

Distinct allosteric pathways in imidazole glycerol phosphate synthase from yeast and bacteria

Federica Maschietto,¹ Aria Gheeraert,² Andrea Piazzini,³ Victor S. Batista,^{1,*} and Ivan Rivalta^{2,3,*}

¹Department of Chemistry, Yale University, New Haven, Connecticut; ²Université de Lyon, CNRS, Institut de Chimie de Lyon, École Normale Supérieure de Lyon, Lyon Cedex 07, France; and ³Dipartimento di Chimica Industriale “Toso Montanari”, Alma Mater Studiorum, Università di Bologna, Bologna, Italia

ABSTRACT Understanding the relationship between protein structures and their function is still an open question that becomes very challenging when allostery plays an important functional role. Allosteric proteins, in fact, exploit different ranges of motions (from sidechain local fluctuations to long-range collective motions) to effectively couple distant binding sites, and of particular interest is whether allosteric proteins of the same families with similar functions and structures also necessarily share the same allosteric mechanisms. Here, we compared the early dynamics initiating the allosteric communication of a prototypical allosteric enzyme from two different organisms, i.e., the imidazole glycerol phosphate synthase (IGPS) enzymes from the thermophilic bacteria and the yeast, working at high and room temperatures, respectively. By combining molecular dynamics simulations and network models derived from graph theory, we found rather distinct early allosteric dynamics in the IGPS from the two organisms, involving significantly different allosteric pathways in terms of both local and collective motions. Given the successful prediction of key allosteric residues in the bacterial IGPS, whose mutation disrupts its allosteric communication, the outcome of this study paves the way for future experimental studies on the yeast IGPS that could foster therapeutic applications by exploiting the control of IGPS enzyme allostery.

SIGNIFICANCE Allosteric regulation is widely present in macromolecules and is essential to coordinate biochemical information transfer between spatially distant sites. Despite the growing interest dedicated to uncovering the mechanism of allosteric processes, the question of how allosteric enzymes from different evolutionary paths achieve the same catalytic function remains elusive. We examine the allosteric pathways of the imidazole glycerol phosphate synthase (IGPS) enzymes from yeast and thermophilic bacteria through the lens of molecular dynamics simulations and graph-theory-based network models. We find that protein-specific cooperative interactions between local and collective modes accomplish the same function of activating the catalytic site upon effector binding to the allosteric site, allowing the two enzymes to optimally function in their (different) natural environments.

INTRODUCTION

Allostery is an essential regulatory process of biological macromolecules of great interest for a wide range of applications, including drug discovery and gene-editing technologies (1–4). Allosteric mechanisms typically transmit the effect of binding of a ligand effector to a distant site, often responsible for catalytic activity (5). Targeting the signal transduction mechanism between the allosteric and catalytic sites can lead to suppression of substrate turnover at the

active site, opening an opportunity for protein engineering or development of non-competitive small molecule inhibitors. An advantage of allosteric drugs is that they selectively tune responses in tissues where the endogenous agonists exert their physiological effects and only when the endogenous agonists are present (6). Such spatial and temporal selectivity cannot be achieved with traditional orthosteric agonists since those modify the receptor function continuously as long as they are present. Another important advantage is the intrinsic safety in overdosage since, once the allosteric sites are occupied, no further allosteric effect can be produced even with excessive doses (5,7). An outstanding challenge, however, is the development of fundamental understanding of allosteric pathways in

Submitted May 10, 2021, and accepted for publication November 29, 2021.

*Correspondence: victor.batista@yale.edu or i.rivalta@unibo.it

Editor: Alexandr Kornev

<https://doi.org/10.1016/j.bpj.2021.11.2888>

© 2021 Biophysical Society.

proteins (1,8–11). In fact, an allosteric mechanism encompasses all steps that are involved in the signal transduction extending from the effector to the active site. These steps include effector binding, allosteric communication (via local contacts and collective motions; i.e., the allosteric pathways) triggering alterations (usually associated to conformational changes) of the active site. It has been proposed that allostery may be an intrinsic property of virtually all proteins (12); however, the extent of conservation of allosteric mechanisms or absence of it across a protein family remains an open question (13). In fact, the similarity of protein structures does not necessarily imply a common function (proteins with different functions can share a common structural framework while the same function can be performed by proteins with different folds), suggesting that the structure/function relationship can be quite complex in terms of allostery (12).

On one side, there are examples of proteins with similar functions and structures retaining similar allosteric pathways that have been reported (14–17), pointing out the role of conserved network of residues in allostery. On the other hand, various studies have reported differences in the allosteric communication between protein homologs. For instance, the structural study of three bacterial chemotaxis protein Y orthologs showed divergent allosteric responses across the protein family, with allosteric signals found to be globally propagated in different, system-dependent, ways (18). Moreover, the characterization of three homologous of the HIV-1 envelope spike allostery has suggested that, despite the common modular structure of the allosteric network that remains highly conserved, the shortest path for communication between distal regions is sensitive to differences in the primary sequences of the individual proteins (19).

Therefore, the assumption that proteins with similar structures would have similar allosteric pathways is not always true, since allosteric communication in protein orthologs is often system specific (18,20,21). So, the extent to which allosteric pathways are conserved among protein orthologs remains an open question (22).

The intrinsic complexity of the question of conservation of allosteric pathways is due to the fact that differences in the allosteric communication between protein homologs can occur at different levels of the allosteric signaling pathways, i.e., involving both changes in local contacts and/or collective motions, suggesting that a detailed knowledge of these communication pathways is required. Here, in response to reviewers, we address this question for the allosteric pathways of imidazole glycerol phosphate synthase (IGPS) enzymes from two different organisms, bacteria (*Thermotoga maritima* [Tm]) and yeast (*Saccharomyces cerevisiae* [Sc]). IGPS enzymes are ideal for our analysis since they are prototypical systems for the study of allostery and have already attracted significant interest as targets for therapeutic applications (23–30). Our study is focused on

understanding how these two allosteric enzymes with different evolutionary paths achieve the same allosteric function despite the significant differences in their primary sequences, and secondary structures. As a consequence of their structural analogy, IGPS enzymes from yeast and bacteria feature the same effector-binding site (31,32) and glutaminase active site (with analogous inactive/active conformations) (30). We focus on the characterization of their allosteric pathways (those of Sc-IGPS being unknown), exploring both local contacts and collective motion contributions to analyze whether or not the two enzymes have the same allosteric mechanism. We find that the early dynamics that initiate allosteric communication are rather different for the two enzymes, resulting in distinct allosteric pathways tailored for activity in the different natural environments of the two enzymes. Thermophiles exhibit robust functionality at high temperatures, while *Saccharomyces* function at room temperature. Their early allosteric dynamics involve differences in both collective motions and inter-residue interactions, which are likely due to the different adaptations of the enzymes to their native conditions.

Structural features of IGPS enzymes from *Thermophiles* and *Saccharomyces*

We begin by summarizing the similarities and structural differences between the two IGPS enzymes. In bacteria, IGPS is a tightly associated heterodimer complex formed by the glutaminase subunit HisH and the cyclase HisF (red and salmon, respectively, in Fig. 1 A) (33). In yeast Sc-IGPS, the two subunits are fused into a single polypeptide chain, His7 (green, in Fig. 1 A) with the two functional domains linked by a short polypeptide (i.e., the connector, circled in Fig. 1 B) (31). The aligned complexes share the same fold, as shown in Fig. 1 A, with a sequence similarity of 52% and 63% for HisH and HisF, and an RMSD of C-alpha carbon atoms of 1.93 and 2.03 Å, respectively (see sequence alignment in Table S2). Throughout this paper, we refer to secondary structural elements by increasing numbering and labeling the residues corresponding to HisH and HisF with prefixes *h* and *f*, respectively, following the standard Tm-IGPS nomenclature (11). The full topography of secondary structural elements of yeast IGPS is reported in the supporting material (Table S1), and for Tm-IGPS is reported in reference (33).

