Emergence of Single- versus Multi-State Allostery

Eric Rouviere [1,2] Rama Ranganathan,[1,3,*] and Olivier Rivoire [4,5,*]

[1]*Center for Physics of Evolving Systems, Department of Biochemistry & Molecular Biology, University of Chicago, Chicago, Illinois 60637, USA*

[2]*Graduate Program in Biophysical Sciences, University of Chicago, Chicago, Illinois 60637, USA*

[3]*The Pritzker School for Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA*

[4]*Center for Interdisciplinary Research in Biology (CIRB), Collège de France, CNRS, INSERM, Université PSL, Paris, France*

[5]*Gulliver UMR CNRS 7083, ESPCI-Paris, Université PSL, Paris, France*

(Received 18 September 2022; revised 22 May 2023; accepted 21 September 2023; published 9 November 2023)

Several physical mechanisms have been proposed to explain allostery in proteins. They differ by the number of internal states that they assume a protein to occupy, leaving open the question of what controls the emergence of these distinct physical forms of allostery. Here, we analyze a simplified model of protein allostery under a range of physical and evolutionary constraints. We find that a continuum of mechanisms between two archetypes emerges through evolution. In one limit, a single-state mechanism exists where ligand binding induces a displacement along a single normal mode, and in the other limit, a multi-state mechanism exists where ligand binding induces a switch across an energy barrier to a different stable state. Importantly, whenever the two mechanisms are possible, the multi-state mechanism confers a stronger allosteric effect and thus a selective advantage. This work defines the essential constraints that distinguish single- and multi-state allostery and sets the stage for a physical theory of its evolutionary origins.

DOI: 10.1103/PRXLife.1.023004

I. INTRODUCTION

Allostery, the change of activity of a macromolecule in response to a perturbation at a distance from its active site, is thought to be a ubiquitous feature of proteins [1]. Initially described in the context of multimeric proteins [2–4], it is now understood to underlie the regulation of proteins with diverse structural architectures, from receptors to signaling proteins and metabolic enzymes [5–9].

Efforts to explain how allostery works date back decades ago with the phenomenological models of Monod, Wyman, and Changeux (the MWC model) and Koshland, Nemethy, and Filmer (the KNF model) [3,4]. These models postulate that allosteric proteins occupy a small number of internal states between which transitions occur either spontaneously or upon interaction with a ligand. The MWC model postulates a thermal equilibrium between two distinct states, while the KNF model postulates that conformational changes are induced by binding events. These models have proved successful at fitting experimental data, and multiple extensions have been developed [4,10–12]. They leave, however, a fundamental question unanswered: Under what physical and evolutionary conditions are one or several internal states expected in allosteric systems?

Strategically, questions about the origins of allosteric mechanisms are facilitated by reduced physical models. In such models, the emergence of one or several states can be rigorously analyzed as a function of applied physical and evolutionary constraints, permitting generalizable insights. Indeed, simplified models of proteins have been studied extensively in the context of protein folding, a problem that they helped to significantly advance [13]. More recently, several such models have been developed to study the physics and evolution of allostery [14–23]. The idea is that by stripping down the complexities of real proteins to the essential physical features that control allostery, we can enable a foundation for better theory and experiment design. Here, we introduce a general model that encompasses previous work but extends it to address the origin and properties of different forms of allostery.

Using this model, we show that two archetypal mechanisms of cooperative allostery can arise, depending on the physical and selective constraints under which evolution takes place. First, we present a single-state mechanism where ligand binding actuates a soft normal mode. Second, we present a multi-state mechanism where ligand binding stabilizes an alternate stable state, resulting in a switchlike conformational change. The former necessarily emerges when the energy landscape is constrained to be smooth. In contrast, when the energy landscape has the possibility to be rugged, we find that the second mechanism provides a statistically more likely and more cooperative mechanism. We elucidate the origin of this evolutionary outcome with a simple analytical theory, and we provide a testable explanation for the pervasiveness of multiple states in allosteric proteins.

*Corresponding authors: ranganathanr@uchicago.edu; olivier.rivoire@espci.fr

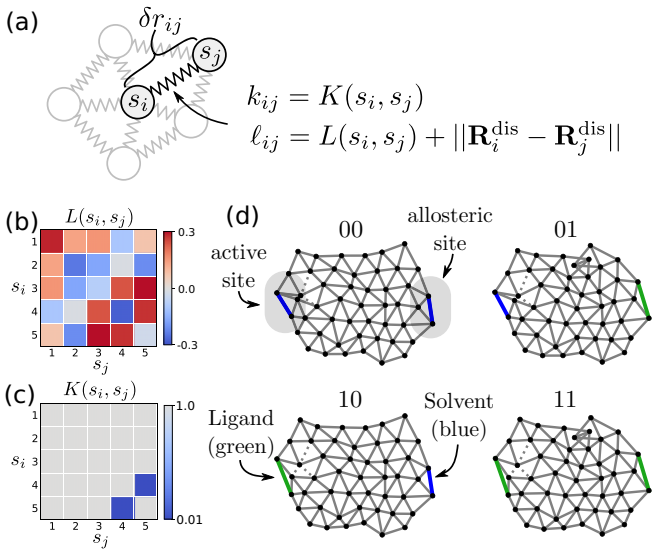

FIG. 1. The elastic network model. (a) The physical parameters of the networks are determined by a sequence $\mathbf{s} = (s_1, \ldots, s_N)$ that specifies the type $s_i \in \{1, \ldots, Q\}$ of each node i. Nodes are organized in a two-dimensional triangular lattice. The spring connecting nodes i and j has stiffness $K(s_i, s_j)$ and rest length $L(s_i, s_j) + ||\mathbf{R}_i^{\rm dis} - \mathbf{R}_j^{\rm dis}||$ (see the Appendix). (b) The table $L(s_i, s_j)$ has entries drawn uniformly in $[-\sigma a, \sigma a]$, where σ controls the disorder of the interactions. (c) The table $K(s_i, s_j)$ has all entries with $K(s_i, s_j) = 1$ except for one soft interaction with $K(s_i, s_j) = 0.01$. (d) Ground-state structures of an elastic network for the four possible ligand combinations. Solid lines represent stiff springs [$K(s_i, s_j) = 1$] and dashed lines represent soft springs [$K(s_i, s_j) = 0.01$]. The ligand binding is modeled by changing the rest length of springs between two pairs of nodes that define the active or allosteric sites. One value of the rest length defines a ligand (in green) and the other is the solvent (in blue).

II. MODEL

Our model abstracts proteins into a two-dimensional elastic network of residues whose interactions depend on their types. Abstractly, the residues are shown as nodes in the network [Fig. 1(a)]. The nodes can take Q values playing the role of the 20 amino acids constituting protein sequences [Fig. 1(a)]. The energy of a network with sequence $\mathbf{s} = (s_1, \ldots, s_N)$ and conformation $\mathbf{r} = (\mathbf{r}_1, \ldots, \mathbf{r}_N)$ is of the form

$$\begin{aligned} &U(\mathbf{r}, \mathbf{s}, \ell_{\rm act}, \ell_{\rm allo}) \\ &\quad = \frac{1}{2}\sum_{\langle i,j\rangle} k_{ij}(\delta r_{ij} - \ell_{ij})^2 + \frac{1}{2}k_{\rm act}(\delta r_{\rm act} - \ell_{\rm act})^2 \\ &\qquad + \frac{1}{2}k_{\rm allo}(\delta r_{\rm allo} - \ell_{\rm allo})^2, \end{aligned} \tag{1}$$

where the sum is over adjacent nodes in the network, and where $\delta r_{ij} = ||\mathbf{r}_i - \mathbf{r}_j||$ is the distance between nodes i and j, k_{ij} is the stiffness of the spring that connects them, and ℓ_{ij} is its rest length. Both of the values of k_{ij} and ℓ_{ij} depend on the types s_i and s_j:

$$k_{ij} = K(s_i, s_j), \quad \ell_{ij} = L(s_i, s_j) + \left|\left|\mathbf{R}_i^{\rm dis} - \mathbf{R}_j^{\rm dis}\right|\right|. \tag{2}$$

The values of $K(s_i, s_j)$ and $L(s_i, s_j)$ are given by $Q \times Q$ interaction tables [Figs. 1(b) and 1(c)], which are fixed during the evolution of the networks. We set all entries of the spring stiffness table $K(s_i, s_j)$ to be stiff ($K = 1$) except for one interaction that is soft ($K = 0.01$). $\mathbf{R}_i^{\rm dis}$ is a conformation generated from a disordered triangular lattice with mean spacing a (see the Appendix). The entries of the rest length deviation table $L(s_i, s_j)$ are drawn uniformly in $[-\sigma a, \sigma a]$, where σ is a dimensionless parameter controlling the variation in the rest length deviations. When $\sigma = 0$, networks are spatially disordered yet have a zero-energy ground state. As σ increases, $L(s_i, s_j)$ become heterogeneous and rest lengths deviate around the base rest lengths $||\mathbf{R}_i^{\rm dis} - \mathbf{R}_j^{\rm dis}||$.

