Determination of causal connectivities of species in reaction networks

William Vance*, Adam Arkin†, and John Ross*‡

*Department of Chemistry, Stanford University, Stanford, CA 94305; and †Department of Bioengineering and Chemistry, University of California, Berkeley, CA 94720
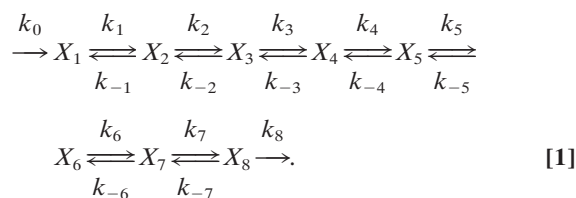
We present an approach to the determination of causal connectivities and part of the kinetics of complex reaction systems. Our approach is based on analytical and computational methods for studying the effects of a pulse change of concentration of a chemical species in a reaction network, either at equilibrium or in a nonequilibrium stationary state. Such disturbances generally propagate through a few species, depending on the values of the kinetic coefficients, before being broadened and dissipated. This short range gives a local probe of the kinetics and connectivity of the reaction network. The range of propagation also indicates species to perturb in further experiments. From piecing together these local connectivities, the global structure of the network can be constructed. The experimental design allows deduction of both reaction orders and rate constants in many cases. An example of the usefulness of the approach is illustrated on a model of a part of glycolysis.

Standard methods used for the study of the structure of reaction networks include determination of the stoichiometry and kinetics of individual steps followed by hypothesizing reaction mechanisms and testing them against such measurements; perturbations of reaction systems with first-order kinetics to determine the constant rate matrix (1); and introduction of radioactive tracers to determine reaction connectivities (2). In addition, several approaches are used to infer causal relationships in genetic networks; however, most of these techniques are applicable only to Boolean networks (3–8). Our work has taken several different approaches to this problem in nonequilibrium systems (9–15), all based on not dissecting the system but maintaining all interactions. Using fluctuating input concentrations, we have shown how correlations among species measured in the entire system may be used to construct distances among species and from this infer information on the reaction pathway (11, 12). This correlation metric construction gives a hierarchy of influence in a system and identifies weakly coupled subsystems; causal connectivities can be obtained with more difficulty. The methods outlined in the present paper give causal reaction connectivities of measured species and regulatory features of a reaction network through analyses of responses of a system to pulse perturbations. For these techniques, the concentrations of species affected by a given pulse need to be measured in time. The measurement resolution required, both in concentration and time, depends on the stationary state concentration values, the rates of reaction, and the connectivity of the reaction network. However, adequate estimates of concentration detection levels may be based on stationary state values, and estimates of time resolution levels may be based on the time of appearance of extrema after a pulse perturbation. Typically, a modest number of time measurements, often 10 or fewer, is sufficient to approximate extrema locations and values. From these measurements, both reaction orders and rate constants of an empirical rate equation may be deduced in many cases. The global structure of the reaction network is obtained by piecing together these local connectivities. We start with straightforward examples to show the simplicity of the approach and then indicate the applicability to complicated cases. An example of the usefulness of the approach is illustrated on a model of a part of glycolysis.

The number of species that need to be pulsed to deduce the causal connectivities of the species, and hence the reaction pathway, depends on the complexity of the reaction system; we do not address the issue of rules applicable to this point. At most, all the species of the system need to be pulsed, which will lead to all connectivities, but frequently fewer than all are necessary. We illustrate in the *Appendix* how connectivity may be obtained by using few perturbations in a simple example.

Method and Theory

The application of the impulse perturbation method to a reaction network consists of changing the concentration of a given species over a short time and then observing the concentration responses X induced in it and other species. A plot of the deviation in concentration from the stationary state value, $X - X_s$, against time shows that for unbranched chains of reactions, the (first) extrema in the concentration of species are ordered along the time axis with decreasing amplitude according to the number of reaction steps separating that species from the initially perturbed species. Pulses propagate strongly, i.e., with large amplitude, in the direction of net reaction velocity and weakly in the reverse direction. As an example, we consider the case of an unbranched chain of reversible first-order reactions:

$$\xrightarrow{k_0} X_1 \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X_2 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} X_3 \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} X_4 \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} X_5 \underset{k_{-5}}{\overset{k_5}{\rightleftharpoons}}$$

$$X_6 \underset{k_{-6}}{\overset{k_6}{\rightleftharpoons}} X_7 \underset{k_{-7}}{\overset{k_7}{\rightleftharpoons}} X_8 \xrightarrow{k_8}. \qquad [1]$$

