Biophysical Journal
Article

Biophysical Society

Chokepoints in Mechanical Coupling Associated with Allosteric Proteins: The Pyruvate Kinase Example

Lewis E. Johnson,[1,2] Bojana Ginovska,[2] Aron W. Fenton,[3] and Simone Raugei[2,*]
[1]Department of Chemistry, University of Washington, Seattle, Washington; [2]Physical and Computational Sciences Directorate, Pacific Northwestern National Laboratory, Richland, Washington; and [3]Department of Biochemistry and Molecular Biology, The University of Kansas Medical Center, Kansas City, Kansas

ABSTRACT Although the critical role of allostery in controlling enzymatic processes is well appreciated, there is a current dearth in our understanding of its underlying mechanisms, including communication between binding sites. One potential key aspect of intersite communication is the mechanical coupling between residues in a protein. Here, we introduce a graph-based computational approach to investigate the mechanical coupling between distant parts of a protein, highlighting effective pathways via which protein motion can transfer energy between sites. In this method, each residue is treated as a node on a weighted, undirected graph, in which the edges are defined by locally correlated motions of those residues and weighted by the strength of the correlation. The method was validated against experimental data on allosteric regulation in the human liver pyruvate kinase as obtained from full-protein alanine-scanning mutagenesis (systematic mutation) studies, as well as computational data on two G-protein-coupled receptors. The method provides semiquantitative information on the regulatory importance of specific structural elements. It is shown that these elements are key for the mechanical coupling between distant parts of the protein by providing effective pathways for energy transfer. It is also shown that, although there are a multitude of energy transfer pathways between distant parts of a protein, these pathways share a few common nodes that represent effective "chokepoints" for the communication.

INTRODUCTION

Allosteric regulation of an enzyme can be viewed as an equilibrium process (1,2) in which an enzyme binds a substrate differently depending on whether an effector is present or absent at a distant site (3,4). Protein dynamics can influence this process and has recently become a focus in studies of allosteric regulation of an enzyme (5–8). Indeed, like the contributions of protein dynamics to other biological functions (9–13), perturbations to an equilibrium protein's dynamics can influence mechanical coupling between different sites (14). Binding of an effector to an allosteric site (1,15), electrochemical activity at a given location such as electron transport through respiratory complexes (16), photoisomerization of a cofactor (17), and slow, large-scale motions of flexible moieties (18) can induce changes in the energy distribution of a protein that, in turn, induce an effect in function at a remote site. One way to characterize dynamic contributions to allostery is to evaluate the mechanical coupling between various residues or structural elements through protein motion and the role of the protein scaffold in channeling kinetic energy through these elements.

Understanding mechanical couplings between sites in a macromolecule requires identifying both 1) the coupled sites and 2) plausible networks of residues through which the sites are mechanically coupled. Coupled sites can be identified by conventional methods for assessing macromolecular dynamics, such as root mean-square fluctuation of atomic positions and principal component analysis (PCA) of the covariance of interatomic motions (19,20), along with systematic mutation studies (21). Normal-mode analysis, whether derived from molecular dynamics (MD) or from coarse-grained models, such as an anisotropic network model (ANM), can provide insight about collective motions capable of transferring energy between sites, as can NMR-derived methods (22). Several methods have been developed for identifying potential allosteric sites based on covariance of atomic motions and normal-mode analysis, including a

Submitted January 28, 2019, and accepted for publication March 21, 2019.
*Correspondence: simone.raugei@pnnl.gov
Editor: Monika Fuxreiter.
https://doi.org/10.1016/j.bpj.2019.03.026

method by Balabin et al. based on the coupling between local structural perturbations and normal modes (8) and a method by Guarnera and Berezovsky (15) in which potential allosteric sites are stiffened to represent the presence of a ligand and normal-mode free energies are compared. A similar method was also previously reported (23).

A complete characterization of the contributions of mechanical couplings to an allosteric mechanism will require a comparison of the mechanical coupling evaluated for the enzyme in complex with all possible combinations of substrate and allosteric effector binding. For a protein with one active site and one allosteric site, four structures would be required to define the allosteric energy cycle: 1) free enzyme, 2) enzyme-substrate complex, 3) enzyme-effector complex, and 4) the ternary substrate-enzyme-effector complex (2). With such data available, an analysis of mechanical coupling could be initiated for respective cocrystalized enzyme complexes to determine how binding of a ligand alters potential mechanical coupling pathways. However, a key initial step is to evaluate the mechanical couplings that exist in currently available protein structures.

Graph-based methods provide an intuitive route for addressing mechanical coupling. Complex chemical systems can be converted to a reduced, graph-like representation of their connectivity based on interatomic distances and other structural parameters, in which nodes are formed from atoms or collections of atoms (e.g., protein residues), and edges connecting them represent interactions (bonded or noncovalent). The resulting graphs can be traversed to identify local structural features, such as noncovalent clusters (24), cyclic structures in glasses (25), or paths between points that minimize a parameter of interest. Graph-based methods have had a long history in analyzing protein structure (26). For instance, a shortest-path (parameter minimization) approach has been applied by Beratan and co-workers for analyzing electron transfer pathways in proteins (27). Graph-based methods (28) have also previously been used for characterizing allosteric coupling in proteins (7), whether based on normal-mode analysis (29–31), PCA (32), NMR-derived covariances of interatomic motions (33), discrete MD-derived contacts and covariances of interatomic motions (7), sampling of multiple MD trajectories (6,17,34,35), clustering based on energies of side-chain interactions (36–38), or evolutionary optimization of ANM-type models (39). Building upon this work, we report a novel, to our knowledge, graph-based method for representing coupling between sites that can be applied to either a crystal structure (via ANM-generated covariances of interatomic motions) or MD data, using either an all-residue or secondary structure representation, that generates a two-dimensional map of mechanical coupling, analogous to those generated by Balabin's method (8). Our principal goals in developing a new graph-based method were 1) to be able to efficiently determine the extent to which structural elements with correlated motion are mechanically coupled and 2) to identify the key structural elements involved in the transduction of the mechanical force between sites, which may proceed through a large ensemble of pathways involving those residues.

Consistent with our interest in allosteric regulation, our primary model system for evaluating our graph-based method for predicting mechanical coupling through proteins is the allosterically regulated human liver pyruvate kinase (hL-PYK). We chose pyruvate kinase because it is a soluble, globular protein (which avoids having to account for the presence of a membrane in the model) with multiple well-characterized allosteric sites (Fig. 1) (21,40–42). Although the enzyme catalyzes a phosphate transfer from phosphoenolpyruvate (PEP) to ADP to form ATP, only the PEP affinity is allosterically regulated. Fructose bisphosphate (FBP) causes increased PEP affinity (43), whereas the alanine effector (Ala^E will be used to distinguish when alanine is the effector ligand) reduces PEP affinity (42). Tang and Fenton recently characterized the contribution of nearly all residues to allosteric regulation in the protein by a systematic mutation study (alanine scan; Ala^M will be used to distinguish when alanine is a replacement amino acid in the protein) (21).

Using hL-PYK as a testbed, first, we demonstrate how our graph-based representation can identify pairs of residues (protein sites or secondary structure elements) that have correlated motions and are mechanically coupled via specific pathways throughout the protein structure. A key advance of our method is the ability to filter out pairs of residues whose covariant motion is purely coincidental. To examine broader applicability, we also apply our method to two membrane proteins previously examined using Balabin's method (8). Taking a step further, we will demonstrate how it is possible to identify residues that are critical for mechanical coupling between two distant parts of a

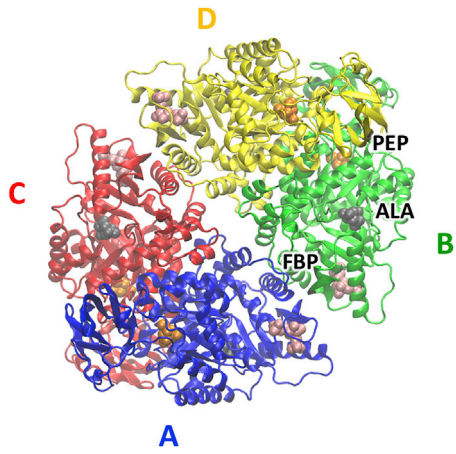

FIGURE 1 Cartoon representation of the structure of hL-PYK (PDB: 4IMA). The binding sites of the phosphoenolpyruvate substrate (PEP, in *orange*) and the activator fructose-1,6-bisphosphate (FBP, in *pink*) and the inhibitor alanine (Ala^E, in *gray*) binding site are highlighted for each chain. The four chains in the protein are color coded and labeled A, B, C and D on the figure. To see this figure in color, go online.

protein. Specifically, we will show that mechanical coupling between binding sites occurs through an ensemble of pathways between the sites. These pathways share a few crucial, common nodes that represent effective "chokepoints." Because these nodes are central to many mechanical coupling pathways, any effort to rationally design proteins with altered dynamics could focus on modifying them. Furthermore, given the observed high correlation between the mechanical couplings and residue positions that contribute to allostery (21), we can speculate that targeting these chokepoints may be an effective way to alter allosteric regulation in enzymes.

METHODS

The proposed graph-based representation of the mechanical coupling is introduced using the crystal structure of the hL-PYK tetramer crystallized with FBP bound, with citrate present in the PEP binding site and with the Ala^{E} site empty (Protein Data Bank (PDB): 4IMA). Because of unresolved residues, particularly in chain D, the crystal structure tetramer was reconstructed and symmetrized based on the nearly complete chain A. Symmetrization has little to no effect on the calculated properties, as discussed in the Supporting Materials and Methods, including Fig. S2.

The method treats each residue as a node on a weighted, undirected graph, in which the edges are defined by locally correlated motions of those residues and weighted by the strength of the correlation. The workflow for building such a graph and extracting quantitative information about the mechanical coupling between distant sites is illustrated in Fig. 2. We start from two input quantities: 1) a correlation matrix (normalized covariance matrix of interatomic motions) **C** (Fig. 3, *top*) and 2) a minimal (Euclidean) distance matrix **D** (for example, see Fig. S3, *first panel*). The correlation matrix can be obtained from an ANM, an MD trajectory, or from experimental (e.g., NMR) data. Thus far, our method has been used with normalized correlation data; however, in principle, it could be used with non-normalized covariances as well. We chose normalized correlation data to facilitate a bounded, linear, and consistent dynamic range that would allow a common set of parameters to be applied to multiple model systems. The correlation matrix elements C_{ij} represent the extent to which the motion of residues i and j are correlated, with a value of 1 indicating perfectly correlated (parallel) motions, a value of -1 indicating perfectly anticorrelated (antiparallel) motions, and a value of 0 indicating uncorrelated (orthogonal) motion. The correlation matrices discussed here were based on C_α ANM models derived from crystal structures and calculated using ProDy (44) using default settings. The distance matrix is obtained from the same structural data.

