

Statistics of cellular signal transduction as a race to the nucleus by multiple random walkers in compartment/phosphorylation space

Ting Lu, Tongye Shen, Chenghang Zong, Jeff Hasty, and Peter G. Wolynes

PNAS 2006;103;16752-16757; originally published online Oct 27, 2006;
doi:10.1073/pnas.0607698103

This information is current as of February 2007.

Online Information & Services	High-resolution figures, a citation map, links to PubMed and Google Scholar, etc., can be found at: www.pnas.org/cgi/content/full/103/45/16752
References	This article cites 17 articles, 11 of which you can access for free at: www.pnas.org/cgi/content/full/103/45/16752#BIBL This article has been cited by other articles: www.pnas.org/cgi/content/full/103/45/16752#otherarticles
E-mail Alerts	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here .
Rights & Permissions	To reproduce this article in part (figures, tables) or in entirety, see: www.pnas.org/misc/rightperm.shtml
Reprints	To order reprints, see: www.pnas.org/misc/reprints.shtml

Notes:

Statistics of cellular signal transduction as a race to the nucleus by multiple random walkers in compartment/phosphorylation space

Ting Lu^{*†}, Tongye Shen^{†‡}, Chenghang Zong^{†‡}, Jeff Hasty^{§¶}, and Peter G. Wolynes^{**†||}

Departments of ^{*}Physics, [‡]Chemistry and Biochemistry, and [§]Bioengineering, [¶]Institute for Nonlinear Science, and [†]Center for Theoretical Biological Physics, University of California at San Diego, La Jolla, CA 92093

Contributed by Peter G. Wolynes, September 8, 2006

Cellular signal transduction often involves a reaction network of phosphorylation and transport events arranged with a ladder topology. If we keep track of the location of the phosphate groups describing an abstract state space, a simple model of signal transduction involving enzymes can be mapped on to a problem of how multiple biased random walkers compete to reach their target in the nucleus yielding a signal. Here, the first passage time probability and the survival probability for multiple walkers can be used to characterize the response of the network. The statistics of the first passage through the network has an asymmetric distribution with a long tail arising from the hierarchical structure of the network. This distribution implies a significant difference between the mean and the most probable signal transduction time. The response patterns for various external inputs generated by our model agree with recent experiments. In addition, the model predicts that there is an optimal phosphorylation enzyme concentration for rapid signal transduction.

stochastic biology | first passage time | NFAT

Achieving a quantitative understanding of the reaction networks that transduce cellular signals is one of the major challenges in biology. Signaling networks are found in a diverse set of organisms, ranging from prokaryotes to eukaryotes, and provide mechanisms for fundamental processes such as gene-regulatory control and cellular communication. Qualitative descriptions of the biomolecular components and mechanisms of cellular signaling have greatly improved our understanding of how cells function and have given insights into how to intervene therapeutically when such signals are miscommunicated. Experimental advances now allow quantitative studies of signal transduction and thereby inspire theoretical treatments. Many networks of nonlinear reactions exhibit interesting behavioral features as ultrasensitivity, adaption, robustness, and discrete “all-or-none” response, which have been quantitatively explored (1–8).

A commonly occurring network topology is the reaction ladder network. This network may be viewed as a generalization of multiple-site phosphorylation/dephosphorylation cascades, such as the pathway governing nuclear factor activation of T cells (NFAT), which regulates the response of T cells to antigen signaling (9–13). To stimulate T cells, NFAT must be transported to the nucleus. This transition occurs in response to a conformational change that exposes a nuclear localization sequence (NLS), which is normally buried in the protein interior in the inactive conformation and thus makes the NFAT inaccessible to transport by importin. The NLS becomes exposed in response to the progressive dephosphorylation of specific serine residues in its regulatory domain. This dephosphorylation occurs in response to an increase of intracellular calcium ions that activate calcineurin, which then dephosphorylates the masking residues. Once a sufficient number of sites have been dephosphorylated, conformational changes expose the NLS so that it can now be transported into the nucleus by the importin. This is

not a one-way process. Inside the nucleus, the NFAT may be progressively rephosphorylated by kinases and subsequently exported to the cytoplasm by the exportin Crm1 (10, 11). To summarize this network generically, the NFAT can exist in a variety of phosphorylation states and at various locations within the cell. Transitions between these phosphorylation/compartimentalization states can be described as a network of reactions consisting of two groups of species C_i and U_i , where the C_i species reside in the cytoplasm and the U_i species reside in the nucleus (Fig. 1). On each side of the cytonuclear barrier, there are $M + 1$ species having different levels of phosphorylation. This network topology can generally be interpreted as representing either processive or distributive mechanisms of phosphorylation (14, 15). If the subscript i labels a specific order of phosphorylation, e.g., phosphorylation of residue A, followed by residue B, followed by residue C, . . . , the network describes the processive (de)phosphorylation; if i represents for the number of (de)phosphorylated residues, the network describes the distributive phosphorylation mechanism, e.g., one residue is first phosphorylated, followed by two residues, then three residues, etc. Of course, the rates connecting i and $i + 1$ will be different for the two mechanisms, but the network topology remains the same. In the following, we study the processive phosphorylation of residues.

