Non-Essential Sites Improve Phosphorylation Switch

Liming Wang, Qing Nie, and German Enciso*

Department of Mathematics

Center for Mathematical and Complex Biological Systems

Center for Complex Biological Systems

University of California

Irvine, CA 92697

*Corresponding author: enciso@uci.edu

Running title: non-essential sites improve switch

Keywords: multi-site phosphorylation, ultrasensitivity, Hill coefficient, Sic1

Subject of Categories: Signal Transduction

Total character count: ~ 22,600

Abstract

Multisite phosphorylation is a common form of post-translational protein regulation which has been used to increase the switch-like behavior of the protein response to increasing kinase concentrations. In this paper, we show that phosphorylation sites that are not necessary for protein activation can strongly enhance the switch-like response. In the case of unordered phosphorylation, this holds when a sufficient amount of sites are phosphorylated regardless of their exact position, as occurs in bulk electrostatic activation. We obtained analytic estimates for the Hill coefficient of the switch-like response, and we observed that a trade-off exists between the switch and the kinase threshold for activation. This finding suggests a possible evolutionary mechanism for the relatively large numbers of phosphorylation sites found in various proteins.

Introduction

Protein phosphorylation is a ubiquitous form of post-translational modification. Since covalently bound phosphate groups are strongly hydrophilic and negatively charged, they can activate or inhibit a protein by changing its conformation or the way it interacts with other proteins. This is a key element of countless molecular systems including many forms of signal transduction, the regulation of the cell cycle and protein degradation. Multiple phosphorylations are often observed on the same protein; for instance, the epidermal growth factor receptor (EGFR) protein becomes activated when it is phosphorylated at multiple tyrosine residues by another EGFR (Alberts, 2002). In addition, many proteins have a surprisingly large number of phosphorylation sites, such as the nine phosphorylation sites identified in the cyclin-dependent kinase inhibitor Sic1 (Verma *et al*, 1997), the more than 30 sites in EGFR and the more than two dozen in p53 (Gnad *et al*, 2007).

It has been argued that a large number of phosphorylation sites could form the basis for a strong switch-like response to an increase in the kinase concentration (Nash *et al*, 2001; Welcker *et al*, 2003). Such a property is desirable in situations where a clear all-or-none response is expected (such as cell division in response to a growth factor), or in the design of multi-stable systems (Angeli *et al*, 2004). However, Gunawardena recently (Gunawardena, 2005) showed that the input-output response of the fully phosphorylated protein is not switch-like. Several plausible mechanisms to enhance the switch-like response have been suggested, including increasing the rates of phosphorylation of the last few sites (Gunawardena, 2005), and allowing the protein to bind to the membrane (Serber and Ferrell, 2007) or scaffold proteins (Liu *et al*, 2009).

Recent experimental results showed that a partially phosphorylated substrate can be as active as a fully phosphorylated one. One example is yeast cell cycle regulator, Sic1, which has nine phosphorylation sites; among those nine sites, any combination of six phosphorylated residues is sufficient to trigger the onset of S phase (Deshaies and Ferrell, 2001; Nash *et al*, 2001; Orlicky *et al*, 2003). In this paper, we propose that the active substrates include those with at least k phosphorylated sites, out of a total of n sites (Box 1). The number *k* designates the *minimal activation number*. At first glance, the inclusion of partially phosphorylated substrates as active proteins seems futile since a system with minimal activation number k out of a total of n sites would be similar to a system of a total of k sites where full phosphorylation is required for activation. Rather surprisingly, the remaining $n-k$ sites markedly improve the ultrasensitivity of the dose-response curve (Fig. 1A). For example, when the minimal activation number *k* reaches about half the total number of sites, and binding and catalytic rates are assumed to be similar at all sites, the effective Hill coefficient is roughly equal to *k*. In contrast, when full phosphorylation is required for activation, the maximum Hill coefficient only approaches two, even for large numbers of sites. This Hill coefficient is also substantially larger than that found in a system based on multi-site phosphorylation and docking to a single receptor site (Klein *et al*, 2003).

In this study, we demonstrated the improvement of the switch-like response upon kinase phosphorylation of the substrate in a specific order (Box 1), such as is the case with auto-phosphorylation of fibroblast growth factor receptor 1 (FGFR1) kinase (Furdui *et al*, 2006), nuclear factor of activated T cells 1, (Crabtree and Olson, 2002), and glycogen synthase kinase 3 (Cohen and Frame, 2001). We also identified an enhancement of the switch-like response for unordered phosphorylation, in which the kinase and phosphatase can act on the substrate in a random order (Strickfaden *et al*, 2007). Finally, we have noticed that inclusion of more non-essential sites may increase ultrasensitivity at the cost of lowering the threshold, leading to a trade-off between the two key characteristics in dose response and suggesting that more non-essential sites may not be always better for creating dose responses.

Results

Better ultrasensitivity through non-essential phosphorylation sites

We first investigated the *perfect balanced* case in which the kinase to phosphatase efficiency, λ (see Text S1A for details) was equal for all phosphorylation sites. When half of the phosphorylation sites were non-essential, i.e. $k = (n+1)/2$, a direct analytical estimate showed the dose-response curve, defined as the steady state fraction of the active substrates, and taking the form of

$$\rho(u) = \frac{(\lambda u)^k}{1+(\lambda u)^k}, \qquad (1)$$

where u is the input given by the ratio of the steady state free kinase to phosphatase (Text S1A). Equation (1) is a Hill function with Hill coefficient k. By comparison, in the traditional model of k-site phosphorylation (Gunawardena, 2005), where only the fully phosphorylated substrate has been considered active, the effective Hill coefficient is then only $2k/(k+1)$, a number always smaller than two. This value was clearly much smaller than k, especially for large k.

For the case of any k and n, the effective Hill coefficient of the response curve, as estimated by the Goldbeter-Koshland Formula (Goldbeter and Koshland, 1981), became (Text S1A)

$$H_s(n,k) \approx 2k\left(1-\frac{k}{n+1}\right) = 2\alpha(1-\alpha)(n+1), \quad (2)$$

where $\alpha = k/(n+1)$. From equation (2), we concluded (Table 1):

1. When the total number of the phosphorylation sites is fixed, as the minimal activation number increases, the ultrasensitivity first increases then decreases. The maximum value is achieved at $k = n/2$ and $k = n/2+1$ when n is even, and at $k = (n+1)/2$ when n is odd (the trend of each curve in Fig. 1B).
2. When the minimal activation number is fixed, the ultrasensitivity is improved by including more non-essential phosphorylation sites (black and red point comparison for the same k in Fig. 1B).
3. For fixed site activation ratio α, the ultrasensitivity increases linearly in the total number of the phosphorylation sites (points on the same ray in Fig. 1B).

We next study the case wherein the relative kinase to phosphatase efficiency for each site i, denoted by λ_i, was allowed to be different. Through exploring large random sets of λ_i we found that, in most of the cases, for a fixed total number of phosphorylation sites, as the minimal activation number increased, the Hill coefficient first increased then decreased (black curves in Fig. 1C and Fig. S1); this trend was similar to that in the perfect balanced case (red curve in Fig. 1C and Fig. S1). In general, the Hill coefficient was found to be high when about half of the total sites were non-essential, whereas low Hill coefficients were more likely to be observed when very few or almost all phosphorylation sites were non-essential (Fig. 1C and Fig. S1). Nevertheless, we observed that the ultrasensitivity improved when partially phosphorylated substrates were considered as active (i.e., $1 < k < n$ vs. $k = n$).

As an illustration of equation (2), we consider the protein kinase Wee1 system in the *Xenopus* egg cell cycle (Kim *et al*, 2005). There, Wee1 is deactivated by Cdk1 through

phosphorylations on five SP/TP sites. It is known that the first three sites (Ser38, Thr53, Ser62) tend to be phosphorylated before the other two (Thr104, Thr150), although the other two are essential to the activity of Wee1 (Kim *et al*, 2005). When the first four sites are mutated, and only Thr150 can be phosphorylated, the activity is equivalent to $n = k = 1$. Thus, the response is hyperbolic with a Hill coefficient of one, and close to the observed experimental value of 1.1 (Kim and Ferrell, 2007). When the Thr104 site is mutated alone (Wee1-T104E), the Hill coefficient increased in experiments to 1.4 (Kim *et al*, 2007). Assuming for simplicity that the first three sites are phosphorylated sequentially, which corresponds in our model to the case of $n = k = 4$, one obtains a comparable value of $H_s(4,4) = 1.6$ under the perfect balance assumption.

Lower threshold as more non-essential phosphorylation sites are included

Roughly speaking, the threshold is the input value around which the response starts to increase significantly. One possible measurement of the threshold is EC10, which is defined as the input value when the output reaches 10% of its maximum (Liu *et al*, 2009).

Our analytical study on the perfect balanced case revealed the following (Table 1 and Text S1C):

1. For a fixed total number of phosphorylation sites, the threshold increases with respect to the minimal activation number, reaching a maximum value when only fully phosphorylated substrate is regarded as active (the trend of the blue and red curves in Fig. 1D).