For a large protein, data on interatomic motions and distances can also be transformed to a further coarse-grained representation, such as using secondary structural elements (identified by STRIDE (45) or similar methods) instead of residues as the matrix elements. Such coarse-graining is motivated by strong internal correlation within structural components in a protein. This allows us to combine many redundant motions into a more readily analyzable set of center-of-mass motions. If a structure is reduced to secondary structural elements, C_{ij} are averaged for each set of residues being merged to generate the reduced correlation matrix $\mathbf{C}_{\mathrm{red}}$, and the smallest distance matrix element D_{ij} between residues is selected for each pair of secondary structural elements to generate the reduced distance matrix $\mathbf{D}_{\mathrm{red}}$. All the subsequent steps in the method are identical whether or not reduced representations are used, and from hereon we will refer to the matrices only as **C** and **D**.

The first step in the process is the generation of an adjacency matrix **A** (Eq. 1) that represents the network of correlated motions in the protein, where

$$A_{ij} = 1 - \left|C_{ij}\right|. \tag{1}$$

Our definition of the adjacency matrix **A** is identical to the definition developed by Dokholyan (7,46). This initial adjacency matrix includes all covariant motions in the protein, regardless of distance or strength, and needs to be pruned to remove interactions that do not involve strong coupling between adjacent residues or structural elements. The pruning is performed using the rule

$$A_{ij} = 0 \text{ if } C_{ij} < \left(\mu_{|\mathbf{C}|} + \sigma_{|\mathbf{C}|}\right) \text{ or } D_{ij} < R_{\mathrm{cut}}, \tag{2}$$

where $\mu_{|\mathbf{C}|}$ and $\sigma_{|\mathbf{C}|}$ are the mean and SD of the elements of **C**, respectively, and R_{cut} is the contact distance cutoff. We set R_{cut} at 7.5 Å based on minimal

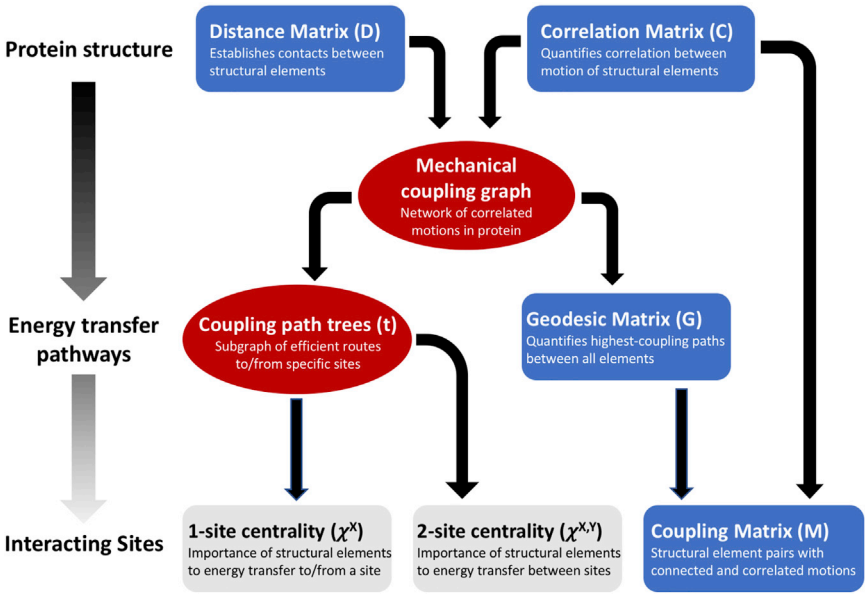

FIGURE 2 Work flow of the protocol for characterizing the mechanical couplings in protein. The distance matrix (**D**) and the normalized covariance (correlation) matrix (**C**) are used to build a mechanical coupling graph. Nodes of the graph represent individual residues, secondary structural elements, or protein domain. The graph can then be traversed to find the highest-coupling (geodesic parameter minimization) route between each pair of nodes (**G** matrix); the mechanical coupling matrix (**M**) can then be calculated based on the correlation matrix and path length matrix, as can the one-site and two-site centrality parameters, χ^X and $\chi^{X,Y}$. To see this figure in color, go online.

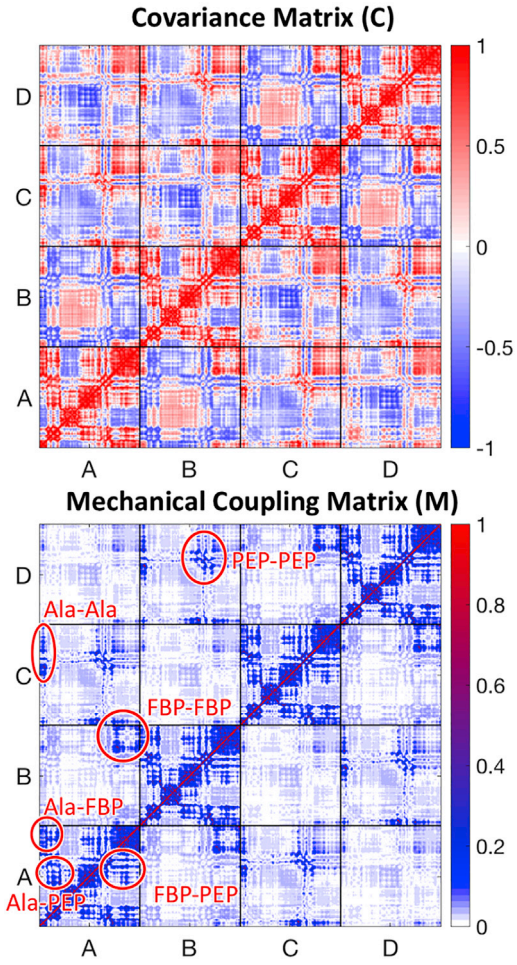

FIGURE 3 (*Top*) All-residue covariance matrix (**C**-matrix) of interatomic motions and (*bottom*) all-residue coupling matrix, **M**, for hL-PYK derived from ANM model using crystal structure (PDB: 4IMA). The resulting **M**-matrix is qualitatively identical with the secondary structure representation (see Supporting Materials and Methods). Coupling can be seen between the alanine and PEP sites, plus between the FBP and PEP sites within each chain, as well as between the FBP sites on two pairs of chains (A and B, C and D), between alanine sites on two pairs of chains (A and C, B and D), and between the PEP sites on two pairs of chains (A and C, B and D). Comparison of two matrices highlights how the path-based weighting (at the heart of the definition of P) substantially reduces noise to better highlight potential allosteric interactions. To see this figure in color, go online.

distance before more than one disconnected node emerges in the graph in any of the systems studied. This distance is consistent with the cutoff distance used by Dokholyan for a related method based on local covariance of interatomic motion (7). Sensitivity analysis to selection of different cutoff distances is reported in Fig. S1. In addition to the minimal correlation and maximal distance cutoffs, our model also applies a maximal coupling threshold in which any correlations $|C_{ij}| > C_{thresh}$ are set to C_{thresh}. We set C_{thresh} to 0.95, representing an approximate value of the correlation within secondary structural elements. The threshold means that nonzero (connected) elements of the adjacency matrix have a lower bound of $A_{thresh} = 1 - C_{thresh}$, which prevents numerical instabilities when correlations approach unity and enables the coupling matrix (described below) to be normalized. Self-loops (diagonal elements in the adjacency matrix) are also removed.

The adjacency matrix is a two-dimensional representation of an undirected graph of the mechanical coupling within the protein (47). Once the graph is defined (Fig. 4), a standard shortest-path (geodesic) method (for example, the "distances" function in MATLAB; The Mathworks, Natick, MA) can be used to find the paths of maximal local covariance of interatomic motion between each pair of elements in the graph to generate the geodesic matrix **G** (Fig. S3, *lower left*), where G_{ij} is the sum of all edge lengths along the path from element i to element j in the graph. With this information at hand, we can identify structural elements pairs that have correlated and connected motions through the mechanical coupling matrix **M** (Fig. 3, *bottom*), whose elements are defined as

$$\mathrm{M}_{ij} = A_{\mathrm{thresh}} \frac{|C_{ij}|}{G_{ij}}, \tag{3}$$

with the diagonal elements M_{ii} set to zero to remove trivial coupling of residues with themselves. A maximal M_{ij} value of unity indicates two adjacent and perfectly correlated residues, whereas a minimal value of zero indicates no significant coupling between residues. The major difference between the **M**-matrix and simple covariance analysis of the motion is that if two residues have correlated motions but do not have an efficient path for mechanical coupling between them (e.g., a large G_{ij}), a small (poorly coupled) M_{ij} will be produced. In this way, geodesic weighting suppresses motions that are merely coincidental. Our process for generating the **M**-matrix and other descriptors can be applied to any protein for which three-dimensional structural information is available. In addition to generating the **M**-matrix, the graph can also be used to calculate standard graph theory metrics such as rank, centrality, path trees (shortest path from a node i to all nodes $j \neq i$), etc. These metrics can be used in conjunction with the **M**-matrix to analyze the coupling within a protein, including the cost-weighted betweenness centrality metrics χ^{X} and $\chi^{X,Y}$ discussed in the next section. The combination of covariance of interatomic motions and contact information from MD trajectories (the information abstracted into the **G** matrix) have previously been applied to determining mechanical coupling pathways by Vishveshwara et al. (48) Our alternative approach is based on condensing the graph representation of the protein into subgraphs that contain only strongly coupled pathways in or out of a site.

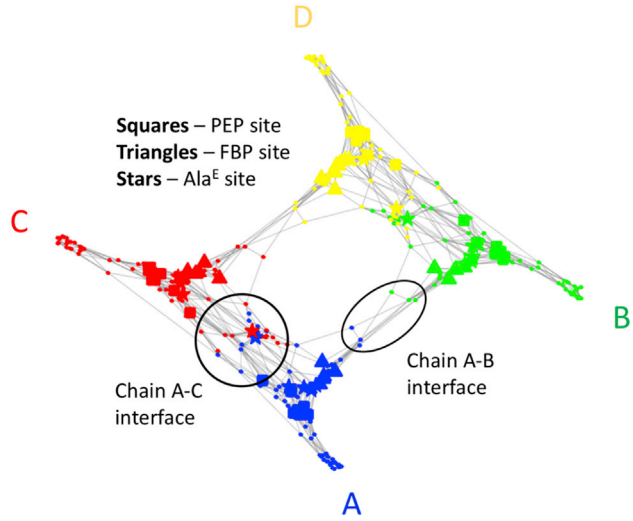

FIGURE 4 Mechanical coupling graph in hL-PYK. Nodes represent secondary structural elements, and edges represent interactions with high local covariance. The edges of the graph represent interactions with high local covariance. The PEP site is noted by squares, the effector FBP site is noted by triangles, and the inhibitor alanine (Ala^{E}) site is noted by stars. To see this figure in color, go online.

