

RESEARCH ARTICLE

Structure-Based Statistical Mechanical Model Accounts for the Causality and Energetics of Allosteric Communication

Enrico Guarnera¹, Igor N. Berezovsky^{1,2*}

1 Bioinformatics Institute (BII), Agency for Science, Technology and Research (ASTAR), Singapore, **2** Department of Biological Sciences (DBS), National University of Singapore (NUS), Singapore

* igorb@bii.a-star.edu.sg

Abstract

Allostery is one of the pervasive mechanisms through which proteins in living systems carry out enzymatic activity, cell signaling, and metabolism control. Effective modeling of the protein function regulation requires a synthesis of the thermodynamic and structural views of allostery. We present here a structure-based statistical mechanical model of allostery, allowing one to observe causality of communication between regulatory and functional sites, and to estimate per residue free energy changes. Based on the consideration of ligand free and ligand bound systems in the context of a harmonic model, corresponding sets of characteristic normal modes are obtained and used as inputs for an allosteric potential. This potential quantifies the mean work exerted on a residue due to the local motion of its neighbors. Subsequently, in a statistical mechanical framework the entropic contribution to allosteric free energy of a residue is directly calculated from the comparison of conformational ensembles in the ligand free and ligand bound systems. As a result, this method provides a systematic approach for analyzing the energetics of allosteric communication based on a single structure. The feasibility of the approach was tested on a variety of allosteric proteins, heterogeneous in terms of size, topology and degree of oligomerization. The allosteric free energy calculations show the diversity of ways and complexity of scenarios existing in the phenomenology of allosteric causality and communication. The presented model is a step forward in developing the computational techniques aimed at detecting allosteric sites and obtaining the discriminative power between agonistic and antagonistic effectors, which are among the major goals in allosteric drug design.



OPEN ACCESS

Citation: Guarnera E, Berezovsky IN (2016) Structure-Based Statistical Mechanical Model Accounts for the Causality and Energetics of Allosteric Communication. *PLoS Comput Biol* 12(3): e1004678. doi:10.1371/journal.pcbi.1004678

Editor: Jin Liu, University of North Texas System College of Pharmacy, UNITED STATES

Received: July 29, 2015

Accepted: November 25, 2015

Published: March 3, 2016

Copyright: © 2016 Guarnera, Berezovsky. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing Interests: The authors have declared that no competing interests exist.

Author Summary

The 50th anniversary of Monod-Changeux-Jacob seminal paper “Allosteric proteins and cellular control systems” became the hallmark of a new wave in the allostery studies and the turning point in our vision of allostery and its implications in protein engineering and drug design. Recent experimental and theoretical works clearly show relevance of allosteric phenomenon to drug design, unraveling advantages of allosteric drugs in comparison to

traditional orthosteric compounds. Remarkable simplicity of allosteric effectors and, at the same time, their potentially high specificity is one of the most important traits. The non conserved nature of allosteric ligands is a basis for avoiding drug resistance, and existence of latent regulatory sites make them attractive drug targets. The model presented in this work provides a theoretical framework for the quantification of the causality and energetics of allosteric regulation, which is a prerequisite for design of effector molecules with required characteristics. The synthesis between the thermodynamics of allostery and the intrinsic atomic nature of proteins and their interactions with the allosteric effectors accomplished in this work is a small initial step in the long endeavor towards future allosteric drugs.

Introduction

Activity modulation through ligand binding in sites other than the catalytic ones is a formidable way of performing biological function in allosteric proteins. Consideration of the experimental activity curves [1] together with the first available hemoglobin X-ray structure [2] had resulted in the introduction of the first theoretical concepts of allostery in the early 1960s. According to the Monod-Wyman-Changeux (MWC) phenomenological model a ligand binding to a symmetric oligomeric protein can shift the equilibrium between active and inactive conformational states, while preserving the symmetry of the oligomer [3]. The cornerstone of Koshland-Nemethy-Filmer (KNF) model is the “induced-fit” mechanism, where ligand binding to one of the monomers triggers conformational changes in its tertiary structure, which, in turn, lead to rearrangements in other parts of the oligomer with no structural symmetry preserved [4]. Despite their conceptual importance, MWC and KNF models share common limitations, such as the requirement of considering only oligomeric structures, and the lack of atomic level description. Additionally, the difference in the kinetics of the process—ligand binding precedes conformational change (KNF) or the opposite (MWC)—is a source of dichotomy between these phenomenological models.

Frauenfelder’s concept of the energy landscape with multiple conformational states [5–9] complemented by the development of NMR methods [10] have motivated a thermodynamic view of allostery as the foundation for the atomic level description [11]. Based on the statistical physics of protein dynamics, Cooper and Dryden showed that binding of the allosteric effector can induce changes in localization and frequency of fluctuations, which may result in cooperative binding even in the absence of substantial conformational changes [12]. Numerous experimental evidences [13, 11, 14], X-ray crystallography of thousands of proteins showed that allosteric regulation is omnipresent, and it should be studied in all types of proteins, including both oligomers and monomers, small single-domain enzymes and huge molecular machines [15, 16], receptors [17, 18], ion channels [19], and proteins with very different functions and cellular roles etc [20, 11].

The current vision of allostery is based on the understanding that intrinsic protein dynamics is modulated by the ligand binding which affects the structure’s function-related degrees of freedom. Hence, the coupling between the ligand binding and protein dynamics was shown to be instrumental in detecting the allosteric and catalytic sites. The strength of this link can be expressed via binding leverage [21] that measures the change in the dynamics of the binding site depending on its original plasticity, structure of the ligand, and the set of interactions that ligand makes with the binding site [22]. It was also shown that allosteric modulation takes place via communication between the allosteric and functional sites. Leverage Coupling

[15, 23] serves as the measure of allosteric communication, reflecting a coherence between degrees of freedom that determine dynamics in corresponding pairs of allosteric and catalytic sites. However, binding leverage and leverage coupling do not provide a quantifiable description of the effect of binding, causality of communication, agonistic/antagonistic nature of the ligands, energetics of the regulation etc.

In this work we present a new method for quantifying the configurational work exerted in different parts of the protein as a result of the ligand binding to the allosteric site—the allosteric free energy. Causal effects of the binding is modeled in a statistical mechanical framework by estimating the per residue free energy difference between the ligand bound and ligand free proteins. An allosteric potential based on normal modes is introduced on a “per residue approximation” for evaluating the energetics of structural changes induced by the overall protein dynamics of the protein in the local environment of a residue. Specifically, starting from the protein’s set of low frequency normal modes, the allosteric free energy of the system is obtained by estimating and comparing the partition functions that describe the conformational ensembles of the ligand free and ligand bound states at the single residue level. Thus, this model allows one to analyze a change in the allosteric free energy of a residue caused by the binding of a single ligand or by the sequential binding of several effector molecules.

We used here a set of proteins collected on the basis of thorough experimental data with well documented presence of the allosteric mechanisms in the regulation of their functions. We found several ways of regulation, dependent on the structures, oligomerization states, and functions of corresponding proteins. The increase of configurational work at the functional site is an archetypal mode of allosteric communication for hetero-oligomers, *e.g.* Aspartate carbamoyltransferase ATCase. A different mode of regulation was found, for example, in NAD-dependent Malic Enzyme (NADME) where binding at the allosteric site induces overall stabilization of the structure, including the functional sites. The inhibition of activity in this protein is supported by the increased dynamics in the catalytic site caused by the binding of ATP effector. The model also successfully discriminates between the negative and positive cooperativity in allosteric regulation, where Catabolite Activator Protein (CAP) and Dihydroxyacetone Kinase (DAK) are examples of the former and Phosphofructokinase (PFK) is an example of the latter. Finally, redistribution of the configurational work in D-3-phosphoglycerate dehydrogenase (PGDH) is strongly influenced by the ring-like shape of this protein, which is an example of the role of topology in allosteric communication.

Results

A structure-based model of allostery

A structure-based model of protein allostery should in general address two related issues: (i) providing a quantitative description of the causal effect of allosteric ligand binding on the protein’s structural dynamics and (ii) evaluating the communication between functional and regulatory sites and the modulation of the protein’s functional activity. Thermodynamical description of the protein states, the effects of the ligand binding, and allosteric cooperativity are globally characterized by the free energy changes of the protein. If, for example, thermodynamics of the two state model of allostery is considered, the ligand free protein populates the inactive R_i and the active R_a states with the Gibbs free energy gain/loss $\Delta G(R_i \rightarrow R_a) = -RT \ln [R_a]/[R_i]$, where $[R_a]/[R_i]$ is the equilibrium constant for the transition $R_i \rightarrow R_a$, R is the ideal gas constant, and T is the temperature. In the event of binding of an allosteric ligand the equilibrium between active and inactive states is shifted, that is the bound states AR_i and AR_a yield the following free energy difference $\Delta G(AR_i \rightarrow AR_a) = -RT \ln [AR_a]/[AR_i]$. The global stabilizing/destabilizing effect of the active state caused by an allosteric effector is thus given by the

free energy difference $\Delta\Delta G(R_a \rightarrow AR_a) = -RT\ln\alpha$, where $\alpha = [AR_a][R_i]/[AR_i][R_a]$ reflects the population shift in the active state.

Protein conformational states are the key determinants in the phenomenological thermodynamic model of allostery [14]. On the other hand, as the phenomenology of the allosteric communication involves coupling between allosteric and functional sites, structure-based models are usually characterized by the “per site/residue” descriptors [24]. The goal is to evaluate the free energy gain/loss $\Delta g_F(\mathcal{P} \rightarrow A\mathcal{P})$ upon transition from the ligand free protein conformational ensemble (for simplicity indicated with \mathcal{P}) to the protein conformational ensemble with the ligand A bound— $A\mathcal{P}$. The simplest case scenario of a protein \mathcal{P} with one allosteric site A and one functional site F , is a structure-based model that should link the ligand binding at the site A to the energetics of the site F . It is important to emphasize that this free energy is precisely the difference between the work exerted in the catalytic site of the ligand free and ligand bound protein states. We express the free energy gain/loss $\Delta g_F(\mathcal{P} \rightarrow A\mathcal{P})$ per site as the average of independent contributions of residues i belonging to a site F such as

$$\Delta g_F(\mathcal{P} \rightarrow A\mathcal{P}) = \frac{1}{n_F} \sum_{i \in F} \Delta g_i(\mathcal{P} \rightarrow A\mathcal{P}) \quad (1)$$

with $\Delta g_i(\mathcal{P} \rightarrow A\mathcal{P})$ the free energy gain/loss per residue caused by allosteric binding, and n_F is the number of residues in the site F . In order to estimate a relative strength of the allosteric effects, the average effect caused by the allosteric binding on all protein residues should be also evaluated:

$$\Delta g_P(\mathcal{P} \rightarrow A\mathcal{P}) = \frac{1}{n_P} \sum_{i \in P} \Delta g_i(\mathcal{P} \rightarrow A\mathcal{P}) \quad (2)$$

with n_P the total number of residues in a protein \mathcal{P} . Thus, the free energies $\Delta g_i(\mathcal{P} \rightarrow A\mathcal{P})$, $\Delta g_F(\mathcal{P} \rightarrow A\mathcal{P})$, and $\Delta g_P(\mathcal{P} \rightarrow A\mathcal{P})$ evaluate the causality and the energetics of allosteric communication at the residue, site, and protein levels respectively. In the case of a protein with multiple allosteric sites, one can also evaluate the free energy difference

$$\Delta\Delta g_i(A^{(n-1)}\mathcal{P} \rightarrow A^{(n)}\mathcal{P}) = \Delta g_i(\mathcal{P} \rightarrow A^{(n)}\mathcal{P}) - \Delta g_i(\mathcal{P} \rightarrow A^{(n-1)}\mathcal{P}) \quad (3)$$

which gives the modulating effect upon sequential binding of allosteric effectors, where the $A^{(n)}\mathcal{P}$ indicates a system with n ligands bound to the protein \mathcal{P} . Thus, similarly to eqs (1) and (2) one has

$$\begin{aligned} \Delta\Delta g_F(A^{(n-1)}\mathcal{P} \rightarrow A^{(n)}\mathcal{P}) &= \frac{1}{n_F} \sum_{i \in F} \Delta\Delta g_i(A^{(n-1)}\mathcal{P} \rightarrow A^{(n)}\mathcal{P}) \\ \Delta\Delta g_P(A^{(n-1)}\mathcal{P} \rightarrow A^{(n)}\mathcal{P}) &= \frac{1}{n_P} \sum_{i \in P} \Delta\Delta g_i(A^{(n-1)}\mathcal{P} \rightarrow A^{(n)}\mathcal{P}) \end{aligned} \quad (4)$$

Considering all the above, the main task of a structure-based model is the evaluation of the free energy gain/loss for any arbitrary residue i upon binding an effector molecule to an allosteric site. The method we propose for estimating the allosteric free energy $\Delta g_i(\mathcal{P} \rightarrow A\mathcal{P})$ consists of three major components. The first component is aimed at characterizing the ligand free \mathcal{P} and ligand bound $A\mathcal{P}$ protein systems by using a C_α harmonic model constructed on the basis of the available structures. The ligand bound system $A\mathcal{P}$ is obtained from the free system \mathcal{P} by harmonic restraining of all the pairs of residues belonging to the allosteric binding site A . In other words, the stabilization of the allosteric binding site A indirectly models the presence of the actual bound molecule, affecting local dynamics of the protein. Fig 1A illustrates both

