Identification of Potential Small Molecule Binding Pockets on Rho Family GTPases

Juan Manuel Ortiz-Sanchez[1]*, Sara E. Nichols[1,2], Jacqueline Sayyah[2], Joan Heller Brown[2], J. Andrew McCammon[1,2,3], Barry J. Grant[4]*

1 Department of Chemistry and Biochemistry and Center for Theoretical Biological Physics, University of California San Diego, La Jolla, California, United States of America, 2 Department of Pharmacology, University of California San Diego, La Jolla, California, United States of America, 3 Howard Hughes Medical Institute, University of California San Diego, La Jolla, California, United States of America, 4 Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, United States of America

Abstract

Rho GTPases are conformational switches that control a wide variety of signaling pathways critical for eukaryotic cell development and proliferation. They represent attractive targets for drug design as their aberrant function and deregulated activity is associated with many human diseases including cancer. Extensive high-resolution structures (>100) and recent mutagenesis studies have laid the foundation for the design of new structure-based chemotherapeutic strategies. Although the inhibition of Rho signaling with drug-like compounds is an active area of current research, very little attention has been devoted to directly inhibiting Rho by targeting potential allosteric non-nucleotide binding sites. By avoiding the nucleotide binding site, compounds may minimize the potential for undesirable off-target interactions with other ubiquitous GTP and ATP binding proteins. Here we describe the application of molecular dynamics simulations, principal component analysis, sequence conservation analysis, and ensemble small-molecule fragment mapping to provide an extensive mapping of potential small-molecule binding pockets on Rho family members. Characterized sites include novel pockets in the vicinity of the conformationaly responsive switch regions as well as distal sites that appear to be related to the conformations of the nucleotide binding region. Furthermore the use of accelerated molecular dynamics simulation, an advanced sampling method that extends the accessible time-scale of conventional simulations, is found to enhance the characterization of novel binding sites when conformational changes are important for the protein mechanism.

Introduction

Rho proteins are eukaryotic intracellular signaling hubs. They function to relay signals from cell-surface receptors to signaling cascades that control diverse cell processes including gene transcription, [1] cell-cycle progression, [2,3] and cytoskeleton reorganization. [4] Like other members of the Ras GTPase superfamily, Rho proteins act as conformational switches, hydrolytically cycling between active GTP-bound and inactive GDP-bound conformations. Three major classes of regulatory proteins modulate Rho's activity: guanine nucleotide exchange factors (GEFs), GTPase-activating proteins (GAPs) and guanine nucleotide dissociation inhibitors (GDIs) (**Figure 1A**). [5] Association with GEFs promotes Rho activation by stimulating the exchange of GDP for GTP. This exchange of nucleotide induces important conformational changes in structural regions termed switch 1 (residues 24–40) and switch 2 (residues 57–75), (see **Figure 1B**). [6,7] This change in conformation allows active Rho to interact with a variety of protein effectors that initiate a network of signals affecting cell functions. Regeneration of the inactive form of Rho is promoted by the action of GAPs, which stimulate GTP hydrolysis and formation of the inactive GDP conformation.

The third class of regulatory proteins, GDIs can sequester inactive GDP-bound Rho proteins and prevent their activation. Disturbing the balance of these GEFs, GAPs and GDIs can affect the fidelity of the activation cycle and have severe consequences for the wide variety of cellular events that are orchestrated by Rho dependent signaling.

Deregulated Rho activity can induce aberrant phenotypes that have been linked to the initiation and progression of multiple cancers [8,9] as well as cardiovascular [10] and neurological disorders. [11] For example, the upregulation of RhoA is observed in breast, colon, lung, gastric bladder and testicular cancer. [12,13] In addition, the overexpression of RhoC, Rac1, Rac2, Rac3 and Cdc42 have been observed in a number of cancers. [9,14] Rearrangement of the RhoH gene that leads to a defective cycling between GDP- and GTP-bound forms is associated with non-Hodgkin's lymphomas. [15,16] Furthermore, the role of Rho family proteins in the expression of NF-kB-dependent genes and the migration of leukocytes along with their interaction with the Angiotensin II pathway indicate that therapeutically targeting Rho proteins may also have applications in the treatment of inflammatory and cardiovascular diseases. [17,18].
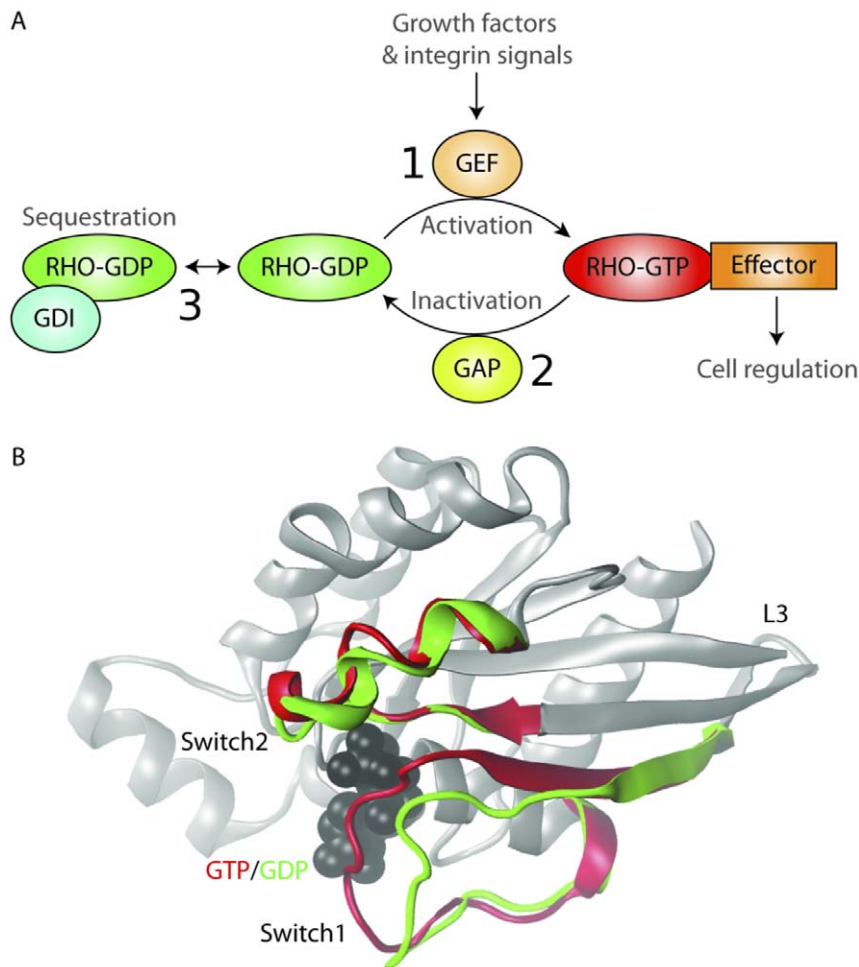
Figure 1. Rho GTPase activation. (A) Schematic representation of the Rho GTPase activation cycle. (1) The GTP-bound active state is generated by guanine exchange factors (GEFs), promoting the exchange of GDP for GTP. (2) GTPases activating proteins (GAPs) catalyze the hydrolysis of GTP to GDP, restoring the GDP-bound inactive state. (3) Stabilization of the GDP-bound inactive state by GDP dissociation inhibitors (GDIs). (B) The molecular structure of representative the GTP and GDP-bound Rho conformations (PDB codes: 1KMQ and 1FTN). The switch regions of the GTP-bound form are shown in red and GDP-bound form in green.
doi:10.1371/journal.pone.0040809.g001

Targeting drugs to Rho GTPases and related signaling pathway members is believed to have significant therapeutic potential. [10,19,20,21,22,23] Inhibitory mutants of RhoA, RhoG, Rac1 and Cdc42 prevent Ras transformation of fibroblasts, and activated mutants of these proteins are transforming. [14] Furthermore, anti-RhoA and anti-RhoC siRNAs have been shown to inhibit the growth and angiogenesis of tumors in mouse models [24]. Existing strategies for targeting deregulated Rho signaling include inhibitors of geranylgeranylation and other post-translational modifications of Rho [25,26,27]. These compounds have the potential to attenuate C-terminal lipid modifications required for plasma membrane localization and subsequent signaling. A drawback of such inhibitors is their poor selectivity as they likely affect many lipid-modified proteins. Another approach involves inhibitors of downstream Rho effectors such as Rho kinase (ROCK). Several ROCK inhibitors have been successful in preclinical studies, highlighting the potential benefit of clinical Rho pathway inhibition. [28,29,30,31] However, because Rho proteins utilize a multitude of downstream effectors, a particular effector inhibitor will likely impair only a subset of Rho malfunctions leading to potentially limited therapeutic benefits. The design of small molecule inhibitors that directly block the nucleotide binding site (NBS) of Rho has thus far been complicated by the conserved nature of this site throughout the larger Ras superfamily. Compounds that interact with this site have a very high likelihood of unspecifically blocking other important G-protein mediated pathways.

