energy is about 1 J (Chapter 2). This tells us that 50 kJ could lift a 100 g mass about 50 km, over 5 times the height of Mt Everest! So $K_{eq} \sim 10^9$ M^{-1} is tight binding. This has been a very brief and consequently superficial treatment of ELISA. But it has been enough to illustrate yet another way the basic ideas of thermodynamics are useful for understanding biological science.

We'll come back to the important subjects of binding and chemical kinetics in Chapter 7 and Chapter 8, respectively.

M. DNA

Throughout this book we have put somewhat more emphasis on proteins than on DNA. In part this is a reflection of the expertise and interests of the author and not a conscious bias against nucleic acids or the people who study them! Besides, we shall see just how important DNA is to the entire story when we reach Chapter 9. Nevertheless, to redress lingering impressions of imbalance, this section looks at the thermostability of DNA and the next one discusses energetic aspects of the polymerase chain reaction.

The structure of the DNA double helix is illustrated schematically in Fig. 5.20. The types of interaction that stabilize the structure are hydrogen bonds and "stacking interactions." Three hydrogen bonds are formed between bases cytosine and guanine, two between adenine and thymine. These are the co-called Watson–Crick base pairs. The adenine of DNA (and RNA) is exactly the same as the adenine of ATP, cAMP, NADH and $FADH_2$. As shown in Fig. 5.14, however, only one of the phosphate groups of ATP actually becomes part of the polynucleotide. You might guess, then, that the stability of double-stranded DNA relative to single-stranded DNA will depend on the proportion of C–G pairs, because this will influence the average number of hydrogen bonds per base pair, and in fact that is correct (Fig. 5.21, Table 5.4). Analysis of structures of nucleic acids has revealed that the bases form extended stacks, interacting with each other by van der Waals forces. Both hydrogen bonds and van der Waals interactions contribute to the overall stability of the double helix.

The equilibrium between double- and single-stranded DNA can be symbolized as

$$D \Leftrightarrow S. \tag{5.48}$$

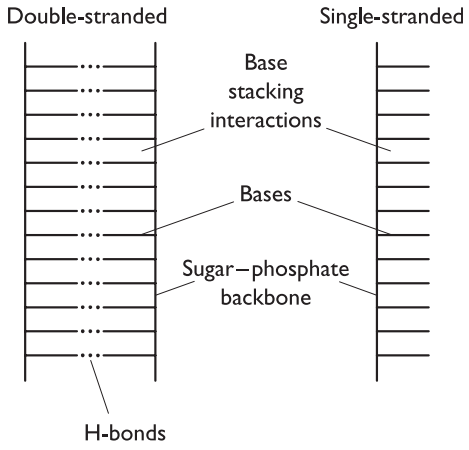


Fig. 5.20 Double-stranded and single-stranded DNA. DNA is composed of bases attached to a sugar–phosphate backbone. (See Fig. 5.14 for a higher resolution view of polynucleic acid.) There are two major types of interaction that stabilize double-stranded DNA: intermolecular hydrogen bonds (polar) and intramolecular base stacking interactions (non-polar). The number of hydrogen bonds depends on the bases involved: three are formed between cytosine (C) and guanine (G), and two between adenine (A) and thymine (T). Intermolecular hydrogen bonds are not present in single-stranded DNA. Based on Fig. 3.16 of van Holde (1985).

Table 5.4. *Association constants for base pair formation*

Base pair	$K(M^{-1})$
Self-association	
A•A	3.1
U•U	6.1
C•C	28
G•G	10^3–10^4
Watson–Crick base pairs	
A•U	100
G•C	10^4–10^5

The measurements were made in deuterochloroform at 25 °C. The data are from Kyoguko *et al.* (1969). Similar values have been obtained for T in place of U. Non-Watson–Crick base pairs are relatively unstable. The entropic component of K is roughly the same for each

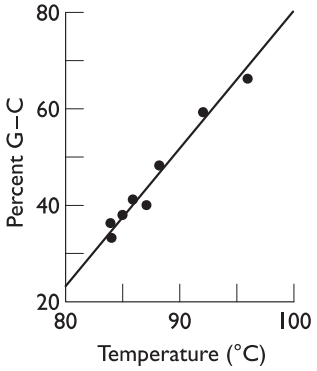


Fig. 5.21 Influence of G–C content on DNA melting temperature. As the percentage of G–C pairs increases, the number of intermolecular hydrogen bonds per base pair increases. The stabilizing effect on double-stranded DNA is reflected in the relationship between G–C content and melting temperature.

The equilibrium constant for this reaction is

$$K = [S]/[D]. \tag{5.49}$$

This equilibrium constant is the product of the Ks for the individual base pairings, as each base pair contributes to the overall stability of the double-stranded molecule.

Thermal denaturation of double-stranded DNA has been studied extensively. As mentioned above, C–G composition is an important determinant of stability of duplex DNA and therefore conditions under which $K = 1$. One means of promoting the dissociation of double-stranded DNA is to add heat. Just as with proteins, heat absorbed by DNA increases its thermal energy, fluctuations of structure become larger, and the disordered state becomes more probable than the ordered one. Measurement of the temperature at which double-stranded DNA is 50% "melted," the melting temperature, is one way of comparing the genetic material of one genome to another (Fig. 5.22). Research in this area has been used to work out empirical rules for the melting temperature of DNA as a function of C–G content, total number of base pairs, and concentration of ions, principally Mg^{2+}. Magnesium ions neutralize the electrostatic repulsion between the negatively charged phosphate groups in the sugar–phosphate backbone by decreasing the range and strength of the repulsive Coulombic interactions[13] between the phosphate groups on opposite strands of the double helix. Decreases in the concentration of such counterions increase the repulsion between strands and reduce the melting temperature of double-stranded DNA.

Figure 5.23 shows percentage of double-helix as a function of temperature for the forward and reverse reactions in Eqn. (5.48).

[13] Coulomb's law is an empirically derived mathematical description of the interaction between charged particles. It is named after the French physicist and military engineer Charles Augustin de Coulomb (1736–1806). The effect of Mg^{2+} counterions is explained by the Debye–Hückel theory of strong electrolytes.

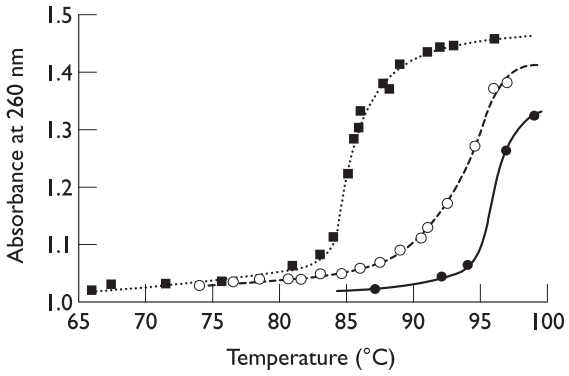


Fig. 5.22 DNA melting curves. The melting temperature varies not only with G–C content but also with size. In other words, a 100 base pair-long double-stranded DNA molecule will have a higher melting temperature than a 50 base pair-long double-stranded DNA for a given percentage and distribution of G–C pairs. Differences in G–C content and distribution and molecular size lead to differences in melting temperature.

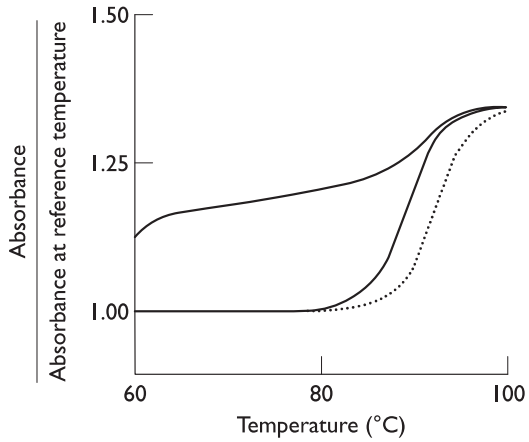


Fig. 5.23 Melting and cooling profile for double-stranded DNA. Solid line. The melting of double-stranded DNA is cooperative. Relatively few base pairs are broken below the melting temperature. Once melting has begun, however, relatively small increases in temperature result in the rupture of a relatively large number of hydrogen bonds. The melting profile differs greatly from the cooling profile. Unless cooling is carried out very slowly, the system will not be in a near-equilibrium state at every temperature value on the reaction pathway. Rapid cooling of melted DNA will not yield perfectly formed double-stranded DNA. Broken line. Melting temperature is influenced not only by G–C content and distribution and molecular size but also by ion concentration, particularly divalent cations. These ions interact favorably with the phosphate groups of the DNA backbone. The consequent reduction of electrostatic repulsion results in increased stability of the DNA duplex. Based on Fig. 3.15 of van Holde (1985) and Fig. 5.15 of Bergethon (1998).

The evident hysteresis (*Greek*, a coming late) in the reverse process arises from a difference in the rates of hydrogen bond breaking and specific annealing of the complementary strands, and the sample's not being at equilibrium throughout the experiment. The dissociation of strands is a much simpler reaction than the formation of perfectly matched double-stranded DNA. And as we have said, reversibility of a process depends on the system being taken through a series of equilibrium or near-equilibrium states. During the reverse reaction, if the system is not given sufficient time to come to equilibrium, some mismatching of bases is likely to occur, preventing or strongly inhibiting the return to the initial conditions.

Thus far we have described DNA in rather general terms. There are different types of DNA, however, not only differences in G–C content, and the different types have different thermodynamic properties. Genomic DNA of higher eukaryotes, for example, is linear: there is a distinct 3′-end and a distinct 5′-end. In plasmid DNA, by contrast, which is of great utility as a vector for carrying "foreign" genes into *E. coli* for production of "recombinant" proteins, there is no distinct 3′-end or 5′-end; this DNA is *circular*. Such DNA can exhibit a variety of conformations ranging from no

supercoiling, or no twisting, to tight supercoiling. This topological characteristic of circular DNA suggests that energetics of plasmid DNA melting will differ from that of linear DNA, even if the basic principles we have discussed thus far apply to both types.

A double-helical DNA molecule with covalently attached ends, as in a plasmid, will have a certain number of "coils." Such coils are analogous to the ones you can introduce in a belt before the buckle is fastened. It is easy to show that the number of coils cannot be changed after fastening the buckle without cutting the belt. In the same way, coils in circular DNA cannot be undone without cutting the polynucleotide strand. From a mathematical point of view, supercoiling can be expressed in terms of three variables as

$$L = T + W. \tag{5.50}$$

L, the *linking number*, is the integral number of times that one DNA strand winds around the other; it is the number of coils in our belt analogy. The *twist*, T, is the number of complete revolutions that one polynucleotide strand makes about the duplex axis (usually the number of base pairs divided by 10.6, the approximate number of base pairs per turn of DNA). T can be positive or negative, depending on the direction of the helix, and it can vary from one part of a molecule to another. W, the *writhe*, is the number of turns that the duplex axis makes about the superhelix axis. Like T, W can be positive or negative.

If the duplex axis of DNA is constrained to lie in a single plane, $W = 0$; there is coiling but no supercoiling, $L = T$, and the twist must be an integral number. From Eqn. (5.50) it is clear that different combinations of W and T are possible for a circular DNA molecule with L, which is a property of the molecule that is constant in the absence of a break in a polynucleotide strand. At equilibrium, one expects a given circular DNA molecule to fluctuate between a variety of conformations, each of which must have linking number L.

