Allostery through the computational microscope: cAMP activation of a canonical signaling domain

Robert D. Malmstrom[1,2], **Alexandr P. Kornev**[3], **Susan S. Taylor**[1,3], and **Rommie E. Amaro**[1,2,*]

[1]Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0340

[2]National Biomedical Computation Resource, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0340

[3]Department of Pharmacology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0340

Abstract

Ligand-induced protein allostery plays a central role in modulating cellular signaling pathways. Here, using the conserved cyclic-nucleotide binding domain of protein kinase A's (PKA) regulatory subunit as a prototype signaling unit, we combine long-timescale, all-atom molecular dynamics simulations with Markov state models to elucidate the conformational ensembles of PKA's cyclic-nucleotide binding domain A for the cAMP-free (apo) and cAMP-bound states. We find that both systems exhibit shallow free-energy landscapes that link functional states through multiple transition pathways. This observation suggests conformational selection as the general mechanism of allostery in this canonical signaling domain. Further, we expose the propagation of the allosteric signal through key structural motifs in the cyclic-nucleotide binding domain and explore the role of kinetics in its function. Our approach integrates disparate lines of experimental data into one cohesive framework to understand structure, dynamics, and function in complex biological systems.

Introduction

Since the introduction of the allosteric effect in *L-threonine deaminase*[1,2], researchers have sought to understand the mechanisms of allostery. The classical phenomenological models proposed by Monod-Wyman-Changeux[3] and (Pauling)-Koshland-Nemethy-Filmer[4,5] have been extended over time to emphasize population shifts or modulations in the conformational ensemble of the proteins[6,7]. Fundamentally, it is the character of the underlying free-energy landscape that determines the mechanism of allostery[6–11].

Corresponding Author: R.E.A (ramaro@ucsd.edu).

While experimental and computational approaches provide insight into allosteric mechanisms[12,13], a robust atomic-level description of the free-energy landscape remains a challenge due to a variety of methodological limitations. Without an atomic-level description of the conformational ensemble and its corresponding free-energy landscape, the details of allostery can remain hidden within ensemble averages or constrained perspectives of protein conformation and dynamics.

The objective of our work is to apply cutting-edge computational approaches to explore the conformational ensemble of a protein with and without a bound ligand, thereby gaining insight into how the ensemble gives rise to protein function. This work focuses on the mechanism of ligand-driven allostery in the cyclic-nucleotide binding domain of protein kinase A's (PKA) regulatory subunit. Despite extensive research conducted on this system[14], questions still remain.

The intercellular activation of PKA by cAMP is a prototypical example of ligand-protein allostery, which occurs through the cooperative binding of cAMP to tandem cyclic-nucleotide binding domains, designated A and B, in each of PKA's regulatory subunits[14]. Crystallographic data shows that the regulatory subunit, isoform Iα, undergoes significant conformational changes upon binding cAMP. It transitions from an extended holo-enzyme conformation that inhibits PKA's catalytic subunits to a compact cAMP-bound conformation that releases and activates PKA's catalytic subunits[15,16]. Within the cyclic-nucleotide binding domain, the inactive holo-enzyme and the active cAMP-bound conformations of the regulatory subunit are characterized by the conformational changes in three structural motifs: the N3A motif, the phosphate-binding cassette (PBC), and the B/C helix[17,18], leaving the core β-barrel sub-domain largely unchanged (Figure 1). These structural motifs represent the fundamental signal transduction component and form the binding interface between the cyclic-nucleotide binding domain and PKA's catalytic subunit.

The existing crystallographic data provides only two conformational states for the cyclic-nucleotide binding domains, the holo-enzyme (a.k.a. H or *inactive*)[16,19] and the cAMP-bound conformation (a.k.a. B or *active*)[15]. NMR dynamics data of the minimal functional component of PKA's regulatory subunit, the A cyclic-nucleotide binding domain (aa. 119–244), indicates the presence of two dominant conformational states[20–22] and suggests conformational selection as the mechanism of allostery[23]. Upon cAMP binding, changes in NMR chemical shift data implicate the phosphate-binding cassette as a major dynamic element, while few changes are observed in the N3A motif or the B/C helix.

However, such experiments are limited in their ability to elucidate details of the full structural ensemble. Instead, they provide averaged metrics that correlate to dynamic motions. Achieving an atomistic description of the underlying conformational ensemble and resolving the corresponding free-energy landscape will clarify the mechanisms by which cAMP modulates the function of the cyclic-nucleotide binding domain.

Because the PKA holo-enzyme contains four cyclic-nucleotide binding domains, cAMP-induced activation is a complicated process. A recent study by Boras et al.[24] provided a

general blueprint of the events during activation at the subcellular level, but the molecular mechanism of activation remains to be elucidated.

To build a foundation on which to address these questions, we simulated fully solvated atomistic models of the conformational ensembles of the A cyclic-nucleotide binding domain (hereafter referred to as the *cyclic-nucleotide binding domain* or *CBD*), with and without cAMP bound. Extensive molecular dynamics (MD) simulations were integrated using Markov state model theory to explore PKA's conformational space and long-timescale dynamics. Markov state models depict the interaction dynamics of discrete interconnected states as a transition probability matrix, at a fixed lag time, assuming that the transitions between states are independent of previous transitions (i.e., Markovian)[25–28]. By assigning individual frames extracted from MD trajectories to discrete conformational states, sampling from many separate trajectories can be integrated into one coherent framework that captures the kinetics and thermodynamics of the conformational ensemble at atomic resolution.

Similar approaches have been used to study protein folding[29], biased transition pathways between functional states[30,31], identification of cryptic allosteric sites[32], and ligand regulation of G-protein coupled receptors[33]. Our work is unique in that it directly assesses the atomic conformational landscape using initial unbiased long-timescale MD simulations augmented with adaptive sampling of one protein in two functional states, as opposed to only combining many short-timescale simulations or sampling along a predetermined reaction coordinate.

We find that the conformational ensembles of cAMP-free (CBD-Apo) and cAMP-bound (CBD-cAMP) functional states exist within a shallow free-energy landscape that allows access to both experimentally determined functional conformations, indicating conformational selection as the principal mechanism of allostery with the isolated cyclic-nucleotide binding domain. We find that the addition of cAMP modifies the principal motions of the cyclic-nucleotide binding domain, which correspond to transitions between active and inactive states. Furthermore, our approach exposes the propagation of the allosteric signal through the cyclic-nucleotide binding domain's signaling motifs.

Our results complement existing structural[15,16] and dynamical[20–23,34] experimental data, and extend our understanding of the mechanism of ligand-induced allostery within the cyclic-nucleotide binding domain. Further, they provide a general framework for understanding the function of the ancient and ubiquitous cyclic-nucleotide binding domain[35,36] and allosteric mechanisms more generally.

Results

Modeling the Conformational Landscape

To explore and map the conformational landscapes of the cyclic-nucleotide binding domain, we integrated multiple molecular dynamics (MD) simulations with Markov state models. Explicit solvent MD simulations of the cyclic-nucleotide binding domain were performed with and without cAMP-bound systems (CBD-cAMP and CBD-Apo, respectively).

The MD simulations of both systems sampled a totaled 74 μs (Supplementary Table 1). The simulations were composed of four long-timescale, ~13 μs simulations on the Anton supercomputer[37] (at the Pittsburgh Supercomputing Center) and multiple parallel short-timescale, 0.5–1 μs simulations using GPU-accelerated AMBER12[38–40]. For both systems, sampling simulations were initiated from experimentally determined atomic coordinates of the cyclic-nucleotide binding domain[15,16] in both the inactive and active conformations. Initial sampling was followed by multiple rounds of directed sampling (10–15 ns simulations using GPU-accelerated AMBER12[38–40]) to refine the Markov state models[41]. Directed sampling simulations were initiated from intermediate conformations selected near poorly sampled transitions within the Markov state model. For both the CBD-cAMP and CBD-Apo systems, no single simulation sampled a complete transition between the experimentally determined inactive and active structures. However, simulations started from either experimental conformation overlapped in conformational space, based on alpha carbon root mean squared distance (RMSD), indicating that the MD trajectories could be integrated meaningfully into a single Markov state model (Supplementary Figure 1).

To build the Markov state models, we needed to define the conformational space, and then divide the space into discrete conformational states. As in protein-folding models[29], the cyclic-nucleotide binding domain conformations are defined by the atomic coordinates of each residue's alpha carbon. We divided the sampled conformational space into discrete conformational states through RMSD clustering using MSMBuilder2[42]. We pre-aligned structures to a common reference frame to calculate RMSD to capture important translational motions lost using the standard minimum RMSD approach, which aligns each pair of compared structures to each other before determining RMSD. Inspection of the individual MD trajectories indicated that the β-barrel of the cyclic-nucleotide binding domain was stable relative to the motions of the N3A motif, the phosphate-binding cassette (PBC), and the B/C helix (Supplementary Figure 2). Therefore, we used the β-barrel as our common reference frame.

All conformations sampled from the MD simulations of both the CBD-Apo and CBD-cAMP systems were clustered together to create a comprehensive framework of conformational states. Once the conformation states were identified, Markov state models were built for each system. This means that the models were built using the same division of conformational space, but each model included only states sampled in its respective simulations. (A comparison between this approach and clustering conformations from CBD-Apo and CBD-cAMP separately is discussed below.)

