

# Allosteric Pathways Originating at Cysteine Residues in Regulators of G-Protein Signaling Proteins

Yong Liu<sup>1</sup> and Harish Vashisth<sup>1,\*</sup>

<sup>1</sup>Department of Chemical Engineering, University of New Hampshire, Durham, New Hampshire

**ABSTRACT** Regulators of G-protein signaling (RGS) proteins play a central role in modulating signaling via G-protein coupled receptors (GPCRs). Specifically, RGS proteins bind to activated  $G\alpha$  subunits in G-proteins, accelerate the GTP hydrolysis, and thereby rapidly dampen GPCR signaling. Therefore, covalent molecules targeting conserved cysteine residues among RGS proteins have emerged as potential candidates to inhibit the RGS/ $G\alpha$  protein-protein interaction and enhance GPCR signaling. Although these inhibitors bind to conserved cysteine residues among RGS proteins, we have previously suggested [J. Am. Chem. Soc. 2018;140:3454–3460] that their potencies and specificities are related to differential protein dynamics among RGS proteins. Using data from all-atom molecular dynamics simulations, we reveal these differences in dynamics of RGS proteins by partitioning the protein structural space into a network of communities that allow allosteric signals to propagate along unique pathways originating at inhibitor binding sites and terminating at the RGS/ $G\alpha$  protein-protein interface.

**SIGNIFICANCE** We reveal correlations between protein dynamics, allosteric communication, and potencies of covalent inhibitors in homologous protein isoforms of the RGS family. Specifically, using molecular dynamics simulations, we discovered that the protein structural space can be partitioned into a network of residue communities that work as hubs for allosteric communication between the binding sites of covalent inhibitors and residues in the protein-protein interfaces. These results explain differential inhibition among RGS proteins and suggest new residues as potential sites for the design of future allosteric modulators.

## INTRODUCTION

Protein-protein interactions (PPIs) are commonly involved in biological functions (1). Given that aberrant PPIs are implicated in several diseases, a traditional approach to inhibit PPIs is to target orthosteric sites using small molecules (2–4). However, it is challenging to inhibit protein-protein interfaces that are often flat, lack well-defined binding pockets, or have pockets that are buried within the interface and require transient exposure for the binding of small molecules (5–9). Alternatively, small molecules can be targeted at allosteric sites with the goal of inhibiting the protein-protein interface by perturbing the conformational dynamics of proteins involved. Moreover, targeting allosteric sites over orthosteric sites has several advantages

because 1) binding at allosteric sites is noncompetitive with the direct binding of endogenous ligands; 2) allosteric binding sites may be more accessible than a buried orthosteric site; and 3) allosteric effects are saturable, whereas the effects of orthosteric binding are concentration dependent (2,7–11). Furthermore, small molecules that covalently modify allosteric sites may increase specificity and also reduce drug dosage (12). Although covalent inhibitors are of concern because of off-target effects, they can significantly decrease the off-rate and thereby improve the potency, as is known for marketed drugs like aspirin (13).

However, targeting allosteric sites is challenging because of a poor understanding of protein conformational dynamics and allosteric communication between various structural motifs. Moreover, it is difficult to discern structural changes based only upon the crystal structures of proteins. In addition, the selectivity of small molecules covalently targeting allosteric sites on proteins from a common family cannot be predicted solely from their structures because of significant similarities in structural folds. Therefore, molecular

Submitted August 28, 2020, and accepted for publication December 14, 2020.

\*Correspondence: [harish.vashisth@unh.edu](mailto:harish.vashisth@unh.edu)

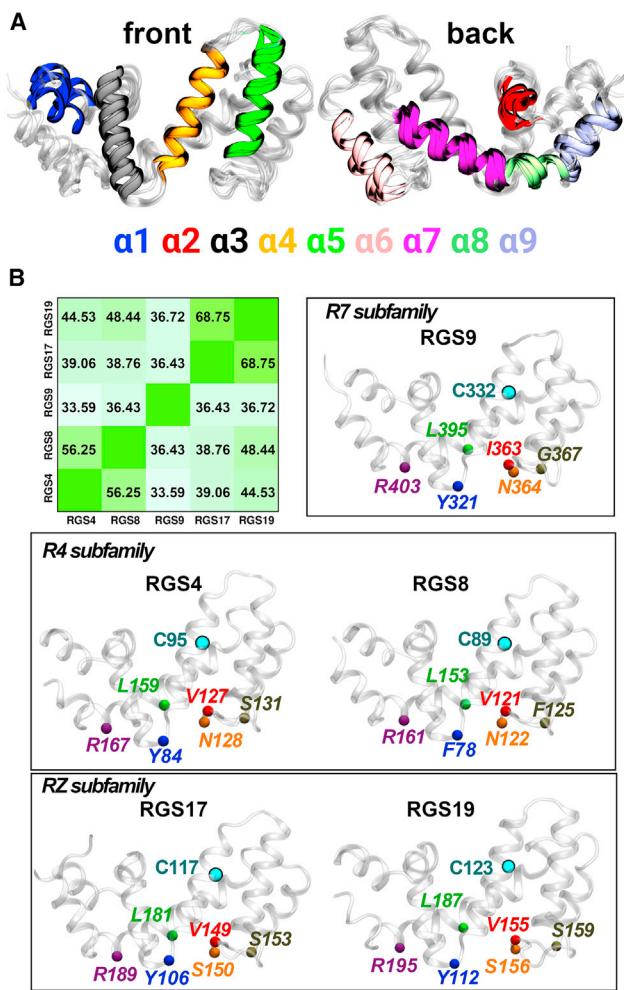
Editor: Marta Filizola.

<https://doi.org/10.1016/j.bpj.2020.12.010>

© 2020 Biophysical Society.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).





**FIGURE 1** Structural similarities and key residues in RGS proteins. (A) Shown are the front and back views of the overlay of structures of five RGS proteins: RGS4 (PDB: 1AGR), RGS8 (PDB: 2ODE), RGS9 (PDB: 1FQI), RGS17 (PDB: 6AM3), and RGS19 (PDB: 1CMZ). The alignment is based on the  $C_{\alpha}$ -atoms of the  $\alpha 4$ -helix. (B) Sequence similarity between RGS proteins is highlighted. The similarity between a pair of RGS proteins was computed based upon the sequence alignment (Fig. S2) and residues 52–178 (RGS4), 46–172 (RGS8), 289–414 (RGS9), 74–200 (RGS17), and 80–206 (RGS19). Shown also are cartoon representations of all RGS proteins, classified by their subfamily, and highlighting the conserved cysteine residues (cyan spheres) that are targeted by covalent inhibitors. Six key residues of each RGS protein that participate in the RGS/G $\alpha$  protein-protein interface are shown by colored spheres and labeled. The structures of three known RGS/G $\alpha$  complexes are shown in Fig. S3. To see this figure in color, go online.

dynamics (MD) simulations are increasingly playing a pivotal role in resolving the details of protein conformational dynamics and allosteric communication pathways (14–16). For example, MD simulations have been successfully applied to map allosteric communication pathways in the kinase family (17–22).

In this work, we aim to probe using MD simulations pathways originating at covalent allosteric sites, specifically at

conserved cysteine residues, in homologous proteins of the regulators of G-protein signaling (RGS) family. Structurally, RGS proteins have a conserved RGS-box domain consisting of nine  $\alpha$ -helices ( $\alpha 1$  through  $\alpha 9$ ) (Fig. 1 A). The functional role of RGS proteins is to bind to activated (GTP-bound) G $\alpha$  subunits of G-proteins and accelerate the rate of GTP hydrolysis, thereby deactivating G $\alpha$  and terminating signaling by G-protein coupled receptors (GPCRs) (2,23–25). Therefore, small-molecule inhibitors of the RGS/G $\alpha$  PPI, which enhance signaling via GPCRs, are potentially useful to develop therapeutics for cancer, cardiovascular diseases, and central nervous system disorders (2,5,26,27).

