

Allosteric pathways in imidazole glycerol phosphate synthase

Ivan Rivalta^{a,1}, Mohammad M. Sultan^a, Ning-Shiuan Lee^a, Gregory A. Manley^a, J. Patrick Loria^{a,b,1}, and Victor S. Batista^{a,1}

^aDepartment of Chemistry and ^bDepartment of Molecular Biophysics and Biochemistry, Yale University, P.O. Box 208107, New Haven, CT 06520-8107

Edited by Michael L. Klein, Temple University, Philadelphia, PA, and approved March 28, 2012 (received for review December 14, 2011)

Protein allosteric pathways are investigated in the imidazole glycerol phosphate synthase heterodimer in an effort to elucidate how the effector (PRFAR, N'-[(5'-phosphoribulosyl)formimino]-5-aminoimidazole-4-carboxamide ribonucleotide) activates glutaminase catalysis at a distance of 25 Å from the glutamine-binding site. We apply solution NMR techniques and community analysis of dynamical networks, based on mutual information of correlated protein motions in the active and inactive enzymes. We find evidence that the allosteric pathways in the PRFAR bound enzyme involve conserved residues that correlate motion of the PRFAR binding loop to motion at the protein-protein interface, and ultimately at the glutaminase active site. The imidazole glycerol phosphate synthase holoenzyme is an important branch point for the histidine and nucleotide biosynthetic pathways and represents a potential therapeutic target against microbes. The proposed allosteric mechanism and the underlying allosteric pathways provide fundamental insights for the design of new allosteric drugs and/or alternative herbicides.

glutamine hydrolysis | protein networks | generalized correlation analysis | network theory

Allostery is a fundamental property that allows for the regulation of function and dynamic adaptability of enzymes and proteins. Allosteric enzymes contain at least two distant binding sites, including the active site responsible for catalytic activity, which binds the substrate, and the allosteric site, which binds the effector and initiates the allosteric signal propagation to the active site. In V-type systems, substrate binding is not affected by the presence of the effector but if the effector is not bound, the allosteric protein is usually catalytically inactive (or poorly active), indicating that the effector binding is coupled to the kinetic and/or thermodynamic parameters of the biochemical reaction in the active site. Allosteric information transfer can range from large, enthalpically driven conformational changes to purely entropically driven motions or a combination of both enthalpic and entropic effects, but in each case the kinetic parameters of the catalyzed reaction at the substrate binding site are altered. At the heart of allostery there is intramolecular thermodynamic coupling over long distances (>10 Å), between the active and allosteric sites. An important challenge for fundamental studies is the elucidation of the allosteric pathways that connect the two ligand-binding sites.

In this work, we combine community network analysis based on molecular dynamics (MD) simulations and NMR studies of protein motion based on relaxation dispersion techniques and chemical shift titrations experiments to provide an atomistic description of allostery in the V-type allosteric enzyme imidazole glycerol phosphate synthase (IGPS) from the thermophile *Thermotoga maritima* (Fig. 1). IGPS is a tightly associated heterodimeric enzyme in which each monomer enzyme catalyzes a different reaction (1–3). The 23 kDa HisH enzyme is a member of the glutamine amidotransferase family that catalyzes the hydrolysis of glutamine (Gln) to produce ammonia. The ammonia travels to the other monomer, HisF, passing through the (β/α)₈ barrel and reacting with the effector PRFAR; i.e., N'-[(5'-phosphoribulosyl)formimino]-5-aminoimidazole-4-carboxamide ribo-

nucleotide, to form imidazole glycerol phosphate (IGP) and 5-aminoimidazole-4-carboxamide ribotide (1, 4, 5). These two products enter the histidine and purine biosynthetic pathways, respectively. IGPS is not present in mammals, but it is involved in essential biosynthetic pathways of microorganisms. In particular, many plant pathogens, and importantly opportunistic human pathogens such as *Cryptococcus*, *Candida*, and *Ajellomyces*, have an IGPS that is highly homologous to the *T. maritima* enzyme. The allosteric IGPS complex is, thus, a potential target for antifungal, antibiotic, and herbicide development (2, 6, 7).

IGPS is a V-type allosteric enzyme because the Gln binding affinity is unaffected by the presence of PRFAR. However, hydrolysis of Gln is accelerated 5,000-fold upon PRFAR binding (8). Here, we explore the effect of PRFAR on correlated motions of amino acid residues and we use this information to evaluate how the effector binding affects the protein network of the apo enzyme. Our methodology is inspired by studies of signaling pathways in tRNA:protein complexes (9). We combine solution NMR techniques to probe the motion of amino acid residues in the ms time scale and chemical shift titrations data with MD simulations for the analysis of mutual information based on Shannon's entropy to quantify the extent of correlated motion of amino acid residues (10). The resulting analysis shows evidence that PRFAR binding alters the community structure of the inactive IGPS protein network, along with the collective pathways linking the effector and active sites, providing valuable insight on the mechanism of information transfer in this V-type allosteric enzyme.

The paper is organized as follows. We first analyze motions in the apo and the effector bound complexes, then compare correlated motions among amino acid residues by measuring the mutual information. We use this measurement to build dynamical protein networks and compare community structures of apo and PRFAR bound IGPS complexes. From the comparison of community structures and the analysis of information paths, IGPS allosteric pathways are obtained. We then propose an allosteric mechanism for the activations of the glutaminase reaction induced by PRFAR binding. Conclusions are summarized in the final section.

Results and Discussion

MD simulations show differences in motion of the flexible loop1 (HisF domain) when comparing apo and PRFAR bound IGPS complexes (Fig. 2). We observed that loop1 in HisF is more ordered and that the two-stranded β -sheet is further extended to the beginning of *fa1* upon PRFAR binding, stabilized by hydro-

Author contributions: I.R., J.P.L., and V.S.B. designed research; I.R., M.M.S., N.-S.L., and G.A.M. performed research; I.R., M.M.S., N.-S.L., G.A.M., J.P.L., and V.S.B. analyzed data; and I.R., J.P.L., and V.S.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence may be addressed. E-mail: ivan.rivalta@yale.edu, patrick.loria@yale.edu, or victor.batista@yale.edu.

See Author Summary on page 8366 (volume 109, number 22).

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1120536109/-DCSupplemental.

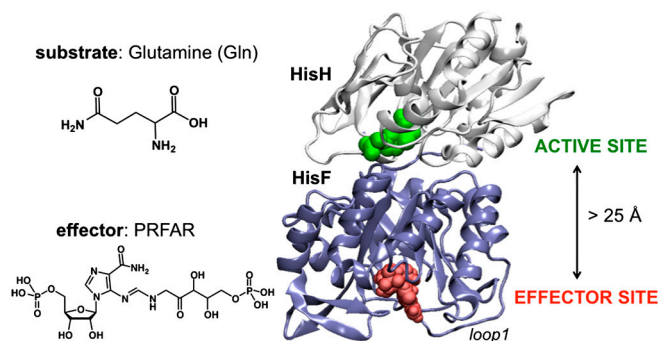


Fig. 1. Structural model of IGPS heterodimer from *T. maritima*, including protein subunits HisH and HisF, glutamine substrate (green) and allosteric effector, PRFAR (red), and flexible loop1 in the cyclase domain.

gen bonding between the highly conserved amino acid residues *f*V17 and *f*G30.