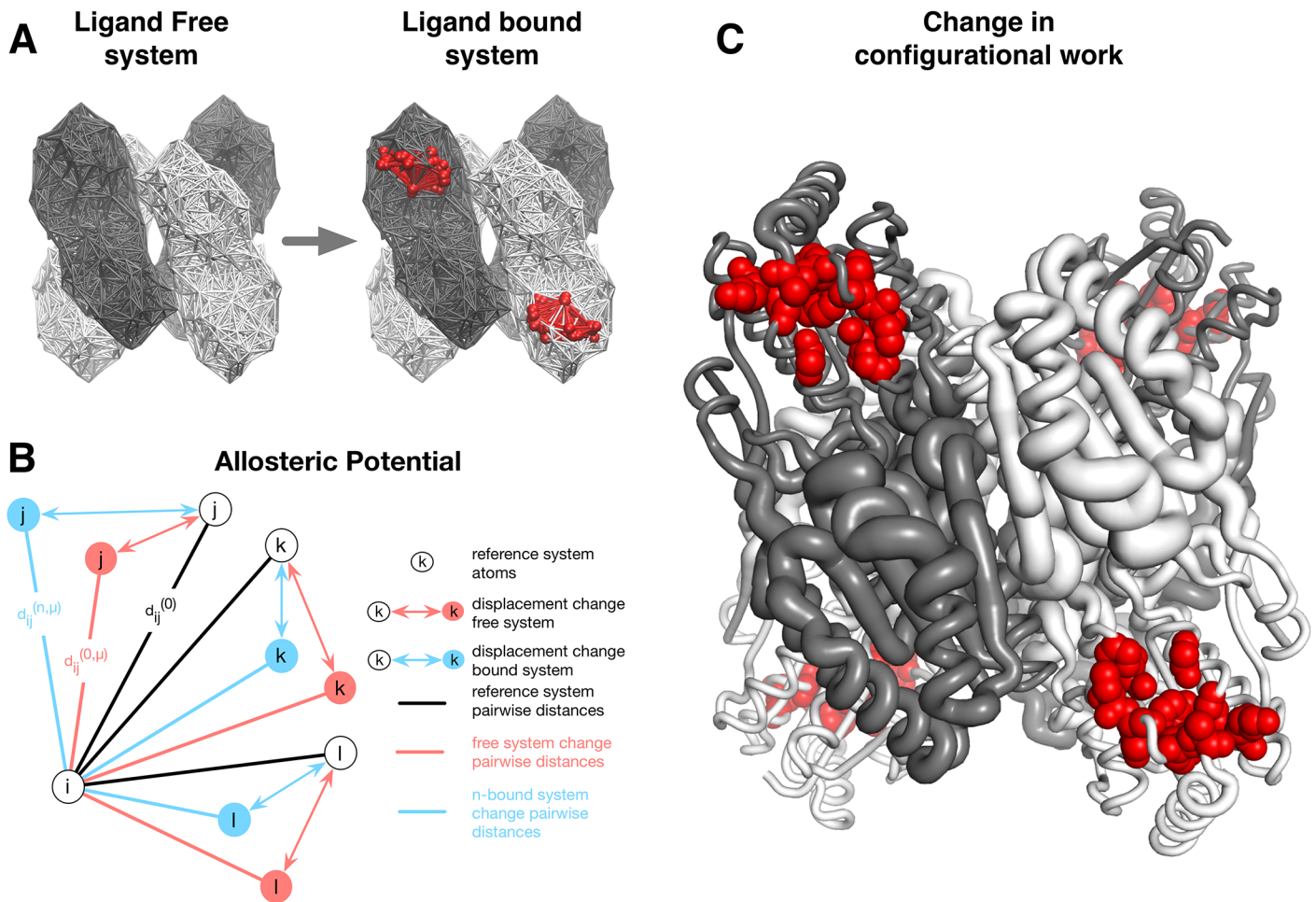


Fig 1. Pictorial representations of three major components of the model developed in this work. (A) The illustration of the ligand free and ligand bound states on the basis of the C_α harmonic approximation. The presence of a ligand at the allosteric binding site is modeled by local restraining of the residue pairs that belong to the binding site. Example of the ligated site is shown in red color. (B) Schematic representation of the concept of allosteric potential: neighbors of a residue i (residues j, k, l), assume different displacements as a consequence of the difference in the structure dynamics of the free and ligated proteins. (C) Cartoon representation of the configurational work gain/loss per residue caused by restraining of the allosteric binding sites. The radius ρ_i per residue i in the tube-like protein representation is determined by the value of the configurational work per corresponding residue, scaled according to the function $\rho_i = (\Delta g_i - \min; \Delta g_i) / (\max; \Delta g_i - \min; \Delta g_i)$. Thus, protein portions (residues, sites, domains) represented with a thick tube show increased dynamics upon restraining of the allosteric binding site and vice versa.

doi:10.1371/journal.pcbi.1004678.g001

ligand free and ligand bound proteins where residues and pairs of interactions belonging to the allosteric binding sites are red colored. Thus, in our definition of the ligand bound system a presence of the ligand is overwhelming, modeling all possible contacts within the binding sites. At the same time, no details of the contact interactions between ligand and binding site residues are accounted for (see [Materials and Methods](#) for the details of the harmonic model and the energy functions involved).

For exploring protein dynamics two sets of normal modes $e_\mu^{(P)}$, $e_\mu^{(AP)}$ are obtained from the harmonic protein model for the ligand free and ligand bound systems, respectively. These sets of normal modes are used as inputs for the *allosteric potential*, which is the second component of our method. The potential is used for evaluating the energetics that reflects structural differences induced by the different dynamics in the ligand free and ligand bound systems at the

single residue level. The per residue allosteric potential used here is:

$$U_i(\sigma) = \frac{1}{2} \sum_{\mu} \varepsilon_{\mu,i} \sigma_{\mu}^2 \quad (5)$$

with the summation running over the frequency modes. The parameters $\varepsilon_{\mu,i}$ are characteristics of the ligand free and ligand bound systems ($\varepsilon_{\mu,i}^{(P)}$ and $\varepsilon_{\mu,i}^{(AP)}$, respectively), and the coefficients σ_{μ} are gaussian variables with variance $1/\varepsilon_{\mu,i}$. Fig 1B provides a pictorial representation of the concept behind the per residue allosteric potential: in a ligand free and a ligand bound systems the neighbors of a residue i assume different displacements. The changes in dynamics of free and bound proteins is reflected in the normal modes $\mathbf{e}_{\mu}^{(P)}$, $\mathbf{e}_{\mu}^{(AP)}$ and, thus, in the mean elastic work that is exerted on residue i (see Materials and Methods for the details).

The last component of the structure-based model is a statistical mechanical approach for the calculation of the per residue allosteric free energy $\Delta g_i(\mathcal{P} \rightarrow AP)$ by comparing the configurational ensembles of the ligand free and ligand bound systems. As the allosteric potential is defined at the single residue level, the ensemble of configurations of a single residue is characterized by all possible displacements assumed by its neighbors. The displacements are obtained from the linear combinations of the low frequency normal modes $\mathbf{e}_{\mu}^{(P)}$, $\mathbf{e}_{\mu}^{(AP)}$ for the ligand free and ligand bound systems, respectively. Thus, the calculation of the canonical partition function per residue i gives the allosteric free energy gain/loss black

$$\Delta g_i(\mathcal{P} \rightarrow AP) = \frac{1}{2} k_B T \sum_{\mu} \ln \frac{\varepsilon_{\mu,i}^{(AP)}}{\varepsilon_{\mu,i}^{(P)}} \quad (6)$$

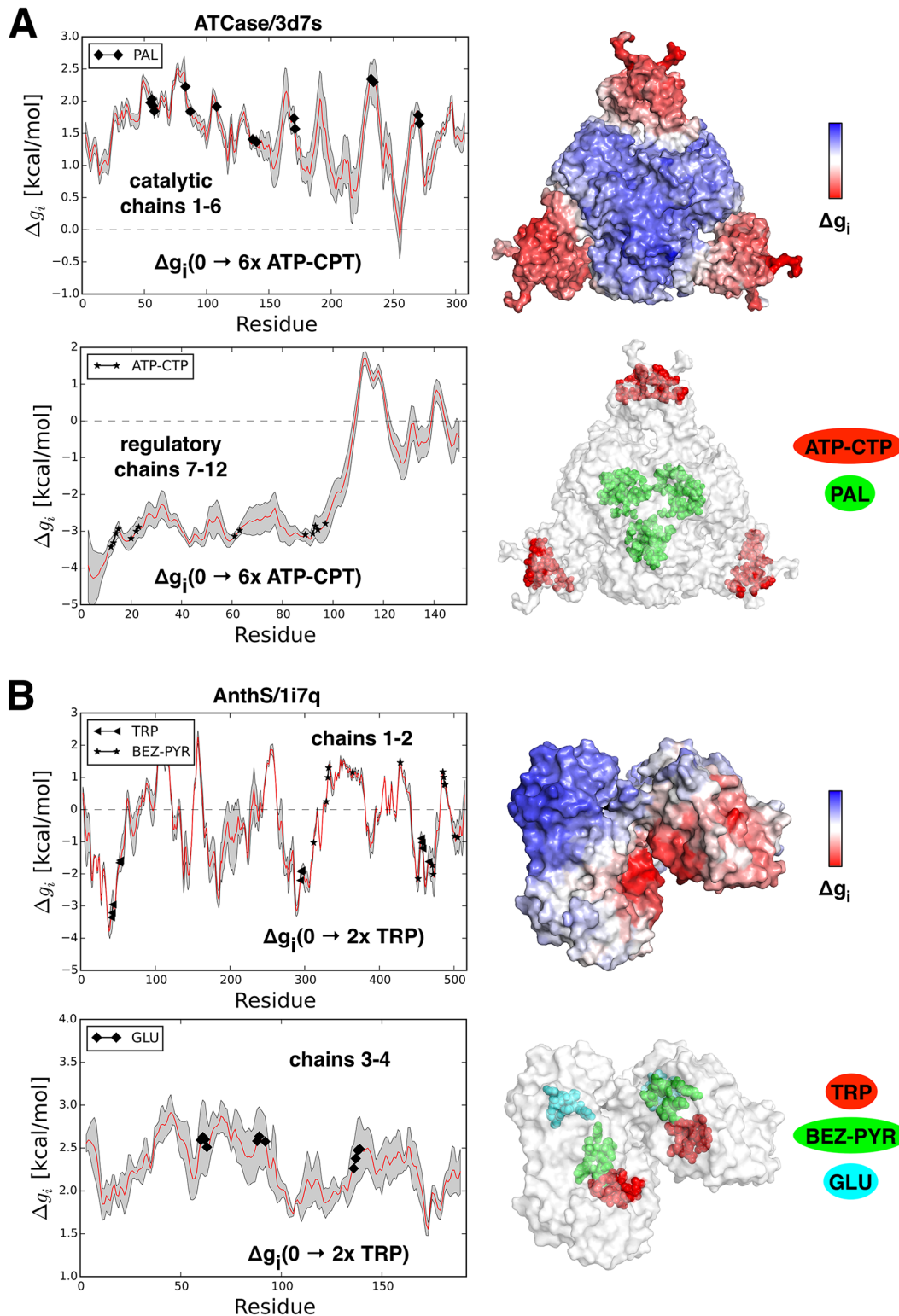
with the summation running over the low frequency modes, k_B is the Boltzmann constant, T is the temperature. Noteworthy, the free energy difference given by eq (6) is a purely configurational entropic contribution to the allosteric free energy, because the reference structure used in a harmonic model of the protein is the same in the ligand free and ligand bound systems. Therefore, the values of $\Delta g_i(\mathcal{P} \rightarrow AP)$ calculated with eq (6) should be interpreted as *the difference in the amount of configurational work* exerted on residue i by its environment due to the changes in the protein dynamics caused by the binding of an allosteric ligand. The resulting free energy profiles $\Delta g_i(\mathcal{P} \rightarrow AP)$ illustrate regions of the protein whose dynamics is affected by the ligand binding. Specifically, negative values of $\Delta g_i(\mathcal{P} \rightarrow AP)$ reflect stabilization in the corresponding regions, positive values, on the contrary, reflect destabilization. Fig 1C shows a tube-like representation of the protein backbone with thickness rescaled as the configurational work per residue. Thus, eqs (1) and (2) constitute a general basis for the analysis of the causality and the energetics of the allosteric communication. Similarly, sequential binding of ligands can result in positive and negative cooperativity. In order to quantify that, we operate with the notion of $\Delta \Delta g_i$ (eqs (3) and (4)), where the sign of the latter determines type of the cooperativity. It is worth noting that throughout the text—particularly in the Results section where protein cases are discussed—the terms allosteric free energy and exerted configurational work are used interchangeably.

Different scenarios and specifics of energetics in allosteric regulation

Analyzing set of proteins below, we show how our model describes different modes of allosteric regulation. In brief, the generic procedure for the analysis of allosteric communication consists of three major steps. First, using the harmonic protein model the set of normal modes that characterize the dynamics of the ligand free and ligand bound states are obtained. The allosteric ligand binding is modeled by restraining the residues that belong to allosteric sites described in

literature. Second, the allosteric potential is used to quantify the deforming effect of normal modes on the environment of protein residues. It evaluates the amount of elastic work that is exerted on protein residues of both the ligand free and ligand bound forms as a result of the proteins dynamics represented by a generic linear combination of normal modes. Third, from the allosteric potential a statistical mechanical approach is finally used for estimating the per residue partition functions. It allows one to obtain the per residue allosteric free energy gain/loss caused by effector binding. Left panels of main text figures including the proteins considered in the work and [S3 Fig](#) show profiles of the per residue allosteric free energy averaged over all homologous monomers (red curves) for the list of proteins analyzed in this work. The grey error bands around the profiles reflect the range of the work exerted per residue in each monomer (standard error), due to the structural differences between homologous monomers in the oligomeric structure. Protein surfaces (backbone in case of PGDH protein) are colored according to the detected allosteric free energy differences per residue. Protein structures are also represented with their binding sites to guide the eye in the analysis of the allosteric free energy profiles.

Aspartate carbamoyltransferase (ATCase). Aspartate carbamoyltransferase (ATCase) [25] is one of the best understood allosteric proteins, the hetero-dodecameric enzyme. The structure is organized in two trimers of the catalytic units and three dimers of the regulatory units ($2 \times 3 + 3 \times 2$). ATCase plays a crucial role in the early steps of pyrimidines biosynthesis. ATP is an allosteric activator of ATCase, which increases the reaction rate of pyrimidine synthesis. CTP is known to be an allosteric inhibitor, and high concentrations of the reaction's end products also negatively regulate pyrimidine yield. Both activator and inhibitor share the same binding site (ATP-CTP), which is located at the outer periphery of the three regulatory dimers. The catalytic binding sites are buried in the two centrally located trimers (PAL, green, [Fig 2A](#)). PAL binding induces a large conformational change from the inactive compact state to the active expanded state. It has been shown that the transition between inactive and active states can be described by the low frequency modes responsible for the large-amplitude motion of the quaternary structure [26]. We analyzed a free energy difference for the ligand free and bound states using four available structures: two inactive forms (apo PDB ID: 3d7s and CTP bound PDB ID: 1rac) and two active forms (PAL and ATP bound form PDB ID: 7ati and the PAL bound structure PDB ID: 1d09). Upon binding to the ATP/CTP site in both inactive structures, stabilization of the allosteric site (3d7s: $\Delta g_{ATP}(0 \rightarrow 6 \times ATP) = -3.35$ kcal/mol and 1rac: $\Delta g_{ATP}(0 \rightarrow 6 \times ATP) = -3.27$ kcal/mol) induces an overall freezing of the regulatory monomers $R - mer$ (3d7s: $\Delta g_{R - mer}(0 \rightarrow 6 \times ATP) = -2.02$ kcal/mol and 1rac: $\Delta g_{R - mer}(0 \rightarrow 6 \times ATP) = -1.61$ kcal/mol, red colored surfaces of the outer dimers, [Fig 2A](#)). This stabilization is compensated by a significant increase in the configurational work (destabilization) at the PAL binding site (3d7s: $\Delta g_{PAL}(0 \rightarrow 6 \times ATP) = 1.98$ kcal/mol and 1rac: $\Delta g_{PAL}(0 \rightarrow 6 \times ATP) = 1.94$ kcal/mol, blue surfaces of the two internal trimers, [Fig 2A](#)). Noteworthy, the configurational work exerted at the catalytic sites is higher than overall destabilization of the catalytic monomers $C - mer$ (3d7s: $\Delta g_{C - mer}(0 \rightarrow 6 \times ATP) = 1.47$ kcal/mol). At the same time, preliminary binding of the inhibitor (CTP in 1rac) practically eliminates a difference between the work exerted at the catalytic sites and the average work exerted in the whole catalytic unit (1rac: $\Delta g_{C - mer}(0 \rightarrow 6 \times ATP) = 1.83$ kcal/mol, see also [S1 Table](#)). Despite that ligand binding to the allosteric sites in activated structures (7ati and 1d09) destabilizes the PAL binding sites (7ati: $\Delta g_{PAL}(0 \rightarrow 6 \times ATP) = 1.88$ kcal/mol and 1d09: $\Delta g_{PAL}(0 \rightarrow 6 \times ATP) = 1.28$ kcal/mol, see [S1 Table](#)), the whole catalytic units are being destabilized almost to the same level (7ati: $\Delta g_{C - mer}(0 \rightarrow 6 \times ATP) = 1.73$ kcal/mol and 1d09: $\Delta g_{C - mer}(0 \rightarrow 6 \times ATP) = 1.18$ kcal/mol). These observations show that corresponding structures (7ati and 1d09) are already in the activated state, while binding to the ATP/CTP site in the inactive state (apo form 3d7s) induces a transition from the inactive to active states.



