Thermodynamically Feasible Kinetic Models of Reaction Networks

Michael Ederer and Ernst Dieter Gilles
Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany

ABSTRACT   The dynamics of biological reaction networks are strongly constrained by thermodynamics. An holistic understanding of their behavior and regulation requires mathematical models that observe these constraints. However, kinetic models may easily violate the constraints imposed by the principle of detailed balance, if no special care is taken. Detailed balance demands that in thermodynamic equilibrium all fluxes vanish. We introduce a thermodynamic-kinetic modeling (TKM) formalism that adapts the concepts of potentials and forces from irreversible thermodynamics to kinetic modeling. In the proposed formalism, the thermokinetic potential of a compound is proportional to its concentration. The proportionality factor is a compound-specific parameter called capacity. The thermokinetic force of a reaction is a function of the potentials. Every reaction has a resistance that is the ratio of thermokinetic force and reaction rate. For mass-action type kinetics, the resistances are constant. Since it relies on the thermodynamic concept of potentials and forces, the TKM formalism structurally observes detailed balance for all values of capacities and resistances. Thus, it provides an easy way to formulate physically feasible, kinetic models of biological reaction networks. The TKM formalism is useful for modeling large biological networks that are subject to many detailed balance relations.

INTRODUCTION

Mathematical models and computer simulations are helpful tools to cope with the increasingly complex reaction networks studied in biology (1). A popular and powerful modeling approach is deterministic kinetic modeling that leads to ordinary differential equation systems (2). The behavior of reaction systems is strongly determined by structural constraints on mass and energy flow that follow from the stoichiometry of the system. The mass consumed by a reaction has to be equal to the mass produced. For an isobaric and isothermal reaction to proceed in the forward direction, the Gibbs energy consumed has to be larger than the produced Gibbs energy. It was shown that key aspects of functionality and regulation of metabolic networks can be understood solely in terms of these constraints (3). Since stoichiometry of most networks is known much better than kinetic rate laws, observation of these constraints is crucial for building useful models. Whereas the constraints concerning mass flow are easily captured in mass balance equations, the correct modeling of energy flow in kinetic models requires the observance of certain constraints on the parameters. Those follow from the second law of thermodynamics and the principle of detailed balance. The second law postulates that a system that is not exposed to external thermodynamic forces will reach a state called thermodynamic equilibrium. For isobaric and isothermal systems (e.g., cells), thermodynamic equilibrium is characterized by a minimum of Gibbs free energy. Detailed balance is also called microscopic reversibility, because it demands that in thermodynamic equilibrium the probability of a microscopic process is equal to the probability of its reverse process, i.e., in thermodynamic equilibrium the

forward rate of a reaction is equal to the backward rate, and thus its overall flux is zero.

Microscopic reversibility restricts the set of physically possible parameter values of thermodynamic systems. For systems near equilibrium, this was discussed in 1931 by Onsager (4,5). Near equilibrium, the constraints have the form of symmetry conditions on phenomenological parameters of the linearized equations. Constraints on stationary, far-from-equilibrium flux distributions of reaction networks were recently discussed in the literature (6–12). The constraints can be written as sign conditions on stationary fluxes. These publications also contain good introductions to the treatment of nonequilibrium reaction systems in thermodynamics. A review concerning detailed balance in dynamic, kinetic modeling can be found in Heinrich and Schuster (13). In the literature (14,15) detailed balancing is discussed in the framework of a formal, mathematical representation of general kinetic systems. Recently, the problem of how to impose detailed balance in complex reaction mechanisms was discussed in two articles (16,17). The former proposes a method based on explicit identification of stoichiometric cycles. The latter relies on the framework of generalized kinetic systems (14) and presents a parameterization of the model in terms of equilibrium concentrations and fluxes.

Here, we attack the problem of imposing detailed balance in kinetic models from a thermodynamic point of view. We confirm and extend the results of Yang et al. (17) and give an intuitive thermodynamic interpretation.

As the following example shows, kinetic models containing true cycles may easily describe systems that violate microscopic reversibility, if no special care is taken. We consider a reaction network describing the random-order complexation of three compounds A, B, and C:
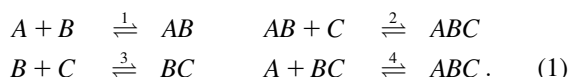
$$A + B \quad \overset{1}{\rightleftharpoons} \quad AB \qquad AB + C \quad \overset{2}{\rightleftharpoons} \quad ABC$$
$$B + C \quad \overset{3}{\rightleftharpoons} \quad BC \qquad A + BC \quad \overset{4}{\rightleftharpoons} \quad ABC \,. \qquad (1)$$

Because the order of complexation is arbitrary, the system contains a cycle:

$$
\begin{array}{ccc}
A + B + C & \overset{1}{\rightleftharpoons} & AB + C \\
\updownarrow \!\downarrow {\scriptstyle 3} & & \updownarrow \!\downarrow {\scriptstyle 2} \\
A + BC & \overset{4}{\rightleftharpoons} & ABC
\end{array}
\,. \qquad (2)
$$
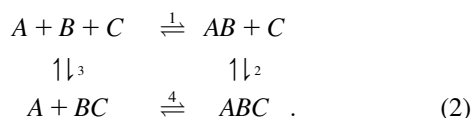
Assuming a homogeneous phase of constant volume, the mole balances read

$$\dot{c}_A = -J_1 - J_4 \qquad \dot{c}_{AB} = +J_1 - J_2$$
$$\dot{c}_B = -J_1 - J_3 \qquad \dot{c}_{BC} = +J_3 - J_4$$
$$\dot{c}_C = -J_2 - J_3 \qquad \dot{c}_{ABC} = +J_2 + J_4 \qquad . \qquad (3)$$

Here c_i [mol m^{-3}] and J_j [mol m^{-3} s^{-1}] denote concentrations and fluxes, respectively. We assume that the reactions obey a mass-action law with kinetic parameters $k_{\pm j}$:

$$J_1 = k_{+1}c_A c_B \quad -k_{-1}c_{AB}$$
$$J_2 = k_{+2}c_{AB}c_C \quad -k_{-2}c_{ABC}$$
$$J_3 = k_{+3}c_B c_C \quad -k_{-3}c_{BC}$$
$$J_4 = k_{+4}c_A c_{BC} \quad -k_{-4}c_{ABC} \,. \qquad (4)$$

The second law and detailed balance demand that thermodynamic equilibrium with $J_{1,eq} = J_{2,eq} = J_{3,eq} = J_{4,eq} = 0$ exists. With $c_{i,eq}$ denoting the equilibrium concentrations this leads to the condition that the product of the equilibrium constants along the cycle has to be one:

$$\frac{c_{AB,eq}}{c_{A,eq}c_{B,eq}} \quad \frac{c_{ABC,eq}}{c_{AB,eq}c_{C,eq}} \quad \frac{c_{BC,eq}c_{A,eq}}{c_{ABC,eq}} \quad \frac{c_{B,eq}c_{C,eq}}{c_{BC,eq}} =$$
$$\frac{k_{+1}}{k_{-1}} \quad \frac{k_{+2}}{k_{-2}} \quad \frac{k_{-4}}{k_{+4}} \quad \frac{k_{-3}}{k_{+3}} = 1 \quad . \qquad (5)$$

The above equation for the parameters is called detailed balance relation or Wegscheider condition. For all other parameter combinations the model describes a physically impossible system with a non-zero steady-state flux and thus a permanent deviation of the concentrations from the equilibrium ratios.

The deviation could be used by an attached system to perform work; i.e., energy-rich products could be formed from energy-poor precursors without consumption of energy-rich substrates. Thus the model appears as a chemical perpetuummobile, which can produce chemical energy. In this sense, a model violating the Wegscheider conditions may be interpreted as a system violating energy conservation (18).

The argumentation above is based on the existence of thermodynamic equilibrium, but is not limited to systems that actually reach thermodynamic equilibrium. The crucial point is, that a system will reach thermodynamic equilibrium, if it is isolated from its surroundings (all boundary fluxes zero). In biological terms this means that the system will die, if feeding stops.

