Noname manuscript No.
(will be inserted by the editor)

The Effect of Site-to-Site Variability in Ultrasensitive Dose Responses

German A. Enciso · Shane Ryerson

Received: date / Accepted: date

Abstract In this paper we study the ultrasensitive behavior of multisite phosphorylation or ligand binding systems, under site-to-site variations in the modification rates. Using computational methods and mathematical analysis, we prove that the Hill coefficient reaches its maximum value when all sites are identical to each other. This is shown for a non-cooperative multisite system with arbitrary activation function as well as for the well known MWC model. We also show that the Hill coefficient of the dose response is locally robust to variations in individual modification rates. The results suggest that maximal ultrasensitivity is reached when sites are similar to each other but not necessarily identical, a conformation found in unstructured modification domains present in many experimental systems.

Keywords Multisite system · Phosphorylation · Signal transduction · Ultrasensitivity · Cooperativity · Allostery

MSC Classification: 92C40, 92C42, 37N25

1 Introduction

Post-translational modifications, such as phosphorylation or ligand binding, are ubiquitous regulators of protein activity. Many proteins are modified at multiple specific locations. For instance, the hemoglobin protein can bind oxygen at four different sites, the Ste5 protein can be phosphorylated at eight

G.A. Enciso (Corresponding Author)
Mathematics Department
University of California, Irvine (USA)
Tel.: +1 949 824 3727
E-mail: enciso@uci.edu

S. Ryerson
Mathematics Department
University of California, Irvine (USA)

sites [25], and the CheY protein can bind the flagellar motor at as many as 34 sites [22]. Many forms of cellular function, from the regulation of the cell cycle [10] to the regulation of gene expression through DNA packaging [27] or the signal transduction of chemotaxis [24], are heavily dependent on the covalent or non-covalent multisite modification of key proteins (Figure 1a,b).

While multisite modifications can serve many different purposes, their best known role lies in their use to create ultrasensitive, or switch-like, dose responses. For instance in the case of hemoglobin, a relatively small difference in O_2 concentration can lead to a large difference in the O_2-hemoglobin binding affinity, which is an essential element in O_2 transport efficiency [12]. In a similar way, multisite phosphorylation is often proposed as a source of sigmoidal, nonlinear responses, which are important for the creation of bistability, oscillations and other nontrivial dynamical behaviors [1, 4, 16, 23].

There are several classical models, and many more follow up models, that explain how multisite modifications can lead to ultrasensitivity [12, 14, 15]. Perhaps the most famous of them is the allosteric Monod-Wyman-Changeux (MWC) model, which assumes that the protein changes back and forth between two different states with different ligand binding affinities [17]. Ligand binding in turn tends to sequester the protein in the high affinity state, which facilitates ligand binding at other sites. More recently, work by the authors has shown that non-allosteric proteins can also behave in an ultrasensitive manner [4, 21].

In this paper we explore how site-to-site variability can affect ultrasensitive responses. Many models in the literature make the simplifying assumption that all sites have equivalent dynamics, in the sense that modification rates are the same at all unmodified sites. For instance, in a hemoglobin molecule with one bound site, all other three sites are modified at the same rate. Similarly, all modified sites of a protein are assumed to lose their modification at the same rate. This assumption is mostly made out of convenience: different dynamics at different sites lead to an unwieldy system, and often the protein in question is one of several proteins in the model, so that there is a need to keep the system as simple as possible. Moreover, using different parameters at each site leads to many more parameters to fit.

This raises interesting questions regarding the ultrasensitive behavior of systems whose sites are not equivalent. For instance, could the ultrasensitivity be *increased* by altering the site dynamics in some clever way, such as dividing the sites into groups with different modification rates? How robust is the Hill coefficient of a dose response with respect to small site-to-site variations?

Using both computational and analytical techniques, we answer these questions and show that maximal ultrasensitivity of a multisite system is obtained when all sites have identical dynamics. Moreover, the Hill coefficient of a dose response is robust to small variations in site-to-site dynamics. This has important biological consequences, as many multisite proteins will have evolutionary pressure to mutate towards roughly equal site-to-site dynamics, i.e. roughly equal access of ligands, kinases and other enzymes to all the individual sites.

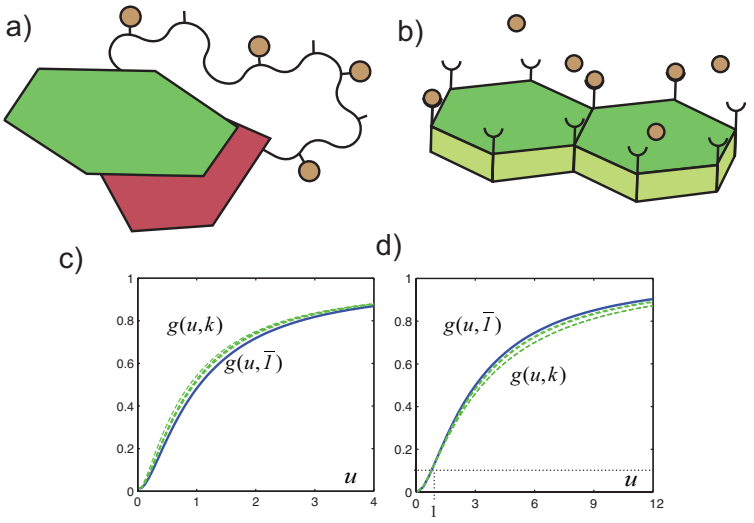

Fig. 1 **a:** Unstructured multisite phosphorylation is common in cell regulation and allows for similar access of regulatory enzymes to the phosphorylation sites. **b:** Multisite ligand binding as an example of non-covalent modification. **c:** Dose response $g(u,k)$ for $k_i = 1$ for all i (solid blue), and several variations with 40% to 50% site-to-site variation (dashed green). Here $n = 14$, $a(x) = e^{\gamma x}/(e^{\gamma/2} + e^{\gamma x})$ and $\gamma = 30$. **d:** Same graphs as in c, but rescaling along the x axis to ensure the value of each graph at $u = 1$ is 0.1.

These predictions are confirmed by several biological experiments. Bioinformatic and crystallographic studies predict that most multisite phosphorylation are unordered [13,19,29]. That is, the phosphorylation sites of proteins such as Ste5 tend to be located along flexible polypeptides without a tertiary or quaternary structure, which facilitates equal access to ligands and kinases (Figure 1a). Moreover, phylogenetic analysis shows that while multisite domains are conserved across species, the individual sites are not conserved, further supporting the idea that these domains are unstructured.

In this way, many proteins appear to have evolved flexible, unordered multisite domains. This picture is consistent with our results, since unordered domains are a simple way to enforce similar access to ligands and modifying enzymes. Unordered domains would likely still contain relatively minor differences in access, but this is not a problem given the robustness of the Hill coefficient to such changes.

In the following sections we study different multisite systems with arbitrary site-to-site dynamics and explore how the Hill coefficient of the system changes with the parameters that regulate each site. For most of the mathematical analysis we use for simplicity a non-cooperative system that generalizes the models in [21,28], using a different parameter for each site. For this system we prove rigorously that identical sites lead to maximal ultrasensitivity (Theorem 2), and that the Hill coefficient of the dose response is robust to small changes in the modification parameters (Theorem 1). We also carry out a computational

analysis of the Hill coefficient as a function of the coefficient of variation of the site parameters, for the non-cooperative model as well as a generalized MWC model. This illustrates that ultrasensitivity is also optimized in this classical model when sites are modified at identical rates.

