Biochemical Coupling Through Emergent Conservation Laws

JOHN C. BAEZ, JONATHAN LORAND, BLAKE S. POLLARD
AND MARU SARAZOLA

June 29, 2018

Abstract

Bazhin has analyzed ATP coupling in terms of quasiequilibrium states where fast reactions have reached an approximate steady state while slow reactions have not yet reached equilibrium. After an expository introduction to the relevant aspects of reaction network theory, we review his work and explain the role of "emergent conservation laws" in coupling. These are quantities, left unchanged by fast reactions, whose conservation forces exergonic processes such as ATP hydrolysis to drive desired endergonic processes.

1 Introduction

In the cell, chemical reactions are often "coupled" so that reactions that release energy drive reactions that are biologically useful but involve an increase in energy. But how, exactly, does coupling work?

Much is known about this question, but the literature is also full of vague explanations and oversimplifications. Coupling cannot occur in equilibrium; it arises in open systems, where the concentrations of certain chemicals are held out of equilibrium due to flows in and out [1, 3, 6]. One might thus suspect that the simplest mathematical treatment of this phenomenon would involve non-equilibrium steady states of open systems. However, Bazhin [2] has shown that some crucial aspects of coupling arise in an even simpler framework. He considers "quasi-equilibrium" states, where fast reactions have come into equilibrium and slow ones are neglected. He shows that coupling occurs already in this simple approximation.

Our goal here is two-fold. First, we review Bazhin's work in a way that readers with no training in biology or chemistry can follow. Second, we explain a fact that seems to have received little attention: in many cases, coupling relies on *emergent conservation laws*.

Conservation laws are important throughout science. Besides those that are built into the fabric of physics, such as conservation of energy and momentum, there are also many "emergent" conservation laws that hold approximately in certain circumstances. Often these arise when processes that change a given quantity happen very slowly. For example, the most common isotope of uranium decays into lead with a half-life of about 4 billion years—but for the purposes

of chemical experiments in the laboratory, it is useful to treat the amount of uranium as a conserved quantity.

The emergent conservation laws involved in biochemical coupling are of a different nature. Instead of making the processes that violate these laws happen *more slowly*, the cell uses enzymes to make other processes happen *more quickly*. At the time scales relevant to cellular metabolism, the fast processes dominate, while slowly changing quantities are effectively conserved. By a suitable choice of these emergent conserved quantities, the cell ensures that certain reactions that release energy can only occur when other "desired" reactions occur. To be sure, this is only *approximately* true, on sufficiently short time scales. But this approximation is enlightening.

Following Bazhin, our main example involves coupling to ATP hydrolysis. We consider the following schema, which abstractly describes a whole family of reactions:

$$X + ATP \;\rightleftharpoons\; ADP + XP_i \tag{1}$$

$$XP_i + Y \;\rightleftharpoons\; XY + P_i \tag{2}$$

Some concrete examples of this schema include:

- The synthesis of glutamine (XY) from glutamate (X) and ammonium (Y). This is part of the important glutamate-glutamine cycle in the central nervous system.

- The synthesis of sucrose (XY) from glucose (X) and fructose (Y). This is one of many processes whereby plants synthesize more complex sugars and starches from simpler building-blocks.

In these and other examples, reactions (1) and (2), taken together, have the effect of synthesizing a larger molecule XY out of two parts X and Y, while ATP is broken down to ADP and P_i. Thus, they have the same net effect as this other pair of reactions:

$$X + Y \;\rightleftharpoons\; XY \tag{3}$$

$$ATP \;\rightleftharpoons\; ADP + P_i \tag{4}$$

where the desired reaction (3) involves an energy increase, while (4), a deliberately simplified version of ATP hydrolysis, releases energy. But in reactions (1) and (2), these reactions are "coupled" so that ATP can only break down to $ADP + P_i$ if X + Y also turns into XY.

As we shall see, this coupling crucially relies on a conserved quantity: the total number of Y molecules plus the total number of P_i's is left unchanged by both reactions above. This fact is not a fundamental law of physics, nor even an approximate law considered fundamental in chemistry, such as conservation of phosphorus. It is an *emergent* conservation law that holds approximately in special situations. Its approximate validity relies on the fact that the cell has enzymes that make the two reactions in the above schema occur more rapidly than reactions that violate this law, such as these:

In what follows, Section 2 provides the background in chemistry needed to follow the paper. Section 3 poses the question "what is coupling?" Section 4

introduces the new reactions required for coupling ATP hydrolysis to the synthesis of XY from components X and Y, and it explains why these reactions are not yet enough for coupling. Section 5 shows that coupling occurs in a "quasiequilibrium" state where these new reactions, assumed much faster than the rest, have reached equilibrium, while the rest are neglected. Section 6 explains the importance of emergent conservation laws. Finally, Section 7 shows that the same pattern is at work in a very different example of coupling: the urea cycle.