As a specific example of DNA supercoiling, consider the circular DNA molecule of the SV40 virus. This molecule is about 5300 base pairs long and is therefore expected to have $L = T \approx 500$ in the absence of supercoiling. The prediction is based on the most energetically favorable number of bases per turn. But in fact, the DNA isolated from SV40 is supercoiled.[14] This probably arises from an untwisted region being present at the end of DNA replication. Such "underwinding" is energetically unfavorable, because the average number of bases per turn is lower than optimal. The conformation of the molecule changes until the lowest free energy state is reached, but regardless of the conformation adopted L is constant (the DNA backbone is not severed). It has been found experimentally that $|W| \approx 25$, so by Eqn. (5.50), $W \approx -25$. The sign of W tells us that the supercoils are negative supercoils, which form to compensate

[14] Under normal salt conditions. The predominant conformation depends on salt concentration and temperature.

for the effects of helix underwinding. Because T is mainly a property of chain length, $T \approx 500$, and $L \approx 475$. Supercoiling increases the elastic strain in circular DNA, just as it does in any other circularized object, for instance, a rubber band! (See Chapters 2 and 3.)

N. Polymerase chain reaction (PCR)

PCR is an extremely useful laboratory process in which double-stranded DNA is replicated rapidly. Under favorable circumstances, a very small amount of starting material can yield a large, readily analyzed product. The technique was developed in the mid 1980s by Kary R. Mullis (1944–) and colleagues at the Cetus Corporation. Mullis, an American, was awarded the Nobel Prize in Chemistry for this work in 1993.

The procedure works as follows. DNA is dissolved in aqueous solution containing a DNA polymerase from a thermophilic bacterium (e.g. *Bacillus stearothermophilus*), polymerase buffer, free nucleotides (dATP, dCTP, dGTP and dTTP, where the "d" means "deoxy"), and oligonucleotide "primers". The primers are short sequences of single-stranded DNA that are designed to bind to either end of the DNA segment of interest. One primer binds one end of one of the complementary DNA strands, and the other primer binds the other end of the other complementary strand (Fig. 5.24).

In a typical PCR experiment, the solution described above is cycled repeatedly through three different temperatures. The first one is usually 95 °C. Thermal energy is used to break the hydrogen bonds and base stacking interactions that stabilize double-stranded DNA. The solution is then rapidly cooled to about 55 °C, at which temperature the primers bind to the complementary sites on the separated strands of the template DNA. The separated strands do not reanneal in the region of the primer binding site because the

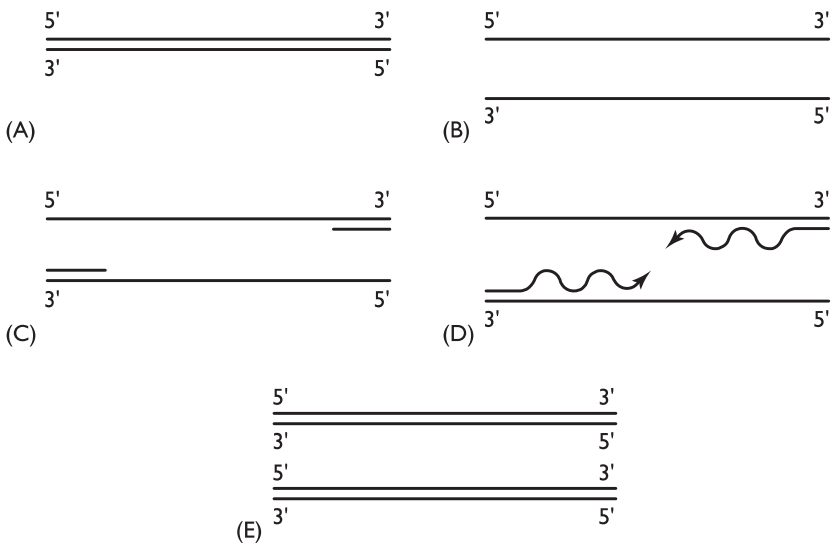


Fig. 5.24 Schematic diagram of PCR. When heated, a double-stranded template DNA (A) melts into two single-stranded molecules (B). If cooling is rapid, the template will not be able to reform. In the excess of complementary oligonucleotide primer, however, binding will occur at temperatures below the primer T_m (C). Increasing the temperature to the optimal value for polymerase activity results in extension of the primer (D). This completes the reaction cycle and yields two double-stranded DNA molecules from one (E). There is an approximately exponential increase in the amount of double-stranded DNA with each cycle.

primers, which are present in great excess, out-compete them. The third temperature of each cycle is usually about 72 °C. At this temperature DNA polymerase activity is high, and it catalyzes the synthesis of a new strand by joining free nucleotide bases to the 3′-end of the primers at a rate of several hundred bases per minute. Each time the thermal cycle is repeated, a strand that was formed with one primer is available to bind the complementary primer, the result being a new two-stranded molecule that is restricted solely to the desired segment of starting material; the region of DNA between the primers is selectively replicated. Further repetitions of the process can produce a billion identical copies of a small piece of DNA in 2–3 h. A well-designed PCR experiment yields the desired product with better than 99% purity.

A question of practical importance to the molecular biologist is: "How long should the oligonucleotide primer be?" There are two main considerations. One is cost: why spend more than is necessary? The other is that although specificity of an oligonucleotide increases with length, size is not necessarily an advantage. In order to answer these questions, let's think about how double-stranded DNA is held together.

Above we said that hydrogen bonds contribute to double-helix stability. How is that? Aren't the bases in single-stranded DNA able to form hydrogen bonds with water? Why should the hydrogen bonds in a double helix be any stronger? In the double helix, the inter-strand hydrogen bonds are not being made and broken constantly, as in single-stranded DNA interacting with water. The hydrogen bond donor or acceptor is, on the average, bonded more of the time in duplex DNA than in single-stranded DNA. Assuming that the enthalpy change of hydrogen bond formation is of roughly the same magnitude in both cases, we should therefore expect a difference in enthalpy between double-stranded DNA and its constituent strands when separated. If heat must be added to "melt" the double helix, then according to Eqn. (2.3) the single strands represent a higher enthalpy state than the double helix at a temperature favoring the double helix. This means that the longer the oligonucleotide primer, the larger ΔH_{d}, where the "d" stands for "denaturation of the double helix." Below T_{m}, base pair formation is energetically favorable with regard to enthalpy. What about entropy?

The double-stranded state has a much lower entropy than the single-stranded one with regard to DNA strands alone. Two particles sticking together is a more orderly situation than two particles floating freely in solution. Formation of a base pair *decreases* the entropy of a strand of DNA and is, therefore, energetically unfavorable. The unfavorability comes not only from requiring that two strands of DNA be in the same place in the sample volume, but also from the restrictions on the shape of both strands that are compatible with helix formation and on the orientation in space of individual bases.

As is often the case in science, thinking will get you only so far. At some stage it becomes necessary to do experiments to find out whether or not the world really is how you imagine it to be. There are two key experimental findings that will help us here. One is that G–C pairs contribute more to duplex-DNA stability than A–T pairs. This cannot be rationalized in terms of base stacking interactions alone, as the surface area of an A–T pair is not appreciably different from that of a G–C pair. The extra stability must come from the G–C pair's extra hydrogen bond. The other empirical finding is that oligonucleotide primers must be about 20 bases long in order for PCR to work well, the exact length depending on the G–C content of the oligo and the temperature at which annealing occurs (usually 55 °C). What this tells us is that we need to form about 50 hydrogen bonds for the favorable enthalpic contribution to the free energy change of double helix formation to exceed the unfavorable entropic contribution.

Now we can see why we will not want to make our oligonucleotide primers too short. We also know that we will not want to make them too long – every afghani, bhat, colon, dollar, euro ... yen, or zaire is dear. But there is another, thermodynamic reason why oligos should not be too long. Assuming a random base sequence in the template DNA strand, the absolute specificity of an oligonucleotide can only increase with length. But if the oligo is long, there will be many, many sites at which partial binding could occur on the template DNA. Moreover, the same oligonucleotide molecule may be able to bind not only to more than one place on the same template at the same time, but also to more than one template molecule! Such a situation, which can be energetically very favorable from an entropic point of view, will promote a huge number of side reactions and yield a very messy PCR product.

O. Free energy of transfer of amino acids

The free energy of transfer is the free energy change on transfer of a compound from one surrounding medium to another, usually, one solvent to another. Suppose we have two perfectly immiscible solvents (ones that do not mix) in the same beaker. There is an interface, with one solvent on one side and the other on the other. There is a large free energy barrier to mixing; this is what it means for the solvents to be immiscible. Now, if a solute is dissolved in one of the solvents, when the three-component system comes to equilibrium, solute will be found in the second solvent as well, if the solute is soluble in it. You will recognize this as a means of doing extraction of chemical compounds with organic solvents. This phenomenon is often grouped together with freezing point depression, boiling point elevation and osmosis, but strictly speaking it is *not* a colligative property. We are interested in it here because it will help us to have a better understanding of the solubility and thermodynamic

stability of biological macromolecules, not because of any particular relationship to colligative properties. A specific example will help to make headway in understanding.

The solubility of phenylalanine (Phe) in water at 25 °C on the molality scale is 0.170 mol (kg solvent)$^{-1}$; in 6 M urea it is 0.263 mol (kg solvent)$^{-1}$. Urea improves the solubility of hydrophobic side chains in aqueous solution, making urea a good chemical denaturant of proteins. Using the given information, one can calculate the standard state free energy of transfer of Phe from water to aqueous urea solution. To keep things simple, we'll assume that the activity coefficient of Phe is approximately the same in both media.

This situation can be pictured as follows (compare Fig. 2.3). In one process, Phe dissolved in water is in equilibrium with crystalline Phe; the solution is saturated. In another process, Phe dissolved in urea solution is in equilibrium with crystalline Phe. Both of these processes can be studied experimentally and the solubilities can be measured. In a third process, which is a sort of thought experiment, Phe in one solution is in equilibrium with Phe in the other solution. We construct a notional boundary between these solutions, and require that it be permeable to Phe but *not* to water or urea. There will be a net flow of Phe across the boundary until equilibrium is reached. The fourth "process" in the thermodynamic cycle is just solid Phe in equilibrium with solid Phe.

In mathematical terms,

$$\mu_{\text{water, sat.}} - \mu_{\text{solid}} = 0 \tag{5.51}$$

$$\mu_{\text{water}, a=1} - \mu_{\text{water, sat.}} = -RT \ln a_{\text{water, sat.}} = +1050 \text{ cal mol}^{-1} \tag{5.52}$$

$$\mu_{\text{urea, sat.}} - \mu_{\text{solid}} = 0 \tag{5.53}$$

$$\mu_{\text{urea}, a=1} - \mu_{\text{urea, sat.}} = -RT \ln a_{\text{water, sat.}} = +790 \text{ cal mol}^{-1} \tag{5.54}$$

$$\mu_{\text{solid}} - \mu_{\text{solid}} = 0. \tag{5.55}$$

The energy barrier between saturation and unit activity is greater in water than urea because the solubility of phenylalanine is lower in water than urea. The difference between Eqns. (5.52) and (5.54), which is what we set out to find, is $\mu_{\text{urea, } a=1} - \mu_{\text{urea, sat.}} - (\mu_{\text{water, } a=1} - \mu_{\text{water, sat.}}) = \mu_{\text{urea, } a=1} - \mu_{\text{solid}} - (\mu_{\text{water, } a=1} - \mu_{\text{solid}}) = \mu_{\text{urea, } a=1} - \mu_{\text{water, } a=1} = [\mu^\circ_{\text{urea, } a=1} + RT \ln(1)] - [(\mu^\circ_{\text{water, } a=1} + RT \ln(1)] = \mu^\circ_{\text{urea, } a=1} - \mu^\circ_{\text{water, } a=1} = 790$ cal mol^{-1} − 1050 cal mol^{-1} = −260 cal mol^{-1}. This is the standard state driving force for transfer of Phe from saturated water to saturated 6 M urea. We can see from the sign of the chemical potential that the transfer is spontaneous; this is exactly what is expected from solubility data.

