Domain-Opening and Dynamic Coupling in the α-Subunit of Heterotrimeric G Proteins

Xin-Qiu Yao and Barry J. Grant*
Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan

ABSTRACT Heterotrimeric G proteins are conformational switches that turn on intracellular signaling cascades in response to the activation of G-protein-coupled receptors. Receptor activation by extracellular stimuli promotes a cycle of GTP binding and hydrolysis on the G protein α-subunit (Gα). Important conformational transitions occurring during this cycle have been characterized from extensive crystallographic studies of Gα. However, the link between the observed conformations and the mechanisms involved in G-protein activation and effector interaction remain unclear. Here we describe a comprehensive principal component analysis of available Gα crystallographic structures supplemented with extensive unbiased conventional and accelerated molecular dynamics simulations that together characterize the response of Gα to GTP binding and hydrolysis. Our studies reveal details of activating conformational changes as well as the intrinsic flexibility of the α-helical domain that includes a large-scale 60° domain opening under nucleotide-free conditions. This result is consistent with the recently reported open crystal structure of Gs, the stimulatory G protein for adenylyl cyclase, in complex with the α2 adrenergic receptor. Sets of unique interactions potentially important for the conformational transition are also identified. Moreover simulations reveal nucleotide-dependent dynamical couplings of distal regions and residues potentially important for the allosteric link between functional sites.

Received for publication 12 April 2013 and in final form 5 June 2013.

*Correspondence: bjgrant@umich.edu

Heterotrimeric G proteins undergo cycles of GTP-dependent conformational rearrangements and alterations of their oligomeric $\alpha\beta\gamma$ form to convey receptor signals to downstream effectors that control diverse cellular processes ranging from movement to division and differentiation. Interaction with activated receptor promotes the exchange of GDP for GTP on the G protein α subunit (Gα) and its separation from its $\beta\gamma$ subunit partners (G$\beta\gamma$). Both isolated Gα and G$\beta\gamma$ then interact with downstream effectors. GTP hydrolysis deactivates Gα, which reassociates with G$\beta\gamma$, becoming ready to restart the cycle. Each of these stages has been subjected to extensive crystallographic studies with high-resolution structures of Gα in complex with GDP, GTP analog, G$\beta\gamma$, and, most recently, the G-protein-coupled receptors now available. These studies have provided extensive mechanistic insight. However, a number of important questions remain, including:

How do the distinct conformations evident in the accumulated structures interconvert?

How do disease-associated mutations affect the fidelity of these transitions?

And, critically, how do distal functional sites responsible for nucleotide and protein partner binding allosterically coordinate their activities?

Here we describe a comprehensive analysis of the accumulated Gα crystallographic structures supplemented with extensive conventional (cMD) and accelerated molecular dynamics (aMD) simulations (1) that together map the structural and dynamical features of Gα in different nucleotide states. These enhanced sampling simulations reveal the spontaneous interconversion between GDP and GTP conformations and also characterize large-scale opening motions of the α-helical domain (HD) that were not accessible to previous simulation studies (2–5). Furthermore, the current simulations results reveal a distinctive pattern of collective motions that provide evidence for a nucleotide-dependent network of dynamic communication between the active site and the receptor and effector binding sites.

Principal component analysis of 53 Gα experimental structures homologous to transducin (Gαt) reveals that the major variation in accumulated structures is the concerted association/disassociation of three nucleotide-binding site loops termed the switch regions (SI, SII, and SIII). An additional small-scale (<10°) rotation of the HD relative to the main catalytic Ras-like domain (RasD) is also apparent (see Fig. S1 in the Supporting Material). The distinct conformation of SI–SIII regions gives rise to nucleotide-associated segregation of GDP- and GTP-analog-bound experimental structures along the PC1-PC2 plane. Interestingly, both GDP- and GTP-bound structures display a skewed distribution along the PC1-PC2 plane that arises from HD rotation. In comparison, the distribution of the GTP-bound structures becomes more restricted and the skew decreases when the

Editor: Nathan Baker.

http://dx.doi.org/10.1016/j.bpj.2013.06.006

mapping is based on a principle component analysis that excludes the HD region (see Fig. S1).

Recently, the HD region of Gαs (the α-subunit of the stimulatory G protein for adenyl cyclase) was shown to adopt a dramatically more open conformation in a crystal structure complex with the β2 adrenergic receptor (β2AR) (6). This clam-shell-like 127° opening in the absence of nucleotide and presence of receptor is consistent with electron microscopy (7) and double electron-electron resonance analysis (8). These results, together with recent hydrogen-deuterium exchange mass spectrometry data (9), indicate that there may be additional functional motions and inherent flexibility in the ensemble of native states beyond those apparent in the accumulated crystal structures of Gαt (9). To address this question, we performed multiple 100-ns aMD simulations of nucleotide-free Gαt. These simulations reveal a spontaneous large-scale opening and closing motion of larger magnitude (>60°) than those evident in the distribution of crystallographic structures (Fig. 1 *A* and see Fig. S2). In addition, the trajectory reveals two dominant modes of HD opening: an out-of-plane shifting (PC1 in Fig. S3) and an in-plane rotation (PC2 in Fig. S3). It is also notable that nucleotide-free aMD simulations sample both active (GTP-like) and inactive (GDP-like) structures (see Fig. S2) in an analogous manner to the spontaneous GDP to GTP interconversion sampled for Ras and Rho small G proteins with similar methods (10–12).

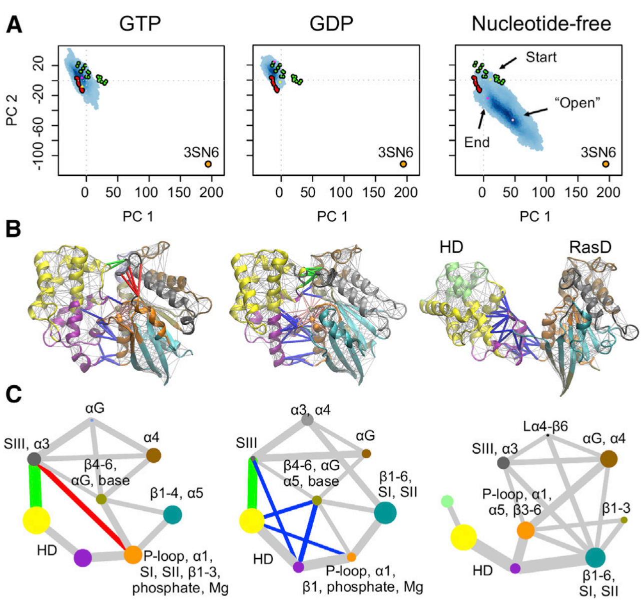

FIGURE 1 Nucleotide-associated differences in flexibility and dynamic coupling. (*A*) Mapping aMD simulation trajectories (*blue points*) onto the principal components obtained from analysis of Gα crystallographic GDP-bound (*green*) and GTP-analog bound (*red*) experimental structures. (*Orange*) Open β2AR-Gαs complex structure. (*B*) Results of dynamic coupling analysis mapped onto the average structure for each nucleotide state. (*Spheres*) Nodes for the nucleotide; the protein cartoon is colored by community structure. (*C*) Community network graph. (*Circles*) Communities, colored as in panel *B*. Radius of the circle indicates the number of residues in the community. Thickness of linking lines is determined by the maximum betweenness of the respective intercommunity edges (see the Supporting Material). (*Red*, *blue*, and *green edges*) Major topological difference between states.

The low sequence identity between Gαt and Gαs (44.5%), as well as the absence of the receptor and Gβγ in the simulations, may explain the difference between the predicted ~60° Gαt-HD rotation and that displayed in the β2AR-Gαs crystallographic structure (see Fig. S3). It is notable that, although the amplitude is much smaller, aMD simulations with bound nucleotide display similar dominant HD motions to those observed in the nucleotide-free simulations (see Fig. S4). This suggests that the interdomain flexibility of RasD and HD is likely an intrinsic feature of Gαt regardless of nucleotide state.