PRFAR binding also alters the breathing motion at the interdomain interface. Fig. 2 shows the low-frequency interdomain motions observed along representative 100 ns trajectories of both apo and PRFAR-bound IGPS complexes. This breathing motion allows for fluctuations between two states (open and closed IGPS

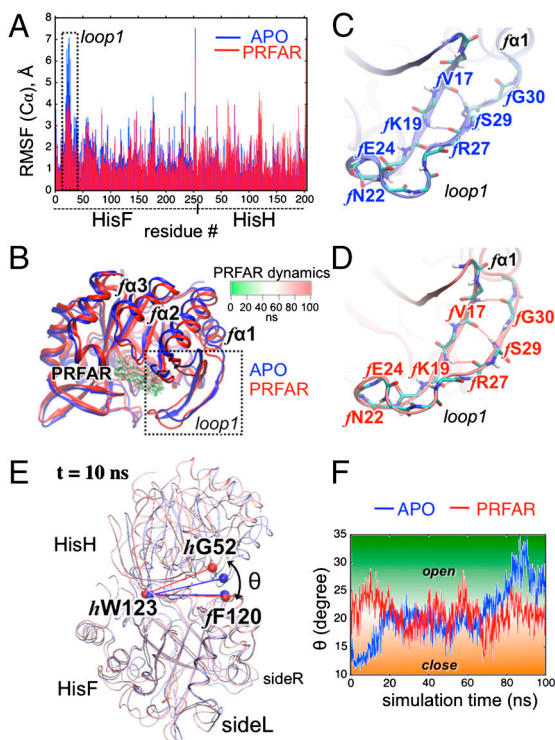


Fig. 2. (A) Root mean square fluctuations (rmsf) of Cα atoms for apo (blue) and PRFAR bound (red) IGPS complexes, calculated with respect to the average structures of the four independent 100 ns MD simulations. (B) HisF average structures of apo (blue) and PRFAR bound (red) enzymes, showing differences induced by PRFAR binding in loop1 and *fa*1, *fa*2, and *fa*3 helices. PRFAR dynamics in the cyclase binding site (colored sticks from red to green for 0–100 ns) are shown. (C) and (D) Hydrogen-bonds in (HisF) loop1 in apo (C, in blue) and PRFAR bound (D, in red) enzymes at 50 ns, showing the extension of the small two-stranded β-sheet via the formation of the *f*V17-*f*G30 hydrogen-bond at the loop1-*fa*1 junction. (E) and (F) Breathing motion in apo (blue) and PRFAR bound (red) IGPS complexes. Instantaneous configuration at *t* = 10 ns (E), showing the larger angle *θ* between Cα atoms (F) of *f*F120, *h*W123, *h*G52 (red and blue balls) in PRFAR bound (red) with respect to the apo complex (blue). Thermal fluctuations of *θ* for representative trajectories of apo and PRFAR-bound complexes (F). Note that *θ* fluctuates back and forth between open and closed configurations 4 times/100 ns for PRFAR bound and 1/100ns for apo.

heterodimer) that are accessible at thermal equilibrium in both the apo and PRFAR complexes. However, opening-closing events are observed more frequently when PRFAR is bound to the cyclase domain. The red trace in Fig. 2F shows that the angle *θ* [i.e., the angle between Cα atoms of *f*F120, *h*W123, *h*G52] opens and closes 4 times/100 ns when PRFAR is bound. Note the maximum at approximately 15 ns followed by a minimum at approximately 30 ns, then a maximum at approximately 40 ns followed by a minimum at approximately 50 ns, a maximum at approximately 60 ns followed by a minimum at approximately 70 ns, and then a maximum at approximately 85 ns followed by a minimum at 100 ns. In contrast, for apo-IGPS (blue trace) *θ* opens and closes only once/100 ns. Similar results are shown for other simulations in *SI Text*. These motions are observed in several independent 100 ns simulations performed, suggesting that PRFAR binding alters the protein-protein interactions at the HisF-HisH interface, consistent with the hinge-like displacement observed in the comparative analysis of crystal structures for the IGPS apo complex from *T. maritima* (5) and the PRFAR bound IGPS from yeast (11) (*SI Text*).

These observations corroborate the hypothesis that the motion of loop1 is coupled to the interdomain “breathing” motion, as previously indicated by steered MD simulations of the effector undocking (12). While a network of conserved residues, including *h*W123-*f*R249 (hinge) and *h*K181-*f*D98 (salt-bridge anchor), has been proposed to play a crucial role (12), the allosteric pathway has yet to be established. In the remainder of this section we address the allosteric mechanism in terms of the analysis of correlations in the motion of amino acid residues, with emphasis on the effect of PRFAR binding on those correlations. We find evidence of the allosteric pathway from the analysis of dynamical networks obtained by weighting the graphs with the mutual information that results from correlations of nuclear displacements in apo and PRFAR bound complexes.

Correlation Analysis of Protein Motions. The analysis of correlations of fluctuations in the positions of Cα-atoms is based on the generalized correlation coefficient $r_{\text{MI}}[\mathbf{x}_i, \mathbf{x}_j] = [1 - \exp(-\frac{2}{3}I[\mathbf{x}_i, \mathbf{x}_j])]^{-\frac{1}{2}}$ computed in terms of the mutual information (10), $I[\mathbf{x}_i, \mathbf{x}_j] = \mathbf{H}[\mathbf{x}_i] + \mathbf{H}[\mathbf{x}_j] - \mathbf{H}[\mathbf{x}_i, \mathbf{x}_j]$. Here, $\mathbf{H}[\mathbf{x}_i]$, $\mathbf{H}[\mathbf{x}_j]$ and $\mathbf{H}[\mathbf{x}_i, \mathbf{x}_j]$ are the marginal and joint (Shannon) entropies (13) for atomic vector displacements ($\mathbf{x}_i, \mathbf{x}_j$), computed as ensemble averages over multiple 50 ns windows along eight independent 100 ns MD simulations for both apo and PRFAR-bound IGPS complexes (see *Materials and Methods*).

Fig. 3 shows our results for r_{MI} of apo-IGPS and differences with respect to the corresponding values for the PRFAR bound complex. Large values ($r_{\text{MI}} > 0.8$) are obtained for highly coupled motions in the apo enzyme, including residues with common secondary structure elements (near off-diagonal elements in Fig. 3A). Amino acid residues at the HisF-HisH interface also show relatively high correlations in the apo complex (r_{MI} around 0.5–0.6), including residues in the *ha*1 helix and the Ω-loop (HisH loop1 connecting *h*β1 and *ha*1) at one side of the complex (Fig. 3A, red box), and interface residues in the *f*β4, *f*β5 and *h*β7 sheets and in the *fa*4 helix at the opposite side of IGPS (Fig. 3A, blue box). In contrast, amino acid residues in the range 201–253 (belonging to *fa*7, *f*β8, *fa*8'-*fa*8 and the C-terminal loop of HisF) are poorly correlated with other residues in the HisH and HisF domains (Fig. 3A, dashed lines).

Comparison of correlated motions between apo and PRFAR bound enzymes reveals a reduction of correlations, induced by PRFAR binding, in agreement with the large, positive entropy change for PRFAR binding observed in isothermal titration calorimetry (ITC) experiments (14). The correlation matrix difference between PRFAR and apo complexes, shown in Fig. 3B, is indeed prevalently characterized by negative values (Fig. 3B,

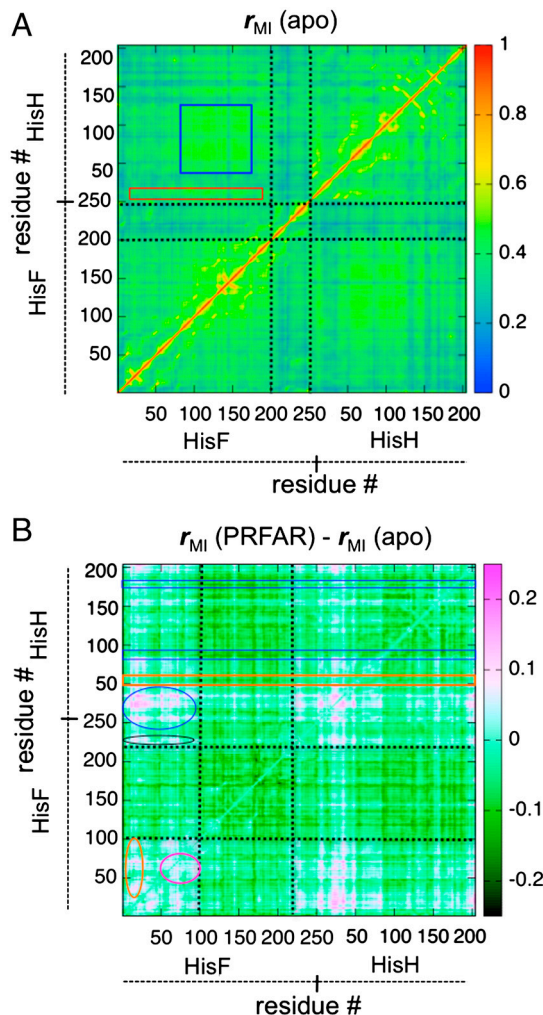


Fig. 3. Comparison of generalized correlation coefficients r_{MI} for apo (A) and PRFAR-minus-apo (B) IGPS complexes. Note low correlations for HisF amino acid residues 200–250 (A, bands within dotted lines), high correlation ($r_{MI} > 0.50$) between HisF residues and HisH residues in $\alpha 1$ and Ω -loop (A, red box) and interface residues in the $f\beta 4$, $f\beta 5$ and $h\beta 7$ sheets and in the $f\alpha 4$ helix (A, blue box) in apo IGPS. Decreased correlation for residues 100–220 in HisF (B, green features within dotted bands) and for residues in the 49-PGVG sequence (orange box) and around the active site (blue boxes) in HisH, increased correlation within residues 50–100 (B, magenta ellipse), between residues 50–100 and HisF loop1 (B, orange ellipse), between $h\beta 1$ - $h\beta 3$ sheets, $h\alpha 1$ helix, Ω -loop, and HisF loop1 (B, blue ellipse), and between residues 222–230 (PRFAR gP-binding site) and residues in the sideR of HisF (B, black ellipse) upon PRFAR binding.

