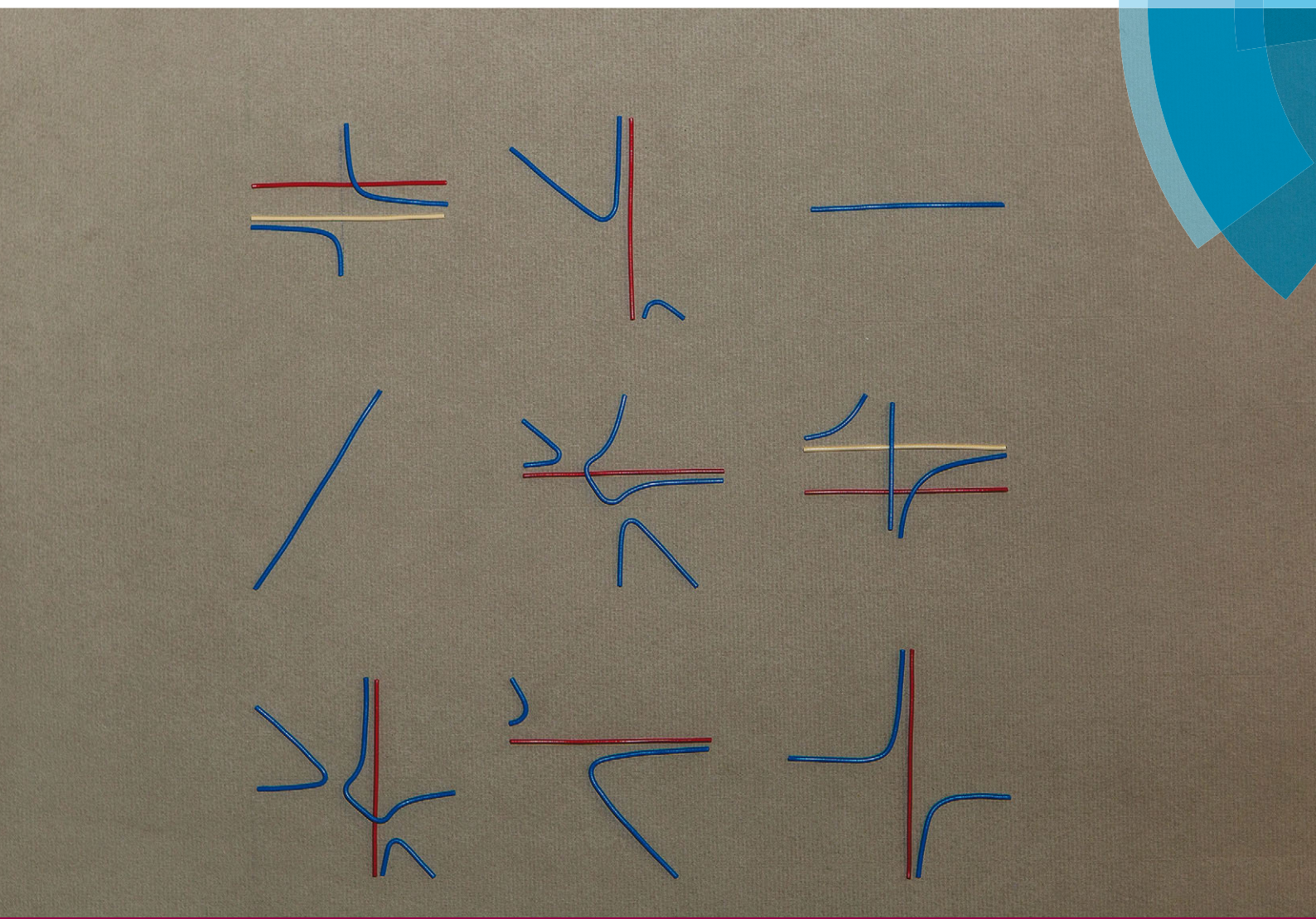
Volume 7 | Number 8 | August 2015 | Pages 837–954

Integrative Biology

Interdisciplinary approaches for molecular and cellular life sciences

www.rsc.org/ibiology



Themed issue: Integrative approaches for signalling and metabolic networks

ISSN 1757-9694

ROYAL SOCIETY OF CHEMISTRY

PAPER
Jeremy Gunawardena *et al.*
Invariants reveal multiple forms of robustness in bifunctional enzyme systems

Indexed in Medline!

Integrative Biology

PAPER

Cite this: *Integr. Biol.*, 2015, 7, 883

Invariants reveal multiple forms of robustness in bifunctional enzyme systems†

Joseph P. Dexter,‡[abc] Tathagata Dasgupta‡[c] and Jeremy Gunawardena*[c]

Experimental and theoretical studies have suggested that bifunctional enzymes catalyzing opposing modification and demodification reactions can confer steady-state concentration robustness to their substrates. However, the types of robustness and the biochemical basis for them have remained elusive. Here we report a systematic study of the most general biochemical reaction network for a bifunctional enzyme acting on a substrate with one modification site, along with eleven sub-networks with more specialized biochemical assumptions. We exploit ideas from computational algebraic geometry, introduced in previous work, to find a polynomial expression (an invariant) between the steady state concentrations of the modified and unmodified substrate for each network. We use these invariants to identify five classes of robust behavior: robust upper bounds on concentration, robust two-sided bounds on concentration ratio, hybrid robustness, absolute concentration robustness (ACR), and robust concentration ratio. This analysis demonstrates that robustness can take a variety of forms and that the type of robustness is sensitive to many biochemical details, with small changes in biochemistry leading to very different steady-state behaviors. In particular, we find that the widely-studied ACR requires highly specialized assumptions in addition to bifunctionality. An unexpected result is that the robust bounds derived from invariants are strictly tighter than those derived by *ad hoc* manipulation of the underlying differential equations, confirming the value of invariants as a tool to gain insight into biochemical reaction networks. Furthermore, invariants yield multiple experimentally testable predictions and illuminate new strategies for inferring enzymatic mechanisms from steady-state measurements.

Received 15th January 2015,
Accepted 21st May 2015

DOI: 10.1039/c5ib00009b

www.rsc.org/ibiology

Insight, innovation, integration

Bifunctional enzymes are known experimentally to maintain robust control of metabolic branch points, osmoregulatory systems, and other important cellular networks, but how this is accomplished by their underlying biochemical reaction networks has remained unclear. We performed a mathematical analysis of a compendium of networks to understand the connection between bifunctionality and robustness at a general level. Our central insight is that concentration robustness is a heterogeneous phenomenon and that subtle changes in the mechanism of a bifunctional enzyme can alter systems-level robustness. Our analysis uses techniques from computational algebraic geometry to reveal underlying polynomial "invariants" that characterize the different forms of robustness. The integration of algebraic calculations and enzymological data suggests that steady-state measurements can provide mechanistic insight into diverse biochemical systems.

1 Introduction

Cellular systems have a remarkable capacity to function correctly even as internal and external conditions change.[1] Understanding the mechanistic basis for this robustness has emerged as a major topic of interest in systems biology.[1–4] Such efforts have been hampered by the difficulty of developing precise, quantitative definitions of the various forms of biological robustness, and the lack of shared molecular or biochemical features between robust systems.

The study of enzymatic bifunctionality, however, has led to considerable progress in understanding a particular kind of robustness. Bifunctional enzymes can be divided into two broad categories, those that catalyze consecutive steps in a metabolic pathway and those that play a more regulatory role by catalyzing opposing modification and demodification of a single substrate.

[a] *Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA*
[b] *Department of Chemistry, Princeton University, Princeton, NJ, USA*
[c] *Department of Systems Biology, Harvard Medical School, Boston, MA, USA. E-mail: jeremy@hms.harvard.edu; Fax: +1 617 432 5012; Tel: +1 617 432 4839*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c5ib00009b

‡ These authors contributed equally to this work.

Little is currently known about the systems-level advantages conferred by the first group of bifunctional enzymes, and there is considerable need for further experimental and theoretical work on such enzymes. Regulatory bifunctional enzymes, in contrast, have been the subject of extensive study from a systems-level perspective and have led to new insights about robust biological control.

Experiments in bacteria and plants have revealed concentration robustness in several systems with bifunctional enzymes.[5–8] A key output of the system, usually the concentration of the substrate in a particular state of modification, was found to remain approximately constant despite changes in the expression levels of system components, including the substrate itself. In many cases, such robustness was maintained even in the face of 10-fold or greater changes in expression of system components. From consideration of these experiments, LaPorte and Koshland, working with the isocitrate dehydrogenase regulatory system, and Russo and Silhavy, working with the EnvZ/OmpR osmosensor, both in *Escherichia coli*, made early attempts to develop a mechanistic, quantitative understanding of the concentration robustness conferred by bifunctional enzymes.[6,9] Their work has been extended by a number of detailed theoretical studies of bifunctional enzyme systems.[7,10–18] In each of these studies, a specific biochemical reaction network was proposed to describe the system of interest, and mathematical calculations or numerical simulations were used to explain how concentration robustness might emerge from the reaction network. Other forms of "robustness," such as sensitivity to fluctuations in parameter values, have been considered in different contexts.[3,19–21]

Regulatory bifunctional enzymes can be subdivided into four classes according to the active site configuration of the bifunctional enzyme and the quaternary structure of the substrate (Fig. 1). Bifunctional enzymes can either have separate, non-interacting active sites for each catalytic activity (*e.g.*, a kinase domain and a distinct phosphatase domain) or a shared active site that can catalyze both activities.[16,17,22] Furthermore, a bifunctional enzyme can modify either a monomeric or multimeric substrate.