2   Some background

In what follows, we will be working with reaction networks. A reaction network consists of a set of *reactions*, for example

$$X + Y \longrightarrow XY.$$
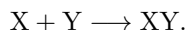
Here X, Y and XY are the *species* involved, and we interpret this reaction as species X and Y combining to form species XY. We call X and Y the *reactants* and XY the *product*. Additive combinations of species, such as $X + Y$, are called *complexes*.

The *law of mass action* states that the rate at which a reaction occurs is proportional to the product of the concentrations of the reactants. The proportionality constant is called the *rate constant*; it is a positive real number associated to a reaction that depends on chemical properties of the reaction along with the temperature, the pH of the solution, the nature of any catalysts that may be present, and so on. Every reaction has a *reverse* reaction; that is, if X and Y combine to form XY, then XY can also split into X and Y. The reverse reaction has its own rate constant.

We summarize this information by writing

$$X + Y \underset{\alpha_{\leftarrow}}{\overset{\alpha_{\rightarrow}}{\rightleftharpoons}} XY$$
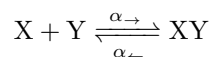
where $\alpha_{\rightarrow}$ is the rate constant for X and Y to combine and form XY, while $\alpha_{\leftarrow}$ is the rate constant for the reverse reaction.

As time passes and reactions occur, the concentration of each species will likely change. We can record this information in a collection of functions $[X] \colon \mathbb{R} \to [0, \infty)$, one for each species X, where $[X](t)$ gives the concentration of the species X at time t. This naturally leads one to consider the *rate equation* of a given reaction, which specifies the time evolution of these concentrations. The rate equation can be read off from the reaction network, and in the above example it is:

$$[\dot{X}] = -\alpha_{\rightarrow}[X][Y] + \alpha_{\leftarrow}[XY]$$
$$[\dot{Y}] = -\alpha_{\rightarrow}[X][Y] + \alpha_{\leftarrow}[XY]$$
$$[\dot{XY}] = \alpha_{\rightarrow}[X][Y] - \alpha_{\leftarrow}[XY]$$

Here $\alpha_{\rightarrow}[X][Y]$ is the rate at which the forward reaction is occurring. Similarly, $\alpha_{\leftarrow}[XY]$ is the rate at which the reverse reaction is occurring.

We say that a system is in *detailed balanced equilibrium*, or simply *equilibrium*, when every reaction occurs at the same rate as its reverse reaction.

This implies that the concentration of each species is constant in time. In our example, the condition for equilibrium is

$$\frac{\alpha_\rightarrow}{\alpha_\leftarrow} = \frac{[\mathrm{XY}]}{[\mathrm{X}][\mathrm{Y}]}$$

and the rate equation then implies that $[\dot{\mathrm{X}}] = [\dot{\mathrm{Y}}] = [\dot{\mathrm{XY}}] = 0$.

The laws of thermodynamics determine the ratio of the forward and reverse rate constants. For any reaction whatsoever, this ratio is

$$\frac{\alpha_\rightarrow}{\alpha_\leftarrow} = e^{-\Delta G^\circ / RT} \tag{5}$$

where T is the temperature, R is the ideal gas constant, and ΔG° is the free energy change under standard conditions.
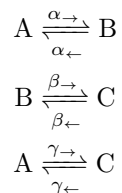
Note that if $\Delta G^\circ < 0$, then $\alpha_\rightarrow > \alpha_\leftarrow$: the rate constant of the forward reaction is larger than the rate constant of the reverse reaction. In this case one may loosely say that the forward reaction "wants" to happen "spontaneously". Such a reaction is called *exergonic*. If on the other hand $\Delta G^\circ > 0$, then the forward reaction is "non-spontaneous" and it is called *endergonic*.

The most important thing for us is that ΔG° takes a very simple form. Each species has a free energy. The free energy of a complex $\mathrm{A}_1 + \cdots + \mathrm{A}_m$ is the sum of the free energies of the species A_i. Given a reaction

$$\mathrm{A}_1 + \cdots + \mathrm{A}_m \longrightarrow \mathrm{B}_1 + \cdots + \mathrm{B}_n,$$

the free energy change ΔG° for this reaction is the free energy of $\mathrm{B}_1 + \cdots + \mathrm{B}_n$ minus the free energy of $\mathrm{A}_1 + \cdots + \mathrm{A}_m$.

