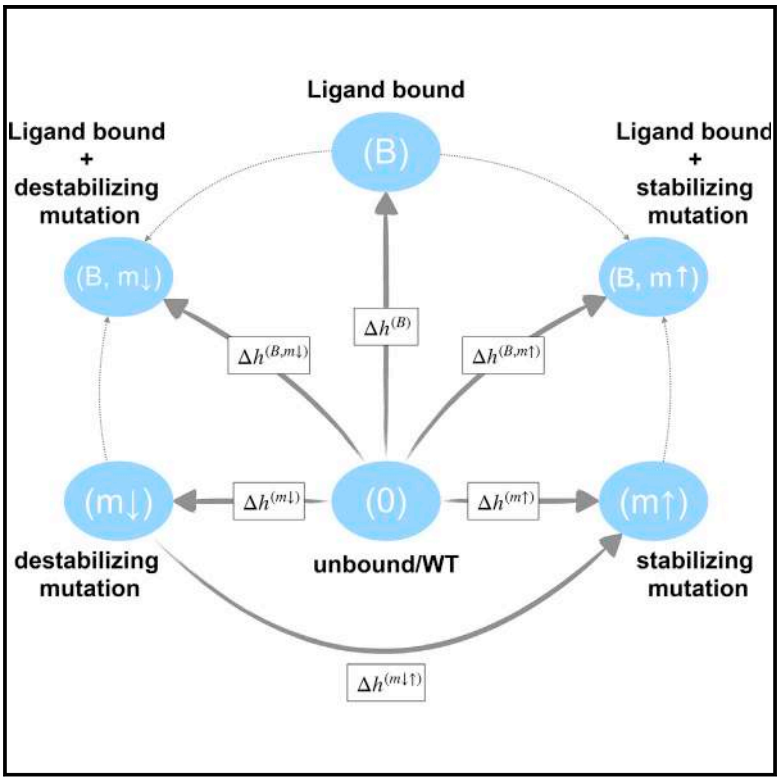
Structure

Toward Comprehensive Allosteric Control over Protein Activity

Graphical Abstract



Highlights

- Structure-based statistical mechanical model of allostery (SBSMMA)
- Perturbations by allosteric ligands and mutations allow control of protein activity
- The allosteric modulation range is a generic descriptor of the allosteric signaling
- Allosteric signaling map gives exhaustive description of allosteric communication

Authors

Enrico Guarnera, Igor N. Berezovsky

Correspondence

igorb@bii.a-star.edu.sg

In Brief

Guarnera and Berezovsky propose a theoretical framework for achieving the comprehensive allosteric control over protein activity. They introduce the allosteric signaling maps, which exhaustively describe the protein allosteric communication at per-residue resolution, allowing the tuning of already existing signaling and the design of new elements of regulation.

Guarnera & Berezovsky, 2019, Structure *27*, 866–878
May 7, 2019
https://doi.org/10.1016/j.str.2019.01.014

CellPress

CellPress

Structure
Theory

Toward Comprehensive Allosteric Control over Protein Activity

Enrico Guarnera,[1] and Igor N. Berezovsky[1,2,3,*]
[1]Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR), 30 Biopolis Street, #07-01, Matrix, Singapore 138671, Singapore
[2]Department of Biological Sciences (DBS), National University of Singapore (NUS), 8 Medical Drive, Singapore 117579, Singapore
[3]Lead Contact
*Correspondence: igorb@bii.a-star.edu.sg
https://doi.org/10.1016/j.str.2019.01.014

SUMMARY

Universality of allosteric signaling in proteins, molecular machines, and receptors complemented by the great advantages of prospected allosteric drugs in the highly specific, non-competitive, and modulatory nature of their actions calls for deeper theoretical understanding of allosteric communication. We present a computational model that makes it possible to tackle the problem of modulating the energetics of protein allosteric communication. In the context of the energy landscape paradigm, allosteric signaling is always a result of perturbations, such as ligand binding, mutations, and intermolecular interactions. The calculation of local partition functions in the protein harmonic model with perturbations allows us to evaluate the energetics of allosteric communication at the single-residue level. In this framework, Allosteric Signaling Maps are proposed as a tool to exhaustively describe allosteric communication in the protein, to tune already existing signaling, and to design new elements of regulation for taking the protein activity under allosteric control.

INTRODUCTION

Starting from phenomenological Monod-Weynman-Changeux (Monod et al., 1965) and Koshland-Nemethy-Filmer (Koshland et al., 1966) considerations, the notion of allostery, i.e., modulation of protein activity by the effector binding to remote regulatory exosites (Berezovsky et al., 2017), had evolved into a well-formulated quantifiable concept (Cui and Karplus, 2008). The current view is based on the key role of protein dynamics (Jardetzky, 1996; Guarnera and Berezovsky, 2016a) in underlying the allosteric signaling regardless of the presence (Jardetzky, 1996; Mitternacht and Berezovsky, 2011) or absence (Cooper, 1976; Cooper and Dryden, 1984) of conformational changes in proteins of all structural types and functions (Gunasekaran et al., 2004; Cui and Karplus, 2008). The non-competitive and modulatory nature of allosteric regulation together with the high specificity of its signaling provide advantages of allosteric drugs (Nussinov and Tsai, 2013) in preventing toxicity and receptor desensitization (Wenthur et al., 2014), and in facilitating functional selectivity (Meharena et al., 2013; Guarnera and Berezovsky, 2016a). The design of allosteric drugs (Berezovsky, 2013) with desired agonist/antagonist activity requires understanding of allosteric mechanisms and ways of their adjustment (Guarnera and Berezovsky, 2016a), resulting in a rapidly growing number of experimental and computational studies of allostery (Nussinov and Tsai, 2013; Lu et al., 2014).

The unifying theme of all kinds of observed allosteric regulation comes from its physical origin: perturbation of the protein dynamics (Guarnera and Berezovsky, 2019). Ligand binding, mutations, post-translational modifications, intermolecular interactions, and their combinations can serve as a source of perturbation. Additionally, according to Cooper's estimate that "small volume fluctuations correspond to very small changes in overall dimensions of a globular protein, but if concentrated in one area would produce cavities or channels in the proteins sufficient to allow entry of solvent or probe molecules," allosteric communication can be driven by pure equilibrium fluctuations (Cooper, 1976). The whole-protein alanine-scanning mutagenesis showed that single mutation can also serve as a perturbation-initiating and/or -alternating allosteric signal (Tang and Fenton, 2017); for example, removal of critical residues disrupts the signal propagation in CFTR (cystic fibrosis transmembrane conductance regulator) (Proctor et al., 2015). Recent theoretical and computational approaches describing the propagation of allosteric signal from a perturbation caused by the mutation/site are discussed elsewhere (Dokholyan, 2016; Guarnera and Berezovsky, 2019).

We have recently proposed the structure-based statistical mechanical model of allostery (SBSMMA), which is based on the harmonic model of a protein and accounts for the causality and energetics of allosteric communication (Guarnera and Berezovsky, 2016b). In the SBSMMA the perturbation caused by the ligand binding is simulated by increasing the stiffness of contacts between residues in the binding site, while the allosteric effect of destabilizing/stabilizing mutations is modeled by weakening/strengthening the couplings in the contact network of the mutated residue (Kurochkin et al., 2017; Guarnera et al., 2017). The model consists of three components: (1) original and perturbed protein states are considered in the context of harmonic approximation and two sets of characteristic normal modes are obtained and (2) used as an input for an allosteric potential, which evaluates the mean elastic work on a residue produced by its neighbors; then, (3) partition functions characterizing the

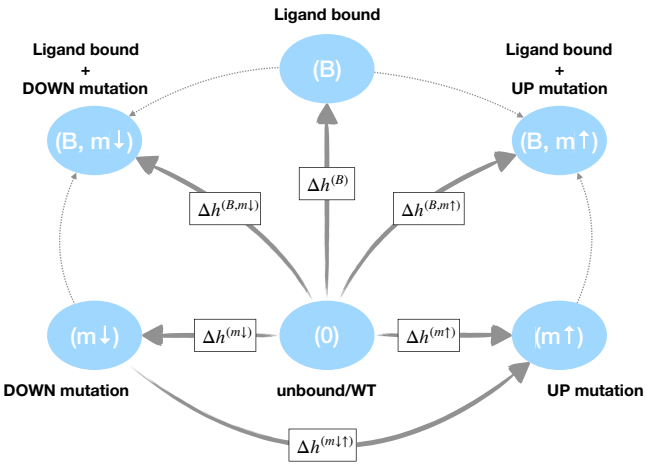


Figure 1. Comprehensive Control of Allosteric Signaling

Allosteric modulations can be quantified for pairs of proteins states defined by different types of perturbations. Starting from a reference state (0, unbound/WT) the paired perturbed state is defined via either ligand binding (B, $\Delta h^{(B)}$), mutation (UP, $\Delta h^{(m\uparrow)}$ and DOWN, $\Delta h^{(m\downarrow)}$), or a combination of the two ($\Delta h^{(B,m\uparrow)}$ and $\Delta h^{(B,m\downarrow)}$). The allosteric effect associated with the transition from destabilizing to stabilizing mutations identifies the allosteric modulation range $\Delta h^{(m\downarrow\uparrow)}$ caused by the corresponding protein sequence position.

ensemble of all possible configurations of a residue neighbors serve for estimating the free energy difference between the original and perturbed (by the ligand binding and/or mutations) states. The obtained free energy difference is the change of configurational work exerted on a residue as a result of the alteration in the ensemble of configurations of its neighbors induced by the perturbation.

In this work we explore the possibility to obtain comprehensive allosteric control over protein activity. To this end, we consider here a potential allosteric effect of every protein residue by introducing the notion of allosteric modulation, which quantifies to what extent a residue can remotely affect protein activity. To obtain a generic description, we quantify the effect of allosteric modulation caused by the residue as a result of the change in allosteric signaling upon replacement of the smallest (Ala/Gly-like) residue to the bulkiest one, irrespective of the native amino acid in the corresponding position. We start from the analysis of allosteric communication between known regulatory and functional sites, followed by the study of allosteric effect of mutations, and the combined impact of ligand binding and mutations. Using case studies of three proteins, tetrameric phosphofructokinase (PFK), ring-shaped tetrameric D-3-phosphoglycerate dehydrogenase (PGDH), and monomeric four-domain insulin-degrading enzyme (IDE), we explore the mechanisms of allosteric signaling in proteins of different structures, oligomeric states, and functions. We investigate the energetics of allosteric signaling caused by the ligand binding, its cooperativity, and a possibility to modulate it via allosterically acting mutations. Additionally, relying on the experimental case study of the PDZ domain, we establish a quantitative link between the calculated allosteric modulation and observed "functional costs of mutations" (Mclaughlin et al., 2012).