Multimeric substrates give rise to multi-site and avidity effects that require special attention. For instance, Hart *et al.* developed models of the nitrogen assimilation system in *E. coli* and the carbon fixation pathway in C_4 plants.[13,14] In both systems, they showed that mathematical robustness arises because the bifunctional enzymes have separate active sites and can therefore form ternary complexes with two units of the respective substrates. Similar analyses have been performed of the bifunctional uridylyltransferase/uridylyl-removing enzyme (UTase/UR),[18,23] also involved in nitrogen assimilation, for which there is now extensive experimental evidence of robustness.[24,25] In contrast, the isocitrate dehydrogenase (IDH) regulatory system in *E. coli* contains a bifunctional enzyme with a shared active site that acts on a dimeric substrate. In recent work we showed that, although the corresponding biochemical network does not exhibit mathematical robustness, there is a realistic parameter regime in which robustness is expected.[16]

Our focus in this paper is on monomeric substrates modified by bifunctional enzymes that have either separate or shared active sites.

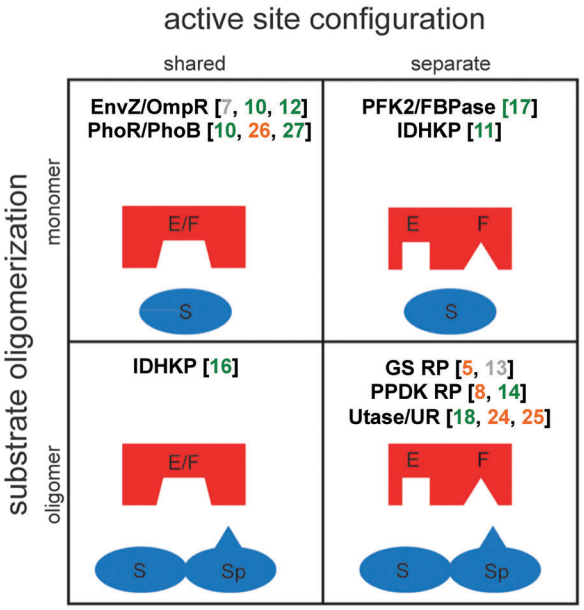


Fig. 1 Classification of bifunctional enzyme systems by biochemical features. Bifunctional enzymes that catalyze opposing post-translational modifications can be sub-divided into four classes according to the active site configuration on the enzyme and the quaternary structure of the substrate. Known examples and a cartoon diagram are provided in each quadrant of the diagram for the appropriate system. References indicate experimental (orange), theoretical (green), or combined experimental and theoretical profiling of robustness (gray). This paper focuses on the robust behavior of systems in the two top quadrants (*i.e.*, monomeric substrates regulated by a bifunctional enzyme).

Several such systems have been studied mathematically,[7,10–12] of which the EnvZ/OmpR two-component osmosensor is the paradigmatic example.[7,10] There is strong experimental evidence for robustness both in EnvZ/OmpR and in other two-component systems such as PhoR/PhoB, which senses phosphorous levels.[7,26,27] No systematic characterization of robustness, however, has been undertaken for these systems.[28] To address this limitation, we developed and analyzed an extensive compendium of reaction networks involving a bifunctional enzyme (with either shared or separate active sites) acting on a monomeric substrate.

Our analysis identified five types of robust behavior: "robust upper bounds on concentration," in which the concentration of the modified or unmodified substrate form is held below a bound set only by the parameter values (*i.e.*, the bound is independent of the total amount of substrate and enzyme); "robust two-sided bounds on the concentration ratio," in which the substrate concentration ratio is held between an upper bound and a lower bound determined only by the parameters; "hybrid robustness," in which the same network exhibits a robust upper bound in concentration for certain parameter values and robust two-sided bounds on concentration ratio for other parameter values; "absolute concentration robustness" (ACR), in which the concentration of a substrate form is fixed at a single value set by the parameters;[12] and "robust concentration ratio," in which the substrate concentration ratio depends only on the parameters. The five behaviors are all variants of concentration

robustness—some quantity involving the concentrations of the different species in the networks is held constant at steady state. We find that the type of concentration robustness is highly dependent on subtle but significant differences in biochemical mechanism. For instance, whether an enzyme–substrate ternary complex can form, and whether the reactions involving the ternary complex are symmetric or asymmetric, has a major impact on the robustness observed. Of particular interest, we demonstrate that ACR, which has dominated recent thinking about enzymatic bifunctionality, requires restrictive biochemical assumptions, and that more complex forms of robustness arise even with minimal assumptions beyond bifunctionality. Our results delineate the complexity of biochemical concentration robustness and suggest experiments that might distinguish one form of robustness from another.

The reaction networks that we study here all involve the interconversion of unmodified (S) and modified (S_P) substrate by a bifunctional regulatory enzyme (E). Assuming mass-action kinetics at steady state yields systems of non-linear polynomial equations in the concentrations of the free enzyme and substrate forms and various enzyme–substrate complexes. Our results are based on mathematical analysis of these equations with the parameters treated symbolically, thereby avoiding numerical simulations that require advance knowledge of parameter values. We use techniques from computational algebraic geometry, which we introduced in previous work, to systematically eliminate variables and to derive polynomial "invariants" involving only [S] and [S_P]. (Throughout the paper, the notation [X] signifies the concentration of X.) These invariants provide the essential insights into robust behavior. Interestingly, we find that invariants yield more accurate bounds than *ad hoc* manipulation of the underlying differential equations.

2 Methods

2.1 Reaction network assumptions

Our networks apply to the modification and demodification of any small molecule or protein in which the modification is by attachment of a small-molecule moiety (phosphate, methyl, acetyl, *etc.*).[29] Polypeptide modifications, such as ubiquitination and SUMOylation, involve more complicated biochemistry and are not modeled by our networks. As is standard, the modification donor (such as ATP in the case of phosphorylation) is assumed to be kept at constant concentration by background cellular processes and is therefore ignored as a dynamical variable. Each catalytic domain of the bifunctional enzyme is assumed to follow a standard Michaelis–Menten reaction scheme, which takes the form

$$E + S \rightleftharpoons ES \rightarrow E + S_P \qquad (1)$$

for the forward modification, with reversible formation of an enzyme–substrate complex (ES) and irreversible formation of product. The limitations of these assumptions are reviewed further in the Discussion.

The most general reaction network (network 0) is shown as biochemical reactions in Fig. 2A and as a directed graph on the corresponding "complexes" (in the sense of chemical reaction network theory[15]) in Fig. 2B. The complexes are enumerated in Fig. 2C. Subsequent networks in Fig. 4 and 5 are shown as sub-graphs of the graph in Fig. 2B with the missing vertices and edges faded. In the text, the rate constants for mass-action kinetics are given in the notation $k_{i\to j}$ for the reaction corresponding to the edge $i \to j$ in the graph on the complexes. The two catalytic activities are not assumed to be independent of each other, so that, for instance, $k_{2\to 3} \neq k_{6\to 7}$.

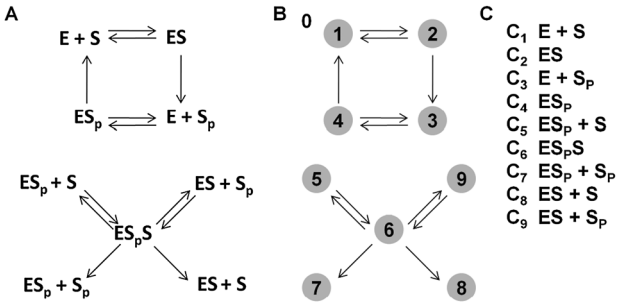


Fig. 2 Full reaction network for a bifunctional enzyme with separate active sites acting on a monomeric substrate. (A) Biochemical reaction network. (B) Directed graph describing the reaction network in A, in which the nodes correspond to the "complexes",[15] as listed to the right. We used the notation $k_{i\to j}$ for the rate constant for the reaction corresponding to the edge $i \to j$ in the graph. (C) List of complexes in the reaction network.

2.2 Algebraic calculations

All conclusions in the paper are based on algebraic analysis of the various chemical reaction networks at steady-state, assuming mass-action kinetics. We make no use of the Michaelis–Menten rate formula or any other form of approximation. Algebraic calculations were either done manually or using Mathematica 9.0 (Wolfram). An accompanying Mathematica notebook containing many of the calculations is provided as ESI.† Different numbering and notation had to be used in the notebook because of Mathematica conventions. The differences are described in tables in the notebook for each network considered in the paper, and these tables should be consulted when comparing expressions between the paper and the notebook. Gröbner basis calculations undertaken in Mathematica were performed as described previously by Manrai and Gunawardena.[30]

3 Results

3.1 Analysis of network 0

3.1.1 *Ad hoc* derivation. We begin with consideration of the full network (network 0), which includes all possible reactions involving the bifunctional enzyme and the two substrate forms (Fig. 2). In the full network, the order of formation of the ternary complex is random, and E can act as either a modifying or demodifying enzyme when bound in any complex. As such, network 0 is the most general model possible of a bifunctional enzyme acting on a monomeric substrate and is completely symmetric. We find that either [S] or [S_P] may be robustly bounded depending on the relative values of the catalytic rate constants for the ternary complex ES_PS.

