Combinatorial protein–protein interactions on a polymerizing scaffold

Andrés Ortiz-Muñoz[a,1], Héctor F. Medina-Abarca[b,1], and Walter Fontana[b,2,3]

[a]Division of Biology and Bioengineering, California Institute of Technology, Pasadena, CA 91125; and [b]Department of Systems Biology, Harvard Medical School, Boston, MA 02115

Edited by Peter Schuster, University of Vienna, Vienna, Austria, and approved December 24, 2019 (received for review July 24, 2019)

Scaffold proteins organize cellular processes by bringing signaling molecules into interaction, sometimes by forming large signalosomes. Several of these scaffolds are known to polymerize. Their assemblies should therefore not be understood as stoichiometric aggregates, but as combinatorial ensembles. We analyze the combinatorial interaction of ligands loaded on polymeric scaffolds, in both a continuum and discrete setting, and compare it with multivalent scaffolds with fixed number of binding sites. The quantity of interest is the abundance of ligand interaction possibilities—the catalytic potential *Q*—in a configurational mixture. Upon increasing scaffold abundance, scaffolding systems are known to first increase opportunities for ligand interaction and then to shut them down as ligands become isolated on distinct scaffolds. The polymerizing system stands out in that the dependency of *Q* on protomer concentration switches from being dominated by a first order to a second order term within a range determined by the polymerization affinity. This behavior boosts *Q* beyond that of any multivalent scaffold system. In addition, the subsequent drop-off is considerably mitigated in that *Q* decreases with half the power in protomer concentration than for any multivalent scaffold. We explain this behavior in terms of how the concentration profile of the polymer-length distribution adjusts to changes in protomer concentration and affinity. The discrete case turns out to be similar, but the behavior can be exaggerated at small protomer numbers because of a maximal polymer size, analogous to finite-size effects in bond percolation on a lattice.

polymerizing scaffold | combinatorial assembly | pleiomorphic ensemble

Protein–protein interactions underlying cellular signaling systems are mediated by a variety of structural elements, such as docking regions, modular recognition domains, and scaffold or adapter proteins (1, 2). These devices facilitate the evolution and control of connectivity within and among pathways. In particular, the scaffolding function of a protein can be conditional upon activation and serve to recruit further scaffolds, thus creating opportunities for network plasticity in real time. Scaffolds are involved in the formation of signalosomes, which are transient protein complexes that process and propagate signals. A case in point is the so-called "destruction complex" that tags β-catenin for degradation in the canonical Wnt pathway. β-Catenin is modified by CK1α and GSK3β without binding any of these kinases directly but interacting with them through an Axin scaffold (3, 4). In addition, the DIX domain in Axin allows for oriented Axin polymers (5, 6), while APC (another scaffold) can bind multiple copies of Axin (7), yielding Axin–APC aggregates to which kinases and their substrates bind. By virtue of their polymeric nature, such scaffold assemblies have no defined stoichiometry and may only exist as a heterogeneous combinatorial ensemble (8, 9)—also called "pleiomorphic ensemble" (10)—rather than a single well-defined complex. Deletion of Axin's DIX domain abolishes degradation of β-catenin (5), and mutations in APC that drive familial adenomatous polyposis map to truncations reducing the number of SAMP repeats at which APC binds Axin (11). These observations suggest a possible link between the size distribution of scaffolding aggregates and disease.

Interest in intracellular phase separation phenomena has increased since the discovery of P body dissolution/condensation in *Caenorhabditis elegans* (12). Much attention has been given to the physics underlying sol–gel transitions and polymerization (13–15). Here, we focus on the combinatorial aspects of ligand interactions on a (noncovalently) polymerizing scaffold, in particular, since scaffold-mediated interactions are subject to the prozone or "hook" effect (16–18): low scaffold concentrations promote interactions between ligands, but high concentrations oppose them by isolating ligands on different scaffold molecules. The main objective of our analysis is to gain insight into how this effect plays out in the context of a polymerizing-scaffold system and whether it affords opportunities for regulation. We proceed by defining and analyzing a simple model at a level of abstraction that only encapsulates combinatorial features without explicitly taking into account spatial constraints arising from polymer conformation.

The Polymerizing-Scaffold System

Let S (the scaffold) be an agent with four distinct binding sites {a,b,x,y}. At site y, agent S can reversibly bind site x of another S with affinity σ, forming (oriented) chains. For the time being, we exclude the formation of rings. Sites a and b can reversibly bind an agent of type A (the enzyme) and of type B (the substrate) with affinities α and β, respectively. All binding interactions are independent. When the system is closed, the total concentrations

Significance

Scaffold proteins play an important role in the control and evolution of cellular signaling processes by brokering interactions among docking proteins. In recent years it has become apparent that many scaffolds can form oligomers. We explore theoretically the combinatorial impact of scaffold polymerization on promoting interactions. It is well known that increasing scaffold abundance first expands opportunities for ligand interaction but then collapses them as ligands become separated on different scaffold instances. The polymerizing system stands out by significantly heightening interaction possibilities in the promoting phase and considerably mitigating their collapse. While the qualitative behavior of the system is robust, parameter changes—achievable through the very processes the system mediates—can considerably sculpt its quantitative behavior.

Author contributions: A.O.-M., H.F.M.-A., and W.F. designed research, performed research, contributed new reagents/analytic tools, and wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]A.O.-M. and H.F.M.-A. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: walter_fontana@hms.harvard.edu.

[3]Present address: Informatique et sciences numeriques, Collège de France, 75231 Paris Cedex 05, France.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1912745117/-/DCSupplemental.

First published January 24, 2020.

of A, B, and S are given by t_A, t_B, and t_S. This setup allows for a variety of configurations, such as shown in Fig. 1*A*. We posit that each enzyme A can act on each substrate B bound to the same complex. We refer to the number pq of potential interactions enabled by a configuration with sum formula $A_pS_nB_q$ as that configuration's "catalytic potential" Q. By extension, we will speak of the catalytic potential Q of a mixture of configurations as the sum of their catalytic potentials weighted by their concentrations.

If we assume that the assembly system equilibrates rapidly, the rate of product formation is given by Qk_{cat}, with k_{cat} the catalytic rate constant and Q the equilibrium abundance of potential interactions between A and B agents. Rapid equilibration is a less realistic assumption than a quasi-steady state but should nonetheless convey the essential behavior of the system. We first provide a continuum description of equilibrium Q in terms of concentrations (which do not impose a maximum polymer length) and then a discrete statistical physics treatment for the average equilibrium Q (where t_S is a natural number and imposes a maximum polymer length).