Currently very little attention has been placed on targeting distal non-nucleotide binding sites that may offer the possibility of modulating Rho activity in a more selective and specific manner. We have recently proposed [32] that conformational selection in Ras like G-proteins including Rho is the dominant mechanism underlying the nucleotide-dependent conformational cycle. [33] Additionally, simulations have predicted allosteric coupling of the NBS with the membrane interacting C-terminus in both Rho and Ras. In agreement with our predictions, Buhrman and collaborators described allosteric modulation of Ras by mutation experiments of residues distal to the NBS. [34] Building on these findings we developed a multi-level computational approach to discover inhibitors of Ras. This approach combined novel pocket identification, conventional molecular simulations, ensemble docking and experimental testing of computationally selected inhibitors. [35] Of particular note, cell-based assays confirmed that a number of the chosen distal binding site directed compounds
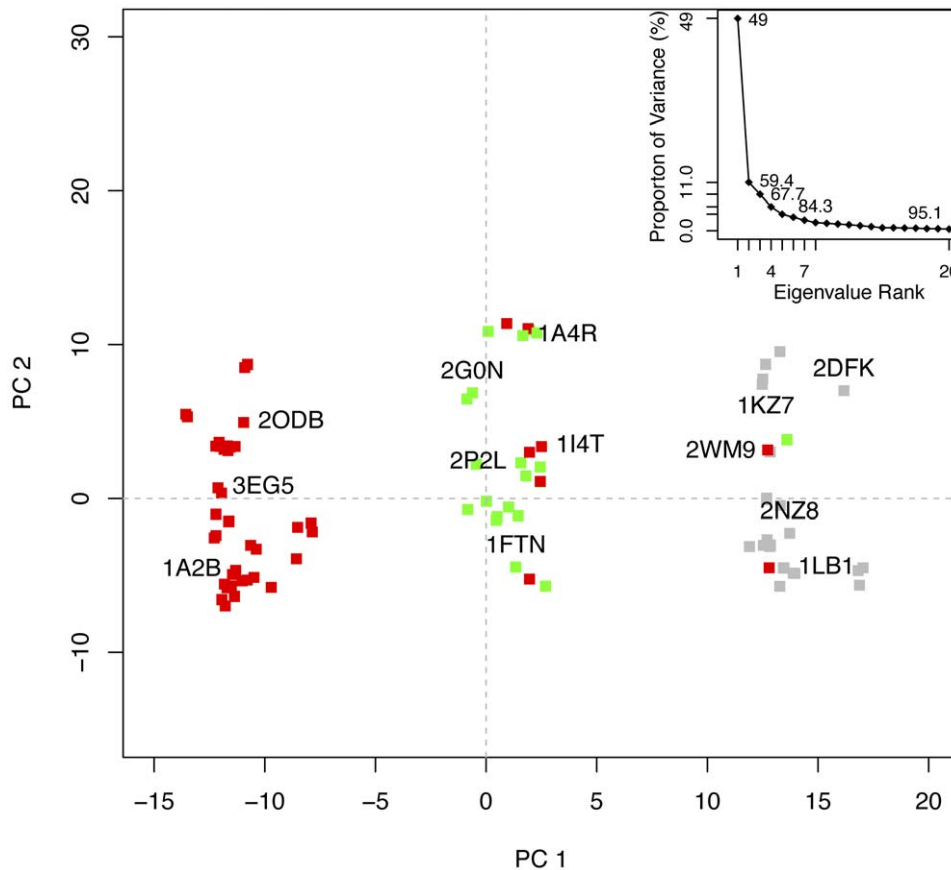
Figure 2. Results of PCA on Rho crystal structures. Conformer plot of available Rho GTPase crystal structures projected in the planes defined by the two most significant principal components (termed PC1 and PC2, see insert). Crystallographic GTP structures are colored red, GDP green, and nucleotide free gray. Inset: Eigenvalue spectrum: results obtained from diagonalisation of the covariance matrix of aligned atom coordinates from the Rho crystal structures. The magnitude of each eigenvalue is expressed as the percentage of the total variance (mean square fluctuation) captured by the corresponding eigenvector. Labels indicate the cumulative sum of the proportion of the total variance accounted for in all preceding eigenvectors.
doi:10.1371/journal.pone.0040809.g002

could inhibit the downstream signaling activity of Ras. Here we build and expand on this approach to exhaustively characterize potential small molecule binding sites on Rho.

Although many binding site search algorithms have been developed, [36] the predictions obtained with these algorithms are often specific to the single structure they were applied to. Here we employ a combination of over one hundred available Rho crystallographic structures and advanced molecular dynamics simulations to provide a more complete mapping of potential binding sites in all major conformational states. Conventional and accelerated [37] molecular dynamics simulations were used together for the first time in the structure-based characterization of binding sites. These simulations allowed us to identify less frequently visited Rho conformations, including those intermediate between GTP and GDP states. Both fragment and grid based mapping algorithms were employed to locate binding sites on both crystal and simulated structures. Our results reveal the presence of novel binding pockets in two relevant regions of Rho: close to the residues responsible for the GTP-GDP switching mechanism but outside the nucleotide binding site, and more distal sites that are allostericaly linked to the nucleotide and effector binding site regions. Some of these novel sites are more accessible in the GDP-bound (inactive) state of Rho, and present a potential novel mechanism of inhibition. Furthermore, we show that the advanced

sampling technique, accelerated molecular dynamics, can be used to enhance sampling of binding sites particularly when important conformational changes are intrinsic to the target protein activity.

Results and Discussion

Crystallographic Structure Analysis

High resolution structural data for the Rho family is particularly rich, given the current availability of 59 crystallographic structures comprising 98 distinct chains in the RCSB Protein Data Bank (date of access 02/10/2011). [38] These structures (see **Table S1**) span the three major nucleotide bound states (with GTP, GDP and nucleotide free). We employed principal component analysis (PCA) to examine the interconformer relationships within this crystallographic dataset (see Methods). Over 80% of the total mean-square displacement (or variance) of atom positional fluctuations was captured in six dimensions, 59.4% in two dimensions and 67.7% in three dimensions (see **Figure 2**). The first few principal components retain most of the variance in the original distribution and thus provide a useful description of the conformational space of the system. **Figure 2** depicts the projection of all Rho crystallographic structures onto the principal planes defined by the two most significant principal components. The first principal component is particularly informative and indicates the existence

of three major conformationally distinct categories. These categories, or conformational clusters, are largely consistent with the nature of the bound nucleotide in each structure and correlate with the results of clustering based on pairwise RMSD values (**Figure S1**). Note that there are some GTP bound structures (including PDB codes 2ATX, 2GCO and 2WMO) that most closely resemble GDP-bound (central green points) or nucleotide free (gray) structures. This is also evident for other structurally related nucleotide binding proteins, such as kinesin and myosin, [39] and suggests a complex underlying dynamic relationship between global conformation and nucleotide state. [39] For the purposes of this study we concluded that the ensemble of over 100 available Rho crystal structures contains significant representatives from all three nucleotide associated conformational states (42% GTP, 28% GDP and 30% APO), thus allowing us to search for potential binding sites in each major Rho conformation. Furthermore, we believe that the size and diversity of the crystallographic ensemble is sufficiently large to support our binding site analyses and comparisons to simulated conformers discussed below.

Binding Site Mapping of Crystal Structures

A fragment-mapping based approach (FTMAP) was used to locate potential small molecule binding hot spots on each of the Rho crystallographic structures. [40] Based on the same principle behind screens for small organic fragment binding by NMR and X-ray crystallography, FTMAP correlates pocket druggability with propensity to bind clusters of small organic compound fragments. This approach was found to outperform a number of other geometric and ligand based pocket identification methods in an earlier study. [35] To further verify mapping results for the current application, we also employed the energetic grid based SiteMap method, [41,42] (see Methods and **Figure S2** for details). In order to characterize the location of each potential binding site the probe occupancy per position was calculated as the fraction of structures in which a given residue can coordinate a fragment. A probe occupancy value of 1 indicates the existence of a proximal binding site in all crystallographic structures. Occupancy values of 0 indicate the complete absence of a binding site in the vicinity of a particular residue. Intermediate values reveal the intrinsic dynamic nature of certain binding pockets, present in some structures but absent in others. Note that a certain fraction of these later pockets would not be identified if only a single crystal structure representative was analyzed (**Figure S3**). **Figure 3** displays the results of this analysis together with Ras superfamily wide sequence conservation, relative solvent exposure and structural variability per position (see full details in Methods). Similar binding site profiles were obtained from SiteMap analysis (correlation value of r = 0.9, see **Figure S2**).

The highest probe occupancy values were found for the known nucleotide binding site region (residues 11–17, 31–37 and 56–59 corresponding to P-loop, switch1 and switch2 regions). This result confirms the presence of this sequence conserved solvent accessible cavity in all the analyzed crystal structures (**Figure 3A–C**). The P-loop region of this pocket (residues 11–17, encompassed by secondary elements β1, loop1 and α1) presents a relatively rigid substructure (with low RMSD values) consistent with its role in coordinating the nucleotide. In contrast, the conformationally responsive switch regions (residues 24–40 and 57–75) display higher structural fluctuations and slightly lower probe occupancy values (most notably in the highly dynamic switch 1 region, residues 31–37) (**Figure 3D**) Targeting this sequence conserved binding pocket could lead to undesirable off-target effects in other GTPases.