2 Non-cooperative Model

Suppose that S is a substrate with n modification sites, which could be covalent or non-covalent and correspond to phosphorylation, ligand binding, etc. We consider non-sequential, first-order modification reactions of this multisite substrate, so that there are a total of 2^n possible protein states. Assume the following:

(H1) Each site is modified independently of the state of the other sites.

This in particular implies that the system is non-cooperative and non-allosteric. The authors have recently studied such a system in the context of chemotaxis signaling and Ste5 phosphorylation [21], and in a practical biological model of bud size control [4]. We denote by S_I the concentration of the protein with phosphorylation state $I \in \{0,1\}^n$. If the protein S_I is not modified at site i, the modification at this site takes place at a rate $\alpha_i u S_I$, where u, the input of the system, represents the concentration of the modifying enzyme (or the concentration of the ligand, as appropriate). If S_I is already modified at site i, then the modification is removed at a rate $\beta_i S_I$. If the removal of the modification is mediated by another enzyme (such as a phosphatase), then we assume that the concentration of that enzyme is constant and is hidden inside the parameter β_i.

If we denote by z_i and w_i the fraction of i-sites that are modified and unmodified at any given time, respectively, we can write ODEs of the form

$$z_i' = \alpha_i u w_i - \beta_i z_i, \quad w_i' = -\alpha_i u w_i + \beta_i z_i, \quad i = 1, \dots, n. \tag{1}$$

Since $z_i + w_i = 1$, one can simplify these ODEs to $z_i' = \alpha_i u(1 - z_i) - \beta_i z_i$. A simple calculation shows that the fraction of modified i-sites at steady state is $z_i = u/(u + k_i)$, where $k_i := \beta_i/\alpha_i$ is the i-th dissociation constant of the system [21]. Thus the dissociation constant k_i incorporates the variability from site to site in this system.

Regarding the activation of the multisite protein, we assume

(H2) The activation level of the protein depends only on the number of modified sites.

This allows to define an activation function $a(x)$, indicating the fractional activity of a substrate molecule that has a fraction x of modified sites. For instance, Ste5 is only active when bound to the cell membrane. If a molecule of Ste5 has 5 phosphorylated sites out of 8, and this leads the protein to

spend 20% of the time bound to the membrane, we say that $a(5/8) = 0.2$. For other proteins, activation might mean the ability to act as an enzyme to other proteins, or being able to bind to a target. For simplicity we assume that the function $a : [0, 1] \to [0, 1]$ is continuous and monotonically increasing, and that $a(0) < 0.1a(1)$.

The work in [21] follows the same assumptions we have here but additionally it includes the third assumption that the modifications are identical from site to site. We relax that assumption here and in this way generalize that work.

Define S_T to be the total concentration of substrate, which is a constant since there is no protein production or degradation in this system. Let us suppose first that all sites are identical, that is $k_i = k$ and $z_i = z = u/(u + k)$ for all i. If we consider a modification state of the substrate by the index $I \in \{0, 1\}^n$, then the independence among the sites allows us to compute the concentration S_I of the protein in state I as

$$\frac{S_I}{S_T} = z^{|I|}(1 - z)^{n-|I|},$$

where $|I| = I_1 + \ldots + I_n$. Finally, this allows to calculate the so-called dose response of the system, which describes the relationship between the input u (the modifying enzyme or ligand) and the total fraction of active protein:

$$\begin{aligned} g(u) &= \frac{1}{S_T}\sum_I S_I a(|I|/n) = \sum_I z^{|I|}(1 - z)^{n-|I|}a(|I|/n) \\ &= \sum_{j=0}^{n} \binom{n}{j} z^j (1 - z)^{n-j} a(j/n) \approx a(z), \end{aligned}$$

where $z = z(u)$. The last approximation follows from the study of Bernstein polynomials and becomes increasingly accurate as $n \to \infty$ [11,21]. This illustrates that the dose response $g(u)$ of this system can be approximated in very simple terms using the activation function.

In case that the dissociation constants k_i are not all equal to each other, the dose response becomes more elaborate.

Lemma 1 *In the case of arbitrary dissociation constants $k_1, \ldots k_n$, the active protein fraction satisfies the formula*

$$g(u, k) = \frac{1}{P(u, k)} \sum_{i=0}^{n} a_i e_{n-i}(k) u^i, \tag{2}$$

where $k := (k_1, \ldots, k_n)$, $P(u, k) := \prod_{j=1}^{n} u + k_j$, $a_i := a(i/n)$, and $e_i(k) := \sum_{|I|=i} \prod_{I_j=1} k_j$ is the elementary symmetric polynomial of degree i over the variables $k = (k_1, \ldots, k_n)$.

Proof. We calculate for any index I,

$$\frac{S_I}{S_T} = \prod_{I_j=1} z_j \prod_{I_j=0} 1 - z_j = \prod_{I_j=1} \frac{u}{u+k_j} \prod_{I_j=0} \frac{k_j}{u+k_j} = \frac{1}{P(u,k)} u^{|I|} \prod_{I_j=0} k_j.$$

The dose response now can be written as

$$\begin{aligned} g(u,k) &= \frac{1}{S_T} \sum_I a_{|I|} S_I = \frac{1}{P(u,k)} \sum_I a_{|I|} u^{|I|} \prod_{I_j=0} k_j \\ &= \frac{1}{P(u,k)} \sum_{i=0}^{n} a_i u^i \sum_{|I|=i} \prod_{I_j=0} k_j = \frac{1}{P(u,k)} \sum_{i=0}^{n} a_i u^i e_{n-i}(k). \end{aligned}$$

□

For example in the case $n = 2$, the active fraction is

$$\begin{aligned} g(u,k) &= a_0(1-z_1)(1-z_2) + a_1 z_1(1-z_2) + a_1(1-z_1)z_2 + a_2 z_1 z_2 \\ &= a_0 \frac{k_1 k_2}{(u+k_1)(u+k_2)} + a_1 \frac{uk_2 + uk_1}{(u+k_1)(u+k_2)} + a_2 \frac{u^2}{(u+k_1)(u+k_2)} \\ &= \frac{1}{P(u,k)} [a_0 e_2 + a_1 e_1 u + a_2 e_0 u^2], \end{aligned}$$

since $e_0 = 1,\ e_1 = k_1 + k_2,\ e_2 = k_1 k_2$.

Regarding the dissociation parameters k_i, we assume that $k_i \geq 0$, $i = 1 \dots n$, but that they are not all identically zero.

3 Robustness of Ultrasensitivity

In this section we show that relatively large variability from site to site does not affect the dose response or its Hill coefficient to a significant extent. We do this both through computational and mathematical analysis. The theoretical results Theorem 1 and Proposition 1 state that if the k_i are similar to each other, then $g(u,(k_1,\dots,k_1))$ equals $g(u,(\bar{k},\dots,\bar{k}))$ to a first order approximation, where $\bar{k} := (k_1 + \dots + k_n)/n$. The Hill coefficients of both dose responses are also the same except for higher order error terms.

