ROLE OF DYNAMICS IN CYCLIC-NUCLEOTIDE-MODULATED ALLOSTERY

ROLE OF DYNAMICS IN CYCLIC-NUCLEOTIDE-MODULATED ALLOSTERY

By BRYAN VANSCHOUWEN, HONOURS B.Sc., M.Sc.

A Thesis Submitted to
the School of Graduate Studies in Partial Fulfilment
of the Requirements for the Degree

Doctor of Philosophy

Doctor of Philosophy (2015) **McMaster University**

(Chemistry) **Hamilton, Ontario**

Thesis Title: Role of Dynamics in Cyclic-Nucleotide-Modulated Allostery

Author: Bryan VanSchouwen, Honours B.Sc. (Brock University), M.Sc. (Brock University)

Supervisor: Professor Giuseppe Melacini

Number of Pages: xv, 205

Lay Abstract:

In this thesis, we examined cyclic-nucleotide-responsive proteins that regulate key physiological processes, and whose malfunction has been linked to cardiovascular and neurological disorders. In particular, in three such proteins we examined dynamics, whose role in cyclic-nucleotide-responsive function is not fully understood. We found that cyclic-nucleotide-dependent variations in dynamics play a critical role in the function of these proteins, with the results for each protein highlighting a different role played by dynamics. Ultimately, we envision that the methods outlined in this thesis will reveal key functional differences among human cyclic-nucleotide-responsive proteins that can eventually lead to the development of novel therapeutics to treat certain diseases such as arrhythmias or epilepsy by selectively targeting a single cyclic-nucleotide-responsive protein.

Abstract:

Cyclic nucleotides such as cAMP and cGMP serve as intracellular second messengers in diverse signaling pathways that control a wide range of cellular functions. Such pathways are regulated by key cyclic nucleotide receptor proteins including protein kinase A (PKA), the exchange protein directly activated by cAMP (EPAC), the hyperpolarization-activated cyclic-nucleotide-modulated (HCN) ion channels, and protein kinase G (PKG), and malfunction of these proteins has been linked to a number of pathologies. While it is known that cyclic nucleotide binding to these proteins leads to structural perturbations that promote their activation, the role played by dynamics in auto-inhibition and cyclic-nucleotide-dependent activation is not fully understood. Therefore, in this thesis we examined dynamics within the cyclic-nucleotide receptor proteins EPAC, HCN and PKG, and found that dynamics are critical for allosteric control of activation and/or autoinhibition of all three proteins. In particular, our findings for EPAC and HCN have highlighted dynamics as a key modulator of the entropic and enthalpic components, respectively, of the free-energy landscape for cAMP-dependent allostery, while our findings for PKG have highlighted dynamics as a key determinant of the cGMP-*vs.*-cAMP selectivity necessary to minimize cross-talk between signaling pathways. Ultimately, we envision that the methods outlined in this thesis will reveal key differences in the regulatory mechanisms of human cyclic nucleotide receptors that can eventually be exploited in the development of novel therapeutics to selectively target a single receptor, and thus treat physiological conditions/diseases linked to malfunction of the target receptor.

Acknowledgements:

First and foremost, thanks to my supervisor, Dr. Giuseppe Melacini, for all of the guidance and mentorship over the course of my time at McMaster. It has been a great pleasure to work as part of your research group.

To my fellow lab members, past and present, for all of their help in making my graduate experience so fruitful: Rajeevan Selvaratnam and Madoka Akimoto, thanks for all the help in getting me acquainted with the wet lab and NMR work, and for all the hard work and helpful discussions that you contributed to my research projects. Rajanish Giri, thanks for the assistance with PKG sample preparation work. Julijana Milojevic, Stephen Boulton, Kody Moleschi and Maryam Sayadi, thanks for the helpful and stimulating discussions.

To the undergraduate thesis/summer students that I had the privilege to work with: Matthew L. Brown, Matthew Yu, Namir Jalanko, Roman Malekzai, and you too, Stephen.

Thanks to my committee members, Drs. Paul Ayers and Paul Berti, for their helpful discussions and constructive criticisms during my committee meetings over the past five years. To the Centre for Microbial Chemical Biology (CMCB) and the Berti research lab, for allowing me to use their equipment for my protein sample preparation work.

I am also grateful to my undergraduate and Master's thesis supervisors, Drs. Stuart M. Rothstein and Heather L. Gordon, for all of their guidance that set me on the path to my current achievements. Good luck in your retirement, Dr. Rothstein.

Last, but not least: To my parents for all of their support throughout my Ph.D. studies, even when they had only the faintest clue about what I was working on at the lab.

Table of Contents

List of Main-Text Figures

Chapter 1: Introduction

Chapter 2: Role of Dynamics in the Auto-Inhibition and Activation of the Exchange Protein Directly Activated by Cyclic AMP (EPAC)

Chapter 3: Role of Dynamics in the Auto-Inhibition and Activation of the Hyperpolarization-Activated Cyclic-Nucleotide-Modulated (HCN) Ion Channels

Chapter 4: Mechanism of cAMP Partial Agonism in Protein Kinase G (PKG)

Chapter 5: Concluding Remarks and Future Outlook

List of Supplementary Figures

Chapter 2: Role of Dynamics in the Auto-Inhibition and Activation of the Exchange Protein Directly Activated by Cyclic AMP (EPAC)

List of Main-Text Tables

Chapter 2: Role of Dynamics in the Auto-Inhibition and Activation of the Exchange Protein Directly Activated by Cyclic AMP (EPAC)

Chapter 3: Role of Dynamics in the Auto-Inhibition and Activation of the Hyperpolarization-Activated Cyclic-Nucleotide-Modulated (HCN) Ion Channels

List of Supplementary Tables

Chapter 2: Role of Dynamics in the Auto-Inhibition and Activation of the Exchange Protein Directly Activated by Cyclic AMP (EPAC)

List of Abbreviations

2-NH_2-cPuMP	cyclic 2-amino-purine riboside monophosphate
AEBSF	4-(2-aminoethyl)benzenesulfonyl fluoride
AI	autoinhibitory linker region
ATP	adenosine triphosphate
BBR	base binding region
cAMP	cyclic adenosine monophosphate
CAP	catabolite activator protein
CBD	cyclic-nucleotide-binding domain (see also “CNB”)
CD	circular dichroism
CDC25-HD	CDC25-homology domain of EPAC
cGMP	cyclic guanosine monophosphate
cIMP	cyclic inosine monophosphate
cPuMP	cyclic purine riboside monophosphate
CHESCA	NMR chemical shift covariance analysis
CHESPA	NMR chemical shift projection analysis
CNB	cyclic-nucleotide-binding domain (see also “CBD”)
CR	catalytic region of EPAC
DBD	DNA-binding domain
DCCM	dynamic cross-correlation matrix
D/D	N-terminal dimerization domain
DEP	Dishevelled Egl-10 Pleckstrin domain of EPAC
DTT	dithiothreitol
EPAC	exchange protein directly activated by cAMP
FPLC	fast protein liquid chromatography
GEF	guanine nucleotide exchange factor
HCN	hyperpolarization-activated cyclic-nucleotide-modulated ion channel
H/D	hydrogen/deuterium exchange NMR
HH	hydrophobic hinge
HMQC	heteronuclear multiple-quantum coherence
HP	helical hairpin of EPAC
HSQC	heteronuclear single-quantum coherence
IL	ionic latch of EPAC
IR	intracellular region of HCN
MD	molecular dynamics
MOPS	3-(N-morpholino)propanesulfonic acid
MW	molecular weight
N3A	N-terminal α-helix bundle of PKA, PKG and HCN CBDs
NADH	nicotinamide adenine dinucleotide (reduced form)
NMR	nuclear magnetic resonance
NMRD	nuclear magnetic relaxation dispersion
NO	nitric oxide
NOE	nuclear Overhauser effect

NTHB	N-terminal α-helix bundle of EPAC CBD
PBC	phosphate binding cassette
PCA	principal component analysis
PF	hydrogen exchange protection factor
PKA	protein kinase A
PKG	protein kinase G
PRE	paramagnetic relaxation enhancement
RA	Ras association domain of EPAC
REM	Ras exchange motif domain of EPAC
RMSD	root-mean-square deviation
RMSF	root-mean-square fluctuation
ROESY	rotating frame nuclear Overhauser effect spectroscopy
R_p-cAMPS	cyclic adenosine monophosphorothioate (R_p-isomer)
R_p-cGMPS	cyclic guanosine monophosphorothioate (R_p-isomer)
RR	regulatory region of EPAC
RSD	reduced spectral density
SB	switchboard region of EPAC
SM	similarity measure
STD	saturation transfer difference NMR
SW	switch helix region of PKG
TD	tetramerization domain of the HCN IR
TM	trans-membrane region of HCN
transfer-NOESY	transferred nuclear Overhauser effect spectroscopy
TROSY	transverse relaxation-optimized spectroscopy
VDW	van der Waals

CHAPTER 1

INTRODUCTION

1.1. Introduction to Cyclic Nucleotide Signaling

Cyclic nucleotides such as cyclic AMP (cAMP) and cyclic GMP (cGMP) serve as intracellular second messengers in diverse signaling pathways that control a wide range of cellular functions.[1,2,3,4,5,6,7,8,9,10,11,12] Stimulation of adenylate or guanylate cyclases by upstream intracellular signals (*e.g.* G-protein-coupled receptor pathway, NO binding) results in increased intracellular concentrations of cAMP or cGMP, respectively, which in turn bind and activate key cyclic nucleotide receptor proteins.[1,4,6,7,13] These receptor proteins include protein kinase A (PKA), the exchange protein directly activated by cAMP (EPAC), the hyperpolarization-activated cyclic-nucleotide-modulated (HCN) ion channels, and protein kinase G (PKG), and are responsible for regulating a multitude of downstream signaling pathways (Figure 1).

PKA and EPAC are involved in regulation of such processes as cell proliferation, cell adhesion, cell-cell junction formation, insulin secretion by pancreatic β-cells, and processing of amyloid proteins. Impaired function of PKA or EPAC has been implicated in cardiovascular pathologies, diabetes, and Alzheimer's disease.[1,2,3,8,9,10,11] The HCN ion channels, meanwhile, play a key role in nerve impulse transmission and heart rate modulation in neuronal and cardiac cells, respectively, and HCN mutations have been implicated in such physiological conditions as bradycardia and epilepsy.[4,5] Finally, PKG is involved in regulation of platelet activation, memory formation, smooth muscle tone and vasodilation, and PKG mutations have been implicated in dysfunctional smooth muscle contraction, and aortic aneurysms and dissections.[6,7,12,13,14,15,16,17,18] Therefore, these cyclic

nucleotide receptors represent attractive therapeutic targets, and thus, the study of cyclic nucleotide recognition and allostery in these proteins is of great interest.

A common functional element involved in the regulation of all cyclic-nucleotide receptor proteins by cyclic nucleotides is an intrinsic protein domain known as a "cyclic-nucleotide-binding domain" ("CBD"; alternatively abbreviated as "CNB").[6,7,10,12,19,20,21,22,23] This domain is responsible for binding the cyclic nucleotide second messenger, and in response undergoes a change in structure and/or dynamics that is coupled to modulation of the function of a catalytic domain or ion channel elsewhere in the host protein – often through changes in the overall structural arrangement of the host protein (Figure 2).[6,7,10,13,19,20,21,12] While it is known that cyclic nucleotide binding to CBDs leads to perturbations that promote activation of key host protein functions, the role played by dynamics in cyclic nucleotide-dependent activation is not fully understood. Therefore, a major focus of this thesis is an in-depth analysis of cyclic-nucleotide-dependent dynamics, which is essential to achieve a more complete understanding of CBD function, and to fully exploit the potential of CBDs as therapeutic targets.

1.2. Dynamics in the Context of Cyclic-Nucleotide Binding Domains (CBDs)

Dynamics in cyclic-nucleotide-binding domains (CBDs) occur at multiple levels encompassing several time- and length-scales. First of all, NMR studies have suggested that apo CBDs exist in a dynamic auto-inhibitory equilibrium between inactive (*i.e.* inhibition-competent) and active (*i.e.* inhibition-incompetent) conformational states (Figure 3). The inactive-*vs.*-active interconversion is fast in the chemical shift NMR time scale and therefore the observed chemical shifts are an average of those of the isolated states. The active and

inactive states differ not only in structure, but also in their profiles of structural fluctuations affecting both the peptide backbone and the side chains ("internal dynamics"). Such fluctuations can occur on short (ps-ns) or longer (ms-μs) time scales, and in some cases, they can result in a transient local unfolding with exposure of protein functional groups to the aqueous solvent. The active *vs.* inactive state differences both at the level of structure and of internal dynamics are critical to explain how cyclic nucleotides and/or mutations regulate protein activity.

Dynamics are paramount not only for the CBD, but also within the cyclic nucleotide ligand itself, which exists in equilibrium between *syn* and *anti* conformational states (Figure 3). In the *syn* conformation the cyclic nucleotide base points toward the ribose moiety, while in the *anti* conformation it is oriented in the opposite direction. It has been found that CBDs typically bind cyclic nucleotides in one of these two conformations, and the *syn*/*anti* conformational tendencies of cyclic nucleotides are thus expected to play a key role in CBD activation. Overall, the allosteric control of eukaryotic CBDs investigated so far appears to conform to a general model of "reciprocal conformational selection", whereby it is not only the cyclic nucleotide that selects for a specific state pre-sampled by the *apo* CBD, but it is also the CBD that selectively binds one of the conformations adopted by the *apo* ligand (Figure 3).[24,25,26,27,28]

1.3. Why Are CBD Dynamics Functionally Significant?

Dynamics play a key role in protein allostery and this is particularly evident in the context of CBDs, as proven by several recent applications of NMR to CBDs. An example of the relevance of dynamics in allostery is provided by PKA, in which a flexible linker region

N-terminal to the regulatory-subunit CBD-A domain was found to serve as a key allosteric element controlling PKA activity, despite the fact that this linker is only partially structured.[29] Specifically, the linker forms weak, conformation-selective interactions with the active conformation of CBD-A, thus tuning the CBD-A inactive/active conformational equilibrium of the *apo* regulatory-subunit to a near-degenerate state that facilitates an optimal response of PKA to cAMP (Figure 4).[29] Furthermore, the linker/catalytic-subunit interactions formed in the inactive state occur at the expense of disrupting the active-conformation-selective linker/CBD-A interactions and *vice versa* (Figure 4), resulting in a structural frustration within the linker that facilitates disruption of the large regulatory-subunit/catalytic-subunit interface in response to the low-molecular-weight (MW) ligand cAMP.[29]

Dynamics are also a key entropic determinant of the free energy changes that control allosteric processes (Figure 3). This general concept has been elegantly illustrated by the group of Kalodimos *et al.* through a series of pioneering publications on the catabolite activator protein (CAP), which is a dimer including two CBDs. In CAP, dynamics controls not only the free energy of allosteric coupling between the two cAMP-binding sites,[30] but also the free energy of DNA binding by CAP. In this respect, a striking example illustrating the relevance of dynamics in allostery is provided by a recently-studied point mutant of CAP, where cAMP binding to the mutant promotes its binding to DNA – forming a complex with a structure very similar to that of the wild-type protein – despite an apparent failure to induce the active, DNA-binding competent conformation in the DNA-binding domain (DBD) of the mutant.[31] NMR analyses revealed that despite the apparent lack of a full shift to the active state, cAMP binding to the mutant CAP resulted in a dynamic exchange between an inactive

DBD conformation and a weakly-populated active DBD conformation.[31] The latter binds DNA with high affinity due to a DNA-induced enhancement of dynamics, which promotes an entropic stabilization of the DNA-bound form – *i.e.* whereas binding of wild-type CAP to DNA is driven by enthalpy, binding of the mutant to DNA is driven by entropy resulting from an increase in structural fluctuations of the mutant upon DNA binding.[31] Since the DNA binding affinity of the active DBD conformation is several orders of magnitude higher than that of the inactive DBD conformation, it was postulated that DNA would selectively bind to the weakly-populated active DBD conformation, thus shifting the conformational equilibrium from the inactive to the active conformation.[31] Conversely, suppression of transient conformational states is useful for eliciting allosteric inhibition.[32] Dynamically-driven allosteric phenomena, although first experimentally proven for the prokaryotic CAP,[30,31,32,33] are also common in eukaryotic CBDs. For example, cAMP binding to the CBD of EPAC was found to result in an increase in the internal dynamics of the domain's N-terminal α-helical bundle (NTHB), with only a minimal change in the secondary or tertiary structure of the NTHB.[34] Such dynamic enhancement was proposed to entropically promote dissociation of a known autoinhibitory interface formed between the NTHB and the catalytic domain of EPAC, thereby promoting EPAC activation.[34]

Besides an involvement in protein allostery, dynamics have also been implicated as a key factor in cyclic nucleotide binding and selectivity of CBDs. For example, the higher cAMP-binding affinity of PKA, compared to other eukaryotic CBDs such as those of EPAC and HCN, was recently attributed to the relative positions of the inactive/active conformational equilibria of the apo-state CBDs.[35] Specifically, while the apo-state CBDs of EPAC and HCN have been proposed to exist predominantly in their inactive conformations,

the apo-state CBD-A of PKA exists in an equilibrium containing a greater population of the active conformation.[29,35] Since the active CBD conformation is expected to have a higher binding affinity than the inactive conformation, the greater apo-state population of active CBD was postulated to result in the greater observed binding affinity of the PKA CBD-A, relative to EPAC and HCN.[35] Indeed, C-terminal deletion mutations of EPAC that were found to increase its apo-active population also significantly increased the observed binding affinity of the EPAC CBD for cAMP,[36] providing further evidence of the relevance of CBD conformational equilibria for binding. However, it is notable that in the prokaryotic CBD of CAP, mutation-induced changes in the CBD inhibitory equilibrium do not always correlate with the observed changes in DNA affinity.[33] This apparent paradox was solved by considering that often mutations that modulate the inhibitory equilibrium are also effective modulators of the entropy of DNA binding, even in the absence of apparent structural changes.[33]

The role of dynamics in cyclic nucleotide selectivity is illustrated by the effects of cAMP and cGMP on the EPAC CBD.[37] Unlike cAMP, cGMP is an EPAC antagonist. This cAMP *vs.* cGMP functional difference with respect to EPAC activation is explained by the observation that while cAMP binds to EPAC in a *syn* conformation, cGMP binds in an *anti* conformation (Figure 3). The alternative binding mode of cGMP leads to a decrease in the internal dynamics of the CBD NTHB, compared to cAMP-bound EPAC.[37] The reduced NTHB dynamics were proposed to hinder the entropically-induced NTHB/catalytic domain dissociation normally promoted by cAMP, thus highlighting a selection against EPAC activation by cGMP at the level of CBD dynamics.[37] Last but not least, it should be noted that in addition to improving our understanding of CBD function, analysis of CBD dynamics

is useful also as an ancillary tool to other experimental techniques. For instance, analysis of flexible N- and C-termini by NMR can be used for the purpose of accurately identifying domain boundaries of a minimal CBD construct to be utilized in further structural characterizations.

The examples outlined above illustrate the pivotal role of dynamics in cyclic-nucleotide-dependent modulation of protein function, and highlight the need to map the dynamic profiles of CBDs across different time- and length-scales. NMR spectroscopy is ideally suited to serve this purpose, and Sections 1.4–1.6 below summarize methods that have been previously utilized to investigate CBD allostery and dynamics by NMR and simulations, as well as examples of notable findings achieved through these methods.

1.4. Selection of CBD Constructs for Examination

Cyclic nucleotide-dependent proteins are often multi-domain systems, and meaningful NMR analyses of cyclic nucleotide-dependent dynamics require the use of a family of protein constructs, which include both single- and multi-domain fragments of the wild-type protein under investigation. Longer constructs, to the ideal limit of the full-length wild-type protein, are obviously preferable in terms of biological relevance, but often pose technical challenges in terms of resolution and sensitivity of solution NMR spectra. Specifically, the large size of the intact protein slows its tumbling in solution, thus accelerating transverse relaxation, and consequently, reducing the efficiency of coherence transfer in multi-dimensional NMR experiments and increasing line broadening. As a result, the sensitivity and resolution of NMR spectra are typically reduced as the protein MW increases. This is a major drawback because high-MW proteins often exhibit limited

solubility, thus hindering the use of protein sample concentrations high enough to yield good signals, as well as extensive peak overlap due to the large number of signals, which can obscure features of interest in the NMR spectra. Although recent advances in isotopic labeling schemes and pulse sequence design[38,39,40,41,42] elegantly circumvent the technical challenges traditionally associated with the NMR investigation of large-MW proteins, it is often advantageous to complement the NMR analyses of multi-domain constructs with studies of shorter fragments, including one-domain segments.

Lower-MW constructs not only provide enhanced sensitivity and resolution in NMR spectra, but often simplify sample preparation through deletion of flexible structural elements that promote self-association (and potentially aggregation), or of structural elements that normally reside in non-solution environments, such as the trans-membrane regions of HCN ion channels.[43,44,45] Low-MW constructs also serve as a stepping-stone in the examination of the intact, full-length protein. Therefore, NMR analyses of CBDs have often focused on fragments of the respective host proteins, usually containing the CBD(s) of interest,[29,34,46,47,48] and occasionally, additional structural elements N- and/or C-terminal to the CBD(s).[29,48] Since such protein NMR investigations rely on examination of protein fragments, well-designed protein constructs with carefully selected domain boundaries and containing the region(s) of interest are a necessity for a successful investigation.

1.5. Validation of CBD Constructs Selected for Examination

Although low-MW protein constructs benefit from increased NMR sensitivity and resolution relative to higher-MW constructs, and hence have the potential to provide access to a wealth of dynamical and structural information on CBDs, it is critical to validate the

functional relevance of short protein constructs early on in the NMR investigation, in order to verify whether the selected constructs will provide a suitable representation of functional properties normally exhibited in the context of the intact protein. Specifically, the validation process is aimed at ensuring that any N- and/or C-terminal truncations present in the construct do not significantly perturb binding and allosteric properties relative to the intact protein.

1.5.1. Validation of Binding Affinities Measured by NMR

One means of validating a selected NMR construct is by measuring its binding affinity for a particular ligand(s), and comparing the measured affinity with that determined for the intact protein. In the case of CBDs, this has been accomplished by NMR either *via* saturation transfer difference (STD) measurements if the ligand binding exhibits micromolar (μM) affinity, or *via* differential unfolding free energy ($\Delta\Delta G_{unfolding}$) measurements if the ligand binding exhibits nanomolar (nM) affinity. For example, in an NMR study of the EPAC CBD performed by Mazhab-Jafari *et al.*,[46] an $EPAC1_h$ fragment composed of the N-terminal helical bundle (NTHB), β-subdomain and hinge helix of the $EPAC1_h$ CBD (*i.e.* $EPAC1_h$ residues 149-305) was used to study the effect of cAMP binding on the CBD. Measurement of the dissociation constant (K_d) by STD resulted in a binding isotherm exhibiting a clear dose-response pattern and a well-defined plateau, indicating that the EPAC construct bound cAMP specifically and with an affinity in the μM range, comparable to that reported previously for full-length EPAC1. This observation provided evidence that despite the truncation, the selected EPAC1 construct preserved key cAMP-responsive features present in the full-length protein.[46] This example also illustrates the potential of STD NMR

experiments for μM K_d measurements using the intact cyclic nucleotide without the need of *ad-hoc* added tags that may perturb the binding affinities.

1.5.2. Validation of Allosteric Responses by the NMR Chemical Shift Projection Analysis (CHESPA)

Another critical step of construct validation pertains to verifying that the extent of activation measured by NMR for a given protein fragment in response to selected perturbations correlates with functional changes observed for the full-length protein in response to the same perturbations. For CBDs, such perturbations may include mutations of the protein and/or covalent modifications of the cyclic nucleotide ligand, and the relative extent of activation in the perturbed CBDs is effectively quantified using a recently developed method called "NMR CHEmical Shift Projection Analysis" (CHESPA).[49] The CHESPA approach probes the position of the inactive-*vs.*-active equilibrium of a perturbed CBD relative to the apo and holo (*i.e.* cyclic-nucleotide-bound) wild-type forms of the same CBD construct, which approximate the inactive and active states, respectively (Figure 3). For this purpose, chemical shifts are used to compute through simple vectorial algebra a fractional activation (X) on a per-residue basis for a perturbed CBD. An X value close to 0 (1) indicates that the position of the auto-inhibitory equilibrium of the perturbed CBD is similar to that of the wild-type apo (holo) CBD. An X value ≤ 0 points to inactivation, while $X \geq 1$ points to activation or superactivation, with X values in between 0 and 1 corresponding to partial activation.

The usefulness of the CHESPA approach in the validation of a CBD construct was illustrated in a recent NMR study of the EPAC CBD performed by Selvaratnam *et al.*,[49] in which an $EPAC1_h$ CBD construct (*i.e.* $EPAC1_h$ residues 149-318) was examined by the

CHESPA. The CHESPA was used to compare apo wild-type, cAMP-bound wild-type and apo L273W-mutant forms of the selected EPAC1$_h$ segment.[49] Notably, while cAMP binding promoted a shift of the construct toward its active state, the L273W mutant promoted a shift further toward the inactive state, in agreement with previous bioassay results for full-length EPAC, demonstrating that the L273W mutant promotes inactivation of full-length EPAC.[49] Furthermore, analysis of the cAMP-bound L273W mutant indicated that the apo-state inactivation by the L273W mutant was preserved upon cAMP binding, in agreement with the previously observed inability of cAMP to activate the L273W mutant of full-length EPAC.[49]

1.5.3. Validation of Structural Integrity

The validation process can be further extended to comparative structural assessments, if the structure of the full-length protein is known. For instance, a simple structural validation is obtained by verifying that the secondary structure of a protein fragment is similar to that observed in the integral system. Overall secondary structure content is effectively measured in solution by Circular Dichroism (CD), and a residue-resolution map of secondary structure elements is reliably obtained through the analysis of secondary chemical shifts.[50,51] Once an NMR-amenable construct is validated at the level of binding, allosteric responses and, when possible, structural integrity, further NMR examination of the same construct can then be performed for the purpose of in-depth explorations of dynamics, interactions and their mutual coupling – *i.e.* allostery.

1.6. Methods for Assessing CBD Dynamics and Allostery by NMR

1.6.1. NMR Relaxation Experiments

Direct assessment at atomic resolution of ps-ns and ms-μs structural fluctuations within proteins is possible through the use of classical NMR relaxation techniques.[52,53] In particular, T_1 and T_2 ^{15}N relaxation and {^{1}H-^{15}N}-NOE experiments are frequently used to probe dynamics in the peptide backbone on a per-residue basis and in N-H containing side chains. The ^{15}N T_1, T_2 and {^{1}H-^{15}N}-NOE relaxation data is typically analyzed in terms of either a model-free formalism or reduced spectral density maps to provide information on both ps-ns and ms-μs time-scale dynamics.[52,53] The slower (ms-μs) motions are best analyzed through measurements at multiple static fields, and through nuclear magnetic relaxation dispersion (NMRD) specifically designed to examine ms-μs dynamics arising from the exchange between ground and excited conformations of proteins.[54,55] The ^{15}N relaxation and dispersion methods have been extensively reviewed elsewhere,[52,53,54,55,56,57,58] and additional NMR relaxation experiments are discussed in Section 1.6.4. Thus, here we will focus primarily on how to use these experiments to extract functionally-relevant dynamical profiles of CBDs.

In order to establish correlations between CBD dynamics and CBD function, it is critical to comparatively analyze the dynamical profiles of a given CBD in several forms that span different degrees of functional activation. Thus, NMR relaxation experiments are best used in conjunction with manipulations designed to trap the protein in a particular functional state. For example, in a recent study of the CBD of EPAC, it was possible to separate cAMP-dependent variations in dynamics due to cAMP binding *vs.* allostery by performing a comparative analysis not only of the apo and cAMP-bound CBDs, which approximate the

wild-type inactive and active states, respectively, but also of an antagonist-bound CBD. In this respect, a useful antagonist of cAMP in relation to PKA and EPAC activation is the phosphorothioate cAMP analog R_p-cAMPS. The antagonism of R_p-cAMPS arises from its inactive-*vs.*-active selectivity pattern, which is different from cAMP. While cAMP binds more tightly the active rather than the inactive state (Figure 3), R_p-cAMPS elicits higher affinity for the inactive state of the PKA and EPAC CBDs due to its bulky exocyclic equatorial sulfur atom (Figure 3).[34] Therefore, the R_p-cAMPS antagonist is a useful tool to trap the inactive-bound state of the CBD and dissect contributions to dynamics arising from binding *vs.* purely allosteric effects, which would have otherwise been convoluted together, had the comparative dynamical analyses been limited exclusively to the pair-wise *apo vs.* cAMP-bound forms of the CBD. These comparative analyses used to correlate dynamics to function are easily extended to several NMR methods, which have proven useful for studying CBD structural fluctuations and allostery. Such methods are not limited to NMR relaxation experiments, and include also hydrogen-exchange and chemical shift analyses, which are discussed in Sections 1.6.2–1.6.3 below.

1.6.2. Hydrogen Exchange NMR Experiments: H/D and H/H

The dynamic structural fluctuations that define the native ensemble accessible to a protein often result in transient structural unfolding, which permits hydrogen exchange between the aqueous solvent and the backbone amides of the affected amino acid residues. The higher the population of the state in which a given residue is transiently solvent exposed, the more rapidly hydrogen exchange at that residue will occur. Therefore, measurement of backbone hydrogen-exchange rates provides an additional means of assessing conformational

dynamics within the protein. Hydrogen-exchange is probed with excellent sensitivity through mass-spectrometry, as shown by several elegant applications to the PKA and EPAC systems.[59,60] However, when hydrogen-exchange is monitored by mass-spectrometry, it is not always possible to go beyond peptide-resolution. In these cases, useful complementary techniques to ensure that hydrogen exchange is probed at residue resolution are H/D NMR and H/H NMR. In H/D NMR, hydrogen-exchange rates are measured through real-time monitoring of {^{1}H,^{15}N}-HSQC cross-peak intensity decays resulting from deuterium incorporation into the backbone amides for a sample dissolved in $^{2}H_2O$.[46] However, H/D exchange rates cannot be quantified for residues whose backbone amides undergo full deuterium incorporation within the dead time of the H/D experiment (usually ~20 minutes). To further characterize these fast-exchanging residues, H/H NMR pulse sequences are used to identify residues that undergo hydrogen-exchange on the 10^0-10^2 ms time scale,[46] with remaining residues classified as exchanging on a seconds-to-minutes time scale.

1.6.2.1. Transient Local Unfolding Probed by Hydrogen Exchange NMR

Residues that are predicted to be highly solvent-exposed, for example based on high solvent-accessible surface areas calculated from the protein structure, generally tend to exhibit rapid hydrogen exchange, while residues predicted to be buried within the protein based on low solvent-accessible surface areas, tend to exhibit slower hydrogen exchange. Indeed, upon examination of the apo CBDs of EPAC and PKA, residues of the solvent-exposed N-terminal α-helical elements exhibited rapid exchange, as evidenced by fast H/D exchange and the observation of H/H NMR signals, while residues of the inner strands of the CBD β-barrels exhibited slower exchange, as evidenced by slower H/D exchange and a lack

of H/H NMR signals.[46,47] However, notable exceptions to this general trend have been observed for the CBDs of EPAC and PKA, and such exceptions have highlighted critical allosteric features of these proteins.[46,47] For example, despite the prediction of low solvent exposure for several residues in the phosphate binding cassette (PBC), hinge helix and β2-β3 loop elements of the apo-state EPAC CBD, the apo-state exchange rates were mostly fast for these residues, suggesting that the PBC, hinge helix and β2-β3 loop elements are subject to significant local structural fluctuations in the apo state.[46] Meanwhile, upon cAMP binding, a dramatic decrease in exchange rate was observed for several residues in these elements, highlighting a quenching of local structural fluctuations within these regions.[46] However, these residues still exhibited more rapid exchange than residues in the inner β-barrel strands, and this was attributed to the existence of residual structural fluctuations in the cAMP-bound state.[46]

Additional changes in local structural fluctuations observed in the PKA and EPAC CBDs by H/D and H/H NMR proved to be significant at the level of autoinhibitory interfaces formed by these CBDs. For example, cAMP removal from the PKA CBD resulted in faster hydrogen exchange for several residues throughout the CBD α-helical subdomain (α-subdomain), suggesting an increased solvent exposure within the affected regions of the α-subdomain.[47] Since the α-subdomain regions in question were known to form key components of the autoinhibitory interface with the PKA C-subunit in the inactive state, the increased solvent exposure in these regions upon cAMP removal was postulated to render these regions more available for establishment of interactions with the C-subunit, with cAMP-dependent CBD perturbations hindering formation of the C-subunit interface *via* increased sequestration of key interaction sites in the CBD.[47]

In the case of EPAC, cGMP binding to the EPAC CBD resulted in slower hydrogen exchange for several residues of the CBD NTHB than did cAMP binding, which was attributed to reduced local structural fluctuations in the cGMP-bound state that were also noted from NMR relaxation experiments.[37] Since the NTHB was known to form key autoinhibitory interactions with the catalytic domain, and since an enhancement of NTHB dynamics upon cAMP binding had been previously postulated to promote NTHB/catalytic domain dissociation,[34] the reduced NTHB dynamics in the cGMP-bound state were proposed to hinder entropically-induced NTHB/catalytic domain dissociation, thus contributing to the selectivity of the EPAC CBD for cAMP versus cGMP at the level of activation.[37]

1.6.2.2. Transient Global Unfolding Probed by Hydrogen Exchange NMR

Hydrogen-exchange NMR experiments are not only sensitive to local structural unfolding, but also to transient global unfolding events. For instance, residues in the inner β-barrel strands of both EPAC and PKA exhibited slower hydrogen exchange upon cAMP binding, despite the very low solvent exposure and slow hydrogen exchange already exhibited by these residues in the apo state.[37,46,47] This observation was attributed to the occurrence of transient global unfolding in the apo CBD which was quenched upon cAMP binding, suggesting an overall stabilization of the CBD structure upon cAMP binding.[37,46,47] The identification of global unfolding H/D exchange pathways for the inner strands of the β-barrel of PKA was also independently confirmed by the agreement between the free energy of unfolding estimated from the corresponding H/D protection factors, and that measured through urea unfolding monitored by intrinsic fluorescence.[47]

The ability to measure free energies of global unfolding ($\Delta G_{unfolding}$) by hydrogen-exchange is useful in several applications involving ligand binding, allostery and inter-domain communication. For example, mutation-induced changes in $\Delta G_{unfolding}$ are useful to probe the free energy landscape of allostery through the application of mutant cycles designed to identify and quantify long-range allosteric couplings.[61] Another application of $\Delta G_{unfolding}$ measurements from H/D exchange pertains to inter-domain interactions. For instance, a PKA construct composed of both tandem CBDs of PKA RIα (referred to here as "PKA RIα (119-379)") was examined using a combination of hydrogen exchange NMR and $\{^{1}H,^{15}N\}$ compounded chemical shifts, in order to gain insight into the mechanism of allosteric cross-talk between the two adjacent CBDs.[62] The wild-type PKA RIα (119-379) construct and two mutants (*i.e.* R209K and R333K, with reduced cAMP binding affinity in the CBD-A or CBD-B domain, respectively) were examined in both the presence and absence of excess cAMP. Comparative analysis of the wild-type and mutant constructs revealed previously unknown details of both the intra- and inter-domain allosteric networks, including notable differences between the two domains despite their sequence and structure homology.[62]

In particular, H/D exchange analysis revealed that CBD-A exerts multiple levels of control on CBD-B, as suggested by losses of H/D exchange protection in both CBDs when cAMP binding to the CBD-A is reduced *via* the R209K mutation.[62] Specifically, a loss of H/D exchange protection for highly buried residues in both CBDs suggested that CBD-B is coupled to CBD-A at the level of global unfolding. In addition, a loss of H/D exchange protection for more exposed residues in both CBDs suggested that CBD-B is also coupled to CBD-A at the level of local unfolding, whereby local unfolding of the CBD-A phosphate-

binding cassette (PBC) leads to increased local unfolding not only for the CBD-A C-terminal α-helices, but also for the N-terminal N3A motif and C-terminal α-helices of CBD-B. The observed coupling of CBD-B to CBD-A was rationalized by the existence of a "tandem trans-domain lid" whereby the CBD-B N3A motif, which is structurally and functionally part of the CBD-B, also forms a cAMP-capping lid region for CBD-A, and is covalently linked to the C-terminal α-helical region of CBD-A, thereby establishing a pathway for allosteric signal transmission between the α-helical subdomains of the two CBDs.[62]

Interestingly, while the H/D exchange analysis suggested that global unfolding originating in CBD-A (as triggered by the R209K mutation) promotes global unfolding in CBD-B, global unfolding originating in CBD-B (as triggered by the R333K mutation) did not seem to promote an appreciable global unfolding in CBD-A.[62] Meanwhile, local unfolding at the CBD-B PBC, as prompted by the R333K mutation, appeared to promote local unfolding of the CBD-A C-terminal α-helices as well as the N3A motif and C-terminal α-helices of CBD-B, in accordance with the "tandem trans-domain lid" model, although the CBD-A local unfolding seemed to occur to a lesser degree than the CBD-B local unfolding that was observed for the R209K mutant.[62] The apparent asymmetry in inter-domain communication at the level of global unfolding was concluded to have evolved to ensure robustness in both PKA activation and deactivation, whereby the stability of CBD-A irrespective of CBD-B ensures that PKA activation is only minimally susceptible to dynamic release of apo-state CBD-B from the catalytic subunit, while the instability of CBD-B as a result of CBD-A global unfolding ensures a prompt deactivation of PKA with release of cAMP from both CBDs before new cAMP molecules need to bind to transduce new incoming signals.[62]

1.6.3. Chemical Shift Covariance Analysis (CHESCA)

The fast-exchanging inactive-*vs.*-active conformational equilibria exhibited by CBDs involves long-range allosteric perturbations that are propagated by changes in structure or dynamics, and while the end points of these allosteric perturbations can be effectively characterized by comparative analyses of the structural and dynamical profiles of the apo and cyclic-nucleotide-bound CBD, it is often difficult to define the networks of amino acid residues involved in allosteric signal propagation using such methods.[63] This is especially true if the allosteric signal propagation involves very subtle changes in residues of the allosteric network, as such changes, although potentially of functional relevance, may fall below the resolution of the aforementioned comparative analysis methods.[63]

To identify otherwise elusive allosteric networks of residues, a method based on NMR CHEmical Shift Covariance Analysis (CHESCA) was recently developed and applied to the CBDs of EPAC and PKA.[29,63] In this method, {^{1}H,^{15}N} chemical shifts are measured for a library of five or more states of the protein, which are selected to represent the protein at varying degrees of activation – *i.e.* at different positions along its intrinsic inactive-*vs.*-active equilibrium.[63] For examination of CBDs, such states may include residue mutations and/or ligand-bound states with chemical modifications in the cyclic nucleotide ligand. Ideally, the perturbations within each CHESCA library are chosen to be confined within a limited spatial region, such that the overall extent of chemical shift variations due to nearest-neighbor effects, as opposed to inactive-*vs.*-active equilibrium shifts, is minimized.[63] In addition, careful referencing of chemical shifts is vital to accurate chemical shift measurements, especially in cases of small chemical shift variations, and is achieved by using an internal referencing compound in the samples (*e.g.* ^{15}N-labelled acetylglycine),

coupled with analysis of the {^{1}H,^{15}N}-HSQC cross-peaks *via* Gaussian line-fitting.[63] The measured chemical shifts are then probed for correlated variations among the examined states. Groups of residues that exhibit closely correlated chemical shift variations among the examined states (referred to as residue "clusters") are identified as undergoing a concerted perturbation in response to the selected set of mutations or ligand modifications. The residues within a given cluster are thus identified as belonging to a common allosteric network.[63] The CHESCA clusters are identified through a combination of correlation matrix and principal component statistical analyses of the "residue x perturbation" matrix of chemical shift values.[63]

The CHESCA method was first implemented to identify the network of residues involved in cAMP-associated allosteric and dynamically-driven signal propagation within the CBD of EPAC.[63] Previously, it had been determined that cAMP binding to the EPAC CBD results in a significant structural shift of the cAMP phosphate binding cassette (PBC) and the adjacent C-terminal α-helix (referred to as the "hinge helix"), as well as an enhancement of dynamics at NTHB residues that form key autoinhibitory interactions with the EPAC catalytic region. However, the means by which the allosteric signal propagates from the PBC and hinge helix to the NTHB residues exhibiting dynamic enhancement had remained elusive, as little structural change had been previously noted in regions of the CBD between these sites. Through CHESCA analysis based on the apo state and four ligand-bound states of the CBD (Figure 3), an allosteric network of residues was identified that links the PBC and hinge helix to the NTHB residues exhibiting dynamic enhancement, providing a pathway for allosteric signal propagation between these sites.[63] Specifically, the pathway originates at the PBC (where cAMP binds), and extends through the hinge helix and NTHB C-terminal

helix, ultimately reaching the NTHB N-terminal helices, where the residues exhibiting dynamic enhancement are located.[63] This pathway was independently validated by several known mutants of EPAC, for which the guanine nucleotide exchange rate had been assessed in full-length EPAC.[63]

While the first CHESCA performed on EPAC utilized a series of ligand-bound states as a perturbation library for the statistical analysis,[63] a more recent implementation of CHESCA took advantage of a series of C-terminal deletion mutants of EPAC.[36] Specifically, a series of progressive truncations within the C-terminal portion of the hinge helix were designed to simulate hinge helix C-terminus unwinding, as observed upon cAMP binding and activation. The wild-type and mutant constructs were then analyzed in their apo states to permit examination of the allosteric network controlled by the hinge helix C-terminus in the apo-inactive *vs.* apo-active equilibrium.[36] The results indicated that hinge helix C-terminal truncation leads to a greater stabilization of the active CBD conformation, surprisingly even in the absence of cAMP, and that the perturbation propagates to all known allosteric sites within the CBD, including sites identified in the first CHESCA analysis of EPAC. Therefore, it was concluded that the hinge helix C-terminus plays a key role in EPAC auto-inhibition, and is tightly coupled to the other allosteric elements of the CBD even in the absence of cAMP.[36]

Besides being implemented as a standalone method, CHESCA can also be used in conjunction with other experimental data to generate or test specific hypotheses of interest. For example, the CHESCA analysis of a PKA fragment containing the CBD-A domain and N-terminal linker revealed that the network of residues involved in allosteric signal propagation within the PKA CBD-A domain also includes residues of the linker, suggesting

that the linker, although quite flexible, forms a key component of the CBD-A allosteric network.[29] Based on these results, it was hypothesized that the N-terminal linker selectively interacts with the active rather than the inactive state of the adjacent CBD-A domain. Although consistent with the CHESCA analysis, this hypothesis appeared counter-intuitive and non-obvious based on the inspection of the X-ray structures of the regulatory subunit of PKA bound to the catalytic subunit (inactive state) or to cAMP (active state). In the former structure the linker is visible and in the vicinity of the CBD, while in the latter structure the linker is only partially visible due to its dynamics, and is involved in crystal contacts. Hence, the hypothesis of active-state-selective linker/CBD interactions needed to be tested by additional independent data in solution. For this purpose, paramagnetic relaxation enhancement (PRE) experiments were performed to probe the linker/CBD interactions, and revealed that the linker significantly interacts with the PKA CBD-A only when the latter is in its active conformation, as predicted based on the CHESCA analysis. Furthermore, mutations designed based on the CHESCA results confirmed the existence of active-state-selective interactions of the linker with the CBD-A.[29] Together, these results suggested that the flexible linker is not simply a passive covalent thread, but rather is an active allosteric element involved in the conversion of the PKA regulatory subunit between inhibition-competent and incompetent states.[29]

1.6.4. Probing Dynamics Beyond the Single Isolated CBD Using Methyl-TROSY and Hybrid NMR-MD Approaches

While NMR has yielded valuable information about CBD allostery in the context of isolated CBDs or CBD-containing protein fragments (as illustrated in Sections 1.6.1–1.6.3), such analysis is only a starting point toward gaining a full understanding of CBD protein

function. In particular, analyzing CBDs in the context of the full-length protein, as well as capturing cyclic-nucleotide-dependent variations in structure and/or dynamics in other domains of CBD-containing proteins, are critical aspects of understanding CBD-associated protein function.[59,64] However, as discussed in Section 1.4, examination of full-length CBD-containing proteins by NMR has often proven experimentally challenging. In addition, while the endpoint states of CBD protein allostery (*i.e.* the apo/inactive-conformation and ligand-bound/active-conformation states; Figure 3) are readily examined by NMR, it is generally difficult to probe the transient intermediate states that are critical to allostery, such as the ligand-bound/inactive-conformation and apo/active-conformation states, which serve as intermediates for the classical induced-fit and conformational-selection allosteric pathways, respectively (Figure 3).[64]

One means of examining larger proteins by NMR is to substitute protein-backbone-based NMR methods (*i.e.* {^{1}H,^{15}N}-HSQC, *etc.*) with methyl-TROSY-based NMR experiments. In this approach, the protein to be examined is selectively {^{1}H,^{13}C}-labeled at the side-chain methyl groups of selected amino acids, but fully deuterated elsewhere in the protein, and the labeled methyl groups are probed *via* {^{1}H,^{13}C}-HMQC correlation NMR spectra.[65] Methyl groups are chosen because they yield intense signals due to their three degenerate ^{1}H spins as well as their internal dynamics, and are typically well dispersed in {^{1}H,^{13}C}-correlation spectra.[65] Furthermore, the methyl-based multiple-quantum coherences probed by HMQC benefit from slow transverse relaxation, thanks to the destructive interference between relaxation mechanisms of the ^{1}H and ^{13}C spins ("methyl TROSY"), as well as to the scarcity of ^{1}H spins achieved by deuteration. The combination of methyl TROSY and deuteration dramatically reduces line broadening and enhances both the

resolution and signal-to-noise ratio of the methyl-based NMR spectra.[65] In addition, by probing only the selected side-chain methyl groups, methyl-TROSY-based methods alleviate problems due to peak overlap that often plague protein-backbone-based NMR analyses of large proteins. The usefulness of methyl NMR in the investigation of CBDs is well illustrated by the CAP studies of the Kalodimos group[30,31,32,33] mentioned in Section 1.3.

In some cases it may be desirable to complement methyl NMR studies with examinations of backbone dynamics, which also probe residues devoid of methyl groups, as well as with the examination of transient intermediate states that remain elusive to experimental investigations. In this respect, it is useful to consider an alternative approach to examine dynamics within full-length CBD-containing proteins, which is based on a hybrid methodology combining NMR and molecular dynamics (MD) simulations.[64] In this hybrid NMR-MD approach, a comparative assessment of results obtained from the MD simulations and from NMR is first performed, in order to determine whether key CBD dynamic features observed from experiment can also be captured by the MD simulations. Once the MD simulations have been validated, the MD-based analysis is then extended to examination of dynamics in the full-length protein and/or in the transient intermediate states.

The application of the hybrid NMR-MD approach to CBDs is illustrated by the recent residue-resolution examination of dynamics in a fully-functional EPAC2 construct containing the essential regulatory CBD domain (*i.e.* CBD-B) and full-length catalytic region of EPAC2 (referred to as "EPAC2 (280-990)").[64] To gain insight into the variations of dynamics in full-length EPAC along the thermodynamic cycle of cAMP-dependent activation (Figure 3), MD simulations were performed starting from all four states of the EPAC2 (280-990) construct (apo/inactive, apo/active, holo/inactive and holo/active; Figure

3).[64] The MD simulations were first validated by comparing CBD backbone order parameters for the apo/inactive-conformation and cAMP-bound/active-conformation states, as obtained from both NMR and MD simulations, which indicated that key experimentally-observed trends in CBD dynamics were successfully captured by the MD simulations.[64] More significantly, the comparative simulation analyses not only confirmed the experimentally-observed trends in CBD dynamics, but also revealed unanticipated dynamic attributes of EPAC, thereby rationalizing previously unexplained aspects of EPAC activation and auto-inhibition.[64] A full description of this work and its findings can be found in Chapter 2 of this thesis.

Besides probing dynamics beyond the experimentally-observable CBD construct and/or states, the hybrid NMR-MD approach can also be used for exploring dynamic attributes of the CBD itself that may not be as readily detected by NMR alone. For example, in a more recent study of the EPAC1$_h$ CBD,[49] a combination of MD and NMR comparative mutational analyses were utilized to dissect the determinants for the EPAC apo-inactive *vs.* apo-active auto-inhibitory equilibrium. It was found by NMR that mutating a highly conserved glycine residue (*i.e.* G238) in the CBD β2-3 loop region to alanine induced a partial shift of the apo CBD toward the active conformation, while producing a relatively negligible effect on the cAMP-bound CBD. It was concluded that G238 contributed to autoinhibition of the EPAC1$_h$ CBD, exerting an inhibitory control on all key allosteric sites of the CBD, but it remained unclear from NMR alone how G238 exerted such an effect.[49] Therefore, dynamics within the β2-3 loop and adjacent phosphate binding cassette (PBC) were further probed by comparative analyses of MD simulations performed on the wild-type

apo/inactive-conformation, G238A-mutant apo/inactive-conformation and wild-type cAMP-bound/active-conformation states of the CBD.

The MD simulations were first validated by computing residue-specific secondary structure probabilities for the simulated states and comparing these results with trends observed in NMR experiments.[49] Further analysis of the simulations revealed that a CH-π interaction between the α-carbon of G238 and the guanidinium moiety of PBC residue R279 is tightened by both the G238A mutation and cAMP binding, as reflected by a similar shortening of G238 α-carbon/R279 guanidinium interaction distances in both the G238A-mutant apo/inactive and wild-type cAMP-bound/active states relative to the wild-type apo/inactive state. Furthermore, Procrustean rotation analyses of the MD simulations indicated that the dynamics of the PBC and β2-3 loop relative to one another were quenched by both the G238A mutation and cAMP binding, and NMR-derived backbone order parameters indicated a quenching of dynamics within both the PBC and β2-3 loop. Together, these results suggested that G238 controls EPAC autoinhibition through its CH-π interaction with R279, which in turn modulates the dynamics of the PBC and β2-3 loop, leading to a loss of conformational entropy upon inactive-to-active conformational transition within the apo form.[49] Based on this idea, it was further concluded that the G238A mutation stabilizes the active conformation of the apo state by pre-quenching dynamics within the PBC/β2-3 loop region, thereby reducing the conformational entropy loss that occurs upon inactive-to-active transition, and explaining the partial shift of the apo CBD toward the active conformation that was observed from NMR.[49] Overall, this example illustrates the potential of combined comparative NMR-MD mutational analyses for dissecting the determinants of auto-inhibitory equilibria.

1.6.5. Assessment of Cyclic Nucleotide Conformational Propensities

Besides examining dynamics within the CBD itself, it is also important to consider dynamics within the cyclic nucleotide ligand, which exists in an equilibrium between *syn* and *anti* conformational states (Figure 3). Indeed, given the selectivity of CBDs for one of these two conformations, the *syn*/*anti* conformational tendencies of cyclic nucleotides are expected to play a functionally significant role in CBD activation. Cyclic nucleotide *syn*/*anti* conformational tendencies are typically probed using two-dimensional off-resonance ROESY experiments for the unbound cyclic nucleotide, or two-dimensional transfer-NOESY experiments for the CBD-bound cyclic nucleotide.[37] If the presence of the protein causes background signals that obscure the cyclic nucleotide signals, transfer-NOESY can be combined with isotope filtering methods, which dramatically reduce contributions from isotopically labeled proteins. Once the cyclic nucleotide signals are clearly detectable, the resulting two-dimensional spectra are analyzed to check for the presence of characteristic ^{1}H-^{1}H NOE cross-peaks that arise due to close proximity of specific pairs of cyclic nucleotide hydrogen atoms to one another.[37] For example, the H8 and H1' cyclic nucleotide hydrogen atoms are significantly closer together in the *syn* conformation than in the *anti* conformation, while the H8 and H3' hydrogen atoms are further apart in the *syn* conformation than in the *anti* conformation (Figure 3).[37] Therefore, the *syn* conformation would give rise to a stronger H8/H1' NOE cross-peak and a weaker H8/H3' NOE cross-peak than the *anti* conformation, thus providing an NOE signature that can be used to identify the predominant conformation of a cyclic nucleotide.[37] A challenging aspect of these applications is the overlap between the H8 and H2 ^{1}H NMR peaks for cAMP. A simple but effective way to circumvent this problem is to pre-incubate the cAMP in $^{2}H_2O$, which promotes exchange of the H8 hydrogen

into deuterium. The deuteration-dependent change in the intensity of the H8 peak is then exploited to separate NOE/ROE contributions from H8 *vs.* H2. Using these approaches, it was determined that the EPAC CBD binds cAMP in a *syn* conformation, but binds cGMP in an *anti* conformation, thus revealing a previously unknown feature of EPAC selectivity for cAMP *vs.* cGMP.[37]

1.7. Thesis Outline

In this thesis, the computational and NMR-based methods described in the preceding sections were applied to the cyclic-nucleotide receptor proteins EPAC (Chapter 2), HCN (Chapter 3) and PKG (Chapter 4) to study the cyclic-nucleotide-modulated dynamics of the intrinsic CBDs, and the regulatory effects of these dynamics on other key functional components of the respective proteins. The ultimate goal of this thesis is to identify key differences in the regulatory mechanisms of human cyclic nucleotide receptors, which can eventually be exploited in the development of novel therapeutics to selectively target a single CBD. Such selectivity is critical to modulate cyclic-nucleotide signaling pathways controlled by a given target receptor, while minimizing side effects due to undesired modulation of other receptors. An overview of the main findings of Chapters 2-4 is reported in the following sections. In Chapter 5, we will summarize the key results obtained in this thesis, and highlight emerging common themes regarding the role of dynamics in CBD allostery. We will also outline what we believe are high-priority emerging questions to be addressed in future work on HCN and PKG.

1.7.1. Chapter 2: Role of Dynamics in the Auto-Inhibition and Activation of the Exchange Protein Directly Activated by Cyclic AMP (EPAC) (VanSchouwen et al. J. Biol. Chem., 2011)

The exchange protein directly activated by cAMP (EPAC) is a key receptor of cAMP in eukaryotes and controls critical signaling pathways. Currently, no residue-resolution information is available on the full-length EPAC dynamics, which are known to be pivotal determinants of allostery. In addition, no information is presently available on the intermediates for the classical induced-fit and conformational-selection activation pathways. Here these questions are addressed through molecular dynamics simulations on five key states along the thermodynamic cycle for the cAMP-dependent activation of a fully functional construct of EPAC2, which includes the cAMP-binding domain (CNB) and the integral catalytic region. The simulations are not only validated by the agreement with the experimental trends in CNB dynamics determined by NMR, but they also reveal unanticipated dynamic attributes, rationalizing previously unexplained aspects of EPAC activation and auto-inhibition. Specifically, the simulations show that cAMP binding causes an extensive perturbation of dynamics in the distal catalytic region, assisting the recognition of the Rap1b substrate. In addition, analysis of the activation intermediates points to a possible hybrid mechanism of EPAC allostery incorporating elements of both the induced-fit and conformational-selection models. In this mechanism an entropy compensation strategy results in a low-free-energy pathway of activation. Furthermore, the simulations indicate that the auto-inhibitory interactions of EPAC are more dynamic than previously anticipated, leading to a revised model of auto-inhibition in which dynamics fine-tune the stability of the auto-inhibited state, optimally sensitizing it to cAMP while avoiding constitutive activation.

1.7.2. Chapter 3: Role of Dynamics in the Auto-Inhibition and Activation of the Hyperpolarization-Activated Cyclic-Nucleotide-Modulated (HCN) Ion Channels (VanSchouwen et al. J. Biol. Chem., 2015)

The hyperpolarization-activated cyclic-nucleotide-modulated (HCN) ion channels control rhythmicity in neurons and cardiomyocytes. Cyclic AMP (cAMP) allosterically modulates HCN through the cAMP-dependent formation of a tetrameric gating ring spanning the intracellular region (IR) of HCN, to which cAMP binds. Although the apo *vs.* holo conformational changes of the cAMP-binding domain (CBD) have been previously mapped, only limited information is currently available on the HCN IR dynamics, which have been hypothesized to play a critical role in the cAMP-dependent gating of HCN. Here, using MD simulations validated and complemented by experimental NMR and CD data, we comparatively analyze HCN IR dynamics in the four states of the thermodynamic cycle arising from the coupling between cAMP-binding and tetramerization equilibria. This extensive set of MD trajectories captures the active-to-inactive transition that had remained elusive for other CBDs, and provides unprecedented insight on the role of IR dynamics in HCN auto-inhibition and its release by cAMP. Specifically, the IR tetramerization domain becomes more flexible in the monomeric states, removing steric clashes that the apo CBD structure would otherwise impose. Furthermore, the simulations reveal that the active/inactive structural transition for the apo-monomeric CBD occurs through a manifold of pathways that are more divergent than previously anticipated. Upon cAMP binding, these pathways become disallowed, pre-confining the CBD conformational ensemble to a tetramer-compatible state. This conformational confinement primes the IR for tetramerization, and thus provides a model of how cAMP controls HCN channel gating.