2. For the fixed minimal activation number, as more non-essential sites are included, the threshold always decreases (blue and red curves comparison for the same k in Fig. 1D).

3. When the site activation ratio α is fixed, as the total site number (equivalently, the minimal activation number) increases, the threshold either increases ($\alpha < 0.9$) or decreases ($\alpha > 0.9$) depending on α.

The threshold plot of randomly chosen λ_i's showed the same trend as k increases for fixed n (black vs. red in Fig. 1D). That is, the threshold increased as fewer sites are considered active.

Similar ultrasensitivity and threshold properties under the unordered mechanism

Here, in the *unordered mechanism*, we assumed that *any* subset of at least *k* phosphorylated sites was sufficient to activate the protein, regardless of the exact position of the sites. The unordered mechanism seems to be especially applicable in the context of bulk electrostatics: when a protein is sufficiently phosphorylated, it may cease to bind to negatively charged or hydrophobic regions, such as the cell membrane, regardless of exactly which sites are involved. This can alter the activity of a protein, as illustrated by the Ste5 scaffold protein in yeast (Strickfaden *et al*, 2007). It has been shown that phosphorylation sites on the substrate tend to be located in poorly conserved (Brown *et al*, 2002) and predominantly disordered (Iakoucheva *et al*, 2004) regions, which suggested that the exact location of the sites can often be unessential for activation and may support mechanisms such as bulk electrostatics.

Because the number of phosphoforms grows exponentially in the total number of phosphorylation sites for the non-sequential case, it is more challenging for analysis.

In the perfect balanced case, the Hill coefficient of the steady state fraction of the active substrates was estimated by (Text S1C)

$$H_r(n,k) \approx 1.71\sqrt{k\left(1-\frac{k}{n+1}\right)} = 1.71\sqrt{\alpha(1-\alpha)}\sqrt{n+1}\,, \qquad (3)$$

which was approximately the square root of the Hill coefficient in the sequential case (Fig. S3A). Therefore, the first two conclusions of the sequential case still held (Fig. 2A and S3B), and the ultrasensitivity increase became $\sqrt{n+1}$ (Text S1D) instead of linear in regards to the total number of the phosphorylation sites.

The first two conclusions on the threshold under the sequential mechanism also carried over to the non-sequential case (Fig. 2B and S3C). In addition, we found that for fixed α, the threshold always increased in the total number of phosphorylation sites, regardless of the value of α (Text S1E). Please see Table 1 for a summary.

Trade-off between ultrasensitivity and threshold

We observed that the introduction of non-essential sites appeared to have opposite effects on the ultransensitivity and the threshold. When the total number of phosphorylation sites was fixed, and as the minimal activation number decreased, the ultrasensitivity first increased then decreased, whereas the threshold always decreased (Fig. 3). Since a good switch is expected to have both high ultrasensitivity and large threshold, there seemed to be an optimal range for the number of non-essential sites when the total number of sites was fixed.

In the example of Sic1, the total number of phosphorylation sites was nine. If the phosphorylation of Sic1 followed a sequential mechanism, then the perfect balanced condition $k = 5$ yielded the largest Hill coefficient (Fig. 3A). When the phosphorylation order of Sic1 was random, k could be further increased to raise the threshold to EC10 (Fig. 3B). Both of these scenarios were consistent with the value $k = 6$ observed from previously published experiments (Nash *et al*, 2001).

Conclusion and Discussion

In the present work, we have proposed a mechanism through the use of non-essential sites that could account for the high ultrasensitivity observed in many multisite phosphorylation systems. For given values of the total number of sites and the minimal number of phosphorylations for activation, we have obtained an estimate of the effective Hill coefficient under both sequential and non-sequential mechanisms. We have studied the effect of non-essential sites on both the effective Hill coefficient and the threshold of a dose-response curve. Our results suggest that increased amounts of non-essential phosphorylation sites improve the ultrasensitivity, but decrease the threshold (Table 1). Thus, to achieve a good activation switch, there is a balance between the number of non-essential sites and the total number of sites.

By decoupling the total number of sites from the number of phosphorylations needed for activation, we may look at the evolution of multisite phosphorylation, since a given mutation can easily change one without altering the other. For instance, consider a mutation which doubles the number of sites while leaving k unchanged. This change may increase ultrasensitivity dramatically (point **a** vs. point **b** in Fig. S4) while slightly decreasing the threshold. For instance, for $n = 10$ and $k = 5$, doubling n barely changed the threshold (blue vs. red at $k = 5$ in Fig. 1D).

To the extent that evolution in this system favored ultrasensitive behavior, such a mutation would be undoubtedly preferred as it would increase the robustness of the on-off switch. Moreover, an advantage is created for k to catch up. In the above example, when n increased to 20, the evolutionary pressure drives k to around 10 (point **b** vs. point **c** in Fig. S4). This process can continue to repeat itself over time, in the form of an arms race between the total number of sites and the minimal activation number. One can speculate that this could sometimes lead to a runaway increase in the number of sites.

Intuitively, why would an addition of sites without increasing the minimal activation number increase the switch-like behavior of a system? Imagine the phosphorylation of an individual protein as a biased random walk. At any given time the protein is in a state between 0 and n phosphorylations, following the rule of a random walk among these states. The propensities for phosphorylation and dephosphorylation are given by the constants E and F, respectively. Suppose that $n = 2k$. If E is even slightly larger than F, the bias in the random walk intuitively makes the protein spend most of its time in one of the states $S_k, \cdots, S_n$, since a biased random walk eventually moves with high probability in the direction of the bias. Similarly, if E is slightly smaller than F, it spends little time in this region. Now, one may show that the fraction of the time that the protein spends in those states is none other than the dose response for $u = E / F$. This would account for the ultrasensitive behavior (see Text S1B for more details).

Of course, the use of non-essential phosphorylation sites may work in combination with other mechanisms in order to achieve better ultrasensitivity. For instance, if the last few

phosphorylations take place at a faster rate in order to increase the switch-like behavior (a form of cooperativity (Gunawardena, 2005)), one can obtain an even higher Hill coefficient through the use non-essential sites (Fig. S5).

Acknowledgments

We would like to thank Karen Sachs for the many stimulating conversations, and Jim Ferrell and Lee Bardwell for useful advice. QN was supported by NIH grants R01GM75309, R01GM67247, P50GM76516 and NSF grant DMS-0917492.

References

Alberts B (2002) *Molecular biology of the cell*, 4th edn. New York: Garland Science.

Angeli D, Ferrell JE, Jr., Sontag ED (2004) Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proc Natl Acad Sci U S A* **101:** 1822-1827.

Brown CJ, Takayama S, Campen AM, Vise P, Marshall TW, Oldfield CJ, Williams CJ, Dunker AK (2002) Evolutionary rate heterogeneity in proteins with long disordered regions. *J Mol Evol* **55:** 104-110.

Cohen P, Frame S (2001) The renaissance of GSK3. *Nat Rev Mol Cell Biol* **2:** 769-776.

Crabtree GR, Olson EN (2002) NFAT signaling: choreographing the social lives of cells. *Cell* **109 Suppl:** S67-79.

Deshaies RJ, Ferrell JE, Jr. (2001) Multisite phosphorylation and the countdown to S phase. *Cell* **107:** 819-822.

Furdui CM, Lew ED, Schlessinger J, Anderson KS (2006) Autophosphorylation of FGFR1 kinase is mediated by a sequential and precisely ordered reaction. *Mol Cell* **21:** 711-717.

Gnad F, Ren S, Cox J, Olsen JV, Macek B, Oroshi M, Mann M (2007) PHOSIDA (phosphorylation site database): management, structural and evolutionary investigation, and prediction of phosphosites. *Genome Biol* **8:** R250.

Goldbeter A, Koshland DE, Jr. (1981) An amplified sensitivity arising from covalent modification in biological systems. *Proc Natl Acad Sci U S A* **78:** 6840-6844.

Gunawardena J (2005) Multisite protein phosphorylation makes a good threshold but can be a poor switch. *Proc Natl Acad Sci U S A* **102:** 14617-14622.

Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res* **32:** 1037-1049.

Kim SY, Ferrell JE, Jr. (2007) Substrate competition as a source of ultrasensitivity in the inactivation of Wee1. *Cell* **128:** 1133-1145.

Kim SY, Song EJ, Lee KJ, Ferrell JE, Jr. (2005) Multisite M-phase phosphorylation of Xenopus Wee1A. *Mol Cell Biol* **25:** 10580-10590.

Klein P, Pawson T, Tyers M (2003) Mathematical Modeling Suggests Cooperative Interactions between a Disordered Polyvalent Ligand and a Single Receptor Site. *Curr Biol* **13**: 1669-1678.

Liu XF, Bardwell L, Nie Q (2009) A combination of multisite phosphorylation and substrate sequestration produces switch-like responses. *Biophysical Journal*.