As a consequence, ΔG° is additive with respect to combining multiple reactions in either series or parallel. In particular, then, the law (5) imposes relations between ratios of rate constants: for example, if we have the following more complicated set of reactions
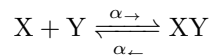
$$\mathrm{A} \underset{\alpha_\leftarrow}{\overset{\alpha_\rightarrow}{\rightleftharpoons}} \mathrm{B}$$

$$\mathrm{B} \underset{\beta_\leftarrow}{\overset{\beta_\rightarrow}{\rightleftharpoons}} \mathrm{C}$$

$$\mathrm{A} \underset{\gamma_\leftarrow}{\overset{\gamma_\rightarrow}{\rightleftharpoons}} \mathrm{C}$$

then we must have

$$\frac{\gamma_\rightarrow}{\gamma_\leftarrow} = \frac{\alpha_\rightarrow}{\alpha_\leftarrow} \frac{\beta_\rightarrow}{\beta_\leftarrow}.$$

So, not only are the rate constant ratios of reactions determined by differences in free energy, but also nontrivial relations between these ratios can arise, depending on the structure of the system of reactions in question.

3   What is coupling?

Suppose that we are in a setting in which the reaction

$$\mathrm{X} + \mathrm{Y} \underset{\alpha_\leftarrow}{\overset{\alpha_\rightarrow}{\rightleftharpoons}} \mathrm{XY}$$

takes place. Let's also assume we are interested in the production of species XY from species X and Y, but that in our system, the reverse reaction is favored to happen. This means that

$$\alpha_{\leftarrow} \gg \alpha_{\rightarrow}$$

and so in equilibrium, the concentrations of the species will satisfy

$$\frac{[\text{XY}]}{[\text{X}][\text{Y}]} \ll 1$$

which we assume undesirable. How can we influence this ratio to get a more desired outcome?

This is where the notion of coupling comes into play. Informally, we think of the coupling of two reactions as a process in which an endergonic reaction—one which does not "want" to happen—is combined with an exergonic reaction—one that *does* "want" to happen—in a way that improves the products-to-reactants concentrations ratio of the first reaction.

An important example of coupling, and one we will focus on, involves ATP hydrolysis

$$\text{ATP} + \text{H}_2\text{O} \underset{\beta_{\leftarrow}}{\overset{\beta_{\rightarrow}}{\rightleftharpoons}} \text{ADP} + \text{P}_\text{i} + \text{H}^+$$
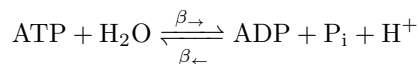
where ATP (adenosine triphosphate) reacts with a water molecule. Typically, this reaction results in ADP (adenosine diphosphate), a phosphate ion P_i and a hydrogen ion H^+. To simplify calculations, we will replace the above equation with

$$\text{ATP} \underset{\beta_{\leftarrow}}{\overset{\beta_{\rightarrow}}{\rightleftharpoons}} \text{ADP} + \text{P}_\text{i}$$
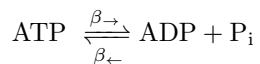
since suppressing the bookkeeping of hydrogen and oxygen atoms in this manner will not affect our main points.

One reason ATP hydrolysis is good for coupling is that this reaction is strongly exergonic:

$$\beta_{\rightarrow} \gg \beta_{\leftarrow}$$

and in fact so much that

$$\frac{\beta_{\rightarrow}}{\beta_{\leftarrow}} \gg \frac{\alpha_{\leftarrow}}{\alpha_{\rightarrow}} \tag{6}$$

Yet this fact alone is insufficient to explain coupling. To see why, suppose our system consists merely of the two reactions

$$\text{X} + \text{Y} \underset{\alpha_{\leftarrow}}{\overset{\alpha_{\rightarrow}}{\rightleftharpoons}} \text{XY} \tag{7}$$

$$\text{ATP} \underset{\beta_{\leftarrow}}{\overset{\beta_{\rightarrow}}{\rightleftharpoons}} \text{ADP} + \text{P}_\text{i} \tag{8}$$

happening in parallel. We can study the concentrations in equilibrium to see that one reaction has no influence on the other. Indeed, the rate equation for