Determination of detailed balance relations requires the consideration of the complete stoichiometry. Many models use a simplified stoichiometry, where the balance equations of

certain compounds are omitted and their concentrations are assumed to be constant. This is done often with the ubiquitous metabolites adenosine triphosphate (ATP), adenosine diphosphate (ADP), and inorganic phosphate (P). Their concentrations are assumed to be tightly regulated and usually only a small part of all reactions transforming ATP, ADP, and P are modeled. The simplified stoichiometry cannot be used for deriving detailed balance relations. We will illustrate this by means of our example network (Eq. 1). Assume that the stoichiometry of reaction 1 was not complete, but that it includes the cleavage of energy-rich ATP into its energy-poor constituents ADP and P. The true stoichiometry of reaction 1 is $A+B+ATP \rightleftharpoons AB+ADP+P$. Because the concentrations of ATP, ADP, and P are assumed to be constant, the form of the rate laws in Eq. 4 needs not to be changed, if $k_{\pm 1}$ values are allowed to depend on the concentrations of ATP, ADP, and P. By clamping these concentrations, an external thermodynamic force is imposed upon the system, preventing it from going to thermodynamic equilibrium, and driving a cyclic flux. Such cycles are called futile cycles, since they are driven by a permanent inflow of energy-rich (ATP) and outflow of energy-poor (ADP, P) compounds. In steady state the energy dissipated by the cycle is equal to the amount of work needed to keep ATP, ADP, and P concentrations constant by rephosphorylating the produced ADP. The argument used to derive the detailed balance relations in Eq. 5 of the original network is not valid here, because it is based on the assumption that thermodynamic equilibrium is reached. This shows that detailed balance relations can only be formulated for true cycles but not for futile cycles. When working with a simplified stoichiometry one cannot derive detailed balance relations, since it is impossible to decide if a certain cycle is a true cycle or a futile cycle. For this reason, we will assume that in the following the considered network stoichiometries are complete.

Whenever a system contains true cycles and its parameters are not known completely and accurately, detailed balance relations have implications on modeling and model analysis. Their explicit consideration can simplify parameter estimation from experimental data, because they lower the number of independent parameters. However, they put constraints on the usual praxis of adopting parameters from other models of similar systems. This can be illustrated with a series of models of epidermal growth factor (EGF) signaling in mammalian cells. During EGF signal transduction phosphorylated Shc, Grb2, and Sos form a complex as A, B, and C in the above example. This subnetwork is embedded in a larger network that contains many phosphorylations and dephosphorylations (especially of Shc and Sos). However, the stoichiometry of the complexation cycle is complete, because none of the reactions requires the cleavage of ATP. In Kholodenko et al. (19), a model of EGF signaling is presented that explicitly acknowledges detailed balance relations and adjusts the parameter values accordingly. The model presented in Schoeberl et al. (20) adopts some of the parameters from Kholodenko et al.

(19) but changes others without observance of detailed balance. Finally Liu et al. (21) analyzes the model of (20) but again changes a few parameters without consideration of detailed balance. Consequently the latter two models are thermodynamically infeasible and a cyclic flux can occur. A more detailed analysis would be necessary to check whether this inconsistency has a major effect on the model behavior or if its effect is negligible.

In Schoeberl et al. (20) as well as Liu et al. (21) a sensitivity analysis was performed, i.e., changes in system behavior induced by small changes in each single parameter are studied. However, because the parameters are not independent from each other, sensitivity analysis may yield misleading results, because impossible parameter variations are also considered. A more detailed analysis would also be necessary here, to judge the consequences.

A problematic aspect of the model in Eqs. 3 and 4 is that it does not formally consider the driving forces of the reactions. Model equations for electrical and mechanical systems explicitly contain voltages and forces, which are physical quantities driving electrical current and momentum changes, respectively; that is, fluxes and conjugated forces point in the same direction. In the case that no external forces are imposed to the system this guarantees the existence of an equilibrium state where all forces and fluxes are zero. Thus the fulfillment of the second law and of detailed balance is ensured by the model structure and it is not necessary to observe special detailed balance equations for the parameters. In this view model equations that violate detailed balance are comparable to the picture ''Waterfall'' by the graphic artist M.C. Escher. It shows a cyclic self-sustained water flow driving a water wheel. This is physically impossible, because water always flows from higher to lower levels; i.e., absolute water levels may serve as potentials for water flow. Graphical representation of a cyclic water flow is possible, because the mapping of a point in the drawing plane to a point in space is not unique and so it is possible to interpret the drawn water surfaces differently. The beholder chooses one of the possible interpretations of a certain surface depending on its local context in the picture. The illusion is based on the fact that every small part of the picture shows a realistic situation with a water flow from higher to lower potentials. Nevertheless as a whole it shows an impossible situation, violating energy conservation.

To structurally avoid such situations, models have to explicitly contain the potentials and driving forces. But as we will see, the quantities usually considered as potentials and driving forces of a reaction system—chemical potentials and negative Gibbs reaction energies—are quite unsuited for kinetic modeling far from equilibrium. Here, we suggest an equivalent alternative driving force that simplifies the model equations.

A holistic understanding of large biochemical networks requires models that describe the functioning of the system under the constraints of mass and energy conservation. Both classes of constraints are determined by the complete stoichiometry of the network. Kinetic modeling structurally accounts for the constraints of mass conservation in the mole balance equations, but can violate energy conservation, if the detailed balance relations are not fulfilled. This situation is unsatisfactory. In the following we will introduce a formalism that is similar to traditional kinetic modeling but structurally forbids models that violate energy conservation. The formalism is useful for modeling large reaction networks with many interconnected stoichiometric cycles. We suggest the term ''thermodynamic-kinetic modeling'' (TKM) for the proposed formalism.

METHODS

Kinetic modeling

Kinetic models can be written as

$$\dot{c} = N J(c, k), \qquad (6)$$

where $c \in \mathbf{R}^n$, $J \in \mathbf{R}^m$, and $k \in \mathbf{R}^l$ are the vectors of concentrations, fluxes, and parameters, respectively. $N \in \mathbf{R}^{n \times m}$ is the stoichiometric matrix. The kinetic properties of the reactions are modeled by the dependency of fluxes J on concentrations c and parameters k. The detailed balance relations that hold in such a system can be written as $B^T \log(K_{eq}) = 0$, where $\log(K_{eq}) \in \mathbf{R}^m$ denotes the vector of the logarithms of the equilibrium constants. The columns of the matrix $B \in R^{m \times d_m}$ span the d_m-dimensional right null space of N, i.e., $NB = 0$. Thus there exist $d_m = m - \text{rank}(N)$ independent detailed balance conditions that restrict the possible values of the equilibrium constants $K_{eq,j}$ of the reactions. These relations are referred to as the ''generalized Wegscheider condition''. For their derivation, see Heinrich and Schuster (13).

Assumptions

In the following we assume that the kinetic rate equations can be described by ideal mass-action rate laws with constant coefficients as in Eq. 4. Later we will relax this assumption and discuss systems of generalized mass-action rate laws. The use of mass-action kinetics implies two thermodynamic assumptions that are discussed in the following:

1. In mass-action networks with constant coefficients (k_{+j}, k_{-j}), the equilibrium constants (k_{+j}/k_{-j}) are, in fact, constant. In nonideal mixtures, equilibrium ''constants'' are not constant, but depend on the composition of the phase. To exclude such cases, we have to assume that the mixture is ideal.
2. In general, irreversible processes can influence each other. An example for this phenomenon is the thermoelectric effect, i.e., a temperature gradient drives an electrical current and an electrical voltage drives a heat flow (4,5). A similar effect may occur between two reactions. This means that the driving force of reaction 1 may drive reaction 2 and vice versa. Here however we assume that the reactions are thermodynamically independent and thus a reaction is only driven by its conjugated force. This is usually fulfilled if we consider elementary reactions.

We apply an additional assumption that simplifies the further considerations:

3. We assume a homogeneous phase with constant temperature T, pressure p, volume V, and molar density c_Σ [mol m^{-3}]. The molar density is the sum of the concentrations of all species in a phase. It is approximatively constant if we assume a high fraction of inert components. In this case, the sum of the concentrations of the reactive species c_i is much smaller than the molar density: $\sum_i c_i \ll c_\Sigma$. The main component of biological reaction systems is usually water. It participates in many biochemical reactions, but their rates are low compared to water concentration. For

this reason, the condition $c_\Sigma \approx const$ is usually fulfilled in biological reaction networks.