Specifically, thiadiazolidinone (TDZD) inhibitors (Fig. S1) that covalently modify cysteine residues in RGS proteins have shown promise in inhibiting the RGS/G $\alpha$  PPI via an allosteric mechanism (24,25,28–31). The selectivities and potencies of TDZD inhibitors are thought to be related to the number of cysteine residues in RGS proteins (32) because the TDZD inhibitor CCG-50014 is selective for RGS4 (with four cysteine residues) over RGS8 (with two cysteine residues) (28,30). However, we have also shown correlations between the inhibitor potency and protein dynamics, especially conformational changes leading to the exposure of buried and conserved cysteine residues in RGS proteins (24,25,31). For example, we showed that RGS proteins retaining only a single shared cysteine residue showed differences in potencies of CCG-50014 (25). Although covalent inhibitors are known to be irreversible, some covalent inhibitors (CCG-63802 and CCG-63808; Fig. S1) of RGS proteins are reversible (29).

Nonetheless, it remains unclear how internal motions in homologous RGS proteins form allosteric networks and what pathways exist through which the allosteric perturbations from covalent binding sites are conveyed to the RGS-G $\alpha$  protein-protein interface. To address these questions, we have studied here five RGS proteins (RGS4, RGS8, RGS9, RGS17, and RGS19) through dynamic allosteric analysis of long timescale MD simulations. This analysis has revealed differences in allosteric pathways originating at the conserved cysteine residues (termed source residues) located on the  $\alpha 4$ -helix of each RGS protein and arriving at each of the six key residues (termed sink residues) in the RGS-G $\alpha$  protein-protein interface.

## METHODS

### System preparation and simulation details

We used Nanoscale Molecular Dynamics (NAMD) (33) software to perform MD simulations using the CHARMM force field (34–36) and Visual Molecular Dynamics (VMD) (37) software to visualize and analyze data. We used the initial coordinates from the Protein Data Bank entries PDB: 1AGR, 2ODE, 1FQI, 6AM3, and 1CMZ for RGS4, RGS8, RGS9, RGS17, and RGS19, respectively. We solvated all systems with TIP3P water molecules and charge neutralized with NaCl. The final simulation

domains were comprised of 28,160 atoms (RGS4), 30,731 atoms (RGS8), 30,777 atoms (RGS17), 29,369 atoms (RGS17), and 29,560 atoms (RGS19). After an initial energy minimization (500 cycles) of all systems, we optimized box volumes in the NPT ensemble for  $\sim 100$  ps using a time step of 2 fs. The pressure was set at 1 atm and controlled using the Nosé-Hoover barostat, and the temperature was controlled at 310 K using the Langevin thermostat. We then conducted long timescale MD simulations of all systems in the NVT ensemble. Each system was subjected to a 2- $\mu$ s-long MD simulation with a time step of 2 fs. The periodic boundary conditions were applied in all simulations. We also analyzed simulation data on RGS4, RGS8, and RGS19 from our previous work (25).

## Residue fluctuations and salt bridge analyses

We analyzed  $C_{\alpha}$ -based root mean-squared fluctuation (RMSF) per residue to identify flexible residues and to compare differences in dynamics among RGS proteins. We also analyzed a network of conserved and non-conserved salt bridges formed between charged amino acids. A salt bridge was considered stable if the distance between the nitrogen atoms of basic residues and any of the oxygen atoms of acidic residues forming the salt bridge was within 3.2 Å. The error bars for the RMSF and salt bridge data were computed based on the block standard error (BSE) analysis (38). Briefly, BSE is given by  $\sigma_n/\sqrt{M}$  where  $M$  denotes the number of blocks in a trajectory with  $N$  frames,  $n$  is the length of each block, and  $\sigma_n$  is the SD computed from averages of an observable (e.g., RMSD, distance) for each block length, which is gradually increased. The function BSE plotted against  $n$  increases monotonically and asymptotically converges to the true standard error associated with the mean of the observable (38).

## Dynamic network analysis

To infer correlated residues and allosteric networks in proteins, several approaches based upon protein sequence and dynamics have been developed (39,40). The sequence-based approaches make use of multiple sequence alignment and the coevolution principle for identifying evolutionarily conserved residues that can be combined into allosteric groups and sectors in proteins (41,42). The dynamics-based approaches rely on data from MD simulations and include principle component analysis (43), mutual information between residues (44), timing correlations (45), interaction correlations (46), and community network analysis (15).

We used the method developed by Sethi et al. (15) to conduct community and allosteric pathway analyses, which has been successfully applied in extensive studies of allosteric in protein kinases (17–22). Based on MD simulations of each protein, we first performed the cross correlation analysis using CARMA (47) by setting the  $C_{\alpha}$ -atoms of residues as nodes. The pairwise correlations are defined by the following equation:

$$C_{ij} = \frac{\langle \Delta \vec{r}_i(t) \cdot \Delta \vec{r}_j(t) \rangle}{(\langle (\Delta \vec{r}_i(t)^2) \rangle \langle (\Delta \vec{r}_j(t)^2) \rangle)^{1/2}}, \quad (1)$$

where  $\Delta \vec{r}_i(t) = \vec{r}_i(t) - \langle \vec{r}_i(t) \rangle$ ,  $\vec{r}_i(t)$  is the position of the node  $i$ , and  $\langle \vec{r}_i(t) \rangle$  is the mean position of the node  $i$ . An edge is formed between two nodes when the two nodes are within a cutoff distance of 4.5 Å for at least 75% of the time of an MD trajectory (15). We have also studied the effect of varying the cutoff distance between 5 and 7 Å (see [Supporting Materials](#)). A length of an edge ( $l_{ij}$ ) is defined as follows:

$$l_{ij} = -\log(|C_{ij}|), \quad (2)$$

where  $C_{ij}$  is the normalized correlation value between the nodes  $i$  and  $j$ . When a  $|C_{ij}| \rightarrow 1$ , the length of an edge goes to 0, whereas the length goes to infinitely large when  $|C_{ij}| \rightarrow 0$ . We then can find the optimal

pathway from a source node to a sink node, which has the shortest length  $l_{opt}$ . The edge betweenness is then defined as the number of shortest pathways crossing that edge.

To identify communities, we used the Girvan-Newman algorithm (48), which removes the edge with the highest betweenness until the last node. Then, the optimal community structure is chosen using a largest modularity value, which measures difference in the probability of intracommunity and intercommunity edges. The modularity value has a maximal value of 1. Generally, the modularity values are in the range of 0.4–0.7 (49). For our systems, the modularity value is  $\sim 0.58$ . In a community network, there may exist several edges connecting any two communities. These edges are termed as critical edges, and the nodes forming these edges are termed as critical nodes.

## RESULTS

In this work, we aim to evaluate differences in structure and dynamics of five distinct RGS proteins (RGS4, RGS8, RGS9, RGS17, and RGS19), especially allosteric communication pathways originating at conserved cysteine residues and ending at the protein-protein interface between RGS proteins and  $G\alpha$  subunits of G-proteins. We first highlight a comparison of sequences and structures of all five RGS proteins, followed by per residue fluctuations as observed from initial structures and subsequent MD simulations. We then discuss conserved salt bridging interactions among various structural motifs, allosteric community network, and allosteric pathways.

