Biophysical Journal
Article

Biophysical Society

Design of Elastic Networks with Evolutionary Optimized Long-Range Communication as Mechanical Models of Allosteric Proteins

Holger Flechsig[1,*]
[1]Department of Mathematical and Life Sciences, Graduate School of Science, Hiroshima University, Higashi-Hiroshima, Hiroshima, Japan

ABSTRACT Allosteric effects often underlie the activity of proteins, and elucidating generic design aspects and functional principles unique to allosteric phenomena represent a major challenge. Here an approach consisting of the in silico design of synthetic structures, which, as the principal element of allostery, encode dynamical long-range coupling among two sites, is presented. The structures are represented by elastic networks, similar to coarse-grained models of real proteins. A strategy of evolutionary optimization was implemented to iteratively improve allosteric coupling. In the designed structures, allosteric interactions were analyzed in terms of strain propagation, and simple pathways that emerged during evolution were identified as signatures through which long-range communication was established. Moreover, robustness of allosteric performance with respect to mutations was demonstrated. As it turned out, the designed prototype structures reveal dynamical properties resembling those found in real allosteric proteins. Hence, they may serve as toy models of complex allosteric systems, such as proteins. Application of the developed modeling scheme to the allosteric transition in the myosin V molecular motor was also demonstrated.

INTRODUCTION

The functional activity of proteins and its precise regulation often relies on allosteric coupling between different functional regions within the macromolecular structure. According to the mechanical perspective, allosteric communication originates from structural changes mediated by a network of physically interacting residues (1). Much resembling the occurrence of a protein quake, local conformational motions initiated, e.g., upon ligand binding to one specific site, propagate themselves across the protein structure to spatially remote regions eventually generating a functional change in the conformation of another site.

Experiments have indeed evidenced the existence of communication networks in proteins, which are formed by only a set of amino acids and constitute allosteric pathways, physically linking remote binding sites (2,3). Various computational strategies aimed at predicting such pathways have been developed, including structure-based network analysis (4–6), stochastic Markov modeling (7,8), and sequence-based statistical methods (9–12). Understanding of allosteric communication at the molecular level has also been widely addressed at atomistic resolution in molecular dynamics (MD) simulations (13–19).

To circumvent the heavy computational burden present in MD simulations, collective conformational motions and allosteric transitions in proteins have been, to a vast extent, investigated at the coarse-grained level using elastic-network models and the analysis of normal modes (20–26). Despite their approximate nature, such simplified descriptions have significantly contributed to understanding the mechanistic underpinnings of allostery in proteins (27,28).

Although, in the majority of cases, the important question is considered to be how allostery works for a particular protein given its specific structure, there may also be more fundamental and general questions to be addressed, such as what generic design and functional principles are requisite to make allostery work, and which dynamical properties are unique to allostery. One possibility to approach such aspects would consist of a systematic screening of the biophysical properties among allosteric proteins with the available structural data, which has been sporadically attempted earlier (6,29).

Here, in an attempt to address this subject, we present an alternative approach that is motivated by the idea to engineer, in silico, artificial spatial structures with dynamical properties resembling those found in real proteins. In

Submitted March 6, 2017, and accepted for publication June 14, 2017.

*Correspondence: holgerflechsig@hiroshima-u.ac.jp

Editor: Monika Fuxreiter.

http://dx.doi.org/10.1016/j.bpj.2017.06.043

previous works, such an approach was employed to design structures of elastic networks that can operate as a model molecular machine (30–32) or swimmer (33). Moreover, we recently used this machine to construct a model motor which, in its function, mimics the myosin protein motor (34).

In this article, we aim to establish a generalized structural model of an allosteric system. To this end, we design, using evolutionary optimization, elastic-network structures that, as the principal element of allostery, encode long-range coupling among two spatially remote local regions. We first explain how the strategy of iterative evolution was developed and then applied to stepwise improve starting from a random elastic network, the allosteric response in the emerging structures. In the designed structures, allosteric communication was then analyzed in terms of the propagation of strain, and its spatial distribution was used to identify pathways through which remote interactions are established. Moreover, the effect of mutations was demonstrated and the robustness of allosteric performance in the designed structures examined. Finally, application of the modeling scheme to allosteric communication in the myosin V molecular motor was demonstrated and commonalities with the designed model systems were exemplified.

METHODS

Construction of random elastic network

The initial two-domain elastic network was constructed by first randomly folding two chains of linked beads, then bringing them into contact and finally determining the network connectivity. One chain consisted of 100 identical beads each and its folded form was constructed as follows. After fixing the position of the first bead, each next bead was placed at random around the position of the previous bead, with the following restrictions: 1) the distance to the preceding bead had to lie within the interval between $l_{\min}$ and $l_{\max}$, 2) the new bead had to be separated from each previous bead by at least the distance $l_{\min}$, and 3) the distance from the new bead to the geometric center of all previous beads should not exceed the threshold $r_{\max}$. In the simulations, the prescribed values $l_{\min} = 4.0$ Å, $l_{\max} = 5.0$ Å, and $r_{\max} = 20.0$ Å were used to generate a compactly folded backbone chain. After constructing two such chains, they were merged in such a way that they came into tight contact but still did not overlap. Positions of the 200 beads in the initial two-domain random structure are denoted by $\vec{R}_i^{(0)}$ and their spatial coordinates are provided as Supporting Materials and Methods. To complete the network construction, we have checked all distances $d_{ij}^{(0)} = \left|\vec{R}_i^{(0)} - \vec{R}_j^{(0)}\right|$ between beads i and j and introduced an elastic spring between those pairs of beads for which the distance $d_{ij}^{(0)}$ was below a prescribed interaction radius of $r_{\text{int}} = 9$ Å. With this choice, the initial two-domain elastic network had 1467 links. It should be noted that in this model, all length units are, in principle, arbitrary. When data from real protein structures is used, the distances between amino acids have the scale of Ångstroms. Hence, throughout the article, this notion is adopted.

Elastic conformational dynamics

The total elastic energy U of the network is the sum of contributions of all elastic links, i.e., $U = (1/2)\sum_{i<j}^{N} \kappa \Gamma_{ij} (d_{ij} - d_{ij}^{(0)})^2$. Here, $N = 200$ is the number of beads; κ is the spring stiffness constant (equal for all springs); $d_{ij} = \left|\vec{R}_i - \vec{R}_j\right|$ is the actual length of a spring connecting beads i and j in some deformed network conformation, with $\vec{R}_i$ being the actual position vector of bead i; and $d_{ij}^{(0)}$ is the corresponding natural spring length. Coefficients Γ_{ij} have value 1 if beads i and j are connected by a spring (i.e., when $d_{ij}^{(0)} < r_{\text{int}}$), and equal 0, otherwise. The dynamics of the elastic network is governed by Newton's equation of motion in the overdamped limit, where the velocity of each network bead is proportional to the forces applied to it. This limit is valid in the low Reynolds number regime of proteins at which inertial movements can be neglected. Hence, the equation for bead i was

$$\gamma \frac{d}{dt}\vec{R}_i = -\frac{\partial}{\partial \vec{R}_i} U + \vec{f}_i = -\kappa \sum_{j}^{N} \Gamma_{ij} \frac{d_{ij} - d_{ij}^{(0)}}{d_{ij}} \left(\vec{R}_i - \vec{R}_j\right) + \vec{f}_i. \tag{1}$$

On the left-hand side, γ is the friction coefficient assumed to be equal for all beads. On the right side are the elastic forces exerted by network springs that are connected to bead i; they depend only on the change in the distance between two connected beads. Additionally, an external force $\vec{f}_i$ could be applied to bead i, which was used in probing allostery (see next section). Note that in the above equations the dependencies on γ and κ can be removed using the rescaled dimensionless time $(\kappa/\gamma)t$. To obtain the positions of network beads, and hence the conformation of the network at any time moment, the set of equations of motion was numerically integrated. In the simulations, a first-order scheme with a time step of 0.1 was employed. It should be noted that the same relaxational elastic-network model was previously applied in several studies of protein conformational dynamics (e.g., (30,35–37)).

Pocket sites and probing of allostery

In the network model, the allosteric site and the regulates site were, for simplicity, each represented by two beads. Those beads have been selected in such a way that the two pockets were sufficiently remote from each other. At the same time, the two beads forming one pocket should be adjacent, but not directly connected by an elastic spring. In the constructed two-domain network, the allosteric pocket was defined by beads with indices 38 and 81, belonging to the first domain; and the regulated pocket was defined by beads with indices 149 and 189, belonging to the second domain.

To probe allosteric communication between the two pockets, a simple force-probe scheme was implemented. Conformational dynamics in the allosteric site was initiated through the application of additional forces, and the subsequent response generated in regulated site was probed by evaluating structural changes there. In the simulations, pair forces have been applied to the beads of the allosteric pocket. The forces were always acting along the direction given by the actual positions of the two pocket beads. Specifically, the force applied to pocket bead 38 was $\vec{f}_{38} = 0.5 \cdot \vec{u}$, with the unit vector $\vec{u} = (\vec{R}_{81} - \vec{R}_{38})/\left|\vec{R}_{81} - \vec{R}_{38}\right|$. The same force, but with the different sign, was acting on the second pocket bead, i.e., $\vec{f}_{81} = -0.5 \cdot \vec{f}_{38}$. Such pair forces would induce only internal network deformations and result in a decrease in the distance between the two pocket beads, thus leading to closing of the allosteric pocket. Because this situation is apparently equivalent to assuming that an additional network bead became bound to the center of the two pocket beads, where it generated attractive forces between itself and each pocket bead, we can also say that the chosen force scheme mimicked binding of a fictitious ligand bead to the allosteric pocket. The action of the additional forces generated deformations of the network, which were first localized in the vicinity of the allosteric pocket, but which gradually spread over the entire network structure. The corresponding process of conformational relaxation was followed by integrating the equations of motion until a steady state of the elastic network (in which all motions were terminated) was reached (at final time T, see Supporting Materials and Methods). In the resulting conformation of the network, the effect of allosteric coupling was examined by

evaluating the distance between the beads corresponding to the regulated site. This distance $d_{149,189}(T) = \left|\vec{R}_{149}(T) - \vec{R}_{189}(T)\right| =: A$ is termed the "allosteric parameter".

Evolutionary optimization

To design networks with perfected allosteric communication, a process of evolutionary optimization, consisting of mutations followed by selection, was applied iteratively, starting from the random network. In particular, the following sequence was carried out. First the allosteric response of the elastic network before the mutation was determined and the allosteric parameter A stored. Then a single structural mutation was performed by randomly selecting one network bead (except for one of the four pocket beads) and changing its equilibrium position. The new equilibrium position was chosen to be randomly oriented within a sphere of radius 2.0 Å around the old equilibrium position. To preserve the backbone chain of each domain, it was additionally required that, after the mutation, the distance between the mutated bead and its left and right neighbor in the chain would still lie within the interval between $l_{\min}$ and $l_{\max}$. Furthermore, distances from the mutated bead to all other network beads should not be smaller than $l_{\min}$. After the mutation, the network connectivity Γ_{ij} was updated by reexamining distances between all beads and the mutated bead; only those pairs that were separated by a distance less than r_{int} were linked by a spring. After a mutation, the elastic network may possess internal free rotations originating from loosely coupled network parts. They can lead to local movements free of energetic cost, which was to be prevented. Therefore, when the number of nonzero eigenvalues in the spectrum of the elastic network was smaller than $3N-6$ (indicating the occurrence of internal rotation zero modes), the mutation was rejected. Once a mutation that fulfilled all criteria was found, probing of allostery in the new elastic network proceeded as described in the previous section, the allosteric parameter after the mutation A^{mut} was determined, and the mutation was evaluated by comparing the allosteric parameter before the mutation A with that after the mutation A^{mut}. Only mutations that were favorable, i.e., those that improved the allosteric response in the network, were selected. Two situations were distinguished. In the evolution process where symmetric coupling between the allosteric and regulated pocket was to be optimized, a mutation was accepted only if $A^{\text{mut}} < A$, and rejected otherwise. In the second independent evolutionary process corresponding to the asymmetric situation, the acceptance criteria for the mutations was $A^{\text{mut}} > A$. As a termination condition for the two evolution processes, we imposed $(A - A^{\text{random}}) < -2.0$ Å for the design of the network with symmetric coupling and $(A - A^{\text{random}}) > 2.0$ Å for the design of the network with asymmetric coupling. A^{random} denotes the allosteric parameter of the initial random network. During both design processes, the improvement of allosteric response in the evolving networks was recorded (see Fig. S1).

