Chapter 5

Mechanical and Chemical Equilibrium in the Living Cell

"The study of particular problems of the calculus of variations, or, as we shall say, particular variational problems, is extremely old. It arises from the fact that for human beings, in many instances, only the best can be good enough."
-L. C. Young

Chapter Overview: In Which We Examine How Cells Manage Energy and How Scientists Compute Energy Transformations

Energy consuming and liberating chemical transformations are one of the hallmarks of living systems. Living cells follow the same principles of conservation of matter and energy as do all other physical systems, though they also operate under an additional set of constraints imposed by their evolutionary history. In this chapter, we first summarize how cells manipulate and store chemical energy in ways that can be used to perform material transformations such as macromolecular synthesis, mechanical work such as muscle contraction or even production of light energy like in a firefly's abdomen. In order to develop the mathematical tools necessary to model these kinds of biological transformations, we exploit the useful simplification that many chemical and mechanical systems can be treated as if they are close to an equilibrium state. As we will see, many real world situations can be surprisingly well modeled using equilibrium assumptions. This perspective alone is enough to provide useful insight into biological phenomena as fundamental and diverse as protein folding, binding reactions and formation of lipid bilayers. This useful oversimplification will form the substance of the next several chapters.

5.1 Energy and the Life of Cells

Much of the business of cellular life involves transformations of energy. Most familiar organisms make their living by eating other living or freshly dead things, thereby consuming energy-carrying organic molecules that have been generated and shepherded by other living organisms. This material transfer process makes all forms of life on Earth an interconnected and interdependent web where a key mode of communication is energy transfer. Humans consume food made up largely of fats, proteins and carbohydrates, initially synthesized in other organisms, mostly plants, animals and fungi. They use the molecules they consume not only to create material but also to fuel the energy-requiring processes of daily life including muscle contraction, heat generation and brain activity.

Ultimately, it would not be possible for life to survive merely by recycling or exchanging energy among organisms - there must be an outside energy source. For most ecosystems on earth, the ultimate energy source is sunlight which is harvested by various cells in plants, in many unicellular eukaryotes and many kinds of bacteria. See Morton (2007) in "Further Reading" for a fascinating discussion of life and light. The light gathered by these cells serves not only their own energy needs, but eventually provides the energy for the remainder of the interdependent web of life. These cells exploit the energy of sunlight using specialized light-harvesting molecules to transfer ions across membranes. The ion gradients store energy in a battery-like form that can then be coupled to enzymes that can, for example, convert CO_2 (and H_2O) from the air into sugar and indirectly into all other biomolecules including proteins, lipids and nucleic acids.

Four key kinds of energy are relevant in biological systems: chemical energy, mechanical energy, electromagnetic energy and thermal energy. Each of these forms of energy may be converted by living organisms into each of the others, with the interesting and important exception that thermal energy is generally a dead end since the second law of thermodynamics prohibits harnessing thermal (random) motions to carry out useful work. For example, electromagnetic energy in the form of a photon from the sun can be harnessed by a cell during photosynthesis to generate chemical energy that may be used for metabolic processes. Conversely, fireflies and other bioluminescent organisms convert chemical energy into photons. Energy gathered by organisms from their environments can be stored for later use, primarily in chemical form. Energy-storing molecules are used by all organisms for a host of important cellular processes. For example, ATP is used to pump molecules across membranes, to create the specialized polymeric apparatus of cellular motility and to power the motors that allow our muscles to twitch.

Throughout the book, we will invoke a wide range of different models to explore how these energy transformations take place. The main goal of this chapter is to demonstrate how the principles of energy minimization and free energy minimization can be used to predict the direction of transformations occurring in living systems. In considering energy minimization calculations, we will often make an explicit or implicit assumption that the system is operating

close to equilibrium, such that any small excursion of the system will typically result in it returning to its original state. How do we reconcile the mathematically convenient equilibrium assumption with the real world observation that biological systems are constantly dynamic and changing? The key insight is that different processes occur at different time scales, and so we can frequently isolate some small part of a biological process occurring at a relatively rapid time scale and pretend that it is at equilibrium with respect to its effects on processes that occur more slowly.

The next several chapters will build up the tools of equilibrium thermodynamics and statistical mechanics for treating equilibrium problems. Before embarking on our journey through the biological uses of statistical mechanics, we begin by taking stock of the interplay of thermal and deterministic forces (and energies) in biology and we examine the chemical basis for biological energy storage.

5.1.1 The Interplay of Deterministic and Thermal Forces

One of the important characteristics of the cellular interior that makes it so different from the world of everyday experience is the fact that thermal and deterministic forces are on equal footing. By thermal forces, we refer to the forces exerted on macromolecular structures as a result of the incessant jiggling of all of the molecules (such as water) that surround them. When we are considering the transformations that a biological system can undergo, it is useful to picture the range of available possibilities in terms of an energy landscape. For example, a protein may exist in a large number of possible conformations but some will be energetically preferred over others. In any given biological system, the shape of the peaks and valleys on the free energy landscape can be changed by an energy input. For example, mechanical stretching of a membrane containing an ion channel will tend to make the open conformation of the channel more favorable relative to the closed conformation. At the same time, the rates at which molecules explore the energy landscape tend to be primarily determined by thermal forces. These different effects can be quantitatively related to one another through use of common units.

Thermal Jostling of Particles Must Be Accounted For in Biological Systems

Perhaps the most famed example of thermal effects is that of Brownian motion (the microscopic basis of diffusion) already introduced in chap. 3 (pg. 160). Observation of small particles ($\approx 1\mu m$), fluorescently labeled molecules and even macromolecules within cells reveals the fact that they suffer excursions which are, to all appearances, completely random. This jostling is a reflection of the fact that in addition to whatever deterministic forces might be applied to the particle or molecule of interest (such as electrostatic interactions, attachment to springs, etc.), they are also subjected to forces due to constant collisions with the molecules that make up the surrounding medium, which in turn constantly collide with one another.

To get a sense of the relative importance of thermal and deterministic forces, we need a numerical measure of the contribution of thermal effects. One way to compare the thermal and deterministic scales is through ratios of the form E_{det}/k_BT, where E_{det} represents the scale of deterministic energies in the problem of interest. For example, we might ask for the energy scale associated with breaking a hydrogen bond, or the energetic cost of bending a DNA molecule. In chap. 6 we will show that the probability of a given "microstate" of a system is proportional to the Boltzmann factor, $exp(-E_{det}/k_BT)$, revealing the quantitative interplay of thermal and deterministic energies. The natural energy unit for a single molecule inside a cell is set by the thermal energy scale at room temperature, namely,

$$k_BT = 4.1\text{pN nm}. \tag{5.1}$$

We can see that this energy scale will be of central importance to the life and times of macromolecules such as lipids, proteins and nucleic acids because the energy delivered by ATP hydrolysis is tens of k_BT and many of the motors that perform the functions of the cell exert piconewton forces over nanometer length scales. For other kinds of biological transformations, it is sometimes more useful to consider the thermal energy scale in different units. For example, for biochemical reactions $k_BT = 0.6$ kcal/mole or 2.5 kJ/mole and when considering thermal motions of charge, we will use $k_BT = 25$ meV.

These ideas on the relative importance of thermal and deterministic forces are made more concrete in fig. 5.1. The horizontal line in the figure corresponds to the thermal energy scale, represented here as $k_BT = 4.1 \times 10^{-21}$ joules. The other lines illustrate the energy cost associated with particular deterministic scenarios such as stripping a fraction of the charge off of spheres of different size, bending rods of different sizes and confining electrons within boxes of different sizes which is meant to convey a feeling for the energy scale of binding (which is also captured explicitly as the energy of hydrogen and van der Waals bonding in the figure). The key point of the figure is to note that at the *nanometer* scale (precisely the scale of the macromolecules of the cell) thermal energies and the deterministic energies of properties like charge rearrangement, bonding and molecular rearrangement are comparable, unlike the familiar centimeter or meter scales where deterministic forces predominate.

Because each of these forms of energy is of comparable scale and effectively interchangeable at the molecular level, a living organism which needs to generate motion, heat, electricity and biomolecular synthesis is expert at energetic interconversions. For the most part, energy used by living cells is derived from chemical energy in food and used to generate all the other forms. An interesting exception is photosynthesis where electromagnetic energy is first converted into chemical energy and then into everything else. To develop a feeling for the numbers we will now consider the molecular basis for generation and storage of chemical energy in cells.

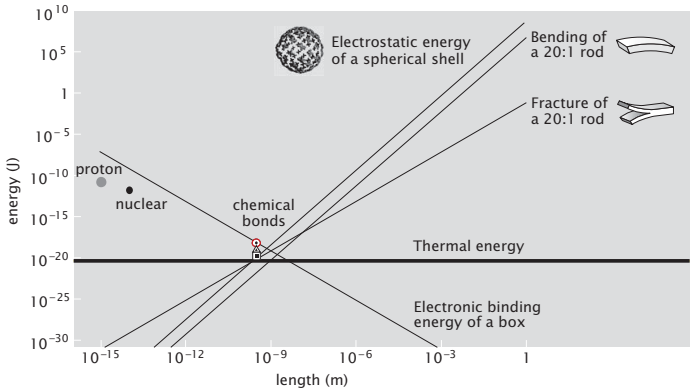

Figure 5.1: Energy as a function of length scale for a number of different energetic mechanisms. The graph shows how thermal, chemical, mechanical and electrostatic energies associated with an object scale with the size of the object. As the characteristic object size approaches that of biological macromolecules, all of the energy scales converge to a single regime. The horizontal line shows the thermal energy scale. The bending energy is estimated by considering an elastic rod with an aspect ratio of 20:1 which is bent into a semicircular arc. The electrostatic energy is estimated for a model spherical protein with polar residues on its surface and for which all of the polar residues are stripped of a single charge (see chap. 9). Chemical energy as a function of length, or binding energy, is estimated approximately by considering the effects of confining a free electron in a box of that length scale. For comparison, measured binding energies are shown for three chemical bonds (hydrogen bonds, phosphate groups in ATP and covalent bonds). On this log-log scale they all appear very similar to one another at the point of convergence. (Adapted from R. Phillips and S. Quake, *Phys. Today*, 59:38, 2006.)

5.1.2 Constructing the Cell: Managing the Mass and Energy Budget of the Cell

In chap. 2 (pg. 59), we estimated the number of all the different kinds of macromolecules in a cell and worked out the number of glucose molecules required to build these constituents if glucose is the sole carbon source. What we neglected was the significant metabolic work that must be performed to transform the carbon atoms of glucose into the carbon atoms of amino acids, nucleotides, fatty acids, etc. Metabolism is the general term used to refer to cellular transformations of one molecule into another. The specific transformation of glucose into the amino acid lysine, for example, requires the ordered action of many enzymes, several of which consume energy while performing the necessary transformations. In living cells, energy is stored and transferred in several forms, most commonly in the form of a high energy chemical bond on the molecule ATP (we will discuss cellular energy in more detail below). The ultimate source of the energy used to synthesize ATP in fact comes from metabolic breakdown of glucose in a pathway known as glycolysis and illustrated schematically in fig. 5.2. For *E. coli* growing in the presence of oxygen, a single molecule of glucose can be metabolically broken down to form up to thirty molecules of ATP from ADP since in this case, the pyruvate emerging from the glycolysis pathway can be used to fuel further energy producing reactions. This process results in carbon dioxide as a waste product. One interesting question is what fraction of the glucose taken on by the cell is used to make new molecules and what fraction is used to provide the energy to make those new molecules? In this section, we will estimate the energy budget of a single *E. coli* proceeding through one round of its cell cycle.

In order to perform this estimate, we need to understand the nature of energy storage in cells, the typical amounts of energy required for metabolic transformations and the ways in which cells allocate their energy and material resources. In most cells, energy is stored in a variety of forms which can be interconverted with very high efficiency. The three most commonly used are ATP, NADH (NADPH) and transmembrane H^+ gradients as shown in fig. 5.3. ATP (adenosine triphosphate) is often referred to as the energy currency of the cell because it can be easily converted into goods and services. The energy liberated by hydrolysis of the γ-phosphate bond on ATP to generate a molecule of ADP (adenosine diphosphate) and an inorganic phosphate ion P_i (PO_4^{2-}) is approximately $20k_BT$ (though it depends upon the concentrations of all of ATP, ADP and P_i as will be shown in section 6.4.4 (pg. 350)). ATP is a useful energy currency because this amount of energy is comparable to the energy consumed in many kinds of biochemical transformations and is intermediate between thermal energy (k_BT) and the energy of a typical covalent bond ($100k_BT$). ATP can be considered as the twenty dollar bill of the cell because of its intermediate value in the overall energy economy of the cell. Spending money in large chunks such as a hundred dollar bill is unwieldy because they are hard to break. On the other hand, paying with just dollar bills is a nuisance because it takes many of them to buy anything useful.

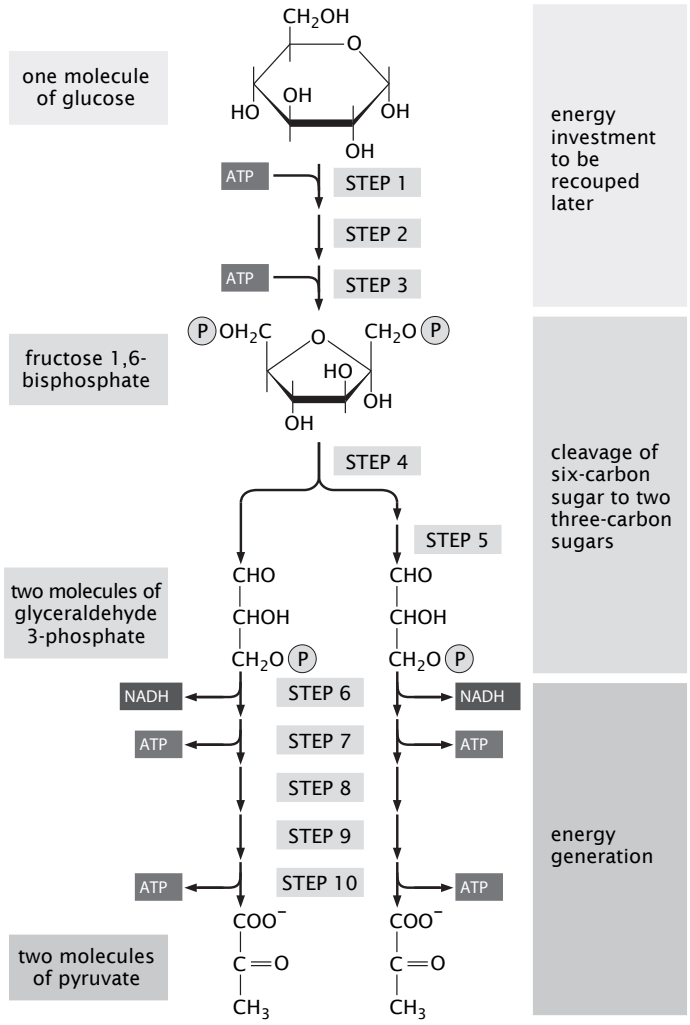

Figure 5.2: A schematic outlining the overall organization of the glycolytic pathway. The outcome of the ten steps of glycolysis is the conversion of a single molecule of glucose into two molecules of pyruvate and the concomitant net production of two molecules of ATP and two of NADH. (Adapted from B. Alberts *et al.*, Molecular Biology of the Cell, 4th ed. New York: Garland Science, 2002.)

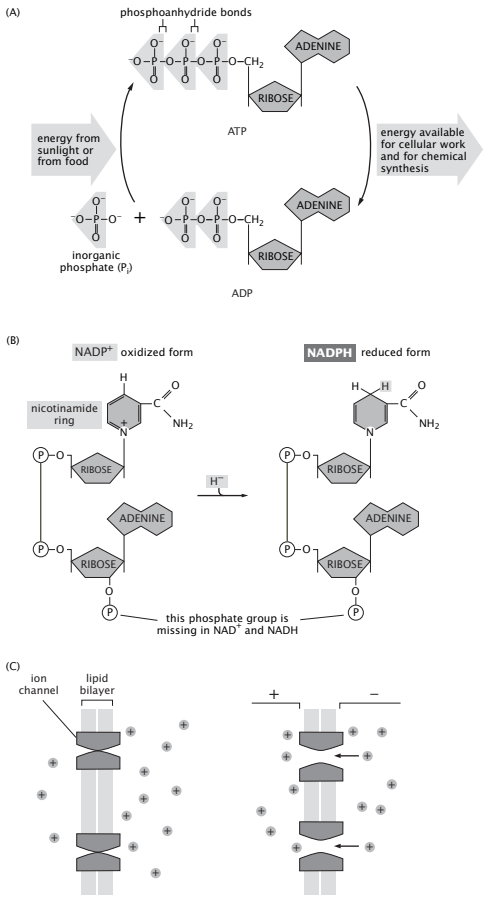

Figure 5.3: Three important forms of biological energy. (A) Energy for chemical synthesis and for force generation is stored in the form of ATP which can be converted to ADP + P_i releasing roughly 20 k_BT of useful energy. ADP + P_i can then be converted back to ATP. While many enzymes use ATP itself, others use GTP, UTP or CTP, but the energies are equivalent. (B) Reducing potential is carried in the form of transferrable high-energy electrons on NADH (or the very similar molecule NADPH). Two electrons can be transferred from NADPH to reduce an oxidized organic compound liberating one hydrogen ion (H^+) and the oxidized form of the carrier molecule $NADP^+$. In this case, the energy liberated by oxidation of one mole of NADH can be used to synthesize roughly three moles of ATP. (C) Transmembrane ion gradients, particularly in the form of H^+ gradients, are also used to store energy. (A-C, adapted from B. Alberts *et al.*, Molecular Biology of the Cell, 4th ed. New York: Garland Science, 2002.)