## Sequence and structural comparison

We have studied five RGS proteins from three different subfamilies: the R4 subfamily (RGS4 and RGS8), the R7 subfamily (RGS9), and the RZ subfamily (RGS17 and RGS19). We report sequence similarity among pairs of RGS proteins in [Fig. 1 B](#) and the sequence alignment for all proteins in [Fig. S2](#). For the same subfamily, we observed a sequence similarity of 56.25% among RGS4 and RGS8 and of 68.75% among RGS17 and RGS19, although sequence similarities are lower ( $\sim 33$ –48%) between RGS proteins from different subfamilies. Contrary to the variation among their sequences, the structures of RGS proteins are highly similar with each protein containing nine  $\alpha$ -helices ([Fig. 1 A](#)) and a highly conserved cysteine residue on the  $\alpha 4$ -helix (labeled in *cyan* in [Fig. 1 C](#) and highlighted in *cyan* in [Fig. S2](#)). The canonical RGS- $G\alpha$  protein-protein interface mainly has three structural motifs on  $G\alpha$  termed as the switch regions (switch I, II, and III; [Fig. S3](#)) that contact those residues in RGS proteins that are buried within the interface ([Fig. 1 C](#)). We aim to understand allosteric coupling between the conserved cysteine residues on the  $\alpha 4$ -helix that are sites of inhibitor binding (31) and key residues in the RGS- $G\alpha$  interface. The comparison of sequences and structures alone is limited in gaining insights into these allosteric couplings. Therefore, we report below metrics aimed at differentiating dynamical features among RGS proteins.

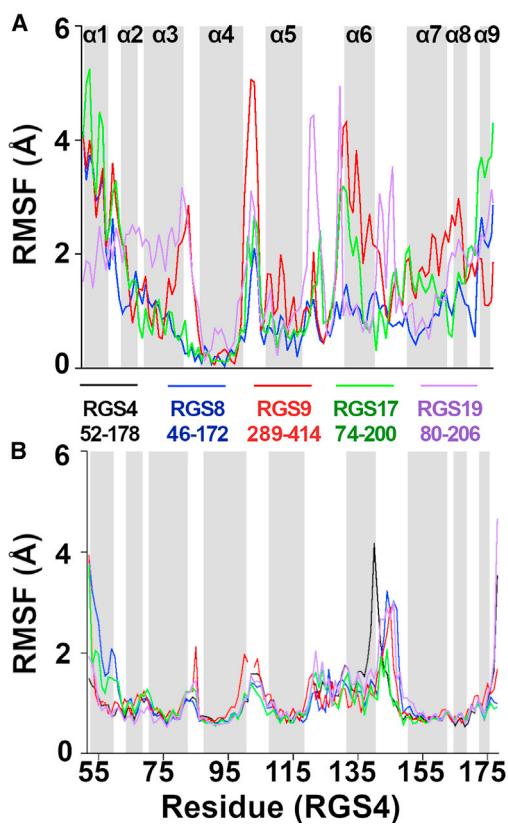


FIGURE 2 Residue Fluctuations. The RMSF values per residue are shown based upon experimental structures (A) and MD simulations (B). The residue numbers on the x-axis are for RGS4 (52–178) that correspond to the following residues in other RGS proteins: 46–172 (RGS8), 289–414 (RGS9), 74–200 (RGS17), and 80–206 (RGS19). A break in the RMSF data for RGS9 indicates a difference in amino-acid sequence from RGS4 (Fig. S2). The error bars corresponding to the RMSF per residue data from MD simulations (B) are shown in Fig. S4. To see this figure in color, go online.

### Residue fluctuations

We report RMSF values per residue based upon the initial structures of all RGS proteins as well as from their subsequent MD simulations (Fig. 2). On comparing experimental structures of other RGS proteins (RGS8, RGS9, RGS17, and RGS19) with RGS4, a canonical member of this family, higher RMSF values for residues in the  $\alpha$ 4– $\alpha$ 5 interhelical loops were observed for all other RGS proteins, with the highest values in the  $\alpha$ 4– $\alpha$ 5 loop of RGS9, which is likely due to the presence of a glycine (G341) residue (Fig. 2). In comparison to other RGS proteins, higher RMSF values for residues in the  $\alpha$ 5– $\alpha$ 6 and  $\alpha$ 6– $\alpha$ 7 interhelical loops and in the  $\alpha$ 3-helix (Fig. 2 A, magenta trace) were observed for RGS19 and for residues in the  $\alpha$ 6 and  $\alpha$ 7 helices of RGS9 and RGS17 (Fig. 2 A, red and green traces).

We further calculated RMSF values based on MD simulations of RGS structures (Fig. 2 B) and observed that residues in the  $\alpha$ 6– $\alpha$ 7 loops showed higher fluctuations in all RGS proteins except in RGS17. Also, the RMSF values

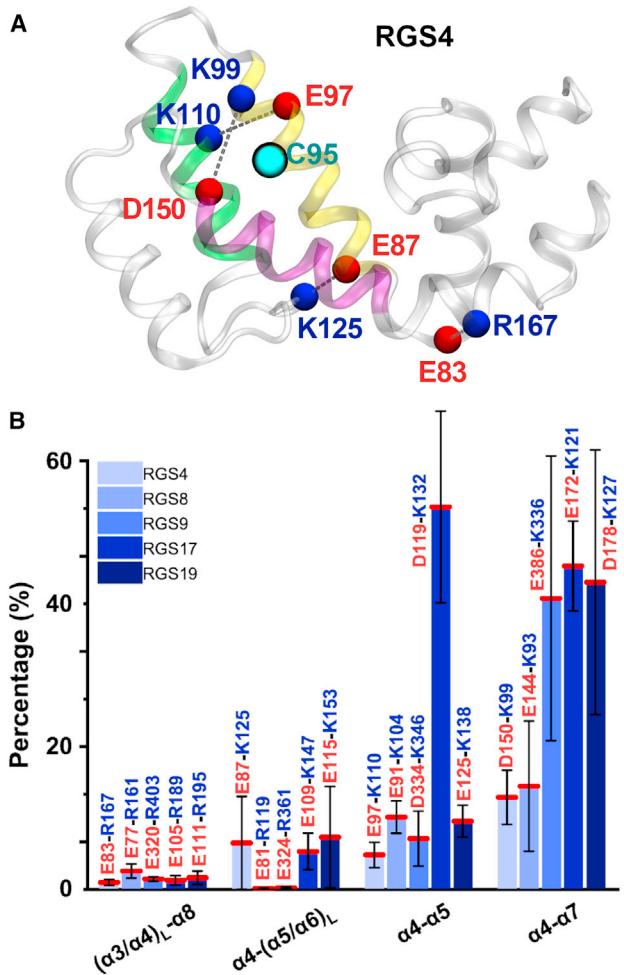
for residues in the  $\alpha$ 3– $\alpha$ 4 loop of RGS9 were higher than in other RGS proteins. Importantly, the structural motifs showing higher flexibility are either located in the RGS/G $\alpha$  protein-protein interface (e.g., residues in the  $\alpha$ 3– $\alpha$ 4 and  $\alpha$ 5– $\alpha$ 6 loops) or in the proximity of cysteine residues (e.g., residues in the  $\alpha$ 4– $\alpha$ 5 and  $\alpha$ 6– $\alpha$ 7 loops) accessed by allosteric inhibitors of RGS proteins. We hypothesize that these differences in conformational fluctuations of residues in the loop motifs of RGS proteins potentially contribute to differences in potencies and selectivities of inhibitory compounds (25,32).

### Correlation between salt bridges and RGS dynamics

We have previously shown that the mutations in residues forming salt bridges in RGS4, RGS8, and RGS19 alter their dynamics and correlate with inhibitor potency (50). Therefore, we further analyzed salt bridging interactions between charged residues in all RGS proteins as changes in these interactions may perturb the conformational flexibility of helical motifs and may play a role in allosteric communication.

