Published in final edited form as:

J Phys Chem B. 2012 March 22; 116(11): 3630–3640. doi:10.1021/jp2102385.

The membrane environment can promote or suppress bistability in cell signaling networks

Steven M. Abel[†], Jeroen P. Roose[‡], Jay T. Groves[¶], Arthur Weiss[§], and Arup K. Chakraborty^{*,||}

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, Department of Anatomy, University of California, San Francisco, CA 94143, Howard Hughes Medical Institute, Department of Chemistry, and Biophysics Graduate Group, University of California, Berkeley, CA 94720, Physical Biosciences and Materials Sciences Divisions, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, Department of Medicine, Division of Rheumatology, Howard Hughes Medical Institute, University of California, San Francisco, CA 94143, and Departments of Chemical Engineering, Chemistry, and Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, Ragon Institute of MGH, MIT, and Harvard, Charlestown, MA 02129

Abstract

Many key biochemical reactions that mediate signal transduction in cells occur at the cell membrane, yet how the 2-dimensional membrane environment influences the collective behavior of signaling networks is poorly understood. We study models of two topologically different signaling pathways that exhibit bistability, examining the effects of reduced protein mobility and increased concentration at the membrane, as well as effects due to differences in spatiotemporal correlations between the membrane environment and 3-dimensional cytoplasm. The two model networks represent the distributive enzymatic modification of a protein at multiple sites and the positive feedback-mediated activation of a protein. In both cases we find that confining proteins to a membrane-like environment can markedly alter the emergent dynamics. For the distributive protein modification network, increased concentration promotes bistability through enhanced protein-protein binding, while lower mobility and membrane-enhanced spatiotemporal correlations suppress bistability. For the positive feedback-mediated activation network, confinement to a membrane environment enhances protein activation, which can induce bistability or stabilize a monostable, active state. Importantly, the influence of the membrane environment on signaling dynamics can be qualitatively different for signaling modules with different network topologies.

*To whom correspondence should be addressed: arupc@mit.edu.

[†]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

[‡]Department of Anatomy, University of California, San Francisco, CA 94143

[¶]Howard Hughes Medical Institute, Department of Chemistry, and Biophysics Graduate Group, University of California, Berkeley, CA 94720, Physical Biosciences and Materials Sciences Divisions, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

[§]Department of Medicine, Division of Rheumatology, Howard Hughes Medical Institute, University of California, San Francisco, CA 94143

^{||}Departments of Chemical Engineering, Chemistry, and Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, Ragon Institute of MGH, MIT, and Harvard, Charlestown, MA 02129

Supporting Information Available

The Supporting Information includes a full description of the kinetic parameters used in the simulations, a discussion of the identification of bistability in the distributive reaction network, the bistability diagrams used to construct Figure 3 in the text, numerical results augmenting the discussion of Figure 5, and additional simulation results for systems with greater confinement and a system with coupled membrane and cytoplasmic regions. This information is available free of charge via the Internet at <http://pubs.acs.org>.

Keywords

signal transduction; chemical master equation; two-dimensional

Introduction

Diverse cell types detect environmental stimuli by means of protein receptors associated with the cell membrane. This results in signal transduction, a process which translates receptor engagement to a cellular response via a network of biochemical reactions. Many of the earliest signaling events occur at the cell membrane, rather than in the cytoplasm. Although the influence of the cell membrane on signal transduction has been difficult to study by traditional experimental methods, recent experiments have begun to reveal that the spatiotemporal dynamics of molecules at the membrane play an important role in shaping the qualitative features of signaling.¹⁻⁵

A mechanistic understanding of how the membrane environment affects the spatiotemporal dynamics of typical signaling reactions is not available. Seeking such an understanding requires confronting interesting theoretical questions that are pragmatically important. At the membrane, diffusion occurs in an effectively two-dimensional environment, the mobility of proteins is low, and recruited proteins can have a higher effective concentration than when in the cytoplasm. Early studies exploring consequences of the membrane on simple reactions were carried out by Adam and Delbrück, who recognized that cytoplasmic proteins could potentially decrease the expected time to find a membrane-bound target protein by first associating with the membrane and then diffusing in two dimensions.^{6,7} Although subsequent studies have explored effects of concentration and two-dimensional diffusion on signaling reactions at the membrane,⁸⁻¹⁵ most theoretical and computational studies of signal transduction have neglected the presence of the membrane, and a comprehensive framework for understanding the spatiotemporal dynamics of signaling at the membrane remains to be established. In particular, it is not understood how stochastic effects, spatiotemporal correlations, and differences in protein concentration and mobility influence signaling modules with different topologies.

In an effort to gain insight into the influence of the membrane on signal transduction, we use computational methods that explicitly account for stochastic fluctuations to study two reaction networks that exhibit bistability, focusing on the roles of diffusion, dimensionality, and confinement on the presence and characteristics of bistability. Bistability is an important feature of many signaling networks, and can be found in signaling pathways that allow cells to make digital choices, such as those relevant to cell division, apoptosis, and T cell activation against foreign antigen.¹⁶ Bistable signaling networks can reside in one of two distinct, stable steady states. Switching between the states can occur by external influences that perturb the signaling network (e.g., receptor stimulation) or by spontaneous, fluctuation-driven transitions.

The two reaction networks considered here are (i) the distributive modification of a substrate protein with two phosphorylation sites, and (ii) a network describing the activation of Ras, which includes a positive feedback loop involving the guanine exchange factor called Son of sevenless (SOS). Both of these networks may be relevant in early T cell signaling events that lead to the decision of whether a T cell will activate against a particular peptide. The first network is a caricature of the Src kinase-mediated phosphorylation of multiple ITAMs in the cytoplasmic tail of the T cell receptor. The second is an important early event in T cell signaling as active Ras proteins can initiate numerous downstream signaling pathways. Signaling networks in diverse cell types also have similar signaling modules.

In both networks, stochastic effects, spatial inhomogeneities that influence protein-protein interactions, and protein mobility may play an important role in shaping collective properties of the dynamics. Traditionally, the effects of diffusion in reaction-diffusion systems have been treated by modifying the kinetic rates in well-mixed kinetic equations to account for the time it takes molecules to diffuse into contact with reaction partners.¹⁷ This framework, originally formulated for infinite three-dimensional systems, becomes more difficult in two dimensions, although much effort has been put into understanding simple reaction schemes.^{18,19} The difficulty in two dimensions, using the language of enzyme-substrate binding, arises from the fact that the rate at which substrate molecules arrive at an enzyme due to diffusion cannot keep up with the rate at which the substrate is bound and transformed, leading to a time-dependent effective reaction rate that decreases with time. Existing theoretical methods are not readily applied to finite systems of arbitrary geometry,²⁰ and recent work has demonstrated that spatiotemporal correlations in particle motion not accounted for by these methods can affect the collective behavior of signaling networks.²¹ Features of the membrane environment are likely to influence effective kinetic rates and correlations in molecular motions, making the qualitative behavior of the signaling networks difficult to capture using traditional methods. Indeed, our results demonstrate that for the two reaction networks we study, the membrane environment can exhibit qualitatively different behavior than a cytoplasmic environment due to the interplay between protein concentration, diffusion, and dimensionality. Moreover, a particular feature that differs between the membrane and cytoplasmic environments can have qualitatively different effects, depending upon the topology of the signaling network.