The Markov state models were built by specifying the maximum distance or *cut-off* for the clustering algorithm, which determined how the conformational space was divided, and a lag time for the model. Model parameters were selected because they were Markovian, as indicated by implied timescale plots, and could maximize the number of conformational states. Additionally, consistency between the Markov state models and the MD simulations were determined by the Chapman-Kolmogorov test[26] (i.e., were the models' generated trajectories similar to the results from the original MD simulations?) (Supplementary Figures 3–10). An evaluation of several RMSD cut-offs and lag times indicated that the best

Markov state model utilized a clustering cut-off of 3.0 Å and a lag time of 9.6 ns (Supplementary Figures 3–4).

Graphs of the Markov state models for CBD-Apo and CBD-cAMP are visualized as a network (Figure 2). Each node corresponds to a cluster of similar conformations or a single conformational state. The diameters of the nodes are proportional to the log of equilibrium population of the conformational state. Thus, smaller nodes indicate conformations that are more rarely sampled. The location of each node is determined by the RMSD of the structure of the cluster's generator, approximately the centroid of the cluster, to the inactive and active crystallographic structures. We note that the location of each node gives a general indication of the conformation of each state relative to the experimental structures, but node locations are a projection from a higher-dimensional space. (Separation of the experimentally determined active node, indicated in Figure 2, from the bulk of the other nodes is an artifact of the clustering algorithm in MSMBuilder2[42], which selects the first molecular dynamics frame [i.e., the equilibrated starting active conformation] as a generator for the first cluster. Notice that the separation of the active conformation node from its neighbors is approximately 3 Å or the RMSD cut-off of the clustering algorithm.)

Ligand-meditated Changes in the Conformational Ensemble

A comparison of the models of the conformational state ensembles for the CBD-Apo and CBD-cAMP systems allowed us to assess how the binding of cAMP modifies the overall free-energy landscape. We compared the distribution of conformational states between the two systems, the overall character of the free-energy landscape, and the kinetics of traversing the free-energy landscape between functional conformations states. For this comparison, and throughout this work, we assumed that the crystallographic structures represent the functional conformations (active or inactive states) of the cyclic-nucleotide binding domain. The nodes containing the functional conformations are indicated in Figure 2.

When comparing the conformational landscapes of CBD-Apo and CBD-cAMP, we noticed two striking features: the large number of shared states (Figure 2) and the character of the unique states. The models indicate that the systems share ~70% of the total number of sampled conformational states (Supplementary Table 2). Notably, the shared conformational states of CBD-Apo and CBD-cAMP comprise 66% and 38%, respectively, of the total population for each ensemble.

Many of the unique CBD-Apo conformational states contain structures in which the cAMP-binding site is formed poorly and, thus, sterically prevented from occurring when cAMP is bound (Supplementary Figure 11). Many of the most populated unique CBD-cAMP conformational states exhibit structures in which the C-terminus is interacting with cAMP (Supplementary Figure 12). These states include the dynamical capping of cAMP's adenosine ring by Y244, part of the C-helix, that substitutes the native capping residue W260 from the "B" CBD in the full regulatory subunit[15,43]. These capping interactions with Y244 have been observed experimentally in the truncated RIα (aa. 1–259)[43] but not previously observed in simulations. Additionally, fluorescent experiments studying the dynamics of R239 (located on the C-helix) indicate that the C-helix has a higher propensity

to interact with the core cyclic-nucleotide binding domain in the presence of cAMP[34]. The unique CBD-cAMP conformational states are not sterically excluded from the CBD-Apo conformation ensemble, and interactions between the C-terminus and the cAMP-binding sites were observed in the CBD-Apo ensemble. Therefore, the absence of unique states in the CBD-cAMP ensemble is likely due to the absence of stabilizing interactions between cAMP and the C-terminus. Importantly, both models sampled the active and inactive states, yet neither state is the most populated conformational state in either system.

Using the most probable conformational state as the reference state, at physiological conditions the depth of free-energy landscape is 6.6 kcal mol^{-1} for CBD-Apo and 7.6 kcal mol^{-1} for CBD-cAMP. The free energies for the individual conformational states follow a Gaussian distribution with a mean depth of the free-energy landscape at 3.1 kcal mol^{-1} +/− 1.4 kcal mol^{-1} for CBD-Apo and 4.2 kcal mol^{-1} +/− 1.6 kcal mol^{-1} for CBD-cAMP. The transition free energy out of a conformational state to any of its neighboring conformational states (i.e., only states that it is connected to) ranges from 5.8 to 8.6 kcals mol^{-1} for both systems. (5.8 kcal mol^{-1} corresponds to a transition made at approximately the lag time of the Markov state model and, thus, is the lowest bound transition free energy calculated from the Markov state model.) The absolute difference in free energy between the active and inactive states is 0.1 kcal mol^{-1} for CBD-Apo and 0.9 kcal mol^{-1} for CBD-cAMP. Overall, the free-energy landscapes of both systems provide access to both functional states at physiological conditions, with cAMP binding deepening the free-energy landscape by ~1 kcal mol^{-1}.

Transition pathway theory[44,45] identified the 10 highest flux paths between the functional conformational states for each ensemble (Supplementary Figure 13). The proportion of the highest flux path relative to the total flux for each functional state ensemble was striking: For CBD-Apo, the highest flux path was only 9% of the total flux, and for CBD-cAMP, it was only 4% of the total flux (Supplementary Table 3). The lack of a dominant path and the variety of paths identified in the analysis indicate that there are multiple pathways between functional conformational states. Inspection of the pathways indicated no common order of conformational changes or shared features at the bottlenecks of the pathways.

Therefore, to understand the transition between active and inactive conformation states across the free-energy landscape, we needed to look at the transition pathways collectively. We measured the kinetics of the transition between active and inactive states using mean first passage times (i.e., the fastest average time it takes for the system to move across the free-energy landscape from active to inactive state, or *vice versa*) (Table 1). Interestingly, the mean first passage times for the transition from inactive to active state are similar with or without cAMP bound: ~30 μs. However, the addition of cAMP slows the transition from active to inactive state by a factor of five, from 16 μs to 84 μs.

These results indicate that CBD-Apo favors the inactive state, which inhibits PKA's catalytic subunit, whereas CBD-cAMP favors the state that activates PKA's catalytic subunit. Thus, the dynamics of the cyclic-nucleotide binding domain derived from the Markov state models are consistent with the function of the regulatory subunit in PKA.

Taken together, our results indicate conformational selection as the governing mechanism of allostery in the cyclic-nucleotide binding domain. The weakest evidence for this mechanism is the number of shared states between both models because it is largely a function of the clustering algorithm. However, when the conformations for each system are clustered separately, the results are the same. In fact, the joint clustering approach better captures the principal motions of the CBD at a 3.0 Å clustering cut-off, as determined by implied timescale plots (Supplementary Figure 14).

The stronger evidence supporting conformational selection is the change in the free-energy landscape and the corresponding transition kinetics. Weinkam and coworkers[8] classified allosteric mechanisms based on the difference between the change in free energy of two functional states and the binding free energy of the ligand. Although the current Markov state models do not allow determination of the free energy of cAMP binding, and there has been no direct experimental measurements of cAMP binding in our construct, we can estimate the binding free energy of cAMP through previous work on the isolated A cyclic-nucleotide binding domain[46]. The free energy of binding of cAMP to the isolated A cyclic-nucleotide binding domain is approximately -3 kcal mol^{-1} [46]. This value is greater than the difference in free energy between the active and inactive states, ~-1 kcal mol^{-1}, supporting conformational selection as the general mechanism of allostery. Additionally, the on-rate for cAMP to the A cyclic-nucleotide binding domain of R, $\sim 4.8 \times 10^4$ s^{-1} [46], is faster than the transition times between active and inactive states in the CBD-cAMP system. And the fact that the transition rates between inactive and active conformations are independent of ligand binding indicate conformational selection as described by Zhou[9]. Experimental results[23,47] support conformational selection as the mechanism of ligand-induced allostery, thus complementing the interpretations of our Markov state model presented here.

One limitation of our models is the potential bias towards conformational selection because the simulations were initiated at both experimentally determined active and inactive states, which were included in and connected through the Markov state models. However, nothing sterically hinders cAMP from binding either functional conformational state in the absence of the catalytic subunit, and no restraints were required to keep cAMP bound during the simulations. Thus, all four systems used to initiate the simulations are potential members of the conformational ensemble. Starting the simulations in each experimental conformation does not predispose them to becoming members of highly populated or networked states or bias the transition kinetics between active and inactive states.

Conformational Space, Principal Motions, and Function

While our results indicate conformational selection as the general mechanism of allostery for the CBD, this finding is not wholly novel due to previous experimental work[23,47]. However, our work provides a foundation to delve further into the mechanisms of allostery at an atomic scale in ways inaccessible to current experimental methods: We can explore structural characteristics of the equilibrium conformational space, and the principal motions of the CBD and its functional macrostates.

To understand the structural characteristics of the conformational landscapes, we built stationary probability density volumes of the positions of the alpha carbons. In other words,

we sought to determine the probability of finding the mass of the alpha carbons in a given volume of space at equilibrium. The volumes were determined by dividing the space in 0.1 Å voxels and calculating the probability of finding an alpha carbon within 3 Å (the RMSD cut-off for the Markov state model) of the voxel for each conformational state. By summing over all states, we obtained a map of the stationary probability density, which provides a visual representation of the conformational ensembles.

Figure 3 shows the probability of density surfaces at 66%, 95%, and 99.9%. The volume enclosed by each surface has a probability greater than the cut-off of being occupied by an alpha carbon. This figure shows the wide range of motion sampled by the α-helical structural subdomain at equilibrium; this range of motion is similar for the CBD-Apo and CBD-cAMP systems. Similar to the graphical representation of the Markov state model, the stationary probability density volumes indicate the similarity of the conformations sampled by CBD-Apo and CBD-cAMP.