1.7.3. Chapter 4: Mechanism of cAMP Partial Agonism in Protein Kinase G (PKG) (VanSchouwen et al. J. Biol. Chem., 2015)

Protein kinase G (PKG) is a major receptor of cGMP and controls signaling pathways often distinct from those regulated by cAMP. Hence, the selective activation of PKG by cGMP *vs.* cAMP is critical. However, the mechanism of cGMP-*vs.*-cAMP selectivity is only limitedly understood. Although the C-terminal cyclic-nucleotide-binding domain of PKG (CNB-B) binds cGMP with higher affinity than cAMP, the intracellular concentrations of cAMP are typically higher than those of cGMP, suggesting that the cGMP-*vs.*-cAMP selectivity of PKG is not controlled uniquely through affinities. Here, we show that cAMP is a partial agonist for PKG, and we elucidate the mechanism for cAMP partial agonism through the comparative NMR analysis of the apo, cGMP- and cAMP-bound forms of PKG CNB-B. We show that although cGMP-activation is adequately explained by a two-state conformational selection model, the partial agonism of cAMP arises from the sampling of a third, partially autoinhibited state.

1.8. Figures

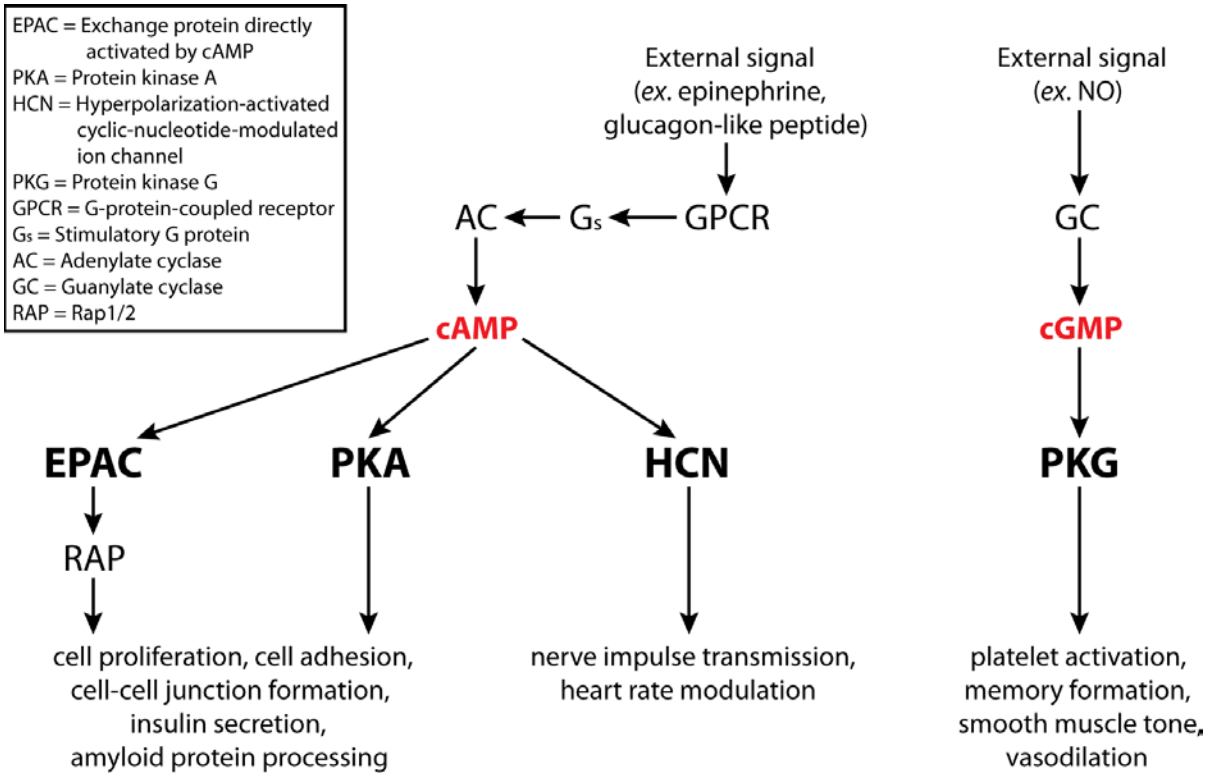

Figure 1: Outline of intracellular signaling pathways propagated by cAMP and cGMP (indicated in red), and their receptors (*i.e.* EPAC, PKA, HCN and PKG). Key elements of the pathways are defined in the figure legend (top left), and physiological processes regulated by each receptor are indicated below the respective receptors.

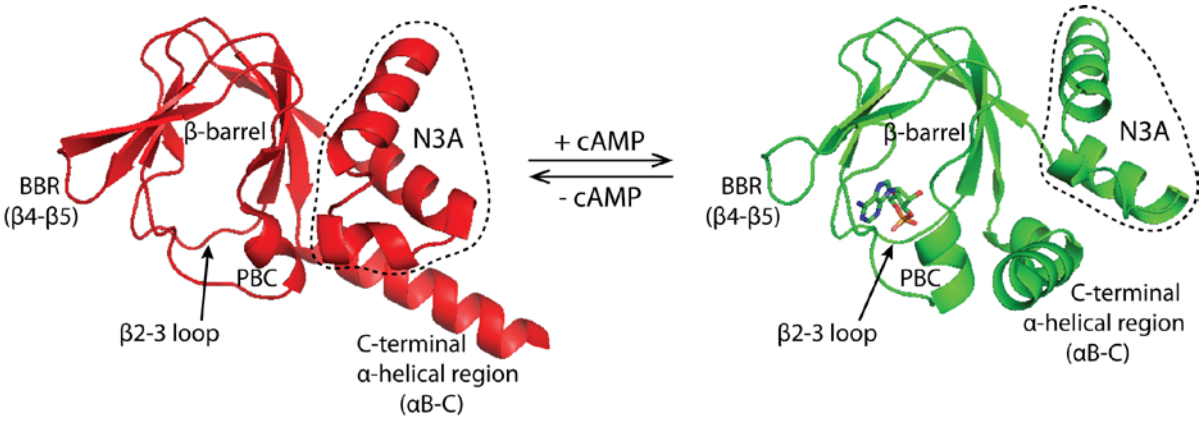

Figure 2: Ribbon-structure illustration of an example of a typical CBD structural rearrangement in response to cyclic nucleotide binding, involving a rearrangement of the N- and C-terminal α-helical components of the CBD. The structures shown represent the cAMP-free (red ribbon) and cAMP-bound (green ribbon) forms of the CBD-A domain of PKA, as previously solved by X-ray crystallography.[10,19] Bound cAMP is shown as sticks, and the following key structural components of the CBD are indicated: the N-terminal α-helix bundle ("N3A", indicated by dotted outlines; referred to as "NTHB" in EPAC);[20,21] the β-barrel (alternatively referred to as the "β-subdomain" or "β-core"), which serves as the central scaffold of the CBD; the base-binding region ("BBR", composed of the β4-β5 segment of the β-barrel) and phosphate-binding cassette ("PBC"), which bind the cyclic nucleotide; the β2-3 loop region adjacent to the PBC; and the C-terminal α-helical segment (αB-C, in PKA).[10,19] The structures were generated using Pymol (Schrödinger, LLC), and to facilitate comparison of the α-helical component arrangements in the two structures, both structures are oriented such that their β-barrels are in the same orientation.

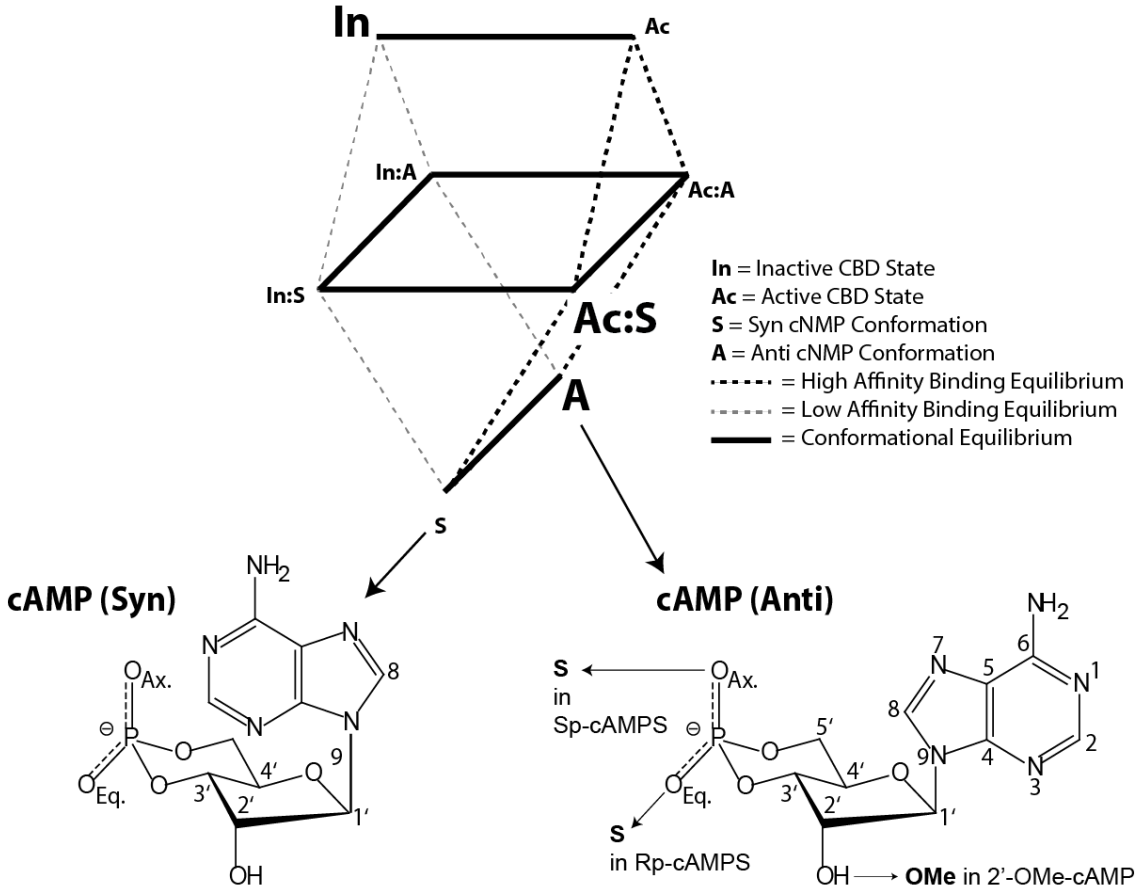

Figure 3: Eight-state thermodynamic cycle for the allosteric control of cyclic-nucleotide binding domains (CBDs) through reciprocal protein-ligand conformational selection. Symbols are defined in the figure legend, and arrows in the equilibria are omitted for simplification purposes. Larger font sizes denote states with higher populations. The apo CBD samples both inactive (auto-inhibitory) and active states, which differ in structure and/or dynamics. In the absence of ligands, the inactive state is typically, although not always (Figure 4), more populated than the active state (top solid line). However, in the holo CBD the reverse applies (square with solid lines), due to the active *vs.* inactive selectivity of the cyclic-nucleotide. Similarly, unbound cAMP samples primarily the *anti* conformation (bottom solid line), while it transitions to the *syn* conformation upon binding to the CBDs of PKA or EPAC. The covalent structure of *anti* cAMP is used to illustrate the library of cAMP analogs originally utilized as a perturbation set for CHEmical Shift Covariance Analysis (CHESCA; Section 1.6.3).

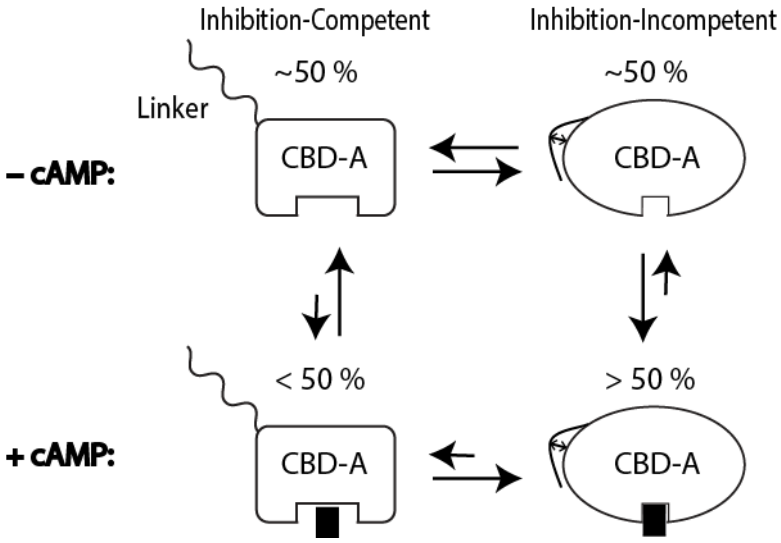

Figure 4: Simplified four-state thermodynamic cycle for the allosteric control of the N-terminal cyclic-nucleotide binding domain of PKA (CBD-A).[29] Single-headed arrows denote conformational or binding equilibria, whereas the double-headed arrow denotes state-selective N-terminal linker/CBD-A interactions. The inhibition-competent state binds the catalytic subunit of PKA with higher affinity than the inhibition-incompetent state. In the absence of cAMP, the two states exhibit nearly degenerate free energies.

1.9. References

[1] Bos, J.L. (2006) *Trends in Biochem. Sci.* **31**: 680-686.
[2] Kawasaki, H., Springett, G. M., Mochizuki, N., Toki, S., Nakaya, M., Matsuda, M., Housman, D. E., and Graybiel, A. M. (1998) *Science* **282**: 2275-2279.
[3] de Rooij, J., Zwartkruis, F. J. T., Verheijen, M. H. G., Cool, R. H., Nijman, S. M. B., Wittinghofer, A., and Bos, J. L. (1998) *Nature* **396**: 474-477.
[4] Biel, M., Wahl-Schott, C., Michalakis, S., and Zong, X. (2009) *Physiol. Rev.* **89**: 847-885.
[5] Craven, K.B., and Zagotta, W.N. (2006) *Annu. Rev. Physiol.* **68**: 375-401.
[6] Kim, J.J., Casteel, D.E., Huang, G., Kwon, T.H., Ren, R.K., Zwart, P., Headd, J.J., Brown, N.G., Chow, D.C., Palzkill, T., and Choel, K. (2011) *PLoS ONE* **6**: e18413(1-15).
[7] Osborne, B.W., Wu, J., McFarland, C.J., Nickl, C.K., Sankaran, B., Casteel, D.E., Woods, V.L., Kornev, A.P., Taylor, S.S., and Dostmann, W.R. (2011) *Structure* **19**: 1317-1327.
[8] Kim, C., Vigil, D., Anand, G., and Taylor, S.S. (2006) *Eur. J. Cell Biol.* **85**: 651-654.
[9] Rehmann, H., Wittinghofer, A., and Bos, J.L. (2007) *Nat. Rev. Mol. Cell Biol.* **8**: 63-73.
[10] Kim, C., Cheng, C.Y., Saldanah, S.A., and Taylor, S.S. (2007) *Cell* **130**: 1032-1043.
[11] Wang, Z., Dillon, T.J., Pokala, V., Mishra, S., Labudda, K., Hunter, B., and Stork, P.J. (2006) *Mol. Cell Biol.* **26**: 2130-2145.
[12] Huang, G.Y., Kim, J.J., Reger, A.S., Lorenz, R., Moon, E.W., Zhao, C., Casteel, D.E., Bertinetti, D., VanSchouwen, B., Selvaratnam, R., Pflugrath, J.W., Sankaran, B., Melacini, G., Herberg, F.W., and Kim, C. (2014) *Structure* **22**: 116-124.
[13] Alverdi, V., Mazon, H., Versluis, C., Hemrika, W., Esposito, G., van den Heuvel, R., Scholten, A., and Heck, A.J. (2008) *J. Mol. Biol.* **375**: 1380–1393.
[14] Taylor, M.S., Okwuchukwuasanya, C., Nickl, C.K., Tegge, W., Brayden, J.E., and Dostmann, W.R.G. (2004) *Mol. Pharmacol.* **65**: 1111-1119.
[15] Ibe, B.O., Ameer, A., Portugal, A.M., Renteria, L., and Raj, J.U. (2007) *J. Pharmacol. Exp. Ther.* **320**: 728-737.
[16] Liu, S., Rao, Y., and Daw, N. (2003) *Cereb. Cortex* **13**: 864-869.
[17] Guo, D.C., Regalado, E., Casteel, D.E., Santos-Cortez, R.L., Gong, L., Kim, J.J., Dyack, S., Horne, S.G., Chang, G., Jondeau, G., Boileau, C., Coselli, J.S., Li, Z., Leal, S.M., Shendure, J., Rieder, M.J., Bamshad, M.J., and Nickerson, D.A. (2013) *Am. J. Hum. Genet.* **93**: 398–404.
[18] Pfeifer, A., Klatt, P., Massberg, S., Ny, L., Sausbier, M., Hirneiß, C., Wang, G.X., Korth, M., Aszódi, A., Andersson, K.E., Krombach, F., Mayerhofer, A., Ruth, P., Fässler, R., and Hofmann, F. (1998) *EMBO J.* **17**: 3045–3051.
[19] Su, Y., Dostmann, W.R., Herberg, F.W., Durick, K., Xuong, N.H., Ten-Eyck, L., Taylor, S.S., and Varughese, K.I. (1995) *Science* **269**: 807-813.
[20] Rehmann, H., Das, J., Knipscheer, P., Wittinghofer, A., and Bos, J.L. (2006) *Nature* **439**: 625-628.
[21] Rehmann, H., Arias-Palomo, E., Hadders, M.A., Schwede, F., Llorca, O., and Bos, J.L. (2008) *Nature* **455**: 124-127.
[22] Zagotta, W.N., Olivier, N.B., Black, K.D., Young, E.C., Olson, R., and Gouaux, E. (2003) *Nature* **425**: 200-205.
[23] Xu, X., Vysotskaya, Z.V., Liu, Q., and Zhou, L. (2010) *J. Biol. Chem.* **285**: 37082-37091.

[24] Gunasekaran, K., Ma, B., and Nussinov, R. (2004) *Proteins: Struct., Funct., Bioinf.* **57**: 433-443.
[25] Nussinov, R., and Tsai, C.J. (2013) *Cell* **153**: 293–305.
[26] Nussinov, R., Tsai, C.J., and Ma, B. (2013) *Annu. Rev. Biophys.* **42**: 169-189.
[27] Ma, B., Shatsky, M., Wolfson, H.J., and Nussinov, R. (2002) *Protein Sci.* **11**: 184-197.
[28] Andrusier, N., Mashiach, E., Nussinov, R., and Wolfson, H.J. (2008) *Proteins: Struct., Funct., Bioinf.* **73**: 271-289.
[29] Akimoto, M., Selvaratnam, R., McNicholl, E.T., Verma, G., Taylor, S.S., and Melacini, G. (2013) *Proc. Natl. Acad. Sci. USA* **110**: 14231-14236.
[30] Popovych, N., Sun, S., Ebright, R.H., and Kalodimos, C.G. (2006) *Nat. Struct. Mol. Biol.* **13**: 831-838.
[31] Tzeng, S.R., and Kalodimos, C.G. (2009) *Nature* **462**: 368-372.
[32] Tzeng, S.R., and Kalodimos, C.G. (2013) *Nat. Chem. Biol.* **9**: 462-465.
[33] Tzeng, S.R., and Kalodimos, C.G. (2012) *Nature* **488**: 236-240.
[34] Das, R., Mazhab-Jafari, M.T., Chowdhury, S., Das, S.S., Selvaratnam, R., and Melacini, G. (2008) *J. Biol. Chem.* **283**: 19691-19703.
[35] Boulton, S., Akimoto, M., VanSchouwen, B., Moleschi, K., Selvaratnam, R., and Melacini, G. (2014) *Biochem. Soc. Trans.* **42**: 302-307.
[36] Selvaratnam, R., Mazhab-Jafari, M.T., Das, R., and Melacini, G. (2012) *PLoS ONE* **7**: e48707(1-9).
[37] Das, R., Chowdhury, S., Mazhab-Jafari, M.T., SilDas, S., Selvaratnam, R., and Melacini, G. (2009) *J. Biol. Chem.* **284**: 23682-23696.
[38] Goto, N.K., and Kay, L.E. (2000) *Curr. Opin. Struct. Biol.* **10**: 585-592.
[39] Tugarinov, V., Hwang, P.M., and Kay, L.E. (2004) *Annu. Rev. Biochem.* **73**: 107-146.
[40] Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M., and Clore, G.M. (1989) *Biochemistry* **28**: 6150-6156.
[41] Ikura, M., Kay, L.E., and Bax, A. (1990) *Biochemistry* **29**: 4659-4667.
[42] Gardner, K.H., and Kay, L.E. (1998) *Annu. Rev. Biophys. Biomol. Struct.* **27**: 357-406.
[43] Craven, K.B., and Zagotta, W.N. (2006) *Annu. Rev. Physiol.* **68**: 375-401.
[44] Zagotta, W.N., Olivier, N.B., Black, K.D., Young, E.C., Olson, R., and Gouaux, E. (2003) *Nature* **425**: 200-205.
[45] Biel, M., Wahl-Schott, C., Michalakis, S., and Zong, X. (2009) *Physiol. Rev.* **89**: 847-885.
[46] Mazhab-Jafari, M.T., Das, R., Fotheringham, S.A., SilDas, S., Chowdhury, S., and Melacini, G. (2007) *J. Am. Chem. Soc.* **129**: 14482-14492.
[47] Das, R., Esposito, V., Abu-Abed, M., Anand, G.S., Taylor, S.S., and Melacini, G. (2007) *Proc. Natl. Acad. Sci. USA* **104**: 93-98.
[48] Huang, G., Kim, J.J., Reger, A.S., Lorenz, R., Moon, E.W., Zhao, C., Casteel, D.E., Bertinetti, D., VanSchouwen, B., Selvaratnam, R., Pflugrath, J.W., Sankaran, B., Melacini, G., Herberg, F.W., and Kim, C. (2014) *Structure* **22**: 116-124.
[49] Selvaratnam, R., VanSchouwen, B., Fogolari, F., Mazhab-Jafari, M.T., Das, R., and Melacini, G. (2012) *Biophys. J.* **102**: 630-639.
[50] Gillard, R.D. (1963) *Analyst* **88**: 825-828.
[51] Eghbalnia, H.R., Wang, L., Bahrami, A., Assadi, A., and Markley, J.L. (2005) *J. Biomol. NMR* **32**: 71-81.

[52] Palmer, A.G. (2004) *Chem. Rev.* **104**: 3623-3640.
[53] Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D., and Kay, L.E. (1994) *Biochemistry* **33**: 5984-6003.
[54] Kleckner, I.R., and Foster, M.P. (2010) *Biochim. Biophys. Acta* **1814**: 942-968.
[55] Korzhnev, D.M., and Kay, L.E. (2008) *Acc. Chem. Res.* **41**: 442-451.
[56] Kay, L.E., Nicholson, L.A., Delaglio, F., Bax, A., and Torchia, D.A. (1992) *J. Magn. Reson.* **97**: 359-375.
[57] Kay, L.E. (1998) *Biochem. Cell Biol.* **76**: 145-152.
[58] Hansen, D.F., Vallurupalli, P., and Kay, L.E. (2008) *J. Phys. Chem. B* **112**: 5898-5904.
[59] Li, S., Tsalkova, T., White, M.A., Mei, F.C., Liu, T., Wang, D., Woods, V.L., and Cheng, X. (2011) *J. Biol. Chem.* **286**: 17889-17897.
[60] Hamuro, Y., Anand, G.S., Kim, J.S., Juliano, C., Stranz, D.D., Taylor, S.S., and Woods, V.L. (2004) *J. Mol. Biol.* **340**: 1185-1196.
[61] Boyer, J.A., Clay, C.J., Luce, K.S., Edgell, M.H., and Lee, A.L. (2010) *J. Am. Chem. Soc.* **132**: 8010-8019.
[62] McNicholl, E.T., Das, R., SilDas, S., Taylor, S.S., and Melacini, G. (2010) *J. Biol. Chem.* **285**: 15523-15537.
[63] Selvaratnam, R., Chowdhury, S., VanSchouwen, B., and Melacini, G. (2011) *Proc. Natl. Acad. Sci. USA* **108**: 6133-6138.
[64] VanSchouwen, B., Selvaratnam, R., Fogolari, F., and Melacini, G. (2011) *J. Biol. Chem.* **286**: 42655-42669.
[65] Ollerenshaw, J.E., Tugarinov, V., and Kay, L.E. (2003) *Magn. Reson. Chem.* **41**: 843-852.

CHAPTER 2

ROLE OF DYNAMICS IN THE AUTO-INHIBITION AND ACTIVATION OF THE EXCHANGE PROTEIN DIRECTLY ACTIVATED BY CYCLIC AMP (EPAC)

2.1. Author's Preface

The work presented in this chapter has previously been published and is reproduced here with permission from the Journal of Biological Chemistry. The full citation is as follows:

VanSchouwen, B., Selvaratnam, R., Fogolari, F. and Melacini, G. (2011) "Role of dynamics in the auto-inhibition and activation of the exchange protein directly activated by cyclic AMP (EPAC)." *Journal of Biological Chemistry* **286**: 42655-42669.

I conducted most of the work necessary for this publication. Dr. Rajeevan Selvaratnam performed the experimental EPAC work that was used to validate the molecular dynamics data. Dr. Federico Fogolari performed the principal component analysis work. I co-wrote the manuscript with Drs. Federico Fogolari and Giuseppe Melacini.

2.2. Abstract

The exchange protein directly activated by cAMP (EPAC) is a key receptor of cAMP in eukaryotes and controls critical signaling pathways. Currently, no residue-resolution information is available on the full-length EPAC dynamics, which are known to be pivotal determinants of allostery. In addition, no information is presently available on the intermediates for the classical induced-fit and conformational-selection activation pathways. Here these questions are addressed through molecular dynamics simulations on five key states along the thermodynamic cycle for the cAMP-dependent activation of a fully functional construct of EPAC2, which includes the cAMP-binding domain (CNB) and the

integral catalytic region. The simulations are not only validated by the agreement with the experimental trends in CNB dynamics determined by NMR, but they also reveal unanticipated dynamic attributes, rationalizing previously unexplained aspects of EPAC activation and auto-inhibition. Specifically, the simulations show that cAMP binding causes an extensive perturbation of dynamics in the distal catalytic region, assisting the recognition of the Rap1b substrate. In addition, analysis of the activation intermediates points to a possible hybrid mechanism of EPAC allostery incorporating elements of both the induced-fit and conformational-selection models. In this mechanism an entropy compensation strategy results in a low-free-energy pathway of activation. Furthermore, the simulations indicate that the auto-inhibitory interactions of EPAC are more dynamic than previously anticipated, leading to a revised model of auto-inhibition in which dynamics fine-tune the stability of the auto-inhibited state, optimally sensitizing it to cAMP while avoiding constitutive activation.

2.3. Introduction

The exchange protein directly activated by cyclic AMP (EPAC) is a major protein involved in intracellular G-protein-mediated signaling in eukaryotic organisms, playing a regulatory role in such processes as cell adhesion, cell-cell junction formation and insulin secretion by pancreatic β-cells.[1,2,3,4,5,6,7,8,9,10] Two isoforms of EPAC have been identified – EPAC1 and EPAC2 – both composed of an amino-terminal regulatory region (RR), and a carboxy-terminal catalytic region (CR; Fig. 1a).[11,12] The regulatory region consists of a Dishevelled Egl-10 Pleckstrin (DEP) domain[13] and one (for EPAC1) or two (for EPAC2) cyclic-nucleotide-binding (CNB) domains (Figs. 1 and S1).[11,12,14,15] The catalytic region consists of the CDC25-homology domain (CDC25-HD) responsible for guanine nucleotide

exchange activity, as well as a Ras exchange motif (REM) domain and a Ras association (RA) domain involved in the sub-cellular localization of EPAC (Fig. 1a).[11,12,14,15]

The structures of EPAC solved in the absence and presence of a non-hydrolysable agonist closely related to cAMP have shown that EPAC adopts two main conformations, which differ with respect to the relative RR/CR orientation (Fig. 1b).[14,15] In the so-called "closed" conformation the RR is in the vicinity of the catalytic domain, blocking access to the Rap1b substrate and effectively inhibiting the guanine nucleotide exchange activity.[14] In the so-called "open" conformation a hinge helix rotation displaces the RR away from the CR, making the catalytic site accessible to the substrate and leading to activation (Fig. 1b).[15] In the absence of cAMP, EPAC populates mainly the closed inactive conformation, which is stabilized by two key sets of interactions commonly referred to as hydrophobic hinge (HH) and ionic latch (IL).[14,15,16,17] The HH involves the phosphate binding cassette (PBC), which has been proposed to sterically block conformational changes in the hinge helix, maintaining this helix in its inactive conformation (Fig. S1). Specifically, it has been hypothesized that the inactive conformation is stabilized by steric hindrance between a leucine side chain from the PBC (L408 in EPAC2) and a phenylalanine side chain from the hinge (F435 in EPAC2).[14,16] The IL involves the CNB and the CDC25-HD domains, which interact through a series of key salt bridges between an N-terminal helical bundle (NTHB) in the CNB domain and the catalytic site of the CDC25-HD domain.[14] In particular, these IL interactions involve the side chains of a glutamine (Q303 in EPAC2) and aspartate (D307 in EPAC2) residue from the DEP-CNB connecting helix ("α1" of the NTHB), a glutamate residue (E332 in EPAC2) from the second NTHB helix ("α2"), and an aspartate (D883 in EPAC2) and arginine (R886 in EPAC2) residue from the CDC25-HD domain catalytic site (Fig. 1b).[14]

When cAMP binds to the CNB domain, a number of changes occur which perturb the HH and IL auto-inhibitory interactions and allow EPAC to adopt its active conformation. First, the interaction of the bound cAMP with the PBC repositions L408, relieving the steric hindrance between the PBC and the hinge helix and allowing the activation-associated hinge rotation to proceed.[15,16] The second major change induced by cAMP binding is a breaking of the ionic latch interactions between the NTHB of the CNB domain and the catalytic site of the CDC25-HD, assisting the displacement of the EPAC regulatory region away from the CDC25-HD *via* the hinge rotation.[14,15] Notably, NMR relaxation experiments have shown that upon cAMP binding, an enhancement of conformational dynamics occurs in the NTHB, including the second and third residues involved in the ionic latch (D307 and E332 in EPAC2), while no major structural changes appeared to occur in the helices that span the ionic latch (*i.e.* α1 and α2).[18] Therefore, cAMP binding triggers an increase in conformational entropy of the ionic latch region, thereby increasing the entropic penalty for preserving the ionic latch, and consequently weakening the ionic-latch-mediated interactions between the CNB and CDC25-HD domains.[18]

Besides causing the HH rearrangement and the IL weakening, cAMP forms through its adenine additional interactions with the base binding region (BBR) of the CNB and the so-called lid region (Fig. S1), thereby stabilizing EPAC into its active conformation.[15,16] The active structure of EPAC is further stabilized by intramolecular interactions that are formed among the CNB and REM domains and the lid, including interactions of PBC residue K405 with lid residues E443, N445 and Y480, and interactions of CNB residues Q369 and Y389 with REM residue Y551.[15] Overall, it is clear that the experimental characterizations of the apo/inactive and holo/active states of EPAC have unveiled key attributes of the mechanisms

of EPAC auto-inhibition and cAMP-dependent activation.[14,15,16,17,18,19,20,21] However, several critical questions remain open. For instance, while NMR has shown that modulations of dynamics by cAMP represent a pivotal feature of EPAC activation and cyclic nucleotide selectivity,[17,18,22,23] the NMR studies have been confined only to the CNB domain. Although dynamics elsewhere in the protein is expected to be critical for allostery,[24,25,26,27,28] no atomic or residue resolution information is currently available on the dynamics of full-length EPAC constructs with integral RR and CR domains. Without a full map of the dynamic profile of EPAC, key aspects of the EPAC function remain unexplained.

Another limitation of the current experimental investigations on EPAC is that they are to a large extent limited to the apo/inactive and holo/active states due to the effective coupling between cAMP binding and the closed-to-open conformational transition. As a result, no information is currently available on the apo/active and holo/inactive cross-states ("metastates"), which represent key intermediates in the thermodynamic cycle that models the coupling between cAMP binding and allosteric conformational changes (Fig. 1b). Specifically, the apo/active metastate is critical when activation proceeds through conformational selection, while the holo/inactive metastate is central when activation proceeds through an induced fit mechanism.[26,29,30,31,32,33] The characterization of the metastates is also essential to dissect the distinct contributions of cAMP binding and of conformational changes to the variations in dynamics between the apo-inactive and the holo-active states.

For the purpose of gaining insight into the dynamics of full length EPAC in all four states of the thermodynamic cycle of EPAC activation (Fig. 1b), we present here molecular dynamics (MD) simulations[34,35,36,37,38] on a functional EPAC2 construct that includes both

the regulatory CNB and the integral catalytic region. The simulations are in agreement with the available NMR data, but they also reveal several ps-ns dynamic features of EPAC allostery that were not anticipated from previous experiments. These features include a dynamic control of critical auto-inhibitory interactions as well as an extensive modulation of both regulatory-region (RR) and catalytic-region (CR) dynamics by both cAMP and Rap1b binding, pointing to the existence of a long-range dynamic RR/CR cross-talk. In addition, the MD simulations reveal a possible low free-energy pathway for the reversible and cAMP-dependent activation of EPAC through a hybrid induced-fit/conformational-selection mechanism.

2.4. Experimental Procedures

Molecular dynamics (MD) simulations in explicit solvent were performed starting from the four states of the thermodynamic cycle for the cAMP-dependent EPAC allostery (Fig. 1b), *i.e.* the experimentally observable apo/inactive and holo/active states as well as the intermediate holo/inactive and apo/active states.[18] The latter two states are unstable under physiological conditions and thus are referred to herein as "metastates" to distinguish them from the experimentally-observable apo/inactive and holo/active states, for which structures have been solved by X-ray crystallography.[14,15] Furthermore, an additional MD trajectory was generated starting from the holo/active state bound to the Rap1b substrate. All simulations were performed on the EPAC2(280-990) fragment (Fig. 1b). This segment contains the essential regulatory CNB domain and the full catalytic region (*i.e.* the REM, RA and CDC25-HD domains).[14,15] The EPAC2(280-990) construct was previously shown to be

fully functional as it exhibited cAMP-dependent GEF activity comparable to that of wild-type EPAC2.[14]

A summary of the MD simulations performed for the three states and the two metastates of EPAC2(280-990) is given in Table S1. The structural dynamics during the last 50 ns of each MD trajectory were examined through multiple and complementary analyses, including root-mean-square deviations (RMSDs; Fig. 2), contact distance time profiles (Figs. S4 and S7), backbone N-H order parameters (S^2; Figs. 3 and 4), dynamic cross-correlation matrices (DCCMs; Fig. 5), differential root-mean-square fluctuations (RMSFs; Figs. 6 and S3), covariance matrices (Fig. S5) and principal component analyses (PCA; Fig. S6).

Each type of analysis examined key aspects of EPAC dynamics: RMSDs from the initial structures were performed to probe the stability of the overall domain topology of EPAC (*i.e.* closed *vs.* open; Fig. 1b); contact distance time profiles were utilized to assess the dynamics of key inter-domain interactions that selectively stabilize either the open or closed topology; backbone N-H order parameters (S^2) were used to probe the amplitudes of local dynamics at residue-resolution. Local dynamics are in principle also probed by RMSFs, provided that the structural superimposition required by RMSF analysis is confined to a single domain in order to minimize contributions from long-range inter-domain motions. The resulting domain-specific RMSF profiles provide a picture of the amplitudes of local fluctuations that is qualitatively similar to the S^2 plots (Fig. S2), but the S^2 analysis benefits from the added advantage of facilitating validation of the MD results against the experimental NMR data available for the CNB domain.[18] Further details about the S^2 *vs.* RMSF comparisons are available as Supplementary Material. Correlations between dynamic fluctuations at different locations, either intra- or inter-domain, were quantified through

DCCM matrices. However, due to the inherent normalization in the DCCM computations, such matrices are not sensitive to the amplitudes of the motions. We therefore complemented the DCCM analysis with covariance and principal component analyses (PCA; Figs. S5 and S6), as well as with differential RMSFs (Figs. 6 and S3). The differential RMSFs refer to the difference between the RMSF values obtained by superimposing all domains and the domain-specific RMSFs. While the former RMSF values reflect both local fluctuations and non-local domain motions, the latter sense mostly local dynamics. Therefore, the differential RMSFs report on the amplitudes of non-local domain motions and are a useful complement to the analysis of inter-domain correlations as identified by the DCCM matrices.

Details about the preparation of the initial structures as well as about the MD simulation protocols are provided in the Supplementary Material. The detailed protocols utilized for the validation and the analysis of the MD simulations are also provided in the Supplementary Material.

2.5. Results

2.5.1. Dynamics of the Apo/Inactive State of EPAC2

The closed-topology of the apo/inactive state and the fold of each individual domain are stable in the ns timescale. As a first assessment of the simulated MD trajectory starting from the apo/inactive structure, the root-mean-square deviations (RMSDs) from both active and inactive initial EPAC2 conformations were computed over the course of the simulation (Fig. 2a). It was found that the RMSD from the inactive-state initial structure increased over the 10-ns initial equilibration period (Fig. 2a, negative time), before settling into fluctuations

within a range of 3-6 Å. These RMSD values are significantly lower than the RMSD between the inactive- and active-state initial structures (~24 Å), indicating that no active/inactive transition occurs during the time span of the simulation. This interpretation is supported by the observation that the apo/inactive-state simulation consistently displayed a ~24 Å RMSD from the active-state structure (Fig. 2a). It is therefore clear that the MD trajectory remains closer to its initial "closed" structure than to the "open" conformation, without a significant change in overall topology. In addition, the fold of each individual domain of the apo/inactive state of EPAC2 does not change significantly during the course of the simulations, as shown by the domain-specific RMSDs from their respective initial structures (Fig. 2a). The domain-specific RMSD values all reached stable maxima of ~3 Å or less early in the initial equilibration period of the simulation (Fig. 2a, negative time), and remained at/near the maxima throughout the rest of the simulation time (Fig. 2a, positive time). Overall, the structures of the individual domains appear to be quite stable. However, the domain-specific RMSDs do not provide insight into the local dynamics, which are better assessed by backbone N-H order parameters.

Patterns of local dynamics in apo/inactive EPAC2 and experimental validation. The backbone N-H order parameters (S^2) for EPAC2 in the apo/inactive state are shown in Fig. 3a and reveal several "hot spots" of local dynamics in each domain. In the CNB domain, the most dynamic regions of the β-core include the PBC, the BBR and the β2-3 loop (Fig. 3a), in agreement with previous NMR relaxation studies.[18] In addition, an appreciable degree of flexibility was evident in the first α-helix of the CNB domain (NTHB α1; Fig. 3a), again in agreement with the NMR data[18]. The extent of the agreement between the NMR data and the

MD results is further appreciated by inspecting Fig. 4a, which compares the backbone amide order parameters (S^2) measured by NMR for the CNB domain of apo EPAC1 versus those computed by MD for the corresponding residues of the homologous EPAC2 in the apo/inactive state. Figure 4a shows that to a large extent the MD simulations capture the qualitative trends of the experimental order parameters. Nine out of 12 local S^2 minima are well reproduced by the MD trajectories (Fig. 4a, black arrows). Some deviations between experimental and calculated S^2 values are apparent especially at the N- and C-termini (Fig. 4a), which is expected considering that the NMR data were acquired on a truncated construct of a different EPAC isoform. No NMR data is currently available on the EPAC2 isoform.

A particularly notable dynamic "hot spot" was observed in the region of the NTHB α1 helix that contains two of the three CNB domain residues involved in the ionic-latch interface between the CNB and CDC25-HD domains (*i.e.* residues Q303 and D307; Figs. 1a,b and S1), suggesting that the ionic latch interface of the apo/inactive state is more dynamic than expected based on the X-ray structure alone.[14] In order to gain further insight into the dynamics of the ionic latch, we examined the time profiles of the ionic-latch distances for the apo/inactive state (Fig. S4). It was found that most of the ionic-latch interactions were indeed very dynamic, as reflected by large and highly-variable contact distances (Fig. S4). In particular, the interaction of Q303 with D883 was highly dynamic, with distances mostly greater than 20 Å (Fig. S4a), while the interactions of Q303 and D307 with R886 were noticeably more stable, but still very dynamic (Fig. S4a,b). The interactions between E332 and R886, on the other hand, consisted of interactions that were considerably more stable than those among any of the other ionic-latch residue pairs (Fig. S4c,d).

Another pattern observed in the order parameter (S^2) profiles shown in Fig. 3a is that regions of lower dynamics were generally clustered in the interior of the protein, while regions of higher dynamics resided primarily at the solvent-exposed exterior. Indeed, such a trend is typical for soluble proteins, and as expected, the most dynamic regions were generally within or near loops (Fig. 3a). For instance, the loops composed of residues 465-476 and 725-731 (Fig. 3a, cyan asterixes) displayed particularly low S^2 values, in agreement with the missing electron density for these loops in the X-ray structure.[14] In addition, relatively low S^2 values for the RA domain (Fig. 3a and Table 1) indicate that the RA module is the most dynamic domain in the catalytic region, in agreement with the poor intra-domain packing and high B-factors previously observed for this domain.[14] Overall, the dynamic profile predicted by the MD simulations for the apo/inactive state of EPAC is well corroborated by the available experimental data. In addition, MD simulations also provide a glimpse into dynamic features of EPAC that are not experimentally accessible, such as those revealed by the DCCM matrix, which probes both intra- and inter-domain correlations.

Intra-domain cross-correlations of apo/inactive EPAC2. Upon examination of the DCCM results for the apo/inactive state (Fig. 5a, lower triangle), it was found that the domains in the catalytic region (*i.e.* REM, RA and CDC25-HD) all exhibited internal correlations that were primarily positive in value. However, the regulatory CNB domain demonstrated several anti-correlated motions between its α-helical and β-sheet subdomains, even though the β-sheet subdomain exhibited internal correlations that were entirely positive in value (Fig. 5a, lower triangle).

Inter-domain cross-correlations of apo/inactive EPAC2. The examination of inter-domain DCCM correlations (Fig. 5a, lower triangle) revealed two key features. First, most of the inter-domain correlations observed in the apo/inactive state were weak or negative in value (Fig. 5a, lower triangle). For instance, it was unexpectedly found that the CNB domain demonstrated largely anti-correlated motion relative to the CDC25-HD domain, despite the fact that these two domains are anchored to each other by the ionic latch (IL) salt bridges (Fig. 5a, lower triangle). This observation is consistent with the dynamic nature of the ionic latch interactions revealed by the time-profiles of the IL distances. Second, the DCCM analysis revealed that the main positive correlations involve residues that either form or are closely associated with the so-called switchboard β-sheet at the boundary between the regulatory and catalytic regions, *i.e.* CNB domain residues 445-461, REM domain residues 480-500 and residues 910-940 from the CDC25-HD domain helical hairpin (Fig. 5a, lower triangle; Fig. 5c). The switchboard site therefore appears to function as a stable interaction-hub that anchors the CNB, REM and CDC25-HD domains to one another.[14] Further insight into the inter-domain motions is provided by the differential RMSF analysis (Figs. 6a and S3a).

The differential RMSF profiles (Figs. 6a and S3a) indicate that the CNB and RA domains are subject to the largest amplitude motions in the apo/inactive state. We therefore anticipate that the anti-correlated motion between the CNB and RA domains revealed by the DCCM analysis (Fig. 5a, lower triangle) dominates the total motional variance. This result is fully confirmed by the examination of inter-domain dynamics through PCA analysis of the covariance matrix (Fig. S5a). Specifically, the extreme projections of the apo/inactive MD trajectory on the first PCA eigenvector (Fig. S6a) indicate a “breathing” motion between the

CNB and RA domains, involving correlated motions of smaller amplitude in the CDC25-HD domain.

2.5.2. Dynamics of the Holo/Active State of EPAC2

The open-topology of the holo/active state and the fold of each individual domain are stable in the ns timescale. Analogously to the apo/inactive state, RMSDs from both active and inactive initial EPAC2 conformations suggested that the holo/active state MD trajectory remains closer to its initial "open" conformation than to the "closed" structure, without a significant change in overall topology. This was reflected by considerably larger RMSDs from the inactive-state initial structure (20-25 Å) than from the active-state initial structure (4-8 Å) over the course of the simulation (Fig. 2d), indicating that no active/inactive transition occurs during the 50 ns time span of the simulation. In addition, the structures of the individual domains appeared to be quite stable, as reflected by lower and less variable RMSDs (~2.5 Å or less) of the individual modules from their respective initial structures during the simulation (Fig. 2d). Notably, the CDC25-HD domain does not deviate significantly from its initial structure despite the absence of the bound Rap1b (Fig. 2d). Overall, the considerable difference between the total and the domain-specific RMSD values points to the presence of enhanced inter-domain motions in the holo/active state relative to the apo/inactive state (Fig. 2a,c).

The CNB domain exhibits a markedly dual pattern of cAMP-dependent dynamics in agreement with the NMR data, with enhanced α-subdomain dynamics and quenched β-subdomain dynamics in the holo/active state relative to the apo/inactive state. Upon examining the backbone order parameters (S^2) computed for the holo/active state, it was

found that the α-subdomain, *i.e.* the NTHB and hinge helix, displayed enhanced dynamics relative to the apo/inactive state (Fig. 3b and Table 1). Particularly notable enhancements were observed within the C-terminal half of the hinge helix and in regions of the NTHB spanning all three residues involved in the ionic latch. In marked contrast to the α-subdomain, the β-core of the CNB displayed quenched dynamics relative to the apo/inactive state, with particularly significant quenching at the phosphate binding cassette (PBC), where cAMP docks, and at the adjacent β2-3 loop (Fig. 3b and Table 1). This pattern of α-subdomain enhancement and β-subdomain quenching of dynamics is qualitatively consistent with the trends observed for the cAMP-dependent changes in both the experimental S^2 values and the high-frequency spectral densities $J(\omega_H+\omega_N)$ measured by ^{15}N NMR for the EPAC1 CNB domain (Fig. 4b-d),[18] thus further validating the MD simulations. We therefore conclude that the MD computations capture the essential features of the cAMP-dependent changes in EPAC dynamics that were experimentally observed for the CNB. In addition, MD provides atomic-resolution insight into the dynamic profile of EPAC domains in the catalytic region that are not currently accessible by NMR.

The catalytic region exhibits multiple changes in dynamics in the holo/active state relative to the apo/inactive state. In addition to the enhanced dynamics of the CNB α-subdomain, enhancements of dynamics in the holo/active *vs.* apo/inactive state were observed at several sites of the catalytic CDC25-HD domain, as well as parts of the REM and RA domains (Fig. 3b). In general, the changes in local dynamics of the holo/active state in the absence of Rap1b relative to the apo/inactive state span the entire protein, with the largest changes generally corresponding to the most dynamic regions of the apo/inactive state (Fig.

3a,b and Table 1). For instance, the most notable dynamic enhancements in the CDC25-HD domain occurred in the loop composed of residues 953-961, in agreement with the missing electron density for this loop in the X-ray structure of holo EPAC2,[15] and in the region spanning both CDC25-HD domain residues that in the apo/inactive state were involved in the ionic latch (Fig. 3b). The transition to the holo/active state also causes quenching of dynamics in other regions of all three catalytic-region domains. In particular, quenching of dynamics was observed in the lid β-sheet and the helical hairpin loop, which interact to form the switchboard, as well as in the loop composed of residues 725-731, in agreement with the presence of this loop in the active-state X-ray structure (Fig. 3b).[15] Furthermore, the RA domain displays the greatest extent of dynamic changes in the catalytic region (Fig. 3b and Table 1), while the CDC25-HD domain exhibits the smallest overall extent of perturbation in the catalytic region (Fig. 3b and Table 1). Interestingly, the observed dynamic perturbations include regions forming some of the interactions that selectively stabilize the holo/active *vs.* the apo/inactive state (*i.e.* the cAMP/CNB, cAMP/REM and CNB/REM contacts; Fig. 3b, solid vertical black lines). In order to further explore the dynamics of these holo/active-specific interactions, several distance time-profiles were analyzed (Fig. S7).

In the holo/active state most of the cAMP/CNB contacts are stable, while the CNB/REM inter-domain interface is dynamic. Upon examination of the time profiles for key interactions of cAMP (Table S2), the hydrogen-bonds of cAMP with residues of the PBC (*i.e.* G404, A407, R414 and A415) and with K450 of the lid β-sheet proved to be quite stable, as suggested by contact distances persistently residing at/near values compatible with formation of hydrogen bonds and/or near the distance values observed in the X-ray structure.

In addition, the hydrophobic contacts of cAMP with the PBC (A415 and A416), lid (L449), F367, and the base binding region (BBR) residue V386 were also quite stable, as indicated by distances persistently residing at/near VDW contact range and/or near the distance values in the X-ray structure (Table S3). The only contacts between cAMP and the CNB that exhibited some degree of dynamics were those involving residues I388 and V394 of the BBR (Fig. S7a). In particular, I388 showed a clear drift in average cAMP contact distance over the course of the simulation, with overall variation over a range of approximately 4 Å (Fig. S7a). The V394 side chain demonstrated two sudden flips in orientation, between which one of the side chain's two terminal methyl groups (γ1 and γ2 methyls) formed a very stable contact with cAMP (Fig. S7a).

Contrary to the majority of the cAMP/CNB contacts, the CNB/REM interface appeared quite dynamic. The PBC→lid hydrogen bonds established by K405 (K405 with E443, N445 and Y480), and the β2-3 loop→REM domain hydrogen bond (Q369 with Y551), were highly dynamic (Fig. S7d,e). Furthermore, the BBR→REM domain hydrophobic contacts identified between Y389 and Y551 were also affected by significant dynamics (Fig. S7c). Finally, the hydrogen-bond interaction of cAMP with K489 of the lid α-helix (α1 of the REM domain) was found to be dynamic, as indicated by a highly variable contact distance (Fig. S7b). Overall, the time-dependent distance profiles of the holo/active state of EPAC suggest that the widespread dynamic changes of the catalytic region affect also its interface with the regulatory CNB domain.

The dynamics of the holo/active state are extensively correlated. Another unique feature of the holo/active state is a markedly distinct DCCM matrix in terms of both patterns

and magnitudes of positive and negative correlations, which are clearly different from any other state of the protein, as illustrated in Fig. 5b (upper half). For instance, marked changes in the correlation magnitudes and patterns are clearly apparent when comparing the DCCM of the holo/active state with the apo/inactive state (Fig. 5b, upper half, *vs.* Fig. 5a, lower half). First, more extensive positive correlations were observed within the REM domain and the CNB domain demonstrated mostly positive internal correlations in the holo/active state (Fig. 5b, upper half), without the extensive negative correlations between the two sub-domains observed for the apo/active state (Fig. 5a, upper half). In addition, mostly positive correlations were also observed between the CNB and REM domains, although they were visibly weaker than the intra-domain correlations (Fig. 5b, upper half). Furthermore, while the CDC25-HD domain demonstrated several markedly negative correlations with the other domains, some well-defined positive correlations were also visible and appeared to correspond to regions of the CDC25-HD domain forming contacts with the other domains (Fig. 5b, upper half). For instance, positive correlations observed with the CNB and REM domains corresponded to the helical hairpin, as well as other CDC25-HD domain elements interacting with the helical hairpin (Fig. 5b, upper half; Fig. 5d). Finally, it is notable that the RA domain is the EPAC2 module with the least positive inter-domain correlations (Fig. 5b, upper half), with the few positive correlations involving the RA domain confined to the RA/CDC25-HD contacts (Fig. 5b, upper half; Fig. 5d).

The overall high degree of correlation in the fluctuations of the holo/active state is also confirmed by PCA analysis of the covariance matrices (Fig. S5d). Specifically, when the MD trajectory was projected onto the space spanned by the first six eigenvectors, with all residues superimposed except for the most flexible regions (*i.e.* the first 20 amino acids of

the NTHB α1 helix and the mostly unstructured loop connecting the CNB and REM domains), the large anti-correlated oscillation that in the apo/inactive state was observed for the CNB and RA domains is now reduced in amplitude, but correlated motions are apparent throughout the structure, particularly in the catalytic region (Fig. S6b). Furthermore, the presence of collective motions involving the entire catalytic region is also independently confirmed by the differential RMSF profiles (Figs. 6d and S3b), which indicate that the REM, RA and CDC25-HD modules are subject to domain motions with significantly increased amplitude in the holo/active state relative to the apo/inactive state (Figs. 6a and S3a).

2.5.3. Dynamics of the Rap1b-Bound Holo/Active State of EPAC2

The open-topology of the Rap1b-bound holo/active state and the fold of each individual domain are stable in the ns timescale. As with the holo/active state, the RMSDs from both active and inactive initial EPAC2 conformations suggest that the Rap1b-bound holo/active state remains closer to its initial "open" structure than to the "closed" structure, without a significant change in overall EPAC2 topology (Fig. S8a). In addition, the structures of the individual domains appeared to be quite stable, as reflected by low (*i.e.* < ~3 Å) RMSDs during the simulation (Fig. S8a, positive time).

The binding of Rap1b substrate to the holo/active state causes a partial reversal of changes in dynamics from the apo/inactive state. Upon examining backbone order parameters (S^2) computed for the holo/active state with bound Rap1b (Fig. 3e and Table 1), it was found that the addition of Rap1b to the holo/active state resulted in a quenching of dynamics not only for most of the CDC25-HD domain, which contacts Rap1b directly, but

also for the distal CNB domain where dynamics in both the α- and β-subdomains are quenched upon Rap1b binding to the CDC25-HD. Unlike the CDC25-HD and CNB domains, an enhancement of dynamics was observed throughout most of the RA domain (Fig. 3e and Table 1). Overall, it appears therefore that Rap1b causes a partial reversal of the dynamic changes observed in the apo/inactive-to-holo/active transition. This trend is also confirmed at the level of the DCCM and differential RMSF analyses.

Upon examination of the DCCM results for the holo/active state with bound Rap1b (Fig. S8b), it was found that the strong correlations and anti-correlations observed within the holo/active state (Fig. 5b, upper half) became weakened upon Rap1b binding. In fact, the Rap1b-bound holo/active state demonstrated DCCM magnitudes that were more comparable to those of the apo/inactive state (Figs. 5 and S8b). In addition, the differential RMSF analysis (Fig. S8c) indicated that the amplitudes of the non-local motions for the REM and CDC25-HD domains are reduced upon substrate binding to values more similar to those observed for the apo/inactive state (Fig. 6a).

2.5.4. Changes in Dynamics Observed for the Holo/Inactive Metastate of EPAC2

The closed-topology of the holo/inactive metastate and the fold of each individual domain are quite stable in the ns timescale. As with the apo/inactive state, the RMSDs from both active and inactive initial EPAC2 conformations suggested that, despite the presence of cAMP, the holo/inactive metastate MD trajectory remains closer to its initial "closed" structure than to the "open" structure, without a significant change in overall topology in the ns timescale (Fig. 2b). In addition, the structures of the individual domains appeared to be quite stable with the exception of the CNB domain, which demonstrated a notable sudden

increase of 1 Å at approximately 37 ns into the 50 ns production run period (Fig. 2b). Visualization of structures before and after this time point indicated a shift in the position of the NTHB helices α1 and α2 relative to the rest of the CNB domain, explaining the observed RMSD increase (Fig. 2b inset). Overall, these observations suggest that the transition from the holo/inactive to the holo/active state occurs in a timescale longer than tens of ns, which is beyond the scope of the MD simulations analyzed here. However, the absence of inter-state transitions during the MD trajectory provides the opportunity of capturing the distinct ns/sub-ns dynamic profile of the holo/inactive metastate, which is otherwise challenging to trap and characterize experimentally.

The CNB and RA dynamics are perturbed in the holo/inactive metastate relative to the apo/inactive state, with lesser overall changes in the other domains. Upon examining the backbone order parameters (S^2), it was found that relative to the apo/inactive state, the holo/inactive metastate demonstrated a quenching of dynamics within the β-core, with particularly significant quenching at the PBC, as expected, and at the adjacent β2-3 loop (Fig. 3d and Table 1). Meanwhile, dynamic enhancements occurred throughout most of the NTHB, including the regions spanning all three CNB domain residues involved in the ionic latch (Fig. 3d and Table 1). Similarly to the CNB α-subdomain, the RA domain exhibited primarily enhanced local dynamics (Fig. 3d and Table 1), while a quenching of dynamics was observed in the lid β-sheet and parts of the CDC25-HD domain, including a partial quench of the helical hairpin loop that forms part of the switchboard (Fig. 3d). Overall, the holo/inactive state displays the lowest average order parameters (S^2) among the five states investigated here (Table 1), suggesting that the holo/inactive state benefits from an entropic

stabilization. Further insight into the dynamics in the holo/inactive state is provided by the comparative analysis of the DCCMs of Fig. 5a.

The holo/inactive metastate promotes a de-correlation of the CNB/CDC25-HD dynamics. Upon examination of the DCCM results for the apo/inactive and holo/inactive states (Fig. 5a), it was found that the anti-correlations between the NTHB and CDC25-HD domain almost vanished and the anti-correlations between the β-core and CDC25-HD domain observed for the apo/inactive state became somewhat diminished in the holo/inactive metastate (Fig. 5a, upper half). In addition, the positive correlations between the NTHB and β-core of the CNB domain became diminished in the holo/inactive metastate (Fig. 5a, upper half). Overall, these observations point to a loosening effect on the IL interface formed by the NTHB upon transition from the apo/inactive state to the holo/inactive metastate. This result is also consistent with the increased α-subdomain local dynamics in the holo/inactive *vs.* apo/inactive state (Table 1) and with the aforementioned shift in the positions of the NTHB α1 and α2 helices relative to the CNB β-core, as revealed by the CNB-specific RMSD time profile of the holo/inactive simulation (Fig. 2b). Furthermore, the apo/inactive-to-holo/inactive transition results in enhanced non-local dynamics for the RA domain and CNB domain β-core, as indicated by the differential RMSF analysis (Figs. 6b and S3d).

