

# Thermodynamic bounds on ultrasensitivity in covalent switching

Jeremy A. Owen,<sup>1</sup> Pranay Talla,<sup>2</sup> John W. Biddle,<sup>3</sup> and Jeremy Gunawardena<sup>3,\*</sup>

<sup>1</sup>Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts; <sup>2</sup>Horace Greeley High School, Chappaqua, New York; and <sup>3</sup>Department of Systems Biology, Harvard Medical School, Boston, Massachusetts

**ABSTRACT** Switch-like motifs are among the basic building blocks of biochemical networks. A common motif that can serve as an ultrasensitive switch consists of two enzymes acting antagonistically on a substrate, one making and the other removing a covalent modification. To work as a switch, such covalent modification cycles must be held out of thermodynamic equilibrium by continuous expenditure of energy. Here, we exploit the linear framework for timescale separation to establish tight bounds on the performance of any covalent-modification switch in terms of the chemical potential difference driving the cycle. The bounds apply to arbitrary enzyme mechanisms, not just Michaelis-Menten, with arbitrary rate constants and thereby reflect fundamental physical constraints on covalent switching.

**SIGNIFICANCE** A common motif in the chemistry of living cells consists of one enzyme that modifies another molecule and another enzyme that removes the modification. By continuously expending energy, these seemingly futile cycles can serve as sensitive switches in biological circuits, allowing cells to amplify external signals and make decisions. However, basic questions about the role of energy expenditure remain open. How much energy must living things consume to make a good switch, and does energy limit sensitivity? In this work, we answer these questions by finding the ultimate limits, set by thermodynamics, on the switching performance of modification cycles and describe the rules for making these simple nonequilibrium response elements as sensitive as possible.

## INTRODUCTION

The covalent modification cycle is a ubiquitous motif in biochemical networks. In this motif, a forward modifying enzyme,  $E$ , covalently attaches a modifying group to a substrate,  $S$ , thereby converting it from an unmodified state,  $S_0$ , to a modified state,  $S_1$ , and a reverse demodifying enzyme,  $F$ , removes the modifying group, converting  $S_1$  back to  $S_0$  (Fig. 1 *a*). Phosphorylation is one the best known forms of covalent modification, but many others are known (1–3), and new forms of modification continue to be uncovered (4). The substrate,  $S$ , may be a protein, in which case mod-

ifications are referred to as posttranslational modifications, but they may also occur on small molecules. For the modification cycles considered here, the modifying group is a small chemical moiety obtained from a donor, such as a phosphate group obtained from ATP. Polypeptide modifications, such as ubiquitin, require a more complex cascade of enzymes for covalent attachment to their substrates and fall outside the scope of this article.

The antagonistic structure of covalent modification cycles was difficult to understand at first—why simultaneously attach a modifying group and also remove it?—and led to them being referred to in the older literature as “futile cycles” (5). In fact, the forward and reverse enzymes allow the balance of  $S_1$  and  $S_0$ , measured, for instance, by the ratio of their steady-state concentrations,  $[S_1]/[S_0]$ , to be maintained away from the value it would have at chemical equilibrium. In other words, a covalent modification cycle can act as a switch (Fig. 1 *b*) in which the value of  $[S_1]/[S_0]$  is modulated by changing the levels of the forward or reverse enzymes (6,7). The idea of a biochemical switch becomes more natural in the context of cellular information processing, and such switches have been found to play key roles in

Submitted December 6, 2022, and accepted for publication April 13, 2023.

\*Correspondence: [jeremy@hms.harvard.edu](mailto:jeremy@hms.harvard.edu)

Jeremy A. Owen's present address is Department of Chemistry, Princeton University, Princeton, New Jersey.

Pranay Talla's present address is Columbia College, Columbia University, New York, New York.

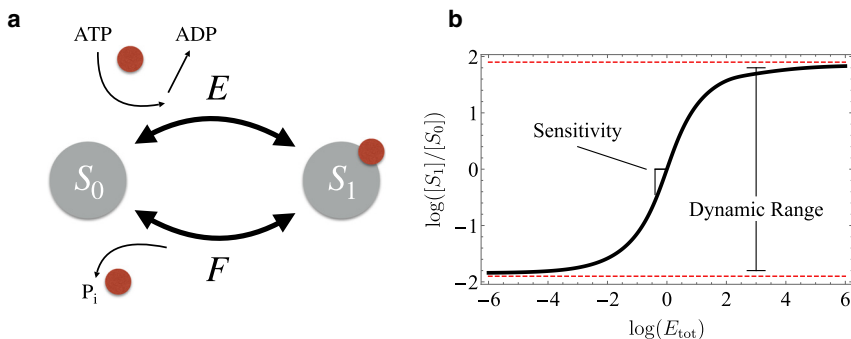
John W. Biddle's present address is Holy Cross College, Notre Dame, Indiana.

Jeremy A. Owen and Pranay Talla contributed equally to this work.

Editor: Jianhua Xing.

<https://doi.org/10.1016/j.bpj.2023.04.015>

© 2023 Biophysical Society.

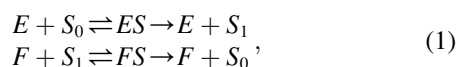


**FIGURE 1** Covalent modification cycle. (a) Schematic of a phosphorylation-dephosphorylation cycle in which kinase  $E$  modifies the substrate  $S$  by covalent addition of a phosphate group (brown disc), donated by ATP, and phosphatase  $F$  removes the modification by hydrolysis to release inorganic phosphate,  $P_i$ .  $S_0$  and  $S_1$  denote the unphosphorylated and phosphorylated forms of  $S$ . (b) Covalent modification gives rise to a switch-like relationship between the total amounts of enzymes, here of the kinase  $E$ , and the steady-state substrate concentrations, which we quantify as illustrated by the sensitivity and dynamic range. To see this figure in color, go online.

signal transduction (8–11), gene regulation (12,13), the cell cycle (14,15), and metabolism (16,17).

The operation of a covalent modification cycle relies on the continued presence of donor molecules to provide modifying groups. The cycle is driven by the chemical potential difference between the donor molecules, such as ATP, and the corresponding molecular residues after modification and demodification, such as ADP and inorganic phosphate ( $P_i$ ). This chemical potential difference is maintained by core metabolic processes within the cell. It is akin to a battery in an electronic circuit, and a modification cycle thereby operates away from thermodynamic equilibrium through continuous dissipation of energy (18,19).

In seminal work, Goldbeter and Koshland performed a mathematical analysis of covalent modification cycles under the assumption of Michaelis-Menten kinetics for the modifying enzymes—catalysis proceeds via a single intermediate enzyme-substrate complex, and product formation is irreversible (6). The chemical reactions are



where  $ES$  and  $FS$  are enzymatic intermediates. When substrate is relatively abundant so that both enzymes are saturated, they found that this system can exhibit unlimited sensitivity to the concentrations of the modifying enzymes. For example, the logarithmic sensitivity,

$$\frac{\partial \log([S_1]/[S_0])}{\partial \log E_{\text{tot}}}, \quad (2)$$

which we refer to hereafter as the sensitivity, can be made as large as desired by varying rate constants. By way of comparison, if the relationship between  $[S_1]/[S_0]$  and  $E_{\text{tot}}$  was imagined to be a Hill function,  $x^H/(1+x^H)$ , then the maximum sensitivity would be the Hill coefficient  $H$ . A sensitivity greater than 1 is said to be “ultrasensitive” (20).

The unlimited sensitivity found by Goldbeter and Koshland is physiologically implausible. It arises from the unrealistic assumption of irreversibility in the Michaelis-Menten reaction mechanism in Eq. 1. Although such an assumption has been nearly universal in quantitative studies of biochem-

ical networks, it would have surprised Michaelis and Menten (21,22), and its dangers have been repeatedly pointed out (8,23,24). In their *in vitro* studies, Michaelis and Menten measured initial reaction rates, when product was not present, so that irreversibility was a reasonable assumption. But, since then, Eq. 1 has been widely used in contexts, such as modification cycles, in which product is very much present and rebinding of product to enzyme is to be expected. Its continuing use has sometimes been justified on the grounds that modification and demodification reactions are often physiologically irreversible in the sense that product is rarely converted back into substrate. However, enzyme mechanisms typically involve greater complexity than the simple Michaelis-Menten mechanism, with multiple intermediates and routes (25), and they may be physiologically irreversible despite product rebinding. The graph-theoretic linear framework for timescale separation (26) allows realistic general reaction mechanisms to be analyzed in which such distinctions can be made. In previous work, we have analyzed modification cycles with realistic enzyme mechanisms and derived formulas for their switching capability in the limit of high substrate (16,27). The switching sensitivity is no longer unbounded but is now limited by the parameters of the switch.

A further difficulty with the irreversibility of Eq. 1 is that it implies infinite entropy production. In reality, every reaction is reversible, although physiological conditions may make the reverse rate extremely low. As we will see below, it is the ratio of forward to reverse rates that yields the finite rate of entropy production. The Michaelis-Menten mechanism is therefore unsuitable for a thermodynamic analysis. Noting this, Qian (19) studied a minimal elaboration of the Goldbeter-Koshland cycle in which both enzymes follow a fully reversible Michaelis-Menten mechanism. This mechanism still proceeds via a single intermediate complex, but product can rebound. Qian found the relationship between the chemical potential difference driving the cycle and figures of merit, such as the sensitivity, of a switch based on the system.