In the present context, molecular species Y_i assemble from $T=3$ distinct building blocks ("atoms") X_j through reversible binding interactions. The Y_i have a graphical (as opposed to geometric) structure (Fig. 1). We denote the number of atoms X_j in species Y_i with $\mu_{i,j}$. The equilibrium concentration y_i of any species Y_i can be obtained by recursion over the assembly reactions as $y_i=\varepsilon_i\prod_{j=1}^{T}x_j^{\mu_{i,j}}$, where $\varepsilon_i=1/\omega_i\prod_{r\in\mathcal{P}}K_r$ is the exponential of the free-energy content of Y_i. ω_i denotes the number of symmetries of Y_i, which, in our case, is always 1 because the polymers are oriented. The product runs over a sequence of reactions r that form an assembly path $\mathcal{P}$ of Y_i. In equilibrium, it is irrelevant which $\mathcal{P}$ one chooses. $K_r\in\{\alpha,\beta,\sigma\}$ is the equilibrium constant of the rth reaction, and the x_j are the equilibrium concentrations of free atoms of type j. Aside from the symmetry correction, ε_i is the exponential of a sum of binding energies, for example, $\varepsilon_i=\alpha^p\beta^q\sigma^r$ for a Y_i that contains p bonds between A and S, q bonds between B and S, and r bonds between S protomers.

Consider first the polymerization subsystem. From what we just laid out, the equilibrium concentration of a polymer of length l is $\sigma^{l-1}s^l$, where s is the equilibrium concentration of monomers of S. Summing over all polymer concentrations yields the total abundance of entities in the system,

$$W(s)=\sum_{l=1}^{\infty}\sigma^{l-1}s^l=s/(1-\sigma s). \qquad [1]$$

$W(s)$ yields the conservation relation, $t_S=s\,dW(s)/ds$, from which we obtain s as:

$$s=\frac{1}{4\sigma}\left(\sqrt{4+1/(\sigma t_S)}-\sqrt{1/(\sigma t_S)}\right)^2. \qquad [2]$$

Using Eq. **2** in $\sigma^{l-1}s^l$ yields the dependence of the polymer-size distribution on parameters t_S and σ. $W(s)$ has a critical point at $s_{\mathrm{cr}}=1/\sigma$, at which the concentrations of all length classes become identical. It is clear from Eq. **2** that s can never attain that critical value for finite σ and t_S.

The Chemostatted Case

In a chemostatted system, s can be fixed at any desired value, including the critical point $1/\sigma$. At this point, ever more protomers are drawn from the S reservoir into the system to compensate for their incorporation into polymers. We next consider the system with ligands A and B held fixed at concentrations a and b. Let $A_pS_nB_q$ be the sum formula of a scaffold polymer of length n with p A agents and q B agents. There are $\binom{n}{p}\binom{n}{q}$ such configurations, each with the same catalytic potential $Q=pq$. Summing up the equilibrium abundances of all configurations yields

$$W(s,a,b)=a+b+\frac{s(1+\alpha a)(1+\beta b)}{1-\sigma s(1+\alpha a)(1+\beta b)}. \qquad [3]$$

Eq. **3** corresponds to the $W(s)$ of ligand-free polymerization, Eq. **1**, by a coarse-graining that erases the ligand-binding state of scaffolds, i.e., by dropping terms not containing s and substituting s for $s(1+\alpha a)(1+\beta b)$. Eq. **3** indicates that, at constant chemical potential for A, B, and S, the presence of ligands lowers the critical point of polymerization to $s_{\mathrm{cr}}=1/(\sigma(1+\alpha a)(1+\beta b))$ because, in addition to polymerization, free S is also removed through binding with A and B.

Q_{poly}, the Q of the system, is obtained by summing up the Q of each configuration weighted by its equilibrium concentration (*SI Appendix*, section 1). Using W, we compute Q_{poly} as

$$Q_{\mathrm{poly}}=ab\frac{\partial^2}{\partial a\partial b}W=\alpha a\beta b\,s\,\frac{1+\sigma\,s(1+\alpha a)(1+\beta b)}{(1-\sigma s(1+\alpha a)(1+\beta b))^3}. \qquad [4]$$

Note that Q_{poly} inherits the critical point of W. The behavior of the chemostatted continuum model is summarized in Fig. 2. Q_{poly} (red) diverges as the polymerization system approaches the critical point. Fig. 2*A*, *Inset* shows the scaffold-length distribution at the black dot on the Q_{poly} profile. The red dotted curve reports the length distribution in the presence

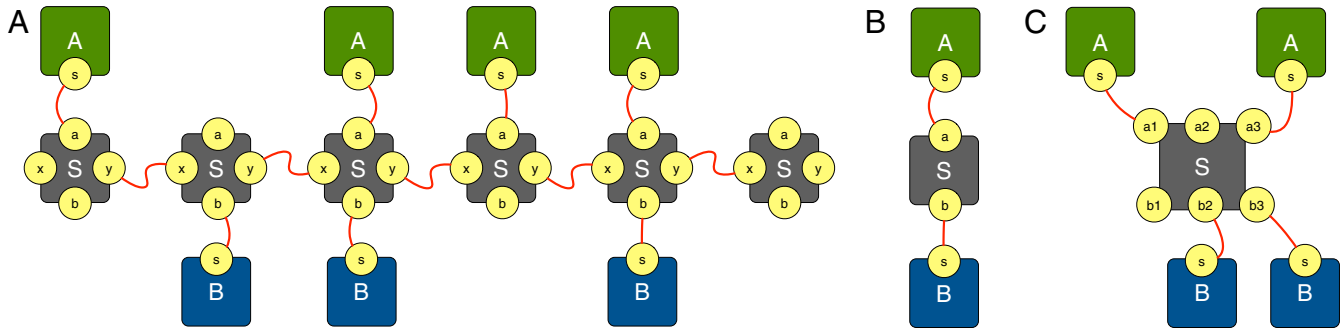

Fig. 1. Scaffold types. (*A*) Protomers *S*, each binding an enzyme *A* and a substrate *B* with affinities α and β, respectively, polymerize with affinity σ to yield a distribution of complexes. (*B*) A monovalent scaffold does not polymerize and has only one binding site for *A* and *B* each. (*C*) A *n*-valent (or multivalent) scaffold is a nonpolymerizing scaffold with *n* binding sites for *A* and *B* each. Here, $n=3$. The catalytic potential *Q* of a configuration is the number of possible interactions between *A* and *B* agents bound to the same complex: 12 in *A*, 1 in *B*, and 4 in *C*.