The same two reactions are catalyzed by the two domains of both IGPS enzymes from thermophiles and yeast. In the glutaminase domain, glutamine (Gln) is hydrolyzed to glutamate, releasing ammonia that migrates (29,31–34) to the cyclase domain, where it reacts with the effector 5-[(5-phospho-1-deoxy-D-ribulos-1-ylimino)methylamino]-1-(5-phospho-beta-D-ribosyl)imidazole-4-carboxamide (PRFAR) to form imidazoleglycerol phosphate (ImGP), a precursor to histidine and 5'-(5-aminoimidazole-4-carboxamide) (AICAR),

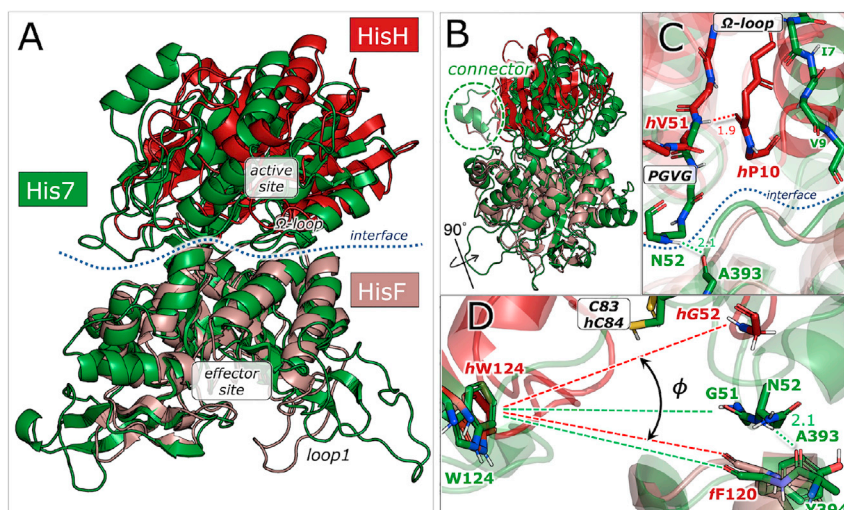


FIGURE 1 Molecular representation of IGPS from thermophile and yeast. (A) Front view of *Tm*-IGPS (red and salmon) as compared with *Sc*-IGPS from yeast *S. cerevisiae* (green). The structures are derived from the PDB models 1GPW and 1OX6 where the missing residues have been reconstructed (as described in the [supporting material](#)). The active site in the glutaminase domain and the effector site in the cyclase domain are more than 25 Å apart in both *Tm*-IGPS and *Sc*-IGPS. (B) Side view of aligned *Tm*-IGPS and *Sc*-IGPS structures, highlighting the position of the connector between the cyclase and glutaminase domains of His7. (C) Close-up of the glutaminase active site in *Tm*- and *Sc*-IGPS (3ZR4 and 1OX5, respectively), showing structural differences next to the active site (Gln substrate not shown). Loop $\beta 3\alpha 2$ with hV51 is tightly bound to P10 in *Tm*-IGPS, but shifted toward the cyclase domain in His7. The $\beta 3\alpha 2$ loop (also known as the PGVG strand) is highly conserved in all IGPS enzymes and is thought to stabilize the oxyanion intermediate formed during the catalytic reaction. (D) Close-up view of the interface between the cyclase and glutaminase domains in yeast and bacterial IGPS. In His7, the interface is closed, with the angle $\phi = 15^\circ$ between G51:C $\alpha \leftrightarrow$ W124:C $\gamma \leftrightarrow$ Y394:C γ , spanned by the green dotted lines, in the crystal structure (PDB: 1OX6) (34). In apo-*Tm* IGPS, however, the HisF:HisH interface is wide open. In the crystal structure (PDB: 1GPW) (33), the corresponding angle (hG52:C $\alpha \leftrightarrow$ hW123:C $\gamma \leftrightarrow$ fF120:C γ) $\phi = 29^\circ$ (between red dotted lines).

zymes and is thought to stabilize the oxyanion intermediate formed during the catalytic reaction. (D) Close-up view of the interface between the cyclase and glutaminase domains in yeast and bacterial IGPS. In His7, the interface is closed, with the angle $\phi = 15^\circ$ between G51:C $\alpha \leftrightarrow$ W124:C $\gamma \leftrightarrow$ Y394:C γ , spanned by the green dotted lines, in the crystal structure (PDB: 1OX6) (34). In apo-*Tm* IGPS, however, the HisF:HisH interface is wide open. In the crystal structure (PDB: 1GPW) (33), the corresponding angle (hG52:C $\alpha \leftrightarrow$ hW123:C $\gamma \leftrightarrow$ fF120:C γ) $\phi = 29^\circ$ (between red dotted lines).

used in the synthesis of purines. While Gln hydrolysis could occur in the absence of the effector, the reaction is accelerated 5000-fold upon PRFAR binding, classifying IGPS as a V-type allosteric enzyme. Recent studies of *Tm*-IGPS (28) have shown that Gln has a different affinity for the enzyme with or without effector, although the major increase in turnover (K_{cat}) is predominant over the change in substrate dissociation constant, K_m^{Gln} .

Experimental and computational studies on PRFAR-bound and PRFAR-free forms of IGPS enzymes have identified flexible parts of the protein with potential allosteric roles in the communication between the effector and catalytic sites (23,26–28,31,34,35). These previous studies have provided evidence of an unformed oxyanion hole as the basis for low glutaminase activity in the effector-free form of the enzyme (31,34,36,37). The term “oxyanion hole” derives from the presence of a negatively charged oxygen on the Gln, generated by the reaction of the cysteine sulfur in the active site and the Gln substrate. The hole generated by the amino acid residues surrounding the anion stabilizes the negative charge before a neutral environment is restored.

The highly conserved sequence in the $\beta 3\alpha 2$ loop of all IGPS enzymes, known as the PGVG (oxyanion) strand next to the glutaminase active site, hosts the charged intermediate. However, the crystal structures of IGPS from both yeast (31,34) and bacteria (33) suggest that the PGVG β strand has an improper conformation in the apo enzymes, with the NH group of hV51/V50 pointing out from the Gln-binding site. Therefore, a 180° turn of the whole oxyanion strand is necessary to stabilize the tetrahedral intermediate and to make the glutaminase enzyme catalytically active. Earlier studies are consistent with the formation of the oxyanion hole as the endpoint of the allosteric mechanism in IGPS enzymes (23,28,30,31,36).

Unlike allostery in bacterial IGPS, the allosteric pathway in IGPS from yeast remains uncertain. The comparative structural analysis of the two enzymes suggests that different allosteric mechanisms might operate in the two systems. For example, the PGVG strand in *Tm*-IGPS is more distant from the cyclase domain compared with *Sc*-IGPS. Further, the hV51-hP10 hydrogen bond (H-bond) that connects the PGVG oxyanion strand with the neighboring Ω -loop has been shown to be crucial in the allosteric mechanism of *Tm*-IGPS (23,38), although it is absent in *Sc*-IGPS (see Fig. 1 C) (31,34).

The cyclase:glutaminase interface in the single-chain *Sc*-IGPS is tighter than in the *Tm*-IGPS heterodimer (see Fig. 1 D) and the only H-bond near the PGVG strand is the N52-A393 interaction (weaker than hV51-hP10 in HisH) that connects the PGVG(N) strand to the $\alpha 4''$ helix in the cyclase domain (see Fig. 1 C). Thus, the oxyanion strand is H-bonded to the HisH glutaminase in bacteria, while it remains at the interface between the two domains in *Sc*-IGPS. Therefore, the communication pathways along the two IGPS domains prior to the reaction at the glutaminase active site are presumably different in the two organisms. The flux of conformational changes associated with the allosteric mechanism of *Tm*-IGPS has been identified by computational studies and verified experimentally (23,25,28,30,38), but a comparative analysis of IGPS from different organisms was missing.

Here, we perform a comparative study of allosteric pathways in *Tm*- and *Sc*-IGPS adopting the same successful methodology used for the studies of bacterial IGPS. In particular, we used graph-theory-derived network models to analyze the correlations of nuclear fluctuations observed in molecular dynamics (MD) simulations of *Sc*-IGPS. This approach involves a set of computational tools that have previously been used to describe different aspects of the

protein dynamics in a variety of systems (39–42), including *Tm*-IGPS (23,38). Notably, in our earlier work on bacterial IGPS, the role of a HisF hydrophobic cluster in transmitting the effector binding signal has been confirmed in NMR titration experiments (23). Besides, mutation experiments coupled with kinetic essays have followed after our predicted allosteric pathways, targeting a few key residues where mutations induced the disruption of the allosteric effects (25). Moreover, our previous computational studies predicted that the allosteric pathways in *Tm*-IGPS involve an opening/closing (breathing) motion of the HisH domain relative to the HisF unit, supported by hinge-like interactions at the HisF:HisH interface (23,38). The crucial role of this interdomain collective motion was recently validated experimentally using an IGPS mutant involving a photo-responsive unnatural amino acid, which could lock the motion at the interface, resulting in modulation of the enzymatic activity (28). Finally, our previous MD simulations, which captured the early dynamics (100 ns) of bacterial IGPS, revealed how, for this time scale, the collective hinge motion is associated with local interresidue interactions that synergistically, and only in presence of the effector, initiate a conformational change in the HisH active site promoting the stabilization of an oxyanion hole. The hypothesis that the allosterically driven formation of such an oxyanion hole is essential for the IGPS catalytic activity, consistent with the active site conformational change seen in our MD simulations, was recently confirmed through experimental studies that finally characterized the pro-active configuration of *Tm*-IGPS (30). Altogether, the various experimental validations of our studies on *Tm*-IGPS allostery strongly support the robustness of our methodology in sampling the early allosteric dynamics of the IGPS enzyme and in characterizing the allosteric pathways (in terms of both local inter-residue interactions and collective protein motions), substantiating its application to the IGPS enzyme in another organism, such as the yeast *Sc*-IGPS.