RESULTS AND DISCUSSION

Identification of mechanically coupled sites in proteins

We start our discussion by analyzing the mechanical coupling between distant parts of pyruvate kinase as identified by the **M**-matrix (*lower panel* of Fig. 3). Within each chain, the structural elements involved in both allosteric sites are strongly coupled to the active site. In addition, the Ala^E site, which is located between the PEP and FBP sites, couples with both. Comparing the **C**-matrix and the **M**-matrix (both shown in Fig. 3), it is evident that the **M**-matrix filters out coincidental correlations in interatomic motion that are prevalent in the **C**-matrix. The underlying structure of the coupling can be seen by examining the mechanical coupling graph, shown in Fig. 4 and reduced to secondary structural elements for clarity. In this figure, structural elements comprising the active site (PEP site) on each chain are indicated using squares; the FBP site (40) is indicated with triangles and the Ala^E site (42) with stars. Proximity on the graph indicates strength of interaction between nodes, with clustered nodes having a greater number of strongly coupled paths between them. The corresponding **M**-matrix for the reduced system of secondary structure elements is shown in Fig. S3, along with the associated **C**- and **G**-matrices, and is qualitatively identical to the all-residue matrix shown in Fig. 3. We studied the energy transfer pathways in hL-PYK at two levels of abstraction: 1) treating the α carbon of each residue as a node and 2) coarse-graining the system by treating the center of mass of each secondary structural element as a node. The all-residue and secondary structural analyses were qualitatively equivalent, as can be seen by comparing Figs. 3 and S3. All-residue results are used for comparing with systematic mutation results, whereas the secondary structure representation is adopted for clarity of visualization of the coupling graph.

Communication networks and chokepoints

The **M**-matrix introduced in the Methods and analyzed in the previous section identifies residues, sites, or secondary structure elements that are mechanically coupled, i.e., parts of a protein or protein complex that have correlated motions and are coupled by efficient (energy transfer) pathways between those endpoints. However, the **M**-matrix does not directly provide any insight on the identity of the sites (residues or secondary structural elements) that contribute to the mechanical coupling between two distant parts because these quantities are abstracted into the **G** matrix. It is possible to obtain this crucial information by calculating the shortest-path tree $\mathbf{t}^{(k)}$ (see Fig. S4 for an example) from each residue in the given site X (e.g., the active site X = FBP) to all other residues in the protein and merging the resulting N_{trees} trees to form a graph of strong coupling routes to or from the site X by summing the inverse of the adjacency matrix elements $t_{ij}^{(k)}$ of the trees and inverting to form the combined adjacency matrix **T**,

$$T_{ij} = \left(\sum_{k=1}^{N_{\text{trees}}} \frac{1}{t_{ij}^{(k)}} \right)^{-1}. \quad (4)$$

The contribution of each residue to conveying a perturbation to or from the site X is then quantified by calculating the betweenness centrality (49,50) (the ratio of number of shortest paths between nodes $i, j \neq k$ that pass through node k over the total number of shortest paths between nodes $i, j \neq k$) of all residues on the graph, weighted by inverse cost to pass through them given by the elements of the matrix **G**. We will refer to the resulting quantity as one-site cost-weighted betweenness centrality (χ). Betweenness centrality is a measure of a node's influence on communication between other nodes (51). It has previously been used as a metric for predicting key proteins in metabolic and protein networks (52). In our work, shortest-path trees and centralities were calculated using the Graph toolbox in MATLAB r2018a; codes for calculating these parameters are included as Data S1, S2, S3, and S4. The one-site cost-weighted betweenness centrality parameter χ^{FBP} is shown in the upper panels of Fig. 5, mapped onto the underlying graph (*left*) and the structure of the protein (*right*). A large χ_k^{X} value indicates that a large number of pathways pass through node k. Therefore, perturbing that node dramatically influences how the mechanical energy from global motions of the protein is transferred to or from the site X. Similarly, we can define the two-site cost-weighted betweenness centrality $\chi_k^{\text{X,Y}}$ between two sites X and Y. A large $\chi_k^{\text{X,Y}}$ value indicates that a large number of pathways between those sites pass through node k. Therefore, perturbing that node is expected to dramatically influences how mechanical energy is transferred between the two sites. The two-site cost-weighted betweenness centrality parameter ($\chi^{\text{FBP,PEP}}$) is shown in the middle panels of Fig. 5, mapped onto the underlying graph (*left*) and the structure of the protein (*right*). In the rest of the work, we will show how the one-site, χ^{X}, and two-site, $\chi^{\text{X,Y}}$, cost-weighted betweenness centralities are able to provide a nearly quantitative understanding of the alanine-scanning data for hL-PYK (21).

The lower panel of Fig. 5 shows χ^{PEP} for the PEP site mapped onto chain A, with the most influential residues noted in red, moderately influential in blue, and noncontributing in white. Coupled residues are observed between the PEP site and both allosteric sites. In addition to the within-chain coupling, the FBP site on each chain is coupled with the neighboring FBP site (chain A to chain B or chain C to chain D). Additionally, the Ala^E site on each chain is coupled with the neighboring Ala^E site (A to C or B to D); couplings are shown in the **M**-matrix in Fig. 3, and

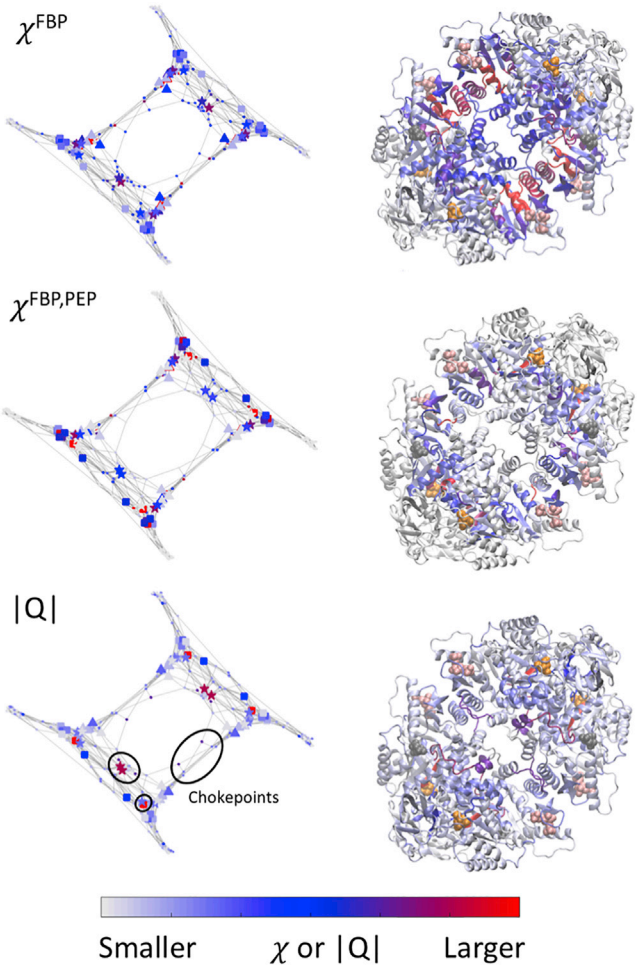

FIGURE 5 Comparison of cost-weighted betweenness centrality parameters χ^{FBP} (*top row*) and $\chi^{FBP,PEP}$ (*middle row*) for FBP interacting with hL-PYK with the allosteric strength |Q| as derived from an alanine scan (21) mapped onto the secondary structure of the protein. The left column shows the data mapped onto the graph representation of the secondary structure generated using the above methods, and the right column shows the same data mapped onto the protein structure. Because the χ parameters provide no information that would indicate the sign of a potential allosteric effect, the $|Q_{ax}|$ parameters represent the factor by which the allosteric effect was observed to change. Substantial overlap is observed between $\chi^{FBP,PEP}$ and |Q|, though several structures in the interior of the protein identified as crucial for allostery by the alanine scan are instead highlighted by χ^{FBP}. Three chokepoint regions highlighted by multiple parameters are also shown, including residues at the A-B and A-C interfaces as well as the ^{370}Cys-^{373}LEU β-sheet in the PEP site. To see this figure in color, go online.

the raw cost-weighted betweenness centrality (without specifying any sites as sources or sinks) is shown in the upper panel of Fig. 5. These interchain couplings involve just a small number of "chokepoints" (*purple* or *red* in Fig. 5) that have few or no significant alternative routes. We can speculate that communication across these chokepoints between chains could provide a mechanism for binding cooperativity (53) between the allosteric sites on different chains. An example of modulation across a chokepoint would be a binding event displacing residues, such as motion of the helix that contains Asp^{499}, Arg^{501}, and Trp^{494} in response to an FBP binding that would change the extent of the interaction of this helix with the Trp^{527} loop on the adjacent subunit, as proposed by Fenton and co-workers (40). This between-chain interaction between FBP sites is shown in the **M**-matrix in Fig. 3.

Modified allosteric regulation resulting from a whole-protein alanine-scanning study

In an alanine scan, mutants of the protein are produced in which each nonalanine and nonglycine residue is individually mutated to Ala^M and the allosteric coupling of the protein is compared to the wild type. The allosteric coupling is defined as

$$Q_{a/x} = \frac{K_a}{K_{a/x}}, \tag{5}$$

where K_a is the apparent affinity of the protein for the substrate (e.g., PEP) in the absence of the effector (e.g., FBP or Ala^E) and $K_{a/x}$ is the apparent affinity of the protein for the substrate at an effector concentration that results in saturation of the allosteric site. The effect of a mutation on one type of allosteric coupling is assessed by the ratio Q_i/Q_0, where Q_i is $Q_{a/x}$ for a protein with an Ala^M substitution at residue i and Q_0 is $Q_{a/x}$ for the wild-type protein. For FBP, which allosterically promotes the affinity of hL-PYK for PEP, $Q_i/Q_0 < 1$ indicates that allosteric promotion by FBP is decreased, and $Q_i/Q_0 > 1$ indicates that allosteric promotion by FBP is increased by mutation of that residue to Ala^M. Conversely, for Ala^E as a ligand, which allosterically inhibits the affinity of hL-PYK for PEP, $Q_i/Q_0 < 1$ indicates that allosteric inhibition by Ala^E is increased, and $Q_i/Q_0 > 1$ indicates that allosteric inhibition by Ala^E is decreased by mutation of that residue to Ala^M.

