Membrane Clustering and the Role of Rebinding in Biochemical Signaling

Andrew Mugler,[†] Aimee Gotway Bailey,[‡] Koichi Takahashi,[§] and Pieter Rein ten Wolde[†*]
[†]FOM Institute AMOLF, Amsterdam, The Netherlands; [‡]American Association for the Advancement of Science, Washington, DC; and [§]RIKEN Quantitative Biology Center, Osaka, Japan

ABSTRACT In many cellular signaling pathways, key components form clusters at the cell membrane. Although much work has focused on the mechanisms behind such cluster formation, the implications for downstream signaling remain poorly understood. Here, motivated by recent experiments, we use particle-based simulation to study a covalent modification network in which the activating component is either clustered or randomly distributed on the membrane. We find that whereas clustering reduces the response of a single-modification network, it can enhance the response of a double-modification network. The reduction is a bulk effect: a cluster presents a smaller effective target to a substrate molecule in the bulk. The enhancement, on the other hand, is a local effect: a cluster promotes the rapid rebinding and second activation of singly activated substrate molecules. As such, the enhancement relies on frequent collisions on a short timescale, leading to an optimal ratio of diffusion to association that agrees with typical measured rates. We complement simulation with analytic results at both the mean-field and first-passage distribution levels. Our results emphasize the importance of spatially resolved models, showing that significant effects of spatial correlations persist even in spatially averaged quantities such as response curves.

INTRODUCTION

Although cells are often modeled as well-mixed chemical reactors, they are highly spatially heterogeneous entities. Beyond merely providing the blueprint for space-dependent processes such as division or patterning, spatial heterogeneities in cellular components are frequently exploited by biochemical networks as additional degrees of freedom in signaling computations (1). The most direct example of this is compartmentalization, in which the same chemical component initiates different phenotypic responses depending on where it is localized within the cell (2,3). In a similar way, the localization of signaling components via scaffolding proteins has effects on signal amplification that depend nontrivially on the surrounding chemical conditions (4). In fact, the colocalization of just two components can have a dramatic response on the amplification properties of an enzyme-driven reaction network (5). Even in spatially uniform systems, spatial correlations between individual molecules can have significant effects on the mean response (6).

One of the most actively studied areas in which spatial heterogeneity is emerging as a key factor is signal transduction at the cell membrane. In addition to imposing a quasi-two-dimensional geometry, the membrane plays host to a large diversity of cellular components, and interactions among these components give rise to a complex spatial organization (7). A central theme of recent work in this field is the prevalence and role of membrane clusters, groups of colocalized molecules that often participate in the detection of external signals and subsequently drive responses within the cell. Perhaps the best-known example of this process is bacterial chemotaxis, in which clusters of receptors detect external ligands, triggering messenger molecules to modulate the activity of flagellar motors (8). Recent studies have also provided evidence for clustering in eukaryotic cell membranes: data from immunoelectron microscopy (9) and single-molecule fluorescence experiments (10) suggest that Ras, a protein that has been implicated in a variety of phenotypic responses (e.g., oncogenesis), forms membrane clusters on which the efficacy of its downstream signaling critically relies. Clustering may also be connected to the partitioning of the membrane itself into spatially segmented domains (11,12), e.g., via interaction with the cytoskeleton (13) or the formation of so-called lipid rafts (14).

Although much modeling work has been done to elucidate the possible mechanisms by which clusters form (15–17), insights into the role that clustering plays in downstream signaling remain largely speculative. Therefore, our primary goal in this study was to quantitatively assess the effect of clustering on the input-output response of a canonical signaling network, using a spatially resolved model. Recognizing its ubiquity in the systems in which clustering is observed (18), we focused on a covalent modification network (often called the push-pull network) in which a substrate is alternately activated and deactivated by two antagonistic components (Fig. 1 *A*). For example, in bacterial chemotaxis, the kinase CheA and the phosphatase CheZ phosphorylate and dephosphorylate the messenger protein CheY, respectively; CheA and CheZ therefore play the roles of the two antagonistic components, and CheY plays the role of the substrate.

Moreover, focusing on a push-pull network naturally permits extension to a double-modification process (Fig. 1 *B*),

Submitted October 21, 2011, and accepted for publication February 3, 2012.

*Correspondence: tenwolde@amolf.nl

Editor: Petra Schwille.

0006-3495/12/03/1069/10 $2.00

doi: 10.1016/j.bpj.2012.02.005

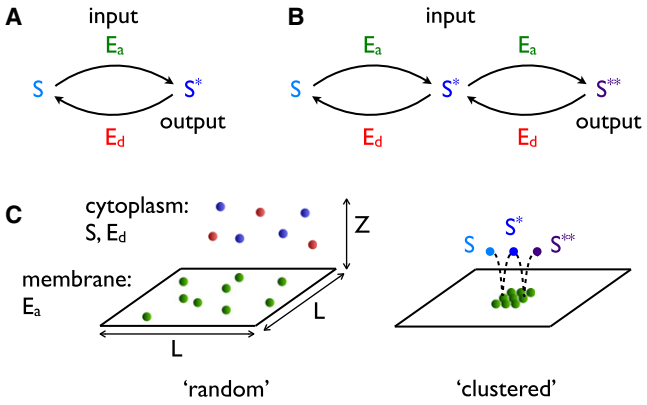

FIGURE 1 Schematics of reaction networks and spatial arrangement of molecules. (*A*) The single-modification network, in which a substrate S is activated and deactivated by components E_a and E_d, respectively. (*B*) The double-modification network, in which the substrate can be doubly activated. (*C*) Whereas S and E_d molecules diffuse freely in the cytoplasm, E_a molecules are fixed on the membrane in either a random (*left*) or clustered (*right*) configuration. The right panel depicts a rebinding event in which a singly activated S^* molecule rapidly returns to the E_a cluster to become a doubly activated S^{**} molecule.

which is a critical step in many membrane-signaling pathways. In eukaryotic cells, for example, active Ras molecules at the membrane initiate a mitogen-activated protein kinase (MAPK) cascade within the cell, each layer of which consists of a dual phosphorylation cycle. In general, dual phosphorylation can be carried out by one of two mechanisms. In a processive mechanism, an enzyme modifies both phosphorylation sites on a substrate molecule before releasing it. In a distributive mechanism, on the other hand, the enzyme must release the substrate after modifying the first site, before rebinding and modifying the second site. It has been shown experimentally that key kinases (19,20) and phosphatases (21) in the MAPK cascade act in a distributive manner, which makes the rebinding process critically important. Therefore, a second goal of this study was to investigate the interplay between clustering and rebinding, and its role in determining the input-output response of a distributive push-pull network.

We provide a spatially resolved description of the system by performing particle-based simulations on a lattice. In parallel, we gain important physical intuition from analytic results derived at both the mean-field and first-passage distribution levels. We find that the input-output response of the network changes depending on whether the activating component is clustered or randomly distributed on the membrane (Fig. 1 *C*). Specifically, whereas clustering reduces the response of a single-modification network, it can enhance the response of a double-modification network. We demonstrate that the reduction is a direct consequence of the fact that a cluster presents a smaller effective target to a substrate molecule in the bulk. By investigating in detail the stochastic nature of the rebinding process, we discover that the enhancement has an entirely different origin: clustering promotes the rapid rebinding and second activation of singly activated substrate molecules (Fig. 1 *C*). We find that such a rapid effect is only exploited when both the activating and deactivating components are sufficiently free to react, such that ultrasensitive networks (22) in which one or the other component is saturated by the substrate do not exhibit the enhancement.

Furthermore, we find that the enhancement with clustering becomes more pronounced as the diffusion coefficient is raised. Underlying this observation is the fundamental advantage that clustering affords as collisions occur more frequently: although the diffusion may be high enough to prevent a substrate molecule from rapidly rebinding an isolated enzyme molecule, it may be insufficient to enable the substrate molecule to escape an entire cluster. Clustering thus prolongs the possibility of rapid rebinding, effectively boosting the association rates of individual molecules, which are often limited by tight orientational constraints (23). Of course, this advantage reaches a limit: at infinite diffusion, all spatial arrangement is forgotten. We are thus led naturally to a ratio of diffusion to association at which the enhancement is optimal.

Together, our results provide a quantitative picture of the nontrivial effects that membrane clustering has on biochemical signaling, for a network that plays a critical role in systems in which clustering has been experimentally observed. More broadly, our results demonstrate the crucial role that spatial correlations play in cellular function, and the associated importance of considering spatial resolution in biophysical models.

METHODS

We consider both a single- and a double-modification push-pull network in which the activating enzyme is localized to the membrane (Fig. 1). To understand the effect of clustering, we compare the situation in which activating enzyme molecules are arranged randomly on the membrane with that in which they are localized at the same surface density to clusters of size N (Fig. 1 *C*). Because we are interested in the effect of clustering on downstream signaling, and not in the dynamics of cluster formation on the membrane, we take activating enzyme molecules to be fixed. Substrate and deactivating enzyme molecules diffuse freely in the cytoplasm with diffusion coefficient D.

Chemical reactions, input-output relation, and sensitivity

The single-modification network (Fig. 1 *A*) is described by the reactions

$$E_a + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} E_aS \overset{k_3}{\rightarrow} E_a + S^*, \tag{1}$$

$$E_d + S^* \underset{k_5}{\overset{k_4}{\rightleftharpoons}} E_dS^* \overset{k_6}{\rightarrow} E_d + S, \tag{2}$$

where S and S^* denote the substrate in its inactive and active forms, respectively. Activation is catalyzed by the activating enzyme E_a, which first

forms a complex before releasing the substrate in its active state, and deactivation is performed similarly by the deactivating enzyme E_d. The double-modification network (Fig. 1 *B*) prescribes additional reactions identical to Eqs. 1 and 2, except with S and S^* replaced by S^* and S^{**}, respectively. We restrict our analysis to networks whose first and second modification processes are identical (i.e., the rates $k_1, k_2, \ldots, k_6$ describing the first modification also describe the second). Furthermore, we assume negligible back reactions: $k_2 = k_5 = 0$. We tested the effects on the main results (Fig. 2) of systematically varying k_2 and k_5. The effects can be understood in terms of the insights provided in the Results section, and are discussed explicitly in the Supporting Material.

The input of the network is defined as the catalytic rate of the activating enzyme k_3, scaled by its counterpart k_6 for the deactivating enzyme: $\chi \equiv k_3/k_6$ (in chemotaxis, for example, k_3 is typically set by the time-averaged ligand occupancy of the receptor cluster (24)). The output is the relative activity of the substrate, i.e., the fraction $\phi \equiv \{[S^*]/[S]_T, [S^{**}]/[S]_T\}$ for the single- or double-modification network, respectively. In a deterministic, well-mixed description, in which rate equations determine the dynamics, the steady-state input-output relation is completely specified by the reaction rates and the conserved total concentrations of substrate $[S]_T$ and enzymes $[E_a]_T$ and $[E_d]_T$. In particular, for both the single- and double-modification networks, one may write the input-output relations entirely in terms of the dimensionless parameters (e.g., see Supporting Material):

$$\alpha \equiv \frac{[E_d]_T}{[E_a]_T},\ \beta \equiv \frac{k_4}{k_1},\ \gamma \equiv \frac{K}{[S]_T},\ \epsilon \equiv \frac{[E_a]_T}{[S]_T}, \tag{3}$$

where $K = k_6/k_4$ is the Michaelis-Menten concentration of the deactivation process, and $[E_a]_T$ is N divided by the volume. The first two parameters determine the bias of the network toward deactivation; $\alpha = \beta = 1$ therefore corresponds to a symmetric network. The last two parameters characterize the sensitivity of the network: in the zero-order (or ultrasensitive) regime, the substrate saturates the enzymes and operates far beyond the Michaelis-Menten concentration ($\{\epsilon, \gamma\} \ll 1$), whereas in the linear regime, both substrate and enzymes operate in the linear regions of their response curves ($\{\gamma^{-1}, \epsilon\,\gamma^{-1}\} \ll 1$).