In the present contribution, we thus compare the early allosteric dynamics and the well-established allosteric pathways of bacterial IGPS (23–26,28,30,38,43) with those of its yeast homolog, here obtained with the same methodology employed for *Tm*-IGPS, in conjunction with new complementary analysis of both *Tm*- and *Sc*-IGPS enzymes.

RESULTS AND DISCUSSION

Changes in correlations induced by PRFAR binding to IGPS from yeast and bacteria

Fig. 2 A shows the effect of PRFAR binding on the structure of correlations in IGPS from bacteria (left panel) and yeast (right panel), respectively. Specifically, Fig. 2 A shows maps of differences of generalized correlation coefficients, $r_{MI}[\mathbf{x}_i, \mathbf{x}_j]$ (39) in PRFAR-bound and apo IGPS of *Tm*-IGPS

(left panel) and *Sc*-IGPS (right panel), respectively. The generalized correlation coefficients $r_{MI}[\mathbf{x}_i, \mathbf{x}_j] = [1 - \exp(-2/3 I[\mathbf{x}_i, \mathbf{x}_j])]^{-1/2}$ provide a quantitative measure of correlations in the positions \mathbf{x}_i and \mathbf{x}_j of C α atoms in residues i and j , based on the mutual information $H[\mathbf{x}_i, \mathbf{x}_j] = H[\mathbf{x}_i] + H[\mathbf{x}_j] - H[\mathbf{x}_i, \mathbf{x}_j]$. Here, $H[\mathbf{x}_i]$ and $H[\mathbf{x}_i, \mathbf{x}_j]$ are the marginal and joint (Shannon) entropies, respectively, for atomic vector displacements (\mathbf{x}_i and \mathbf{x}_j) computed as ensemble averages over MD simulations of apo IGPS and PRFAR-bound states. The resulting correlation patterns reflect the early dynamics of *Sc*-IGPS (and *Tm*-IGPS), obtained by averaging the generalized correlation coefficients computed on six independent replicas of 100 ns (four replicas for *Tm*-IGPS), thus allowing for direct comparisons with earlier studies of *Tm*-IGPS (23,38) (further details provided in the [supporting material](#)). In addition, we performed a similar comparative analysis of correlations obtained instead using a gaussian network model (44) and based on the crystallographic structures of the IGPS enzyme from the two organisms (see Fig. S1). Notably, the resulting correlation matrices show evident differences, indicating that only part of the changes in correlations sampled with MD simulations are encoded in the structural differences between the two systems.

The distinct patterns of correlations, shown in Fig. 2 A for yeast and bacterial IGPS, suggest distinct allosteric motions triggered by PRFAR binding in the two enzymes. In particular, *Tm*-IGPS (Fig. 2 A, left) shows various domains within HisH and HisF where the residues are more correlated among themselves than with residues in other parts of the protein. This indicates a sort of internal division within HisH and HisF domains that clearly appears as blocks of reduced correlations (magenta features in Fig. 2 A, left panel) in one side of the *Tm*-IGPS (namely, sideL) and increased correlation (green features in Fig. 2 A, left panel) on the opposite side of the protein (namely, sideR). In *Tm*-IGPS, weaker correlations in the PRFAR-bound complex correspond to weaker interfacial HisH-HisF interactions upon effector binding. Reduced correlations affect the interdomain hinge-like breathing motion, as observed in MD simulations of apo and PRFAR-bound enzymes (23,38).

The hinge-like breathing motion plays a central role in the allosteric regulation of *Tm*-IGPS, as recently confirmed by experiments (28). The effector-induced internal division within HisH and HisF domains of *Tm*-IGPS is essentially absent in the *Sc*-IGPS cyclase and glutaminase domains of His7 (see Fig. 2 A, right panel), with a sizable increase of correlations observed only between cyclase residues 345–400 (belonging to $f\alpha 3$, $f\beta 4$, and $f\alpha 4$) and the rest of the enzyme. Moreover, in contrast with *Tm*-IGPS, binding of PRFAR in *Sc*-IGPS induces milder effects on the correlations of motions in the whole enzyme. Therefore, it is clear that PRFAR binding to *Sc*-IGPS does not affect a hinge-like breathing motion as in *Tm*-IGPS, consistent with the