To provide intuition about what σ represents, consider the absolute minimum energy of the network, known as the ground-state energy,

$$E(\mathbf{s}, \ell_{\rm act}, \ell_{\rm allo}) = \min_{\mathbf{r}} U(\mathbf{r}, \mathbf{s}, \ell_{\rm act}, \ell_{\rm allo}). \tag{3}$$

The energy is substantially lower than other local minima associated with excited states when $\sigma = 0$. Increasing σ increases the heterogeneity of the rest lengths, which leads to networks where no single conformation can minimize the energy of all springs simultaneously. This phenomenon, known as frustration, increases the energy of the ground state and decreases the energy difference between the ground and excited states. At large values ($\sigma = 0.3$) the energy is rugged with many nearly degenerate minima [see Fig. S1 of the Supplemental Material (SM) [25]].

We define two binding sites—an active and an allosteric site—on opposite sides of the network. The two sites correspond to springs of stiffnesses $k_{\rm act}$ and $k_{\rm allo}$ and rest lengths $\ell_{\rm act}$ and $\ell_{\rm allo}$ between nodes separated by distances $\delta r_{\rm act}$ and $\delta r_{\rm allo}$. Binding at these sites is modeled by changing the rest length of one of these springs from one value representing the "solvent" ℓ_0 to another value representing a "ligand" ℓ_1 [colored bonds in Fig. 1(d); their stiffness is, for simplicity, the same, $k_{\rm act} = k_{\rm allo} = 1$]. Four ground-state energies are therefore defined, depending on whether the two sites are unbound, $E_{00} = E(\ell_0, \ell_0)$, one site is bound, $E_{01} = E(\ell_0, \ell_1)$ and $E_{10} = E(\ell_1, \ell_0)$, or both are bound, $E_{11} = E(\ell_1, \ell_1)$. These energies depend on the sequence $\mathbf{s}$ and are estimated using a variant of the Basin-Hopping algorithm [24] (see the Appendix). Given the four energies, we quantify allostery by the binding cooperativity, the extent to which the binding energy at the active site depends on the presence of a ligand at the allosteric site, that is,

$$\Delta\Delta E(\mathbf{s}) = [E_{10}(\mathbf{s}) - E_{00}(\mathbf{s})] - [E_{11}(\mathbf{s}) - E_{01}(\mathbf{s})]. \tag{4}$$

We evolve the sequence $\mathbf{s}$ of a network using a Monte Carlo algorithm with the analog of mutations consisting of randomly changing the type of the node and a fitness defined by $\Delta\Delta E$ (see the Appendix).

III. RESULTS

A. Physical constraints on allostery

Previous elastic network models used to evolve allostery [15,17–20] shared two key design properties. First, they lacked frustration, the condition in which stressed springs

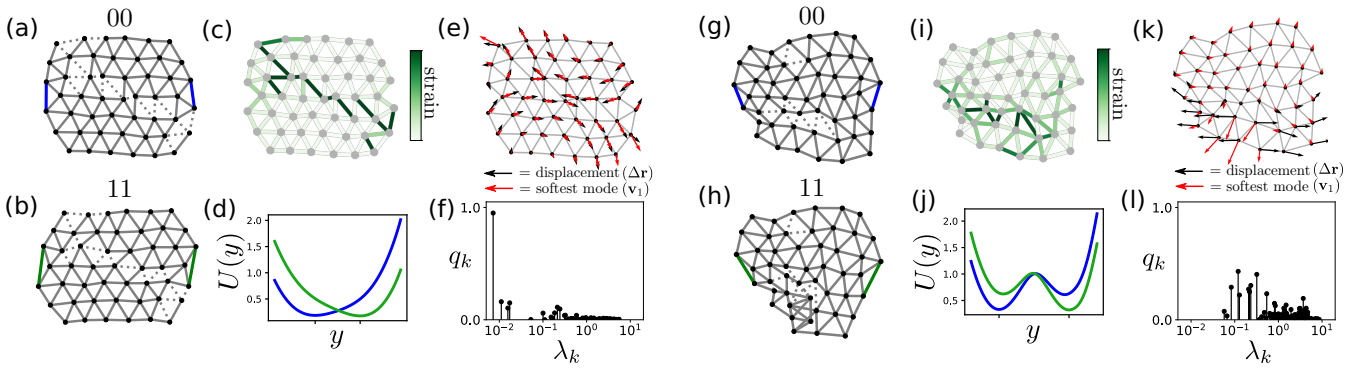

FIG. 2. Archetypes of networks with two different allosteric mechanisms. Parts (a)–(f) show a network evolved with homogeneous interactions ($\sigma = 0$), which displays a single-state mechanism. Parts (g)–(l) show a network evolved with disordered interactions ($\sigma = 0.3$), which displays a two-state mechanism. (a),(b) Ground-state structures of the fully solvated (00) and fully bound (11) networks. (c) Bond strain between the structures in (a) and (b). (d) Energy of the fully solvated (blue) and fully bound (green) networks along the conformational coordinate y interpolating between the structures in (a) and (b) (see the Appendix). (e) Structural displacements between the structures in (a) and (b) (00 $\rightarrow$ 11, in black) and motion of the softest mode of the unbound state (00, in red), showing that the two are aligned. (f) Overlap q_k between the structural displacement upon binding ligand $\Delta\mathbf{r}$ and each normal mode of the network $\mathbf{v}_k$ ($q_k = |\mathbf{v}_k \cdot \Delta\mathbf{r}/\|\Delta\mathbf{r}\|\,|$) as a function of the mode stiffness λ_k, which shows, as in (e), a strong overlap along the first mode. (g)–(l) The same analyses for a network with a two-state mechanism, showing, in contrast, a major conformational change [(h) vs (g)], an energy with two minima along the conformational coordinate, and no overlap between the displacement induced by the ligand and the softest normal modes (k)–(l).

occur in the ground state. This design results in energy landscapes with effectively one minimum. Second, mutations only changed the spring stiffness of interactions. Our model implements these design properties when $\sigma = 0$.

In this limit, evolving networks for allostery reproduce the mechanism described in previous works [16,17,19]—a soft normal mode connects the active site to the allosteric site, and ligand binding induces a strain that aligns with this soft mode [Figs. 2(a)–2(f)]. To explain, a soft normal mode defines a direction of motion along which the system can fluctuate substantially with near isoenergetic changes. We quantify the implication of this motion in allostery by computing the overlap between the network's softest mode $\mathbf{v}_1$ and the allosteric displacement $\Delta\mathbf{r}$ ($q_1 = |\mathbf{v}_1 \cdot \Delta\mathbf{r}/\|\Delta\mathbf{r}\|\,|$) (see the Appendix). The energy along the conformational coordinate, $U(y) = U(\mathbf{R}_{00} + y\Delta\mathbf{r})$, where $\mathbf{R}_{00}$ represents the conformation in the unbound ground state, has a single minimum that shifts upon ligand binding, corresponding to a conformational change [Fig. 2(d)]. For $\sigma = 0$, the ground-state structures in the unbound $\mathbf{R}_{00}$ and bound $\mathbf{R}_{11}$ conditions belong to the same basin of attraction [Fig. 2(d)]. Thus, we say that ligand binding induces a single-state conformational change (see the Appendix), in contrast to cases where the basins of attraction differ [e.g., see Fig. 2(j)]. Together, a large overlap of the conformational change with a single soft mode and the presence of a single energetic basin defines "single-state allostery."