Targeting a distal allosteric pocket may offer significant selectivity and affinity advantages for Rho GTPase drug design. The probe occupancy analysis presented in **Figure 3A** provides a general landscape for such pocket locations over all crystallographic structures. However, information about the differences between the binding site distributions in the three major conformational states of Rho is also informative. Significant differences in the binding site landscape in the GTP, GDP and APO states could present an opportunity to stabilize one nucleotide-bound state by interacting with state specific binding pocket features. To investigate this possibility, we quantified the differences in absolute probe occupancy between the three major Rho states.

Figure 4A–B shows the absolute probe occupancy comparison between GTP/GDP and GDP/APO states. Orange squares along the x-axis indicate the position of residues with statistically significant differences (with a *p*-value less than 0.05). As shown, most of the significantly different regions, in terms of probe occupancy, between GTP and GDP structures lay in the NBS region. This is a useful positive control indicating the conformational differences in this pocket associated with the interconversion between the GTP and GDP forms. There is a noticeable lack of difference between GDP/APO states with respect to the previous GDP/GTP comparison. Since APO crystal structures are usually obtained by isolating Rho from the GDP/GTP exchange process, [43] after the GDP nucleotide is released (**Figure 1A**), it is reasonable that the NBS arrangement of the APO and GDP states share extensive similarities. **Figure 4C** displays probe occupancy results mapped to the molecular structure, indicating the most populated areas with an increasing red chain radii representation. Thick red regions represent high occupancy, whereas thin white segments correspond to low occupancy areas. This analysis clearly highlights the presence of distinct binding sites on each state of Rho. For example, sites centered the loop2 (GTP), α2 (GDP) and α1 (APO) regions.

Several regions exhibit a marked preference for higher probe occupancy in GDP structures than GTP and APO. The most prominent correspond to the clusters of Trp-55, Asp-56, Thr-57, Ala-58, Gly-59 and Asp-67, Thr-68, Ala-69, Gly-70, Gln-71, located in the β3 and α2 structural elements, respectively. These groups of residues lay in close proximity to loop 4 (Switch 2), one of the regulatory regions for GTP-GDP exchange in Rho. In addition, residues Thr-107, Pro-108 and Ile-109 also form a group of amino acids with significant probe occupancy differences between GTP and GDP states, and favored in the latter. These residues reside in Loop 7 close to the pocket formed by residues 149, 161 and 153 (also detected in this analysis) and are believed to be an allosteric modulator region in Ras. [34] Indeed, residues in the vicinity of these pockets are involved in correlated motions with the active site (**Figure S4**), suggesting that ligand binding at these pockets may allosterically modulate the active site architecture, as shown in other proteins. [44,45,46] This is consistent with the report that divalent ion-cyclen binding at an analogous pocket to that flanked by Loop 7 in Ras stabilizes a conformation that has weak effector-binding potential. [47].

In summary, the large number of available Rho crystal structures has allowed us to locate more binding pockets than could be found from analysis of individual structures. In addition, the comprehensive nature of the dataset allowed us to track how the accessibility of these pockets varies as a function of conformation, consistent with their distinct functionally relevant nucleotide state. Given the clear nucleotide related clustering of crystal structures evident in the PCA analysis and their distinct pocket properties evident from fragment mapping (**Figures 2** and **4**), it was hypothesized that additional
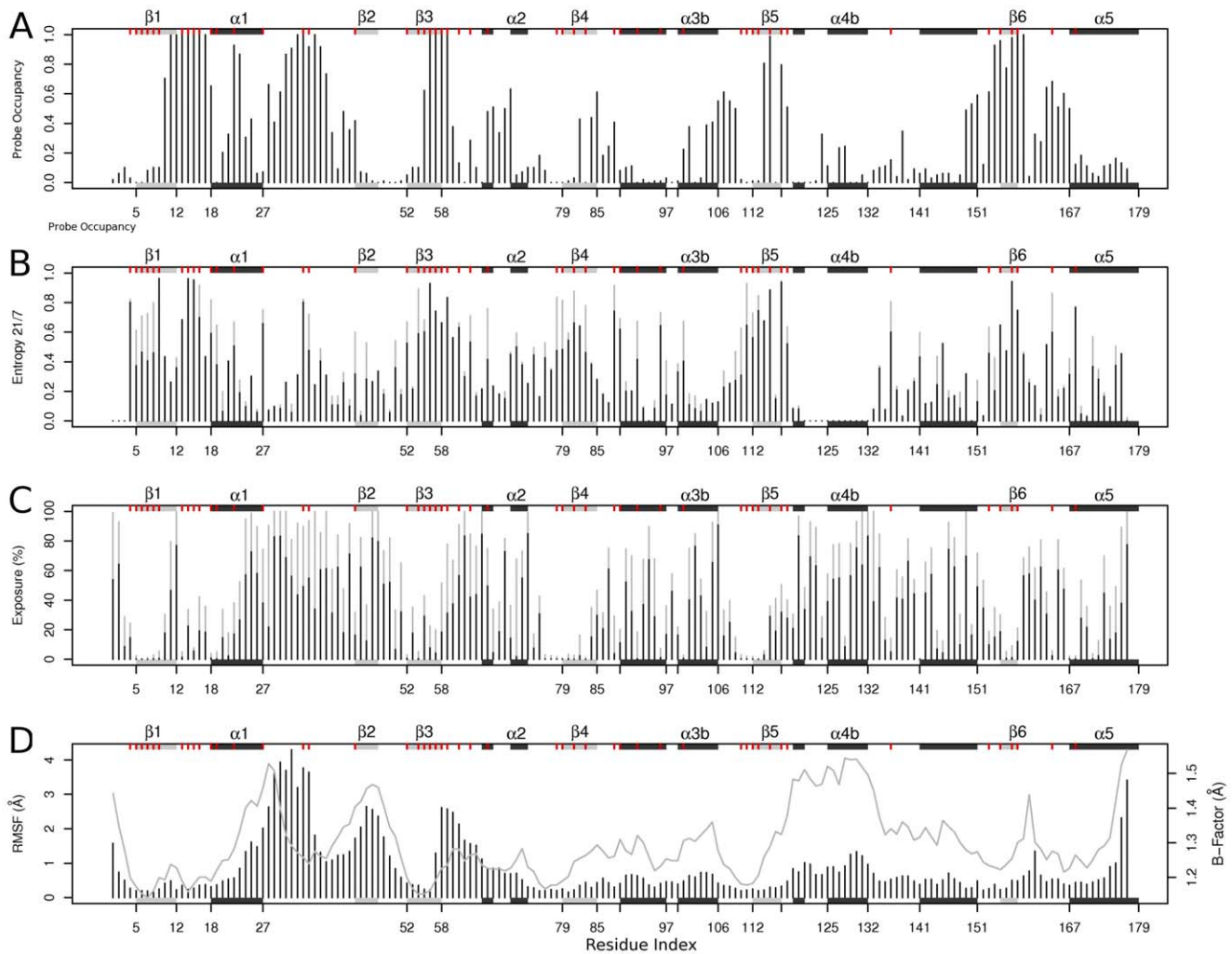
Figure 3. Fragment binding site characterization in the crystallographic dataset. (A) Fragment probe occupancy per position in Rho across all available structures. (B) Sequence conservation entropy scores for a 21-letter alphabet (20 amino acids and a gap, in black) and 7-letter alphabet (where amino acids are grouped into six classes based on their physicochemical properties, in light gray). (C) The mean (black) and maximum (light gray) solvent exposure per position in all structures. (D) Flexibility per position in terms of RMSF (bars) and mean *B*-factor (line) over all structures. Major elements of secondary structure (shaded rectangles) and positions with a high degree of sequence conservation (red ticks) are indicated in the marginal areas of each plot to facilitate comparison. Both the secondary structure and residue numbering are according to the RhoA-GDP complex from *Homo Sapiens* (PDB code 1FTN).
doi:10.1371/journal.pone.0040809.g003

conformations not evident in the crystallographic dataset (including those intermediate between the three major conformational states), may also be relevant for pocket determination. In order to investigate this further, and hence perform a more exhaustive search of binding sites, we have employed both conventional and accelerated molecular dynamics simulations.

Conventional Molecular Dynamics

We sought to investigate whether a single or multiple conventional molecular dynamics simulations (cMD) could resolve new pockets in addition to those evident in the crystallographic dataset. We first carried out three 30 ns long cMD simulations, commencing from each of the three major conformational states (PDB codes for the initial structure models are 1A2B, 1FTN and 1LB1 for the GTP GDP and APO states respectively, see **Figure 2**). Conformations from each trajectory were analyzed with the FTMap method and the probe occupancy per position computed as described previously (**Figure 5**).