At steady state, the rate of formation of S_P from S must be balanced by the rate of formation of S from S_P, which leads to the following relationship between the three complexes in the network:

$$k_{2\to3}[ES] + k_{6\to7}[ES_PS] = k_{4\to1}[ES_P] + k_{6\to8}[ES_PS]. \quad (2)$$

Solving for the ternary complex in this expression yields

$$[ES_PS] = \frac{k_{4\to1}[ES_P] - k_{2\to3}[ES]}{k_{6\to7} - k_{6\to8}}. \quad (3)$$

Here we ignore the possibility that $k_{6\to7} = k_{6\to8}$, since such fine-tuning of rate constants is unlikely to occur naturally. Production and consumption of the ternary complex must also balance at steady state, leading to the following expression:

$$[ES_PS] = \frac{k_{5\to6}[S][ES_P] + k_{9\to6}[S_P][ES]}{k_{6\to5} + k_{6\to9} + k_{6\to7} + k_{6\to8}}. \quad (4)$$

Subtracting eqn (3) from eqn (4) eliminates the ternary complex and gives

$$A[ES_P] + B[ES] = 0, \quad (5)$$

where the coefficients A and B are

$$A = \frac{k_{5\to6}}{k_{6\to5} + k_{6\to9} + k_{6\to7} + k_{6\to8}}[S] - \frac{k_{4\to1}}{k_{6\to7} - k_{6\to8}} \quad (6)$$

and

$$B = \frac{k_{9\to6}}{k_{6\to5} + k_{6\to9} + k_{6\to7} + k_{6\to8}}[S_P] + \frac{k_{2\to3}}{k_{6\to7} - k_{6\to8}}. \quad (7)$$

There are two possibilities for the relative rates of modification and demodification in the ternary complex. Either the modification rate is greater and $k_{6\to8} < k_{6\to7}$, or the demodification rate is greater and $k_{6\to7} < k_{6\to8}$. If the modification rate is greater, $B > 0$, so to satisfy eqn (5) we must have $A < 0$. This implies the following bound on [S]:

$$[S] < \frac{k_{4\to1}(k_{6\to5} + k_{6\to9} + k_{6\to7} + k_{6\to8})}{k_{5\to6}(k_{6\to7} - k_{6\to8})}. \quad (8)$$

When $k_{6\to7} < k_{6\to8}$, by an analogous argument we have the following bound on $[S_P]$:

$$[S_P] < \frac{k_{2\to3}(k_{6\to5} + k_{6\to9} + k_{6\to7} + k_{6\to8})}{k_{9\to6}(k_{6\to8} - k_{6\to7})}. \quad (9)$$

Thus, when the modification rate exceeds the demodification rate, [S] is robustly held below a threshold that is determined only by the rate constants, whereas $[S_P]$ is bounded when the demodification rate is greater.

3.1.2 Improved upper bounds using invariants. The *ad hoc* argument above is straightforward and follows directly from the underlying differential equations. We turn now to an alternative method that is less direct but reveals that the upper bounds in eqn (8) and (9) are not optimal.

The method of Gröbner bases for polynomial systems, which can be thought of as a generalization of Gaussian elimination for linear systems, can be used to obtain the simplest possible algebraic expression between specified variables that holds at steady state. We have previously used Gröbner bases to study multisite protein phosphorylation and robustness in *E. coli* isocitrate dehydrogenase regulation,[16,30] and readers interested in the mathematical details of the approach should consult the paper by Manrai and Gunawardena. From a Gröbner basis calculation on the polynomial system describing network 0 (the details of which are provided in the accompanying Mathematica notebook), it is possible to obtain the following polynomial expression in the steady-state concentrations of S and S_P:

$$(a_1[S]^2[S_P] + a_2[S][S_P]^2 + a_3[S]^2 + a_4[S_P]^2 + a_5[S][S_P] + a_6[S] + a_7[S_P])[E] = 0, \quad (10)$$

where $a_1,\ldots,a_7$ are algebraic combinations of the rate constants given in the Mathematica notebook. At this point, we restrict attention to those steady states in which each species has a positive steady-state concentration, which we call positive steady states. Since then $[E] \neq 0$, we can divide by [E] in the expression above to obtain a polynomial in just the two substrate forms:

$$a_1[S]^2[S_P] + a_2[S][S_P]^2 + a_3[S]^2 + a_4[S_P]^2 + a_5[S][S_P] + a_6[S] + a_7[S_P] = 0. \quad (11)$$

Eqn (11) is an example of what we have called in previous work an ''invariant''—a polynomial expression on selected variables that holds in any positive steady state.[30] This cubic invariant can be used to obtain robust upper bounds on [S] or $[S_P]$ by the following argument.

The invariant defines an algebraic curve, which is a locus of points in the ([S],$[S_P]$) plane that satisfy the invariant. An initial way to get a sense of the shape of such a curve is to consider its asymptotes (the tangent lines at infinity), which are given by the highest-degree terms in the polynomial. For eqn (11) the highest-degree terms are

$$a_1[S]^2[S_P] + a_2[S][S_P]^2 = 0, \quad (12)$$

which can be factored into the following linear factors:

$$[S][S_P](a_1[S] + a_2[S_P]). \quad (13)$$

These factors determine the slopes of the possible asymptotes: there can be a vertical asymptote ($[S] = c_1$), a horizontal asymptote ($[S_P] = c_2$), or a diagonal asymptote ($a_1[S] + a_2[S_P] = c_3$). Lower-order terms in the invariant must be taken into account to characterize the asymptotes fully. This quick calculation, however, is helpful to provide a sense of the possibilities. As we will see, the vertical asymptote sets the robust upper bound for [S], while the horizontal asymptote sets the bound for $[S_P]$. The existence of horizontal or vertical asymptotes for the invariant is therefore an indication of robust upper bounds. (The diagonal asymptote may be indicative of a robust ratio, which is not relevant for this network but becomes so for others.)

We now rigorously show the existence of robust upper bounds using a classical technique to study a curve by examination of its intersections with horizontal and vertical lines. If we take [S] to be constant, which corresponds to the intersection of the curve with a vertical line in the ([S],$[S_P]$) plane, then the invariant in eqn (11) reduces to a quadratic equation in $[S_P]$:

$$A[S_P]^2 + B[S_P] + C = 0, \quad (14)$$

where $A = a_2[S] + a_4$, $B = a_1[S]^2 + a_5[S] + a_7$, and $C = a_3[S]^2 + a_6[S]$. The solutions of a quadratic equation are given by its discriminant, but it is easier to consider the normalized constant term C/A. If this term is negative, the quadratic equation must have one positive and one negative solution. Inspection of the expressions for $a_1,\ldots,a_7$ in the supplementary Mathematica notebook shows that $a_3,a_6 > 0$, $a_4,a_7 < 0$, a_5 is indeterminate in sign, and the sign of both a_1 and a_2 is the same as that of $k_{6\to7} - k_{6\to8}$. Hence, if $k_{6\to7} - k_{6\to8} > 0$, we have $a_1,a_2,a_3,a_6 > 0$ and $a_4,a_7 < 0$. For any positive steady state (for which $[S],[S_P] > 0$), we have that $C/A < 0$ if, and only if, $0 < [S] < -a_4/a_2$. In this case, eqn (14) has only one positive solution for $[S_P]$.

To confirm that this expression sets an upper bound for [S] it is necessary to show that no part of the curve in the positive quadrant has $[S] > -a_4/a_2$. In this region, the normalized constant term C/A is positive and eqn (14) can have two positive solutions, two negative solutions, or two complex-conjugate solutions. We need to rule out the two positive solutions. We consider the intersection of the curve with a horizontal line, so that $[S_P]$ is taken to be constant. This procedure yields a quadratic equation in [S]:

$$A'[S]^2 + B'[S] + C' = 0, \qquad (15)$$

in which $A' = a_1[S_P] + a_3$, $B' = a_2[S_P]^2 + a_5[S_P] + a_6$, and $C' = a_4[S_P]^2 + a_7[S_P]$. Assuming again that $k_{6\to8} < k_{6\to7}$, the normalized constant term C'/A' is negative for all $[S_P] > 0$. It follows that a horizontal line intersects the curve in at most one point in the positive quadrant. Since we already know that there is such a point with $0 < [S] < -a_4/a_2$, there cannot also be another point with $[S] > -a_4/a_2$. Hence, the curve in the positive quadrant is restricted to $0 < [S] < -a_4/a_2$, which sets an upper bound for [S]. Expanding the expressions for a_4 and a_2 given in the Mathematica notebook, we find that

$$[S] < -\frac{a_4}{a_2} = \frac{k_{4\to1}(k_{6\to5} + k_{6\to7})}{k_{5\to6}(k_{6\to7} - k_{6\to8})}. \qquad (16)$$

This upper bound is strictly tighter than eqn (8). In particular, the numerator of the bound in eqn (16) does not contain the term $k_{6\to9} + k_{6\to8}$, which appears in the numerator of the bound obtained from *ad hoc* calculations. We return to this surprising discrepancy in the Discussion.