Making the energy landscape more rugged by taking $\sigma > 0$ qualitatively changes this picture [Figs. 2(g)–2(l)]. A network evolved for allostery in this context can now switch upon binding to a different stable state that is separated by an energy barrier, a multi-state conformational change [Fig. 2(j)]. Additionally, the allosteric displacement no longer necessarily overlaps with the softest mode [Fig. 2(l)]. This mechanism defines "multi-state allostery."

Which of these mechanisms emerges over evolution depends critically on the diversity of available interactions imposed by σ: as σ increases, networks become more frustrated, resulting in more rugged energy landscapes. The fraction of evolved networks with a multi-state conformational change (symbolized by n_{cc}/n) increases while the overlap between the softest mode and the allosteric displacement decreases [Figs. 3(a) and 3(b)]. This is also the case for random networks [gray curves in Figs. 3(a) and 3(b)], but evolved networks show significant enrichment of multi-state mechanisms for large σ and of single-state mechanisms for low σ (green curves). In between these limits exists a continuum of mechanisms that are characterized by intermediate values of n_{cc}/n and q_1. The results of Figs. 3(a) and 3(b) are robust to parameter choice [see Fig. S2 of the SM [25]] and hold for 3D networks as well [see Fig. S3 of the SM [25]].

The difference between mechanisms is also illustrated by representing the ground-state energy of evolved networks as a function of different binding ligands, corresponding to varying the rest lengths ℓ_{act} and ℓ_{allo} of the springs defining the active and allosteric sites [Figs. 3(c) and 3(d)]. In this representation, the ground-state energy landscape associated with a single-state allosteric network typically takes the form of an anisotropic basin elongated in the $\ell_{\text{act}} = \ell_{\text{allo}}$ direction such that $E_{11} + E_{00} < E_{10} + E_{01}$. On the other hand, the energy landscape associated with a multi-state allosteric network can display two minima around $\ell_{\text{act}} = \ell_{\text{allo}} = \ell_0$ and $\ell_{\text{act}} = \ell_{\text{allo}} = \ell_1$, which also achieves $E_{11} + E_{00} < E_{10} + E_{01}$.

B. Evolutionary constraints on allostery

The previous results show that the diversity of available physical interactions determines the mechanism of allostery. Under conditions in which both mechanisms can evolve, which does evolution favor? We address this question by considering a set of interactions that contains both homogeneous ($\sigma = 0$) and disordered ($\sigma = 0.3$) rest length deviation distributions [Fig. 4(a)]. Networks evolved with such an in-

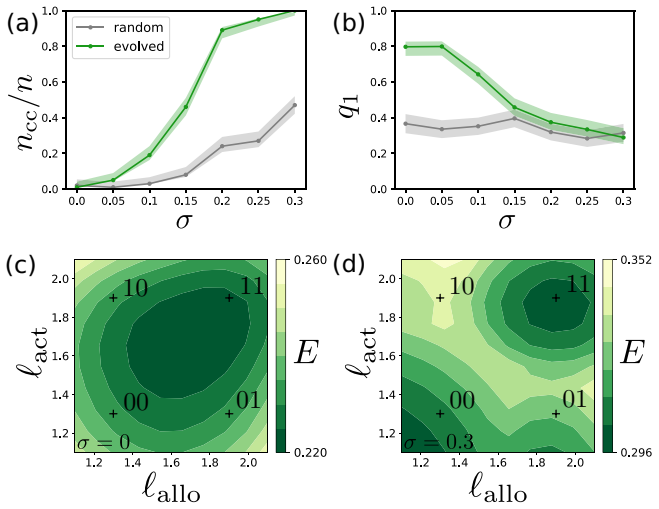

FIG. 3. Interaction disorder controls the mechanism of allostery. (a),(b) Statistics over $n = 100$ networks evolved to maximize $\Delta\Delta E$ over 500 Monte Carlo iterations as a function of the disorder σ of the interaction table $L(s_i, s_j)$. (a) Fraction of networks that undergo a multi-state conformational change n_{cc}/n (errors bars are 95% Wilson CI). (b) Mean overlap between the softest mode of the network $\mathbf{v}_1$ and the allosteric displacement upon ligand binding $\Delta\mathbf{r}$ computed as $q_1 = |\mathbf{v}_1 \cdot \Delta\mathbf{r}/\|\Delta\mathbf{r}\|\,|$ (error bars are 95% Wald CI). (c),(d) Examples of ground-state energy surfaces $E(\ell_{act}, \ell_{allo})$ for networks evolved with homogeneous and disordered interactions, respectively. The "+" marks the locations of different binding combinations, $\ell_{act}, \ell_{allo} = (\ell_0, \ell_0), (\ell_1, \ell_0), (\ell_0, \ell_1), (\ell_1, \ell_1)$.

teraction table have effectively the "choice" to populate a part of sequence space with either smooth or rugged local energy landscapes. Simulations show that networks evolved under a selection for cooperativity have a greater number of disordered interactions (those between nodes of types 1–5) than those of random networks (random, 24%; evolved, 34%). Correspondingly, the vast majority of these evolved networks display multi-state conformational changes (random, 10%; evolved, 90%) and an overlap with the softest mode comparable to those of random networks (random, $q_1 = 0.32$; evolved, $q_1 = 0.36$). These results are represented by the data point $E_{unfolded} = \infty$ in Fig. 4. They indicate that the multi-state mechanism is more competitive than the single-state mechanism. That is, when the interactions provide sufficient disorder, networks tend to evolve towards multi-state allostery.

In real proteins, the analog of our interaction tables is interactions between the amino acids. The physics of these interactions are largely fixed, and σ is therefore not subject to physiological control. But some tunable parameters may play a role similar to σ and effectively control the ruggedness of the energy landscape. One such parameter is thermal stability, which is itself subject to natural selection. As a proxy for thermal stability, we introduce here an arbitrarily defined energy $E_{unfolded}$ to represent the energy of a nonfunctional unfolded state. To impose a stability constraint, we again select networks based on $\Delta\Delta E$, but now we restrict the evolution to sequences satisfying $E_{00} < E_{unfolded}$.

By evolving networks for varying values of $E_{unfolded}$, we verify that smaller values of $E_{unfolded}$ result in more stable networks [Fig. 4(f)]. The most stable networks are less likely to undergo a multi-state conformational change upon ligand binding [Fig. 4(c)] and, on average, show a greater overlap between the allosteric displacement and the softest mode [Fig. 4(e)]. Additionally, the usage of disordered interactions decreases with decreasing $E_{unfolded}$ [Fig. 4(d)]. These results show how an additional selective pressure, here stability, can alter the likelihood to evolve a single- or multi-state mechanism of allostery. The underlying phenomenon is the same as before—stable networks minimize frustration and thus tend to populate smoother parts of the landscape, effectively corresponding to a smaller σ. Thus, thermal stability is a parameter that can, in principle, control the mechanism of allostery.

C. 1D model

The two archetypal mechanisms displayed by our two-dimensional (2D) network model can be illustrated in an even simpler one-dimensional (1D) model. This model consists of two rigid bars connected by three springs—two harmonic springs represent the active and allosteric sites, and a single elastic spring with the potential to flip mediates the mechanism of allostery [Fig. 5(a)]. The energy of this model is

$$U(x) = \tfrac{1}{2}[k_a(x - \ell_{act})^2 + k_m(|x| - \ell_m)^2 + k_a(x - \ell_{allo})^2]. \tag{5}$$

The ground-state energy $E(\ell_{act}, \ell_{allo}) = \min_x U(x)$ is harmonic with a single minimum when $\ell_m = 0$ and has two minima when $\ell_m > 0$ [Figs. 5(c) and 5(d)]. As in the 2D elastic network, we represent solvent and ligand by the algebraic values ℓ_0 and ℓ_1, respectively. For simplicity, we assume $\ell_0 = -\ell_1$. Under these assumptions, the cooperativity is

$$\Delta\Delta E = \frac{k_a(k_a\delta + 2k_m\ell_m)\delta}{2k_a + k_m}, \tag{6}$$

with $\delta = |\ell_1 - \ell_0|$ [see the SM [25]]. The mechanism is strictly single-state only for $\ell_m = 0$, in which case $\Delta\Delta E$ cannot exceed $k_a\delta^2/2$. On the other hand, when $\ell_m > 0$, $\Delta\Delta E$ can take arbitrarily large values and scales as $\Delta\Delta E \sim k_a\ell_m\delta$ when $k_m \to \infty$. A continuum of mechanisms exists between these two extremes [Fig. 5(e)], but, when viewing ℓ_m as a physical property of the landscape and k_m as an evolutionary parameter, two distinct regimes emerge: when $\ell_m < \delta/4$, a single-state mechanism with a soft mode ($k_m = 0$) is optimal, while when $\ell_m > \delta/4$, a two-state mechanism with a large barrier ($k_m = \infty$) is optimal. Finally, if ℓ_m is itself subject to evolution, maximal cooperativity $\Delta\Delta E$ is achieved by the two-state mechanism ($\ell_m = \infty, k_m = \infty$).