The rates of both phosphorylation (by kinases) and dephosphorylation (by phosphatases) are naturally modeled as Michaelis–Menten reactions with rates that depend on the availability of enzymes. Although dephosphorylations dominantly occur in the cytoplasm and phosphorylations dominantly occur in the nucleus in the case of NFAT signaling (9–11), in our model we allow them both to occur in either environment with different rates. The protein phosphorylation state by affecting the conformation of that protein determines how easily it is translocated into or out of the nucleus. In our model k_+^i represents the reaction rate from C_i to U_i and k_-^i is the rate of going from U_i to C_i . An individual NFAT molecule thus makes random walks through its phosphorylation/location space according to these microscopic reaction rates.

The NFAT signaling network shares its topology with many other networks in the cell. It simplifies to a typical two-state system if there is only a single phosphorylation state; the network resembles an enzymatic futile cycle when there are two phosphorylation states but no compartmental transport (16); and the network is also similar to the Monod–Wyman–Changeux model (17). The reaction ladder network thus presents a paradigm for the interplay of spatial heterogeneity and posttranscriptional

Author contributions: T.L., T.S., J.H., and P.G.W. designed research; T.L., T.S., and C.Z. performed research; T.L., T.S., and C.Z. analyzed data; and T.L., T.S., C.Z., J.H., and P.G.W. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: NFAT, nuclear factor activation of T cells; NLS, nuclear localization sequence; FPT, first passage time.

^{||}To whom correspondence should be addressed. E-mail: pwolynes@ucsd.edu.

© 2006 by The National Academy of Sciences of the USA

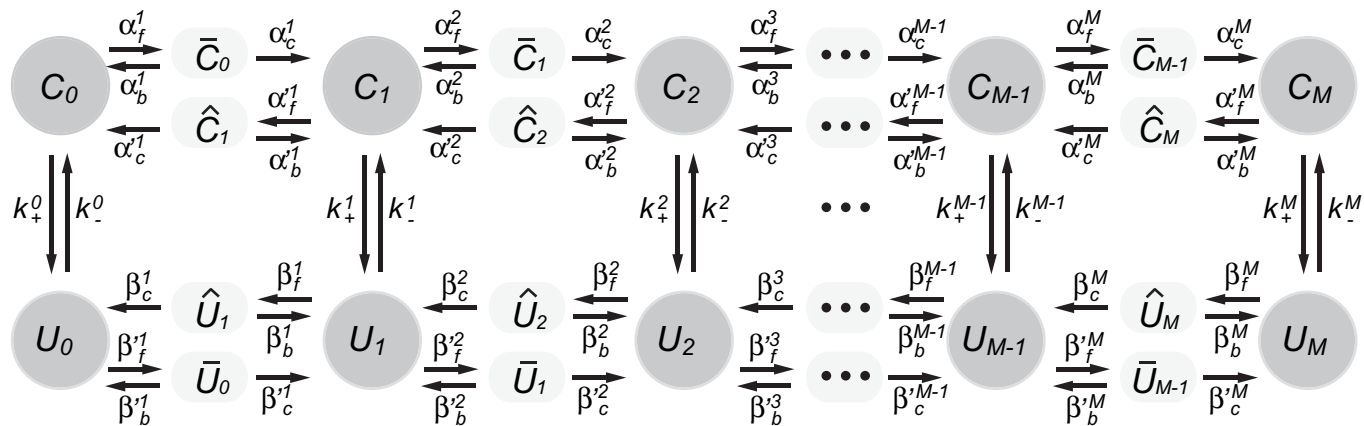


Fig. 1. The reaction ladder network. The C and U represent two distinct compartmental locations of the signaling molecules, say, cytoplasmic and nuclear regions. The subscript i ($i = 0, 1, 2, \dots, M$) indicates the dephosphorylation states. The \hat{C}_i and \hat{U}_i (\hat{U}_i and \hat{U}_i) are the signaling protein–enzyme complex forms. The subscripts f , b , and c represent forward, backward, and catalyzed rates of each reaction. The k_+ and k_- are the transport rates for transitions between C and U .

modification in the flow of biological information. The modification reactions, conformational changes and intercompartmental transports are intrinsically stochastic events. On the ladder network, each signaling molecule follows a path through the network, causing transcription to occur at random times. When the modifying enzymes are abundant, the network is effectively linear and each molecule walks through location/phosphorylation space independently. When enzymes are limited in number, the network becomes nonlinear and the walks of different signaling molecules interact by competing for enzymes. Ultimately, it takes only one NFAT to turn on its target gene. Thus we can say, somewhat anthropomorphically, that the individual NFAT's are competing in a race to the DNA. We thus have a problem of determining the statistics of mean first passage for multiple walkers.

The statistical problem of calculating the mean first passage time (FPT) for random walkers has a distinguished history (18–22). Here we will find the first passage time and the survival probability in terms of a dynamic probability distribution. We will then show that the exact solution of this problem of multiple random walkers having the same goal does indeed agree with the results of Monte Carlo simulation of the network when enzymes are unlimited. The mean FPT distribution is found to be asymmetric and has a long tail. The solution also shows there is an optimal forward reaction rate yielding the most rapid arrival of a viable signaling protein to the target.

