Non-steady state mass action dynamics without rate constants: dynamics of coupled reactions using chemical potentials

William R Cannon and Scott E Baker

1 Biological Sciences Division, Pacific Northwest National Laboratory, 902 Battelle Blvd, Richland, WA 99352, United States of America
2 Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, 902 Battelle Blvd, Richland, WA 99352, United States of America
3 Author to whom any correspondence should be addressed

E-mail: William.Cannon@pnnl.gov

Keywords: reaction dynamics, reaction kinetics, reaction affinities, statistical thermodynamics, chemical potentials, maximum entropy

Supplementary material for this article is available online

Abstract

Comprehensive and predictive simulation of coupled reaction networks has long been a goal of biology and other fields. Currently, metabolic network models that utilize enzyme mass action kinetics have predictive power but are limited in scope and application by the fact that the determination of enzyme rate constants is laborious and low throughput. We present a statistical thermodynamic formulation of the law of mass action for coupled reactions at both steady states and non-stationary states. The formulation uses chemical potentials instead of rate constants. When used to model deterministic systems, the method corresponds to a rescaling of the time dependent reactions in such a way that steady states can be reached on the same time scale but with significantly fewer computational steps. The relationships between reaction affinities, free energy changes and generalized detailed balance are central to the discussion. The significance for applications in systems biology are discussed as is the concept and assumption of maximum entropy production rate as a biological principle that links thermodynamics to natural selection.

Introduction

One hundred and fifty years ago Peter Waage and Cato Maximillian Guldberg published their first article describing the law of mass action, that the rate of a chemical reaction is proportional to the concentration of the reacting species \( \gamma_1 \). For a reversible reaction, labeled by \( +\gamma_1 \) for the forward reaction and \( -\gamma_1 \) for the reverse reaction, in which \( \gamma_1 \) moles of chemical \( A \) are transformed into \( \gamma_2 \) moles of chemical \( B \),

\[
\gamma_2 A + \gamma_1 B \rightleftharpoons \gamma_1 A + \gamma_2 B, \quad \text{(Scheme A)}
\]

the forward rate due to reaction \( 1 \) is simply,

\[
\text{forward rate} = k_1 [A]^{\gamma_1}. \quad \text{(1)}
\]

The brackets indicate the concentration, and the constant of proportionality \( k_1 \) is known as the rate constant. A similar relation exists for the reverse reaction \( -1 \). The net rate is given by,

\[
\text{net rate} = k_1 [A]^{\gamma_1} - k_2 [B]^{\gamma_2}. \quad \text{(2)}
\]

All introductory chemistry texts describe the law of mass action in one form or another. Although the relationship is simple and can easily be applied to many reactions, the application to more complex systems such as biological metabolism is challenging because most rate constants are not available and measuring the missing rate constants is very labor intensive.

Typically rate constants are measured by first isolating and purifying the individual enzymes and then determining the rate constants spectrophotometrically in vitro. The process is challenging to implement in a high-throughput manner and in principle, measurements would have to be done for each enzyme in the system of interest because even enzymes having high sequence similarity to the orthologous enzyme in another species can still have widely differing rate parameters. For example, for the enzyme dihydrofolate reductase, the turnover rates for the substrate 7,8-dihydrofolate measured in vitro vary five orders of magnitude across species—from 284 s\(^{-1}\) to less than 1 s\(^{-1}\). Consequently, thousands of rate parameters would be needed for each new genome, and even then the traditional kinetic analysis provides measurements of rates in vitro rather than in vivo. Thus, while a large number of enzyme rate parameters have been determined for a few well-studied model systems such as Escherichia coli and Saccharomyces cerevisiae, this is
not true for most systems of interest. As a result of the difficulty in obtaining rate constants, constraint-based flux models have been the method of choice for large-scale modeling of biological processes such as metabolism. Unfortunately, the lack of physics in regard to the law of mass action limits the ability of these approaches to narrow the solution space to only physically likely solutions. A considerable amount of work has focused on extending constraint-based models to include the physics of mass action in terms of thermodynamics constraints [5], hybrid models [6] or dual variables [7].

Thermodynamic [8–12] and other approaches [13–17] have been proposed for the study of elementary reactions using the law of mass action that do not use rate constants, and instead have used the concept of reaction affinities at steady state. In fact, a clear fundamental understanding of reaction affinities is essential for understanding these methods. IUPAC defines the reaction affinity $A$ as the derivative of the total free energy $G$ with respect to a reaction coordinate, or the extent of the reaction, $\xi$, such that $A = \frac{\delta G}{\delta \xi}$. Nevertheless, there are conflicting definitions of reaction affinity in the literature where it is common to see the reaction affinity defined as $A = -\Delta G_{\text{rxn}}$ [10], where $\Delta G_{\text{rxn}}$ is the free energy change in the system due to a reaction. Other non-standard definitions in the literature directly employ the standard chemical potential $\mu_i^0$ or the concentration-dependent chemical potential $\mu_i$ such that $\frac{\delta G}{\delta \xi} = \Delta G_{\text{rxn}} = \sum M_i \mu_i^{0} \nu_i$ [18] or \[
\left(\frac{\delta G}{\delta \xi}\right)_{T,p} = \Delta G_{\text{rxn}} = \sum M_i \mu_i \nu_i \quad \text{[19],}
\]
where the signed stoichiometric coefficients $\nu_i = -\gamma_i$ for reactants and $\nu_i = +\gamma_i$ for products are used. In these definitions of reaction affinity, the loose use of calculus (a derivative is not a difference) is used to associate reaction affinity with free energy changes. Moreover, adding to the confusion is the statement that $\Delta G_{\text{rxn}} = \sum M_i \mu_i \nu_i$. While this is true for a closed system characterized by constant total number of particles (in the limit of large numbers), many reactive systems of high interest, such as biological systems, are open systems. In an open system, the chemical potential $\mu_i$ is defined as a function of the average count or concentration of the respective chemical species [20]; that is, the chemical potential is a parameter—a constant—rather than a variable. In this case, the value given by $\sum \mu_i \nu_i$ is also a constant; it does not vary with reaction conditions. But the change in the reaction free energy does vary with reaction conditions; it is only approximately constant if a reaction doesn’t change the abundance of the reactants or products significantly, as in the case of the thermodynamic limit of large concentrations at steady state. If this is the case, then the usual kinetic description of the reaction rate can be formulated to include a free energy term. In the case of reaction scheme $A$, \[
et rate \approx k_i[A]^\gamma (1 - e^{-\Delta G_{\text{rxn}}/RT}),
\]
That is, in the limit that the concentrations do not change (i.e. a deterministic steady state), equation (4) is exact, as originally suggested by Temkin [11]—but only for macroscopic and deterministic steady state systems. A linear relationship between rates and $e^{-\Delta G_{\text{rxn}}/RT}$ exists when $\Delta G_{\text{rxn}}$ is highly favorable and in that case the approximation is useful [8, 9]. But this is an extreme case and is not appropriate for the variety of reaction conditions in many systems. To see that this is the case, one only need to consider the initial phase of a reaction when only reactants for the forward reaction are present. In this case, for the forward reaction free energy $\Delta G_{1}$ and a reverse reaction free energy $\Delta G_{1}$, equation (4) results in a finite reverse reaction rate because $\Delta G_{-1} = -\Delta G_{1}$ and $\Delta G_{1}$ has a finite value. One could argue that if the reactant concentrations for the reverse reaction are at zero, then the exponent of the free energy change for the reverse reaction should be zero (i.e. $\Delta G_{-1}$ would have an infinitely large positive value). However, the logic of this argument results in the contradiction that $\Delta G_{-1} = -\Delta G_{1}$ and therefore free energy is no longer a function of state. Even at steady state, when the number of particles of the reaction intermediates drop below ~10, equation (4) will not produce the correct results. A single firing of a reaction in this case can easily result in changes in the abundance of chemical species of 10–100%. This discussion suggests that a microscopic function related to the free energy change is needed that does not require any assumptions about linear relationships between rates and thermodynamic driving forces.