2.5.5. Changes in Dynamics Observed for the Apo/Active Metastate of EPAC2

The open-topology of the apo/active metastate and the fold of each individual domain are stable in the ns timescale. As with the holo/active state, the RMSDs from both active and inactive initial EPAC2 conformations suggested that the apo/active metastate MD trajectory remains closer to its initial “open” structure than to the “closed” structure, without

a significant change in overall topology despite the absence of cAMP (Fig. 2c). In addition, the structures of the individual domains appeared to be quite stable, as reflected by lower RMSDs during the simulation (Fig. 2c, positive time). Again, the absence of inter-state transitions during the course of the MD trajectory provides the opportunity to capture the dynamic features of the otherwise elusive apo-active metastate.

The apo/active metastate demonstrates mostly quenched dynamics, relative to the apo/inactive state. Upon examining the backbone order parameters (S^2), it was found that relative to the apo/inactive state, the apo/active metastate demonstrated a quenching of dynamics in the β-core and a partial enhancement of dynamics confined to most of the α-subdomain, *i.e.* the NTHB and hinge helix (Fig. 3c and Table 1). Meanwhile, a quenching of dynamics was observed in the RA domain, as well as in the lid β-sheet and helical hairpin loop (Fig. 3c and Table 1). Overall, the apo/active state displays the highest average order parameters (S^2) among the states investigated here (Table 1), suggesting that the apo/active state is associated with an entropic penalty.

Correlations among residue movements are less extensive in the apo/active metastate than in the holo/active state. Upon examination of the DCCM results, it was found that unlike the holo/active state, the apo/active metastate demonstrated correlations somewhat comparable in magnitude to those observed for the two overall weakly correlated inactive states (Fig. 5a,b). Specifically, the correlation of CNB and lid/REM domain movements observed in the holo/active state was abolished in the apo/active metastate, suggesting a de-correlation across the CNB/REM domain interface upon cAMP dissociation (Fig. 5b). Furthermore, the large amplitude domain motions observed for the CNB and CR in the

holo/active state are also to a large extent lost in the apo/active state (Figs. 6c,d and S3b,c), thus confirming the loss of inter-domain fluctuations upon cAMP dissociation and further supporting the existence of an entropic penalty associated with the apo/active metastate.

2.6. Discussion

The MD simulations are validated by the available experimental data and also provide novel insight into functional dynamic features that have remained elusive to experimental methods. The first notable feature of the MD simulations presented here is that they capture the key trends in the dynamic profiles experimentally probed by NMR for the CNB domain in both the apo and holo states. Such agreement has two key implications. First, it suggests that the NMR data previously obtained for a single-domain construct[18] are relevant also in the context of full length EPAC. Second, the ability of MD simulations to reproduce critical experimental trends validates the reliability of MD trajectories in qualitatively capturing the EPAC dynamic profile and its cAMP-dependence. The MD simulations thus provide an unprecedented view of EPAC dynamics, which is particularly valuable for the motions that have remained so far challenging to probe experimentally at residue resolution, such as those in the catalytic region and in the apo/active and holo/inactive metastates (Fig. 1b). The major results obtained from the five simulations of EPAC2(280-990) (Table S1) are summarized schematically in Figs. 7 and S8d. The dynamic fluctuations of EPAC summarized in Figs. 7 and S8d have several pivotal implications for both the auto-inhibition and the cAMP-dependent activation of EPAC. We will first address the role of dynamics in the latter.

2.6.1. Dynamics in the cAMP-Dependent Activation of EPAC

EPAC activation involves a reciprocal long-range RR/CR dynamic cross-talk, assisted by highly-correlated motions in the switchboard region and relevant for substrate binding. Upon conversion from the apo/inactive state to the holo/active state, several changes in local dynamics occur throughout EPAC (Fig. 7), indicating that the cAMP-dependent changes in dynamics that were previously observed for the CNB domain[18] are actually not confined solely to the RR, but extend to the CR as well (Fig. 7). Specifically, for both the RA domain and the helical hairpin (HP) in the CR, the local dynamics are quenched in the holo/active state relative to the apo/inactive state (Fig. 7). The reduction of HP dynamics explains the reduced solvent accessibility observed for this region upon cAMP binding in recent H/D exchange MS experiments.[39] Furthermore, the cAMP-dependent dynamic quenching in the HP region reduces the entropic penalty for the binding of the Rap1b substrate to the HP,[15] which is likely to promote the Rap1b/CDC25-HD interaction.[40]

The perturbations in the CR caused by cAMP binding to the RR are not limited only to local dynamics, but they affect also the collective inter-domain dynamics. For instance, Figs. 6 and 7 show that cAMP binding amplifies the amplitude of the non-local RA and CDC25-HD domain dynamics. Remarkably, such long-range cAMP-dependent effects on non-local CR dynamics occur irrespective of whether cAMP binds to the inactive or active conformations of EPAC (Figs. 6 and 7), suggesting that cAMP controls the function of the catalytic region of EPAC not only by stabilizing the open active structure as previously thought,[14,15,16,17] but also by contributing directly to the modulation of CR dynamics.

The MD simulations also indicate that the cross-talk between the RR and CR of EPAC is reciprocal. Not only does cAMP-binding to the RR modulate CR dynamics, but

Rap1b-binding to the CR also modulates the RR dynamics (Figs. 3e and S8c,d; Table 1). These long-range effects of Rap1b-binding are fully consistent with the highly-correlated character of the dynamics in the holo/active state (Fig. 5b, upper half), which serves the purpose of promoting allosteric signal propagation across the different EPAC domains. In particular, the switchboard region emerges as a stable hub of interactions that anchor together three critical EPAC domains (the CNB, REM and catalytic CDC25-HD domains) and remain highly correlated in all EPAC states simulated here (Fig. 5), mediating the dynamical cross-talk between the regulatory and catalytic regions of EPAC. Additional insight into the mechanism of EPAC activation is obtained through a further analysis of the dynamic profiles of the two metastates (*i.e.* holo/inactive and apo/active), which are critical to dissect the relative contributions of cAMP-binding and of closed-to-open conformational transitions to the cAMP-dependent EPAC activation.

EPAC activation involves dynamic contributions from both cAMP binding and the transition to the active conformation. The dynamic profiles of the holo/inactive and apo/active metastates (Table 1; Figs. 3 and 7) suggest that both cAMP binding and transition of EPAC to the active conformation contribute to the changes in local dynamics occurring upon EPAC activation. However, the magnitudes of the distinct contributions of the two metastates appear to be domain specific. For instance, in the case of the CNB domain, the pattern of α-subdomain enhancement and β-core quenching in local dynamics reported for the apo/inactive to holo/active transition is also observed in both metastates (Figs. 3 and 7; Table 1), while in the case of the CR domains, the quenching in local dynamics associated with the apo/inactive to holo/active transition (Table 1) appears to arise predominantly from

the apo/active metatstate (Table 1). Such insight provided by the MD simulations into the metastate dynamics has important implications for the significance of the induced-fit *vs.* conformational-selection pathways of allostery, which proceed through the holo-inactive and apo-active intermediates, respectively (Fig. 1b).[26,29,30,31,32,33]

Induced-fit vs. conformational-selection mechanisms of allostery. The MD simulations of the metastates (Figs. 1b and 7) suggest that the induced-fit and conformational-selection pathways can both contribute to the activation of EPAC by cAMP. The conformational-selection pathway is a plausible model of EPAC allostery because the apo/active metastate has a higher binding affinity for cAMP than the apo/inactive state, due to better geometric complementarity[29,30,31] and reduced binding site dynamics (Fig. 7). However, it should be considered that the RMSD, order parameter (S^2) and DCCM results consistently point to a decoupling of dynamics across the ionic-latch interface in the holo/inactive metastate, *i.e.* the binding of cAMP to the inactive state triggers a destabilization of the ionic-latch interface between the CNB and CDC25-HD domains. This means that under conditions in which the induced-fit pathway prevails, *i.e.* when cAMP binding precedes the conformational transition from inactive to active structures, cAMP primes the EPAC system for the 'closed' to 'open' transition by loosening the ionic-latch interface prior to the actual conversion from inactive to active conformations (Fig. 7). Furthermore, the holo/inactive metastate, unlike the apo/active metastate, is entropically stabilized by local dynamics as indicated by a minimal average S^2 value (Table 1). We therefore conclude that, while the conformational selection pathway can account for allostery in EPAC, the induced fit mechanism should not be *a priori* ruled out either. Indeed, it has

been shown that both pathways are viable depending on the experimental conditions,[29,32,33] and previous kinetic and NMR studies have suggested that ligand-associated conformational changes occur significantly through both pathways under a range of experimental conditions.[41,42] A possible hybrid mechanism for EPAC allostery should then be considered and it is discussed below.

A hybrid induced-fit/conformational-selection mechanism for EPAC allostery provides a low-free energy activation pathway. The finding that in the holo/inactive state the IL interface is weakened by cAMP binding due to an increase in its entropic penalty suggests a possible low free energy path for the apo/inactive-to-holo/active transition. Specifically, low-affinity binding of cAMP to the apo/inactive state leads to the transient formation of a holo/inactive intermediate, which may serve as a transitory complex in which the IL is weakened and primed for transition to the active conformation (Fig. 7). Due to the low binding affinity of cAMP for the inactive conformation of EPAC, it cannot be ruled out that during the conformational transition cAMP may dissociate from EPAC, leading to the formation of the apo/active intermediate (Fig. 7). However, this metastate can in turn re-associate with cAMP with significantly higher affinity, eventually resulting in the holo/active state (Fig. 7).

A remarkable feature of the proposed hybrid activation pathway is that in each successive step unfavorable entropic losses are compensated, at least in part, by favorable entropic gains. For instance, in the apo/inactive-to-holo/inactive transition a partial quenching of local β-core and switchboard dynamics is compensated by an enhancement in the local dynamics of the NTHB and the RA domain (Fig. 7a,b). In the holo/inactive-to-

apo/active step, entropic losses in the β-core, CDC25-HD and RA domains are partially offset by entropic gains linked to the release of cAMP and to the increased dynamics of the CNB α-subdomain (Fig. 7b,c). Last but not least, in the apo/active-to-holo/active transition the entropic penalty due to the loss of free cAMP and to the quenching of REM domain dynamics is at least partially compensated by enhancements in the dynamics of the CDC25-HD and RA domains (Fig. 7c,d). It is notable that such entropic compensations not only apply to cAMP binding/release and local protein dynamics, but extend to protein inter-domain motions as well (Figs. 6 and 7). We therefore propose that such stepwise compensatory changes in protein dynamics serve the purpose of reducing excessive entropic losses along the apo/inactive-to-holo/active conversion pathway, thereby facilitating the minimization of free energy during the cAMP-dependent activation (Fig. 7). Of course, a similar consideration would apply for the reverse transitions that occur during the holo/active-to-apo/inactive conversion and therefore the hybrid pathway proposed above may also promote the reversibility of EPAC activation in response to cAMP. Such reversibility is critical for the timely termination of the signal relayed by cAMP in the EPAC-dependent signaling pathways.

2.6.2. Dynamics in the Auto-Inhibition of EPAC

Another notable conclusion emerging from the MD simulations presented here is that the two key sets of auto-inhibitory interactions that selectively stabilize the apo-inactive state, *i.e.* the hydrophobic hinge and the ionic latch, are both significantly affected by dynamics. The mechanism of EPAC auto-inhibition should thus be revisited in light of the dynamic profiles revealed by the MD trajectories.

A dynamic hydrophobic hinge. The high degree of dynamics predicted for the PBC in the apo/inactive state calls for a re-examination of the 'hydrophobic hinge' hypothesis previously proposed to rationalize the stabilization of the 'closed' EPAC conformation. The original hypothesis posited that in the inactive ('closed') topology the PBC residue L408 and the hinge helix residue F435 are oriented such that they create inhibitory steric hindrance between the PBC and hinge helix. As a result, the rotation of the hinge helix towards the PBC that occurs during conversion to the active ('open') conformation is blocked in the absence of cAMP, inhibiting the conformational shift.[14,15] However, this static view of the L408/F435 interaction does not take into account the highly dynamic nature of the PBC in the apo/inactive state (Fig. 3a). An additional explanation reconciling the pivotal inhibitory role of the conserved L408/F435 residues with the high degree of dynamics of the PBC is that the tight L408/F435 packing is unfavorable in the absence of cAMP as it would result in a significant entropic loss for the PBC. This interpretation is supported by the observation that the dynamics within the PBC, as well as other key β-core sites, were found to be quenched not only by cAMP binding, but also by the transition of EPAC to its active conformation (Figs. 3 and 7a,c; Table 1). Remarkably, the apo/active state displays the highest average S^2 value for the β-core as well as for the full length construct (Table 1), further confirming that the transition from inactive to active conformations in the absence of cAMP is associated with a significant penalty in terms of decreased conformational entropy.

Dynamics sensitize the IL to cAMP by causing a partial weakening of selected functional IL salt bridges. The MD simulation of the apo/inactive state suggests that at least part of the ionic-latch (IL) interface between the CNB and CDC25-HD domains is quite

labile. The order parameter (S^2) in the IL region and the time-profiles of the IL interaction distances consistently point to the IL interface being overall quite dynamic in the apo/inactive state, with the IL interactions mediated by Q303 and D307 being considerably less stable than the E332/R886 IL salt bridge (Figs. 3a and S4). This pattern is likely due at least in part to the presence of a Gly-Pro residue pair at positions 305 and 306, which destabilizes the α1 helix, including the flanking residues. However, it should be noted that despite their flexibility, the IL residues Q303 and D307 still play a role in the ionic latch interface, because when the transient interactions formed by these residues are perturbed through the Δ306 deletion mutation, the maximal activity of EPAC2 is enhanced fivefold.[14] Thus, the apparent lability of the contacts mediated by Q303 and D307 may serve the purpose of sensitizing the IL RR/CR inhibitory interface of EPAC to dynamic perturbations arising from cAMP binding, while at the same time avoiding the complete severance of the IL interface, thereby promoting efficient yet cAMP-dependent EPAC activation.[43,44] This interpretation is also consistent with the anti-correlated fluctuations of the CNB and CDC25-HD domains (Fig. 5a, lower triangle), which suggest the presence of "breathing" motions between the CNB domain and the catalytic region. Such motions may assist the initial phase of the opening of the CNB/CDC25-HD interface that occurs during the cAMP-dependent transition to the active conformation.

2.6.3. Concluding Remarks

We have performed five 50 ns MD simulations in explicit solvent on a ~700 residue functionally integral construct of EPAC2, starting from the key states of the thermodynamic cycle that describes the coupled binding/activation equilibria of EPAC. The MD results are

not only consistent with the currently available experimental data on the dynamics of EPAC, but they also provide an unprecedented insight on ps-ns dynamic attributes that have so far remained largely elusive to experimental approaches. Such dynamic features provide the basis for a molecular model explaining several key aspects of EPAC activation and auto-inhibition. Specifically, it was found that cAMP binding causes an extensive long-range perturbation of dynamics in the distal catalytic region, which assists the Rap1b substrate recognition. In addition, analysis of the apo/active and holo/inactive cross-states ("metastates") suggested a possible hybrid mechanism of EPAC allostery that incorporates elements of both induced-fit and conformational-selection mechanisms, resulting in a low-free-energy pathway by which EPAC is effectively and reversibly activated through an entropy compensation strategy. Finally, the simulations revealed that the auto-inhibitory interactions stabilizing the apo/inactive state of EPAC are significantly more dynamic than previously anticipated, suggesting that dynamics play a key role in controlling the stability of the apo/inactive state. Such dynamic fine-tuning is critical for the optimal sensitization of the auto-inhibitory interactions to the cAMP allosteric effector, because it ensures that the auto-inhibitory interactions are stable enough to avoid constitutive activation in the absence of cAMP, while at the same time preventing an excessive stabilization that would compromise an effective cAMP-dependent activation of EPAC. It is anticipated that the models proposed here are of general applicability to multi-domain signaling proteins controlled by allosteric effectors.

2.7. Acknowledgements

We thank Dr. Holger Rehmann (Utrecht University) and Prof. M. Zaccolo (University of Glasgow) for helpful discussions. We acknowledge the Shared Hierarchical Academic Research Computing Network (SHARCNET) for use of their high-performance computing resources. This work was supported by grants from the Canadian Institutes of Health Research (CIHR) and the Natural Sciences and Engineering Research Council of Canada (NSERC) to G.M. We are also indebted to the Heart and Stroke Foundation of Canada for a M. Andrew New Investigator Award to G.M.

2.8. Table and Figures

Table 1. *Average Backbone N-H Order Parameter (S^2) Values and Differences Computed for the Simulated States of EPAC2(280-990)*

Region	Apo/ Inactive State[a]	Holo/Active – Apo/Inactive[b]	Holo/Inactive – Apo/Inactive[b]	Apo/Active – Apo/Inactive[b]	Holo/Active with Rap1b – Holo/Active[b]
Full construct	0.81[c]	0.01 (0.21; -0.29)[e]	-0.05 (0.18; -0.15)[e]	0.07 (0.20; -0.13)[e]	-0.04 (0.13; -0.23)[e]
CNB α-subdomain	0.79[d]	-0.38 (0.00; -0.38)[f]	-0.12 (0.09; -0.13)[f]	-0.15 (0.00; -0.15)[f]	0.12 (0.13; -0.28)[f]
CNB β-core	0.79[d]	0.15 (0.17; -0.32)[f]	0.14 (0.20; -0.28)[f]	0.17 (0.19; -0.14)[f]	0.21 (0.21; 0.00)[f]
REM domain	0.84[d]	0.08 (0.17; -0.18)[f]	0.01 (0.16; -0.18)[f]	0.06 (0.17; -0.15)[f]	-0.04 (0.13; -0.12)[f]
RA domain	0.76[d]	0.12 (0.23; -0.17)[f]	-0.08 (0.19; -0.15)[f]	0.14 (0.22; -0.11)[f]	-0.16 (0.08; -0.17)[f]
CDC25-HD domain	0.85[d]	0.05 (0.17; -0.17)[f]	0.00 (0.19; -0.20)[f]	0.11 (0.20; -0.11)[f]	0.02 (0.13; -0.20)[f]

[a]Average apo/inactive S^2 values.
[b]Average S^2 differences (ΔS^2) computed between pairs of states of EPAC2(280-990). See footnotes *e* and *f* for further details.
[c]Average value computed across the complete EPAC2(280-990) construct.
[d]Average values computed for each region of the EPAC2(280-990) construct.
[e]Average values computed across the complete EPAC2(280-990) construct, only considering ΔS^2 values determined to be statistically significant (*i.e.* values where $|\Delta S^2| \geq 0.07$). The respective averages of all positive and all negative ΔS^2 values that satisfy the $|\Delta S^2| \geq 0.07$ criterion are shown in parentheses.
[f]Average values computed for each region of the EPAC2(280-990) construct, only considering values where $|\Delta S^2| \geq 0.07$. The respective averages of all positive and all negative ΔS^2 values that satisfy the $|\Delta S^2| \geq 0.07$ criterion are shown in parentheses.

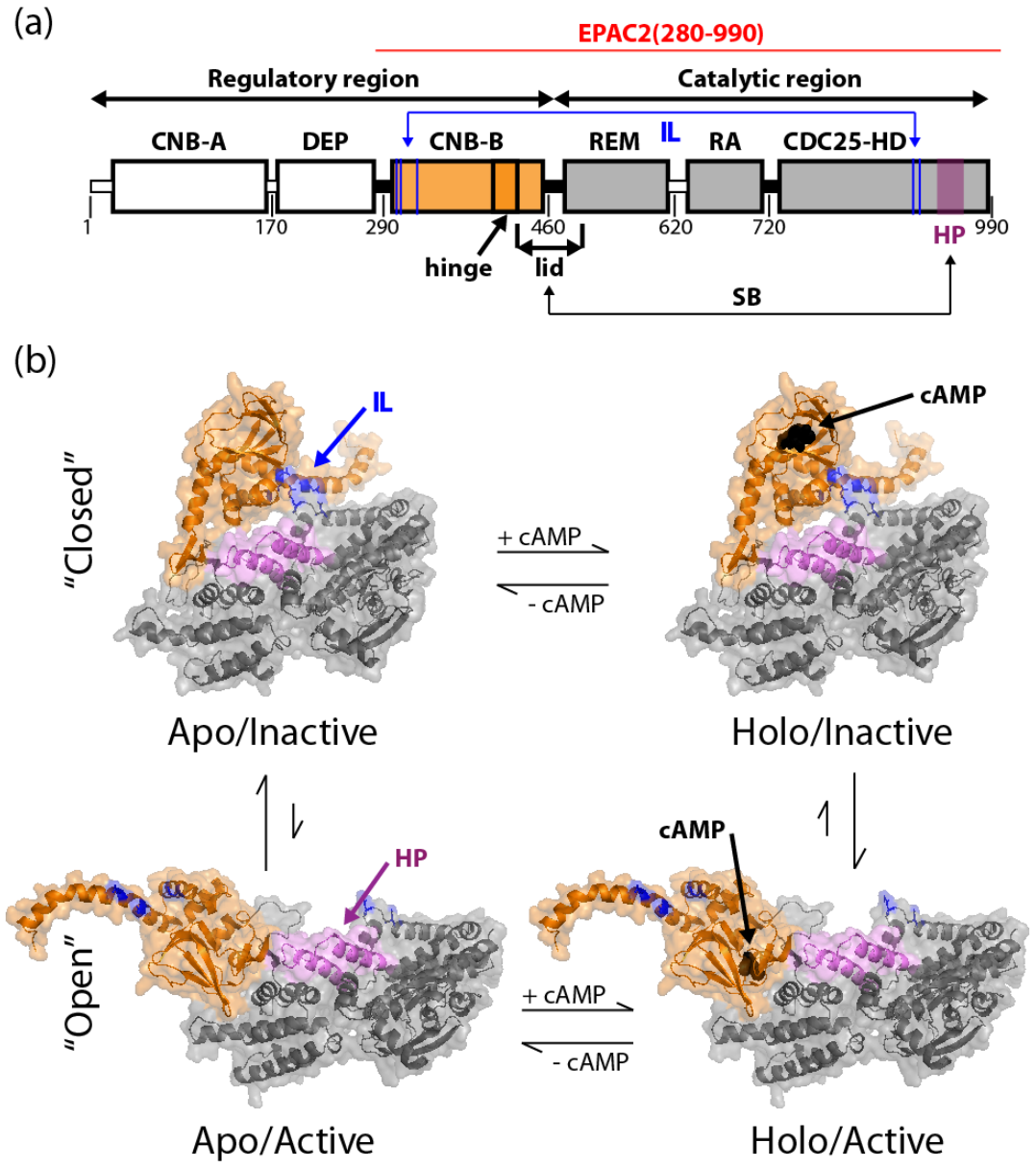

Figure 1: (a) Schematic representation of the primary structure of EPAC2, the EPAC isoform examined in the current work. The regulatory and catalytic regions of EPAC2 are indicated, as are the major underlying regions of EPAC2 structure: the cAMP binding domains ("CNB-A" and "CNB-B"); the DEP domain; the hinge helix ("hinge"); the lid region ("lid"); the REM domain; the RA domain; the catalytic CDC25-HD; and the helical hairpin (HP; indicated in purple) whose central loop, together with the lid β-sheet, is part of a β-sheet known as the "switchboard" (SB). The key ionic latch (IL) interface residues (Q303, D307, E332, D883 and R886) are indicated in blue. The EPAC2(280-990) fragment utilized in the current study is indicated by a red line and the allosteric CNB domain is highlighted in

orange. Residue numbers are also given to indicate the approximate domain boundaries within the primary structure. The structural elements excluded from the current work (*i.e.* the CNB-A and DEP domains, and the REM-RA connecting loop) are shown in white. **(b)** Ribbon/surface-structure model of EPAC allostery, illustrating the four states and the two coupled equilibria involved in the thermodynamic cycle for the cAMP-dependent EPAC activation. The EPAC2(280-990) fragment utilized in the current study is shown for all four states. The allosteric CNB domain (CNB-B) is highlighted in orange, the helical hairpin (HP) in purple and the bound cAMP is shown as a black space-filling structure. The key ionic latch (IL) interface residues are indicated in blue. All structures were generated using Pymol (Schrödinger, LLC), based on structures 2BYV and 3CF6 from the RCSB Protein Data Bank.

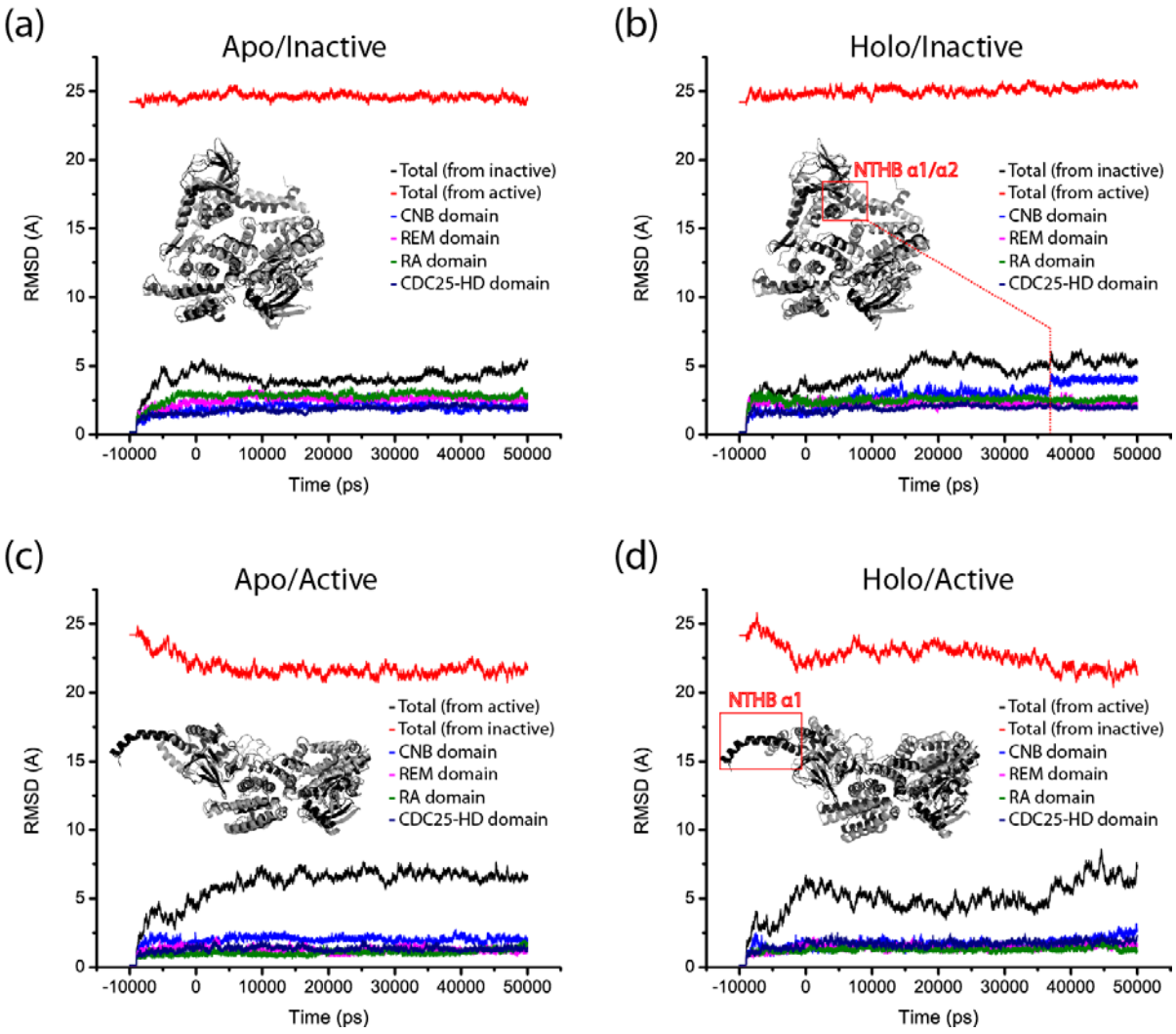

Figure 2: Root-mean-square deviations (RMSDs) from the initial structures computed over the course of each MD simulation: **(a)** apo/inactive; **(b)** holo/inactive; **(c)** apo/active; **(d)** holo/active. For each simulation, full-protein RMSDs from both active/inactive initial apo structures (for apo-state simulations) or both active/inactive initial holo structures (for holo-state simulations) are shown, along with RMSDs of each domain from its respective initial structure. The color-coding is shown in each panel. The initial equilibration period of each simulation is indicated as negative times, while the production run (used in subsequent analyses) is indicated as positive times. The overlaid initial and final structures from each simulation are shown as insets (black ribbon structure = initial structure; gray ribbon structure = final structure). For the holo/inactive-metastate simulation, the time at which a sudden conformational shift was observed in the NTHB α1 and α2 helices of the CNB domain is also indicated (red dashed line in panel b).

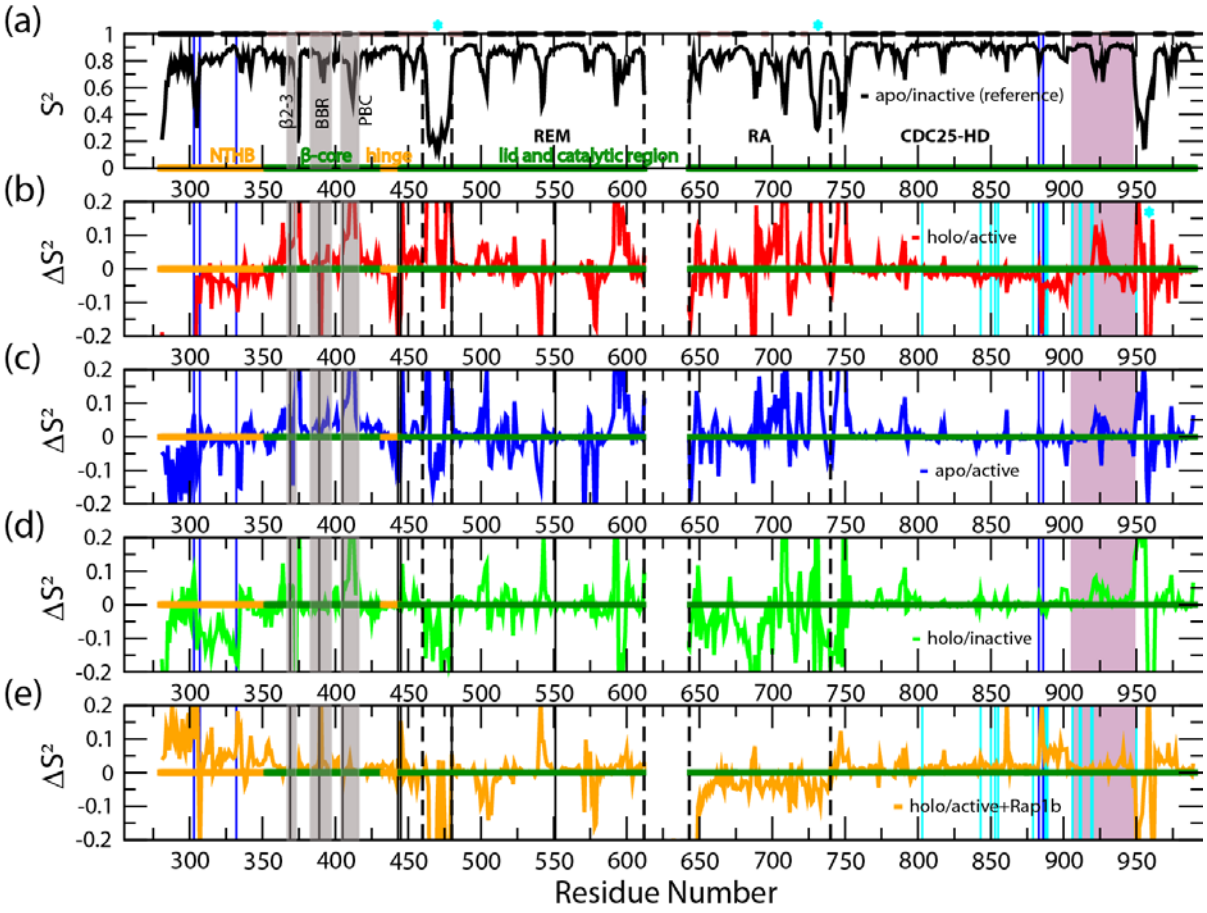

Figure 3: (a) Backbone N-H order parameters (S^2) for the apo/inactive state of EPAC2(280-990). **(b-d)** Differences in order parameters (ΔS^2) relative to the apo/inactive state computed for the **(b)** holo/active state, **(c)** apo/active metastate and **(d)** holo/inactive metastate. **(e)** Order parameter (S^2) differences relative to the holo/active state computed for the holo/active state with bound Rap1b. In all panels, the domain boundaries are marked by vertical dashed lines and key structural regions are indicated by orange and dark green bars: N-terminal helical bundle (NTHB), β-core, hinge helix, and the lid and catalytic region. The three key β-core elements (β2-3 loop, BBR, and PBC) are highlighted by gray bars and the helical hairpin (HP; located in the CDC25-HD domain) by a purple bar. The five ionic-latch (IL) residues are indicated by vertical blue lines, and residues involved in CNB domain-lid/REM domain contacts of the active conformation are indicated by vertical, solid black lines. Residues that form contacts with bound Rap1b are indicated by vertical cyan lines. The secondary structure elements are indicated over the top panel (black rectangles = α-helices; brown rectangles = β-strands), and loops missing from the apo/inactive and holo/active X-ray structures are indicated by cyan asterixes in panels (a) and (b), respectively.

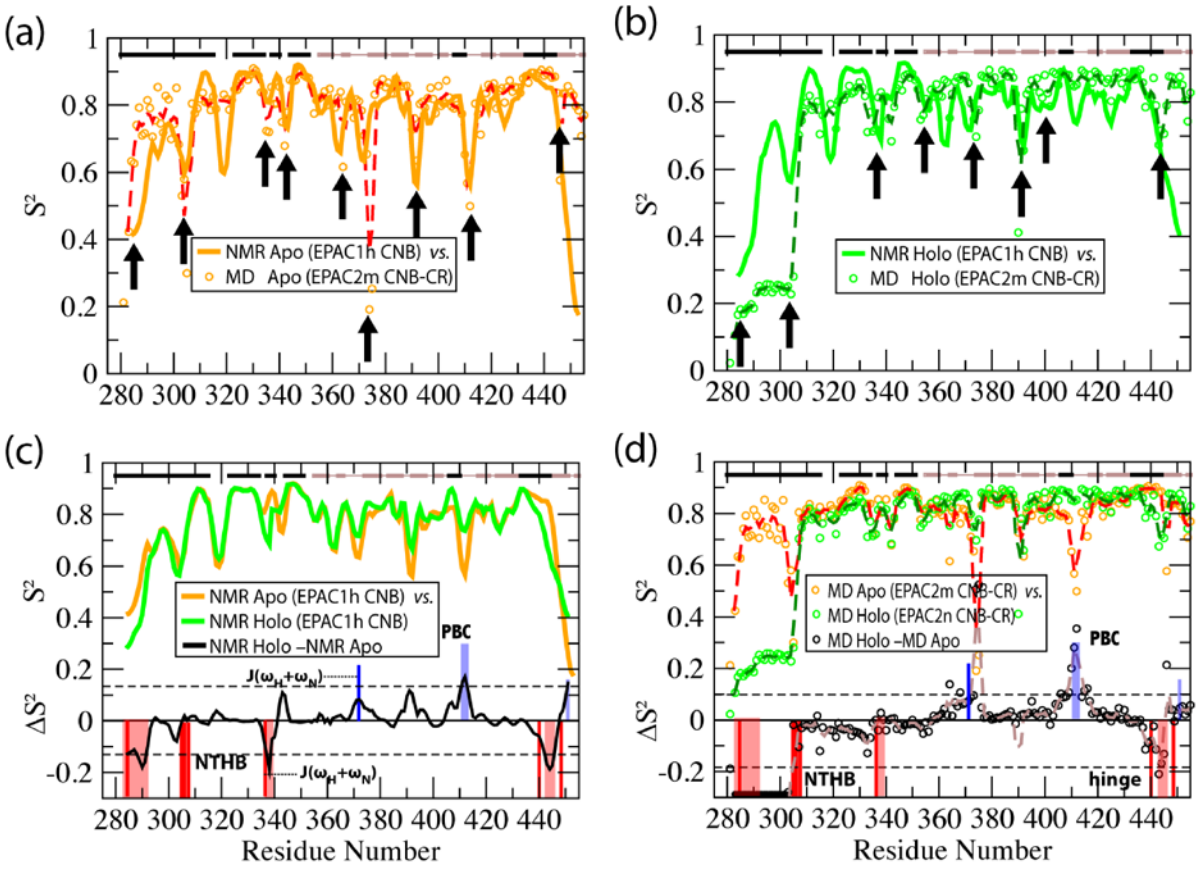

Figure 4: Comparison of the CNB backbone N-H order parameters (S^2) from NMR and from MD. The MD order parameters were computed based on the 50 ns simulations of the EPAC2(280-990) construct. Symbol legends are embedded in the figure and, to clarify trends in the MD data, three-point-average smoothed trend-lines are displayed for each MD data set. **(a)** NMR *vs.* MD S^2 comparison for the apo state of EPAC. Local minima in the NMR data that also appear in the MD data are indicated by black arrows. **(b)** Similar to panel (a), but for the holo state of EPAC. **(c)** Effect of cAMP-binding on the NMR order parameters as illustrated by S^2 values, as well as S^2 differences (ΔS^2). Dashed lines are shown at ΔS^2 values equal to the average value ± ε, where ε = one standard deviation + estimated error for the ΔS^2 experimental values (*i.e.* 0.07). The solid horizontal line is at ΔS^2 = 0. The regions affected by the most significant enhancements (decreases) in dynamics upon cAMP-binding, as assessed based on the experimental S^2 difference (ΔS^2) between holo and apo states, are highlighted with red (blue) background. Red (blue) vertical bars denote residues subject to enhancements (decreases) in ps-ns dynamics upon cAMP-binding, as previously assessed based on high frequency reduced spectral density analyses (*i.e.* $J(\omega_H+\omega_N)$). **(d)** Effect of cAMP-binding on the MD order parameters. Dashed lines are shown at ΔS^2 values equal to the 5% trimmed average +/- one standard deviation. The solid horizontal line is at ΔS^2 = 0. The same color coding for the patterns of dynamic

enhancement and quenching as in panel (c) are reported in panel (d) to facilitate the comparison of the qualitative trends of cAMP-dependent dynamics. Black circles for residues 281-302 are off-scale and are just reported at a ΔS^2 of -0.28. In the top of all panels, the secondary structure is reported as black and brown rectangles representing α-helices and β-strands, respectively.

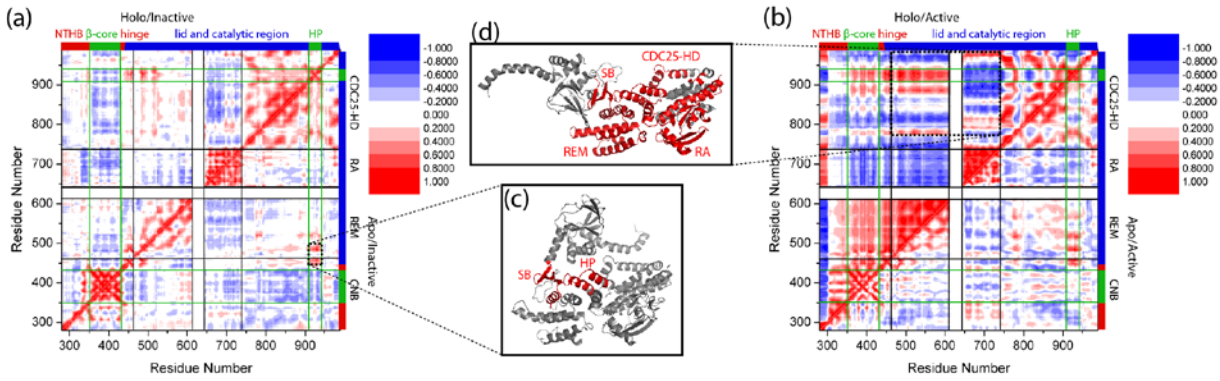

Figure 5: Dynamic cross-correlation matrices (DCCMs) computed for the apo/inactive (a, lower half) and holo/active (b, upper half) states as well as holo/inactive (a, upper half) and apo/active (b, lower half) metastates of EPAC2(280-990). The correlation coefficient for each pair of α-carbon atoms is represented as a point plotted for the corresponding pair of residues. Red points indicate positive correlation of residue movements, while blue points indicate negative correlation (anti-correlation) of residue movements. The four domains (CNB, REM, RA, and CDC25-HD) are separated by solid black lines and key structural regions are indicated: N-terminal helical bundle (NTHB) in red, β-core in green (delineated by flanking green lines for clarity), hinge helix in red, lid and catalytic region in blue and the helical hairpin in green (HP; delineated by flanking green lines). Key positive correlations with the CDC25-HD domain are indicated by dashed black rectangles. The structural regions involved in the key correlations of the apo/inactive (a) and holo/active (b) states are illustrated in panels (c) and (d), respectively. The correlated structural elements are highlighted in red and in the apo/inactive state correspond to residues within/near the switchboard β-sheet (SB). HP stands for helical hairpin. All structures were generated using Pymol (Schrödinger, LLC), based on the 2BYV and 3CF6 coordinates from the RCSB Protein Data Bank.

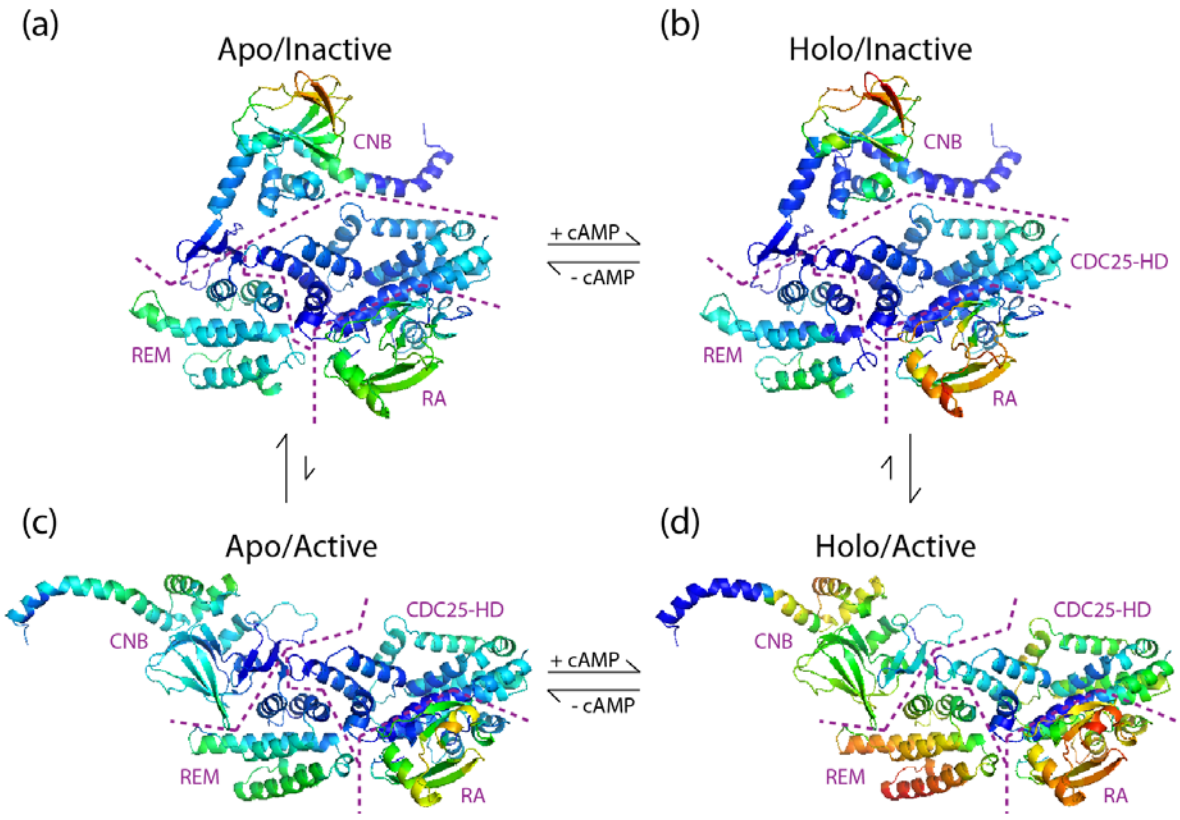

Figure 6: Ribbon-structure map of the differential RMSF values computed for each state of EPAC2(280-990): **(a)** apo/inactive; **(b)** holo/inactive; **(c)** apo/active; **(d)** holo/active. The differential RMSF values for each state are indicated in the respective structures using a continuous colour scale, which ranges from dark blue (RMSF = 0.0 Å) to red (RMSF = 3.0 Å). For clarity, the domain boundaries are indicated by purple dashed lines. All structures were generated using Pymol (Schrödinger, LLC), based on structures 2BYV and 3CF6 from the RCSB Protein Data Bank.

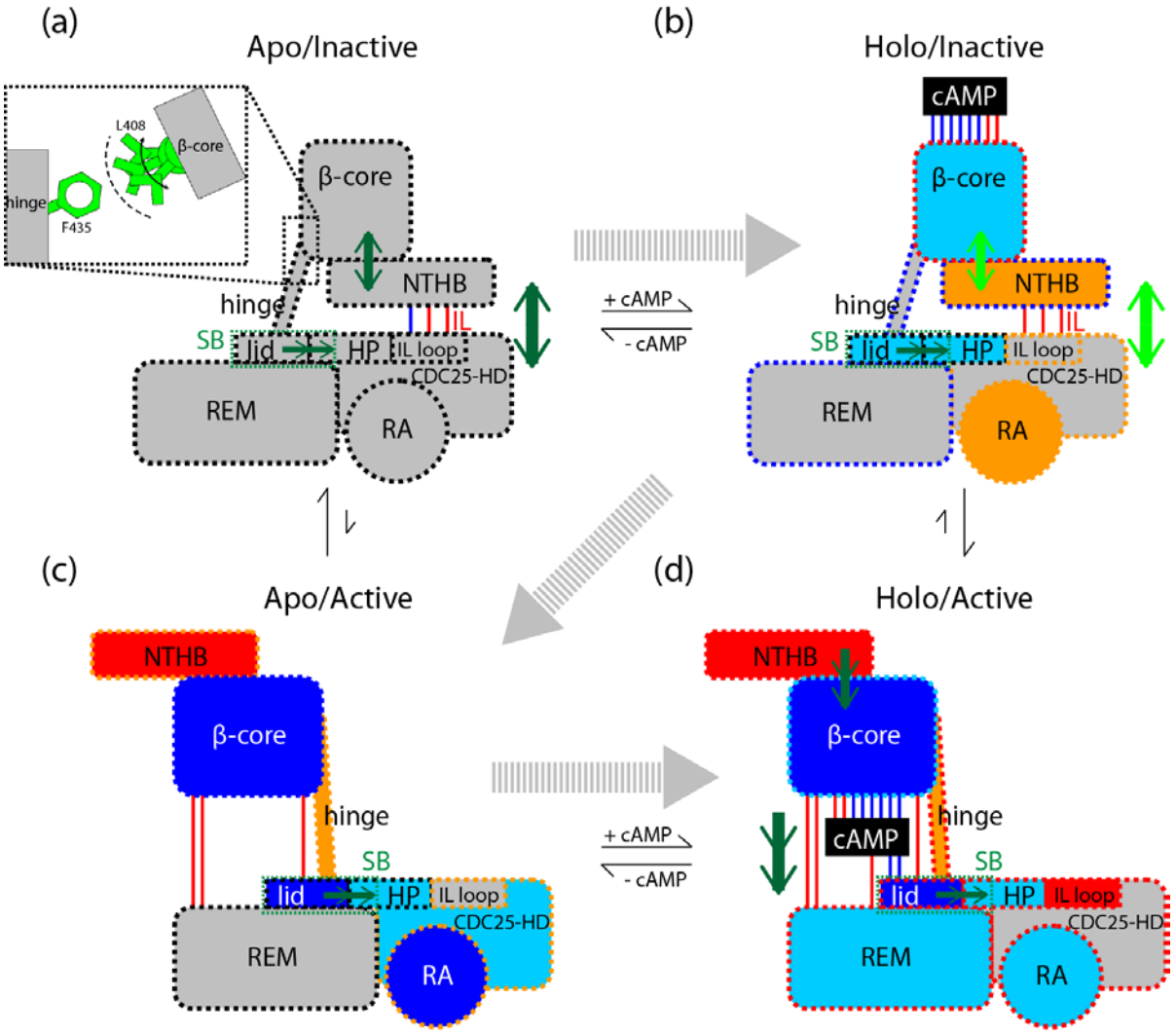

Figure 7: Schematic outline of the major results obtained from the analysis of the four states of EPAC2(280-990): **(a)** apo/inactive; **(b)** holo/inactive; **(c)** apo/active; **(d)** holo/active. The following key structural elements are examined: the N-terminal helical bundle (NTHB), β-core and hinge helix of the CNB domain; the switchboard (SB; highlighted by a green, dotted rectangle), which is formed by the lid β-sheet and by the helical hairpin (HP) and anchors the CNB, REM and CDC25-HD domains together; the ionic latch (IL); and the REM, RA and CDC25-HD domains of the catalytic region. Changes in the local dynamics of key structural regions (relative to the apo/inactive state) are indicated by color-coding the interior of each region as follows: gray = no statistically-significant overall change (*i.e.* $|\text{average } \Delta S^2| < 0.07$); red = enhanced dynamics with average $\Delta S^2 \leq -0.14$; orange = enhanced dynamics with $-0.07 \geq \text{average } \Delta S^2 > -0.14$; blue = quenched dynamics with average $\Delta S^2 \geq 0.14$; cyan = quenched dynamics with $0.07 \leq \text{average } \Delta S^2 < 0.14$. Changes in the long-range inter-domain dynamics of the key structural regions (relative to the apo/inactive state) are indicated by color-coding

the element outline as follows: black = no signifcant overall change; red = enhanced dynamics; orange = partially-enhanced dynamics; blue = quenched dynamics; cyan = partially-quenched dynamics. The residue-residue contacts of the CNB domain with the catalytic region and the contacts between the protein and the bound cAMP (shown as a solid black rectangle) are each indicated as blue or red lines depending on whether they are stable or dynamic, respectively. Green arrows indicate correlated motions among the structural elements that were identified from the DCCM analysis, where dark green = stronger correlation; light green = weaker correlation; one-way arrow = positive correlation; two-way arrow = negative correlation (anti-correlation). The inset for the apo/inactive state illustrates the hypothesis that the phosphate binding cassette (PBC) dynamics in the apo/inactive state hinder the proper packing of PBC residue L408 and hinge helix residue F435 against one another. In this state, movement of L408 creates "dynamic hindrance" (indicated as a dashed line) with F435, entropically blocking their packing. Large, dashed gray arrows between states indicate the inter-state transitions involved in the proposed induced-fit/conformational selection hybrid mechanism of EPAC allostery.

2.9. References

[1] Kawasaki, H., Springett, G. M., Mochizuki, N., Toki, S., Nakaya, M., Matsuda, M., Housman, D. E. and Graybiel, A. M. (1998) *Science* **282**, 2275-2279.
[2] de Rooij, J., Zwartkruis, F. J. T., Verheijen, M. H. G., Cool, R. H., Nijman, S. M. B., Wittinghofer, A. and Bos, J. L. (1998) *Nature* **396**, 474-477.
[3] Bos, J.L. (2006) *Trends Biochem. Sci.* **31**, 680-686.
[4] Roscioni, S.S., Elzinga, C.R.S. and Schmidt, M. (2008) *N-S Arch. Pharmacol.* **377**, 345-357.
[5] Rangarajan, S., Enserink, J.M., Kuiperij, H.B., de Rooij, J., Price, L.S., Schwede, F. and Bos, J.L. (2003) *J. Cell Biol.* **160**, 1-7.
[6] Enserink, J.M., Price, L.S., Methi, T., Mahic, M., Sonnenberg, A., Bos, J.L. and Taskén, K. (2004) *J. Biol. Chem.* **279**, 44889-44896.
[7] Cullere, X., Shaw, S.K., Andersson, L., Hirahashi, J., Luscinskas, F.W. and Mayadas, T.N. (2005) *Blood* **105**, 1950-1955.
[8] Schmidt, M., Sand, C., Jakobs, K.H., Michel, M.C. and Weernink, P.A.O. (2007) *Curr. Opin. Pharmacol.* **7**, 193-200.
[9] Holz, G.G. (2004) *Diabetes* **53**, 5-13.
[10] Robert, S., Maillet, M., Morel, E., Launnay, J.M., Fischmeister, R., Mercken, L. and Lezoualch, F. (2005) *FEBS Lett.* **579**, 1136-1142.
[11] Gloerich, M. and Bos, J.L. (2010) *Annu. Rev. Pharmacol.* **50**, 355-375.
[12] Kannan, N., Wu, J., Anand, G.S., Yooseph, S., Neuwald, A.F., Venter, J.C. and Taylor, S.S. (2007) *Genome Biol.* **8**, R264(1-13).
[13] de Rooij, J., Rehmann, H., van Triest, M., Cool, R.H., Wittinghofer, A. and Bos, J.L. (2000) *J. Biol. Chem.* **275**, 20829-20836.
[14] Rehmann, H., Das, J., Knipscheer, P., Wittinghofer, A. and Bos, J.L. (2006) *Nature* **439**, 625-628.
[15] Rehmann, H., Arias-Palomo, E., Hadders, M.A., Schwede, F., Llorca, O. and Bos, J.L. (2008) *Nature* **455**, 124-127.
[16] Harper, S.M., Wienk, H., Wechselberger, R.W., Bos, J.L., Boelens, R. and Rehmann, H. (2008) *J. Biol. Chem.* **283**, 6501-6508.
[17] Tsalkova, T., Blumenthal, D.K., Mei, F.C., White, M.A. and Cheng, X. (2009) *J. Biol. Chem.* **284**, 23644-23651.
[18] Das, R., Mazhab-Jafari, M.T., Chowdhury, S., Das, S.S., Selvaratnam, R. and Melacini, G. (2008) *J. Biol. Chem.* **283**, 19691-19703.
[19] Mazhab-Jafari, M.T., Das, R., Fotheringham, S.A., Das, S.S., Chowdhury, S. and Melacini, G. (2007) *J. Am. Chem. Soc.* **129**, 14482-14492.
[20] Rehmann, H., Prakash, B., Wolf, E., Rueppel, A., de Rooij, J., Bos, J.L. and Wittinghofer, A. (2003) *Nat. Struct. Biol.* **10**, 26-32.
[21] Rehmann, H., Rueppel, A., Bos, J.L. and Wittinghofer, A. (2003) *J. Biol. Chem.* **278**, 23508-23514.
[22] Selvaratnam, R., Chowdhury, S., VanSchouwen, B. and Melacini, G. (2011) *P. Natl. Acad. Sci. USA* **108**, 6133-6138.
[23] Das, R., Chowdhury, S., Mazhab-Jafari, M.T., SilDas, S., Selvaratnam, R. and Melacini, G. (2009) *J. Biol. Chem.* **284**, 23682-23696.

[24] Kern, D. and Zuiderweg, E.R. (2003) *Curr. Opin. Struct. Biol.* **13**, 748-757.
[25] Tsai, C.J., del Sol, A. and Nussinov, R. (2008) *J. Mol. Biol.* **378**, 1-11.
[26] Ma, B. and Nussinov, R. (2010) *Curr. Opin. Chem. Biol.* **14**, 652-659.
[27] Nicholson, L.K. (2002) *Biophys. J.* **82**, 3A-3A.
[28] Masterson, L.R., Cheng, C., Yu, T., Tonelli, M., Kornev, A., Taylor, S.S. and Veglia, G. (2010) *Nat. Chem. Biol.* **6**, 821-828.
[29] Boehr D.D. and Wright, P.E. (2008) *Science* **320**, 1429-1430.
[30] Grünberg, R., Leckner, J. and Nilges, M. (2004) *Structure* **12**, 2125-2136.
[31] Bosshard, H.R. (2001) *News Physiol. Sci.* **16**, 171-173.
[32] Weikl T.R. and von Deuster, C. (2009) *Proteins* **75**, 104-110.
[33] Boehr, D.D., Nussinov, R. and Wright, P.E. (2009) *Nat. Chem. Biol.* **5**, 789-796.
[34] Grant, B.J., Gorfe, A.A. and McCammon, J.A. (2010) *Curr. Opin. Struc. Biol.* **20**, 142-147.
[35] Adcock, S.A. and McCammon, J.A. (2006) *Chem. Rev.* **106**, 1589-1615.
[36] McCammon, M.G. and Robinson, C.V. (2004) *Curr. Opin. Chem. Biol.* **8**, 60-65.
[37] Wong, C.F. and McCammon, J.A. (2003) *Annu. Rev. Pharmacol.* **43**, 31-45.
[38] Patel, N., Dubins, D.N., Pomès, R. and Chalikian, T.V. (2011) *J. Phys. Chem. B* **115**, 4856-4862.
[39] Li, S., Tsalkova, T., White, M.A., Mei, F.C., Liu, T., Wang, D., Woods, V.L. and Cheng, X. (2011) *J. Biol. Chem.* **286**, 17889-17897.
[40] Jayaraman, B. and Nicholson, L.K. (2007) *Biochemistry* **46**, 12174-12189.
[41] Hammes, G.G., Chang, Y.C. and Oas, T.G. (2009) *P. Natl. Acad. Sci. USA* **106**, 13737-13741.
[42] Masterson, L.R., Shi, L., Metcalfe, E., Gao, J., Taylor, S.S. and Veglia, G. (2011) *P. Natl. Acad. Sci. USA* **108**, 6969-6974.
[43] Deindl, S., Kadlecek, T.A., Cao, X., Kuriyan, J. and Weiss, A. (2009) *P. Natl. Acad. Sci. USA* **106**, 20699-20704.
[44] Cho, J.H., Muralidharan, V., Vila-Perello, M., Raleigh, D.P., Muir, T.W. and Palmer, A.G. (2011) *Nat. Struct. Mol. Biol.* **18**, 550-556.