However, the differences between the stationary probability density volumes give us new structural insight into the mechanisms of allostery within the CBD. The first notable difference is the unique density formed by the B/C helix proximal to cAMP in the CBD-cAMP ensemble (Figure 4A). This density corresponds to the interactions between the C-terminus and the adenosine ring of cAMP discussed above. A second difference is the higher probability that the N3A motif extends over the B/C helix in the CBD-Apo ensemble (Figure 4B). This density corresponds with the experimentally determined binding interface between PKA's catalytic and regulatory subunits.[16] Interestingly, the CBD-Apo system only increased density by about 10% in forming the interface. This small variation shows how seemingly small changes in a conformation ensemble give rise to inhibitory function of the CBD-Apo system by increasing the probability of formation of the binding interface between PKA's catalytic and regulatory subunits.

To determine the principal motions of the CBD, we built kinetic coarse-grained models from the Markov state models of each system to identify metastable states or the principal motions of CBD arising from the conformational ensemble. These models were generated by complementarily employing principal component and Markov clustering analysis[25,28,32,42,48,49].

The CBD-Apo implied timescale plots (Supplementary Figure 3) suggest at least one dominant slow motion followed by numerous secondary faster motions. The dominant slow motion of the CBD-Apo ensemble is between conformations similar to the active and conformations similar to the inactive conformational states (Figure 5). Of note, the second most dominant motion is integration of the unfolded N-helix (the N3A motif) into the core β-barrel (Supplementary Figure 15). These conformations are only observed in the CBD-Apo ensemble as they deform the cAMP-binding site, sterically excluding cAMP. Integration of the unfolded N-helix requires partial unfolding of the core β–barrel, which was observed in previous NMR work[21]. We recognize that these conformations may be an artifact of the truncated CBD and, therefore, may not be part of the mechanisms of allostery for the regulatory subunit *in vivo*. Still, the consistency between the results of the

simulations and the NMR work indicates that the models are able to replicate structural events seen experimentally.

Like the CBD-Apo system, the implied timescale plot for the CBD-cAMP system indicates one dominant slow motion followed by a number of faster motions (Supplementary Figure 4). Although the dominant motion is structurally similar to that of CBD-Apo (i.e., it is also the transition between active and inactive conformations [Figure 5]), it is slower than that of CBD-Apo (compare Supplementary Figures 3 and 4). Interestingly, the second slowest motion within the CBD-cAMP system corresponds to those states where the C-terminus is interacting with cAMP.

To strengthen the relationship between motions of the CBD and its function, we generated a functional coarse-grained model of the conformational ensembles. The model identifies which conformational stats have a higher probability of transitioning to either the active of inactive state, thus indicating the role of each conformational state in the function of the system. Using committer analysis from transition pathway theory[44,45], the ensembles were divided into two functional macrostates. Each macrostate was composed of conformational states with a >50% probability of transitioning to the active or inactive state. In other words, the functional coarse-grain model can be conceptualized as the "continental divide" of the free-energy landscape, indicating to which functional state (active or inactive) an individual conformational state is connected to most strongly.

The functional coarse-grain model of CBD-Apo (Figure 4) contains similar populations for the active and inactive macrostates, with a preference for the inactive macrostate. CBD-cAMP (Figure 4), however, prefers the active macrostate (Table 2). These results agree with NMR studies that show similar inactive and active conformational populations at very-low-cAMP concentration and domination of the active macrostate at high-cAMP concentrations[21]. Additionally, the relationship of the half-life of either functional macrostate is similar to the mean first passage time between functional conformational states in that transitions from inactive to active states have similar kinetics in both systems, whereas the presence of cAMP slows the transition from active to inactive state (Table 2).

The functional coarse-grain model and the kinetic coarse-grain model closely correlate with each other as determined by conformational state membership in the functional or kinetic macrostates. Correlation between the two coarse-grained models is expected as they both depend on the transition kinetics of the Markov state models; however, which motion corresponded to the change in functional states is not clear *a priori*. This correlation indicates that the functional macrostates determined by transition pathway theory correspond to the slowest motion of the cyclic-nucleotide binding domain (Figure 4). It also supports the assumption that the crystallographic conformations represent the active and inactive states of the cyclic-nucleotide binding domain and that its slowest motion is the transition between active and inactive states.

The coarse-grained models combined with observations of the change in stationary probability densities for the two systems support a general explanation that the modification of the transition kinetics by cAMP gives rise to function. In CBD-Apo, transitions between

states occur with a tendency towards the inactive conformation, which increases the likelihood of generating the interface that allows for the regulatory subunit to bind the catalytic subunit leading to inactivation of PKA. With cAMP bound, the transition between inactive and active states is slower, which is likely the result of interactions formed between the C-terminus and cAMP. The preference for the active-like conformations in the CBD-cAMP system decreases the probability of forming the binding interface and allowing activation of PKA.

Motions of the Signaling Motifs

To gather more detail on the allosteric mechanism, we examined the propagation of an allosteric signal through the helical signaling motifs of the cyclic-nucleotide binding domain by building new Markov state models for the three principal signaling motifs using the same MD trajectories (apo and cAMP-bound). While the motions of the signaling motifs are captured in the Markov state models of the full CBD, as discussed above, individual models elucidate the conformational ensembles of focused structural elements and provide an unobstructed view of their modulation by cAMP. The Markov state models for the individual motifs were developed and analyzed using the same methods for the full system (Supplementary Figures 5–10). They indicate conformational selection as the general mechanism of allostery, due to well-connected active and inactive conformational states (Figure 7) with multiple pathways across the free-energy landscape for each of the three key motifs (Supplementary Table 3).

When we compared the conformational state distribution between the apo and cAMP-bound systems, we observed that the N3A motif and the B/C helix share a majority of their conformational states (Figure 7 and Supplementary Table 2), similar to the full cyclic-nucleotide binding domain. The B/C helix generally explores the conformational space between the experimentally determined active and inactive structures (Figure 7A). However, the N3A motif explores conformations dissimilar to either experimental conformation (Figure 7C). These models are consistent with the observed motions of these domains (Supplementary Figure 2), particularly with the unfolding of the N-helix (Supplementary Figure 11).

These results support and help explain why there is no observed change in chemical shifts in NMR experiments[21] for the N3A motif or the B/C helix despite the changes in the crystallographic structures (Figure 1). Simply put, in solution the conformational space explored by the N3A motif and the B/C helix is largely independent of cAMP, though cAMP does modulate the relative populations of the conformational states and their kinetics.

In contrast, the phosphate-binding cassette showed significant changes in chemical shifts upon binding cAMP.[21] While these changes appear to be driven by interaction with cAMP, the Markov state model of the phosphate-binding cassette illuminates another aspect of these changes: that, upon binding of cAMP, the phosphate-binding cassette loses about half of its conformational states, representing a significant loss in conformational entropy (Figure 7B, Supplementary Table 2). Additionally, the kinetic and functional coarse-grain models of the phosphate-binding cassette show significant changes between apo and cAMP systems (Supplementary Figure 17 and Supplementary Table 5) unlike the other signaling motifs

(Supplementary Figures 16, 18 and Supplementary Tables 4, 6). Without cAMP, the phosphate-binding cassette has a shallow free-energy landscape with quick transition between metastable macrostates (Table 1, Supplementary Figure 17, and Supplementary Table 5). With cAMP bound, there is one dominant (slow) transition between two metastable macrostates (Table 1, Supplementary Figure 17, and Supplementary Table 6; also compare implied timescale plots in Supplementary Figures 7 and 8). Unlike the other models (Figure 5 and Supplementary Figures 16 and 18), the conformational states that comprise these metastable macrostates do not correspond to the division of microstates seen in the functional coarse-grain model (Supplementary Figure 17). Instead the functional coarse-grain model indicates a strong propensity of the phosphate-binding cassette to assume a cAMP-binding conformation or active conformation even in the absence of cAMP (Supplementary Figure 17 and Supplementary Table 5). Overall, the model shows that, without cAMP bound, the phosphate-binding cassette explores a variety of conformations, but the binding site has a tendency to assume the active cAMP-bound conformation. cAMP binding selects the active conformation while allowing transition between active and inactive state but significantly slows the transition between these states (Table 1, Supplementary Table 16).

The spatial arrangement of the helical signaling motifs suggests that the propagation of the allosteric signal would be from the phosphate-binding cassette to the B/C helix to the N3A motif (Figure 1). Even the natures of the conformation ensembles, described by the Markov state models, suggest propagation of the allosteric signal from a switch like the phosphate-binding cassette out to the conformationally diverse N3A motif.

However, the kinetics of the signaling motifs suggest another mechanism. The mean first passage times between active and inactive states suggest that the movement of the B/C helix is the rate-limiting step for the functional transition because it has the slowest transitions between active and inactive states (Table 1). Like the full CBD, the B/C helix has similar transition times between active and inactive states whether or not cAMP is bound. However, cAMP significantly decreases the rate of the active-to-inactive transition (Table 1). This rate reduction is due to stabilization of the active conformation by cAMP through interactions between the C helix and cAMP (Figure 4), as observed and discussed above.