On the basis of the above preliminary analysis, we propose a residue-by-residue scanning of allosteric effects of mutations to obtain the Allosteric Signaling Maps (ASMs). ASMs allow us to establish a clear relationship between residues of the protein with a potential to allosterically modulate protein activity and those residues that mainly affect overall structure stability or act orthosterically, being located in the functional sites and/or directly interacting with its residues. We show that ASMs provide complete information on the allosteric communication in the protein, which allows one to control protein activity at per-residue resolution by complementing and tuning already existing allosteric signaling with effects of mutations, as well as designing potentially new elements of regulation on the basis of allosteric signals from individual protein residues and their combinations. Additionally, the change of mode in the allosteric signaling observed in the analysis of ASMs as a function of distances allows to examine the role of domains/subunits in allosteric communication within multidomain/oligomeric proteins (Fan et al., 2018) and to identify critical signaling inside and outside domains/monomers with the sign of modulation opposite to that characteristic for these structural units.

RESULTS

We propose here a theoretical model of allosteric control of protein activity aimed at the quantification of allosteric effects at the single-residue level. Within the framework of SBSMMA (Guarnera and Berezovsky, 2016b; Kurochkin et al., 2017; Guarnera et al., 2017), the effect of a perturbation P (ligand binding, mutations, and their combinations) is evaluated for single residues as a free energy difference $\Delta g_i^{(P)}$ between the perturbed and unperturbed protein states. The free energy difference $\Delta g_i^{(P)}$ gives the change in the work applied on the residue i as a result of the applied perturbation. The portion of work exerted due to purely allosteric effect is defined as the allosteric modulation $\Delta h_i^{(P)}$, which unlike $\Delta g_i^{(P)}$ distinguishes between local and global effects that are induced upon a perturbation P (see Equation 8). In other words, the allosteric modulation $\Delta h_i^{(P)}$ is calculated as a background-free effect, where protein-average per-residue allosteric free energy is subtracted from that detected on the residue/site of interest, showing the extent at which the allosteric signal is stronger than the background signaling in the protein. Positive allosteric modulation, expressed as the increase of work acting on a residue/site, may result in a local conformational change, whereas negative modulation, corresponding to a decrease in the work, may prevent conformational change.

Figure 1 shows a scheme of the comprehensive allosteric control of the protein activity, illustrating possible modes of allosteric regulation originating from different types of perturbations, such as ligand binding (B, $\Delta h^{(B)}$), stabilizing (UP, $\Delta h^{(m\uparrow)}$), and destabilizing (DOWN, $\Delta h^{(m\downarrow)}$) mutations, and combinations of them ($\Delta h^{(B,m\uparrow)}$ and $\Delta h^{(B,m\downarrow)}$). To generically characterize the allosteric effect of mutations regardless of the native residue in a selected position, we obtain the allosteric modulation range $\Delta h^{(m\downarrow\uparrow)}$ as a difference between modulations obtained upon stabilizing (UP) and destabilizing (DOWN) mutations. Furthermore, the ASM is introduced as an exhaustive description of the allosteric signaling that can originate from potentially every residue of

the protein. The ASM data can be directly used for the design and tuning of allosteric sites (Tee et al., 2018), prediction of latent regulatory exosites, and of allosteric effects of mutations.

Comprehensive Model of Allostery

Starting from the problems of causality and energetics in allosteric communication (Guarnera and Berezovsky, 2016b), we present a computational model with the goal of achieving comprehensive allosteric control over the protein activity.

Given a protein reference conformational state, denoted as "0" (unperturbed state), the C_α harmonic model energy function is

$$E^{(0)}(\boldsymbol{r}-\boldsymbol{r}^0)=\sum_{\langle i,j\rangle}k_{ij}\left(d_{ij}-d_{ij}^0\right)^2, \qquad \text{(Equation 1)}$$

where $\boldsymbol{r}$ is the 3N-dimensional vector of coordinates of the C_α atoms, $\boldsymbol{r}^0$ is the vector of C_α positions of the reference structure, d_{ij} is the distance between the C_α atoms i and j, d_{ij}^0 is the corresponding distance in the reference structure, and $k_{ij}\sim(1/d_{ij}^0)^6$ is a distance-dependent force constant with a global cutoff d_c = 25 Å, with the summation running over the pairs of neighbors i,j within the global distance cutoff (Hinsen et al., 2000; Hinsen, 2000).

For the protein in a perturbed state caused by the combination of ligand binding and mutations, the energy function associated with this perturbed P state is

$$E^{(P)}(\boldsymbol{r}-\boldsymbol{r}^0,S,m)=\sum_{\langle i,j\rangle,i\neq m}k_{ij}\left(d_{ij}-d_{ij}^0\right)^2+\alpha\sum_{\langle i,j\rangle\in S}k_{ij}\left(d_{ij}-d_{ij}^0\right)^2+\theta\sum_{\langle mj\rangle}k_{mj}\left(d_{mj}-d_{mj}^0\right)^2. \qquad \text{(Equation 2)}$$

First, given the binding site S the ligand binding is modeled by introducing an additional harmonic restraining term between all the residue pairs composing the occupied binding site, with $\alpha \gg 1$ a stiffening parameter ($\alpha = 10^2$, see Guarnera and Berezovsky, 2016b for details). Similarly, a mutation is modeled by altering the strength of the force constants associated with the contacts between the mutated residue m and its neighbors. Two types of mutations are defined: UP (↑, stabilizing) mutations mimic substitutions with bulkier and strongly interacting residues, while DOWN (↓, destabilizing) mutations model substitutions to small Ala/Gly-like residues. To obtain a significant effect comparable with ligand-binding perturbations ($\alpha = 10^2$), the scaling factor of the interaction of mutated residue with its neighbors is set to $\theta = 10^2$ for stabilizing (UP) and $\theta = 10^{-2}$ for destabilizing (DOWN) mutations. Despite the only qualitative nature of factor θ and because of the lack of explicit sequence considerations, the chosen values satisfactorily work in discriminating between two extreme regimes of stabilizing and destabilizing mutations. Equation 2 can be easily generalized to the case of multiple sites and mutations.

For the reference protein state 0 described in Equation 1 and the generally perturbed state P described in Equation 2, the Hessian matrix $\mathbf{K}=\partial^2E/\partial\boldsymbol{r}_i\partial\boldsymbol{r}_j$ is considered and two sets of orthonormal normal modes $(\boldsymbol{e}_\mu^{(0)},\boldsymbol{e}_\mu^{(P)})$, characterizing the configurational ensembles of corresponding protein states are obtained. The modes $\boldsymbol{e}_\mu$ are used in the allosteric potential to evaluate the elastic work exerted on a particular residue i as a result of configurational changes described by the corresponding low-frequency normal modes

$$U_i(\sigma)=\frac{1}{2}\sum_\mu\varepsilon_{\mu,i}\sigma_\mu^2, \qquad \text{(Equation 3)}$$

where $\varepsilon_{\mu,i}$ are parameters defined from the normal modes as

$$\varepsilon_{\mu,i}=\sum_j\left|\boldsymbol{e}_{\mu,i}-\boldsymbol{e}_{\mu j}\right|^2. \qquad \text{(Equation 4)}$$

Essentially, Equation 3 accounts for the change of displacement of the neighbors of residue i over the modes with a given set of Gaussian distributed amplitudes $\sigma=(\sigma_1,\ldots,\sigma_\mu,\ldots)$ with variance $1/\varepsilon_{\mu,i}$. Since the generic displacement of a residue i is written as $\boldsymbol{r}_i(\sigma)-\boldsymbol{r}_i^0=\sum_\mu\sigma_\mu\boldsymbol{e}_{\mu,i}$, the vector σ is a representation of a configurational state of the residue i.

Integrating over the ensemble of all possible configurations σ of a residue, the per-residue partition function is obtained as

$$\begin{aligned}z_i&=\int d\sigma e^{-U_i(\sigma)/k_BT}\\&=\prod_\mu\int d\sigma_\mu e^{-\sigma_\mu^2\varepsilon_{\mu,i}/2k_BT}\\&=\prod_\mu\left(\pi\frac{2k_BT}{\varepsilon_{\mu,i}}\right)^{1/2}\end{aligned} \qquad \text{(Equation 5)}$$

and, correspondingly, the free energy $g_i=-k_B\ln z_i+const$. Thus, the free energy change on the residue i originating from the perturbation P as a combination of ligand binding and mutations is

$$\Delta g_i^{(P)}=\frac{1}{2}k_BT\sum_\mu\ln\frac{\varepsilon_{\mu,i}^{(P)}}{\varepsilon_{\mu,i}^{(0)}}. \qquad \text{(Equation 6)}$$

A positive value $\Delta g_i^{(P)}>0$ indicates an increase of work applied on residue i, which may result in its conformational change. A negative $\Delta g_i^{(P)}<0$ value shows a stabilization of residue i, which, on the contrary, may prevent its conformational change.

Cooperativity in the effect of perturbations in the case of oligomeric and multidomain proteins should also be considered. For example, in case of a homo-oligomer with identical effector-binding sites in each chain, one can define several bound states with an intermediate number of occupied binding sites (partially perturbed state "$P_{1/2}$") and one with all sites fully occupied (fully perturbed state P). Thus, the cooperativity associated with the sequential perturbation by the binding of additional ligand can be obtained via the relation

$$\Delta\Delta g_i^{(P)}=\frac{1}{2}k_BT\sum_\mu\ln\frac{\varepsilon_{\mu,i}^{(P)}\varepsilon_{\mu,i}^{(0)}}{\left(\varepsilon_{\mu,i}^{(P_{1/2})}\right)^2}. \qquad \text{(Equation 7)}$$

It should be noted that the cooperativity relation in Equation 7 also holds in the case of mutations.