Nash P, Tang X, Orlicky S, Chen Q, Gertler FB, Mendenhall MD, Sicheri F, Pawson T, Tyers M (2001) Multisite phosphorylation of a CDK inhibitor sets a threshold for the onset of DNA replication. *Nature* **414:** 514-521.

Orlicky S, Tang X, Willems A, Tyers M, Sicheri F (2003) Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase. *Cell* **112:** 243-256.

Serber Z, Ferrell JE, Jr. (2007) Tuning bulk electrostatics to regulate protein function. *Cell* **128:** 441-444.

Strickfaden SC, Winters MJ, Ben-Ari G, Lamson RE, Tyers M, Pryciak PM (2007) A mechanism for cell-cycle regulation of MAP kinase signaling in a yeast differentiation pathway. *Cell* **128:** 519-531.

Verma R, Annan RS, Huddleston MJ, Carr SA, Reynard G, Deshaies RJ (1997) Phosphorylation of Sic1p by G1 Cdk required for its degradation and entry into S phase. *Science* **278:** 455-460.

Welcker M, Singer J, Loeb KR, Grim J, Bloecher A, Gurien-West M, Clurman BE, Roberts JM (2003) Multisite phosphorylation by Cdk2 and GSK3 controls cyclin E degradation. *Mol Cell* **12:** 381-392.

Figure Legends

Box 1: n-site phosphorylations and dephosphorylations.

(A) A schematic diagram of n-site phosphorylations and dephosphorylations under the sequential and distributive mechanism by kinase *E* and phosphatase *F*, respectively. Inside the black dashed box is the conventional definition of active substrates. The generalized definition of active substrates includes all substrates inside the red dashed box.

(B) The fraction of active substrates by the conventional definition at the steady state.

(C) The fraction of active substrates by the generalized definition at the steady state.

Figure 1: Ultrasensitivity and threshold under the sequential mechanism.

(A) Dose-response curves for fixed $k = 10$ and various n's.

(B) Comparing the Hill coefficients computed directly from the Goldbeter-Koshland formula (dots) and from equation (2) (curves) for different k's when $n = 5$ (red) and $n = 11$ (black). For example, point "a" denotes the Hill coefficient of the dose-response curve when $n = k = 5$; point "b": $n = 11, k = 5$; point "c": $n = 5, k = 4$.

(C) The Hill coefficients against different k's for ten sets of random λ_i's (black) and $\lambda_i = 1$ (red). Here, $n = 20$ and λ_i randomly varies between 0.1 and 10. See Fig. S1 for simulations of 100 samples of λ_i's.

(D) The threshold against different k's. Green curve: $n = 10, \lambda_i = 1$; red curve: $n = 20, \lambda_i = 1$; black curves: $n = 20$ and λ_i randomly varying between 0.1 and 10. See Fig. S2 for simulations of 100 samples of λ_i s.

Figure 2: Ultrasensitivity and threshold under the unordered mechanism.

(A) Comparing the Hill coefficients computed directly from the Goldbeter-Koshland formula (dots) and from equation (3) (curves) for different k's when $n = 20$ (red) and $n = 40$ (black).

(B) The threshold against different k's . Red: $n = 20$; black: $n = 40$. In both curves, $\lambda_i = 1$.

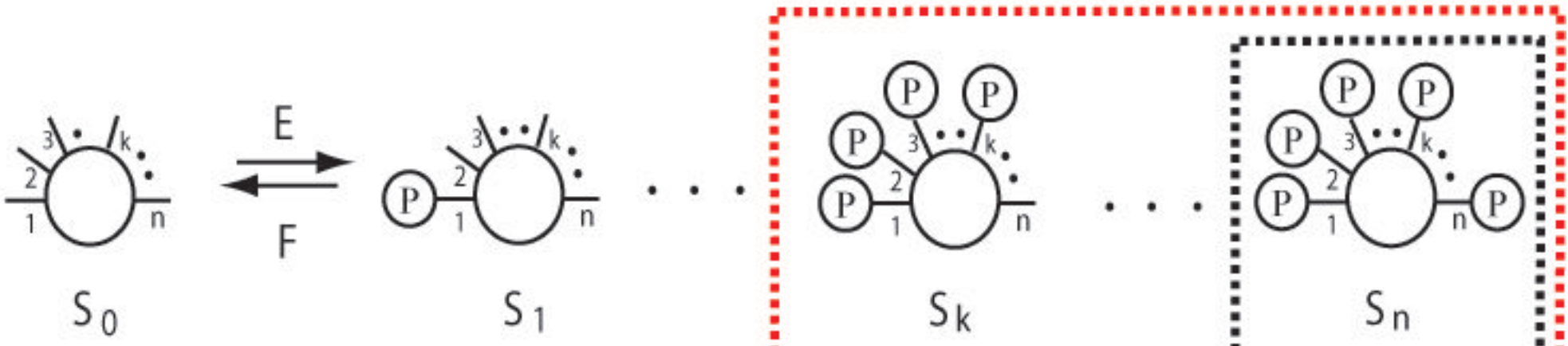
Figure 3: The trade-off between ultrasensitivity and threshold.

(A) Ultrasensitivity (black) and threshold (blue) under the sequential mechanism.

(B) Ultrasensitivity (black) and threshold (blue) under the unordered mechanism.

In both plots, $n = 9,\ \lambda_i = 1$.