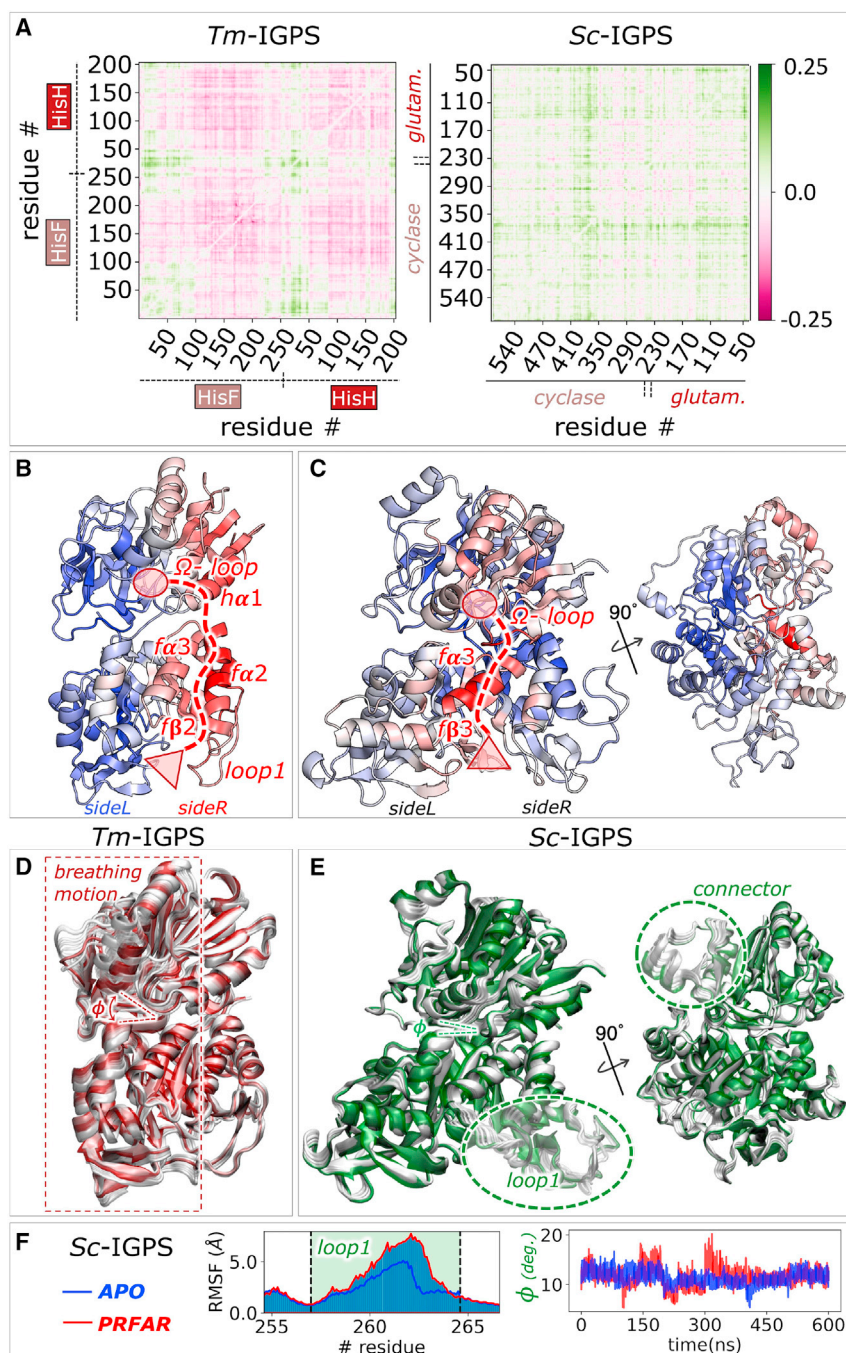


FIGURE 2 Analysis of correlated motions in IGPS from yeast and bacteria. (**A**) Comparison of generalized correlation coefficients $r_{M[x_i, x_j]}$ for PRFAR-minus-apo *Tm*-IGPS (left) and *Sc*-IGPS (right). In *Tm*-IGPS, PRFAR induces changes in both HisF and HisH, leading to innerly correlated domains (green features) as well as uncorrelated residues (magenta features), with amino acid residues 100–220 in HisF (sideL) featuring a decrease in correlations with the rest of the enzyme (black dotted lines). The PRFAR-minus-apo correlation matrix in *Sc*-IGPS does not exhibit similar features to those found in *Tm*-IGPS but rather milder changes of correlations due to effector binding, except for a sizable increase in correlations observed between cyclase residues 345 and 400 (belonging to $\alpha 3$, $\beta 4$, and $\alpha 4$) and the rest of *Sc*-IGPS (black dotted lines). The abbreviation glutam. refers to the glutaminase domain. (**B** and **C**) EC differences (PRFAR-minus-apo) projected onto the apo structure of *Tm*- (**B**) and *Sc*-IGPS (**C**), computed for local correlation values (damping-distance parameter $\lambda = 5$), featuring gains (in red) and loss (in blue) of centrality upon effector binding. The allosteric pathways from the effector site (red triangle) to the active site (red circle) in both enzymes are marked with red dotted lines. The main secondary structure elements along the pathways are labeled. (**D** and **E**) Differential (PRFAR-minus-apo) essential dynamics from the first PC of *Tm*- and *Sc*-IGPS MD trajectories. A rotated view of yeast IGPS is reported to visualize the motion of the connector. (**F**) RMSF of loop1 (left) and time evolution of the hinge breathing motion (right) in apo (blue lines) and PRFAR-bound (red lines) *Sc*-IGPS. The breathing motion is monitored by the G51(C α)–W124(C γ)–Y394(C γ) angle (ϕ) over the concatenated (600 ns) MD simulations.

hypothesis of different allosteric mechanisms in the two organisms.

Long- and short-range allosteric communication in IGPS from yeast and bacteria

Our analysis of correlations in *Tm*-IGPS and *Sc*-IGPS shows distinct changes in correlated motions induced by PRFAR binding that result from changes in both long- and short-

range interactions and enable allosteric activation of yeast and bacterial IGPS. Fig. 2 shows the principal component analysis (PCA) (40,45,46) and eigenvector centrality (EC) network analysis (38) of correlated motions. PCA selects the principal collective motions sampled from MD simulations by diagonalization of the covariance matrix of atomic displacements (see details in the supporting material), although it is limited to linear correlations. Thus, we employ the EC analysis to include non-linear correlations in an effort to disentangle long- and short-range contributions.

The EC methodology represents a cost-effective approach that yields fundamental understanding of allosteric mechanisms at the molecular level (38,47,48). Our implementation is based on a weighted graph with nodes corresponding to C α atoms and weights between pairs of C α atoms i and j determined by the corresponding generalized correlation coefficient $r_{\text{MI}}[\mathbf{x}_i, \mathbf{x}_j]$, as discussed above (see Fig. 2 A and B).

The centrality c_i of residue i is a real positive number defined by the i -th entry of the leading eigenvector of the weighted adjacency matrix $A_{ij} = (1 - \delta_{ij}) r_{\text{MI}}[\mathbf{x}_i, \mathbf{x}_j] \exp(-d_{ij}/\lambda)$. The damping parameter allows for the analysis of local correlations by simply dumping out the contributions from pairs of residues beyond a given range (see Fig. S5).

We initially focus on local centrality changes $\Delta c_i = c_i^{\text{PRFAR}} - c_i^{\text{APO}}$, induced by PRFAR, analyzed by defining A_{ij} with $\lambda = 5$ Å. Panels B and C in Fig. 2 show the normalized centrality differences Δc_i induced by PRFAR binding to *Tm*-IGPS and *Sc*-IGPS, respectively, with a color scale from minimal (blue) to maximal (red) values of Δc_i (details in the supporting material). The computed centrality differences reveal significant differences in the two organisms. For *Tm*-IGPS (38), only sideR transfers the allosteric signal through a pathway that involves multiple secondary structural elements: loop1, f β 2, f α 2, and f α 3 in HisF and h α 1, Ω -loop in HisH. The signal reaches the active site at the hC84 residue via alteration of H-bonding interactions with the highly conserved PGVG (oxyanion) strand, adjacent to the Ω -loop (23,38). In *Sc*-IGPS, however, the increased centralities induced by PRFAR binding are not localized on the sideR of the protein and involve a smaller number of secondary structure elements than in *Tm*-IGPS.

In fact, the PRFAR allosteric signal in His7 involves mainly f β 3 and f α 3 in the cyclase domain (where most of the increased values are found) with a direct link to the Ω -loop in the glutaminase domain that allows the signal to reach the active site (PGVG and C83) more directly than in *Tm*-IGPS.

Short-range correlations are affected by local contacts while long-range correlations involve collective modes that relate to slow protein motions. Here, we combine PCA and EC analysis to characterize the main collective modes and long-range correlations involved in the allosteric mechanisms. For *Tm*-IGPS, we have shown that the comparison of centrality differences obtained with $\lambda = \infty$ and $\lambda = 5$ allows for the characterization of long-range correlations in allosteric mechanisms that directly relate to the breathing motion of bacterial IGPS (38). Notably, we observed that the results agree with the essential motions induced by the effector as obtained by PCA (see Fig. S7). The essential motions are obtained by projecting the MD trajectories onto the main PRFAR-minus-apo difference principal components (PCs; ΔPC_1 and ΔPC_2 for first and second components, respectively). Fig. 2 D and E show the effector-induced essential motions described by the ΔPC_1 in both bacteria and yeast IGPS, indicating that there are significant differ-

ences in the two organisms. Indeed, the alteration of the breathing hinge motion in *Tm*-IGPS (see Fig. 2 D), upon effector binding, is replaced by a large motion of the loop1 (residues 250–275) and the connector site (residues 206–236) in *Sc*-IGPS. Analogously, the PRFAR-minus-apo difference for the second PC (ΔPC_2) reveals additional differences in the effector-induced essential dynamics of the two systems (see Figs. S8 and S9), with a mild movement of loop1 accompanying the *Tm*-IGPS hinge motion. In contrast, for *Sc*-IGPS, spring-like motion of the surface secondary structural elements of His7 was detected (see Videos S1–S3).

Overall, these results indicate that loop1 is involved in short-range interactions in *Tm*-IGPS allostery. However, in *Sc*-IGPS the loop1 is part of the long-range communication, becoming freer to fluctuate upon effector binding (see the root-mean-square fluctuations [RMSFs], reported in Fig. 2 F, left panel and in Fig. S8 compared with those in *Tm*-IGPS).

We note that loop1 is much shorter in *Tm*-IGPS than in *Sc*-IGPS so it might play different functional roles in the two systems. In fact, inspection of our MD trajectories suggests that loop1 of *Tm*-IGPS might play a 2-fold role in the *Tm*-IGPS by being involved in short-range allosteric communication and at the same time functioning as a gatekeeper to keep the effector in the binding pocket under high-temperature conditions. In *Sc*-IGPS, however, changes in the motion of loop1 induced upon effector binding are accompanied by the motion of the cyclase-glutaminase interdomain connector (see Fig. 2 E), alternatively to the breathing motion observed in *Tm*-IGPS (see Fig. 2 F, right panel), which is not present in *Sc*-IGPS (28). Moreover, the long and highly mobile loop1 of *Sc*-IGPS might facilitate PRFAR binding under room temperature conditions. In the absence of a prominent hinge-like motion as observed in *Tm*-IGPS, the role of the connector in *Sc*-IGPS is more related to the propagation of low-vibrational motions across the two domains. In this sense, while it was possible to successfully suggest point mutagenesis experiments targeting specific local contacts for loop1 in *Tm*-IGPS, the same is hard to do for loop1 and the connector site in *Sc*-IGPS, as their role is not associated with allosteric local contact changes but rather with the collective motions initiating allosteric communication.

Clearly, the combination of EC and PCA is a powerful methodology for identifying protein domains that are significantly affected upon binding of an allosteric effector and for characterization of essential motions, providing evidence of collective modes and inter-residue interactions that control the underlying allosteric mechanisms.