In Figure 1c we simulate the dose response $g(\cdot,k)$ for multiple sets of dissociation rates $k_1,\dots,k_n$. We plot a dose response with identical parameters $k_i = 1$ using a solid blue line. Here $n = 14$ and the activation function is of the form $a(x) = e^{\gamma x}/(\delta + e^{\gamma x})$ as detailed in [4]. We also plot several random variations with a coefficient of variation $CV = \sigma(k)/E(k)$ between 0.4 and 0.5.

In Figure 1d we display the same curves, but rescaling each curve along the x-axis so that they have value 0.1 at the same input $u = 1$. Intriguingly, the solid curve now has maximal values among the different functions, which will have important consequences for the analysis of the next section.

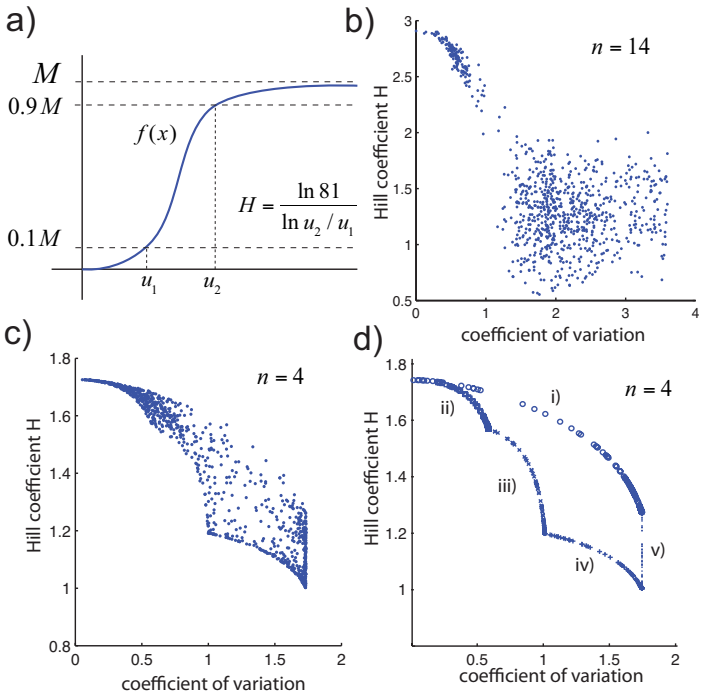

Fig. 2 **a:** Defining the generalized Hill coefficient for arbitrary sigmoidal curves. **b:** Each set of dissociation constants $k = (k_1 \dots k_n)$ is assigned a blue dot according to the coefficient of variation $\sigma(k)/E(k)$ and the generalized Hill coefficient of the dose response. Here the same value of n and activity function as in Figure 1c. **c:** Same plot for $n = 4$ and $a_0 = a_1 = a_2 = 0$ and $a_3 = a_4 = 1$. **d:** Restricted parameter sets illustrate what parameter combinations increase or decrease the Hill coefficient of the dose response. i) $k_1 > 1, k_2 = k_3 = k_4 = 1$. ii) $k_1 = 1, k_2 = k_3 = k_4 > 1$. For $a = 1000$: iii) $k_1 = 1$, $1 < k_2 < a$, $k_3 = k_4 = a$. iv) $k_1 = k_2 = 1$, $a < k_3 < a^2$, $k_4 = a^2$. v) $k_1 = k_2 = 1$, $1 < k_3 < a$, $k_4 = a^2$.

In order to quantify ultrasensitive behavior we use the definition of a generalized Hill coefficient given in [7] (see also [3]). For any sigmoidal function $f : \mathbb{R}^+ \to \mathbb{R}^+$ with saturation value M, set

$$H := \frac{\ln 81}{\ln \frac{f^{-1}(0.9M)}{f^{-1}(0.1M)}}.$$

See Figure 2a for an illustration. The smaller the ratio between $f^{-1}(0.9M)$ and $f^{-1}(0.1M)$, the higher the ultrasensitivity and the larger H. Moreover, if $f(x) = x^h/(K^h + x^h)$ is any Hill function, then $H = h$. Therefore H generalizes the standard Hill coefficient to arbitrary sigmoidal functions.

Figure 2b shows a plot of the Hill coefficient of multiple dose responses as in Figure 1c, plotted against the coefficient of variation $\sigma(k)/E(k)$ of the dissociation rates. The individual values of k_i were now varied over 8 orders of magnitude each for every single data point, leading to a much larger range in

the coefficient of variation CV. Notice that the ultrasensitivity appears to be highest when the noise is equal to zero, that is when all sites have the same value. As the noise increases, there is a wider variability in the ultrasensitivity. However when the noise is relatively small, the Hill coefficient changes little. We carried out the same analysis for a simpler system in the case $n = 4$, and found qualitatively similar results in a simpler graph (Figure 2c). Here $a_0 = a_1 = a_2 = 0$ and $a_3 = a_4 = 1$ for simplicity.

In Figure 2d we aimed at identifying the parameter sets that lead to the highest and lowest ultrasensitive behavior for a given amount of noise. The parameter sets in curve i) consistently had the highest ultrasensitive behavior, and they were modeled using the constraint $k_1 > 1, k_2 = k_3 = k_4 = 1$. This corresponds to a system where most sites are identical, except for one site that is easier to dissociate (due to a higher dissociation constant). For a small amount of noise, the lowest ultrasensitivity is found when $k_1 = 1, k_2 = k_3 = k_4 > 1$. This is a system in which one site dominates because it has much higher affinity than the rest, and therefore it resembles a system where $n = 1$. Other parameter restrictions closely approach the edges found in the case of random parameter combinations.

We prove mathematically one of the observations in these simulations, namely that starting from a system with identical dissociation constants $k_i = \bar{k}$, $i = 1 \dots n$, small variations in the value of these parameters do not change the ultrasensitivity of the system to a first order approximation. This can be seen in Figure 2b,c and d, where the initial slope of the Hill coefficient with respect of parameter variations is *zero*. Notice that this is a stronger claim than to simply state that $g(\cdot, k)$ and $H(k)$ change continuously with respect to k. This analysis holds for arbitrary activation gradients $a(x)$. Using a slight abuse of notation, we write $\bar{k}$ for the average value as well as for the vector $(\bar{k}, \dots, \bar{k})$. Also, using standard asymptotic notation, given a vector v and a scalar-valued function $R(v)$ we say that $R(v) = o(v)$ if $R(v)/|v| \to 0$ as $v \to 0$.

Proposition 1 *Given $k = (k_1, \dots, k_n)$, set $\bar{k} = (k_1 + \dots + k_n)/n$, and $\Delta k_i = k_i - \bar{k}$. Then*

$$g(u, k) = g(u, \bar{k}) + o(\Delta k).$$

Proof. Notice that

$$\Delta k_1 + \dots + \Delta k_n = k_1 + \dots + k_n - n\bar{k} = 0.$$

Since the function $g(u, \bar{k})$ is symmetric on each of the variables k_i,

$$\frac{\partial}{\partial k_i} g(u, \bar{k}) = \frac{\partial}{\partial k_j} g(u, \bar{k}) = c$$

for some real number $c = c(\bar{k})$ and all $i, j = 1 \dots n$. The result follows from calculating the directional derivative

$$D_{\Delta k}\ g(u, \bar{k}) = \nabla_k\ g(u, \bar{k}) \circ \Delta k = c(\Delta k_1 + \dots + \Delta k_n) = 0.$$

□

We denote by g_{min} and g_{max} the minimum and maximum values of the dose response, respectively. Since $g(0,k) = a_0$ and $g(\cdot,k)$ saturates at a_n, we have $g_{min} = a_0$, $g_{max} = a_n$.