Methods

Spatial and temporal dynamics of the signaling networks are modeled as discrete-space, continuous-time, stochastic reaction-diffusion processes. Space is discretized into a cubic lattice, with each lattice site having a size commensurate with a few proteins. Each site is assumed to be well mixed, and reactions occur in a stochastic manner between molecules occupying the same lattice site. We assume diffusive motion of proteins on scales larger than a lattice site, and each molecule hops with a rate determined by its diffusion coefficient to a randomly chosen nearest neighbor lattice site. The lattice spacing is $0.01 \mu\text{m}$, and reflecting boundary conditions are used. The state of the system is completely specified by the number and type of molecules at each spatial location. Given an initial state of the system, the time evolution of the state of the system is governed by the spatially resolved chemical master equation,²² which provides an equation for the time evolution of the probability distribution over all possible states. Simulations are performed using our previously developed Stochastic Simulation Compiler,²³ which exactly samples trajectories for systems of arbitrary geometry using rule-based methods (rule based methods for homogeneous systems are used in Ref. 24).

To explore the effects of confining reactions near the membrane, we simulate reactions in a slab in which one of the dimensions is systematically varied while keeping the other dimensions and the number of molecules fixed. The system dimensions are $1\mu\text{m} \times 1\mu\text{m} \times l$, with l , the “confinement length,” allowed to vary. Confining the system to a single layer of lattice sites gives an effectively two-dimensional system. We additionally vary the diffusion coefficient to mimic the effects of the membrane on diffusion. Typical membrane diffusion coefficients, which reflect membrane-associated protein motion in a cellular environment, range from 0.01 to $0.1 \mu\text{m}^2/\text{s}$,^{15,17} while typical cytoplasmic diffusion coefficients range from 1 to $10 \mu\text{m}^2/\text{s}$.²⁵

The slab-like system geometry provides a conceptual framework by which to gain insight into how physical features of the membrane environment affect signaling networks. By

independently controlling protein mobility and the confinement length, we can make the system either more cytoplasm-like or more membrane-like. Additionally, the confined geometry provides a model for cells such as T cells and B cells, in which the cell nucleus can occupy a large fraction of the cell interior. In such cells, the gap between the nuclear and plasma membranes can be small ($< 0.4 \mu\text{m}$), leading to slab-like environments in which the spatial extent of one dimension is much smaller than the extent of the other dimensions.

Results

Distributive reaction network

A distributive reaction network is one in which a substrate molecule is catalytically modified multiple times by an enzyme that must unbind from the molecule between each catalytic step. Here we consider the simple example of a substrate protein containing two residues that can each be either phosphorylated or unphosphorylated. A kinase protein catalyzes the phosphorylation of each site, and a phosphatase protein catalyzes the dephosphorylation of each site. We denote the state of the substrate protein by distinguishing between three states: no sites phosphorylated (S_0), one of the sites phosphorylated (S_1), or both sites phosphorylated (S_2). Denoting the kinase by E , the phosphatase by P , and inactive forms of the enzymes by E^* and P^* , the reaction network considered is:

This generic distributive reaction network has been used by Takahashi et al. as a model for a portion of the mitogen-activated protein kinase cascade.²¹ We choose kinetic parameters to be consistent with their work. We vary the value of k_d , which reflects the time required for a kinase to exchange ADP with ATP, over several orders of magnitude as it is unknown for most protein kinases. We find its value to be qualitatively important when considering the effects of the membrane environment, which highlights the importance of experimentally determining its value for protein kinases.

This network model is sufficiently general to provide insight into signaling pathways that contain opposed distributive modification cycles, provided appropriate kinetic parameters are chosen. For example, the phosphorylation of ITAM regions on the ζ -chain of the T cell receptor by the tyrosine kinase Lck involves multiple kinase modifications. It is unknown whether this reaction occurs by a distributive or the alternative processive mechanism, in which an enzyme does not have to unbind from a substrate protein between catalytic events. Additionally, it is unknown how the membrane environment might affect qualitative cellular-level behavior of distributive reaction networks, so principles gleaned from simulations such as those in this paper may prove useful in designing experiments to distinguish whether a phosphorylation cascade occurs via a distributive mechanism.

It was demonstrated by Markevich et al. that a distributive reaction network can exhibit bistable behavior.²⁶ The physical argument is based on the sequestration of enzymes by the substrate molecule, and a necessary condition for the presence of bistability is that the

number of substrate molecules must exceed the number of at least one type of enzyme.²⁷ Here we consider the case in which the substrate molecules outnumber each type of enzyme. When bistability is present, in one steady state, fully phosphorylated proteins (S_2) outnumber the unphosphorylated proteins (S_0), while in the other steady state, the relative numbers are reversed. Consider a system that resides in the steady state with mostly unphosphorylated proteins. Most kinases will be bound and sequestered by S_0 proteins. Since there are relatively few S_2 proteins, there will be an excess of free phosphatase. Hence, when a kinase creates a singly phosphorylated substrate protein, the new S_1 protein is much more likely to bind to a phosphatase than a kinase in a well-mixed system. This provides a driving force to return the protein to the unphosphorylated state, thereby stabilizing the S_0 state. An analogous argument holds for the other steady state.

In three dimensions, Takahashi et al. recently showed that a smaller diffusion coefficient can change the characteristics of the distributive reaction network due to enhanced rebinding of substrate protein and enzyme after the first catalytic reaction.²¹ Rebinding of enzyme and substrate can result in a second catalytic event quickly following the first, leading to an effective processive-like step taking S_0 directly to S_2 . Processive enzyme catalysis cannot exhibit bistable behavior. It was also noted that increasing k_a served to weaken features associated with the distributive reaction network. Recent experimental results provide evidence that molecular crowding in mammalian cells, which presumably increases the likelihood of rebinding between proteins, converts the double phosphorylation of MAP kinase proteins from a distributive to a processive mechanism.²⁸

Our focus is on the effects of confining the distributive reaction network to a 2-dimensional membrane-like environment. Features of the membrane include increased concentration and smaller distances between proteins, lower protein mobility, and differences in the properties of random walks in two and three dimensions. Our simulation results parse the consequences of the interplay between these effects, which are difficult to intuit.

In our simulations, we use a typical membrane protein concentration of 100 molecules/ μm^2 for the substrate protein, using initial conditions of $N_{S_2} = N_{S_0} = 50$ and $N_E = N_P = 25$, with all other species and complexes initially unpopulated. We simulate many independent dynamic trajectories in which the position and state of all proteins are tracked in time. Initially, the molecules are placed uniformly at random in space. Each trajectory corresponds to an experiment with an isolated patch of membrane. Starting simulations with $N_{S_2} = N_{S_0}$ and all enzymes unbound gives an unbiased starting point: for a bistable system, half of the trajectories (on average) should go to each of the steady states.

