MAPPING ALLOSTERY IN UNCONVENTIONAL SYSTEMS

MAPPING ALLOSTERIC SITES AND PATHWAYS IN SYSTEMS UNAMENABLE TO TRADITIONAL STRUCTURE DETERMINATION

By STEPHEN BOULTON, B.Sc. Honours

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

Doctor of Philosophy (2018) **McMaster University**

(Biochemistry) **Hamilton, Ontario**

Thesis Title: Mapping Allosteric Sites and Pathways in Systems Unamenable to Traditional Structure Determination

Abbreviated Title: Mapping Allostery in Unconventional Systems

Author: Stephen Boulton, Honours B.Sc. (McMaster University)

Supervisor: Professor Giuseppe Melacini

Number of Pages: xix, 201

Lay Abstract

Allostery is a regulatory mechanism for proteins, which controls functional properties of one distinct site through the perturbation of another distinct, and often distant, site. The two sites are connected via a series of residues that undergo conformational changes once perturbed by the allosteric effector. Mapping these communication pathways reveals mechanisms of protein regulation, which are invaluable for developing pharmacological modulators to target these pathways or for understanding the mechanisms of disease mutations that disrupt these pathways. Allosteric pathways have been traditionally determined using structure determination approaches that provide a static snapshot of the protein's structure. However, these approaches are typically not effective when allostery relies extensive changes in dynamics. The goal of this thesis was to develop methods to characterize systems that are dynamic or otherwise unsuitable for traditional structure determination. Herein, we utilize NMR spectroscopy to analyze the allosteric mechanisms of three cAMP-binding proteins involved in cardiovascular health.

Thesis Abstract

Allostery is a regulatory process whereby a perturbation by an effector at one discrete locus creates a conformational change that stimulates a functional change at another. The two sites communicate through networks of interacting residues that respond in a concerted manner to the allosteric perturbation. These allosteric networks are traditionally mapped with high resolution structure determination techniques to understand the conformational changes that regulate protein function as well as its modulation by allosteric ligands and its dysfunction caused by disease-related mutations (DRMs). However, high resolution structural determination techniques, such as X-ray crystallography, cryo-electron microscopy and nuclear Overhauser effect NMR spectroscopy are not always amenable for systems plagued by poor solubility and line broadening caused by μs-ms dynamics or systems where allostery relies primarily on dynamical rather than structural changes. This dissertation discusses methodologies to map the allosteric sites and pathways for such challenging systems. The foundation of this approach is to model allosteric pathways in the context of their respective thermodynamic cycles. In chapter 2, the thermodynamic cycle of a DRM in the hyperpolarization-activated cyclic nucleotide-gated ion channel 4 (HCN4) is analyzed with respect to structure, dynamics and kinetics, revealing how the DRM remodels the free energy landscape of HCN4 and results in a loss-of-function disease phenotype. In chapter 3, the mechanism of action of an uncompetitive inhibitor for the exchange protein activated by cAMP is elucidated by characterizing its selectivity for distinct conformations within the thermodynamic cycle that are trapped using a combination of mutations and ligand analogs. In chapter 4, we discuss two new protocols for the chemical shift covariance analysis (CHESCA). The CHESCA is an approach that identifies allosteric signaling pathways by measuring concerted residue responses to a library of chemical perturbations that stabilize conformational equilibria at different positions. Overall, the approaches discussed in this dissertation are widely applicable for mapping the mechanisms of allosteric perturbations that arise from ligand binding, post-translational modifications and mutations, even in systems where traditional structure determination techniques remain challenging to implement.

Acknowledgements

First and foremost, I would like to thank my advisor, Professor Giuseppe Melacini, for the continued guidance over the past six years. I was able to learn so much from your mentorship that I will take with me for the rest of my life. More than anything, I will remember your many proverbs and the important messages they convey. I will continue to 'hope for the best, but prepare for the worst', 'not let the perfect get in the way of the good' and 'go after the low hanging fruit' whenever possible. It has been a pleasure working in your lab that I will never forget.

Next, I would like to thank all the members of the Melacini lab that I have had the fortune of working with. You have made these long years all the more enjoyable. I will always treasure our discussions in the lab or at group outings. You all helped me grow as both a scientist and a person, so for that I thank you. To my undergraduate thesis advisor, Dr. Bryan VanSchouwen, I'd like to thank you for teaching me the importance of following a protocol 'to the T'. To my predecessor, Dr. Rajeevan Selvaratnam, whose projects I've inherited, I'd like to thank you for your continued advice and friendship. To Dr. Madoka Akimoto, I've learned more from you than anyone else in the Melacini lab. I'm sorry for all the questions over the years. I would never have achieved so much if it weren't for your help. To my good friend, the self-proclaimed cloning master, Kody Moleschi, I'll always remember the good times we had in the lab and out. You were always down for a drink when I needed to blow off some steam. To Rashik Ahmed, it was fun gossiping with you over coffee breaks. I believe I was able to learn a lot more from you than you could from me. I'm sure you're going to be successful in the future, so if you ever start a company, hire me on as your CSO. For everybody else - Naeimeh Jafari, Olivia Byun, Moustafa Algamal, Melanie Gloyd, Pascale Swanson, Honzhou Shao and the many undergrads (I'm sorry I couldn't mention everyone), thank you for all the good times.

To my committee members, Drs. Alba Guarne and Nathan Magarvey, I valued your advice and comments over the years. Dr. Guarne, thank you for always providing access to your lab. It was difficult adjusting once you left. I hope you do well at McGill. Dr. Magarvey, I had fun times TA'ing your course. I still have a bunch of the swag that the students tried to bribe us with. I wish I could have tapped your knowledge and experience more effectively.

I would also like to thank the great staff at the NMR facility, Centre for Microbial Chemical Biology and Biointerfaces Institute for all your assistance and access to your facilities. In particular I would like to thank Dr. Bob Berno, who taught me much about NMR spectroscopy, Susan

McCusker who trained me on many of the instruments in the CMCB, and Marta Princz, who trained me at the BI facilities. Thank you again, for your patience and support.

Lastly, I need to express my sincere gratitude to my family. Without your support, I would never have gotten this far. You provided financial support when I was broke, motivation when I was lazy, relaxation when I was stressed and laundry services when I was… lazy. You would always listen to me talk about my work even when you didn't fully understand what I actually did. I thank you for that and for giving me the inspiration to pursue a scientific career. I owe my present success to you all more than anybody.

Table of Contents

List of Main Text Figures

Chapter 1: Advances in NMR Methods to Map Allosteric Sites: From Models to Translation

Chapter 2: Free Energy Landscape Remodeling for the Cardiac Pacemaker Channel Explains the Molecular Basis of Familiar Sinus Bradycardia

Chapter 3: Mechanism of Selective Enzyme Inhibition Through Uncompetitive Regulation of an Allosteric Agonist

List of Supplementary Figures

Chapter 2: Free Energy Landscape Remodeling for the Cardiac Pacemaker Channel Explains the Molecular Basis of Familiar Sinus Bradycardia

Chapter 3: Mechanism of Selective Enzyme Inhibition Through Uncompetitive Regulation of an Allosteric Agonist

Chapter 4: A Tool Set to Map Allosteric Networks Through the NMR Chemical Shift Covariance Analysis

List of Tables

List of Supplementary Tables

List of Abbreviations

AC	agglomerative clustering
BBR	base binding region
cAMP	adenosine 3'5' cyclic monophosphate
CBD	cAMP-binding domain
CCS	compounded chemical shift
CEST	chemical-exchange saturation transfer
CHESPA	chemical shift covariance analysis
CHESCA	chemical shift projection analysis
CR	catalytic region
CS	conformational selection
CTL	C-terminal lid
CPMG	Carr-Purcell-Meiboom-Gill
DDA	dynamically driven allostery
DMSO	dimethyl sulfoxide
GEF	guanine exchange factor
EM	electron microscopy
EPAC	exchange protein activated by cAMP
ES	enzyme-substrate
ESI	EPAC selective inhibitor
HCN	hyperpolarization-activated cyclic nucleotide-gated
HSA	human serum albumin
HSQC	heteronuclear single quantum coherence spectroscopy
HTS	high throughput screening
IDP	intrinsically disordered protein
IF	induced fit
IL	ionic latch
IPC	inter-residue pairwise correlation
IR	intracellular region
ITC	isothermal titration calorimetry
MD	molecular dynamics
MTSL	*S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl) methyl methanesulfonothioate
NMR	nuclear magnetic resonance

NMRD	NMR relaxation dispersion
NOESY	nuclear Overhauser effect spectroscopy
NTHB	N-terminal helical bundle
PBC	phosphate binding cassette
PCA	principle component analysis
PKA	protein kinase A
PRE	paramagnetic relaxation enhancement
PTM	post translational modification
RDC	residual dipolar couplings
RDSM	root-mean-square-deviation
ROESY	rotating frame Overhauser effect spectroscopy
RR	regulatory region
SI	supplementary/supporting information
SLAPSTIC	spin-labels attached to protein side chains as a tool to identify interacting compounds
STD	saturation transfer difference
SVD	singular value decomposition
TMD	transmembrane domain
WT	wild type

Chapter 1

Advances in NMR Methods to Map Allosteric Sites: From Models to Translation

1.1 Author's Preface

The work discussed in this chapter has previously been published and is reproduced here with permission from the American Chemical Society. The full citation is as follows:

Boulton S and Melacini G. Advances in NMR Methods to Map Allosteric Sites: From Models to Translation. *Chem. Rev.* **116**(11), 6267-6304 (2016).

This publication encompasses a review of methodologies used to map allosteric sites and pathways that were discussed in articles primarily between 2011-2016. Our review is focused largely on NMR experiments, but includes sections on X-ray crystallography, small angle X-ray scattering, cryo-electron microscopy, molecular dynamics simulations and bioinformatics. It provides a thorough overview of both traditional and emerging methodologies to probe allosteric pathways and therefore serves as a suitable introductory chapter for my thesis about experimental approaches to analyze allosteric signaling in non-conventional systems, such as those in which structural changes fail to explain allosteric phenotypes or structural determination methods cannot be applied. Many of the experimental approaches that were adopted for the systems investigated in the thesis were adapted, in part, from pre-existing experiments that are discussed within this review. The review also summarizes the impact of allostery in drug development and disease-related mutations, thereby rationalizing the translational potential of the work discussed in this thesis. An additional section (1.8) follows the review, which summarizes the connective themes of this dissertation and provides an overview of the ensuing chapters.

1.2 Abstract

The last five years have witnessed major developments in the understanding of the allosteric phenomenon, broadly defined as coupling between remote molecular sites. Such advances have been driven not only by new theoretical models and pharmacological applications of allostery, but also by progress in the experimental approaches designed to map allosteric sites and transitions. Among these techniques, NMR spectroscopy has played a major role given its unique near-atomic resolution and sensitivity to the dynamics that underlie allosteric couplings. Here, we highlight recent progress in the NMR methods tailored to investigate allostery with the goal of offering an overview of which NMR approaches are best suited for which allosterically relevant questions. The picture of the allosteric "NMR toolbox" is provided starting from one of the simplest models of allostery (*i.e.* the four-state thermodynamic cycle) and continuing to more complex multi-state mechanisms. We also review how such an "NMR toolbox" has assisted the elucidation of the allosteric molecular basis for disease-related mutations and the discovery of novel leads for allosteric drugs. From this overview, it is clear that NMR plays a central role not only in experimentally validating transformative theories of allostery, but also in tapping the full translational potential of allosteric systems.

1.3 Introduction

Allostery is broadly defined as the presence of couplings between distinct sites within a molecular system[1–6]. Typical allosteric systems are proteins, monomeric or oligomeric, involved in multiple interactions, in which perturbations at one interaction site modulate other site(s). A classic allosteric perturbation is ligand binding, as often observed in the case of enzymes and carrier proteins. In enzymes, activity is regulated by effectors binding at locations remote from the active sites, and in multi-site carrier proteins, such as tetrameric hemoglobin or monomeric human serum albumin, ligand binding at one site controls affinities at other sites. However, allosteric perturbations are not limited only to binding interactions, but they include also a wide range of post-translational modifications (*e.g.* phosphorylation, glycation, nitration just to cite a few)[7–14] as well as mutations[15–21]

Allostery is a fundamental mechanism to modulate biological function. It controls a broad range of physiological processes, including homeostasis through regulatory feedback loops[22], response to stimuli and stressors through cellular signaling[21,23,24], protein folding[25,26] , protein degradation[27] and the metabolism of xenobiotic and endogenous compounds[28] . In addition, allostery is of interest not only in physiology, but also in pathology, as defective allosteric signaling often results in deregulation and disease, and in immunology, as allosteric coupling between different antigen-binding epitopes modulates the immune response[29]. Furthermore, allosteric sites have opened new opportunities for selective drug targeting of systems previously considered "undruggable", as allosteric loci are subject to reduced evolutionary pressure for conservation compared to their orthosteric counterparts and they are not necessarily pre-occupied by competing endogenous ligands[30–34]. Hence, allosteric models and methods to exhaustively and accurately map allosteric sites have attracted increasing attention both in basic and applied research[35].

Modeling allostery is a challenging task, as it requires that structural analyses be extended beyond the ground state of the conformational ensemble to the excited states, which are only scarcely populated. Hence, allosteric models require a full map of structure and dynamics for the ground and excited states accessible within the complex free energy landscape of an allosteric system[36–40]. In spite of these complexities in modeling allostery, it is often advantageous to take an "Occam's razor" approach and focus first on one of the simplest possible models of allostery, *i.e.* the four-state thermodynamic cycle that arises from the coupling of two equilibria, a two-state conformational equilibrium and a ligand-binding equilibrium with 1:1 stoichiometry (Fig. 1a)[41,42]. Similar thermodynamic cycles apply to other allosteric perturbations different from ligand binding,

e.g. post-translational modifications and mutations that alter the active *vs.* inactive free energy difference through state-selective stabilization or destabilization[4,43–45]. Hence, the thermodynamic cycle of Fig. 1a based on coupled conformational and binding equilibria provides a simple but general conceptual framework to start rationalizing allostery.

An example of a two-state conformational equilibrium is the exchange between the tensed (T) and relaxed (R) conformations of the haemoglobin protomers, which exhibit low and high affinities for the O_2 ligand, respectively, as originally proposed by Monod, Wyman and Changeux (MWC) more than fifty years ago[46]. In more general terms, the two allosteric states are two sub-sets of conformations that differ at least at the level of the binding site for the allosteric effector (*i.e.* the "allosteric site") and at a functional site, which is typically an additional ligand/protein binding site or an auto-inhibitory site, distinct and remote from the allosteric site. In this respect, the two allosteric states are often referred to as the inactive and active states. The differences between the two states at the level of the allosteric site ensure that the allosteric effector binds the two states with different affinities (Fig. 1a,b). The differential affinities of the allosteric effector for the two states (*i.e.* state-selectivity) are essential for ensuring that the binding and conformational equilibria are reciprocally coupled with a free energy of coupling equivalent to:

$$\Delta G_{Allosteric\ Coupling} = RT\ ln\left(\frac{K_{a,Inactive}}{K_{a,Active}}\right) \qquad (1)$$

where R is the gas constant, T is the temperature in Kelvins and $K_{a,Inactive}$ and $K_{a,Active}$ denote the association constants of the allosteric effector for the inactive and active states, respectively[47]. The binding of the allosteric effector will selectively stabilize the state exhibiting the highest affinity (Fig. 1a,b). Since the two states differ also at the level of the functional site, the change in state populations caused by the allosteric effector results in a modulation of function.

Mapping the inactive *vs.* active state differences is often the first step towards understanding allostery in molecular terms. The inactive *vs.* active differences at the level of the binding and functional sites reveal the determinants of inactive *vs.* active selectivity and of functional modulation, respectively. The inactive *vs.* active differences at other sites provide further mechanistic insight and are collectively referred to here as the 'allosteric network'. Although this definition of an allosteric network does not imply any specific temporal sequence of events for the allosteric signal, it often provides insight into the molecular determinants of allosteric coupling between the binding and functional sites. Hence, a full map of inactive *vs.* active state differences is the foundation of projects aimed at dissecting the molecular mechanisms of allostery. The first step towards reconstructing such a map typically involves comparative apo *vs.* holo analyses.

Traditional structural determination methods, based on either X-ray crystallography or NMR spectroscopy, often capture the most populated state of each form (*i.e.* the apo-inactive and the holo-active; Fig. 1a,b). The apo-inactive and holo-active structures provide essential initial and final snapshots of the allosteric transition, respectively, and their comparison typically reveals key allosteric hot-spots, as explained in Section 2.1. However, one limitation of the apo-inactive *vs.* holo-active structural comparison is that it is not necessarily informative on the nature of the allosteric transition, *i.e.* the free-energy barriers and the local minima where the transition intermediates reside (Fig. 1c). Two key allosteric intermediates are the apo-active and the holo-inactive states (Fig. 1a-c).

According to the conformational selection (CS) model, as originally proposed by MWC for haemoglobin, the preferred pathways from the apo-inactive to the holo-active states are through the apo-active intermediate, whereby the apo protein pre-samples the active state prior to ligand binding, which selects for the pre-existing active conformers (Fig. 1d). However, according to the induced fit (IF) model, as originally proposed by Koshland, Nemethy and Filmer (KNF) for hemoglobin[48,49], the preferred apo-inactive to holo-active pathway is through the holo-inactive intermediate, *i.e.* binding occurs before the conformational changes, unlike what was posited by the CS model (Fig. 1c).

While the available experimental data support the generality of the CS model[50,51], the IF model is useful to explain how allosteric ligands access binding sites that are occluded in the holo-active state, for which binding is contingent upon opening a gating "lid" or portal region. In addition, it should be noted that allosteric transitions may occur through a manifold of pathways that involve a mixture of both IF and CS intermediates, depending on the specific allosteric system under investigation, the ligand concentration and on the specific experimental conditions[52–56]. In general, the IF and CS models are part of a continuum of free energy landscape scenarios[39,40]. Such scenarios range from instances in which the active *vs.* inactive free energy difference of the apo system is >> RT and the apo-active population is negligible (Fig. 1c), as posited by the IF model, to situations in which the apo-active *vs.* apo-inactive free energy difference is ≥ RT and the apo-active population is detectable (Fig. 1d), as expected for the CS case. Hence, the IF *vs.* CS distinction may depend also on the specific experimental techniques available to detect the apo-active state and the respective detection thresholds. Furthermore, the CS *vs.* IF differences are not limited to the equilibrium thermodynamics (*i.e.* ΔG between free energy local minima and relative state populations), but they pertain also to the kinetics of binding and conformational transitions (*i.e.* free

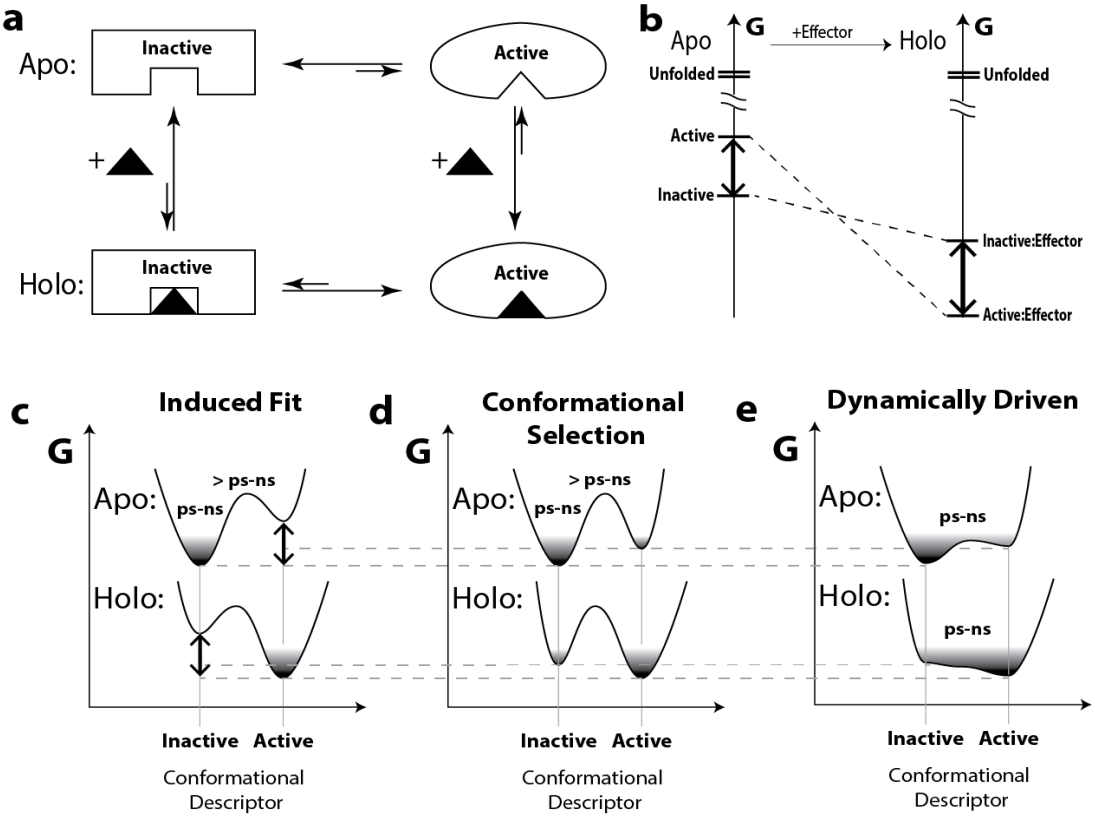

Figure 1. *Two-state models of allosteric coupling.* **(a)** Thermodynamic cycle arising from the coupling of a two-state inactive *vs.* active equilibrium and the effector (filled triangle) binding equilibrium. The allosteric coupling arises from the state-selectivity of the effector ligand and from the inactive *vs.* active differences in structure and/or dynamics that extend beyond the effector-binding site. **(b)** Free energy diagram of the cycle in panel (a). This is a purely thermodynamic representation, devoid of kinetic detail. **(c-e)** More complete description of the free energy landscape, including free energy barriers and hence the kinetics for the inactive *vs.* active equilibria in the apo and holo forms. Darker shading within a free energy basin denotes higher populations in agreement with the Boltzmann distribution. If the inactive *vs.* active $|\Delta G| >> RT$, the excited state is only negligibly populated, as expected in the case of the induced fit model **(c)**. When the inactive *vs.* active $|\Delta G|$ is reduced, the excited state becomes partially thermally accessible and a conformational selection scenario is possible **(d)**. In both panels (c) and (d), when the free energy barrier that separates the two states is sufficiently high, the inactive *vs.* active transitions typically occur in time-scales > ns, which include the fast, intermediate (ms-μs) or slow exchange regimes. Faster (ps-ns) motions usually still exist within the rugged surface of the free energy well of each state. However, if the inactive *vs.* active free energy barrier decreases, *e.g.* panel **(e)**, fast ps-ns motions span a wider region of conformational space. The amplitude of ps-ns motions reports on the width of the merged free energy well **(e)**, which in turn is a measure of conformational entropy. In this context, it is possible that allostery be driven primarily by dynamics. Alternatively, if the apo-inactive *vs.* apo-active free energy barrier increases, it is possible that the apo-active state becomes kinetically inaccessible even when the inactive *vs.* active $|\Delta G|$ is comparable to that shown in panel (b), resulting in an mechanism that more closely resemble the induced fit than the conformational selection. In general, the distinction between the induced fit and conformational selection fluxes depends on both thermodynamics and kinetics, *i.e.* on both local minimum *vs.* local minimum free energy differences and on local maximum *vs.* local minimum free energy differences, as well as on the experimental conditions, as explained in the text. Panels **(c-e)** were adapted with permission from Weinkam et al.[59]. Copyright 2012 National Academy of Sciences, U.S.A.

energy barriers between local minima and transition rates)[56]. The differences between the thermodynamics of the apo-inactive *vs.* apo-active equilibria of Fig. 1c,d provide only a necessary, but insufficient, condition for the CS *vs.* IF pathway distinction, which ultimately relies on the prevailing flux[56]. The latter depends on both thermodynamics and kinetics, *i.e.* both equilibrium and rate constants define the fluxes through the CS and IF pathways under a given set of experimental conditions[56].

Another reason why the detection and structural characterization of the apo-active and holo-inactive intermediates is an informative exercise is that the apo-active and holo-inactive intermediates are critical reference states necessary to reliably separate binding from purely allosteric effects. The simple apo-inactive *vs.* holo-active binary comparison is typically not sufficient for obtaining a reliable binding *vs.* allostery separation, whereas contrasting the apo-inactive *vs.* apo-active or the holo-inactive *vs.* holo-active singles out purely allosteric effects. Similarly, the apo-inactive *vs.* holo-inactive and the apo-active *vs.* holo-active comparisons report on perturbations that arise uniquely from binding. In addition, the apo-active intermediate is critical to understand the mechanism of endogenous auto-inhibition or constitutive activation, while the holo-inactive state reveals the interactions that are important for effector binding without stimulating activation, which help facilitate the design of exogenous inhibitors. The latter include both competitive allosteric inhibitors, which target the same allosteric site as the endogenous ligand, and uncompetitive allosteric inhibitors, which bind at distinct allosteric sites present uniquely in the holo-inactive state[57,58]. Hence, characterizing the holo-inactive and apo-active intermediates is critical to understand allostery and map allosteric binding sites with therapeutic potential. Recent advances in this direction are outlined in Section 2.2.

The previous discussion assumes that the active and inactive states correspond to distinct minima in the free energy landscape of the IF and CS models (Fig. 1c,d). However, when the free energy barrier between the active and inactive states is not >> RT, the free energy basins of the CS scenario (Fig. 1d) merge into a wider free energy well where allostery may be driven primarily by entropy[59] (Fig. 1e). In this context, a key determinant of entropy is structural dynamics, hence the term "dynamically driven allostery" [60,61]. Dynamically driven allostery (DDA) (Fig. 1e) explains how it is possible for allostery to occur without detectable changes in structure[15,60,61]. Dynamics not only are key determinants of the state entropy and hence of the free energy[62–71], but they also provide a means for transiently sampling neighbouring states in the free energy landscape of the allosteric cycle through synchronous collective motions (Fig. 1c-e)[72,73]. Furthermore, dynamics provide a

mechanism to explain the ability of a single binding site to bind multiple ligands[2,65,66,73–83]. The rates of dynamic conformational equilibria are also critical to identify rate-determining steps in the kinetics of binding and enzymatic catalysis[72,84–87]. Overall, dynamical changes are as relevant as structural variations[86,88–90] and mapping structural changes is necessary but not sufficient to understand allostery[8,42,91–95,95]. Profiling allosteric sites and dissecting the driving forces for allosteric transitions require a full atomic-resolution map of both structural and dynamical changes along the four-state thermodynamic cycle of Fig. 1a, which defines the free-energy landscape for the simplest allosteric coupling (Fig. 1). In this respect, NMR spectroscopy is an extremely useful and versatile tool to investigate allosteric systems. It provides the necessary resolution and probes both structure and dynamics, yielding an experimental map of the structural and dynamical changes that occur within the free energy landscape of allosteric systems.

Here, we review recent advances in NMR methods for characterizing allosteric systems. The main goal of this review is to provide an overview of which allosterically-relevant questions are addressed by which methods (Table 1), rather than providing technical details of different approaches. For this purpose, we outlined representative applications to allosteric systems of state-of-the-art methods, focusing primarily on contributions within the past ~four years, as excellent reviews are already available for prior years[81,89,96–99]. The review is articulated in four major sections. After an introductory section, we provide a survey of the approaches currently available to map allostery in the context of the four-state thermodynamic cycle (Section 2). Then, we summarize key advances in mapping allosteric transitions beyond the four-state thermodynamic cycle (Section 3). Last but not least, we review how the translational potential of allosteric system is tapped by extending the approaches outlined in Sections 2 and 3 to the discovery of new allosteric effectors and to disease-related mutations (Section 4).

1.4 Mapping Allostery in the Context of the Four-State Thermodynamic Cycle

1.4.1 Comparative Structural Analyses

1.4.1.1 Comparison of Apo vs. Holo Structures

The most traditional and still widely applied approach to allosteric proteins is the comparative analysis of discrete functional states for which structures are solved by X-ray crystallography, NMR spectroscopy and/or electron microscopy. Such comparisons, which in the simplest cases are between the apo and holo forms of the allosteric system[100–114], reveal how changes at functional

Table 1. Overview of NMR techniques discussed in this review[a]

Method	Description	Prime Time Scale[b]	Information Provided	Sections	Selected References
Perturbation Mapping					
Chemical Shift Perturbation (CSP) NMR	A binary comparative analysis of chemical shift changes before and after a perturbation (*i.e.* ligand binding, PTM, mutation, etc.)	slow-fast	ligand binding sites, K_d measurements, conformational changes, structural distortions	2.1.2, 2.1.3, 2.2.2, 3.2.3, 4.2	140,141,389
Chemical Shift Projection Analysis (CHESPA)	A ternary comparative analysis of direction and magnitude of chemical shift changes induced by a perturbation in relation to two reference values	fast	Measurement of fractional activations, identification of ligand binding sites and three-state equilibria	2.2, 3.2.3, 4.2	47,121,226
Chemical Shift Covariance Analysis (CHESCA)	Chemical shift analyses involving four or more samples to identify correlated residue responses to a targeted library of perturbations	fast	Mapping functional allosteric networks, clusters and hot spots.	2.2.2.1	42,115,120,226
Structure					
Chemical Shifts	^{1}H, ^{15}N, ^{13}C Assignments		Structures of Protein Domains	2.1, 2.2.2.3	141-143, 146, 243, 246
NOEs & J	Through-space ^{1}H-^{1}H cross-relaxation (NOE) and through-bound scalar coupling constants (J)	-	Short range distances and dihedral angles, binding interfaces by intermolecular NOEs	2.1.2, 2.1.3, 3.3	279,390,391
Residual Dipolar Couplings (RDCs)	Measurement of dipolar couplings not fully averaged out due to partial alignment		Orientation of domains and structural motifs and dynamics	2.1.2, 2.1.3, 3.2.3, 3.3	161,162,392, 393
Paramagnetic Relaxation Enhancement (PRE)	Induction of enhanced T_2 relaxation by a covalently linked spin label (SL, *i.e.* a paramagnetic molecule with unpaired electrons)	-	Long-range (< ~35Å) distances from position of spin label, long-range conformational changes, domain orientation, ligand binding sites	2.1.2, 2.1.3	169,393
Pseudo-contact chemical shifts (PCS)	Chemical shift changes arising from the dipolar coupling between the magnetic moments of a nucleus and an unpaired electron (metal)	-	Long-range (< ~40Å) distance from metal, conformational changes	2.1.3	170,172
Dynamics[c]					
Chemical Shifts	Deviations from reference 'random coil' values	fast	Approximate S^2 order parameters, rotamer populations	2.2.2.3	139, 144, 242, 247
T_1, T_2, HN-NOE	Relaxation rates sensitive to ps-ns motions[d]	fast[d]	Correlation times, S^2 order parameters, entropy changes	2.1.3, 2.2.3.1, 2.2.3.2, 4.1, 4.2, 4.3	393–395
T1ρ, NMR Relaxation Dispersion (NMRD), TROSY Hahn-Echo	Relaxation rates sensitive to µs-ms motions	intermediate	Identification and characterization of scarcely populated states, exchange rates, state populations	2.2, 3.2.2, 4.2	201,393–395
Chemical Exchange Saturation Transfer (CEST)	Saturation transfer between excited state (minor population) and ground state (major population) through chemical exchange	slow	Identification and characterization of lowly populated states, exchange rates, state populations	2.2	37,38,86,200
Longitudinal exchange (e.g. z or zz exchange)	Longitudinal magnetization or two-spin order transfer during mixing time	slow	Chemical shift assignment, kinetics (k_{ex})	2.2, 3.2.1	195,196,198, 393
Hydrogen Exchange	Detection of hydrogen exchange with solvent H/H exchange (~ 1-50 s^{-1}); H/D exchange (< $10^{3} s^{-1}$) Equilibrium Fractionation factors	-	Solvent accessibility, hydrogen bond strength, transient thermal unfolding free energies	2.2.4	275,396,397
Ligand-Based Methods[e]					
Saturation Transfer Difference (STD)	Transfer of magnetization from a saturated receptor to a bound ligand or viceversa (Kd range ~µM – ~mM)	Intermediate--fast	Epitope mapping, average and state-specific K_d measurements, drug screening	2.3.1, 2.3.2, 4.1	47,283,398
Transfer-NOESY	^{1}H-^{1}H distances for receptor-bound ligands (Kd range ~µM – ~mM)	Intermediate-fast	Bound-ligand conformation	2.3.1	277,282
Target Immobilized NMR Screening (TINS)	The target protein is immobilized onto a solid support and a mixture of test compounds is added. Ligand binding is identified by intensity losses.	slow-fast	Ligand binding, drug screening	4.1	342
Fluorine chemical shift Anisotropy and eXchange for Screening (FAXS)	A competition experiment in which compounds are screened for their ability to displace a low affinity ligand (spy) from a receptor, causing its signal to increase due to reduced line-broadening	slow-fast	Ligand binding, drug screening	4.1	344,345,343

[a]This table is not a comprehensive list, but serves as an introductory guide to rapidly identify sections of interest within this review. Further details are available in the text. [b]Fast, intermediate and slow refer to kinetic regimes in which the exchange rate (Kex) is >, ~ or < relative to the chemical shift difference in Hz between two exchanging states (Δω). For ligand-based methods, the time scale refers to the on/off kinetics. Prime time scale refers to the time scale best sensed by a specific method listed here. [c]Most of the relaxation experiments listed in this section apply to either ^{15}N or ^{13}C nuclei. [d]T_2 is also sensitive to dynamics in the µs-ms (intermediate) time scale. [e]The methods above are protein-based methods as they are based on protein signal detection.

sites result from conformational changes driven by binding of allosteric effectors. Sites subject to such conformational changes are here referred to as allosteric "hot spots". An effective approach to compare structures for the purpose of mapping allosteric hot spots typically starts with the identification of a structural moiety that remains invariant upon ligand binding (Fig. 2, red shape). Superimposition of such invariant moiety facilitates the apo *vs.* holo comparison and the identification of the local conformations accessible to the structural elements affected by ligand binding (Fig. 2, green rectangles). Another simple approach to identify sites subject to local conformational changes is the computation of apo *vs.* holo local (*e.g.* three-residue) root-mean-square-deviation (RMSD) profiles[115]. However, it is advisable to take advantage also of more advanced algorithms developed to identify hinge axes and bending sites that describe the transition between two given conformations in terms of 'rigid body' movements of "dynamic" domains. The latter are defined as compact structural fragments that move as coherent units[116,117]. The DynDom software serves this purpose and is available online at http://fizz.cmp.uea.ac.uk/dyndom/[116–118].

Once the accessible conformations for each structural element or "dynamic" domain are identified, a useful exercise is to build hypothetical "mixed" intermediate models in which a subset of structural elements retains the apo structure, while another subset of structural elements adopts the holo structure (Fig. 2, bottom panel). Steric clashes within these mixed models often reveal non-trivial allosteric sites, which are not subject to apparent changes in local conformation and hence would have escaped analyses based uniquely on the apo *vs.* holo structural comparisons. For example, in the cAMP-binding domain (CBD) of the exchange protein directly activated by cAMP (EPAC), a β-barrel was identified as an invariant sub-domain not significantly affected by cAMP-binding, whereas both a cAMP phosphate binding cassette (PBC) and an hinge helix adjacent to it were affected by cAMP[115,119]. The apo *vs.* holo superimposition of the β-barrel revealed that the apo PBC sterically clashed against the adjacent helix in the holo conformation, revealing a key but otherwise non-obvious allosteric site, referred to as 'hydrophobic hinge' that controls the cAMP-dependent hinge helix[119]. Since the rotation of the hinge helix controls access to the EPAC catalytic site, the hypothesis on the 'hydrophobic hinge' could be tested by enzymatic assays and mutations that modulate the extent of steric hindrance between the PBC and the hinge helix[120,121].

Considering that hinge rotations of helices are a recurring motif in allosteric transitions, this example illustrates the power of even relatively simple apo *vs.* holo structural comparative analyses, including the evaluation of steric-hindrance in "mixed" hypothetical intermediates. In this respect, it is critical to note that when steric-clashes are evaluated using NMR structures, it is imperative to

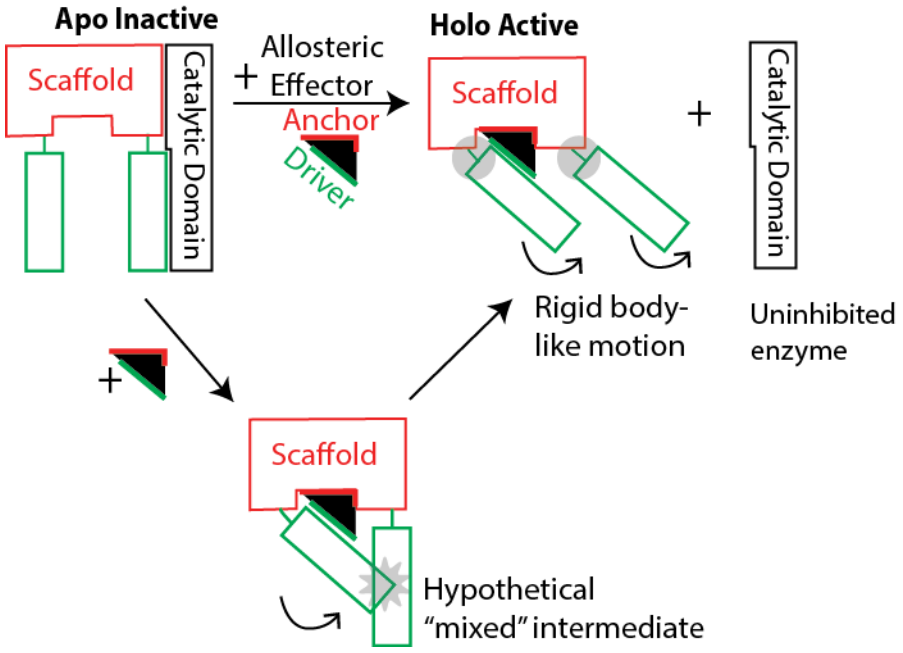

Figure 2. *Schematic illustration of an allosteric structural transition recapitulated in terms of rigid body-like motions of "dynamic domains" (green)[118] relative to a structurally invariant scaffold (red).* The case illustrated here pertains to a conformational switch that controls the inhibition of a functional domain or subunit (*e.g.* enzyme; black). Within the allosteric effector ligand that drives the transition of the conformational switch, "anchor" (red) and "driver" (green) regions are identified. The former interact with the receptor scaffold, while the latter interact with the receptor's dynamic domains[3]. Regions subject to local conformational changes in the apo-inactive *vs.* holo-active transition (*e.g.* hinges for rigid-body rotations) are highlighted by grey circles. Regions affected by steric clashes in hypothetical mixed intermediates, which include elements of both the initial and final states, are marked by a grey star. Both types of sites (*i.e.* grey circles and stars) may include residues that are indispensable for the allosteric transition and represent loci to be tested through mutations and functional assays. Deletion of side chains at these sites may lead to constitutionally active or inactive variants of the allosteric receptor.

extend such evaluation to all the NMR-structures in the ensemble to verify that steric-hindrance is a conserved feature recurrent in most frames of the NMR ensemble. In addition, it is essential to take into consideration the flexibility of the regions involved in steric-clashes, as dynamics may effectively reduce steric hindrance. Overall, it is essential to revisit comparative structural analyses in light of the atomic-resolution profiles of dynamics provided by NMR. For this purpose, it is ideal to complement X-ray crystallography with NMR spectroscopy.

1.4.1.2 Complementarity between X-Ray Crystallography and NMR Spectroscopy

The apo *vs.* holo structural comparison is informative only when the resolution is sufficient to resolve functionally relevant apo *vs.* holo differences. In this respect, the availability of high-resolution structures solved, for example, by X-ray crystallography is invaluable. However, the comparative analysis of crystal structures should take into account crystal packing and the undesired

co-crystallization of buffer ions or additives, which may bias the structure of dynamic proteins, such as those typically involved in allostery[2,122]. When crystal packing remains unaccounted for, it may lead to misinterpretations of conformational changes[42,123]. For example, a system that benefitted from the complementarity between X-ray crystallography and NMR structural determination methods is the hyperpolarization-activated cAMP-gated (HCN) ion channel, which controls cardiac and neuronal excitability[124,125]. Crystal structures of both the apo and cAMP-bound states were solved previously[126,127], but revealed only subtle differences (RMSD ~ 0.5 Å). However, when Akimoto et al. solved the apo solution structure of a monomeric HCN CBD by NMR using a combination of chemical shifts, RDCs, NOEs, PREs and homology restraints, a large reorientation of the α-helical domains relative to the cAMP-bound structure was found[123]. This observation indicates that crystal packing forced the apo crystal structure to adopt a tetrameric conformation, which is more reflective of the cAMP-bound active state. In addition, a bromine ion co-crystallized within the cAMP-binding pocket mimicked the effects of cAMP. The apo CBD structure solved in solution helped understand the mechanism of cAMP-regulated channel tetramerization by illustrating that the conformational changes experienced upon cAMP binding relieve potential steric clashes that would otherwise occur between the rigid core of the CBD and the α-helices that comprise the tetramerization domain[123]. This example illustrates the importance of taking into account crystal packing and dynamics in comparative structural analyses. Dynamics is especially relevant when evaluating steric clashes in hybrid models, in which active and inactive sub-components of the systems are superimposed.

Another reason why it is critical to complement crystallography and NMR spectroscopy is that electron density is often missing in intrinsically flexible regions such as loops, linkers or tail regions, which often play important roles in regulating protein function[42,128–133]. For example, the regulatory subunit of protein kinase A (PKA-R) includes an N-terminal flexible linker, which binds and inhibits the catalytic subunit[134,135]. Parts of this dynamic linker lack electron density in several crystal structures and linker truncation is occasionally necessary prior to structure determination to facilitate crystallization[136,137]. However, NMR spectroscopy revealed that the linker dynamically samples conformations that preferentially interact with the active state of an adjacent cAMP-binding domain (CBD) and hence the linker is an integral component of the allosteric network controlling PKA-R[42]. Given the dynamic nature of linker regions, it is essential to complement crystallographic investigations of allosteric systems with NMR spectroscopy. Another excellent illustration of the synergies between crystallography and NMR is the work of Harada et al.[138] on

heme oxygenase-1 (HO-1). HO-1 catalyzes the degradation of heme, which is toxic to cells in its free form. However, the underlying mechanism is not fully understood as X-ray structures lack electron density in the vicinity of the heme ligation site. Using NMR chemical shifts and relaxation experiments, Harada et al. found that the heme binding site fluctuates in concert with a surface exposed loop, transiently sampling a partially unfolded state[138]. Interestingly, mutations in the surface exposed loop that affects its dynamics but not its structure changed the enzymatic activity of HO-1[138]. Overall, these examples illustrate that although X-ray crystallography remains an invaluable tool for structural allostery, the dynamic nature of allosteric systems requires that crystallography be complemented by NMR spectroscopy.

An additional limitation of the apo *vs.* holo structural comparative analyses that warrants the combination of crystallography and NMR spectroscopy arises when conformational changes relevant for allosteric transitions are subtle[72,115,139]. Such functionally significant changes may fall below the resolution of conventional structure determination methods, thus passing unnoticed. For example, oligomeric regulatory enzymes frequently rely on slight rigid body rotations of subunits by only ~5° to allosterically control function[27]. Hence, it is critical to complement structural comparisons with other types of comparative analyses, which focus on observables that are highly sensitive to even minimal conformational variations. The NMR chemical shifts are well suited for this purpose, as they are among the most accurate and precise reporters of variations in the local chemical and spatial environment[140]. Hence, NMR chemical shifts are excellent atomic-resolution sensors of structural and dynamic changes in a wide range of applications. They can be used to measure affinities for rapid binding processes, map binding sites and otherwise elusive allosteric perturbations as well as predict protein structures[140–146]. The simplest and most common approach involves the apo *vs.* holo comparative analysis of chemical shifts. This type of chemical shift perturbation (CSP) mapping, is commonly utilized to identify binding sites and interfaces as well as other sites affected by binding perturbations[27,112,113,130,147–157]. More advanced approaches to chemical shift mapping are discussed below (Section 2.2.2.). Overall, X-ray crystallography and NMR spectroscopy are highly complementary[158,159] and both approaches are needed for understanding allostery. However, even when combined, X-ray crystallography and NMR spectroscopy may not be sufficient for complex allosteric systems, such as integral multi-domain proteins, and additional structural methods must be integrated in the structural determination process.

1.4.1.3 Integrative Structural Approaches for Multi-Domain Systems

Allosteric regulation of signaling pathways often relies on multi-domain proteins[160], in which ligand-dependent domain re-orientations control transitions between 'open' and 'closed' topologies where active sites are accessible or occluded, respectively. Hence, the comparative apo *vs.* holo structural analysis of multi-domain proteins is of central relevance to allostery. However, given their complexity, structure determination of multi-domain systems frequently requires the combination of high and low-resolution techniques. The former, such as crystallography and NOE or chemical shift-based NMR methods are suitable to solve structures of single domains or subunits, while the latter, such as small and wide angle X-ray scattering (SAXS & WAXS respectively), small angle neutron scattering (SANS) and rigid-body fitted residual dipolar couplings (RDCs), are necessary to reconstruct the overall inter-domain topology. The integration of high and low-resolution approaches results in reliable structural models of full-length multi-domain proteins and is well illustrated by the work of Schwieters et al.[161], who elucidated the full length structures of the *E. coli* Enzyme I (EI) in its ligand free and histidine phosphocarrier protein (HPr)-bound states[161].

EI is a component of the phosphotransferase system, which catalyzes the transfer of a phosphate group from phosphoenolpyruvate (PEP) to H189 in EI and subsequently to H15 in HPr. The distance between the PEP substrate and H189 is controlled by the position of the N-terminal phosphoryl transfer domain of EI (EIN), where H189 is located, relative to the C-terminal dimerization domain of EI (EIC), where PEP binds. The EIN-EIC topology was determined using conjoined rigid body/torsional angle/Cartesian simulated annealing with experimental restraints from residual dipolar couplings (RDCs) and SAXS/WAXS and SANS data. RDCs defined the relative orientation of the EIN and EIC domains, although with translational degeneracy, while SAXS/WAXS provided details about the overall shape and size of the molecule, including the solvent layer, which was validated by independent SANS measurements. A clustering algorithm was utilized to find converged structures and those with the best fit to the RDC and SAXS/WAXS data and minimal total energy were retained for analysis. Based on this approach, Schweiters et al. were able to propose a model in which, following phosphoryl transfer from PEP to H189 in a closed EIN-EIC topology, HPr binds to the EIN domain, the pyruvate product is released and the protein adopts an open topology wherein phosphoryl transfer occurs between EI and HPr. Another example illustrating the importance of integrating NMR and SAXS studies is the work of Skora et al. on the multi-domain c-Abl kinase[162]. Skora *et al*. combined SAXS and RDC orientation tensor measurements to differentiate between open (active) *vs.* closed (inactive) domain topologies, which

differ in the proximity of the SH and kinase domains (Fig. 3)[162]. Furthermore, Skora et al. showed that domain-averaged ^{15}N T_1 and T_2 measurements are very effective at capturing enhanced flexibility in the individual domains upon transition to the open conformation[162] (Fig. 3). Overall, SAXS, SANS, RDCs and spin-relaxation measurements are effective means of probing inter-domain orientations[162–165].

The previous examples are excellent illustrations of how scattering techniques (SAXS, WAXS and SANS), in conjunction with NMR, are instrumental in reconstructing the overall topology of multi-domain proteins. However, due to their inherently low resolution, scattering approaches alone cannot provide detailed maps of inter-domain interfaces. Hence, when such interfaces are not accessible by traditional structure determination methods, alternative approaches must be sought. An effective method to map inter-domain interfaces relies on the integration of homology modeling and chemical shift changes induced by soft-mutations. This approach is well illustrated by the work of Zhuravleva et al.[25], who used a series of "soft" mutations to test domain interfaces predicted through homology modeling for DnaK, a multi-domain molecular chaperone. Soft mutations are mutations that are sufficiently conservative to minimize long-range effects, but sizeable enough to create local perturbations resulting in detectable chemical shift changes in proximal residues. Examples of "soft mutations" are L to I, M to I, E to D and D to N. Chemical shifts induced by such mutations provided a critical test of domain interfaces in homology models[25].

Chemical shift changes also provide a means to probe the overall domain organization. The analysis of chemical shift changes induced by domain deletion mutants is a simple but effective approach to test hypothetical "domain-resolution" alternative models of binding or quaternary structure[156,166]. For example, by comparing Met methyl TROSY NMR spectra of full length p53 and of its isolated core and tetramerization domains, Bista et al. could differentiate between different proposed quaternary structures and models of DNA binding[166] . Since the folded domains of p53 are connected and flanked by intrinsically unstructured regions, probing inter-domain interactions in p53 through more conventional structural determination methods would have been challenging. However, chemical shifts sense not only domain interfaces, but also longer-range perturbations occurring upon inter-domain interactions. Hence, chemical shifts based methods to map inter-domain interfaces are often complemented by other means to probe more directly inter-domain interfaces, which include paramagnetic relaxation enhancements (PREs)[167–169] and pseudo-contact chemical shifts (PCSs)[170]. Both PREs and PCSs report on larger distances compared to NOEs (*i.e.* <~35 Å for the spin-label for PREs[171] and <~40 Å from the metal for PCSs[172]) and hence

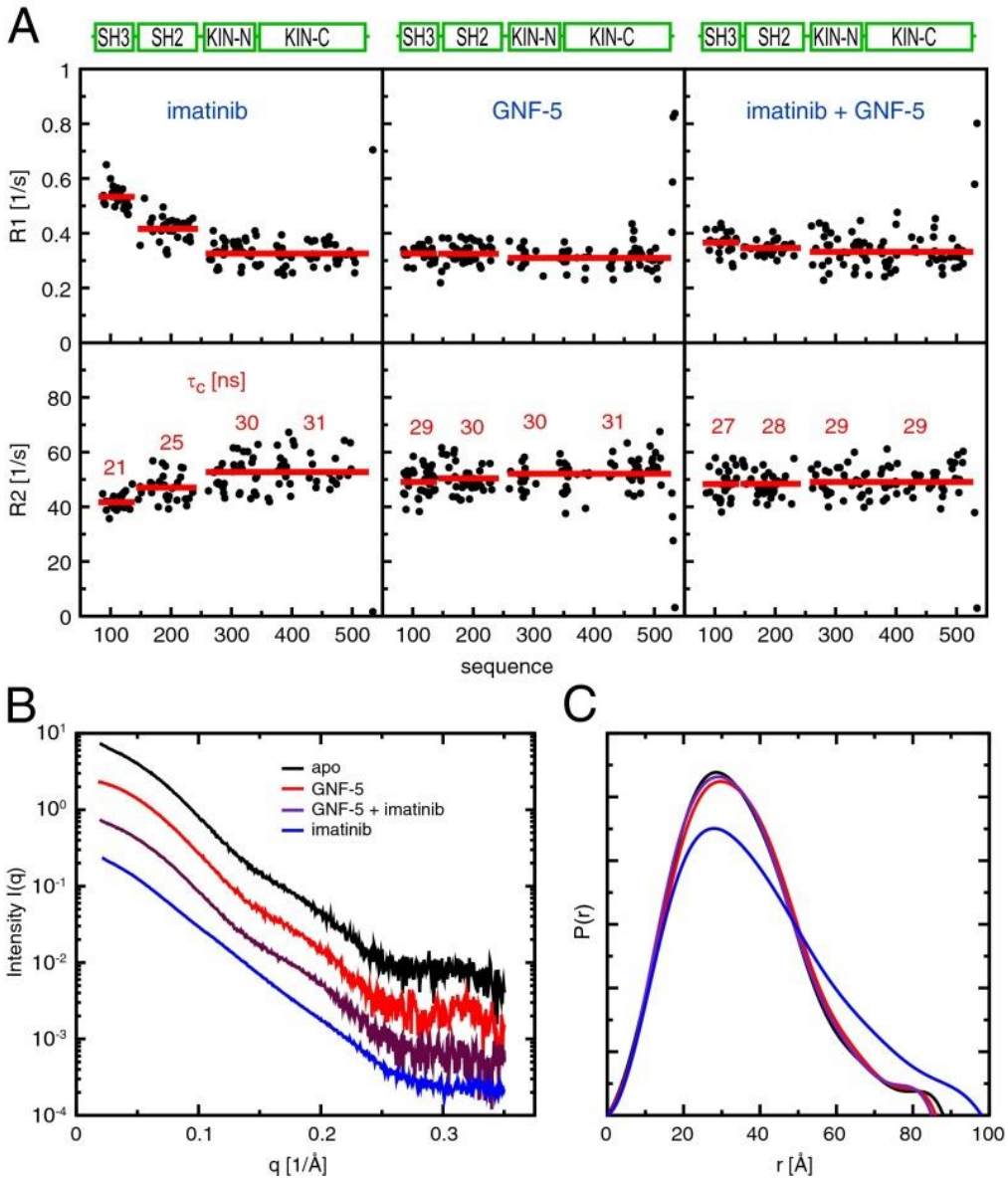

Figure 3. *Mapping changes in inter-domain topology by combining NMR ^{15}N relaxation (R_1 and R_2) and SAXS*. The data presented here were acquired for the c-Abl kinase bound to inhibitors that bind at distinct sites (*i.e.* Imatinib and/or GNF-5). In the absence of GNF-5 an "open" topology is observed with the two N-terminal domains tumbling faster than the two C-terminal domains, as supported by their higher (lower) ^{15}N R_1 (R_2) relaxation rates **(a)**. In the presence of GNF-5 a "closed" topology is detected in which the two N-terminal domains tumble together with the two C-terminal domains, as supported by more uniform ^{15}N R_1 and R_2 relaxation rates **(a)**. Domain average relaxation rates are marked by red lines in panel **(a)**. Panels **(b)** and **(c)** depict the SAXS intensity plots, which were offset for the sake of clarity, and the corresponding distance distributions, respectively, confirming the different inter-domain topology adopted when only Imatinib is bound. Reproduced with permission from Skora et al.[162]. Copyright 2013 National Academy of Sciences, U.S.A.

they are suitable to detect contacts between domains.[173,174] In conclusion, the elucidation of the structural basis of allosteric regulation relies on integrative approaches spanning a wide range of resolutions, from low-resolution methods (*e.g.* SAXS, WAXS, SANS, domain deletion) to medium-resolution techniques (*e.g.* PREs) to high-resolution classical structure determination of the component domains. In the context of this resolution spectrum, it is worth noting the "resolution revolution" currently taking place in cryo-electron microscopy (EM).

Recent developments in direct electron detector cameras for cryo-EM have raised the potential of this technique for determining the structure of macromolecules at atomic resolution[175–177]. For example, recently Bartesaghi et al. solved the structure of β-galactosidase in complex with a small molecule to a resolution of 2.2 Å[178]. In addition, several other protein structures have been solved within the past few years with resolutions approaching those of X-ray crystallography[176,179–183]. These include the 3.4 Å structure of human γ-secretase[179], the 3.5 and 3.6 Å models of the ribosome[180], and the 3.4 Å structure of the ion channel TRPV1[181], just to name a few. The advantages of cryo-EM over X-ray crystallography are that protein crystallization is not required and that low quantities (micrograms) of sample are often sufficient for structure elucidation. These advances presently place cryo-EM as an atomic resolution structural biology technique that opens new opportunities for determining the structure of macromolecules that are difficult to obtain in large quantities or which are refractory to crystallization. In addition, in cryo-EM the specimen is flash frozen in its own physiological buffer, maintained in a hydrated state and imaged using minimal electron dose protocols. Therefore, the elucidated structure is free of artifacts induced by crystal packing or radiation damage[175,176]. Solution NMR spectroscopy still permits structural elucidation under the most native-like conditions of the three, without minimum molecular weight limits[179–181,184].

1.4.2 Comparative Dynamical Analyses

Changes in the ground-state structure alone are not always sufficient to explain how a ligand allosterically controls interactions at remote sites[60]. The apo form of an allosteric system is inherently dynamic due to the presence of auto-inhibitory equilibria (Fig. 1) and, even after binding of allosteric effectors, residual dynamics may remain because the auto-inhibitory equilibria are still present even in the holo form (Fig. 1)[185]. In addition, binding may result in both formation of new interactions and loss or strain of pre-existing interactions, leading to modulations of dynamics[186–189] that are key determinants of the entropic component of the free energy and also of the kinetics of binding and catalysis[65,66,72,82,84–86,190–192]. Hence, dynamics is frequently correlated with biological

function. In order to elucidate the mechanisms underlying dynamic-function correlations, it is critical to comparatively analyze the atomic-resolution dynamic profiles of allosteric systems at different levels of functional activation, starting from the apo and holo forms. The approaches needed to obtain such dynamic maps depend to a large extent on the time-scale for the active *vs.* inactive exchange process relative to the evolution time-scale of the NMR chemical shifts (Fig. 1). In this respect, it should be noted that the time-scale for the active *vs.* inactive exchange does not necessarily coincide with that for the apo *vs.* holo exchange. The former can be faster than the latter[42] or vice versa[185,193] and of course the latter depends also on the ligand concentration.

The apo-inactive *vs.* apo-active exchange can occur within a wide window of time-scales ranging from the slow (*e.g.* ≥ ms)[7,193,194], the intermediate (*e.g.* ms-μs)[100] to the fast exchange (*e.g.* ≤ μs)[42] regimes. When slow, distinct sets of resonances for the apo-inactive and apo-active states may be directly observed in the NMR spectra and the apo-inactive *vs.* apo-active exchange rate can be quantified through N_z or C_z exchange experiments[195–199]. When the apo-inactive *vs.* apo-active exchange rate starts to approach the intermediate regime, the apo-active state is often not directly observable, but is efficiently probed through NMR relaxation dispersion (NMRD) methods suitable for the ~ms - sub-ms time-scale window and chemical-exchange saturation transfer (CEST) techniques ideal for the longer ~tens of ms - ms time-scale window. The NMRD and CEST approaches are ideally suited to probe 'invisible' excited states, providing both the kinetics (*i.e.* rates) and the thermodynamics (*i.e.* populations) of the ground to excited state transition, as reviewed in depth elsewhere[37,38,86,200,201]. In addition, based on the NMRD and CEST family of experiments it is possible to infer also the structure of the excited state[37,38,200–206].

One of the experimental approaches to probe the structure of the excited states through NMRD or CEST is by measuring ^{15}N or ^{13}C chemical shifts, which are sufficient to critically test the hypothesis that apo ensembles pre-sample the bound conformation in the slow-intermediate time-scale. For this purpose, the ground *vs.* excited state chemical shift changes are compared to the apo *vs.* holo chemical shift changes (excluding residues at or near the binding site) to establish whether the apo excited state resembles the holo-active state[207–209]. Similar approaches are applicable to the dynamics of holo samples. An example of how the residual dynamics of a ligand-bound state of a protein underlie heterotropic allosteric coupling is provided by siderocalin Q83, which binds two ligands, *i.e.* enterobactin and arachidonic acid[210]. Coudevylle et al. compared apo *vs.* holo and ground *vs.* excited state chemical shift changes determined through NMR monitored titrations and NMRD experiments, respectively, and showed that when enterobactin binds Q83, the arachidonic

acid binding site samples an excited state similar the arachidonic acid-bound state, explaining the positive cooperativity observed between the binding of the two ligands to Q83[210]. One experimental challenge when probing intermediate exchange is that the line-broadening intrinsic to this exchange regime reduces signal-to-noise ratios. However, even in the limit of broadening beyond detection it is still possible to obtain functionally relevant information. For example, Larion et al. constructively utilized line-broadening beyond detection in Ile Methyl and Trp TROSY spectra to prove that a domain in the apo form of human glucokinase is subject to ~ms dynamics, which is rate-determining for catalytic turnover at low glucose concentrations[85].

When the apo-inactive *vs.* apo-active exchange rate falls in the fast exchange regime, the observed chemical shifts are population-weighted averages of the ppm values for the inactive and active conformations. Hence, under the assumption of a two-state exchange, a linear pattern is expected for the peak positions of samples in which state populations are modulated through ligand binding and/or mutations[115,211]. Since the chemical shifts directly encode state populations, by comparing the apo chemical shifts to those of samples in which the active and inactive states are selectively stabilized, it is possible to gauge the state populations, which, assuming a Boltzmann distribution, reflect the depth of the respective free energy basins (Fig. 1c-d)[42,43,65]. While in principle this approach provides an accurate and precise read-out of state populations, in practice two major caveats should be taken into consideration. First, isolation of pure states is not always feasible. Even when the active and inactive states are stabilized through addition of saturating amounts of endogenous effectors and allosteric inhibitors, respectively, the apo inhibitory equilibrium is only modulated by the ratio of the state-specific dissociation constants. In this respect, it is useful to complement ligand-based state-stabilization with mutations designed to trap the inactive or active states, *i.e.* constitutively active/inactive mutants, including molecular stapling through cross-linking[47]. Nevertheless, these approaches often provide only approximations, as stabilization of pure states frequently remains a challenge even after combining ligand and mutational perturbations. When pure reference states are not available, only relative populations are extracted from chemical shifts. However, if residual ms-µs dynamics is observed, relaxation dispersion experiments provide a means to determine the populations of the residual state, opening the possibility of obtaining absolute populations[43,61].

The second caveat is that the perturbations utilized to modulate the position of the inhibitory equilibrium also result in nearest neighbour effects and possibly in other distortions of the apo-active and apo-inactive conformations. These may lead to residue-specific deviations from linearity

in the relative peak positions and in general to residue-specific changes in the relative peak positions, resulting in a distribution of state populations as opposed to a single well defined population[123]. Hence, it is advisable to extend the comparative chemical shift analyses to more than one residue. A systematic approach to extend comparative chemical shift analyses of three samples to all assigned residues is the CHEmical Shift Projection Analysis (CHESPA; Fig. 4a)[27,120,121]. A useful complement to the projection analysis of chemical shifts is provided by chemical shift correlations (Fig. 4b)[212]. These types of three-fold chemical shift comparative analyses are an example of a more general experimental strategy to interrogate allosteric systems through libraries of perturbations with the goal of establishing dynamics-function relationships, which is the focus of the following section.

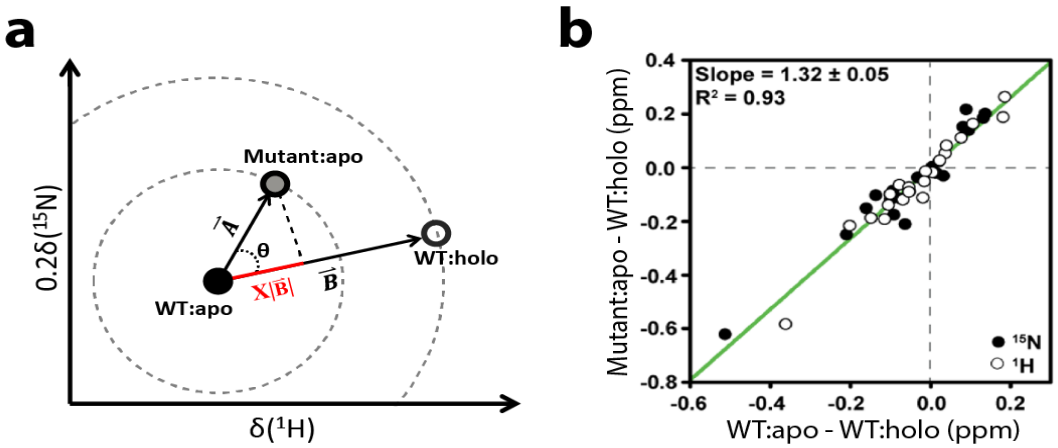

Figure 4: *Simple Three-Sample Chemical Shift Perturbation Analyses*. **(a)** Schematic diagram of the CHEmical Shift Projection Analysis (CHESPA)[121]. The CHESPA analysis is a vectorial comparative chemical shift analysis between three samples, two of which serve as reference (*e.g.* WT:apo *vs.* WT:holo). The third sample represents a unique perturbation (*e.g.* mutant:apo), which is evaluated relative to the reference vector. The scaled chemical shifts of the three samples defined two vectors as shown in panel (a), which are used to assess the direction (θ) and relative magnitude (X) of the chemical shift change caused by the perturbation. The angle (θ) reveals whether the perturbation is linear, as expected for a fast-exchanging, two-state equilibrium, or non-linear, as expected in the presence of more than two conformers or nearest neighbor effects (*e.g.* ligand binding interactions, localized structural distortions, etc.). Meanwhile, the fractional activation (X) reports on the relative populations of the two reference states in the third (perturbed) sample. **(b)** Representative example of a chemical shift correlation plot[47,212]. Similar to the CHESPA analysis, chemical shift correlation plots require three samples, two that serve as reference and a third that represents a distinct perturbation. The slope of the correlation reports on the relative population of conformers in the perturbed sample. For instance, if the chemical shift changes between the reference states (e.g. WT:apo – WT:holo) are plotted along the x-axis and the chemical shift differences between the perturbed state and one of the reference states are plotted along the y-axis (e.g. Mutant:apo – WT:holo) the slope of the linear correlation represents the ratio of the inactive populations in the mutant *vs.* WT apo forms [212]. The CHESPA analysis is useful to pre-screen residues that report primarily on the two-state inactive *vs.* active equilibrium, thus increasing the linearity of correlation plots, such as that shown in panel (b). Panel (b) was adapted from Moleschi et al.[47]. Copyright 2015 American Chemical Society.

1.4.2.1 Interrogating Allosteric Systems through Perturbation Libraries

Understanding how dynamics relates to biological function in allosteric systems requires the comparative analysis of dynamic profiles mapped at different stages of functional modulation[27,61,72,89,115,128]. While apo *vs.* holo comparisons are a necessary starting point, they are not sufficient to establish dynamics – function correlations and to reliably separate binding *vs.* allosteric contributions. Hence, it is crucial to interrogate allosteric biomolecules through a library of perturbations designed to modulate both dynamics and biological function through shifts in the position of inhibitory equilibria. Perturbations typically include analogs of the endogenous allosteric effector ligand as well as post-translational modifications or mutations[15,115,120,213–215]. The design and selection of the perturbation library is critical to successfully correlate dynamics to function and it should be carefully tailored to the target allosteric system and question at hand.

The perturbation set utilized to interrogate a given allosteric system typically meets four key requirements (Fig. 5a): (a) All perturbations are functionally profiled through bioassays (*e.g.* enzymatic assays); (b) The perturbation library spans a wide-range of activities; (c) Perturbations are minimally invasive to minimize distortions of the system under investigation; (d) Perturbations are spatially clustered in the same region of the structure to minimize spreading of nearest-neighbour effects[115]. The design of perturbation sets that meet conditions (a-d) is often informed by structural analyses (Section 2.1). For example, covalent modifications (Fig. 5b) may target non-covalent interactions between the allosteric effector and the allosteric acceptor (Fig. 5c)[115]. When diverse covalent changes affect "driver" groups of the endogenous allosteric effector, as defined in Section 2.3.2.[3], it is likely that the resulting ligand library includes the full spectrum of reverse-agonists, antagonists, partial agonists and perhaps even super-agonists, reflecting a full modulation of the holo-inactive *vs.* holo-active equilibrium[3,115,216]. In addition, fragment-based ligand libraries may also provide informative perturbation sets[217].

Similarly to the design of ligand libraries, mutation libraries are typically designed to perturb interfaces where allosteric effectors bind. However, allosteric binding sites are not the only loci where mutations are likely to result in maximal functional changes with minimal structural perturbations (*i.e.* conditions (b) and (c) above). Other loci are regions subject to local conformational variations and/or to steric hindrance (grey sites in Fig. 2). Especially, hinge segments are likely to be sensitive to relatively subtle mutations[218], such as "soft" mutations, which decrease the risk of excessive unfolding[73,128]. When the allosteric regions of interest are located at the N- or C-termini of a protein construct, deletion mutations are another effective tool to generate

perturbation sets. For example, progressive N- or C-terminal deletions are useful to mimic partial unfolding[120,219]. Conversely, secondary structure-promoting mutations may increase the population of states in which a given secondary structure element is selectively stabilized[220]. Besides, point and deletion mutations, another source of informative allosteric modulation is provided by cross-linking of residue pairs whose distance varies during the allosteric transition[47,221].

When designing perturbation sets with the goal of establishing dynamics – function relationships, it is not always trivial to predict *a priori* how a given mutation or ligand modification alters structure and dynamics. For instance, it is important to consider that evolutionary conservation of a given residue does not necessarily imply that the residue is critical to allosterically control function, while it is possible that mutations of non-conserved residues lead to dramatic modulations of function[75]. Hence, perturbation screening is often a valuable alternative to rationally designing mutations that elicit desired allosteric effects. In this respect, mutant libraries obtained through high-throughput mutagenesis[16] could be another source of mutational perturbations to interrogate allosteric systems. A valuable screening "read-out" tool to select mutations resulting primarily in changes in dynamics rather than in structure is provided by the entropic contribution to binding as measured by isothermal titration calorimetry (ITC)[222]. Overall, the choice of an informative perturbation library is critical for the successful investigation of allosteric sensors and, while the general guidelines above are meant to assist the design of perturbation sets, it is paramount to carefully customize the library to the specific system under investigation.

Once an informative perturbation library is selected, the response of the allosteric system to the perturbation set is mapped at atomic-resolution through NMR experiments, which typically include chemical shift and spin-relaxation measurements (Fig. 5a). The comparative analysis of chemical shifts and spin-relaxation rates assists the identification of functionally relevant dynamic changes and the underlying network of allosteric sites (Fig. 5a), as outlined in greater detail in the following sections. Based on the map of functional dynamics and allosteric sites, it is often possible to design a "second-generation" library of perturbations to further interrogate the allosteric system of interest in a more targeted manner (Fig. 5a).

1.4.2.2 Mapping Allosteric Sites through Comparative Chemical Shift Analyses

An excellent illustration of the allosteric insight provided by comparative chemical shift analysis of a set of perturbations is the investigation of the 20S proteasome core particle (CP) by Ruschak and Kay[27]. The 20S CP consists of four heptameric protein rings in which activators bind ~75 Å away from the sites that catalyze proteolysis. Given the subtle nature of the underlying

conformational changes, the molecular basis for this long-range allosteric cross-talk had remained elusive to classical structure determination approaches and even to comparative RDC analyses[27,223]. However, significant chemical shift changes were observed between apo and activator bound 20S CP[27]. By further modulating the regulatory equilibria through either an allosteric inhibitor or mutations at the allosteric or active sites, Ruschak and Kay identified multiple linear chemical shift changes that arise from shifts in the populations of interchanging conformers linked to different distributions of proteolytic cleavage products[27]. The most significant chemical shift changes are localized at subunit interfaces, consistent with the interconverting conformers differing for slight subunit rotations. Furthermore, the differences between the conformers, as revealed by chemical shift changes, cluster in contiguous structural regions, which define an allosteric network connecting the allosteric and active sites[27]. Critical to this milestone study of allostery was the use of methyl TROSY (HMQC), which was essential given the high MW of the system (MW ~670 kDa and 1 MDa without and with activator, respectively). The ^{1}H-^{13}C HMQC spectra of the ILV or M methyls in the 20S CP were assigned through a combination of mutations, NOEs and triple-resonance spectra acquired for isolated subunits conjugated to an N-terminal D_9 solubility tag[27].

Chemical shift changes were instrumental also for mapping allosteric pathways in another proteasome, the 230 kDa HslU-HslV complex[224]. Single point mutations were engineered around the HslU binding site or at locations subject to conformational change upon HslU binding and the resulting chemical shift changes were measured by methyl-TROSY. The HslV variants exhibiting the largest chemical shift changes throughout the protein also commonly exhibited the greatest effect on the catalytic efficiency. In addition, complementary mutations of residues at opposite positions of the symmetry axis, revealed that the allosteric pathway is unidirectional. These seminal examples illustrate the effectiveness of chemical shift changes in mapping allosteric pathways in solution. Similar chemical shift changes are equally useful in solid-state NMR experiments and are invaluable to elucidate the mechanism underlying trans-membrane allosteric coupling[225]. For example Wylie et al. assigned the backbone and side-chain resonances of the full-length 160 amino-acid bacterial potassium channel KcsA through a combination of 3D (NCACX, NCOCX, CAN(CO)CX) and 4D (CANCOCA) MAS-NMR experiments[225]. Analysis of chemical shift changes revealed that removal of K^+ ions results in perturbations that reach the pH gate at the opposite side of the membrane, pointing to the presence of trans-membrane allosteric coupling[225]. Overall, the previous examples illustrate the power of chemical shift changes in mapping long-range effects of perturbations. However, perturbations typically elicit both long- and short-range

effects, *i.e.* they not only modulate the populations of pre-existing conformers, but they also result in nearest-neighbour effects (NNEs) and other distortions, which may propagate beyond the immediate binding pocket. An approach that has been proposed to separate such effects relies on the CHEmical Shift Covariance Analysis (CHESCA)[115].

1.4.2.3 CHEmical Shift Covariance Analysis (CHESCA)

CHESCA is an NMR-based method for dissecting the contribution of each residue to binding *vs.* allostery[115,120,226]. In the context of CHESCA, the response of the allosteric system to the perturbation library is modeled considering that the chemical shifts of residue pairs that report on the same two-state fast-exchanging equilibrium are linearly correlated in the absence of significant nearest neighbour effects (Fig. 5d,e). The pairwise residue correlations are then mapped into a correlation matrix (Fig. 5f,g). If all residues conform to an ideal two-state model, correlations are expected for all residue-pairs subject to significant chemical shift variations, with the only source of de-correlation being experimental noise. However, in practice other sources of de-correlations are observed, including nearest neighbour effects and in general deviations from the two-state model. As a result, the chemical shift correlation matrix is often sparser than predicted by a two-state model[42,227] and subsets of tightly correlated residues exhibiting concerted responses to the perturbation library are identified through hierarchical clustering algorithms[115,226]. In addition, an independent singular value decomposition (SVD) analysis is performed to validate the allosteric clusters determined from the analysis of the chemical shift correlation matrix[115]. Selvaratnam et al. applied this method to the exchange protein activated by cAMP (EPAC), revealing a previously elusive allosteric network that originates at the conserved cAMP binding cassette and extends to remote allosteric loci, including those controlled through modulations of dynamics. Furthermore, mutations at the EPAC allosteric sites identified by the CHESCA result in modulations of catalytic activity[115].

Several other creative implementations of the CHESCA have been developed and utilized to investigate not only allosteric regulation[228], but also ligand binding[229], enzyme catalysis[227,230,231] and flexible linkers[42]. For example, Kim et al. showed that CHESCA together with a method based on a residue-specific principal component analysis (PCA) of chemical shifts[232] is useful to monitor the distribution of open and closed topologies in the catalytic subunit of PKA[231]. Another clever application of CHESCA to understand the allosteric control of enzymatic catalysis is the work of Axe et al. on tryptophan synthase (TS), which catalyzes the conversion of indole-3-glycerol phosphate to glyceraldehyde-3-phosphate (G3P) and indole[227,230]. Axe et al. utilized a library of

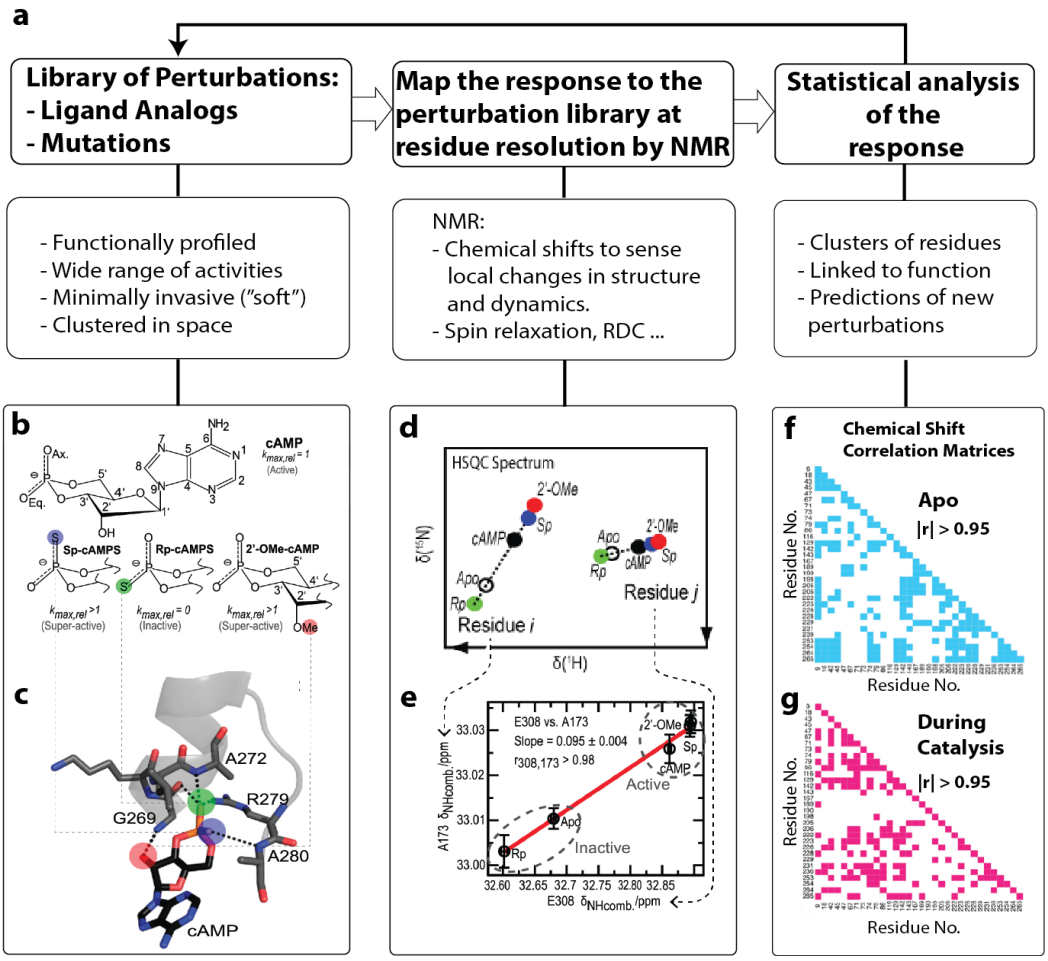

Figure 5. *Interrogating allosteric systems through perturbation libraries.* **(a)** General scheme to probe allostery via libraries of perturbations. The first step is the selection of an informative perturbation set of either mutations and/or analogs of the allosteric effector. An example of an analog library is shown in panel **(b)** for the cAMP allosteric effector. Other libraries include cyclic nucleotides with different bases (*e.g.* cIMP). The covalent modifications of cAMP selectively perturb the non-covalent interactions that anchor cAMP to its receptor, as illustrated in **(c)**. The second step focuses on mapping the response of the allosteric system to the perturbation library. NMR is an excellent tool to obtain such a map at atomic/residue resolution and in particular NMR chemical shifts are exquisite probes of even subtle, but functionally relevant, structural and/or dynamical changes. When the active *vs.* inactive exchange is in the fast regime, linear chemical shift patterns are observed for residues that sense primarily the two-state active *vs.* inactive exchange, as shown in panel **(d)**. Hence, the chemical shifts of residues that sense the same cooperative two-state transitions are linearly correlated **(e)**. Such inter-residue chemical shift correlations are the basis of the CHEmical Shift Covariance Analysis (CHESCA)[115,226]and are systematically recapitulated by the chemical shift correlation matrix reporting the Pearson correlation coefficient for each residue pair. For example, panel **(f)** shows the correlation matrix for apo α-Trp Synthase (αTS) resulting from a library of Ala-to-Gly perturbations[227]. Interestingly, when the same perturbation set was utilized to probe αTS during active turnover, a different correlation pattern was observed **(g)**, suggesting that catalysis requires an allosteric network distinct from that present in the absence of ligands[227]. CCS stands for compounded chemical shift, which is a linear combination of ^{1}H and ^{15}N chemical shifts in ppm. Panels (a-e) were adapted from Selvaratnam et al.[115] (Copyright 2011 National Academy of Sciences, U.S.A) and panels (f) and (g) were reproduced from Axe et al.[227] (Copyright 2013 Elsevier).