In this report, we first review the relationship between reaction affinity and free energy for isolated reactions using concepts from statistical thermodynamic integration to explicitly demonstrate the relationship between kinetic formulations of rates that use rate constants and a statistical thermodynamics formulation that uses chemical potentials. (Readers intimately familiar with the relationship between reaction affinities, free energies and multinomial statistical models can read ahead to the discussion of coupled reactions following equation (12).) Subsequently, a theorem for coupled reactions is presented based on chemical potentials which can provide relative rates for all sequentially coupled reactions under all conditions—steady state or not—obviating the need for rate constants in many applications. However, if just one rate constant is known for the system of sequentially coupled reactions, then exact rates can be obtained, as well. The only requirement is that steady state levels of reaction intermediates be measurable and available to formulate chemical potentials. Unlike previous thermodynamic formulations, the approach is capable of modeling equilibrium, near-equilibrium and far from equilibrium conditions. We discuss the current work in the context of generalized detailed balance, and show that the work described herein extends the
applicability of generalized detailed balance. The ability to model the dynamics of systems for which accurate rate constants are not available, such as biological systems, would be a considerable advance for predictive modeling in biology. At macroscopic scale, this approach is shown to be the rescaled deterministic rate law based on a rigorous definition of reaction affinities and a coupling constant is introduced that couples the rates of independent reactions. Furthermore, in the case that steady state concentrations are not available, it is shown that a maximum entropy assumption results in both a thermodynamically and kinetically optimal system in which the rate of entropy production is maximized.

Theory

Review of reaction affinity and closed systems

The IUPAC Gold Book defines the reaction affinity as,

$$A = -\frac{dG}{d\xi},$$

where $\xi$ is the extent of a reaction and $G$ is the free energy, usually taken to be the Gibbs energy. We will use instead $G$ to indicate a generalized free energy function that can be specifically defined by ensemble parameters such as total number of particles ($N$), volume ($V$) and temperature ($T$). During an incremental increase in a reaction by $d\xi$, the amount $n_i$ of a substance $i$ changes by $dn_i = \nu_i d\xi$ where $\nu_i$ is the signed stoichiometric coefficient such that $\gamma_i = |\nu_i|$. Over the course of a reaction, an initial amount of substance changes from $n_{i,0}$ to $n_i = n_{i,0} + \Delta n_i$ where $\Delta n_i = \int \nu_i d\xi$. During the course of the reaction, the free energy change is given by,

$$\Delta_{\text{rxn}} G(\alpha, 0) = G(\xi = \alpha) - G(\xi = 0) = \int_{\xi=0}^{\xi=\alpha} dG d\xi = -\int_{\xi=0}^{\xi=\alpha} A d\xi.$$

For the extent of the reaction, we use the accepted criteria for thermodynamic integration [21, 22] such that $\xi$ varies from $\xi_{\text{initial}} = 0$ (no reaction) to $\xi_{\text{final}} = 1$ representing a complete reaction in the sense that a stoichiometric unit of products is produced from a stoichiometric unit of reactants. When the extent of the reaction is greater than one, then more than a stoichiometric amount of material is processed through a reaction. Since the free energy is a state function, the actual path of the integral does not matter.

The reaction affinity as a function of concentrations or counts can be found by formulating the free energy as the multinomial Boltzmann distribution and explicitly taking the derivative of the free energy with respect to the extent of the reaction or reaction coordinate, $\xi$. The multinomial Boltzmann distribution describes how a mixed population of molecules is distributed among $M$ chemical species $i$, each with a count of $n_i$, such that $\sum n_i = N_T$. The species are not equally likely, but are distributed according to the Boltzmann criteria of the molecular partition function $q_i = e^{-\beta_i \mu_{i}^{\text{sys}}/k_B T}$ where $\mu_{i}^{\text{sys}}$ is the free energy of formation of compound $i$ averaged over all solvent configurations and $k_B T$ is the ambient thermal energy ($k_B$ is Boltzmann’s constant and $T$ is temperature). The justification for using the Boltzmann criteria is similar to that used by De Decker \textit{et al}, that the heat released due to chemical reactions is either moderate or averaged over, such that microscopic processes like velocity distributions of individual atoms and molecules are decoupled from the statistics of the counts [23]. The microscopic analogue $G$ of the free energy is then,

$$\frac{-G}{k_B T} = \log \left[ \frac{N_T!}{\prod_{i=1}^{M} \frac{1}{n_i!} q_i^n} \right] = \log (N_T!) - \sum_i \log(n_i!) + \sum_i n_i \log q_i,$$

$$\approx [N_T \cdot \log(N_T) - N_T] - \sum_i [n_i \cdot \log(n_i) - n_i] + \sum_i n_i \log q_i,$$

where Stirling’s approximation is used in equation (6). $G$ is a microscopic analogue of the macroscopic Gibbs energy $G$ because it does not involve averaging the counts $n_i$ over all configurations. (The macroscopic Gibbs energy $G$ is a function of total number of particles ($N_{\text{total}}$), temperature ($T$) and pressure ($P$) and is an average over all microstates at fixed $N_{\text{total}}, T$ and $P$.) Thus, once initial transients are averaged out, the macroscopic free energy is a constant for a closed system.) Taking the derivative of the microscopic free energy $G$ with respect to the extent of the reaction gives,

$$\frac{A}{k_B T} = -\frac{1}{k_B T} \frac{dG}{d\xi} \approx \frac{d}{d\xi} \left\{ \left[ N_T \cdot \log(N_T) - N_T \right] - \sum_i [n_i \cdot \log(n_i) - n_i] + \sum_i n_i \log q_i \right\} = -\sum_i \frac{d}{d\xi}(n_i \cdot \log(n_i) - n_i) + \sum_i \frac{d}{d\xi}(n_i \log q_i).$$

Here, $N_T$ is an ensemble parameter—a constant—but for an open system $dN_T/d\xi$ is small but not generally zero. In the next section we deal with open systems; here, for the sake of simplicity we will assume that $dN_T/d\xi = 0$. Next, we substitute the identity

$$\frac{d}{d\xi} = \frac{d}{dn_i} \frac{dn_i}{d\xi}$$

and subsequently substitute $\frac{dn_i}{d\xi} = \nu_i$,

$$= -\sum_i \frac{d}{dn_i} (n_i \cdot \log(n_i) - n_i) + \sum_i \frac{d}{dn_i} \log q_i$$

$$= -\sum_i \nu_i \cdot (n_i \cdot \log(n_i) - n_i) + \sum_i \nu_i \log q_i$$

$$= -\sum_i \nu_i \cdot \log(n_i) + \sum_i \nu_i \log q_i.$$

In a closed system at constant $N_T$, the quantity $-k_B T \log(q_i/n_i)$ is defined as the chemical potential $\mu_i$. 

W R Cannon and S E Baker
which is unfortunate because this makes the chemical potential a variable rather than a parameter which can lead to confusion when discussing an open system where the chemical potential is strictly a parameter (see below). Likewise, one could argue that the open system parameter $\mu_i$ should be referred to explicitly as an average chemical potential. While these distinctions are clear in discussions of statistical thermodynamics and ensembles, they are less clear in discussions of metabolic modeling. Regardless, the reaction affinity and ensembles, they are less clear in discussions of statistical thermodynamics and ensembles.