The transition between distinct conformations (structural clusters; see Fig. S5) was observed to correspond to significant dynamical changes in side-chain contacts (see Fig. S6). Specifically, we found sequential contacts breaking during the HD in-plane rotation motion starting from the region between HD helix αD and RasD helix αG toward that between HD helix αE and RasD SIII and the P-loop. In comparison, for the out-of-plane shift, we found simultaneous breaking and formation of contacts in the region containing the loop between helices αB and αC, the N-terminus of αA, αE, and αF of HD; α1, SI, and the loop between strand β6 and helix α5 of RasD. Interactions highlighted in these regions as potentially important for the conformational transitions include D137::K276, S140::K273, S140::D227, Q143::R238, N145::E39, and D146::K266, the effect of which can be further evaluated by mutagenesis experiments and simulations.

Dynamic network analysis methods developed by Sethi et al. (13) were used to examine whether the motions of one residue were correlated to the motions of another (distant) residue. In this approach, a weighted graph is constructed where each residue represents a node and the weight of the connection between nodes represents their respective correlation value. A clustering of edges is then used to define local communities of highly correlated residues that represent substructures that are highly intraconnected, but loosely interconnected. Applying this approach to multiple 40-ns cMD simulations initiated from GTP-, GDP-, and aMD-derived APO conformations revealed a consistent community composition as well as a distinct pattern of intercommunity connection between nucleotide states (Fig. 1, *B* and *C*).

The dynamics of the RasD region can be decomposed into two main communities that stem from the nucleotide base and phosphate regions in GDP and GTP states: The first community is composed of residues from the P-loop, helix α1, strands β1–β3, and the phosphates of the nucleotide (*orange* in Fig. 1, *B* and *C*). The second community comprises residues from helix αG, strands β4–β6, and the nucleotide base region (*tan* in Fig. 1, *B* and *C*). This dynamic partitioning of the central β-sheet and central role of the nucleotide is consistent with the bilobal structure and dynamics previously reported for Ras (14). In the presence of GTP, the first community includes or is dynamically coupled to SI, SII, and SIII regions (see the *orange node* and the *red edge* in Fig. 1 *C*). Removal of the γ-phosphate of GTP disrupts this region, leading to decoupling of the switch regions from the nucleotide. Also evident for GDP states is an apparent tighter coupling of RasD and HD regions (*blue edges* in Fig. 1 *C*). We note that these findings are robust to the choice of initial simulation conditions and are observed in both cMD and aMD simulations (see Fig. S7 and Fig. S8). Nucleotide-free Gαt simulations display an altered dynamical network with respect to those of nucleotide bound states. In particular, RasD and HD regions lose connecting edges consistent with the large-scale opening of these domains (e.g., SIII-HD *green edges* in Fig. 1 *C*).

A number of residues highlighted here as potentially important for mediating the coupling between prominent communities (see Table S1 in the Supporting Material) have been shown by previous mutagenesis studies to affect GDP release. For example, the double mutation A322S/R174M was found to significantly enhance the rate of GDP release (15). The current results indicate that these positions are involved in coupling the nucleotide and RasD. Also, mutations R144A and L232Q caused a faster basal GDP release rate in Gαi1 (16). The current analysis indicates that the equivalent positions in Gαt (S140 and M228) couple the RasD and HD, and suggests that their mutation could promote domain-domain motions. We also note the apparent coupling of α5 with the nucleotide base and Ploop-β1 with the phosphate regions of GDP. These direct connections of the receptor connecting N- and C-terminus to GDP are suggestive of potential routes for receptor-mediated GDP release. We expect further study of these sites and of receptor-bound dynamics to be informative in this regard.

In conclusion, simulations suggest a flexible HD in Gαt similar to that found for Gαs. In particular, in the absence of nucleotide we observed the spontaneous large-scale opening and closing of HD relative to RasD, which was unseen in previous computational studies. Moreover, we found that the functional states of Gαt are associated with the distinct dynamical couplings of functional regions including SI–SIII, P-loop, α5, and the HD region. Finally, our results indicate that nucleotide may not directly induce large-scale conformational changes but, instead, act as a modulator of intrinsically accessible conformations and as a central participant in their associated dynamical couplings.

SUPPORTING MATERIAL

Eight figures, and two tables are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(13)00680-2.

ACKNOWLEDGMENTS

We thank Drs. G. Scarabelli and R. K. Sunahara for valuable discussions and we gratefully acknowledge support from the University of Michigan.

REFERENCES and FOOTNOTES

1. Hamelberg, D., J. Mongan, and J. A. McCammon. 2004. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J. Chem. Phys.* 120:11919–11929.
2. Ceruso, M. A., X. Periole, and H. Weinstein. 2004. Molecular dynamics simulations of transducin: interdomain and front to back communication in activation and nucleotide exchange. *J. Mol. Biol.* 338:469–481.
3. Khafizov, K., G. Lattanzi, and P. Carloni. 2009. G protein inactive and active forms investigated by simulation methods. *Proteins* 75:919–930.
4. Louet, M., D. Perahia, …, N. Floquet. 2011. A concerted mechanism for opening the GDP binding pocket and release of the nucleotide in heterotrimeric G-proteins. *J. Mol. Biol.* 411:298–312.
5. Raimondi, F., M. Seeber, …, F. Fanelli. 2008. Mechanisms of inter- and intramolecular communication in GPCRs and G proteins. *J. Am. Chem. Soc.* 130:4310–4325.
6. Rasmussen, S. G., B. T. DeVree, …, B. K. Kobilka. 2011. Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature* 477:549–555.
7. Westfield, G. H., S. G. Rasmussen, …, G. Skiniotis. 2011. Structural flexibility of the Gαs α-helical domain in the β2-adrenoceptor Gs complex. *Proc. Natl. Acad. Sci. USA* 108:16086–16091.
8. Van Eps, N., A. M. Preininger, …, W. L. Hubbell. 2011. Interaction of a G protein with an activated receptor opens the interdomain interface in the α-subunit. *Proc. Natl. Acad. Sci. USA* 108:9420–9424.
9. Chung, K. Y., S. G. Rasmussen, …, R. K. Sunahara. 2011. Conformational changes in the G protein Gs induced by the β2 adrenergic receptor. *Nature* 477:611–615.
10. Grant, B. J., A. A. Gorfe, and J. A. McCammon. 2009. Ras conformational switching: simulating nucleotide-dependent conformational transitions with accelerated molecular dynamics. *PLOS Comput. Biol.* 5:e1000325.
11. Grant, B. J., J. A. McCammon, and A. A. Gorfe. 2010. Conformational selection in G-proteins: lessons from Ras and Rho. *Biophys. J.* 99:L87–L89.
12. Ortiz-Sanchez, J. M., S. E. Nichols, …, B. J. Grant. 2012. Identification of potential small molecule binding pockets on Rho family GTPases. *PLoS ONE* 7:e40809.
13. Sethi, A., J. Eargle, …, Z. Luthey-Schulten. 2009. Dynamical networks in tRNA:protein complexes. *Proc. Natl. Acad. Sci. USA* 106:6620–6625.
14. Gorfe, A. A., B. J. Grant, and J. A. McCammon. 2008. Mapping the nucleotide and isoform-dependent structural and dynamical features of Ras proteins. *Structure* 16:885–896.
15. Zielinski, T., A. J. Kimple, …, R. G. Lowery. 2009. Two Gα(i1) rate-modifying mutations act in concert to allow receptor-independent, steady-state measurements of RGS protein activity. *J. Biomol. Screen.* 14:1195–1206.
16. Remmers, A. E., C. Engel, …, R. R. Neubig. 1999. Interdomain interactions regulate GDP release from heterotrimeric G proteins. *Biochemistry* 38:13795–13800.