A second form of chemical energy important for metabolic transformations is carried in the form of easily transferrable electrons on the molecules NADH and NADPH. Many metabolic transformations require that an organic molecule be altered in its level of oxidation. Oxidation and reduction reactions refer to the transfer of electrons between compounds. A compound is oxidized when electrons are removed and reduced when electrons are added. For organisms growing in an oxygen-rich environment of the modern Earth, oxidation reactions are usually spontaneous. However, reduction reactions require energy input. For reductive biosynthesis, a pair of electrons are usually donated by NADPH creating an oxidized form of this carrier molecule $NADP^+$. Hence, NADPH gives up its hydride ion (H^-) in the same way that ATP gives up P_i, in both cases liberating energy for doing useful biochemical work.

Another use of reducing energy in cells is to establish H^+ gradients across membranes. This is an example of the third major form of biological energy storage. The electrical consequences of charge separation by transmembrane ion gradients will be the focus of chap. 17. Ion gradients are easily interconvertible with either ATP energy or NADH energy. NADH can donate its high energy electrons to electron carrier molecules in the plasma membrane of bacterial cells or in the inner mitochondrial membrane of eukaryotic cells, that ultimately liberate H^+ ions on the opposite side generating a gradient. The energy stored in this kind of ion gradient can be converted into ATP through the action of the enzyme F1-F0 ATP synthase. Energy release is effected here by letting ions flow across the membrane through transmembrane proteins, such as ATP synthase. This is very analogous to the way a hydroelectric plant uses the kinetic energy of water supplied by gravity. Here the role of water is played by H^+ ions. However, instead of gravity, electrostatic and entropic forces drive the flow of ions and the ATP-synthase plays the role of the turbine.

How do all of these forms of energy contribute to the synthesis of the biological molecules that make up the cell? Conversion of a carbon source such as glucose into carbon skeletons of any of the other necessary organic molecules (amino acids, nucleotides, phospholipids, etc) proceeds through an intricate series of stepwise chemical transformations where metabolic enzymes catalyze the rearrangement of atoms within a substrate molecule, the cleavage of covalent chemical bonds in the substrate and the formation of new covalent bonds. Fig. 5.4 gives a schematic of the chain of reactions connecting the food source (for example, glucose) to the final product, namely, two cells. The product of one biochemical reaction in this kind of biochemical pathway goes on to be the substrate in the next reaction. In the diagram, this chain of reactions is denoted by “fueling products”, “building blocks”, “macromolecules” and “structures”. All organic molecules within the cell are linked to one another through an intricate and highly-interconnected network or web of metabolic reactions. The overall architecture of the network is dominated by the existence of critical nodes represented by important intermediate molecules (such as the precursor metabolites shown in fig. 5.4) that can be in turn converted into various final products.

A highly schematized diagram of one of the most important metabolic net-

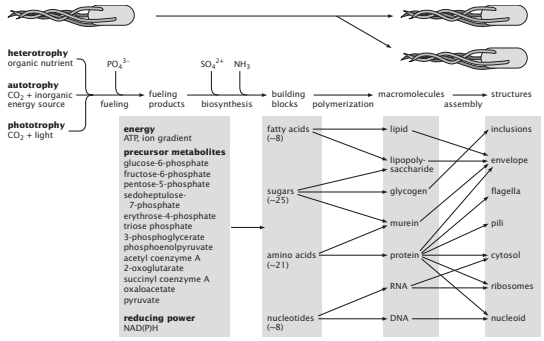

Figure 5.4: Energy and mass costs to make a new bacterial cell. This diagram illustrates the flow of materials and energy required for bacterial duplication. Nutrients are taken from the environment, either organic molecules provided by other organisms or carbon dioxide and light in the case of photosythetic bacteria. Together with a few inorganic ions such as phosphate, sulfate and ammonium, the carbon sources consumed by the bacterium are converted into precursor metabolites and then into the fatty acids, sugars, amino acids and nucleotides that are used to build macromolecules. The macromolecules are further assembled into large scale structures of the cell. The numbers shown in the "building blocks" column correspond to the rough number of molecular building blocks of each type. (Adapted from M. Schaechter *et al.*, Microbe, Washington DC, ASM Press, 2006.)

works (glycolysis) of *E. coli* and other cells is shown in fig. 5.2. After the six-carbon molecule glucose is taken up by the bacterial cell, it is broken down by the process of glycolysis to form two copies of the three-carbon molecule pyruvate. This overall set of transformations takes place through ten distinct chemical steps as shown at the molecular level in fig. 5.5. Pyruvate, in turn, can be used to synthesize a variety of amino acids or fatty acids. As glucose is broken down to form pyruvate, some of the chemical energy stored in its covalent bonds is used to synthesize ATP and NADH. These high energy carrier molecules can then donate their energy to drive forward biosynthetic reactions that are not intrinsically energetically favorable.

A useful way to envision the energetic transformations during glycolysis is to picture each molecular species as having a characteristic energy and as glucose goes through its series of transformations, the molecule travels up and down on a hilly energy landscape. This idea is related to the treatment of enzymes that we introduced in fig. 3.24 (pg. 161). There we showed that the bent form of the substrate molecule resides at a slightly higher energy level than the straight form. For real molecules such as the intermediates in the glycolytic pathway, the energy of particular species depends not only on its molecular structure but also on its concentration in the cell. We will consider the exact definition of these molecular energies in chap. 6.

This general framework has prepared us to estimate the amount of glucose needed to provide energy for a single round of cell division compared to the amount required to provide structural building blocks. Accurate estimation of this number requires a detailed examination of each of the biosynthetic pathways in *E. coli* and tabulation of the energy consumed or liberated at every enzymatic step. Such calculations have been undertaken by brave biochemists. For our purposes, we will instead attempt a cruder but much simpler scheme in which we posit "typical" costs and gains associated with each class of molecule as counted in chap. 2 (proteins, nucleic acids, etc.).

- **Estimate: The Energy Budget Required to Build a Cell.** Our estimates of the inventory of a cell given in chap. 2 (pg. 59) provided a feeling for the numbers of each kind of macromolecule needed to make a new cell. If the cell has glucose as its sole carbon source, the carbons in the sugar need to be taken apart and reassembled as useful building blocks such as amino acids and nucleotides which make the construction of the macromolecules of the cell possible.

  The concept of the estimate we undertake here is represented by the analogy of considering the cost of constructing a building. Overall costs can be subdivided into the costs of the physical construction materials themselves and the cost of the labor required to put them together. In the cell, both the construction material (in the form of organic molecules) and the energy source ultimately are derived from nutrients taken up by the cell. As with our earlier estimates regarding *E. coli* in chaps. 2 and 3, we will consider cells growing in a medium where glucose is its sole source of carbon and biosynthetic energy. Previously, in chap. 2, we estimated that the

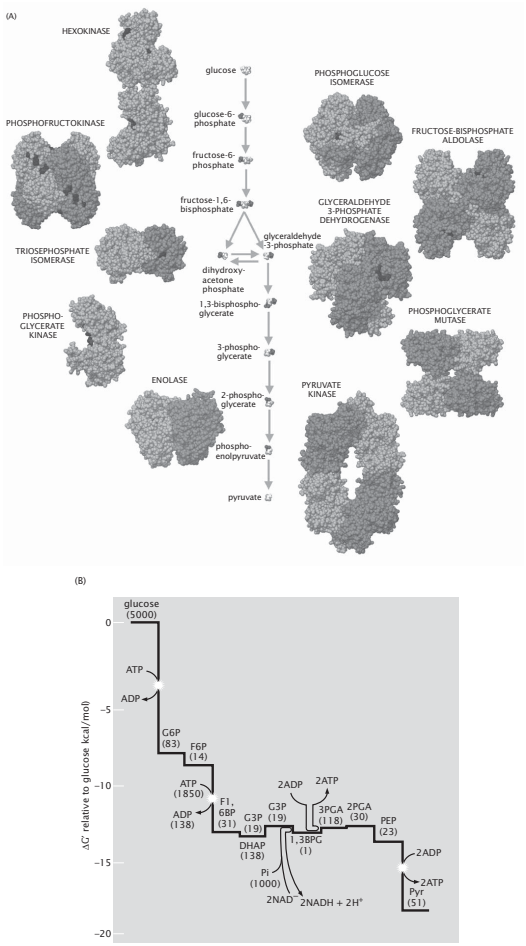

Figure 5.5: The molecules of the glycolytic pathway and the energy landscape for their transformations. (A) By a series of ten chemical steps, one molecule of glucose is converted into two molecules of pyruvate. Each step is catalyzed by a specific enzyme, all of which are shown here as spacefilling models. The enzymes are substantially larger than the small-molecule substrates on which they act. (Illustration from David Goodsell) (B) The downward energetic progression of the glycolytic pathway is illustrated graphically where each horizontal bar represents the relative energy level of one of the glycolytic intermediates. Overall, the transformation of glucose to pyruvate is extremely energetically favorable. Some of the energy liberated during each of these transformation steps is captured by the high-energy carrier molecules, ATP and NADH. Three of the steps in glycolysis have such large negative energy changes associated with them that they are considered irreversible: phosphorylation of glucose to glucose-6-phosphate, phosphorylation of fructose-6-phosphate to fructose-1,6-bisphosphate, and conversion of phosphoenolpyruvate to pyruvate with the concomitant synthesis of ATP. Many of the other steps take place with little net energy change. (A, courtesy of David Goodsell; B, adapted from C. K. Mathews *et al.*, Biochemistry, San Francisco, Addison Wesley Longman, Inc., 2000.)

total number of carbon atoms required to construct a new *E. coli* cell is approximately 10^{10}. At 6 carbon atoms per glucose molecule, this means the cell must take on roughly 2×10^9 glucose molecules simply to provide the raw construction materials for doubling its mass so that it can divide. How many additional glucose molecules must the cell take up to convert to the biosynthetic energy required to refashion all those carbon skeletons into cellular material?

In *E. coli*, there are seven major classes of macromolecular components whose synthesis we must consider: protein, DNA, RNA, phospholipid, lipopolysaccharide, peptidoglycan and glycogen. Because each of these kinds of components involves its own elaborate biosynthesis pathways, we must consider them separately. Rather than going through all, we will start with the illustrative example of proteins, briefly discuss DNA and RNA and then assert the final outcome of the energy budget calculation.

For biosynthesis of proteins when glucose is the sole carbon source, the glucose carbon skeletons must first be converted into amino acids, and then those amino acids must be polymerized to form new proteins. As can be easily appreciated by a glance at fig. 2.23 (pg. 96), amino acids vary significantly in their structure and some are more complicated to synthesize than others. Over the past 100 years, the metabolic pathways for synthesis of each of these amino acids has been determined and the responsible enzymes identified using methods like the pulse-chase method (see fig. 3.3 on pg. 127) and generation of auxotrophic mutants, which need to be fed precursor molecules to survive. All of the amino acid synthetic pathways are connected directly or indirectly to the glycolytic pathway shown in fig. 5.5. Indeed, all metabolites in the cell are connected to all others through the elaborate metabolic web which can be graphically represented in a summary diagram that covers most of a wall and which resembles the Tokyo subway map but is substantially more intricate. Alanine, for example, can be synthesized from pyruvate in a single step, by a single enzyme. Tryptophan, in contrast, requires the coordinated action of twelve enzymes. The net synthesis cost for making each of the amino acids is summarized in table 5.1. For purposes of calculating the energy budget, we must also take into account the fact that some amino acids are much more abundant than others. For example, glycine is approximately tenfold more abundant than tryptophan. By multiplying the energetic cost to make each amino acid by its relative abundance in the cell, we can estimate that the average energetic cost to synthesize an amino acid is roughly 1.2 ATP equivalents for cells growing aerobically and 4.7 ATP equivalents for cells growing anaerobically.

After the amino acids are synthesized, they must be strung together to make proteins. This painstaking assembly work requires a large input energy to the tune of roughly four ATP equivalents for each amino acid, including the cost to attach the amino acids to tRNAs and to power the movement of the ribosome. As a result, the cost for adding each amino

Amino Acid	Abundance	ATP equivalent (anaerobic)	ATP equivalent (aerobic)
alanine (A)	9.8×10^7	-1.0	1.5
arginine (R)	5.6×10^7	8.5	11.0
asparagine (N)	4.6×10^7	3.0	5.5
aspartate (D)	4.6×10^7	0.0	2.5
cysteine (C)	1.7×10^7	11.5	14.0
glutamate (E)	5.0×10^7	-3.5	-1.0
glutamine (Q)	5.0×10^7	-2.5	0.0
glycine (G)	12.0×10^7	-2.5	0.0
histidine (H)	1.8×10^7	7.0	7.0
isoleucine (I)	5.5×10^7	8.5	13.5
leucine (L)	8.6×10^7	-10.5	-3.0
lysine (K)	6.5×10^7	6.0	11.0
methionine (M)	2.9×10^7	24.5	27.0
phenylalanine (F)	3.5×10^7	2.0	7.0
proline(P)	4.2×10^7	2.5	5.0
serine (S)	4.1×10^7	-2.5	0.0
threonine (T)	4.8×10^7	7.0	9.5
tryptophan (W)	1.1×10^7	7.0	9.5
tyrosine (Y)	2.6×10^7	-0.5	4.5
valine (V)	8.1×10^7	-2.0	3.0

Table 5.1: Amino acid abundance and energetic cost for making the amino acids under both aerobic and anaerobic growth conditions. A negative value implies that synthesis of the amino acid from glucose is favorable so energy is generated rather than consumed. (Data from F. C. Neidhardt *et al.*, Physiology of the Bacterial Cell, Sunderland, Sinauer Associates, Inc, 1990 and M. Schaechter *et al.*, Microbe, Washington DC, ASM Press, 2006.)

acid is 5.2 ATP equivalents corresponding to the 1.2 ATP equivalents it costs to make the average amino acid and the 4 ATP equivalents it takes to add the amino acid onto the peptide chain. Multiplying this by the total number of amino acids that need to be strung together to make a cell, we find

$$\text{protein energy cost} \approx 5.2 ATP \times 300 \times 3 \times 10^6 \approx 45 \times 10^8 \text{ATP equivalents.} \tag{5.2}$$

We have taken 300 as the number of amino acids in the "average" protein, and used an approximate number of 3×10^6 proteins per bacterium.

It is possible to perform similar calculations for each of the other six classes of macromolecules. Interestingly, while we found for proteins that synthesis of the amino acid precursors is relatively energetically inexpensive and assembly into proteins, relatively costly, the situation is the opposite for DNA and RNA. Here, the energy required to synthesize a nucleotide triphosphate precursor is large, on the order of 10-20 ATP equivalents depending upon growth conditions, but the additional cost required to assemble the polymers is small. Whereas amino acid synthesis consumed less than a quarter of the total energy required to make proteins, nucleotide synthesis requires nearly 90 % of the energy required to make nucleic acids.

Table 5.2 summarizes the biosynthetic cost for each of the major classes of macromolecule. As noted above, the exact numbers will vary depending upon growth conditions, but these will serve as a reasonable estimate for our standard *E. coli* growing under standard conditions. Recall that an *E. coli* cell must take up roughly 2×10^9 glucose molecules for building materials to double its mass. Growing with maximum efficiency under aerobic conditions a single molecule of glucose can generate up to 30 molecules of ATP with carbon dioxide as the waste product. Comparing the total amount of biosynthetic energy required by adding up all of the components in table 5.2, about 2×10^{10} ATP equivalents are required or about 6×10^8 molecules of glucose. Thus, it requires about one-third as much glucose just to pay for labor as it does to provide the actual building materials for constructing a new cell. Under less efficient growth conditions the cost of biosynthesis can actually exceed the cost of materials by as much as tenfold. Furthermore, in the estimates above we have ignored the fact that macromolecules are constantly degraded and replenished inside the cell. This will surely increase the overall energy budget, but will not affect our estimates by more than an order of magnitude.

We have seen that much of the useful energy available to cells is stored in the energy of phosphate bonds. This energy can be released in a variety of different ways for processes such as those associated with the central dogma, for cell motility and for setting up ion gradients across cells. One of the tools we will use to examine these transformations of energy is the theory of mechanical and chemical equilibrium as embodied in the laws of thermodynamics and statistical mechanics. That is, the calculus of equilibrium to be set up in this and the

Class	Biosynthetic Cost (aerobic) - ATP equiv.
protein	1.2×10^{10}
DNA	3.5×10^{8}
RNA	1.6×10^{9}
phospholipid	3.2×10^{9}
lipopolysaccharide	3.8×10^{8}
peptidoglycan	1.7×10^{8}
glycogen	3.1×10^{7}

Table 5.2: Biosynthetic cost in ATP equivalents to synthesize the macromolecules of the cell

following chapter is an abstract tool that permits us to predict the direction and extent of important energy transactions in the cell. With a sense of the energy scales associated with important biological transformations now in hand, we turn to an analysis of the tools used to characterize these transformations.