this reaction network is

$$[\dot{X}] = -\alpha_\rightarrow[X][Y] + \alpha_\leftarrow[XY]$$

$$[\dot{Y}] = -\alpha_\rightarrow[X][Y] + \alpha_\leftarrow[XY]$$

$$[\dot{XY}] = \alpha_\rightarrow[X][Y] - \alpha_\leftarrow[XY]$$

$$[\dot{ATP}] = -\beta_\rightarrow[ATP] + \beta_\leftarrow[ADP][P_i]$$

$$[\dot{ADP}] = \beta_\rightarrow[ATP] - \beta_\leftarrow[ADP][P_i]$$

$$[\dot{P_i}] = \beta_\rightarrow[ATP] - \beta_\leftarrow[ADP][P_i]$$

When concentrations are constant, these are equivalent to the relations

$$\frac{[XY]}{[X][Y]} = \frac{\alpha_\rightarrow}{\alpha_\leftarrow} \quad \text{and} \quad \frac{[ADP][P_i]}{[ATP]} = \frac{\beta_\rightarrow}{\beta_\leftarrow}$$
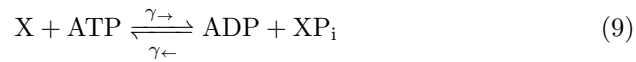
We thus see that ATP hydrolysis is in no way affecting the ratio of [XY] to [X][Y]. Intuitively, there is no coupling because the two reactions proceed independently. This "independence" is clearly visible if we draw the reaction network as a so-called Petri net [1]:
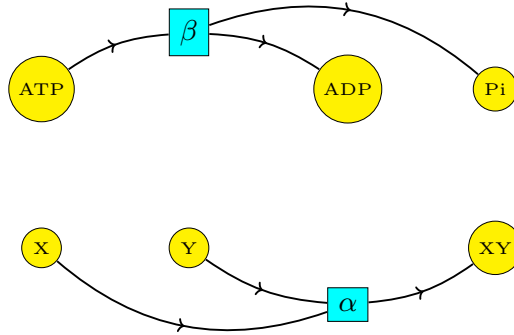

So what really happens when we are in the presence of coupling?

4   Interactions

For coupling to occur, the reactant species in both reactions must interact in some way. Indeed, in real-world examples that follow the above schema and involve coupling, it is observed that, aside from the reactions represented in equations (7) and (8), two other reactions (and their reverses) take place:

$$X + ATP \underset{\gamma_\leftarrow}{\overset{\gamma_\rightarrow}{\rightleftharpoons}} ADP + XP_i \tag{9}$$

$$XP_i + Y \underset{\delta_\leftarrow}{\overset{\delta_\rightarrow}{\rightleftharpoons}} XY + P_i \tag{10}$$

We can picture all four reactions (7)–(10) in a single Petri net as follows:
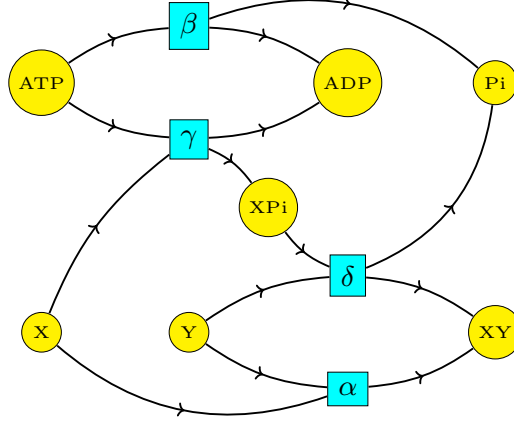
Taking into account this more complicated set of reactions, which are interacting with each other, is *still not enough* to explain the phenomenon of coupling. To see this, consider the rate equation for the system comprised of all four reactions:

$$[\dot{X}] = -\alpha_\rightarrow[X][Y] + \alpha_\leftarrow[XY] - \gamma_\rightarrow[X][ATP] + \gamma_\leftarrow[ADP][XP_i]$$
$$[\dot{Y}] = -\alpha_\rightarrow[X][Y] + \alpha_\leftarrow[XY] - \delta_\rightarrow[XP_i][Y] + \delta_\leftarrow[XY][P_i]$$
$$[\dot{XY}] = \alpha_\rightarrow[X][Y] - \alpha_\leftarrow[XY] + \delta_\rightarrow[XP_i][Y] - \delta_\leftarrow[XY][P_i]$$
$$[\dot{ATP}] = -\beta_\rightarrow[ATP] + \beta_\leftarrow[ADP][P_i] - \gamma_\rightarrow[X][ATP] + \gamma_\leftarrow[ADP][XP_i]$$
$$[\dot{ADP}] = \beta_\rightarrow[ATP] - \beta_\leftarrow[ADP][P_i] + \gamma_\rightarrow[X][ATP] - \gamma_\leftarrow[ADP][XP_i]$$
$$[\dot{P}_i] = \beta_\rightarrow[ATP] - \beta_\leftarrow[ADP][P_i] + \delta_\rightarrow[XP_i][Y] - \delta_\leftarrow[XY][P_i]$$
$$[\dot{XP}_i] = \gamma_\rightarrow[X][ATP] - \gamma_\leftarrow[ADP][XP_i] - \delta_\rightarrow[XP_i][Y] + \delta_\leftarrow[XY][P_i]$$