Strain propagation and pathways

In the initial random network, the two designed networks, and the selected mutant networks, the propagation of strain after ligand binding to the allosteric pocket was monitored. The strain of an elastic link connecting beads i and j was defined as $s_{ij}(t) = d_{ij}(t) - d_{ij}^{(0)}$. In the employed model, conformational changes corresponded to relaxation processes of the elastic network structure (see Eq. 1). Therefore, the energy injected locally at the allosteric site, as a consequence of ligand binding there, would not only be converted into deformations of elastic bead connections but would also dissipate. In particular, deformations of elastic springs become significantly damped the farther away they are located from the pocket. Because we still wanted to discuss the anisotropy of strain distribution in the network, a method in which the strain was normalized with respect to the distance from the allosteric site (in terms of the minimal path) was employed. Details are described in the Supporting Materials and Methods. For the visualization of strain, a link was colored blue (if $s_{ij} > 0$) or red ($s_{ij} < 0$) and the width corresponded to $\left|\tilde{s}_{ij}\right|$ (superscript $\sim$ refers to the normalized strain). To determine the communication chains shown in Fig. 5 the maximum absolute strain of each link during the simulation was stored and only those links whose normalized strain exceeded a prescribed threshold $\tilde{s}_t$ were considered (for details see Supporting Materials and Methods). For both designed networks, a threshold value of $\tilde{s}_t = 60\%$ was imposed. To obtain the communication skeleton in the 52-196-197 double mutant of the designed network with symmetric cooperativity (shown in Fig. 6 *B*), a threshold of $\tilde{s}_t = 45\%$ was used. For the 52-185 mutant of the designed network with asymmetric cooperativity (see Fig. 6 *b*), a threshold of $\tilde{s}_t = 25\%$ was used.

Robustness

All mutations performed are listed in Table S1. For each mutant network, a robustness coefficient was computed as the ratio of the change in the pocket size of the regulated pocket in the considered mutant network and the pocket size change in the wild-type network, i.e., $(\tilde{d}_{149,189}^{(0)} - \tilde{d}_{149,189}^{\text{final}})/(d_{149,189}^{(0)} - d_{149,189}^{\text{final}})$. Here, the superscript "$\sim$" denotes distances in the mutant network and superscript "final" denotes distances in the corresponding steady state of the elastic network after ligand binding to the allosteric site (i.e., at time T).

Application to allostery in the myosin V molecular motor

Details of the elastic network construction, modeling of ATP-binding, and quantification of allosteric coupling in the myosin V motor domain can be found in the Supporting Materials and Methods.

RESULTS

Random elastic network and evolutionary optimization

Our structural model of an allosteric system was based on evolutionary optimization starting from a random elastic network. The initial random network was constructed as a two-domain structure. It was obtained by randomly folding two polymer chains independently, each consisting of 100 elastically linked identical beads and representing one domain, and then merging them such that they form a common domain interface. In the model, physical interactions between all beads (beyond those acting between neighboring beads in each chain) were introduced by connecting those two beads by an elastic link that had a separation smaller than a prescribed interaction radius. Thus the complete elastic network was obtained. Details of the construction are described in the Methods.

When elastic networks are constructed based on the structures of real proteins, the beads correspond to atoms or typically represent entire amino acid residues, and elastic links between them empirically mimic the effective interactions between them (38,39).

Here, we proceed with the constructed random network of two compactly folded coupled domains shown in Fig. 1. In each network domain a pocket site, for simplicity formed by two beads only, was chosen such that a sufficient separation of the two sites was ensured (see Methods).

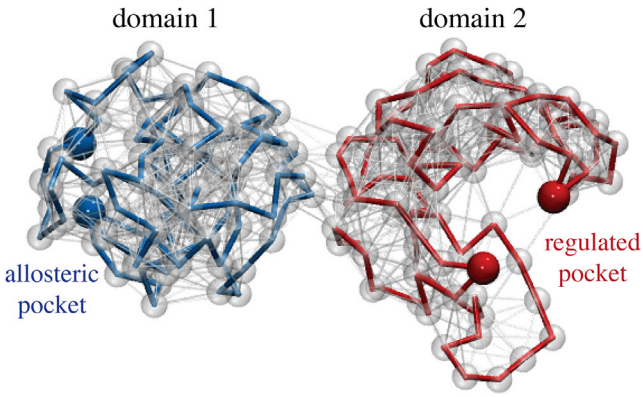


FIGURE 1 Random elastic-network structure. Given here is the initial two-domain network consisting of 200 beads connected by elastic links. The randomly folded chains, each forming one domain, are shown as blue (respectively, *red*) traces of thick bonds. Other links are displayed as thin, in gray. In each domain the two beads representing the remotely placed pockets are highlighted as colored beads. To see this figure in color, go online.

One site represented the allosteric pocket whereas the other mimicked the regulated pocket, respectively. As shown in Fig. 1, the pockets were located opposite to each other near the surface of the respective domain.

With this setup, the ability of the elastic network to conduct allosteric communication between the pockets was examined using a simple force-probe scheme, in which conformational dynamics in the allosteric site was initiated through local binding of an additional ligand bead, and the subsequent response generated in the regulated site has been detected. In particular, attractive pair forces that were acting on the pocket beads of the allosteric site have been applied to mimic binding of the fictitious ligand bead to its center. The dynamics of protein elastic-networks consists of processes of overdamped relaxation motions (see Methods). Hence, as a result of such forces, all beads underwent coupled relaxation motions, bringing the elastic network from its original conformation (without the forces) to a deformed steady state of the network, with the ligand bead tightly bound to the allosteric site (i.e., with the forces applied for a long time). During this process, the additional forces tend to close this site, i.e., its two beads move toward each other, and first the elastic links localized in their vicinity became deformed. Eventually deformations propagated through the entire network structure until a final steady network conformation was reached. In the simulations, the conformational motions inside the network were followed by numerically integrating the equations of motion for all network beads (see Methods). The response generated inside the regulated pocket in the second domain was quantified in terms of distance changes between its two corresponding beads.

As we found, the random network structure did not reveal any allosteric coupling between the two sites. Whereas the allosteric site closed upon ligand binding, with the distance between pocket beads changing by ~ -4 Å, no response was detected at the regulated site, with a change of $<10^{-4}$ Å (for the choice of lengths units, see Methods). Proceeding with this observation, an algorithm of evolutionary optimization aimed to design networks with pronounced allosteric communication, was established. The optimization process consisted of sequences of mutations followed by selection, which were applied iteratively starting from the initial random network. In each evolution step, the following sequence was carried out: 1) mutation: a single structural mutation was performed by randomly selecting one network bead (excluding one of the four pocket beads) and changing its equilibrium position, which corresponded to an alteration of elastic connections in the vicinity of the mutation spot; 2) probing of allostery: the force-probe scheme was applied in a simulation of the new mutant network structure and its conformation with the ligand bead bound to the allosteric site was obtained; and 3) selection: the allosteric response in the regulated pocket of the network before and after the mutation was compared and the mutation was scored favorable, and was accepted if the mutant network showed improved allosteric coupling between the two sites, otherwise it was rejected. The optimization procedure was iteratively applied and terminated, once a network structure with a prescribed level of sufficient allosteric coupling was obtained. Details of the implementation are described in the Methods.

In this work two prototype examples of elastic network structures with allosteric coupling were designed and analyzed. In two independent simulations of evolutionary optimization, starting both times with the same initial random network, two different network structures with optimized allostery have been designed. The first one was designed under the side condition that allosteric coupling between the remote sites had symmetric cooperativity, i.e., binding of the ligand bead to the allosteric site and the concomitant closing of that pocket resulted in consequent closing of the regulated site. In the second case, the requirement was vice versa; allosteric communication was optimized under the premise that ligand-induced closing of the allosteric pocket triggered opening in the regulated site, i.e., asymmetric cooperative coupling was realized. In both cases, as a result of several hundreds of mutations that were needed during evolution to improve allosteric coupling, the networks underwent significant structural remodeling and concomitant rewiring of elastic interactions between their beads. In Fig. 2, snapshots of network architectures along the processes of evolution are displayed. For both designed networks, it could be observed that besides mutation-induced changes taking place inside the individual domains, their common interface became significantly remodeled. As evolution progressed, the allosteric response generated inside the regulated pocket upon ligand binding to the allosteric pocket was gradually enhanced during both design processes, as can be seen by comparing corresponding snapshots of network conformations in Fig. 2.

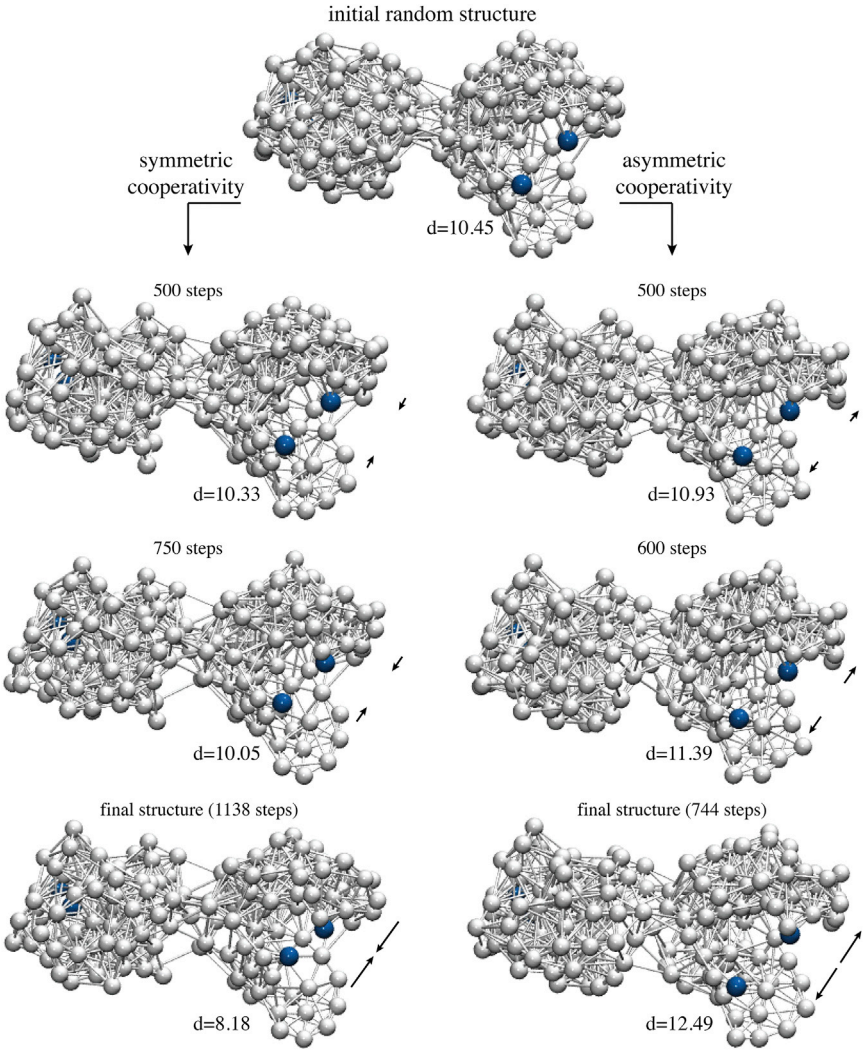


FIGURE 2 Evolutionary optimization. Given here are snapshots of network conformations taken at different steps of accepted mutations along the processes of evolutionary optimization toward designing the two prototype allosteric networks. The initial random network structure is shown in the top row. Each depicted network conformation corresponds to the steady state with the allosteric pocket being closed after ligand binding and the allosteric response in the regulated pocket becoming progressively improved. In the design process of the network with symmetric cooperativity, the regulated pocket increasingly closed as evolution proceeds, whereas during the design of the network with asymmetric cooperativity, its ability to open emerged (indicated by *black arrows*). Pocket beads are highlighted in blue color and the size of the regulated pocket in Å is given for all snapshots. Mutation-induced structural changes and concomitant remodeling of elastic interactions can be seen by comparing snapshots. Final structures of designed networks are shown in the bottom row. To see this figure in color, go online.