Ala-to-Gly mutations and ^{15}N-Ala-selective labeling to compare the CHESCA correlation-matrices measured in the absence of ligands and in the presence of indole and G3P, *i.e.* under conditions of active turnover (Fig. 5f,g)[227]. The comparison of the two CHESCA matrices revealed that the allosteric networks of tryptophan synthase (TS) change during catalysis. Such changes are accompanied by modulations of ps-ns and ms-μs dynamics during catalytic turnover[227]. While the examples above illustrate the scope of CHESCA applications, it is clear that CHESCA is just one of the several tools available to separate binding from allosteric effects and map allosteric networks. For example, binding sites are also reliably and sensitively mapped through cross-saturation and PRE methods[7,233–235], while other complementary experimental methods are available to detect allosteric networks. For example, double- and triple-mutant cycles are classical methods to identify and quantify allosteric couplings[236,237] and in general mutagenesis in combination with functional assays is pivotal to validate the relevance of allosteric networks mapped through biophysical methods[115,120,154].

1.4.2.4 COntact Networks Through Alternate Conformation Transitions (CONTACT)

Another approach to map allosteric networks underlying long-range chemical shift changes was proposed by van den Bedem *et al*. who developed an algorithm called CONTACT (COntact Networks Through Alternate Conformation Transitions) to identify groups of conformationally heterogeneous residues through X-ray crystallographic data[238]. CONTACT initially models heterogeneous features of the X-ray electron density by computing an optimal fit of one to four conformations for each residue. CONTACT then calculates steric clashes between atoms of different residues and repositions residues recursively until such clashes are eliminated. Through this approach, multiple allosteric pathways can be identified. Van den Bedem *et al*. tested the CONTACT algorithm on two well-established systems, cyclophilin A and dihydrofolate reductase (DHFR). For cyclophilin A, a nine residue contact network was identified that connects the hydrophobic core to the active site. A mutation of one residue within this network resulted in chemical shift changes of several other residues spread across the network and caused a reduction of the catalytic activity by 70%. For DHFR, the CONTACT results are in agreement with long-range NMR chemical shift change caused by an allosteric mutation[15]. The comparison of CONTACT pathways derived from cryogenic vs. room temperature crystal structures[239], suggests that the latter sample broader conformational ensembles and thus produce more extensive CONTACT pathways[238]. In addition, based on the multiconformer models obtained from room-temperature crystallography it is possible to compute backbone and side-chain order parameters for

ps-ns motions that are in agreement with the order parameters obtained from NMR relaxation rates[240]. Overall, the CONTACT method[238] provides a promising approach to predict allosteric pathways from X-ray crystallographic data, complementing the CHESCA NMR method. The CONTACT algorithm provides direct structural insight, but the CHESCA NMR approach is applicable to partially unfolded systems that are typically elusive to crystallization. In addition, since CHESCA relies simply on chemical shift measurements, it is potentially applicable to *in cell* NMR as well[241]. Both CHESCA and CONTACT help identify allosteric sites that may not be obvious based on comparative analyses of structures solved through more traditional methods.

1.4.2.5 Other Uses of Chemical Shifts

While the CHESCA approach utilizes chemical shifts as exquisite sensors of variations of the local spatial and chemical environment, it is important to note that chemical shifts also report on structure and dynamics[144–146,242–246]. For example, Berjanskii and Wishart devised an empirical method to quantify backbone and side-chain dynamics through chemical shifts[144,242]. The approach is based on the idea of modeling the order parameters as a function of the differences between experimentally measured chemical shifts and the corresponding values for a random coil ensemble. Each of these differences is weighed by scaling factors that were optimized according to motions determined by a training set of representative MD simulations[144,242]. Another approach to probe side chain dynamics through chemical shifts, relies on ILV methyl ppm values to reveal inter-conversions between trans and gauche (+/-) conformations upon binding[139,247]. Using methyl ^{13}C chemical shifts to determine rotameric conformations of side chains is an effective method to probe dynamics within hydrophobic cores and detect population differences as low as ~5%, which are often elusive to other structural determination methods[139]. While chemical shift-based approaches provide a quick assessment of dynamic profiles over a possibly wide range of time-scales, a more quantitative evaluation of dynamics requires NMR spin-relaxation measurements, as discussed in the following section.

1.4.2.6 Functional Dynamics from Comparative Spin-Relaxation Analyses

Functional dynamics refer to motions that are critical to control biological function. For example, the dynamic switch between ‘closed’ and ‘open’ topologies of multi-domain proteins controls accessibility of active sites, as outlined in Section 2.1.3 on rigid-body-like domain orientation. However, rigid-body motions alone do not fully capture changes in intra-domain dynamics coupled to inter-domain movements. Such variations in intra-domain flexibility are often

critical to explain how allosteric effectors control inter-domain topology and/or how the latter modulates function[248]. Hence, an important class of functional dynamics changes includes modulations of domain dynamics coupled to changes in inter-domain orientation[249]. In addition, as discussed above, biological function can also be modulated in the absence of apparent structural changes, as in the case of dynamically driven allostery. These two important classes of functional dynamic variations are the subjects of the next two sections.

1.4.2.7 Internal Dynamics Coupled to Inter-Domain Orientation

Changes in internal dynamics are frequently coupled to variations in the relative orientation of domains and to biological function. For example, Venditti et al. comparatively analyzed through ^{13}C-Single Quantum (SQ) and ^{1}H-^{13}C-Multiple Quantum (MQ) methyl NMR dispersion measurements the μs-ms dynamics of bacterial Enzyme I (EI) in the apo, inhibitor- and substrate-bound forms[250]. The latter form was stabilized by silencing auto-phosphorylation through a histidine to alanine mutation in the EI active site. Venditti et al. showed that μs-ms dynamics at the C-terminal domain are quenched when the N- and C-terminal domains adopt a closed topology (*i.e.* in the substrate-bound form), but not when the inter-domain topology is open (*i.e.* in the apo and inhibitor-bound form), thus establishing a correlation between changes in intra-domain dynamics and in inter-domain orientation, which controls phosphorylation. Another example illustrating how domain re-orientation is coupled to intra-domain changes in dynamics is provided by the work of Zhuravleva and Gierasch on the Hsp70 chaperone, which shows that contacts between the nucleotide- and the β-substrate binding domains lead to enhanced dynamics at the substrate binding site, explaining how multiple chaperone clients are bound to Hsp70[73].

Internal dynamics is coupled not only to changes in domain topology, but also to re-orientations of sub-domains (or lobes). This is well illustrated by the work of Masterson et al. on the catalytic (C) subunit of protein kinase A, in which a correlation between dynamics and function is effectively established through a library of perturbations[72]. The PKA C subunit was investigated in five different forms: the apo form, the binary complex with nucleotide and three ternary complexes, all including the nucleotide and either the phospholamban substrate or the PKI inhibitor with or without excess Mg^{2+}. To stabilize the substrate-bound sample, phosphorylation was inhibited by using a non-hydrolysable analog of ATP (AMP-PNP). The five forms of PKA-C investigated by Masterson et al. exhibit different ps-ns an ms-μs dynamic profiles and varying relative orientations of the small- *vs.* large-lobes of the kinase domain[72]. The ps-ns dynamics were probed primarily through HN-NOEs, while the ms-μs dynamics were assessed through R_{ex} rates.

The R_{ex} rates were measured exploiting the cross-correlation between the ^{15}N-^{1}H dipolar-dipolar and the ^{15}N chemical shift anisotropy relaxation mechanisms to avoid high pulsing rates required by CPMG-based experiments, which are not always tolerated by cryoprobes at high salt-concentrations[251]. The inter-lobe orientation was gauged through chemical shift changes and MD simulations. Due to the rapid exchange between open, intermediate or closed topologies, the distribution of populations for these states was quantified through PCA of ppm variations[115,232]. PCA was also utilized to capture the essential dynamics in the MD trajectories of PKA-C, revealing that the first principal component (PC1) corresponds to inter-lobe re-orientations. The plot of representative inter-lobe distances *vs.* the PC1 projection was then used to assess the distribution of open, intermediate or closed conformers[72].

The comparative analyses of dynamics and inter-lobe orientation across the five PKA-C samples investigated by Masterson et al. revealed that nucleotide binding synchronizes the lobe dynamics to catalytic turn over ("dynamically committed state") and that substrate binding leads to a ternary complex that is still dynamically committed and accessing a broad distribution of inter-lobe topologies. However, when the substrate peptide is replaced by a pseudo-substrate inhibitor (PKI) with excess Mg^{2+} ions, the closed topology is significantly stabilized relative to the open and intermediate states, thus introducing free energy barriers that lead to quenched ps-ms dynamics and inhibition of catalytic turn over[72]. Another seminal investigation on allosteric regulation of kinases is the work of Tokunaga et al.[252]. Tokunaga et al. utilized NMR-monitored titrations and comparative chemical shift analyses of Met and Ile methyls measured through SOFAST-HMQC spectra to show that docking interactions in the mitogen-activated protein kinase (MAPKs) control substrate phosphorylation through both anchoring and allosteric effects[252]. Specifically, substrate docking allosterically increases the affinity of ATP for the kinase active sites and ATP binding in turn leads to a shift towards the active conformation, thus up-regulating p38α's enzymatic activity[252].

Another influential contribution that revealed functional coupling between inter-domain interfaces and an active site is the work of Wilson et al. on Pin1, which is composed of a peptidyl-prolyl isomerase (PPIase) domain linked to a WW-domain[253,254,255]. In order to probe the functional role of the PPIase-WW inter-domain interface, Wilson et al. designed an I-to-A mutation that selectively disrupts inter-domain interactions, as confirmed by comparing the chemical shift changes caused by this mutation to those of a domain-deletion mutant[254]. Upon disruption of the PPIase-WW domain-interface, Wilson et al. observed not only changes in substrate affinity and

isomerase activity at the distal PPIase active site, but also changes in intra-domain dynamics as probed by a combination of backbone ^{15}N relaxation, analyzed through N-H reduced spectral density mapping, and methyl ^{2}H relaxation, from which axial order parameters were derived[254]. The determination of the latter requires a reliable description of the overall tumbling motions, which, in the case of the Pin1 I-to-A mutant with disrupted domain interface, was obtained through domain-specific correlation times. These were derived from the $R_2(^{15}N)/R_1(^{15}N)$ ratios measured for backbone N-H vector not affected by significant internal motions, as gauged from the reduced spectral density at zero frequency[254].

The examples reported above show how internal dynamics in the ps-ns or ms-μs time-scales is coupled to changes in overall domain topology. However, it is also possible for inter-domain re-orientations to be coupled with slow (> 10 ms) internal dynamics, as in the case of Crk adaptor proteins controlled by a *cis* – *trans* Pro isomerization switch[256]. Crk includes two tandem SH_3 domains that sample either open ('on') or closed ('off') topologies. However, the latter can form only when a Pro in the inter-domain linker and adjacent to the C-terminal SH_3 domain populates the *cis* isomer[256]. Another allosteric sensor with tandem domains adopting open *vs.* closed orientations is the regulatory subunit of PKA (PKA-R). In the absence of the allosteric effector, cAMP, the tandem cAMP-binding domains are decoupled and sample active *vs.* inactive states independently of each other[257]. Upon binding of cAMP, the active state is selected within each domain and the active-active inter-domain interactions stabilize the closed topology ('double conformational selection')[257].

Overall, the studies reviewed in this section highlight that changes of inter-domain topology (*e.g.* "open" *vs.* "closed" transitions) are critical not only because they control access to active sites proximal to inter-domain interfaces, but also because they are often coupled to changes in functional intra-domain dynamics that extend beyond inter-domain interfaces. These effects cannot be appreciated by the simple rigid-body-like representation of inter-domain re-orientation discussed in Section 2.1.3, whereby domain topology affects exclusively the exposure of active sites at or near inter-domain interfaces. However, it should be noted that it is also possible to observe dynamics that are not correlated to function[258]. Clear signatures of such "non-functional" dynamics are a mismatch between the time-scales or the state-populations of the dynamic process under investigation and the inactive *vs.* active inter-conversion[224,258].

1.4.2.8 Dynamically Driven Allostery

'Dynamically driven' allostery (Fig. 1e) provides an effective mechanism to modulate conformational entropy and control both inter- and intra-molecular interactions, such as protein:DNA[61], protein:ligand interactions[60] and interactions between domains of the same polypeptide chain[248,259]. Although the possibility of purely dynamic allostery in the absence of structural changes was hypothesized in 1984[260], direct experimental proof of dynamically driven allostery became available more recently through elegant NMR experiments on the dimeric cAMP-receptor protein (CRP)[60,89,90,97] . Binding of one cAMP equivalent to the CRP dimer causes structural changes confined to a single protomer, while the other is subject to enhancements in ps-ns dynamics (Fig. 6a). Binding of a second cAMP equivalent quenches such dynamics, thus imposing an entropic penalty, which was confirmed by ITC measurements and explains the negative cooperativity between the two cAMP binding events[60,89]. Another clear proof of the role of dynamics as key determinant of the entropy of binding was provided by the $cAMP_2$:CRP-S62F mutant complex, which samples only ~2% of DNA-binding competent conformers, but still binds DNA with high-affinity due to enhancements in dynamics, and corresponding conformational entropy, upon DNA binding[261].

A valuable meter of conformational entropy changes is provided by variations in order parameters derived from ^{15}N amide and ^{13}C methyl relaxation experiments designed to probe ps-ns time-scale dynamics[61,62,67,68]. Although such changes in internal dynamics tend to decay quite sharply with the distance from the binding site[62], they account for long-range allosteric effects that would be challenging to rationalize otherwise[62] and they result in conformational entropies that scale linearly with the total binding entropies[61,68]. For example, Tzeng and Kalodimos utilized a library of eleven CRP variants to show that the changes in the conformational entropy upon DNA binding evaluated by NMR exhibit an excellent linear correlation with the entropy of DNA binding calorimetrically measured by ITC (Fig. 6b)[61]. The changes in conformational entropy were gauged by Tzeng and Kalodimos starting from the methyl order parameters (S^2_{axis}) through an empirical calibration approach that relies on estimations of solvation entropies based on the differences of solvent accessible solvent area upon binding[262]. Similar estimations of conformational entropy are obtained by analyzing the methyl order parameters using the inventory approach[61,263].

Tzeng and Kalodimos measured the methyl order parameters (S^2_{axis}) through ^1H-based cross-correlated relaxation experiments of per-deuterated CRP samples with the exception of the ILVAM methyls, which were $^{13}CH_3$ labeled[61]. Methyl order parameters can also be measured through ^2H-

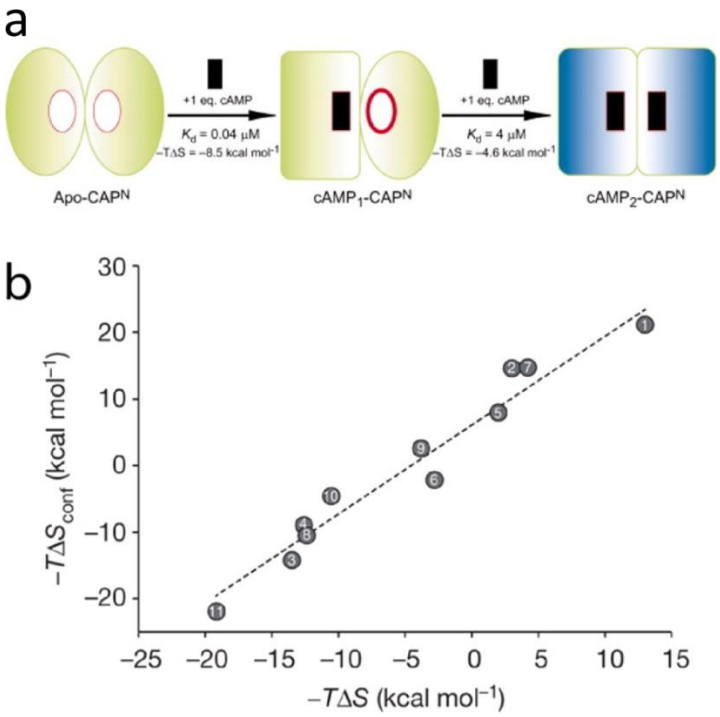

Figure 6. *Fast (ps-ns) dynamics underlies dynamically driven allostery.* **(a)** Dynamically driven anti-cooperativity in the cAMP-binding domain of the Catabolite Activator Protein (CAP). Binding of one cAMP molecule causes structural changes only within the bound protomer, while subjecting the adjacent protomer to enhanced ps-ns dynamics. Binding of a second cAMP molecule quenches dynamics within both protomers creating an entropic penalty. **(b)** Correlation between total binding entropy (-TΔS) determined *via* ITC and conformational entropy ($-T\Delta S_{conf}$) computed by NMR spin relaxation experiments for 11 CAP mutants. Adapted with permission from references 89 (Copyright 2011 Wiley) and 61 (Copyright 2012 Nature Publishing Group).

based relaxation experiments, which however are less sensitive than the ^{1}H-based experiments[61]. This proof-of-principle investigation on CRP corroborates that conformational entropy differences estimated by NMR are an excellent proxy for total entropy changes upon binding[61,264,265]. However, in order to establish such correlations, it is critical to ensure that non-conformational sources of entropy variations, *i.e.* changes in solvation and roto-translational motions upon binding, are constant within error among the variants utilized to establish the linear correlation[61]. Overall, considering that the variants utilized to correlate the conformational and total entropy changes exhibit DNA/CRP interfaces that are essentially invariant, the work of Tzeng and Kalodimos[61] is an excellent illustration of how modulations of internal ps-ns dynamics are essential to rationalize variations in binding free energies that cannot be otherwise explained by structural changes or variations in the active *vs.* inactive populations.

Another "meter" of conformational entropy, is provided by the backbone order parameters that quantify the amplitude of ps-ns fluctuations of the N-H vectors[60]. Order parameters of N-H vectors are typically obtained through the traditional model free (MF) analysis of ^{15}N T_1, T_2 and HN-NOEs relaxation data or through generalizations of the MF approach, such as the two-body coupled-rotator method called slowly relaxing local structure (SRLS)[266]. The SRLS analysis models local protein motions in terms of a local ordering tensor and a local diffusion tensor, a description that is more realistic than that obtained through the original MF method alone. However, the best-fit SRLS parameters are often affected by large uncertainties. This limitation of the SRLS approach is circumvented by integrating the SRLS analysis of relaxation data with MD simulations[266]. The combined method, referred to as SRLSMD, was successfully applied to the comparative analysis of dynamics of a Rho-GTPase[266] and illustrates the synergy between NMR relaxation and MD simulations.

Several other illustrations of the complementarity between NMR and MD simulations are available[53,55,267,268]. For example, VanSchouwen et al. first validated MD simulations using NMR-derived order parameters, and then exploited MD simulations to explore aspects of allostery that are not readily accessible through experiments[51,53]. For example, it was possible to assess the presence of dynamically driven allostery in EPAC constructs that more closely approach the full-length integral protein than those amenable to NMR analyses[51]. In addition, through MD simulations it is possible to explore transition pathways between the experimentally determined apo-inactive and holo-active structures[54,55,269], hydration changes relevant for allostery[221] as well as dynamics that is elusive to experimental characterization, *e.g.* flexibility in systems with poor solubility or large MWs and in transient intermediates, such as the apo-active or holo-inactive states (Fig. 1a)[53,55].

1.4.2.9 Other NMR Methods to Probe Conformational Ensembles

Although NMR chemical shift perturbations and spin relaxation experiments are invaluable to map allosteric networks and the dynamical basis of allostery, as outlined above, other NMR observables provide an often-indispensable complement. For example, comparative analyses of hydrogen exchange-based protection factors are excellent probes of solvent accessibility and have been utilized to understand how occlusion of binding and active sites controls autoinhibition[270] and to detect minute populations of conformers with solvent exposed sites[271]. In more general terms, H/D exchange protection factors sense transient local and global unfolding events[272,273], providing experimental validation of ensemble models, such as those provided by the Corex algorithm[36]. The

maximal protection factors measured for buried sites are directly related to the free energy of global unfolding, which is also accessible through independent control experiments in which chemical or thermal denaturation is monitored, for example, by fluorescence or differential scanning calorimetry, respectively. Free energies of global unfolding are useful to gauge the depth of wells in the free energy landscape that defines allosteric transitions (Fig. 1). This is because the unfolded states of different conformers are assumed to correspond to a similar free energy level, providing a common reference frame to evaluate the differential free energies of the folded conformers (Fig. 1b)[42,43]. By measuring mutant *vs.* WT unfolding free energy differences it is possible to establish the mechanism though which a mutant increases the population of a given state, *i.e.* active or inactive: either by selectively stabilizing that state or by selectively destabilizing the other state or by a combination of the two mechanisms.

H/D NMR experiments are also an effective tool to spot non-additivity in double-mutant cycles, in which one perturbation is a mutation and the other is the thermal transition from exchange-incompetent to exchange-competent states of a single amide proton[274]. In addition, two H/D exchange methods that extend beyond classical H/D exchange kinetics and provide important details about allosteric systems are equilibrium H/D fractionation factors and H/D isotope effects on NMR chemical shifts. Both techniques probe the stability of backbone and side chain hydrogen bonds and examine how they change in response to ligand binding or conformational exchange[189,275]. Last but not least, NMR provides also a means to monitor chemical unfolding at residue resolution. For example, Chen et al. monitored the urea-induced unfolding of the CesAB chaperone using normalized $^{1}H,^{15}N$-HSQC cross-peak intensities and extracted residue-specific free energies of unfolding and apparent urea concentrations at the midpoint of unfolding transition (C_m)[276].

1.4.3 Ligand-Based View of Allostery

While the previous discussion focuses primarily on the allosteric receptor, the allosteric effector deserves an equally important level of scrutiny. The resulting ligand-based view of allostery is complementary to the protein-based maps outlined above and hinges upon two main themes: reciprocal conformational selection and dissection of "driver" *vs.* "anchor" groups. These topics are outlined in the following sections.

1.4.3.1 Multiple Ligand Conformations - Reciprocal Conformational Selection

While protein conformational changes are central to understanding allosteric coupling, it should be considered that often conformational changes upon binding occur also at the level of allosteric ligands: not only does the ligand select a specific conformation within the protein structural ensemble, but also the protein selects a specific conformation within the ligand structural ensemble, *i.e.* conformational selection is often reciprocal[277]. Hence, it is critical to map binding-coupled conformational changes also for ligands[277,278]. In this respect, traditional structure determinations of complexes[279] are frequently complemented by Transfer-NOESY experiments, which are applicable if the ligand of interest does not bind tightly (K_D > nM). Transfer-NOESY data are combined with STD spectra, which are useful not only to monitor binding isotherms and determine dissociation constants[47,280,281], but also to provide a map of the interaction epitopes within the ligand[282–284]. Another ligand-based NMR method that allows mapping of interactions for phosphorous-containing effectors (*e.g.* nucleotides, phospholipids etc.) in a label-free manner and without requiring isotopic enrichment is ^{31}P-NMR. The ^{31}P chemical shifts and line-widths are exquisitely sensitive to binding[104,285]. Similarly, ^{19}F NMR is a valuable ligand-based sensor of binding and allosteric equilibria[286,287]. For example, Nicks et al. utilized ^{19}F-labeled analogs of the tryptophan synthase (TS) substrate to probe the equilibrium between low and high-activity states of the TS β-subunit[287]. The absence of ^{31}P and ^{19}F in most proteins makes these nuclei excellent tools to selectively monitor binding mechanisms from a ligand-centered perspective.

1.4.3.2. Mapping "Driver" *vs.* "Anchor" Functional Groups

Ligand-centered analyses of allostery are instrumental also to dissect the binding *vs.* allosteric contributions of specific atoms or functional groups within an allosteric effector. Differentiating between binding and allosteric contributions is critical to design ligand analogs that act as agonist, antagonist or reverse agonist. In this respect, it is useful to classify the atoms of allosteric ligands as "drivers" *vs.* "anchors"[288,289] . Both types of atoms elicit specific interactions with the allosteric system, but the former interact with receptor regions subject to inactive *vs.* active variations, while the latter interact with receptor sites that are invariant in the inactive-active transition (Fig. 2). Hence, drivers but not anchors are expected to mediate state-selective interactions and contribute directly to allosteric coupling (efficacy). However, anchors are critical for binding (potency) and for ensuring that drivers are correctly positioned to actually 'drive' conformational transitions[288]. The identification of drivers *vs.* anchors is expected to define a major challenge for future studies of allosteric interactions, especially considering that multiple coupled drivers and anchors may

simultaneously act at a single interface between an allosteric effector and the host protein[288]. Embracing these challenges will enhance the design of novel therapeutic allosteric effectors.

While this is still an emerging area, an effective method to quantitatively assess the "driver" *vs.* "anchor" role of a given atom or group of atoms within an allosteric effector relies on the measurement of state-specific associations constants, *i.e.* $K_{a,Active}$ and $K_{a,Inactive}$. Covalent modifications of purely "driver" sites are expected to alter the $K_{a,Active}/K_{a,Inactive}$ ratio and hence the free energy of allosteric coupling (equation 1) and the efficacy. Covalent modifications of purely "anchor" sites are expected to maintain the $K_{a,Active}/K_{a,Inactive}$ ratio to a large extent unaltered, but affect the average affinity, computed, based on the thermodynamic cycle of Fig. 1a, as:

$$K_{a,Observed} = X_{Inactive,Apo}\, K_{a,Inactive} + X_{Active,Apo}\, K_{a,Active} \qquad (2)$$

where $X_{Inactive,Apo}$ and $X_{Active,Apo}$ denote the molar fractions of inactive and active state prior to ligand binding[47]. An experimental approach to measure state-specific association constants combining molecular stapling, ligand-based NMR (*e.g.* STD) and protein-based NMR (*e.g.* chemical shift correlations) has been recently proposed[47]. Alternatively, state-specific association constants can be quantified through non-equilibrium (*e.g.* stopped-flow) fluorescence measurements that rely on rapid mixing and possibly the use of low temperatures to resolve different kinetic phases in the thermodynamic cycle of allosteric coupling[185,290]. Equation (2) also reveals that the affinity for allosteric systems depends on the position of the inactive *vs.* active equilibrium of the apo form. When such equilibrium is skewed towards the inactive state, the observed average affinity may reflect primarily the $K_{a,inactive}$ state-specific association constant, *i.e.* the stability of the holo-inactive complex (Fig. 1a).

1.5 Mapping Allostery Beyond the Four-State Thermodynamic Cycle

While the four-state thermodynamic cycle arising from the coupling of conformational and 1:1 binding equilibria (Fig. 1a) provides an initial framework to rationalize allostery, it is clear that the simple view of Fig. 1a does not always capture the complexity of allosteric control[291]. For example, although allostery is now recognized as a property of both monomeric and oligomeric proteins[10,292], oligomerization remains an effective mechanism to amplify allosteric responses. This type of amplification involves binding stoichiometries higher than 1:1 in systems subject to homotropic allostery and goes beyond the scope of the simple scheme of Fig. 1a. The importance of stoichiometries higher than 1:1 is easily appreciated also by considering that one of the simplest and most common allosteric ligands is the proton ion[293–296]. Another source of deviations from the

four-state thermodynamic cycle in Fig. 1a is the presence of more than two states in the conformational ensembles accessible to allosteric systems. This is often the case when the allosteric host includes weakly coupled moieties. Weak allosteric coupling can still be functionally relevant, as in the case of G protein coupled receptors (GPCRs)[297]. One of most general treatments of weak coupling is the "Ensemble Allosteric Model" (EAM)[36], in which the energetics of an apo allosteric system is recapitulated in terms of at least two states accessible to each sub-component (*e.g.* domain or sub-domain) and state-specific inter-sub-component couplings. Allosteric ligands alter the free energy hierarchy of the apo ensemble by eliciting state specific interactions with specific sub-components[36]. Hence, the EAM provides an effective conceptual framework to model a wide range of allosteric systems, including those deviating from the simple four-state cycle of Fig. 1a, such as oligomers and weakly coupled systems that access complex multi-state ensembles. The following sub-sections focus on these types of allosteric receptors, starting from *in vitro* systems and progressing towards environments that more closely recapitulate allostery in the context of cellular complexity.

1.5.1 Investigation of Homotropic Allostery in Oligomeric Systems

The original MWC model for tetrameric hemoglobin is an excellent illustration of how oligomerization cooperatively amplifies allosteric couplings. Due to the conservation of symmetry, the MWC model posits that all protomers within the oligomer must always exist in the same state, either tensed (T) or relaxed (R), resulting in a T_n *vs*. R_n conformational equilibrium, where n is the number of protomers in the oligomer (Fig. 7a). Given the higher R *vs*. T affinity for molecular oxygen, O_2 binding to a single subunit shifts the equilibrium towards R_4, increasing the effective concentration of the three unoccupied R subunits and resulting in a net increase in oxygen affinity for subsequent binding. It is clear that this effect is amplified when applied to larger oligomers. Hence, oligomeric proteins have historically been considered prototypical allosteric sensors.

When investigating allostery in oligomeric systems, it is critical to characterize binding intermediates as they help differentiate between different allosteric models. For example, unlike the MWC model, the KNF model allows for mixed R/T binding intermediates (Fig. 7b). Although it is possible for even simple ^{1}H NMR spectra to provide sensitive markers of structural perturbations and quaternary structure[298–300], the investigation of binding intermediates in allosteric oligomeric systems is facilitated by the joint analysis of NMR, ITC and CD data in the context of the ensemble model of allostery[36,301,302]. ITC provides the fractions of singly and doubly ligand-bound enzyme that are utilized to model the HSQC intensities during ligand titration, taking into account residue-

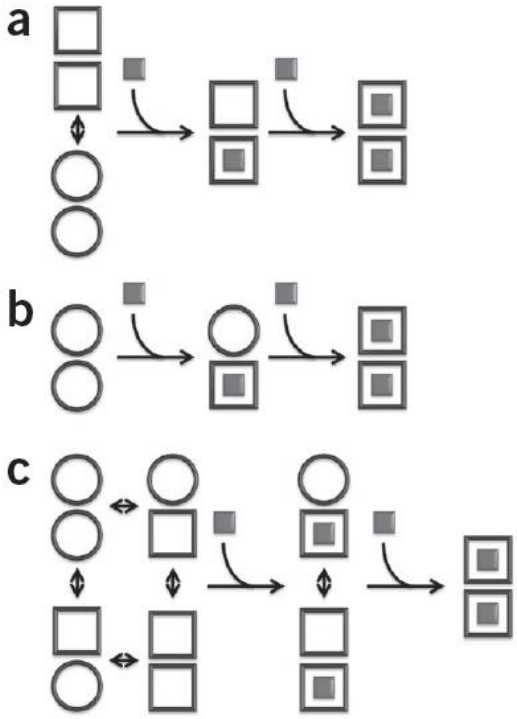

Figure 7. *Models of homotropic allostery as exemplified by a dimer*. Each protomer exists in two states, the relaxed and tensed states (square and circle) with high and low affinity for the allosteric effector ligand. For the sake of simplicity, the affinity of the tensed state for the ligand is assumed to be negligible. In the case of the Monod-Wyman-Changeux (MWC) model (**a**), the dimer symmetry is conserved at every stage of ligand binding, ruling out mixed relaxed-tense combinations. The conservation of symmetry posited by MWC explains positive homotropic cooperativity. In the case of the Koshland-Nemethy-Filmer (KNF) model (**b**) no tensed *vs.* relaxed pre-equilibrium is hypothesized (*i.e.* induced fit) and no conservation of symmetry is assumed. The tensed to relaxed conversion occurs upon ligand binding and cooperativity, either positive or negative, depends on inter-subunit interactions. In the most general case, *i.e.* the Ensemble Allosteric Model (EAM) (**c**), pre-equilibria in the apo form are assumed and no conservation of symmetry is imposed. The population of each apo dimer state depends on the intrinsic stability of each protomer state as well as on the inter-protomer coupling. Upon binding, populations are further modulated by the state-specific affinities of the ligand as well as the ligand concentration. The MWC and KNF models can be seen as specific examples of the EAM[36,91,160]. Reproduced with permission from Freiburger et al.[302]. Copyright 2011 Nature Publishing Group.

specific transfer of broadening from singly- to doubly-bound forms in the parameter fitting[301,302]. CD molar ellipticities are critical to monitor thermal unfolding and their concentration dependence provides a control for the oligomerization state[301,302]. For example, through the joint-NMR-ITC-CD analysis Frieburger et al. elegantly showed that a dimeric enzyme conforms to the MWC model when ligands target the dimer interface, but to the KNF model when ligands bind at distal sites[301](Fig. 7a,b). In addition, it is also possible that the binding and the monomer *vs.* dimer equilibria are coupled to each other, as shown in the case of a chemokine through chemical shift

monitored titrations in conjunction with intrinsic fluorescence polarization to monitor dimerization[303].

1.5.2 Beyond the Two-State Exchange Model

The active-to-inactive transition often occurs through intermediates. Identifying such intermediates is essential to access the degree of allosteric cooperativity and to map the transition pathway(s). The choice of NMR approaches utilized to identify such allosteric intermediates depends to a large extent on the time scale of the exchange between the inter-converting conformers. Hence, we will discuss separately the case of slow, intermediate and fast exchanging 'third' states. However, it should be considered that a single allosteric system may include different sites subject to different exchange rates and regimes (*e.g.* fast and slow exchange)[27] and such multiplicity of timescales is itself evidence of deviation from the simple two-state exchange model [304,305]. This scenario of a single system for which different sites are subject to dynamics in multiple time-scales should not be confused with another common case in which a single two-state exchange (Fig. 8a) appears in the fast-regime for residues with small inter-state chemical shift differences, but in the intermediate or slow regime for residues with larger inter-state chemical shift variations (Fig. 8b). It is clear that

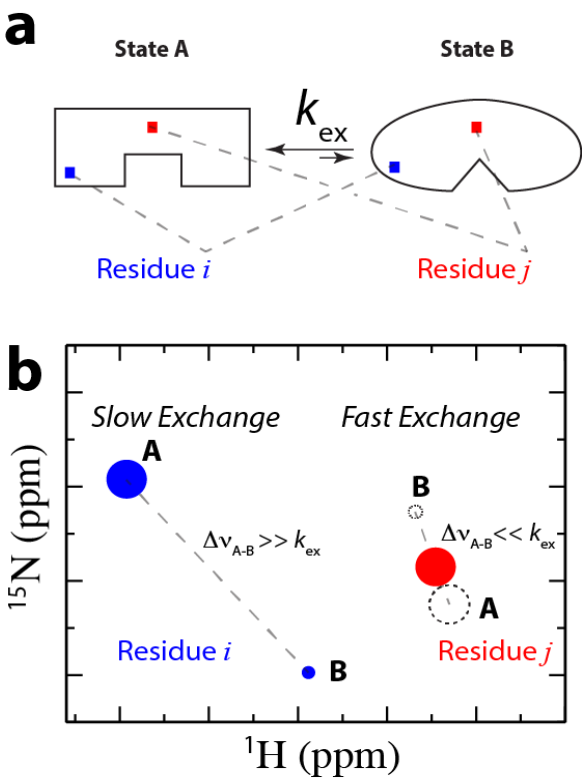

Figure 8: *Illustration of how a single two-state conformational transition may lead to different exchange regimes for different probe residues.* ***(a)*** Schematic model for a single equilibrium between two states (*i.e.* A and B) with exchange kinetics quantified by the rate constant $k_{ex} = k_1 + k_{-1}$. Residues *i* and *j* provide two distinct probes reporting on the same exchange event. ***(b)***. If $\Delta\nu_{AB}$ for residue *i* is sufficiently large to be >> k_{ex}, the N-H for this residue will appear in the slow exchange regime giving rise to two separate cross-peaks (blue). In this case, the intensity of the cross-peaks is indicative of the population of the respective states. If $\Delta\nu_{AB}$ for residue *j* is sufficiently small to be << k_{ex}, the N-H for this residue will appear in the fast exchange regime, giving rise to a single cross-peak (red) for which chemical shifts are population weighted-averages of the chemical shifts of the two pure states (dotted circles).

such partitioning of different residues into distinct exchange regimes is field dependent, as frequency differences are[306].

1.5.2.1 Third State in Slow Exchange

A clear example of the limitation of the two-state model and the relevance of additional states is provided by recent investigations of GPCRs[297,307]. For example, Manglik et al. mapped the free energy landscape of the β_2-adrenergic receptor (β_2AR) using a combination of double electron-electron resonance (DEER) spectroscopy, ^{19}F NMR relaxation dispersion and saturation transfer[297]. By mutating selected residues to cysteine, Manglik et al. were able to covalently link either a spin label for use in DEER spectroscopy or a fluorinated label for ^{19}F NMR relaxation dispersion and saturation transfer experiments aimed at characterizing the dynamic ensemble of conformations selected by either inverse agonists or full agonists. Manglik et al. found that apo β_2AR samples two distinct states (S1 and S2) exchanging in the sub-ms time scale, whereas in the fully active state, in which β_2AR was bound to both a full agonist and Nb80, *i.e.* a mimic of the G protein, G_s, it was found that the GPCR sampled mainly a single state (S4) (Fig. 9a). Interestingly, when Manglik et al. examined β_2AR in the presence of agonist only (*i.e.* in the absence of Nb80), they observed an heterogeneous conformational ensemble in which the S1 and S2 states were still significantly populated together with an intermediate state (S3), with chemical shifts distinct from the other states, as confirmed by line-shape deconvolution, and in slow exchange with the S1 and S2 states, as indicated by ^{19}F saturation transfer experiments (Fig. 9b-d)[297]. The S3 on-pathway intermediate is consistent with the presence of a loose allosteric coupling between the agonist-binding pocket and the G-protein-coupling domain[297]. Overall, the study of Manglik et al. illustrates how the combination of tailored NMR probes of dynamics and other spectroscopies that sample distance distributions (*e.g.* DEER) is often sufficient to map in a targeted but effective manner complex multi-state dynamic equilibria accessed by allosteric systems.

Another excellent example of the functional relevance of deviations from the two-state model is the work of Gustavsson et al.[7] and Reining et al.[308]. Gustavsson utilized [^{13}C,^{13}C]-DARR MAS NMR experiments to show that SERCA-bound phospholamban accesses three states in slow exchange, including two that are inhibitory (T and R) and one that is non-inhibitory (B) and stabilized by phosphorylation[7]. Reining et al. found evidence that the *Vibrio vulnificus* add ribowitch dynamically samples three slowly exchanging, distinct states referred to as apoA, apoB and holo[308]. The riboswitch is regulated by adenine binding, which results in a conformational switch exposing its Shine-Dalgarno sequence and AUG start codon. The exchange rates among the

states were measured via Nz-exchange spectroscopy and stopped-flow kinetics and were used, in combination with binding affinities derived from ITC, to define the switching efficiency (*i.e.* the difference between the holo state populations at two different concentrations of adenine). Reining et al. compared the switching efficiencies between the add riboswitch and a mutated riboswitch, which lacks the ability to form the apoB state, and found that the third state results in a wider range of switching efficiencies over a broader range of temperatures, *i.e*. deviation from a two-state model leads to a "temperature-compensated" regulatory switch [308].

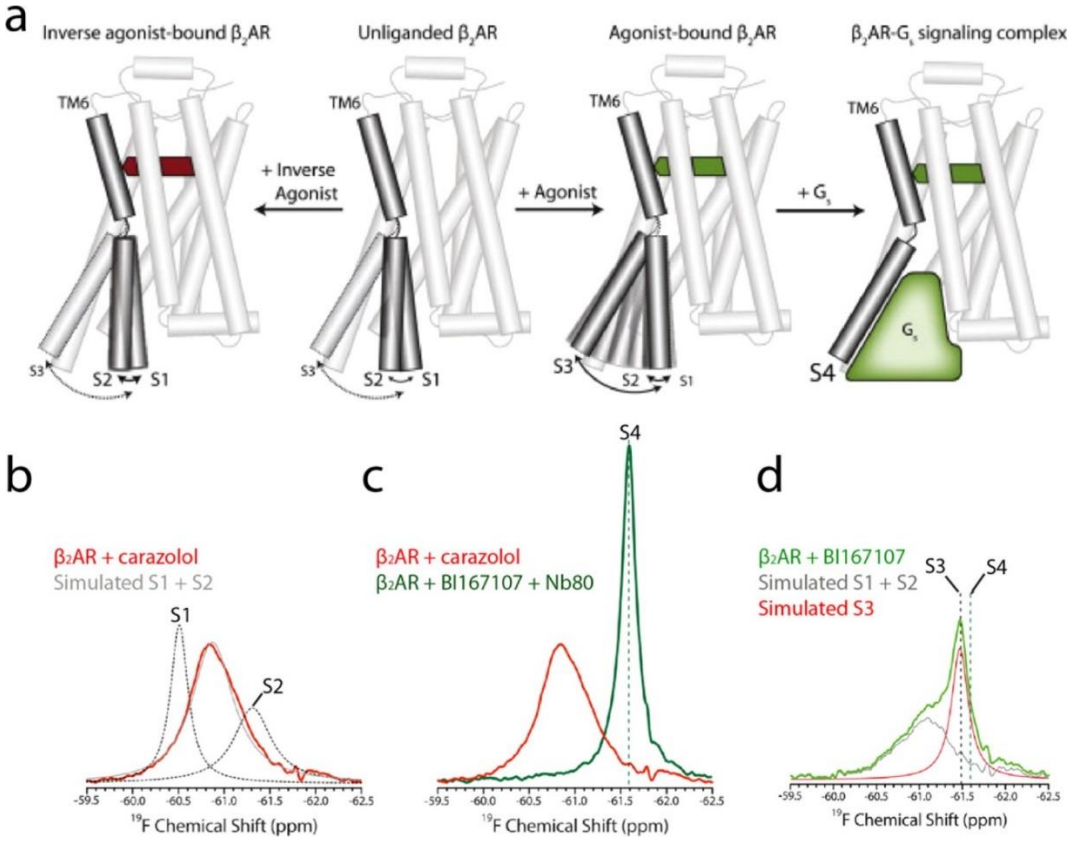

Figure 9: *A multi-state conformational switch*. **(a)** The dynamic conformational equilibria of the β_2 Adrenergic Receptor (β_2AR). The TM6 helix of β_2AR exists in a dynamic equilibrium between its closed (S1), partially closed (S2) and open (S3/S4) conformations. By introducing a fluorine probe in the cytoplasmic end of the TM6 helix, it was possible to selectively probe how this helix shifts among the S1-4 conformers in response to a set of ligand perturbations. For instance, the ^{19}F NMR spectra are shown for β_2AR in the presence of an inverse agonist (carazolol, panels b and c, red), an agonist (BI167107, panel d, green) as well as an agonist and the G protein mimic Nb80 (panel c, green). In addition, line-shape simulations guided by NMR relaxation dispersion parameters provided a means to gauge individual state contributions (panel b, black dashed peaks; panel d, grey and red spectra). Through this approach, Manglik et al. were able to deconvolute spectra like the one shown in panel d and identify a new intermediate state (S3) distinct from S4, although its chemical shift is similar to the S4 state. This figure adapted with permission from Manglik et al.[297]. Copyright 2015 Elsevier.

1.5.2.2 Third State in Intermediate Exchange

Intermediates in allosteric transitions are often sampled in the μs-ms time-scale. For example, McDonald et al. using ^{15}N CPMG NMR dispersion (NMRD) experiments proved that this is the case for the *E. coli* CheY, a regulatory protein that controls bacterial chemotaxis[304]. McDonald et al. observed a poor correlation between the $\Delta\omega$ values extracted from NMRD data and the $\Delta\delta$ chemical shift changes between CheY in the presence and absence of saturating amounts of the phsophoryl-mimic BeF_x and 10 mM Mg^{2+}, which stabilize the active state. This suggested that the transitions sampled by NMRD do not conform to a simple concerted two-state model for the inactive *vs.* active switch. Overall, the μs-ms dynamics of CheY mapped at residue-resolution through NMRD experiments are consistent with asynchronous segmental motions that give rise to local, as opposed to global, switching between inactive and active-like conformers of CheY[304]. Based on these and similar observations[304,305,309], the two main signatures of deviations from the two-state model for allosteric systems subject to μs-ms dynamics appear to be the heterogeneity in site-specific exchange rates from NMRD and lack of correlations between the ground *vs.* excited $\Delta\omega$ from NMRD and the apo *vs.* holo chemical shift differences ($\Delta\delta$). When significant deviations from the two-state model are detected, fitting of CPMG relaxation dispersion data using more complex multi-state models is possible, although additional experimental data may be needed to avoid dealing with an underdetermined problem[310,311]. Alternatively, viscosity-dependent kinetics is an elegant approach to probe the length scale of segmental motions underlying solvent-driven processes and to measure the related effective hydrodynamic radii (EHR), as was shown for the gating of the 20S core particle[223]. Furthermore, the EHR size is a useful predictor of how exchange kinetics is affected by cellular crowding[223].

1.5.2.3 Third State in Fast Exchange

A promising approach to characterize fast-exchanging multi-state systems is based on the combination of residual dipolar couplings (RDCs) and Monte-Carlo simulations to explore the protein energy landscape[312]. The RDCs measured for distal sites subject to conformational changes are sensitive to the extent of correlation between such changes[312]. This effect was exploited to show that the inter-domain equilibria of adenylate kinase include not only fully closed and open topologies, but also a conformationally heterogeneous intermediate state in which one domain adopts an open-like orientation, while another is in a closed-like conformation[312]. Another method to investigate conformational ensembles with three rapidly interconverting states capitalizes on the observed chemical shifts being population-weighted linear averages of the pure state values. In this

case, it is instrumental to expand the apo *vs.* holo comparison to a third sample in which the third state is at least partially populated. For example, one of the cyclic-nucleotide binding domains of protein kinase G samples three states, *i.e.* an active state, prevailing in the apo form, an inactive state, selected by cGMP, as well as an intermediate states with mixed active/inactive features, partially stabilized by cAMP[313]. The comparative chemical shift analysis of the apo, cGMP- and cAMP-bound forms through CHESPA was critical to differentiate between two- *vs.* three-state models, as it systematically revealed non-linear HSQC patterns for selected residues distal from the cyclic nucleotide binding sites and sub-sets of residues exhibiting different degrees of fractional activation[313]. Furthermore, based on the CHESPA analysis it was possible to estimate the populations of the three states[313]. This example illustrates the importance of complementing spectroscopic approaches with an at least partial stabilization of the intermediate through chemical means (*e.g.* replacement of cGMP with cAMP). When covalent modifications achieve an essentially quantitative stabilization of the intermediate, the trapped third-state becomes amenable to more classical structural and dynamical determination methods.

An excellent example of how to utilize mutations to trap transient intermediates is provided by the work of Zhuravleva et al. on DnaK[25]. DnaK functions by cycling between the ATP- and the substrate and ADP-bound states. Substrate binding to the Substrate Binding Domain (SBD) enhances ATP hydrolysis, while ATP binding to the Nucleotide Binding Domain (NDB) reduces the affinity of the substrate for the SBD. To dissect the mechanism of allosteric cross-talk between substrate binding and ATP hydrolysis, Zhuravleva et al. trapped an intermediate of the DnaK allosteric transition in which both substrate and ATP are simultaneously bound to DnaK. This was achieved by identifying a point mutation of DnaK with impaired ATP hydrolysis. The comparative NMR analysis of the ATP-bound mutant DnaK in the presence and absence of substrate revealed that the latter 'undocks' the interface between the nucleotide binding domain (NDB) and the b-subdomain of the substrate binding domain (SBD) of DnaK, while preserving interactions between the NDB and the NDB-SBD linker, which are lost only upon hydrolysis of ATP. Hence, the use of a point mutation to silence ATP hydrolysis and capture the catalytically inactive intermediate of DnaK was instrumental in dissecting which ligand controls which interaction among those perturbed in the allosteric transition. Furthermore, the work of Zhuravleva et al. on DnaK highlights how inter-domain linkers may serve not only as passive inter-domains tethers, but also as active allosteric elements coupled to ligand binding, as discussed in Section 3.4. The investigation of

DnaK also emphasizes how understanding functionally relevant allosteric processes hinges upon thorough maps of multi-state conformational ensembles, which is the subject of the next section.

1.5.3 Multi-State Conformational Ensembles

Understanding allostery in the context of conformational selection in the broadest sense requires an atomic-resolution structural description of the conformers that dynamically interconvert within the ensemble accessible to a protein. However, multi-conformer structural determination implies that the number of degrees of freedom easily exceeds the number of experimental restraints. Hence, relying exclusively on experimental data would lead to an underdetermined problem, which means that integration of experiment and simulations is advisable. For example, in the case of sub-ms dynamics, it is possible to model the conformational ensemble through NOE- and RDC-restrained ensemble molecular dynamics (MD) simulations. Given the complexity of the ensemble a large number of experimental restraints are needed even for a relatively small protein such as ubiquitin, *i.e.* > 2500 NOEs and ~2000 NH RDCs from 36 alignment media[314]. Furthermore additional data (*i.e.* RDC for different bond vectors, cross-correlated relaxation rates and trans-hydrogen-bond scalar couplings) are needed for validation purposes. This approach provides a validated ensemble of hundreds of conformers called "Ensemble Refinement for Native proteins using a Single alignment Tensor" or ERNST[314]. Once the ERNST ensemble is available, it is interrogated through correlation analyses (*e.g.* correlation matrices and principal value analysis) with the goal of identifying correlated motions in the sub-ms time-scale[314]. Through this approach Fenwick et al. revealed collective motions involving β-strands separated by ~15 Å in ubiquitin and defining a path for signal propagation across the ubiquitin structure. In addition, such collective motions may assist the sampling of conformations adopted by ubiquitin when in complex with other proteins, suggesting a role in protein-protein recognition[314]. Another highly promising avenue to exhaustively sample conformational space and map the underlying free energy landscape is through replica-averaged meta-dynamics simulations in combination with NMR chemical shifts and RDCs, which sense ms dynamics[315,316]. For example Kukic et al. using this approach revealed that the apo I-domain of the Lymphocyte Function-associated Antigen 1 (LFA-1) accesses three conformational states[315]. These examples on folded proteins show the effectiveness of hybrid simulation-NMR approaches in recapitulating multi-state conformational ensembles, a problem that is exacerbated in systems that are not fully structured, as discussed in the next section.

1.5.4 Investigation of Allostery in Intrinsically Disordered Proteins (IDPs) and Linkers

While allostery is a concept originally developed for globular proteins, recent evidence suggests that long-range cross-talk is present also in intrinsically disordered proteins (IDPs), which include amyloidogenic peptides, such as α-synuclein (αS) and amyloid beta (Ab), as well as partially unstructured regions in multi-domain proteins, such as linkers[8,42,91–95,95]. Allostery in IDPs is expected to be critical to understand the effect of post-translational modifications, such as oxidative nitration and phosphorylation[8,13]. Such covalent modifications often selectively stabilize conformers with residual structure that are sampled within the complex free-energy landscape of IDPs, thus giving rise to allostery. Hence, one of the critical challenges is to identify strategies to single out functionally relevant conformational equilibria from the background of non-functionally relevant conformational "noise" arising from the dynamic ensemble of interconverting conformers accessible to IDPs.

The NMR investigation of IDPs has been recently covered by excellent monographies and reviews[317–320]. Hence, here we will just briefly mention one possible approach to identify allosteric couplings involving IDPs and linkers based on double-mutant cycles. For example, the allosteric coupling between a linker and an adjacent ligand-binding globular domain of the regulatory subunit of protein kinase A (PKA-RIa) was investigated by preparing four samples: agonist and antagonist-bound, each of them with or without deletion of the linker. For all four-samples unfolding free energies were measured through urea titration monitored by intrinsic fluorescence of Trp residues located in the globular domain[42]. When the globular domain binds to the antagonist (*i.e.* Rp-cAMPS), it is locked in the inactive state and the linker deletion does not result in significant changes of the unfolding free energy, pointing to negligible linker-domain interactions. When the globular domain binds the agonist (*i.e.* cAMP), it is locked in the active state and the linker deletion causes a significant reduction of the unfolding free energy[42]. Hence, the linker is allosterically coupled to the adjacent domain through weak but functionally significant state-selective interactions, which were independently confirmed through PREs and CHESCA.

Binding of cAMP to the CBD shifts the linker's equilibrium towards the states that preferentially interact with the active CBD. Hence, this linker is an integral part of the allosteric network controlled by cAMP and mutations within this allosteric linker modulate the kinase inhibition and cAMP-dependent activation[42]. The allosteric linker of PKA is representative of a more general class of linkers that elicit state-selective interactions with adjacent domains thus tuning regulatory equilibria in allosteric systems, such as DnaK, K-RAS4B and Crk[128,256,321]. Similarly, in the X-ray structures of the copper-sensing operon repressor (CsoR) the N-terminal tail

is folded, but solution NMR experiments show that in the apo state the tail is unfolded and undergoes an allosteric disorder-to-order transition upon ligand binding[322]. Thus partially unfolded regions flanking folded domains are often key regulators of function. However, it should be considered that interactions involving linkers and IDPs are likely to be significantly affected by environmental conditions, which are the subject of the next section.

1.5.5 Contextualizing Allostery: Recapitulating Cellular Complexity

While *in vitro* investigations remain a necessary fundamental starting point to dissect and understand the complexity of allosteric systems, it is critical to take into account that allostery within the cellular context involves additional layers of complexity. These include the presence of multiple competing interactions, the lipid membrane as well as crowding and quinary effects[323]. For example, multiple binders are common for signalling systems that function as hubs, such as RAS GTPases. A single RAS protein often competitively binds multiple effectors, referred to as RAS binding domains (RBDs). The competition between RBDs for binding to RAS is a central determinant of the hierarchy of downstream signaling pathways. Hence, Smith and Ikura[324] developed an elegant NMR-based approach to directly probe the competition within pairs of RBDs (*i.e.* RBDa and RBDb) for RAS binding. In this scheme, HSQC spectra are acquired for 2:2:1 mixtures of RBDa:RBDb:^{15}N-RAS, with the latter bound to a non-hydrolysable analog of GTP (Fig. 10). Since only RAS is isotopically labeled, the intensities of well-resolved HSQC cross-peaks specific for each RBDx:RAS complex provide a direct and quantitative readout of the extent of RBDa *vs.* RBDb competition for binding to wt RAS[324] (Fig. 10). Through this method a competition matrix was built for all the RBD pairs. A similar matrix was built for an oncogenic mutant of RAS, revealing that the mutation alters the hierarchy of downstream signaling pathways (Fig. 10). Overall, the approach of Smith and Ikura provides a general and robust method to probe direct competition in ternary mixtures without the need of indirect inferences from data on binary mixtures[324]. This type of application is an excellent illustration of how NMR can be utilized to dissect the allosteric complexity of signaling hubs, which regulate multiple competing interactions. Another allosteric effector is the lipid bilayer. For example, Gustavsson et al. using MAS NMR provided proof of the principle that lipid bilayer composition affects the kinetics and the position of regulatory equilibria, such as that between the ground and excited state of the membrane protein phospholamban, which controls the sarcoplasmic reticulum calcium ATPase[7,325].

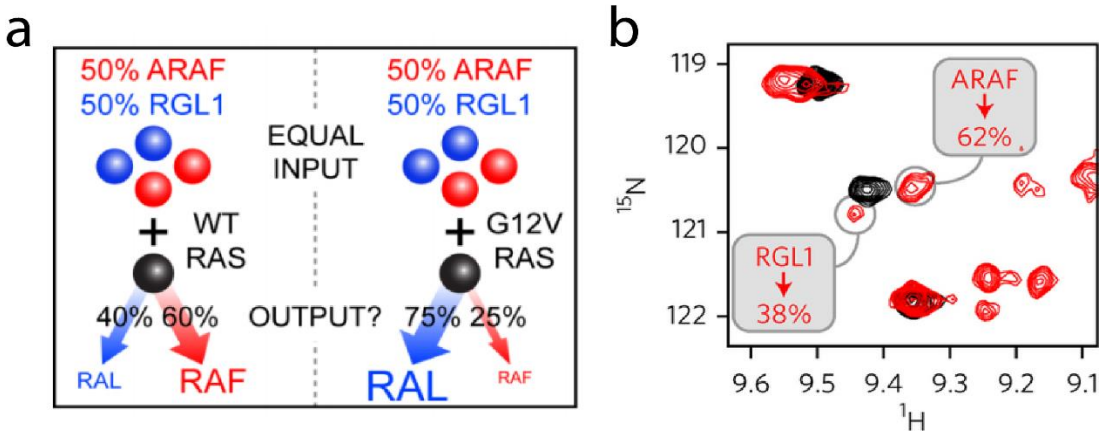

Figure 10. *Similar binding "inputs" result in different WT vs. mutant "outputs".* **(a)** The input is defined by the competitive binding of two-fold excess RAS-binding domains (RBDs), such as unlabeled ARAF and RGL1, to ^{15}N-labeled RAS, while the output is the fraction of each binary RBD:RAS complex formed, as quantified by well-resolved HSQC cross-peaks **(b)**. The "outputs" RAF and RAL denote the GTPases that the guanine exchange factors ARAF and RGL1 respectively regulate. Adapted with permission from reference 324. Copyright 2014 Nature Publishing Group.

1.6 Towards NMR-Assisted Translation of Allostery

NMR is a versatile spectroscopic tool suitable not only to elucidate fundamental allosteric mechanisms, but also to tap the translational potential of allosteric site targeting. Specifically, NMR has played a pivotal role in the screening of new allosteric effectors and the elucidation of their underlying mechanism of action. Furthermore, NMR is ideally suited to map the allosteric effects elicited by disease related mutations (DRMs), which are often elusive to other structural determination methods[326]. The following sections review recent contributions of NMR to the search for new allosteric ligands and their mechanisms, as well as those of DRMs.

1.6.1 Screening for New Allosteric Effectors

The targeting of allosteric sites, either covalently or non-covalently, permits a degree of potency and selectivity that is often not accessible by targeting their orthosteric counterparts. As such, screening for allosteric effectors is a rapidly growing field[31,217,327–331]. The discovery of novel allosteric effectors frequently relies on traditional ligand screening methods, including protein- and ligand-based NMR methods[332–334]. However, the extension of such approaches to allosteric systems also poses new challenges that require customization of the original screening methods to ensure an efficient search for ligands targeting allosteric sites with minimal false positives and negatives. For example, saturation transfer difference (STD) NMR experiments are commonly utilized for ligand screening due to their robustness and label-free nature[333,335–337]. STD NMR methods are

effectively adapted to the search of allosteric effectors through competitive binding experiments implemented in the presence of excess ligands targeting the orthosteric site, with the goal of differentiating between competitive and allosteric inhibitors[338]. While STD experiments work best if the K_D > nM, the affinity window of applicability can be expanded through competition experiments with other allosteric ligands[47,339]. Competition experiments also offer the advantage of being less sensitive to non-specific binding (NSB) than direct binding assays. However, if a clear plateau is observed in the binding isotherm monitored through the STD amplification factor[340], contributions from NSB, which is typically non-saturable, are likely negligible. When full saturation cannot be reached due to poor solubility, NSB can be ruled out through protein-based NMR chemical shift changes, which are localized to a specific binding-site in the case of specific binding, as opposed to being scattered throughout the protein surface as expected for NSB[341].

An additional screening strategy to efficiently identify allosteric inhibitors is the fragment based approach of Chen et al.[342] known as target immobilized NMR screening (TINS). TINS requires the immobilization of both the screening target and a reference protein onto a solid support. A reference protein is needed as a control for NSB. 1D ^{1}H NMR spectra are then acquired for the target and reference in the presence of a mixture of fragments and the resulting spectra are overlaid. Fragments that bind specifically to the target will experience extreme line broadening due to enhanced R_2 relaxation rates caused by binding to an immobilized protein. Hence, significantly reduced target (T) vs. reference (R) intensity ratios indicate selective binding to the target. Using TINS, Chen et al. identified 94 fragments from their initial library of 531 compounds that preferentially bound to the adenosine A_{2A} GPCR (A_{2A}R). Chen et al. then filtered out fragments that interacted with the orthosteric site *via* competition experiments with both an orthosteric inhibitor and an agonist. The remaining leads were screened for their effect on the dissociation rate of the orthosteric effectors, with fragments that increased or reduced the k_{off} rate more than 30% being defined as negative and positive allosteric modulators, respectively. Chen et al. found from their initial library of 531 compounds, three ligands that acted as negative allosteric modulators and eight compounds that acted as positive allosteric modulators. Overall, the approach of Chen et al. describes a robust screening strategy for unstable membrane proteins with promise for identifying both allosteric inhibitors and agonists.

Another efficient NMR-based screening method is FAXS (Fluorine chemical shift Anisotropy and eXchange for Screening), which relies on the reduction of ^{19}F R_2 relaxation rates occurring upon competitive displacement of a bound ^{19}F-labeled spy ligand[343–345]. The high-sensitivity and

reliability of the FAXS method are illustrated by its application to the screening of human hexokinase activators with only minimal false positives[346]. The use of ^{19}F NMR in screening for allosteric effectors is not limited to ^{19}F-labeled ligands[346], but applies also to ^{19}F-labeled proteins. For example, Ge et al. introduced 5F-Trp at multiple sites within a ligand responsive loop of a malaria drug target to map ligand binding and allosteric transitions[347] and Pomerantz et al. introduced 3F-Tyr to monitor ligand binding to KIX[348]. Both approaches capitalize on the abundance of aromatic amino acids at protein interfaces and on the exquisite sensitivity of 1D ^{19}F NMR spectra to ligand binding and conformational changes[347,348].

Last but not least, a promising new dimension in screening for allosteric drug leads has been opened by high-pressure NMR studies, which allow the population of conformers with new druggable sites that would have been otherwise hidden[349]. These "non-classical" allosteric sites provide new opportunities to control protein-protein interfaces through small molecules, complementing the "classical" allosteric loci available at ambient pressure[349] . An alternative approach to identify novel allosteric sites relies on computational solvent mapping (FTMap).[88] The FTMap predictions can be tested through NMR titrations, which are exquisitely sensitive to weak interactions and are not affected by crystal packing artifacts.

1.6.2 Mechanisms of New Allosteric Effectors

Once an allosteric inhibitor is identified through screening, elucidation of the underlying mechanism of action through further biophysical characterization is likely to assist further lead optimization[288,350,351]. One of the simplest forms of biophysical characterization is a 2D NMR-monitored titration of the newly identified allosteric effector into a solution of the isotopically labeled receptor, which provides not only a direct validation of the target protein, but also an initial determination of the location of the inhibitor binding site as sensed by residue-specific cross-peak intensity losses and/or chemical shift changes[30,32,329,352–363]. However, the magnitude of residue specific intensity and chemical shift changes may not be sufficient to obtain an unambiguous identification of the binding sites for allosteric effectors, which, by definition, cause both short- and long-range perturbations. Hence, it is critical to corroborate the location of binding sites for allosteric effectors using mutagenesis[362,364,365] or more direct means, such as PREs or reverse saturation transfer[27,366]. For example the selective saturation of an allosteric inhibitor in the presence of a deuterated ILV-methyl ^{13}C,^{1}H labeled 20S proteasome core particle resulted in attenuation of the ^{13}C-^{1}H HMQC (Methyl TROSY) cross-peaks corresponding to a Valine in the inhibitor binding site (*i.e.* V101)[27] (Fig. 11). This result was further confirmed by the observation

that upon inhibitor binding V101 exhibits a non-linear chemical shift pattern distinctly different from those detected for other residues subject to conformational changes only[27]. This type of non-linear *vs.* linear singularities is conveniently and systematically identified through the chemical shift projection analysis or CHESPA[120,121], assisting the separation of contributions from nearest-neighbour effects *vs.* conformational equilibria shifts to ppm variations[121,367,368]. Once the binding pocket for the allosteric inhibitor is reliably identified, titration of the endogenous ligand in the absence and presence of the allosteric inhibitor provides a quantitative assessment of the extent of coupling between the two ligands[352].

Further mechanistic insight is obtained through structural elucidations[369–374]. For example, the inhibitory mechanism of a compound 4EGI-1, that targets eukaryotic translation initiation, was determined by elucidating the structure of 4EGI-1 bound to the translation initiation factor elF4E. This structure revealed that 4EGI-1 interacted with a site distal from the interface at which elF4E binds with its translation initiation partner elF4G. Furthermore, it was revealed that 4EGI-1 causes an elongation of a helix within elF4E by one-turn, which interferes with its ability to interact with elF4G[375]. Another example underscoring how structural information, even at low domain resolution, assists in dissecting the molecular mechanism of allosteric drugs is provided by the work of Skora et al.[162]. Skora et al. comparatively analyzed the inter-domain orientation of the c-Abl kinase in its binary and ternary complexes with Imatinib, which is a competitive inhibitor targeting the canonical ATP binding site, and GNF-5, which is an allosteric inhibitor that binds the myristoyl pocket and helps to overcome resistance to inhibitors that bind at the ATP binding site (*e.g.* Imatinib). Skora et al. found that Imatinib promotes an open (active) conformation, while GNF-5 locks the c-Abl in its closed (inactive) topology (Fig. 3) [54]. Interestingly, when GNF-5 is added to the binary complex of c-Abl:Imatinib, the c-Abl shifts back to its closed state (Fig. 3)[162].

The previous examples illustrate how structural information helps understand allosteric mechanisms of inhibition. However, often structures alone are not sufficient, because apo allosteric systems access active *vs.* inactive equilibria that are skewed towards the latter state, with only minimal populations of the former. This implies that inhibitors that further stabilize the inactive state may result in essentially undetectable structural changes[209] (Fig. 12). However, the reduction in the active state population occurring upon inhibitor binding is likely to result in major changes in the dynamic profile as probed by NMRD, especially when the active *vs.* inactive exchange occurs in the intermediate regime[209,271]. Another example illustrating how changes in dynamics upon inhibitor binding help elucidate the underlying inhibitory mechanism is provided by the work of

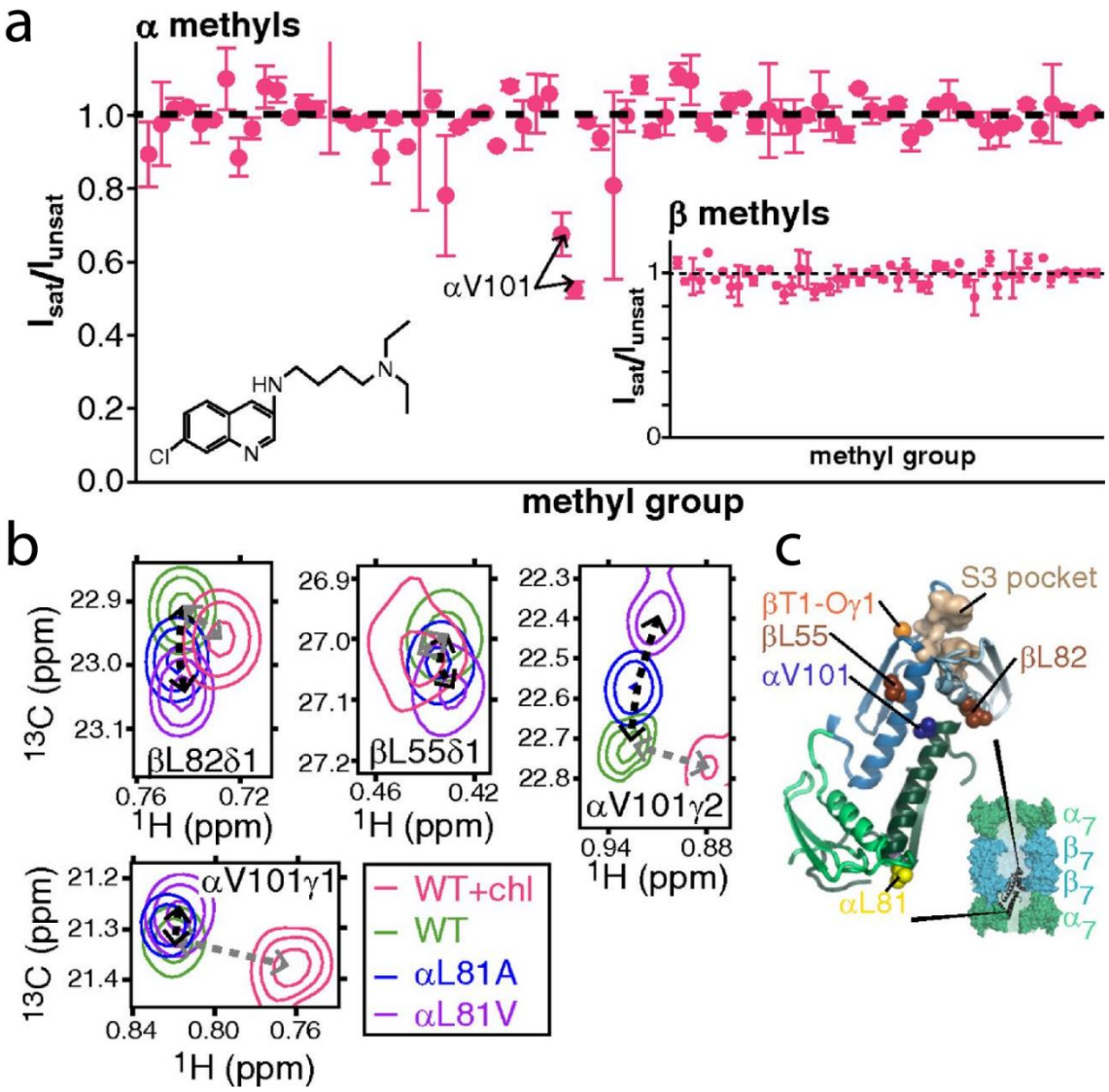

Figure 11: *Mapping the binding interface of a ligand to its allosteric receptor through reverse saturation transfer*. **(a)** Saturation transfer from the chloroquine ligand (structure in insert) to the proteasome 20S core particle (CP) ILV methyls in either the α or β subdomains. Saturation transfer was quantified through on *vs.* off-resonance intensity ratios, where on-resonance refers to saturation of chloroquine's aromatic resonances. Significant saturation transfer was observed for αV101 indicating that chloroquine bound in close proximity to this Valine. **(b)** Non-linear chemical shift patterns confirm that chloroquine binds close to αV101. For residues close to αV101, the binding of chloroquine (light pink) results in non-linear chemical shift changes (gray dashed arrows) unlike the cross-peak shifts that reflect purely allosteric changes (black dashed arrows). **(c)** Residues with non-linear chemical shift changes highlighted in (b) mapped onto the crystal structure of CP as spheres. Figure was reproduced with permission from Ruschak and Kay[27]. Copyright 2012 National Academy of Sciences, U.S.A.

Hammoudeh et al.[369]. Hammoudeh et al. first performed a screen against dihydropteroate synthase, a dimeric bacterial enzyme involved in folic acid synthesis, and discovered an allosteric inhibitor. Then, using NMR and crystallography, Hammoudeh et al. showed that the inhibitor binds at the dimer interface and does not prevent substrate binding, but rather exerts its effect at a later step in the catalytic cycle, by controlling active site loops that modulate product release. MD simulations further revealed how the inhibitor binding at the dimer interface affects the dynamics of the active site loops[369].

Another example in which structural and dynamical investigations provided insight instrumental for further lead optimization is the work by Meli et al. on comparative apo *vs.* holo analyses of NMR chemical shifts, NMR relaxation and MD simulations, which revealed how a small molecule inhibitor binds to a relatively flat and at first sight "undruggable" surface of FGF2[376,377]. The work of Meli et al. illustrates how the integration of NMR and simulations is critical to elucidate the mechanism of action of new allosteric drug leads. For instance, Rousaki et al. examined the action of the anti-cancer compound, MKT-077, which targets the Hsp70 chaperone, and found that it interacts with Hsp70 allosterically. Through NMR chemical shifts in combination with NMR-restrained docking simulations and MD simulations, the drug's binding site was identified[355]. The investigation of the mechanism of new allosteric effectors identified through screening is expected to be an area in which hybrid NMR and computational methods are not only helpful, but often essential as drug leads are frequently poorly soluble and full saturation is not reachable experimentally leading to sparse NMR data that can only be interpreted through complementary simulations.

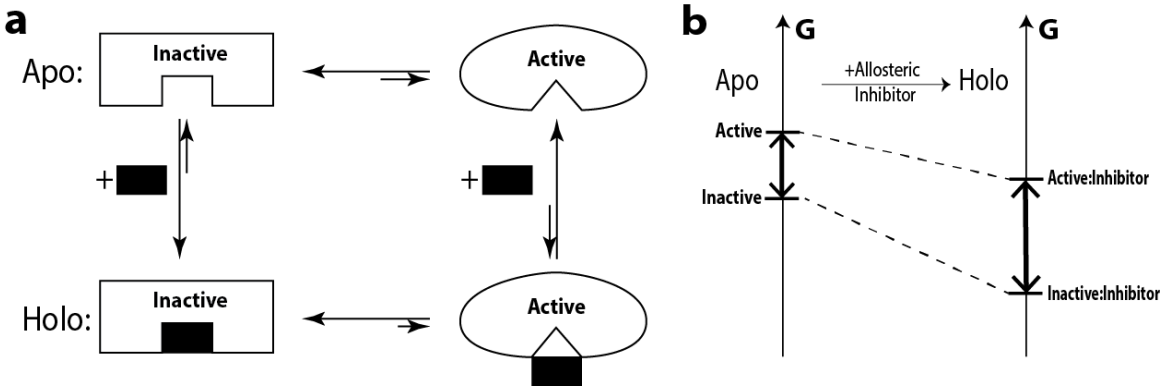

Figure 12. *Simplified mechanism of allosteric inhibition.* Unlike allosteric effectors (Fig. 1a), ligands that elicit allosteric inhibition are selective for the inactive state, as shown in panel **(a)**, leading to further stabilization of such state, as shown in panel **(b)**. Hence, inhibition may occur without apparent change of structure[209].

1.6.3 Disease Related Mutations

Allostery is critical to understand how mutations affect function and result in disease. Often disease-related mutations (DRMs) lead to both short and long-range perturbations and it is necessary to map both types of effects to rationalize the gain or loss-of-function phenotype of a given mutant[18–20,378,379]. In this respect, mutation-induced structural changes are a central aspect of understanding disease at the molecular level. However, often DRMs do not cause apparent changes in structure, but in stability with respect to unfolding[380] and/or in dynamics[231]. Mutation induced destabilization is quantified through equilibrium unfolding or H/D exchange experiments monitored by NMR and discrepancies between the two approaches are often an indication of the presence of unfolding intermediates[381]. Mutation induced changes in dynamics are also effectively probed by NMR. For example, Kim et al. analyzed by NMR a lethal arginine deletion in phospholamban, which is phosphorylated by protein kinase A (PKA)[231]. This phospholamban mutation causes reduced phosphorylation rates and PKA affinities. The underlying mechanism was revealed by the comparative WT *vs.* mutant NMR analysis, showing that, unlike WT phospholamban, the mutant is unable to drive the full transition from intermediate to fully closed states of the ATP-bound catalytic subunit of PKA, locking the catalytic subunit in a partially closed state subject to enhanced but dysfunctional active site dynamics that is not catalytically committed, thus explaining the reduced catalytic efficiency[231] (Fig. 13).

The work of Kim et al. is an excellent illustration of how modulation of dynamics, as revealed by comparative WT *vs.* mutant analyses of NMR chemical shifts, NMR relaxation and MD trajectories are essential to mechanistically relate mutations to disease phenotypes. The investigation of DRMs can be expanded by designing mutant libraries to further interrogate the allosteric system, as outlined in Section 2.2.1. For example, the analysis of an Arg deletion DRM could be complemented by "softer" mutants, in which parts of the Arg side chain are eliminated or modified (*e.g.* mutations from R to K, M, A). Such library of mutations would provide opportunities to initiate CHESCA-like analyses aimed at dissecting nearest neighbour *vs.* long-range effects of DRMs. Furthermore, a mechanistic understanding of how DRMs and related mutations perturb allosteric regulation in the context of even the simplest models of allostery (Fig. 1a) requires an assessment of the DRM-induced changes not only in the position of the apo inhibitory equilibria (Section 2.2.), but also in the state-specific association constants (Section 2.3.2.). The former assessment is essential to understand how DRMs control auto-inhibition or constitutive activation, while the latter helps rationalize how DRMs affect ligand-dependent

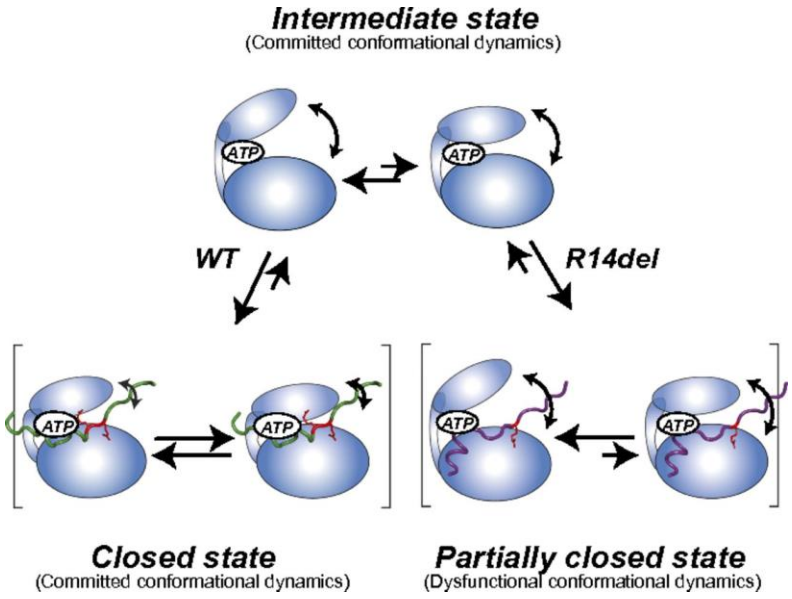

Figure 13: *Modulation of functional dynamics by a disease related mutation (DRM).* The catalytic cycle of protein kinase A's catalytic domain (PKA-C, blue lobes) is characterized by a range of dynamic conformational states defined by the relative positions of its N- and C- terminal lobes. In its nucleotide and substrate-bound form, the wildtype (WT) enzyme forms a closed conformation, which is stabilized by electrostatic interactions between glutamates in PKA-C and R14 in the phospholamban substrate. Dynamics are still present within this form, but they are significantly reduced compared to the nucleotide-bound enzyme. In contrast, the R14 deletion mutant perturbs the interactions between PKA-C and the substrate, limiting the transition of the N- and C-terminal lobes into the closed complex. Instead, the enzyme adopts a partially closed conformation with enhanced dynamics with respect to the WT. This figure was reproduced with permission from Kim et al.[231]. Copyright 2015 National Academy of Sciences, U.S.A.

activation. Both types of assessment provide a basis to dissect the contribution of DRMs to ligand affinity in the context of equation (2).