Kinetic and thermodynamics of a reaction

We consider a reaction with educts E, products P, and stoichiometric coefficients ν_i:

$$\sum_{i \in E} |\nu_i| X_i \rightleftharpoons \sum_{i \in P} |\nu_i| X_i. \qquad (7)$$

The reaction rate follows a mass-action law

$$J = k_+ \prod_{i \in E} c_i^{|\nu_i|} - k_- \prod_{i \in P} c_i^{|\nu_i|} \qquad (8)$$

with constant coefficients k_+ and k_-. As can be derived from the fundamental equations of thermodynamics and the mole balances, the density of entropy production $\sigma[S]$ [J K^{-1} m^{-3} s^{-1}] of a reaction is given by $T \sigma[S] = \Delta\mu J$, where J [mol m^{-3} s^{-1}] is the reaction flux, T [K] the thermodynamic temperature and $\Delta\mu$ [J mol^{-1}] is the negative Gibbs reaction energy

$$\Delta\mu = \sum_{i \in E} |\nu_i| \mu_i - \sum_{i \in P} |\nu_i| \mu_i, \qquad (9)$$

which is a linear combination of the chemical potentials μ_i. E.g., for reaction 1 in Eq. 1: $\Delta\mu_1 = \mu_A + \mu_B - \mu_{AB}$.

For ideal mixtures, the chemical potentials μ_i are given by

$$\mu_i = \mu_{i,0}(T,p) + R^* T \log(c_i/c_\Sigma), \qquad (10)$$

with the chemical standard potential $\mu_{i,0}$ of pure X_i, and R^* the universal gas constant. So, the Gibbs reaction energy of reaction 1 is $\Delta\mu_1 = \mu_{A,0} + \mu_{B,0} - \mu_{AB,0} + R^* T \log(c_A c_B c_{AB}^{-1} c_\Sigma^{-1})$.

Because we assumed thermodynamic independence of the reactions, the entropy production density $\sigma[S]$ of each single reaction is necessarily positive or zero. This means that a reaction can only proceed if the Gibbs energy of its products is smaller than the Gibbs energy of the educts. Thus for every reaction it holds that *sign* $\Delta\mu = sign\ J$. This condition assures observance of the second law ($\sigma[S] \geq 0$) and detailed balance ($J_{eq} = 0$ for equilibrium with $\Delta\mu_{eq} = 0$). In analogy to electrical theory where electrical resistances are the ratio of voltage drop and current, there exists a positive thermodynamic resistance $\bar{R}$ with $\bar{R} = \Delta\mu/J > 0$. So, $\Delta\mu$ may be interpreted as the thermodynamic force driving flux J through resistance $\bar{R}$. Unfortunately, not even for the simplest kinetic laws is the resistance $\bar{R}$ constant; instead, it takes a complicated nonlinear form containing polynomial and logarithmic terms. E.g., the resistance of reaction 1 in Eq. 1 is $\bar{R}_1 = (\mu_{A,0} + \mu_{B,0} - \mu_{AB,0} + R^* T \log(c_A c_B c_{AB}^{-1} c_\Sigma^{-1}))/(k_{+1} c_A c_B - k_{-1} c_{AB})$. Observe, that the zero sets of numerator and denominator of $\bar{R}_1$ are equal, since $\bar{R}_1 > 0$. Thus the expression contains removable singularities.

The characterization of a reaction kinetic by its resistance $\bar{R}$ is possible, but due to its complex form, unpracticable. Hence this representation is usually only used in its linearized form in the vicinity of thermodynamic equilibrium. Since biology usually studies systems far from equilibrium, it is quite unsuited for modeling.

RESULTS

Thermokinetic potentials and capacities

We introduce a different representation of the chemical potentials μ_i that is needed later on:

$$\mu_i = R^* T \log(\xi_i) \quad \text{or} \quad \xi_i = \exp(\mu_i/(R^* T)). \qquad (11)$$

We call ξ_i the thermokinetic potential of X_i. Because it is derived from μ_i, the thermokinetic potential ξ_i also describes

the potential of a reactive species X_i. A nice property of ξ_i is that for ideal mixtures (see Eq. 10) it is proportional to c_i:

$$c_i = C_i \xi_i. \qquad (12)$$

The factor C_i [mol m^{-3}] is the capacity of i. It depends on the chemical standard potential $\mu_{i,0}$:

$$C_i(T,p) = c_\Sigma \exp(-\mu_{i,0}(T,p)/(R^* T)). \qquad (13)$$

The Gibbs reaction energies can be written in dependence on the ξ_i, e.g., for reaction 1 in the example network it is $\Delta\mu_1 = R^* T (\log(\xi_A \xi_B) - \log(\xi_{AB}))$.

An interpretation of the above-introduced quantities C_i and ξ_i can be gained by considering thermodynamic equilibrium states with μ_{eq}, ξ_{eq}, and c_{eq}. A reaction is in thermodynamic equilibrium, if $\Delta\mu_{eq} = 0$. Thus the condition for thermodynamic equilibrium of the whole reaction system is in matrix notation

$$0 = N^T \mu_{eq}, \ 0 = N^T \log(\xi_{eq}), \ 0 = N^T \log\left(\frac{c_{eq}}{C}\right), \qquad (14)$$

where the logarithm and the fraction c_{eq}/C have to be understood elementwise.

This means that the capacities determine the equilibrium constants:

$$K_{eq} = \prod_{i \in P} C_i^{|\nu_i|} \Big/ \prod_{i \in E} C_i^{|\nu_i|} = \frac{k_+}{k_-}. \qquad (15)$$

Forces and resistances

By a redefinition of the thermodynamic force along a reaction, we can achieve a constant resistance for simple mass-action kinetics. The key property of $\Delta\mu$ that has to be conserved is its sign distribution in dependence on the concentrations. We may easily construct a quantity F with $sign(F) = sign(\Delta\mu)$ by applying the exponential function to both terms in Eq. 9,

$$F = \exp\left(\frac{\sum_{i \in E} |\nu_i| \mu_i}{R^* T}\right) - \exp\left(\frac{\sum_{i \in P} |\nu_i| \mu_i}{R^* T}\right), \qquad (16)$$

and thus arriving at

$$F = \prod_{i \in E} \xi_i^{|\nu_i|} - \prod_{i \in P} \xi_i^{|\nu_i|}, \qquad (17)$$

equivalent to

$$F = \prod_{i \in E} C_i^{-|\nu_i|} \prod_{i \in E} c_i^{|\nu_i|} - \prod_{i \in P} C_i^{-|\nu_i|} \prod_{i \in P} c_i^{|\nu_i|}, \qquad (18)$$

which has the form of a mass-action law. E.g., for reaction 1 it is $F_1 = \xi_A \xi_B - \xi_{AB} = C_A^{-1} C_B^{-1} c_A c_B - C_{AB}^{-1} c_{AB}$. We call F the thermokinetic force along a reaction. The according resistance R [$mol^{-1}\ m^3\ s$] is defined by

$$R = F/J. \qquad (19)$$

A nice property of R is that it is constant for mass-action kinetics. A comparison of coefficients of the kinetic rate law in Eq. 8 and the relation $J = R^{-1} F$ yields that

$$R = k_+^{-1} \prod_{i \in E} C_i^{-|\nu_i|} = k_-^{-1} \prod_{i \in P} C_i^{-|\nu_i|}. \qquad (20)$$

Both expressions for R are equivalent, because the capacities C_i implicitly define the equilibrium constants (Eq. 15). For reaction 1 in the example network it holds that $R_1 = k_{+1}^{-1} C_A^{-1} C_B^{-1} = k_{-1}^{-1} C_{AB}^{-1}$.

Non-mass-action kinetics

The formalism can easily be extended to non-mass-action rate laws. Consider the generalized mass-action rate law

$$J = f(c, q) \left(k_+ \prod_{i \in E} c_i^{|\nu_i|} - k_- \prod_{i \in P} c_i^{|\nu_i|} \right), \qquad (21)$$

where $f(c, q)$ is an arbitrary positive function dependent on concentrations c and further parameters q (13,22). This is a quite general expression, since all kinetic laws in ideal mixtures can be written in this form. The corresponding resistance is given by $R^{-1} = k_+ f(c, q) \prod_{i \in E} C_i^{|\nu_i|} = k_- f(c, q) \prod_{i \in P} C_i^{|\nu_i|}$. This is an arbitrary, positive function of c. Thus for a non-mass-action kinetic the resistance is no longer constant.

We will demonstrate this for the single reaction $X \rightleftharpoons Y$ with a reversible Michaelis-Menten kinetic with enzyme concentration c_E:

$$J = \frac{c_E(c_X/k_{X,0} - c_Y/k_{Y,0})}{1 + c_X/k_{X,1} + c_Y/k_{Y,1}}. \qquad (22)$$

This is a generalized mass-action kinetic with $k_+ = 1/k_{X,0}$, $k_- = 1/k_{Y,0}$, and $f(c, q) = c_E/(1 + c_X/k_{X,1} + c_Y/k_{Y,1})$ with $q = (k_{X,1}, k_{Y,1})$. With an appropriate scaling one can use the $k_{i,0}$ as capacities $C_i = k_{i,0}$. The corresponding resistance is $R = (1 + c_X/k_{X,1} + c_Y/k_{Y,1})/c_E$ or $R = (1 + k_{X,0}/k_{X,1} \xi_X + k_{Y,0}/k_{Y,1} \xi_Y)/c_E$. This may be written as $R = \rho_0 + \rho_X \xi_X + \rho_Y \xi_Y$ where ρ_i are new parameters. Thus for complex kinetics the resistances are not constant, but are dependent on the concentrations of educts and products. For typical kinetic laws of catalyzed reactions the dependency is polynomial or rational.