Figure 5A depicts the probe occupancy per residue profile obtained for the combined cMD trajectory conformers (orange) together with that obtained from analysis of the complete ensemble of crystal structures (gray). Also shown in **Figure 5B** are results obtained for a single GTP initiated trajectory (red). Blue and green points at the bottom of each plot indicate residue positions found to interact directly with probe molecules in one dataset but not in the other (blue for those absent in the trajectory conformers and green for those absent in crystal structures). Clearly multiple cMD simulations are required to obtain a binding site distribution comparable to that evident in the large distribution of crystal structures (note the large number of blue points obtained when only a single trajectory is considered, **Figure 5B**). This is also evident in the higher Pearson correlation coefficient for the probe occupancy per residue between crystal structures and the combined trajectories (0.92) and obtained with individual GTP (0.84), GDP (0.82) or APO (0.89) state initiated trajectories. The combined cMD data also highlights a number of
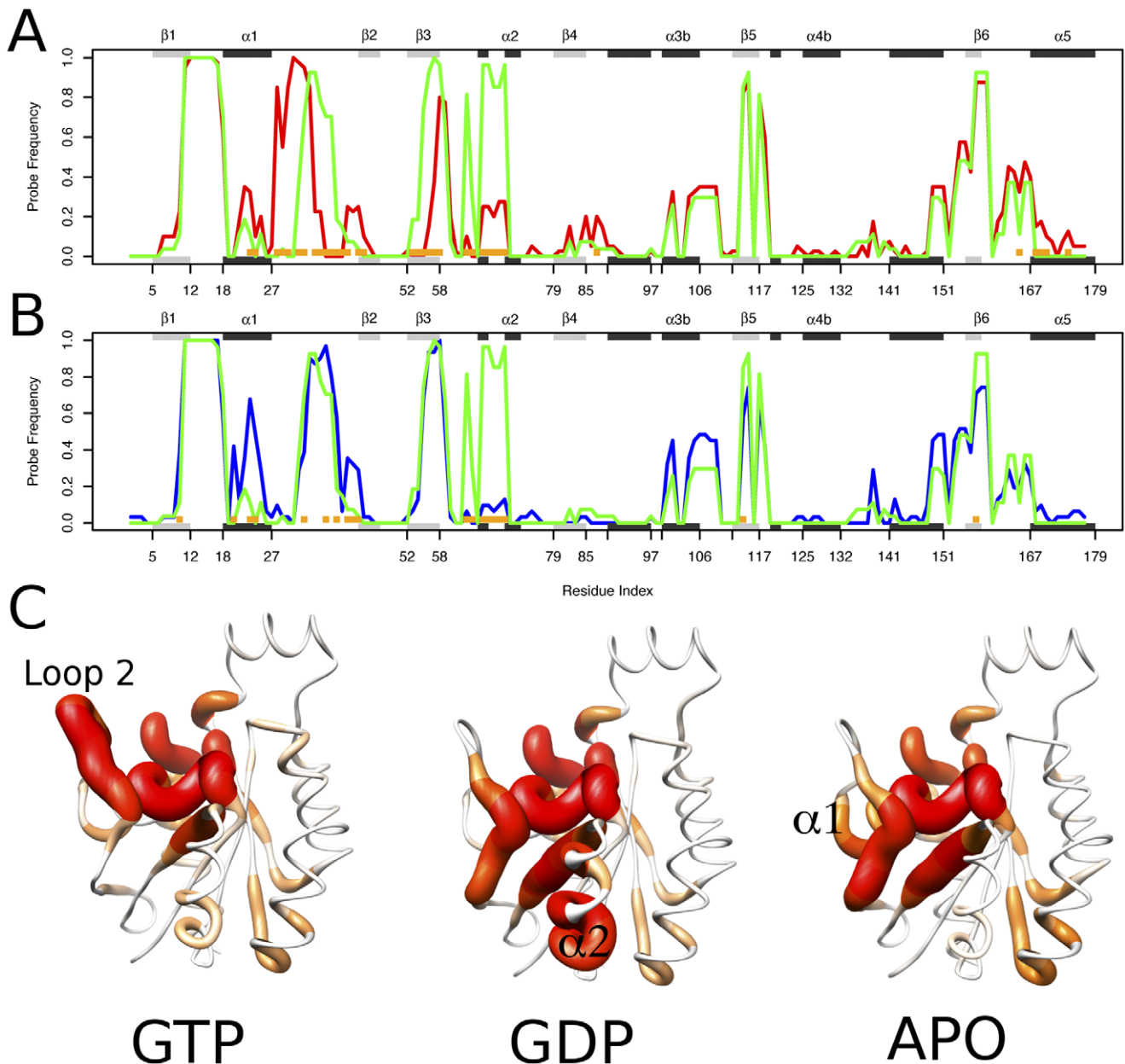
Figure 4. Nucleotide state associated differences in probe occupancy. (A) Comparison of the absolute probe occupancy per position between GDP/GTP and (B) GDP/APO structure sets. Probe occupancies are colored in red, green and blue for GTP, GDP and APO states. Dashed lines indicate positions in the nucleotide binding site. Orange squares indicate residues with statistically significant differences in probe occupancy (p<0.05) between each structure set. Major elements of secondary structure (shaded rectangles) are indicated in the marginal top areas of each plot. (C) Increasing colored radii representation of the absolute probe occupancy per residue in the GTP, GDP and APO crystal structure ensembles.
doi:10.1371/journal.pone.0040809.g004

new binding site positions not evident in the crystal structures (green points in **Figure 5A**). These positions (including residues 68 to 106) were found to cluster in the three dimensional structure defining an enlarged pocket between the α2 and α3 structural elements distal to the NBS (**Figure 5C**).

Principal component and RMSD analysis was used to assess the conformational space sampled in our cMD simulations and its relation to available crystal structures. **Figure 6A** shows the distribution of sampled conformers (density-shaded red, green and blue points for GTP, GDP and APO simulated systems) along with available crystal structures (black points) projected onto the dominant eigenvectors obtained from analysis of the crystal

structure ensemble described previously. The GTP and GDP simulations were found to exhibit a relatively restricted sampling that is localized to regions around the corresponding cluster of crystallographic structures. The lack of overlap of these distributions is consistent with the high minimal inter trajectory RMSD value (1.33 Å) between GDP and GTP simulations. This data indicate an absence of interconversion between the GTP and GDP states under these simulation conditions. In contrast, the APO trajectory (blue points in **Figure 6A**) displays a wider sampling pattern that encompasses GDP like conformers and displays a low minimal RMSD to the GDP trajectory (0.73 Å). This is consistent with results from our previous studies of Ras where conformational
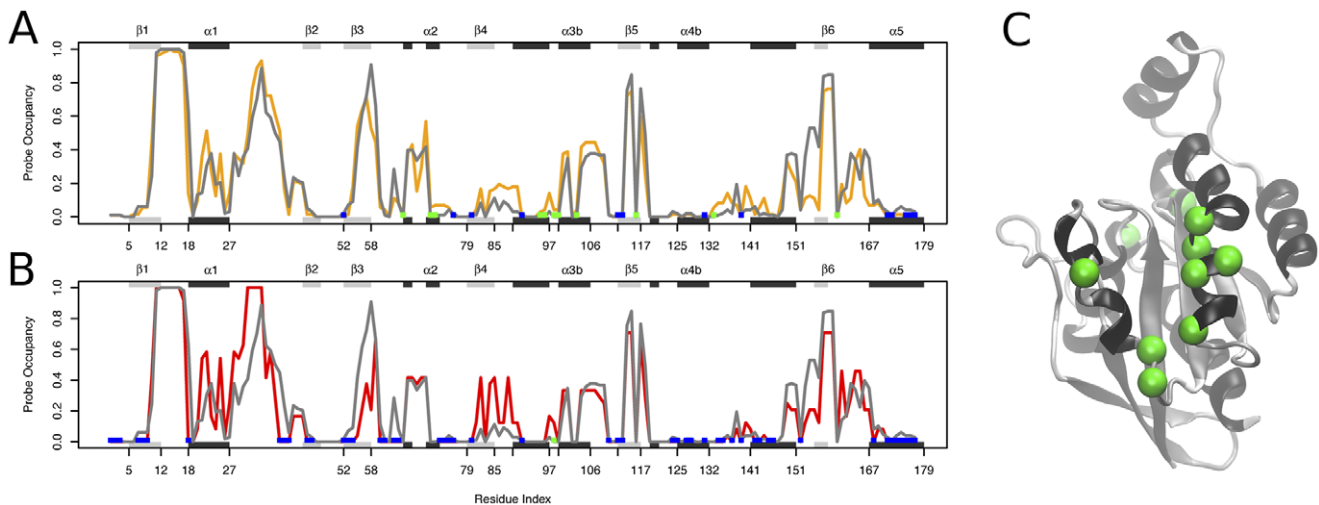
Figure 5. Probe binding results from conventional molecular dynamics simulations. The probe occupancy for the ensemble of crystal structures (gray) and the **A** combined conventional molecular dynamics (cMD) conformers (orange), and **B** an individual cMD GTP trajectory (red). Blue and green squares at the bottom of each plot highlight those residues that are involved in a binding pocket only in the crystal structure data set (blue) or the trajectory dataset (green). **C** Binding site positions found in cMD that are not observed in any crystallographic structure, see text for details.
doi:10.1371/journal.pone.0040809.g005

sampling was observed to be more restricted in the presence of the bound nucleotide than the sampling obtained in the absence of nucleotide. [32].