We observed eight salt bridges conserved across all five RGS proteins, four of which connect different helices or loops (Fig. 3 A), whereas the remaining four salt bridges reside within the same helix or a loop (Fig. S5). In Fig. 3 A, we show four conserved salt bridges for RGS4 that connect two different helices or a helix with a loop: E83-R167 links the  $\alpha$ 3– $\alpha$ 4 interhelical loop with the  $\alpha$ 8-helix, E87-K125 links the  $\alpha$ 4-helix with the  $\alpha$ 5– $\alpha$ 6 interhelical loop, E97-K110 links the  $\alpha$ 4-helix with the  $\alpha$ 5-helix, and K99-D150 links the  $\alpha$ 4-helix with the  $\alpha$ 7-helix. The percentage occupancy of each of the four conserved salt bridges across all five RGS proteins is shown in Fig. 3 B.

These results suggest the following: 1) the E77-R161 salt bridge in RGS8, corresponding to E83-R167 in RGS4, is marginally stronger than in other four RGS proteins; 2) salt bridges homologous to E87-K125 in RGS4 show marginally stronger interactions in RGS4, RGS17, and RGS19 than in RGS8 and RGS9; 3) among salt bridges homologous to E97-K110 in RGS4 as well as among three other conserved salt bridges, the D119-K132 salt bridge in RGS17 shows the highest occupancy; and 4) the K99-D150 salt bridge in RGS4, which connects the  $\alpha$ 4-helix and the  $\alpha$ 7-helix, shows a stronger connection in the R7 subfamily (RGS9) and the RZ subfamily (RGS17 and RGS19) than in the R4 subfamily (RGS4 and RGS8). Overall, salt bridge analyses reveal differences in the interhelical interactions between the  $\alpha$ 4– $\alpha$ 5 and  $\alpha$ 4– $\alpha$ 7 pairs of helices (Fig. 3 A). Importantly, the salt bridges affecting the conformational flexibility of the  $\alpha$ 4– $\alpha$ 5 helical pair will lead to allosteric perturbations because one of these salt bridges is located near the conserved cysteine residue recognized by covalent inhibitors (e.g., E97-K110; Fig. 3) and the other



**FIGURE 3** Conserved salt bridges in RGS proteins. (A) Four salt bridges conserved across five RGS proteins are highlighted on the structure of RGS4, a canonical member of the RGS family. The conserved cysteine residue is highlighted as a cyan sphere and labeled C95. Three key helices connected via interhelical salt bridges are also colored uniquely in cartoon representations:  $\alpha_4$  (yellow),  $\alpha_5$  (green), and  $\alpha_7$  (magenta). (B) For conserved salt bridges, shown is the percentage occupancy (with error bars) computed based on fractional time of the simulation trajectory during which a given salt bridge was intact based on a distance criterion. The error bars were computed based on the block standard error analysis (see Methods) (38). The subscript “L” in salt bridge labels for helices indicates a loop connecting two helices. For example,  $(\alpha_3/\alpha_4)_L$  signifies the loop connecting the  $\alpha_3$  and  $\alpha_4$  helices. To see this figure in color, go online.

salt bridge is located near the protein-protein interface (e.g., E87-K125; Fig. 3). We further analyzed many unique and nonconserved salt bridges across all RGS proteins (see Supporting Results; Fig. S6).

### Community network in RGS proteins

By using data from MD simulations spanning 10  $\mu$ s, we obtained residue-residue correlation maps for all RGS proteins (Fig. S7). We then used the Girvan-Newman algorithm (48) to find communities of correlated residues, thereby revealing

the underlying community network (Fig. 4). In this algorithm (see Methods), the community structure is probed based on a key metric termed “edge betweenness” of an edge, which is the number of shortest pathways between pairs of vertices that run along it; the edge is a bridge between two nodes/vertices in a network. By definition, the edge betweenness for intercommunity nodes is higher because of the existence of unique shortest pathways and lower for intracomunity nodes because of many alternative pathways (48,51). We further hypothesized that intercommunity communication can be established through bridging via critical nodes leading to the propagation of perturbations originating at the shared cysteine residue to residues in the RGS-G $\alpha$  interface. We highlight such intercommunity connections in schematic maps shown in Fig. 4, where a thicker line indicates a stronger connection between communities.

We observed that RGS proteins from the same subfamily partition into the same number of communities, seven communities for the R4 subfamily and eight communities for the RZ subfamily. Moreover, RGS9, a member of the R7 subfamily, partitions into seven communities similar to the R4 subfamily. However, the composition of various communities varies across proteins from the same family or different family: 1) the community 1 (C1) contains the  $\alpha_1$  helix in all five RGS proteins but additionally the  $\alpha_2$  helix in RGS4, RGS9, and RGS19 and the  $\alpha_9$  helix in RGS8; 2) the community 2 (C2) contains the  $\alpha_4$ -helix in all five RGS proteins, although it may also involve parts of the  $\alpha_5$ -helix, as seen in RGS17, or the  $\alpha_7$ -helix, as seen in RGS4, RGS8, RGS9, and RGS19; 3) the community 3 (C3) only contains the  $\alpha_6$ -helix in all five RGS proteins; 4) the community 4 (C4) is composed of the  $\alpha_3$ -helix, the  $\alpha_8$ -helix, and/or the  $\alpha_9$ -helix, as seen in RGS4, RGS8, RGS9, and RGS17, but it contains fewer residues on the  $\alpha_8$  and  $\alpha_9$  helices in RGS19; 5) the community 5 (C5) contains the flexible  $\alpha_5$ - $\alpha_6$  interhelical loop located near the RGS/G $\alpha$  interface, as seen in the R4 subfamily (RGS4 and RGS8) and in the RZ subfamily (RGS17 and RGS19); however, in RGS9, C5 is distinct and it contains residues from the terminal  $\alpha_9$ -helix; 6) the community 6 (C6) contains parts of the  $\alpha_7$ -helix, the  $\alpha_8$ -helix, and/or the  $\alpha_9$  helix in RGS proteins; for example, in the RZ subfamily members, C6 contains the  $\alpha_7$ -helix in RGS17 but also includes the  $\alpha_8$  and  $\alpha_9$  helices in RGS19; 7) the community 7 (C7) mainly contains the  $\alpha_5$ -helix in all RGS proteins except in RGS17; and 8) the community 8 is only observed in the RZ subfamily members, in which in RGS17, it contains the terminal  $\alpha_9$ -helix, and in RGS19, it contains only one residue (P172) on the  $\alpha_6$ - $\alpha_7$  interhelical loop.

To understand the allosteric perturbations originating at the binding site of covalent inhibitors, it is useful to examine the links of C2 to other communities because the conserved cysteine residue on the  $\alpha_4$ -helix is located within C2. Particularly significant are links of C2 to communities harboring residues located in the RGS-G $\alpha$  protein-protein interface

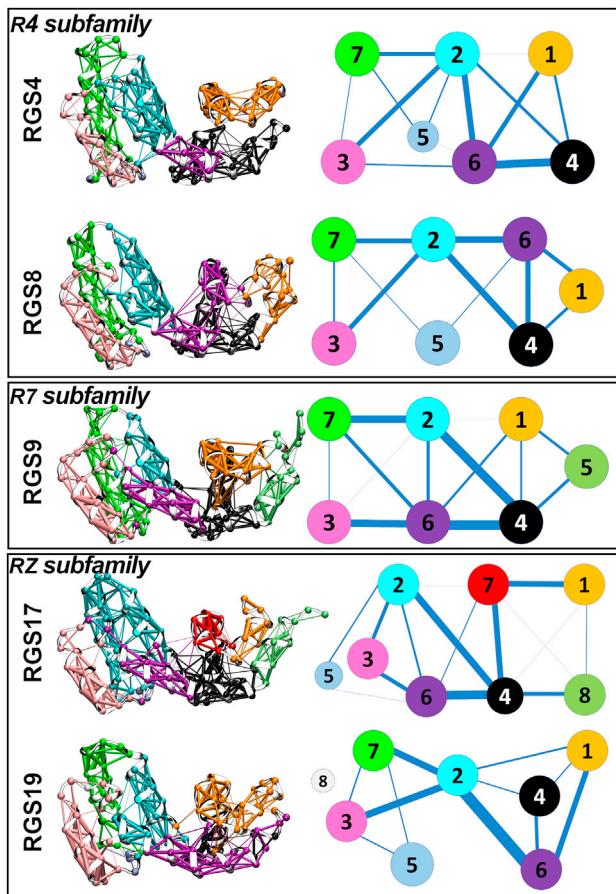


FIGURE 4 Community network in RGS proteins. The networks of communities and their schematic community maps are shown for RGS proteins, as organized by their subfamily memberships. The  $C_{\alpha}$ -atoms of residues forming each community are uniquely colored and mapped on respective protein structures. The communities are labeled 1 through 7 or 8 in corresponding community schematic maps. The cumulative edge betweenness is represented as the width of intercommunity links. Additional details about critical nodes and listing of residue memberships for each community are shown in Tables S1–S5. To see this figure in color, go online.