Because the Q_i/Q_0 quantifies the influence of a single residue mutation on the allosteric regulation between two binding events, this single parameter can easily be correlated with computational metrics based on residue mobility or interactions (i.e., the mechanical coupling metrics discussed in the prior sections). However, whereas the ratio Q_i/Q_0 indicates whether a mutation of a site induces an increase or a decrease of the allosteric effect, the cost-weighted betweenness centrality only informs us that a site is important for the mechanical coupling between distant parts of the protein, and it does not provide any direct information regarding biochemical activity (i.e., increase or decrease). As such, we focus our comparison on the magnitude of the change in allosteric activity |Q|, which is shown for FBP in the lower panels of Fig. 5, mapped onto both the graph representation of the protein (*left*) and the structure (*right*), highlighting the regions of the protein that are key to allostery and their correspondence with pathways by which sites can be mechanically coupled. A corresponding representation for Ala^E is shown in

Fig. S5; the underlying numerical values are also available as Table S1.

For alanine as an effector (Fig. 6, *bottom left*), substitutions in the Ala^E binding site (*asterisks*) reduce the allosteric inhibition, substitutions adjacent to the phosphate binding loop (residues 444–449) in the FBP site enhance the Ala^E allosteric effect, and substitutions adjacent to the allosteric loop (residues 527–533) in the FBP site reduce the Ala^E allosteric effect (inhibition), as do substitutions of some polar residues in the PEP site regions (e.g., at His^{286}, Thr^{340}, and His^{391}). Far broader regions of substitution affect FBP (Fig. 6, *bottom right*) binding, with substantial shifts occurring because of substitutions within the FBP, PEP, and Ala^E sites along with other residues between them.

Correlation between mechanical coupling and allostery

A strong Ala^M substitution effect for a residue correlates well with the one-site χ^X and the two-site $\chi^{X,PEP}$ cost-weighted betweenness centralities of that residue with $X = Ala^E$ or FBP (Fig. 6, *middle* and *top panels*). As can be seen from Fig. 6, the $\chi^{X,PEP}$ centralities are able to capture nearly the totality of the major features in the Q_i/Q_0 trends for both $X = Ala^E$ and FBP (highlighted as *green areas* in Fig. 6). Nevertheless, some features in Q_i/Q_0 trends are absent in $\chi^{X,PEP}$ (highlighted as *yellow areas* in Fig. 6). These features are instead present in χ^X. This finding implies that the majority of the Ala substitutions with strong effects on Q_i/Q_0 are located along pathways of mechanical communication between the effector sites and the PEP site. Interestingly, some Ala^M substitutions that alter the Q_i/Q_0 ratios are located along pathways that convey energy between two Ala^E sites or between two FBP site but are not along paths to the PEP site. These positions are only reflected in χ^{Ala} or in χ^{FBP}. This set of Ala^M substitutions are consistent with the hypothesis that an Ala^M substitution can alter the cooperativity of effector binding. It has previously been shown that change in the cooperative binding of either effector or substrate can indirectly influence $Q_{a/x}$ (54) and therefore the Q_i/Q_0 ratios used in this study. This same type of three-ligand interaction can be observed when the binding of a ligand induces cooperativity in a binding response for a second ligand (55). The possibility of a

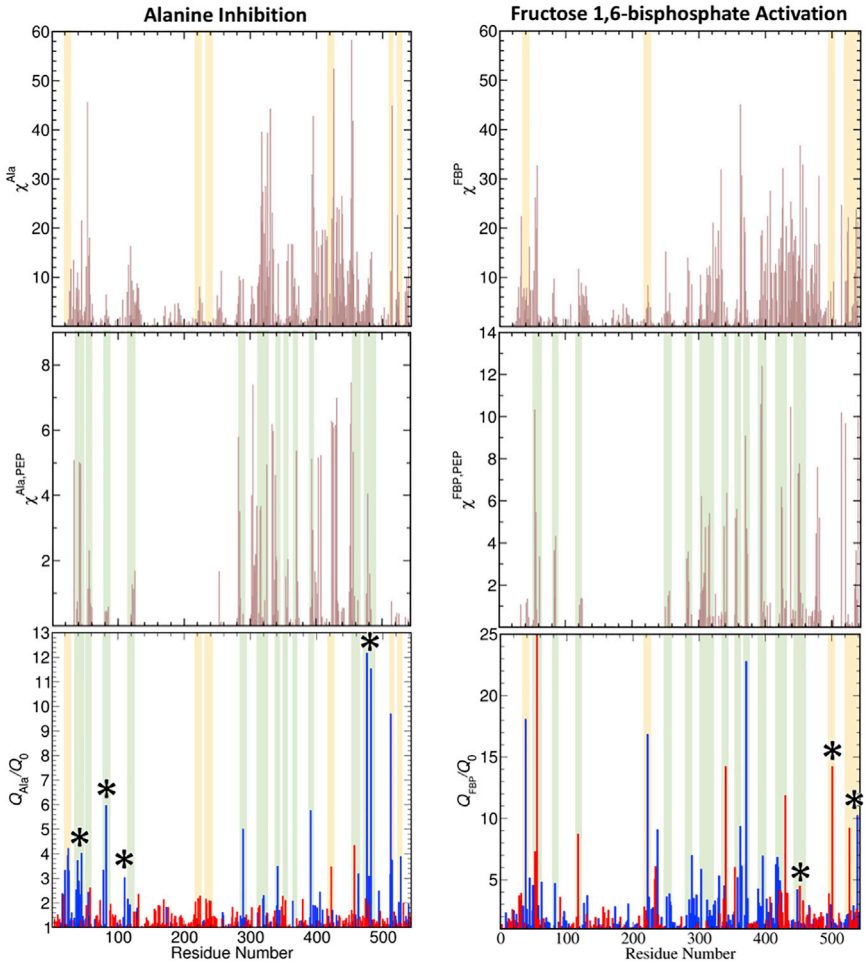

FIGURE 6 Comparison of one-site, χ^X, and two-site, $\chi^{X,PEP}$, cost-weighted betweenness centrality (likelihood of a node being on a path leading to $X = Ala^E$ and FBP or connecting X to the PEP site, weighted by the degree of coupling along the path) elements for the PEP site on hL-PYK (*light red line*) with the Q_i/Q_0 values from the alanine scan (*bottom panels*) (21). In the bottom panels, blue bars for Q_i/Q_0 indicate that the mutant has a reduced response to Ala^E (*left*) or increased response to FBP (*right*), whereas red bars are inverse values and indicate that the mutant has an increased response to Ala^E (*left*) or reduced response to FBP (*right*). The color convention is from Tang and Fenton and reflects Ala^E as an allosteric inhibitor and FBP as an allosteric activator (21). The green and yellow areas indicate common features between Q_i/Q_0 and $\chi^{X,PEP}$ or χ^X, respectively. Asterisks (*) mark residues within either the Ala^E or FBP allosteric site. χ^X and $\chi^{X,PEP}$ values are scaled by a factor of 1/1000. To see this figure in color, go online.

three-ligand coupling (i.e., $Q_{a/x}$ dependence on concentration of a third ligand) was also acknowledged in the original whole-protein Ala^M-scanning mutagenesis study (21). Thus, a secondary benefit of the mechanical coupling versus Ala^M-scanning data may be a mechanism to predict these indirect influences on allosteric mechanisms. We also note that the proposed two-site cost-weighted betweenness centralities are able to predict the larger magnitude of the Ala^M substitution effect for FBP than for Ala^E.

Another key observation from the alanine-scanning data that is further backed by cost-weighted betweenness centrality data is a strong degree of complementarity between residues involved in pathways to each allosteric site (e.g., substitutions that reduce the Ala^E allosteric effect often boost the FBP effect and vice versa). This complementarity is further shown in Fig. S6. This observation supports the hypothesis of key "chokepoint" residues influencing allostery by altering mechanical coupling through the protein.

Our results demonstrate that, using readily available structural data, a graph-based method can provide predictions on the importance of residues in allosteric coupling that correlate well with rigorous data from mutation-based studies, both for effects on the active site and between the allosteric sites. However, these correlations come with several caveats, that we discuss below. As previously stated, contributions of mechanical coupling to an allosteric mechanism would best be evaluated by comparing changes in observed mechanical coupling pathways between the four enzyme complexes that define the allosteric energy cycle (i.e., free enzyme, enzyme-substrate complex, enzyme-effector complex, and the ternary substrate-enzyme-effector complex). Those four structures are not currently available for hL-PYK. As an alternative, we evaluated the mechanical couplings in one available crystal structure and compared the identified mechanical couplings to the full-protein alanine-scanning mutagenesis data. Relying on a single crystal structure presents the potential to introduce several biases into the mechanical coupling evaluation. Allosteric coupling ($Q_{a/x}$) in the wild-type hL-PYK protein is pH dependent (43), with an increase in regulation as pH increases from 6.0 to 8.0 but negligible allosteric regulation below pH 6.0. The available structures of hL-PYK have been determined at pH below 6.0 (56). The calculations of the mechanical coupling are dependent on the topology included in the structure. Therefore, if the topology of the protein is maintained over the pH range (i.e., the pH dependence of allostery is due to an altered ionization state) and if ligand binding events do not substantially modify that topology, then the mechanical couplings should be maintained. However, ligand-induced changes in topology (and therefore changes in mechanical coupling) could themselves contribute to allosteric mechanisms. Given these considerations, the strength of this study is the correlation of the calculated mechanical couplings with experimental results.