2.10. Supplementary Information: Experimental Procedures

2.10.1. Molecular Dynamics Simulation Protocol

Initial Structure Preparation. Mouse EPAC2 was used for all MD simulations as this is the EPAC isoform for which experimental structures have been determined by X-ray crystallography in the absence and presence of the agonist S_p-cAMPS, which is a non-hydrolysable analog of cAMP.[1,2,3] Initial structures of the apo/inactive and holo/active states were obtained from X-ray crystal structures of mouse EPAC2 (PDB IDs 2BYV and 3CF6, respectively)[2,3,4] by deleting all amino acids other than residues 280-990 and all water molecules from both structures. In addition, SwissPDB Viewer[5] was used to reconstruct the residue 280-309 segment for the holo/active state, since the 3CF6 crystal structure was determined based on a Δ306 mutant of EPAC2.[3] To accomplish this, the apo/inactive and holo/active structures were superimposed using the backbone atoms of the residues in the segment of the α1 helix present in the holo/active structure (*i.e.* residues 311-314). This approach to the modeling of the 280-309 segment is justified by the fact that no significant structural changes were observed by NMR in this region upon cAMP binding.[1] Furthermore, the bound Sp-cAMPS molecule was replaced with the endogenous cAMP and an extra surface-bound Sp-cAMPS molecule was deleted from the structure as it was found in a crystal packing contact. Finally, the loops missing from the X-ray structures were reconstructed using the online ModLoop module of Modeller.[6] However, it was found that the REM-RA connecting loop could not be reconstructed due to its large size (*i.e.* 30 residues). Therefore, this loop was excluded from the simulations, although this was not expected to bias significantly the MD results, since the loop projects out from the protein surface.[2,3] To eliminate abnormal charges resulting from free N- and C-termini in this

region, the termini at either end of the missing loop were replaced by neutral capping groups (acetyl and N-methylamine). For the holo/active state, two trajectories were generated with and without the bound Rap1b substrate (which is included in the 3CF6 structure).

The initial structure for the apo/active metastate simulation was obtained from the initial structure of the holo/active state in the absence of Rap1b simply by deleting the bound cAMP. The initial structure for the holo/inactive metastate trajectory was obtained by docking a cAMP molecule into the initial structure of the apo/inactive state, by optimally superimposing the 3CF6 crystal structure onto the apo/inactive state initial structure using SwissPDB Viewer.[5] Superimposition was performed using the phosphate-binding cassette (PBC) loop (residues 403-416), since this segment forms critical hydrogen bonds with the ribose-phosphate moiety of bound cAMP.[3] A summary of the MD simulations performed for EPAC2(280-990) is given in Table S1.

MD Simulation Protocol. All MD simulations were performed using the NAMD 2.7 software[7] on the Shared Hierarchical Academic Research Computing Network (SHARCNET).[8] The CHARMm27 force field[9] was used for all simulations and the simulations were set up to mimic experimental conditions utilized in previous solution-state studies of EPAC:[1] a pH of 7.6; explicit water (with periodic boundary conditions) with a 50 mM concentration of NaCl; a constant temperature of 34°C (307 K); and a constant external pressure of 1 atm. Protein structure coordinate and parameter files (with hydrogen atoms) for the EPAC2 structures were constructed using the "Psfgen" module of VMD 1.8.6.[10] Parameters for cAMP were constructed from the parameters for the adenine ribonucleotide ("ADE" in the CHARMm force field) by applying the force field's intrinsic "CY35" patch

function and parameters for the acetyl and N-methylamine capping groups (see above) were constructed using the "ACE" and "CT3" patches, respectively.[9] In order to mimic a pH of 7.6, hydrogen atoms were added such that all His side chains were in their unionized τ-state and the N-/C-termini and all Asp, Glu, Arg and Lys side chains were in their ionized states. The structures were then immersed in a cubic box of TIP3P water molecules using the Solvate module of VMD 1.8.6,[10] ensuring a minimum distance of 12 Å between the protein and the edge of the solvent box. Salt ions (Na^+ and Cl^-) were added to the solvent box using the Autoionize module of VMD 1.8.6,[10] such that the system was neutralized and the effective NaCl concentration in the solvent was 50 mM.

Initial energy minimizations were performed using the conjugate gradient algorithm of NAMD. Minimization was performed for 5000 steps with harmonic position restraints on the protein backbone (force constant of 300.0 kcal/mol·Å^2), followed by an additional 2000 steps without restraints. During minimization, a cutoff of 15 Å was utilized for all non-bonded energy calculations. Electrostatic interactions beyond the cutoff distance were computed using the Particle Mesh Ewald (PME) algorithm,[11] with a tolerance of 10^{-6} and a maximum grid spacing of 1.0 Å. Molecular dynamics simulations were then performed under cubic periodic boundary conditions, beginning from the energy-minimized initial structures. A time-step of 1.0 fs was implemented throughout the simulations. All water molecules were constrained to their equilibrium geometries using the SETTLE algorithm[12] and all covalent bonds to hydrogen were constrained using the SHAKE algorithm.[13] A cutoff of 12 Å with PME implementation was utilized for non-bonded energy calculations during the simulations. Short-range non-bonded and long-range electrostatic interactions were evaluated every 2.0 fs and 4.0 fs, respectively, using the RESPA multiple timestep

integrator.[14] All minimizations and simulations were executed on a 2.83 GHz octuple-core Xeon cluster, using 72 CPUs per run. Execution of each simulation required a total of one month for each inactive state (~170,000 atoms with solvent) or two months for each active state (>300,000 atoms with solvent).

The structures were heated linearly from 0 K to 307 K over 200 ps at constant volume, using the velocity reassignment protocol of NAMD. The heated structures were then simulated at 307 K and constant volume (NVE ensemble) for another 1.0 ns, to allow a period of temperature equilibration prior to introducing pressure regulation. Next, the structures were simulated at a constant temperature and pressure (NPT ensemble) for 1.0 ns, to allow a period of temperature and pressure equilibration prior to the NPT production run. A constant temperature of 307 K was maintained using the Langevin dynamics algorithm,[15] with a Langevin damping coefficient of 1.0 ps^{-1}. A constant pressure of 1 atm (1.01325 bar) was maintained using the Nosé-Hoover Langevin piston method,[7] with a barostat oscillation period of 200.0 fs and a barostat damping time scale of 100.0 fs. Throughout the heating and equilibration runs, weak harmonic position restraints were imposed on the protein backbone (force constant of 5.0 kcal/mol·$Å^2$), to permit equilibration of the protein side chains and solvent without altering the protein backbone. Finally, production-run simulations were performed at a constant temperature and pressure (NPT ensemble) without restraints. These runs were executed for 60 ns, in order to obtain a 50 ns trajectory for analysis, while allowing for a final unrestrained equilibration period of 10 ns before the 50 ns trajectory. A constant temperature and pressure were maintained using the NPT protocol described above. During the production runs, structures were saved every 10000 timesteps (*i.e.* every 10.0 ps) for subsequent analysis.

2.10.2. Analysis of EPAC Structural Dynamics

Backbone root-mean-square deviations: stability of the overall protein topology. As an initial assessment of the MD simulations, full-protein backbone root-mean-square deviations (RMSDs) from the initial active ("open") and inactive ("closed") EPAC2 structures were examined over the course of the simulations. In addition, RMSDs of the individual domains were computed from their respective initial structures.

Backbone N-H order parameters: local dynamics and comparison with NMR data. To quantify the amplitudes of local structural fluctuations and permit a qualitative/semi-quantiative validation of the simulations against experiment, an analysis of backbone N-H order parameters (S^2) was performed for each trajectory. First, the generated structures from each simulation were optimally superimposed onto the energy-minimized initial simulation structure based on α-carbon coordinates, using the "orient" module of CHARMm 31.1.[16] Then, calculation of the S^2 order parameters was performed using the "NMR" module of CHARMm 31.1.[16] During the calculation, the NMR magnetic field strength parameter ("hfield") was set to a value of 16.44 T, in accordance with the magnetic field strength of the NMR spectrometer utilized in the NMR experiments.[1] In addition, the characteristic correlation time of overall protein tumbling ("rtumbl") was set to a value of 45820.0 ps for the inactive-conformation simulations, or 51470.0 ps for the active-conformation simulations. The aforementioned values of "rtumbl" were calculated based on the respective X-ray structures using the HYDRONMR 7c software.[17] The extensive flexibility at the EPAC N-terminus was accounted for in the HYDRONMR computations as previously

shown.[1] In the HYDRONMR calculations the temperature was set to 307 K and the magnetic field strength to 16.44 T.

A corresponding set of experimentally-derived order parameters (S^2) was obtained from the RCI algorithm,[18] using experimentally measured $^{13}C\alpha$, $^{13}C\beta$, ^{13}CO, ^{1}H, and ^{15}N chemical shifts for the apo and cAMP-bound states.[1] The RCI-based order parameters do not rely on specific models of overall rotation and therefore are ideally suited for constructs with extensive flexible regions that may affect the rate and anisotropy of the overall tumbling, such as the EPAC construct used for NMR.[1] Another distinct advantage of experimentally assessing dynamics through the RCI index is that the RCI is measurable also for the critical phosphate binding cassette (PBC) region, which in the apo state is subject to considerable line-broadening compromising the signal-to-noise ratio of traditional ^{15}N relaxation experiments. In addition, the trends in the cAMP-dependent variations of S^2 were also supplemented by previously measured high-frequency reduced spectral densities (*i.e.* $J(\omega_H+\omega_N)$),[1] which sense the fast sub-ns dynamics that are well sampled by the ns trajectories of the MD simulations. All experimental NMR data were collected for an EPAC construct composed of residues 149-318 of the homologous EPAC1h isoform, which span the EPAC1 CNB domain.[1]

Time profiles of key inter-atomic contacts: EPAC2 interfacial dynamics. In order to assess the stability of key interfaces between the CNB domain and the catalytic region, time profiles of several selected inter-atomic distances were computed over the course of the simulations. The distances examined included a number of hydrogen bond interactions identified previously in the X-ray crystal structures,[2,3] as well as several hydrophobic

contacts identified in the holo/active X-ray structure[3] using the Ligand Explorer software with an hydrophobic contact distance cutoff of 4.0 Å.[19]

Dynamic cross-correlation matrices: correlated fluctuations. To examine patterns of correlation among residue movements, a dynamic cross-correlation matrix (DCCM)[20,21] analysis was performed on each MD trajectory. The analysis was carried out as follows. First, the generated structures from each simulation were optimally superimposed onto the energy-minimized initial simulation structure based on α-carbon coordinates, using the "orient" module of CHARMm 31.1.[16] In the subsequent DCCM calculation, the average Cartesian coordinates of the α-carbon atoms were computed over all structures in the trajectory. Then, the DCCM of correlation coefficients was calculated as follows:[21]

$$C_{i,j} = \frac{\left\langle (r_i - \langle r_i \rangle) \cdot (r_j - \langle r_j \rangle) \right\rangle}{\sqrt{\left\langle (r_i - \langle r_i \rangle)^2 \right\rangle \cdot \left\langle (r_j - \langle r_j \rangle)^2 \right\rangle}} \qquad \text{(S1)}$$

where $C_{i,j}$ is the correlation coefficient describing the extent of correlation between the movements of α-carbon atoms *i* and *j*; r_i and r_j are the vectors of the Cartesian coordinates of the α-carbon atoms *i* and *j*, respectively, at a given time point during the trajectory; $\langle r_i \rangle$ and $\langle r_j \rangle$ are the average Cartesian coordinates of α-carbon atoms *i* and *j*, respectively, computed over the whole trajectory; the angle brackets denote averages computed over the whole trajectory. In the DCCM plots, a correlation coefficient of 1.0 corresponds to perfect correlation of α-carbon atom motions, while a correlation coefficient of -1.0 corresponds to perfect anti-correlation.[21] A correlation coefficient of 0.0 corresponds to perfectly uncorrelated motions.[21]

Differential root-mean-square fluctuation analysis: long-range domain motions. To quantify the amplitudes of long-range inter-domain motions, an analysis of differential root-mean-square fluctuations (RMSFs) was performed on each MD trajectory. In this analysis, the generated structures from each simulation were optimally superimposed onto the energy-minimized initial simulation structure based on α-carbon coordinates, using the "orient" module of CHARMm 31.1.[16] The RMSF calculation was subsequently performed by first computing the average Cartesian coordinates of the α-carbon atoms over all structures in the trajectory. Then, the RMSFs around the average coordinates were calculated as follows:

$$RMSF(i) = \sqrt{\left\langle (r_i - \langle r_i \rangle)^2 \right\rangle} \qquad \text{(S2)}$$

where *RMSF(i)* is the RMSF describing the fluctuations of residue *i* and, similarly to equation S1, r_i is the vector of the Cartesian coordinates of the α-carbon atom of residue *i* at a given time point during the trajectory; $<r_i>$ is the average Cartesian coordinates of the α-carbon atom of residue *i*, computed over the whole trajectory; the angle brackets denote averages computed over the whole trajectory. This process was performed first for the complete EPAC2 construct to obtain global RMSF values, which report on both local and non-local motions, and then for each of the individual EPAC2 domains to obtain domain-specific RMSF values, which report predominantly on local motions. For the domain-specific RMSF calculations, the following domain boundaries were utilized in the superimpositions: residues 300-460 for CNB; 480-612 for REM; 643-740 for RA; and 741-990 for CDC25-HD. Finally, the differential RMSF values for each residue were obtained by subtracting the domain-specific RMSF values from the respective global RMSF values. The

differential RMSF values are therefore expected to report mainly on non-local 'domain' motions.

It should be noted that, although the domain-specific RMSF values display an overall trend similar to that of the S^2 order parameters (Fig. S2), the RMSF and S^2 report on slightly different features of the local dynamics. The order parameters quantify the dynamics of the N-H bond vector orientation, while the RMSFs probe the oscillations of the α-carbon atom position. As a result motions that do not affect the orientation of the N-H bond vector are mute to the S^2 order parameters but not necessarily to the RMSFs, while fluctuations that do not alter the α-carbon position escape the RMSF detection but may still be registered by the S^2 values. For the sake of completeness we have therefore reported both the S^2 and the RMSF analyses of the MD simulations (Figs. 3 and S2).

Principal component analysis: collective residue movements. To further examine long-range movements within the protein, an α-carbon-coordinate principal component analysis (PCA) was performed on each MD trajectory. The analysis was carried out as follows. Each simulation was analyzed using the "g_covar" and "g_anaeig" analysis tools of GROMACS, based on α-carbon coordinates of the generated structures after optimal superposition onto the respective initial structures. The 3N x 3N mass-weighted covariance matrix for the N α-carbon atoms considered was computed as follows:

$$C_{i,j} = \sqrt{M_i M_j}\left\langle r_i - \left\langle r_i \right\rangle \right\rangle \left\langle r_j - \left\langle r_j \right\rangle \right\rangle \qquad \text{(S3)}$$

where $C_{i,j}$ is the covariance of coordinates r_i and r_j during the trajectory; M_i and M_j are the masses of α-carbon atoms *i* and *j*; similarly to the previous equations r_i and r_j are the x-, y-, or z-coordinate of α-carbon atoms *i* and *j*, respectively, at a given time point during the

trajectory; $<r_i>$ and $<r_j>$ are the corresponding average x-, y-, or z-coordinates of α-carbon atoms *i* and *j*, respectively, computed over the whole trajectory. During the analysis, it was found that the choice of protein fragment used for structural superimposition significantly affected the results. Therefore, the highly-floppy N-terminal helix and CNB-REM domain connecting loop were not considered during superimposition (*i.e.* only residues 300-460 and 480-990 were superimposed), in order to facilitate examination of dynamics in other regions.

The mass-weighted covariance matrix was then diagonalized and eigenvalues and eigenvectors sorted by eigenvalue magnitude were obtained. It was found that examination of the first six eigenvectors (which accounted for 68-92% of the total variance in the different simulations, as reflected by the trace of the covariance matrix) were sufficient for examination of long-range collective motions during the simulations. In order to get a better understanding of the motions corresponding to these largest eigenvectors, the trajectory was projected onto the multidimensional space spanned by the eigenvectors and visually examined. In addition, extreme conformations (*i.e.* those having the largest and smallest projection on each of the six largest eigenvectors) were computed and visualized.

2.11. Supplementary Information: Tables and Figures

Table S1. *Summary of the MD Simulations Performed for EPAC2(280-990)*

State	PDB X-ray Crystal Structure	N-terminal Construction Required? (Y/N)	Contains Bound cAMP? (Y/N)	Contains Bound Rap1b? (Y/N)	Simulation Length (ns)	No. of CPUs Used	Execution Time (months)[a]
Apo/ Inactive	2BYV	N	N	N	50.0	72	1
Holo/ Active	3CF6	Y	Y	N	50.0	72	2
Apo/ Active	3CF6	Y	N	N	50.0	72	2
Holo/ Inactive	2BYV	N	Y	N	50.0	72	1
Holo/ Active + Rap1b	3CF6	Y	Y	Y	50.0	72	2

[a] The large difference in execution times for inactive- and active-state simulations was due to the fact that the solvent box around the protein was larger for the active-state simulations. This was necessary in order to accommodate the more extended active-state conformation within the solvent box.

Table S2. *Stable Hydrogen-Bond Interactions between EPAC2 and cAMP in the Holo/Active State of EPAC2(280-990)*

Hydrogen Bond[a]	X-ray Distance (Å)	MD Distance Average (Å)[b]	MD Distance Range (Å)
G404 N→cAMP O2'	3.24	3.0 (0.19)	2.6 – 3.5
A407 N→cAMP O_{Eq}	2.80	3.7 (0.39)	2.7 – 4.7
R414 NH1→cAMP O_{Eq}	2.77	2.8 (0.14)	2.5 – 3.1
R414 NH2→cAMP O_{Eq}	5.07	4.9 (0.15)	4.6 – 5.3
A415 N→cAMP O_{Ax}	3.34	3.0 (0.29)	2.5 – 3.7
K450 O→cAMP N6	2.83	2.9 (0.15)	2.5 – 3.3

[a] All interactions are described as atom pairs, which are listed in accordance with PDB atom notations.
[b] The corresponding standard deviations are given in brackets.

Table S3. *Stable Hydrophobic Interactions between EPAC2 and cAMP in the Holo/Active State of EPAC2(280-990)*

Hydrophobic Contact[a]	X-ray Distance (Å)	MD Distance Average (Å)[b]	MD Distance Range (Å)
A415 CB→cAMP C2	3.96	3.9 (0.37)	3.2 – 4.8
A416 CB→cAMP C2	3.85	3.9 (0.34)	3.1 – 4.7
A416 CB→cAMP C5’	3.90	4.1 (0.28)	3.3 – 4.8
L449 CB→cAMP C6	3.61	4.0 (0.24)	3.4 – 4.5
L449 CB→cAMP C5	3.80	4.0 (0.33)	3.4 – 4.5
L449 CD1→cAMP C2	3.98	4.5 (0.50)	3.3 – 5.7
L449 CD2→cAMP C5	3.92	4.0 (0.30)	3.2 – 4.8
L449 CD2→cAMP C4	3.66	3.9 (0.31)	3.1 – 4.7
L449 CD2→cAMP C2’	3.89	4.1 (0.45)	3.5 – 4.9
F367 CE1→cAMP C5’	3.71	4.0 (0.27)	3.4 – 4.6
F367 CE2→cAMP C5’	5.29	5.8 (0.42)	4.7 – 7.0
F367 CZ→cAMP C5’	3.90	4.5 (0.42)	3.4 – 5.5
V386 CG1→cAMP C5	3.77	4.1 (0.28)	3.3 – 4.9
V386 CG2→cAMP C5	5.02	5.2 (0.44)	4.0 – 6.4
V386 CG1→cAMP C4	3.62	3.9 (0.25)	3.2 – 4.5
V386 CG2→cAMP C4	4.86	5.1 (0.40)	4.0 – 6.2

[a] All interactions are described as atom pairs, which are listed in accordance with PDB atom notations.

[b] The corresponding standard deviations are given in brackets.

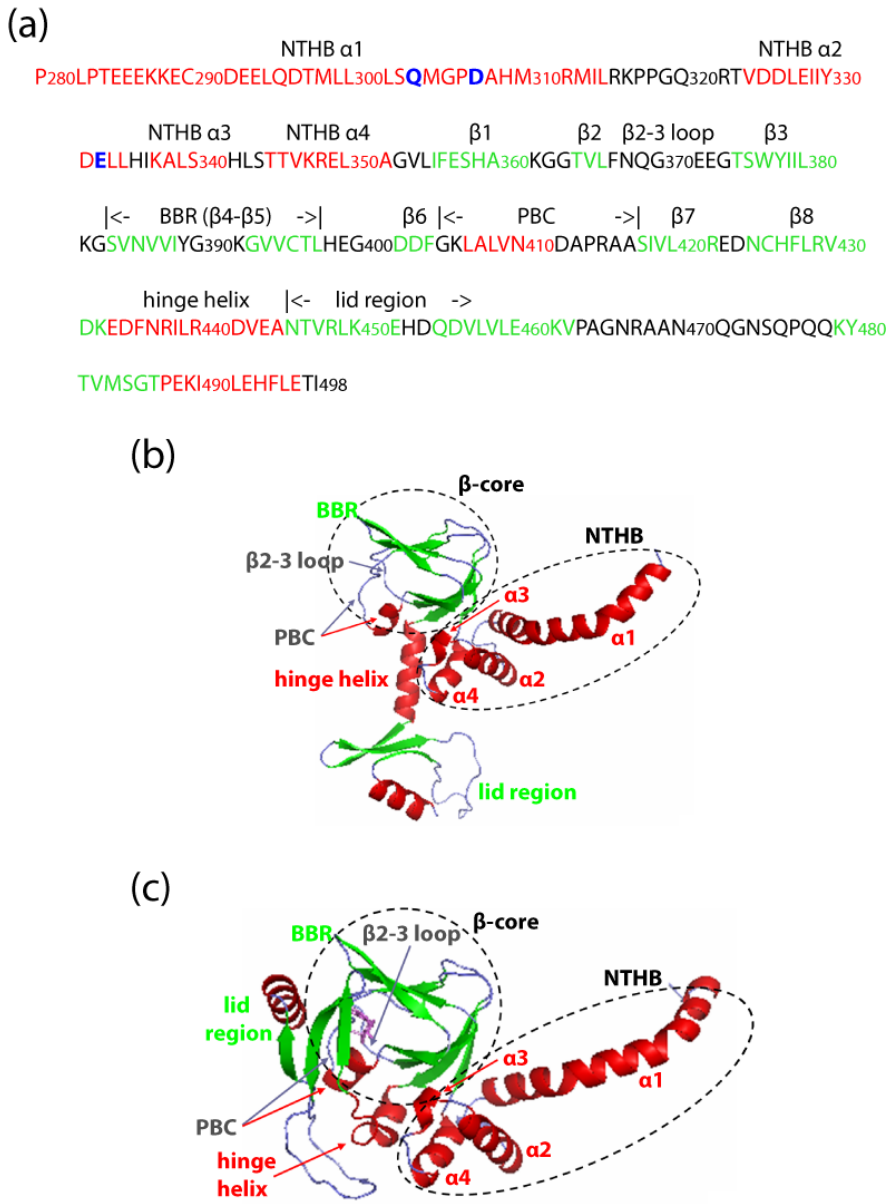

Figure S1: (a) Amino acid sequence of the allosteric CNB domain and lid. The α-helix elements are highlighted in red and the β-strand elements in green. The positions of key structural elements are indicated above the sequence: the N-terminal helical bundle (NTHB) α-helices (with ionic-latch residues highlighted in blue); the eight β-strands of the β-core; the β2-3 loop; the phosphate binding cassette (PBC) and base-binding region (BBR), which bind the cAMP; the hinge helix and the lid region. Note that the BBR is formed by residues from the β4 and β5 strands and that the last β-strand and α-helix of the lid are considered part of the REM domain. **(b,c)** Ribbon-structure illustrations of the **(b)** apo/inactive and **(c)** holo/active states of the allosteric CNB domain and lid, showing the arrangements of key

structural elements in the inactive and active conformations. As in panel (a), α-helices are highlighted in red and β-strands in green. Also, the bound cAMP is shown as a purple stick model and the positions of key structural elements are indicated similarly to panel (a): the N-terminal helical bundle (NTHB) α-helices; the β-core; the β2-3 loop; the base-binding region (BBR); the phosphate binding cassette (PBC); the hinge helix and the lid region. Note that the C-terminal end of the hinge helix becomes unraveled in the active conformation. The structures were generated using PyMol (Schrödinger, LLC), based on the 2BYV and 3CF6 structures from the RCSB Protein Data Bank.

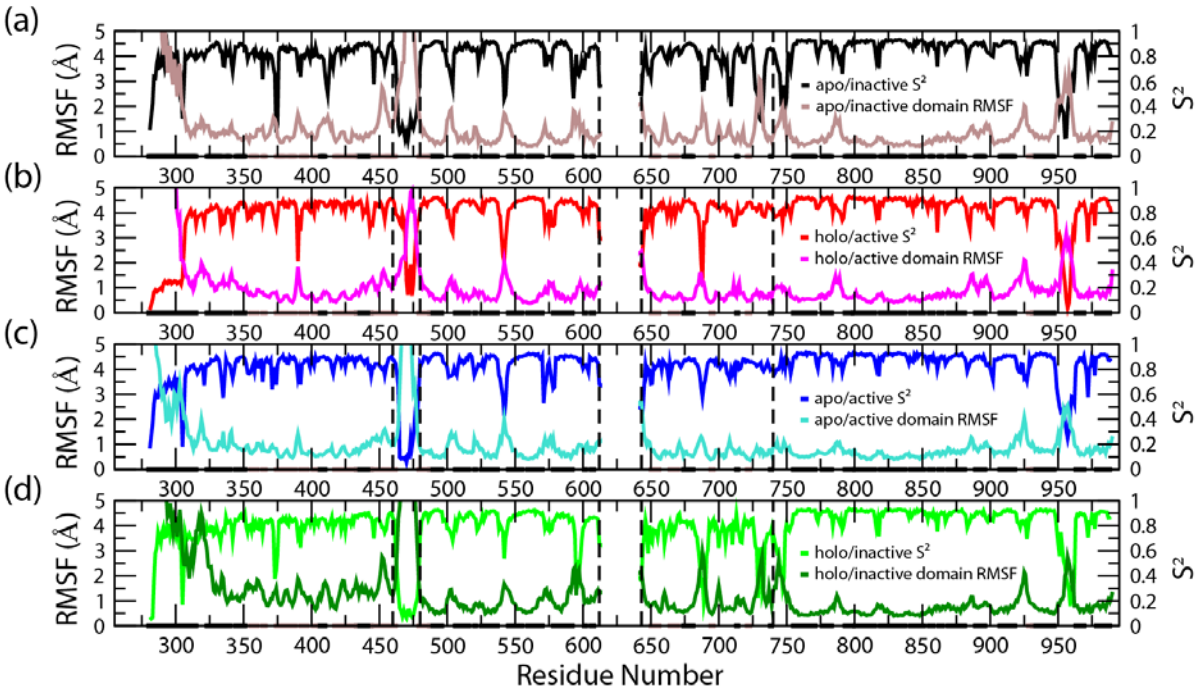

Figure S2: Comparison of the backbone N-H order parameter (S^2) values and the root-mean-square fluctuation (RMSF) values obtained for the EPAC2(280-990) construct by superimposition of each individual domain (*i.e.* domain-specific RMSF values): **(a)** apo/inactive state; **(b)** holo/active state; **(c)** apo/active metastate; **(d)** holo/inactive metastate. The domain boundaries are indicated by vertical dashed lines on all plots. It should be noted that although some minor discrepancies were observed between the two sets of results, such discrepancies were not unexpected since the calculated RMSF values probe α-carbon atom coordinate variations in three-dimensional space, while the calculated S^2 values probe orientational variations of backbone N-H bond vectors.

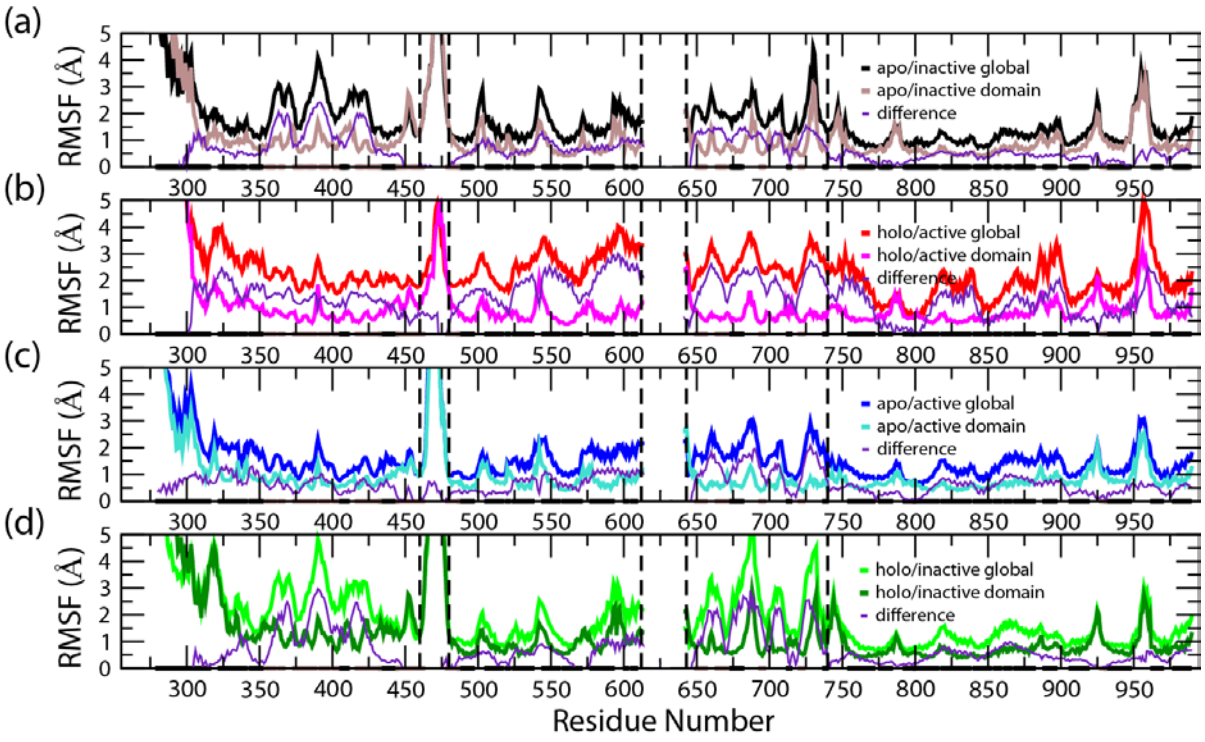

Figure S3: Comparison of root-mean-square fluctuation (RMSF) values obtained for EPAC2(280-990) by superimposing all domains of the protein (*i.e.* global RMSF values) *vs.* superimposing each domain individually (*i.e.* domain-specific RMSF values). The corresponding differential RMSF values (*i.e.* the global RMSF values minus the domain-specific RMSF values) are also reported. **(a)** Apo/inactive state; **(b)** holo/active state; **(c)** apo/active metastate; **(d)** holo/inactive metastate. The domain boundaries are indicated by vertical dashed lines on all plots.

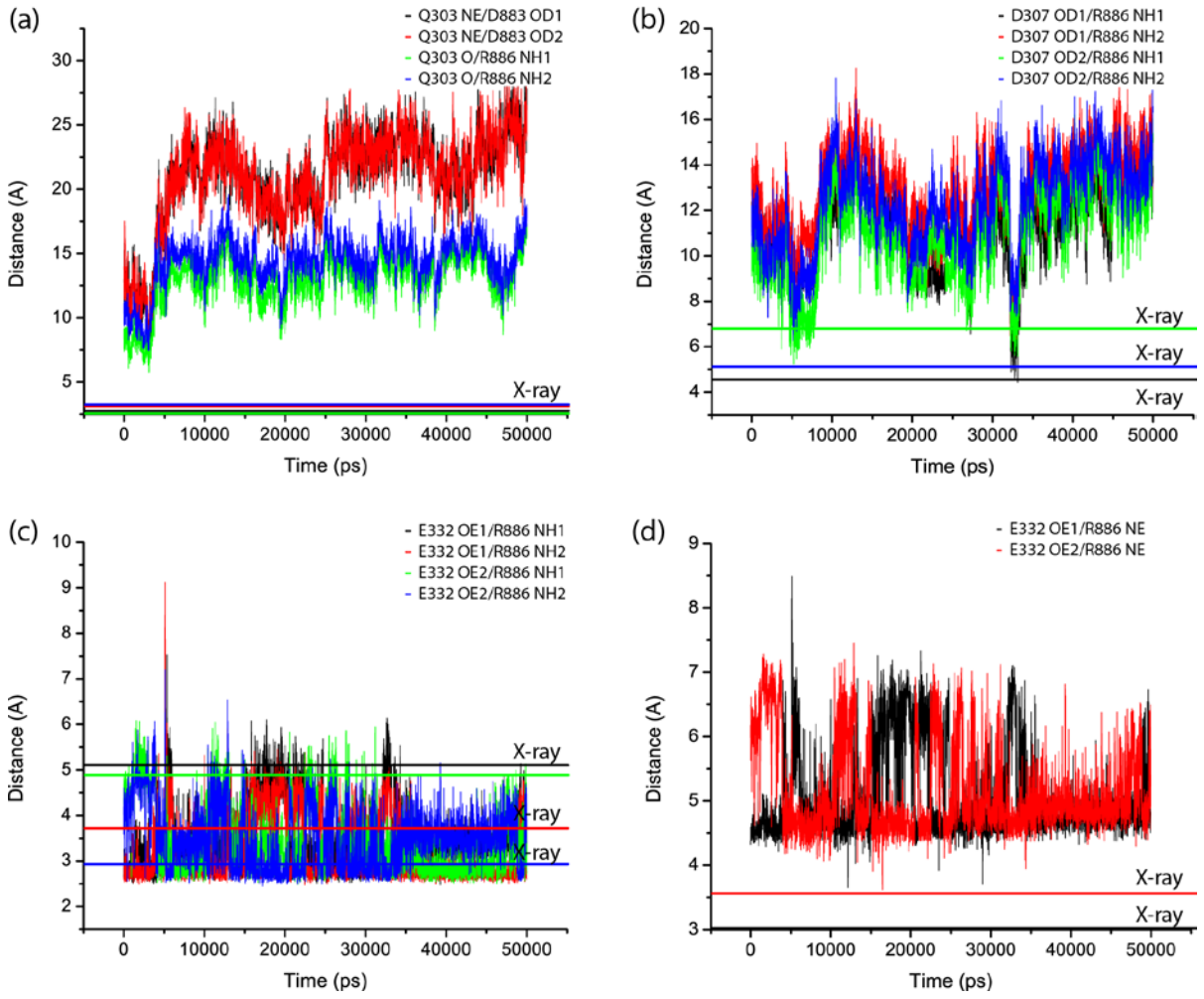

Figure S4: Contact distance *vs.* time plots for interactions in the ionic-latch (IL) interface of the apo/inactive state of EPAC2(280-990). For each contact, the corresponding distance value in the X-ray structure is also indicated on the graph as a straight, horizontal line, color-coded in the same manner as the respective simulation data. The color code is reported in each panel. **(a)** Interactions of NTHB residue Q303 with CDC25-HD domain residues D883 and R886; **(b)** interactions of NTHB residue D307 with CDC25-HD domain residue R886; **(c,d)** interactions of NTHB residue E332 with CDC25-HD domain residue R886.

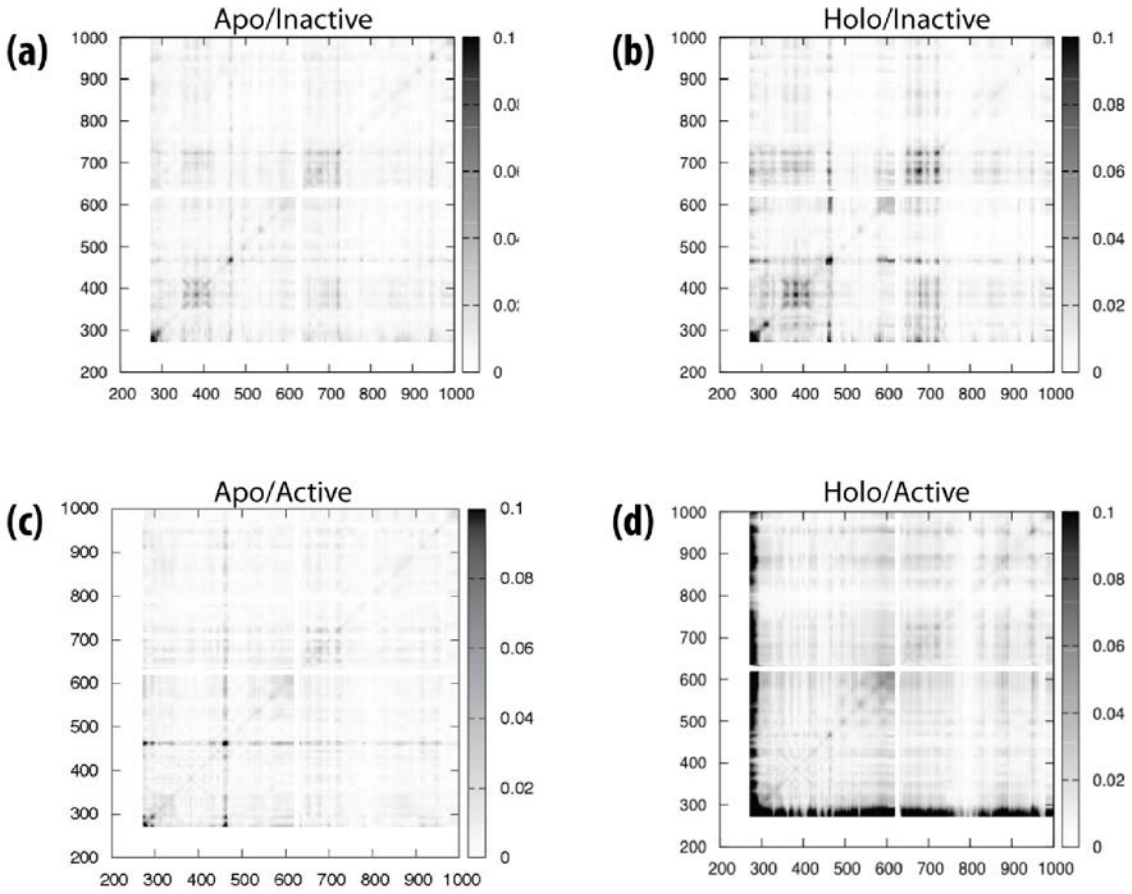

Figure S5: Covariance matrices for the **(a)** apo/inactive, **(b)** holo/inactive, **(c)** apo/active and **(d)** holo/active MD trajectories of EPAC2(280-990). Covariances of α-carbon coordinates (in nm^2) are plotted for each pair of amino acid residues. Values between 0.0 and 0.1 nm^2 are plotted in grayscale (where white and black correspond to 0.0 nm^2 and 0.1 nm^2, respectively) and values outside this range are displayed in black.

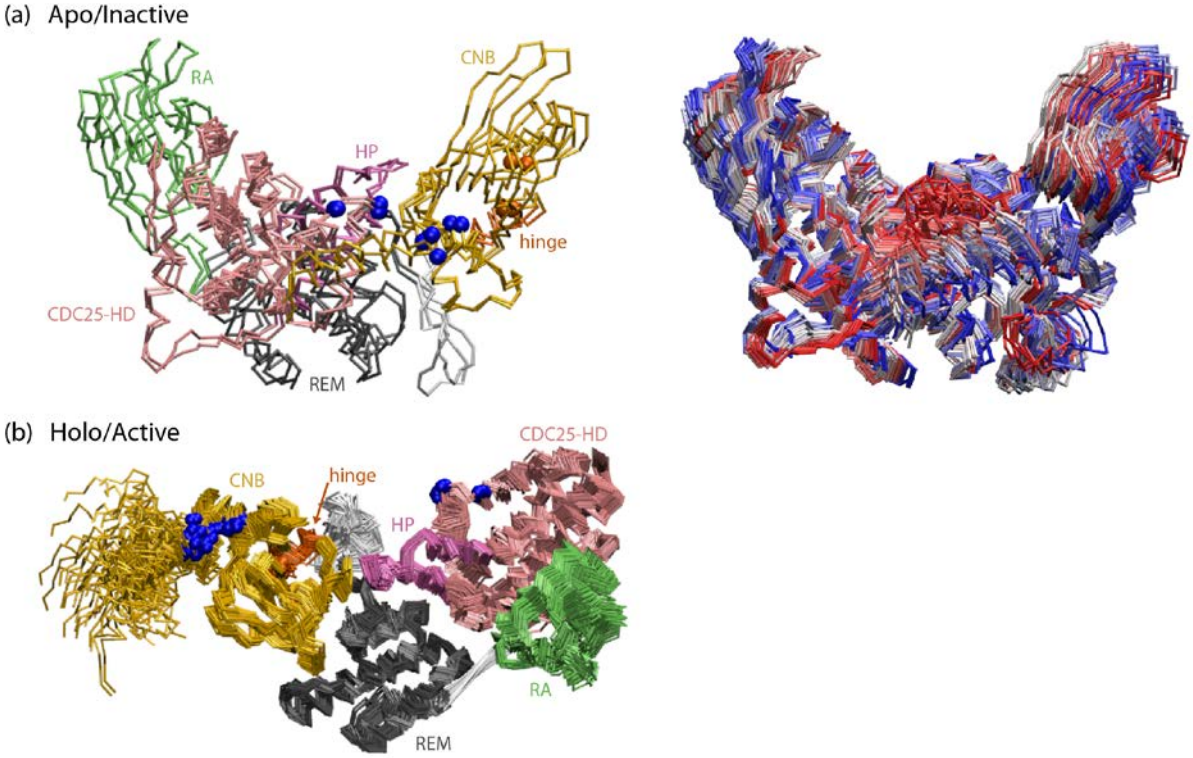

Figure S6: (a, left panel) Backbone trace representation of the extreme projections of the apo/inactive-state trajectory onto the first eigenvector computed by PCA. The CNB domain is indicated in yellow, the hinge helix in orange, the REM domain in dark grey, the RA domain in green, the CDC25-HD domain in light pink and the helical hairpin (HP) in dark pink. The ionic-latch residues are indicated as blue spheres and the PBC-hinge interface residues (*i.e.* L408 and F435) are indicated as orange spheres. An anti-correlated breathing motion of the CNB and RA domains is apparent from the projections and is further illustrated by the variations in structure over the course of the trajectory, as shown in panel **(a, right)**, where the frame color varies from red to blue with progression of the simulation time. **(b)** Snapshots from the holo/active-state trajectory projected onto the multidimensional space spanned by the first six eigenvectors. The color coding is the same as in panel (a, left).

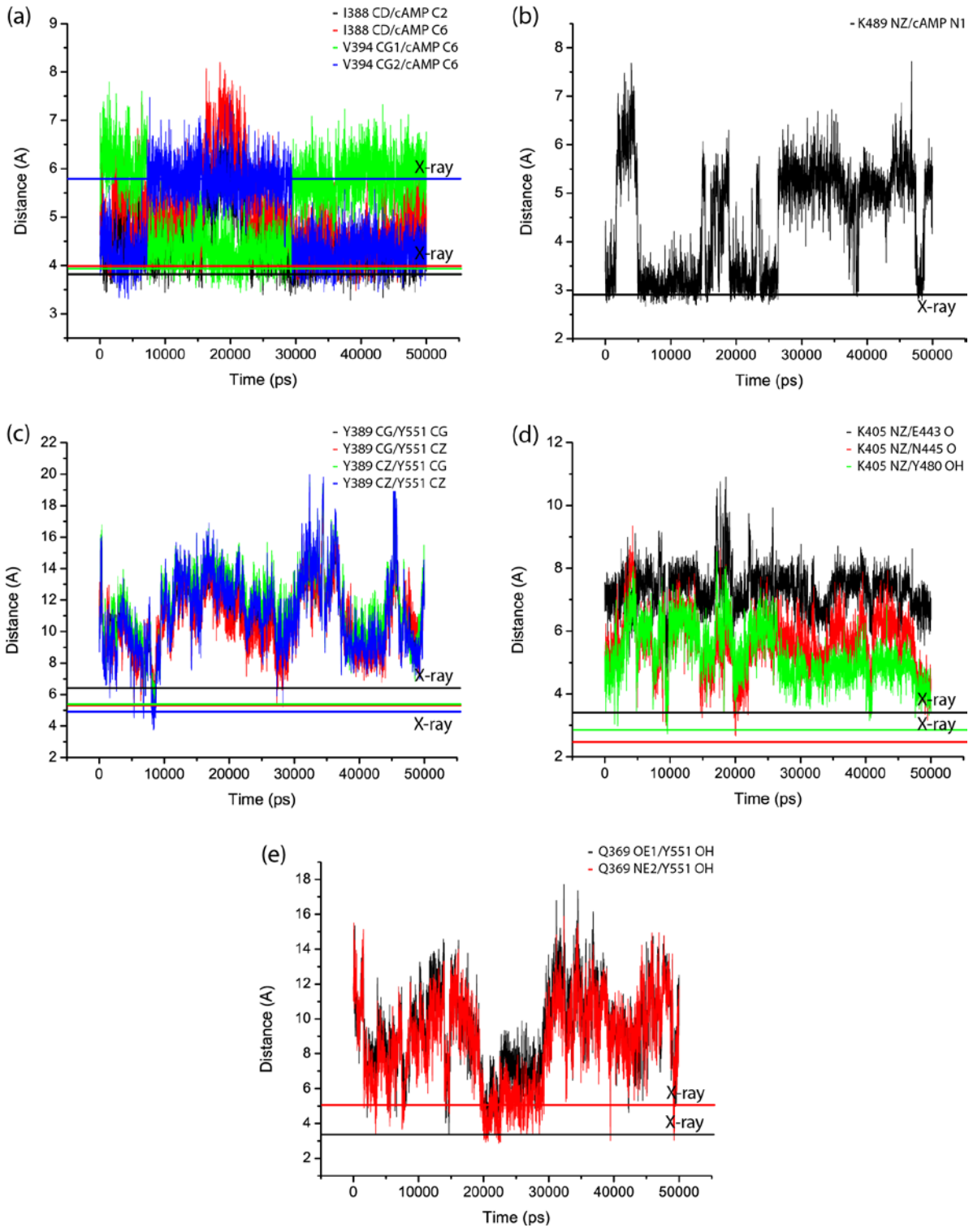
(a)
I388 CD/cAMP C2
I388 CD/cAMP C6
V394 CG1/cAMP C6
V394 CG2/cAMP C6
X-ray
X-ray
Distance (A)
Time (ps)
(b)
K489 NZ/cAMP N1
X-ray
Distance (A)
Time (ps)
(c)
Y389 CG/Y551 CG
Y389 CG/Y551 CZ
Y389 CZ/Y551 CG
Y389 CZ/Y551 CZ
X-ray
X-ray
Distance (A)
Time (ps)
(d)
K405 NZ/E443 O
K405 NZ/N445 O
K405 NZ/Y480 OH
X-ray
X-ray
Distance (A)
Time (ps)
(e)
Q369 OE1/Y551 OH
Q369 NE2/Y551 OH
X-ray
X-ray
Distance (A)
Time (ps)

Figure S7: Contact distance *vs.* time plots for dynamic interactions in the cAMP/CNB and CNB/REM interfaces of the holo/active state of EPAC2(280-990). For each contact, the corresponding distance value in the X-ray structure is also indicated on the graph as a straight, horizontal line, color-coded in the same manner as the respective simulation data. The color coding is reported in each panel. **(a)** Hydrophobic contacts of BBR residues I388 and V394 with cAMP; **(b)** interaction of K489 with cAMP; **(c)** hydrophobic contacts between BBR residue Y389 and REM domain residue Y551; **(d)** interactions of PBC residue K405 with lid residues E443, N445 and Y480; **(e)** interactions of β2-3 loop residue Q369 with REM domain residue Y551.

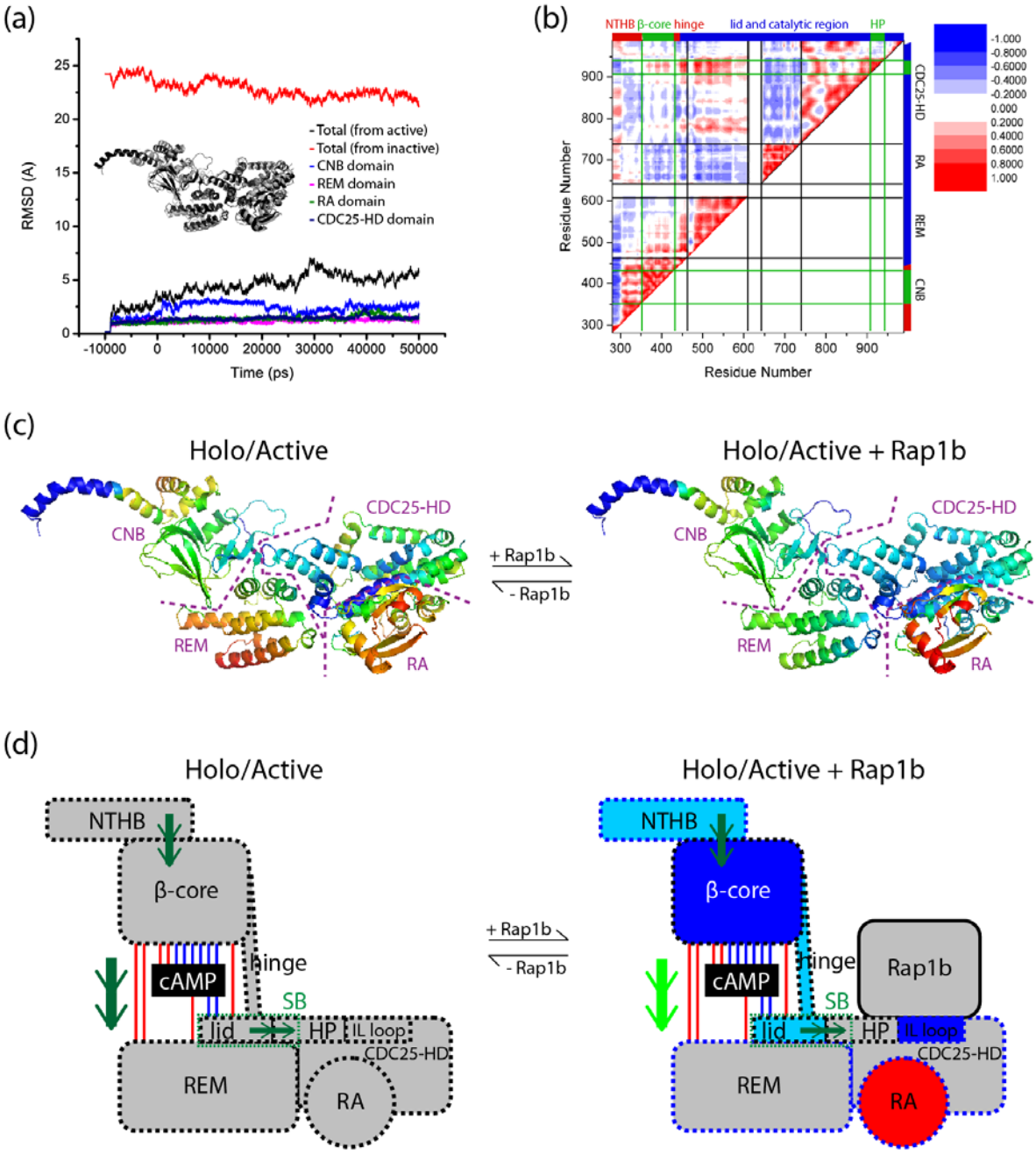

Figure S8: (a) Root-mean-square deviations (RMSDs) of EPAC2(280-990) from the initial structures computed over the course of the MD simulation for the holo/active state with bound Rap1b. The figure layout is similar to that of Figure 2. **(b)** Dynamic cross-correlation matrix (DCCM) computed for the holo/active state of EPAC2(280-990) with bound Rap1b. Terminology follows a similar scheme to that of Figure 5. **(c)** Ribbon-structure illustration of the differential RMSF values computed for the holo/active state of EPAC2(280-990) with

and without bound Rap1b. The structures are displayed in a manner similar to that of Figure 6. **(d)** Schematic outline of the major results obtained from analysis of the holo/active state of EPAC2(280-990) with and without bound Rap1b. The following key structural elements are examined: the N-terminal helical bundle (NTHB), β-core and hinge helix of the CNB domain; the switchboard (SB; highlighted by a green, dotted rectangle), which is formed by the lid β-sheet and helical hairpin (HP) and anchors the CNB, REM and CDC25-HD domains together; the ionic latch (IL); the REM, RA and CDC25-HD domains and the bound Rap1b. The terminology follows a similar scheme to that of Figure 7, except that changes in the local and inter-domain dynamics due to Rap1b binding are now illustrated and reported relative to the holo/active state, rather than to the apo/inactive state.

2.12. Supplementary References

[1] Das, R., Mazhab-Jafari, M.T., Chowdhury, S., Das, S.S., Selvaratnam, R. and Melacini, G. (2008) *J. Biol. Chem.* **283**, 19691-19703.
[2] Rehmann, H., Das, J., Knipscheer, P., Wittinghofer, A. and Bos, J.L. (2006) *Nature* **439**, 625-628.
[3] Rehmann, H., Arias-Palomo, E., Hadders, M.A., Schwede, F., Llorca, O. and Bos, J.L. (2008) *Nature* **455**, 124-127.
[4] The RCSB Protein Data Bank: http://www.rcsb.org/.
[5] The Swiss Institute of Bioinformatics. Swiss-PdbViewer Version 4.0.1. (2008).
[6] Fiser, A. and Sali, A. (2003) *Bioinformatics* **19**, 2500-2501.
[7] Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L. and Schulten, K. (2005) *J. Comput. Chem.* **26**, 1781-1802.
[8] The Shared Hierarchical Academic Research Computing Network (SHARCNET): www.sharcnet.ca.
[9] MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D. and Karplus, M. (1998) *J. Phys. Chem. B* **102**, 3586-3616.
[10] Theoretical Biophysics Group, Beckman Institute. VMD Version 1.8.6. (2007).
[11] Darden, T., York, D. and Pederson, L. (1993) *J. Chem. Phys.* **98**, 10089-10092.
[12] Miyamoto, S. and Kollman, P.A. (1992) *J. Comput. Chem.* **13**, 952-962.
[13] Ryckaert, J.P., Ciccotti, G. and Berendsen, H.J.C. (1977) *J. Comput. Phys.* **23**, 327–341.
[14] Tuckerman, M., Berne, B.J. and Martyna, G.J. (1992) *J. Chem. Phys.* **97**, 1990-2001.
[15] Izaguirre, J.A., Catarello, D.P., Wozniak, J.M. and Skeel, R.D. (2001) *J. Chem. Phys.* **114**, 2090-2098.
[16] Brooks, B.R., Brooks III, C.L., Mackerell Jr., A.D., Nilsson, L., Petrella, R.J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A.R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R.W., Post, C.B., Pu, J.Z., Schaefer, M., Tidor, B., Venable, R.M., Woodcock, H.L., Wu, X., Yang, W., York, D.M. and Karplus, M. (2009) *J. Comput. Chem.* **30**, 1545-1614.
[17] García de la Torre, J., Huertas, M.L. and Carrasco, B. (2000) *J. Magn. Reson.* **147**, 138-146.
[18] Berjanskii M.V. and Wishart, D.S. (2005) *J. Am. Chem. Soc.* **127**, 14970-14971.
[19] Ligand Explorer. Accessed at the RCSB Protein Data Bank: http://www.rcsb.org/.
[20] Arnold, G.E. and Ornstein, R.L. (1997) *Biophys. J.* **73**, 1147-1159.
[21] Moritsugu, K. and Smith, J.C. (2008) *Biophys. J.* **95**, 1639-1648.

CHAPTER 3

ROLE OF DYNAMICS IN THE AUTO-INHIBITION AND ACTIVATION OF THE HYPERPOLARIZATION-ACTIVATED CYCLIC-NUCLEOTIDE-MODULATED (HCN) ION CHANNELS

3.1. Author's Preface

The work presented in this chapter has previously been published and is reproduced here with permission from the Journal of Biological Chemistry. The full citation is as follows:

VanSchouwen, B., Akimoto, M., Sayadi, M., Fogolari, F. and Melacini, G. (2015) "Role of dynamics in the auto-inhibition and activation of the hyperpolarization-activated cyclic-nucleotide-modulated (HCN) ion channels." *Journal of Biological Chemistry* **290**: 17642-17654.

I conducted most of the work necessary for this publication. Madoka Akimoto performed the experimental HCN work that was used to validate and complement the molecular dynamics data. I co-wrote the manuscript with Drs. Maryam Sayadi, Federico Fogolari and Giuseppe Melacini.