Though this example has involved a complete amino acid, there is in principle no reason why the experiment could not be done with

Table 5.5. | *Thermodynamics of transfer at 25 °C from nonpolar solvent to water of various chemical groups*

Chemical group	ΔG_{tr} (cal mol^{-1} Å^{-2})	ΔH_{tr} (cal mol^{-1} Å^{-2})	ΔC_p (cal K^{-1} mol^{-1} Å^{-2})
Aliphatic: –CH_3, –CH_2–, CH	+8	−26	0.370
Aromatic	−8	−38	0.296
Hydroxyl	−172	−238	0.008
Amide & amino: –NH–, NH_2	−132	−192	−0.012
Carbonyl C: C=	+427	+413	0.613
Carbonyl O: =O	−38	−32	−0.228
Thiol and sulfur: –SH, –S–	−21	−31	−0.001

The data are from Ooi, T. and Oobataka, M. (1988) *J. Biochem.* 103, 114–120.

various small organic molecules "components" of the larger chemical. Comparison of the thermodynamic data with structural information would then provide clues to the thermodynamics of transfer of individual chemical groups. Table 5.5 gives thermodynamic values for the transfer of various chemical groups from nonpolar organic solvent to water.

There are at least two practical lessons we can draw from the above analysis. One is that the hydrophobic surface of phenylalanine, or indeed of any amino acid side chain which interacts with itself in the solid state by means of hydrophobic interactions, forms more favorable interactions with urea than water. Looked at another way, urea could be said to weaken hydrophobic interactions. Which leads to point number two. Empirical studies have shown that urea is a good chemical denaturant of proteins. We mentioned something about this in the context of dialysis but did not elaborate. The example of this section helps to rationalize the empirical finding. We know from X-ray analysis of the folded states of proteins that, although some hydrophobic side chains do appear on the protein surface, the core is mainly hydrophobic. In the presence of urea, where the solubility of the hydrophobic side chains is considerably increased relative to the absence of urea, the unfolded state of the protein is correspondingly more thermodynamically favorable. This fact can be used to investigate protein stability, as we shall see below and in Chapter 6.

P. | Protein solubility

Here we are interested not so much in solubility of a substance *per se* but in solubility of proteins and protein–nucleic acid complexes. Our approach is qualitative and practical rather than quantitative and theoretical.

A protein molecule is a very complex polyion; there are numerous ionizable groups and a variety of pK_as. The solubility of a protein in aqueous solution will depend strongly on ionic strength and pH (Fig. 5.25). This is of the greatest practical significance for the choice of techniques that one might use to study a protein molecule. For instance, nuclear magnetic resonance (NMR) spectroscopy is a very high-resolution structural technique, making it valuable for protein structure determination and other aspects of biochemical research. But NMR is also an extremely insensitive technique, meaning that a very large concentration of protein is needed for a good signal-to-noise ratio, on the order of 1 mM or higher. At the same time, NMR structural studies generally require a relatively low rate of exchange of labile protons in tryptophan and tyrosine side chains, a pH-dependent phenomenon (see Chapter 8). In other words, the protein must not only be highly soluble, it must be soluble in a suitable pH range (near neutral pH). These requirements (and others!) restrict which protein structures can be studied by NMR.

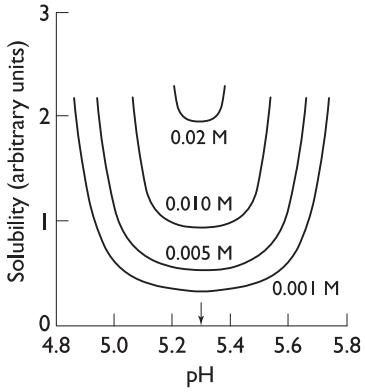


Fig. 5.25 Protein solubility. This depends not only on the net charge of the protein, which varies with pH, but also on ionic strength and temperature. The data shown are for a milk protein called β-lactoglobulin at different concentrations of NaCl. Solubility is very low at pH 5.3, the approximate isoelectric point of the protein. At this pH, the solubility of β-lactoglobulin increases exponentially with increasing ionic strength (solubility $\approx 0.255e^{101pH}$). The data are from Fox and Foster (1957).

Here's an example of the effect of pH on protein solubility. The PTB domain of chicken tensin is highly soluble at pH 2, where its net charge is about +30. The net charge on each PTB molecule at acidic pH is so great that electrostatic repulsion inhibits the formation of protein aggregates, as long as the ionic strength is low. At pH 7, by contrast, the net charge on the PTB domain is 0 and it is not very soluble at all. This pH is the so-called *isoelectric* point, or p*I*, of the PTB domain. The isoelectric point of a protein depends primarily on *amino acid composition*; more specifically, the number of amino acids with ionizable side chains, and to some extent the location of each ionizable residue in the folded protein structure. If the number of basic side chains is relatively large and the number of acidic side chains relatively small, as with hen egg white lysozyme, the isoelectric point of the protein will be high, and the net charge is likely to be positive through most of the usual pH range (2–12). All that prevents a protein from aggregating at its p*I* is its solubility, which depends on the actual number of charged groups present, even if the sum of charges is 0. Charge properties can be used to purify a protein. For instance, recombinant PTB domain can be separated from some bacterial proteins by adjusting the cell lysate to pH 7. Separation of the precipitate effectively isolates recombinant PTB domain from the many bacterial proteins which remain soluble at neutral pH.

In general, the situation with protein solubility is more complex than we have made it seem thus far. At a given pH, a typical protein will have both positive and negative charges. Depending on the location of the charges on the protein surface, if the ionic strength is low, proteins can interact with each other by electrostatic attraction. It is often found therefore that *the solubility of a protein at low ionic strength increases with salt concentration*. This phenomenon is known as "*salting in*," and because it depends on the protein being charged, the effect is least pronounced at the isoelectric point. By contrast, at high ionic strength the protein charges are strongly shielded. Electrostatic

repulsion is negligible. Solubility is reduced. This effect is known as "*salting out*." The shape of a solubility curve with ionic strength is thus, roughly speaking, U-shaped (Fig. 5.25). Salting out is thought to arise from the screening of charges and from the numerous interactions between the salt ions and water resulting in a *decrease* in the water molecules available to solvate the protein.

Salting out is a useful means of purifying proteins. For instance, ammonium sulfate is often used to purify antibodies. Below a certain ionic strength, antibodies and some other proteins are soluble, but many other proteins are insoluble. The insoluble proteins can be removed from solution by centrifugation. Above a certain ion strength, the antibodies themselves precipitate. They can be separated from the rest of the solution and subjected to further purification. A similar procedure can be used to purify many different proteins. Once a protein is sufficiently pure, it is sometimes possible to crystallize it by dissolving it in a salt solution near the solubility limit of the protein. From a thermodynamic point of view, crystallization occurs because the crystalline state of the protein has a lower Gibbs free energy than the solution state.

Q. Protein stability

This section is on cooperative and reversible order–disorder transitions in proteins. It builds on several of the several previous sections, including those on DNA and PCR. A key difference between protein stability and duplex DNA stability is the size of ΔC_p between the ordered and disordered states: in proteins it is relatively large, in DNA relatively small. As we shall see, the magnitude of ΔC_p can have a marked effect on the thermostability of a protein.

When protein folding/unfolding is cooperative, effectively only two states are populated at equilibrium: the folded (native) state and the unfolded (denatured) state. The transition occurs over a relatively narrow range of the independent variable, be it temperature, pH or chemical denaturant concentration. In such cases, the equilibrium can be represented as

$$\mathrm{F} \Leftrightarrow \mathrm{U}. \tag{5.56}$$

The equilibrium constant (Eqn. (4.11)) is then

$$K_{eq} = [\mathrm{U}]/[\mathrm{F}]. \tag{5.57}$$

Note that Eqns. (5.56) and (5.57) are at least consistent with the idea that all the information required for a protein molecule to fold into its native form will be present in the amino acid sequence. The free energy difference between the folded state of a protein and its unfolded state is independent of the path! Regardless of the process by which a protein folds – in the cell or in a test tube – the free energy difference between folded and unfolded forms is the same (given the same temperature, ion concentrations, pH, etc.). But is a

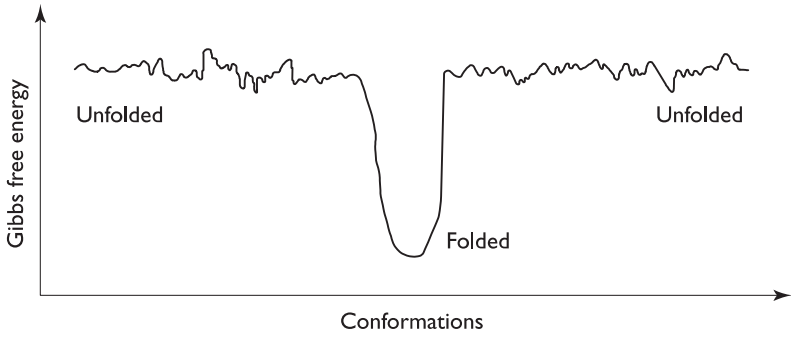


Fig. 5.26 Free energy profile of a "well-behaved" small protein. There are only two stable states: the folded state and the unfolded state. The number of unfolded conformations is vastly greater than the number of folded conformations. Although the unfolded conformations differ in energy, these differences are relatively small. Therefore the collection of unfolded conformations can be thought of as a single state. The energy difference between the unfolded state and the folded state is comparatively very large.

catalyst needed to get the reaction to proceed on a biologically relevant time scale?

Early attempts to give a thermodynamic description of reversible protein denaturation and coagulation appeared in the 1920s and 1930s in the work of American physical biochemists Alfred Ezra Mirsky (1900–1974), Mortimer Louis Anson (1901–1968), and Linus Carl Pauling.[15] In the 1950s, Rufus Lumry (1920-) and Henry Eyring (1901–1981), also both Americans, provided a more substantial mathematical analysis of reversible protein denaturation. This experimental and theoretical work modeled the folded state of a protein as corresponding to a (local) minimum of free energy, also known as an *energy well*, when conditions favor the folded state (Fig. 5.26). An important lingering difficulty, however, was the generality of the applicability of the thermodynamic description. Did it work for some proteins and not others? Did it apply to proteins with disulfide bonds? Christian Boehmer Anfinsen's investigations of the reversible denaturation of proteins showed conclusively that the native state of a protein with disulfide bonds could be recovered spontaneously, even when the disulfides were not formed in the denatured protein. This led to the general acceptance of the "thermodynamic hypothesis" for the folding of proteins, according to which attainment of the native structure rests solely upon the amino acid sequence. Anfinsen (1916–1995), an American, was awarded the Nobel Prize in Chemistry for this work in 1972. Since then, and particularly since the late 1980s, the goal of working out the structural basis of protein folding and thermostability has been pursued with considerable intensity throughout the world.