Validation of generality

Although our study primarily focused on hL-PYK, we also performed further validation of the model by comparing prediction of allosteric sites in two G-protein coupled receptors, human β-2-adrenoreceptor (B2AR) and bovine rhodopsin, with predictions from Balabin's method (8). Like Balabin's method, this **M**-matrix approach provides a two-dimensional representation of potential interactions between portions of a protein. **M**-matrices for B2AR and bovine rhodopsin are shown in Fig. 7. Regions in which both analyses indicated interactions are indicated by orange triangles. Only two regions flagged by the Balabin analysis were not detected by the **M**-matrix analysis (indicated by *red X*); however, the **M**-matrix analysis identified coupling involving residues adjacent to these regions. One significant difference between the methods is that the **M**-matrix showed far fewer potential interactions in the terminal regions. Because both proteins were modeled in the absence of a membrane environment using both analyses, it is

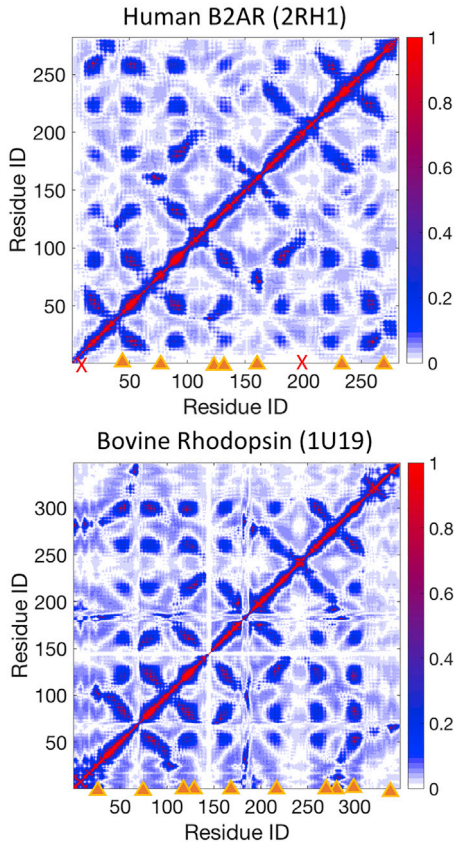

FIGURE 7 Coupling matrices derived from ANM models of human B2AR (PDB: 2RH1) (*top*), and bovine rhodopsin (PDB: 1U19) (*bottom*). Results are consistent with Beratan and co-workers' method (8), with the exception that the **M**-matrix method shows less sensitivity to fluctuations of the high-mobility regions near the C- and N- termini of the proteins. Orange triangles indicate regions identified by both methods; red X symbols indicate regions indicated as coupled in (8) that are not found coupled by the **M**-matrix method. To see this figure in color, go online.

difficult to assign significance to this difference. However, the difference is likely due to the large amplitude of terminal residue motions in the ANM model, which receives much higher weight in a normal-mode-based model as opposed to a local-coupling-based model like the **M**-matrix.

The utility of our method is further illustrated by calculating $\chi^{X,Y}$ between the active site of each receptor (AS) and the G-protein binding site on each receptor. The latter contains a conserved E(D)RY sequence (residues 102–104 (α130–α132) for B2AR and 134–136 for rhodopsin, setting the first residue in the crystal structure as 1) (8). Calculated $\chi^{AS,E(D)RY}$ values are shown mapped onto both an all-residue graph and the protein backbone structure in Fig. 8. Clear coupling pathways are observed in both cases, with a sequence of high-centrality residues propagating up an α-helix from the E(D)RY moiety and bifurcating into multiple routes into the active site (retinal site in rhodopsin, epinephrine site in B2AR). The localization of these pathways is once again suggestive of potential chokepoints whose alteration could disrupt communication between the sites.

CONCLUSIONS

Graph-based methods (7) have recently gained popularity as a powerful tool for analyzing biological macromolecules, complementary to conventional metrics such as covariance of interatomic motions and PCA of these correlations. We have developed a novel, to our knowledge, graph-based method that, like several existing methods, is derived from covariance of interatomic motions and, like others, can identify communication pathways within a protein. However, this method combines three crucial advantages into a single approach: 1) it can be performed with arbitrary structural data (e.g., from a crystal structure, NMR, MD, or other simulation techniques); 2) it incorporates the influence of multiple possible paths between sites and produces a simple two-dimensional representation (the **M**-matrix) of mechanically coupled interactions between residues or secondary structural elements in a protein that can identify sites that may interact with each other; and 3) it can produce simple one-dimensional representations (χ^{X} and $\chi^{X,Y}$) of the importance of residues or secondary structural elements

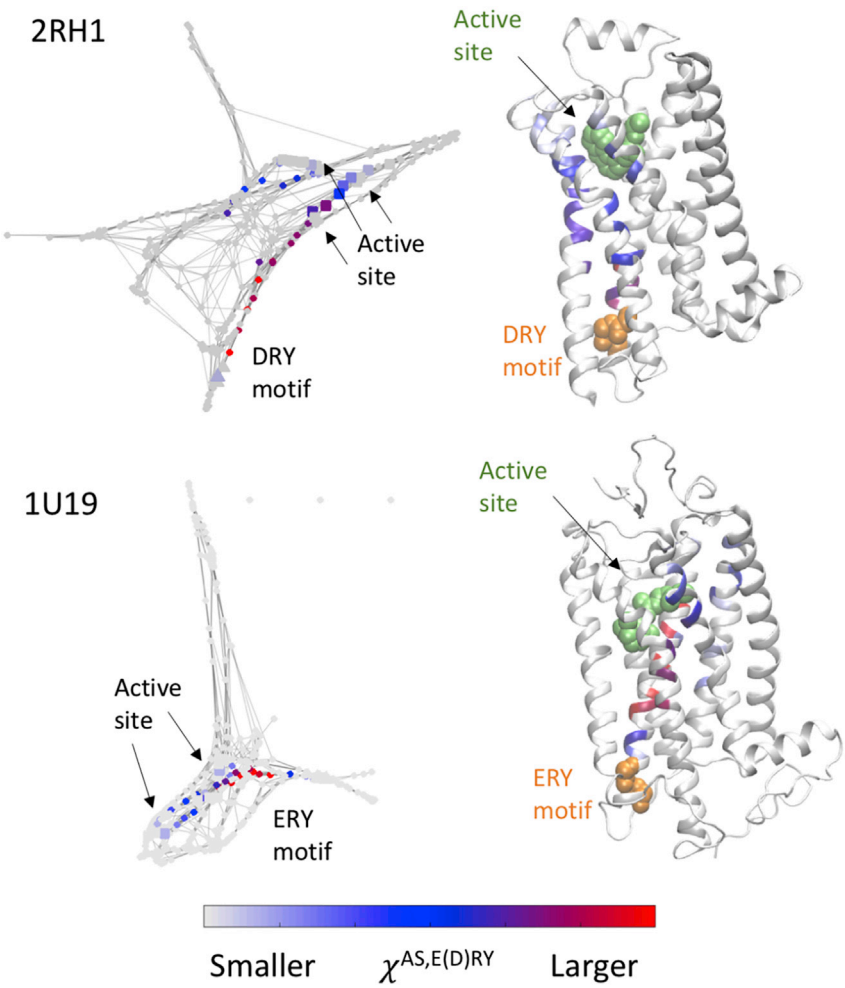

FIGURE 8 Comparison of cost-weighted betweenness centrality parameters $\chi^{AS,E(D)RY}$ between the active site and the E(D)RY motif on the G-protein binding site for two G-protein-coupled receptors, human B2AR (PDB: 2RH1, *top*) and bovine rhodopsin (PDB: 1U19, *bottom*). The ligand in the active site (the β-blocker carazolol for 2RH1 and retinol for 1U19) is shown in green, and the E(D)RY motif is shown in orange. $\chi^{AS,E(D)RY}$ is mapped to an all-residue graph representation on the left and to the protein backbone on the right. Discrete pathways of high-χ residues can be observed for both proteins. To see this figure in color, go online.

for the mechanical coupling with a site. These parameters are calculated from cost-weighted betweenness centralities for pathways involving one or two sites. They provide us with semiquantitative information on the importance of specific structural elements (single residues, secondary structure elements, or full protein domains) to funnel energy between distant parts of a protein or protein complex. The three outputs of our method provide complementary information for determining connected sites and the influence of residues on the strength of those connections.

In summary, our analysis clearly shows that the mechanical communication between remote protein sites is the result of an ensemble of internal motions of a protein, which originate a multitude of energy transfer pathways between sites. However, these pathways share a few common nodes that represent effective "chokepoints" (31). The proposed computational methodology was correlated with outcomes from a full-protein alanine-scanning mutagenesis study for the allosteric regulation of hL-PYK (21). It was further validated against prior computational predictions of allosteric coupling on two G-protein coupled receptors (8). Because of the high correlation between the key residues identified using our method and the mutagenesis study, we propose that the chokepoint nodes identified in this study can be targeted to modulate the protein's allosteric behavior. Overall, our results are suggestive that the **M**-matrix and cost-weighted betweenness centralities (x) are useful tools for rapidly predicting sites that can be affected by perturbing a macromolecule motion at a site of interest and for elucidating the key residues involved in the energy transduction between those sites.

SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2019.03.026.

AUTHOR CONTRIBUTIONS

S.R. conceived the idea. L.E.J. and S.R. developed the method. L.E.J. wrote all of the analysis codes. All authors performed research, analyzed data, and wrote the manuscript.

ACKNOWLEDGMENTS

This work was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences (L.E.J., B.G., and S.R.) and the National Institutes of Health grant GM115340 (A.W.F.). Pacific Northwest National Laboratory is operated by Battelle for the US Department of Energy under Contract DE-AC05-76RL01830.