Fig. 3 shows a plot of the invariant in the ([S],[S$_P$]) plane with all parameters set to 1 except $k_{6\to7}$, which is set to 2 to ensure that $k_{6\to8} < k_{6\to7}$. The resulting curve has three disjoint lobes (blue curves). The branch that lies in the positive quadrant is confined to the left of the line $[S] = -a_4/a_2$ (red line). The upper bound in eqn (8) is shown as the green line, revealing the striking discrepancy between the two bounds. The bound obtained from the invariant describes a vertical asymptote to the curve and is thus the best possible upper bound on [S].

An analogous argument under the assumption that $k_{6\to7} < k_{6\to8}$ reveals a horizontal asymptote that gives the optimal upper bound for $[S_P]$:

$$[S_P] < -\frac{a_3}{a_1} = \frac{k_{2\to3}(k_{6\to8} + k_{6\to9})}{k_{9\to6}(k_{6\to8} - k_{6\to7})}. \qquad (17)$$

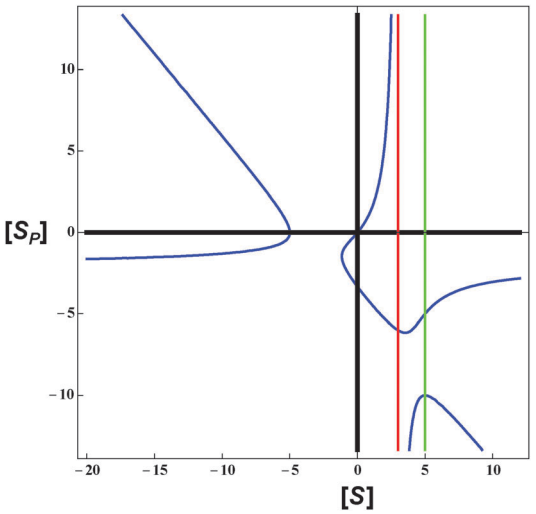


Fig. 3 Asymptotic behavior of the cubic invariant. The cubic invariant describes a curve with three distinct lobes. Only portions of the second lobe correspond to positive [S] and [S$_P$], and the vertical asymptote in the first quadrant corresponds to the robust upper bound on [S]. The vertical red line indicates the bound calculated from the invariant in eqn (16), and the vertical green line indicates the looser bound calculated in eqn (8) by using the *ad hoc* procedure. To generate the plot, all parameters except $k_{6\to7}$ were set to 1; $k_{6\to7}$ was set to 2 to ensure that [S] rather than [S$_P$] is bounded.

As before, this bound is strictly tighter than eqn (9). We see from this analysis that network 0 exhibits different behavior depending on the choice of certain parameter values.

3.2 Other sub-networks with robust upper bounds (networks i, ii, iv, and viii)

Fig. 4 shows eight related sub-networks that can be constructed by eliminating one or more reactions involving the ternary complex from the full network. Each of these networks contains all possible reactions involving binary enzyme–substrate complexes (nodes 1–4), but makes assumptions about ternary complex formation and catalytic activity. There are eight networks that meet these requirements because there are three possible binding orders for the ternary complex (random, ordered with S then S$_P$ binding to E, ordered with S$_P$ then S binding) and three possible catalytic activities (modification and demodification, modification only, demodification only). One combination (random binding, modification and demodification activity) is the full network.

The eight networks can be divided into symmetric pairs by reflection through the ternary complex (node 6). As such, networks i and ii, iii and vi, iv and viii, and v and vii in Fig. 4 are related by interchange of the roles of S and S$_P$. The algebraic analysis of one member of each pair can be applied readily to the other member. It is thus sufficient to explain the analysis of networks i, iii, iv, and v in detail. The results for all networks are summarized in Table 1.

Network i (and the symmetric network ii) have random-order binding to the ternary complex. These are the only two networks aside from the full network that yield cubic invariants. The remaining networks in Fig. 4 all have quadratic invariants. The details of the invariants are provided in the accompanying Mathematica notebook.

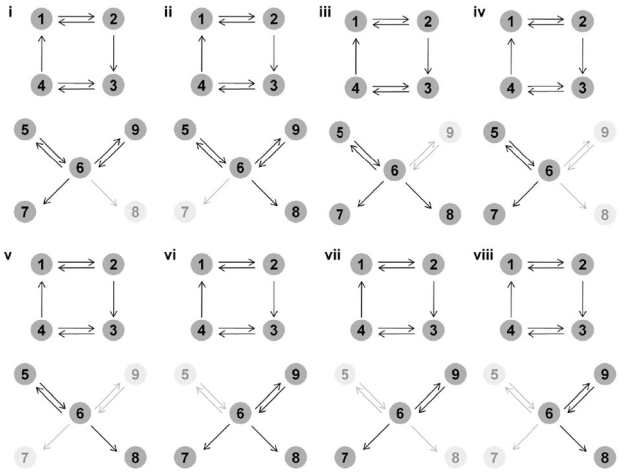


Fig. 4 Sub-networks of the full network. Reaction networks are shown for the eight sub-networks of the full network that can be obtained by eliminating one or two reactions involving the ternary complex. Excluded reactions are shown in light gray; the actual reactions in the sub-network are black. Nodes are labeled as in Fig. 2.

The cubic invariant for network i has a similar structure to that of eqn (11) (the invariant for network 0), but the pattern of signs is simpler. Using the same notation for the coefficients as in eqn (11), we see from inspection of the notebook that $a_1,a_2,a_3,a_6 > 0$, $a_4,a_7 < 0$, and a_5 is indeterminate. It follows by a similar analysis that [S] is robustly bounded above by $-a_4/a_2$:

$$[\mathrm{S}] < -\frac{a_4}{a_2} = \frac{k_{4\to1}(k_{6\to5}+k_{6\to7})}{k_{5\to6}k_{6\to7}}. \tag{18}$$

This bound, however, holds for all parameter values. In contrast, network ii has a robust upper bound on [S_P], which also holds irrespective of the values of the parameters.

Networks iii, iv, and v all have quadratic invariants, whose analysis is simpler than for the cubic invariants from networks 0, i, and ii. We turn first to network iv, which shows behavior similar to that of network i but for different reasons. Inspection of the quadratic invariant for network iv indicates that the highest-degree term is [S][S_P], suggesting the possible existence of a vertical or horizontal asymptote associated with robust upper bounds. The invariant can be rewritten in the form

$$[\mathrm{S_P}] = \frac{a_8[\mathrm{S}]}{a_9 + a_{10}[\mathrm{S}]}, \tag{19}$$

where $a_8,a_9 > 0$ and $a_{10} < 0$. The denominator of this expression becomes infinite when $[\mathrm{S}] = -a_9/a_{10} > 0$, revealing a vertical asymptote for positive values of [S]. For both [S] and [S_P] to be positive, [S] must be confined to the left of that asymptote, so that $0 < [\mathrm{S}] < -a_9/a_{10}$. As such, [S] is robustly bounded above:

$$[\mathrm{S}] < \frac{k_{4\to1}(k_{6\to5}+k_{6\to7})}{k_{5\to6}k_{6\to7}}. \tag{20}$$

The remaining networks iii and v exhibit different behaviors, which are studied in the next two sections.

3.3 Robust two-sided bounds on the concentration ratio (networks v and vii)

Network v has a quadratic invariant, whose highest-degree terms are $[\mathrm{S}]^2$ and [S][S_P]. This suggests the possible existence of a vertical asymptote and a diagonal asymptote, the former of which may have an accompanying robust upper bound. A vertical asymptote does exist, but in this case it does not imply an upper bound. Rearrangement of the quadratic invariant yields the following expression:

$$\frac{[\mathrm{S_P}]}{[\mathrm{S}]} = \frac{a_{11}+a_{12}[\mathrm{S}]}{a_{13}+a_{14}[\mathrm{S}]}, \tag{21}$$

where a_{11},a_{12},a_{13}, and a_{14}, which are given in the supplemental notebook, are all positive. The denominator of the expression on the right of eqn (21) becomes infinite at $[\mathrm{S}] = -a_{13}/a_{14} < 0$, revealing a vertical asymptote. This asymptote plays no physiological role, however, because it occurs at a negative value of [S].