We have already discussed protein denaturation (in Chapters 2 and 3) in the context of illustrating the physical meaning of H and S and showing the utility of a van't Hoff analysis of equilibrium

[15] Pauling (1901–1994) was awarded the Nobel Prize in Chemistry in 1954 for his work on protein structure. His model of DNA structure, which had the bases pointing outwards, was no longer tenable after publication of the famous work of Watson and Crick. Something worth remembering: Nobel laureates are unusually accomplished rational animals, but they are capable of error. At the same time, though, it is fair to say that no one's discovered anything who hasn't also made a mistake.

constant data. Now let's consider protein denaturation in terms of free energy. At constant temperature, Eqn. (4.2) becomes

$$\Delta G^\circ_d = \Delta H^\circ_d - T\Delta S^\circ_d, \tag{5.58}$$

where as before the subscript signifies "denaturation." ΔG°_d is the difference in Gibbs free energy between the unfolded state and the folded state of the protein. In most cases, the energy of the unfolded state is measured relative to the energy of the folded state; i.e. the folded state is the reference state. There are two main reasons for the convention: the folded state has the least ambiguous conformation, and more often than not *equilibrium* studies investigate transitions *from* the folded state *to* the unfolded state; the folded state is often the starting state (not true of kinetic protein refolding experiments!). ΔG°_d alone tells us nothing about the relative magnitudes of ΔH°_d or ΔS°_d; an *infinite* number of combinations of these thermodynamic functions would be consistent with a given value of ΔG°_d. Of course, many of these combinations of ΔH°_d and ΔS°_d will have little or no physical meaning for the system under study, and only one combination will actually describe the system under study. In order to fix the values, we'll have to do at least one more experiment. It has been found that ΔH°_d and ΔS°_d for proteins can be very large in comparison with ΔG°_d. For instance, it is common for the maximum value of ΔG°_d for a protein in solution to be about 15 kcal mol^{-1}, and for ΔH°_d at the denaturation temperature to be more than an order of magnitude greater. ΔG°_d for proteins is thus a delicate balance of ΔH°_d and ΔS°_d.

At the *melting temperature*, also called the *heat-denaturation temperature*, the fraction of molecules in the folded state equals that in the unfolded state; the free energy difference between them, ΔG°_d, is 0. This leads to Eqn. (3.21) and enables you to calculate the entropy of unfolding from measurement of ΔH°_d. Including the temperature dependence of ΔH and ΔS explicitly, Eqn. (5.58) becomes

$$\Delta G^\circ_d(T) = \Delta H^\circ_d(T_r) + \Delta C_p(T - T_r) - T[\Delta S^\circ_d(T_r) + \Delta C_p \ln(T/T_r)], \tag{5.59}$$

where the subscript "r" means "reference." $\Delta G^\circ_d(T)$ is not $(\Delta G^\circ_d \times T)$ but ΔG°_d evaluated at temperature T. As an example, suppose that our reference temperature is 25 °C and that both ΔH°_d and ΔS°_d are known at this temperature. What is ΔG°_d at 35 °C? If ΔH°(25 °C) = 51 kcal mol^{-1}, ΔS°(25 °C) = 100 cal mol^{-1} K^{-1}, and ΔC_p = 1500 cal mol^{-1} K^{-1}, then ΔG°(35 °C) = 51 kcal mol^{-1} + 1500 cal mol^{-1} K^{-1} × (308 K − 298 K) − 308 K + [100 cal mol^{-1} K^{-1} + 1500 cal mol^{-1} K^{-1} + ln(308 K/298 K)] = 20 kcal mol^{-1}. $\Delta G^\circ_d(T)$ is known as the *stability* of a protein. It tells you how much energy must be expended (more specifically, the minimum amount of work that must be done) to unfold the protein at a given temperature. A plot of $\Delta G^\circ_d(T)$ versus temperature (or any other independent variable, e.g. pH or concentration of chemical denaturant) is called a *stability curve*.

The stability curve as a function of temperature resembles a parabola and has a peak at which ΔG°_d is a maximum. It can be shown

(using Eqn. (5.58) and a little calculus) that at this temperature, called the *temperature of maximum stability*, $\Delta S^\circ_d = 0$ (compare Fig. 4.4). That is, the stability of the folded state of a protein is a maximum when the entropy of the folded state and surrounding solution is equal to the entropy of the unfolded state and the surrounding solution. At this temperature, which is often 40 or 50 K below the heat-denaturation temperature, enthalpic interactions alone hold the folded state together. Just *below* T_m (for heat denaturation), ΔG°_d is positive (if the folded state is the reference state). On the average, unfolding will not occur spontaneously, because $\Delta G^\circ_d > 0$.[16] To bring about unfolding by a further temperature increase, we expect ΔH°_d to be positive; this is roughly the energy required to disrupt non-covalent interactions in the folded state. We also expect ΔS°_d to be positive, as the polypeptide chain will be much more disordered in the unfolded state than in the folded one, and we don't expect the order of the surrounding solution to increase with temperature. But *on the balance*, $\Delta G^\circ_d > 0$ below T_m. Above T_m, the balance is shifted towards the entropy, $|T\Delta S^\circ_d| > |\Delta H^\circ_d|$, and there is net unfolding of the protein.

ΔC_p plays a key role in protein stability. Both the enthalpy change and the entropy change of denaturation depend on ΔC_p, so the free energy change depends on ΔC_p. Figure 5.27 shows how the Gibbs free energy difference between the unfolded and folded states changes as the magnitude of ΔC_p changes. If $H^\circ_d(T_r)$ and $S^\circ_d(T_r)$ are held constant, *decreasing* ΔC_p *increases* the breadth of the stability curve, and *increasing* ΔC_p *decreases* the breadth; all the curves intersect at the reference temperature. This tells us that if a protein has a small heat capacity change on unfolding, it is likely to have a relatively high transition temperature, and this is exactly what is observed experimentally. By contrast, when ΔC_p is relatively large, the stability curve becomes sharply peaked and can cross the temperature axis in more

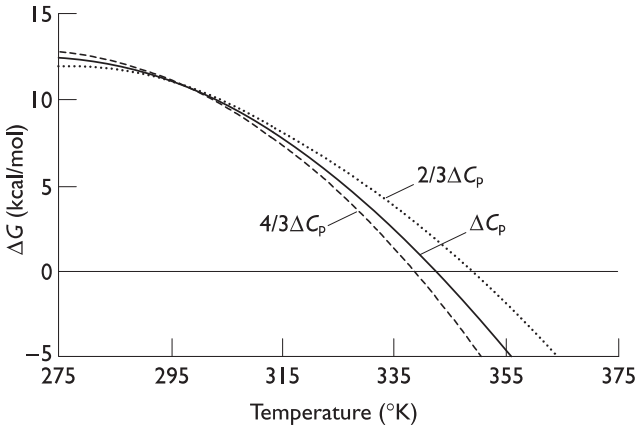


Fig. 5.27 Protein stability curves. Each one is a plot of ΔG versus T (or some other independent variable). This gives the free energy difference between the unfolded state and the folded state (the minimum amount of work that must be done to induce a transition from the folded state to the unfolded state). The curvature in ΔG v. T arises from the positive and relatively large ΔC_p of protein unfolding. The stability curve crosses the T-axis at two points, the heat- and cold-denaturation temperatures. In the figure only the heat denaturation temperatures are seen.

[16] This statement needs some qualification. In fact, unfolding *can and does* occur spontaneously when $\Delta G > 0$, but not in greater abundance than spontaneous refolding of unfolded protein (see Chapter 6). The more positive ΔG, the less probable spontaneous unfolding. The situation is just the opposite when $\Delta G < 0$.

than one place in the experimentally accessible range (when the solvent is in the liquid state).

The second intersection of ΔG°_d with the temperature axis, which occurs well below the heat-denaturation temperature, is known as the *cold-denaturation temperature*. The mathematical form of the stability curve, which is based on solid experimental evidence of *heat*-denaturation, suggests that protein unfolding can be induced by heating or, strange as it may seem, by cooling. This prediction has been confirmed by experimental studies in a number of cases, greatly underscoring the value of good mathematical modeling of experimental results for prediction of the behavior of biochemical systems. Cold denaturation seems rather counterintuitive. For in order to melt a crystal, one expects to have to *add* heat, in which case $\Delta H > 0$. The entropy change on protein unfolding, $\Delta S^\circ_d = \Delta H^\circ_d / T_m$, is therefore positive, in accord with intuition. By contrast, in cold denaturation $\Delta H^\circ_d < 0$! It follows that $\Delta S^\circ_d < 0$ for cold denaturation. Weird! Some peptide aggregates are known to exhibit cold denaturation on heating from room temperature. In other words, the peptides tend not to be aggregated in aqueous solution at 25°C, but heating the solution leads to spontaneous gelation, a type of peptide aggregation which resembles the condensation of hydrophobic residues in the folding of a small protein. In all these cases, ΔG°_d can pass through 0 more than once because both ΔH and ΔS depend on T.

So far we have been discussing ΔG as a function of temperature. There are other independent variables we could consider, for example, pH and chemical denaturant concentration. Let's look at the latter first. As chemical denaturant is added to a protein solution, the folded state becomes destabilized *relative to* the unfolded state, and the protein unfolds. At the so-called midpoint concentration of denaturant, $\Delta G = 0$, and the fraction of molecules in the folded state is equal to the fraction in the unfolded state. The superscript on G has disappeared because now we are considering the protein in solution in the presence of denaturant. Note how the midpoint concentration of chemical denaturation closely resembles the transition temperature of thermal denaturation. In pH denaturation, either acid or base is added to the protein solution to induce unfolding. At the midpoint pH, half of the molecules are in one state and half are in the other, and again, $\Delta G = 0$.

Chemical denaturation and pH denaturation are such common forms of studying protein stability that further details can be provided here. The stability of the folded state of a protein in the presence of a chemical denaturant is often modeled as

$$\Delta G_d = \Delta G^\circ_d - mc, \tag{5.60}$$

where c, the concentration of denaturant (usually in molar units), is the only independent variable, and m is a parameter that depends on temperature, pH, buffer, and – is it surprising? – the protein. Note that at the concentration midpoint, which is determined experimentally, $\Delta G^\circ_d = mc$. So if ΔG°_d is known independently, for

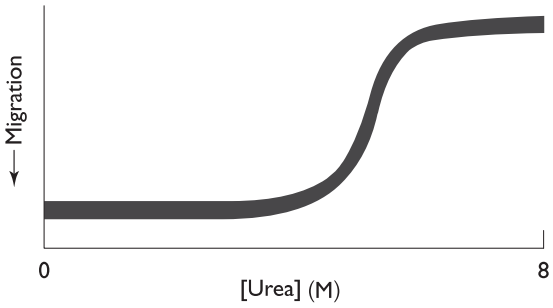


Fig. 5.28 Urea denaturation gel electrophoresis. Structural transitions in protein can be studied by polyacrylamide gel electrophoresis. The horizontal axis is a linear gradient in urea concentration. When the concentration is sufficiently high, the protein unfolds. This is detected as a change in the mobility of the protein in the gel matrix. Compact folded proteins generally migrate faster in the gel than extended unfolded proteins.

example, by calorimetry, m can be determined experimentally. Eqn. (5.60) "works"; it can be used to study protein denaturation. But it should be added that the physical meaning of m is not entirely clear, though it seems to correspond roughly to the increase in surface area of the protein in contact with the solvent upon unfolding. Figure 5.28 illustrates the effect on protein stability of changing the denaturant concentration. The denaturant is urea, and a change in the conformation of the protein is detected as a change in electrophoretic mobility. The more compact the protein, the higher the mobility in a network of non-migrating polymers. An unfolded protein migrates more slowly than a folded one. The model of protein stability outlined here is plausible for a number of reasons. At the same time, however, it says nothing at all about what the denaturant does to make the protein unfold. Moreover, experiments show that, although in many cases the dependence of ΔG_d on denaturant concentration is approximately linear, in some cases it is distinctly non-linear. The physical meaning of such non-linear behavior is usually uncertain.