For modeling reaction networks, irreversible kinetic laws, with an always-positive reaction flux, are often used. An example is the Michaelis-Menten kinetic $J = J_{max} c_S/(k_m + c_S)$ where c_S is the substrate concentration and J_{max} and k_m are parameters. Physically, however, there are no such irreversible reactions because, according to the principle of microscopic reversibility, the direction of any reaction can be reversed by providing high product and low educt concentrations. Consequently, irreversible kinetic laws cannot be modeled with the proposed formalism, because they are thermodynamically infeasible. Reactions appear to be irreversible, if the potentials of the educts $\prod_{i \in E} \xi^{|\nu_i|}$ are always much larger than those of the products $\prod_{i \in P} \xi^{|\nu_i|}$. Thus the apparent irreversibility of a reaction is a systemic property and not a property of a single reaction. When gaining for a physically sound model, the irreversibility should follow from the complete model equations and not be modeled explicitly.

Let us review what we gained by defining thermokinetic potentials ξ_i (Eq. 11) and forces F_j (Eq. 17). Irreversible thermodynamics uses chemical potentials μ_i and negative Gibbs reaction energy $\Delta\mu$ for the description of a reaction. This, however, leads to complex expressions for the resistances $\tilde{R}$. By introducing the thermokinetic force F, we heavily simplified the resistance R, which is now constant for mass-action kinetics and polynomial or rational for kinetics of catalyzed reactions. To fulfill the constraints imposed by energy conservation the modeler only has to assure that R is positive. To be able to conveniently express the thermokinetic force F, we introduced the thermokinetic potentials ξ_i. The thermokinetic force of a reaction is the difference of two power law expressions in the ξ_i whose exponents are the stoichiometric coefficients. Thermokinetic potentials ξ_i and concentrations c_i are proportional, such that a constant capacity relating both can be defined.

Model equations

With the above definitions we can write a mathematical model of a reaction system as

$$\begin{aligned} \dot{c} &= N J \\ C_i \xi_i &= c_i & \text{for} \quad i = 1 \ldots n \\ R_j J_j &= F_j(\xi) & \text{for} \quad j = 1 \ldots m. \end{aligned} \qquad (23)$$

By N we denote the stoichiometric matrix and by c and J the vectors of concentrations and reaction fluxes, respectively. We will illustrate this by writing down the model equations of our example (Eq. 1). They consist of the mole balances already formulated in Eq. 3, the relations between thermokinetic potentials and concentrations

$$\begin{aligned} C_A \xi_A &= c_A & C_{AB} \xi_{AB} &= c_{AB} \\ C_B \xi_B &= c_B & C_{BC} \xi_{BC} &= c_{BC} \\ C_C \xi_C &= c_{CC} & C_{ABC} \xi_{ABC} &= c_{ABC} \end{aligned} \qquad (24)$$

and the kinetic rate equations

$$\begin{aligned} R_1 J_1 &= \xi_A \xi_B - \xi_{AB} & R_2 J_2 &= \xi_{AB} \xi_C - \xi_{ABC} \\ R_3 J_3 &= \xi_B \xi_C - \xi_{BC} & R_4 J_4 &= \xi_{BC} \xi_A - \xi_{ABC}. \end{aligned} \qquad (25)$$

An analogy to electrical networks becomes obvious. The concentrations c_i are the chemical counterpart of electrical charges. They are the integral of the incoming flux and current, respectively. The thermokinetic potentials ξ_i are analogs of the electrical voltages. By the capacities C_i they are linked to concentrations and charges, respectively. Forces F_j correspond to voltage drops along resistances R_j. However, whereas electrical forces are always linear in the potentials,

thermokinetic forces may be nonlinear, since for reactions of higher order the forces F_j follow a nonlinear mass-action law.

The three parts of Eq. 23 correspond to the basic constraints that determine the behavior of a reaction network. The first part models the constraints introduced by mass conservation. The parameters are collected in the stoichiometric matrix N. In the second part the position of thermodynamic equilibrium is defined by the capacities C_i. Implicitly, for given concentrations this determines the signs of the reaction fluxes consistent with energy conservation and the second law. The third part represents the kinetic part: it describes the behavior of the flux in nonequilibrium. Its parameters are the resistances R_j.

Ordinary differential equations

Above we stated the model equation in the form of a differential-algebraic equation system in the variables c, ξ, and J. For practical reasons it is often advantageous to use a representation as an ordinary differential equation system. Simple operations allow the representation of J as a function of ξ or c, respectively. Thus it is possible to reduce the differential-algebraic system to an ordinary differential system in ξ or c, respectively. In the latter case, the only difference to models of the traditional kinetic modeling formalism is that the parameters are C_i and R_j and not $k_{\pm j}$.

Open systems

Usually kinetic models do not describe closed systems that actually approach thermodynamic equilibrium. The most common assumption is that several concentrations or fluxes are clamped, i.e., are fixed to a specified value. As discussed in the Introduction, this is often done with ATP, ADP, and P concentrations. When clamping these concentrations in an ATP consuming network, one implicitly assumes the existence of an attached unmodeled system that rephosphorylates the produced ADP. In TKM models clamping can be done analogously, by fixing the respective quantities. The resulting model is thermodynamically feasible, since internal cyclic fluxes are not possible. However, it does not go to thermodynamic equilibrium, since a continuous flow through the system occurs.

It is also possible to model parts of the network in the TKM formalism and parts in the traditional kinetic modeling formalism. Then some of the reaction fluxes J_j are not determined by forces and resistances, but by a traditional concentration-dependent kinetic rate equation. This may be reasonable in order to include subnetworks with partly unknown stoichiometry.

Parameters and their values

Degrees of freedom

Although capacities and resistances are well-defined physical parameters, their measurement may be difficult. The capacities C_i depend on the chemical potentials of the pure compounds $\mu_{i,0}$, but $\mu_{i,0}$ cannot easily be determined absolutely.

However, the key property is that the capacities define equilibrium concentrations and in this way the equilibrium constants. Any set of equilibrium concentrations $c_{i,eq}$ of the network may serve as capacities $C_i = c_{i,eq}$ and still the condition sign $F = $ sign J is fulfilled (see Eq. 18) and thus positive resistances R_j exist. The different possible choices of the parameters change the scaling of the thermokinetic potentials ξ_i and the forces F_j, but leave the concentrations c_i and fluxes J_i invariant.

One may ask the question: How many degrees of freedom exist for the choice of the capacities? For given initial conditions, thermodynamic equilibrium is unique and two initial conditions lead to the same equilibrium state, if stoichiometrically possible. Thus the number of independent conservation relations d_n that restrict the stoichiometric abilities of the network (e.g., $c_A + c_{AB} + c_{ABC} = const$) is equal to the number of degrees of freedom for the choice of capacities and resistances. To formalize this we reformulate the equilibrium condition in Eq. 14. The vector c_{eq} is a vector of equilibrium concentrations and thus can be used as an alternative capacity vector, if

$$N^T \log(c_{eq}) = N^T \log(C). \qquad (26)$$

These are rank(N) conditions for n concentrations c_i and thus there are $d_n = n - \text{rank}(N)$ degrees of freedom to choose new capacities c_{eq}.

After determining the capacities C_i the resistances R_j can be identified from dynamic measurement data or from a comparison of coefficients with a known kinetic law.