Theorem 1 *Let $H(k)$ be the generalized Hill coefficient of $g(\cdot,k)$, and suppose $g_{min} < 0.1g_{max}$. Given $k = (k_1,\ldots,k_n)$, let $\bar{k} = (k_1+\ldots+k_n)/n$, $\Delta k_i = k_i - \bar{k}$. Then*

$$H(k) = H(\bar{k}) + o(\Delta k).$$

Proof. Since $g_{min} < 0.1g_{max}$, there exist $u_1(k), u_2(k)$ such that $g(u_1(k),k) = 0.1g_{max}$ and $g(u_2(k),k) = 0.9g_{max}$, respectively. In fact $u_1(k)$, $u_2(k)$ are differentiable by the Implicit Function Theorem, since $\partial/\partial u\ g(u,k) > 0$. Using the chain rule to calculate the gradient of $u_i(k)$, we obtain

$$D_{\Delta k}\ u_i(\bar{k}) = \nabla u_i(\bar{k}) \circ \Delta k = -\left(\frac{\partial g}{\partial u}\right)^{-1} \nabla_k\ g(u_i(\bar{k}),\bar{k}) \circ \Delta k = 0$$

by the proof of Proposition 1. Now notice that $H(k) = \ln 81/\ln \frac{u_2(k)}{u_1(k)}$. Again using the chain rule we calculate

$$D_{\Delta k} H(\bar{k}) = \nabla H \circ \Delta k = \left(\frac{\partial H}{\partial u_1}\nabla u_1(\bar{k}) + \frac{\partial H}{\partial u_2}\nabla u_2(\bar{k})\right) \circ \Delta k = 0.$$

□

One could argue that if the parameters k_i are varied from a given value $\bar{k}$, then the average of the new parameters can deviate from $\bar{k}$, leading to changes in the dose response. However, as we will see below, the Hill coefficient of a dose response is invariant under parameter rescaling. In particular H has the same value given identical parameters $k_i = \bar{k}$, regardless of the actual value of $\bar{k}$.

4 Maximum Hill coefficient

We prove that the Hill coefficient of the non-cooperative dose response $g(\cdot,k)$ (2) is maximized when all dissociation rates k_i are identical to each other, another basic feature observed in simulations (Theorem 2). It might seem a daunting task to maximize the function $H(k) = \ln 81/\ln \frac{g(\cdot,k)^{-1}(0.9g_{max})}{g(\cdot,k)^{-1}(0.1g_{max})}$. But this problem can be simplified through a series of steps as outlined below.

First, we will assume without loss of generality that k satisfies $g(1,k) = 0.1g_{max}$. Then, rather than maximizing the Hill coefficient $H(\cdot)$, we focus on maximizing $g(c,\cdot)$ for a fixed value $c > 1$. It will turn out that $g(c,\cdot)$ is maximized when all dissociation constants are identical to each other, $k_i = \bar{k}$, regardless of the value of c (Figure 3a). That is, $g(c,k) < g(c,\bar{k})$ for any $k \neq \bar{k}$ and any $c > 1$. Finally, notice that maximizing the dose response for all $c > 1$ given $g(1,k) = 0.1g_{max}$ is the same as minimizing $g(\cdot,k)^{-1}(0.9g_{max})$, which in turn maximizes the Hill coefficient.

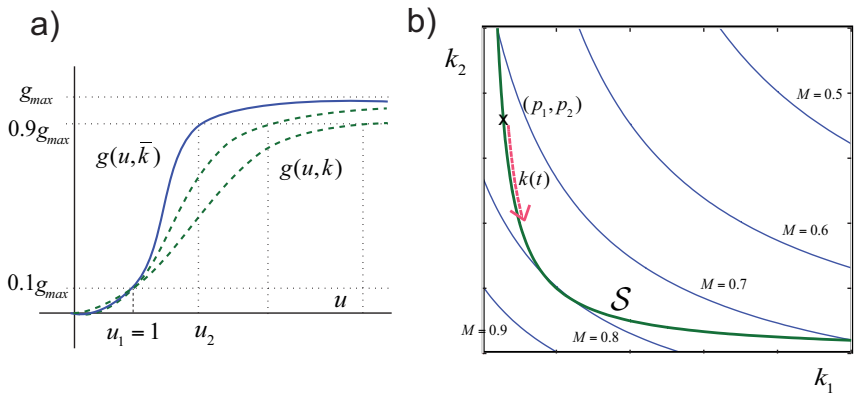

Fig. 3 a: In order to maximize the Hill coefficient, assume that $g(1,k) = 0.1g_{max}$ and maximize $g(c,k)$ for $c > 1$. **b:** Proof of Proposition 3 for the case $n = 2$, $a_0 = 0, a_1 = 0.5, a_2 = 1$.

Before proving the main results, we state and prove several technical lemmas. The following lemma shows that when k is multiplied by a scalar α, the dose response $g(\cdot,k)$ is also rescaled by α, in particular its Hill coefficient is unchanged.

Lemma 2 *For every $u \geq 0$, $\alpha > 0$, and $k = (k_1,\dots,k_n)$, it holds that $g(\alpha u, \alpha k) = g(u,k)$. That is, the function $g(u,k)$ is homogeneous of degree 0.*

Proof. Notice $P(\alpha u, \alpha k) = \alpha^n P(u,k)$, $e_j(\alpha k) = \alpha^j e_j(k)$, and

$$g(\alpha u, \alpha k) = \frac{1}{P(\alpha u, \alpha k)} \sum_{i=0}^{n} a_i \alpha^{n-i} e_{n-i}(k)(\alpha u)^i = g(u,k).$$

□

Define the set $\mathcal{S} = \{k \in [0,\infty)^n | g(1,k) = 0.1g_{max}\}$. We show in the following result that optimization of the Hill coefficient can be carried out without loss of generality on this set.

Lemma 3 *For any nonzero $k \in [0,\infty)^n$ there exists $k' \in \mathcal{S}$ such that $H(k') = H(k)$.*

Proof. Let $u_1, u_2 \in \mathbb{R}^+$ be such that $g(u_1,k) = 0.1g_{max}$, $g(u_2,k) = 0.9g_{max}$. Set $k' = k/u_1$, and notice that $g(1,k') = 0.1g_{max}$, $g(u_2/u_1,k') = 0.9g_{max}$ by homogeneity. Therefore $k' \in \mathcal{S}$ and both dose responses have the same Hill coefficient $H = \ln(81)/\ln(u_2/u_1)$. □

Lemma 4 *There exists a unique scalar $\bar{k} > 0$ such that $(\bar{k},\dots,\bar{k}) \in \mathcal{S}$.*

Proof. To prove existence, consider the set of dissociation rates $k_i = 1$, $i = 1,\dots,n$, and rescale as in the proof of Lemma 2 to find $\bar{k} = 1/u_1$. For uniqueness, suppose $g(1,(\mu\dots\mu)) = g(1,(\nu\dots\nu)) = 0.1g_{max}$, for $0 < \mu < \nu$. Then

$g(\nu/\mu,(\nu\dots\nu)) = g(1,(\nu\dots\nu))$, which violates strict monotonicity of the dose response. □

Define $\Delta a_i := a_i - a_{i-1}$, for $i = 1\dots n$, and introduce the following notation:

$$e_i^{\gamma=0}(k) := \sum_{\substack{|I|=i\\ I_\gamma=0}} \prod_{I_j=1} k_j, \qquad e_i^{\gamma,\delta=0} := \sum_{\substack{|I|=i\\ I_\gamma=0, I_\delta=0}} \prod_{I_j=1} k_j,$$

and similarly for $e_i^{\gamma=1}(k)$ etc. The following technical result is based on the Newton inequalities and will be used in the proof of Proposition 2.