green spots). In particular, the largest reduction of correlations is observed for residues in the $f\alpha 4$ '- $f\alpha 7$ helices and in the $f\beta 5$ - $f\beta 7$ β -sheets on one side of the HisF domain (residues 100–220) (Fig. 3B, dashed lines). Note that the features highlighted in Fig. 3B result from thermal averaging, including all the possible combinations of PRFAR-minus-apo differences available from multiple MD simulations (SI Text). The green color code indicates features with $r_{MI}(\text{PRFAR}) - r_{MI}(\text{apo}) = -0.1$, including a decrease of correlation for residues 100–220 (Fig. 3B, green bands). These changes are 5–10 times larger than the rmsds of the average matrix elements (SI Text). Magenta features, indicating an increase of correlation upon PRFAR binding, are also much larger than the rmsds.

The depletion of correlations indicates a certain degree of isolation of the decorrelated HisF amino acid residues from the rest of the HisF domain and from most of the residues in the HisH enzyme. The reduction of correlations in the PRFAR

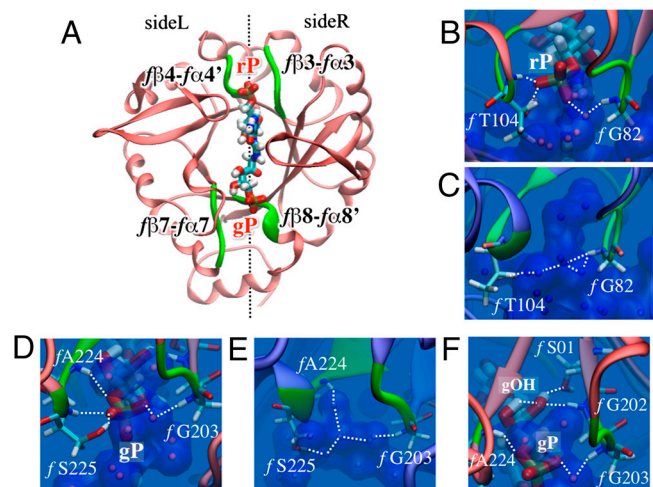


Fig. 4. PRFAR pose in the cyclase domain (HisF) and binding sites of the PRFAR phosphate groups. (A) Phosphate groups at the PRFAR ribose (rP) and glycerol (gP) sides are located between $f\beta 3$ - $f\alpha 3$ and $f\beta 4$ - $f\alpha 4'$ turns (green) and $f\beta 7$ - $f\alpha 7$ and $f\beta 8$ - $f\alpha 8'$ turns (green), respectively. (B) and (C) H-bonds (white dashed lines) at the interstitial site between $f\beta 3$ - $f\alpha 3$ and $f\beta 4$ - $f\alpha 4'$ turns (green), showing direct interactions between rP and fT104 ($f\beta 4$ - $f\alpha 4'$ turn) and water-mediated interactions between rP and fG83 ($f\beta 3$ - $f\alpha 3$ turn) in the PRFAR bound complex (B), water oxygen atoms shown as red balls. In the apo complex (C) only water molecules (blue balls) are present in the interstitial zone forming an hydrogen-bonding network. (D–F) H-bonds at the interstitial site between $f\beta 7$ - $f\alpha 7$ and $f\beta 8$ - $f\alpha 8'$ turns (green), showing hydrogen bonds between gP and fA224-S225 ($f\beta 8$ - $f\alpha 8'$ turn) and water-mediated interactions between gP and fG203 ($f\beta 7$ - $f\alpha 7$ turn) in the PRFAR bound complex (D). The extended hydrogen-bonding network in the apo complex at the $f\beta 7$ - $f\alpha 7$ and $f\beta 8$ - $f\alpha 8'$ site is shown in (E). (F) H-bonds formed by the hydroxyl groups of the PRFAR glycerol side (gOH) with amino acid residues in the $f\beta 7$ - $f\alpha 7$ turn, including fS201, fG202 and fG203.

bound complex is also indicative of a weakening of interface interactions in this region upon effector binding. Changes in correlations at the interface are associated with the previously discussed changes in breathing motions, observed in MD simulations of apo and PRFAR bound complexes (Fig. 2). To understand the molecular origin of these changes in motion, we characterized the detailed PRFAR-HisF interactions and how these interactions affect the structure of correlations.

The pose of PRFAR in the cyclase active sites extends across the bottom of the $(\beta/\alpha)_8$ barrel, lying on a long cleft that separates the underside of HisF in two regions (Fig. 4).

Phosphate groups at the glycerol side (gP) and at the ribose side (rP) of PRFAR are bound to the cyclase domain, occupying interstitial sites between the $f\beta 7$ - $f\alpha 7$ and $f\beta 8$ - $f\alpha 8'$ turns and between the $f\beta 3$ - $f\alpha 3$ and $f\beta 4$ - $f\alpha 4'$ turns, respectively. Hydrogen bonds between the rP group and both the side chain and the backbone of the highly conserved fT104 (in the $f\beta 4$ - $f\alpha 4'$ turn) are the strongest interactions between PRFAR and HisF, as also suggested by undocking simulations of the effector (12). PRFAR-HisF interactions via the rP-fT104 hydrogen bonds induces desolvation in the interstitial region of the $f\beta 3$ - $f\alpha 3$ and $f\beta 4$ - $f\alpha 4'$ turns. Water-mediated hydrogen bond interactions present in the apo complex at this site are replaced by PRFAR-mediated interactions (Fig. 4). On the glycerol side of PRFAR, hydrogen bonds of the gP group with the backbone of fA224 and hydroxyl group of fS225 side chain are the first contacts to be broken during undocking of the effector (12), indicating a formation of weaker interactions with respect to interactions in the ribose side. Our MD simulations also indicate that hydroxyl groups of the PRFAR glycerol moiety create a network of hydrogen bonds with highly conserved fS201 and fG202 residues at the end of $f\beta 7$ (Fig. 4), contributing to the consistent number of protein-effector interactions in the interstitial region between the $f\beta 7$ - $f\alpha 7$ and $f\beta 8$ - $f\alpha 8'$

turns. These effector-protein contacts substitute for protein-solvent interactions that modulate interresidue communication in this region of the apo complex. The observed desolvation upon PRFAR binding is consistent with an entropically driven binding of the effector, as observed in the ITC experiments (14). Desolvation around $f\beta3$ - $f\alpha3$, $f\beta4$ - $f\alpha4'$, $f\beta7$ - $f\alpha7$, and $f\beta8$ - $f\alpha8'$ turns (residues 100–220) and formation of selective effector-protein interactions at residues $fT104$, $fS201$ - $fG202$, and $fA224$ - $fS225$, upon PRFAR binding, induce the observed decrease of correlations within amino acid residues in the range 100–220.

Notably, PRFAR binding at the bottom of the HisF barrel induces a reduction of correlations in the HisH domain, which is more than 20 Å away from the effector binding site. In particular, amino acid residues of the IGPS catalytic triad ($hC84$, $hH178$, $hE180$) (1, 4, 5) and their closest neighborhoods show lower r_{MI} values in the PRFAR bound complex than in the apo complex (Fig. 2B, blue boxes). Motions of amino acid residues in the $hP49$ - $hG50$ - $hV51$ - $hG52$ sequence (49-PGVG, also called the oxyanion strand) are also affected by PRFAR binding, with a significant reduction of correlations induced by effector binding (Fig. 3B, orange box). This result is consistent with the previously proposed allosteric mechanism invoking conformational changes of the 49-PGVG sequence induced by PRFAR binding (2, 14), and more generally, with the notion that effector binding triggers specific motions in the (distant) active site that favor catalytically active conformations.