Confining the distributive network promotes bistability through increased concentration

Results in Figure 1 show the effects of decreasing the confinement length. For each simulation, we hold constant the number of proteins, the diffusion coefficient, and the rate of enzyme activation. The diffusion coefficient is characteristic of cytoplasmic diffusion and the value of k_a is of the same order of magnitude as that estimated by recent experiments measuring the half-life of ADP release after phosphorylation for a serine-threonine protein kinase.²⁹ At large l , the distribution of the number of S_2 molecules at steady state is unimodal, and as the system is confined, the distribution at first broadens. At sufficiently small l , the distribution becomes bimodal, which implies that a population of isolated experiments under these conditions would partition into two subpopulations, one containing cells with most substrate proteins fully phosphorylated and one containing cells with most substrate proteins unphosphorylated. This trend continues for values of l smaller than those considered in Figure 1, as is demonstrated in the Supporting Information. Stochastic fluctuations at short times determine which state a trajectory reaches.

A dominant feature as the system becomes confined is that the concentration of molecules increases. This favors reactions involving the binding of enzyme and substrate proteins, leading to a higher fraction of bound enzyme-substrate pairs. This enhances the sequestration effect that is necessary for bistability in distributive reaction networks. This results in the emergence of bistability as the system is confined, with more confinement resulting in less frequent fluctuation-driven transitions between states.

Decreasing protein mobility and/or increasing the rate of enzyme activation suppresses bistability

The membrane environment can also suppress bistability because of smaller diffusion coefficients. A transition from bistable to monostable behavior can be seen by decreasing the diffusion coefficient while holding the confinement length (l) and other parameters fixed. At sufficiently small values of D , the system is monostable, with a unimodal distribution of N_{S_2} . Increasing D , when l is sufficiently small, eventually leads to bistable behavior, with larger D giving rise to more separated states that are less likely to undergo fluctuation-driven switching. The overall dependence of the bistability on D and l , with $k_a = 0.7 \text{ s}^{-1}$, is summarized in the “bistability diagram” in Figure 2. The effect of increasing k_a , and hence reducing the refractory time of enzymes, is shown in Figure 3. Experimental studies measuring the rate of ADP release for the protein kinase PKA estimate a rate 30 to 50 times larger than $k_a = 0.7 \text{ s}^{-1}$, suggesting that enzyme turnover rates may vary over orders of magnitude, depending on the kinase.³⁰ In general, as k_a increases, greater confinement and higher mobility are necessary to support bistability.

Protein mobility affects features of bistability in part because the transport of molecules influences the reaction kinetics.¹⁷ As mobility decreases, effective diffusion-influenced kinetic rates are reduced, leading to a smaller fraction of bound enzyme-substrate pairs. This weakens the sequestration effect necessary for bistability. At sufficiently high mobility, effective kinetic rates approach their well-mixed values, yielding the maximum sequestration effect.

Both protein mobility and the rate of enzyme reactivation (or refractory time of the enzyme) can affect features of bistability through their effect on the rebinding between enzyme and substrate molecules. The half-life of an inactive enzyme associated with $k_a = 0.7 \text{ s}^{-1}$ is approximately 1 s. In this time, the characteristic distance traveled by the enzyme in an unbounded 2-dimensional domain is $(4Dt)^{1/2} = 2 \text{ }\mu\text{m}$. This distance is of the same order as the system size simulated, which implies that the location of the enzyme upon reactivation will be weakly correlated with its position immediately after catalysis. Thus, a processive-like step is unlikely to occur as spatial correlation between the enzyme and substrate molecule does not persist on time scales associated with enzyme reactivation. For $k_a = 0.7 \text{ s}^{-1}$, rebinding plays a minimal role compared with the influence of transport on the effective kinetic rates for the values of D considered. The effects of rebinding become more prominent as k_a increases and D decreases, since spatial correlation between the enzyme and substrate molecule is more likely to persist for times associated with the reactivation of the enzyme. For a given value of l and D , the effects of diffusion on the transport of molecules are expected to be similar. This suggests that differences in the line of conditions separating bistable and monostable responses as k_a is varied (Figure 3) are due primarily to rebinding effects. Re-binding plays a more prominent role with increasing k_a , converting the distributive mechanism to an effectively processive one (which does not support bistability). A larger value of k_a at a given value of l , requires higher protein mobility to minimize rebinding and support bistability, which is similar to the results in Ref. 21.

Decreasing protein mobility enhances rebinding for multiple reasons: (i) the enzyme is more likely to become active before hopping occurs; (ii) once an enzyme is active and in contact

with S_1 , a binding event is more likely to occur; and (iii) if hopping occurs before enzyme activation, the enzyme and substrate protein are closer together on average when the enzyme becomes active. To gain insight into the importance of rebinding in the distributive reaction network, we developed a quantitative estimate of the probability of rebinding and catalysis between one E^* and one S_1 protein, given they just reacted and occupy the same lattice site (see the Appendix for details). Figure 4 shows that the effects of rebinding are much more pronounced with $k_d = 700 \text{ s}^{-1}$, in both two and three dimensions, and that rebinding is more significant in two dimensions.

At fixed concentration, shape of the signaling volume influences bistability

To emphasize that the emergence of bistability upon confinement (change in geometry) is not simply a concentration effect, we examined the steady state distribution of N_{S_2} in two slab-shaped systems that have equal volume and contain equal numbers of molecules. However, they differ in shape, with dimensions of $2\mu\text{m} \times 2\mu\text{m} \times 0.01\mu\text{m}$ and $1\mu\text{m} \times 1\mu\text{m} \times 0.04\mu\text{m}$. Although both systems have the same concentration of proteins, differences in shape affect the distribution of distances between proteins, with the average distance between neighboring molecules smaller in the second case. This holds since, in general, the expected distance from a molecule to its nearest neighbor is well approximated by setting $\rho V(r) \sim 1$, where $V(r)$ is the volume of the region contained within a sphere of radius r , ρ is the concentration of molecules, and $\rho V(r)$ gives the expected number of molecules within the region. For the two systems considered in this section, the average distance between neighboring molecules is larger than the slab thickness, so we can approximate the volume within radius r of a point at the center of the slab as $V(r) \sim \pi r^2 \lambda$, where λ denotes the slab thickness. Using $\rho V(r) \sim 1$ and solving for r , it follows that the expected distance to a nearest neighbor scales as $\lambda^{-1/2}$. Thus, at constant concentration, as the slab thickness increases, the average distance between neighboring molecules decreases. This result holds for confined regions in which the slab thickness is smaller than the average distance between neighboring molecules. At sufficiently large values of λ , the average distance between nearest neighbors approaches a constant value.