In PKA and other proteins, the experimental free energy of binding of cAMP to the cyclic-nucleotide binding domain is dominated by enthalpic contributions[50–52], suggesting that the contacts between cAMP and the phosphate-binding cassette and B/H helix are drivers for this phenomenon. Our model indicates that the binding of cAMP near the phosphate-binding cassette dampens the dynamics of that signaling motif, greatly increasing its propensity to adopt an active-like conformation, and that the motions of the B/C helix are the rate-limiting step in this transition.

Discussion

Overall, we find that the prototype cyclic-nucleotide binding domain is a highly dynamic system and that cAMP modulates its function by modifying the dynamics of the signaling motifs. By comparing CBD-Apo and CBD-cAMP ensembles, we see that the free-energy

landscape allows access to both active and inactive states through an ensemble of transition pathways (i.e., there is no one dominant pathway between states). Collectively, our analysis supports conformational selection as the general mechanism of allostery in this system. The addition of cAMP modulates the transition rates between the functional states but only between active-to-inactive states and not *vice versa*. This modulation of rates occurs by slowing motion in the CBD-cAMP ensemble, which is caused by interactions with cAMP. Finally, our work indicates that the change in dynamics of the B/C helix is the rate-limiting step in adopting a fully active conformation. Furthermore, the motion of the B/C helix is essential for signal propagation to the N3A motif and formation of the interface with the catalytic subunit. This mechanism could be tested experimentally though NMR by observing shifts in the relative populations of active and inactive ensembles caused by mutations in the C-terminus, specifically, mutations of residue 244, which should modify interactions with cAMP.

While the general mechanism of allostery for the cyclic-nucleotide binding domain is conformational selection, some elements of the mechanism appear to behave more like induced fit. Particularly notable in this regard are the interactions between the C-helix and cAMP, which generate multiple conformational states that are not sampled in the apo state, as compared to the motions of the individual signaling motifs that appear more similar to entropic mechanisms of allostery. Furthermore, in the context of activation of the full kinase, these mechanisms may change, as interaction with the catalytic subunit and the remaining portions of the regulatory subunit will almost certainly modulate the conformational ensemble of the cyclic-nucleotide binding domain. This work suggests that a clearer picture of the free-energy landscape, achieved through an integrated experimental and computational approach, is required to fully understand allosteric behavior.

Although a number of our conclusions were already observed experimentally, the power of our computational Markov-state-model-based approach is that it unifies multiple lines of existing experimental data into a single, cohesive framework and, thereby, achieves a more complete understanding of allostery. This approach builds on an emerging paradigm for understanding protein function and dynamics based on computational exploration and visualization of protein conformational ensembles and their underlying functional free-energy landscapes in atomic detail.

Methods

System Preparation

Parameterization of cAMP—We parameterized cAMP for the MD simulations as a small organic molecule using the Generalized Amber Force Field (GAFF)[53] to allow for consistent treatment of ligands in future studies. The coordinates for cAMP were taken from the 1RGS[15] crystal structure. Partial atomic charges for cAMP were determined by single-point energy calculations using Schrodinger's QM module Jaguar (Suite 2012: Jaguar, version 7.9, Schrödinger, LLC, New York, NY, 2012) using Hartree-Fock level of theory and 6–311g** bases. GAFF atom types were assigned using antechamber[54], and the parameters files were prepared with AMBER's xleap.

MD System Preparation—Using Schrodinger's Maestro's (Suite 2012: Maestro, version 9.3, Schrödinger, LLC, New York, NY, 2012) PDB prep and the Desmond System Preparer, we prepared four systems for MD simulations, CBD-Apo and CBD-cAMP in both the active and inactive conformations. The heavy-atom atomic coordinates for the active and inactive conformations of CBD-Apo were taken for crystal structures 1RGS[15] and 2QCS[16] resides 119 to 244, respectively. We added cAMP to the inactive conformation by aligning the cAMP-binding site of 1RGS and 2QCS and "cutting and pasting" the coordinates. For the CBD-Apo active conformation, the coordinates for cAMP were deleted from 1RGS. Within PBD prep, each system was capped to remove terminal charges, the systems were protonated at pH 7.0 with the pK_a titratable residues determined with the Maestro-integrated PROPKA, and, due to the low resolution of 1RGS, all crystallographic waters were removed. Using the Desmond System Preparer, we solvated each system in a cubic water box using TIP3 waters[55], counter ions, and 120 mM NaCl. No restraints were placed on cAMP.

Molecular Dynamics Simulations

System Parameterization—System coordinates from Maestro were converted into an xleap-readable format using in-house Python scripts. Each system was parameterized in xleap using the Amber99SB[56] force field, and periodic boundary conditions were implemented.

GPU-enabled Molecular Dynamics Simulations—All systems were minimized and equilibrated using the GPU version of Amber12[38–40]. We minimized the system in four stages: 1) 500 steps of hydrogen-only minimization, 2) 500 steps of solvent minimization, 3) 500 steps with only the backbone constrained, and 4) 5,000 steps of full minimization. We equilibrated the system using harmonic equilibration at 310K over four sequential 500 ps runs, decreasing the restraint potential on the backbone each step. GPU-enabled AMBER12 production runs were carried out as an NTP ensemble at 310K and 1 bar with a 2 fs time-step and partial mesh Ewald electrostatic approximation. MD input files are provided on dryad.org[57].

Anton Molecular Dynamics Simulations—MD simulations on Anton[37] were performed on the same parameterized, minimized, and equilibrated systems as GPU-enabled AMBER12 simulations. The Anton simulation was run in the NTP ensemble, using Anton's Berendsen thermostat-barostat, at 310K and 1 bar with a 2 fs time-step and partial mesh Ewald electrostatic approximation.

Initial and Adaptive Sampling—We sampled the conformational ensemble from the equilibrated conformation of each system, starting each run with a different set of initial velocities. Supplementary Table 1 contains the total sampling time for each system used in this analysis.

In conjunction with model building, we performed multiple adaptive sampling runs. These runs were initiated from preliminary Markov state model conformational states with <3 members. These conformational states corresponded to clustered regions of the

conformational ensemble poorly sampled by the MD simulations and, therefore, the transitions to and from those states were poorly sampled. For each under-sampled conformational state, one conformation was selected. From that conformation, three new 10–15 ns MD simulations were run, each with a new initial velocity using GPU-enabled AMBER12. The short MD simulations were sufficient to explore the local conformational landscape and return to a local energy minimum.

Building Markov State Models

Trajectory Preparation—MD trajectories were processed using CPPTRAJ[58] and VMD[59]. All frames were aligned using Cα of residues 153 to 199 and 211 to 223, the stable β-barrel, to crystallographic active conformation. The alignment provided a common reference frame to compare all conformations while maximizing the difference between conformations. The frame rate for the trajectories was unified to 120 ps due to different conformational sampling rates between the GPU-enabled AMBER12 (0.5 ps) and Anton (120 ps) MD simulations. The trajectories were converted into NAMD's[60] ".dcd" trajectory format for analysis with MSMBuilder2.

Building the Markov State Model—We used MSMBuilder2[32,42], release 2.51, to preform cluster analysis, Markov state model building, and Markov state model selection. To have a common set of microstates for both the CBD-Apo and CBD-cAMP systems, all trajectories were clustered together using the hybrid k-centers k-medoids clustering algorithm using the custom distance metric option to calculate RMSD without first performing an alignment on the atomic coordinates (an option not available in MSMBuilder2). MSMBuilder2 allows for refinement of clustering through refinement of frame membership within a cluster as well as global refinement of cluster generators. For our systems, we found that any refinement only made the space discretization worse, as determined by implied timescale plots. Therefore, we didn't use any refinement options for our clustering. Markov state models were then built on the CBD-Apo and CBD-cAMP trajectories separately using cluster generators to assign state membership to each frame of the trajectories. When the best conformational landscape partition and lag time were determined (see discussion below), a final Markov state model was built using the maximal likelihood estimator option in MSMBuilder2.

Model Selection—For the Markov state model of the whole systems, we selected a clustering cut-off distance of 3.0 Å and a lag time of 80 steps or 9.6 ns. For consistency in comparing models, a lag time of 9.6 ns was used for all.

Markov State Model Visualization—The Markov state models were visualized using in-house Python scripts leveraging the publically available NetworkX module[61].

Markov State Model Analysis

Equilibrium Population—The equilibrium population was determined from the first eigenvalue of the Markov state model's transition probability matrix that corresponds to the equilibrium distribution of the conformational states.

Transition Pathway Theory Analysis and Transition Pathway Theory Clustering—Transition pathway theory analysis was preformed using MSMBuilder2 transition pathway theory scripts[42]. Total flux was calculated as described in Prinz et al.[45]. Transition-pathway-theory-based clustering was preformed using in-house Python scripts leveraging MSMBuilder2 modules. The script assigned transition pathway theory macrostates to each microstate of the original Markov state model based on the probability of transitioning to either the inactive or the active conformation (cut-off 50%). The conformational state assignments for each frame of the trajectories were reassigned using in-house Python scripts and a new Markov state model based on the macrostate assignments was built with MSMBuilder2.

Markov Cluster Analysis Macrostate Models—Markov cluster analysis clustered the states of the Markov state model based on their kinetic relationships within the Markov state model using a random walk algorithm to determine local sinks in the Markov state model[48]. We performed the Markov cluster analysis using in-house Python scripts that assigned new Markov cluster analysis state assignments to the original state assignments and built a Markov state model. Additionally, we used MSMBuilders2's PCCA+[42] module to validate the division identified through Markov clustering.

Mean First Passage Time—Absolute MFPT between the putative functional microstates was calculated as described in Singhal et al.[62].

