December 30, 2024 Draft

Physical Principles in Biology

Biology 3550/3551
Spring 2025

Chapter 5: Thermodynamics

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

Chapter 5

Thermodynamics

Thermodynamics is the branch of chemistry and physics that is concerned with the interconversion of different forms of energy. The laws of thermodynamics place absolute constraints on how much work can be obtained from different forms of energy and whether or not specific processes will occur spontaneously. Nobody gets to break these laws! Thermodynamics does not tell us how a process will occur, but only whether or not the process will occur spontaneously.

Thermodynamics is definitely one of the most challenging subjects for nearly all students, for several reasons:

- The ideas are fundamentally abstract and subtle. Even people who have thought about thermodynamics for many years easily get tripped up.
- It uses math! Even if many of the ideas seem to be conceptually simple, a deep and useful understanding depends on mathematics.
- The language can be confusing. Different disciplines sometimes use different terms. In addition, the language has historical origins, and the history is convoluted.
- Historical confusion. Unlike classical mechanics, which almost all came from one person, Newton, over a short period of time, thermodynamics was developed over a long period of time by several generations of scientists, and there were periods of profound confusion.

But, it's worth it! The current issues regarding energy and climate change offer a dramatic examples of how important the interconversion of energy forms can be. In the context of our course, we will use these principles to understand how biological systems become organized and do amazing things.

5.1 Energy, Work and Heat

I. Units of energy.

Before going on, it is worthwhile to review again the definition of energy and the units we use. A general definition of energy is that it is the "ability to do work". We define mechanical work as the integral of force applied over distance:

$$w = \int_{x_1}^{x_2} F dx$$

If the force is constant, then

$$w = F(x_2 - x_1)$$

The units of work must be force times distance. The SI unit of force is the newton, which is the force required to accelerate a mass of 1 kg by 1 m/s per s. The SI unit of work or energy is then the Nm, or joule:

$$1\,\mathrm{joule} = 1\,\mathrm{newton} \cdot \mathrm{meter} = 1\,\mathrm{kg} \cdot \mathrm{m}^2\mathrm{s}^{-2}$$

Another commonly used unit of energy is the calorie, which is defined as the amount of energy required to raise the temperature of 1 g of water by 1 °C, or 1 K. One problem with defining the calorie in this way is that the heat required to raise the temperature of water depends on the starting temperature. 4 °C and 15 °C have been used to define the unit, sometimes with a subscript indicating the temperature. The standard definition now is in terms of the Joule, as defined above:

$$1\,\mathrm{cal} = 4.184\,\mathrm{J}$$

Another source of confusion is that the calorie has been defined at different times in terms of the energy to heat either 1 g of water or 1 kg. The current convention is to use a lower case "c" to designate the "gram calorie" and an uppercase "C" to designate the "kilogram calorie". The energy content of foods is expressed in kilogram calories. This is one reason that the idea of loosing weight by eating ice cream (which should cool the body and result in negative calorie intake) is doomed to failure!

II. An important distinction: Temperature versus heat.

Before considering some simple examples, it is important to be sure that we are using language carefully. One common cause of confusion is the difference between temperature and heat.

- Temperature has a relatively straight-forward definition. It is a property of a given mass of matter that is directly related to average kinetic energy of the molecules making up the matter. For an ideal gas, the temperature determines the relationship between volume and gas according to the ideal gas law:

  $$PV = nRT$$

  For an ideal gas, temperature is related to the average kinetic energy of the ideal gas molecules according to:

  $$E = 3kT/2$$

- Heat is often defined as "a form of energy". More specifically, heat is the flow of energy from a warm object to a cooler one, with the result that the temperatures become more equal.

At one time, heat was thought to be a massless substance, often called “caloric”, that flowed invisibly between objects. We now understand that heat is not something that objects contain, but the language still seems to imply this. This is one of the historical origins of some of the confusion in thermodynamics.

It is possible to formulate the laws of thermodynamics without invoking “heat”, but that seems rather artificial to me and only shifts the difficulties to other words.

- In summary: temperature is a property of matter that can be directly measured, while heat is a flow of energy associated with temperature changes. The heat flow is always from a warm object to a cooler one.

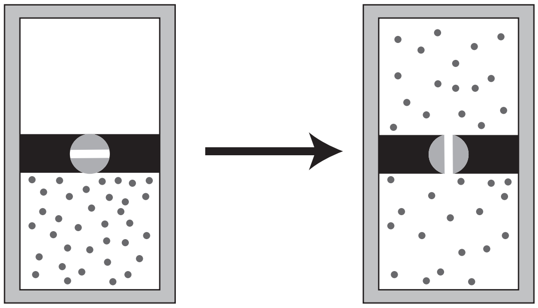
III. Some examples based on the expansion and compression of gasses.

Historically, one of the most important motivations for the development of thermodynamics was the invention of the steam engine. Once steam engines started to be used, there was great interest in getting the maximum amount of work from them. Since steam engines work through the expansion of a gas, many of the ideas in thermodynamics were formulated in this context.

At the time, though, the relationship between the pressure of a gas and the motion of molecules was not understood, and the theory was developed without any explicit model for what actually generated forces or work. This treatment is usually referred to as “classical thermodynamics”. Later, a molecular interpretation was developed, what we now call “statistical mechanics”. Both can stand on their own and are completely consistent with one another. But, I think that it is often easier to merge the two approaches when trying to understand things.

1. Gas expansion without work

   Suppose we build an apparatus as shown below:



   This apparatus should have the following features:

   - It is completely insulated from its surroundings, so that heat can’t flow in or out of it.
   - There are two chambers separated by a valve that can be opened or closed without perturbing anything else and without generating any heat, that is there can be no friction.
   - One chamber is filled with a gas at some arbitrary pressure, P, and temperature,T.

- The other chamber is evacuated.

Once the system has equilibrated, we open the valve. As we have discussed at length, the molecules are moving about due to kinetic energy in random directions. Eventually (quite quickly, actually) we expect the molecules to distribute themselves throughout the vessel, with roughly equal numbers on the two sides.

What can we say about what has changed and what hasn't?

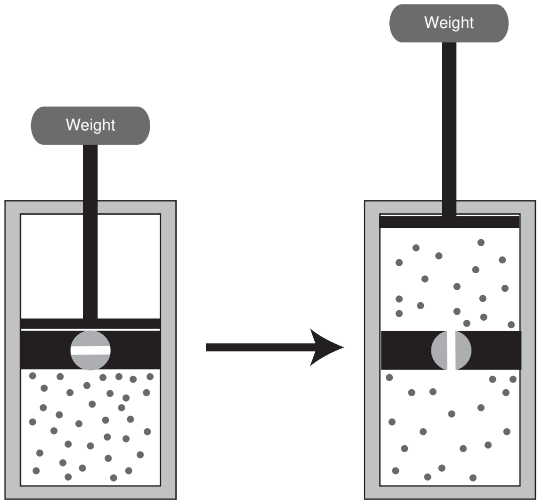
- Nothing has happened to change the average velocities of the molecules. Therefore, the temperature should be the same, and the average kinetic energy is the same.
- The volume has increased. So, by the ideal gas law, $PV = nRT$, the pressure should decrease.
- Because of the insulation, no heat has flowed in or out of the chambers.
- Nothing, except the molecules, has been moved, so no work has been done.

But, something else has happened. Even though no work was done and the energy has stayed the same, we know that we would have to do work, *i.e.,* spend energy, to restore the system to its original state. You probably already know the fancy word for this, "entropy", but let's put off discussing this concept. What we can say, though, is that somehow or other we have wasted the potential to do work, even though the total thermal energy has stayed the same.

Processes in which there is no exchange of heat are called "adiabatic".

2. Gas expansion with work, but without heat flow.

Now, let's think about another type of apparatus. Again we will keep the apparatus fully insulated. But, now a movable piston has been added to the upper chamber. The upper chamber is still evacuated, and our engineers have made a perfect, frictionless seal to the outside, through which a rod connects the piston to a weight.



When we open the valve, the gas molecules begin to collide with the piston, and the pressure from the gas molecules pushes the piston and the weight upward. Mechanical work is now being done. What is changing as the piston moves?

- Does the temperature change? Consider what happens when the piston is pushed upwards. The gas molecules collide with the piston and transfer some of their energy to it. Unlike when they collide with a fixed wall, their velocity is not quite as great when they bounce off in the other direction. As a consequence the average kinetic energy of the molecules decreases, meaning that the temperature decreases as well.
- If the volume increases by the same amount as in the previous example and the temperature decreases, then the pressure must decrease *more* than in the previous example.

The key point about this example is that some of the kinetic energy of the gas molecules has been converted to mechanical work. From the conservation of momentum, the change in energy should be equal to the amount of work done.

Let's put some labels on the quantities involved:

- E = energy of the molecules. For an ideal gas, this is entirely translational kinetic energy. We will define the change in energy associated with a process as:

  $$\Delta E = E_{\text{final}} - E_{\text{start}}$$

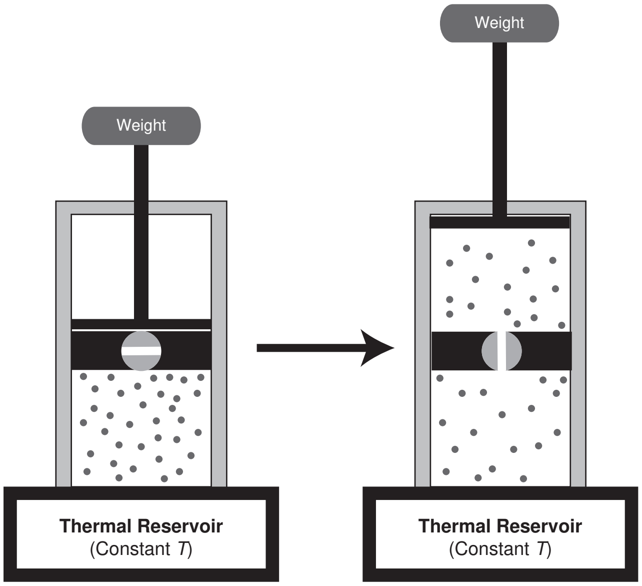
  It's important to keep track of the signs! We will define the work involved, w, so that it is positive when work is done on the gas (or, more generally the "system"), and negative when the system does work on the outside world, as in this case.
  For this example, we know that the amount of work must equal the change in energy:

  $$\Delta E = w$$

  Is this always true?