In the next paragraphs the two prototype designed network structures are presented and their dynamical properties analyzed.

Designed prototype networks

The designed networks are shown in Fig. 2 (*bottom panel*). Both networks retained the initial two-domain architecture after the evolution; however, their structures were clearly different from the initial random network. In both designed networks a pronounced domain interface with many interdomain links had emerged; such interface was rather sparse in the initial random network (see Fig. 2). Remodeling of that region under the process of evolution apparently points toward the importance of the pattern of elastic connections in the domain contact region for allosteric communication in our model systems. Those aspects will be further discussed in the next section.

In Fig. 3, *A* and *a*, the two designed networks are shown, each in their equilibrium conformation and in the respective steady conformations, with the ligand bound to the allosteric pocket and the allosteric response having been triggered in the remote regulated sites. Traces showing changes in the size of both pockets upon binding of the ligand bead to the allosteric pockets are also provided (see Fig. 3, *B* and *b*). In both designed networks, the allosteric pocket closed rapidly when the ligand bead bound there, with the size changing by -3.8 Å in both cases. In striking contrast, motions inside the regulated pocket were much slower, when changes in its size only gradually set in and the pocket smoothly approached its closed conformation in the network with symmetric cooperativity (change by -2.3 Å), or open conformation in the network with asymmetric cooperativity, respectively (change by $+2.0$ Å). Movies visualizing conformational motions during the allosteric transition in the designed networks are provided as Movies S2 and S3.

Comparing the ligand-bound network conformation with the initial structure reveals that, in both designed networks, the conformational changes underlying allosteric dynamics

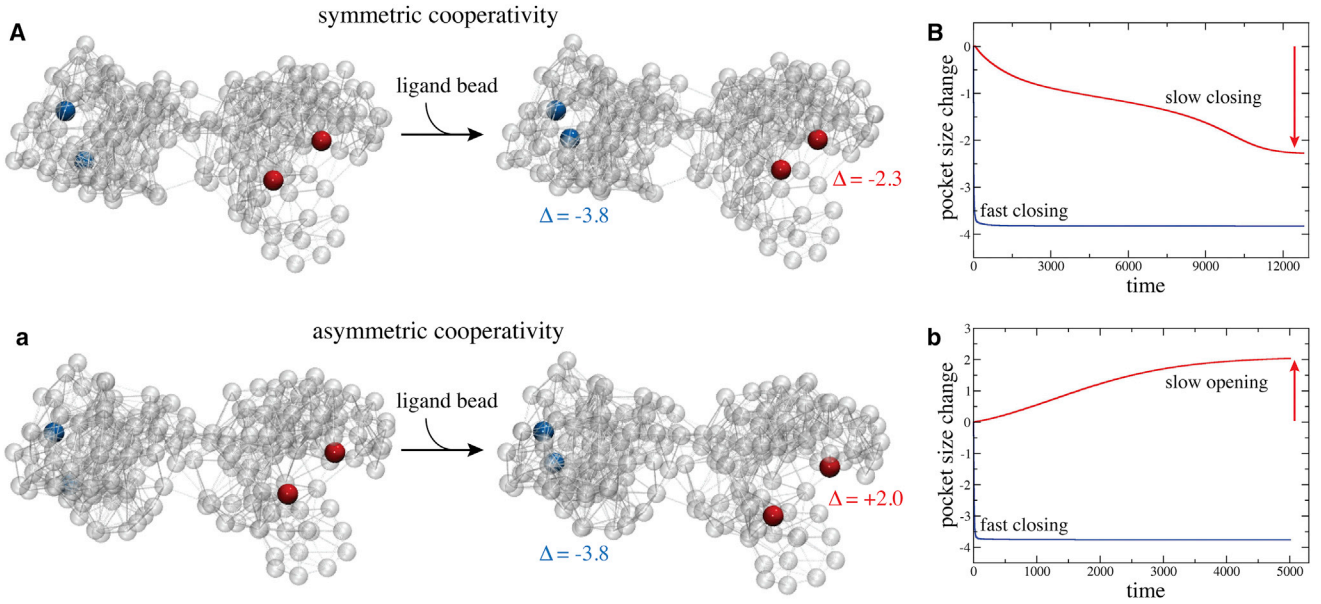


FIGURE 3 Allosteric coupling in designed network structures. (*A* and *a*) For both designed networks, the initial structure and the final structure with the ligand bound to the allosteric pocket are shown. Elastic connections between network beads are shown as thin transparent lines and pocket beads are colored blue (allosteric pocket) and red (regulated pocket), respectively. In both designed networks the initial and respective final structure is very similar (see text), but the large-amplitude motions of pocket beads evidencing intrastructure allosteric communication are apparent. In the network with symmetric cooperativity, ligand-induced closing of the allosteric pocket triggers allosteric closing of the regulated pocket, whereas, in the network with asymmetric cooperativity, the allosteric response consisted of opening of the regulated site. (*B* and *b*) Given here are traces showing changes in the size of both pockets upon binding of the ligand bead to the allosteric pockets (*blue* color for allosteric the pocket, *red* for the regulated pocket). To see this figure in color, go online.

did not involve any large-scale structural rearrangements but were rather governed by small-amplitude subtle motions. In fact, initial and ligand-bound structures compared by root mean squared deviations of only 0.6 Å in the symmetric case and 0.46 Å in the asymmetric case, respectively. Although local motions in the two pockets were indeed pronounced, the communication between them must therefore have resulted from cascades of small-amplitude displacements of network beads located in the regions connecting both pockets.

By the application of evolutionary optimization to the random network, we could successfully design special networks whose two-domain structure encodes enhanced allosteric interactions between two remote pockets. All three network structures are provided as Supporting Materials and Methods. Next, to unravel signatures that may underlie allosteric communication in the two prototype networks, we have analyzed the conformational dynamics in terms of mechanical strain propagating through the network structures.

Propagation of strain and communication pathways

Apparently, allosteric interactions in our model systems result from conformational changes propagating from the allosteric site across the domain interface to the regulated site. Therefore, in both designed structures, the elastic strain inside the networks was computed during the entire simulation after ligand binding to the allosteric site until the steady state of the respective network was reached. The strain inside a network was determined in terms of the deformation of elastic links connecting the beads, i.e., deviations from their natural lengths in the initial ligand-free network. Details of the computation are found in the Methods and Supporting Materials and Methods. To conveniently visualize strain propagation, the network links are displayed bicolored, to distinguish stretching or compression, and their width is proportional to the magnitude of the respective deformation. To compare to the designed networks, the strain propagation was also followed in the initial random network, which was allosterically inactive. Snapshots of the strain distribution in all three networks taken at different time moments during the simulation are shown in Fig. 4. Movies visualizing conformational motions and the propagation of strain are provided as Movies S1, S2, and S3.

In all three networks the strain was first localized in the vicinity of the allosteric pocket where the ligand was bound. After that, interactions between other beads also rapidly set in and the majority of links in domain 1 became deformed. In the random network, the distribution of strain in domain 1 was rather homogeneous; most links were stretched or compressed at similar levels and there was no distinctive feature present. In this network, no significant deformations of links located in the central domain interface region could be detected, and therefore spreading of strain into the second

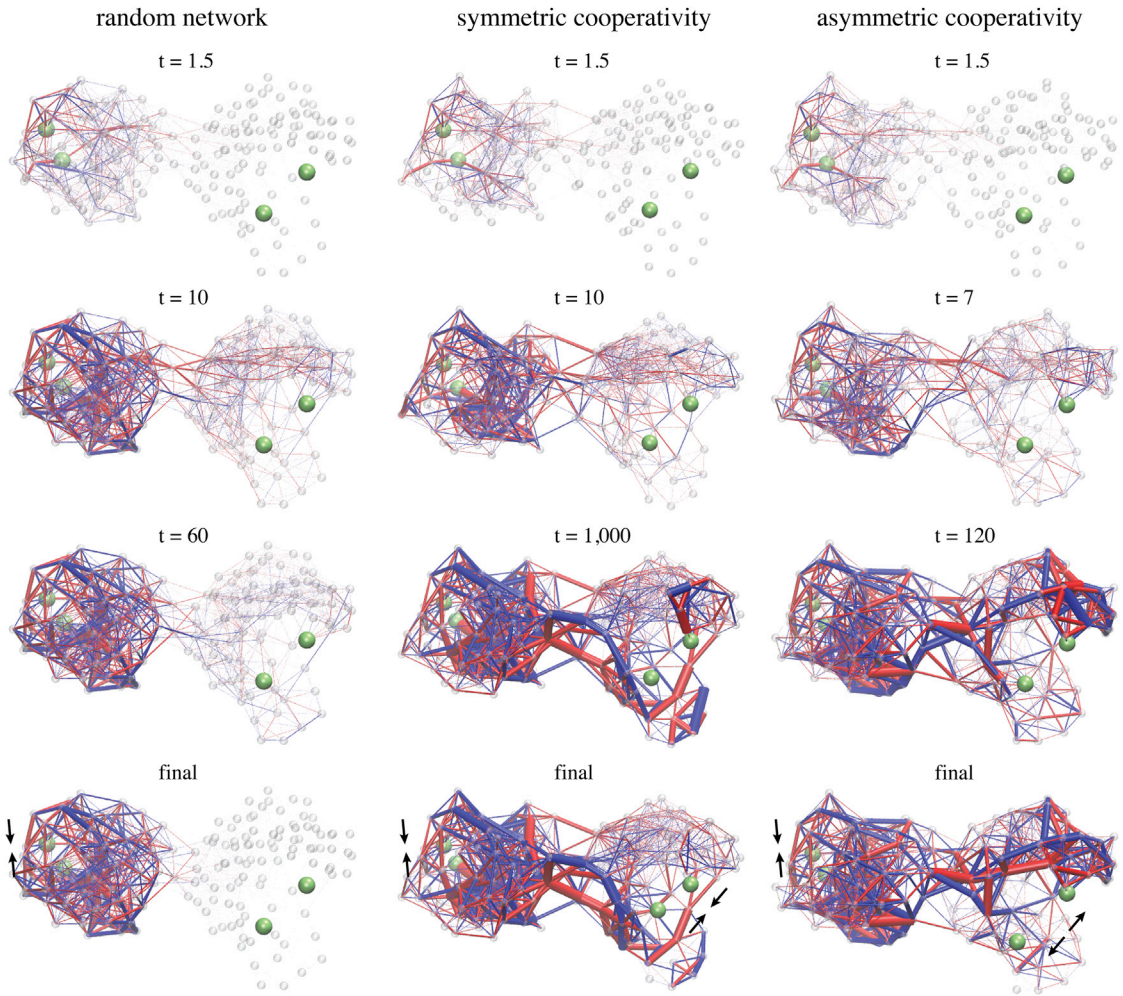


FIGURE 4 Strain propagation. Propagation of strain subsequent to ligand-binding to the allosteric pocket is visualized in the random network and in the two designed structures. Snapshots taken at different time moments during each simulation are shown in similar perspectives. Network bonds are displayed bicolored, with blue indicating stretching of the bond and red denoting compression as compared to the corresponding natural bond length. The thickness encodes the absolute value of strain the link was subject to. To see this figure in color, go online.

domain during the simulation was practically absent. The structure of this domain is apparently very stiff, preventing any internal conformational motions from occurring. Hence, the random elastic network cannot conduct allosteric communication. The situation in both designed networks was completely different. First, after ligand binding to the allosteric site, conformational motions spread across the domain interface and generated deformations in the second domain and in the vicinity of the regulated site. Second, not all elastic pair interactions in the designed structures were equally involved in the process of allosteric coupling. Instead, the temporal distribution of strain shows that communication between the remote pockets proceeded through specific subnetworks that were apparently critical for allostery in the designed structures. Those networks are formed by a subset of beads connected by springs that undergo major deformations and, hence, contribute essential pathways for the spreading of conformational motions from the allosteric site across the structure to the regulated site. In both designed network structures, such remarkably strained springs were found in domain 1 and, in contrast to the random network, in the central domain interface region and domain 2, where they form sparse clusters of important pair interactions between beads.