3.2. Abstract

The hyperpolarization-activated cyclic-nucleotide-modulated (HCN) ion channels control rhythmicity in neurons and cardiomyocytes. Cyclic AMP (cAMP) allosterically modulates HCN through the cAMP-dependent formation of a tetrameric gating ring spanning the intracellular region (IR) of HCN, to which cAMP binds. Although the apo-*vs.*-holo conformational changes of the cAMP-binding domain (CBD) have been previously mapped, only limited information is currently available on the HCN IR dynamics, which have been hypothesized to play a critical role in the cAMP-dependent gating of HCN. Here, using MD

simulations validated and complemented by experimental NMR and CD data, we comparatively analyze HCN IR dynamics in the four states of the thermodynamic cycle arising from the coupling between cAMP-binding and tetramerization equilibria. This extensive set of MD trajectories captures the active-to-inactive transition that had remained elusive for other CBDs, and provides unprecedented insight on the role of IR dynamics in HCN auto-inhibition and its release by cAMP. Specifically, the IR tetramerization domain becomes more flexible in the monomeric states, removing steric clashes that the apo CBD structure would otherwise impose. Furthermore, the simulations reveal that the active/inactive structural transition for the apo-monomeric CBD occurs through a manifold of pathways that are more divergent than previously anticipated. Upon cAMP-binding, these pathways become disallowed, pre-confining the CBD conformational ensemble to a tetramer-compatible state. This conformational confinement primes the IR for tetramerization, and thus provides a model of how cAMP controls HCN channel gating.

3.3. Introduction

The hyperpolarization-activated cyclic-nucleotide-modulated (HCN) ion channels are key determinants of nerve impulse transmission and heart rate modulation in neuronal and cardiac cells, respectively.[1,2,3,4,5,6,7,8,9,10,11,12] The HCN channels exist in four known isoforms, HCN1-4, of which HCN2 and HCN4 are the isoforms most significantly regulated by cAMP.[1,2,5,7,9,10,13] All isoforms contain an N-terminal trans-membrane region (TM), which assembles as a tetramer harboring the ion pore, and a C-terminal intracellular region (IR) that confers channel regulation by cAMP (Figure 1a-b).[1,2,5,7,9,10] The cAMP-induced up-regulation of ion channel opening has been suggested to be closely linked to

tetramerization of the HCN IR monomers in response to cAMP binding. However, it is currently not fully understood how cAMP promotes tetramerization of the HCN IR.[3,14,15,16]

The HCN IR tetramer adopts an 'elbow-shoulder' topology[3] involving four contiguous α-helices (αA'-αD') at the IR N-terminus, which are referred to as tetramerization domain (TD) (Figure 1a-b). The TD is followed by a C-terminal cAMP-binding domain (CBD), which allosterically controls the self-association of the TD. In the absence of cAMP the TD tetramer is unstable, while in the presence of cAMP the TD tetramerizes.[3] In a recent study performed by Akimoto *et al.* (2014),[17] the apo-state structure of the HCN4 CBD was solved by NMR, and was found to differ from the structure of the cAMP-bound CBD[5] by a rearrangement of the CBD α-helical structure elements similar to that observed for CBDs of other eukaryotic cAMP receptors.[18,19,20,21,22,23,24,25,26] The study proposed a model of HCN4 allostery in which the apo-state CBD conformation destabilizes the tetramer through steric clashes with the TD, arising from the incompatibility of the apo-state CBD conformation with the tight packing imposed by the tetramer assembly.[17] It has been hypothesized that the CBD/TD steric clashes are released through an increase in TD structural dynamics occurring upon dissociation of the tetramer.[17] This hypothetical dynamics-based model provides a simple and effective mechanism to explain how cAMP controls IR tetramerization and contributes to ion channel gating. Nevertheless, only limited information is currently available on the TD dynamics, and on the inactive/active transition within the CBD, even though such attributes play a critical role in the proposed model.

Monomer TD dynamics are not easily experimentally assessed in the currently available HCN IR monomer constructs, since the αA'-αB' helices were deleted to stabilize the monomeric state, and the αC'-αD' helices are possibly affected by the construct's N-

terminal truncation as well as by line-broadening in the NMR spectra.[17] Furthermore, while the cAMP-bound and apo CBD structures of the HCN4 IR were experimentally determined, providing the start- and end-points of the active/inactive transition, the pathways and structural variations that occur between these structures have eluded experimental analysis so far.

Here we examine the structural dynamics of the HCN4 intracellular region through molecular dynamics (MD) simulations in explicit solvent. Multiple MD trajectories were simulated for each of the four states of the thermodynamic cycle describing the coupling between cAMP-binding and tetramerization of the HCN4 intracellular region (*i.e.* the apo-monomer or inactive state, the apo-tetramer, the holo-monomer and the holo-tetramer or active state; Figure 1c), resulting in a total of 1.2 μs of simulated trajectory time (Table 1). The comparative analysis of these four sets of MD simulations provides unprecedented opportunities to dissect the distinct contributions of cAMP binding and tetramerization to the structural dynamics of the HCN IR and its domains (*i.e.* TD and CBD). In addition, we show that our MD simulations capture the active-to-inactive CBD conformational transition, which occurs in the sub-μs timescale for the apo-monomeric HCN IR, but which had remained elusive for other eukaryotic CBDs.[27] Our results on HCN reveal that CBDs sample a manifold of active/inactive pathways that are significantly more divergent than previously anticipated. The occurrence of the active-to-inactive conversion in our MD trajectories is also critical to probe the TD dynamics in the inactive form of the auto-inhibitory apo-monomer HCN IR. The MD simulations were validated and complemented by experimental NMR and CD data, and reveal that dynamics play a pivotal role in the allosteric coupling between cAMP-binding to the CBD and self-association of the TD.

3.4. Experimental Procedures

3.4.1. Overview

Molecular dynamics (MD) simulations were performed for a protein construct spanning residues 521-717 of the intracellular region of human HCN4 (referred to herein as "HCN4 IR"), which contains the full TD and CBD regions of the HCN4 intracellular region (Figure 1b). Although the intact HCN4 IR construct also contains an additional segment C-terminal to the CBD (a segment absent from the X-ray structure), previous functional data suggests that deletion of this C-terminal segment has negligible effect on the function of the intact HCN protein.[28] Initial structures for the simulations of all four states in the thermodynamic cycle of Figure 1c were constructed based on the X-ray crystal structure of the cAMP-bound intracellular region (PDB ID "3OTF").[5] Hence, the simulations for the apo-monomer HCN IR state were initiated from a non-equilibrium state, *i.e*. the active HCN IR protomer conformation. Details about the preparation of the initial structures as well as the MD simulation protocols, analyses and validation by experimental data are described below.

3.4.2. Molecular Dynamics Simulation Protocol

Initial Structure Preparation. The HCN4 IR construct spanning residues 521-717 of the HCN4 intracellular region was used for all MD simulations. The initial structure for the holo-monomer state (Figure 1c) was obtained from the X-ray crystal structure of the cAMP-bound intracellular region (PDB ID "3OTF")[5] by deleting all water molecules from the structure, and using SwissPDB Viewer[29] to reconstruct partially missing side chains on the protein surface. The initial structure for the apo-monomer state (Figure 1c) was obtained

from the initial structure for the holo-monomer state simply by deleting the bound cAMP. The initial structures for the apo-tetramer and holo-tetramer states (Figure 1c) were obtained from the initial structures for the apo-monomer and holo-monomer states, respectively, by using the SwissPDB Viewer[29] to generate four copies of the monomer structure by applying "BIOMT" rotation/translation structure transformations specified in the header lines of the "3OTF" PBD text file.

MD Simulation Protocol. All MD simulations were performed using the NAMD 2.8 software[30] on the Shared Hierarchical Academic Research Computing Network (SHARCNET).[31] The CHARMM27 force field with CMAP correction[32,33,34,35] was implemented for all simulations, and the simulations were set up to mimic the following experimental conditions: a pH of 7.0; explicit water (with periodic boundary conditions) with a 50 mM concentration of NaCl; a constant temperature of 34°C (307 K); and a constant external pressure of 1 atm. Protein structure coordinate and parameter files with hydrogen atoms for the HCN4 IR structures were constructed using the "Psfgen" module of VMD 1.8.6.[36] Parameters for cAMP were constructed from the parameters for the adenine ribonucleotide ("ADE" in the CHARMM force field) by applying the force field's intrinsic "CY35" patch function.[32] In order to mimic a pH of 7.0, hydrogen atoms were added such that all His side chains were in their unionized τ-state, and the N-/C-termini and all Asp, Glu, Arg and Lys side chains were in their ionized states. The structures were then immersed in a cubic box of TIP3P water molecules (with box dimensions of 90 Å for monomers, or 120 Å for tetramers) using the Solvate module of VMD 1.8.6,[36] such that there was a minimum distance of 12 Å between the protein and the edge of the solvent box. Salt ions (Na^+ and Cl^-)

were added to the solvent box using the "Autoionize" module of VMD 1.8.6,[36] such that the system was neutralized and the number of ions corresponded to an NaCl concentration of 50 mM.

Initial energy minimizations were performed using the conjugate gradient algorithm of NAMD. Minimization was performed for 5000 steps with harmonic position restraints on the protein backbone (force constant of 300.0 kcal/mol·Å^2), followed by an additional 2000 steps without restraints. During minimization, a cutoff of 15 Å was utilized for all non-bonded energy calculations. Electrostatic interactions beyond the cutoff distance were computed using the Particle Mesh Ewald (PME) algorithm,[37] with a tolerance of 10^{-6} and a maximum grid spacing of 1.0 Å. Molecular dynamics simulations were then performed under cubic periodic boundary conditions, beginning from the energy-minimized initial structures. A time-step of 1.0 fs was implemented throughout the simulations. All water molecules were constrained to their equilibrium geometries using the SETTLE algorithm[38] and all covalent bonds to hydrogen were constrained using the SHAKE algorithm.[39] A cutoff of 12 Å with PME implementation was utilized for non-bonded energy calculations during the simulations. Short-range non-bonded and long-range electrostatic interactions were evaluated every 2.0 fs and 4.0 fs, respectively, using the RESPA multiple timestep integrator.[40] All minimizations and simulations were executed on a 2.83 GHz octuple-core Xeon cluster, using 64 CPUs per run.

The structures were heated linearly from 0 K to 307 K over 200 ps at constant volume, using the velocity reassignment protocol of NAMD. The heated structures were then simulated at 307 K and constant volume (NVE ensemble) for another 1.0 ns, to allow a period of temperature equilibration prior to introducing pressure regulation. Next, the

structures were simulated at a constant temperature and pressure (NPT ensemble) for 1.0 ns, to allow a period of temperature and pressure equilibration prior to the NPT production run. A constant temperature of 307 K was maintained using the Langevin dynamics algorithm,[41] with a Langevin damping coefficient of 1.0 ps^{-1}. A constant pressure of 1 atm (1.01325 bar) was maintained using the Nosé-Hoover Langevin piston method,[30] with a barostat oscillation period of 200.0 fs and a barostat damping time scale of 100.0 fs. Throughout the heating and equilibration runs, harmonic position restraints were imposed on the protein backbone (force constant of 5.0 kcal/mol·$Å^2$) to permit equilibration of the protein side chains and solvent without altering the protein backbone. Finally, production-run simulations were performed at a constant temperature and pressure (NPT ensemble) without restraints. These runs were executed for 120 ns in order to obtain a 100 ns trajectory for analysis, while allowing for a final unrestrained equilibration period of 20 ns before the 100 ns trajectory, and were performed at constant temperature and pressure using the NPT protocol described above. The simulations of all four states were performed in triplicate, and during the production runs, structures were saved every 10000 timesteps (*i.e.* every 10.0 ps) for subsequent analysis. A summary of the MD simulations performed for the HCN4 IR construct is given in Table 1.

3.4.3. Analysis of HCN4 IR Structural Dynamics

Backbone root-mean-square deviations. As an assessment of HCN4 IR structural propensities within the simulations, root-mean-square deviations (RMSDs) from the initial 3OTF monomer structure were computed for the constituent HCN4 IR protomers within each simulation, as well as for selected highly dynamic structural regions within each protomer

(Figure 1b): the CBD; the tetramerization domain (*i.e.* αA'-αD' region); the N3A (*i.e.* αE'-αA region); the PBC; and the αB-αC region. The RMSDs for the tetramerization domain, N3A, PBC, and αB-αC regions were computed *via* a protomer backbone overlay onto the 3OTF protomer reference structure at the β-core, in order to permit examination of the movement of these regions relative to the central scaffold of the CBD (*i.e.* the β-core). The RMSDs for the full-length protomer and CBD were computed *via* a protomer backbone overlay onto the 3OTF protomer reference structure at the region under examination, in order to permit examination of the structural variation within that region. In addition, RMSD calculations were performed for the N3A and αB-αC regions using the average atomic coordinates for the previously-solved apo HCN4 CBD structure ensemble (solved by NMR; *i.e.* the average structure derived from PDB entry "2MNG")[17] as a reference structure, in order to assess the extent of tendency toward the inactive-state CBD structure that was exhibited by these regions during the simulations.

RMSD-based structural distributions to map the pathways of active/inactive transition in the CBD of HCN4. To further assess CBD structural propensities within the simulations, RMSD-based structure similarity measures were computed for the constituent HCN4 IR protomers within each simulation. First, the RMSDs of the N3A and αB-αC regions from both the 3OTF and 2MNG protomer reference structures were computed as described above. Then, the RMSD-based similarity measure for the N3A in any given structure ("SM_{N3A}") was computed from the RMSD values for each constituent protomer as follows:

$$SM_{N3A} = (RI - RA) / R_{ref} \quad (1)$$

with "RA" denoting the N3A RMSD between the given structure and the *active*-state structure (*i.e.* 3OTF); "RI" denoting the N3A RMSD between the given structure and the *inactive*-state structure (*i.e.* average structure from 2MNG); and "R_{ref}" denoting the N3A RMSD between the *active*-state and *inactive*-state structures. For comparison, the SM_{N3A} values for the individual structures from the 2MNG ensemble were also computed. Similar definitions were used to compute the corresponding similarity measures for the αB-αC region ("$SM_{\alpha B\text{-}\alpha C}$"). Two-dimensional plots of the computed similarity measures were then constructed to examine how concertedly the similarities for the N3A and αB-αC regions vary.

Backbone N-H order parameters: local dynamics and comparison with NMR data. To quantify the amplitudes of local structural fluctuations, and permit a validation of the simulations against NMR data available for the HCN4 CBD, an analysis of backbone N-H order parameters (S^2) was performed for the constituent HCN4 IR protomers within each simulation. First, the generated structures of each constituent protomer were optimally superimposed onto the initial protomer structure based on α-carbon coordinates, using the "orient" module of CHARMM 31.1.[42] Then, calculation of the S^2 order parameters across the simulation trajectory was performed using the "NMR" module of CHARMM 31.1.[42] During the calculation, the NMR magnetic field strength parameter ("hfield") was set to a value of 16.44 Tesla, in accordance with the magnetic field strength of the NMR spectrometer utilized in the NMR experiments. In addition, the characteristic correlation time of overall protein tumbling ("rtumbl") was set to a value of 12310.0 ps. This value of "rtumbl" was calculated based on the 3OTF protomer X-ray structure using the

HYDRONMR 7c software.[43] In the HYDRONMR calculation, the temperature was set to 307 K, and the magnetic field strength to 16.44 Tesla.

To assess whether key CBD dynamic features were successfully captured by the simulations, the CBD order parameters computed from the holo-monomer simulations were compared to a corresponding set of experimentally-derived order parameters (S^2) obtained for the cAMP-bound monomeric CBD, since in both cases, the CBD was expected to exist predominantly in its active-state conformation. The experimentally-derived order parameters were computed from experimentally measured $^{13}C\alpha$, $^{13}C\beta$, ^{13}CO, ^{1}H, and ^{15}N chemical shifts for the cAMP-bound HCN4 IR-ΔA'B' construct (which is monomeric due to deletion of the αA'-αB' segment of the TD) using the RCI algorithm.[44] The RCI-based order parameters do not rely on specific models of overall rotation and therefore are ideally suited for constructs with extensive flexible regions that may affect the rate and anisotropy of the overall tumbling, such as the HCN4 IR and HCN4 IR-ΔA'B' constructs.

Backbone root-mean-square fluctuations (RMSFs). To further assess structural fluctuations with residue-resolution, an analysis of root-mean-square fluctuations (RMSFs) was performed for the constituent HCN4 IR protomers within each simulation. In this analysis, the generated structures of each constituent protomer were first optimally superimposed onto the initial protomer structure based on α-carbon coordinates. The RMSF calculation was subsequently performed by first computing the average Cartesian coordinates of the α-carbon atoms over all structures in each overlaid structure set. Then, the RMSFs around the average coordinates of each structure set were calculated as follows:

$$RMSF(i) = \sqrt{\left\langle (r_i - \langle r_i \rangle)^2 \right\rangle} \qquad (2)$$

where "*RMSF(i)*" denotes the RMSF describing the fluctuations of residue "*i*", and "r_i" denotes the Cartesian coordinates of the α-carbon atom of residue "*i*" for a given structure; "$<r_i>$" denotes the average Cartesian coordinates of the α-carbon atom of residue "*i*", computed over the whole structure set; and the angle brackets denote averages computed over the whole structure set. This process was performed with all structure overlays done at the β-core, in order to permit examination of residue-specific movements relative to the central scaffold of the CBD (*i.e.* the β-core).

Circular Dichroism. Circular Dichroism (CD) spectra were recorded using an AVIV spectrometer model 215. All measurements were taken in a 1-mm quartz cuvette using a slit width of 1.053 cm. The CD spectra were measured at 25 ^{o}C for solutions of the HCN4 αA'-αB' (521-562) and αC'-αD' (566-588) peptides (GenScript USA Inc.) at a 0.1 mg/mL concentration in a 10 mM potassium phosphate buffer (pH 6.5) with and without 40% TFE. The estimated percentage of helical structure was calculated from the $[\theta]_{222}$. Each point in the CD spectra was corrected for background reading using a blank buffer solution.

3.5. Results

3.5.1. Overall IR Dynamics

As an initial assessment of the IR structural dynamics, RMSDs of the full IR and its constituent domains (*i.e.* the TD and CBD; Figure 1a-b) were computed from the cAMP-bound structure, which is assumed to represent the active-state IR.[3,5,45] All RMSD calculations were performed for the constituent protomers of all four simulated states (Figure 2). To obtain insight on how these states vary across the thermodynamic cycle of Figure 1c,

we mapped the RMSD distributions within each set of simulations (Figure 3). The RMSD values for the full IR exhibit a trend of progressive decrease in the order apo-monomer > holo-monomer > apo-tetramer > holo-tetramer (Figure 3a), suggesting that stabilization of the overall active-state IR structure involves contributions from both cAMP binding and IR tetramerization. However, the degrees of contribution by cAMP binding and tetramerization were found to vary significantly between the TD and CBD regions (Figure 3b-c). For the TD, the RMSD values of the monomers generally exceed those of the tetramers, irrespective of whether they are apo or holo (Figure 3b), suggesting that the active-state TD configuration is stabilized primarily by tetramerization. For the CBD, a different pattern is observed (Figure 3c). The CBD RMSDs from the active-state structure are only marginally reduced in each tetramer state relative to its respective monomer state (Figure 3c), suggesting that unlike the TD, the main determinant of active-state CBD structure stabilization is not tetramerization.

3.5.2. CBD Dynamics

To further examine the CBD structural dynamics, RMSDs of the major CBD α-helical structural elements (*i.e.* the PBC, αB-αC and N3A; Figure 1b) relative to the β-core were computed (Figure 3d-f). The RMSDs for both the PBC and αB-αC elements demonstrated a reduction in the extent of deviation from their active-state configurations only in the two holo states (Figure 3d-e), suggesting that the active-state configurations of these elements are stabilized mainly by cAMP binding, with little contribution from tetramerization. In marked contrast with the PBC and αB-αC regions, the N3A demonstrated an enhancement in its extent of deviation from the active structure only in the apo-monomer

state, while the other three states all exhibit similarly quenched dynamics (Figure 3f). This result suggests that the stabilization of the N3A active-state configuration is strongly influenced by both cAMP binding and tetramerization.

To assess the extent to which the CBD structural dynamics span active-to-inactive conformational transitions, active *vs.* inactive similarity measures (SM) were defined according to Equation 1 (see Experimental Procedures section) based on the RMSDs from both the experimental active and inactive conformations for two key CBD elements involved in the transition, *i.e.* the αB-αC helices and the N3A. When the SM approaches a value of 1 (-1), the CBD element for which the SM was computed adopts a conformation similar to the active (inactive) state. In the active CBD conformation, the αB-αC helices are proximal to the β-barrel, while the N3A moves away from it (*i.e.* αB-αC in/N3A out topology, or "in-out" for short; Figure 1d). In the inactive CBD conformation, however, the reverse CBD topology prevails, with the αB-αC helices moving away from the β-barrel, and the N3A residing near it (*i.e.* αB-αC out/N3A in topology, denoted as "out-in"; Figure 1d).

The active *vs.* inactive structure similarity measures (SM) were computed for the αB-αC helices and N3A in each frame of the simulated trajectories, and compiled as two-dimensional plots (Figure 4). Each two-dimensional plot is composed of four quadrants corresponding to the αB-αC in/out and N3A in/out independent combinations. The αB-αC in/N3A out topology of the active CBD conformation falls in the top-right quadrant (*i.e.* SM(αB-αC) > 0 and SM(N3A) > 0), while the αB-αC out/N3A in topology of the inactive CBD conformation falls in the bottom-left quadrant (*i.e.* SM(αB-αC) < 0 and SM(N3A) < 0; Figure 4). The other two quadrants correspond to mixed "out/out" and "in/in" CBD

topologies (*i.e.* SM(αB-αC) < 0 and SM(N3A) > 0, or SM(αB-αC) > 0 and SM(N3A) < 0; Figure 4).

The two-dimensional plots of active *vs.* inactive structure similarity measures were constructed for each state of the thermodynamic cycle of cAMP-dependent tetramerization (Figure 1c) as shown in Figure 4a-d, using the same color code as in Figure 1c. It should be noted that both panels (a) and (b) of Figure 4 include the two-dimensional SM plots for the holo states, but in panel (a) the plot for the holo-tetramer is in the front layer, while in panel (b) it is the plot for the holo-monomer that is in the front layer. Similarly, both panels (c) and (d) of Figure 4 display the two-dimensional SM plots for the apo states, but with different superimposition orders.

Figure 4a-b shows that both holo states (*i.e.* the holo-monomer and holo-tetramer) are confined primarily to the active CBD quadrant, consistent with the experimental evidence that cAMP stabilizes the active CBD conformation irrespective of the IR oligomerization state.[17] However, marked and unanticipated differences exist between the similarity distributions of the αB-αC and N3A elements. Specifically, while the αB-αC similarity distributions are narrow and entirely confined to the active CBD quadrant, the N3A distributions are significantly broader and partially exit from the active CBD quadrant (Figure 4a-b). These distribution patterns change dramatically in the apo states (*i.e.* the apo-monomer and apo-tetramer), as shown in Figure 4c-d.

Both apo states exhibit an αB-αC distribution significantly broader than that of the holo states (Figure 4). Unlike the holo states, the apo states span both in and out αB-αC orientations relative to the β-core, with the latter orientation sampled slightly more in the apo-monomer than the apo-tetramer simulations (Figure 4c-d). In addition, the apo-

monomer and apo-tetramer differ in terms of the N3A distributions. In the apo-tetramer, the range of N3A similarity measures (Figure 4c-d) remains confined to values comparable to those observed for the two holo states (Figure 4a-b). In the apo-monomer, however, the N3A distribution shifts to lower values, suggesting a further transition towards the inactive CBD conformation (Figure 4c-d). Indeed, the simulated ensemble generated for the apo-monomer includes a sub-set of conformations exhibiting the inactive out-in topology, with multiple conformations approaching or falling within the range of similarity measures calculated for the experimental inactive-state CBD NMR structure ensemble[17] (dashed red lines in Figure 4c-d). This observation is significant because the MD trajectories started from the opposite topology (active in-out), and suggests that the active-to-inactive transition was captured by the simulations, in agreement with the experimental observation that the inactive CBD state is the most stable conformation in the apo-monomeric IR.[17]

To confirm that inactive-state-like CBD configurations were being explored within the apo-monomer HCN4 IR simulations, frames exhibiting minimal N3A RMSDs (*i.e.* < 3 Å) from the average apo-monomer CBD structure previously solved by NMR (2MNG) were selected for further assessment (Figure 5a). The N3A RMSDs were used as a selection criterion because the orientation of the N3A relative to the β-core in the CBD is a key determinant of the inactive/active conformational shift, as noted previously,[17] while the inherent apo-state flexibility in the αB-αC region could bias the selection. Once the inactive-state-like structures from each apo-monomer simulation were selected, the structures exhibiting the smallest, median and largest N3A RMSDs out of all of the selected inactive-state-like structures were overlaid onto the average 2MNG structure to assess their extent of similarity to the average 2MNG CBD conformation (Figure 5b). In addition, inter-α-carbon

contact distances were computed for selected pairs of residues that bridge the interface between the N3A and β-core, to further confirm whether the selected structures exhibited N3A/β-core relative orientations close to or approaching that of the average 2MNG structure (Figure 5c-d). To properly assess the N3A movement, the contact distance analysis was performed for representative residue pairs that are farther apart in the 2MNG structure than in the 3OTF structure, as well as representative residue pairs that are closer together in the 2MNG structure than in the 3OTF structure (Figure 5c-d). Overall, the results of Figure 5 confirm the occurrence of active-to-inactive CBD conversion in our apo-monomer simulations as also suggested by Figure 4c-d, and hence, the two-dimensional plots in Figure 4c-d provide invaluable clues about the pathways for the active-inactive CBD transition of the apo-monomeric HCN IR.

Figure 4c-d shows that the active *vs.* inactive similarity measures (SM) for the N3A and αB-αC elements in the apo-monomer trajectories are significantly less correlated than expected for a fully concerted in-out to out-in transition (Figure 4c-d). The N3A and αB-αC SM distributions sample a wide range of values, indicating the existence of a manifold of pathways for the transition between active-like and inactive-like structures within the apo-monomer ensemble, and suggesting that the apo-monomer CBD also transiently samples non-canonical mixed CBD topologies (*e.g.* in-in, and to a smaller extent, out-out topologies; Figure 4c-d).

3.5.3. TD Dynamics

To further assess the internal TD structural dynamics, backbone N-H order parameters (S^2; Figure 6a) and root-mean-square fluctuations (RMSFs; Figure 6b) were

computed for the constituent protomers of all four simulated states. The S^2 and RMSF profiles in Figure 6 are validated by their agreement with the experimentally observed quenching of dynamics observed for the αB-αC regions upon cAMP binding in both monomeric and tetrameric states.[17,45] As an additional means of validating the simulations, backbone N-H order parameters (S^2) computed for the CBD from the holo-monomer simulations were compared with order parameters determined experimentally for the cAMP-bound monomeric state (Figure 7a). In this state, the CBD is expected to exist predominantly in its active-state structure, which is represented by the initial simulation structure for this state (Figure 1). Comparison of the two S^2 data sets revealed that the trends in CBD local dynamics computed from the simulations were comparable to the trends observed from experiment (Figure 7a). This observation suggests that key CBD dynamic features were successfully captured by the simulations, thus lending further credibility to the MD trajectories.

The S^2 and RMSF plots (Figure 6) are not only in agreement with the available experimental data, but they also reveal new dynamic attributes that are currently not fully accessible to experiments. For example, the S^2 values in the TD region (*i.e.* αA'-αD') are consistently lower in the monomeric states than in the tetrameric states (Figure 6a). A distinct monomer *vs.* tetramer difference is also observed for the RMSF values in the TD region (Figure 6b). These observations point to an enhancement of TD dynamics in the monomeric states irrespective of whether they are apo or holo (Figure 6), in agreement with the trends observed in Figure 3b for the whole-TD RMSDs. The monomer-*vs.*-tetramer increase in TD flexibility suggests that the former are significantly less structured than the latter. In order to gauge the degree of unstructuring of the monomeric TD region, we

acquired CD spectra of two peptides spanning the αA'-αB' and αC'-αD' helices, which constitute the TD region. Two separate αA'-αB' and αC'-αD' peptides (as opposed to a single αA'-αD' peptide) were utilized to circumvent tetramer formation and better probe the monomeric state. The CD spectra of both peptides show only marginal helical content in the absence of helix inducers such as TFE (Figure 7b-c and Table 2). While the αA'-αB' and αC'-αD' peptides may not retain the same degree of structuring as the full αA'-αD' TD region, these results suggest that in the absence of tertiary interactions between the αA'-αB' and αC'-αD' segments of the TD, the helical secondary structure of the TD is to a large extent lost (Figure 7b-c and Table 2).

3.6. Discussion

We have examined how HCN4 IR dynamics vary along the thermodynamic cycle arising from the coupling between the cAMP-binding and tetramerization equilibria (Figure 1c). The key results are summarized in Table 3 and Figure 8. The reliability of these results is supported by the fact that several changes in the simulated dynamics are in agreement with previous experimental observations. For instance, our MD simulations correctly capture the quenching of dynamics of the C-terminal CBD αB-αC helices upon cAMP binding to either the monomeric or tetrameric IR (Table 3 and Figure 8), as previously determined by NMR, DEER and fluorescence spectroscopy.[17,45,46,47] Furthermore, our MD simulations correctly predict that the N3A topology in the apo-monomer IR undergoes a transition from the "out" to the "in" orientation, in agreement with the recently solved NMR structure of the apo-monomeric CBD.[17] Our MD simulations are also corroborated by the fact that they reproduce the order parameter (S^2) trends derived from the NMR data for the holo-

monomeric CBD. Overall, our MD-based results appear in agreement with the available experimental information on the HCN4 IR dynamics, lending credibility to our simulations.

The MD simulations presented here are not only supported by previous measurements, but they also reveal several functionally relevant dynamic attributes of the HCN4 IR that have so far remained elusive to experimental approaches. First, the MD trajectories of the inactive apo-monomeric IR clearly indicate that the TD is highly dynamic (Figure 6). The TD flexibility in the apo-monomer plays a central role in stabilizing the monomeric form of the apo IR, because it effectively removes the steric hindrance that would otherwise arise between the αA'-αB' helices and the β-subdomain of the inactive-state CBD. Hence, the TD dynamics is critical to explain how cAMP-binding and IR tetramerization are coupled.

Second, the MD simulations show that, overall, the active holo-tetrameric IR is less dynamic than the inactive apo-monomeric state (Figures 3, 6 and 8; Table 3). Both cAMP-binding and tetramerization contribute to the quenching of IR dynamics that occurs upon activation. However, the extent of influence by cAMP-binding *versus* tetramerization varies between structural regions within the IR, as revealed by the comparative analysis of the dynamic profiles for the four states of the thermodynamic cycle in Figure 1c (Table 3 and Figure 8). The TD structural dynamics are influenced primarily by tetramerization, while the CBD structural dynamics are influenced mainly by cAMP binding, with a lesser contribution from tetramerization (Table 3 and Figure 8). In particular, closer examination of the CBD revealed that the PBC and αB-αC helices are influenced primarily by cAMP binding, but the N3A motif is influenced by both cAMP binding and tetramerization. The latter observation suggests that the N3A may play a critical allosteric role in the synergistic coupling between

cAMP-binding and tetramerization, whereby perturbation of the N3A upon cAMP binding may facilitate tetramerization, and *vice versa*.

Third, our MD simulations reveal that the transition between the inactive (*i.e.* N3A "in"/αB-αC "out") and active (*i.e.* N3A "out"/αB-αC "in") CBD topologies occurs through a manifold of divergent pathways (Figure 4c-d and Table 3). The active-to-inactive transitions captured by our HCN simulations could not be detected in previous simulations of other CBDs[27,48] and indicate that the conversion between the in/out and out/in CBD topologies is significantly less concerted than previously anticipated.[49] Our results clearly show that transient exploration of non-canonical mixed topologies is possible for the apo CBD (Figure 4c-d). Notably, while the non-canonical CBD topologies appear accessible to both apo-monomers and apo-tetramers, in the former the "in" N3A topology is more populated than in the latter – *i.e.* tetramerization of the apo IR causes a shift to a preferred "out" N3A topology. In fact, the apo-tetramer distribution of N3A orientations is similar to that observed for the holo IR states, despite the absence of cAMP (Figures 4 and 8; Table 3). The preference for the "out" N3A topology in the apo-tetramer is explained by the reduction in TD dynamics that occurs upon TD self-association (Figures 4 and 8; Table 3), which in turn destabilizes the "in" N3A topology due to intra- and inter-molecular steric clashes between the TD and the CBD β-barrel.

The preference of the apo-tetramers for the "out" N3A topology revealed by our simulations rationalizes why the relative N3A/β-barrel orientation remains virtually unaffected irrespective of whether the HCN IR tetramer is crystallized in the presence or absence of cAMP.[45] This prior observation based on X-ray crystallography was puzzling given the markedly different functional gating properties of the apo and holo HCN IR, but is

well explained in light of our MD results. Furthermore, our simulations reveal that, although minor populations of the two holo ensembles still sample the N3A "in" topology preferred by the apo-monomer, the ranges of possible conformations for the two holo states exhibit a greater degree of overlap than is observed for the two apo states (Figure 4 and Table 3). These findings suggest that cAMP binding selectively stabilizes the CBD into tetramer-compatible configurations, thereby priming the IR for tetramerization and providing a viable mechanism for the cAMP-dependent gating of HCN ion channels.

In conclusion, the current study provides an unprecedented view of TD and CBD structural dynamics within the HCN IR, and their role in promoting HCN channel gating through tetramerization. First, the TD is dynamic in the monomer states, but becomes structured in the tetramer states. The TD flexibility in the monomeric states removes the steric clashes that the structure of the apo CBD would otherwise impose, explaining the auto-inhibitory role of the apo IR. Second, the apo CBD undergoes an active/inactive transition through a manifold of pathways. These pathways are more divergent than previously anticipated, and involve a combinatorial array of in/out topologies for the orientation of the N- and C-terminal CBD α-helices relative to the β-subdomain. Third, the conformational ensemble accessible to the apo CBD is confined in the tetrameric state due to steric constraints imposed by the tetramer, which prevents a complete active-to-inactive transition. The binding of cAMP pre-confines the CBD conformational ensemble to a tetramer-compatible state, thereby priming the IR for tetramerization, and providing an explanation for how cAMP controls HCN channel gating through IR tetramerization. We anticipate that the concepts proposed here may be of general applicability for gating rings controlled by allosteric effectors as well as for multi-domain signaling systems.[50,51,52,53,54,55,56,57,58,59,60,61]

3.7. Acknowledgements

We thank Professors Eric Accili (U. British Columbia), Anna Moroni (University of Milan) and Susan S. Taylor (UCSD), as well as S. Boulton and K. Moleschi for helpful discussions. This project was funded by a grant to G.M. from the Canadian Institutes of Health Research (CIHR).

3.8. Tables and Figures

Table 1. *Summary of the MD Simulations Performed for the HCN4 IR*

Simulated IR State [a]	Number of HCN4 IR Protomers	Contains Bound cAMP (yes/no)?	Simulation Length per Replicate (ns)	Number of Simulation Replicates Executed	Number of CPUs Used
Apo-Monomer	1	No	100.0	3	64
Apo-Tetramer	4	No	100.0	3	64
Holo-Monomer	1	Yes	100.0	3	64
Holo-Tetramer	4	Yes	100.0	3	64

[a] The initial structures for all states were obtained from the X-ray crystal structure of the cAMP-bound intracellular region of human HCN4 (PDB ID "3OTF").

Table 2. *Helical Content (%)*

	0 % TFE	40 % TFE
HCN4 αA'-αB' [a]	6%	46%
HCN4 αC'-αD' [a]	7%	66%

[a] The peptide concentration was 0.1 mg/mL. The estimated percentage of helical structure was calculated from the $[\theta]_{222}$ in Figure 7b,c.

Table 3. *Summary of Dynamic Changes Along the Thermodynamic Cycle for the Coupling between Tetramerization and cAMP-Binding to the HCN4 IR*

Simulated IR States		TD (αA'-αD')	CBD	
			N3A (αE'-αA)	αB-αC
Apo-Monomer	IR_1	High [a]	In/Out [b]	High,[a] In/Out [c]
Apo-Tetramer	IR_4	Low [a]	Out (In) [b]	High,[a] In/Out [c]
Holo-Monomer	IR_1:$cAMP_1$	High [a]	Out (In) [b]	Low, [a] In [c]
Holo-Tetramer	IR_4:$cAMP_4$	Low [a]	Out (In) [b]	Low, [a] In [c]

[a] "High"/"low" refers to enhanced/quenched internal dynamics as quantified by RMSD, S^2 and RMSF changes in Figures 3 and 6.

[b] "In" refers to the in-topology whereby the N3A is oriented toward the CBD β-barrel (Figure 1d). "Out" refers to the out-topology whereby the N3A is oriented away from the CBD β-barrel (Figure 1d). The topology in parenthesis is less populated than the one without parenthesis, but is still sampled by the N3A as shown in Figure 4.

[c] "In" refers to the in-topology whereby the αB-αC helices are oriented toward the CBD β-barrel (Figure 1d). "Out" refers to the out-topology whereby they are oriented away from the CBD β-barrel (Figure 1d). "In/Out" indicates that both topologies are sampled by the αB-αC helices, as shown in Figure 4.

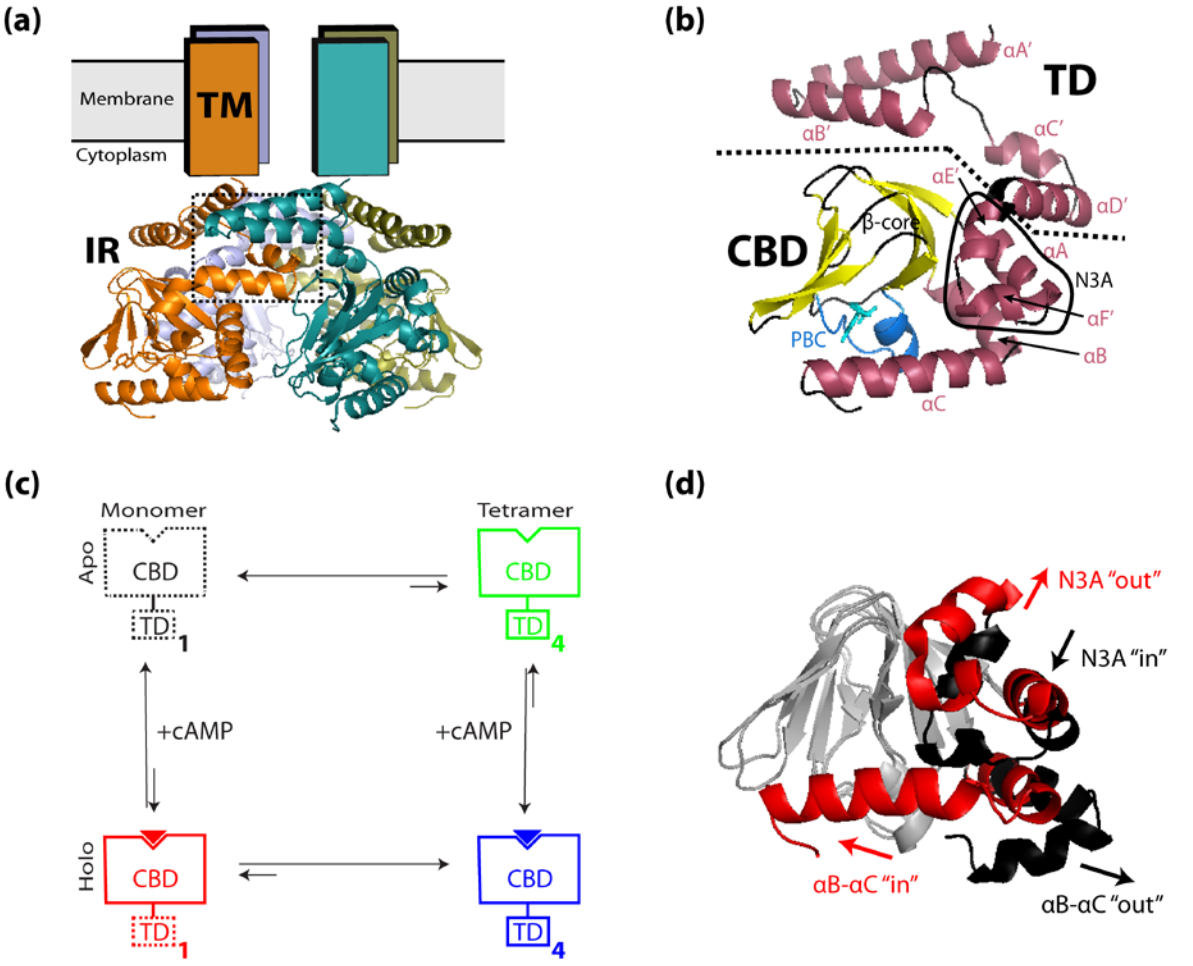

Figure 1: (a,b) Overview of the HCN architecture. **(a)** Active tetrameric structure of HCN4, based on the structure of the HCN4 IR bound to cAMP (RCSB Protein Data Bank code "3OTF"), viewed parallel to the plane of the cell membrane. The four HCN4 monomers are indicated in orange, olive green, blue-gray and teal, and the tetramerization interface between the C-linkers of neighbouring monomers is indicated (dotted rectangle). The four N-terminal transmembrane (TM) regions, whose atomic-resolution structure is currently unknown, are indicated as rectangles, while the four C-terminal intracellular regions (IR) are shown as ribbon structures. **(b)** Structural details of the HCN4 intracellular region for a single protomer illustrated as a ribbon structure, with α-helices indicated in maroon, β-strands in yellow, and bound cAMP as cyan sticks. Boundaries between major structural regions are delineated by dotted lines, and key structural elements are indicated. The αA'-αD' helices of the C-linker form the "tetramerization domain" (TD) of the IR. The αA helix N-terminal to the β-subdomain, together with the αE' and αF' C-linker helices, forms the N-terminal "N3A" motif of the cAMP-binding domain (CBD). The phosphate-binding cassette (PBC), where cAMP binds, is indicated in light blue. **(c)** Outline of the thermodynamic cycle for the

coupling between cAMP-binding and tetramerization of the HCN intracellular region. The tetramerization domain (TD) and cAMP-binding domain (CBD) regions are schematically indicated as rectangles, and bound cAMP is indicated as a solid triangle. Solid contour lines indicate domain states with structures similar to the 3OTF X-ray structure, while dashed contour lines indicate domain states with structures that are possibly different from the 3OTF X-ray structure. The subscripts "1" and "4" refer to the number of protomers within each state. **(d)** Overlay of the 3OTF (red ribbon) and average 2MNG (black ribbon) structures (overlaid at their β-cores), illustrating the conformational changes that occur within the CBD α-subdomain during cAMP-associated activation. Upon cAMP binding, the αB-αC region shifts from an "out" position to an "in" position, while the N3A shifts from an "in" position to an "out" position. For clarity, the β-cores of both structures are shown in gray, and the bound cAMP and αA'-αD' helices are omitted. All ribbon structures were generated using Pymol (Schrödinger, LLC).

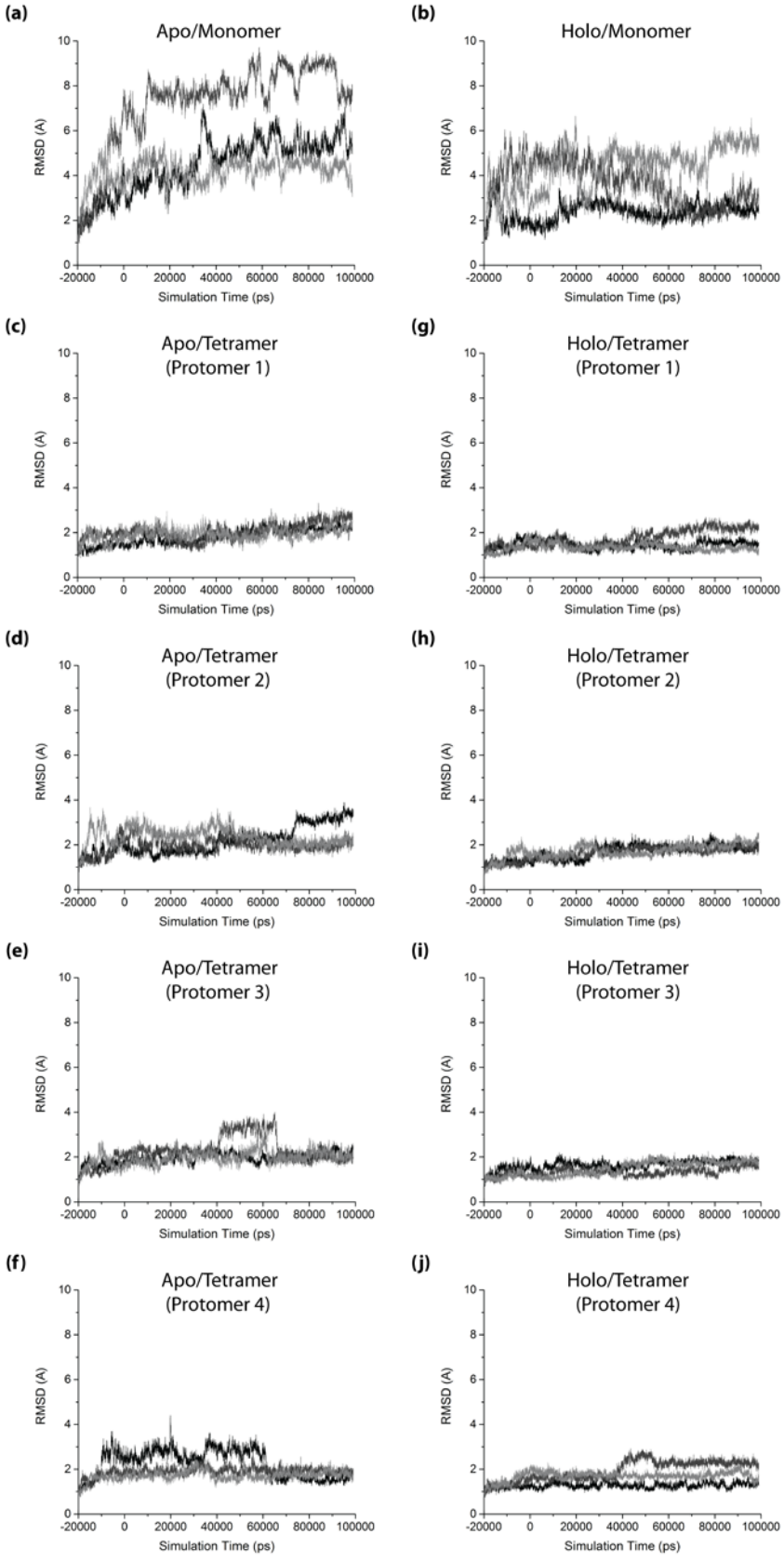
(a)
Apo/Monomer
(b)
Holo/Monomer
(c)
Apo/Tetramer
(Protomer 1)
(g)
Holo/Tetramer
(Protomer 1)
(d)
Apo/Tetramer
(Protomer 2)
(h)
Holo/Tetramer
(Protomer 2)
(e)
Apo/Tetramer
(Protomer 3)
(i)
Holo/Tetramer
(Protomer 3)
(f)
Apo/Tetramer
(Protomer 4)
(j)
Holo/Tetramer
(Protomer 4)
RMSD (A)
Simulation Time (ps)
0
2
4
6
8
10
-20000
0
20000
40000
60000
80000
100000

Figure 2: Root-mean-square deviations (RMSDs) from the initial 3OTF structure computed for each HCN4 IR protomer over the course of the respective HCN4 IR simulations: **(a)** apo-monomer state; **(b)** holo-monomer state; **(c-f)** apo-tetramer state; and **(g-j)** holo-tetramer state. RMSD trajectories are shown for all individual protomers (individual panels) and replicates (black, dark gray and light gray plots in each panel) for the respective states. The initial equilibration period of each simulation is indicated as negative times, while the production run used in the other analyses is indicated as positive times.

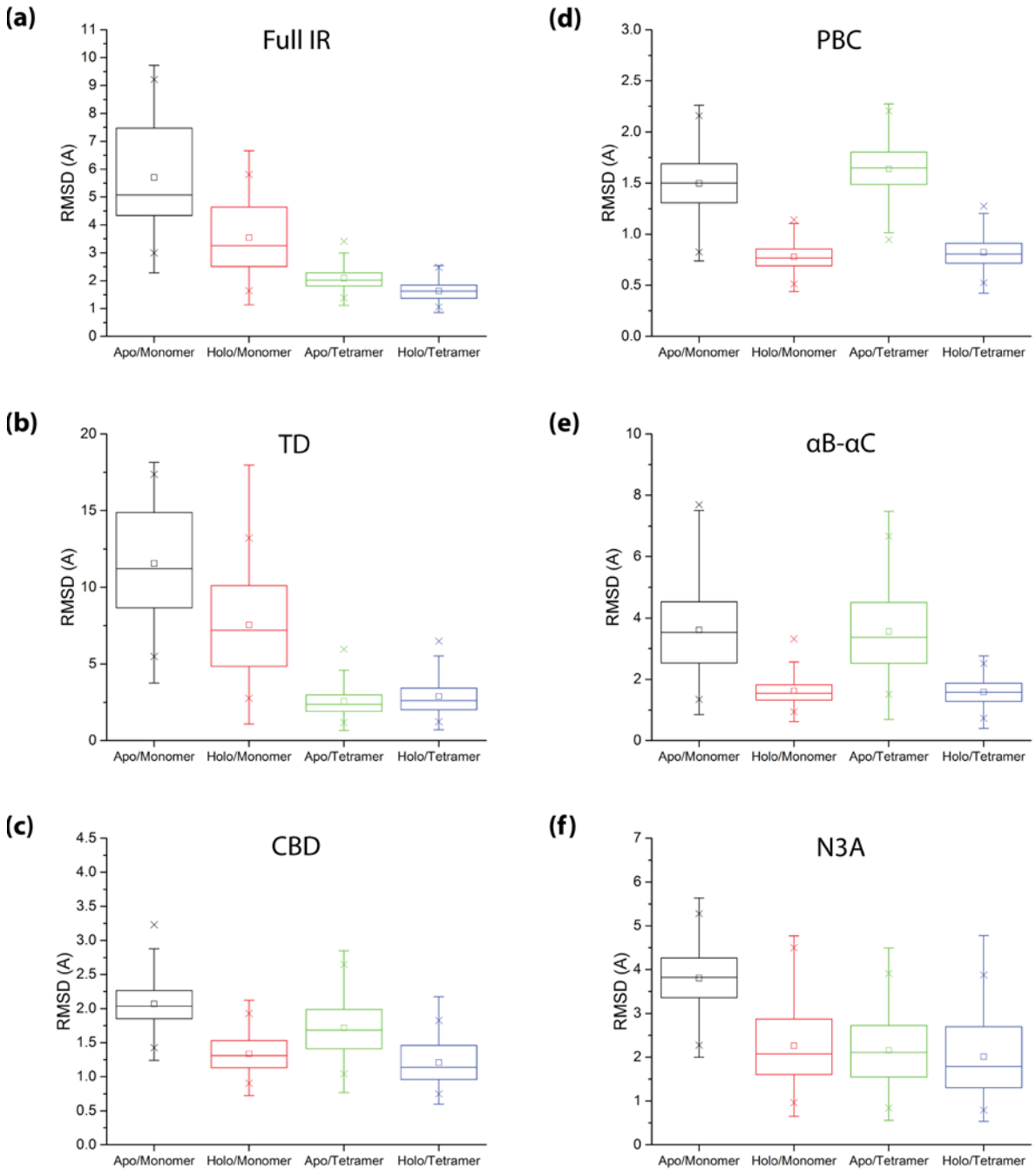

Figure 3: (a-c) Distribution of backbone root-mean-square deviations (RMSDs) of the major protomer structural regions from their configurations in the initial 3OTF structure, as observed in the context of the apo-monomer (black), holo-monomer (red), apo-tetramer (green) and holo-tetramer (blue) states of the HCN4 IR. **(a)** RMSDs for the entire HCN4 IR; **(b)** RMSDs for the tetramerization domain (TD), computed with structure overlay at the β-

core; **(c)** RMSDs for the CBD. **(d-f)** Distributions of RMSDs from the initial 3OTF structure computed for the major CBD α-subdomain structural regions: **(d)** the PBC; **(e)** the αB-αC region; and **(f)** the N3A (*i.e.* αE'-αA region); all computed with structure overlay at the β-core. The color code is as in Figure 1c. All boxplots were constructed using Origin 9.1 (OriginLab Corporation), based on the RMSD data from all protomers and replicates for the respective states. The statistics reported in each boxplot are as follows: the middle, bottom and top lines of the central box represent the median, 25th percentile and 75th percentile of the data set, respectively; the whiskers represent additional data falling within 1.5*IQR above the 75th percentile or below the 25th percentile (where IQR is the difference between the 75th and 25th percentiles); the "□" symbol represents the mean of the data set; and the two "×" symbols represent the 1st and 99th percentiles of the data set.

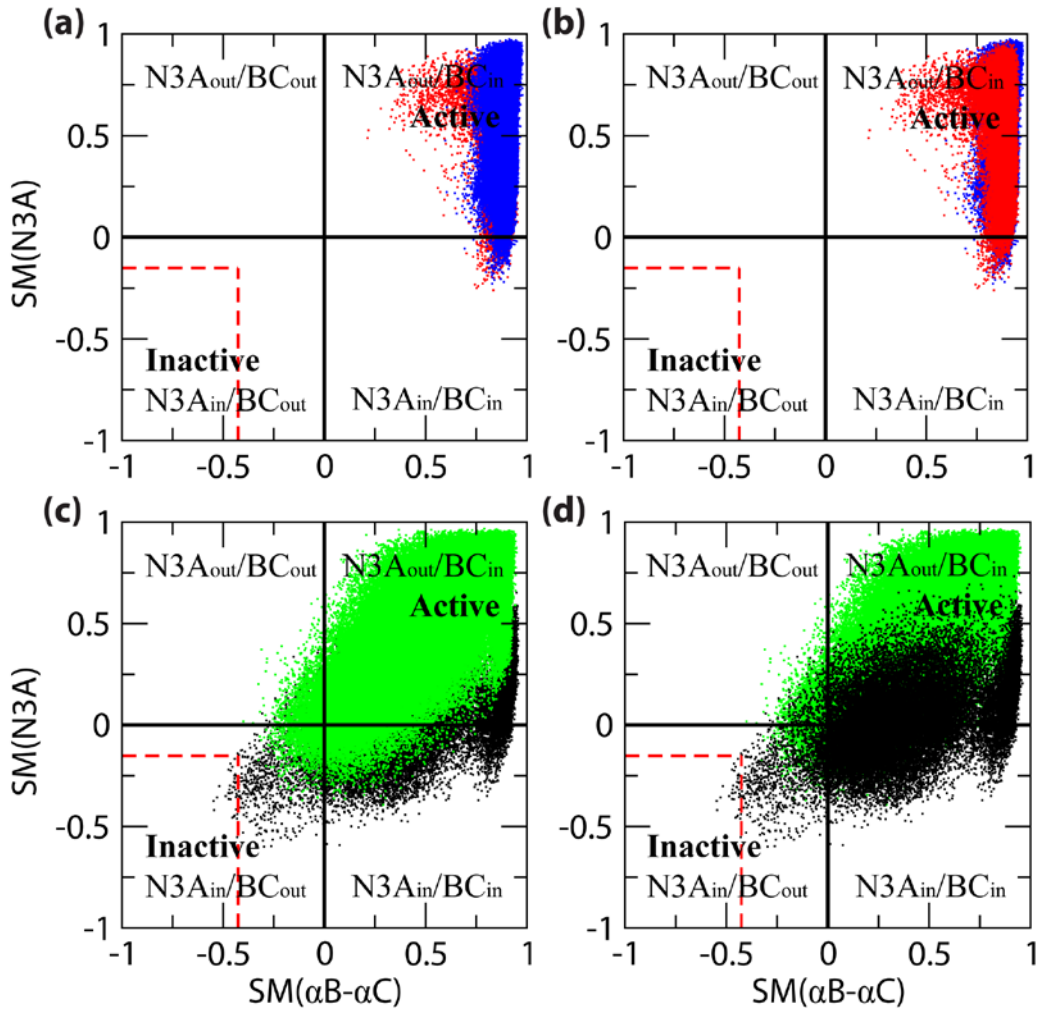

Figure 4: Two-dimensional plots of the computed active *vs.* inactive structure similarity measures (SM) for the N3A (*i.e.* αE'-αA) and αB-αC regions, as observed in the context of the four states of the HCN4 IR. The SM values were computed based on equation 1 in the Experimental Procedures section. **(a-b)** holo-monomer (red) and holo-tetramer (blue); **(c-d)** apo-monomer (black) and apo-tetramer (green). Note that each top panel (*i.e.* a and b) includes both holo simulation data, but with a different order of front *vs.* back layers. Similarly, each bottom panel (*i.e.* c and d) includes both apo simulation data, again with different orders of front *vs.* back layers. The color code is as in Figure 1c, and the N3A and αB-αC structural topologies (*i.e.* "in" *versus* "out", and "Inactive" *versus* "Active") represented by the quadrants of each graph are indicated. For comparison, the boundaries of the range of similarity measure values computed from the individual structures of the inactive-state NMR structure ensemble (*i.e.* 2MNG) are also indicated (red dashed lines). Details of the similarity calculations are provided in the Experimental Procedures section (Equation 1).