NMR is an excellent tool not only for elucidating the molecular mechanism of action of DRMs, but also for functionally profiling disease related mutants. For example, Smith et al. utilized cross-peak intensities in HSQC spectra of ^{15}N-labeled RAS to directly monitor in real time the kinetics of RAS cycling within mammalian tumor cell lysates[382]. Significant catalysis was detected with 10-100 μg of protein per sample using a 1.7 mm cryoprobe NMR[382]. The dead time to set up an NMR experiment is typically of the order of minutes. This robust and sensitive approach when applied in combination with cell-based assays, such as immunoblotting to monitor protein expression levels, allows to differentially profile how individual mutants affect catalysis, allosteric regulation, and membrane targeting[382]. The investigation of Smith et al. is an excellent illustration of how NMR–based assays provide a new dimension to traditional enzymology and cell biology studies. In this respect, further impact from NMR-based enzymatic assays are anticipated when protein-based NMR detection, as pioneered by Smith et al.[382], is complemented with ligand-based

NMR detection of ^{1}H and ^{31}P nuclei, which circumvent the need of isotopic enrichment[104,383,384], or ^{13}C, when isotopic labeling is possible[373]. ^{1}H NMR in conjunction with rapid mixing and quenching methods, such as acidification, has proven particularly useful also for the detection of reaction intermediates[385] .

1.7 Concluding Remarks and Future Perspective

We have summarized key questions that have emerged in the field of allostery in the last ~four-five years as well as the NMR approaches that have been proposed to address them (Table 1). The scope of these questions ranges from fundamental allosteric models to translational drug discovery. As the theories and applications of allostery continue to evolve, new questions will be defined and these will in turn inform the development on new experimental approaches. For example, the emerging concepts of driver *vs.* anchor groups[3,288,289] will require new methods to measure state-specific association constants. Another exemplary rising trend pertains to allosteric systems that cannot be adequately described by a single two-state transition as they are subject to "soft" allosteric couplings more appropriately described by the Ensemble Allosteric Model (EAM). The EAM will warrant the development of new techniques to measure transition and inter-state coupling energies that define the hierarchies of states in allosteric systems. Similarly, other emerging theories related to allostery, *e.g.* frustration[386–388] of free energy landscapes, will undoubtedly drive ingenuity in the experimental arena as well.

While there is no doubt that new transformative experimental methods will continue to be proposed to map allosteric sites, a promising general pattern that seems to have emerged in the last few years is one in which allosteric systems are interrogated through targeted libraries of perturbations, tailored to address the functional question at hand. The response to such perturbation libraries is then mapped by taking advantage of the resolution and sensitivity to dynamic processes typical of NMR. In this respect, it will be critical to integrate different "types" of NMR: ligand- and protein-based NMR methods as well as in vitro and in cell NMR data. The former are necessary to obtain the resolution and sensitivity required to validate quantitative allosteric models, while the latter are invaluable to recapitulate cellular complexity and contextualize allosteric mechanisms. It is clear that such integrative approaches should not be limited to NMR alone, but should extend to complementary experimental techniques, ranging from enzyme kinetics or other bio-assays needed to functionally profile allosteric perturbation libraries, to other biophysical

methods, such as fluorescence, DEER, CD, ITC and scattering methods, as illustrated by multiple examples in this review.

Last but not least, in order to fully reconstruct the free energy landscape within which allosteric systems function, experimental techniques, such as NMR, must be complemented through computational methods, such as molecular dynamics (MD) simulations, bioinformatics and thermodynamic models. Furthermore, *in silico* approaches offer access to allosteric transition pathways, *i.e.* transition states and transient intermediates, which may remain elusive to experimental characterization. However, computational methods alone are not sufficient, as experimental validation is needed. When computational and NMR methods are synergistically combined, the net result is a depth of understanding of the allosteric phenomenon that by far exceeds the sum of what could be achieved by using either approach independently of the other. In general, it will be critical to continue to integrate experimental data with simulations, which remain invaluable to connect the "experimental dots" and obtain a sharper picture of allostery.

1.8 Thesis Outlook

Structure determination methods are invaluable for understanding allosteric signalling. They provide an atomic resolution picture of the interaction between an allosteric effector and its target receptor, as well as the resulting conformational changes that propagate an allosteric perturbation to an orthosteric site. In section 1.4.1 we discuss traditional structure determination methods, such as x-ray crystallography, SAXS, NMR spectroscopy and cryo-EM, and provide examples of how they can be used to characterize allosteric sites and pathways. We also discuss some cases where traditional structure determination methods fail to capture important allosteric conformational changes and highlight the importance of employing an integrative approach that collects structural information from a variety of techniques to compensate for their weaknesses. However, some systems are not amenable to traditional structural determination methods. In some cases, there can be a complete absence of structural changes despite substantial differences in biochemical properties. In other cases, limited solubility, low ligand affinity and/or non-specific effects restrict experimental conditions to a window unsuitable for most traditional structural determination methods. The goal of this thesis is to develop methods and experimental strategies to elucidate the molecular mechanisms of allosteric perturbations in systems where traditional structural determination methods cannot be applied.

In chapter 2, I discuss the bradycardia-inducing S672R variant of the hyperpolarization-activated cyclic nucleotide-gated ion channel isoform 4 (HCN4)[399-401]. The mutation results in a loss-of-function phenotype, in which activation requires more negative (hyperpolarized) membrane voltages and channel deactivation occurs significantly faster[399-400]. However, X-ray crystallography and our own NMR analysis revealed that the mutation did not significantly alter the protein's structure[400-401]. Furthermore, the mutation does not significantly affect the affinity of HCN4's allosteric effector, cAMP[400-401]. The discrepancy between the altered electrophysiological phenotype and the lack of significant conformational changes was resolved by comparatively analyzing the free energy landscapes (FELs) of both wild type and S672R HCN4[401]. The FEL changes were mapped using a combination of NMR experiments, including saturation transfer difference (STD) to determine the free energies of cAMP binding, the chemical shift projection analysis (CHESPA) to measure the relative population of inactive and active states and NMR spin relaxation experiments to calculate the conformational entropy of the systems. Our dynamic analysis of the S672R mutant revealed that the mutation shifts the auto-inhibitory equilibrium towards the inactive state, increases the conformational entropy of the apo form and enhances internal dynamics at several sites surrounding the cAMP binding pocket. The latter two effects contribute to faster cAMP binding kinetics without significantly altering its affinity. Overall, the remodelled FEL of the S672R variant, rationalizes the change in electrophysiological phenotype that results in sinus bradycardia. The approaches we discuss for modeling the FEL of HCN and its S672R variant are applicable to disease-causing mutations in other systems, including those in which classical structure determination is challenging.

In chapter 3, I discuss our efforts to characterize the mechanism of action of the uncompetitive EPAC selective inhibitor, CE3F4R[57,58]. CE3F4R has poor solubility and forms soluble aggregates in aqueous solutions at high concentrations. The formation of these aggregates competitively interferes with binding of monomeric CE3F4R to EPAC, so the maximum CE3F4R concentration, which can be added to EPAC, is restricted within a specific window that minimalizes aggregate formation. Unfortunately, this workable window prevents the system from reaching saturation, resulting in a heterogenous sample of free and bound states, which is difficult to characterize by traditional structural determination approaches. In addition, the small ligand excess permitted by these workable conditions results in intermediate exchange broadening, which hampers the application of traditional NMR based structure determination experiments. Therefore, to solve the mechanism of this inhibitor's action, we adopted a chemical biologist's approach in which a library

of mutations and ligand derivatives were utilized to stabilize specific conformations of EPAC and the effect of CE3F4R on these distinct conformations was analyzed using simple and robust NMR methods, such as protein-based NH-HSQC spectra analyzed with the CHESPA. This strategy revealed that the inhibitor stabilizes a mixed-state intermediate that contains attributes of both inactive and active EPAC. The inactive conformations are stabilized in regions that regulate the activation of the enzyme. Furthermore, we identified the binding site of CE3F4R using a combination of chemical shifts, ligand-to-protein STD experiments and paramagnetic relaxation enhancement (PRE) experiments. Independently, each method provided only a rough estimate of the CE3F4R binding site, but when combined, the three methods identified a distinct binding pocket. Given that hydrophobic compounds similar to CE3F4R are common in commercially available screening libraries the simple and robust methods discussed in this chapter could be an effective strategy to characterize the mechanisms of hydrophobic drug leads when traditional structural determination methods cannot be applied.

The formation of soluble aggregates by hydrophobic compounds, as seen for CE3F4R, is a common challenge observed in drug screening[402-404]. The types of compounds that form such aggregates are often promiscuous towards a range of targets and may inhibit them through non-specific mechanisms, such as protein unfolding[402-404]. Hence, we have been following up the work of chapter 3, by investigating means to prevent aggregate formation. A common approach, applied in high throughput screening, utilizes small amounts of non-ionic detergents, such as triton X-100, to dissociate the aggregates[404]. However, we have observed that this approach can strip selective inhibitors away from their target protein and result in false negatives. In chapter 5.4, I discuss our progress on characterizing these detergent-inhibitor complexes and developing alternative approaches to reduce aggregation-based, non-specific interactions, while minimizing the perturbation of specific enzyme-inhibitor interactions.

Chapter 4 discusses the development of two new protocols for the chemical shift covariance analysis (CHESCA). The theory of CHESCA is discussed in section 1.4.2.3, but to summarize, it is a method that identifies networks of residues that respond in a concerted manner to a set of perturbations[115,120,226]. The set of perturbations is carefully selected to probe a specific equilibrium and to identify residues that participate in that equilibrium. Amide chemical shifts are used as a readout for the position of each state in the equilibrium and covariance between sets of residues is used to identify networks of residues that sense or contribute to that equilibrium. Since chemical shifts are exquisite sensors of protein structure and dynamics, the CHESCA can provide a detailed

map of residues involved in allosteric conformational changes. Amide chemical shifts are also robust and easy to measure in conformationally heterogenous samples, such as intrinsically disordered proteins or unsaturated receptors, where traditional structure determination methods remain challenging. For instance, Akimoto et al. used the CHESCA on protein kinase A (PKA) to discover that an intrinsically disordered linker is allosterically involved in the cAMP-modulated auto-inhibitory equilibrium[42]. This was a novel finding despite the availability of several X-ray crystal structures of the same region, in which the linker was often affected by crystal packing artefacts or poor electron density[42]. In the original iteration of the CHESCA, the identification of these amino acid networks was achieved using a single-linkage agglomerative clustering algorithm. However, we discovered using this algorithm that several residues did not display a concerted response to the same set of perturbations (*i.e.* it resulted in false positives). In chapter 4, we developed a diagnostic tool based on the CHESPA to identify false positives in the CHESCA and implemented two new versions of the CHESCA analysis with a reduced propensity for detecting false positives. The new iterations of the CHESCA utilize complete-linkage agglomerative clustering and independent CHESPA analyses of proton and nitrogen chemical shifts to overcome correlation chaining and chemical shift projection artefacts imposed by single-linkage clustering and linear combination of proton and nitrogen chemical shifts.

In the final chapter, I discuss the methodologies that were developed and implemented throughout this thesis to address questions about allosteric systems that are challenging to characterize with traditional structure determination techniques. The approaches are reviewed not just in the context of the systems described here, but also in terms of their potential application and their significance in the field of allostery. The HCN4 S672R mutation and EPAC inhibition by CE3F4R are also revisited to discuss the potential therapeutic applications that emerge from their newly discovered mechanisms. Finally, I discuss future perspectives for characterizing mechanisms of action for hydrophobic inhibitors that are unamenable to traditional structure determination.

1.9 Acknowledgements

We thank Professors G. Veglia (U. of Minnesota), Susan S. Taylor (UCSD), Jin Liu (UNTHSC), R. Nussinov (NCI-NIH), J. Ortega (McMaster) and M. Akimoto (Keio U.) as well as Dr. B. VanSchouwen and K. Moleschi (McMaster) for helpful discussions.

1.10 References

(1) Tsai, C. J.; Nussinov, R. A Unified View of "How Allostery Works". *PLoS Comput. Biol.* **2014**, *10* (2), e1003394.

(2) Nussinov, R.; Tsai, C. J. Allostery Without a Conformational Change? Revisiting the Paradigm. *Curr. Opin. Struct. Biol.* **2015**, *30*, 17–24.

(3) Nussinov, R.; Tsai, C. J.; Liu, J. Principles of Allosteric Interactions in Cell Signaling. *J. Am. Chem. Soc.* **2014**, *136* (51), 17692–17701.

(4) Nussinov, R.; Ma, B.; Tsai, C. J.; Csermely, P. Allosteric Conformational Barcodes Direct Signaling in the Cell. *Structure* **2013**, *21* (9), 1509–1521.

(5) Nussinov, R. How Do Dynamic Cellular Signals Travel Long Distances? *Mol. Biosyst.* **2012**, *8* (1), 22–26.

(6) Nussinov, R.; Tsai, C. J. Allostery in Disease and in Drug Discovery. *Cell* **2013**, *153* (2), 293–305.

(7) Gustavsson, M.; Verardi, R.; Mullen, D. G.; Mote, K. R.; Traaseth, N. J.; Gopinath, T.; Veglia, G. Allosteric Regulation of SERCA by Phosphorylation-mediated Conformational Shift of Phospholamban. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (43), 17338–17343.

(8) Sevcsik, E.; Trexler, A. J.; Dunn, J. M.; Rhoades, E. Allostery in a Disordered Protein: Oxidative Modifications to Alpha-Synuclein Act Distally To Regulate Membrane Binding. *J. Am. Chem. Soc.* **2011**, *133* (18), 7152–7158.

(9) Murciano-Calles, J.; Corbi-Verge, C.; Candel, A. M.; Luque, I.; Martinez, J. C.; Oliva, B. Post-Translational Modifications Modulate Ligand Recognition by the Third PDZ Domain of the MAGUK Protein PSD-95. *PloS one* **2014**, *9* (2), e90030.

(10) Skinner, J.; Ong, H.; Rosner, M.; Sosnick, T. Structural Characterization of a Phospho-Switch in the Metastasis Suppressor RKIP (LB189). *FASEB J.* **2014**, *28*, LB189-.

(11) Subedi, G. P.; Hanson, Q. M.; Barb, A. W. Restricted Motion of the Conserved Immunoglobulin G1 N-glycan is Essential for Efficient Fc-gammaRIIIa Binding. *Structure* **2014**, *22* (10), 1478–1488.

(12) Cannon, J. F. How Phosphorylation Activates the Protein Phosphatase-1 • Inhibitor-2 Complex. *Biochim. Biophys. Acta* **2013**, *1834* (1), 71–86.

(13) Bah, A.; Vernon, R. M.; Siddiqui, Z.; Krzeminski, M.; Muhandiram, R.; Zhao, C.; Sonenberg, N.; Kay, L. E.; Forman-Kay, J. D. Folding of an Intrinsically Disordered Protein by Phosphorylation as a Regulatory Switch. *Nature* **2015**, *519* (7541), 106–109.

(14) Zhang, J.; Petit, C. M.; King, D. S.; Lee, A. L. Phosphorylation of a PDZ Domain Extension Modulates Binding Affinity and Interdomain Interactions in Postsynaptic Density-95 (PSD-95) Protein, a Membrane-associated Guanylate Kinase(MAGUK). *J. Biol. Chem.* **2011**, *286* (48), 41776–41785.

(15) Boehr, D. D.; Schnell, J. R.; McElheny, D.; Bae, S.-H.; Duggan, B. M.; Benkovic, S. J.; Dyson, H. J.; Wright, P. E. A Distal Mutation Perturbs Dynamic Amino Acid Networks in Dihydrofolate Reductase. *Biochemistry* **2013**, *52* (27), 4605–4619.

(16) Starita, L. M.; Pruneda, J. N.; Lo, R. S.; Fowler, D. M.; Kim, H. J.; Hiatt, J. B.; Shendure, J.; Brzovic, P. S.; Fields, S.; Klevit, R. E. Activity-enhancing Mutations in an E3 Ubiquitin Ligase Identified by High-Throughput Mutagenesis. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (14), E1263-72.

(17) Abu-Abed, M.; Das, R.; Wang, L.; Melacini, G. Definition of an Electrostatic Relay Switch Critical for the cAMP-Dependent Activation of Protein Kinase A as Revealed by the D170A Mutant of RIalpha. *Proteins* **2007**, *69* (1), 112–124.

(18) Egesborg, P.; Carlettini, H.; Volpato, J. P.; Doucet, N. Combinatorial Active-site Variants Confer Sustained Clavulanate Resistance in BlaC β-lactamase from Mycobacterium tuberculosis. *Protein Sci.* **2015**, *24* (4), 534–544.

(19) Wals, P.-Y. de; Doucet, N.; Pelletier, J. N. High Tolerance to Simultaneous Active-site Mutations in TEM-1 Beta-Lactamase: Distinct Mutational Paths Provide More Generalized Beta-lactam Recognition. *Protein Sci.* **2009**, *18* (1), 147–160.

(20) Westman, E. L.; Yan, M.; Waglechner, N.; Koteva, K.; Wright, G. D. Self Resistance to the Atypical Cationic Antimicrobial Peptide Edeine of Brevibacillus brevis Vm4 by the N-acetyltransferase EdeQ. *Chem. Biol.* **2013**, *20* (8), 983–990.

(21) Smock, R. G.; Gierasch, L. M. Sending Signals Dynamically. *Science* **2009**, *324* (5924), 198–203.

(22) Li, J.; Ilangovan, U.; Daubner, S. C.; Hinck, A. P.; Fitzpatrick, P. F. Direct Evidence For a Phenylalanine Site in the Regulatory Domain of Phenylalanine Hydroxylase. *Arch. Biochem. Biophys.* **2011**, *505* (2), 250–255.

(23) Taylor, S. S.; Ilouz, R.; Zhang, P.; Kornev, A. P. Assembly of allosteric macromolecular switches: lessons from PKA. *Nat. Rev. Mol. Cell Biol.* **2012**, *13* (10), 646–658.

(24) Taylor, S. S.; Zhang, P.; Steichen, J. M.; Keshwani, M. M.; Kornev, A. P. PKA: Lessons Learned After Twenty Years. *Biochim. Biophys. Acta* **2013**, *1834* (7), 1271–1278.

(25) Zhuravleva, A.; Clerico, E. M.; Gierasch, L. M. An Interdomain Energetic Tug-of-war Creates the Allosterically Active State in Hsp70 Molecular Chaperones. *Cell* **2012**, *151* (6), 1296–1307.

(26) Gao, X.-C.; Zhou, C.-J.; Zhou, Z.-R.; Wu, M.; Cao, C.-Y.; Hu, H.-Y. The C-terminal Helices of Heat Shock Protein 70 Are Essential For J-domain Binding and ATPase Activation. *J. Biol. Chem.* **2012**, *287* (8), 6044–6052.

(27) Ruschak, A. M.; Kay, L. E. Proteasome allostery as a population shift between interchanging conformers. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (50), E3454-62.

(28) Estrada, D. F.; Laurence, J. S.; Scott, E. E. Substrate-modulated Cytochrome P450 17A1 and Cytochrome B5 Interactions Revealed By NMR. *J. Biol. Chem.* **2013**, *288* (23), 17008–17018.

(29) Dutta, K.; Varshney, A. K.; Franklin, M. C.; Goger, M.; Wang, X.; Fries, B. C. Mechanisms Mediating Enhanced Neutralization Efficacy of Staphylococcal Enterotoxin B by Combinations of Monoclonal Antibodies. *J. Biol. Chem.* **2015**, *290* (11), 6715–6730.

(30) Foster, T. L.; Thompson, G. S.; Kalverda, A. P.; Kankanala, J.; Bentham, M.; Wetherill, L. F.; Thompson, J.; Barker, A. M.; Clarke, D.; Noerenberg, M., et al. Structure-Guided Design Affirms Inhibitors of Hepatitis C Virus p7 as a Viable Class of Antivirals Targeting Virion Release. *Hepatology (Baltimore, Md.)* **2014**, *59* (2), 408–422.

(31) Christopoulos, A.; Changeux, J.-P.; Catterall, W. A.; Fabbro, D.; Burris, T. P.; Cidlowski, J. A.; Olsen, R. W.; Peters, J. A.; Neubig, R. R.; Pin, J.-P., et al. International Union of Basic and Clinical Pharmacology. XC. multisite pharmacology: recommendations for the nomenclature of receptor allosterism and allosteric ligands. *Pharmacol. Rev.* **2014**, *66* (4), 918–947.

(32) Udi, Y.; Fragai, M.; Grossman, M.; Mitternacht, S.; Arad-Yellin, R.; Calderone, V.; Melikian, M.; Toccafondi, M.; Berezovsky, I. N.; Luchinat, C., et al. Unraveling Hidden Regulatory Sites in Structurally Homologous Metalloproteases. *J. Mol. Biol.* **2013**, *425* (13), 2330–2346.

(33) Christopoulos, A. Allosteric Binding Sites on Cell-surface Receptors: Novel Targets for Drug Discovery. *Nat. Rev. Drug Discovery* **2002**, *1* (3), 198–210.

(34) Nussinov, R.; Tsai, C. J. The Design of Covalent Allosteric Drugs. *Annu. Rev. Pharmacol. Toxicol.* **2015**, *55*, 249–267.

(35) Smith, R. E.; Tran, K.; Richards, K. M.; Luo, R. Using NMR to Develop New Allosteric and Allo-Network Drugs. *Curr. Drug Discov. Technol.* **2015**, *12* (4), 193–204.

(36) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. The ensemble nature of allostery. *Nature* **2014**, *508* (7496), 331–339.

(37) Baldwin, A. J.; Kay, L. E. Nmr Spectroscopy Brings Invisible Protein States into Focus. *Nat. Chem. Biol.* **2009**, *5* (11), 808–814.

(38) Mittermaier, A.; Kay, L. E. New tools provide new insights in NMR studies of protein dynamics. *Science* **2006**, *312* (5771), 224–228.

(39) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. The Energy Landscapes and Motions of Proteins. *Science (New York, N.Y.)* **1991**, *254* (5038), 1598–1603.

(40) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis. *Proteins* **1995**, *21* (3), 167–195.

(41) Akimoto, M.; Moleschi, K.; Boulton, S.; VanSchouwen, B.; Selvaratnam, R.; Taylor, S. S.; Melacini, G. Allosteric Linkers in cAMP Signalling. *Biochem. Soc. Trans.* **2014**, *42* (1), 139–144.

(42) Akimoto, M.; Selvaratnam, R.; McNicholl, E. T.; Verma, G.; Taylor, S. S.; Melacini, G. Signaling Through Dynamic Linkers as Revealed by PKA. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (35), 14231–14236.

(43) Gardino, A. K.; Villali, J.; Kivenson, A.; Lei, M.; Liu, C. F.; Steindel, P.; Eisenmesser, E. Z.; Labeikovsky, W.; Wolf-Watz, M.; Clarkson, M. W., et al. Transient non-native hydrogen bonds promote activation of a signaling protein. *Cell* **2009**, *139* (6), 1109–1118.

(44) Niu, W.-N.; Yadav, P. K.; Adamec, J.; Banerjee, R. S-glutathionylation enhances human cystathionine beta-synthase activity under oxidative stress conditions. *Antioxid. Redox. Signal.* **2015**, *22* (5), 350–361.

(45) Desjardins, G.; Meeker, C. A.; Bhachech, N.; Currie, S. L.; Okon, M.; Graves, B. J.; McIntosh, L. P. Synergy of Aromatic Residues and Phosphoserines Within the Intrinsically Disordered DNA-binding Inhibitory Elements of the Ets-1 Transcription Factor. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* (30), 11019–11024.

(46) Monod, J.; Wyman, J.; Changeux, J.-P. On the nature of allosteric transitions: A plausible model. *J. Mol. Biol.* **1965**, *12* (1), 88–118.

(47) Moleschi, K. J.; Akimoto, M.; Melacini, G. Measurement of State-Specific Association Constants in Allosteric Sensors through Molecular Stapling and NMR. *J. Am. Chem. Soc.* **2015**, *137* (33), 4094–4113.

(48) Koshland, D. E.; Némethy, G.; Filmer, D. Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits *. *Biochemistry* **1966**, *5* (1), 365–385.

(49) Long, D.; Bruschweiler, R. Atomistic Kinetic Model for Population Shift and Allostery in Biomolecules. *J. Am. Chem. Soc.* **2011**, *133* (46), 18999–19005.

(50) LeMaster, D. M.; Mustafi, S. M.; Brecher, M.; Zhang, J.; Héroux, A.; Li, H.; Hernández, G. Coupling of Conformational Transitions in the N-terminal Domain of the 51-kDa FK506-binding Protein (FKBP51) Near Its Site of Interaction with the Steroid Receptor Proteins. *J. Biol. Chem.* **2015**, *290* (25), 15746–15757.

(51) Changeux, J. P.; Danchin, A. Selective Stabilisation of Developing Synapses as a Mechanism for the Specification of Neuronal Networks. *Nature* **1976**, *264* (5588), 705–712.

(52) Boehr, D. D.; Nussinov, R.; Wright, P. E. The Role of Dynamic Conformational Ensembles in Biomolecular Recognition. *Nat. Chem. Biol.* **2009**, *5* (11), 789–796.

(53) VanSchouwen, B.; Selvaratnam, R.; Fogolari, F.; Melacini, G. Role of dynamics in the autoinhibition and activation of the exchange protein directly activated by cyclic AMP (EPAC). *J. Biol. Chem.* **2011**, *286* (49), 42655–42669.

(54) Malmstrom, R. D.; Kornev, A. P.; Taylor, S. S.; Amaro, R. E. Allostery Through the Computational Microscope: cAMP Activation of a Canonical Signalling Domain. *Nat. Commun.* **2015**, *6*, 7588.

(55) VanSchouwen, B.; Akimoto, M.; Sayadi, M.; Fogolari, F.; Melacini, G. Role of Dynamics in the Autoinhibition and Activation of the Hyperpolarization-activated Cyclic Nucleotide-modulated (HCN) Ion Channels. *J. Biol. Chem.* **2015**, *290* (29), 17642–17654.

(56) Hammes, G. G.; Chang, Y.-C.; Oas, T. G. Conformational selection or induced fit: a flux description of reaction mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106* (33), 13737–13741.

(57) Courilleau, D.; Bisserier, M.; Jullian, J.-C.; Lucas, A.; Bouyssou, P.; Fischmeister, R.; Blondeau, J.-P.; Lezoualc'h, F. Identification of a Tetrahydroquinoline Analog as a Pharmacological Inhibitor of the cAMP-binding Protein Epac. *J. Biol. Chem.* **2012**, *287* (53), 44192–44202.

(58) Courilleau, D.; Bouyssou, P.; Fischmeister, R.; Lezoualc'h, F.; Blondeau, J.-P. The (R)-enantiomer of CE3F4 is a Preferential Inhibitor of Human Exchange Protein Directly Activated by Cyclic AMP Isoform 1 (Epac1). *Biochem. Biophys. Res. Commun.* **2013**, *440* (3), 443–448.

(59) Weinkam, P.; Pons, J.; Sali, A. Structure-based model of allostery predicts coupling between distant sites. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (13), 4875–4880.

(60) Popovych, N.; Sun, S.; Ebright, R. H.; Kalodimos, C. G. Dynamically Driven Protein Allostery. *Nat. Struct. Mol. Biol.* **2006**, *13* (9), 831–838.

(61) Tzeng, S.-R.; Kalodimos, C. G. Protein Activity Regulation by Conformational Entropy. *Nature* **2012**, *488* (7410), 236–240.

(62) Moorman, V. R.; Valentine, K. G.; Bédard, S.; Kasinath, V.; Dogan, J.; Love, F. M.; Wand, A. J. Dynamic and thermodynamic response of the Ras protein Cdc42Hs upon association with the effector domain of PAK3. *J. Mol. Biol.* **2014**, *426* (21), 3520–3538.

(63) Laine, J. M.; Amat, M.; Morgan, B. R.; Royer, W. E.; Massi, F. Insight into the allosteric mechanism of Scapharca dimeric hemoglobin. *Biochemistry* **2014**, *53* (46), 7199–7210.

(64) Law, S. M.; Gagnon, J. K.; Mapp, A. K.; Brooks, C. L. Prepaying the entropic cost for allosteric regulation in KIX. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* (33), 12067–12072.

(65) Takeuchi, K.; Tokunaga, Y.; Imai, M.; Takahashi, H.; Shimada, I. Dynamic multidrug recognition by multidrug transcriptional repressor LmrR. *Sci. Rep.* **2014**, *4*, 6922.

(66) Regan, M. C.; Horanyi, P. S.; Pryor, E. E.; Sarver, J. L.; Cafiso, D. S.; Bushweller, J. H. Structural and dynamic studies of the transcription factor ERG reveal DNA binding is allosterically autoinhibited. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (33), 13374–13379.

(67) Wand, A. J. The Dark Energy of Proteins Comes to Light: Conformational Entropy and its Role in Protein Function Revealed by NMR Relaxation. *Curr. Opin. Struct. Biol.* **2013**, *23* (1), 75–81.

(68) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. Conformational Entropy in Molecular Recognition by Proteins. *Nature* **2007**, *448* (7151), 325–329.

(69) Lee, A. L.; Wand, A. J. Microscopic Origins of Entropy, Heat Capacity and the Glass Transition in Proteins. *Nature* **2001**, *411* (6836), 501–504.

(70) Lee, A. L.; Kinnear, S. A.; Wand, A. J. Redistribution and loss of side chain entropy upon formation of a calmodulin-peptide complex. *Nat. Struct. Biol.* **2000**, *7* (1), 72–77.

(71) Arumugam, S.; Gao, G.; Patton, B. L.; Semenchenko, V.; Brew, K.; van Doren, S. R. Increased Backbone Mobility in Beta-Barrel Enhances Entropy Gain Driving Binding of N-TIMP-1 to MMP-3. *J. Mol. Biol.* **2003**, *327* (3), 719–734.

(72) Masterson, L. R.; Shi, L.; Metcalfe, E.; Gao, J.; Taylor, S. S.; Veglia, G. Dynamically Committed, Uncommitted, and Quenched States Encoded in Protein Kinase A Revealed by NMR Spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (17), 6969–6974.

(73) Zhuravleva, A.; Gierasch, L. M. Substrate-binding Domain Conformational Dynamics Mediate Hsp70 Allostery. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112* (22), E2865-73.

(74) Berlow, R. B.; Swain, M.; Dalal, S.; Sweasy, J. B.; Loria, J. P. Substrate-Dependent Millisecond Domain Motions in DNA Polymerase Beta. *J. Mol. Biol.* **2012**, *419* (3-4), 171–182.

(75) Zayner, J. P.; Antoniou, C.; French, A. R.; Hause, R. J.; Sosnick, T. R. Investigating Models of Protein Function and Allostery With a Widespread Mutational Analysis of a Light-activated Protein. *Biophys. J.* **2013**, *105* (4), 1027–1036.

(76) Lechtenberg, B. C.; Freund, S. M.; Huntington, J. A. An Ensemble View of Thrombin Allostery. *Biol. Chem.* **2012**, *393* (9), 889–898.

(77) Kimlicka, L.; van Petegem, F. The Structural Biology of Ryanodine Receptors. *Sci. China Life Sci.* **2011**, *54* (8), 712–724.

(78) Maler, L.; Blankenship, J.; Rance, M.; Chazin, W. J. Site-site communication in the EF-hand Ca2+-binding protein calbindin D9k. *Nat. Struct. Biol.* **2000**, *7* (3), 245–250.

(79) Whitley, M. J.; Lee, A. L. Exploring the role of structure and dynamics in the function of chymotrypsin inhibitor 2. *Proteins* **2011**, *79* (3), 916–924.

(80) Whittier, S. K.; Hengge, A. C.; Loria, J. P. Conformational motions regulate phosphoryl transfer in related protein tyrosine phosphatases. *Science* **2013**, *341* (6148), 899–903.

(81) Kern, D.; Zuiderweg, E. R. P. The role of dynamics in allosteric regulation. *Curr. Opin. Struct. Biol.* **2003**, *13* (6), 748–757.

(82) Oyen, D.; Fenwick, R. B.; Stanfield, R. L.; Dyson, H. J.; Wright, P. E. Cofactor-Mediated Conformational Dynamics Promote Product Release From Escherichia coli Dihydrofolate Reductase via an Allosteric Pathway. *J. Am. Chem. Soc.* **2015**, *137* (29), 9459–9468.

(83) Lipchock, J. M.; Loria, J. P. Nanometer Propagation of Millisecond Motions in V-type Allostery. *Structure* **2010**, *18* (12), 1596–1607.

(84) Cho, M.-K.; Gayen, A.; Banigan, J. R.; Leninger, M.; Traaseth, N. J. Intrinsic Conformational Plasticity of Native EmrE Provides a Pathway for Multidrug Resistance. *J. Am. Chem. Soc.* **2014**, *136* (22), 8072–8080.

(85) Larion, M.; Salinas, R. K.; Bruschweiler-Li, L.; Miller, B. G.; Brüschweiler, R. Order-disorder transitions govern kinetic cooperativity and allostery of monomeric human glucokinase. *PLoS Biol.* **2012**, *10* (12), e1001452.

(86) Palmer, A. G. Enzyme Dynamics From NMR Spectroscopy. *Acc. Chem. Res.* **2015**, *48* (2), 457–465.

(87) Kokkinidis, M.; Glykos, N. M.; Fadouloglou, V. E. Protein Flexibility and Enzymatic Catalysis. *Adv. Protein Chem.* **2012**, *87*, 181–218.

(88) Buhrman, G.; O'Connor, C.; Zerbe, B.; Kearney, B. M.; Napoleon, R.; Kovrigina, E. A.; Vajda, S.; Kozakov, D.; Kovrigin, E. L.; Mattos, C. Analysis of Binding Site Hot Spots on the Surface of Ras GTPase. *J. Mol. Biol.* **2011**, *413* (4), 773–789.

(89) Kalodimos, C. G. NMR reveals novel mechanisms of protein activity regulation. *Protein Sci.* **2011**, *20* (5), 773–782.

(90) Kalodimos, C. G. Protein function and allostery: a dynamic relationship. *Ann. N.Y. Acad. Sci.* **2012**, *1260*, 81–86.

(91) Hilser, V. J. Structural biology: Signalling from disordered proteins. *Nature* **2013**, *498* (7454), 308–310.

(92) Bista, M.; Petrovich, M.; Fersht, A. R. MDMX Contains an Autoinhibitory Sequence Element. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (44), 17814–17819.

(93) Yan, Y.; McCallum, S. A.; Wang, C. M35 oxidation induces Abeta40-like structural and dynamical changes in Abeta42. *J. Am. Chem. Soc.* **2008**, *130* (16), 5394–5395.

(94) Parigi, G.; Rezaei-Ghaleh, N.; Giachetti, A.; Becker, S.; Fernandez, C.; Blackledge, M.; Griesinger, C.; Zweckstetter, M.; Luchinat, C. Long-range correlated dynamics in intrinsically disordered proteins. *J. Am. Chem. Soc.* **2014**, *136* (46), 16201–16209.

(95) Ferreon, A. C. M.; Ferreon, J. C.; Wright, P. E.; Deniz, A. A. Modulation of Allostery by Protein Intrinsic Disorder. *Nature* **2013**, *498* (7454), 390–394.

(96) Manley, G.; Rivalta, I.; Loria, J. P. Solution NMR and computational methods for understanding protein allostery. *J. Phys. Chem* **2013**, *117* (11), 3063–3073.

(97) Tzeng, S.-R.; Kalodimos, C. G. Protein dynamics and allostery: an NMR view. *Curr. Opin. Struct. Biol.* **2011**, *21* (1), 62–67.

(98) Masterson, L. R.; Cembran, A.; Shi, L.; Veglia, G. Allostery and Binding Cooperativity of the Catalytic Subunit of Protein Kinase A by NMR Spectroscopy and Molecular Dynamics Simulations. *Adv. Protein Chem.* **2012**, *87*, 363–389.

(99) Manley, G.; Loria, J. P. NMR insights into protein allostery. *Arch. Biochem. Biophys.* **2012**, *519* (2), 223–231.

(100) Karanth, N. M.; Sarma, S. P. The coil-to-helix transition in IlvN regulates the allosteric control of Escherichia coli acetohydroxyacid synthase I. *Biochemistry* **2013**, *52* (1), 70–83.

(101) Schünke, S.; Stoldt, M.; Lecher, J.; Kaupp, U. B.; Willbold, D. Structural Insights Into Conformational Changes of a Cyclic Nucleotide-Binding Domain in Solution From Mesorhizobium loti K1 Channel. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (15), 6121–6126.

(102) Samanta, S.; Situ, A. J.; Ulmer, T. S. Structural characterization of the regulatory domain of brain carnitine palmitoyltransferase 1. *Biopolymers* **2014**, *101* (4), 398–405.

(103) Saponaro, A.; Pauleta, S. R.; Cantini, F.; Matzapetakis, M.; Hammann, C.; Donadoni, C.; Hu, L.; Thiel, G.; Banci, L.; Santoro, B., et al. Structural basis for the mutual antagonism of cAMP and TRIP8b in regulating HCN channel function. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* (40), 14577–14582.

(104) Wiedemann, C.; Szambowska, A.; Häfner, S.; Ohlenschläger, O.; Gührs, K.-H.; Görlach, M. Structure and regulatory role of the C-terminal winged helix domain of the archaeal minichromosome maintenance complex. *Nucleic Acids Res.* **2015**, *43* (5), 2958–2967.

(105) Zhang, S.; Huang, T.; Ilangovan, U.; Hinck, A. P.; Fitzpatrick, P. F. The Solution Structure of the Regulatory Domain of Tyrosine Hydroxylase. *J. Mol. Biol.* **2014**, *426* (7), 1483–1497.

(106) Das, R.; Liang, Y.-H.; Mariano, J.; Li, J.; Huang, T.; King, A.; Tarasov, S. G.; Weissman, A. M.; Ji, X.; Byrd, R. A. Allosteric Regulation of E2:E3 Interactions Promote a Processive Ubiquitination Machine. *EMBO J.* **2013**, *32* (18), 2504–2516.

(107) Miller, T. C. R.; Mieszczanek, J.; Sánchez-Barrena, M.; Rutherford, T. J.; Fiedler, M.; Bienz, M. Evolutionary adaptation of the fly Pygo PHD finger toward recognizing histone H3 tail methylated at arginine 2. *Structure* **2013**, *21* (12), 2208–2220.

(108) Mowrey, D. D.; Cui, T.; Jia, Y.; Ma, D.; Makhov, A. M.; Zhang, P.; Tang, P.; Xu, Y. Open-channel structures of the human glycine receptor \alpha1 full-length transmembrane domain. *Structure* **2013**, *21* (10), 1897–1904.

(109) Rana, R.; Carroll, C. E.; Lee, H.-J.; Bao, J.; Marada, S.; Grace, C. R. R.; Guibao, C. D.; Ogden, S. K.; Zheng, J. J. Structural Insights Into the Role of the Smoothened Cysteine-Rich Domain in Hedgehog Signalling. *Nat. Commun.* **2013**, *4*, 2965.

(110) Blech, M.; Peter, D.; Fischer, P.; Bauer, M. M. T.; Hafner, M.; Zeeb, M.; Nar, H. One Target-Two Different Binding Modes: Structural Insights into Gevokizumab and Canakinumab Interactions to Interleukin-1 Beta. *J. Mol. Biol.* **2013**, *425* (1), 94–111.

(111) Chakravorty, D. K.; Wang, B.; Lee, C. W.; Guerra, A. J.; Giedroc, D. P.; Merz, K. M. Solution NMR Refinement of a Metal Ion Bound Protein Using Metal Ion Inclusive Restrained Molecular Dynamics Methods. *J. Biolmol. NMR* **2013**, *56* (2), 125–137.

(112) Tanaka, H.; Akagi, K.-i.; Oneyama, C.; Tanaka, M.; Sasaki, Y.; Kanou, T.; Lee, Y.-H.; Yokogawa, D.; Dobenecker, M.-W.; Nakagawa, A., et al. Identification of a new interaction mode between the Src homology 2 domain of C-terminal Src kinase (Csk) and Csk-binding protein/phosphoprotein associated with glycosphingolipid microdomains. *J. Biol. Chem.* **2013**, *288* (21), 15240–15254.

(113) Siddiqui, N.; Tempel, W.; Nedyalkova, L.; Volpon, L.; Wernimont, A. K.; Osborne, M. J.; Park, H.-W.; Borden, K. L. B. Structural insights into the allosteric effects of 4EBP1 on the eukaryotic translation initiation factor eIF4E. *J. Mol. Biol.* **2012**, *415* (5), 781–792.

(114) Pruneda, J. N.; Littlefield, P. J.; Soss, S. E.; Nordquist, K. A.; Chazin, W. J.; Brzovic, P. S.; Klevit, R. E. Structure of an E3:E2~Ub complex reveals an allosteric mechanism shared among RING/U-box ligases. *Mol. Cell* **2012**, *47* (6), 933–942.

(115) Selvaratnam, R.; Chowdhury, S.; VanSchouwen, B.; Melacini, G. Mapping Allostery Through the Covariance Analysis of NMR Chemical Shifts. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (15), 6133–6138.

(116) Romanowska, J.; Nowiński, K. S.; Trylska, J. Determining Geometrically Stable Domains in Molecular Conformation Sets. *J. Chem. Theory Comput.* **2012**, *8* (8), 2588–2599.

(117) Taylor, D.; Cawley, G.; Hayward, S. Quantitative Method for the Assignment of Hinge and Shear Mechanism in Protein Domain Movements. *Bioinformatics* **2014**, *30* (22), 3189–3196.

(118) Lee, R. A.; Razaz, M.; Hayward, S. The DynDom database of protein domain motions. *Bioinformatics* **2003**, *19* (10), 1290–1291.

(119) Rehmann, H.; Arias-Palomo, E.; Hadders, M. A.; Schwede, F.; Llorca, O.; Bos, J. L. Structure of Epac2 in complex with a cyclic AMP analogue and RAP1B. *Nature* **2008**, *455* (7209), 124–127.

(120) Selvaratnam, R.; Mazhab-Jafari, M. T.; Das, R.; Melacini, G. The auto-inhibitory role of the EPAC hinge helix as mapped by NMR. *PloS one* **2012**, *7* (11), e48707.

(121) Selvaratnam, R.; VanSchouwen, B.; Fogolari, F.; Mazhab-Jafari, M. T.; Das, R.; Melacini, G. The Projection Analysis of NMR Chemical Shifts Reveals Extended EPAC Autoinhibition Determinants. *Biophys. J.* **2012**, *102* (3), 630–639.

(122) Cronin, M.; Coolbaugh, M. J.; Nellis, D.; Zhu, J.; Wood, D. W.; Nussinov, R.; Ma, B. Dynamics Differentiate Between Active and Inactive Inteins. *Eur. J. Med. Chem.* **2015**, *91*, 51–62.

(123) Akimoto, M.; Zhang, Z.; Boulton, S.; Selvaratnam, R.; VanSchouwen, B.; Gloyd, M.; Accili, E. A.; Lange, O. F.; Melacini, G. A Mechanism for the Auto-inhibition of Hyperpolarization-activated Cyclic

Nucleotide-gated (HCN) Channel Opening and its Relief by cAMP. *J. Biol. Chem.* **2014**, *289* (32), 22205–22220.

(124) Baruscotti, M.; Barbuti, A.; Bucchi, A. The Cardiac Pacemaker Current. *J. Mol. Cell. Cardiol.* **2010**, *48* (1), 55–64.

(125) Craven, K. B.; Zagotta, W. N. CNG and HCN Channels: Two Peas, One Pod. *Annu. Rev. Physiol.* **2006**, *68*, 375–401.

(126) Zagotta, W. N.; Olivier, N. B.; Black, K. D.; Young, E. C.; Olson, R.; Gouaux, E. Structural Basis for Modulation and Agonist Specificity of HCN Pacemaker Channels. *Nature* **2003**, *425* (6954), 200–205.

(127) Taraska, J. W.; Puljung, M. C.; Olivier, N. B.; Flynn, G. E.; Zagotta, W. N. Mapping the Structure and Conformational Movements of Proteins with Transition Metal Ion FRET. *Nat. Methods* **2009**, *6* (7), 532–537.

(128) Zhuravleva, A.; Gierasch, L. M. Allosteric Signal Transmission in the Nucleotide-binding Domain of 70-kDa Heat Shock Protein (Hsp70) Molecular Chaperones. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (17), 6987–6992.

(129) Fernández-de-Alba, C.; Berrow, N. S.; Garcia-Castellanos, R.; Garcús; Pons, M. On the Origin of the Selectivity of Plasmidic H-NS Towards Horizontally Acquired DNA: Linking H-NS Oligomerization and Cooperative DNA Binding. *J. Mol. Biol.* **2013**, *425* (13), 2347–2358.

(130) Piserchio, A.; Warthaka, M.; Devkota, A. K.; Kaoud, T. S.; Lee, S.; Abramczyk, O.; Ren, P.; Dalby, K. N.; Ghose, R. Solution NMR insights into docking interactions involving inactive ERK2. *Biochemistry* **2011**, *50* (18), 3660–3672.

(131) Tripathy, C.; Zeng, J.; Zhou, P.; Donald, B. R. Protein Loop Closure Using Orientational Restraints from NMR Data. *Proteins* **2012**, *80* (2), 433–453.

(132) Brosey, C. A.; Soss, S. E.; Brooks, S.; Yan, C.; Ivanov, I.; Dorai, K.; Chazin, W. J. Functional Dynamics in Replication Protein a DNA Binding and Protein Recruitment Domains. *Structure* **2015**, *23* (6), 1028–1038.

(133) De, S.; Chan, A. C. K.; Coyne, H. J. 3.; Bhachech, N.; Hermsdorf, U.; Okon, M.; Murphy, M. E. P.; Graves, B. J.; McIntosh, L. P. Steric Mechanism of Auto-inhibitory Regulation of Specific and Non-specific DNA Binding by the ETS Transcriptional Repressor ETV6. *J. Mol. Biol.* **2014**, *426* (7), 1390–1406.

(134) Su, Y.; Dostmann, W.; Herberg, F. W.; Durick, K.; Xuong, N. H.; Eyck, L. ten; Taylor, S. S.; Varughese, K. I. Regulatory Subunit of Protein Kinase A: Structure of Deletion Mutant with cAMP Binding Domains. *Science (New York, N.Y.)* **1996**, *269* (5225), 807–813.

(135) Badireddy, S.; Yunfeng, G.; Ritchie, M.; Akamine, P.; Wu, J.; Kim, C. W.; Taylor, S. S.; Qingsong, L.; Swaminathan, K.; Anand, G. S. Cyclic AMP Analog Blocks Kinase Activation by Stabilizing Inactive Conformation: Conformational Selection Highlights a New Concept in Allosteric Inhibitor Design. *Mol. Cell. Proteomics* **2011**, *10* (3), M110.004390.

(136) Knighton, D. R.; Zheng, J. H.; Eyck, L. F. ten; Xuong, N. H.; Taylor, S. S.; Sowadski, J. M. Structure of a Peptide Inhibitor Bound to the Catalytic Subunit of Cyclic Adenosine Monophosphate-dependent Protein Kinase. *Science* **1991**, *253* (5018), 414–420.

(137) Anand, G. S.; Krishnamurthy, S.; Bishnoi, T.; Kornev, A.; Taylor, S. S.; Johnson, D. A. Cyclic AMP- and (Rp)-cAMPS-Induced Conformational Changes in a Complex of the Catalytic and Regulatory (RIalpha) Subunits of Cyclic AMP-Dependent Protein Kinase. *Mol. Cell. Proteomics* **2010**, *9* (10), 2225–2237.

(138) Harada, E.; Sugishima, M.; Harada, J.; Fukuyama, K.; Sugase, K. Distal regulation of heme binding of heme oxygenase-1 mediated by conformational fluctuations. *Biochemistry* **2015**, *54* (2), 340–348.

(139) Brüschweiler, S.; Konrat, R.; Tollinger, M. Allosteric Communication in the KIX Domain Proceeds Through Dynamic Repacking of the Hydrophobic Core. *ACS Chem. Biol.* **2013**, *8* (7), 1600–1610.

(140) Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.

(141) Shen, Y.; Bax, A. Homology Modeling of Larger Proteins Guided by Chemical Shifts. *Nat. Methods* **2015**, *12* (8), 747–750.

(142) Shen, Y.; Bax, A. Protein structural information derived from NMR chemical shift with the neural network program TALOS-N. *Methods Mol. Biol.* **2015**, *1260*, 17–32.

(143) Berjanskii, M.; Arndt, D.; Liang, Y.; Wishart, D. S. A Robust Algorithm for Optimizing Protein Structures with NMR Chemical Shifts. *J. Biolmol. NMR* **2015**, *63* (3), 255–264.

(144) Berjanskii, M. V.; Wishart, D. S. A Simple Method to Measure Protein Side-chain Mobility Using NMR Chemical Shifts. *J. Am. Chem. Soc.* **2013**, *135* (39), 14536–14539.

(145) Boomsma, W.; Tian, P.; Frellsen, J.; Ferkinghoff-Borg, J.; Hamelryck, T.; Lindorff-Larsen, K.; Vendruscolo, M. Equilibrium Simulations of Proteins Using Molecular Fragment Replacement and NMR Chemical Shifts. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* (38), 13852–13857.

(146) Cavalli, A.; Salvatella, X.; Dobson, C. M.; Vendruscolo, M. Protein Structure Determination from NMR Chemical Shifts. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (23), 9615–9620.

(147) Rivalta, I.; Sultan, M. M.; Lee, N.-S.; Manley, G. A.; Loria, J. P.; Batista, V. S. Allosteric pathways in imidazole glycerol phosphate synthase. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (22), E1428-36.

(148) Karagöz, G. E.; Duarte, A. M. S.; Ippel, H.; Uetrecht, C.; Sinnige, T.; van Rosmalen, M.; Hausmann, J.; Heck, A. J. R.; Boelens, R.; Rüdiger, S. G. D. N-terminal domain of human Hsp90 triggers binding to the cochaperone p23. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (2), 580–585.

(149) Farina, B.; Doti, N.; Pirone, L.; Malgieri, G.; Pedone, E. M.; Ruvo, M.; Fattorusso, R. Molecular Basis of the PED/PEA15 Interaction with the C-terminal Fragment of Phospholipase D1 Revealed by NMR Spectroscopy. *Biochim. Biophys. Acta, Proteins Proteomics* **2013**, *1834* (8), 1572–1580.

(150) Lee, D.-H.; Ha, J.-H.; Kim, Y.; Jang, M.; Park, S. J.; Yoon, H. S.; Kim, E.-H.; Bae, K.-H.; Park, B. C.; Park, S. G., et al. A Conserved Mechanism for Binding of p53 DNA-Binding Domain and Anti-Apoptotic Bcl-2 Family Proteins. *Mol. Cells* **2014**, *37* (3), 264–269.

(151) Oyen, D.; Wechselberger, R.; Srinivasan, V.; Steyaert, J.; Barlow, J. N. Mechanistic analysis of allosteric and non-allosteric effects arising from nanobody binding to two epitopes of the dihydrofolate reductase of Escherichia coli. *Biochim. Biophys. Acta* **2013**, *1834* (10), 2147–2157.

(152) Takacs, M.; Petoukhov, M. V.; Atkinson, R. A.; Roblin, P.; Ogi, F.-X.; Demeler, B.; Potier, N.; Chebaro, Y.; Dejaegere, A.; Svergun, D. I., et al. The asymmetric binding of PGC-1\alpha to the ERR\alpha and ERR\gamma nuclear receptor homodimers involves a similar recognition mechanism. *PloS one* **2013**, *8* (7), e67810.

(153) Tanio, M.; Nishimura, K. Intramolecular allosteric interaction in the phospholipase C-\delta1 pleckstrin homology domain. *Biochim. Biophys. Acta* **2013**, *1834* (6), 1034–1043.

(154) Watson, L. C.; Kuchenbecker, K. M.; Schiller, B. J.; Gross, J. D.; Pufall, M. A.; Yamamoto, K. R. The glucocorticoid receptor dimer interface allosterically transmits sequence-specific DNA signals. *Nat. Struct. Mol. Biol.* **2013**, *20* (7), 876–883.

(155) Borthakur, S.; Hibbert, R. G.; Pang, M. O. Y.; Yahya, N.; Bax, H. J.; Kao, M. W.; Cooper, A. M.; Beavil, A. J.; Sutton, B. J.; Gould, H. J., et al. Mapping of the CD23 Binding Site on Immunoglobulin E (IgE) and Allosteric Control of the IgE-fc EpsilonRI Interaction. *J. Biol. Chem.* **2012**, *287* (37), 31457–31461.

(156) Kang, S.; Kwon, H.; Wen, H.; Song, Y.; Frueh, D.; Ahn, H.-C.; Yoo, S. H.; Wagner, G.; Park, S. Global dynamic conformational changes in the suppressor domain of IP3 receptor by stepwise binding of the two lobes of calmodulin. *FASEB J.* **2011**, *25* (3), 840–850.

(157) Sami, F.; Smet-Nocca, C.; Khan, M.; Landrieu, I.; Lippens, G.; Brautigan, D. L. Molecular basis for an ancient partnership between prolyl isomerase Pin1 and phosphatase inhibitor-2. *Biochemistry* **2011**, *50* (30), 6567–6578.

(158) Berger, I.; Blanco, A. G.; Boelens, R.; Cavarelli, J.; Coll, M.; Folkers, G. E.; Nie, Y.; Pogenberg, V.; Schultz, P.; Wilmanns, M., et al. Structural Insights into Transcription Complexes. *J. Struct. Biol.* **2011**, *175* (2), 135–146.

(159) Gulerez, I. E.; Gehring, K. X-ray Crystallography and NMR as Tools for the Study of Protein Tyrosine Phosphatases. *Methods* **2014**, *65* (2), 175–183.

(160) Choi, J. H.; Laurent, A. H.; Hilser, V. J.; Ostermeier, M. Design of protein switches based on an ensemble model of allostery. *Nat. Commun.* **2015**, *6*, 6968.

(161) Schwieters, C. D.; Suh, J.-Y.; Grishaev, A.; Ghirlando, R.; Takayama, Y.; Clore, G. M. Solution structure of the 128 kDa enzyme I dimer from Escherichia coli and its 146 kDa complex with HPr using residual dipolar couplings and small- and wide-angle X-ray scattering. *J. Am. Chem. Soc.* **2010**, *132* (37), 13026–13045.

(162) Skora, L.; Mestan, J.; Fabbro, D.; Jahnke, W.; Grzesiek, S. NMR reveals the allosteric opening and closing of Abelson tyrosine kinase by ATP-site and myristoyl pocket inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (47), E4437-45.

(163) Salinas, R. K.; Bruschweiler-Li, L.; Johnson, E.; Brüschweiler, R. Ca2+ binding alters the interdomain flexibility between the two cytoplasmic calcium-binding domains in the Na+/Ca2+ exchanger. *J. Biol. Chem.* **2011**, *286* (37), 32123–32131.

(164) Jankowski, W.; Saleh, T.; Pai, M.-T.; Sriram, G.; Birge, R. B.; Kalodimos, C. G. Domain organization differences explain Bcr-Abl's preference for CrkL over CrkII. *Nat. Chem. Biol.* **2012**, *8* (6), 590–596.

(165) Safaee, N.; Kozlov, G.; Noronha, A. M.; Xie, J.; Wilds, C. J.; Gehring, K. Interdomain allostery promotes assembly of the poly(A) mRNA complex with PABP and eIF4G. *Mol. Cell* **2012**, *48* (3), 375–386.

(166) Bista, M.; Freund, S. M.; Fersht, A. R. Domain-Domain Interactions in Full-Length p53 and a Specific DNA Complex Probed by Methyl NMR Spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (39), 15752–15756.

(167) Almeida, R. M.; Geraldes, Carlos F G C; Pauleta, S. R.; Moura, J. J. G. Gd(III) Chelates as NMR Probes of Protein-Protein Interactions. Case Study: Rubredoxin and Cytochrome c3. *Inorg. Chem.* **2011**, *50* (21), 10600–10607.

(168) Tang, C.; Schwieters, C. D.; Clore, G. M. Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. *Nature* **2007**, *449* (7165), 1078–1082.

(169) Clore, G. M.; Iwahara, J. Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-population States of Biological Macromolecules and Their Complexes. *Chem. Rev.* **2009**, *109* (9), 4108–4139.

(170) Otting, G. Protein NMR Using Paramagnetic Ions. *Annu. Rev. Biophys.* **2010**, *39*, 387–405.

(171) Sekhar, A.; Kay, L. E. NMR Paves the Way for Atomic Level Descriptions of Sparsely Populated, Transiently Formed Biomolecular Conformers. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (32), 12867–12874.

(172) Schmitz, C.; Vernon, R.; Otting, G.; Baker, D.; Huber, T. Protein Structure Determination From Pseudocontact Shifts Using ROSETTA. *J. Mol. Biol.* **2012**, *416* (5), 668–677.

(173) Ueda, T.; Takeuchi, K.; Nishida, N.; Stampoulis, P.; Kofuku, Y.; Osawa, M.; Shimada, I. Cross-saturation and Transferred Cross-saturation Experiments. *Q. Rev. Biophys.* **2014**, *47* (2), 143–187.

(174) Shimada, I. NMR Techniques for Identifying the Interface of a Larger Protein–Protein Complex: Cross-Saturation and Transferred Cross-Saturation Experiments. In *Nuclear magnetic resonance of biological*

macromolecules; James, T. L., Ed.; Methods in Enzymology 394; Acad. Press: San Diego [u.a.], (2005); pp 483–506.

(175) Kuehlbrandt, W. Cryo-EM Enters a New Era. *eLife* **2014**, *3*, e01963.

(176) Callaway, E. The revolution will not be crystallized: a new method sweeps through structural biology. *Nature News* **2015**, *525* (7568), 172.

(177) Villarreal, S. A.; Stewart, P. L. CryoEM and image sorting for flexible protein/DNA complexes. *Journal of structural biology* **2014**, *187* (1), 76–83.

(178) Bartesaghi, A.; Merk, A.; Banerjee, S.; Matthies, D.; Wu, X.; Milne, J. L. S.; Subramaniam, S. 2.2 A resolution cryo-EM structure of -galactosidase in complex with a cell-permeant inhibitor. *Science* **2015**, *348* (6239), 1147–1151.

(179) Bai, X.-c.; Yan, C.; Yang, G.; Lu, P.; Ma, D.; Sun, L.; Zhou, R.; Scheres, S. H. W.; Shi, Y. An atomic structure of human γ-secretase. *Nature* **2015**, *525* (7568), 212–217.

(180) Khatter, H.; Myasnikov, A. G.; Natchiar, S. K.; Klaholz, B. P. Structure of the Human 80S Ribosome. *Nature* **2015**, *520* (7549), 640–645.

(181) Liao, M.; Cao, E.; Julius, D.; Cheng, Y. Structure of the TRPV1 Ion Channel Determined by Electron Cryo-Microscopy. *Nature* **2013**, *504* (7478), 107–112.

(182) Amunts, A.; Brown, A.; Bai, X.-c.; Llacer, J. L.; Hussain, T.; Emsley, P.; Long, F.; Murshudov, G.; Scheres, S. H. W.; Ramakrishnan, V. Structure of the Yeast Mitochondrial Large Ribosomal Subunit. *Science* **2014**, *343* (6178), 1485–1489.

(183) Jiang, J.; Pentelute, B. L.; Collier, R. J.; Zhou, Z. H. Atomic structure of anthrax protective antigen pore elucidates toxin translocation. *Nature* **2015**, *521* (7553), 545–549.

(184) Wong, W.; Bai, X.-c.; Brown, A.; Fernandez, I. S.; Hanssen, E.; Condron, M.; Tan, Y. H.; Baum, J.; Scheres, S. H. W. Cryo-EM structure of the Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine. *eLife* **2014**, *3*, e01963.

(185) Kerns, S. J.; Agafonov, R. V.; Cho, Y.-J.; Pontiggia, F.; Otten, R.; Pachov, D. V.; Kutter, S.; Phung, L. A.; Murphy, P. N.; Thai, V., et al. The energy landscape of adenylate kinase during catalysis. *Nat. Struct. Mol. Biol.* **2015**, *22* (2), 124–131.

(186) Danielsson, J.; Awad, W.; Saraboji, K.; Kurnik, M.; Lang, L.; Leinartaite, L.; Marklund, S. L.; Logan, D. T.; Oliveberg, M. Global Structural Motions from the Strain of a Single Hydrogen Bond. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (10), 3829–3834.

(187) Ermakova, E.; Miller, M. C.; Nesmelova, I. V.; López-Merino, L.; Berbé, S.; Cañada, F. J.; Jiménez-Barbero, J.; Sol\'ıs, D.; Gabius, H.-J.; Mayo, K. H. Lactose binding to human galectin-7 (p53-induced gene 1) induces long-range effects through the protein resulting in increased dimer stability and evidence for positive cooperativity. *Glycobiology* **2013**, *23* (5), 508–523.

(188) Anderson, J. S.; Mustafi, S. M.; Hernández, G.; LeMaster, D. M. Statistical Allosteric Coupling to the Active Site Indole Ting Flip Equilibria in the FK506-Binding Domain. *Biophys. Chem.* **2014**, *192*, 41–48.

(189) Xu, N.; Tochio, N.; Wang, J.; Tamari, Y.; Uewaki, J.-I.; Utsunomiya-Tate, N.; Igarashi, K.; Shiraki, T.; Kobayashi, N.; Tate, S.-I. The C113D mutation in human Pin1 causes allosteric structural changes in the phosphate binding pocket of the PPIase domain through the tug of war in the dual-histidine motif. *Biochemistry* **2014**, *53* (34), 5568–5578.

(190) Garriga, D.; Ferrer-Orta, C.; Querol-Audúria. Role of motif B loop in allosteric regulation of RNA-dependent RNA polymerization activity. *J. Mol. Biol.* **2013**, *425* (13), 2279–2287.

(191) McDonald, L. R.; Whitley, M. J.; Boyer, J. A.; Lee, A. L. Colocalization of Fast and Slow Timescale Dynamics in the Allosteric Signaling Protein CheY. *J. Mol. Biol.* **2013**, *425* (13), 2372–2381.

(192) Mueller, L. J.; Dunn, M. F. NMR crystallography of enzyme active sites: probing chemically detailed, three-dimensional structure in tryptophan synthase. *Acc. Chem. Res.* **2013**, *46* (9), 2008–2017.

(193) Velyvis, A.; Yang, Y. R.; Schachman, H. K.; Kay, L. E. A solution NMR study showing that active site ligands and nucleotides directly perturb the allosteric equilibrium in aspartate transcarbamoylase. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (21), 8815–8820.

(194) Bondarenko, V.; Mowrey, D.; Liu, L. T.; Xu, Y.; Tang, P. NMR Resolved Multiple Anesthetic Binding Sites in the TM Domains of the Alpha4 Beta2 nAChR. *Biochim. Biophys. Acta* **2013**, *1828* (2), 398–404.

(195) Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E. A Heteronuclear Correlation Experiment for Simultaneous Determination of 15N Longitudinal Decay and Chemical Exchange Rates of Systems in Slow Equilibrium. *J. Biolmol. NMR* **1994**, *4* (5), 727–734.

(196) Rodriguez, J. C.; Jennings, P. A.; Melacini, G. Using chemical exchange to assign non-covalent protein complexes in slow exchange with the free state: enhanced resolution and efficient signal editing. *J. Biolmol. NMR* **2004**, *30* (2), 155–161.

(197) Sahu, D.; Clore, G. M.; Iwahara, J. TROSY-based z-exchange spectroscopy: application to the determination of the activation energy for intermolecular protein translocation between specific sites on different DNA molecules. *J. Am. Chem. Soc.* **2007**, *129* (43), 13232–13237.

(198) Li, Y.; Palmer, A. G. TROSY-Selected ZZ-Exchange Experiment For Characterizing Slow Chemical Exchange in Large Proteins. *J. Biolmol. NMR* **2009**, *45* (4), 357–360.

(199) John, M.; Schmitz, C.; Park, A. Y.; Dixon, N. E.; Huber, T.; Otting, G. Sequence-specific and stereospecific assignment of methyl groups using paramagnetic lanthanides. *J. Am. Chem. Soc.* **2007**, *129* (44), 13749–13757.

(200) Vallurupalli, P.; Bouvignies, G.; Kay, L. E. Studying "invisible" excited protein states in slow exchange with a major state conformation. *J. Am. Chem. Soc.* **2012**, *134* (19), 8148–8161.

(201) Massi, F.; Johnson, E.; Wang, C.; Rance, M.; Palmer, A. 3. NMR R1 rho rotating-frame relaxation with weak radio frequency fields. *J. Am. Chem. Soc.* **2004**, *126* (7), 2247–2256.

(202) Kay, L. E.; Frydman, L. A special "JMR Perspectives" issue: foresights in biomolecular solution-state NMR spectroscopy - from spin gymnastics to structure and dynamics. *J. Magn. Reson.* **2014**, *241*, 1–2.

(203) Rosenzweig, R.; Kay, L. E. Bringing dynamic molecular machines into focus by methyl-TROSY NMR. *Annu. Rev. Biochem.* **2014**, *83*, 291–315.

(204) Long, D.; Bouvignies, G.; Kay, L. E. Measuring Hydrogen Exchange Rates in Invisible Protein Excited States. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* (24), 8820–8825.

(205) Sanchez-Medina, C.; Sekhar, A.; Vallurupalli, P.; Cerminara, M.; Muñoz, V.; Kay, L. E. Probing the free energy landscape of the fast-folding gpW protein by relaxation dispersion NMR. *J. Am. Chem. Soc.* **2014**, *136* (20), 7444–7451.

(206) Rennella, E.; Huang, R.; Velyvis, A.; Kay, L. E. (13)CHD2-CEST NMR spectroscopy provides an avenue for studies of conformational exchange in high molecular weight proteins. *J. Biolmol. NMR* **2015**, *63* (2), 187–199.

(207) Horváth, G.; Király, P.; Tárkányi, G.; Toke, O. Internal motions and exchange processes in human ileal bile acid binding protein as studied by backbone (15)N nuclear magnetic resonance spectroscopy. *Biochemistry* **2012**, *51* (9), 1848–1861.

(208) Farber, P. J.; Mittermaier, A. Concerted Dynamics Link Allosteric Sites in the PBX Homeodomain. *J. Mol. Biol.* **2011**, *405* (3), 819–830.

(209) Tzeng, S.-R.; Kalodimos, C. G. Allosteric Inhibition Through Suppression of Transient Conformational States. *Nat. Chem. Biol.* **2013**, *9* (7), 462–465.

(210) Coudevylle, N.; Geist, L.; Hoetzinger, M.; Tollinger, M.; Konrat, R. Siderocalin Q83 Exhibits Differential Slow Dynamics upon Ligand Binding. *J. Biolmol. NMR* **2011**, *51* (1-2), 83–88.

(211) Selvaratnam, R.; Akimoto, M.; VanSchouwen, B.; Melacini, G. cAMP-Dependent Allostery and Dynamics in Epac: An NMR View. *Biochem. Soc. Trans.* **2012**, *40* (1), 219–223.

(212) Lorieau, J. L.; Louis, J. M.; Schwieters, C. D.; Bax, A. pH-Triggered, Activated-State Conformations of the Influenza Hemagglutinin Fusion Peptide Revealed by NMR. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (49), 19994–19999.

(213) Yadav, P. K.; Xie, P.; Banerjee, R. Allosteric communication between the pyridoxal 5'-phosphate (PLP) and heme sites in the H2S generator human cystathionine beta-synthase. *J. Biol. Chem.* **2012**, *287* (45), 37611–37620.

(214) Lofgren, M.; Banerjee, R. Loss of Allostery and Coenzyme B12 Delivery by a Pathogenic Mutation in Adenosyltransferase. *Biochemistry* **2011**, *50* (25), 5790–5798.

(215) Doucet, N.; Khirich, G.; Kovrigin, E. L.; Loria, J. P. Alteration of Hydrogen Bonding in the Vicinity of Histidine 48 Disrupts Millisecond Motions in RNase A. *Biochemistry* **2011**, *50* (10), 1723–1730.

(216) Tam, M. F.; Rice, N. W.; Maillett, D. H.; Simplaceanu, V.; Ho, N. T.; Tam, T. C. S.; Shen, T.-J.; Ho, C. Autoxidation and oxygen binding properties of recombinant hemoglobins with substitutions at the \alphaVal-62 or \betaVal-67 position of the distal heme pocket. *J. Biol. Chem.* **2013**, *288* (35), 25512–25521.

(217) Krimm, I.; Lancelin, J.-M.; Praly, J.-P. Binding evaluation of fragment-based scaffolds for probing allosteric enzymes. *J. Med. Chem.* **2012**, *55* (3), 1287–1295.

(218) Thomas-Chollier, M.; Watson, L. C.; Cooper, S. B.; Pufall, M. A.; Liu, J. S.; Borzym, K.; Vingron, M.; Yamamoto, K. R.; Meijsing, S. H. A naturally occurring insertion of a single amino acid rewires transcriptional regulation by glucocorticoid receptor isoforms. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (44), 17826–17831.

(219) Nelson, M. L.; Kang, H.-S.; Lee, G. M.; Blaszczak, A. G.; Lau, D. K. W.; McIntosh, L. P.; Graves, B. J. Ras signaling requires dynamic properties of Ets1 for phosphorylation-enhanced binding to coactivator CBP. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107* (22), 10026–10031.

(220) Dementiev, A.; Swanson, R.; Roth, R.; Isetti, G.; Izaguirre, G.; Olson, S. T.; Gettins, P. G. W. The Allosteric Mechanism of Activation of Antithrombin as an Inhibitor of Factor IXa and Factor Xa: Heparin-independent Full Activation Through Mutations Adjacent to Helix D. *J. Biol. Chem.* **2013**, *288* (47), 33611–33619.

(221) Buchli, B.; Waldauer, S. A.; Walser, R.; Donten, M. L.; Pfister, R.; Blöchliger, N.; Steiner, S.; Caflisch, A.; Zerbe, O.; Hamm, P. Kinetic Response of a Photoperturbed Allosteric Protein. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (29), 11725–11730.

(222) Tinberg, C. E.; Khare, S. D.; Dou, J.; Doyle, L.; Nelson, J. W.; Schena, A.; Jankowski, W.; Kalodimos, C. G.; Johnsson, K.; Stoddard, B. L., et al. Computational Design of Ligand-binding Proteins with High Affinity and Selectivity. *Nature* **2013**, *501* (7466), 212–216.

(223) Latham, M. P.; Sekhar, A.; Kay, L. E. Understanding the mechanism of proteasome 20S core particle gating. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* (15), 5532–5537.

(224) Shi, L.; Kay, L. E. Tracing an allosteric pathway regulating the activity of the HslV protease. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* (6), 2140–2145.

(225) Wylie, B. J.; Bhate, M. P.; McDermott, A. E. Transmembrane allosteric coupling of the gates in a potassium channel. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* (1), 185–190.

(226) Boulton, S.; Akimoto, M.; Selvaratnam, R.; Bashiri, A.; Melacini, G. A Tool Set to Map Allosteric Networks through the NMR Chemical Shift Covariance Analysis. *Sci. Rep.* **2014**, *4*, 7306.

(227) Axe, J. M.; Boehr, D. D. Long-Range Interactions in the Alpha Subunit of Tryptophan Synthase Help to Coordinate Ligand Binding, Catalysis, and Substrate Channeling. *J. Mol. Biol.* **2013**, *425* (9), 1527–1545.

(228) Dawson, J. E.; Farber, P. J.; Forman-Kay, J. D. Allosteric Coupling Between the Intracellular Coupling Helix 4 and Regulatory Sites of the First Nucleotide-binding Domain of CFTR. *PloS one* **2013**, *8* (9), e74347.

(229) Stollar, E. J.; Lin, H.; Davidson, A. R.; Forman-Kay, J. D. Differential dynamic engagement within 24 SH3 domain: peptide complexes revealed by co-linear chemical shift perturbation analysis. *PloS one* **2012**, *7* (12), e51282.

(230) Axe, J. M.; O'Rourke, K. F.; Kerstetter, N. E.; Yezdimer, E. M.; Chan, Y. M.; Chasin, A.; Boehr, D. D. Severing of a Hydrogen Bond Disrupts Amino Acid Networks in the Catalytically Active State of the Alpha Subunit of Tryptophan Synthase. *Protein Sci.* **2015**, *24* (4), 484–494.

(231) Kim, J.; Masterson, L. R.; Cembran, A.; Verardi, R.; Shi, L.; Gao, J.; Taylor, S. S.; Veglia, G. Dysfunctional conformational dynamics of protein kinase A induced by a lethal mutant of phospholamban hinder phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112* (12), 3716–3721.

(232) Cembran, A.; Kim, J.; Gao, J.; Veglia, G. NMR Mapping of Protein Conformational Landscapes Using Coordinated Behavior of Chemical Shifts upon Ligand Binding. *Phys. Chem. Chem. Phys.* **2014**, *16* (14), 6508.

(233) Anthis, N. J.; Clore, G. M. Visualizing Transient Dark States by NMR Spectroscopy. *Q. Rev. Biophys.* **2015**, *48* (1), 35–116.

(234) Clore, G. M. Exploring Sparsely Populated States of Macromolecules by Diamagnetic and Paramagnetic NMR Relaxation. *Protein Sci.* **2011**, *20* (2), 229–246.

(235) Shimada, I.; Ueda, T.; Matsumoto, M.; Sakakura, M.; Osawa, M.; Takeuchi, K.; Nishida, N.; Takahashi, H. Cross-saturation and transferred cross-saturation experiments. *Prog. Nucl. Magn. Reson. Spectrosc.* **2009**, *54* (2), 123–140.

(236) Serrano, L.; Horovitz, A.; Avron, B.; Bycroft, M.; Fersht, A. R. Estimating the contribution of engineered surface electrostatic interactions to protein stability by using double-mutant cycles. *Biochemistry* **1990**, *29* (40), 9343–9352.

(237) Maksay, G. Allostery in Pharmacology: Thermodynamics, Evolution and Design. *Prog. Biophys. Mol. Biol.* **2011**, *106* (3), 463–473.

(238) van den Bedem, H.; Bhabha, G.; Yang, K.; Wright, P. E.; Fraser, J. S. Automated identification of functional dynamic contact networks from X-ray crystallography. *Nat. Methods* **2013**, *10* (9), 896–902.

(239) Fraser, J. S.; van den Bedem, H.; Samelson, A. J.; Lang, P. T.; Holton, J. M.; Echols, N.; Alber, T. Accessing Protein Conformational Ensembles Using Room-Temperature X-ray Crystallography. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (39), 16247–16252.

(240) Fenwick, R. B.; van den Bedem, H.; Fraser, J. S.; Wright, P. E. Integrated Description of Protein Dynamics From Room-Temperature X-ray Crystallography and NMR. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* (4), E445-54.

(241) Majumder, S.; DeMott, C. M.; Burz, D. S.; Shekhtman, A. Using Singular Value Decomposition to Characterize Protein-Protein Interactions by In-Cell NMR Spectroscopy. *ChemBioChem* **2014**, *15* (7), 929–933.

(242) Berjanskii, M. V.; Wishart, D. S. A Simple Method to Predict Protein Flexibility Using Secondary Chemical Shifts. *J. Am. Chem. Soc.* **2005**, *127* (43), 14970–14971.

(243) Kontaxis, G.; Delaglio, F.; Bax, A. Molecular fragment replacement approach to protein structure determination by chemical shift and dipolar homology database mining. *Methods Enzymol.* **2005**, *394*, 42–78.

(244) Kohlhoff, K. J.; Robustelli, P.; Cavalli, A.; Salvatella, X.; Vendruscolo, M. Fast and Accurate Predictions of Protein NMR Chemical Shifts from Interatomic Distances. *J. Am. Chem. Soc.* **2009**, *131* (39), 13894–13895.

(245) Shen, Y.; Lange, O.; Delaglio, F.; Rossi, P.; Aramini, J. M.; Liu, G.; Eletsky, A.; Wu, Y.; Singarapu, K. K.; Lemak, A., et al. Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (12), 4685–4690.

(246) Wishart, D. S.; Arndt, D.; Berjanskii, M.; Tang, P.; Zhou, J.; Lin, G. CS23D: a web server for rapid protein structure generation using NMR chemical shifts and sequence data. *Nucleic Acids Res.* **2008**, *36* (Web Server issue), W496-502.

(247) Brüschweiler, S.; Schanda, P.; Kloiber, K.; Brutscher, B.; Kontaxis, G.; Konrat, R.; Tollinger, M. Direct Observation of the Dynamic Process Underlying Allosteric Signal Transmission. *J. Am. Chem. Soc.* **2009**, *131* (8), 3063–3068.

(248) Das, R.; Mazhab-Jafari, M. T.; Chowdhury, S.; SilDas, S.; Selvaratnam, R.; Melacini, G. Entropy-driven cAMP-Dependent Allosteric Control of Inhibitory Interactions in Exchange Proteins Directly Activated By cAMP. *J. Biol. Chem.* **2008**, *283* (28), 19691–19703.