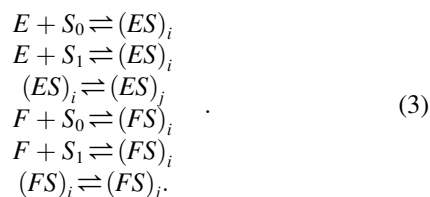
In the linear framework, realistic enzyme mechanisms can be analyzed with reversibility assumed throughout. We use this approach here to establish bounds on the switching dynamic range and sensitivity (Fig. 1 b) of any covalent

modification cycle in which the forward and reverse enzyme each follow their own realistic enzyme mechanism. We also show explicitly that these thermodynamic bounds can be approached as closely as desired. Our work generalizes the analysis of Goldbeter and Koshland (6,18) and the subsequent work of Qian (19,28) and reveals fundamental physical constraints free from restrictive enzymological assumptions.

## RESULTS

### Covalent modification cycles

For the present purposes, a covalent modification cycle is any chemical reaction network with mass action kinetics, built out of any number of reactions of the form (27,29)



Any such network can be viewed as a detailed realization of the schematic modification cycle illustrated in Fig. 1 *a*. This class of models encompasses the irreversible cycle studied by Goldbeter and Koshland but also much more biochemically realistic ones reflecting complex enzymology, with any number of intermediates, and, in particular, the reversibility of enzymes.

The reactions in Eq. 3 imply the conservation of the total concentration of substrate  $S_{\text{tot}}$ :

$$S_{\text{tot}} = [S_0] + [S_1] + \sum_i [(ES)_i] + \sum_j [(FS)_j], \quad (4)$$

as well as the total concentrations of both enzymes,  $E_{\text{tot}}$  and  $F_{\text{tot}}$ :

$$\begin{aligned} E_{\text{tot}} &= [E] + \sum_i [(ES)_i] \\ F_{\text{tot}} &= [F] + \sum_j [(FS)_j]. \end{aligned} \quad (5)$$

The assumption of mass-action kinetics gives a system of polynomial differential equations for the time evolution of the concentration of each chemical species. The equations are arrived at by summing the individual contributions to the rate of formation/destruction of each species due to each reaction.

Given any fixed choice of rate constants and the conserved quantities  $S_{\text{tot}}$ ,  $E_{\text{tot}}$ , and  $F_{\text{tot}}$ , a covalent modification cycle admits a unique steady state or dynamical fixed point (16). This fact allows us to view steady-state quantities, such as the steady-state value of the ratio  $[S_1]/[S_0]$ , as functions of the conserved quantities. For a general covalent modification cycle,

the polynomial equations satisfied at steady state, implicitly defining this functional relationship, can be very complicated—having arbitrarily many terms and rate constants appearing in them. Nevertheless, they possess a basic structure, set by the schema in Eq. 3, that will enable us to apply a powerful algebraic approach—the linear framework—to make general statements about them.

### Background on the linear framework

Here, we briefly introduce the linear framework. In the following subsection, we will discuss the example of covalent modification cycles. The framework was introduced in (26). It provides a unified graph-theoretic approach to analyzing those molecular systems that have been previously studied by ad hoc methods of timescale separation, including protein allostery, receptor pharmacology, and gene regulation, as reviewed in (30). A more recent review (31) discusses the material needed for the present article and should be consulted for further details and background.

The framework centers on finite, simple, directed graphs with labeled edges; an example graph is shown in Fig. 2 *a*. For the application considered here, the graph vertices, denoted by  $1, 2, \dots, n$ , represent chemical species; the edges, denoted  $i \rightarrow j$ , represent reactions; and the labels, denoted  $\ell(i \rightarrow j)$ , represent positive rates with dimensions of  $(\text{time})^{-1}$ . The labels may be algebraic expressions that include time-varying concentrations of chemical species. We will discuss such labels further in the next section; in the present, to explain the machinery of the framework, we regard the labels as constant symbols.

Let  $G$  be a linear framework graph. Such a graph can be naturally given a dynamics by assuming that each edge is a chemical reaction with the corresponding label as the rate constant for mass-action kinetics. Since an edge has only a single source vertex, the dynamics must be linear and is therefore described by a matrix differential equation,

$$\frac{dx}{dt} = \mathcal{L}(G) \cdot x, \quad (6)$$

where  $x = (x_1, \dots, x_n)^T \in \mathbb{R}^n$  is the time-dependent column vector of vertex concentrations (here,  $^T$  denotes transpose) and the linear operator,  $\mathcal{L}(G)$ , is the Laplacian matrix of  $G$ . Since material is only moved around the graph, without being created or destroyed, there is a conservation law: at all times,  $x_1 + \dots + x_n = x_{\text{tot}}$ ; equivalently, each column of  $\mathcal{L}(G)$  sums to zero.

We will be concerned with steady states,  $x^*$ , of systems described by graphs so that  $(dx/dt)|_{x=x^*} = 0$ . Accordingly,  $x^* \in \ker \mathcal{L}(G)$ . It can be shown that if  $G$  is strongly connected, then  $\dim \ker \mathcal{L}(G) = 1$ . Recall that a graph is strongly connected if any two distinct vertices,  $i$  and  $j$ , can be connected by a directed path,  $i = i_1 \rightarrow i_2 \rightarrow \dots \rightarrow i_k = j$ . A canonical basis element  $\rho(G) \in \ker \mathcal{L}(G)$  may be determined in terms of the edge labels by using the matrix-tree

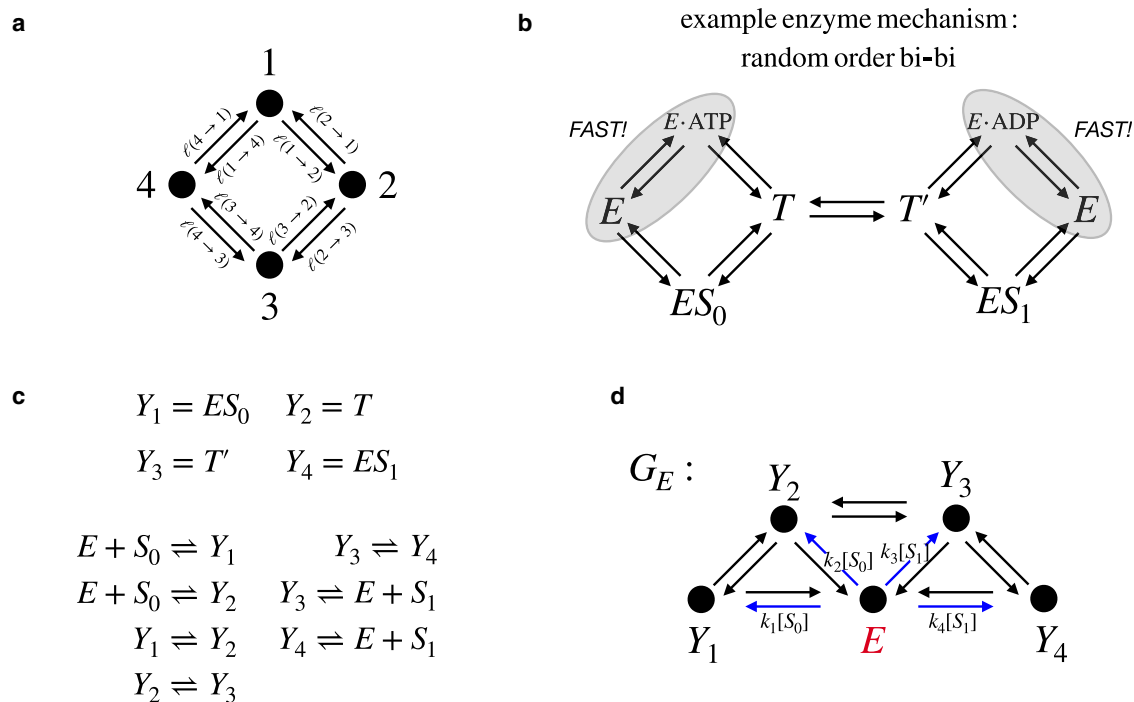


FIGURE 2 Modeling enzyme mechanisms using the linear framework. (a) An example of a linear framework graph. (b) An example of a realistic enzyme mechanism, in which two substrates (ATP and  $S_0$ ) bind to an enzyme  $E$  in any order to form a ternary complex  $T$ , which is transformed into a complex  $T'$  from which two products (ADP and  $S_1$ ) are released in any order. To be able to model this mechanism using the grammar of Eq. 9, the binding/unbinding of at least one of the substrates and one of the products to the bare enzyme must come to a rapid equilibrium (denoted by the gray ovals). (c) Subject to this timescale assumption, the network in (b) can be cast in terms of the grammar of Eq. 9. The details are explained in (16). For example, the reactions between  $E \cdot \text{ATP}$  and  $T$  are incorporated into the reactions between  $E + S_0$  and  $Y_2$ , with the rate of the forward reaction carrying the concentration of ATP. (d) A linear framework graph  $G_E$  corresponding to the mechanism shown in (c). Edges whose labels include concentration terms are shown in blue. To see this figure in color, go online.

theorem (MTT) of graph theory (32–35). If  $H$  is any subgraph of  $G$ , let  $w(H)$  denote the product of the edge labels over the edges in  $H$ ,

$$w(H) = \prod_{i \rightarrow j \in H} \ell(i \rightarrow j).$$

Recall that a spanning tree is a connected subgraph of  $G$  that contains each vertex of  $G$  (spanning) and has no cycles if edge directions are ignored (tree). It is said to be rooted at  $i$  if  $i$  is the only vertex with no outgoing edge (which orients the tree). Let  $\Theta_i(G)$  denote the set of spanning trees of  $G$  that are rooted at  $i$ . Then, the MTT shows that

$$\rho_i(G) = \sum_{T \in \Theta_i(G)} w(T). \quad (7)$$

Since  $x^*$  must be proportional to  $\rho(G) \in \ker \mathcal{L}(G)$ , it is straightforward to obtain the ratio of steady states in terms only of the edge labels,

$$\frac{x_i^*}{x_j^*} = \frac{\sum_{T_i \in \Theta_i(G)} w(T_i)}{\sum_{T_j \in \Theta_j(G)} w(T_j)}, \quad (8)$$

and we will exploit this below.