As we have said, a change in pH can also cause a protein to unfold. Such denaturation usually occurs for two reasons. One is that a charged side chain can be partly buried, giving it an anomalous pK_a, and a change in the ionization state of the side chain can destabilize folded structure. Another is that at extremes of pH, the net charge on the protein can be very large, resulting in an unusually large destabilizing contribution to the overall thermostability. The change in stability of the folded state, $\Delta(\Delta G_d)$, varies with a change in pH as

$$\Delta(\Delta G_d)/\Delta(\mathrm{pH}) = 2.3RT\Delta Q, \tag{5.61}$$

where ΔQ is the difference in number of bound protons between the unfolded state and the folded state, Δ(pH) corresponds to a difference in $[H_3O^+]$, and 2.3 is a conversion factor related to logarithms. Using these hints, can you derive Eqn. (5.61)?

It has frequently been observed that point mutations in proteins lead to relatively large changes in ΔS° and ΔH° but a relatively small change in ΔG°. The changes in ΔS° and ΔH° are often difficult to rationalize in terms of changes in protein structure as assessed by NMR spectroscopy or X-ray crystallography. This "cancellation" of

changes in the enthalpy and entropy terms is known as *enthalpy–entropy compensation*. Apparently, subtle differences in structure and solvation can have significant thermodynamic consequences. This points up both the remarkable plasticity of the folded state of a protein and the still shallow understanding of the microscopic origins of macroscopic properties of materials and systems. We'll return to this topic in Chapter 9.

Now let's bring a bit of biology into the picture. We wish to cover two subtopics to round out this section: the engineering of enzymes to enhance their thermostability without altering specificity and the role of stability in protein degradation.

Protein engineering: enzymes are biological catalysts (Chapter 8). The ability to produce massive quantities of an enzyme by recombinant DNA technology has made it feasible to consider the use of enzymes in biomedical, chemical, and industrial applications. Often, though, the physical environment of the enzyme in a practical application will be different from the environment in which it is found in nature. An enzyme can be engineered by standard molecular biological techniques to tailor properties to specific applications. For instance, a way in which the stability of an enzyme can be increased is to decrease the disorder of its unfolded state. This will lead to a substantial decrease in ΔS_d but in some cases effectively no change in ΔH_d. The change can be brought about by replacing a Gly residue with any other residue type. Some Gly residues, for instance, ones in turns, won't work well for this purpose, because replacement has too large an impact on the structure of the folded state. In general, though, Gly residues make the polypeptide backbone very flexible, while the amino acids with side chains restrict bond rotations in the backbone. Pro residues allow practically no backbone flexibility, and these can be introduced into turns. Yet another approach is to add disulfide bonds. These link different parts of the polypeptide chain and reduce its mobility in the unfolded protein (see Chapter 6, Section C). A complementary approach to stabilization of a protein is to increase the enthalpy of the folded state. This can lead to a substantial increase in ΔH_d but effectively no change in ΔS_d. Unfavorable electrostatic interactions in the folded state can be replaced by favorable ones, and negatively charged side chains (particularly that of Asp) can be placed at the beginning of an α-helix to interact favorably with the helix dipole. Amino acid substitutions can be made within helices to increase their strength, and within the protein core to increase the hydrophobic surface area. All such changes can lead to a protein of increased stability.

Protein degradation: as we have said, protein stability can be measured by ΔG_d° or K_{eq}. Because K_{eq} measures the ratio of the forward and reverse rate constants (Chapter 8), when the folded state of a protein is stable (i.e. when ΔG_d° is large, if the folded state is the reference state), the rate of folding must be greater than the rate of unfolding. When the folded state is the more stable one, there will still be some molecules in the unfolded state, even though that

proportion will be small except in the region of the transition (see Chapter 6). When $K_{eq} = 1$, not only is there no free energy difference between the states, there is an equimolar mixture of folded state and unfolded state. In other words, the bottom of the energy well is at the same level as the ground surrounding the well! The proportion of molecules in one state or the other changes as the conditions are adjusted. Even when the stability of the folded state is relatively large, some unfolded proteins will be around, though the relative proportion of unfolded molecules could be 10^{-6} or less. A question we'd like to be able to answer is this: because the turnover rate of endogenous protein in an organism is high, i.e. because the body is constantly recycling its protein, (see Chapter 9), does the body clear protein by having a specialized degradation protein bind to and digest unfolded proteins? Does the degradation of an unfolded protein then diminish the population of folded proteins by mass action and thereby stimulate the synthesis of replacement protein? Or, does the body make "degradation proteins" that actively unfold proteins at random? If the former, it is thermal energy and the specific solution conditions of the body (salt concentration, pH, etc.) that play some role in clearing proteins. If the latter, then energy must be supplied to clear proteins, since work must be done to denature a stable folded protein. In fact, there appear to be proteins that can unfold and degrade stable, native proteins in the presence of ATP. One such protein in bacteria, ClpA, is a member of the Clp/Hsp100 "chaperone" family. Are there corresponding proteins in mammals? Another question we'd like to be able to answer is, if the body continually recycles protein, it must continually make it, and because proteins are synthesized on ribosomes from mRNA templates, genes must continually be transcribed, and if mutations in genes can lead to pathological proteins, and if mutations accumulate as the body ages, does the body somehow "program" its own death?

R. Protein dynamics

In an earlier chapter we described the folded state of a protein as an organic crystal. Indeed, if this were not so, it probably would not be possible to crystallize proteins! More importantly, without a relatively fixed geometry in the catalytic site, how could an enzyme carry out a specific function? Such considerations might give the false impression that the folded state of a protein has a rigidly fixed structure. Instead, folded states of proteins, though sturdy and crystal-like, are nevertheless flexible, and they exhibit many very rapid small-scale fluctuations. Evidence for dynamical behavior which arises from thermal motion tells us that the native structure of a protein is a large ensemble of similar and rapidly inter-converting conformations that have the same or nearly the same free energy. As we shall see, structural mobility in the native state has crucial functional significance.

First, let's see how this section links to the one on protein stability. The melting temperature of a protein showing two-state behavior depends on the balance of ΔH_d and ΔS_d. For given values of these thermodynamic functions, if the folded state of a protein were extremely rigid, then the change in entropy on unfolding would be large, and the protein would never fold; T_m would simply be too low (see Eqn. (3.21)). And if the entropy of the unfolded state of a protein were not very different from the entropy of the folded state, ΔS_d would be small, and $\Delta H_d/\Delta S_d$ would be so large that the protein would never unfold. This could be disastrous for the cell if a protein became modified in such a way as to be pathogenic.

Now, you might find it interesting that protein flexibility is demonstrated by the very fact that proteins can be crystallized! There are two points we can make. One is that high-resolution X-ray diffraction data provide valuable information on the motions of atoms more massive than hydrogen. Modeling of the protein structure is a matter of fitting a molecule with the known covalent constraints to an electron density map. The map does not reveal precisely where the center of mass of an atom will be, but only a volume of space where an atom is likely to be found. Analysis of such data shows that some atoms in a protein move very little while others move a great deal – when the protein is folded and in a crystal. There is another way in which protein crystals reveal that such motions exist, and it shows that the motions have physiological relevance. X-ray studies of the oxygen transport and storage proteins hemoglobin and myoglobin show that there is no obvious route for O_2 to move from the solvent to the binding site; oxygen takes up space. One concludes that O_2 (and CO_2) binding and release depend on fluctuations in structure known as "*breathing motions*."

The length and time scales of such motions depend on free energy differences relative to the minimum free energy structure. There are three basic types of dynamical motion: *atomic fluctuations*, *collective motions* and *triggered conformational changes*. Atomic fluctuations occur on a time scale on the order of picoseconds and are relatively small in magnitude, while conformational changes are typically much slower and larger. X-ray analysis of the active site cleft of hen lysozyme, for example, shows that some of its atoms move by $\sim$1 Å on substrate binding. Small but significant.

Other experimental methods that reveal the motions of folded proteins are NMR spectroscopy and hydrogen exchange. NMR can be used to measure the rate of 180°-flipping of the ring of a Phe or Tyr side chain about the C_β–C_α bond. The rate varies, but it generally falls in the μs–s range. Recent developments in NMR data collection and analysis permit a more general analysis of polypeptide backbone and amino acid side chain dynamics. NMR can also be coupled with the exchange of labile protons in the polypeptide backbone of a protein to gain information on protein dynamics. Exchange of such protons is temperature-dependent for two reasons: the stability of the native state of a protein varies with temperature, as we

saw above, and the exchange rate depends on temperature (as well as pH). These experiments involve D_2O, and deuterium is exchanged in for solvent hydrogen. The approach is particularly attractive from the point of view of experimental design and analysis of results, because hydrogen and deuterium are chemically identical but have completely different NMR characteristics (see Chapter 8).

S. Non-equilibrium thermodynamics and life

At some point in your study of thermodynamics you may have wondered: if the First Law requires the total energy of the universe to be constant, and the Second Law requires that every process be accompanied by an increase in the entropy of the universe, then how is life possible at all? Do the tremendously complex forms of matter we call living organisms violate the laws of thermodynamics? Clearly, the answer must be no, if the laws of thermodynamics as formulated actually do describe our universe.

In Chapters 4 and 5 we have used a combination of the First and Second Laws to look at a number of biological processes at equilibrium. We have seen how useful the Gibbs free energy function is for describing these processes. Although aspects of biochemistry can be described in this way, we should always remember that no living organism is at equilibrium! This holds not only for the organism as a whole but each of its cells. Moreover, it applies to every last bacterium inhabiting the cosy environment of your gut. Important for us, a non-equilibrium process is by definition irreversible (though possibly not completely irreversible)! Let's look at this topic somewhat more carefully.

An *in vitro* biochemical reaction is a closed system (Chapter 1). As such, it will change until equilibrium is reached. A living organism, on the other hand, be it an amoeba, a bombardier beetle, or a wildebeest, is an open system. An organism is therefore *never* at equilibrium. An organism takes in high-enthalpy and low-entropy compounds from its surroundings, transforms them into a more useful form of chemical energy, and returns low-enthalpy and high-entropy compounds to its surroundings. By means of such energy flows, living organisms degrade the quality of the energy of the universe. *Non-equilibrium systems "dissipate" the useful energy of the universe*.

Energy flow through an organism is like water flow through a channel. (But this does not mean that energy is a material particle or a collection of particles!). The rate of flow through an organism in adulthood is approximately constant, a situation known as *steady state*. A steady-state system changes continually, but there is no *net change* in the system – its physical makeup, the amount of matter present. (What changes occur in the brain when an adult learns something new and remembers it?) Steady state in an open system is the analog of equilibrium in a closed system. A steady inward flow of energy is the most stable state an open system can achieve. As

depicted in Fig. 2.5B, if the inward flow of energy differs from the rate at which energy is consumed, a change in weight occurs. A voracious reader is clearly something very different from a voracious eater.