Introducing the *reaction velocities*

$$J_\alpha = \alpha_\rightarrow[X][Y] - \alpha_\leftarrow[XY]$$
$$J_\beta = \beta_\rightarrow[ATP] - \beta_\leftarrow[ADP][P_i]$$
$$J_\gamma = \gamma_\rightarrow[ATP][X] - \gamma_\leftarrow[ADP][XP_i]$$
$$J_\delta = \delta_\rightarrow[XP_i][Y] - \delta_\leftarrow[XY][P_i]$$

we can write the rate equation as

$$[\dot{X}] = -J_\alpha - J_\gamma$$
$$[\dot{Y}] = -J_\alpha - J_\delta$$
$$[\dot{XY}] = J_\alpha + J_\delta$$
$$[\dot{ATP}] = -J_\beta - J_\gamma$$
$$[\dot{ADP}] = J_\beta + J_\gamma$$
$$[\dot{P}_i] = J_\beta + J_\delta$$
$$[\dot{XP}_i] = J_\gamma - J_\delta$$

In a steady state, all these time derivatives are zero, so

$$J_\alpha = J_\beta = -J_\gamma = -J_\delta$$

Furthermore, in a detailed balanced equilibrium every reaction occurs at the same rate as its reverse reaction, so all four reaction velocities vanish. This implies the relations

$$\frac{[\text{XY}]}{[\text{X}][\text{Y}]} = \frac{\alpha_\rightarrow}{\alpha_\leftarrow}, \quad \frac{[\text{ADP}][\text{P}_\text{i}]}{[\text{ATP}]} = \frac{\beta_\rightarrow}{\beta_\leftarrow}, \quad \frac{[\text{ADP}][\text{XP}_\text{i}]}{[\text{ATP}][X]} = \frac{\gamma_\rightarrow}{\gamma_\leftarrow}, \quad \frac{[\text{XY}][\text{P}_\text{i}]}{[\text{XP}_\text{i}][Y]} = \frac{\delta_\rightarrow}{\delta_\leftarrow}.$$

Thus, even when the reactants interact to form the new species XP_i, *there can be no coupling if the whole system is in equilibrium*, since then the ratio $[\text{XY}]/[\text{X}][\text{Y}]$ is forced to be $\alpha_\rightarrow/\alpha_\leftarrow$. Coupling can only arise out of equilibrium. But how, precisely, does coupling occur?
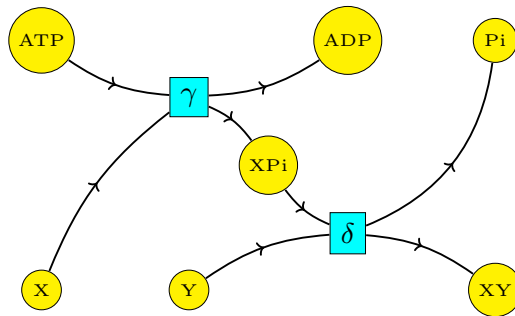
5 Quasiequilibria

Coupling is achieved through the action of enzymes. An enzyme can increase the rate constant of a reaction. However, it cannot change the ratio of forward to reverse rate constants, since that is fixed by the difference of free energies as in Equation (5), and the presence of an enzyme does not change this. Indeed, if an enzyme *could* change this ratio, there would be no need for coupling! Increasing the ratio $\alpha_\rightarrow/\alpha_\leftarrow$ in the reaction

$$\text{X} + \text{Y} \underset{\alpha_\leftarrow}{\overset{\alpha_\rightarrow}{\rightleftharpoons}} \text{XY}$$
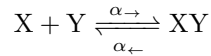
would favor the formation of XY, as desired. But this option is not available.

Instead, the cell uses enyzmes to greatly increase the rate constants for (9) and (10) while leaving those for (7) and (8) essentially unchanged. In this situation we can ignore reactions (7) and (8) and still have a good approximate description of the dynamics, at least for sufficiently short time scales. Thus, we shall study *quasiequilibria*, namely steady states of the rate equation for reactions (9) and (10) but not (7) and (8). In this approximation, the relevant Petri net becomes this:

Now it is impossible for ATP to turn into $\text{ADP} + \text{P}_\text{i}$ without $\text{X} + \text{Y}$ also turning into XY. As we shall see, this is the key to coupling. In quasiequilibrium states, we shall find a nontrivial relation between the ratios $[\text{XY}]/[\text{X}][\text{Y}]$ and $[\text{ATP}]/[\text{ADP}][\text{P}_\text{i}]$.

Of course, this is just part of the full story. Over longer time scales, reactions (7) and (8) become important. They would drive the system toward a true equilibrium, and destroy coupling, if there were not an inflow of the reactants ATP, X and Y and an outflow of the products P_i and XY. To take these inflows and outflows into account, we need the theory of 'open' reaction networks [1, 3, 6].