Example

To illustrate the determination of capacities and resistances we will compute them for the example network in Eq. 1 using the parameters given in Kholodenko et al. (19) for the complexation of phosphorylated *Shc* (A), *Grb2* (B), and *Sos* (C):

$$
\begin{aligned}
k_{+1} &= 3.0 \times 10^{+6}\,\text{M}^{-1}\text{s}^{-1} & k_{-1} &= 1.0 \times 10^{-1}\text{s}^{-1} \\
k_{+2} &= 3.0 \times 10^{+7}\,\text{M}^{-1}\text{s}^{-1} & k_{-2} &= 6.4 \times 10^{-2}\text{s}^{-1} \\
k_{+3} &= 1.0 \times 10^{+5}\,\text{M}^{-1}\text{s}^{-1} & k_{-3} &= 1.5 \times 10^{-3}\text{s}^{-1} \\
k_{+4} &= 2.1 \times 10^{+7}\,\text{M}^{-1}\text{s}^{-1} & k_{-4} &= 1.0 \times 10^{-1}\text{s}^{-1}.
\end{aligned} \qquad (27)
$$

First we determine equilibrium concentrations for A, B, C, AB, BC, and ABC to use them as capacities. Observe that for this we have three degrees of freedom, because the system contains three conservation relations ($d_n = 3$), one for each conserved moiety A, B, and C, respectively. For determining the capacities we can freely choose their total concentrations. To arrive at meaningful values we use the total concentrations given in Kholodenko et al. (19) for the total concentrations of *Shc*, *Grb2*, and *Sos*:

$$
\begin{aligned}
c_A & & + c_{AB} & & + c_{ABC} &= 150 \times 10^{-9}\text{M} \\
& c_B & + c_{AB} & + c_{BC} & + c_{ABC} &= 85 \times 10^{-9}\text{M}. \\
& & c_C & + c_{BC} & + c_{ABC} &= 34 \times 10^{-9}\text{M}
\end{aligned}
$$
$$(28)$$

With the above values we compute the steady state $\dot{c}_i = 0$. Up to rounding errors this is an equilibrium state with $J_1 = J_2 = J_3 = J_4 = 0$, because the detailed balance relations are fulfilled. The equilibrium concentrations can be used as capacities:

$$C_A \approx 82.0 \times 10^{-9} M \quad C_{AB} \approx 37.5 \times 10^{-9} M$$
$$C_B \approx 15.2 \times 10^{-9} M \quad C_{BC} \approx 1.77 \times 10^{-9} M$$
$$C_C \approx 1.74 \times 10^{-9} M \quad C_{ABC} \approx 30.5 \times 10^{-9} M. \quad (29)$$

Using Eq. 24 we now can determine the resistances R_j by a comparison of coefficients of Eq. 25 and Eq. 4:

$$R_1 = k_{+1}^{-1} C_A^{-1} C_B^{-1} = k_{-1}^{-1} C_{AB}^{-1} = 2.7 \times 10^8 M^{-1} s$$
$$R_2 = k_{+2}^{-1} C_{AB}^{-1} C_C^{-1} = k_{-2}^{-1} C_{ABC}^{-1} = 5.1 \times 10^8 M^{-1} s$$
$$R_3 = k_{+3}^{-1} C_B^{-1} C_C^{-1} = k_{-3}^{-1} C_{BC}^{-1} = 3.8 \times 10^{11} M^{-1} s$$
$$R_4 = k_{+4}^{-1} C_{BC}^{-1} C_A^{-1} = k_{-4}^{-1} C_{ABC}^{-1} = 3.3 \times 10^8 M^{-1} s. \quad (30)$$

Observe that the two alternatives for calculating the resistances R_j in the above equation are only equivalent, since the capacities C_i define a thermodynamic equilibrium. The kinetic parameters $k_{\pm j}$ of a model violating detailed balance cannot be converted into capacities C_i and resistances R_j.

A closer look at the above-computed capacities and resistances reveals large differences of their values. The low value of C_{BC} means that an accumulation of BC is thermodynamically unfavorable, because a high concentration leads to very high ξ_{BC}. This strongly activates BC-consuming reactions. The high value of R_3 indicates that additionally the formation of BC is kinetically unfavorable. This suggests that the complexation usually proceeds in the order $A + B + C \rightarrow AB + C \rightarrow ABC$. Similar considerations can be made on the basis of the parameters k_{+j} and k_{-j}. There, however, the split into thermodynamic and kinetic reasoning is difficult. The above considerations are confirmed by a numerical simulation (Fig. 1): Because of our choice of the capacities as the equilibrium state that is actually approached, all thermokinetic potentials ξ_i approach 1. Concentration c_{BC} is very low, but its thermokinetic potential ξ_{BC} is comparable to the thermokinetic potential ξ_{AB}. Force F_3 is quite high, but J_3 is very small.

In general a comparison of capacities and resistances of different reactions and compounds may be misleading, because there is some freedom in choosing these parameters. However, if the capacities are chosen to be the actually approached equilibrium concentrations, a direct comparison is possible. In case of open systems with in- and outflows that do not reach equilibrium, one can choose an equilibrium state that would be approached if the system would be closed in a certain typical state.

Parameter numbers

Let us compare the number of parameters of the traditional and the new TKM formalism. The example-model requires eight parameters in the traditional formalism (two for each reaction), but 10 in the TKM formalism (six capacities + four resistances). The eight parameters in the traditional formalism are constrained by one detailed balance equation. That is,
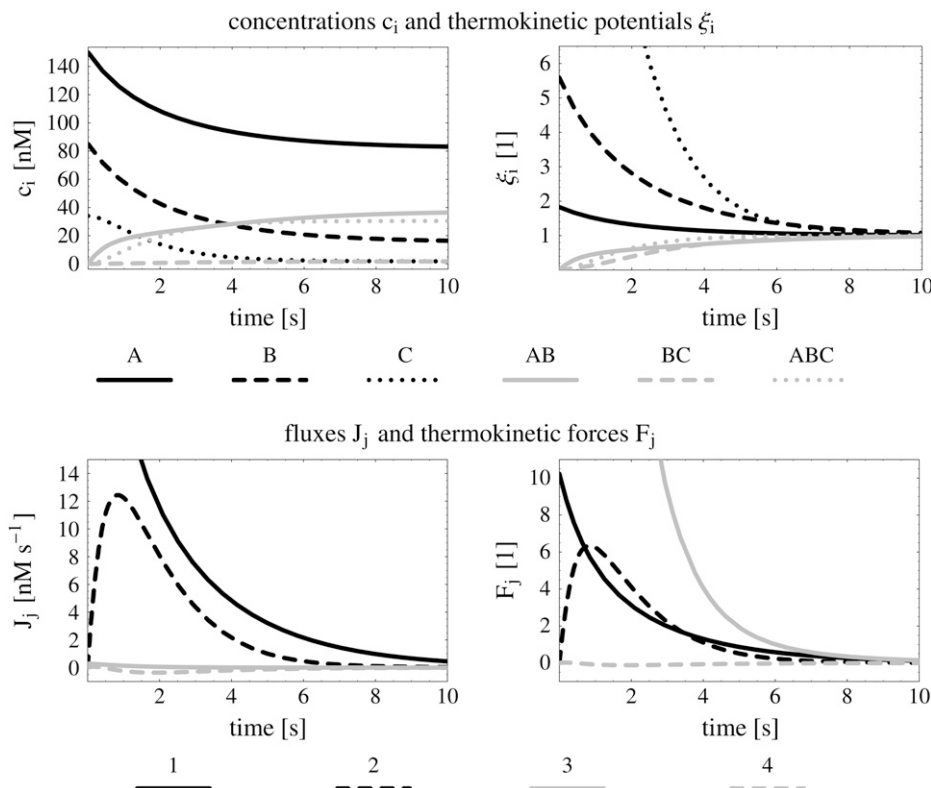

FIGURE 1 Simulation results for the example network.

there are seven independent parameters. To choose the 10 parameters of the TKM formalism we have three degrees of freedom. Thus there are effectively seven independent parameters. To compare the parameter numbers in general (see Table 1), we consider a network consisting of n compounds and m reactions with mass-action type kinetics. The left and the right null space of N have dimensions $d_n = n - \text{rank}(N)$ and $d_m = m - \text{rank}(N)$, respectively. This means there are d_n conservation relations and d_m independent cycles. The usually applied formalism requires $2\,m$ formal parameters, one for each forward k_{+j} and backward k_{-j} reaction. These parameters are partly related to each other by d_m detailed balance relations. Thus there are $2\,m - d_m = m + \text{rank}(N)$ independent parameters. Our formalism requires $m + n$ parameters. Every compound has a capacity C_i and every reaction has a resistance R_j. Since usually—in contrast to the example—the number of reactions m exceeds the number of compounds n, $m + n$ is normally $<2\,m$. Additionally every conservation relation adds an degree of freedom to the choice of parameters. Thus the number of independent parameters for a network consisting of mass-action type reactions is here, de facto, equal to $m + n - d_n = m + \text{rank}(N)$.