Distribution of FPTs and Survival Probability

Different phosphorylation and dephosphorylation processes are catalyzed by specific enzymes (23, 24). For simplicity of treatment, we assume a universal kinase and a universal phosphatase, K_c and F_c , in the cytoplasm and another set, K_u and F_u , in the nucleus. Also for simplicity we assume only apo proteins can be transported between the cytoplasm and the nucleus. The enzymes (K_c , F_c , K_u , and F_u) as well as the signaling protein phosphatase complexes (\hat{C}_i or \hat{U}_i) and the signaling protein kinase complexes (\hat{C}_i or \hat{U}_i) cannot be transported.

Suppose that there are c_i , \bar{c}_i , and \hat{c}_i proteins in the C_i , \bar{C}_i , and \hat{C}_i states and u_i , \bar{u}_i , and \hat{u}_i proteins in the U_i , \bar{U}_i , and \hat{U}_i states respectively ($i = 0, 1, 2, \dots, M$). Then, the numbers of proteins in each state is described by a $6M + 6$ dimensional vector $\vec{n} \equiv (\hat{c}_0, c_0, \bar{c}_0, \hat{u}_0, u_0, \bar{u}_0; \dots; \hat{c}_M, c_M, \bar{c}_M, \hat{u}_M, u_M, \bar{u}_M)$, where $\hat{c}_0, \bar{u}_0, \hat{u}_M$, and \bar{c}_M are zeros due to the boundaries of the network. The numbers of the enzymes in the system are defined by a four dimensional vector $\vec{E} = (F_c, K_c, F_u, K_u)$. We define a state

$|\Psi(t)\rangle$ as $|\Psi(t)\rangle \equiv \sum_{\vec{n}, \vec{E}} P(\vec{n}, \vec{E}, t) |\vec{n}, \vec{E}\rangle$, where $P(\vec{n}, \vec{E}, t)$ is the probability having \vec{n} and \vec{E} numbers of proteins and enzymes in the network. The master equation describing the network can then be written as $\partial/\partial t |\Psi(t)\rangle = W |\Psi(t)\rangle$, where W is the transition rate matrix whose dimension depends on the total numbers of enzymes and substrates.

Generally solving this master equation represents a challenging many-body problem. However, when the numbers of enzymes in the cytoplasm and the nucleus are very large compared with the total number of signaling proteins, as often happens in real biological systems, the phosphorylation and dephosphorylation processes that lead to the transitions of the signaling molecules are uncorrelated. Each protein can then be modeled as an independent random walker. Our assumption of an enzyme-saturated situation makes the mathematics of the network relative simple and the problem of multiple but independent random walkers can be solved exactly. This exact solution allows several interesting properties of the network to be explored.

A key aspect characterizing signaling pathways is the time to achieve a response after receiving an upstream signal, i.e., the typical delay time between a stimulus and the corresponding response. This is a stochastic quantity. The response occurs when one of the random walkers successfully binds to the DNA. To quantify this, we may consider the FPT for a random walker starting from the initial position \vec{r}_i , arriving at the final position \vec{r}_f for the first time. $F(\vec{r}_i, \vec{r}_f, t)$ is the probability distribution of such a random walker, initially in \vec{r}_i , whose FPT of reaching the final position \vec{r}_f is time t . $F(\vec{r}_i, \vec{r}_f, t)$ is related to the occupancy probability $P(\vec{r}_i, \vec{r}_f, t)$, which is the probability that a particle is found at the position \vec{r}_f at time t irrespective of when it arrived. This relation is

$$P(\vec{r}_i, \vec{r}_f, t) = \int_0^t d\tau F(\vec{r}_i, \vec{r}_f, \tau) P^{(s)}(\vec{r}_f, \vec{r}_f, t - \tau). \quad [1]$$

Both the FPT probability $F(\vec{r}_i, \vec{r}_f, \tau)$ and occupancy probability $P(\vec{r}_f, \vec{r}_f, t)$ are normalized through $\int dt F(\vec{r}_i, \vec{r}_f, t) = 1$ and $\int d\vec{r}_f P(\vec{r}_i, \vec{r}_f, t) = 1$. $P^{(s)}(\vec{r}, \vec{r}, t)$ is the occupancy probability with the identical initial and final position, i.e., the chance of a particle staying at and returning to the same position \vec{r} after time t . In terms of Laplace transforms, the above equation can then be rewritten as

$$F(\vec{r}_i, \vec{r}_f, t) = \mathcal{L}^{-1} \left\{ \frac{\mathcal{L}\{P(\vec{r}_i, \vec{r}_f, t)\}}{\mathcal{L}\{P^{(s)}(\vec{r}_f, \vec{r}_f, t - \tau)\}} \right\}. \quad [2]$$