The analytical prediction of the 1D model (6) can be compared with data from the 2D model by defining 2D analogs, $\tilde{k}_m$ and $\tilde{\ell}_m$, of the 1D model's reduced parameters, k_m/k_a and ℓ_m/δ. We define $\tilde{k}_m = \lambda_1/\tilde{\lambda}$, where λ_1 is the stiffness of the softest nonzero mode and $\tilde{\lambda}$ is the median mode stiffness. This quantity captures the stiffness of the allosteric mode in single-state allostery that k_m/k_a measures in the 1D model. In the 1D model, ℓ_m is related to the height of the barrier of $U(x)$ when $k_a = 0$: $\Delta U_{barrier,1D} = k_m\ell_m^2/2$. To define an equivalent quantity for the 2D elastic network model, we consider $\tilde{\ell}_m = \delta^{-1}\sqrt{2\Delta U_{barrier}/\lambda_1}$ when $k_{act} = k_{allo} = 0$, and $\Delta U_{barrier}$

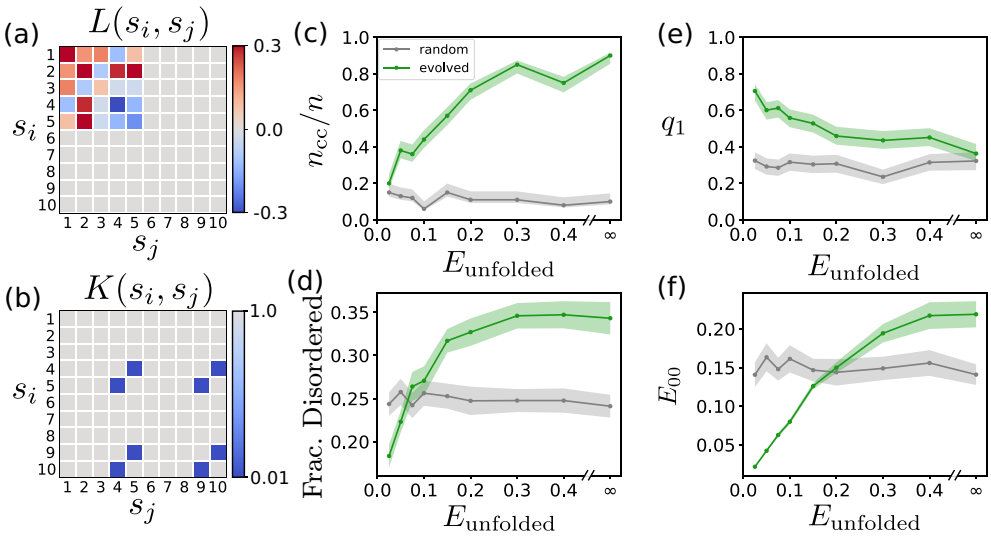

FIG. 4. Additional selective pressures can modulate the mechanism of allostery. (a) An example of a rest length deviation table $L(s_i, s_j)$ designed to support both homogeneous and disordered interactions. (b) Spring-constant interaction table $K(s_i, s_j)$. (c)–(f) Properties of networks evolved under a joint pressure for cooperativity (large $\Delta\Delta E$) and stability (small E_{00}) as a function of the intensity E_{unfolded} of the selective pressure for stability (statistics over 100 networks evolved through 500 Monte Carlo iterations). (c) Fraction of networks that undergo a two-state conformational change (errors bars are 95% Wilson CI). (d) Fraction of interactions between nodes of types 1–5 $[(s_i, s_j) \in \{1, 2, 3, 4, 5\}^2$; errors bars are 95% Wilson CI]. (e) Mean overlap of the allosteric displacement upon ligand binding ($00 \rightarrow 11$) and the softest nonzero mode of the network in the unbound state (errors bars are 95% Wald CI). (f) Ground-state energy of the unbound state (errors bars are 95% Wald CI). In the absence of constraint for stability ($E_{\text{unfolded}} = \infty$), networks evolve a multi-state mechanism using the disordered interactions. As the constraint for stability is increased (E_{unfolded} decreased), the networks tend to evolve a single-state mechanism relying on nondisordered interactions.

is the height of the network's barrier (see the Appendix). We find that nonallosteric random networks have large values of $\tilde{k}_m$ and small values of $\tilde{\ell}_m$ for all values of σ [Fig. 6(a)]. At small values of sigma, evolved, allosteric networks localize to the single-state region where $\tilde{k}_m$ and $\tilde{\ell}_m$ are both small as predicted by the 1D model. Increasing σ enables networks to have larger energy barriers and larger values of $\tilde{\ell}_m$ [see Fig. S4A of the SM [25]]. Evolved networks at large σ take advantage of this by switching to a multi-state mechanism and localizing to the large $\tilde{k}_m$, large $\tilde{\ell}_m$ region. Comparing Figs. 5(e) and 6(a) shows the remarkable ability of the 1D theory to qualitatively describe the complex behavior from the 2D network.

Equation (6) predicts that cooperativity of the two mechanisms should scale differently with the magnitude of ligand binding perturbation $\delta = |\ell_1 - \ell_0|$ ($\Delta\Delta E \sim \delta^\gamma$: single-state, $\gamma = 2$; and multi-state, $\gamma = 1$). When $\sigma = 0$ and networks use the single-state mechanism, we find $\gamma \approx 1.95$ for evolved networks [Fig. 6(b)]. Increasing σ results in decreasing values of γ toward unity. This result suggests a testable hypothesis: If smaller ligands induce more subtle perturbations, then allosteric proteins that bind small molecules should use the multi-state mechanism more often than allosteric proteins whose ligands are other large proteins.

To summarize, the 1D model identifies the key elements necessary for understanding the connection between single-state and multi-state allostery. It suggests that the specific details of the 2D network, which are absent in the 1D model, can vary with little effect [Fig. S2 of the SM [25]].

IV. DISCUSSION

In this work, we study the evolutionary origins of single or multiple states in systems selected for allostery using a minimal yet generic physical model of protein allostery. Depending on the ruggedness of the underlying energy landscapes, two archetypal mechanisms emerge: in smooth landscapes a single-state mechanism evolves, while in rugged landscapes a multi-state mechanism evolves. In single-state systems, allostery is mediated through the actuation of a soft global mode induced by ligand binding, whereas in multi-state systems, ligand binding stabilizes an alternate preexisting state. We find that the multi-state mechanism has a greater potential for cooperativity and is therefore evolutionarily favored whenever it has the possibility to evolve. We also find that additional selective pressures can modulate the outcome of evolution; in particular, a strong selection for stability favors the evolution of single-state mechanisms. These results are demonstrated by numerical simulations in a two-dimensional elastic network model and recapitulated by analytical calculations in a simpler one-dimensional model.