For first-order reactions, it is convenient to study the relative change in concentration $u = (X - X^s) = X^s$ as a function of time, $u(t)$. These reduced variables exhibit relatively simple relations at extrema. Equating the time derivative of $u_i(t)$ to zero at the time of an occurrence of an extrema t_i, we obtain a relation between the extremum value $u_i(t_i)$ and the values of the relative concentrations of other species $u_{j \neq i}(t_i)$. In the following, we often abbreviate the extremum value as u_i^* and omit explicit reference to time t_i. From the deterministic kinetic equations for Eq. **1**, we obtain

$$u_i^* \equiv u_i(t_i) = \alpha u_{i-1}(t_i) + (1 - \alpha)u_{i+1}(t_i), \qquad [2]$$

where the coefficient α is less than one,

$$\alpha = \frac{j_f}{j_T} \leq 1; \qquad [3]$$

$j_f = k_{i-1}X_{i-1}^s$ is the steady state forward flow into X_i; and $j_T = j_f + j_r$ is the total flow into X_i. This relation shows that a maximum in u_i occurs between the curves for the preceding and succeeding
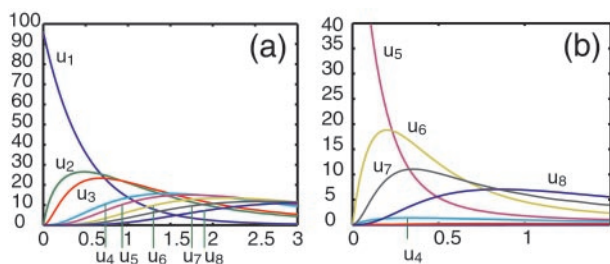
Fig. 1. Plots of the relative deviation in concentration from the stationary state versus time for all the species of the mechanism in Eq. **1**. The maxima are ordered according to the number of reaction steps separating that species from the initially perturbed species. In *b*, the pulse propagates with large amplitude in the direction of the overall reaction velocity and weakly in the opposite direction.
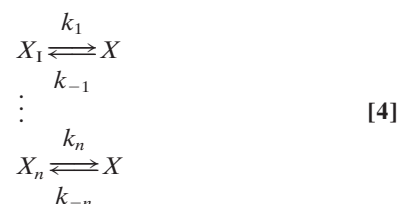
species; i.e., in addition to time ordering of extrema, the values at maxima of (relative deviations of) species are also ordered according to the number of reaction steps separating that species from the initially perturbed species. The usefulness of these relations follows from their geometric interpretation and independence from rate values. Measurements to investigate connectivity of species through reactions need only determine times of occurrences of extrema and concentration values of species at them. Fig. 1*a* shows the propagation of a pulse following a pulse perturbation of the first species X_1 in the chain of reactions. Using relative deviations of concentrations, we verify that maxima, as well as values at maxima, are ordered according to the number of reaction steps separating that species from the initially perturbed species. The effect of perturbing a species in the middle of the chain is shown in Fig. 1*b*. In this case, the pulse propagates with large amplitude in the direction of the overall reaction velocity and weakly in the opposite direction. For segments of systems in which there is not a large net reaction velocity, e.g., for subsystems close to equilibrium, a pulse perturbation induces extrema in neighboring species that are similar in magnitude: a pulse propagates equally in both directions away from the perturbation source.

More complicated mechanisms are studied through an analysis of times and values of extrema in concentration changes of species. These values give information about local connectivities and kinetics (e.g., feedback, higher-order kinetics). Several observations, which often hold, may be used as guidelines in the initial construction of connectivities of species in a reaction network. (*i*) The time of an extremum increases as the number of reaction steps separating that species from the initially perturbed species increases, unless some species act as effectors in distant reactions. (*ii*) Conversely, the initial curves of concentration changes of a species with time approach the time axis as the number of reaction steps separating that species from the initially perturbed species increases (in the absence of effectors). (*iii*) Species that are directly connected through reactions to the initially perturbed species exhibit nonzero initial slopes. (*iv*) Species that are not directly connected through reactions to the initially perturbed species exhibit zero initial slopes. (*v*) All responses are positive deviations from the stationary state unless there is a feedback, feedforward, or higher-order (>1) kinetic step. (*vi*) For short times, prior to the exit of material from the pulse to the surroundings of the reaction system, the concentration change of the pulse is conserved: the sum of deviations of concentrations (weighted by stoichiometric coefficients) must be constant and equal to the change in concentration of the initial pulse. This property is useful in confirming that all species produced from the pulse through reactions have been detected and can help in determining correct stoichiometric coefficients. (*vii*) When it is possible to identify the rate

expressions for reaction steps, then kinetic constants may be estimated from the values of the concentrations of species at an extremum. In the above example, the value of α, Eq. **3**, may be estimated from extrema, which gives the ratio k_{i-1}/k_{-i}.

We have investigated the effects of some geometrical arrangements and rate expressions on the conditions for extrema. The geometries considered include (*i*) straight chains, (*ii*) branched chains, i.e., several chains converging and diverging from a given species, (*iii*) positive and negative feedback and feedforward, and (*iv*) a cycle of reactions. The rate expressions treated include: (*i*) first- and second-order kinetics, (*ii*) generalized mass-action kinetics (for small perturbations), and (*iii*) uni-uni enzyme catalyzed reactions (Michaelis–Menten). The analytic solutions for these simplify considerably for the case of irreversible reactions.