In summary, individual cMD simulations did not recover the full binding site distribution evident in the large set of crystallographic structures. Multiple combined cMD simulations were more successful in mapping potential binding sites and highlighted a number of additional potential pocket interactions not recovered in the crystallographic structures. These novel potential interactions are directly related to the observation of new binding site configurations not evident in the ensemble of available crystal structures. Most notably, the cMD simulations provide informa-

tion about the intermediate conformations between the GDP and APO states. The inclusion of these intermediate conformations reveals novel potential pocket interactions that are favored in the GDP/APO states. However, the current cMD simulations provide no information about conformations intermediate between GTP and GDP like states. Given the enhanced predictive performance of multiple cMD simulations and their ability to map novel pocket features, we pursued a more exhaustive conformational search by enhancing the exploration of conformational space using accelerated molecular dynamics simulations.
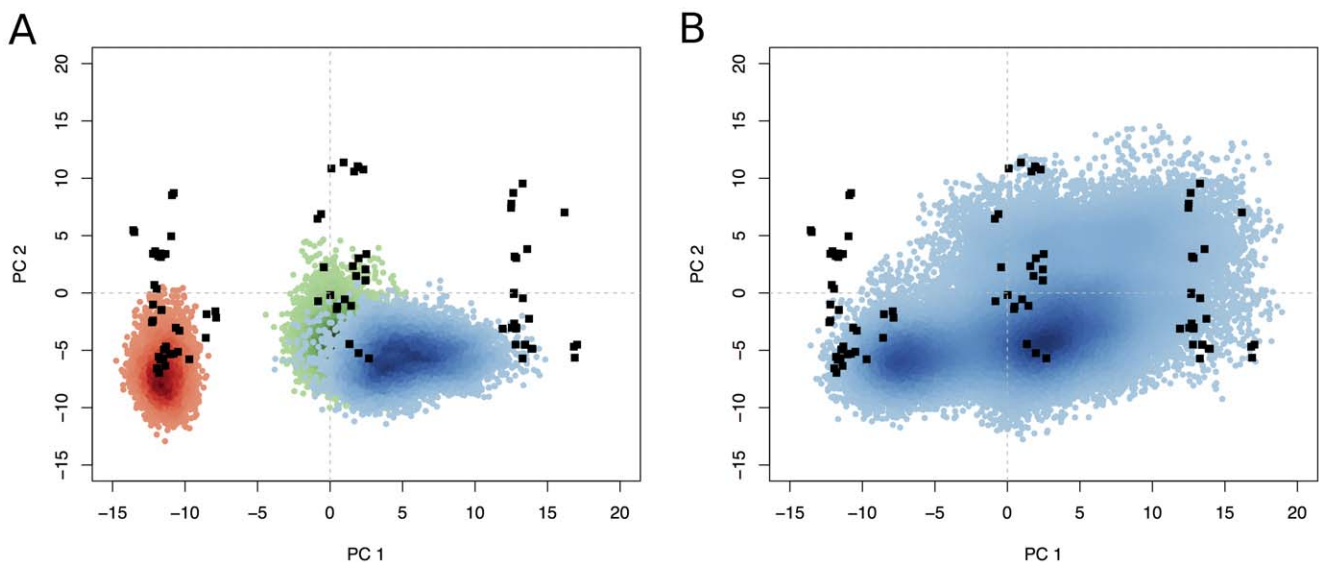
Figure 6. Conformational sampling in conventional and accelerated molecular dynamics simulations. The projection of (A) cMD and (B) aMD conformers onto the first two principal components of the crystallographic ensemble. The distribution of simulated conformers is depicted with density-shaded red, green and blue points for the GTP, GDP and APO simulations respectively. Crystal structures conformations are in black (see **Figure 2** and main text for details).
doi:10.1371/journal.pone.0040809.g006

Accelerated Molecular Dynamics

Accelerated MD (aMD) simulations modify the potential energy landscape explored in cMD by raising the energy minima thus lowering the effective barriers between states resulting in an enhanced degree of conformational sampling (see Methods for details). In order to test the utility of aMD simulations for pocket identification with respect to our cMD and crystallographic datasets, we performed a single 60 ns long aMD simulation, commencing from a single nucleotide free structure. Principal component analysis was again used to assess the conformational space explored by the aMD simulation (**Figure 5B** density-shaded blue points). Results indicate that the aMD simulation explores a wider region of conformational space than individual, or indeed cumulative, cMD simulations (see **Figure 5**A). In addition, the aMD simulation provides details on the intermediate regions between all three major conformational states of Rho.

The results of applying FTMap and subsequent probe analysis to aMD derived conformers are shown in **Figure 7**. Also shown are the results obtained for the ensemble of crystal structures, a single cMD trajectory and the combined cMD trajectories. The overall similarity of results between the four datasets is clearly apparent. The major hot spots detected in the ensemble of crystal structures are also observed in our three cMD trajectories and single aMD trajectory. For example the Asp-67, Thr-68, Ala-69, Gly-70, Gln-71 cluster in α2. However, in general, probe occupancies are notably higher in the identified distal regions of the aMD conformations. This provides evidence that the accelerated simulation of a single structure can sample more accessible hotspots, by exploring a larger region of the conformational space, than that contained in either our large set of crystal structures or combined cMD trajectories. In addition, some binding pockets observed in the aMD conformations were present at very low or nearly negligible probe occupancy in the crystal or cMD datasets. There is a noticeably higher occupancy in the α3, β5, α3 and Loop 3 regions in **Figure 7D** compared to **Figures 7A** and **7C** reflecting the sampling of more open pocket conformations in these regions. For example the crystal structures (**Figures 7A**) present a binding pocket in the region of α2 facing β1. In the aMD ensemble (**Figure 7D**), a similar binding pocket is also observed in α2, but expanded onto β5 (the opposite direction).

Targeting Predicted Binding Sites

Using the Schrödinger Glide package, [48] compounds from the NCIDS II were docked against the identified α2 and Loop 3 pockets in each structure that reported fragment-binding activity in these regions. Four high-ranking compounds (compound identifiers: 2561, 25740, 157767 and 335504) were selected and their effect on thrombin stimulated RhoA activation in 1321N1 glioblastoma cells assessed (see Methods). Thrombin is a potent mitogen for 1321N1 glioblastoma cells and elicits its cellular responses through activation of PAR-1 receptors, which in turn leads to activation of RhoA. [49,50,51] Cells were pre-treated with the compounds at a concentration of 30 μM, subsequently stimulated with thrombin for 15 minutes and RhoA pull down assays were performed to assess increases in GTP-bound RhoA. As shown in **Figure 8**, three of the four compounds (335504, 25740 and 2561) significantly inhibited thrombin-induced RhoA activation by approximately 70–80% relative to vehicle control, while compound 157767 had no significant effect. These inhibitors did not significantly affect resting (unstimulated) levels of active RhoA. These data suggest that computationally selected compounds directed against our

novel pockets have the ability to block RhoA activation induced by agonist stimulation in glioblastoma cells.

In summary, our results demonstrate that both cMD and aMD simulations can facilitate the identification of potential small molecule binding sites on Rho proteins. Application of aMD is expected to be particularly valuable when large distributions of crystal structures are not available and when the target protein is known to undergo important conformational changes as a result of its biological activity. Furthermore, the results of virtual screening and subsequent preliminary experimental testing support the hypothesis that distal pockets can potentially interact with small drug-like compounds that may attenuate Rho activity.

Conclusions

In this work we have conducted an exhaustive search of potential binding pockets of the G-protein Rho, focusing on less conserved binding pockets, distal to the NBS. We have performed our study in an incremental fashion, progressively increasing the sophistication of our methodological approach, whilst reducing the number of analyzed structures. We have proceeded in this way in order to find an optimal compromise between the maximum number of results obtained and the amount of initial information required. To this end, we first sampled the binding site landscape of Rho considering every available crystallographic structure from the RCSB Protein Data Bank, [38] taking into account the three major conformational states revealed by PCA (GTP, GDP and APO). Secondly, we performed cMD simulations on three representative Rho structures, one for each state, and analyzed each simulation for new binding sites. Finally, we analyzed binding sites from an aMD simulation commenced from a single Rho structure.

Our results demonstrate that the binding site landscape of Rho is highly dependent on its conformational state, which is in turn modulated by nucleotide turnover. This necessitates the identification of binding pockets and cognate ligands that account for multiple distinct conformations. Here we have identified a number of novel binding pockets, some of them not present in any resolved crystallographic structure known to date, including sites distal to the NBS and proximal to Loop 3 and Loop 7. In addition, some of these pockets are more accessible in the GDP state, compared with GTP or APO conformations. We postulate that the interaction of drug-like compounds with these pockets may alter the distribution of active and inactive Rho conformations and thereby deregulate Rho's activity. A more exhaustive virtual screening and rational drug design study, as well as experimental validation is being performed currently to further assess these novel pockets. We note that our preliminary results from these studies are encouraging in that micromolar concentrations of candidate inhibitors directed against these pockets block thrombin stimulated RhoA activation by approximately 70–80% in cell lines. This suggests that the predicted compounds could serve as starting points for lead generation targeted against ligand-induced (or constitutively elevated) Rho activity in cancer cells. We are currently comparing the affinity of these compounds using full dose response curves and have determined that at least for compound 335504, 70% inhibition of thrombin-induced RhoA activation at 10 μM. However, further work is required to definitively characterize the drugability of our predicted sites and compounds.