(249) Peng, J. W. Investigating Dynamic Interdomain Allostery in Pin1. *Biophys. Rev.* **2015**, *7* (2), 239–249.

(250) Venditti, V.; Tugarinov, V.; Schwieters, C. D.; Grishaev, A.; Clore, G. M. Large interdomain rearrangement triggered by suppression of micro- to millisecond dynamics in bacterial Enzyme I. *Nat. Commun.* **2015**, *6*, 5960.

(251) Wang, C.; Rance, M.; Palmer, A. G. 3. Mapping chemical exchange in proteins with MW 50 kD. *J. Am. Chem. Soc.* **2003**, *125* (30), 8968–8969.

(252) Tokunaga, Y.; Takeuchi, K.; Takahashi, H.; Shimada, I. Allosteric Enhancement of MAP Kinase p38 Alpha's Activity and Substrate Selectivity by Docking Interactions. *Nat. Struct. Mol. Biol.* **2014**, *21* (8), 704–711.

(253) Greenwood, A. I.; Kwon, J.; Nicholson, L. K. Isomerase-catalyzed binding of interleukin-1 receptor-associated kinase 1 to the EVH1 domain of vasodilator-stimulated phosphoprotein. *Biochemistry* **2014**, *53* (22), 3593–3607.

(254) Wilson, K. A.; Bouchard, J. J.; Peng, J. W. Interdomain interactions support interdomain communication in human Pin1. *Biochemistry* **2013**, *52* (40), 6968–6981.

(255) De, S.; Greenwood, A. I.; Rogals, M. J.; Kovrigin, E. L.; Lu, K. P.; Nicholson, L. K. 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* **2012**, *51* (43), 8583–8596.

(256) Sarkar, P.; Saleh, T.; Tzeng, S.-R.; Birge, R. B.; Kalodimos, C. G. Structural basis for regulation of the Crk signaling protein by a proline switch. *Nat. Chem. Biol.* **2011**, *7* (1), 51–57.

(257) Akimoto, M.; McNicholl, E. T.; Ramkissoon, A.; Moleschi, K.; Taylor, S. S.; Melacini, G. Mapping the Free Energy Landscape of PKA Inhibition and Activation: A Double-Conformational Selection Model for the Tandem cAMP-Binding Domains of PKA RIalpha. *PLoS Biol.* **2015**, *13* (11), e1002305.

(258) Villali, J.; Pontiggia, F.; Clarkson, M. W.; Hagan, M. F.; Kern, D. Evidence against the "Y-T coupling" mechanism of activation in the response regulator NtrC. *J. Mol. Biol.* **2014**, *426* (7), 1554–1567.

(259) Fogolari, F.; Corazza, A.; Fortuna, S.; Soler, M. A.; VanSchouwen, B.; Brancolini, G.; Corni, S.; Melacini, G.; Esposito, G. Distance-Based Configurational Entropy of Proteins from Molecular Dynamics Simulations. *PloS one* **2015**, *10* (7), e0132356.

(260) Cooper, A.; Dryden, D. T. Allostery Without Conformational Change. A Plausible Model. *Eur. Biophys. J.* **1984**, *11* (2), 103–109.

(261) Tzeng, S.-R.; Kalodimos, C. G. Dynamic Activation of an Allosteric Regulatory Protein. *Nature* **2009**, *462* (7271), 368–372.

(262) Marlow, M. S.; Dogan, J.; Frederick, K. K.; Valentine, K. G.; Wand, A. J. The role of conformational entropy in molecular recognition by calmodulin. *Nat. Chem. Biol.* **2010**, *6* (5), 352–358.

(263) Yang, D.; Kay, L. E. Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: application to protein folding. *J. Mol. Biol.* **1996**, *263* (2), 369–382.

(264) Kasinath, V.; Sharp, K. A.; Wand, A. J. Microscopic insights into the NMR relaxation-based protein conformational entropy meter. *J. Am. Chem. Soc.* **2013**, *135* (40), 15092–15100.

(265) Sharp, K. A.; O'Brien, E.; Kasinath, V.; Wand, A. J. On the relationship between NMR-derived amide order parameters and protein backbone entropy changes. *Proteins* **2015**, *83* (5), 922–930.

(266) Zerbetto, M.; Anderson, R.; Bouguet-Bonnet, S.; Rech, M.; Zhang, L.; Meirovitch, E.; Polimeno, A.; Buck, M. Analysis of 15N-1H NMR Relaxation in Proteins by a Combined Experimental and Molecular Dynamics Simulation Approach: Picosecond-nanosecond Dynamics of the Rho GTPase Binding Domain of Plexin-b1 in the Dimeric State Indicates Allosteric Pathways. *J. Phys. Chem* **2013**, *117* (1), 174–184.

(267) Buchenberg, S.; Knecht, V.; Walser, R.; Hamm, P.; Stock, G. Long-range Conformational Transition of a Photoswitchable Allosteric Protein: Molecular Dynamics Simulation Study. *J. Phys. Chem* **2014**, *118* (47), 13468–13476.

(268) Palazzesi, F.; Barducci, A.; Tollinger, M.; Parrinello, M. The allosteric communication pathways in KIX domain of CBP. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (35), 14237–14242.

(269) Pontiggia, F.; Pachov, D. V.; Clarkson, M. W.; Villali, J.; Hagan, M. F.; Pande, V. S.; Kern, D. Free Energy Landscape of Activation in a Signalling Protein at Atomic Resolution. *Nat. Commun.* **2015**, *6*, 7284.

(270) Coyne, H. J.; De, S.; Okon, M.; Green, S. M.; Bhachech, N.; Graves, B. J.; McIntosh, L. P. Autoinhibition of ETV6 (TEL) DNA Binding: Appended Helices Sterically Block the ETS Domain. *J. Mol. Biol.* **2012**, *421* (1), 67–84.

(271) Das, R.; Abu-Abed, M.; Melacini, G. Mapping Allostery Through Equilibrium Perturbation NMR Spectroscopy. *J. Am. Chem. Soc.* **2006**, *128* (26), 8406–8407.

(272) Gangadhara, B. N.; Laine, J. M.; Kathuria, S. V.; Massi, F.; Matthews, C. R. Clusters of Branched Aliphatic Side Chains Serve as Cores of Stability in the Native State of the HisF TIM Barrel Protein. *J. Mol. Biol.* **2013**, *425* (6), 1065–1081.

(273) Xu, J.; Lee, Y.; Beamer, L. J.; van Doren, S. R. Phosphorylation in the catalytic cleft stabilizes and attracts domains of a phosphohexomutase. *Biophys. J.* **2015**, *108* (2), 325–337.

(274) Boyer, J. A.; Clay, C. J.; Luce, K. S.; Edgell, M. H.; Lee, A. L. Detection of Native-state Nonadditivity in Double Mutant Cycles via Hydrogen Exchange. *J. Am. Chem. Soc.* **2010**, *132* (23), 8010–8019.

(275) Li, G. C.; Srivastava, A. K.; Kim, J.; Taylor, S. S.; Veglia, G. Mapping the Hydrogen Bond Networks in the Catalytic Subunit of Protein Kinase A Using H/D Fractionation Factors. *Biochemistry* **2015**, *54* (26), 4042–4049.

(276) Chen, L.; Balabanidou, V.; Remeta, D. P.; Minetti, Conceicao A S A; Portaliou, A. G.; Economou, A.; Kalodimos, C. G. Structural Instability Tuning as a Regulatory Mechanism in Protein-protein Interactions. *Mol. Cell* **2011**, *44* (5), 734–744.

(277) Das, R.; Chowdhury, S.; Mazhab-Jafari, M. T.; SilDas, S.; Selvaratnam, R.; Melacini, G. Dynamically Driven Ligand Selectivity in Cyclic Nucleotide Binding Domains. *J. Biol. Chem.* **2009**, *284* (35), 23682–23696.

(278) Atkinson, A. P.; Baguet, E.; Galland, N.; Le Questel, J.-Y.; Planchat, A.; Graton, J. Structural Features and Hydrogen-Bond Properties of Galanthamine and Codeine: An Experimental and Theoretical Study. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2011**, *17* (41), 11637–11649.

(279) Habazettl, J.; Allan, M. G.; Jenal, U.; Grzesiek, S. Solution Structure of the PilZ Domain Protein PA4608 Complex with Cyclic di-GMP Identifies Charge Clustering as Molecular Readout. *J. Biol. Chem.* **2011**, *286* (16), 14304–14314.

(280) Sindhuwinata, N.; Grimm, L. L.; Weißbach, S.; Zinn, S.; Munoz, E.; Palcic, M. M.; Peters, T. Thermodynamic signature of substrates and substrate analogs binding to human blood group B galactosyltransferase from isothermal titration calorimetry experiments. *Biopolymers* **2013**, *99* (10), 784–795.

(281) Vrontaki, E.; Leonis, G.; Papadopoulos, M. G.; Simcic, M.; Golic Grdadolnik, S.; Afantitis, A.; Melagraki, G.; Hadjikakou, S. K.; Mavromoustakos, T. Comparative binding effects of aspirin and anti-inflammatory Cu complex in the active site of LOX-1. *J. Chem. Inf. Model.* **2012**, *52* (12), 3293–3301.

(282) Canales, A.; Nieto, L.; Rodránchez-Murcia, P. A.; Coderch, C.; Cortés-Cabrera, A.; Paterson, I.; Carlomagno, T.; Gago, F.; Andreu, J. M.; Altmann, K.-H., et al. Molecular Recognition of Epothilones by Microtubules and Tubulin Dimers Revealed by Biochemical and NMR Approaches. *ACS Chem. Biol.* **2014**, *9* (4), 1033–1043.

(283) Assrir, N.; Richez, C.; Durand, P.; Guittet, E.; Badet, B.; Lescop, E.; Badet-Denisot, M.-A. Mapping the UDP-N-Acetylglucosamine Regulatory Site of Human Glucosamine-6P Synthase by Saturation-Transfer Difference NMR and Site-Directed Mutagenesis. *Biochimie* **2014**, *97*, 39–48.

(284) Zhang, W.; Li, R.; Shin, R.; Wang, Y.; Padmalayam, I.; Zhai, L.; Krishna, N. R. Identification of the Binding Site of an Allosteric Ligand Using STD-NMR, Docking, and CORCEMA-ST Calculations. *ChemMedChem* **2013**, *8* (10), 1629–1633.

(285) Wei, Y.; Stec, B.; Redfield, A. G.; Weerapana, E.; Roberts, M. F. Phospholipid-binding sites of phosphatase and tensin homolog (PTEN): exploring the mechanism of phosphatidylinositol 4,5-bisphosphate activation. *J. Biol. Chem.* **2015**, *290* (3), 1592–1606.

(286) Morizono, D. 19F NMR study on the complex of fluorinated vitamin D derivatives with vitamin D receptor: elucidation of the conformation of vitamin D ligands accommodated in the receptor. *J. Med. Dent. Sci.* **2011**, *58* (4), 103–112.

(287) Niks, D.; Hilario, E.; Dierkers, A.; Ngo, H.; Borchardt, D.; Neubauer, T. J.; Fan, L.; Mueller, L. J.; Dunn, M. F. Allostery and substrate channeling in the tryptophan synthase bienzyme complex: evidence for two subunit conformations and four quaternary states. *Biochemistry* **2013**, *52* (37), 6396–6411.

(288) Nussinov, R.; Tsai, C. J. Unraveling Structural Mechanisms of Allosteric Drug Action. *Trends. Pharmacol. Sci.* **2014**, *35* (5), 256–264.

(289) Nussinov, R.; Tsai, C. J. 'Latent Drivers' Expand the Cancer Mutational Landscape. *Curr. Opin. Struct. Biol.* **2015**, *32*, 25–32.

(290) Daniels, K. G.; Suo, Y.; Oas, T. G. Conformational Kinetics Reveals Affinities of Protein Conformational States. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112* (30), 9352–9357.

(291) Khirich, G.; Loria, J. P. Complexity of protein energy landscapes studied by solution NMR relaxation dispersion experiments. *J. Phys. Chem* **2015**, *119* (9), 3743–3754.

(292) Tsai, C.-J.; Del Sol, A.; Nussinov, R. Protein Allostery, Signal Transmission and Dynamics: a Classification Scheme of Allosteric Mechanisms. *Mol. Biosyst.* **2009**, *5* (3), 207–216.

(293) Dawson, J. E.; Seckute, J.; De, S.; Schueler, S. A.; Oswald, A. B.; Nicholson, L. K. Elucidation of a pH-folding Switch in the Pseudomonas syringae Effector Protein AvrPto. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106* (21), 8543–8548.

(294) Fitch, C. A.; Platzer, G.; Okon, M.; Garcia-Moreno, B. E.; McIntosh, L. P. Arginine: Its pKa Value Revisited. *Protein Sci.* **2015**, *24* (5), 752–761.

(295) Ludwiczek, M. L.; D'Angelo, I.; Yalloway, G. N.; Brockerman, J. A.; Okon, M.; Nielsen, J. E.; Strynadka, N. C. J.; Withers, S. G.; McIntosh, L. P. Strategies for Modulating the pH-Dependent Activity of a Family 11 Glycoside Hydrolase. *Biochemistry* **2013**, *52* (18), 3138–3156.

(296) McIntosh, L. P.; Naito, D.; Baturin, S. J.; Okon, M.; Joshi, M. D.; Nielsen, J. E. Dissecting Electrostatic Interactions in Bacillus circulans Xylanase Through NMR-Monitored pH Titrations. *J. Biolmol. NMR* **2011**, *51* (1-2), 5–19.

(297) Manglik, A.; Kim, T. H.; Masureel, M.; Altenbach, C.; Yang, Z.; Hilger, D.; Lerch, M. T.; Kobilka, T. S.; Thian, F. S.; Hubbell, W. L., et al. Structural Insights into the Dynamic Process of \beta2-Adrenergic Receptor Signaling. *Cell* **2015**, *161* (5), 1101–1111.

(298) Kanaori, K.; Tajiri, Y.; Tsuneshige, A.; Ishigami, I.; Ogura, T.; Tajima, K.; Neya, S.; Yonetani, T. T-quaternary structure of oxy human adult hemoglobin in the presence of two allosteric effectors, L35 and IHP. *Biochim. Biophys. Acta* **2011**, *1807* (10), 1253–1261.

(299) Nagatomo, S.; Hamada, H.; Yoshikawa, H. Elongation of the Fe-His bond in the \alpha subunit induced by binding of the allosteric effector bezafibrate to hemoglobins. *J. Phys. Chem* **2011**, *115* (44), 12971–12977.

(300) Yuan, Y.; Shen, T.-J.; Gupta, P.; Ho, N. T.; Simplaceanu, V.; Tam, T. C. S.; Hofreiter, M.; Cooper, A.; Campbell, K. L.; Ho, C. A biochemical-biophysical study of hemoglobins from woolly mammoth, Asian elephant, and humans. *Biochemistry* **2011**, *50* (34), 7350–7360.

(301) Freiburger, L.; Miletti, T.; Zhu, S.; Baettig, O.; Berghuis, A.; Auclair, K.; Mittermaier, A. Substrate-dependent switching of the allosteric binding mechanism of a dimeric enzyme. *Nat. Chem. Biol.* **2014**, *10* (11), 937–942.

(302) Freiburger, L. A.; Baettig, O. M.; Sprules, T.; Berghuis, A. M.; Auclair, K.; Mittermaier, A. K. Competing Allosteric Mechanisms Modulate Substrate Binding in a Dimeric Enzyme. *Nat. Struct. Mol. Biol.* **2011**, *18* (3), 288–294.

(303) Ziarek, J. J.; Getschman, A. E.; Butler, S. J.; Taleski, D.; Stephens, B.; Kufareva, I.; Handel, T. M.; Payne, R. J.; Volkman, B. F. Sulfopeptide Probes of the CXCR4/CXCL12 Interface Reveal Oligomer-specific Contacts and Chemokine Allostery. *ACS Chem. Biol.* **2013**, *8* (9), 1955–1963.

(304) McDonald, L. R.; Boyer, J. A.; Lee, A. L. Segmental Motions, Not a Two-State Concerted Switch, Underlie Allostery in CheY. *Structure* **2012**, *20* (8), 1363–1373.

(305) Kleckner, I. R.; Gollnick, P.; Foster, M. P. Mechanisms of Allosteric Gene Regulation by NMR Quantification of Microsecond-Millisecond Protein Dynamics. *J. Mol. Biol.* **2012**, *415* (2), 372–381.

(306) Millet, O.; Loria, J. P.; Kroenke, C. D.; Pons, M.; Palmer, A. G. The Static Magnetic Field Dependence of Chemical Exchange Linebroadening Defines the NMR Chemical Shift Time Scale. *J. Am. Chem. Soc.* **2000**, *122* (12), 2867–2877.

(307) Trzaskowski, B.; Latek, D.; Yuan, S.; Ghoshdastider, U.; Debinski, A.; Filipek, S. Action of Molecular Switches in GPCRs - Theoretical and Experimental Studies. *Curr. Med. Chem.* **2012**, *19* (8), 1090–1109.

(308) Reining, A.; Nozinovic, S.; Schlepckow, K.; Buhr, F.; Fürtig, B.; Schwalbe, H. Three-state Mechanism Couples Ligand and Temperature Sensing in Riboswitches. *Nature* **2013**, *499* (7458), 355–359.

(309) Korzhnev, D. M.; Bezsonova, I.; Lee, S.; Chalikian, T. V.; Kay, L. E. Alternate binding modes for a ubiquitin-SH3 domain interaction studied by NMR spectroscopy. *J. Mol. Biol.* **2009**, *386* (2), 391–405.

(310) Li, P.; Martins, I. R. S.; Rosen, M. K. The Feasibility of Parameterizing Four-State Equilibria Using Relaxation Dispersion Measurements. *J. Biolmol. NMR* **2011**, *51* (1-2), 57–70.

(311) Farber, P. J.; Slager, J.; Mittermaier, A. K. Local Folding and Misfolding in the PBX Homeodomain from a Three-state Analysis of CPMG Relaxation Dispersion NMR Data. *J. Phys. Chem* **2012**, *116* (34), 10317–10329.

(312) Esteban-Martén, J.; Cossins, B.; Bertoncini, C. W.; Guallar, V.; Wolf-Watz, M.; Salvatella, X. Correlated Inter-domain Motions in Adenylate Kinase. *PLoS Comput. Biol.* **2014**, *10* (7), e1003721.

(313) VanSchouwen, B.; Selvaratnam, R.; Giri, R.; Lorenz, R.; Herberg, F. W.; Kim, C.; Melacini, G. Mechanism of cAMP Partial Agonism in Protein Kinase G (PKG). *J. Biol. Chem.* **2015**, *290* (48), 28631–28641.

(314) Fenwick, R. B.; Esteban-MartWalter, K. F. A.; Milovanovic, D.; Becker, S.; Lakomek, N. A.; Griesinger, C.; Salvatella, X. Weak Long-Range Correlated Motions in a Surface Patch of Ubiquitin Involved in Molecular Recognition. *J. Am. Chem. Soc.* **2011**, *133* (27), 10336–10339.

(315) Kukic, P.; Alvin Leung, H. T.; Bemporad, F.; Aprile, F. A.; Kumita, J. R.; Simone, A. d.; Camilloni, C.; Vendruscolo, M. Structure and dynamics of the integrin LFA-1 I-domain in the inactive state underlie its inside-out/outside-in signaling and allosteric mechanisms. *Structure* **2015**, *23* (4), 745–753.

(316) Leung, H. T. A.; Kukic, P.; Camilloni, C.; Bemporad, F.; Simone, A. d.; Aprile, F. A.; Kumita, J. R.; Vendruscolo, M. NMR Characterization of the Conformational Fluctuations of the Human Lymphocyte Function-Associated Antigen-1 I-Domain. *Protein Sci.* **2014**, *23* (11), 1596–1606.

(317) Kragelj, J.; Ozenne, V.; Blackledge, M.; Jensen, M. R. Conformational propensities of intrinsically disordered proteins from NMR chemical shifts. *ChemPhysChem* **2013**, *14* (13), 3034–3045.

(318) Bouvignies, G.; Blackledge, M. Structural Biology Turned on Its Head. *ChemBioChem* **2015**, *16* (7), 1033–1034.

(319) Felli, I. C.; Pierattelli, R. *Intrinsically Disordered Proteins Studied by NMR Spectroscopy*; Advances in experimental medicine and biology volume 870; Springer: Cham, 2015.

(320) Jensen, M. R.; Zweckstetter, M.; Huang, J.-r.; Blackledge, M. Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy. *Chemical reviews* **2014**, *114* (13), 6632–6660.

(321) Lu, S.; Banerjee, A.; Jang, H.; Zhang, J.; Gaponenko, V.; Nussinov, R. GTP Binding and Oncogenic Mutations May Attenuate Hypervariable Region (HVR)-Catalytic Domain Interactions in Small GTPase K-Ras4B, Exposing the Effector Binding Site. *J. Biol. Chem.* **2015**, *290* (48), 28887–28900.

(322) Chang, F.-M. J.; Coyne, H. J.; Cubillas, C.; Vinuesa, P.; Fang, X.; Ma, Z.; Ma, D.; Helmann, J. D.; Garc\'ıa-de los Santos, A.; Wang, Y.-X., et al. Cu(I)-mediated Allosteric Switching in a Copper-sensing Operon Repressor(CsoR). *J. Biol. Chem.* **2014**, *289* (27), 19204–19217.

(323) Wirth, A. J.; Gruebele, M. Quinary protein structure and the consequences of crowding in living cells: leaving the test-tube behind. *Bioessays* **2013**, *35* (11), 984–993.

(324) Smith, M. J.; Ikura, M. Integrated RAS signaling defined by parallel NMR detection of effectors and regulators. *Nat. Chem. Biol.* **2014**, *10* (3), 223–230.

(325) Gustavsson, M.; Traaseth, N. J.; Veglia, G. Probing Ground and Excited States of Phospholamban in Model and Native Lipid Membranes by Magic Angle Spinning NMR Spectroscopy. *Biochim. Biophys. Acta, Biomembr.* **2012**, *1818* (2), 146–153.

(326) Xu, X.; Marni, F.; Wu, S.; Su, Z.; Musayev, F.; Shrestha, S.; Xie, C.; Gao, W.; Liu, Q.; Zhou, L. Local and global interpretations of a disease-causing mutation near the ligand entry path in hyperpolarization-activated cAMP-gated channel. *Structure* **2012**, *20* (12), 2116–2123.

(327) Cardoso, R.; Love, R.; Nilsson, C. L.; Bergqvist, S.; Nowlin, D.; Yan, J.; Liu, K. K.-C.; Zhu, J.; Chen, P.; Deng, Y.-L., et al. Identification of Cys255 in HIFf-1 Alpha as a Novel Site for Development of Covalent Inhibitors of HIF-1 Alpha ARNT PasB Domain Protein-protein Interaction. *Protein Sci.* **2012**, *21* (12), 1885–1896.

(328) Cheney, D. L.; Bozarth, J. M.; Metzler, W. J.; Morin, P. E.; Mueller, L.; Newitt, J. A.; Nirschl, A. H.; Rendina, A. R.; Tamura, J. K.; Wei, A., et al. Discovery of Novel P1 Groups for Coagulation Factor VIIa Inhibition Using Fragment-based Screening. *J. Med. Chem.* **2015**, *58* (6), 2799–2808.

(329) Han, S.; Czerwinski, R. M.; Caspers, N. L.; Limburg, D. C.; Ding, W.; Wang, H.; Ohren, J. F.; Rajamohan, F.; McLellan, T. J.; Unwalla, R., et al. Selectively targeting an inactive conformation of interleukin-2-inducible T-cell kinase by allosteric inhibitors. *Biochem. J.* **2014**, *460* (2), 211–222.

(330) Zhuo, X.; Huang, X. S.; Degnan, A. P.; Snyder, L. B.; Yang, F.; Huang, H.; Shu, Y.-Z.; Johnson, B. M. Identification of Glutathione Conjugates of Acetylene-containing Positive Allosteric Modulators of Metabotropic Glutamate Receptor Subtype 5. *Drug Metab. Dispos.* **2015**, *43* (4), 578–589.

(331) Gregory, K. J.; Sexton, P. M.; Christopoulos, A. Overview of Receptor Allosterism. *Curr. Protoc. Pharmacol.* **2010**, *Chapter 1*, Unit 1.21.

(332) Harner, M. J.; Frank, A. O.; Fesik, S. W. Fragment-based drug discovery using NMR spectroscopy. *J. Biolmol. NMR* **2013**, *56* (2), 65–75.

(333) Dalvit, C.; Caronni, D.; Mongelli, N.; Veronesi, M.; Vulpetti, A. NMR-based Quality Control Approach for the Identification of False Positives and False Negatives in High Throughput Screening. *Curr. Drug Discov. Technol.* **2006**, *3* (2), 115–124.

(334) Pellecchia, M.; Bertini, I.; Cowburn, D.; Dalvit, C.; Giralt, E.; Jahnke, W.; James, T. L.; Homans, S. W.; Kessler, H.; Luchinat, C., et al. Perspectives on NMR in drug discovery: a technique comes of age. *Nat. Rev. Drug Discovery* **2008**, *7* (9), 738–745.

(335) Laurence, J. S.; Latham, C. F.; Tinetti, R. N.; Johnson, A.; Tyssen, D.; Huber, K. D.; Sluis-Cremer, N.; Simpson, J. S.; Headey, S. J.; Chalmers, D. K., et al. Identification of mechanistically distinct inhibitors of HIV-1 reverse transcriptase through fragment screening. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112* (22), 6979–6984.

(336) Dalvit, C. NMR Methods in Fragment Screening: Theory and a Comparison with Other Biophysical Techniques. *Drug Discov. Today* **2009**, *14* (21-22), 1051–1057.

(337) Cala, O.; Guillière, F.; Krimm, I. NMR-based Analysis of Protein-ligand Interactions. *Anal. Bioanal. Chem.* **2014**, *406* (4), 943–956.

(338) Carvalho, LuSeixas, Raquel S. G. R.; Silva, A. M. S.; Nave, M.; Martins, A. C.; Erhardt, S.; Fernandes, E.; Cabrita, E. J.; Marques, M. M. B. Synthesis and Evaluation of New Benzimidazole-based COX Inhibitors: A Naproxen-like Interaction Detected By STD-NMR. *RSC Adv.* **2015**, *5* (61), 49098–49109.

(339) Dalvit, C. Ligand- and Substrate-based 19F NMR Screening: Principles and Applications to Drug Discovery. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51* (4), 243–271.

(340) Mayer, M.; Meyer, B. Group Epitope Mapping by Saturation Transfer Difference NMR To Identify Segments of a Ligand in Direct Contact with a Protein Receptor. *J. Am. Chem. Soc.* **2001**, *123* (25), 6108–6117.

(341) Zhu, Y.; Chen, H.; Boulton, S.; Mei, F.; Ye, N.; Melacini, G.; Zhou, J.; Cheng, X. Biochemical and Pharmacological Characterizations of ESI-09 Based EPAC Inhibitors: Defining the ESI-09 "Therapeutic Window". *Sci. Rep.* **2015**, *5*, 9344.

(342) Chen, D.; Errey, J. C.; Heitman, L. H.; Marshall, F. H.; IJzerman, A. P.; Siegal, G. Fragment Screening of GPCRs Using Biophysical Methods: Identification of Ligands of the Adenosine A 2A Receptor with Novel Biological Activity. *ACS Chem. Biol.* **2012**, *7* (12), 2064–2073.

(343) Vulpetti, A.; Hommel, U.; Landrum, G.; Lewis, R.; Dalvit, C. Design and NMR-based screening of LEF, a library of chemical fragments with different local environment of fluorine. *J. Am. Chem. Soc.* **2009**, *131* (36), 12949–12959.

(344) Dalvit, C.; Mongelli, N.; Papeo, G.; Giordano, P.; Veronesi, M.; Moskau, D.; Kummerle, R. Sensitivity Improvement in 19F NMR-based Screening Experiments: Theoretical Considerations and Experimental Applications. *J. Am. Chem. Soc.* **2005**, *127* (38), 13380–13385.

(345) Dalvit, C.; Fagerness, P. E.; Hadden, D. T. A.; Sarver, R. W.; Stockman, B. J. Fluorine-NMR Experiments for High-throughput Screening: Theoretical Aspects, Practical Considerations, and Range of Applicability. *J. Am. Chem. Soc.* **2003**, *125* (25), 7696–7703.

(346) Assemat, O.; Antoine, M.; Fourquez, J.; Wierzbicki, M.; Charton, Y.; Hennig, P.; Perron-Sierra, F.; Ferry, G.; Boutin, J. A.; Delsuc, M. ^{19}F Nuclear Magnetic Resonance Screening of Glucokinase Activators. *Anal. Biochem.* **2015**, *477*, 62–68.

(347) Ge, X.; MacRaild, C. A.; Devine, S. M.; Debono, C. O.; Wang, G.; Scammells, P. J.; Scanlon, M. J.; Anders, R. F.; Foley, M.; Norton, R. S. Ligand-induced conformational change of Plasmodium falciparum AMA1 detected using 19F NMR. *J. Med. Chem.* **2014**, *57* (15), 6419–6427.

(348) Pomerantz, W. C.; Wang, N.; Lipinski, A. K.; Wang, R.; Cierpicki, T.; Mapp, A. K. Profiling the dynamic interfaces of fluorinated transcription complexes for ligand discovery and characterization. *ACS Chem. Biol.* **2012**, *7* (8), 1345–1350.

(349) Kalbitzer, H. R.; Rosnizeck, I. C.; Munte, C. E.; Narayanan, S. P.; Kropf, V.; Spoerner, M. Intrinsic Allosteric Inhibition of Signaling Proteins by Targeting Rare Interaction States Detected by High-Pressure NMR Spectroscopy. *Angew. Chem. Int. Ed.* **2013**, *52* (52), 14242–14246.

(350) Rosnizeck, I. C.; Filchtinski, D.; Lopes, R. P.; Kieninger, B.; Herrmann, C.; Kalbitzer, H. R.; Spoerner, M. Elucidating the Mode of Action of a Typical Ras State 1(T) Inhibitor. *Biochemistry* **2014**, *53* (24), 3867–3878.

(351) Mizoue, L. S.; Chazin, W. J. Engineering and design of ligand-induced conformational change in proteins. *Curr. Opin. Struct. Biol.* **2002**, *12* (4), 459–463.

(352) Dings, R. P. M.; Miller, M. C.; Nesmelova, I.; Astorgues-Xerri, L.; Kumar, N.; Serova, M.; Chen, X.; Raymond, E.; Hoye, T. R.; Mayo, K. H. Antitumor Agent Calixarene 0118 Targets Human Galectin-1 as an Allosteric Inhibitor of Carbohydrate Binding. *J. Med. Chem.* **2012**, *55* (11), 5121–5129.

(353) Lee, G. M. Enzyme Inhibition by Allosteric Capture of an Inactive Conformation. *J. Mol. Biol.* **2011**, *411* (5), 999–1016.

(354) Pagano, K.; Torella, R.; Foglieni, C.; Bugatti, A.; Tomaselli, S.; Zetta, L.; Presta, M.; Rusnati, M.; Taraboletti, G.; Colombo, G., et al. Direct and allosteric inhibition of the FGF2/HSPGs/FGFR1 ternary complex formation by an antiangiogenic, thrombospondin-1-mimic small molecule. *PloS one* **2012**, *7* (5), e36990.

(355) Rousaki, A.; Miyata, Y.; Jinwal, U. K.; Dickey, C. A.; Gestwicki, J. E.; Zuiderweg, E. R. Allosteric Drugs: The Interaction of Antitumor Compound MKT-077 with Human Hsp70 Chaperones. *J. Mol. Biol.* **2011**, *411* (3), 614–632.

(356) Li, X.; Srinivasan, S. R.; Connarn, J.; Ahmad, A.; Young, Z. T.; Kabza, A. M.; Zuiderweg, E. R. P.; Sun, D.; Gestwicki, J. E. Analogs of the Allosteric Heat Shock Protein 70 (Hsp70) Inhibitor, MKT-077, as Anti-Cancer Agents. *ACS Med. Chem. Lett.* **2013**, *4* (11), 1042–1047.

(357) Vashisth, H.; Storaska, A. J.; Neubig, R. R.; Brooks, C. L. Conformational dynamics of a regulator of G-protein signaling protein reveals a mechanism of allosteric inhibition by a small molecule. *ACS Chem. Biol.* **2013**, *8* (12), 2778–2784.

(358) Reverdatto, S.; Rai, V.; Xue, J.; Burz, D. S.; Schmidt, A. M.; Shekhtman, A. Combinatorial library of improved peptide aptamers, CLIPs to inhibit RAGE signal transduction in mammalian cells. *PloS one* **2013**, *8* (6), e65180.

(359) Rosnoblet, C.; Fritzinger, B.; Legrand, D.; Launay, H.; Wieruszeski, J.-M.; Lippens, G.; Hanoulle, X. Hepatitis C Virus NS5B and Host Cyclophilin A Share a Common Binding Site on NS5A. *J. Biol. Chem.* **2012**, *287* (53), 44249–44260.

(360) Christen, M. T.; Menon, L.; Myshakina, N. S.; Ahn, J.; Parniak, M. A.; Ishima, R. Structural Basis of the Allosteric Inhibitor Interaction on the HIV-1 Reverse Transcriptase RNase H Domain. *Chem. Biol. Drug Des.* **2012**, *80* (5), 706–716.

(361) Chang, L.; Miyata, Y.; Ung, P. M. U.; Bertelsen, E. B.; McQuade, T. J.; Carlson, H. A.; Zuiderweg, E. R. P.; Gestwicki, J. E. Chemical Screens Against a Reconstituted Multiprotein Complex: Myricetin Blocks DnaJ Regulation of DnaK Through an Allosteric Mechanism. *Chem. Biol.* **2011**, *18* (2), 210–221.

(362) Placzek, W. J.; Sturlese, M.; Wu, B.; Cellitti, J. F.; Wei, J.; Pellecchia, M. Identification of a novel Mcl-1 protein binding motif. *J. Biol. Chem.* **2011**, *286* (46), 39829–39835.

(363) San Sebastián, E.; Zimmerman, T.; Zubia, A.; Vara, Y.; Martin, E.; Sirockin, F.; Dejaegere, A.; Stote, R. H.; Lopez, X.; Pantoja-Uceda, D., et al. Design, synthesis, and functional evaluation of leukocyte function associated antigen-1 antagonists in early and late stages of cancer development. *J. Med. Chem.* **2013**, *56* (3), 735–747.

(364) Dings, R. P. M.; Kumar, N.; Miller, M. C.; Loren, M.; Rangwala, H.; Hoye, T. R.; Mayo, K. H. Structure-based Optimization of Angiostatic Agent 6DBF7, an Allosteric Antagonist of Galectin-1. *J. Pharmacol. Exp. Ther.* **2013**, *344* (3), 589–599.

(365) Ragnarsson, L.; Wang, C.-I. A.; Andersson, Å.; Fajarningsih, D.; Monks, T.; Brust, A.; Rosengren, K. J.; Lewis, R. J. Conopeptide Rho-TIA Defines a New Allosteric Site on the Extracellular Surface of the Alpha1B-Adrenoceptor. *J. Biol. Chem.* **2013**, *288* (3), 1814–1827.

(366) Huang, H.; Vogel, H. J. Structural basis for the activation of platelet integrin αIIbβ3 by calcium- and integrin-binding protein 1. *J. Am. Chem. Soc.* **2012**, *134* (8), 3864–3872.

(367) VanSchouwen, B.; Selvaratnam, R.; Giri, R.; Lorenz, R.; Herberg, F. W.; Kim, C.; Melacini, G. Mechanism of cAMP Partial Agonism in Protein Kinase G (PKG). *J. Biol. Chem.* **2015**, *290* (48), 28631–28641.

(368) Tomaselli, S.; Pagano, K.; Boulton, S.; Zanzoni, S.; Melacini, G.; Molinari, H.; Ragona, L. Lipid Binding Protein Response to a Bile Acid Library: a Combined NMR and Statistical Approach. *FEBS J.* **2015**, *282* (21), 4094–4113.

(369) Hammoudeh, D. I.; Daté, M.; Yun, M.-K.; Zhang, W.; Boyd, V. A.; Viacava Follis, A.; Griffith, E.; Lee, R. E.; Bashford, D.; White, S. W. Identification and characterization of an allosteric inhibitory site on dihydropteroate synthase. *ACS Chem. Biol.* **2014**, *9* (6), 1294–1302.

(370) Zoetewey, D. L.; Ovee, M.; Banerjee, M.; Bhaskaran, R.; Mohanty, S. Promiscuous Binding at the Crossroads of Numerous Cancer Pathways: Insight From the Binding of Glutaminase Interacting Protein With Glutaminase L. *Biochemistry* **2011**, *50* (17), 3528–3539.

(371) Zhang, P.; Leger, A. J.; Baleja, J. D.; Rana, R.; Corlin, T.; Nguyen, N.; Koukos, G.; Bohm, A.; Covic, L.; Kuliopulos, A. Allosteric Activation of a G Protein-coupled Receptor with Cell-penetrating Receptor Mimetics. *J. Biol. Chem.* **2015**, *290* (25), 15785–15798.

(372) Zhang, Z.; Cellitti, J.; Teriete, P.; Pellecchia, M.; Stec, B. New Crystal Structures of HSC-70 Atp Binding Domain Confirm the Role of Individual Binding Pockets and Suggest a New Method of Inhibition. *Biochimie* **2015**, *108*, 186–192.

(373) Balakrishnan, A.; Jordan, F.; Nathan, C. F. Influence of Allosteric Regulators on Individual Steps in the Reaction Catalyzed by Mycobacterium tuberculosis 2-hydroxy-3-oxoadipate synthase. *J. Biol. Chem.* **2013**, *288* (30), 21688–21702.

(374) Kaczor, A. A.; Karczmarzyk, Z.; Fruziński, A.; Pihlaja, K.; Sinkkonen, J.; Wiinämaki, K.; Kronbach, C.; Unverferth, K.; Poso, A.; Matosiuk, D. Structural studies, homology modeling and molecular docking of novel non-competitive antagonists of GluK1/GluK2 receptors. *Bioorg. Med. Chem.* **2014**, *22* (2), 787–795.

(375) Papadopoulos, E.; Jenni, S.; Kabha, E.; Takrouri, K. J.; Yi, T.; Salvi, N.; Luna, R. E.; Gavathiotis, E.; Mahalingam, P.; Arthanari, H., et al. Structure of the eukaryotic translation initiation factor eIF4E in complex with 4EGI-1 reveals an allosteric mechanism for dissociating eIF4G. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111* (31), E3187-95.

(376) Meli, M.; Pagano, K.; Ragona, L.; Colombo, G. Investigating the dynamic aspects of drug-protein recognition through a combination of MD and NMR analyses: implications for the development of protein-protein interaction inhibitors. *PloS one* **2014**, *9* (5), e97153.

(377) Ronca, R.; Giacomini, A.; Di Salle, E.; Coltrini, D.; Pagano, K.; Ragona, L.; Matarazzo, S.; Rezzola, S.; Maiolo, D.; Torrella, R., et al. Long-Pentraxin 3 Derivative as a Small-Molecule FGF Trap for Cancer Therapy. *Cancer Cell* **2015**, *28* (2), 225–239.

(378) Sugitani, N.; Shell, S. M.; Soss, S. E.; Chazin, W. J. Redefining the DNA-binding domain of human XPA. *J. Am. Chem. Soc.* **2014**, *136* (31), 10830–10833.

(379) Tsai, C. J.; Nussinov, R. The Free Energy Landscape in Translational Science: How Can Somatic Mutations Result in Constitutive Oncogenic Activation? *Phys. Chem. Chem. Phys.* **2014**, *16* (14), 6332–6341.

(380) Yun, J.-H.; Kim, M.; Kim, K.; Lee, D.; Jung, Y.; Oh, D.; Ko, Y.-J.; Cho, A. E.; Cho, H.-S.; Lee, W. Solution structure of the transmembrane 2 domain of the human melanocortin-4 receptor in sodium dodecyl sulfate (SDS) micelles and the functional implication of the D90N mutant. *Biochim. Biophys. Acta* **2015**, *1848* (6), 1294–1302.

(381) Chandak, M. S.; Nakamura, T.; Makabe, K.; Takenaka, T.; Mukaiyama, A.; Chaudhuri, T. K.; Kato, K.; Kuwajima, K. The H/D-exchange Kinetics of the Escherichia coli Co-chaperonin GroES Studied by 2D NMR and DMSO-quenched Exchange Methods. *J. Mol. Biol.* **2013**, *425* (14), 2541–2560.

(382) Smith, M. J.; Neel, B. G.; Ikura, M. NMR-based functional profiling of RASopathies and oncogenic RAS mutations. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (12), 4574–4579.

(383) Guo, B.; Gurel, P. S.; Shu, R.; Higgs, H. N.; Pellegrini, M.; Mierke, D. F. Monitoring ATP hydrolysis and ATPase inhibitor screening using (1)H NMR. *Chem. Commun.* **2014**, *50* (81), 12037–12039.

(384) Eicher, J. J.; Snoep, J. L.; Rohwer, J. M. Determining Enzyme Kinetics for Systems Biology With Nuclear Magnetic Resonance Spectroscopy. *Metabolites* **2012**, *2* (4), 818–843.

(385) Belenky, I.; Steinmetz, A.; Vyazmensky, M.; Barak, Z.; Tittmann, K.; Chipman, D. M. Many of the Functional Differences Between Acetohydroxyacid Synthase (AHAS) Isozyme I and Other AHASs Are a Result of the Rapid Formation and Breakdown of the Covalent Acetolactate-Thiamin Diphosphate Adduct in AHAS I. *FEBS J.* **2012**, *279* (11), 1967–1979.

(386) Ferreiro, D. U.; Komives, E. A.; Wolynes, P. G. Frustration in Biomolecules. *Q. Rev. Biophys.* **2014**, *47* (4), 285–363.

(387) Ferreiro, D. U.; Hegler, J. A.; Komives, E. A.; Wolynes, P. G. On the Role of Frustration in the Energy Landscapes of Allosteric Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (9), 3499–3503.

(388) Zheng, W.; Schafer, N. P.; Wolynes, P. G. Frustration in the Energy Landscapes of Multidomain Protein Misfolding. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (5), 1680–1685.

(389) Lisi, G. P.; Loria, J. P. Solution NMR Spectroscopy for the Study of Enzyme Allostery. *Chem. Rev.* **2016**.

(390) Vogeli, B.; Orts, J.; Strotz, D.; Chi, C.; Minges, M.; Walti, M. A.; Guntert, P.; Riek, R. Towards a true protein movie: a perspective on the potential impact of the ensemble-based structure determination using exact NOEs. *Journal of magnetic resonance (San Diego, Calif. : 1997)* **2014**, *241*, 53–59.

(391) Guerry, P.; Herrmann, T. Advances in Automated NMR Protein Structure Determination. *Quarterly reviews of biophysics* **2011**, *44* (3), 257–309.

(392) Venditti, V.; Egner, T. K.; Clore, G. M. Hybrid Approaches to Structural Characterization of Conformational Ensembles of Complex Macromolecular Systems Combining NMR Residual Dipolar Couplings and Solution X-ray Scattering. *Chem. Rev.* **2016**.

(393) Kleckner, I. R.; Foster, M. P. An Introduction to NMR-Based Approaches For Measuring Protein Dynamics. *Biochimica et biophysica acta* **2011**, *1814* (8), 942–968.

(394) Guo, J.; Zhou, H.-X. Protein Allostery and Conformational Dynamics. *Chem. Rev.* **2016**.

(395) Palmer, A. G. NMR Characterization of the Dynamics of Biomacromolecules. *Chem. Rev.* **2004**, *104* (8), 3623–3640.

(396) Redfield, C. Using Nuclear Magnetic Resonance Spectroscopy to Study Molten Globule States of Proteins. *Methods (San Diego, Calif.)* **2004**, *34* (1), 121–132.

(397) Mazhab-Jafari, M. T.; Das, R.; Fotheringham, S. A.; SilDas, S.; Chowdhury, S.; Melacini, G. Understanding cAMP-Dependent Allostery by NMR Spectroscopy: Comparative Analysis of the EPAC1 cAMP-Binding Domain in its Apo and cAMP-Bound States. *Journal of the American Chemical Society* **2007**, *129* (46), 14482–14492.

(398) Bhunia, A.; Bhattacharjya, S.; Chatterjee, S. Applications of Saturation Transfer Difference NMR in Biological Systems. *Drug discovery today* **2012**, *17* (9-10), 505–513.

(399) Milanesi, R.; Baruscotti, M.; Gnecchi-Ruscone, T.; DiFrancesco, D. Familial sinus bradycardia associated with a mutation in the cardiac pacemaker channel. *New Engl. J. Med.* **2006,** 354, 151–157.

(400) Xu, X.; Marni, F.; Wu, S.; Su, Z.; Musayev, F.; Shrestha, S.; Xie, C.; Gao, W.; Liu, Q.; Zhou, L. Local and Global Interpretations of a Disease-Causing Mutation near the Ligand Entry Path in Hyperpolarization-Activated cAMP-Gated Channel. *Structure* **2012,** 20, 2116–2123.

(401) Boulton, S.; Akimoto, M.; Akbarizadeh, S.; Melacini, G. Free Energy Landscape Remodeling for the Cardiac Pacemaker Channel Explains the Molecular Basis of Familiar Sinus Bradycardia. *J. Biol. Chem.* **2017**, 292(15), 6414-6428.

(402) Feng, B. Y.; Shelat, A.; Doman, T. N.; Guy, R. K.; Shoichet, B. K. High-throughput assays for promiscuous inhibitors. *Nature chemical biology* **2005**, *1*, 146–148.

(403) McGovern, S. L.; Helfand, B. T.; Feng, B.; Shoichet, B. K. A specific mechanism of nonspecific inhibition. *J. Med. Chem.* **2003**, *46*, 4265–4272.

(404) Feng, B. Y.; Simeonov, A.; Jadhav, A.; Babaoglu, K.; Inglese, J.; Shoichet, B. K.; Austin, C. P. A high-throughput screen for aggregation-based inhibition in a large compound library. *Journal of medicinal chemistry* **2007**, *50*, 2385–2390.

Chapter 2

Free Energy Landscape Remodeling for the Cardiac Pacemaker Channel Explains the Molecular Basis of Familiar Sinus Bradycardia

2.1 Author's Preface

The work presented in this chapter has previously been published and is reproduced here with permission from the American Society for Biochemistry and Molecular Biology. The full citation is as follows:

Boulton S, Akimoto M, Akbarizadeh S and Melacini G. Free Energy Landscape Remodeling for the Cardiac Pacemaker Channel Explains the Molecular Basis of Familiar Sinus Bradycardia. *J. Biol. Chem.* **292**(15), 6414-6428 (2017).

I conducted most of the experiments necessary for this publication. Dr. Madoka Akimoto and I assigned the backbone amide chemical shifts for the apo and cAMP-bound states of wild type HCN4. Sam Akbarizadeh assisted with the preparation of the S672R HCN4 variant and with the acquisition of some NH-HSQC spectra. I co-wrote the manuscript with Dr. Giuseppe Melacini.

2.2 Abstract

The hyperpolarization-activated and cyclic nucleotide-modulated ion channel (HCN) drives the pacemaker activity in the heart, and its malfunction can result in heart disorders. One such disorder, familial sinus bradycardia, is caused by the S672R mutation in HCN4, whose electrophysiological phenotypes include a negative shift in the channel activation voltage and an accelerated HCN4 deactivation. The outcomes of these changes are abnormally low resting heart rates. However, the molecular mechanism underlying these electrophysiological changes is currently not fully understood. Crystallographic investigations indicate that the S672R mutation causes limited changes in the structure of the HCN4 intracellular gating tetramer, but its effects on protein dynamics are unknown. Here, we utilize comparative S672R versus WT NMR analyses to show that the S672R mutation results in extensive perturbations of the dynamics in both apo- and holo-forms of HCN4, reflecting how S672R remodels the free energy landscape for the modulation of HCN4 by cAMP, i.e. the primary cyclic nucleotide modulator of HCN channels. We show that the S672R mutation results in a constitutive shift of the dynamic auto-inhibitory equilibrium toward inactive states of HCN4 and broadens the free-energy well of the apo-form, enhancing the millisecond to microsecond dynamics of the holo-form at sites critical for gating cAMP binding. These S672R-induced variations in dynamics provide a molecular basis for the electrophysiological phenotypes of this mutation and demonstrate that the pathogenic effects of the S672R mutation can be rationalized primarily in terms of modulations of protein dynamics.

2.3 Introduction

The cardiac pacemaker current is regulated by hyperpolarization-activated cyclic nucleotide-modulated (HCN) ion channels, and in particular, by the isoform HCN4, which is the most abundant in the sinoatrial node(1, 2). These channels open during the diastolic phase resulting in an ion influx that cause a depolarizing shift in the membrane's resting potential towards its activation threshold. The channel opening is predominantly regulated by a voltage sensor in its trans-membrane domain (TMD) that responds to membrane hyperpolarization. However, the intracellular region (IR) provides an additional means of modulation via cyclic nucleotide-regulated auto-inhibition(3–7) The IR includes two domains critical for regulating channel gating: the C-linker and the cAMP-binding domain (CBD) (Fig 1A). The C-linker comprises six helices (A'-F') that facilitate cAMP dependent activation by promoting tetramerization through intermolecular interactions between the A' and B' helices of one protomer and the C' and D' helices of another (Fig 1A)(8, 4, 7, 6). The tetramerization competency of the C-linker is influenced by conformational changes within the CBD that occur upon cAMP binding(9, 10).

The CBD consists of a contiguous β-subdomain and a non-contiguous α-subdomain (Fig 1B). The β-subdomain spans the primary cAMP-binding site, which includes a variable β-hairpin motif referred to as the base binding region (BBR; Fig 1B) and a highly conserved anchoring site, known as the phosphate binding cassette (PBC; Fig 1B, Fig. S1). In addition, cAMP binding is gated by a C-terminal helix (or C-helix in short) and a segment following the C-helix (or C-terminal Lid, CTL). Together the C-helix and the CTL define the CBD lid (Fig 1B), which is indispensable for cAMP signalling via HCN channels(9, 11). Deletion of the lid, even partially, is sufficient to render HCN channel gating insensitive to cAMP(9, 11). This is because the cAMP-driven lid closure is coupled to conformational changes in the N-terminal α-subdomain (N3A) (Fig 1B), which controls the IR tetramerization by removing steric clashes between the C-linker and the rigid β-subdomain of the apo CBD(9, 8, 12, 13). In the absence of the steric clashes arising from the apo CBD, the C-linker forms a stable IR tetramer and the tonic HCN auto-inhibition is removed(9).

The coupling between cAMP binding, CBD activation and channel gating is of central importance to the physiological function of HCN(9, 14). Mutations that perturb or disrupt this coupling cause pathological conditions for the heart and pacemaker current(15, 11, 16–19). For example, alteration of the highly conserved S672 residue in the HCN4 CBD by the heterozygous S672R mutation results in bradycardia, i.e. abnormally low resting heart rates that are reduced by ~30%(17, 18). Electrophysiological profiling using integral HCN4 channels revealed that the

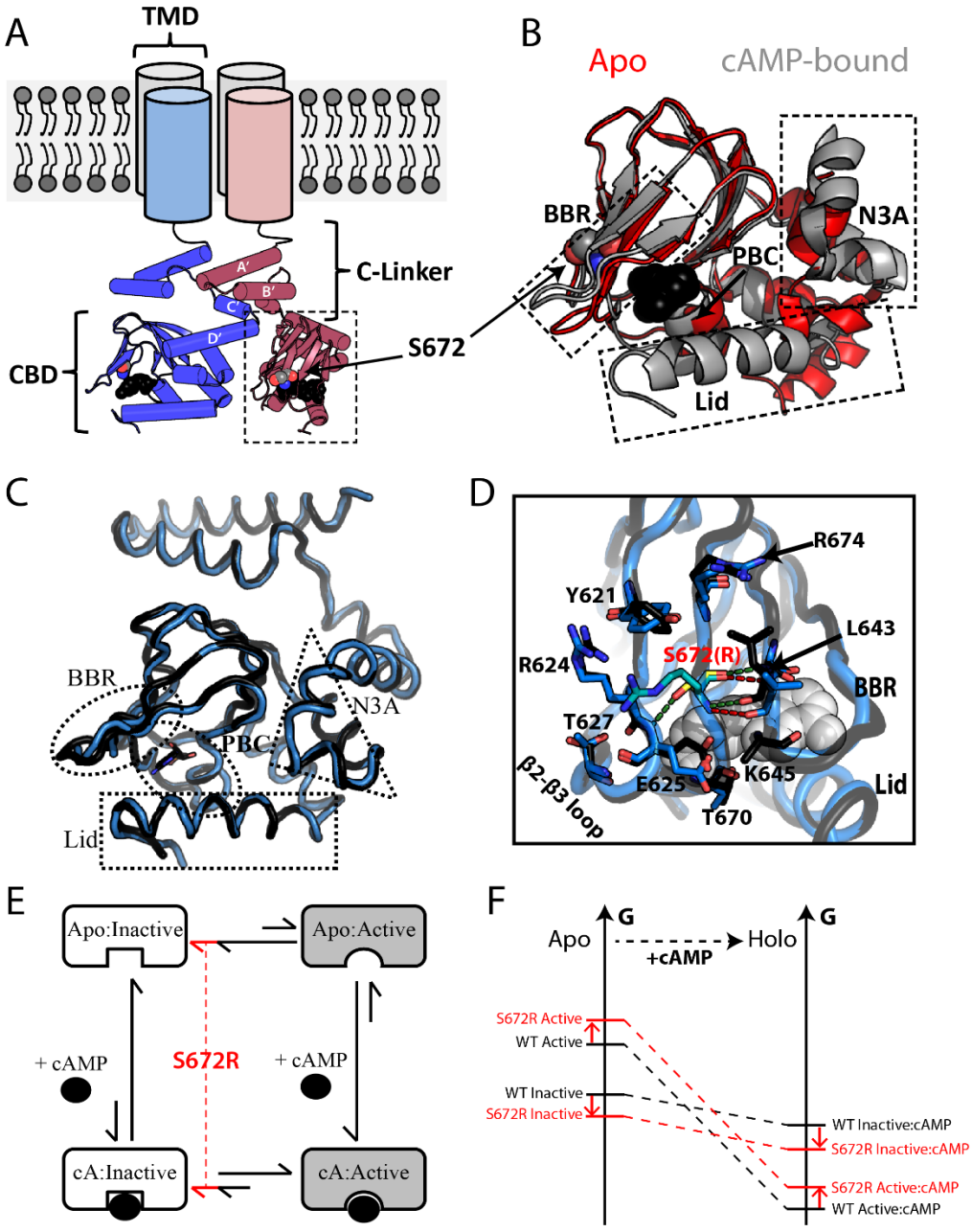

Figure 1: Structural Architecture of the HCN4 Intracellular Region. **A)** Topology of tetrameric HCN4 channels. The trans-membrane domain (TMD) is connected to the cAMP-binding domain (CBD) through an α-helical domain known as the C-linker, which also serves as the main point of contact between protomers within the gating tetramer. S672 and cAMP are shown as gray and black spheres, respectively. **B)** Overlay of the apo (red; PDB ID: 2MNG) and cAMP-bound (gray; PDB ID: 3OTF) CBDs of WT HCN4. Select regions (ie. N3A, BBR, and Lid) that undergo cAMP-dependent conformational changes are indicated with dashed boxes. cAMP and S672 are depicted as in panel A. **C)** Structural overlay of holo WT (black; PDB ID: 3OTF) and holo S672R (blue; PDB ID: 4HBN) HCN4 CBDs. **D)** Changes in key contacts mediated by S672 (black ribbon, yellow sticks) and S672R (blue ribbon, cyan sticks). Polar contacts for S672 and S672R are depicted with green and red dashed lines, respectively. Selected residues that form interactions with

S672(R) or that experience side-chain reorientations are shown as sticks. cAMP is shown as white spheres. **E)** Four state thermodynamic cycle to model allostery for the HCN cAMP binding domain (CBD), in which the auto-inhibitory equilibrium is coupled to the cAMP binding equilibrium through the active *vs.* inactive state selectivity of cAMP. A hypothetical mechanism for the bradycardia-induced S672R mutant posits that the auto-inhibitory CBD equilibrium is shifted by the mutation further towards the inactive state (red lines). **F)** Simplified free energy diagram illustrating that the hypothesized mutation-induced shift of the auto-inhibitory equilibrium toward the inactive state may arise either from a stabilization of the inactive state and/or from a destabilization of the active state. This scheme assumes that the state-specific association constants are not significantly affected by the mutation.

S672R HCN4 variant is activated at voltages that are ~10% more negative than wild type in either the absence or the presence of endogenous cAMP(18). The negative S672R vs. WT voltage shift was independently confirmed using chimeric constructs in which the HCN2 trans-membrane channel was combined with the HCN4 IR(20), indicating that S672R represents a loss-of-function mutation with enhanced resistance to channel opening, i.e. S672R causes a reduced inward current flow in the diastolic phase, explaining the slower than normal heart rate observed in patients affected by the mutation(18, 17). Furthermore, another S672R-incuded change consistently revealed by electrophysiology is a reduction on the deactivation time constant, i.e. the S672R mutation causes a significant acceleration of HCN channel deactivation(20, 18).

As a first step towards elucidating the molecular mechanism for the observed electrophysiological phenotypes of the bradycardia S672R mutation, the structure of the S672R HCN4 IR mutant was solved in the presence of cAMP (Fig 1C)(20). However, the crystal structure of the cAMP-bound S672R HCN4 IR tetramer did not reveal any major change relative to wild type (Fig 1C; RMSD 0.41 Å), with the exception of the BBR (Fig 1C), which exhibited increased local disorder for four loop residues, as indicated by missing electron density and side chain re-packing of selected surrounding residues (Fig. 1D)(5, 8, 20). No appreciable variations in subunit assembly as a result of the mutation were detected(5, 8, 20). While these initial structural results provide an essential framework for further investigations of the S672R variant, several questions about the molecular mechanism underlying the S672R bradycardia-inducing mutation remain still open.

It is still not clear how the S672R mutation perturbs the apo CBD and in general the CBD dynamics, which has been recognized as a key determinant of the auto-inhibitory CBD function in HCN gating and in other cAMP-dependent systems(21–29, 9, 13). One of the aspects of CBD dynamics that is most relevant for auto-inhibition is the equilibrium between inactive and active conformations (30–38) (Fig 1B and 1E). The coupling of this two-state dynamic conformational equilibrium with the cAMP-binding equilibrium results in a four-state thermodynamic cycle (Fig 1E), which provides the simplest allosteric model to rationalize the cAMP-dependent modulation

of HCN channels, i.e. a shift to less negative activation voltages, as well as to rationalize the effect of disease related mutations. Considering that the S672R mutant leads to a shift to more negative activation voltages, i.e. an effect opposite to cAMP, while still preserving similar K1/2 values for integral HCN4 channels (17, 18), we hypothesize that the S672R mutation compromises the cAMP-dependent activation either by giving rise to a partial reversal of the two-state auto-inhibitory equilibrium towards the inactive conformation (Fig 1E and 1F) and/or by partially stabilizing additional states, distinct from the fully active or fully inactive states, in which the activation function of the CBD is compromised. To test these hypotheses, here we comparatively analyze the S672R and WT HCN4 CBDs in both apo and holo forms using NMR spectroscopy, which reports on dynamics at residue resolution and over multiple time-scales (30, 39). Furthermore, we examined the S672R vs. WT relaxation and chemical shift changes, which are exquisitely sensitive atomic reporters of subtle, but functionally relevant allosteric perturbations that are often elusive to other structural techniques (30, 27).

2.4 Results

2.4.1 The S672R Mutation Leads to Pervasive Perturbations of the HCN4 CBD in Both Apo and Holo Forms.

As a first step towards understanding how the bradycardia-linked S672R mutation affects the apo and holo HCN4 CBD, we assigned the ^{1}H-^{15}N HSQC spectra of the S672R HCN4 (563-724) construct (Fig. S2) in the absence of cAMP (Fig 2A) and in the presence of saturating amounts of cAMP (Fig 2B). The comparison with the corresponding spectra of WT HCN4 (563-724)(9) reveals that the S672R mutation results in pervasive changes for both apo and holo forms (Fig 2A-C), well beyond what previously anticipated based on X-ray structure comparisons (Fig 1C)(20). In both apo and holo forms, the most pronounced chemical shift changes are observed not only at the site of the mutation, but extend to loci located within a 15 Å radius from S672 (Fig 2C and 2D), including the β2-3 loop, the PBC, the BBR and, in the case of the holo sample, the C-terminal lid as well (Fig 2C). In addition, several residues throughout the β-subdomain and the C-terminal helix in both apo and holo forms exhibits notable ppm changes (> 0.1 ppm; Fig 2C,E,F).

Some of the largest S672R *vs.* WT ppm variations arise either from losses of key interactions mediated by the S672 hydroxyl, as in the case of E625 and adjacent residues in the β2-3 loop (Fig 1D), or from local structural changes, as in the case of the BBR that exhibits prominent S672R *vs.*

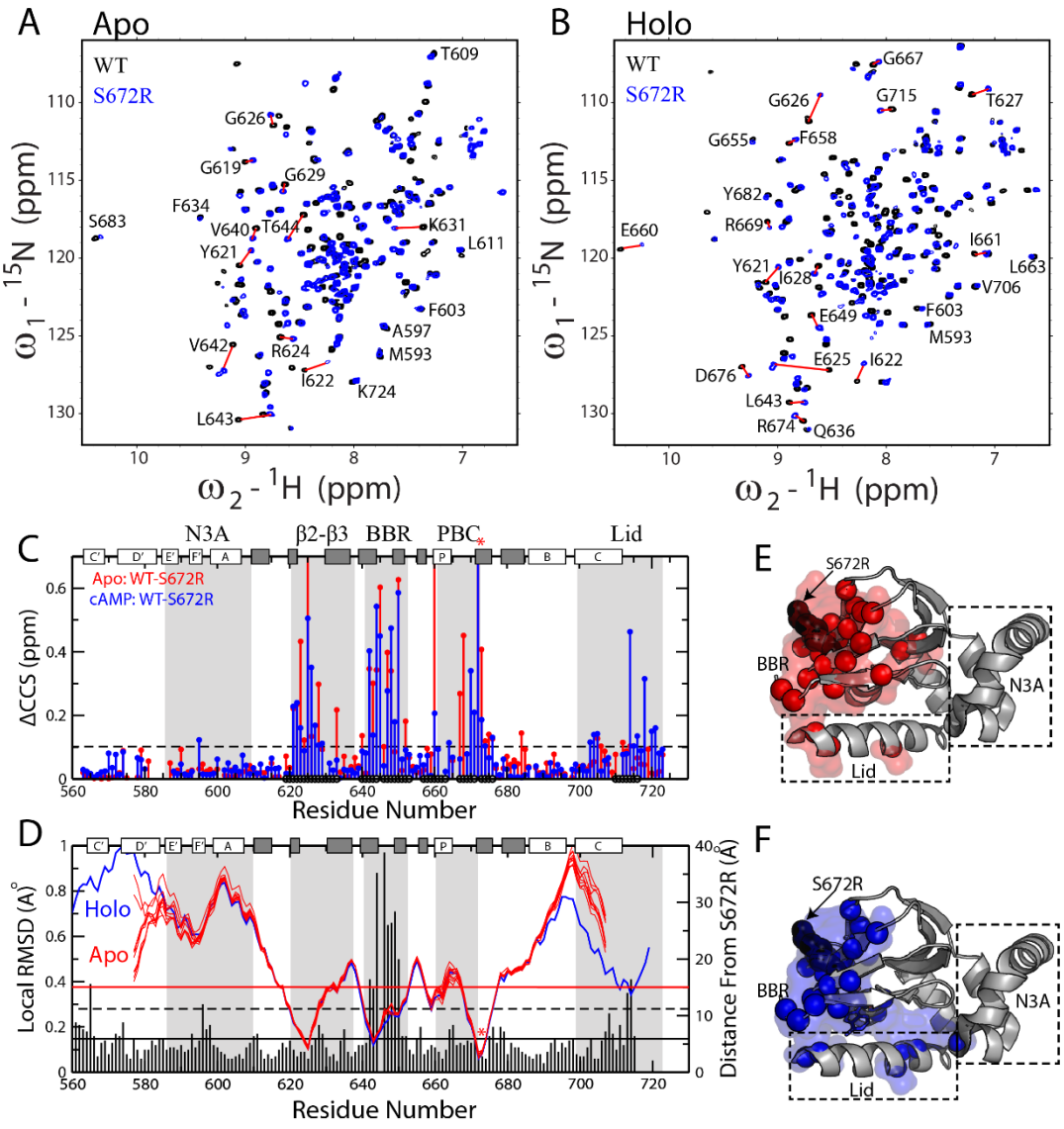

Figure 2: WT vs. S672R Chemical Shift Perturbations in the Apo and Holo HCN4 CBD. Overlay of NH-HSQC spectra for WT (black) and S672R (blue) HCN4 (563-724) in the apo **(A)** and cAMP-bound **(B)** samples. **C)** Compounded Chemical shift (CCS) differences between the WT and S672R variant for the apo (red) and cAMP-bound (blue) samples. The black circles along the bottom of the plot depict residues within 15 Å of the mutation site. The protein's secondary structure is depicted along the top as white (α-helices) and gray (β-strands) rectangles, while the BBR and several other motifs are highlighted with gray background. **D)** Residue-specific local WT *vs.* S672R RMSDs (vertical bars) computed in Molmol**(56)** by aligning three adjacent residues from the WT and S672R cAMP bound X-ray structures (PDB IDs: 3OTF(8) and 4HBN(20), respectively), as was previously described(41, 27). The solid and dashed black horizontal lines depict the average and average + SD RMSD values, while the horizontal red line represents a 15 Å distance from S672R. The blue and red lines represent the backbone amide nitrogen distances from residue S672(R) for the cAMP bound structure (4HBN) and the apo NMR ensemble (2MNG), respectively. The mutation site (S672R) is highlighted with a red asterisk in both panels (C) and (D). **(E-F)** Residues with CCS changes greater than the dashed line in panel C were mapped onto the crystal structure as spheres and surface for both the apo **(E)** and the cAMP-bound **(F)** samples.

WT RMSD values based on the available crystal structures (Fig 2D). However, most residues outside of the BBR display only relatively subtle S672R *vs.* WT structural changes, as confirmed by marginal RMSD values (Fig 2D)(8, 20) and by the lack of significant S672R *vs.* WT differences in the secondary structure profiles computed based on the secondary chemical shifts (Fig S3). These observations suggest that the S672R *vs.* WT ppm variations of Fig 2C may also reflect mutation-dependent modulations in dynamics. In fact, chemical shifts of CBDs are known to report on the position of the dynamic auto-inhibitory equilibrium of CBDs (Fig 1E), *i.e.* they encode for the relative inactive *vs.* active state populations. Specifically, NMR peak positions reflect population-weighted averages of the respective ppm values for the pure inactive and active CBD states, which exchange rapidly on the chemical shift NMR time-scale(9, 40). A simple but effective means to extract state-populations from average chemical shifts is the projection analysis(29, 21, 9, 30, 41, 28), in which changes in HSQC cross-peak positions arising from the mutation are evaluated relative to a reference vector defined by the WT apo and WT holo samples. For example, Fig 3A-C shows how the residue-specific S672R holo *vs.* WT holo vectors are analyzed in terms of the angles defined with the reference vectors (θ) and of the normalized projections onto the reference vectors (X). The cos(θ) values (Fig 3A and 3B) report on the direction of the mutation-induced shifts in the inhibitory equilibrium (*i.e.* $\cos(\theta) < 0$ for inactivation) and on deviations from the linearity expected based on pure two-state equilibria (*i.e.* $|\cos(\theta)| < 1$), while the X values (Fig 3C) quantify the degree of mutation-induced fractional inhibition or activation(21, 9, 30, 41, 28, 29).

2.4.2 The S672R Mutation Shifts the Dynamic Inhibitory Equilibrium of the HCN4 CBD Towards an Inactive, but Partially De-Correlated State.

The chemical shift projection analysis (CHESPA) of holo S672R shows that negative cos(θ) values are observed for the majority of the residues (Fig 3A), with a cos(θ) distribution highly skewed towards -1 (Fig 3B), reflecting a clear prevalence of mutation-induced shifts in cross-peak positions from the WT holo towards the WT apo, *i.e.* the auto-inhibitory state. However, the extents of theses shifts, *i.e.* the magnitudes of the X values, are highly residue-dependent (Fig 3B). Residue-specific variations in X values are expected for sites proximal to either the mutation and/or cAMP, such as the BBR and the β2-3 loop (Fig 3A), because these loci are affected by nearest neighbour effects (NNEs). Nevertheless, in our case residue-specific variations in X values are also observed for sites more removed from the mutation and cAMP, such as the N-terminal helices and the C-terminal region (Fig 3C). For example, the X values of CTL residues 714-724 range from -

0.2 to ~ -1 with an average X value of -0.4 (Fig 3C and 3E), while in the N-terminal N3A significantly lower X values close to -0.1 are measured (Fig 3C and 3F). The variations in residue-specific X values represent a clear deviation from the pattern of nearly residue independent X values expected for these regions based on the two-state model (Fig 1E), indicating that the response of the holo CBD to the S672R mutation cannot be fully recapitulated by a simple two-state auto-inhibitory equilibrium (Fig 1E). Deviation from the two-state model are also confirmed by the appearance of $|\cos(\theta)|$ values < 1 for several residues, including those not affected by NNE(27) (Fig 3A,D), pointing to the presence of non-linear patterns (Fig. 3G). Overall, the CHESPA analysis of holo S672R reveals that S672R causes a partial reversal of the inhibitory equilibrium towards a less activated state (Fig 3B), but such reversion is not as correlated as expected based on a purely two-state model because it is significantly more pronounced in the C-terminal lid than in other regions largely insensitive to NNEs (Fig 3C,E,F).

To investigate whether the mutation-induced de-correlation is unique to the holo sample or applies also to the apo form, we extended the CHESPA analysis to the apo S672R sample (Fig 3I-L). Although the S672R holo CHESPA analysis clearly shows a shift towards inactivation, this shift is not expected to affect the auto-inhibitory S672R apo CHESPA. The S672R-induced auto-inhibitory shift arises from a partial stabilization of the inactive state and/or from partial destabilization of the active state (Fig 1F). However, in either case, the WT apo inhibitory equilibria of our construct are known to be already almost quantitatively skewed toward the inactive state (Fig 1E)(40, 9). Hence the S672-induced auto-inhibitory shift is expected to be negligible under our experimental conditions, suggesting that the S672R apo CHESPA analysis should reflect primarily de-correlative effects of the mutation, *i.e.* deviations from the two-state model. This prediction is confirmed by the $\cos(\theta)$ values measured for S672R apo (Fig 3I), exhibiting a broad scattered distribution spanning the full -1 to 1 range (Fig 3J), which is in marked contrast to the inactivation-skewed pattern observed for S672R holo (Fig 3B). Furthermore, the abundance of residues with $|\cos(\theta)| < 1$ (Fig 3H and 3J) and the scatter detected in the residue-specific X values (Fig 3K) confirm that the perturbations caused by S672R cannot be modeled exclusively based on a two-state model (Fig 1E). These deviations from the two-state transition point to mutation-induced de-correlations and possibly to enhanced dynamics. The latter were gauged through S672R *vs.* WT comparative ^{15}N-relaxation analyses (Fig 4; Fig S4-S7; Tables S1-S3). Internal dynamics were surveyed using spin relaxation experiments (^{15}N-T_1, ^{15}N-T_2 and ^{1}H^{15}N-NOEs) analyzed in terms of N-H S^2 order parameters (Fig 4A and 4B) and reduced spectral densities (Fig S4-S6; Fig 4C-J).

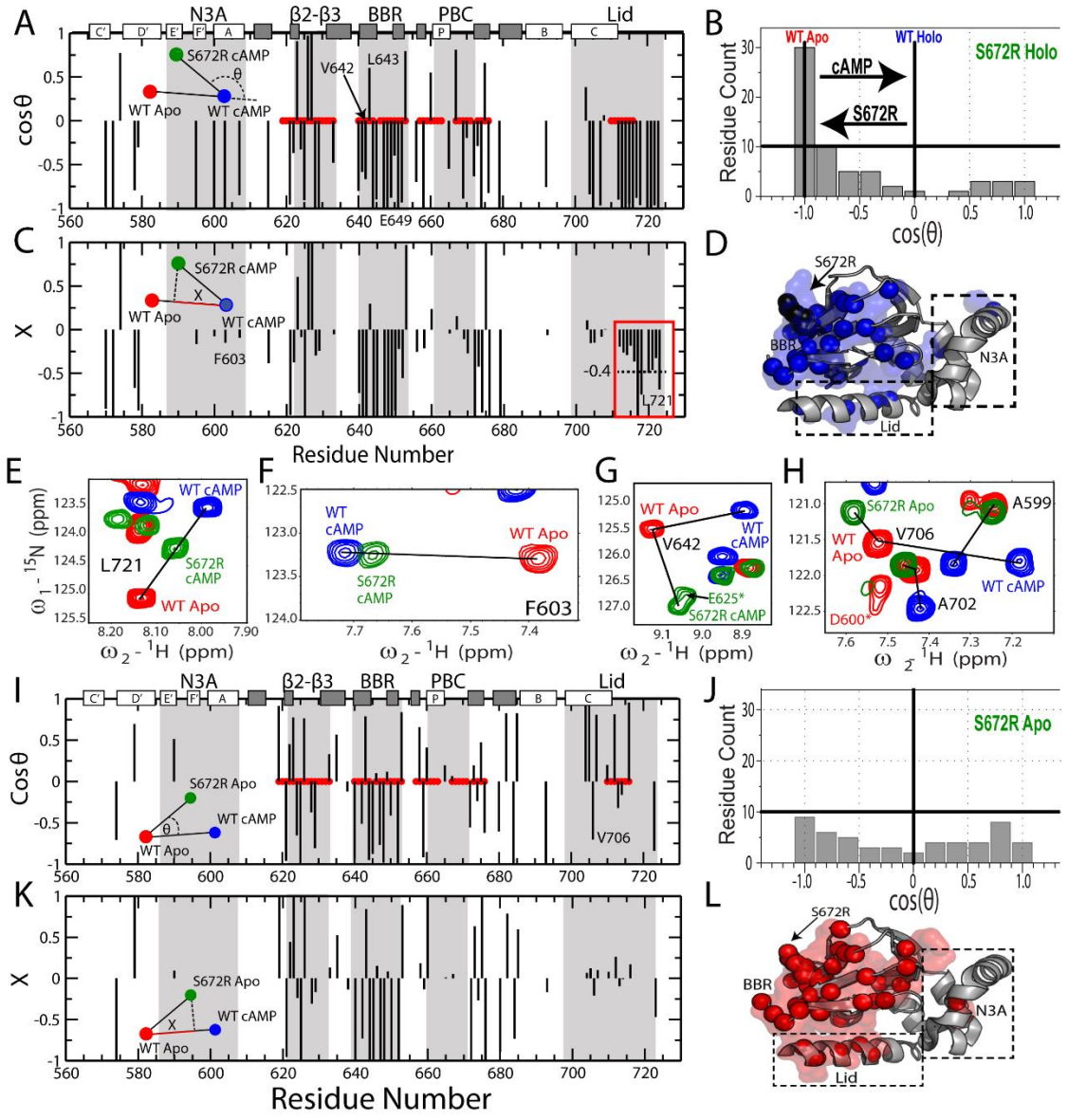

Figure 3: Chemical Shift Projection Analysis (CHESPA) of the HCN4 S672R Mutation. **A)** The cos(θ) values computed as in equation (3), as explained in the text. Selected residues reported in panels E-H are labeled. **B)** Distribution of cosθ values from panel A, revealing a net predominance of negative cos(θ) values. **C)** Fractional activation (X) values for the perturbation vector relative to the activation vector, computed as per equation (4). The red box highlights a subset of lid residues with markedly negative X values, pointing to a significant shift toward the inactive state. Selected residues reported in panels E-H are labeled. **D)** 3D map of CHESPA outliers, defined as residues with either $|\cos\theta| < 0.9$ or $|X| > 1.1$. Such outliers reveal deviations from the two-state equilibrium, such as those arising from nearest neighbour effects or population of a third state. **E)** Representative NH-HSQC cross-peaks for a residue from the C-terminal lid region with a holo S672R X value close to -0.4 (panel C red box). **F)** Similarly to panel (E), but for a representative residue with a holo S672R X value close to -0.1. **G)** A representative residue (V642) from the S672R holo spectra that exhibited a non-linear chemical shift change relative to the WT apo-holo states. **H)** Representative residues from the S672R apo spectra compared with WT apo and holo states that displayed non-linear chemical shift

changes. This non-linearity is also evident from the CHESPA analysis of the apo S672R HCN4 CBD, shown in panels **(I)** and **(J)**, revealing a nearly uniform cos(θ) distribution **(I)** between the -1 and 1 extremes. Selected residues reported in panels E-H are labeled in panel I. **(K)** X values for apo S672R residues C. **(L)** Similar to panel (D) but for the apo S672R CHESPA outliers.

Significant changes in ps-ns and μs-ms dynamics, identified as explained in the Materials and Methods section, were mapped on the crystal structure of S672R HCN4 (Fig 4K-N). The S^2 distributions shown in Fig 4A,B provide a global picture of the mutant-induced variations in fast (ps-ns) dynamics, while Fig 4C-J and Fig 4K-L offer a region and residue-resolution dissection of dynamics, respectively.

2.4.3 The Comparative S672R vs. WT Analysis of Internal Dynamics for the Apo CBD Reveals that the S672R Mutation Results in a Significant Overall Enhancement in the Amplitude of the ps-ns Internal Motions with Local Increases in ms-μs Dynamics Near cAMP Phosphate Recognition Sites.