Neither in the traditional nor in the TKM formalism does the number of formal parameters ($2\,m$, $m + n$) match the de facto number of independent parameters $m + \text{rank}(N)$. But whereas the parameters of the traditional formalism are constrained by detailed balance relations, the parameters of the TKM formalism can be adjusted independently from each other. However, the parameters cannot be determined uniquely on the basis of measurements of concentrations and fluxes, but a given system can be described by several different pa-

rameter sets. This freedom can be used constructively by choosing typical equilibrium concentrations as capacities. In this way capacities and resistances of different compounds and reactions become comparable.

Merging models

The nonuniqueness of the parameters has to be observed, when merging two TKM models into a single model. The parameters C_i and R_j of the two models can only be combined directly, if the capacities refer to a compatible equilibrium state. Otherwise the parameters have to be converted accordingly (see Table 1). A simpler way to merge two models is to first convert the C_i and R_j to the rate parameters $k_{\pm j}$. Observe, that the $k_{\pm j}$ are not constant, if R_j correspond to a non-mass-action kinetic. Then the $k_{\pm j}$ of the two models can be merged without problems. The resulting model can be converted back to the TKM formalism. Necessary conversion laws are listed in Table 1.

Comparison to Yang et al. (17)

In Yang et al. (17) a method for imposing detailed balance in complex reaction networks is introduced. Considering Markov models of single molecule dynamics and a formal representation of general mass action kinetics, they arrive at an elegant formulation of the detailed balance relations and use those to suggest an alternative parameterization for kinetic models. They suggest parameterizing kinetic models by the equilibrium concentrations $c_{\text{eq},i}$ and the unidirectional equilibrium rates $J_{\text{eq},j}^{\pm} = k_{+}\prod_{i\in E}c_{\text{eq},i}^{|\nu_i|} = k_{-}\prod_{i\in P}c_{\text{eq},i}^{|\nu_i|}$. Here,

TABLE 1 Comparison of variables and parameters of mass-action networks

	Formalism	
	Traditional	**TKM**
Model variables		
n compounds	c_i	c_i, ξ_i
m reactions	J_j	J_j, F_j
Model parameters		
Number	$d_{\text{trad}} = 2\,m$	$d_{\text{TKM}} = m + n$
Compounds	—	$C_i > 0$
Reactions	$k_{+j}, k_{-j} > 0$	$R_j > 0$
Detailed balance relations		
Number	$d_m = m - \text{rank}(N)$	0
Equations	$B^T \log(\frac{k_+}{k_-}) = 0$ (with $N\,B = 0$)	—
Degrees of freedom in choice of parameter values		
Number	0	$d_n = n - \text{rank}(N)$
Equation	—	$N^T \log(C') = N^T \log(C'')$ (with C', C'' alternative capacity vectors)
De facto number of parameters		
	$d = d_{\text{trad}} - d_m$	$d = d_{\text{TKM}} - d_n$
	$d = m + \text{rank}(N)$	$d = m + \text{rank}(N)$
Parameter conversion		
	$k_{+j} = R_j^{-1}\prod_{i\in E_j}C_i^{-|\nu_{ij}|}$	$N^T\log(C) = \log(k_+/k_-)$
	$k_{-j} = R_j^{-1}\prod_{i\in P_j}C_i^{-|\nu_{ij}|}$	$R_j = k_{+j}^{-1}\prod_{i\in E_j}C_i^{-|\nu_{ij}|} = k_{-j}^{-1}\prod_{i\in P_j}C_i^{-|\nu_{ij}|}$

motivated from the formalism of irreversible thermodynamics and its notion of potentials and forces, we arrive at a parameterization with capacities $C_i = c_{eq,i}$ and resistances $R_j = 1/J_{eq,j}^{\pm}$ (see Eq. 20). Our approach provides an intuitive, thermodynamic interpretation of these parameters. We discuss their relation to thermodynamic potentials and forces. The use of thermokinetic potentials ξ_i and forces F_j provides a physically sound but intuitive way for the notation of model equations. The formalism is applicable to non-mass-action kinetics.

DISCUSSION

In the usual kinetic modeling formalism it is possible to build models that describe physically infeasible systems, if the Wegscheider conditions are violated. For every true cycle in the network, there exists a Wegscheider condition that constrains the set of physically feasible kinetic parameters. If the parameters violate these equations, the model describes a system that violates detailed balance. In particular, then the model does not contain a state of thermodynamic equilibrium where all fluxes vanish. This would correspond to a chemical perpetuum-mobile that maintains nonequilibrium concentrations and thus is able to perform work without consuming an equivalent amount of energy-rich substrates. Thus a violation of the Wegscheider conditions can be interpreted as a violation of energy conservation.

In the following we first discuss cases where a violation of detailed balance is critical. Then we review the TKM formalism and show how it may be useful in such cases.

Examples of large models

Recently, several large models of biochemical networks were developed that consider the complete stoichiometry of the modeled reactions. The behavior of such networks is constrained by thermodynamics.

In the literature (23,24) an approach for modeling signal transduction systems is presented that accounts for the whole combinatorial complexity emerging from the high number of different possible protein complexes. This approach is extended in Conzelmann et al. (25) to reduce the number of necessary model equations. In addition to futile cycles, networks of protein complexation contain many true cycles (see Eq. 2) as well, and thus their parameters are subject to detailed balance relations. This issue is crucial for a correct parameterization of signal transduction networks. However, it is not explicitly discussed in the above-mentioned publications and no strategies are offered to come to a valid parameterization. We expect that the behavior of signal transduction networks is constrained by thermodynamics and that some of their design principles can only be understood from this viewpoint.

For several large metabolic networks the complete stoichiometry of each reaction is available mainly due to the efforts of the group around Palsson (*Escherichia coli* (26), *Helicobacter pylori* (27), *Staphylococcus aureus* (28), *Methanosarcina barkeri* (29), and *Saccharomyces cerevisiae* (30)). These models are mainly used in the context of constraint-based modeling (31) (e.g., flux balance and elementary mode analysis). In future they may serve as the basis of kinetic models. The power of constraint-based modeling shows that the behavior of metabolic networks is strongly determined by their stoichiometry and by the constraints of mass and energy conservation (3). Thus for their kinetic modeling it is essential to explicitly consider the energy balance constraints.

Modeling, identification, and analysis

Models, as described above, contain many true cycles. Additionally, their kinetic parameters are often not known exactly. Thus if no special care is taken, the actual chosen parameter set very likely violates detailed balance relations and the model may show a behavior that is physically impossible. The situation gets more involved, if one considers parameter variations.

For parameter identification often numerical optimization methods are used. The model parameters are varied to find the parameter set that leads to the best fit of model behavior and measurement data. If detailed balance relations are not explicitly considered, the optimization method will spend most of the time examining impossible parameter sets. If some parameters are not identifiable (e.g., due to noisy or missing measurement data) the best fit is likely to be attained at infeasible parameter values. On the other hand if the detailed balance relations are considered, less measurement data is necessary to identify the parameters uniquely.

Model analysis often involves sensitivity or robustness analysis, i.e., the effects of small parameter perturbations are studied. This is also critical, because one has to take care to exclude thermodynamically infeasible parameter combinations. Otherwise, impossible parameter variations are considered and the result is biased.

The goal of systems biology is to gain a holistic understanding of biological systems. In particular, this means we must understand how biological systems function under the unalterable constraints imposed by basic physical laws. It is desirable that the model equations fulfill these constraints structurally, i.e., that with any choice of the kinetic parameters, the model describes a physically feasible system. Whereas in traditional kinetic modeling the mole balance equations structurally account for mass conservation, energy conservation has to be assured by an appropriate choice of the kinetic parameters.

TKM formalism

We introduced a novel formalism called thermodynamic-kinetic modeling (TKM) for building kinetic differential equation models of reaction networks that structurally

account for energy conservation. For this purpose the full stoichiometry of the network has to be known. The formalism explicitly considers driving forces of reactions. The driving forces are directed toward thermodynamic equilibrium and thus toward lowering the Gibbs energy in the system. The simple condition that flux and force are always directed into the same direction assures the existence of thermodynamic equilibrium and thus the observance of energy conservation. Our method differs from former approaches in the choice of potentials and forces. Traditionally the chemical potentials μ_i and consequently the negative Gibbs reaction energies $\Delta\mu_j$ are used. In Heinrich and Schuster (13) it is discussed that thermodynamic flow-force relationships of this kind are often inferior to kinetic rate equations. Kinetic rate equations allow us to incorporate detailed mechanistic knowledge and thus give deeper insight. Further, thermodynamic forces $\Delta\mu_j$ do not contain the full information that determines the reaction rates. A force $\Delta\mu$ depends only on the reaction quotient $\prod_{i\in P} c_i^{|\nu_i|} / \prod_{i\in E} c_i^{|\nu_i|}$ and thus is invariant to variations of concentrations that do not change the reaction quotient. However, reactions rates are usually not invariant to such variations. An example will clarify this: If the concentrations c_X and c_Y in the reaction $X \rightleftharpoons Y$ double, the reaction quotient c_Y/c_X and thus $\Delta\mu$ remains equal. However, if we assume mass-action kinetics the reaction rate $J = k_+ c_X - k_- c_Y$ doubles. This scaling information is not present in $\Delta\mu$ and has to be captured by the dependence of the resistance $\tilde{R}$ on the concentrations c_i. Thus the use of $\Delta\mu$ as a force for reactions far from equilibrium is unpracticable, since the according resistance—the quotient of force and flux—is highly nonlinear.