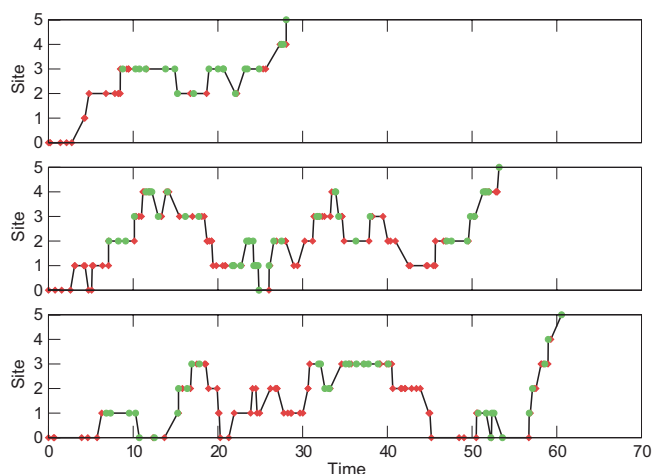


Fig. 2. Three typical trajectories of a random walker traveling from C_0 to U_M are plotted as time vs. site number, i.e., the state of dephosphorylation. The red diamonds label a protein in the cytoplasmic form (C_i , \bar{C}_i or \hat{C}_i), whereas the green dots label the nuclear form (U_i , \bar{U}_i , or \hat{U}_i). The parameters are chosen as follows: the phosphorylation site number $M = 5$, the forward reaction rates $\alpha_f = \beta_f = \beta_f^c = \alpha_f = 0.2$, the backward rates $\alpha_b = \beta_b = \beta_b^c = \alpha_b = 1.0$, the catalyzed rates $\alpha_c = \beta_c = \beta_c^c = \alpha_c = 1.0$, the transport rates $k_+^M = k_-^M = 0.2$, and the ratio of transport rates $\gamma_+ = \gamma_- = 2.718$.

The survival probability is the probability that, up to time t , the random walker still has never reached the target position \bar{r}_f , which is represented as $S(\bar{r}_i, \bar{r}_f, t)$. By the definition, the survival probability S is

$$S(\bar{r}_i, \bar{r}_f, t) = 1 - \int_0^t d\tau F(\bar{r}_i, \bar{r}_f, \tau). \quad [3]$$

The above FPT probability and survival probability are formulated for a single particle; however, it is straightforward to expand this to the multiparticle case if there is no interaction between random walkers. In the case of large number of enzymes, multiple particles move independently and thus the probability for having all N particles can be obtained by multiplying the survival probabilities for each single particle, i.e.

$$S(\bar{r}_i, \bar{r}_f, t; N) = S^N(\bar{r}_i, \bar{r}_f, t; 1) = \left(1 - \int_0^t d\tau F(\bar{r}_i, \bar{r}_f, \tau; 1) \right)^N, \quad [4]$$

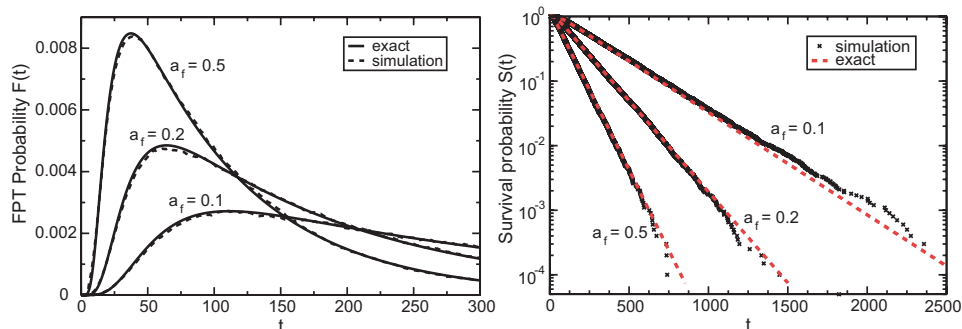


Fig. 3. Comparisons of the probabilities of the FPT and survival probabilities. (Left) Probability vs. the FPT. (Right) The survival probability vs. time. Parameters are the same as those in Fig. 2.

Table 1. Probability distribution of a random walker at time $t = 50$

$C_i + \bar{C}_i + \hat{C}_i$	$U_i + \bar{U}_i + \hat{U}_i$
0.2236	0.0294
0.2098	0.0358
0.1540	0.0410
0.0983	0.0424
0.0571	0.0404
0.0358	0.0325

For a given i , the three terms C_i , \bar{C}_i , and \hat{C}_i with same phosphorylation states are collected together to represent the total probability of a protein at the i phosphorylation state in the C group. The three terms U_i , \bar{U}_i , and \hat{U}_i are also collected together for the same reason. Parameters are chosen the same as those in Fig. 2.

where $F(\bar{r}_i, \bar{r}_f, t; 1)$ and $S(\bar{r}_i, \bar{r}_f, t; 1)$ are the FPT and the survival probability for a single particle. The probability of having exactly z of total N particles in the position \bar{r}_f at time t irrespective of their arrivals is

$$P(\bar{r}_i, \bar{r}_f, t; N, z) = \frac{N!}{z!(N-z)!} P^z(\bar{r}_i, \bar{r}_f, t) [1 - P(\bar{r}_i, \bar{r}_f, t)]^{N-z}. \quad [5]$$

One defines the accumulated FPT probability $F^{ac}(\bar{r}_i, \bar{r}_f, t; N, z)$ as the probability that at time t z of the total N particles have all arrived at the destination for the first time by time t . The expression is

$$F^{ac}(\bar{r}_i, \bar{r}_f, t; N, z) = \frac{zN!}{z!(N-z)!} S^{N-z}(\bar{r}_i, \bar{r}_f, t; 1) \cdot [1 - S(\bar{r}_i, \bar{r}_f, t; 1)]^{z-1} \cdot F(\bar{r}_i, \bar{r}_f, t; 1). \quad [6]$$