Lemma 5 *The function $T(x)$ defined by*

$$T(x) = \frac{\sum_{i=1}^{n} \Delta a_i x^i e_{n-i}^{l,m=0}}{\sum_{i=1}^{n} \Delta a_i x^i e_{n-i-1}^{l,m=0}}$$

is strictly increasing on $(0,\infty)$.

Proof.

Define $f_1(x), f_2(x)$ as the numerator and denominator of $T(x)$, respectively. Since $T' = (f_1' f_2 - f_2' f_1)/(f_2)^2$, we calculate $f_1'(x)f_2(x) - f_2'(x)f_1(x)$:

$$\sum_{i=1}^{n} \Delta a_i\, i\, x^{i-1} e_{n-i}^{l,m=0} \sum_{j=1}^{n} \Delta a_j\ x^j e_{n-j-1}^{l,m=0} - \sum_{i=1}^{n} \Delta a_i\ x^i e_{n-i}^{l,m=0} \sum_{j=1}^{n} \Delta a_j\, j\, x^{j-1} e_{n-j-1}^{l,m=0}$$

$$= \sum_{i,j=1}^{n} \Delta a_i \Delta a_j x^{i+j-1} e_{n-i}^{l,m=0} e_{n-j-1}^{l,m=0} (i-j).$$

Denoting each sumand of this expression by S_{ij}, and noting that $S_{ii} = 0$ for all $i = 1\dots n$, we write the sum as

$$\sum_{i=1}^{n}\sum_{j=1}^{i-1} S_{ij} + S_{ji} = \sum_{i=1}^{n}\sum_{j=1}^{i-1} \Delta a_i \Delta a_j x^{i+j-1}(i-j)\left[e_{n-i}^{l,m=0} e_{n-j-1}^{l,m=0} - e_{n-j}^{l,m=0} e_{n-i-1}^{l,m=0}\right].$$

Since each $e_*^{l,m=0}$ can be described as a full symmetric polynomial on the variables $\{k_1,\dots,k_n\} - \{k_l, k_m\}$, it follows from Newton's inequalities [18] that $e_{n-i}^{l,m=0} e_{n-j-1}^{l,m=0} - e_{n-j}^{l,m=0} e_{n-i-1}^{l,m=0} > 0$ whenever $i > j$. Thus $T'(x) > 0$ and the proof is complete. □

As mentioned before, we will use the notation $\bar{k}$ to denote the scalar as well as its associated vector for convenience. We now fix a value $c > 1$ and work to maximize $g(c,k)$ among all $k \in \mathcal{S}$. We show that the largest value of $g(c,k)$ is obtained when all rates k_i are identical to each other. Define the functions

$$M(k) := g(c,k), \quad L(k) := g(1,k).$$

The constraint $k \in \mathcal{S}$ can now be rephrased as $L(k) = 0.1 g_{max}$. The next result provides the key ingredient for the maximization of $g(c, \cdot)$.

Proposition 2 *For any nonzero $k \in [0, \infty)^n$ such that $k_l < k_m$,*

$$\frac{\partial M}{\partial k_l} \frac{\partial L}{\partial k_m} < \frac{\partial M}{\partial k_m} \frac{\partial L}{\partial k_l}. \tag{3}$$

Proof: Calculate

$$\begin{aligned} \frac{\partial M}{\partial k_l} = \frac{\partial}{\partial k_l} g(c,k) &= \left(\frac{\partial}{\partial k_l} \frac{1}{P(c,k)} \right) \sum_{i=0}^{n} a_i c^i e_{n-i}(k) \quad + \quad \frac{1}{P(c,k)} \frac{\partial}{\partial k_l} \sum_{i=0}^{n} a_i c^i e_{n-i}(k) \\ &= \frac{-M}{(c+k_l)} + \frac{1}{P(c,k)} \sum_{i=0}^{n-1} a_i c^i \frac{1}{k_l} e_{n-i}^{l=1}(k). \end{aligned}$$

Multiplying each side by $k_l(c+k_l)P(c,k)$,

$$\begin{aligned} k_l(c+k_l)P(c,k)\frac{\partial M}{\partial k_l} =& - k_l \sum_{i=0}^{n-1} a_i c^i e_{n-i}^{l=1} - k_l \sum_{i=1}^{n} a_i c^i e_{n-i}^{l=0} + (c+k_l) \sum_{i=0}^{n-1} a_i c^i e_{n-i}^{l=1} \\ =& k_l \sum_{i=1}^{n} (a_{i-1} - a_i) c^i e_{n-i}^{l=0}, \end{aligned}$$

where the final equality is obtained by a re-indexing of terms and the use of the identity $k_l \, e_s^{l=0} = e_{s+1}^{l=1}$. Thus,

$$\frac{\partial M}{\partial k_l} = \frac{-1}{(c+k_l)P(c,k)} \sum_{i=1}^{n} \Delta a_i c^i e_{n-i}^{l=0} < 0. \tag{4}$$

A similar calculation applies for $\partial L / \partial k_l$ by setting $c = 1$. Let us examine the term $\sum_{i=1}^{n} \Delta a_i c^i e_{n-i}^{l=0}$, which does not contain k_l, and rewrite it as a sum of those terms involving k_m and those that do not:

$$\sum_{i=1}^{n} \Delta a_i c^i e_{n-i}^{l=0} = \sum_{i=1}^{n} \Delta a_i c^i e_{n-i}^{l,m=0} + k_m \sum_{i=1}^{n} \Delta a_i c^i e_{n-i-1}^{l,m=0} =: A + k_m B.$$

Carrying out the same procedure to define C and D as above for $c = 1$, we can write

$$\frac{\frac{\partial M}{\partial k_l}}{\frac{\partial L}{\partial k_l}} = \frac{(1+k_l) \sum_{i=1}^{n} \Delta a_i c^i e_{n-i}^{l=0}}{(c+k_l) \sum_{i=1}^{n} \Delta a_i e_{n-i}^{l=0}} = \left(\frac{1+k_l}{c+k_l} \right) \left(\frac{A + B k_m}{C + D k_m} \right). \tag{5}$$