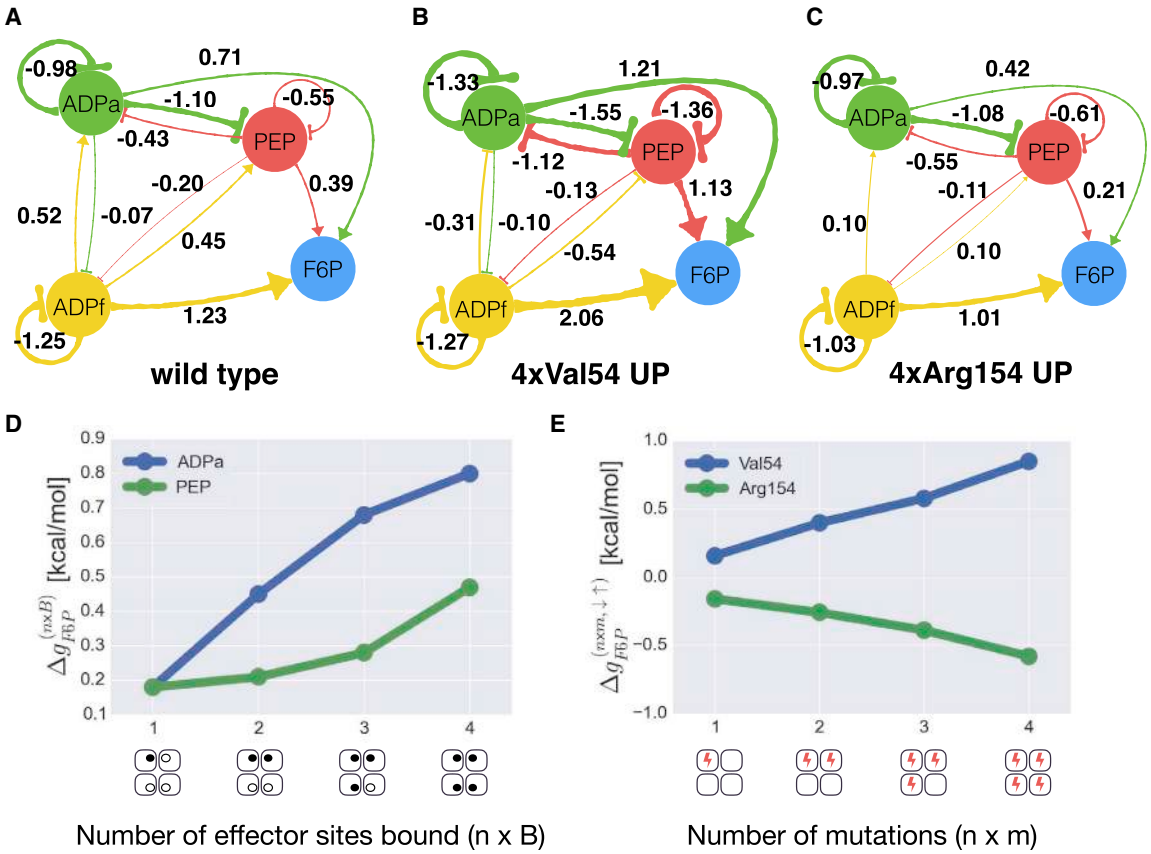


Figure 2. Allosteric Communication in the Phosphofructokinase and its Cooperativity

(A–C) Graphs of allosteric signaling in the wild-type (A) and mutated structures with UP mutations of 4xVal54 (B) and 4xArg154 (C). Allosteric and functional sites are represented by nodes with links associated with the values of the allosteric modulation $\Delta h_F^{(B)}$ of the regulated site *F* upon the binding to site *B*.

(D) Free energy response $\Delta g_{F6P}^{(n\times B)}$ at the F6P sites as a function of the number *n* of bound allosteric sites *B* (PEP or ADPa): opposite concavities in the curves suggest negative cooperativity at the F6P sites upon ADPa binding and positive upon PEP binding, respectively.

(E) Free energy response $\Delta g_{F6P}^{(n\times m\downarrow\uparrow)} = \Delta g_{F6P}^{(n\times m\uparrow)} - \Delta g_{F6P}^{(n\times m\downarrow)}$ evaluates the free energy difference between UP and DOWN mutated protein states with *n* residues *m* mutated sequentially – one for each of the protein subunits. Very weak positive and negative cooperativities are observed at the F6P sites upon mutating residues Val54 and Arg154, respectively.

To estimate a purely allosteric effect on a residue *i* due to the perturbation *P*, the deviation of the free energy difference $\Delta g_i^{(P)}$ from its mean value over all the residues of the protein chain containing the residue *i* is considered

$$\Delta h_i^{(P)} = \Delta g_i^{(P)} - \left\langle \Delta g_i^{(P)} \right\rangle_{\text{Chain}}. \qquad \text{(Equation 8)}$$

This background-free free energy change is called allosteric modulation $\Delta h_i^{(P)}$ (positive or negative) caused by the perturbation *P*. The difference

$$\Delta h_i^{(m\downarrow\uparrow)} = \Delta h_i^{(m\uparrow)} - \Delta h_i^{(m\downarrow)}, \qquad \text{(Equation 9)}$$

evaluates the "allosteric modulation range" caused by the mutation from the smallest (Ala/Gly-like) to the bulkiest residues (for example, Phe or Trp). The modulation range $\Delta h_i^{(m\downarrow\uparrow)}$ is a generic descriptor of the strength of the allosteric signal on a residue *i* originating from substitutions of the residue *m*, which can be calculated for any residue position *i* of the protein. It is clear that for reverse mutations, namely from the largest to smallest residues, the modulation range is $\Delta h_i^{(m\uparrow\downarrow)} = -\Delta h_i^{(m\downarrow\uparrow)}$. Finally, the allosteric modulation at the level of sites is evaluated by the averaging of per-residue modulations $\Delta h_i^{(P)}$ over the residues belonging to the site of interest,

$$\Delta h_{\text{Site}}^{(P)} = \left\langle \Delta h_i^{(P)} \right\rangle_{\text{Site}}. \qquad \text{(Equation 10)}$$

Importantly, the absolute values of allosteric modulation should be considered in relation to the thermal fluctuations. While values exceeding k_BT should be regarded as a strong manifestation of allosteric communication, combinations of the low-value allosteric responses may also result in significant allosteric modulation in the affected protein regions.

Phosphofructokinase

PFK phosphorylates fructose-6-phosphate (catalytic site F6P) in the process of glycolysis (Blangy et al., 1968; Traut, 2007), being activated by ADP (or GDP) and inhibited by phosphoenolpyruvate PEP. The energetics of allosteric regulation in PFK is represented by the network schemes (Figure 2) that show direction

of signaling and allosteric free energies applied in corresponding sites. Figure 2A shows a summarizing graph of the allosteric signaling in PFK originating from the binding perturbation in the corresponding sites of all four subunits of the tetramer (apo form was used; PDB: 3pfk). It shows that perturbation of the inhibitor (PEP)-binding site induces a positive allosteric modulation at the catalytic site F6P ($\Delta h_{F6P}^{(4xPEP)} = 0.39$ kcal/mol) and a negative modulation at the substrate site ADPf ($\Delta h_{ADPf}^{(4xPEP)} = -0.20$ kcal/mol). Similarly, perturbation at the activator (ADPa)-binding site causes a positive modulation at the catalytic site F6P ($\Delta h_{F6P}^{(4xADPa)} = 0.71$ kcal/mol), which is an order of magnitude higher in absolute value than weak negative modulation at the substrate site ADPf ($\Delta h_{ADPf}^{(4xADPa)} = -0.07$ kcal/mol). Binding sites of the activator ADPa and inhibitor PEP overlap significantly, and the difference in the results from their binding perturbation is expressed in the higher positive modulation at the F6P functional sites and less negative modulation at the substrate-binding site, ADPf, as a result of the activator (ADPa) binding. Therefore, one can assume that the mechanism of activation upon ADPa binding is based on the structural reorganization/adjustment resulting from the work exerted at the F6P site. The inhibitor PEP, on the other hand, apparently acts by preventing the F6P and ADPf sites from achieving the functional state, as it initiates weaker conformational changes in the F6P site and causes rearrangement in the ADPf site (see the diagram in Figure 2A for corresponding allosteric energies). To further explore this scenario of regulation, we consider the combined perturbation effect of ADPf substrate site together with either the ADPa-activator or PEP-inhibitor binding. In the case of the activator, a stronger positive modulation on the catalytic site ($\Delta h_{F6P}^{(4xADPf,4xADPa)} = 1.64$ kcal/mol) is detected rather than in the case when ADPf only is perturbed (the difference is 0.41 kcal/mol). When PEP and ADPf are bound together, modulation at the F6P site is unchanged compared with the effect of ADPf binding ($\Delta h_{F6P}^{(4xADPf,4xPEP)} = 1.21$ kcal/mol). Altogether, one can conclude that ADPa acts with extra 0.41 kcal/mol in addition to 1.23 kcal/mol from ADPf, and the activating effect of ADPa binding is similar when it binds alone ($\Delta h_{F6P}^{(4xADPa)} = 0.71$ kcal/mol) in comparison with the inhibiting PEP action ($\Delta g_{F6P}^{(4xPEP)} = 0.39$ kcal/mol). Noteworthy, positive allosteric modulation at the catalytic ($\Delta h_{F6P}^{(4xADPf)} = 1.23$ kcal/mol) and both sites of effectors ($\Delta h_{ADPa}^{(4xADPf)} = 0.52$ kcal/mol, $\Delta h_{PEP}^{(4xADPf)} = 0.45$ kcal/mol) is also detected upon perturbation at the ADPf substrate site (Figure 2A).

The different nature of ADPa and PEP modes of action is also indicated in the cooperativity of F6P binding associated with the protein states with an increasing number of perturbed effector sites. Figure 2D (see also Table S1) shows the free energy response at the catalytic site $\Delta g_{F6P}^{(nxB)}$ as a function of the number *n* of occupied effector sites (PEP or ADPa). The opposite concavities of the response curves indicate that PEP causes positive and ADPa negative cooperativities in the F6P site. The mean cooperativity of the substrate binding (see Equation 7) is estimated for all pairs of protein states *n* and *n + 1* upon sequential perturbation of PFK subunits is $\Delta\Delta g_{F6P}^{(PEP)} = 0.08$ kcal/mol in the case of PEP and $\Delta\Delta g_{F6P}^{(ADPa)} = -0.07$ kcal/mol in the case of ADPa.

Because only few residues make a difference between the binding sites of activator ADPa and inhibitor PEP, the functional role of individual residues in these sites has attracted a lot of attention. Following earlier experimental work (Lau and Fersht, 1989), we started from analyzing stabilizing (UP) and destabilizing (DOWN) mutations on the set of eight residues, Arg21, Arg25, Val54, Asp59, Arg154, Glu187, Arg211, and Lys213, belonging to the ADPa and PEP sites (Table S2). Complete data on the allosteric modulation in the functional and allosteric sites of PFK are presented in Figure S1. For each residue, we considered protein states with single-residue mutations in one subunit and four-residue mutations in all subunits simultaneously, calculating respective modulation ranges $\Delta h^{(1\times m\downarrow\uparrow)}$ and $\Delta h^{(4\times m\downarrow\uparrow)}$. Mutations of residues Val54 and Arg154 yield the strongest and, at the same time, opposite effect on the F6P site (Figures 2B and 2C; Table S2). Noteworthy, very weak negative modulation range of Arg154 ($\Delta h_{F6P}^{(1xArg154\downarrow\uparrow)} = -0.07$ kcal/mol) can be explained by the bulkiness of native Arg154, which apparently allosterically prohibits the conformational changes in the functional F6P site. Mutation of Val54 ($\Delta h_{F6P}^{(1xVaa54\downarrow\uparrow)} = 0.47$ kcal/mol) shows, on the contrary, positive modulation range, pointing to the conformational changes in the F6P site caused by the substitution into bulky amino acids.