Besides, the residues showing the largest centrality values include those in f β 3 and f α 3 in the cyclase domain and those in the Ω -loop and vicinity (Fig. 2 C, highlighted in red), which represent promising targets for site-directed mutagenesis studies since they exhibit the highest increase in centrality upon PRFAR binding. The impact of mutants on *Tm*-IGPS has been evaluated experimentally through mutagenesis studies coupled to kinetic experiments (25,32), confirming

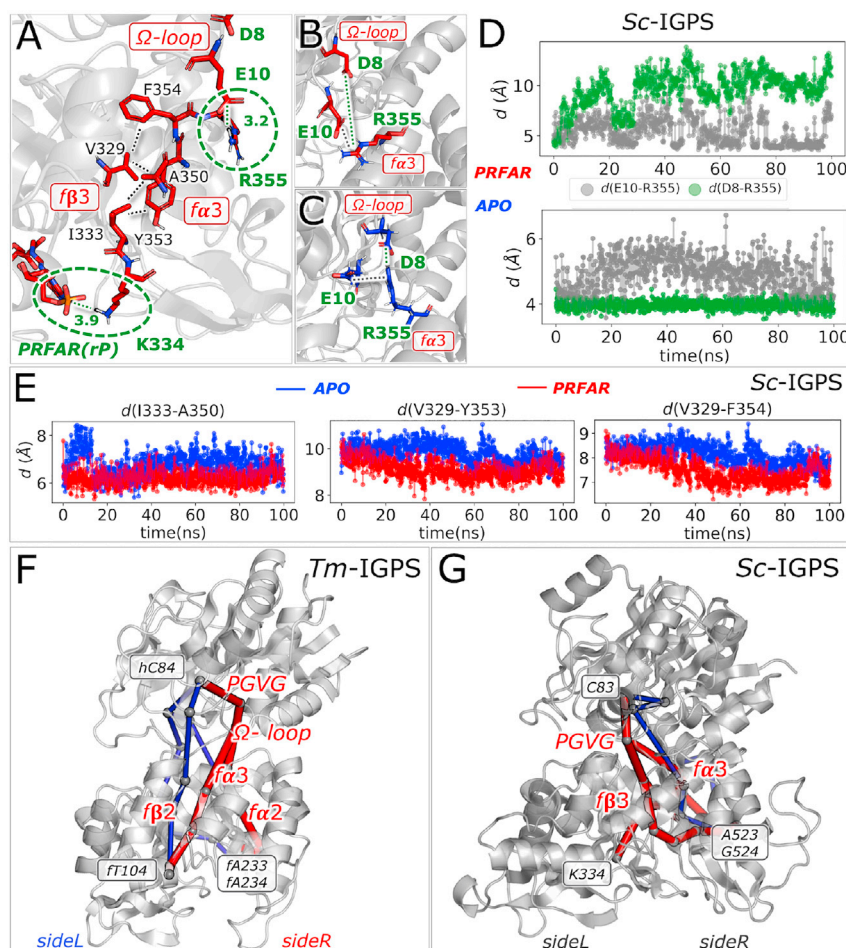


FIGURE 3 Allosteric communication between the effector and the glutaminase active site. (A) Local contacts spanning from the PRFAR binding site to the cyclase:glutaminase interface, involving at the extremes two salt bridges (green circles) between residue K334 in fβ3 and the ribose-side phosphate of PRFAR (rP) and between R355 in fα3 and D8 and E10 in the glutaminase Ω-loop, bridged by a cluster of hydrophobic interactions between fβ3 and fα3 residues (i.e., I333-A350-Y353-V329-F354). (B and C) Representative MD snapshots of the average R355-D8 and R355-E10 salt-bridge picture in the effector-bound (red sticks residues, B) and apo (blue sticks residues, C) complexes. (D) The effect of PRFAR binding on the time evolution of the R355-D8 and R355-E10 salt-bridge distances, along a representative 100 ns MD trajectory. (E) Tightening of the interactions in the fβ3-fα3 hydrophobic cluster upon PRFAR binding, along a representative 100-ns MD trajectory. (F and G) Shortest communication pathways connecting the fT104, fA223, and fA224 residues and the K334, A523, and G524 residues in the PRFAR binding sites of Tm-IGPS (F) and Sc-IGPS (G), respectively, and the Gln substrate binding site, i.e., hC84 and C83, respectively.

that mutants that directly target the allosteric pathway have a strong impact on the allosteric communication. We anticipate that similar site-directed mutagenesis studies on Sc-IGPS targeting the residues along the highest centrality pathway could shed light on the adaptation of the allosteric pathways in these protein homologs. We emphasize that mutants that lie outside of predicted allosteric pathways have been found to be less disruptive of the allosteric function in other systems (49), suggesting that future mutagenesis studies targeting random mutations of both Tm- and Sc-IGPS would be very informative for further insights that foster therapeutic applications aimed at altering the functionality of IGPS enzymes by targeting residues that control the enzyme's dynamics.

While we performed 12 independent 100-ns runs, one for each model of the apo and holo systems Sc-IGPS, the results discussed above are obtained by averaging the calculated properties over all model replicas (see [materials and methods](#) section). Hence, the average picture discussed above (involving differences between apo and holo dynamics) is representative of the allosteric process, although the individual simulations would present different EC (and PCA) profiles (as reported in [Fig. S5](#)).

Remarkably, the average correlation and EC profiles over the different replicas resemble one of them (labeled as *sim₁* in [Figs. S2, S4, and S5](#)), which seems to capture more clearly the allosteric effect (see additional comments in the [supporting material](#) documentation), so it has been selected as the most representative model replica in the following analysis.

In the next section, we analyze the allosteric pathways by inspecting those residues that are involved in short-range interactions responsible for information transfer across the catalytic units of IGPS. We do so by focusing on 100-ns snapshots that encompass most of the allosteric traits as identified by EC and PCA. The analysis provides understanding at the molecular level of the differences of the allosteric mechanism in the two organisms.

Allosteric pathways in IGPS from yeast and bacteria

[Fig. 3](#) shows the analysis of allosteric pathways in Tm-IGPS and Sc-IGPS as determined by the influence of PRFAR on the correlations of thermal nuclear fluctuations. We find that

optimal communication pathways from the effector to the active sites are distinct in the two systems since PRFAR affects specific interactions in the two systems. In *Sc*-IGPS, the phosphate group at the ribose side (rP) of the effector forms a tight salt bridge with K334 in the β 3 sheet (see Fig. 3 A) that is favored over the D335-K334 H-bond present in the apo state (see Fig. S10). In *Tm*-IGPS, such ionic interaction with PRFAR is absent (23) as there is no residue capable of establishing a salt bridge with the effector in the bacterial enzyme. In *Sc*-IGPS, K334 is adjacent to I333, which belongs to a network of hydrophobic contacts (I333-A350-Y353-V329-F354) spanning over the whole β 3- α 3 region (see Fig. 3 A). Notably, these hydrophobic interactions are significantly strengthened upon PRFAR binding (see Fig. 3 E), and thus a hydrophobic cluster is most responsible for transmitting the effector signal through the cyclase domain (i.e., HisF in *Tm*-IGPS), similarly to the process in bacterial IGPS (23). However, the activation of the hydrophobic cluster in *Tm*-IGPS (comprising the β 48- β 50- β 52- β 23 residues) involves the β 2 sheet (not the β 3- α 3 region as in *Sc*-IGPS). More importantly, the activation mechanism involves the loop1, which is engaged in short-range allosteric interactions in *Tm*-IGPS. Furthermore, we note that changes in hydrophobic contacts due to PRFAR binding are primarily driven by interactions with the π -system of the imidazolecarboxamide group of PRFAR (Fig. S13). In *Sc*-IGPS, however, the allosteric signal is initiated upon formation of the K334-PRFAR(rP) salt bridge.

Changes in the hydrophobic contacts in *Tm*-IGPS induced by PRFAR binding affect a network of salt bridges on the surface of the IGPS sideR, involving ionic interactions between the charged residues β R59, β E67, β E71, β E91, and β R95 in the α 2 and α 3 helices (at HisF) and the β R18 residue in α 1 (at HisH) (23). In *Sc*-IGPS, however, there are no corresponding charged surface residues that can create a salt-bridge network and, thus, the signal travels from PRFAR through the β 3- α 3 hydrophobic cluster until it reaches the charged residue R355 (at the end of α 3), which interfaces the glutaminase domain (Fig. 3 A). As shown in Fig. 3 B and C, indeed, the R355 charged sidechain could engage in interface ionic interactions with either D8 or E10 sidechains, belonging to the Ω -loop of the glutaminase subunit. Notably, as shown in Fig. 3 D, the R355-D8 salt bridge is stably formed throughout the MD trajectories of apo *Sc*-IGPS. However, PRFAR binding induces a change in the R355 partner, favoring formation of the R355-E10 salt bridge, which is weaker than the apo R355-D8 bond. These results indicate that the effector alters the α 3/ Ω -loop ionic interactions at the cyclase:glutaminase interface in *Sc*-IGPS, while in *Tm*-IGPS the affected salt bridges at the HisF/HisH interface involve the α 2/ α 1 helices of sideR. We suggest that future mutagenesis studies of *Sc*-IGPS can target the important residues highlighted in our analysis; i.e. those along the allosteric pathway (I333-A350-Y353-V329-F354, R355, E10 and D8).