In contrast to the decrease in correlations observed for some groups of amino acid residues upon PRFAR binding, increased correlations (Fig. 3B, magenta features) are observed for residues on the opposite side of the HisF protein, sideR (Fig. 1, Right). This side of the heterodimer includes the flexible loop1 at the bottom of the $(\beta/\alpha)_8$ barrel, the $f\alpha2$ - $f\alpha3$ helices and the $f\beta2$ - $f\beta4$ sheets in HisF, the $ha1$ helix and the Ω -loop at the HisF-HisH interface, and the $h\beta1$ - $h\beta3$ sheets in HisH. Interestingly, an increment of correlated motions is observed between residues belonging to $f\alpha2$ - $f\alpha3$ helices and $f\beta2$ - $f\beta4$ sheets (Fig. 3B, magenta ellipse) and residues in the flexible loop1, including residues $fN22$ and $fF23$ (Fig. 3B, orange ellipse). Moreover, an increase of correlations due to effector binding is observed between HisF and HisH residues at the sideR (Fig. 3B, blue ellipse). In particular, motions of residues in the $h\beta1$ - $h\beta3$ sheets in HisH are correlated with motions of the HisF loop1 through interface residues in the $ha1$ helix and the Ω -loop. Finally, larger r_{MI} values between residues in the PRFAR gP-binding site (at the $f\beta8$ - $f\alpha8'$ turn, residues 222–230) and residues in the sideR of HisF are obtained in the presence of PRFAR (Fig. 3B, black ellipse), indicating the presence of other correlated motions involving the sideR of the $(\beta/\alpha)_8$ barrel.

In summary, the mutual information analysis indicates that PRFAR binding induces changes in correlations of motions that are heterogeneously distributed within the heterodimer complex. On one side of the IGPS complex, sideR, the effector binding induces an increase of correlations for residues within the HisF subunit and between HisF and HisH residues. Because amino acid residues in sideR such as the molecular hinge and the salt bridge anchor and provide structural stability to the IGPS heterodimer, the increase of correlations in this region upon PRFAR binding indicates that effector binding would not destabilize the heterodimeric form of IGPS during Gln hydrolysis. On the other hand, a significant decrease of correlations is observed in the opposite side of the IGPS complex (hereafter named sideL) indicating that the entropically driven PRFAR binding promotes the separation of the HisF-HisH domains in the sideL. More importantly, PRFAR binding induces a significant reduction of correlations for motion of residues belonging to the Gln active site and adjacent 49-PGVG β -strand, more than 25 Å away from the effector-binding site.

The correlation analysis does show evidence of significant changes in the protein motion induced by effector binding. However, a more complete analysis of the structure of correlations is required to elucidate the allosteric pathways and residues that are crucial for signal propagation from the allosteric to the active sites. For this purpose, we use a network theory method that provides understanding of the underlying allosteric mechanism in the IGPS heterodimer.

Networks and Communities of Residues. Dynamical weighted networks for both apo and PRFAR bound complexes were built from MD calculations of generalized correlation coefficients (r_{MI}) (see *Materials and Methods*). The weighted graphs are obtained using the average values of r_{MI} calculated for all the MD simulations, as reported in Fig. 2, representing the average protein networks. Analysis of the time-dependent evolution of the weighted networks along 100 ns indicates that the time-averaged networks adequately describe the dynamical networks, capturing the protein conformational changes induced by effector binding during the 50 ns dynamics of the IGPS complexes. While rearrangement of residues in the communities and changes in the total number of communities are observed among different community structures, due to limitations of the modularity methodology (*SI Text*), many common features are present in different community networks and consequently in the average community structures. As previously reported for the analysis of allosteric pathways in tRNA:protein complexes (9), this methodology can provide valuable insight not readily available from the measures of correlation alone. Here, we sensibly improve the evaluation of the correlated motions by using the generalized correlation coefficient instead of the Pearson coefficient, overcoming the underestimation of correlation associated with the conventional covariance analysis (10). We focus on the generalized correlation coefficients based

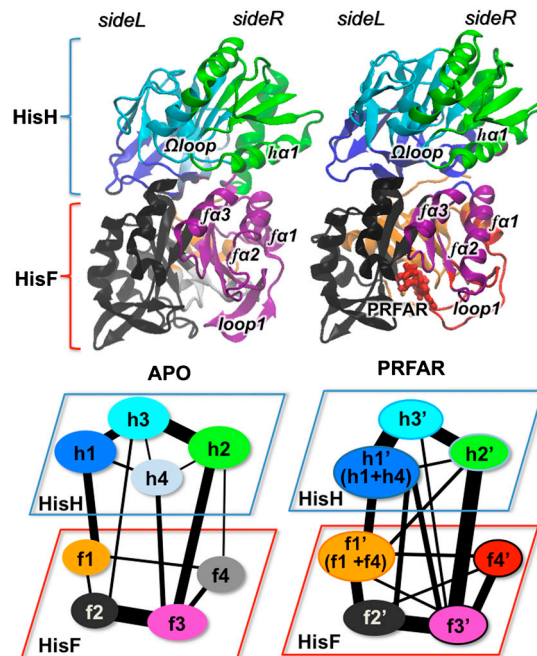


Fig. 5. Color coded optimal community network of the apo (Left) and PRFAR bound (Right) IGPS complexes. Color Scheme: h1, h1' (dark blue), h2, h2' (green), h3, h3' (cyan), h4 (light blue), f1, f1' (orange), f2, f2' (black), f3, f3' (purple), f4 (gray), and f4' (red), with predominant contributions from HisH (h) and HisF (f) subunits. The apo communities reported in brackets in the PRFAR network indicate those communities that contribute the most to the newly formed PRFAR communities. Widths of connectivity lines are proportional to the ITB with connections smaller than 10% of the maximum ITB value not shown, for clarity.

residue in providing structural rigidity and stability to the architecture of the IGPS heterodimer (15). Our community analysis consistently indicates that *hN15* plays an important role in the communication network of the apo complex, providing stable contacts between the *ha1* helix and the *fa2-fβ3* and *fβ3-fa2* turns at the sideR of IGPS. Interestingly, *hN15* is connected to the rest of the *h2* community via *hN12* in the Ω -loop.

PRFAR binding induces an increase in the communication flow between communities *f3'* and *h2'* with respect to the flow between the apo communities *f3* and *h2*. This increment in information flow is due to the enhanced correlations in the sideR of the PRFAR-bound complex, with increased communication between the helix *ha1* and the Ω -loop in HisH and the HisF interface residues belonging to the *fa2-fa3* helices. Interestingly, PRFAR binding induces significant changes in the connections within communities *h2'* and *h3'* in the glutaminase domain, with the *hN12-hP10* connection in the Ω -loop having the highest betweenness in the new *h2'* community, and connections between the Ω -loop and the Gln binding site (including the conserved 49-PGVG sequence) being loosened with respect to that in the apo community *h2*. These observations reflect a decrease in correlation, upon PRFAR binding, for the 49-PGVG sequence and the HisH active site (Fig. 2, blue and orange boxes) and the concomitant increase of correlations for residues at the sideR of HisF, including the *fa2-fa3* helices (Fig. 2, magenta circle).