Considering Val54 and Arg154 positions in PFK, we studied a combined effect of allosteric modulation caused by the ligand binding and mutations. The difference in sizes of arginine and valine manifested in the opposite sign of the modulation range observed for residues Val54 and Arg154 (Table S2) prompted us to analyze the effect of stabilizing mutations of these residues on the allosteric signaling caused by the ligand binding. Figures 2B and 2C contains network representations of the allosteric signaling associated with four-residue Val54 and Arg154 stabilizing mutations (UP), explaining details of the difference in the effects of these mutations in combination with ligand binding. Specifically, increased positive allosteric signaling from all regulatory sites to F6P (e.g. $\Delta g_{F6P}^{(4xPEP,4xVal54\uparrow)} = 1.13$ kcal/mol and $\Delta h_{F6P}^{(4xADPa,4xVal54\uparrow)} = 1.21$ kcal/mol) is observed as a result of the 4xVal54 stabilizing mutations (compare allosteric free energies of signaling in mutated and non-mutated forms in Figures 2A and 2B). It is coupled with the prevention of conformational changes that presumably repress binding to the regulatory ADPa and PEP sites as a result of ADPf perturbation ($\Delta g_{PEP}^{(4xADPf,4xVal54\uparrow)} = -0.54$ kcal/mol and $\Delta g_{ADPa}^{(4xADPf,4xVal54\uparrow)} = -0.31$ kcal/mol), and no significant changes in signaling from the regulatory ADPa and PEP sites to ADPf were detected (Figures 2A and 2B). Thus, stabilizing mutations of Val54 shows that increased positive signaling to the functional site can be achieved via the recruitment of bulkier residues in place of Val54. Given that Val54 belonging to PEP site shows a positive modulation range ($\Delta h_{F6P}^{(4xVal54\downarrow\uparrow)} = 0.95$ kcal/mol, Table S2) and stabilizing 4xVal54 mutations increase a positive signaling to F6P from both ADPa and PEP regulatory sites (Figures 2A and 2B), one can presume that it may change the mode of PEP site action from an inhibiting to activating one.

A different picture of signaling in PFK is observed upon 4xArg154 stabilizing (UP) mutations, including some decrease in positive signaling from activating ADPa and inhibiting PEP

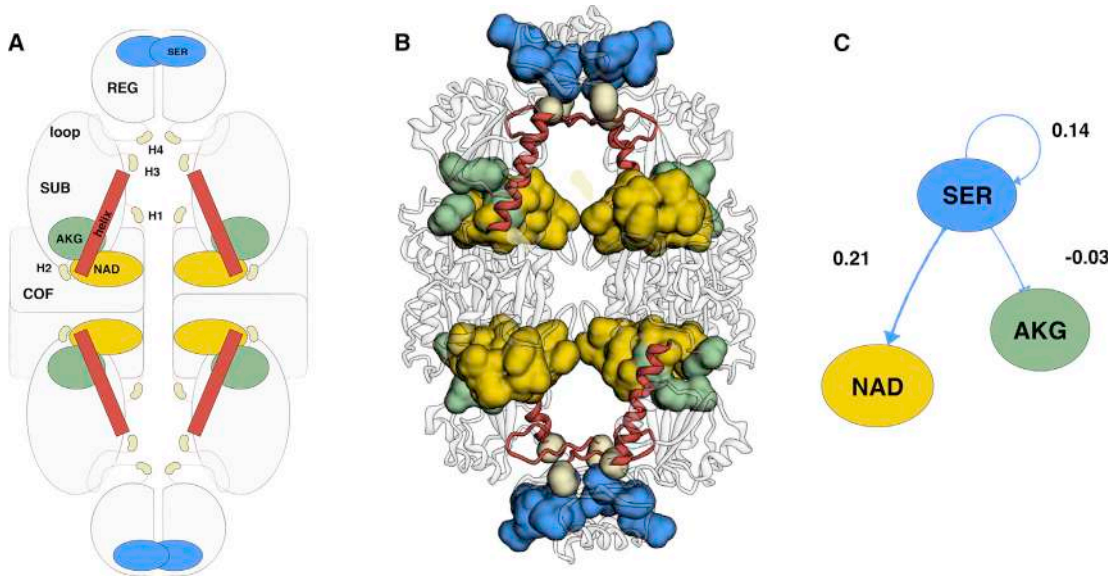


Figure 3. Domain-Based Allosteric Regulation in D-3-Phosphoglycerate Dehydrogenase

(A) Cartoon representation of the D-3-phosphoglycerate dehydrogenase (PGDH) with indicated hinges and helices, which are crucial for allosteric signaling between regulatory (contains SER-binding site), substrate (AKG), and cofactor (NAD) domains.

(B) Ribbon representation of the PGDH, showing allosterically relevant structural connections (helices and hinges) between regulatory and functional domains.

(C) Network scheme representation of the allosteric signaling in PGDH upon L-serine binding.

sites to F6P, as well as from ADPf to PEP sites (see Figures 2A and 2C for corresponding allosteric modulations). Since Arg154 is a bulky residue by itself, the only noticeable effect of its UP mutation is lower in absolute value allosteric signaling between the regulatory and functional sites. Decreased positive allosteric modulation on F6P upon ADPa binding compared with wild-type regulation together with overall negative modulation simply show that bulky substitutions in Arg154 are rather disruptive for the overall allosteric signaling in PFK. Finally, opposite and weak cooperativity, positive in the case of Val54 and negative in Arg154 mutations (Figure 2E and Table S1), respectively, further supports the difference in the ways of modulation of allosteric signaling in PFK encoded in these two residue positions.

D-3-Phosphoglycerate Dehydrogenase

The ring-shaped tetrameric PGDH is the allosterically regulated enzyme that catalyzes formation of 3-phosphohydroxypyruvate from 3-phospho-D-glycerate with NAD as cofactor. Contrary to the majority of allosterically regulated enzymes PGDH is a V-type enzyme, that is, the rate of reaction but not the substrate binding is allosterically modulated (Al-Rabiee et al., 1996). Each PGDH subunit contains three well defined structural domains: the cofactor-binding domain (NAD sites), the substrate-binding domain (AKG sites), and the regulatory domain. PGDH is allosterically inhibited by the L-serine (SER site) via the binding at the interface between the regulatory domains of adjacent units. Structural analysis suggests that PGDH catalytic activity and its allosteric inhibition occurs through the hinge-mediated rigid motion of domains (Al-Rabiee et al., 1996). The hinges H2 and H3, as the caps of helix Gln298-Asn318, and an additional hinge H4, constitute a pin-shaped spring, which performs the function of a signal-transferring device from the regulatory to the substrate domains (Figures 3A and 3B). The L-serine binding causes simultaneous opening of the interface between adjacent regulatory domains and of the active-site cleft, inhibiting the catalytic activity (Grant et al., 1996). Our calculations (active form was used; PDB: 1yba) show that binding perturbation to the SER effector site causes different modes of allosteric modulation on the effector ($\Delta h_{SER}^{(4xSER)} = 0.14$ kcal/mol), cofactor ($\Delta h_{NAD}^{(4xSER)} = 0.21$ kcal/mol), and substrate-binding (AKG, $\Delta h_{AKG}^{(4xSER)} = -0.03$ kcal/mol) sites located in corresponding PGDH domains. These results corroborate the modular response of PGDH discussed elsewhere (Al-Rabiee et al., 1996), which is in agreement with the idea that SER binding induces a more open form of the interface between regulatory domains (Grant et al., 2002) reflected in a positive modulation observed in them. Additionally, weak allosteric modulation detected at the substrate site AKG can apparently serve as an indication of the V-type nature of PGDH enzyme (Al-Rabiee et al., 1996) with modulation of the reaction rate, but not the substrate binding (Grant et al., 1996).

The major question we address in the PGDH analysis is whether and to what extent mutations can affect domain-based mode of allosteric communication and regulation of protein activity. Complete data on the effects of mutations on the functional and allosteric sites of PGDH are presented in Figure S2. We specifically considered mutations (Table 1) in hinges H1 (Pro105, Phe106), H2 (Gly294, Gly295), H3 (Gly319, Ser320), and H4 (Gly336, Gly337), and in the SER-binding site (Ile365 and Leu370). Mutations of residues in H1 result in the negative modulation range in the cofactor and substrate domains (Table 1), which suggests an overall decrease of motion of these domains

Table 1. List of Residues Mutated in PGDH Protein

Mutation *m*		$\Delta h_{SER}^{(m\downarrow\uparrow)}$ (kcal/mol)	$\Delta h_{AKG}^{(m\downarrow\uparrow)}$ (kcal/mol)	$\Delta h_{NAD}^{(m\downarrow\uparrow)}$ (kcal/mol)
Pro105 (H1)	1x	−0.02	−0.07	−0.03
	4x	0.05	−0.41	−0.14
Phe106 (H1)	1x	0.15	−0.07	−0.06
	4x	0.21	−0.32	−0.16
Gly294 (H2)	1x	−0.07	−0.08	−0.01
	4x	−0.16	−0.12	−0.02
Gly295 (H2)	1x	−0.02	−0.02	−0.02
	4x	−0.06	−0.05	−0.06
Gly319 (H3)	1x	−0.05	0.01	0.06
	4x	−0.35	−0.07	0.36
Ser320 (H3)	1x	−0.04	0	0
	4x	−0.09	−0.01	0.22
Gly336 (H4)	1x	−0.15	−0.04	0.18
	4x	−2.22	0.37	0.69
Gly337 (H4)	1x	−0.28	−0.11	0.08
	4x	−1.14	0.15	0.49
Ile365 (SER)	1x	−0.19	0	0.09
	4x	−1.13	0.04	0.50
Leu370 (SER)	1x	0.04	0	0.02
	4x	0.22	−0.13	0.17

Single-residue (1x) and four-residue (4x) UP/DOWN mutations are considered, and the corresponding allosteric modulation ranges $\Delta h^{(m\downarrow\uparrow)}$ (Equation 9) are reported for the sites SER, AKG, and NAD.