The expression on the right of eqn (21) is defined and positive for all non-negative value of [S]. It has the value a_{11}/a_{13} when [S] = 0 and tends to the value a_{12}/a_{14} as [S] gets large. From calculation of the derivative, it is clear that the expression decreases monotonically as [S] increases if $a_{12}/a_{14} < a_{11}/a_{13}$, whereas the expression increases monotonically if $a_{11}/a_{13} < a_{12}/a_{14}$. In either case, the concentration ratio [S_P]/[S] is trapped between two-sided robust upper bounds. Expanding the expressions for a_{11},a_{12},a_{13}, and a_{14} we find that, if

$$k_{4\to1}(k_{2\to1}+k_{2\to3}) < k_{2\to1}(k_{4\to1}+k_{4\to3}), \tag{22}$$

then

$$\frac{a_{11}}{a_{13}} = \frac{k_{1\to2}k_{2\to3}}{k_{2\to1}k_{3\to4}} < \frac{[\mathrm{S_P}]}{[\mathrm{S}]} < \frac{k_{1\to2}k_{2\to3}(k_{4\to1}+k_{4\to3})}{k_{3\to4}k_{4\to1}(k_{2\to1}+k_{2\to3})} = \frac{a_{12}}{a_{14}}. \tag{23}$$

Similarly, if

$$k_{2\to1}(k_{4\to1}+k_{4\to3}) < k_{4\to1}(k_{2\to1}+k_{2\to3}), \tag{24}$$

then

$$\frac{a_{11}}{a_{13}} = \frac{k_{1\to2}k_{2\to3}}{k_{2\to1}k_{3\to4}} > \frac{[\mathrm{S_P}]}{[\mathrm{S}]} > \frac{k_{1\to2}k_{2\to3}(k_{4\to1}+k_{4\to3})}{k_{3\to4}k_{4\to1}(k_{2\to1}+k_{2\to3})} = \frac{a_{12}}{a_{14}}. \tag{25}$$

These results suggest that certain biochemical combinations are unsuitable for implementing a robust upper bound on one of the substrate forms. In particular, robust upper bounds are lost when the only substrate form produced from the ternary complex is required to bind second in the formation of the ternary complex (*e.g.*, in network v only S_P is produced from the ternary complex, but the binding order is S then S_P). Interestingly, we find that such conditions do lead to bounds on the concentration ratio of the substrate modification forms, a type of robustness that has not been noted previously. As such, this asymmetry appears to have important consequences for robustness.

3.4 Hybrid robustness (networks iii and vi)

We turn now to network iii. Inspection of its invariant shows that it resembles that for network v, analyzed in the previous section. It has the same terms, and the invariant can be reorganized into the same form as eqn (21). There is, however,

Table 1 Summary of robust behavior in the networks. Table 1 gives the robust upper bound, robust two-sided bound on concentration ratio, ACR constant, or robust concentration ratio for networks 0 through xi (numbered as in Fig. 2, 4 and 5), along with the corresponding parameter conditions that must be satisfied. The degree of the underlying invariant for each network is also listed. Only the tighter bound is listed for networks for which the *ad hoc* derivation and the invariant yield different results. As indicated in the Methods, the numbering and notation for parameters differs between the paper and the accompanying Mathematica notebook. The tables in the notebook should be consulted to ensure correct translation between the notebook and the paper

Net.	Robustness	Conditions	Invariant
0	$[S] < \frac{k_{4\to1}(k_{6\to5}+k_{6\to7})}{k_{5\to6}(k_{6\to7}-k_{6\to8})}$	$k_{6\to8} < k_{6\to7}$	Cubic
0	$[S_P] < \frac{k_{2\to3}(k_{6\to8}+k_{6\to9})}{k_{9\to6}(k_{6\to8}-k_{6\to7})}$	$k_{6\to7} < k_{6\to8}$	Cubic
i	$[S] < \frac{k_{4\to1}(k_{6\to5}+k_{6\to7})}{k_{5\to6}k_{6\to7}}$	None	Cubic
ii	$[S_P] < \frac{k_{2\to3}(k_{6\to8}+k_{6\to9})}{k_{6\to8}k_{9\to6}}$	None	Cubic
iii	$[S] < \frac{k_{4\to1}(k_{2\to1}+k_{2\to3})(k_{6\to5}+k_{6\to7}+k_{6\to8})}{k_{5\to6}(k_{6\to7}(k_{2\to1}+k_{2\to3})-k_{2\to1}k_{6\to8})}$	$k_{2\to1}k_{6\to8} < k_{6\to7}(k_{2\to1}+k_{2\to3})$	Quadratic
iii	$\frac{k_{1\to2}k_{2\to3}k_{6\to8}}{k_{3\to4}(k_{2\to1}k_{6\to8}-k_{6\to7}(k_{2\to1}+k_{2\to3}))} < \frac{[S_P]}{[S]} < \frac{k_{1\to2}k_{2\to3}(k_{4\to1}+k_{4\to3})}{k_{3\to4}k_{4\to1}(k_{2\to1}+k_{2\to3})}$	$k_{6\to7}(k_{2\to1}+k_{2\to3}) < k_{2\to1}k_{6\to8}$ and $k_{4\to1}k_{6\to8}(k_{2\to1}+k_{2\to3}) < (k_{4\to1}+k_{4\to3})(k_{2\to1}k_{6\to8}-k_{6\to7}(k_{2\to1}+k_{2\to3}))$	Quadratic
iii	$\frac{k_{1\to2}k_{2\to3}k_{6\to8}}{k_{3\to4}(k_{2\to1}k_{6\to8}-k_{6\to7}(k_{2\to1}+k_{2\to3}))} > \frac{[S_P]}{[S]} > \frac{k_{1\to2}k_{2\to3}(k_{4\to1}+k_{4\to3})}{k_{3\to4}k_{4\to1}(k_{2\to1}+k_{2\to3})}$	$k_{6\to7}(k_{2\to1}+k_{2\to3}) < k_{2\to1}k_{6\to8}$ and $(k_{4\to1}+k_{4\to3})(k_{2\to1}k_{6\to8}-k_{6\to7}(k_{2\to1}+k_{2\to3})) < k_{4\to1}k_{6\to8}(k_{2\to1}+k_{2\to3})$	Quadratic
iv	$[S] < \frac{k_{4\to1}(k_{6\to5}+k_{6\to7})}{k_{5\to6}k_{6\to7}}$	None	Quadratic
v	$\frac{k_{1\to2}k_{2\to3}}{k_{2\to1}k_{3\to4}} < \frac{[S_P]}{[S]} < \frac{k_{1\to2}k_{2\to3}(k_{4\to1}+k_{4\to3})}{k_{3\to4}k_{4\to1}(k_{2\to1}+k_{2\to3})}$	$k_{4\to1}(k_{2\to1}+k_{2\to3}) < k_{2\to1}(k_{4\to1}+k_{4\to3})$	Quadratic
v	$\frac{k_{1\to2}k_{2\to3}}{k_{2\to1}k_{3\to4}} > \frac{[S_P]}{[S]} > \frac{k_{1\to2}k_{2\to3}(k_{4\to1}+k_{4\to3})}{k_{3\to4}k_{4\to1}(k_{2\to1}+k_{2\to3})}$	$k_{2\to1}(k_{4\to1}+k_{4\to3}) < k_{4\to1}(k_{2\to1}+k_{2\to3})$	Quadratic
vi	$[S_P] < \frac{k_{2\to3}(k_{4\to1}+k_{4\to3})(k_{6\to7}+k_{6\to8}+k_{6\to9})}{k_{9\to6}(k_{6\to8}(k_{4\to1}+k_{4\to3})-k_{4\to3}k_{6\to7})}$	$k_{4\to3}k_{6\to7} < k_{6\to8}(k_{4\to1}+k_{4\to3})$	Quadratic
vi	$\frac{k_{1\to2}(k_{4\to3}k_{6\to7}-k_{6\to8}(k_{4\to1}+k_{4\to3}))}{k_{4\to1}k_{3\to4}k_{6\to7}} < \frac{[S_P]}{[S]} < \frac{k_{1\to2}k_{2\to3}(k_{4\to1}+k_{4\to3})}{k_{3\to4}k_{4\to1}(k_{2\to1}+k_{2\to3})}$	$k_{6\to8}(k_{4\to1}+k_{4\to3}) < k_{4\to3}k_{6\to7}$ and $(k_{2\to1}+k_{2\to3})(k_{4\to3}k_{6\to7}-k_{6\to8}(k_{4\to1}+k_{4\to3})) < k_{2\to3}k_{6\to7}(k_{4\to1}+k_{4\to3})$	Quadratic
vi	$\frac{k_{1\to2}(k_{4\to3}k_{6\to7}-k_{6\to8}(k_{4\to1}+k_{4\to3}))}{k_{4\to1}k_{3\to4}k_{6\to7}} > \frac{[S_P]}{[S]} > \frac{k_{1\to2}k_{2\to3}(k_{4\to1}+k_{4\to3})}{k_{3\to4}k_{4\to1}(k_{2\to1}+k_{2\to3})}$	$k_{6\to8}(k_{4\to1}+k_{4\to3}) < k_{4\to3}k_{6\to7}$ and $k_{2\to3}k_{6\to7}(k_{4\to1}+k_{4\to3}) < (k_{2\to1}+k_{2\to3})(k_{4\to3}k_{6\to7}-k_{6\to8}(k_{4\to1}+k_{4\to3}))$	Quadratic
vii	$\frac{k_{1\to2}k_{4\to3}}{k_{3\to4}k_{4\to1}} < \frac{[S_P]}{[S]} < \frac{k_{1\to2}k_{2\to3}(k_{4\to1}+k_{4\to3})}{k_{3\to4}k_{4\to1}(k_{2\to1}+k_{2\to3})}$	$k_{4\to3}(k_{2\to1}+k_{2\to3}) < k_{2\to3}(k_{4\to1}+k_{4\to3})$	Quadratic
vii	$\frac{k_{1\to2}k_{4\to3}}{k_{3\to4}k_{4\to1}} > \frac{[S_P]}{[S]} > \frac{k_{1\to2}k_{2\to3}(k_{4\to1}+k_{4\to3})}{k_{3\to4}k_{4\to1}(k_{2\to1}+k_{2\to3})}$	$k_{2\to3}(k_{4\to1}+k_{4\to3}) < k_{4\to3}(k_{2\to1}+k_{2\to3})$	Quadratic
viii	$[S_P] < \frac{k_{2\to3}(k_{6\to8}+k_{6\to9})}{k_{6\to8}k_{9\to6}}$	None	Quadratic
ix	$[S] = \frac{k_{4\to1}(k_{6\to5}+k_{6\to7})}{k_{5\to6}k_{6\to7}}$	None	Linear
x	$[S_P] = \frac{k_{2\to3}(k_{6\to9}+k_{6\to8})}{k_{9\to6}k_{6\to8}}$	None	Linear
xi	$\frac{[S_P]}{[S]} = \frac{k_{1\to2}k_{2\to3}(k_{4\to1}+k_{4\to3})}{k_{3\to4}k_{4\to1}(k_{2\to1}+k_{2\to3})}$	None	Linear