5.2 Biological Systems as Minimizers

In the previous section, we considered the energetics of macromolecular synthesis, only one of many activities undertaken by busy cells. In a more general sense, our consideration of the role of energy and energy transformations in the processes of life can also be applied to many other kinds of problems beyond biosynthesis. What determines the shape of a red blood cell? Given a particular oxygen partial pressure in the lungs, what is the fractional binding occupancy of the hemoglobin within red blood cells? How much force is required to package the DNA within the capsid of a bacteriophage? What fraction of Lac repressor molecules in an *E. coli* cell are bound to DNA and what is the probability that one such molecule is bound specifically? Each of these questions is ultimately a question about energy transactions and can be couched in the form of a minimization problem in which we seek the least value of some function. For example, as will be shown in chap. 11, the question of the shape of red blood cells will be formulated mathematically as the problem of minimizing the free energy of the membrane and associated architectural filaments which bound the cell. Similarly, our discussion of chemical equilibrium and equilibrium constants for problems ranging from the occupancy of hemoglobin by oxygen to the binding of Lac repressor to DNA will be founded upon equality of chemical potentials, which is a simple consequence of minimizing the free energy. As we will see, questions in both mechanical and chemical equilibrium can be stated in the language of minimization principles.

The remainder of this chapter commences our efforts to develop mathematical models of biological energy transformations viewed through the prism of minimization principles. A complementary view of transformations will be developed in chap. 15, when we explicitly consider rates and dynamics. The key

point of the present discussion is how our understanding of the equilibrium configurations of systems ranging from DNA-protein complexes to bones under stress can be built around the idea of minimizing an appropriate energy quantity. The quest to develop this intuition will lead us to the mathematics of the calculus of variations and will culminate in the elucidation of Gibbs' calculus of equilibrium in the form of the principle of minimum free energy.

5.2.1 Equilibrium Models for Out of Equilibrium Systems

Given that living organisms are one of the quintessential examples of systems that are out of equilibrium, it is natural to ask to what extent the tools of equilibrium physics are of any use in biology. Perhaps surprisingly, in fact there is a wealth of examples where the use of equilibrium ideas is well justified.

Equilibrium Models Can Be Used for Nonequilibrium Problems If Certain Processes Happen Much Faster Than Others

The decision of whether an equilibrium description is appropriate for a given problem often comes down to a question of time scales. As a simplest example, we examine the validity of treating a cell as though it is in mechanical equilibrium. Mechanical equilibrium is characterized by the absence of any unbalanced forces in a system. However, a more nuanced description of mechanical equilibrium appropriate for some biological problems is the idea that all of the forces in the system are balanced on the time scales at which the biological process is taking place. For example, as a cell crawls across a surface, the cytoskeleton is pushing on parts of the plasma membrane. In some cases, the response of the membrane can be thought of as so fast on the time scale of the underlying cytoskeletal dynamics that at every instant the membrane has equilibrated mechanically with respect to the forces produced by the cytoskeleton.

Similar arguments apply in the case of chemical equilibrium. For concreteness, consider the reaction,

$$A \underset{k_-}{\overset{k_+}{\rightleftarrows}} B \xrightarrow{r} C \ . \tag{5.3}$$

where we have assumed for simplicity that the backwards reaction from C to B has a negligible rate (this approximation is useful for thinking about processes such as transcription and translation). The basic argument being made verbally in this section and mathematically later (in chap. 15) is that if the rates associated with the conversion between A and B are sufficiently fast in comparison with the rate at which B is depleted as a result of conversion into C, then we can think of the reaction $A \rightleftharpoons B$ as being in equilibrium. The concrete signature of this rapid preequilibrium is that the amount of A and B occurs in a fixed ratio determined by the ratio of the forward and backward rates for the reaction.

The outcome of this kind of analysis is shown in fig. 5.6. The key point of the calculation is embodied in the fact that after an initial transient period, the ratio $[A]/[B]$ (we use the notation $[A]$ to mean "concentration of A") is constant for all subsequent times even though the absolute number of A and B

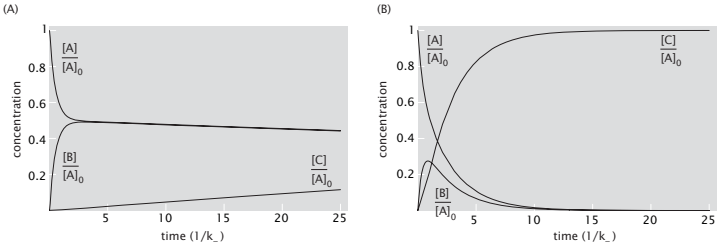

Figure 5.6: Rapid approach to equilibrium of a subprocess. Plot of the time dependence of the concentrations in the reaction $A \rightleftharpoons B \rightharpoonup C$ of $A(t)$, $B(t)$ and $C(t)$. (A) For the case in which the rate for converting B to C is slow in comparison to the rates for the reaction between A and B, after an initial transient period, A and B reach their equilibrium values relative to each other for the remainder of the process. (B) Plot showing the case in which there is no rapid pre-equilibrium.

molecules is decreasing over time. This fixed ratio is the equilibrium constant for the reaction $A \rightleftharpoons B$. If the rate of conversion to the product C is too fast, the rapid preequilibrium condition is no longer satisfied yielding a situation like that shown in fig. 5.6(B).

As will become clear in subsequent chapters, there are many cases in which the *numbers* associated with various kinetic processes justify the use of equilibrium arguments like those to be developed in this chapter. The mindset that justifies this approach is one of time scales; namely, when the rate constants for some initial reaction in a series of reactions are fast (in a way that can be evaluated mathematically), then that reaction can be treated as an equilibrium reaction.

5.2.2 Proteins in "Equilibrium"

To set the stage concretely for some of the ways in which we will invoke equilibrium models to think about problems of biological interest, fig. 5.7 shows some examples where we treat proteins from an equilibrium perspective. In some instances (figs. 5.7(A) and (B)), our analysis can be built strictly around the notion of mechanical equilibrium. The examples of chemical equilibrium begin with the claim that it is useful to think of the folded state of a protein as a free energy minimizer. The next example of protein properties from an equilibrium perspective is the treatment of the way in which the charge state of a protein depends upon the pH of the solution. Here the idea is that the charge state of the protein reflects a competition between the entropy gained by permitting charges to wander in solution and the corresponding energy cost associated with removing those charges from their protein host. Yet another example that will

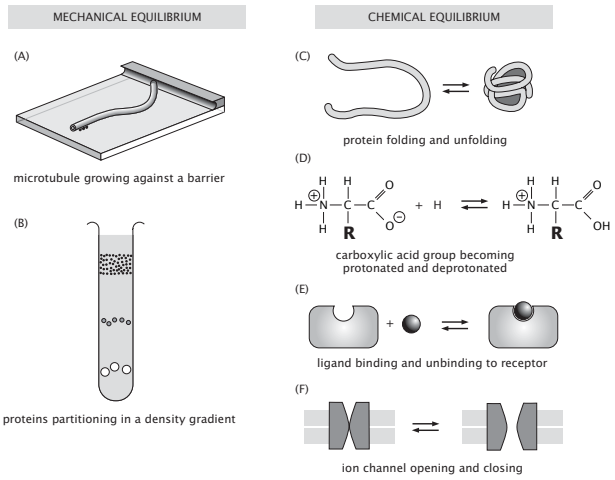

Figure 5.7: Proteins in equilibrium. Schematic showing many examples of the way in which proteins are approximated as being in equilibrium.

arise repeatedly throughout the book is the treatment of binding where in the case of a protein we can think of it as being complexed with some ligand of interest. Here too the basic picture is one of an interplay between the entropy associated with free ligands and the energetic gain they garner as a result of being bound to their protein host. A final example where at times it is convenient to think of a protein as being in equilibrium is when that protein coexists in an active and inactive form and where the relative probability of these states is dictated by some external influence. For example, as will be discussed in chap. 7, phosphorylation of a protein can shift it from an inactive to an active state. A second example also to be examined in chap. 7 is the gating of ion channels. Here too, channel gating can sometimes be treated as an equilibrium problem where some tuning parameter such as the external tension in the membrane or an applied voltage can alter the probability that the channel is open.

Protein Structures Are Free Energy Minimizers

As a result of the sequencing of an ever-increasing number of genomes, the challenge to assign meaning to that genomic information has also increased. In particular, with the genetic sequence in hand, what can be said about the structure and function of the various proteins coded for in these genomes? Assuming that a particular gene within a genome has been identified, the question can be posed differently. We have already seen in fig. 1.4 (pg. 30) that the languages of nucleic acids and proteins are related by the universal genetic code which tells us how to translate the DNA sequence into a corresponding amino acid *sequence*. However, once the relevant amino acid sequence has been determined, we still

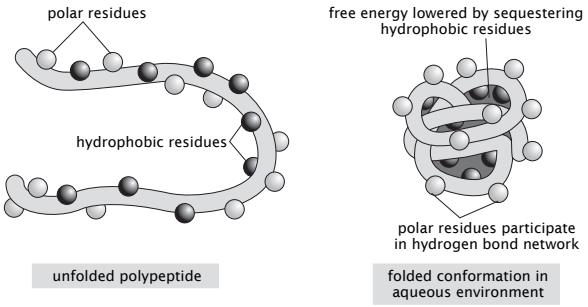

Figure 5.8: Schematic of the way in which protein folding sequesters hydrophobic amino acids while leaving their polar counterparts in contact with the surrounding solution.

don't know the structure implied by that given primary sequence.

A first step in solving this problem corresponds to answering the question: of all of the possible ways that that particular set of amino acids can fold up, which has the lowest free energy? From an intuitive perspective, we already possess heuristic ideas for thinking about protein folding as illustrated in fig. 5.8. In particular, the key idea is that certain amino acid side chains can happily participate in the hydrogen bonding network of the surrounding solution, while those residues with hydrophobic side chains are sequestered from the surrounding solution. From a quantitative perspective, these structural preferences have a corresponding free energy benefit.

A second way in which proteins are conveniently viewed from the equilibrium perspective has to do with their charge state. As the pH of the solution is varied, the charge on different amino acid residues in a particular polypeptide chain will vary. An example from the amino acid glycine is shown in fig. 5.9. We can think about the liberation of charge in solution as a result of the competition between the energetic favorability of keeping unlike charges near to each other and the entropic benefit of letting the charges stray from their protein host. The reason for bringing up these protein examples is to highlight the way in which equilibrium ideas are often a starting point for the analysis of important biological problems.

5.2.3 Cells in "Equilibrium"

We have seen that there are many circumstances in which the molecules and macromolecular assemblies of the cell can be viewed from an equilibrium perspective. At the larger scales representative of cells themselves, there are many cases in which we can consider some particular part of the cell (such as the

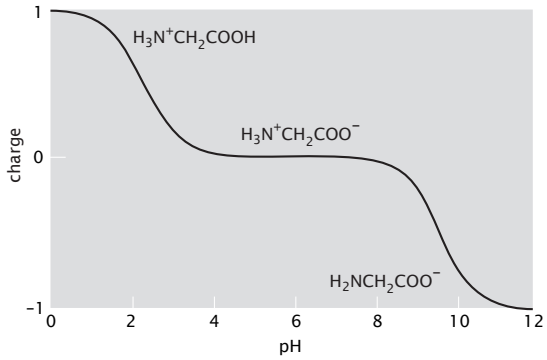

Figure 5.9: Titration curve showing the charge state of the amino acid glycine as a function of the pH of the solution. Low pH corresponds to a high concentration of H^+ ligands resulting in saturation of the glycines. (Adapted from K. Dill and S. Bromberg, Molecular Driving Forces, New York, Garland Press, 2003.)

membrane) as being in local mechanical or chemical equilibrium. One example of this kind of thinking is that of the equilibrium shapes of red blood cells. As shown in fig. 5.10, the shapes of such cells have been precisely characterized experimentally and can similarly be calculated.

As will be introduced in the remainder of the chapter and driven home as a key part of the rest of the book, in problems of free energy minimization there are two key steps: first, the selection of a class of competitors and second, the determination of the free energy associated with each such competitor. In the setting of red blood cells of interest here, the class of competitors is the set of all shapes satisfying two geometric constraints, namely, that the overall area of the red blood cell surface be the same from one shape to the next and also, that the volume enclosed by that area be the same. Fig. 5.10 shows in the right panel the shapes that have the lowest free energy for different choices of a control parameter which is the difference in area between the two leaflets of the membrane.

5.2.4 Mechanical Equilibrium From a Minimization Perspective

As argued above, there are a variety of different biologically interesting examples which, when examined in physical terms, amount to problems in minimization. One class of problems which can be thought of in this way center on mechanical equilibrium.

The Mechanical Equilibrium State Is Obtained by Minimizing the Potential Energy

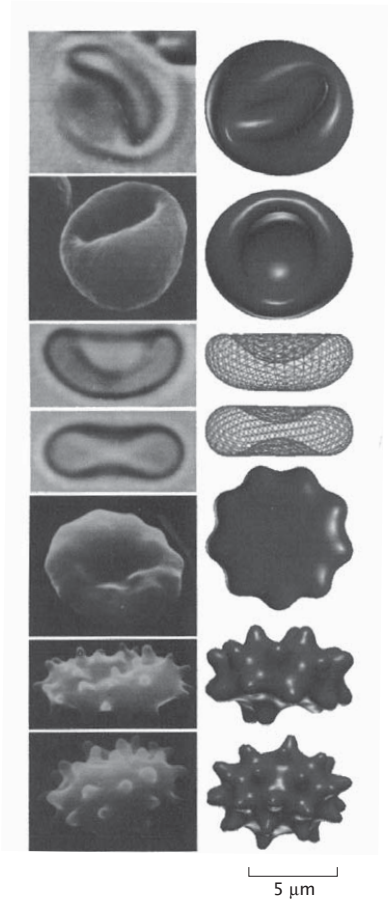

Figure 5.10: Red blood cell shapes. The left column shows shapes of red blood cells as observed experimentally and the right column shows calculations of the shapes. (Adapted from G. Lim *et al.*, *Proc. Nat. Acad. Sci.*, 99:16766, 2002.)

One way to think about the mechanics of bodies at rest is Newton's first law of motion, namely, that if a body is in equilibrium, then there are no unbalanced forces on that body. Stated mathematically, the condition of translational equilibrium is

$$\sum_i \mathbf{F}_i = 0, \tag{5.4}$$

where $\mathbf{F}_i$ is the ith force acting on the body. The use of the bold face letter in writing the forces $\mathbf{F}_i$ reflects the fact that the force is a vector quantity. For example, if we consider the hook shown in fig. 5.11, there is a force acting on that hook due to a spring and a second force due to the hanging weight and these forces balance each other. Their force vectors have the same length and point in opposite directions. However, it is not always most convenient or enlightening to consider equilibrium problems in the vectorial language of forces. The alternative that will often be favored throughout the remainder of the book is the equivalent formulation of the problem of mechanical equilibrium as one of minimization. The principle of minimum potential energy asserts that the mechanical state of equilibrium is the one (out of all of the possible alternatives) that has the lowest potential energy.

To write the equilibrium of the system shown in fig. 5.11 in terms of energy, we can write the potential energy as a sum of two terms, one of which captures the energy of the stretched spring and the other of which describes the "loading device", namely, the lowering of the weight. Given these concepts, the potential energy can be written as

$$U(x) = \underbrace{\frac{1}{2}k(x-x_0)^2}_{\text{PE of spring}} - \underbrace{mg(x-x_0)}_{\text{PE of weight}}, \tag{5.5}$$

where x_0 is the length of the spring when it is unstretched and will also serve as our zero point for the potential energy of the hanging weight. We use the label "PE" for potential energy. These two terms are shown in fig. 5.11(B) and we see that their sum has a minimum (i.e. the equilibrium point). To actually find the point x_{eq} at which the minimum occurs, we note that at x_{eq}, the slope of the function $U(x)$ is zero - this condition corresponds to the mathematical statement $dU/dx = 0$ as will be shown in more detail below. Minimization of the potential energy in this case corresponds physically to finding that choice of the displacement x_{eq} that leads to the lowest energy and is determined by the condition

$$\frac{dU}{dx} = k(x_{eq} - x_0) - mg = 0. \tag{5.6}$$

This result can be rewritten as

$$x_{eq} = x_0 + \frac{mg}{k}, \tag{5.7}$$

which tells us the size of the excursion made by the spring about its equilibrium postion. The result jibes with our intuition in the sense that larger weights (mg

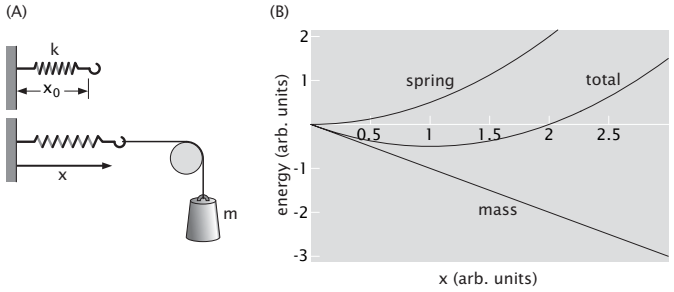

Figure 5.11: Mechanical equilibrium as potential energy minimization. (A) Schematic showing how the mechanical equilibrium of a system can be thought of from the point of view of minimization of the potential energy. (B) Potential energy of spring and weight and their sum as a function of the displacement.

big) leads to larger excursions and a stiffer spring (k large) results in a smaller excursion.

The idea of the potential energy of the loading device introduced above is pervasive and will be used repeatedly in the book. As shown in fig. 5.12 we will think about the energy associated with deforming cantilevers such as in the atomic-force microscope, polymers and membranes. In all of these cases, when we write down the total energy (or free energy) of the system, we will have to account for the way in which the deformation of our system of interest (i.e. the polymer or membrane) leads to an attendant change in the energy of the loading device, as depicted here by the lowering of a weight.