First-Order Kinetics. We consider a set of reactions that may be described by first-order or pseudo-first-order kinetics; e.g., reversible Michaelis–Menten kinetics in which the substrate concentration is low. Specifically, the reactions either produce or consume a species X:

$$
\begin{matrix}
& k_1 & \\
X_I & \rightleftharpoons & X \\
& k_{-1} & \\
\vdots & & \\
& k_n & \\
X_n & \rightleftharpoons & X \\
& k_{-n} &
\end{matrix}
\qquad [4]
$$

Then the rate expression for X is

$$
\frac{d}{dt} X = \sum_{i=1}^{n} k_i X_i - X \sum_{i=1}^{n} k_{-i}. \qquad [5]
$$

The evolution of the relative deviation from the stationary state, $u = (X - X^s) = X^s$, may be written

$$
\frac{1}{\sum_{i=1}^{n} k_{-i}} \frac{d}{dt} u = \sum_{i=1}^{n} \alpha_i u_i - u, \qquad [6]
$$

where $\alpha_i = j_i/j_T$, $j_i = k_i X_i^s$, and $j_T = \Sigma j_i$. We note that the sum of coefficients is equal to one: $\Sigma \alpha_i = 1$. An extremum in u occurs when the above derivative with respect to time is zero, which for this system results in the expression:

$$
u^* = \sum_{i=1}^{n} \alpha_i u_i. \qquad [7]
$$

This abbreviated form for the relation between values of variables at an extremum of u omits explicit inclusion of the time variable; implicit in this expression is the evaluation of all variables at time t^* at which $du/dt = 0$.

From this expression, we have the result that an extremum u^* occurs within (the convex hull of) the curves u_i for the species that produce X. For example, in the case of a chain of irreversible reactions $X_1 \rightarrow X \rightarrow X_2$, the extremum u^* occurs on u_1; the curve for the species that produces X. Another example is the case of irreversible reactions in which two species, X_1 and X_2, separately produce X; and X produces one product X_3. Then u^* occurs between the curves for the species that produce X, u_1, and u_2, and at such a point the coefficients sum to one: $\alpha_1 + \alpha_2 = 1$.
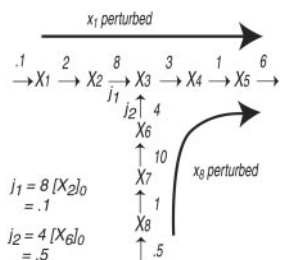
Fig. 2. Chemical reaction mechanism for converging chains of irreversible first-order reactions. The rates of production of species X_1 and X_8 are held constant at 0.1 and 0.5, respectively. $[X_i]_0$ denotes the stationary state concentration of species X_i.

Fig. 2 shows a diagram of converging chains of irreversible first-order reactions. Perturbation of species X_1 causes a pulse to propagate through the sequence $X_2; X_3; X_4; X_5$, Fig. 3a; perturbation of species X_8 causes the pulse to propagate through the sequence $X_7; X_6; X_3; X_4; X_5$, Fig. 3b. In each of these figures, we note that aside from the branch species X_3, the maximum value (of the relative deviation) of a species occurs on the curve for the species that produces it. For the branch species, the maximum value of u_3 is equal to $\alpha_i u_i$, where i is 2 or 6, depending on whether the pulse propagate through species X_2 or X_6, i.e., whether X_1 or X_8 is perturbed. The sum of the coefficients $\alpha_2 + \alpha_6$ is equal to one and is a useful check on the correct number of converging chains at X_3. These relations may be extended to networks of nearly irreversible reactions, where they hold to within a correction that is the order of the reverse rate divided by the forward rate.

Cyclic reaction systems may be investigated by using the above methods. Such systems are composed of chains of reactions that form a closed loop; inputs and outputs are branch points of the loop. Using Eq. **7**, the linear and branch points of the system may be identified through appropriate perturbations. Another class of systems for which the above methods and results are directly applicable is those in which radioactive tracer species are introduced. If a tracer is added so that the total concentration of the species is constant, then the response of the tracer obeys first-order kinetics.

Generalized Mass-Action Kinetics. The propagation of pulses, and in particular the conditions for extrema, may also be described for higher-order reaction kinetics.
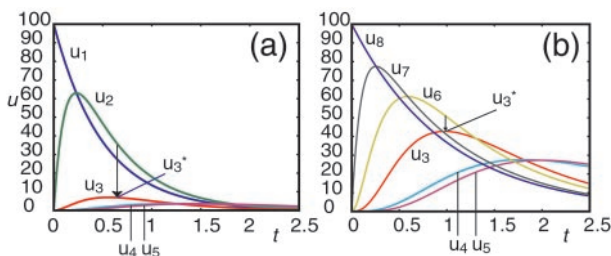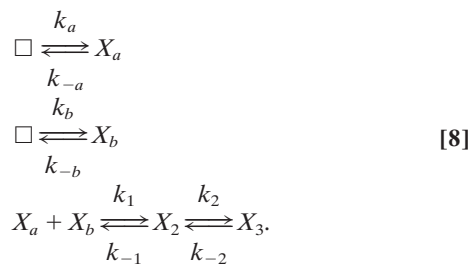

Fig. 3. Plots of relative deviation in concentration versus time for species of the mechanism in Fig. 2. In a, a perturbation of the concentration of species X_1 causes a pulse to propagate through the sequence $X_2; X_3; X_4; X_5$; in b, perturbation of species X_8 causes the pulse to propagate through the sequence $X_7; X_6; X_3; X_4; X_5$. With the exception of the branch species X_3, the maximum value (of the relative deviation) of a species occurs on the curve for the species that produces it. For the branch species, the maximum value of u_3 is equal to $\alpha_i u_i$, where i is 2 or 6, depending on whether the pulse originates from species X_2 or X_6. The coefficients α_2 and α_6 are the relative contributions of the branch fluxes to the total flux: $\alpha_2 = j_1/(j_1 + j_2)$ and $\alpha_6 = 1 - \alpha_2$.