Regardless, the reaction affinity can now be written as,

$$\frac{A}{k_B T} \approx - \sum_i n_i \mu_i,$$

where the approximation is again due to the use of Stirling’s approximation in equation (7). To avoid later confusion with the chemical potential in an open system, we will express the reaction affinity in terms of the equilibrium constant and counts (or similarly concentrations). Taking the exponential of the reaction affinity and writing the counts $n_i$ as an explicit function of the extent of the reaction at the initial condition $\xi = 0$,

$$e^{\Delta A(\xi = 0)/k_B T} \approx \prod_{i}^{M} n_i(\xi = 0)^{\gamma_i} \text{ products} \prod_{i}^{M} q_i^{\gamma_i} \text{ reactants}$$

$$= \prod_{i}^{M} n_i(\xi = 0)^{\gamma_i} \text{ products} \prod_{i}^{M} q_i^{\gamma_i} \text{ reactants}$$

$$= \prod_{i}^{M} n_i(\xi = 0)^{\gamma_i} \text{ products} \prod_{i}^{M} q_i^{\gamma_i} \text{ reactants}$$

$$= \prod_{i}^{M} n_i(\xi = 0)^{\gamma_i} \text{ products} \prod_{i}^{M} q_i^{\gamma_i} \text{ reactants}$$

$$= \prod_{i}^{M} n_i(\xi = 0)^{\gamma_i} \text{ products} \prod_{i}^{M} q_i^{\gamma_i} \text{ reactants}$$

$$= \prod_{i}^{M} n_i(\xi = 0)^{\gamma_i} \text{ products} \prod_{i}^{M} q_i^{\gamma_i} \text{ reactants}$$

The constant $K_3$ is the equilibrium constant for reaction 1, defined as the ratio of the molecular partition functions $q_i$ of the species $i$ involved in the reaction, each raised to the power of its stoichiometric coefficient in the reaction [24]. It is important to note that the partition functions $q_0, \ldots, q_M$ are the exponents of standard chemical potentials $\mu_0, \ldots, \mu_M$ (scaled by the available thermal energy), which in this case are averaged over all lower degrees of freedom—solvation, translational, rotational, vibrational and electronic.

In the case of reaction scheme A, the reaction affinity becomes,

$$e^{\Delta A(\xi = 0)|_{/h} T} \approx K_3^{\gamma_i} n_i(\xi = 0)^{\gamma_i}$$

In contrast to the reaction affinity, the microscopic analogue of the free energy change is the log odds of the densities of two different states, which for an reactant state at $\xi = 0$ and a product state at $\xi = 1$ its exponential is [25],

$$e^{-\Delta G(\xi = 1, \xi = 0)|_{/h} T} \approx K_3^{\gamma_i} n_i(\xi = 0)^{\gamma_i}$$
together that transform initial reactants into final products. Consider the simple coupled reaction system,

\[
\begin{align*}
\gamma_{A;1} A &\rightarrow \gamma_{B;1} B \\
\gamma_{B;2} B &\rightarrow \gamma_{C;2} C
\end{align*}
\]  
(Scheme B)

where \(A, B\) and \(C\) are chemical species and \(\gamma_{ij}\) is the stoichiometric coefficient for species \(i\) of reaction \(j\).

For an open system, the microscopic free energy is given by a multinomial Boltzmann distribution similar to equation (6) but without any restriction on the total number of particles. In this case, the derivative of the total number of particles in equation (7) must be included in the integral of equations (7) and (10). If the total number of particles before the reaction \((\xi = 0)\) is \(N_0^n\) and after the reaction \((\xi = 1)\) is \(N_0^n + \sum \nu_i\) then the respective reaction affinity and free energy change are,

\[
\begin{align*}
\frac{\Delta A}{k_B T} &= \frac{d}{d\xi} \log(N_0^n) - \sum_i \frac{d}{d\xi} \log(n_i) + \sum \nu_i \log q_i \\
\frac{\Delta \Xi}{k_B T} &= -\log\left(\frac{N_0^n + \sum \nu_i}{N_0^n}\right) + \sum \log\left(\frac{(n_i^0 + \nu_i)1}{n_i^01}\right) \\
&= \sum \nu_i \log q_i.
\end{align*}
\]

Here the symbol \(\Xi\) is used to represent the microscopic free energy of an open system and \(\Delta \Xi\) to indicate a change in the microscopic free energy with dependence on the total number of particles. In the equation for the free energy change, \(\log\left(\frac{N_0^n + \sum \nu_i}{N_0^n}\right)\) is negligible for systems with a large number of particles undergoing a single reaction. For the rest of the discussion, we will assume that the total number of particles does not change simply for convenience and clarity, and will consequently use formulations analogous to equations (8) and (9) instead.

In order to model coupled reactions in a manner analogous to the use of reaction affinities for isolated reactions, a function is needed for the odds of one reaction relative to the next reaction. That is, for reaction scheme B information is needed on the relative rate to equation (6) but without any restriction on the system-dependent and have no system-independent standard value. Next, we demonstrate how non-equilibrium steady state measurements can be used to determine coupling constants and how sequentially-coupled reactions can be modeled using chemical potentials instead of rate constants. In all cases discussed below, the extent of each reaction \(i\) is \(\xi_i = 0\) and so we will drop the explicit notation for the extent of the reaction.

For reactions shown in scheme B, the deterministic time dependence of each chemical species is governed by the usual set of ordinary differential equations. For species \(B\) this time dependence is,

\[
\begin{align*}
\frac{d n_B}{dt} &= \gamma_{b_A} (\mathcal{R}_A n_A^{\gamma_A}) - \gamma_{b_B} n_B^{\gamma_B} \\
&- \gamma_{b_C} (\mathcal{R}_C n_C^{\gamma_C}).
\end{align*}
\]

The deterministic rate equation can be rescaled in time by dividing \((\mathcal{R}_A n_A^{\gamma_A}) = \tau_1\) and replacing the ratios of rate constants with equilibrium constants,

\[
\tau_1 \cdot \frac{d n_B}{dt} = \gamma_{b_A} (1 - K_{-1} n_B^{\gamma_B})
\]

(14)

where \(K_{-1} = \mathcal{R}_C / \mathcal{R}_A\) and \(K_{-2} = \mathcal{R}_B / \mathcal{R}_C\) are equilibrium constants such that the second term in each parenthetical expression corresponds to the reaction affinity for each reverse reaction. The factor \(\mathcal{R}_B n_B^{\gamma_B} / \mathcal{R}_C n_C^{\gamma_C}\) is the value of the desired coupling term \(K_{21} n_B^{\gamma_B} / n_C^{\gamma_C}\).

Unless one can determine the coupling term \(K_{21} n_B^{\gamma_B} / n_C^{\gamma_C}\), equation (15) is not useful for modeling systems of coupled reactions without rate constants. However, this coupling term can be determined by solving equation (15) at steady state. For the reactions in scheme B at steady state and with steady state concentrations denoted by SS,

\[
K_{21} \frac{n_B(\text{SS})^{\gamma_B}}{n_C(\text{SS})^{\gamma_C}} = \gamma_{b_A} \left(1 - K_{-1} \frac{n_B(\text{SS})^{\gamma_B}}{n_B(\text{SS})^{\gamma_C}}\right)
\]

(16)

Equations (15) and (16) tells us that if we measure concentrations of metabolites at a steady state, we can determine the relative dynamics of sequentially-coupled reactions without rate constants, whether at steady state or not. The dynamics are relative to the characteristic time of some base reaction, determined by the choice of \(n = (\mathcal{R}_A n_A^{\gamma_A})^{-1}\).
The parameters that are needed are the steady state concentrations and equilibrium constants $K_{-1}$ and $K_{-2}$ (or equivalently the molecular partition functions $q_{Z_0}, q_{Z_1}$ and $q_{Z_2}$). With these parameters in hand, it is easy to determine the coupling constants for the reactions using chemical potentials,

$$
K_{21} = \frac{n_{A}(SS)\gamma_{1}}{n_{B}(SS)\gamma_{2}} \left[ 1 - \left( \frac{\gamma_{2}}{\gamma_{1}} \right) \left( \frac{n_{A}(SS)\gamma_{1}}{n_{B}(SS)\gamma_{2}} + \frac{\gamma_{1}}{\gamma_{2}} \right) \right] \frac{1}{1 - \left( \frac{\gamma_{2}}{\gamma_{1}} \right) \left( \frac{n_{A}(SS)\gamma_{1}}{n_{B}(SS)\gamma_{2}} + \frac{\gamma_{1}}{\gamma_{2}} \right)}.
$$