The overall S^2 distributions observed in the absence of cAMP for the WT and S672R CBD reveal that the mutation results in a marked net shift towards lower S^2 values (Fig 4A). This result points to increased amplitudes of ps-ns motions in S672R vs. WT apo CBD. Furthermore, considering that order parameters for ps-ns motions are a proxy for conformational entropy(30), the S672R-induced shift towards reduced S^2 values suggests a widening of the well in the free-energy landscape minimum representing the apo CBD (Fig 5A). The S672R vs. WT dynamics enhancements of the apo CBD are not limited to the ps-ns time scale, but they include also μs-ms motions (Fig 4K). In particular, the β2-β3 loop proximal to the mutated site (Fig 1D) displays some of the greatest residue-specific enhancements of dynamics, as observed for both low and high frequency spectral density functions (Fig S4; Fig 4C,K). Specifically, residue E625, which forms a hydrogen bond with S672 (Fig 1D), exhibits a remarkable ~three-fold increase in J(0) (Fig S4A; Fig 4C), consistent with a large increase in μs-ms dynamics. In addition, several residues adjacent to E625 display enhanced ps-ns dynamics (Fig 4K; Fig S4C). These data suggest that by disrupting the interaction between S672 and E625, the S672R mutation destabilizes the entire β2-β3 loop. Since the β2-β3 loop lies adjacent to the PBC, we checked if this destabilization could extend to this cAMP-binding motif. Although the dynamic characterization of the apo PBC is sparse due to line-broadening, several residues at the edges of the PBC do indeed exhibit enhanced μs-ms dynamics (Fig 4G; Fig S4A), confirming that the β2-β3 enhancement of dynamics does at least partially extend to the PBC, which is a primary cAMP recognition site. So next we checked if the

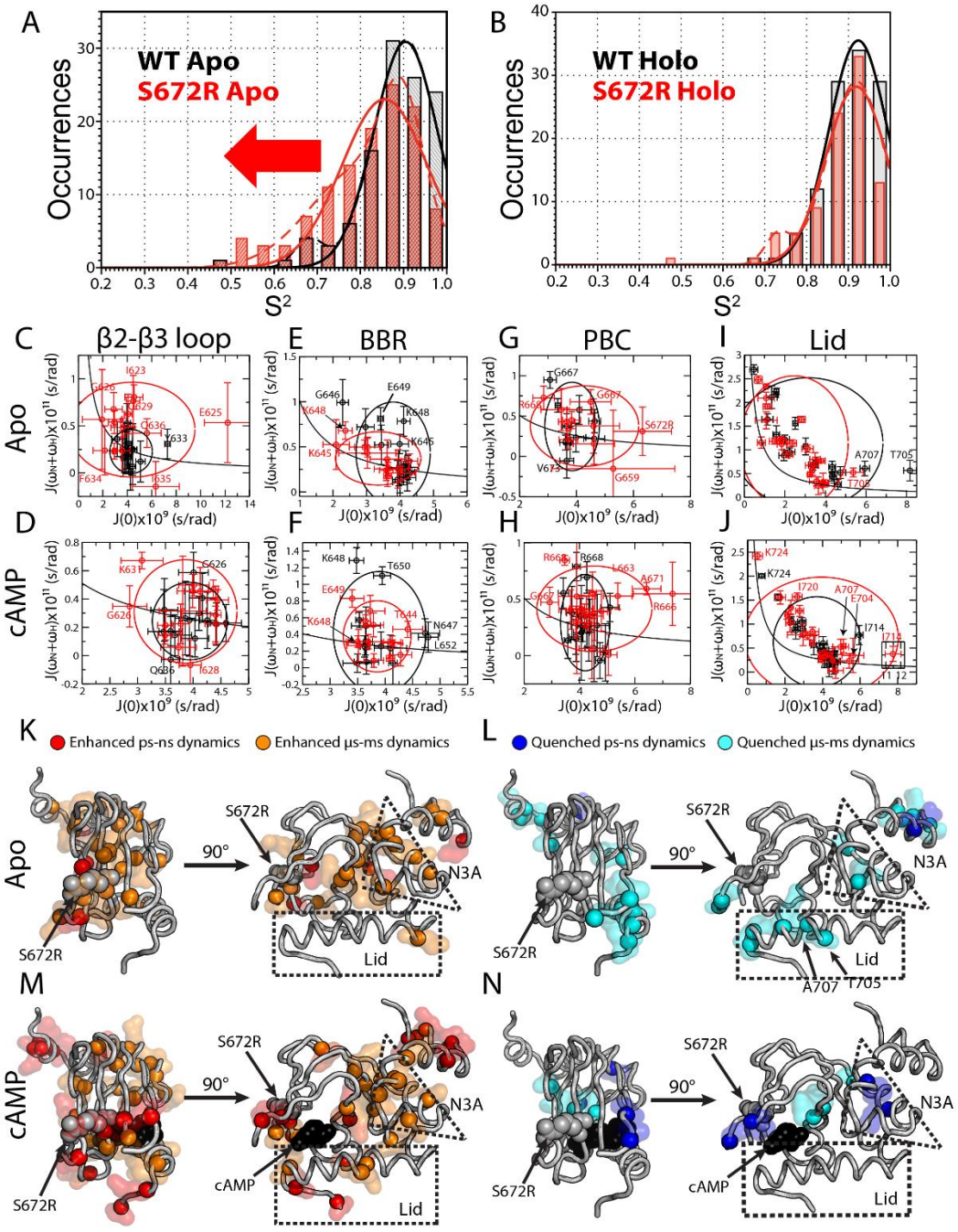

Figure 4: Comparative S672R vs. WT Analysis of HCN4 CBD Dynamics. Order parameter (S^2) distribution for WT (black) and S672R (red) in the absence **(A)** or presence of saturating amounts of cAMP **(B)**. Data was fit with either one (solid line) or two (dashed line) Gaussian distributions and as summarized in Table S1. **C-J)** Two-dimensional reduced spectral densities plots for the apo WT (black) and S672R (red) HCN4 CBD. The following four regions were probed: the β2-β3 loop (a.a. 621-636), BBR (a.a. 639-652), extended PBC (a.a. 659-676) and Lid region (a.a. 699-724), for both apo (top row) and cAMP-bound (bottom row) samples. The Lid is further dissected into the Lid and CTL in Fig S6. The solid black line represents the expected spectral densities for an isotropic rigid rotor, while the black and red ellipsoids are centered on

the average spectral density values with axes defined by +/- two standard deviations for the WT and mutant, respectively. **K)** Map of apo residues with S672R *vs.* WT enhanced ps-ns (red spheres and surface) and μs-ms (orange spheres and surface) dynamics. **L)** Map of apo residues with S672R *vs.* WT quenched ps-ns (blue sphere and surface) and μs-ms (cyan sphere and surface) dynamics. **M,N)** as in panels (K) and (L) except for the cAMP-bound sample.

S672R causes significant changes in dynamics at the other main cAMP contact sites, i.e. the BBR and the lid.

As to the BBR, while the 645-650 region in the WT apo CBD displays higher $J(\omega_N+\omega_H)$ values than adjacent residues, pointing to flexibility in the ps-ns time scale, no further significant enhancements of dynamics could be detected for S672R apo under our experimental conditions (Fig S4C; Fig 4E). While BBR residues K645 and K648 did exhibit slight decreases in J(0) (Fig 4E; Fig S4A), K648 displayed quenched dynamics at the ω_N frequency (Fig S4B), and for either residues no significant changes at the $J(\omega_N+\omega_H)$ spectral density were detected. On the other hand, the apo C-helix exhibited quenching of μs-ms dynamics as seen by decreased J(0) values for T705 and A707 (Fig 4L; Fig S4A; Fig S6A). Meanwhile, our probes for the BBR and lid do not support a major net S672R *vs.* WT enhancement in ps-ns dynamics in these regions (Fig 4E,I). However, it should be considered that these regions are already flexible in the ps-ns time-scale in the WT apo CBD, as shown by significantly increased $J(\omega_N+\omega_H)$ value relative to those expected for overall tumbling (Fig S4C). Beyond the cAMP binding sites, significant dynamic changes were found at the interface between α- and β- subdomains and even in parts of the N-terminal C-linker (Fig 4K,L; Fig S4). The net effect of these pervasive variations in dynamics in the apo CBD is that the S672R mutation results in an enhancement in the amplitude of fast (ps-ns) motions (Fig 4A) and in the associated entropy of the conformational ensemble sampled by the apo CBD (Fig. 5A). In order to evaluate the effect of the S672R mutation on the holo CBD dynamics, the comparative S672R vs. WT ^{15}N-relaxation analyses were extended to the data acquired in the presence of excess cAMP (Fig 4B,D-J,M,N; Fig S5 and S6).

2.4.4 In the Holo CBD, the S672R Mutation Results in Localized Enhancements in the Dynamics of Key cAMP Recognition and Gating Regions (PBC and Lid).

Unlike the apo CBD (Fig 4A), in the case of the holo CBD, no net S672R vs. WT shift to larger amplitude ps-ns motions is observed in the overall distribution of S^2 order parameters (Fig 4B). However, several localized S672R vs. WT enhancements of dynamics were detected also for the holo CBD (Fig 4H,J,M,N; Fig S5 and S6). The most dramatic increase in dynamics was found in two loci within the cAMP-binding pocket, i.e. PBC and lid, but significant enhancements of

dynamics were also identified in selected sites within the N3A and C-linker (Fig 4H,J,M; Fig S5). The PBC, which is nearly completely assigned in the cAMP-bound state and better sampled than in the apo state, experiences large enhancements of μs-ms dynamics, as indicated by several significant increases in the J(0) (Fig 4H; Fig S5A). Unlike the apo CBD, the mutation induced increase in dynamics for the PBC did not extent to the adjacent β2-β3 loop (Fig 4D). Similarly to the apo CBD, for the BBR, i.e. the other cAMP contact region in the β-subdomain, only a few significant changes in dynamics were recorded and these exhibited a mixed pattern, including both enhancements and quenches (Fig 4F; Fig S5).

Besides the PBC, the other region of the holo CBD experiencing major dynamic enhancements is the C-helix and the adjacent C-terminal lid segment (Fig 4J,M,N; Fig S5 and S6). The S672R mutation caused a significant enhancement of μs-ms dynamics in the C-helix, as seen by increased J(0) values for E704, T705, A707 and L708 (Fig S6B). In addition, I714, which forms the capping interactions with the adenine base of cAMP, displayed a remarkable two-fold increase in the corresponding J(0) value (Fig S5A, Fig S6D; Fig 4J,M). Furthermore, several residues in the C-terminal lid (CTL) exhibited significant enhancements in ns dynamics as seen by $J(\omega_N)$ (Fig S5B) and ps-ns dynamics as revealed by increased $J(\omega_N+ \omega_H)$ values (Fig 4M; Fig S5C).

Interestingly, for the S672R holo sample the ps-ns dynamics in the lid region, as probed by the average $J(\omega_N+ \omega_H)$ spectral densities for the C-helix and CTL lies in between those of the WT apo and WT holo samples (Tables S2 and S3). This result is in agreement with the fractional activations observed by CHESPA for this region, which also shows a partial reversion towards the apo-inactive state (Fig 3C). Together, the ^{15}N-relaxation data and the CHESPA comparative analyses for the holo CBD consistently point to a partial mutation-induced disengagement of the CTL, which is part of a key gating element for cAMP. Furthermore, the mutation also enhances the dynamics of key cAMP recognition elements, such as the PBC (Fig 4G,H), suggesting that S672R may lead also to altered on-off cAMP exchange kinetics. This hypothesis is also supported by the overall S672R *vs.* WT net shift toward lower order parameters, *i.e.* higher conformational entropy, observed for the apo CBD (Fig 4A). The wider free energy well of the mutant is in agreement with reduced free energy barriers for the apo-holo exchange (Fig 5A), corroborating the hypothesized S672R *vs.* WT acceleration of cAMP binding kinetics. In order to further test this hypothesis, we analyzed the cAMP binding dynamics utilizing protein-based and ligand-based NMR experiments.

2.4.5 The S672R Mutation Accelerates the On/Off Rates of the HCN4 CBD:cAMP Complex.

The cAMP-binding kinetics were probed by comparative analyses of chemical exchange regimes in HSQC spectra (Fig 5B-F) and ROESY build up curves (Fig 5G,H). The HSQC analysis involved monitoring residue-specific transitions from slow to intermediate and from intermediate to fast for apo-holo exchange regimes in HSQC spectra acquired at sub-stoichiometric ligand:protein ratios, as illustrated in Fig 5B. At a given protein:ligand stoichiometric ratio, the effective apo-holo exchange constant, k_{ex}, is defined by:

$$k_{ex} = [\text{cAMP}]_{free} * k_{on} + k_{off} \qquad (1)$$

A particular value of k_{ex} results in different exchange regimes for different residues, depending on the respective differences in Hz between the apo and holo resonance frequencies ($\Delta\nu_{A\text{-}H}$). If $k_{ex} << \Delta\nu_{A\text{-}H}$ for a selected residue, that amino acid will appear in the slow exchange regime, which is characterized by two peaks appearing at the positions of the apo and holo states with intensities proportional to the populations of ligand-bound and unbound protein (Fig 5B, blue peaks). However, under the same experimental conditions, and therefore with the same k_{ex}, for other residues that experience lower $\Delta\nu_{A\text{-}H}$ values, it is possible that $k_{ex} \approx \Delta\nu_{A\text{-}H}$, *i.e.* the residues fall within the intermediate exchange regime in which peaks broaden significantly, often beyond detection (Fig 5B, black peak). For amino acid with lower $\Delta\nu_{A\text{-}H}$ so that $k_{ex} >> \Delta\nu_{A\text{-}H}$, the fast exchange regime applies and only a single peak is observed whose chemical shift is a population weighted average of the apo and holo chemical shifts (Fig 5B, red peak). The difference in residue-specific exchange regimes provides an opportunity for a simple estimation of k_{ex} by assessing the range of $\Delta\nu_{A\text{-}H}$ values that result in intermediate chemical exchange. For this purpose, the $\Delta\nu_{A\text{-}H}$ values were sorted in decreasing order and the chemical exchange regime was determined for each residue to define the boundaries between slow, intermediate and fast exchange (Fig 5D and F). The k_{ex} was then assessed based on the range of $\Delta\nu_{A\text{-}H}$ values that resulted in intermediate chemical exchange (Fig 5D,F). Overall, Fig 5D and 5F illustrate how the ranked $\Delta\nu_{A\text{-}H}$ frequency differences define a 'built-in' scale for gauging k_{ex}. This approach is not as quantitative as other methods (*e.g.* N_z-exchange spectroscopy(26, 42)), but it is applicable to both HCN4 WT and mutant, unlike N_z-exchange experiments, as explained in Supp. Text.

The HSQC-based k_{ex} assessment was performed for both the WT and S672R HCN4 CBD in the presence of sub-stoichiometric amounts of cAMP (Fig 5C-F). Evidence that k_{ex} increases in going from S672R to WT is already provided by residues that experience similar apo *vs.* holo frequency differences, but fall in the slow exchange regime in WT and in the fast exchange regime

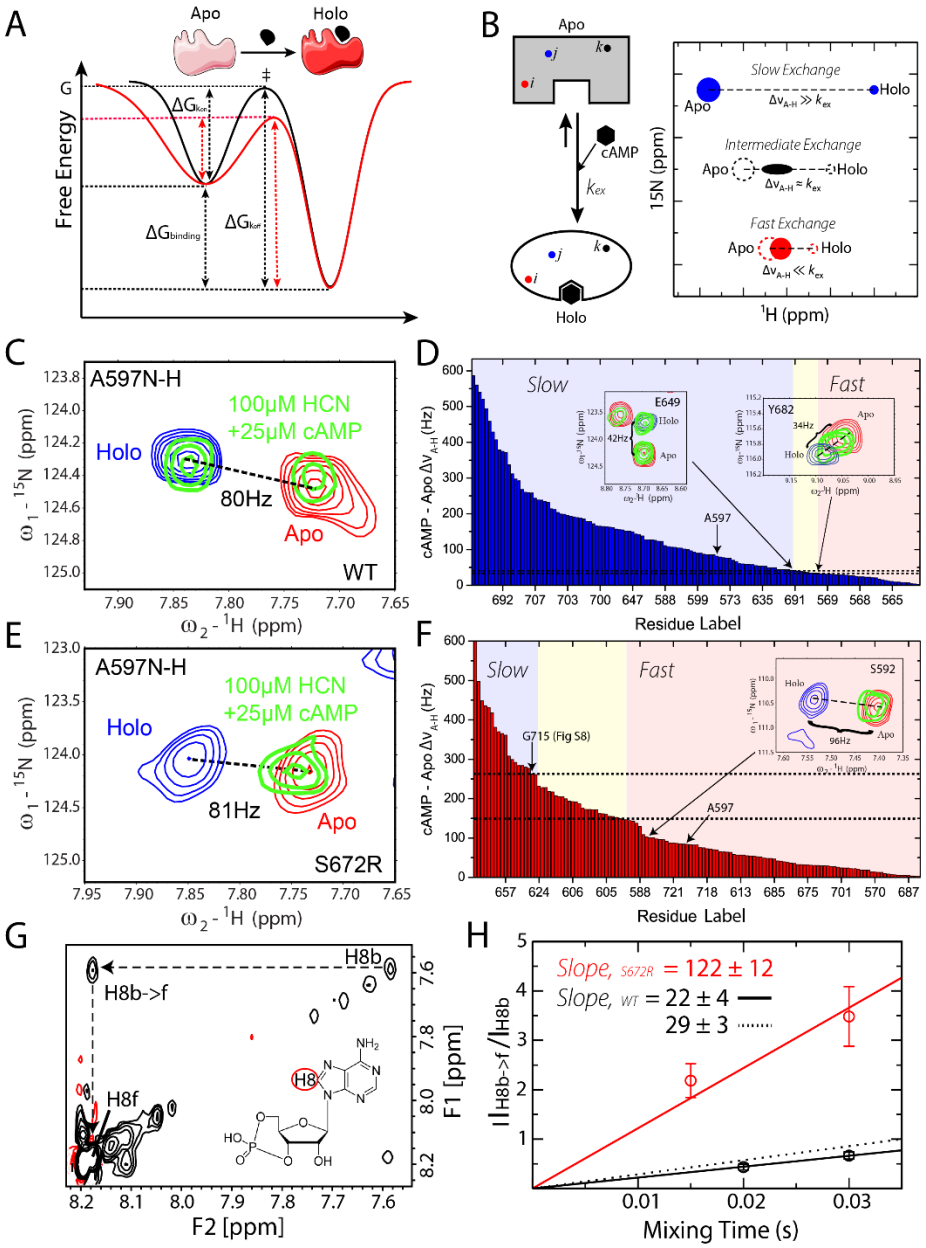

Figure 5: S672R vs. WT cAMP-Binding Kinetics from Protein- and Ligand-Based NMR Experiments. **A)** Schematic diagram depicting the effects of the reduced S672R *vs.* WT order parameters (Fig 4A) and higher conformational entropy on the free energy landscape of binding. **B)** Schematic illustration of the rationale for measuring apo-holo exchange rates through HSQC spectra at sub-stoichiometric ligand concentrations. Further details are available in the text. **C,E)** Representative cross-peak from the NH-HSQC spectra of the WT (C) and S672R HCN4 CBD (E). The red and blue peaks represent the apo and fully bound states, respectively, while the green peaks are from a sample containing 100 μM HCN4 CBD and 25 μM cAMP. For both WT and mutant CBDs, the $\Delta\nu_{A\text{-}H}$ is similar (80 *vs.* 81 Hz) but the exchange regime changes from slow for the WT (C) to fast for S672R (E). **D,F)** Sorted plots of $\Delta\nu_{A\text{-}H}$ *vs.* residue label arranged from largest to smallest for WT (D) and S672R (F). Indicated within the plots with colours are the chemical exchange regimes: blue – slow, yellow – intermediate and red – fast. Insets show representative residues

close to the boundaries between slow-intermediate and fast-intermediate exchange regimes. G715 spectra are shown in Fig S8. The dashed lines highlight the $\Delta\nu_{A-H}$ range corresponding to the intermediate exchange regime. **G)** ROESY spectra of 50 mM WT HCN4 CBD in the presence of excess cAMP (500 mM). The diagonal peaks indicated in the spectra correspond to the H8 proton of cAMP in its bound (H8b) and free (H8f) states. The cross-peak for the chemical exchange from bound to free H8 is also indicated as H8b->f. Black contours indicate positive peaks; red contours indicate negative peaks. The H8b->f preserves the same sign as the diagonal, as expected for chemical exchange cross-peaks. **H)** Plot of the H8b->f *vs.* H8b intensity ratios vs. ROESY mixing times. The initial slopes of the build-ups based on the first order approximation are reported in the figure. For the S672R mutant the bound ligand diagonal peak was detectable only up to a ROESY mixing time of 30 ms due to rapid decay. In contrast, the WT bound H8 signal was detectable up to a ROESY mixing time of 50ms. The solid black line depicts the slope of the two WT points shown in the plot, while the dashed line depicts the slope that includes two additional points at mixing times of 40 and 50 ms.

in S672R. This is the case, for example, for A597 as shown in Fig 5C and 5E. In WT HCN4, two separate peaks are detected for A597 (Fig 5C), consistent with slow apo-holo exchange, while in S672R a single peak is observed (Fig 5E), consistent with fast apo-holo exchange. Considering that the $\Delta\nu_{A-H}$ values of WT and S672R are comparable (Fig 5C and 5E), the slow *vs.* fast exchange regimes observed for A597 in WT *vs.* S672R, suggests that the mutation results in an increased k_{ex} value. The two- *vs.* one-peak patterns shown in Fig 5C and 5E are the basis for determining whether a given residues experiences slow or fast apo-holo exchange (see Supp. Text and Fig S8 for further details), so that the plots of the sorted $\Delta\nu_{A-H}$ values are partitioned into the three different kinetic regimes (Fig 5D and F). The comparison of Figures 5D and 5F reveals marked WT *vs.* S672R differences in the positions of the apo – holo exchange regimes within the scales defined by the sorted $\Delta\nu_{A-H}$ values (Fig 5D and 5F). For WT the slow to fast transition occurs in the 30-40 Hz range (Fig 5D), whereas for S672R the intermediate regimes falls in the 150-260 Hz window (Fig 5F), confirming that the mutation leads to a significantly enhanced k_{ex} value.

The S672R-induced k_{ex} enhancement revealed by the HSQC analyses (Fig 5B-F) reflects a primarily kinetic effect, as no significant S672R *vs.* WT differences were observed for the thermodynamic affinities of the HCN4 CBD, with K_d values in the 2-10 μM range for both mutant and WT (Fig 6). Specifically, based on equation (S3), the assessed k_{off} values for WT and S672R are expected to fall in the 23 – 31 s^{-1} and 113 – 202 s^{-1} ranges, respectively, pointing to a major S672R-induced acceleration in the cAMP on/off dynamics. This result was independently confirmed through ligand-based NMR experiments, *i.e.* ROESY spectra (Fig 5G and 5H), which complement the protein-based NMR data, *i.e.* the semi-quantitative HSQC analysis outlined above. The ROESY data were acquired in the presence of excess ligand and slow exchange was observed between the free and bound ligand peaks, as confirmed by ROESY cross-peaks with the same sign as the diagonal (Fig 5G). The initial slopes of the mixing time build up measured for the bound-to-

free exchange cross-peak fall in the 18-32 s^{-1} and 110-134 s^{-1} ranges for WT and S672R, respectively (Fig 5H). These values are consistent with the k_{off} ranges independently estimated based on the HSQC data and confirm that the S672R mutation leads to an at least ~three-fold enhancement in the on – off exchange rates of cAMP.

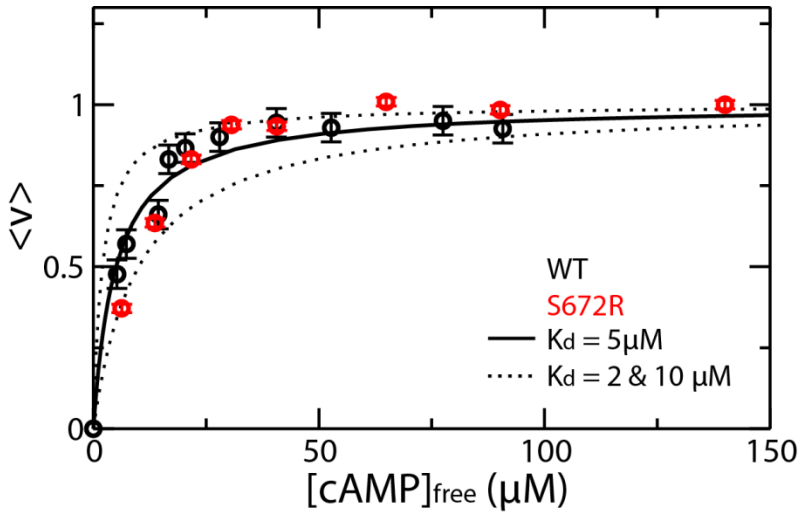

Figure 6: Binding isotherms of S672R (red) human HCN4 (563-724) with cAMP monitored by STD NMR. Data was fitted with a rectangular hyperbola for a 1:1 binding stoichiometry providing a K_d in the 2-10 μM range. The fractional saturation was measured by normalizing the STD amplification factor (9). The corresponding data for WT human HCN4 (563-724) (9) is shown (black dots) for the convenience of comparison. The low μM affinity reported here for the S672R variant is in agreement with the range of K_d values of 0.8 and 11.2 μM measured by fluorescence anisotropy and ITC by Xu *et al*. (20), respectively. In addition, considering that the apo WT state is already primarily sampling the inactive state, based on the scheme in Fig. 1F a mutation induced-shift towards inactivation is not expected to result in major variations in the observed affinity for cAMP.

2.5 Discussion

2.5.1 The Bradycardia-Related S672R Mutation Remodels the Free Energy Landscape (FEL) and the Dynamic Profiles Critical for the cAMP-Dependent Modulation of HCN4 Channels.

The comparative NMR analyses presented here (Fig 2-6; Fig S3-S8) have revealed how the S672R bradycardia-inducing mutation remodels the free energy landscape (FEL) for the cAMP-dependent regulation of the HCN4 channels, as illustrated in Fig 7A. The FEL remodelling caused by S672R (Fig 7A) is recapitulated by two main perturbations. First, the mutation results in a shift of the auto-inhibitory CBD equilibrium towards inactivation, as shown by the CHESPA data (Fig

3), consistent with the stabilization of inactive conformations and/or the destabilization of the active state (Fig 7A, vertical red arrows). Second, the S672R variant leads to a widening of the apo-inactive free energy basin (Fig 7A, horizontal red arrows), as indicated by a redistribution of the apo CBD order parameters for fast (ps-ns) motions towards lower values (Fig 4A). The broadening of the apo-inactive conformational sub-set is also confirmed by the CHESPA data (Fig 3), which indicate that the mutation-induced shift to inactive states cannot be fully captured by a simple two-state exchange model, i.e. rather than a single cooperative transition, S672R causes a partial de-correlation, which accounts for the observed residue dependence of the fractional activation values (X) and for the $|\cos(\theta)| < 1$ values observed at multiple sites (Fig 3). Considering that the apo CBD samples primarily inactive conformations, the broadening of the free energy well caused by S672R (Fig 7A) is expected to lower the free energy barrier for apo-holo exchange (Fig 5A). This prediction was confirmed by the observation that S672R accelerates the cAMP binding kinetics by a factor of at least three-fold, as measured by both protein and ligand-based NMR methods (Fig 5C-H).

The mutation-induced enhancement of dynamics is not limited only to the fast (ps-ns) time-scale, but extends also to the μs-ms dynamic window (Fig 4C-N). The S672R variant exhibits enhanced ms-μs dynamics at sites critical for cAMP recognition and binding, such as the PBC (Fig 7B). The PBC enhancement of ms-μs dynamics was detected in both apo and holo forms and extends to the adjacent β2-3 loop in the apo CBD, while for the holo CBD it affects the C-terminal lid, referred to as CTL (Fig 7B). The CTL is part of the lid region critical for the cAMP-dependent HCN channel activation(14) and for gating cAMP binding in WT (Fig 1B and 1C). However, in the holo S672R the CTL becomes partially disengaged, as consistently indicated by both the negative CHESPA fractional activation values (Fig 3B, C, E) and the concurrent mutation-induced enhancement of both fast and slow dynamics (Fig 4; Fig S6D; Table S3). The enhanced lid dynamics in S672R vs. WT is also in full agreement with the faster binding kinetics observed for the mutated variant. Overall, the picture emerging from the NMR investigation presented here suggests that the S672R-induced perturbations extend well beyond what was previously anticipated based on X-ray crystallography of the cAMP-bound intracellular region(20). The crystal structure of S672R and WT holo HCN4 IR tetramers differ primarily at the base-binding region (BBR), while NMR chemical shift changes show that for both apo and holo forms S672R affects multiple sites within a ~15 Å radius, spanning not only the BBR but also several other loci in the β-subdomain and lid regions (Fig 2D-F).

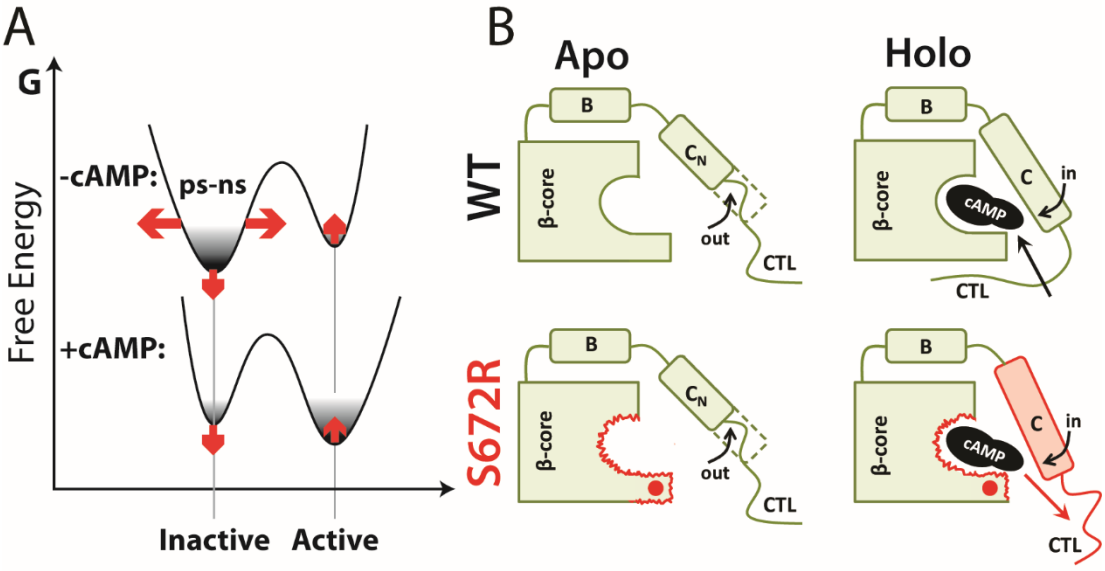

Figure 7: Proposed Mechanism of S672R-Induced Bradycardia Based on Free Energy Landscape Re-Modeling. **A)** Schematic free energy landscape for the allosteric cAMP-dependent modulation of the HCN4 CBD and its perturbation by the S672R mutation (red arrows). The horizontal red arrows denote the S672R *vs.* WT enhanced conformational entropy of the inactive state as indicated by the S672R *vs.* WT decreased S^2 parameters of the apo samples (Fig 4A). The vertical red arrows illustrate that the S672R mutation results in a CBD shift toward inactivation either by destabilizing the active state and/or by stabilizing the inactive state. **B)** Summary of critical S672R *vs.* WT changes in dynamics. Only the β-core and the C-terminal α-subdomain (B and C helices and the adjacent CTL) are shown for simplicity. Red wavy lines indicate S672R *vs.* WT enhanced dynamics in the HCN4 CBD. According to the proposed model, the enhanced dynamics at sites lining the cAMP binding pocket as well as the disengagement of the CTL in the cAMP-bound sample lead to an acceleration of the kinetics of cAMP binding and unbinding. The red dot indicates the site of the S672R mutation. The red outline and red shading denote areas with enhanced S672R *vs.* WT dynamics, leading to an increased k_{off} rate for cAMP (red arrow).

2.5.2 The S672R-Induced FEL Remodelling Provides an Initial Molecular Explanation for the Electrophysiological Changes Underlying Familial Bradycardia.

The S672R-induced remodelling of the free energy landscape for the cAMP-dependent modulation of HCN (Fig 7) provides a molecular framework to rationalize the two main electrophysiological signatures of this bradycardia-related mutation, i.e. the negative shift in the activation voltage and the faster deactivation. The former phenotype may arise as a result of a reduction in cAMP affinity (20) and/or of a constitutive S672R-induced variation in the position of the auto-inhibitory CBD equilibrium towards inactivation(18, 17). Our data support the notion that the second mechanism significantly contributes to the negative S672R-induced shift in the activation voltage observed by electrophysiology. Considering that activation voltages of full length HCN4 have been shown to correlate with the position of the auto-inhibitory CBD equilibrium as

probed by CHESPA(9), the partial inactivation observed for the S672R variant (vertical red arrows in Fig 7A) is consistent with the negative S672R vs. WT change of the activation voltage observed by electrophysiology for the full-length channel(18, 20). Furthermore, the position of the auto-inhibitory CBD equilibrium in reconstituted apo full-length channels may be affected by the presence of dimers known to be populated in the absence of cAMP(43, 44). If dimerization causes a partial shift of the WT apo CBD towards activation, the S672R mutation may result in an appreciably increased sampling of inactive states also in the absence of cAMP, explaining the constitutive negative shifts in activation voltages observed for apo channels(18).

The other putative mechanism, i.e. a reduction of cAMP affinity, does not appear to be sufficient alone to explain the observed negative S672R-induced shift in the activation voltage without any constitutive mutation-induced partial inactivation. This conclusion is supported by several independent observations. First, S672R is known to cause a negative shift in activation voltages even in the absence of cAMP(18). Second, we did not observe any major S672R vs. WT increases in the cAMP K_d measured under our experimental conditions by STD for the monomeric HCN4 CBD, as both K_d values fall in the 2-10 μM range (Fig 6). Third, no significant S672R-induced enhancements in the EC50 values were reported for the full-length HCN4 channels either(18). Fourth, although a S672R vs. WT $K_{1/2}$ increase was reported for either HCN2 or hybrid HCN2/4 channels(20), a significant mutation-induced reduction in integral HCN4 affinity would decrease the sensitivity to physiological levels of cAMP and result in chronotropic incompetence, which however was not observed in patients affected by the S672R mutation(18, 17). These observations consistently suggest that the negative S672R-induced shift in the activation voltage cannot be fully accounted for by a reduction of cAMP affinity alone without any constitutive S672R-induced variation in the position of the auto-inhibitory CBD equilibrium.

The other electrophysiological signature of the bradycardia-related mutation, i.e. the faster deactivation, is explained at least in part by the S672R-induced acceleration of the cAMP off kinetics, as seen by our combined analyses of protein- and ligand-based NMR experiments (Fig. 5). This result is in agreement with the S594R vs. WT k_{off} enhancement reported for the unbinding of a fluorescent cAMP analog from full-length mHCN2 channels and is rationalized by both the S672R-induced widening of the apo free energy well (Fig 5A) and the enhancement of dynamics observed in the holo CBD for the PBC and the lid, two critical elements of the cAMP binding sites (Fig 7). Unlike the PBC and the lid, the remaining cAMP recognition region, i.e. the BBR, does not appear to be subject to significant mutation-induced enhancements of dynamics (Fig 4).

In conclusion, we have shown how the bradycardia-related S672R mutation remodels the free energy landscape (FEL) for the cAMP-dependent modulation of HCN4 channel gating. The S672R-induced FEL remodeling is recapitulated by two main effects. First, S672R results in the constitutive stabilization of inactive vs. active states, contributing to the negative S672R vs. WT shift in the activation voltage observed by electrophysiology for HCN4 channels. Second, S672R leads to enhanced dynamics at sites pivotal for cAMP binding, such as the PBC and the lid, which in turn leads to an acceleration of the cAMP unbinding kinetics, explaining at least in part the faster deactivation reported for S672R vs. WT HCN4 channels. Hence, our results provide a mechanistic explanation of the main electrophysiological determinants of the familial bradycardia diagnosed in patients carrying the S672R mutation. In addition, the NMR approaches outlined here will be useful to map how other disease related mutations remodel functionally relevant free energy landscapes in other systems as well (45–48).

2.6 Experimental Procedures

2.6.1 HCN4 Sample Preparation.

The expression and purification of HCN4 (563-724), referred to here as HCN4 CBD for simplicity, was adapted from a previously established protocol (9). A rare codon optimized SUMO-HCN4 fusion construct was sub-cloned into a pET302NT-His vector (Invitrogen) and expressed in *Escherichia coli* BL-21(DE3) cells in M9 minimal media isotopically enriched with ^{15}N ammonium chloride (Sigma Aldrich). Expression was induced with 0.5mM Isopropyl β-D-1-thiogalactopyranoside (IPTG, Fisher Scientific) at an OD_{600} of 0.6 and incubated for a further 16-18 hours at 20°C. Cells were lysed with a cell disruptor and the cell lysate was centrifuged at 20,000xg for 1 hour. The cell lysate was then filtered with 0.8/0.2μm syringe filters (VWR) and purified by Ni-NTA affinity chromatography following the same protocol as in Akimoto et al. (9). Following proteolytic cleavage of the His_6-SUMO tag with His_6-tagged TEV protease the sample was purified again with Ni-NTA affinity chromatography to remove both the SUMO tag and the TEV protease. The cleaved protein was then diluted three fold with Sp Wash buffer (10 mM MES, 10 mM KCl and 10 mM 2-mercaptoethanol, pH 6.0) and further purified by cation exchange chromatography with a 5 mL HiTrap Sp column (GE Healthcare Life Sciences). Following injection the protein was eluted using a 100% gradient over 100 minutes with Sp elution buffer (10 mM MES, 1 M KCl, and 10 mM 2-mercaptoethanol, pH 6.0). The purified protein was then

dialyzed overnight in NMR buffer (20 mM MES, 75 mM KCl, 2 mM EDTA, 2 mM EGTA, 1 mM DTT and 0.02% sodium azide, pH 6.5) and stored at 4 °C. HCN4 (563-724) S672R was created by site-directed mutagenesis using a variation of the Quikchange protocol. The reaction used the KOD Hot Start master mix (Novagen) with primers synthesized in-house. The reaction protocol followed the manufacturer's specifications for the KOD Hot Start kit. Following the reaction, parental DNA was degraded with DpnI for 1 hour at 37 °C and the reaction product was transformed in *E. coli* top10 competent cells. Plasmids were later isolated from single colonies grown in liquid LB media using a GeneJET Plasmid Miniprep kit (Thermo Scientific) and sequenced to confirm the mutation. HCN4 (563-724) S672R was expressed and purified similarly to the WT with the exception that the cation exchange gradient was changed to 100% Sp elution buffer over 25 minutes.

2.6.2 NMR Spectroscopy.

NMR experiments were acquired on a Bruker Avance 700MHz spectrometer equipped with a 5mm TCI cryo-probe, unless otherwise specified. All experiments were acquired at 300 K in the NMR buffer defined above with 5 % D_2O unless otherwise specified. Spectra were processed either with NMRpipe (49) or directly in Bruker's Topspin and analyzed in NMRFAM-SPARKY (50). Chemical shift assignments of S672R were obtained through spectral comparison with WT spectra (9), if no ambiguities were present, and through triple resonance experiments (i.e. HNCO, HNCA and HN(CO)CACB) otherwise. The TALOS+ server was used to predict the protein's secondary structures using HN, $C\alpha$, $C\beta$ and C' chemical shifts and to cross check assignments based on acceptable backbone torsional angles (51).

2.6.3 Chemical Shift Analyses.

Compounded chemical shift differences (ΔCCS) were computed as described previously (41, 9, 40, 27, 29), using the following equation:

$$\Delta CCS = \sqrt{((\delta_{H,1} - \delta_{H,2})^2 + (0.2*(\delta_{N,1} - \delta_{N,2}))^2)} \qquad (2)$$

Where the δ symbols represent the proton (H) and nitrogen (N) chemical shifts (in ppm) for states 1 and 2. The CHESPA protocol has been described elsewhere (29, 23, 9). Briefly, the WT apo and WT holo cross-peaks for a given residue define a reference vector A, while the S672R (either apo or holo) and its corresponding WT state define a perturbation vector B. The cosθ and fractional activation (X) values are then computed for residues with mutant vs. WT ΔCCS > 0.05 ppm, using the following equations:

$$cos\,\theta = \frac{\vec{A} \bullet \vec{B}}{|\vec{B}||\vec{A}|} \qquad (3)$$

$$X = \frac{\vec{A} \bullet \vec{B}}{|\vec{B}|^2} \qquad (4)$$

2.6.4 NMR Spin Relaxation Measurements and Reduced Spectral Density Mapping.

The longitudinal and transverse relaxation rates, as well as the {^{1}H-^{15}N} steady state NOEs were measured for 150 μM WT and S672R HCN4 (563-724) samples in their apo and cAMP-bound (2 mM cAMP) forms. Pseudo-3D versions of the inversion recovery and CPMG pulse sequences modified to include a water flip-back pulse and sensitivity enhancement were used to measure T_1 and T_2 rates, respectively (52, 53). Both experiments included 1024 complex points with a spectral width of 14.28 ppm in the t_2 dimension and 256 complex points with a spectral width of 31.82 ppm in the t_1 dimension. The T_1 experiment was acquired with 16 scans, 128 dummy scans, a recycle delay of 2.5s and the following relaxation delays: 100 (x2), 200, 300, 400 (x2), 500, 600, 800 and 1000 (x2) ms (where x2 indicates duplicate spectra). The T_2 experiment was acquired with 32 scans, 128 dummy scans, a recycle delay of 1.8s and the following CPMG relaxation delays: 8.42, 16.96 (x2), 25.44, 33.92, 42.40 (x2), 50.88, 59.36, 76.32 and 93.28 (x2) ms. The HN NOE data for WT had been previously acquired (9) and for the S672R mutant the HN NOE were measured similarly to WT. In brief, the HN-NOE experiment used a 10s recycle delay with a 5s proton saturation period. NOE experiments were acquired in 10-12 sets of saturated and unsaturated spectra, which were co-added and processed similarly to the HSQC spectra noted above. The HN-NOEs were then computed from the intensity ratios of saturated/non-saturated peaks.

Error bars were determined from replicate measurements. T_1 and T_2 experiments were processed in Topspin using cosine squared window functions and residual water deconvolution. Spectra were then transferred to Sparky and fitted to an exponential decay to extract T_1 and T_2 rates. Reduced spectral densities were computed as described previously (24, 25). Significant changes in dynamics were identified by computing the average spectral density values for all residues along with their standard deviations (SDs) and by excluding from subsequent averaging calculations any residues outside the average +/- one SD. This computation of progressively trimmed averages and standard deviations was completed three times, after which any further iteration resulted only in

marginal changes, i.e. ~1%. At this point, the average spectral density was considered to arise primarily from overall molecular tumbling as opposed to internal motions.

The assessment of spectral densities reflecting overall tumbling served as a reference to identify significant S672R vs. WT variations in internal dynamics based on the following criteria: i) the spectral density of a residue for either WT or S672R has to be outside the average +/- five SDs; ii) the S672R vs. WT spectral density change for a given residue must exceed five times the SD obtained from the overall tumbling assessment or, if greater, the sum of the respective experimental errors. Residues that met such criteria were deemed subject to significant S672R vs. WT changes in internal dynamics and mapped onto the crystal structure of S672R HCN4. Residues undergoing S672R vs. WT enhancement (suppression) in ps-ns backbone dynamics were identified through increases (decreases) in $J(\omega_H+\omega_N)$, while residues with S672R vs. WT enhancement (suppression) in ms-μs backbone dynamics were identified through increases (decreases) in J(0), typically with values above those computed for overall tumbling, in the absence of $J(\omega_H+\omega_N)$ changes, as well as through 850 vs, 700 MHz R_2 comparisons. The T_2 relaxation rates were also measured on a Bruker Avance 850 MHz spectrometer equipped with a TXI probe to further examine residues with significant contributions from R_{ex} arising from dynamics on the ms-μm time-scale. The pulse sequence employed for this analysis was a pseudo-3D version of the CPMG experiment with sensitivity enhancement and a water flip-back pulse. Spectra were acquired with 2048 complex points and a spectral width of 14 ppm in the direct dimension and 256 complex points and a spectral width of 36 ppm in the indirect dimension. Six CPMG relaxation delays were acquired in duplicate: 16.96, 33.92, 50.88, 67.84, 84.80 and 101.76 ms. The comparison of R_2 rates at 850 and 700 MHz are shown in Fig S7. The difference in R_2 rates at 850 and 700 MHz was then computed and significant changes in μs-ms dynamics were identified using an approach similar to the one described above for the reduced spectral density analysis. The S^2 order parameters were calculated from J(0) and $J(\omega_N + \omega_H)$ as previously described (54, 55). The S^2 distributions were modeled through non-linear fitting with both one and two Gaussian curves. Further details are available in Table S1.

2.6.5 Measurements of Binding Thermodynamics and Kinetics.

Binding affinities of cAMP for the WT and S672R HCN4 CBD were determined by measuring binding isotherms through the saturation transfer difference (STD) amplification factor, as previously explained (9). For the assessment of the binding kinetics through protein based

experiments N_z-experiments could not be implemented due to weak holo peaks at cAMP concentrations sufficiently low to ensure slow exchange (Supp. Text). Hence, a simpler and less quantitative approach was taken based on the analysis of gradient and sensitivity enhanced {^{15}N,^{1}H}-heteronuclear single quantum coherence (HSQC) spectra, as explained in the Results section. These HSQC datasets were acquired with 1024 complex points and a spectral width of 14.1 ppm in the directly detected ^{1}H dimension and 128 complex points and a spectral width of 31.8 ppm in the indirectly detected ^{15}N dimension. Spectra were processed employing a 60° shifted sine squared bell window function, zero filling, and linear prediction in the indirect dimension. Titrations were acquired with 100 μM total protein concentrations and total cAMP concentrations of 25, 50, 100, 200, 400, 600, 1000 and 1500 μM. The HSQC-based assessment of the cAMP-binding kinetics was complemented by acquiring 2D-ROESY spectra, recorded with mixing times in the 15-50 ms range in the presence of excess ligand. Spectra were acquired with spectral widths of 12 ppm in both dimensions and 2048 and 512 complex t_2 and t_1 points, respectively. The strength of the continuous wave spin-lock strength was set at 2.5 kHz. The cAMP off kinetics was probed through the initial slope of the plot of the ratio of the bound-to-free (b->f) ROE cross-peak intensity to the bound ligand diagonal peak intensity vs. the ROESY mixing times.

2.7 Acknowledgements:

We thank Professors Eric Accili (University of British Columbia), L. Zhou (Virginia Commonwealth University), W.N. Zagotta (U. Washington), Anna Moroni (University of Milan), D. Di Francesco (University of Milan) and Susan S. Taylor (University of California at San Diego) as well as B. VanSchouwen, J. Byun, R. Ahmed, N. Jafari, and M. Gloyd (McMaster) for helpful discussion.

2.8 References

1. Benarroch, E. E. (2013) HCN channels: function and clinical implications. *Neurology* **80**, 304–310

2. Craven, K. B., and Zagotta, W. N. (2006) CNG and HCN channels: two peas, one pod. *Ann. Rev. Physiol.* **68**, 375–401

3. Wicks, N. L., Wong, T., Sun, J., Madden, Z., and Young, E. C. (2011) Cytoplasmic cAMP-sensing domain of hyperpolarization-activated cation (HCN) channels uses two structurally distinct mechanisms to regulate voltage gating. *Proc. Nat. Acad. Sci.* **108**, 609–614

4. 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

5. 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

6. Ng, L. C., Putrenko, I., Baronas, V., van Petegem, F., and Accili, E. A. (2016) Cyclic Purine and Pyrimidine Nucleotides Bind to the HCN2 Ion Channel and Variably Promote C-Terminal Domain Interactions and Opening. *Structure (London, England : 1993)* **24**, 1629–1642

7. Lolicato, M., Nardini, M., Gazzarrini, S., Moller, 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

8. Xu, X., Vysotskaya, Z. V., Liu, Q., and Zhou, L. (2010) Structural Basis for the cAMP-dependent Gating in the Human HCN4 Channel. *J. Biol. Chem.* **285**, 37082–37091

9. 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

10. Seifert, R. (2016) Distinct Signaling Roles of cIMP, cCMP, and cUMP. *Structure* **24**, 1627–1628

11. Schweizer, P. A., Duhme, N., Thomas, D., Becker, R., Zehelein, J., Draguhn, A., Bruehl, C., Katus, H. A., and Koenen, M. (2010) cAMP Sensitivity of HCN Pacemaker Channels Determines Basal Heart Rate But Is Not Critical for Autonomic Rate Control. *Circulation: Arrhythmia and Electrophysiology* **3**, 542–552

12. 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 Meth* **6**, 532–537

13. VanSchouwen, B., Akimoto, M., Sayadi, M., Fogolari, F., and Melacini, G. (2015) Role of Dynamics in the Autoinhibition and Activation of the Hyperpolarization-activated Cyclic Nucleotide-modulated (HCN) Ion Channels. *J. Biol. Chem.* **290**, 17642–17654

14. Zhou, L., and Siegelbaum, S. A. (2007) Gating of HCN channels by cyclic nucleotides: residue contacts that underlie ligand binding, selectivity, and efficacy. *Structure (London, England : 1993)* **15**, 655–670

15. Verkerk, A. O., and Wilders, R. (2014) Pacemaker activity of the human sinoatrial node: Effects of HCN4 mutations on the hyperpolarization-activated current. *Europace* **16**, 384–395

16. Baruscotti, M., Bucchi, A., Milanesi, R., Paina, M., Barbuti, A., Gnecchi-Ruscone, T., Bianco, E., Vitali-Serdoz, L., Cappato, R., and DiFrancesco, D. (2015) A gain-of-function mutation in the cardiac pacemaker HCN4 channel increasing cAMP sensitivity is associated with familial Inappropriate Sinus Tachycardia. *Eur. Heart J.*

17. DiFrancesco, D. (2013) Funny channel gene mutations associated with arrhythmias. *J. Physiol.* **591**, 4117–4124

18. Milanesi, R., Baruscotti, M., Gnecchi-Ruscone, T., and DiFrancesco, D. (2006) Familial sinus bradycardia associated with a mutation in the cardiac pacemaker channel. *New Engl. J. Med.* **354**, 151–157

19. Schulze-Bahr, E., Neu, A., Friederich, P., Kaupp, U. B., Breithardt, G., Pongs, O., and Isbrandt, D. (2003) Pacemaker channel dysfunction in a patient with sinus node disease. *J. Clin. Invest.* **111**, 1537–1545

20. Xu, X., Marni, F., Wu, S., Su, Z., Musayev, F., Shrestha, S., Xie, C., Gao, W., Liu, Q., and Zhou, L. (2012) Local and Global Interpretations of a Disease-Causing Mutation near the Ligand Entry Path in Hyperpolarization-Activated cAMP-Gated Channel. *Structure* **20**, 2116–2123

21. 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). *J. Biol. Chem.* **290**, 28631–28641

22. 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. Nat. Acad. Sci.* **110**, 14231–14236

23. Boulton, S., Akimoto, M., VanSchouwen, B., Moleschi, K., Selvaratnam, R., Giri, 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. *Biochm. Soc. Trans.* **42**, 302–307

24. 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

25. 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

26. 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

27. Selvaratnam, R., Chowdhury, S., VanSchouwen, B., and Melacini, G. (2011) Mapping allostery through the covariance analysis of NMR chemical shifts. *Proc. Nat. Acad. Sci.* **108**, 6133–6138

28. 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

29. 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

30. Boulton, S., and Melacini, G. (2016) Advances in NMR Methods To Map Allosteric Sites: From Models to Translation. *Chem. Rev.* **116**, 6267–6304

31. Nussinov, R. (2016) Introduction to Protein Ensembles and Allostery. *Chem. Rev.* **116**, 6263–6266

32. Ma, B., and Nussinov, R. (2016) Protein dynamics: Conformational footprints. *Nature Chem. Biol.* **12**, 890–891

33. Guo, J., and Zhou, H.-X. (2016) Protein Allostery and Conformational Dynamics. *Chem. Rev.* **116**, 6503–6515

34. Guo, C., and Zhou, H. X. (2016) Unidirectional allostery in the regulatory subunit RIalpha facilitates efficient deactivation of protein kinase A. *Proc. Nat. Acad. Sci.* **113**, E6776-E6785

35. Guo, J., and Zhou, H.-X. (2015) Dynamically Driven Protein Allostery Exhibits Disparate Responses for Fast and Slow Motions. *Biophys. J.* **108**, 2771–2774

36. Greenwood, A. I., Rogals, M. J., De, S., Lu, K. P., Kovrigin, E. L., and Nicholson, L. K. (2011) Complete determination of the Pin1 catalytic domain thermodynamic cycle by NMR lineshape analysis. *J. Biomol. NMR* **51**, 21–34

37. Cembran, A., Kim, J., Gao, J., and Veglia, G. (2014) NMR mapping of protein conformational landscapes using coordinated behavior of chemical shifts upon ligand binding. *Phys. Chem. Chem. Phys.* **16**, 6508–6518

38. Omidvar, R., Xia, Y., Porcelli, F., Bohlmann, H., and Veglia, G. (2016) NMR structure and conformational dynamics of AtPDFL2.1, a defensin-like peptide from Arabidopsis thaliana. *Biochim. Biophys. Acta* **1864**, 1739–1747

39. Kleckner, I. R., and Foster, M. P. (2011) An introduction to NMR-based approaches for measuring protein dynamics. *Biochim. Biophys. Acta* **1814**, 942–968

40. Moleschi, K. J., Akimoto, M., and Melacini, G. (2015) Measurement of State-Specific Association Constants in Allosteric Sensors through Molecular Stapling and NMR. *J. Am. Chem. Soc.*, 150817143428000

41. 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

42. Farrow, N. A., Zhang, O., Forman-Kay, J. D., and Le Kay (1994) A heteronuclear correlation experiment for simultaneous determination of 15N longitudinal decay and chemical exchange rates of systems in slow equilibrium. *J. Biomol. NMR* **4**, 727–734

43. Ulens, C., and Siegelbaum, S. A. (2003) Regulation of hyperpolarization-activated HCN channels by cAMP through a gating switch in binding domain symmetry. *Neuron* **40**, 959–970

44. VanSchouwen, B., and Melacini, G. (2016) Structural Basis of Tonic Inhibition by Dimers of Dimers in Hyperpolarization-Activated Cyclic-Nucleotide-Modulated (HCN) Ion Channels. *J. Phys. Chem.*

45. Schuetz, A. K., and Kay, L. E. (2016) A Dynamic molecular basis for malfunction in disease mutants of p97/VCP. *eLife* **5**

46. Sekhar, A., Rumfeldt, J. A. O., Broom, H. R., Doyle, C. M., Sobering, R. E., Meiering, E. M., and Kay, L. E. (2016) Probing the free energy landscapes of ALS disease mutants of SOD1 by NMR spectroscopy. *Proc. Nat. Acad. Sci.*

47. 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. Nat. Acad. Sci.* **112**, 3716–3721

48. Nussinov, R., and Tsai, C.-J. (2013) Allostery in disease and in drug discovery. *Cell* **153**, 293–305

49. 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

50. Goddard, T. D., and Kneller, D. G. *Sparky3*, University of California, San Francisco

51. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223

52. Loh, A. P., Guo, W., Nicholson, L. K., and Oswald, R. E. (1999) Backbone dynamics of inactive, active, and effector-bound Cdc42Hs from measurements of (15)N relaxation parameters at multiple field strengths. *Biochemistry* **38**, 12547–12557

53. Wang, C., Pawley, N. H., and Nicholson, L. K. (2001) The role of backbone motions in ligand binding to the c-Src SH3 domain. *J. mol. biol.* **313**, 873–887

54. Bracken, C., Carr, P. A., Cavanagh, J., and Palmer, A. 3. (1999) Temperature dependence of intramolecular dynamics of the basic leucine zipper of GCN4: implications for the entropy of association with DNA. *J. mol. biol.* **285**, 2133–2146

55. Popovych, N., Sun, S., Ebright, R. H., and Kalodimos, C. G. (2006) Dynamically driven protein allostery. *Nat. Struct. Mol. Biol.* **13**, 831–838

56. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics* **14**, 51–55

2.9 Supplementary Information: Experimental Procedures

2.9.1 Fast *vs.* Slow Exchange Regime Assignment for Residues with Weak HSQC Intensities

For residues with weak HSQC peak intensities it is possible that at sub-stoichiometric ratios of cAMP the minor holo peak for slow exchanging residues may pass undetected, resulting in a single observable peak similarly to fast exchanging residues. To eliminate these ambiguities the assignment of fast *vs.* slow exchange regimes was corroborated through HSQC spectra acquired at a higher ligand:protein concentration ratio in order to increase the free ligand concentration (*i.e.* $[cAMP]_{free}$) and, based on equation (1), k_{ex}. The resultant increase in the k_{ex} value affects residues in the slow and fast exchange regimes differently. Slowly exchanging residues experience a shift toward the intermediate regime, causing line-broadening and a decrease in signal intensity (Fig S8A,B). In contrast, when a residue is in the fast exchange regime, an increase in the k_{ex} rate would not decrease its intensity (Fig S8C,D).

It is also notable that the faster cAMP on-off exchange observed for the S672R mutant relative to the WT limits the applicability of Nz-exchange experiments, which could in principle provide quantitative measurements of the on/off rates for cAMP. Furthermore, if higher cAMP concentrations were utilized to increase the holo fraction, the k_{ex} value would increase as well bringing most residues in the S672R mutant in the intermediate/fast exchange regime, where the Nz-exchange experiment does not apply. Due to these limitations, we could not utilized Nz-spectroscopy for measuring the cAMP binding kinetics and we reverted to the HSQC analysis outlined in the main text.

2.9.2 Calculation of On/Off Rates Starting from the k_{ex} And K_d Values.

On and off rates of cAMP binding (k_{on} & k_{off}) were computed from the k_{ex} and the K_d values as follows. Since:

$$K_d = k_{off}/k_{on} \qquad \text{(S1)}$$

Equation (1) for k_{ex} is re-arranged as:

$$k_{on} = k_{ex} / ([cAMP]_{free} + K_d) \qquad \text{(S2)}$$

$$k_{off} = k_{ex} / (([cAMP]_{free} / K_d) + 1) \qquad \text{(S3)}$$

The k_{ex} value is estimated based on the identification of the intermediate exchange regime within the scale defined by the sorted apo-holo Hz differences, as shown in Fig 5D-F. The K_d value was

measured through the STD amplification actors (Fig S2), while $[cAMP]_{free}$ is obtained based on the total protein and ligand concentrations as well as the K_d value.

2.10 Supplementary Tables

Table S1. Gaussian Fitting of S^2 Order Parameter Distributions[a]

		Apo		*Holo*	
	Parameters	*WT*	*S672R*	*WT*	*S672R*
Single Gaussian	A (Height)	30.9	23.0	35.5	28.2
	B (Position)	0.90	0.86	0.92	0.92
	C (Width)	0.07	0.10	0.07	0.07
Two Gaussians	A1 (Height)	31.1	15.7	35.5	29.1
	B1 (Position)	0.91	0.90	0.92	0.92
	C1 (Width)	0.07	0.05	0.07	0.07
	A2 (Height)	3.9	13.5	$< 10^{-8}$	4.7
	B2 (Position)	0.68	0.80	0.18	0.73
	C2 (Width)	0.04	0.12	0.09	0.03

[a] The functions subject to non-linear fitting were: $A*\exp(-(S^2-B)^2/2C^2)$ for a single Gaussian or $A1*\exp(-(S^2-B1)^2/2C1^2) + A2*\exp(-(S^2-B2)^2/2C2^2)$ for two Gaussians.

Table S2. Reduced Spectral Density Statistics for Apo WT and Apo S672R HCN4 (563-724) Samples

		$J(0) \times 10^9$ (s/rad)		$J(\omega_N) \times 10^{10}$ (s/rad)		$J(\omega_N + \omega_H) \times 10^{11}$ (s/rad)	
Region	**Residues**	**WT**	**S672R**	**WT**	**S672R**	**WT**	**S672R**
Total	563-724	3.82 ± 0.21	3.35 ± 0.45	1.69 ± 0.12	1.72 ± 0.11	0.56 ± 0.15	0.61 ± 0.13
β2-β3	621-636	4.32 ± 0.28	4.41 ± 0.94	1.56 ± 0.15	1.55 ± 0.19	0.23 ± 0.20	0.46 ± 0.23
BBR	639-652	3.79 ± 0.20	3.35 ± 0.48	1.62 ± 0.13	1.65 ± 0.12	0.43 ± 0.16	0.37 ± 0.13
PBC	659-676	3.85 ± 0.24	4.28 ± 0.75	1.65 ± 0.10	1.74 ± 0.20	0.37 ± 0.17	0.38 ± 0.26
C-Helix	699-713	4.58 ± 0.21	3.23 ± 0.25	1.76 ± 0.11	1.90 ± 0.08	0.64 ± 0.13	0.76 ± 0.09
CTL	714-724	1.72 ± 0.19	1.23 ± 0.24	2.06 ± 0.12	1.97 ± 0.08	1.69 ± 0.13	1.82 ± 0.09

Table S3. Reduced Spectral Density Statistics for Holo WT and Holo S672R HCN4 (563-724) Samples

		$J(0) \times 10^9$ (s/rad)		$J(\omega_N) \times 10^{10}$ (s/rad)		$J(\omega_N + \omega_H) \times 10^{11}$ (s/rad)	
Region	**Residues**	**WT**	**S672R**	**WT**	**S672R**	**WT**	**S672R**
Total	563-724	4.04 ± 0.20	4.13 ± 0.29	1.64 ± 0.12	1.70 ± 0.13	0.39 ± 0.17	0.48 ± 0.14
β2-β3	621-636	3.94 ± 0.23	3.86 ± 0.23	1.62 ± 0.13	1.68 ± 0.12	0.24 ± 0.19	0.31 ± 0.13
BBR	639-652	3.70 ± 0.17	3.86 ± 0.20	1.66 ± 0.12	1.64 ± 0.09	0.42 ± 0.16	0.38 ± 0.10
PBC	659-676	4.21 ± 0.22	4.53 ± 0.43	1.62 ± 0.13	1.63 ± 0.17	0.29 ± 0.20	0.41 ± 0.18
C-helix	699-713	4.38 ± 0.17	4.65 ± 0.28	1.68 ± 0.14	1.77 ± 0.17	0.27 ± 0.19	0.39 ± 0.15
CTL	714-724	2.92 ± 0.18	3.23 ± 0.32	1.84 ± 0.10	1.99 ± 0.15	1.02 ± 0.12	1.21 ± 0.14

2.11 Supplementary Figures

	PBC	*	*	*	*							*	*					
HCN4	659	G	E	I	C	L	L	T	R	G	R	R	T	A	S	V	R	A	D
HCN1	535	G	E	I	C	L	L	T	K	G	R	R	T	A	S	V	R	A	D
HCN2	604	G	E	I	C	L	L	T	R	G	R	R	T	A	S	V	R	A	D
HCN3	488	G	E	I	C	L	L	T	R	G	R	R	T	A	S	V	R	A	D
EPAC1	269	G	Q	L	A	L	V	N	D	A	P	R	A	A	T	I	I	L	R
EPAC2	404	G	K	L	A	L	V	N	D	A	P	R	A	A	S	I	V	L	R
PKA RIα CBD-A	199	G	E	L	A	L	I	Y	G	T	P	R	A	A	T	V	K	A	K
PKA RIα CBD-B	323	G	E	I	A	L	L	M	N	R	P	R	A	A	T	V	V	A	R
PKG RIβ CBD-A	182	G	E	L	A	I	L	Y	N	C	T	R	T	A	T	V	K	T	L

Fig S1. Multiple sequence alignment of the extended phosphate binding cassettes (PBCs) of mammalian CBDs. Residues that are fully conserved within at least three HCN isoforms are highlighted in gray, while those with similar side chain properties are highlighted in blue and those that are conserved among non-HCN CBDs are shown in green. The S672R mutation site is outlined with a red rectangle and the residues that form hydrogen bonds with cAMP in HCN4 are marked by asterisks along the top.

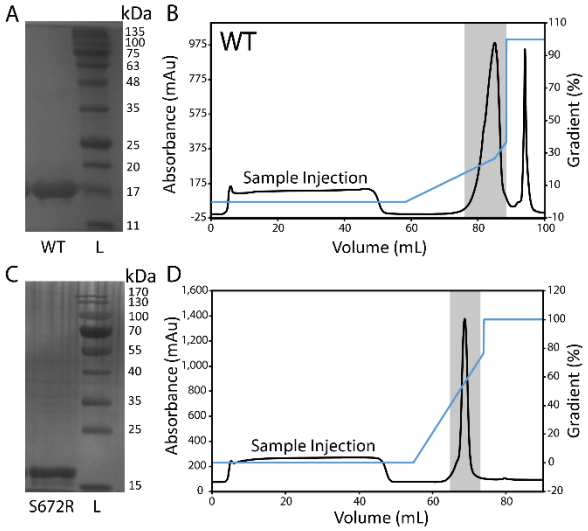

Fig S2. Purity of the WT and S672R HCN4 (563-724) constructs as assessed by SDS-PAGE and FPLC cation exchange chromatography. A) SDS-PAGE gel of WT HCN4 (563-724) with ladder (L). The expected molecular weight of this construct is ~18 kDa. **B)** FPLC chromatogram from cation exchange purification for WT HCN4 (563-724). The black line represents the absorbance at 254 nm, while the blue line depicts the gradient of Sp elution buffer (1M KCl). The gray box highlights the collected fractions of purified HCN that were used for NMR experiments. **C)** SDS-PAGE gel for a ~one year old S672R HCN4 (563-724) sample. **D)** Cation exchange chromatogram for S672R HCN4 (563-724).

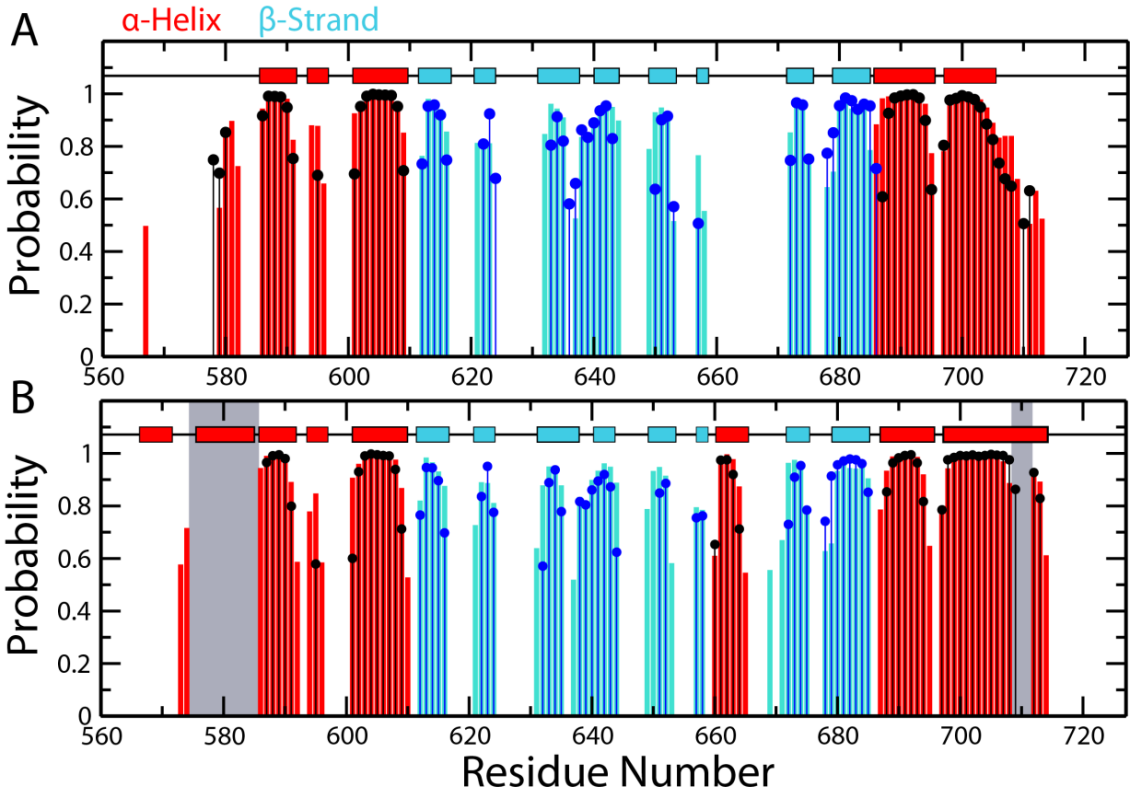

Fig S3. Secondary structure prediction for the wt and S672R CBDs of the HCN4 CBD using TALOS+ [51] **with NH, Cα, Cβ and C' chemical shifts**. **A)** Secondary structure probabilities for the apo HCN4 CBD. The α-helices and β-strands are shown as red and cyan bars for the S672R mutant and as black and blue lines for the wt. The secondary structures from the apo solution structure of HCN4 (a.a. 579-707; PDB ID: 2MNG [9]) are depicted along the top using the same colour code as for the S672R secondary structure probabilities. **B)** Secondary structure prediction for the cAMP-bound state of the HCN4 CBD. Colour codes are the same as panel A. The secondary structure highlighted along the top of the plot was derived from the X-ray crystal structure of cAMP-bound HCN4 (a.a. 521-717; PDB ID: 3OTF [8]). Gray boxes highlight regions with incomplete chemical shift assignment.

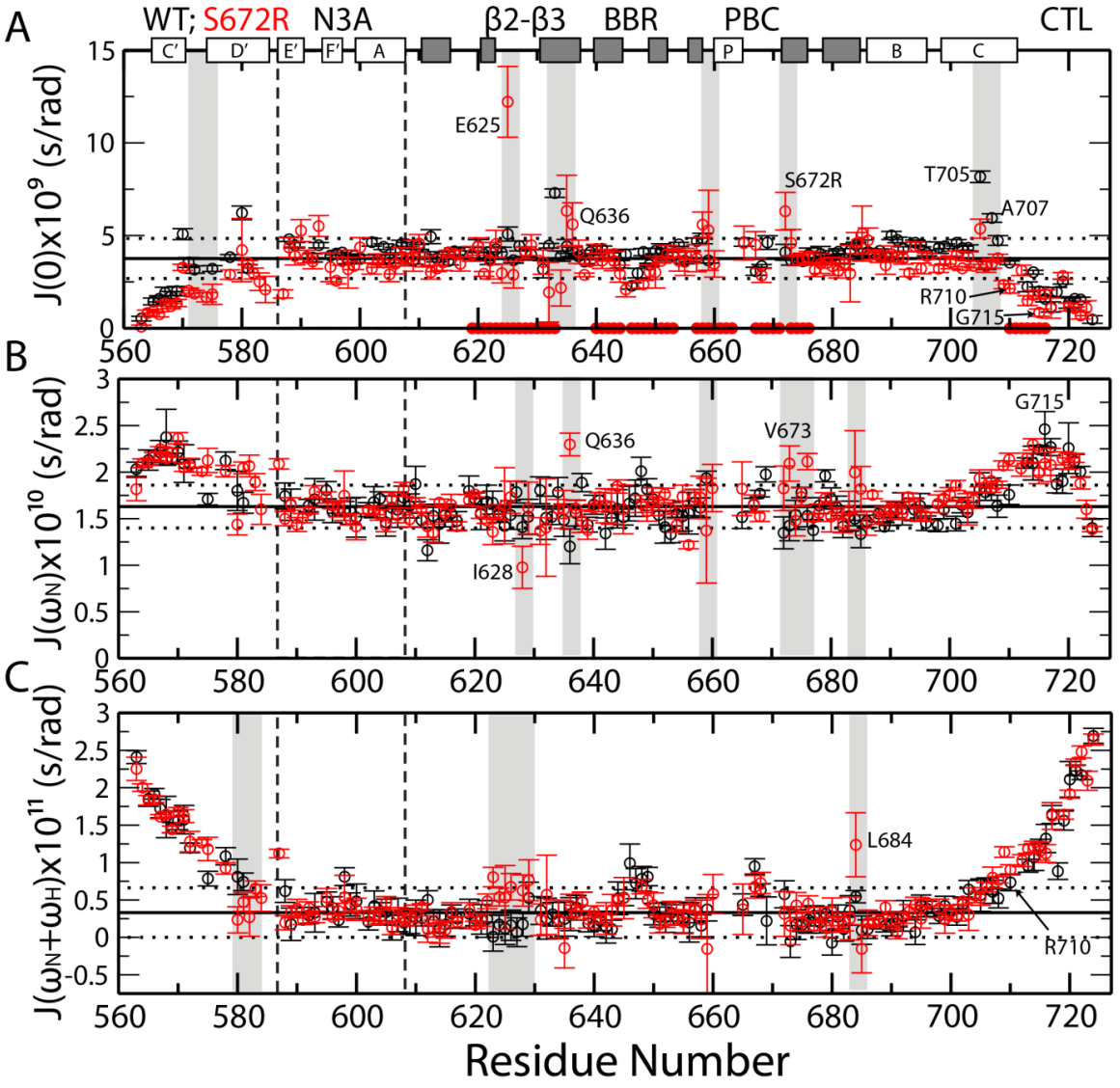

Fig S4. Reduced spectral densities for apo WT (black) and S672R (red) HCN4 (563-724). These spectral densities were computed from the R_1 and R_2 values, as well as heteronuclear {^{1}H-^{15}N}-NOEs (HN-NOEs) [9]. The secondary structure is depicted along the top of the J(0) plot as white (α-helices) and gray (β-strands) rectangles. The solid and dashed horizontal lines represent the average spectral density and the average value +/- five standard deviations, respectively, as explained in the main text. Selected residues subject to significant changes in dynamics are highlighted with gray boxes. The dashed vertical lines define the boundaries for the N3A motif.

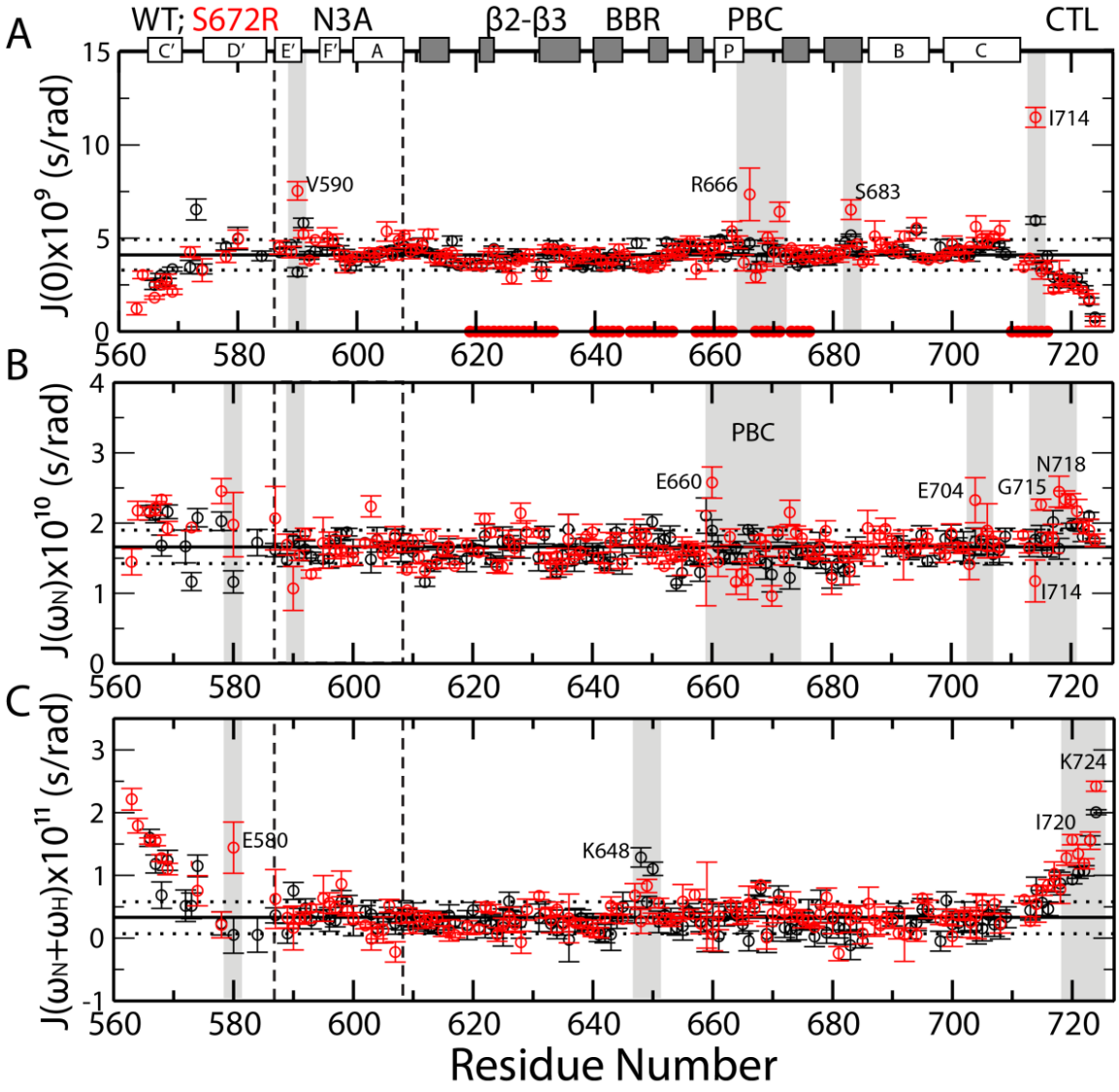

Fig S5. Reduced spectral densities for cAMP-bound WT (black) and S672R (red) HCN4 (563-724). Notations and color codes are as in Figure S4.