(e.g., C3 through C7). We observed a stronger communication between C2 and C3 in all RGS proteins except in RGS9, suggesting that the perturbations originating in C2 can be directly transmitted to C3 in the R4 and RZ subfamily members (RGS4, RGS8, RGS17, and RGS19). In RGS9, C2 can communicate with C3 via the C4–C6 bridge or via the C7–C6 bridge. We also observed in all RGS proteins that C2 communicates with C4 or C6 to a varying extent. Compared with RGS4 and RGS19, a stronger communication between C2 and C4 can be found in RGS8, RGS9, and RGS17. Similarly, C2 and C6 have a stronger communication in the R4 subfamily (RGS4 and RGS8) and one member of the RZ subfamily (RGS19) but not in other members of the RZ (RGS17) and R7 (RGS9) subfamilies. A direct communication between C2 and C5 in all RGS proteins is weak or not observed. Importantly, C2 serves as a hub of connectivity with several other communities in all RGS proteins.

## Allosteric communication pathways

Using the network of communities, we further analyzed allosteric pathways originating at the conserved cysteine residue in each RGS protein (termed as a source residue) and ending at six different residues in the RGS-G $\alpha$  interface (each termed as a sink). Among these six residues (Fig. 1 B; Fig. S3), one residue (Y84 in RGS4, F78 in RGS8, F321 in RGS9, Y106 in RGS17, and Y112 in RGS19) resides on the  $\alpha$ 3- $\alpha$ 4 loop and contacts the switch I region of the G $\alpha$  subunit; three residues (V127, N128, and S131 in RGS4 and counterparts in other RGS proteins; Fig. S2) are located in the  $\alpha$ 5- $\alpha$ 6 loop interacting with the switch II (residues V127 and N128 in RGS4) or switch III (residue S131 in RGS4) regions of G $\alpha$ ; and the remaining two residues, residing on the  $\alpha$ 7-helix and  $\alpha$ 8-helix, contact the switch I region of G $\alpha$  (Fig. 1 B; Fig. S3; (31)). These residues are key participants in the RGS-G $\alpha$  interface because they directly contact the G $\alpha$  subunit (Fig. S3; (52,53)). Importantly, mutations in Y84, N128, L159, and R167 in RGS4 significantly decreased the GAP activity of RGS4 (52). In addition, we also found that N128 and L159 in RGS4, N122 and L153 in RGS8, I363 in RGS9, and S150 in RGS17 are critical nodes (Tables S1–S5), which are of importance for intercommunity communication.

Employing the method developed by Sethi et al. (15) that transforms the residue-residue correlation data (Fig. S7) to the length of a pathway (see Methods), we discovered allosteric pathways from the source to sink residues and their corresponding lengths (Figs. 5 and S8). Given that we have one source residue (the conserved cysteine at the  $\alpha$ 4-helix; Fig. 1 B) and six sink residues (each located in the protein-protein interface; Fig. 1 B), we have obtained six unique pathways (termed P1–P6) for each RGS protein that originates at the source residue and terminates at each of the sink residues. We examined and concluded that the allosteric pathways cannot be predicted by trivial physical distance analysis, which predicts the order of pathways (shortest to longest) as P2 < P3 < P5 < P4 < P1 < P6, but the shortest allosteric pathway is P5 and the longest is P3 (RGS8, RGS9, and RGS19) or P4 (RGS4 and RGS17) (Fig. S8). We briefly describe all allosteric pathways below.

Among allosteric pathways, a shorter pathway length indicates that the sink residue is easily affected by the source residue. Based on the pathway length analysis, our rankings for pathways from the shortest to longest in five RGS proteins are as follows: 1) RGS17 < RGS8 < RGS9 < RGS19 < RGS4 (P1); 2) RGS19 < RGS8 < RGS4 < RGS17 < RGS9 (P2); 3) RGS4 < RGS17 < RGS9 < RGS19 < RGS8 (P3); 4) RGS19 < RGS4 < RGS8 < RGS9 < RGS17 (P4); 5) RGS9 < RGS8 = RGS17 < RGS4 < RGS19 (P5); and 6) RGS9 < RGS8 = RGS17 < RGS4 < RGS19 (P6). Examining these rankings for the fastest allosteric perturbation pathway (shortest pathway length) originating at the conserved cysteine

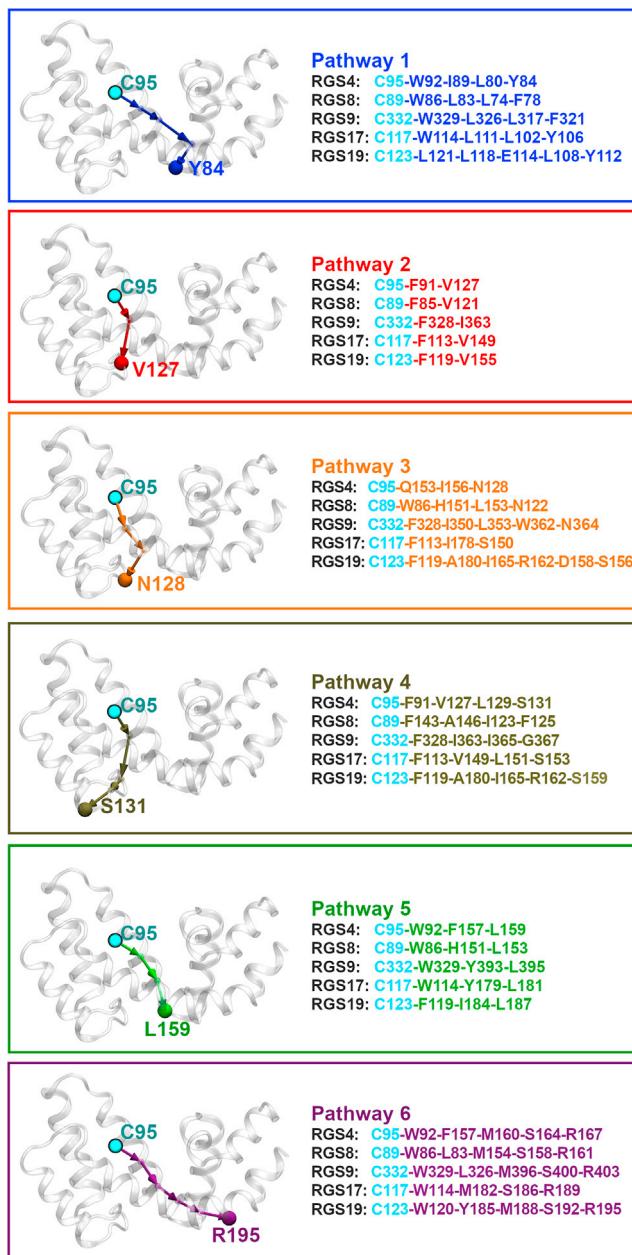


FIGURE 5 Allosteric pathways between source and sink residues. Shown are optimal pathways from the conserved cysteine residue (cyan sphere) to six sink residues. The panels shown depict each of the six pathways (*colored uniquely*) on the structure of RGS4 along with the details of residues for each pathway in all RGS proteins, where the first residue serves as a source residue and the last residue serves as a sink residue. See also Fig. S8. Additional pathways originating at a second cysteine residue conserved only among RGS4 and RGS8 are shown in Fig. S9 and discussed in Supporting Results. To see this figure in color, go online.

residue, which is the source residue and the binding site of covalent inhibitors, reveals that the fastest pathway to any of the sink residues located in the protein-protein interface is distinct in each of the three subfamilies of RGS proteins. For the pathways P1, P2, or P4, the fastest perturbations will be in the RZ subfamily; for the pathway P3, the fastest

perturbations will be in the R4 subfamily; and for the pathways P5 or P6, the fastest perturbations will be in the R7 subfamily.