For the case of two different species reacting, we take the following reaction scheme:

$$\square \underset{k_{-a}}{\overset{k_a}{\rightleftharpoons}} X_a$$

$$\square \underset{k_{-b}}{\overset{k_b}{\rightleftharpoons}} X_b \qquad \qquad [8]$$

$$X_a + X_b \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X_2 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} X_3.$$

The condition for an extremum in one of the species u_b involved in the bimolecular reaction may easily be derived. For small reverse velocities, $j_{-1}/j_T \ll 1$; the condition reduces to

$$u_b^* = -\frac{\gamma u_a}{1 + \gamma u_a}, \qquad \qquad [9]$$

where

$$\gamma = \frac{j_1}{j_T} = \frac{1}{1 + k_{-b}X_b^s/k_1 X_a^s X_b^s} \leq 1, \qquad [10]$$

and $j_{-1} = k_{-1}X_2^s$, $j_T = j_{-b} + j_1 = k_{-b}X_b^s + k_1 X_a^s X_b^s$. If the second-order reaction rate is much greater than the rate of the other reaction that destroys X_b, then we have $j_{-b}/j_1 \ll 1$, and Eq. **9** reduces to $u_b^* = -u_a/(1 + u_a)$. The approximate condition for an extremum in the product u_2, for small reverse rate coefficient $j_{-b}/j_1 \ll 1$

$$u_2^* = u_a + u_b + u_a u_b. \qquad \qquad [11]$$

To illustrate an application of these conditions, we consider an impulse change of concentration of species X_a in Eq. **8**. Following this perturbation, the concentration of species X_b decreases below the stationary state value and reaches a minimum value given by Eq. **9**. The concentration of the product of the reaction increases and reaches a maximum value given by Eq. **11**. For small perturbations, the maximum value u_2^* is the sum of u_a and u_b.

Generalized mass-action kinetics may be treated similarly. We consider the case in which the creation and consumption of a species X_1 are each single reactions with power rate laws (we assume the reverse rates of these reactions are small):

$$\frac{d}{dt} X_1 = \nu_1 k_1 X_a^{\eta_a} X_b^{\eta_b} \cdots X_l^{\eta_l} - \nu_2 k_2 X_1^{\eta_1} X_2^{\eta_2} \cdots X_\eta^{\eta_\eta}. \quad [12]$$

The sets of species that produce X_1; $(X_a, X_b, \ldots, X_l)$; and the set of products, $(X_1, X_2, \ldots, X_n)$; are not necessarily disjoint. These types of rate equations, called S-systems, have been applied to the analysis of biochemical pathways and genetic networks (16). Then for small deviations from the stationary state, the equation may be linearized, which leads to the following condition at an extremum in u_1 ($du_1/dt = 0$)

$$u_1^* \approx \frac{1}{\eta_1}[(\eta_a u_a + \cdots + \eta_l u_l) - (\eta_2 u_2 + \cdots + \eta_\eta u_\eta)]. \quad [13]$$

That is, as in the case of irreversible first- and second-order kinetics, at an extremum the values of the reduced variables are linearly related with coefficients determined only by the exponents η. For these and similar systems, the exponents η may be determined through experiments that probe the effects of perturbing different species that are involved in reactions that create X_1.

We illustrate the application of this relation by using a sequence of first- and second-order reactions given in Fig. 4. In this example,
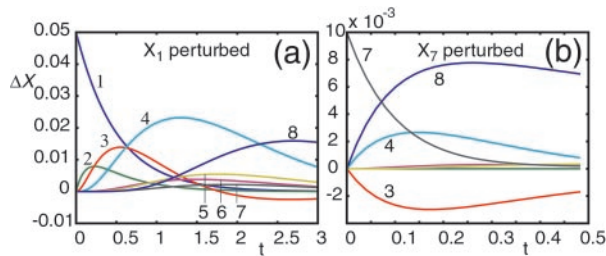
Fig. 4. Chemical reaction mechanism of a linear chain of coupled first- and second-order reactions. The order of a reaction is given by the stoichiometry.

where a single species is produced and destroyed at each reaction step, the above relation simplifies to $u_1^* \approx (\eta_a/\eta_1)u_a$. An impulse perturbation of X_1 results in the responses shown in Fig. 5. Connectivities between species are reflected in the times of extrema of the relative deviations: u_2 follows $u_1, \ldots, u_8$ follows u_7. The relations between the variables at extrema give ratios of exponents. For example, X_2 is produced through a first-order reaction involving X_1 and consumed through a second-order reaction of X_2. This results in the relation $u_2^* \approx (\eta_1/\eta_2)u_1 = (1/2)u_1$ at an extremum in X_2; i.e., the order of the second reaction is determined from the extremum relation. The other relations are similarly determined at extrema of Fig. 4 and are $u_3^* \approx 2u_2$, $u_6^* \approx (1/2)u_5$, and $u_7^* \approx 2u_6$. Using these relations along with other information, such as initial rate studies, the rate coefficients k_i may be determined.