Half-life—The half-life of coarse-grained microstates was calculated using a kinetic Monte Carlo scheme on the Markov state model's transitions matrix as described in Shukla et al.[63]. Trajectories were generated by randomly selecting a starting state within a macrostate, then moving between states based on the transition's probability as determined by the Markov state model. The passage time of the trajectory was the number of steps in the trajectory multiplied by the lag time (9.6 ns). The half-life was calculated by taking the average of 10 median passage times from a collection of 1,000 trajectories.

Calculating Kinetics and Free Energies of the Conformational Landscape—
Assuming that transitioning out of one conformational state into its neighboring conformational states can be modeled as parallel reactions, we can treat the kinetics for the conformational changes between connected conformational states as first-order reactions. Therefore, we determined the rate constant using

$$k = - \frac{\ln(p_1/p_0)}{\tau}$$

where k in the first-order rate constant, p_0 is the probability of starting in a given state (i.e., the starting concentration or 1), p_1 is the probability of still being in that state at a given lag time (i.e., the concentration after a give time), and τ is the lag time (9.6 ns). The free energy of transition was calculated using Eyring's equation as follows:

$$\Delta G^{++} = -RTln\left(\frac{h}{\kappa k_B T}k\right)$$

where R is the gas constant, h is Plank's constant, k_B is Boltzmann's gas constant, κ is the transmission coefficient assumed to be 1, T is temperature (310 K), and k is the rate constant calculated above.

The free energy of a single microstate was calculated using

$$\Delta G = -RTln\left(\frac{p_i}{p_r}\right)$$

where p_i is the probability of a given microstate, p_r is the probability of the reference microstate (the most probable microstate in the ensemble), R is the gas constant (0.0012 kcal mol^{-1} K^{-1}), and T is temperature (310 K).[64]

Markov State Model Signaling Motifs

The Markov state models of the signaling motifs were built and analyzed using the same methods described above except we only used a subset of Cα. The residues used for the analysis of each subdomain were: 119 to 151 for the N3A motif, 199 to 211 for the phosphate-binding cassette, and 226 to 244 for the B/C helix. Markov state models were built with a clustering cut-off of 4.5 Å for the N3A motif, 1.5 Å for the phosphate-binding cassette, and 4.5 Å for the B/C helix. The lag time used was 9.6 ns.

Data Sharing

MD trajectories, sample MD input scripts, and the Markov state model analysis scripts are available for download through datadryad.org[57].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Changeux J-P. The feedback control mechanisms of biosynthetic L-threonine disminase by L-isoleucine. Cold Spring Harb. Symp. Quant. Biol. 1961; 26:313–318. [PubMed: 13878122]

2. Monod J, Jacob F. Teleonomic mechanisms in cellular metabolism, gowth, and differentiation. Cold Spring Harb. Symp. Quant. Biol. 1961; 26:389–401. [PubMed: 14475415]

3. Monod J, Wyman J, Changeux J. On the Nature of Allosteric Transitions? J. Mol. Biol. 1965; 12:88–118. [PubMed: 14343300]

4. Koshland DE, Némethy G, Filmer D. Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry. 1966; 5:365–385. [PubMed: 5938952]

5. Pauling L. The oxygen equilibrium of hemoglobin and its structural interpretation. Proc. Natl. Acad. Sci. U. S. A. 1935; 21:186–191. [PubMed: 16587956]

6. Cui Q, Karplus M. Allostery and cooperativity revisited. Protein Sci. 2008; 17:1295–1307. [PubMed: 18560010]

7. Motlagh HN, Wrabl JO, Li J, Hilser VJ. The ensemble nature of allostery. Nature. 2014; 508:331–339. [PubMed: 24740064]

8. Weinkam P, Pons J, Sali A. Structure-based model of allostery predicts coupling between distant sites. Proc. Natl. Acad. Sci. U. S. A. 2012; 109:4875–4880. [PubMed: 22403063]

9. Zhou H-X. From induced fit to conformational selection: a continuum of binding mechanism controlled by the timescale of conformational transitions. Biophys. J. 2010; 98:L15–L17. [PubMed: 20303846]

10. Tsai CJ, Kumar S, Ma B, Nussinov R. Folding funnels, binding funnels, and protein function. Protein Sci. 1999; 8:1181–1190. [PubMed: 10386868]

11. Boehr DD, Nussinov R, Wright PE. The role of dynamic conformational ensembles in biomolecular recognition. Nat. Chem. Biol. 2009; 5:789–796. [PubMed: 19841628]

12. Feher VA, Durrant JD, Van Wart AT, Amaro RE. Computational approaches to mapping allosteric pathways. Curr. Opin. Struct. Biol. 2014; 25:98–103. [PubMed: 24667124]

13. Collier G, Ortiz V. Emerging computational approaches for the study of protein allostery. Arch. Biochem. Biophys. 2013; 538:6–15. [PubMed: 23933229]

14. Taylor SS, Ilouz R, Zhang P, Kornev AP. Assembly of allosteric macromolecular switches: lessons from PKA. Nat. Rev. Mol. Cell Biol. 2012; 13:646–658. [PubMed: 22992589]

15. Su Y, et al. Regulatory subunit of protein kinase A: structure of deletion mutant with cAMP binding domains. Science. 1995; 269:807–813. [PubMed: 7638597]

16. Kim C, Cheng CY, Saldanha SA, Taylor SS. PKA-I holoenzyme structure reveals a mechanism for cAMP-dependent activation. Cell. 2007; 130:1032–1043. [PubMed: 17889648]

17. Sjoberg TJ, Kornev AP, Taylor SS. Dissecting the cAMP-inducible allosteric switch in protein kinase A RIalpha. Protein Sci. 2010; 19:1213–1221. [PubMed: 20512974]

18. Kornev AP, Taylor SS, Ten Eyck LF. A generalized allosteric mechanism for cis-regulated cyclic nucleotide binding domains. PLoS Comput. Biol. 2008; 4:e1000056. [PubMed: 18404204]

19. Kim C, Xuong N-HH, Taylor SS. Crystal structure of a complex between the catalytic and regulatory (RIalpha) subunits of PKA. Science. 2005; 307:690–696. [PubMed: 15692043]

20. Das R, Melacini G. A model for agonism and antagonism in an ancient and ubiquitous cAMP-binding domain. J. Biol. Chem. 2007; 282:581–593. [PubMed: 17074757]

21. Das R, et al. cAMP activation of PKA defines an ancient signaling mechanism. Proc. Natl. Acad. Sci. U. S. A. 2007; 104:93–98. [PubMed: 17182741]

22. Das R, Abu-Abed M, Melacini G. Mapping allostery through equilibrium perturbation NMR spectroscopy. J. Am. Chem. Soc. 2006; 128:8406–8407. [PubMed: 16802799]

23. Akimoto M, et al. Signaling through dynamic linkers as revealed by PKA. Proc. Natl. Acad. Sci. U. S. A. 2013; 110:14231–14236. [PubMed: 23946424]

24. Boras BW, Kornev A, Taylor SS, McCulloch AD. Using Markov State Models to Develop a Mechanistic Understanding of Protein Kinase A Regulatory Subunit RIα Activation in Response to cAMP Binding. J. Biol. Chem. 2014; 289:30040–30051. [PubMed: 25202018]

25. Pande VS, Beauchamp K, Bowman GR. Everything you wanted to know about Markov State Models but were afraid to ask. Methods. 2010; 52:99–105. [PubMed: 20570730]

26. Prinz JH, et al. Markov models of molecular kinetics: Generation and validation. J. Chem. Phys. 2011; 134:174105. [PubMed: 21548671]

27. Chodera JDJ, Noé F. Markov state models of biomolecular conformational dynamics. Curr. Opin. Struct. Biol. 2014; 25C:135–144. [PubMed: 24836551]

28. Malmstrom RD, Lee CT, Van Wart AT, Amaro RE. Application of molecular-dynamics based markov state models to functional proteins. J. Chem. Theory Comput. 2014; 10:2648–2657. [PubMed: 25473382]

29. Lane TJ, Bowman GR, Beauchamp K, Voelz Va, Pande VS. Markov state model reveals folding and functional dynamics in ultra-long MD trajectories. J. Am. Chem. Soc. 2011; 133:18413–18419. [PubMed: 21988563]

30. Yang S, Banavali NK, Roux B. Mapping the conformational transition in Src activation by cumulating the information from multiple molecular dynamics trajectories. Proc Natl Acad Sci U S A. 2009; 106:3776–3781. [PubMed: 19225111]

31. Shukla D, Meng Y, Roux B, Pande VS. Activation pathway of Src kinase reveals intermediate states as targets for drug design. Nat. Commun. 2014; 5:3397. [PubMed: 24584478]

32. Bowman GR, Huang XH, Pande VS. Using generalized ensemble simulations and Markov state models to identify conformational states. Methods. 2009; 49:197–201. [PubMed: 19410002]

33. Kohlhoff KJ, et al. Cloud-based simulations on Google Exacycle reveal ligand modulation of GPCR activation pathways. Nat. Chem. 2014; 6:15–21. [PubMed: 24345941]

34. Anand GS, et al. Cyclic AMP- and (Rp)-cAMPS-induced conformational changes in a complex of the catalytic and regulatory (RI{alpha}) subunits of cyclic AMP-dependent protein kinase. Mol. Cell. Proteomics. 2010; 9:2225–2237. [PubMed: 20167947]

35. Berman HM, et al. The cAMP binding domain: an ancient signaling module. Proc. Natl. Acad. Sci. U. S. A. 2005; 102:45–50. [PubMed: 15618393]

36. Kannan N, et al. Evolution of allostery in the cyclic nucleotide binding module. Genome Biol. 2007; 8:R264. [PubMed: 18076763]

37. Shaw DE, et al. Atomic-Level Characterization of the Structural Dynamics of Proteins. Science. 2010; 330:341–346. [PubMed: 20947758]

38. Case, DA., et al. Amber 12. San Francisco: University of California; 2012.

39. Gotz AW, et al. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. J. Chem. Theory Comput. 2012; 8:1542–1555. [PubMed: 22582031]

40. Salomon-Ferrer R, Goetz AW, Poole D, Le Grand S, Walker RC. Routine micorsecond molecular dynamics simuations with AMBER - Part II: Particle Mesh Ewald. J. Chem. Theory Comput. 2012 In Review.