The last equality is obtained by substituting for the definition of the chemical potential for an open system. There are many biological systems for which the stoichiometric coefficients are all unity. In this case, the coupling constant has the simple form,

$$
K_{21} = \frac{n_{A}(SS)}{n_{B}(SS)} \left[ 1 - \frac{1}{1 - \left( \frac{\gamma_{2}}{\gamma_{1}} \right) \left( \frac{n_{A}(SS)\gamma_{1}}{n_{B}(SS)\gamma_{2}} + \frac{\gamma_{1}}{\gamma_{2}} \right)} \right].
$$

The rate equation (14) expressed entirely in terms of chemical potentials and in terms of time relative to $\tau_{1}$ is then,

$$
\tau_{1} \frac{dn_{A}}{dt} = \gamma_{1} \left( \frac{1}{1 - \left( \frac{\gamma_{2}}{\gamma_{1}} \right) \left( \frac{n_{A}(SS)\gamma_{1}}{n_{B}(SS)\gamma_{2}} + \frac{\gamma_{1}}{\gamma_{2}} \right)} \right) n_{A}(SS) - \gamma_{1} \frac{n_{B}}{n_{A}} \left( 1 - \frac{1}{1 - \left( \frac{\gamma_{2}}{\gamma_{1}} \right) \left( \frac{n_{A}(SS)\gamma_{1}}{n_{B}(SS)\gamma_{2}} + \frac{\gamma_{1}}{\gamma_{2}} \right)} \right) n_{B}(SS),
$$

where again the extent of the reaction is taken to be $\xi = 0$ for all reactions such that the counts $n_{i}$ are the instantaneous counts for each of the chemical species $i$. The rate equation can be expressed more succinctly in terms of equilibrium and coupling constants,

$$
\tau_{1} \frac{dn_{A}}{dt} = \gamma_{1} \left( 1 - K_{-1} \frac{n_{B}^{A}}{n_{A}^{*}} \right) - \gamma_{1} K_{21} \frac{n_{B}^{A}}{n_{A}^{*}} \frac{1}{1 - K_{-2} \frac{n_{B}^{C}}{n_{B}^{*}}},
$$

where again $K_{-1}$ and $K_{-2}$ are the equilibrium constants for the reverse reactions in scheme B. Equations (16) and (19) are the key concepts of this report. It will be demonstrated in the Results that for sequentially coupled reactions, using these equations with steady state concentrations or chemical potentials in lieu of rate constants enables one to determine the relative time dependence of the chemical species. In the next section, we generalize these equations to an arbitrary chain of sequential reactions.

**General form of coupling**

A set of sequentially coupled reactions can be represented by,

$$
\gamma_{Z_{0}}Z_{0} \xrightarrow{1} \gamma_{Z_{1}}Z_{1}
$$

$$
\gamma_{Z_{1}}Z_{1} \xrightarrow{2} \gamma_{Z_{2}}Z_{2}
$$

$$
\cdots
$$

$$
\gamma_{Z_{v-1}}Z_{v-1} \xrightarrow{v-1} \gamma_{Z_{v}}Z_{v-1}
$$

(Scheme C)

where $Z_{v}$ is the product of the $v$th reaction with a stoichiometric coefficient of $\gamma_{Z_{v}}$. For the stoichiometric coefficient $\gamma_{Z_{v}}$, the first subscript $Z_{v}$ indicates the chemical species and the second subscript $v$ indicates the reaction. Equation (19) can be generalized using the last two reactions of scheme C as a general case. The coupling term for the $v$th reaction can be derived in terms of the previous reaction $(v - 1)$ from the steady state condition, as follows. The relative time dependence of the counts $n_{Z_{v-1}}$ of the reactant of the $v$th reaction, $Z_{v-1}$, is given by,

$$
\frac{1}{K_{v-1}n_{Z_{v-1}}^{2-1}} \frac{dn_{Z_{v-1}}}{dt} = \gamma_{Z_{v-1}} \left( 1 - \frac{\gamma_{v-1}n_{Z_{v-1}}^{2-1}}{\gamma_{v-1}n_{Z_{v-1}}^{2-1}} \right) \frac{d}{dt} \left( \frac{\gamma_{v-1}n_{Z_{v-1}}^{2-1}}{\gamma_{v-1}n_{Z_{v-1}}^{2-1}} \right)
$$

where $c_{v,v-1}$ is the coupling term that will be derived. In equation (20) the time dependence is relative to that of the $(v - 1)$th reaction. At steady state $\frac{dn_{Z_{v-1}}}{dt} = 0$ such that, similarly to equation (16),

$$
\frac{SS}{\gamma_{v-1}} \frac{dn_{Z_{v-1}}}{dt} = \left( 1 - \frac{\gamma_{v-1}n_{Z_{v-1}}^{2-1}}{\gamma_{v-1}n_{Z_{v-1}}^{2-1}} \right) \frac{d}{dt} \left( \frac{\gamma_{v-1}n_{Z_{v-1}}^{2-1}}{\gamma_{v-1}n_{Z_{v-1}}^{2-1}} \right)
$$

where SS again indicates that the values of $n_{Z_{v-1}}, n_{Z_{v-1}}^{2-1}$ and $n_{Z_{v-1}}^{2-1}$ were obtained at steady state. The numerical value of $c_{v,v-1}$ at any state $I$ is easily determined,

$$
c_{v,v-1}(I) = \frac{SS}{c_{v,v-1}} \frac{n_{Z_{v-1}}^{2-1}}{n_{Z_{v-1}}^{2-1}} \frac{n_{Z_{v-1}}^{2-1}}{n_{Z_{v-1}}^{2-1}} = \frac{K_{v-1}n_{Z_{v-1}}^{2-1}}{n_{Z_{v-1}}^{2-1}} \frac{n_{Z_{v-1}}^{2-1}}{n_{Z_{v-1}}^{2-1}} = K_{v-1} \frac{n_{Z_{v-1}}^{2-1}}{n_{Z_{v-1}}^{2-1}}.
$$

(21)
The sequential nature of the reactions suggests an iterative process, in which the coupling between the first reaction in the sequence and the $i$th reaction is given by,

$$c_{i,1} = \prod_{j=1}^{i-1} c_{i,j-1},$$

where $c_{1,0} \equiv 1$. Consequently, the change in the counts of the product of the $(\nu - 1)$th reaction, $n_{Z_{\nu-1}}^p$ when rescaled by the rate of the first reaction in the sequence is,

$$\frac{1}{k_0 n_{Z_0}^p} \frac{dn_{Z_{\nu-1}}}{dt} = \left(1 - K_{\nu-1} n_{Z_{\nu-1}}^p n_{Z_{\nu-2}}^{p-1} \prod_{i=1}^{\nu-1} c_{i,i-1} \right)$$

Using this formulation, all reactions in a reaction network can be rescaled by the transition probability of any specific reaction. In principle, then knowledge of any one rate constant can be used to derive the absolute time-dependence of the entire system.