3. Gas expansion at constant temperature with work.

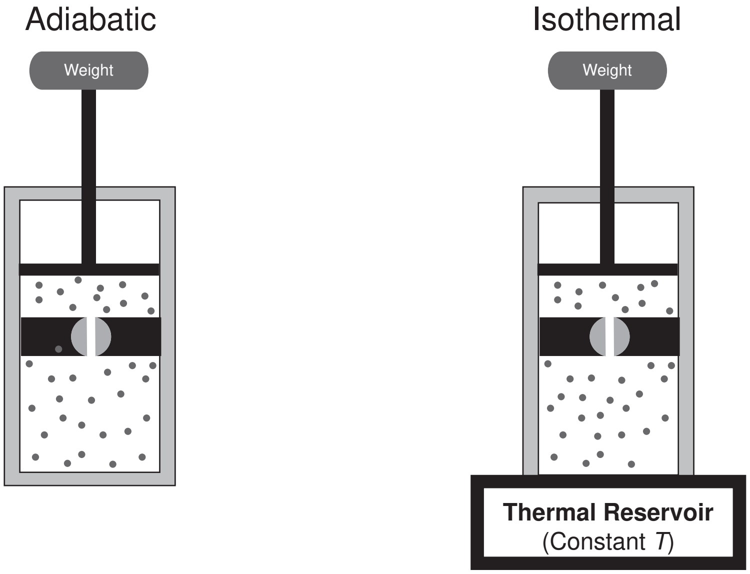
   Next, we have the engineers build an even fancier device by adding a thermal reservoir at the bottom of the cylinder:



Basically, the reservoir is a large mass equilibrated at the same temperature as the gas. If the temperature of the gas drops, heat will flow to equalize the two temperatures, without significantly lowering the temperature of the reservoir.

What happens, now, when we allow the piston to move and do work?

- As the piston is pushed upwards, the gas molecules loose some energy, and the temperature starts to drop. But, as soon as that happens, heat flows from the reservoir.
- At the end of the process, the temperature is the same as when it started, so the energy must be the same as well. In this respect, the result is the same as in the adiabatic expansion without work.
- Again, work has been done. Is more work done in the adiabatic or isothermal process? Consider a point part way through each process, where the volumes have increased by the same amount, as illustrated below:



Because the adiabatic process does not allow heat to flow to the gas, the temperature at this intermediate point must be lower for the adiabatic process than the isothermal one. As a consequence, the pressure must be lower for the adiabatic process. Since pressure represents force divided by the area over which it is exerted, the total force on the piston must also be less for the adiabatic process than for the isothermal one. We can thus conclude that the isothermal process can produce more work than the adiabatic process.

The energy for the additional work from the isothermal process is drawn from the heat of the reservoir.

In this example, $\Delta E = 0$, but work has been done, and there has been a flow of heat.

We represent the heat flow by the symbol q and define it so that it is positive when heat flows from the surroundings into the system of interest. In this case, q is positive and w is negative, and:

$$q = -w$$

IV. The first law of thermodynamics.

The common statement of the first law is that the energy of the universe is constant. But, the more formal statement is that for any process, the change in energy, E, is the sum of the work done on the system and the heat absorbed from the surroundings:

$$\Delta E = q + w$$

We can see how this applies to the three examples from above:

- Adiabatic expansion without work.

$$q = 0$$
$$w = 0$$
$$\Delta E = 0$$

- Adiabatic expansion with work

$$q = 0$$
$$w < 0$$
$$\Delta E < 0$$
$$\Delta E = w$$

- Isothermal expansion with work

$$q > 0$$
$$w < 0$$
$$\Delta E = q + w = 0$$
$$q = -w$$

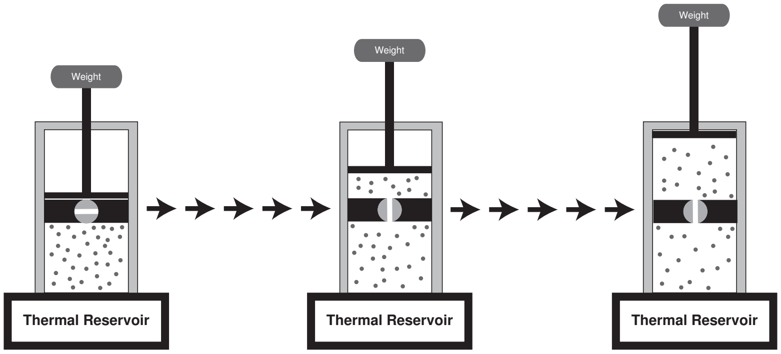
Is there a proof of the first law? No! The laws of thermodynamics are postulates, and our confidence in them comes from the fact that no one has ever found an exception.

In 1775 the French Royal Academy of Science effectively made the first law of thermodynamics (before it was called that) a real law by declaring that it would no longer consider patent applications for perpetual motion machines.

V. Reversible expansion and compression

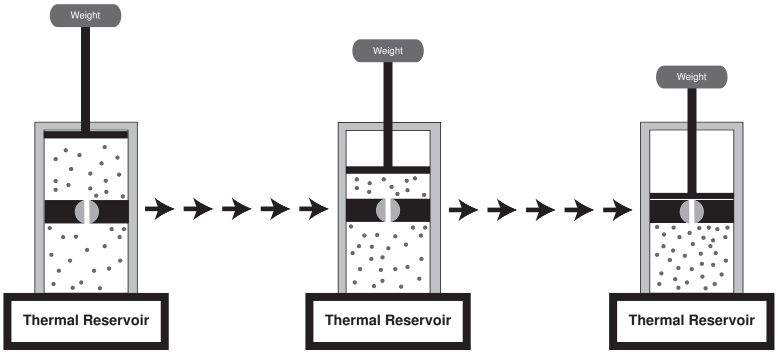
We have already seen that we can get extra work from the expansion of the gas by allowing it to draw heat from its surroundings as it expands. Even if the temperature is the same at the beginning and end of the expansion processes, different paths can lead to different amounts of work being produced and different amounts of heat absorbed.

The maximum amount of work that can be generated from the expansion of a gas is obtained by controlling the movement of the piston so that it is slow enough that the temperature never actually drops. This cannot be done in practice, but it can be approached as a limit, in the sense of limits in calculus, by allowing the piston to move only in infinitesimal steps, as suggested in the drawing below:



Because the piston moves in infinitesimally small steps, the temperature never falls below that of the reservoir. If a larger step ever does take place, the temperature will drop, causing the pressure to drop. As a consequence, less work will be produced by the expansion. This is why we can argue that this is the path that will lead to the maximum production of work (the most negative value of w).

The reverse of this process, diagrammed below, is the one that requires the *least* amount of work to compress the gas to its original volume.



Again, the volume is changed in infinitesimally small steps. As the gas is compressed, the piston imparts extra kinetic energy on the gas molecules, but (in the ideal case) the excess energy is instantly transferred to the thermal reservoir and the temperature remains constant. Because the volume of the gas decreases, the pressure increases, and progressively more work must be done for each decrease in the volume. But, if larger steps were taken, the temperature would increase temporarily, causing a larger increase in pressure and requiring more work for the next step.

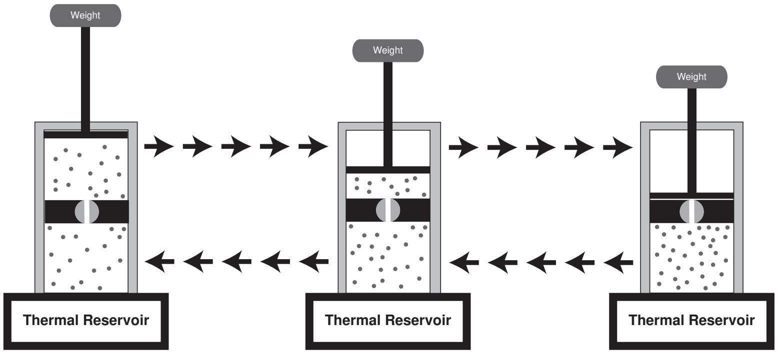
The two processes described above are said to be reversible, and that term can be taken in two senses. In the first sense, the two processes are reversible because they are exact opposites. The maximum work that can be done during expansion is also the minimum amount of work that is required for compression. Thus:

$$w_{\text{exp,rev}} = -w_{\text{comp,rev}}$$

Since $\Delta E = 0$ for both processes, and $q = -w$ for each, we can also conclude that

$$q_{\text{exp,rev}} = -q_{\text{comp,rev}}$$

We can also imagine a cycle formed by the two processes, as illustrated below:



Because $w_{\text{exp,rev}} = -w_{\text{comp,rev}}$ and $q_{\text{exp,rev}} = -q_{\text{comp,rev}}$, w and q for the complete cycle must be zero. Thus, the system is restored to its original volume, pressure, temperature and energy with no net input or output of either work or heat.

Of course, no real process can meet this ideal. As noted above, any real expansion process will produce less work than the ideal, and any real compression will require more work than the ideal. As a consequence, there is a net amount of work applied to the system and w for the cycle will be greater than zero. If the original temperature is restored $\Delta E = 0$ and $q = -w$, and q will be negative, meaning that heat is transferred to the surroundings. This is the general nature of less-than-ideal processes, they require more work (or produce less) than the ideal process and they release more heat.