Must a living organism be a non-equilibrium system? If it were not, it could not do useful work. An equilibrium system cannot do useful work. This is because at equilibrium, there is no free energy difference between reactants and products. An equilibrium process cannot be directed. It is "rudderless." The schematic diagrams of earlier chapters highlighted the machine-like qualities of living organisms; indeed, there are many similarities. For instance, both organisms and machines are made of matter, and the processes carried out by both are, at least to some extent, irreversible. Organisms and machines can do work. Because body temperature is not very different from and often greater than the temperature of the surroundings, an organism can do very little work by means of heat transfer, practically none. Instead, organisms do work by taking in free energy from their surroundings – food. Like machines, organisms "wear out" with use. Excessive physical activity can damage the body. Extensive bicycling on insufficient conditioning, for example, can result in a damaged nervous system, with neurons in the leg tingling constantly, possibly until death. Machines, by contrast, are basically "static" structures. The plastic in a computer keyboard, for example, is not changing dramatically from moment to moment. The keyboard molecules are not simultaneously being degraded by enzymes and replaced by newly synthesized ones. Living organisms are different. They need free energy because they must renew themselves to live. Their proteins are constantly being destroyed and new ones must be made to take their place. DNA is constantly being replicated. Certainly, organisms display machine-like qualities. But organisms are different from lifeless machines. A living organism cannot be at equilibrium.

Where does the thermodynamic irreversibility of a living organism come from? If many individual biochemical reactions are reversible, at what length scale does irreversibility set in? What is the microscopic origin of irreversibility in biological organisms? These are hard questions! We do not pretend to answer them fully here. There are numerous contributions to the overall fact of irreversibility, but a particularly important one is non-productive hydrolysis of ATP. As we have seen, ATP is hydrolyzed spontaneously in water. If hydrolysis is not coupled to a metabolic reaction, the energy released will go off as heat – irreversibly. There are also three "irreversible" steps in the metabolism of glucose to pyruvate. These occur between glucose and G6P, F6P and FDP and PEP and pyruvate. This turns out to be extremely important for cellular function, for it is really only at the irreversible steps of a process that "machine-like" control can be exerted: it permits regulation of the speed of the reaction. Such regulation is of considerable importance to reactions that occur downstream.

Lastly, non-equilibrium systems present a number of problems for the quantification of thermodynamic functions. The First Law has been verified experimentally and quantitatively for living organisms. It's harder to do this for the Second Law for reasons outlined above. Entropy, free energy and chemical potential cannot be measured for non-equilibrium systems. There is nevertheless a way of connecting a non-equilibrium system with something more amenable to study and analysis, and that is the internal energy, *U*. Suppose we wish to measure the internal energy of a non-equilibrium system. This can be done by isolating the system and waiting for it to come to equilibrium. Because the system is isolated, the internal energy will be the same at equilibrium as in any non-equilibrium state. If *U* of the equilibrium state is then measured with respect to some reference value, then *U* of the non-equilibrium state is known. Can you think of other ways of approaching the problem of quantifying thermodynamic functions and verifying the Laws for living organisms? Doing so might lead to new insights on what it is to be alive. Best wishes!

T. References and further reading

Adair, G. (1925). The osmotic pressure of hæmoglobin in the absence of salts. *Proceedings of the Royal Society of London A*, **109**, 292–300.

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science*, **181**, 223–30.

Arakawa, T. & Timasheff, S. N. (1985). Theory of protein solubility. *Methods in Enzymology*, **114**, 49–77.

Atkins, P. W. (1994). *The Second Law: Energy, Chaos, and Form*, ch. 8. New York: Scientific American.

Baker, T. A. (1999). Trapped in the act. *Nature*, **401**, 29–30.

Barth, R. H. (1992). Dialysis. In *Encyclopedia of Applied Physics*, ed. G. L. Trigg, vol. 4, pp. 533–5. New York: VCH.

Ben-Shem, A., Frolow, F. & Nelson, N. (2003). Crystal structure of plant photosystem I. *Nature*, **426**, 630–5.

Bergethon, P. R. (1998). *The Physical Basis of Biochemistry: the Foundations of Molecular Biophysics*, ch. 13.1. New York: Springer-Verlag.

Brandts, J. F. (1964). The thermodynamics of protein denaturation. I. The denaturation of chymotrypsinogen. *Journal of the American Chemical Society*, **86**, 4291–301.

Brandts, J. F. (1964). The thermodynamics of protein denaturation. II. A model of reversible denaturation and interpretations regarding the stability of chymotrypsinogen. *Journal of the American Chemical Society*, **86**, 4302–14.

Bridger, W. A. & Henderson, J. F. (1983). *Cell ATP*. New York: John Wiley.

Brønsted, J. N. (1923). *Recueil des Travaux Chimiques des Pays-Bas*, **42**, 718–28.

Chothia, C. (1984). Principles that determine the structure of proteins, *Annual Review of Biochemistry*, **53**, 537–72.

Christensen, H. N. & Cellarius, R. A. (1972). *Introduction to Bioenergetics: Thermodynamics for the Biologist: A Learning Program for Students of the Biological and Medical Sciences.* Philadelphia: W. B. Saunders.

Conti, B., Sanchez-Alavez, M., Winsky-Sommerer, R., Morale, M.C., Lucero, J., Brownell, S., Fabre, V., Huitron-Resendiz, S., Henriksen, S., Zorrilla, E. P., de Lecea, L. & Bartfai, T. (2006). Transgenic mice with a reduced core body temperature have an increased life span. *Science*, **314**, 825–8.

Cooper, A. (2001). Thermodynamics of protein-protein and peptide interactions. In Aggeli, A., Boden, N. & Zhang, S. (eds.) *Self-assembling Peptide Systems in Biology, Medicine and Engineering*, pp. 47–64. Dordrecht: Kluwer Academic.

Cramer, W. A. & Knaff, D. B. (1991). *Energy Transduction in Biological Membranes. A Textbook of Bioenergetics*. New York: Springer-Verlag.

Creighton, T. E. (1991). Stability of folded proteins. *Current Opinion in Structural Biology*, **1**, 5–16.

Dawes, E. A. (1962). *Quantitative Problems in Biochemistry*, 2nd edn, ch. 1. Edinburgh: E. & S. Livingstone.

Donnan, F. G. (1911). Title. *Zeitschrift für Elektrochemie und Angewandte Physikalische Chemie*, **17**, 572.

Encyclopædia Britannica CD 98, "Colligative Property," "Dialysis," "Metabolism," "Photosynthesis," "Saturation," and "Vapour Pressure."

Epstein, I. R. (1989). The role of flow in systems far-from-equilibrium, *Journal of Chemical Education*, **66**, 191–5.

Fersht, A. R. (1999). *Structure and Mechanism in Protein Science: a Guide to Enzyme Catalysis and Protein Folding*. New York: W. H. Freeman.

Fox, S. & Foster, J. S. (1957). *Introduction to Protein Chemistry*, p. 242. New York: John Wiley.

Franks, F. (1995). Protein destabilization at low temperatures. *Advances in Protein Chemistry*, **46**, 105–39.

Frauenfelder, H., Parak, F. & Young, R. D. (1988). Conformational substates in proteins. *Annual Review of Biophysics and Biophysical Chemistry*, **17**, 451–79.

Freeman, B. (1995). Osmosis. In *Encyclopedia of Applied Physics*, ed. G. L. Trigg, vol. 13, pp. 59–71. New York: VCH.

Fröhlich, H. (1969). Quantum mechanical concepts in biology. In *Theoretical Physics and Biology*, ed. M. Marios. Amsterdam: North-Holland.

Fruton, J. S. (1999). *Proteins, Enzymes, Genes: the Interplay of Chemistry and Biology*. New Haven: Yale University Press.

Garrett, J. (1990). Thermodynamics in sheep. *Education in Chemistry*, **27**, 127.

George, P. & Rutman, R. J. (1960). The 'high energy phosphate bond' concept. *Progress in Biophysics and Biophysical Chemistry*, **10**, 1–53.

Gillispie, Charles C. (ed.) (1970). *Dictionary of Scientific Biography*. New York: Charles Scribner.

Girandier, L. & Stock, M. J. (eds.) (1983). *Mammalian Thermogenesis*. London: Chapman & Hall.

Gutfreund, H. (1949). In *Hæmoglobin: a Symposium Based on a Conference Held at Cambridge in June 1948 in Memory of Sir John Barcroft*, ed. F. J. W. Roughton and J. C. Kendrew, p. 197. London: Butterworths.

Haase, R. (1969). *Thermodynamics of Irreversible Processes*. New York: Dover.

Harold, F. M. (1986). *The Vital Force: a Study of Bioenergetics*. New York: W. H. Freeman.

Harris, D. A. (1995). *Bioenergetics at a Glance*, ch. 1. Oxford: Blackwell Science.

Hatefi, Y. (1985). The mitochondrial electron transport and oxidative phosphorylation system. *Annual Review of Biochemistry*, **54**, 1015–69.

Haynie, D. T. (1993). *The Structural Thermodynamics of Protein Folding*, ch. 4. Ph.D. thesis, The Johns Hopkins University.

Hinckle, P. C. & McCarty, R. E. (1978). How cells make ATP. *Scientific American*, **238**(3), 104–23.

Karplus, M. & McCammon, J. A. (1986). Protein dynamics. *Scientific American*, **254**(4), 30–9.

Katchalsky, A. & Curran, P. F. (1967). *Nonequilibrium Thermodynamics in Biophysics*. Cambridge, Massachusetts: Harvard University Press.

Kauzmann, W. (1958). Some factors in the interpretation of protein denaturation. *Advances in Protein Chemistry*, **14**, 1–63.

Klotz, I. M. (1986). *Introduction to Biomolecular Energetics*, cc. 3–7. Orlando: Academic Press.

Kondepudi, D. & Prigogine, I. (1998). *Modern Thermodynamics: from Heat Engines to Dissipative Structures*, ch. 8.2. Chichester: John Wiley.

Kyoguko, Y., Lord, R. C. & Rich, A. (1969). An infrared study of the hydrogen-bonding specificity of hypoxanthine and other nucleic acid derivatives. *Biochimica et Biophysica Acta*, **179**, 10–17.

Lamm, M. S., Rajagopal, K., Schneider, J. P. & Pochan, D. J. (2005). Laminated morphology of nontwisting b-sheet fibrils constructed via peptide self-assembly. *Journal of the American Chemical Society*, **127**, 16 692–700.

Lodish, H., Baltimore, D., Berk, A., Zipursky, S. L., Matsudaira, P. & Darnell, J. (1995). *Molecular Cell Biology*, 3rd edn, cc. 2, 4 & 21. New York: W. H. Freeman.

Lowell, B. B. & Spiegelman, B. M. (2000). Towards a molecular understanding of adaptive thermogenesis. *Nature*, **404**, 652–60.

Makhatadze, G. I. & Privalov, P. L. (1995). Energetics of protein structure. *Advances in Protein Chemistry*, **47**, 307–425.

Matthews, B. W. (1995). Studies on protein stability with T4 lysozyme. *Advances in Protein Chemistry*, **46**, 249–78.

McCammon, J. A. & Harvey, S. C. (1987). *Dynamics of Proteins and Nucleic Acids*. Cambridge: Cambridge University Press.

Millar, D., Millar, I., Millar, J. & Millar, M. (1989). *Chambers Concise Dictionary of Scientists*. Cambridge: Chambers.