Feedback and Feedforward. Feedback and feedforward effects arise in chemical reaction networks from a variety of mechanisms. The rate of a reaction may be either increased or decreased by an effector species not directly involved in the reaction; the location of this effector species in the network may either precede (feedforward) or succeed (feedback) the reaction. A general mechanism may be treated by linearizing the equations about a stationary state. For these cases, absolute deviations from the stationary state are more useful than relative deviations, because the species directly affected by the reaction show symmetric deviations. We consider an effector species X_{eff} either promoting or inhibiting the reaction $X_i \rightarrow X_{i+1}$, and all reverse processes are assumed negligible. We assume a general

Fig. 5. Plots of relative deviation in concentration versus time for species of the mechanism in Fig. 4. A pulse perturbation of the concentration of species X_1 results in the responses shown. The pulse propagates through the chain with maxima of relative deviations of species ordered according to the positions of species in the chain. From the plot we find the relations: $u_2^* \approx (1/2)u_1$, $u_3^* \approx 2u_2$, $u_6^* \approx (1/2)u_5$, and $u_7^* \approx 2u_6$; and $u_7^* \approx 2u_6$; the values of these coefficients may be related to the orders of the reactions through Eq. **14**.

Fig. 6. Chemical reaction mechanism illustrating positive feedback in a linear chain of irreversible first-order reactions. The rate of production of species X_1 is held constant at v_f.

rate function $f(X_i; X_{eff})$ for the feedforward or feedback that has nonzero derivatives with respect to X_i and X_{eff} at the stationary state. Then the linearized rate equations are

$$\frac{d}{dt}\Delta X_i = a_{i-1}\Delta X_{i-1} - (a_i\Delta X_i + a_{eff}\Delta X_{eff}) \quad [14]$$

$$\frac{d}{dt}\Delta X_{i+1} = a_i\Delta X_i + a_{eff}\Delta X_{eff} - a_{i+1}\Delta X_{i+1}, \quad [15]$$

where ΔX_i denotes the deviation in concentration of species X_i from the stationary state value and $a_i = \partial f/\partial X_i$, $a_{eff} = \partial f/\partial X_{eff}$. If $a_{eff} > 0$ ($a_{eff} < 0$), then X_{eff} acts as an activator (inhibitor) of the reaction. The effector species X_{eff} and its influence on the reaction may be identified through a (direct or indirect) perturbation of the concentration of X_{eff}. After such a perturbation, for short times such that changes in X_i; X_{i-1}, and X_{i+1} are small in comparison to ΔX_{eff}, deviations in species X_i and X_{i+1} are mirror images about the time axis:

$$\frac{d}{dt}\Delta X_i \approx -a_{eff}\Delta X_{eff}, \quad \frac{d}{dt}\Delta X_{i+1} \approx a_{eff}\Delta X_{eff}. \quad [16]$$

[Assuming ΔX_{eff} and the coefficients a are $O(1)$ and that $\Delta X_i(\Delta t)/\Delta X_f$ is small $O(\varepsilon)$, then $\Delta t = O(\varepsilon)$ for this relation to hold.] The extrema are ordered with the extremum of ΔX_{i+1} occurring before that of ΔX_i, independent of the sign of a_{eff} (assuming that X_{eff} does not influence X_{i-1} significantly, e.g., X_{eff} does not directly produce X_{i-1}). This may be shown as follows: if $a_{eff} > 0$, then $a_i\Delta X_i + a_{eff}\Delta X_{eff}$ starts out positive, which implies a consumption of X_i by Eq. **15**, and decreases to zero at the minimum of ΔX_i; Eq. **15** (the influence of X_{i-1} is negligible if it is not influenced directly by X_{eff}). According to Eq. **16**, ΔX_{i+1} initially increases, because $a_i\Delta X_i + a_{eff}\Delta X_{eff} > 0$, and reaches a maximum value at the crossing of the curves $a_{i+1}\Delta X_{i+1}$ and $a_i\Delta X_i + a_{eff}\Delta X_{eff}$ which occurs before the minimum of ΔX_i (at $a_i\Delta X_i + a_{eff}\Delta X_{eff} = 0$). Similarly, if $a_{eff} < 0$; then $a_i\Delta X_i + a_{eff}\Delta X_{eff}$ starts out negative and increases to zero at the maximum of ΔX_i; following the previous argument, the minimum of ΔX_{i+1} occurs before the maximum of ΔX_i. This ordering of extrema allows the identification of the direction of the net reaction velocity $X_i \rightarrow X_{i+1}$.

An example of a system with positive feedback is shown in Fig. 6. The associated rate equations are integrated for a perturbation in concentration of X_1 and X_7; the absolute deviation of concentration from the stationary state is given in Fig. 7 *a* and *b*, respectively. Fig. 7*a* shows maxima of species occurring in the order of species in the chain of reactions $X_1 \rightarrow X_8$. Species X_3 falls below the stationary state value, which indicates a possible feedback effect from one of the later species. A trial perturbation of X_7, Fig. 7*b*, shows that X_3 and X_4 are mirror images about the time axis for short times, and their initial slopes are nonzero. These observations indicate that X_7 activates the reaction from X_3 to X_4 (the peak of X_4 occurs before that of X_3, implying that X_3 precedes X_4 in the reaction sequence).