For the two systems considered here, the average distance between neighboring molecules in the less confined system is roughly half the distance between molecules in the more confined system. A more complete analysis, presented in the Supporting Information, shows that the average distance from a molecule to its k^{th} nearest neighbor is smaller in the less confined system. The smaller average distance between molecules reduces the effect of diffusion on the effective kinetic rates and reduces the likelihood of rebinding, since molecules must diffuse a smaller distance before they become effectively well mixed. This is reflected in the steady state distributions shown in Figure 5: while bistability is present in both systems, the stable states are more separated when $\lambda = 0.04 \mu\text{m}$, with less frequent fluctuation-induced switching between steady states. Notice that, in the case in which concentration is fixed, a difference in shape (less confinement) promotes bistability. This result shows that differences in the time (or distance) required for a protein to encounter other proteins after its last signaling act can result in qualitative differences in the input-output characteristics of signaling modules on cell membranes. In the limit of very high mobility, the systems are well mixed and shape does not affect the dynamics.

Positive feedback regulation of Ras activation

Ras is a small G protein (GTPase) that plays an important role in many signaling pathways (e.g., T cell receptor and growth factor receptor signaling).³¹ Ras is inactive when bound to GDP and active when bound to GTP. In its active state it can enzymatically modify numerous substrate proteins involved in signal transduction, particularly those involved in cell proliferation.³²

Ras activity is regulated by the GTPase-activating protein RasGAP and the guanine nucleotide exchange factors RasGRP and SOS. RasGAP, when bound to Ras, enhances the rate of conversion of GTP to GDP, thus promoting the inactivation of Ras. RasGRP and SOS both catalyze the exchange of GDP with GTP, thus activating Ras. SOS has been shown to have both an allosteric and a catalytic pocket that bind to Ras. The activation of Ras leads to more SOS molecules with RasGTP bound to the allosteric pocket, and an enhancement in the rate of catalysis.³³ Thus, there is a SOS-mediated positive feedback loop in the Ras activation network, with RasGTP enhancing the production of additional RasGTP. This positive feedback loop can lead to bistability, hysteresis, and digital signaling.³⁴

Ras is a membrane-anchored protein, and RasGRP and SOS are recruited to the membrane by diacylglycerol (DAG) and adaptor proteins, respectively. The degree to which they are recruited to the membrane is regulated by upstream signaling pathways. The importance of the membrane environment in Ras activation by SOS is illustrated by recent experiments in which the rate of Ras activation is greatly enhanced when membrane-bound Ras is used instead of Ras in solution.³⁵ We consider the effects of membrane confinement and diffusion on the following simple representation of the Ras activation network (adapted from Das et al.³⁴). We denote RasGDP by D and RasGTP by T, and we assume the allosteric pocket of SOS must be occupied for catalytic activity.

This network exhibits bistability in appropriate parameter regimes. When the network is bistable, the steady states can be characterized by the number of active Ras molecules. One state is “active” and contains a relatively large number of RasGTP, while the other state is “inactive” and contains relatively few RasGTP. To test for the presence of bistability, trajectories are started from two different sets of initial conditions. In each case, all molecules begin unbound and are uniformly distributed in space. In one case all Ras are in the active form and in the other case all Ras are in the inactive form. If the system is bistable, trajectories starting from the all-RasGDP state will likely fall into the basin of attraction for the inactive state, while trajectories starting from the all-RasGTP state will likely fall into the basin for the active state. If the system does not exhibit bistability, then all trajectories will evolve in time toward a common state. In bistable systems, it is common for one steady state to be more stable than the other. In our stochastic simulations, this is reflected by stochastic switching from the less stable state to the more stable state on the time scale of the simulations. As such, in simulations with steady states of disparate stability, the less stable steady state is only transiently populated by trajectories starting in its basin of attraction.

Membrane confinement of the Ras network and/or decreasing protein mobility promotes Ras activation

As before, we simulate the Ras network in a slab geometry which is confined in one spatial dimension, varying the diffusion coefficient as well as the number of DAG and RasGRP molecules. The simulations are performed with 150 Ras molecules, 80 SOS molecules, 8 RasGAPs, and a variable number of DAG and RasGRP molecules ($N_{\text{DAG}} = N_{\text{GRP}} = 0$; $N_{\text{DAG}} = 6$, $N_{\text{GRP}} = 20$; and $N_{\text{DAG}} = 24$, $N_{\text{GRP}} = 80$). The three most confined systems all exhibit bistable behavior (Figure 6), but with $l = 0.01\mu\text{m}$, trajectories from the inactive state spontaneously switch to the active state, and with $l = 0.03\mu\text{m}$, trajectories from the active state spontaneously switch to the inactive state. This gives an indication of the relative stability of each steady state. At $l = 0.04\mu\text{m}$, the system is at the limit of bistability, and most trajectories from the active regime quickly populate the inactive state.

These results can be understood in terms of concentration effects: as the system is confined, the concentration of proteins increases, promoting binding between proteins. This is the same principle that resulted in enhanced binding between substrate proteins and enzymes in the distributive reaction network. Here, more concentrated conditions favor the binding of Ras molecules to the allosteric and catalytic pockets of SOS, which enhances the production of RasGTP. At sufficiently high levels, RasGTP can initiate and sustain positive feedback, which is necessary for bistability.

Along with the results for $D = 1\mu\text{m}^2/\text{s}$ (Figure 6), Figure 7 shows that reducing the mobility of molecules decreases the stability of the inactive state relative to the active state. At high mobility, the system is bistable, with the active state more stable than the inactive state, while at low mobility, the system is monostable and resides in the active state. Lower mobility allows more cells to achieve the highly active state in the bistable system. Two effects promote stability of the active state as the mobility of proteins decreases. The first is that RasGAPs, which have the highest binding rate (k_{11}) to Ras, experience effects of lowering diffusion at higher mobility than other binding reactions; i.e., the fastest reaction becomes diffusion-influenced first. This serves to decrease the effective binding rate between RasGAP and RasGTP, thus increasing the survival time of an active Ras protein. This promotes positive feedback through SOS and can enhance the further production of RasGTP. This result highlights that different reactions are influenced differently upon reduction in protein mobility, and predicting the effects of diffusion on the network requires knowledge of rates of the biochemical reactions. Additionally, as the rate of diffusion decreases, rebinding between recently bound proteins becomes more likely. For example, RasGTP bound to the allosteric site of SOS becomes more likely to rebind, increasing its effective lifetime.

The average time that RasGTP is bound to the SOS allosteric pocket, per binding event, is $\tau = k_4^{-1}$. Two effects increase the effective lifetime: multiple bindings before hopping and multiple diffusive excursions and returns before the molecules become mixed. We treat the probability of returning as in the distributive reaction network and assume that if RasGTP and SOS diffuse beyond a characteristic distance apart, they are effectively mixed. This distance is set roughly by the expected distance to the nearest RasGDP molecule, which competes for the allosteric pocket of SOS. The number of bindings before hopping is geometrically distributed, as is the number of returns before diffusion beyond a cutoff distance. The expected number of each type of event is

$$\begin{aligned} E[\text{number of bindings before hopping}] &= \frac{p_{\text{on}}}{1-p_{\text{on}}} \\ E[\text{number of returns before mixing}] &= \frac{p_{\text{return}}}{1-p_{\text{return}}} \end{aligned}$$

Note that here p_{on} is the probability that RasGTP binds SOS before either molecule hops away. Putting these results together, the effective lifetime is

$$\begin{aligned} \tau_{\text{effective}} &= \left(1 + \frac{p_{\text{on}}}{1-p_{\text{on}}}\right) \tau + p_{\text{on}} \left(\frac{p_{\text{return}}}{1-p_{\text{return}}}\right) \left(1 + \frac{p_{\text{on}}}{1-p_{\text{on}}}\right) \tau \\ &= \left(1 + \frac{p_{\text{on}}}{1-p_{\text{on}}}\right) \left(1 + \frac{p_{\text{on}} p_{\text{return}}}{1-p_{\text{return}}}\right) \tau \end{aligned}$$

With $D = 0.01 \mu\text{m}^2/\text{s}$, the effective time (including rebinding) that RasGTP is bound to the allosteric pocket of SOS, compared to the case without rebinding, is 1.9 times longer in two dimensions and 1.3 times longer in three dimensions. Additionally, as mobility decreases, it is more likely for a recently produced RasGTP to stay in the proximity of the SOS molecule for long enough to enhance its binding to the allosteric pocket when RasGDP unbinds. Both of these effects serve to enhance the effects of positive feedback.