around hinge H1 (Figures 3A and 3B). At the same time, mutations of Phe106 originate some structuring/orientation in the regulatory site caused by the positive modulation ($\Delta h_{SER}^{(4xPhe106\downarrow\uparrow)} = 0.21$ kcal/mol). Mutations of the hinge H2 do not significantly affect either of the sites/domains. Both hinges H3 and H4 are located between the regulatory and cofactor domains before the helix of the pin-shaped spring, which is crucial for conformational changes involving these domains. Mutations of residues in these hinges lead to the substantial allosteric modulation of allosteric response in PGDH caused by the L-serine binding (Figure 3): (1) they result in positive modulation on the cofactor domain and negative modulation on the regulatory sites (highest in the case of mutations in H4); and (2) mutations in H4 change the sign of modulation on the substrate-binding site compared with the effect of the L-serine binding, resulting in a positive work exerted on it in the case of mutations in all four subunits. All stabilizing mutations (Gly319/H3, Gly336/H4, and Gly337/H4) apparently lead to canceling the inhibition mode, as they may prevent binding of the inhibitor (negative modulation, $\Delta h_{SER}^{(4xGly336\downarrow\uparrow)} = -2.22$ kcal/mol and $\Delta h_{SER}^{(4xGly337\downarrow\uparrow)} = -1.14$ kcal/mol, Table 1) and, on the contrary, may enable binding of the substrate (positive modulation, $\Delta h_{AKG}^{(4xGly336\downarrow\uparrow)} = 0.37$ kcal/mol and $\Delta h_{AKG}^{(4xGly337\downarrow\uparrow)} = 0.15$ kcal/mol, Table 1). Additional positive work exerted in the cofactor site can additionally support binding of NAD ($\Delta h_{NAD}^{(4xGly336\downarrow\uparrow)} = 0.69$ kcal/mol and $\Delta h_{NAD}^{(4xGly336\downarrow\uparrow)} = 0.49$ kcal/mol, Table 1). Noteworthy, three of the four mutated residues are glycines, hence the effect of mutation is expected to be more pronounced especially in cases of four-residue mutation (Ser320/H3 mutation shows the weakest effect, see Table 1). Finally, the PGDH residues Ile365 and Leu370 induce opposite allosteric modulation on the substrate (AKG) and regulatory (SER) sites (Table 1). In addition to the importance of single-residue mutations in modulation of allosteric signaling observed for positions in difference hinges, this drastic difference in the results of mutations of similar in size and physical-chemical characteristics and the closely located Ile365 and Ile370 protein residues show that any residue position in the protein can potentially be a source of allosteric modulation.

Insulin-Degrading Enzyme (IDE)

IDE is a monomeric Zn^{2+}-dependent protease organized in four domains, with the interface between the amino-terminal (domains 1 and 2) and carboxy-terminal (domains 3 and 4) halves forming the degradation chamber. The cleavage (CLS), Zn^{2+}-binding (ZN), and β-recognition sites (AB) are spatially close in the structure, constituting the complex active site where the substrate binding, recognition, and hydrolysis take place. The substrate acquisition is supported by its anchoring in the additional exosite (EXO) (Shen et al., 2006), and its proteolysis is allosterically regulated by the ATP bound to remote (ATP) site (Kurochkin et al., 2018). Although insulin was the first molecule recognized as the substrate, hence ''insulin-degrading enzyme'', it appeared that this protein has a unique specificity toward β-structure-forming aggregation-prone molecules, including Aβ and other amyloidogenic peptides, hormones, chemokines, and growth factors to name a few (Kurochkin et al., 2018). It is of great importance for the context of this work that in addition to IDE allosteric effector ATP, protein activity can be also allosterically modulated and targeted against specific substrates by different molecular groups, small molecules, post-translational modifications, and mutations (Kurochkin et al., 2017, 2018).

Since the EXO is critical for anchoring substrates and the ATP-binding (ATP) site provides allosteric modulation of IDE activity, we first explore effects of binding to these sites on the CLS, ZN, and AB sites (the Zn-bound form of IDE in complex with insulin B chain was used; PDB: 2g54, see Figure 4A). In general, perturbation of the EXO and ATP sites provides rather modest allosteric signaling to the catalytic sites ($\Delta h_{CLS}^{(EXO)} = -0.23$ kcal/mol and $\Delta h_{CLS}^{(ATP)} = 0.09$ kcal/mol), except the relatively high positive modulations of ZN ($\Delta h_{ZN}^{(EXO)} = 0.94$ kcal/mol) and AB ($\Delta h_{AB}^{(EXO)} = 0.57$ kcal/mol) sites upon binding perturbation of the EXO site. Positive modulation on the AB and ZN sites becomes stronger ($\Delta h_{ZN}^{(EXO,ATP)} = 1.23$ kcal/mol, $\Delta h_{CLS}^{(EXO,ATP)} = 0.08$ kcal/mol, $\Delta h_{AB}^{(EXO,ATP)} = 0.83$kcal/mol) when both EXO and ATP sites are perturbed, pointing to the importance of mutual effect of binding to these sites for modulation of the catalytic site activity (see Figure 4B). Binding to these sites also allosterically affects the energetics of domains, which was shown to be important for IDE activity and its modulation (Kurochkin et al., 2017). Perturbation of the EXO site causes positive modulation in domains 1 and 4 and negative modulation in domains 2 and 3 (see table in Figure 4C). ATP binding originates from a positive modulation in domain 2, apparently facilitating binding to the EXO site located in this domain Figure 4C. Mutual binding to EXO and ATP sites leaves positive

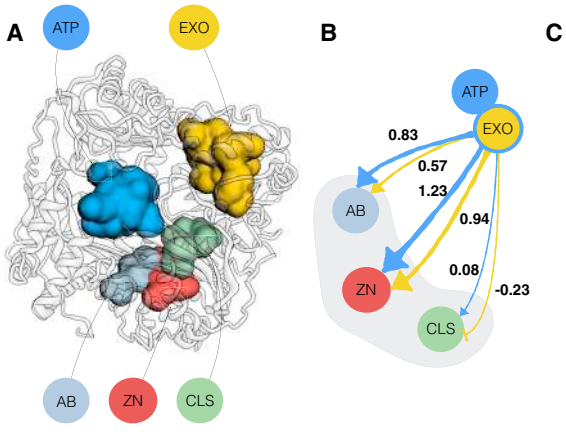


Figure 4. Allosteric Signaling between Functional and Regulatory Sites of Insulin-Degrading Enzyme

(A) Ribbon representation of insulin-degrading enzyme (IDE) shows location of functional and regulatory sites: AB, recognition site; CLS, cleavage site; EXO, exosite; ZN, Zn^{2+}-binding site; ATP, ATP-binding site.

(B) Allosteric signaling from ATP and EXO sites to the components of the functional site (AB, CLS, and ZN).

(C) The values of modulation detected in IDE domains (shown below as a graphical cartoon) as a result of the binding perturbation at the sites—EXO, ATP, and combined EXO + ATP—are reported in a table.

modulation only on domain 1 (Figure 4C), characterizing the state of the protein in which the substrate is bound and is being processed.

Recent studies of five allosteric mutations, including both computational analysis and experimental verification, showed that specific modulation of IDE activity against Aβ substrate can be achieved by allosterically affecting the energetics of domains and components of the functional site (Kurochkin et al., 2017). It seems promising, therefore, to use allosteric effects of mutations for activation of IDE against targets of interest, leaving its activity against other substrates unchanged. As the first step in this direction we obtained here complete data on the allosteric effect of mutations on the energetics of functional and regulatory sites (sites EXO, CLS, ZN, AB, ATP; Figure S3) and on the allosteric modulation of domains (domains 1–4, Figure S4). Mutations in positions located in ATP and EXO sites clearly show an activating nature of these sites on IDE activity (Figure S3). In particular, all elements of the functional site (AB, CLS, EXO, and ZN) yield positive allosteric modulation range, which is an indicator of conformational changes facilitating protein activity as a result of allosteric signaling from the ATP site. Mutations in the exosite EXO provide positive modulation range in AB, ATP, and ZN sites, and negative range in CLS site, which apparently points to the importance of recognizing β-forming substrate before its processing in the CLS site (Figure S4). Mutations in AB and ZN sites induce a positive modulation range in ATP and EXO sites. Because of their relative closeness (see Figures 4C and S3), these sites work against each other and the CLS site, which may also suggest that the sequence of events in the substrate processing concludes with the involvement of the CLS site. Sequential mutations in IDE domains show a consistent picture of the positive modulation range in the diagonal domains, that is, in the pairs of domains 1-3/3-1 and 2-4/4-2 (Figure 4C). Mixed positive-negative modulation range is systematically observed in domains flanking domain n, namely domains $(n-1)$ and $(n+1)$, in which mutations are performed. The patterns of mixed modulation range in flanking domains reflect a structural similarity between IDE domains, which are all homologous Aβ-roll folds. At the same time, all domains share at maximum 15% of sequence identity, which is reflected in numerous local diversity of observed modulation in terms of both the sign and the value.

Linking Allosteric Modulation and Functional Cost of Mutations in the PDZ Domain

Our next goal is to establish a quantitative link between the allosteric effects of mutations and experimentally observed data on the "functional cost of mutations", and to put these data in the context of direct mutational effects. To this end, we considered a widely studied mini-protein with well-documented allosteric regulation, the PDZ domain (Mclaughlin et al., 2012), which plays an essential role in the organization of complex signaling. PDZ recognizes disordered peptides and binds them with allosterically modulated strength. We have considered the PDZ domain structure of rat $PSD95^{pdz3}$ (PDB: 1be9) and performed exhaustive scanning of stabilizing (UP) and destabilizing (DOWN) mutations, calculating corresponding allosteric modulations, $\Delta h_{sub}^{(m\uparrow)}$ and $\Delta h_{sub}^{(m\downarrow)}$, for the substrate-binding regions. The values of modulation are averaged over the residues that are in contact with the substrate peptide with the contact distance up to 4.5 Å. In Figure 5, a comparison between the calculated allosteric modulations and the experimental functional cost (gain/loss of function) averaged over all possible mutations for each PDZ residue position is shown. The experimental data $\langle \Delta E_i^x \rangle_x$ were taken and adapted from Figure 2B in Mclaughlin et al. (2012) and are shown in Figure 5A along with the protein secondary structure profile. Data are also fully reported in Table S3. Most of the residue positions have a neutral effect on PDZ function upon mutations. However, 20 out of 83 residue positions (yellow) induce a significant functional loss in PDZ upon mutations (dark-blue color in the heatmap plot in Figure 5A). These residue positions are either in direct contact with the peptide substrate (strands β_2, β_3, and helix α_1) or are located within the allosteric protein region between helix α_1 and strand β_4 (Mclaughlin et al., 2012), suggesting the prominent role of corresponding position in orthosterically and allosterically affecting PDZ function. The "distance to substrate" chart at the bottom of Figure 5B shows that in some cases it is a clear orthosteric effect of mutations on residues directly binding the substrate, whereas in other cases it is the result of allosteric signaling from the residues distant to those that bind the substrate. In case of UP mutations, negative allosteric modulation (yellow circles located in the strands β_2, β_3, and helix α_1 positions in Figure 5B) is observed. A mild positive modulation at