We may also define the simultaneous FPT probability $F^{si}(\bar{r}_i, \bar{r}_f, t; N, z)$, which is the probability that at time t z of the total N particles simultaneously arrived at the destination for the first time. The corresponding expression is the same as that for single particle case, i.e.,

$$F^{si}(\bar{r}_i, \bar{r}_f, t; N, z) = \mathcal{L}^{-1} \left\{ \frac{\mathcal{L}\{P(\bar{r}_i, \bar{r}_f, t; N, z)\}}{\mathcal{L}\{P^{(s)}(\bar{r}_i, \bar{r}_f, t - \tau; N, z)\}} \right\}.$$

Results

It is easy to assign different rates to each step and carry out the calculations. For simplicity, we will first assume uniform forward reaction rates α_f as well as uniform backward rates α_b and catalyzed rates α_c for all phosphorylation and dephosphorylation

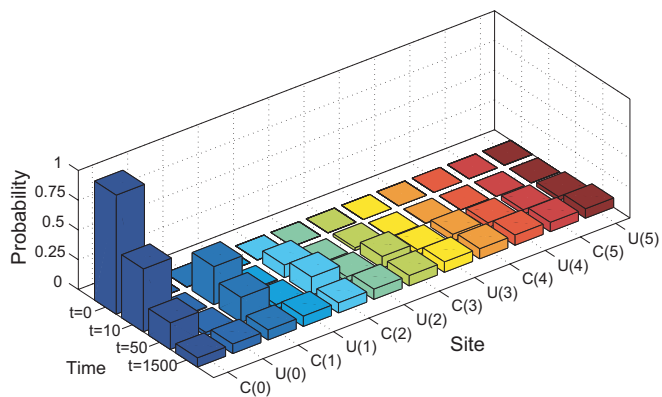


Fig. 4. Dynamic propagation of the probability distribution. In the site axis, the $C(i)$ and $U(i)$ ($i = 0, 1, \dots, 5$), defined as $C(i) = C_i + \bar{C}_i + \hat{C}_i$ and $U(i) = U_i + \bar{U}_i + \hat{U}_i$, are collected terms representing the probabilities of a protein at the i phosphorylation state in the C and U group, respectively. Four time slices $t = 0, 10, 50$, and $1,500$ are chosen to show the dynamic propagation. The parameters are chosen the same as those in Fig. 2.

events. We also assume that the transportation rates k_+^i increase evenly and k_-^i decrease evenly with the increase of the number of unphosphorylated sites i , i.e., $k_+^i = k_+^M(\gamma_+)^{i-M}$, $k_-^i = k_-^0(\gamma_-)^{-i}$ ($i = 0, 1, 2, \dots, M$). This assumption captures the empirical observation that fully dephosphorylated NFAT is much easier to transport from the cytoplasm into the nucleus than phosphorylated NFAT. In the nucleus, the fully phosphorylated NFAT is most easily transported to the cytoplasm (23).

Comparison of the Exact Solution with Simulation. Fig. 2 illustrates three trajectories taken from a Monte Carlo simulation of a signaling protein traveling from the initial fully phosphorylated state C_0 in the cytoplasm to the fully dephosphorylated state U_M in the nucleus (25, 26). The red diamonds in Fig. 2 indicate that the protein is found in the cytoplasm regardless of its specific form (C_i , \bar{C}_i , or \hat{C}_i), whereas the green dots indicate that the protein is found in the nucleus regardless of its specific form (U_i , \bar{U}_i , or \hat{U}_i). Transitions between red and green sites indicate a transverse across the cytoplasm–nucleus barrier, whereas up and down transitions indicate phosphorylation and dephosphorylation events.

In Fig. 3, we compare the mean FPT probability and survival probability from the exact solution with those from the Monte

Carlo simulations. Fig. 3 *Left* shows several distributions of the FPT probability computed from the exact solutions (solid lines) and those computed from the stochastic simulations (broken lines) with the same parameters; Fig. 3 *Right* shows the corresponding survival probabilities from the exact solutions (red broken lines) and those from simulations (crosses). These simulation results agree very well with the exact solutions.

An exact solution can be found not only for the steady state distribution but also for the dynamics away from the steady states. Table 1 shows the occupancy probability distribution at time $t = 50$ s when the network is far from the steady state. The difference between the exact solution and simulation is $\approx 2\%$, which is essentially the sampling error. Fig. 4 further explores the network dynamics. For a network with size $M = 5$, the walker initially resides in the C_0 state ($t = 0$ s) but propagates to other states by time $t = 10$ s and $t = 50$ s. Eventually the probability reaches a steady profile. The last time shown, $t = 1,500$ s, is much later than the mean FPT, 180 s.

We can also compare the model's predications with laboratory experiments carried by Dolmetsch *et al.* (27). These experiments measure differential NFAT activation as a function of the amplitude and duration of a calcium stimulus. Their work uncovered three different response patterns of the nuclear fraction of total NFATs that result from different stimulus (spike followed by plateau, a single spike and a low-level plateau). In Fig. 5, stimuli similar to experimentally used inputs (Fig. 5a) are entrained to our model and are also shown to result in three response patterns (Fig. 5b). The predicted patterns agree well with those seen in the experimental studies (Fig. 5b *Inset*).