Mitchell, P. (1976). Vectorial chemistry and the molecular mechanisms of chemiosmotic coupling: power transmission by proticity *Biochemical Society Transactions*, **4**, 398–430.

Morowitz, H. J. (1978). *Foundations of Bioenergetics*, ch. 3E. New York: Academic Press.

Nicholls, D. G. & Ferguson, S. J. (1992). *Bioenergetics 2*, ch 3. London: Academic Press.

Noyes, R. M. (1996). Application of the Gibbs function to chemical systems and subsystems. *Journal of Chemical Education*, **73**, 404–8.

O'Brien, C. (1994). New enzyme structure reveals cell's rotary engine. *Science*, **265**, 1176–7.

Osyczka, A., Moser, C. C., Daldal, F. & Dutton, P. L. (2004). Reversible redox energy coupling in electron transfer chains. *Nature*, **427**, 607–12.

Pardee, G. S. & Ingraham, L. L. (1960). Free energy and entropy in metabolism. In *Metabolic Pathways*, ed. D.M. Greenberg, vol. I. New York: Academic Press.

Pauling, L. C. (1970). Structure of high energy molecules. *Chemistry in Britain*, **6**, 468–72.

Pepys, M. B., Hawkins, P. N., Booth, D. R., Vigushin, D. M., Tennent, G. A., Soutar, A. K., Totty, N., Nguyent, O., Blake, C. C. F., Terry, C. J., Feest, T. G., Zalin, A. M. & Hsuan, J. J. (1993). Human lysozyme gene mutations cause hereditary systemic amyloidosis. *Nature*, **362**, 553–7.

Peusner, L. (1974). *Concepts in Bioenergetics*, cc. 3, 5, 6, 7 & 10–8. Englewood Cliffs: Prentice-Hall.

Plum, G. E. & Breslauer, K. J. (1995). Calorimetry of proteins and nucleic acids. *Current Opinion Structural Biology*, **5**, 682–90.

Prigogine, I. (1967). *Introduction to Thermodynamics of Irreversible Processes*. New York: John Wiley.

Prigogine, I. (1969). Structure, dissipation and life. In *Theoretical Physics and Biology*, ed. M. Marios. Amsterdam: North-Holland.

Prigogine, I., Nicolis, G. & Babloyants, A. (1972). Thermodynamics of evolution. *Physics Today*, **25** (11), 23–8.

Prigogine, I., Nicolis, G. & Babloyants, A. (1972). Thermodynamics of evolution. *Physics Today*, **25** (12), 38–44.

Record, M. T., Jr, Zhang, W. & Anderson, C. F. (1998). Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: a practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects. *Advances in Protein Chemistry*, **51**, 281–353.

Rich, P. (2003). Chemiosmotic coupling: The cost of living. *Nature*, **421**, 583.

Roberts, T. J., Marsh, R. L., Weyland, P. G., & Taylor, C. R. (1997). Muscular force in running turkeys: the economy of minimizing work. *Science*, **275**, 1113–15.

Rondelez, Y., Tresset, G., Nakashima, T., Kato-Yamada, Y., Fujita, H., Takeuchi, S. & Noji, H. (2005). Highly conserved ATP synthesis by F1-ATPase single molecules. *Nature*, **433**, 773–7.

Rovner, S. L. (2004). The secrets of aging. *Chemical & Engineering News*, **82**(34), 30–5.

Ruelle, D. (2001). A departure from equilibrium. *Nature*, **414**, 263–5.

Schellman, J. A. (1987). The thermodynamic stability of proteins. *Annual Review of Biophysics and Biophysical Chemistry*, **16**, 115–37.

Schnitzer, M. J. (2001). Doing a rotary two-step. *Nature*, **410**, 878–81.

Secrest, D. (1996). Osmotic pressure and the effects of gravity on solutions. *Journal of Chemical Education*, **73**, 998–1000.

Segal, I. H. (1976). *Biochemical Calculations: How to Solve Mathematical Problems in General Biochemistry*, 2nd edn, ch. 3. New York: John Wiley.

Shavit, N. (1980). Energy transduction in chloroplasts. *Annual Review of Biochemistry*, **49**, 111–39.

Shaw, A. & Bott, R. (1996). Engineering enzymes for stability. *Current Opinion in Structural Biology*, **6**, 546–50.

Shortle, D. (1996). The denatured state (the other half of the folding equation) and its role in protein stability. *Federation of the American Societies for Experimental Biology Journal*, **10**, 27–34.

Smith, C. A. & Wood, E. J. (1991). *Energy in Biological Systems*, cc. 1.3 & 1.4. London: Chapman & Hall.

Snell, F. M., Shulman, S., Spencer, R. P. & Moos, C. (1965). *Biophysical Principles of Structure and Function*, ch. 8. Reading, Massachusetts: Addison-Wesley.

Spolar, R., Livingstone, J. & Record, M. T., Jr. (1992). Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. *Biochemistry*, **31**, 3947–55.

Tanford, C. (1968). Protein denaturation (parts A and B). *Advances in Protein Chemistry*, **23**, 121–282.

Timasheff, S. N. (1993). The control of protein stability and association by weak interactions with water: how do solvents affect these processes? *Annual Review of Biophysics and Biomolecular Structure*, **22**, 67–97.

Timasheff, S. N. (1998). Control of protein stability and reactions by weakly interacting cosolvents: the simplicity and the complicated. *Advances in Protein Chemistry*, **51**, 355–432.
Timbrell, J. A. (1991). *Principles of Biochemical Toxicology*, 2nd edn, ch. 7. London: Taylor & Francis.
van Holde, K. E. (1985). *Physical Biochemistry*, 2nd edn, cc. 2.1, 2.3, 2.4, 3.4 & 3.5. Englewood Cliffs: Prentice-Hall.
Voet, D. & Voet, J. G. (1995). *Biochemistry*, 2nd edn, cc. 3, 4, 15–4–15–6, 16, 18–1, 19–1, 20, 22, 28–3, 28–5A & 34–4B. New York: John Wiley.
Weber-Ban, E. U., Reid, B. G., Miranker, A. D. & Horwich, A. L. (1999). Global unfolding of a substrate protein by the Hsp100 chaperone ClpA. *Nature*, **401**, 90–3.
Williams, S. (1999). Life as a part-time plant. *Wellcome News*, **20**, 38.
Williams, T. I. (ed.) (1969). *A Biographical Dictionary of Scientists*. London: Adam & Charles Black.
Woodcock, A. & Davis, M. (1978). *Catastrophe Theory*. Harmondsworth: Penguin.
Wrigglesworth, J. (1997). *Energy and Life*, cc. 3, 5.7.2, 7.1, 7.3 & 7.5.1. London: Taylor & Francis.
Yasuda, R., Noji, H., Yoshida, M., Kinosita, K., Jr & Itoh, H. (2001). Resolution of distinct rotational substeps by submillisecond kinetic analysis of F_1-ATPase. *Nature*, **410**, 898–904.
Youvan, D. C. & Marrs, B. L. (1987). Molecular mechanisms of photosynthesis. *Scientific American*, **256** (6), 42–8.
Zouni, A., Witt, H.-T., Kern, J., Fromme, P., Krauß, N., Saenger, W. & Orth, P. Crystal structure of photosystem II from Synechococcus elongates at 3.8 Å resolution. *Nature*, **409**, 739–43.

U. Exercises

1. Speculate in broad terms on the effect on Earth of the cessation of photosynthesis.

2. The energy conversion process by which sunlight is converted into biomass is not completely efficient. What happens to the energy that does not become biomass? Rationalize your answer in terms of the First and Second Laws of thermodynamics.

3. Animal life as part-time plant? Sue Williams of the Department of Botany, University of Western Australia, says that the green-tinged sea slugs she studies "enslave" chloroplasts from the seaweed they ingest, and use them as a means of capturing up to 25% of their energy. Explain how this might work.

4. Use the following information to determine the standard free energy change of ATP hydrolysis.

$$\text{Glucose} + \text{ATP} \rightarrow \text{glucose-6-phosphate} + \text{ADP} \qquad \Delta G^{\circ\prime} = -16.7 \text{ kJ mol}^{-1}$$
$$\text{Glucose-6-phosphate} \rightarrow \text{glucose} + P_i \qquad \Delta G^{\circ\prime} = -13.8 \text{ kJ mol}^{-1}$$

Show all work.

5. Buffers containing ATP are ordinarily made up fresh and not stored as a stock solution. When a stock solution is made, it must usually be kept at 4°C (short term storage) or at −20°C (long term storage). Rationalize these practices. What bearing does this have on the necessary molecular machinery of a cell?

6. ATP is the energy currency of the cell. ATP is essential for life as we know it. Comment on the stability of ATP in aqueous solution and the constraints this may place on theories of the origin of life.

7. The free energy status of a cell can be described in various ways. One of these, called the *adenylate energy charge* (AEC), was first proposed by Daniel Edward Atkinson (1921–). The AEC is defined as

$$\mathrm{AEC} = ([\mathrm{ATP}] + 0.5[\mathrm{ADP}])/([\mathrm{ATP}] + [\mathrm{ADP}] + [\mathrm{AMP}])$$

and it varies between 1.0, when all the adenine nucleotide is ATP, and 0, when all the β- and γ-phosphoanhydride bonds have been hydrolyzed. The relative amounts of ATP, ADP, and AMP can be determined by comparing the sizes of the respective peaks in a high-performance liquid chromatography (HPLC) profile. The AEC of a healthy cell is about 0.90–0.95. Malignant hypothermia is an inherited muscle disease in humans and pigs. Patients suffer rapid rises in body temperature, spasms in skeletal muscle, and increases in the rate of metabolism, which can be fatal if not treated with a suitable muscle relaxant. The following data were obtained before the onset of symptoms and just prior to the death of a pig afflicted with the disease.

	[ATP]	[ADP] μmol g^{-1} tissue	[AMP]
Before symptoms	4.2	0.37	0.029
Before death	2.1	0.66	0.19

Calculate the AEC before the symptoms began to occur and just before death. Comment on the magnitude of the values and what they indicate.

8. A 1 M solution of a glucose gives a pressure more than 25 times greater than that of the atmosphere. A 1 M solution of a salt gives an even larger osmotic pressure. Explain.

9. Suppose we have an osmometer that is constructed from a capillary tube with a membrane in the middle, and that the tube is oriented *horizontally* (why?). Now let some osmotic particles suddenly appear on one side of the tube only. Explain what happens.

10. You have a U-tube osmometer with cells of equal shape and volume. On one side, you place a sphere of volume V, and at the

same time and on the other side you place a cube of volume V. Neither particle is membrane-permeant. Suppose that these particles are able to interact with the solvent. Explain what will happen in the following situations: (a) the particles are so dense that they sit on the bottom of the cells of the osmometer; (b) the density of the particles is such that they are able to diffuse throughout the volume of their respective cells.

11. What causes the membrane of a red blood cell to burst when the cell is placed in hypotonic solution? Be as specific as possible.

12. Suppose you have an osmometer in which the solute particles are confined to a fixed volume, for instance an indestructible membrane of fixed volume. What happens? Why?

13. Suppose you have an osmometer with a membrane that is permeable to water but not to larger molecules. Add glucose to one side to a final concentration of 0.001 M and hemoglobin to the other side to a concentration of 0.001 M. Will a pressure head develop? If yes, on which side will the water level be higher? If no, why not?