However, this is beyond our scope here. We only consider reactions (9) and (10), which give the following rate equation:

$$[\dot{X}] = -\gamma_\rightarrow[X][ATP] + \gamma_\leftarrow[ADP][XP_i]$$
$$[\dot{Y}] = -\delta_\rightarrow[XP_i][Y] + \delta_\leftarrow[XY][P_i]$$
$$[\dot{XY}] = \delta_\rightarrow[XP_i][Y] - \delta_\leftarrow[XY][P_i]$$
$$[\dot{ATP}] = -\gamma_\rightarrow[X][ATP] + \gamma_\leftarrow[ADP][XP_i]$$
$$[\dot{ADP}] = \gamma_\rightarrow[X][ATP] - \gamma_\leftarrow[ADP][XP_i]$$
$$[\dot{P_i}] = \delta_\rightarrow[XP_i][Y] - \delta_\leftarrow[XY][P_i]$$
$$[\dot{XP_i}] = \gamma_\rightarrow[X][ATP] - \gamma_\leftarrow[ADP][XP_i] - \delta_\rightarrow[XP_i][Y] + \delta_\leftarrow[XY][P_i].$$

Quasiequilibria occur when all these time derivatives vanish. This happens when

$$\gamma_\rightarrow[X][ATP] = \gamma_\leftarrow[ADP][XP_i]$$
$$\delta_\rightarrow[XP_i][Y] = \delta_\leftarrow[XY][P_i].$$

This pair of equations is equivalent to

$$\frac{\gamma_\rightarrow}{\gamma_\leftarrow}\frac{[X][ATP]}{[ADP]} = [XP_i] = \frac{\delta_\leftarrow}{\delta_\rightarrow}\frac{[XY][P_i]}{[Y]}$$

and it implies

$$\frac{[XY]}{[X][Y]} = \frac{\gamma_\rightarrow}{\gamma_\leftarrow}\frac{\delta_\rightarrow}{\delta_\leftarrow}\frac{[ATP]}{[ADP][P_i]}$$

If we forget about the species XP_i (whose presence is crucial for the coupling to happen, but whose concentration we do not care about), the quasiequilibrium does not impose any conditions other than the above relation. We conclude that, under these circumstances and assuming we can increase the ratio of [ATP] to $[ADP][P_i]$, it is possible to increase the ratio of [XY] to [X][Y].

We can say a bit more, since we can express the ratios of forward and reverse rate constants in terms of exponentials of free energy differences using equation (5). Reactions (7) and (8), taken together, convert X + Y + ATP to XY + ADP + Pi. So do reactions (9) and (10) taken together. Thus, these two pairs of reactions involve the same total change in free energy, so

$$\frac{\alpha_\rightarrow}{\alpha_\leftarrow}\frac{\beta_\rightarrow}{\beta_\leftarrow} = \frac{\gamma_\rightarrow}{\gamma_\leftarrow}\frac{\delta_\rightarrow}{\delta_\leftarrow}. \tag{11}$$

Together with (6), this implies

$$\frac{\gamma_\rightarrow}{\gamma_\leftarrow}\frac{\delta_\rightarrow}{\delta_\leftarrow} \gg 1$$

Thus,

$$\frac{[XY]}{[X][Y]} \gg \frac{[ATP]}{[ADP][P_i]}.$$

This is why coupling to ATP hydrolysis is so good at driving the synthesis of XY from X and Y. Ultimately, this inequality arises from the fact that the *decrease* in free energy for the reaction $ATP \to ADP + P_i$ greatly exceeds the *increase* in free energy for the reaction $X + Y \to XY$. But this fact can only be "put to use" in the presence of coupling; that is, when we observe interactions as described in section 4 that are in a quasiequilibrium as explained in this present section.

6   Emergent conservation laws

The rate equations for reactions (9–10) have the following *conserved quantities*; in other words, the following quantities are constant in time.

1. $[X] + [XP_i] + [XY]$, due to the conservation of X,

2. $[Y] + [XY]$, due to the conservation of Y,

3. $3[ATP] + [XP_i] + [P_i] + 2[ADP]$, due to the conservation of phosphorus,

4. $[ATP] + [ADP]$, due to the conservation of adenosine.

In fact, these conserved quantities were already present in the larger system involving reactions (7-10). Moreover, these quantities, and their linear combinations, are the *only* conserved quantities for the larger system.