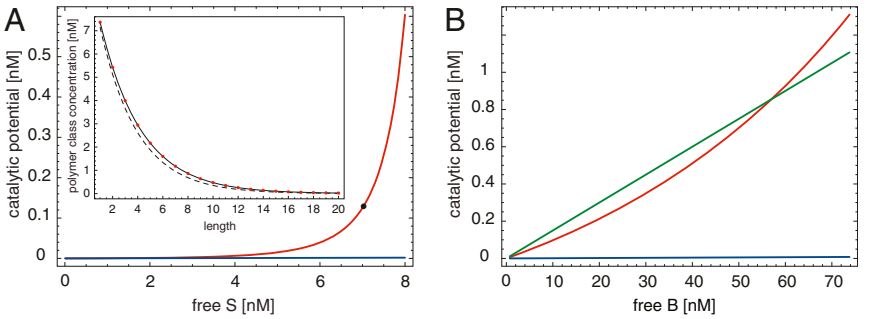

Fig. 2. Catalysis in a chemostatted polymerizing-scaffold system. (*A*) The red graph shows the catalytic potential Q as a function of chemostatted s according to Eq. **4** for $\alpha = \beta = 10^6$ M^{-1}, $\sigma = 10^8$ M^{-1}, and $a = b = 15 \cdot 10^{-9}$ M (about 9,000 molecules in a volume of 10^{-12} L). The blue curve is the special case of $\sigma = 0$, which is the monovalent scaffold system, $Q = \alpha a \beta b\, s$. *Inset* shows the scaffold-length distribution at $s = 7.15$ nM, corresponding to Q at the black filled circle. The critical point in this example is $s_{cr} \sim 9.7$ nM. (*B*) The catalytic potential at $s = 7.15$ nM as a function of clamped b (the substrate); other parameters are as in *A*. Red: polymerizing-scaffold system; blue: monovalent scaffold; green: chemostatted Michaelis–Menten in which *A* binds directly to *B* with affinity α.

of ligands, $[\{A_* S_k B_*\}] = \sigma^{-1}(\sigma s(1+\alpha a)(1+\beta b))^k$, whereas the black dotted curve reports the length distribution in the absence of ligands, $s_k \equiv [S_k] = \sigma^{k-1} s^k$. The presence of A and B shifts the distribution to longer chains. The blue curve in Fig. 2*A* shows the catalytic potential of the monovalent scaffold, $\sigma = 0$. It increases linearly with s but at an insignificant slope compared with the polymerizing case, which responds by raising the size distribution, thus drawing in more S from the reservoir to maintain a fixed s; this, in turn, draws more A and B into the system. In Fig. 2*B*, s is fixed and b, the substrate concentration, is increased. The green straight line is the Michaelis–Menten case, which consists in the direct formation of an AB complex and whose $Q = \alpha\, a\, b$ is linear in b. The red line is the polymerizing-scaffold system whose $s_{\rm cr}$ can be attained by just increasing b (Eq. **4**). All else being equal, there is a b at which more substrate can be processed than through direct interaction with an enzyme. The slope of the monovalent scaffold (blue) is not noticeable on this scale.

The Continuum Case in Equilibrium

We turn to the system with fixed resources t_S, t_A, and t_B, expressed as real-valued concentrations. Eq. **4** for $Q_{\rm poly}$ is now evaluated at the equilibrium concentrations s, a, and b of the free atoms. These are obtained by solving the system of conservation equations, $t_S = s\,\partial W/\partial s$, $t_A = a\,\partial W/\partial a$, and $t_B = b\,\partial W/\partial b$ (solutions in *SI Appendix*, section 1). The orange curve in Fig. 3*A* depicts the saturation curve of the catalytic potential $Q_{\rm direct}$ of the Michaelis–Menten mechanism for a fixed concentration t_A

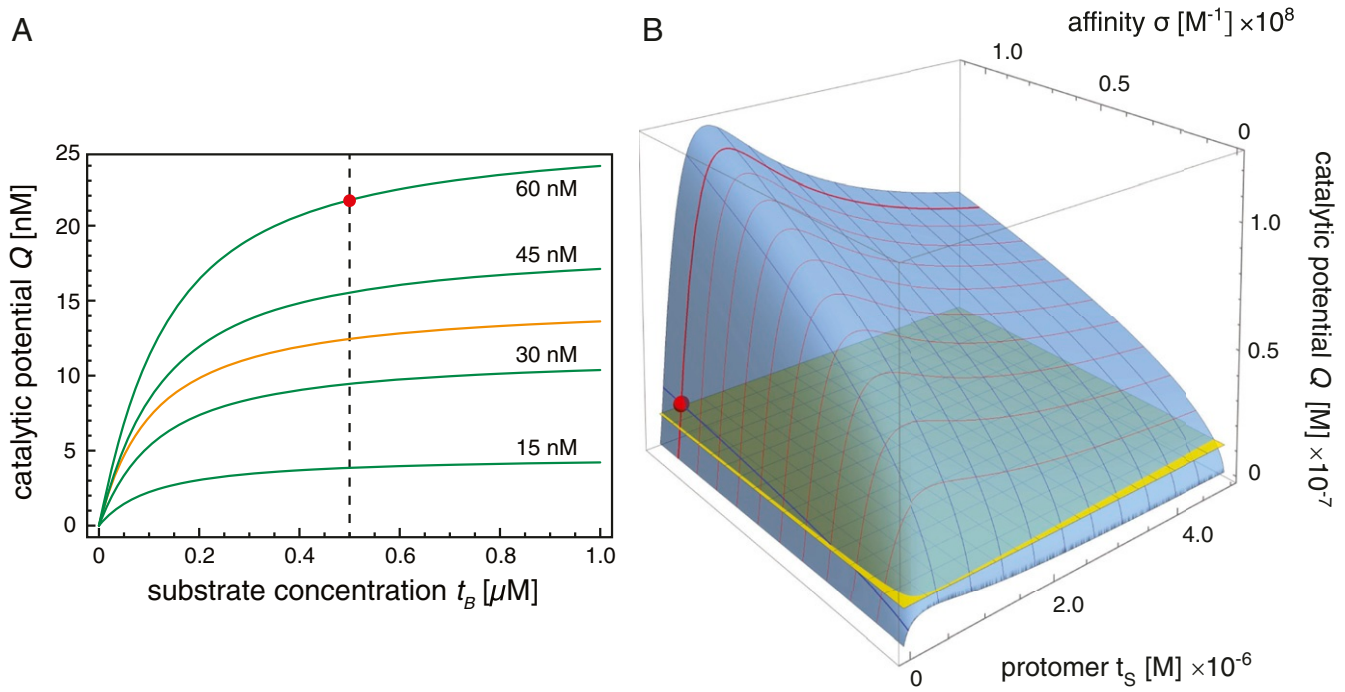

Fig. 3. Catalysis in an equilibrated polymerizing-scaffold system. (*A*) The orange curve shows the saturation of catalytic potential Q of the direct Michaelis–Menten type enzyme–substrate interaction as a function of total substrate t_B for $\beta = 10^7$ M^{-1} and $t_A = 15 \cdot 10^{-9}$ M. The green curves depict the saturation curves for Q of the polyscaffold with affinities $\alpha = \beta = 10^7$ M^{-1} and $\sigma = 10^8$ M^{-1} at various protomer abundances t_S. (*B*) The catalytic-potential surface for the polyscaffold as a function of t_S and σ; other parameters are as in *A*. The red ball corresponds to the conditions marked by the red dot in *A* ($t_B = 5 \cdot 10^{-7}$ M). The flat yellow surface is the Q for the direct enzyme–substrate interaction (i.e., the intersection of the vertical dotted line in *A* with the orange curve).