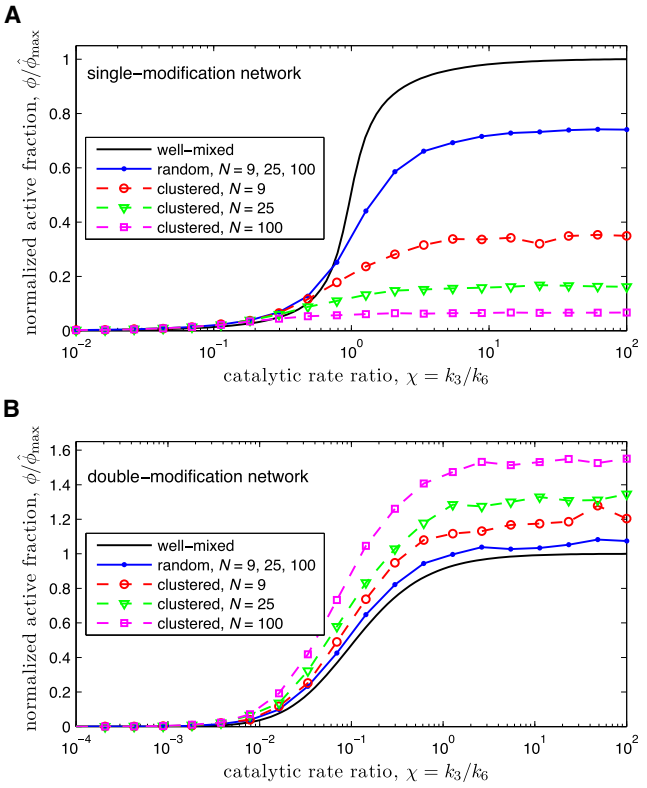

FIGURE 2 Input-output response of single- and double-modification networks. Cluster size N is varied at constant surface density $\mu = 0.01$ and depth $\zeta = 25$, leaving the response unchanged with N in the random configuration. The well-mixed curve is established according to rate equations in steady state (Supporting Material). Curves are normalized by the maximal value of the well-mixed response, $\hat{\phi}_{\max}$. (*A*) A symmetric ($\alpha = \beta = 1$) single-modification network with low diffusion ($\delta = 1/4\ \pi$) and zero-order sensitivity ($\gamma = \epsilon = 0.1$); here $\phi = [S^*]/[S]_T$. (*B*) A deactivation-biased ($\alpha = 5$, $\beta = 1$) double-modification network with intermediate diffusion ($\delta = 10$) and linear sensitivity ($\gamma^{-1} = 0.05$, $\epsilon\,\gamma^{-1} = 0.01$); here $\phi = [S^{**}]/[S]_T$. It can be seen that in the clustered configuration, the response is reduced with N for the single-modification network, and enhanced with N for the double-modification network.

Lattice model

We perform spatially resolved simulations with excluded volume interactions on a regular three-dimensional lattice. We make the approximation that all molecules have equal diameter ℓ, and we let this diameter define the lattice spacing, such that molecules neighboring each other on the lattice are in contact. Clustered molecules are placed in contact in a square arrangement on the membrane (we tested that a circular arrangement does not change the results). The membrane comprises the $x - y$ plane and extends for a length L in each direction, beyond which periodic boundaries are imposed. The cytoplasm has depth Z, with reflective boundaries at both the membrane ($z = 0$) and the farthest point from it ($z = Z$). The Supporting Material provides a detailed account of how reactions and diffusion are implemented on the lattice, in particular such that detailed balance is obeyed. All source code, written in C++ and MATLAB (The MathWorks, Natick, MA), is freely available at http://rebind.sourceforge.net.

Spatial resolution introduces new parameters into the problem beyond those of the well-mixed system (Eq. 3), which are captured by the dimensionless quantities given below. In addition to the cluster size N, one has

$$\delta \equiv \frac{\ell D}{k_1}, \quad \mu \equiv \frac{N\ell^2}{L^2}, \quad \zeta \equiv \frac{Z}{\ell}. \tag{4}$$

The quantity $1/4\pi\delta = k_1/4\pi\ell D$ is the ratio of the activating enzyme's intrinsic association rate k_1, which is the association rate given that the molecules are in contact, to the corresponding diffusion-limited value $4\pi\ell D$; as such, δ captures the strength of diffusion relative to association. The parameters μ and ζ describe the surface density of activating enzymes and the cytoplasmic depth, respectively.

Parameter selection

In the following section, we discuss in detail the effects of varying the parameters that govern network symmetry (α, β), sensitivity (γ, ϵ), and diffusion (δ). In all results that establish network characteristics (see Figs. 2, 3, and 5, and Fig. S1, Fig. S2, and Fig. S3), the surface density of activating enzymes and the cytoplasmic depth are set using estimates from experimentally studied systems. We arrive at the typical values $\mu = 0.01$ and $\zeta = 25$, as described in the Supporting Material.

RESULTS

We begin by presenting and explaining the main difference between the single- and double-modification networks:

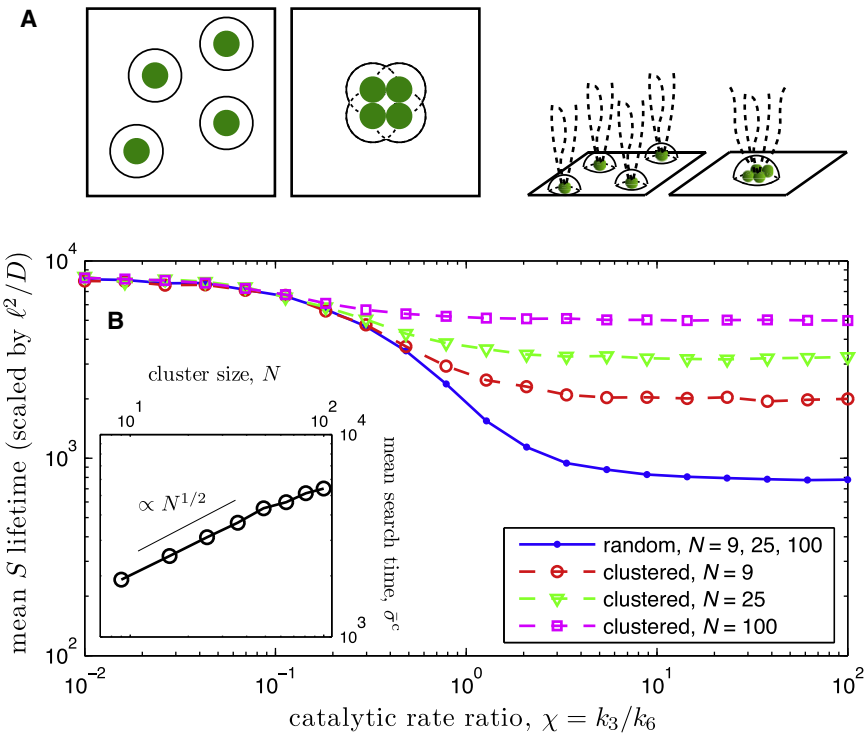

FIGURE 3 Target size effect. (*A*) Left: Cartoon depicting neighborhoods (*solid lines*) surrounding activating enzyme molecules (*solid circles*) in random and clustered configurations. The total neighborhood volume (the target size) is smaller in the clustered configuration due to the overlap of individual molecules' volumes. Right: A smaller target admits fewer distinct paths from the bulk (*dashed lines*), thereby increasing the mean time required to find the target by a diffusive search. (*B*) The main plot shows the mean lifetime of S molecules in the single-modification network; parameters are as in Fig. 2 *A*. The inset shows that the value to which the mean lifetime asymptotes at high input χ in the clustered configuration (the mean search time $\bar{\sigma}^c$) grows as the square root of the cluster size N.

whereas clustering reduces the response of a single-modification network, it can enhance the response of a double-modification network. The magnitude of each effect scales with the cluster size N (Fig. 2). The reduction for single-modification networks is generic and persists with changes in network symmetry (α, β), sensitivity (γ, ϵ), and diffusion (δ). The enhancement for double-modification networks, on the other hand, is more specific and occurs in deactivation-biased linear-sensitivity networks with intermediate diffusion. Results presented later in this section will explain this specificity.

Fig. 2 also illustrates more generally the effect of localizing the activating enzymes to the membrane by comparing the spatially averaged response with the response in the well-mixed case. As shown in Fig. 2 *A*, localization reduces the maximal response of a single-modification network (compare the well-mixed curve with the random curve). Such a reduction was seen in previous work (5) and is the result of the concentration gradients that form due to the asymmetric localization of activating and deactivating enzymes. As shown in Fig. 2 *B*, the double-modification network can avoid this reduction and can in fact achieve an amplification beyond the well-mixed response instead.

Clustering reduces the effective target size

How does clustering reduce the response of a single-modification network? The key is that a cluster presents a smaller effective target to a molecule in the bulk. The target size reduction leads to substrate molecules spending more time in the inactive state, and thus to a reduced output (Fig. 2 *A*). One can understand this target size reduction by imagining that each molecule possesses a local *neighborhood*, i.e., a volume into which another molecule can diffuse and eventually react (Fig. 3 *A*). When the E_a molecules are arranged in a random configuration at sufficiently low density, the total volume of such neighborhoods (the target size) is simply N times an individual E_a molecule's neighborhood. However, when the E_a molecules are clustered, the individual neighborhoods overlap and the target size is reduced (Fig. 3 *A*).

To understand quantitatively the impact of the target size reduction on the response of the network, we consider the time it takes an S molecule, released from the bulk, to bind an E_a molecule on the membrane (the *lifetime* of the S molecule). If the E_a molecules are free with high probability (i.e., unoccupied by other substrate molecules), the lifetime is dominated by the search time s, the time to find and bind an E_a molecule. The mean search time from the bulk can be estimated as the inverse of the association rate over the volume of the box. An excursion from the bulk is dominated by the diffusive trajectory, such that we can estimate the association rate by its diffusion-limited value: $\bar{s} \approx (4\pi\ell D/L^2 Z)^{-1} = L^2 Z/4\pi\ell D$. A random distribution of E_a molecules presents N targets of diameter ℓ, which reduces the mean search time by a factor of N: $\bar{s}^r = \bar{s}/N = L^2 Z/4\pi N\ell D$. A cluster, on the other hand, presents one target with effective diameter $\ell_{\text{eff}} \sim \sqrt{N}\ell$, making the mean search time $\bar{s}^c = L^2 Z/4\pi\,(\sqrt{N}\ell)\,D$. Scaling by the natural timescale ℓ^2/D (the time to diffuse approximately

one molecular diameter), and recalling Eq. 4, these times read $\overline{\sigma}^{\mathrm{r}} \equiv \overline{s}^{\mathrm{r}}/(\ell^2/D) = \zeta/4\pi\mu$ and $\overline{\sigma}^{\mathrm{c}} \equiv \overline{s}^{\mathrm{c}}/(\ell^2/D) = \sqrt{N}\zeta/4\pi\mu$, which makes clear that at constant surface density the search time is independent of N for the random configuration, but scales with $N^{1/2}$ for the clustered configuration.

Fig. 3 *B* shows the mean lifetime of S molecules as a function of the input χ for a single-modification network with zero-order sensitivity. At high input, the mean lifetime asymptotes to the value corresponding to the search from the bulk, consistent with the above analysis. It is clear that for the clustered configuration, this asymptotic value depends on the cluster size N, and the inset shows that it indeed scales with $N^{1/2}$, as predicted.

Two important conditions of the above analysis are that the E_a molecules are free and that the S molecule is released randomly from the bulk (and not, say, still within the neighborhood of the cluster). The first condition is met at high input ($\chi \equiv k_3/k_6 \gg 1$), when the high catalytic rate of the activating enzymes leaves the E_a molecules free with high probability. The second condition is also met at high input for networks with zero-order sensitivity, in which saturation of the deactivating enzymes leaves the E_d molecules occupied with high probability. High occupation of E_d molecules means that a typical S^* molecule has ample time to randomize its position before ultimately binding a free E_d molecule and being released as an S molecule. Interestingly, we observe that in networks with linear sensitivity, in which E_d molecules remain free even at high input (Supporting Material), the reduction in the output upon clustering persists (not shown), which leads us to conclude that the portion of S molecules that do originate in the bulk continue to contribute to a reduction in the response.

It is important to emphasize that the target size effect is a bulk effect, not a local effect, in the sense that clustering not only reduces the number of neighboring sites from which a substrate molecule can bind, it more generally reduces the number of distinct paths that lead to the target from a point in the bulk (Fig. 3 *A*). This intuition is confirmed by a simple test: under an alternative implementation, in which a substrate molecule can only bind an E_a molecule from the neighboring lattice site perpendicular to the membrane, we observe an increase in the search time with cluster size N that is only slightly less pronounced than that in the inset of Fig. 3 *B* (not shown). Because this alternative implementation has the property that clustering the E_a molecules does not change the number of available neighboring sites, the increase in search time with N is strictly due to a reduction in the number of paths from which the target is accessible.

It is also important to point out that the target size effect is a generic property of diffusive random walks, and as such it is just as present for double-modification networks as it is for single-modification networks. However, as we will describe next, in a particular parameter regime the effects of rapid rebinding can overcome the target size effect, leading to an enhancement of the response rather than a reduction.