Regions in which the elastic links remained only marginally populated by strain were rather stiff as compared to the rest of the network. In that regard we observed that the structure of domain 2 in the designed networks is special. In the network with symmetric allosteric cooperativity, the two lobes in the tweezerlike domain structure, each harboring one of the pocket beads of the regulated site, were found to be quite flexible whereas the other parts in domain 2 were stiff (see Fig. 4). In the network with asymmetric allosteric coupling, the hinge region connecting both lobes appears to be very flexible instead (see Fig. 4). The specific pattern of flexibility in the designed networks has apparently

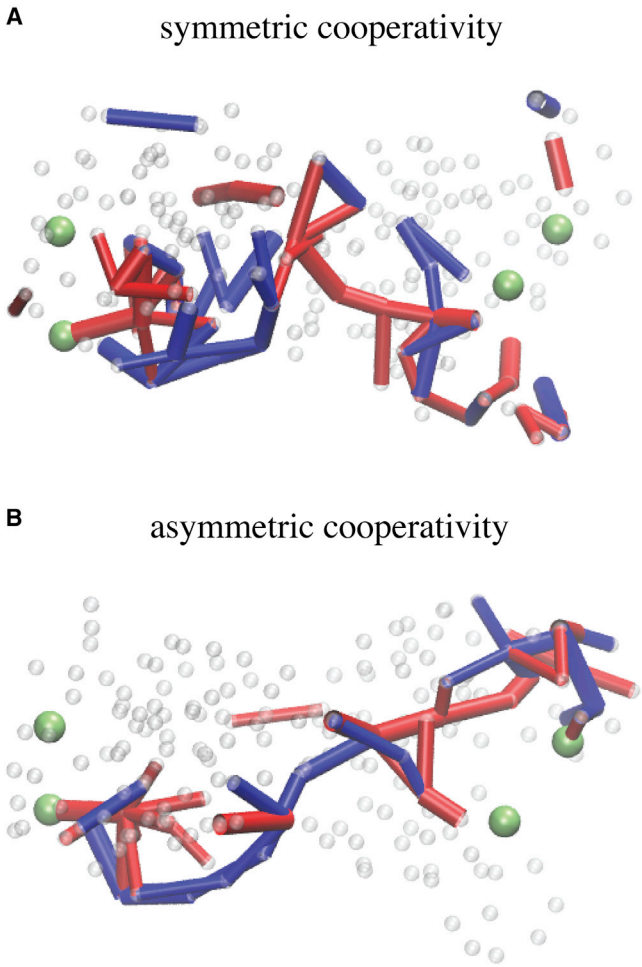


FIGURE 5 Communication chains in designed structures. (*A* and *B*) Elastic links, which are significantly involved in the transport of strain, are found to form simple connected chains that meander through the structure and connect the allosteric pocket in domain 1 (*green beads*, *left side*) with the regulated pocket in the remote second domain, domain 2 (*green beads*, *right side*). To see this figure in color, go online.

emerged during the process of evolutionary optimization to make opening/closing motions of the regulated site possible, and hence enable the designed network architectures to conduct allosteric communication.

The analysis of strain propagation in the designed networks clearly reveals functionally important signatures underlying the investigated types of allosteric coupling. Nonetheless, the strain distribution was rather complex, and we therefore aimed to deduce simpler pathways that could characterize communication between the pockets, by focusing only on those network springs that carry major strain (see Methods). Results are shown in Fig. 5. In both networks, simple chains formed by significantly deformed adjacent network links were found to meander from the allosteric pocket in domain 1 via the domain interface to the regulated pocket in domain 2. Such chains are composed of a series of stretched springs in domain 1, that are potentially caused by the closing motion of the allosteric site. They are further connected to a linkage of compressed springs located in domain 2, and trigger the respective allosteric response there.

The identified communication chains are obviously critical for the transmission of conformational changes between the domains and therefore may represent a functional skeleton, critical for allosteric activity in the designed network architectures. Immediately, one may pose the question whether, and to what extent, allosteric coupling can be maintained if structural changes would be applied to such motifs.

Robustness of allosteric communication

With regard to the robustness of allosteric communication in the two designed networks, the intention was not to perform a systematic screening of the effect of structural perturbations. Instead, the aim was to present a demonstration. Therefore the analysis of robustness was limited to a few exemplary structural mutations: those applied to beads belonging to the identified communication pathways and on examining their influence on allosteric interactions. In a first set of simulations, mutations consisting of the deletion of only a single elastic link between two network beads were considered. In the realm of proteins, such type of perturbation would roughly correspond to the mutation of a single amino acid residue, resulting in specific local interactions disappearing. For each specific mutant network, a single simulation starting with ligand binding to the allosteric pocket was carried out and after completion, the response in the regulated pocket was quantified and the level of allosteric communication determined. The results are listed in Table S1. It is found that allostery in both designed networks is generally robust with respect to the removal of single interactions. To understand how long-range coupling between both pockets is maintained in the mutant networks, we have visualized the propagation of strain and determined communication pathways for some exemplary cases, similar to what has been performed for the designed wild-type networks (see previous section).

For the network with symmetric cooperativity, three cases were focused on in more detail, all of which maintained allosteric activity (see Table S1). They corresponded to a mutant with one major interaction at the central domain interface having been removed and another mutant with a link deleted in domain 2, closer to the regulated pocket. The computed communication chains in those two mutants are shown in Fig. S3. Movies of strain propagation in those two mutant networks are provided as Movies S4 and S5. To highlight robustness in the first designed network, a third example, consisting of a double mutant network in which both selected links were deleted, was also considered (see Movie S6). For all three mutant networks, we find that the strain propagation networks as well as the extracted

communication chains are very similar to that of the wild-type networks, except in the vicinity of the respective mutation site (see Fig. 6, *A* and *B*; Fig. S3). There it is found that in the neighborhood of the deleted link, a few bead interactions that in the wild-type network have not played a major role in the transmission of strain, become important, forming a bridge through which strain was able to side-track the mutation site (see Figs. 6 *B* and S3). This result demonstrates that in the mutant structures, similar communication pathways are available which, together with the bridge motifs that compensate the mutation defect, are employed for the propagation of conformational changes and hence may provide the foundation for the robustness of allosteric coupling in this designed network.

For the designed network with asymmetric allosteric coupling, one mutant structure with a single deleted link located at the domain interface has been selected for illustration. Despite the mutation, it revealed full allosteric activity. Propagation of strain subsequent to ligand binding to the allosteric pocket is shown in Movie S7 and the extracted communication pathways are depicted in Fig. 6 *b*. It is found that the network of bead interactions through which communication between both pockets was transmitted is rather complex, with the strain populating large parts of the mutant structure. Hence, in contrast to the wild-type network, allosteric coupling in this mutant does not proceed via simple communication chains.

Although, generally, allosteric communication in both designed networks was found to be robust with respect to the deletion of a single bead interaction, there were still a few critical mutations that involved a drastic decrease in the coupling between both pockets. They can correspond to perturbations in the vicinity of either pocket, or are located at the domain interface (see Fig. S2, *A* and *B*). In the final step of investigations, the effect of stronger mutations was considered, where an entire bead was deleted in a simulation of each network. Those perturbations typically led to a remarkable drop or even complete knockout of allosteric communication (see Table S1).

Application to allostery in proteins

In this article, two prototype coarse-grained architectures of fictitious allosteric proteins have been successfully designed and investigated. In the final part of the article, it is demonstrated that the developed modeling approach, the analysis of strain propagation, and the methods to quantify and visualize communication pathways, can be applied in the

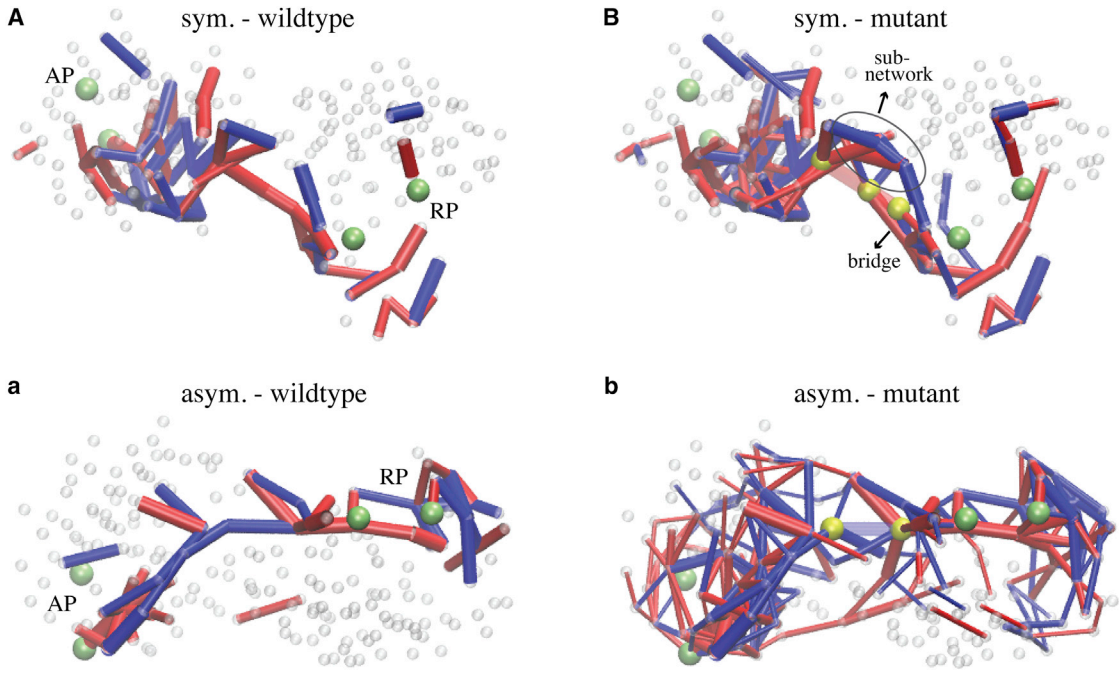


FIGURE 6 Communication pathways in mutant structures. Communication chains in the designed wild-type network structures (*A* and *a*) are compared with the communication pattern obtained from strain propagation in a corresponding selected mutant structure (*B* and *b*). Communication chains in (*A*) and (*a*) are the same as shown in Fig. 5, but here another perspective is chosen for better comparison with the mutant structures. (*B*) Shown here are communication pathways in the 52-197-196 double mutant of the network with symmetric allosteric coupling. Beads with index 52, 197 and 196 are highlighted in yellow and the two bonds removed between them are shown as transparent. Links that, in this mutant, became important to redirect strain over the mutation site, are indicated by an ellipse. (*b*) Given here are communication pathways in the 52-185 mutant of the designed network with asymmetric allosteric coupling. Beads 52 and 185 are displayed in yellow and the deleted link between them is transparent. Allosteric communication in this mutant was maintained via a complex strain transportation network, and a simple chain, as found in the wild-type, was not available (compare *a* and *b*). To see this figure in color, go online.

structure-based investigation of real allosteric proteins. For the purpose of demonstration, the modeling scheme was applied to the popular myosin V (myoV) molecular motor, which steps along actin filaments operating as a processive cargo transporter (40). In molecular motors, the nucleotide pocket (where chemical events, consisting of ATP-binding and hydrolysis, take place) is typically spatially remote from the site where mechanical forces are exerted, e.g., on filaments, and allosteric regulation between them is essential for chemo-mechanical coupling and the motors' functional activity.