**Results**

**Simulations of mass action using coupling constants in lieu of rate constants**

Like equilibrium constants, coupling constants between reactions are the ratio of the rate constants of the respective elementary reactions. Unlike equilibrium constants, however, coupling constants between reactions such as $K_{21}$ are system dependent just as are rate constants and chemical potentials in non-equilibrium steady state systems. For convenience, we refer to equations (18) and (19) as differential equations using coupled reaction theorem. For the reactions of scheme B, figure 1 shows the comparison between a simulation solving the stochastic equivalent of equation (14) and one solving for stochastic equivalent of equation (19). Both simulations used the Gillespie stochastic simulation method [26]. In the kinetic stochastic simulation, reaction/transition probabilities were determined using rate constants and counts, while in the stochastic thermodynamic simulation reaction reaction probabilities were determined using chemical potentials and counts of species. Reactions were selected randomly based on the formulation of transition probabilities shown in table 1. In these simulations, the absolute time is not modeled—all results pertain to the characteristic time of the $+1$ reaction in scheme B. However, if just one rate constant is available then the coupled reaction theorem can be used to reproduce the exact same time-dependent trajectory as the trajectory governed by the (un-scaled) ordinary differential equation (14).

The top row in figure 1 compares the steady state trajectories of the reaction intermediate $B$ from stochastic simulations using the coupled reaction theorem (blue) with trajectories from stochastic kinetic simulations (red). Simulations were carried out under a range of conditions, from mildly non-equilibrium conditions with a total free energy change from initial reactants to final products of $-2$ kcal mol$^{-1}$, to strongly non-equilibrium conditions with an overall free energy change of $-20$ kcal mol$^{-1}$. The steady state solutions of the deterministic (real valued) ordinary differential equation are shown as dotted lines for comparison. When using the same set of random numbers, the trajectories from the coupled reaction theorem simulations are exactly the same as that for the stochastic kinetic simulations (the trajectories are offset by $+5$ counts for clarity), indicating that fluctuations around the steady state are identical. We can also demonstrate that the reaction probabilities are proportional, which is to be expected from the mathematical equivalence of equations (15) and (19).

The use of the coupled reaction theorem is valid away from the steady state, as well. Shown in row 2 in figure 1 are the transient decays from non-equilibrium states to the steady states obtained by setting the initial counts of the intermediate species $B$ at a multiple of the steady state count. The transient is then the decay of the count of the intermediate species $B$ to its steady state level. As expected, both the stochastic kinetic simulations and the coupled reaction theorem simulations produce exactly the same trajectory when using the same set of random numbers. In the case of rows 1 and 2 in figure 1, the forward rate constants for the two reactions differ in scale by four orders of magnitude ($R_1/R_2 = 10^{-4}$).

The multi-scale nature of the approach is demonstrated by holding $R_1/R_2 = K_1$ and $R_1/R_2 = K_2$ constant while varying $R_1/R_2$ over a large range. Row 3 in figure 1 shows the average steady state concentrations calculated for both coupled reaction theorem and stochastic kinetic simulations when $R_1/R_2$ is varied over eight orders of magnitude ($10^{-4}$ to $10^{5}$) in 100000 different simulations, each represented as a point in the plot. As the ratio $R_1/R_2$ increases, the concentration of the reaction intermediate increase as the intermediate is produced faster than it can be taken away until a steady state is reached. Once again, both methods produce exactly the same average steady state concentration.

Steady state concentrations, however, are less sensitive to variations in the rate constants than the rate of material flow through the system of reactions. The net flow of material through reaction 1 in scheme B during a simulation with $n_{\text{steps}}$ timesteps is given by,
\[ \pi_\pi^n = - f_\pi^n,1,\text{net} \]

where \( n_{\text{step}} \) is the total number of reaction steps and the reaction probabilities \( \pi_\pi^n \) and \( \pi_{\pi-1} \) are given in Table 1. At steady state, the net flow of material through each reaction is equal such that \( f_{\pi,\text{net}} = f_{\pi-1,\text{net}} \). The net material flow through either reaction during the simulation is plotted as a function of the ratio of the rate constants, \( \mathcal{R}_1/\mathcal{R}_2 \), in the bottom row in Figure 1. There is a many-to-one relationship between the ratio \( \mathcal{R}_1/\mathcal{R}_2 \) and the overall material flow because each value of \( \mathcal{R}_1/\mathcal{R}_2 \) can correspond to multiple combinations of rate constants.

Figure 1. (A)–(D) Comparisons between stochastic kinetic simulations (red) and simulations using the coupled reaction theorem (blue) as described in the text in which the same set of random numbers were used in both sets of simulations. Simulations of the coupled reactions of scheme B with all stoichiometric coefficients equal to unity were carried out at different driving forces by fixing \( C \) and setting the boundary concentrations of \( A \) to an appropriate level. In each case, the equilibrium constant for reaction 1 (\( K_1 \)) is 25-fold greater than that for reaction 2 (\( K_2 \)). In both (A) and (B) the coupled reaction theorem trajectory is offset by +5 counts so that it can be distinguished. (A) Steady state trajectory from representative simulations using the same set of random numbers. (B) Non-equilibrium transient trajectories from representative simulations using the same set of random numbers. (C) Steady state counts of the intermediate \( B \) from 100,000 simulations of both stochastic kinetic and stochastic thermodynamic simulations in which the ratio of the rate constants \( \mathcal{R}_1/\mathcal{R}_2 \) in the kinetic simulation is varied from \( 10^{-4} \) to \( 10^4 \) while keeping \( \mathcal{R}_1/\mathcal{R}_2 = K_1 \) and \( \mathcal{R}_1/\mathcal{R}_2 = K_2 \). In the stochastic thermodynamic simulation, the parameters were varied using different steady state counts instead of varying rate constants. (D) Steady state material flow over the same set of simulations as in (C).

Table 1. Transition probabilities used in stochastic simulations. For the simulations based on rate constants, the denominator

\[
\begin{align*}
\mathcal{D} &= \mathcal{R}_A \pi_A + \mathcal{R}_B \pi_B + \mathcal{R}_C \pi_C
\end{align*}
\]

while for the simulations based on the coupled reaction theorem

\[
\begin{align*}
\mathcal{D} &= 1 + K_{-1} \cdot \frac{t_1}{s_1} + K_{21} \cdot \frac{t_2}{s_2} + K_{22} \cdot \frac{t_2}{s_2} - \frac{t_1}{s_1}.
\end{align*}
\]

| Method | Transition probability |
|--------|------------------------|
| Kinetic | \( \pi_1 \) | \( \pi_{\pi-1} \) | \( \pi_2 \) | \( \pi_{\pi-2} \) |
| Coupled reaction theorem | \( \frac{\mathcal{R}_A}{\mathcal{D}} \cdot t_1 \cdot K_{-1} \cdot \frac{t_1}{s_1} \) | \( \frac{\mathcal{R}_B}{\mathcal{D}} \cdot t_2 \cdot K_{21} \cdot \frac{t_2}{s_2} \) | \( \frac{\mathcal{R}_C}{\mathcal{D}} \cdot \frac{t_2}{s_2} \cdot K_{-2} \cdot \frac{t_2}{s_2} \) | \( \frac{\mathcal{R}_D}{\mathcal{D}} \cdot \frac{t_1}{s_1} \cdot K_{21} \cdot \frac{t_2}{s_2} \) |
of $R_1$ and $R_2$. Once again the coupled reaction theorem and stochastic kinetic simulations produce exactly the same results when the same set of random numbers are used to select which reaction to fire. In fact, the correlation between the trajectories from stochastic kinetic and the coupled reaction theorem simulations across all values of $R_1/R_2$ is 1.0 within the numerical precision of the software.