Enzyme-Catalyzed Reactions. Many segments of biochemical pathways may be modeled as unbranched chains consisting of enzyme-catalyzed reactions. We concentrate on a chain in which the kinetics are described by the irreversible Michaelis–Menten equation:

Fig. 7. Plots of deviation in concentration ΔX from the stationary state value versus time for the species of the mechanism in Fig. 6. In *a*, a perturbation of the concentration of species X_1 causes a pulse to propagate through the linear chain $X_2, X_3, \cdots, X_8$; the deviation of species X_3 falls below the stationary state value, which indicates a possible feedback effect from one of the species farther in the chain. In *b*, a perturbation of species X_7 causes the initial slopes of the deviations in X_3, X_4, and X_8 to be nonzero and the deviations in X_3 and X_4 to be mirror images about the time axis (for short times); these observations indicate that X_7 activates the reaction from X_3 to X_4 (because the peak of X_4 occurs before that of X_3, implying that X_3 precedes X_4 in the reaction sequence) and is the precursor to X_8. The maximum deviations are approximately 10% of the stationary state values or larger, except X_4 in *b*, which is 2% (the stationary state values of $X_1 - X_8$ are 0.05, 0.0125, 0.0166, 0.1, 0.0166, 0.025, 0.01, and 0.1, respectively).

$$ \to S_1 \xrightarrow{E_1} S_2 \xrightarrow{E_2} S_3 \to \qquad [17] $$
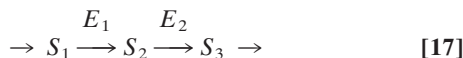
The rate of the relative deviation of S_2 is given by

$$ S_2^s \frac{d}{dt} u_2 = \frac{V_{m1}^+(1 + u_1)S_1^s/K_{m1}}{1 + (1 + u_1)S_1^s/K_{m1}} - \frac{V_{m2}^+(1 + u_2)S_2^s/K_{m3}}{1 + (1 + u_2)S_2^s/K_{m3}}. \qquad [18] $$

For small deviations from the stationary state, we linearize the equations, use the stationary state condition for the relation between S_1^s and S_2^s, and obtain the following expression for an extremum in u_2:

$$ u_2^* = \frac{1 + S_2^s/K_{m3}}{1 + S_1^s/K_{m1}} u_1, \qquad [19] $$

where $K_{m1} = (k_{-1} + k_2)/k_1$; $K_{m3} = (k_{-2} + k_3)/k_2$: We see that, unlike the case of first-order reactions, the value of u_2^* may be greater or less than u_1. If the concentrations of the substrates are much smaller than the corresponding Michaelis constants, i.e., $S_1^s/K_{m1} \ll 1$; $S_2^s/K_{m3} \ll 1$, then kinetics reduce to first-order, and the above relation becomes that for the case of irreversible first-order kinetics, $u_2^* = u_1$.
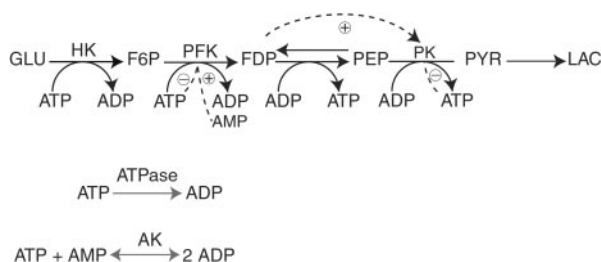
Fig. 8. Chemical reaction mechanism for an abbreviated model of the glycolytic pathway. The mechanism includes many of the known activations and inhibitions of enzymes by metabolites. Broken lines indicate activation ⊕ or inhibition ⊖ of enzymes by metabolites. The abbreviations for the enzymes are hexokinase (HK), phosphofructokinase (PFKase), and pyruvate kinase (PKase); the abbreviations for the five independent variables are fructose 6-phosphate (F6P), fructose 1,6-bisphosphate (FDP), phosphoenolpyruvate (PEP), pyruvate (PYR), and ATP (in the following figure, these are labeled 1–5, respectively).
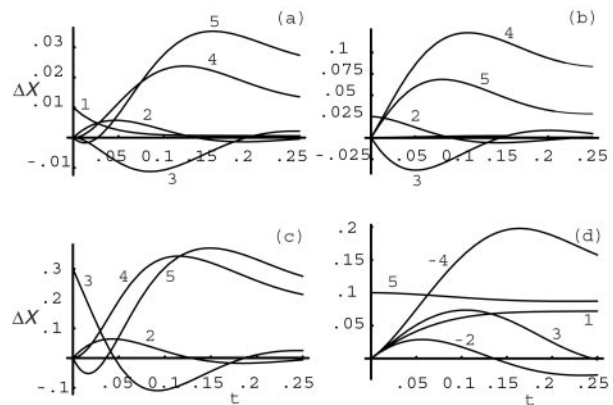

Fig. 9. Plots of time series for the absolute deviations in concentrations of species of the mechanism in Fig. 8 with impulse perturbations of F6P (1), FDP (2), PEP (3) and ATP (5). From an analysis of the responses to these perturbations, we deduce: 1 produces 2, 2 produces 3, 2 activates the conversion of 3 to 4, 3 produces 4, and 5 is consumed in the conversion of 1 to 2 (with the stoichiometric ratio of species 5 to species 2 being 1:1); the conversion of 1 to 2 is highly irreversible; 5 is produced at the same rate as 4 in the reaction 3 to 4; 5 is consumed at the same rate as 2 is produced, which indicates that 5 is a cosubstrate for the reverse reaction of 3 to 2; 5 inhibits the reaction of 1 to 2 and the reaction of 3 to 4. In each of the response plots, the perturbation is 10% of the stationary state value; responses are approximately several percent of the stationary state values or larger, except X_5, which is in *a* and *b* (the stationary state values of $X_1 - X_5$ are 0.091, 0.200, 2.70, 3.00, and 30.0, respectively).