work (blue part of the Δg scale) as a result of allosteric communication between sites. Catalytic (PAL) and regulatory (ATP-CTP) sites are shown in green and red, respectively. In (B) the averaged free energy $\Delta g_i(0 \rightarrow 2 \times TRP)$ profiles are shown for chains 1–2 (BEZ and TRP site) and chains 3–4 (GLU site) of Anthranilate Synthase AnthS. Similarly to the ATCase allosteric sites, the restraining in chains 1–2 induces high configurational work in the chains 3–4 that contain catalytic sites. Here and in the following figures containing data on proteins the red curve in the chart shows allosteric free energy profiles, the grey error band reflects the range in the amount of work exerted per residue in each monomer, because of the structural differences between homologous monomers in the oligomeric structure. Surface representations are colored according to the conformational work exerted per residues in corresponding part of the protein (red—negative values of the conformational work, showing local stabilization; blue—positive values of the work, pointing to increase of the local dynamics). We also show representative structures with color-marked locations of the corresponding ligands.

doi:10.1371/journal.pcbi.1004678.g002

Anthranilate Synthase (AnthS). Anthranilate Synthase (AnthS) is a hetero-tetrameric enzyme catalyzing reaction that produces anthranilate (BEZ), pyruvate (PYR), and glutamate from two substrates—chorismate and L-glutamine (GLU). Substrates of this enzyme bind to two distinct sites, BEZ, PYR, and GLU (Fig 2B), mutually affecting each others activity. Tryptophan (TRP) is the allosteric down regulator for this enzyme [27] which binds at its designated site. Products BEZ and PYR share most of the residues composing their binding site, hence the BEZ-PYR is considered as a common site for both products and merged in our calculations. Two X-ray structures were analyzed for this protein: the active state structure (PDB ID: 1i7q) with three bound substrates (BEZ, PYR and GLU) and the inactivated structure (PDB ID: 1i7s) with a bound allosteric inhibitor (TRP). In both active and inactive states binding of the inhibitor leads to the stabilization of the catalytic binding site BEZ-PYR (see 1i7q in Table 1 and 1i7s in S1 Table). In the active form these effects are stronger, apparently reflecting the presence of bound substrates in the analyzed active structure (1i7q).

Bovine Glutamate Dehydrogenase (BGDH). With about 3000 residues the bovine glutamate dehydrogenase (BGDH) is the largest complex analyzed in our protein list (see Fig 3).

This homo-hexamer protein catalyses the oxidative deamination of l-glutamate to 2-oxoglutarate. It is up-regulated by ADP and down-regulated by GTP. The coenzyme NADP⁺ that binds in the NAD site is an antagonist. Catalytic sites (where substrates GLU and NDP bind) are located very close to each other in the catalytic cleft. ADP is found in the active and inactive conformations. It supposedly favors transition between two states by facilitating the opening of the catalytic cleft [28]. GTP binds only to the inactive conformation [29]. Calculations were performed on the three structures: the apo form (PDB ID: 1nr7) and two forms with a bound allosteric ligand, GTP (PDB ID: 1hwz) and ADP (PDB ID: 1nqt), respectively. In particular, we compared allosteric free energy in the ligand free state with those in structures with bound allosteric ligands (GTP or ADP). In the active apo form (1nr7) we observed that significant decrease ($\Delta g_{ADP}(0 \rightarrow 6 \times ADP) = -1.51$ kcal/mol) of the configurational work at the activator site ADP induces a small overall stabilization of the protein ($\Delta g_{1nr7}(0 \rightarrow 6 \times ADP) = -0.16$ kcal/mol)—see Fig 3A for the corresponding allosteric free energy profile. As a result, GLU site is strongly stabilized ($\Delta g_{GLU}(0 \rightarrow 6 \times ADP) = -0.88$ kcal), whereas NDP site yields lightly increased configurational work ($\Delta g_{NDP}(0 \rightarrow 6 \times ADP) = 0.11$ kcal). If the inhibitor (GTP) is bound, the active state is destabilized ($\Delta g_{1nr7}(0 \rightarrow 6 \times GTP) = 0.61$ kcal), while the NDP and GLU substrate sites are destabilized less than the units containing them (Table 1). Free energy changes associated with the substrate binding sites point to the activation of the protein via relaxation of the NDP binding site. Another active form of the protein, with ADP activator bound (1nqt), shows qualitatively similar results. High structural similarity (RMSD ~ 0.4 Å) between the apo (1nr7) and holo (1nqt) active forms of BGDH hints that binding of the ADP activator mostly stabilizes the inherently active apo form of the protein. The results of calculations for the inactive holo form (1hwz) with bound inhibitor (GTP) and both substrates (NDP and GLU) are very different compared to those obtained for the active forms. Binding of the activator results in the overall stabilization of the whole structure ($\Delta g_{1hwz}(0 \rightarrow 6 \times ADP) = -0.41$ kcal/mol) along with

Table 1. Results on allosteric causality and energetics obtained for proteins analyzed in this work. Data for only one representative structure is given in this table (for the complete list of proteins see [S1 Table](#)). First column provides protein names, degree of oligomerization and total number of residues. Second column: PDB ID of the protein and names of ligands (if not the apo form). Third column: number, name, and type (⁽¹⁾(A) allosteric activator, (I) allosteric inhibitor, and (S) substrate) of ligated binding sites in our model. Fourth column: the mean allosteric free energy difference of the ligated binding site $\Delta G_A(\mathcal{P} \rightarrow \mathcal{A}\mathcal{P})$ averaged over all the residues belonging to \mathcal{A} in all protein monomers, and the mean allosteric free energy difference averaged over all the residues of the protein monomers (in parentheses). Fifth column: names of the binding site under allosteric regulation, typically catalytic sites. Sixth column: the mean allosteric free energy difference (or work exerted at) of the regulated binding site $F \Delta G_F(\mathcal{P} \rightarrow \mathcal{A}\mathcal{P})$ averaged over all the residues belonging to F in all protein monomers, and the mean allosteric free energy difference averaged over all the residues of the protein monomers (in parentheses $\Delta G_{mer}(\mathcal{P} \rightarrow \mathcal{A}\mathcal{P})$). A TC case: ^(*) and ^(**) the values corresponding to the regulatory ΔG_{R-mer} and catalytic ΔG_{C-mer} monomers, respectively. AnthS: ⁽⁺⁾ and ⁽⁺⁺⁾ the values corresponding to the monomers containing TRP $\Delta G_{TRP-mer}$ and GLU $\Delta G_{GLU-mer}$. CAP: ^(*) and ^(**) the values corresponding to the bound monomer ΔG_{B-mer} and free monomer ΔG_{F-mer} .

Protein (#ch, #res)	PDB ID (ligands)	Ligated Site (A/I/S) ^(P)	$\Delta G_A(\rightarrow A) \times (\Delta G_{mer}(\rightarrow A))$ kcal/mol	Regulated Site	$\Delta G_F(\rightarrow A) \times (\Delta G_{mer}(\rightarrow A))$ kcal/mol
ATCase (2x3+3x2, 2736)	3d7s (apo)	6 x ATP/CTP (A/I)	-3.35 (-2.02) ^(*)	PAL	1.98 (1.47) ^(**)
AnthS (2+2, 1426)	1i7q (BEZ, PYR, GLU)	2 x TRP (I)	-2.87 (-0.42) ^(*)	BEZ+PYR GLU	-0.60 (-0.42) ^(*) 2.13 (2.29) ^(**)
BGDH (6, 2976)	1nr7 (apo)	6 x ADP (A)	-1.51 (-0.16)	NDP	0.11 (-0.16)
		6 x GTP (I)	-0.32 (0.61)	GLU	-0.88 (-0.16)
		6 x ADP (A)	-0.64 (0.60)	NDP	0.14 (0.61)
		6 x GTP (I)	-0.40 (0.60)	GLU	0.47 (0.61)
CAP (2, 418)	2wc2 (apo)	1 x cAMP (A)	-1.49 (-0.13) ^(#)	NDP	0.32 (0.60)
		2 x cAMP (A)	-2.37 (-0.30)	GLU	-0.01 (0.60)
DAHPS (4, 1401)	1ggl (PGA)	4 x PHE (I)	-2.33 (0.25)	DNA	0.88 (0.24) ^(##)
DAK (2, 735)	3iu5 (apo)	1 x ADP (S)	-1.93 (-0.72)	cAMP	-1.06 (0.24) ^(##)
		1 x ARG (S)	-0.76 (-0.72)	DNA	0.85 (-0.30)
		2 x ADP (S)	-0.85 (-0.05)	PGA	0.33 (0.25)
		2 x ARG (S)	-0.11 (-0.05)	ADP	0.49 (0.74)
G6PD (6, 1596)	1cd5 (apo)	6 x 16G (A)	-0.04 (0.31)	ARG	1.03 (0.74)
NADME (4, 2232)	1efk (NAD)	4 x FUM (A)	-0.53 (-0.14)	ADP	-0.85 (-0.05)
		4 x ATP (S/I)	-1.42 (-0.25)	ARG	-0.11 (-0.05)
PFK (4, 1284)	3plk (apo)	4 x PEP/ADPa (I/A)	-0.78 (0.16)	AGP	0.37 (0.31)
		4 x ADPf (S/A)	-0.74 (0.50)	NAD	-0.14 (-0.14)
		4 x PEP/ADPa (I/A)	0.35 (0.68)	NAD	0.11 (-0.25)
		4 x ADPf (S/A)	-1.08 (0.68)	F6P	0.71 (0.16)
PGDH (4, 1624)	1psd (SER, NAD)	8 x SER (I)	-0.58 (0.03)	ADPf	-0.11 (0.16)
		1 x ATP (I)	-4.28 (-0.10)	F6P	1.74 (0.50)
PKA (1, 336)	1atp (ATP)	1 x ATP (I)	-3.63 (-0.05)	F6P	2.22 (0.68)
PTP1B (1, 278)	2hnp (apo)	1 x 892 (I)	-2.89 (-0.55)	AKG	0.08 (0.03)
SSUPRT (4, 868)	1xtu (U5P, CTP)	4 x CTP (I)	-1.31 (0.02)	NAD	0.38 (0.03)
ThrS (2, 884)	1e5x (apo)	1 x SAM (A)	-3.67 (-0.24)	MPD	1.97 (-0.10)
		2 x SAM (A)		BPM	0.36 (-0.05)
				U5P	0.14 (-0.53)
				TRS	-0.31 (0.02)
				PLP	0.35 (0.02)
				TRS	-0.54 (-0.24)
				PLP	0.52 (-0.24)

doi:10.1371/journal.pcbi.1004678.t001

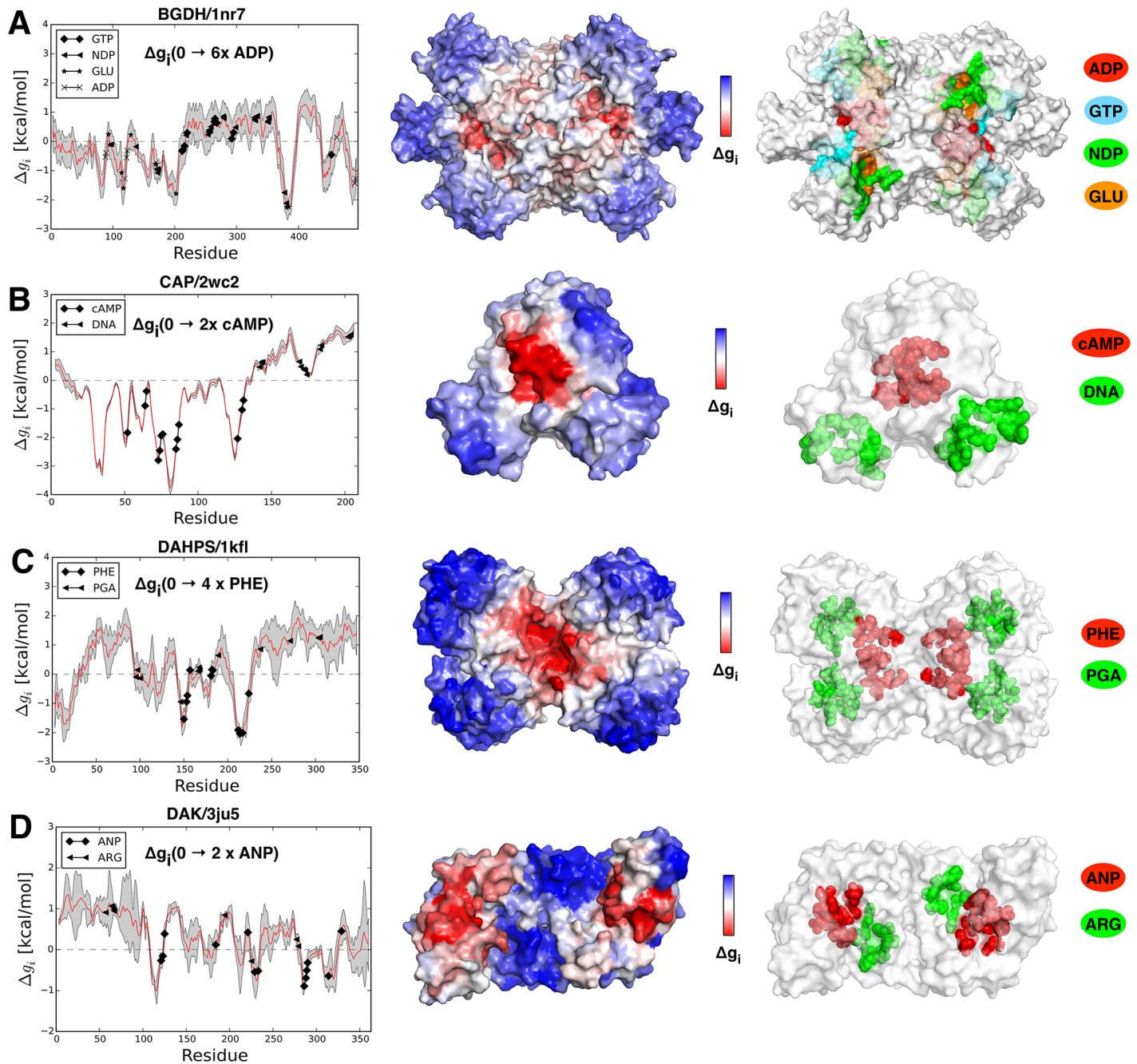


Fig 3. Free energy profiles and structural representatives. (A) bovine glutamate dehydrogenase BGDH free energy profile $\Delta g_i(0 \rightarrow 6xADP)$ with marked residue positions of the GTP, NDP, GLU, and ADP binding sites. (B) Catabolite activator protein (CAP) free energy profile $\Delta g_i(0 \rightarrow 2xcAMP)$ with marked locations of the cAMP and DNA binding sites. (C) The 3-Deoxy-D-arabinoheptulosonate 7-phosphate synthase DAHPS free energy profile $\Delta g_i(0 \rightarrow 4xPHE)$ with marked positions of the PHE and PGA binding sites. (D) The dihydroxyacetone kinase DAK free energy profile $\Delta g_i(0 \rightarrow 2xANP)$, ANP and ARG are ligands.