Here, for the myoV motor domain (41), allosteric communication between the ATP-binding pocket and the actin-binding cleft was investigated. In the following, concepts and results are focused on whereas an in-detail description of the undertaken modeling can be found in the Supporting Materials and Methods. First, an elastic network of the nucleotide-free motor domain of myoV was constructed. Then, binding of an ATP molecule was emulated by applying precise forces only locally to the nucleotide-binding pocket, dynamically steering it to the experimentally known conformation that was complexed to an ATP analog (41). During this simulation, the entire conformational dynamics in myoV was monitored, and allosteric transduction of motions from the nucleotide-binding pocket to the actin cleft detected.

It is found that localized ATP-related dynamics, consisting of relative motions of the P-loop and switch I and II inside the nucleotide-binding pocket, was remarkably well reproduced. Initially, residues of the free pocket compared well with an RMSD of 2.4 Å to those corresponding to the ATP-bound conformation, although during the simulation, the RMSD dropped to zero (see Fig. S4), indicating that, locally, motions that corresponded to ATP-binding in myoV were perfectly reproduced. Furthermore, in response to ATP-binding to its pocket, it is found that conformational changes within the entire motor domain slowly set in and large-amplitude relative motions of the L50 and U50 were generated. As a result of these motions, the actin-binding cleft also underwent changes and significantly opened up (changes of 3.5 Å; see Fig. S4). In Movie S8, conformational motions during the allosteric transition in myoV are visualized, and in Fig. 7 *A*, the initial free conformation and that obtained after ATP-binding are superimposed, highlighting domain rearrangements and the coupled allosteric conformational changes between the nucleotide pocket and the actin site.

Next, to elucidate mechanical aspects of allostery in myoV, strain propagation was analyzed, employing the same methods developed for the designed artificial protein networks. In Movie S9, propagation of strain subsequent to ATP-ligand binding is visualized, and snapshots taken at different time moments during the simulation are provided in Fig. S5. It was found that ATP-binding generated significant strain that was localized near the ATP motifs and further spread into the regions around the actin cleft. We observed that propagation of strain from the ATP-pocket to the actin site proceeded through well-defined subnetworks formed by spring connections that underwent significant deformations. The identified pathways and chains found to be essential for strain transport are shown in Fig. 7 *B*. Remarkably, regions in which strain was accumulated primarily coincided with the structural motifs of conserved residues that are critical for the mechanochemistry of the myoV motor. First, due to ATP-binding, considerable strain is localized to the P-loop, switch I, and switch II motifs, and in the regions bridging them, e.g., near the entrance to the ATP-pocket. Second, there are definite and pronounced transport pathways reaching from the switch I and switch II motifs into the L50 domain from where they meander further into the actin site (see Fig. 7 *B*). There, strain occupied critical motifs involved in interactions with actin filaments, e.g., the activation loop, the HCM loop, and the strut. The latter region provides the flexible

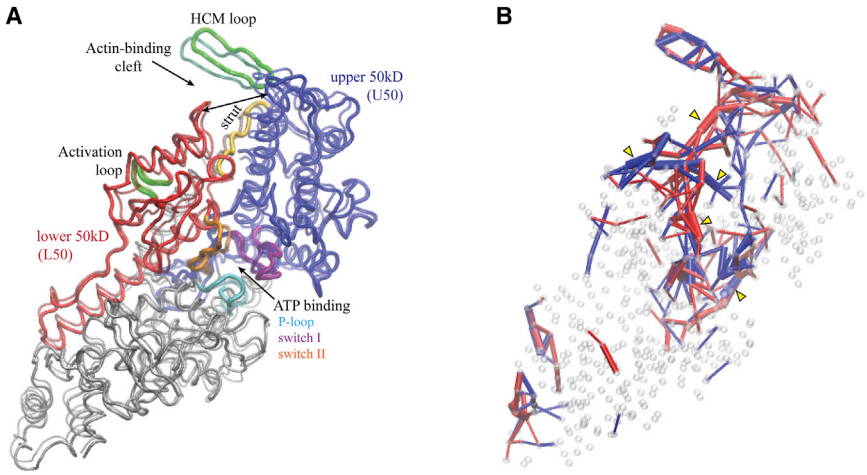


FIGURE 7 Allosteric regulation in myosin V. In (*A*), the nucleotide-free structure of the myoV motor domain (shown as *transparent*) and the conformation obtained after modeling of ATP-binding (as *solid*) are superimposed. Significant opening of the actin-binding cleft resulting from allosteric interactions with the ATP-binding pocket was observed. Major domains are colored red and blue, and structural motifs involved in ATP-binding and interactions with actin are also highlighted. In (*B*), the identified communication pathways underlying allosteric regulation in myoV are shown. Coding of color and bond strength is similar to that used in the analysis of the designed networks (Figs. 4 and 5). Localization of elastic strain in the ATP and actin interaction motifs together with transport chains between them are clearly visible. Essential regions are marked by yellow arrows. To see this figure in color, go online.

linker between L50 and U50 domains and its strained state may explain opening of the actin cleft.

Exemplary mutations applied near the conserved motifs turned out to significantly reduce allosteric coupling in myoV (details are given in Supporting Materials and Methods).

Application of the modeling scheme, developed to design prototype structures with optimal allosteric coupling, was demonstrated here for the so-called "rigor state" (nucleotide-free, actin-bound) to "postrigor" state (ATP-bound, detached from actin) transition in the myosin V protein. This transition is well investigated experimentally, and based on previous computational studies (e.g., (37,42)). In our simulations, the local conformational dynamics related to binding of an ATP ligand as well as the concomitant actin-site opening resulting from allosteric coupling was remarkably well reproduced, evidencing the applicability of the developed approach to the realm of protein allostery (see Discussion). Analysis of strain propagation revealed distinct functional signatures of intramolecular communication, and hence may provide interesting novel aspects for a mechanistic-level explanation of the allosteric regulation in myosin V.

DISCUSSION

The phenomenon of allosteric coupling between different functional regions within a macromolecular structure is ubiquitously present in proteins and therefore raises important questions about the fundamental nature of the underlying mechanisms. In contrast to typically employed structure-based modeling of protein dynamics, a different methodology toward approaching such essential problems is presented in this article. Instead of considering real protein structures, artificial analog structures that encode pronounced long-range allosteric coupling between two spatially remote pockets were designed. The structures were represented by elastic networks, similar to the coarse-grained models widely used to describe conformational dynamics of real proteins.

A force-probe scheme consisting of ligand-binding to the allosteric pocket, following conformational motions spreading over the structure, and detecting the response generated in the regulated pocket, was implemented to evaluate the ability of the network to conduct allosteric communication. Initially starting with a random elastic network, which did not reveal any allostery, a scheme of evolutionary optimization was iteratively applied to design two prototype elastic-network structures with perfected allosteric coupling, one with symmetric and the second with asymmetric cooperativity.

In the designed networks, well-defined pathways and simple chains of important interactions, established by only a few network beads, were identified to constitute the signatures that underlie allosteric communication. Remote interactions were robust even if minor structural perturbations were applied. However, a single critical mutation could knock out completely allosteric communication in the designed networks.

Whereas the first descriptions of allosteric systems, namely the MWC and the KNF models (43,44), were formulated more than 50 years ago in the absence of any structural data and were thus of a phenomenological nature, the development of very sophisticated experimental techniques and the vast amount of protein structures becoming available in the last decades have since allowed us to investigate allosteric communication in proteins on the molecular level and led to a decent understanding of the mechanism underlying allostery in several role model proteins. In the recent past, structure-based computational modeling, aiming to follow the conformational motions as the underpinning of allosteric effects in proteins, has gained a lot of attention. In particular, due to the timescale gap present in atomistic-level MD simulations, coarse-grained elastic network models that are limited to the mechanical aspect of protein operation became very popular to investigate slow allosteric transitions with timescales beyond the microsecond range. In those studies, the analysis of conformational changes is typically based on the computation of normal modes, and allosteric effects are discussed in terms of short- and long-ranged correlations of amino-acid residue displacements.

The aim in this study was to present a model that emphasizes the mechanical picture of allosteric systems. Therefore, the elastic network model was used. Here, however, the full elastic dynamics of the network was considered by always numerically integrating the nonlinear equations of motion and monitoring processes of conformational relaxation; no linearization was performed and the conformational dynamics beyond the limiting normal mode approximation could be followed. As a consequence, this model has the obvious advantage that it can resolve the temporal order of events that eventually establish allosteric communication in the network structures, starting from forces and strains that were first generated locally at the allosteric site as a consequence of ligand binding, followed by the subsequent propagation of conformational motions via the domain interface, to finally induce structural changes in the remote regulated pocket. This model, therefore, naturally includes causality as the guiding principle to manifest allosteric communication, and therefore is richer in its explanatory power compared to correlation-based analyses.

This model emphasizes the structural viewpoint of allostery (as described in (1)) in which allostery is regarded as a consequence of optimized communication between the remote functional sites, established through the propagation of strain along pathways that are formed by a set of interacting residues. The employed elastic network description clearly implies limitations. All network particles

are of the same kind and physical interactions between them are incorporated into empirical effective potentials; thermal fluctuations were also neglected for simplicity. Despite such gross simplifications, the dynamical properties of the designed prototype allosteric structures reveal remarkable similarities with those found in real allosteric proteins: 1) specific parts of the structure are flexible, whereas other regions form stiff clusters; 2) the propagation of conformational changes that results in the long-range coupling of the remote sites proceeds through communication pathways involving only a few of the many intrastructural interactions; and 3) single critical mutations can knock out allosteric coupling.

Application of the modeling scheme to the rigor-to-post-rigor transition in the myosin V molecular motor correctly reproduced the conformational dynamics corresponding to the ATP-binding related detachment from actin filaments via allosteric transduction. Despite the complexity of chemo-mechanical operation, simple communication pathways were identified that underlie allosteric coupling in this protein motor. Hence, the applicability of the developed approach to investigate protein allostery is demonstrated. Moreover, commonalities between the designed prototype networks and a real protein is exemplified. Although the artificial network structures have emerged as a result of simplified in silico evolution, they share functional properties with a protein structure that underwent real biological evolution to encode allosteric dynamics.

In summary, we may conclude that the designed elastic network structures can provide suitable physical models for the mechanics of complex allosteric biomolecules, such as proteins.

It should be remarked that the evolutionary pressure imposed in the design process consisted only in magnifying changes in the regulated pocket in response to ligand-induced changes in the allosteric pocket; no other requirements were imposed and the dynamical properties that actually improved allosteric communication in the evolving structures emerged autonomously. However, in future studies, design algorithms that involve coevolution can also be implemented and multiple optimization criteria can be imposed. Moreover, in future studies, the design and analysis of a larger number of allosteric elastic networks can be undertaken. That would allow us to compare properties such as communication pathways among them and possibly draw conclusions on the generality of such dynamical motifs.

Previously, the relaxational elastic-network approach employed in this study was already applied to model ATP-related dynamics in protein machines and to investigate allosteric coupling (35,36,45). Whereas in those studies allosteric coupling was only qualitatively discussed, the approach presented here may generally be applied in structure-based modeling and help in understanding important mechanistic aspects of protein allostery.