As noted as early as fig. 1.12 (pg. 45) in chap. 1, "springs" show up in a surprising variety of circumstances. One example that we are particularly fond of is the use of laser light to make a spring in the form of optical tweezers (introduced in fig. 4.11 on pg. 198) such as shown in fig. 5.13. In particular, we consider the case of a bead in an optical trap which is subject to a load due to a piece of tethered DNA. We ask what displacement the bead suffers in the trap as a result of the applied load? We can write down the potential energy function in the form

$$U(x) = \frac{1}{2}k_{trap}x^2 - Fx, \tag{5.8}$$

where we have assumed that the optical trap can be treated as a spring with stiffness k_{trap} and that the applied force is characterized by a magnitude F. If we now seek the energy minimizing choice of x, obtained through solving $dU/dx = 0$, we find that the equilibrium displacement in this case is given by

$$x_{eq} = \frac{F}{k_{trap}}. \tag{5.9}$$

The characteristic scales for an experiment like this are forces in the range of tens of piconewtons. This kind of experiment permits the measurement of

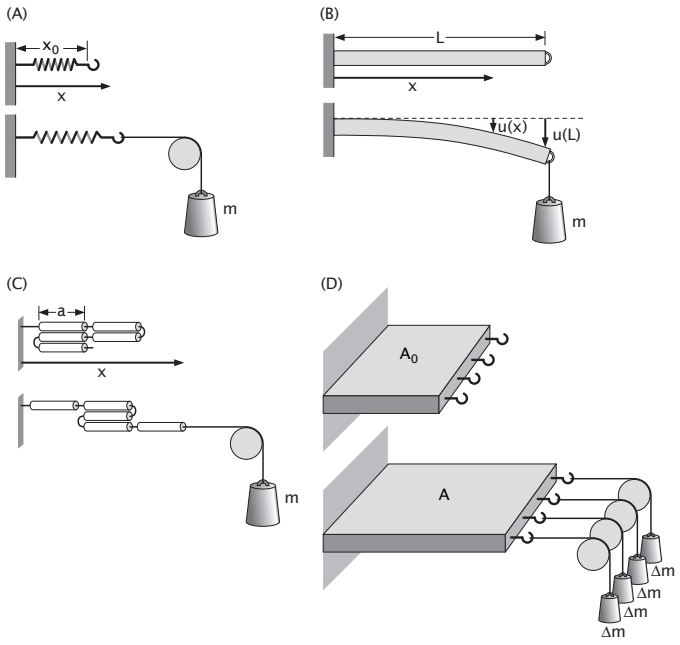

Figure 5.12: Mechanics of loading devices. (A) mass-spring system, (B) beam under the action of an applied force, (C) polymer chain subjected to a load, (D) membrane subjected to an applied tension.

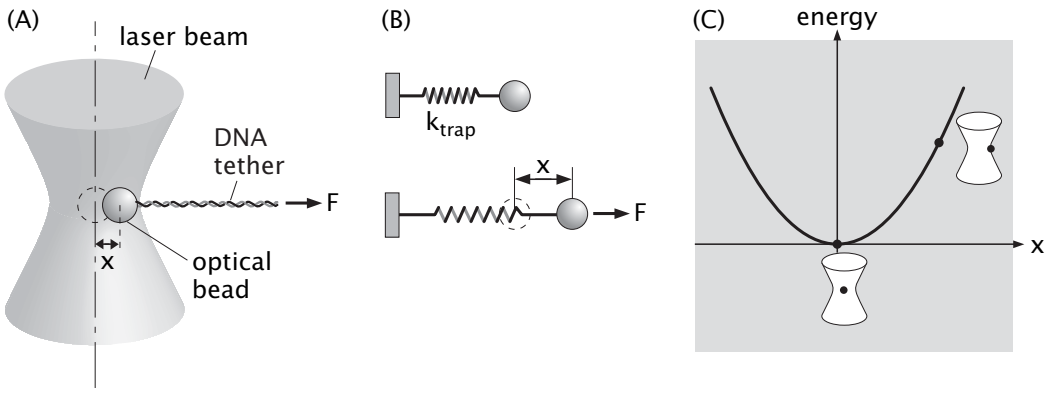

Figure 5.13: Representation of an optical trap as a mass-spring system. (A) Schematic showing how force removes the bead from the center of the trap. (B) Replacement of the optical problem with a corresponding effective spring. (C) Energy of the bead in the trap as a function of its position in the trap. Note that the energy is only quadratic for sufficiently small displacements.

a variety of interesting single-molecule properties such as the force-extension characteristics of macromolecules (DNA, RNA, proteins, etc..) and the force-velocity characteristics of molecular motors. Several examples of this kind of experiment were already introduced in fig. 4.12 (pg. 198). An example of force-extension data for DNA obtained by using single-molecule methods (in this case a magnetic tweezers rather than an optical tweezers) is shown in fig. 5.14. This experiment allows for applying a range of different forces to DNA and examining the corresponding elongation of the DNA molecule. At low forces, the extension increases linearly with force while at high forces the extension saturates since the molecule has been stretched to its full contour length. As will be seen in chap. 8, these experiments can be compared directly with our theoretical understanding of DNA mechanics.

These examples on mechanical equilibrium provide an introduction to the key precept of the present chapter which is the idea that equilibrium structures are minimizers of potential energy (zero temperature) or free energy (finite temperature). To perform such a minimization, we need to write the energy or free energy in terms of some set of variables that characterize the geometric state (i.e. the structure) of the system. Once we have written the energy or free energy in terms of the parameters characterizing the system, then our task is reduced to the mathematics of determining which out of all of the various structural competitors leads to the lowest value of the potential energy or free energy. For the simple mass-spring systems introduced in this section, the minimization required the evaluation of a derivative. However, more generally we must address the mathematical question: given a function, out of all of the possible competitors, how do we find the one that minimizes the value of that function?

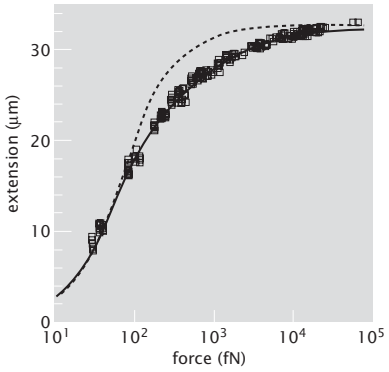

Figure 5.14: Force-extension curve for double-stranded DNA. Data for force vs extension for double stranded DNA from λ-phage (with the DNA molecule here resulting from linking two such molecules for a total length of 97 kbp) illustrating the distinction between the freely-jointed chain model (dotted line) and the worm-like chain model (solid line). The freely-jointed chain model will be discussed in detail in chap. 8 and the worm-like chain model will be discussed in chap. 10. (Adapted from C. Bustamante *et al.*, *Science*, 265:1599, 1994.)

5.3 The Mathematics of Superlatives

The search for extrema is a mathematical embodiment of the human instinct for superlatives. In casual conversation, rarely an hour goes by without injecting words such as "best" and "worst" into our speech. Our technologies similarly reflect the pressure to make things faster, smaller, lighter, safer, etc. The development of modern mathematics included tools for finding functions that could be characterized by superlatives such as biggest and smallest. The present section is a mathematical excursion which aims to show how to replace the verbal and intuitive case-by-case discriminations with precise mathematical tools that permit us to search over what amounts to an infinite set of competitors. The reason such a mathematical interlude is necessary is that the study of equilibrium demands that we minimize functions such as the potential energy or the free energy and as a result, we need the mathematics that permits us to effect such minimizations.

5.3.1 The Mathematization of Judgement: Functions and Functionals

The translation from everyday language, where superlatives are characterized by words such as "best", "fastest", etc. to the mathematical form of these same concepts requires the introduction of a scheme for attaching numbers to the

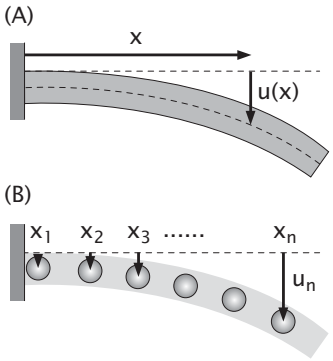

Figure 5.15: Two different representations of the geometry of a beam subjected to a load on its end. (A) continuous representation of the beam geometry, (B) discretization of the geometry of the beam.

degree of "bestness".

As a concrete example in the mathematization of superlatives, we consider the bending of a beam as shown in fig. 5.15. This particular example will arise repeatedly througout the remainder of the book in many disguises. For example, when we think about the geometry of deformed DNA, the buckling of microtubules under force and the use of cantilevers as tools for applying force to macromolecules, in each case we will write the energy of the system in terms of the geometry of these bent beams and will seek the configuration that leads to the lowest energy cost. The question we are interested in answering is: what choice of the displacement function $u(x)$ leads to the lowest value of the potential energy of the beam and the loading device? Note that in this case, the potential energy depends upon the specification of an entire function, $E_{tot}[u(x)]$, where we have introduced the square bracket notation $[\ldots]$ to call attention to the fact that the energy depends upon a function rather than a finite set of parameters. An alternative that sometimes comes in handy is to discretize the geometry of the beam as shown in fig. 5.15(B). In this case, we treat the beam as a series of discrete masses where now there is a set $(u_1, u_2, \cdots u_N)$ of displacements which determine the potential energy. In this case, the energy is a *function* of the unknowns $(u_1, u_2, \cdots u_N)$ and can be written as $E_{tot}(u_1, u_2, \cdots u_N)$.

Functionals Deliver a Number For Every Function They Are Given

When we write the energy in the form $E_{tot}(u_1, u_2, \cdots u_N)$, we are on familiar mathematical turf. A discrete set of parameters u_1, u_2 etc. suffice to describe the geometry of the system and the energy is a *function* of these geometric parameters. In writing the energy in the form $E_{tot}[u(x)]$ we have implicitly

introduced a new mathematical idea (a functional), since in this case it takes a function $u(x)$ to characterize the geometry of the deformed beam and the energy depends upon the function. An energy functional assigns an energy to each configuration, where the configuration itself is characterized by an entire function.

To be concrete in our thinking, fig. 5.16 shows several examples where the free energy depends upon the disposition of the system as characterized by a function. Fig. 5.16(A) shows several different structures for a deformed beam. Each deformed configuration of the beam is described by a *different* function $u(x)$. Further, each such $u(x)$ corresponds to a different energy. The figure shows the energy minimizing configuration as well as a particularly bad guess (i.e. high energy) for the deformed geometry. The energy meter icon aims to show how the energy of the latter configuration is higher than that of the energy minimizing structure. A more subtle example to be taken up again in chap. 9 concerns the distribution of ions around a protein. In this case, the unknown function is $\rho(\mathbf{r})$, the density of ions as a function of position in space. Here too, it is possible to write down a free energy functional that delivers a free energy for each and every guess we might make for the density of ions. Fig. 5.16(B) shows both the free energy minimizing distribution of ions as well as a less than optimal distribution of ions and the energy meter reports their respective overall free energies.

The overarching theme of this section is the idea of a cost function or functional. The key point is that we want to compute some quantity that we are interested in minimizing. In many cases, this "cost function" is the energy or free energy. If that cost function depends upon the disposition of a finite set of parameters such as the $(u_1, u_2, \cdots u_N)$ that characterized our beam represented discretely in fig. 5.15, then indeed, the cost function is a function. On the other hand, if it takes an entire function such as $u(x)$ to characterize the state of the beam, then we have a cost functional since we have to specify a function in order to determine the energy or free energy.

5.3.2 The Calculus of Superlatives

The previous discussion showed the way in which we can cast our ideas about the extent to which some quality or quantity of interest is best or worst, biggest or smallest and so on. This led us naturally to the idea of functions and functionals. Now that we are able to say how good or bad, big or small a particular quantity is as a function of some control parameters (or control function), we pose the question of how to discriminate amongst all the competitors to find the winner. We begin by discussing the implementation of these ideas in the context of ordinary calculus. The more general question of finding the extreme values of functionals is treated in the appendix at the end of the chapter.

Finding the Maximum and Minimum Values of a Function Requires That We Find Where the Slope of the Function Equals Zero

As our first foray into the question of how to cast our search for superlatives

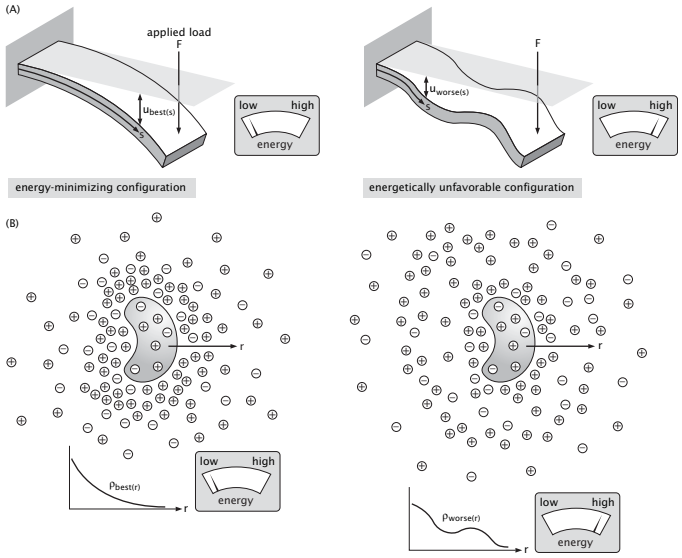

Figure 5.16: Example of two functionals. (A) Energy as a function of the shape, $u(x)$ of a beam. Different shapes have different strain energies. (B) Free energy as a function of the distribution of ions in solution in the vicinity of a protein. The density of ions is characterized by the function $\rho(\mathbf{r})$.

in mathematical terms, we recall a few ideas from the ordinary calculus of maxima and minima. We consider a function $f(u_1, u_2, \cdots u_N) = f(\{u_i\})$, which depends upon the N variables $\{u_i\}$, where we have introduced the notation $\{\}$ to indicate a set of objects. We imagine the variables $\{u_i\}$ are allowed to range over some set of values, and we ask the question: what choice of the values $\{u_i\}$ renders the function $f(\{u_i\})$ maximum or minimum? To be concrete, we remind the reader of the discussion surrounding fig. 5.11. In this case, our minimization problem involves one parameter, the displacement x.

To find the maxima and minima of functions, we find those values of the function for which the slope is zero as embodied in

$$\frac{\partial f}{\partial u_i} = 0, (i = 1, 2, \cdots N). \tag{5.10}$$

That is, our problem amounts to solving the N-equations in N unknowns given by eqn. 5.10. The notation $\partial f/\partial u_i$ refers to the partial derivative and is explained in "The Math Behind the Models" box below. We have said nothing about how we might go about solving such equations, but the prescription for obtaining them is now clear. Though we will not have the space to go into the subtlety of solving such equations for a generic nonlinear problem, we refer the reader to the entertaining cautionary tales of Acton (1990).

- **The Math Behind the Models: the Partial Derivative.** Throughout the book, it will be of interest to find out how functions vary as we change a variable. Often, however, we will be interested in functions that depend upon more than one variable simultaneously. For example, in minimization problems, often the energy (or free energy) will depend upon more than one parameter. For example, the free energy can depend both upon the volume of the system and the number of particles. Another important example is functions $f(x, t)$ that depend upon both position (x) and time (t) simultaneously. For example, we might like to know the deflection of a beam characterized by the function $u(x, t)$ which tells us how much deflection there is a distance x along the beam at a time t. Alternatively, we might interest ourselves in the concentration of some molecule $c(\mathbf{r}, t)$ at every position in space. In this case, the function depends upon four variables since the vector $\mathbf{r}$ is really (x, y, z).

  In these cases, the notion of a derivative is more subtle because we have to say with respect to what variable. The mathematical tool that arises in this case is the partial derivative. The idea is explained in fig. 5.17. The derivative of ordinary calculus tells how a function changes as a result of a small excursion. The partial derivative generalizes that idea by telling us how a function changes when we make an excursion in *one* of the variables in the function while leaving the others constant. An intuitive example from everyday experience is illustrated by walking off of a mountain pass. The shape of a mountain pass is like a saddle. In particular, walking in one direction leads us down whereas walking in a perpendicular direction leads us up the peaks that bound that mountain pass. In these two cases,

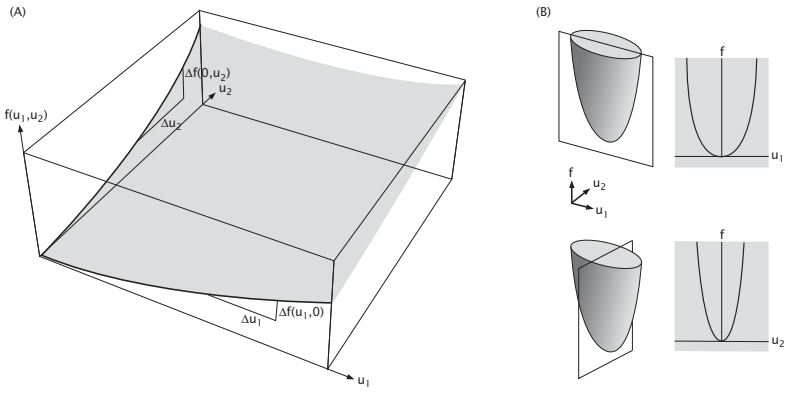

Figure 5.17: Illustration of the concept of a partial derivative. The plot shows the function $f(u_1, u_2)$ which depends upon the variables u_1 and u_2. If u_2 is held fixed, the surface is reduced to a curve and the partial derivative is nothing more than the ordinary derivative familiar from calculus, but on this particular curve.

the partial derivatives actually have different signs since in one case the curve is sloping downward and in the other, it is sloping upward.