doi:10.1371/journal.pcbi.1004678.g003

quite pronounced freezing of the catalytic sites. These observations are in agreement with the role of activator as a stabilizer of the active apo form (in 1nr7 and 1nqt structures). Fig 4A illustrates complex mechanisms of the BGDH regulation (left scheme in Fig 4A). The right scheme in Fig 4A illustrates a toy experiment in which both allosteric regulators are bound (ADP and

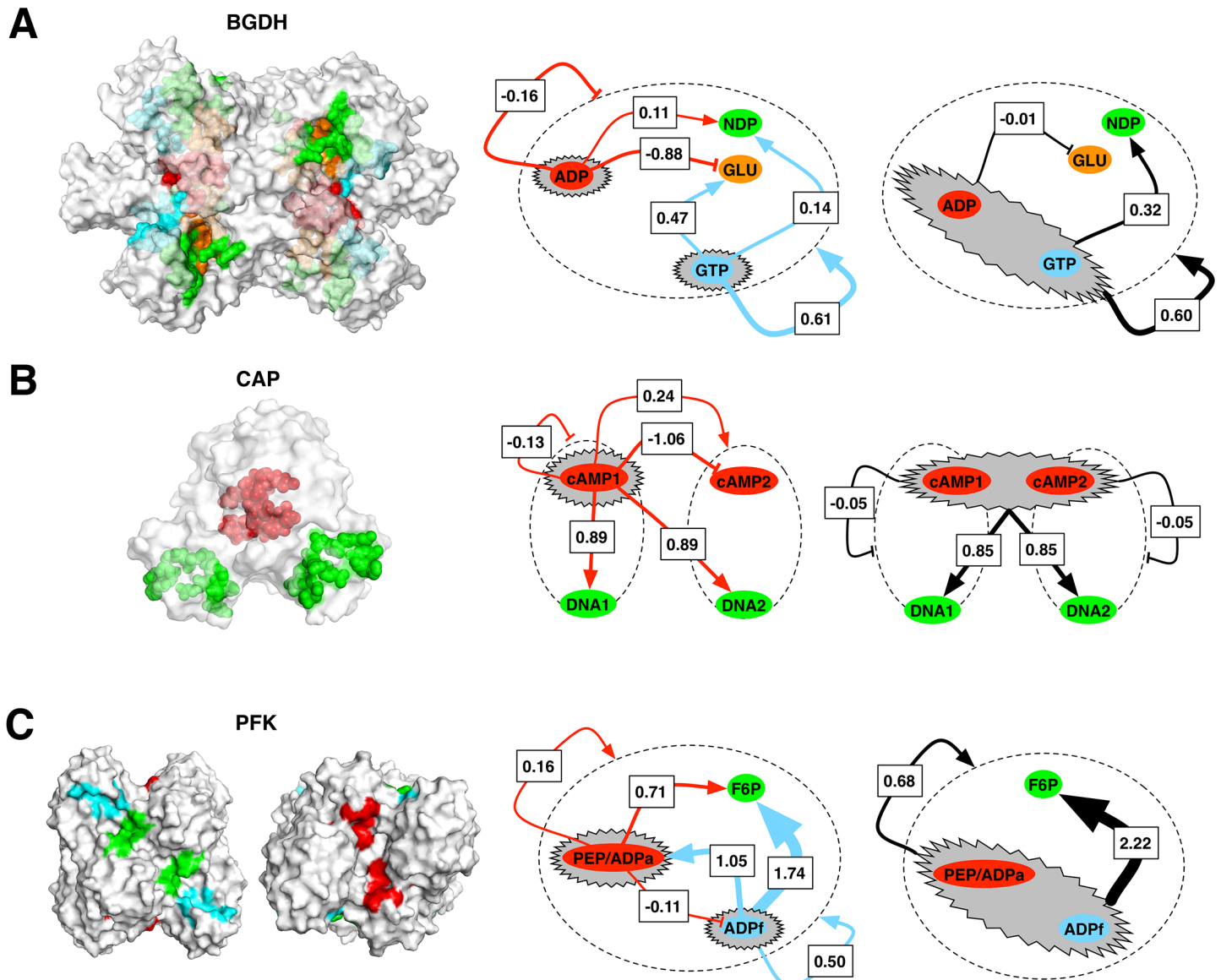


Fig 4. Schematic representation of the details of allosteric communication. (A) bovine glutamate dehydrogenase, BGDH; (B) catabolite activator protein, CAP; (C) phosphofructokinase, PFK. Structures show locations of the binding sites, which have node shapes in the diagrams on the right. Grey ovals show ligated allosteric sites in corresponding modes of regulation. Arrows illustrate the causal relations between the allosteric and functional sites. Numbers on the arrows provide the allosteric free energy (or work exerted, in kcal/mol) at the catalytic site F as a result of binding at the allosteric site, $\Delta g_F(P \rightarrow AP)$. (A) BGDH. Simultaneous binding of ADP and GTP is compared to independent ones. In both PFK and BGDH schemes the dashed line ovals represent the whole protein complex. (B) CAP. There are two possible scenarios of regulation, depending on whether one ($1xcAMP$) or two ($2xcAMP$) allosteric ligands are bound. The dashed line ovals represents the protein monomers. (C) PFK. The mode of regulation in which both $PEP/ADPa$ and ADP_f ligands are independently bound is compared to the case in which both effector and substrate are simultaneously bound, revealing the modulating role of ADP_f in the configurational work exerted at the $F6P$ site.

doi:10.1371/journal.pcbi.1004678.g004

GTP), which shows the prevailing effect of the inhibitor (see Table 1). Interestingly, the pronounced changes in the dynamics of the structure reflect the conformational changes caused by the binding of the inhibitor, whereas binding of the activator only stabilizes the apo form.

Catabolite Activator Protein (CAP). Catabolite Activator Protein (CAP), dimeric transcription factor, is a very well studied example of DNA-binding activity modulation by means of what is commonly understood as dynamically driven (or governed by conformational

entropy) allostery [30]. Allosteric regulation is performed via negatively cooperative cAMP-binding in two binding sites (cAMP)—one for each protein monomer. NMR studies have clearly shown that binding at only one of the two sites increases fluctuations in the structure without changing its conformation, whilst binding at both sites greatly reduces fluctuations in the whole complex [30]. We studied all three available CAP structures: the apo form (PDB ID: 2wc2), the structure with one allosteric site ligated (PDB ID: 1run), and the form with both allosteric sites ligated and DNA bound (PDB ID: 1o3q). The apo form was obtained as a mean conformer from the 20-structure bundle NMR data [31]. We applied our method in two steps by considering the allosteric free energies of states with one and with both cAMP sites ligated, and by comparing them with allosteric free energies in the apo form. Differences in configurational work are estimated for the allosteric sites and for the DNA-bound region (Table 1 and Fig 3B). First, we consider the model based on the apo structure (2wc2). If a single cAMP site is ligated, it stabilizes the whole bound monomer ($\Delta g_{B-mer}(0 \rightarrow 1 \times cAMP) = -0.13$ kcal/mol), along with a significant destabilization ($\Delta g_{F-mer}(0 \rightarrow 1 \times cAMP) = 0.24$ kcal/mol) of the free monomer, as well as destabilization ($\Delta g_{DNA}(0 \rightarrow 1 \times cAMP) = 0.88$ kcal/mol) of the DNA-binding regions. If the cAMP sites are occupied in both monomers the overall conformational work of the whole protein is further reduced $\Delta g_{2wc2}(0 \rightarrow 2 \times cAMP) = -0.30$ kcal/mol, reflecting the loss of structural flexibility in qualitative agreement with the experimental observations. The conformational work at the DNA-binding regions ($\Delta g_{DNA}(0 \rightarrow 2 \times cAMP) = 0.85$ kcal/mol) is similar to the value observed for the structure with the only one allosteric site ligated. The difference $\Delta \Delta g_{2wc2} = \Delta g_{2wc2}(0 \rightarrow 2 \times cAMP) - \Delta g_{2wc2}(0 \rightarrow 1 \times cAMP) = -0.54$ kcal/mol is consistent with the known negative cooperativity of CAP upon cAMP multiple binding. In the other structures with cAMP bound (1run) and both cAMP and DNA bound (1o3q) a qualitatively similar picture is observed. Analysis of these two forms shows that binding of both cAMPs relaxes the DNA binding site (an increase in flexibility) and facilitates DNA binding. If DNA is already bound, cAMPs binding provides further stabilization of the DNA binding site. In Fig 4B two schematic graphs give a pictorial summary of CAP's regulation mechanisms in the situations of single (left) and double (right) ligated cAMP binding sites.

3-deoxy-D-arabinoheptulosonate 7-phosphate Synthase (DAHPS). 3-deoxy-D-arabinoheptulosonate 7-phosphate synthase (DAHPS) is a key enzyme involved in the biosynthesis of aromatic amino acids such as phenylalanine, tyrosine, and tryptophan. The amino acid products work as allosteric effectors that provide a feedback regulation and inhibit the catalytic activity. We study two DAHPS forms: inactive (PDB ID: 1kfl) and active (PDB ID: 1gg1). The former is crystallized in presence of the allosteric inhibitor-product L-phenylalanine (allosteric site PHE) and substrate phosphoenolpyruvate (active site PGA); the latter—in the presence of the substrate only. The over-stabilization of the PHE site ($\Delta g_{PHE}(0 \rightarrow 4 \times PHE) = -2.33$ kcal/mol) in the active form induces an increase of configurational work at the PGA functional site ($\Delta g_{PGA}(0 \rightarrow 4 \times PHE) = 0.33$ kcal/mol), which is higher than the overall increase of the mean configurational work $\Delta g_{1gg1}(0 \rightarrow 4 \times PHE) = 0.25$ kcal/mol (see Fig 3C for the corresponding allosteric profiles). This suggests that the inhibition effect caused by PHE binding may facilitate the substrate release from the catalytic site. This observation is in a qualitative agreement with the results obtained for the inactive form (1kfl), where all the observed effects are less pronounced (S1 Table).

Dihydroxyacetone Kinase (DAK). The dihydroxyacetone kinase (DAK) is a protein kinase that catalyzes the phosphoryl transfer between phosphagen and ADP necessary for the energy exchange in cell metabolism. The enzymatic activity in DAK is homotropically regulated by its substrate that works as a negatively cooperative inhibitor [32]. There is an interaction between active sites of two monomers, which is archetypal for homo-dimeric enzymes with negative cooperativity [32]. In particular, ligand binding in one monomer affects the

active site of the other monomer. Two ligands, substrate L-arginine (site ARG) and product ADP, bind to the catalytic site of the protein. Bound to one of the monomers, both ligands can also be regarded as allosteric effectors for another part of the protein. We studied the active form (PDB ID: 3ju5) and the inactive one (PDB ID: 3ju6), which contains ARG and the ADP analog—ANP (Table 1 and Fig 3D). Occupation of sites ADP and ARG in one monomer of the active form rigidify it ($\Delta g_{B-mer}(0 \rightarrow 1 \times (ADP + ARG)) = -0.72$ kcal/mol), increasing the configurational work in the free monomer ($\Delta g_{F-mer}(0 \rightarrow 1 \times (ADP + ARG)) = 0.74$ kcal/mol). As a result, the free energy loss by the ligated allosteric site monomer is compensated by the work exerted at the active site monomer. When both active sites are ligated the overall configurational work of the whole protein is greatly reduced ($\Delta g_{3ju5}(0 \rightarrow 2 \times (ADP + ARG)) = -0.05$ kcal/mol). The difference, $\Delta \Delta g_{3ju5}(1 \times (ADP + ARG) \rightarrow 2 \times (ADP + ARG)) = \Delta g_{3ju5}(0 \rightarrow 2 \times (ADP + ARG)) - \Delta g_{3ju5}(0 \rightarrow 1 \times (ADP + ARG)) = -0.79$ kcal/mol, reflects a typical phenomenology of the negative cooperativity, and it is consistent with the inhibition mechanisms proposed elsewhere [32]. The results obtained for the inactive form (1ju6) are in qualitative agreement (SI Table) with those obtained for the active structure.

Glucose-6-phosphate Dehydrogenase (G6PD). Hexaoligomeric glucose-6-phosphate dehydrogenase (G6PD) catalyzes a regulatory step of N-acetylglucosamine catabolism, producing fructose 6-phosphate and ammonia via isomerisation and deamination of the glucosamine 6-phosphate [33]. Activity of G6PD is regulated by the binding of N-acetylglucosamine-6-phosphate, which acts as an allosteric activator. We studied three available forms of the protein: the inactive apo form (PDB ID: 1cd5), the ligated form (PDB ID: 1hor) with a competitive inhibitor (site AGP) bound to the active site, and the active form (1hot) with bound allosteric ligand (site 16G). When applied to the apo form of the protein our model shows that binding at the allosteric sites (16G) does not result in their stabilization ($\Delta g_{16G}(0 \rightarrow 6 \times 16G) = -0.04$ kcal/mol). At the same time, we detected an overall destabilization of the whole complex ($\Delta g_{1cd5}(0 \rightarrow 6 \times 16G) = 0.31$ kcal/mol) and the functional sites ($\Delta g_{AGP}(0 \rightarrow 6 \times 16G) = 0.37$ kcal/mol). The destabilization effect $\Delta g_i(0 \rightarrow 6 \times 16G)$ is illustrated in Fig 5A. The large errors per monomer are also evident, reflecting the asymmetry in the X-ray structure. The forms with competitive inhibitor (1hor) and with activator (1hot) show different behavior with respect to the apo form, but they yield qualitatively similar behavior between themselves. This similarity reflects the fact that both forms 1hot and 1hor are likely in the same conformational state, which is different from that of the apo form. It has been shown that low frequency modes explain the allosteric transition for G6PD from the apo form to the active one [21], and the destabilization of the protein is apparently a result of large quaternary structure reorganization caused by the binding of the allosteric ligand [34].