Clustering promotes rapid rebinding

How does clustering enhance the response of a double-modification network? The key is that a cluster promotes rapid rebinding of singly activated substrate molecules. Rapid rebinding results in more doubly activated substrate and thus an enhanced output (Fig. 2 *B*). Rebinding only occurs in the double-modification network. In the single-modification network, once a substrate molecule is released by an enzyme, it can only bind to an enzyme of the opposite type. To clearly understand the rapid rebinding effect in double-modification networks, we first consider the distribution of rebinding times for a reduced system: a single S^* molecule is released from one of N E_a molecules, with no E_d molecules present. It rebinds to any E_a molecule in a time r, whose dimensionless analog we define as $\rho \equiv r/(\ell^2/D)$. Previous studies considered rebinding to a planar geometry in the context of ligand-receptor binding (25–27) (even specifically for clustered chemoreceptors in bacteria (27)), and the analytic results derived below extend the results of those studies, particularly for double-modification networks.

As shown in Fig. 4, for both a random and a clustered configuration of E_a molecules, the rebinding time distribution contains three regimes. Short times (the *molecular* regime) correspond to short excursions, after which the S^* molecule rebinds to the same E_a molecule (or cluster) from which it came. Intermediate times (the *planar* regime) correspond to excursions that are sufficiently far for the S^* to see the membrane as a plane uniformly populated with E_a molecules (or, due to the periodicity, with clusters), yet not far enough to see the reflective boundary. The granularity of individual E_a molecules is thus lost, and the membrane appears as a uniform semiabsorbent plane. Long times (the *bulk* regime) correspond to long excursions, during which the S^* molecule randomizes its position completely, returning as if from the bulk. These three regimes are discussed in detail below.

Bulk regime

The bulk regime exhibits an exponential distribution because it describes the time required to find an E_a molecule given a random starting position in the bulk. An exponential distribution is expected from a well-mixed system, in which reactions obey exponential waiting time statistics. Here, however, the molecular and planar regimes emerge due entirely to spatial correlations, affecting the network even at the level of the mean response (Fig. 2 *B*). Moreover, in the bulk regime, the substrate molecule has strayed far enough from the membrane that it effectively returns from the bulk. This return time is therefore equivalent to the search time defined above for S molecules in single-modification

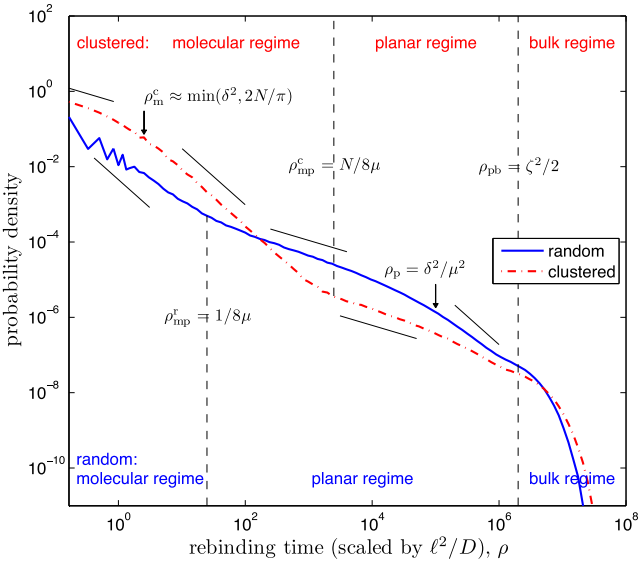

FIGURE 4 Rebinding time distributions. Distribution of times ρ for a single S^* molecule, released from an E_a molecule, to rebind to an E_a molecule when the E_a molecules are either randomly distributed (*solid*) or clustered (*dash-dot*) on the membrane. No E_d molecules are present. Short straight lines scale with $\rho^{-1/2}$ (shallow) and $\rho^{-3/2}$ (steep). To elucidate scalings, high values are taken for the dimensionless parameters describing cluster size ($N = 100$), inverse surface density ($\mu^{-1} = 200$), and cytoplasmic depth ($\zeta = 2000$); here $\delta = 1.6$. For each of the two configurations, there emerge three regimes that are distinguishable by scalings. Times separating the regimes, as well as times marking scaling crossovers within regimes, are derived in the text and indicated on the figure: ρ^{r}_{mp} and ρ^{c}_{mp} separate the molecular from the planar regime in the random and clustered configurations, respectively; ρ_{bp} separates the planar from the bulk regime in both configurations; ρ^{c}_{m} marks the scaling crossover within the molecular regime in the clustered configuration; and ρ_{p} marks the scaling crossover within the planar regime in both configurations.

networks. Accordingly, one can see in Fig. 4 that the time constant characterizing the exponential is larger in the clustered case, precisely due to the target size effect discussed above. Finally, the onset of the bulk regime is determined by the time it takes the S^* molecule to randomize its position, which is approximately the time required to diffuse the full cytoplasmic depth: $r_{pb} = Z^2/2D$, or $\rho_{pb} = \zeta^2/2$.

Planar regime

In the planar regime, the substrate molecule has not diffused far enough to enter the bulk and lose memory of its starting position, but it has diffused far enough that the membrane appears as a uniform semiabsorbent plane. The problem can be reduced to an effectively one-dimensional one in the z direction with a radiation boundary at $z = 0$. The one-dimensional rate k_{eff} (with dimensions of length per time) describing association at the boundary follows from a renormalization of the three-dimensional rate k_1. Clearly, k_{eff} should scale with the surface density N/L^2, and we find good agreement with the simplest dimensionally consistent definition, $k_{eff} \equiv k_1\, N/L^2$.

As shown in the Supporting Material, the rebinding time distribution for this one-dimensional problem is readily obtained from the Green's function and exhibits scalings of $\rho^{-1/2}$ at short times, $\rho^{-3/2}$ at long times, and a crossover time of $\rho_p = (D/\ell k_{eff})^2 = \delta^2/\mu^2$. Short times comprise a collision-dominated subregime in which the excursion is dominated by many unsuccessful reflections, and thus inherits the $t^{-1/2}$ scaling from the Gaussian Green's function of a particle freely diffusing in one dimension. Long times comprise a search-dominated subregime in which after a long excursion the particle returns to an effectively absorbing boundary, producing the $t^{-3/2}$ scaling characteristic of a one-dimensional random walker returning to an absorbing origin. Further details are provided in the Supporting Material.

The transition between the molecular and planar regimes occurs when the S^* molecule diffuses far enough perpendicular to the membrane that it no longer detects the granularity of the E_a molecules, a distance roughly equal to half the mean spacing between E_a molecules in the random configuration, or between clusters in the clustered configuration. In the random configuration, the mean spacing between E_a molecules is set by the surface density, yielding a separating time of $r^{r}_{mp} = [\sqrt{(L^2/N)}/2]^2/2D$, or $\rho^{r}_{mp} = 1/8\mu$. In the clustered configuration, the spacing between clusters is L, yielding a separating time of $r^{c}_{mp} = (L/2)^2/2D$, or $\rho^{c}_{mp} = N/8\mu$.

Molecular regime

The molecular regime is defined by short excursions in which the substrate molecule rebinds to the E_a molecule or cluster from which it came. The molecular regime exhibits $\rho^{-1/2}$ and $\rho^{-3/2}$ scalings whose origins are the same as those in the planar regime: the scalings arise from a collision-dominated or search-dominated return, respectively, to a single molecule or cluster. For a return to single molecule, which applies to the random configuration, these scalings were described in previous work (6). The crossover time was derived to be

$$r_m = \frac{\ell^2/D}{(1 + k/4\pi\ell D)^2}, \tag{5}$$

where $k = 2k_1$, with the factor of 2 arising from reflection of the E_a across the membrane. However, one can see from Fig. 4 that in the random configuration, the crossover time is obscured by alternations in the probability density at short times, which is an artifact of the lattice implementation. To be precise, an S^* molecule starting next to an E_a molecule can only rebind in an odd number of time steps (assuming it moves diffusively every time step, which is true at short times for large δ); the exception occurs when another E_a molecule is placed next to or very near the first E_a molecule, but at low surface densities such a placement occurs with low probability. We validated the distributions in Fig. 4

using Green's function reaction dynamics (6), verifying that lattice artifacts do not quantitatively change the probability densities.

In the clustered configuration, the crossover time within the molecular regime is indeed resolvable and can be described in terms of the previously considered results. A large, absorbent cluster ($N \gg 1, \delta \ll 1$) can be approximated as a plane with an effective one-dimensional association rate $k_{\mathrm{eff}} \equiv k_1/\ell^2$, yielding a dimensionless crossover time of $(D/\ell k_{\mathrm{eff}})^2 = \delta^2$. In the opposite limit, a small, reflective cluster ($N \sim 1, \delta \gg 1$) can be approximated as a spherical object whose effective diameter is obtained by equating surface areas: $4\pi\,(\ell_{\mathrm{eff}}/2)^2 = 2N\ell^2$ (neglecting cluster edges). In the limit of large δ, the denominator in Eq. 5 approaches unity, making the crossover time approximately ℓ_{eff}^2/D, or $2\,N/\pi$ in dimensionless units. Because the expressions in both the plane and sphere limits scale with parameters that are large in the opposite limits, we use the minimum as an estimate of the crossover time: $\rho_{\mathrm{m}}^{\mathrm{c}} \approx \min(\delta^2, 2N/\pi)$.

Fig. 4 corroborates all scalings and crossover times derived above using an illustrative set of sample parameters. Because we have analytic estimates for the crossover times, they can be tuned to expand or contract the various regimes, a fact we used to confirm the validity of the scalings beyond the confidence implied by Fig. 4 alone.

Fig. 4 also directly displays the advantage that clustering affords in the rebinding problem: at short times, the probability of rebinding is enhanced, leading to a probability gap over the random configuration. In fact, the characteristic time that determines the extent of this gap, $\rho_{\mathrm{m}}^{\mathrm{c}}$, reveals the parameter regimes that give rise to enhanced rebinding, and thus ultimately to an enhanced signal output for the network. Specifically, the gap increases as $\rho_{\mathrm{m}}^{\mathrm{c}}$ increases, by either increasing the cluster size N or increasing diffusion relative to association, δ. Increasing the cluster size is a straightforward way of enhancing rebinding, and the associated enhancement of the output is demonstrated in Fig. 2 *B*.

The reason that increasing diffusion increases the probability gap is perhaps less straightforward but can be understood at the molecular level. Increasing diffusion induces more unsuccessful collisions before rebinding eventually occurs. Rapid rebinding to a single E_a molecule (which is the task when the E_a molecules are randomly distributed) therefore becomes unlikely. Rapid rebinding to a cluster, on the other hand, remains less unlikely, owing to the presence of neighbors. The number of collisions in the neighborhood of a cluster is simply larger than that for a single molecule by virtue of the former's increased size. The probability of ultimately achieving a successful collision is thus higher for the clustered configuration than for the random configuration, by a factor that increases as diffusion increases.

At a more detailed mechanistic level, we may consider the fate of an S^* molecule that has just been released by an E_a molecule, and now resides at a neighboring lattice point. As diffusion increases, the probability increases that the S^* molecule will take a step away from the E_a molecule. In the random configuration, it is then increasingly likely for diffusion to carry the S^* molecule away from the immediate vicinity of the E_a molecule. In the clustered configuration, on the other hand, several of these diffusive paths will lead directly to another E_a molecule. Clustering therefore poses an advantage when immediate rebinding is unlikely but rebinding after several diffusive steps is more probable.

Deactivation connects rebinding to the network response

Interestingly, despite the probability gap elucidated above, we observe that the means of the two rebinding distributions in Fig. 4 are the same: the enhancement conferred to the clustered configuration in the molecular regime is compensated for by the target size effect in the bulk regime. The equivalence of means is a consequence of the fact that we isolated the rebinding process. Alone, the rebinding process is equivalent to one dissociation and subsequent association event of the equilibrium reaction $A + B \rightleftharpoons C$. Because it is an equilibrium reaction, detailed balance implies that the occupancy (i.e., the fraction of time in which an A molecule is bound to a B molecule) is independent of the spatial configuration of A molecules (excluded volume effects can be safely neglected given the low densities). Because we know that the mean time that an A molecule is bound to a B molecule (the mean bound time) is the inverse of the dissociation rate and thus independent of the spatial configuration of A molecules, the mean unbound time must also be independent of the spatial configuration of A molecules. Therefore, although the push-pull network as a whole prescribes a nonequilibrium process, the rebinding process alone is effectively in equilibrium, and the mean rebinding time is the same for a random and a clustered configuration.

Why then does the probability gap translate to an enhancement with clustering at the level of the mean response, as in Fig. 2 *B*? Indeed, it is precisely because thus far in the discussion we have not reintroduced the E_d molecules. The effect of the E_d molecules is to bind and deactivate the S^* molecules with the longest excursion times, removing them from the rebinding problem and thereby truncating the rebinding distributions beyond a characteristic timescale, which we call the *capture* time. The truncation alleviates the target size effect, imparting the clustered configuration with a shorter mean rebinding time than the random configuration. The capture effect is discussed in more detail in the Supporting Material and illustrated in Fig. S1.