### Modeling a covalent modification cycle

In previous work, posttranslational modification systems, like the covalent modification cycle of Fig. 1 a, were modeled as interacting systems of linear framework graphs (16,27). This approach is also reviewed in (31). An important feature of the approach is that enzyme reaction mechanisms can be substantially more general than the conventional Michaelis-Menten mechanism in Eq. 1, allowing, in particular, for reversibility and multiple intermediate complexes and thereby addressing the problems described in the introduction. Specifically, an enzyme mechanism may be composed of any appropriate reactions from the “grammar,”



which transforms between substrate and product. Here, we have used “ $*$ ” as a generic subscript to avoid index proliferation.  $S_*$  denotes a substrate form, such as  $S_0$  or  $S_1$  in Eq. 1, and  $Y_*$  denotes an intermediate complex, such as  $ES$  in Eq. 1. The only requirement we place on the mechanism is that there must be no “dead-end” intermediate complexes that can be formed but not destroyed. We will see that this amounts to supposing that

a certain corresponding graph, described below, is strongly connected.

The Michaelis-Menten mechanism in Eq. 1, or the reversible version used by Qian in (19), can be constructed from the grammar in Eq. 9. But the grammar can also capture more complicated mechanisms, such as the “random order bi-bi” mechanism (36) pictured in Fig. 2 *b* in which an enzyme has two substrates that are bound in either order and forms two products that are released in either order. Such mechanisms are important for forward modifying enzymes, like kinases, which use a secondary substrate, like ATP, to donate the modifying group (Fig. 1 *a*) and release a secondary product, like ADP. Similarly, the enzyme mechanisms implied by the covalent modification cycle in Eq. 3 can also be accommodated in Eq. 9.

In order to cast realistic enzyme mechanisms in our grammar, we must make an approximation because, in Eq. 9, substrates are not permitted to bind to intermediate complexes, so secondary substrates cannot be explicitly represented. Instead, as explained in more detail in (27), we will assume that secondary-substrate binding occurs as a “rapid equilibrium” with either the free enzyme  $E$  or some intermediate complex  $Y_*$  so that the concentration of the secondary substrate can be absorbed into the appropriate rate constant. We additionally assume that the concentration of this secondary substrate is held constant—a reasonable assumption in the case of many forward modifying enzymes, especially kinases, where ATP concentration is held constant by background metabolic processes despite fluctuations in demand. Subject to these assumptions, the mechanism in Fig. 2 *b* can be well approximated by the reaction network shown in Fig. 2 *c*, which can be specified by the grammar in Eq. 9.

Any mechanism for an enzyme  $E$  expressed in the grammar may be represented by a linear framework graph  $G_E$  (Fig. 2 *d*), in which the vertices are the free enzyme  $E$  and the intermediate complexes  $Y_*$ , the edges correspond to the reactions in the mechanism, and the labels correspond to the reaction rates. The time-dependent concentrations of the substrate forms  $S_*$  appear in the labels for those edges in which a substrate form binds to the free enzyme. Accordingly, as can be seen in the example in Fig. 2 *d*, only those edges outgoing from the vertex  $E$ , which are colored blue, have edge labels depending on the substrate concentrations. In keeping with the reversibility of the enzyme mechanism in Fig. 2 *c*, every transition in the graph  $G_E$  in Fig. 2 *d* is reversible. We will always require that  $G_E$  be strongly connected and in particular that no intermediates form irreversibly.

The linear Laplacian dynamics on  $G_E$ , as given by Eq. 6, is merely a rewriting of the dynamics of the enzyme mechanism under mass-action kinetics. But this construction will allow us to apply the MTT, in the form of Eq. 8, to algebra-

ically express the variables associated with the vertices in terms of the edge labels.

The general form of the covalent modification cycle in Eq. 3 can easily be constructed within the grammar in Eq. 9: it consists of two enzymes,  $E$  and  $F$ , with arbitrary mechanisms obeying the grammar of Eq. 9, and two interconverting substrate forms,  $S_0$  and  $S_1$ . The reverse modifying enzyme,  $F$ , may follow a different mechanism from that of  $E$ . This is typically the case in reality, as the removal of a modifying group is often a single-substrate hydrolysis reaction. The ability to realistically represent the difference between the mechanisms of  $E$  and  $F$  is an important benefit of our approach.

Applying the linear framework to a covalent modification cycle, we get two graphs:  $G_E$ , whose vertices are  $E$  and its intermediates, and  $G_F$ , whose vertices are  $F$  and its intermediates. Since  $[S_0]$  and  $[S_1]$  appear in the labels of the edges of  $G_E$  and  $G_F$  directed out of  $E$  and  $F$ , respectively, they will also appear on the right hand side of Eq. 8, but they do so in a limited way. Specifically, since each rooted spanning tree has at most one edge directed out of any vertex, they appear only linearly. This observation leads to the following simple relations at steady state (27):

$$\sum_i \frac{[(ES)_i]}{[E]} = \frac{[S_0]}{K_0^E} + \frac{[S_1]}{K_1^E}, \quad (10)$$

$$\sum_j \frac{[(FS)_j]}{[F]} = \frac{[S_0]}{K_0^F} + \frac{[S_1]}{K_1^F},$$

where  $K_0^E, K_1^E, K_0^F$ , and  $K_1^F$  are the total generalized Michaelis-Menten constants (tgMMCs) of the covalent modification cycle (27,29). The tgMMCs depend only on rate constants and have dimensions of concentration. The specific algebraic expressions for the tgMMCs in terms of the rate constants depend on the choice of enzyme mechanism from the grammar in Eq. 9. An informal explanation of how the tgMMCs are calculated is given in (31).

Substituting expressions like Eq. 10 into the original polynomial steady-state equations yields a similarly compact expression for the substrate ratio (Eq. 13; (27)):

$$\frac{[S_1]}{[S_0]} = \frac{c_0^E[E] + c_0^F[F]}{c_1^E[E] + c_1^F[F]}, \quad (11)$$

where the quantities  $c_0^E, c_1^E, c_0^F$ , and  $c_1^F$  are the total generalized catalytic efficiencies (tgCEs) of the covalent modification cycle (27,29). These quantities depend only on rate constants and have dimensions of  $(\text{concentration} \times \text{time})^{-1}$ . As with the tgMMCs, the specific algebraic expressions for the tgCEs in terms of the rate constants depend on the choice of enzyme mechanism from the grammar in Eq. 9. An informal explanation of how the tgCEs are calculated is given in (31).



The three conservation laws and the relations of Eqs. 10 and 11 can now be combined to yield two equations involving only  $[S_0]$  and  $[S_1]$  as variables:

$$S_{\text{tot}} = [S_0] + [S_1] + E_{\text{tot}} \left( \frac{[S_0]/K_0^E + [S_1]/K_1^E}{1 + [S_0]/K_0^E + [S_1]/K_1^E} \right) + F_{\text{tot}} \left( \frac{[S_0]/K_0^F + [S_1]/K_1^F}{1 + [S_0]/K_0^F + [S_1]/K_1^F} \right), \text{ and} \quad (12)$$

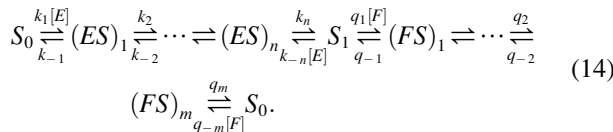
$$\left( \frac{E_{\text{tot}}}{F_{\text{tot}}} \right) \left( 1 + \frac{[S_1]}{K_1^E} + \frac{[S_0]}{K_0^E} \right) (c_0^E [S_0] - c_1^E [S_1]) = \left( 1 + \frac{[S_0]}{K_0^F} + \frac{[S_1]}{K_1^F} \right) (c_1^F [S_1] - c_0^F [S_0]). \quad (13)$$

For the purposes of understanding the steady-state dependence of  $[S_0]$  and  $[S_1]$  on the conserved quantities, all the possible complexity permitted by the schema in Eq. 3, and all the freedom to choose rate constants, have been reduced to eight generalized parameters—the four tgMMCs and the four tgCEs.