An Optimal Forward Reaction Rate Favors the Passage. Many studies have highlighted the efficiency, sensitivity, and robustness of signal transduction networks (1, 4, 28). In this regard, the ladder network exhibits an interesting property, the existence of an optimal value for the forward reaction rates. Fig. 6 shows the mean, the most probable, and the root-mean-square of the FPT of a signaling protein from the C_0 state to the U_M . Clearly, there is an optimal forward reaction rate for the passage: The optimum occurs at $\alpha_f \approx 1$ for the mean FPT, but the optimal values of α_f for the most probable and the root-mean-square passage times are ≈ 2 .

The existence of this optimum may seem to conflict with the intuition that the higher enzyme concentration is, the shorter the passage time is. When the forward reaction rates are very slow, the forward reactions do indeed constitute a bottleneck. How-

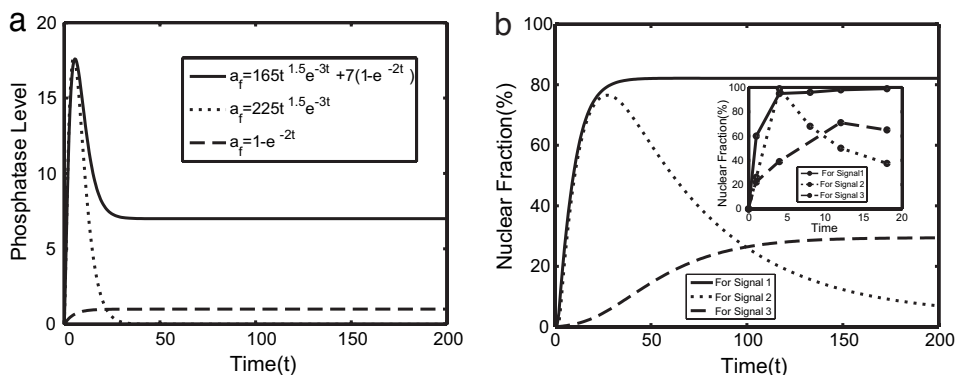


Fig. 5. Comparisons of experiments with the model. (a) The concentration of phosphatases in cytoplasm is as the input with three different patterns: spike followed by plateau (solid line), a single spike (dotted line), and a low-level plateau (broken line), which are the same as experimental stimuli (27). (b) Fractions of proteins in nuclear forms predicted from the model and those from experiments. Different lines correspond to different stimuli. The curves in the inset are experimental results from Dolmetsch *et al.* (27). Parameters are chosen as: the phosphorylation site number $M = 5$, the forward reaction rates $\beta_f = 1.0$, $\alpha_f = \beta_f = 0.1$, the backward rates $\alpha_b = \beta_b = 1.0$, $\alpha_b' = \beta_b' = 0.1$, the catalyzed rates $\alpha_c = \beta_c = 10.0$, $\alpha_c' = \beta_c' = 10.0$, the transport rates $k_+^M = 0.6$, $k_-^0 = 0.1$, and the ratio of transport rates $\gamma_+ = \gamma_- = 2.718$.

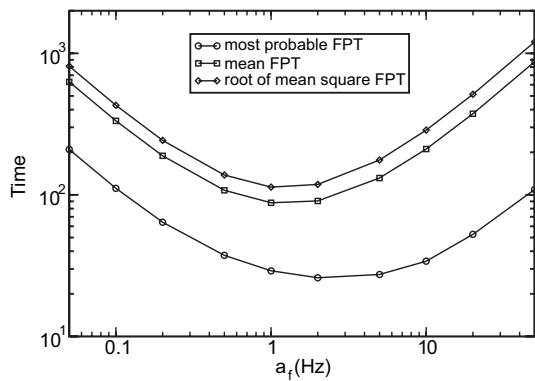


Fig. 6. An optimal forward reaction rate α_f favors efficient passage. The horizontal axis is the forward reaction rate swept from 0.05 to 50. The three curves represent the mean FPT, the root mean square FPT, and the most probable FPT according to the forward rate α_f . Other parameters are fixed as the phosphorylation site number $M = 5$, the forward reaction rates $\alpha_f = \beta_f = \beta_f = \alpha_f$, the backward rates $\alpha_b = \beta_b = \alpha_b = 1.0$, the catalyzed rates $\alpha'_c = \beta'_c = \alpha_c = 1.0$, the transport rates $k_+^M = k_-^0 = 0.2$, and the ratio of transport rates $\gamma_+ = \gamma_- = 2.718$.

ever, with increasing forward reaction rates, the walkers will be found more and more often in the signaling protein–enzyme complex forms; this helps signaling proteins move toward dephosphorylated states. Eventually, when the forward reaction rate is too large compared with the transportation rates k_+ and k_- , the signaling proteins will then spend most of their time in the transport incompetent complex forms (\bar{C}_i , \bar{C}_i , \bar{U}_i , or \bar{U}_i) rather than in the apo forms required for transport. This ultimately leads to slower transport between the cytoplasm and the nucleus.