If we think of the height of the local topography of a mountain as $f(u_1, u_2)$, where u_1 and u_2 correspond to two perpendicular axes, then the partial derivative tells us how the function changes when we walk along these two directions. These ideas are represented mathematically through the definitions

$$\frac{\partial f(u_1, u_2)}{\partial u_1} = \lim_{\Delta u_1 \to 0} \frac{f(u_1 + \Delta u_1, u_2) - f(u_1, u_2)}{\Delta u_1}, \tag{5.11}$$

and

$$\frac{\partial f(u_1, u_2)}{\partial u_2} = \lim_{\Delta u_2 \to 0} \frac{f(u_1, u_2 + \Delta u_2) - f(u_1, u_2)}{\Delta u_2}. \tag{5.12}$$

For the sake of concreteness in finding minima, we consider the simple example of quadratic functions like those shown in fig. 5.18. The two-dimensional example has the functional form

$$f(u_1, u_2) = \frac{1}{2}(A_{11}u_1^2 + A_{22}u_2^2 + 2A_{12}u_1u_2). \tag{5.13}$$

If we now implement the injunction of eqn. 5.10 (i.e. $\partial f/\partial u_1 = 0$ and $\partial f/\partial u_2 = 0$), we find

$$\begin{aligned} A_{11}u_1 + A_{12}u_2 &= 0 \\ A_{21}u_1 + A_{22}u_2 &= 0, \end{aligned} \tag{5.14}$$

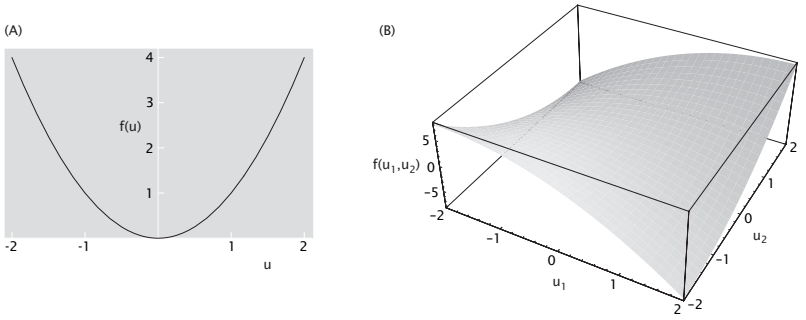

Figure 5.18: Quadratic energy functions. (A) Case of $f(u) = \frac{1}{2}Au^2$ and (B) case of $f(u_1, u_2) = \frac{1}{2}(A_{11}u_1^2 + A_{22}u_2^2 + 2A_{12}u_1u_2)$.

a pair of coupled, linear equations for the minimizing values of u_1 and u_2. Assuming the equations have a unique solution (which is true if the determinant of the matrix A is non-zero), $u_1 = u_2 = 0$ is clearly such a solution, indicating that the function f has a minimum (or maximum) at $(0, 0)$.

5.4 Configurational Energy

In Mechanical Problems, Potential Energy Determines the Equilibrium Structure

Our brief foray into the mathematical machinery used to find minimizers leaves us poised now to ask physically motivated questions of biological interest. In particular, we return to the way in which biological structures can be thought of either as minimizers of the potential energy (this in cases where thermal effects can be ignored) or of the free energy. In this section, we attack the strictly mechanical question of what determines the potential energy of structures and, how the potential energy minimizing structure may be selected from the class of all structural competitors. These ideas will be used in subsequent chapters in thinking about deformations of DNA, cytoskeletal filaments and membranes.

In order to apply the mathematics of superlatives, we must first be able to pass energetic judgement on the relative goodness or badness of a given structure. In particular, to pass this judgement we require an energy function (or functional) which delivers an energy for each and every value of the structural parameters for the structure of interest. The nature of such energy functions forms the backdrop for much of the history of physics.

One class of energy functions with deep significance are those which posit a quadratic dependence of the energy on the departure from equilibrium. The

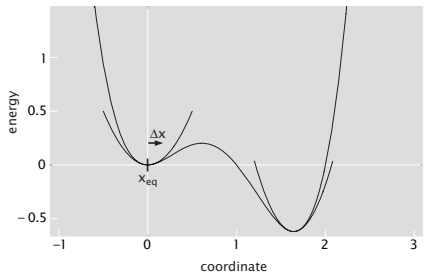

Figure 5.19: Potential energy as a function of the coordinate x. A quadratic representation of that energy landscape is shown in the vicinity of two equilibrium points.

motivation for this class of energy function is the idea that, regardless of the detailed features of a given energy landscape, near equilibria any such function can be treated as a quadratic function of the variables that describe the excursion from equilibrium. Concretely, if we consider the one-dimensional case where the potential energy is of the form $U(x)$ as shown in fig. 5.19 and there is a point of equilibrium at x_{eq}, then we may expand the function $U(x)$ in a Taylor series, keeping terms only up to quadratic order. The idea of the Taylor series is pervasive and is explained in the "Math Behind the Models" box after this section. The Taylor series for our potential is of the form

$$U(x) = U(x_{eq} + \delta x) \approx U(x_{eq}) + \frac{dU}{dx}|_{eq}\delta x + \frac{1}{2}\frac{d^2U}{dx^2}|_{eq}\delta x^2, \tag{5.15}$$

where we have introduced the notation δx to characterize the excursion about the equilibrium point. In this one-dimensional case, δx can be thought of as the distance traveled away from the equilibrium point x_{eq}. This situation is shown in fig. 5.19. Equilibrium demands that $\frac{dU}{dx}|_{eq} = 0$ since at the equilibrium point there are no unbalanced forces and hence we are left with

$$U(x_{eq} + \delta x) \approx U(x_{eq}) + \frac{1}{2}\frac{d^2U}{dx^2}|_{eq}\delta x^2, \tag{5.16}$$

which is of the form $U(x) = \frac{1}{2}kx^2$, where the 'stiffness' of the 'spring' holding the system at equilibrium is given by $k = d^2U/dx^2$. This same idea can be generalized to higher dimensions in which excursions are permissible in multiple directions (for example, on a mountain top, we can choose to walk off in two orthogonal directions and the energy cost of doing so is quadratic in the excursion variables).

One of the most powerful incarnations of the idea developed above is provided by the theory of elasticity which teaches us how to write down the energy of a continuous body, such as a rod or a membrane, as a quadratic function of

the strain which measures the amount of deformation. We take up the elastic energy of deformation (and Hooke's law) in the next section.

- **The Math Behind the Models: The Beauty of the Taylor Expansion** A very important tool invoked in the mathematical analysis of physical models is the use of the so-called Taylor expansion. Series expansions of this kind will be one of our primary mathematical tools in the remainder of the book. The idea is very simple and amounts to replacing a function $f(x)$ in some neighborhood with a simple polynomial. As will be seen repeatedly throughout the book, the virtue of these approximations is that it allows us often to replace intractable nonlinear expressions with simple algebraic surrogates which we can handle analytically and give an intuitive sense of the mathematics.

  The idea of the Taylor expansion is embodied in the simple formula

  $$f(x) \approx a_0 + a_1 x + a_2 x^2 + \dots . \tag{5.17}$$

  The symbol $\approx$ refers to the fact that in the neighborhood of the point x, the left and right sides of this equation are *approximately* equal. Most of the time, we will only keep terms up to second order and as a result, the Taylor series algorithm reduces to the question: what three coefficients a_0, a_1 and a_2 should we use to best approximate the function $f(x)$?

  For concreteness, let's consider the case in which we are interested in the behavior of the function $f(x)$ near $x = 0$. If we set $x = 0$ in both sides of eqn. 5.17, we see that $a_0 = f(0)$. But we already know the function $f(x)$ so all we have to do is find its value at $x = 0$ to obtain the first coefficient. Next, let's take the derivative of both sides of eqn. 5.17 with respect to x. We are left with the equation

  $$f'(x) \approx a_1 + 2a_2 x + \dots . \tag{5.18}$$

  Once again, if we set $x = 0$, we are left with $a_1 = f'(0)$. We can continue to play the same game, this time evaluating the second derivative, with the result

  $$f''(x) \approx 2a_2 + \dots , \tag{5.19}$$

  which leads to $a_2 = f''(0)/2$. This same basic analysis can be carried on indefinitely if one is interested in higher order terms. Most of the time we will be content with the expression

  $$f(x) \approx f(0) + f'(0)x + \frac{1}{2} f''(0) x^2 . \tag{5.20}$$

  The conclusion of this little analysis is that if we want to find a simple quadratic surrogate for our function of interest, all we need to know is the value of the function and its first two derivatives at the point around which we are expanding. An example of this kind of analysis for the case

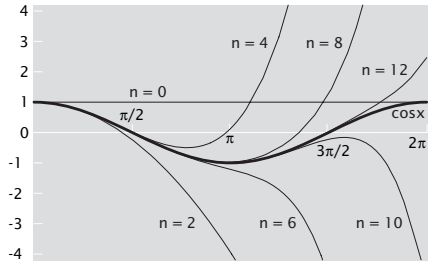

Figure 5.20: Comparison of the function cos x and its Taylor expansion. The curves are labeled by the order of the highest term kept in the Taylor series. For example, $n = 2$ means that the series goes to quadratic order, etc. The cosine function we are approximating is shown in bold for comparison to the approximate expressions.

of cos x is shown in fig. 5.20. In particular, using the rules given above, the Taylor series for this function is given by

$$\cos x \approx 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \frac{x^{10}}{10!} + \cdots . \tag{5.21}$$

Fig. 5.20 compares the function cos x to various approximations based upon the Taylor series. We see that as more terms are included, the approximation is good for a wider range of values of x. Of course, there are mathematical subtleties that arise when considering a generic function, such as the question of convergence of the Taylor series. For example the function $1/(1-x)$ has the Taylor series, $1 + x + x^2 + x^3 + \cdots$, which is finite only for values of x such that $-1 < x < 1$.

5.4.1 Hooke's Law: Actin to Lipids

There Is a Linear Relation Between Force and Extension of a Beam

To see how these ideas about small departures from equilibrium can be applied to continuous bodies of biological significance such as DNA, cytoskeletal filaments and membranes, we begin by examining the elasticity of a stretched rod. This subject will be taken up in detail in chap. 10 and our aim here is to present the conceptual underpinnings of the ideas of elasticity theory. Consider a beam of undeformed length L which is stretched by an amount ΔL as shown in fig. 5.21. The geometric state of deformed objects is most naturally captured in terms of a quantity known as the strain and defined in the current setting as

$$\varepsilon = \frac{\Delta L}{L}. \tag{5.22}$$

The central idea captured by the notion of strain is that adjacent points of the material suffer *different* displacements. In our current example, the displacement of a given point depends upon how far it is from the origin. The result of such relative displacements is that the bonds in the material are stretched as depicted schematically in fig. 5.21. Though our thought experiment considers the case of extension, one can just as easily consider the case of compression in which case $\Delta L < 0$. Note that for simplicity we ignore the small displacements perpendicular to the direction of stretch known as the Poisson effect.

To garner an idea of the mechanical interpretation of these deformations, fig. 5.21 suggests that we can think of the overall macroscopic deformation as imposing the stretching of a huge set of microscopic springs which correspond to the bonds between the atoms making up that beam. We recall that the relation between force and stretch for a spring is given by

$$F = -k\Delta a, \tag{5.23}$$

where k is the spring constant, Δa is the extension of the spring and F is the force it engenders. Macroscopically, this same idea is written as

$$\frac{F}{A} = E\frac{\Delta L}{L}, \tag{5.24}$$

where F is the applied force, A is the cross sectional area of the beam and E is a material property known as the Young modulus which reflects the stiffness of the beam. Note that the Young modulus has units of force/area or energy/volume, since the strain is dimensionless. The quantity F/A is known as the stress and has dimensions of force per unit area.

For the simple model of a beam composed of many microscopic springs shown in fig. 5.21, where two nearby springs are separated by distance a_0, eqn. 5.24 can be derived from eqn. 5.23. In particular, if a force F is applied to the beam it is balanced by all the springs in the cross-section of the beam which each stretch by the same amount. Springs in "parallel" (like resistors) share the load. Since the number of such springs is A/a_0^2, where a_0^2 is area taken up by an individual spring, each spring will be extended by $\Delta a = \frac{F/k}{A/a_0^2}$. This result is a reflection of the fact that n equivalent springs in parallel will each suffer a displacement $(F/k)/n$. Since all the springs that make up the beam have the same extension, the net extension of the beam will be $\Delta L = (L/a_0)\Delta a$, where L/a_0 is the number of springs along the length of the beam, each contributing amount Δa to ΔL. If we substitute for Δa in the last equation the expression derived in the previous one, we arrive at eqn. 5.24, with $E = k/a_0$. In this simple model of an elastic solid, the Young modulus is the ratio of the spring constant associated with a bond between two atoms, divided by the typical distance between them.

The Energy to Deform an Elastic Material Is a Quadratic Function of the Strain

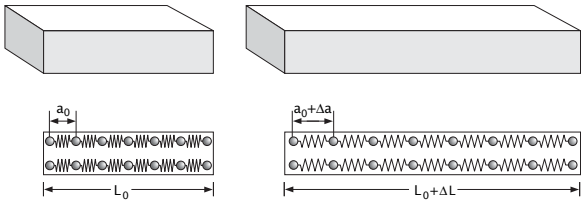

Figure 5.21: Illustration of the interpretation of beam stretching in terms of deformation of microscopic springs. The continuum description of the beam as a deformable solid can be interpreted in terms of the stretching of the individual atomic bonds.

As yet, we have presented Hooke's law as a statement about the forces that result from deforming elastic materials. However, in many circumstances it is more useful to characterize the elastic properties of a deformable material through reference to its energy. If we refer back to simple ideas about springs, the elastic energy stored in a spring by virtue of displacing it a distance Δa from its equilibrium position is given by

$$E_{strain} = \frac{1}{2}k(\Delta a)^2, \tag{5.25}$$

where once again, k is the spring constant. The more general statement that is applicable to an elastic material that has suffered an extensional strain like that shown in fig. 5.21 can be obtained by a divide and conquer strategy in which the material is divided up into a bunch of little volume elements. In each such volume element, we compute the strain energy density and multiply by the volume of that element to obtain the energy for that little chunk of material. By summing (in fact, integrating) over all of the material elements in the material, we find the total strain energy as

$$E_{strain} = \frac{EA}{2}\int_0^L (\frac{\Delta L}{L})^2 dx, \tag{5.26}$$

where A is the cross-sectional area of the beam. This equation can be derived for the simple model of a beam shown in fig.5.21, in the same way eqn. 5.24 was derived above, by adding up the elastic energy of all the microscopic springs that make up the beam.

If we consider the more general case in which the relative stretch is a function of position along the axis of the beam, the energy associated with deformation is given by

$$E_{strain} = \frac{EA}{2}\int_0^L (\frac{du(x)}{dx})^2 dx. \tag{5.27}$$

The key point of all of this is the existence of an energy function that penalizes *relative* changes in length of adjacent material points in a body.

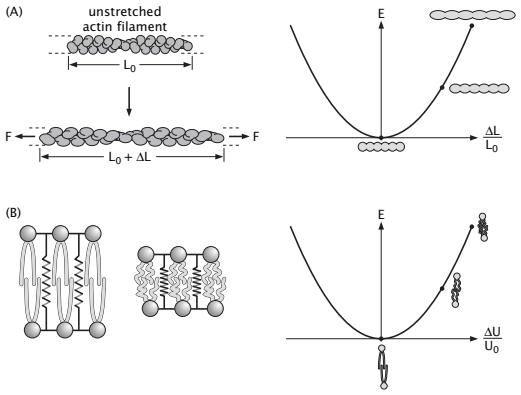

Figure 5.22: Deformation of macromolecular assemblies and the corresponding elastic energy cost associated with these deformations. (A) Schematic of F-actin stretching in response to a force applied along the filament axis. The energy curve shows a quadratic cost to either elongate or shrink the filament relative to its equilibrium length. (B) Schematic of deformation in which the thickness of the lipid bilayer is changed relative to its equilibrium value. The energy curve shows the elastic energy cost to change the thickness of a lipid bilayer from its equilibrium thickness.

In the remainder of the book, we will appeal to elastic arguments like those described above. The virtue of these elastic arguments is that they will permit us to probe the energy cost of processes such as DNA packing (in nucleosomes and viruses), the buckling of microtubules at high force and the deformation of lipid bilayers in the neighborhood of ion channels, to name but a few examples. Several examples are illustrated in fig. 5.22 where we see that analysis of elements of the cytoskeleton will be couched in the language of elasticity theory. In addition, the figure also foreshadows our examination of lipid bilayer membranes in chap. 11. In both cases, the basic idea is the same, namely, that there is a quadratic energy cost associated with small excursions of the system about its equilibrium configuration.

So far, we have argued that in many instances it is convenient to represent mechanical equilibrium as the condition of minimum potential energy. Of course, to carry out such a minimization, we must first have a way of assigning potential energy to different configurations. We have seen that for systems *near* equilibrium the energy cost can be written as a quadratic function of the excursions about that equilibrium. These quadratic energies emerge both when characterizing the elastic response of materials treated as a continuous medium and when carrying out an atom-by-atom reckoning of the energy of configuration. However, there is often more to the delicate balance that determines structures than their potential energy alone. Thermal forces also make their presence known and we take up the apparatus to handle this part of the free

energy budget in the remainder of the chapter.