NAD-dependent Malic Enzyme (NADME). Malate dehydrogenase (decarboxylating) enzyme (NADME) is a homo-tetrameric oxidoreductase that catalyzes decarboxylation of L-malate to pyruvate and the concomitant reduction of the cofactor NAD⁺ or NADP⁺. ATP works as a competitive inhibitor of the NAD active site, and it can also bind at the tetrameric interface (site ATP). Fumarate binds at the dimeric interface (site FUM), working as an allosteric activator [35]. We considered two forms of the NADME protein: the open active form (1efk) ligated with NADP⁺ substrate at the NAD binding site, and the closed inactive form (1gz3) in complex with the competitive inhibitor ATP and activator FUM. Binding at the FUM site in the model based on the open form (1efk) has a mild stabilizing effect on the whole complex ($\Delta g_{1efk}(4 \times FUM) = -0.14$ kcal/mol), as well as on the active site NAD (Fig 5B). Interestingly, restraining the ATP site induces an overall stabilization of the protein ($\Delta g_{1efk}(4 \times ATP) = -0.25$ kcal/mol), but a mild increase in configurational work at the NAD active site ($\Delta g_{NAD}(4 \times ATP) = 0.11$ kcal/mol). One can presume, therefore, that the allosteric activation by the fumarate is coupled with a stabilization of the active site, thus providing a proper binding of a

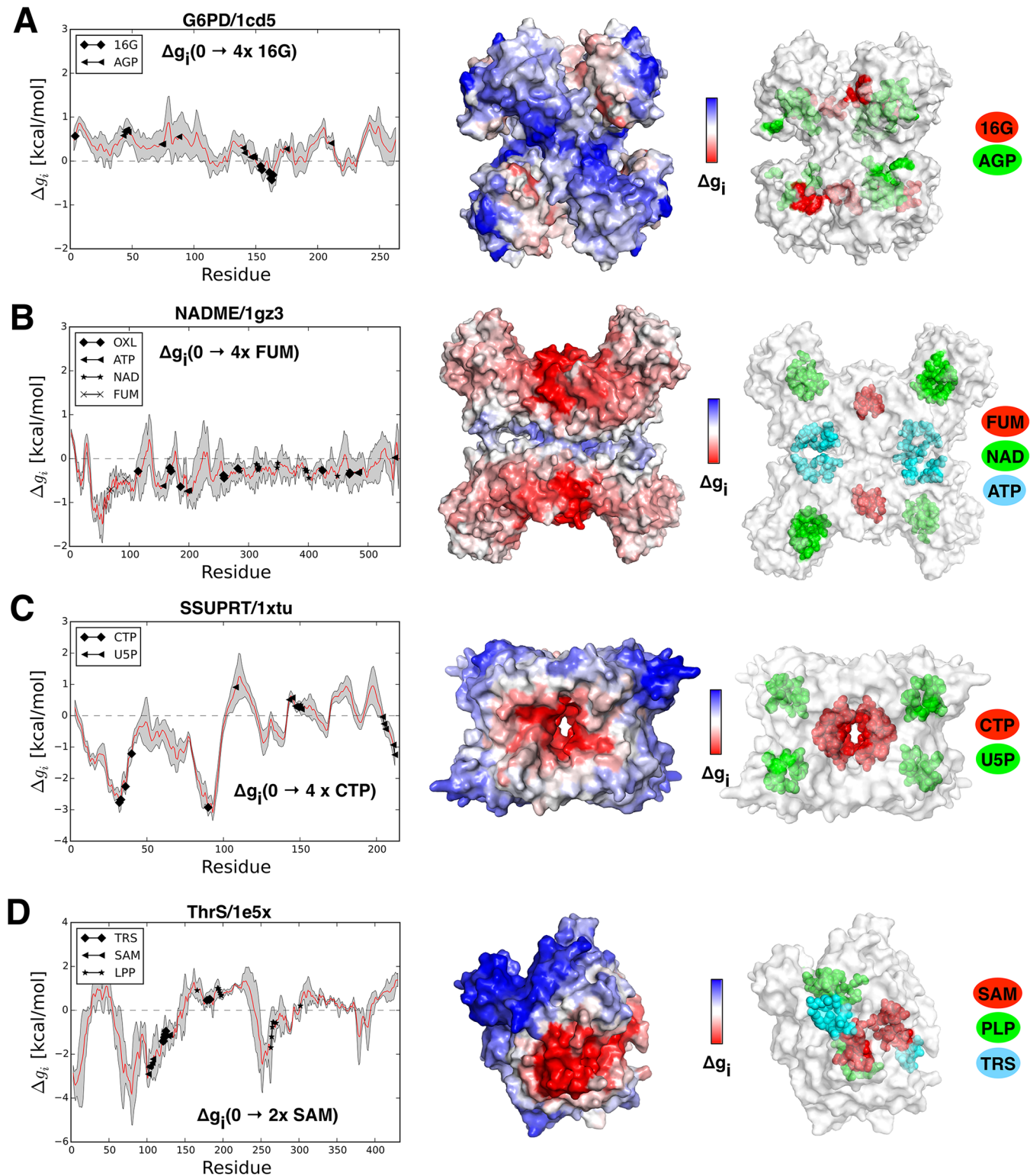


Fig 5. Free energy profiles and structural representatives of (A) G6PD, (B) NADME, (C) SSUPTR, and (D) ThrS. (A) 6-mer glucose-6-phosphate dehydrogenase G6PD free energy profile $\Delta g_i(0 \rightarrow 6x16G)$ with marked locations of the 16G and AGP binding sites. (B) Malate dehydrogenase (decarboxylating) enzyme NADME free energy profile $\Delta g_i(0 \rightarrow 4xFUM)$ with indicated positions of ATP, NAD and FUM binding sites. (C) Sulfolobus solfataricus uracil phosphoribosyltransferase SSUPTR free energy profile $\Delta g_i(0 \rightarrow 4xCTP)$ with marked CTP and U5P binding sites. (D) Threonine synthase ThrS free energy profile $\Delta g_i(0 \rightarrow 2xSAM)$ with shown locations of the TRS, SAM and LPP binding sites.

doi:10.1371/journal.pcbi.1004678.g005

substrate. On the other hand, the work increasing at the active site caused by the ATP binding might partially induce inhibition as it would facilitate the substrate release. Similar qualitative results were obtained in the model based on the inactive form (see 1gz3 in [S1 Table](#)). The availability of the NADME's apo form would be important for finalizing the scenario of NAMDE's regulation.

Sulfolobus Solfataricus Uracil Phosphoribosyltransferase (SSUPTR). The *Sulfolobus solfataricus* uracil phosphoribosyltransferase (SSUPTR) is a homo-tetramer that catalyzes the transformation of 5-phosphate-alpha-1-diphosphate and uracil into uridine 5'-monophosphate (site U5P) and diphosphate. The complex is allosterically inhibited by CTP (site CTP). It has been suggested that CTP binding and consequent structural changes of the complex stabilize the flexible loop in a closed conformation [36]. Four CTP binding sites are located at the central interface between two dimers (see [Fig 5C](#)), while four functional sites (U5P) are placed at the outer corners of the complex. Two structures were considered in our model: the form with bound substrate U5P and effector CTP (PDB ID: 1xtu), and with only substrate bound to the protein (PDB ID: 1xtt). Restraining the allosteric sites in the 1xtu form induces significant stabilization of the whole complex ($\Delta g_{1xtu}(0 \rightarrow 4 \times CTP) = -0.53$ kcal/mol), as well as an increase in the configurational work at the functional sites. An increase in the configurational work at the active sites, along with unchanged overall protein stability in the other form (1xtt, see [S1 Table](#)) suggest that allosteric inhibition acts by decreasing plasticity of the active site.

Threonine Synthase (ThrS). Threonine synthase (ThrS) is a homo-dimeric enzyme that is in charge of the final steps of the threonine synthesis (TRS), using the pyridoxal-L-phosphate PLP substrate. It has been shown that allosteric activation caused by the S-adenosylmethionine (SAM) binding triggers a large reorganization of the substrate site in one monomer that induces a conformation change of PLP to an active conformation [37]. Restraining SAM binding sites in a model based on the protein's apo form (1e5x) does not affect an overall dimer stability, while it increases work at both PLP sites ($\Delta g_{PLP}(0 \rightarrow 1 \times SAM) = 0.35$ kcal/mol). Restraining both SAM sites causes significant stabilization of the dimer ($\Delta g_{1e5x}(0 \rightarrow 2 \times SAM) = -0.24$ kcal/mol), and, at the same time, it increases configurational work at the PLP sites ($\Delta g_{PLP}(0 \rightarrow 2 \times SAM) = 0.52$ kcal/mol, [Fig 5D](#)). A mild positive cooperativity at the PLP site ($\Delta \Delta g_{PLP}(1 \times SAM \rightarrow 2 \times SAM) = 0.17$ kcal/mol) was also observed.

D-3-phosphoglycerate Dehydrogenase (PGDH). The D-3-phosphoglycerate dehydrogenase (PGDH) is a ring-shaped tetrameric enzyme that catalyzes formation of 3-phosphohydroxypyruvate from 3-phospho-D-glycerate with NAD as a cofactor. PGDH is allosterically regulated by serine as a cooperative inhibitor. Binding of the pair of serines to dimer interfaces inhibits PGDH, affecting mainly the rate of catalytic reaction [38]. We analyzed two available X-ray forms: the first structure includes the allosteric inhibitor SER and cofactor NAD (PDB ID: 1psd); the second structure (PDB ID: 1yba) contains substrates AKG, NAD bound at the catalytic site. Upon binding to the allosteric sites in the inhibited structure, the tetramer is weakly destabilized ($\Delta g_{1psd}(0 \rightarrow 8 \times SER) = 0.03$ kcal/mol), while the substrate sites, in particular NAD ($\Delta g_{NAD}(0 \rightarrow 8 \times SER) = 0.38$ kcal/mol), yield some increase of the configurational work. As shown in [Fig 6](#), increase of the work exerted at four central domains of the tetramer is compensated by the over-stabilization of the peripheral domains where the allosteric binding takes place (see [S1 Table](#)). Qualitatively similar results for overall stability of the protein and for the work exerted at the functional sites are obtained by restraining the allosteric sites in the active form (1yba). However, restraining of the allosteric sites in the active form of the protein does not result in stabilization of the domains that contain them ([S1 Table](#)). On the contrary, the work exerted in the corresponding domains of the inhibited form (1psd) is positive (bottom structure in [Fig 6](#)). Additionally, the work exerted at the functional sites of the active form is lower than the one obtained for the inactive structure. Comparison of the results obtained for

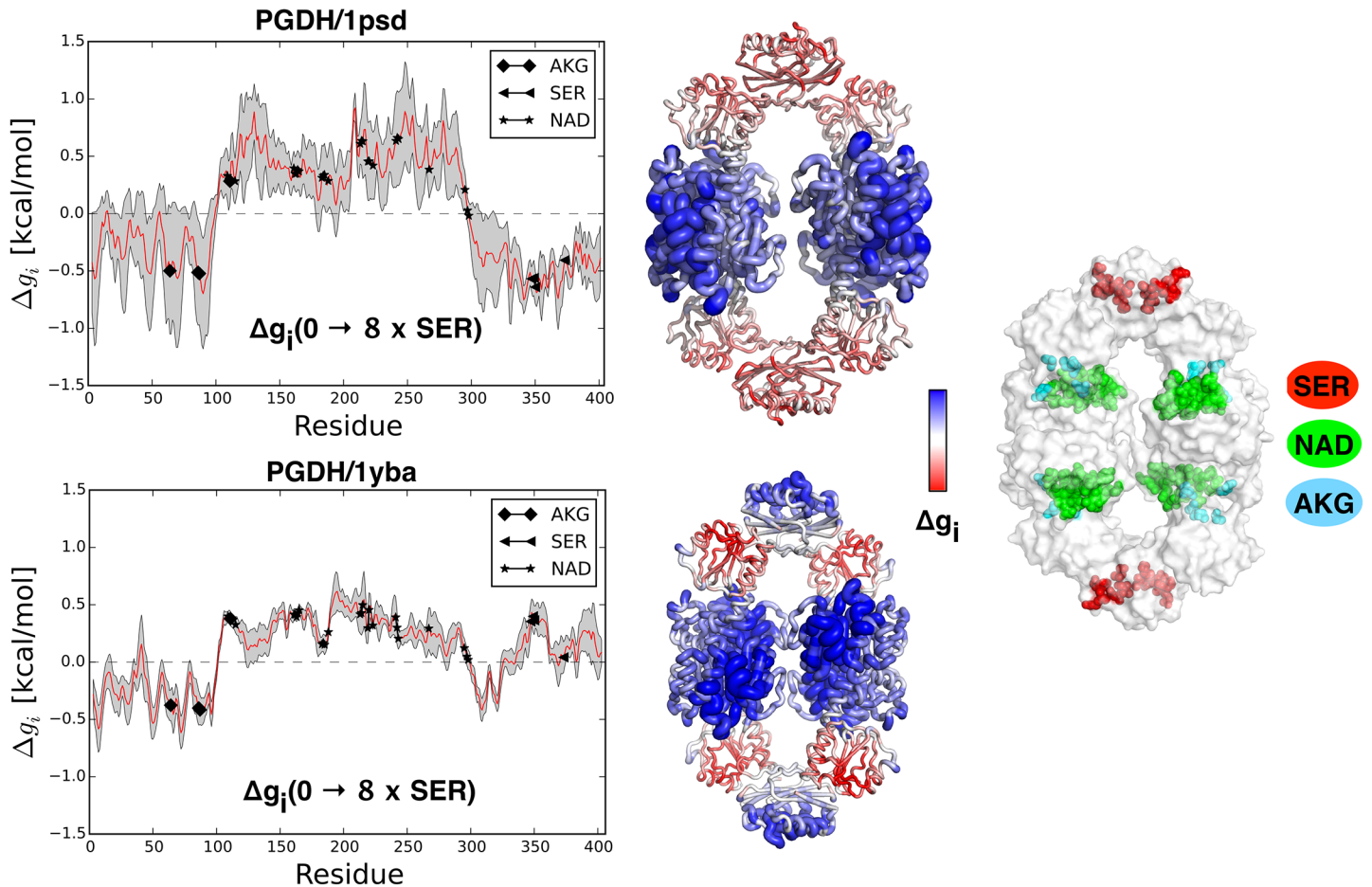


Fig 6. Free energy profiles and structural representations of the D-3-phosphoglycerate dehydrogenase PGDH. The profiles $\Delta g_i(0 \rightarrow 8 \times SER)$ are based on calculations performed on two protein forms, 1psd and 1yba, respectively. Protein structures are shown using a colored tube-like representation with both colors, and radius of the tube scaled according to the configurational work exerted per corresponding residue (the radius ρ_i per residue i is scaled according to the function $\rho_i = (\Delta g_i - \min_j \Delta g_j) / (\max_j \Delta g_j - \min_j \Delta g_j)$). Structures in the middle represent two forms of the protein analyzed in this work: inactive (1psd, top) and active (1yba, bottom). The structure in the right panel illustrates locations of the protein's ligands: SER (red), NAD (green), and AKG (blue).

doi:10.1371/journal.pcbi.1004678.g006

inactive and active forms shows that the PGDH's active state is supported by the global dynamics of the whole structure rather than by increase of the local dynamics in the catalytic monomers.