An additional feature of interest is that within the monostable regime, the expected time to reach the active steady state from the all-RasGDP initial condition varies non-monotonically with the diffusion coefficient, as is demonstrated by the response times reported in Figure 7. As the diffusion coefficient decreases from $0.1 \mu\text{m}^2/\text{s}$, the average time to reach the active state first decreases (faster response time); however, for sufficiently small D , the average time increases (slower response time). Above the crossover diffusivity, decreasing mobility decreases the response time by reducing the effectiveness of RasGAPs and enhancing positive feedback through rebinding effects. Below the crossover diffusivity, the response time is dominated by the time it takes for molecules to diffuse into contact. Hence, reducing D in this regime increases the overall response time. The non-monotonic response time is similar to a result from Ref. 21, in which lowering mobility first decreased response times due to enhanced rebinding but eventually increased response times due to longer molecular encounter times.

The bistability diagram for the Ras activation network is shown in Figure 8 and summarizes the influence of confinement, the diffusion coefficient, and the number of DAG and RasGRP on bistability (previous studies have explored the influence of RasGRP on Ras activation in well-mixed systems^{34,36}). There are three dominant effects at play in the results: (i) confinement increases the concentration of molecules, which enhances protein-protein binding and the effects of positive feedback; (ii) decreasing D enhances production of RasGTP through reduced effective RasGAP activity and enhanced positive feedback effects due to rebinding; and (iii) increasing N_{GRP} enhances the activation of Ras, thus biasing the system toward an active state.

Discussion

Our results reveal how certain general features of the membrane environment can qualitatively change the dynamics and collective behavior of signaling networks. The specific features we studied, compared with a 3-dimensional cytoplasmic environment, are increased protein concentration, decreased protein mobility, and different spatiotemporal correlations between molecules. The character of the influence of these features can depend upon the topology of the signaling network.

For the distributive reaction network, we find that effects of membrane confinement can either enhance or suppress bistability. Increased concentration at the membrane enhances binding between enzymes and substrate molecules, promoting the sequestration of enzymes and supporting bistability. Protein mobility and spatiotemporal correlations are closely intertwined, with correlations arising from the spatial proximity between molecules at times soon after they first interact and mobility influencing the time scale over which spatial correlations persist. The nature of the spatiotemporal correlations is different in the membrane environment because of lower protein mobilities and differences in the character of 2-dimensional and 3-dimensional random walks. The importance of spatiotemporal correlations is highlighted by the strong dependence of bistability on rebinding between enzymes and substrate proteins in the distributive reaction network. Additionally, protein mobility and the distribution of distances between proteins influence the time required for a molecule to diffuse to a reaction partner. This influences the effective kinetic parameters governing the system, with slower diffusion leading to smaller effective binding rates and relatively fast reactions being influenced before slower reactions upon lowering protein mobility. For example, in the Ras activation network, RasGAP binding experiences the effects of diffusion at the highest mobility, biasing the system toward active Ras as mobility decreases.

The last result also highlights that the effect of changing the same parameter due to membrane confinement can be different, depending upon the topology of the signaling network. Reduced protein mobility due to membrane confinement abrogates bistability for both reaction networks studied, but leads to qualitatively different outcomes. In the distributive reaction network, loss of bistability due to lower protein mobility results in a new monostable state with lower levels of active molecules compared to the more active bistable state. In sharp contrast, in the Ras activation network, loss of bistability due to lower mobility results in a monostable state with levels of active molecules similar to the previously more active bistable state. In one case, lower mobility leads to fewer active molecules, while in the other case, lower mobility promotes activation.

In our analysis of the two signaling networks, we used a lattice-based model in which proteins, considered to be well-mixed within a lattice site, diffuse by hopping between neighboring lattice sites. Other methods include continuous-space, particle-based methods,^{13,21} which could potentially give different rebinding properties on molecular length scales. The diffusion coefficients we use to model protein motion are typically reported for motion over length scales large compared to protein size, and treating motion on molecular length scales as diffusive may not be appropriate for membrane-bound proteins. Additionally, it is unlikely that the lattice and continuum methods would give qualitatively different behavior in the distributive reaction network, as the values of k_a governing the refractory time of the enzymes are typically much smaller than the rate at which proteins hop to nearest neighbor lattice sites. Hence, it is unlikely that the well-mixed assumption of the lattice site leads to an overestimate of rebinding at short times, as might be the case if k_a were sufficiently large. Additionally, in the Ras activation network, the two methods are likely to produce qualitatively similar results, since in both cases the time that Ras is effectively bound to the allosteric site of SOS is enhanced in two dimensions.

Studying the dynamics of signaling networks within a slab-like geometry gave physical insight into how the membrane environment can affect signaling networks. It will be of considerable interest to study systems in which molecules confined to a membrane interact with molecules that diffuse within the cytoplasm. When introducing the distributive reaction network, we discussed the phosphorylation of ITAM regions on the cytoplasmic tail of T cell receptors by the kinase Lck. In this biologically motivated network, the substrate molecule and kinase are confined the membrane, but the phosphatase can diffuse within the

cytoplasm. In Figure S5 in the Supporting Information, we demonstrate that such a system can exhibit bistability, although it is expected that details of the signaling output will depend on details of the system geometry.

A deep understanding of the role of the membrane in cell signaling will result from synergy between theory, computation, and experiments. Many of our results can be tested by studying model lipid bilayers with well-defined concentrations of tethered proteins and comparing their behavior with experiments in which the proteins interact in solution (such as in a cuvette) and in live cells. Various other effects of the membrane environment, such as those due to cytoskeletal dynamics and receptor orientation,³⁷ remain to be carefully studied. It will be especially important to use theory to disentangle specific effects from general principles in studying these influences of the membrane. Theoretical studies would benefit from techniques to account for diffusion-influenced kinetic parameters in arbitrary geometries, more sophisticated techniques to account for rebinding and spatiotemporal correlations between molecules, and methods to account for the spatially inhomogeneous and noisy environment in which signaling takes place.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by a NIH Director's Pioneer award (AKC) and NIH grant 1P01AI091580-01 (SMA, AKC, AW, JPR, JTG). We are grateful to Prof. John Kuriyan for important input and discussions.