An Example: The Glycolytic Pathway

We present an application of some of the methods outlined above to a model of the glycolytic pathway. The model we use includes many of the known activations and inhibitions of enzymes by metabolites, Fig. 8. The HKase reaction is assumed to operate at constant rate, which neglects the influence of ATP and glucose 6-phosphate. The concentrations of glucose, lactic acid, and the total adenine nucleotide pool are kept constant; the AKase reaction is assumed to be at equilibrium. Prior to the application of a pulse, the system is in a nonequilibrium stationary state. We refer to ref. 17 for details concerning the assumptions and the derivation of rate equations. There are five independent variables: fructose 6-phosphate (F6P), fructose 1,6-bisphosphate (FDP), phosphoenolpyruvate (PEP), pyruvate (PYR), and ATP; in the following, we label the variables 1–5, respectively. Our parameter values are the same as those in ref. 17, except for the input rate, which we take as $V = 1.5$ mM/min.

Fig. 9 *a–d* shows time series for the absolute deviations of species with impulse perturbations of F6P (1), FDP (2), PEP (3) and ATP (5); solutions were obtained from numerical integration of the model equations. In Fig. 9*a*, the response to an impulse perturbation of (1) is shown. The variables exhibit extrema in the following order in time 2, 3, 4; the variable 5 exhibits both a minimum and a maximum. Variables 3 and 4 have zero initial slopes, which is
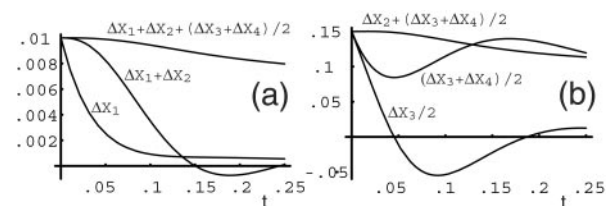

Fig. 10. Plots of time series for sums of deviations in concentrations for the species of the mechanism in Fig. 8. In *a*, the concentration of species X_1 is perturbed; in *b*, the concentration of species X_3 is perturbed. Plot *a* shows that the sum of the deviations $\Delta X_1 + \Delta X_2 + (\Delta X_3 + \Delta X_4)/2$ is slowly decaying in time. From this conservation, we infer that species 2 produces two molecules of species 3, which produces species 4. Plot *b* shows a similar transient conservation of mass in species 2, 3, and 4 following a perturbation of 3; the initial flow is from 3 to (1/2) 2 and over a longer time from 3 to 4.
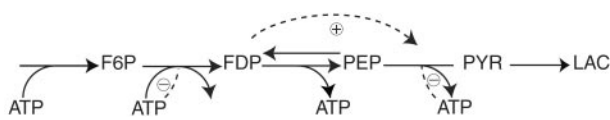
Fig. 11. A simplified mechanism constructed from the analysis of Figs. 9 and 10. This diagram captures the topology of the reaction network in Fig. 8 and many of the effectors.

Fig. 12. Chemical reaction mechanism for diverging chains of first-order reactions. The rates of production of species X_1 are held constant at 0.1.

consistent with their production from 2: the deviation in 2 starts at zero, which implies that the velocity of 2 into any of its products, e.g., (3, 4), is initially zero. Also, for short times, the deviations in 3 and 4 are mirror images about the time axis, which indicates that the conversion between 3 and 4 is affected by 2 (if 1 were to affect this conversion, then the initial slope of 3 and 4 would be nonzero). A separate perturbation of 4, which is not shown, does not produce a significant response in 3; therefore, it is inferred that 3 produces 4. A verification that 1 is converted to 4 is shown in Fig. 10*a*, where the sum of the deviations in concentrations $\Delta X_1 + \Delta X_2 + (\Delta X_3 + \Delta X_4)/2$ is slowly decaying in time. This indicates an approximate conservation of mass among these species. We note from this conservation that each molecule of species 2 produces two molecules of species 3, which in turn produces (one molecule of) species 4. The initial slope of 5 in Fig. 9*a* is equal to minus that of 2, which shows that 5 is consumed in equal numbers to 2 produced. In summary, from the responses to a perturbation of 1 (and 4), we have the results: 1 produces 2, 2 produces 3, 2 activates the conversion of 3 to 4, 3 produces 4, and 5 is consumed in the conversion of 1 to 2 (with the stoichiometric ratio of species 5 to species 2 being 1:1).

Fig. 9*b* shows responses to an impulse perturbation of species 2. The initial slopes of 3 and 4 are nonzero, which indicates that 2 has a direct effect on (activates) the conversion of 3 to 4; the initial effect of 2 is to activate the reaction from 3 to 4 (and slowly produce 3). The minimum in 3 precedes the maximum in 4 following a perturbation of 2, which indicates that 2 directly produces 3 and shifts the minimum of 3 (which would occur after the maximum in 4 without this direct production) to an earlier time. Species 1 does not respond to the perturbation, which shows that the conversion of 1 to 2 is highly irreversible. The initial slope of species 5 is equal to that of 4, which shows that 5 is produced at the same rate as 4 in the reaction 3 to 4. The new information from this perturbation is the production of 5 in this latter reaction.