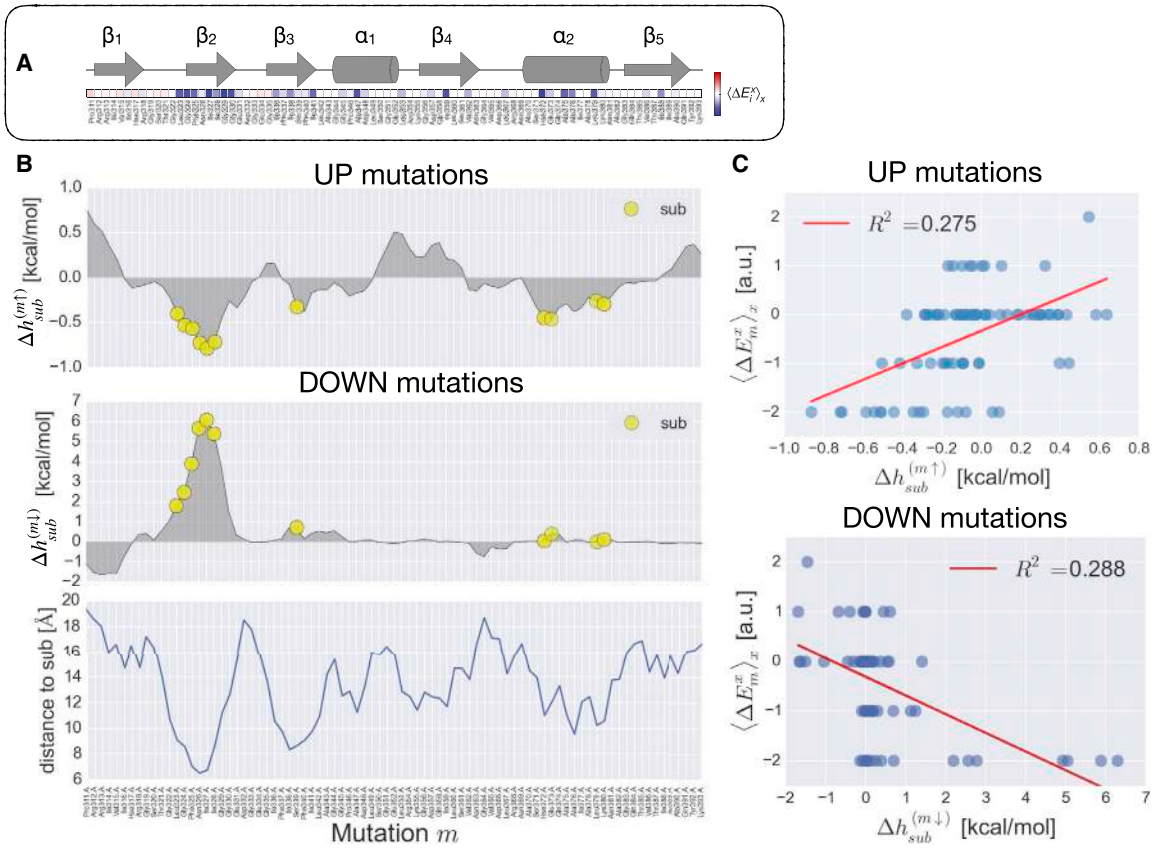


Figure 5. Allosteric and Orthosteric Modulation upon Mutations in the PDZ Domain

(A) Comparison between the allosteric modulations of the PDZ domain substrate-binding regions as a function of stabilizing and destabilizing mutations with the experimental functional cost $\langle \Delta E_i^x \rangle_x$ derived from the mean functional cost of mutations reported in Figure 2B of Mclaughlin et al. (2012).

(B) The allosteric modulation profiles upon stabilizing (UP) and destabilizing (DOWN) mutations $\Delta h_{\text{sub}}^{(m\uparrow)}$ and $\Delta h_{\text{sub}}^{(m\downarrow)}$, respectively, along with the mean distance of a residue position with substrate site (bottom chart).

(C) A positive correlation trend is detected between the experimental data of mean functional cost $\langle \Delta E_i^x \rangle_x$ and the calculated modulation upon stabilizing mutations, whereas the opposite is observed with the modulation upon destabilizing mutations.

the substrate site is observed for mutations in the known allosteric site at the boundary between helix α_1 and strand β_4. In the case of destabilizing (DOWN) mutations, a strong positive modulation of the substrate site is obtained for mutations in the substrate-binding site at the strand β_2. Similar to the case of stabilizing mutations, some residue positions affect function directly (orthosterically), while others do this via potential allosteric signaling.

The experimental data of functional cost $\langle \Delta E_i^x \rangle_x$ taken from Mclaughlin et al. (2012) were binned in five discrete values (-2, strong loss of function; -1, mild loss of function; 0, neutral; 1, mild gain of function; 2, strong gain of function, in arbitrary units) and used in linear regressions against the calculated allosteric modulation data. The results of these tests are shown in Figure 5C. Despite the low R^2 values, 0.275 and 0.288 for stabilizing and destabilizing mutations, respectively, a clear trend in both cases suggests that the allosteric modulation detected at the substrate site upon mutations is a good indicator of the relative functional importance of mutations in the allosterically relevant positions of the protein.

Exhaustive Analysis of Allosteric Signaling: Allosteric Signaling Maps (ASMs)

The analysis of three case studies, PFK, PGDH, and IDE, along with the additional analysis of experimentally characterized orthosteric and allosteric effects of mutations in PDZ domain, show that despite structural and functional differences, allosteric signaling is inherent in these proteins. It can be generically described using formalism of the allosteric modulation in per-residue approximation and can be quantified for functional and regulatory sites, (sub)domains, chains, and other structural units. Figure 6 contains distributions of allosteric modulation ranges as a result of point mutations Val54 in PFK, Gly336 in PGDH, and Phe807 in IDE. Positive allosteric modulation is exemplified by the sites F6P, NAD, and EXO, as well as by the individual residues (200 in PFK and PGDH, and 500 in IDE) in these proteins. Although, in general, distributions are skewed to the positive values, the shapes and ranges of negative/positive modulations are specific for each protein, depending on their individual structural characteristics, such as oligomerization state, types of folds/domains forming the protein, and linkers/hinges between them. Mutations of small residues, such as Val54 (PFK) and

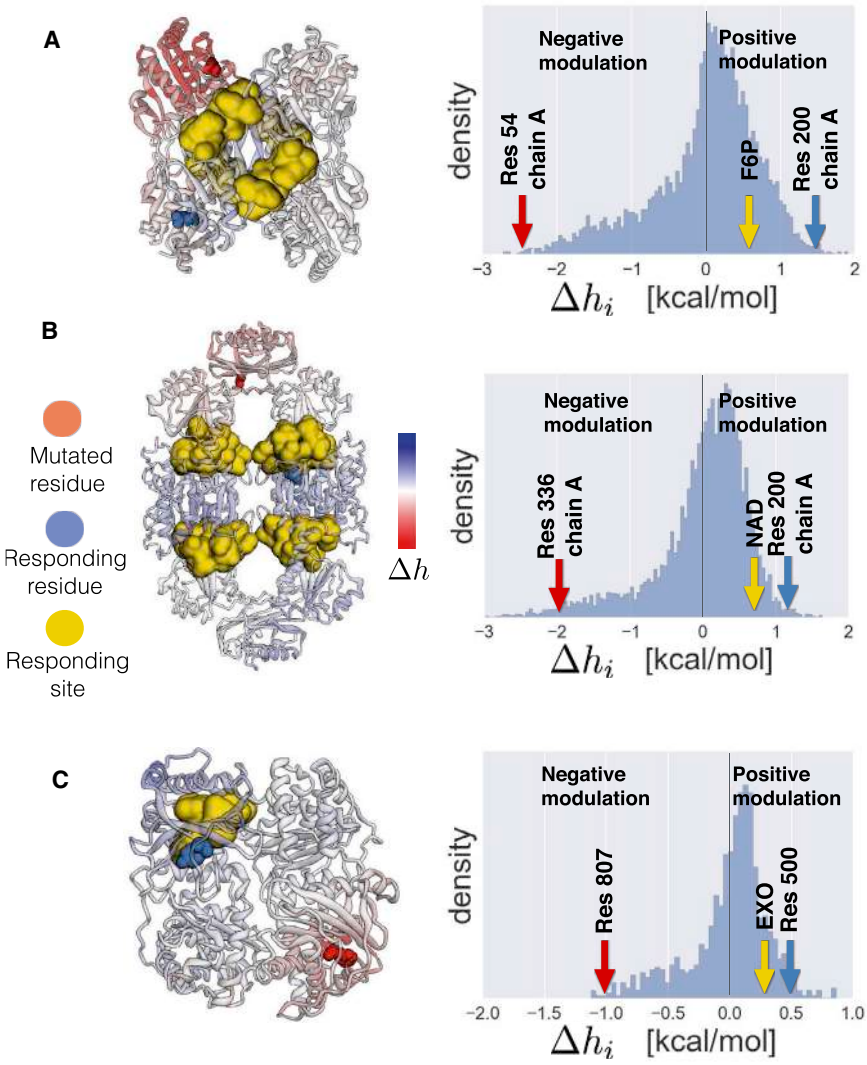


Figure 6. Distributions of Allosteric Modulation Ranges Caused by Single-Residue Mutations in phosphofructokinase, PGDH, and IDE

(A) Mutation of residue 54 in phosphofructokinase (PFK).

(B) Mutation of residue 336 in PGDH.

(C) Mutation of residue 807 in IDE.

Gly336 (PGDH), are expected to result in a stronger modulation imposed on responding sites and residues as well as stronger local negative modulation around mutated residues compared with substitution of rather bulky Phe807 (IDE). In other words, the sequence dependence of allosteric modulation is also at play in addition to structure dependence shown above. Therefore, to obtain a complete picture of the regulation and to be able to consider any individual or combined effect(s) of the ligand(s) binding and/or mutation(s) in the context of structure-function relationship, one has to perform an exhaustive scan of mutations and obtain their sequence-structure-dependent modulatory effects.

The calculation of the allosteric modulation range $\Delta h_i^{(m\downarrow\uparrow)}$ (Equation 9) upon residue-by-residue mutations provides the complete pattern of the allosteric response upon perturbation, which we call the ASM. Importantly, because of the nature of the allosteric potential, the elastic energy applied to a residue depends on the configurational states of the neighboring residues; hence, the ASM is an inherently asymmetric matrix. Figure 7 contains ASMs for the proteins studied in this work, PFK (Figure 7A), PGDH (Figure 7B), and IDE (Figure 7C), respectively. Major patterns of negative and positive modulations are clearly distinguishable in ASMs, revealing the domain and oligomeric structures of proteins. Specifically, negative modulation ranges delineate compact domains and monomers located along the diagonals of corresponding ASMs. In tetrameric PFK each monomer consists of two domains strongly interacting with each other via the last quarter of the second domain (Figure 7A). Each of the PGDH subunits consists of the cofactor, substrate, and regulatory domains clearly visible along the matrix diagonal (Figure 7B). Four domains of the single-chain IDE also form a characteristic pattern of negative modulation (Figure 7C).