None of the expressions A, B, C, D contain k_l or k_m, and one can repeat the same procedure to calculate the partial derivatives with respect to k_m, which results in exchanging k_l and k_m on the right hand side of (5). Now, notice that (3) is equivalent to

$$\frac{\frac{\partial M}{\partial k_l}}{\frac{\partial L}{\partial k_l}} < \frac{\frac{\partial M}{\partial k_m}}{\frac{\partial L}{\partial k_m}} \tag{6}$$

$$\left(\frac{1+k_l}{c+k_l}\right)\left(\frac{A+Bk_m}{C+Dk_m}\right) < \left(\frac{1+k_m}{c+k_m}\right)\left(\frac{A+Bk_l}{C+Dk_l}\right) \tag{7}$$

$$\left(\frac{1+k_l}{c+k_l}\right)\left(\frac{C+Dk_l}{A+Bk_l}\right) < \left(\frac{1+k_m}{c+k_m}\right)\left(\frac{C+Dk_m}{A+Bk_m}\right). \tag{8}$$

The result follows immediately from showing that the following function is strictly increasing:

$$F(x) = \left(\frac{1+x}{c+x}\right)\left(\frac{C+Dx}{A+Bx}\right)$$

It is easy to see that the term $\left(\frac{1+x}{c+x}\right)$ is an increasing function, since $c > 1$. So, to prove that F is strictly increasing, we need only show that $\left(\frac{C+Dx}{A+Bx}\right)$ is a strictly increasing function. This is true if and only if $AD - BC > 0$ or $A/B > C/D$. Using Lemma 5 it follows that

$$A/B = T(c) > T(1) = C/D,$$

which completes the proof. □

A similar approach as that followed in Proposition 2 can be carried out from the Lagrange multiplier equations, $\partial M/\partial k_l = \lambda \partial L/\partial k_l$, $i = 1, \ldots, n$. Solving for λ we find $\lambda = (\partial M/\partial k_l)/(\partial L/\partial k_l)$, for all l, and hence

$$\frac{\frac{\partial M}{\partial k_l}}{\frac{\partial L}{\partial k_l}} = \frac{\frac{\partial M}{\partial k_m}}{\frac{\partial L}{\partial k_m}},$$

for every l, m. The same argument as in equations (7), (8) shows that $F(k_l) = F(k_m)$ for all l, m. Since F is one-to-one, this implies that $k_l = \bar{k}$ for all l. Notice, however, that we cannot conclude from this equation whether the critical point is a maximum or a minimum, hence the stronger result in Proposition 2.

We use Proposition 2 to show that the point $\bar{k}$ is in fact the global maximum of M on $\mathcal{S}$.

Proposition 3 *Let $p \in \mathcal{S}$, $p \neq \bar{k}$. Then there exists $p' \in S$ such that $M(p) < M(p')$. In particular p cannot be a maximum of M in $\mathcal{S}$.*

Proof: Suppose that $p \in \mathcal{S}$ and l, m are such that $p_l < p_m$. We will prove that there exists a differentiable function $k(t) \in \mathcal{S}$ defined for all t on a neighborhood of 0 such that $k(0) = p$ and $M(k(t))$ is strictly increasing. This in particular will prove the proposition.

Without loss of generality we can assume $i = 1, j = 2$ since $M(k)$ is a symmetric function. See Figure 3b for an illustration. Consider the set of all points (k_1, k_2) that satisfy the equation

$$L(k_1, k_2, p_3, \dots, p_n) = 0.1 g_{max}.$$

We describe this set as the graph of a function $k_2 = r(k_1)$ in some neighborhood of the point (p_1, p_2). The Implicit Function Theorem guarantees that the function $r(t)$ exists locally around p as long as

$$\left.\frac{\partial L(k_1, k_2, p_3, \dots, p_n)}{\partial k_2}\right|_{(p1,p2)} \neq 0.$$

By (4) this derivative is always negative, so the condition holds. Define

$$k(t) = (p_1 + t, r(t), p_3, \dots, p_n),$$

for t in a neighborhood of 0. Let us differentiate both sides of the equation $L(k(t)) = 0.1 g_{max}$:

$$0 = \frac{\partial L}{\partial k_1}\frac{dk_1}{dt} + \frac{\partial L}{\partial k_2}\frac{dk_2}{dt} = \frac{\partial L}{\partial k_1} + \frac{\partial L}{\partial k_2} r'(t).$$

Solving for $r'(t)$, we obtain

$$r'(t) = \frac{-\frac{\partial L}{\partial k_1}}{\frac{\partial L}{\partial k_2}}.$$

Notice that $r'(t)$ is negative since both partial derivatives of L are negative. Now differentiate $M(k(t))$ with respect to t to obtain

$$\frac{dM(k(t))}{dt} = \frac{\partial M}{\partial k_1} + \frac{\partial M}{\partial k_2} r'(t) = \frac{\partial M}{\partial k_1} + \frac{\partial M}{\partial k_2}\frac{-\frac{\partial L}{\partial k_1}}{\frac{\partial L}{\partial k_2}}.$$

Then $\frac{d}{dt} M(k(t)) > 0$ is equivalent to $\frac{\partial M}{\partial k_1}\frac{\partial L}{\partial k_2} < \frac{\partial M}{\partial k_2}\frac{\partial L}{\partial k_1}$, which holds by Proposition 2. □

If $\mathcal{S}$ is bounded (and therefore compact), M must have a maximum value on $\mathcal{S}$, and by the above proposition $\bar{k}$ maximizes M. For the case of unbounded $\mathcal{S}$ the same conclusion follows directly from the next result.

Proposition 4 *The real function*

$$\sigma(x) = max\{M(p) \,|\, p \in \mathcal{S},\ ||p||_\infty = x\}$$

is strictly decreasing for all $x > \bar{k}$.

Proof: Fix $x > \bar{k}$. The set $A = \{M(p) \,|\, p \in \mathcal{S},\ ||p||_\infty = x\}$ is bounded and closed, and M achieves its maximum value on A at some point $p \in A$.

Since $||p||_\infty > \bar{k}$, there exists m such that $||p||_\infty = p_m > \bar{k}$. Also, there exists l such that $p_l < p_m$. By Proposition 2, there exists $k(t)$ on $\mathcal{S}$ such that $k(0) = p$ and $M(k(t))$ is strictly increasing. Moreover it is clear from the proof of the proposition that $||k(t)||_\infty \leq ||p||_\infty$ for all sufficiently small $t > 0$. In fact $||k(t)||_\infty \neq ||p||_\infty$ for all small $t > 0$ by the choice of p, so the inequality is strict. It follows that there exists $\epsilon > 0$ such that for all $y \in (x - \epsilon, x)$, $\sigma(y) > \sigma(x)$. This implies that σ is strictly decreasing for any $x > \bar{k}$.

Theorem 2 *Suppose $k \in [0, \infty)^n$, $k \neq 0$. Then $H(k) \leq H(1, \ldots, 1)$. Moreover, equality holds if and only if k is a scalar multiple of $(1, \ldots, 1)$.*

Proof.

First, notice that $H(1, \ldots, 1) = H(\bar{k})$, since rescaling $\bar{k}$ merely rescales the dose response by Lemma 2 and doesn't alter the Hill coefficient. By the same argument, the Hill coefficient doesn't change under any scalar multiple of $(1, \ldots, 1)$. Given $k \in [0, \infty)^n$, $k \neq 0$, we can assume without loss of generality that $k \in \mathcal{S}$ by Lemma 3.

Fix $c > 1$. By Proposition 3 and Proposition 4, $\bar{k}$ maximizes the function M on $\mathcal{S}$. In particular $g(c, k) = M(k) < M(\bar{k}) = g(c, \bar{k})$. Since $g(1, k) = g(1, \bar{k}) = 0.1 g_{max}$, and $c > 1$ is arbitrary, it follows that $H(k) < H(\bar{k})$.