The ideal expansion and compression processes described above are also reversible in a second sense: Either process can be reversed at a given point by an infinitesimal force in the opposite direction. This is the meaning usually implied by the term in thermodynamics.

VI. The maximum work from expanding a gas at constant temperature

We can calculate the work from the reversible isothermal expansion with a bit of calculus, starting with the integral for any work function:

$$w = \int_{x_1}^{x_2} f dx$$

where x_1 and x_2 are the initial and final positions of the piston, and f is the force. During the reversible gas expansion, the force does not remain constant, but rather drops as the pressure decreases. At any instant, the force is proportional to the pressure:

$$f = -P \cdot A$$

where A is the area of the piston. The negative sign reflects the fact that the work done on the system by expansion will be negative. For each small increment of x, there is a corresponding small increment in the volume of the gas:

$$dV = Adx$$

or, we can write:

$$dx = dV/A$$

We can replace dx in the integral with dV/A, and f with $-P \cdot A$, so that:

$$\begin{aligned} w &= -\int_{V_1}^{V_2} (P \cdot A) dV/A \\ &= -\int_{V_1}^{V_2} P dV \end{aligned}$$

(recall that the product PV has the units of energy, or work!) We can also express pressure in terms of volume, $P = nRT/V$:

$$w = -\int_{V_1}^{V_2} \frac{nRT}{V} dV$$

n, R and T are all constant, so they can be moved outside of the integral.

$$\begin{aligned} w &= -nRT \int_{V_1}^{V_2} \frac{1}{V} dV \\ &= -nRT \ln(V)\Big|_{V_1}^{V_2} \\ &= -nRT \ln\left(\frac{V_2}{V_1}\right) \end{aligned}$$

Since we have the same number of molecules at the beginning and end of the process, the concentration is inversely related to the volume. So, we can write this expression in terms of concentration:

$$w = -nRT\ln\left(\frac{C_1}{C_2}\right)$$

The form of this expression will likely be familiar to you, or will be soon. This is the origin of all of the expressions like $-RT\ln(C_1/C_2)$ that occur so frequently in chemistry!

Since the energy of the gas molecules is the same at the beginning and end of the process, $\Delta E = 0$, and we can write:

$$\Delta E = 0 = q + w$$

$$q = -w = nRT\ln\left(\frac{V_2}{V_1}\right) = nRT\ln\left(\frac{C_1}{C_2}\right)$$

VII. State functions versus path functions

The quantities E, on the one hand, and q and w, on the other are fundamentally different. The energy of a system depends only on the state of that system, irrespective of how it got there. For an ideal gas, the energy depends only on the temperature and the number of molecules. A change in energy, ΔE, depends only on the starting and ending states.

But, will we always get the same amount of work for a given change in state? Our examples clearly show that we won't. For both the adiabatic expansion without work and the isothermal expansion, ΔE is zero, but q and w are different for the two processes. Furthermore, the amount of work produced by an isothermal expansion (where the temperature is the same at the beginning and end) depends on just how the expansions is carried out.

We say that the energy is a "state function", while heat and work are path-dependent functions. For a gas, the pressure, P and volume, V, are also state functions.

Although work and heat are not state functions, the changes in these quantities associated with the ideal, reversible process separating two states do represent changes in state function. Thus, there are quantities, w_{rev} and q_{rev}, that are associated with any two states that can, in principle, be interconverted by a reversible process.

The quantity w_{rev} is described as being the "free energy change", ΔF, for the conversion of one state to another.[1] ΔF represents the maximum amount of work that can be obtained from a favorable process, or the minimum amount of work required to drive an unfavorable process. The energy change is "free" in the sense that it is the maximum amount of energy that is available to do work. In a more practical sense, free energy is anything but free, since it is the kind of energy that we pay for when we buy, for

---

[1] More specifically, ΔF is the change in Helmholz free energy, as distinguished from the Gibbs free energy that will be introduced later.

instance, electrical power or gasoline. On the other hand, molecular kinetic energy is "free" in the sense that it is always there, but it can't actually be used to do any work.

The quantity q_{rev} for any two states is also very important, as we will soon see when we discuss the elusive concept of entropy.

5.2 Entropy and the Second Law

The term entropy is commonly associated with the idea of randomness, and this general meaning is used very widely, even entering every-day language. In the context of thermodynamics, where the term originated, a much more specific definition is required, however. There are, in fact, two different and precise ways in which entropy is defined, reflecting the classical and statistical approaches to thermodynamics. Here, we will look at both definitions and show that they are equivalent for at least one process, the expansion of a gas.

I. The classical definition of entropy.

The maximum work that we calculated for the expansion of a gas is simply a function of the temperature and the starting and finishing volumes. This allows us to define this quantity as a state function. It also describes what is *lost* when a gas expands, the ability to do work, and it seems to be related to the increased disorder of the gas, *i.e.,* its entropy.

In fact, the thermodynamic definition of entropy is based on the heat absorbed, q, via the path leading to maximum work, that is the reversible process:

$$\Delta S = \frac{q_{\text{rev}}}{T}$$

where q_{rev} is the heat absorbed by the system in the reversible process, which is $-w_{\text{rev}}$.

We can apply this definition to the case of a gas expansion, when the temperature is the same at the beginning and end of this process. Among the examples we have considered, this could be either the adiabatic expansion without work or the isothermal expansion with work. Because entropy is a state function, the change in entropy is independent of the path between states. *But*, the heat quantity used to calculate the entropy change is the one associated with the reversible process. For the reversible isothermal expansion of a gas, we showed that the heat absorbed is:

$$q_{\text{rev}} = -w_{\text{rev}} = nRT \ln\left(\frac{V_2}{V_1}\right)$$

The entropy change is then calculated as:

$$\Delta S = \frac{q_{\text{rev}}}{T} = nR \ln\left(\frac{V_2}{V_1}\right)$$

The entropy change is easiest to calculate for a process in which the temperature stays constant, but there is still a maximum amount of work that can be obtained for the transition between two states that have different temperatures, and this maximum amount of work is also obtained through a reversible, (*i.e.,* infinitely slow) path. In this case, the entropy change is calculated as an integral:

$$\Delta S = \int_{T_1}^{T_2} \frac{q_{\mathrm{rev}}}{T} dT$$

where T_1 is the initial temperature and T_2 is the final temperature, and q_{rev} may be a continuously changing function of temperature.

Now, entropy is supposed to be a quantity that increases spontaneously and is related to disorder. In the case of the expanding gas, the quantity we calculated above does seem to be related to disorder, since the molecules are less ordered in a larger volume.

The entropy of a system, as defined here, is a state function. The entropy changes by the same amount irrespective of how much work, for instance, was obtained from the expansion.

But, it is important to point out that this definition only applies to the isolated system. We haven't said anything about how the entropy of the surroundings change. We will come back to this when we talk about the second law, which concerns the entropy of the system *and* its surroundings.

II. The statistical definition of entropy.

The classical definition of entropy is probably not the one that you are most familiar with, which probably refers to the idea of randomness or disorder. In fact, the classical definition is is not an easy concept to work with and is only useful in rather restricted cases, like heat engines.

The important point from the earlier discussions of gas expansion and contraction is that the energy, E, is a function that helps define a state, but the change in energy for a change from one state to another is not sufficient to tell us whether or not the change is favorable or how much work can be obtained. In addition to the internal energy, there is this other quantity that defines the states and determines how much work can be obtained.

In the case of the expanding gas, it appears that the driving force is the tendency of the individual molecules to occupy as large a volume as is made available to them. As we have repeatedly discussed, this is a purely probabilistic phenomenon. If the molecules begin on one side of the container and are allowed to move freely to the other, there is a 50% probability that any one molecule will wind up on the other side. There will be a net flow to the other side until the concentrations on the two sides are equal.

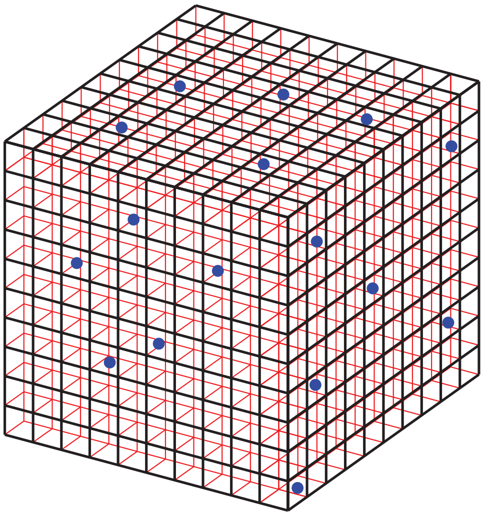
The statistical definition of entropy of a state can be expressed as:

$$S = k \ln \Omega$$

where k is the Boltzmann constant, and Ω is the number of equally probable microstates that make up the state. In other words, Ω is the number of different ways of arranging the components of the system. This is the fundamental postulate of statistical mechanics. It was deduced (not proven!) by Ludwig Boltzmann, and it is engraved on his tombstone in Vienna. Like the laws of thermodynamics, we believe it because it works!

So, all that we have to do to calculate the entropy is count up the ways of arranging the components. How do we do *that*? This is not at all trivial, and all we can usually do is apply this definition to idealized simple systems. But, the problem is made a little bit easier if we only try to calculate the *change* in entropy for a process. For instance, we can go back to our example of a gas expanding. For even a modest number of molecules, there are a vast number of microstates, each defined by the position and momentum of each molecule. We obviously can't count them all. But, if the temperature stays constant, then the number of possible values for the momentum of each molecule stays the same. So, that all we have to consider is the number of different positions for each molecule.