The capture effect also underlies the observation that only double-modification networks with linear sensitivity, not zero-order sensitivity, benefit from clustering, as demonstrated in Fig. S2. The reason is that the capture effect relies

on the E_d molecules being free. The zero-order regime, on the other hand, corresponds to saturation of the enzymes by the substrate, such that at high input the E_d molecules are not free and instead are occupied with high probability. The relationship between the capture effect and network sensitivity is also discussed in more detail in the Supporting Material.

Clustering leads to an optimal ratio of diffusion to association

In studying the rebinding distributions, we discovered that increasing diffusion relative to association enhances rapid rebinding to a cluster more strongly than to a random configuration, because it increases the probability of unsuccessful collisions. Therefore, one would expect the enhancement of the output with clustering to increase with the ratio of diffusion to association, δ. However, we also know that at high diffusion, the network is well mixed and the spatial arrangement of the molecules is irrelevant; clustering should therefore confer no enhancement at high δ. In fact, these competing effects lead to a value of δ at which the enhancement is optimal, as shown in Fig. 5.

Fig. 5 illustrates that as δ increases, the enhancement (i.e., the difference in maximal output between the clustered and the random configuration $\Delta\,\phi_{\max} \equiv \phi^{c}_{\max} - \phi^{r}_{\max}$) first increases and then decreases. The inset shows this nonmonotonic behavior for several values of the cluster size N. The optimal enhancement increases with N; moreover, the value δ^* at which the optimum occurs also increases with N. These observations are consistent with the notion that a larger cluster more effectively confers the advantage associated with frequent unsuccessful collisions.

Quantitatively, δ^* approaches ~10 for the largest cluster size ($N = 100$), corresponding to an association rate $k_1 \sim \ell D/10$, roughly $10 \cdot 4\pi \approx 125$ times less than the diffusion-limited value. In fact, this is precisely the regime in which many protein–protein association reactions operate within cells. Most measured association rates are on the order of $10^6 - 10^7 \mathrm{M}^{-1}\mathrm{s}^{-1}$ (23), whereas the diffusion-limited rate is typically $10^8 - 10^9 \mathrm{M}^{-1}\mathrm{s}^{-1}$. Our results suggest that for these reactions, clustering can optimally enhance the output. Therefore, we expect the mechanism identified here to have quite significant biological relevance.

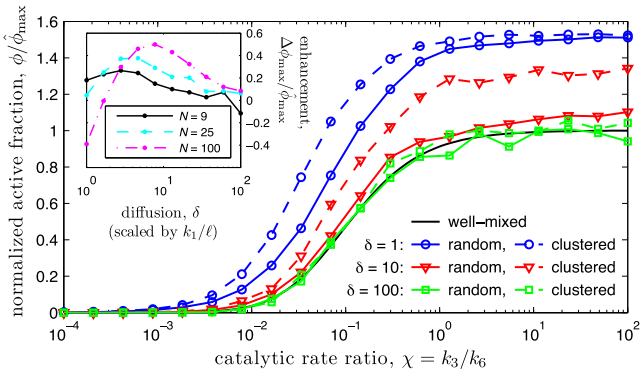

FIGURE 5 Effect of varying diffusion on the response of a double-modification network. The network is linear ($\gamma^{-1} = 0.05$, $\epsilon\,\gamma^{-1} = 0.01$) and deactivation-biased ($\alpha = 5$, $\beta = 1$) with $N = 25$, $\mu = 0.01$, and $\zeta = 25$. The inset shows the normalized difference between the maximal output in the clustered and random configurations (the enhancement) versus δ, exhibiting an optimum for each of several values of the cluster size N.

DISCUSSION

We have provided a detailed view of the varied effects that membrane clustering can have on the signaling properties of a canonical biochemical network. The network under study and the values of relevant biophysical parameters were drawn from experimentally studied systems, both prokaryotic and eukaryotic, in which membrane clustering was recently observed. We implemented a spatially resolved model, appealing to both simulation and analytic results to demonstrate that spatial correlations can have nontrivial effects, even at the level of the mean input-output response. In particular, we have shown that spatial effects at both the bulk scale (in terms of a diffusive target search process) and the molecular scale (in terms of rapid stochastic rebinding events) affect the response of a network in ways that are not captured by a well-mixed, spatially uniform description.

Our results make it clear that the effect of clustering depends on both the network topology and the biochemical parameters. For example, we identify a general property of diffusive random walks, i.e., that clustering the target increases the search time from the bulk, which leads generically to a reduced response in a single-modification network. However, when the topology of the network is extended to double-modification, the reduction can be overcome by a local effect: clustering promotes rapid rebinding of singly activated substrate molecules to the activating enzyme molecules. When the concentration of free deactivating enzyme molecules is sufficiently high to isolate these rapid rebinding events, the result is an enhancement of the response. Importantly, this enhancement is specific to networks with linear sensitivity. Ultrasensitive networks, in which the deactivating enzyme molecules are saturated by the substrate at high input, do not confer the enhancement because the mechanism relies on the deactivating enzyme molecules being free. Moreover, the enhancement is most pronounced in the presence of unsuccessful collisions, when the probability to escape a single activating enzyme molecule is much higher than that to escape a cluster. The specificity of the enhancement to both linear sensitivity and a certain ratio of diffusion to association highlights the importance of biochemical parameters in predicting the effects of clustering.

The result that clustering is most beneficial in the presence of unsuccessful collisions has important functional

implications. We find that the diffusion coefficient at which the cluster-induced enhancement is optimal corresponds to an intrinsic association rate roughly 100 times smaller than its diffusion-limited value. Such a result is quantitatively consistent with intracellular conditions because typical protein–protein association rates are measured to be roughly two orders of magnitude lower than the diffusion limit (23). Indeed, intrinsic association rates are inherently constrained by the tight orientational precision required to achieve a successful binding event, because binding requires the alignment of small reactive patches (23,28). This fact naturally induces unsuccessful collisions. Clustering may therefore have evolved as a functional way around such tight molecular geometric constraints, allowing a cell to boost signal output despite the low intrinsic association rates of individual molecules.

The effect of molecular orientation on intrinsic association rates is a subtle but important issue, especially with regard to rapid rebinding events. During very rapid rebinding between a pair of molecules, one might expect rotational alignment to be preserved, leading to an increase in the effective intrinsic association rate. On the other hand, if the molecules diffuse only a molecular diameter away from each other, then orientational diffusion will already significantly reduce the effective intrinsic association rate. Moreover, many enzymes are only capable of rebinding a substrate molecule after a characteristic reactivation time because, for example, ADP/ATP exchange must take place before rebinding can occur (6). In this study we take a coarse-grained view, neglecting these opposing effects and simply assuming that orientation is randomized immediately upon dissociation, even for very rapid rebinding events. Although this assumption may not be accurate for rapid rebinding to the same particle, it is important to emphasize that the benefit provided by clustering relies on rebinding to neighboring particles, for which the assumption of random orientation is more reasonable. Nonetheless, the relationship between orientation and rebinding rates would surely benefit from further detailed study.

Finally, our findings emphasize the general importance of the role of the rebinding process in biochemical signaling. The importance of rebinding in related systems has been discussed, with interesting consequences for the mean response. For example, in studying how the diffusive motion of a repressor protein effects gene expression, investigators observed that rebinding boosts noise, leading to bursts in gene expression, and that this effect can be captured by renormalizing the on- and off-rates in a well-mixed model (29). On the other hand, simulations (6) and experiments (30) on signaling via a MAPK cascade revealed that spatial correlations due to rapid rebinding introduce qualitative changes in the mean response that cannot be captured by the well-mixed theory. Our results here are more resonant with the second case, because it is clear that membrane localization and subsequent clustering introduce changes to the rebinding statistics (Fig. 4) that go beyond the exponential distributions expected from a well-mixed description. More broadly, the importance of rebinding in explaining the potency of T-cell ligand binding, in which a long aggregated binding time arises from a sequence of many fast rebinding events, has been recognized (31).

This study represents a first step in using simulation and analytic techniques to understand the role of spatial organization in signaling. It is our view that spatially resolved models, as well as a sharp theoretical framework, can help formalize and make more quantitative the inferences that are being made from the wealth of experimental data on systems that exhibit clustering, colocalization, and other nontrivial spatial heterogeneity.

SUPPORTING MATERIAL

Five sections, including three figures, plus references (32–38) are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(12)00205-6.

We thank Nils Becker for useful discussions, Christopher Govern for a critical reading of the manuscript, and Martijn Wehrens for validation with Green's function reaction dynamics.

This work is part of the research program of the Stichting voor Fundamenteel Onderzoek der Materie, which is financially supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek.

REFERENCES

1. Kinkhabwala, A., and P. I. H. Bastiaens. 2010. Spatial aspects of intracellular information processing. *Curr. Opin. Genet. Dev.* 20:31–40.
2. Onken, B., H. Wiener, …, E. C. Chang. 2006. Compartmentalized signaling of Ras in fission yeast. *Proc. Natl. Acad. Sci. USA.* 103:9045–9050.
3. Matallanas, D., V. Sanz-Moreno, …, P. Crespo. 2006. Distinct utilization of effectors and biological outcomes resulting from site-specific Ras activation: Ras functions in lipid rafts and Golgi complex are dispensable for proliferation and transformation. *Mol. Cell. Biol.* 26:100–116.
4. Locasale, J. W., A. S. Shaw, and A. K. Chakraborty. 2007. Scaffold proteins confer diverse regulatory properties to protein kinase cascades. *Proc. Natl. Acad. Sci. USA.* 104:13307–13312.
5. van Albada, S. B., and P. R. ten Wolde. 2007. Enzyme localization can drastically affect signal amplification in signal transduction pathways. *PLOS Comput. Biol.* 3:1925–1934.
6. Takahashi, K., S. Tanase-Nicola, and P. R. ten Wolde. 2010. Spatio-temporal correlations can drastically change the response of a MAPK pathway. *Proc. Natl. Acad. Sci. USA.* 107:2473–2478.
7. Radhakrishnan, K., A. Halász, …, J. S. Edwards. 2010. Quantitative understanding of cell signaling: the importance of membrane organization. *Curr. Opin. Biotechnol.* 21:677–682.
8. Sourjik, V., and J. P. Armitage. 2010. Spatial organization in bacterial chemotaxis. *EMBO J.* 29:2724–2733.
9. Plowman, S. J., C. Muncke, …, J. F. Hancock. 2005. H-ras, K-ras, and inner plasma membrane raft proteins operate in nanoclusters with differential dependence on the actin cytoskeleton. *Proc. Natl. Acad. Sci. USA.* 102:15500–15505.
10. Murakoshi, H., R. Iino, …, A. Kusumi. 2004. Single-molecule imaging analysis of Ras activation in living cells. *Proc. Natl. Acad. Sci. USA.* 101:7317–7322.

11. Goswami, D., K. Gowrishankar, …, S. Mayor. 2008. Nanoclusters of GPI-anchored proteins are formed by cortical actin-driven activity. *Cell.* 135:1085–1097.

12. van Zanten, T. S., A. Cambi, …, M. F. Garcia-Parajo. 2009. Hotspots of GPI-anchored proteins and integrin nanoclusters function as nucleation sites for cell adhesion. *Proc. Natl. Acad. Sci. USA.* 106:18557–18562.

13. Machta, B. B., S. Papanikolaou, …, S. L. Veatch. 2011. Minimal model of plasma membrane heterogeneity requires coupling cortical actin to criticality. *Biophys. J.* 100:1668–1677.

14. Simons, K., and M. J. Gerl. 2010. Revitalizing membrane rafts: new tools and insights. *Nat. Rev. Mol. Cell Biol.* 11:688–699.

15. Das, J., M. Kardar, and A. K. Chakraborty. 2009. Positive feedback regulation results in spatial clustering and fast spreading of active signaling molecules on a cell membrane. *J. Chem. Phys.* 130:245102.

16. Tian, T., S. J. Plowman, …, J. F. Hancock. 2010. Mathematical modeling of K-Ras nanocluster formation on the plasma membrane. *Biophys. J.* 99:534–543.

17. Gurry, T., O. Kahramanoğullari, and R. G. Endres. 2009. Biophysical mechanism for ras-nanocluster formation and signaling in plasma membrane. *PLoS ONE.* 4:e6148.

18. Sauro, H. M., and B. N. Kholodenko. 2004. Quantitative analysis of signaling networks. *Prog. Biophys. Mol. Biol.* 86:5–43.

19. Burack, W. R., and T. W. Sturgill. 1997. The activating dual phosphorylation of MAPK by MEK is nonprocessive. *Biochemistry.* 36:5929–5933.

20. Ferrell, Jr., J. E., and R. R. Bhatt. 1997. Mechanistic studies of the dual phosphorylation of mitogen-activated protein kinase. *J. Biol. Chem.* 272:19008–19016.