REFERENCES

1. Fenton, A. W. 2008. Allostery: an illustrated definition for the 'second secret of life'. *Trends Biochem. Sci.* 33:420–425.
2. Carlson, G. M., and A. W. Fenton. 2016. What mutagenesis can and cannot reveal about allostery. *Biophys. J.* 110:1912–1923.
3. Reinhart, G. D. 1983. The determination of thermodynamic allosteric parameters of an enzyme undergoing steady-state turnover. *Arch. Biochem. Biophys.* 224:389–401.
4. Reinhart, G. D. 1988. Linked-function origins of cooperativity in a symmetrical dimer. *Biophys. Chem.* 30:159–172.
5. Goodey, N. M., and S. J. Benkovic. 2008. Allosteric regulation and catalysis emerge via a common route. *Nat. Chem. Biol.* 4:474–482.
6. La Sala, G., S. Decherchi, …, W. Rocchia. 2017. Allosteric communication networks in proteins revealed through pocket crosstalk analysis. *ACS Cent. Sci.* 3:949–960.
7. Dokholyan, N. V. 2016. Controlling allosteric networks in proteins. *Chem. Rev.* 116:6463–6487.
8. Balabin, I. A., W. Yang, and D. N. Beratan. 2009. Coarse-grained modeling of allosteric regulation in protein receptors. *Proc. Natl. Acad. Sci. USA.* 106:14253–14258.
9. Papaleo, E., G. Saladino, …, R. Nussinov. 2016. The role of protein loops and linkers in conformational dynamics and allostery. *Chem. Rev.* 116:6391–6423.
10. Bhabha, G., J. T. Biel, and J. S. Fraser. 2015. Keep on moving: discovering and perturbing the conformational dynamics of enzymes. *Acc. Chem. Res.* 48:423–430.
11. Klinman, J. P. 2015. Dynamically achieved active site precision in enzyme catalysis. *Acc. Chem. Res.* 48:449–456.
12. Kohen, A. 2015. Role of dynamics in enzyme catalysis: substantial versus semantic controversies. *Acc. Chem. Res.* 48:466–473.
13. Beratan, D. N., C. Liu, …, Y. Zhang. 2015. Charge transfer in dynamical biosystems, or the treachery of (static) images. *Acc. Chem. Res.* 48:474–481.
14. Liu, J., and R. Nussinov. 2017. Energetic redistribution in allostery to execute protein function. *Proc. Natl. Acad. Sci. USA.* 114:7480–7482.
15. Guarnera, E., and I. N. Berezovsky. 2016. Structure-based statistical mechanical model accounts for the causality and energetics of allosteric communication. *PLoS Comput. Biol.* 12:e1004678.
16. Yu, H., C. H. Wu, …, H. Li. 2018. Structure of an ancient respiratory system. *Cell.* 173:1636–1649.e16.
17. Buchenberg, S., F. Sittel, and G. Stock. 2017. Time-resolved observation of protein allosteric communication. *Proc. Natl. Acad. Sci. USA.* 114:E6804–E6811.
18. Danyal, K., S. Shaw, …, E. Antony. 2016. Negative cooperativity in the nitrogenase Fe protein electron delivery cycle. *Proc. Natl. Acad. Sci. USA.* 113:E5783–E5791.
19. Ichiye, T., and M. Karplus. 1991. Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins.* 11:205–217.
20. Balsera, M. A., W. Wriggers, …, K. Schulten. 1996. Principal component analysis and long time protein dynamics. *J. Phys. Chem.* 100:2567–2572.
21. Tang, Q., and A. W. Fenton. 2017. Whole-protein alanine-scanning mutagenesis of allostery: a large percentage of a protein can contribute to mechanism. *Hum. Mutat.* 38:1132–1143.
22. Copperman, J., and M. G. Guenza. 2015. Predicting protein dynamics from structural ensembles. *J. Chem. Phys.* 143:243131.
23. Jacobs, D. J., D. R. Livesay, …, D. Verma. 2012. Ensemble properties of network rigidity reveal allosteric mechanisms. *Methods Mol. Biol.* 796:279–304.
24. Ozkanlar, A., and A. E. Clark. 2014. ChemNetworks: a complex network analysis tool for chemical systems. *J. Comput. Chem.* 35:495–505.
25. Le Roux, S., and P. Jund. 2010. Ring statistics analysis of topological networks: new approach and application to amorphous GeS2 and SiO2 systems. *Comput. Mater. Sci.* 49:70–83.
26. Vishveshwara, S., K. V. Brinda, and N. Kannan. 2002. Protein structure: insights from graph theory. *J. Theor. Comput. Chem.* 01:187–211.

27. Beratan, D. N., J. N. Betts, and J. N. Onuchic. 1991. Protein electron transfer rates set by the bridging secondary and tertiary structure. *Science.* 252:1285–1288.

28. Daily, M. D., and J. J. Gray. 2009. Allosteric communication occurs via networks of tertiary and quaternary motions in proteins. *PLoS Comput. Biol.* 5:e1000293.

29. Yao, X. Q., L. Skjærven, and B. J. Grant. 2016. Rapid characterization of allosteric networks with ensemble normal mode analysis. *J. Phys. Chem. B.* 120:8276–8288.

30. Guzel, P., and O. Kurkcuoglu. 2017. Identification of potential allosteric communication pathways between functional sites of the bacterial ribosome by graph and elastic network models. *Biochim. Biophys. Acta, Gen. Subj.* 1861:3131–3141.

31. 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.

32. Marino, V., and D. Dell'Orco. 2016. Allosteric communication pathways routed by Ca^{2+}/Mg^{2+} exchange in GCAP1 selectively switch target regulation modes. *Sci. Rep.* 6:34277.

33. Kurzbach, D. 2016. Network representation of protein interactions: theory of graph description and analysis. *Protein Sci.* 25:1617–1627.

34. Laine, E., C. Auclair, and L. Tchertanov. 2012. Allosteric communication across the native and mutated KIT receptor tyrosine kinase. *PLoS Comput. Biol.* 8:e1002661.

35. Wriggers, W., K. A. Stafford, …, D. E. Shaw. 2009. Automated event detection and activity monitoring in long molecular dynamics simulations. *J. Chem. Theory Comput.* 5:2595–2605.

36. Vijayabaskar, M. S., and S. Vishveshwara. 2010. Interaction energy based protein structure networks. *Biophys. J.* 99:3704–3715.

37. Brinda, K. V., S. Vishveshwara, and S. Vishveshwara. 2010. Random network behaviour of protein structures. *Mol. Biosyst.* 6:391–398.

38. Brinda, K. V., and S. Vishveshwara. 2005. A network representation of protein structures: implications for protein stability. *Biophys. J.* 89:4159–4170.

39. Flechsig, H. 2017. Design of elastic networks with evolutionary optimized long-range communication as mechanical models of allosteric proteins. *Biophys. J.* 113:558–571.

40. Ishwar, A., Q. Tang, and A. W. Fenton. 2015. Distinguishing the interactions in the fructose 1,6-bisphosphate binding site of human liver pyruvate kinase that contribute to allostery. *Biochemistry.* 54:1516–1524.

41. Fenton, A. W., and Q. Tang. 2009. An activating interaction between the unphosphorylated n-terminus of human liver pyruvate kinase and the main body of the protein is interrupted by phosphorylation. *Biochemistry.* 48:3816–3818.

42. Alontaga, A. Y., and A. W. Fenton. 2011. Effector analogues detect varied allosteric roles for conserved protein-effector interactions in pyruvate kinase isozymes. *Biochemistry.* 50:1934–1939.

43. Fenton, A. W., and M. Hutchinson. 2009. The pH dependence of the allosteric response of human liver pyruvate kinase to fructose-1,6-bisphosphate, ATP, and alanine. *Arch. Biochem. Biophys.* 484:16–23.

44. Bakan, A., L. M. Meireles, and I. Bahar. 2011. ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics.* 27:1575–1577.

45. Heinig, M., and D. Frishman. 2004. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res.* 32:W500–W502.

46. Proctor, E. A., P. Kota, …, N. V. Dokholyan. 2015. Rational coupled dynamics network manipulation rescues disease-relevant mutant cystic fibrosis transmembrane conductance regulator. *Chem. Sci. (Camb.).* 6:1237–1246.

47. Clark, A. E. 2015. Intermolecular network theory: a general approach for understanding the structural and dynamic properties of liquids and solutions. *Annu. Rep. Comput. Chem.* 11:313–359.

48. Vishveshwara, S., A. Ghosh, and P. Hansia. 2009. Intra and intermolecular communications through protein structure network. *Curr. Protein Pept. Sci.* 10:146–160.

49. Freeman, L. C. 1977. A set of measures of centrality based on betweenness. *Sociometry.* 40:35–41.

50. Pavlopoulos, G. A., M. Secrier, …, P. G. Bagos. 2011. Using graph theory to analyze biological networks. *BioData Min.* 4:10.

51. Di Paola, L., M. De Ruvo, …, A. Giuliani. 2013. Protein contact networks: an emerging paradigm in chemistry. *Chem. Rev.* 113:1598–1613.

52. Jeong, H., S. P. Mason, …, Z. N. Oltvai. 2001. Lethality and centrality in protein networks. *Nature.* 411:41–42.

53. Jurica, M. S., A. Mesecar, …, B. L. Stoddard. 1998. The allosteric regulation of pyruvate kinase by fructose-1,6-bisphosphate. *Structure.* 6:195–210.

54. Fenton, A. W., and G. D. Reinhart. 2002. Isolation of a single activating allosteric interaction in phosphofructokinase from Escherichia coli. *Biochemistry.* 41:13410–13416.

55. Johnson, J. L., and G. D. Reinhart. 1992. MgATP and fructose 6-phosphate interactions with phosphofructokinase from Escherichia coli. *Biochemistry.* 31:11510–11518.

56. Holyoak, T., B. Zhang, …, A. W. Fenton. 2013. Energetic coupling between an oxidizable cysteine and the phosphorylatable N-terminus of human liver pyruvate kinase. *Biochemistry.* 52:466–476.

Biophysical Journal, Volume *116*

Supplemental Information

Chokepoints in Mechanical Coupling Associated with Allosteric Proteins: The Pyruvate Kinase Example

Lewis E. Johnson, Bojana Ginovska, Aron W. Fenton, and Simone Raugei

Validation of cutoff distances for local correlation

One of the key parameters in building the local correlation graph underlying the **M**-matrix is the cutoff distance between residues (R_{cut}) beyond which residues are considered to be non-interacting. If this distance is set too short, strong non-covalent interactions between side chains may be ignored, but if it is set too long, residues that are identified as interacting may not be interacting directly, with their interactions instead mediated by a third residue. For the method to function optimally, it should include ***all*** pairwise correlations, and ***only*** pairwise correlations. Previous work [1-2] by Dokholyan and co-workers found that an inter-residue distance of 7.5 Å provided a reasonable cutoff. Increasing R_{cut} to 10 Å (Figure S1) results in a much higher density of paths, likely including some non-local correlations, while reducing R_{cut} to 5 Å (Figure S1) leaves a very sparse graph that only includes the strongest correlations. Reducing below 5 Å results in some isolated nodes (residues not indicated as interacting with any other residues).