Phosphofructokinase (PFK). Phosphofructokinase (PFK) is an important regulatory enzyme of glycolysis and is one of the best studied allosteric proteins [39]. PFK's substrate is fructose-6-phosphate (site F6P) and its four subunits are controlled by the variety of activators and inhibitors. The protein is allosterically activated by ADP and inhibited by phosphoenolpyruvate PEP whose binding sites essentially overlap (Fig 7, common site PEP/ADPa). Another ADP binding site (site ADPf) is located in the vicinity of the functional one (F6P) at the dimer-dimer interface. We analyzed three PFK structures: the apo form (PDB ID: 3pfk), the structure with bound ADPa and ADPf (PDB ID: 4pfk), and the structure with bound inhibitor PEP (PDB ID: 6pfk). Both the apo and the ADP-bound forms are the active ones, and their high structural similarity indicates that binding of the activator stabilizes the apo conformation. Transition from the active to the inactive form is characterized by the rotation of tetramer subunits [40], which can be fairly well described by the low frequency normal modes [21]. Upon

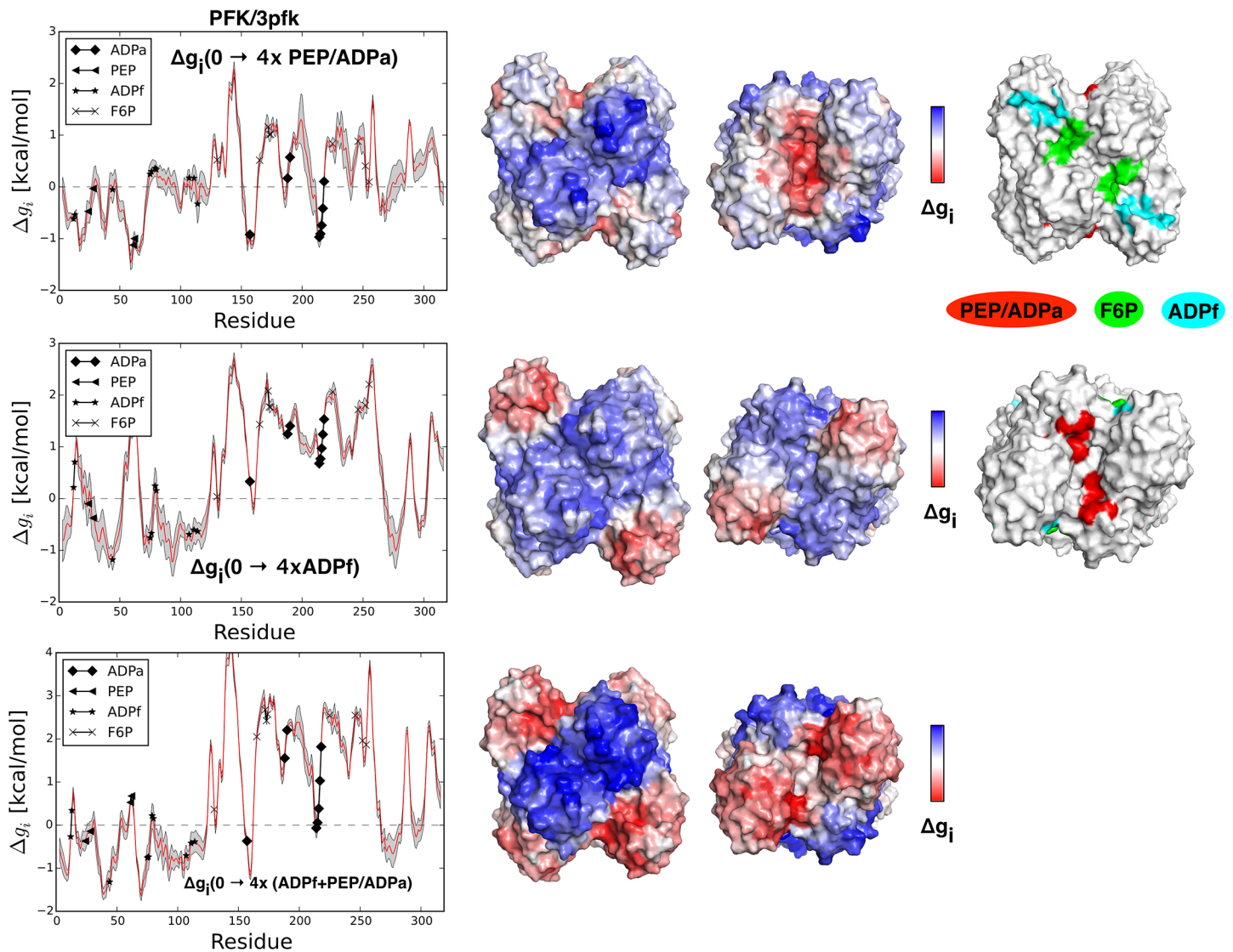


Fig 7. Free energy profiles and representative structures of Phosphofructokinase PFK upon restraining its various sites. All the results shown here are obtained from the analysis of the protein apo form (PDB ID: 3pfk). Three situations are shown: the inhibitor/activator PEP/ADPa is bound $\Delta g_i(0 \rightarrow 4 \times \text{PEP/ADPa})$; the substrate (ADPf) is bound $\Delta g_i(0 \rightarrow 4 \times \text{ADPf})$; and both (PEP/ADPa and ADPf) ligands are bound $\Delta g_i(0 \rightarrow 4 \times (\text{PEP/ADPa} + \text{ADPf}))$. The free energy profiles are shown with the colored protein surfaces according to the configurational work exerted per residue (middle column), and with representative structures where the locations of the bound ligands are marked by different colors (right column).

doi:10.1371/journal.pcbi.1004678.g007

binding at the PEP/ADPa site the whole protein is only weakly destabilized ($\Delta g_{3pfk}(0 \rightarrow 4 \times \text{PEP/ADPa}) = 0.16$ kcal/mol), while the F6P sites yield an increase of the mean configurational work ($\Delta g_{F6P}(0 \rightarrow 4 \times \text{PEP/ADPa}) = 0.71$ kcal/mol) along with mild stabilization of the functional ADPf binding sites ($\Delta g_{ADPf}(0 \rightarrow 4 \times \text{PEP/ADPa}) = -0.11$ kcal/mol). Interestingly, having the only substrate ADPf binding sites ligated not only causes an overall destabilization of the protein ($\Delta g_{3pfk}(0 \rightarrow 4 \times \text{ADPf}) = 0.50$ kcal/mol), but it also results in a significant increase of configurational work at the F6P site ($\Delta g_{F6P}(0 \rightarrow 4 \times \text{ADPf}) = 1.74$ kcal/mol). Binding at both PEP/ADPa and ADPf sites also leads to a major increase of configurational work at the F6P site ($\Delta g_{F6P}(0 \rightarrow 4 \times (\text{PEP/ADPa} + \text{ADPf})) = 2.22$ kcal/mol). Fig 7 contains the allosteric free energy profiles in cases of binding at the PEP/ADPa, ADPf, and PEP/ADPa+ADPf sites,

respectively. The effect of occupying PEP/ADPa and functional ADPf sites unravels the double-fold role of the PEP/ADPa site. First, it can provide an inhibition upon binding the antagonist via decrease of the configurational work at the functional ADPf site ($\Delta g_{ADPf}(0 \rightarrow 4 \times PEP/ADPa) = -0.11$ kcal/mol). At the same time, it can work as a modulator increasing configurational work in the functional sites upon binding the ADP to both PEP/ADPa and functional ADPf sites ($\Delta \Delta g_{F6P}(4 \times PEP/ADPa \rightarrow 4 \times (PEP/ADPa + ADPf)) = 1.51$ kcal/mol). Modulating activity of the PEP/ADPa site agrees with an early detected positive cooperativity in binding of two substrates (F6P and ADP) in presence of the PEP [41]. PFK's regulation is schematically illustrated in Fig 4C. Both cases, with the binding sites PEP/ADPa (inhibition) and ADPf (activation) individually ligated, and with effectors bound to both sites (positive cooperativity), are shown in figure. The other forms of the proteins show qualitatively similar results (S1 Table).

Protein Kinase A (PKA) and protein tyrosine phosphatase 1B (PTP1B). These are two small monomeric proteins for which we analyzed the apo forms and structures with bound inhibitors (see the allosteric free energy profiles in S3 Fig). In both cases we found that binding of the inhibitor causes relaxation in the catalytic site. One can hypothesize, therefore, that inhibition in these proteins happens via the release of the substrate from the destabilized catalytic site.

Brief overview of the analyzed proteins

There are different scenarios of allosteric regulation in 14 proteins analyzed in this work. Both activation and inhibition modes are observed for BGDH and PFK; only activation modes for ATCase, CAP, DAK, G6PD, NADME, and ThrS; and only inhibition modes for AnthS, DAHPS, SSUPTR, PGDH, PKA, and PTP1B. In CAP and DAK activation modes are coupled with negative cooperativity; in ThrS—with positive cooperativity; and in PFK there are both modes of regulation and positive cooperativity. Since there is no experimental data on free energy changes associated with allosteric signaling at a single residue level, the allosteric free energy obtained in our model can be only qualitatively compared with experimental observations. Computationally obtained free energy changes allow one to conclude on the modulation of dynamics in the catalytic sites under regulation. Positive and negative values of Δg reflect increased/decreased dynamics in the corresponding catalytic sites. We found that changes in the site dynamics are always opposite in proteins with both modes of regulation (BGDH and PFK). Our data also indicated positive cooperativity in PFK. In all cases of activation except BGDH, the activation mode is coupled with increased dynamics at the catalytic site. A combination of the activation with negative cooperativity in CAP and DAK and with positive cooperativity in ThrS was detected. An increase in catalytic site dynamics in BGDH, DAHPS, SSUPRT, PKA, and PTP1B provides an inhibition mode. In AnthS and PFK inhibition is based on stabilization of the catalytic sites (decreased dynamics). We found that either increase or decrease in dynamics of the catalytic site can be associated with activation or inhibition. For example, increase of dynamics can result in activation favoring the binding of substrate, or contribute to inhibition supporting release of the substrate. At the same time, a decrease in dynamics can stabilize the active state of the catalytic site, or can prevent binding of the substrate (inhibition). As a result, any change in a catalytic site's dynamics is indicative of the regulation, albeit the mode of regulation should be further explored. Another important aspect that should be studied in the future is the issue of agonists/antagonists that bind to the same allosteric site. For example, opposite modes of the regulation were detected, depending on the bound allosteric ligand in ATCase and PFK. Current and future work in this direction includes development of the atomic resolution model that can determine agonistic/antagonistic nature of individual residues and their combinations within the same binding site. Merged with state-of-the-art

experimental techniques, e.g. disulfide trapping [42], this approach could provide a foundation for design of allosteric effectors with required agonistic/antagonistic activity [22].

Discussion

The rise in interest in allosteric processes and, as a consequence, turning the focus to the design of allosteric drugs [43, 44, 22] calls for the development of theoretical models that allow one to quantify protein energetics affected/modulated by the binding of allosteric effectors. In this context, an ideal model should be the result of the synthesis between the thermodynamic and structural views of allostery [14]. Such a comprehensive model should provide an allosteric free energy at the single residue level, a key quantity for evaluating the free energy associated with the processes of allosteric regulation. The model should also provide quantification of the free energy difference between sequential steps of binding of identical and/or different types of effector molecules, as well as to distinguish effects of agonists and antagonists on the protein's functional activity.

The structure-based statistical mechanical model of allostery introduced here consists of three key components: (i) it models the effector binding via local perturbation/stabilization of the protein harmonic model; (ii) per residue allosteric potential estimates structural changes due to the effect of protein dynamics on the residue environment; (iii) statistical mechanical ensemble treatment provides a quantification of the allosteric free energy obtained from the per-residue partition functions in the ligand free and bound states. The model is applied to a large set of allosteric proteins previously analyzed in the context of allosteric site prediction [21]. The list of proteins is heterogeneous in terms of their functions, degree of oligomerization, sizes, and allosteric phenomenologies. We found a variety of modes of allosteric regulation, e.g. binding in allosteric sites causes different levels of tension or relaxation at functional sites. These observations show that allosteric communication between protein sites is a direct consequence of the configurational work redistribution according to structural peculiarities of the allosterically regulated proteins. The calculated changes in stability of catalytic sites, i.e. configurational work exerted at these sites, are found to be strongly dependent on the topology, oligomerization state, and other structural/functional traits. A paradigm case is, for example, D-3-phosphoglycerate dehydrogenase (PGDH) which shows how the redistribution of configurational work strongly resembles the domain structures of the protein and its functional mechanisms. Stability modulation, in turn, appeared to be more dependent on the type of the allosteric regulation: in case of Phosphofructokinase (PFK) it is a result of the interplay between binding of different ligands. In particular, PEP (allosteric inhibitor) and ADP (substrate and/or activator depending on the location) are both responsible for the stability regulation of the F6P binding site and for the positive cooperativity of the F6P binding. In the case of catabolite activator protein (CAP) the sequential scenario of binding to two allosteric cAMP sites delineates the negative cooperativity scenario in allosteric regulation of the DNA binding. Overall, a clear picture with good qualitative agreement between our observations and experimentally described modes of regulation was obtained. Looking forward to the next steps in this work, it is important to briefly discuss the major advantages of the model as well as its drawbacks. The first obvious limitation stems from its coarse-grained nature (C_α harmonic model), since the lack of atomic detail affects the quality of the allosteric effects' estimation. An indirect modeling of the binding effect can be considered as both a limitation and an advantage of the model. First, it does not take into account the structure of the ligands or the actual set of interactions in the binding sites. On the other hand, the indirect way of mimicking the ligand binding is a generic framework that can be applied to different types of proteins without any preliminary knowledge of the allosteric sites. While the model could benefit from empirical parametrization

on the basis of experimental observations, it provides an energy estimate of the allosteric signaling.

To conclude, this approach should be regarded as a *ground state model*, as the reference structure used in a harmonic model of a protein is the same in the ligand free and ligand bound systems. The free energy difference between the ligand free and ligand bound states is chiefly determined by the contribution of the configurational entropy. Allosteric communication detected by the model is indicative of the signaling between the regulatory and functional sites in both directions. Therefore, perturbation of the functional site can point to the potential locations of relevant the allosteric ones. As a result, the perturbation of the functional sites would allow one to determine the allosteric sites that modulate activity of corresponding catalytic sites, to identify specifics of the allosteric communication at the atomic level resolution, and to discriminate between the agonism and antagonism of allosteric effectors.

Materials and Methods

Harmonic model of free and bound states

A proper choice for the description of protein dynamics is a key component in modeling of the allosteric processes. For the sake of computational simplicity we adopt here a description of the protein configurational states based on a C_α harmonic model. Despite this radical simplification, harmonic models are surprisingly powerful in describing the large-amplitude protein motions based on the consideration of the low-frequency normal modes [45–47]. An effective harmonic potential introduced in [48] is used in this work within the implementation given in MMTK simulation package [49]. The pairwise effective energy function between all C_α atoms is

$$E^{(0)}(\mathbf{r} - \mathbf{r}^0) = \sum_{\text{pairs } i,j} k_{ij} (d_{ij} - d_{ij}^0)^2 \quad (7)$$

where \mathbf{r} is the 3N-dimensional vector of coordinates of the C_α atoms, \mathbf{r}^0 is a vector of C_α positions of the reference structure, d_{ij} is the Euclidian distance between the C_α atoms i and j , d_{ij}^0 is the corresponding distance in the reference structure, and k_{ij} is the distance-dependent force constant decaying as $(1/d_{ij}^0)^6$ with a global 25Å distance cutoff (see default settings in [48]). In order to estimate effects of the binding to different sites (or their combinations) on the protein energetics, we consider two harmonic systems: the free system \mathcal{P} with energy given by eq (7) and the system $A^{(n)}\mathcal{P}$ with n ligands bound to sites s_1, \dots, s_n . We define the latter by adding a binding site's stiffening harmonic term to the energy function of the free system, that is

$$E^{(n)}(\mathbf{r} - \mathbf{r}^0) = E^{(0)}(\mathbf{r} - \mathbf{r}^0) + \alpha \sum_k V_{s_k}(\mathbf{r} - \mathbf{r}^0) \quad (8)$$

with the additional terms $V_s(\mathbf{r} - \mathbf{r}^0) = 1/2 \sum_{\text{pairs } i,j \in s} k_{ij} (d_{ij} - d_{ij}^0)^2$ corresponding to the restrained binding site s and $\alpha > 0$ is the stiffening factor. The effect of ligand binding is indirectly modeled as a stabilization effect on the atoms that belong to the binding site of interest. Thus, the factor α gives the extent to which these indirect interactions affect the atoms of the binding site. To ensure a pronounced allosteric effect, a stiffening factor $\alpha = 100$ was chosen based on the analysis of the allosteric free energy profiles obtained for the list of proteins (see S1 Fig). From the Hessian matrices $\mathbf{K}^{(0)} = \partial^2 E^{(0)}/\partial \mathbf{r}_i \partial \mathbf{r}_j$ and $\mathbf{K}^{(n)} = \partial^2 E^{(n)}/\partial \mathbf{r}_i \partial \mathbf{r}_j$ the orthonormal normal modes $\mathbf{e}_\mu^{(0)}$, $\mathbf{e}_\mu^{(n)}$ were calculated for the ligand free and ligand bound systems \mathcal{P} , $A^{(n)}\mathcal{P}$, respectively.