This intriguing phenomenon could also be qualitatively explained as follows. The passage time of a signaling protein from the fully phosphorylated state in the cytoplasm (C_0) to the fully dephosphorylated state in the nuclear (U_M) consists of two parts: the time for dephosphorylation (T_d) and the time for transport (T_t). The typical time from site A to its neighbor B can be estimated by adding the inverse rates using the “one-step-forward” approximation. Assuming there are total l neighbors C_i , $i = 1, \dots, l$ (including $B = C_j$) that A can directly jump to with rate k_i , the typical time is $((k_i/\sum k_i) \times k_j)^{-1}$. For each dephosphorylation step from state i to $i + 1$, this estimate yields a time $(2\alpha_f + k_+^i)/\alpha_f^2 + (\alpha_b + \alpha_c)/\alpha_c^2$. The total dephosphorylation time is $T_d = M \times [(2\alpha_f + \bar{k}_+)/\alpha_f^2 + (\alpha_b + \alpha_c)/\alpha_c^2]$, where \bar{k}_+ is the average of k_+^i values. The transport time from C_i to U_i can

be approximated as $(2\alpha_f + k_+^i)/k_+^i$ based on similar approximations. Because the transport could happen in any C_i state, the mean transportation time requires averaging over all i , which results in a time $T_t = (2\alpha_f + \bar{k}_+)/\bar{k}_+^2$. The total mean passage time can therefore be approximated as

$$T = \left\{ M \left[\frac{2\alpha_f + \bar{k}_+}{\alpha_f^2} + \frac{\alpha_b + \alpha_c}{\alpha_c^2} \right] + \frac{2\alpha_f + \bar{k}_+}{\bar{k}_+^2} \right\}. \quad [7]$$

The existence of an optimal forward reaction rate requires the existence of a solution to the equation $\partial T/\partial \alpha_f = 0$ has solution(s). As a result, we must have the equation $\alpha_f^3 - M\bar{k}_+^2\alpha_f - M\bar{k}_+^3 = 0$. With the parameters shown in Fig. 6, the optimal forward rate is estimated to be 0.49, which is comparable with the exact solution (≈ 1) as shown in Fig. 6.

Asymmetry of the Reaction Network Affects the Passage. Although the ladder network is symmetrical in topology, the reaction rates in the cytoplasm and nucleus are quite different owing to different availabilities of appropriate enzymes. To illustrate the effect of the network’s asymmetry on signal transduction, we employ two parameters to probe this: the phosphorylation asymmetry parameter and the transport asymmetry parameter. The phosphorylation asymmetry parameter is defined as $\theta = a'_f/a_f = a'_b/a_b = a'_c/a_c = b'_f/b_f = b'_b/b_b = b'_c/b_c$, where $\theta \leq 1$, which characterizes the preference for undergoing dephosphorylation compared with the phosphorylation processes in the different environments. The other parameter, transport asymmetry parameter ϕ , is defined as $\phi = a_f/b_f = a_b/b_b = a_c/b_c = a'_f/b'_f = a'_b/b'_b = a'_c/b'_c$, which characterizes the relative activities of reactions in the cytoplasm compared with the corresponding reactions in the nucleus.

Fig. 7 illustrates the effects of these asymmetries on the network behavior. Fig. 7 *Left* shows that increasing the phosphorylation asymmetry slows down signaling; Fig. 7 *Right* shows that increasing the transport asymmetry speeds up signaling.

The FPT Distribution Has a Long Tail. The probability distribution of the FPT has a long tail due to the network’s hierarchical structure. In the limits of either large or small forward reaction rates, the probability distribution is very flat and the tail can be extremely long.

With the long-tail distribution, the mean FPT can be greatly different from the most probable FPT. In Fig. 3, the most probable FPTs for $\alpha_f = 0.5, 0.2, 0.1$ are 40, 60, and 100 s, respectively. However, the ratio of the mean FPT and the most probable FPT ranges from on the order of 1 to the order of 3 with the chosen set

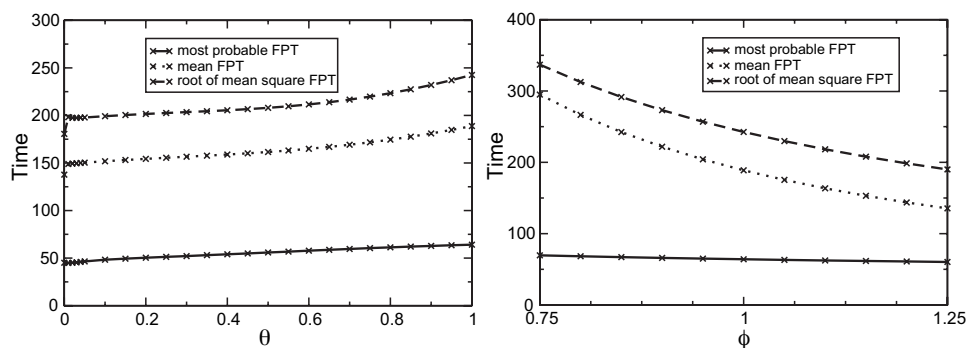


Fig. 7. Effect of the asymmetry of the reaction networks to the passage time. (*Left*) A plot of characteristic times as functions of the parallel asymmetry parameter θ . The other parameters are fixed at the phosphorylation site number $M = 5$, the forward reaction rates $\alpha_f = \beta_f = 0.2$, the backward rates $\alpha_b = \beta_b = 1.0$, the catalyzed rates $\alpha_c = \beta_c = 1.0$, the transport rates $k_+^M = k_-^0 = 0.2$, and the ratio of transport rates $\gamma_+ = \gamma_- = 2.718$. (*Right*) The plot of characteristic times as functions of rung asymmetry parameter ϕ . The other parameters are fixed at $M = 5$, $\alpha_f = \beta_f = 0.2$, $\alpha_b = \beta_b = 1.0$, $\alpha_c = \beta_c = 1.0$, $k_+^M = k_-^0 = 0.2$, $\gamma_+ = \gamma_- = 2.718$.