### Thermodynamic constraints on the tgCEs

Now, we turn to thermodynamics. We will assume from now on that every graph is reversible, in the sense that if  $i \rightarrow j$ , then also  $j \rightarrow i$ . It is also important that the reverse edge  $j \rightarrow i$  represents the process that is the time reverse of that represented by  $i \rightarrow j$  and not just some other process for returning from  $j$  to  $i$  (37).

At thermodynamic equilibrium, each reaction occurs with the same frequency as its time reverse. This principle of detailed balance (38) implies a relation between the rate constants of reactions that form a cycle. Consider, for example, an arbitrary cycle of reactions in a covalent modification cycle taking  $S_0$  to  $S_1$  via the enzyme  $E$  and back to  $S_0$  via the the enzyme  $F$ :



For this cycle, detailed balance at thermodynamic equilibrium implies that

$$\frac{k_1 k_2 \dots q_2 q_m}{k_{-1} k_{-2} \dots q_{-2} q_{-m}} = 1. \quad (15)$$

Recall that some of the rate constants  $k_i, q_i$  appearing above may conceal concentrations of cofactors we do not explicitly model. Holding these concentrations at fixed values away from their equilibrium values drives the system out of equilibrium into a nonequilibrium steady state. The log ratio of rates on the left hand side of Eq. 15—known as the thermodynamic force, or cycle affinity—quantifies the breaking of detailed balance and can often be identified

as the entropy produced in the environment when the cycle is traversed in units of the Boltzmann constant,  $k_B$  (39–41). For a chemical system held out of equilibrium by the presence of species assumed to have fixed concentration (either because there is a very large reservoir of them or because a fixed concentration is actively maintained), this will depend on a chemical potential difference  $\Delta\mu$ . For example, suppose  $E$  is a kinase and  $F$  a phosphatase so that one complete realization of the cycle entails the binding of one molecule of ATP, the release of one molecule of ADP, and the release of one molecule of  $P_i$ . Then,

$$\log \left( \frac{k_1 k_2 \dots q_2 q_m}{k_{-1} k_{-2} \dots q_{-2} q_{-m}} \right) = \frac{\mu_{\text{ATP}} - \mu_{\text{ADP}} - \mu_{P_i}}{k_B T} = \frac{\Delta\mu}{k_B T} \approx 20 \text{ to } 30 \quad (16)$$

under typical physiological conditions. In this work, we focus on the natural case in which the thermodynamic force is the same around any cycle in which  $E$  makes the modification of  $S$  and  $F$  removes it. In that case, we can show that

$$\log \left( \frac{c_0^E c_1^F}{c_1^E c_0^F} \right) = \frac{\Delta\mu}{k_B T}. \quad (17)$$

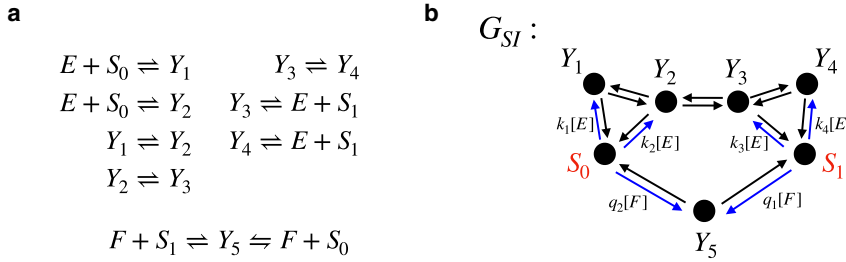
In the general case, where the force is different around different cycles, the lefthand side is bounded by the largest force. Without loss of generality, we take  $c_0^F/c_1^F < c_0^E/c_1^E$  so that  $\Delta\mu > 0$ .

Equation 17 is the physical constraint on the parameters of a general covalent modification cycle from which all our results descend. To prove it, it will be convenient to consider another kind of linear framework graph, denoted  $G_{SI}$ , whose vertices are the substrate forms and the intermediate complexes.  $G_{SI}$  is an amalgam of  $G_E$  and  $G_F$  together with the substrate forms  $S_0$  and  $S_1$  and it is more convenient for a thermodynamic analysis of the cycle ( $G_{SI}$  is distinct from the graph on the substrate forms introduced in prior work involving the linear framework (27)).  $G_{SI}$  is shown for an example covalent modification cycle in Fig. 3.  $G_{SI}$  is strongly connected as long as  $G_E$  and  $G_F$  are strongly connected, as we assume.

To prove Eq. 17, we begin with Eq. 11 from above, which says that

$$\frac{[S_1]}{[S_0]} = \frac{c_0^E [E] + c_0^F [F]}{c_1^E [E] + c_1^F [F]} = \frac{c_0^E ([E]/[F]) + c_0^F}{c_1^E ([E]/[F]) + c_1^F}. \quad (18)$$

An alternative expression for this ratio can be obtained from the MTT applied to the graph  $G_{SI}$ . Since  $G_{SI}$  is assumed to be reversible, it is helpful—to organize the expressions that arise from the MTT—to consider undirected spanning trees in the corresponding undirected graph  $G_{SI}^u$ , which is  $G_{SI}$ , but with the directions of the edges ignored. Given any vertex  $i$ , any spanning tree  $T$  of  $G_{SI}^u$  can be uniquely “lifted” to a directed spanning tree  $T_i$  rooted at  $i$ ,



**FIGURE 3** An example covalent modification cycle and the graph  $G_{SI}$ . (a) A covalent modification cycle where the enzyme  $E$  obeys the random order bi-bi mechanism from Fig. 2 and the enzyme  $F$  acts by a reversible Michaelis-Menten mechanism. (b) The associated linear framework graph  $G_{SI}$  involving the substrates and the intermediates. Edges whose labels include concentration terms are shown in blue. To see this figure in color, go online.

and all rooted spanning trees arise in this way. This means, in particular, that there is a bijection between  $\Theta_i(G_{SI})$  and  $\Theta_j(G_{SI})$  for any two distinct vertices  $i$  and  $j$ . As explained further in (42), this bijection can be made explicit by taking any rooted spanning tree  $T \in \Theta_i(G_{SI})$  and reversing the direction of the edges on the unique directed path in  $T$  from  $j$  to the root  $i$ , to give a rooted spanning tree  $T' \in \Theta_j(G_{SI})$ . Fig. 4 c illustrates this construction between the spanning trees  $T_{S_0}$  and  $T_{S_1}$ . Notice that, as a consequence of this transformation, the edges of  $T$  that are not on the unique directed path from  $j$  to  $i$  are also found in  $T'$ .

Applying the MTT to the spanning trees arising from  $G_{SI}$ , we get

$$\frac{[S_1]}{[S_0]} = \frac{\sum_T w(T_{S_1})}{\sum_T w(T_{S_0})}, \quad (19)$$

where the sum is over the spanning trees of  $G_{SI}^u$  and  $T_{S_1}$  and  $T_{S_0}$  are the directed spanning trees rooted at  $S_1$  and  $S_0$ , respectively.

Every spanning tree of  $G_{SI}^u$  contains a unique path between  $S_0$  and  $S_1$ —if there were multiple paths, it would fail to be a tree, and if there were no path, the tree would fail to include every vertex (it would fail to be “spanning”). Let  $U$  be the set of spanning trees where the path between  $S_0$  and  $S_1$  involves intermediates containing  $E$ . Let  $V$  be the set of spanning trees where the path between  $S_0$  and  $S_1$  involves intermediates containing  $F$ . Every spanning tree of  $G_{SI}^u$  lies in either  $U$  or  $V$ , never both (Fig. 4 b).

Furthermore, if  $T \in U$ , the weights  $w(T_{S_0})$  and  $w(T_{S_1})$  must be linear in  $[E]$  because the directed trees must leave  $S_1$  (resp.  $S_0$ ) by exactly one edge, and it must be an edge carrying  $[E]$  on its label since  $T$  lies in  $U$ . Similarly, if  $T$  lies in  $V$ , the weights  $w(T_{S_0})$  and  $w(T_{S_1})$  must be linear in  $[F]$ .

Consequently, we can write

$$\begin{aligned}
 \frac{[S_1]}{[S_0]} &= \frac{\sum_{T \in U} w(T_{S_1}) + \sum_{T \in V} w(T_{S_1})}{\sum_{T \in U} w(T_{S_0}) + \sum_{T \in V} w(T_{S_0})} \\
 &= \frac{A[E] + B[F]}{C[E] + D[F]} = \frac{A([E]/[F]) + B}{C([E]/[F]) + D}, \quad (20)
 \end{aligned}$$

where  $A$ ,  $B$ ,  $C$ , and  $D$  are positive constants that do not depend on  $[E]$  or  $[F]$ . We also have

$$\frac{AD}{BC} = \frac{A[E]D[F]}{C[E]B[F]} = \frac{\sum_{T \in U} w(T_{S_1}) \sum_{T \in V} w(T_{S_0})}{\sum_{T \in U} w(T_{S_0}) \sum_{T \in V} w(T_{S_1})}. \quad (21)$$

When comparing Eq. 20 with Eq. 18, it is worth noting that we cannot simply conclude that the coefficients of the expressions on the righthand sides are equal. To proceed, we need the following claim.