These observations suggest that allosteric perturbations propagate 1) in the R4 subfamily via P3 that connects the source cysteine residue (C95 in RGS4) to a sink residue on the  $\alpha$ 5- $\alpha$ 6 loop (N128 in RGS4); 2) in the RZ subfamily via P1, which connects the source cysteine residue (C117 in RGS17) to a sink residue on the  $\alpha$ 3- $\alpha$ 4 loop (Y106 in RGS17), or via P2 and P4, which connect the source cysteine residue (C123 in RGS19) to two sink residues on the  $\alpha$ 5- $\alpha$ 6 loop (V155 and S159 in RGS19); and 3) in the R7 subfamily via P5 or P6, which connect the source cysteine residue (C332 in RGS9) to sink residues on the  $\alpha$ 7-helix (L395 in RGS9) or on the  $\alpha$ 8-helix (R403 in RGS9). Because the preferred pathways for each subfamily terminate at distinct structural motifs in RGS proteins that contact distinct regions in  $\text{G}\alpha$  subunits (the  $\alpha$ 3- $\alpha$ 4 loop contacts switch I, the  $\alpha$ 5- $\alpha$ 6 loop contacts switch III, and  $\alpha$ 7/ $\alpha$ 8 helices contact switch I/II; Fig. S3), a differential inhibitory effect is expected because of binding of covalent molecules at conserved cysteine residues.

## DISCUSSION

In this work, we have studied differences in dynamical features of five RGS proteins from three subfamilies, including the R4 subfamily (RGS4 and RGS8), the RZ subfamily (RGS17 and RGS19), and the R7 subfamily (RGS9). We highlight differences in sequences, structures, residue fluctuations, salt bridging interactions, and allosteric communities and pathways. We hypothesize that, collectively, these differences in dynamics of RGS proteins are correlated with the differential inhibitory effect observed in binding of covalent inhibitors at a conserved cysteine residue on the  $\alpha$ 4-helix of each RGS protein (Fig. 1).

We observed key differences in dynamics of a bundle of helices ( $\alpha$ 4,  $\alpha$ 5,  $\alpha$ 6, and  $\alpha$ 7) that are connected by three flexible loop motifs (the  $\alpha$ 4- $\alpha$ 5,  $\alpha$ 5- $\alpha$ 6, and  $\alpha$ 6- $\alpha$ 7 loops). Importantly, two of these loop motifs ( $\alpha$ 4- $\alpha$ 5 and  $\alpha$ 6- $\alpha$ 7) are located near the conserved cysteine residue, which serves as the binding site for covalent inhibitors, whereas the third loop motif ( $\alpha$ 5- $\alpha$ 6) is located in the RGS- $\text{G}\alpha$  protein-protein interface and therefore directly contacts the  $\text{G}\alpha$  subunit. These helices and loops are held together by several conserved salt bridges (Fig. 3), and the differential strength of these salt bridges underlie the flexibility of each RGS protein. For example, the salt bridges connecting the  $\alpha$ 4- $\alpha$ 5 helical pair and the  $\alpha$ 6- $\alpha$ 7 helical pair have a higher occupancy in proteins of the R7 or RZ subfamilies (RGS9, RGS17, and RGS19) in comparison to proteins of the R4 subfamily (RGS4 and RGS8).

Within the same subfamily (e.g., the R4 subfamily), we identified two nonconserved salt bridges (D136-K155 and R139-E151) connecting the  $\alpha$ 6-helix and the  $\alpha$ 7-helix in

RGS4, although no similar salt bridge pair was found in RGS8 (Fig. S6). However, in a different subfamily (e.g., the RZ subfamily), we identified again two nonconserved salt bridges (K168-D178 and K168-D179) connecting the  $\alpha$ 6-helix and the  $\alpha$ 7-helix in RGS19 (Fig. S6). The differential flexibilities of these helical and loop motifs due to distinct salt bridging interactions in RGS proteins are consistent with differential inhibitory effect (5,25,28,29) because of variability in the exposure of the side chains of conserved cysteine residues, as reported in our previous work (25). As an example, we showed in our previous work that introducing a negative charge (L111D) on the  $\alpha$ 4 helix of RGS19 resulted in a new salt bridging interaction with the  $\alpha$ 5- $\alpha$ 6 loop that increased the thermal stability of RGS19 and decreased inhibitor potency by severalfold, likely because of difficulties in inhibitor access to the side chain of the conserved cysteine residue on the  $\alpha$ 4-helix (50).

We also compared pairs of RGS proteins across all three subfamilies using a difference cross correlation analysis (Fig. S10). On comparing proteins from the same subfamily, for example the R4 subfamily (RGS8 versus RGS4), in RGS8, we observed a decreased correlation between the  $\alpha$ 6- $\alpha$ 7 loop and the  $\alpha$ 4/ $\alpha$ 7 helices, whereas an increased correlation between the  $\alpha$ 6- $\alpha$ 7 loop and the  $\alpha$ 6 helix. Similarly, comparing proteins of the RZ subfamily showed in RGS19 a decreased correlation between the  $\alpha$ 4- $\alpha$ 5 and  $\alpha$ 6- $\alpha$ 7 loops. Among different subfamily members (RGS9/RGS17/RGS19 versus RGS4), we observed increased correlations between the  $\alpha$ 6 and  $\alpha$ 7 helices in RGS9, no significant change in correlations for RGS17, and decreased correlations between the  $\alpha$ 4- $\alpha$ 5 and  $\alpha$ 6- $\alpha$ 7 loops in RGS19. However, using the R4 subfamily member RGS8 as a reference, we found increased correlations between the  $\alpha$ 6 and  $\alpha$ 7 helices in RGS9, increased correlations between the  $\alpha$ 4- $\alpha$ 5 and  $\alpha$ 6- $\alpha$ 7 loops in RGS17, and no significant change in correlations for RGS19.

From our residue-residue correlation data (Fig. S7), we also observed that the motions in the  $\alpha$ 4-helix and the  $\alpha$ 7-helix are highly correlated in all RGS proteins, although this correlation is weaker in RGS19. The observation of weaker correlations in the  $\alpha$ 4/ $\alpha$ 7 helical pair of RGS19 relative to other RGS proteins is consistent with the higher hydrogen-deuterium exchange (HDX) rates in these motifs of RGS19 and lower HDX rates of these motifs in other RGS proteins (e.g., RGS4 and RGS8), as reported in our previous work (25). Furthermore, weaker correlations and higher HDX rates in the  $\alpha$ 4/ $\alpha$ 7 helical pair of RGS19 suggest easier accessibility of the conserved inhibitor binding cysteine residue located on the  $\alpha$ 4 helix, which is consistent with the observation that RGS19 is more potently inhibited by CCG-50014 than RGS4/RGS8 (25). Collectively, these differences highlight that because of differential dynamics in helical and loop motifs surrounding the inhibitor binding sites, the access to inhibitors is distinct among RGS proteins of the same or different subfamilies, and as a result, the inhibitory effect is distinct.