Elastic network models have been extensively studied as coarse-grained models for proteins and have been shown to accurately capture thermal fluctuations [26,27]. Here, we use these models to implement a generic energy landscape in which the allosteric role of soft modes and multiple states can be understood. Multistate conformational changes take the form of buckling instability in the elastic network model, while in proteins, state transitions may take many different

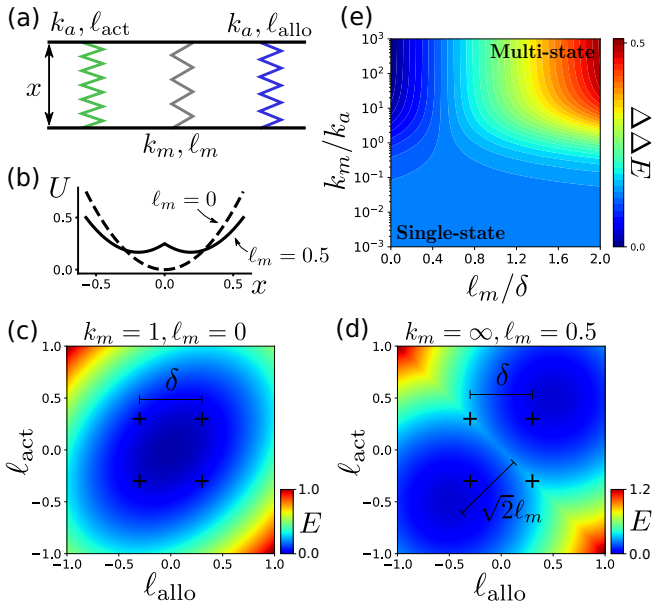

FIG. 5. A 1D model recapitulates the results obtained with 2D elastic network. (a) The model consists of three springs connecting two rigid bars constrained to move in one dimension. Two springs are analogous to the ligand springs of the elastic network. The third spring, whose contribution to the energy is $k_m(|x| - \ell_m)^2/2$, mediates the mechanism of allostery and can flip. (b) Total energy $U(x)$ as a function of x for two values of ℓ_m: the energy landscape has a single minimum $\ell_m = 0$ and two minima when $\ell_m > 0$. (c) The ground-state energy of the 1D model, $E(\ell_{\rm act}, \ell_{\rm allo})$. Small values of k_m, ℓ_m correspond to a single-state mechanism. (d) The ground-state energy of the 1D model, $E(\ell_{\rm act}, \ell_{\rm allo})$. Large values of k_m, ℓ_m correspond to a multi-state mechanism. (e) Cooperativity $\Delta\Delta E$ as a function of the normalized quantities k_m/k_a and ℓ_m/δ, showing a continuum between purely single-state and strongly two-state mechanisms.

physical forms [5,28–30]. Regardless, the essential feature of multi-state conformational switching is a barrier between two energy minima, which is shared in simple models and real proteins. The fact that our results extend to 3D networks [Fig. S3 of the SM [25]] and they are recapitulated by a 1D model of three springs (Figs. 5 and 6) indicates that they are robust to details of the geometry of the network and to the nature of the potential. Indeed, the only feature necessary to control single- or multi-state outcomes is tunable ruggedness. An implication is that as seen by simulations [Fig. 4(f)], multi-state proteins that varied through directed evolution protocols to optimize stability should lose their multi-state character due to loss of ruggedness, a prediction that can be experimentally tested. These findings illustrate the clarifying value of simple physical models in isolating the essential parameters from the large ensemble of complex properties exhibited by natural proteins.

For computational expedience, the simulations here are at zero temperature, ignoring changes in system entropy. However, the distinction between single-state and multi-state mechanisms extends to finite temperatures in a straightforward manner by simply considering the free energy instead of the ground-state energy. We carried out this generalization for the one-dimensional model, showing that temperature does not modify the general conclusions reported here [see the SM [25]]. In future work, it will be valuable to extend this work to consider "dynamic allostery" in which long-range interactions are driven purely by propagated changes in entropy [31–33]. The route to this extension is by enabling ligand binding to stiffen ligand binding sites rather than applying a force or a displacement. Our model also does not explicitly account for partial or global unfolding [34], but allosteric mechanisms where ligand binding differentially stabilizes the folded and unfolded ensembles are also instances of multi-state mechanisms when folding is a cooperative process [35].

The distinction between single- and multi-state mechanisms of allostery reflects the two classes of phenomenological models that have been introduced to describe allostery, namely the MWC and KNF models [3,4]. The KNF model, where a state changes only upon ligand binding, assumes a single-state mechanism, while the MWC model, where two states exist before ligand binding, assumes a two-state mechanism. Indeed, the physical model described here can be reformulated from a thermodynamic perspective to be analogous to MWC and KNF models (see the SM [25]). The MWC and KNF models are archetypes at the two ends of a continuum of models for allostery that interpolate between them [11]. Thus the work presented here provides a foundation for understanding the assumptions underlying these models, and perhaps more importantly, the evolutionary origin of specific models for allostery.

The central parameter controlling single- or multi-state allostery is the heterogeneity σ of interactions, leading to frustration and ruggedness in the energy landscape of the model networks. Indeed, a significant prior work shows that real proteins are frustrated networks [36,37] and that this property is important for allostery [38]. Since conformational switching between stable states [5,39,40], folding intermediates [41], and misfolding [42] are commonplace in proteins, we conclude that amino acid interactions display the kind of heterogeneity in natural proteins that enables the evolution of multi-state allostery. Indeed, these considerations together with the work presented here suggest that a multi-state mechanism is in fact statistically more likely in natural allosteric proteins.

This statement is consistent with the diversity of mechanisms underlying allostery observed in proteins. Previous work using normal mode analysis to estimate the largest overlap between the allosteric displacement and any normal mode ($\max_i\ q_i$) of proteins has found a broad range of values. Several studies converge to find that the largest overlaps in allosteric proteins are of the order of 0.5 with a standard deviation of the order 0.2 [21,43]. When we measure the largest overlap $\max_i\ q_i$ rather than the overlap with the softest mode q_1, we find overlaps of a similar order ($\gtrsim$0.5) even for nonallosteric random networks [Fig. S6EF of the SM [25]]. Additionally, we find that many evolved allosteric networks lie in the continuum between the two archetypal mechanisms, with both an energy barrier separating multiple states and some overlap with the softest mode $q_1 > 0.5$ [Fig. S6A-D,K,L of the SM [25]]. Thus, the empirical finding that the allosteric conformational change overlaps with the softest mode does not signify the absence of a multi-state conformational change. This point has been made by Gur, Zomot, and Bahar

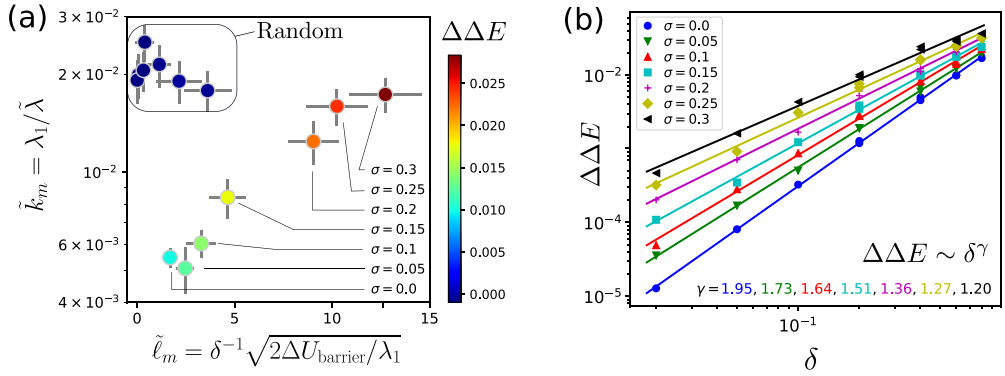

FIG. 6. 1D model predicts the behavior of the 2D elastic networks. $\tilde{k}_m$ and $\tilde{\ell}_m$ are the 2D elastic network analogs of the 1D model's parameters k_m and ℓ_m. (a) The mean $\tilde{k}_m$ and $\tilde{\ell}_m$ and $\Delta\Delta E$ are plotted for 100 random and 100 evolved networks at different σ. Error bars are 95% CI. Consistent with the 1D model, nonallosteric random networks localize to the large $\tilde{k}_m$ and small $\tilde{\ell}_m$ corner. When σ is small, networks approach the single-state mechanism limit ($\ell_m = 0$, $k_m = 0$). As σ increases, networks localize towards the multi-state mechanism limit of large k_m, large ℓ_m. (b) Scaling of allosteric cooperativity $\Delta\Delta E$ with the magnitude of ligand binding perturbation $\delta = |\ell_1 - \ell_0|$. Each data point shows the mean $\Delta\Delta E$ averaged over 100 evolved networks with parameters σ and δ. For each σ, data are fit to the power law $\Delta\Delta E \sim \delta^{\gamma}$, and fitted γ values are shown at the bottom. The 1D model predicts that in the single-state limit, $\gamma = 2$, while in the multi-state limit $\gamma = 1$. 2D networks evolved at $\sigma = 0$ use a single-state mechanism and show a scaling of $\gamma \approx 1.95$. As σ increases, networks increasingly use the multi-state mechanism, and values of γ limit toward one.

based on analysis of elastic networks by molecular-dynamics simulations [44].