References

1. Groves JT, Kuriyan J. *Nat Struct Mol Biol.* 2010; 17:659–665. [PubMed: 20495561]
2. Fooksman DR, Vardhana S, Vasiliver-Shamis G, Liese J, Blair D, Waite J, Sacristán C, Victora G, Zanin-Zhorov A, Dustin ML. *Annu Rev Immunol.* 2010; 28:79–105. [PubMed: 19968559]
3. Kholodenko BN, Hancock JF, Kolch W. *Nat Rev Mol Cell Biol.* 2010; 11:414–426. [PubMed: 20495582]
4. Singleton KL, Roybal KT, Sun Y, Fu G, Gascoigne NRJ, van Oers NSC, Wülfing C. *Sci Signal.* 2009; 2:ra15. [PubMed: 19351954]
5. Lee K, et al. *Science.* 2003; 302:1218–1222. [PubMed: 14512504]
6. Adam, G.; Delbrück, M. *Structural Chemistry and Molecular Biology.* Rich, A.; Davidson, N., editors. Freeman: San Francisco; 1968.
7. McCloskey MA, Poo MM. *J Cell Biol.* 1986; 102:88–96. [PubMed: 3001105]
8. Haugh JM, Lauffenburger DA. *Biophys J.* 1997; 72:2014–2031. [PubMed: 9129805]
9. Shea LD, Omann GM, Linderman JJ. *Biophys J.* 1997; 73:2949–2959. [PubMed: 9414209]
10. Kholodenko BN, Hoek JB, Westerhoff HV. *Trends Cell Biol.* 2000; 10:173–178. [PubMed: 10754559]
11. Haugh JM. *Biophys J.* 2002; 82:591–604. [PubMed: 11806904]
12. Elf J, Ehrenberg M. *Syst Biol.* 2004; 1:230–236.
13. Monine MI, Haugh JM. *Biophys J.* 2008; 95:2172–2182. [PubMed: 18502802]
14. Alonso S, Bär M. *Phys Biol.* 2010; 7:046012. [PubMed: 21149929]
15. Dushek O, van der Merwe PA, Shahrezaei V. *Biophys J.* 2011; 100:1189–1197. [PubMed: 21354391]
16. Ferrell JE. *Curr Opin Cell Biol.* 2002; 14:140–148. [PubMed: 11891111]
17. Lauffenburger, DA.; Linderman, JJ. *Receptors: Models for Binding, Trafficking, and Signaling.* Oxford University Press; USA: 1996.
18. Torney DC, McConnell HM. *Proc R Soc London, Series A.* 1983; 387:147–170.

19. Szabo A. *J Phys Chem.* 1989; 93:6929–6939.
20. Bénichou O, Chevalier C, Klafter J, Meyer B, Voituriez R. *Nat Chem.* 2010; 2:472–477. [PubMed: 20489716]
21. Takahashi K, Tanase-Nicola S, ten Wolde PR. *Proc Natl Acad Sci USA.* 2010; 107:2473–2478. [PubMed: 20133748]
22. Hattne J, Fange D, Elf J. *Bioinformatics.* 2005; 21:2923–2924. [PubMed: 15817692]
23. Lis M, Artyomov MN, Devadas S, Chakraborty AK. *Bioinformatics.* 2009; 25:2289–2291. [PubMed: 19578038]
24. Faeder JR, Blinov ML, Hlavacek WS. *Methods Mol Biol.* 2009; 500:113–167. [PubMed: 19399430]
25. Luby-Phelps K. *Int Rev Cytol.* 1999; 192:189–221. [PubMed: 10553280]
26. Markevich NI, Hoek JB, Kholodenko BN. *J Cell Biol.* 2004; 164:353–359. [PubMed: 14744999]
27. Ortega F, Garcés JL, Mas F, Kholodenko BN, Cascante M. *FEBS J.* 2006; 273:3915–3926. [PubMed: 16934033]
28. Aoki K, Yamada M, Kunida K, Yasuda S, Matsuda M. *Proc Natl Acad Sci USA.* 2011; 108:12675–12680. [PubMed: 21768338]
29. Keshwani MM, Harris TK. *J Biol Chem.* 2008; 283:11972–11980. [PubMed: 18326039]
30. Zhou J, Adams JA. *Biochemistry.* 1997; 36:15733–15738. [PubMed: 9398302]
31. Genot E, Cantrell DA. *Curr Opin Immunol.* 2000; 12:289–294. [PubMed: 10781411]
32. Campbell SL, Khosravi-Far R, Rossman KL, Clark GJ, Der CJ. *Oncogene.* 1998; 17:1395–1413. [PubMed: 9779987]
33. Margarit S, Sondermann H, Hall BE, Nagar B, Hoelz A, Pirruccello M, Bar-Sagi D, Kuriyan J. *Cell.* 2003; 112:685–695. [PubMed: 12628188]
34. Das J, Ho M, Zikherman J, Govern C, Yang M, Weiss A, Chakraborty AK, Roose JP. *Cell.* 2009; 136:337–351. [PubMed: 19167334]
35. Gureasko J, Galush WJ, Boykevisch S, Sondermann H, Bar-Sagi D, Groves JT, Kuriyan J. *Nat Struct Mol Biol.* 2008; 15:452–461. [PubMed: 18454158]
36. Riese MJ, Grewal J, Das J, Zou T, Patil V, Chakraborty AK, Koretzky GA. *J Biol Chem.* 2011; 286:5254–5265. [PubMed: 21138839]
37. Wu Y, Vendome J, Shapiro L, Ben-Shaul A, Honig B. *Nature.* 2011; 475:510–513. [PubMed: 21796210]
38. Redner, S. *A Guide to First-Passage Processes.* Cambridge University Press; 2001. This material is available free of charge via the Internet at <http://pubs.acs.org/>

Appendix: Probability of rebinding and catalysis

In this section, we present a quantitative estimate of the probability of rebinding and catalysis between one E^* and one S_1 protein, given they just reacted and occupy the same lattice site. These results also apply to the phosphatase reactions and include dependence on the dimensionality of the system, the rate of enzyme activation, and the concentration of molecules.

The basic idea behind the calculations is that once the initial reaction takes place, the enzyme and substrate molecule must remain “sufficiently close” to avoid becoming well-mixed and sequestered by other molecules in the system. We implement this condition by assuming that the enzyme and S_1 cannot diffuse beyond a characteristic distance apart, with the distance set by the expected distance to the nearest particle that can bind (and hence sequester) one of the molecules. Let ρ denote the concentration of such “traps” in this section. For simplicity, we assume the lattice is infinitely extended in all directions.

Active enzyme

We first calculate the probability that an active enzyme and S_1 molecule that occupy the same lattice site react before diffusing beyond a characteristic distance apart. In a subsequent section, we extend the result to the case in which the enzyme is initially inactive.