To analyze the interactions that generate the changes in the community network, we must focus on specific hydrophobic, ionic, and hydrogen-bonding interactions associated with the amino acid residues identified by the community network analysis. Such study indicates that the changes in the communication pathways induced by PRFAR binding are triggered by alteration of hydrophobic and ionic interactions at the sideR of the IGPS complex. Fig. 6 shows the time-dependent distances between hydrophobic residues at the bottom of the HisF (β/α)₈ barrel. There are clear differences between the apo and PRFAR bound

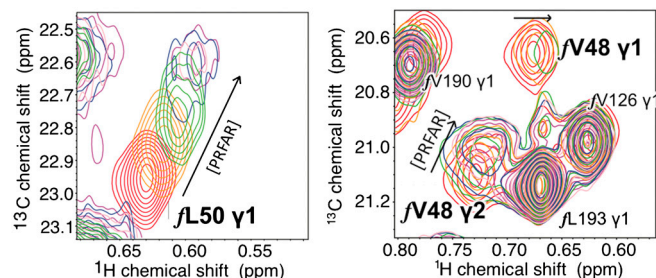


Fig. 7. The ^1H - ^{13}C HSQC spectral overlays of the titration of PRFAR into IGPS (0.89 mM). Spectra were collected for 0 mM (red), 0.56 mM (yellow), 0.74 mM (green), 1.81 mM (blue), and 3.47 mM (purple). Large chemical shift perturbations and significant line broadening are observed for V48 and L50 of HisF.

complexes. The residues primarily affected are those in the *fβ2* strand that line the ammonia tunnel. The residues closest to the effector binding site, including *fL50*, *fI52*, and *fF23*, become more compact upon PRFAR binding, forming a hydrophobic cluster. The *fL50-fI52-fF23* cluster formation is primarily driven by interactions with the π -system of the imidazolecarboxamide group of PRFAR (Fig. 6). The distance between *fV48* and *fL50* residues elongates considerably in the PRFAR bound complex while both *fL50-fF23* and *fL50-fI52* shorten upon PRFAR binding.

Our NMR titration experiments (Fig. 7) show prominent changes, induced by PRFAR binding, in the ^1H and ^{13}C chemical shifts for the methyl groups of key residues in the *fβ2* strand. Fig. 7 shows that PRFAR induces a large change in both the chemical shifts and the line widths for the resonances of *fL50*

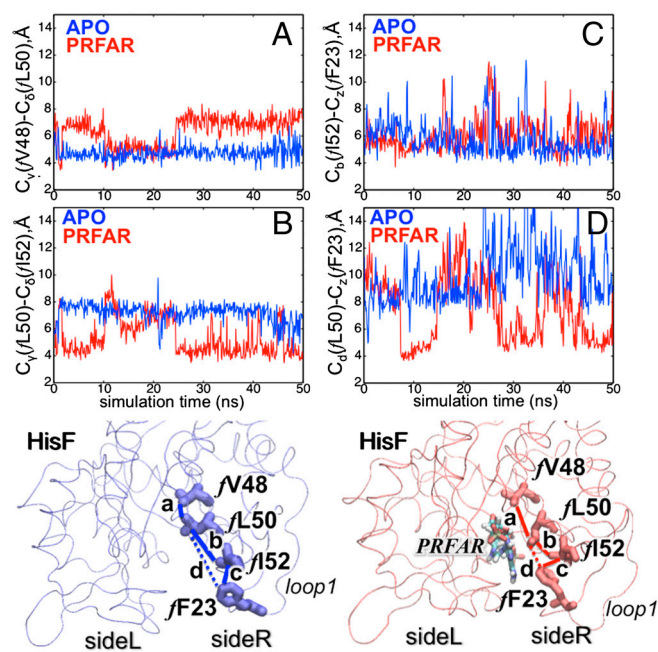


Fig. 6. Changes in hydrophobic interactions induced by PRFAR in the HisF β/α ₈ barrel. (A–D) show specific interaction distances as a function of simulation time for apo (blue) and PRFAR (red), for the hydrophobic residues shown in the bottom box. All highlighted residues but *fF23* (in loop1) belong to the *fβ2* strand, making a direct link between loop 1 and the ammonia tunnel. The representative MD snapshot ($t = 40$ ns) shown in the bottom box includes only HisF for simplicity. Interresidue distances (a, b, c, and d) are indicated in analogy to time-dependent distances reported in (A–D).

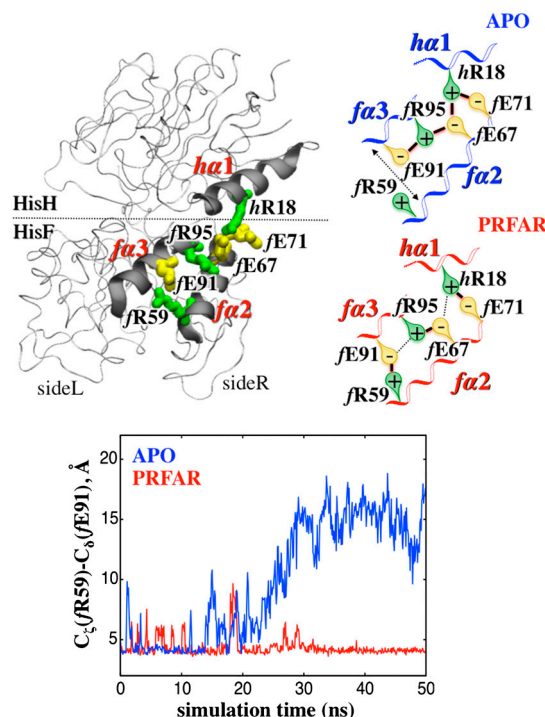


Fig. 8. (Top) Changes in the network of salt-bridges at sideR of IGPS, induced by PRFAR binding. (Left) The *fa2*, *fa3*, and *ha1* helices are highlighted in gray, with positively and negatively charged residues in green and yellow sticks, respectively. (Right) Schematic representations of ionic interactions in apo (blue) and PRFAR bound (red) complexes with the most stable ion-pair interactions highlighted with solid black lines. (Bottom) Thermal fluctuations of the C_α - C_α distance between the charged side chains of *fR59* and *fE91*, respectively, along a representative 50 ns MD simulation of both apo (blue) and PRFAR bound (red) IGPS complexes.

and fV48, indicating a significant change in the local chemical environment. This result is consistent with the conformational changes due to formation of the fL50-fI52-fF23 cluster. The disruption of the fV48-fL50 interaction on the ns timescale may represent the initial loosening of the protein core that precedes the wholesale enhancement of ms motions observed by solution NMR. Moreover, the alteration of hydrophobic interactions in the f β 2 strand, upon PRFAR binding, can also affect the opening of the ammonia tunnel gate favoring the NH₃ transfer from the glutaminase subunit to the bottom of the (β/α)₈ barrel. More relevantly, formation of the fL50-fI52-fF23 cluster in the f β 2 region also involves extension of the β -hairpin in loop1 and rearrangement of the adjacent secondary structures elements, such as the f α 1 and the f α 2 helices. Some of these rearrangements involve ionic interactions at the HisF/HisH interface, consistent with previous studies (12, 16).

Our MD simulations show that upon binding of PRFAR there is an alteration of the network of salt-bridges belonging to the ha1, fa2, and fa3 helices. As shown in Fig. 8, these α -helices are located in the sideR of the IGPS complex, with charged side chains placed at the protein surface and fa2 packing against the hydrophobic f β 2 region. As with hydrophobic interactions, these Coulombic contacts span over the PRFAR and Gln binding sites.

MD simulations indicate that the network of salt bridges formed by residues fR59, fE67, fE71, fE91, fR95, and hR18 is significantly altered by PRFAR binding (Fig. 8). In particular, the fR59-fE91 salt bridge located at the bottom of HisF breaks in less than 20 ns in the apo complex while remaining stable for the entire simulation (50 ns) when PRFAR is bound (Fig. 8). Due to the weak fR59-fE91 interaction in the apo complex, fa2 and fa3 helices can be disrupted by thermal fluctuations, indicating a certain degree of flexibility of this region, adjacent to loop1. In the presence of PRFAR, instead, fa2 and fa3 helices are less flexible and their motion is more tightly coupled, as indicated by the increased correlations of these residues in the PRFAR bound complex (Fig. 3). Considering the proximity of loop1 to the f β 2 and that of f β 2 to fa2, the formation of the hydrophobic cluster fL50-fI52-fF23 and the extension of the small two-stranded β -sheet in loop1 (fV17-fG30 H-bond formation, *SI Text*) should favor a conformation of HisF with fa2 near fa3, promoting the stabilization of the fR59-fE91 salt bridge.