An interesting perspective is the potential value of simplified models for a better understanding of computational approaches for inferring allosteric networks from extant sequence data [45]. Indeed, statistical analyses of homologous protein sequences has demonstrated the existence of sparse, collectively evolving networks of amino acids within proteins that mediate allosteric communication [46–48]. It will be interesting to use the simplified models presented here as a ground truth to study how the sequence-based inference process depends on both the sampling of input data and the parameters that control the emergence of different forms of allostery.

While our work focuses on allostery in the context of proteins, the implications of these results extend more broadly to the emerging field of learning in physical systems [49]. The idea of physical learning is to start with a randomly configured (and nonfunctional) physical system and to iteratively optimize its performance for a chosen task by varying the many physical parameters [50–52]. For example, recent studies have taken this approach to train an elastic network to perform an allosteric task [17,23,52,53]. Typically, networks are designed with minimal frustration, resulting in single-state allostery, but multi-state behavior has been empirically shown to emerge when rest lengths are allowed to vary [52,53]. The work presented here provides a clear rationale for these observations and predicts that by design with frustration, more efficient and varied allosteric functions may be achieved. It will be interesting to test these ideas using now emerging methods to prototype physical materials that embody the simulated networks [17,54,55]. Such experiments in macroscopic materials would provide a clear test of the ideas presented in this work even as we await practical methods for engineering allosteric proteins.

ACKNOWLEDGMENTS

We thank Kabir Husain, Riccardo Ravasio, Nachi Stern, Ayanna Matthews, and members of the Ranganathan and Rivoire Labs for helpful discussions. R.R. and O.R. are grateful for funding from the FACCTS program of the France Chicago Center. R.R. acknowledges support from NIH RO1GM12345 and RO1GM141697, the University of Chicago, and O.R. from ANR-17-CE30-0021.

APPENDIX: METHODS

To build a network, nodes are first placed on a triangular lattice (with spacing a) with their positions given by $\mathbf{R}_i^{\rm lat}$. An adjacency matrix is defined by nearest neighbors, $A_{ij} = 1$ if $||\mathbf{R}_i^{\rm lat} - \mathbf{R}_j^{\rm lat}|| = a$ and $A_{ij} = 0$, otherwise. The nodes are then displaced by a random uniform perturbation $\Delta\mathbf{R}_i \sim ([-a\epsilon, a\epsilon], [-a\epsilon, a\epsilon])$ to achieve a disordered structure $\mathbf{R}_i^{\rm dis} = \mathbf{R}_i^{\rm lat} + \Delta\mathbf{R}_i$. The dimensionless parameter ϵ controls the extent of the spatial disorder and is taken to be $\epsilon = 0.1$ unless stated otherwise. Each node is assigned a value $s_i \in \{1, 2, \ldots, Q\}$ giving the network a sequence $\mathbf{s} = (s_1, \ldots, s_N)$. Springs are connected between nodes with stiffness and rest lengths that depend on the sequence $\mathbf{s}$,

$$k_{ij} = A_{ij}K(s_i, s_j), \tag{A1}$$

$$\ell_{ij} = L(s_i, s_j) + \left|\left|\mathbf{R}_i^{\rm dis} - \mathbf{R}_j^{\rm dis}\right|\right|. \tag{A2}$$

The values of $K(s_i, s_j)$ and $L(s_i, s_j)$ are given by $Q \times Q$ symmetric interaction tables. Specifically, we consider all interactions in $K(s_i, s_j)$ to be stiff ($K = 1$) except for one randomly chosen entry that is soft ($K = 0.01$). The entries of $L(s_i, s_j)$ are drawn uniformly in $[-a\sigma, a\sigma]$, where σ is a dimensionless parameter controlling the disorder in the rest length deviations.

The full energy of a network with conformation $\mathbf{r} = (\mathbf{r}_1, \ldots, \mathbf{r}_N)$ sequence s is

$$\begin{aligned} U(\mathbf{r}, \mathbf{s}, \ell_{\text{act}}, \ell_{\text{allo}}) &= \frac{1}{2}\sum_{i>j} k_{ij}(\delta r_{ij} - \ell_{ij})^2 \\ &+ \frac{1}{2}k_{\text{rep}}\sum_{i>j}(1 - A_{ij})\Theta(\ell_{\text{rep}} - \delta r_{ij})(\delta r_{ij} - \ell_{\text{rep}})^2 \\ &+ \frac{1}{2}k_{\text{act}}(\delta r_{\text{act}} - \ell_{\text{act}})^2 + \frac{1}{2}k_{\text{allo}}(\delta r_{\text{allo}} - \ell_{\text{allo}})^2, \end{aligned} \tag{A3}$$

where $\mathbf{r}_i = (x_i, y_i)$ is the position of node i, and $\delta r_{ij} = ||\mathbf{r}_i - \mathbf{r}_j||$ is the distance between nodes i and j. To prevent nonphysical network collapse behaviors, the second term of (A3) defines a harmonic repulsion term between nonadjacent nodes. Θ is the Heaviside function so that $\Theta(\ell_{\text{rep}} - \delta r_{ij}) = 1$ if $\ell_{\text{rep}} > \delta r_{ij}$ and 0 otherwise. $\ell_{\text{rep}} = 0.7$ and $k_{\text{rep}} = 2$ are the distance cutoff and spring constant for the harmonic repulsion term, respectively. All networks simulated in this work have a size of $N = 49$ nodes.

The third and fourth terms in (A3) detail the contribution of ligand binding to the active and allosteric sites, respectively. The active site is chosen by picking two nodes on one side of the network that are previously unconnected and connecting them with a spring of rest length ℓ_{act} and stiffness k_{act}. δr_{act} measures the distance between these two nodes. The allosteric site is chosen in the same way but on the opposite side of the network. For this work, $k_{\text{act}} = k_{\text{allo}} = 1$. Importantly, ℓ_{act} and ℓ_{allo} are independent of $\mathbf{s}$.

The ground-state energy of the network is computed as

$$E(\mathbf{s}, \ell_{\text{act}}, \ell_{\text{allo}}) = \min_{\mathbf{r}} U(\mathbf{r}, \mathbf{s}, \ell_{\text{act}}, \ell_{\text{allo}}). \tag{A4}$$

Binding is modeled by changing the rest length of a binding site spring from one value representing the "solvent" ℓ_0 to another value representing the "ligand" ℓ_1 [colored bonds in Fig. 1(d)]. The binding cooperativity is computed as

$$\begin{aligned} \Delta\Delta E(\mathbf{s}) = &[E(\mathbf{s}, \ell_1, \ell_0) - E(\mathbf{s}, \ell_0, \ell_0)] \\ &- [E(\mathbf{s}, \ell_1, \ell_1) - E(\mathbf{s}, \ell_0, \ell_1)]. \end{aligned} \tag{A5}$$

Disorder enters in three forms in our model, the first from variability in the stiffnesses of the springs, which is enabled by the heterogeneity of the stiffness table $K(s_i, s_j)$. A second form enters through the spatial disorder in the base conformation $\mathbf{R}_j^{\text{dis}}$. In the case of $\epsilon > 0$ and $\sigma = 0$, the rest lengths reduce to the pairwise distances of the disordered lattice conformation $\ell_{ij} = ||\mathbf{R}_i^{\text{dis}} - \mathbf{R}_j^{\text{dis}}||$, which results in $U(\mathbf{s}, \mathbf{R}^{\text{dis}}) = 0$ (ignoring the effect of ligands). This implies that spatial disorder originating from $\epsilon > 0$ does not contribute to the frustration of the network. Finally, when $\sigma > 0$ the rest lengths can deviate from $||\mathbf{R}_i^{\text{dis}} - \mathbf{R}_j^{\text{dis}}||$ by an amount given by $L(s_i, s_j)$. Only the disorder in $L(s_i, s_j)$ gives rise to frustration in the network. The realizations of $K(s_i, s_j)$, $L(s_i, s_j)$, and $\mathbf{R}_i^{\text{dis}}$ are fixed during the evolution of network.