As mentioned above, if just one rate constant is available then coupled reaction theorem simulations can reproduce the exact time-dependent trajectory as the trajectory governed by the (un-scaled) ordinary differential equation (14). Measurement of $R_a$ allows one to formulate the time dependence of the reaction intermediates such as $B$ as (see equation (19)),

$$
\frac{d n_B}{d t} = \tau_1^{-1} \left[ \gamma_a \left( 1 - K_{n_n}^{n_n} n_B^{n_B} \right) \right. \\
\left. - \gamma_a K_{n_1} n_1^{n_1} n_B^{n_B} \right] - \gamma_a K_{n_2} n_2^{n_2} n_B^{n_B} - \gamma_a K_{n_3} n_n^{n_n} n_B^{n_B} \right].
$$

(23)

Simulations in relative time

More conveniently, simulations can be carried out in relative time using equation (19) and then adjusted post-simulation to absolute time. Furthermore, relative to equations (23) or (14), the use of equation (19) can accelerate the convergence of the non-equilibrium transients according to the time scale of the rescaling reaction. If the dynamic equations are rescaled by the reaction with the fastest dynamics or highest probability, then the values of the respective time derivatives for the set of differential equations governing the dynamics of the system is reduced accordingly. As a consequence, simulations may converge to the steady state in a fewer number of steps. Regardless, the simulated time-to-steady state will be equivalent when time rescaling is used in this manner since reaching the steady state depends on the time scale of the reaction with the lowest probability or slowest dynamics.

To illustrate this time dependence, consider a pair of reactions similar to those in scheme B but catalyzed by a pair of enzymes, as shown in figure 2. Enzyme 1 ($E_1$) binds substrate $A$ and produces product $B$, which is released and in turn bound by enzyme $E_2$ to produce product $C$. Starting from a highly non-equilibrium state consisting of only substrate $A$ and enzymes $E_1$ and $E_2$, initially the fastest reaction will be the binding of $A$ to $E_1$ to form the complex $E_1A$. The slowest reaction will be that producing the final product, $C$. Using a kinetic simulation based on rate constants and coupled reaction theorem simulation based on chemical potentials and rescaling, the dynamics of the unscaled and time-scaled system over the course window are shown in figure 3(A) for the first enzymatic reaction and in figure 3(B) for the second enzymatic reaction. The differential equations used to model the reactions are analogous to equations (19) and (14), and are provided in the supplementary information along with all of the details of the simulation in the form of a browser-based computational notebook. A stiff, adaptive solver using was used to solve the differential equations.

Although both simulations start from the same set of concentrations, the simulation using the coupled reaction theorem and rescaling converges to the steady state more rapidly under all reaction conditions that were tested. The kinetic simulation spends a considerable amount of time simulating the fast dynamics that occur below $10^{-4}$ s (figures 3(A) and (B)), while the simulation using the coupled reaction theorem effectively accelerates the convergence of the fast dynamics and spends relatively little time below the millisecond time scale. Regardless, the steady state is reached in a relative time scale according to the time scale of the reaction simulating the coupled reaction theorem simulation at approximately the same cumulative time point as the kinetic simulation because the dynamics of the slowest reactions are unchanged in the vicinity of the steady state.

That the rescaled ODEs resulting from implementing the coupled reaction theorem require less computational work to reach steady state solution is implied by figure 3(C) which shows that the rescaled ODE’s require fewer number of time steps to reach the steady state than the full ODE using rate parameters. Specifically how much computational work is reduced and under what conditions will be a topic of future reports where it will be tested on a sufficiently complex system of equations.

Course graining using summary reactions

In addition to being able to produce the correct dynamics without the need for rate constants, the coupled reaction theorem has an advantage over kinetic formulations of reaction dynamics for multi-scale modeling of the steady state: course-graining the dynamics such that the equations are ‘telescopic’ is relatively easy. One can zoom in or out of the details of the reaction system as needed, which can be a considerable advantage for modeling. If one is only interested in steady state phenomena, it is not necessary to model elementary reactions; summary reactions can be modeled using chemical potentials instead and still produce the correct steady state dynamics. The course-grained dynamics will be the composite dynamics of the collapsed system, however.

To demonstrate the ability to collapse reaction schemes to coarse-grained summary reactions, one only needs to consider a Kirchoff’s loop relationship for cycles [28, 29]. Consider the first reaction cycle shown in figure 2 from reactant $A$ through reactions $2$–$4$ to product $B$ and then through reaction $1$ back to initial reactant $A$. The free energy change for traversing around a cycle is zero. In the deterministic limit (large number of particles) and at steady state, the relation between rate and free energy is $\Delta G = -RT \log(r_i/r_f)$.
Figure 2. Coupled enzyme catalyzed reactions corresponding to the uncatalyzed reactions of scheme B. The uncatalyzed reactions are 1 and 5. In the catalyzed reactions enzyme 1 ($E_1$) binds substrate $A$ in reaction 2, converting the substrate to intermediate $B$ in reaction 3, and releasing $B$ in reaction 4, while enzyme 2 ($E_2$) binds the intermediate $B$ in reaction 6, converts the intermediate to product $C$ in reaction 7, and releases the product in reaction 8.

Figure 3. Time dependence of enzymatic reaction 1 (A) and reaction 2 (B) using rate parameters (labeled ‘kinetic’), the coupled reaction theorem with time-rescaling (labeled ‘CRT with time rescaling’) and the coupled reaction theorem formulation of summary reactions that do not include catalyst dynamics (labeled ‘CRT Summary Rxn’). The latter can be thought of as the statistical course graining of the reaction dynamics of the detailed system in which the summary reaction represents the course grained system. In both coupled reaction theorem simulations the processes occurring on small time scales (fast dynamics) converge much faster than in the simulation based on rate parameters. (C) Plot of the time scale of the dynamics as a function of the simulation step. The time rescaling results in larger time steps taken per simulation step for both CRT simulations. (D) The ratio of the rates for each of the reactions in figure 2. Reactions 2–4 and 6–8 involve enzyme dynamics while reactions 1 and 4 are summary reactions representing the overall process. At steady state, a Kirchhoff loop law is obeyed such that equation (24) holds. All subscripts refer to the reaction scheme shown in figure 2. Reaction parameters for figure are as follows. Equilibrium constants: $K_1 = 100$, $K_2 = 0.2$, $K_3 = 500$, $K_4 = 10^6$, $K_5 = 10^{-4}$, $K_6 = 500$, and $K_7 = 10^{-6}$. Rate parameters (volume-independent): $k_2 = 10^6$, $k_3 = 10^4$, $k_4 = 10^5$, $k_5 = 5 \cdot 10^4$, $k_6 = 10^3$, and $k_{-4} = 10^3$, $k_{-2} = 2$, $k_9 = 10^3$, $k_{-4} = 10^6$. Concentrations: figures (A)–(C), $[A] = 10^{-4}$ and $[C] = 1.4373 \cdot 10^{-21}$ (both fixed); for figure (D), the initial condition is $[C] = 1.4373 \cdot 10^{-7}$ and the initial concentrations for enzymes is $[E_1] = 0.5 \cdot 10^{-3}$ and $[E_2] = 0.5 \cdot 10^{-3}$. 
where $R$ is the gas constant, $r_1$ is the forward reaction rate and $r_1$ is the reverse reaction rate [11, 30, 31]. For such a cycle at steady state,

$$\frac{r_1}{r_1} = \frac{r_2 r_3 r_4}{r_2 r_3 r_4}. \quad (24)$$

For an enzyme catalyzed reaction of a single substrate, $r_{r2}, r_{r3}$ and $r_4$ correspond to the rates of binding of the substrate, the catalytic conversion of the substrate to product, and the rate of release of product, respectively. The ratio $r_1/r_1$ then corresponds to the ratio of rates of the combined steps. The dynamics of the full set of reactions (reactions 2, −2, 3, −3, 4, −4, and 6, −6, 7, −7, 8, −8) and the summary reactions (reactions 1, −1, 5, −5) in figure 2 are compared in figure 3(D), which shows the rate ratio of each reaction as the system approaches steady state, as well as the product of rate ratios in equation (24). In this figure, comparison of the product of the ratio of rates of reaction 1 (Rxn 1, blue dashed line). While the combined trajectory for reactions 2, 3, and 4 (in the product of the ratio of the rates) start at $10^{-8}$ s, the trajectory doesn’t converge until $10^{-4}$ s. In contrast, the trajectory of the ratio of forward and reverse rates for the equivalent reaction 1 takes minimal time to converge in the first few time steps and has converged by $10^{-6}$ s. Likewise, the combined trajectory for reactions 6, 7, and 8 (Rxn 678, blue solid line) start at $10^{-8}$ s and converges by $10^{-4}$ s, while the trajectory for equivalent reaction 5 (Rxn 5, blue dashed line) converges immediately in the first few time steps at $10^{-6}$ s.