*Claim 1*

Suppose  $a_1, b_1, c_1, d_1, a_2, b_2, c_2$ , and  $d_2$  are nonzero real numbers such that

$$\frac{a_1x + b_1}{c_1x + d_1} = \frac{a_2x + b_2}{c_2x + d_2} \quad (22)$$

for all  $x$ . Then,

$$\frac{a_1d_1}{b_1c_1} = \frac{a_2d_2}{b_2c_2}. \quad (23)$$

*Proof*

Equation 22 implies that

$$(a_1x + b_1)(c_2x + d_2) = (a_2x + b_2)(c_1x + d_1). \quad (24)$$

Expanding yields  $(a_1c_2 - a_2c_1)x^2 + (a_1d_2 - a_2d_1 + b_1c_2 - b_2c_1)x + (b_1d_2 - b_2d_1) = 0$ . Since this holds for all  $x$ , we find that  $b_1d_2 = b_2d_1$  and  $a_1c_2 = a_2c_1$ . This means that

$$1 = \frac{b_1d_2}{b_2d_1} = \frac{a_1c_2}{a_2c_1}, \quad (25)$$

from which the desired result follows by rearrangement.

Equating Eqs. 18 and 20 and applying the claim yields

$$\frac{c_0^E c_1^F}{c_1^E c_0^F} = \frac{\sum_{T \in U} w(T_{S_1}) \sum_{T \in V} w(T_{S_0})}{\sum_{T \in U} w(T_{S_0}) \sum_{T \in V} w(T_{S_1})}. \quad (26)$$

Now, let  $T$  be a tree of  $G_{SI}^u$ , and suppose that  $P$  is the unique directed path from  $S_0$  to  $S_1$  in  $T_{S_1}$ . Let  $P^*$  denote the reverse directed path from  $S_1$  to  $S_0$ . As noted previously, the edges that do not lie on  $P$  in  $T_{S_1}$ , and the edges that do not lie on  $P^*$  in  $T_{S_0}$  are identical (Fig. 4 c). It therefore follows that  $w(T_{S_1})/w(T_{S_0}) = w(P)/w(P^*)$ . By assuming that the thermodynamic force is the same about any cycle in our reaction network, this quantity depends only on whether  $T$  lies in  $U$  or  $V$ .

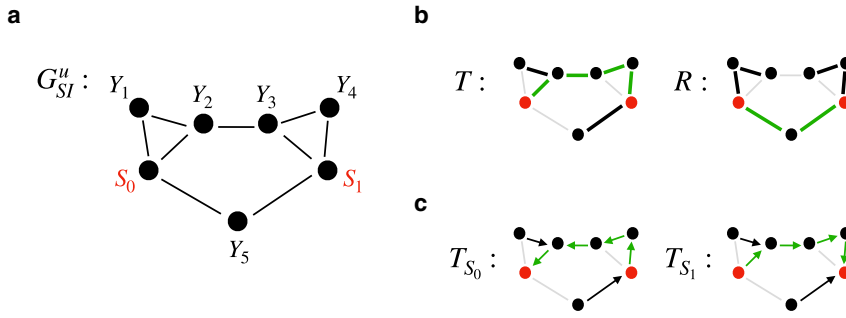


FIGURE 4 Lifting the undirected spanning trees of  $G_{S_I}^u$ . (a) The graph  $G_{S_I}^u$  associated with the covalent modification cycle of Fig. 3, together with (b) two example spanning trees of  $G_{S_I}^u$ ,  $T$  and  $R$ . In each tree, the unique path between  $S_0$  and  $S_1$  is indicated in green. In  $T$ , the intermediates on the unique path are in  $E$  so that  $T \in U$ ; in  $R$ , the intermediates are in  $F$  so that  $R \in V$ . (c) The directed spanning trees  $T_{S_0}$  and  $T_{S_1}$  rooted at  $S_0$  and  $S_1$ , respectively. Note how the unique directed path between  $S_0$  and  $S_1$  is reversed to convert between  $T_{S_0}$  and  $T_{S_1}$ , as explained in the text. To see this figure in color, go online.

Consequently,

$$\frac{c_0^E c_1^F}{c_1^E c_0^F} = \frac{w(P)w(Q)}{w(P^*)w(Q^*)}, \quad (27)$$

where  $P$  is any directed path from  $S_0$  to  $S_1$  through the intermediates containing  $E$ ,  $Q$  is any directed path from  $S_1$  to  $S_0$  through the intermediates containing  $F$ , and  $P^*$  and  $Q^*$  are the respective reverses of those paths.

Accordingly,

$$\log\left(\frac{c_0^E c_1^F}{c_1^E c_0^F}\right) = \log\left(\frac{w(P)w(Q)}{w(P^*)w(Q^*)}\right) = \log\left(\frac{w(C)}{w(C^*)}\right), \quad (28)$$

where  $C$  is the directed cycle in  $G_{S_I}$  formed by the concatenation of  $P$  and  $Q$ . But the term  $\log\left(\frac{w(C)}{w(C^*)}\right)$  is exactly a log ratio of rates about a cycle, as we considered in our physical discussion above. It is therefore equal to the thermodynamic force, or chemical potential difference  $\Delta\mu/k_B T$ , holding the system out of equilibrium. This proves Eq. 17.

### The dynamic range

The groundwork we have laid now allows us to establish bounds—in terms of  $\Delta\mu/k_B T$ —on the characteristics of a switch based on any covalent modification cycle, taking the “output” variable of the switch to be the logarithm of steady-state ratio  $[S_1]/[S_0]$  and “input” variables to be the conserved enzyme totals.

The simplest characteristic of a switch is the difference between the largest and smallest values its output variable can assume—the dynamic range. From Eq. 18, we have

$$\log\left(\frac{c_0^F}{c_1^F}\right) < \log\left(\frac{[S_1]}{[S_0]}\right) < \log\left(\frac{c_0^E}{c_1^E}\right), \quad (29)$$

and these limits can be approached, the larger when  $F_{\text{tot}} \rightarrow 0$  and the smaller when  $E_{\text{tot}} \rightarrow 0$ . The dynamic range of the switch is then the difference of these extremes, which, by Eq. 17, is simply equal to  $\Delta\mu/k_B T$ . In the special case where each enzyme forms only one intermediate enzyme-substrate

complex, this thermodynamic bound on the dynamic range was found by Qian (19).

### The high substrate regime

We now turn to the sensitivity of a switch—how sharply can the output respond to a small change in the input? Our main result is that in the limit where  $S_{\text{tot}}$  is very large compared with the enzyme totals  $E_{\text{tot}}$  and  $F_{\text{tot}}$  and the generalized Michaelis-Menten constants, the sensitivity is bounded by a simple function of the switch parameters. To be precise, the ratio  $[S_1]/[S_0]$  has a well-defined limit as  $S_{\text{tot}} \rightarrow \infty$ . It is this limiting value, which depends on the rate constants and conserved enzyme totals, whose sensitivity we bound in the following result,

$$\left| \frac{\partial}{\partial \log E_{\text{tot}}} \log\left(\lim_{S_{\text{tot}} \rightarrow \infty} \frac{[S_1]}{[S_0]}\right) \right| \leq \frac{\sqrt{c_0^E c_1^F / c_1^E c_0^F} - 1}{2}, \quad (30)$$

whose proof will follow below. By Eq. 17, the righthand side can be expressed in terms of the thermodynamic force alone:

$$\left| \frac{\partial}{\partial \log E_{\text{tot}}} \log\left(\lim_{S_{\text{tot}} \rightarrow \infty} \frac{[S_1]}{[S_0]}\right) \right| \leq \frac{\exp(\Delta\mu/2k_B T) - 1}{2}. \quad (31)$$

Equation 31 is our main result. We note that in the limit of thermodynamic irreversibility  $\Delta\mu/k_B T \rightarrow \infty$ , the righthand side of Eq. 31 diverges, consistent with the finding of “unlimited” sensitivity by Goldbeter and Koshland, as discussed in the introduction.

To prove the bound in Eq. 30, our starting point is Eq. 13. In prior work (16), it was shown that Eq. 13 implies the existence of the limit

$$\sigma = \lim_{S_{\text{tot}} \rightarrow \infty} [S_1] / [S_0]$$

and that the limiting value  $\sigma$  is the unique positive solution of the quadratic equation (Eq. 18; (16)):

$$\frac{E_{\text{tot}}}{F_{\text{tot}}} \left( \frac{1}{K_0^F} + \frac{\sigma}{K_1^F} \right) (c_0^E - c_1^E \sigma) = \left( \frac{1}{K_0^E} + \frac{\sigma}{K_1^E} \right) (c_1^F \sigma - c_0^F). \quad (32)$$



This equation can be rearranged to express  $\frac{E_{\text{tot}}}{F_{\text{tot}}}$  as a rational function of  $\sigma$ ,

$$\frac{E_{\text{tot}}}{F_{\text{tot}}} = \frac{\left(\frac{1}{K_0^E} + \frac{\sigma}{K_1^E}\right)(c_1^F \sigma - c_0^F)}{\left(\frac{1}{K_0^F} + \frac{\sigma}{K_1^F}\right)(c_0^E - c_1^E \sigma)}, \quad (33)$$

with numerator and denominator both positive for any value of  $\sigma$  within the bounds on its value set by our dynamic range result in Eq. 29. This means that  $E_{\text{tot}}/F_{\text{tot}}$  can be viewed as a continuously differentiable function of  $\sigma$ .