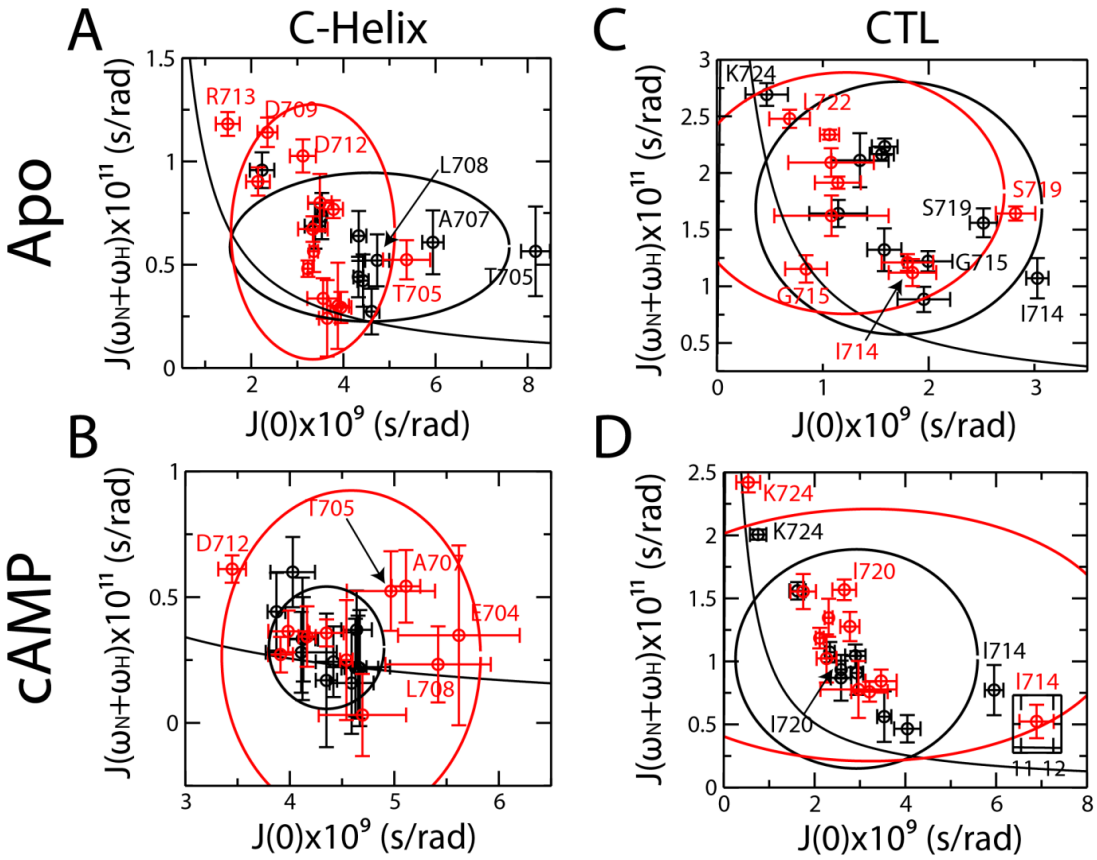

Fig S6. Two-dimensional reduced spectral densities plots for the C-Helix (a.a. 699 – 713) and CTL (a.a. 714 – 724). The layout is the same as in Fig 4, panels C-J.

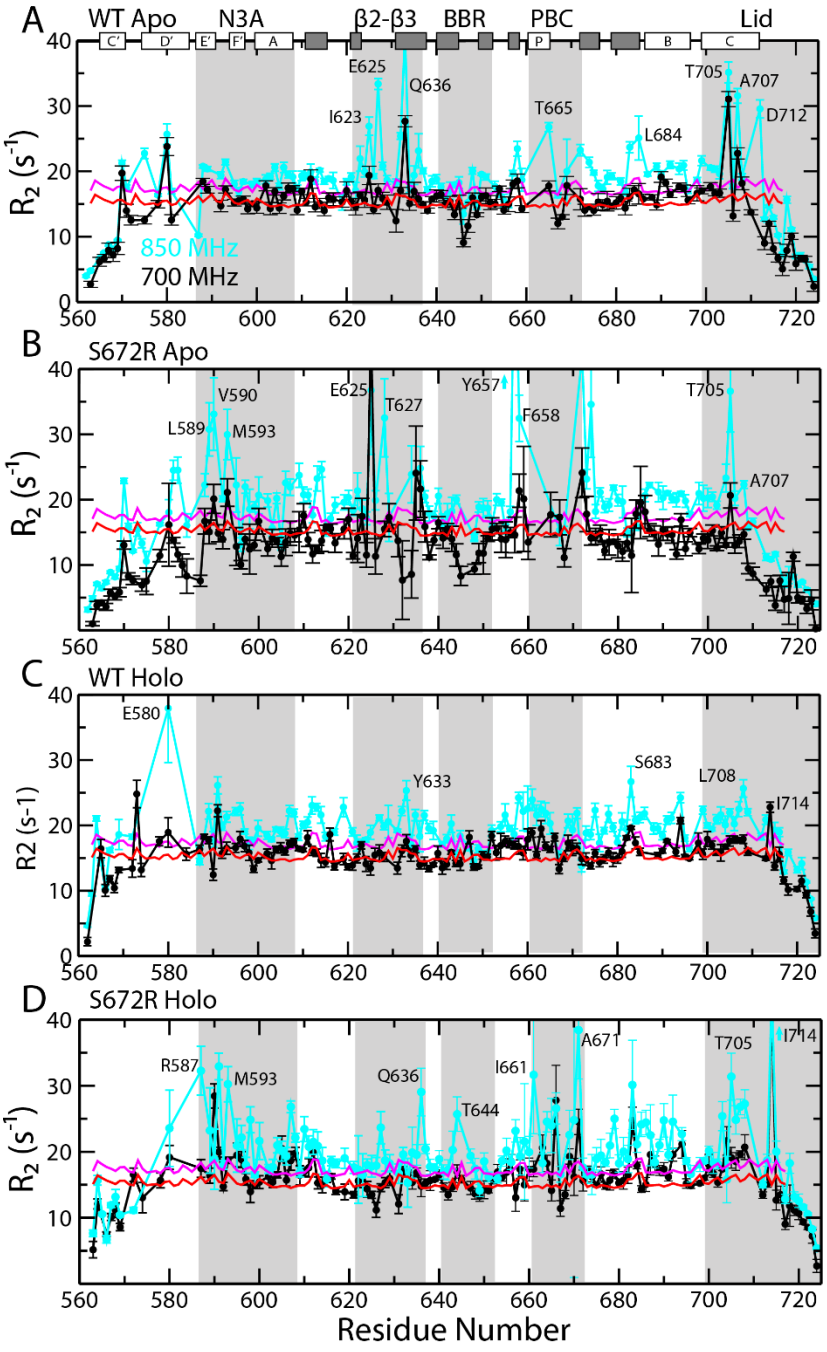

Fig S7. Comparison of R_2 relaxation rates at 850 MHz and 700 MHz for WT Apo (A), S672R Apo (B), WT cAMP-bound (C) and S672R cAMP-bound (D) HCN4. R_2 rates at 700 and 850 MHz are shown in black and cyan, respectively. The red and magenta lines depict the simulated R_2 rates, computed with HydroNMR[56], at 700 and 850 MHz, respectively.

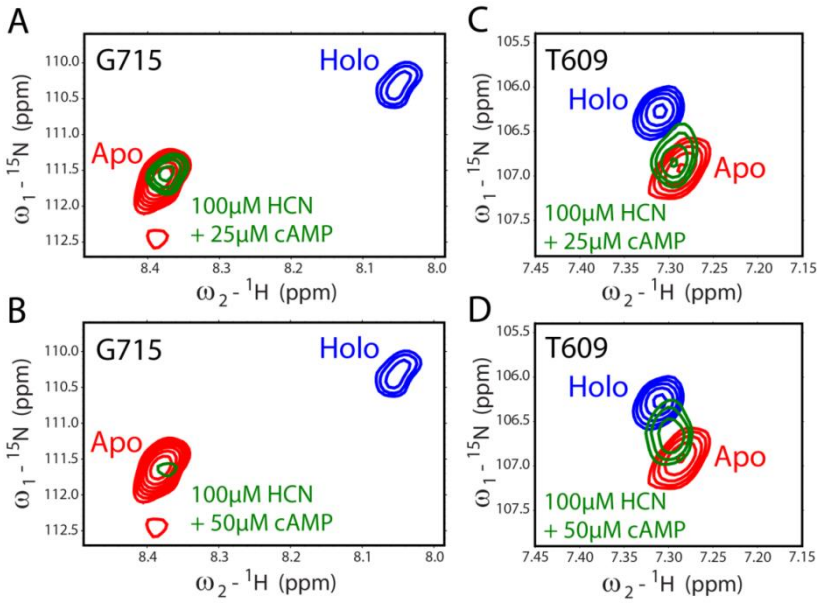

Fig S8. Evidence that residue G715 undergoes slow apo-holo exchange in the S672R HNC4 CBD. **A)** NH-HSQC spectra showing residue G715's apo (red), holo (blue) and partially bound (green) states. The lack of a partially bound peak at the position of the holo state raises the question of whether the G715 residue is in slow or fast exchange. **B)** The same spectra as panel A but after increasing the cAMP concentration from 25 to 50 μM. An increase in ligand concentration causes an increase in the exchange rate, which, if in the fast exchange regime, would cause sharper and more intense cross-peaks. In this spectra, however, there is a further broadening of the partially bound signal (green cross-peak), suggesting that in panel A, G715 was in the slow exchange and the increase in cAMP concentration drove the protein towards the intermediate exchange regime. The absence of the G715 holo peak for the partially bound sample in panel A likely reflects the fact that the holo state is only minimally populated and partially broadened. **C,D)** A similar analysis was performed for residue T609, which is already in the fast exchange regime at 25μM cAMP and therefore serves as a control. When the cAMP concentration was increased to 50μM (D), there was no major signal loss, unlike for G715.

Chapter 3

Mechanism of Selective Enzyme Inhibition Through Uncompetitive Regulation of an Allosteric Agonist

3.1 Author's Preface

The work presented in this chapter has been accepted for publication and is in press. It is reproduced here with permission from the American Chemical Society. The full citation is as follows:

Boulton S, Selvaratnam R, Blondeau JP, Lezoualc'h F and Melacini G. Mechanism of Selective Enzyme Inhibition Through Uncompetitive Regulation of an Allosteric Agonist. *J. Am. Chem. Soc.* (2018) [In Press].

I conducted most of the experiments necessary for this publication. Dr. Rajeevan Selvaratnam acquired preliminary data on CE3F4-EPAC interactions. Dr. Jean-Paul Blondeau provided GDP exchange rate data used in Figure 3H. Dr. Frank Lezoualc'h provided the CE3F4 compounds. I co-wrote the manuscript with Dr. Giuseppe Melacini.

3.2 Abstract

Classical uncompetitive inhibitors are potent pharmacological modulators of enzyme function that selectively target enzyme-substrate complexes (E:S). Therefore, their inhibitory potency is amplified by increasing substrate concentrations. Recently, an unconventional uncompetitive inhibitor, called CE3F4R, was discovered for the exchange protein activated by cAMP isoform 1 (EPAC1). Unlike classical uncompetitive inhibitors, CE3F4R is uncompetitive with respect to an allosteric effector, cAMP, as opposed to the substrate, i.e. CE3F4R targets the E:cAMP rather than the E:S complex. However, the mechanism of CE3F4R as an uncompetitive inhibitor is currently unknown. Here, we elucidate the mechanism of CE3F4R's action using NMR spectroscopy. Due to limited solubility and line broadening, which pose major challenges for traditional structural determination approaches, we resorted to a combination of protein- and ligand-based NMR experiments to comparatively analyze EPAC mutations, inhibitor analogs and cyclic nucleotide derivatives that trap EPAC at different stages of activation. We discovered that CE3F4R binds within the EPAC cAMP-binding domain (CBD) at a sub-domain interface distinct from the cAMP binding site, acting as a wedge that stabilizes a cAMP-bound mixed-intermediate. The latter includes attributes of both the apo/inactive and cAMP-bound/active states. The mixed-intermediate targeted by CE3F4R traps the CBD's hinge helix in its inactive conformation, locking EPAC into a closed domain topology that restricts substrate access to the catalytic domain. The proposed mechanism of action also explains the isoform selectivity of CE3F4R in terms of a single EPAC1 vs. EPAC2 amino acid difference that destabilizes the active conformation of the hinge helix.

3.3 Introduction

Uncompetitive inhibition is a pharmacologically effective means of modulating enzyme activity[1,2]. Conventional uncompetitive inhibitors specifically target the enzyme-substrate (ES) complex without competiting with the substrate for binding[1,2]. Therefore, the potency of uncompetitive inhibitors increases with increasing substrate concentrations and inhibition is amplified when the enzyme is at its most active state[1,2]. Furthermore, the recognition of a binary ES complex by uncompetitive inhibitors provides an additional layer of selectivity that is not present when targeting an isolated enzyme[1,2]. Hence, uncompetitive inhibition provides unique opportunities to simultaneously maximize both potency and specificity.

A non-classical type of uncompetitive enzyme inhibition, where the inhibitor is uncompetitive with respect to an allosteric agonist rather than the substrate, was recently discovered for the exchange protein activated by cAMP (EPAC)[3,4]. EPAC is a guanine exchange factor (GEF) for Rap GTPases, which is allosterically modulated by the secondary messenger, cAMP[5–9]. EPAC proteins regulate a broad array of physiological processes, including calcium homeostasis in cardiomyocytes, which affects cardiac excitation and contraction, insulin secretion from pancreatic β cells, integrin-mediated cell adhesion, long-term potentiation and depression in the hippocampus[5,10–13]. Therefore, EPAC proteins are considered targets for a wide range of therapeutic applications[10] and their inhibition has shown potential for treating pancreatic cancer[11,12], breast cancer[13], atherosclerosis[14], chronic inflammatory pain[15], diabetes[16,17] and infectious diseases such as rickettsioses[18], Coronavirus[19], and malaria[20]. The novel uncompetitive inhibitor, denoted as CE3F4R (Fig. 1A), was identified in a high throughput screen for its ability to inhibit EPAC mediated GDP exchange both *in vitro* and in cardiac myocytes[3]. The unusual aspect of the CE3F4R inhibitor is that it is uncompetitive with respect to the allosteric effector, cAMP, as opposed to the enzyme substrate, Rap (Fig. 1B)[3,4]. Specifically, the CE3F4R inhibitor does not act by interacting with the traditional ES complex, which in this case is EPAC:Rap, but instead it binds the enzyme-allosteric effector complex (EPAC:cAMP) (Fig. 1B). In this study, we seek to determine the mechanism of action for CE3F4R as a non-classical uncompetitive inhibitor for the cAMP-dependent activation of EPAC.

To understand the mechanism of action for CE3F4R it is essential to consider that the cAMP-dependent activation of EPAC is regulated by the relative orientation of an N-terminal regulatory region (RR) and a C-terminal catalytic region (CR). The RR and CR share a conserved architecture in the two primary mammalian EPAC isoforms (EPAC1 and EPAC2, Fig. 1C)[5,7,9,21,22]. The RR

contains a disheveled-Egl-10-plekstrin (DEP) domain, which associates the protein to the plasma membrane, and a cAMP-binding domain (CBD), which allosterically regulates GEF activity[5,9,21,22]. The CR of both isoforms (Fig. 1C) consists of a Ras exchange motif (REM), a Ras association domain (RA) and a CDC25 homology domain (CDC25HD)[5,9,21,22]. The CDC25HD catalyzes GDP-GTP exchange for Rap, while the REM domain helps stabilize a catalytic helix and the RA domain influences subcellular localization[5,9,21,22]. EPAC2 houses an additional CBD at its N-terminus, which is dispensable for cAMP-dependent GEF activity[23]. On the contrary, the conserved regulatory CBD at the RR C-terminus[22,24,25] controls the relative orientations of the RR and CR with respect to each other. When cAMP is absent, the two regions adopt a closed topology, in which the RR restricts access of Rap GTPases to the CR, resulting in constitutive inhibition. Upon binding of cAMP to the conserved CBD, the two regions adopt an open (active) topology, where the Rap substrate can access the CR and GDP-GTP exchange is promoted.

The CBD is composed of a discontinuous α-helical subdomain interjected by a β-subdomain (Fig. 1B)[21,22,24–27]. The N-terminus of the α-subdomain spans four helices (Fig. 1D-E), *i.e.* α1, which, together with α2, forms a salt bridge with the CR in the closed topology, known as the ionic latch (IL), and α2-4, which define a conserved motif referred to as 'N3A'. The β-subdomain houses two cAMP binding elements, the base binding region (BBR) and the phosphate binding cassette (PBC). The latter is essential for cAMP binding and for driving the cAMP-dependent conformational changes[21,22,24–27]. The C-terminal α-subdomain contains a hinge helix that is critical for both the auto-inhibition and the cAMP-dependent activation of EPAC[21,22,24–27].

The relative orientations of the PBC and hinge helices (Fig. 1E) controls the transition from closed to open CR / RR topologies. The position of the PBC and hinge helices is in turn controlled by a hydrophobic contact between a leucine in the PBC and a phenylalanine in the hinge helix[27–29] (Fig. 1F). In the apo/inactive state, the PBC is in an "out" position, which consequently forces the hinge helix into an "out" position as well (Fig. 1E-F). In this PBC out/hinge out conformation, the RR occludes the catalytic site of the CDC25HD domain ("closed inactive topology"). When cAMP binds, the PBC shifts to an "in" orientation, which relieves the steric clash that would otherwise occur between the PBC leucine and the hinge phenylalanine in a "PBC out/hinge in" conformation (Fig. 1E-F). The "PBC in" conformation then results in the hinge-helix rotating to an "in" orientation and partially unfolding. In the PBC in/hinge in conformation (stabilized by cAMP), the RR is displaced away from the CR ("open active topology"), making the catalytic site accessible to Rap binding.

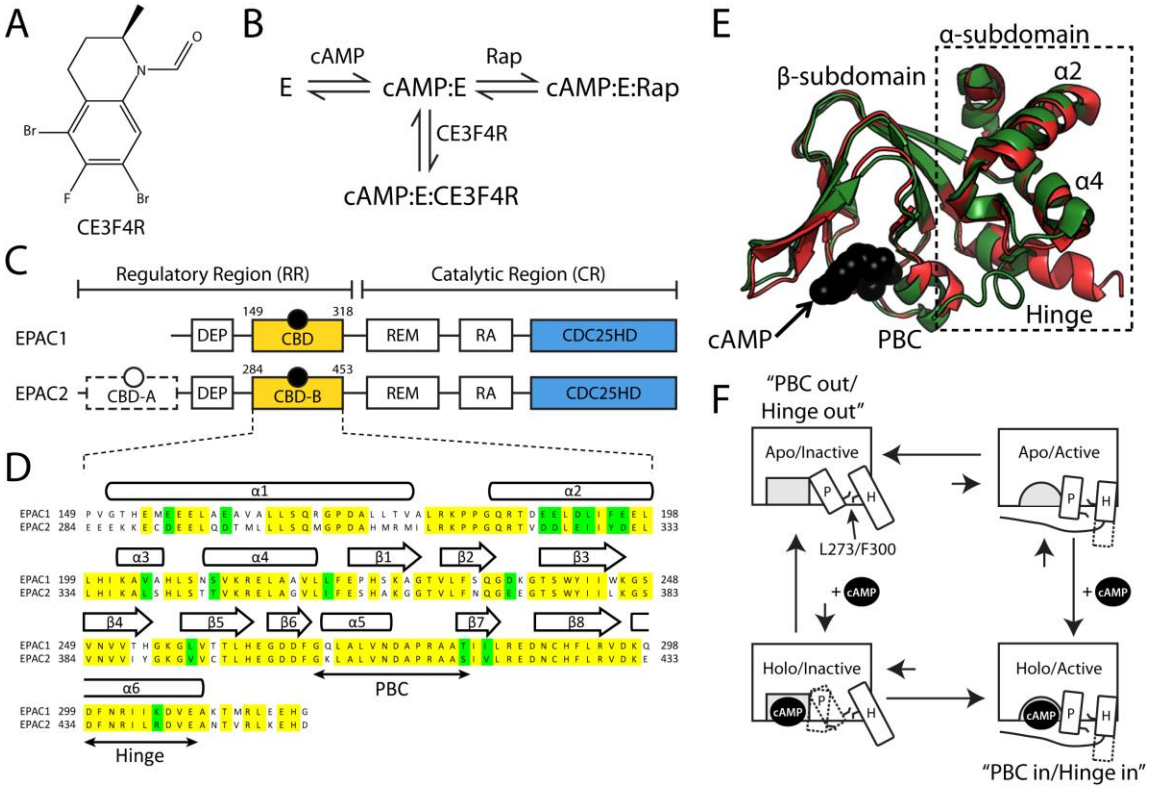

Figure 1. Structure of the uncompetitive EPAC inhibitor CE3F4R and domain topology of EPAC. **A)** Molecular structure of CE3F4R. **B)** Schematic representation of the binding equilibria of CE3F4, which is uncompetitive with respect to cAMP, the endogenous allosteric effector of EPAC, denoted here as E, as opposed to the EPAC substrate, Rap. **C)** Domain organization of EPAC proteins. The regulatory region (RR) consists of a DEP (disheveled Egl-10 pleckstrin) domain and cAMP-binding domains (CBDs). The essential CBD for regulation is coloured in orange, with a molecule of bound cAMP depicted as a black circle. The amino acid range for these domains in EPAC1 and EPAC2 are indicated above the orange rectangle. In EPAC2, there is an additional CBD (CBD-A) that is not essential for regulation and is therefore shown with a dashed rectangle and a white circle for bound cAMP. The catalytic region (CR) contains the RAS exchange motif (REM), the RAS association (RA) domain and the CDC25 homology domain (CDC25HD). The CDC25HD domain (blue) is the main component of EPAC's GEF activity. **D)** Sequence alignment of the EPAC1 CBD and the EPAC2 CBD-B. Yellow indicates a conserved residue, while green indicates amino acids with similar side chain properties. The secondary structure is indicated above the sequence as cylinders (helices) and arrows (strands) and select regions important for cAMP binding or activation are specified with double sided arrows. **E)** Overlay of the apo-inactive (red) and holo-active (green) crystal structures of the EPAC2's CBD-B[21,22]. **F)** The four-state thermodynamic cycle of the conserved EPAC CBD, which arises from the coupling of the auto-inhibitory equilibrium and the cAMP binding equilibrium. The P and H rectangles correspond to the PBC and Hinge helices and the lines between them represent the hydrophobic hinge defined by L273 in the PBC and F300 in the hinge helix. The L273/F300 hydrophobic hinge controls the helices' positions in relation to each other. Dotted lines indicate partial unfolding or unknown states.

The four-state thermodynamic cycle created from the coupling between the CBD auto-inhibitory equilibrium and the cAMP binding equilibrium (Fig. 1F) provides a foundation for understanding the mechanism of CE3F4R action[30–41]. This is because, when the EPAC1 CBD is removed, CE3F4R is unable to inhibit GEF activity, suggesting that CE3F4R interferes with the activation of the CBD by cAMP, rather than acting directly on the EPAC catalytic site[3]. Furthermore, CE3F4R is selective for EPAC1 *vs.* EPAC2, suggesting that the additional CBD of EPAC2 is not required for inhibition[4]. Hence, we focused our analyses on the effect of CE3F4R on the allosteric thermodynamic cycle of the EPAC1 CBD (Fig. 1F), spanning residues 149-318 (Fig. 1D).

We hypothesized that the uncompetitive EPAC inhibitor shifts the cAMP-bound equilibria towards the holo-inactive state (Fig. 1F) through the formation of a cAMP:EPAC1 CBD: CE3F4R ternary complex. To test this hypothesis and identify CE3F4R's mechanism of action, we utilized NMR spectroscopy to map the interactions of CE3F4R with the CBD of EPAC1. However, CE3F4R suffers from limited solubility and intermediate exchange line broadening occurring upon binding EPAC, as commonly observed for drug leads from screening efforts[42–46]. These experimental circumstances typically require the use of non-saturating and sub-stoichiometric concentrations and pose significant limitations to the application of traditional structural determination methods, such as X-ray crystallography and NOESY-based structure determination methods[47]. Classical NOESY experiments can be informative even for partially saturated proteins. However, the limited sensitivity caused by line-broadening and low concentrations imposed by poor solubility, did not permit traditional NOESY-based structure determination approaches in our case. To overcome these challenges, we relied on an experimental design centered around the comparative analysis of ligand- and protein-based NMR experiments applied to a library of EPAC mutations, cyclic nucleotide derivatives and CE3F4R analogs. The combination of NMR experiments includes the chemical shift projection analysis (CHESPA)[28], ligand-to-protein and protein-to-ligand saturation transfer difference (STD) experiments[24,48–51] as well as paramagnetic relaxation enhancement (PRE) measurements[52–54] with spin labels located at multiple sites to triangulate the inhibitor binding pocket. This experimental design is generally suitable for hydrophobic drug leads and the resulting mechanism of action reveals how non-classical uncompetitive inhibition occurs with respect to an allosteric activator as opposed to a substrate. Furthermore, the proposed inhibitory mechanism rationalizes the isoform selectivity of CE3F4R for EPAC1 *vs.* EPAC2.

3.4 Results

3.4.1 CE3F4R Preferentially Binds the cAMP-Bound *vs*. Apo EPAC1 CBD.

As a first step towards probing the interactions between CE3F4R and the EPAC1 CBD, NH-HSQC spectra of both apo and cAMP-bound EPAC1 (149-318) were acquired, each in the absence and presence of CE3F4R (Fig. 2A, B). For the apo state, the chemical shift variations upon CE3F4R addition are marginal with the majority of cross-peaks in the two HSQC spectra remaining superimposed (Fig. 2A). In contrast, the chemical shift changes observed upon addition of CE3F4R to the cAMP-bound CBD are significantly larger (Fig. 2B), suggesting that CE3F4R interacts preferentially with the cAMP-bound state of EPAC1. Although the CE3F4R-induced chemical shift changes observed for EPAC1 in the absence of cAMP are on average lower than those observed in the presence of cAMP, residual ppm shifts are still observed for apo EPAC1 due to weak residual binding of CE3F4R. These results are consistent with CE3F4R functioning as a non-classical uncompetitive inhibitor with respect to cAMP[3]. Therefore, we focused subsequent analyses on the interaction of CE3F4R with cAMP-bound EPAC1 CBD.

Some of the largest chemical shift differences upon addition of CE3F4R to cAMP-bound EPAC1 are in the α4 helix, β3 strand, N-terminal phosphate binding cassette (PBC) and α6 (hinge) helix (Fig. 2C). When the residues with CCS differences greater than the average plus one standard deviation are mapped onto the crystal structure, they form an almost continuous surface at the interface of the α- and β-subdomains (Fig. 2D). A notable feature of this surface is that it does not include key cAMP binding sites. For example, the surface map of Fig. 2D does not include the BBR, the C-terminal PBC and the β2-β3 loop, which is adjacent to the PBC (Fig. 2C, D). Hence, it is unlikely that CE3F4R binds in direct contact with cAMP. Instead, the CE3F4R binding site is likely within the confines of the CCS-based surface shown in Fig. 2D. This result implies that CE3F4R is an allosteric uncompetitive inhibitor. However, the exact location of the CE3F4R binding-site within the surface outlined in Fig. 2D cannot be accurately predicted at this stage of the analysis, since CCS chemical shifts report on both ligand binding and allosteric conformational changes and these two contributions need to be separated.

3.4.2 Initial Separation of Binding and Allosteric Effects.

In order to separate the binding *vs.* allosteric contributions in the CE3F4R derived CCS changes (Fig. 2C, D) and investigate the type of conformational change driven by CE3F4R, we

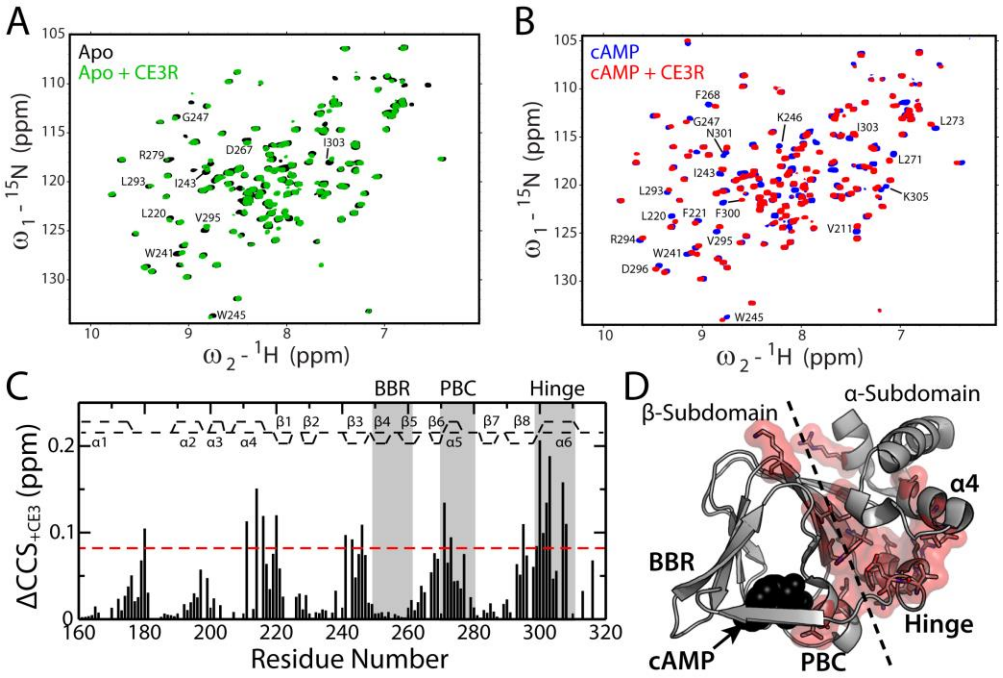

Figure 2. CE3F4R preferentially binds the cAMP-bound CBD of EPAC1. **A)** Overlay of Apo (black) and Apo + CE3F4R (green) NH-HSQC spectra. Selected residues are labeled to facilitate the comparison with panel B. The DMSO concentration was kept constant at 1% v/v for all samples. **B)** As panel A, but for the cAMP-bound EPAC1 CBD without (blue) and with CE3F4R (red). **C)** The compounded chemical shift differences (ΔCCS) *vs.* residue plot for the spectra in panel B. The horizontal dashed lines represent the average CCS + one standard deviation (SD). The secondary structure of EPAC is denoted along the top of the plot. **D)** Surface map of residues in which the CCS differences for the cAMP:EPAC1 CBD complex +/- CE3F4R are greater than the average plus one standard deviation.

we complemented the CCS data of Fig. 2 with the chemical shift projection analysis (CHESPA; Fig. 3)[28,49–51]. The CHESPA measures both the magnitude and the direction of chemical shift changes of a perturbed state relative to a set of reference states (Fig. 3A)[28]. For example, a perturbation vector is defined by the chemical shift changes between the EPAC:cAMP and the EPAC:cAMP:CE3F4R states and a reference vector is determined by the chemical shift changes between the EPAC:cAMP and the apo EPAC states (Fig. 3A). Fractional activations (X) are then computed for each residue from the normalized projection of the perturbation vector onto the reference vector (Fig. 3A) and the direction of the chemical shift change is determined by calculating the angle (θ) between the two vectors (Fig. 3A)[28]. For residues sensing primarily allosteric conformational changes and adhering to a fast-exchanging two-state model, the CE3F4R-induced chemical shift variations are expected to be nearly aligned with the reference states (*i.e.* $|\cos\theta| \geq 0.9$)[28]. In addition, assuming that the apo and cAMP-bound states serve as proxies for the fully inactive and active states, respectively, the X values for linear shifts in the -1 to 0 range reflect

CE3F4R-induced changes in the relative inactive *vs.* active populations. The projection analysis therefore provides directional information at residue-resolution on the nature of the allosteric conformation shifts (*e.g.* inhibition *vs.* activation). In contrast, residues near the CE3F4R binding site do not sense only allosteric conformational changes, but also report on the unique chemical environment caused by their proximity to CE3F4R. The latter may result in deviations from the two-state model and non-linear chemical shift changes (*i.e.* $|\cos\theta| < 0.9$) are expected as well as X values that may fall outside the -1 to 0 range, offering a simple means to start separating binding *vs.* allosteric contributions on a per-residue basis and thus to begin narrowing down the location of the CE3F4R binding site.

3.4.3 CE3F4R Targets a Mixed Intermediate Distinct from both Apo and cAMP-Bound EPAC CBD Structures.

The CHESPA analysis reveals that upon addition of CE3F4R, most residues shift towards the inactive (apo) state (Fig. 3B, C, E, F), consistent with the overall inhibitory function of CE3F4R revealed by enzymatic assays[3,4]. In particular, the hinge helix, which exhibits the largest overall chemical shift changes, shifts towards the inactive (apo) state by an average of ~50% (Fig. 3C, E). In addition, the α4 helix, which is adjacent to the hinge helix, also shifts back by ~45% on average (Fig. 3B,E). Both α4 and α6 regions feature linear chemical shift changes as indicated by cosθ values close to -1, consistent with a fast-exchanging two-state equilibrium (Fig. 3F). However, not all EPAC1 CBD regions display the same extent of inhibitory shift (X, Fig. 3E). For example, the PBC did not significantly change compared to the active (cAMP-bound) state (Fig. 3D, E).

Overall, the fractional activation analysis (Fig. 3E) reveals that the inhibition elicited by CE3F4R does not arise from a simple reversal of the inactive *vs.* active equilibrium, but rather from the stabilization of an intermediate with mixed active and inactive features in different regions of the EPAC1 CBD. CE3F4R binding shifts the hinge helix towards the inactive conformation, referred to as "hinge out" topology, while keeping the PBC in its active, cAMP-bound-like conformation, denoted as "PBC in" (Fig. 3E). This mixed "PBC in/hinge out" topology reflects a unique intermediate that is selected for by CE3F4R. Hence, we hypothesized that modifications, such as EPAC mutations or cAMP analogs, that stabilize this mixed intermediate will increase the affinity of CE3F4R, whereas those that destabilize the intermediate will decrease the affinity, because the observed affinity is a population weighted average of state specific association constants[24,51,55].

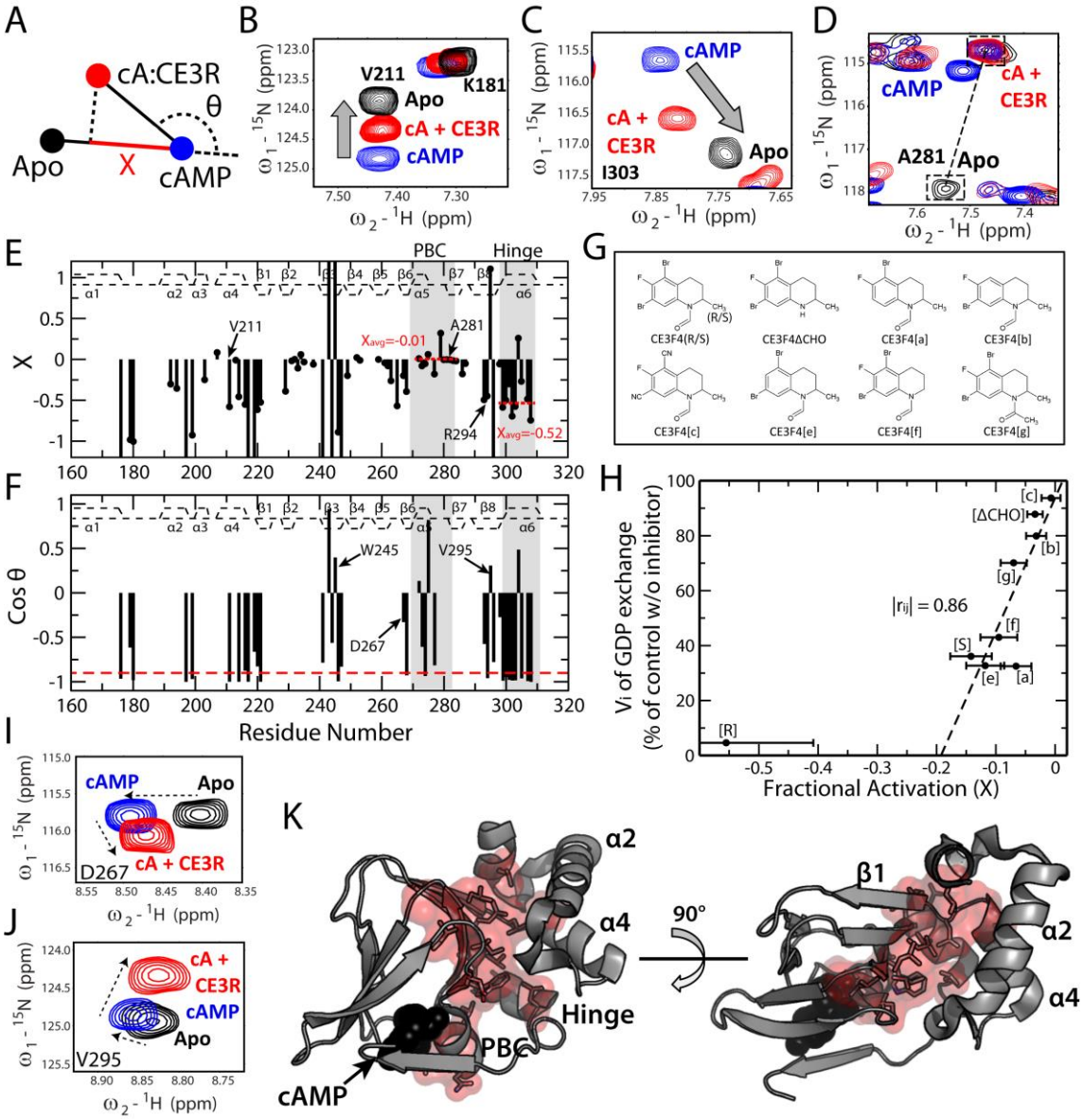

Figure 3. CE3F4R deactivates cAMP-bound EPAC by stabilizing a distinct inactive "PBC in/Hinge out" mixed intermediate. **A)** Schematic illustration of the chemical shift projection analysis that shows how the CE3F4R-derived fractional activations (X) and cosθ values are calculated with respect to the apo and cAMP-bound states. **(B-D)** Representative residues from the NH-HSQC spectra, which illustrate the CE3F4R-dependent CCS changes in the α4 helix **(B)** hinge helix **(C)** and PBC **(D)**. **E)** CHESPA fractional activation analysis of CE3F4R binding to the cAMP:CE3F4R complex. **F)** CHESPA cos(θ) *vs*. residue plot for the cAMP:EPAC1 CBD:CE3F4R complex. The dashed red line is the cutoff (-0.9) utilized here for distinguishing linear *vs*. non-linear chemical shift changes. **G)** Molecular structures of CE3F4 analogs. **H)** Plot of rates of GDP exchange *vs.* average fractional activations (X) of the hinge helix (residues 298-308) for the CE3F4 analogs in panel G. The GDP exchange rates are from Courilleau *et al.*[4]. All compounds with the exception of CE3F4R, CE3F4S and CE3F4[f] are racemic mixtures of the R and S enantiomers. **I-J)** Representative residues from NH-HSQC spectra that show non-linear chemical shift changes caused by CE3F4R binding. **K)** Resides with cos(θ) values greater than -0.9 or X values outside one standard deviation of the -1 to 0 range were mapped on the crystal structure as red surface representation.

To test our hypothesis about CE3F4R targeting the "PBC in/hinge out" mixed intermediate of the EPAC1 CBD, we prepared the L273W EPAC mutation. The L273W mutation introduces steric hindrance between the PBC P-helix and the adjacent hinge through the bulky indole group (Fig. 4A)[27,28]. As a result, the L273W EPAC1 mutant locks the hinge helix in the inactive topology (*i.e.* "hinge out") even in the presence of cAMP, suppressing the ability of cAMP to activate EPAC1[27,28]. Meanwhile, the PBC stays in the "in" state, similarly to the cAMP-bound WT (Fig. 4A-C)[27,28]. Therefore, the L273W mutation stabilizes the mixed "PBC in/hinge out" intermediate prior to addition of CE3F4R, which we predict will increase CE3F4R binding affinity relative to the WT. This prediction was confirmed by measuring the affinity of CE3F4R for L273W through chemical shift titrations, which show that the affinity of CE3F4R towards the mutant is more than five times higher than for the WT (Fig. 4E, F).

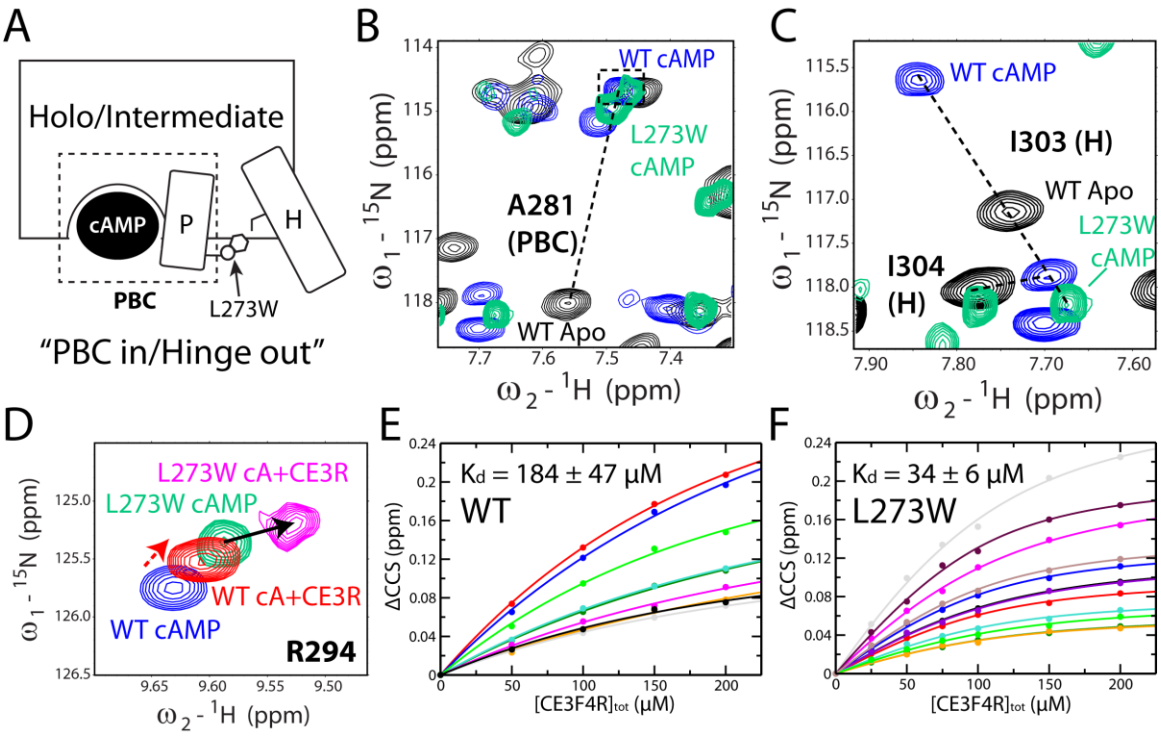

Figure 4. Stabilization of the "PBC in/Hinge out" mixed intermediate by the L273W mutation increases the CE3F4R Binding affinity for the cAMP-bound EPAC1 CBD. **A)** Schematic representation of the L273W mutation and its effect on the PBC and hinge positions. **(B and C)** Representative HSQC spectra of residues from the PBC and hinge that illustrate the stabilization by L273W of the "PBC in/hinge out" intermediate even in the absence of CE3F4R. **D)** Representative HSQC peak for cAMP-bound WT and L273W EPAC1 CBD in the absence and presence of CE3F4R. **(E and F)** CE3F4R titration curves for WT (E) and L273W (F) EPAC in the presence of cAMP. Each curve corresponds to the ΔCCS of a different residue. K_d values were determined from the average of all fitted residues.

Further independent evidence supporting the role of the "hinge out" topology in the uncompetitive inhibition by CE3F4R is provided by CE3F4R derivatives that exhibit lower inhibitory levels, as revealed through enzymatic assays of a fully functional EPAC1 construct[3,4] (Fig. 3G). Stereochemical and/or covalent modifications of CE3F4R that lead to increased rates of GDP exchange, also result in significant reductions of the fractional shifts of the hinge helix towards the inactive state (X ~ -1), as detected through our NMR CHESPA analysis on EPAC1 (149-318) (Fig. 3H). A linear regression (Fig. 3H, dashed line) was computed for all analogs without a pure R stereocenter. Resulting in a Pearson's correlation coefficient of 0.86. In addition, the Spearman's rank correlation coefficients were determined for all analogs including CE3F4R and the cAMP-bound state. The ρ value was 0.85 with a two-tailed P value of 0.2%, indicating that the correlation between the GDP exchange rates and the fractional activation (X) values determined through NMR chemical shifts is statistically significant. The ability of EPAC1 (149-318) to differentiate between CE3F4R and its analogs (Fig. 3G, H) not only indicates that the interactions between CE3F4R and EPAC1 (149-318) analyzed here are specific, but it also corroborates the relevance of the "hinge out" orientation in the uncompetitive inhibition by CE3F4R.

We also complemented the positive control offered by L273W, with a negative control to show that when the mixed "PBC in/hinge out" intermediate is destabilized, the affinity for CE3F4R is reduced. For this purpose, we analyzed the interaction of CE3F4R with Rp-cAMPS (Rp)-bound EPAC. Rp is a thiophosphate derivative of cAMP that sterically locks the PBC in its "out" conformation (Fig. 5A)[56]. The CHESPA analysis of Rp-bound EPAC confirm that both the PBC and hinge helices shift completely towards an apo/inactive-like state (Fig. 5A-F), hence destabilizing the mixed intermediate hypothesized to be targeted by CE3F4R. As expected, when CE3F4R was added to Rp-bound EPAC, it resulted in negligible chemical shift changes (average of 0.01 ppm; Fig. 5G, H), which are indicative of negligible CE3F4R binding. hence destabilizing the mixed intermediate hypothesized to be targeted by CE3F4R. This negative control further confirms our CHESPA-based hypothesis that CE3F4R targets the "PBC in/hinge out" intermediate as opposed to an apo-inactive-like "PBC out/hinge out" state. Furthermore, the reduced CE3F4R binding observed upon replacement of cAMP with Rp-cAMPS, also confirms that CE3F4R is unlikely to bind adjacent to the adenine base, which is common to the two cyclic nucleotides. However, the identification of the binding site of CE3F4R within this mixed EPAC1 intermediate still needs to be refined.

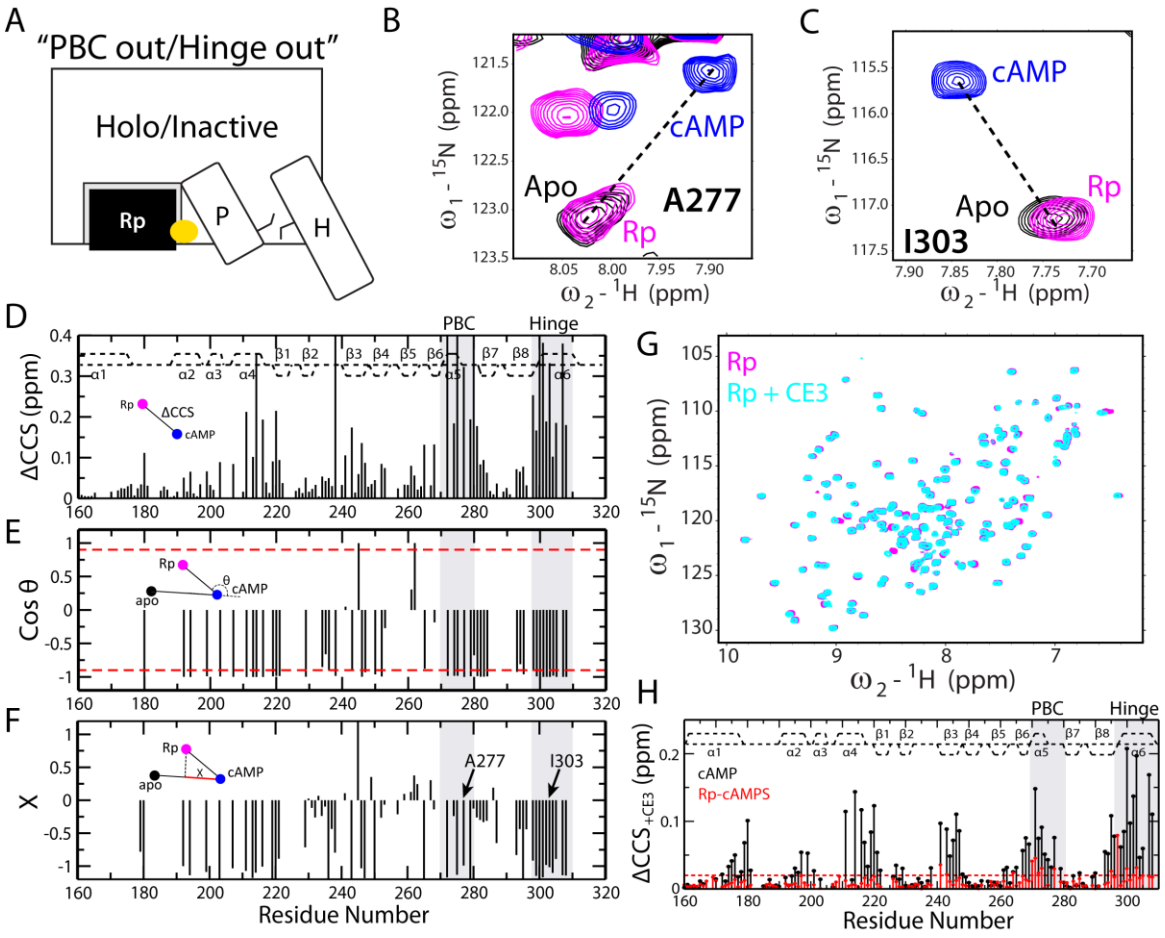

Figure 5. Stabilization of the "PBC out/Hinge out" Conformation with the Rp-cAMPS analog silences CE3F4R Binding. **A)** Schematic of Rp-cAMPS (Rp) mediated inhibition of EPAC. The yellow circle represents the position of the sulfur atom in the thiophosphate group. **B-C)** Representative HSQC spectral regions of residues from the PBC and hinge helix that illustrate the stabilization by Rp of the "PBC out/hinge out" conformation. **D)** CCS differences between cAMP-bound and Rp-bound EPAC. The secondary structure is indicated along the top of the plot. **E)** CHESPA cos(θ) analysis for Rp relative to apo and cAMP-bound states (see inset). The red dashed lines indicate cos(θ) values of -0.9 and 0.9. **F)** CHESPA fractional activation (X) analysis for Rp relative to apo and cAMP-bound states (see inset). **G)** NH HSQC spectra of Rp-bound EPAC1 CBD with (cyan) and without (magenta) CE3F4R. **H)** CCS difference plot for CE3F4R bound and unbound EPAC in the presence of cAMP (black) and Rp (red). The red dashed line indicates the average CCS difference plus one standard deviation for Rp.

3.4.4 The CE3F4R Inhibitor Binds at the α- and β-Subdomain Interface of the EPAC1 CBD.

The CHESPA analysis also helped refine the location of the CE3F4R binding pocket. Residues with non-linear chemical shift changes ($|\cos\theta| < 0.9$) (Fig. 3F, I, J) were mapped onto the X-ray crystal structure of the homologous EPAC2 isoform (Fig. 3K). In contrast to the simple ΔCCS magnitude map (Fig. 2D), in the CHESPA-based binding site map (Fig. 3K) several sites at the hinge, α4 helix and β-subdomains are absent due to the high linearity of their chemical shifts. As a

result, all residues in Fig. 3K are now quite narrowly clustered at the interface of the α- and β-subdomains. To further confirm the position of the CE3F4R binding site, we also acquired STD-(^{13}C-^{1}H) HSQC experiments with saturation of the CE3F4R formyl resonance. Saturation transfer from CE3F4R to the protein was then detected through the ^{13}C^{1}H cross-peaks, which were assigned utilizing HCCH-TOCSY experiments. The STD measurements were performed in D_2O to minimize saturation of amides, although some highly protected amides resonating close to the saturation frequency did not fully exchange. In addition, the H2 and H8 resonances of cAMP are separated from the CE3F4R formyl resonance only by ~0.25 ppm. To account for any potential saturation leak-through arising from protein amides or cAMP resonances, we compared STD-HSQCs in the absence and presence of CE3F4R. Several residues exhibited a significant STD enhancement upon CE3F4R addition (Fig. 6A-C). The STD signals were normalized to a reference spectrum (STR) and the STD/STR differences between the samples with and without CE3F4R were computed (Fig. 6D). The largest increases in STD/STR ratios were observed for residues Y242, I243, D267, and R294 (Fig. 6D-E). Some residues displayed slightly higher STD/STR ratios in the absence of CE3F4R (Fig. 6D), likely because of differences in the extent of residual H_2O content between the two samples. Regardless, the residues that exhibited significant increases in STD/STR ratios in the presence of CE3F4R (Fig. 6D, E) were overall consistent with the independent CHESPA cosθ analysis (Fig. 3H).

As a further independent validation of the proposed CE3F4R binding site, we performed paramagnetic relaxation enhancement (PRE) experiments known as SLAPSTIC (Spin-Labels Attached to Protein Side chain as a Tool to identify Interacting Compounds)[52–54]. In the SLAPSTIC method, a spin-label is covalently conjugated to the protein for the purpose of causing line-broadening for ligands that bind in its vicinity (< ~25 Å). The protein sample is subsequently reduced to separate the spin-label from the protein and ligand peaks that were broadened from the spin-label should sharpen. In our case, the SLAPSTIC approach was implemented as a triangulation technique by engineering two cysteine mutations (S210C and T261C) at opposing ends of the CBD (Fig. 6F), so that both spin-labels are less than 20 Å away from the alleged CE3F4R binding site at the α/β-subdomain interface, as independently proposed based on the CHESPA and STD-HSQC analyses. The covalent attachment of the MTSL spin-label to each mutant was confirmed by comparing the oxidized *vs*. reduced HN-HSQC cross-peak intensities (I_{ox}/I_{red}) to the respective distances from the mutation site, where the spin-label was attached (Fig. S1A, B). This comparison also confirmed that the maximum distance for which a significant broadening effect could be

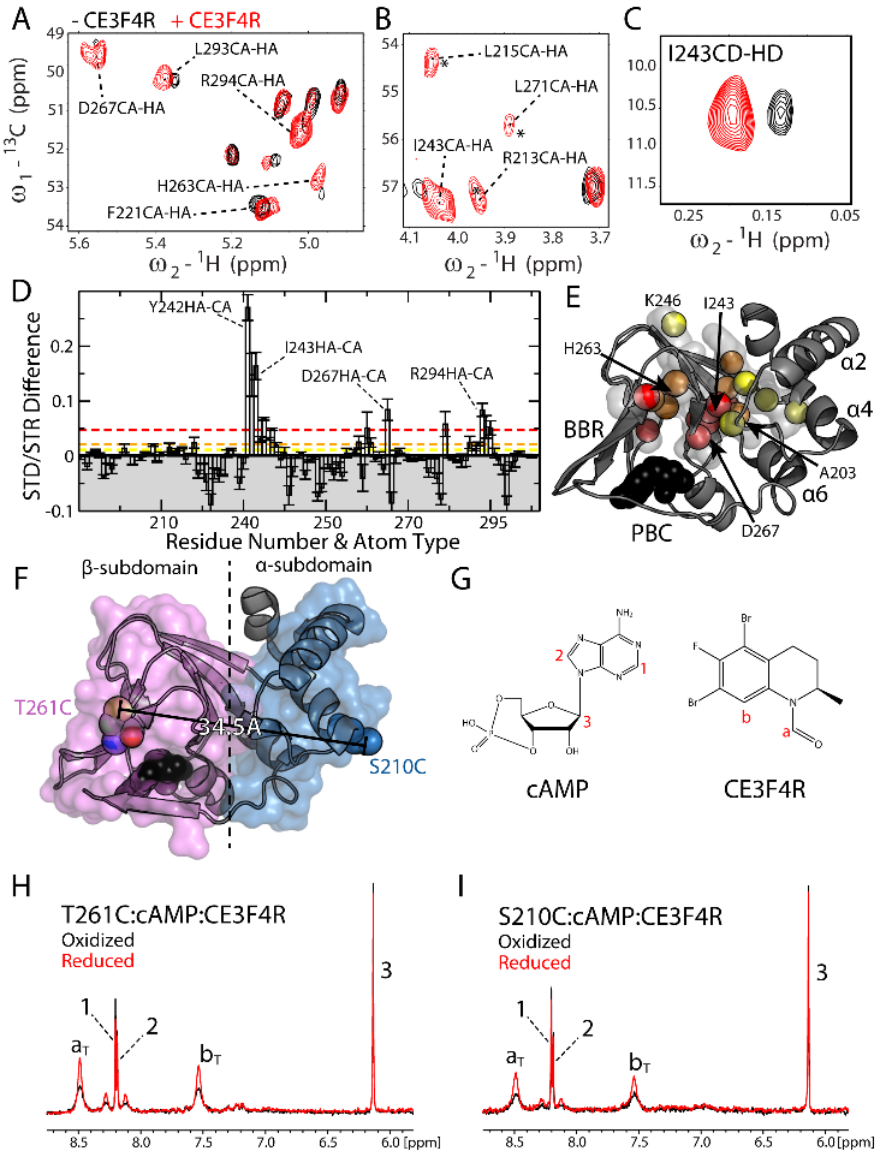

Figure 6. CE3F4R binds the EPAC1 CBD at the interface of the α- and β-subdomains. **(A-C)** Representative peaks from 2D STD spectra that show increased signal in the presence of CE3F4R (red) in comparison to cAMP alone (black). Black asterisks indicate the position of cross-peaks in the reference spectra that are missing in the STD spectra. **D)** The STD/STR differences between the samples with and without CE3F4R. The red dashed line indicates the average + standard deviation, while the yellow and orange lines indicate the average + 2x and 3x the average STD/STR difference error, respectively. **E)** Map of the CH bonds with STD/STR differences greater than the cut-offs described for panel D (*i.e.* red spheres represent CH bonds greater than the red lines). **F)** Crystal structure of EPAC2 with mutation sites (S210C and T261C) engineered to link the spin label indicated as spheres. The blue and pink surfaces depict residues that are within 20 Å from the S210C and T261C spin labels, respectively. **G)** structures of cAMP and CE3F4R. The labels indicate the assigned peak in the spectra of panels (H, I). **H)** 1D spectra of spin labeled (SL) T261C EPAC1-CBD in complex with cAMP and CE3F4R. Spectra were acquired for both the oxidized (black) and a reference sample reduced with 5 mM DTT (red). **I)** similar to (H), but for SL-S210C EPAC1 CBD.

observed was ~20 Å. 1D ^{1}H-NMR spectra were then acquired for the MTSL-labeled mutants in complex with cAMP and CE3F4R (Fig. 6H, I). The CE3F4R peaks broadened significantly when added to both MTSL-labelled mutants and the broadening was reduced upon reduction with DTT (Fig. 6H, I, Table 1). These results suggest that EPAC-bound CE3F4R senses both spin labels. Our PRE-control experiments (Supporting Information) indicate that the line-broadening of CE3F4R caused by the spin-labels arises from specific interactions. Furthermore, the PRE-induced signal losses for CE3F4R are comparable in the T261C and S210C samples (Fig. 6H, I), suggesting that CE3F4R binds the EPAC1 CBD at a site that is approximately equidistant from the two spin-labels. This PRE-based finding is consistent with the CE3F4R binding pocket being located at the α/β-subdomain interface and confirms our independent CE3F4R binding site assessments based on the CHESPA (Fig. 3E-H) and STD-HSQC analyses (Fig. 6A-E). Furthermore, the 1D PRE data of Fig. 6H,I illustrate the importance of complementing protein-based NMR experiments with ligand-based NMR experiments. In the next section, we further characterize the EPAC1:CE3F4R interaction using ligand-based NMR experiments, such as transfer-NOESY and protein-to-ligand STDs.

Table 1. SLAPSTIC I_{ox}/I_{red} Ratios

Site[a]	S210C: cA+CE3F4R	T261C: cA+CE3F4R	S210C: CE3F4[c]
a_T	0.45 ± 0.04	0.49 ± 0.03	0.87 ± 0.01
a_C	0.50 ± 0.10	0.52 ± 0.06	0.87 ± 0.03
b_T	0.53 ± 0.05	0.52 ± 0.03	1.01 ± 0.03
b_C	n/a	0.52 ± 0.06	1.06 ± 0.09
e_T	0.57 ± 0.01	0.58 ± 0.01	0.86 ± 0.01

[a]Assignment abbreviations are as in the peak labels of Fig. 7A.

3.4.5 The EPAC1 CBD Binds Preferentially the Trans CE3F4R Conformer.

As a first step towards characterizing the CE3F4R-EPAC1 interaction through ligand-based NMR, we acquired the 1D ^{1}H NMR spectrum of CE3F4R in the absence of EPAC1 (Fig. 7A). Fig. 7A shows that the peaks of CE3F4R are doubled into two sets with intensity ratios of ~3:1 due to the slow rotation around the N-formyl group[57], as confirmed by the analysis of ROESY (Fig. 7B) and NOESY spectra, also acquired in the absence of the EPAC1 CBD (Fig. 7C, orange cross-peaks). The presence of ROE cross- peaks with the same sign as the diagonal (Fig. 7B) indicates that the

additional peaks arise from slow chemical exchange between two conformations. The NOE cross-peaks between the formyl proton and the H8 of CE3F4, which are observed only for the major conformation (Fig. 7C, orange cross-peaks), are consistent with the major and minor populations reflecting the *trans* and *cis* conformations of CE3F4R, respectively, in the absence of EPAC1:cAMP[57]. Under our experimental conditions, the major (*trans*) conformer of CE3F4R preferentially interacts with EPAC1, as indicated by the comparison of the 1D ^{1}H NMR spectra of CE3F4R in the absence and presence of the cAMP-bound EPAC1 CBD (Fig. 7D). Figure 7D shows that upon addition of EPAC1 the peak positions for the minor population are only slightly affected, while the peaks of the major conformation exhibit more appreciable chemical shift changes and line-broadening, suggesting that the major (*trans*) conformation of CE3F4R is the preferred binding partner for the EPAC1 CBD. This result is confirmed by the transfer-NOE cross-peaks observed in the presence of the EPAC1 CBD with the same sign as the diagonal between the H8 and formyl protons for the major CE3F4R population (Fig. 7C, black spectra). However, residual interactions of EPAC1 with the *cis* CE3F4R conformer cannot be ruled out. Subtle, yet detectable, shifts and line-broadening are observed also for the minor CE3F4R peaks (Fig. 7D, top traces) and saturation transfer is detected for both major and minor CE3F4R signals (Fig. 7D, bottom traces). While it is possible that this saturation pattern is due to chemical exchange relay between the major and minor conformations, the STD signals are also consistent with the minor conformer still exhibiting residual interactions with EPAC1. Hence, EPAC1 may interact with both trans and cis conformers of CE3F4R, but preferentially binds the trans conformation.

The STD spectra of Fig. 7D also reveal that residual CE3F4R saturation occurs even in the absence of EPAC (Fig. 7D, purple spectrum). A possible explanation is saturation leak through caused by a CE3F4R methyl resonating close to the saturation frequency (~0.4 ppm difference), but saturation transfer is inefficient for an unbound monomeric small molecule. However, hydrophobic drug leads, such as CE3F4R, may form soluble aggregates in aqueous solutions[42,45,58–61]. Under such circumstances, saturation transfer between protons in a small molecule may be observed, explaining why residual STD remains for CE3F4R in the absence of EPAC. The possibility of such CE3F4R aggregates poses a limitation on the maximum CE3F4R concentration and on our ability to fully saturate the EPAC1 binding site with CE3F4R. Nevertheless, addition of EPAC1 still results in a significant enhancement of the STD signal (Fig. 7D; Table 2), indicating that the observed STD signal primarily reflects interactions between EPAC and CE3F4R. Furthermore, the STD/STR ratios of CE3F4R are significantly higher for the cAMP-bound state compared to the apo EPAC1

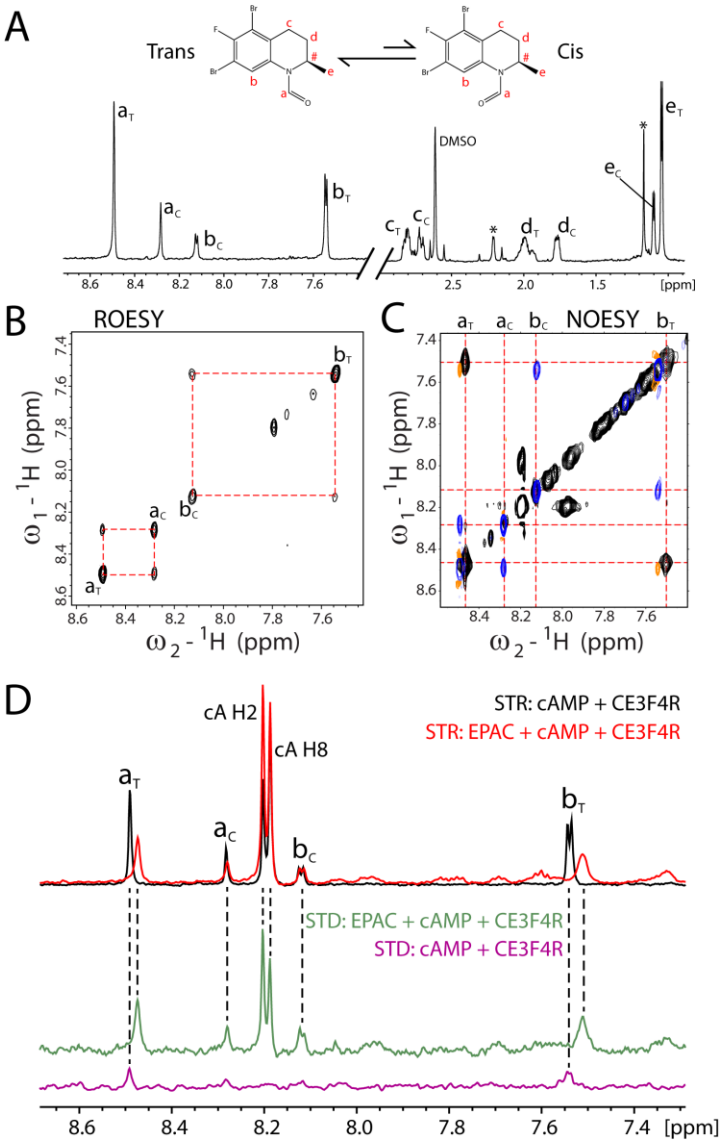

Figure 7. EPAC1 specifically interacts with the major conformer of CE3F4R. **A)** 1D Watergate spectra of free CE3F4R in sodium phosphate buffer. The letter code is used to indicate the assignment of CE3F4 signals. The resonance of the proton adjacent to the methyl group (#) partially overlaps with H_2O and is therefore eliminated by the Watergate suppression scheme. The major and minor populations of each proton are denoted with a subscript T or C, respectively. The asterisk marks an impurity. No appreciable line-broadening was observed for these peaks upon addition of EPAC. **B)** Off-resonance ROESY spectrum of free CE3F4R. The positive phase of the ROE cross-peaks relative to the diagonal indicates that they arise from chemical exchange. **C)** Transfer NOESY spectra of CE3F4R in the absence (blue & orange) or in the presence of the EPAC1 CBD and cAMP (black). In the former spectra, the blue represents positive phases and the orange represents negative phases for cross-peaks relative to diagonal peaks. The red dashed lines indicate the position and assignment of the CE3F4R peaks in the presence of the EPAC1 CBD. **D)** Saturation transfer difference (STD) of cAMP and CE3F4R in the absence (purple) and presence (green) of the EPAC1 CBD. The reference spectra (STR) are included for comparison.

sample (Table 2), confirming that CE3F4R is more selective for the cAMP-bound conformation, as expected for an uncompetitive inhibitor[3].

Table 2. Relative STD/STR Ratios of cAMP and CE3F4R Peaks

Sample			cAMP			CE3F4R	
EPAC	cAMP	CE3F4R	STD/STR H2	STD/STR H8	STD/STR H1'	STD/STR (a_T)[a]	STD/STR (b_T)[a]
–	+	+	N/A	N/A	N/A	0.48 ± 0.07	0.5 ± 0.1
+	+	–	1.00 ± 0.06[c]	0.84 ± 0.05	0.84 ± 0.04	N/A	N/A
+	–	+	N/A	N/A	N/A	0.7 ± 0.1 [b]	0.9 ± 0.2 [b]
+	+	+	1.02 ± 0.06	0.86 ± 0.05	0.86 ± 0.04	1.9 ± 0.2	2.1 ± 0.2

[a] Assignment abbreviations are as in the peak labels of Fig. 7A. [b] Acquired at 850 MHz.
[c] The STD/STD ratios were normalized to the STD/STR ratio of this cell.

3.4.6 Molecular Basis of the CE3F4R Selectivity for the EPAC1 *vs.* EPAC2 Isoform.

In order to elucidate the molecular basis for the EPAC1 vs. EPAC2 selectivity of CE3F4R, we looked for sequence differences between the two isoforms in the regions surrounding the CE3F4R binding pocket identified above. We found two significant amino acid differences between EPAC1 and EPAC2 in regions surrounding CE3F4R. The first is W245 in the β3 strand (L380 in EPAC2) and the second is Q270 in the PBC (K405 in EPAC2) (Fig. 1D, 7A, S6A). In the cAMP-bound crystal structure of EPAC2, K405 stabilizes the hinge helix in the active (in) conformation by forming a salt bridge with E443 (Fig. 8A). Although E443 of EPAC2 is conserved in EPAC1 (E308, Fig. 1D), this ionic interaction is expected to be weaker or non- existent in EPAC1 *vs.* EPAC2, given the glutamine in place of the lysine. The EPAC2 *vs.* EPAC1 stabilization of the hinge helix in the 'in' conformation is expected to decrease the population of the 'PBC in/hinge out' mixed intermediate selectively targeted by CE3F4R. Hence, we hypothesized that the Q270K mutation would decrease the affinity for CE3F4R and explain the EPAC1 vs. EPAC2 selectivity of the inhibitor.

To test this hypothesis, we created the Q270K EPAC1 mutant. As expected, the HSQC chemical shifts of the apo and cAMP-bound Q270K EPAC1 mutant are quite similar to those of WT EPAC1 with the exception of the hinge helix region and the adjacent α4 helix (Fig. 8B, C). For the apo state, the Q270K CHESPA analysis shows that the mutation causes a minor overall shift towards activation (Fig. S2). The effect is most predominant in the hinge helix, for which the average fractional activation (X) is ~40 % higher than WT (Fig. S2B). In the cAMP-bound state,

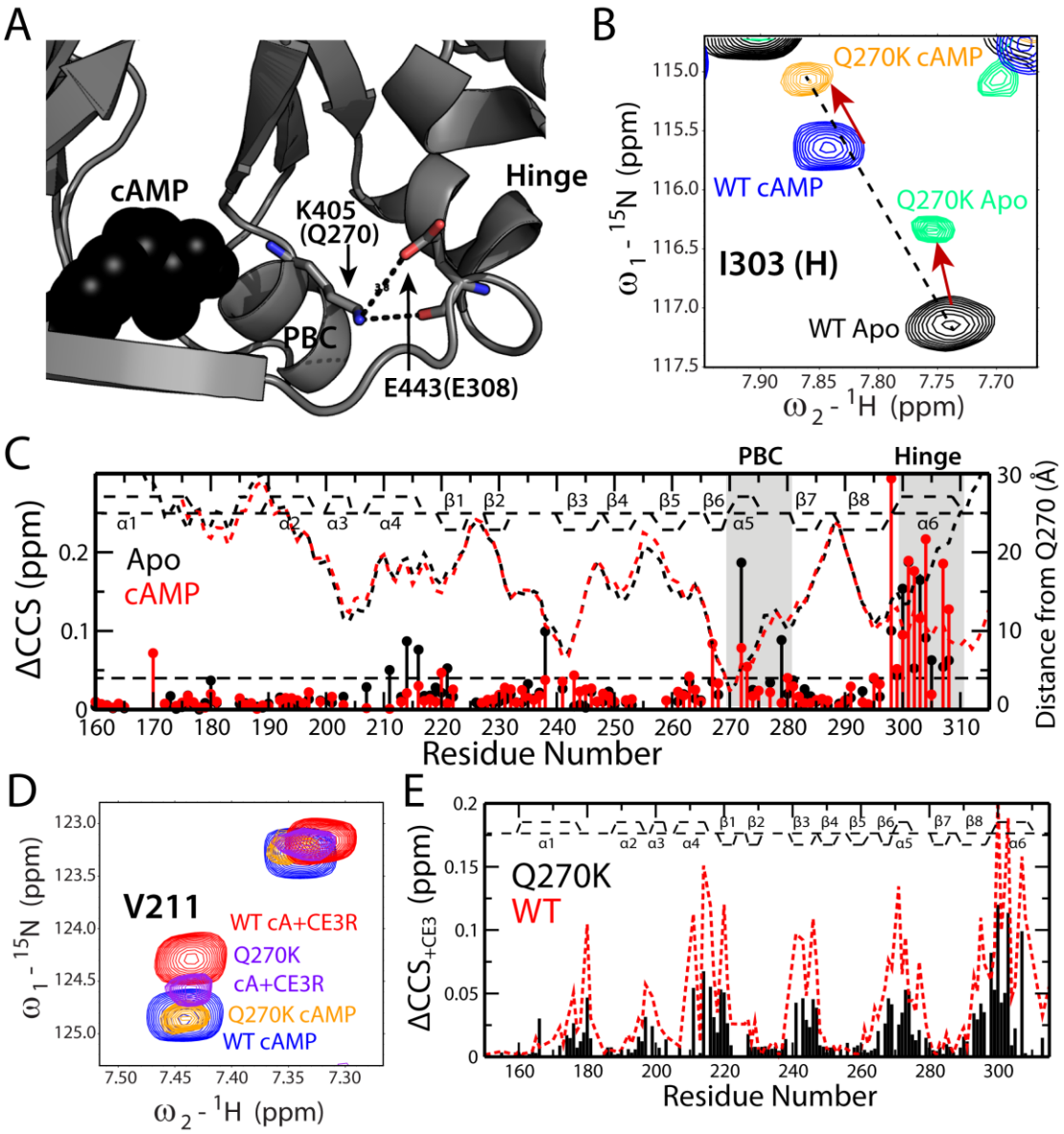

Figure 8. Q270 in the PBC confers to CE3F4R selectivity for EPAC1 *vs*. EPAC2. **A)** Interaction between K405 in the PBC (Q270 in EPAC1) and E443 in the hinge (E308 in EPAC1), as observed in the cAMP-bound X-ray crystal structure of EPAC2. cAMP is shown as black spheres. **B)** Representative hinge residue from HSQC spectra showing the increased stabilization of the "hinge in" conformation for Q270K relative to WT. **C)** The CCS differences between WT and Q270K EPAC for both apo (black) and cAMP-bound (red) states. The black and red dashed lines indicate the Cβ to N distances in Å from Q270 for the apo and cAMP-bound crystal structures, respectively. **D)** Representative HSQC cross-peaks for cAMP-bound WT and Q270K EPAC in the absence and presence of CE3F4R. **E)** The CCS differences between CE3F4R bound and unbound EPAC1 CBD for both WT and Q270K in the presence of cAMP. The EPAC, cAMP and CE3F4R concentrations were similar for both WT and Q270K.

the hinge helix is also more active than WT (Fig. S2B). These results confirm the notion that a salt bridge between K270 and E308 stabilizes the hinge helix in the active 'in' conformation. As expected, the addition of CE3F4R to the cAMP-bound Q270K mutant resulted in significantly reduced chemical shift changes compared to WT (Fig. 8D, E), indicating weaker binding. This result is consistent with the anticipated role of the "hinge in" conformation in CE3F4R recognition, although direct interactions between Q270 and CE3F4R cannot be ruled out. We also examined the W245A EPAC1 mutant (Fig. S3) and in this case the chemical shift changes induced by CE3F4R binding to cAMP-bound EPAC1 (149-318) are comparable or higher to those measured for WT (Fig. S3E), indicating that W245A EPAC1 retains the ability to bind CE3F4R. This is consistent with the observation that there is no appreciable stabilization of the "hinge in" state relative to WT and no significant W245A *vs.* WT changes were detected for the PBC (Fig. S3B, C). Overall, these results confirm our hypothesis that Q270, unlike W245, is a major determinant of the EPAC1 *vs.* EPAC2 selectivity of CE3F4R, as the replacement of K with Q disengages the hinge helix from the 'in' conformation.

3.5 Discussion

3.5.1 A Model for the Mechanism of Action of CE3F4R as EPAC1-Selective Uncompetitive Inhibitor.

The results of our NMR analyses on CE3F4R are summarized in Figure 9. Figure 9 illustrates a model that rationalizes the non-classical uncompetitive EPAC1 inhibition by CE3F4R in the context of the thermodynamic cycle for the cAMP-dependent allosteric regulation of EPAC1. A central feature of the proposed mechanism of action is that CE3F4R preferentially targets a mixed cAMP-bound intermediate of EPAC1, which is distinct from the structures of both apo and cAMP-bound (holo) EPAC1, but combines elements from both apo and holo conformations (Fig. 9). In the mixed intermediate targeted by CE3F4R, the PBC, which recognizes the cyclic phosphate, is maintained in the 'in' orientation by cAMP, similarly to holo EPAC1, while the hinge C-terminal helix, which regulates the transition from closed (inactive) to open (active) topologies, shifts from the 'in' of holo EPAC1 to the 'out' orientation of apo EPAC1, as indicated by our comparative chemical shift analyses (Fig. 2 - 5). The preferential targeting of the mixed PBC 'in'/hinge 'out' EPAC1 intermediate by CE3F4R explains why it is uncompetitive with respect to cAMP, but not with respect to Rp-cAMPS, which is an effective allosteric inhibitor of EPAC1 that binds at the same site as cAMP[62]. Unlike cAMP, Rp-cAMPS locks the PBC in the 'out' orientation, thus

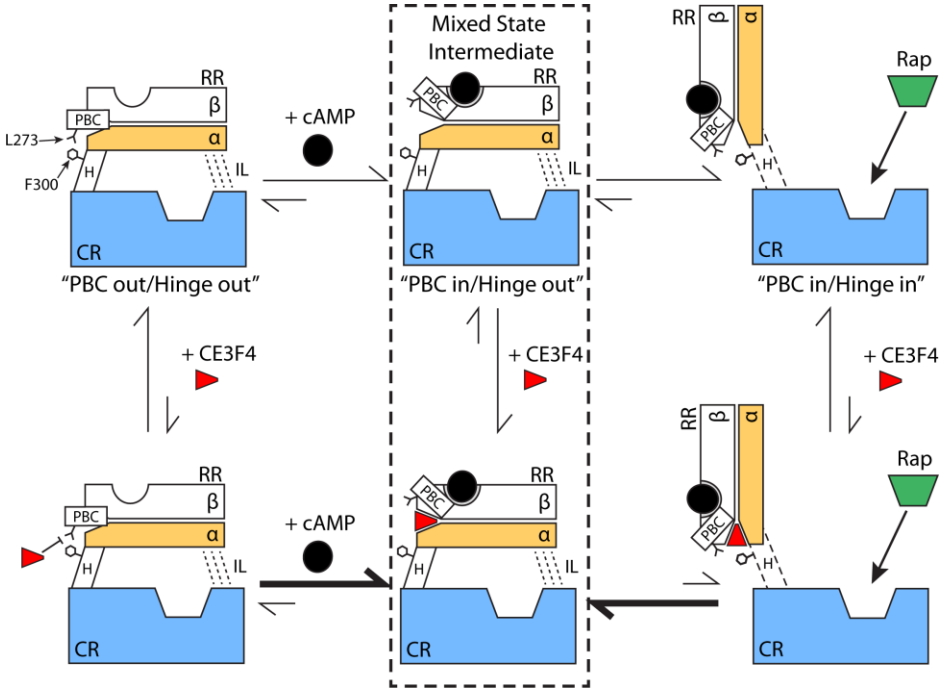

Figure 9. Proposed mechanism of action for the CE3F4R uncompetitive EPAC1 inhibitor. In the apo state, the regulatory region (RR), which includes both α- (white) and β- (orange) subdomains, occludes the Rap binding site in the catalytic region (CR). This position of the RR is stabilized by the "hinge helix (H) out" conformation and by the ionic latch (IL) interactions with the CR. The "hinge out" conformation is a result of a steric clash between L273 in the PBC helix (P) and F300 in the hinge helix. When cAMP binds, the PBC helix rotates in, which relieves the steric clash between L273 and F300 and allows the hinge helix to rotate to the 'in' orientation and partially unfold (dashed lines in right panels). This causes the RR to be displaced from the CR, which becomes then accessible to Rap substrates. CE3F4R hinders this closed-to-open transition by binding to and stabilizing the "PBC in/Hinge out" mixed intermediate (dashed rectangle), which traps the RR in a position that blocks the Rap binding site.

destabilizing the mixed intermediate selectivity targeted by CE3F4R. Hence, the proposed mechanism (Fig. 9) recapitulates the counter-intuitive lack of synergy between two allosteric inhibitors of EPAC1 that bind at distinct bind sites, *i.e.* CE3F4R and Rp-cAMPS.