The additional harmonic terms in eq 8 induce a reorganization of the normal modes of the free system, consistent with the observation that ligand binding causes an overall change in the

protein dynamics by shifting from low to high values in the normal mode frequency spectrum [12, 50, 51]. Formulated in energy terms, restraint of the degrees of freedom in a binding site essentially mimics an additional energy that is supplied to the system in the case of the actual ligand binding. This energy supply can be “communicated” to other functionally relevant protein regions, and it can be expressed in terms of the gain/loss of the work exerted in corresponding parts of the protein. Similar *perturbation-based* harmonic models have been employed in previous works [52–55] with a special emphasis on the role of dynamics of individual residues in relation to allosteric binding and mutations. Fig 1A contains a pictorial illustration of the harmonic protein model used in this work for defining the ligand free and ligand bound protein systems. In many cases, there can be several effector and substrate ligands that can bind to a protein independently or in different combinations. In order to give a generalized description of the above cases, we switch here from the binary terminology of the apo (free) and holo (bound) forms of the protein to the notations free and ligated forms of the protein (or binding site).

Allosteric potential. In the previous section the sets of normal modes were obtained for the ligand free and bound protein states. Here, the normal modes are used as a basis for defining an allosteric potential in the context of a statistical mechanical framework. We are interested in evaluating the effect of the ligand binding on an arbitrary chosen residue i of the protein. In order to do that, we introduce an allosteric potential of residue i determined by the normal mode μ as:

$$U_{\mu,i} = \frac{1}{2} \sum_{j:d_{ij}^0 < d_c} c(d_{ij}^{(\mu)} - d_{ij}^0)^2 \tag{9}$$

with $c = 1 \text{ kcal/mol/\AA}^2$, $d_c = 11\text{\AA}$ a distance cutoff, d_{ij}^0 are C_α distances in the reference structure R , and $d_{ij}^{(\mu)}$ are C_α distances as a result of the normal mode μ . Thus, the potential evaluates the total elastic work that is exerted on the residue i as a result of the deforming action of a normal mode μ on its neighboring residues j . The idea behind the allosteric potential per residue i and mode μ is to compare the different deforming effects on a residue i induced by the normal mode of the free system ($e_\mu^{(0)}$) and of the system with n bound ligands ($e_\mu^{(n)}$), respectively. Fig 1B illustrates how the neighbors of an residue i (residues j, k, l) undergo different displacements upon a normal mode μ , changing the inter-residue distances $d_{ij}^{(0,\mu)}$ and $d_{ij}^{(n,\mu)}$ in the free and ligated sites.

The displacement of the residue i due to mode e_μ is $\mathbf{r}_i^{(\mu)} - \mathbf{r}_i^0 = \sigma_\mu \mathbf{e}_{\mu,i}$ where \mathbf{r}_i^0 are the coordinates of residue i and σ_μ is a gaussian coefficient. Thus, the allosteric potential in eq (9) can be rewritten as (to avoid redundancy we focus on the n ligated binding sites only)

$$U_{\mu,i}^{(n)}(\sigma_\mu) = \frac{1}{2} \varepsilon_{\mu,i}^{(n)} \sigma_\mu^2 \tag{10}$$

in terms of the coefficients σ_μ and residue intensity parameters

$$\varepsilon_{\mu,i}^{(n)} = \sum_{j:d_{ij}^0 < d_c} c |\mathbf{e}_{\mu,i}^{(n)} - \mathbf{e}_{\mu,j}^{(n)}|^2 \tag{11}$$

Summing over the modes of interest, the total allosteric potential of the residue i is

$$U_i^{(n)}(\sigma) = \frac{1}{2} \sum_{\mu} \varepsilon_{\mu,i}^{(n)} \sigma_\mu^2 \tag{12}$$

where $\sigma = (\sigma_1, \sigma_2, \dots)$ is a set of gaussian coefficients that can be regarded as a *configurational state* of a residue i .

Free energy of ligand binding and allosteric communication. In a statistical thermodynamic fashion the configurational free energy of a harmonic model is simply given by the expression $G = 1/2k_B T \sum_{\mu} \ln \lambda_{\mu} + \text{const}$, with λ_{μ} the eigenvalues of the Hessian matrix \mathbf{K} of the model ($\lambda_{\mu} = \omega_{\mu}^2$ with ω_{μ} the frequency associated to the mode μ). Thus, the free energy difference between the state with n bound ligands $A^{(n)}\mathcal{P}$ and the ligand free state \mathcal{P} is:

$$\Delta G(\mathcal{P} \rightarrow A^{(n)}\mathcal{P}) = \frac{1}{2}k_B T \sum_{\mu} \ln \frac{\lambda_{\mu}^{(n)}}{\lambda_{\mu}^{(0)}} \quad (13)$$

and the modulating free energy

$$\Delta \Delta G(\mathcal{P} \rightarrow A^{(m)}\mathcal{P} \rightarrow A^{(n)}\mathcal{P}) = \frac{1}{2}k_B T \sum_{\mu} \ln \frac{\lambda_{\mu}^{(0)} \lambda_{\mu}^{(n)}}{(\lambda_{\mu}^{(m)})^2} \quad (14)$$

where an intermediate case with m ligands bound $A^{(m)}\mathcal{P}$ state is defined by the $m < n$. The relations in Eqs (13) and (14) are global parameters, which quantify the configurational work (allosteric free energy) gain/loss of the structure's conformational ensemble that is exerted by the protein as a consequence of having multiple binding sites occupied [51].

Contrary to the above generic approach, our goal here is to quantify the allosteric effects at the single residue level, estimating the allosteric free energy per residue. To this end, the introduced allosteric potential eq (12) is now used for obtaining the canonical partition functions $z_i^{(0)}, z_i^{(n)}$ of residue i in the ligand free and n -ligated states of the protein, respectively. The set of coefficients $\sigma = (\sigma_1, \sigma_2, \dots)$ in eq (10) identifies an atomic configuration that corresponds to a particular linear combination of the low frequency normal modes. The coefficients σ_{μ} are gaussian variables with zero mean and variance $1/\epsilon_{\mu, i}$ (eq 11). For example, the per residue partition function for the state with n bound ligands, $z_i^{(n)}$, that corresponds to the ensemble of all possible configurations is

$$\begin{aligned} z_i^{(n)} &= \int d\sigma e^{-U_i^{(n)}(\sigma)/k_B T} \\ &= \prod_{\mu} \int d\sigma_{\mu} e^{-\sigma_{\mu}^2 \epsilon_{\mu, i}^{(n)}/2k_B T} \\ &= \prod_{\mu} \left(\pi \frac{2k_B T}{\epsilon_{\mu, i}^{(n)}} \right)^{1/2} \end{aligned} \quad (15)$$

The free energy per residue is $g_i = -k_B T \ln z_i + \text{const}$ so that the allosteric free energy difference per residue between the n -ligated state $A^{(n)}\mathcal{P}$ and the free state \mathcal{P} is

$$\Delta g_i(\mathcal{P} \rightarrow A^{(n)}\mathcal{P}) = \frac{1}{2}k_B T \sum_{\mu} \ln \frac{\epsilon_{\mu, i}^{(n)}}{\epsilon_{\mu, i}^{(0)}} \quad (16)$$

and the corresponding modulating free energy is

$$\Delta \Delta g_i(\mathcal{P} \rightarrow A^{(m)}\mathcal{P} \rightarrow A^{(n)}\mathcal{P}) = \frac{1}{2}k_B T \sum_{\mu} \ln \frac{\epsilon_{\mu, i}^{(n)} \epsilon_{\mu, i}^{(0)}}{(\epsilon_{\mu, i}^{(m)})^2} \quad (17)$$

with the intermediate m -ligated $A^{(m)}\mathcal{P}$ state $m < n$. The free energies in eqs (16) and (17)

quantify the maximal configurational work $\Delta g_i(\mathcal{P} \rightarrow A^{(n)}\mathcal{P})$ that is exerted on a residue i as a consequence of the change in the ensemble of configurations assumed by the residue's neighbors as a result of ligand binding. It is clear that the free energy in eq (16) represents the change in stability of a residue i caused by the allosteric signaling. Similarly, the expression in eq (17) evaluates the modulating effect on the stability (configurational work) of a residue i caused by changing the number of ligated allosteric sites. In Fig 1C the values of the free energies per residue, are represented via the thickness of the protein backbone (a tube-like representation is used).

The obtained free energies in the eqs (16) and (17) clearly depend on the number of low frequency normal modes. The dependence of the free energy profiles per residue on the number of normal modes (we used 5, 10, 25, and 50 modes) was analyzed for a subset of proteins (CAP, BGDH, and PFK). The results of these calculations are shown in the S2 Fig. There is no strong qualitative difference between the free energy profiles obtained with 10 and 50 modes. We concluded, therefore, that ten normal modes can already provide qualitatively correct picture, which can be further refined by the inclusion of additional modes. The choice of using the first ten low frequency modes is also corroborated by the previous studies on the same set of proteins [21, 15], where it was shown that the first ten normal modes fairly well described conformational transitions and allosteric communication.

Protein data set and computational framework. The method introduced in this work is applied to the list of allosteric enzymes used in previous studies in the context of binding leverage and leverage coupling calculations [15, 21]. Most of the proteins in the list (ATCase, AnthS, BGDH, CAP, DAHPS, DAK, G6PD, NADME, PFK, PGDH, PKA, PTP1B, SSUPRT, ThrS) are homo-oligomeric complexes, except two hetero-oligomers (ATCase, AnthS) and two monomeric forms (PKA, PTP1B). For each of the analyzed proteins at least two forms were used (apo form whenever that was available). All the calculations considered the structures cleaned of the all bound ligands. The actual protein assemblies (oligomeric structures) used as an input in this work were obtained from the PDBePISA (Proteins, Interfaces, Structures and Assemblies) interactive tool [56].

The protein modeling as well as the normal mode calculations were carried out with the MMTK package [49]. The C_α effective harmonic model described in [48] was used to characterize the dynamics of the ligand free and ligand bound protein systems. Specifically, the combination of the Harmonic Distance Restraint module with the Compound Force Field feature were used to construct the harmonic model associated to the ligand bound protein. The Fourier approximation was used for the calculations of the normal modes [57]. The MMTK calculations as well as the implementation of the three steps of the model presented above were coded in an iPython notebook which is available upon request.

Supporting Information

S1 Fig. The plots showing the dependence of the allosteric free energy profiles on the different values of the stiffening parameter α that characterize the ligand bound protein states. (PDF)

S2 Fig. The plots showing the dependence of the allosteric free energy profiles on the number of used normal modes. (PDF)

S3 Fig. The allosteric free energy profiles of PKA and PTP1B proteins. (PDF)

S1 Table. Results on allosteric causality and energetics obtained for proteins analyzed in this work. The table includes the complete data obtained for all available forms of proteins. For designation of columns see [Table 1](#) in the main text.

(PDF)

Acknowledgments

The authors thank Stephen Fox for critical reading of the manuscript, valuable comments, and editing.

Author Contributions

Conceived and designed the experiments: EG INB. Performed the experiments: EG. Analyzed the data: EG INB. Wrote the paper: EG INB.

References

1. Monod J, Changeux JP, Jacob F. Allosteric proteins and cellular control systems. *Journal of Molecular Biology*. 1963 Apr; 6:306–329. doi: [10.1016/S0022-2836\(63\)80091-1](https://doi.org/10.1016/S0022-2836(63)80091-1) PMID: [13936070](https://pubmed.ncbi.nlm.nih.gov/13936070/)
2. Perutz MF, Bolton W, Diamond R, Muirhead H, Watson HC. Structure of haemoglobin. An x-ray examination of reduced horse haemoglobin. *Nature*. 1964 Aug; 203:687–690. doi: [10.1038/203687a0](https://doi.org/10.1038/203687a0) PMID: [14207261](https://pubmed.ncbi.nlm.nih.gov/14207261/)
3. Monod J, Wyman J, Changeux JP. On the nature of allosteric transitions: A plausible model. *Journal of Molecular Biology*. 1965 May; 12(1):88–118. doi: [10.1016/S0022-2836\(65\)80285-6](https://doi.org/10.1016/S0022-2836(65)80285-6) PMID: [14343300](https://pubmed.ncbi.nlm.nih.gov/14343300/)
4. Koshland DE, Némethy G, Filmer D. Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry*. 1966 Jan; 5(1):365–385. doi: [10.1021/bi00865a047](https://doi.org/10.1021/bi00865a047) PMID: [5938952](https://pubmed.ncbi.nlm.nih.gov/5938952/)
5. Austin R, Beeson KW, Eisenstein L, Frauenfelder H, Gunsalus I. Dynamics of ligand binding to myoglobin. *Biochemistry*. 1975; 14(24):5355–5373. doi: [10.1021/bi00695a021](https://doi.org/10.1021/bi00695a021) PMID: [1191643](https://pubmed.ncbi.nlm.nih.gov/1191643/)
6. Beece D, Eisenstein L, Frauenfelder H, Good D, Marden M, Reinisch L, et al. Solvent viscosity and protein dynamics. *Biochemistry*. 1980; 19(23):5147–5157. doi: [10.1021/bi00564a001](https://doi.org/10.1021/bi00564a001) PMID: [7448161](https://pubmed.ncbi.nlm.nih.gov/7448161/)
7. Fenimore PW, Frauenfelder H, McMahon B, Young R. Bulk-solvent and hydration-shell fluctuations, similar to α - and β -fluctuations in glasses, control protein motions and functions. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101(40):14408–14413. doi: [10.1073/pnas.0405573101](https://doi.org/10.1073/pnas.0405573101) PMID: [15448207](https://pubmed.ncbi.nlm.nih.gov/15448207/)
8. Frauenfelder H, Petsko GA, Tsernoglou D. Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature*. 1979; 280:558–563. doi: [10.1038/280558a0](https://doi.org/10.1038/280558a0) PMID: [460437](https://pubmed.ncbi.nlm.nih.gov/460437/)
9. Frauenfelder H, Sligar SG, Wolynes PG. The energy landscapes and motions of proteins. *Science*. 1991; 254(5038):1598–1603. doi: [10.1126/science.1749933](https://doi.org/10.1126/science.1749933) PMID: [1749933](https://pubmed.ncbi.nlm.nih.gov/1749933/)
10. Tzeng SR, Kalodimos CG. Protein dynamics and allostery: an NMR view. *Current opinion in structural biology*. 2011; 21(1):62–67. doi: [10.1016/j.sbi.2010.10.007](https://doi.org/10.1016/j.sbi.2010.10.007) PMID: [21109422](https://pubmed.ncbi.nlm.nih.gov/21109422/)
11. Cui Q, Karplus M. Allostery and cooperativity revisited. *Protein Science*. 2008 Aug; 17(8):1295–1307. doi: [10.1110/ps.03259908](https://doi.org/10.1110/ps.03259908) PMID: [18560010](https://pubmed.ncbi.nlm.nih.gov/18560010/)
12. Cooper A, Dryden DTF. Allostery without conformational change. *European Biophysics Journal*. 1984 Oct; 11(2):103–109. doi: [10.1007/BF00276625](https://doi.org/10.1007/BF00276625) PMID: [6544679](https://pubmed.ncbi.nlm.nih.gov/6544679/)
13. Hardy JA, Wells JA. Searching for new allosteric sites in enzymes. *Current opinion in structural biology*. 2004; 14(6):706–715. doi: [10.1016/j.sbi.2004.10.009](https://doi.org/10.1016/j.sbi.2004.10.009) PMID: [15582395](https://pubmed.ncbi.nlm.nih.gov/15582395/)
14. Tsai CJ, Nussinov R. A unified view of “how allostery works”. *PLoS computational biology*. 2014 Feb; 10(2):e1003394. doi: [10.1371/journal.pcbi.1003394](https://doi.org/10.1371/journal.pcbi.1003394) PMID: [24516370](https://pubmed.ncbi.nlm.nih.gov/24516370/)
15. Mitternacht S, Berezovsky IN. Coherent conformational degrees of freedom as a structural basis for allosteric communication. *PLoS computational biology*. 2011 Dec; 7(12):e1002301. doi: [10.1371/journal.pcbi.1002301](https://doi.org/10.1371/journal.pcbi.1002301) PMID: [22174669](https://pubmed.ncbi.nlm.nih.gov/22174669/)
16. Saibil HR, Fenton WA, Clare DK, Horwich AL. Structure and allostery of the chaperonin GroEL. *Journal of molecular biology*. 2013; 425(9):1476–1487. doi: [10.1016/j.jmb.2012.11.028](https://doi.org/10.1016/j.jmb.2012.11.028) PMID: [23183375](https://pubmed.ncbi.nlm.nih.gov/23183375/)
17. Smith NJ, Milligan G. Allostery at G protein-coupled receptor homo- and heteromers: uncharted pharmacological landscapes. *Pharmacological reviews*. 2010; 62(4):701–725. doi: [10.1124/pr.110.002667](https://doi.org/10.1124/pr.110.002667) PMID: [21079041](https://pubmed.ncbi.nlm.nih.gov/21079041/)