The alteration of the fR59-fE91 interaction, due to PRFAR binding, propagates through the adjacent fE91-fR95 and fR95-fE67 ion-pairs and in turn through the hR18-fE67 and the hR18-fE71 salt bridges at the HisH-HisF interface. Breaking of the fR59-fE91 salt bridge in the apo complex stabilizes interactions between charged side chains with alternating sign, composed by fE91-fR95-fE67-hR18-fE71 (Fig. 8), which remains stable for >35 ns along the MD simulation. As shown in Fig. 8, this chain of ionic interactions is altered upon PRFAR binding, with a stable fR59-fE91 salt bridge limiting the interaction between fE91 and fR95 residues and stabilizing specific ion pairs such as fR95-fE67 and hR18-fE71. This rearrangement of ionic interactions upon PRFAR binding induces the increased communication flow between sideR communities f3' and h2' by coupling more tightly the motions of α -helices ha1, fa2, and fa3. Overall, the conformational rearrangement of helices ha1, fa2, fa3, and Ω -loop, upon effector binding, do not damage the structural stability of the IGPS heterodimer as indicated by preservation of clustering of interface residues in the f3 and f3' communities. However, the reorganization of ionic interactions at the HisF-HisH interface is coupled to the breathing motion of the IGPS complex, with the PRFAR bound complex assuming a conformation that facilitates motions at the heterodimer interface. The breathing motion present in the PRFAR bound complex exposes more frequently the interface region to the solvent water molecules, which reduces the correlations in motions at the bottom of the HisH domain, including correlations between motions

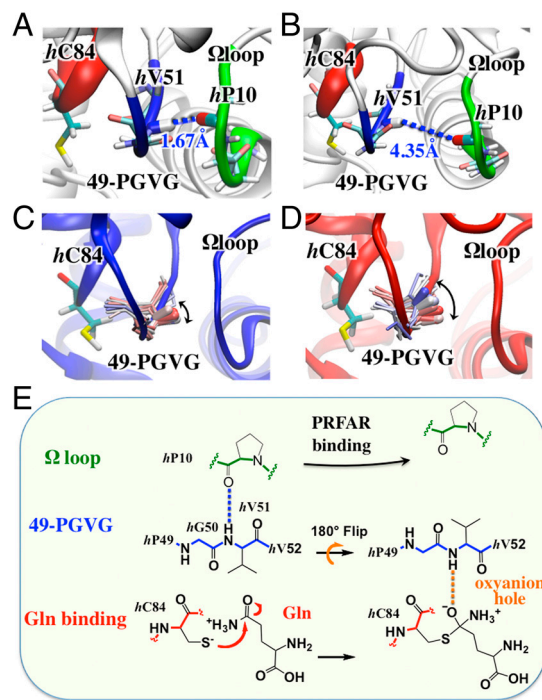


Fig. 9. (A) and (B) MD representative snapshots (at $t = 50$ ns) of apo (A) and PRFAR bound (B) IGPS complexes, showing H-bonds between hP10 (Ω -loop, in green) and hV51 (oxyanion strand, in blue). (C) and (D) Partial reorientation of the oxyanion strand triggered by PRFAR binding. Thermal fluctuations of the 49-PGVG backbone are shown (sticks colored from red to blue for 0–50 ns) highlighting the NH amide group of hV51 (larger sticks). Apo (C) and PRFAR bound (D) average structures (colored cartoons) and Gln binding site (hC84, colored sticks) are also shown. (E) Schematic representation of the oxyanion hole formation during the Gln hydrolysis reaction.

of the Ω -loop and the oxyanion strand at sideR, and motions between h β 7 and h β 8 β -sheets at the sideL of HisH. As we describe below, these results have significant implications on the allosteric mechanism of IGPS.

Allosteric Mechanism of IGPS. A recurring feature among enzymes with the catalytic triad motif Cys-His-Glu, such as guanine monophosphate synthetase (GMPS) (17), carbamoyl phosphate synthetase (CPS) (18), and anthranilate synthase (AS) (19), is the so-called “oxyanion strand,” a β -strand adjacent to the Gln binding site (the nucleophilic cysteine, hC84, of the catalytic triad). The importance of the oxyanion strand is due to its crucial role in forming an “oxyanion hole” that stabilizes a tetrahedral intermediate during the Gln hydrolysis, assisting the enzymatic activity. The oxyanion strand in the crystal structures of GMPS, CPS, and AS (17–19) has been resolved in its “active conformation.” The orientation of the backbone of amino acid residues hG50 and hV51 (*T. maritima* IGPS numbering) leaves the NH group of hV51 pointing into the catalytic site, thus becoming part of the oxyanion hole. In IGPS from *T. maritima*, the oxyanion strand is constituted by the highly conserved sequence hP49-hG50-hV51-hG52 (49-PGVG) in the HisH subunit. The crystal structures of IGPS from *T. maritima* (5) indicates that the PGVG β -strand has an improper (inactive) conformation, with the carbonyl group of hG50 pointing into the Gln binding site (Fig. 9). Consequently, the NH group of hV51 is not able to form the hydrogen bond with the negatively charged oxygen of the tetrahedral intermediate and thus is not in optimal position for taking part in the oxyanion hole. This “inactive conformation” of the PGVG backbone has also been observed in the IGPS from *Saccharomyces cerevisiae* (2). These observations led to the hypothesis that a conformational change of the IGPS oxyanion strand, associated with a

180° rotation of the 49-PGVG backbone, is induced by effector binding (Fig. 9). This hypothesis is recently supported by our ^1H , ^{15}N NMR analysis of the 49-PGVG sequence, revealing conformational exchange motions of *hG50* on the μs -ms timescale, upon titration with PRFAR (14).

In the inactive conformation of the 49-PGVG backbone, the orientation of the NH group of *hV51* is stabilized by a hydrogen bond with the carbonyl oxygen of *hP10* in the Ω -loop. Disruption of this interaction is, therefore, a prerequisite for allowing the conformational change associated with the flip of the oxyanion strand backbone. All MD simulations indicate that this hydrogen bond is stable during the entire 100 ns trajectories (*SI Text*) in the apo IGPS complex. On the other hand, this important hydrogen bond breaks within 50 ns upon PRFAR binding, as indicated by all four independent MD simulations of the PRFAR-bound IGPS complex (*SI Text*). The *hV51-hP10* hydrogen bond is responsible for connecting the Ω -loop with the HisH active site, including the 49-PGVG strand and the catalytic triad. The community network analysis reveals that breaking of the *hV51-hP10* hydrogen bond is coupled with changes in the interface communications at the IGPS sideR and between the Ω -loop and the glutaminase active site. Fig. 9 *A* and *B* show the difference in the *hV51-hP10* interaction as observed in a representative MD simulation of the apo and PRFAR bound IGPS complexes. The *hV51-hP10* hydrogen bond begins to weaken after 12 ns and it completely breaks after 27 ns. Interestingly, we observed that the 49-PGVG strand partially rotates in the ns time scale upon breaking of the *hV51-hP10* H-bond (Fig. 9 *C* and *D*). The displacement of the 49-PGVG sequence from the Ω loop confers more rotational degrees of freedom to the oxyanion strand, enabling it to rotate with respect to glutaminase binding site. The complete rotation of the oxyanion strand requires overcoming a large activation barrier $>20 \text{ kcal mol}^{-1}$ on the ms time scale, as suggested by NMR solution experiments (14).

In summary, PRFAR binding modifies the protein network of IGPS, affecting the two sides of the heterodimer complex in two different ways. At sideL, entropically driven PRFAR binding decreases the degree of correlation in the motion of residues, inducing the splitting of the large HisF community *f2* and favoring the transfer of entropy that increases the overall breathing motion of the IGPS complex. At sideR, PRFAR binding affects hydrophobic interactions at the bottom of the $(\beta/\alpha)_8$ barrel, inducing conformational changes that propagate through ionic interactions of α -helices *f* α 2, *f* α 3, and *h* α 1 and, finally, affecting hydrogen-bonding interactions between the Ω -loop and the oxyanion strand adjacent to the HisH active site.