one significant difference. We have a_{11},a_{12},$a_{13} > 0$, but a_{14} is of indeterminate sign. We must therefore distinguish two possibilities. If $a_{14} < 0$, there is a vertical asymptote at a positive value of [S] and hence a robust upper bound with $0 < [S] < -a_{13}/a_{14}$. In this parameter region, network iii behaves like network iv. If $a_{14} > 0$, however, the vertical asymptote occurs at a negative, unphysiological value of [S]. Network iii then behaves like network v and exhibits robust two-sided bounds on the concentration ratio. (We ignore the case $a_{14} = 0$, again on the grounds that such fine tuning is unlikely to occur.) Network iii and (by symmetry) network vi therefore exhibit hybrid robustness, with different robust behaviors in different regions of the parameter space. The precise details of the parameter regions and the bounds are given in Table 1.

3.5 Absolute concentration robustness (networks ix and x)

The full network and the eight sub-networks in Fig. 4 all contain a complete set of reactions involving the binary enzyme–substrate complexes. We now consider the two networks in Fig. 5A, in which only one binary enzyme–substrate complex is formed and, accordingly, the ternary complex is formed in only one way. These networks are symmetric under interchange of S and S_P, and it is only necessary to analyze one of them. The Gröbner basis (or a simple *ad-hoc* calculation) indicates that network ix has a linear invariant of the form

$$[S] = \frac{k_{4\to1}(k_{6\to5} + k_{6\to7})}{k_{5\to6}k_{6\to7}}. \quad (26)$$

Eqn (26) shows ACR in S, which can also be deduced using the Shinar–Feinberg theorem.[12] The deficiency of the network is one, and a pair of non-terminal nodes (4 and 5) differ by only the single species S. The network thus exhibits ACR in S.

We note that very small changes in reaction network structure can mean the difference between ACR and a robust upper bound. For instance, the only difference between network ix (which has ACR) and network iv (which has a robust upper bound) is the absence of node 2.

3.6 Robust concentration ratio (network xi)

Up to this point, our focus has been on bifunctional enzymes that have separate, non-interacting active sites that are far enough apart to permit the simultaneous binding of two substrate molecules. Many regulatory bifunctional enzymes, however, have only a single active site for both activities or have active sites that are too close together to permit simultaneous binding.[22,31] Such enzymes are not able to form a ternary complex when acting on a monomeric substrate and therefore give rise to a full network that contains only four nodes (Fig. 5B). As shown in the accompanying Mathematica notebook, the following linear invariant in [S] and [S_P] can be derived using a Gröbner calculation:

$$k_{1\to2}k_{2\to3}(k_{4\to1} + k_{4\to3})[S] - k_{3\to4}k_{4\to1}(k_{2\to1} + k_{2\to3})[S_P] = 0. \quad (27)$$

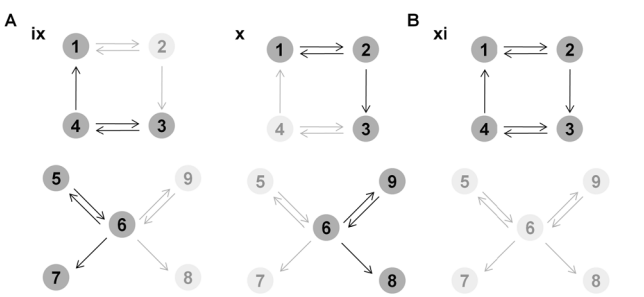


Fig. 5 Reaction networks leading to ACR and robust ratios. (A) Reaction networks that exhibit ACR in S or S_P. (B) Reaction network involving only binary complexes that exhibits a robust ratio of the substrate forms. This network is the "full network" for a bifunctional enzyme with a shared active site acting on a monomeric substrate. Excluded reactions are shown in light gray; the actual reactions in the sub-network are black. Nodes are labeled as in Fig. 2.

Rearrangement of eqn (27) yields

$$\frac{[S_P]}{[S]} = \frac{k_{1\to2}k_{2\to3}(k_{4\to1} + k_{4\to3})}{k_{3\to4}k_{4\to1}(k_{2\to1} + k_{2\to3})}. \quad (28)$$

Eqn (28) indicates that the relative amounts of modified and unmodified substrate will stay the same in all positive steady states. This ratio is a form of robustness that has not been noted before and could indicate a previously unrecognized advantage conferred by certain bifunctional enzyme systems.

3.7 Absence of ACR in networks other than ix and x

It is logically possible that a network might exhibit ACR in addition to another form of robustness such as an upper bound on concentration. We therefore checked that there is no possibility of ACR in any of the networks in the compendium besides ix and x. We find that, if a network has an invariant involving both S and S_P, then neither of these variables can exhibit ACR. For instance, if S exhibits ACR, then substituting for the absolute value of [S] in the invariant imposes a polynomial constraint on [S_P], which in turn can take only a finite number of values. However, the total amount of substrate in the system can be made arbitrarily large. Since the concentrations of enzyme–substrate complexes are limited by the total amount of enzyme, the only way for this to happen is if either [S] or [S_P] becomes arbitrarily large. This contradiction shows that S cannot exhibit ACR. Accordingly, neither the full network nor any of the networks in Fig. 4 exhibit ACR in either variable.

3.8 Experimental tests

To the best of our knowledge, there are no existing experimental demonstrations of a robust substrate ratio in a bifunctional enzyme system. As such, identification and characterization of bifunctional enzyme systems satisfying the assumptions of network xi should be of particular future interest. The two-component EnvZ/OmpR osmoregulatory system in *E. coli* does meet the basic assumptions of quadrant II (shared active site, monomeric substrate) but has been shown experimentally to exhibit ACR.[7] Imposition of additional assumptions on network xi, including initial autophosphorylation of EnvZ, can lead to networks with ACR.[7,10]

The framework developed here also offers the attractive possibility of distinguishing bifunctional enzyme mechanisms by measurements undertaken at steady state, without use of advanced biochemical or structural techniques and without the need for fitting parameter values to data. Here we outline a candidate experimental strategy. Our suggested approach involves measurement of [S] and [S_P] following incubation *in vitro* of a purified bifunctional enzyme with its substrate, under conditions in which the system can reach a steady state,[30] and testing whether the appropriate invariant is satisfied. Quantification of [S] and [S_P] can be accomplished by western blotting with antibodies specific for each substrate form or by mass spectrometry with an exogenous heavy-isotope or internal peptide standard.[32–34] Mass spectrometry has become a standard method for detection of post-translational modifications and has been used to quantify modification form distributions in several systems, including CDK1/2 and ERK.[32,35,36] Because the invariants found here do no

depend on the total amounts of enzyme or substrate, these can be chosen arbitrarily over a wide range, without the need for high precision in setting up the reaction conditions. This feature also permits the substrate to be taken in sufficient excess over the enzyme so that dissociation of the enzyme–substrate complexes, which is difficult to prevent under normal experimental conditions, does not introduce much error into the measurement of [S] and [S_P]. An advantage of bifunctional enzymes for such "systems biochemistry" is that it is not necessary to incubate two competing enzymes. Although straightforward in principle, simultaneous incubation can lead to technical complications in practice because of unexpected interactions between the two enzymes *in vitro*.

The statistical issues that arise in testing invariants against experimental data have been addressed by Harrington *et al.*,[37] who built upon our previous work.[30] Their technique involves transformation of model variables so that the locus of points satisfying the invariant lies on a hyperplane regardless of parameter values. The transformed input data, which is obtained from the measured [S] and [S_P] values, can then be deemed consistent or inconsistent with a given steady-state invariant depending on distance from this hyperplane. Such parameter-independent model assessment enables confirmation or rejection of robustness even in systems for which measurements of rate constants have not been made.

A natural first experiment would be to undertake such measurements for a well-studied system that satisfies the assumptions of the full model (network 0), such as the mammalian 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase involved in glycolytic regulation, and confirm that eqn (11) is satisfied. It might eventually be possible to make mechanistic predictions about poorly characterized bifunctional enzyme systems by measuring [S] and [S_P] and determining which invariant in the compendium holds. For instance, measurement of a robust ratio could suggest that a regulatory enzyme has a shared active site even in the absence of detailed structural data.