We construct the thermokinetic force F that has the same sign distribution as $\Delta\mu$, but is proportional to mass-action reaction rates. Thus the resistance $R = F/J$ of a mass-action reaction is constant. The dependence of the resistance on concentrations has to capture only effects that come from a deviation from ideal mass-action behavior. The thermo-kinetic force F can be conveniently expressed using thermo-kinetic potentials ξ_i of educts and products. They are the ratios of concentrations c_i to constant capacities C_i. Any set of equilibrium concentrations may serve as capacities C_i.
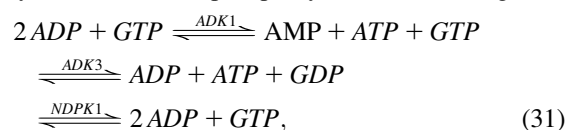
The TKM formalism provides an adequate way to introduce the thermodynamic concepts of potentials and forces to kinetic modeling and thus resolves the dilemma whether to use thermodynamic flux-force relationships or kinetic rate equations for describing biochemical reaction networks.

If, as it is the case for most natural reaction networks that the number of reactions exceeds the number of compounds, then the TKM formalism will require fewer parameters than the traditional formalism and thus simplify parameter identification and sensitivity analysis (see Table 1). A slight complication is caused by the nonuniqueness of capacities. Parameter values cannot be determined uniquely, but different parameter sets are equivalent with respect to the description of concentrations and fluxes. The different choices affect only the

scaling of potentials and forces. In praxis the nonuniqueness is not an obstacle for using the TKM formalism, because the different possible choices can be easily converted.

TKM of large metabolic systems

To demonstrate the use of the TKM formalism, we consider a hypothetical, kinetic model of the metabolism of *E. coli* K-12 iJR904 with stoichiometry as published in Reed et al. (26). The following computations were made with Mathematica (32); for the code see the Supplementary Material. The network contains $n = 762$ compounds and $m = 931$ metabolic reactions plus one reaction describing cell growth. The stoichiometric matrix $N \in \mathbf{R}^{762\times932}$ has rank(N) = 722. Thus there are $d_m = 210$ independent cycles in the network, corresponding to 210 independent Wegscheider conditions (see Table 1). A closer look reveals cycles, that contain phosphorylation of ADP, e.g.,

$$2\,ADP + GTP \xrightleftharpoons{ADK1} AMP + ATP + GTP$$
$$\xrightleftharpoons{ADK3} ADP + ATP + GDP$$
$$\xrightleftharpoons{NDPK1} 2\,ADP + GTP, \qquad (31)$$

with guanosine-diphosphate (GDP), guanosine-triphosphate (GTP), and reaction names as in Reed et al. (26). Detailed balance in such cycles is critical when studying energy metabolism. To get an overview over all cycles in the network we compute a kernel matrix $B \in \mathbf{R}^{932\times210}$ with $N\,B = 0$ using exact arithmetics. Since the kernel matrix is not unique and kernel matrices with a lot of zero entries will give shorter cycles, we bring B to column-reduced echelon form. Then ~99% of its elements are zero. The 210 columns of B correspond to 210 independent cycles. This means that 23% of the equilibrium constants cannot be adjusted freely, but are determined by Wegscheider conditions. Cycles contain between 2 and 49 reactions (mean 9.2). A reaction participates, on average, in 2.1 cycles (minimum 0, maximum 107). The number of reactions that do not participate in any cycle is equal to the number of zero rows in B and thus is independent of the actual choice of B. Only 48% of the reactions do not participate in any cycle. This means that changes in 52% of the equilibrium constants in the model have side effects on other equilibrium constants, possibly distributed over several functional units. This makes it very difficult to assess the effect of parameter changes during modeling.

Let us assume that the network is modeled by reversible, generalized mass-action kinetics (see Eq. 21). The number of parameters in the traditional formalism is $d_{trad} = 1864$ plus the parameters describing the deviation from ideal mass-action behavior. In the TKM formalism we have $d_{TKM} = 1694$ parameters plus the parameters describing the dependence of the resistances on the concentrations. The de facto number of mass-action parameters is $d = 1654$ (see Table 1). This means that the traditional formalism requires 13% and the TKM formalism 2.4% more mass-action parameters than minimally required. Further 40 capacities can be chosen freely.

TKM of large signal transduction systems

Let us assume a scaffold protein that binds k different ligands at k different binding sites. This is a typical motif in many signal-transducing pathways, where the ligands are signaling proteins, kinases, or phosphatases. Observe that the example network in Eq. 1 is of this kind with $k = 2$. Every binding site may be occupied or unoccupied. Thus there are 2^k possible complexes of the scaffold protein and the number of all compounds in the system is $n = 2^k + k$. The number of different reactions per ligand is 2^{k-1} and thus the number of all reactions is $m = 2^k k/2$. There are $d_n = k + 1$ conservation relations and thus the number of cycles is $d_m = (k/2 - 1)\, 2^k + 1$ and the number of de facto parameters is $d = (k/2 + 1)\, 2^k - 1$ (see Table 1). In Table 2 cycle numbers and parameter numbers are compared. The ratio of equilibrium constants that can be determined by Wegscheider conditions (d_m/m) grows with k. Already at $k = 4$ more than half of the equilibrium constants are determined by the rest. When comparing the minimal number of de facto mass-action parameters d with the number of mass-action parameters necessary for a model in the traditional or the TKM formalism, we get the over-parameterization of the respective formalisms. At $k = 1$ the traditional formalism has the minimal number of parameters and the TKM formalism needs twice as many parameters. But already at $k = 3$ the use of the TKM formalism is advantageous in terms of parameter numbers.

The above considerations show that the TKM formalism can significantly simplify the kinetic modeling of large systems. In large systems the fulfillment of all Wegscheider conditions is not trivial. Changes in the model that affect a cycle require adjustments in model parts that are not directly related to the original changes. Such changes are parameter variations and the introduction of additional reactions. Since cycles are ubiquitous and interwoven, the necessary adjustments may be nonintuitively and distributed over large parts of the model. The TKM formalism allows us to adjust all model parameters independently and thus frees the modeler of an explicit consideration of a high number of Wegscheider conditions. Further it requires less formal parameters than the traditional formalism and allows some capacities to be chosen freely. For the modeler, this is more convenient than to observe equality constraints between a higher number of parameters.

TKM and model parameterization

A major feature of the TKM formalism is that—in contrast to the traditional formalism—the parameters of TKM are natu-rally associated with compounds and reactions: The capacities C_i determine the equilibrium of the network. Low capacities C_i indicate a high free energy content of the respective compound. An accumulation of this compound is energetically unfavorable.

Capacities are thermodynamic quantities based on the standard chemical potentials that are organism-independent. The standard Gibbs energies of formation of many chemical species are tabulated (33). For biochemical reactions it is of advantage when working with reactants that are sums of species (e.g., ATP = ATP^{4-} + HATP^{3-} + MgATP^{2-}). For reactants, standard transformed Gibbs energies of formation can be computed and are tabulated (34). These numbers may be used as chemical standard potentials of the respective species or reactants and thus capacities can be directly calculated. Such tables are particularly useful for the modeling of metabolic networks.

Resistances R_j and capacities C_i determine the behavior in nonequilibrium. Low resistances R_i indicate that a reaction proceeds very fast, if its not in equilibrium. Resistances in metabolic networks depend on the characteristics of the catalyzing enzymes and thus are organism-dependent.

The only thermodynamic condition on the new parameters R_j and C_i is their positiveness.