Supporting Material

Domain-opening and dynamic coupling in the alpha subunit of heterotrimeric G proteins

Xin-Qiu Yao[*] and Barry J. Grant[*§]

[*]Department of Computational Medicine and Bioinformatics,

University of Michigan, Ann Arbor, MI 48109, USA

[§]Correspondence to B. J. G. bjgrant@umich.edu

Methods

Unless otherwise indicated, all crystallographic structure and post simulation analysis was performed with the Bio3D package (1) available from < http://thegrantlab.org/bio3d/ >. Atomic coordinates for all crystallographic structures homologues to the α-subunit of transducin (Gαt) were obtained from the RCSB Protein Data Bank (2). Structures with unresolved residues in the switch regions were excluded from analysis leading to a dataset containing 53 structural species (See Table S2 for full details). Prior to assessing the variability of the structures, iterated rounds of structural superposition were performed to identify the most structurally invariant region. During the procedure, residues with the largest positional differences (measured as the volume of an ellipsoid determined from the Cartesian coordinates of the Cα atoms) were removed, before each round of superposition, until only invariant core residues remained (3). The identified "core" structure, which consisted of residues 32-52 (β1-Ploop-α1), 195 (β3), 216-226 (β4), 239 and 242-247 (α3), 260-274 (β5-αG), 279 and 282-283 (the loop between αG and α4), 295-304 (α4), and 317-336(β6-α5), was used as the reference frame for the superposition of both crystal structures and conformations from MD simulations.

Atomic coordinates for the Gs protein α subunit (Gαs) from the β2AR-Gs complex structure (PDB ID: 3SN6; (4)) were also used to evaluate simulation results. Structural information for missing residues in Gαs was generated with MODELLER 9.11 (5), using the average structure of the Gαt dataset as a template.

Molecular dynamics

All simulations were performed with the AMBER12 package (6) and the all-atom force field ff99SB (7). Additional parameters for guanine nucleotides were taken from Meagher et al. (8). The Mg^{2+}·GDP-bound transducin crystal structure (PDB ID: 1TAG; (9)) was employed as the starting point for both GDP bound and nucleotide free simulations. In addition, the Mg^{2+}·GSP (PDB ID: 1TND; (10)) structure was used as the start for GTP bound simulations, where the sulfur atom (S1γ) in the GTP analog, GSP (5'-guanosine-diphosphate-monothiophosphate), was replaced with the corresponding oxygen (O1γ) of GTP. In our model, basic residues like Arg and Lys were protonated, while acidic residues like Asp and Glu were deprotonated. The protonation states for His residues were determined based on an inspection of residues local environment and their pKa values calculated with PDB2PQR (11). To examine the sensitively of our results to altered His protonation, we performed an additional set of simulations using different initial protonation state assignments. These results, detailed in Fig S8, indicate an overall robustness to our initial protonation state assignments.

Simulation structures were solvated in a truncated cubic box of pre-equilibrated TIP3P water molecules, which extended 12 Å in each dimension from the surface of protein atoms. Sodium counterions (Na^+) were added to neutralize the systems. Energy minimization was performed in four stages, with each stage employing 500 steps of steepest decent followed by 1500 steps of conjugate gradient using constant-volume periodic boundary conditions. First, solvent only minimization with fixed protein and ligand solute atoms. Second, fixed backbone with free side-chain and ligand atoms. Third, fixed solvent with free solute atoms. Finally, all atoms were relaxed without restraints. Following minimization, 10ps of molecular dynamics (MD) simulation was performed to heat the system from 0K to 300K in a NVT ensemble. To bring systems to the correct density, a further 1ns of equilibration simulation was then performed using an NPT ensemble (T=300K, P=1bar). Production phase 40-ns conventional MD (cMD) and 100-ns accelerated MD (aMD) simulations were then performed at constant temperature (300K) and constant pressure (1bar). For both energy minimization and MD simulations, the particle-mesh Ewald (PME) summation method was adopted to treat long-range electrostatic interactions. In addition, an 8Å cutoff was used to truncate the short-range nonbonded VDW interactions. Additional operational parameters for molecular dynamics included a 2fs time step, removal of the center-of-mass motion at every 1000 steps, and update of the nonbonded neighbor list every 25 steps. The SHAKE algorithm was also used to constrain all covalent bonds involving hydrogen atoms.

Accelerated molecular dynamics

Accelerated molecular dynamics (aMD) is an enhanced sampling method that aims to capture the long-time dynamics of solvated biomolecules. aMD has been shown to increase conformational sampling over conventional MD (cMD) and has been successfully applied to a wide range of applications and biological systems (12-18). This method adds a non-negative boost potential, $\Delta V(r)$, to the original potential energy $V(r)$ every time $V(r)$ is below an energy threshold E (Equation 1). This has the effect of easing the crossing of energy barriers and increase the rate of escape from energy basins (19),

$$\Delta V(r) = \begin{cases} 0, & V(r) \geq E \\ \dfrac{(E - V(r))^2}{\alpha + (E - V(r))}, & V(r) < E \end{cases} \tag{1}$$

where α modulates the depth and the local roughness of the energy basins in the modified potential. In this work, to enhance the sampling of both torsional degree of freedom and diffusive motions, we employed the dual-boosting version of aMD, which is based on applying boost potentials separately to torsional and total potential energy terms (20),

$$\begin{aligned} V(r) &= V_0(r) + V_t(r) \\ V^*(r) &= \left\{V_0(r) + \left[V_t(r) + \Delta V_t(r)\right]\right\} + V_T(r) \end{aligned} \tag{2}$$

where $V_t(r)$ is the total potential of the torsional terms, $\Delta V_t(r)$ and $\Delta V_T(r)$ are the boost potentials for the torsional terms, $V_t(r)$, and the total potential $V_T(r)$, respectively. Here $V_T(r)=V_0(r)+V_t(r)+\Delta V_t(r)$. The parameters E and α of both torsional and total boost potentials were set empirically. Specifically, for total potential, we set $\alpha_T=0.2N_{atom}\varepsilon$ and $E_T=\langle V_T(r)\rangle+\alpha_T$, where N_{atom} is the number of atoms in the system and $\varepsilon=1.0\text{kcal}\cdot\text{mol}^{-1}$. For torsional terms, we set $\alpha_t=0.7N_{res}\varepsilon$ and $E_t=3.5N_{res}\varepsilon+\langle V_t(r)\rangle+\alpha_t$, where N_{res} is the number of protein residues. The average potential energy used in above equations was derived from the 1ns equilibration cMD simulations (See above).

Principal components analysis

Principal component analysis (PCA) was performed for both crystallographic structures and MD trajectory snapshots to capture and characterize inter-conformer relationship. The application of PCA to both distributions of experimental structures and MD trajectories, along with its ability to provide considerable insight into the nature of conformational differences in a range of protein families has been previously discussed (21-24). Briefly, PCA is based on the diagonalization of the covariance matrix, COV, with elements COV_{ij} calculated from the Cartesian coordinates of Cα atoms, r, after the superposition of all structures under analysis:

$$COV_{ij} = \left\langle \left(r_i - \langle r_i \rangle\right) \cdot \left(r_j - \langle r_j \rangle\right) \right\rangle \tag{3}$$

where i and j enumerate all possible pairs of 3N Cartesian coordinates (N is the number of atoms being analyzed). The eigenvectors of COV, referred to principal components (PCs), form a linear basis set matching the distribution of structures. The variance of the distribution along each eigenvector or PC is given by the corresponding eigenvalue. Projection of the distribution onto the subspace defined by the largest PCs (along which the structural variance is the largest) provides a low-dimensional representation of structures facilitating inter-conformer analysis. The residues used to calculate the covariance matrix and subsequent PCs are dependent on cases: For the analysis of entire protein, we chose the residues where there is no gap at the same position in the alignment of available experimental structures; On the other hand, for analyzing the variance in one specific domain, we used a subset of the residues defined above that belong to that domain. We defined the Ras-like domain (RasD) as residues 27-56 and 174-340 and the α-helical domain (HD) by residues 57-173.

Conformer clustering

PCA for the HD of conformers predicted by all the nucleotide free aMD and cMD simulations was performed. For each trajectory, snapshots of protein structure were taken every 25 frames (i.e. a time-interval of 50ps). After superposition of all selected conformers based on the "core structure" (See above), Cα atoms of HD were chosen to perform the PCA. The results showed consistent "out-of-plane shifting" and "in-plane rotation" motions to those revealed by PCA of each individual trajectory (Fig. S3 and Fig. S5). We then clustered the conformers with the *k*-means method, with structural dissimilarity defined by Euclidean distance based on the first 10 PCs. Note that the first 10 PCs account for almost 99% structural variation of the HD. We obtained six clusters representing six metastable states visited by Gαt during the domain opening: One close form, two open forms along the two dominant modes of the HD motion, and three intermediate half-open forms (Fig. S5). Representatives of clusters were selected as the conformers that were closest (by root-mean square deviation, RMSD) to the centers of clusters, and they were considered for further dynamical network analysis (See below).