of enzyme as a function of substrate t_B. The green curves are saturation profiles of the polymerizing-scaffold system at varying protomer abundances t_S under the same condition. As in the chemostatted case, beyond some value of t_S, the catalytic potential of the polymerizing system exceeds that from direct interaction.

Q_{poly} can be modulated not only by the protomer concentration t_S but also the protomer affinity σ (Fig. 3*B*). Increasing t_S improves Q_{poly} dramatically at all affinities up to a maximum after which enzyme and substrate become progressively separated due to the prozone effect. At all protomer concentrations, in particular, around the maximizing one, Q_{poly} always increases with increasing affinity σ.

Comparison with Multivalent Scaffold Systems. With regard to Q, a polymer chain of length n is equivalent to a multivalent scaffold agent $S_{(n)}$ with n binding sites for A and B each. It is therefore illuminating to compare the polymerizing system with multivalent scaffolds and their mixtures.

The equilibrium concentration of configurations $A_p S_{(n)} B_q$ for an n-valent scaffold can be calculated by exploiting the independence of binding interactions (*SI Appendix*, section 2). The calculation yields as a general result that the catalytic potential for an arbitrary scaffolding system consists of two factors:

$$Q = \underbrace{p(t_{\mathrm{sit}}, t_A, \alpha)p(t_{\mathrm{sit}}, t_B, \beta)}_{I}\ \underbrace{Q_{\mathrm{max}}(\vec{t}_S)}_{II}. \qquad [5]$$

The dimensionless function $p(t_{\mathrm{sit}}, t_X, \gamma)$ denotes the equilibrium fraction of X-binding sites, with total concentration t_{sit}, that are occupied by ligands of type X, with total concentration t_X, interacting with affinity γ:

$$p(t_{\mathrm{sit}}, t_X, \gamma) = \frac{\gamma t_X - \gamma t_{\mathrm{sit}} - 1 + \sqrt{4\gamma t_X + (\gamma t_X - \gamma t_{\mathrm{sit}} - 1)^2}}{\gamma t_X - \gamma t_{\mathrm{sit}} + 1 + \sqrt{4\gamma t_X + (\gamma t_X - \gamma t_{\mathrm{sit}} - 1)^2}}.$$

This expression is the well-known dimerization equilibrium, computed at the level of sites rather than scaffolds and taken relative to t_{sit} (*SI Appendix*, section 2).

Factor *I* depends on the total concentration of ligand-binding sites (for each type) but not on how these sites are partitioned across the agents providing them. For example, a multivalent scaffold $S_{(n)}$, present at concentration $t_{S_{(n)}}$, provides $t_{\mathrm{sit}} = n t_{S_{(n)}}$ binding sites and the probability that a site of any particular agent is occupied is the same as the probability that a site in a pool of $n t_{S_{(n)}}$ sites is occupied. For a heterogeneous mixture of multivalent scaffold agents, we have $t_{\mathrm{sit}} = \sum_{i=1}^{n} i\, t_{S_{(i)}}$; for a polymerizing system in which each protomer S exposes one binding site, we have $t_{\mathrm{sit}} = t_S$.

Factor *II* is the maximal Q attainable in a scaffolding system. This factor depends on how sites are partitioned across scaffold agents with concentrations $\vec{t}_S = (t_{S_{(1)}}, \ldots, t_{S_{(n)}})$ but does not depend on ligand-binding equilibria. For example, a system of multivalent agents at concentrations $\vec{t}_S$ has $Q_{\mathrm{max}} = \sum_{i=1}^{n} i^2 t_{S_{(i)}}$. The polymerizing-scaffold system is analogous, but $n = \infty$ and the $t_{S_{(i)}}$ are determined endogenously by aggregation: $t_{S_{(i)}} = s_i = \sigma^{i-1} s^i$. This yields simple expressions for the catalytic potential of a polymerizing scaffold, Q_{poly}, and multivalent scaffold, Q_{multi}:

$$Q_{\mathrm{poly}} = p(t_S, t_A, \alpha)p(t_S, t_B, \beta)\frac{s(1+\sigma s)}{(1-\sigma s)^3} \qquad [6]$$

$$Q_{\mathrm{multi}} = p(n\, t_{S_{(n)}}, t_A, \alpha)p(n\, t_{S_{(n)}}, t_B, \beta) n^2 t_{S_{(n)}},$$

with s in Eq. **6** given by Eq. **2**. Eq. **6** is equivalent to Eq. **4**. While Eq. **4** requires solving a system of mass conservation equations to obtain a, b, and s, Q_{poly}, as given by Eq. **6**, does not refer to a and b but only to s, as determined by the ligand-free polymerization subsystem. The Q that shapes the Michaelis–Menten rate law under the assumption of rapid equilibration of enzyme–substrate binding has the same structure as Eq. **5**: $Q_{\mathrm{direct}} = p(t_A, t_B, \alpha)t_A$, where t_A and t_B are the total enzyme and substrate concentration, respectively. The presence of a second concurrent-binding equilibrium in Eq. **5** characterizes the prozone effect.

Adding sites, all else being equal, necessarily decreases the fraction p of sites bound. Specifically, factor I tends to zero like $1/t_{\mathrm{sit}}^2$ for large t_{sit}. In contrast, Q_{max} increases monotonically, since adding sites necessarily increases the maximal number of interaction opportunities between A and B. For a multivalent scaffold, Q_{max} diverges linearly with t_{sit}. For the polymerizing system, Q_{max} diverges like $t_{\mathrm{sit}}^{3/2}$ (*SI Appendix*, section 5).