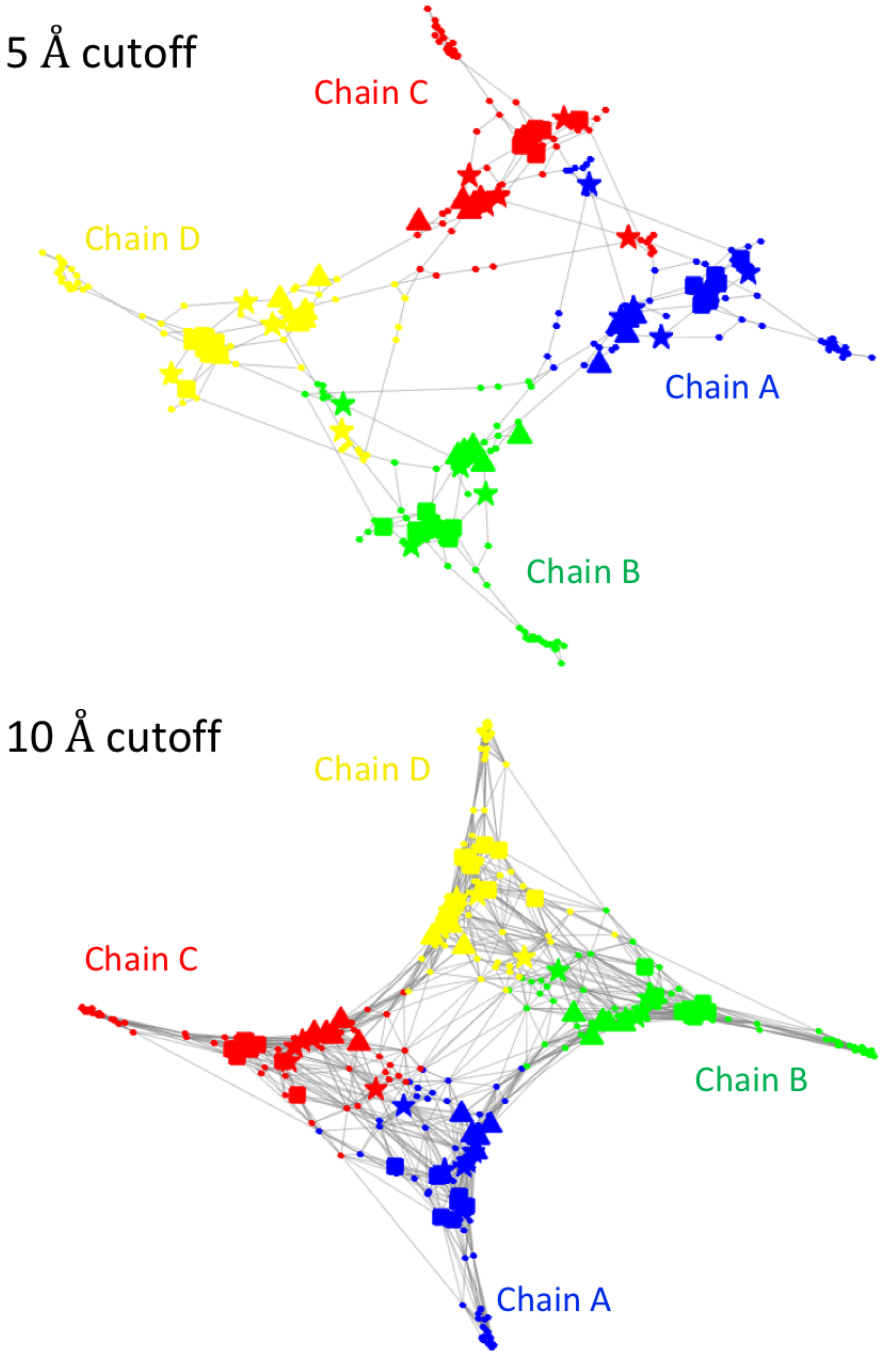

Figure S1. Mechanical coupling graphs for human liver pyruvate kinase using Euclidian distance cutoffs R_{cut} for building the adjacency matrix. Chains are distinguished by color; PEP sites are shown as squares, FBP sites as triangles, and Ala^{E} sites as stars. Reducing the cutoff reduces the density of possible paths, though the essential configuration of the protein is preserved.

Effects of crystal structure symmetrization

A number of residues were absent or poorly resolved in the source crystal structure for hL-PYK (PDB: 4IMA). Missing or poorly resolved residues were not consistent between the four chains in the hL-PYK homotetramer, translating to an asymmetric mechanical coupling graph (Figure S2, left). The asymmetries were resolved by reconstructing the protein from applying symmetry operations to the structure of chain A to build a nearly complete, symmetric model. While differences could easily be visually resolved in the graph, effects on the M-matrix and χ parameters were minimal and were primarily localized to chain D, which had the most unresolved residues (83 out of 543, vs 21 for chain A). The non-symmetrized M-matrix for 4IMA in terms of secondary structural elements is shown in Figure S2, right.

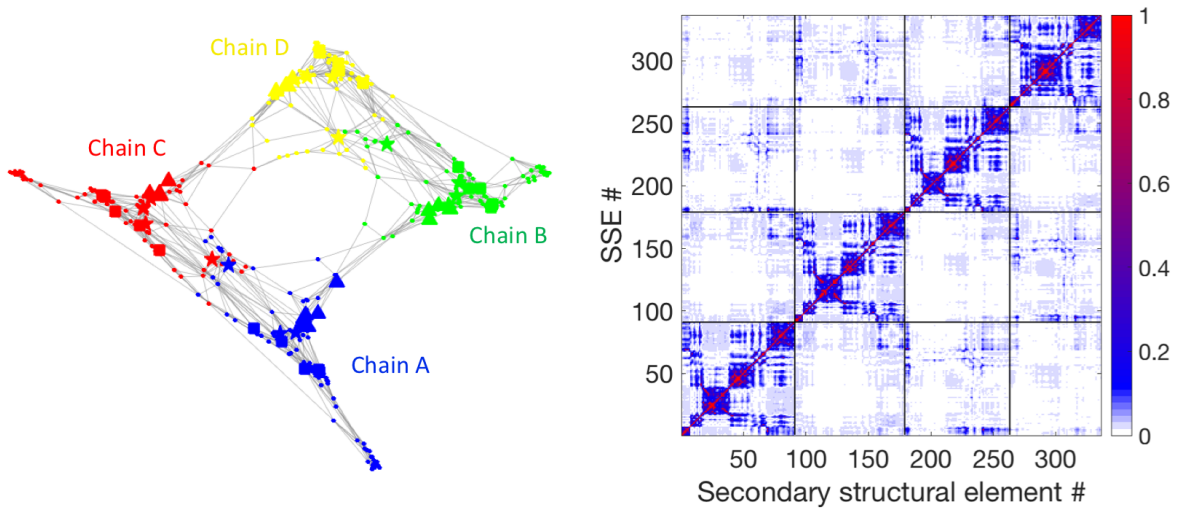

Figure S2. Non-symmetrized mechanical coupling graph for 4IMA (left) and corresponding M-matrix (right)

Reduced D, C, G, and P-matrices for hL-PYK

Graphs shown in the main manuscript have been reduced to secondary structural elements for clarity in illustrating and understanding mechanical couplings. As the residues within each secondary structural element are strongly coupled to each other, coarse-graining the structure from residues to secondary structural elements does not have any substantial qualitative effect on the interactions shown by the **C** (Figure S3, top right) and **M** matrix (Figure S3, bottom right), along with the Euclidian distance matrix **D** (Figure S3, top left) used for deriving the graph and geodesic **G** (Figure S3, bottom left) distance matrix for the graph that is used to derive the **M**-matrix.

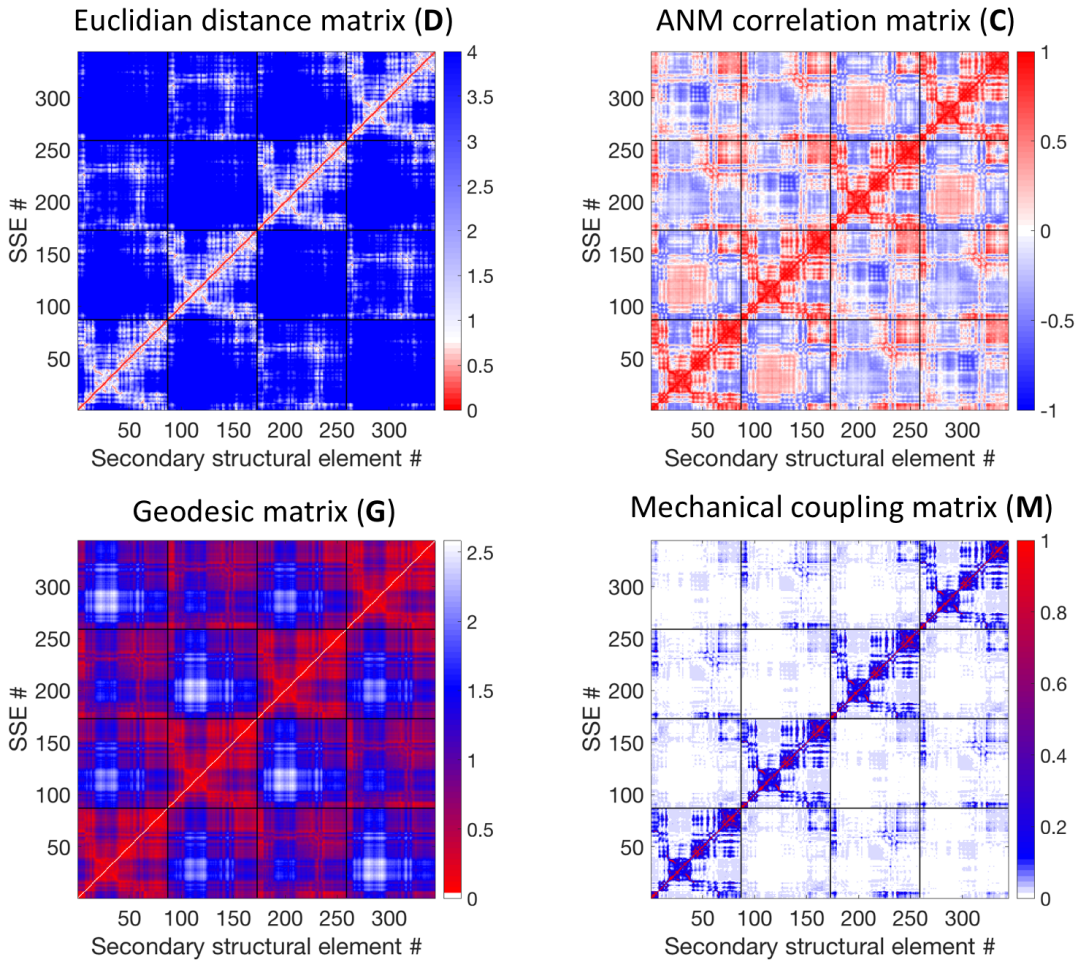

Figure S3. Reduced Euclidian distance (D, **top left**) correlation (C, **top right**), geodesic distance (G**, bottom left**) and mechanical coupling (M, **bottom right**) matrices for hL-PYK, crystal structure 4IMA.

Example shortest path tree

Shortest path trees are a type of subgraph that represents the geodesic path from one node *i* on a graph to every other node $j \neq i$ on the graph. The betweenness centrality metric for evaluating coupling between sites uses a subgraph built from residues that are identified as of interest based on the **M**-matrix and/or experimental data. An example of a shortest path tree is shown in Figure S4.