21. Zhao, Y., and Z. Y. Zhang. 2001. The mechanism of dephosphorylation of extracellular signal-regulated kinase 2 by mitogen-activated protein kinase phosphatase 3. *J. Biol. Chem.* 276:32382–32391.

22. Goldbeter, A., and D. E. Koshland, Jr. 1981. An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci. USA.* 78:6840–6844.

23. Northrup, S. H., and H. P. Erickson. 1992. Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc. Natl. Acad. Sci. USA.* 89:3338–3342.

24. Sourjik, V., and H. C. Berg. 2002. Binding of the *Escherichia coli* response regulator CheY to its target measured in vivo by fluorescence resonance energy transfer. *Proc. Natl. Acad. Sci. USA.* 99:12669–12674.

25. Gopalakrishnan, M., K. Forsten-Williams, …, U. C. Täuber. 2005. Ligand rebinding: self-consistent mean-field theory and numerical simulations applied to surface plasmon resonance studies. *Eur. Biophys. J.* 34:943–958.

26. Lagerholm, B., and N. Thompson. 2000. Temporal dependence of ligand dissociation and rebinding at planar surfaces. *J. Phys. Chem. B.* 104:863–868.

27. Andrews, S. S. 2005. Serial rebinding of ligands to clustered receptors as exemplified by bacterial chemotaxis. *Phys. Biol.* 2:111–122.

28. Camacho, C. J., S. R. Kimura, …, S. Vajda. 2000. Kinetics of desolvation-mediated protein-protein binding. *Biophys. J.* 78:1094–1105.

29. van Zon, J. S., M. J. Morelli, …, P. R. ten Wolde. 2006. Diffusion of transcription factors can drastically enhance the noise in gene expression. *Biophys. J.* 91:4350–4367.

30. Aoki, K., M. Yamada, …, M. Matsuda. 2011. Processive phosphorylation of ERK MAP kinase in mammalian cells. *Proc. Natl. Acad. Sci. USA.* 108:12675–12680.

31. Govern, C. C., M. K. Paczosa, …, E. S. Huseby. 2010. Fast on-rates allow short dwell time ligands to activate T cells. *Proc. Natl. Acad. Sci. USA.* 107:8724–8729.

32. Morelli, M. J., and P. R. ten Wolde. 2008. Reaction Brownian dynamics and the effect of spatial fluctuations on the gain of a push-pull network. *J. Chem. Phys.* 129:054112.

33. Tian, T., A. Harding, …, J. F. Hancock. 2007. Plasma membrane nanoswitches generate high-fidelity Ras signal transduction. *Nat. Cell Biol.* 9:905–914.

34. Li, M., and G. L. Hazelbauer. 2004. Cellular stoichiometry of the components of the chemotaxis signaling complex. *J. Bacteriol.* 186:3687–3694.

35. Phillips, R., J. Kondev, and J. Theriot. 2009. Physical Biology of the Cell. Garland Science, New York.

36. Alberts, B., D. Bray, …, J. D. Watson. 1994. Molecular Biology of the Cell. Garland Publishing, New York.

37. Carslaw, H. S., and J. C. Jaeger. 1959. Conduction of Heat in Solids, 2nd ed. Clarendon Press, Oxford.

38. Beck, J. V., K. D. Cole, …, B. Litkouhi. 1992. Heat Conduction Using Green's Functions. Hemisphere Publishing, Philadelphia.

Supporting material for "Membrane clustering and the role of rebinding in biochemical signaling"

Andrew Mugler[1], Aimee Gotway Bailey[2], Koichi Takahashi[3], Pieter Rein ten Wolde[1,*]

[1]FOM Institute AMOLF, Science Park 104, 1098 XG Amsterdam, The Netherlands
[2]American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005, USA
[3]RIKEN Quantitative Biology Center, 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan

This document contains five sections. The first section supplements the methods of the main text, describing the implementation of reactions and diffusion on the lattice, and the estimation of parameters from experimentally studied systems. The second section presents analytic results for the single- and double-modification networks in the deterministic, well-mixed limit. The third section presents analytic results for the rebinding process in terms of first-passage distributions. The fourth section describes the influence of the deactivating enzymes on the rebinding process and the associated benefit of clustering in networks with linear sensitivity; this section contains Figs. S1 and S2. The fifth section discusses the results of systematically varying the back-reaction rates k_2 and k_5, which are assumed to be zero in the main text; this section contains Fig. S3.

1 Supplementary methods

In this section, we provide two supplements to the methods described in the main text: the implementation of reactions and diffusion on the lattice, and the estimation of certain parameters from experimentally studied systems.

1.1 Reaction-diffusion implementation on the lattice

Here we describe how reactions and diffusion are implemented for particles on the lattice. In particular, the implementation obeys detailed balance and ensures that the deterministic results (next section) are recovered by spatially averaged quantities in the high-diffusion limit.

We remind the reader that we consider a regular three-dimensional lattice with excluded volume interactions. We make the approximation that all molecules have equal diameter ℓ, and we let this diameter define the lattice spacing, such that molecules neighboring each other on the lattice are in contact. In the clustered configuration, N molecules are placed in contact in a square arrangement on the membrane; in the random configuration, N molecules are arranged randomly on the membrane. We have tested that the clustered results are not significantly affected by instead placing molecules in an arrangement that is circular (up to the lattice resolution). The membrane comprises the x-y plane and extends for a length L in each direction, beyond which periodic boundaries are imposed. The cytoplasm has depth Z, with reflective boundaries at both the membrane ($z = 0$) and the farthest point from it ($z = Z$). Reflection at $z = 0$ implies that cytoplasmic molecules do not bind directly to the membrane, but rather only bind to the cytoplasmic domain of membrane-bound molecules.

Particle numbers used in the simulation box of volume L^2Z are expressible in terms of the dimensionless parameters in Eqns. 3-4 of the main text: the number of activating enzyme molecules is N, the number of deactivating enzyme molecules is αN, and the number of substrate molecules is N/ϵ. Activating enzyme molecules are membrane-bound and fixed. Other particles are initialized randomly in the simulation box

*Correspondence: tenwolde@amolf.nl

and diffuse with coefficient D. At each time step, particles are looped over, with each particle diffusing to a neighboring site or participating in a reaction with certain probabilities as described below.

Over timescales longer than the time to diffuse a few molecular diameters, rotational diffusion sufficiently randomizes a molecule's orientation. Thus although we imagine each deactivating enzyme as possessing a catalytic domain to which the substrate binds, we model its reaction propensities as isotropically distributed over its surface. Since the activating enzymes, on the other hand, are membrane-bound, the situation is more subtle: we suppose that the cytoplasmic domain of each activating enzyme molecule traces out the hemispherical solid angle inside the membrane, except when blocked by neighboring activating enzymes (a consideration particularly relevant when the activating enzymes are clustered). Neighbors thus have the effect of reducing the molecule's cross-section: the reaction propensity of each activating enzyme molecule is distributed over the portion of its surface both inside the membrane and unblocked by neighbors.

We note that the system as described can be mapped to a statistically equivalent system with periodic boundaries in the z direction, which offers both simpler implementation and, in some cases, more direct analytic interpretation. Specifically, the reflective boundaries at $z = 0$ and Z are equivalent to a periodic boundary at $z = 0$ and $2Z$, so long as we recognize that the cytoplasmic molecules then double in number and the membrane-bound molecules (having been reflected across the membrane) become twice as reactive. The periodic boundaries confer the advantage that the membrane no longer needs to be explicitly implemented in the simulation: cytoplasmic molecules can occupy the plane in which the activating enzyme molecules reside, and substrate molecules can bind to activating enzyme molecules from any free neighboring site.

We describe the implementation of reactions and diffusion on the lattice using a simple example, then extend to the push-pull reactions (Eqns. 1-2 of the main text). We consider the binary reversible reaction $A + B \rightleftharpoons C$, in which association and dissociation are described by intrinsic rates k_a (with dimensions of length cubed per time) and k_d (with dimensions of inverse time), respectively. Dissociation is modeled as a first-order reaction event with an exponential waiting time distribution; the probability for a C molecule to dissociate in a time step dt is thus $p_d \approx k_d dt$ for small p_d. Association is set by detailed balance, which equates the ratio of microscopic probabilities to enter and leave a reaction state to the ratio of macroscopic rates (1); on a lattice with spacing ℓ, the detailed balance condition reads $p_a \ell^3 / p_d = k_a / k_d$, yielding $p_a = (k_a/\ell^3)dt$ for the association probability of an A and B molecule at contact. Finally, diffusion is implemented according to its microscopic definition, namely that the mean squared distance traveled in a time dt is $6Ddt$; the six possible moves on the lattice result in a mean squared distance of $6p_D\ell^2$ in one time step, making $p_D = (D/\ell^2)dt$ the probability for a molecule to diffuse to a neighboring site.

We now extend the above expressions to the push-pull reactions and write them in terms of the dimensionless parameters (Eqns. 3-4 of the main text), the cluster size N, the input $\chi \equiv k_3/k_6$, and the dimensionless time $\tau \equiv t/(\ell^2/D)$. The probability to diffuse to a neighboring site in a time step dt is $p_D = (D/\ell^2)dt = d\tau$. There are two association reactions, with probabilities of occurring from contact $p_1 = (k_1/\ell^3)dt = (1/\delta)d\tau$ (activation) and $p_4 = (k_4/\ell^3)dt = (\beta/\delta)d\tau$ (deactivation). Lastly, there are two dissociation reactions, with probabilities of occuring $p_3 = k_3 dt = (\chi\beta\gamma\mu/\delta\epsilon\zeta)d\tau$ (activation) and $p_6 = k_6 dt = (\beta\gamma\mu/\delta\epsilon\zeta)d\tau$ (deactivation), where in addition to using the definitions of the dimensionless parameters, we have recognized explicitly that $[E_a]_T = N/L^2Z$. The time step $d\tau$ is chosen small enough that the sum of each molecule's diffusion and reaction probabilities is bounded from above by one at all times.

Both association and dissociation probabilities are divided uniformly over the faces of each molecule (or, in the case of membrane-bound molecules, the faces not blocked by fixed neighbors). Furthermore, the choice of which molecule actually moves during a dissociation event is determined by diffusion: in the binary reaction, A would move with probability $D_A/(D_A+D_B)$ and B would move with probability $D_B/(D_A+D_B)$. In practice, then, when a substrate molecule dissociates from an activating enzyme molecule, the substrate molecule always moves, because the activating enzymes are fixed. When a substrate molecule dissociates from a deactivating enzyme molecule, on the other hand, each molecule moves with probability 1/2, because the diffusion coefficients are equal. These choices ensure that the total probability of associating or dissociating in each time step sums to p_a or p_d, respectively, and therefore that detailed balance is obeyed.

1.2 Parameter estimation

In all results establishing network characteristics (Figs. 2, 3, and 5 of the main text and S1-3 here), the surface density of activating enzymes and the cytoplasmic depth are set using estimates from experimentally studied

systems, as described here. In eukaryotic cells, clustered Ras has been measured to occupy a membrane surface fraction of $\mu \lesssim 1\%$ (2). A similar value arises in bacterial chemotaxis: the 'long' form of CheA (the form which both associates to receptors and phosphorylates CheY) is present in roughly 4500 copies per *Escherichia coli* cell (3); taking a cell surface area of 6 μm^2 and a typical protein diameter of 4 nm (4), one obtains $\mu \sim 0.012$. We therefore take $\mu = 0.01$. The cytoplasmic depth is a measure of the maximum distance from the membrane that a molecule diffuses before encountering a reflective barrier (or, at the most, half the distance to the opposite membrane). In bacteria, this distance is upper bounded by half the smallest cell lengthscale, or ∼500 nm. In eukaryotic cells, this depth is instead dictated by the presence of large organelles near the membrane. An extreme upper bound can be obtained by noting that organelles comprise roughly half the cell volume (5), and imagining they are spherically packed in the center of a spherical cell of radius $R \sim 5$ μm implies a maximum depth of $Z = [1 - (1/2)^{1/3}]R \sim 1000$ nm. Organelles are, of course, more loosely distributed within the cell, such that a depth on the order of 100 nm might be more realistic. We therefore take $Z = 100$ nm, which for a molecular diameter of 4 nm gives $\zeta = 25$.

2 Analytic results in the deterministic, well-mixed limit

In this section, we derive key analytic results for both the single- and double-modification push-pull network in the deterministic, well-mixed limit (i.e. invoking rate equations). Strictly speaking, these results are exact for averaged quantities in the limit of infinite diffusion. More broadly, however, they lend powerful intuition to the spatially resolved results, even when diffusion plays a significant role.