A

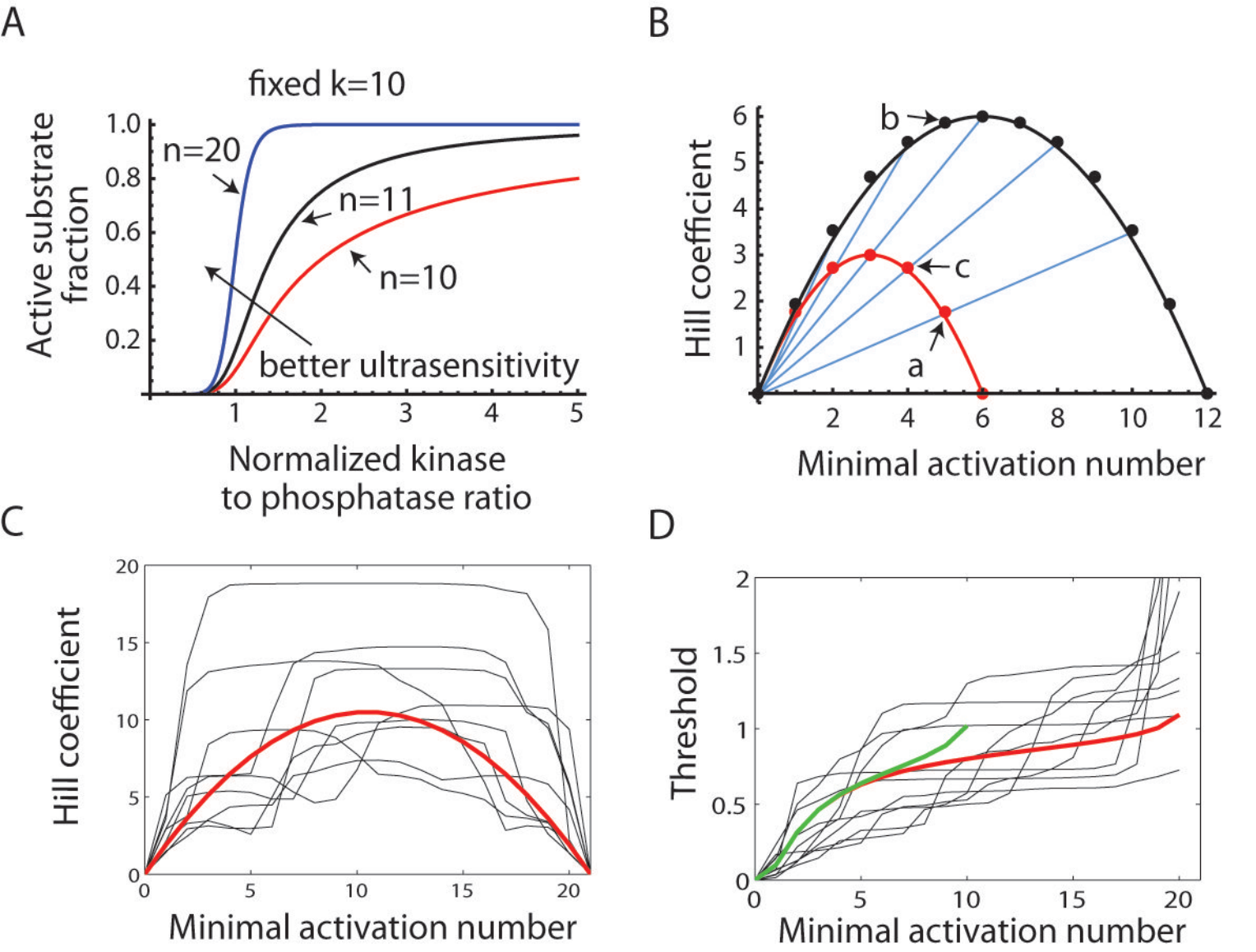


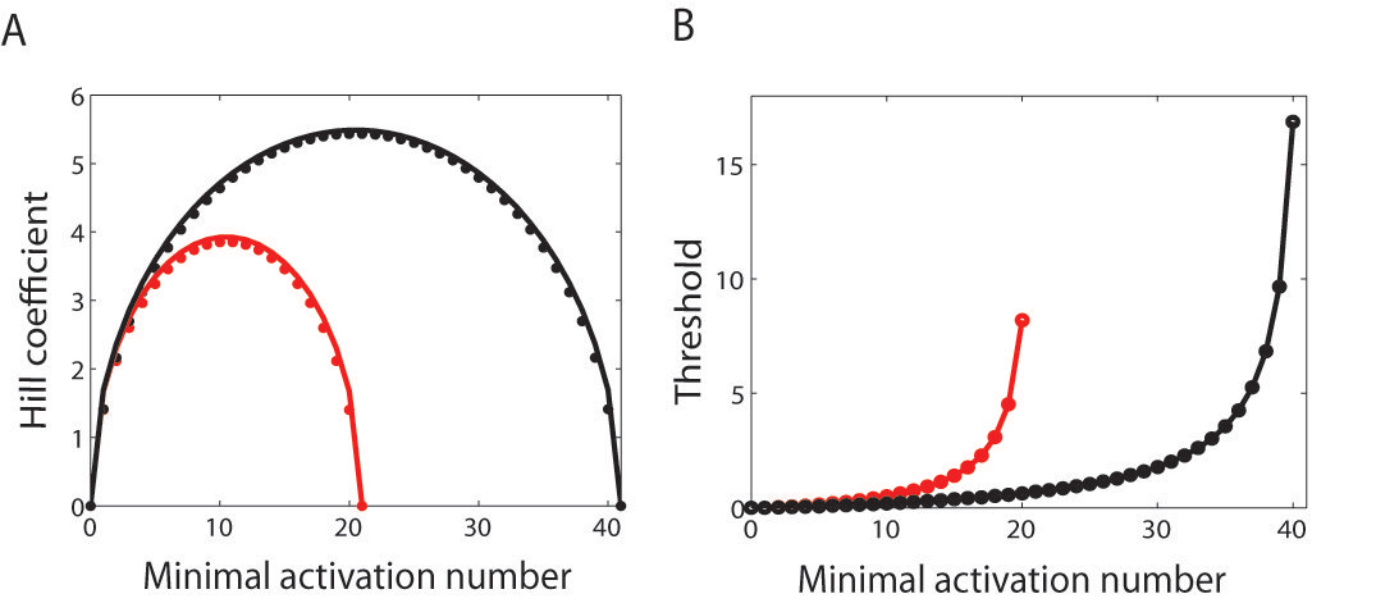
B

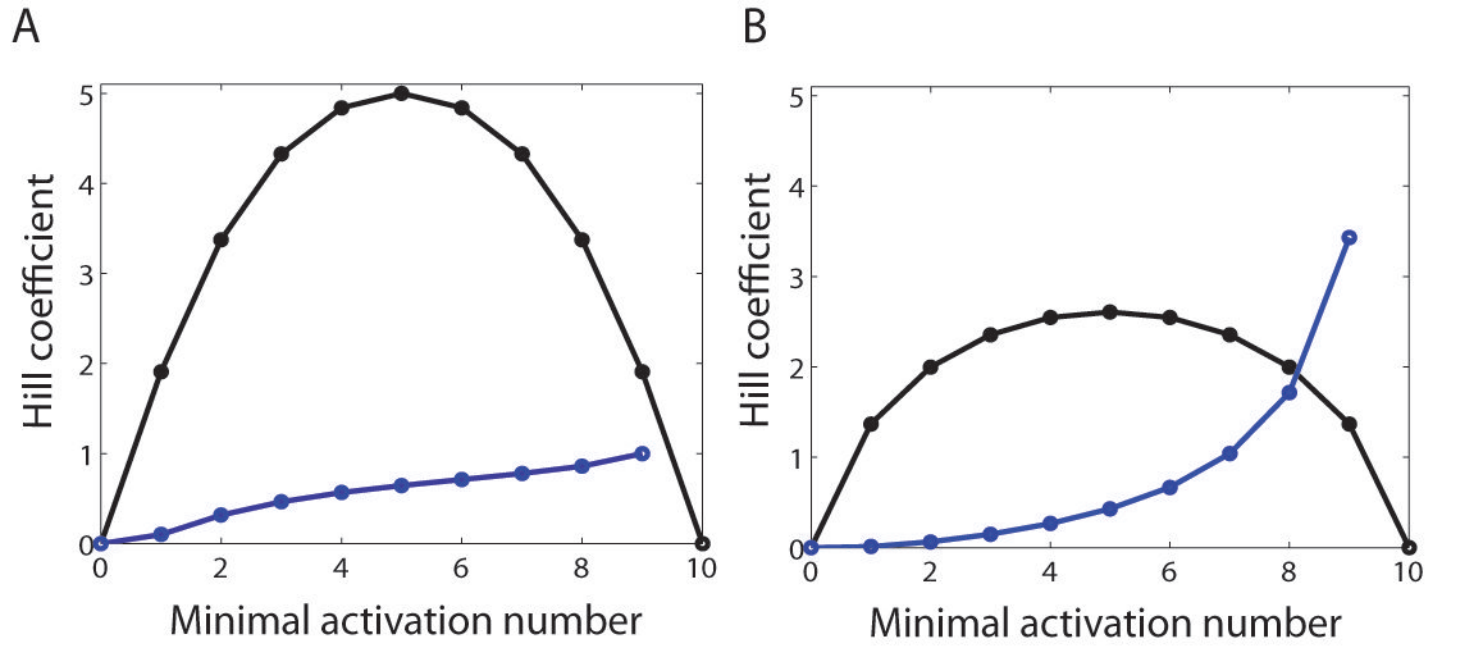
$$\rho_n = \frac{s_n}{1 + \cdots + s_n}$$

C

$$\rho_{n,k} = \frac{s_k + \cdots + s_n}{1 + \cdots + s_k + \cdots + s_n}$$


A
fixed k=10
Active substrate fraction
1.0
0.8
0.6
0.4
0.2
n=20
n=11
n=10
better ultrasensitivity
1
2
3
4
5
Normalized kinase to phosphatase ratio
B
Hill coefficient
6
5
4
3
2
1
b
c
a
2
4
6
8
10
12
Minimal activation number
C
Hill coefficient
20
15
10
5
0
0
5
10
15
20
Minimal activation number
D
Threshold
2
1.5
1
0.5
0
0
5
10
15
20
Minimal activation number


A
Hill coefficient
Minimal activation number
0
1
2
3
4
5
6
0
10
20
30
40
B
Threshold
Minimal activation number
0
5
10
15
0
10
20
30
40


A
B
Hill coefficient
Minimal activation number
0
1
2
3
4
5
6
8
10

Non-Essential Sites Improve Phosphorylation Switch

Liming Wang, Qing Nie, and German Enciso

Text S1: Supplementary Material

Contents

A Computing the Hill coefficient under sequential phosphorylation **1**

B Intuition using biased random walks **3**

C Unordered phosphorylation **4**

D Ultrasensitivity with respect to k, n, and α **7**

D.1 Fix n, vary k 7

D.2 Fix k, vary n 7

D.3 Fix α, vary n 8

E Threshold with respect to k, n, and α **8**

E.1 Fix n, vary k 8

E.2 Fix k, vary n 9

E.3 Fix α, vary n 10

A Computing the Hill coefficient under sequential phosphorylation

In the sequential case, each phosphorylation and dephosphorylation consist of the following elementary chemical reactions respectively,

$$S_i + E \underset{k_{\text{off}_i}}{\overset{k_{\text{on}_i}}{\rightleftarrows}} ES_i \overset{k_{\text{cat}_i}}{\rightarrow} S_{i+1} + E,$$

$$S_{i+1} + F \underset{l_{\text{off}_i}}{\overset{l_{\text{on}_i}}{\rightleftarrows}} FS_{i+1} \overset{l_{\text{cat}_i}}{\rightarrow} S_i + F.$$

Here, k_{on_i} and k_{off_i} are the binding and unbinding rates of kinse E and substrate S_i, respectively; k_{cat_i} is the catalytic rate for the complex ES_i to produce S_{i+1}. Simliarly, l_{on_i} and l_{off_i} are the binding and unbinding rates of the phosphatase F and the substrate S_{i+1}, respectively; l_{cat_i} is the catalytic rate for the complex FS_{i+1} to produce S_i.