Added note

After finishing the article I became aware of two quite recently appeared works, in which the design of mechanical networks with allosteric coupling had also been undertaken (46,47). There, however, different computational algorithms were used and optimization of allosteric coupling was performed within the linear response limitation; the underlying network architectures were also very different (e.g., 2D on-lattice models (46)). Both works principally demonstrate the applicability of design processes to obtain desired responses. The design process introduced in my article was developed completely independently and all results were presented already in January 2016 (see (48)). Besides other differences, its principal distinction is that during the design process, the allosteric response was optimized by always considering the full nonlinear elastic dynamics of the networks. The importance of nonlinearities for protein function involving allostery has previously been evidenced (e.g., in (49,50)). The network architectures used in my study resemble more closely the 3D compact fold of real proteins; in fact, the designed networks can be regarded as coarse-grained representations of fictitious protein structures. Most importantly, this work goes beyond developing the design process. In fact, the main emphasis here was put on understanding mechanical properties underlying long-range communication in the designed networks; this was achieved by visualizing and analyzing propagation of strain, extracting communication pathways and chains that were critical for allosteric coupling, and investigating robustness of the designed functional properties.

SUPPORTING MATERIAL

Supporting Materials and Methods, nine movies, and three data files are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(17)30693-8.

AUTHOR CONTRIBUTIONS

H.F. designed and performed research, analyzed data, and wrote the manuscript.

ACKNOWLEDGMENTS

The author is grateful to Alexander S. Mikhailov and Yuichi Togashi for helpful discussions.

All figures of networks and the movies have been prepared using the VMD software (51). This work was supported by JSPS KAKENHI grant No. JP16K05518.

REFERENCES

1. Tsai, C.-J., and R. Nussinov. 2014. A unified view of "how allostery works". *PLOS Comput. Biol.* 10:e1003394.

2. Brüschweiler, S., P. Schanda, …, M. Tollinger. 2009. Direct observation of the dynamic process underlying allosteric signal transmission. *J. Am. Chem. Soc.* 131:3063–3068.
3. Grutsch, S., S. Brüschweiler, and M. Tollinger. 2016. NMR methods to study dynamic allostery. *PLOS Comput. Biol.* 12:e1004620.
4. del Sol, A., H. Fujihashi, …, R. Nussinov. 2006. Residues crucial for maintaining short paths in network communication mediate signaling in proteins. *Mol. Syst. Biol.* 2:2006;2:2006.0019. Epub 2006 May 2.
5. del Sol, A., M. J. Araúzo-Bravo, …, R. Nussinov. 2007. Modular architecture of protein structures and allosteric communications: potential implications for signaling proteins and regulatory linkages. *Genome Biol.* 8:R92.
6. Daily, M. D., T. J. Upadhyaya, and J. J. Gray. 2008. Contact rearrangements form coupled networks from local motions in allosteric proteins. *Proteins.* 71:455–466.
7. Chennubhotla, C., and I. Bahar. 2006. Markov propagation of allosteric effects in biomolecular systems: application to GroEL-GroES. *Mol. Syst. Biol.* 2:36.
8. Chennubhotla, C., and I. Bahar. 2007. Signal propagation in proteins and relation to equilibrium fluctuations. *PLOS Comput. Biol.* 3:1716–1726.
9. Lockless, S. W., and R. Ranganathan. 1999. Evolutionarily conserved pathways of energetic connectivity in protein families. *Science.* 286:295–299.
10. Süel, G. M., S. W. Lockless, …, R. Ranganathan. 2003. Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat. Struct. Biol.* 10:59–69.
11. Dima, R. I., and D. Thirumalai. 2006. Determination of network of residues that regulate allostery in protein families using sequence analysis. *Protein Sci.* 15:258–268.
12. Tang, S., J. C. Liao, …, J. P. Schmidt. 2007. Predicting allosteric communication in myosin via a pathway of conserved residues. *J. Mol. Biol.* 373:1361–1373.
13. Ma, J., P. B. Sigler, …, M. Karplus. 2000. A dynamic model for the allosteric mechanism of GroEL. *J. Mol. Biol.* 302:303–313.
14. Ghosh, A., and S. Vishveshwara. 2007. A study of communication pathways in methionyl- tRNA synthetase by molecular dynamics simulations and structure network analysis. *Proc. Natl. Acad. Sci. USA.* 104:15711–15716.
15. Cui, Q., and M. Karplus. 2008. Allostery and cooperativity revisited. *Protein Sci.* 17:1295–1307.
16. Dixit, A., and G. M. Verkhivker. 2011. Computational modeling of allosteric communication reveals organizing principles of mutation-induced signaling in ABL and EGFR kinases. *PLOS Comput. Biol.* 7:e1002179.
17. Laine, E., C. Auclair, and L. Tchertanov. 2012. Allosteric communication across the native and mutated KIT receptor tyrosine kinase. *PLOS Comput. Biol.* 8:e1002661.
18. Naithani, A., P. Taylor, …, M. D. Walkinshaw. 2015. A molecular dynamics study of allosteric transitions in *Leishmania mexicana* pyruvate kinase. *Biophys. J.* 109:1149–1156.
19. Hertig, S., N. R. Latorraca, and R. O. Dror. 2016. Revealing atomic-level mechanisms of protein allostery with molecular dynamics simulations. *PLOS Comput. Biol.* 12:e1004746.
20. Xu, C., D. Tobi, and I. Bahar. 2003. Allosteric changes in protein structure computed by a simple mechanical model: hemoglobin T↔R2 transition. *J. Mol. Biol.* 333:153–168.
21. Zheng, W., B. R. Brooks, and D. Thirumalai. 2006. Low-frequency normal modes that describe allosteric transitions in biological nanomachines are robust to sequence variations. *Proc. Natl. Acad. Sci. USA.* 103:7664–7669.
22. Chennubhotla, C., Z. Yang, and I. Bahar. 2008. Coupling between global dynamics and signal transduction pathways: a mechanism of allostery for chaperonin GroEL. *Mol. Biosyst.* 4:287–292.
23. Yang, Z., P. Májek, and I. Bahar. 2009. Allosteric transitions of supramolecular systems explored by network models: application to chaperonin GroEL. *PLOS Comput. Biol.* 5:e1000360.
24. Morra, G., G. Verkhivker, and G. Colombo. 2009. Modeling signal propagation mechanisms and ligand-based conformational dynamics of the Hsp90 molecular chaperone full-length dimer. *PLOS Comput. Biol.* 5:e1000323.
25. Erman, B. 2013. A fast approximate method of identifying paths of allosteric communication in proteins. *Proteins.* 81:1097–1101.
26. Krieger, J., I. Bahar, and I. H. Greger. 2015. Structure, dynamics, and allosteric potential of ionotropic glutamate receptor N-terminal domains. *Biophys. J.* 109:1136–1148.
27. Bahar, I., C. Chennubhotla, and D. Tobi. 2007. Intrinsic dynamics of enzymes in the unbound state and relation to allosteric regulation. *Curr. Opin. Struct. Biol.* 17:633–640.
28. Bahar, I., T. R. Lezon, …, E. Eyal. 2010. Global dynamics of proteins: bridging between structure and function. *Annu. Rev. Biophys.* 39:23–42.
29. Daily, M. D., and J. J. Gray. 2007. Local motions in a benchmark of allosteric proteins. *Proteins.* 67:385–399.
30. Togashi, Y., and A. S. Mikhailov. 2007. Nonlinear relaxation dynamics in elastic networks and design principles of molecular machines. *Proc. Natl. Acad. Sci. USA.* 104:8697–8702.
31. Cressman, A., Y. Togashi, …, R. Kapral. 2008. Mesoscale modeling of molecular machines: cyclic dynamics and hydrodynamical fluctuations. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 77:050901 (R).
32. Huang, M.-J., R. Kapral, …, H.-Y. Chen. 2013. Coarse-grain simulations of active molecular machines in lipid bilayers. *J. Chem. Phys.* 138:195101.
33. Sakaue, T., R. Kapral, and A. S. Mikhailov. 2010. Nanoscale swimmers: hydrodynamic interactions and propulsion of molecular machines. *Eur. Phys. J. B.* 75:381–387.
34. Sarkar, A., H. Flechsig, and A. S. Mikhailov. 2016. Towards synthetic molecular motors: a model elastic-network study. *New J. Phys.* 18:043006.
35. Flechsig, H., and A. S. Mikhailov. 2010. Tracing entire operation cycles of molecular motor hepatitis C virus helicase in structurally resolved dynamical simulations. *Proc. Natl. Acad. Sci. USA.* 107:20875–20880.
36. Flechsig, H., D. Popp, and A. S. Mikhailov. 2011. In silico investigation of conformational motions in superfamily 2 helicase proteins. *PLoS One.* 6:e21809.
37. Düttmann, M., Y. Togashi, …, A. S. Mikhailov. 2012. Myosin-V as a mechanical sensor: an elastic network study. *Biophys. J.* 102:542–551.
38. Tirion, M. M. 1996. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* 77:1905–1908.
39. Haliloglu, T., I. Bahar, and B. Erman. 1997. Gaussian dynamics of folded proteins. *Phys. Rev. Lett.* 79:3090–3093.
40. Houdusse, A., and H. L. Sweeney. 2016. How myosin generates force on actin filaments. *Trends Biochem. Sci.* 41:989–997.
41. Coureux, P.-D., H. L. Sweeney, and A. Houdusse. 2004. Three myosin V structures delineate essential features of chemo-mechanical transduction. *EMBO J.* 23:4527–4537.
42. Cecchini, M., A. Houdusse, and M. Karplus. 2008. Allosteric communication in myosin V: from small conformational changes to large directed movements. *PLOS Comput. Biol.* 4:e1000129.
43. Monod, J., J. Wyman, and J. P. Changeux. 1965. On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* 12:88–118.
44. Koshland, D. E., Jr., G. Némethy, and D. Filmer. 1966. Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry.* 5:365–385.
45. Flechsig, H. 2016. Nucleotide-induced conformational dynamics in ABC transporters from structure-based coarse grained modeling. *Front. Phys.* 4:3.

46. Yan, L., R. Ravasio, …, M. Wyart. 2017. Architecture and coevolution of allosteric materials. *Proc. Natl. Acad. Sci. USA.* 114:2526–2531.

47. Rocks, J. W., N. Pashine, …, S. R. Nagel. 2017. Designing allostery-inspired response in mechanical networks. *Proc. Natl. Acad. Sci. USA.* 114:2520–2525.

48. Flechsig, H. 2016. Evolutionary optimization of simple polymer networks: models of synthetic allosteric proteins. *In* NIG International Symposium 2016 Japan Q-Bio week, Tokyo Symposium, Jan 9–11, 2016. https://q-bio.jp/wiki/IIS_Sympo#NIG_International_Symposium_2016.

49. Miyashita, O., J. N. Onuchic, and P. G. Wolynes. 2003. Nonlinear elasticity, protein quakes, and the energy landscapes of functional transitions in proteins. *Proc. Natl. Acad. Sci. USA.* 100:12570–12575.

50. Togashi, Y., T. Yanagida, and A. S. Mikhailov. 2010. Nonlinearity of mechanochemical motions in motor proteins. *PLOS Comput. Biol.* 6:e1000814.

51. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:27–38.

Biophysical Journal, Volume *113*

Supplemental Information

Design of Elastic Networks with Evolutionary Optimized Long-Range Communication as Mechanical Models of Allosteric Proteins

Holger Flechsig

PROBING OF ALLOSTERY AND EVOLUTIONARY OPTIMISATION

The force-probe employed scheme to score the allosteric response in the networks during the design processes was explained in the main text. To follow conformational motions in a network subsequent to ligand-binding to the allosteric site, the set of equations (Eq. (1), main text) was numerically integrated, always until a steady state of that network in which motions were sufficiently terminated was reached (the time needed was referred to as T). As a termination condition for the process of numerical integration a requirement for the instantaneous bead velocity (averaged over all beads) to drop below a prescribed threshold was imposed in the simulations. The condition was $\frac{1}{N}\sum_i^N |v_i| < 10^{-6}$.