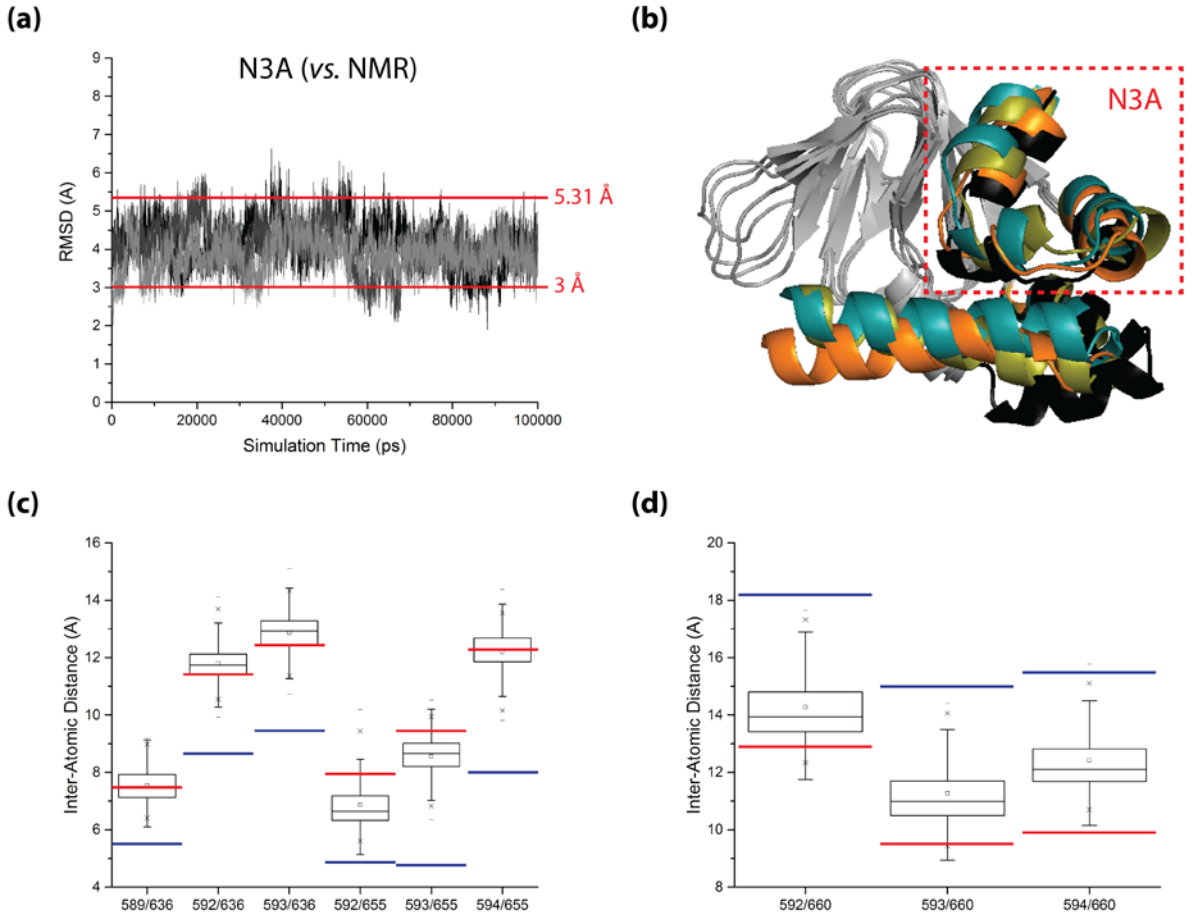

Figure 5: Sampling of inactive-state-like CBD topologies within the apo-monomer HCN4 IR simulations. **(a)** Backbone root-mean-square deviation (RMSD) trajectories of the N3A (*i.e.* αE'-αA region) from its configuration in the average apo-state NMR structure of the HCN4 CBD (*i.e.* the average structure derived from RCSB PDB entry "2MNG"), as observed in each of the three replicate apo-monomer HCN4 IR simulations (black, dark gray and light gray plots, respectively). All RMSDs were computed with structure overlay at the β-core. The threshold RMSD value used for selecting inactive-state-like structures for further assessment (*i.e.* 3 Å), and the RMSD between the 3OTF and average 2MNG structures (*i.e.* 5.31 Å), are indicated by horizontal lines. **(b)** Ribbon-structure illustration of the selected inactive-state-like structures exhibiting the smallest (orange ribbon), median (olive green ribbon) and largest (teal ribbon) N3A RMSDs from the average 2MNG structure, overlaid at their β-cores onto the average 2MNG structure (black ribbon). For clarity, the β-cores of all structures are shown in gray, and the αA'-αD' helices of all structures are omitted. **(c,d)** Boxplots of inter-α-carbon distances for selected pairs of residues that bridge the interface between the N3A and β-core, as observed for the selected ensemble of inactive-state-like structures. Residue pairs that are farther apart in the average 2MNG structure (panel c), and residue pairs that are closer together in the average 2MNG structure (panel d), were

considered. All boxplots were constructed using Origin 9.1 (OriginLab Corporation), and for reference, the corresponding distance values observed in the 3OTF structure (blue lines) and in the average 2MNG structure (red lines) are indicated for all plots.

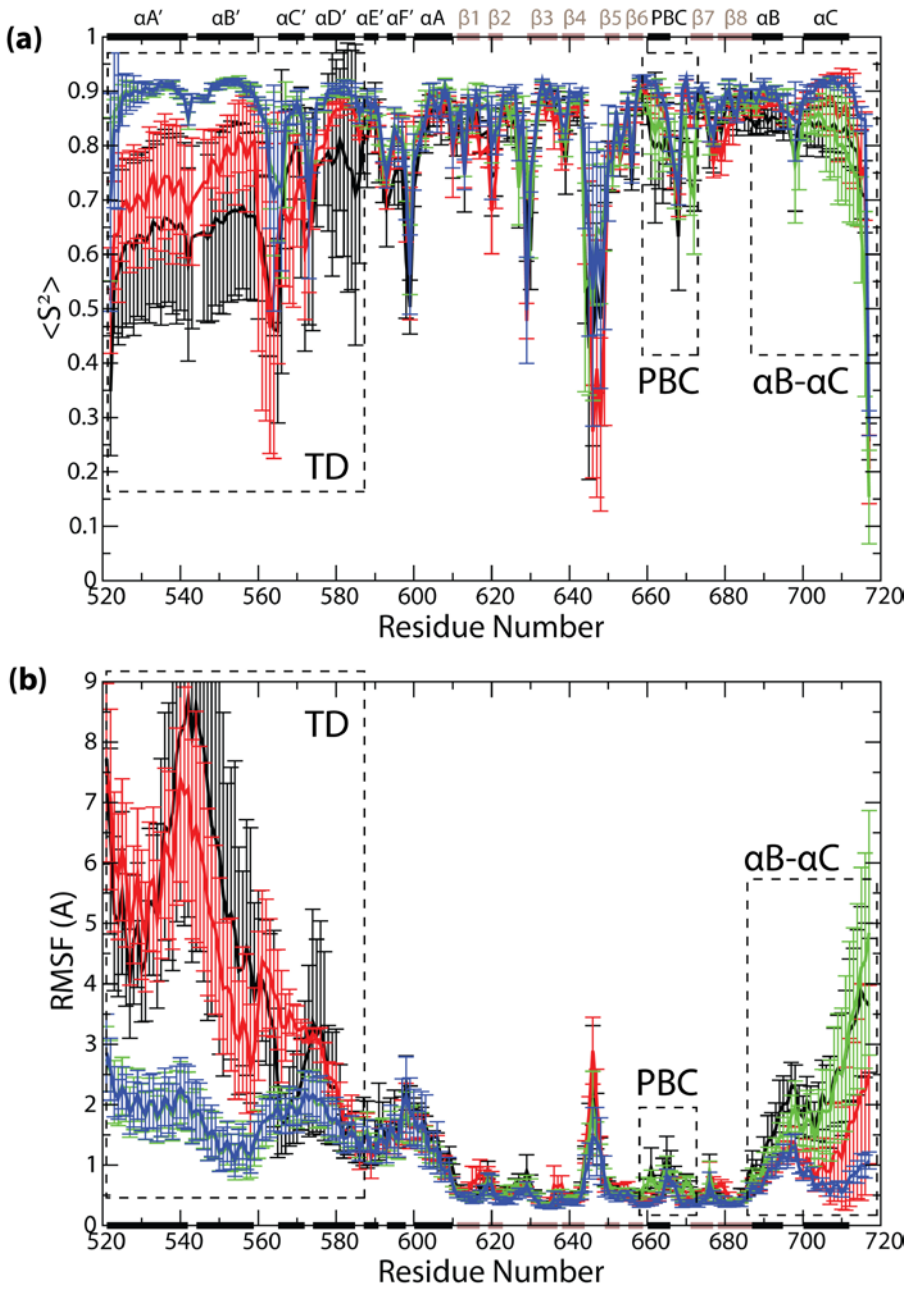

Figure 6: (a) Backbone N-H order parameters (S^2) and **(b)** root-mean-square fluctuations (RMSFs) *vs.* residue number plots, computed with structure overlay at the β-core, and averaged across all protomers and replicates for the respective states. The secondary structure elements are indicated along the horizontal dimensions of both graphs (black bars = α-helices; brown bars = β-strands), and structural regions exhibiting notable differences among states are indicated (dotted rectangles).

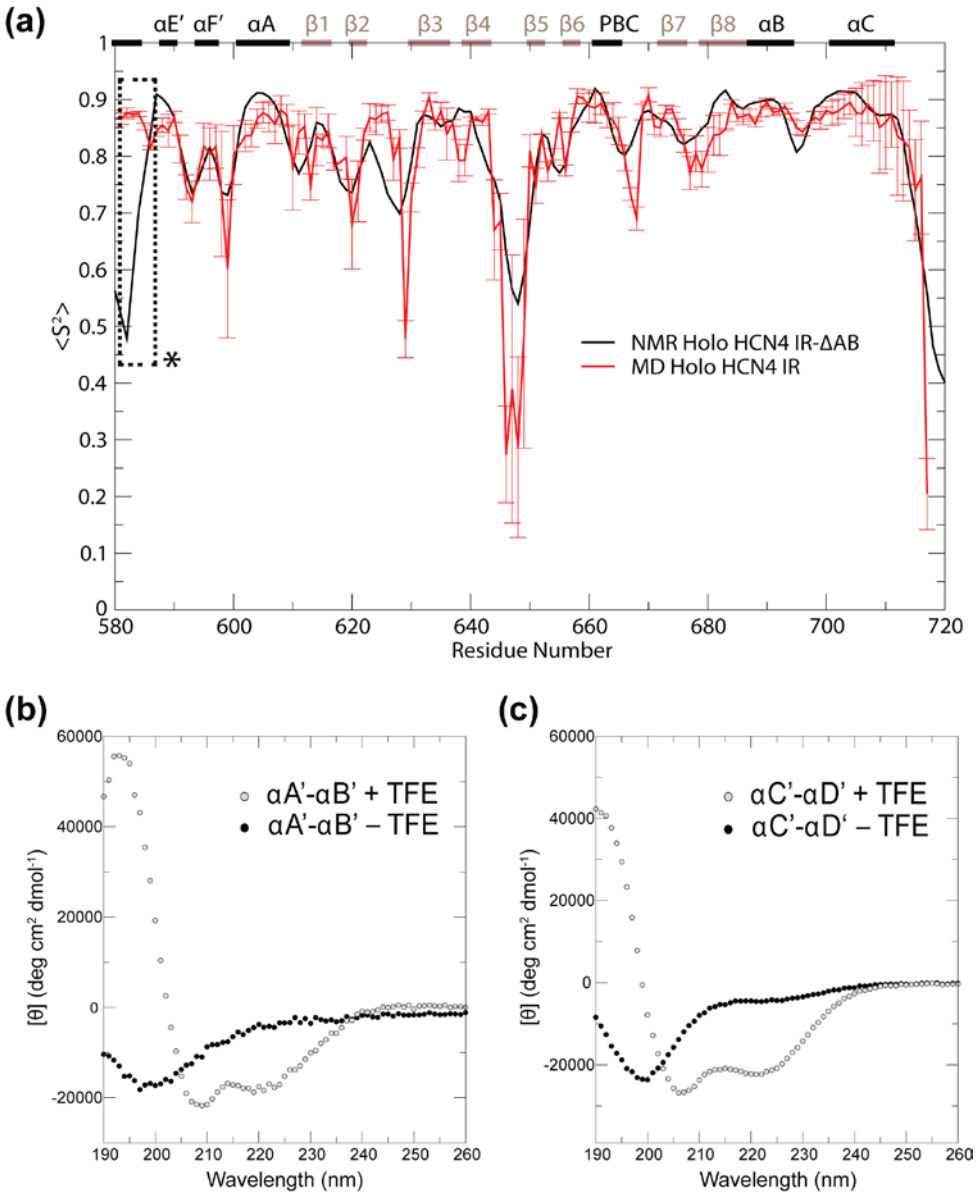

Figure 7: Experimental data supporting the MD simulations of HCN4. **(a)** Comparison of the CBD-region backbone N-H order parameters (S^2) obtained for the monomeric, cAMP-bound HCN4 intracellular region from NMR (black plot) and from MD (red plot). The secondary structure elements are indicated along the top of the graph (black bars = α-helices; brown bars = β-strands). The dashed box marked with an asterisk (*) signifies that the experimental S^2 values for the αD' helix are not reliable due to lack of assignment in this region, in which multiple residues are subject to broadening beyond detection. **(b-c)** CD spectra of synthetic peptides spanning the HCN4 αA'-αB' **(b)** and αC'-αD' **(c)** helices in 10 mM potassium phosphate buffer (pH 6.5) in the presence (open circles) and absence (solid circles) of 40% TFE.

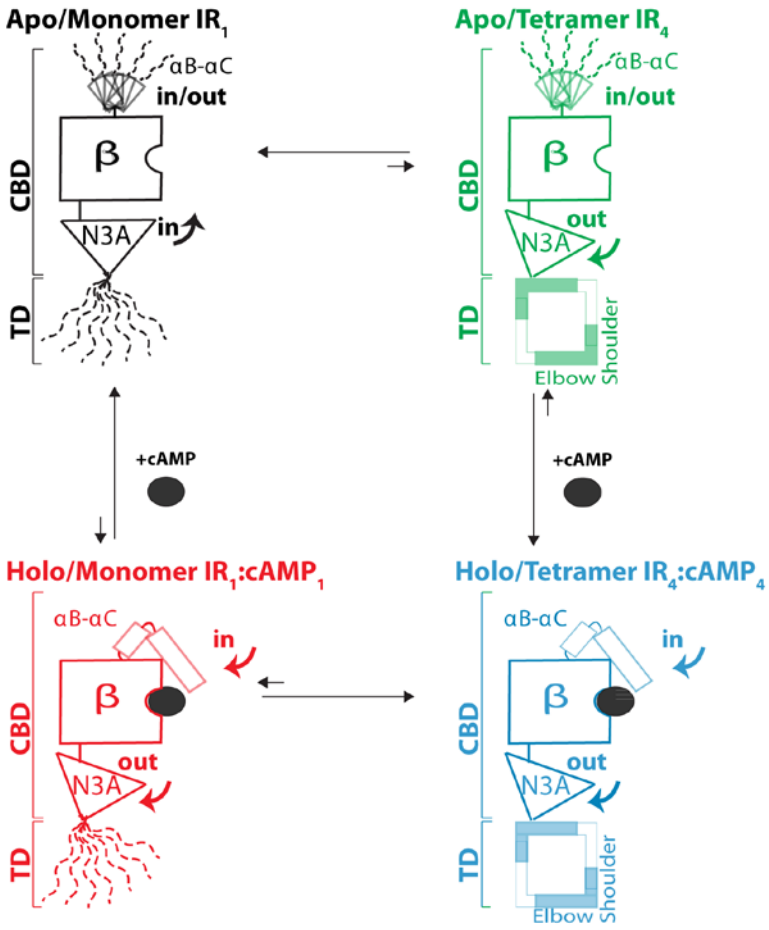

Figure 8: Scheme summarizing the dynamic changes in the HCN4 IR that occur along the thermodynamic cycle for the coupling between tetramerization and cAMP-binding (Figure 1c). Dashed lines denote segments that are at least partially unstructured and highly dynamic. The β-subdomain is denoted as "β" and remains largely invariant across the thermodynamic cycle. For the sake of clarity, only the most populated in/out topologies are shown for the N3A motif. Further details are available in Table 3.

3.9. References

[1] Biel, M., Wahl-Schott, C., Michalakis, S. and Zong, X. (2009) Hyperpolarization-activated cation channels: from genes to function. *Physiol. Rev.* **89**: 847-885.
[2] Craven, K.B. and Zagotta, W.N. (2006) CNG and HCN channels: two peas, one pod. *Annu. Rev. Physiol.* **68**: 375-401.
[3] Zagotta, W.N., Olivier, N.B., Black, K.D., Young, E.C., Olson, R. and Gouaux, E. (2003) Structural basis for modulation and agonist specificity of HCN pacemaker channels. *Nature* **425**: 200-205.
[4] Kusch, J., Biskup, C., Thon, S., Schultz, E., Nache, V., Zimmer, T., Schwede, F. and Benndorf, K. (2010) Interdependence of receptor activation and ligand binding in HCN2 pacemaker channels. *Neuron* **67**: 75-85.
[5] Xu, X., Vysotskaya, Z.V., Liu, Q. and Zhou, Lei. (2010) Structural basis for the cAMP-dependent gating in the human HCN4 channel. *J. Biol. Chem.* **285**: 37082-37091.
[6] Wang, J., Chen, S., Nolan, M.F. and Siegelbaum, S.A. (2002) Activity-dependent regulation of HCN pacemaker channels by cyclic AMP: signaling through dynamic allosteric coupling. *Neuron* **36**: 451-461.
[7] Bois, P., Chatelier, A., Bescond, J. and Faivre, J.F. (2011) In *Ion Channels and Their Inhibitors* (Gupta, S.P., ed.) pp. 33-51, Springer-Verlag, Berlin.
[8] Rosenbaum, T. and Gordon, S.E. (2004) Quickening the pace: looking into the heart of HCN channels. *Neuron* **42**: 193-196.
[9] Johnson, J.P. and Zagotta, W.N. (2005) The carboxyl-terminal region of cyclic nucleotide-modulated channels is a gating ring, not a permeation path. *Proc. Natl. Acad. Sci. USA* **102**: 2742-2747.
[10] Biel, M., Schneider, A. and Wahl, C. (2002) Cardiac HCN channels: structure, function, and modulation. *Trends Cardiovasc. Med.* **12**: 206-213.
[11] Qu, Y., Whitaker, G.M., Hove-Madsen, L., Tibbits, G.F. and Accili, E.A. (2008) Hyperpolarization-activated cyclic nucleotide-modulated 'HCN' channels confer regular and faster rhythmicity to beating mouse embryonic stem cells. *J. Physiol.* **586**: 701-716.
[12] Hegle, A.P., Nazzari, H., Roth, A., Angoli, D. and Accili, E.A. (2010) Evolutionary emergence of N-glycosylation as a variable promoter of HCN channel surface expression. *Am. J. Physiol. Cell. Physiol.* **298**: C1066-C1076.
[13] Berrera, M., Pantano, S. and Carloni, P. (2006) CAMP modulation of the cytoplasmic domain in the HCN2 channel investigated by molecular simulations. *Biophys. J.* **90**: 3428-3433.
[14] Craven, K.B. and Zagotta, W.N. (2004) Salt bridges and gating in the COOH-terminal region of HCN2 and CNGA1 channels. *J. Gen. Physiol.* **124**: 663-677.
[15] Chow, S.S., VanPetegem, F. and Accili, E.A. (2012) Energetics of cyclic AMP binding to HCN channel C terminus reveal negative cooperativity. *J. Biol. Chem.* **287**: 600-606.
[16] Lolicato, M., Nardini, M., Gazzarini, S., Möller, S., Bertinetti, D., Herberg, F.W., Bolognesi, M., Martin, H., Fasolini, M., Bertrand, J.A., Arrigoni, C., Thiel, G. and Moroni, A. (2011) Tetramerization dynamics of C-terminal domain underlies isoform-specific cAMP gating in hyperpolarization-activated cyclic nucleotide-gated channels. *J. Biol. Chem.* **286**: 44811-44820.

[17] Akimoto, M., Zhang, Z., Boulton, S., Selvaratnam, R., VanSchouwen, B., Gloyd, M., Accili, E.A., Lange, O.F. and Melacini, G. (2014) A mechanism for the auto-inhibition of hyperpolarization-activated cyclic nucleotide-gated (HCN) channel opening and its relief by cAMP. *J. Biol. Chem.* **289**: 22205-22220.
[18] Huang, G.Y., Kim, J.J., Reger, A.S., Lorenz, R., Moon, E.W., Zhao, C., Casteel, D.E., Bertinetti, D., VanSchouwen, B., Selvaratnam, R., Pflugrath, J.W., Sankaran, B., Melacini, G., Herberg, F.W. and Kim, C. (2014) Structural basis for cyclic-nucleotide selectivity and cGMP-selective activation of PKG I. *Structure* **22**: 116-124.
[19] Kim, C., Cheng, C.Y., Saldanha, S.A. and Taylor, S.S. (2007) PKA-I holoenzyme structure reveals a mechanism for cAMP-dependent activation. *Cell* **130**: 1032-1043.
[20] Su, Y., Dostmann, W.R., Herberg, F.W., Durick, K., Xuong, N.H., Ten-Eyck, L., Taylor, S.S. and Varughese, K.I. (1995) Regulatory subunit of protein kinase A: structure of deletion mutant with cAMP binding domains. *Science* **269**: 807-813.
[21] Das, R., Esposito, V., Abu-Abed, M., Anand, G.S., Taylor, S.S. and Melacini, G. (2007) cAMP activation of PKA defines an ancient signaling mechanism. *Proc. Natl. Acad. Sci. USA* **104**: 93-98.
[22] Das, R. and Melacini, G. (2007) A model for agonism and antagonism in an ancient and ubiquitous cAMP-binding domain. *J. Biol. Chem.* **282**: 581-593.
[23] Akimoto, M., Selvaratnam, R., McNicholl, E.T., Verma, G., Taylor, S.S. and Melacini, G. (2013) Signaling through dynamic linkers as revealed by PKA. *Proc. Natl. Acad. Sci. USA* **110**: 14231-14236.
[24] Akimoto, M., Moleschi, K., Boulton, S., VanSchouwen, B., Selvaratnam, R., Taylor, S.S. and Melacini, G. (2014) Allosteric linkers in cAMP signalling. *Biochem. Soc. Trans.* **42**: 139-144.
[25] Abu-Abed, M., Das, R., Wang, L. and Melacini, G. (2007) Definition of an electrostatic relay switch critical for the cAMP-dependent activation of protein kinase A as revealed by the D170A mutant of RI alpha. *Proteins* **69**: 112-124.
[26] McNicholl, E.T., Das, R., SilDas, S., Taylor, S.S. and Melacini, G. (2010) Communication between tandem cAMP binding domains in the regulatory subunit of protein kinase A-I alpha as revealed by domain-silencing mutations. *J. Biol. Chem.* **285**: 15523-15537.
[27] VanSchouwen, B., Selvaratnam, R., Fogolari, F. and Melacini, G. (2011) Role of dynamics in the autoinhibition and activation of the exchange protein directly activated by cyclic AMP (EPAC). *J. Biol. Chem.* **286**: 42655-42669.
[28] Wainger, B.J., DeGennaro, M., Santoro, B., Siegelbaum, S.A. and Tibbs, G.R. (2001) Molecular mechanism of cAMP modulation of HCN pacemaker channels. *Nature* **411**: 805-810.
[29] Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**: 2714-2723.
[30] Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L. and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**: 1781-1802.
[31] The Shared Hierarchical Academic Research Computing Network (SHARCNET): www.sharcnet.ca.

[32] MacKerell Jr., A. D., Bashford, D., Bellott, M., Dunbrack Jr., R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D. and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**: 3586-3616.
[33] MacKerell Jr., A.D., Feig, M. and Brooks III, C.L. (2004) Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **25**: 1400-1415.
[34] Foloppe, N. and MacKerell Jr., A.D. (2000) All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem.* **21**: 86-104.
[35] MacKerell Jr., A.D. and Banavali, N.K. (2000) All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *J. Comput. Chem.* **21**: 105-120.
[36] Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Molec. Graphics* **14**: 33-38.
[37] Darden, T., York, D. and Pederson, L. (1993) Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**: 10089-10092.
[38] Miyamoto, S. and Kollman, P.A. (1992) Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **13**: 952-962.
[39] Ryckaert, J.P., Ciccotti, G. and Berendsen, H.J.C. (1977) Numerical algorithms of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**: 327–341.
[40] Tuckerman, M., Berne, B.J. and Martyna, G.J. (1992) Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **97**: 1990-2001.
[41] Izaguirre, J.A., Catarello, D.P., Wozniak, J.M. and Skeel, R.D. (2001) Langevin stabilization of molecular dynamics. *J. Chem. Phys.* **114**: 2090-2098.
[42] Brooks, B.R., Brooks III, C.L., Mackerell Jr., A.D., Nilsson, L., Petrella, R.J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A.R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R.W., Post, C.B., Pu, J.Z., Schaefer, M., Tidor, B., Venable, R.M., Woodcock, H.L., Wu, X., Yang, W., York, D.M. and Karplus, M. (2009) CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30**: 1545-1614.
[43] García de la Torre, J., Huertas, M.L. and Carrasco, B. (2000) HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.* **147**: 138-146.
[44] Berjanskii, M.V. and Wishart, D.S. (2005) A simple method to predict protein flexibility using secondary chemical shifts. *J. Am. Chem. Soc.* **127**: 14970-14971.
[45] Taraska, J.W., Puljung, M.C., Olivier, N.B., Flynn, G.E. and Zagotta, W.N. (2009) Mapping the structure and conformational movements of proteins with transition metal ion FRET. *Nat. Methods* **6**: 532-537.

[46] Puljung, M.C., DeBerg, H.A., Zagotta, W.N. and Stoll, S. (2014) Double electron-electron resonance reveals cAMP-induced conformational change in HCN channels. *Proc. Natl. Acad. Sci. USA* **111**: 9816-9821.
[47] Saponaro, A., Pauleta, S.R., Cantini, F., Matzapetakis, M., Hammann, C., Donadoni, C., Hu, L., Thiel, G., Banci, L., Santoro, B. and Moroni, A. (2014) Structural basis for the mutual antagonism of cAMP and TRIP8b in regulating HCN channel function. *Proc. Natl. Acad. Sci. USA* **111**: 14577-14582.
[48] Selvaratnam, R., VanSchouwen, B., Fogolari, F., Mazhab-Jafari, M.T., Das, R. and Melacini, G. (2012) The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants. *Biophys. J.* **102**: 630-639.
[49] Rehmann, H., Wittinghofer, A. and Bos, J.L. (2007) Capturing cyclic nucleotides in action: snapshots from crystallographic studies. *Nat. Rev. Mol. Cell. Biol.* **8**: 63-73.
[50] Nussinov, R., Tsai, C.J. and Liu, J. (2014) Principles of allosteric interactions in cell signaling. *J. Am. Chem. Soc.* **136**: 17692-17701.
[51] Nussinov, R. and Tsai, C.J. (2014) Unraveling structural mechanisms of allosteric drug action. *Trends Pharmacol. Sci.* **35**: 256-264.
[52] Tsai, C.J. and Nussinov, R. (2014) A unified view of "how allostery works". *PLoS Comput. Biol.* 10.1371/journal.pcbi.1003394.
[53] Nussinov, R., Ma, B. and Tsai, C.J. (2014) Multiple conformational selection and induced fit events take place in allosteric propagation. *Biophys. Chem.* **186**: 22-30.
[54] Srivastava, A.K., McDonald, L.R., Cembran, A., Kim, J., Masterson, L.R., McClendon, C.L., Taylor, S.S. and Veglia, G. (2014) Synchronous opening and closing motions are essential for cAMP-dependent protein kinase A signaling. *Structure* **22**: 1735-1743.
[55] Veglia, G. and Cembran, A. (2013) Role of conformational entropy in the activity and regulation of the catalytic subunit of protein kinase A. *FEBS J.* **280**: 5608-5615.
[56] Cembran, A., Masterson, L.R., McClendon, C.L., Taylor, S.S., Gao, J. and Veglia, G. (2012) Conformational equilibrium of N-myristoylated cAMP-dependent protein kinase A by molecular dynamics simulations. *Biochemistry* **51**: 10186-10196.
[57] De, S., Greenwood, A.I., Rogals, M.J., Kovrigin, E.L., Lu, K.P. and Nicholson, L.K. (2012) Complete thermodynamic and kinetic characterization of the isomer-specific interaction between Pin1-WW domain and the amyloid precursor protein cytoplasmic tail phosphorylated at Thr668. *Biochemistry* **51**: 8583-8596.
[58] Blyuss, K.B. and Nicholson, L.B. (2012) The role of tunable activation thresholds in the dynamics of autoimmunity. *J. Theor. Biol.* **308**: 45-55.
[59] Jayaraman, B. and Nicholson, L.K. (2007) Thermodynamic dissection of the ezrin FERM/CERMAD interface. *Biochemistry* **46**: 12174-12189.
[60] Ishiyama, N., Lee, S.H., Liu, S., Li, G.Y., Smith, M.J., Reichardt, L.F. and Ikura, M. (2010) Dynamic and static interactions between p120 catenin and E-cadherin regulate the stability of cell-cell adhesion. *Cell* **141**: 117-128.
[61] Smith, M.J. and Ikura, M. (2014) Integrated RAS signaling defined by parallel NMR detection of effectors and regulators. *Nat. Chem. Biol.* **10**: 223-230.

CHAPTER 4

MECHANISM OF cAMP PARTIAL AGONISM IN PROTEIN KINASE G (PKG)

4.1. Author's Preface

The work presented in this chapter has been accepted for publication in the Journal of Biological Chemistry, and is reproduced here with permission from the Journal of Biological Chemistry. The current citation is as follows:

VanSchouwen, B., Selvaratnam, R., Giri, R., Lorenz, R., Herberg, F.W., Kim, C. and Melacini, G. (2015) "Mechanism of cAMP partial agonism in protein kinase G (PKG)." *Journal of Biological Chemistry*: published online ahead of print (doi: 10.1074/jbc.M115.685305).

I conducted most of the work necessary for this publication. Dr. Rajeevan Selvaratnam performed initial optimization work for the protein construct used in the experiments, as well as initial NMR analysis and spectrum assignment work. Dr. Rajanish Giri performed additional NMR sample preparation work. Robin Lorenz and Dr. Friedrich W. Herberg performed and analyzed the kinase activity assays for intact PKG. Dr. Choel Kim designed and constructed the plasmid DNA vectors used for protein expression. I co-wrote the manuscript with Drs. Friedrich W. Herberg, Choel Kim and Giuseppe Melacini.

4.2. Abstract

Protein kinase G (PKG) is a major receptor of cGMP and controls signaling pathways often distinct from those regulated by cAMP. Hence, the selective activation of PKG by cGMP *vs.* cAMP is critical. However, the mechanism of cGMP-*vs.*-cAMP selectivity is only limitedly understood. Although the C-terminal cyclic-nucleotide-binding domain of PKG

(CNB-B) binds cGMP with higher affinity than cAMP, the intracellular concentrations of cAMP are typically higher than those of cGMP, suggesting that the cGMP-*vs.*-cAMP selectivity of PKG is not controlled uniquely through affinities. Here, we show that cAMP is a partial agonist for PKG, and we elucidate the mechanism for cAMP partial agonism through the comparative NMR analysis of the apo, cGMP- and cAMP-bound forms of PKG CNB-B. We show that although cGMP-activation is adequately explained by a two-state conformational selection model, the partial agonism of cAMP arises from the sampling of a third, partially autoinhibited state.

4.3. Introduction

Protein kinase G (PKG) is a major receptor of the cGMP second messenger. By binding to PKG, cGMP regulates intracellular signaling pathways that control a wide range of intracellular processes, such as cell differentiation, platelet activation, memory formation and vasodilation.[1,2,3,4,5] The PKG signaling pathways are often distinct from those regulated by cAMP-dependent proteins such as protein kinase A (PKA).[6,7,8,9,10,11,12] Thus, a key element of the cross-talk between PKG and PKA signaling pathways is the selective activation of PKG by cGMP rather than by cAMP. However, the mechanism of cAMP-*versus*-cGMP selectivity in PKG is currently only limitedly understood.

One possible determinant of cyclic nucleotide selectivity is the lower binding affinity of PKG for cAMP than for cGMP.[13,14,15,16] Indeed, the C-terminal cyclic-nucleotide-binding domain (CNB-B) of PKG Iβ was previously found to exhibit greater binding affinity for cGMP, which has been attributed to specific interactions between the base-binding region (BBR) of CNB-B and the cGMP base moiety, as observed in previously solved

crystallographic X-ray structures (Figure 1e).[17,18] However, the intracellular concentration of cAMP is typically significantly greater than that of cGMP,[19,20,21,22,23] thus potentially allowing cAMP to bind to PKG despite its lower binding affinity and suggesting that other means of cyclic nucleotide selectivity must be present in PKG.

Another key difference in the response of PKG Iβ to cAMP *vs.* cGMP is that in contrast to cGMP, cAMP is only a partial agonist of PKG Iβ, as shown by Cook photometric kinase assays of PKG Iβ activity (Figure 1f). Such partial agonism of cAMP is not fully explained by previously solved structures of PKG Iβ, as they reveal very similar structural rearrangements of both CNB domains in response to either cAMP or cGMP binding (Figure 1b-d)[2,17,18] that are analogous to structural rearrangements observed for other CNB domains.[8,12,24,25,26,27,28,29] Hence, we hypothesize that the partial-agonist response of PKG Iβ to cAMP may reflect differences in the dynamics of cAMP- *vs.* cGMP-bound PKG Iβ. For example, in the context of the two-state conformational selection mechanism, the partial agonism of cAMP could simply arise from a reduced inactive-to-active shift of the dynamic conformational equilibrium underlying PKG activation.[30,31,32,33] Indeed, such conformational dynamics have been shown to play a role in the varying responses of other CNB domains to cyclic nucleotides.[13,34] Alternatively, in the presence of cAMP the regulatory region of PKG may dynamically sample a third conformation, which is distinct from both the apo-inactive and cGMP-bound active structures,[30] and is at least partially auto-inhibitory of the kinase function of PKG.

As a first step towards testing our hypotheses and differentiating between the two- *vs.* three-state hypothetical models for the partial agonism of cAMP, here we report the comparative NMR analyses of the apo, cAMP- and cGMP-bound forms of the PKG Iβ CNB-

B domain (*i.e.* "PKG Iβ(219-369)"; see Figure 1a). This domain is adjacent to the catalytic domain (Figure 1a), and has been previously identified as a critical control unit for auto-inhibition and cGMP selectivity in PKG Iβ.[17,18] In addition, we examined the CNB-B domain because, although the CNB-A domain contributes to the cyclic nucleotide modulation of PKG I function,[17,35,36,37] the CNB-A domain undergoes only minor structural changes upon cyclic nucleotide binding.[2,17] We found that the partial agonism of cAMP cannot be explained by a simple reversal of the two-state inactive/active conformational equilibrium that rationalizes cGMP-activation.[2,17] Rather, our data show that the partial agonism exhibited by cAMP is the result of cAMP-bound CNB-B sampling a distinct third conformational state. In this third state, the lid region that caps the cyclic nucleotide base in the fully active state is disengaged and kinase inhibition is only partially effective. By populating this third conformation, and to a lesser degree the inactive conformation, cAMP reduces the overall level of PKG activity compared to cGMP, thereby providing an explanation for the observed partial agonism of cAMP.

4.4. Experimental Procedures

Preparation of the PKG Iβ(219-369) construct. The PKG Iβ(219-369) construct was expressed with an N-terminal poly-L-histidine tag, in *Escherichia coli* strain BL-21(DE3). The *E. coli* cells were grown at 37 °C in isotopically enriched minimal media supplemented with trace metals, D-biotin and thiamine-HCl. Expression was induced with 0.4 mM isopropyl β-D-1-thiogalactopyranoside at an optical density of 1.0 (at λ = 600 nm), and the cells were further incubated for 16 hours at 20 °C before being harvested by centrifugation.

The harvested cells were lysed using a cell disruptor, and cell debris was subsequently removed by centrifugation (20,000xg for one hour). For this process, the cells were re-suspended in 50 mM Tris buffer (pH 7.6) with 150 mM NaCl, 1 mM β-mercaptoethanol and 0.2 mM AEBSF. An initial purification of the PKG construct was then performed using a Ni^{2+}–Sepharose resin in a 5-mL gravity column. After passing the cell lysate through the column, the column was rinsed with more cell resuspension buffer, followed by 50 mM Tris buffer (pH 7.6) with 500 mM NaCl, 20 mM imidazole and 1 mM β-mercaptoethanol. The protein was then eluted from the column using 50 mM Tris buffer (pH 7.6) with 50 mM NaCl, 300 mM imidazole and 1 mM β-mercaptoethanol.

The collected gravity column eluant was dialyzed in 50mM Tris buffer (pH 7.6) with 100mM NaCl and 1mM β-mercaptoethanol, and cleaved with TEV protease for 48 hours to remove the N-terminal poly-L-histidine tag from the PKG Iβ(219-369) construct. The cleaved poly-L-histidine tag and TEV protease were removed from the dialyzed eluant by passing the eluant through a Ni^{2+}–Sepharose gravity column, followed by a rinse with dialysis buffer. All column flow-through and rinse were collected, and the PKG Iβ(219-369) construct was further purified by size-exclusion FPLC (HiLoad Superdex 120-mL column). The FPLC was performed using a 50 mM Tris running buffer (pH 7.0) with 100 mM NaCl, 1 mM DTT and 0.02%(w/v) NaN_3, which also served as the NMR sample buffer for most of the NMR experiments (except for H/D hydrogen exchange, which was performed in a pH 7.0 ^2H_2O-based phosphate buffer).

Spectrophotometric kinase activity assay. Kinase activity was measured using the coupled spectrophotometric assay as described by Cook et al. (1982).[38] The

phosphotransferase activity of full-length PKG Iβ purified from Sf9 cells was determined in 100 µL reaction volumes containing 35 nM PKG Iβ, 1 mM Kemptide (LRRASLG), varying concentrations of cGMP or cAMP, respectively, and the assay mix. The depletion of NADH directly correlating with the conversion of ATP was monitored using a spectrophotometer (SPECORD® 205; Analytik Jena). The buffer used for the spectrophotometric assay consisted of 100 mM MOPS (pH 7), 10 mM $MgCl_2$, 1 mM phosphoenolpyruvate, 1 mM ATP, 150 U lactate dehydrogenase, 84 U pyruvate kinase, 220 µM NADH and 5 mM β-mercaptoethanol. Phosphotransferase activity was plotted against the logarithmic cyclic nucleotide concentration, and data points were fitted using a sigmoidal dose-response fit in GraphPad Prism 6.01.

NMR data acquisition and analysis. All NMR spectra for the PKG Iβ(219-369) construct were acquired at 306 K with a Bruker Avance 700-MHz NMR spectrometer equipped with a 5 mm TCI cryoprobe. For the chemical-shift-based analyses of PKG Iβ(219-369), two-dimensional (^{1}H,^{15}N) HSQC NMR spectra were acquired for apo, cAMP- and cGMP-bound samples of PKG Iβ(219-369). A PKG Iβ(219-369) concentration of 20 µM was used for all HSQC samples, and 80 µM of ^{15}N-labelled N-acetylglycine was added to the samples as an internal reference for subsequent alignment of the HSQC spectra with one another. The experiments for cAMP- and cGMP-bound PKG Iβ(219-369) were performed with a cyclic nucleotide concentration of 2 mM in order to achieve saturated ligand binding, thus minimizing the influence of differing ligand binding affinities on the results of the comparative analyses. The spectra were processed with NMRPipe,[39] and analyzed using Sparky.[40] Peak assignments for the HSQC spectra were obtained from

standard three-dimensional triple-resonance NMR spectra (*i.e.* HNCACB, CBCA(CO)NH, HNCA, HN(CO)CA), using the automated PINE-NMR server[41] to facilitate the initial assignment process. The assigned HSQC spectra were aligned using Sparky, and the resulting (^{1}H,^{15}N) chemical shift data for apo, cAMP- and cGMP-bound PKG Iβ(219-369) were utilized to compute (^{1}H,^{15}N) combined chemical shifts, perform chemical shift projection analyses (CHESPA), and build chemical shift correlation plots, as previously described.[13,42,43,44,45] The CHESPA "outlier" residues are defined as those deviating significantly from the two-state equilibrium between conformations resembling the apo and cGMP-bound PKG Iβ(219-369) structures (*i.e.* residues with projection angle cosine values less than 0.8). The chemical shift correlation plots were confined to residues with CHESPA fractional activation values less than 1.2 and projection angle cosine values greater than 0.95.

To further examine PKG Iβ(219-369) dynamics, backbone ^{15}N relaxation (R_1, R_2, HN-NOE) and hydrogen-exchange (H/D and H/H) NMR data for the apo, cAMP- and cGMP-bound samples of PKG Iβ(219-369) were acquired and analyzed using a protocol similar to that described previously,[13,46,47] with the T_1 and T_2 data sets acquired as pseudo-3D matrices.[48] Reduced spectral densities (RSDs) were then computed from the ^{15}N relaxation rates as described previously.[13,46,47] Considering the rigidity typically exhibited by the inner β-strands (*i.e.* β3, β4, β7 and β8) of the β-barrel of CNB domains,[27,46] R_1R_2 and/or J(0) values significantly greater than the inner-β-strand average were considered indicative of the presence of ms-μs backbone dynamics, while R_1R_2, J(0) and/or HN-NOE values significantly less than the inner-β-strand average, and/or $J(\omega_H + \omega_N)$ values significantly greater than the inner-β-strand average, are indicative of the presence of ps-ns backbone dynamics. Differences in the R_1R_2 and J(0) values between different samples of PKG Iβ(219-369) were

deemed significant if the error bars of the compared data points did not overlap one another. Meanwhile, due to the larger errors affecting the HN-NOE and $J(\omega_H + \omega_N)$ values, differences in these values were deemed significant if the error bar of one of the compared data points did not overlap the other data point, and *vice versa*. Finally, significant differences in hydrogen exchange rates were indicated by non-overlapping protection factor (PF) value error bars, or transitions between different time scales of hydrogen exchange. The latter are defined, from fast to slow exchange, as: H/H-timescale exchange, H/D-dead-time exchange, quantifiable exchange (*i.e.* measurable PF), and finally, exchange that is too slow for reliable PF quantification.

Determination of state populations in the three-state model. The state populations were estimated based on the fractional activations measured through CHESPA. Specifically, the state populations were determined based on the average fractional activations observed for the pre-lid region, which is reflective of the population of the state in which this region adopts an active conformation (*i.e.* the active state), and for the CNB-B domain region preceding residue 340, which is reflective of the total population of states in which this region is active-like (*i.e.* the intermediate and active states). Thus, the population of the intermediate state is reflected by the difference between the two average fractional activations, while the population of the inactive state is the percentage not accounted for by the other two states.

4.5. Results

Initial assessment of cAMP vs. cGMP differences in PKG Iβ(219-369) through compounded chemical shift changes. As an initial assessment of the differences between cAMP- and cGMP-bound PKG Iβ(219-369), we examined the pairwise compounded chemical shift changes between apo, cAMP- and cGMP-bound PKG Iβ(219-369) (Figure 2c). Chemical shifts from the apo state revealed that binding of both cAMP and cGMP results in major, widespread structural changes within the CNB-B domain, including part of the switch helix region (Figure 2c, black and green points), in agreement with the available X-ray structures.[17,18] In addition, as expected, the cAMP-*vs.*-cGMP chemical shifts detected differences for residues at/near the base-binding pocket, including residues in the β5 strand of the BBR, and the cGMP-capping residue from the switch helix region (*i.e.* residue Y351; Figure 2c, red points). However, Figure 2c also shows that the cAMP- and cGMP-bound states exhibit differences that are not apparent from previous structural comparisons. For example, unanticipated cAMP-*vs.*-cGMP chemical shift differences were observed for residues outside the binding site, including the residues spanning the 340-350 region, and the N-terminal N3A motif spanning residues 220-250 (Figure 2c, red points marked with dotted boxes). To better understand the origin of these long-range differences, we analyzed the cAMP-bound chemical shifts using chemical shift projection analysis (CHESPA; Figure 2b,d,e).

The chemical shift projection analysis (CHESPA) of cAMP-bound PKG Iβ(219-369) reveals three major types of long-range perturbations caused by the replacement of cAMP with cGMP. The CHESPA analysis was performed for cAMP-bound PKG Iβ(219-369),

utilizing apo and cGMP-bound PKG Iβ(219-369) as reference states (Figure 2b). The resulting computed fractional activations (Figure 2d) and projection angle cosine values (Figure 2e) were complemented by a cAMP-apo *vs.* cGMP-apo chemical shift correlation plot analysis, whose slope provides an alternative estimation of the relative cAMP-*vs.*-cGMP fractional activation (Figure 2f). Examination of the CHESPA (Figure 2d,e) and chemical shift correlation results (Figure 2f) revealed three groups of non-binding-site residues demonstrating notable cAMP-*vs.*-cGMP differences. The first group includes residues 340-350 (referred to herein as the "pre-lid" region), which link the αB helix to the C-terminal capping residue (*i.e.* Y351; a residue referred to herein as the "lid"). The residues in this first group exhibited an average fractional activation of only ~40% (Figure 2d), as was also evident from partial HSQC peak shifts for these residues (Figure 2h), suggesting that in the presence of cAMP the pre-lid region is subject to a partial disengagement from its active-state structural arrangement. The second group of residues is composed of the residues N-terminal to residue 340, which exhibited an average fractional activation of ~87% in the presence of cAMP (Figure 2d,f,g), reflecting a markedly different degree of activation compared to the pre-lid region. The third notable group spans multiple "outlier" residues, which deviate from a two-state conformational equilibrium in the presence of cAMP, as seen through low projection angle cosine values (Figure 2e, purple highlights) and non-linear arrangements of apo, cAMP- and cGMP-bound state HSQC cross-peaks (Figure 2i,j). The identified "outlier" residues include cGMP-binding residue R297 (from the β5 strand), as well as multiple non-binding-site residues in the pre-lid region N-terminus and the N3A motif (Figure 2e, purple highlights). When mapped on the structure of cGMP-bound PKG Iβ(219-369), the non-binding-site "outlier" residues form a cluster spanning the interface

between the pre-lid region and the N3A motif (Figure 3a, gray surface), indicating that the partial disengagement of the pre-lid region caused by the replacement of cGMP with cAMP may in turn influence the N3A motif as well. In addition, the low degree of fractional activation observed for the pre-lid region (Figure 2d) points to a loss of structure in this segment, suggesting that dynamics are involved in the observed perturbations caused by cAMP. In order to test this hypothesis, we examined PKG Iβ(219-369) dynamics more closely through ^{15}N relaxation NMR measurements.

Assessment of apo-vs.-cGMP differential dynamics though backbone ^{15}N relaxation rates and reduced spectral densities. As a first assessment of PKG Iβ(219-369) dynamics, we measured backbone ^{15}N relaxation rates and reduced spectral densities (RSDs) for the apo, cAMP- and cGMP-bound states (Figures 4 and 5). The apo *vs.* cGMP-bound state comparison revealed several unanticipated changes in dynamics. First, the β4-β5 loop in the base binding region (BBR) exhibited consistent ps-ns dynamics in both the apo and cGMP-bound states (Figure 4c,d and Figure 5a,c, gray highlights). In addition, the phosphate binding cassette (PBC) exhibited a shift from ms-μs to ps-ns dynamics upon cGMP binding (Figure 4c,d and Figure 5a,c, gray highlights), suggesting a retention of residual PBC dynamics. This is in marked contrast to the more quantitative quenching of PBC dynamics observed previously for the bound states of other CNB domains, which unlike PKG, include a proline residue in the PBC.[8,12,27,46,49,50] In addition, our data indicated that the region C-terminal to the αB helix (*i.e.* C-terminal to residue 340) is subject to ps-ns dynamics in the apo state (Figure 4c,d and Figure 5a,c).[17] The binding of cGMP produces a dramatic quench of ps-ns dynamics in the pre-lid region residues (*i.e.* residues 340-350) that bridge the αB

helix to the lid residue Y351 (Figure 4c,d and Figure 5a,c). However, despite this dramatic quenching of dynamics, Y351 retains partial ps-ns flexibility (Figure 4c,d and Figure 5a,c).

Assessment of cAMP-vs.-cGMP differential dynamics though backbone ^{15}N relaxation rates and reduced spectral densities. The region C-terminal to the αB helix is also a site of previously-unanticipated cAMP-*vs.*-cGMP differences in dynamics. In particular, partially enhanced ps-ns dynamics were observed in the cAMP-bound state throughout the region C-terminal to residue 340, including the lid residue Y351 (Figure 4c,d and Figure 5a,c, orange highlights). Furthermore, several other residues exhibited enhanced ms-μs dynamics in the cAMP-bound state (Figure 4c and Figure 5a, red highlights). The latter enhancements include binding-site residues from the β5 strand and the PBC N-terminus, as well as non-binding-site residues such as those in the N3A motif and the pre-lid region N-terminus (Figure 4c and Figure 5a, red highlights). These cAMP-*vs.*-cGMP differences cannot be explained by variations in diffusional anisotropy, given the similarity of the cAMP- and cGMP-bound structures of the CNB-B domain (Figure 1b-d). Therefore, the pre-lid region emerges as a site affected by cAMP-*vs.*-cGMP enhancements in both ps-ns and ms-μs dynamics. This conclusion is further supported by the correlated changes in high- and low-frequency spectral densities for this region (*i.e.* $J(\omega_H+\omega_N)$ *vs.* $J(0)$; Figure 5d).

Figure 5d shows that the cAMP-*vs.*-cGMP changes deviate from the simple rigid rotor model, suggesting the simultaneous presence of both ps-ns and ms-μs flexibility in the pre-lid region of the cAMP-bound state. These observations corroborate the partial disengagement of the lid, as suggested by the CHESPA analysis (Figure 2d). Furthermore, the map of non-binding site loci subject to enhanced ms-μs dynamics in the presence of

cAMP coincides quite well with that of the "outlier" residues identified from CHESPA (Figure 3b, red surfaces *vs.* Figure 3a, gray surfaces). Overall, the data of Figures 4 and 5 confirm that cGMP replacement with cAMP leads to increased structural dynamics not only at binding site residues, but also at selected allosteric sites of PKG Iβ(219-369) (*e.g.* the pre-lid and N3A loci; Figure 3b). Unlike the cAMP-*vs.*-cGMP changes, the apo-*vs.*-cGMP variations in the correlated $J(\omega_H+\omega_N)$ *vs.* $J(0)$ spectral densities follow, within error, the values expected for a rigid rotor in the absence of ms-μs dynamics (Figure 5d, black line), further confirming that cGMP-binding quenches ps-ns dynamics in the pre-lid region.

Further assessment of dynamics through hydrogen exchange rates. Next, we measured the backbone hydrogen exchange rates for apo, cAMP- and cGMP-bound PKG Iβ(219-369) (Figure 6). The comparison of the apo and cGMP-bound states revealed unanticipated changes in dynamics. For example, multiple residues of both the αB helix and the N3A motif showed significantly reduced hydrogen exchange rates, pointing to a reduced access to partially unfolded and solvent exposed excited states upon cGMP binding (Figure 6). However, when cGMP is replaced by cAMP, a partial recovery of solvent exposure is observed, as several residues exhibited greater hydrogen exchange rates in the cAMP-bound state than in the cGMP-bound state (Figure 6). The affected areas include residues in the vicinity of the binding site (*e.g.* the β5 and β6 strands, the β2-3 loop adjacent to the PBC, and the PBC N- and C-termini), as well as non-binding-site residues, such as the N-terminus of the pre-lid region and the N3A motif (Figure 6, red highlights). Overall, the identified non-binding-site residues affected by cAMP overlap quite well with the "outlier" residues identified from CHESPA (Figure 3c, red surfaces with dotted outline, *vs.* Figure 3a, gray

surfaces with dotted outline), suggesting that in agreement with the ^{15}N relaxation and RSD results, the cAMP-*vs.*-cGMP perturbations observed for PKG Iβ(219-369) involve an increase in dynamics at the interface between the α-helical regions flanking the β-barrel (Figure 3). Furthermore, the pre-lid and lid regions exhibited high solvent exposure in all three states, as suggested by fast hydrogen exchange (Figure 6), thus confirming the finding from ^{15}N relaxation and RSDs that the lid retains significant dynamics even in the cGMP-bound state.

4.6. Discussion

The comparative NMR analysis of the CNB-B construct (*i.e.* PKG Iβ(219-369)) in its apo, cGMP- and cAMP-bound states revealed that the cAMP-*vs.*-cGMP differences are more extensive than previously anticipated, reaching well beyond the cGMP contact sites to crucial allosteric structural elements. In particular, the chemical shift, ^{15}N relaxation and hydrogen-exchange data (Figures 2 and 4-6) consistently show that cAMP binding results in only a partial recruitment of the pre-lid region to its active-state conformation, which in turn leads to perturbations of the adjacent N3A region (Figure 3). Furthermore, when bound to the PKG Iβ CNB-B domain, cAMP exhibits a higher propensity for an *anti* base orientation compared to cGMP, which mainly adopts a *syn* conformation.[17,18] A possible explanation for the differing *syn*/*anti* conformational propensities of bound cAMP *vs.* cGMP is the intrinsic preference of free cAMP for the *anti* orientation together with a lack of stable interactions between the 6-NH_2 moiety of cAMP, which is replaced by a less bulky hydrogen-bond-accepting 6-*oxo* moiety in cGMP, and the residues of the cyclic nucleotide binding site responsible for cGMP selectivity, including residue R297 (in β5), which interacts with the 6-

oxo moiety of cGMP (Figure 1e). The *anti vs. syn* difference may reduce the efficacy with which the lid residue (*i.e.* Y351) is recruited to its active-state capping interaction with the bound cyclic nucleotide, thus providing an explanation for the partial pre-lid region disengagement observed in the cAMP-bound state. Indeed, the analogous capping residue of the cAMP-bound PKA CNB-B domain (*i.e.* Y371) adopts a structural arrangement similar to that of Y351 in the cGMP-bound PKG CNB-B domain, and the bound cyclic nucleotides in both of these domains adopt a *syn* conformation,[17] suggesting that replacement of cGMP by cAMP without a *syn*-to-*anti* transition is not sufficient to achieve the observed lid disengagement.

The inability of cAMP to recruit the pre-lid region into the active state to the same extent as observed for the majority of the CNB-B domain residues preceding residue 340 (Figure 2d) suggests that the partial agonism of cAMP cannot be explained by a simple two-state active-inactive model of conformational selection. A purely two-state model dictates that all of the CNB-B domain residues should exhibit a similar fractional activation, which is not the case here (Figure 2d). Hence, the dual pattern of fractional activations observed in Figure 2d (horizontal arrows) implies the presence of a third state in which kinase inhibition is only partially effective. The deviation from a two-state model is also independently confirmed by the observation for multiple non-binding-site residues of clear deviations from a linear apo-cAMP-cGMP CHESPA pattern (Figure 2e,i,j). These results consistently suggest that the partial-agonist response of PKG Iβ to cAMP (Figure 1f) cannot be explained by a simple reversal of the two-state inactive/active conformational equilibrium (Figure 7, red and green structures). Instead, our results point to a dynamic exchange of the cAMP-

bound CNB-B domain with a third, "intermediate" conformational state (Figure 7, orange structure), leading to a three-state model (Figure 7).

In the proposed three-state model (Figure 7), the population of the intermediate state is ~47%, while that of the apo-like inactive state is ~13%, as assessed based on the fractional activations measured through CHESPA (Figure 2d). The intermediate state sampled in the presence of cAMP maintains an active-like conformation with the exception of the pre-lid region (*i.e.* residues 340-350). Upon cAMP binding, the pre-lid region becomes disengaged and more dynamic in the ps-ns and ms-μs time-scales (Figure 7, orange structure). Notably, the disengagement of the pre-lid region would alter the relative positions of the PKG regulatory and catalytic regions, which are connected to one another by the CNB-B domain/switch helix region (Figure 1a).[2,4,17,18] Such repositioning is likely to affect the access of substrates to the PKG catalytic site,[6] explaining why the intermediate state retains at least partial kinase inhibition. Hence, by promoting population of the "intermediate" conformation, and to a lesser degree of the inactive conformation (Figure 7), cAMP reduces the overall level of PKG activity compared to cGMP, explaining the partial agonism observed for cAMP (Figure 1f). Such partial agonism represents a key determinant of the cGMP-*vs.*-cAMP selectivity of PKG, and together with differing binding affinities for cAMP *vs.* cGMP,[17,18] contributes to the minimization of cross-talk between cGMP- and cAMP-controlled signaling pathways.

Finally, it is worth noting that while the partial agonism of cAMP contributes to cyclic nucleotide selectivity for cGMP-controlled signaling pathways, previous studies have suggested that selected signaling responses are influenced by cAMP *via* modulation of PKG activity, rather than cAMP-dependent proteins such as PKA.[5,51] A possible explanation for

this phenomenon is that as the intracellular cAMP concentration rises, the minimal cGMP concentration necessary for full PKG activation increases, and the shape of the cGMP-dependent activation profile of PKG may also change. Such modulation of PKG activity by cAMP represents a notable contribution to the control of PKG, and may thus have relevant physiological implications for cGMP-dependent intracellular signaling.