Conclusions

The allosteric pathway of the V-type IGPS heterodimer has been elucidated by combining NMR, MD simulations, and the correlation analysis of protein motions based network theory methods. PRFAR binding changes the structure of interactions that are essential for the allosteric mechanism. Dynamical rearrangements involve the flexible loop1 and the hydrophobic region (*f* β 2) at the HisF allosteric site, ionic interactions in the interdomain region next to the loop1 (sideR), including residues in the *f* α 2 and the *f* α 3 helices in HisF, and in the helix *h* α 1 in HisH, as well as hydrogen bonding between the Ω -loop and the conserved 49-PGVG sequence (oxyanion strand) adjacent to the glutaminase active site. The entropically driven effector binding enhances the overall HisF-HisH interdomain motion, with the sideR portion of the protein complex remaining more rigidly closed than the sideL part. The resulting protein motions partially promote rotation of the conserved oxyanion strand responsible for the inactive-to-active allosteric transition. Single amino acid residues, such as *fL50*, *fR59*, *fE91*, and *hP10*, play an important role in the allosteric mechanism and should be good candidates for unique mutagenesis studies. Therefore, the proposed allosteric pathway provides

useful information for the design of new molecules that might hamper glutaminase activity in plant or human pathogens by interrupting the allosteric signal propagation of IGPS.

Materials and Methods

Modeling and Molecular Dynamics Simulations. The computational structural models for apo and PRFAR bound IGPS complexes are based on the crystal structure of the bienzyme complex from *T. maritima* at 2.4 Å resolution (Protein Data Bank code 1GPW) (1, 4, 5). The HisH-HisF apo-complex having the C-terminal loop of the HisF unit resolved in a closed (active) conformation (i.e., chains C and D) was chosen as starting structure. All crystal waters belonging to this protein complex were kept. The active-site mutation (N11D) present in the original crystal structure was mutated back to its wild-type form. The positions of the phosphate groups in the *T. maritima* IGPS crystal structure and in the PRFAR bound complex from yeast (PDB code 1OX5) (11) match nicely after superposition of the two X-ray structures. Therefore, the initial pose of the PRFAR effector in the HisF cavity was chosen according to its position in the yeast structure (*SI Text*). MD simulations of the apo and PRFAR bound structures of IGPS are based on the AMBER-ff99SB (20) force field for the protein and Generalized Amber Force Field (21) for the PRFAR ligand (see *SI Text*), as implemented in the NAMD2 software package (22). Four independent MD simulations have been performed for each complex (apo and PRFAR-bound) for a total simulation time of 0.8 μs . Further details of pre-equilibration procedure and MD production runs are described in *SI Text*.

Correlation Analysis of Protein Motions. The cross-correlation analysis is an established method used for quantifying correlated motion in proteins from MD simulations (23, 24). According to this approach, correlations between residues motions are obtained from MD trajectories by calculating the Pearson correlation coefficient

$$r[\mathbf{x}_i, \mathbf{x}_j] = \langle \mathbf{x}_i \cdot \mathbf{x}_j \rangle / (\langle \mathbf{x}_i^2 \rangle \langle \mathbf{x}_j^2 \rangle)^{1/2}, \quad [1]$$

where $\mathbf{x}_{i,j}$ are the fluctuation vectors of atoms *i* and *j*, defined as the deviation from the mean; i.e., $\mathbf{x} = \mathbf{v} - \langle \mathbf{v} \rangle$, of the atomic position vector \mathbf{v} , with $\langle \rangle$ denoting the ensemble average. Estimate of correlations from the Pearson coefficient is not independent from the relative orientation of the atomic fluctuations and is limited to the detection of linear correlations only, with relevant artifacts in the quantification of motion correlations of amino acid residues (10). Lange, et al. (10) recently proposed a method to measure generalized correlations in biological systems, which overcomes the problems related to the use of the Pearson coefficient, as coefficient of nondetermination. The joint probability distribution $p(\mathbf{x})$ of (*N*) random variables \mathbf{x}_i , with $i = 1, \dots, N$, equals the product of their marginal distribution, $p_i(\mathbf{x}_i)$, if and only if the components \mathbf{x}_i are independent; i.e., uncorrelated. Therefore, the ensemble-averaged deviation from the uncorrelated distribution is given by the mutual information (MI), defined as

$$I[\mathbf{x}_1, \dots, \mathbf{x}_N] = \int p(\mathbf{x}) \ln \frac{p(\mathbf{x})}{\prod_i^N p_i(\mathbf{x}_i)} d\mathbf{x}. \quad [2]$$

MI vanishes in the case of independent random variables; i.e., it vanishes in the bivariate case of the fluctuation vectors \mathbf{x}_i and \mathbf{x}_j for fully uncorrelated motions of residues *i* and *j*. The above defined MI yields values in the range $[0, \infty)$ leading to a less intuitive interpretation of the correlations with respect to the Pearson coefficient, which values vary from -1 to 1 . For an easier interpretation of the MI, it is thus convenient to define a generalized correlation coefficient, r_{MI} , as the Pearson-like coefficient derived from the mutual information. Considering colinear Gaussian distributions of unit variance ($d = 3$), the generalized correlation coefficient r_{MI} is defined as

$$r_{\text{MI}}[\mathbf{x}_i, \mathbf{x}_j] = \{1 - \exp(-2I[\mathbf{x}_i, \mathbf{x}_j]/d)\}^{-1/2}. \quad [3]$$

The coefficient r_{MI} is zero for fully uncorrelated variables and assumes values up to 1 for fully correlated variables.

MI is closely related to the Shannon entropy (13), $H[\mathbf{x}]$, which is defined as the expected value of the information content of a discrete random variable \mathbf{x} , having probability mass function $p(\mathbf{x})$,

$$H[\mathbf{x}] = - \int p(\mathbf{x}) \ln p(\mathbf{x}) d\mathbf{x}. \quad [4]$$

The mutual information in the case of bivariate random variables can be calculated as

$$I[\mathbf{x}_i, \mathbf{x}_j] = H[\mathbf{x}_i] + H[\mathbf{x}_j] - H[\mathbf{x}_i, \mathbf{x}_j], \quad [5]$$

where the marginal, $H[\mathbf{x}_i]$ and $H[\mathbf{x}_j]$, and the joint, $H[\mathbf{x}_i, \mathbf{x}_j]$, entropies could be estimated with different methods (25).

In this work, the information content $H[\mathbf{x}]$ is calculated using the atomic positions fluctuations derived from the MD simulations by means of the k -nearest neighbor distances algorithm (25) as implemented in GROMACS (10, 26).

The generalized correlation coefficient, r_{MI} , for each pair of amino acid residues of the IGPS complexes are, then, obtained according to Eq. 5 considering the motion of C α atoms, averaging for each independent 100 ns MD simulation over six 50 ns windows, and then averaging over four independent MD simulations for both apo and PRFAR-bound complexes (for a total simulation time of 0.8 μ s).

Community Network Analysis. The IGPS network is defined as the set of amino acid residues; i.e., the nodes (or vertices), connected by edges whose length is weighted using the generalized correlation coefficient r_{MI} . For a fair comparison between the apo and the PRFAR bound IGPS networks, PRFAR is not considered to be a node, with apo and PRFAR protein networks having the same number of nodes. Two nodes are considered connected if any heavy atoms from the two residues are within 5.0 Å of each other (distance cutoff) for at least 75% of the frames (percentage cutoff) analyzed (*SI Text*).

The edge lengths in the network are obtained using the generalized correlation coefficients associated to the mutual information of each pair of residues. We first obtained the matrix of direct communication 'distances' $\mathbf{w}_{(ij)}^{(0)} = -\log[r_{MI}(\mathbf{x}_i, \mathbf{x}_j)]$ between all pairs of residues (ij). The matrix of shortest distances $\mathbf{w}_{(ij)}^{(N)}$ was determined by using the Floyd-Warshall algorithm (27), considering direct distances $\mathbf{w}_{(ij)}^{(0)}$ as well as up to N possible intermediate residues mediating indirect communication pathways (where N is the number of residues in the system). The matrix of shortest pathways $\mathbf{w}_{(ij)}^{(N)}$ provides the edge-betweenness matrix, with elements $\mathbf{b}_{(ij)}$ defined as the number of shortest paths that include pair (ij) as one of its communication segments. Therefore, the edge betweenness matrix $\mathbf{b}_{(ij)}$ is a measure of 'traffic' passing through edges within the network and can be used for partitioning the communication network into communities. We use the Girvan-Newmann algorithm (27, 28), maximizing the modularity measure (Q) (29) to optimize the quality of the community structure (*SI Text*).