Contact activity analysis

Analysis of rare contact formation and breaking events associated with conformational changes was performed with TimeScapes (version 1.2.2) (25). TimeScapes employs a contact matrix built from distances between residues along with a median filter and Gaussian kernel to monitor the fraction of significant contact formation and breaking events per trajectory segment. For the current analysis, we used the same residue subset used for the PCA on crystal substructures (i.e. equivalent positions found in all crystal structures). In addition, only the residue pairings from different domains were considered (inter-domain contacts). Side chains were considered in contact if their distance was between 6Å and 7Å. The half width median filter was set to a value of 6ns.

Cross-correlation and dynamical network analysis

To identify protein segments with correlated atomic motions the cross-correlation coefficient, *Cij*, for the displacement of all residue pairs, *i* and *j*, was calculated. In addition to analysis of aMD trajectories, we performed three 40ns cMD simulations, with initial conformations taken from GTP bound crystallographic structure (PDB ID: 1TND), GDP bound crystallographic structure (PDB ID: 1TAG), and the open conformer predicted by nucleotide free simulations. To further verify the robustness of results, two additional independent 40ns cMD simulations for GTP and GDP states were also performed. From these simulations we calculated the maximum cross-correlation between heavy atoms belonging to each residue pair *i* and *j*:

$$C_{ij} = \max\left\{ C^{a}_{mn} = \langle \Delta r_m \cdot \Delta r_n \rangle / \left(\langle \Delta r_m^2 \rangle \langle \Delta r_n^2 \rangle \right)^{1/2} \middle| m \in A(i), n \in A(j) \right\} \tag{4}$$

where Δr_m is the displacement from the mean position of the *m*th atom of residue *i* determined over the length of the simulation and $A(i)$ the set of all heavy atoms belonging to residue *i*. The nucleotide and Mg^{2+} were also included in the calculation; For the nucleotide, all heavy atoms were split into two parts, treated as two residues, which represented the base and the phosphate regions, respectively.

A network of residue-residue and residue-ligand coupling was built, with nodes defined by the Cα atoms for each amino acid, Mg^{2+}, PA atom for the phosphate and N9 atom for the base region of the nucleotide. Using these data dynamical networks were constructed following the method of Luthey-Schulten and colleagues (26). In this approach a weighted graph is constructed where each residue represents a node. Two nodes are connected in the network if they are in contact during the trajectory segment under analysis; i.e., their closest heavy atoms are within 6.0 Å for 75% of simulation frames. Edges between nodes *i* and *j* are weighted (w_{ij}) by their respective correlation value (C_{ij}):

$$W_{ij} = -\log\left(\left|C_{ij}\right|\right) \quad (5)$$

Hierarchical clustering was used to generate aggregate nodal clusters, or communities, that are highly correlated and within close physical proximity. Network analysis concepts (i.e., shortest path, centrality, and suboptimal path analysis) were used to identify prominent nodes and paths in the network using the VMD dynamical network analysis plugin (27). In addition, network topology graphs were generated with Cytoscape 2.8.3 (28), in which circles represented communities and lines the connections between communities. The circle radius indicated the number of residues in the community and the lines width was scaled by the maximal betweenness of the edges that connected the two communities.

Supporting References

1. Grant, B. J., A. P. Rodrigues, K. M. ElSawy, J. A. McCammon, and L. S. Caves. 2006. Bio3d: an R package for the comparative analysis of protein structures. Bioinformatics 22:2695-2696.
2. Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. 2000. The Protein Data Bank. Nucleic Acids Res 28:235-242.
3. Gerstein, M., and R. B. Altman. 1995. Average core structures and variability measures for protein families: application to the immunoglobulins. J Mol Biol 251:161-175.
4. Rasmussen, S. G., B. T. DeVree, Y. Zou, A. C. Kruse, K. Y. Chung, T. S. Kobilka, F. S. Thian, P. S. Chae, E. Pardon, D. Calinski, J. M. Mathiesen, S. T. Shah, J. A. Lyons, M. Caffrey, S. H. Gellman, J. Steyaert, G. Skiniotis, W. I. Weis, R. K. Sunahara, and B. K. Kobilka. 2011. Crystal structure of the beta2 adrenergic receptor-Gs protein complex. Nature 477:549-555.
5. Eswar, N., B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M. Y. Shen, U. Pieper, and A. Sali. 2007. Comparative protein structure modeling using MODELLER. Curr Protoc Protein Sci Chapter 2:Unit 2.9.
6. Case, D., T. Darden, T. I. Cheatham, C. Simmerling, J. Wang, R. Duke, R. Luo, R. Walker, W. Zhang, K. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Wails, A. Goetz, I. Kolossvary, K. Wong, F. Paesani, J. Vanicek, R. Wolf, J. Liu, X. Wu, S. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, M. Hsieh, G. Cui, D. Roe, D. Mathews, M. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, and P. Kollman. 2012. AMBER 12. University of California, San Francisco.
7. Hornak, V., R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling. 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 65:712-725.
8. Meagher, K. L., L. T. Redman, and H. A. Carlson. 2003. Development of polyphosphate parameters for use with the AMBER force field. J Comput Chem 24:1016-1025.
9. Lambright, D. G., J. P. Noel, H. E. Hamm, and P. B. Sigler. 1994. Structural determinants for activation of the alpha-subunit of a heterotrimeric G protein. Nature 369:621-628.
10. Noel, J. P., H. E. Hamm, and P. B. Sigler. 1993. The 2.2 A crystal structure of transducin-alpha complexed with GTP gamma S. Nature 366:654-663.
11. Dolinsky, T. J., J. E. Nielsen, J. A. McCammon, and N. A. Baker. 2004. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res 32:W665-667.
12. Ortiz-Sanchez, J. M., S. E. Nichols, J. Sayyah, J. H. Brown, A. McCammon, and B. J. Grant. 2012. Identification of Potential Small Molecule Binding Pockets on Rho Family GTPases. Plos One 7:e40809.
13. Thomas, J. R., P. C. Gedeon, B. J. Grant, and J. D. Madura. 2012. LeuT Conformational Sampling Utilizing Accelerated Molecular Dynamics and Principal Component Analysis. Biophysical Journal 103:L01-L03.
14. Bucher, D., B. J. Grant, P. R. Markwick, and J. A. McCammon. 2011. Accessing a Hidden Conformation of the Maltose Binding Protein Using Accelerated Molecular Dynamics. Plos