Finally, we find that dynamic simulations, both conventional and accelerated, can be used as a complementary tool to perform exhaustive binding site identification. Both techniques can enhance sampling of binding pockets, and accelerated MD seems particularly useful given extensive conformational changes. These
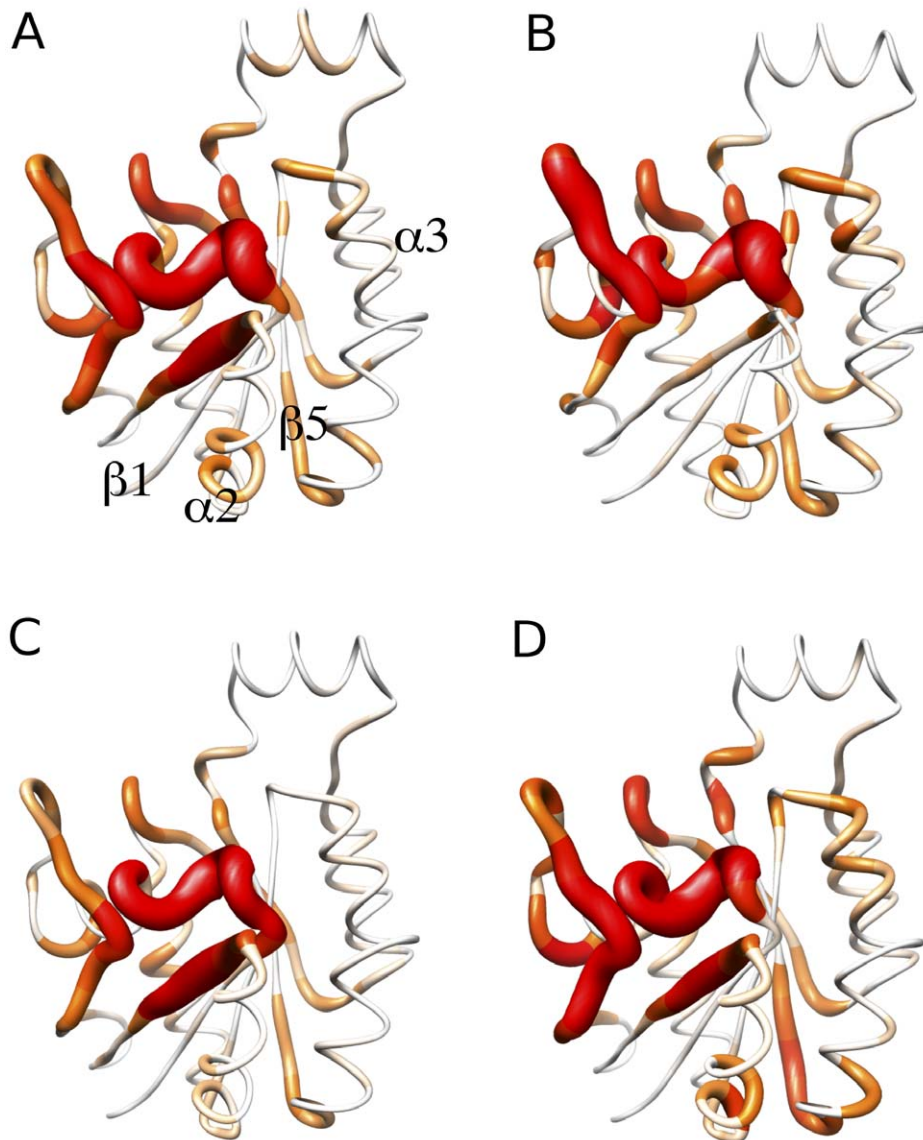
Figure 7. Absolute probe occupancy results from fragment mapping analysis. (A) results from the ensemble of crystal structures, (B) the combined GTP cMD conformers, (C) the combined cMD conformers and (D) the single aMD trajectory conformers. The protein structure of reference corresponds to the RhoA-GDP complex from *Homo Sapiens* (PDB code 1FTN).
doi:10.1371/journal.pone.0040809.g007

methods aid in finding the maximal amount of significant binding sites with a minimal amount of input information.

Materials and Methods

Crystal Structures Analysis

The Bio3D package [52] was used to query and analyze all available Rho structures in the RCSB Protein Data Bank. [38] Principal component analysis was employed to examine the conformational relationships between superposed crystal structures and simulated conformers. The application of PCA to distributions of experimental structures and molecular dynamics 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. [53,54,55,56,57,58] Briefly, PCA is based on the diagonalization of the covariance matrix, C, with elements C_{ij} built from the Cartesian coordinates, r, of the superimposed Rho structures (eq. 1):

$$C_{ij} = \langle (r_i - \langle r_i \rangle)(r_j - \langle r_j \rangle) \rangle \qquad (1)$$

where i and j represent all possible pairs of 3N Cartesian coordinates (where N is the number of atoms) being considered. The eigenvectors of the covariance matrix correspond to a linear basis set of the distribution of structures, referred to as PCs, whereas the eigenvalues provide the variance of the distribution along the corresponding eigenvectors.

Sequence Conservation Analysis

The PFAM alignment PF00071 was used as a basis for examining sequence conservation within the wider Ras superfamily. [59] To assess the level of conservation at each position in the alignment, the entropy per position was calculated. "Entropy" is based on Shannon's information entropy for both a 21-letter
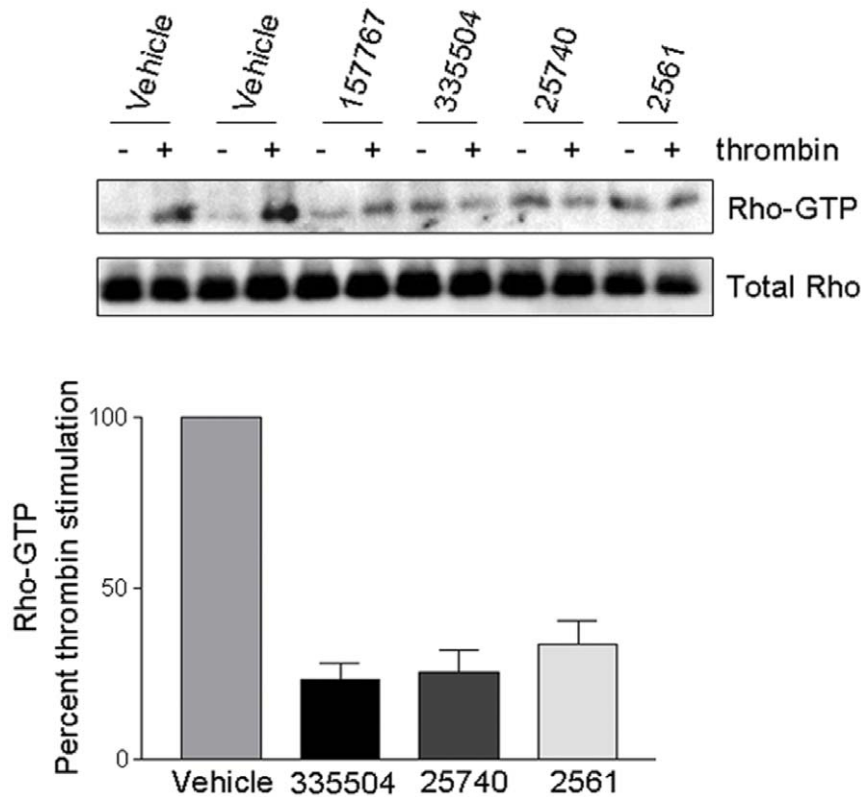
Figure 8. Putative Rho compounds inhibit thrombin stimulated RhoA activation in glioblastoma cells. Compounds were selected from a pocket directed virtual screen and tested on Glioblastoma cells (1321N1). Cells were pre-treated for 1 hr with either DMSO control or 30 μM of compound 157767, 335504, 25740 or 2561 and subsequently stimulated for 15 minutes with either vehicle or 0.5 U/mL thrombin. RhoA activity was assessed via pull-down by rhotekin and total RhoA was determined using an anti-RhoA antibody. Bar graph shows densitometric results from 4 independent experiments. Data are presented as means ± SEM; *P<0.001 vs. vehicle control, one-way ANOVA.
doi:10.1371/journal.pone.0040809.g008

alphabet (20 amino acids and a gap character) and a seven-letter alphabet (six groups of amino acids and a gap character) [60,61] (eq. 2):

$$S = -\sum_{i}^{N} p_i \log_2 p_i \qquad (2)$$

where S is Shannon's entropy, p_i is the frequency of each alphabet's letter at position i and N is the alphabet's size (7 or 21). The six groups chosen were aliphatic (A, V, L, I, M and C), aromatic (F, W, Y and H), polar (S, T, N and Q), positive (K and R), negative (D and E), and finally special conformations (G and P). Entropy scores plotted in **Figure 3** are normalized so that conserved (low entropy) columns score 1 and diverse (high entropy) columns score 0 (eq. 3):

$$C = \frac{-\sum_{i}^{N} p_i \log_2 p_i}{\log_2\left(\min\left(N_{seq}, N\right)\right)} \qquad (3)$$

where C is the normalized entropy, p_i is the frequency of each alphabet's letter at position i, N is the alphabet's size and N_{seq} is the length of the sequence. We define a position to be conserved if the entropy 21 or entropy 7 at a position is >0.6. Positions in which more than 30% of the sequences have gaps were excluded from all sequence conservation analysis.

Exposed Surface Area

Percent solvent exposure per position was calculated with the NACCESS program. [62] A residue was considered to be exposed when the accessible surface area of the residue was more than 40% of the measured accessible surface area of that residue in an extended Ala-X-Ala tripeptide context.