4.7. Acknowledgements

We thank Professors Susan S. Taylor (UCSD) and W. Dostmann (University of Vermont), as well as M. Akimoto, S. Boulton and K. Moleschi (McMaster University), for helpful discussions. This project was funded by a grant to G.M. from the Canadian Institutes of Health Research (CIHR) and from the Natural Sciences and Engineering Research Council of Canada (NSERC). F.W.H. is funded by the Federal Ministry of Education and Research (fund number 0316177F No Pain), and R.L. by the graduate program “Functionomics” at Kassel University.

4.8. Figures

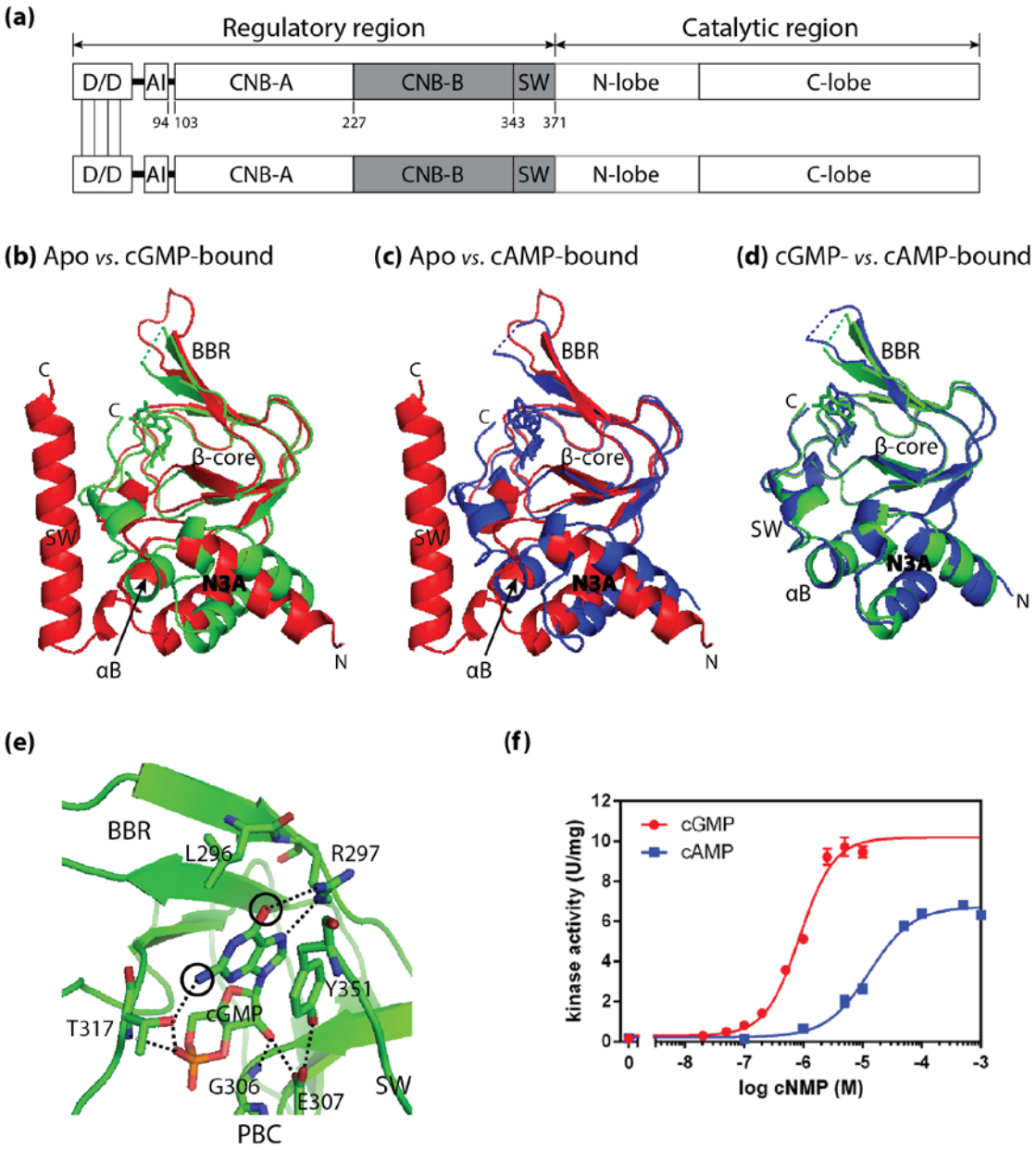

Figure 1: Introduction to PKG Iβ domain and structural architecture and to cGMP *vs.* cAMP binding and activation. **(a)** Schematic overview of the domain organization of PKG Iβ. The regulatory and catalytic regions of PKG are indicated, as are the major underlying structural domains: the N-terminal dimerization domain (D/D); the autoinhibitory linker region (AI); the cyclic-nucleotide-binding domains (CNB-A and CNB-B); the switch helix region (SW); and the two lobes of the catalytic region (N-lobe and C-lobe). Residue numbers for the regulatory region domain boundaries are indicated, and the monomer fragment examined in

the current study (referred to herein as "PKG Iβ(219-369)") is highlighted in gray. **(b-d)** Ribbon-structure overlays of the previously-solved apo (red ribbon), cAMP-bound (blue ribbon), and cGMP-bound (green ribbon) structures of PKG Iβ(219-369).[17,18] Bound cAMP and cGMP are shown as blue and green sticks, respectively, and the following key structural elements are indicated: the N-terminal α-helix bundle (N3A), the β-barrel (β-core); the base binding region (BBR) element involved in cGMP binding; and the C-terminal αB helix (αB) and switch helix region (SW). To clearly demonstrate the differences in structure, the structures are overlaid at their β-barrels. **(e)** Zoomed-in view of the cGMP-binding site of the cGMP-bound structure. Key cGMP-interacting residues[17] are indicated as sticks, hydrogen bonds as black dotted lines, and cGMP functional groups that are different/absent in cAMP are highlighted with black circles. **(f)** Activation of PKG Iβ by cGMP and cAMP, highlighting the difference in maximal activity levels in the presence of saturating amounts of cAMP *versus* cGMP. The phosphotransferase activity of PKG Iβ was monitored at various concentrations of cAMP (blue plot) and cGMP (red plot) by a spectrophotometric kinase assay. Here, 1 U is equal to a substrate conversion rate of 1 μmol per min. Each data point was measured at least in duplicate, and the data sets were fitted using a sigmoidal dose-response algorithm.

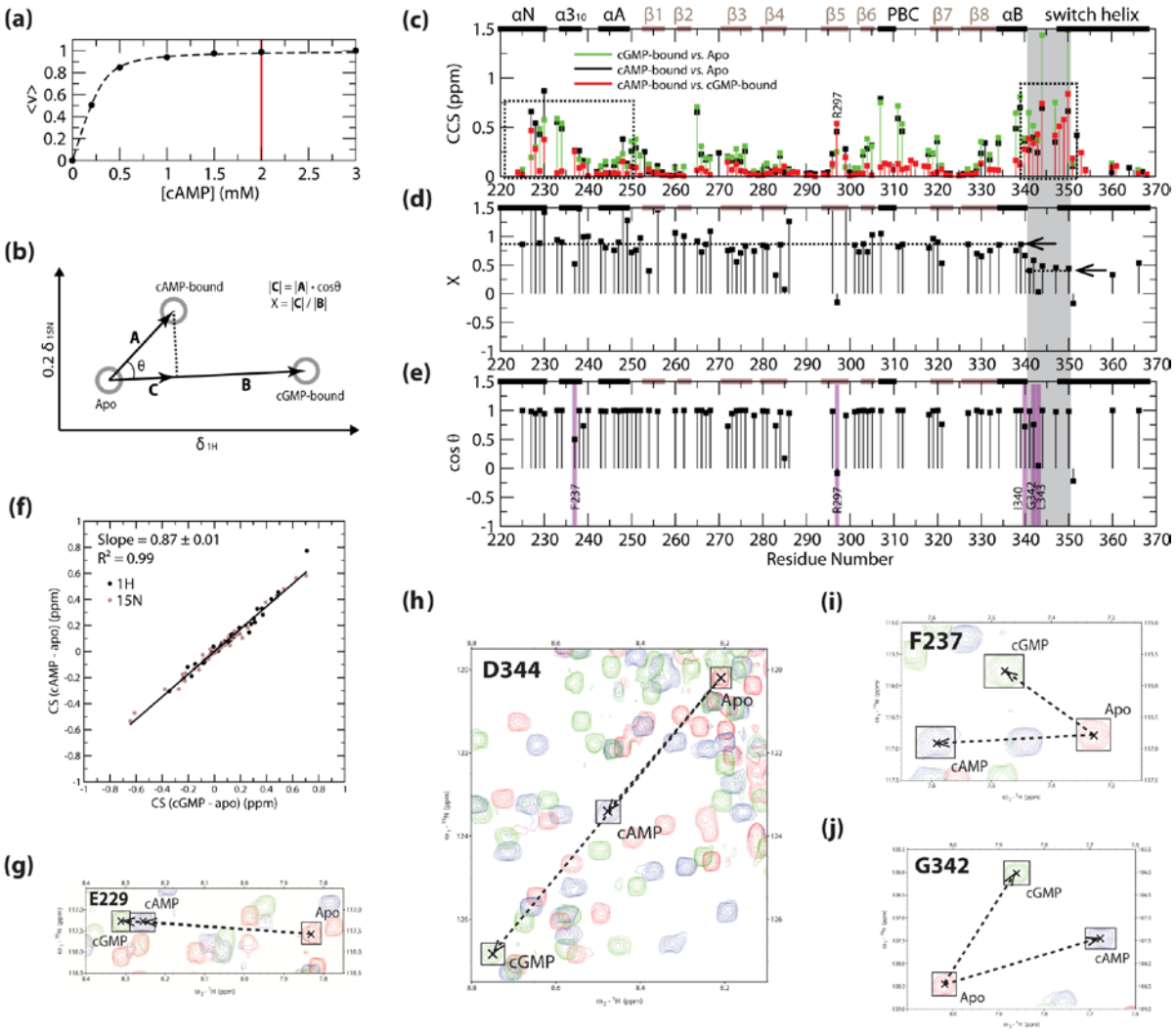

Figure 2: Effects of cAMP binding on PKG Iβ(219-369), as observed from CHESPA analysis. **(a)** Ligand binding isotherm for cAMP-bound PKG Iβ(219-369), as computed from the combined chemical shifts obtained by (^{1}H,^{15}N)-HSQC. The fitted curve is shown as a black dashed line, and the cAMP concentration selected to achieve saturated binding is indicated by a vertical red line. This concentration was selected based on the lack of further shift in HSQC peak positions during ligand titration, as determined by HSQC overlays. **(b)** Schematic illustration of the chemical shift projection analysis (CHESPA) calculation methodology, using (^{1}H,^{15}N)-HSQC chemical shifts as input.[44] The ^{1}H and ^{15}N chemical shift values for the reference (*i.e.* apo and cGMP-bound) and perturbation (*i.e.* cAMP-bound) states are plotted as gray circles, and key vectors and parameters used in the analysis are indicated. Positive values of X and cosθ indicate a shift in favor of activation in the perturbation state, while negative values indicate a shift in favor of inactivation. **(c-e)** Results of the CHESPA analysis of cAMP-bound PKG Iβ(219-369), with apo and cGMP-bound PKG Iβ(219-369) as reference states, as shown in panel (b). The secondary structure

elements from the apo-state X-ray structure are indicated across the top of each graph: black bars = α-helices; brown bars = β-strands. **(c)** Compounded chemical shifts of cAMP-bound (black points) and cGMP-bound (green points) PKG Iβ(219-369) from apo PKG Iβ(219-369), and of cAMP-bound from cGMP-bound PKG Iβ(219-369) (red points). Black dotted boxes indicate notable cAMP-*vs*.-cGMP chemical shift differences observed for residues outside the cGMP binding site. **(d,e)** Fractional shifts toward activation ("X"; panel d) and projection angle cosines ("cosθ"; panel e) achieved by cAMP binding, relative to apo and cGMP-bound PKG Iβ(219-369). The average fractional shift toward activation computed for the residue 340-350 region (average of X values from panel d), and the overall fractional shift toward activation computed for residues N-terminal to residue 340 (computed slope in panel f), are marked by arrows and black dotted lines. The residue 340-350 region is indicated in all plots by a gray highlight, and notable low-cosθ-value residues ("outliers," as defined in the Experimental Procedures section) are indicated by purple highlights. **(f)** Chemical shift correlation plot of ^{1}H (black points) and scaled ^{15}N (brown points) chemical shifts of cAMP-bound *versus* cGMP-bound PKG Iβ(219-369) (both relative to the apo state). The black line delineates the line of best fit as determined from linear regression, and the corresponding slope and correlation coefficient are indicated. **(g-j)** Expansions of the overlaid (^{1}H,^{15}N)-HSQC spectra of apo (red contours), cAMP-bound (blue contours) and cGMP-bound (green contours) PKG Iβ(219-369), illustrating the cAMP-associated perturbations of representative high-cosθ-value residues from the residue 340-350 region (panel h) and from N-terminal to residue 340 (panel g), as well as representative low-cosθ-value ("outlier") residues from the N3A (panel i) and residue 340-350 region (panel j).

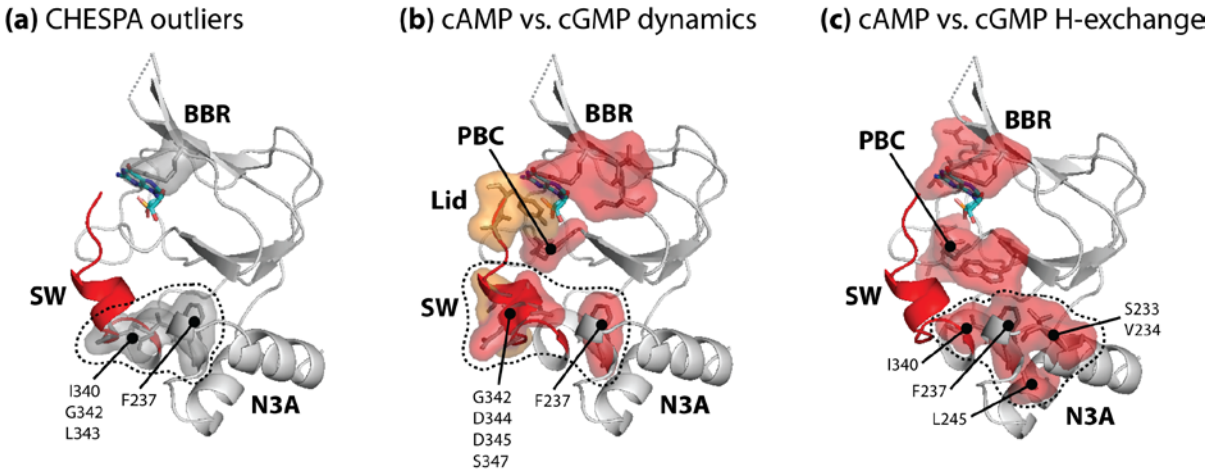

Figure 3: Three-dimensional maps of cAMP *vs.* cGMP differential dynamics. The maps shown illustrate residues exhibiting **(a)** low cosθ values from CHESPA ("outlier" residues), **(b)** differences in cAMP- *vs.* cGMP-bound dynamics, and **(c)** differences in cAMP- *versus* cGMP-bound hydrogen exchange rates, as identified in Figures 2 and 4-6. All residue sets are illustrated in the context of the cGMP-bound structure of PKG Iβ(219-369), with bound cGMP shown as colored sticks, and the following structural elements are indicated: the N3A region; the base binding region (BBR) and phosphate binding cassette (PBC) elements involved in cGMP binding; and the C-terminal switch helix region (SW) and cGMP-capping lid (both highlighted as red ribbon). The identified residues are shown as sticks and surfaces, and the surfaces in panels (b) and (c) follow the same color codes as the residue highlights in Figures 4-6. The cluster of non-binding-site residues identified in each case is delineated by a black dotted line, and the identities of the constituent residues are indicated.

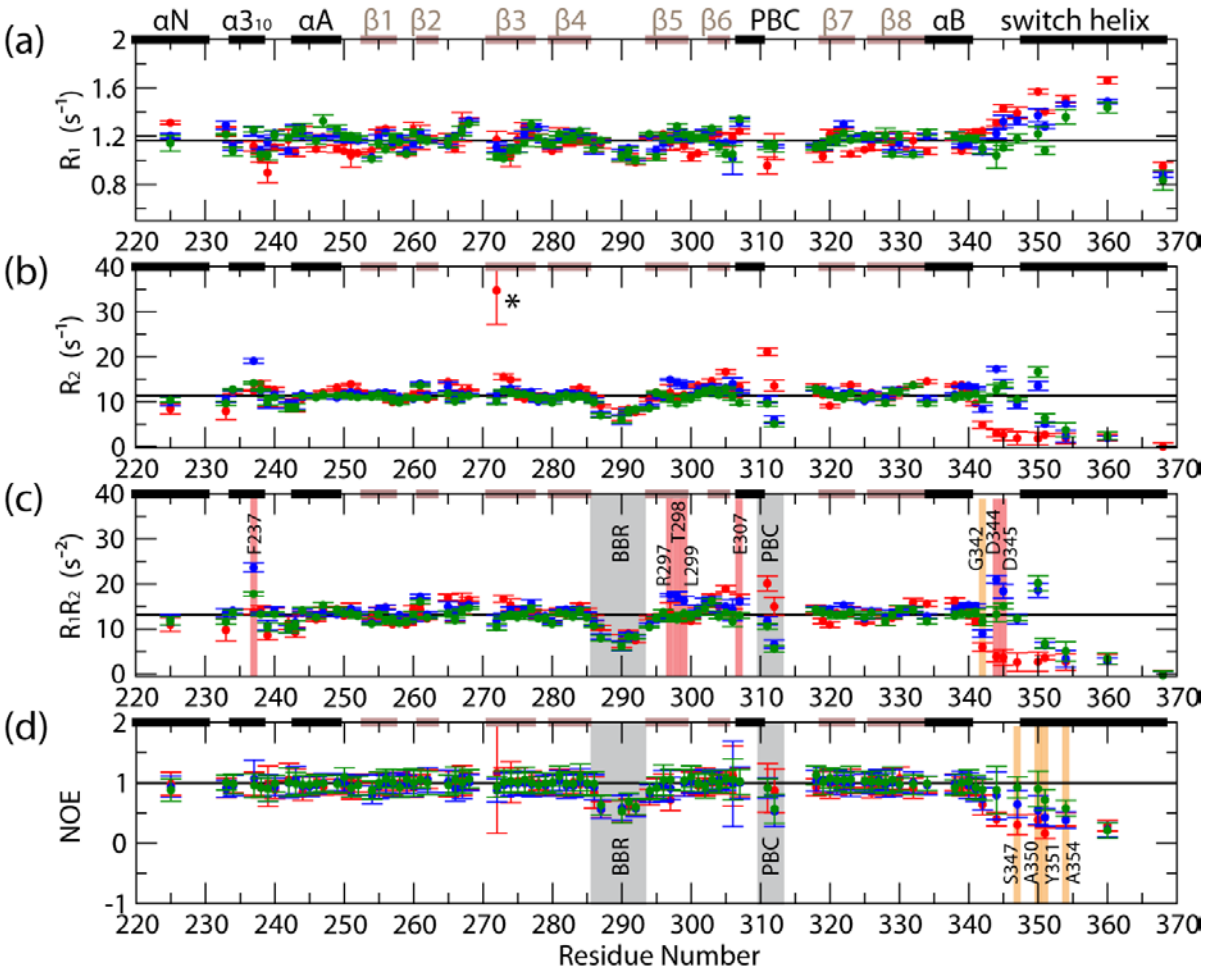

Figure 4: Backbone ^{15}N relaxation rates for apo (red points), cAMP-bound (blue points) and cGMP-bound (green points) PKG Iβ(219-369). **(a)** R_1 relaxation rates; **(b)** R_2 relaxation rates; **(c)** product of R_1 and R_2; **(d)** (^{1}H,^{15}N)-NOE values, computed as I_{sat}/I_{nonsat}. Black horizontal lines denote the average values computed for the inner β-strands of the β-barrel, which, due to the rigidity of this region of the protein, are assumed to represent the overall tumbling motion of the protein in solution. Residues exhibiting notable enhancements of ms-μs/ps-ns dynamics in cAMP- *versus* cGMP-bound PKG Iβ(219-369) are indicated by red/orange highlights, respectively, while other regions exhibiting notable dynamics are indicated by gray highlights, and an apo-state residue for which the R_2 rate could not be properly quantified (due to a poor signal-to-noise ratio) is marked with an asterisk. The secondary structure elements from the apo-state X-ray structure are indicated across the top of each graph (black bars = α-helices, brown bars = β-strands), and residues for which no data is shown are prolines or were not successfully assigned in the relaxation NMR spectra for one or more states.

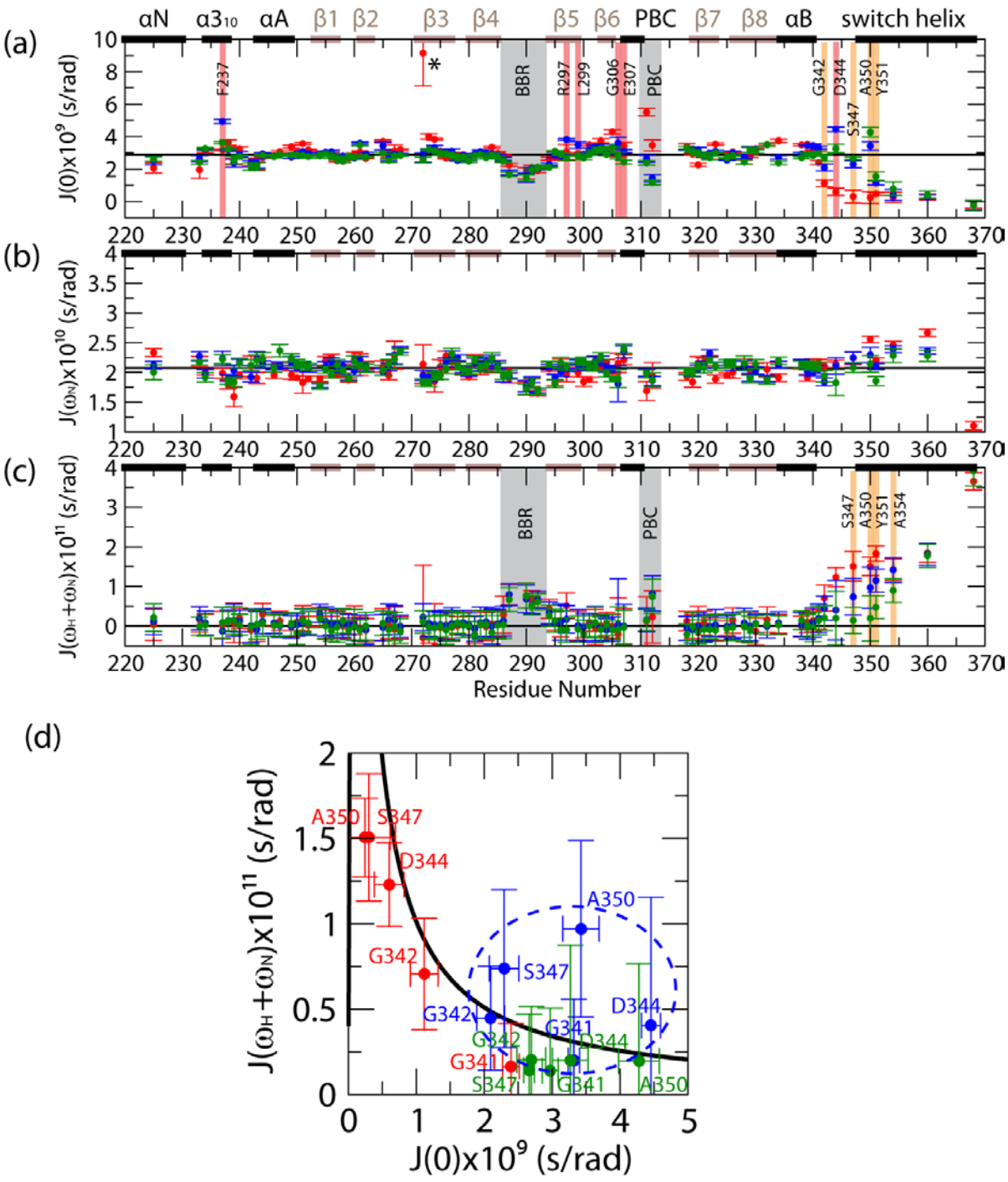

Figure 5: Reduced spectral densities (RSDs) for apo (red points), cAMP-bound (blue points) and cGMP-bound (green points) PKG Iβ(219-369). **(a)** J(0) values; **(b)** J(ω_N) values; **(c)** J(ω_H + ω_N) values. The RSDs were computed as described previously,[13,46,47] based on the relaxation rates reported in Figure 4. Black horizontal lines denote the average values

computed for the inner β-strands of the β-barrel, which, due to the rigidity of this region of the protein, are assumed to represent the overall tumbling motion of the protein in solution. All color codes illustrated are as in Figure 4. **(d)** Two-dimensional $J(\omega_H + \omega_N)$ *versus* $J(0)$ plots for residues 341-350. The values expected for a rigid rotor in the absence of ms-μs dynamics are indicated by the black plot, and the cAMP-bound state data are indicated by a blue, dashed oval.

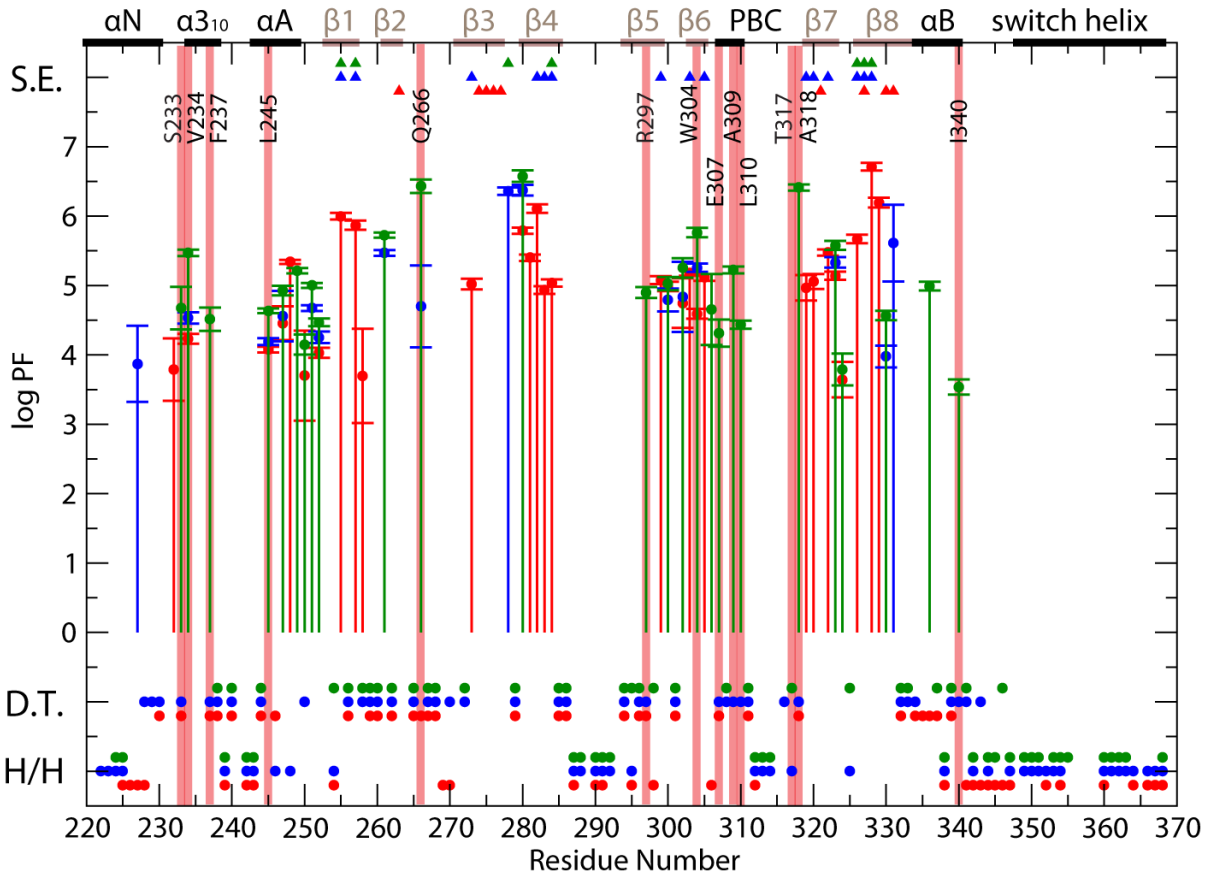

Figure 6: Backbone hydrogen exchange results for apo (red points), cAMP-bound (blue points) and cGMP-bound (green points) PKG Iβ(219-369). Residues are grouped into four types as follows: residues that undergo hydrogen exchange on the 10^0-10^2 ms time scale (as identified by H/H NMR) are denoted by "H/H"; other residues that undergo hydrogen exchange within the dead time of the H/D NMR experiment, and thus exhibit hydrogen exchange on a seconds-to-minutes time scale, are denoted by "D.T."; residues with hydrogen exchange rates quantifiable by H/D NMR are plotted as the logarithms of protection factors (PFs) computed from the H/D NMR data, whereby higher values indicate slower hydrogen exchange; and residues whose hydrogen exchange rates were too slow to be properly quantified are denoted by "S.E." and triangles. Residues exhibiting notable enhancements of hydrogen exchange rates in cAMP- *versus* cGMP-bound PKG Iβ(219-369) are indicated by red highlights, and the secondary structure elements from the apo-state X-ray structure are indicated across the top of each graph: black bars = α-helices; brown bars = β-strands. Residues for which no data is shown are prolines or were not successfully assigned in the NMR spectra.

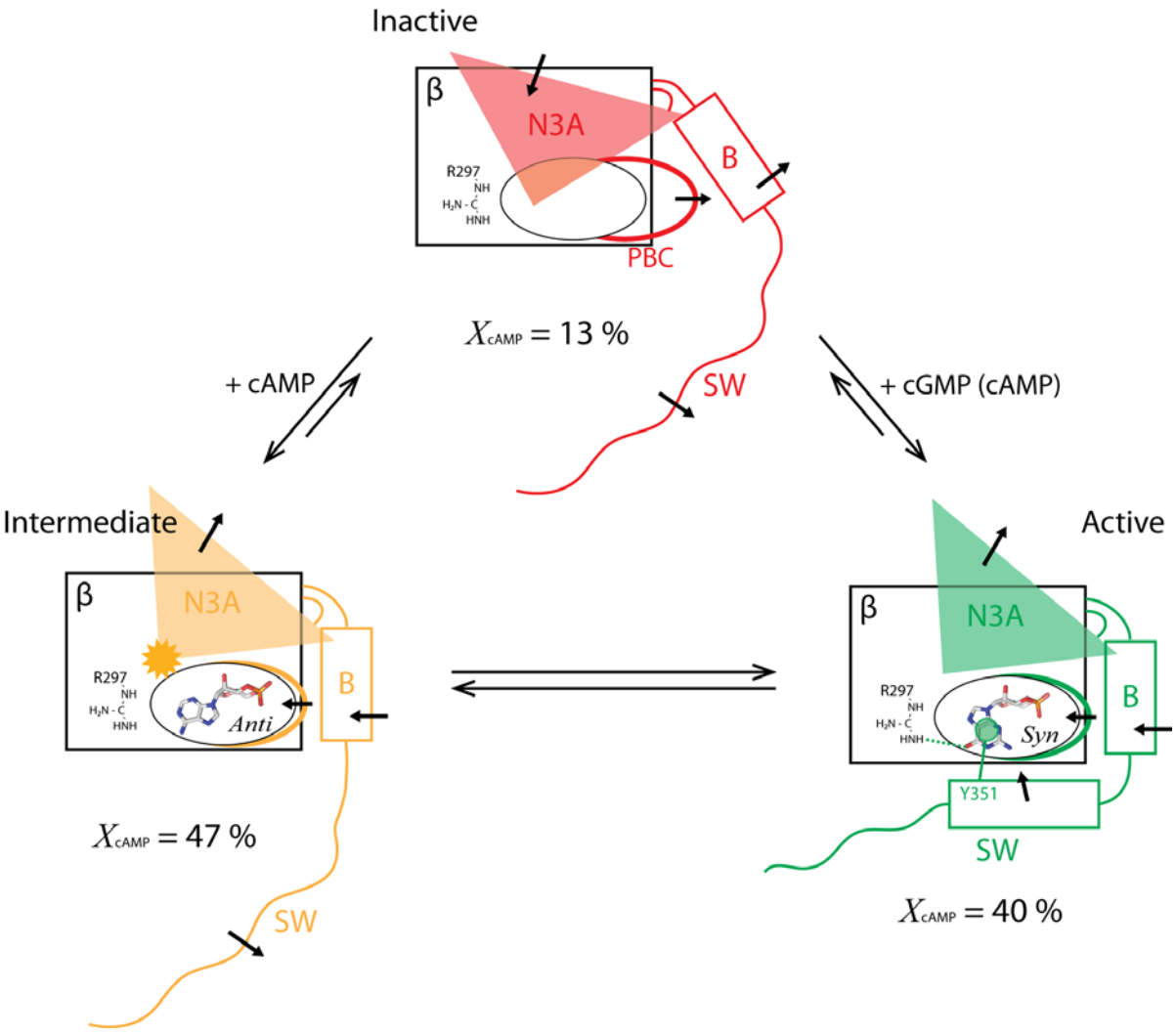

Figure 7: Schematic summary of the proposed CNB-B model for cAMP partial agonism, illustrating the "inactive" (red) and "active" (green) conformational states explored in the cGMP-modulated two-state activation equilibrium, and the additional "intermediate" state (orange) stabilized in the presence of cAMP. The structural shifts of the α-helical subdomain elements in each conformational state are indicated by black arrows, and the N3A perturbation unique to the intermediate state is indicated by an orange starburst. The bound ligands and their conformations are shown as labelled stick structures, and the key active-state cGMP interactions with residues R297 (cGMP-selective hydrogen bond) and Y351 (cGMP-capping lid interaction) are schematically indicated. The percentages given below each state identify the populations of each state in the presence of cAMP, as calculated from the average fractional shift toward activation computed for the residue 340-350 region (average of X values from Figure 2d), and the overall fractional shift toward activation computed for residues N-terminal to residue 340 (computed slope in Figure 2f).

4.9. References

[1] Francis, S.H., Busch, J.L., Corbin, J.D. and Sibley, D. (2010) cGMP-dependent protein kinases and cGMP phosphodiesterases in nitric oxide and cGMP action. *Pharmacol. Rev.* **62**: 525-563.
[2] Kim, J.J., Casteel, D.E., Huang, G., Kwon, T.H., Ren, R.K., Zwart, P., Headd, J.J., Brown, N.G., Chow, D.C., Palzkill, T. and Choel, K. (2011) Co-crystal structures of PKG Iβ (92–227) with cGMP and cAMP reveal the molecular details of cyclic-nucleotide binding. *PLoS ONE* **6**: e18413(1-15).
[3] Hofmann, F., Bernhard, D., Lukowski, R. and Weinmeister, P. (2009) In *Handbook of Experimental Pharmacology* (Rosenthal, W., ed) pp. 137-162, Springer-Verlag, Berlin.
[4] Osborne, B.W., Wu, J., McFarland, C.J., Nickl, C.K., Sankaran, B., Casteel, D.E., Woods, V.L., Kornev, A.P., Taylor, S.S. and Dostmann, W.R. (2011) Crystal structure of cGMP-dependent protein kinase reveals novel site of interchain communication. *Structure* **19**: 1317-1327.
[5] Francis, S.H. and Corbin, J.D. (1999) Cyclic nucleotide-dependent protein kinases: intracellular receptors for cAMP and cGMP action. *Crit. Rev. Clin. Lab. Sci.* **36**: 275-328.
[6] Blumenthal, D.K., Copps, J., Smith-Nguyen, E.V., Zhang, P., Heller, W.T. and Taylor, S.S. (2014) The roles of the RIIβ linker and N-terminal cyclic nucleotide-binding domain in determining the unique structures of the type IIβ protein kinase A: a small angle x-ray and neutron scattering study. *J. Biol. Chem.* **289**: 28505-28512.
[7] Cheng, C.Y., Yang, J., Taylor, S.S. and Blumenthal, D.K. (2009) Sensing domain dynamics in protein kinase A-I{alpha} complexes by solution X-ray scattering. *J. Biol. Chem.* **284**: 35916-35925.
[8] Kim, C., Cheng, C.Y., Saldanha, S.A. and Taylor, S.S. (2007) PKA-I holoenzyme structure reveals a mechanism for cAMP-dependent activation. *Cell* **130**: 1032-1043.
[9] Kim, J., Masterson, L.R., Cembran, A., Verardi, R., Shi, L., Gao, J., Taylor, S.S. and Veglia, G. (2015) Dysfunctional conformational dynamics of protein kinase A induced by a lethal mutant of phospholamban hinder phosphorylation. *Proc. Natl. Acad. Sci. USA* **112**: 3716-3721.
[10] Masterson, L.R., Yu, T., Shi, L., Wang, Y., Gustavsson, M., Mueller, M.M. and Veglia, G. (2011) cAMP-dependent protein kinase A selects the excited state of the membrane substrate phospholamban. *J. Mol. Biol.* **412**: 155-164.
[11] Srivastava, A.K., McDonald, L.R., Cembran, A., Kim, J., Masterson, L.R., McClendon, C.L., Taylor, S.S. and Veglia, G. (2014) Synchronous opening and closing motions are essential for cAMP-dependent protein kinase A signaling. *Structure* **22**: 1735-1743.
[12] Su, Y., Dostmann, W.R., Herberg, F.W., Durick, K., Xuong, N.H., Ten-Eyck, L., Taylor, S.S. and Varughese, K.I. (1995) Regulatory subunit of protein kinase A: structure of deletion mutant with cAMP binding domains. *Science* **269**: 807-813.
[13] Das, R., Chowdhury, S., Mazhab-Jafari, M.T., SilDas, S., Selvaratnam, R. and Melacini, G. (2009) Dynamically driven ligand selectivity in cyclic nucleotide binding domains. *J. Biol. Chem.* **284**: 23682-23696.
[14] De, S., Greenwood, A.I., Rogals, M.J., Kovrigin, E.L., Lu, K.P. and Nicholson, L.K. (2012) Complete thermodynamic and kinetic characterization of the isomer-specific

interaction between Pin1-WW domain and the amyloid precursor protein cytoplasmic tail phosphorylated at Thr668. *Biochemistry* **51**: 8583-8596.
[15] Greenwood, A.I., Kwon, J. and Nicholson, L.K. (2014) Isomerase-catalyzed binding of interleukin-1 receptor-associated kinase 1 to the EVH1 domain of vasodilator-stimulated phosphoprotein. *Biochemistry* **53**: 3593-3607.
[16] Jayaraman, B. and Nicholson, L.K. (2007) Thermodynamic dissection of the ezrin FERM/CERMAD interface. *Biochemistry* **46**: 12174-12189.
[17] Huang, G.Y., Kim, J.J., Reger, A.S., Lorenz, R., Moon, E.W., Zhao, C., Casteel, D.E., Bertinetti, D., VanSchouwen, B., Selvaratnam, R., Pflugrath, J.W., Sankaran, B., Melacini, G., Herberg, F.W. and Kim, C. (2014) Structural basis for cyclic-nucleotide selectivity and cGMP-selective activation of PKG I. *Structure* **22**: 116-124.
[18] Huang, G.Y., Gerlits, O.O., Blakeley, M.P., Sankaran, B., Kovalevsky, A.Y. and Kim, C. (2014) Neutron diffraction reveals hydrogen bonds critical for cGMP-selective activation: insights for cGMP-dependent protein kinase agonist design. *Biochemistry* **53**: 6725-6727.
[19] Beavo, J.A. and Brunton, L.L. (2002) Cyclic nucleotide research – still expanding after half a century. *Nat. Rev. Mol. Cell Biol.* **3**: 710-718.
[20] Börner, S., Schwede, F., Schlipp, A., Berisha, F., Calebiro, D., Lohse, M.J. and Nikolaev, V.O. (2011) FRET measurements of intracellular cAMP concentrations and cAMP analog permeability in intact cells. *Nat. Protoc.* **6**: 427-438.
[21] Wolter, S., Golombek, M. and Seifert, R. (2011) Differential activation of cAMP- and cGMP-dependent protein kinases by cyclic purine and pyrimidine nucleotides. *Biochem. Biophys. Res. Commun.* **415**: 563-566.
[22] Zaccolo, M., De Giorgi, F., Cho, C.Y., Feng, L., Knapp, T., Negulescu, P.A., Taylor, S.S., Tsien, R.Y. and Pozzan, T. (2000) A genetically encoded, fluorescent indicator for cyclic AMP in living cells. *Nat. Cell Biol.* **2**: 25-29.
[23] Zaccolo, M. and Pozzan, T. (2002) Discrete microdomains with high concentration of cAMP in stimulated rat neonatal cardiac myocytes. *Science* **295**: 1711-1715.
[24] Akimoto, M., Selvaratnam, R., McNicholl, E.T., Verma, G., Taylor, S.S. and Melacini, G. (2013) Signaling through dynamic linkers as revealed by PKA. *Proc. Natl. Acad. Sci. USA* **110**: 14231-14236.
[25] Akimoto, M., Zhang, Z., Boulton, S., Selvaratnam, R., VanSchouwen, B., Gloyd, M., Accili, E.A., Lange, O.F. and Melacini, G. (2014) A mechanism for the auto-inhibition of hyperpolarization-activated cyclic nucleotide-gated (HCN) channel opening and its relief by cAMP. *J. Biol. Chem.* **289**: 22205-22220.
[26] Das, R., Esposito, V., Abu-Abed, M., Anand, G.S., Taylor, S.S. and Melacini, G. (2007) cAMP activation of PKA defines an ancient signaling mechanism. *Proc. Natl. Acad. Sci. USA* **104**: 93-98.
[27] Das, R. and Melacini, G. (2007) A model for agonism and antagonism in an ancient and ubiquitous cAMP-binding domain. *J. Biol. Chem.* **282**: 581-593.
[28] VanSchouwen, B., Selvaratnam, R., Fogolari, F. and Melacini, G. (2011) Role of dynamics in the auto-inhibition and activation of the exchange protein directly activated by cyclic AMP (EPAC). *J. Biol. Chem.* **286**: 42655-42669.

[29] VanSchouwen, B., Akimoto, M., Sayadi, M., Fogolari, F. and Melacini, G. (2015) Role of dynamics in the auto-inhibition and activation of the hyperpolarization-activated cyclic-nucleotide-modulated (HCN) ion channels. *J. Biol. Chem.* **290**: 17642-17654.
[30] Boehr, D.D., Nussinov, R. and Wright, P.E. (2009) The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**: 789-796.
[31] Nussinov, R. and Tsai, C.J. (2014) Unraveling structural mechanisms of allosteric drug action. *Trends Pharmacol. Sci.* **35**: 256-264.
[32] Nussinov, R., Ma, B. and Tsai, C. (2014) Multiple conformational selection and induced fit events take place in allosteric propagation. *J. Biophys. Chem.* **186**: 22-30.
[33] Nussinov, R., Tsai, C.J. and Liu, J. (2014) Principles of allosteric interactions in cell signaling. *J. Am. Chem. Soc.* **136**: 17692-17701.
[34] Boulton, S., Akimoto, M., VanSchouwen, B., Moleschi, K., Selvaratnam, R. and Melacini, G. (2014) Tapping the translation potential of cAMP signalling: molecular basis for selectivity in cAMP agonism and antagonism as revealed by NMR. *Biochem. Soc. Trans.* **42**: 302-307.
[35] Alverdi, V., Mazon, H., Versluis, C., Hemrika, W., Esposito, G., van den Heuvel, R., Scholten, A. and Heck, A.J. (2008) cGMP-binding prepares PKG for substrate binding by disclosing the C-terminal domain. *J. Mol. Biol.* **375**: 1380-1393.
[36] Smith, J.A., Reed, R.B., Francis, S.H., Grimes, K. and Corbin, J.D. (2000) Distinguishing the roles of the two different cGMP-binding sites for modulating phosphorylation of exogenous substrate (heterophosphorylation) and autophosphorylation of cGMP-dependent protein kinase. *J. Biol. Chem.* **275**: 154-158.
[37] Wall, M.E., Francis, S.H., Corbin, J.D., Grimes, K., Richie-Jannetta, R., Kotera, J., Macdonald, B.A., Gibson, R.R. and Trewhella, J. (2003) Mechanisms associated with cGMP binding and activation of cGMP-dependent protein kinase. *Proc. Natl. Acad. Sci. USA* **100**: 2380-2385.
[38] Cook, P.F., Neville, M.E., Vrana, K.E., Hartl, F.T. and Roskoski, R. (1982) Adenosine cyclic 3',5'-monophosphate dependent protein kinase: kinetic mechanism for the bovine skeletal muscle catalytic subunit. *Biochemistry* **21**: 5794-5799.
[39] Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**: 277-293.
[40] Goddard, T.D. and Kneller, D.G. SPARKY – NMR Assignment and Integration Software. University of California, San Francisco, 2006.
[41] Bahrami, A., Assadi, A.H., Markley, J.L. and Eghbalnia, H.R. (2009) Probabilistic interaction network of evidence algorithm and its application to complete labeling of peak lists from protein NMR spectroscopy. *PLoS Comput. Biol.* **5**: e1000307(1-15).
[42] Boulton, S., Akimoto, M., Selvaratnam, R., Bashiri, A. and Melacini, G. (2014) A tool set to map allosteric networks through the NMR chemical shift covariance analysis. *Sci. Rep.* **4**: 7306.
[43] Selvaratnam, R., Chowdhury, S., VanSchouwen, B. and Melacini, G. (2011) Mapping allostery through the covariance analysis of NMR chemical shifts. *Proc. Natl. Acad. Sci. USA* **108**: 6133-6138.

[44] Selvaratnam, R., VanSchouwen, B., Fogolari, F., Mazhab-Jafari, M.T., Das, R. and Melacini, G. (2012) The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants. *Biophys. J.* **102**: 630-639.
[45] Selvaratnam, R., Mazhab-Jafari, M.T., Das, R. and Melacini, G. (2012) The auto-inhibitory role of the EPAC hinge helix as mapped by NMR. *PLoS ONE* **7**: e48707(1-9).
[46] Das, R., Mazhab-Jafari, M.T., Chowdhury, S., SilDas, S., Selvaratnam, R. and Melacini, G. (2008) Entropy-driven cAMP-dependent allosteric control of inhibitory interactions in exchange proteins directly activated by cAMP. *J. Biol. Chem.* **283**: 19691-19703.
[47] Mazhab-Jafari, M.T., Das, R., Fotheringham, S.A., SilDas, S., Chowdhury, S. and Melacini, G. (2007) Understanding cAMP-dependent allostery by NMR spectroscopy: comparative analysis of the EPAC1 cAMP-binding domain in its apo and cAMP-bound states. *J. Am. Chem. Soc.* **129**: 14482-14492.
[48] Mulder, F.A.A., Schipper, D., Bott, R. and Boelens, R. (1999) Altered flexibility in the substrate-binding site of related native and engineered high-alkaline Bacillus subtilisins. *J. Mol. Biol.* **292**: 111-123.
[49] Rehmann, H., Das, J., Knipscheer, P., Wittinghofer, A. and Bos, J.L. (2006) Structure of the cyclic-AMP-responsive exchange factor Epac2 in its auto-inhibited state. *Nature* **439**: 625-628.
[50] Rehmann, H., Arias-Palomo, E., Hadders, M.A., Schwede, F., Llorca, O. and Bos, J.L. (2008) Structure of Epac2 in complex with a cyclic AMP analogue and RAP1B. *Nature* **455**: 124-127.
[51] White, R.E., Kryman, J.P., El-Mowafy, A.M., Han, G. and Carrier, G.O. (2000) cAMP-dependent vasodilators cross-activate the cGMP-dependent protein kinase to stimulate BK_{Ca} channel activity in coronary artery smooth muscle cells. *Circ. Res.* **86**: 897-905.

CHAPTER 5

CONCLUDING REMARKS AND FUTURE OUTLOOK

5.1. Summary of Thesis and Emerging Common Themes

In this thesis, we examined cyclic-nucleotide-modulated dynamics within the cyclic-nucleotide receptor proteins EPAC (Chapter 2), HCN (Chapter 3) and PKG (Chapter 4). We found that in all three cyclic-nucleotide receptor proteins, dynamics are critical for allosteric control of protein activation and/or autoinhibition. However, the findings for each system revealed a different role of dynamics in allostery.

Our work on EPAC has highlighted dynamics as a key modulator of the entropic component of the free-energy landscape for cAMP-dependent allostery. In particular, the entropic contributions from dynamics lead to a low-free-energy pathway for EPAC activation, and play a role in fine-tuning the stability of the auto-inhibited state, thus optimally sensitizing it to cAMP while avoiding constitutive activation. Meanwhile, our work on HCN has highlighted dynamics as a key modulator of the enthalpic component of the free-energy landscape for cAMP-dependent allostery. Specifically, it was determined that an increased flexibility of the tetramerization domain (TD) N-terminal to the CBD in the monomeric state alleviates steric clashes with the TD that the apo CBD structure would otherwise impose in the tetrameric state. The binding of cAMP confines the CBD to conformations that will not impose such steric clashes, thereby promoting tetramerization. Finally, our work on PKG has highlighted dynamics as a key determinant of cGMP-*vs.*-cAMP selectivity that arises due to cAMP exerting only a partial-agonist activity against PKG. In particular, while cGMP-associated activation is readily modeled by a two-state conformational equilibrium within the C-terminal cGMP-binding domain (CNB-B),

accounting for the partial agonism of cAMP requires the dynamic sampling of a third CNB-B conformational state in which kinase inhibition is at least partially still effective.

While our work has shed light on key attributes of cyclic-nucleotide receptor protein dynamics, and the role of such dynamics in protein allostery, more work remains to be done in order to achieve a more complete understanding of CBD function and to fully exploit the potential of CBDs as therapeutic targets. In this respect, we outline in the following sections several high-priority emerging questions and perspectives for HCN and PKG, along with some key preliminary findings that we have compiled as of this thesis.

5.2. Emerging Questions and Perspectives for HCN

5.2.1. Open Questions on HCN Auto-Inhibition

In Chapter 3, we found that the HCN4 tetramerization domain (TD) becomes more flexible in the monomeric states, thus removing steric clashes with the TD that the apo-state CBD structure would otherwise impose. Furthermore, we found that cAMP binding confines the CBD conformational ensemble to a tetramer-compatible state, thereby priming the intracellular region (IR) for tetramerization, and providing an explanation for how cAMP controls IR tetramerization. However, the potential structural consequences of the previously observed steric clashes,[1] which arise within the IR tetramer when the CBD is in an inactive-state conformation, were not fully explored. In particular, while tetramer dissociation would release such steric clashes, we wondered if tetramer dissociation is the only possible mechanism for steric clash release.

We hypothesized that the steric clashes of the inactive-state IR tetramer may be released not only through dissociation, but alternatively through tetramer distortion and/or an

induced shift of the CBD toward its active-state conformation. To test this hypothesis, we have performed molecular dynamics (MD) simulations starting from several "hybrid" tetramer structures of the HCN4 IR construct examined in Chapter 3. Each of these hybrid structures consists of an HCN4 IR tetramer structure in which the CBD regions were substituted with a CBD conformation obtained from one of the structures in the previously-solved NMR-based inactive-state CBD ensemble,[1] and were thus designed to simulate a situation in which an inactive-state CBD conformation is placed into the context of an HCN4 IR tetramer. The simulation trajectories for each tetramer have so far been analyzed through computation of root-mean-square deviations (RMSDs) from the active-state structure (Figure 1), as well as distances between the centers-of-mass for selected segments of the four constituent protomers in each tetramer (Figures 2 and 3), and the results of these analyses are discussed below.

Notably, for a sub-set of inactive-state CBD conformations, the IR tetramer appears to induce a shift toward the active-state CBD structure, in agreement with our previous observation that the conformational ensemble accessible to the CBD becomes confined in the tetramer due to steric constraints imposed by the tetramer (Chapter 3). For example, the NMR #2 and NMR #10 hybrid tetramers demonstrated a shift of their protomer TD and CBD structural arrangements closer to the active-state conformation (Figure 1), while maintaining the tetrameric arrangement of their TD regions (Figure 2). Similarly, the NMR #9 hybrid tetramer demonstrated a shift of its protomer TD and CBD structural arrangements closer to the active-state conformation (Figure 1). However, the NMR #9 tetramer also demonstrated greater fluctuation from the active-state tetrameric arrangement, in the form of a diagonal inter-protomer "breathing" motion within the TD and CBD regions (Figures 2 and 3) – a

motion also observed to a lesser degree in the NMR #10 tetramer CBD region (Figure 3) – that appears analogous to a tetramer breathing motion noted previously for HCN2.[2] This result suggests that in addition to an inactive-to-active CBD conformational shift, tolerance of the steric clashes from some inactive-state CBD conformations requires amplification of an intrinsic breathing motion within the tetramer.

In contrast, the NMR #6 and NMR #7 hybrid tetramers both failed to undergo shifts closer to the active-state conformation (Figure 1). For example, the NMR #7 tetramer demonstrated a visible destabilization of the 1-4 and 2-3 inter-protomer interfaces within its TD region (Figure 2h), with its CBDs seemingly shifting toward an arrangement with dimer-of-dimers symmetry (Figure 3h). Indeed, visual examination of structures from the NMR #7 tetramer simulation revealed a dissociation of the tetramer's CBD region into dimers, as well as a visible distortion of the TD region from its active-state tetramer arrangement – including a large distortion in the αC'-αD' segment (Chapter 3 Figure 1b) that was expected to ultimately favor dissociation of the tetramer. These results are consistent with the notion that IR dimers may serve as intermediates in the IR tetramerization equilibrium, as suggested by previous functional mutagenesis[3] and size-exclusion chromatography[4] experiments. In this context, the observed structural shifts of the NMR #6 and NMR #7 tetramers may represent an early stage of tetramer dissociation *via* dimers.

Overall, the above results suggest that the steric clashes within the hybrid tetramers can be removed either by preserving the tetrameric structure and shifting the CBD toward its active-state conformation, or by preserving the inactive-state conformation of the CBD and distorting the tetramer. Such distortions resemble an amplification of intrinsic breathing motions within the tetramer, and lead to the formation of a dimer-of-dimers structure that

may serve as a precursor to tetramer dissociation. Therefore, the inactive-state CBD imposes a free-energy barrier to IR tetramer assembly that attenuates HCN activation in the absence of cAMP, and is expected to be alleviated by cAMP-associated confinement of the CBD to more active-like conformations (as discussed in Chapter 3).

Meanwhile, the observed appearance of a dimer-of-dimers symmetry in some of the hybrid-tetramer simulations prompted us to more closely examine the structural dynamics of the dimer state, and any potential contribution of such dynamics to HCN4 allostery. Thus, we are now performing MD simulations starting from both active-state and "hybrid" dimer structures of the HCN4 IR construct, with the goal of assessing dimer dynamics.

5.2.2. cAMP-Modulated HCN Allostery in the Context of Full-Length HCN: From Tetramerization to Ion Channel Modulation

Altogether, our HCN work thus far (Chapter 3 and Section 5.2.1) has revealed key attributes of HCN TD and CBD structural dynamics, and how such dynamics confer cAMP-modulated regulation of IR tetramerization *via* TD self-association. Still, since our HCN construct lacks the ion-conducting trans-membrane (TM) region (Chapter 3 Figure 1), the work does not address the question of how IR tetramerization in turn influences ion channel gating. However, even though our HCN construct lacks the TM region, previous electrophysiology-based functional assays performed on HCN mutants (Figure 4), together with the close proximity of the αA'-αB' region to the TM region (Chapter 3 Figure 1), provide an explanation for how IR tetramerization influences ion channel gating in intact HCN.

In particular, alteration of the HCN tail region N-terminal to the TM region, or deletion of the HCN tail region C-terminal to the IR, did not significantly affect ion channel

gating, thus highlighting the TM and IR regions as the essential functional unit of HCN.[5,6] Meanwhile, deletion of the αA-αC segment of the CBD (while leaving the TD intact) resulted in a constitutive up-regulation of ion channel opening comparable to that resulting from cAMP binding, suggesting that the inactive-state CBD exerts an autoinhibitory effect on ion channel opening (Figure 4b),[5] in agreement with the findings of our HCN work (Chapter 3). Furthermore, deletion of the αD'-CBD segment of the IR – which would disrupt the TD tetramer by eliminating interactions with most of the αC'-αD' region (Chapter 3 Figure 1),[7,8] and eliminate any interactions with the CBDs – resulted in a constitutively enhanced inhibition of ion channel opening,[9,10] suggesting that TD tetramer disruption favors ion channel autoinhibition, and that interactions with the αC'-αD' region or CBDs may not be necessary for autoinhibition (Figure 4c). In addition, deletion of the entire IR resulted in constitutive ion channel up-regulation (similar to the CBD truncation mutant), suggesting that the TD is not strictly necessary for ion channel opening (Figure 4d).[5,6] Finally, point mutations[11] have suggested that the αA' helix N-terminal segment forms interactions with a TM-region cytoplasmic loop (*i.e.* the S4-S5 linker) which is known to be important in normal hyperpolarization-dependent ion channel gating,[12] and that disruption of these interactions hinders channel closure (Figure 4e).