Our goal is to compute the steady state proportion of the substrates with at least k phosphorylated sites,

$$\frac{s_k + \cdots + s_n}{s_0 + \cdots + s_k + \cdots + s_n}.$$

Here, small letters denote the steady state concentrations of their corresponding proteins; the number k is referred as the *minimal activation number*. The steady state concentrations of different phosphoforms turn out to satisfy the following relations [2, 3, 4]:

$$s_{i+1} = \lambda_i u s_i, \tag{1}$$

where λ_i and u are defined as

$$u := \frac{e}{f}, \quad \lambda_i := \frac{k_{\text{cat}_i} L_{M_i}}{K_{M_i} l_{\text{cat}_i}}, \quad K_{M_i} := \frac{k_{\text{cat}_i} + k_{\text{off}_i}}{k_{\text{on}_i}}, \quad L_{M_i} := \frac{l_{\text{cat}_i} + l_{\text{off}_i}}{l_{\text{on}_i}}.$$

Note that in general the steady states are not unique for arbitrary given total concentrations of the kinase and the phosphatase [2, 3, 4], however, it is unique when considering the free kinase to phosphatase ratio as an input. Based on (1), the steady state fraction of substrates with at least k phosphorylated sites equals

$$\frac{\lambda_1\lambda_2\cdots\lambda_k u^k + \cdots + \lambda_1\lambda_2\cdots\lambda_n u^n}{1 + \lambda_1 u + \lambda_1\lambda_2 u^2 + \cdots + \lambda_1\lambda_2\cdots\lambda_k u^k + \cdots + \lambda_1\lambda_2\cdots\lambda_n u^n}.$$

When the relative kinase to phosphatase efficiencies are similar for each site, i.e., $\lambda_i \approx \lambda$, the above formula can be simplified to

$$r_{n,k}(x) := \frac{x^k + \cdots + x^n}{1 + x + \cdots + x^k + \cdots + x^n} = \frac{x^{n+1} - x^k}{x^{n+1} - 1}, \tag{2}$$

where $x := \lambda u$. This function $r_{n,k}(x)$ defines the input-output (also called the dose-response) curve with $\lambda e/f$ as the input and the proportion of active substrates as the output. When $n = 2k - 1$, equation (2) simplifies to

$$r_{2k-1,k}(x) = \frac{x^k}{1 + x^k},$$

which is a Hill function with Hill coefficient k. The effective Hill coefficient of the function $r_{n,k}(x)$, $0 < k < n$, can be estimated by the Goldbeter-Koshland Formula [1],

$$H_s(n, k) = \frac{\ln 81}{\ln(v_{n,k}/u_{n,k})}, \tag{3}$$

where

$$u_{n,k} = r_{n,k}^{-1}(0.1), \quad v_{n,k} = r_{n,k}^{-1}(0.9). \tag{4}$$

If the dose-response curve is a Hill function $\frac{x^m}{K^m + x^m}$, formula (3) recovers the Hill coefficient m. For the convenience of notation, when there is no confusion, we omit the subscripts of r, u, and so on. Since $r(x)$ is an increasing function in x, both u and v in (4) are well-defined. When $k = 0$ and $n + 1$, naturally, we define

$$H_s(n, 0) = H_s(n, n + 1) = 0.$$

Let α be the ratio of k and $n+1$, termed as the *site activation ratio*. We next calculate $H_s(n,k)$ for arbitrary n and k. Define an auxiliary function

$$\bar{r}_\alpha(x) = \frac{x - x^\alpha}{x - 1}.$$

Let us denote the Hill coefficient of $\bar{r}_\alpha$ by $\bar{H}$. Naturally, $\bar{H}$ equals zero when $\alpha = 0$ and $\alpha = 1$. For $0 < \alpha < 1$, we perform a change of variable with $\bar{x} = x^{n+1}$, then

$$r_{n,k}(x) = \frac{\bar{x} - \bar{x}^\alpha}{\bar{x} - 1} = \bar{r}_\alpha(\bar{x}).$$

If we let $\bar{u} = u_{n,k}^{n+1}$ and $\bar{v} = v_{n,k}^{n+1}$, then $\bar{r}_\alpha(\bar{u}) = 0.1$, $\bar{r}_\alpha(\bar{v}) = 0.9$, and

$$\bar{H}(\alpha) = \frac{\ln 81}{\ln(\bar{v}/\bar{u})} = \frac{\ln 81}{(n+1)\ln(v_{n,k}/u_{n,k})} = \frac{1}{n+1} H_s(n,k).$$

Multiplying both sides by $n+1$, we obtain

$$H_s(n,k) = \bar{H}(\alpha)\,(n+1). \tag{5}$$

Next, we show that $\bar{H}(\alpha)$ is symmetric with respect to $\alpha = 1/2$. Let $\bar{u}$ be such that $\bar{r}_\alpha(\bar{u}) = 0.1$, then it is equivalent to show that $\bar{r}_{1-\alpha}(\bar{u}^{-1}) = 0.9$, which indeed holds from a straightforward computation,

$$\frac{\bar{u}^{-1} - \bar{u}^{-(1-\alpha)}}{\bar{u}^{-1} - 1} = 0.9.$$

Similiarly, let $\bar{v}$ be such that $\bar{r}_\alpha(\bar{v}) = 0.9$, then $\bar{r}_{1-\alpha}(\bar{v}^{-1}) = 0.1$, and thus the Hill coefficient of $\bar{r}_{1-\alpha}$ is

$$\bar{H}(1-\alpha) = \frac{\ln 81}{\ln(\bar{u}^{-1}/\bar{v}^{-1})} = \frac{\ln 81}{\ln(\bar{v}/\bar{u})} = \bar{H}(\alpha).$$

Our numerical simulations further reveal that $\bar{H}(\alpha)$ is well approximated by the quadratic function $2\alpha(1-\alpha)$, that is,

$$H_s(n,k) \approx 2\alpha(1-\alpha)(n+1) = 2k\left(1 - \frac{k}{n+1}\right). \tag{6}$$

B Intuition using biased random walks

It helps to have an intuition for why simply reducing the minimal activation number (or adding additional sites without increasing this number) increases the switch-like behavior of the system. Imagine the phosphorylation of an individual protein as a discrete stochastic event. At any given time t the protein is in a state between 0 and n phosphorylations, and it follows a random walk between these states. The propensities for phosphorylation and dephosphorylation are given by the constants e and f respectively. The probabilities $P_0(t), \ldots, P_n(t)$ for being in a specific state at time t satisfy the system of differential equations

$$\begin{aligned}
P_0' &= -eP_0 + fP_1 \\
P_1' &= eP_0 - fP_1 - eP_2 + fP_0 \\
&\vdots \\
P_n' &= eP_{n-1} - fP_n,
\end{aligned}$$

and the steady state probabilities P_i are given by the same formula (2) as the steady state concentrations for the original continuous system, in the perfect balanced case.

Now, if e is even slightly larger than f, then the bias in the random walk will intuitively make the protein spend most of its time in the top half of the states, since any biased random walk eventually moves with high probability in the direction of the bias. Similarly it will spend little time in this region if e is slightly smaller than f. This accounts for the ultrasensitive behavior. Moreover, if e is slightly larger than f, then the ball will spend slightly more time at the most phosphorylated state than if $e = f$, but not much more, since the random process itself will constantly kick it out of this location. This means that P_n as a function of e/f is much less ultrasensitive than $P_k + \cdots + P_n$.

C Unordered phosphorylation

In the unordered case, the sites are phosphorylated and dephosphorylated in a random order. Once a substrate-enzyme complex is formed, different products can be made based on their catalytic rates. The number of phosphoforms grows exponentially with n in the unordered mechanism in contrast to linearly in the sequential mechanism, and we thus expect to see pronounced difference between sequential and unordered cases as n increases.

Introduce an index vector $\vec{a}$, consisting of only zeros and ones, to represent substrates in different phosphoforms. For example, S_{001} denote the substrate with three phosphorylation sites, of which the first two sites are empty, and the last one is occupied. A general phosphorylation and dephosphorylation reaction can be decomposed into elementary reactions as

$$S_{\vec{a}} + E \underset{k^{\vec{a}}_{\rm off}}{\overset{k^{\vec{a}}_{\rm on}}{\rightleftarrows}} ES_{\vec{a}} \overset{k^{\vec{a},\vec{b}}_{\rm cat}}{\to} S_{\vec{b}} + E,$$

$$S_{\vec{b}} + F \underset{l^{\vec{b}}_{\rm off}}{\overset{l^{\vec{b}}_{\rm on}}{\rightleftarrows}} FS_{\vec{b}} \overset{l^{\vec{b},\vec{a}}_{\rm cat}}{\to} S_{\vec{a}} + F.$$

Here, the index vector $\vec{b}$ could be any vector obtained by replacing a zero in vector $\vec{a}$ by a one. For example, when $\vec{a} = (0,0,1)$, $\vec{b}$ could be $(0,1,1)$ or $(1,0,1)$, but not $(1,1,1)$. We assume that the kinase-substrate complex is determined by the reacting substrate and kinase, but not by the releasing product. That is, when $\vec{a}$ is given, different choices of $\vec{b}$ share the same kinase-substrate complex, $ES_{\vec{a}}$. This is especially suitable for the situation when the kinase has a docking site. Once the substrate binds to the docking site, any unphosphorylated residue on the substrate is a candidate to be phosphorylated.