18. Wenthur CJ, Gentry PR, Mathews TP, Lindsley CW. Drugs for allosteric sites on receptors. Annual review of pharmacology and toxicology. 2014; 54:165. doi: [10.1146/annurev-pharmtox-010611-134525](https://doi.org/10.1146/annurev-pharmtox-010611-134525) PMID: [24111540](https://pubmed.ncbi.nlm.nih.gov/24111540/)
19. Auerbach A. The energy and work of a ligand-gated ion channel. Journal of molecular biology. 2013; 425(9):1461–1475. doi: [10.1016/j.jmb.2013.01.027](https://doi.org/10.1016/j.jmb.2013.01.027) PMID: [23357172](https://pubmed.ncbi.nlm.nih.gov/23357172/)
20. Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? Proteins: Structure, Function, and Bioinformatics. 2004; 57(3):433–443.
21. Mitternacht S, Berezovsky IN. Binding leverage as a molecular basis for allosteric regulation. PLoS computational biology. 2011 Sep; 7(9):e1002148. doi: [10.1371/journal.pcbi.1002148](https://doi.org/10.1371/journal.pcbi.1002148) PMID: [21935347](https://pubmed.ncbi.nlm.nih.gov/21935347/)
22. Guarnera E, Berezovsky IN. Allosteric sites: remote control in regulation of protein activity. Curr Opin Struct Biol. 2016; 37:1–8. doi: [10.1016/j.sbi.2015.10.004](https://doi.org/10.1016/j.sbi.2015.10.004)
23. Goncarenco A, Mitternacht S, Yong T, Eisenhaber B, Eisenhaber F, Berezovsky IN. SPACER: Server for predicting allosteric communication and effects of regulation. Nucleic acids research. 2013 Jul; 41 (Web Server issue):W266–72. doi: [10.1093/nar/gkt460](https://doi.org/10.1093/nar/gkt460) PMID: [23737445](https://pubmed.ncbi.nlm.nih.gov/23737445/)
24. del Sol A, Tsai CJ, Ma B, Nussinov R. The origin of allosteric functional modulation: multiple pre-existing pathways. Structure. 2009; 17(8):1042–1050. doi: [10.1016/j.str.2009.06.008](https://doi.org/10.1016/j.str.2009.06.008) PMID: [19679084](https://pubmed.ncbi.nlm.nih.gov/19679084/)
25. Berg JM, Tymoczko JL, Stryer L. Biochemistry: international edition. WH Freeman & Company Limited; 2006.
26. Thomas A, Field MJ, Perahia D. Analysis of the low-frequency normal modes of the R state of aspartate transcarbamylase and a comparison with the T state modes. Journal of molecular biology. 1996; 261(3):490–506. doi: [10.1006/jmbi.1996.0478](https://doi.org/10.1006/jmbi.1996.0478) PMID: [8780788](https://pubmed.ncbi.nlm.nih.gov/8780788/)
27. Spraggon G, Kim C, Nguyen-Huu X, Yee MC, Yanofsky C, Mills SE. The structures of anthranilate synthase of *Serratia marcescens* crystallized in the presence of (i) its substrates, chorismate and glutamine, and a product, glutamate, and (ii) its end-product inhibitor, L-tryptophan. Proceedings of the National Academy of Sciences. 2001; 98(11):6021–6026. doi: [10.1073/pnas.111150298](https://doi.org/10.1073/pnas.111150298)
28. Banerjee S, Schmidt T, Fang J, Stanley CA, Smith TJ. Structural studies on ADP activation of mammalian glutamate dehydrogenase and the evolution of regulation. Biochemistry. 2003; 42(12):3446–3456. doi: [10.1021/bi0206917](https://doi.org/10.1021/bi0206917) PMID: [12653548](https://pubmed.ncbi.nlm.nih.gov/12653548/)
29. Smith TJ, Schmidt T, Fang J, Wu J, Siuzdak G, Stanley CA. The structure of apo human glutamate dehydrogenase details subunit communication and allostery. Journal of molecular biology. 2002; 318(3):765–777. doi: [10.1016/S0022-2836\(02\)00161-4](https://doi.org/10.1016/S0022-2836(02)00161-4) PMID: [12054821](https://pubmed.ncbi.nlm.nih.gov/12054821/)
30. Popovych N, Sun S, Ebright RH, Kalodimos CG. Dynamically driven protein allostery. Nature structural & molecular biology. 2006 Sep; 13(9):831–838. doi: [10.1038/nsmb1132](https://doi.org/10.1038/nsmb1132)
31. Popovych N, Tzeng SR, Tonelli M, Ebright RH, Kalodimos CG. Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. Proceedings of the National Academy of Sciences. 2009; 106(17):6927–6932. doi: [10.1073/pnas.0900595106](https://doi.org/10.1073/pnas.0900595106)
32. Wu X, Ye S, Guo S, Yan W, Bartlam M, Rao Z. Structural basis for a reciprocating mechanism of negative cooperativity in dimeric phosphagen kinase activity. The FASEB Journal. 2010; 24(1):242–252. doi: [10.1096/fj.09-140194](https://doi.org/10.1096/fj.09-140194) PMID: [19783784](https://pubmed.ncbi.nlm.nih.gov/19783784/)
33. Horjales E, Altamirano MM, Calcagno ML, Garratt RC, Oliva G. The allosteric transition of glucosamine-6-phosphate deaminase: the structure of the T state at 2.3 Å resolution. Structure. 1999; 7(5):527–537. PMID: [10378272](https://pubmed.ncbi.nlm.nih.gov/10378272/)
34. Oliva G, Fontes MR, Garratt RC, Altamirano MM, Calcagno ML, Horjales E. Structure and catalytic mechanism of glucosamine 6-phosphate deaminase from *Escherichia coli* at 2.1 Å resolution. Structure. 1995; 3(12):1323–1332. doi: [10.1016/S0969-2126\(01\)00270-2](https://doi.org/10.1016/S0969-2126(01)00270-2) PMID: [8747459](https://pubmed.ncbi.nlm.nih.gov/8747459/)
35. Yang Z, Lanks CW, Tong L. Molecular mechanism for the regulation of human mitochondrial NAD(P)⁺-dependent malic enzyme by ATP and fumarate. Structure. 2002; 10(7):951–960. PMID: [12121650](https://pubmed.ncbi.nlm.nih.gov/12121650/)
36. Arent S, Harris P, Jensen KF, Larsen S. Allosteric regulation and communication between subunits in uracil phosphoribosyltransferase from *Sulfolobus solfataricus*. Biochemistry. 2005; 44(3):883–892. doi: [10.1021/bi048041l](https://doi.org/10.1021/bi048041l) PMID: [15654744](https://pubmed.ncbi.nlm.nih.gov/15654744/)
37. Mas-Droux C, Biou V, Dumas R. Allosteric Threonine Synthase reorganization of the pyridoxal phosphate site upon asymmetric activation through s-adenosylmethionine binding to a novel site. Journal of Biological Chemistry. 2006; 281(8):5188–5196. doi: [10.1074/jbc.M509798200](https://doi.org/10.1074/jbc.M509798200) PMID: [16319072](https://pubmed.ncbi.nlm.nih.gov/16319072/)
38. Schuller DJ, Grant GA, Banaszak LJ. The allosteric ligand site in the Vmax-type cooperative enzyme phosphoglycerate dehydrogenase. Nature structural biology. 1995; 2(1):69–76. doi: [10.1038/nsb0195-69](https://doi.org/10.1038/nsb0195-69) PMID: [7719856](https://pubmed.ncbi.nlm.nih.gov/7719856/)
39. Traut TW. Regulatory allosteric enzymes. Springer Science, New York; 2007.

40. Schirmer T, Evans PR. Structural basis of the allosteric behaviour of phosphofructokinase. *Nature*. 1990; 343(6254):140–145. doi: [10.1038/343140a0](https://doi.org/10.1038/343140a0) PMID: [2136935](https://pubmed.ncbi.nlm.nih.gov/2136935/)
41. Valdez BC, French BA, Younathan ES, Chang S. Site-directed mutagenesis in *Bacillus stearothermophilus* fructose-6-phosphate 1-kinase. Mutation at the substrate-binding site affects allosteric behavior. *Journal of Biological Chemistry*. 1989; 264(1):131–135. PMID: [2521215](https://pubmed.ncbi.nlm.nih.gov/2521215/)
42. Sadowsky JD, Burlingame MA, Wolan DW, McClendon CL, Jacobson MP, Wells JA. Turning a protein kinase on or off from a single allosteric site via disulfide trapping. *Proceedings of the National Academy of Sciences*. 2011; 108(15):6056–6061. doi: [10.1073/pnas.1102376108](https://doi.org/10.1073/pnas.1102376108)
43. Berezovsky IN. Thermodynamics of allostery paves a way to allosteric drugs. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*. 2013; 1834(5):830–835. doi: [10.1016/j.bbapap.2013.01.024](https://doi.org/10.1016/j.bbapap.2013.01.024)
44. Nussinov R, Tsai CJ. Unraveling structural mechanisms of allosteric drug action. *Trends in pharmacological sciences*. 2014; 35(5):256–264. doi: [10.1016/j.tips.2014.03.006](https://doi.org/10.1016/j.tips.2014.03.006) PMID: [24742712](https://pubmed.ncbi.nlm.nih.gov/24742712/)
45. Brooks B, Karplus M. Normal modes for specific motions of macromolecules: application to the hinge-bending mode of lysozyme. *Proceedings of the National Academy of Sciences of the United States of America*. 1985 Aug; 82(15):4995–4999. doi: [10.1073/pnas.82.15.4995](https://doi.org/10.1073/pnas.82.15.4995) PMID: [3860838](https://pubmed.ncbi.nlm.nih.gov/3860838/)
46. Tirion MM. Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic Analysis. *Physical Review Letters*. 1996 Aug; 77(9):1905–1908. doi: [10.1103/PhysRevLett.77.1905](https://doi.org/10.1103/PhysRevLett.77.1905) PMID: [10063201](https://pubmed.ncbi.nlm.nih.gov/10063201/)
47. Yang L, Song G, Jernigan RL. How Well Can We Understand Large-Scale Protein Motions Using Normal Modes of Elastic Network Models? *Biophysical Journal*. 2007 Aug; 93(3):920–929. doi: [10.1529/biophysj.106.095927](https://doi.org/10.1529/biophysj.106.095927) PMID: [17483178](https://pubmed.ncbi.nlm.nih.gov/17483178/)
48. Hinsen K, Petrescu AJ, Dellerue S, Bellissent-Funel MC, Kneller GR. Harmonicity in slow protein dynamics. *Chemical Physics*. 2000; 261(1):25–37. doi: [10.1016/S0301-0104\(00\)00222-6](https://doi.org/10.1016/S0301-0104(00)00222-6)
49. Hinsen K. The molecular modeling toolkit: A new approach to molecular simulations. *Journal of Computational Chemistry*. 2000 Jan; 21(2):79–85. doi: [10.1002/\(SICI\)1096-987X\(20000130\)21:2%3C79::AID-JCC1%3E3.0.CO;2-B](https://doi.org/10.1002/(SICI)1096-987X(20000130)21:2%3C79::AID-JCC1%3E3.0.CO;2-B)
50. Piazza F, Sanejouand YH. Long-range energy transfer in proteins. *Physical biology*. 2009; 6(4):046014. doi: [10.1088/1478-3975/6/4/046014](https://doi.org/10.1088/1478-3975/6/4/046014) PMID: [19910672](https://pubmed.ncbi.nlm.nih.gov/19910672/)
51. McLeish TC, Rodgers T, Wilson MR. Allostery without conformation change: modelling protein dynamics at multiple scales. *Physical biology*. 2013; 10(5):056004. doi: [10.1088/1478-3975/10/5/056004](https://doi.org/10.1088/1478-3975/10/5/056004) PMID: [24021665](https://pubmed.ncbi.nlm.nih.gov/24021665/)
52. Zheng W, Brooks BR, Doniach S, Thirumalai D. Network of Dynamically Important Residues in the Open/Closed Transition in Polymerases Is Strongly Conserved. *Structure (London, England: 1993)*. 2005 Jan; 13(4):565–577. doi: [10.1016/j.str.2005.01.017](https://doi.org/10.1016/j.str.2005.01.017)
53. Ming D, Wall ME. Quantifying allosteric effects in proteins. *Proteins: Structure, Function, and Bioinformatics*. 2005 Jun; 59(4):697–707. doi: [10.1002/prot.20440](https://doi.org/10.1002/prot.20440)
54. Rodgers TL, Townsend PD, Burnell D, Jones ML, Richards SA, McLeish TCB, et al. Modulation of global low-frequency motions underlies allosteric regulation: demonstration in CRP/FNR family transcription factors. *PLoS biology*. 2013 Sep; 11(9):e1001651. doi: [10.1371/journal.pbio.1001651](https://doi.org/10.1371/journal.pbio.1001651) PMID: [24058293](https://pubmed.ncbi.nlm.nih.gov/24058293/)
55. Eren D, Alakent B. Frequency response of a protein to local conformational perturbations. *PLoS computational biology*. 2013; 9(9):e1003238. doi: [10.1371/journal.pcbi.1003238](https://doi.org/10.1371/journal.pcbi.1003238) PMID: [24086121](https://pubmed.ncbi.nlm.nih.gov/24086121/)
56. Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology*. 2007; 372(3):774–797. doi: [10.1016/j.jmb.2007.05.022](https://doi.org/10.1016/j.jmb.2007.05.022) PMID: [17681537](https://pubmed.ncbi.nlm.nih.gov/17681537/)
57. Hinsen K. Analysis of domain motions by approximate normal mode calculations. *Proteins: Structure, Function, and Bioinformatics*. 1998 Nov; 33(3):417–429. doi: [10.1002/\(SICI\)1097-0134\(19981115\)33:3%3C417::AID-PROT10%3E3.0.CO;2-8](https://doi.org/10.1002/(SICI)1097-0134(19981115)33:3%3C417::AID-PROT10%3E3.0.CO;2-8)