The ability of CE3F4R to selectively recognize the PBC "in/hinge out" holo EPAC1 intermediate is also an indication of the specificity of the EPAC1-CE3F4R interactions. This specificity is further confirmed by the preferential recognition by EPAC1 of CE3F4R vs. other CE3F4 derivatives (Fig. 3H) and by the selectivity for the trans vs. cis isomer of CE3F4R, as illustrated by our ligand-based NMR analyses (Fig. 7). However, Sonawane et al. recently proposed that EPAC preferentially interacts with the minor rotamer (cis) of CE3F4[57], based on the increased IC_{50} value of a thioformyl CE3F4R derivative that disrupts rotamer exchange and stabilizes the major (trans) conformation[57]. A possible explanation to reconcile this apparent discrepancy is that the reduced potency of the thioformyl CE3F4R derivative is simply caused by the presence of the

bulky sulfur atom. In fact, similar oxygen-to-sulfur isolobal replacements have been previously shown to reduce affinity even in the absence of ligand conformational changes[63,64].

Another unique attribute of the proposed mechanism of action (Fig. 9) is that the uncompetitive inhibitor CE3F4R does not bind EPAC1 in contact with cAMP, but rather at a distinct binding site located at the interface of the α- and β-subdomains of the EPAC1 CBD. The location of the CE3F4R binding site between the EPAC1 CBD subdomains is consistently supported by multiple independent analyses, including our STD-HSQC (Fig. 6D, E), triangulated PREs (Fig. 6F-I; Fig. S1) and CHESPA cos(θ) data (Fig. 3F, K). A notable feature of this proposed binding pocket is that it is not visibly apparent in the apo and cAMP-bound X-ray crystal structures[22,27], which indicate that the binding site may arise from the transient sampling of the "PBC in/hinge out" intermediate[65]. Regardless, the interaction of CE3F4R with the α/β subdomains interface leads to two main implications. First, it offers a viable explanation as to why the hinge helix shifts from the 'in' to the 'out' orientation upon CE3F4R binding. The binding of CE3F4R hinders the recruitment of the C-terminal hinge helix into the 'in' orientation towards the interface between the CBD's α- and β-subdomains. By targeting a similar inter-subdomain interface, CE3F4R may outcompete the hinge helix and stabilizes the hinge 'out' orientation. Meanwhile, the PBC remains in the 'in' state stabilized by the bound cAMP. Hence, the location of the CE3F4R binding site at the inter-subdomain interface is in the full agreement with the notion that the CE3F4R uncompetitive inhibitor preferentially targets the "PBC in/hinge out" intermediate, as discussed above. The second notable implication of CE3F4R binding at the α/β-subdomain interface is the identification of a key residue that enables isoform selective inhibition, *i.e*. Q270. Q270 is unique to EPAC1 and located at the PBC N-terminus. In EPAC2, the residue at this position is a lysine that recruits the hinge helix in an 'in' orientation through a salt-bridge, as supported by our comparative Q270K mutational analyses (Fig. 8; Fig. S2). The Q270K-induced destabilization of the hinge out orientation explains the EPAC1 *vs*. EPAC2 selectivity of CE3F4R. Interestingly, this glutamine at the PBC N-terminus of EPAC1 is replaced by a glutamate in both CBDs of PKA, where it stabilizes the C-terminal CBD lids in a 'in' orientation, offering a possible rationale for the EPAC1 *vs*. PKA selectivity of CE3F4R[3]. Hence, the proposed mechanism of action (Fig. 9) provides a viable explanation for the remarkable selectivity of CE3F4R for EPAC1 *vs*. EPAC2 and PKA. The EPAC *vs*. PKA and isoform selectivity of CE3F4R makes CE3F4R a highly promising drug lead. Furthermore, the uncompetitive nature of the EPAC1-selective inhibition with respect to cAMP, ensures that the inhibitory function of CE3F4R is potentiated when and

where the subcellular concentration of cAMP is locally increased in response to tight spatio-temporal control, thus further increasing the specificity of CE3F4R.

3.6 Conclusion

We have shown how a combination of solution NMR methods is effective in elucidating the mechanism of allosteric inhibition by an hydrophobic drug lead that precludes the possibility of forming a saturated protein:ligand complex amenable to traditional structure determination methods. Comparative chemical shift projection analyses were instrumental in identifying the activation intermediate targeted by the drug lead and obtaining an initial assessment of the binding site, which was then independently confirmed through PRE triangulation and ligand-to-protein STD experiments, while protein-to-ligand STD together with TrNOESY, ROESY and line-broadening analyses allowed an independent ligand-based assessment of the interaction.

The resulting mechanism of action was validated through mutations as well as cyclic nucleotide and CE3F4 analogs, revealing that the uncompetitive inhibitor, CE3F4R, binds at the CBD α- and β-subdomain interface and selectively stabilizes a mixed intermediate that is distinct from the fully apo and holo structures, but combines elements of both. The proposed model rationalizes the non-classical uncompetitive inhibition of EPAC1 by CE3F4R and provides a viable molecular basis for its isoform and EPAC *vs.* PKA selectivity.

The mechanism of action of CE3F4R elucidated here explains how uncompetitive inhibition occurs with respect to allosteric effectors as opposed to substrates and is anticipated to serve as a foundation for designing a new generation of EPAC1 specific inhibitors, which are potential therapeutics for treating tumor invasion, inflammation and heart failure. In addition, the experimental design utilized in this study is transferable to mechanistic investigations of other lipophilic drug leads.

3.7 Experimental Section

3.7.1 Sample Preparation.

The CBD of human EPAC1 spanning residues 149-318 was expressed and purified according to previously established protocols[26,62,66]. The final purified protein was dialyzed in NMR buffer (50mM Tris, pH 7.6, 50mM NaCl, 2mM EDTA, 2mM EGTA, 1mM DTT and 0.02% NaN_3) and 5% D_2O was added to all NMR samples before data acquisition. For ligand-based NMR

experiments, EPAC was exchanged into sodium phosphate buffer (20mM Na_2HPO_4, pH 7.6, 50mM NaCl in >99.8% D_2O) using either PD10 columns (GE Healthcare) or Zeba 7K MWCO desalting columns (Thermo Scientific). All mutations were created using a variation of the Quikchange site-directed mutagenesis procedure as described previously[50,51] and verified by Sanger sequencing (MOBIX). The concentration of EPAC was determined by A_{280} measurements with a theoretical extinction coefficient derived from the amino acid sequence by EXPASY ProtParam[67]. The lyophilized CE3F4 compounds were dissolved in 99.9 % d_6-DMSO (Sigma) to stock concentrations of 20-35 mM. Proton NMR spectra were then acquired to verify the integrity and purity of all compounds. CE3F4R peak assignment was obtained mainly from 1D chemical shifts, peak multiplicity and integration and confirmed with 2D-COSY experiments. cAMP was purchased from Sigma and all other cyclic nucleotides were purchased from Biolog Life Sciences.

3.7.2 NMR Data Acquisition and Analysis.

NMR experiments were acquired on a Bruker Avance 700 MHz spectrometer equipped with a 5 mm TCI cryo-probe or on a Bruker Avance 850 spectrometer equipped with a TXI probe. All protein-based experiments were acquired at 306 K, while ligand-based experiments were acquired at 298 K, unless otherwise specified. One dimensional experiments were processed directly in Topspin using an exponential multiplication window function with line broadening of 3 Hz. Multi-dimensional NMR experiments were processed with NMRpipe[68] and analyzed in SPARKY[69]. For NH-HSQC spectra, all samples contained ^{15}N-acetyl glycine, which was used an internal reference to align spectra for precise chemical shift measurements. Chemical shift differences for samples with and without CE3F4R were measured using a constant DMSO concentration, which typically was kept at or below 1% v/v, since this was a range known to not cause significant chemical shift changes in EPAC[70]. Peaks were assigned by comparison with previously assigned EPAC1 spectra[26,62,66]. Aliphatic CH resonances were assigned using a combination of newly acquired hCCH and HCcH-TOCSY and TOCSY-HSQC(NH) 3D spectra and previously acquired CBCACONH and HBHACONH data[28,66,71]. Additional NMR experimental procedures can be found in the supporting information[72–78].

3.8 Acknowledgements

We would like to dedicate this work in memory of our friend and colleague, Greg Bahun. We also thank Dr. M. Akimoto, Dr. B. VanSchouwen, Dr. X. C. Cheng (UTHealth) and R. Ahmed for helpful discussions.

3.9 References

(1) Herries, D. G. Principles of Enzyme Kinetics. *Biochemical Education* **1976**, *4*, 36.

(2) Cornish-Bowden, A. Why is uncompetitive inhibition so rare? A possible explanation, with implications for the design of drugs and pesticides. *FEBS letters* **1986**, *203*, 3–6.

(3) Courilleau, D.; Bisserier, M.; Jullian, J.-C.; Lucas, A.; Bouyssou, P.; Fischmeister, R.; Blondeau, J.-P.; Lezoualc'h, F. Identification of a tetrahydroquinoline analog as a pharmacological inhibitor of the cAMP-binding protein Epac. *The Journal of biological chemistry* **2012**, *287*, 44192–44202.

(4) Courilleau, D.; Bouyssou, P.; Fischmeister, R.; Lezoualc'h, F.; Blondeau, J.-P. The (R)-enantiomer of CE3F4 is a preferential inhibitor of human exchange protein directly activated by cyclic AMP isoform 1 (Epac1). *Biochemical and Biophysical Research Communications* **2013**, *440*, 443–448.

(5) Schmidt, M.; Dekker, F. J.; Maarsingh, H. Exchange protein directly activated by cAMP (epac): A multidomain cAMP mediator in the regulation of diverse biological functions. *Pharmacological reviews* **2013**, *65*, 670–709.

(6) Rooij, J. de; Zwartkruis, F. J.; Verheijen, M. H.; Cool, R. H.; Nijman, S. M.; Wittinghofer, A.; Bos, J. L. Epac is a Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP. *Nature* **1998**, *396*, 474–477.

(7) Kawasaki, H.; Springett, G. M.; Mochizuki, N.; Toki, S.; Nakaya, M.; Matsuda, M.; Housman, D. E.; Graybiel, A. M. A Family of cAMP-Binding Proteins That Directly Activate Rap1. *Science* **1998**, *282*, 2275–2279.

(8) Cheng, X.; Ji, Z.; Tsalkova, T.; Mei, F. Epac and PKA: A tale of two intracellular cAMP receptors. *Acta biochimica et biophysica Sinica* **2008**, *40*, 651–662.

(9) Bos, J. L. Epac proteins: Multi-purpose cAMP targets. *Trends in biochemical sciences* **2006**, *31*, 680–686.

(10) Robichaux, W. G.; Cheng, X. Intracellular cAMP Sensor EPAC: Physiology, Pathophysiology, and Therapeutics Development. *Physiological reviews* **2018**, *98*, 919–1053.

(11) Almahariq, M.; Tsalkova, T.; Mei, F. C.; Chen, H.; Zhou, J.; Sastry, S. K.; Schwede, F.; Cheng, X. A novel EPAC-specific inhibitor suppresses pancreatic cancer cell migration and invasion. *Molecular pharmacology* **2013**, *83*, 122–128.

(12) Wang, X.; Luo, C.; Cheng, X.; Lu, M. Lithium and an EPAC-specific inhibitor ESI-09 synergistically suppress pancreatic cancer cell proliferation and survival. *Acta biochimica et biophysica Sinica* **2017**, *49*, 573–580.

(13) Kumar, N.; Gupta, S.; Dabral, S.; Singh, S.; Sehrawat, S. Role of exchange protein directly activated by cAMP (EPAC1) in breast cancer cell migration and apoptosis. *Molecular and cellular biochemistry* **2017**, *430*, 115–125.

(14) Wang, H.; Robichaux, W. G.; Wang, Z.; Mei, F. C.; Cai, M.; Du, G.; Chen, J.; Cheng, X. Inhibition of Epac1 suppresses mitochondrial fission and reduces neointima formation induced by vascular injury. *Scientific reports* **2016**, *6*, 36552.

(15) Singhmar, P.; Huo, X.; Eijkelkamp, N.; Berciano, S. R.; Baameur, F.; Mei, F. C.; Zhu, Y.; Cheng, X.; Hawke, D.; Mayor, F. *et al.* Critical role for Epac1 in inflammatory pain controlled by

GRK2-mediated phosphorylation of Epac1. *Proceedings of the National Academy of Sciences of the United States of America* **2016**, *113*, 3036–3041.

(16) Hu, Y.; Robichaux, W. G.; Mei, F. C.; Kim, E. R.; Wang, H.; Tong, Q.; Jin, J.; Xu, M.; Chen, J.; Cheng, X. Role of Exchange Protein Directly Activated by Cyclic AMP Isoform 1 in Energy Homeostasis: Regulation of Leptin Expression and Secretion in White Adipose Tissue. *Molecular and cellular biology* **2016**, *36*, 2440–2450.

(17) Yan, J.; Mei, F. C.; Cheng, H.; Lao, D. H.; Hu, Y.; Wei, J.; Patrikeev, I.; Hao, D.; Stutz, S. J.; Dineley, K. T. *et al.* Enhanced leptin sensitivity, reduced adiposity, and improved glucose homeostasis in mice lacking exchange protein directly activated by cyclic AMP isoform 1. *Molecular and cellular biology* **2013**, *33*, 918–926.

(18) Gong, B.; Shelite, T.; Mei, F. C.; Ha, T.; Hu, Y.; Xu, G.; Chang, Q.; Wakamiya, M.; Ksiazek, T. G.; Boor, P. J. *et al.* Exchange protein directly activated by cAMP plays a critical role in bacterial invasion during fatal rickettsioses. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110*, 19615–19620.

(19) Tao, X.; Mei, F.; Agrawal, A.; Peters, C. J.; Ksiazek, T. G.; Cheng, X.; Tseng, C.-T. K. Blocking of exchange proteins directly activated by cAMP leads to reduced replication of Middle East respiratory syndrome coronavirus. *Journal of virology* **2014**, *88*, 3902–3910.

(20) Dawn, A.; Singh, S.; More, K. R.; Siddiqui, F. A.; Pachikara, N.; Ramdani, G.; Langsley, G.; Chitnis, C. E. The central role of cAMP in regulating Plasmodium falciparum merozoite invasion of human erythrocytes. *PLoS pathogens* **2014**, *10*, e1004520.

(21) Rehmann, H.; Arias-Palomo, E.; Hadders, M. A.; Schwede, F.; Llorca, O.; Bos, J. L. Structure of Epac2 in complex with a cyclic AMP analogue and RAP1B. *Nature* **2008**, *455*, 124–127.

(22) Rehmann, H.; Das, J.; Knipscheer, P.; Wittinghofer, A.; Bos, J. L. Structure of the cyclic-AMP-responsive exchange factor Epac2 in its auto-inhibited state. *Nature* **2006**, *439*, 625–628.

(23) Niimura, M.; Miki, T.; Shibasaki, T.; Fujimoto, W.; Iwanaga, T.; Seino, S. Critical role of the N-terminal cyclic AMP-binding domain of Epac2 in its subcellular localization and function. *Journal of cellular physiology* **2009**, *219*, 652–658.

(24) Selvaratnam, R.; Mazhab-Jafari, M. T.; Das, R.; Melacini, G. The auto-inhibitory role of the EPAC hinge helix as mapped by NMR. *PloS one* **2012**, *7*, e48707.

(25) VanSchouwen, B.; Selvaratnam, R.; Fogolari, F.; Melacini, G. Role of dynamics in the autoinhibition and activation of the exchange protein directly activated by cyclic AMP (EPAC). *The Journal of biological chemistry* **2011**, *286*, 42655–42669.

(26) Mazhab-Jafari, M. T.; Das, R.; Fotheringham, S. A.; Sildas, S.; Chowdhury, S.; Melacini, G. Understanding cAMP-dependent allostery by NMR spectroscopy: Comparative analysis of the EPAC1 cAMP-binding domain in its apo and cAMP-bound states. *Journal of the American Chemical Society* **2007**, *129*, 14482–14492.

(27) Rehmann, H.; Prakash, B.; Wolf, E.; Rueppel, A.; Rooij, J. de; Bos, J. L.; Wittinghofer, A. Structure and regulation of the cAMP-binding domains of Epac2. *Nature structural biology* **2003**, *10*, 26–32.

(28) Selvaratnam, R.; VanSchouwen, B.; Fogolari, F.; Mazhab-Jafari, M. T.; Das, R.; Melacini, G. The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants. *Biophysical journal* **2012**, *102*, 630–639.

(29) Tsalkova, T.; Blumenthal, D. K.; Mei, F. C.; White, M. A.; Cheng, X. Mechanism of Epac activation: Structural and functional analyses of Epac2 hinge mutants with constitutive and reduced activities. *J. Biol. Chem.* **2009**, *284*, 23644–23651.

(30) Tsai, C.-J.; Nussinov, R. The molecular basis of targeting protein kinases in cancer therapeutics. *Seminars in cancer biology* **2013**, *23*, 235–242.

(31) Nussinov, R. Introduction to Protein Ensembles and Allostery. *Chemical reviews* **2016**, *116*, 6263–6266.

(32) Huang, W.; Nussinov, R.; Zhang, J. Computational Tools for Allosteric Drug Discovery: Site Identification and Focus Library Design. *Methods in molecular biology (Clifton, N.J.)* **2017**, *1529*, 439–446.

(33) Narayanan, C.; Bernard, D. N.; Bafna, K.; Gagné, D.; Chennubhotla, C. S.; Doucet, N.; Agarwal, P. K. Conservation of Dynamics Associated with Biological Function in an Enzyme Superfamily. *Structure (London, England : 1993)* **2018**, *26*, 426-436.e3.

(34) Narayanan, C.; Gagné, D.; Reynolds, K. A.; Doucet, N. Conserved amino acid networks modulate discrete functional properties in an enzyme superfamily. *Scientific reports* **2017**, *7*, 3207.

(35) Ha, K. N.; Traaseth, N. J.; Verardi, R.; Zamoon, J.; Cembran, A.; Karim, C. B.; Thomas, D. D.; Veglia, G. Controlling the inhibition of the sarcoplasmic Ca2+-ATPase by tuning phospholamban structural dynamics. *J. Biol. Chem.* **2007**, *282*, 37205–37214.

(36) Traaseth, N. J.; Ha, K. N.; Verardi, R.; Shi, L.; Buffy, J. J.; Masterson, L. R.; Veglia, G. Structural and dynamic basis of phospholamban and sarcolipin inhibition of Ca(2+)-ATPase. *Biochemistry* **2008**, *47*, 3–13.

(37) De, S.; Greenwood, A. I.; Rogals, M. J.; Kovrigin, E. L.; Lu, K. P.; Nicholson, L. K. 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* **2012**, *51*, 8583–8596.

(38) Kukic, P.; Alvin Leung, H. T.; Bemporad, F.; Aprile, F. A.; Kumita, J. R.; Simone, A. de; Camilloni, C.; Vendruscolo, M. Structure and dynamics of the integrin LFA-1 I-domain in the inactive state underlie its inside-out/outside-in signaling and allosteric mechanisms. *Structure (London, England : 1993)* **2015**, *23*, 745–753.

(39) Simone, A. de; Richter, B.; Salvatella, X.; Vendruscolo, M. Toward an accurate determination of free energy landscapes in solution states of proteins. *Journal of the American Chemical Society* **2009**, *131*, 3810–3811.

(40) Yu, C. H.; Lee, W.; Nokhrin, S.; Dmitriev, O. Y. The Structure of Metal Binding Domain 1 of the Copper Transporter ATP7B Reveals Mechanism of a Singular Wilson Disease Mutation. *Scientific reports* **2018**, *8*, 581.

(41) Boulton, S.; Melacini, G. Advances in NMR Methods To Map Allosteric Sites: From Models to Translation. *Chemical reviews* **2016**, *116*, 6267–6304.

(42) McGovern, S. L.; Caselli, E.; Grigorieff, N.; Shoichet, B. K. A Common Mechanism Underlying Promiscuous Inhibitors from Virtual and High-Throughput Screening. *J. Med. Chem.* **2002**, *45*, 1712–1722.

(43) McGovern, S. L.; Helfand, B. T.; Feng, B.; Shoichet, B. K. A specific mechanism of nonspecific inhibition. *J. Med. Chem.* **2003**, *46*, 4265–4272.

(44) Feng, B. Y.; Shelat, A.; Doman, T. N.; Guy, R. K.; Shoichet, B. K. High-throughput assays for promiscuous inhibitors. *Nature chemical biology* **2005**, *1*, 146–148.

(45) Feng, B. Y.; Simeonov, A.; Jadhav, A.; Babaoglu, K.; Inglese, J.; Shoichet, B. K.; Austin, C. P. A high-throughput screen for aggregation-based inhibition in a large compound library. *Journal of medicinal chemistry* **2007**, *50*, 2385–2390.

(46) Rehmann, H. Epac-inhibitors: Facts and artefacts. *Scientific reports* **2013**, *3*, 3032.

(47) Reibarkh, M.; Malia, T. J.; Hopkins, B. T.; Wagner, G. Identification of individual protein-ligand NOEs in the limit of intermediate exchange. *Journal of biomolecular NMR* **2006**, *36*, 1–11.

(48) Ruschak, A. M.; Kay, L. E. Proteasome allostery as a population shift between interchanging conformers. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, E3454-62.

(49) Akimoto, M.; Zhang, Z.; Boulton, S.; Selvaratnam, R.; VanSchouwen, B.; Gloyd, M.; Accili, E. A.; Lange, O. F.; Melacini, G. A mechanism for the auto-inhibition of hyperpolarization-activated cyclic nucleotide-gated (HCN) channel opening and its relief by cAMP. *The Journal of biological chemistry* **2014**, *289*, 22205–22220.

(50) Boulton, S.; Akimoto, M.; Akbarizadeh, S.; Melacini, G. Free energy landscape remodeling of the cardiac pacemaker channel explains the molecular basis of familial sinus bradycardia. *The Journal of biological chemistry* **2017**, *292*, 6414–6428.

(51) Moleschi, K. J.; Akimoto, M.; Melacini, G. Measurement of State-Specific Association Constants in Allosteric Sensors through Molecular Stapling and NMR. *Journal of the American Chemical Society* **2015**, *137*, 10777–10785.

(52) Dias, D. M.; Ciulli, A. NMR approaches in structure-based lead discovery: Recent developments and new frontiers for targeting multi-protein complexes. *Progress in biophysics and molecular biology* **2014**, *116*, 101–112.

(53) Jahnke, W.; Rüdisser, S.; Zurini, M. Spin label enhanced NMR screening. *Journal of the American Chemical Society* **2001**, *123*, 3149–3150.

(54) Jahnke, W. Spin labels as a tool to identify and characterize protein-ligand interactions by NMR spectroscopy. *Chembiochem : a European journal of chemical biology* **2002**, *3*, 167–173.

(55) Boulton, S.; Akimoto, M.; VanSchouwen, B.; Moleschi, K.; Selvaratnam, R.; Giri, R.; Melacini, G. Tapping the translation potential of cAMP signalling: Molecular basis for selectivity in cAMP agonism and antagonism as revealed by NMR. *Biochemical Society transactions* **2014**, *42*, 302–307.

(56) Badireddy, S.; Yunfeng, G.; Ritchie, M.; Akamine, P.; Wu, J.; Kim, C. W.; Taylor, S. S.; Qingsong, L.; Swaminathan, K.; Anand, G. S. Cyclic AMP analog blocks kinase activation by stabilizing inactive conformation: Conformational selection highlights a new concept in allosteric inhibitor design. *Molecular & cellular proteomics : MCP* **2011**, *10*, M110.004390.

(57) Sonawane, Y. A.; Zhu, Y.; Garrison, J. C.; Ezell, E. L.; Zahid, M.; Cheng, X.; Natarajan, A. Structure–Activity Relationship Studies with Tetrahydroquinoline Analogs as EPAC Inhibitors. *ACS Medicinal Chemistry Letters* **2017**, *8*, 1183–1187.

(58) Frenkel, Y. V.; Clark, A. D.; Das, K.; Wang, Y.-H.; Lewi, P. J.; Janssen, P. A. J.; Arnold, E. Concentration and pH dependent aggregation of hydrophobic drug molecules and relevance to oral bioavailability. *Journal of medicinal chemistry* **2005**, *48*, 1974–1983.

(59) Schreier, S.; Malheiros, S. V.P.; Paula, E. de. Surface active drugs: Self-association and interaction with membranes and surfactants. Physicochemical and biological aspects. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2000**, *1508*, 210–234.

(60) Seidler, J.; McGovern, S. L.; Doman, T. N.; Shoichet, B. K. Identification and prediction of promiscuous aggregating inhibitors among known drugs. *Journal of medicinal chemistry* **2003**, *46*, 4477–4486.

(61) Liu, H.-Y.; Wang, Z.; Regni, C.; Zou, X.; Tipton, P. A. Detailed kinetic studies of an aggregating inhibitor; inhibition of phosphomannomutase/phosphoglucomutase by disperse blue 56. *Biochemistry* **2004**, *43*, 8662–8669.

(62) Selvaratnam, R.; Chowdhury, S.; VanSchouwen, B.; Melacini, G. Mapping allostery through the covariance analysis of NMR chemical shifts. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108*, 6133–6138.

(63) Das, R.; Melacini, G. A model for agonism and antagonism in an ancient and ubiquitous cAMP-binding domain. *J. Biol. Chem.* **2007**, *282*, 581–593.

(64) Frey, P. A.; Sammons, R. D. Bond order and charge localization in nucleoside phosphorothioates. *Science* **1985**, *228*, 541–545.

(65) Simone, A. de; Aprile, F. A.; Dhulesia, A.; Dobson, C. M.; Vendruscolo, M. Structure of a low-population intermediate state in the release of an enzyme product. *eLife* **2015**, *4,* DOI: 10.7554/eLife.02777.

(66) Das, R.; Chowdhury, S.; Mazhab-Jafari, M. T.; Sildas, S.; Selvaratnam, R.; Melacini, G. Dynamically driven ligand selectivity in cyclic nucleotide binding domains. *The Journal of biological chemistry* **2009**, *284*, 23682–23696.

(67) Gasteiger, E.; Hoogland, C.; Gattoler, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. *Protein Identification and Analysis Tools on the ExPASy Server*; Humana Press, 2005.

(68) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *Journal of biomolecular NMR* **1995**, *6*, 277–293.

(69) Lee, W.; Tonelli, M.; Markley, J. L. NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy. *Bioinformatics (Oxford, England)* **2015**, *31*, 1325–1327.

(70) Zhu, Y.; Chen, H.; Boulton, S.; Mei, F.; Ye, N.; Melacini, G.; Zhou, J.; Cheng, X. Biochemical and pharmacological characterizations of ESI-09 based EPAC inhibitors: Defining the ESI-09 "therapeutic window". *Scientific reports* **2015**, *5*, 9344.

(71) Das, R.; Mazhab-Jafari, M. T.; Chowdhury, S.; Sildas, S.; Selvaratnam, R.; Melacini, G. Entropy-driven cAMP-dependent allosteric control of inhibitory interactions in exchange proteins directly activated by cAMP. *The Journal of biological chemistry* **2008**, *283*, 19691–19703.

(72) Akimoto, M.; Selvaratnam, R.; McNicholl, E. T.; Verma, G.; Taylor, S. S.; Melacini, G. Signaling through dynamic linkers as revealed by PKA. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110*, 14231–14236.

(73) Boulton, S.; Akimoto, M.; Selvaratnam, R.; Bashiri, A.; Melacini, G. A tool set to map allosteric networks through the NMR chemical shift covariance analysis. *Scientific reports* **2014**, *4*, 7306.

(74) Boulton, S.; Selvaratnam, R.; Ahmed, R.; Melacini, G. Implementation of the NMR CHEmical Shift Covariance Analysis (CHESCA): A Chemical Biologist's Approach to Allostery. *Methods in molecular biology (Clifton, N.J.)* **2018**, *1688*, 391–405.

(75) Byun, J. A.; Melacini, G. NMR Methods to Dissect the Molecular Mechanisms of Disease-Related Mutations (DRMs): Understanding How DRMs Remodel Functional Free Energy Landscapes. *Methods (San Diego, Calif.)* **2018,** DOI: 10.1016/j.ymeth.2018.05.018.

(76) VanSchouwen, B.; Ahmed, R.; Milojevic, J.; Melacini, G. Functional dynamics in cyclic nucleotide signaling and amyloid inhibition. *Biochimica et biophysica acta* **2017**, *1865*, 1529–1543.

(77) VanSchouwen, B.; Melacini, G. Cracking the allosteric code of NMR chemical shifts. *Proceedings of the National Academy of Sciences of the United States of America* **2016**, *113*, 9407–9409.

(78) VanSchouwen, B.; Selvaratnam, R.; Giri, R.; Lorenz, R.; Herberg, F. W.; Kim, C.; Melacini, G. Mechanism of cAMP Partial Agonism in Protein Kinase G (PKG). *The Journal of biological chemistry* **2015**, *290*, 28631–28641.

3.10 Supplementary Results

PRE Control Experiments. Interestingly, unlike the CE3F4R peaks, the cAMP signals did not significantly change upon reduction (Fig. 6H, I), although T261C is located within 10-15 Å of most cAMP protons. The absence of cAMP signal broadening is due to the very large (20-fold) excess of cAMP used in these experiments. In fact, when the molar excess of cAMP is reduced to ten-fold, the signals of cAMP become broadened for the T261C mutant, but not for S210C (Fig. S1D, E). The S210C site is at least 25 Å away from cAMP protons. Hence, the observation of cAMP signal broadening for the T261C mutant, but not the S210C mutant, corroborates the distance dependency of the spin-label effects and confirms that the PRE effect in the SLAPSTIC method complies with the expected maximum distance determined based on the HSQC analysis (Fig. S1A, B).

As a negative control to rule out contributions from potential non-specific interactions at multiple sites close to both PRE Cys mutations, we repeated the experiment with the CE3F4[c] analog, in which the two bromine atoms are replaced by cyanide substituents[1]. Since the structure of CE3F4[c] is similar to that of CE3F4R, we would expect similar non-specific interactions with EPAC. However, since CE3F4[c] is the weakest EPAC inhibitor of all CE3F4 analogs as seen by guanine exchange functional assays[1] and our chemical shifts analyses (Fig. 3H), we wouldn't expect to observe any significant line broadening in the absence of non-specific interactions. When CE3F4[c] was added to the MTSL-labelled S210C mutant, there were only marginal changes in peak intensities between oxidized and reduced samples (Fig. S1F; Table 1). This confirmed that the PRE line-broadening observed for CE3F4R reflects specific interactions.

3.11 Supplementary Information: Experimental Procedures

3.11.1 NOESY and 1D STD Experiments.

Chemical exchange between cis and trans conformers of CE3F4R was probed by off resonance ROESY-experiments, as described previously[2], with a mixing time of 158 ms. The sample contained 200 μM CE3F4R in $NaPO_4$ buffer with 1% DMSO. NOESY spectra were acquired for the same sample as well as a sample containing 50 μM EPAC with 300 μM cAMP and 300 μM CE3F4R. NOESY spectra had a mixing time of 250 ms, 64 scans and 64 dummy scans, with 2048 complex points in the direct dimension and 512 complex points in the indirect dimension and a spectral width in both dimensions of 12ppm. Saturation transfer difference (STD) experiments

were performed as described previously[3-7] with 16K points, a spectral width of 13 ppm centered at 4.7 ppm, and 256 scans. The saturation frequency was 0.66 ppm unless otherwise stated. A 30 ms spin-lock filter was also utilized to minimize interference from protein signals. A 20 μM concentration of EPAC was used for these experiments, while cAMP and CE3F4R were kept at 200 and 100 μM, respectively. The DMSO concentration was kept constant at 0.5 %.

3.11.2 Chemical Shift Analyses.

Unless otherwise specified, all NH-HSQC used for chemical shift analyses were acquired on samples with 100 μM EPAC, 1 mM cAMP (or other cyclic nucleotides), 200 μM CE3F4R and 1 % DMSO. Compounded chemical shift changes (ΔCCS) between two samples (denoted as one and two) were computed as:

$$\Delta CCS = \sqrt{(\delta_{H1} - \delta_{H2})^2 + (0.2 \times (\delta_{N1} - \delta_{N2}))^2}$$

The chemical shift projection analysis (CHESPA) was performed as described previously[3,4,7,8] with one exception. For the cos(θ) measurements, residues were only included in the analysis if the ΔCCS of the perturbation vector (*e.g.* the vector between cAMP-bound EPAC and cAMP:CE3F4R-bound EPAC in Fig. 3D inset) was greater than 0.04 ppm. Whereas, for fractional activation (X) values, residues were considered if the ΔCCS of either the perturbation vector or reference vector (e.g. the vector between the apo and cAMP-bound EPAC cross-peaks in Fig. 3D inset) was greater than 0.04 ppm.

The dissociation constants (K_D) for the binding of CE3F4R to WT and L273W EPAC were computed from the chemical shift changes in NH-HSQCs upon titration of CE3F4R. During the titrations full saturation could not be achieved due to the limited solubility of CE3F4R. Using lower protein concentrations to achieve a higher fraction bound was also not an option, as EPAC helps to solubilize the ligand through binding. Residues with well resolved peaks that did not overlap with any other resonances during the titration and with a ΔCCS between 0 and 200 μM CE3F4R concentrations greater than 0.05 ppm were used to build the binding isotherms. The curves were fit using a 1:1 binding model with fast on/off exchange, using the maximum ΔCCS value as one of the fitted parameters. The K_D values from each residue were then averaged and the error was computed from the standard deviation of the average.

3.11.3 2D Saturation Transfer Experiments.

HSQC-based saturation transfer experiments were performed on samples of uniformly ^{13}C- and ^{15}N-labelled EPAC for both the binary (cAMP-bound) and ternary (cAMP and CE3F4R bound) complexes. Concentrations of EPAC, cAMP, CE3F4R and DMSO were 125 μM, 1.5 mM, 500 μM and 1.4% v/v, respectively. The samples were saturated for 2s at the CE3F4R formyl resonance (8.86 ppm). Experiments were performed in $NaPO_4$ buffer so that protein amides could exchange with D_2O. The difference between STD/STR ratios for binary and ternary complexes was also calculated to account for peaks arising from non-selective saturation of the nearby H2/H8 resonances of cAMP or from saturation of highly protected amides that did not fully exchange with D_2O. Errors were determined from propagation of the S/N ratios.

3.11.4 Paramagnetic Relaxation Enhancement (PRE) Experiments.

Cysteine mutations at S210 and T261 were engineered by site-directed mutagenesis to serve as spin-label attachment sites at opposite sides of the protein. The spin-label, (1-Oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate (MTSL), was purchased from Toronto Research Chemicals and covalently attached to the cysteine mutants described above using the same protocol described previously for PKA and HCN[4,9]. To confirm spin-label attachment at the mutation site and to rule out its attachment to a native cysteine at position 290 that is buried within the β-core, NH-HSQC were acquired under oxidizing and reducing (5mM DTT after 1 hour incubation at room temp) conditions and the oxidized vs reduced intensity ratios (I_{ox}/I_{red}) were calculated, normalized to the largest I_{ox}/I_{red} and compared to the distances from the spin-label attachment site. To determine if the I_{ox}/I_{red} ratios were significantly decreased, the average and standard deviation of all I_{ox}/I_{red} ratios were calculated twice, excluding residues that fell outsides the bounds of the avg +/- SD of the initial analysis. Residues with lower I_{ox}/I_{red} ratios than the average – SD of the secondary analysis were considered significantly decreased due to the presence of the spin-label.

Binding of CE3F4R to EPAC was assayed via the SLAPSTIC (Spin Labels Attached to Protein Side chain as a Tool to identify Interacting Compounds) method[10-12]. The SLAPSTIC samples contained 5 μM of spin-labelled EPAC, with 100 μM cAMP and 150 μM CE3F4R. For controls with cAMP alone, the molar ratio of EPAC to cAMP was changed to 1:10. 1D Watergate spectra with and without 10 or 30 ms spin-locks were acquired for all samples. After acquiring spectra under oxidized conditions, samples were reduced as explained above and NMD spectra

were reacquired. Oxidized *vs*. reduced peak intensities were then compared to determine if there CE3F4R bound at EPAC sites within 20-25 Å of the spin label attachment location.

3.12 Supplementary Figures

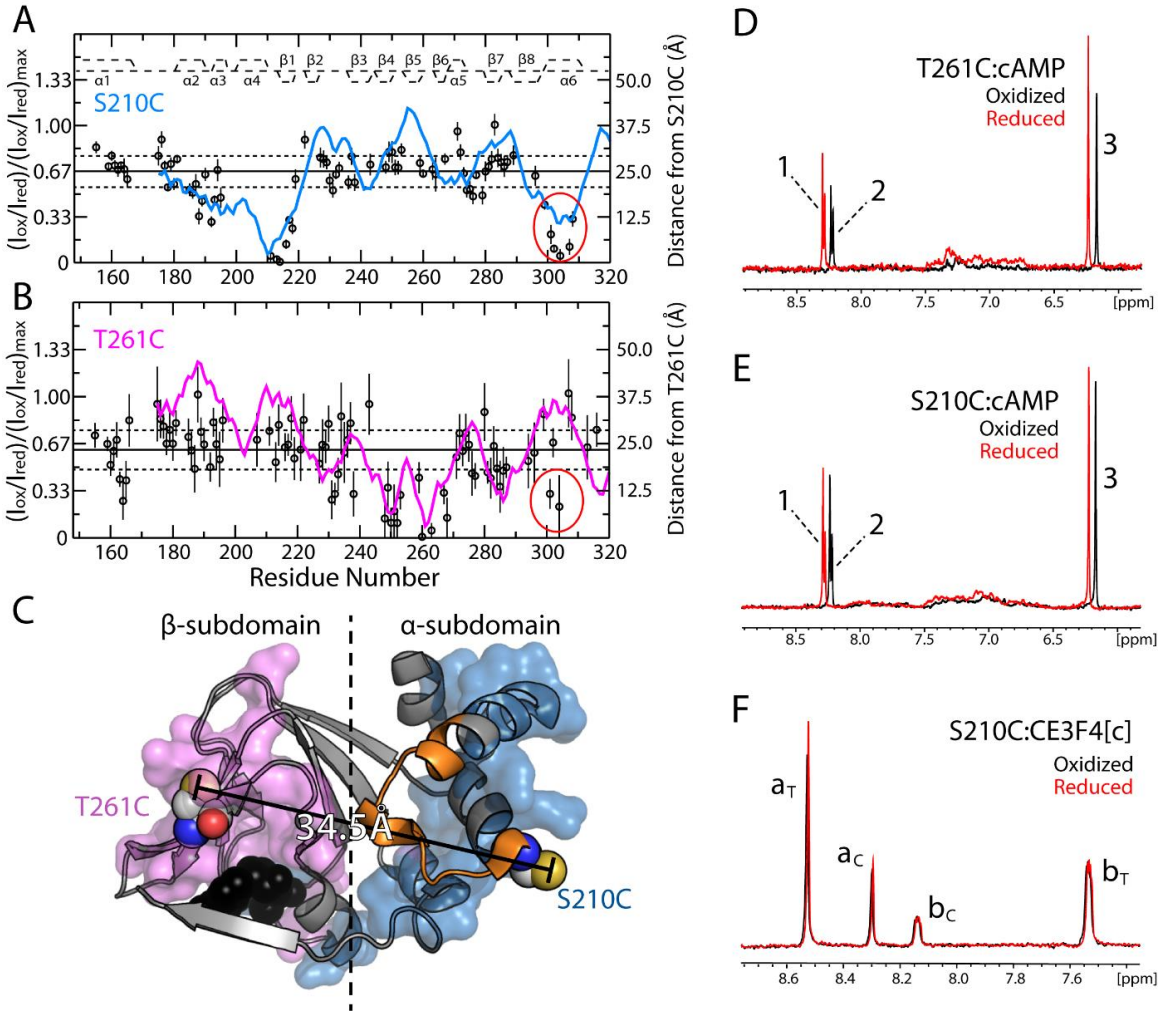

Figure S1. *Control Paramagnetic Relaxation Enhancement (PRE) Experiments*. **A)** The oxidized/reduced intensity ratios for MTSL-labeled S210C normalized to the maximum oxidized/reduced intensity ratio, $(I_{ox}/I_{red})_{max}$. The blue line indicates the distance (right axis) from S210C to the amide of each residue in the cAMP-bound EPAC2 crystal structure. The solid and dashed horizontal lines, represent the average oxidized/reduced intensity ratio +/- SD calculated using an iterative averaging procedure (See experimental procedures). The secondary structure is indicated along the top of the plot. **B)** Similar to panel A except for the T261C mutant. The red circles indicate regions undergoing fast (ps-ns) motions[1,12], explaining why outliers were observed in the oxidized/reduced intensity ratios. **C)** Crystal structure of the EPAC2 CBD-B with spin labeled mutation sites (S210C and T261C) indicated as spheres. The range of the paramagnetic relaxation enhancement was measured by mapping the residues with decreased oxidized/reduced intensity ratios (below the lower dashed line in panels A and B) onto the crystal structure as blue (S210C) and pink (T261C) surfaces. The region depicted in orange is an area prone to intermediate µs-ms dynamics, which results in sparse chemical shift assignment. The distances between S210C and T261C is shown for reference. Oxidized and reduced 1D spectra of SL-T261C **(D)** and SL-S210C **(E)** in a 1:10 ratio of EPAC to cAMP. The reduced spectra (red) was frameshifted to facilitate the comparison of peak intensities. **F)** Oxidized and reduced 1D spectra of SL-S210C in complex with CE3F4[c] (in the absence of cAMP).

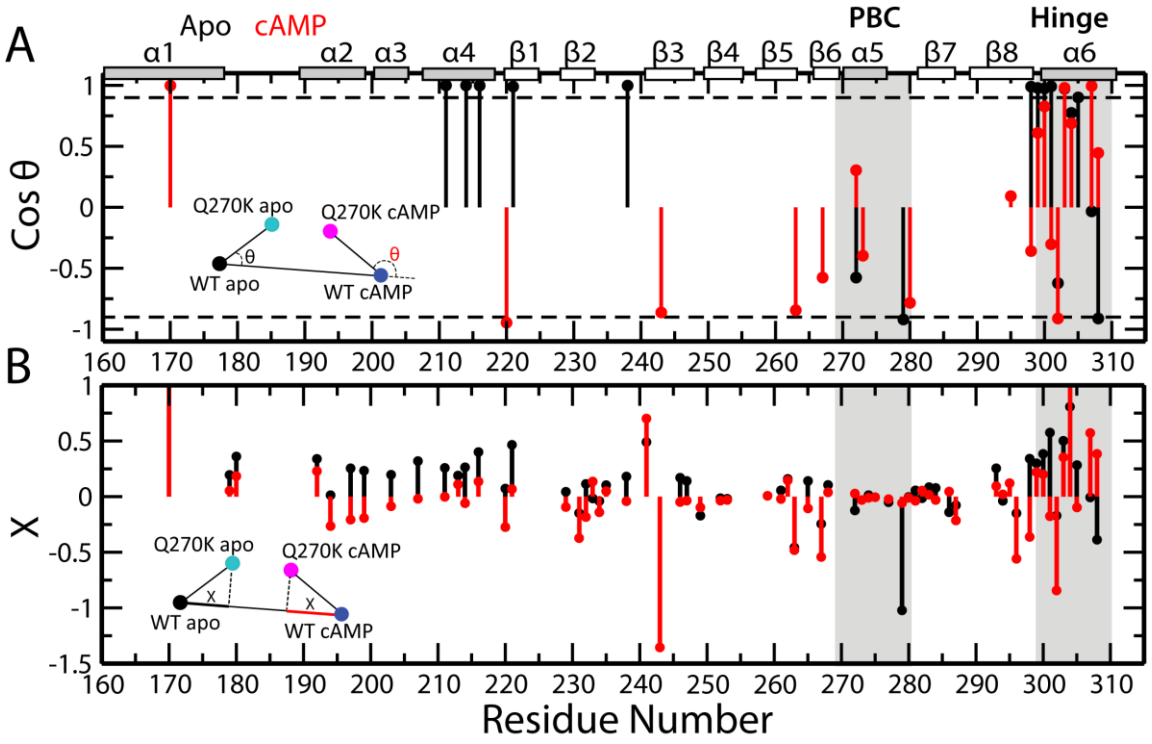

Figure S2. *The Q270K Mutation Enhances the Hinge Recruitment for both apo and cAMP-Bound States*. **A)** CHESPA cos(θ) *vs*. residue plot for the Q270K mutant in the apo (black) and cAMP-bound (red) state. The inset schematic explains how the cos(θ) values were computed for Q270K relative to the WT apo and cAMP-bound states. The dashed lines are the cutoff (-/+ 0.9) for distinguishing linear vs non-linear chemical shift changes. The secondary structure of EPAC is indicated along the top of the plot. **B)** CHESPA fractional activation (X) analysis of Q270K. The inset schematic explains how the X is measured relative to the WT apo and cAMP-bound states.

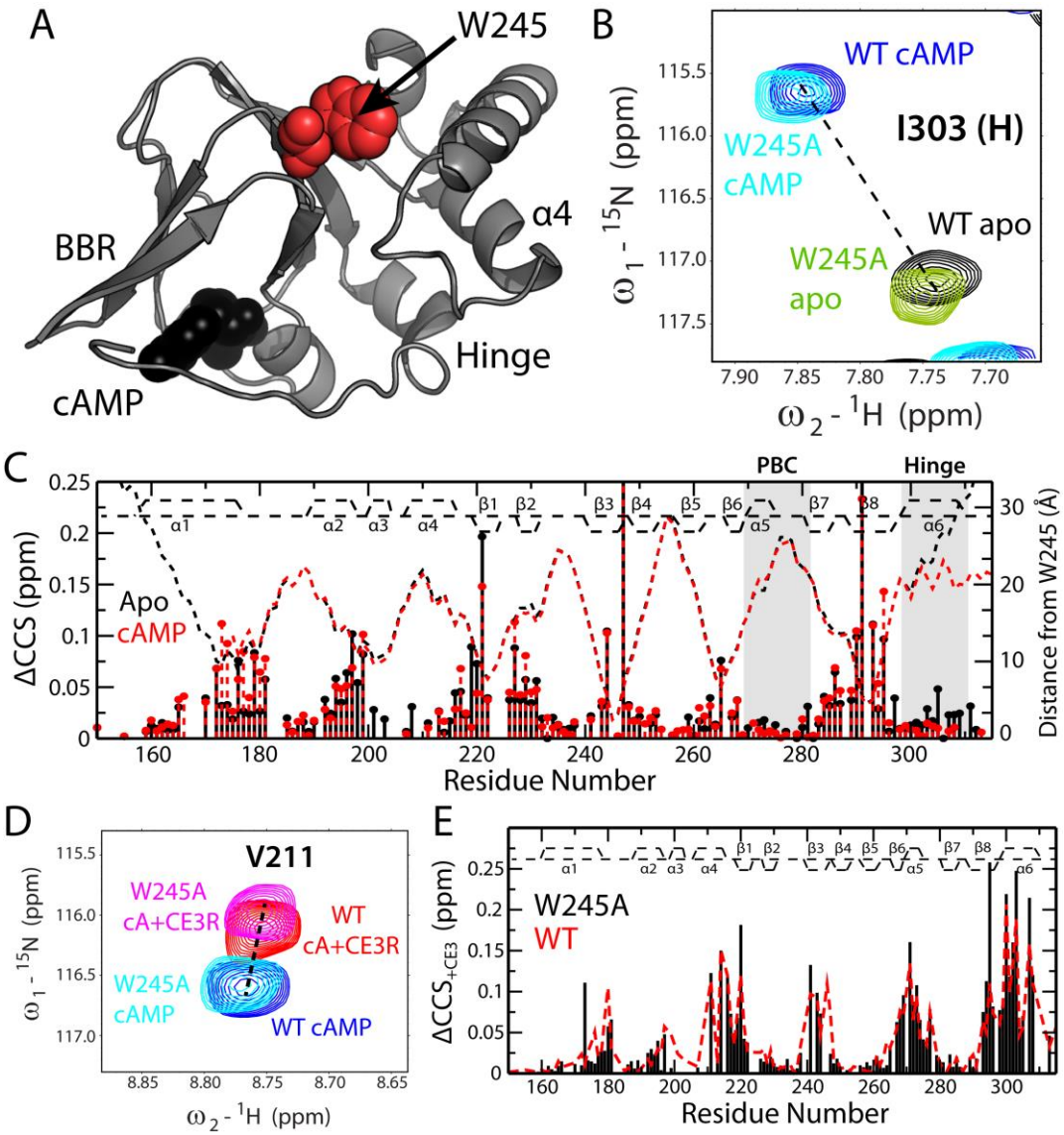

Figure S3. W245 *is not a Key Determinant of the CE3F4R EPAC1 vs, EPAC2 Selectivity*. **A)** An EPAC1 homology model of the cAMP-bound EPAC2 X-ray crystal structure. W245 is shown as red spheres. cAMP is shown as black spheres. **B)** Representative hinge residue from HSQC spectra expansions, showing that W245A does not significantly disturb the hinge helix conformations in either the apo or cAMP-bound states. **C)** The CCS differences between the WT and W245A EPAC1 CBD for both apo (black) and cAMP-bound (red) states. The black and red dashed lines indicate the Cβ to N distances in Å from W245A for the apo and cAMP-bound crystal structures, respectively. **D)** Representative HSQC cross-peak for cAMP-bound WT and W245A EPAC1 CBD in the absence and presence of CE3F4R. **E)** The CCS differences upon addition of CE3F4R to the EPAC1 CBD for both WT and W245A in the presence of cAMP.

3.13 Supplementary References

(1) Courilleau, D.; Bouyssou, P.; Fischmeister, R.; Lezoualc'h, F.; Blondeau, J.-P. The (R)-enantiomer of CE3F4 is a preferential inhibitor of human exchange protein directly activated by cyclic AMP isoform 1 (Epac1). *Biochemical and Biophysical Research Communications* **2013**, *440*, 443–448.

(2) Das, R.; Chowdhury, S.; Mazhab-Jafari, M. T.; Sildas, S.; Selvaratnam, R.; Melacini, G. Dynamically driven ligand selectivity in cyclic nucleotide binding domains. *The Journal of biological chemistry* **2009**, *284*, 23682–23696.

(3) Boulton, S.; Akimoto, M.; Akbarizadeh, S.; Melacini, G. Free energy landscape remodeling of the cardiac pacemaker channel explains the molecular basis of familial sinus bradycardia. *The Journal of biological chemistry* **2017**, *292*, 6414–6428.

(4) Akimoto, M.; Zhang, Z.; Boulton, S.; Selvaratnam, R.; VanSchouwen, B.; Gloyd, M.; Accili, E. A.; Lange, O. F.; Melacini, G. A mechanism for the auto-inhibition of hyperpolarization-activated cyclic nucleotide-gated (HCN) channel opening and its relief by cAMP. *The Journal of biological chemistry* **2014**, *289*, 22205–22220.

(5) Selvaratnam, R.; Mazhab-Jafari, M. T.; Das, R.; Melacini, G. The auto-inhibitory role of the EPAC hinge helix as mapped by NMR. *PloS one* **2012**, *7*, e48707.

(6) Mazhab-Jafari, M. T.; Das, R.; Fotheringham, S. A.; Sildas, S.; Chowdhury, S.; Melacini, G. Understanding cAMP-dependent allostery by NMR spectroscopy: Comparative analysis of the EPAC1 cAMP-binding domain in its apo and cAMP-bound states. *Journal of the American Chemical Society* **2007**, *129*, 14482–14492.

(7) Moleschi, K. J.; Akimoto, M.; Melacini, G. Measurement of State-Specific Association Constants in Allosteric Sensors through Molecular Stapling and NMR. *Journal of the American Chemical Society* **2015**, *137*, 10777–10785.

(8) Selvaratnam, R.; VanSchouwen, B.; Fogolari, F.; Mazhab-Jafari, M. T.; Das, R.; Melacini, G. The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants. *Biophysical journal* **2012**, *102*, 630–639.

(9) Akimoto, M.; Selvaratnam, R.; McNicholl, E. T.; Verma, G.; Taylor, S. S.; Melacini, G. Signaling through dynamic linkers as revealed by PKA. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110*, 14231–14236.

(10) Dias, D. M.; Ciulli, A. NMR approaches in structure-based lead discovery: Recent developments and new frontiers for targeting multi-protein complexes. *Progress in biophysics and molecular biology* **2014**, *116*, 101–112.

(11) Jahnke, W.; Rüdisser, S.; Zurini, M. Spin label enhanced NMR screening. *Journal of the American Chemical Society* **2001**, *123*, 3149–3150.

(12) Jahnke, W. Spin labels as a tool to identify and characterize protein-ligand interactions by NMR spectroscopy. *Chembiochem : a European journal of chemical biology* **2002**, *3*, 167–173.

(13) Das, R.; Mazhab-Jafari, M. T.; Chowdhury, S.; Sildas, S.; Selvaratnam, R.; Melacini, G. Entropy-driven cAMP-dependent allosteric control of inhibitory interactions in exchange proteins directly activated by cAMP. *The Journal of biological chemistry* **2008**, *283*, 19691–19703.

Chapter 4

A Tool Set to Map Allosteric Networks Through the NMR Chemical Shift Covariance Analysis

4.1 Author's Preface

The work presented in this chapter has previously been published and is reproduced here with permission from the Nature Publishing Group. The full citation is as follows:

Boulton S, Akimoto M, Selvaratnam R, Bashiri A and Melacini G. A tool set to map allosteric networks through the NMR chemical shift covariance analysis. *Sci. Rep.* **8**, 07306 (2014).

I conducted most of the analyses required for this publication. Dr. Madoka Akimoto acquired the NMR spectra and assigned the backbone resonances of PKA. Dr Rajeevan Selvaratnam and Amir Bashiri performed preliminary analyses with the complete-linkage algorithm and provided chemical shifts for EPAC. I co-wrote the manuscript with Dr. Giuseppe Melacini.

4.2 Abstract

Allostery is an essential regulatory mechanism of biological function. Allosteric sites are also pharmacologically relevant as they are often targeted with higher selectivity than orthosteric sites. However, a comprehensive map of allosteric sites poses experimental challenges because allostery is driven not only by structural changes, but also by modulations in dynamics that typically remain elusive to classical structure determination methods. An avenue to overcome these challenges is provided by the NMR chemical shift covariance analysis (CHESCA), as chemical shifts are exquisitely sensitive to redistributions in dynamic conformational ensembles. Here, we propose a set of complementary CHESCA algorithms designed to reliably detect allosteric networks with minimal occurrences of false positives or negatives. The proposed CHESCA toolset was tested for two allosteric proteins (PKA and EPAC) and is expected to complement traditional comparative structural analyses in the comprehensive identification of functionally relevant allosteric sites, including those in otherwise elusive partially unstructured regions.

4.3 Introduction

Physiological function and homeostasis are tightly regulated by allostery. Allosteric regulation is also often exploited pharmacologically for enhancing target selectivity [1–11]. One of the most common mechanisms of allosteric regulation relies on the coupling of binding and conformational equilibria, which is modeled by a four-state thermodynamic cycle (Figure 1A). For example, in one of the prototypical allosteric systems, *i.e.* the regulatory subunit of Protein Kinase A (PKA R), activation is controlled by a dynamic equilibrium between inactive and active conformations that differ not only at the binding site of the allosteric effector, *i.e.* cAMP, but also at remote loci essential for inhibition of the catalytic subunit (Figure 1A) [7,8,12,13]. The allosteric effector cAMP binds with higher affinity to the active *vs.* the inactive conformation of PKA R. Because of the active *vs.* inactive state selectivity of cAMP, binding of cAMP to the regulatory subunit shifts the conformational equilibrium towards the active state (Figure 1A), weakening the association of the regulatory and catalytic subunits and releasing kinase inhibition. Allosteric cycles (Figure 1A) provide therefore a simple but effective thermodynamic model of how ligands allosterically regulate remote inhibitory sites.

In order to manipulate the thermodynamics of allostery for therapeutic purposes, it is essential to fully map at atomic resolution the active *vs.* inactive state differences, which will be collectively referred to here as 'allosteric networks'. When each discrete functional state (*e.g.* inactive and active) is structurally homogeneous, *i.e.* adopts a well-defined and distinct structure as determined through classical structure determination methods [14–16], comparative structural analyses effectively map allosteric networks of interactions that link distal protein sites. For example, in the cAMP binding domain (CBD) of PKA R, which is composed of an α- and a β-subdomain, active *vs.* inactive conformational differences are observed primarily in the α-subdomain and in the cAMP binding motif known as the phosphate binding cassette (PBC) (Figure S1) [17,18]. Upon cAMP binding, the PBC and the C-terminal helices shift inwards towards the β-subdomain, while the N-terminal helical bundle (NTHB) moves away from the β-subdomain (Figure S1B). In this case, the comparative structural analysis of R bound to either the kinase subunit (C) or to cAMP is effective in revealing the cAMP-dependent allosteric networks within the globular CBDs of PKA.

While comparative structural analyses are an invaluable approach for the elucidation of allosteric networks, a growing body of evidence indicates that allostery relies not only on structural changes, but also on modulations of dynamics [19–24]. In addition, critical inhibitory sites under allosteric control are often found in partially unstructured regions, such as flexible linkers that

remain elusive to classical structure determination methods [12–14,19,25]. Two examples that illustrate the importance of dynamics and flexible linkers in allostery are provided by the Exchange Protein directly activated by cAMP (EPAC) and PKA R. In EPAC1, a region known to mediate multiple inhibitory cAMP-dependent interactions with the catalytic domain was found to be subject to only minimal cAMP-dependent structural changes [24,26–28]. However, cAMP binding to EPAC1 causes an enhancement of dynamics in this region, weakening the underlying auto-inhibitory interactions by imposing an entropic penalty [24,26]. In PKA R, the comparative analysis of crystal structures could not reliably identify allosteric sites in a dynamic linker critical for kinase inhibition, because electron density was either missing or affected by crystal packing [12,17,18]. However, the linker was later found to elicit state selective interactions that allosterically couple it to cAMP [12]. For both EPAC and PKA R, these otherwise elusive dynamic allosteric sites were detected using an alternative approach known as the CHEmical Shift Covariance Analysis (CHESCA) [12,25]. In addition, CHESCA has been applied to other systems, revealing amino acid networks underlying enzyme catalysis and inhibition, which have been confirmed by independent mutational analyses, and shows promise for *in vivo* applications [12,25,29–35].

The CHESCA method is particularly effective in mapping functionally relevant allosteric sites within dynamic and partially unstructured regions, which are common in signalling systems but often escape detection through classical structure determination methods [3,9,12,23,25,29–32]. CHESCA relies on the covariance analysis of NMR chemical shifts to identify and functionally categorize allosteric networks of residues eliciting concerted responses to a small library of analogs of the allosteric effector ligand. The analogs feature covalent modifications that perturb the non-covalent interactions anchoring the endogenous allosteric effector (*i.e.* cAMP) to its receptor (*i.e.* PKA RIα) and typically include reverse-agonists (*i.e.* Rp-cAMPS, Figure 1B), antagonists, partial (*i.e.* 2'-OMe-cAMP, Figure 1B) and full agonists (*i.e.* Sp-cAMPS, Figure 1B). These functionally diverse ligands are utilized under fully saturating conditions to effectively lock the inactive *vs.* active equilibrium at different degrees of activation. {^{15}N,^{1}H}-HSQC spectra are then acquired for the allosteric protein under investigation either in the *apo* form or saturated by each selected ligand (Figure 1C). To reduce the dimensionality of the HSQC chemical shifts, the nitrogen and proton chemical shifts of each residue (δ_N and δ_H, respectively) are linearly combined as:

$$CCS = 0.2\delta_N + \delta_H \qquad (1)$$

where CCS is the compounded chemical shift. When the inactive *vs.* active exchange is fast in the NMR chemical shift time scale, as often the case [12,25], the chemical shifts observed for residues sensing exclusively the allosteric conformational equilibrium are linear weighed averages between those of the pure active and inactive states. Under these conditions, the modulation of the inactive *vs.* active equilibrium by the ligands in the CHESCA library results in residue-specific CCS changes that are linearly correlated (Figure 1D) [25].[18] Hence, linear inter-residue pairwise CCS correlations (IPCs, Figure 1D) serve as effective signatures for residue pairs exhibiting a concerted response to the perturbations implemented by the CHESCA library. In this respect, IPCs provide the foundation for the systematic elucidation of allosteric networks [25].

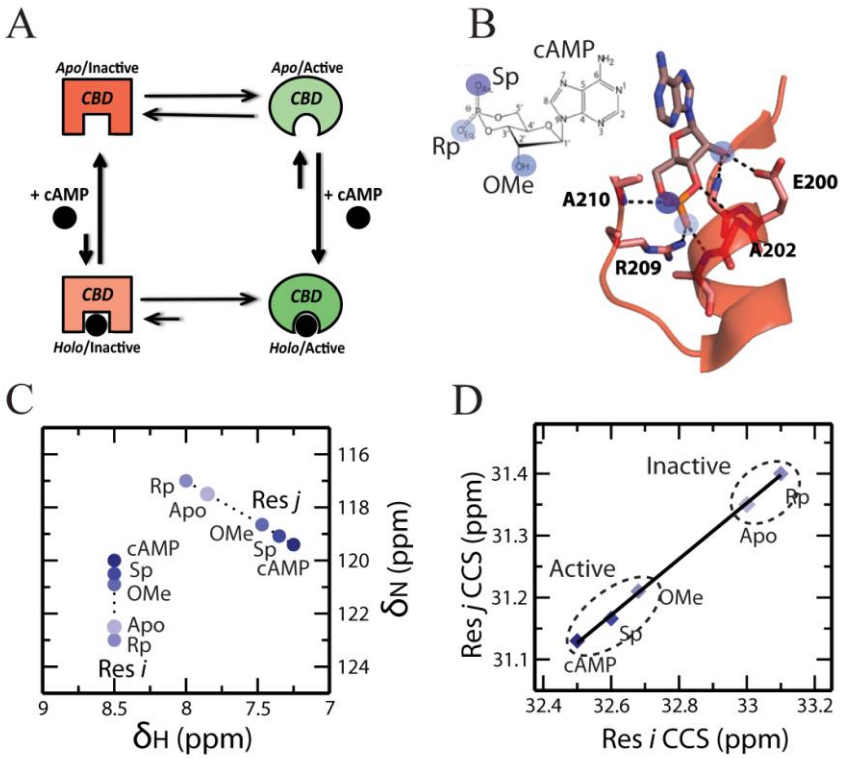

Figure 1: *Allosteric thermodynamic cycle and the CHESCA experimental design.* **(A)** Allosteric thermodynamic cycle of PKA RIα based on the coupling of auto-inhibitory and cAMP-binding equilibria. **(B)** Chemical library selection for the CHEmical Shift Covariance Analysis (CHESCA). The blue circles mark cAMP sites that interact directly with the CBD and are modified in the CHESCA library (*e.g.* S_p-cAMPS, R_p-cAMPS and 2'-OMe-cAMP). **(C)** Representative HSQC peaks of two residues, *i* and *j*, subject to the CHESCA perturbations outlined in **(B)**. **(D)** The corresponding pairwise inter-residue correlation plot for the compounded chemical shifts (CCS) of residues *i* and *j*.

Another critical feature of the CHESCA method is that the point-distribution in the IPCs provides a means to assign a function to the residue networks identified through CHESCA. For instance, networks defined by IPCs similar to the one shown in Figure 1D, in which inhibited forms (*i.e. apo* or bound to reverse-agonists and antagonists) are segregated from the active forms (*i.e.* bound to agonists), are assigned an allosteric function. Whereas networks featuring IPCs in which the separation is between *apo* and bound forms, irrespective of the degree of activation, are assigned a primarily binding function [12,25]. An exhaustive map of allosteric and binding networks requires therefore the identification of all possible IPCs defined by a given library of perturbations (*i.e.* ligand analogs and/or mutations).

In order to systematically identify all possible residue pairs involved in linear IPCs, the correlation matrix (**R**) is computed. **R** is the correlation matrix of **M** transpose, where **M** is a matrix obtained by compiling the observed CCS values, in which rows correspond to residues and columns to ligands in the CHESCA library. In the original implementation of CHESCA, denoted here as CHESCA-SL, the **R** matrix was utilized to identify residue clusters based on single linkage hierarchical clustering [12,25,31], an agglomerative algorithm in which a single linear IPC is sufficient for assigning a residue to a cluster, *i.e.* cluster growth relies on local criteria [12,25,31]. Single linkage hierarchical clustering is effective in exhaustively reconstructing allosteric networks, minimizing false negatives even when all possible linear IPCs are not experimentally observed. However, single linkage methods are also known to be biased by chaining effects that lead to false positives [36–39].

Here, we show how false positives arising from the use of single linkage clustering in CHESCA-SL are identified and minimized. The identification of false positives relies on cross-checking the CHESCA results through an independent method to analyse chemical shifts, *i.e.* the chemical shift projection analysis (CHESPA) [40]. The CHESPA is a simple vector analysis of the (^{1}H,^{15}N) chemical shifts for three states, typically the *apo*, the allosteric effector-bound and an analog-bound form (Figure S2). Hence, CHESPA, unlike CHESCA, alone is not suitable to identify allosteric networks, but it is useful to monitor at residue-resolution the response to selected ligands in the CHESCA library and cross-check the CHESCA-derived clusters for false positives, *i.e.* residues with analog-responses differing from the majority of the remaining residues in the same cluster. We also show that once false positives are identified through CHESPA, they can be minimized through the use of a clustering algorithm complementary to single linkage, *i.e.* complete linkage agglomerative clustering (CHESCA-CL).

Our results indicate that CHESCA-CL significantly reduces the number of false positives, but does not completely eliminate them. We show that the residual false-positives arise from the degeneracy intrinsic to the combination of (^{1}H,^{15}N) chemical shifts into a single CCS, as defined in equation (1). The residual false-positives are eliminated by a modified CHESCA scheme based on the covariance analysis of separate ^{1}H and ^{15}N chemical shifts. The combination of complete-linkage and separate ^{1}H and ^{15}N chemical shifts results in a robust algorithm (CHESCA-I) for the reliable identification of allosteric networks. The robustness of CHESCA-I was confirmed by the overall agreement with the allosteric networks independently mapped through the inactive *vs.* active comparative structural analysis of well-folded globular domains. Furthermore, for less structured and dynamic regions, CHESCA-I preserves the ability to identify otherwise elusive functional allosteric sites. Overall, the proposed algorithms (CHESCA -SL, -CL and -I) define a CHESCA tool set for reliably mapping allosteric networks and we provide a 'user guide' flow chart for the effective implementation of this CHESCA tool set, through which potential false positives and false negatives are detected and minimized.

4.4 Results and Discussion

4.4.1 The Chemical Shift Projection Analysis (CHESPA) Reveals False Positives in CHESCA-SL.

The CHESPA analysis of R_p-cAMPS (or Rp in short) was used to identify false positives in the allosteric clusters defined through CHESCA-SL as applied to RIα 91-244, which spans the critical CBD of PKA. Rp was chosen over the other analogs in the CHESCA library since it is a reverse agonist[41–43]. Therefore, residues sampling primarily the allosteric inactive *vs.* active equilibrium are expected to exhibit ppm changes reflecting an opposite shift in the activation equilibrium relative to cAMP (fractional activation $X < 0$; Figure S2), whereas residues affected by Rp binding but not allosteric conformational changes would experience chemical shift changes similar to cAMP ($X > 0$) as well as unique ppm shifts influenced by the replacement of the equatorial phosphate oxygen with a bulkier sulphur atom in Rp (*i.e.* NNEs). The residue-specific X values observed for Rp in PKA RIα 91-244 are shown in Figure 2. As expected, the CHESPA analysis of Rp results in a splitting of chemical shift changes between those that shift in a direction similar to cAMP ($X > 0$) and those that shift in the opposite direction ($X < 0$; Figure 2). The negative X values are observed primarily for residues in the α-subdomain, which was previously

predicted to play an integral role in the allosteric activation of PKA, while the positive X values are mainly localized in the β-subdomain, a region that contains two important cAMP binding elements, the base binding region (BBR) and the phosphate binding cassette (PBC) (Figure S1).

Residues from the CHESCA-SL allosteric network are highlighted in Figure 2 as solid vertical lines to cross-check whether the CHESCA-SL analysis could distinguish reliably between allosteric and binding elements. Allosteric residues are expected to sense the same conformational equilibrium and hence to share similar fractional activations (X). While this is proved true for the majority of the residues in the allosteric cluster identified through CHESCA-SL (Figure 2C), a subset of residues within the same cluster (*e.g.* β-core residues 162-165, 178, 213, 216) exhibit positive fractional activations (Figure 2C). Among these, some correspond to marginal X values close to zero, suggesting that they are within the noise of the CHESPA analysis, but other residues, such as I163, Q164 and K216, feature significant fractional activations (Figure 2C) and clearly capture false positives of the CHESCA-SL analysis, which is based on single linkage clustering, a method notorious for the presence of chaining effects.

4.4.2 One Source of False Positives in CHESCA-SL is the Single-Linkage "Chaining" of Weakly Correlated Residues.

We hypothesized that residues with opposite fractional activations (X) are clustered together by CHESCA-SL because of an inherent chaining property of the single linkage hierarchical agglomerative clustering method utilized in CHESCA-SL. Single-linkage clustering links two clusters together if there is a high correlation between any of the residues within either of the clusters[36,37]. Even if the majority of correlations for residues between those two clusters are poor, a single high correlation will still cause the two clusters to be linked together. For example, residues K216 and L221 belong to the same allosteric cluster as defined by CHESCA-SL, but exhibit fractional activations with opposite signs (Figure 2C) and hence are very poorly correlated, as shown in the IPC of Figure 3A. The correlation coefficient between K216 and L221 is 0.83, which is significantly lower than the 0.98 cutoff typically used in CHESCA-SL[12,25]. Furthermore, Figure 3A shows that along the K216 axis the Rp state is positioned roughly midway between the *apo* state (inactive) and the three other *holo* states cAMP, Sp and OMe (active), whereas along the L221 axis it is the *apo* state that is found between the Rp and the three active states. Hence, the example illustrated in Figure 3A reveals that single-linkage may cluster together two residues, such as K216 and L221, with markedly different responses to the ligands in the CHESCA perturbation library.

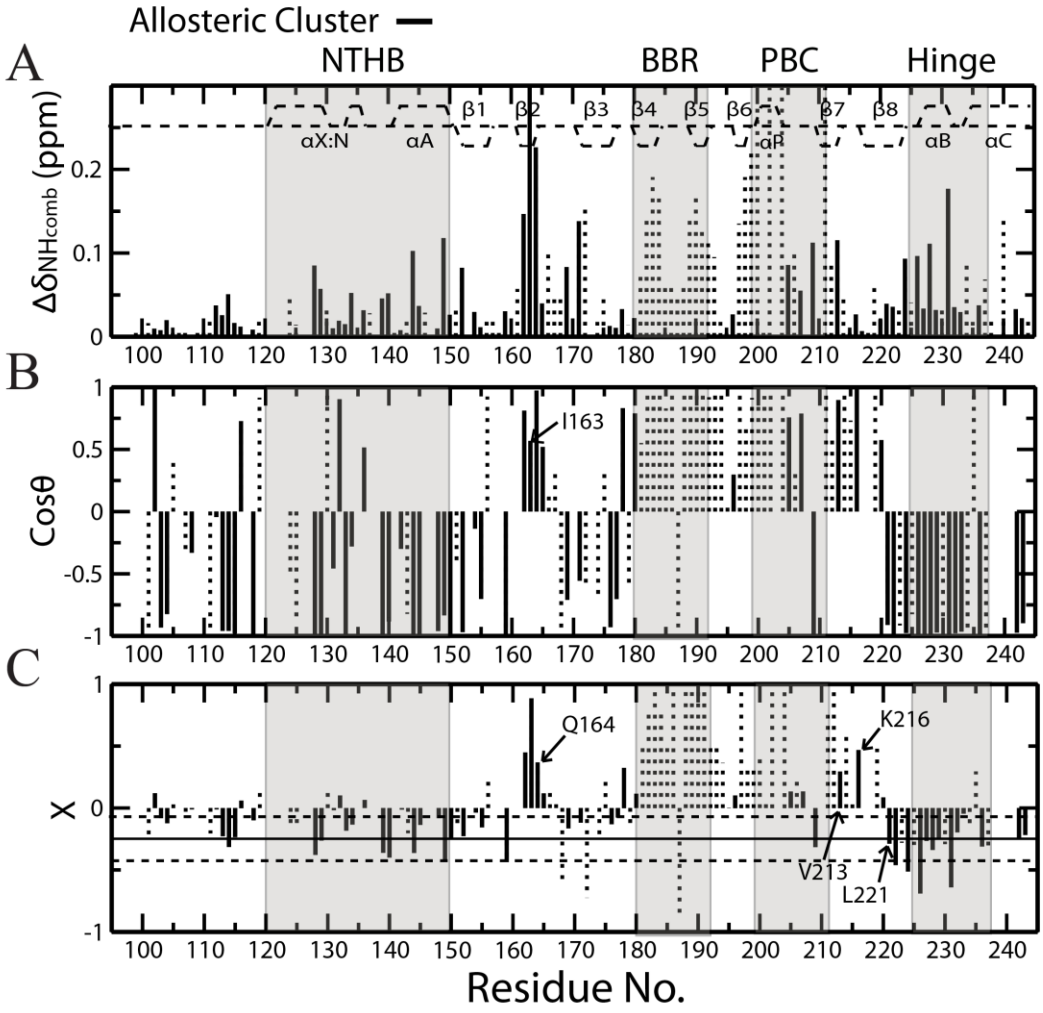

Figure 2: *The chemical shift projection analysis (CHESPA) of R_p-cAMPS (Rp) bound PKA RIα CBD-A.* **(A)** The $\Delta\delta_{NHcomb}$ *vs.* residue plot for R_p-cAMPS bound *vs. apo* PKA RIα (91-244). Solid black lines represent residues belonging to the allosteric cluster identified from the CHESCA-SL analysis, while the $\Delta\delta_{NHcomb}$ values of other residues are reported as dotted vertical bars. The secondary structure of PKA is shown as dashed lines at the top of the plot. **(B)** The cosθ *vs.* residue plot. The θ angle is defined in Figure S2. **(C)** Plot of fractional activations (X) *vs.* residue. Fractional activations are defined in Figure S2. The solid horizontal line represents the average of all the negative X values with adjacent horizontal dashed lines corresponding to the average +/- one standard deviation. Selected residues are highlighted.

We hypothesized that the partitioning of two functionally distinct residues, such as K216 and L221, to the same cluster arose from a chain of strong correlations that links together residues for which the direct correlation is very weak. We further expected to see in this chain of correlations very subtle, but consistently incremental shifts in the distributions of states that would explain the large net difference in distributions that are observed between K216 and L221. As expected, a chain of strong pairwise correlations linking the weakly correlated K216 and L221 residues was

identified (Figure 3B-G). This chain begins with the K216-Q164 pair (Figure 3B), then continues with the Q164-V213, V213-Y205 and Y205-V115 sequential correlations (Figure 3C-E) and eventually ends with the V115-L221 IPC (Figure 3F,G). At each step in this chain, the Rp state is subject to a subtle progressive shift towards the *apo* state, until it eventually crosses it (Figure 3B-F). The distribution of states at the beginning of the chain, with Rp midway between the *apo* and *holo* states, is reminiscent of a correlation between residues affected by binding and nearest neighbour perturbations unique to the interactions with Rp. In contrast, the distribution of states at the end of the chain more closely reflects the expected positions for allosteric residues. It is notable that a similar shift is observed for the partial agonist 2'-OMe-cAMP (or OMe in short) bound state [41,44,45]. Near the beginning of the chain, its position is close to the other bound states, such as cAMP and the full agonist Sp-cAMPS (Sp), but as the chain progresses, it slowly shifts towards the *apo* state. This observation corroborates that residues at the beginning of the chain play a role in binding, while those near the end report primarily on the allosteric equilibrium. Overall, the example in Figure 3A illustrates the possibility that single linkage clustering chains together within a single cluster residues with divergent responses to the CHESCA library (*i.e.* chaining effect). Hence, the chaining effect results in false positives within the allosteric cluster identified by CHESCA-SL.