Fig. 4*A* provides a wide-range comparison of Q_{poly} (red) with Q_{multi} for various valencies (blue) at the same site concentration $t_{\mathrm{sit}} = t_S$. On a log–log scale, scaffolds of arbitrary valency n exhibit a Q_{multi} whose slope as a function of t_{sit} is 1, with offset proportional to n, until close to the peak. For the polymerizing scaffold, the first-order term of the series expansion of Q_{poly} is independent of the affinity σ (*SI Appendix*, section 5), whereas the second-order term is linear in σ. Hence, for small t_{sit}, the polymerizing system behaves like a monovalent scaffold, and any multivalent scaffold offers a better catalytic potential. However, as t_S increases, the equilibrium shifts markedly toward polymerization, resulting in a slope of 2, which is steeper than that of any multivalent scaffold. The steepening of Q_{poly} is a consequence of longer chains siphoning off ligands from shorter ones (*SI Appendix*, section 4). All n-valent scaffolds reach their maximal Q_{multi} at the same abundance of sites $t_{\mathrm{sit}} = n\, t_{S_{(n)}} = t_S$ and before Q_{poly}. The superlinear growth in Q_{max} of the polymerizing system softens the decline of Q_{poly} to an order $t_S^{-1/2}$ for large t_S. In contrast, the decline of Q_{multi} is of order t_{sit}^{-1}. In sum, the polymerizing-scaffold system catches up with any multivalent scaffold, reaches peak-Q later, and declines much slower.

The mitigation of the prozone effect begs for a mechanistic explanation, in particular, since a prozone could occur not only within each length class but also between classes. To assess the within-class prozone, we think of a length class k as if it were an isolated k-valent scaffold population at concentration $t_{S_{(k)}} = s_k = \sigma^{k-1} s^k$ with $Q_{\mathrm{multi}} = p(k\, s_k, t_A, \alpha)p(k\, s_k, t_B, \beta)k^2 s_k$. Assuming equal affinity α for both ligands A and B, Q_{multi} peaks at $\hat{t}_{S_{(k)}} = k^{-1}(\alpha^{-1} + (t_A + t_B)/2)$. However, when established through a polymerization system, $t_{S_{(k)}} = s_k \leq 1/\sigma$ for any k and any t_S (*SI Appendix*, section 2 and Fig. S1*A*). This means that for k up to $\sigma/\alpha + \sigma(t_A + t_B)/2$, the concentration s_k of polymers of length k can never exceed the concentration required for the prozone peak $\hat{t}_{S_{(k)}}$. For the parameters used in the red curve of Fig. 4*B*, this value of k is about 35. To put this in perspective, in Fig. 4*B* at the yellow marker and at peak-Q_{poly}, 98 and 68%, respectively, of all sites are organized in length classes below 10. Thus, the most populated lengths avoid the within-class prozone entirely (for example, $k = 3$ in Fig. 4*C*, green solid line). However, the actual behavior of the kth length class occurs in the context of all other classes, i.e., at site concentration t_S, not just $k\, s_k$. In that frame, the class does exhibit a prozone (Fig. 4*C*, red solid line). Hence, the overall prozone of the polymerizing-scaffold system is mainly due to the distribution and ensuing isolation of ligands across length classes, not within. This "heterogeneity prozone"

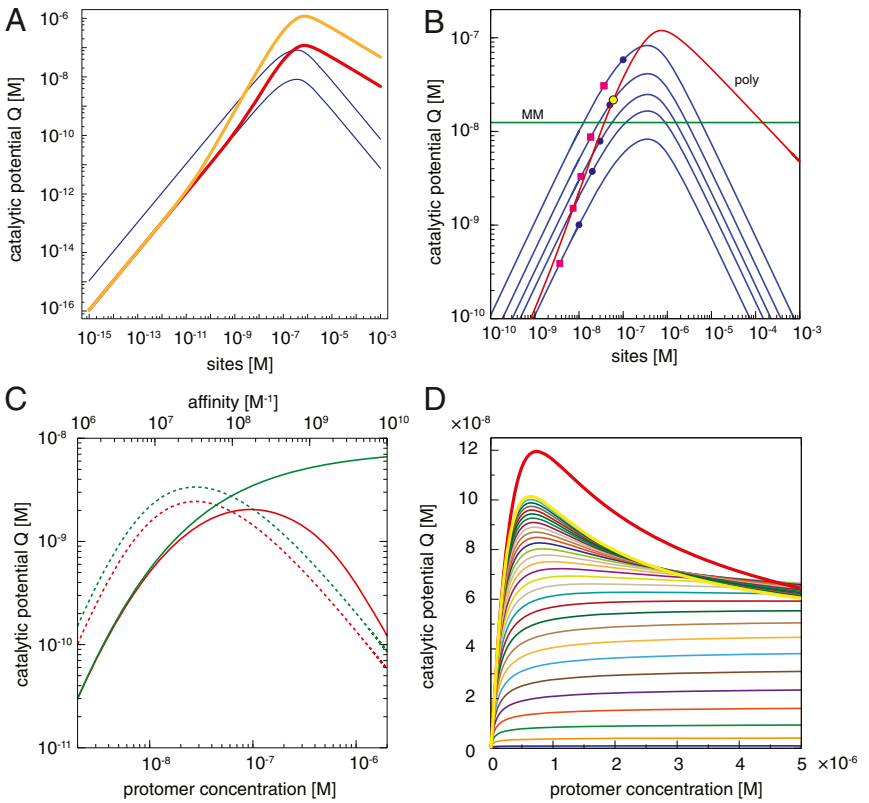

Fig. 4. Multivalent scaffolds and polymerizing scaffold. (*A*) Large-scale view of the catalytic potential Q as a function of site concentration t_S. The blue curves depict Q_{multi} for a monovalent (lower) and 10-valent scaffold. The location of the Q_{multi} peak is independent of the valency n when expressed as a function of $t_{\text{sit}} = t_S$ (*SI Appendix*, section 5, Eq. **38**). The red and orange curves depict Q_{poly} for two affinities, $\sigma = 10^8$ M^{-1} and $\sigma = 10^{10}$ M^{-1}, respectively. Other parameters: $\alpha = \beta = 10^7$ M^{-1}, $t_A = 1.5 \cdot 10^{-8}$ M, and $t_B = 5 \cdot 10^{-7}$ M. On a log–log scale, the up-slope of Q_{poly} is 1 initially—as for multivalent scaffolds—and increases to 2 prior to reaching the prozone peak. The down-slope is $-1/2$, whereas it is -1 for multivalent scaffolds (*SI Appendix*, section 5). (*B*) Close-up of the peak region in *A* for the red curve; multivalent scaffolds were added for $n = 2, 3, 5$. The slight asymmetry in the Q profiles of multivalent scaffolds stems from the differences in ligand concentrations of our running example (*SI Appendix*, section 11). The yellow dot on the Q_{poly} curve corresponds to the red dot in Fig. 3. A pink square on a blue curve of valency n marks Q_{multi} when the scaffold concentration $t_{S(n)}$ is the same as s_n at the t_S at which the length class n dominates the polymerizing system (*SI Appendix*, section 3 and Fig. S2*B*). The blue dots mark Q_{multi} at scaffold concentrations $t_{S(n)} = 1/\sigma$, the asymptotic (and maximal) value of s_n for all n in the limit of infinite t_S. These markers show that the prozone peak is never reached within the most populated length classes. "MM" labels the Michaelis–Menten case of Fig. 3 for comparison. (*C*) The solid lines in the graph exemplify the absence of a prozone within the isolated length class $n = 3$ and the occurrence of a prozone in the context of all other classes. Green solid: Q_{multi} for $n = 3$ using $t_{S_{(3)}} = s_3$ and $t_{\text{sit}} = 3\, t_{S_{(3)}}$. Red solid: Q_{multi} for $n = 3$ using $t_{S_{(3)}} = s_3$ but $t_{\text{sit}} = t_S$. The dotted lines illustrate the situation for $n = 3$ as a function of affinity σ (upper abscissa). In this dimension, the bending of the curves is not due to a prozone effect, since the number of sites does not increase. (*D*) Cumulative sums from $i = 1$ to $n = 30$ of Q_{multi} with $t_{S_{(i)}} = s_i$ and $t_{\text{sit}} = \sum_{i=1}^{n} i\, t_{S_{(i)}}$.