Our analyses of a network of communities within RGS proteins further revealed that the community C2, which harbors the inhibitor binding site (a conserved cysteine residue), forms a hub of connectivity with many other communities in RGS proteins. For example, C2 and C3 had stronger connectivity in the R4 and RZ subfamilies but not in the R7 subfamily. Moreover, a stronger communication between C2 and C4 was found in at least one protein member of each subfamily (RGS8 in the R4 subfamily; RGS9 in the R7 subfamily; and RGS17 in the RZ subfamily), whereas we did not observe any direct communication between C2 and C5. Furthermore, our analyses of allosteric pathways highlighted unique pathways, along which allosteric perturbations propagate from the inhibitor binding site to residues in the protein-protein interface. For the pathways P1, P2, or P4, the fastest perturbations are in the RZ subfamily; for the pathway P3, the fastest perturbations are in the R4 subfamily; and for the pathways P5 or P6, the fastest perturbations in the R7 subfamily. These differences suggest that the binding of covalent inhibitors to RGS proteins differentially perturbs distinct regions (switch I, II, and III) in the  $\text{G}\alpha$  subunit, thereby resulting in a distinct inhibitory effect.

Additionally, besides the conserved cysteine residue that serves as a binding site of allosteric inhibitors in RGS proteins, we found that two more conserved residues that lined most of the pathways may be important for allosteric regulation: F91 and W92 in RGS4, F85 and W86 in RGS8, F328 and W329 in RGS9, F113 and W114 in RGS17, and F119 and W120 in RGS19. This observation is consistent with our previous NMR data that showed significant perturbations in the residue F91 in RGS4 on CCG-50014 binding (31). However, the importance of these two residues may vary in different RGS proteins. For instance, four optimal pathways in RGS8 (P1, P3, P5, and P6) crossed the residue W86, whereas four optimal pathways in RGS19 (P2, P3, P4, and P5) crossed the residue F119. Because both of these residues are located away from the RGS- $\text{G}\alpha$  protein-protein interface, targeting them (potentially using noncovalent compounds) may provide an alternative route to achieve allosteric modulation in RGS proteins. It is supported by the fact that a binding pocket for inhibitors has been proposed near the phenylalanine residue previously (30). Moreover, we have shown that noncovalent analogs of TDZD compounds can dock and stably reside in the vicinity of these two residues (24). Therefore, our findings are potentially useful in the future design of inhibitors with enhanced selectivity among protein members of the RGS family.

We also point out that although the community partitioning analyses can vary to some extent among independent simulations or analyses based on the first and second halves of an MD trajectory or a cutoff distance, the allosteric pathways including the shortest pathways remain largely preserved (Figs. S11–S15). Furthermore, we note that the communities and allosteric pathways reported in our work

are computed based on the cross correlation functions consistent with the previous work (15,17–22). However, in future applications of these methods, it may be useful to examine the limitations of established correlation analysis, as highlighted by a previous study (54). It may also be useful to utilize information theory-based approaches to study allosteric mechanisms (55).

## CONCLUSIONS

We have studied the differences in dynamics of five RGS proteins (RGS4, RGS8, RGS9, RGS17, and RGS19) from three subfamilies (R4, R7, and RZ) with an aim to probe their differential inhibition by covalent inhibitors that target a conserved cysteine residue located on the  $\alpha$ 4-helix in each protein. Via analyses of residue fluctuations, salt bridging interactions, allosteric communities, and pathways in RGS proteins, we highlight differences in dynamics of helical and loop motifs surrounding the inhibitor binding site and near the RGS-G $\alpha$  protein-protein interface. Our results reveal that preferred allosteric pathways exist among RGS members from distinct families that allow the propagation of allosteric signals from the inhibitor binding site to distinct regions in the RGS-G $\alpha$  protein-protein interface. We also suggest another pair of conserved residues on the  $\alpha$ 4-helix (a Phe and a Trp residue) as potential docking sites for noncovalent inhibitors, given that these two residues lined several allosteric pathways.

## SUPPORTING MATERIAL

Supporting Material can be found online at <https://doi.org/10.1016/j.bj.2020.12.010>.

## AUTHOR CONTRIBUTIONS

Y.L. and H.V. designed the research. Y.L. performed the research and analyzed data. Y.L. and H.V. wrote the article.

## ACKNOWLEDGMENTS

We acknowledge financial support provided by the National Science Foundation under grant CHE-1508595 (H.V.) and the National Institutes of Health (NIH) Centers of Biomedical Research Excellence Center of Integrated Biomedical and Bioengineering Research (grant P20GM113131) through an Institutional Development Award (to H.V.). We also acknowledge financial support from the NIH under grant R35GM138217 (H.V.). Research reported in this publication was supported in part by the National Institute Of General Medical Sciences of the NIH under Award Numbers P20GM113131 and R35GM138217. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. We thank computational support through the following resources: Premise, a central shared high-performance computing cluster at University of New Hampshire supported by the Research Computing Center; Bio-Made, a heterogeneous CPU/GPU cluster supported by the NSF EPSCoR award (OIA-1757371); and the NSF-supported (ACI-1548562) Extreme Science and Engineering Discovery Environment (XSEDE) (56) Comet

resource at San Diego Supercomputer Center under grant TG-MCB160183 (H.V.).

## REFERENCES

1. Jones, S., and J. M. Thornton. 1996. Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. USA.* 93:13–20.
2. Neubig, R. R., and D. P. Siderovski. 2002. Regulators of G-protein signalling as new central nervous system drug targets. *Nat. Rev. Drug Discov.* 1:187–197.
3. Wells, J. A., and C. L. McClendon. 2007. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature.* 450:1001–1009.
4. Arkin, M. R., Y. Tang, and J. A. Wells. 2014. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chem. Biol.* 21:1102–1114.
5. Sjögren, B. 2017. The evolution of regulators of G protein signalling proteins as drug targets - 20 years in the making: IUPHAR Review 21. *Br. J. Pharmacol.* 174:427–437.
6. Kessler, D., M. Gmachl, ..., D. B. McConnell. 2019. Drugging an undruggable pocket on KRAS. *Proc. Natl. Acad. Sci. USA.* 116:15823–15829.
7. Christopoulos, A. 2002. Allosteric binding sites on cell-surface receptors: novel targets for drug discovery. *Nat. Rev. Drug Discov.* 1:198–210.
8. Dokholyan, N. V. 2016. Controlling allosteric networks in proteins. *Chem. Rev.* 116:6463–6487.
9. Wentur, C. J., P. R. Gentry, ..., C. W. Lindsley. 2014. Drugs for allosteric sites on receptors. *Annu. Rev. Pharmacol. Toxicol.* 54:165–184.
10. Hopkins, A. L., and C. R. Groom. 2002. The druggable genome. *Nat. Rev. Drug Discov.* 1:727–730.
11. Lindsley, C. W. 2014. 2013 Philip S. Portoghese Medicinal Chemistry Lectureship: drug discovery targeting allosteric sites. *J. Med. Chem.* 57:7485–7498.
12. Nussinov, R., and C.-J. Tsai. 2015. The design of covalent allosteric drugs. *Annu. Rev. Pharmacol. Toxicol.* 55:249–267.
13. Roth, G. J., N. Stanford, and P. W. Majerus. 1975. Acetylation of prostaglandin synthase by aspirin. *Proc. Natl. Acad. Sci. USA.* 72:3073–3076.
14. Papaleo, E., G. Saladino, ..., R. Nussinov. 2016. The role of protein loops and linkers in conformational dynamics and allostery. *Chem. Rev.* 116:6391–6423.
15. Sethi, A., J. Eargle, ..., Z. Luthey-Schulten. 2009. Dynamical networks in tRNA:protein complexes. *Proc. Natl. Acad. Sci. USA.* 106:6620–6625.
16. VanWart, A. T., J. Eargle, ..., R. E. Amaro. 2012. Exploring residue component contributions to dynamical network models of allostery. *J. Chem. Theory Comput.* 8:2949–2961.
17. Chopra, N., T. E. Wales, ..., A. H. Andreotti. 2016. Dynamic allostery mediated by a conserved tryptophan in the Tec family kinases. *PLoS Comput. Biol.* 12:e1004826.
18. Kornev, A. P., and S. S. Taylor. 2015. Dynamics-driven allostery in protein kinases. *Trends Biochem. Sci.* 40:628–647.
19. McClendon, C. L., A. P. Kornev, ..., S. S. Taylor. 2014. Dynamic architecture of a protein kinase. *Proc. Natl. Acad. Sci. USA.* 111:E4623–E4631.
20. Ahuja, L. G., A. P. Kornev, ..., S. S. Taylor. 2017. Mutation of a kinase allosteric node uncouples dynamics linked to phosphotransfer. *Proc. Natl. Acad. Sci. USA.* 114:E931–E940.
21. Srivastava, A. K., L. R. McDonald, ..., G. Veglia. 2014. Synchronous opening and closing motions are essential for cAMP-dependent protein kinase A signaling. *Structure.* 22:1735–1743.