2.1 The single-modification network

We begin with the single-modification network (Eqns. 1-2 of the main text), which in steady state is described by the rate equations

$$\begin{aligned} 0 &= \frac{d[S]}{dt} = -k_1[E_a][S] + k_6[E_d S^*], && (1)\\ 0 &= \frac{d[S^*]}{dt} = -k_4[E_d][S^*] + k_3[E_a S], && (2)\\ 0 &= \frac{d[E_a]}{dt} = -\frac{d[E_a S]}{dt} = -k_1[E_a][S] + k_3[E_a S], && (3)\\ 0 &= \frac{d[E_d]}{dt} = -\frac{d[E_d S^*]}{dt} = -k_4[E_d][S^*] + k_6[E_d S^*]. && (4) \end{aligned}$$

Here, as in the main text, we neglect back reactions ($k_2 = k_5 = 0$). The rate equations are complemented by the conservation laws

$$\begin{aligned} [E_a]_T &= [E_a] + [E_a S], && (5)\\ [E_d]_T &= [E_d] + [E_d S^*], && (6)\\ [S]_T &= [S] + [S^*] + [E_a S] + [E_d S^*]. && (7) \end{aligned}$$

As implied by the conservation laws, the rate equations contain three redundancies from the zero net flux of activating enzyme, deactivating enzyme, and substrate; two are made explicit in Eqns. 3-4, and the third is revealed by the fact that the sum of Eqns. 1-2 equals the sum of Eqns. 3-4. Eqns. 1-7 thus constitute six independent equations for six unknowns. Scaling concentrations by the Michaelis-Menten concentration of the deactivation process, $K \equiv k_6/k_4$,

$$x_1 \equiv \frac{[E_a]}{K}, \qquad x_2 \equiv \frac{[S]}{K}, \qquad x_3 \equiv \frac{[E_d]}{K}, \qquad x_4 \equiv \frac{[E_a S]}{K}, \qquad x_5 \equiv \frac{[E_d S^*]}{K}, \tag{8}$$

and recalling the definitions of the dimensionless parameters introduced in Eqn. 3 of the main text,

$$\alpha \equiv \frac{[E_d]_T}{[E_a]_T}, \qquad \beta \equiv \frac{k_4}{k_1}, \qquad \gamma \equiv \frac{K}{[S]_T}, \qquad \epsilon \equiv \frac{[E_a]_T}{[S]_T}, \tag{9}$$

the six independent equations may be written

$$\begin{aligned} x_1 x_2 &= \beta x_5, && (10)\\ \chi^{-1} x_3 \phi &= \gamma x_4, && (11)\\ x_3 \phi &= \gamma x_5, && (12)\\ \epsilon &= \gamma(x_1 + x_4), && (13)\\ \alpha\epsilon &= \gamma(x_3 + x_5), && (14)\\ 1 &= \phi + \gamma(x_2 + x_4 + x_5), && (15) \end{aligned}$$

where $\chi \equiv k_3/k_6$ and $\phi \equiv [S^*]/[S]_T$ are the input and output, respectively.

Combining Eqns. 10-15 yields a third-degree polynomial equation for ϕ (6):

$$\begin{aligned} 0 = {} & \chi(\chi-\alpha)\phi^3 + [\alpha(1+\chi)(\chi-\alpha)\epsilon + \chi(2\chi+\alpha\beta\chi-\alpha)\gamma - \chi(\chi-\alpha)]\,\phi^2 \\ & + \chi\,[\alpha(1+\chi)\epsilon + \chi(1+\alpha\beta)\gamma - (2\chi-\alpha)]\,\gamma\phi - \chi^2\gamma^2. \end{aligned} \quad (16)$$

In principle, Eqn. 16 can be solved for the input-output relation $\phi(\chi)$. However, the solution to such a cubic equation is quite unwieldy, and we therefore focus on limits of Eqn. 16 (or the original Eqns. 10-15) at high input ($\chi \gg 1$), and in the zero-order ($\{\epsilon, \gamma\} \ll 1$) and linear regimes ($\{\eta, \nu\} \ll 1$). Here, for notational convenience, we have defined

$$\eta \equiv \frac{1}{\gamma} = \frac{[S]_T}{K}, \qquad \nu \equiv \frac{\epsilon}{\gamma} = \frac{[E_a]_T}{K}. \quad (17)$$

We find the maximal output value, $\phi(\chi \gg 1) \equiv \phi_{\max}$, directly from Eqns. 10-15. The leading order result is obtained when $\chi^{-1} = 0$ exactly; by Eqn. 11 this leads to $x_4 = 0 \Rightarrow [E_a S] = 0$, which makes sense because at infinite catalytic rate k_3 the complex $E_a S$ has zero lifetime. Combining the five remaining equations yields a quadratic equation,

$$0 = \phi_{\max}^2 + [\alpha\epsilon + (1+\alpha\beta)\gamma - 1]\phi_{\max} - \gamma, \quad (18)$$

whose solution directly gives $\phi_{\max}$.

In the zero-order regime, to zeroth order in the small parameters ($\gamma = \epsilon = 0$), Eqn. 18 reads $0 = \phi_{\max}(\phi_{\max} - 1)$, giving the maximal output value $\phi_{\max} = 1$: it is possible to activate all substrate molecules at high input. Further insight is revealed by Eqn. 16, which for $\gamma = \epsilon = 0$ reduces to

$$0 = (\chi - \alpha)\phi^2(\phi - 1). \quad (19)$$

Here, when $\chi \neq \alpha$, ϕ must be either 0 or 1, implying switch-like behavior around the threshold $\chi^* = \alpha$. This switch-like behavior is the hallmark of zero-order sensitivity (6).

In the linear regime, we obtain $\phi_{\max}$ by rewriting Eqn. 18 in terms of η and ν:

$$0 = \eta\phi_{\max}^2 + [\alpha\nu + (1+\alpha\beta) - \eta]\phi_{\max} - 1. \quad (20)$$

To zeroth order in the small parameters ($\eta = \nu = 0$), Eqn. 20 gives

$$\phi_{\max} = \frac{1}{1+\alpha\beta}. \quad (21)$$

Interestingly, for a symmetric network ($\beta = \alpha = 1$), we see that in the linear regime it is only possible to activate half the substrate molecules at high input. We also obtain the threshold value from Eqn. 16, which in terms of η and ν reads

$$\begin{aligned} 0 = {} & \chi(\chi-\alpha)\eta^2\phi^3 + [\alpha(1+\chi)(\chi-\alpha)\nu + \chi(2\chi+\alpha\beta\chi-\alpha) - \chi(\chi-\alpha)\eta]\,\eta\phi^2 \\ & + \chi\,[\alpha(1+\chi)\nu + \chi(1+\alpha\beta) - (2\chi-\alpha)\eta]\,\phi - \chi^2. \end{aligned} \quad (22)$$

Taking Eqn. 22 to first order in η and ν yields

$$0 = \eta(2\chi+\alpha\beta\chi-\alpha)\phi^2 + [\alpha(1+\chi)\nu + \chi(1+\alpha\beta) - (2\chi-\alpha)\eta]\,\phi - \chi, \quad (23)$$

into which we insert $\phi = \phi_{\max}/2 = 1/[2(1+\alpha\beta)]$ and solve for χ, yielding the threshold value

$$\chi^* = \frac{\alpha(1+2\alpha\beta)\eta + 2\alpha(1+\alpha\beta)\nu}{2(1+\alpha\beta)^2 + (1+3\alpha\beta)\eta - 2\alpha(1+\alpha\beta)\nu} \approx \left[\frac{\alpha(1+2\alpha\beta)}{2(1+\alpha\beta)^2}\right]\eta + \left[\frac{\alpha}{1+\alpha\beta}\right]\nu. \tag{24}$$

We see that while in the zero-order regime the threshold is set by the ratio of activating to deactivating enzymes, α, in the linear regime the threshold vanishes in proportion to the small parameters that define the regime, η and ν. We find that Eqn. 24 also serves as a good estimate for the threshold in a double-modification network with linear sensitivity, and thus explains why the threshold shifts to small χ in Fig. S2 below.

2.2 The double-modification network

We now consider the double-modification network, which prescribes additional reactions identical to Eqns. 1-2 of the main text, except with S and S^* replaced by S^* and S^{**}, respectively. As in the main text, we restrict our analysis to networks whose first and second modification processes are identical (i.e. the rates $k_1, k_2, \ldots, k_6$ describing the first modification also describe the second). There are now nine species, described by the dimensionless variables in Eqn. 8, the new variables

$$x_6 \equiv \frac{[S^*]}{K}, \qquad x_7 \equiv \frac{[E_a S^*]}{K}, \qquad x_8 \equiv \frac{[E_d S^{**}]}{K}, \tag{25}$$

and the redefined output $\phi \equiv [S^{**}]/[S]_T$. As before the nine rate equations contain three redundancies from the zero net flux of activating enzyme, deactivating enzyme, and substrate; together with the three conservation laws we thus have nine independent equations for nine unknowns:

$$\begin{aligned}
x_1 x_2 &= \beta x_5, && (26)\\
\chi^{-1} x_6 x_3 &= x_4, && (27)\\
x_6 x_3 &= x_5, && (28)\\
x_1 x_6 &= \beta x_8, && (29)\\
\chi^{-1}\phi x_3 &= \gamma x_7, && (30)\\
\phi x_3 &= \gamma x_8, && (31)\\
\epsilon &= \gamma(x_1 + x_4 + x_7), && (32)\\
\alpha\epsilon &= \gamma(x_3 + x_5 + x_8), && (33)\\
1 &= \phi + \gamma(x_2 + x_4 + x_5 + x_6 + x_7 + x_8). && (34)
\end{aligned}$$

Although it is no longer straightforward to combine Eqns. 26-34 into a single polynomial in ϕ, results in certain limits can be obtained directly from the equations themselves. For example, we may immediately seek the maximal output value $\phi_{\max}$ by taking the limit $\chi \gg 1$ to zeroth order ($\chi^{-1} = 0$). By Eqns. 27 and 30 this limit leads to $x_4 = 0 \Rightarrow [E_a S] = 0$ and $x_7 = 0 \Rightarrow [E_a S^*] = 0$, respectively, which make sense because at infinite catalytic rate k_3 the complexes $E_a S$ and $E_a S^*$ have zero lifetime. Although the remaining seven equations still do not lead easily to a single equation for $\phi_{\max}$, it is possible to derive an expression for $\phi_{\max}$ directly in the linear regime.

The linear regime implies that the parameters η and ν are small (Eqn. 17); in terms of these parameters, the remaining seven equations read:

$$\begin{aligned}
x_1 x_2 &= \beta x_5, && (35)\\
x_6 x_3 &= x_5, && (36)\\
x_1 x_6 &= \beta x_8, && (37)\\
\eta\phi_{\max} x_3 &= x_8, && (38)\\
\nu &= x_1, && (39)\\
\alpha\nu &= x_3 + x_5 + x_8, && (40)\\
\eta &= \eta\phi_{\max} + x_2 + x_5 + x_6 + x_8. && (41)
\end{aligned}$$

Furthermore, since the linear regime is defined by the fact that both substrate and enzyme concentrations are much smaller than K, the dimensionless variables x_i are also small (Eqns. 8, 25). We may then recognize that Eqns. 35-36 and 37-38 imply that x_5 and x_8, respectively, are small to second order. To first order, then, $x_5 = x_8 = 0$, and

$$\begin{aligned} x_1 x_2 &= \beta x_6 x_3, & (42)\\ x_1 x_6 &= \beta\eta\phi_{\max} x_3, & (43)\\ \nu &= x_1, & (44)\\ \alpha\nu &= x_3, & (45)\\ \eta &= \eta\phi_{\max} + x_2 + x_6, & (46) \end{aligned}$$

where Eqns. 42 and 43 come from combining Eqns. 35-36 and 37-38, respectively. It is now trivial to solve Eqns. 42-46 for $\phi_{\max}$, yielding

$$\phi_{\max} = \frac{1}{1 + \alpha\beta + \alpha^2\beta^2}. \quad (47)$$

Upon comparing this expression with that for the single-modification network (Eqn. 21), we see that the maximal output for the double-modification network is suppressed by an additional term $\alpha^2\beta^2$ in the denominator. Indeed, with respect to the deactivating enzymes, if the activating enzymes are fewer ($\alpha > 1$) or associate more weakly to the substrate ($\beta > 1$), suppression of the output is severe in the linear regime.

Additionally, the result that x_5 and x_8 are small to second order implies that $[E_d S^*]$ and $[E_d S^{**}]$ are much smaller than $[E_d]$. To leading order, then, $[E_d]_T \approx [E_d]$, i.e. in the linear regime the deactivating enzymes are approximately all free, *even at maximal input*. Indeed, this feature is a primary difference between the two sensitivity regimes: in the zero-order regime either the activating or deactivating enzymes are saturated by the substrate, while in the linear regime both activating and deactivating enzymes are free. The implications of this freedom for signaling are discussed at several points in the main text.