becomes noticeable only when including all length classes up to relatively high k because the majority of sites are concentrated at low k, where they are even jointly insufficient to cause a prozone (Fig. 4*D*).

At constant t_S and in the limit $\sigma \to \infty$, s_k tends toward zero for all k (*SI Appendix*, Fig. S3*C*). In the σ dimension, unlike in the t_S dimension, the class s_k itself has a peak. As σ increases, the k of the class that peaks at a given σ increases. Consequently, the Q_{multi} of each length class in isolation will show a "fake" prozone with increasing σ, due entirely to the polymerization wave passing through class k as it moves toward higher k while flattening (Fig. 4*C*, dotted lines). Since there is no site inflation, the overall Q_{poly} increases monotonically.

Effects of ligand imbalance and unequal ligand-binding affinities are discussed in the *SI Appendix*, section 11.

Interaction Horizon. The assumption that every A can interact with every B on the same scaffold can be relaxed by introducing an "interaction horizon," $q_{max}(l, h)$, defined as the number h of scaffold bonds within which a bound A can interact with a bound B on a polymer of size l. Thus, an A can interact with at most $2h + 1$ substrate agents B: h to its "left," h to its "right," and the one bound to the same protomer. The interaction horizon only affects the $Q_{\max}$ of a polymer of length l, thus replacing the interaction factor l^2 with (*SI Appendix*, section 6):

$$q_{max}(l,h) = \begin{cases} l(2h+1) - h(h+1), & \text{for } 0 \le h \le l-1 \\ l^2, & \text{for } h \ge l \end{cases}.$$

In the most restrictive scenario, we assume a fixed horizon h, independent of l. With this assumption, Eq. **6** becomes (*SI Appendix*, section 6)

$$Q_{\text{poly}} = p(t_S, t_A, \alpha)\, p(t_S, t_B, \beta)\, \frac{s\left(1 + \sigma s - 2(\sigma s)^{h+1}\right)}{(1 - \sigma s)^3}. \quad [7]$$

In Eq. **7**, the numerator of the $Q_{\max}$ term of Eq. **6** is corrected by $-2s(\sigma s)^{h+1}$. Since $\sigma s < 1$ for all finite t_S and σ, even

moderate values of h yield only a small correction to the base case of a limitless horizon.

The Discrete Case in Equilibrium

In the discrete case, we replace concentrations with particle numbers $t_S, t_A, t_B \in \mathbb{N}$ in a specified reaction volume V. In this setting, we must convert deterministic equilibrium constants, such as σ, to corresponding "stochastic" equilibrium constants σ_s through $\sigma_s = \sigma/(\mathcal{A}V)$, where $\mathcal{A}$ is Avogadro constant. For simplicity, we overload notation and use σ for σ_s.

The basic quantity we need to calculate is the average catalytic potential $\langle Q_{\text{poly}} \rangle = \sum_{l,i,j} i\, j \langle n_{lij} \rangle$, where $\langle n_{lij} \rangle$ is the average number of occurrences of a polymer of length l with i and j ligands of type A and B, respectively. Conceptually, $\langle n_{lij} \rangle$ counts the occurrences of an assembly configuration $A_i S_l B_j$ in every possible state of the system weighted by that state's Boltzmann probability. In *SI Appendix*, section 7, we show that $\langle n_{lij} \rangle$ is given by the number of ways of building one copy of $A_i S_l B_j$ from given resources (t_S, t_A, t_B) times the ratio of two partition functions—one based on a set of resources reduced by the amounts needed to build configuration $A_i S_l B_j$, the other based on the original resources. The posited independence of all binding processes in our model implies that the partition function is the product of the partition functions of polymerization and dimerization, which are straightforward to calculate (*SI Appendix*, section 8). While exact, the expressions we derive for $\langle Q_{\text{poly}} \rangle$ (*SI Appendix*, section 8, Eq. **66**) and $\langle Q_{\text{multi}} \rangle$ (*SI Appendix*, section 8, Eq. **69**) are sums of combinatorial terms and therefore not particularly revealing. For numerical evaluation of these expressions, we change the size of the system by a factor ξ (typically $\xi = 0.01$), i.e., we multiply volume and particle numbers with ξ and affinities with $1/\xi$. Such resizing preserves the average behavior. Our numerical examples therefore typically deal with 10 to 1,000 particles and stochastic affinities on the order of 10^{-2} to 10^{-1} molecules.

The key aspect of the discrete case is the existence of a largest polymer consisting of all t_S protomers. We refer to it as the "maximer"; no maximer exists in the continuum case because of the infinite fungibility of concentrations (Fig. S9). Since there is only one maximer for a given t_S, its expectation is the probability of observing it: $\langle s_{\max} \rangle = t_S!\, \sigma^{t_S - 1} / Z_{t_S}^{(\text{poly})}$, where $Z_{t_S}^{(\text{poly})}$ is the partition function of polymerization (*SI Appendix*, sections 8 and 9). This probability is graphed as a function of t_S and σ in Fig. 5*A*. At any fixed t_S, the probability of observing the maximer will tend to 1 in the limit $\sigma \to \infty$. This puts a ceiling to $Q_{\max}$ that is absent from the continuum description. In the t_S dimension, the maximer probability decreases as t_S increases at constant σ.