The comparison of *Tm*- and *Sc*-IGPS active sites in the crystallographic structures highlights how the effector-induced β V51- β P10 H-bond breaking (23) (a crucial allosteric step observed for the bacterial enzyme; see Fig. S15) is not plausible *Sc*-IGPS where H-bonding interactions near the PGVG oxyanion strand, stable throughout the dynamics, are limited to the A393-N52 H-bond at the interface (see Fig. 1 C). The interface H-bond in apo *Sc*-IGPS is weaker than the (buried) β V51- β P10 bond in apo *Tm*-IGPS and, despite weakening of the A393-N52 interaction upon effector binding (see Fig. S16), dynamical fluctuations are more related to the (quite narrow) breathing motion in His7 (see Fig. 2 F) than to allosteric signal propagation through local contacts. Therefore, the observation that PRFAR binding in the yeast affects the α 3/ Ω -loop ionic interactions is not sufficient to explain how the effector signal is transferred from the interdomain interface to the active site of *Sc*-IGPS (i.e., there is no direct, allosterically modulated connection between the PGVG oxyanion strand and the Ω -loop in *Tm*-IGPS).

We analyze the communication pathways that link the effector site in the cyclase domain with the glutaminase active site and the activation mechanism toward the catalytically active state in both yeast and bacteria. The enzymatic communication pathways are computed as the optimal paths (i.e., paths with stronger correlation) connecting specific pairs of physically distant residues. Amino acid residues correspond to the nodes of a graph with edges defined by the strength of correlations between pairs of residues (11) (i.e., higher correlated pairs correspond to shorter bonds and are more likely to belong to the optimal communication path).

The communication pathways start at the PRFAR binding site with residues β T104, β A223, and β A224 of *Tm*-IGPS, and K334, A523, and G524 of *Sc*-IGPS. The target final node is the Gln substrate binding site (i.e., β C84 and C83 in *Tm*-IGPS and *Sc*-IGPS, respectively). As shown in Fig. 3 F and G, the resulting communication channels are affected by the effector binding (apo pathways in blue and PRFAR-bound in red), featuring significant differences between the two organisms. In accordance with our EC analysis (Fig. 2 B and C), the signal from the effector to the active site is triggered by PRFAR binding and is preferentially transferred through sideR in *Tm*-IGPS, involving PGVG and the Ω -loop. In *Sc*-IGPS, however, the allosteric pathway is more internal, allowing direct communication between the PGVG oxyanion strand and the cyclase domain, enhancing a spring-like PC motion of protein expansion and contraction.

The final step of allosteric activation at the IGPS active site is the rearrangement of the PGVG strand associated with the flipping of the amide N-H group of residue β V50/ β V51 toward the Gln-binding site (in *Tm*/*Sc*, respectively), which allows formation of an oxyanion hole (Fig. 4). In *Tm*-IGPS, we demonstrated that the initiation of the PGVG flipping requires breaking of the β V51- β P10

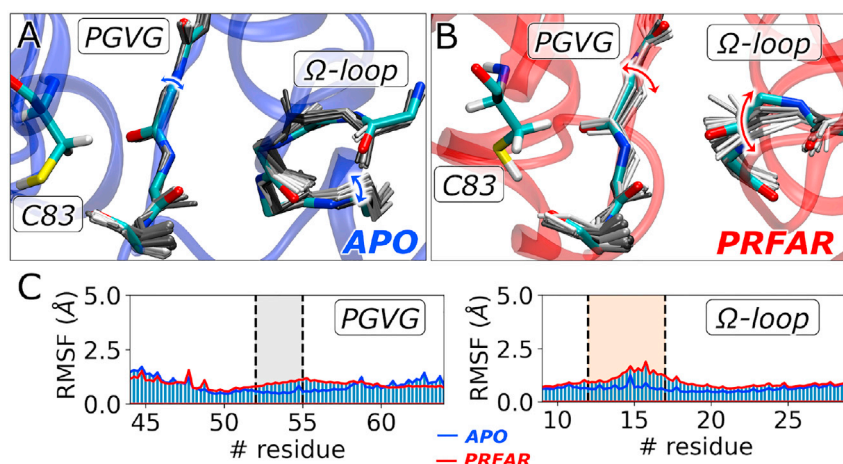


FIGURE 4 (A and B) Enhanced thermal fluctuations of the PGVG oxyanion strand and Ω-loop triggered by PRFAR binding in the glutaminase active site of *Sc*-IGPS. Average secondary structure in apo (blue), PRFAR-bound (red), and Gln-binding site (C83, colored sticks) are also depicted. (C) The RMSF profile of the PGVG oxyanion strand and Ω-loop in a representative (100 ns) trajectory in the apo (blue lines) and PRFAR-bound (red lines) complexes.

H-bond to separate the strand from the nearby Ω-loop (supporting material, Fig. S15). Notably, the breaking of this H-bond interaction has been resolved in the X-ray structure of the *Tm*-IGPS pro-active conformation (28), along with rearrangement of the PGVG strand and formation of the oxyanion hole. However, it remains to be established how the final allosteric step is initiated in *Sc*-IGPS, where PGVG and the Ω-loop are not linked by an H-bond.

Fig. 4 shows the early dynamics of PGVG and Ω-loop in *Sc*-IGPS and the differences observed (within 100 ns of a representative MD trajectory) between the apo and the PRFAR-bound complexes. The secondary structure elements PGVG and Ω-loop are not directly connected (e.g., by H-bonding) and are found to be more separated in *Sc*-IGPS than in *Tm*-IGPS (see Fig. 1 C). Nevertheless, both structural elements exhibit enhanced motion upon effector binding (Fig. 4 A–D), showing that changes in ionic interactions at the cyclase:glutaminase interface (e.g., R355-D8/E10 salt-bridge exchange; Fig. 3 B–D) correlates directly with motions in both secondary structural elements as the effector binds and promotes the interdomain signal transduction toward the active site.

In *Tm*-IGPS, changes in ionic contacts promote the HisF-HisH breathing motion that breaks the PGVG/Ω-loop H-bond and facilitates the PGVG flipping. In contrast, allostery in *Sc*-IGPS involves directly the Ω-loop, a structural element that affects the interface and enables the PGVG rearrangement in the absence of a hinge-like breathing motion. The limited interdomain motion in *Sc*-IGPS suggests that effector binding does not affect water accessibility to the glutaminase active site. Nevertheless, it is important to note that the reduced interdomain motion in *Sc*-IGPS is accompanied by enhanced collective motions of both the loop1 and the interdomain covalent connector, not present in *Tm*-IGPS.

Conclusions

We have characterized the early dynamics that involve the allosteric pathways of the IGPS enzyme in yeast and ther-

mophilic bacteria by combining MD simulations and graph network analysis of correlated motions influenced by effector binding. We have found rather distinct allosteric pathways in the two enzymes, with specific inter-residues interactions and collective protein motions associated with conformational changes that initiate the communication between the allosteric and catalytic sites.

We speculate that the structural differences between yeast and bacterial IGPS are tailored to allow the proteins to function in their respective natural environments, leading to different allosteric mechanisms communicating distant sites in the IGPS enzymes of the two organisms. The heterodimer *Tm*-IGPS adapts the allosteric pathways to exploit a larger flexibility at high temperatures by allowing ample hinge-like motions of the two protein subunits. In contrast, the single-chain enzyme *Sc*-IGPS, which functions at room temperature, establishes more internal allosteric pathways in terms of inter-residues interactions, allowing for more direct communication between the PGVG oxyanion strand and the cyclase domain, enhanced by an overall spring-like motion of protein expansion and contraction, driven by flexible portions of the protein (loop1 and connector site). These predictions pave the way for future experimental validation (by mutagenesis, NMR, and kinetic essays) of the proposed differences between the allostery in the two organisms.

Our study contributes to understanding how proteins absolving for the same function, but from different evolutionary pathways, preserve their functionality in different environments by adapting their signaling pathway.

MATERIALS AND METHODS

Correlation matrices for *Tm*-IGPS are obtained from the same trajectories and following the same protocol as in reference (23), while yeast models are built *ex novo*.

The computational structural models for apo and PRFAR-bound yeast IGPS complexes are based on the crystal structure of the bienzyme complex from *Sc*-IGPS at 2.4 Å resolution (PDB: 1OX6-B) (31). The HisH-HisF apo-complex having several missing residues (261–275, 301–304, and

551–552) and three extra residues at the beginning of the chain required modeling prior to simulation. To complete the structure, first, we stripped the first three residues, then we aligned and added residues 256–260 and 299–310 from 1OX4-B (removing overlapping residues from 1OX6 due to poor alignment). Finally, we added residues 550–552 from 1JVN-A (removing residue 550 from 1OX6-B). We constructed the remaining residues (256–275) using different tools available online, using which we produced six different structural models. One model was generated using Modeller (50), a second one using Swiss-Model (51), and four suitable homology models were found on modbase. PRFAR was bound to each model by aligning each structure to the effector-bound crystal structure of yeast IGPS (34) (PDB: 1OX5).