Computational Biology 7:e1002034.
15. Bucher, D., B. J. Grant, and J. A. McCammon. 2011. Induced Fit or Conformational Selection? The Role of the Semi-closed State in the Maltose Binding Protein. Biochemistry 50:10530-10539.
16. de Oliveira, C. A. F., B. J. Grant, M. Zhou, and J. A. McCammon. 2011. Large-Scale Conformational Changes of Trypanosoma cruzi Proline Racemase Predicted by Accelerated Molecular Dynamics Simulation. Plos Computational Biology 7:e1002178.
17. Grant, B. J., A. A. Gorfe, and J. A. McCammon. 2010. Large conformational changes in proteins: signaling and other functions. Current Opinion in Structural Biology 20:142-147.
18. Grant, B. J., A. A. Gorfe, and J. A. McCammon. 2009. Ras Conformational Switching: Simulating Nucleotide-Dependent Conformational Transitions with Accelerated Molecular Dynamics. Plos Computational Biology 5:e1000325.
19. Hamelberg, D., J. Mongan, and J. A. McCammon. 2004. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. J Chem Phys 120:11919-11929.
20. Hamelberg, D., C. A. de Oliveira, and J. A. McCammon. 2007. Sampling of slow diffusive conformational transitions with accelerated molecular dynamics. J Chem Phys 127:155102.
21. Caves, L. S., J. D. Evanseck, and M. Karplus. 1998. Locally accessible conformations of proteins: multiple molecular dynamics simulations of crambin. Protein Sci 7:649-666.
22. Gorfe, A. A., B. J. Grant, and J. A. McCammon. 2008. Mapping the nucleotide and isoform-dependent structural and dynamical features of Ras proteins. Structure 16:885-896.
23. Grant, B. J., J. A. McCammon, L. S. Caves, and R. A. Cross. 2007. Multivariate analysis of conserved sequence-structure relationships in kinesins: coupling of the active site and a tubulin-binding sub-domain. J Mol Biol 368:1231-1248.
24. van Aalten, D. M., D. A. Conn, B. L. de Groot, H. J. Berendsen, J. B. Findlay, and A. Amadei. 1997. Protein dynamics derived from clusters of crystal structures. Biophys J 73:2891-2896.
25. Wriggers, W., K. A. Stafford, Y. Shan, S. Piana, P. Maragakis, K. Lindorff-Larsen, P. J. Miller, J. Gullingsrud, C. A. Rendleman, M. P. Eastwood, R. O. Dror, and D. E. Shaw. 2009. Automated Event Detection and Activity Monitoring in Long Molecular Dynamics Simulations. Journal of Chemical Theory and Computation 5:2595-2605.
26. Sethi, A., J. Eargle, A. A. Black, and Z. Luthey-Schulten. 2009. Dynamical networks in tRNA:protein complexes. Proc Natl Acad Sci U S A 106:6620-6625.
27. Eargle, J., and Z. Luthey-Schulten. 2012. NetworkView: 3D display and analysis of protein RNA interaction networks. Bioinformatics 28:3000-3001.
28. Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498-2504.
29. Zielinski, T., A. J. Kimple, S. Q. Hutsell, M. D. Koeff, D. P. Siderovski, and R. G. Lowery. 2009. Two Galpha(i1) rate-modifying mutations act in concert to allow receptor-independent, steady-state measurements of RGS protein activity. J Biomol Screen 14:1195-1206.
30. Remmers, A. E., C. Engel, M. Liu, and R. R. Neubig. 1999. Interdomain interactions regulate GDP release from heterotrimeric G proteins. Biochemistry 38:13795-13800.

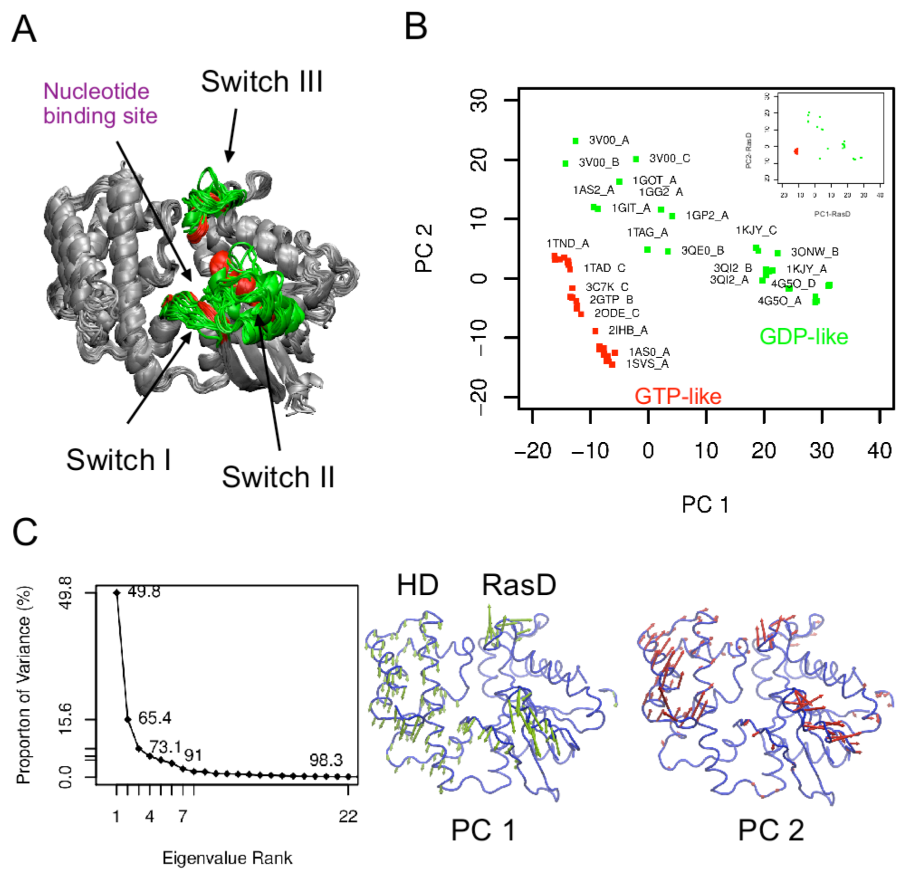

Figure S1. Results of principal component analysis on Gαt's crystallographic structures. (*A*) Superposition of the crystallographic structures, with switch regions colored by nucleotide state (red for GTP, green for GDP). (*B*) Mapping of the crystallographic structures onto the PC1-PC2 planes. *Inset* is the mapping of the structures onto the PC1-PC2 plane constructed on the RasD regions alone (i.e. with the exclusion of the α-helical domain). (*C*) Proportion of variance for the top eigenvalues (left) and the dominant modes of motion (arrows) revealed by PCA (right).

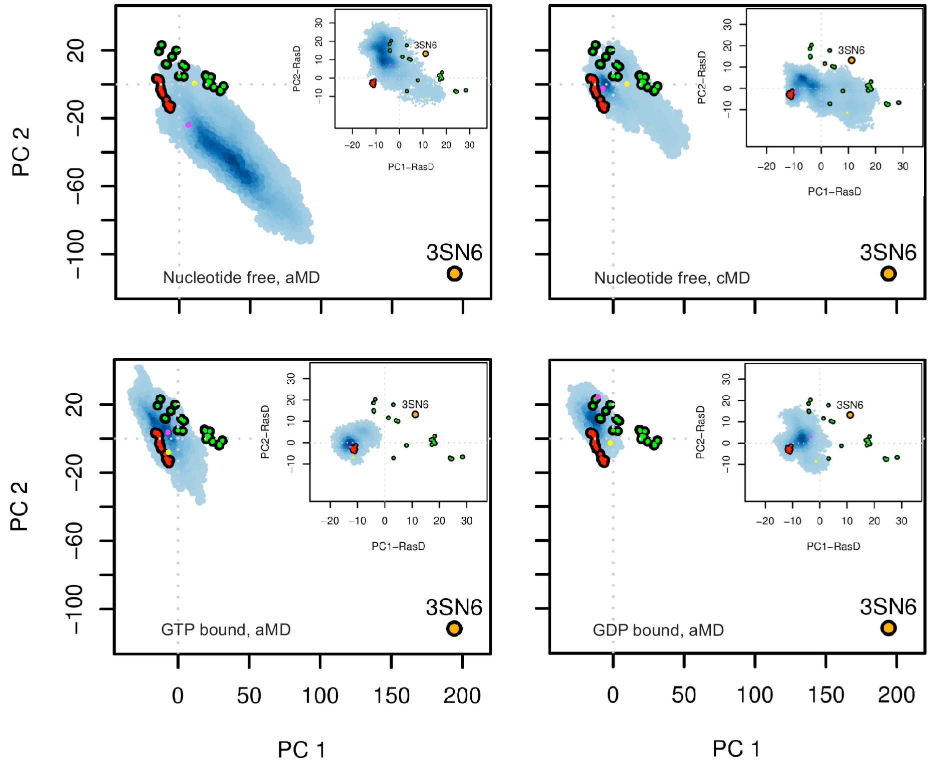

Figure S2. Comparison of sampling among simulations. Conformers from each 100ns simulation trajectory (blue points) are mapped onto the first two principal components (PCs) obtained from analysis of the crystallographic structures. Yellow and magenta points depict the first and last frames of each trajectory. *Inset*, mapping based on the PCA with excluding the α-helical domain.