Polymerization as considered here has a natural analogy to bond percolation on a one-dimensional lattice. In the case of polymerization, the probability p that any two protomers are linked by a bond is a function of t_S and σ: $p = 1 - 2/(1 + \sqrt{1 + 4\sigma t_S})$ (Fig. 5*B* and *SI Appendix*, section 9). The salient observation is that for small t_S, the maximer has a significant probability of already occurring at modest affinities: for 10 protomers and $\sigma = 1$, p is already 0.78 and the maximer probability a respectable 0.06. For larger t_S, the maximer loses significance unless the affinity is scaled up correspondingly (*SI Appendix*, section 10). This is also reflected in the mass distribution (Fig. 5*C*).

Fig. 6*A* compares the discrete polymerizing-scaffold system with discrete multivalent scaffolds, much like Fig. 4*A* for the continuum case. The behavior of the discrete case is similar to that of the continuum case—with a few nuances that are prominent at low particle numbers and high affinities, such as the topmost orange curve. Its $\langle Q_{\text{poly}} \rangle$ profile does not hug the monovalent profile (bottom green chevron curve) to then increase its slope into the prozone peak as in the continuum case (Fig. 4*A*). A behavior like in the continuum case is observed for the lower orange and red curves, for which σ is much weaker. In the continuum case, the affinity does not affect slope—the slope always changes from 1 to 2—but determines where that change occurs (Fig. 4*A*). The higher the affinity, the earlier the change. The topmost orange curve could be seen as realizing an extreme version of the continuum behavior in which an exceptionally high affinity causes a change to slope 2 at unphysically low protomer concentrations. That such a scenario can be easily realized in the discrete case is due to the significant probability with which the maximer occurs at low particle numbers, similar to finite-size percolation. It bears emphasis that, as the number t_S of protomers increases, the maximer probability decreases

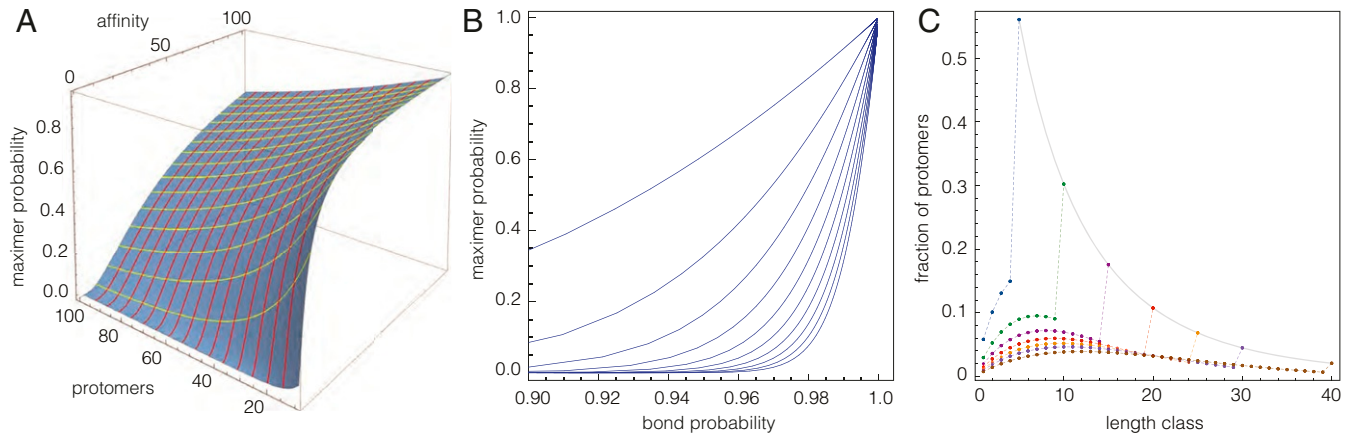

Fig. 5. Maximer. (*A*) The surface depicts the probability of observing the maximer as a function of t_S and σ. (*B*) Here, the maximer probability is graphed as function of the probability $p(t_S, \sigma)$ that a bond exists between two protomers. Each curve corresponds to a particular t_S with varying σ. t_S ranges from 10 (topmost curve) to 100 (bottom curve) in increments of 10, while σ ranges from 1 to 1,000. (*C*) Mass distributions in the polymerizing-scaffold model. Any curve depicts the fraction of protomers in all length classes n, computed as $n\, \sigma^{n-1} t_S!/(t_S - n)!\, Z_{t_S - n}^{(\text{poly})} / Z_{t_S}^{(\text{poly})}$, with $Z_{t_S}^{(\text{poly})}$ the partition function for polymerization with t_S protomers (*SI Appendix*, section 8). Each curve corresponds to a given number of protomers: $t_S = 5$ (blue), 10 (green), 15 (plum), 20 (red), 25 (orange), 30 (purple), 40 (brown); affinity $\sigma = 3$ in all cases. When t_S is small, the longest possible polymer—the "maximer"—is realized with appreciable frequency and dominates the mass distribution. As t_S increases, at fixed σ, the maximal length class increases too, but its dominance fades.

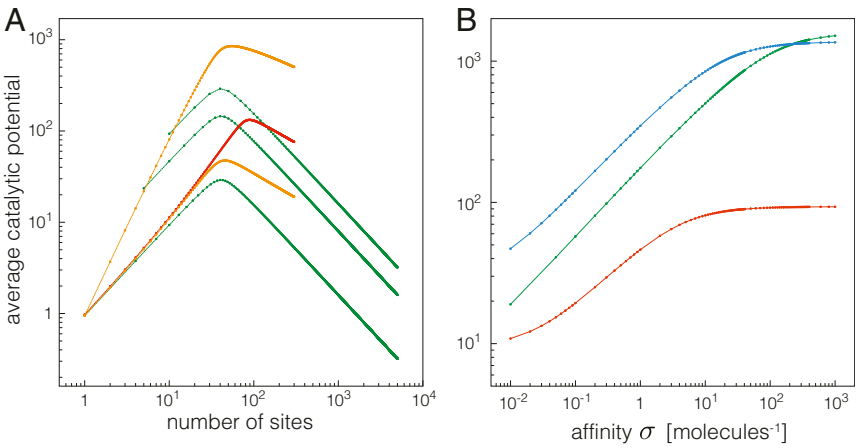

Fig. 6. Multivalent and polymerizing scaffolds in the discrete case. (*A*) Comparison of polymerizing scaffold (orange and red) with multivalent systems of various valencies (green). Orange: $t_A = t_B = 40$, $\alpha = \beta = 0.9$, $\sigma = 10$ (upper), and $\sigma = 0.01$ (lower). All affinities are in units of molecules^{-1}. Red: $t_A = t_B = 80$, $\alpha = \beta = 0.9$, and $\sigma = 0.01$. Green: $t_A = t_B = 40$, $\alpha = \beta = 0.9$, valency $n = 10$ (top), $n = 5$ (middle), and $n = 1$ (bottom). (*B*) $\langle Q_{\mathrm{poly}} \rangle$ as a function of affinity σ. $t_A = t_B = 40$, $\alpha = \beta = 0.9$, $t_S = 300$ (green), $t_S = 10$ (red), and $t_S = 50$ (blue).