NMR Titration. HisF and HisH (with a C-terminal His₆ tag) codon optimized (Genscript) *T. maritima* genes for *Escherichia coli* were individual cloned into a pET43.1b vector between the NdeI-HindIII and NdeI-XhoI sites, respectively. ²H, ¹⁵N, and ¹³C methyl groups of Ile, Leu, and Val residues (ILV) HisF-²H HisH isotope enriched samples were prepared by transforming the HisF and HisH plasmids into separate *E. coli* BL21(DE3) cells, which were grown in 100% D₂O M9 minimal media with ¹⁵NH₄Cl as the nitrogen source and the isoleucine, leucine, and valine precursors 2-ketobutyric acid-4-¹³C (Isotec) and 2-keto-3-(methyl-D3)-butyric acid-4-¹³C (Isotec) for HisF. The cells were grown to an OD₆₀₀ of 0.9–1.0 at 37 °C and were induced with a final concentration of 1 mM IPTG. To achieve δ -¹³CH₃ isoleucine and leucine as well as γ -¹³CH₃ valine, 60 mg 2-ketobutyric acid-4-¹³C and 100 mg 2-keto-3-(methyl-D3)-butyric acid-4-¹³C per liter of M9 were added 1 h prior to induction (30). Carbon-3 bound hydrogens for 2-ketobutyric acid-4-¹³C and 2-keto-3-(methyl-D3)-butyric acid-4-¹³C were exchanged with deuterium by dissolving the compounds in 30 mL and 20 mL of D₂O, respectively, with one KOH pellet. The high pH mixture was then incubated overnight at 37 °C to facilitate exchange. Before addition to the culture, the pH was lowered to 7.0. The cells were harvested after 8 h of induction. HisF and HisH cells were resuspended in 5 mL 100 mM potassium phosphate, 100 mM Tris-HCl (pH 9.0), and 1.0 M NaCl per gram of cell. The resuspended cells were colysed by sonication with a 1.5 times growth culture volume excess of deuterated HisH. Cell waste was removed through centrifugation and the supernatant was mixed with Ni/NTA resin. The Ni/NTA resin was washed with 100 mM potassium phosphate (pH 7.5), 10 mM Tris-Base (pH 9.5), and 10 mM imidazole. The HisF-HisH complex was eluted in 10 mM N-cyclohexyl-3-aminopropanesulfonic acid (CAPS), 10 mM Tris-Base (pH 9.5) and 250 mM imidazole and dialyzed against 4.0 L of 10 mM Hepes (pH 7.3), 10 mM KCl, 1.0 mM EDTA and 1.0 mM DTT. PRFAR was dissolved in 10 mM Hepes (pH 7.3), 10 mM KCl, 1.0 mM EDTA, 1.0 mM DTT and the pH was adjusted to 7.3. ¹H-¹³C spectra were collected for 0, 0.19, 0.56, 0.74, 1.81, and 3.47 mM PRFAR with a starting IGPS concentration of 0.89 mM.

ACKNOWLEDGMENTS. We thank J. Lipchock for acquisition of the NMR titration data, and O. Lange and V. Limongelli for very useful discussions. V.S.B. acknowledges supercomputer time from the National Energy Research Scientific Computing Center (NERSC) and support from the National Institutes of Health Grant 1R01-GM-084267-01 for the developments of methods and models implemented in this study. G.M. acknowledges support from National Institutes of Health training Grant T32GM008283.

- Klem TJ, Chen Y, Davisson VJ (2001) Subunit interactions and glutamine utilization by *Escherichia coli* imidazole glycerol phosphate synthase. *J Bacteriol* 183:989–996.
- Chaudhuri BN, et al. (2001) Crystal structure of imidazole glycerol phosphate synthase: a tunnel through a (beta/alpha)₈ barrel joins two active sites. *Structure* 9:987–997.
- Omi R, et al. (2002) Structure of imidazole glycerol phosphate synthase from *Thermus thermophilus* HB8: open-closed conformational change and ammonia tunneling. *J Biochem* 132:759–765.
- Klem TJ, Davisson VJ (1993) Imidazole glycerol phosphate synthase: the glutamine amidotransferase in histidine biosynthesis. *Biochemistry* 32:5177–5186.
- Douangamath A, et al. (2002) Structural evidence for ammonia tunneling across the (beta/alpha)₈ barrel of the imidazole glycerol phosphate synthase holoenzyme complex. *Structure* 10:185–193.
- Breitbach K, Kohler J, Steinmetz I (2008) Induction of protective immunity against *Burkholderia pseudomallei* using attenuated mutants with defects in the intracellular life cycle. *Trans R Soc Trop Med Hyg* 102:S89–S94.
- Gomez MJ, Neyfakh AA (2006) Genes involved in intrinsic antibiotic resistance of *Acinetobacter baylyi*. *Antimicrob Agents Chemother* 50:3562–3567.
- Myers RS, Jensen JR, Deras ML, Smith JL, Davisson VJ (2003) Substrate-induced changes in the ammonia channel for imidazole glycerol phosphate synthase. *Biochemistry* 42:7013–7022.
- Sethi A, Eargle J, Black AA, Luthey-Schulten Z (2009) Dynamical networks in tRNA: protein complexes. *Proc Natl Acad Sci USA* 106:6620–6625.
- Lange OF, Grubmüller H (2006) Generalized correlation for biomolecular dynamics. *Proteins* 62:1053–1061.
- Chaudhuri BN, Lange SC, Myers RS, Davisson VJ, Smith JL (2003) Toward understanding the mechanism of the complex cyclization reaction catalyzed by imidazole glycerol-phosphate synthase: crystal structures of a ternary complex and the free enzyme. *Biochemistry* 42:7003–7012.
- Amaro RE, Sethi A, Myers RS, Davisson VJ, Luthey-Schulten ZA (2007) A network of conserved interactions regulates the allosteric signal in a glutamine amidotransferase. *Biochemistry* 46:2156–2173.
- Shannon CE (1948) A mathematical theory of communication. *AT&T Tech J* 27:623–656.
- Lipchock JM, Loria JP (2010) Nanometer propagation of millisecond motions in V-type allostery. *Structure* 18:1596–1607.
- Myers RS, Amaro RE, Luthey-Schulten ZA, Davisson VJ (2005) Reaction coupling through interdomain contacts in imidazole glycerol phosphate synthase. *Biochemistry* 44:11974–11985.
- Amaro RE, Myers RS, Davisson VJ, Luthey-Schulten ZA (2005) Structural elements in IGPS synthase exclude water to optimize ammonia transfer. *Biophys J* 89:475–487.
- Tesmer JG, Klem TJ, Deras ML, Davisson VJ, Smith JL (1996) The crystal structure of GMP synthetase reveals a novel catalytic triad and is a structural paradigm for two enzyme families. *Nat Struct Biol* 3:74–86.
- Thoden JB, Holden HM, Wesenberg G, Raushel FM, Rayment I (1997) Structure of carbamoyl phosphate synthetase: a journey of 96 angstrom from substrate to product. *Biochemistry* 36:6305–6316.
- Knochel T, et al. (1999) The crystal structure of anthranilate synthase from *Sulfolobus solfataricus*: Functional implications. *Proc Natl Acad Sci USA* 96:9479–9484.
- Case DA, et al. (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26:1668–1688.
- Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. *J Comput Chem* 25:1157–1174.
- Phillips JC, et al. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802.
- Hünenberger PH, Mark AE, Vangunsteren WF (1995) Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular-dynamics simulations. *J Mol Biol* 252:492–503.
- Ichihye T, Karplus M (1991) Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins* 11:205–217.
- Kraskov A, Stogbauer H, Grassberger P (2004) Estimating mutual information. *Phys Rev E* 69:066138.
- Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4:435–447.
- Floyd RW (1962) Algorithm-97—Shortest Path. *Commun ACM* 5:345–345.
- Girvan M, Newman MEJ (2002) Community structure in social and biological networks. *Proc Natl Acad Sci USA* 99:7821–7826.
- Newman MEJ (2006) Modularity and community structure in networks. *Proc Natl Acad Sci USA* 103:8577–8582.
- Goto NK, Gardner KH, Mueller GA, Willis RC, Kay LE (1999) A robust and cost-effective method for the production of Val, Leu, Ile (delta 1) methyl-protonated N-15-, C-13-, H-2-labeled proteins. *J Biomol Nmr* 13:369–374.