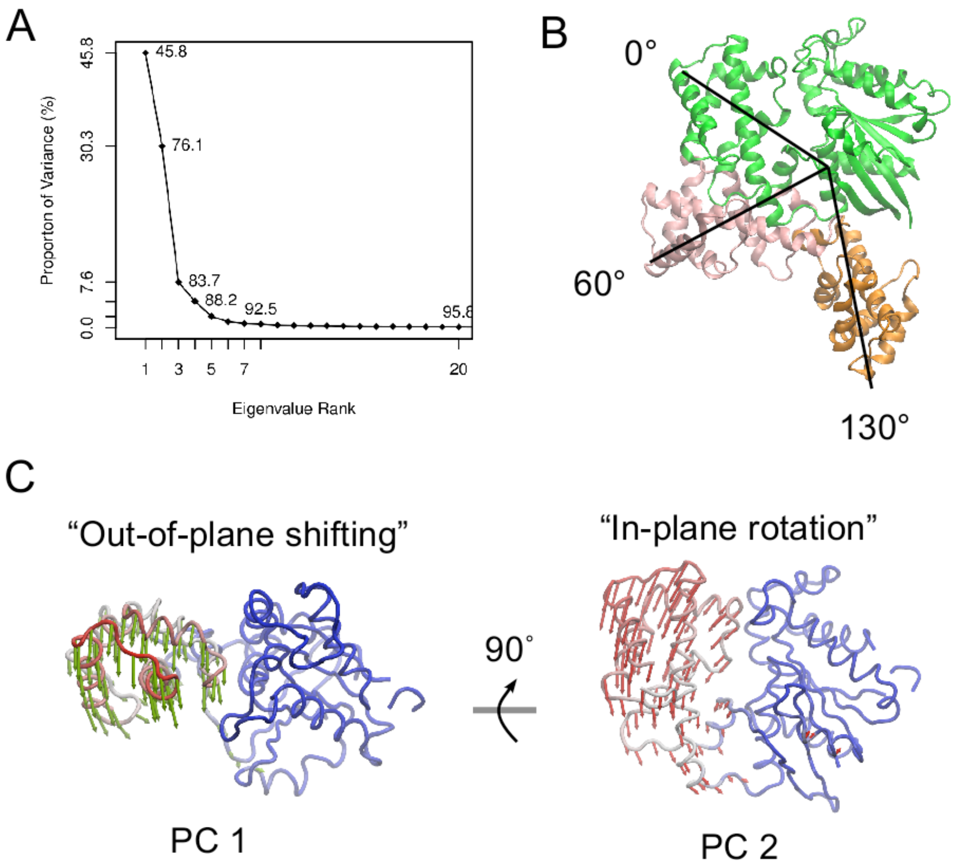

Figure S3. Principal component analysis of a single nucleotide-free aMD simulation. (A) The proportion of variance for the top 20 eigenvalues. (B) Superimposition of the initial (green) and open form (pink) conformers from the simulation onto the open Gαs from the β2AR-Gs complex structure (orange). The Ras-like domain is relatively stable and is only shown for the initial MD conformer for clarity. (C) Dominant modes of motion (arrows) from PCA. Protein is represented as tubes with color indicating the mobility: Blue (low) to red (high).

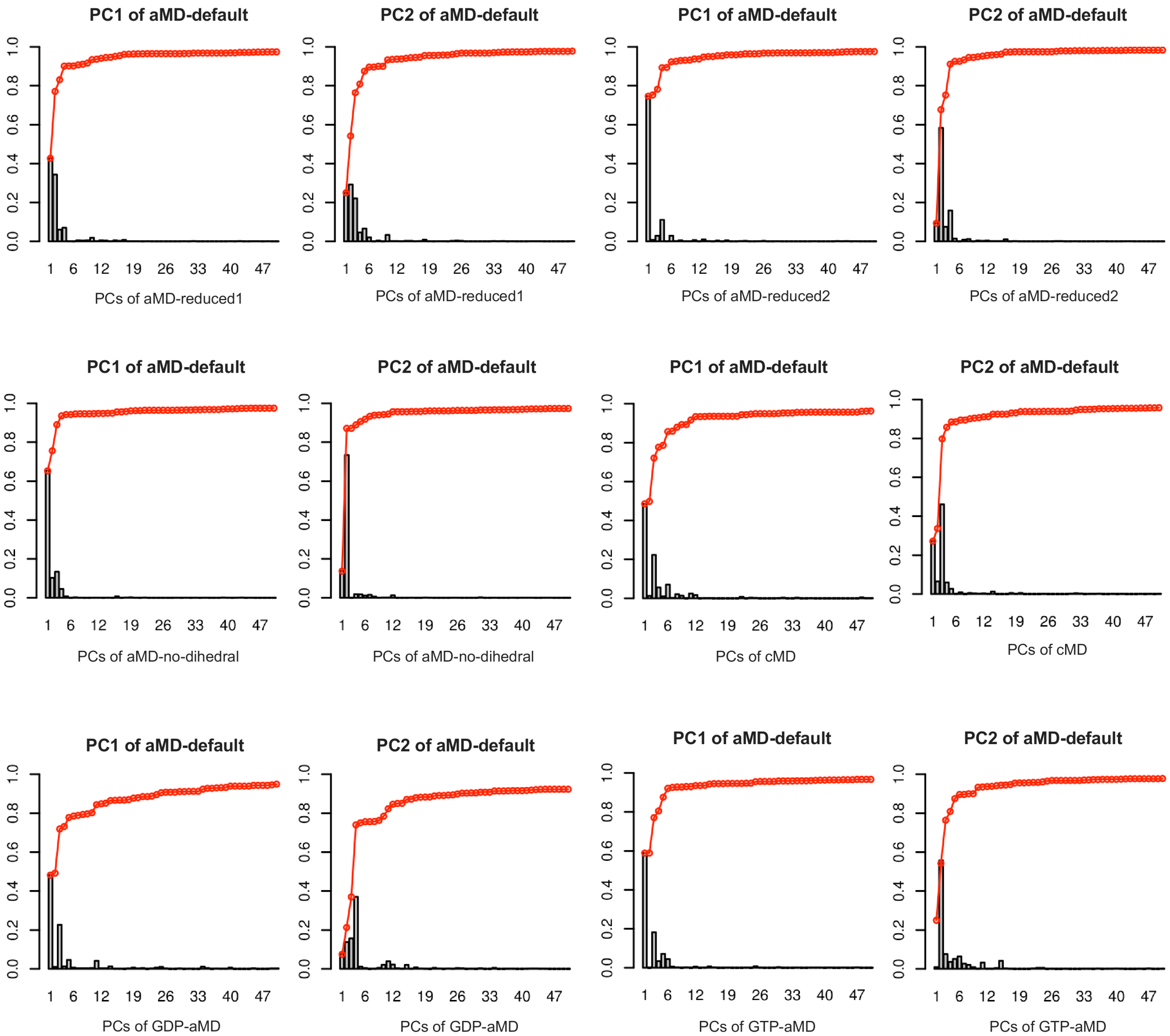

Figure S4. Comparison of dominant motions among simulations. The absolute value of the inner-product between the principal components (PCs) derived from PCA of different simulation trajectories are shown (gray bars). Red lines are the geometric accumulation of the product along the x-axis. Simulations are performed under nucleotide-free conditions except for those explicitly indicated on the x-axis. Besides that with the default parameters, multiple nucleotide free aMD simulations were performed, with boost potential for torsional angle either reduced or completely removed.