Suppose that we divide up the initial volume into a grid of small cubes, say 1 nm on a side, as diagrammed below:



If the number of molecules is N and the number of cubes is N_c, then we can calculate the number of ways of arranging the N molecules into the N_c cubes. To simplify things, we can assume that N_c is much larger than N, so that once we place a molecule in a given cube, it doesn't significantly reduce the number of places available to the next molecule. So, taking one molecule at a time, the number of ways of arranging them is:

$$N_\mathrm{c} \cdot N_\mathrm{c} \cdot N_\mathrm{c} \cdot \cdots = N_\mathrm{c}^N$$

But, if all of the molecules are indistinguishable, then we have to divide by number of ways of choosing the molecules sequentially, $N!$. So:

$$\Omega = \frac{N_\mathrm{c}^N}{N!}$$

So, the initial entropy would be:

$$S = k \ln \Omega = k \ln \left(\frac{N_{\mathrm{c}}^N}{N!} \right)$$

But, this is a little suspect, since the result depends on the number of cubes that I divided the volume up to. Would the entropy suddenly become larger if I decided to use smaller cubes? That doesn't sound right. But, it turns out to be OK if we consider a *change* in volume.

To describe a change in volume, we keep the size of the cubes the same, and the volume change is specified by the change in the number of cubes. If we call the initial number of cubes $N_{\mathrm{c},1}$ and the final number $N_{\mathrm{c},2}$, the change in entropy is given by:

$$\begin{aligned} \Delta S = S_2 - S_1 &= k \ln \left(\frac{N_{\mathrm{c},2}^N}{N!} \right) - k \ln \left(\frac{N_{\mathrm{c},1}^N}{N!} \right) \\ &= k(N \ln(N_{\mathrm{c},2}) - \ln(N!)) - k(N \ln(N_{\mathrm{c},1}) - \ln(N!)) \\ &= kN \ln \left(\frac{N_{\mathrm{c},2}}{N_{\mathrm{c},1}} \right) \end{aligned}$$

Thus, the change in the entropy change depends on the *ratio* of the number of cubes at the beginning and end of the process. If we make the volume of the individual cubes larger or smaller, this ratio is not affected.

If we want to express the entropy change on a molar basis, n, the number of moles, is equal to N divided by Avogadro's number, N_{A}, and $N = nN_{\mathrm{A}}$:

$$\Delta S = nN_{\mathrm{A}}k \ln \left(\frac{N_{\mathrm{c},2}}{N_{\mathrm{c},1}} \right)$$

Since k is the gas constant, R divided by Avogadro's number, we can write:

$$\Delta S = nR \ln \left(\frac{N_{\mathrm{c},2}}{N_{\mathrm{c},1}} \right)$$

Finally, $N_{\mathrm{c},2}$ and $N_{\mathrm{c},1}$ are proportional to the volumes before and after and we can write the expression as:

$$\Delta S = nR \ln \left(\frac{V_2}{V_1} \right)$$

Notice that this is exactly the same as the expression we derived from the classical definition of the entropy change for this process.

III. Microstates with different probabilities

What if the state is defined by microstates that do not all have equal probabilities? For instance, what if we were to divide up the total volume into cubes of different sizes, so that they had different probabilities of being occupied. Would the calculated entropy be different? More significantly, what if we were considering a more complicated environment, like the interior of a cell, where certain molecules have specific affinities for different compartments Could we still use this approach to calculate entropy, or at least entropy changes?

The more general expression for the statistical entropy is:

$$S = -k \sum_{i=1}^{N} p_i \ln p_i$$

where N is the number of microstates, and p_i is the probability of microstate i. We won't try to prove this, but we can consider a couple of extreme cases. At one extreme, if we have N states with equal probabilities, the probability of each state must be $p_i = 1/N$. The entropy is then calculated as:

$$\begin{aligned} S &= -k \sum_{i=1}^{N} p_i \ln p_i = -k \sum_{i=1}^{N} \frac{1}{N} \ln \frac{1}{N} \\ &= -kN \frac{1}{N} \ln \frac{1}{N} \\ &= k \ln N \end{aligned}$$

consistent with the original expression for Ω microstates with equal probabilities. At the other extreme, if there is only a single state, its probability must be 1, and the entropy is:

$$\begin{aligned} S &= -k \sum_{i=1}^{N} p_i \ln p_i \\ &= -k 1 \ln(1) \\ &= 0 \end{aligned}$$

So, this state has an entropy of zero. In general, for a given number of microstates, if some of the microstates are more probable than others, the entropy will be lower than if all of the microstates have equal probabilities. So, for instance, if some configurations of a molecule are more probable because certain atoms tend to interact with each other, the molecule will have a lower entropy than if all of the possible conformations had equal probabilities.

IV. Entropy and information

There is another way of thinking about entropy that comes from a quite different discipline, information theory, which is concerned about how to efficiently and accurately transmit and manipulate information. This was a subject that was developed during the middle decades of the 20th Century, largely at the late great Bell Laboratories, which did basic research for the major U.S. telephone company of the time (AT&T).

In this view, entropy can be described as the amount of information required to specify the exact state of a system. So, for instance, a crystal of a pure substance can be described with relatively little information, because each molecule is in an equivalent position of the crystal. If we know the structure of the molecule and the parameters describing the crystal lattice, then we can completely reconstruct the structure of the crystal. On the other hand, to completely describe a gas of the same molecules requires that we individually specify the position, orientation and momentum of each molecule.

This idea can be quantified, to give a function that calculates, for instance, the number of bits required to transmit a particular message or other information such as images or audio signals. Some messages contain a great deal of repetition and so can be encoded with relatively few bits, meaning that they have low entropy. A message composed of randomly-chosen letters has a high entropy and requires more bits to communicate. The equations for calculating information content have exactly the same form as the statistical definition of thermodynamic entropy, and information content is often described as information entropy.

The science of information theory has become very important in the digital age, as it defines the minimum amount of resources (including time) required to transmit a message or other information. The algorithms that are used to compress computer files, images and mobile-phone transmissions take advantage of the fact that the information is *not* random, *i.e.,* has a relatively low entropy, to represent it with less data.

There is another interesting aspect to the relationship between information and entropy. Reducing thermodynamic entropy requires work, or energy. Similarly, manipulating information and reducing its entropy (such as in compressing it) requires work. There are people working on the theory of quantifying exactly what is the minimum amount of energy required for specific computational tasks.

We don't always think about it, but the information technologies that we take advantage of use huge amounts of energy. Google is a major energy consumer. In 2009, the London Times published an article saying that a Google search produced about 1 g of CO_2, and this number is widely cited. This claim was quickly denied by Google, which a couple of years later released its own estimates of its energy consumption, indicating that a single search uses 0.3 watth (the energy required to power a 60-watt light bulb for about 20 s, or 3.600 J). If produced by burning coal, this would produce about 0.2 g of CO_2, far less than the original estimate. To its credit, Google has also invested heavily in renewable energy sources.

V. The second law

Now that we have defined this new state function, entropy (twice in fact), we can state the second law of thermodynamics. There are several equivalent ways to state this law, but the most common one is probably the following:

> For a spontaneous process, the total entropy of the system and its surroundings increases.

This leaves us with two questions, though:

- What do we mean by spontaneous?
- How do we define, measure or calculate the entropy change for the surroundings?

By spontaneous, we mean that the process will occur without any mechanical work being applied to the system. In other words, $w \leq 0$. If $w < 0$, then we can actually use the process to *do* work. A gas expansion is an example of a spontaneous process, and one that can do work. Compressing a gas, on the other hand, requires work.

For a specified process, the entropy change for the surroundings is defined as:

$$\Delta S_{\text{surr}} = -\frac{q}{T}$$

The important point here is that q is the heat absorbed by the system for the actual process, *not* the maximum-work reversible process. The entropy of the surroundings is *not* a function of the state of the system.

The entropy of the surroundings increases whenever heat flows from the system outwards. This only occurs when the system is warmer than the surroundings, and the flow of heat represents an increase in the disorder of the surroundings. If the system has a lower temperature than the surroundings, heat will flow inward, and the entropy of the surroundings will decrease. These are the only ways that the system can affect the entropy of the surroundings.

If there is a net flow of heat from the system to the surroundings, the entropy of the system can *decrease* in a spontaneous process.

Consider two cases for the expansion of an ideal gas:

1. Adiabatic expansion with no work. $\Delta E = 0$, $q = w = 0$.

   $$\Delta S_{\text{univ}} = \Delta S_{\text{sys}} + \Delta S_{\text{surr}} = nR \ln\left(\frac{V_2}{V_1}\right)$$

   So long as $V_2 > V_1$, $\Delta S_{\text{univ}} > 0$, and the process is spontaneous, which we knew already! But, now we have a way of quantifying the tendency of the gas to expand.

2. Reversible isothermal expansion with maximum work. $\Delta E = 0$, $q = q_{\text{rev}}$

   $$\begin{aligned}\Delta S_{\text{univ}} &= \Delta S_{\text{sys}} + \Delta S_{\text{surr}} \\ &= \frac{q_{\text{rev}}}{T} + \frac{-q_{\text{rev}}}{T} \\ &= 0\end{aligned}$$

   So, the reversible process is right on the edge of being spontaneous, which we can take as the meaning of a reversible process.

The key thing to keep in mind about the second law is that it refers to the entropy of universe, not just the system. This means that the entropy of the system can decrease in a spontaneous process, as long as there is a flow of heat to the surroundings. Heat flow to the surroundings can only occur if the system is warmer than its surroundings, and the overall effect is to equilibrate the temperature of the universe, which represents an increase in entropy of the universe.

Although we can use the second law in this form, this approach becomes quite awkward when it is applied to things like chemical reactions, which is the major context in which thermodynamics is used in biology. In the next section, we consider the thermodynamics of chemical reactions and introduce a more practical way of applying the second law.

5.3 Thermodynamics of Chemical Reactions

I. E and ΔE reconsidered

Most generally, the change in internal energy for a process, ΔE, is defined by the first law of thermodynamics, stated as $\Delta E = q + w$, where q is the heat absorbed by the system and w is the work done on the system. Heat, in turn, is defined as the flow of energy that can increase temperature (but doesn't necessarily for a given process), and work is the integral of force with respect to distance. The internal energy of the system, E, is a state function, and ΔE is independent of the path taken from the beginning state to the ending state. The heat, q, and work, w, however, are path dependent. But, whatever the path ΔE must equal $q + w$.