Compressing reactions 2, 3, and 4 into summary reaction 1 and reactions 6,7, and 8 into summary reaction 5 are the mass action equivalent of course-graining in atomistic simulations: the degrees of freedom are averaged over and compressed to fewer degrees of freedom, enabling the removal of high-frequency dynamics from the trajectories. Since the course-grained summary reactions 1 or 5 do not model the dynamics of enzyme binding, the rate ratio converges to the steady state value rapidly.

That the rate ratios can be predicted using course-grained summary reactions and without the need for rate constants can dramatically affect the individual and overall steps of the enzyme-catalyzed reactions, including rate constants.

**Branched reactions**

For branched reactions, however, steady state concentrations alone are not sufficient to derive the dynamics. Consider the branched reaction where a product of the first reaction $A \Rightarrow B$ is then the reactant for two reactions occurring in parallel,

$$A \Rightarrow B \Rightarrow C \Rightarrow D$$

When the concentrations of each species is large enough, the deterministic rate law for species $B$ is,

$$\frac{d[R]}{dt} = \frac{K_{11}}{K_{11}} R_A - \frac{K_{11}}{K_{11}} R_B$$

Using the approach described above and solving for rate ratios at steady state results in two unknown coupling ratios and one equation, which cannot be solved without additional information. However, if either the flux $\frac{r_3}{r_1}$ or $\frac{r_3}{r_4}$ is known, through reactions 1 and 3, respectively, is measured using isotopic labeling assays (metabolic flux analysis [33]), then this additional information allows for determination of the necessary coupling parameters,

$$\frac{r_3}{r_1} = \frac{K_{11}}{K_{11}} R_A - \frac{K_{11}}{K_{11}} R_B$$

or,

$$\frac{R_B}{R_A} = \frac{r_3}{r_1} \left(1 - \frac{K_{11}}{K_{11}} \right)$$

In fact, since all reactions in a non-branching pathway all have the same net flux at steady state, any steady state flux through a reaction in a non-branching portion of a pathway containing reaction 1 and any steady state flux through a non-branching pathway containing reaction 3 are all that are needed to determine the coupling between reactions 1 and 3.

**Useful assumptions when data on reaction affinities are not available**

What can be done if steady state measurements of metabolite levels or reaction fluxes for branched reactions are incomplete or not available? To answer this question, it is necessary to know how the coupling terms (e.g. $K_{11}$ in reaction (19)) affect the reaction dynamics—in particular, the net rate through a reaction (flux) and the energetic cost to a biological cell. As can be seen in figure 1(D), changing the ratio of the rate constants $\frac{R_1}{R_2}$ and the overall rate because each value of $\frac{R_1}{R_2}$ can correspond to multiple combinations of $R_1$ and $R_2$.)
Net rate of formation is equal to or greater than 1. Within each plot, the net rate of production of product is a sigmoidal function of the coupling term, specifically the rate parameter \( k_1 \), with the highest rates occurring when \( \exp(A(20.0\,\text{K})/\text{RT}) = K_2 n_B^{2.0}/n_A^{3.1} \) is equal to or greater than 1.0. As the driving force on the overall reaction is increased from left to right in each of the four plots, the range of values for the coupling term converges to 1.0, as predicted from equation (16) when the reverse reaction is not the range of values for the coupling term converges to 1.0, as predicted from equation (16) when the reverse reaction is not.

As can be seen in the middle row of figure 4, the average value 1.0 for the coupling term has particular significance—it is the maximum entropy change (or production) steady state out of all the possible steady states. The definition of entropy production used here is the the entropy change in going from initial products to final reactants due to reactions (rxns) \([25, 34]\),

\[
\Delta S_{\text{rxn}} = -\sum_{\text{rxns} \, \text{i states} \, \text{j}} Pr(C_i \rightarrow C_j + \delta S_i) \log Pr(C_i \rightarrow C_j + \delta S_i),
\]

(26)

where \( C_j \) is the configuration of state \( J, \delta S_i \) is a change of state due to reaction \( i, C_j + \delta S_i \) is the configuration of state \( J + \delta S_i \), and \( Pr(C_i \rightarrow C_j + \delta S_i) \) is the time independent state-to-state transition probability given by,

\[
Pr(C_i \rightarrow C_j + \delta S_i) = \frac{A_i}{\sum_{\text{rxns} \, j} A_j}.
\]

The total free energy of the system is also represented in the figures by color-coding. The state of lowest free energy is also the state of maximum entropy production.

Thus, the maximum entropy production steady state is the thermodynamically most efficient steady state (least heat dissipation) and is characterized by being the steady state of lowest total free energy and highest thermodynamic likelihood. By comparison of the top and middle rows, and especially for a non-equilibrium driving force of 2 kcal mol\(^{-1}\), it is apparent that it is possible to maximize the net rate but at the cost of thermodynamic efficiency. That is, while a somewhat higher net rate can be achieved, the system must be maintained at a higher energy. However, at even moderately high driving forces of 5 kcal mol\(^{-1}\) for the overall reaction, there is no speed advantage—the net rate is maximized very close to the thermodynamic optimum. As long as both reactions individually have...
moderate non-equilibrium driving forces, there is no speed advantage to tweaking rate constants. Rate and thermodynamic efficiency are meaningfully combined in the concept of entropy production rate, hinted at by Lotka [35, 36] and proposed by Schrödinger and many others since [37] as the operating principle for biological systems. The definition of the entropy production rate used here is,

$$\sigma = r_{\text{net}} \cdot \Delta S_{\text{net}}. \quad (27)$$

As can be seen from the bottom row of figure 4, the entropy production rate is also maximized when $K_{21} n_1^{21} / n_2^{21} = 1.0$, as are the entropy production and total free energy of the system. This is significant in that it suggests that on the larger scale of biological systems there does not necessarily need to be a tradeoff between growth rate and the thermodynamic efficiency of growth. Furthermore, the bottom row of figure 4 shows that there is a dramatic decrease in steady state entropy production rates as one moves away from the thermodynamically optimal state in which \( \exp(A_c \xi / R T) = K_{21} n_1^{21} / n_2^{21} = 1.0 \). If one assumes that adaptation favors organisms with greater entropy production, then one can create reasonable models without even the need for steady state values of reaction intermediates by setting the coupling term to the value that maximizes the entropy production rate.

The definition of entropy production used above differs from that used recently by De Decker et al, in describing entropy production rates for mass action processes [23],

$$\sigma_i = \frac{1}{T} \sum_i A_i \cdot r_i,$$

where \( T \) is the temperature. Besides the difference in units, one important difference is that the definition of entropy production rate in equation (27) uses the overall rate going from initial reactants to final products of the system rather than the rate of individual processes. One definition of entropy production rates is not necessarily more correct than any other, but rather, like any descriptive statistic, the appropriate entropy production rate formulation depends on the phenomena one is characterizing. For biological systems undergoing natural selection in which one wants to characterize competition among species, it is the overall rate of entropy production via growth that is important and not necessarily the entropy production rates of sub-processes.

**Discussion**

In the work described above, we show that the state of the system at steady state is sufficient to define the material and energy fluxes through a system of serially coupled reactions. This is particularly important for biology in that both steady state concentrations and fluxes can be determined for many chemical species.