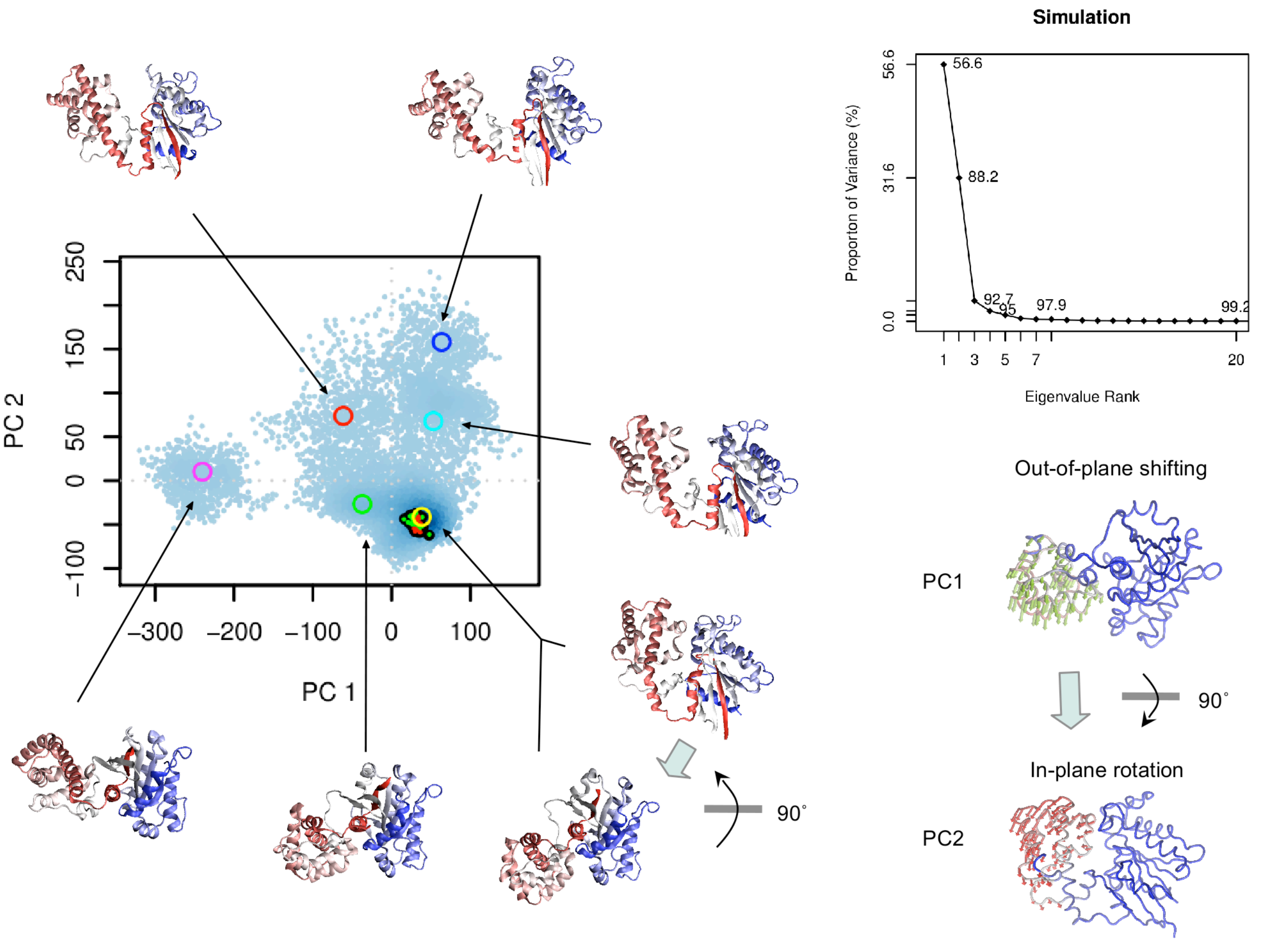

Figure S5. Clustering of conformers from nucleotide-free simulations. Centers of clusters are indicated as circles with representative structures depicted as cartoon and colored by residue index. The proportion of variance for top eigenvalues and the dominant modes of motion of the trajectory principal components (arrows mapped onto the protein structures) are also shown.

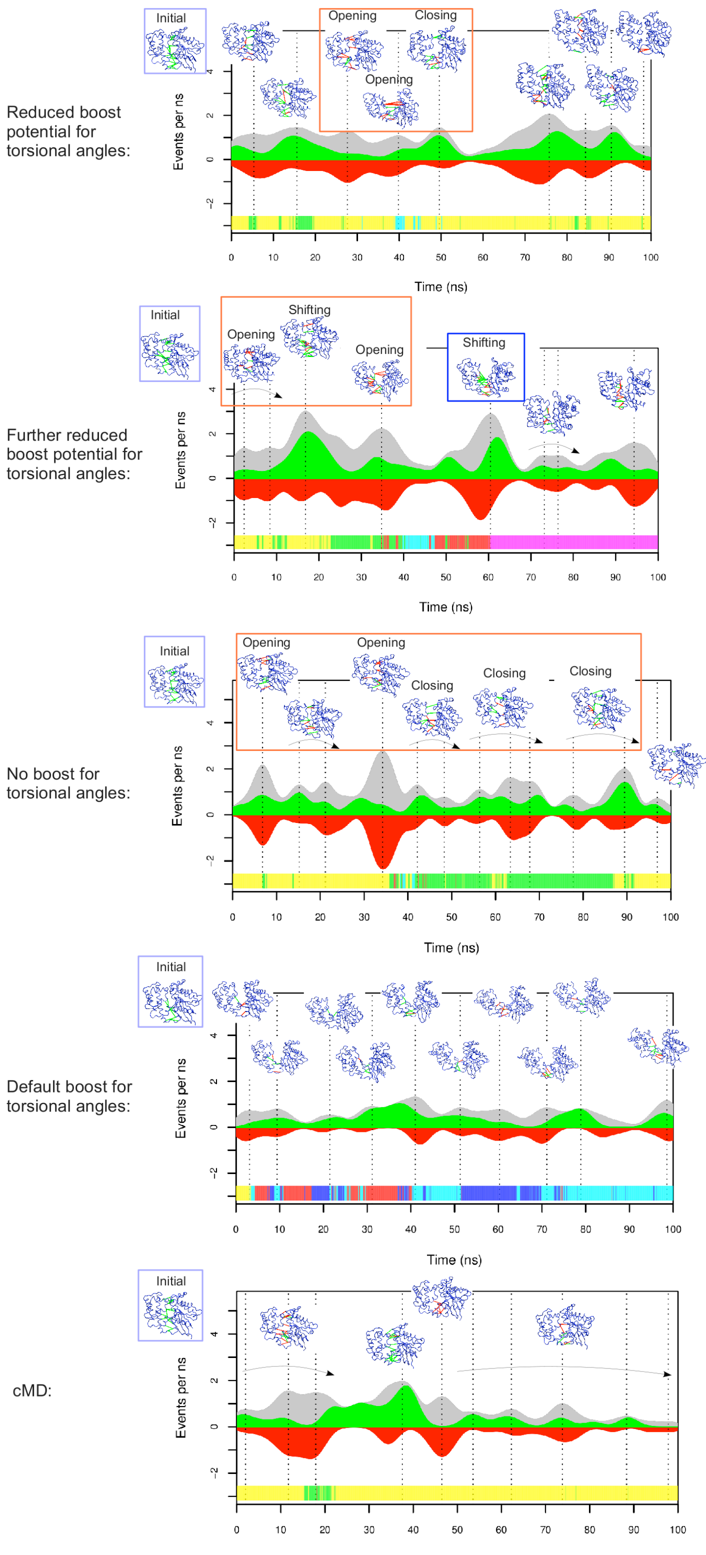

Figure S6. Dynamics of nucleotide-free aMD simulations characterized by the breaking and formation of inter-domain amino acid side-chain contacts. The red and green areas show the rate of contact breaking and formation respectively, while the gray area depict the rate of contact change as a whole (contact forming and breaking). The green and red lines in the molecular structure snapshots indicate the location of contact forming and breaking events as determined by the TimeScapes package (See Methods for further details). The color bar displayed on the bottom shows the transitions of the protein among distinct structural clusters, with color code the same as used in Fig. S5. Besides that with the default parameters, multiple nucleotide free aMD simulations were performed, with boost potential for torsional angle either reduced or completely removed.