4 Discussion

Modification/demodification bifunctional enzymes are found in diverse cellular networks. Many of these networks regulate core transitions, such as flux partitioning at a metabolic branch point, which require coherent regulation. As such, bifunctional enzymes are used to maintain robust control of network output. Our analysis has confirmed that bifunctionality confers robustness to total enzyme levels regardless of mechanistic details. As we have discussed previously,[17,38] this robustness can have important physiological consequences, including preventing incoherent behavior in tissues due to cell-to-cell variations in protein expression levels.[39]

We have also identified several further insights, including that it is an oversimplification to think of biochemical concentration robustness as a single concept. Rather, it appears that multiple types of robust control can be implemented by biochemical reaction networks involving bifunctional enzymes. We have identified the mechanistic requirements for achieving five different types of robustness. The invariant involving [S] and [S_P] provides a general means to determine the type of robustness that a network exhibits. Furthermore, the type of robustness is highly dependent on the biochemical mechanism. Adding or removing just a single node from a reaction network can radically change the type of robustness. Additionally, a single network may exhibit different types of robustness depending on the numerical values of its parameters, as in networks iii and vi. We now consider several extensions and implications of the analysis, including the possibility of engineering biochemical robustness, treatment of non-Michaelis–Menten reaction mechanisms, and the advantages of adopting an "invariant-centric" viewpoint in theoretical biochemistry.

4.1 Engineered robustness

Section 3.8 outlines a minimal set of experiments to validate the core predictions of the invariants. Once this validation has been achieved, it might be possible to use our results to design modified or synthetic networks that exhibit a particular type of robustness. For instance, modification of the *E. coli* glyoxylate bypass regulatory machinery could yield a novel system exhibiting a robust substrate ratio. Reversible phosphorylation of IDH by the bifunctional isocitrate dehydrogenase kinase/phosphate (IDHKP) partitions flux between the glyoxylate bypass and the full TCA cycle.[6,16] *E. coli* IDHKP has a single active site for both its kinase and phosphatase activities,[22] and IDH forms an obligate homodimer,[40] which places it outside the scope of the networks in our compendium. However, mutation of one of the IDH phosphorylation sites, which potentially could be achieved by constructing a tandem dimer of one wild-type and one mutant subunit,[41] would make it a functional monomer. This modified system would be described by network xi and therefore exhibit a robust ratio (*i.e.*, constant [IDH]/[IDH_P]), whereas our previous steady-state analysis predicted different robustness for the wild-type network.[16] Biochemical analysis of the mutant system would thus provide a stringent test of a key prediction from the compendium.

4.2 Treatment of non-Michaelis–Menten reaction networks

We have assumed here that both forward and reverse modification follow the Michaelis–Menten reaction scheme. This is the standard assumption in the literature and is nearly universally followed. As we have pointed out elsewhere,[17,38,43] however, if the enzyme–substrate complex can release product, on thermodynamic grounds the product must be able to rebind the enzyme. Michaelis and Menten could assume a simplified reaction scheme because they measured reaction rates in the absence of product.[42] In the context of modification and demodification cycles, each enzyme may encounter substantial amounts of product with some attendant degree of rebinding. A more realistic scheme is

$$E + S \rightleftharpoons ES \rightarrow EP \rightleftharpoons E + P, \qquad (29)$$

which allows for product rebinding while still being irreversible overall, as would be expected under many physiological conditions for forward post-translational modification and reverse demodification. In previous work, we have developed mathematical

methods for analyzing systems with such physiologically-realistic reaction schemes, allowing for multiple intermediates and branching pathways.[17,38,43] For a bifunctional enzyme system, there is still an invariant between [S] and [S_P] even when modification and demodification take place with arbitrarily complicated reactions schemes. This invariant, however, can become very complicated, with its degree increasing with the complexity of the reaction mechanism. In particular, the invariant may no longer be cubic, and analysis of its shape by the methods used for network 0 becomes more difficult. The analysis presented here, despite being restricted to the classical Michaelis–Menten scheme, should be considered an essential first step towards the technically more challenging problem of dealing with more realistic biochemistry. We stress, moreover, that analysis of reaction schemes beyond Michaelis–Menten is still in its infancy, and the vast majority of studies in the literature continue to assume Michaelis–Menten kinetics.

4.3 Invariants in biochemical analysis

Our analysis of the reaction network compendium demonstrates that invariants can be a helpful tool for studying biochemical systems at steady state. For each of the twelve networks, the invariant in [S] and [S_P] reveals the kind of robustness that the network exhibits. Cubic invariants are found for networks 0, i, and ii, in which the ternary complex is formed by random-order binding. Quadratic invariants are found for networks iii to viii, in which the ternary complex is formed by ordered binding. Linear invariants are found for networks ix and x, in which only one of the binary complexes is formed, and for network xi, in which no ternary complex is formed. Invariants involving both [S] and [S_P] rule out the possibility of ACR but allow for robust upper bounds on concentration (networks 0, i, ii, iv and viii), robust two-sided bounds on concentration ratio (networks v and vii), hybrid behavior between these two types of robustness (networks iii and vi), or a robust concentration ratio (network xi). ACR arises only for the two networks in which the invariant involves just one of the substrate forms (networks ix and x).

ACR has been widely discussed in the literature following the pioneering insights of Shinar and Feinberg.[4] We found that asymmetry in the reactions involving the binary complexes is required for ACR. In the two networks with ACR (ix and x in Fig. 5A), the bifunctional enzyme forms a ternary complex but can only form a binary complex with S or S_P, not both. To the best of our knowledge, no known biochemical networks exhibit these characteristics. It is possible, however, that such a network structure could be implemented if each active site on the bifunctional enzyme had a massively different affinity for the substrate depending on whether the other site was already occupied. This scenario would be analogous to the avidity effect observed for bifunctional enzymes acting on multimeric substrates.[13] It might also be possible to modulate relative affinity artificially with a small molecule, suggesting the possibility that drugs could be developed to promote or interfere with robust control.

Despite the limited number of networks that give rise to ACR, ACR is a useful conceptual framework for understanding one common type of biological robustness. To that end, demonstrations of approximate ACR (*i.e.*, when the steady-state concentration of a variable is only approximately, not exactly, constant) have been made in several networks, including network iv.[11] To show approximate ACR, Shinar *et al.* assumed that the substrate is present in large excess over the bifunctional enzyme, $S_T \gg E_T$, and that total substrate is large compared to certain aggregated parameters. (They proposed network iv as a model of the IDH/IDHKP system, for which there is experimental evidence that $S_T \gg E_T$.[44,45]) In our context, approximate ACR can be obtained by substituting

$$[S_P] \approx S_T - [S], \tag{30}$$

which holds when enzyme–substrate complexes can be ignored ($S_T \gg E_T$), into eqn (19). It is then possible to show using a Taylor expansion that [S] is approximately constant when S_T is sufficiently large. The invariants we derived can therefore be used as starting points for investigating approximate ACR in many specific systems.

In addition to providing a systematic mathematical language for classifying the different types of robustness, invariants also yield tighter bounds on the robustness than we were able to obtain from *ad hoc* algebraic manipulation of the underlying differential equations (Fig. 3). We were puzzled by this unexpected discrepancy. Our best explanation for it is that the method of Gröbner bases distills all the algebraic information that is present in the underlying equations and yields the simplest possible algebraic relationship between [S] and [S_P]. It appears that *ad hoc* manipulation of the equations, while intuitively reasonable in terms of our biochemical understanding, may not take into account all the information in the system as a whole. It is possible that some sequence of manipulations not motivated by biochemical understanding might yield the optimal bounds. We were unable to find any such sequence, however, and the ability to pick out the variables of interest (in this case [S] and [S_P]) and to eliminate all others using Gröbner bases provides a powerful capability for analyzing biological systems at steady state.

The difficulty with using invariants is calculating them. The method of Gröbner bases can always be used in principle, but in practice it becomes computationally infeasible for networks that are only slightly more complicated than the ones studied here. This is particularly so when parameters are treated symbolically, as we have done. In other work, we have developed efficient methods for generating a restricted class of invariants.[15] For modification and demodification systems, such as the Goldbeter–Koshland loop, invariants can be calculated using the so-called linear framework,[46] which allows invariants to be written down irrespective of the complexity of the enzyme reaction mechanisms.[17,38] Software packages designed specifically for computational algebraic geometry, such as Singular and Macaulay2,[47,48] have also enabled efficient calculation of invariants for more complicated polynomial systems. Development of new methods for calculating invariants remains an important challenge. Efficient methods could allow for steady-state algebraic analysis of entire metabolic or

signaling networks, complementing existing probabilistic landscape approaches.[21] An ''invariant-centric'' viewpoint should enable a deeper investigation of robustness in biochemical networks than has been possible previously.

Acknowledgements

We thank Yasemin Sancak and Sudhakaran Prabakaran for comments on the manuscript, and Jeremy Owen, David Croll, and Tom Muir for scientific discussions. This work was supported by National Science Foundation Grant 0856285 (T.D. and J.G.). J.P.D. was supported in part by a National Science Foundation Graduate Research Fellowship.