(Fig. 5*C*), since the length of the maximer is t_S. However, once the maximer has receded in dominance, the increased number of length classes below it have gained occupancy and control the catalytic potential much like in the continuum case. Likewise, affinity does not appear to affect the slope of the downward leg as t_S increases.

The discrete multivalent scaffold system behaves much like its continuum counterpart. In the affinity dimension (Fig. 6*B*), the discrete system shows a behavior similar to the continuum case with the qualification that $\langle Q_{\mathrm{poly}} \rangle$ must level off to a constant, rather than increasing indefinitely. This is because, at constant t_S, an ever increasing affinity will eventually drive the system into its maximer ceiling. Because of the volume dependence of stochastic equilibrium constants, such an increase in affinity at constant protomer number can be achieved by any reduction of the effective reaction volume, for example, by confinement to a vesicle or localization to a membrane raft.

We determined SDs using stochastic simulations of the cases presented in Fig. 6*A* (*SI Appendix*, section 12). For a given $\langle Q \rangle$, the SD is larger after the prozone peak than before. Upon adding ligand-binding sites, the ratio of SD to mean (noise) increases much slower for the polymerizing system than for multivalent scaffolds.

Main Conclusions

Our theoretical analysis of a polymerizing-scaffold system shows that, at constant chemical potential, the system can be driven into criticality not only by increasing protomer concentration or affinity but by just increasing ligand concentrations. In equilibrium, polymerizing scaffolds exhibit a different type of prozone effect than multivalent ones: the polymerizing system provides a greater increase in catalytic potential on the upward leg above a certain protomer concentration t_S, delays the prozone peak, and significantly mitigates the collapse at high t_S. This behavior is caused by the response of the polymer-length distribution to changes in t_S and could be exploited to identify the presence of a polymerizing scaffold in an experimental setting. When particles are present in discrete numbers, not continuous concentrations, system behavior is affected by the existence of a maximal polymer length. Behavior easily attainable at small protomer numbers requires extreme parameter values in the concentration-based description. Other than that, the discrete and the concentration-based systems behave similarly.

A polymerizing scaffold concentrates ligands locally and facilitates their interaction. In this it functions like a compartment, but through a mechanism that can be readily regulated by varying protomer concentration t_S and polymerization affinity σ. We surmise that the regulation of catalytic potential in such systems is best modulated through the affinity σ, as this would not incur a prozone effect (Fig. 3*B* and Eq. **5**) while being presumably faster and less costly than modulating t_S. Regulation of σ could occur through posttranslational modification of the scaffold protomer. Regulation through σ is most effective at a t_S close to the prozone peak, whose location with respect to t_S is robust to changes in σ (Fig. 3*B*). Our analysis suggests that conformational constraints may not qualitatively subvert these observations, since adding a bond-distance constraint to ligand interactions did not fundamentally alter the combinatorial picture. Still, taking into account polymer conformation would increase realism. A significant extension of this work would consider scaffolding units of distinct types that form multiply interconnected aggregates. Such aggregates could control a larger diversity of ligand interactions and affect the aggregation/disaggregation transition.

ACKNOWLEDGMENTS. We gratefully acknowledge discussions with Tom Kolokotrones, Eric Deeds, and Daniel Merkle.

1. R. P. Bhattacharyya, A. Reményi, B. J. Yeh, W. A. Lim, Domains, motifs, and scaffolds: The role of modular interactions in the evolution and wiring of cell signaling circuits. *Annu. Rev. Biochem.* **75**, 655–680 (2006).
2. M. C. Good, J. G. Zalatan, W. A. Lim, Scaffold proteins: Hubs for controlling the flow of cellular information. *Science* **332**, 680–686 (2011).
3. C. Liu *et al.*, Control of β-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell* **108**, 837–847 (2002).
4. S. Ikeda *et al.*, Axin, a negative regulator of the wnt signaling pathway, forms a complex with GSK-3β and β-catenin and promotes GSK-3β-dependent phosphorylation of β-catenin. *EMBO J.* **17**, 1371–1384 (1998).
5. S. Kishida *et al.*, DIX domains of dvl and axin are necessary for protein interactions and their ability to regulate ?-catenin stability. *Mol. Cell. Biol.* **19**, 4414–4422 (1999).
6. M. Fiedler, C. Mendoza-Topaz, T. J. Rutherford, J. Mieszczanek, M. Bienz, Dishevelled interacts with the dix domain polymerization interface of axin to interfere with its function in down-regulating β-catenin. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1937–1942 (2011).
7. J. Behrens *et al.*, Functional interaction of an axin homolog, conductin, with β-catenin, apc, and gsk3β. *Science* **280**, 596–599 (1998).
8. E. J. Deeds, J. Krivine, J. Feret, V. Danos, W. Fontana, Combinatorial complexity and compositional drift in protein interaction networks. *PLoS One* **7**, e32032 (2012).

9. R. Suderman, E. Deeds, Machines vs. ensembles: Effective MAPK signaling through heterogeneous sets of protein complexes. *PLoS Comput. Biol.* **9**, e1003278 (2013).
10. B. J. Mayer, M. L. Blinov, L. M. Loew, Molecular machines or pleiomorphic ensembles: Signaling complexes revisited. *J. Biol.* **8**, 81 (2009).
11. J. Behrens *et al.*, Functional interaction of an axin homolog, conductin, with -catenin, APC, and GSK3. *Science* **280**, 596–599 (1998).
12. C. Brangwynne *et al.*, Germline p granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729–1732 (2009).
13. P. Li *et al.*, Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336–340 (2012).
14. L. Bergeron-Sandoval, N. Safaee, S. Michnick, Mechanisms and consequences of macromolecular phase separation. *Cell* **165**, 1067–1079 (2016).
15. S. Boeynaems *et al.*, Protein phase separation: A new phase in cell biology. *Trends Cell Biol.* **28**, 420–435 (2018).
16. D. Bray, S. Lay, Computer-based analysis of the binding steps in protein complex formation. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13493–13498 (1997).
17. J. E. Ferrell Jr, What do scaffold proteins really do. *Science STKE* **2000**, pe1 (2000).
18. A. Levchenko, J. Bruck, P. W. Sternberg, Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5818–5823 (2000).