of parameters. Fig. 6 shows that the ratio of the most probable FPT and mean FPT is large even on a logarithm scale.

Conclusions and Discussions

In this paper, we studied a general signal transduction network-reaction ladder network that models multiple-site phosphorylation and cytonuclear transport. As often happens for real networks, the enzymes are assumed to be abundant, and so each signaling protein independently wanders through its various states of phosphorylation and location in compartments. Except for the last binding step, the network is effectively linear and therefore can be solved exactly. This exact solution is confirmed by using the Monte Carlo simulation. Even this simple network exhibits several interesting stochastic features. It exhibits a long tail of the probability distribution for the signaling time and there is an optimal forward reaction rate that favors speedy signaling.

This optimum suggests there may be an optimal amounts of enzymes for efficient signal transduction.

When available enzymes are not abundant, the random walkers on the network become correlated because they must share the limited resources of available enzymes. Nonlinearity then starts to play an important role throughout the process not just in the acquiring last step of the final target. In this situation when an enzyme is bound, several phosphorylation or dephosphorylation events will occur before unbinding, similar to the processive aspect of transcription (29). It will be interesting to study such scenarios and to study the effects of oscillatory upstream signaling molecules on the network. These will help us further understanding the phenomenology and quantitative design criteria of effective signal transduction mechanisms.

This work is supported by the Center for Theoretical Biological Physics through National Science Foundation Grants PHY0216576 and PHY0225630 and National Institutes of Health Grant GM69811-01.

- Huang C, Ferrell J (1996) *Proc Natl Acad Sci USA* 93:10078–10083.
- Changeux J, Edelstein S (2005) *Science* 308:1424–1428.
- Ferrel J (1996) *Trends Biosci* 21:460–466.
- Ferrel J, Machleder E (1998) *Science* 280:895–898.
- Roux PP, Blenis J (2004) *Microbiol Mol Biol Rev* 68:320–344.
- Schaeffer HJ, Weber MJ (1999) *Mol Cell Biol* 19:2435–2444.
- Doan T, Mendez A, Detwiler P, Chen J, Rieke F (2006) *Science* 313:530–533.
- Bhalla US, Lyengar R (1999) *Science* 283:381–387.
- Crabtree GR, Olson EN (2002) *Cell* 109:S67–S79.
- Hogan GD, Chen L, Nardone J, Rao A (2003) *Genes Dev* 17:2205–2232.
- Horsley V, Pavlath GK (2002) *J Chem Biol* 156:771–774.
- Shen T, Zong C, Hamelberg D, McCammon JA, Wolynes PG (2005) *FASEB J* 19:1389–1395.
- Salazar C, Höfer T (2003) *J Mol Biol* 327:31–45.
- Gunawardena J (2005) *Proc Natl Acad Sci USA* 102:14617–14622.
- Aubol BE, Chakrabarti S, Ngo J, Shaffer J, Nolen B, Fu X, Ghosh G, Adams JA (2003) *Proc Natl Acad Sci USA* 100:12601–12606.
- Samoilov M, Plyasunov S, Arkin AP (2004) *Proc Natl Acad Sci USA* 102:2310–2315.
- Monod J, Wyman J, Changeux JP (1965) *J Mol Biol* 12:88–118.
- McQuarrie D (1967) *Stochastic Approach to Chemical Kinetics* (Methuen, London), 1st Ed.
- Van Kampen NG (2003) *Stochastic Processes in Physics and Chemistry* (Elsevier, Oxford), Revised Ed.
- Lindenberg K, Seshadri V, Shuler K, Weiss G (1980) *J Stat Phys* 23:11–25.
- Bedeaux D, Lindenberg K, Shuler K (1971) *J MathPhys* 12:2116–2123.
- Yuste SB, Lindenberg K (1996) *J Stat Phys* 85:501–512.
- Rao A, Luo C, Hogan PG (1997) *Annu Rev Immunol* 15:707–747.
- Henkelman G, LaBute M, Tung C, Fenimore P, McMahon BH (2004) *Proc Natl Acad Sci USA* 102:15347–15351.
- Gillespie D (1976) *J Comput Phys* 22:403–434.
- Lu T, Volfson D, Tsimring L, Hasty J (2004) *IEE Systems Biol* 1:121–127.
- Dolmetsch RE, Lewis RS, Goodnow CC, Healy JJ (1997) *Nature* 386:855–858.
- Li G, Qian H (2003) *Cell Biochem Biophys* 39:45–59.
- Paulsson J, Ehrenberg M (2001) *Q Rev Biophys* 34:1–59.