Estimating ground states. We estimate the ground state of a network with a version of the Basin-Hopping algorithm [24] where a genetic algorithm is used instead of a Monte Carlo algorithm to search conformation space. The algorithm (i) initializes a population of p copies of a network; (ii) perturbs ($\mathbf{r} \to \mathbf{r} + \boldsymbol{\eta}$, $\boldsymbol{\eta} \sim \mathcal{N}$) and relaxes each structure to a local minimum using the FIRE algorithm [56]; (iii) removes the $p/2$ structures with the highest relaxed energy; (iv) replicates each remaining structure; (v) repeats steps (ii)–(iv) $N_{\text{iterations}}$ times; (vi) outputs the structure with the lowest energy. We take $p = 20$ and $N_{\text{iterations}} = 100$, which we find to be sufficient for estimating the ground state (See Fig S1D of the SM [25] for an estimate of the error in the algorithm).

Definition of single-state and multi-state mechanisms. Let $\mathbf{R}_{00}$ and $\mathbf{R}_{11}$ denote the unbound ($\ell_{\text{act}} = \ell_{\text{allo}} = \ell_0$) and bound ($\ell_{\text{act}} = \ell_{\text{allo}} = \ell_1$) ground-state conformations, respectively,

$$\mathbf{R}_{ii} = \arg\min_{\mathbf{r}} U(\mathbf{r}, \mathbf{s}, \ell_i, \ell_i), \quad i = 0, 1. \tag{A6}$$

We define $\mathbf{R}'_{00}$ as the structure relaxed from $\mathbf{R}_{11}$ in the absence of a ligand ($\ell_{\text{act}} = \ell_{\text{allo}} = \ell_0$), and $\mathbf{R}'_{11}$ as the structure relaxed from $\mathbf{R}_{00}$ in the presence of the two ligands ($\ell_{\text{act}} = \ell_{\text{allo}} = \ell_1$). We say that the mechanism is single-state if $\mathbf{R}_{00} = \mathbf{R}'_{00}$ and $\mathbf{R}_{11} = \mathbf{R}'_{11}$, and multi-state otherwise.

Normal modes. Normal modes are computed by eigenvalue decomposition of the Hessian matrix H of the energy in the unbound state $\mathbf{R}_{00}$ when the active and allosteric site springs are removed, $k_{\text{act}} = k_{\text{allo}} = 0$. The components of H are

$$H_{ij} = \frac{\partial^2}{\partial x_i \partial x_j}\bar{U}\Big|_{\mathbf{r}=\mathbf{R}_{00}}, \tag{A7}$$

$$\bar{U} = U(\mathbf{r}, \mathbf{s}, \ell_0, \ell_0, k_{\text{act}} = 0, k_{\text{allo}} = 0). \tag{A8}$$

There are three zero modes ($\lambda_k = 0$) corresponding to global rotation $\mathbf{v}_{\text{rot}}$ and two global translations $\mathbf{v}_x$ and $\mathbf{v}_y$. The nonzero modes and their stiffnesses are denoted $\mathbf{v}_k$ and λ_k with $\lambda_i \leqslant \lambda_j$ if $i > j$, with $\mathbf{v}_1$ being the softest nonzero mode.

Allosteric displacement. Given $\mathbf{d} = \mathbf{R}_{11} - \mathbf{R}_{00}$, the allosteric displacement $\Delta\mathbf{r}$ is computed as

$$\Delta\mathbf{r} = \mathbf{d} - (\mathbf{d}\cdot\mathbf{v}_{\text{rot}})\mathbf{v}_{\text{rot}} - (\mathbf{d}\cdot\mathbf{v}_x)\mathbf{v}_x - (\mathbf{d}\cdot\mathbf{v}_y)\mathbf{v}_y. \tag{A9}$$

Overlaps. The overlap between a nonzero normal mode $\mathbf{v}_k$ and the allosteric displacement $\Delta\mathbf{r}$ is computed as

$$q_k = \left|\frac{\Delta\mathbf{r}}{\|\Delta\mathbf{r}\|}\cdot\mathbf{v}_k\right|. \tag{A10}$$

Conformational coordinate. The conformational coordinate y is defined by $\mathbf{r}(y) = \mathbf{R}_{00} + y\Delta\mathbf{r}$, and its associated energy is defined by $U(y) = U(\mathbf{r}(y), \mathbf{s})$.

Monte Carlo evolution. Networks are "evolved" using a Metropolis Monte Carlo method with an error-catching step due to the stochastic nature of the ground-state finding algorithm. A candidate sequence $\mathbf{s}'$, one mutation away from the current sequence $\mathbf{s}$, is accepted with probability $p = \exp([\Delta\Delta E(\mathbf{s}') - \Delta\Delta E(\mathbf{s})]/T_{\text{evo}})$, where $T_{\text{evo}} = 10^{-5}$. Following the acceptance or rejection step, the ground states of the current sequence, E_{00}, E_{10}, E_{01}, and E_{11}, are systematically reevaluated. If a lower ground-state energy is found, the current sequence is set to the previous sequence, and ground states are recomputed. This step prevents inaccurate ground-state estimates from causing fictitiously large values of $\Delta\Delta E$. Importantly, the error in the ground state finding algorithm does not affect the results as demonstrated by Fig. S2P of the SM [25]. The same results are found with more exhaustive sampling and thus lower rates of error.

Elastic network analogs of 1D model parameters. In the 1D model, k_m can be interpreted as the stiffness of the allosteric mode. To define an equivalent quantity for the 2D elastic network model, we consider $\tilde{k}_m = \lambda_1/\tilde{\lambda}$, where λ_1 is the stiffness of the softest nonzero mode and $\tilde{\lambda}$ is the median stiffness. In the 1D model, ℓ_m is related to the height of the barrier of $U(x)$ when $k_a = 0$: $\Delta U_{\text{barrier}} = k_m \ell_m^2/2$. To define an equivalent quantity for the 2D elastic network model, we consider $\tilde{\ell}_m = \delta^{-1}\sqrt{2\Delta U_{\text{barrier}}/\lambda_1}$ when $k_{\text{act}} = k_{\text{allo}} = 0$. We define $\Delta U_{\text{barrier}} = 0$ if $U(y)$ has one minimum [$U'(y) = 0, U''(y) > 0$]. If $U(y)$ has two minima, we take y_1 and y_2 to be their locations and y_3 to be the location of the maximum in between. We define $\Delta U_{\text{barrier}} = U(y_3) - \max(U(y_1), U(y_2))$. Scenarios with more than two minima are not observed.

[1] K. Gunasekaran, B. Ma, and R. Nussinov, Is allostery an intrinsic property of all dynamic proteins? Proteins Struct. Funct. Bioinf. **57**, 433 (2004).

[2] J. Monod, J.-P. Changeux, and F. Jacob, Allosteric proteins and cellular control systems, J. Mol. Biol. **6**, 306 (1963).

[3] J. Monod, J. Wyman, and J.-P. Changeux, On the nature of allosteric transitions: A plausible model, J. Mol. Biol. **12**, 88 (1965).

[4] D. E. Koshland, Jr., G. Nemethy, and D. Filmer, Comparison of experimental binding data and theoretical models in proteins containing subunits, Biochemistry **5**, 365 (1966).

[5] B. F. Volkman, D. Lipson, D. E. Wemmer, and D. Kern, Two-state allosteric behavior in a single-domain signaling protein, Science **291**, 2429 (2001).

[6] K. Kamata, M. Mitsuya, T. Nishimura, J.-i. Eiki, and Y. Nagata, Structural basis for allosteric regulation of the monomeric allosteric enzyme human glucokinase, Structure **12**, 429 (2004).

[7] L. T. May, K. Leach, P. M. Sexton, and A. Christopoulos, Allosteric modulation of G protein–coupled receptors, Annu. Rev. Pharmacol. Toxicol. **47**, 1 (2007).

[8] J.-P. Changeux and S. J. Edelstein, Allosteric mechanisms of signal transduction, Science **308**, 1424 (2005).

[9] K. Helmstaedt, S. Krappmann, and G. H. Braus, Allosteric regulation of catalytic activity: Escherichia coli aspartate transcarbamoylase versus yeast chorismate mutase, Microbiol. Mol. Biol. Rev. **65**, 404 (2001).

[10] A. Szabo and M. Karplus, A mathematical model for structure-function relations in hemoglobin, J. Mol. Biol. **72**, 163 (1972).