Fig. 9*c* shows responses to an impulse perturbation of species 3. Both 2 and 4 have nonzero initial slopes, which shows that 3 produces both 2 and 4. The small amplitude of 2 indicates that the reaction 2 to 3 is reversible. New information is also contained in the initial time series for 2 and 5: 5 is consumed at the same rate as 2 is produced, which indicates that 5 is a cosubstrate for the reverse reaction of 3 to 2. Fig. 10*b* shows a transient conservation of mass in species 2, 3, and 4; the initial flow is from 3 to (1/2) 2 and over a longer time from 3 to 4.

Fig. 9*d* shows responses to an impulse perturbation of species 5. The minus signs in the figure denote reflection of the deviations about the time axis. The increase of 1 and decrease of 2 indicate that 5 inhibits the reaction 1 to 2, without affecting the rate of 2 to 3. The opposite rates of 3 and 4 indicate that 5 also inhibits the conversion of 3 to 4.
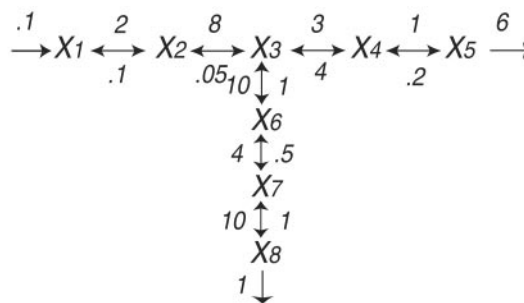
From the above analysis of the propagation of pulses in the glycolytic pathway, we may construct the reaction scheme shown in Fig. 11. This diagram captures the topology of the reaction network and many of the effectors. Further studies of the adenine nucleotides would help determine the role of the other effector (AMP) and other substrates, such as ADP, for each reaction.

Discussion

From an analysis of responses of reaction systems to pulse perturbations in concentrations of chemical species (applied to stationary state), we are able to construct the dominant kinetic structure of reaction networks. The level of information obtained depends on the number of species that are accessible to measurement and external perturbation. While correct deductions about the order of a particular reaction necessitates all involved species be identified and measured, much of the network structure can be obtained with less information. For example, in some biochemical systems, the kinetics of reactions may be treated as first-order or pseudo first-order; in these cases, identification of ordering of species in linear chains of reactions and branch points follows from analysis of extrema following appropriate perturbations. In these systems, it is also possible to determine the location of species that have not been measured in the network and the number of chains converging or diverging from a branch point from locations of extrema and sums of relative fluxes, respectively. Blind tests of the methods presented here are needed. These studies will be performed on realistic models of biological pathways and should help determine the numbers of measurements needed to construct networks of a given level of complexity.

Appendix

Much can be learned, by simple deduction, from measuring the responses to a pulse perturbation of a given species in a reaction mechanism. Consider the following case: a flow from the main branch into a side branch given in Fig. 12. Perturbations of only three chemical species, X_1, X_2, X_6, and calculations of the responses of the other species suffice to deduce the causal connectivities of the species and the reaction mechanism.

1. Wei, J. & Prater, J. D. (1962) in *Advances in Catalysis* (Academic, New York), Vol. 13, pp. 204–390.
2. Boudart, M. (1968) *Kinetics of Chemical Processes* (Prentice-Hall, Englewood Cliffs, NJ).
3. Liang, S., Fuhrman, S. & Simogyi, R. (1998) *Pacific Symposium on Biocomputing* 3, 18–29.
4. Weaver, D. C., Workman, C. T. & Stormo, G. D. (1999) *Pacific Symposium on Biocomputing* 4, 112–123.
5. Akutsu, T., Miyano, S. & Kuhara, S. (1999) *Pacific Symposium on Biocomputing* 4, 17–28.
6. Akutsu, T., Miyano, S. & Kuhara, S. (2000) *Pacific Symposium on Biocomputing* 5, 293–304.
7. Ideker, T. E., Thorsson, V. & Karp, R. M. (2000) *Pacific Symposium on Biocomputing* 5, 305–316.
8. D'Haeseleer, P., Liang, S. & Simogyi, R. (2000) *Bioinformatics* 16, 707–726.
9. Eiswirth, M., Freund, A. & Ross, J. (1991) *J. Phys. Chem.* 95, 1294–1299.
10. Gilman, A. & Ross, J. (1995) *Biophys. J.* 69, 1321–1333.
11. Arkin, A. & Ross, J. (1995) *J. Phys. Chem.* 99, 970–979.
12. Arkin, A., Shen, P. & Ross, J. (1997) *Science* 277, 1275–1279.
13. Ross, J. & Vlad, M. (1999) *Annu. Rev. Phys. Chem.* 50, 51–78.
14. Vlad, M. O., Moran, M. & Ross, J. (2000) *Physica A* 278, 504–525.
15. Tsuchiya, T. & Ross, J. (2001) *J. Phys. Chem. A* 105, 4052–4058.
16. Irvine, D. H. & Savageau, M. A. (1990) *SIAM J. Num. Anal.* 27, 704–735.
17. Termonia, Y. & Ross, J. (1981) *Proc. Natl. Acad. Sci. USA* 78, 2952–2956.

CHEMISTRY