41. Bowman GR, Ensign DL, Pande VS. Enhanced Modeling via Network Theory: Adaptive Sampling of Markov State Models. J. Chem. Theory Comput. 2010; 6:787–794. [PubMed: 23626502]

42. Beauchamp KA, et al. MSMBuilder2: Modeling Conformational Dynamics at the Picosecond to Millisecond Scale. J. Chem. Theory Comput. 2011; 7:3412–3419. [PubMed: 22125474]

43. Ringheim GE, Saraswat LD, Bubis J, Taylor SS. Deletion of cAMP-binding site B in the regulatory subunit of cAMP-dependent protein kinase alters the photoaffinity labeling of site A. J. Biol. Chem. 1988; 263:18247–18252. [PubMed: 2848031]

44. Metzner P, Schutte C, Vanden-Eijnden E. Transition Path Theory for Markov Jump Processes. Multiscale Model. Simul. 2009; 7:1192–1219.

45. Prinz JH, Keller B, Noe F. Probing molecular kinetics with Markov models: metastable states, transition pathways and spectroscopic observables. Phys. Chem. Chem. Phys. 2011; 13:16912–16927. [PubMed: 21858310]

46. Herberg FW, Taylor SS, Dostmann WR. Active site mutations define the pathway for the cooperative activation of cAMP-dependent protein kinase. Biochemistry. 1996; 35:2934–2942. [PubMed: 8608131]

47. Badireddy S, et al. Cyclic AMP analog blocks kinase activation by stabilizing inactive conformation: conformational selection highlights a new concept in allosteric inhibitor design. Mol. Cell. Proteomics. 2011; 10 M110.004390.

48. Berezovska G, Prada-Gracia D, Mostarda S, Rao F. Accounting for the kinetics in order parameter analysis: Lessons from theoretical models and a disordered peptide. J. Chem. Phys. 2012; 137:194101. [PubMed: 23181288]

49. Chodera JD, Singhal N, Pande VS, Dill KA, Swope WC. Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics. J. Chem. Phys. 2007; 126:155101. [PubMed: 17461665]

50. Moll D, Schweinsberg S, Hammann C, Herberg FW. Comparative thermodynamic analysis of cyclic nucleotide binding to protein kinase A. Biol. Chem. 2007; 388:163–172. [PubMed: 17261079]

51. Möller S, et al. Cyclic Nucleotide Mapping of Hyperpolarization-Activated Cyclic Nucleotide-Gated (HCN) Channels. ACS Chem. Biol. 2014; 9:1128–1137. [PubMed: 24605759]

52. Son I, Selvaratnam R, Dubins DN, Melacini G, Chalikian TV. Ultrasonic and densimetric characterization of the association of cyclic AMP with the cAMP-binding domain of the exchange protein EPAC1. J. Phys. Chem. B. 2013; 117:10779–10784. [PubMed: 23968295]

53. Wang JM, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J. Comput. Chem. 2004; 25:1157–1174. [PubMed: 15116359]

54. Wang JM, Wang W, Kollman PA, Case DA. Automatic atom type and bond type perception in molecular mechanical calculations. J. Mol. Graph. Model. 2006; 25:247–260. [PubMed: 16458552]

55. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. J. Chem. Phys. 1983; 79:926–935.

56. Hornak V, et al. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins. 2006; 65:712–725. [PubMed: 16981200]

57. Malmstrom RD, Kornev A, Taylor SS, Amaro RE. Data from: Allostery through conformational selection: cAMP regulation of the cyclic-nucleotide binding domain in PKA. 2014

58. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. J. Chem. Theory Comput. 2013; 9:3084–3095. [PubMed: 26583988]

59. Humphery W, Dalke A, Schulten K. VMD-Visual Molecular Dynamics. J. Molec. Graph. 1996; 14:33–38. [PubMed: 8744570]

60. Phillips JC, et al. Scalable molecular dynamics with NAMD. J. Comput. Chem. 2005; 26:1781–1802. [PubMed: 16222654]

61. Hagberg, AA.; Schult, DA.; Swart, PJ. Exploring network structure, dynamics, and functions using NetworkX. In: Varoquaux, G.; Vaught, T.; Millman, J., editors. 7th Python Sci. Conf. 2008. p. 11-15.

62. Singhal N, Pande VS. Error analysis and efficient sampling in Markovian state models for molecular dynamics. J. Chem. Phys. 2005; 123:204909. [PubMed: 16351319]

63. Shukla D, Meng Y, Roux B, Pande VS. Activation pathway of Src kinase reveals intermediate states as targets for drug design. Nat. Commun. 2014; 5:3397. [PubMed: 24584478]

64. Altis A, Otten M, Nguyen PH, Hegger R, Stock G. Construction of the free energy landscape of biomolecules via dihedral angle principal component analysis. J. Chem. Phys. 2008; 128:245102. [PubMed: 18601386]
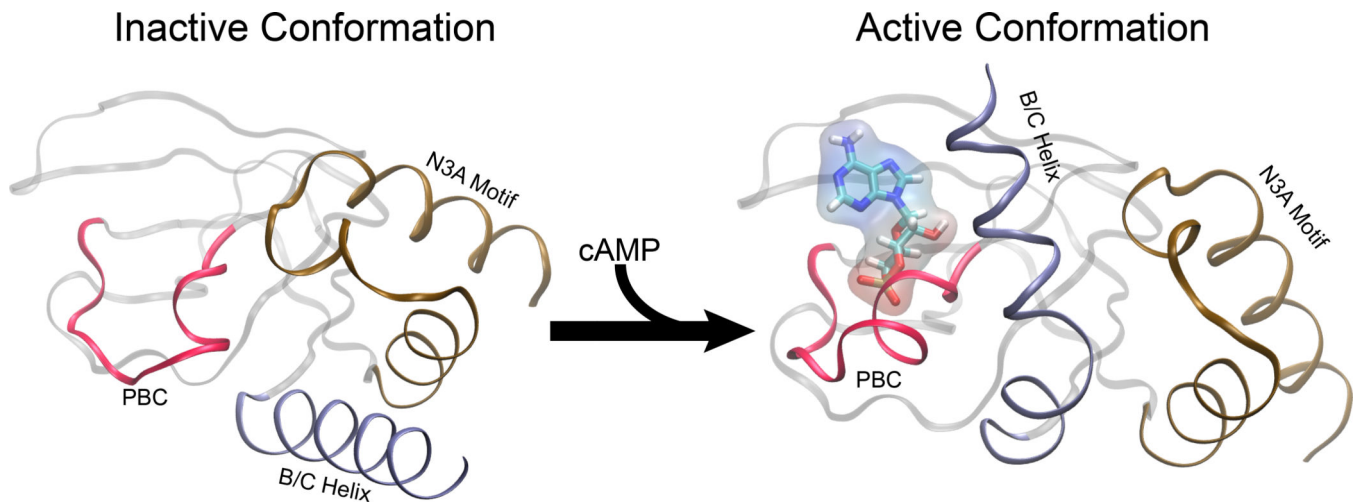
Figure 1. Experimentally Determined Conformational Changes in the Cyclic-nucleotide Binding Domain