During the process of evolutionary optimisation in the two designed networks the improvement of the allosteric response was recorded. It is shown in Fig. S1. Details are given in the figure caption.

STRAIN PROPAGATION AND COMMUNICATION PATHWAYS

To discuss the anisotropy of the strain distribution in the random and designed networks we have introduced a method to normalise the strain of elastic links with respect to the distance from the allosteric site. Therefore, we introduced the shortest *graph distance* between a network bead with index i and a bead of the allosteric pocket with index p as the minimal number of links of a path connecting the two beads, denoted by D_{ip}. The shortest *graph distance* of an elastic spring connecting beads i and j to the allosteric pocket was then defined as $D_{ij} := \min\{D_{ip1}, D_{ip2}, D_{jp1}, D_{jp2}\}$, where $p1$ and $p2$ were the indices of beads of the allosteric pocket. Next we defined shells S_n of links, each of which contained all those elastic links (ij) that had the same shortest *graph distance* n ($n = 1, 2, \dots$) to the allosteric pocket, i.e. $S_n := \{\text{all links (ij) for which } D_{ij} = n\}$. Hence, each elastic network link (ij) was uniquely assigned to one shell S_n and its elastic strain at time t (defined in the Methods section of the main text) was denoted by $s_{ij}(t, n)$. After this procedure, obviously all elastic links had been sorted into their respective shells; all links with shortest *graph distance* 1 were in shell S_1, those with shortest *graph distance* 2 were in shell S_2, etc. Now, during a first simulation of the allosteric operation of a designed network - following ligand binding to the allosteric site until the steady conformation was reached - for each shell S_n a maximum

absolute strain value $m_n = \max\{|s_{ij}(t,n)|\}$ of all links belonging to the same shell S_n was determined and stored. Then, in a repeated simulation of the same network, the strain of each elastic link (ij) was normalised by dividing its value by the maximum absolute strain value m_n of the shell n the particular link belonged to, i.e. $\tilde{s}_{ij}(t) = s_{ij}(t,n)/m_n$. The normalised strain $\tilde{s}_{ij}(t)$ was used to visualise the strain propagation in both designed networks in a time-resolved fashion, presented in Movies S2 and S3. Corresponding snapshots are shown in the main text Fig. 4. In the same simulation we have memorised for each link (ij) the maximum absolute value m_{ij} of its normalised strain, i.e. $m_{ij} = \max|\tilde{s}_{ij}(t)|$. Those values were employed to determine the communication pathways (shown in the main text Fig. 5), which were constructed by considering only those links that were significantly involved in the strain propagation, imposed by the condition $m_{ij} > \tilde{s}_t$, where $\tilde{s}_t$ was a prescribed threshold value for the normalised strain (with its values given in the main text Methods section).

In the random elastic network a similar procedure of strain normalisation was undertaken. However, to compare the propagation of strain to that in the two designed networks, the link strain was normalised with respect to the coefficients m_n from the designed network with symmetric allosteric coupling.

For the selected mutant networks the procedure of strain normalisation was also applied to visualise strain propagation during the allosteric transition (Movies S4-S7) and to construct communication pathways (Fig. 6 main text and Fig. S3).

It should be noted that for the normalisation procedure the shortest *graph distance* between a network bead and the a bead of the allosteric pocket was determined with a standard algorithm using the powers of the adjacency matrix.

ROBUSTNESS OF ALLOSTERIC COMMUNICATION

Robustness of allosteric communication with respect to exemplary single structural mutations applied to each of the two designed prototype networks was analysed. To determine the communication chains for the two mutants of the designed network with symmetric allosteric communication shown in Fig. S3, a threshold of $\tilde{s}_t = 55\%$ was used for the $52-197$ mutant, and $\tilde{s}_t = 60\%$ was used for the $196-197$ mutant.

APPLICATION TO THE MYOSIN V MOLECULAR MOTOR

Network construction and elastic dynamics

The two designed prototype structures serve as coarse-grained representations of artificial proteins and allosteric conformational changes corresponded to relaxational motions of their respective elastic networks. In the application to real allosteric proteins the same elastic network model was used, similar to our previous publications (see references in the main text). The elastic network of the myosin V motor domain (myoV) was constructed for the nucleotide-free structure [1] (PDB code: 1W8J; residues: Ala2-Leu42, Asp47-Ser184, Asn191-Gln594, Lys634-Lys766). Each of the amino acid residues was replaced by a single bead. Bead positions $\vec{R}_i^{(0)}$ corresponded to the alpha-carbon position of the respective residue and the network of elastic springs connecting the beads was constructed by comparing all distances $d_{ij}^{(0)}$ with the prescribed interaction radius r_{int}, thus determining the adjacency matrix (see main text Methods section). For myoV an interaction radius of 7.5Å was used. In elastic-network simulations of real proteins the choice of interaction radius is widely debated and values between 7.0Å and 14.0Å are commonly employed (see e.g. [2]). The constructed elastic network of the nucleotide-free myoV had $N = 716$ beads connected by 3114 springs. For myoV the same notations as for the artificial networks was used and similarly the same equations of motions described the elastic dynamics of the myoV network (see main text Methods section and equations (1) there).

Modelling of ATP-binding and allosteric coupling between nucleotide binding pocket and actin cleft

In the designed networks the modelling scheme consisted in local binding of a ligand bead to the allosteric pocket, following subsequent conformational motions through the network structure, and, evaluating the allosteric response in the remote regulated pocket. A similar scheme shall be implemented in the investigation of allosteric dynamics in myoV. Nonetheless, simplifications applied in the designed networks, such as a two-bead pocket and pair-forces to mimic ligand binding, cannot be maintained and shall be modified in order to account for the circumstances in real proteins.

The nucleotide pocket in myoV is an extended structure with many residues involved

and interactions with nucleotides are very complex and involve binding of ATP, its hydrolysis reaction, and, the removal of chemical products. Here, to remain consistent with the description developed for the artificial networks, we focused on the nucleotide binding process and implemented at the coarse-grained level a scheme to emulate binding of an ATP ligand to the pocket. In myoV the majority of the nucleotide binding pocket is composed of the conserved residue motifs referred to as the P-loop (Gly163-Lys169), switch I (Gly207-Arg219), and switch II (Asp437-Asn447). In the constructed myoV elastic network the beads corresponding to all those residues were considered to be the ATP-binding pocket in our simulation. Next, from the crystal structure of the myoV motor complexed with the ATP-analog MgADP.BeFx [1] (PDB code: 1W7J), the structure of the nucleotide-binding pocket which corresponds to the ATP-bound conformation was extracted. The alpha-carbon positions of the corresponding residues were stored in vectors $\vec{R}_k^{ATP}$ (indices k indicate pocket beads). To emulate dynamically binding of an ATP-ligand to the free myoV elastic network the following strategy was employed. Through the application of dynamically chosen local forces to beads of the nucleotide binding pocket, its structure was steered from the free to the known final ATP-bound conformation. To implement this, the restraint energy $U_p = \frac{1}{2}kN_p(RMSD(t) - RMSD^*(t))^2$ was added for beads of the nucleotide binding pocket. Here, k was a stiffness parameter and N_p was the number of pocket beads. $RMSD(t)$ was the root mean square displacement of the pocket beads in the final ATP-bound conformation with respect to the actual pocket bead positions at time t after superposition (note that $RMSD(t) = RMSD(\vec{R}_{k_1}(t), \ldots, \vec{R}_{k_{N_p}}(t), \vec{R}_{k_1}^{ATP}, \ldots, \vec{R}_{k_{N_p}}^{ATP})$. $RMSD^*(t)$ was the prescribed RMSD at time t, chosen to linearly decrease with time from the initial RMSD (denoted by $RMSD^{(0)}$) to zero within the period T, i.e. $RMSD^*(t) = RMSD^{(0)} - RMSD^{(0)}\frac{t}{T}$. The steering forces $\vec{f}_k$ acting on the pocket beads (see main text Methods Eq. (1)) were computed as $\vec{f}_k = -\frac{\partial}{\partial \vec{R}_k}U_p$. Therefore, during the simulation, the conformation of the final ATP-bound pocket was always superimposed with that of the actual pocket and the steering forces were constantly updated. The set of equations of motions (see Eq. (1) in main text Methods) was integrated for a total number of 1.000.000 integration steps and $T = 500.000 \cdot dt$ was chosen. That means that during the first half of the simulation the nucleotide binding pocket was actively steered towards the ATP-bound conformation, and, after that, trapped there for the remaining half, when slower conformational motions in the rest of the motor domain were monitored. For the numerical integration a time-step of $dt = 0.01$ was used, $k = 100$

was chosen and the number of pocket beads was $N_p = 31$. It should be noted that the approach developed here corresponds to the coarse-grained analogue of widely used targeted molecular dynamics simulations.

To elucidate allosteric coupling in myoV we have particularly focused on monitoring conformational changes in the actin interaction cleft which were generated in response to ATP-binding to the nucleotide binding pocket. The actin cleft has a tweezer-like shape (see main text Fig. 7A) and in order to quantify changes there we have measured during the entire simulation the distance between just two labels; network beads corresponding to residues Ser345 and Pro516 were selected as labels. While this is a simplification, the full conformation changes are displayed in Fig. 7A and motions were recorded in Movie S8. In Fig. S4 the RMSD of the nucleotide binding pocket with respect to the ATP-bound target pocket, as it changed during the simulation is shown. Distance changes between the two labels in the actin cleft are also plotted.

Strain propagation and communication pathways

To discuss the anisotropy of the strain distribution in myoV the same method of strain normalisation which was developed to analyse the designed network was employed. Hence, in the myoV elastic network the strain of links was normalised with respect to the distance from the ATP-binding pocket. In the myoV network the shortest *graph distance* of an elastic spring connecting beads i and j to the ATP-binding pocket was defined as $D_{ij} := \min\{D_{ip}, D_{jp}\}$, where p was the index of the network bead corresponding to residue Ser165 from the P-loop motif and D_{ip}, D_{jp}, were the shortest *graph distances* between bead i and bead p, bead j and bead p, respectively. The same normalisation procedure as described for the designed networks was then followed.

To determined the strain communication pathways and chains displayed in main text Fig. 7B, a threshold value of $\tilde{s}_t > 25\%$ was considered for the maximum absolute value of normalised strain of elastic links.

In Movie S9 conformational motions subsequent to ATP-binding and the propagation of strain are visualised. Snapshots taken at different time moments during the simulation are provided in Fig. S5.

Exemplary mutations

A comprehensive analysis of mutations in myoV was beyond the scope of this study. Hence, a few mutations applied to specific spots along the identified communication pathways have been considered and their effect on allosteric coupling checked. As a simplification, similar to the case of the designed networks, a mutation consisted in altering some spring interaction around the mutated residue, or in the removal of the corresponding network bead.

Two mutations involving residue Asn398 from the U50 domain, which is located at the interface to the L50 domain near the actin interaction cleft, have been considered. In mutant 1 an additional spring interaction between Asn398 and residue Cys513 at the interface to the L50 domain was present. In mutant 2 spring interactions from Asn398 to Glu566 and to Lys567 were added. In mutant 1 the allosteric coupling was significantly reduced (dropped by 30%) and in mutant 2 it dropped by 15%. The influence of residue Asn398 on allosteric activity in myosin V can possibly be investigated in site-directed mutagenesis studies.

In the mutant 3 structure the residues Glu445, Ile446, and Asn447, which were located at the interface between the ATP-binding pocket and the L50 domain were deleted, resulting in a substantial reduction of allosteric coupling by 40%.

On the other side the deletion of the short sequence Lys634-Ser643, whose structure was not involved in strain transport through the myoV elastic network, was found to be without considerable consequences for allosteric coupling.