The 12 generated structures (six in the apo state, six bound to the effector) align with RMSD < 5 Å. To allow for a direct comparison between the dynamics of IGPS enzymes from *Tm*- and *Sc*-IGPS, we kept the simulation conditions analogous to the one used for bacterial IGPS in reference (23). Our choice of keeping the simulation conditions identical was motivated by recent studies demonstrating that PRFAR is a weaker allosteric activation at growth temperature than it is at room temperature (52). For the sake of clarity, we report some essential details below. MD simulations of the apo and PRFAR-bound structures of yeast IGPS are based on the AMBER-ff99SB (53) force field for the protein and generalized amber force field (54) for the PRFAR ligand (see [supporting material](#)), as implemented in the Amber20 software package (55). We performed 12 independent MD simulations, one for each complex (apo and PRFAR bound) for a total simulation time of 1.2 μs. Further details of the pre-equilibration procedure and MD production runs are described in the [supporting material](#). Details on the computation of generalized correlation coefficients and covariances between pairs of residues and their analysis through the EC metrics and PCA as well as the description of how to compute allosteric pathways across yeast and bacterial IGPS are provided in the [supporting material](#). Protein representations are obtained using the Pymol (56) software, with the exception of time-evolution representations, which are produced using VMD (56,57).

Determination of the allosteric pathways

The allosteric pathway for information transfer has been investigated by employing mutual information-based correlation analysis and network models from graph theory (39,40). Generalized correlations $r_{MI}[\mathbf{x}_i, \mathbf{x}_j]$ capture noncollinear correlations between pairs of residues i and j , and are helpful in pointing out the residues that are most affected by the binding of an effector, and with it the information channels that govern the allosteric control. $r_{MI}[\mathbf{x}_i, \mathbf{x}_j]$ alone can be hard to decipher and require some post-processing to interpret protein behavior. Network analysis tools (11,58), including different centrality metrics (59), can be applied for the interpretation of correlated protein motions and their allosteric behavior. Here, the Cα-atoms of the proteins' amino acid residues constitute the nodes of a dynamical network graph, connected by edges (residue pair connection in terms of $r_{MI}[\mathbf{x}_i, \mathbf{x}_j]$). An adjacency matrix is then constructed such that it can be used to identify the key amino acid residues of IGPS with high susceptibility to effector binding. A simple, yet effective metric that extracts central nodes in the adjacency matrix is the EC (38). The basic idea behind this measure is the assumption that the centrality index of a node is not only determined by its position in the network but also by the neighboring nodes, hence it measures how well connected a node is to other well-connected nodes in the network. The protein network can be used to determine the optimal pathways for the information transfer between two nodes, defined as the shortest paths connecting a specific pair of nodes. In this context, edge lengths (i.e., the internode distances in the graph) are defined using the coefficients according to $-\log(r_{MI}[\mathbf{x}_i, \mathbf{x}_j])$, implying that highly correlated pairs (featuring good communication) are close in distance in the graph. In particular, we applied the Dijkstra algorithm to calculate the shortest

pathways between residues *f*A233–*f*A234–A523/G524–R528 and hC84–C83, where each set of residues belongs to a different domain of bacterial and yeast IGPS, respectively. Hence, the computed pathways are composed of residue-to-residue steps that optimize the overall correlation (i.e., the momentum transport) between residues *f*A223–*f*A224 (at the effector site) and hC84 (in the glutaminase active site) in *Tm*-IGPS, and similarly residues K334, A523, G524, and C83 in His7. Additional details on the methods are included in the [supporting material](#). As mentioned above, all analyses are performed on six different models (*sim*₀, *sim*₁, ... *sim*₅) for yeast and four for bacterial IGPS, retrieved from reference (23), for which we examine both the apo and PRFAR-bound dynamics. Generalized correlation coefficients and covariances of atomic displacements are computed independently on each apo and PRFAR-bound 100-ns simulations. We compute the average PRFAR-bound-minus-apo correlation and covariance over each different model (four for bacteria and six for yeast). Remarkably, the average pictures depicted in [Fig. 2](#), obtained as the average apo-minus-holo correlation (or covariance) computed across the different models, are representative of the allosteric process, although the individual simulations present different correlations matrices, EC, and PCA profiles (as shown in [Figs. S2–S5](#)). Among the six apo and PRFAR-bound replicas, the dynamics of *sim*₁ clearly resembles that of the average pictures, as illustrated in [Fig. S5](#). Therefore, the characterization of shortest pathways and specific effector-induced contact changes has been reported in [Figs. 3 and 4](#) using data from the representative model (i.e., *sim*₁).

PCs of protein dynamics

PCA (40) is a recognized approach to capture the essential motions of the simulated systems. In PCA, the covariance matrix of the protein Cα atoms is calculated and diagonalized to obtain a new set of coordinates (eigenvectors) to describe the system motions. Each eigenvector—or PC—is associated with an eigenvalue, which denotes how much each eigenvector is representative of the system dynamics.

To avoid translational artifacts, we set the center of mass of each frame at the origin and rotate each frame to its optimally aligned orientation relative to the average structure—computed over all apo trajectories—which also has its center of mass at the origin. Next, we evaluate the covariances of the positional fluctuations of each system over the apo and PRFAR-bound trajectories obtained by concatenation of the independent apo and effector-bound replicas. Because the motion of sidechains is mostly independent of the essential dynamics of IGPS, we restrict the covariance to the backbone atoms only. Projecting the original (centered) data onto the eigenvectors results in the PCs, whose associated eigenvalue (variance) is indicative of the portion of motion that the eigenvector describes. Together, the first two PCs relative to *Tm*-IGPS incorporate 44% and 33% of the total motion of the bacterial apo and PRFAR-bound trajectories, respectively ([Fig. S3 A](#)), while the percentages become 42% and 44% for His7 ([Fig. S3 B](#)). The contribution added by the third PC is much smaller, hence we limited our analysis to the first two.

By projecting the trajectory coordinates onto the PCs, one can visualize the essential motions induced by effector binding in yeast and bacterial IGPS on the protein structure, along the trajectory. The corresponding motions are shown in [Figs. 2 D, E, S8 A](#), and [B](#).

SUPPORTING MATERIAL

Supporting material can be found online at <https://doi.org/10.1016/j.bpj.2021.11.2888>.

AUTHOR CONTRIBUTIONS

F.M., A.G., and A.P. performed the research. I.R. and V.S.B. designed the research. F.M., V.S.B., and I.R. wrote the paper.

ACKNOWLEDGMENTS

I.R. and A.G. gratefully acknowledge the use of HPC resources of the Pôle Scientifique de Modélisation Numérique (PSMN) of the ENS-Lyon, France.

This work was supported by the NIH grant GM106121 (V.S.B.) and a generous allocation of high-performance computing time from NERSC.