Although these definitions stand on their own, it is helpful to have a molecular interpretation. When considering an ideal gas, the only form of energy present is the translational kinetic energy, and E can be calculated from the relationship $E_{\mathrm{k}} = 3kT/2$. Any change in internal energy, ΔE, can only be due only to a change in temperature.

The two distinguishable components of the internal energy are the kinetic energy and potential energy. In the simplest case, an ideal gas, there is only the kinetic energy associated with translational motion, and $E = 3nRT/2$, for n moles at temperature T. Since the particles making up the gas have no internal structure, there is no potential energy or internal thermal motions. If the temperature is the same at the beginning and end of the process, $\Delta E = 0$.

When we deal with real molecules, even in the gas phase, things get more complicated. There are now internal motions that contribute to the kinetic energy, and the internal structures of the molecules give rise to potential energy. The potential energy represents energy that can be converted to heat or mechanical energy if the structures of the molecules change, such as by forming or breaking chemical bonds. Potential energy can also be absorbed or released by changes that do not involve covalent bonds, such as the formation or breaking of hydrogen bonds or other "weak" interactions.

In liquids and solids, there are additional forms of potential energy, due to interactions among molecules.

For systems involving real molecules, the first law ($\Delta E = q + w$) still holds for any process, but there may be redistributions of kinetic and potential energy within the system that do not change the total internal energy, E. These redistributions will not be reflected in ΔE.

II. Enthalpy (H)

For practical problems involving chemical reactions, we usually apply the second law of thermodynamics using an approach developed by Josiah Willard Gibbs (1839–1903), who was arguably the first really great American physical or theoretical scientist. Gibbs, working mostly on his own, put together the ideas of thermodynamics into a consistent system. He spent his entire career at Yale College (as it was then called), and published all of his work in the Transactions of the Connecticut Academy of Sciences. His work was very slow to be fully appreciated, especially in his home country.

One of Gibbs' major accomplishments was formulating the application of thermodynamics to chemical reactions. In doing so, he introduced two new state functions, which we now call enthalpy (H) and the Gibbs free energy (G).

The formal definition of enthalpy is:

$$H = E + PV$$

where P and V are pressure and volume respectively. Like E, P and V are state functions, so H must be also. At first glance, however, determining H or ΔH doesn't look any easier than determining E or ΔE!

But, for chemical reactions, especially in relatively dilute solutions, we can often impose additional restrictions. First, and most important, we can specify that the the process occurs under constant pressure. Using the subscripts 1 and 2 to indicate the beginning and ending states, respectively, the change in enthalpy can then be written:

$$\begin{aligned}\Delta H &= H_2 - H_1 = E_2 - E_1 + PV_2 - PV_1 \\ &= \Delta E + P\Delta V\end{aligned}$$

If ΔV is not zero, the volume change represents work, as we discussed in the case of gas expansion or compression. For most chemical reactions, this is the only form of work associated with the reaction.[2] If the volume of the system increases, then work is done *by* the system, and w is negative. If this form of work is designated w_{p}(the subscript "p" indicating constant pressure), then we can write:

$$w_{\mathrm{p}} = -P\Delta V$$

Assuming that this is the only form of work associated with the process at constant pressure, and designating the associated heat absorbed q_{p}, we have:

$$\begin{aligned}\Delta H &= \Delta E + P\Delta V = q_{\mathrm{p}} + w_{\mathrm{p}} - w_{\mathrm{p}} \\ &= q_{\mathrm{p}}\end{aligned}$$

---

[2] Notable exceptions are found in the molecular motors in living organisms, as well as some synthetic systems, but here we are focusing on relatively simple chemical reactions.

This gives the usual working definition for ΔH: The heat absorbed by the system at constant pressure. Although q is not, in general a state function, because q_p refers to a specific path between two states, it is a state function. In this respect, the use of q_p to define ΔH is analogous to the use of q_rev to define ΔS. This makes it relatively easy to measure ΔH for a reaction, using a calorimeter.

With the above restrictions, we can also express ΔH in terms of ΔE:

$$\Delta H = \Delta E - w_\mathrm{p}$$

III. ΔG, the change in Gibbs free energy.

The free energy function named for Gibbs, G, is defined as:

$$G = H - TS_\mathrm{sys}$$

and the change in G *at constant temperature* is:

$$\Delta G = \Delta H - T\Delta S_\mathrm{sys}$$

As noted earlier, the term $T\Delta S_\mathrm{sys}$ represents the heat absorbed during the reversible path. As argued above, ΔH is the heat absorbed by the process at constant pressure, and with no work other than w_p. Recall that the entropy change for the surroundings for a specific process is defined as:

$$\Delta S_\mathrm{surr} = -\frac{q}{T}$$

For the constant pressure process (again assuming constant temperature):

$$\Delta S_\mathrm{surr} = -\frac{q_\mathrm{p}}{T} = -\frac{\Delta H}{T}$$

and, for any constant temperature process:

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

If we then restrict ourselves to processes where both pressure and temperature are constant, the total entropy change for the universe is:

$$\begin{aligned}\Delta S_\mathrm{univ} &= \Delta S_\mathrm{surr} + \Delta S_\mathrm{sys} = -\frac{\Delta H}{T} - \Delta S_\mathrm{sys} \\ &= -\frac{\Delta G}{T}\end{aligned}$$

This result demonstrates that the sign of ΔG indicates whether or not the process is spontaneous: If ΔG is negative, ΔS_univ is positive and the process is favorable. If ΔG is positive, ΔS_univ is negative and the process is unfavorable. Because ΔG is a state function, as are ΔH, T and ΔS, we can determine whether or not a process is spontaneous without consideration of the surroundings, given the constraint of constant pressure. (Here we are assuming that temperature is also constant, but this assumption is not necessary in order to use the Gibbs free energy to determine if a process is spontaneous; The math just gets more complicated if temperature changes.)

IV. Free energy and available energy for work.

Earlier (page 141), another free energy function, the Helmholz free energy, was introduced and identified with the work associated by a process carried out along a reversible pathway:

$$\Delta F = w_{\text{rev}}$$

Since the reversible process yields the maximum amount of work obtainable from a transition between two states, $-\Delta F$ is the maximum work available from a favorable process. For an unfavorable process, ΔF is the minimum amount of work required to drive the process.

Though it's not very obvious from the relationship above, the Helmoltz free energy, F, is closely related to the Gibbs free energy, G. From the first law, we can write:

$$\Delta E = q_{\text{rev}} + w_{\text{rev}}$$
$$w_{\text{rev}} = \Delta E - q_{\text{rev}}$$

Therefore:

$$\Delta F = \Delta E - q_{\text{rev}}$$

Recall that ΔS, for a process at constant temperature, is q_{rev}/T. Substituting, we have:

$$\Delta F = \Delta E - T\Delta S$$

The formal definition of the Helmholtz free energy is:

$$F = E - TS$$

For comparison, the definition of the Gibbs free energy is:

$$G = H - TS$$

and the change in Gibbs free energy at constant temperature is:

$$\Delta G = \Delta H - T\Delta S$$

Thus, the difference between F and G is the difference between E and H. From before, the definition of H is:

$$H = E + PV$$

where P and V are the pressure and volume, respectively. For a process at constant pressure:

$$\Delta H = \Delta E + P\Delta V$$

Therefore, ΔG and ΔF are related according to:

$$\Delta G = \Delta F + P\Delta V$$

The term $P\Delta V$ represents the work due to any change in volume, and can be either positive or negative. Since $-\Delta F$ is the maximum amount of work available from a process, $-\Delta G$ is that amount of work minus any work due to a volume change.

If the the volume change is negligible, then $\Delta H = \Delta E$, and $\Delta G = \Delta F$. With this provision, ΔG can be interpreted as the minimum work required to drive an unfavorable process (for $\Delta G > 0$) or the negative of the maximum work that can be obtained from a favorable process ($\Delta G < 0$). This condition is often satisfied for chemical reactions in dilute solution and is frequently assumed to be valid for such reactions.

V. Free energy changes for chemical reactions

Most of the transformations of energy in biology involve chemical reactions. When we considered gas expansion problems, one of our underlying assumptions was that the internal energy of the system was only a function of temperature. But, when chemical reactions are involved, there is usually a change in the internal energy of the molecules. Forming and breaking chemical bonds almost always involves a change in potential energy. Bonds form because the atoms involved find a lower energy state in which they "share" electrons. The changes in internal potential energy must be accompanied by either changes in the thermal (kinetic) energy of the system, heat flow to or from the surroundings or mechanical work, or some combination of the above.

In biological systems, the energy changes associated with reactions are frequently coupled to other processes. In that way, the energy change for a favorable process can be used to drive an unfavorable one. The most commonly used tool for keeping track of, and understanding, these energy conversions is the Gibbs free energy.