5.5 Structures as Free Energy Minimizers

In the previous section, we have seen how mechanics can provide insights into the equilibrium configurations of systems at all different scales. However, our arguments were incomplete because we neglected the role of thermal fluctuations in dictating equilibria. The aim of the present section is to explore the extension of our discussion of equilibrium to supplement energy minimization with the often conflicting demand of maximizing the entropy.

Though we will derive the result in its full glory later in the chapter, for the moment we examine the notion of free energy qualitatively. The concept of the free energy is embodied in the equation

$$\text{free energy} = \text{energy} - \text{temperature} \times \text{entropy}, \tag{5.28}$$

where the entropy (as will be shown below) is a measure of the number of different ways of rearranging the system. The fundamental argument of the remainder of the chapter and one of the foundational tools of the rest of the book is the idea that *the equilibrium state of a system is that choice out of all states available to the system that minimizes the free energy.*

The Entropy Is a Measure of the Microscopic Degeneracy of a Macroscopic State

From a mathematical perspective, the ideas introduced above about thermal forces are codified in the notion of the entropy. Though we are coached in thinking about energy from our earliest exposures to science, in fact, there is a much more intuitive state variable that provides deep insight into the factors determining the equilibrium states of complex systems such as a solution with a number of interacting species. In particular, the entropy of a closed system provides a measure of the number of different microscopic ways that we can realize a given observed macroscopic state, and can be written as

$$S = k_B \ln W, \tag{5.29}$$

where W is the number of microstates compatible with the macrostate of interest and k_B is the Boltzmann constant. In light of this definition, we see that when minimizing the free energy, the energetic terms tend to favor lower energy while the entropy contribution favors the macrostates that can be realized in the most ways.

As a concrete example relevant to our attempt to quantitatively unravel gene expression (see chaps. 6 and 19), we consider the role of entropy in the context of DNA-binding proteins. In particular, the entropy in this case reveals the number of distinct ways that we can arrange the bound proteins (nonspecifically) along the entire DNA molecule, as shown in fig. 5.23. We imagine that our DNA molecule has a total of N binding sites, N_p of which are occupied by the protein

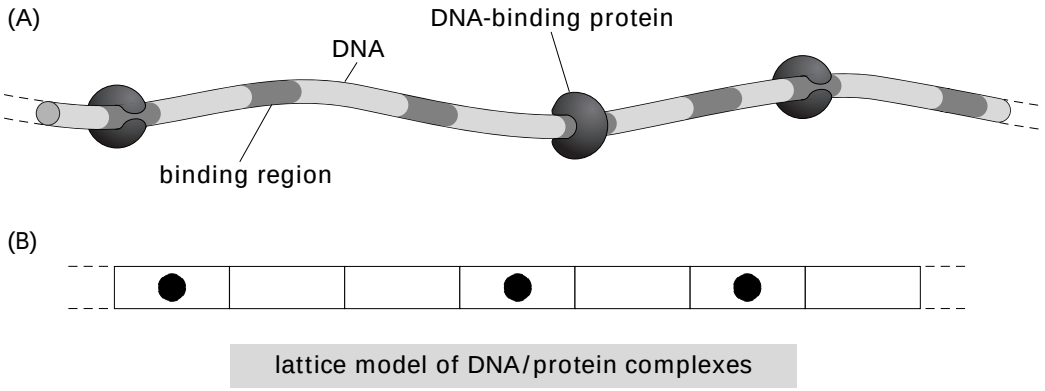

Figure 5.23: Possible arrangements of proteins on a DNA molecule. (A) The cartoon schematizes a DNA molecule on which are a series of binding sites which are shaded dark gray. The DNA binding proteins can occupy any of these sites. (B) The lattice model represents a further idealization in which we imagine the DNA molecule as a series of boxes into which we can put the DNA-binding proteins.

of interest. Further, we assume that the binding energies when the proteins are bound nonspecifically are the same regardless of which nonspecific sites are occupied (though in reality, even the energetics of nonspecific binding varies from site to site). As stated above, the entropy is a measure of the number of distinct ways of realizing a given macroscopic situation, in this case characterized by the number of possible binding sites and the number of binding proteins and is given in most general terms as

$$S = k_B \ln W(N_p; N), \tag{5.30}$$

where S is the entropy and $W(N_p; N)$ is the multiplicity factor which reflects the number of ways of rearranging the N_p proteins on the N binding sites. This definition of the entropy results from key consistency conditions such that the entropy of a composite system should be additive (although the total number of microstates for such a composite system is multiplicative).

For our example of DNA-binding proteins, we note that we have N choices as to where we lay down the first of the N_p proteins. Once this protein has been put down, we only have $N-1$ remaining sites where we might elect to put down the second protein. The third protein may now be put down on the DNA in any one of the remaining $N-2$ binding sites. Hence, the total number of ways of laying down our N_p proteins is $N \times (N-1) \times (N-2) \cdots \times (N-N_p+1)$. However, we have ignored the fact that these are not distinct configurations since we have no way to distinguish the case in which the first protein landed on site 10 and the second protein on site 15 and vice versa. As a result, we have overcounted and must divide by the number of rearrangements of those N_p

proteins on the occupied sites, which, following the same argument as above, is $N_p \times (N_p - 1) \cdots \times 1$. This product of all integers from one to N_p is N_p-factorial, which is denoted $N_p!$. We are now in a position to write the total number of *microscopic* arrangements as

$$W(N_p; N) = \frac{N \times (N-1) \times (N-2) \cdots \times (N - N_p + 1)}{N_p \times (N_p - 1) \cdots \times 1}. \tag{5.31}$$

If we now multiply top and bottom of the equation by $(N - N_p)!$, it results in the more pleasingly symmetric form

$$W(N_p; N) = \frac{N!}{N_p!(N - N_p)!}. \tag{5.32}$$

For a DNA-binding protein such as Lac repressor, there are roughly 10 copies of this protein bound on the roughly 5×10^6 DNA-binding sites within the *E. coli* genome. The formula above then tells us that there are roughly 3×10^{60} distinct arrangements of the Lac repressor bound to the *E. coli* genome.

Now that the counting has been effected, we are prepared to invoke Boltzmann's equation for the entropy given in eqn. 5.30. To compute this entropy we need to evaluate

$$S = k_B \ln \frac{N!}{N_p!(N - N_p)!}. \tag{5.33}$$

One of the key approximations needed in cases like this is known as the Stirling approximation which in its simplest form can be written as

$$\ln N! \approx N \ln N - N. \tag{5.34}$$

The origins of this approximation are taken up in "The Math Behind the Models" below and in the problems at the end of the chapter. In the context of our DNA-protein problem, if we invoke the Stirling approximation we find

$$S = -k_B N[c \ln c + (1 - c) \ln (1 - c)], \tag{5.35}$$

where we have introduced the more convenient concentration variable, $c = N_p/N$. The entropy as a function of concentration is shown in fig. 5.24. The key insight to emerge from this expression is the way in which the number of different ways of arranging the two species of interest depends upon their relative numbers. We see that the entropy is maximal when half of the sites are occupied - this situation reflects the fact that this concentration permits the most distinct arrangements.

- **The Math Behind the Models: The Stirling Approximation.** The Stirling approximation arises as a result of the need to evaluate expressions of the form $\ln N!$. The simplest heuristic argument to derive the result is based on the observation that

$$\ln N! = \ln [N(N-1)(N-2) \cdots 1]. \tag{5.36}$$

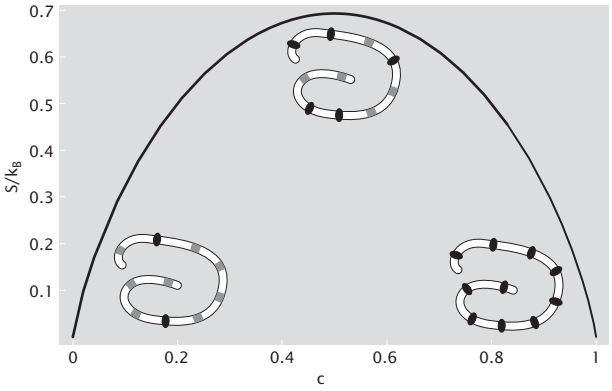

Figure 5.24: Entropy as a function of concentration of DNA binding proteins. The schematics show a DNA molecule with binding sites labeled in gray and with DNA binding proteins as black ovals on the binding sites. In going from left to right, the fraction of sites occupied by proteins increases.

On the other hand, by virtue of the property of logarithms that $\ln AB = \ln A + \ln B$, we can rewrite eqn. 5.36 as

$$\ln N! = \sum_{n=1}^{N} \ln n. \tag{5.37}$$

We can now replace this sum with the approximate integral

$$\sum_{n=1}^{N} \ln n \approx \int_{1}^{N} \ln x \, dx = N \ln N - N. \tag{5.38}$$

In the problems at the end of the chapter, this approximation is treated more carefully.

5.5.1 Entropy and Hydrophobicity

To gain a little more practice in the use of the entropy idea we consider a toy model of one of the most important molecular driving forces in biological systems, namely, the hydrophobic effect. The qualitative idea is that when a hydrophobic molecule is placed in water it deprives the water molecules in its vicinity from participating in some of the hydrogen bonds that they would have in the hydrophobic molecule's absence. An example of the highly idealized hydrogen bonding network in water is illustrated in fig. 5.25.

Hydrophobicity Results From Depriving Water Molecules of Some of Their Configurational Entropy

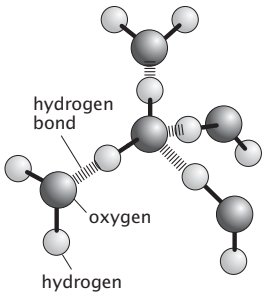

Figure 5.25: The hydrogen bonding network in water. Water molecules participate in hydrogen bonding (illustrated by the dashed lines joining adjacent water molecules). A given water molecule can be idealized as having neighbors arranged in a tetrahedral structure.

The objective of the present section is to make an estimate of the magnitude of these hydrophobic effects. The basic thrust of the argument will be to describe how nonpolar molecules in solution deprive water molecules of the capacity to engage in hydrogen bonding and thereby steal away part of their orientational entropy. This simple model borrows from a model originally formulated by Pauling (1935) to capture the entropy of ice. With this mechanism in hand, we then carry out numerical estimates of the size of this effect.

The structural idea suggested by fig. 5.25 is that the oxygen atoms of neighboring water molecules form a tetrahedral network. As further suggested by fig. 5.25, these water molecules form a dynamic network of hydrogen bonds, where each oxygen, on average, makes two hydrogen bonds with the four water molecules surrounding it. A useful conceptual framework for thinking about hydrophobicity is that when nonpolar molecules are placed in solution, the water molecules that neighbor the nonpolar molecule of interest have a restricted set of choices for effecting such hydrogen bonding. We can coarse grain the continuum of possible orientations available to a water molecule to the six distinct orientations shown in fig. 5.26. As a result, it is possible to estimate the entropic disadvantage associated with the presence of nonpolar molecules (see Dill and Bromberg, 2002 for a clear description of this effect).

The six orientations that a water molecule can assume derive from the six ways of choosing to point the hydrogen atoms associated with the water molecule of interest towards the vertices of a tetrahedron. If one of the four water molecules in its immediate vicinity is replaced by a nonpolar molecule then the number of available orientations drops to three since one of the possible hydrogen bonding partners is now gone. For example, if we assume that the neighboring water molecule in the direction of the lower right hand vertex of

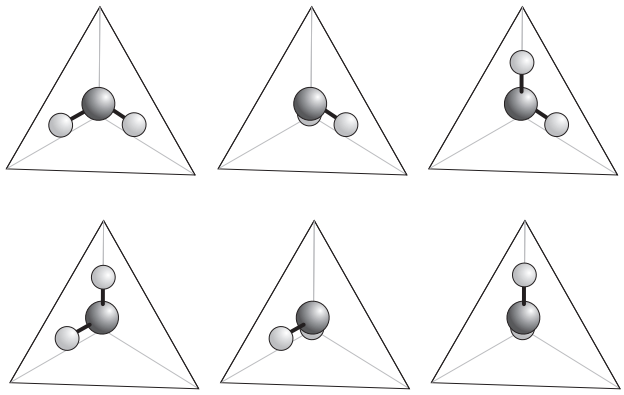

Figure 5.26: Orientations of water molecules in a tetrahedral network. Each image shows a different arrangement of the water molecule that permits the formation of hydrogen bonds with neighboring water molecules. The hydrogen bonds are in the directions of the vertices that are *not* occupied by hydrogens in the figure. (Adapted from K. Dill and S. Bromberg, Molecular Driving Forces, New York, Garland Press, 2003.)

fig. 5.26 is removed, this means that hydrogen bonds can no longer be formed with the oxygen on the water shown in the figure and the hydrogen atoms on the missing water molecule. As a result, the three configurations in the bottom of fig. 5.26 are now forbidden. This simple model predicts that the presence of the nonpolar molecule deprives each neighboring water molecule of half of its possible orientations as a participant in the hydrogen bonding network. The entropy change of each such water molecule is given by

$$\Delta S_{\text{hydrophobic}} = \underbrace{k_B \ln 3}_{\text{constrained } H_2O} - \underbrace{k_B \ln 6}_{\text{unconstrained } H_2O} = -k_B \ln 2. \quad (5.39)$$

Thus far we have determined the entropy loss per water molecule. To make our estimate useful, we now need to estimate the number of water molecules that are impacted by the presence of the nonpolar (i.e. hydrophobic) molecule of interest.

We can obtain a quantitative description of the hydrophobic cost to place a hydrophobic molecule in water as

$$\Delta G_{\text{hydrophobic}}(n) = \text{nk}_B\text{T} \ln 2, \quad (5.40)$$

where n is the number of water molecules adjacent to the nonpolar molecule of interest. Here we have accounted only for the entropic contribution to the free energy cost, which is given by $-T\Delta S_{\text{hydrophobic}}$. One particularly useful way of characterizing our result is to say that the presence of hydrophobic molecules incurs some free energy cost per unit area ($\gamma_{hydrophobic}$) and hence

that the free energy cost to embed a given hydrophobic molecule in water is obtained as $\Delta G_{\text{hydrophobic}} = \gamma A$, where A is the effective area of the interface between the hydrophobic molecule and the surrounding water. As said above, a more convenient representation of this result is to assign a free energy per unit area which can be obtained by determining the area per water molecule that is contributed. Using the simple estimate that ten water molecules cover an area of approximately 1 nm^2 and that $\ln 2 \approx 0.7$, we can see that the interfacial free energy required to submerse a hydrophobic object in water is roughly 7 $k_BT/\ nm^2$. For a small molecule such as oxygen (O_2) that has an approximate surface area between 0.1 nm^2 and 0.2 nm^2, the energetic cost of putting this molecule in water costs roughly 1 k_BT. As a result, oxygen can be readily dissolved in water, even though it is nonpolar and cannot form hydrogen bonds. However, larger hydrophobic molecules comparable in size to a protein or even a sugar molecule would require significant free energy input to be dissolved in water.

The hydrophobic effect is responsible for the everyday observation that oil and water do not mix. The free energy cost resulting from this simple model for putting an individual hydrocarbon molecule such as octane into a watery environment is on the order of 15 k_BT. Each addition of a new molecule of octane to water costs the same amount of free energy additively. However, if the octane molecules clump together, the total surface area of the clump may be much less than the sums of their individual surface areas. In water at room temperature where individual molecules can jiggle around rapidly it usually takes no more than a few seconds for the molecules to sort themselves out such that the interfacial surface is minimized.

Amino Acids Can Be Classified According to Their Hydrophobicity

The energies associated with the hydrophobic effect are extremely important at both the molecular and cellular scale in dictating the formation of structures. For example, consider a protein that contains a variety of amino acid side chains, some of which are hydrophilic (able to form hydrogen bonds with water) and others of which are hydrophobic. From the argument outlined above, it is clear that there must be a free energy cost for the hydrophobic side chains to exist in an aqueous environment. As a first approximation, the folding of proteins into defined three-dimensional structures can be thought of as an application of the principle of the separation of oil and water. The protein is made up of an elastic backbone from which dangle a mixture of hydrophobic and hydrophilic amino acid side chains. As described above, the entropic demands of the system will tend to force the hydrophobic side chains to gather together in a sequestered internal oil droplet at the heart of the protein. Hydrophilic amino acid side chains will tend to remain on the protein surface where they can form hydrogen bonds with water. This concept was illustrated in fig. 5.8 (pg. 258).

To make an accurate quantitative model describing the role of the hydrophobic effect in protein folding we would have to know the relative free energy cost for water exposure for each of the 20 amino acids. However, as a useful simplified strategy for building intuition, we will start off by simply dividing the

amino acids into two broad groups, one that includes all hydrophobic residues (H) and the other that includes all hydrophilic or polar residues (P). As we will explore in more detail in chap. 8, this drastically oversimplified model provides useful estimates for many aspects of structure.

As a result of the arguments given above, we can rank the various hydrophobic amino acids most simply through reference to the effective area that they present to the surrounding water. Within this framework, the hydrophobic cost of exposing such a residue is of the form

$$\Delta G_{hydrophobic} \approx \underbrace{\gamma_{hydrophobic}}_{\text{cost/area}} \quad \underbrace{A_{hydrophobic}}_{\text{hydrophobic area}} \quad . \tag{5.41}$$

The detailed implementation of this strategy is left to the reader in the problems at the end of the chapter.