REFERENCES

- Wodak, S. J., E. Paci, ..., T. McLeish. 2019. Allostery in its many disguises: from theory to applications. *Structure*. 27:566–578.
- Greener, J. G., and M. J. Sternberg. 2018. Structure-based prediction of protein allostery. *Curr. Opin. Struct. Biol.* 50:1–8.
- Loutchko, D., and H. Flechsig. 2020. Allosteric communication in molecular machines via information exchange: what can be learned from dynamical modeling. *Biophys. Rev.* 12:443–452.
- East, K. W., E. Skeens, ..., G. P. Lisi. 2020. NMR and computational methods for molecular resolution of allosteric pathways in enzyme complexes. *Biophys. Rev.* 12:155–174.
- Birdsall, N. J., T. Farries, ..., M. Sugimoto. 1999. Subtype-selective positive cooperative interactions between brucine analogs and acetylcholine at muscarinic receptors: functional studies. *Mol. Pharmacol.* 55:778–786.
- Klein, J., and K. Löffelholz. 1996. Cholinergic Mechanisms: From Molecular Biology to Clinical Significance. Elsevier.
- Christopoulos, A. 2002. Allosteric binding sites on cell-surface receptors: novel targets for drug discovery. *Nat. Rev. Drug Discov.* 1:198–210.
- Guo, J., and H.-X. Zhou. 2016. Protein allostery and conformational dynamics. *Chem. Rev.* 116:6503–6515.
- Bozovic, O., J. Ruf, ..., P. Hamm. 2021. The speed of allosteric signaling within a single-domain protein. *J. Phys. Chem. Lett.* 12:4262–4267.
- Stock, G., and P. Hamm. 2018. A non-equilibrium approach to allosteric communication. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 373:20170187.
- Rivalta, I., and V. S. Batista. 2021. Community network analysis of allosteric proteins. *Methods Mol. Biol.* 2253:137–151.
- Suplatov, D., and V. Švedas. 2015. Study of functional and allosteric sites in protein superfamilies. *Acta Naturae*. 7:34–45.
- Hwang, P. K., and R. J. Fletterick. 1986. Convergent and divergent evolution of regulatory sites in eukaryotic phosphorylases. *Nature*. 324:80–84.
- Micheletti, C. 2013. Comparing proteins by their internal dynamics: exploring structure-function relationships beyond static structural alignments. *Phys. Life Rev.* 10:1–26.
- Shulman, A. I., C. Larson, ..., R. Ranganathan. 2004. Structural determinants of allosteric ligand activation in RXR heterodimers. *Cell*. 116:417–429.
- Süel, G. M., S. W. Lockless, ..., R. Ranganathan. 2003. Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat. Struct. Biol.* 10:59–69.
- Chen, E., K. Reiss, ..., G. P. Lisi. 2021. A structurally preserved allosteric site in the MIF superfamily affects enzymatic activity and CD74 activation in D-dopachrome tautomerase. *J. Biol. Chem.* 297:101061.
- Mottonen, J. M., D. J. Jacobs, and D. R. Livesay. 2010. Allosteric response is both conserved and variable across three CheY orthologs. *Biophys. J.* 99:2245–2254.
- Sethi, A., J. Tian, ..., S. Gnanakaran. 2013. A mechanistic understanding of allosteric immune escape pathways in the HIV-1 envelope glycoprotein. *PLoS Comput. Biol.* 9:e1003046.
- Gruber, R., and A. Horovitz. 2016. Allosteric mechanisms in chaperonin machines. *Chem. Rev.* 116:6588–6606.
- Royer, W. E., Jr., H. Zhu, ..., J. E. Knapp. 2005. Allosteric hemoglobin assembly: diversity and similarity. *J. Biol. Chem.* 280:27477–27480.
- Livesay, D. R., K. E. Kreth, and A. A. Fodor. 2012. A critical evaluation of correlated mutation algorithms and coevolution within allosteric mechanisms. *Methods Mol. Biol.* 796:385–398.
- Rivalta, I., M. M. Sultan, ..., V. S. Batista. 2012. Allosteric pathways in imidazole glycerol phosphate synthase. *Proc. Natl. Acad. Sci. U S A.* 109:E1428–E1436.
- Rivalta, I., G. P. Lisi, ..., V. S. Batista. 2016. Allosteric communication disrupted by a small molecule binding to the imidazole glycerol phosphate synthase protein–protein interface. *Biochemistry*. 55:6484–6494.
- Lisi, G. P., K. W. East, ..., J. P. Loria. 2017. Altering the allosteric pathway in IGPS suppresses millisecond motions and catalytic activity. *Proc. Natl. Acad. Sci. U S A.* 114:E3414–E3423.
- Botello-Smith, W. M., and Y. Luo. 2019. Robust determination of protein allosteric signaling pathways. *J. Chem. Theor. Comput.* 15:2116–2126.
- Lake, P. T., R. B. Davidson, ..., M. McCullagh. 2020. Residue-level allostery propagates through the effective coarse-grained hessian. *J. Chem. Theor. Comput.* 16:3385–3395.
- Kneuttinger, A. C., C. Rajendran, ..., R. Sterner. 2020. Significance of the protein interface configuration for allostery in imidazole glycerol phosphate synthase. *Biochemistry*. 59:2729–2742.
- Chittur, S. V., T. J. Klem, ..., V. J. Davisson. 2001. Mechanism for acivicin inactivation of triad glutamine amidotransferases. *Biochemistry*. 40:876–887.
- Wurm, J. P., S. Sung, ..., R. Sprangers. 2021. Molecular basis for the allosteric activation mechanism of the heterodimeric imidazole glycerol phosphate synthase complex. *Nat. Commun.* 12:2748.
- Chaudhuri, B. N., S. C. Lange, ..., J. L. Smith. 2003. Toward understanding the mechanism of the complex cyclization reaction catalyzed by imidazole glycerolphosphate synthase: crystal structures of a ternary complex and the free enzyme. *Biochemistry*. 42:7003–7012.
- List, F., M. C. Vega, ..., M. Wilmanns. 2012. Catalysis uncoupling in a glutamine amidotransferase bienzyme by unblocking the glutaminase active site. *Chem. Biol.* 19:1589–1599.
- Douangamath, A., M. Walker, ..., M. Wilmanns. 2002. Structural evidence for ammonia tunneling across the ($\beta\alpha$)8 barrel of the imidazole glycerol phosphate synthase bienzyme complex. *Structure*. 10:185–193.
- Chaudhuri, B. N., S. C. Lange, ..., J. L. Smith. 2001. Crystal structure of imidazole glycerol phosphate synthase: a tunnel through a ($\beta\alpha$)8 barrel joins two active sites. *Structure*. 9:987–997.
- Kneuttinger, A. C., K. Straub, ..., R. Sterner. 2019. Light regulation of enzyme allostery through photo-responsive unnatural amino acids. *Cell Chem. Biol.* 26:1501–1514.e9.
- Amaro, R. E., A. Sethi, ..., Z. A. Luthey-Schulten. 2007. A network of conserved interactions regulates the allosteric signal in a glutamine amidotransferase. *Biochemistry*. 46:2156–2173.
- Amaro, R. E., R. S. Myers, ..., Z. A. Luthey-Schulten. 2005. Structural elements in IGP synthase exclude water to optimize ammonia transfer. *Biophys. J.* 89:475–487.
- Negre, C. F. A., U. N. Morzan, ..., V. S. Batista. 2018. Eigenvector centrality for characterization of protein allosteric pathways. *Proc. Natl. Acad. Sci. U S A.* 115:E12201–E12208.
- Lange, O. F., and H. Grubmüller. 2006. Generalized correlation for biomolecular dynamics. *Proteins*. 62:1053–1061.
- Lange, O. F., and H. Grubmüller. 2008. Full correlation analysis of conformational protein dynamics. *Proteins*. 70:1294–1312.
- Palermo, G. 2019. Structure and dynamics of the CRISPR-cas9 catalytic complex. *J. Chem. Inf. Model.* 59:2394–2406.
- Melo, M. C. R., R. C. Bernardi, ..., Z. Luthey-Schulten. 2020. Generalized correlation-based dynamical network analysis: a new high-performance approach for identifying allosteric communications in molecular dynamics trajectories. *bioRxiv* <https://doi.org/10.1101/2020.06.18.160572>.

Maschietto et al.

43. Gheeraert, A., L. Pacini, ..., I. Rivalta. 2019. Exploring allosteric pathways of a V-type enzyme with dynamical perturbation networks. *J. Phys. Chem. B*. 123:3452–3461.
44. Atilgan, A. R., S. R. Durell, ..., I. Bahar. 2001. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* 80:505–515.
45. David, C. C., and D. J. Jacobs. 2014. Principal component analysis: a method for determining the essential dynamics of proteins. *Methods Mol. Biol.* 1084:193–226.
46. Amadei, A., A. B. Linssen, and H. J. Berendsen. 1993. Essential dynamics of proteins. *Proteins*. 17:412–425.
47. Jalili, M., A. Salehzadeh-Yazdi, ..., K. Alimoghaddam. 2016. Evolution of centrality measurements for the detection of essential proteins in biological networks. *Front. Physiol.* 7:375.
48. Ashtiani, M., A. Salehzadeh-Yazdi, ..., M. Jafari. 2018. A systematic survey of centrality measures for protein-protein interaction networks. *BMC Syst. Biol.* 12:80.
49. Wang, J., A. Jain, ..., N. V. Dokholyan. 2020. Mapping allosteric communications within individual proteins. *Nat. Commun.* 11:3862.
50. Webb, B., and A. Sali. 2016. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* 86:2.9.1–2.9.37.
51. Waterhouse, A., M. Bertoni, ..., T. Schwede. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46:W296–W303.
52. Lisi, G. P., A. A. Currier, and J. P. Loria. 2018. Glutamine hydrolysis by imidazole glycerol phosphate synthase displays temperature dependent allosteric activation. *Front. Mol. Biosci.* 5:4.
53. Wang, J., P. Cieplak, and P. A. Kollman. 2000. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* 21:1049–1074.
54. Wang, J., R. M. Wolf, ..., D. A. Case. 2004. Development and testing of a general amber force field. *J. Comput. Chem.* 25:1157–1174.
55. Case, D. A., H. M. Aktulga, ..., P. A. Kollman. 2021. Amber 2020. University of California.
56. Schrödinger, L., and W. DeLano. 2020. PyMOL, Available at: <http://www.pymol.org/pymol>.
57. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38.
58. Yang, Z., R. Algesheimer, and C. J. Tessone. 2016. A comparative analysis of community detection algorithms on artificial networks. *Sci. Rep.* 6:30750.
59. Oldham, S., B. Fulcher, ..., A. Fornito. 2019. Consistency and differences between centrality measures across distinct classes of networks. *PLoS One*. 14:e0220061.