To see this, we use some standard ideas from reaction network theory [4, 5]. Consider the 7-dimensional space with orthonormal basis ATP, ADP, P_i, XP_i, X, Y, and XY. We can think of complexes as vectors in this space. An arbitrary choice of the concentrations of all species also defines a vector in this space. Furthermore, any reaction involving these species defines a vector in this space, namely the sum of the products minus the sum of the reactants. This is called the *reaction vector* of this reaction. The four reactions (7-10) give these reaction vectors:

$$v_\alpha = XY - X - Y$$
$$v_\beta = P_i + ADP - ATP$$
$$v_\gamma = XP_i + ADP - ATP - X$$
$$v_\delta = XY + P_i - XP_i - Y$$

Any change in concentrations caused by these reactions must lie in the *stoichiometric subspace*: that is, the space spanned by the reaction vectors. Since these vectors obey one nontrivial relation, $v_\alpha + v_\beta = v_\gamma + v_\delta$, the stochiometric subspace is 3-dimensional. Therefore, the space of conserved quantities must be 4-dimensional, since these specify the constraints on allowed changes in concentrations.

In contrast to this situation, if we consider only reactions (9) and (10), the stoichiometric subspace is 2-dimensional, since v_γ and v_δ are linearly independent. Thus, the space of conserved quantities becomes 5-dimensional. Indeed, we can find an additional conserved quantity:
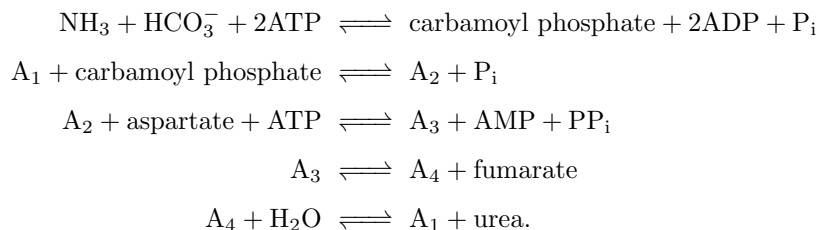
5. $[Y] + [P_i]$

that is linearly independent from the four conserved quantities we had before. It does not derive from the conservation of a particular molecular component. In other words, conservation of this quantity is not a fundamental law of chemistry. Instead, it is an *emergent* conservation law that arises only in situations where the rate constants of reactions (9) and (10) are so much larger than those of (7) and (8) that we can ignore the latter two reactions.

This emergent conservation law captures the phenomenon of coupling. The only way for ATP to form ADP + P_i without violating this law is for Y to be consumed in the same amount as P_i is created, thus forming the product XY.

7 Another example

ATP hydrolysis is a simple example of coupling through emergent conservation laws, but the phenomenon is more general. A slightly more complicated example is the urea cycle. The first metabolic cycle to be discovered, it is used by land-dwelling vertebrates to convert ammonia, which is highly toxic, to urea for excretion.

The urea cycle consists of these reactions:

$$NH_3 + HCO_3^- + 2ATP \rightleftharpoons \text{carbamoyl phosphate} + 2ADP + P_i$$

$$A_1 + \text{carbamoyl phosphate} \rightleftharpoons A_2 + P_i$$

$$A_2 + \text{aspartate} + ATP \rightleftharpoons A_3 + AMP + PP_i$$

$$A_3 \rightleftharpoons A_4 + \text{fumarate}$$

$$A_4 + H_2O \rightleftharpoons A_1 + \text{urea}.$$

Ammonia (NH_3) and carbonate (HCO_3^-) enter in the first reaction, along with ATP. The four remaining reactions form a cycle in which various species $A_1, \ldots, A_4$ "cycle around," each transformed into the next[1]. One atom of nitrogen from carbamoyl phosphate and one from aspartate enter this cycle, and they are incorporated in urea, which leaves the cycle along with fumarate.

All this is powered by two exergonic reactions: the hydrolysis of ATP to ADP and P_i and the hydrolysis of ATP to adenosine monophosphate (AMP) and a compound with two phosphorus atoms, pyrophosphate (PP_i). Thus, we are seeing a more elaborate example of an endergonic process coupled to ATP hydrolysis. The most interesting new feature is the use of a cycle.

Since inflows and outflows are crucial to the purpose of the urea cycle, a full analysis requires treating this cycle as an open chemical reaction network [1, 3, 6]. However, we can gain some insight into coupling just by studying the emergent conservation laws present in this network, ignoring inflows and outflows.

There are a total of 16 species in the urea cycle. There are 5 forward reactions, which are easily seen to have linearly independent reaction vectors. Thus, the stoichiometric subspace has dimension 5. There must therefore be 11 linearly independent conserved quantities.

---

[1]For the curious reader, these species are A_1 = ornithine, A_2 = citrulline, A_3 = argininosuccinate, and A_4 = arginine.