[11] J. Herzfeld and H. E. Stanley, A general approach to cooperativity and its application to the oxygen equilibrium of hemoglobin and its effectors, J. Mol. Biol. **82**, 231 (1974).

[12] Q. Cui and M. Karplus, Allostery and cooperativity revisited, Protein Sci. **17**, 1295 (2008).

[13] J. N. Onuchic and P. G. Wolynes, Theory of protein folding, Curr. Opin. Struct. Biol. **14**, 70 (2004).

[14] M. Hemery and O. Rivoire, Evolution of sparsity and modularity in a model of protein allostery, Phys. Rev. E **91**, 042704 (2015).

[15] L. Yan, R. Ravasio, C. Brito, and M. Wyart, Architecture and coevolution of allosteric materials, Proc. Natl. Acad. Sci. (USA) **114**, 2526 (2017).

[16] T. Tlusty, A. Libchaber, and J.-P. Eckmann, Physical model of the genotype-to-phenotype map of proteins, Phys. Rev. X **7**, 021037 (2017).

[17] J. W. Rocks, N. Pashine, I. Bischofberger, C. P. Goodrich, A. J. Liu, and S. R. Nagel, Designing allostery-inspired response in mechanical networks, Proc. Natl. Acad. Sci. (USA) **114**, 2520 (2017).

[18] H. Flechsig, Design of elastic networks with evolutionary optimized long-range communication as mechanical models of allosteric proteins, Biophys. J. **113**, 558 (2017).

[19] L. Yan, R. Ravasio, C. Brito, and M. Wyart, Principles for optimal cooperativity in allosteric materials, Biophys. J. **114**, 2787 (2018).

[20] S. Dutta, J.-P. Eckmann, A. Libchaber, and T. Tlusty, Green function of correlated genes in a minimal mechanical model of protein evolution, Proc. Natl. Acad. Sci. (USA) **115**, E4559 (2018).

[21] R. Ravasio, S. M. Flatt, L. Yan, S. Zamuner, C. Brito, and M. Wyart, Mechanics of allostery: Contrasting the induced fit and population shift scenarios, Biophys. J. **117**, 1954 (2019).

[22] O. Rivoire, Parsimonious evolutionary scenario for the origin of allostery and coevolution patterns in proteins, Phys. Rev. E **100**, 032411 (2019).

[23] J. W. Rocks, H. Ronellenfitsch, A. J. Liu, S. R. Nagel, and E. Katifori, Limits of multifunctionality in tunable networks, Proc. Natl. Acad. Sci. (USA) **116**, 2506 (2019).

[24] D. J. Wales and J. P. Doye, Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms, J. Phys. Chem. A **101**, 5111 (1997).

[25] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PRXLife.1.023004 for additional information.

[26] I. Bahar, A. R. Atilgan, and B. Erman, Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential, Folding Des. **2**, 173 (1997).

[27] A. R. Atilgan, S. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin, and I. Bahar, Anisotropy of fluctuation dynamics of proteins with an elastic network model, Biophys. J. **80**, 505 (2001).

[28] B. Schuler, E. A. Lipman, and W. A. Eaton, Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy, Nature (London) **419**, 743 (2002).

[29] L. Ma and Q. Cui, Activation mechanism of a signaling protein at atomic resolution from advanced computations, J. Am. Chem. Soc. **129**, 10261 (2007).

[30] Y. Xu, J. P. Colletier, H. Jiang, I. Silman, J. L. Sussman, and M. Weik, Induced-fit or preexisting equilibrium dynamics? lessons from protein crystallography and md simulations on acetylcholinesterase and implications for structure-based drug design, Protein Sci. **17**, 601 (2008).

[31] A. Cooper and D. Dryden, Allostery without conformational change, Eur. Biophys. J. **11**, 103 (1984).

[32] T. C. McLeish, T. Rodgers, and M. R. Wilson, Allostery without conformation change: Modelling protein dynamics at multiple scales, Phys. Biol. **10**, 056004 (2013).

[33] N. Popovych, S. Sun, R. H. Ebright, and C. G. Kalodimos, Dynamically driven protein allostery, Nat. Struct. Mol. Biol. **13**, 831 (2006).

[34] D. M. Mitrea and R. W. Kriwacki, Regulated unfolding of proteins in signaling, FEBS Lett. **587**, 1081 (2013).

[35] V. J. Hilser and E. B. Thompson, Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins, Proc. Natl. Acad. Sci. (USA) **104**, 8311 (2007).

[36] D. U. Ferreiro, J. A. Hegler, E. A. Komives, and P. G. Wolynes, Localizing frustration in native proteins and protein assemblies, Proc. Natl. Acad. Sci. (USA) **104**, 19819 (2007).

[37] D. U. Ferreiro, J. A. Hegler, E. A. Komives, and P. G. Wolynes, On the role of frustration in the energy landscapes of allosteric proteins, Proc. Natl. Acad. Sci. (USA) **108**, 3499 (2011).

[38] S. Gianni and P. Jemth, Allostery frustrates the experimentalist, J. Mol. Biol. **435**, 167934 (2022).

[39] H. Frauenfelder, F. Parak, and R. D. Young, Conformational substates in proteins, Annu. Rev. Biophys. Biophys. Chem. **17**, 451 (1988).

[40] L. C. James, P. Roversi, and D. S. Tawfik, Antibody multispecificity mediated by conformational diversity, Science **299**, 1362 (2003).

[41] Y. Bai, T. R. Sosnick, L. Mayne, and S. W. Englander, Protein folding intermediates: Native-state hydrogen exchange, Science **269**, 192 (1995).

[42] F. Chiti and C. M. Dobson, Protein misfolding, functional amyloid, and human disease, Annu. Rev. Biochem. **75**, 333 (2006).

[43] W. Zheng, B. R. Brooks, and D. Thirumalai, Low-frequency normal modes that describe allosteric transitions in biological nanomachines are robust to sequence variations, Proc. Natl. Acad. Sci. (USA) **103**, 7664 (2006).

[44] M. Gur, E. Zomot, and I. Bahar, Global motions exhibited by proteins in micro-to milliseconds simulations concur with anisotropic network model predictions, J. Chem. Phys. **139**, 121912 (2013).

[45] O. Rivoire, K. A. Reynolds, and R. Ranganathan, Evolution-based functional decomposition of proteins, PLoS Comput. Biol. **12**, e1004817 (2016).

[46] S. W. Lockless and R. Ranganathan, Evolutionarily conserved pathways of energetic connectivity in protein families, Science **286**, 295 (1999).

[47] G. M. Süel, S. W. Lockless, M. A. Wall, and R. Ranganathan, Evolutionarily conserved networks of residues mediate allosteric communication in proteins, Nat. Struct. Biol. **10**, 59 (2003).

[48] N. Halabi, O. Rivoire, S. Leibler, and R. Ranganathan, Protein sectors: evolutionary units of three-dimensional structure, Cell **138**, 774 (2009).

[49] M. Stern and A. Murugan, Learning without neurons in physical systems, Annu. Rev. Condens. Matter Phys. **14**, 417 (2023).

[50] M. Stern, M. B. Pinson, and A. Murugan, Continual learning of multiple memories in mechanical networks, Phys. Rev. X **10**, 031044 (2020).

[51] M. Stern, C. Arinze, L. Perez, S. E. Palmer, and A. Murugan, Supervised learning through physical changes in a mechanical system, Proc. Natl. Acad. Sci. (USA) **117**, 14843 (2020).

[52] M. Stern, D. Hexner, J. W. Rocks, and A. J. Liu, Supervised learning in physical networks: From machine learning to learning machines, Phys. Rev. X **11**, 021045 (2021).

[53] D. Hexner, A. J. Liu, and S. R. Nagel, Periodic training of creeping solids, Proc. Natl. Acad. Sci. (USA) **117**, 31690 (2020).

[54] N. Pashine, D. Hexner, A. J. Liu, and S. R. Nagel, Directed aging, memory, and nature's greed, Sci. Adv. **5**, eaax4215 (2019).

[55] N. Pashine, Local rules for fabricating allosteric networks, Phys. Rev. Mater. **5**, 065607 (2021).

[56] E. Bitzek, P. Koskinen, F. Gähler, M. Moseler, and P. Gumbsch, Structural relaxation made simple, Phys. Rev. Lett. **97**, 170201 (2006).