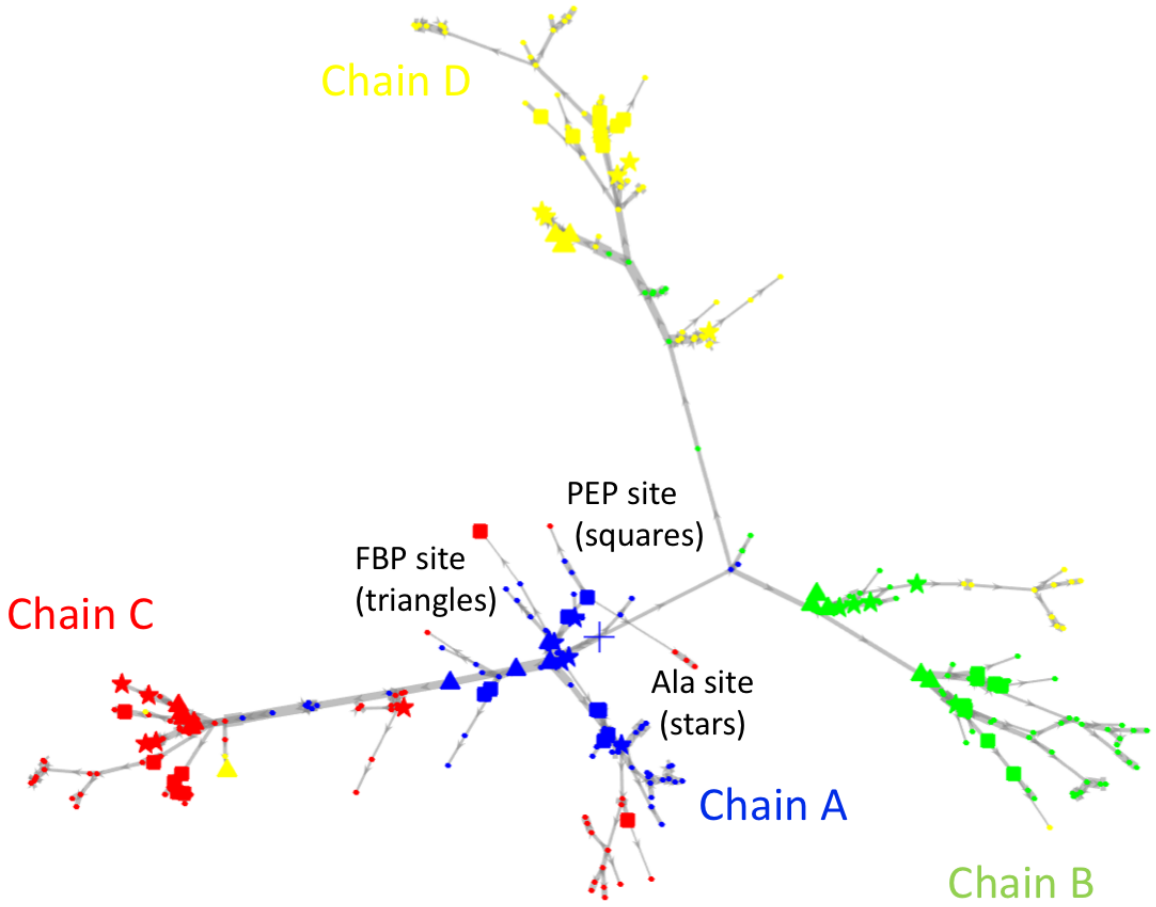

Figure S4. Shortest path tree of mechanical coupling network showing interaction of the chain A FBP site (blue triangles) with all other secondary structural elements in the proteins. Chains are colored; secondary structural elements in PEP sites are shown as squares, FBP sites as triangles, and Ala^{E} sites as stars. The origin of the search (SSE 85 in the FBP site in Chain A) is shown as a large plus (+). The strength of the coupling is represented by the thickness of lines in between the chain A FBP site is strongly coupled with the PEP and alanine sites on Chain A, with some coupling to the FBP site on chain B and the PEP site on chain C with much longer paths required to reach any other sites.

Comparison of χ and |Q| for Ala^{E}

A comparison of cost-weighted betweenness centrality parameters and alanine scan results for Ala^{E} as a ligand, analogous to the comparison shown for FBP in Figure 5, is shown below, with data averaged over secondary structural elements. While the quantitative, residue-by-residue comparison shown in Figure S5 is of similar quality to the comparison observed for FBP, the quality of the visual comparison is reduced by coarse-graining due to the higher degree of locality (e.g. single residue) for mutations affecting Ala^{E} than for FBP, as mentioned in the main text. Several key regions influencing are identified by the χ analysis of mechanical coupling; these regions are indicated with black arrows on the lower panel of the figure.

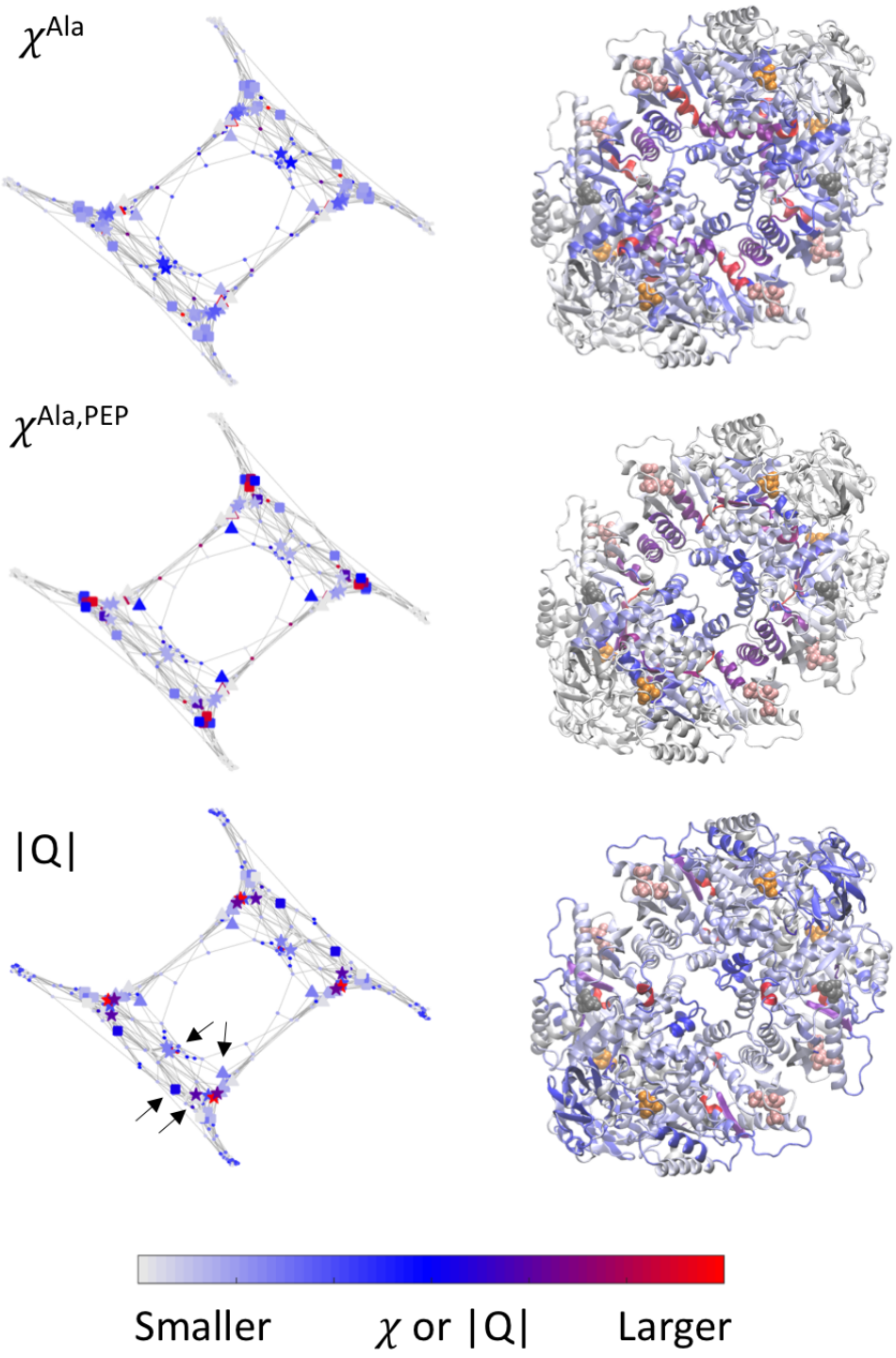

Figure S5. Comparison of χ^{Ala}, $\chi^{Ala,PEP}$, and |Q| for hL-PYK, averaged over secondary structural elements.

Relationship between FBP and Ala^{E} allostery

In pyruvate kinase, FBP functions as an allosteric activator of PEP affinity and Ala^{E} as an allosteric inhibitor of PEP affinity. Of the residues identified by the alanine scan as key to each allosteric effect, some show effects of opposite sign for each ligand, indicating that they are involved in both pathways. The complementarity between allosteric effects is shown in Figure S6.

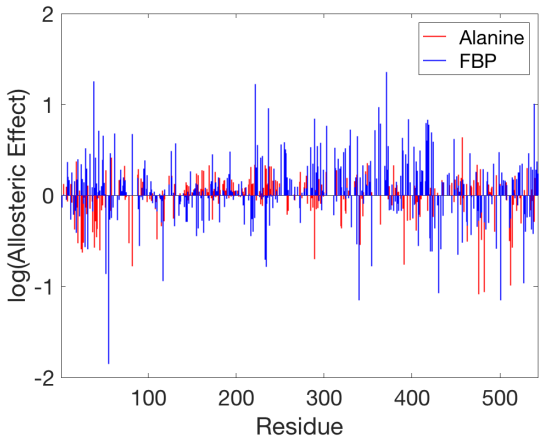

Figure S6. Allosteric effect as represented by alanine scan Q_i/Q_0 for Ala and FBP by residue, showing complementarity between mutations that change the negative allosteric effect of alanine with those that change the positive allosteric effect of FBP.

References for Supporting Information:

1. Proctor, E. A.; Kota, P.; Aleksandrov, A. A.; He, L.; Riordan, J. R.; Dokholyan, N. V., Rational Coupled Dynamics Network Manipulation Rescues Disease-Relevant Mutant Cystic Fibrosis Transmembrane Conductance Regulator. *Chem Sci* **2015,** *6*, 1237-1246. doi:10.1039/C4SC01320D

2. Dokholyan, N. V., Controlling Allosteric Networks in Proteins. *Chem Rev* **2016,** *116*, 6463-6487. doi:10.1021/acs.chemrev.5b00544