SUPPORTING MOVIES AND DATA

Time-dependent propagation of strain in the random network, the two designed networks, the selected mutant networks, and in the myoV elastic network, are provided as Movies S1-S9. In each of the movies the frame rate is not constant and has been adapted such the fast dynamics inside the allosteric pocket, the inter-domain propagation, as well as the slow motions inside the regulated pocket, can be conveniently followed. The actual time during the simulation is always given at the bottom right in the movies.

- **Movie S1** Conformational motions and strain propagation in the allosterically inactive initial random elastic network.
- **Movie S2** Conformational motions and strain propagation in the designed network with symmetric allosteric coupling.
- **Movie S3** Conformational motions and strain propagation in the designed network with asymmetric allosteric coupling.
- **Movie S4** Conformational motions and strain propagation in the $52-197$ mutant of the designed network with symmetric allosteric coupling.
- **Movie S5** Conformational motions and strain propagation in the $196-197$ mutant of the designed network with symmetric allosteric coupling.
- **Movie S6** Conformational motions and strain propagation in the $52-197-196$ double-mutant of the designed network with symmetric allosteric coupling.
- **Movie S7** Conformational motions and strain propagation in the $52-185$ mutant of the designed network with asymmetric allosteric coupling.
- **Movie S8** Conformational motions obtained from the simulation of the ATP-binding induced rigor to post-rigor state transition in the myosin V motor domain.
- **Movie S9** Conformational motions and strain propagation obtained from the simulation of the ATP-binding induced rigor to post-rigor state transition in the myosin V motor domain.

As Supplementary Data the set of spatial coordinates $\vec{R}_i^{(0)}$ of the initial random network and the two designed networks are provided in the respective equilibrium conformation.

- random_network.dat: text file containing spatial coordinates of the initial random network. 1st column: bead index (0 to 199); 2nd,3rd,4th columns: x,y,z coordinate of the bead position.
- designed_network1.dat: text file containing spatial coordinates of the designed network with symmetric allosteric coupling. 1st column: bead index (0 to 199); 2nd,3rd,4th columns: x,y,z coordinate of the bead position.
- designed_network2.dat: text file containing spatial coordinates of the designed network with asymmetric allosteric coupling. 1st column: bead index (0 to 199); 2nd,3rd,4th columns: x,y,z coordinate of the bead position.

---

[1] Coureux, P.-D., H. Lee Sweeney, and A. Houdusse. 2004. Three myosin V structures delineate essential features of chemo-mechanical transduction. EMBO J. 23:4527-4537.

[2] Alilgan, C., Z.N. Gerek, S.B. Ozkan, and A.R. Atilgan. 2010. Manipulation of conformational change in proteins by single-residue perturbations. Biophys. J. 99:933-943.

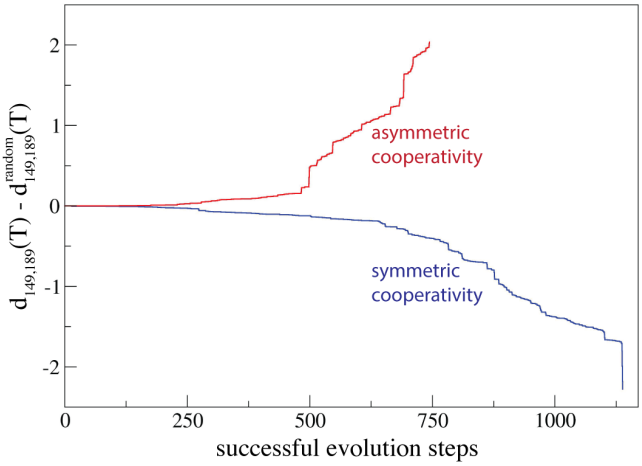
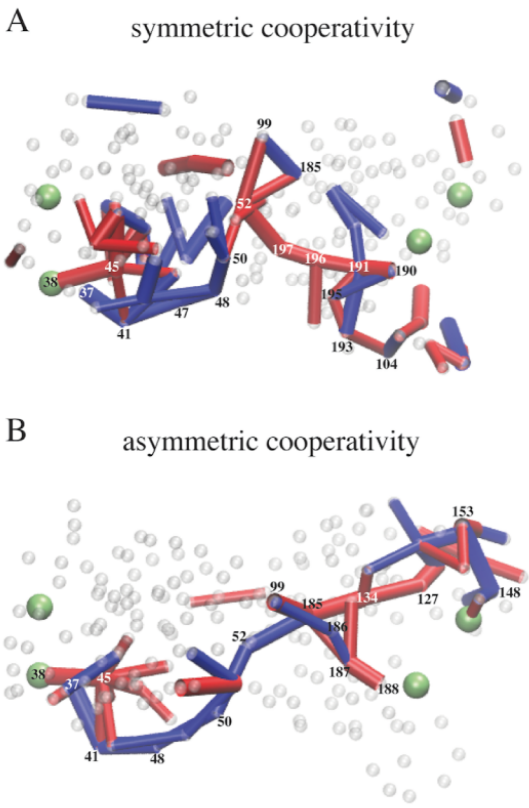


FIG. S 1. Improvement of the allosteric response in the evolving networks during the two independent processes of evolutionary optimisation; $d_{149,189}(T)$ is the allosteric parameter of the elastic network at the current stage of evolution and $d^{random}_{149,189}(T)$ is the value in the initial random elastic network. A successful evolution step corresponded to an accepted mutation, i.e. one after which the allosteric response of the network had improved. The total number of successful evolution steps was $1,138$ in the design process of the network with symmetric cooperativity and 744 in the design process of the network with asymmetric cooperativity, respectively.



	network with symmetric cooperativity		network with asymmetric cooperativity	
single bond deletion	bead indices	robustness coefficient	bead indices	robustness coefficient
	38-45	0.31	38-45	0.73
	37-41	1.00	37-41	0.98
	41-47	1.00	41-48	0.97
	47-48	1.00	48-50	0.96
	48-50	1.00	50-52	0.92
	50-52	1.00	52-185	1.07
	99-185	0.34	185-187	0.16
	52-197	1.00	134-185	1.09
	196-197	0.98	127-134	1.00
	191-196	0.97	148-153	1.00
	190-191	0.97		
	191-193	1.00		
	193-195	1.00		
	104-193	0.33		
single bead deletion	bead index	robustness coefficient	bead index	robustness coefficient
	45	0.33	45	0.79
	41	1.00	41	0.97
	50	0.39	48	0.94
	52	0.07	50	0.87
	99	0.36	52	0.66
	185	0.30	99	0.07
	196	0.96	185	0.03
	197	0.07	186	-0.02
	191	0.92	187	0.36
	105	0.23	188	-0.02

FIG. S 2. and Table S1. **Mutations and robustness of allosteric communication.** A,B): Communication chains in the designed *wildtype* networks are shown in the same perspective as in Fig. 5 (main text), with the beads which were subjected to mutations being indicated by the respective bead indices. Mutations performed in both designed networks are listed in the table on the right side. A colour scale from blue to red shall illustrate neutral to fatal mutations, according to the respective robustness coefficient (defined in the main text Methods section). For the $52-197-196$ double mutant of the network with symmetric coupling (which is not listed in the table) the robustness coefficient was 0.99.

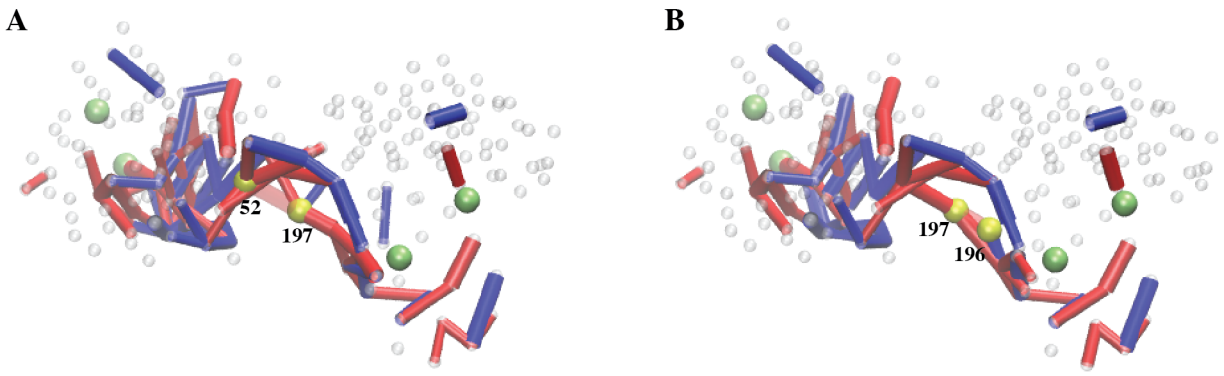


FIG. S 3. **Communication chains in selected mutant structures.** For the designed network with symmetric cooperativity the communication chains in two representative mutants, each in which a single elastic link was deleted, are shown. The $52 - 197$ mutant (A) as well as the $196 - 197$ mutant (B) both retained full allosteric activity, with robustness coefficients of 1.00 and 0.98, respectively (see Table T1). In each mutant the deleted link is shown in transparent and the corresponding beads are highlighted in yellow.

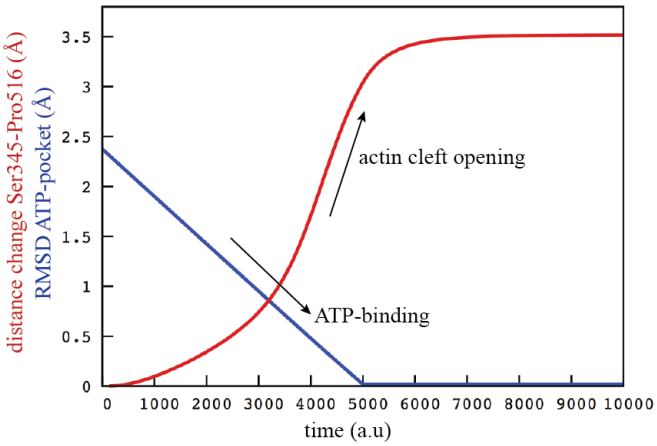


FIG. S 4. **ATP-induced actin-cleft opening.** The RMSD of the nucleotide binding pocket with respect to the target conformation is shown as it changed during the simulation (blue curve). Distance changes between labels in the actin cleft are plotted as red curve.

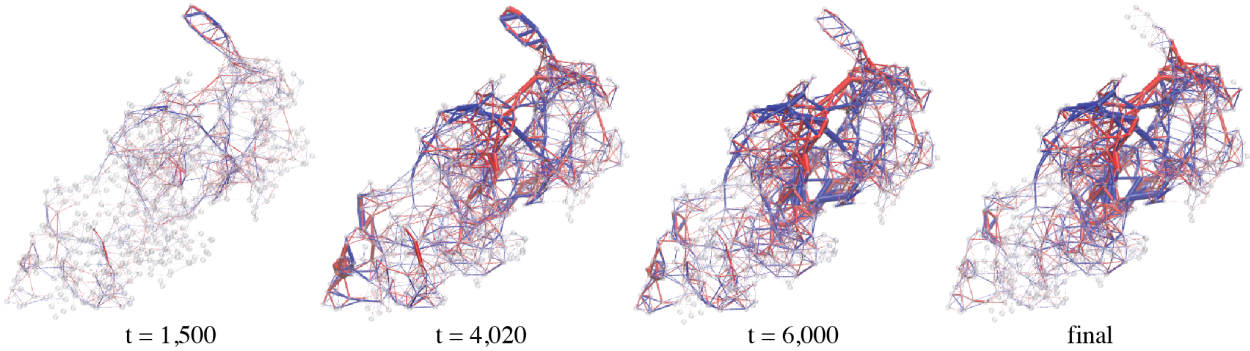


FIG. S 5. **Strain propagation in myoV.** Propagation of strain subsequent to ATP-binding to the nucleotide binding pocket is visualised. Snapshots taken at different time moments during the simulation are shown. Coding of colour and bond thickness is similar to that used for the designed networks (see main text Fig. 4).