Some of these conserved quantities can be explained by fundamental laws of chemistry. All the species involved are made of five different atoms: carbon, hydrogen, oxygen, nitrogen and phosphorus. The conserved quantity

$$3[\text{ATP}] + 2[\text{ADP}] + [\text{AMP}] + 2[\text{PP}_i] + [\text{P}_i] + [\text{carbamoyl phosphate}]$$

expresses conservation of phosphorus. The conserved quantity

$$[\text{NH}_3] + [\text{carbamoyl phosphate}] + [\text{aspartate}] + 2[\text{urea}] + 2[\text{A}_1] + 3[\text{A}_2] + 4[\text{A}_3] + 4[\text{A}_4]$$

expresses conservation of nitrogen. Conservation of oxygen and carbon give still more complicated conserved quantities. Conservation of hydrogen and conservation of charge are not really valid laws in this context, because all the reactions are occurring in water, where it is easy for protons (H^+) and electrons to come and go. So, four linearly independent "fundamental" conserved quantities are relevant to the urea cycle.

There must therefore be seven other linearly independent conserved quantities that are "emergent": that is, not conserved in every possible reaction, but conserved by those in the urea cycle. A computer calculation shows that we can use these:

1. $[\text{ATP}] + [\text{ADP}] + [\text{AMP}]$, due to conservation of adenosine by all reactions in the urea cycle.

2. $[\text{H}_2\text{O}] + [\text{urea}]$, since the only reaction in the urea cycle involving either H_2O or urea has H_2O as a reactant and urea as a product.

3. $[\text{aspartate}] + [\text{PP}_i]$, since the only reaction involving either aspartate or PP_i has aspartate as a reactant and PP_i as a product.

4. $2[\text{NH}_3] + [\text{ADP}]$, since the only reaction involving either NH_3 or ADP has NH_3 as a reactant and 2ADP as a product

5. $2[\text{HCO}_3^-] + [\text{ADP}]$, since the only reaction involving either HCO_3^- or ADP has HCO_3^- as a reactant and 2ADP as a product.

6. $[\text{A}_3] + [\text{fumarate}] - [\text{PP}_i]$, since these species are involved only in the third and fourth reactions of the urea cycle, and this quantity is conserved in both those reactions.

7. $[\text{A}_1] + [\text{A}_2] + [\text{A}_3] + [\text{A}_4]$, since these species "cycle around" the last four reactions, and they are not involved in the first.

These emergent conservation laws prevent either form of ATP hydrolysis from occurring on its own: the reaction

$$\text{ATP} + \text{H}_2\text{O} \rightleftharpoons \text{ADP} + \text{P}_i$$
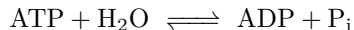
violates conservation of quantities 2, 4 and 5, while

$$\text{ATP} + \text{H}_2\text{O} \rightleftharpoons \text{AMP} + \text{PP}_i$$
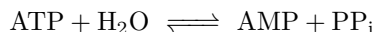
violates conservation of quantities 2, 3 and 6. (In these reactions we are neglecting H^+ ions, since as mentioned these are freely available in aqueous solution.)

Indeed, any linear combination of these two forms of ATP hydrolysis is prohibited. Since this requires only two emergent conservation laws, the presence of seven is a bit of a puzzle. Conserved quantity 3 prevents the destruction of aspartate without the production of an equal amount of PP_i, conserved quantity 4 prevents the destruction of NH_3 without the production of an equal amount of ADP, and so on. But there seems to be more coupling than is strictly "required". Of course, many factors besides coupling are involved in an evolutionarily advantageous reaction network.

Acknowledgements

References

[1] J. Baez and B. Pollard, A compositional framework for reaction networks, *Rev. Math. Phys.* **29** (2017), 1750028.

[2] N. Bazhin, The essence of ATP coupling, *ISRN Biochemistry* **2012** (2012), article 827604.

[3] M. Esposito and M. Polettini, Irreversible thermodynamics of open chemical networks I: emergent cycles and broken conservation laws, *J. Chem. Phys.* **141** (2014), 024117.

[4] M. Feinberg, *Lectures on Chemical Reaction Networks*, 1979.

[5] F. Horn and R. Jackson, General mass action kinetics, *Arch. Ration. Mech. Anal.* **47** (1972), 81–116.

[6] D. Kondepudi and I. Prigogine, *Modern Thermodynamics: From Heat Engines to Dissipative Structures*, Wiley, 2014.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, RIVERSIDE CA, 92521, USA AND CENTRE FOR QUANTUM TECHNOLOGIES, NATIONAL UNIVERSITY OF SINGAPORE, 117543, SINGAPORE

E-mail address: `baez@math.ucr.edu`

INSTITUTE FOR MATHEMATICS, UNIVERSITY OF ZURICH, 8057 ZURICH, SWITZERLAND

E-mail address: `jonathanlorand@protonmail.com`

ENGINEERING RESEARCH ACCELERATOR, CARNEGIE MELLON UNIVERSITY, PITTSBURGH PA, 15213, USA

E-mail address: `blake561@gmail.com`

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA NY, 14853, USA

E-mail address: `mes462@cornell.edu`