For the simple case of the interconversion of one compound to another, we can write a generic chemical reaction as:

$$\mathrm{A} \rightleftharpoons \mathrm{B}$$

In principle, any chemical reaction is reversible, in that there is a finite probability that it can occur in either direction. The probability that any molecule of A will be converted to B is the same as for any other molecule of A. As a consequence, the total rate of conversion of A to B will be proportional to the number of A molecules, or equivalently, their concentration. Similarly, the rate of conversion of B to A will be proportional to the concentration of molecules of B. The net rate in the change in the concentration of A will be the difference between these two rates, which we can write as a differential equation:

$$\frac{d[\mathrm{A}]}{dt} = -k_\mathrm{f}[\mathrm{A}] + k_\mathrm{r}[\mathrm{B}]$$

where k_f and k_r are rate constants that relate the probability of each reaction to the concentration of the molecules. The net rate of conversion from B to A is the negative of the rate of conversion from A to B, which we can write as:

$$\frac{d[\mathrm{B}]}{dt} = -\frac{d[\mathrm{A}]}{dt} = k_\mathrm{f}[\mathrm{A}] - k_\mathrm{r}[\mathrm{B}]$$

We won't worry about solving these equation, except to note that if we start out with all A, some of it will be converted to B. As the concentration of A decreases, the rate of conversion to B will decrease, and the rate of conversion of B to A will increase. At some point, the flow in each direction will be equal, and the rate of change in concentration will be zero. At this point:

$$\frac{d[\mathrm{A}]}{dt} = 0$$

$$k_\mathrm{f}[\mathrm{A}] = k_\mathrm{r}[\mathrm{B}]$$

$$\frac{[\mathrm{B}]}{[\mathrm{A}]} = \frac{k_\mathrm{f}}{k_\mathrm{r}}$$

At this point, we say that the reaction is at equilibrium, and we can define an equilibrium constant:

$$K_\mathrm{eq} = \frac{k_\mathrm{f}}{k_\mathrm{r}} = \frac{[\mathrm{B}]_\mathrm{eq}}{[\mathrm{A}]_\mathrm{eq}}$$

where the subscripts indicate that the concentrations are those that are measured at equilibrium. This is not the most rigorous derivation of an equilibrium constant, but it will do.

One of the implications of the equilibrium state is that there is no way in which to obtain work from the reaction. If the system is at equilibrium, then an infinitesimal amount of work could shift it in either direction, by an infinitesimal amount. Looked at another way, if we were to have an immense volume of the reactants at the equilibrium concentrations, one mole of A could be converted to one mole of B (or vise versa) without doing any work, provided that the concentrations didn't change significantly. Therefore, the free energy change for the reaction must be zero.

Suppose, though, that we were to change the concentrations from their equilibrium values, say by adding one or the other of the reactants. Then, we would expect the concentrations to shift towards those that would satisfy the equilibrium constant, and during this process we could, at least in principle, extract some work. At concentrations that do not satisfy the equilibrium conditions, ΔG is negative if the concentrations favor the forward reaction, and ΔG is positive if the concentrations favor the reverse reactions.

For any set of reactant and product concentrations, ΔG is defined as the free energy for converting one mole of reactants to one mole of products, A to B in the example used here, in a volume large enough that the concentrations do not change significantly.

We can derive an expression to calculate ΔG for an arbitrary set of concentrations by considering the sum of three processes:

1. $\mathrm{A}([\mathrm{A}]_{\mathrm{eq}}) \longrightarrow \mathrm{B}([\mathrm{B}]_{\mathrm{eq}})$
2. $\mathrm{A}([\mathrm{A}]) \longrightarrow \mathrm{A}([\mathrm{A}]_{\mathrm{eq}})$
3. $\mathrm{B}([\mathrm{B}]_{\mathrm{eq}}) \longrightarrow \mathrm{B}([\mathrm{B}])$

where the concentrations of A and B in the three processes are indicated in the parentheses: $[\mathrm{A}]_{\mathrm{eq}}$ and $[\mathrm{B}]_{\mathrm{eq}}$ represent the equilibrium concentrations and [A] and [B] are the concentrations at which ΔG is to be calculated. The sum of the three reactions is:

4. $\mathrm{A}([\mathrm{A}]) \longrightarrow \mathrm{B}([\mathrm{B}])$

A general principle of chemical thermodynamics states that the changes in state functions (including ΔE, ΔH, ΔG and ΔS) associated with individual reactions can be added to yield the change in state functions for the corresponding sum of the reactions. Thus, we can write:

$$\Delta G_4 = \Delta G_1 + \Delta G_2 + \Delta G_3$$

Since process 1 represents the equilibrium condition, $\Delta G_1 = 0$. Process 2 represents a change in the concentration of A, from the concentration of interest, [A], to the equilibrium concentration, $[\mathrm{A}]_{\mathrm{eq}}$. If we assume the results for ideal gasses can be extended to the reaction in dilute solution, the reversible work for the change in concentration, w_{rev}, is calculated as:

$$w_{\mathrm{rev}} = RT \ln \frac{[\mathrm{A}]_{\mathrm{eq}}}{[\mathrm{A}]}$$

If we further assume that the volume change at constant pressure is negligible, then $w_{\mathrm{rev}} = \Delta G$, and we can write the free energy change for process 2 as:

$$\Delta G_2 = RT \ln \frac{[\mathrm{A}]_{\mathrm{eq}}}{[\mathrm{A}]}$$

From analogous arguments, the free energy change for process 3 is:

$$\Delta G_3 = RT \ln \frac{[\mathrm{B}]}{[\mathrm{B}]_{\mathrm{eq}}}$$

The free energy change for process 4, the reaction at concentrations [A] and [B] is then given by:

$$\begin{aligned}
\Delta G_4 &= \Delta G_1 + \Delta G_2 + \Delta G_3 \\
&= 0 + RT \ln \frac{[\mathrm{A}]_{\mathrm{eq}}}{[\mathrm{A}]} + RT \ln \frac{[\mathrm{B}]}{[\mathrm{B}]_{\mathrm{eq}}} \\
&= RT \ln \frac{[\mathrm{B}][\mathrm{A}]_{\mathrm{eq}}}{[\mathrm{A}][\mathrm{B}]_{\mathrm{eq}}} \\
&= RT \ln \frac{[\mathrm{B}]}{[\mathrm{A}]} - RT \ln \frac{[\mathrm{B}]_{\mathrm{eq}}}{[\mathrm{A}]_{\mathrm{eq}}} \\
&= RT \ln \frac{[\mathrm{B}]}{[\mathrm{A}]} - RT \ln K_{\mathrm{eq}}
\end{aligned}$$

In order to define and calculate the free energy change for a reaction at different concentrations, it is useful to define a standard state for measuring and reporting free energy changes, and the most widely used standard is to specify that all of the reactants and products are at 1 M concentration. (Or, for the gas phase, that the pressures of each component is 1 atm.) The free energy change under these conditions is defined as the standard free energy change, ΔG°. From the equation above, the standard free energy change is calculated as:

$$\begin{aligned}
\Delta G^\circ &= RT \ln \frac{1\,\mathrm{M}}{1\,\mathrm{M}} - RT \ln K_{\mathrm{eq}} \\
&= -RT \ln K_{\mathrm{eq}}
\end{aligned}$$

This then provides a link between the standard free energy change and the equilibrium constant. Both represent essentially the same thing: the extent to which one side of the reaction is favored. If the forward reaction ($\mathrm{A} \longrightarrow \mathrm{B}$) is favored, then $K_{\mathrm{eq}} > 1$ and $\Delta G^\circ < 0$. A negative value of ΔG° also implies that work can be obtained from the reaction if the reactants are initially at their standard-state conditions.

Measuring the equilibrium constant is generally the most straight forward way of measuring ΔG°. Once this value is known, the free energy change for other concentrations can be easily calculated. If ΔH is also know, for instance by measurement with a calorimeter, the entropy change, ΔS, can then be calculated from the relationship $\Delta G = \Delta H - T\Delta S_{\mathrm{sys}}$.

The commonly used form of the equation for calculating the free energy change at concentrations other than the standard state ones is:

$$\Delta G = \Delta G^\circ + RT \ln \frac{[\mathrm{B}]}{[\mathrm{A}]}$$

Especially for reactions that involve more than one reactant or product, it is convenient to introduce a quantity called the *reaction quotient*, Q. This quantity is defined as the product of the reaction product concentrations divided by the product of reactant concentrations. For the reaction:

$$\mathrm{aA} + \mathrm{bB} \rightleftharpoons \mathrm{cC} + \mathrm{dD}$$

The reaction quotient is:

$$Q = \frac{[\mathrm{C}]^c[\mathrm{D}]^d}{[\mathrm{A}]^a[\mathrm{B}]^b}$$

When the reaction is at equilibrium, Q is equal to the equilibrium constant, K_{eq}, and the free energy associated with the reaction is zero. At other concentrations, the free energy change is calculated as:

$$\Delta G = \Delta G^\circ + RT \ln Q$$

VI. Concentrations and standard states

Things become a little more complicated when the number of reactant and product molecules are not equal. For instance, consider the reaction:

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{C}$$

We can write an equilibrium constant:

$$K_{\mathrm{eq}} = \frac{[\mathrm{C}]_{\mathrm{eq}}}{[\mathrm{A}]_{\mathrm{eq}}[\mathrm{B}]_{\mathrm{eq}}}$$

And the expression for the free energy change is:

$$\Delta G = \Delta G^\circ + RT \ln \frac{[\mathrm{C}]}{[\mathrm{A}][\mathrm{B}]}$$

But, there are some funny things here:

- The equilibrium constant has dimensions of inverse concentration. If different units are used for concentration, the value of the equilibrium constant will change.
- The reaction quotient (Q) also has units of inverse concentration. How can we take the logarithm of a quantity with units?
- The value of ΔG will also depend on the concentration units chosen.

We can solve the problem of the logarithms by stipulating that the reaction quotient should be written in terms of the ratios of the actual concentrations and the standard state concentrations. If the standard state concentrations are 1 M, the the reaction quotient is written as:

$$Q = \frac{[\mathrm{C}]/1\,\mathrm{M}}{([\mathrm{A}]/1\,\mathrm{M})([\mathrm{B}]/1\,\mathrm{M})}$$

Now, Q is dimensionless, and it is OK to take a logarithm. But, it is still true that K_{eq}, ΔG and ΔG° will all depend on the units of concentration chosen.

The underlying reason for this is that when the number of molecules changes in a reaction there is an intrinsic change in entropy. In the example above, there will be a reduction of entropy when two molecules are converted to one, and this will disfavor the reaction. If the standard state is, for instance, 1 mM instead of 1 M, then the loss of entropy will be larger, and this will be reflected in the numerical value of the equilibrium constant and the free energy change.