A chiefly negative modulation range within compact structural units (distance matrices are shown in Figure S5) is a consequence of different effects of UP and DOWN mutations (Figure 7). UP mutations dominate the modulation, varying from strongly negative typical for overstabilizing interactions of the mutated residue environment to weakly positive at longer distances (Figures 8 and S5). The change of sign in modulation by UP mutations and almost complete decay from originally weak signal from DOWN mutations results in the change of the modulation mode from negative to positive at the distances 29 Å in PFK, 40 Å in PGDH, and 34 Å in IDE, respectively (Figures 7 and S5). These critical lengths at which allosteric signaling changes its sign match the typical radius of gyration for protein domain sizes ranging from 50 to 350 residues (Lobanov et al., 2008). Remarkably, recent analysis of the chemical-shift perturbation datasets originating from the ligand binding and mutations (Rajasekaran et al., 2017) revealed a universal distance-dependent decoy of the allosteric signal percolation within 20–25 Å. Figures 7 and 8 supported by the above experimental data emphasize, therefore, the two-faceted nature of allosteric modulation: (1) negative modulation dominates inside compact structural units where perturbation takes place, causing (2) mostly positive modulation in other parts of the protein starting at the borders of perturbed monomers/domains and percolating through the rest of the protein as a result of the low-frequency normal mode dynamics. It should be noted, however, that modulation of opposite signs observed in both cases points to the importance of

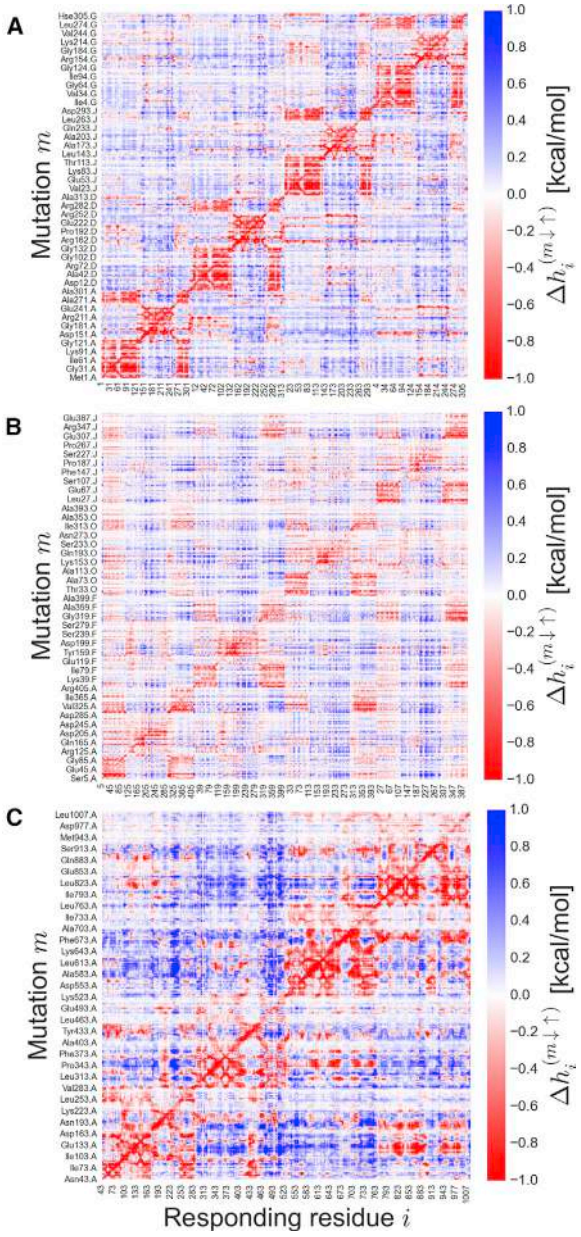


Figure 7. Allosteric Signaling Maps

Allosteric signaling maps for the proteins PFK (A), PGDH (B), and IDE (C). The allosteric effect caused by the generic mutation *m* (y axis) on a responding residue *i* (x axis) is expressed via the allosteric modulation range $\Delta h_i^{(m\downarrow\uparrow)}$ (Equation 9). Negative modulation (red) on responding residues indicates a decrease of work applied on them: stabilization. The increase in the work associated with positive modulations (blue) is indicative of local conformational changes.

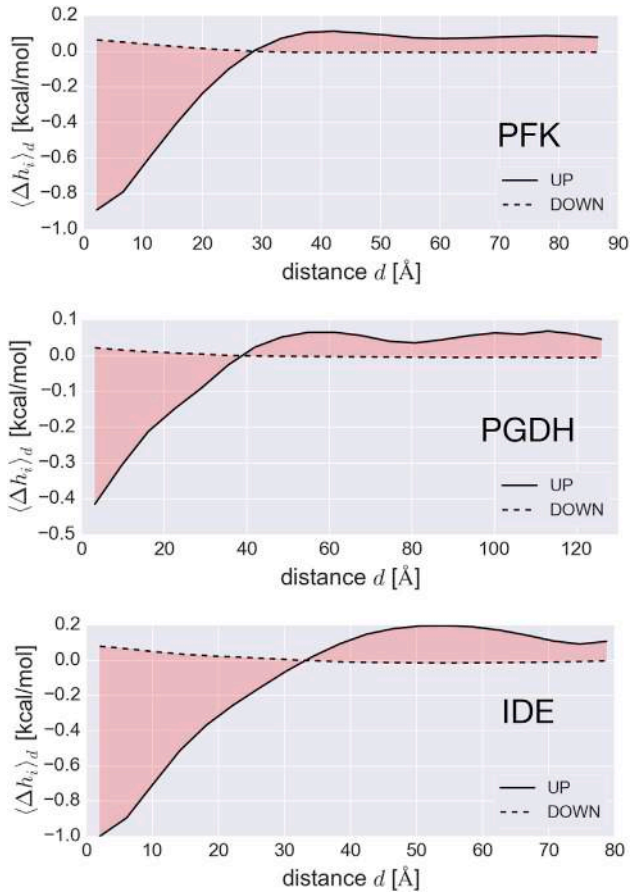


Figure 8. Critical Lengths of Allosteric Signaling

Critical lengths of the allosteric signaling in PFK (A), PGDH (B), and IDE (C). UP mutations cause strong mean negative modulation at short distances from the perturbation, starting from −0.93 kcal/mol in PFK, −0.42 kcal/mol in PGDH, and −0.98 kcal/mol in IDE, which converges toward positive residual value (0.11 kcal/mol in PFK, 0.06 kcal/mol in PGDH, and 0.09 kcal/mol in IDE). The mean modulation upon DOWN mutations is generally weak and positive (starts from 0.06 kcal/mol in PFK, 0.02 kcal/mol in PGDH, and 0.08 kcal/mol in IDE), decaying to a very small residual value. A critical length for the propagation of allosteric signals is observed from the distance where the mean allosteric modulations upon mutation switch their mode (29 Å in PFK, 40 Å in PGDH, and 34 Å in IDE).

corresponding signaling inside and outside of compact structural units.

The signaling observed between different sites of proteins (Figures 2A, 3C, and 4B), its modulation by mutations (Figures 2B and 2C; Table S1), and cooperativity of their actions (Figures 2 and S1; Table S2) are complemented in ASMs by the exhaustive analysis of the modulation ranges caused by the allosteric mutations of all residues in the protein. As a result, the global picture of allosteric modulation, and, at the same time, all relevant details in per-residue resolution can be quantified and analyzed. For example, in addition to the dependence of the effect of mutation on a certain site averaged over corresponding sites in all domains/monomers (Figures S1–S4), ASMs (Figure 7) show how individual mutation affects the energetics of corresponding sites in every domain/monomer. The difference between regulation inside homological domains and subunits of oligomers also becomes visible. Additionally, inspection of ASMs allows one to determine mutations of which positions located outside regulatory exosite originate allosteric signaling similar to that caused by mutations in the regulatory site of interest, thus allowing to obtain a required allosteric regulation caused by an alternative residue(s).

DISCUSSION

An existing consensus on the omnipresence and importance of allosteric mechanisms in the regulation of functional activity of proteins and molecular machines (Gunasekaran et al., 2004; Berezovsky et al., 2017) and an increasing demand on the design of allosteric drugs (Guarnera and Berezovsky, 2016a; Lu et al., 2014; Nussinov and Tsai, 2013) require an accurate theoretical understanding of allostery and development of computational models for achieving the high-resolution control (on the level of individual amino acids or atomic groups) of allosteric signaling and regulation. In this work, we combine protein harmonic modeling with statistical mechanical formalism in a perturbation-based approach, which allows estimation of per-residue allosteric free energy as a result of ligand(s) binding, mutation(s), and their combinations. We introduce the notion of allosteric modulation in order to distinguish targeted allosteric signaling to the sites/residues from the background effect caused by the same perturbation. We also formulate the notion of allosteric mutation, which we propose to consider as the basic element of the modulation of protein activity. We generically define the allosteric effect of mutation, "allosteric modulation range", as the difference between the responses in regulated sites/residues obtained from stabilizing and destabilizing mutations, respectively. Finally, we also propose the Allosteric Signalling Map (ASM) as an exhaustive description of allosteric signaling in the protein at per-residue resolution.

The generality of the model is shown here on three proteins with different degrees of oligomerization, functions, and mechanisms of regulation. Specifically, using the example of homotetrameric PFK we analyze a switch between the activating and inhibiting modes originating from the binding of activator/inhibitor ADPa/PEP in overlapping binding sites (Tee et al., 2018), as well as positive/negative cooperativity of their actions. The synergy of these effectors with a cosubstrate ADPf and opposite modulation provided by Val54 and Arg154 mutations are also discussed. The ring-shaped tetrameric PGDH was used to show how domain-based allosteric regulation works and how mutations in hinges between regulatory and functional domains affect protein activity. The dependence of allosteric regulation on the location of mutations, such as mutation Ile365 and Leu370, and type of the amino acid substitution, for instance stabilizing mutations Gly319/H2, Gly336/H4, and Gly337/H4, is also studied. IDE is a single-chain four-domain protein with the multicomponent catalytic site allosterically regulated via binding to EXO and ATP sites. A specific pattern of allosteric signaling with positive modulation in the diagonal domain and mixed modulation in flanking ones caused by sequential mutations in IDE domains was detected. We believe that this pattern of signaling can be explained by the structural homology of four IDE domains, whereas the diversity of sign and value of signaling in individual domains originates from the low sequence identity (at maximum 15%) between them. Additionally, by performing full mutational scanning on the PDZ domain mini-protein, we show that allosteric modulation can serve as a quantitative indicator of allosterically relevant positions in the protein.