3 Rebinding time distribution in one dimension

In this section, we consider the problem of a particle diffusing in a one-dimensional space that is free on one side and has a radiation boundary condition on the other. The distribution of first-passage times at the boundary is directly obtainable from the Green's function for this problem, which is known. The result provides a good approximation for the distribution of rebinding times to a planar membrane populated with absorbing constituents, for excursions long enough that the plane appears as a uniform, semi-absorbent sink. The results in this section are not novel, but are provided for intuition and completeness. We refer the reader to (7, 8) for comprehensive treatments of many diffusion problems, including this one, and to (9) for an intuitive discussion of the analogous problem in three dimensions.

We consider a particle diffusing along the positive z-axis with a radiation boundary at $z = 0$. The diffusion equation describes the evolution of the probability density $p(z|t)$:

$$\frac{\partial p(z|t)}{\partial t} = D\frac{\partial^2 p(z|t)}{\partial z^2}. \quad (48)$$

The radiation boundary condition states that the flux leaving the boundary is due to a reaction, which requires both that the particle is at the boundary, with probability $p(0|t)$, and that the reaction fires, with intrinsic rate k (dimensions length per time):

$$D\left.\frac{\partial p(z|t)}{\partial z}\right|_{z=0} = kp(0|t). \quad (49)$$

The solution given that the particle starts at point z_0, i.e.

$$p(z|0) = \delta(z - z_0), \quad (50)$$

which is called the Green's function, is known (7):

$$\begin{aligned} p(z|t,z_0) &= \frac{1}{\sqrt{4\pi Dt}}\left[e^{-(z-z_0)^2/(4Dt)} + e^{-(z+z_0)^2/(4Dt)}\right] \\ &\quad -\frac{k}{D}e^{k^2t/D}e^{k(z+z_0)/D}\mathrm{erfc}\left(\frac{z+z_0}{\sqrt{4Dt}} + \sqrt{\frac{k^2t}{D}}\right). \end{aligned} \tag{51}$$

The first term is the solution for a reflecting boundary, and the second term describes the loss of probability incurred by the reaction; erfc denotes the complementary error function.

The distribution of first-passage times through the boundary $P(t|z_0)$ is equal to the flux out of the boundary at time t, which by Eqn. 49 is

$$P(t|z_0) = kp(0|t,z_0). \tag{52}$$

We obtain $p(0|t,z_0)$ directly from Eqn. 51, yielding

$$P(t|z_0) = \frac{k}{\sqrt{\pi Dt}}e^{-z_0^2/(4Dt)} - \frac{k^2}{D}e^{k^2t/D}e^{kz_0/D}\mathrm{erfc}\left(\frac{z_0}{\sqrt{4Dt}} + \sqrt{\frac{k^2t}{D}}\right). \tag{53}$$

We specialize to the distribution of *rebinding* times $r \equiv t$ by demanding that the particle starts at the boundary, $z_0 = 0$. Eqn. 53 then becomes

$$P(r) = \frac{k}{\sqrt{\pi Dr}} - \frac{k^2}{D}e^{k^2r/D}\mathrm{erfc}\left(\sqrt{\frac{k^2r}{D}}\right) \tag{54}$$

$$= \frac{1}{r_\mathrm{p}}\left[\frac{1}{\sqrt{\pi r/r_\mathrm{p}}} - e^{r/r_\mathrm{p}}\mathrm{erfc}\left(\sqrt{r/r_\mathrm{p}}\right)\right]. \tag{55}$$

In the second line we have recognized that reaction and diffusion define a characteristic timescale $r_\mathrm{p} \equiv D/k^2$. The meaning of this timescale is elucidated by considering times much shorter or longer than r_p, as described below.

At short times ($r \ll r_\mathrm{p}$), the second term in Eqn. 55 is unity to leading order, and the first term dominates, producing a $r^{-1/2}$ scaling:

$$P(r) \approx \frac{1}{\sqrt{\pi r_\mathrm{p}}}r^{-1/2} \qquad r \ll r_\mathrm{p}. \tag{56}$$

Such short times correspond to a collision-dominated regime: the particle does not diffuse appreciably far from the boundary; instead, it makes quick bounces against the boundary, getting reflected until ultimately becoming absorbed.

The above intuition can be sharpened in two ways. First, we may quantify the notion of "appreciably far" by realizing that reaction and diffusion define a characteristic length $d \equiv D/k$. The speed at which a particle travels this length in time r_p is $d/r_\mathrm{p} = k$, revealing that the reaction rate k may also be interpreted as the mean velocity at which particles are "pulled" into the boundary. Particles diffusing farther than d escape this radiative pull, while particles remaining within d stay close to the boundary until absorbed.

Second, we may appeal to Bayes's rule to understand the scaling in Eqn. 56. Supposing that a short excursion is comprised of a number of unsuccessful reflections, ultimate absorption requires the radiation reaction to fire *given* that the particle is at the boundary ($z = 0$). The probability of this event occurring at time r can be written using Bayes's rule as

$$p(r|z=0) = \frac{p(z=0|r)p(r)}{p(z=0)} \propto p(z=0|r)p(r), \tag{57}$$

where $p(z=0) = \int_0^\infty dr\, p(z=0|r)p(r)$ normalizes the distribution and is independent of time. The first term on the right, for a reflecting particle, is equivalent to the free-particle solution in one dimension: $p(z=0|r) = (4\pi Dr)^{-1/2}e^{-(0)^2/4Dr} \propto r^{-1/2}$. The second term is described by an exponential waiting time

distribution, whose time constant must be given by the only timescale in the problem, r_p: $p(r) = r_\mathrm{p}^{-1} e^{-r/r_\mathrm{p}}$. For $r \ll r_\mathrm{p}$, $p(r) \approx r_\mathrm{p}^{-1}$ is constant to leading order, and $p(r|z=0) \propto r^{-1/2}$, as in Eqn. 56.

At long times ($r \gg r_\mathrm{p}$), the erfc in Eqn. 55 can be approximated by its asymptotic limit, yielding

$$P(r) \approx \frac{1}{r_\mathrm{p}} \left[\frac{1}{\sqrt{\pi r/r_\mathrm{p}}} - e^{r/r_\mathrm{p}} \frac{e^{-r/r_\mathrm{p}}}{\sqrt{\pi r/r_\mathrm{p}}} \left(1 + \sum_{n=1}^{\infty} (-1)^n \frac{1 \cdot 3 \cdot 5 \cdot \ldots (2n-1)}{(2r/r_\mathrm{p})^n} \right) \right] \tag{58}$$

$$= \sqrt{\frac{r_\mathrm{p}}{4\pi}} r^{-3/2} \qquad r \gg r_\mathrm{p}, \tag{59}$$

to leading order. Such long times correspond to a search-dominated regime: the time spent far from the boundary diffusing is much greater than the time spent close to the boundary making short reflections. The process is therefore well approximated by a one-dimensional random walker returning to an absorbing origin, which scales as $r^{-3/2}$. Indeed, explicitly imposing an absorbing boundary by taking $k \to \infty$ makes the crossover time $r_\mathrm{p} \to 0$, such that the distribution scales as $r^{-3/2}$ for all times.

In the main text, we rescale r by the characteristic time to diffuse a molecular diameter ℓ, yielding the dimensionless rebinding time $\rho \equiv r/(\ell^2/D)$ and the associated crossover time $\rho_\mathrm{p} = r_p/(\ell^2/D) = (D/\ell k)^2$. The $\rho^{-1/2}$ and $\rho^{-3/2}$ scalings, as well as the crossover time ρ_p, are observed in Fig. 4 of the main text in the *planar* regime, in which a substrate molecule diffuses far enough from the membrane that the problem can be approximated as one-dimensional, but not so far that it encounters the reflective boundary opposite the membrane.

4 The capture effect and the benefit of linear sensitivity

When considering the rebinding of an S^* molecule to the membrane-bound E_a molecules in the double-modification network, the effect of the cytoplasmic E_d molecules is to bind and deactivate the S^* molecules with the longest excursion times, removing them from the rebinding problem. This leads to a truncation of the rebinding distributions beyond a characteristic timescale, which we call the *capture* time. The truncation alleviates the target size effect, imparting the clustered configuration with a shorter mean rebinding time than the random configuration.

The capture effect is illustrated in Fig. S1, in which E_d molecules are gradually introduced into the system. With E_d molecules present, an S^* molecule has two fates f: it may rebind to an E_a molecule ($f = +$) or be captured by an E_d molecule ($f = -$); measuring the time τ for either fate samples the joint distribution $p(\tau, f)$. The mean rebinding time $\bar{\rho}$ is then computed from the conditional distribution $p(\tau|+) = p(\tau, +)/p(+)$, where $p(+)$ is the total probability of rebinding as opposed to capture. Figure S1A shows that the difference in mean rebinding times between the random and clustered configurations, $\Delta\bar{\rho} \equiv \bar{\rho}^\mathrm{r} - \bar{\rho}^\mathrm{c}$, indeed increases as the ratio α of E_d to E_a molecules is increased. The increase subsides at large α, when the truncation at the capture time dominates the rebinding distribution, such that $\bar{\rho}$ approaches the capture time. The capture time can be estimated as the inverse of the association rate times the concentration of E_d molecules, $(k_4[E_d]_T)^{-1}$, or $\delta\zeta/\alpha\beta\mu$ in dimensionless units; accordingly, Fig. S1A demonstrates that $\bar{\rho}$ scales with α^{-1} at large α.

The result of the capture effect for the double-modification network is illustrated in Fig. S1B, which shows the maximal output ϕ_max as a function of α for a network with linear sensitivity. One observes that the clustered configuration produces a larger ϕ_max than the random configuration beyond an enzyme ratio of $\alpha^* \sim 1.5$ (Fig. S1B inset), indicating that clustering enhances the output when there are more deactivating enzymes than activating enzymes, such that the capture effect is strong. Indeed, at α^* the difference in mean rebinding times ($\Delta\bar{\rho} \sim 9 \times 10^3$) surpasses the corresponding difference in mean search times ($\Delta\bar{\sigma} \equiv \bar{\sigma}^\mathrm{c} - \bar{\sigma}^\mathrm{r} \approx 3 \times 10^3$) sufficiently to enhance to the output — that is, the capture effect overtakes the target size effect.

Figure S1B illustrates a clear trade-off: flooding the system with deactivating enzymes (i.e. increasing α) allows clustering to enhance the output beyond that of the random configuration, but it reduces the maximal output in general. In fact, the reduction with α was derived in the deterministic, well-mixed limit; the result (Eqn. 47) is shown in Figure S1 and provides a baseline from which the data diverge due to spatial effects. The intuition behind the reduction is straightforward: biasing the network toward deactivation reduces the

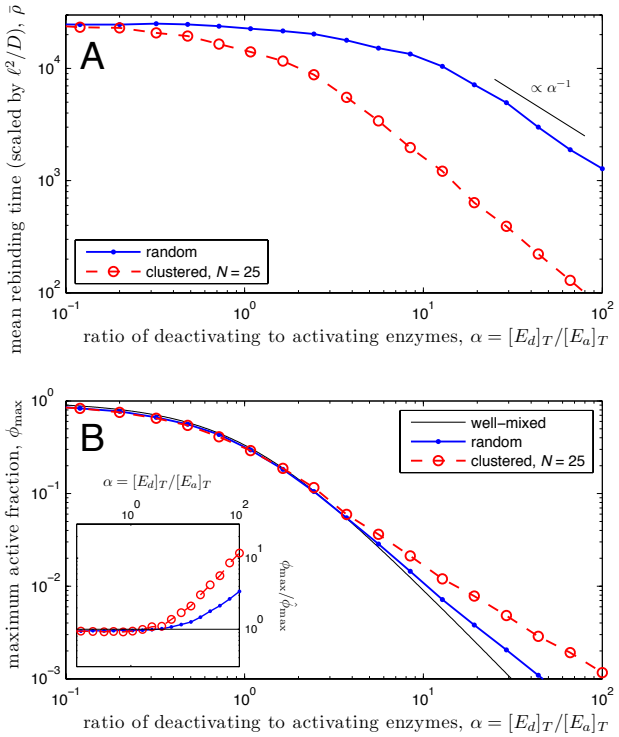

Figure S1: The capture effect. (A) As deactivating enzymes are introduced in the double-modification network, which increases the enzyme ratio α, the mean rebinding time $\bar{\rho}$ decreases more quickly for a clustered configuration than for a random configuration. Here $\mu = 0.01$, $\zeta = 25$, $N = 25$, $\delta = 10$, and $\beta = 1$. (B) Correspondingly, the maximal output $\phi_{\max}$ for the clustered configuration is higher than that for the random configuration at large α. Solid line shows the analytic prediction in the well-mixed limit (Eqn. 47), from which the data deviate due to spatial effects. Inset shows same data normalized by the well-mixed prediction $\hat{\phi}_{\max}$. Sensitivity is linear ($\gamma^{-1} = 0.05$, $\epsilon\gamma^{-1} = 0.01$) with other parameters as in A.