Given that the enzyme and substrate molecule occupy the same lattice site, the probability of binding before either molecule hops to a neighbor is

$$p_{\text{on}} = \frac{\gamma_{\text{on}}}{\gamma_{\text{on}} + \gamma_{\text{hop}}} = \frac{(k_4/a^3)}{(k_4/a^3) + 2n(D/a^2)},$$

with γ a total rate, k_4 the bulk on-rate between E and S_1 , a the lattice spacing, D the diffusion co-efficient, and n the number of nearest neighbor lattice sites. The probability of binding depends on the dimensionality, since $n = 2d$, where d denotes the dimensionality. The probability of hopping before binding is $p_{\text{hop}} = 1 - p_{\text{on}}$. Once the enzyme and substrate protein are bound, the probability of unbinding before reacting is

$$p_{\text{off}} = \frac{\gamma_{\text{off}}}{\gamma_{\text{off}} + \gamma_{\text{rxn}}} = \frac{k_5}{k_5 + k_6},$$

The probability of reacting before unbinding is $p_{\text{rxn}} = 1 - p_{\text{off}}$. For the distributive reaction network, $p_{\text{off}} \approx 0.10$.

Now assume that a hopping event occurs before a reaction takes place and that the enzyme and substrate protein occupy nearest neighbor lattice sites. Let p_{return} be the probability that, as they undergo random walks, they co-occupy the same lattice site before diffusing beyond the prescribed cutoff distance. We call this a “re-encounter” or “return” event and construct an approximation for p_{return} in the next section. Treating p_{return} as known, we can write a formally exact expression for the probability that E and S_1 , starting from the same lattice site, react before diffusing beyond the cutoff distance. We denote this probability by $\text{Pr}\{\text{react}\}$.

One way to construct $\text{Pr}\{\text{react}\}$ is to consider all possible sequences of events leading to a reaction. It is clear that the final two events must be the sequence in which the two molecules bind and then react (this occurs with probability $p_{\text{on}} p_{\text{rxn}}$). Before this event, any sequence of events leading to E and S_1 occupying the same lattice site without the molecules diffusing too far apart is allowed. The two molecules, given they start at the same lattice site, return to the same configuration after binding and then unbinding (this occurs with probability $p_{\text{on}} p_{\text{off}}$) or after hopping and then undergoing a diffusive excursion that returns to a common lattice site (this occurs with probability $p_{\text{hop}} p_{\text{return}}$). Accounting for all possible sequences ending with catalysis gives the probability of reaction

$$\begin{aligned} \text{Pr}\{\text{react}\} &= (p_{\text{on}} p_{\text{rxn}}) \sum_{l=0}^{\infty} \sum_{k=0}^l \binom{l}{k} (p_{\text{hop}} p_{\text{return}})^k (p_{\text{on}} p_{\text{off}})^{l-k} \\ &= (p_{\text{on}} p_{\text{rxn}}) \sum_{l=0}^{\infty} (p_{\text{hop}} p_{\text{return}} + p_{\text{on}} p_{\text{off}})^l \\ &= \frac{p_{\text{on}} p_{\text{rxn}}}{1 - p_{\text{hop}} p_{\text{return}} - p_{\text{on}} p_{\text{off}}} \end{aligned}$$

The following section is devoted to developing an approximation for p_{return} .

Probability of return

Begin with E and S_1 occupying nearest neighbor lattice sites. We wish to estimate p_{return} , the probability that they eventually occupy the same lattice site without diffusing beyond a fixed distance apart. We begin by making the assumption that space is continuous, with molecules distributed uniformly at random. Using properties of spatial Poisson processes, it is straightforward to show that the expected distance to a nearest “trap” molecule (distributed with concentration ρ) is

$$\begin{aligned} E_2[r] &= \frac{1}{2\sqrt{\rho}} \\ E_3[r] &= \frac{\Gamma(1/3)}{3} ((4/3)\pi\rho)^{-1/3} \approx (0.55)\rho^{-1/3}, \end{aligned}$$

where the subscripts denote two and three dimensions (the dimensions of the concentrations are not explicitly distinguished, but are assumed to be consistent with the dimensionality).

As an approximation to the lattice random walk problem, consider a particle diffusing in continuous space between two concentric, absorbing spherical boundaries. The particle diffuses with diffusion coefficient $D_{\text{tot}} = 2D$ and represents the relative distance between the two molecules. We wish to compute the probability that the particle reaches the inner sphere before it reaches the outer sphere, which corresponds to the enzyme and S_1 molecule reaching a common lattice site before diffusing beyond a cutoff distance. Let the radius of the inner sphere be half the lattice spacing, $a/2$, and let the particle start at position $r = a$ to mimic the nearest neighbor condition on the lattice. We use the expected distance to the closest trap as the radius of the outer sphere. This calculation falls into the category of first passage problems, a thorough review of which can be found in Ref. 38.

Noting the spherical symmetry in the problem, let $\Phi_{-}(r)$ denote the probability that the diffusing particle, starting at radial position r , reaches the inner boundary before reaching the outer boundary. With this notation, $p_{\text{return}} = \Phi_{-}(a)$. The function Φ_{-} is known to satisfy the Laplace equation, $\nabla^2 \Phi_{-} = 0$, subject to the boundary conditions $\Phi_{-}(R_{-}) = 1$ and $\Phi_{-}(R_{+}) = 0$ (here, $R_{-} = a/2$ and $R_{+} = E_d[r]$).³⁸

It is straightforward to solve these equations in two and three dimensions.³⁸ The return probability can be written

$$p_{\text{return}} = \begin{cases} \frac{\ln(R_{+}/a)}{\ln(R_{+}/R_{-})} & d=2 \\ \frac{\frac{1}{a} - \frac{1}{R_{+}}}{\frac{1}{R_{-}} - \frac{1}{R_{+}}} & d=3 \end{cases}$$

Inactive enzyme

Now consider the full case in which the inactive enzyme E^* and S_1 co-occupy the same lattice site. The rate of enzyme activation, k_a , is an important parameter, since it controls how long the enzyme is expected to remain inactive. The longer it is inactive, the less likely it is for rebinding to occur. The probability that the enzyme becomes active in the time interval from t_1 to t_2 is

$$\begin{aligned}\Pr\{\text{activation in } (t_1, t_2)\} &= \int_{t_1}^{t_2} dt k_a e^{-k_a t} \\ &= e^{-k_a t_1} (1 - e^{-k_a(t_2 - t_1)})\end{aligned}$$

We will use this result later.