Despite the simplicity of the harmonic model considered in this work, which does not consider any sequence specificity, a wide spectrum of allosteric phenomenology originating from ligand binding, mutations, and their combinations can be captured and estimated. For the improvement of quantitative evaluation while retaining the computationally efficient coarse-graining approach, the introduction of sequence-based parametrization would be a natural future development of the model.

Examples of allosteric signaling considered here show that the diversity in modulatory modes of regulatory sites and competitiveness/cooperativity in their work can be allosterically affected by the mutations located anywhere in the protein. We argue that complete mutational scanning can be instrumental in taking the protein activity under allosteric control. The resulting ASM can help to reveal a specific role of protein structural units (chains/domains) and associated critical length at which the mode switch in allosteric signaling takes place. The ASM identifies the strongest modulations, both positive and negative, inside these structures and between them. It also allows one to infer alternative ways of allosteric regulation. Starting from the analysis of ASMs and comparison of the effects of known allosteric sites and mutations, one can design new sites and use residue substitutions to obtain the required mode and value of allosteric signaling to sites/residues of interest. Mutual adjustment of regulatory exosites and ligands interacting with them on the basis of ASM-documented signaling is a step toward the design of allosteric drugs with the required mode of action and efficacy.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCE TABLE
- CONTACT FOR RESOURCE SHARING
- METHOD DETAILS
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and three tables and can be found with this article online at https://doi.org/10.1016/j.str.2019.01.014.

ACKNOWLEDGMENTS

E.G. thanks Prof. Martin Karplus for the insightful discussion on this work and Dr. Türkan Haliloglu for suggesting to compare the PDZ experimental dataset with results obtained with the model.

AUTHOR CONTRIBUTIONS

Conceptualization, E.G. and I.N.B.; Methodology, E.G. and I.N.B.; Investigation, E.G. and I.N.B.; Software, E.G.; Formal Analysis, E.G. and I.N.B.; Visualization, E.G.; Writing – Original Draft, E.G. and I.N.B.; Writing – Review & Editing, E.G. and I.N.B.; Supervision, I.N.B.; Project Administration, I.N.B.; Funding Acquisition, I.N.B.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 17, 2018
Revised: November 23, 2018
Accepted: January 25, 2019
Published: February 28, 2019

REFERENCES

Al-Rabiee, R., Lee, E.J., and Grant, G.A. (1996). The mechanism of velocity modulated allosteric regulation in D-3-phosphoglycerate dehydrogenase cross-linking adjacent regulatory domains with engineered disulfides mimics effector binding. J. Biol. Chem. *271*, 13013–13017.

Berezovsky, I.N. (2013). Thermodynamics of allostery paves a way to allosteric drugs. Biochim. Biophys. Acta *1834*, 830–835.

Berezovsky, I.N., Guarnera, E., Zheng, Z., Eisenhaber, B., and Eisenhaber, F. (2017). Protein function machinery: from basic structural units to modulation of activity. Curr. Opin. Struct. Biol. *42*, 67–74.

Blangy, D., Buc, H., and Monod, J. (1968). Kinetics of the allosteric interactions of phosphofructokinase from *Escherichia coli*. J. Mol. Biol. *31*, 13–35.

Cooper, A. (1976). Thermodynamic fluctuations in protein molecules. Proc. Natl. Acad. Sci. U S A *73*, 2740–2741.

Cooper, A., and Dryden, D.T.F. (1984). Allostery without conformational change. Eur. Biophys. J. *11*, 103–109.

Cui, Q., and Karplus, M. (2008). Allostery and cooperativity revisited. Protein Sci. *17*, 1295–1307.

Dokholyan, N.V. (2016). Controlling allosteric networks in proteins. Chem. Rev. *116*, 6463–6487.

Fan, Y., Cross, P.J., Jameson, G.B., and Parker, E.J. (2018). Exploring modular allostery via interchangeable regulatory domains. Proc. Natl. Acad. Sci. U S A *115*, 3006–3011.

Grant, G.A., Hu, Z., and Xu, X.L. (2002). Cofactor binding to *Escherichia coli* D-3-phosphoglycerate dehydrogenase induces multiple conformations which alter effector binding. J. Biol. Chem. *277*, 39548–39553.

Grant, G.A., Schuller, D.J., and Banaszak, L.J. (1996). A model for the regulation of D-3-phosphoglycerate dehydrogenase, a Vmax-type allosteric enzyme. Protein Sci. *5*, 34–41.

Guarnera, E., and Berezovsky, I.N. (2016a). Allosteric sites: remote control in regulation of protein activity. Curr. Opin. Struct. Biol. *37*, 1–8.

Guarnera, E., and Berezovsky, I.N. (2016b). Structure-based statistical mechanical model accounts for the causality and energetics of allosteric communication. PLoS Comput. Biol. *12*, e1004678.

Guarnera, E., and Berezovsky, I.N. (2019). On the perturbation nature of allostery: sites, mutations, and signal modulation. Curr. Opin. Struct. Biol. *56*, 18–27.

Guarnera, E., Tan, Z.W., Zheng, Z., and Berezovsky, I.N. (2017). AlloSigMA: allosteric signaling and mutation analysis server. Bioinformatics *33*, 3996–3998.

Gunasekaran, K., Ma, B., and Nussinov, R. (2004). Is allostery an intrinsic property of all dynamic proteins? Proteins *57*, 433–443.

Hinsen, K. (1998). Analysis of domain motions by approximate normal mode calculations. Proteins *33*, 417–429.

Hinsen, K. (2000). The molecular modeling toolkit: a new approach to molecular simulations. J. Comput. Chem. *21*, 79–85.

Hinsen, K., Petrescu, A.-J., Dellerue, S., Bellissent-Funel, M.-C., and Kneller, G.R. (2000). Harmonicity in slow protein dynamics. Chem. Phys. *261*, 25–37.

Jardetzky, O. (1996). Protein dynamics and conformational transitions in allosteric proteins. Prog. Biophys. Mol. Biol. *65*, 171–219.

Koshland, D.E., Némethy, G., and Filmer, D. (1966). Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry *5*, 365–385.

Kurochkin, I.V., Guarnera, E., and Berezovsky, I.N. (2018). Insulin-degrading enzyme in the fight against Alzheimer's disease. Trends Pharmacol. Sci. *39*, 49–58.

Kurochkin, I.V., Guarnera, E., Wong, J.H., Eisenhaber, F., and Berezovsky, I.N. (2017). Toward allosterically increased catalytic activity of insulin-degrading enzyme against amyloid peptides. Biochemistry *56*, 228–239.

Lau, F., and Fersht, A.R. (1989). Dissection of the effector-binding site and complementation studies of *Escherichia coli* phosphofructokinase using site-directed mutagenesis. Biochemistry *28*, 6841–6847.

Lobanov, M.Y., Bogatyreva, N., and Galzitskaya, O. (2008). Radius of gyration as an indicator of protein structure compactness. Mol. Biol. *42*, 623–628.

Lu, S., Li, S., and Zhang, J. (2014). Harnessing allostery: a novel approach to drug discovery. Med. Res. Rev. *34*, 1242–1285.

Mclaughlin, R.N., Poelwijk, F.J., Raman, A., Gosal, W.S., and Ranganathan, R. (2012). The spatial architecture of protein function and adaptation. Nature *491*, 138–142.

Meharena, H.S., Chang, P., Keshwani, M.M., Oruganty, K., Nene, A.K., Kannan, N., Taylor, S.S., and Kornev, A.P. (2013). Deciphering the structural basis of eukaryotic protein kinase regulation. PLoS Biol. *11*, e1001680.

Mitternacht, S., and Berezovsky, I.N. (2011). Binding leverage as a molecular basis for allosteric regulation. PLoS Comput. Biol. *7*, e1002148.

Monod, J., Wyman, J., and Changeux, J.-P. (1965). On the nature of allosteric transitions: a plausible model. J. Mol. Biol. *12*, 88–118.

Nussinov, R., and Tsai, C.-J. (2013). Allostery in disease and in drug discovery. Cell *153*, 293–305.

Proctor, E.A., Kota, P., Aleksandrov, A.A., He, L., Riordan, J.R., and Dokholyan, N.V. (2015). Rational coupled dynamics network manipulation rescues disease-relevant mutant cystic fibrosis transmembrane conductance regulator. Chem. Sci. *6*, 1237–1246.

Rajasekaran, N., Sekhar, A., and Naganathan, A.N. (2017). A universal pattern in the percolation and dissipation of protein structural perturbations. J. Phys. Chem. Lett. *8*, 4779–4784.

Shen, Y., Joachimiak, A., Rosner, M.R., and Tang, W.-J. (2006). Structures of human insulin-degrading enzyme reveal a new substrate recognition mechanism. Nature *443*, 870–874.

Tang, Q., and Fenton, A.W. (2017). Whole-protein alanine-scanning mutagenesis of allostery: a large percentage of a protein can contribute to mechanism. Hum. Mutat. *38*, 1132–1143.

Tee, W.-V., Guarnera, E., and Berezovsky, I.N. (2018). Reversing allosteric communication: from detecting allosteric sites to inducing and tuning targeted allosteric response. PLoS Comput. Biol. *14*, e1006228.

Traut, T.W. (2007). Regulatory Allosteric Enzymes (Springer Science).

Wenthur, C.J., Gentry, P.R., Mathews, T.P., and Lindsley, C.W. (2014). Drugs for allosteric sites on receptors. Annu. Rev. Pharmacol. Toxicol. *54*, 165.

STAR★METHODS

KEY RESOURCE TABLE

RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
MMTK	(Hinsen, 2000)	http://dirac.cnrs-orleans.fr/MMTK.html
Combined effect of ligand binding and mutation perturbations	This paper and (Guarnera et al., 2017)	http://allosigma.bii.a-star.edu.sg

CONTACT FOR RESOURCE SHARING

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Igor N. Berezovsky (igorb@bii.a-star.edu.sg).

METHOD DETAILS

For the characterization of the configurational ensembles associated with the reference and perturbed states, two sets of normal modes were calculated using the MMTK Python package (Hinsen, 2000) and utilizing the Fourier approximation (Hinsen, 1998). For the calculation of the free energy changes in Equation 6 only the 10 lowest frequency normal modes were used according to the considerations done in (Guarnera and Berezovsky, 2016b).

QUANTIFICATION AND STATISTICAL ANALYSIS

The combined effect of ligand binding and mutation perturbations according to Equation 2 is implemented in the updated version of the webserver AlloSigMA (Guarnera et al., 2017). Linear regression calculations were performed using the LinearRegression package implemented in the Sklearn Python library.

DATA AND SOFTWARE AVAILABILITY

The calculations of the Allosteric Signaling Maps were performed via Python scripts, which are implemented in the AlloSigMA server (Guarnera et al., 2017, see also Key Resource Table) and can be requested by e-mail.