□

5 Site Variability in the MWC Model

We describe a version of the Monod-Wyman-Changeux model that allows for site-to-site variability and generalizes the original MWC system. We carry out a similar computational analysis as in previous sections, by calculating the effect of site variability in the ultrasensitivity of the dose response.

The multisite protein in this system can also be in one of 2^n possible modification states, described by a binary vector I of length n. For instance, $I = (1, 0, 1)$ indicates that the first and third positions are modified. A basic assumption of the MWC model is that the multisite substrate can be in one of two different states, one of which has a larger ligand affinity than the other. The lower affinity state is called *tense* and denoted by T_I, and the higher affinity state R_I is called *relaxed* (Figure 4a).

In Figure 4b we describe the reactions of this system, each of which takes place at a linear rate to form the system of differential equations. For a given index I, suppose that a new index J has magnitude $|J| = |I| + 1$ and is such that $I(i) = 0$, $J(i) = 1$, and $I(m) = J(m)$ for every $m \neq i$. Protein R_I can be reversibly modified at site i to obtain protein R_J, and similarly for T_I, T_J. The rates of transition between relaxed and tense forms are such that the system satisfies detailed balance at steady state [5]. Constant α_i represents the binding affinity of the ligand to the i-th site. Since the system is detail balanced and every reaction is in equilibrium at steady state, we set the reverse reaction

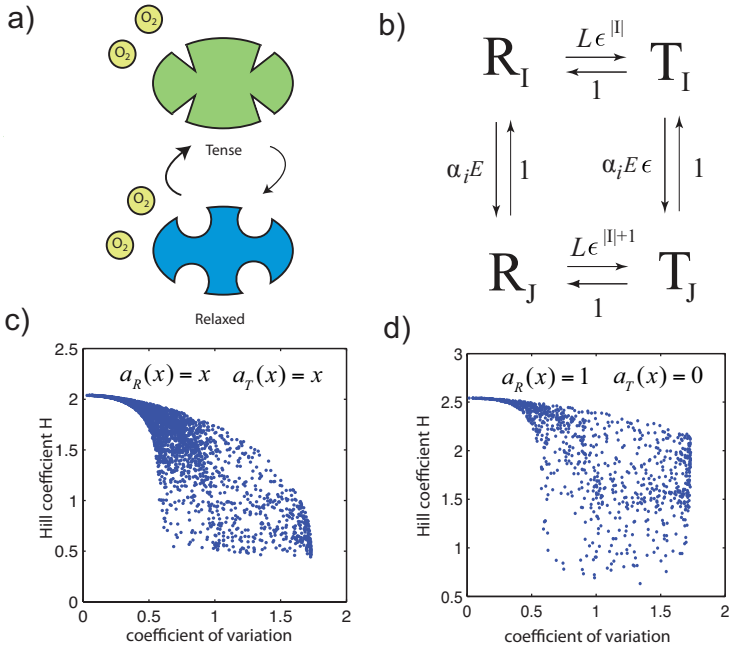

Fig. 4 **a:** In the MWC model a multisite protein changes allosterically between two states with different ligand affinity. **b:** The reaction network of the generalized MWC model, where I is any index vector and J is the result of adding a single modification in position i. **c,d:** Ultrasensitivity of dose response for MWC model under parameter variations, for $n = 4$, $\epsilon = 0.01$, $L = 100$, using two different sets of activation functions.

rates to 1 without loss of generality. If parameters for reverse reactions are added, then α_i can be interpreted as the ratio of forward to reverse rates.

Using detailed balance of every reaction at steady state, we obtain three distinct steady-state equations for the system:

$$R_J = \alpha_i E R_I, \qquad T_J = \alpha_i \epsilon E T_I, \qquad T_I = L\epsilon^{|I|} R_I,$$

for I and J as described above. By induction it is clear that

$$R_I = E^{|I|} R_0 \prod_{I_j=1} \alpha_j, \qquad T_I = L\epsilon^{|I|} R_I = L\epsilon^{|I|} E^{|I|} R_0 \prod_{I_j=1} \alpha_j,$$

for every index vector I. The total amount of protein in the system can be calculated as

$$S_T = \sum_I R_I + T_I = R_0 \sum_I (1 + L\epsilon^{|I|}) E^{|I|} \prod_{I_j=1} \alpha_j = R_0 \sum_{i=0}^{n} (1 + L\epsilon^i) E^i e_i(\alpha).$$

This can be used to solve for R_I as a function of the input E and the model parameters,

$$R_I = E^{|I|} R_0 \prod_{I_i=1} \alpha_i = S_T E^{|I|} \prod_{I_i=1} \alpha_i \left(\sum_{i=0}^{n} (1 + L\epsilon^i) E^i e_i(\alpha) \right)^{-1}.$$

Similarly for the tense protein T_I. This allows to calculate the dose response of the system,

$$g(E, \alpha) = \frac{1}{S_T} \sum_I a_R(|I|/n) R_I + \sum_I a_T(|I|/n) T_I,$$

as a function of E, α. Here the activity functions $a_R(x), a_T(x)$ generalize the function defined for the non-cooperative model. In the original MWC model, for example, the purpose of the system is to transport oxygen, and hence activation is proportional to modification, $a_R(x) = a_T(x) = x$. If however this is a modification system such as a multisite receptor complex, it is possible that only one of the two forms, tense or relaxed, is active for downstream signaling, so that e.g. $a_R(x) = 1,\ a_T(x) = 0$.

The graphs in Figure 4c,d describe the ultrasensitive behavior of the generalized MWC system under parameter variations. Both graphs are carried out for $n = 4$. In the first graph the activation functions are $a_R(x) = a_T(x) = x$, i.e. it considers parameter variations of the well known MWC model. The second graph assumes that only the relaxed state is active, $a_R(x) = 1,\ a_T(x) = 0$, as might be the case in a signaling system. Both graphs are ultrasensitive, and in both cases the ultrasensitive decreases with increasing parameter noise, although the second graph appears to be more robust to parameter variations.

6 Conclusions

It is interesting that in all the simulations of this paper the Hill coefficient is robust with respect to changes in modification rates up to a coefficient of variation of around 0.5. This holds for both the independent multisite model and for the MWC model, as well as for different values of the number of sites n. Not much is known about the site-to-site variability of modification sites in actual experimental systems, in large part due to the difficulties of carrying out such experiments. But in Harvey et al [10] the fractional occupancy of Swe1 sites was measured after stimulation of the protein with the Cdk1 kinase (see Table 1 in that reference), and the results are largely consistent with this rough estimate. This coefficient of variation is also consistent with data on the Gli1 multisite protein (Lee Bardwell, personal communication). Once again, the idea is that as proteins are allowed to evolve and vary, evolutionary forces counteract variability only once this limit is reached.

The mathematical proofs of this paper, while somewhat elaborate, use only elementary multivariable calculus techniques, together with symmetric

polynomials and Newton's inequalities, first published in 1732. A number of recent papers in the literature have also used algebraic techniques to analyze post translational modification systems of chemical reactions using algebraic methods [6,20,30], whether to establish bounds on the shape of dose responses or update standard methods such as zero-order ultrasensitivity in the context of more realistic assumptions.