VII. Calculating the entropy change for a bimolecular reaction.

Following the approach used when considering the expansion of a gas (pages 144–145), we can use the statistical definition of entropy to estimate the entropy loss due to converting two molecules into one.

As in the example of gas expansion, we begin by dividing up the volume of interest into small cubes with volume V_c. The total volume is $N_{\mathrm{c}}V_{\mathrm{c}}$, where N_{c} is the number of small cubes.

Considering just one molecule each of A and B, the number of ways of placing these two molecules in the lattice of cubes, $\Omega_{\mathrm{A,B}}$ is N_{c}^2. This again assumes that the number of cubes is much larger than the total number of molecules. Because, the two molecules are assumed to be different, N_{c}^2 is not divided by two, since swapping the positions of A and B results in a distinct arrangement. If the two reacting molecules were identical, then $\Omega_{\mathrm{A,B}}$ would equal $N_{\mathrm{c}}^2/2$

The number of arrangements of a single molecule of C in the volume is $\Omega_{\mathrm{C}} = N_{\mathrm{c}}$.

The entropy change for a single pair of A and B molecules being converted to C is then given by:

$$\begin{aligned}\Delta S &= k\ln\left(\frac{\Omega_{\mathrm{C}}}{\Omega_{\mathrm{A,B}}}\right) = k\ln\left(\frac{N_{\mathrm{c}}}{N_{\mathrm{c}}^2}\right)\\ &= -k\ln N_{\mathrm{c}}\end{aligned}$$

This result indicates that the entropy decreases for the reaction, as we expect. But, the result is also somewhat problematic, because the number of cubes, N_{c}, does not cancel out, as it did for the example of a gas expansion. Therefore, the size of the cubes does matter in this case.

Though there is no absolutely certain way to define the size of the cubes, we can at least make a reasonable estimate. The cubes should be just big enough to hold one of the reactant or product molecules. The product is probably larger than either of the reactants, making it difficult to specify a single size, but a reasonable estimate would be a cube 1 nm on each side, or a volume of 1 nm^3. The number of cubes is then calculated by dividing the reaction volume, V, by the volume of a single cube, V_{c}. We will specify that the total volume is given in liters. Taking into account the necessary conversion factors, N_c can be expressed as:

$$N_\mathrm{c} = \frac{V(\mathrm{L})}{1\,\mathrm{nm}^3} \times \frac{10^3\,\mathrm{m}^3}{L} \times \left(\frac{10^9\,\mathrm{nm}}{\mathrm{m}}\right)^3 \approx V(\mathrm{L}) \times 10^{24}$$

The entropy change is then:

$$\Delta S = -k\ln(N_\mathrm{c}) = -k\ln(V(\mathrm{L}) \times 10^{24})$$

Note that the entropy change is closely related to the reaction volume, V, and becomes more negative with larger volume. This can be understood by recognizing that all of the molecules gain entropy when the volume is increased, but the two reactant molecules, together, lose more than the single product molecule.

For one mole each of A and B converted to C, the result above is multiplied by Avogadro's number:

$$\begin{aligned}\Delta S &= -N_\mathrm{A}k\ln(V(\mathrm{L}) \times 10^{24})\\ &= -R\ln(V(\mathrm{L}) \times 10^{24})\end{aligned}$$

If the volume is specified as 1 L, a numerical result can be calculated:

$$\begin{aligned}\Delta S &= -R\ln(10^{24}) = -8.31\,\mathrm{J/K} \times 53\\ &= -460\,\mathrm{J/K}\end{aligned}$$

The contribution to the free energy change at 298 K is:

$$-T\Delta S = -460\,\mathrm{J/K} \times 440\,\mathrm{J/K} = 140\,\mathrm{kJ}$$

In order for the reaction to be favorable, this large contribution would have to be compensated for by other factors, such a favorable enthalpy change due to covalent changes in the molecule or other entropic factors.

This simple calculation shows that the change of entropy associated with a reaction in which the number of molecules changes can be very significant. It is particularly significant in the context of biological systems, where very large molecules, such as proteins, nucleic acids and polysaccharides, are assembled from small building blocks (amino acids, nucleotides and sugars). However, accurately calculating these entropy changes is challenging and a source of some controversy. The calculation above considers only the loss of translational degrees of freedom, and a large reduction in rotational freedom is also to be expected in a bimolecular reaction. Other factors, including the changes in internal motions (bond rotations and vibrations) and interactions with solvent molecules may also contribute to the total entropy change. Among various estimates, the one above represents the lower range of calculated entropy changes for the association of two molecules.

VIII. Activity versus concentration

The expression relating the free energy and concentrations is strictly valid only for ideal gas molecules. If there is any tendency for the molecules to interact with one another or with solvent molecules, this can affect the free energy change. The formal way of dealing with this problem is to replace the concentrations with "activities" for each of the reactants and products:

$$\Delta G = \Delta G^{\circ} + RT \ln \frac{a_A a_B}{a_C a_D}$$

The activities are dimensionless quantities, but they are defined relative to a specific standard state (*e.g.,* 1 M). At relatively low concentrations, the activities are often close to being proportional to concentration and are often expressed in terms of activity coefficients, γ:

$$a_A = \gamma_A[\mathrm{A}]$$

The activity coefficient can depend on concentration. If $\gamma = 1/(1\,\mathrm{M})$, then the species behaves ideally, and the activity is equal to the concentration divided by the standard state concentration.

For most biochemical reactions, at relatively low concentrations, the deviations from ideal behavior are not very large (or at least we hope that they aren't), and we rarely worry about them. As a practical matter, we usually assume that we can use concentrations in the equilibrium expression.

5.4 "Chemical Energy" and Metabolism

It is common to hear people talk about "chemical energy" or refer to certain compounds as having "high energy" or "high energy bonds". We can even eat "energy bars" and drink "energy drinks". This language can be rather vague, or even misleading. Here we consider some specific examples of biochemical processes and the meaning of chemical energy in this context.

I. Glucose oxidation

The most useful measure of chemical energy is the free energy change. For instance, the oxidation of glucose by molecular oxygen is written as:

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

Like any reaction, this reaction is reversible and has an equilibrium constant. The extent to which it is favorable depends on the concentrations of all of the reactants and products. The standard free energy change for this reaction is about $-2.700\,\mathrm{kJ/mol}$, meaning that it is an extremely favorable reaction when all of the reactants are at their standard states (1 M or 1 atm). But, if the O_2 concentration is very low, the reaction is much less favorable. Prior to about 2.5 billion years ago, the concentration of oxygen in earth’s atmosphere was about 10^5-fold less than it is now. Under those conditions, the oxidation of glucose was not favorable at all. Glucose was used as a source of “energy” through glycolysis (a partial enzymatic breakdown of glucose), and this is still an important metabolic reaction. But glycolysis is a much less favorable overall reaction than complete oxidation (at the present atmospheric concentrations of O_2 and CO_2), meaning that glucose or other carbohydrates can provide much less “energy” under anaerobic conditions.

The nutritional calories that are listed for foods are *not* free energy changes. They are measured by burning the foods (or their digestible ingredients) and directly measuring the heat in a calorimeter at constant volume, with an excess of oxygen. So, these calories actually represent ΔE. For glucose, the dietary value is about $4\,\mathrm{kcal/g}$, which corresponds to about $175\,\mathrm{kJ/mol}$, much less than the standard free energy change. This is because there is a large entropy change in the reaction as well, as 7 molecules are converted to 12.

The nutritionists then estimate how much work must be done by the body in order to offset the metabolism of, for instance, $1\,\mathrm{g}$ of sugar. Most of the free energy is lost as heat. Relating food calories to free energy changes is not straight forward (to me at least!)

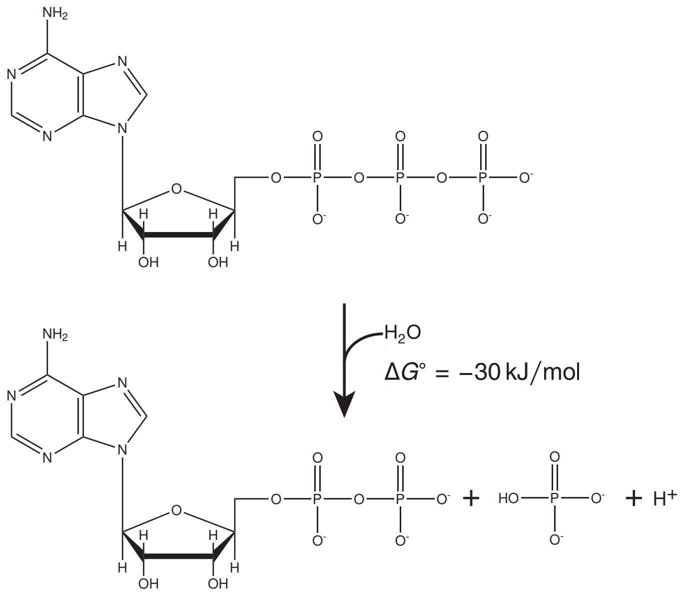
The important point is that in our atmosphere, the oxidation of glucose and related molecules is very favorable, and organisms can couple this very favorable reaction to processes that otherwise would not be favorable.

The synthesis of glucose and related molecules is *unfavorable* and depends on coupling to other favorable reactions, which are driven by absorbing the energy of light from the sun.

For our cars and some electric power plants, we (mostly) use as a fuel hydrocarbons or coal, which also are oxidized with a large negative free energy change. All of the energy in hydrocarbons and coal, was originally captured by photosynthesis. Plants and animals that ate plants died and the carbon was converted into the forms we now extract and burn. We are quickly consuming the vast amounts that were accumulated over billions of years, and we are moving the carbon into the atmosphere.