References

1 H. Kitano, Towards a theory of biological robustness, *Mol. Syst. Biol.*, 2007, **3**, 137.
2 U. Alon, M. Surette, N. Barkai and S. Leibler, Robustness in bacterial chemotaxis, *Nature*, 1999, **397**, 168–171.
3 A. Larhlimi, S. Blachon, J. Selbig and Z. Nikoloski, Robustness of metabolic networks: a review of existing definitions, *BioSystems*, 2011, **106**, 1–8.
4 G. Shinar and M. Feinberg, Design principles for robust biochemical reaction networks: what works, what cannot work, and what might almost work, *Math. Biosci.*, 2011, **231**, 39–48.
5 P. Senior, Regulation of nitrogen metabolism in *Escherichia coli* and *Klebsiella aerogenes*: studies with the continuous-culture technique, *J. Bacteriol.*, 1975, **123**, 407–418.
6 D. LaPorte, P. Thorsness and D. Koshland, Compensatory phosphorylation of isocitrate dehydrogenase, A mechanism for adaptation to the intracellular environment, *J. Biol. Chem.*, 1985, **260**, 10563–10568.
7 E. Batchelor and M. Goulian, Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 691–696.
8 S. Ohta, Y. Ishida and S. Usami, High-level expression of cold-tolerant pyruvate, orthophosphate dikinase from a genomic clone with site-directed mutations in transgenic maize, *Mol. Breed.*, 2006, **18**, 29–38.
9 F. Russo and T. Silhavy, The essential tension: opposed reactions in bacterial two-component regulatory systems, *Trends Microbiol.*, 1993, **1**, 306–310.
10 G. Shinar, R. Milo, M. Martinez and U. Alon, Input-output robustness in simple bacterial signaling systems, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 19931–19935.
11 G. Shinar, J. Rabinowitz and U. Alon, Robustness in glyoxylate bypass regulation, *PLoS Comput. Biol.*, 2009, **5**, e1000297.
12 G. Shinar and M. Feinberg, Structural sources of robustness in biochemical reaction networks, *Science*, 2010, **327**, 1389–1391.
13 Y. Hart, D. Madar, J. Yuan, A. Bren, A. Mayo, J. Rabinowitz and U. Alon, Robust control of nitrogen assimilation by a bifunctional enzyme in *E. coli.*, *Mol. Cell*, 2011, **41**, 117–127.
14 Y. Hart, A. Mayo, R. Milo and U. Alon, Robust control of PEP formation in the carbon fixation pathway of C_4 plants by a bi-functional enzyme, *BMC Syst. Biol.*, 2011, **5**, 171.
15 R. Karp, M. Perez-Millan, T. Dasgupta, A. Dickenstein and J. Gunawardena, Complex-linear invariants of biochemical reaction networks, *J. Theor. Biol.*, 2012, **311**, 130–138.
16 J. Dexter and J. Gunawardena, Dimerization and bifunctionality confer robustness to the isocitrate dehydrogenase regulatory system in *Escherichia coli*, *J. Biol. Chem.*, 2013, **288**, 5770–5778.
17 T. Dasgupta, D. Croll, J. Owen, M. Vander Heiden, J. Locasale, U. Alon, L. Cantley and J. Gunawardena, A fundamental trade off in covalent switching and its circumvention by enzyme bifunctionality in glucose homeostasis, *J. Biol. Chem.*, 2014, **289**, 13010–13025.
18 R. Straube, Sensitivity and robustness in covalent modification cycles with a bifunctional converter enzyme, *Biophys. J.*, 2013, **105**, 1925–1933.
19 J. Stelling, U. Sauer, Z. Szallasi, F. J. Doyle and J. Doyle, Robustness of cellular functions, *Cell*, 2004, **118**, 675–685.
20 A. Rizk, G. Batt, F. Fages and S. Soliman, A general computational method for robustness analysis with applications to synthetic gene networks, *Bioinformatics*, 2009, **25**, i169–i178.
21 C. Li, E. Wang and J. Wang, Landscape topography determines global stability and robustness of a metabolic network, *ACS Synth. Biol.*, 2012, **1**, 229–239.
22 J. Zheng and Z. Jia, Structure of the bifunctional isocitrate dehydrogenase kinase/phosphatase, *Nature*, 2010, **465**, 961–965.
23 A. C. Ventura, P. Jiang, L. Van Wassenhove, D. Del Vecchio, S. D. Merajver and A. J. Ninfa, Signaling properties of a covalent modification cycle are altered by a downstream target, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 10032–10037.
24 P. Jiang, Y. Zhang, M. R. Atkinson and A. J. Ninfa, The robustness of the *Escherichia coli* signal-transducing UTase/UR-PII covalent modification cycle to variation in the PII concentration requires very strong inhibition of the UTase activity of UTase/UR by glutamine, *Biochemistry*, 2012, **51**, 9032–9044.
25 P. Jiang, A. C. Ventura and A. J. Ninfa, Characterization of the reconstituted UTase/UR-PII-NRII-NRI bicyclic signal transduction system that controls the transcription of nitrogen-regulated (Ntr) genes in *Escherichia coli*, *Biochemistry*, 2012, **51**, 9045–9057.
26 R. Gao and A. M. Stock, Probing kinase and phosphatase activities of two-component systems *in vivo* with concentration-dependent phosphorylation profiling, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 672–677.
27 R. Straube, Reciprocal regulation as a source of ultrasensitivity in two-component systems with a bifunctional sensor kinase, *PLoS Comput. Biol.*, 2014, **10**, e1003614.
28 Y. Hart and U. Alon, The utility of paradoxical components in biological circuits, *Mol. Cell*, 2013, **49**, 213–221.

29 S. Prabakaran, G. Lippens, H. Steen and J. Gunawardena, Post-translational modification: nature's escape from genetic imprisonment and the basis for cellular information processing, *Wiley Interdiscip. Rev.: Syst. Biol. Med.*, 2012, **4**, 565–583.
30 A. Manrai and J. Gunawardena, The geometry of multisite phosphorylation, *Biophys. J.*, 2008, **95**, 5533–5543.
31 H. Astley, K. Parsley, S. Aubry, C. Chastain, J. Burnell, M. Webb and J. Hibberd, The pyruvate, orthophosphate dikinase regulatory proteins of *Arabidopsis* are both bifunctional and interact with the catalytic and nucleotide-binding domains of pyruvate, orthophosphate dikinase, *Plant J.*, 2010, **68**, 1070–1080.
32 V. Mayya, K. Rezual, L. Wu, M. Fong and D. Han, Absolute quantification of multisite phosphorylation by selective reaction monitoring mass spectrometry, *Mol. Cell. Proteomics*, 2006, **5**, 1146–1157.
33 S. Gerber, J. Rush, O. Stemman, M. Kirschner and S. Gygi, Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 6940–6945.
34 H. Steen, J. Jebanathirajah, M. Springer and M. Kirschner, Stable isotope-free relative and absolute quantitation of protein phosphorylation stoichiometry by MS, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 3948–3953.
35 M. Mann, S.-E. Ong, M. Gronberg, H. Steen, O. Jensen and A. Pandey, Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome, *Trends Biotechnol.*, 2002, **20**, 261–268.
36 S. Prabakaran, R. Everley, I. Landrieu, J. Wieruszeski, G. Lippens, H. Steen and J. Gunawardena, Comparative analysis of Erk phosphorylation suggests a mixed strategy for measuring phospho-form distributions, *Mol. Syst. Biol.*, 2011, **7**, 482.
37 H. Harrington, K. Ho, T. Thorne and M. Stumpf, Parameter-free model discrimination criterion based on steady-state coplanarity, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 15746–15751.
38 Y. Xu and J. Gunawardena, Realistic enzymology for post-translational modification: zero-order ultrasensitivity revisited, *J. Theor. Biol.*, 2012, **311**, 136–152.
39 A. Sigal, R. Milo, A. Cohen, N. Geva-Zatorsky, Y. Klein, Y. Liron, N. Rosenfeld, T. Danon, N. Perzov and U. Alon, Variability and memory of protein levels in human cells, *Nature*, 2006, **444**, 643–646.
40 J. H. Hurley, P. E. Thorsness, V. Ramalingam, N. H. Helmers, D. E. Koshland and R. M. Stroud, Structure of a bacterial enzyme regulated by phosphorylation, isocitrate dehydrogenase, *Proc. Natl. Acad. Sci. U. S. A.*, 1989, **86**, 8635–8639.
41 R. E. Campbell, O. Tour, A. E. Palmer, P. A. Steinbach, G. S. Baird, D. A. Zacharias and R. Y. Tsien, A monomeric red fluorescent protein, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 7877–7882.
42 J. Gunawardena, Some lessons about models from Michaelis and Menten, *Mol. Biol. Cell*, 2012, **23**, 517–519.
43 J. Gunawardena, Time-scale separation: Michaelis and Menten's old idea, still bearing fruit, *FEBS J.*, 2014, **281**, 473–488.
44 A. C. Borthwick, W. H. Holms and H. G. Nimmo, Isolation of active and inactive forms of isocitrate dehydrogenase from *Escherichia coli* ML308, *Eur. J. Biochem.*, 1984, **141**, 393–400.
45 G. A. Nimmo, A. C. Borthwick, W. H. Holms and H. G. Nimmo, Partial purification and properties of isocitrate dehydrogenase kinase/phosphatase from *Escherichia coli* ML308, *Eur. J. Biochem.*, 1984, **141**, 401–408.
46 J. Gunawardena, A linear framework for time-scale separation in nonlinear biochemical systems, *PLoS One*, 2012, **5**, e36321.
47 W. Decker, G.-M. Greuel, G. Pfister and H. Schönemann, *Singular 4-0-2—a computer algebra system for polynomial computations*, http://www.singular.uni-kl.de.
48 D. R. Grayson and M. E. Stillman, *Macaulay2, a software system for research in algebraic geometry*, http://www.math.uiuc.edu/Macaulay2.