We begin by estimating τ_{return} , the average time for a diffusive excursion in which the two molecules diffuse apart but return together before diffusing beyond the cutoff distance. To do this, we return to the problem of a particle diffusing between two concentric spheres and compute the expected time for the particle to reach the inner sphere, conditioned on not hitting the outer sphere. This time, starting from position r , is denoted by $t_{-}(r)$. Following Ref. 38, this quantity satisfies

$$D_{\text{tot}} \nabla^2 [\Phi_{-}(r) t_{-}(r)] = -\Phi_{-}(r)$$

subject to the boundary conditions $\Phi_{-}(r) t_{-}(r) = 0$ for $r \in \{R_{-}, R_{+}\}$. We have the solutions for Φ_{-} from above. In three dimensions, the solution for $\tau_{\text{return}} = t_{-}(a)$ is

$$\tau_{\text{return}} = \frac{(a - R_{-})(2R_{+} - R_{-} - a)}{6D_{\text{tot}}}$$

In two dimensions, the solution is more complicated but well-behaved:

$$\tau_{\text{return}} = \frac{R_{+}^2 \ln(a/R_{-}) + R_{-}^2 [1 + \ln(R_{+}/R_{-})] \ln(R_{+}/a) - a^2 [1 + \ln(R_{+}/a)] \ln(R_{+}/R_{-})}{4D_{\text{tot}} \ln(R_{+}/R_{-}) \ln(R_{+}/a)}$$

These equations are strictly increasing as a function of increasing R_{+} (decreasing ρ).

For each excursion taken by the molecules, we have the probability that they return and the average time for them to return, given they do not diffuse beyond a cutoff distance. Using these results, along with the distribution of times that the enzyme activates, we can estimate the probability that the proteins react before they diffuse apart.

Let $\Delta t = \tau_{\text{hop}} + \tau_{\text{return}}$, where $\tau_{\text{hop}} = (2nD/a^2)^{-1}$ is the average time for the molecules to hop apart once they co-occupy a site. For most values of k_a and D , the average time for the enzyme to become active is much longer than Δt , which implies that the two molecules are likely to diffuse apart many times before they can bind and react. Enzymes that activate earlier than average are more likely to contribute significantly to rebinding, since the molecules must undergo fewer excursions and returns. Once the enzyme is active and occupies the same lattice site as S_1 , we can use the probability of binding and catalysis from the previous sections.

In the remainder of the calculation, we break time into increments of size Δt . For each increment, we have the probability that the enzyme activates as well as the probability that the enzyme and substrate molecule have remained sufficiently close throughout the process. Let α denote the probability that E and S_1 react, given that they start at the same lattice site (computed previously). Then the probability that E^* and S_1 react can be approximated as

$$\Pr\{S_1 \rightarrow S_2\} = (1 - e^{-k_a \tau_{\text{hop}}}) \alpha + \sum_{k=0}^{\infty} \left[e^{-k_a(k\Delta t)} e^{-k_a \tau_{\text{hop}}} (1 - e^{-k_a \Delta t}) \right] \left[(p_{\text{return}})^{k+1} \right] \alpha$$

The first term accounts for cases in which the enzyme activates before any hops have occurred. Terms within the sum account for the probability of $k + 1$ successful returns without diffusing beyond the cutoff distance (second term in square brackets), as well as the probability that the enzyme activated within the time interval from $\tau_{\text{hop}} + k\Delta t$ to $\tau_{\text{hop}} + (k + 1)\Delta t$ (first term in square brackets). Summing the geometric series and substituting the expression for α , we obtain the desired result

$$\Pr\{S_1 \rightarrow S_2\} = \left[(1 - e^{-k_a \tau_{\text{hop}}}) + e^{-k_a \tau_{\text{hop}}} (1 - e^{-k_a \Delta t}) \left(\frac{p_{\text{return}}}{1 - p_{\text{return}} e^{-k_a \Delta t}} \right) \right] \times \left(\frac{p_{\text{on}} p_{\text{rxn}}}{1 - p_{\text{hop}} p_{\text{return}} - p_{\text{on}} p_{\text{off}}} \right)$$

This probability is plotted as a function of ρ , D , k_a and d in Figure 4.

Figure 1. Decreasing the confinement length promotes the emergence of bistability in the distributive reaction network

(A–E) Each histogram summarizes the distribution of the final number of S_2 molecules from 500 independent trajectories (1000 s each). The diffusion coefficient is $D = 1\mu\text{m}^2/\text{s}$, the rate of enzyme activation is $k_a = 0.7\text{ s}^{-1}$, and the initial conditions are the same for each trajectory. Bimodality, which reflects an underlying bistability, emerges as the system becomes more confined. Within the bistable regime, additional confinement reduces the frequency of spontaneous state switching. (F) The number of S_2 molecules as a function of time for two stochastic trajectories which reach different steady states ($l = 0.01\mu\text{m}$).

Figure 2. Bistability diagram for the distributive reaction network

At fixed number of molecules and with $k_d = 0.7 \text{ s}^{-1}$, the confinement length (l) and diffusion coefficient (D) together determine whether the network exhibits bistability. The dashed line is simply intended to guide the eye in distinguishing between the bistable (circles) and monostable (diamonds) regions. The color scheme indicates the degree of separation between the two modes in the bistable region, with hotter colors indicating greater separation between modes. The degree of separation is measured by the standard deviation of the N_{S_2} distribution (see Supporting Information for details). There exists a confinement length l^* , such that for $l > l^*$, the system is monostable for all D . Analysis in the well-mixed limit ($D \rightarrow \infty$) gives l^* to be approximately $0.6 \mu\text{m}$.

Figure 3. Approximate dividing lines between monostable and bistable regions for various values of k_a

For a given choice of D and l , productive rebinding between an enzyme and substrate protein is more likely for larger values of k_a . As a result, the transition from bistable to monostable behavior occurs at larger values of the diffusion coefficient for larger values of k_a , which takes on values of 0.7 s^{-1} , 7 s^{-1} , 70 s^{-1} , and 700 s^{-1} . The full bistability diagrams are presented in the Supporting Information.

Figure 4. Rebinding is more pronounced in two dimensions and for smaller enzyme refractory times (larger k_a)

The color scheme indicates the approximate probability for an enzyme and substrate protein to rebind and react given a first catalytic event just occurred. The molecules are assumed to be unable to rebind if they diffuse farther apart than the mean distance separating nearest neighbor molecules in the system.

Figure 5. The shape of the confining region, at fixed volume and concentration, affects features of the bistability

The average distance between particles is larger in the system corresponding to the top figures, which leads to slower effective kinetic rates and a greater likelihood of rebinding between an enzyme and substrate protein after the first catalytic step.

Figure 6. Confining the Ras activation network affects properties of the bistability at fixed protein mobility

Here, $D = 1 \mu\text{m}^2/\text{s}$ and $N_{\text{GRP}} = 0$. Each figure displays the time dependence of numerous trajectories. Trajectories in red start with all Ras inactive and trajectories in blue start with all Ras active. Twenty five trajectories from each initial condition are displayed for each value of l . Increasing l decreases the relative stability of the active Ras state.

Figure 7. Decreasing the diffusion coefficient stabilizes the active Ras state relative to the inactive Ras state

Here the system is 2-dimensional ($l = 0.01 \mu\text{m}$) with no RasGRP. The value of the response time (τ_{resp}) reported in each figure gives the average time for trajectories starting from the all-RasGDP state (red) to first reach a state in which N_{RasGTP} is half the value of its steady state average. An intermediate value of D minimizes the average time to transition from all RasGDP to the long-time value of RasGTP.

Figure 8. Bistability diagrams for the Ras-SOS reaction network

The active state is promoted by increased confinement, decreased mobility, and increased numbers of N_{GRP} .