The result we show in this paper for nonsequential systems contrasts with the situation for the sequential case, which was discussed in the paper by Wang, Nie and Enciso [28]. In that paper the efficiency of the i-th modification was described with the parameter λ_i, and it was shown that certain combinations of nonidentical parameters λ_i could yield a higher ultrasensitivity than when all λ_i are equal (see Figure 2C and Figure S1 of that paper). Specifically, ultrasensitivity becomes larger when the last few modifications have a higher λ_i than the first few modifications, which creates an effect very similar to cooperative behavior. In the nonsequential case this behavior cannot be replicated because the sites are unordered, and therefore one cannot guarantee that the most efficient modifications take place last. If anything the opposite effect is observed, in that the most efficient modifications naturally take place before the least efficient ones, which intuitively explains why varying the modification rates reduces ultrasensitivity.

Site-to-site variability has been used in the past to facilitate other phenomena such as bistability [1,9,26] or in the analysis of multisite dynamics and dose responses [2,8]. For instance, in [1] it is shown that the mere presence of a protein scaffold that binds to the multisite protein can significantly increase the likelihood of bistable behavior, when each modification rate is randomized. Notice that the systems considered in this paper are monostable, which follows from the derivation of the (unique) steady state concentrations for each value of the input. Bistability would require additional complexity such as Michaelis-Menten enzymatic reactions, additional feedback loops, or additional proteins such as a scaffold. At the same time, the presence of ultrasensitive behavior can significantly facilitate bistability once such additional elements are included.

Acknowledgments

We would like to thank Michael Cranston and Lee Bardwell for their help and suggestions. This material is based upon work supported by the National Science Foundation under Grants Nos. DMS-1122478 and 1129008.

References

1. Chan, C., Liu, X., Wang, L., Bardwell, L., Nie, Q., Enciso, G.: Protein scaffolds can enhance the bistability of multisite phosphorylation. PLoS Comp. Biol. **8**, 1–9 (2012)
2. Conradi, C., Shiu, A.: A global convergence result for processive multisite phosphorylation systems. Bull. Math. Biol. **77**(1), 126–155 (2015)

3. Enciso, G.: Multisite mechanisms for ultrasensitivity in signal transduction. In: Lecture Notes in Mathematics, vol. 2102, chap. 6. Springer Verlag (2013)
4. Enciso, G., Kellogg, D., Vargas, A.: Compact modeling of allosteric multisite proteins: Application to a cell size checkpoint. PLoS Comp. Biol. **10**(2), 1–12 (2014)
5. Feinberg, M.: Necessary and sufficient conditions for detailed balancing in mass action systems of arbitrary complexity. Chem. Eng. Sci. **44**(9), 1819–1827 (1989)
6. Feliu, E., Knudsen, M., Andersen, L., Wiuf, C.: An algebraic approach to signaling cascades with n layers. Bull. Math. Biol. **74**(1), 45–72 (2012)
7. Goldbeter, A., Koshland, D.: An amplified sensitivity arising from covalent modification in biological systems. Proc. Natl. Acad. Sci. USA **78**, 6840–6844 (1981)
8. Gunawardena, J.: Multisite protein phosphorylation makes a good threshold but can be a poor switch. Proc. Natl. Acad. Sci. USA **102**, 14,617–14,622 (2005)
9. Harrington, H., Feliu, E., Wiuf, C., Stumpf, M.: Cellular compartments cause multistability and allow cells to process more information. Biophys. J. **104**, 1824–1831 (2013)
10. Harvey, S., Enciso, G., Dephoure, N., Gygi, S., Gunawardena, J., Kellogg, D.: A phosphatase threshold sets the level of Cdk1 activity in early mitosis in budding yeast. Mol. Biol. Cell **22**(19), 3595–3608 (2012)
11. Herzog, F., Hill, J.: The Bernstein polynomials for discontinuous functions. Amer. J. Math. **68**(1), 109–124 (1946)
12. Hill, A.: The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. Proceedings of the Physiological Society **40**(suppl.), iv–vii (1910)
13. Iakoucheva, L., Radivojac, P., Brown, C., O'Connor, T., Sikes, J., Obradovic, Z., Dunker, A.: The importance of intrinsic disorder for protein phosphorylation. Nucleic Acid Res. **32**, 1037–1049 (2004)
14. Koshland, D., Nemethy, G., Filmer, D.: Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry **5**(1), 365–385 (1966)
15. Lenz, P., Swain, P.: An entropic mechanism to generate highly cooperative and specific binding from protein phoshporylations. Current Biology **16**, 2150–2155 (2006)
16. Liu, X., Bardwell, L., Nie, Q.: A combination of multisite phosphorylation and substrate sequestration produces switch-like responses. Biophysical Journal **98**(8), 1396–1407 (2010)
17. Monod, J., Wyman, J., Changeux, J.P.: On the nature of allosteric transitions: a plausible model. Journal of Molecular Biology **12**, 88–118 (1965)
18. Niculescu, C.: A new look at Newton's inequalities. J. Inequal. Pure Appl. Math. **1**(2), 1–14 (2000)
19. Peng, K., Vucetic, S., Radivojac, P., Brown, C., Dunker, A., Obradovic, Z.: Optimizing long intrinsic disorder predictors with protein evolutionary information. J. Bioinform. Comput. Biol **3**(1), 35–60 (2005)
20. Perez-Millan, M., Dickenstein, A.: Implicit dose-response curves. J. Math. Biol. **70**(7), 1669–1684 (2014)
21. Ryerson, S., Enciso, G.: Ultrasensitivity in independent multisite systems. J. Math. Biol. **69**(4), 977–999 (2014)
22. Sarkar, M., Blair, D.: Chemotaxis signaling protein Che Y binds to the rotor protein FliN to control the direction of flagellar rotation in Escherichia coli. Proc. Natl. Acad. Sci USA **107**(20), 9370–9375 (2010)
23. Serber, Z., Ferrell, J.: Tuning bulk electrostatics to regulate protein function. Cell **128**, 441–444 (2007)
24. Sourjik, V., Berg, H.: Functional interactions between receptors in bacterial chemotaxis. Nature **428**, 437–441 (2004)
25. Strickfaden, S., Winters, M.J., Ben-Ari, G., Lamson, R., Tyers, M., Pryciak, P.: A mechanism for cell-cycle regulation of MAP kinase signaling in a yeast differentiation pathway. Cell **128**, 519–531 (2007)
26. Thomson, M., Gunawardena, J.: Unlimited multistability in multisite phosphorylation systems. Nature **460**, 274–277 (2009)
27. Vignali, M., Steger, D., Neely, K., Workman, J.: Distribution of acetylated histones resulting from Gal4-VP16 recruitment of SAGA and NuA4 complexes. EMBO Journal **19**(11), 2629–2640 (2000)
28. Wang, L., Nie, Q., Enciso, G.: Nonessential sites improve phosphorylation switch. Biophysical Letters **99**(10), 41–43 (2010)

29. Wright, P., Dyson, J.: Intrinsically disordered proteins in cellular signaling and regulation. Nature Rev. Mol. Cell. Biol. **16**, 18–29 (2015)
30. Xu, Y., Gunawardena, J.: Realistic enzymology for post-translational modification: Zero-order ultrasensitivity revisited. J. Theor. Biol. **311**, 139–152 (2012)