(RIα aa. 119 to 244) in cAMP bound active, right, and holoenzyme inactive, left, with key signaling structural sub-domains highlighted: phosphate binding cassette (PBC) – red, B/C helix – Ice blue, and N3A motif – ochre.
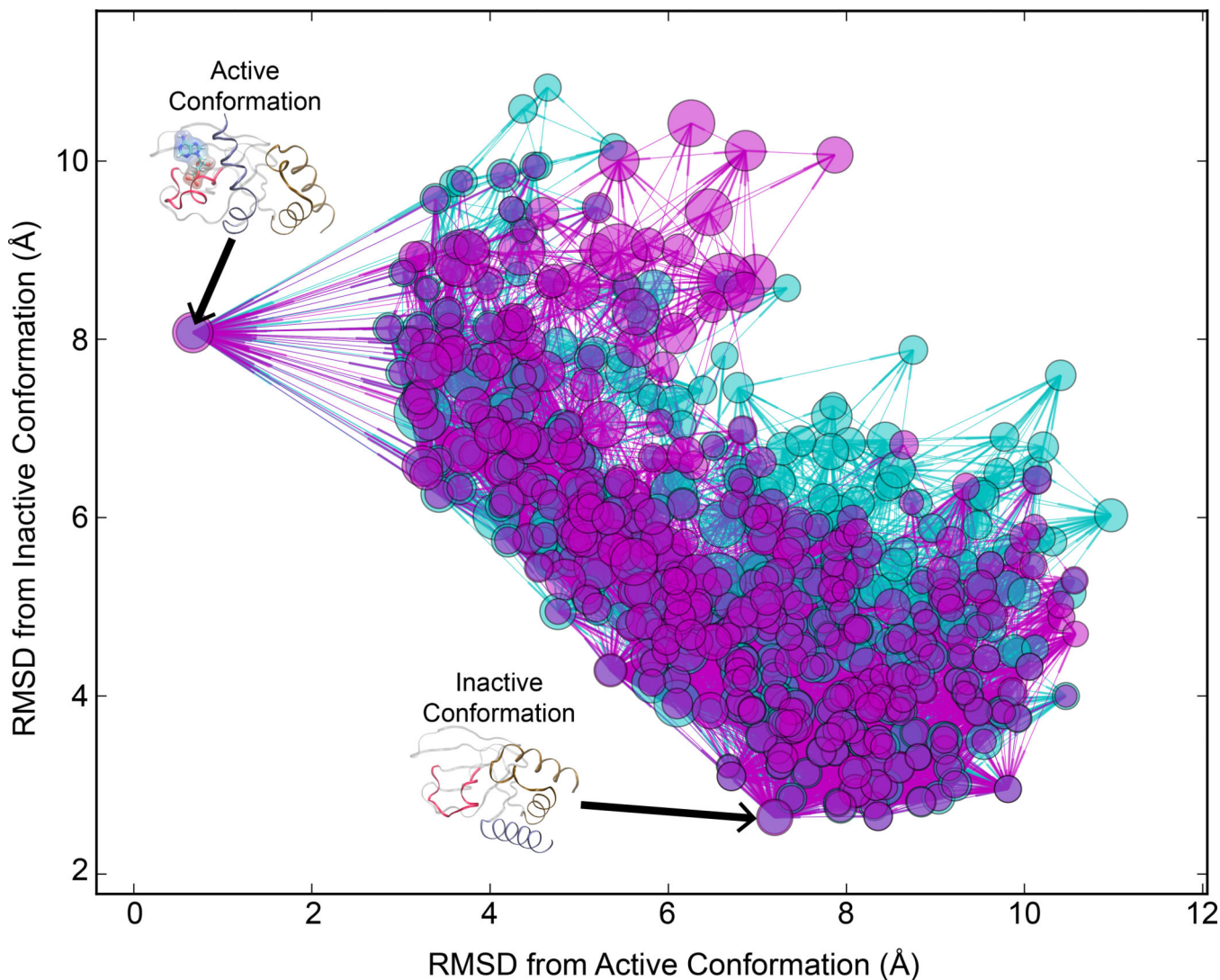
Figure 2. Overlay of Markov State Models State Graphs for the Cyclic-nucleotide Binding Domain with (CBD-cAMP, magenta) and without (CBD-Apo, cyan) cAMP bound

The position of each conformational state node is the RMSD of the generator of each cluster relative to the reference conformations. The diameter of each conformational state node (cyan, magenta) is proportional to the relative the log of the equilibrium population. (i.e. The lager the node the more probable the state at equilibrium.) Cyan and magenta nodes are transparent; therefore dark purple nodes represent nodes that have occupancy in both CBD-Apo and CBD-cAMP Markov state models (i.e., dark purple nodes overlap).
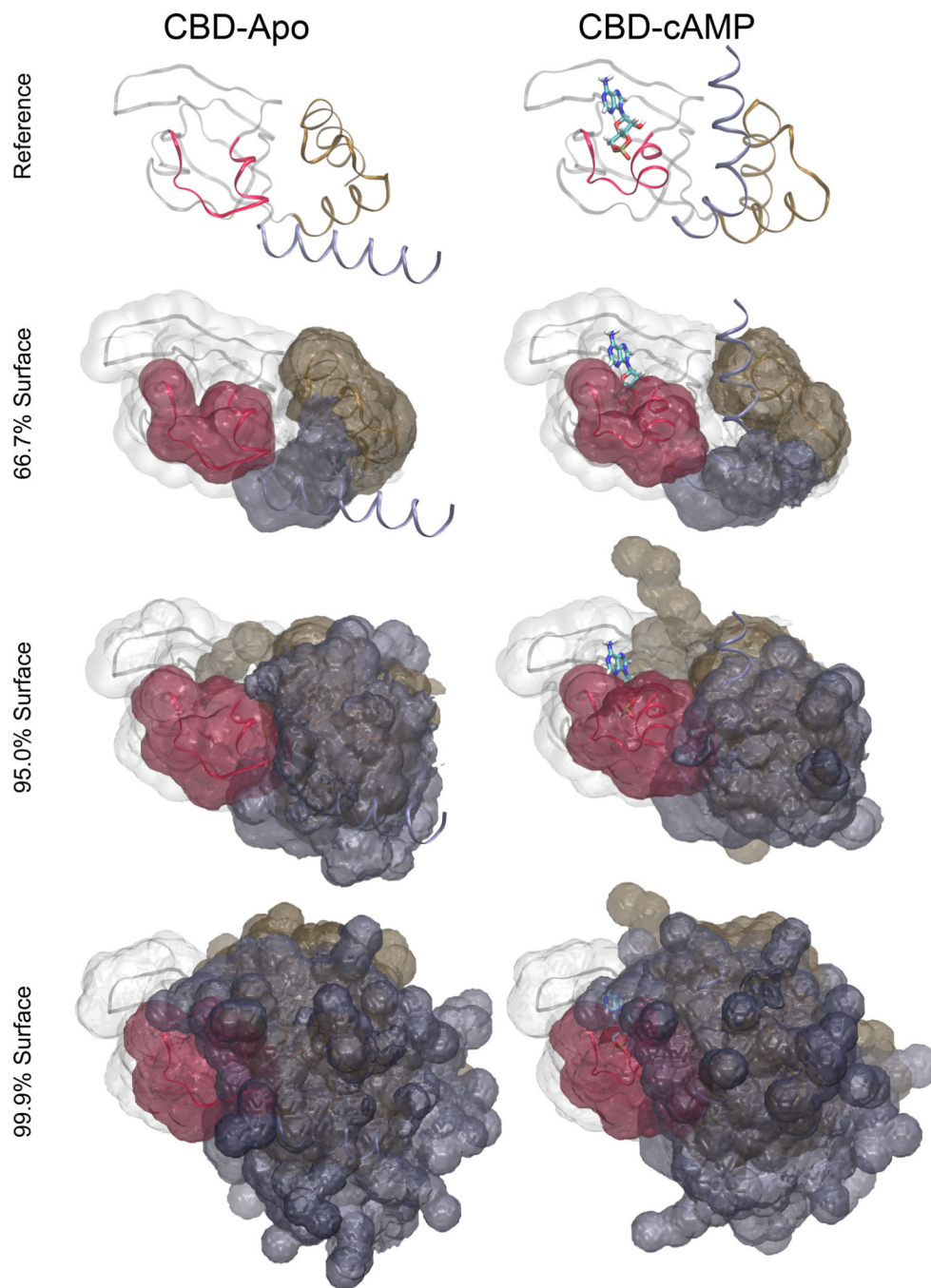
Figure 3. Equilibrium Probability Density Surfaces for CBD-Apo and CBD-cAMP at 66%, 95%, and 99.9% Cutoffs