Similarly as the sequential case, the steady state concentrations of different phosphoforms satisfy [3],

$$s_{\vec{b}} = \lambda_{\vec{a},\vec{b}}\, u\, s_{\vec{a}}, \tag{7}$$

where

$$\lambda_{\vec{a},\vec{b}} := \frac{k^{\vec{a},\vec{b}}_{\rm cat}}{K^{\vec{a},\vec{b}}_M} \frac{L^{\vec{b},\vec{a}}_M}{l^{\vec{b},\vec{a}}_{\rm cat}}, \quad K^{\vec{a},\vec{b}}_M := \frac{k^{\vec{a},\vec{b}}_{\rm off} + \sum_{\vec{a} \overset{E}{\to} \vec{b}} k^{\vec{a},\vec{b}}_{\rm cat}}{k^{\vec{a},\vec{b}}_{\rm on}}, \quad L^{\vec{b},\vec{a}}_M := \frac{l^{\vec{b},\vec{a}}_{\rm off} + \sum_{\vec{b} \overset{F}{\to} \vec{a}} l^{\vec{b},\vec{a}}_{\rm cat}}{l^{\vec{b},\vec{a}}_{\rm on}}.$$

Let us denote by s_i the total concentration of different phosphoform substrates with i sites being phosphorylated. For example, s_1 represents the sum of s_{100}, s_{010}, and s_{001}. Under the perfect balanced condition ($\lambda_{\vec{a},\vec{b}} = \lambda$), we have

$$s_i = \left(\begin{array}{c} n \\ i \end{array} \right) x^i,$$

since there are n choose k different phosphoforms with exactly k phosphorylated sites. Thus, the steady state proportion of the active substrates is,

$$g_{n,k}(x) = \frac{\left(\begin{array}{c} n \\ k \end{array} \right) x^k + \cdots + \left(\begin{array}{c} n \\ n \end{array} \right) x^n}{1 + \left(\begin{array}{c} n \\ 1 \end{array} \right) x + \cdots + \left(\begin{array}{c} n \\ k \end{array} \right) x^k + \cdots + \left(\begin{array}{c} n \\ n \end{array} \right) x^n} = \frac{\sum_{i=k}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) x^i}{(1+x)^n},$$

where $x = \lambda u$. Next, we prove that there exists a function σ_r that only depends on α such that $H_r(n,k)$ can be written as

$$H_r(n,k) \approx \sigma_r(\alpha)\sqrt{n+1}. \tag{8}$$

To show this, we first interpret the function $g_{n,k}(x)$ in terms of the random process of tossing coins. Define i.i.d. random variabled $Y_i \in \{0,1\}$, $i = 1, \ldots, n$, with

$$\text{Prob}(Y_i = 1) = p, \quad \text{Prob}(Y_i = 0) = q,$$

where $p + q = 1$ and $p, q > 0$. Let $Y_i = 1$ denote the ith toss being head, and $Y_i = 0$ denote the ith toss being tail. The expectation and the variance of Y_i are

$$E(Y_i) = p, \quad Var(Y_i) = E(Y_i^2) - (EY_i)^2 = p - p^2.$$

The sum of Y_i's, $W_n := \sum_{i=1}^{n} Y_i$, counts the total number of heads. That is, $\text{Prob}(W_n = k)$ represents the probability of seeing k heads in n independent experiments, which can be computed as

$$\text{Prob}(W_n \geq k) = \sum_{i=k}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) p^i q^{n-i}.$$

If we let $x = \frac{p}{q}$, then

$$\text{Prob}(W_n \geq k) = \sum_{i=k}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) \left(\frac{x}{1+x} \right)^i \left(\frac{1}{1+x} \right)^{n-i} = g_{n,k}(x). \tag{9}$$

On the other hand, the central limit theorem says

$$Z_n := \frac{W_n - np}{\sqrt{p - p^2}\sqrt{n}} \sim N(0,1).$$

Therefore,

$$\begin{aligned} g_{n,k}(x) &= \text{Prob}(W_n \geq k) \\ &= \text{Prob}\left(Z_n \geq \frac{k - np}{\sqrt{p - p^2}\sqrt{n}} \right) \\ &\approx 1 - \Phi\left(\frac{k - np}{\sqrt{p - p^2}\sqrt{n}} \right) \\ &= 1 - \Phi\left(\frac{k(1+x) - nx}{\sqrt{xn}} \right), \end{aligned} \tag{10}$$

where function Φ is the cumulative distribution function of $N(0,1)$. The definition of v says

$$0.9 \approx 1 - \Phi\left(\frac{k(1+v) - nv}{\sqrt{vn}}\right). \tag{11}$$

For the simplicity of notations, we use equal sign from now on. Rearranging (11), we have

$$k(v+1) - nv = -\xi\sqrt{nv},$$

where $\xi = -\Phi^{-1}(0.1)$. In the above equation, define $\gamma := \sqrt{v}$ and replace k by $\alpha(n+1)$,

$$(\alpha - \frac{n}{n+1})\gamma^2 + \xi\sqrt{\frac{n}{(n+1)^2}}\gamma + \alpha = 0.$$

For large n, the above equation is approximately

$$(\alpha - 1)\gamma^2 + \frac{\xi}{\sqrt{n+1}}\gamma + \alpha = 0.$$

The roots are

$$\gamma_{1,2} = \frac{-\frac{\xi}{\sqrt{n+1}} \pm \sqrt{\frac{\xi^2}{n+1} - 4(\alpha-1)\alpha}}{2(\alpha-1)}.$$

Because $\alpha < 1$ and $\gamma > 0$, we have

$$\sqrt{v} = \gamma = \frac{-\frac{\xi}{\sqrt{n+1}} - \sqrt{\frac{\xi^2}{n+1} - 4(\alpha-1)\alpha}}{2(\alpha-1)}.$$

Similarly, the equation of u, where $g(u) = 0.1$, is

$$k(u+1) - nu = \xi\sqrt{nu}, \tag{12}$$

and the solution is

$$\sqrt{u} = \frac{\frac{\xi}{\sqrt{n+1}} - \sqrt{\frac{\xi^2}{n+1} - 4(\alpha-1)\alpha}}{2(\alpha-1)}. \tag{13}$$

Therefore

$$\begin{aligned}
\frac{\sqrt{v}}{\sqrt{u}} &= \frac{\frac{\xi}{\sqrt{n+1}} + \sqrt{\frac{\xi^2}{n+1} - 4(\alpha-1)\alpha}}{-\frac{\xi}{\sqrt{n+1}} + \sqrt{\frac{\xi^2}{n+1} - 4(\alpha-1)\alpha}} \\
&= \frac{\frac{\xi^2}{n+1} + 2\alpha(1-\alpha) + \frac{\xi}{\sqrt{n+1}}\sqrt{\frac{\xi^2}{n+1} - 4(\alpha-1)\alpha}}{2\alpha(1-\alpha)} \\
&= 1 + \frac{\frac{\xi}{\sqrt{n+1}}\sqrt{\frac{\xi^2}{n+1} - 4(\alpha-1)\alpha}}{2\alpha(1-\alpha)} + \frac{\frac{\xi^2}{n+1}}{2\alpha(1-\alpha)}.
\end{aligned} \tag{14}$$

Define a shorthand

$$A = \frac{\xi}{\sqrt{\alpha(1-\alpha)}}\frac{1}{\sqrt{n+1}}.$$

The third term in (14) becomes $A^2/2$, and the second term in (14) becomes

$$A\sqrt{1+\frac{\xi^2}{4\alpha(1-\alpha)(n+1)}} \approx A,$$

for fixed α and large n. Thus,

$$\frac{\sqrt{v}}{\sqrt{u}} \approx e^{\frac{\xi}{\sqrt{\alpha(1-\alpha)}}\frac{1}{\sqrt{n+1}}}, \tag{15}$$

and

$$\ln\frac{v}{u} = \frac{2\xi}{\sqrt{\alpha(1-\alpha)}}\frac{1}{\sqrt{n+1}}.$$

By the Goldbeter-Koshland Formula, we have

$$H_r(n,k) \approx \frac{\ln 81}{2\xi}\sqrt{\alpha(1-\alpha)}\sqrt{n+1} \approx 1.71\sqrt{\alpha(1-\alpha)}\sqrt{n+1} = 1.71\sqrt{k\left(1-\frac{k}{n+1}\right)}, \tag{16}$$

which is (8) with

$$\sigma_r(\alpha) = 1.71\sqrt{\alpha(1-\alpha)}.$$

D Ultrasensitivity with respect to k, n, and α

D.1 Fix n, vary k

The sequential case

$$\frac{\partial H_s}{\partial k} = 2\left(1-\frac{2k}{n+1}\right).$$

Thus, when n is fixed, as k increases, the Hill coefficient first increases, then decreases, and the maximum is achieved at $2k = n+1$.