Rewriting Eq. 32 in terms of variables  $x = \log\left(\frac{E_{\text{tot}}}{F_{\text{tot}}}\right)$  and  $y = \log \sigma$ , we have

$$x = \log(K_1^E + K_0^E e^y) + \log(c_1^F e^y - c_0^F) - \log(K_1^F + K_0^F e^y) - \log(c_0^E - c_1^E e^y) + \log\left(\frac{K_0^F K_1^F}{K_0^E K_1^E}\right). \quad (34)$$

Our goal is to bound the derivative  $dy/dx$ . We will do this by studying the derivative of the inverse, which is

$$\frac{dx}{dy} = \left(\frac{e^y}{e^y - \frac{c_0^F}{c_1^F} + \frac{c_0^E}{c_1^E} e^y}\right) + \left(\frac{e^y}{\frac{K_1^E}{K_0^E} + e^y} - \frac{e^y}{\frac{K_1^F}{K_0^F} + e^y}\right). \quad (35)$$

The second term in brackets is the difference of two positive quantities no larger than one, so it can be no smaller than  $-1$ . It approaches  $-1$  when  $\frac{K_1^E}{K_0^E} \rightarrow \infty$  and  $\frac{K_1^F}{K_0^F} \rightarrow 0$ .

The first term in brackets is minimized, for fixed values of the tgCEs, when  $y = \log\sqrt{\frac{c_0^F c_0^E}{c_1^F c_1^E}}$ , when it takes the value

$$\left(\sqrt{\frac{c_1^F c_0^E}{c_0^F c_1^E} + 1}\right) / \left(\sqrt{\frac{c_1^F c_0^E}{c_0^F c_1^E} - 1}\right).$$

Therefore,

$$\frac{dx}{dy} \geq \frac{\sqrt{\frac{c_1^F c_0^E}{c_0^F c_1^E} + 1}}{\sqrt{\frac{c_1^F c_0^E}{c_0^F c_1^E} - 1}} - 1 > 0. \quad (36)$$

Since the function relating  $x$  and  $y$  is continuously differentiable and this derivative is never zero, this implies that

$$\frac{dy}{dx} \leq \left(\frac{\sqrt{\frac{c_1^F c_0^E}{c_0^F c_1^E} + 1}}{\sqrt{\frac{c_1^F c_0^E}{c_0^F c_1^E} - 1}} - 1\right)^{-1} = \frac{\sqrt{c_0^E c_1^F / c_1^E c_0^F} - 1}{2}, \quad (37)$$

which establishes Eq. 30. Finally, applying Eq. 17 yields Eq. 31. See Fig. 5 a for a numerical illustration of this result.

In prior work, Qian (19)—studying a reversible covalent-modification switch with a single intermediate for each enzyme—gave an asymptotic formula for the derivative  $\partial([S_1]/S_{\text{tot}})/\partial \log E_{\text{tot}}$  in the high substrate regime and at a point where  $[S_1]/S_{\text{tot}} = 1/2$ . Starting from our Eq. 35, and taking  $y = 0$ , we can develop a similar expression:

$$\frac{\partial}{\partial \log E_{\text{tot}}} \log\left(\lim_{S_{\text{tot}} \rightarrow \infty} \frac{[S_1]}{S_{\text{tot}}}\right) = \frac{1}{4} \left[ \left(\frac{1}{1 - \frac{c_0^F}{c_1^F} + \frac{c_0^E}{c_1^E} - 1}\right) + \left(\frac{1}{\frac{K_1^E}{K_0^E} + 1} - \frac{1}{\frac{K_1^F}{K_0^F} + 1}\right) \right]^{-1}, \quad (38)$$

when  $[S_1]/S_{\text{tot}} = 1/2$ . This result may be compared with Eq. 17 of Qian (19), but it is not equivalent to it. Our equation reduces, under the multiple limits simultaneously taken by Qian, to the thermodynamic parts of his expression. However, Qian's equation additionally includes terms

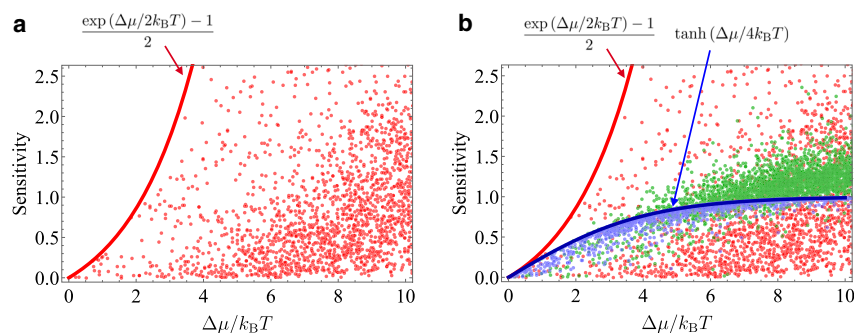


FIGURE 5 Thermodynamic force bounds sensitivity. Numerical illustration of our bounds by random sampling of rate constants for a covalent modification cycle in which both enzymes have single intermediates (see materials and methods for details). (a) Points show values of the thermodynamic force and sensitivity achieved when  $S_{\text{tot}} = 20,000$  (red). In (b), the same is plotted together with  $S_{\text{tot}} = 2$  (green), and  $S_{\text{tot}} = 0.01$  (blue). In all cases,  $E_{\text{tot}} = F_{\text{tot}} = 1$ . These numerical results are compared with the bound Eqs. 31 (solid red curve) and 39 (solid blue curve). To see this figure in color, go online.

proportional to  $1/S_{\text{tot}}$ , which cannot appear in our expression because we have already taken the limit  $S_{\text{tot}} \rightarrow \infty$ .

### The low substrate regime

The sensitivity bound Eq. 31 for the high substrate regime can be viewed as a companion to a corresponding result in the low substrate regime (43) given by the following formula:

$$\left| \frac{\partial}{\partial \log E_{\text{tot}}} \log \left( \lim_{S_{\text{tot}} \rightarrow 0} \frac{[S_1]}{[S_0]} \right) \right| \leq \tanh(\Delta\mu / 4k_B T). \quad (39)$$

Note that the righthand side is less than 1: there can be no ultrasensitivity in the low substrate regime. This result is a manifestation of a recently identified (43) universal thermodynamic bound on the response of nonequilibrium systems to perturbations. It also follows directly from the approach described here, as we now show.

In the limit where  $S_{\text{tot}}$  is very small, we have  $[E] \rightarrow E_{\text{tot}}$ ,  $[F] \rightarrow F_{\text{tot}}$ , which together with Eq. 18 implies that

$$\lim_{S_{\text{tot}} \rightarrow 0} \frac{[S_1]}{[S_0]} = \frac{c_0^E (E_{\text{tot}}/F_{\text{tot}}) + c_0^F}{c_1^E (E_{\text{tot}}/F_{\text{tot}}) + c_1^F}. \quad (40)$$

Let  $\varphi$  denote this limiting quantity and set  $z = E_{\text{tot}}/F_{\text{tot}}$ . Taking the derivative, we find that

$$\begin{aligned} \frac{\partial}{\partial \log E_{\text{tot}}} \log(\varphi) &= \frac{c_0^E c_1^F z - c_0^F c_1^E z}{(c_0^E + c_0^F z)(c_1^E + c_1^F z)} \\ &= \frac{(\sqrt{c_0^E c_1^F z} - \sqrt{c_0^F c_1^E z})(\sqrt{c_0^E c_1^F z} + \sqrt{c_0^F c_1^E z})}{(c_0^E + c_0^F z)(c_1^E + c_1^F z)}. \end{aligned}$$

Using the inequality of the arithmetic and geometric means,

$$(c_0^E + c_0^F z)(c_1^E + c_1^F z) \geq \left( \sqrt{c_0^E c_1^F z} + \sqrt{c_0^F c_1^E z} \right)^2.$$

Recalling that  $\tanh(x) = (e^{2x} - 1)/(e^{2x} + 1)$ , we see that

$$\begin{aligned} \left| \frac{\partial}{\partial \log E_{\text{tot}}} \log(\varphi) \right| &\leq \frac{\sqrt{c_0^E c_1^F z} - \sqrt{c_0^F c_1^E z}}{\sqrt{c_0^E c_1^F z} + \sqrt{c_0^F c_1^E z}} \\ &= \tanh \left( \frac{1}{4} \log \left( \frac{c_0^E c_1^F}{c_0^F c_1^E} \right) \right). \end{aligned} \quad (41)$$

Finally, applying relation Eq. 17 yields Eq. 39. See Fig. 5 b for a comparison of this bound with that for the high substrate regime. The bound on the sensitivity in the high substrate regime is larger than the bound in the low substrate regime for all nonzero values of  $\Delta\mu/k_B T$ . For small  $\Delta\mu/k_B T$ , the two bounds are equal to first order.