TKM and sensitivity analysis

Consider a sensitivity analysis of the steady-state of the example network in Eq. 1. Remember, that the steady-state of the example is thermodynamic equilibrium. The interpretation of the steady-state sensitivity matrices $\partial c_{ss}/\partial k$ and $\partial J_{ss}/\partial k$ is difficult, since variations of the eight parameters $k_{\pm j}$ are restricted to a seven-dimensional nonlinear set. In particular it is not easily visible, that the sensitivity of the steady-state fluxes J_{ss} is zero. When one explicitly considers the Wegscheider condition, the variations of a reduced parameter vector $k' \in \mathbf{R}^7$ are studied. The eighth parameter is determined by the Wegscheider condition. Then $\partial J_{ss}/\partial k' = 0$, but the matrix $\partial c_{ss}/\partial k'$ is hard to interpret, since a variation of an element of k' is not exclusively associated with a reaction. In the TKM formalism it is $\partial c_{ss}/\partial R = 0$, reflecting the fact, that the equilibrium state is independent of the reaction mechanisms. Further $\partial J_{ss}/\partial C = 0$ and $\partial J_{ss}/\partial R = 0$, since in thermodynamic equilibrium fluxes vanish. Only the matrix $\partial c_{ss}/\partial C$ is not a zero matrix, since the equilibrium composition is determined by the capacities C_i. This example

TABLE 2 Relative cycle and parameter numbers in models of ligand bindings

k	1	2	3	4	5	6	7	8	9	10	∞
d_m/m	0%	25%	42%	53%	61%	67%	72%	75%	78%	80%	100%
$d_{trad}/d - 1$	0%	14%	26%	36%	44%	51%	56%	60%	64%	67%	100%
$d_{TKM}/d - 1$	100%	43%	21%	11%	5.4%	2.7%	1.4%	0.7%	0.36%	0.18%	0%

Equation: d_m/m is the ratio of equilibrium constants that is determined by the Wegscheider condition, $d_{trad}/d - 1$ is the over-parameterization of the traditional formalism and $d_{TKM}/d - 1$ is the over-parameterization of the TKM formalism.

shows that a correct interpretation of sensitivities with respect to mass-action parameters $k_{\pm j}$ may be difficult. Variations of capacities C_i and resistances R_j provide a natural way to study the sensitivity of reaction networks.

In systems biology one faces the problem of mathematical modeling of large reaction networks. Often their structure expressed by the stoichiometric matrix is completely known, but initially only insufficient information on kinetic parameters is available. In such cases kinetic modeling and model analysis is complicated by the need to observe detailed balance relations. By using thermokinetic potentials and forces the TKM formalism structurally avoids violation of detailed balance. Additionally, fewer parameters are required. Thus it supports the modeler in building physically sound models.

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at http://www.biophysj.org.

REFERENCES

1. Kitano, H. 2002. Computational systems biology. *Nature.* 420:206–210.

2. Olivier, B. G., and J. L. Snoep. 2004. Web-based kinetic modeling using JWS Online. *Bioinformatics.* 20:2143–2144.

3. Stelling, J., S. Klamt, K. Bettenbrock, S. Schuster, and E. D. Gilles. 2002. Metabolic network structure determines key aspects of functionality and regulation. *Nature.* 420:190–193.

4. Onsager, L. 1931. Reciprocal relations in irreversible processes I. *Phys. Rev.* 37:405–426.

5. Onsager, L. 1931. Reciprocal relations in irreversible processes II. *Phys. Rev.* 38:2265–2279.

6. Yang, F., H. Qian, and D. A. Beard. 2005. Ab initio prediction of thermodynamically feasible reaction directions from biochemical network stoichiometry. *Metab. Eng.* 7:251–259.

7. Qian, H., D. A. Beard, and S. dan Liang. 2003. Stoichiometric network theory for nonequilibrium biochemical systems. *Eur. J. Biochem.* 270:415–21.

8. Qian, H., and D. A. Beard. 2005. Thermodynamics of stoichiometric biochemical networks in living systems far from equilibrium. *Biophys. Chem.* 114:213–220.

9. Price, N. D., I. Famili, D. A. Beard, and B. O. Palsson. 2002. Extreme pathways and Kirchhoff's second law. *Biophys. J.* 83:2879–2882.

10. Beard, D. A., S. dan Liang, and H. Qian. 2002. Energy balance for analysis of complex metabolic networks. *Biophys. J.* 83:79–86.

11. Beard, D. A., E. Babson, E. Curtis, and H. Qian. 2004. Thermodynamic constraints for biochemical networks. *J. Theor. Biol.* 228:327–333.

12. Beard, D. A., and H. Qian. 2005. Thermodynamic-based computational profiling of cellular regulatory control in hepatocyte metabolism. *Am. J. Physiol. Endocrinol. Metab.* 288:E633–E644.

13. Heinrich, R., and S. Schuster. 1996. The Regulation of Cellular Systems. Chapman & Hall, New York.

14. Horn, F., and R. Jackson. 1972. General mass action kinetics. *Arch. Ration. Mech. Anal.* 47:81–116.

15. Feinberg, M. 1972. Complex balancing in general kinetic systems. *Arch. Ration. Mech. Anal.* 49:187–194.

16. Colquhoun, D., K. A. Dowsland, M. Beato, and A. J. R. Plested. 2004. How to impose microscopic reversibility in complex reaction mechanisms. *Biophys. J.* 86:3510–3518.

17. Yang, J., W. J. Bruno, W. S. Hlavacek, and J. Pearson. 2006. On imposing detailed balance in complex reaction mechanisms. *Biophys. J.* 91:1136–1141.

18. Qian, H. 2005. Cycle kinetics, steady state thermodynamics and motors—a paradigm for living matter physics. *J. Phys. Condens. Matter.* 17:S3783–S3794.

19. Kholodenko, B., O. Demin, G. Moehren, and J. Hoek. 1999. Quantification of short term signaling by the epidermal growth factor receptor. *J. Biol. Chem.* 274:30169–30181.

20. Schoeberl, B., C. Eichler-Jonsson, E. D. Gilles, and G. Müller. 2002. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat. Biotechnol.* 20:370–375.

21. Liu, G., M. T. Swihart, and S. Neelamegham. 2005. Sensitivity, principal component and flux analysis applied to signal transduction: the case of epidermal growth factor mediated signaling. *Bioinformatics.* 21:1194–1202.

22. Schauer, M., and R. Heinrich. 1983. Quasi-steady-state approximation in the mathematical modeling of biochemical reaction networks. *Math. Biosci.* 79:425–442.

23. Blinov, M. L., J. R. Faeder, B. Goldstein, and W. S. Hlavacek. 2004. BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. *Bioinformatics.* 20:3289–3291.

24. Blinov, M. L., J. R. Faeder, B. Goldstein, and W. S. Hlavacek. 2005. A network model of early events in epidermal growth factor receptor signaling that accounts for combinatorial complexity. *Biosystems.* 83:136–151.

25. Conzelmann, H., J. Saez-Rodriguez, T. Sauter, B. N. Kholodenko, and E. D. Gilles. 2006. A domain-oriented approach to the reduction of combinatorial complexity in signal transduction networks. *BMC Bioinformatics.* 7:34.

26. Reed, J. L., T. D. Vo, C. H. Schilling, and B. O. Palsson. 2003. An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol.* 4:R54.

27. Schilling, C. H., M. W. Covert, I. Famili, G. M. Church, J. S. Edwards, and B. O. Palsson. 2002. Genome-scale metabolic model of *Helicobacter pylori* 26695. *J. Bacteriol.* 184:4582–4593.

28. Becker, S. A., and B. O. Palsson. 2005. Genome-scale reconstruction of the metabolic network in *Staphylococcus aureus* N315: an initial draft to the two-dimensional annotation. *BMC Microbiol.* 5:8.

29. Feist, A. M., J. C. M. Scholten, B. O. Palsson, F. J. Brockman, and T. Ideker. 2006. Modeling methanogenesis with genome-scale metabolic reconstruction of *Methanosarcina barkeri. Mol. Sys. Biol.* 2): msb4100046-E1–E14.

30. Duarte, N. C., M. J. Herrgard, and B. O. Palsson. 2004. Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model. *Genome Res.* 14:1298–1309.

31. Reed, J. L., and B. O. Palsson. 2003. Thirteen years of building constraint-based in silico models of *Escherichia coli. J. Bacteriol.* 185:2692–2699.

32. Wolfram Research, Inc. 2003. Mathematica, Version 5.0. Champaign, IL.

33. Yaws, C. L. 1999. Chemical Properties Handbook. McGraw-Hill, Englewood Cliffs, NJ.

34. Alberty, R. A. 1998. Calculation of standard transformed Gibbs energies and standard transformed enthalpies of biochemical reactants. *Arch. Biochem. Biophys.* 353:116–130.