The unordered case

$$\frac{\partial H_r}{\partial k} = 1.7145\frac{1-\frac{2k}{n+1}}{\sqrt{\frac{k}{n+1}\left(1-\frac{k}{n+1}\right)}}.$$

Similarly, when n is fixed, as k increases, the Hill coefficient first increases, then decreases, and the maximum is achieved at $2k = n+1$.

D.2 Fix k, vary n

Notice that H_s in (6) and H_r in (16) can be written either in terms of α and n or in terms of k and n. In the following computations, we use the expressions of H_s and H_r involving only k and n.

The sequential case

$$\frac{\partial H_s}{\partial n} = \frac{2k^2}{(n+1)^2} > 0.$$

Thus, when k is fixed, as n increases, the Hill coefficient always increases.

The unordered case

$$\frac{\partial H_r}{\partial n} = \frac{1.7145}{2\sqrt{k\left(1-\frac{k}{n+1}\right)}}\frac{k^2}{(n+1)^2} > 0.$$

So, when k is fixed, as n increases, the Hill coefficient always increases.

D.3 Fix α, vary n

In the following computations, we use the expressions of H_s and H_r in (6) and (16) invloving only α and n.

The sequential case

$$\frac{\partial H_s}{\partial n} = 2\alpha(1-\alpha) > 0.$$

Thus, when α is fixed, as n increases, the Hill coefficient always increases.

The unordered case

$$\frac{\partial H_r}{\partial n} = \frac{1.7145\alpha(1-\alpha)}{2\sqrt{n+1}} > 0.$$

Therefore, when α is fixed, as n increases, the Hill coefficient always increases.

E Threshold with respect to k, n, and α

Let us first prove a fact that will be repeatedly used in our analysis. Define a function

$$y(x) = \frac{a_k x^k + \cdots + a_n x^n}{1 + a_1 x + \cdots + a_{k-1}x^{k-1} + a_k x^k + \cdots + a_n x^n},$$

and we claim that

$$\frac{dy}{dt} > 0 \tag{17}$$

for non-negative x and for arbitrary integers $n > k$, equivalent to proving that the function $1/y$ is decreasing in x. The derivative of $1/y$ with respect to x is,

$$\frac{\sum_{i=1}^{k-1}\sum_{j=k}^{n}(i-j)a_i a_j x^{i+j-1} - \sum_{j=k}^{n} a_j x^{j-1}}{(a_k x^k + \cdots + a_n x^n)^2} < 0.$$

Thus, we have proved the claim.

E.1 Fix n, vary k

The sequential case

For fixed n, rewrite the function $r_{n,k}(x)$ as $l(k,x)$. Thus, the threshold is the solution of

$$0.1 = l(k,x). \tag{18}$$

Taking derivative of both sides of equation (18) with respect to k, we obtain

$$\frac{dx}{dk} = -\frac{\partial l/\partial k}{\partial l/\partial x}.$$

Based on (17), $\partial l/\partial x$ is positive. On the other hand,

$$\frac{\partial l}{\partial k}(k,x) = -\frac{x^k \ln x}{x^{n+1}-1} < 0$$

on both intervals $x > 1$ and $0 < x < 1$. Also, it is easy to see that $\partial l/\partial k$ is continuous at $x = 1$ with

$$\frac{\partial l}{\partial k}(k,1) = -\frac{1}{n+1}.$$

Thus, dx/dk is always positive, i.e., for fixed n, the threshold is increasing in k.

The unordered case

The threshold in the unordered case is solved from equation (12). For fixed n, taking derivatives with respect to k on both sides of equation (12), we obtain

$$\frac{du}{dk} = \frac{u+1}{n-k+\frac{\xi\sqrt{n}}{2\sqrt{u}}} > 0.$$

Thus, for fixed n, the threshold increases in k.

E.2 Fix k, vary n

The sequential case

For fixed k, rewrite the function $r_{n,k}(x)$ as $h(n,x)$. Thus, the threshold is the solution of

$$0.1 = h(n,x). \tag{19}$$

Taking derivative of both sides of equation (19) with respect to n, we obtain

$$\frac{dx}{dn} = -\frac{\partial h/\partial n}{\partial h/\partial x}.$$

Based on (17), $\partial h/\partial x$ is positive. On the other hand,

$$\frac{\partial h}{\partial n}(n,x) = \frac{x^k-1}{(x^{n+1}-1)^2} x^{n+1} \ln x > 0$$

on both intervals $x > 1$ and $0 < x < 1$. Also, it is easy to see that $\partial h/\partial n$ is continuous at $x = 1$ with

$$\frac{\partial h}{\partial n}(n,1) = \frac{k}{(n+1)^2}.$$

Thus, dx/dn is always negative, i.e., for fixed k, the threshold is decreasing in n.

The unordered case

The threshold in the unordered case is solved from equation (12). For fixed k, taking derivatives with respect to n on both sides of equation (12), we obtain

$$\frac{du}{dn} = -\frac{\left(\frac{\xi}{2\sqrt{nu}}+1\right)u}{n-k+\frac{\xi\sqrt{n}}{2\sqrt{u}}} < 0.$$

Thus, for fixed k, the threshold decreases in n.

E.3 Fix α, vary n

The sequential case

For fixed α, $\bar{u} = \bar{f}^{-1}(0.1)$ is fixed, and the threshold $u = \bar{u}^{1/(n+1)}$. So, the monotonicity of u with respect to n depends on whether $\bar{u}$ is greater than one. On the other hand, $\bar{u} > 1$ if and only if $\alpha > 0.9$. Therefore, the threshold increases in n when $\alpha < 0.9$ and decreases when $\alpha > 0.9$.

The unordered case

In the nonsequential case, the threshold is given in (13). Rewrite $\sqrt{u}$ as

$$\sqrt{u} = \frac{2\alpha}{\frac{\xi}{\sqrt{n+1}} + \sqrt{\frac{\xi^2}{n+1} - 4(\alpha-1)\alpha}}.$$

It is easy to see that for fixed α, u is increasing n, i.e., the threshold is increasing in n.

References

[1] A. Goldbeter and D.E. Koshland. An amplified sensitivity arising from covalent modification in biological systems. *Proceedings of the National Academy of Sciences*, 78(11):6840–6844, 1981.

[2] J. Gunawardena. Multisite protein phosphorylation makes a good threshold but can be a poor switch. *Proc. Natl. Acad. Sci.*, 102:14617–14622, 2005.

[3] M. Thomson and J. Gunawardena. Unlimited multistability in multisite phosphorylation systems. *Nature*, 460(7252):274–277, 2009.

[4] L. Wang and E.D. Sontag. On the number of steady states in a multiple futile cycle. *Journal of Mathematical Biology*, 57:29–52, 2008.

Figure S1

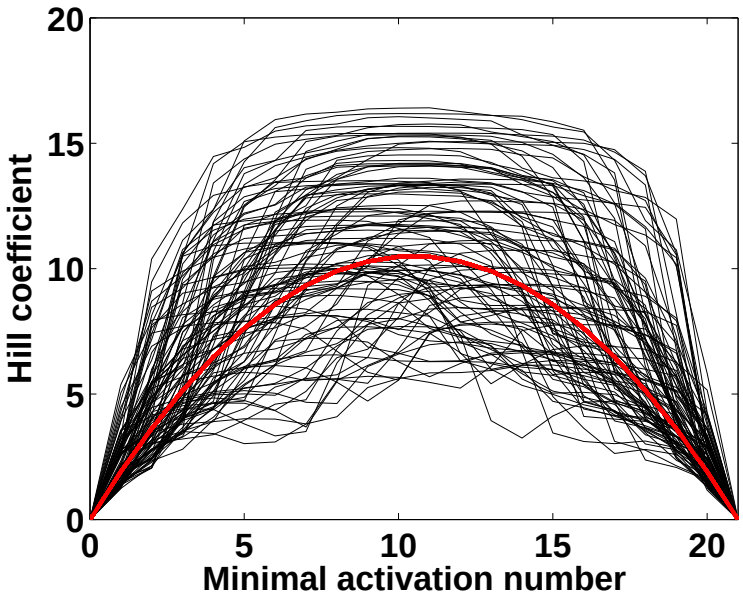


Figure S1: Plot of the Hill coefficients for random λs. Each black curve corresponds to one set of λ_is. In total, 100 sets of λ_is are generated, where each $\log_{10}\lambda_i$ follows a uniform distribution on $[-1, 1]$. The red curve represents the perfect balanced case when $\lambda_i = 1$.