When in Water, Hydrocarbon Tails on Lipids Have an Entropy Cost

These same ideas can also be used to give an approximate description of the free energy associated with lipids when they are isolated in solution. Lipid molecules are characterized by polar head groups that are attached to long, fatty acid tails which are hydrophobic. The simple and useful idea in this case is to consider each such tail as though it presents a cylinder of hydrophobic material and to assign a free energy cost to isolated lipids given by the product of the hydrophobic free energy cost per unit area computed above and the area presented by the "cylinder" from the lipid tails. The free energy cost associated with isolated lipids leads to the key driving force resulting in the formation of lipid bilayers.

5.5.2 Gibbs and the Calculus of Equilibrium

We have already observed that the principle of minimum potential energy presides over questions of the mechanical equilibrium configurations of systems at zero temperature. We now enter into a discussion of a principle that plays precisely the same role for systems in equilibrium at finite temperature. Once again, we will see that the equilibrium edict can be couched in variational language (i.e. as a minimization problem). In particular, our discussion will culminate with the statement that out of all competing states of a system, the equilibrium state minimizes the relevant free energy. However, before discussing free energy minimization, we reflect on the even more fundamental embodiment of the second law of thermodynamics, namely, that for a closed system in equilibrium the entropy is a *maximum*.

Thermal and Chemical Equilibrium Are Obtained by Maximizing the Entropy

The study of thermal and chemical equilibrium are presided over by the second law of thermodynamics. In words, this law can be stated as the assertion that the macroscopic equilibrium state of an *isolated* system is that state that

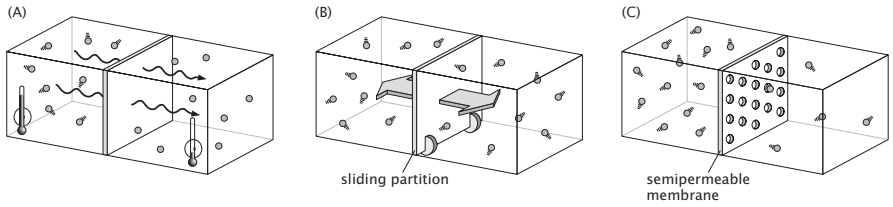

Figure 5.27: Schematic representation of an isolated system with two subcompartments and a barrier between these two compartments which permits transfer of (A) energy, (B) volume and (C) particles.

can occur in the largest number of microscopic ways (i.e. which maximizes the entropy). Stated differently, when faced with the question of choosing from the space of all macroscopic competitors, choose that one that has the most microscopic representatives. This is the governing principle of all of thermodynamics in the same sense as all of mechanics derives from Newton's second law, $F = ma$.

How the injunction of entropy maximization plays out in real but simple circumstances is illustrated in an isolated system like that shown in fig. 5.27 which has an internal partition. Our use of the word "isolated" refers to the fact that the contents of the container are entirely indifferent to anything and everything that we do outside - hence, we are unable to communicate with that system by doing work on it, by heating it or by applying any sorts of fields such as magnetic or electric fields. The thought experiment of interest here involves the idea of spontaneously removing the constraint implied by the internal partition. For example, as shown in fig. 5.27(A), if we permit the flow of energy between the two compartments, there will be a transfer of energy until the entropy is maximized (which corresponds to equality of temperature). As shown in fig. 5.27(B), if the brakes which hold the partition fixed are released, this partition is free to slide until the overall entropy of the system has reached a maximum. Depending upon which side has the greater pressure, the partition will roll either to the left or the right until the pressures on the two sides are equal. Finally, the case of most biological interest is that shown in fig. 5.27(C) in which the internal partition is rendered permeable to the flow of particles. In this case, the particles will flow across the partition until the entropy is maximized, a state we will show later corresponds to equality of chemical potentials in the two regions.

To show that the idea of entropy maximization leads to consequences that are consistent with our physical intuition, we reason quantitatively about the isolated system with a partition shown in fig. 5.27. We claimed that upon removal of the constraints represented by the partition, there would be a redistribution of energy (heat will flow), an adjustment in the volume (the partition will roll in one direction or the other) and a redistribution of particles (particles

will diffuse) until the entropy of the closed system reaches a maximum. Mathematically, we can see this by examining $S_{tot} = S_1(E_1, V_1, N_1) + S_2(E_2, V_2, N_2)$, the total entropy which is an additive function of the entropy on the two sides of the partition. Note that because our system is isolated, there are constraints of the form $E_{tot} = E_1 + E_2$, $V_{tot} = V_1 + V_2$ and $N_{tot} = N_1 + N_2$, where E_{tot} is the total energy of the closed system, V_{tot} is the total volume and N_{tot} is the total number of particles. As noted above, when the conditions implied by the initial constraints are relaxed (e.g. the brake is released and the partition can roll), there will be a spontaneous change in the state of the system until the system entropy is maximized. Mathematically, for the case in which the partition permits exchange of energy, the entropy maximization takes the form

$$dS = (\frac{\partial S_1}{\partial E_1})dE_1 + (\frac{\partial S_2}{\partial E_2})dE_2 = 0, \tag{5.42}$$

If we now invoke the fact that the total energy is conserved, we have $dE_2 = -dE_1$, which when substituted into eqn. 5.42 yields

$$(\frac{\partial S_1}{\partial E_1} - \frac{\partial S_2}{\partial E_2})dE_1 = 0. \tag{5.43}$$

To see our derivation through to the end, we now introduce the thermodynamic definition of temperature, $dS/dE = 1/T$, which reveals that our result is equivalent to the statement that $T_1 = T_2$. That is, when the partition permits energy transfer, heat will flow until the temperature on the two sides is equal. Note that we cannot derive every result from thermodynamics here and encourage readers unfamiliar with thermodynamic identities to consider the "Further Reading" at the end of the chapter.

The argument goes in precisely the same way when we consider the case where the brakes are removed and the partition is permitted to slide. In this case it is the volume which will be adjusted in such a way as to maximize the system entropy. In particular, the condition of entropy maximization is

$$dS = (\frac{\partial S_1}{\partial V_1})dV_1 + (\frac{\partial S_2}{\partial V_2})dV_2 = 0, \tag{5.44}$$

Once again, we exploit the constraint which tells us that $dV_2 = -dV_1$, resulting in

$$(\frac{\partial S_1}{\partial V_1} - \frac{\partial S_2}{\partial V_2})dV_1 = 0. \tag{5.45}$$

At this point, we use the thermodynamic identity that $p/T = (\partial S/\partial V)_{E,N}$, resulting in the observation that entropy maximization corresponds to equality of pressure.

The case which is probably of greatest biological interest is that in which the partition permits the flow of particles. In this case, entropy maximization corresponds to the statement

$$dS = (\frac{\partial S_1}{\partial N_1})dN_1 + (\frac{\partial S_2}{\partial N_2})dN_2 = 0, \tag{5.46}$$

Exploiting the constraint that the overall number of particles is fixed, we have that $dN_1 = -dN_2$, resulting in

$$(\frac{\partial S_1}{\partial N_1} - \frac{\partial S_2}{\partial N_2})dN_1 = 0. \tag{5.47}$$

Here we use the result that the chemical potential is defined in terms of the entropy change as $\mu/T = -(\partial S/\partial N)_{E,V}$, and hence, entropy maximization implies equality of chemical potentials on both sides of the partition.

The key point of these arguments has been to highlight the variational description of thermodynamic equilibrium. That is, the privileged equilibrium state of a system can be found by maximizing the entropy. The driving forces implied by entropy maximization have a variety of interesting consequences which we take up presently. Though many problems of interest will require a more sophisticated implementation of the second law in the form of the principle of minimum free energy, there are a number of problems where entropy maximization can be used directly. Important examples include the notion of an entropic spring which describes the force-extension characteristics of molecules like DNA, the notion of depletion forces between macromolecular assemblies in solution and the origins of osmotic pressure. In anticipation of the role of entropy maximization in coming sections, it is of interest here to show how order can arise from entropy maximization.

5.5.3 Structure as a Competition

Thus far, we have asserted that equilibrium structures reflect energy minima (zero temperature) and entropy maxima (finite temperature). However, the case of greatest interest for biological model building is associated with a variational middle ground between the strictly mechanical ambition of minimizing energy and the statistical ambition of maximizing entropy. In particular, both the *in vitro* assays of solution biochemistry and the *in vivo* chemical action of cellular life reflect a more subtle situation in which the system of interest can exchange energy or matter (or both) with the surroundings. The variational injunction in these cases is to minimize the free energy, which can be thought of intuitively as teasing out the competition between maximizing multiplicity and minimizing energy. The variational principle that is equipped to permit the playing out of this competition is the principle of minimum free energy introduced in words earlier in the chapter.

Free Energy Minimization Can Be Thought of as an Alternative Formulation of Entropy Maximization

As developed above, Gibbs' calculus of equilibrium asserts that when contemplating *isolated* systems, our best guess as to the equilibrium state is that macroscopic state that can happen in the most ways microscopically. This injunction is translated into mathematical terms by virtue of the introduction of the entropy which we are asked to maximize in order to find equilibria. On the other hand, there are a number of problems of interest for which the system

may not be thought of as isolated. Indeed, most interesting systems (like macromolecules in the cell, macromolecular assemblies such as RNA polymerase or the ribosome, or indeed, cells themselves) are in contact with an external medium with which they can exchange energy and matter. Whether we think of a cell in the ocean or DNA polymerase in a thermal cycler used for doing the polymerase chain reaction, our system of interest is in contact with the rest of the world. Interestingly, the problem of maximizing the entropy of a system plus the reservoir with which it is in contact is equivalent to minimizing the free energy of just the system itself. The beauty of this principle is that it allows us to dismiss the huge potential complexity engendered by the fact that our system is in contact with a reservoir and to consider only those degrees of freedom that describe the system itself.

To see how this discussion goes, consider fig. 5.28 which shows examples of both isolated and closed systems. Note that a closed system is characterized by the ability to exchange energy with its environment. An open system is free to exchange both energy and matter with its environment. The theoretical trick that allows us to exploit the principle of entropy maximization in the context of these systems is to turn a closed system into an isolated system by putting our original system in contact with a reservoir. From this perspective, our overall system consists of the original system and its associated reservoir and its equilibrium is now dictated by entropy maximization. For example, in the *in vitro* situation of solution biochemistry, a test tube in contact with a water bath is an example of a closed system that can exchange energy with its environment. Similarly, within the confines of a cell, we can think of a particular site on a DNA molecule as the system of interest, and the reservoir as the DNA-binding proteins in the cytoplasm or on other nonspecific sites on the DNA.

To see the analysis of this problem through to the end, we now maximize the entropy of our composite system made up of the original system and reservoir. An alternative way of thinking of our requirement that the entropy be maximized is to say that during any spontaneous process that follows the removal of a constraint, the entropy will increase. We may write this statement as

$$dS_{tot} = dS_r + dS_s \geq 0. \tag{5.48}$$

where S_s refers to the entropy of our system and S_r to the entropy of the reservoir. We may borrow from our knowledge of the first law of thermodynamics, which permits us to write the change in energy of the reservoir as

$$dE_r = TdS_r - pdV_r, \tag{5.49}$$

where TdS_r is the heat added to the reservoir and pdV_r is the work done by the reservoir. If we substitute this relation into our entropy inequality, we have

$$dS_s + \frac{dE_r}{T} + \frac{p}{T}dV_r \geq 0. \tag{5.50}$$

Like before, since the energy and volume are conserved in our overall isolated

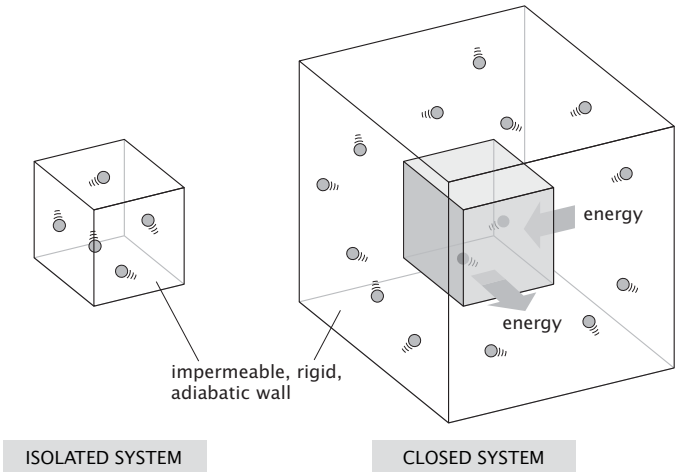

Figure 5.28: Isolated and closed systems. The isolated system is unable to exchange energy and matter with the rest of the world. In the closed system, there is an exchange of energy between the system and the surrounding reservoir.

system, we have $dE_r = -dE_s$ and $dV_r = -dV_s$, resulting in

$$dS_{tot} = dS_s - \frac{1}{T}dE_s - \frac{p}{T}dV_s \geq 0. \tag{5.51}$$

Finally, if we multiply both sides of our equation by $-T$, resulting in a change in direction of the inequality, we are left with

$$dG = d(E_s + pV_s - TS_s) \leq 0, \tag{5.52}$$

where we have introduced the Gibbs free energy $G = E - TS + pV$ and have shown that in a spontaneous process, the free energy will be reduced or, what is the same, that the Gibbs free energy will take its minimum value in the equilibrium state. We remind the reader that the statement that the Gibbs free energy is minimized in equilibrium is founded upon the more fundamental statement of entropy maximization for the entire system which is built up from the subsystem of interest and the reservoir. Further, note that once we elect to invoke the Gibbs free energy, there is no reference to the coordinates associated with the reservoir, other than to say that our system of interest is kept at fixed temperature T and pressure p by virtue of its contact with the reservoir.

5.5.4 An Ode to ΔG

The Free Energy Reflects a Competition Between Energy and Entropy

The interesting physical insight tied to the free energy is that equilibrium structures reflect a competition between energetic and entropic influences. If the temperature goes to zero, we recover the minimization of the energy E, which is the principle that presides over questions of strictly *mechanical* equilibrium. On the other hand, as the temperature rises, the entropic term makes itself heard with increasing forcefulness until at sufficiently high temperatures, it dominates the decision concerning the equilibrium state.

To be specific, this battle between energy and entropy is well illustrated by the examples presented in fig. 5.7 (pg. 257). For example, in the context of the folding of a protein into its native state, the energetic component of the problem has to do with the formation of various native contacts between amino acids that result in a net energy lowering, while the entropic part of the free energy budget has to do with the number of alternative conformations available to the protein when it is *not* folded. At high enough temperatures, the entropic imperative must be obeyed and the protein denatures. A similar competition is seen in the context of the charges on proteins where the energy of electrostatic interactions tends to keep charges localized to their molecular hosts while the entropic part of the free energy budget prefers to see these charges delocalized. As a final case study in deconstructing the free energy in the context of the examples of fig. 5.7, consider the case of binding of ligands to a protein. In this case, the ligands are afforded an entropic advantage if they wander around in solution. On the other hand, binding to their molecular hosts confers an energetic advantage and the interplay between these competing demands is the province of free energy minimization.

In the coming chapters, we will invoke this idea of a competition between energetic and entropic factors repeatedly in contexts ranging from protein folding to the distribution of ions around a ribosome to the deflection of biofunctionalized cantilevers in the presence of particular antigens. In each case, our arguments will be formulated first in terms of an energetic term which tends to pull the system in one direction - for example, the Coulomb attraction between ions in solution and some macromolecule will tend to localize those ions near the macromolecule. The second competing term will reflect the will of the entropic term which will favor the spreading out of the ions around the macromolecule. Such arguments ultimately serve as the concrete outcome of the present chapter which has argued for the idea that the question of equilibrium structures can be seen as one of minimization of the relevant potential. As such, the free energy presides over all questions requiring that we determine the equilibrium state. Note that there are different free energies that are most convenient (Helmholtz free energy, Gibbs free energy) and for simplicity, we will ignore these subtleties and always use the symbol G for the free energy in the remainder of the book.

5.6 Summary and Conclusions

Much of the busy activity of cellular life involves transformations of matter and energy. In this chapter we showed that the study of these physical and chemical

transformations can be couched in the language of finding the maximum (entropy) or least (potential energy) value of some physical quantity. In particular, we argued that Newtonian statics can be reformulated through the idea that we seek the lowest value of the potential energy of the system that is consistent with whatever constraints are imposed. Though these mechanical principles are important, we saw that the principle of minimum free energy is of even greater biological importance. In particular, we have argued that isolated systems containing many particles (and thus which can be realized by astronomical numbers of microscopically distinct but macroscopically identical states) can be thought of as satisfying a different variational imperative, namely, the maximization of their multiplicity. The calculus of equilibrium for many-particle systems is founded on the idea that macroscopic states are those which can be realized in the largest number of ways microscopically. Coming chapters will show how this simple idea can be used to understand a huge variety of different biological phenomena.

5.7 Appendix: The Euler-Lagrange Equations, Finding the Superlative

Finding the Extrema of Functionals Is Carried Out Using the Calculus of Variations

This chapter centered on our ability to find the maxima and minima of energies, entropies and free energies. On the other hand, as we saw in fig. 5.16, often the energy and free energy dictated by our biological problems will be functionals rather than functions. This raises the mathematical question of: given an infinite set of competitor functions, how can we find that function that minimizes the energy or free energy functional? The aim of the present discussion is to generalize the earlier discussion based on ordinary calculus and to examine the functional analog of finding the extremum of a function. For a deeper discussion of these issues, we refer the reader to both Lanczos (1970) and Gelfand and Fomin (1963).