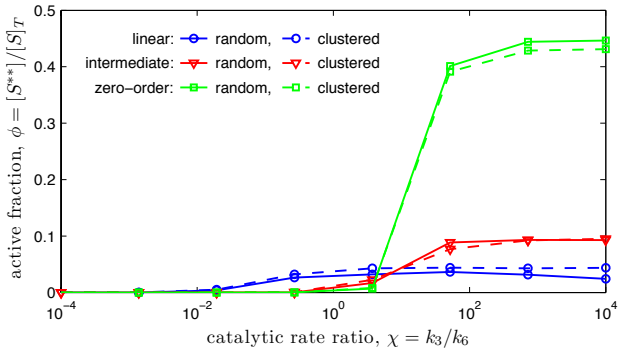

Figure S2: The effect of varying sensitivity on the response of a double-modification network. Network is deactivation-biased ($\alpha = 5$, $\beta = 1$) with intermediate diffusion ($\delta = 10$) and $N = 25$, $\mu = 0.01$, and $\zeta = 25$. Sensitivity is varied from linear ($\gamma^{-1} = 0.05$, $\epsilon\gamma^{-1} = 0.01$) to intermediate ($\gamma = 0.2$, $\epsilon = 0.1$) to zero-order ($\gamma = 0.05$, $\epsilon = 0.05$).

fraction of active substrate. The reduction, however, is unique to networks with linear sensitivity, which leads to the question: can clustering enhance the network response in the zero-order regime, in which the maximal output remains high?

The answer, revealed in Fig. S2, is no: as the sensitivity is varied from linear to zero-order, the enhancement with clustering vanishes, then reverses. The reason is that the capture effect, which underlies the enhancement with clustering, relies on the E_d molecules being free. The zero-order regime, on the other hand, corresponds to saturation of the enzymes by the substrate, such that at high input the E_d molecules are not free, but rather they are occupied with high probability. The mean capture time is then exceedingly long, such that the mean rebinding time is independent of the configuration of E_a molecules. The target size effect takes over, and the random configuration produces a higher output. The end result, as illustrated in Fig. S2, is that the benefit of clustering is specific to double-modification networks with linear sensitivity; in ultrasensitive networks the benefit is lost.

5 Finite back-reaction rates

Here we discuss the effects of systematically varying the back-reaction rates k_2 and k_5, which are assumed to be zero in the above sections and the main text. Specifically, we investigate how the main results of the study (Fig. 2 of the main text) change as k_2 and k_5 are individually varied across five orders of magnitude around k_6, which sets the timescale of reactions in the system. The findings are illustrated in Fig. S3 and can be understood as described below.

The most straightforward consequence of changing k_2 and k_5 is to change the Michaelis–Menten constants for the activation and deactivation processes, $K_a \equiv (k_2 + k_3)/k_1$ and $K \equiv (k_5 + k_6)/k_4$, respectively. As a result, increasing k_2 will increase the threshold of the response (Fig. S3A, B), and increasing k_5 will decrease the threshold of the response (Fig. S3C, D).

In the single-modification network, k_2 describes the rate at which an S molecule dissociates from an E_a molecule "prematurely", i.e. before it is activated. Fig. S3A demonstrates that increasing k_2 introduces no change in the response of the single-modification network besides the threshold shift; in particular, the reduction in the output upon clustering remains unchanged. One might be tempted to think that when k_2 is high, such that premature dissociation is frequent, clustering would promote more rapid rebinding of the S molecule to an E_a molecule and potentially change the response. However, unlike the rebinding of S^* molecules in the double-modification network, where trajectories can be terminated by deactivation, for S molecules here rebinding is the only fate. As described in the main text, without the possibility of termination, the mean rebinding time is independent of the spatial configuration of E_a molecules. We conclude that taking nonzero k_2 does not introduce any further dependence on configuration, but rather has the effect of renormalizing the catalytic rate k_3 to a reduced value, changing only the threshold of the response.

In the double-modification network, k_2 describes the rate of each of two processes: the premature dissociation of an S molecule from an E_a molecule and the premature dissociation of an S^* molecule from an E_a molecule. While the mean rebinding time of an S molecule is independent of spatial configuration as just discussed, the mean rebinding time of an S^* molecule is dependent on spatial configuration — this is, of course, the source of the enhancement in the response, as described in the main text. Nonetheless, we observe that increasing k_2 introduces no change in response of the double-modification network besides the threshold shift (Fig. S3B); in particular, the enhancement in the output upon clustering remains unchanged. One might hypothesize that when k_2 is high, such that premature S^* dissociation is frequent, the rapid rebinding of S^* molecules promoted by clustering would be compounded by the multiple rebinding events, thus enhancing the response further than for $k_2 = 0$. If such an effect were present, it would only be observable in the threshold region: below the threshold the output is suppressed, and above the threshold k_3 dominates over k_2, making premature dissociation rare. Fig. S3B demonstrates that, in fact, further enhancement with k_2 is not observed, even in the threshold region. We conclude that either the magnitude of the effect or the range of χ over which it occurs is negligible for these data.

In the single-modification network, k_5 describes the rate at which an S^* molecule dissociates from an E_d molecule prematurely, i.e. before it is deactivated. Fig. S3C demonstrates that the reduction in the output upon clustering persists as k_5 increases. This result is expected since the origin of the reduction, the target

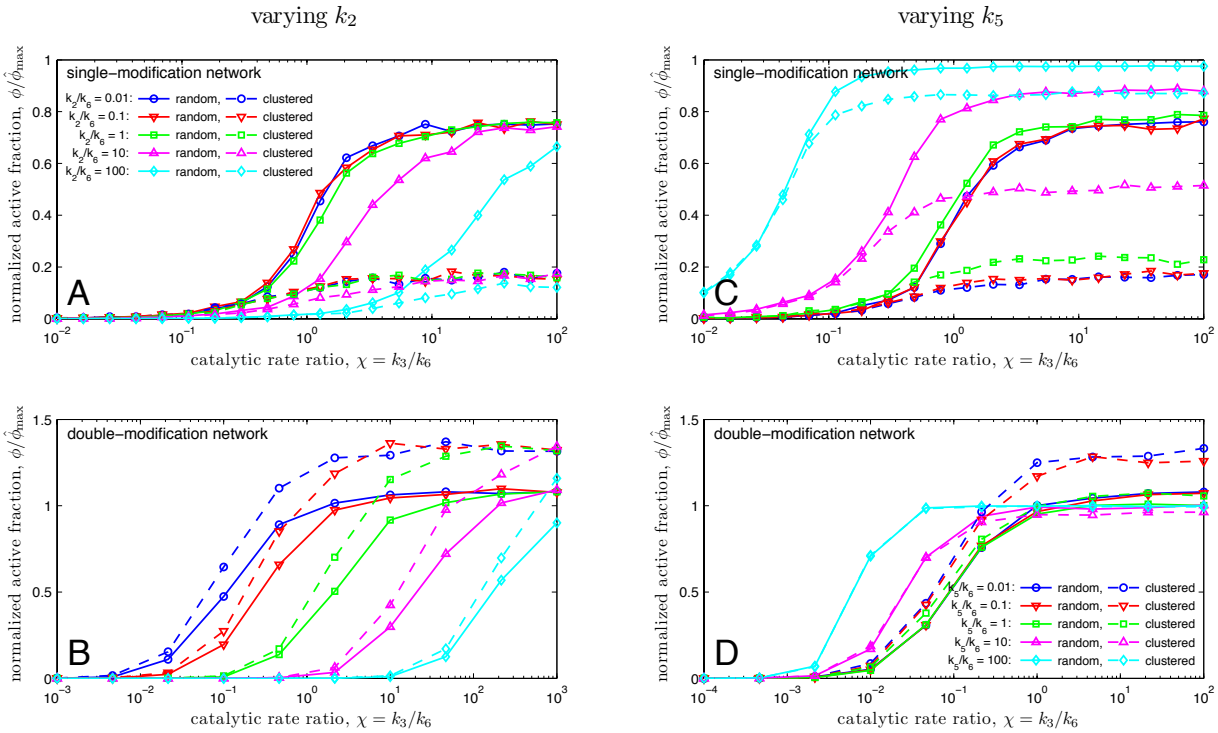

Figure S3: Input-output response of a single- (top) and double-modification network (bottom) as back-reaction rates k_2 (left) and k_5 (right) are varied across five orders of magnitude as indicated (legend in A also describes B; legend in D also describes C). Cluster size is $N = 25$, surface density is $\mu = 0.01$, and depth is $\zeta = 25$. Curves are normalized by the maximal value of the well-mixed response, $\hat{\phi}_{\max}$. (A, C) A symmetric ($\alpha = \beta = 1$) single-modification network with low diffusion ($\delta = 1/4\pi$) and zero-order sensitivity ($\gamma = \epsilon = 0.1$); here $\phi = [S^*]/[S]_T$. (B, D) A deactivation-biased ($\alpha = 5$, $\beta = 1$) double-modification network with intermediate diffusion ($\delta = 10$) and linear sensitivity ($\gamma^{-1} = 0.05$, $\epsilon\gamma^{-1} = 0.01$); here $\phi = [S^{**}]/[S]_T$. A note for visual clarity: in D, the random and clustered curves overlap for $k_5/k_6 = 100$.

size effect, is a general property of diffusive trajectories. Fig. S3C also demonstrates that, in addition to the threshold shift, the maximal output in both the random and the clustered configuration increases with k_5, ultimately approaching that of the well-mixed response, $\hat{\phi}_{\max}$. This result also makes sense, since increasing k_5 makes the E_d molecules less effective at rapidly deactivating the S^* molecules, which in turn weakens the spatial gradients of $[S^*]$ and $[E_d S^*]$ perpendicular to the membrane. As described in previous work (10) and noted in the main text, these gradients are responsible for the difference between the maximal outputs of the well-mixed and the membrane-localized (both random and clustered) responses; weakening the gradients therefore suppresses this difference.

In the double-modification network, k_5 describes the rate of each of two processes: the premature dissociation of an S^* molecule from an E_d molecule and the premature dissociation of an S^{**} molecule from an E_d molecule. Just as a prematurely dissociated S molecule has only one fate (to rebind to an E_a molecule), a prematurely dissociated S^{**} also has only one fate: to rebind to an E_d molecule. Therefore, the mean rebinding time of an S^{**} molecule is independent the spatial configuration, and premature S^{**} dissociation caused by nonzero k_5 does not introduce any further configuration dependence in the response. Premature S^* dissociation, on the other hand, does change the response: it reduces the ability of E_d molecules to successfully deactivate S^* molecules upon capture. Because the enhancement in the output upon clustering relies upon sufficient deactivation to terminate trajectories and truncate the rebinding time distribution, the enhancement vanishes for sufficiently high k_5, as demonstrated in Fig. S3D.

Supporting references

1. Morelli, M. J., and P. R. ten Wolde. 2008. Reaction Brownian dynamics and the effect of spatial fluctuations on the gain of a push-pull network. *J Chem Phys*. 129:054112.

2. Tian, T., A. Harding, K. Inder, S. Plowman, R. G. Parton, and J. F. Hancock. 2007. Plasma membrane nanoswitches generate high-fidelity Ras signal transduction. *Nat Cell Biol*. 9:905–14.

3. Li, M., and G. L. Hazelbauer. 2004. Cellular stoichiometry of the components of the chemotaxis signaling complex. *J Bacteriol*. 186:3687–94.

4. Phillips, R., J. Kondev, and J. Theriot. 2009. Physical biology of the cell. Garland Science.

5. Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. 1994. Molecular biology of the cell. Garland Publishing.

6. Goldbeter, A., and D. E. Koshland. 1981. An amplified sensitivity arising from covalent modification in biological systems. *Proc Natl Acad Sci USA*. 78:6840–4.

7. Beck, J. V., K. D. Cole, A. Haji-Sheikh, and B. Litkouhi. 1992. Heat conduction using Green's functions. Hemisphere.

8. Carslaw, H. S., and J. C. Jaeger. 1959. Conduction of heat in solids. Oxford: Clarendon Press, 2nd Ed.

9. Takahashi, K., S. Tanase-Nicola, and P. R. ten Wolde. 2010. Spatio-temporal correlations can drastically change the response of a MAPK pathway. *Proc Natl Acad Sci USA*. 107:2473–8.

10. van Albada, S. B., and P. R. ten Wolde. 2007. Enzyme localization can drastically affect signal amplification in signal transduction pathways. *PLoS Comput Biol*. 3:1925–34.