Figure S2

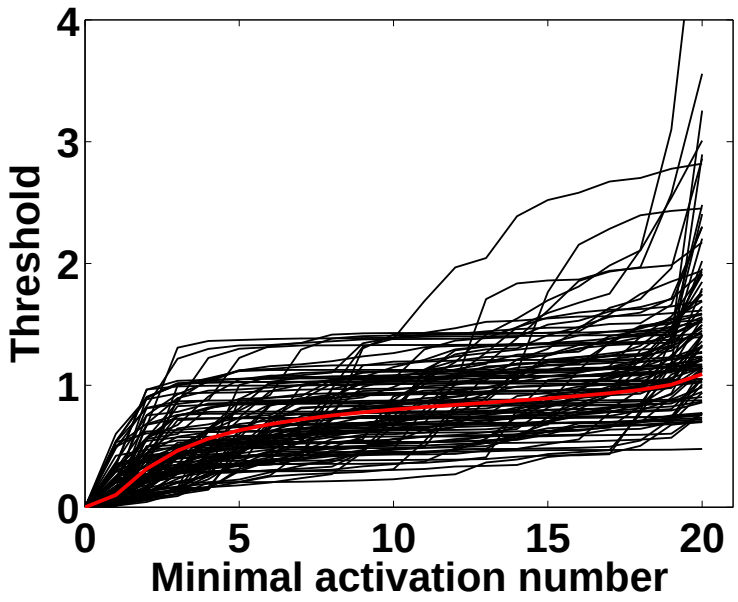


Figure S2: Plot of the thresholds for random λs. Each black curve corresponds to one set of λ_is. In total, 100 sets of λ_is are generated, where each $\log_{10}\lambda_i$ follows a uniform distribution on $[-1, 1]$. The red curve represents the perfect balanced case when $\lambda_i = 1$.

Figure S3

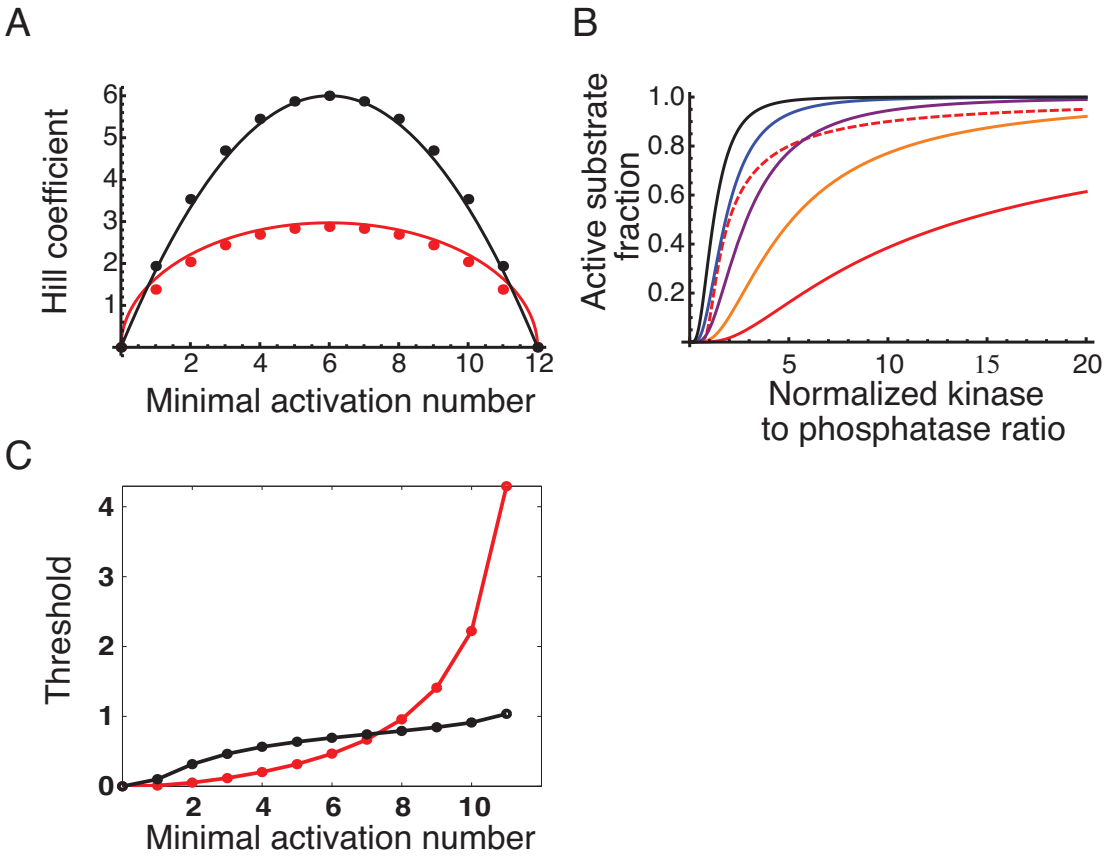


Figure S3: Comparison between sequential and non-sequential mechanisms. (A) The Hill coefficient under the sequential (black) and non-sequential (red) mechanisms. Here, the dots are computed directly from the Goldbeter-Koshland formula, and the curves are estimations from equations (2) and (3) in the main text. In both plots, $n = 11$ and $\lambda_i = 1$. **(B)** Plot of the does-response curves for the sequential case $k = 10$ (red dashed), non-sequential case $k = 10$ (red solid), non-sequential case $k = 9$ (orange solid), 8 (purple solid), 7 (blue solid), 6 (black solid). In all plots, n is fixed at 10. **(C)** The threshold under the sequential (black) and non-sequential (red) mechanisms. Here, $n = 11$ and $\lambda_i = 1$.

Figure S4

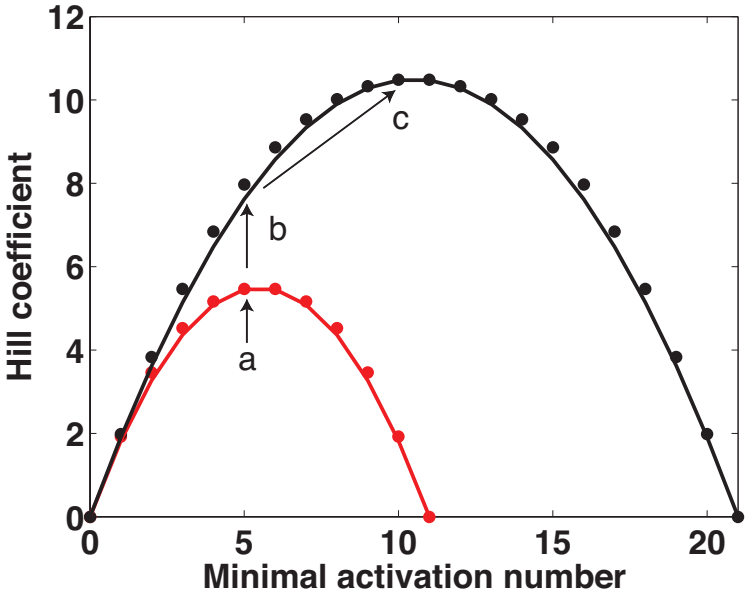


Figure S4: Evolutionary race between k and n. The plot of the Hill coefficient for $n = 10$ (red) and $n = 20$ (black). The point a: $k = 5, n = 10$; point b: $k = 5, n = 20$; point c: $k = 10, n = 20$. In all simulations, $\lambda_i = 1$.

Figure S5

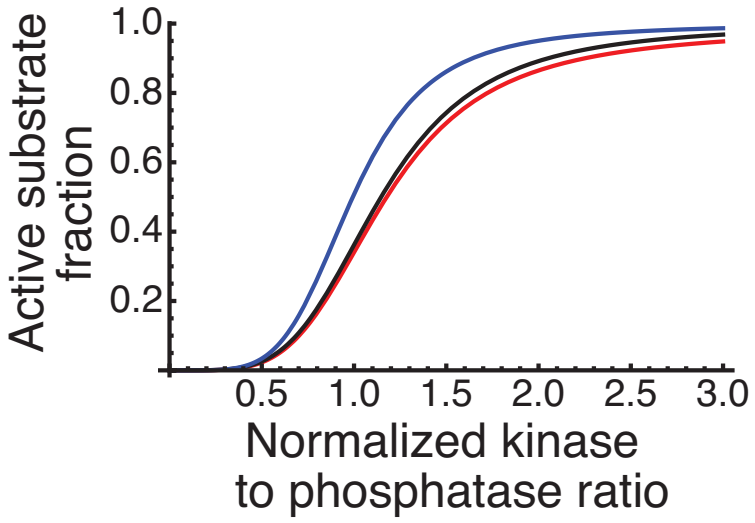


Figure S5: Combination of cooperativity and non-essential sites. The original system (red) shows high ultrasensitivity due to cooperativity with $\lambda_1, \ldots, \lambda_4 = 0.5, \lambda_5 = 16$, $n = k = 5$. There are many ways of combining cooperativity and non-essential sites. For example, $n = 6, k = 5$ with $\lambda_6 = 0.5$ (blue); $n = 5, k = 4$ (black).