The surfaces correspond to the probability density surface for each residues alpha carbon. Densities contributed from key structural subdomains are colored: phosphate binding cassette (PBC) – red, B/C helix – Ice blue, and N3A motif – ochre. See figure 1 for description of CBD representations.
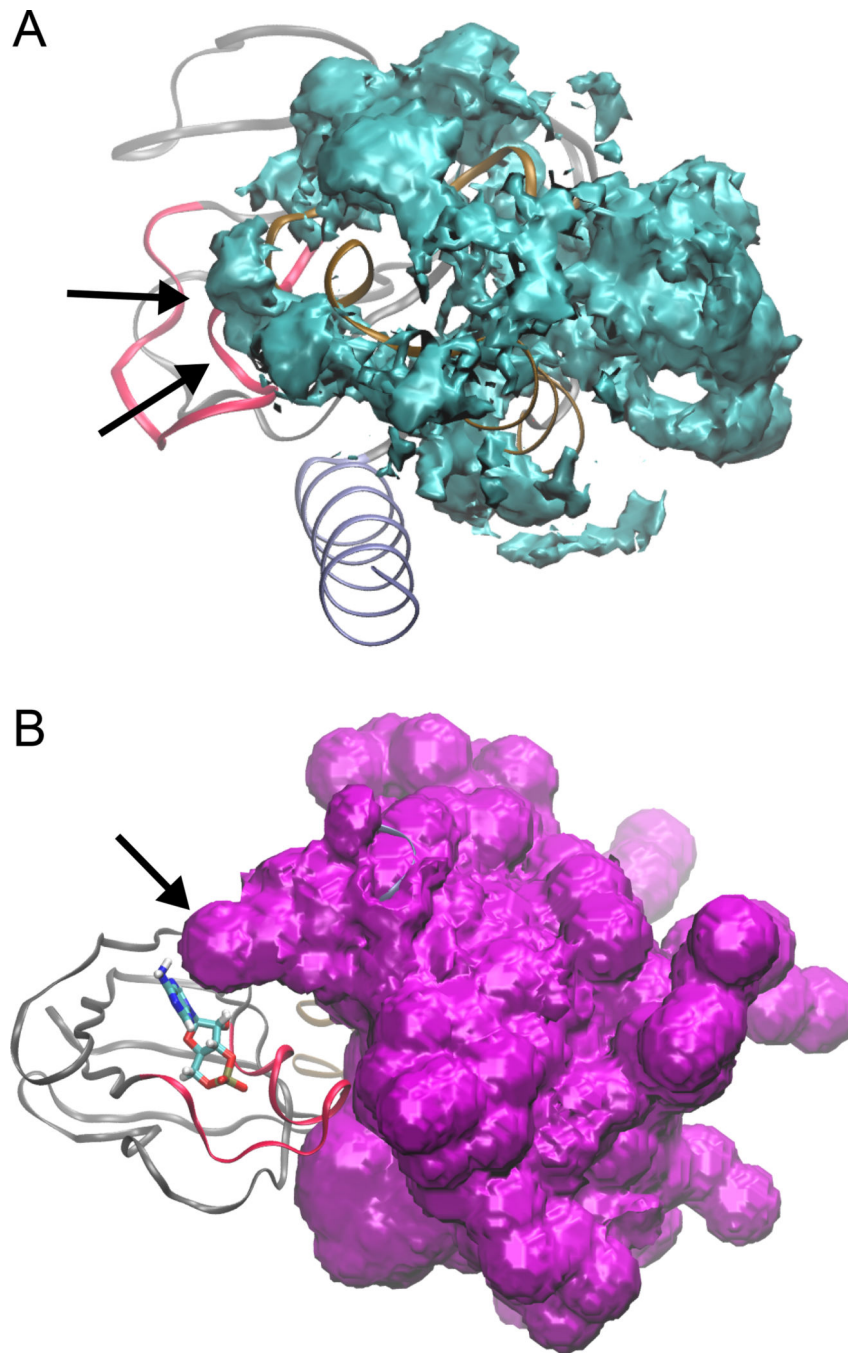
Figure 4. Selected Difference Surfaces Between Equilibrium Probability Densities of CBD-Apo and CBD-cAMP
A) The different surface for 10% or more density for the N3A Motif of CBD-Apo over CDB-cAMP. Arrows indicate interface between PKA's catalytic and regulatory subunits B) The different surface for 1% or more density for the B/C Helix of CBD-Apo over CDB-cAMP.
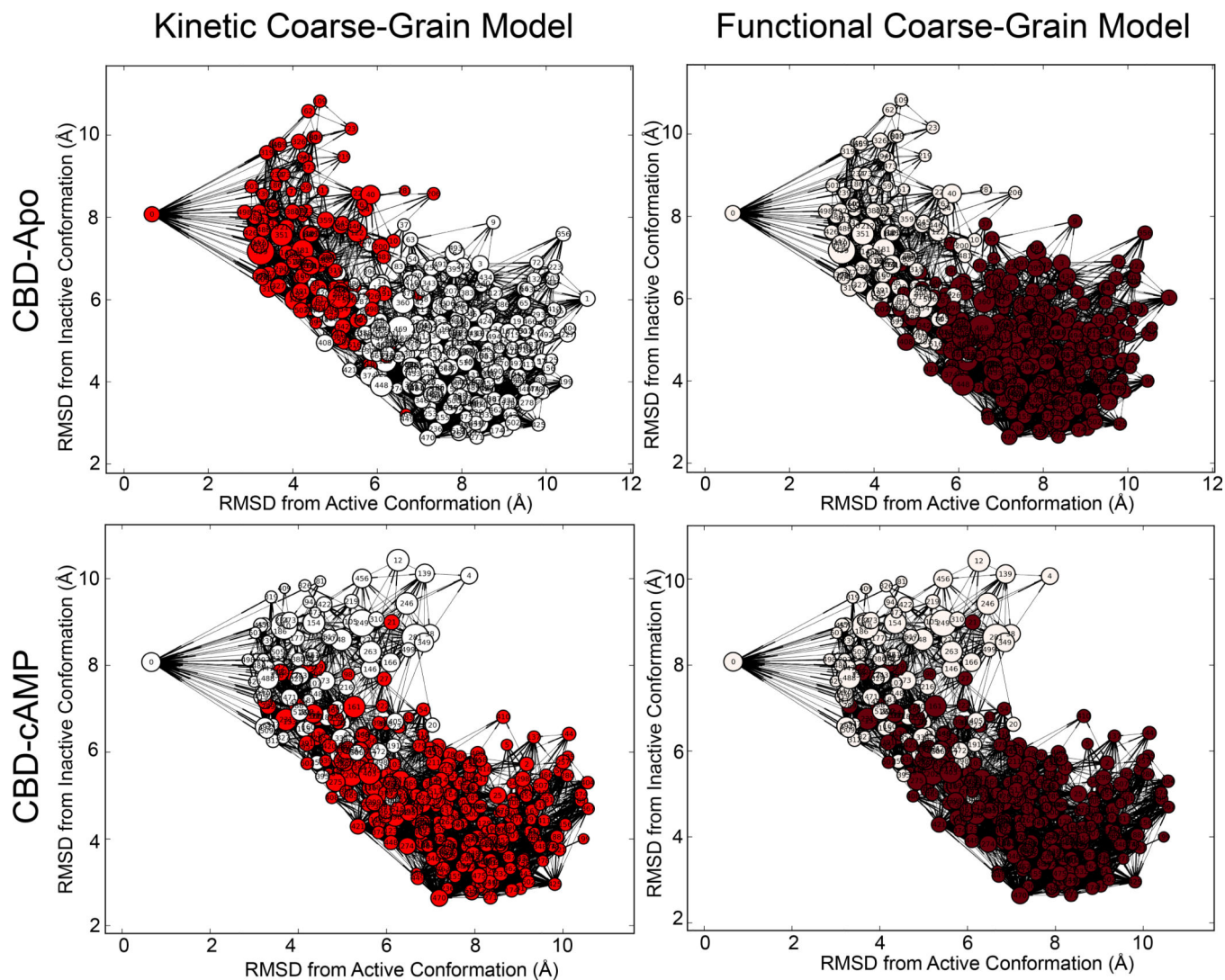
Figure 5. Comparison of Conformational State Assignment for Kinetic and Functional Coarse-Grain Models

Conformational state meta-stable states for the first/slowed motions in the kinetic coarse-grain Markov state model graph are colored below according to macrostate membership (red, white). Functional coarse-grain models are colored either white or maroon, depending on whether the state "would go" to the active or inactive conformations, respectively.
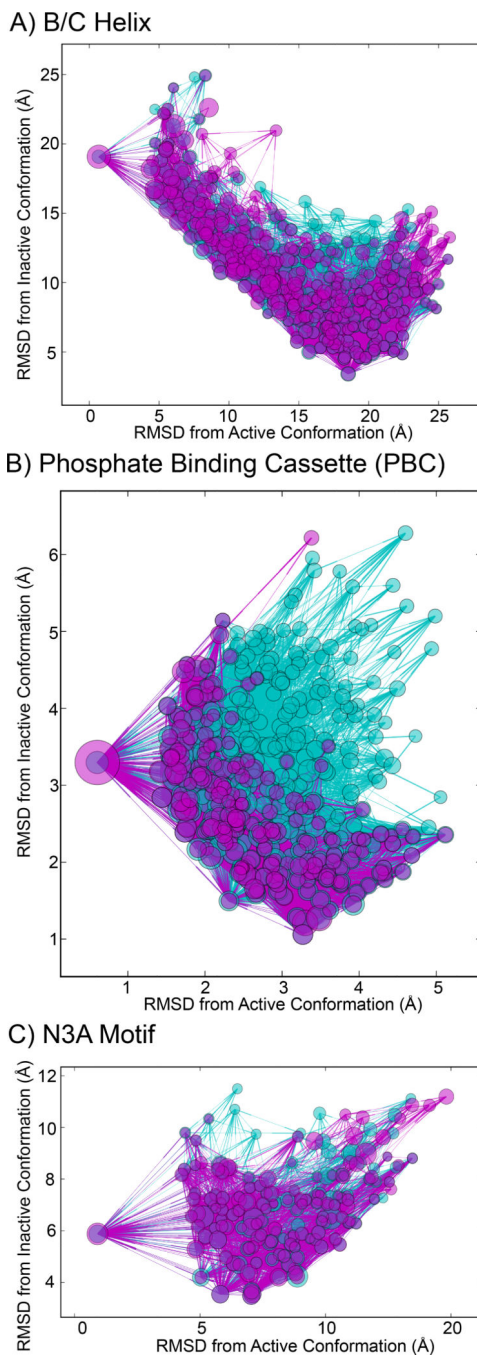
Figure 6. Overlay of Markov State Model Conformational State Graphs for the Sub-Domains of Cyclic-nucleotide Binding Domain with (CBD-cAMP, magenta) and without (CBD-Apo, cyan) cAMP Bound

Markov state models conformational state graphs use the same conventions defined in Figure 2.

Table 1

Mean First Passage Time Between Putative Functional Microstates.

	Without cAMP		With cAMP	
System	**A→I**	**A←I**	**A→I**	**A←I**
CBD	16.1 μs	34.0 μs	83.7 μs	30.3 μs
PBC	2.9 μs	0.9 μs	4.9 μs	0.2 μs
B/C Helix	8.2 μs	13.4 μs	24.3 μs	10.6 μs
N3A	1.3 μs	2.9 μs	2.2 μs	1.6 μs

Table 2

Population and Half-life for CBD-Apo and CBD-cAMP Functional Macrostates

| Macrostate | CBD-Apo | | CBD-cAMP | |
	Population	$t_{1/2}$ Mean(STD)	Macrostate	Population	$t_{1/2}$ Mean(STD)
Active	36%	2.29(0.09) µs	Active	77%	36.59(1.12) µs
Inactive	64%	8.62(0.41) µs	Inactive	23%	7.41(0.45) µs