### Saturating our bounds

Remarkably, at any fixed force, there are covalent modification cycles that, with the same kinetic parameters, can saturate either bound Eq. 39 or Eq. 31, as the substrate

concentration is varied. Maximal sensitivity, in both  $S_{\text{tot}}$  regimes, can be achieved when the enzymes act as “mirrors” of each other (i.e., the network is invariant under the exchange  $E \leftrightarrow F$  and  $S_0 \leftrightarrow S_1$ ).

For example, consider a covalent modification cycle with  $c_0^E = c_1^F = \exp(\Delta\mu/2k_B T)$ ,  $c_1^E = c_0^F = 1$ , and the “forward” generalized Michaelis-Menten constants  $K_0^E = K_1^F \equiv K_M$ , which are very small compared with the “reverse” constants  $K_1^E = K_0^F \equiv K_R$ . For such a covalent modification cycle, Eq. 35 gives the maximal sensitivity in the high  $S_{\text{tot}}$  regime explicitly as

$$\frac{\exp(\Delta\mu/2k_B T) - 1}{2} \times \frac{1 + K_M/K_R}{1 + \exp(\Delta\mu/2k_B T)K_M/K_R}, \quad (42)$$

saturating our bound Eq. 31 when  $K_R \gg K_M$ . For the same parameters, in the low  $S_{\text{tot}}$  regime, we have

$$\frac{[S_1]}{[S_0]} \approx \frac{c_0^E (E_{\text{tot}}/F_{\text{tot}}) + c_0^F}{c_1^E (E_{\text{tot}}/F_{\text{tot}}) + c_1^F} = \frac{\exp(\Delta\mu/2k_B T)(E_{\text{tot}}/F_{\text{tot}}) + 1}{(E_{\text{tot}}/F_{\text{tot}}) + \exp(\Delta\mu/2k_B T)}, \quad (43)$$

leading to a maximal sensitivity of  $\tanh(\Delta\mu/4k_B T)$ , saturating the low  $S_{\text{tot}}$  bound Eq. 39. These results are illustrated in Fig. 6.

## DISCUSSION

The bounds we have presented provide a quantitative picture of how general covalent modification cycles are

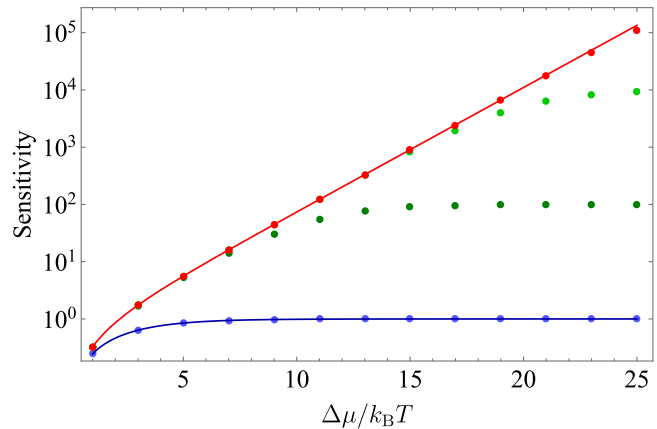


FIGURE 6 The same parameter choice can saturate the low and high  $S_{\text{tot}}$  bounds. The sensitivity bound given by Eq. 31 in the high  $S_{\text{tot}}$  regime is shown by the solid red line, and the sensitivity bound given by Eq. 39 in the low  $S_{\text{tot}}$  regime is shown by the solid blue line. The colored dots represent sensitivities for a covalent modification cycle with  $K_0^E = K_1^F = 10^{-4}$ ,  $K_1^E = K_0^F = 10^3$ ,  $c_0^E = c_1^F = \exp(\Delta\mu/2k_B T)$ ,  $c_1^E = c_0^F = 1$ , and  $E_{\text{tot}} = F_{\text{tot}} = 1$  (see materials and methods for details). Different colors for the dots signify different values of  $S_{\text{tot}}$ : 100 (red), 4 (light green), 2 (dark green), and 0.01 (blue). Note that  $S_{\text{tot}} = 100$  and  $S_{\text{tot}} = 0.01$  already achieve, within the limits of resolution, the higher and lower bounds, respectively. To see this figure in color, go online.

constrained by thermodynamics. We have framed our results in terms of the chemical potential  $\Delta\mu$ , which is the nonequilibrium driving force of the system and is a natural way to quantify the energetic requirements of the switch. To see this, it is important to note that there can be no direct trade-off between the rate of (free) energy expenditure—the power—and steady-state properties of the system like its sensitivity or dynamic range. This is because the power can be scaled arbitrarily by scaling all reaction rates by the same constant or, equivalently, by rescaling time, whereas this operation leaves the steady state completely unchanged. To compare it with steady-state quantities, power must be divided by some other rate relevant to the system. The quantity  $\Delta\mu$  is equal to the energy consumed per cycle completion, i.e., it is the power divided by the rate of cycle completion.

Our results highlight the fact that under common circumstances, thermodynamics does not tightly constrain the sensitivity of biochemical switches. For example, in a phosphorylation-dephosphorylation cycle driven by ATP hydrolysis, typical physiological values of  $\Delta\mu \approx 20 - 30 k_B T$  are deep in the saturated regime of the low  $S_{\text{tot}}$  bound and, in the high  $S_{\text{tot}}$  case, lead to a maximum possible sensitivity of  $\sim 10^4$  to  $10^6$ , far in excess of what is needed to account for typical Hill coefficients measured in ultrasensitive systems (20). The sensitivity of such strongly driven switches may be constrained by other factors, such as kinetics or the abundance of substrate.

Our bounds also show that, from the perspective of making a good switch based on covalent modification, enzymological complexity provides no benefit. No matter the number of enzymatic intermediates or how elaborate their reactions, the same thermodynamic bounds on sensitivity hold. This is in contrast to numerous examples in biophysics where having more states or “steps” provides some advantage. For example, in kinetic proofreading, having more proofreading steps allows for a degree of discrimination that can never be achieved with fewer, even as  $\Delta\mu \rightarrow \infty$  (43–45). As another example, the maximum possible coherence of biochemical oscillations is conjectured to depend strongly on the number of states available, together with the strength of nonequilibrium driving (46).

A number of basic questions remain. First, what is the bound on sensitivity in terms of  $\Delta\mu$  and  $S_{\text{tot}}$ ? We have only studied the limiting cases in which substrate is very scarce or abundant. It is natural to conjecture that the maximum possible sensitivity is increasing in  $S_{\text{tot}}$  and that, therefore, our high substrate bound Eq. 31 in fact holds for all  $S_{\text{tot}}$ .

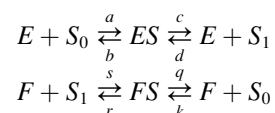
Second, in this article, we have focused on steady-state behavior. It would be very interesting to also understand the constraints on the transient behavior. For example, in vision, the exceptional amplification that

enables rod cells to respond one or a few photons involves a transient response of a modification cycle involving rhodopsin and transducin, driven by GTP hydrolysis (47,48).

Finally, we note that even simpler properties of general covalent modification cycles remain incompletely understood. For example, to our knowledge, it remains an open problem to prove the monotonicity of the steady-state ratio  $[S_1]/[S_0]$  as a function of the enzyme totals. Such matters may at first seem only of mathematical interest, but we think understanding them carefully could bear fruit—especially in the study of systems, such as signaling cascades, in which covalent modification cycles appear as parts (49).

## MATERIALS AND METHODS

To generate Fig. 5, the sensitivity was evaluated numerically for the simplest reversible covalent modification cycle



for random choices of the rate constants  $a, b, c, d, s, r, q,$  and  $k$ . Explicitly, for the  $S_{\text{tot}} = 20,000$  (red) and  $S_{\text{tot}} = 2$  (green) points, all rate constants were drawn uniformly from the interval  $(0, 10)$  except for  $d$  and  $k$ , which were drawn uniformly from  $(0, 0.1)$ . For the  $S_{\text{tot}} = 0.01$  (blue) points, all rate constants were drawn uniformly from the interval  $(0, 10)$  except for  $d$  and  $k$ , which were drawn uniformly from  $(0, 0.5)$ . The polynomial equations defining the steady state were solved numerically (using `NSolve` in Wolfram Mathematica), and the derivative defining the sensitivity at  $E_{\text{tot}} = F_{\text{tot}} = 1$  was estimated by taking a finite difference. For this covalent modification cycle,  $\Delta\mu/k_B T = \log\left(\frac{acsa}{bdrk}\right)$ .

To generate Fig. 6, Eqs. 12 and 13 were solved numerically, with  $K_0^E = K_1^F = 10^{-4}$ ,  $K_1^E = K_0^F = 10^3$ ,  $c_0^E = c_1^F = \exp(\Delta\mu/2k_B T)$ ,  $c_1^E = c_0^F = 1$ , and varying values of  $S_{\text{tot}}$ . The sensitivity with respect to  $E_{\text{tot}}$  at  $E_{\text{tot}} = F_{\text{tot}} = 1$  was estimated by taking a finite difference, i.e., by finding  $S_1/S_0$  for  $E_{\text{tot}} = 1$  and  $E_{\text{tot}} = 1 + \epsilon$ , taking the difference of the resulting values, and dividing by  $\epsilon$ . We used  $\epsilon = 0.000001$ .