Related to this work, Zia and Schmittmann developed a framework for generalized detailed balance which extended the concept of detailed balance to the non-equilibrium domain [38]. The idea is that any steady state is uniquely characterized by the pair \( \{Pr^*, V^*\} \), where \( Pr \) is the probability of a configuration \( C = \{n_1, \ldots, n_M\} \) and \( Pr^* \) is the value at the steady state configuration (the most probable state in a system at steady state). \( V^* \) is the set of probability currents which are proportional to the net steady state fluxes; that is, an element \( v_i \) of \( V^* \) is such that \( v_i \propto r_i - r_{-i} \), the difference of the forward and reverse reaction rates for reaction \( i \). Equilibrium is the special case of \( \{Pr^*, V^*\} \).

Zia and Schmittmann demonstrated that the pair \( \{Pr^*, V^*\} \) uniquely characterize a steady state for both equilibrium and non-equilibrium systems. In contrast, for the special case of sequentially coupled reactions described above, we demonstrated that configuration, \( C^* = \{n_1^*, \ldots, n_M^*\} \), the set of counts/concentrations at steady state, alone is sufficient to uniquely characterize a non-equilibrium steady state in terms of relative time, and that the absolute time-dependence can be determined with knowledge of just one rate. The reason for this is that the standard chemical potentials of species in a reaction constrain the dynamics of forward and reverse rates of the same reaction, while the dynamics between reactions are likewise constrained by the observed steady state chemical potentials. The steady state chemical potentials as parameters for an open system determine the probability of any chemical species \( i \),

$$\theta_i = \frac{e^{-\mu_i / k_B T}}{\sum_j e^{-\mu_j / k_B T}},$$

which in turn determines \( Pr^* \), the probability density of the steady state configuration \( C^* \) [25],

$$Pr(C^*|\theta_1, \ldots, \theta_M) = N_T^M \prod_{i=1}^M \frac{1}{n_i!} \theta_i^{n_i}. \quad (28)$$

Equation (28) is directly related to the microscopic free energy of the system [25]. It is clear from the work described above that for branched reaction pathways, knowledge of \( \Theta = \{\theta_1, \ldots, \theta_M\} \) (or \( \{\mu_1, \ldots, \mu_M\} \)) and only a subset of \( V^* \) is sufficient to characterize a unique non-equilibrium steady state. This begs the question, how many fluxes are needed to uniquely characterize a steady state in addition to the \( M \) chemical potentials/concentrations

Fleming et al [7] have recently addressed a this question—what conditions are necessary and sufficient for duality between unidirectional fluxes and the concentrations counts of chemical species in systems of coupled reactions? Their conclusion was that for a stoichiometric matrix of dimensions \( M \times P \) consisting of \( M \) rows corresponding to chemical species and \( P \) columns corresponding to chemical reactions, the rank of the stoichiometric matrix determines the number of reactions whose dynamics can be determined based on species concentrations alone. Surprisingly, they found that 26 of 29 genome-scale
metabolic models had stoichiometric matrices that were full rank. Of the remaining three stoichiometric matrices of the models, all were less than full rank by only three or less. Since the steady state concentrations are used to set chemical potentials for use as parameters in an open system, the implication for the work presented above is that the relative reaction dynamics could in principle be simulated using chemical potentials in place of rate constants for as many reactions as there were metabolites in the full rank models. This suggests that only \( M_p \)-fluxes are needed in addition to the \( M \) chemical potentials/concentrations to uniquely characterize any steady state system without rate constants.

The demonstration that the maximum entropy production rate solution to the system of equations in scheme B is in many cases also the maximum rate solution prompts discussion on the nature of natural selection. Previously, a correlation between the standard entropy production rate solution to the system of equations in a simple metabolic model was demonstrated [39]. If metabolite measurements are available, the hypothesis of maximum entropy for steady states of evolutionary optimized systems is testable [40, 41]. It is important to note however, that while entropy production may be maximized at steady state, steady state entropy production rates are constrained by the rates achievable by the respective enzymes [42].

From an evolutionary perspective, conservation of a specific rate constant would be very hard and would require strict error tolerances in replication, which would lead to decreased ability to adapt. More likely, fitness fluctuates around the state of maximum entropy production. However, it must be kept in mind that the coupled reactions of metabolism do not represent the total thermodynamics of the cell and it is even possible for an individual reaction entropy to decrease rather than increase. Feedback regulation of enzyme activities may occur for many reasons, such as regulating the production of metabolites to synchronize the cell’s network of coupled reactions with those of the environment [43]. Regardless, predictive models of complex adaptive systems such as those found in biology and elsewhere that do not depend on hard-to-measure parameters are urgently needed to accelerate research in systems relevant to medicine, climate and energy challenges. We have outlined a statistical thermodynamic framework that would enable such large-scale simulations. The approach does not rely on kinetic parameters, but rather on standard free energy values and metabolite concentrations for which robust measurement methods are being developed [44].

Acknowledgments

The development of new discrete/stochastic methods was the basis for this work and was supported by the US Department of Energy’s Office of Biological and Environmental Research and the Environmental Molecular Sciences Laboratory (EMSL). Application to continuous/deterministic systems was supported by the Microbiomes in Transition Laboratory Directed Research and Development Initiative at PNNL. EMSL is a national scientific user facility operated by PNNL for the Office of Biological and Environmental Research at the US Department of Energy. PNNL is operated by Battelle for the US Department of Energy under Contract DE-AC06-76RLO.

Supplementary material

The details of the computations regarding the coupled enzyme simulations (see figures 2 and 3) are available as a literate computing (Jupyter) notebook with Matlab code at https://github.com/wrcannon/CoupledEnzymeSimulations and can be viewed online at http://nbviewer.jupyter.org/github/wrcannon/CoupledEnzymeSimulations/blob/master/CoupledEnzymeSimulations_PhysicalBiology.ipynb.

Alternatively, a pdf version of the notebook is provided as a supplementary file.

ORCID iDs

William R Cannon https://orcid.org/0000-0003-3789-7889
Scott E Baker https://orcid.org/0000-0001-5085-3106

References

[1] Waage P and Guldberg C M 1986 Studies concerning affinity J. Chem. Educ. 63 1044–47
[2] Guldberg C M and Waage P 1879 Über die chemische Affinität Erdmann’s J. Pract. Chem. 127 69–114
[3] Guldberg C M and Waage P 1867 Etudes sur les affinités chimiques (Christiania: Brogger & Christie)
[4] Voit E O, Martens H A and Omholt S W 2015 150 years of the mass-action law PloS Comput. Biol. 11 e1004012
[5] Henry C S, Broadbelt L J and Hatzimanikatis V 2007 Thermodynamics-based metabolic flux analysis Biophys. J. 92 1792–805
[6] Chowdhury A, Zomorrodhi A R and Maranas C D 2014 k-OptForce: integrating kinetics with flux balance analysis for strain design PLoS Comput. Biol. 10 e1003487
[7] Fleming R M, Vlassis N, Thiele I and Saunders M A 2016 Conditions for duality between fluxes and concentrations in biochemical networks J. Theor. Biol. 409 1–10
[8] Nielsen J 1997 Metabolic control analysis of biochemical pathways based on a thermokinetic description of reaction rates Biochem. J. 321 133–8
[9] van der Meer R, Westerhoff H V and Vandam K 1980 Linear relation between rate and thermodynamic force in enzyme-catalyzed reactions Biochim. Biophys. Acta 591 488–93
[10] Boudart M 1976 Consistency between kinetics and thermodynamics J. Phys. Chem. 80 2869–70
[11] Temkin M I 1963 Kinetics of stationary reactions Dokl. Akad. Nauk. SSSR 152 156 https://arxiv.org/abs/1001.2861
[12] Vlad M O, Popa V T and Ross J 2011 Consistency between kinetics and thermodynamics: general scaling conditions for reaction rates of nonlinear chemical systems without constraints far from equilibrium J. Phys. Chem. A 115 507–13