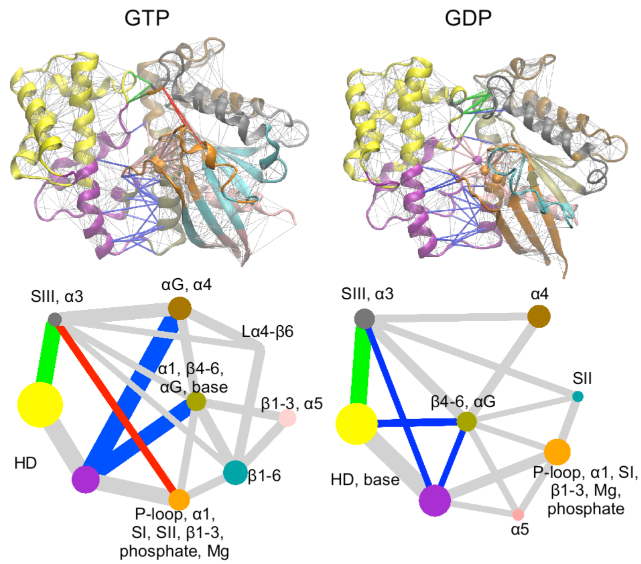

Figure S7. Nucleotide associated differences in dynamic coupling calculated from 40-ns cMD simulations independent of those discussed in the main text (Fig. 1B-C).

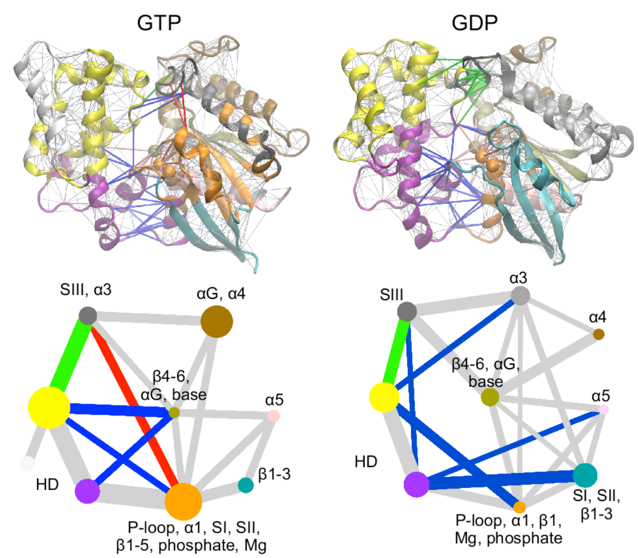

Figure S8. Nucleotide associated differences in dynamic coupling calculated from 40-ns cMD simulations with protonation states of histidine residues different from those discussed in the main text (Fig. 1B-C).

Movies are available from <https://vimeo.com/user17469580/videos>

Movie S1. The opening of HD predicted by nucleotide free aMD simulation.

Movies S2-S4. Dynamic views of the dynamical coupling and community networks for the GTP, GDP, and APO states, respectively.

Table S1. Key residues identified from the dynamical network analysis

	Location	Nucleotide state	References
Nucleotide-residue coupling			
A37	P-loop	GTP	-
G38-I45	P-loop	Both	-
D146	N-terminus of αE	GDP	-
R172, S173	αF	GDP	-
R174	Switch I	Both	Zielinski, et al., 2009(29)
V175	Switch I	Both	-
K176, T177	Switch I	GTP	-
D196	C-terminus of β3	Both	-
V197	C-terminus of β3	GTP	-
N265-D268	Lβ5-αG	Both	-
C321	Lβ6-α5	Both	-
A322	Lβ6-α5	Both	Zielinski, et al., 2009(29)
Inter-domain residue-residue coupling			
T44::L171, Q48::L171	α1::αF	GDP	-
K47::F65, K50::Y57, K50::E61, G56::E61	α1::αA	GDP	-
K47::V170, K47::L171	α1::αF	Both	-
K47::S173	α1::αF	GTP	-
D55::Y57, G56::S58	α1::αA	Both	-
I68::R174, I68::V175, N72::R174	αA::Switch I	GDP	-
A139::M228, S140::D227, Q143::M228, Q143::V229, L144:M228	αD::Switch III	GDP	-
S140::M228	αD::Switch III	GDP	Remmers, et al., 1999(30)
D146::K266	αE::Lβ5-αG	GDP	-
Y151::R174	αE::Switch I	GTP	-
R172::R174, S173::V175	αF::Switch I	Both	-

Table S2. Analyzed crystallographic structures of G protein α subunit

PDB ID	Chain	Ligand	Source	Resolution	Reference
1FQJ	A, D	ALF, GDP, MG	*Bos Taurus, Rattus Norvegicus*	2	Slep et al. (2001)
1FQK	A, C	ALF, GDP, MG	*Bos Taurus, Rattus Norvegicus*	2.3	Slep et al. (2001)
1GOT	A	MSE, GDP	*Bos Taurus*	2	Lambright et al. (1996)
1TAD	A, B, C	ALF, CA, CAC, GDP	*Bos Taurus*	1.7	Sondek et al. (1994)
1TAG	A	GDP, MG	*Bos Taurus*	1.8	Lambright et al. (1994)
1TND	A, B, C	CAC, GSP, MG	*Bos Taurus*	2.2	Noel et al. (1993)
3V00	A, B, C	GDP	*Bos Taurus, Rattus Norvegicus*	2.9	Singh et al. (2012)
1AGR	A, D	ALF, CIT, GDP, MG	*Rattus Norvegicus*	2.8	Tesmer et al. (1997)
1AS0	A	GSP, MG, SO4	*Rattus Norvegicus*	2	Raw et al. (1997)
1AS2	A	GDP, PO4	*Rattus Norvegicus*	2.8	Raw et al. (1997)
1BH2	A	GSP, MG	*Rattus Norvegicus*	2.1	Posner et al. (1998)
1CIP	A	GNP, MG	*Rattus Norvegicus*	1.5	Coleman et al. (1999)
1GFI	A	ALF, GDP, MG	*Rattus Norvegicus*	2.2	Coleman et al. (1994)
1GG2	A	GDP	*Rattus Norvegicus, Bos Taurus*	2.4	Wall et al. Cell (1995)
1GIA	A	GSP, MG	*Rattus Norvegicus*	2	Coleman et al. (1994)
1GIL	A	GSP, MG	*Rattus Norvegicus*	2.3	Coleman et al. (1994)
1GIT	A	GDP, PO4	*Rattus Norvegicus*	2.6	Berghuis et al. (1996)
1GP2	A	GDP	*Rattus Norvegicus, Bos Taurus*	2.3	Wall et al. (1995)
1KJY	A, C	CS, GDP, MG	*Homo Sapiens*	2.7	Kimple et al. (2002)
1SVK	A	ALF, GDP, MG	*Rattus Norvegicus*	2	Thomas et al. (2004)
1SVS	A	GNP, MG	*Rattus Norvegicus*	1.5	Thomas et al. (2004)
2GTP	B	ALF, GDP, MG	*Homo Sapiens*	2.5	Soundararajan et al. (2008)
2IHB	A	ALF, GDP, MG	*Homo Sapiens*	2.7	Soundararajan et al. (2008)
2ODE	A, C	ALF, GDP, MG	*Homo Sapiens*	1.9	Soundararajan et al. (2008)
2OM2	A, C	GDP, MG	*Homo Sapiens*	2.2	Sammond et al. (2007)
2V4Z	A	ALF, GDP, MG	*Homo Sapiens*	2.8	Kimple et al. (2009)
2XNS	A, B	GDP, SRT, SO4	*Homo Sapiens*	3.4	Sammond et al. (2011)
2ZJY	A	ALF, GDP, MG	*Rattus Norvegicus*	2.8	Morikawa et al. To be Published
3C7K	C	ALF, GDP, MG	*Mus Musculus*	2.9	Slep et al. (2008)
3FFA	A	GSP, MG, SO4	*Rattus Norvegicus*	2.3	Kapoor et al. (2009)
3ONW	A, B	GDP, SO4	*Homo Sapiens*	2.4	Bosch et al. (2011)
3QE0	B	GDP, MG	*Homo Sapiens*	3	Bosch et al. (2012)
3QI2	A, B	GDP, GOL, SO4	*Homo Sapiens*	2.8	Bosch et al. (2012)
4G5O	A, D	CIT, GDP, SO4	*Homo Sapiens*	2.9	Jia et al. To be Published
4G5Q	A, D	GDP, SO4, CIT	*Homo Sapiens*	2.9	Jia et al. To be Published
4G5R	A	GDP, SO4, CIT	*Homo Sapiens*	3.5	Jia et al. To be Published