## AUTHOR CONTRIBUTIONS

Conceptualization, J.A.O., P.T., J.W.B., and J.G.; investigation, J.A.O., P.T., J.W.B., and J.G.; writing, J.A.O., P.T., J.W.B., and J.G.; supervision, J.G.; funding acquisition, J.G.

## ACKNOWLEDGMENTS

We thank two anonymous reviewers for their helpful comments. J.W.B. and J.G. were supported by grant 1462629 from the US National Science Foundation.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## REFERENCES

1. Prabakaran, S., G. Lippens, ..., J. Gunawardena. 2012. Post-translational modification: nature's escape from from genetic imprisonment and the basis for cellular information processing. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 4:565–583.
2. Ree, R., S. Varland, and T. Arnesen. 2018. Spotlight on protein N-terminal acetylation. *Exp. Mol. Med.* 50:1–13.
3. Walsh, C. T. 2006. Posttranslational Modification of Proteins. Roberts and Company.
4. Farrelly, L. A., R. E. Thompson, ..., I. Maze. 2019. Histone seronylation is a permissive modification that enhances TFIID binding to H3K4me3. *Nature.* 567:535–539.
5. Scrutton, M. C., and M. F. Utter. 1968. The regulation of glycolysis and gluconeogenesis in animal tissues. *Annu. Rev. Biochem.* 37:249–302.
6. Goldbeter, A., and D. E. Koshland. 1981. An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci. USA.* 78:6840–6844.
7. Stadtman, E. R., and P. B. Chock. 1977. Superiority of interconvertible enzyme cascades in metabolic regulation: analysis of monocyclic systems. *Proc. Natl. Acad. Sci. USA.* 74:2761–2765.
8. Blüthgen, N., F. J. Bruggeman, ..., B. N. Kholodenko. 2006. Effects of sequestration on signal transduction cascades. *FEBS J.* 273:895–906.
9. Chen, C. A., and D. R. Manning. 2001. Regulation of G proteins by covalent modification. *Oncogene.* 20:1643–1652.
10. Torrecilla, I., and A. B. Tobin. 2006. Co-ordinated covalent modification of G-protein coupled receptors. *Curr. Pharm. Des.* 12:1797–1808.
11. Ventura, A. C., P. Jiang, ..., A. J. Ninfa. 2010. Signaling properties of a covalent modification cycle are altered by a downstream target. *Proc. Natl. Acad. Sci. USA.* 107:10032–10037.
12. Berger, S. L. 2001. An embarrassment of niches: the many covalent modifications of histones in transcriptional regulation. *Oncogene.* 20:3007–3013.
13. Lehembre, F., P. Badenhorst, A. Dejean, ..., 2000. Covalent modification of the transcriptional repressor tramtrack by the ubiquitin-related protein Smt3 in *Drosophila* flies. *Mol. Cell Biol.* 20:1072–1082.
14. Fisher, D., L. Krasinska, ..., B. Novák. 2012. Phosphorylation network dynamics in the control of cell cycle transitions. *J. Cell Sci.* 125:4703–4711.
15. Okamura, H., M. Doi, ..., M. Matsuo. 2010. Mammalian circadian clock system: molecular mechanisms for pharmaceutical and medical sciences. *Adv. Drug Deliv. Rev.* 62:876–884.
16. Dasgupta, T., D. H. Croll, ..., J. Gunawardena. 2014. A fundamental trade-off in covalent switching and its circumvention by enzyme bifunctionality in glucose homeostasis. *J. Biol. Chem.* 289:13010–13025.
17. Kustu, S., J. Hirschman, ..., J. C. Meeks. 1984. Covalent modification of bacterial glutamine synthetase: physiological significance. *Mol. Gen. Genet.* 197:309–317.
18. Goldbeter, A., and D. E. Koshland, Jr. 1987. Energy expenditure in the control of biochemical systems by covalent modification. *J. Biol. Chem.* 262:4460–4471.
19. Qian, H. 2003. Thermodynamic and kinetic analysis of sensitivity amplification in biological signal transduction. *Biophys. Chem.* 105:585–593.
20. Ferrell, J. E., Jr., and S. H. Ha. 2014. Ultrasensitivity part I: Michaelian responses and zero-order ultrasensitivity. *Trends Biochem. Sci.* 39:496–503.
21. Gunawardena, J. 2012. Some lessons about models from Michaelis and Menten. *Mol. Biol. Cell.* 23:517–519.
22. Michaelis, L., and M. Menten. 1913. Die kinetik der Invertinwirkung. *Biochem. Z.* 49:333–369.
23. Nam, K.-M., B. M. Gyori, ..., J. Gunawardena. 2020. Robustness and parameter geography in post-translational modification systems. *PLoS Comput. Biol.* 16:e1007573.
24. Ortega, F., L. Acerenza, ..., M. Cascante. 2002. Product dependence and bifunctionality compromise the ultrasensitivity of signal transduction cascades. *Proc. Natl. Acad. Sci. USA.* 99:1170–1175.
25. Fersht, A. 1985. Enzyme Structure and Mechanism. W. H. Freeman and Company.
26. Gunawardena, J. 2012. A linear framework for time-scale separation in nonlinear biochemical systems. *PLoS One.* 7:e36321.
27. Xu, Y., and J. Gunawardena. 2012. Realistic enzymology for post-translational modification: zero-order ultrasensitivity revisited. *J. Theor. Biol.* 311:139–152.
28. Qian, H. 2007. Phosphorylation energy hypothesis: open chemical systems and their biological functions. *Annu. Rev. Phys. Chem.* 58:113–142.
29. Thomson, M., and J. Gunawardena. 2009. The rational parameterisation theorem for multisite post-translational modification systems. *J. Theor. Biol.* 261:626–636.
30. Gunawardena, J. 2014. Time-scale separation: Michaelis and Menten's old idea still bearing fruit. *FEBS J.* 281:473–488.
31. Nam, K.-M., R. Martinez-Corral, and J. Gunawardena. 2022. The linear framework: using graph theory to reveal the algebra and thermodynamics of biomolecular systems. *Interface Focus.* 12:20220013.
32. Hill, T. L. 1966. Studies in irreversible thermodynamics IV. diagrammatic representation of steady state fluxes for unimolecular systems. *J. Theor. Biol.* 10:442–459.
33. King, E. L., and C. Altman. 1956. A schematic method of deriving the rate laws for enzyme-catalyzed reactions. *J. Phys. Chem.* 60:1375–1378.
34. Mirzaev, I., and J. Gunawardena. 2013. Laplacian dynamics on general graphs. *Bull. Math. Biol.* 75:2118–2149.
35. Tutte, W. T. 1948. The dissection of equilateral triangles into equilateral triangles. *Math. Proc. Camb. Phil. Soc.* 44:463–482.
36. Segel, I. H. 1993. Enzyme Kinetics: Behaviour and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems. Wiley Interscience.
37. Astumian, R. D., C. Pezzato, ..., J. F. Stoddart. 2020. Non-equilibrium kinetics and trajectory thermodynamics of synthetic molecular pumps. *Mater. Chem. Front.* 4:1304–1314.
38. Lewis, G. N. 1925. A new principle of equilibrium. *Proc. Natl. Acad. Sci. USA.* 11:179–183.
39. Esposito, M., and C. Van den Broeck. 2010. Three faces of the second law. I. Master equation formulation. *Phys. Rev. E.* 82:011143.
40. Schnakenberg, J. 1976. Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Mod. Phys.* 48:571–585.
41. Seifert, U. 2011. Stochastic thermodynamics of single enzymes and molecular motors. *Eur. Phys. J. E.* 34:26.
42. Wong, F., A. Dutta, ..., J. Gunawardena. 2018. Structural conditions on complex networks for the Michaelis-Menten input-output response. *Proc. Natl. Acad. Sci. USA.* 115:9738–9743.
43. Owen, J. A., T. R. Gingrich, and J. M. Horowitz. 2020. Universal thermodynamic bounds on nonequilibrium response with biochemical applications. *Phys. Rev. X.* 10:011066.

44. Murugan, A., D. A. Huse, and S. Leibler. 2014. Discriminatory proof-reading regimes in nonequilibrium systems. *Phys. Rev. X*. 4:021016.
45. Ouldridge, T. E., C. C. Govern, and P. R. ten Wolde. 2017. Thermodynamics of computational copying in biochemical systems. *Phys. Rev. X*. 7:021004.
46. Barato, A. C., and U. Seifert. 2017. Coherence of biochemical oscillations is bounded by driving force and network topology. *Phys. Rev. E*. 95:062409.
47. Arshavsky, V. Y., and M. E. Burns. 2014. Current understanding of signal amplification in phototransduction. *Cell. Logist.* 4:e29390.
48. Baylor, D. 1996. How photons start vision. *Proc. Natl. Acad. Sci. USA*. 93:560–565.
49. Barabanschikov, A., and J. Gunawardena. 2019. Monostationarity and multistationarity in tree networks of Goldbeter–Koshland loops. *Bull. Math. Biol.* 81:2463–2509.