14. Suppose you are involved in preparing recombinant protein for a physical biochemistry experiment. The approach involves 8 M urea, formation of which from urea crystals and water is highly endothermic. The bacteria are lysed in 8 M urea, a chemical denaturant, and the recombinant protein is separated from the bacterial proteins by column chromatography. Separation of the recombinant protein from urea is done by dialysis in two stages. In each, *c.* 100 ml of lysate is dialyzed against 5 l of water. The dialysis membrane allows the passage of water and urea but not protein. Will the volume of the protein preparation change in this procedure, and if so, how? Assuming that the volume of the protein solution at the end of dialysis is 100 ml, what is the final concentration of urea? Explain, in enthalpic and entropic terms, the driving force for the reduction in urea concentration in the first step of dialysis. Explain from a thermodynamic point of view what drives the further reduction in urea concentration in the second step of dialysis.

15. Recall what happens to a red blood cell when it's placed in a hypotonic solution. What must be done to ensure that dialysis tubing doesn't burst?

16. Prove that Eqns. (5.19) and (5.20) follow from the preceding equations.

17. Show that Eqn. (5.38) follows from Eqn. (5.36).

18. Derive Eqn. (5.40). (Hint: start with $R_T = R + R \bullet I + R \bullet L$, express R and $R \bullet I$ in terms of $R \bullet L$, and solve for $R \bullet L$.)

19. Equation (5.41) is 0 for all concentrations of I_i when there is no inhibition (compound I_i has no effect), and it is 1 at 100% inhibition. Explain.

20. Analysis of gene regulation involves study of structural and thermodynamic aspects of how proteins bind nucleic acid. One area of such research is the recognition of DNA operator sites by repressor molecules. Suppose protein P binds a single specific sequence on a molecule of DNA D. This is a common mechanism for the regulation of gene expression. At equilibrium, $P + D \Leftrightarrow P \bullet D$. A bacterial cell contains one molecule of DNA. Assume that cell is cylindrical, and that its diameter and length are 1 μm and 2 μm, respectively. Calculate the total concentration of D. Assume that $K_{eq} = 10^{-10}$ M. Calculate the $[P \bullet D]$, assuming that $[P] = [D]$. The concentration of *bound* D is just $[P \bullet D]$. Calculate the concentration of *unbound* D. Calculate $[P \bullet D]/[P]$. Give an interpretation of this quantity. The subject of binding will be discussed in detail in Chapter 7.

21. The previous problem involved the association and dissociation of two types of macromolecule, proteins and DNA. A basic feature of such situations is the dependence of the equilibrium on the total concentrations of the interacting species. The concept can be illustrated by means of the monomer–dimer equilibrium. Consider the equilibrium

$$2M \Leftrightarrow D. \tag{5.61}$$

The total concentration of monomer, $[M]_T$, is $[M] + 2[D]$, where the factor 2 accounts for there being two monomers in each dimer. This equation can be solved for [D]. Write down an expression for the equilibrium constant for the reaction in Eqn. (5.61). Combine this with your equation for [D] and solve the resulting quadratic equation for [M]. Show that $[M]/[M]_T \to 1$ as $[M]_T \to 0$, and that $[M]/[M]_T \to 0$ as $[M]_T \to \infty$. How does one interpret these limiting conditions?

22. What might be the structural basis for the low stability of Georgia hemoglobin relative to normal hemoglobin?

23. Hemocyanin is a Cu-containing oxygen-binding protein that is found in some invertebrates. In squid hemocyanin, when the partial pressure of oxygen gas is 0.13 atm at 25 °C, the oxygen binding sites are 33% saturated. Assuming that each hemocyanin molecule binds one molecule of oxygen gas, calculate the equilibrium constant. What are the units of the equilibrium constant? Calculate the standard state free energy change when hemocyanin interacts with O_2(aq). The solubility of pure oxygen in water at 1 atm and 25 °C is 0.001 17 mol $(\text{kg } H_2O)^{-1}$.

24. In ELISA, what type of interactions are likely to be most important for protein adsorption to the solid support? Why are antibodies able to bind to partially denatured protein?

25. Explain in thermodynamic terms how a single 30-cycle PCR experiment can yield billions of copies of double-stranded DNA.

26. Under normal conditions, complementary strands of DNA form a double helix. In the section on PCR we provided a way of rationalizing the stability of DNA. Compare and contrast our view with that put forward by Voet and Voet, authors of a popular biochemistry textbook (see pp. 866–70 of the second edition, published in 1995). Can the data in their Table 28–4 be trusted? Why or why not?

27. Equation (5.50) for DNA supercoiling resembles the First Law of Thermodynamics. List and explain the similarities and differences.

28. A certain machine of a biotechnology company provides a controlled environment for the automation of sequence-specific DNA analysis and performs all the reaction steps required for capture and detection of nucleic acids. A main feature of the product is its capture specificity. For instance, suppose a 300 bp PCR fragment derived from the filamentous bacteriophage M13 was specifically captured by using a series of complementary oligonucleotide probes 24 residues in length, and that the capture probes incorporated 0–6 mismatches with the target. Explain how optimizing the hybridization conditions (i.e. by adjusting the temperature) could distinguish sequences differing by a single base.

29. "Hot start." When plasmid DNA is used as the template in a PCR reaction, the enzyme buffer, plasmid, and oligonucleotide primers are often incubated at 95 °C for several minutes before starting thermal cycling. Why?

30. The release of insulin from pancreatic β-cells on uptake of glucose is a complex process. The steps of the process in rough outline are as follows. The resting membrane potential of a β-cell is determined by open ATP-sensitive K^+ channels in the plasma membrane. After a meal, glucose is taken into the cell and phosphorylated. Eventually, there is an increase in [ATP]/[ADP] ratio in the cell, and this closes the K^+ channels. The membrane depolarizes, stimulating the opening of Ca^{2+} channels. Calcium enters the cell, stimulating the release of insulin through exocytosis of secretory granules. Describe each step of this process in moderately detailed thermodynamic terms.

31. Isothermal titration calorimetry. The key condition underlying this technique is thermodynamic equilibrium. When an aliquot of titrant is injected, the Gibbs free energy of the system increases. A spontaneous chemical reaction occurs until G reaches a new

minimum and equilibrium is established once again. An ITC study of a ligand binding to a macromolecule was carried out at three temperatures, T_1, T_2 and T_3, where $T_1 < T_2 < T_3$. At T_1, $\Delta H_b > 0$; at T_2, $\Delta H_b = 0$; and at T_3, $\Delta H_b > 0$. The ligand is known to bind the macromolecule at all three temperatures by means of independent experiments. Explain what is happening in the reaction cell at each stage of a general ITC experiment, viz. before an injection and during an injection. Rationalize the results obtained.

32. Speculate on the possibility of observing the cold denaturation of DNA. What about tRNA?

33. The folded and unfolded states of a protein are in equilibrium as shown in Eqn. (5.57). Suppose that you are working with a solution of RNase A at a concentration of 2.0×10^{-3} M, and that fractions of protein in the *unfolded* state are as follows: 50 °C: 0.002 55; 100 °C: 0.14. In the thermal denaturation of this protein, there are essentially just two states, the folded one and the unfolded one, so the fraction of protein in the folded state is just one minus the fraction in the unfolded state. Calculate ΔH° and ΔS° for *unfolding* of RNase A. What key assumption must be made about temperature-dependence? Calculate ΔG° for *unfolding* of RNase A at 37 °C. Is this process spontaneous at this temperature? Determine the melting temperature of RNase A under standard state conditions (for a two-state reaction, at T_m half of the proteins are folded and half are unfolded).

34. The role of ΔC_p in protein stability and its molecular origin was discussed in publications by the American biochemist John Brandts as early as 1964. Use Eqn. (4.3) to investigate the role of ΔC_p in the thermostability of a protein. One relatively easy way to do this is to assume values for ΔH and ΔS at some reference temperature, say 298 K, and then to use a spreadsheet to calculate ΔG throughout a temperature range that includes 0–100 °C. Plot ΔG v. T for several different values of ΔC_p. Note that the curve crosses the T-axis at two points. What are the names of these intercepts? What if $\Delta C_p < 0$? Is this physically meaningful? Is it relevant to biological macromolecules?

35. Suppose you have designed a four-helix bundle. A four-helix bundle is just a polypeptide that folds into four helices of approximately equal length, and whose helices are bundled together. The helices interact with each other in the core of the protein. Various structural techniques show that at room temperature the structure is highly dynamic and not very much like an organic crystal, though all four helices are intact. Thermal denaturation studies, however, indicate that the unfolding temperature of your designed protein is over 100 °C! Explain. How could the design be modified to reduce the melting temperature and increase the specificity of interactions in the protein core?

36. Living organisms have been described as "relatively stable" systems that "show an organized collective behavior which cannot be described in terms of an obvious (static) spatial order" and are "not near thermal equilibrium." Explain.

37. The synthesis of ATP under standard conditions requires 7.7 kcal mol^{-1}, and this is coupled to the movement of $2H^+$ across a mitochondrial membrane. Calculate the pH difference across the inner mitochondrial membrane needed to drive ATP synthesis at 25 °C.

38. Oxidation–reduction reactions in *E. coli* generate a pH gradient of +1 (outside to inside) and a voltage gradient of –120 mV (outside to inside). What free energy is made available by this proton motive force? β-Galactosides are transported along with H + ions. Calculated the maximum concentration ratio of β-galactoside that can result from the coupling of its transport to the proton motive force.

39. An empirical expression for the melting temperature of double-stranded DNA in the presence of NaCl is

$$T_m = 41.1X_{G+C} + 16.6 \log[Na^+] + 81.5, \quad (5.62)$$

where X_{G+C} is the mole fraction of G–C pairs. Given a 1000 base pair gene with 293 Gs and 321 Cs, calculate the sodium ion concentration at which it will have a melting temperature of 65 °C.

40. Use the following osmotic pressure data for horse hemoglobin in 0.2 M phosphate and at 3 °C to determine the molecular mass of the protein.

Concentration of hemoglobin (g/100 ml)	Osmotic pressure (cm H_2O)
0.65	3.84
0.81	3.82
1.11	3.51
1.24	3.79
1.65	3.46
1.78	3.82
2.17	3.82
2.54	3.40
2.98	3.76
3.52	3.80
3.90	3.74
4.89	4.00
6.06	3.94
8.01	4.27
8.89	4.36

41. The effect of pH on the osmotic pressure of sheep hemoglobin was investigated by Gilbert Adair (Chapter 7). The following data were obtained.

pH	Osmotic pressure (mmHg/1 g protein/100 ml)*
5.0	21.5
5.4	13.4
6.5	3.2
6.7	2.4
6.8	2.4
6.8	3.5
6.8	4.5
7.2	5.0
9.6	15.6
10.2	21.4

*1 mmHg = 133.322 . . . Pa.

Plot the data and use them to deduce the isoelectric point of sheep hemoglobin.

42. Why would it not be a good idea to water your houseplants with boiling water?

43. Suggestion biochemical means by which one might test the origin of the heat produced by *Arum maculatum* (see Chapter 3). (Hint: use tissue extracts of the spadix and appendix of the plant and consider poisons that block either electron transport or oxidative phosphorylation).

44. It is sometimes said that the two terminal phosphoanhydride bonds of ATP are "high-energy" bonds. This implies that the energy released as free energy when the bond is cleaved is stored within the bond itself. Why is the term *high-energy bond* misleading?

45. Mg^{2+} ions interact with ATP under physiological conditions. What is the likely effect of this on the free energy of hydrolysis of ATP? Why?