II. ATP hydrolysis

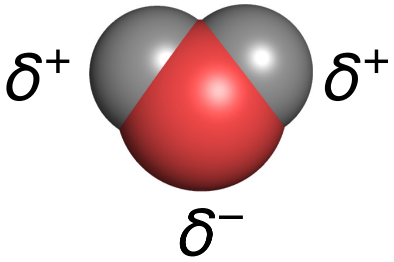
Another compound that is often described as having “high energy” is ATP, with the hydrolysis reaction shown below:



The bond linking the second and third phosphate groups is often described as a “high energy bond”. In fact, there is nothing very special about this bond. What is important is that the reaction has a large negative free energy change under physiological conditions, about $-30\,\mathrm{kJ/mol}$.

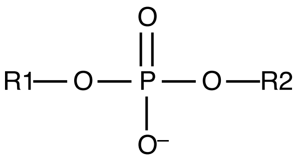
We won’t worry for now how this compound gets made in the first place. Basically, ATP serves as a kind of energy currency, it is formed during the oxidation of glucose (and other favorable reactions) and it is hydrolyzed to provide energy for other processes.

In considering why this reaction is so favorable, the first thing to emphasize about this reaction is that it occurs in the presence of water, and water is one of the reactants. We will talk more about water later, but for now, the important thing is that water is a polar molecule, meaning that the electrons of the molecule are unevenly distributed leading to partial charges on the hydrogen and oxygen atoms:



The partial charges of the water molecule can interact with charges on other molecules making the charged forms much more stable than they would be otherwise.

ATP, as suggested by its name, has three phosphate groups:

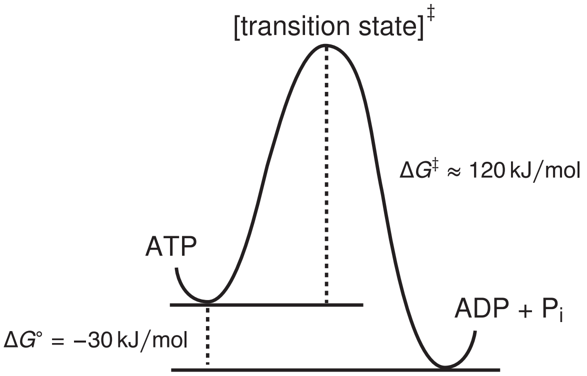


$$\mathrm{R1{-}O{-}\underset{\underset{O^-}{|}}{\overset{\overset{O}{\|}}{P}}{-}O{-}R2}$$

In water, at neutral pH, one or more of the oxygen atoms will be in an ionized, negatively charged, state. In ATP, there are four negative charges on the three phosphates (one each on the first and second and two on the third phosphate). This results in a quite high density of negative charge. When the third phosphate is removed by hydrolysis, the charges are not so close together and are shielded by the water molecules, leading to a reduction in potential energy. There are other factors, but this is the major one. Further hydrolysis, of the remaining phosphate-phosphate ester bond, is also favorable, but not by quite so much.

Even if a reaction is thermodynamically favored, it may not occur very rapidly. This is obviously true for oxidation of glucose, and it is also true for ATP hydrolysis, for which the half time is about 20 days at neutral pH and 60 °C.

It's useful to represent the relationship between thermodynamics and kinetics as a reaction coordinate, or energy profile:



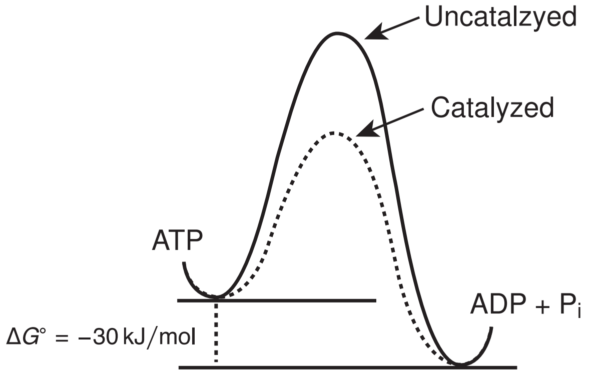
The central idea here is that the molecules have to acquire a high energy state in order for the reaction to proceed. This high energy state is called a "transition state" or "activation complex" and is a bit of a mythical beast. It is imagined to be in equilibrium with the products and reactants, but present at only very low concentrations. As soon as it forms, it breaks down to either products or reactants, with equal probability. The higher the transition-state energy, the slower the reaction.

The theory for relating reaction rates to energies was developed by Henry Eyring. See the painting in the lobby of building named for him!

In the absence of a catalyst, the reaction rate under physiological conditions would be insignificant. This is important: If ATP spontaneously hydrolyzed, it wouldn't be an effective way of storing energy. It would be like trying to use a fuel that spontaneously combusts.

The key to using chemical energy in biology is controlling the reaction by catalysts, enzymes. Enzyme can increase the rate of a reaction by several orders of magnitude.

A somewhat glib way of saying what enzymes, or other catalysts, do, is to say that they lower the transition state energy[3]:

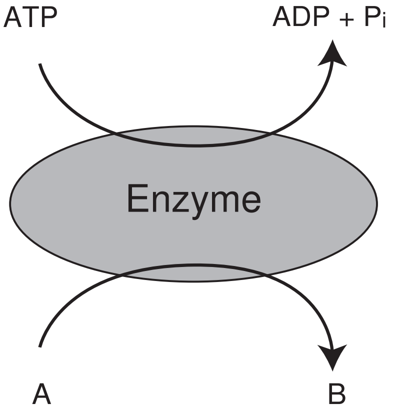


How do enzymes increase reaction rates? This is a major sub-discipline of biochemistry, and there are many different effects, not all of which are fully understood or agreed on. But, the basic idea is that the enzyme, which is a protein (or sometimes RNA) molecule, binds to the reactants and creates a local chemical environment that makes the reaction much more likely to proceed. One important part of this is that the local concentrations are much higher than they are when the molecules are free in solution. The enzyme can also help stabilize charges that form in the transition state.

An enzyme cannot alter the thermodynamics of a reaction! If it catalyzes the forward reaction, it must also catalyze the reverse reaction by the same factor. Any scheme that suggests that both reactions are not catalyzed equally violates the first law!

III. Enzymatic coupling

An enzyme that simply catalyzed ATP hydrolysis would not be very useful. Essentially all enzymes that catalyze ATP hydrolysis do so in a way that couples the favorable reaction with another reaction that would otherwise be unfavorable. The structure and mechanism of the enzyme is such that the hydrolysis reaction can only occur if the other reaction also occurs.



---

[3] The statement is glib because it doesn't really provide any additional information about the mechanism of the catalyst. But, considering the transition state and what determines its free energy is a valuable way of framing mechanistic questions

The enzyme links the two reactions together so that one can’t occur without the other, so that the overall reaction is:

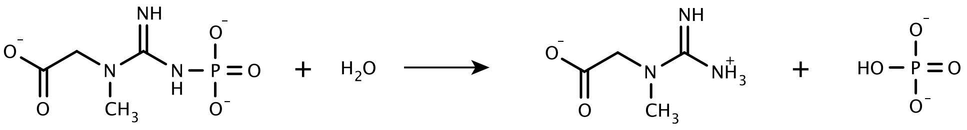
$$\mathrm{A} + \mathrm{ATP} \rightleftharpoons \mathrm{B} + \mathrm{ADP} + \mathrm{P_i}$$

The free energy changes for the individual reactions are added together. Even if ΔG for the second reaction is positive, the overall reaction can be favorable.

The coupled process can be a chemical reaction, but it can also be a physical process such as movements of molecules across a membrane against a concentration gradient or the generation of mechanical force, as in muscle.

The details of how this coupling occurs depends on the structures of the proteins that catalyze the reactions.

The enzyme creatine kinase offers a good example of enzymatic coupling. Like ATP, creatine phosphate (also called phosphocreatine) has a large negative free energy of hydrolysis



$$\Delta G^{\circ} = -43\,\mathrm{kJ/mol}$$

The enzyme creatine kinase catalyzes the exchange of phosphate between ATP/ADP and creatine phosphate/creatine (CP/C):

$$\mathrm{ADP} + \mathrm{CP} \rightleftharpoons \mathrm{ATP} + \mathrm{C}$$

The standard free energy change for the overall reaction is: $-43\,\mathrm{kJ/mol} + 30\,\mathrm{kJ/ml} = -13\,\mathrm{kJ/mol}$

Creatine, creatine phosphate and the enzyme creatine kinase are found in a variety of animal tissues, including muscle and brain cells, where creatine phosphate serves as a short-term source of reserve chemical energy.

In resting muscle cells, the typical concentrations are:

4 mM ATP
0.013 mM ADP
25 mM creatine phosphate
13 mM creatine

We can calculate the free energy change for the formation of ATP from CP under these conditions:

$$\begin{aligned}
\Delta G &= \Delta G^\circ + RT \ln \frac{[\mathrm{ATP}][\mathrm{C}]}{[\mathrm{ADP}][\mathrm{C}]} \\
&= -13\,\mathrm{kJ/mol} + RT \ln \frac{4\,\mathrm{mM} \cdot 13\,\mathrm{mM}}{0.013\,\mathrm{mM} \cdot 25\,\mathrm{mM}} \\
&= -13\,\mathrm{kJ/mol} + 8.314\,\mathrm{J/K} \cdot 310\,\mathrm{K} \ln(160) \\
&= -13\,\mathrm{kJ/mol} + 13\,\mathrm{kJ/mol} \\
&\approx 0
\end{aligned}$$

The four compounds are at equilibrium, because the enzyme quickly equilibrates them. But, when there is a large demand for ATP, the concentration of ATP goes down, the concentration of ADP goes up, and the forward reaction, as written above, becomes favorable, to restore ATP concentration.

But, the reserve of creatine phosphate is relatively limited and lasts about 4 seconds in a sprinter. After that, either ATP is restored by oxidative phosphorylation, or glucose is metabolized by glycolysis. The latter provides much less ATP than oxidative metabolism and leads to the accumulation of lactic acid.