Binding Site Mapping

We used the FTMap method of Brenke and co-works to highlight protein surface regions that have the potential to bind the highest number of small molecular probes. [40] Both crystal structures and each cluster representative form cMD and aMD were subject to fragment mapping. Hot-spot residues (those that comprise prominent fragment binding sites) were analyzed across all structures. A residue was assumed to be in contact with a probe molecule (and assigned a probe occupancy value of 1 for a given structure) if any two heavy atoms from the probe and residue were closer than 5.0 Å. The probe occupancy per residue in a given set of structures was obtained by normalizing the summed probe occupancy count for each residue by the number of structures in the input dataset.

We also analyzed crystal structure and trajectory conformers with the SiteMap method from Schrödinger, Inc. New York. [41,42] SiteMap identifies potential binding sites based on geometric and energetic properties, linking together proximal grid points based on how close the grid points are to the protein surface and how well sheltered they are from solvent. [41,42] For our purposes, a protein residue from a particular structure was

assumed to be part of a potential binding site if a favorable SiteMap grid point was within 5.0 Å of any heavy atom on that residue. The occupancy per residue for a given structure set was then obtained by normalizing the counts for each equivalent residue by the number of structures in the input dataset.

Molecular Dynamics Simulations

Simulations were performed using the AMBER10 package [63] and corresponding all-atom potential function ff99SB. [64] Operational parameters include periodic boundary conditions, TIP3P water and charge-neutralizing counter ions, with full particle-mesh Ewald electrostatics. A 2fs time step and a 10 Å cutoff were used for the truncation of VDW non-bonded interactions. Constant volume heating (to 300 K) was performed over 10 ps, followed by constant temperature (300 K), constant pressure (1 atm) equilibration for an additional 200 ps. Finally, constant pressure constant temperature production dynamics was performed for both conventional and accelerated MD implementations. The SHAKE algorithm was used to constrain all covalent bonds involving hydrogen atoms. In order to simultaneously enhance the sampling of internal and diffusive degrees of freedom a dual boosting aMD approach was employed, based on separate torsional and total boost potentials. [37,65] aMD starting structures and standard operational parameters were identical to those used for cMD. The energy level, E, below which the boost is applied and tuning parameter, α, that modulates the depth and local roughness of basins in the modified potential, were based on a previous work. [65].

Average-linkage hierarchical clustering based on pairwise RMSD analysis was used to inspect the dominant conformations sampled by each simulation. Inspection of the clustering dendogram was used to partition structures from each trajectory into 24 dominant groups (ranked according to their populations). The closest structure to the average structure from each cluster, in terms of RMSD, was chosen as a representative conformation for further fragment mapping analysis.

Virtual Screening

Schrödinger's Glide package [48] was used to screen a subset of the NCIDS II small molecule library [66] against $\alpha 2$ and Loop 3 predicted distal binding pockets. The full NCIDS II (comprising 3881 compounds) was first filtered (to yield a subset of 2291 compounds) by applying the following criteria: molecular weight <553 Da, number of rotatable bonds ≤14, number of potential donor Hydrogen bond atoms ≤6, number of potential acceptor Hydrogen bond atoms ≥3, Hydrophilic polar surface area ≥30 Å, and predicted octanol/water partitioning ≥ −0.5 QlogP. Screening against pocket grids generated for each major conformation in crystallographic and simulated conformers (78 total grids) was performed at the XP level as described previously. [35] Only the compounds with a docking score of −8 or lower were collected and ranked from all grids for each pocket.

Rho Activation Assay

An affinity pull-down assay using a glutathione S-transferase (GST) fusion protein of the RhoA binding domain of the RhoA effector rhotekin was performed to determine RhoA activation in 1321N1 glioblastoma cells. Cells were lysed in buffer containing 50 mM Tris, pH 7.4, 0.1% Triton X-100, 150 mM NaCl, 5 mM MgCl2 and 10% glycerol, supplemented with protease and phosphatase inhibitors. Cell lysates were clarified by centrifugation and subsequently incubated with the sepharose-bound GST-rhotekin-RhoA binding domain for 50 minutes at 4°C. The beads and the precipitated proteins were washed, boiled and resolved by SDS-PAGE. Total RhoA was detected in the whole cell lysate by immunoblotting with a RhoA antibody (Santa Cruz). GTP-bound RhoA was normalized to total RhoA.

Supporting Information

Figure S1 The results of RMSD based clustering of available Rho structures. Structure labels are colored by nucleotide state (red for GTP, green for GDP, pink for GXP and gray for nucleotide free).
(JPG)

Figure S2 Binding site characterization with FTMap and SiteMap. FTMap calculated probe occupancy (black) and SiteMap calculated grid occupancy (gray) per position across all available Rho crystal structures (correlation value, r = 0.9). See main text for further details.
(JPG)

Figure S3 Binding site characterization in single *versus* multiple structures. Fragment probe occupancy per position in a single crystal structure (PDB code 1FTN, gray bars) and averaged across all available crystal structures (black lines, see main text for details).
(JPG)

Figure S4 Residue-residue plot of correlated motions. The extent of correlation for all equivalent residue pairs (of Cα atomic displacement) during Rho (lower triangle) and Ras (upper triangle) aMD simulations. The correspondence of Rho and Ras residues was determined from structural alignment with gap positions indicated with a broken axis line (see Ras positions 122–136 that represent a Rho specific insert). The color scale runs from pink (for values ranging between −1 to −0.75), through white (−0.25 to 0.25) to cyan (0.75 to 1). Negative values are indicative of displacements along opposite directions, namely anticorrelated motions, whereas positive values depict correlated motions occurring along the same direction. Major secondary structure elements of Rho are indicated schematically with helices in black and strands in gray. All calculations were were performed with the Bio3D package.
(JPG)

Table S1 Ensemble of resolved Rho crystal structures used in this work.
(DOC)

Author Contributions

Conceived and designed the experiments: JMO BJG JAM. Performed the experiments: JMO BJG JS. Analyzed the data: JMO SEN BJG JS. Contributed reagents/materials/analysis tools: JMO BJG JS. Wrote the paper: JMO SEN BJG JAM JS JHB.

References

1. Treisman R, Miralles F, Posern G, Zaromytidou AI (2003) Actin dynamics control SRF activity by regulation of its coactivator MAL. Cell 113: 329–342.

2. Hall A (2005) Rho GTPases and the control of cell behaviour. Biochem. Soc. Trans. 33: 891–895.

3. Jaffe AB, Hall A (2005) Rho GTPases: Biochemistry and biology. Annu. Rev. Cell Dev. Biol. 21: 247–269.

4. Carlier MF, Ducruix A, Pantaloni D (1999) Signalling to actin: the Cdc42-N-WASP-Arp2/3 connection. Chem. Biol. 6: R235-R240.

5. Hall A, Etienne-Manneville S (2002) Rho GTPases in cell biology. Nature 420: 629–635.

6. Sprang SR (1997) G proteins, effectors and GAPs: structure and mechanism. Curr. Opin. Struct. Biol. 7: 849–856.

7. Vetter IR, Wittinghofer A (2001) Signal transduction - The guanine nucleotide-binding switch in three dimensions. Science 294: 1299–1304.

8. Boettner B, Van Aelst L (2002) The role of Rho GTPases in disease development. Gene 286: 155–174.

9. Vega FM, Ridley AJ (2008) Rho GTPases in cancer cell biology. Febs. Letters 582: 2093–2101.

10. Li HL, Deng J, Feng EG, Ma S, Zhang Y, et al. (2011) Design and Synthesis of Small Molecule RhoA Inhibitors: A New Promising Therapy for Cardiovascular Diseases? J. Med. Chem. 54: 4508–4522.

11. Kitaoka Y, Kitaoka Y, Kumai T, Lam TT, Kuribayashi K, et al. (2004) Involvement of RhoA and possible neuroprotective effect of fasudil, a Rho kinase inhibitor, in NMDA-induced neurotoxicity in the rat retina. Brain Res. 1018: 111–118.

12. Fritz G, Just I, Kaina B (1999) Rho GTPases are over-expressed in human tumors. Int. J. Cancer 81: 682–687.

13. Ellenbroek SIJ, Collard JG (2007) Rho GTPases: functions and association with cancer. Clin. Exp. Metastas. 24: 657–672.

14. Sahai E, Marshall CJ (2002) Rho-Gtpases and Cancer. Nat. Rev. Cancer 2: 133–147.

15. Preudhomme C, Roumier C, Hildebrand MP, Dallery-Prudhomme E, Lantoine D, et al. (2000) Nonrandom 4p13 rearrangements of the RhoH/TTF gene, encoding a GTP-binding protein, in non-Hodgkin's lymphoma and multiple myeloma. Oncogene 19: 2023–2032.

16. Grise F, Bidaud A, Moreau V (2009) Rho GTPases in hepatocellular carcinoma. Biochimica Et Biophysica Acta-Rev. Cancer 1795: 137–151.

17. Zhao DZ, Pothoulakis C (2003) Rho GTPases as therapeutic targets for the treatment of inflammatory diseases. Expert Opin. Ther. Targets 7: 583–592.