22. Meng, Y., L. G. Ahuja, ..., B. Roux. 2018. A catalytically disabled double mutant of Src tyrosine kinase can be stabilized into an active-like conformation. *J. Mol. Biol.* 430:881–889.
23. Mohammadiarani, H., V. S. Shaw, ..., H. Vashisth. 2018. Interpreting hydrogen-deuterium exchange events in proteins using atomistic simulations: case studies on regulators of G-protein signaling proteins. *J. Phys. Chem. B.* 122:9314–9323.
24. Mohammadi, M., H. Mohammadiarani, ..., H. Vashisth. 2019. Interplay of cysteine exposure and global protein dynamics in small-molecule recognition by a regulator of G-protein signaling protein. *Proteins*. 87:146–156.
25. Shaw, V. S., H. Mohammadiarani, ..., R. R. Neubig. 2018. Differential protein dynamics of regulators of G-protein signaling: role in specificity of small-molecule inhibitors. *J. Am. Chem. Soc.* 140:3454–3460.
26. Sjögren, B., L. L. Blazer, and R. R. Neubig. 2010. Regulators of G protein signaling proteins as targets for drug discovery. *Prog. Mol. Biol. Transl. Sci.* 91:81–119.
27. Sjögren, B., and R. R. Neubig. 2010. Thinking outside of the "RGS box": new approaches to therapeutic targeting of regulators of G protein signaling. *Mol. Pharmacol.* 78:550–557.
28. Turner, E. M., L. L. Blazer, ..., S. M. Husbands. 2012. Small molecule inhibitors of regulators of G Protein signaling (RGS) proteins. *ACS Med. Chem. Lett.* 3:146–150.
29. Blazer, L. L., D. L. Roman, ..., R. R. Neubig. 2010. Reversible, allosteric small-molecule inhibitors of regulator of G protein signaling proteins. *Mol. Pharmacol.* 78:524–533.
30. Blazer, L. L., H. Zhang, ..., R. R. Neubig. 2011. A nanomolar-potency small molecule inhibitor of regulator of G-protein signaling proteins. *Biochemistry*. 50:3181–3192.
31. Vashisth, H., A. J. Storaska, ..., C. L. Brooks, III. 2013. Conformational dynamics of a regulator of G-protein signaling protein reveals a mechanism of allosteric inhibition by a small molecule. *ACS Chem. Biol.* 8:2778–2784.
32. Hayes, M. P., C. R. Bodle, and D. L. Roman. 2018. Evaluation of the selectivity and cysteine dependence of inhibitors across the regulator of G protein-signaling family. *Mol. Pharmacol.* 93:25–35.
33. Phillips, J. C., R. Braun, ..., K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–1802.
34. MacKerell, A. D., D. Bashford, ..., M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.
35. Vanommeslaeghe, K., E. Hatcher, ..., A. D. MacKerell, Jr. 2010. CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* 31:671–690.
36. Huang, J., S. Rauscher, ..., A. D. MacKerell, Jr. 2017. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods*. 14:71–73.
37. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27–28..
38. Grossfield, A., and D. M. Zuckerman. 2009. Quantifying uncertainty and sampling quality in biomolecular simulations. *Annu. Rep. Comput. Chem.* 5:23–48.
39. Wodak, S. J., E. Paci, ..., T. McLeish. 2019. Allostery in its many guises: from theory to applications. *Structure*. 27:566–578.
40. Gautier, C., L. Laursen, ..., S. Gianni. 2018. Seeking allosteric networks in PDZ domains. *Protein Eng. Des. Sel.* 31:367–373.
41. Lockless, S. W., and R. Ranganathan. 1999. Evolutionarily conserved pathways of energetic connectivity in protein families. *Science*. 286:295–299.
42. Halabi, N., O. Rivoire, ..., R. Ranganathan. 2009. Protein sectors: evolutionary units of three-dimensional structure. *Cell*. 138:774–786.
43. Igaev, M., and H. Grubmüller. 2018. Microtubule assembly governed by tubulin allosteric gain in flexibility and lattice induced fit. *eLife*. 7:e34353.
44. Bowman, G. R., and P. L. Geissler. 2012. Equilibrium fluctuations of a single folded protein reveal a multitude of potential cryptic allosteric sites. *Proc. Natl. Acad. Sci. USA*. 109:11681–11686.
45. Lin, M. M. 2016. Timing correlations in proteins predict functional modules and dynamic allostery. *J. Am. Chem. Soc.* 138:5036–5043.
46. Kong, Y., and M. Karplus. 2009. Signaling pathways of PDZ2 domain: a molecular dynamics interaction correlation analysis. *Proteins*. 74:145–154.
47. Glykos, N. M. 2006. Software news and updates. Carma: a molecular dynamics analysis program. *J. Comput. Chem.* 27:1765–1768.
48. Girvan, M., and M. E. Newman. 2002. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA*. 99:7821–7826.
49. Newman, M. E., and M. Girvan. 2004. Finding and evaluating community structure in networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 69:026113.
50. Shaw, V. S., M. Mohammadi, ..., R. R. Neubig. 2019. An interhelical salt bridge controls flexibility and inhibitor potency for regulators of G-protein signaling proteins 4, 8, and 19. *Mol. Pharmacol.* 96:683–691.
51. Guo, J., and H.-X. Zhou. 2016. Protein allostery and conformational dynamics. *Chem. Rev.* 116:6503–6515.
52. Srinivasa, S. P., N. Watson, ..., K. J. Blumer. 1998. Mechanism of RGS4, a GTPase-activating protein for G protein alpha subunits. *J. Biol. Chem.* 273:1529–1533.
53. Kosloff, M., A. M. Travis, ..., V. Y. Arshavsky. 2011. Integrating energy calculations with functional assays to decipher the specificity of G protein-RGS protein interactions. *Nat. Struct. Mol. Biol.* 18:846–853.
54. Lange, O. F., and H. Grubmüller. 2006. Generalized correlation for biomolecular dynamics. *Proteins*. 62:1053–1061.
55. LeVine, M. V., and H. Weinstein. 2014. NbIT—a new information theory-based analysis of allosteric mechanisms reveals residues that underlie function in the leucine transporter LeuT. *PLoS Comput. Biol.* 10:e1003603.
56. Towns, J., T. Cockerill, ..., N. Wilkins-Diehr. 2014. XSEDE: accelerating scientific discovery. *Comput. Sci. Eng.* 16:62–74.