The pattern of ion channel activities exhibited by these mutants, together with the close proximity of the αA'-αB' region to the TM region (Chapter 3 Figure 1), are consistent with a scenario in which the non-tetramerized αA'-αB' region exerts an autoinhibitory effect on ion channel opening *via* an interface with the TM region, and formation of the TD tetramer attenuates this autoinhibitory interface to promote ion channel opening – a scenario analogous to the competing autoinhibition- and activation-favoring interfaces formed by the

flexible N-terminal linker region of PKA (as noted in Section 1.3). Indeed, the greater TD dynamics in the non-tetramerized IR that were noted in our HCN work (Chapter 3) would liberate the αA'-αB' region for formation of the proposed αA'-αB'/TM interface, and may thus signify that similar to the N-terminal linker of PKA, the αA'-αB' region of HCN undergoes a dynamic exchange between autoinhibition- and activation-favoring structural arrangements. Furthermore, the constitutively enhanced inhibition exhibited by the αD'-CBD deletion mutant would be explained by promotion of the αA'-αB'/TM interface due to a disruption of the TD tetramer that concurrently preserves the αA'-αB' region, while the constitutive up-regulation exhibited by the full-IR deletion mutant would be explained by a disruption of the αA'-αB'/TM interface due to removal of the αA'-αB' region. Finally, C-terminal truncation of the isolated TD segment was found to reduce its self-association propensity (Chapter 3 unpublished data), suggesting that the constitutively enhanced inhibition exhibited by the αD'-CBD deletion mutant is likely due to interaction of the truncated TD regions with the TM region, rather than with one another.

These findings suggest that the TD and CBD structural dynamics described in our HCN work (Chapter 3 and Section 5.2.1) form part of a multi-layer allosteric mechanism in which CBD conformational shifts modulate ion channel gating by tuning the propensity of the activation-favoring TD tetramer interface relative to an autoinhibitory αA'-αB'/TM interface that is in competition with the TD tetramer (Figure 5). Thus, when the CBD is in a tetramer-incompatible conformation, the autoinhibitory αA'-αB'/TM interface is favoured over the TD tetramer interface, hindering ion channel opening (Figure 5). However, when the CBD adopts a tetramer-compatible conformation, the TD tetramer interface is favored, thereby attenuating the autoinhibitory interface and promoting ion channel opening (Figure

5). Therefore, by stabilizing the CBD in a tetramer-compatible conformational state, cAMP binding enhances the overall propensity for preference of the TD tetramer interface over the autoinhibitory interface, thus promoting ion channel opening (Figure 5), and providing a more complete explanation for how cAMP controls HCN ion channel gating.

5.3. Emerging Questions and Perspectives for PKG

5.3.1. Question on PKG Partial Agonism by cAMP

In Chapter 4, we found that the cGMP-*vs.*-cAMP selectivity of PKG relies on a reduced degree of activation elicited by cAMP *vs.* cGMP, which arises due to differences in the structural dynamics of the cAMP- *vs.* cGMP-bound forms of the CNB-B domain of PKG. Specifically, while activation by cGMP is well explained by a two-state conformational equilibrium, the partial agonism of cAMP arises from sampling of a third conformational state in which the lid region is partially disengaged, and kinase inhibition is at least partially still effective. However, we were left with the question of how differing interactions with the cyclic nucleotide base moiety of cAMP *vs.* cGMP (Figure 6) may contribute to the observed partial agonism of cAMP.

In order to dissect the role of the cGMP-specific base substituents (Figure 6) in the observed cGMP selectivity, we have performed CHESPA analyses of the PKG Iβ(219-369) construct bound to modified cyclic nucleotides in which one or both of the cGMP-specific base substituents have been deleted (*i.e.* cIMP, 2-NH_2-cPuMP and cPuMP, which define a double-mutant cycle as illustrated in Figure 6), following a protocol similar to that used in Chapter 4. In addition, we performed transfer-NOESY analyses (Chapter 1 Section 1.6.5) on the cyclic-nucleotide-bound PKG Iβ(219-369) constructs to assess whether the base

substituent deletions produce a change in the *syn*/*anti* conformational propensity of the bound cyclic nucleotide, as was noted previously for bound cAMP.[13,14]

Notably, the deletions of individual base substituents produced differing effects on the degree of activation depending on which substituent was affected. Specifically, deletion of the 2-NH_2 substituent (*i.e.* cIMP) produced very little overall change in the observed degree of activation (Figure 7a-b). Meanwhile, deletion of the 6-*oxo* substituent (*i.e.* 2-NH_2-cPuMP) produced a visibly greater effect on the observed degree of activation (Figure 7c-d), with perturbations comparable to those observed when the 2-NH_2 and 6-*oxo* substituents are both deleted (*i.e.* cPuMP; Figure 7e-f). Furthermore, the results observed for 2-NH_2-cPuMP-bound and cPuMP-bound PKG Iβ(219-369) both demonstrated perturbations comparable to those observed for cAMP-bound PKG Iβ(219-369) (compare Figure 7c-f with Chapter 4 Figure 2b-c). Therefore, proper interaction of the cyclic nucleotide binding pocket with the 6-*oxo* substituent of cGMP (*i.e.* hydrogen-bonding with residue R297; Chapter 4 Figure 1e) is critical to the observed cGMP-*vs.*-cAMP selectivity of the CNB-B domain of PKG.

Interestingly, none of the base substituent deletions seemed to produce the *syn*-to-*anti* conformational transition that was noted previously for bound cAMP.[13,14] This result suggests that the full *syn*-to-*anti* conformational transition of cAMP requires the presence of the 6-NH_2 substituent of cAMP, and may thus involve clashes between the 6-NH_2 substituent and residue R297, since the 6-NH_2 substituent is bulkier and has opposite hydrogen-bond donor/acceptor character compared to the 6-*oxo* substituent of cGMP (Figure 6 and Chapter 4 Figure 1e). Furthermore, this result suggests that a full *syn*-to-*anti* conformational transition may not be necessary for partial agonism (Chapter 4).

5.3.2. Question on PKG Inhibition by Cyclic Nucleotide Analogues

Another question that has arisen for PKG pertains to its response to phosphorothioate analogues of the canonical cGMP ligand (Figure 8a). Specifically, a preliminary CHESPA analysis of R_p-cGMPS-bound PKG Iβ(219-369) revealed that R_p-cGMPS exerts a partial-agonist activity on the CNB-B domain of PKG (Figure 8b-c). In contrast, the corresponding analogue of cAMP (*i.e.* R_p-cAMPS) was previously found to exert a reverse-agonist activity on the CNB domains of both PKA and EPAC, shifting the inactive/active equilibria of these domains closer to their inactive states.[15,16,17,18] Thus, we posed the question of why R_p-phosphorothioate analogues of the canonical cyclic-nucleotide ligands exert such differing activities toward PKG *versus* toward the structurally homologous PKA and EPAC.

In the case of PKA and EPAC, it was determined that the steric bulk of the R_p-phosphorothioate moiety of R_p-cAMPS stabilized the CNB domain phosphate-binding cassette (PBC) in its inactive-state arrangement, thus stabilizing an inactive-state steric interaction between the PBC and the α-helix C-terminal to the CNB domain β-barrel (a phenomenon referred to as the "hydrophobic hinge") that inhibits inactive-to-active transition of the CNB domain.[15,16,17,18,19,20] Thus, we hypothesized that the partial-agonist response of the PKG CNB-B domain to R_p-cGMPS may arise as a result of other activation-favoring phenomena that partially override the inhibitory influence of the hydrophobic hinge. The active-state capping interaction of switch-helix residue Y351 with the bound cyclic nucleotide (Chapter 4 Figure 1e) emerged as one possibility for such an activation-favoring phenomenon, due to the apparent importance of this capping interaction in PKG activation as noted in Chapter 4. However, it was previously determined that movement of the α-helix C-terminal to the β-barrel would be necessary for the switch-helix region to assume a suitable

position for establishment of the Y351 capping interaction,[13] suggesting that the PBC itself would need to allow for partial attenuation of the hydrophobic hinge. Thus, another possible activation-favoring phenomenon contributing to partial agonism may be the presence of PBC flexibility that allows the PBC to accommodate the steric bulk of R_p-cGMPS, thereby attenuating the steric interaction of the hydrophobic hinge. Indeed, a sequence alignment performed for the CNB domains revealed that in PKA and EPAC, the PBC C-terminal segment (where the R_p-phosphorothioate moiety would bind)[15,16] contains a proline residue,[20,21,22,23,24] while in PKG this proline residue is absent[13] (Figure 8d). Furthermore, cyclic nucleotide binding to PKA and EPAC quenched the dynamics of the PBC,[17,18] while in PKG the PBC retained residual dynamics even after cyclic nucleotide binding (Chapter 4 Figures 3c,d and 4a,c).

To test our hypothesis, we performed CHESPA analyses of the R_p-cGMPS-bound forms of Y351A and V315P point-mutants of PKG Iβ(219-369) (Figure 9). These mutants were designed to assess the effects on the partial-agonist response to R_p-cGMPS resulting from disengagement of the Y351 capping interaction (Chapter 4 Figure 1e), and PBC rigidification *via* insertion of a proline residue analogous to that found in PKA and EPAC (Figure 8d), respectively. Analysis of the R_p-cGMPS-bound Y351A mutant revealed a reduced overall degree of activation compared to R_p-cGMPS-bound wild-type PKG Iβ(219-369) (Figure 9a-b), suggesting that as hypothesized, the Y351 capping interaction may play a role in promoting the observed partial-agonist response to R_p-cGMPS. However, the R_p-cGMPS-bound Y351A mutant did not exhibit a shift from partial agonism to reverse agonism, as indicated by the retention of positive fractional activation values in the CHESPA results (Figure 9a-b). Therefore, the Y351 capping interaction alone does not fully account

for the partial-agonist response to R_p-cGMPS. Indeed, as with the R_p-cGMPS-bound Y351A mutant, analysis of the R_p-cGMPS-bound V315P mutant revealed a reduced overall degree of activation compared to R_p-cGMPS-bound wild-type PKG Iβ(219-369) (Figure 9c-d), suggesting that as hypothesized, a greater PBC flexibility in PKG also plays a role in promoting the observed partial-agonist response to R_p-cGMPS. Together, these results suggest that the partial-agonist response of the PKG CNB-B domain to R_p-cGMPS arises due to a partial override of R_p-phosphorothioate-associated stabilization of the hydrophobic hinge through a combination of PBC flexibility that allows the PBC to accommodate the steric bulk of R_p-cGMPS, and stabilization of the active state by the capping interaction of residue Y351 with the bound R_p-cGMPS.

5.3.3. Question on the Role of Dynamics in PKG Allostery

As a further question for PKG, we wondered what other roles dynamics might play in the allosteric thermodynamic cycle of cGMP-associated PKG activation. Indeed, as was the case for EPAC (Chapter 2), no residue-resolution information is currently available on the intermediate states ("metastates") for the induced-fit and conformational-selection activation pathways, whose dynamics may be pivotal determinants of PKG allostery. We hypothesize that dynamics of the states and metastates along the thermodynamic cycle may serve as key entropic determinants of PKG allostery, as was found to be the case for EPAC (Chapter 2). As a first step toward testing this hypothesis, we will perform and analyze molecular dynamics (MD) simulations of the CNB-B domain of PKG, starting from the four key states (*i.e.* apo/inactive, apo/active, holo/inactive and holo/active) along the thermodynamic cycle

for cGMP-associated activation, and following a protocol analogous to that implemented for EPAC (Chapter 2).

5.4. Concluding Remarks

The analysis of CBDs through a combination of multiple computational and NMR-based methods has significantly advanced our understanding of the role that dynamics play in allostery and in turn in the function of cyclic-nucleotide-dependent proteins. This work has revealed that, although eukaryotic CBDs share a structurally conserved fold, they exhibit a wide range of dynamic profiles. Such differences provide a starting point for the development of ligands that are selective for specific eukaryotic CBDs, as is needed to fully exploit the therapeutic potential of CBD targeting. Hence, we envision that the methods outlined in this thesis will be extended to map how drug-leads selected through screening perturb the multiple states of the allosteric cycle for the cyclic-nucleotide-dependent activation of cyclic-nucleotide-dependent proteins.

5.5. Figures

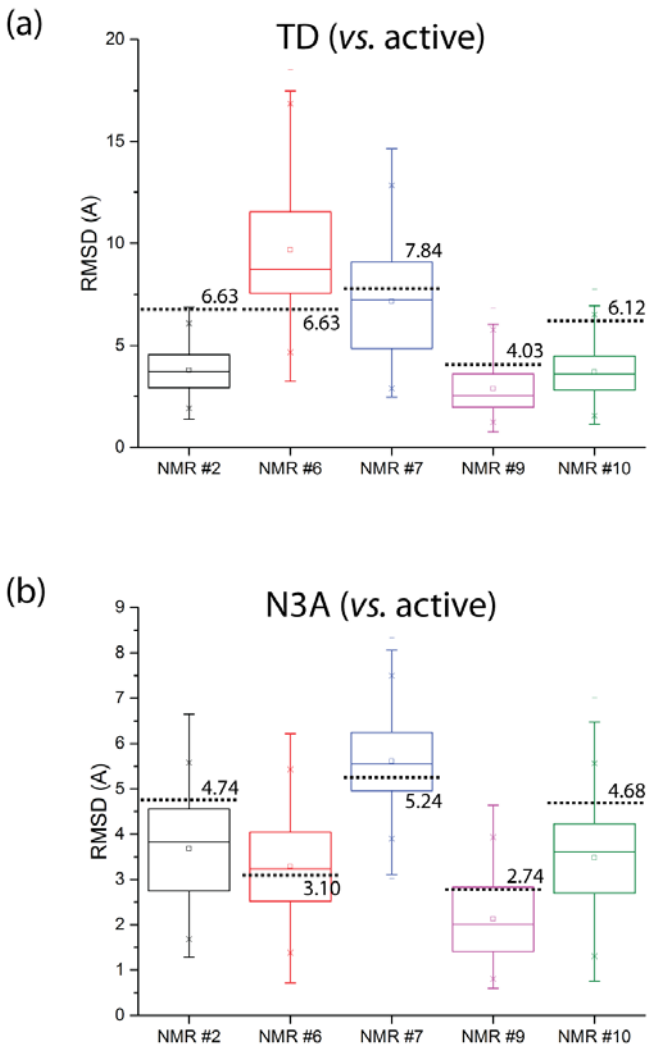

Figure 1: (a) Boxplots of backbone root-mean-square deviations (RMSDs) of the TD region from its position in the active-state structure (represented by the "3OTF" structure utilized in Chapter 3), as observed for each of the simulated "hybrid" tetramers. **(b)** Boxplots of backbone RMSDs of the N3A region from its configuration in the active-state structure, as observed for each of the simulated "hybrid" tetramers. The boxplots were constructed using Origin 9.1 graphing software (OriginLab Corporation) based on the RMSD data for all four protomers of each "hybrid" tetramer structure, and all RMSDs were computed with the structures overlaid at the β-core (as explained in Chapter 3). For reference, the corresponding RMSDs between the respective initial "hybrid" structures and the active-state structure are indicated by dotted horizontal lines.

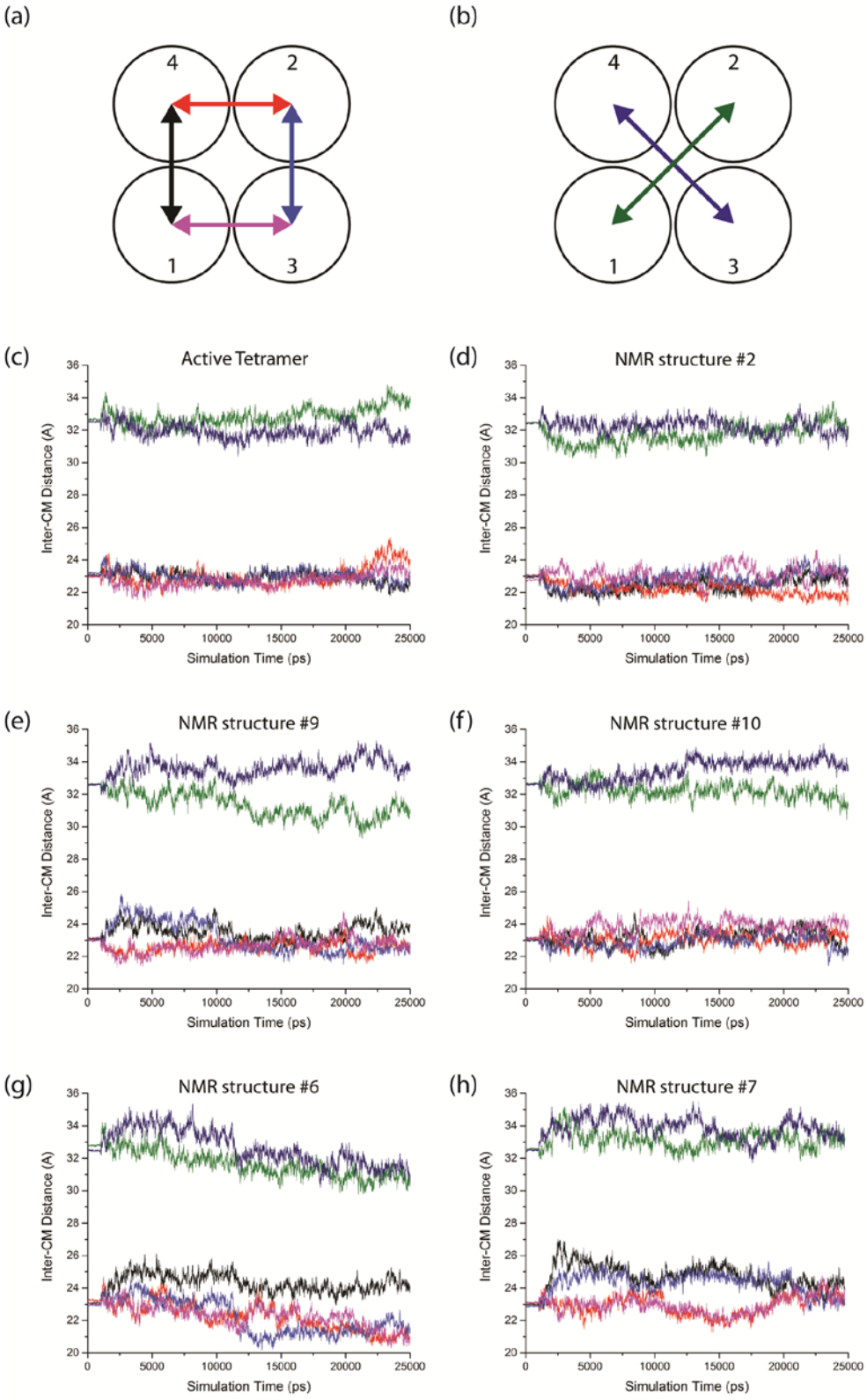
(a)
4
2
1
3
(b)
4
2
1
3
(c)
Active Tetramer
Inter-CM Distance (A)
Simulation Time (ps)
(d)
NMR structure #2
Inter-CM Distance (A)
Simulation Time (ps)
(e)
NMR structure #9
Inter-CM Distance (A)
Simulation Time (ps)
(f)
NMR structure #10
Inter-CM Distance (A)
Simulation Time (ps)
(g)
NMR structure #6
Inter-CM Distance (A)
Simulation Time (ps)
(h)
NMR structure #7
Inter-CM Distance (A)
Simulation Time (ps)

Figure 2: Inter-center-of-mass distance results for the TD portion of the tetramer, as observed for each of the simulated "hybrid" tetramers. **(a,b)** Outline of the inter-center-of-mass distances (indicated as arrows) between protomers that are situated **(a)** adjacent and **(b)** opposite one another in the tetramer. Individual protomers are represented as numbered circles, and the colour codes shown for the distances correspond to the colour codes used in the plots. **(c-h)** Inter-center-of-mass distance trajectories computed for the distances outlined in panels a-b, with colour coding as in panels a-b: **(c)** cAMP-bound active-state tetramer (used here as a control); **(d)** NMR structure #2 hybrid tetramer; **(e)** NMR structure #9 hybrid tetramer; **(f)** NMR structure #10 hybrid tetramer; **(g)** NMR structure #6 hybrid tetramer; and **(h)** NMR structure #7 hybrid tetramer.

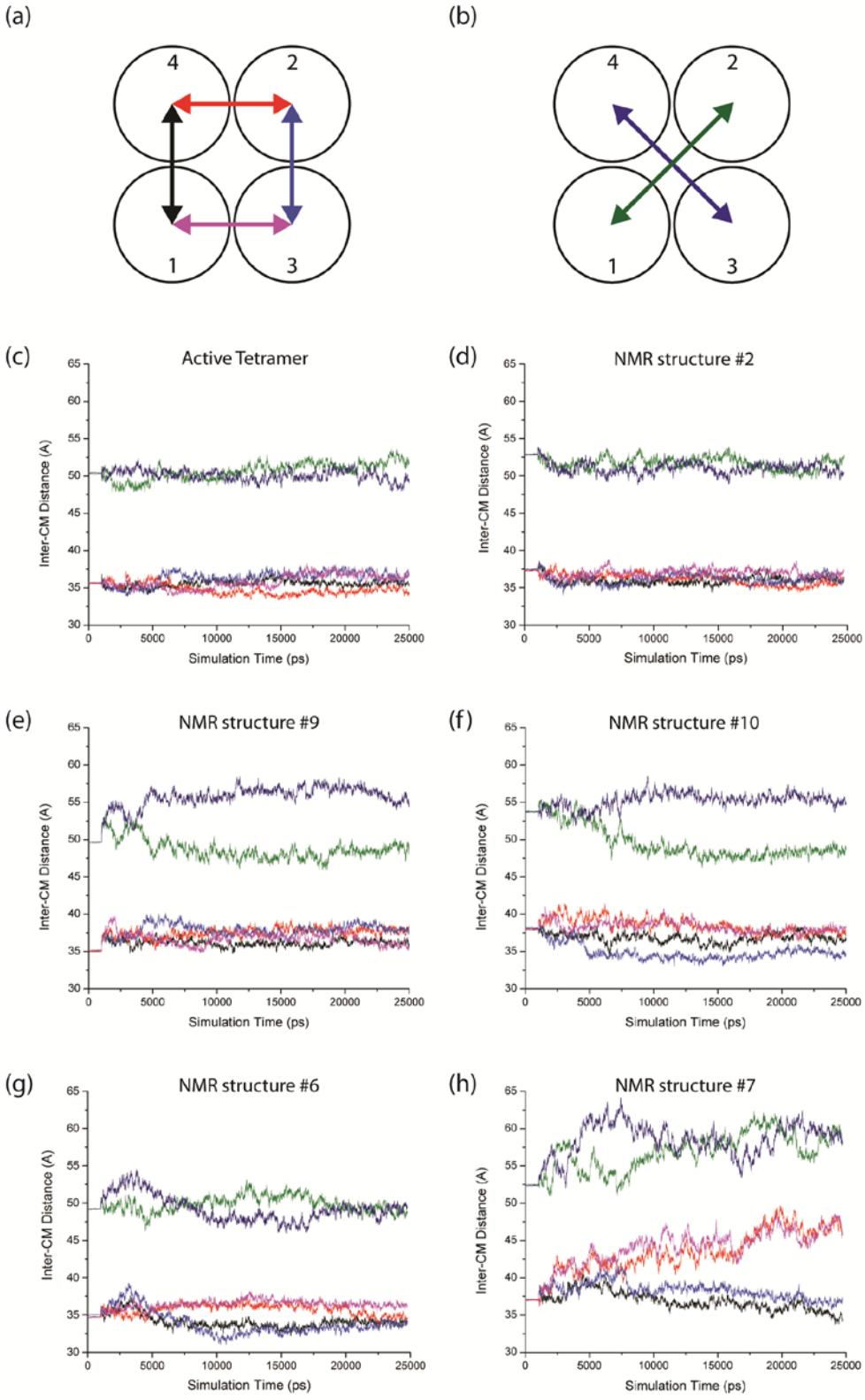
(a)
4
2
1
3
(b)
4
2
1
3
(c)
Active Tetramer
Inter-CM Distance (A)
Simulation Time (ps)
(d)
NMR structure #2
Inter-CM Distance (A)
Simulation Time (ps)
(e)
NMR structure #9
Inter-CM Distance (A)
Simulation Time (ps)
(f)
NMR structure #10
Inter-CM Distance (A)
Simulation Time (ps)
(g)
NMR structure #6
Inter-CM Distance (A)
Simulation Time (ps)
(h)
NMR structure #7
Inter-CM Distance (A)
Simulation Time (ps)

Figure 3: Inter-center-of-mass distance results for the CBD portion of the tetramer, as observed for each of the simulated "hybrid" tetramers. **(a,b)** Outline of the inter-center-of-mass distances (indicated as arrows) between protomers that are situated **(a)** adjacent and **(b)** opposite one another in the tetramer. Individual protomers are represented as numbered circles, and the colour codes shown for the distances correspond to the colour codes used in the plots. **(c-h)** Inter-center-of-mass distance trajectories computed for the distances outlined in panels a-b, with colour coding as in panels a-b: **(c)** cAMP-bound active-state tetramer (used here as a control); **(d)** NMR structure #2 hybrid tetramer; **(e)** NMR structure #9 hybrid tetramer; **(f)** NMR structure #10 hybrid tetramer; **(g)** NMR structure #6 hybrid tetramer; and **(h)** NMR structure #7 hybrid tetramer.

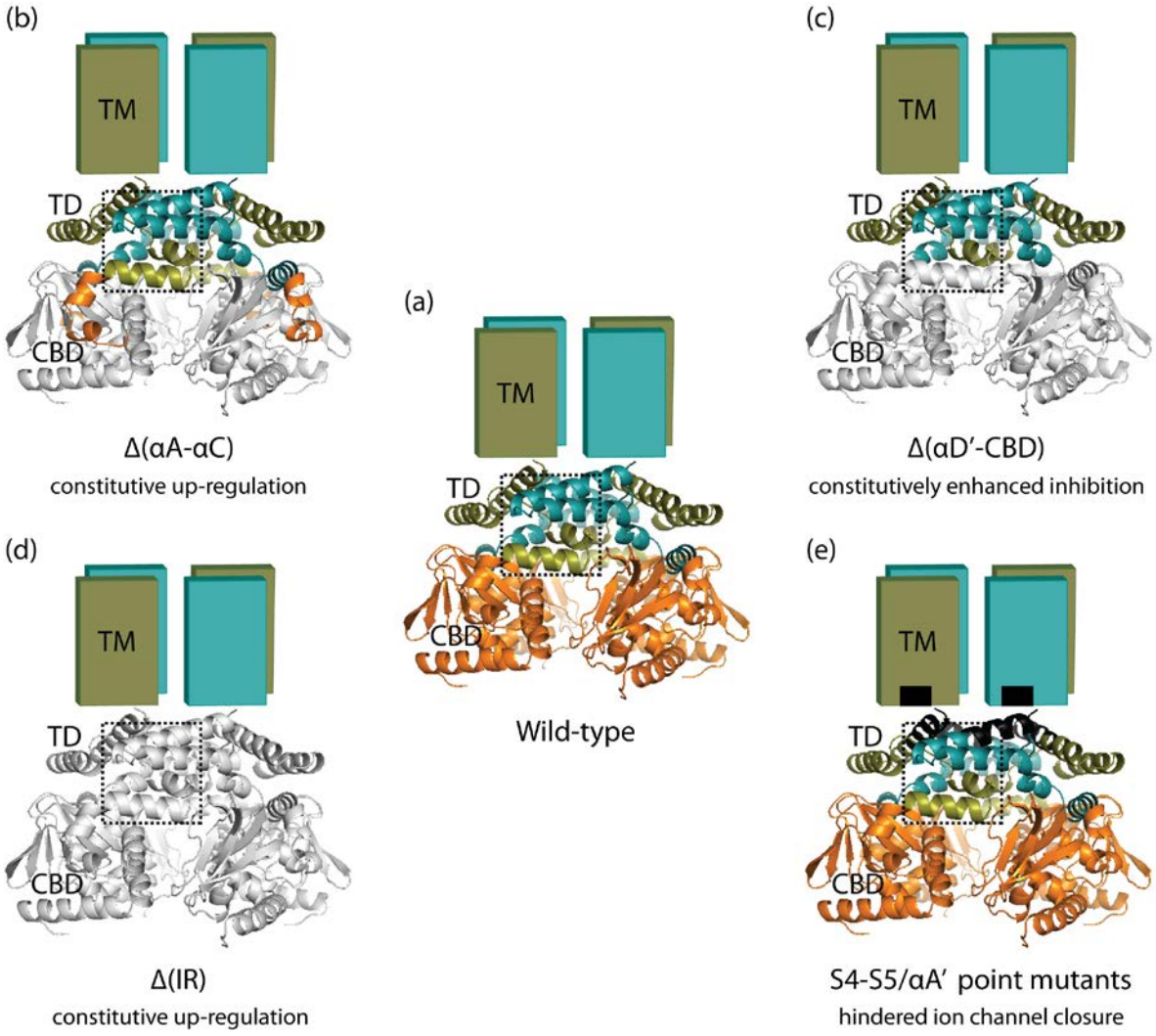

Figure 4: Outline of notable HCN mutants, and the resulting perturbations of HCN functional behavior (relative to the apo wild-type protein) as reported from previous electrophysiology-based functional assays.[5,6,9,10,11,12] All mutants are mapped in the active tetrameric structure of HCN4 (RCSB Protein Data Bank code "3OTF"), viewed parallel to the plane of the cell membrane. The N-terminal trans-membrane (TM) regions, whose atomic-resolution structure is currently unknown, are indicated as rectangles, while the tetramerization-domain (TD) and CBD segments of the C-terminal intracellular regions (IR) are shown as ribbon structures (for further details on the structure, see Chapter 3 Figure 1). For clarity, the TM and TD regions of alternating protomers are indicated in olive green and teal, the CBD regions are highlighted in orange, and the TD-tetramer interface between the αA'-αB' and αC'-αD' segments from adjacent protomers is indicated by a black dotted box. All segments deleted from the three deletion mutants (panels b-d) are shown in gray, and the locations of the regions subjected to point mutations (panel e) are highlighted in black (black rectangle = S4-S5 linker of TM region; black ribbon = αA' helix N-terminal segment).

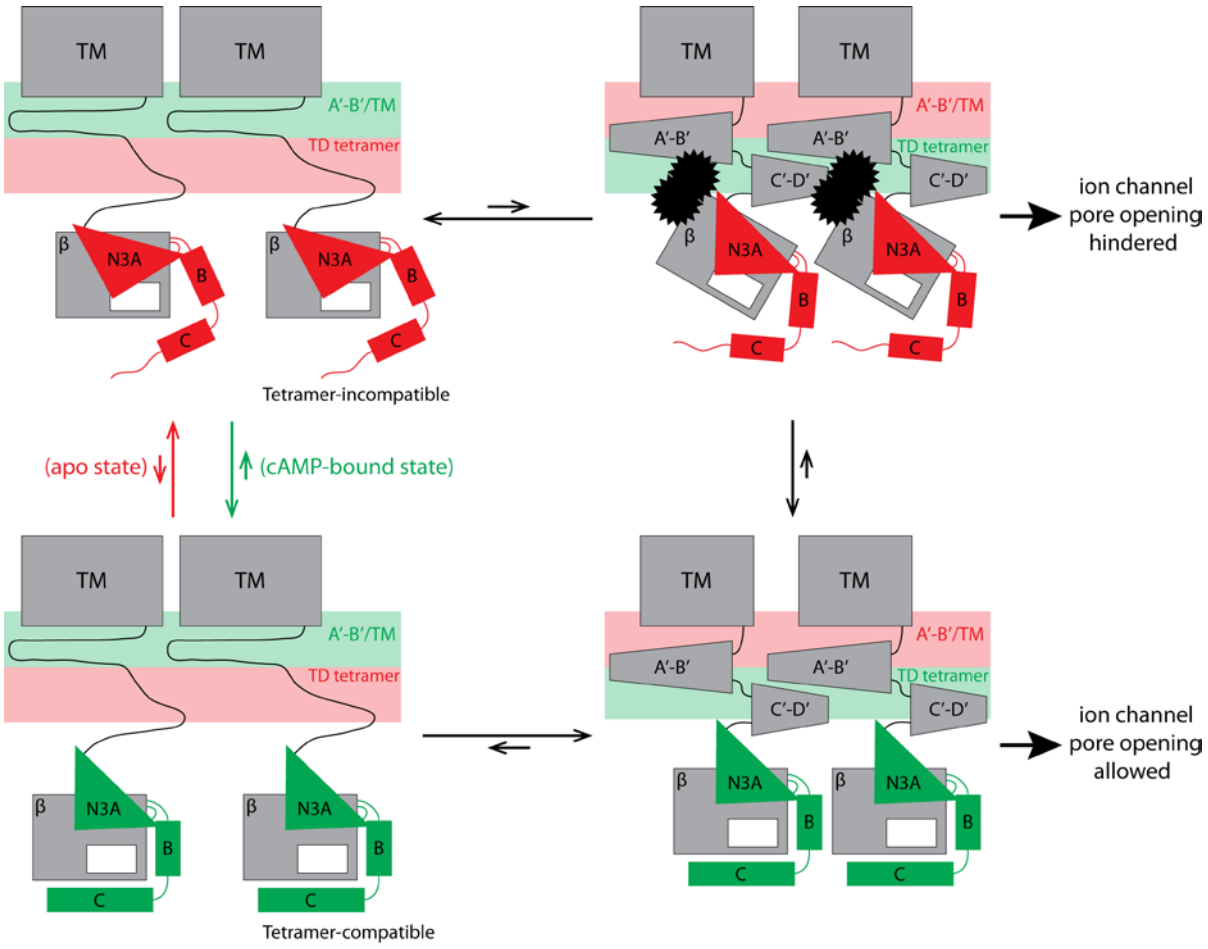

Figure 5: Schematic summary of the compiled allosteric model of HCN ion channel modulation by CBD conformational shifts. The trans-membrane (TM) regions and CBD β-subdomains ("β") are shown as gray rectangles, and the cAMP binding sites as white rectangles. The non-tetrameric tetramerization domain (TD) is represented as a black line between the TM and the CBD N3A, and is presumed to be less structured than in the tetrameric state (based on TD structural analyses described in Chapter 3); while the αA'-αB' and αC'-αD' segments of the more ordered tetrameric TD are shown as gray trapezoids. The N3A, αB ("B") and αC ("C") elements of the CBD are shown in red for tetramer-incompatible CBD conformations, and in green for tetramer-compatible CBD conformations; and TD/β-subdomain steric clashes brought on by tetramer-incompatible CBD conformations are indicated by black starbursts. The auto-inhibitory αA'-αB'/TM interface ("A'-B'/TM") and the competing TD-tetramer interface are each marked by red highlights if they are disengaged in the respective structural state, or by green highlights if they are engaged in the respective structural state; and the state of the CBD conformational equilibrium within the non-tetrameric state of HCN (left panels) is indicated for both the apo (red arrows) and cAMP-bound (green arrows) forms of the protein.

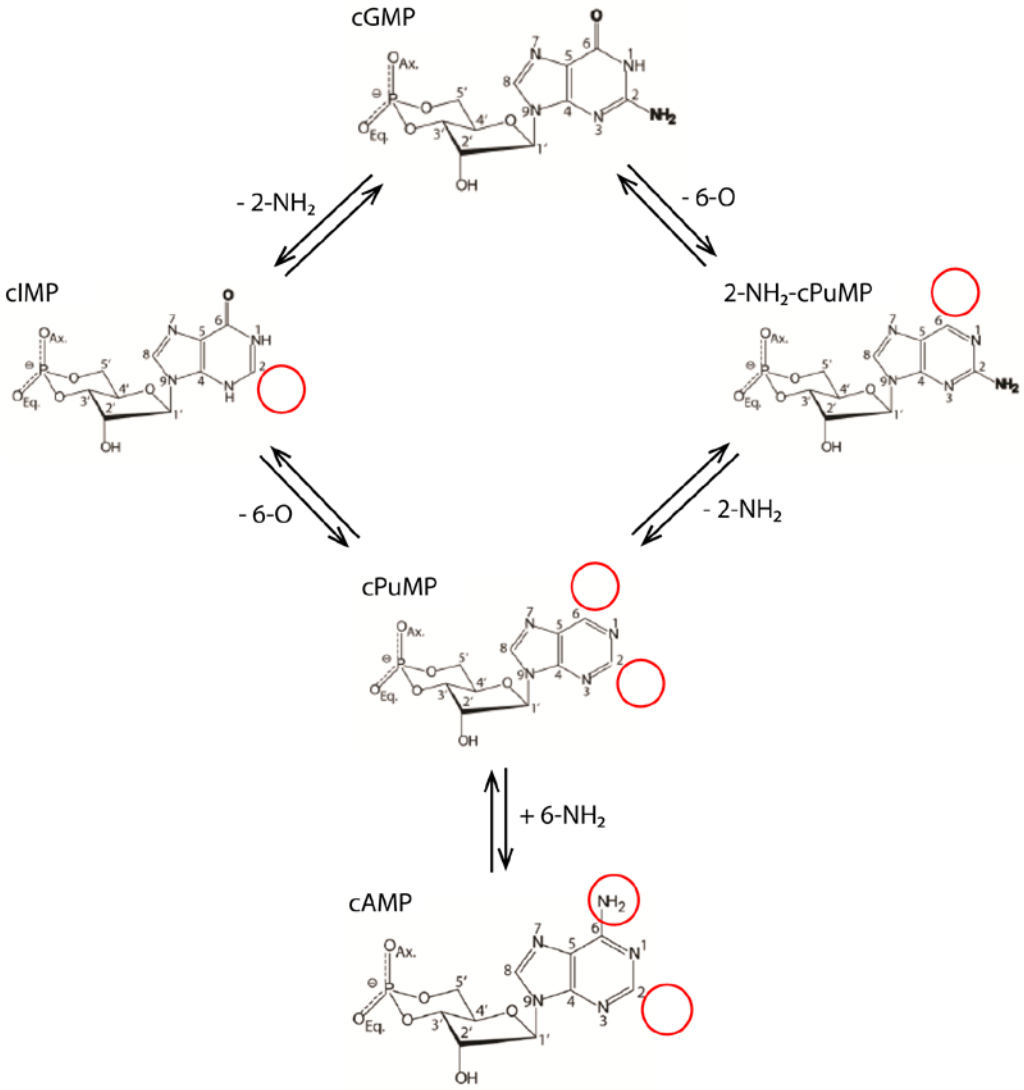

Figure 6: Double-mutant cycle for the conversion of cGMP into cAMP *via* the modified cyclic nucleotides cIMP, 2-NH_2-cPuMP and cPuMP. The substituent addition/removal at each step is indicated, and substituent differences from cGMP are highlighted in the structures by red circles.

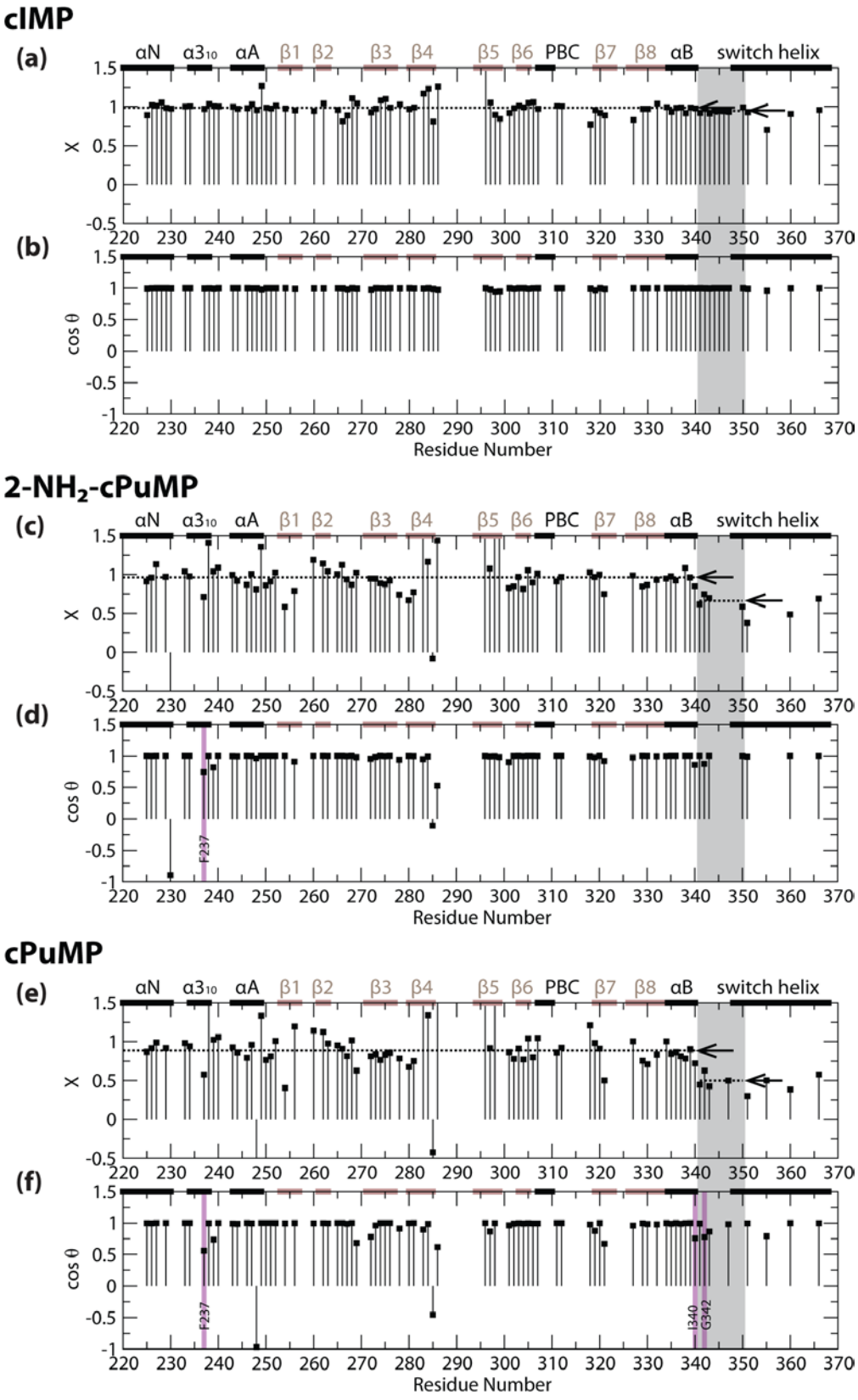
cIMP
(a)
αN α3₁₀ αA β1 β2 β3 β4 β5 β6 PBC β7 β8 αB switch helix
X
(b)
cos θ
Residue Number
2-NH₂-cPuMP
(c)
(d)
F237
cPuMP
(e)
(f)
F237
I340
G342

Figure 7: (a-b) Effects of cIMP binding on PKG Iβ(219-369), as observed from CHESPA analysis. The secondary structure elements from the apo-state X-ray structure are indicated across the top of each graph: black bars = α-helices; brown bars = β-strands. **(a)** Fractional shifts toward activation ("X") and **(b)** projection angle cosines ("cosθ") achieved by cIMP binding, relative to apo and cGMP-bound PKG Iβ(219-369). The average fractional shift toward activation for the residue 340-350 region, and overall fractional shift toward activation for residues N-terminal to residue 340 (both computed as in Chapter 4 Figure 2), are marked by black dotted lines and arrows. The residue 340-350 region is indicated in all plots by a gray highlight, and notable low-cosθ-value ("outlier") residues (identified as in Chapter 4 Figure 2) are indicated by purple highlights. **(c-d)** As for (a-b), but illustrating the effects of 2-NH_2-cPuMP binding on PKG Iβ(219-369). **(e-f)** As for (a-b), but illustrating the effects of cPuMP binding on PKG Iβ(219-369).

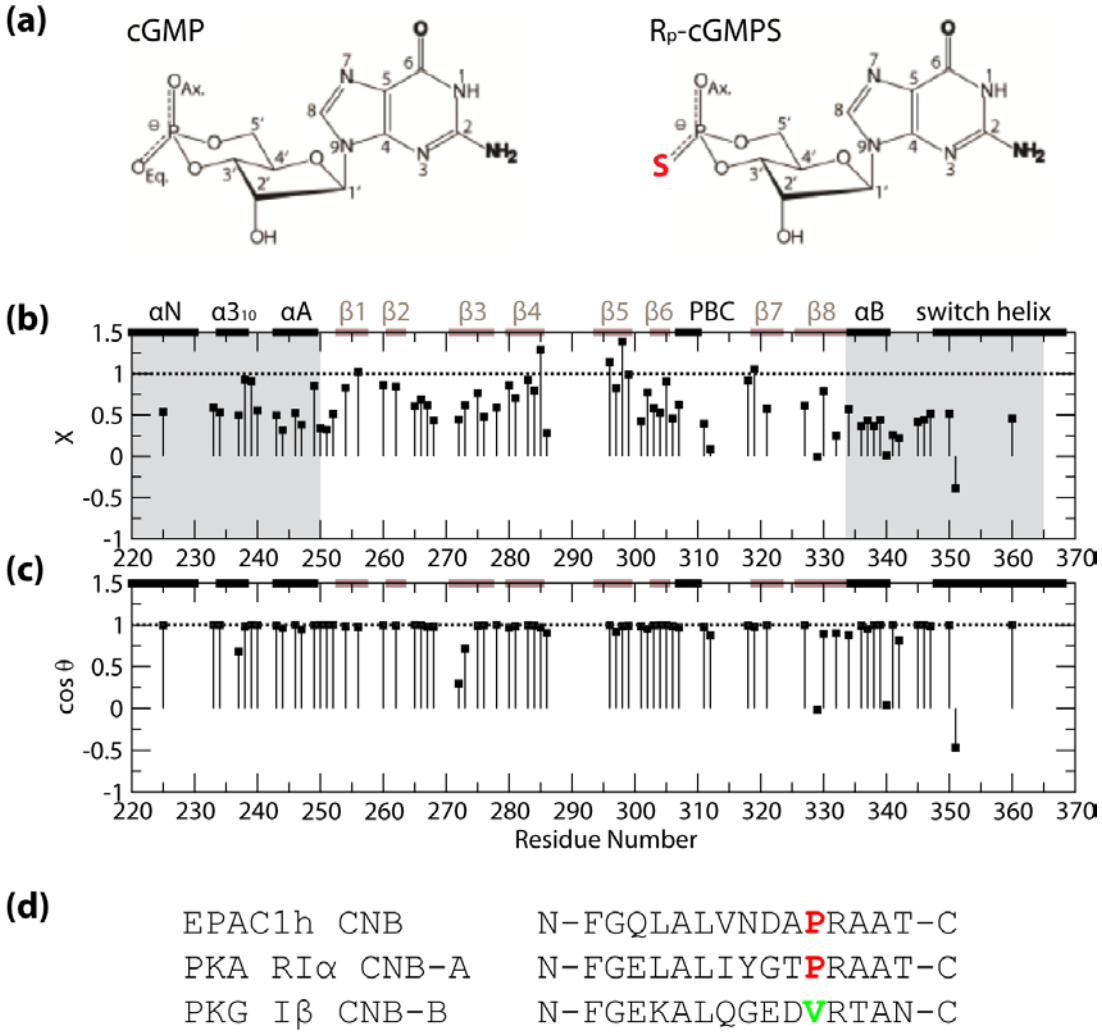

Figure 8: (a) Comparison of the structures of the canonical cGMP ligand of PKG, and its R_p-phosphorothioate analogue (*i.e.* R_p-cGMPS). The sulfur atom substituted for the equatorial oxygen atom of cGMP is highlighted in red in the R_p-cGMPS structure. **(b-c)** Effects of R_p-cGMPS binding on PKG Iβ(219-369), as observed from CHESPA analysis. The secondary structure elements from the apo-state X-ray structure are indicated across the top of each graph: black bars = α-helices; brown bars = β-strands. **(b)** Fractional shifts toward activation ("X") and **(c)** projection angle cosines ("cosθ") achieved by R_p-cGMPS binding, relative to apo and cGMP-bound PKG Iβ(219-369). Fractional shift and projection angle cosine values of 1 (corresponding to full cGMP-like activation) are marked by black dotted lines, and the N- and C-terminal α-helical elements (which undergo structural rearrangement in response to cGMP binding)[13,14] are indicated by gray highlights. **(d)** Amino acid sequence alignment of the PBC regions of CNB domains from EPAC, PKA and PKG. The proline residues present in the C-terminal segment of the PBC regions of EPAC and PKA are highlighted in red, and the corresponding residue in PKG (*i.e.* residue V315) is highlighted in green.

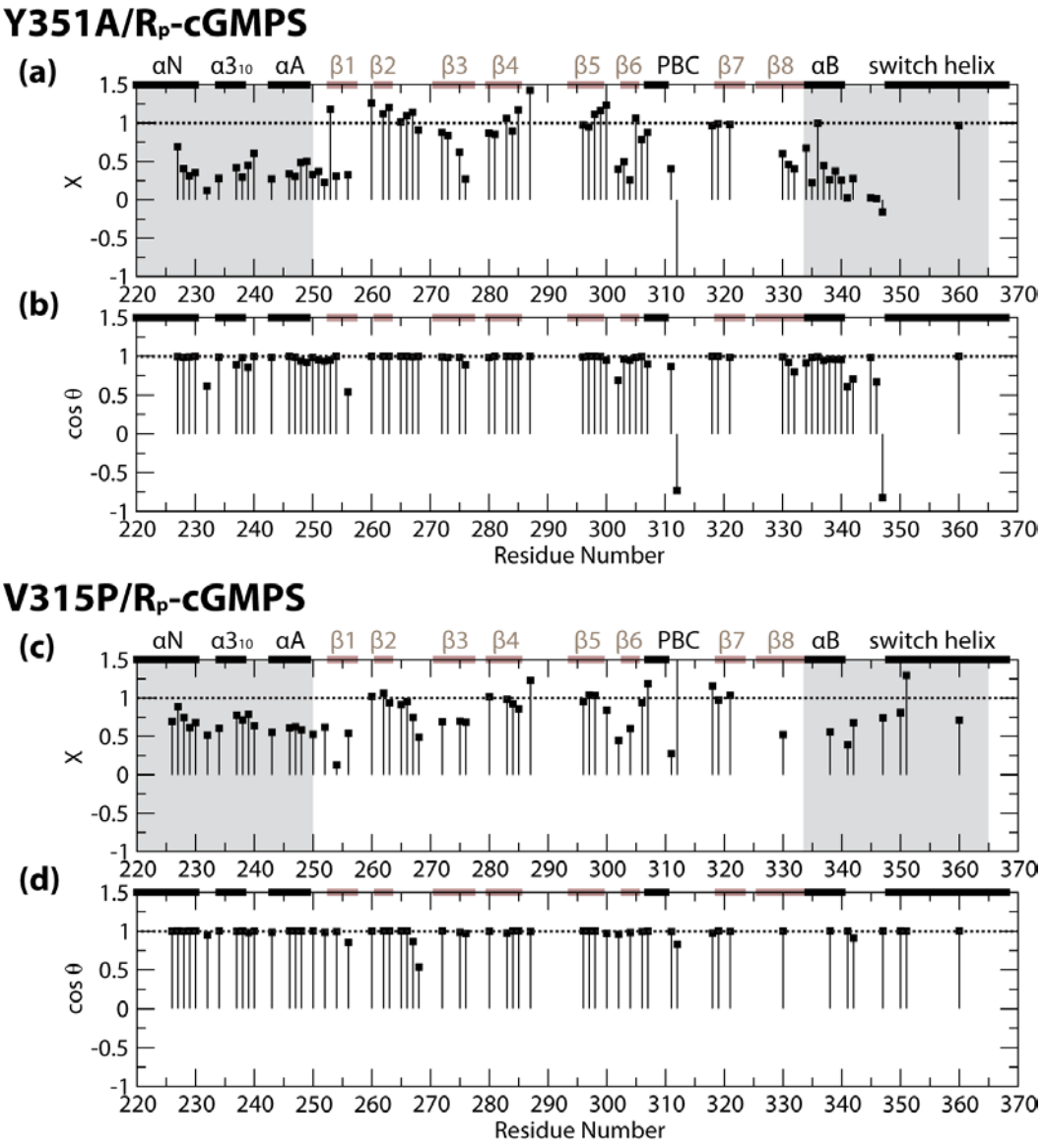

Figure 9: (a-b) Effects of the Y351A point-mutation on R_p-cGMPS-bound PKG Iβ(219-369), as observed from CHESPA analysis. The secondary structure elements from the apo-state X-ray structure are indicated across the top of each graph: black bars = α-helices; brown bars = β-strands. **(a)** Fractional shifts toward activation ("X") and **(b)** projection angle cosines ("cosθ") achieved by the R_p-cGMPS-bound Y351A mutant, relative to apo and R_p-cGMPS-bound wild-type PKG Iβ(219-369). Fractional shift and projection angle cosine values of 1 (corresponding to full wild-type R_p-cGMPS-like activation) are marked by black dotted lines, and the N- and C-terminal α-helical elements (which undergo structural rearrangement in response to cGMP binding)[13,14] are indicated by gray highlights. **(c-d)** As for (a-b), but illustrating the effects of the V315P point-mutation on R_p-cGMPS-bound PKG Iβ(219-369).

5.6. References

[1] Akimoto, M., Zhang, Z., Boulton, S., Selvaratnam, R., VanSchouwen, B., Gloyd, M., Accili, E.A., Lange, O.F. and Melacini, G. (2014) *J. Biol. Chem.* **289**: 22205-22220.
[2] Berrera, M., Pantano, S. and Carloni, P. (2006) *Biophys. J.* **90**: 3428–3433.
[3] Ulens, C. and Siegelbaum, S.A. (2003) *Neuron* **40**: 959-970.
[4] Lolicato, M., Nardini, M., Gazzarini, S., Möller, S., Bertinetti, D., Herberg, F.W., Bolognesi, M., Martin, H., Fasolini, M., Bertrand, J.A., Arrigoni, C., Thiel, G. and Moroni, A. (2011) *J. Biol. Chem.* **286**: 44811-44820.
[5] Wainger, B.J., DeGennaro, M., Santoro, B., Siegelbaum, S.A. and Tibbs, G.R. (2001) *Nature* **411**: 805-810.
[6] Wang, J., Chen, S. and Siegelbaum, S.A. (2001) *J. Gen. Physiol.* **118**: 237–250.
[7] Zagotta, W.N., Olivier, N.B., Black, K.D., Young, E.C., Olson, R. and Gouaux, E. (2003) *Nature* **425**: 200-205.
[8] Xu, X., Vysotskaya, Z.V., Liu, Q. and Zhou, Lei. (2010) *J. Biol. Chem.* **285**: 37082-37091.
[9] Herrmann, S., Hofmann, F., Stieber, J. and Ludwig, A. (2012) *Br. J. Pharmacol.* **166**: 501-509.
[10] Alig, J., Marger, L., Mesirca, P., Ehmke, H., Mangoni, M.E. and Isbrandt, D. (2009) *Proc. Natl. Acad. Sci. USA* **106**: 12189–12194.
[11] Decher, N., Chen, J. and Sanguinetti, M.C. (2004) *J. Biol. Chem.* **279**: 13859-13865.
[12] Chen, J., Mitcheson, J.S., Tristani-Firouzi, M., Lin, M. and Sanguinetti, M.C. (2001) *Proc. Natl. Acad. Sci. USA* **98**: 11277-11282.
[13] Huang, G.Y., Kim, J.J., Reger, A.S., Lorenz, R., Moon, E.W., Zhao, C., Casteel, D.E., Bertinetti, D., VanSchouwen, B., Selvaratnam, R., Pflugrath, J.W., Sankaran, B., Melacini, G., Herberg, F.W. and Kim, C. (2014) *Structure* **22**: 116-124.
[14] Huang, G.Y., Gerlits, O.O., Blakeley, M.P., Sankaran, B., Kovalevsky, A.Y. and Kim, C. (2014) *Biochemistry* **53**: 6725−6727.
[15] Selvaratnam, R., Chowdhury, S., VanSchouwen, B. and Melacini, G. (2011) *Proc. Natl. Acad Sci. USA* **108**: 6133-6138.
[16] Akimoto, M., Selvaratnam, R., McNicholl, E.T., Verma, G., Taylor, S.S. and Melacini, G. (2013) *Proc. Natl. Acad. Sci. USA* **110**: 14231-14236.
[17] Das, R., Mazhab-Jafari, M.T., Chowdhury, S., Das, S.S., Selvaratnam, R. and Melacini, G. (2008) *J. Biol. Chem.* **283**: 19691-19703.
[18] Das, R. and Melacini, G. (2007) *J. Biol. Chem.* **282**: 581-593.
[19] Mazhab-Jafari, M.T., Das, R., Fotheringham, S.A., Das, S.S., Chowdhury, S. and Melacini, G. (2007) *J. Am. Chem. Soc.* **129**: 14482-14492.
[20] Harper, S.M., Wienk, H., Wechselberger, R.W., Bos, J.L., Boelens, R. and Rehmann, H. (2008) *J. Biol. Chem.* **283**: 6501-6508.
[21] Rehmann, H., Das, J., Knipscheer, P., Wittinghofer, A. and Bos, J.L. (2006) *Nature* **439**: 625-628.
[22] Rehmann, H., Arias-Palomo, E., Hadders, M.A., Schwede, F., Llorca, O. and Bos, J.L. (2008) *Nature* **455**: 124-127.
[23] Kim, C., Cheng, C.Y., Saldanha, S.A. and Taylor, S.S. (2007) *Cell* **130**: 1032-1043.
[24] Su, Y., Dostmann, W.R., Herberg, F.W., Durick, K., Xuong, N.H., Ten-Eyck, L., Taylor, S.S. and Varughese, K.I. (1995) *Science* **269**: 807-813.