The Euler-Lagrange Equations Let Us Minimize Functionals By Solving Differential Equations

In many cases, the type of functional minimization described above can be written in a very specialized form, namely, as the search for that function which leads to an extremum for an integral. In this case, we are asked to minimize a functional of the form

$$E[u(s)] = \int_{a_1}^{a_2} f(u(s), u'(s))ds. \tag{5.53}$$

As with the calculation of extrema of functions, the defining condition is that for any "small excursion" about the extrema, there should be no first order change

in the value of the functional. To make this idea more concrete, we consider excursions of the form $\eta(s)$, and demand that

$$\frac{\delta E[u(s)]}{\delta u(s)} = \lim_{\varepsilon \to 0} \frac{E[u(s) + \varepsilon\eta(s)] - E[u(s)]}{\varepsilon} = 0. \tag{5.54}$$

If we reason by analogy with ordinary calculus, what this expression tells us is that if we have found the minimizing function $u(s)$, then any small excursion (namely, $\varepsilon\eta(s)$) about that minimum will lead to no change in the functional. The notation $\delta E[u(s)]/\delta u(s)$ reminds us that we are taking the "functional derivative" as opposed to the derivative of ordinary calculus. The class of admissible excursions, $\eta(s)$ is further specified by boundary conditions imposed on the competitor functions. For example, for the case when the values of the competitor functions are fixed at the boundaries, the admissible excursions must satisfy $\eta(a_1) = \eta(a_2) = 0$. The condition expressed in eqn. 5.54 may now be written as

$$\frac{\delta E}{\delta u(s)} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}[\int_{a_1}^{a_2} f(u(s) + \varepsilon\eta(s), u'(s) + \varepsilon\eta'(s))ds - \int_{a_1}^{a_2} f(u(s), u'(s))ds]. \tag{5.55}$$

We now consider a Taylor series expansion of the integrand $f(u(s)+\varepsilon\eta(s), u'(s)+\varepsilon\eta'(s))$, and in particular, we consider such an expansion to first order, resulting in

$$f(u(s) + \varepsilon\eta(s), u'(s) + \varepsilon\eta'(s)) \approx f(u(s), u'(s)) + \varepsilon\frac{\partial f}{\partial u'}\eta'(s) + \varepsilon\frac{\partial f}{\partial u}\eta(s). \tag{5.56}$$

As a result, we may now write

$$\frac{\delta E}{\delta u(s)} = \int_{a_1}^{a_2} (\frac{\partial f}{\partial u'}\eta'(s) + \frac{\partial f}{\partial u}\eta(s))ds. \tag{5.57}$$

Until now, our results have shown us how to reexpress our original problem, in terms of the rate of change of the function $f(u(s), u'(s))$. At this point, we have two essential steps which remain. First, we rearrange eqn. 5.57 by exploiting a single integration by parts. In particular, we note

$$\int_{a_1}^{a_2} \frac{\partial f}{\partial u'}\eta'(s)ds = \eta(s)\frac{\partial f}{\partial u'}|_{a_1}^{a_2} - \int_{a_1}^{a_2} \frac{d}{ds}\frac{\partial f}{\partial u'}\eta(s)ds, \tag{5.58}$$

where the first term on the right hand side of the equation is zero because $\eta(a_1) = \eta(a_2) = 0$ as dictated by boundary conditions. As a result of these manipulations, we may rewrite eqn. 5.57 as

$$\frac{\delta E}{\delta u(s)} = \int_{a_1}^{a_2} (-\frac{d}{ds}(\frac{\partial f}{\partial u'}) + \frac{\partial f}{\partial u})\eta(s)ds. \tag{5.59}$$

We now carry out the second step, which is to acknowledge that the condition for an extremum really corresponds to the statement $\frac{\delta E}{\delta u(s)} = 0$. In light of this condition and as a result of the fact that $\eta(s)$ is arbitrary, we are left with

$$\frac{d}{ds}(\frac{\partial f}{\partial u'}) - \frac{\partial f}{\partial u} = 0 \tag{5.60}$$

What we have learned is that the problem of minimizing a functional is equivalent to that of solving a differential equation. In particular, the differential equation associated with variational problems like that given above is the so-called Euler-Lagrange equation of the problem of interest. The beauty of this result is that it has replaced a problem that we don't know how to do, namely, minimizing a functional, with another that we might know how to do, namely, solving a differential equation. The minimization of a functional to solve problems of interest will show up again in chap. 9 when we think about the charge distribution around a protein, chap. 10 when thinking about the atomic-force microscope and biofunctionalized cantilevers and in chap. 11 when we work out the deformation in a lipid bilayer membrane induced by membrane proteins such as ion channels.

5.8 Problems

1. The sugar budget revisited.

In chap. 3 we worked out the rate of sugar uptake to provide the construction materials for a dividing bacterium. However, as shown in this chapter, sugar molecules also provide the *energy* needed to perform macromolecular synthesis. Amend the estimate of chap. 3 to include the fact that sugar supplies construction materials and the energy needed to assemble them. How many sugars are needed to provide the energy and construction materials for making a new cell? Make an estimate for the average rate of sugar uptake for a dividing bacterium in light of this amendment to our earlier estimates.

2. A feeling for the numbers: covalent bonds.

(a) Based on their typical energies and distances estimate the frequency of vibration of covalent bonds.

(b), Based on your result from part (a), estimate the time step required to do a classical mechanical simulation of protein dynamics.

3. Stretching DNA

Fig. 5.29 shows the experimental force-extension curve for stretching single stranded DNA (ssDNA). Here we consider the model of ssDNA as N springs which represent the covalent bonds connecting all the phosphorous atoms along the backbone of the ssDNA molecule.
(A) From the graph we see that for forces exceeding 50pN the ssDNA behaves like a spring (i.e. there is a simple linear relation between the force and the extension. By drawing a straight line through these data points estimate the spring constant of the ssDNA molecule (in pN/μm) as well as its unstretched length (in μm).
(B) Using the estimate for the spring constant of a covalent bond, $k_b = 20$N/m, and the distance between neighboring P atoms (or nucleotides) in unstretched ssDNA, $a = 5\mathring{A}$, compute the spring constant of the ssDNA molecule used in

the experiment. How does your computed value compare to the experimentally determined value from part a)?

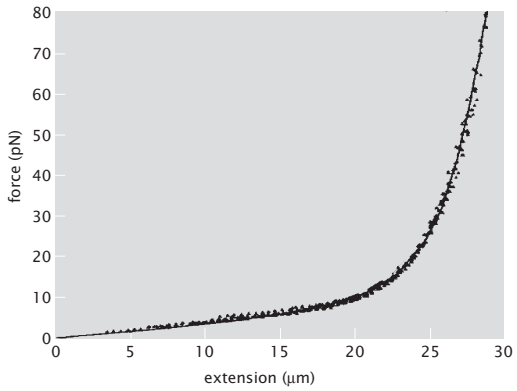

Figure 5.29: Force extension curve for single stranded DNA. (Adapted from S. B. Smith *et al.*, *Science*, 271:795, 1996.)

4. Taylor expansions.

(a) Repeat the derivation of the Taylor expansion for *cos* x given in the chapter, but now expand around the point $x = \pi/2$.
(b) Do a Taylor expansion of the function e^x. Generate a plot of the fractional error as a function of x for different orders of the approximation and a comparison of the function and the different order expansions such as was shown in fig. 5.20.
(c) A one-dimensional potential energy landscape is give by the equation $f(x) = x^4 - 2x^2 - 1$. Find the two minima and do a Taylor expansion around one of them to second order. Show the original function and the approximation on the same plot in the style of fig. 5.19 (pg. 272).

5. A feeling for the numbers: comparing multiplicities.

Boltzmann's equation for the entropy (eqn. 5.29) tells us that the entropy difference between a gas and liquid is given by

$$S_{gas} - S_{liquid} = k_B \ln \frac{W_{gas}}{W_{liquid}}. \tag{5.61}$$

From the macroscopic definition of entropy as $dS = dQ/T$ we can make an estimate of the ratios of multiplicities by noting that boiling of water takes place at fixed T at 373 K.
(a) Consider a cubic centimeter of water and use the result that the heat needed to boil water (the latent heat of vaporization) is given by $Q_{vaporization} = 40.66$ kJ/mole (at 100 degrees centigrade) to estimate the ratio of multiplicities of

water and water vapor for this number of molecules. Write your result as 10 to some power. If we think of multiplicities in terms of an ideal gas at fixed T, then

$$\frac{W_1}{W_2} = \left(\frac{V_1}{V_2}\right)^N. \tag{5.62}$$

What volume change would one need to account for the liquid/vapor multiplicity ratio? Does this make sense?

(b) In the chapter we discussed the Stirling approximation and the fact that our results are incredibly tolerant of error. Let's pursue that in more detail. We have found that the typical types of multiplicities for a system like a gas are of order $W \approx exp(10^{25})$. Now, let's say we are off by a factor of 10^{1000} in our estimate of the multiplicities, namely, $W = 10^{1000}exp(10^{25})$. Show that the difference in our evaluation of the entropy is utterly negligible whether we use the first or second of these results for the multiplicity. This is the error tolerance that permits us to use the Stirling approximation so casually!

6. Stirling approximation revisited.

The Stirling approximation is useful in a variety of different settings. The goal of the present problem is to work through a more sophisticated treatment of this approximation than the simple heuristic argument given in the chapter. Our task is to find useful representations of $n!$ since terms of the form ln $n!$ arise often in reasoning about entropy.

(a) Begin by showing that

$$n! = \int_0^\infty x^n e^{-x} dx. \tag{5.63}$$

To demonstrate this, use repeated integration by parts. In particular, demonstrate the recurrence relation

$$\int_0^\infty x^n e^{-x} dx = n \int_0^\infty x^{n-1} e^{-x} dx, \tag{5.64}$$

and then argue that repeated application of this relation leads to the desired result.

(b) Make plots of the integrand $x^n e^{-x}$ for various values of n and observe the peak width and height of this integrand. We are interested now in finding the value of x for which this function is a maximum. The idea is that we will then expand about that maximum. To carry out this step, consider $\ln(x^n e^{-x})$ and find its maximum - argue why it is okay to use the logarithm of the original function as a surrogate for the function itself - that is, show that the maxima of both the function and its logarithm are at the same x. Also, argue why it might be a good idea to use the logarithm of the integrand rather than the integrand

itself as the basis of our analysis. Call the value of x for which this function is maximized x_0. Now expand the logarithm about x_0. In particular, examine

$$\ln\left((x_0+\delta)^n e^{-(x_0+\delta)}\right) = n\ln\left(x_0+\delta\right) - (x_0+\delta) \tag{5.65}$$

and expand to second order in δ. Exponentiate your result and you should now have an approximation to the original integrand which is good in the neighborhood of x_0. Plug this back into the integral (be careful with limits of integration) and by showing that it is acceptable to send the lower limit of integration to $-\infty$, show that

$$n! \approx n^n e^{-n} \int_{-\infty}^{\infty} e^{-\frac{\delta^2}{2n}} d\delta. \tag{5.66}$$

Evaluate the integral and show that in this approximation

$$n! = n^n e^{-n} (2\pi n)^{1/2}. \tag{5.67}$$

Also, take the logarithm of this result and make an argument as to why we can get away with dropping the discussion of the $(2\pi n)^{1/2}$ term.

7. Energy cost of macromolecular synthesis.

Visit the website "ecocyc.org" and find the metabolic pathways for synthesis and breakdown of all the small molecules found in *E. coli*. Look at two pathways, glycolysis and serine synthesis. As you will see, the amino acid serine is constructed from the small molecule 3-phosphoglycerate which is an intermediate of the glycolytic pathway. Several energy-requiring steps and energy-generating steps occur along the way. How many molecules of glucose must be taken up to provide the carbon skeleton used to make serine? How many molecules of ATP are consumed and created along the way? How many reducing equivalents of NADH and NADPH are consumed or created along the way? What is the overall energy cost to synthesize one molecule of serine in units of ATP and units of k_BT? How many molecules of glucose must be metabolized in order to generate this amount of energy?

8. Counting and diffusion.

In this chapter, we began practicing with counting arguments. One of the ways we will use counting arguments is in thinking about diffusive trajectories.
Consider 8 particles, 4 are black and 4 are white. 4 particles can fit left of a permeable membrane and 4 can fit right of the membrane. Imagine that due to random motion of the particles every arrangement of the 8 particles is equally likely. Some possible arrangements are: BBBB|WWWW, BBBW|BWWW, WBWB|WBWB; the membrane position is denoted by |.
(a) How many different arrangements are there?
(b) Calculate the probability of all 4 black particles on the left of the permeable membrane. What is the probability of having 1 white particle and 3 black

particles on the left of the membrane. Finally, calculate the probability that 2 white and 2 black particles are left of the membrane. Compare these three probabilities. Which arrangement is most likely?
(c) Imagine that in one time instant a random particle from the left side exchanges places with a random particle on the right hand side. Starting with 3 black particles and 1 white particle on the left of the membrane, compute the probability that after one time instant there are 4 black particles on the left? What is the probability that there are 2 black and 2 white particles on the left, after that same time instant? Which is the more likely scenario of the two?

(adapted from Dill and Bromberg 2003)

9. Molecular driving forces In section 5.5.2 we showed that entropy maximization leads to our intuitive ideas about equilibrium. However, that discussion can be extended to reveal the direction of spontaneous processes. In particular, during any spontaneous process, we know that the entropy will increase. Use this fact in the form of the statement that $(\mu_2 - \mu_1)dN_1 \geq 0$ to deduce the role of differences chemical potential as a "driving force" for mass transport. If $\mu_2 > \mu_1$, which direction will particles flow? Make analogous arguments for the flow of energy and changes in volume.

5.9 Further Reading

O. Morton, **Eating the Sun**, Fourth Estate, London: England, 2007. This excellent book describes the story of how our modern understanding of photosynthesis was developed.

F. C. Neidhardt, J. L. Ingraham and M. Schaechter, **Physiology of the Bacterial Cell**, Sinauer Associates, Inc., Sunderland: Massachusetts, 1990. Chap. 5 on "Biosynthesis and Fueling" is particularly relevant for the present chapter.

G. Gottschalk, **Bacterial Metabolism**, Springer-Verlag, New York: New York, 1986. This book is full of interesting insights into the census and energy budget of bacterial cells.

N. C. Price, R. A. Dwek, R. G. Ratcliffe and M. R. Wormald, **Principles and Problems in Physical Chemistry for Biochemists**, Oxford University Press, Oxford: England, 2001. This excellent book describes many of the important chemical reactions of biology from a thermodynamic perspective.

D. S. Lemons, **Perfect Form**, Princeton University Press, Princeton: New Jersey, 1997. Lemon's book is a pedagogical delight and offers a variety of interesting, yet simple, insights into how mechanics can be couched in the language of minimization.

P. J. Nahin, **When Least is Best**, Princeton University Press, Princeton: New Jersey, 2004. Nahin gives a sense of the wide range of problems that can be formulated as questions of minimization.

F. S. Acton, **Numerical Methods That Work**, The Mathematical Association of America, Washington, D. C., 1990. Acton is thoughtful and very amusing and describes some of the pitfalls of numerical mathematics. See also his book **Real Computing Made Real** for more fun and insights.

C. Kittel and H. Kroemer, **Thermal Physics**, W. H. Freeman and Company, San Francisco: California, 1980. This book will provide background on thermodynamics and statistical mechanics for the interested reader.

H. B. Callen, **Thermodynamics and an Introduction to Thermostatistics**, John Wiley and Sons, New York: New York, 1985. Callen champions the idea of maximum entropy as the basis for finding equilibrium states. Our treatment mirrors his chap. 2. on "The Conditions of Equilibrium".

C.Tanford, "Contribution of Hydrophobic Interactions to the Stability of the Globular Conformation of Proteins", J. Am. Chem. Soc. **84**, 4240 (1962) and **The Hydrophobic Effect: Formation of Micelles and Biological Membranes**, Krieger Publishing Company, Malabar: Florida, 1991. Tanford has a nice touch in describing hydrophobicity.

5.10 References

C. Bustamante, J. F. Marko, E. D. Siggia and S. Smith, "Entropic Elasticity of λ-Phage DNA", Science **265**, 1599 (1994).

K. Dill and S. Bromberg, **Molecular Driving Forces**, Garland Press, New York: New York, 2003.

G. Lim, M. Wortis and R. Mukhopadhyay, "Stomatocyte-discocyte-echinocyte sequence of the human red blood cell: Evidence for the bilayer-couple hypothesis from membrane mechanics", Proc. Nat. Acad. Sci. **99**, 16766 (2002).

C. K. Mathews, K. E. van Holde and K. G. Ahern, **Biochemistry**, Addison Wesley Longman, Inc, San Francisco: California, 2000. Their fig. 13.6 gives an energy profile of the glycolysis pathway.

L. Pauling, "The Structure and Entropy of Ice and of Other Crystals with Some Randomness of Atomic Arrangement", J. Amer. Chem. Soc. **57**, 2680 (1935). Pauling's paper clearly outlines the structural assumptions behind his ice model which can also be used to estimate the hydrophobic cost of placing molecules in

water.

R. Phillips and S. Quake, "The Biological Frontiers of Physics", Physics Today, May 2006.

M. Schaechter, J. L. Ingraham and F. C. Neidhardt, **Microbe**, ASM Press, Washington DC, 2006.

S. B. Smith, Y. Cui and C. Bustamante, "Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules", Science **271**, 795 (1996).