18. Budzyn K, Marley PD, Sobey CG (2006) Targeting Rho and Rho-kinase in the treatment of cardiovascular disease. Trends Pharmacol. Sci. 27: 97–104.

19. Aznar S, Fernandez-Valeron P, Espina C, Lacal JC (2004) Rho GTPases: potential candidates for anticancer therapy. Cancer Lett. 206: 181–191.

20. Zheng Y, Nassar N, Cancelas J, Zheng J, Williams DA (2006) Structure-function based design of small molecule inhibitors targeting Rho family GTPases. Curr. Top. Med. Chem. 6: 1109–1116.

21. Fritz G, Kaina B (2006) Rho GTPases: Promising cellular targets for novel anticancer drugs. Curr. Cancer Drug Targets 6: 1–14.

22. Riganti C, Aldieri E, Doublier S, Bosia A, Ghigo D (2008) Statins-mediated inhibition of Rho GTPases as a potential tool in anti-tumor therapy. Mini-Rev. Med. Chem. 8: 609–618.

23. Lu Q, Longo FM, Zhou HC, Massa SM, Chen YH (2009) Signaling Through Rho GTPase Pathway as Viable Drug Target. Curr. Med. Chem. 16: 1355–1365.

24. Sun HW, Tong SL, He J, Wang Q, Zou L, et al. (2007) RhoA and RhoC-siRNA inhibit the proliferation and invasiveness activity of human gastric carcinoma by Rho/PI3K/Akt pathway. World Journal of Gastroenterology 13: 3517–3522.

25. Sebti SM, Der CJ (2003) Opinion - Searching for the elusive targets of farnesyltransferase inhibitors. Nat. Rev. Cancer 3: 945–951.

26. Sebti SM (2005) Protein farnesylation: Implications for normal physiology, malignant transformation, and cancer therapy. Cancer Cell 7: 297–300.

27. Gelb MH, Brunsveld L, Hrycyna CA, Michaelis S, Tamanoi F, et al. (2006) Therapeutic intervention based on protein prenylation and associated modifications. Nat. Chem. Biol. 2: 518–528.

28. Liao DK, Seto M, Noma K (2007) Rho kinase (ROCK) inhibitors. J. Cardiovasc. Pharmacol. 50: 17–24.

29. Chen YT, Bannister TD, Weiser A, Griffin E, Lin L, et al. (2008) Chroman-3-amides as potent Rho kinase inhibitors. Bioorg. Med. Chem. Lett. 18: 6406–6409.

30. Schirok H, Kast R, Figueroa-Perez S, Bennabi S, Gnoth MJ, et al. (2008) Design and Synthesis of Potent and Selective Azaindole-Based Rho Kinase (ROCK) Inhibitors. ChemMedChem 3: 1893–1904.

31. Fang XG, Yin Y, Chen YT, Yao L, Wang B, et al. (2010) Tetrahydroisoquino-line Derivatives As Highly Selective and Potent Rho Kinase Inhibitors. J. Med. Chem. 53: 5727–5737.

32. Grant BJ, McCammon JA, Gorfe AA (2010) Conformational Selection in G-Proteins Lessons from Ras and Rho. Biophys. J. 99: L87–L89.

33. Grant BJ, Gorfe AA, McCammon JA (2010) Large conformational changes in proteins: signaling and other functions. Curr. Opin. Struct. Biol. 20: 142–147.

34. Buhrman G, Holzapfel G, Fetics S, Mattos C (2010) Allosteric modulation of Ras positions Q61 for a direct role in catalysis. PNAS USA 107: 4931–4936.

35. Grant BJ, Lukman S, Hocker HJ, Sayyah J, Brown JH, et al. (2011) Novel Allosteric Sites on Ras for Lead Generation. PLoS One 6: e25711 1–10.

36. Perot S, Sperandio O, Miteva MA, Camproux AC, Villoutreix BO (2010) Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. Drug Discov. Today 15: 656–667.

37. Hamelberg D, Mongan J, McCammon JA (2004) Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. J. Chem. Phys. 120: 11919–11929.

38. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The Protein Data Bank. Nucleic Acids Rese. 28: 235–242.

39. Grant BJ, McCammon JA, Caves LSD, Cross RA (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.

40. Brenke R, Kozakov D, Chuang GY, Beglov D, Hall D, et al. (2009) Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. Bioinf. 25: 621–627.

41. Halgren T (2007) New method for fast and accurate binding-site identification and analysis. Chem. Bio. Drug Des. 69: 146–148.

42. Halgren TA (2009) Identifying and Characterizing Binding Sites and Assessing Druggability. J. Chem. Inf. Model. 49: 377–389.

43. Snyder JT, Worthylake DK, Rossman KL, Betts L, Pruitt WM, et al. (2002) Structural basis for the selective activation of Rho GTPases by Dbl exchange factors. Nat. Struct. Biol. 9: 468–475.

44. Formaneck MS, Ma L, Cui Q (2006) Reconciling the "old" and "new" views of protein allostery: A molecular simulation study of chemotaxis Y protein (CheY). Proteins: Struct., Funct., Bioinf. 63: 846–867.

45. Gsponer J, Christodoulou J, Cavalli A, Bui JM, Richter B, et al. (2008) A coupled equilibrium shift mechanism in calmodulin-mediated signal transduc-tion. Structure 16: 736–746.

46. McClendon CL, Friedland G, Mobley DL, Amirkhani H, Jacobson MP (2009) Quantifying Correlations Between Allosteric Sites in Thermodynamic Ensem-bles. J. Chem. Theor. Comput. 5: 2486–2502.

47. Rosnizeck IC, Graf T, Spoerner M, Trankle J, Filchtinski D, et al. (2010) Stabilizing a Weak Binding State for Effectors in the Human Ras Protein by Cyclen Complexes. Angew. Chem. Int. Ed. 49: 3830–3833.

48. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, et al. (2004) Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J. Med. Chem. 47: 1739–1749.

49. Post GR, Collins LR, Kennedy ED, Moskowitz SA, Aragay AM, et al. (1996) Coupling of the thrombin receptor to G12 may account for selective effects of thrombin on gene expression and DNA synthesis in 1321N1 astrocytoma cells. Mol. Biol. Cell 7: 1679–1690.

50. Majumdar M, Seasholtz TM, Buckmaster C, Toksoz D, Brown JH (1999) A Rho exchange factor mediates thrombin and Gα12-induced cytoskeletal responses. J. Biol. Chem. 274: 26815–26821.

51. Martin CB, Mahon GM, Klinger MB, Kay RJ, Symons M, et al. (2001) The thrombin receptor, PAR-1, causes transformation by activation of Rho-mediated signaling pathways. Oncogene 20: 1953–1963.

52. Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD (2006) Bio3d: an R package for the comparative analysis of protein structures. Bioinf. 22: 2695–2696.

53. Caves LSD, Nguyen DT, Hubbard RE (1991) Conformational variability of insulin: a molecular dynamics analysis. In: Goodfellow JM, editor. Molecular Dynamics: Applications in molecular biology. London: The Macmillan Press Ltd. 27–68.

54. vanAalten DMF, Conn DA, deGroot BL, Berendsen HJC, Findlay JBC, et al. (1997) Protein dynamics derived from clusters of crystal structures. Biophys. J. 73: 2891–2896.

55. Abseher R, Horstink L, Hilbers CW, Nilges M (1998) Essential spaces defined by NMR structure ensembles and molecular dynamics simulation show significant overlap. Proteins: Struct., Funct., Bioinf. 31: 370–382.

56. Caves LSD, Evanseck JD, Karplus M (1998) Locally accessible conformations of proteins: Multiple molecular dynamics simulations of crambin. Protein Sci 7: 649–666.

57. Elsawy KM, Hodgson MK, Caves LSD (2005) The physical determinants of the DNA conformational landscape: an analysis of the potential energy surface of single-strand dinucleotides in the conformational space of duplex DNA. Nucleic Acids Res. 33: 5749–5762.

58. Grant BJ, Gorfe AA, McCammon JA (2009) Ras Conformational Switching: Simulating Nucleotide-Dependent Conformational Transitions with Accelerated Molecular Dynamics. PLoS Comput. Biol. 5(3): e1000325.

59. Finn RD, Mistry J, Tate J, Coggill P, Heger A, et al. (2010) The Pfam protein families database. Nucleic Acids Res. 38: D211–D222.

60. Shannon CE (1948) A Mathematical Theory of Communication. Bell System Technical Journal 27: 379–423.

61. Shenkin PS, Erman B, Mastandrea LD (1991) Information-Theoretical Entropy as a Measure of Sequence Variability. Proteins: Struct., Funct., Gen. 11: 297–313.

62. Hubbard S, Thornton JM (1993) NACCESS, computer program. London: Department of Biochemistry and Molecular Biology. University College London.

63. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, et al. (2005) The Amber biomolecular simulation programs. J. Comput. Chem. 26: 1668–1688.

64. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, et al. (2006) Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins: Struct., Funct., Bioinf. 65: 712–725.

65. Hamelberg D, de Oliveira CAF, McCammon JA (2007) Sampling of slow diffusive conformational transitions with accelerated molecular dynamics. J. Chem. Phys. 127: 155102–155111.

66. Developmental Therapeutics Program NCI/NIH website. Available: http://dtp.nci.nih.gov/branches/dscb/div2_explanation.html. Accessed 2012 January 15.