4.4.3 Complete-Linkage Clustering Overcomes the Chaining Effect.

To overcome the chaining effect caused by single-linkage clustering, we considered other types of clustering methods, such as complete linkage clustering. Complete linkage clustering examines the correlations of every pair of residues between groups and, unlike single linkage, will only link groups together if the lowest correlation coefficient among all pairs is above a designated cutoff [46,47]. This ensures that all residues within a given cluster are highly correlated with each other. Due to the high stringency of this method, the number of false positives in the correlations is expected to be considerably reduced relative to single linkage. Hence, we re-analyzed using complete linkage clustering the chemical shift data of PKA RIα previously utilized for CHESCA-SL [12]. Figure S3 shows the dendrogram representing the complete-linkage agglomerative clustering of PKA RIα (91-244). Figure S3 shows that the stringency of complete linkage fragments the single linkage clusters into sub-clusters with reduced size due to the sparse nature of the **R** matrix. For instance, using a correlation coefficient cutoff of 0.98, as for the single-linkage of CHESCA-SL [12,25], the maximum cluster size is nine residues, almost one order of magnitude less than the maximum cluster size obtained through single-linkage (*i.e.* ~60 residues; Figure 3G) [12]. However,

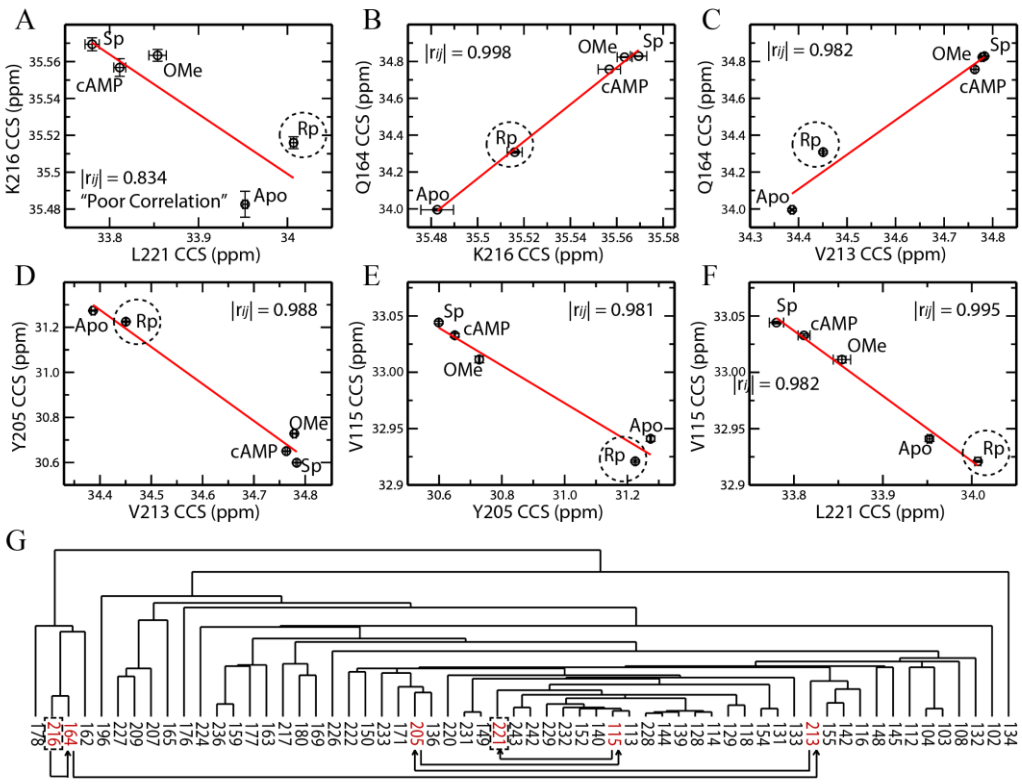

Figure 3: *Analysis of inter-residue CCS pairwise correlations reveals the "chaining effect" of single-linkage agglomerative clustering*. **(A)** Inter-residue pairwise correlation between L221 and K216. Despite the poor K216 *vs.* L221 correlation, single-linkage clustering assigns L221 and K216 to the same cluster because of a chain of strong correlations that links these two residues together, as shown in panels **(B-F)**. **(G)** The single-linkage dendrogram for the allosteric residue cluster reveals the branch locations of residues in panels **(A-F)** (red residues). The black arrows outline the chain of correlations connecting K216 and L221 (dashed rectangles).

complete linkage generates multiple clusters with more than three residues (clusters I-VII, Figure S3 and S4A), which share similar distributions of their functional states (*i.e.* active *vs.* inactive), as proven by the corresponding state dendrograms (Figure S4B). Furthermore, all the residues included in the smallest dendrogram branch that spans clusters I-VII (blue box in Figure S3) exhibit singular value decomposition (SVD) scores aligned along the same principal component (PC), as shown in Figure 4B.

The SVD analysis offers an independent approach to separate residues that contribute to each of the major equilibria (*i.e.* binding and allosteric) [25]. The positions of the loadings in the PC plot reveal the functional role of residues spaced along a given axis. For example, in Figure 4B the loadings progress along PC1 from Sp, cAMP and 2'OMe to *apo* and Rp. This is the pattern expected for the allosteric equilibrium, thereby assigning to PC1 and to all residues with scores

spaced along it a role in allosteric activation. In this respect, it is remarkable that all the amino acids within the selected branch of the complete-linkage dendrogram (blue box in Figure S3) correspond to scores that are aligned along PC1 with minimal variation along PC2 (Figure 4B), suggesting they are all associated to a similar allosteric function.

Based on the combined SVD (Figure 4B) and state dendrogram analyses (Figure S4B), the fragmented sub-clusters generated by complete-linkage were reassembled into a single allosteric cluster (Figure S3, blue box), which will be referred to here as the CHESCA-CL allosteric cluster. The protocol for reconstructing the CHESCA-CL allosteric cluster is summarized in Figure S5 and the corresponding allosteric network is depicted in Figure 4C as a grid of blue lines overlaid to the correlation matrix **R**. Figure 4C also illustrates that despite the added stringency of complete-linkage clustering, the previously identified critical allosteric sites in the flexible linker region spanning residues 100-120 are still captured. Finally, the allosteric residues from complete-linkage clustering were mapped on the fractional activation plot of Rp-bound PKA to verify that there is a significant reduction in the number of false positives (blue lines in Figure 4D).

Figure 4D shows that, as expected, the number of residues with positive fractional activations was greatly reduced. For instance, Q164 and K216, which exhibit positive fractional activations and were classified as allosteric by single linkage clustering (Figure 2C), are now correctly excluded from the allosteric cluster defined through complete linkage (Figure 4D). However, there are still a few residues with positive fractional activations that are assigned to the allosteric cluster by CHESCA-CL (Figure 4D). The most significant of these is residue I163, with a positive fractional activation of ~0.9. To determine why this correlation remained, its IPCs were examined (Figure 5A), but surprisingly the IPCs for this residue matched the expected pattern that was observed for the majority of other complete-linkage allosteric residues, with both the inactive states (*i.e. apo* and Rp) well separated from the active ones (*i.e.* cAMP, 2'OMe, Sp) (Figure 5A). To understand this discrepancy between the CHESPA and CHESCA analyses, we turned to the other CHESPA parameter, the cosθ value (Figure 2B). Interestingly, for I163 the cosθ value is close to 0.5, revealing that the chemical shifts of these states were non-linear in the HSQC spectra (Figure 5B) and suggesting that the CHESPA *vs.* CHESCA discrepancy observed at the level of I163 arises from the use of the CCS (defined in equation 1) as opposed to the separate ^{1}H and ^{15}N ppm values in the covariance analysis.

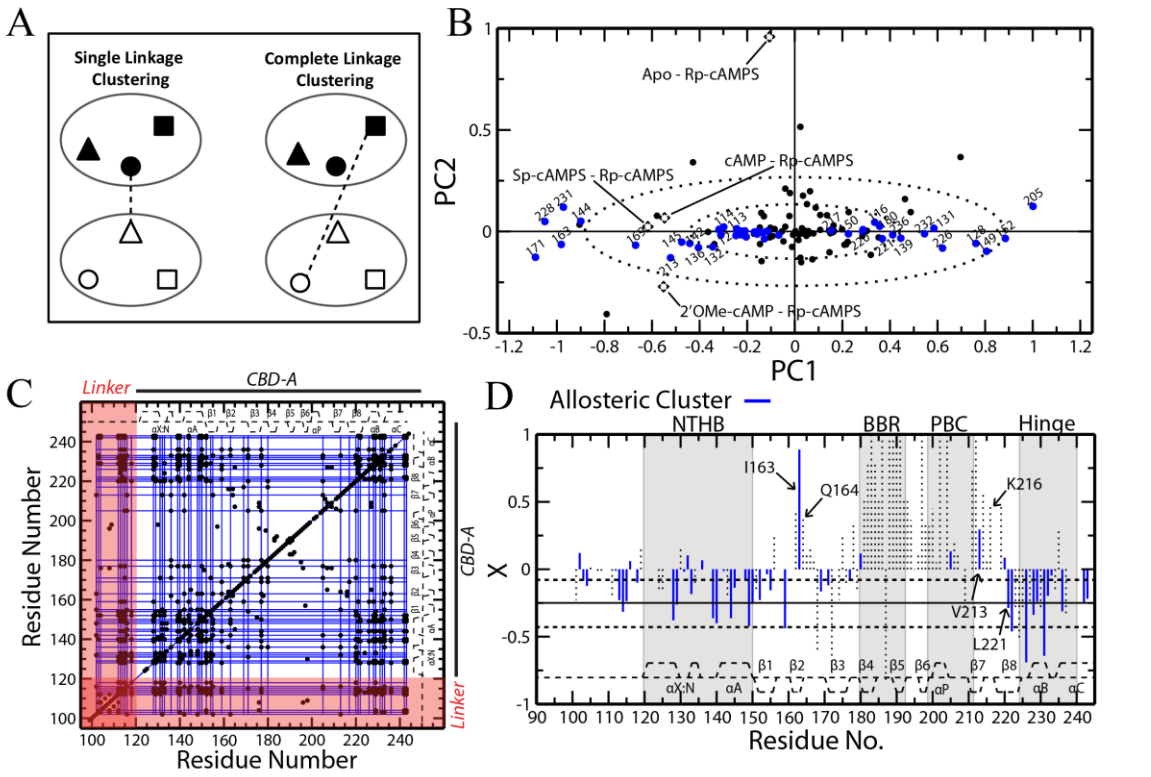

Figure 4: *Complete-linkage agglomerative clustering maps allosteric networks without "chaining effects".* **(A)** Schematic comparison of single *vs*. complete linkage clustering. The shapes (triangle, circle and square) represent residues within existing clusters (large ovals) and the dashed line depicts the method by which the two clusters are linked. **(B)** Singular value decomposition (SVD) analysis of the combined chemical shifts of PKA RIα (91-244). Loadings are shown as black dashed diamonds and scores are shown as circles. Blue scores represent residues from the reconstructed allosteric cluster determined by complete-linkage clustering (Figures S3 and S4). The dashed ovals correspond to the standard deviations of PC1 and PC2. **(C)** The correlation matrix for PKA RIα (91-244) with complete-linkage clusters shown as blue lines. Only correlations with $|r_{ij}| \geq 0.98$ are shown (black dots). The secondary structure is displayed as dashed lines along the top and side of the plot. **(D)** A plot of Rp-bound fractional activations similar to Figure 2C, but with residues from the complete-linkage cluster shown as blue lines. Selected residues are labelled, including I164 and V213 that exhibit positive X values unlike the majority of the residues in the allosteric cluster identified through complete linkage.

4.4.4 Another Source of False Positives Is the Combined Chemical Shift Projection Compression.

Chemical shifts from different nuclei belonging to the same residue are often combined into a single linearly weighed average, in order to reduce the dimensionality of the chemical shift analysis and provide a single residue-specific descriptor utilized to map interactions at residue-resolution. Hence in our original CHESCA-SL implementation, we had calculated the combined ^{1}H and ^{15}N chemical shift using equation (1). In geometrical terms, the linear combination of equation (1) closely resembles a projection of the nitrogen and proton chemical shifts onto an axis rotated by an

angle β relative to the original ^{1}H axis of the ^{1}H,^{15}N HSQC plane. This is appreciated by rewriting equation (1) as:

$$CCS = \alpha(\delta_H \cdot \cos\beta + \delta_N \cdot \sin\beta) \quad (2)$$

where α:

$$\alpha = \sqrt{(1 + SF^2)} \quad (3)$$

and:

$$\beta = \tan^{-1}(SF) \quad (4)$$

with SF being the scaling factor for the nitrogen chemical shifts, *i.e.* 0.2 in the case of equation (1), which results in $\beta = 11.3^{o}$ and $\alpha = 1.02$. Interpreting the CCS equation (1) in terms of a projection onto a rotated axis visually illustrates how HSQC peaks that fall on the same projection line result in the same CCS (Figure 5C). In other words, it is possible that HSQC peaks with distinct chemical shifts lead to identical CCS values. We will refer to this type of CCS degeneracy as "projection compression".

The CCS projection compression provides an explanation as to why I163, which exhibits positive fractional activations in the CHESPA analysis, was assigned by the CHESCA-SL and -CL schemes to the allosteric cluster populated primarily by residues with negative fractional activations (Figures 2C and 4D). The CHESPA relies on a vectorial analysis in a two-dimensional Cartesian coordinate system, thereby removing any effects from the CCS projection compression. On the contrary, the previous CHESCA implementations rely on a mono-dimensional CCS scale and thereby projection compression is a potential concern. For example, the HSQC peaks corresponding to the Rp and *apo* states of residue I163 are clearly well separated (Figure 5B), but they exhibit similar CCS values due to the projection compression (Figure 5B, insert). This results in a linear CCS distribution that fits the pattern expected for the allosteric cluster (Figure 5A).

To overcome the effect of the CCS projection compression, two independent CHESCA analyses were performed on separate ^{15}N and ^{1}H chemical shift matrices. Allosteric clusters were determined for each CHESCA analysis using complete-linkage similarly to the protocol discussed above and residues conserved in both the ^{15}N and ^{1}H–based allosteric clusters were selected (Figure 5D). The Rp fractional activations for these residues, defined as CHESCA-I allosteric ensemble, are displayed in Figure 5E as solid blue lines, showing that the residues with the largest positive fractional activations, such as I163 and V213, are now completely eliminated from the allosteric cluster (Figure 5E). However, there was also a significant reduction in the number of residues with

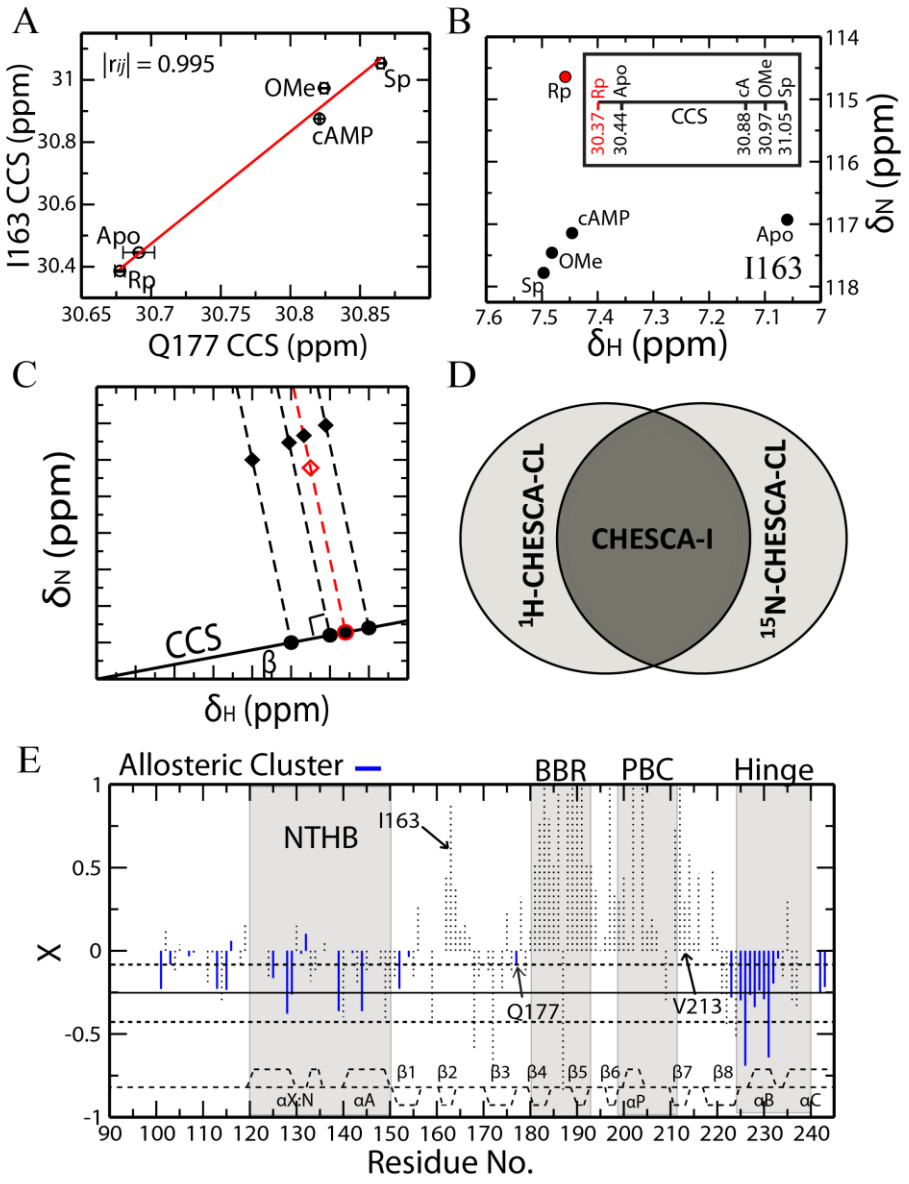

Figure 5: *Potential artefacts arising from the use of the combined chemical shift (CCS) projection and an alternative CHESCA algorithm to account for them*. **(A)** Selected inter-residue pairwise CCS correlation between two residues, I163 and Q177, with different fractional activations for the Rp-bound state. **(B)** HSQC cross-peak positions for residue I163. The CCS values of the five states are displayed in the insert. **(C)** Illustration of the similarity between the use of CCS and the projection of 2D (^{1}H, ^{15}N) cross-peaks into a single axis rotated by an angle β (equation 2). Different 2D (H,N) cross-peaks (black filled and red open diamonds) may result in similar CCS values (circles). **(D)** Scheme for a CHESCA approach designed to circumvent the projection compression effect ("CHESCA-I"). Two independent CHESCA-CL implementations are applied to the nitrogen and proton chemical shifts using a 0.95 $|r_{ij}|$ cutoff value (H- and N-CHESCAs) and only the residues that are conserved between the two complete-linkage clusters generated by the H- and N-CHESCAs are included in the functional network. **(E)** Rp-bound fractional activations for PKA RIα (91-244) with residues from the CHESCA-I method shown as blue lines. Solid and dashed horizontal lines are defined as in Figure 2C.

negative fractional activations (Figure 5E) relative to the previous implementation of complete linkage using CCS (Figure 4D). One possible explanation is that inter-residue correlations were lost for those residues exhibiting linear HSQC variations largely along a single axis, either ^{15}N or ^{1}H, but not both. Thereby, these residues are detected by only the proton or nitrogen analysis and are consequently removed when collecting the residues conserved between the two analyses. It is therefore clear that, while the added value of this method is the stringent determination of pairwise correlations and the drastic reduction in the number of false positives, it also results in an increased number of false negatives and it should be used in conjunction with less stringent and complementary methods such as those relying on the combined chemical shifts (*i.e.* CHESCA-SL and -CL).

4.4.5 Comparison of Structure-Based and CHESCA-Based Allosteric Networks within Folded Domains.

In order to further gauge the reliability of the different CHESCA options outlined here, we sought to compare the allosteric networks derived through CHESCA with those established through traditional structural comparisons. Although CHESCA offers the additional advantage of accurately mapping allosteric networks for partially unstructured regions, the comparative structural analyses of the inactive *vs.* active states for well-structured globular regions, such as the CBDs of PKA, provide a reliable CHESCA-independent map of allosteric networks, which is useful to cross-validate the CHESCA results [13,48,49]. For this purpose, residue-specific RMSD values were measured between the active (*i.e.* cAMP-bound) and inactive (*i.e.* C-subunit bound) structures of PKA RIα CBD-A [17,18] and residues with RMSDs greater than or equal to 1 Å (*i.e.* 10 % of the maximum RMSD) were mapped onto the cAMP-bound crystal structure [8,18] to highlight the regions that undergo major structural changes during activation (Figure 6A). The majority of these residues occur in the α-subdomain, spanning the NTHB and hinge helices as well as the PBC region (Figure 6A).

The RMSD-based allosteric cluster of Figure 6A was compared to that obtained through the traditional CHESCA-SL analysis (Figure 6B), which relies on single linkage clustering. The two allosteric clusters are comparable at the level of the α-subdomain (Figure 6A,B), but in the case of the single-linkage CHESCA-SL cluster (Figure 6B) there are also several residues identified in the β-subdomain. Although the β-subdomain structure is to a large extent invariant upon cAMP-binding, selected residues from the β-core are indeed expected if they are in the vicinity of the α-subdomain and experience local changes in their spatial environment upon transition from the

inactive to the active state. Additional correlations involving the β-core may also be anticipated if they play a significant role in the activation of PKA, but arise from structural and/or dynamical perturbations that fall below the resolution of changes detectable by crystallography. For example, several loops in the β-core play important roles in ligand binding and activation [5,50], but due to their dynamic nature they are poorly structurally defined. Nonetheless, these explanations cannot account for all β-core residues observed in Figure 6B. For instance, single-linkage clustering detects several residues far from the α-subdomain and in rigid β-strand regions (Figure 6B). Such residues are clearly false positives caused by the chaining effects intrinsic to the single linkage clustering of CHESCA-SL.

The β-core residues identified as false positives in the single linkage CHESCA-SL are partially removed by the complete-linkage CHESCA-CL (Figure 6C). As shown in Figure 6C, CHESCA-CL still captures the α-subdomain and the β-core residues that are either in loops or adjacent to the α-subdomain, as anticipated. Very few residues from the PBC were captured, but this is expected since unique PBC perturbations from each of the different cAMP analogs induce ligand specific nearest-neighbour chemical shift changes, which compromise the linearity in at least some of the chemical shift correlations. Finally, the β-core residues identified as false positives in the CHESCA-SL are almost completely removed by the CHESCA-I (Figure 6D). Most of the α-subdomain is still captured as well as adjacent β-core residues (Figure 6D), but the allosteric cluster is now significantly sparser and some regions that were expected to generate CHESCA correlations were absent altogether (*i.e*. residues 148-151 in the NTHB and residues 168,169 and 171-174 in the β2-β3 loop). As a further validation, the proposed CHESCA methods were also applied to EPAC (Figure S6 and supplementary text). Overall, in both PKA R and EPAC the allosteric networks mapped by CHESCA-CL are consistent with those independently defined through comparative structural and/or mutational analyses of the well folded globular domains, while at the same time also capturing dynamic inhibitory sites that would have otherwise remained elusive.

4.5 Concluding Remarks.

The CHEmical Shift Covariance Analysis (CHESCA) relies on a library of perturbations with known functional profiles to comprehensively map at residue resolution allosteric networks. The CHESCA method is especially useful for identifying allosteric residues belonging to dynamic regions, such as linkers and loops, which may remain elusive to traditional structure determination methods. Here, we have identified two sources of false positives in the detection of allosteric

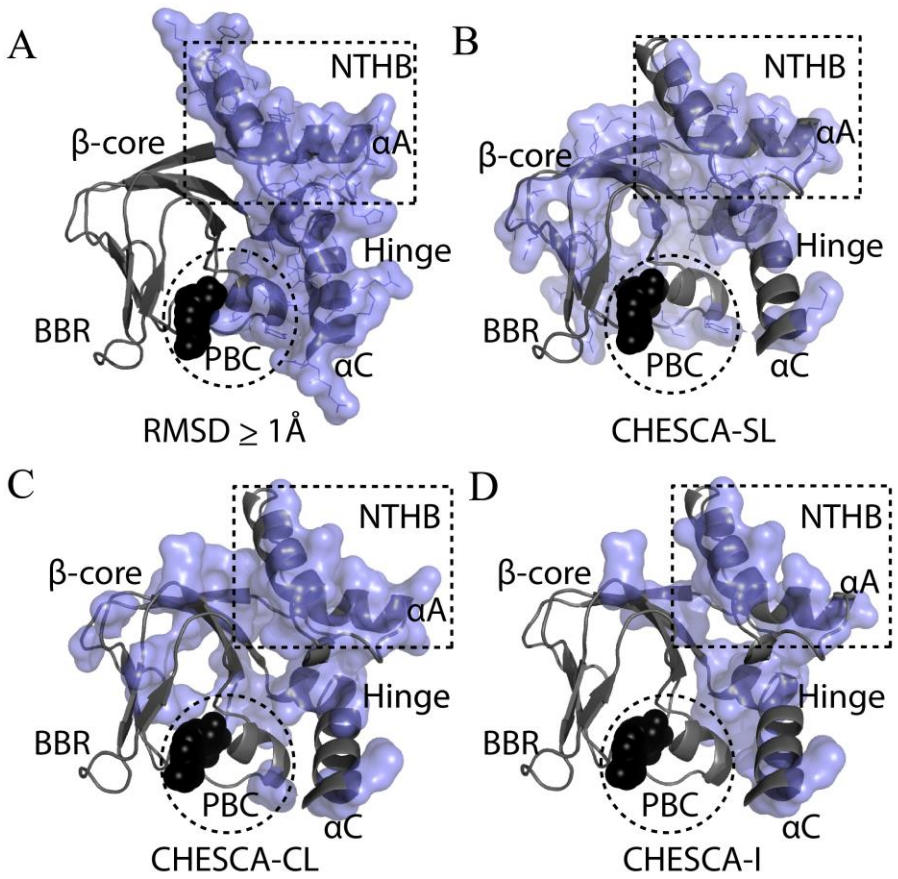

Figure 6: *Summary maps of CHESCA-based allosteric ensembles identified for PKA RIα CBD-A.* **(A)** Structure-based allosteric residues for the folded CBD-A. Residues with a local RMSD between the C-bound and cAMP-bound crystal structures of PKA R (PDB IDs: 3FHI and 3PNA). greater than 1 Å are shown as a blue surface. **(B-D)** Allosteric residues from the CHESCA-SL, -CL and -I analyses of PKA CBD-A, respectively, mapped onto its crystal structure. cAMP is shown as black spheres.

residues through the original CHESCA algorithm (*i.e.* CHESCA-SL). One source of false positive is single linkage clustering, which tends to cluster together poorly correlated residues exhibiting different responses to the CHESCA library ('chaining effects'). Another cause of false positives is the use of combined ^{1}H and ^{15}N chemical shifts, as residues with different ^{1}H and ^{15}N chemical shifts may result in similar combined ppm values ('projection compression'). We have shown that both types of false positives are effectively identified using the CHEmical Shift Projection Analysis (CHESPA). The CHESPA signature that reveals false positives from the CHESCA analysis, is the presence of outlier residues with fractional activations that differ markedly from the majority of other residues in the allosteric cluster.

We have proposed two new CHESCA methods to minimize the presence of false positives. One method (*i.e.* CHESCA-CL) is based on complete rather than single linkage clustering, while

the other (*i.e.* CHESCA-I) relies also on the use of separate rather than combined ^{1}H and ^{15}N chemical shifts. Both CHESCA-CL and -I complement the original CHESCA-SL scheme. CHESCA-SL provides an effective approach for the comprehensive detection of networks of residues involved in conformational equilibria underlying allostery. Although CHESCA-SL may lead to false positives due to chaining effects, it reduces the chances of false negatives especially for systems in which the chemical shift correlation matrix **R** is sparse. As the stringency of the analysis is increased by imposing complete linkage clustering (CHESCA-CL), the number of false positives arising from chaining effects is drastically reduced and the correlations appear to be more directed towards the core groups of residues involved in allostery. If the allosteric residues identified by CHESCA-CL still include residual false positives, as revealed by markedly different residue-specific fractional activations measured through CHESPA, the CHESCA-I is available to minimize the risk of projection compression.

CHESCA-I offers the highest level of stringency, which is advantageous as it not only reduces false positive artefacts, but it also 'zooms in' on core allosteric residues in the protein assisting the prioritization of sites to be tested by mutagenesis. However, the added stringency of CHESCA-I also results in increased false negatives. In this respect, a balanced application of the CHESCA method is likely to be the CHESCA-CL. However, it is important to consider that CHESCA-CL is part of a larger tool set of complementary chemical shift analyses (*i.e*. CHESCA–SL, -CL and -I as well as CHESPA) and it is advised that multiple analyses from this tool set be implemented for a single chemical shift data set in order to obtain a reliable reconstruction of the underlying allosteric networks. A 'user guide' for the proposed CHESCA tool set is summarized by the flow chart shown in Figure 7.

4.6 Methods

The {^{15}N,^{1}H}-HSQC datasets and chemical shifts used for PKA RIα CBD-A (91-244) were acquired at a protein concentration of 10 μM and a ligand concentration of 3 mM [12]. The error bars were derived from the standard deviation of chemical shifts at saturation, *i.e*. ligand concentrations of 1-3 mM. The $EPAC1_h$ (149-318) chemical shifts were as in Selvaratnam *et. al.* [25]. The protocol for CHESCA-SL and CHESPA have been described elsewhere [40]. Further details on the methods are available as Supplementary Material.

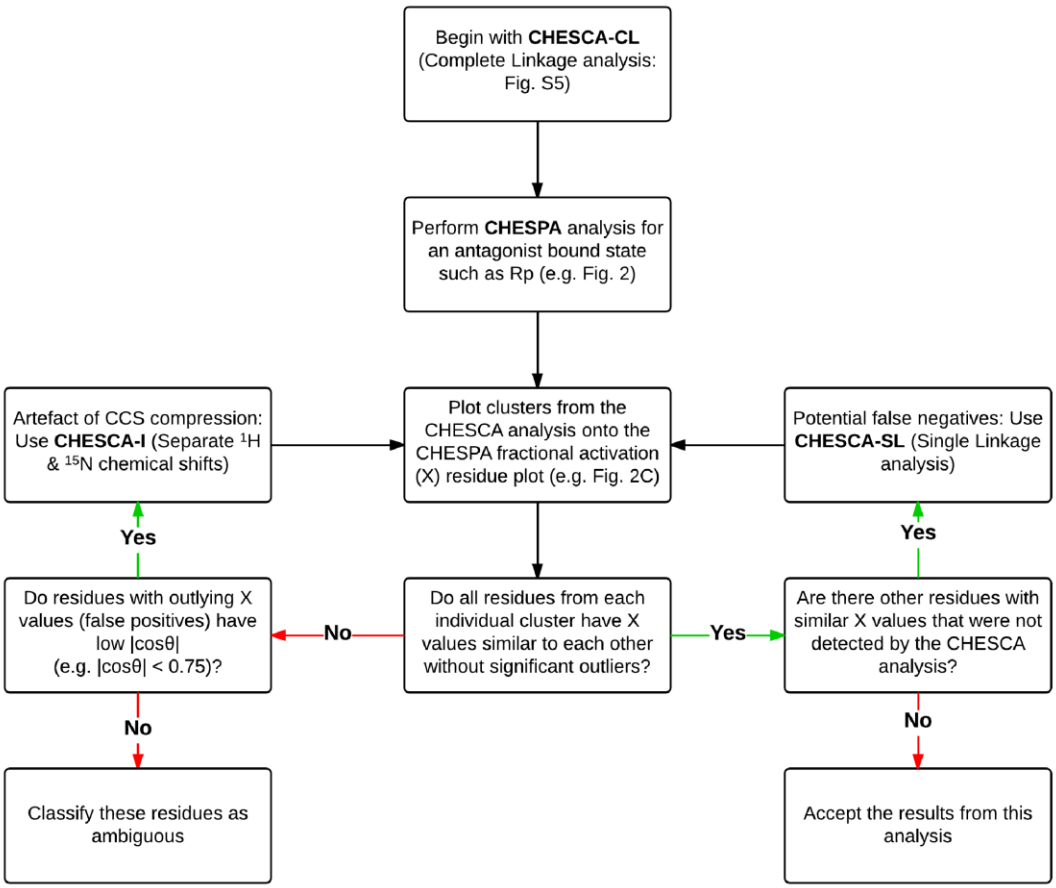

Figure 7: *A flow chart 'user guide' for the implementation of the proposed CHESCA tool set*. Initially, the chart starts with CHESCA-CL, since it provides an optimal balance between the minimization of false positives and false negatives. If residual false positives are detected through CHESPA, the application of the CHESCA-I approach is recommended to minimize those arising from the CCS compression. If false negatives are revealed through CHESPA, the implementation of the CHESCA-SL algorithm is also advised. It is sufficient to implement each type of CHESCA analysis only once. However, multiple types of CHESCA analyses (*i.e.* SL, CL, I) are often needed for a single chemical shift data set.

4.7 Acknowledgements

We thank Dr. Forman-Kay for helpful discussions and CIHR (MOP-68897) and NSERC for funding.

4.8 References

1. Kuriyan, J. & Eisenberg, D. The origin of protein interactions and allostery in colocalization, *Nature* **450,** 983–990 (2007).
2. Das, R. *et al.* cAMP activation of PKA defines an ancient signaling mechanism, *Proc Natl Acad Sci U S A* **104,** 93–98 (2007).
3. Vendruscolo, M. Protein regulation: the statistical theory of allostery, *Nat. Chem. Biol.* **7,** 411–412 (2011).
4. Smock, R. G. & Gierasch, L. M. Sending signals dynamically, *Science* **324,** 198–203 (2009).
5. Kannan, N. *et al.* Evolution of allostery in the cyclic nucleotide binding module, *Genome Biol.* **8,** R264 (2007).
6. Berman, H. M. *et al.* The cAMP binding domain: an ancient signaling module, *Proc. Natl. Acad. Sci. U.S.A.* **102,** 45–50 (2005).
7. McNicholl, E. T. Das, R. SilDas, S. Taylor, S. S. & Melacini, G. Communication between tandem cAMP binding domains in the regulatory subunit of protein kinase A-Ialpha as revealed by domain-silencing mutations, *J. Biol. Chem.* **285,** 15523–15537 (2010).
8. Su, Y. *et al.* Regulatory subunit of protein kinase A: structure of deletion mutant with cAMP binding domains, *Science* **269,** 807–813 (1995).
9. Nussinov, R. Ma, B. Tsai, C.-J. & Csermely, P. Allosteric conformational barcodes direct signaling in the cell, *Structure* **21,** 1509–1521 (2013).
10. Gustavsson, M. *et al.* Allosteric regulation of SERCA by phosphorylation-mediated conformational shift of phospholamban, *Proc. Natl. Acad. Sci. U.S.A.* **110,** 17338–17343 (2013).
11. Cembran, A. *et al.* Conformational equilibrium of N-myristoylated cAMP-dependent protein kinase A by molecular dynamics simulations, *Biochemistry* **51,** 10186–10196 (2012).
12. Akimoto, M. *et al.* Signaling through dynamic linkers as revealed by PKA, *Proceedings of the National Academy of Sciences* **110,** 14231–14236 (2013).
13. Akimoto, M. *et al.* Allosteric linkers in cAMP signalling, *Biochem. Soc. Trans.* **42,** 139–144 (2014).
14. Daily, M. D. & Gray, J. J. Local motions in a benchmark of allosteric proteins, *Proteins* **67,** 385–399 (2007).
15. Frauenfelder, H. McMahon, B. H. Austin, R. H. Chu, K. & Groves, J. T. The role of structure, energy landscape, dynamics, and allostery in the enzymatic function of myoglobin, *Proc. Natl. Acad. Sci. U.S.A.* **98,** 2370–2374 (2001).
16. Francis, S. H. Turko, I. V. & Corbin, J. D. (Elsevier2000), pp. 1–52.
17. Kim, C. Xuong, N.-H. & Taylor, S. S. Crystal structure of a complex between the catalytic and regulatory (RIalpha) subunits of PKA, *Science* **307,** 690–696 (2005).
18. Badireddy, S. *et al.* Cyclic AMP analog blocks kinase activation by stabilizing inactive conformation: conformational selection highlights a new concept in allosteric inhibitor design, *Mol. Cell Proteomics* **10,** M110.004390 (2011).
19. Tsai, C.-J. del Sol, A. & Nussinov, R. Allostery: absence of a change in shape does not imply that allostery is not at play, *J. Mol. Biol.* **378,** 1–11 (2008).
20. Tzeng, S.-R. & Kalodimos, C. G. Dynamic activation of an allosteric regulatory protein, *Nature* **462,** 368–372 (2009).
21. Tzeng, S.-R. & Kalodimos, C. G. Protein dynamics and allostery: an NMR view, *Curr. Opin. Struct. Biol.* **21,** 62–67 (2011).

22. Popovych, N. Sun, S. Ebright, R. H. & Kalodimos, C. G. Dynamically driven protein allostery, *Nat Struct Mol Biol* **13,** 831–838 (2006).

23. Nussinov, R. Tsai, C.-J. & Ma, B. The underappreciated role of allostery in the cellular network, *Annu Rev Biophys* **42,** 169–189 (2013).

24. Das, R. *et al.* Dynamically Driven Ligand Selectivity in Cyclic Nucleotide Binding Domains, *Journal of Biological Chemistry* **284,** 23682–23696 (2009).

25. Selvaratnam, R. Chowdhury, S. VanSchouwen, B. & Melacini, G. Mapping allostery through the covariance analysis of NMR chemical shifts, *Proceedings of the National Academy of Sciences* **108,** 6133–6138 (2011).

26. Das, R. *et al.* Entropy-driven cAMP-dependent allosteric control of inhibitory interactions in exchange proteins directly activated by cAMP, *J. Biol. Chem.* **283,** 19691–19703 (2008).

27. Rehmann, H. *et al.* Structure of Epac2 in complex with a cyclic AMP analogue and RAP1B, *Nature* **455,** 4–127 (2008).

28. Rehmann, H. Das, J. Knipscheer, P. Wittinghofer, A. & Bos, J. L. Structure of the cyclic-AMP-responsive exchange factor Epac2 in its auto-inhibited state, *Nature* **439,** 625–628 (2006).

29. Axe, J. M. & Boehr, D. D. Long-range interactions in the α subunit of tryptophan synthase help to coordinate ligand binding, catalysis, and substrate channeling, *J. Mol. Biol.* **425,** 1527–1545 (2013).

30. Axe, J. M. *et al.* Amino acid networks in a $(\beta/\alpha)_8$ barrel enzyme change during catalytic turnover, *J. Am. Chem. Soc.* **136,** 6818–6821 (2014).

31. Selvaratnam, R. Mazhab-Jafari, M. T. Das, R. Melacini, G. & Hofmann, A. The Auto-Inhibitory Role of the EPAC Hinge Helix as Mapped by NMR, *PLoS ONE* **7,** e48707 (2012).

32. Dawson, J. E. Farber, P. J. & Forman-Kay, J. D. Allosteric coupling between the intracellular coupling helix 4 and regulatory sites of the first nucleotide-binding domain of CFTR, *PLoS ONE* **8,** e74347 (2013).

33. Cembran, A. Kim, J. Gao, J. & Veglia, G. NMR mapping of protein conformational landscapes using coordinated behavior of chemical shifts upon ligand binding, *Phys Chem Chem Phys* **16,** 6508–6518 (2014).

34. Stollar, E. J. Lin, H. Davidson, A. R. & Forman-Kay, J. D. Differential dynamic engagement within 24 SH3 domain: peptide complexes revealed by co-linear chemical shift perturbation analysis, *PLoS ONE* **7,** e51282 (2012).

35. Majumder, S. DeMott, C. M. Burz, D. S. & Shekhtman, A. Using singular value decomposition to characterize protein-protein interactions by in-cell NMR spectroscopy, *Chembiochem* **15,** 929–933 (2014).

36. Sibson, R. SLINK: An optimally efficient algorithm for the single-link cluster method, *The Computer Journal* **16,** 30–34 (1973).

37. Tabakis, E. in *Robust Statistics, Data Analysis, and Computer Intensive Methods,* edited by P. Bickel, *et al.* (Springer New York, New York, NY, 1996), pp. 375–389.

38. Zhao, W. Serpedin, E. & Dougherty, E. R. Spectral preprocessing for clustering time-series gene expressions, *EURASIP J Bioinform Syst Biol,* 713248 (2009).

39. Gaspar, J. M. & Thomas, W. K. Assessing the consequences of denoising marker-based metagenomic data, *PLoS ONE* **8,** e60458 (2013).

40. Selvaratnam, R. *et al.* The Projection Analysis of NMR Chemical Shifts Reveals Extended EPAC Autoinhibition Determinants, *Biophysical Journal* **102,** 630–639 (2012).

41. Poppe, H. *et al.* Cyclic nucleotide analogs as probes of signaling pathways, *Nat. Methods* **5,** 277–278 (2008).

42. Wu, J. Jones, J. M. Nguyen-Huu, X. Ten Eyck, Lynn F & Taylor, S. S. Crystal structures of RIalpha subunit of cyclic adenosine 5'-monophosphate (cAMP)-dependent protein kinase complexed with (Rp)-adenosine 3',5'-cyclic monophosphothioate and (Sp)-adenosine 3',5'-cyclic monophosphothioate, the phosphothioate analogues of cAMP, *Biochemistry* **43,** 6620–6629 (2004).

43. Christensen, A. E. cAMP Analog Mapping of Epac1 and cAMP Kinase: Discriminating analogs demonstrate that Epac and cAMP kinase act synergistically to promote PC-12 cell neurite extension, *Journal of Biological Chemistry* **278,** 35394–35402 (2003).

44. Brown, Simon H J *et al.* Implementing fluorescence anisotropy screening and crystallographic analysis to define PKA isoform-selective activation by cAMP analogs, *ACS Chem. Biol.* **8,** 2164–2172 (2013).

45. Dao, K. K. Epac1 and cAMP-dependent Protein Kinase Holoenzyme Have Similar cAMP Affinity, but Their cAMP Domains Have Distinct Structural Features and Cyclic Nucleotide Recognition, *Journal of Biological Chemistry* **281,** 21500–21511 (2006).

46. Willett, P. Recent trends in hierarchic document clustering: A critical review, *Information Processing & Management* **24,** 577–597 (1988).

47. Defays, D. An efficient algorithm for a complete link method, *The Computer Journal* **20,** 364–366 (1977).

48. Boulton, S. *et al.* 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 (2014).

49. Selvaratnam, R. Akimoto, M. VanSchouwen, B. & Melacini, G. cAMP-dependent allostery and dynamics in Epac: an NMR view, *Biochem. Soc. Trans.* **40,** 219–223 (2012).

50. Boettcher, A. J. *et al.* Realizing the allosteric potential of the tetrameric protein kinase A RIα holoenzyme, *Structure* **19,** 265–276 (2011).

4.9 Supplementary Information: Methods

4.9.1 CHESPA Analysis and Fractional Activations of PKA RIα (91-244).

Briefly, the CHESPA analysis relies on {^{15}N,^{1}H}-HSQC spectra acquired for three samples of the protein of interest. In the case of PKA RIα (91-244), CHESPA focused on the *apo*, cAMP-bound and R_p-cAMPS-bound samples under fully saturating conditions. Two CHESPA vectors are then defined: a reference vector for the endogenous allosteric effector (*i.e.* vector **B** in Figure S2) and a vector for the perturbed state (*i.e.* vector **A** in Figure S2) to be compared to the reference state. The magnitude of vector **B** measures the 'distance' between the *apo* and cAMP-bound peaks calculated for each residue using the equation:

$$\triangle\delta_{NHcomb} = \sqrt{[(\delta_{H,apo} - \delta_{H,cAMP})^2 + (0.2*(\delta_{N,apo} - \delta_{N,cAMP}))^2]} \qquad (5)$$

The distance between the *apo* and R_p-cAMPS bound states was also calculated and defined as the magnitude of vector **A** in the CHESPA analysis. The angle between the **A** and **B** vectors is denoted as θ so that:

$$\cos\theta = \frac{\vec{A}\bullet\vec{B}}{|\vec{A}||\vec{B}|} \qquad (6)$$

and the normalized projection of **A** onto **B** is defined as the 'fractional activation' (X), which is calculated as:

$$X = \frac{\vec{A}\bullet\vec{B}}{|B|^2} \qquad (7)$$

The fractional activation reveals whether a given perturbation (*e.g.* Rp-cAMPS) shifts the protein towards activation (X > 0) or inactivation (X < 0).

4.9.2 CHESCA Analyses of PKA RIa CBD-A (91-244) and EPAC1$_h$ (149-318).

The CHESCA protocol has been described elsewhere [3,5,6]. Briefly, CHESCA relied on {^{15}N,^{1}H}-HSQC spectra acquired for the *apo* and four *holo* states (cAMP, R_p-cAMPS, S_p-cAMPS

and 2'-OMe-cAMP) under fully saturating conditions. Spectra were assigned and the combined chemical shifts (CCS) for each residue were calculated using equation (1). The inter-residue CCS pairwise correlations in the correlation matrix (**R** matrix) with an absolute value of the Pearson's correlation coefficients greater than a selected threshold (*e.g.* 0.95-0.98) were selected. Hierarchical agglomerative clustering (AC) of correlated residues was performed using the Cluster 3.0 software[7] and visualized as dendrograms with Java TreeView[8]. Singular value decomposition analysis of HSQC chemical shifts was performed using Octave.

4.9.3 Structure-Based Allostery Predictions for PKA RIα (91-244).

Local RMSD calculations were performed in MolMol[9] using the PDB structures for C-bound PKA (PDB ID 3FHI)[1] and cAMP-bound PKA (PDB ID 3PNA)[2]. The structures were aligned at their β-cores (150-225) and the local RMSDs were measured for each residue in the structured region spanning residues 112-236. The surface of the residues affected by the cAMP-dependent allosteric conformational change (*i.e.* with RMSD > 1 Å, which is 10% of the maximum RMSD) was mapped onto the structures using Pymol (http://www.pymol.org).

4.10 Supplementary Information: Validation of the EPAC CHESCA Analysis

In Figure S6 the residues from the allosteric clusters in each CHESCA analysis are mapped in green in both the chemical shift plots and the cAMP-bound crystal structure of EPAC[4]. Previously, using the single linkage based CHESCA-SL (Figure S6A,B), we identified an allosteric network that stemmed from the PBC and extended to the α4 and α6 helices *via* a series of hydrophobic spines, culminating at the inhibitory ionic latch in the NTHB (Figure S6A,B). The inhibitory function of the ionic latch (IL) of EPAC was independently supported by site-directed mutagenesis[10]. However, single-linkage clustering also incorporated a number of residues in the β-subdomain region that were not expected to participate in the allosteric network based on the invariance of the β-core[3,10] (Figure S6A,B). When the complete-linkage CHESCA-CL was applied, several unexpected residues from the β-core were removed, while residues that connected the PBC with the ionic latch were still detected, as expected (Figure S6C,D). The results from the complete-linkage $^{15}N/^{1}H$ intersection CHESCA-I analysis (Figure S6E,F) were in general similar to the complete-linkage CCS clustering (Figure S6C,D), but with the PBC residues now missing.

4.11 Supplementary References

1. Kim, C. Xuong, N.-H. & Taylor, S. S. Crystal structure of a complex between the catalytic and regulatory (RIalpha) subunits of PKA, *Science* **307,** 690–696 (2005).
2. Badireddy, S. *et al.* Cyclic AMP analog blocks kinase activation by stabilizing inactive conformation: conformational selection highlights a new concept in allosteric inhibitor design, *Mol. Cell Proteomics* **10,** M110.004390 (2011).
3. Selvaratnam, R. Chowdhury, S. VanSchouwen, B. & Melacini, G. Mapping allostery through the covariance analysis of NMR chemical shifts, *Proceedings of the National Academy of Sciences* **108,** 6133–6138 (2011).
4. Rehmann, H. *et al.* Structure of Epac2 in complex with a cyclic AMP analogue and RAP1B, *Nature* **455,** 124–127 (2008).
5. Selvaratnam, R. Mazhab-Jafari, M. T. Das, R. Melacini, G. & Hofmann, A. The Auto-Inhibitory Role of the EPAC Hinge Helix as Mapped by NMR, *PLoS ONE* **7,** e48707 (2012).
6. Akimoto, M. *et al.* Signaling through dynamic linkers as revealed by PKA, *Proceedings of the National Academy of Sciences* **110,** 14231–14236 (2013).
7. de Hoon, M J L, Imoto, S. Nolan, J. & Miyano, S. Open source clustering software, *Bioinformatics* **20,** 1453–1454 (2004).
8. Saldanha, A. J. Java Treeview--extensible visualization of microarray data, *Bioinformatics* **20,** 3246–3248 (2004).
9. Koradi, R. Billeter, M. & Wuthrich, K. MOLMOL: a program for display and analysis of macromolecular structures, *J Mol Graph* **14,** 51-5, 29-32 (1996).
10. Rehmann, H. Das, J. Knipscheer, P. Wittinghofer, A. & Bos, J. L. Structure of the cyclic-AMP-responsive exchange factor Epac2 in its auto-inhibited state, *Nature* **439,** 625–628 (2006).

4.12 Supplemental Figures

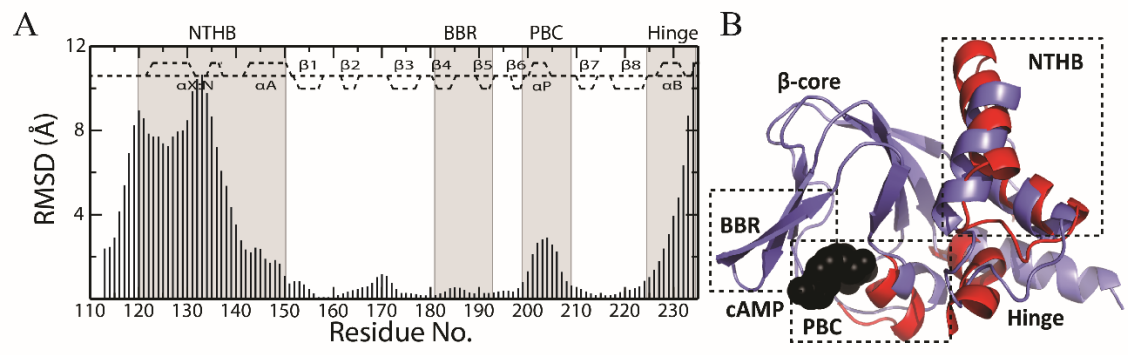

Figure S1: *Allosteric cAMP-dependent structural changes in the well folded cAMP-Domain A of PKA RIα (91-244) as revealed by RMSDs*. **(A)** RMSDs of individual residues measured between the structures of C-subunit bound PKA RIa CBD-A (PDB code 3FHI)[1] and cAMP bound PKA (PDB code 3PNA)[2] aligned at their β-cores (residues 150-225), as shown in panel (B). The dashed lines along the top represent the secondary structure of the protein. **(B)** An overlay of the crystal structures for C-bound (blue) and cAMP-bound (red) PKA aligned at their β-cores. cAMP is shown as black spheres.

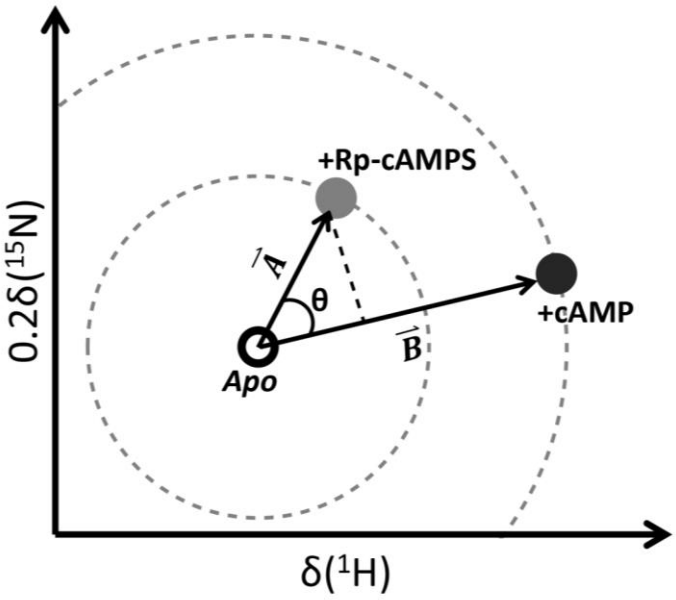

Figure S2: *Schematic representation of the vectors utilized in the CHEmical Shift Projection Analysis (CHESPA)*. The small solid circles represent HSQC peaks for the *apo* (Black, Open), R_p-cAMPS (Rp) bound (Gray, Closed) and cAMP bound (Black, Closed) states. The $\Delta\delta_{NHcomb}$ between the Rp -bound and *apo* states was calculated as the magnitude of vector **A**. Similarly, the $\Delta\delta_{NHcomb}$ between the cAMP-bound and *apo* states was calculated as the magnitude of vector **B**. The angle between vectors **A** and **B** is defined as θ (Equation 6) and the fractional activation (X) is measured by projecting vector **A** onto vector **B** and normalizing the projection to the magnitude of vector **B** (Equation 7).

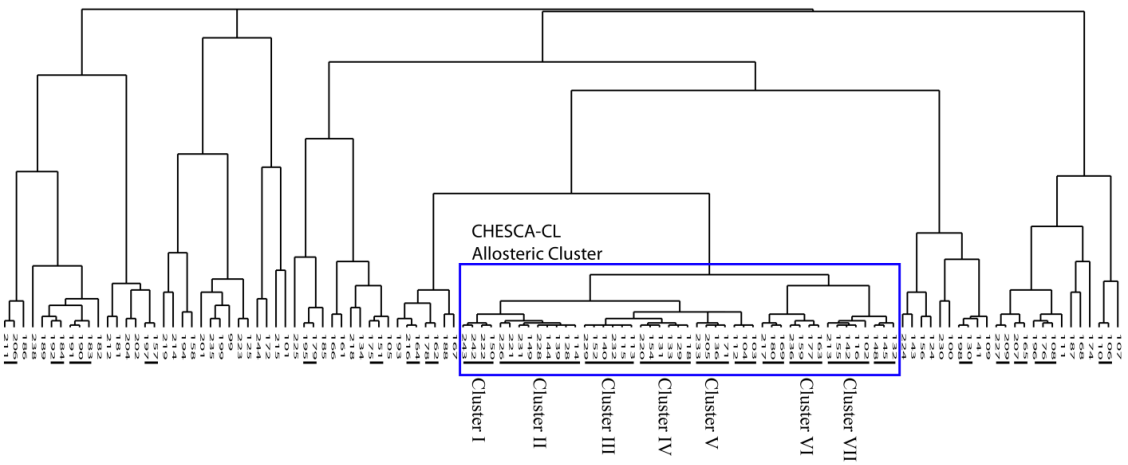

Figure S3: *Whole dendrogram representing the complete-linkage agglomerative clustering of PKA RIα (91-244).* The horizontal lines beneath the dendrogram represent individual fragmented clusters that were computed from complete-linkage clustering using a cutoff of 0.98 for the absolute value of the correlation coefficient. Out of these clusters, seven include more than three residues. These seven clusters are further analyzed in Figure S4 and the smallest branch in the complete dendrogram including them is marked by a blue rectangle. All residues within this branch fit the criteria that define them as a single network (Figure S5) and were therefore reconstituted into a single allosteric cluster.

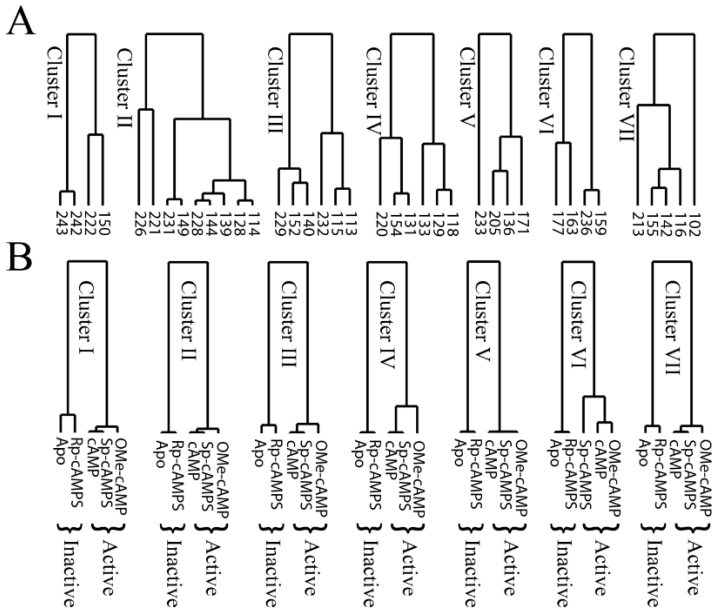

Figure S4: *Complete-linkage clustering results in fragmentation of the allosteric cluster into functionally similar sub-clusters.* **(A)** Residue-based dendrograms generated through complete linkage clusters. The branch height is proportional to the correlation coefficients between clusters of residues. **(B)** State-based dendrograms of the clusters shown in **(A)**. Even though complete-linkage clustering results in the formation of several new sub-clusters, they all share similar functional state distributions (*i.e.* inactive states: *apo* and Rp-cAMPS, separated from the active states: cAMP, Sp-cAMPS and 2'-OMe-cAMP).

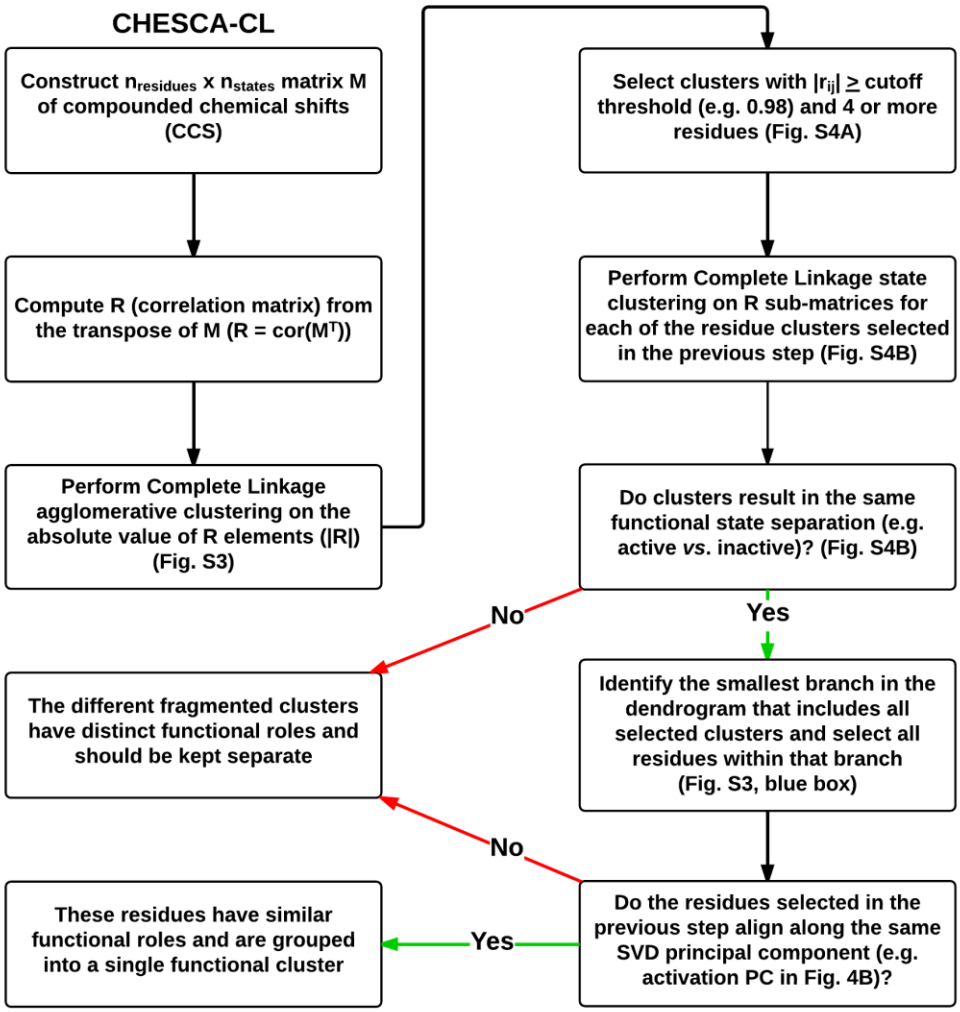

Figure S5: *A protocol for the implementation of the CHESCA-CL analysis including algorithms for the reconstruction of allosteric clusters starting from fragmented sub-clusters generated by complete-linkage*. Selection of allosteric clusters from complete-linkage clustering is achieved by first selecting groups with more than three residues and a Pearson's correlation coefficient ($|r_{ij}| \geq 0.98$ (Figure S4A). If the groups from the previous step have the same functional separation (*i.e.* active *vs.* inactive) in the respective state dendrograms (Figure S4B), the smallest branch of the complete-linkage dendrogram that spans them is identified (blue box in Figure S3) and all residues in this branch are grouped together. If all residues within this group align along the same functional PC (*i.e.* activation in Figure 4B) from the SVD analysis, they are defined as a single allosteric cluster. Correlation and singular value decomposition (SVD) computations are performed similarly to Selvaratnam *et al.* [3].

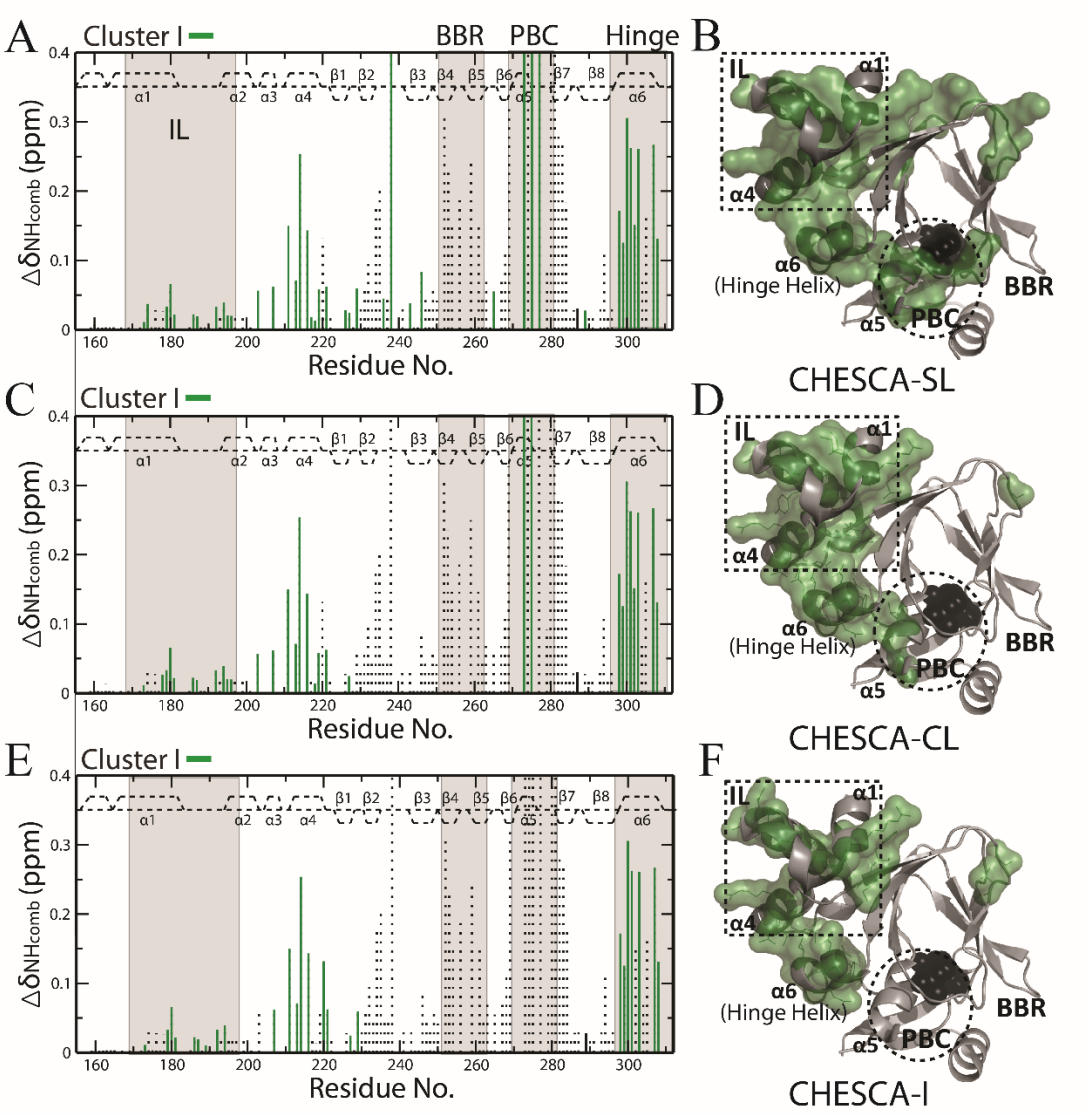

Figure S6: *CHESCA toolset applied to EPAC1$_h$.* **(A)** CCS plot for cAMP binding to EPAC1 (149-318) with residues from the CHESCA-SL allosteric cluster displayed as solid green lines. **(B)** Green residues from panel (A) mapped onto the crystal structure of the homologous EPAC2 (PDB ID: 3CF6)[4]. Dashed lines mark the ionic latch (IL), critical for EPAC auto-inhibition, and the Phosphate Binding Cassette (PBC), where cAMP binds. **(C)** CCS plot for cAMP binding to EPAC with residues from the CHESCA-CL allosteric cluster displayed as solid green lines. **(D)** Green residues from panel (C) mapped onto the crystal structure of EPAC2. **(E)** CCS plot for cAMP binding to EPAC with residues from the CHESCA-I allosteric cluster displayed as solid green lines. **(F)** Green residues from panel (E) mapped onto the crystal structure of EPAC2.

Chapter 5

Conclusions and Future Directions

5.1 Overview of the Previous Chapters

Mapping allosteric interactions and pathways with high resolution structure determination techniques is instrumental for understanding not only the function of allosteric proteins, but also their regulation by allosteric effectors and their dysfunction by disease causing mutations. High resolution structures permit simple comparisons between two or more states that reveal proponents of allosteric conformational changes. It is also an effective way to map the interactions between an allosteric effector and its target receptor. Therefore, structure determination holds enormous predictive power for drug development as it permits rationale-based design of ligand derivatives with potential to target specific allosteric hotspots that modulate biological activity. However, modeling allostery based entirely on structure determination is challenging from both a biological and experimental standpoint. For instance, dynamics can negatively impact the implementation and/or analysis of traditional structural determination methods[1–7]. Intrinsically disordered regions (IDRs) are difficult to crystallize due to their conformational disorder that creates structural variability and impairs self-organization into crystals[8–10]. Even when it is possible to crystallize IDRs, the disordered regions can be poorly represented by electron density or influenced by crystal packing artefacts to adopt conformations that are not native in solution[8–10]. IDRs are not strictly unsuitable for structure determination though as NMR spectroscopy is generally effective for IDR structure elucidation[9,10]. However, it is partly dependent on the time scale of motions involved. Dynamics in the μs-ms range cause line broadening that impede chemical shift and NOESY based structure determination experiments[2,11,12].

From the experimental standpoint, most structure determination methods require high, non-physiological concentrations of protein and substrate. In some instances when the protein is conformationally unstable, these high concentrations can result in precipitation or forced oligomerization[8]. The high concentrations required for structure determination are also a major limitation in systems with low affinity or poorly soluble ligands, such as EPAC and its uncompetitive inhibitor CE3F4R (discussed in chapter 3). CE3F4R's poor solubility limits the maximum concentration that can be added to EPAC and its weak affinity implies that the system can not reach saturation under conditions that are necessary for structure determination.

Conformational heterogeneity, such as that between free and bound states, often hinders crystallography and cryo-EM[8]. NMR spectroscopy is sometimes effective at characterizing such systems since chemical shifts and NOEs are population weighted averages for fast exchanging equilibria[2,10,12,13]. However, CE3F4R's binding is close to the intermediate exchange regime and therefore subject to line broadening, which reduces sensitivity and makes intermolecular NOEs undetectable.

To overcome such challenges and to effectively model allostery in systems unamenable to traditional structure determination, the best strategy is to comprehensively model the thermodynamic cycle with a combination of structure, dynamics and kinetics. Structure does not always require high resolution approaches. In challenging systems, resolution can be traded for sensitivity. For example, EPAC inhibition by CE3F4R was initially probed by bioluminescence resonance energy transfer to detect changes in the relative orientations of the regulatory and catalytic domains[14,15]. In our work, because of the low sensitivity of intermolecular NOESY experiments caused by intermediate exchange broadening, we used ligand to protein STDs to map the binding site of CE3F4R. The experiment results in a larger range of transfer due to additional spin diffusion and reduces the information that could be potentially gained about specific ligand to protein NOE that would otherwise be obtained with traditional NOESY experiments, but it provides a substantial increase in sensitivity and selectivity because of the resonance-dependent saturation of CE3F4R[11]. Similarly, the PRE SLAPSTIC experiments provide extremely sensitive detection of ligand binding, but cover a broad range up to 25Å[16]. However, we describe an approach to increase the resolution of these PRE experiments by including additional spin label sites to triangulate the position of the binding pocket.

Chemical shift changes are perhaps one of the most reliable detectors of structural changes[2,13]. They are readily measured for heterogenous samples, they can potentially be obtained for nearly every atom in the protein and they provide accurate information about secondary structure[2,13]. If high resolution structures are available for some states in the thermodynamic cycle, chemical shifts changes can provide detailed structural information about a novel perturbed state, such as that arising from a mutation or bound ligand. For instance, if the conformation of state A is reflected by a unique set of chemical shifts and the conformation of state B contains its own unique set of chemical shifts, then the conformation of state C can be predicted based on whether the chemical shifts resemble more closely state A or B. This is the essence of the CHESPA analysis described at several points throughout this thesis[2,17–20]. For the HCN4 S672R mutation, the CHESPA analysis

revealed that the C-terminal lid became disengaged in the cAMP-bound state resembling more closely the conformation of the apo state. In chapter 3, the CHESPA was utilized at multiple points to probe the position of EPAC's PBC and hinge helices with regards to mutations, cyclic nucleotide derivatives and the inhibitor CE3F4R.

Chemical shift changes are also effective at identifying novel ligand binding sites, especially when used in conjunction with the CHESPA. For instance, the initial map of CE3F4R-induced chemical shift changes (chapter 3, Fig. 2D) encompassed the binding site that was later determined from STDs and PREs. Independently, chemical shift perturbations can not distinguish between ligand binding and allosteric conformational changes, but when compared to a suitable set of reference states, chemical shifts provide a reliable map of a ligand's binding site. This is because novel perturbations (such as those arising from the binding of a unique ligand) create chemical shift changes in surrounding areas that do not strictly adhere to the equilibrium of the reference states. This results in non-linear chemical shift changes, which can aptly be measured by the CHESPA cosθ analysis. When applied to CE3F4R, it provided a narrow map of the predicted binding site that was in agreement with STDs and PREs.

When structures are unavailable, chemical shifts can still provide information about protein tertiary interactions using the CHESCA analysis (discussed in chapter 4)[2,18,21–23]. The main reporters of the CHESCA are amide chemical shifts, which are excellent sensors of protein structure[13]. When a set of perturbations cause concerted structural changes to regions in a protein, the chemical shifts reflect this in a correlated manner. This is exemplified by the comparisons of CHESCA networks in PKA and EPAC with the local RMSDs from their corresponding crystal structures (chapter 4, Fig. 6)[23]. The additional advantage of the CHESCA is that it works well for dynamic and intrinsically disordered regions/proteins, which are difficult to solve by traditional structure determination. For instance, the biological function of the PKA's disordered N-terminal linker (discussed in chapter 4) was poorly characterized due to the lack of reliable structural information from x-ray crystallography[21]. However, its removal causes extensive perturbations to the adjacent CBD marked by a constitutive shift in the cAMP-bound equilibrium towards inactivation[21]. The CHESCA discovered parts of the linker that responded in a concerted manner to regions in the adjacent CBD that controlled the auto-inhibitory equilibrium[21]. The CHESCA further guided the design of PKA mutations in the linker that disrupted the network between the linker and the CBD and shifted the regulatory domain towards inactivation, thereby inhibiting catalytic activity[18,21].

When structure cannot explain function, it is necessary to examine dynamics. Allostery is a dynamic process that requires analysis of both highly populated ground states, amenable to structure determination, and transient, scarcely populated excited states that are difficult to observe experimentally[1–7]. When the free energy difference between the ground and excited states is relatively small, allostery can be driven by conformational entropy[2,6]. This phenomenon, known as dynamically driven allostery, explains how biological properties of a protein can change without detectable structural changes[2,6]. The S672R HCN4 mutation, discussed in chapter 2, provides a good example of this effect. Neither the X-ray crystal structure nor the NMR secondary structure analysis reveal any significant structural changes relative to the wild type, other than those within the immediate vicinity of the mutation site, but the mutation results in long range allosteric control over the channel's gating properties[17,24,25]. The dynamic analysis of the HCN4 S672R mutation revealed that it shifts the equilibrium of the cAMP-bound state towards inactivation by ~10%. It also increases the conformational entropy of the apo form and enhances internal motions of many residues around the cAMP binding pocket. The increased conformational entropy of the apo state led us to predict that it would decrease the free energy barrier between the apo and cAMP-bound states and result in faster exchange kinetics of binding. This hypothesis was proven using NMR derived measurements of k_{on} and k_{off} rates for cAMP, thus explaining how this mutation modulates the channel's gating kinetics despite its distance from the channel gating domain and its insignificant effect on cAMP binding affinity. This latter point also indicates why it is important to measure kinetics, since it shows that changes in k_{on} and k_{off} rates, even in the absence of changes in K_d can drastically alter the phenotype of the protein and result in disease. Kinetic constants of binding are also important pharmacological properties to describe the efficacy of drug leads[26,27].

5.2 Significance of this Thesis

There are two significant implications from this thesis. The first important contribution is the methodologies that were developed and employed to map the allosteric sites and pathways in challenging systems that are otherwise unsuitable for traditional structure determination. The development and refinement of chemical shift analyses such as the CHESPA and CHESCA were instrumental in understanding the difficult systems discussed in this thesis. For instance, mapping the residues exhibiting non-linear chemical shift changes with the CHESPA cosθ analysis was integral in narrowing down the binding site of CE3F4R (chapter 3) and for filtering out residues subject to nearest neighbor effects for the S672R mutation when calculating the inactive and active

state populations (chapter 2). It was also invaluable for discovering the chemical shift projection artefacts that were causing false positives in the CHESCA (chapter 4). We also developed a quick and easy approach for measuring exchange rates of ligand binding from HSQC titration data (chapter 2). Unlike other protein-based NMR experiments for measuring exchange rates, such as Nz-exchange, CPMG relaxation dispersion and chemical exchange saturation transfer (CEST), this HSQC-based approach, although less quantitative, can cover a broader range of time scales. Finally, the new protocols that were designed for the CHESCA (chapter 4) drastically increase the stringency of the analysis reducing false positives and other artefacts that cause ambiguities in the maps of allosteric hotspots.

The experimental approaches discussed in chapter 3 are also suitable for characterizing the mechanism of action for drug leads, including hydrophobic compounds that have poor solubility. A noteworthy approach from these experiments was the use of mutations and cyclic nucleotide analogs to circumvent the experimental limitations imposed by CE3F4R. Since it was challenging to directly unveil the effects caused by CE3F4R on the structural ensemble of EPAC, the mutations and cAMP analogs provided a way of selectively trapping EPAC conformations in a known manner to observe how CE3F4R would interact with those conformations. The PRE triangulation approach was also an important development, which is analogous to GPS, where multiple satellites (spin labels) at different positions provide precise positioning of the target site. In our experiments, we only used two PRE sites to confirm the position of the CE3F4R binding pocket, proposed from chemical shifts and ligand-to-protein STDs, but with additional sites it may be possible to precisely identify the bounds of the binding pocket even without sufficient supporting evidence.

The other important implication of this thesis is from the elucidated mechanisms of the S672R mutation in HCN4 and the CE3F4R inhibition of EPAC. The HCN4 S672R mutation is a loss-of-function, disease-causing mutation with no current treatment. However, the mechanism of this mutation deduced from this study provides a potential avenue for pharmacologically correcting the disease phenotype. One of the main characteristics of this mutation is that it enhances the binding kinetics of HCN's allosteric effector, cAMP, which results in faster channel deactivation kinetics. Therefore, a potential approach to compensate for the faster binding kinetics is to utilize a cAMP analog with slower on/off binding kinetics.

The mechanism of EPAC inhibition by CE3F4R also has several translational implications. First, EPAC is a target for wide range of diseases, including pancreatic cancer and heart failure[28–30], so pharmacological modulators have several therapeutic applications. Multiple classes of EPAC

selective inhibitors have been discovered[14,15,30–33], but CE3F4R is the only one with mechanism of action that is not competitive with cAMP. Therefore, it will not lose efficacy during times when cAMP concentrations are increased. The identification of the CE3F4R binding site and the structure activity relationships for CE3F4R analogs can also assist rationale-based modifications to develop next generation CE3F4R derivatives with potentially stronger potencies. Finally, to the best of our knowledge, the study reports the first mechanism of action for an uncompetitive inhibitor that is uncompetitive with respect to an allosteric effector as opposed to an enzyme's substrate.

5.3 References

(1) Ward, A. B.; Sali, A.; Wilson, I. A. Biochemistry. Integrative structural biology. *Science (New York, N.Y.)* **2013**, *339*, 913–915.

(2) Boulton, S.; Melacini, G. Advances in NMR Methods To Map Allosteric Sites: From Models to Translation. *Chemical reviews* **2016**, *116*, 6267–6304.

(3) van den Bedem, H.; Fraser, J. S. Integrative, dynamic structural biology at atomic resolution—it's about time. *Nature Methods* **2015**, *12*, 307.

(4) Baldwin, A. J.; Kay, L. E. NMR spectroscopy brings invisible protein states into focus. *Nature chemical biology* **2009**, *5*, 808–814.

(5) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. The ensemble nature of allostery. *Nature* **2014**, *508*, 331.

(6) Popovych, N.; Sun, S.; Ebright, R. H.; Kalodimos, C. G. Dynamically driven protein allostery. *Nature structural & molecular biology* **2006**, *13*, 831–838.

(7) Cooper, A.; Dryden, D. T. Allostery without conformational change. A plausible model. *European biophysics journal : EBJ* **1984**, *11*, 103–109.

(8) Deller, M. C.; Kong, L.; Rupp, B. Protein stability: A crystallographer's perspective. *Acta crystallographica. Section F, Structural biology communications* **2016**, *72*, 72–95.

(9) Dyson, H. J.; Wright, P. E. Intrinsically unstructured proteins and their functions. *Nature reviews. Molecular cell biology* **2005**, *6*, 197–208.

(10) Konrat, R. NMR contributions to structural dynamics studies of intrinsically disordered proteins. *Journal of magnetic resonance (San Diego, Calif. : 1997)* **2014**, *241*, 74–85.

(11) Reibarkh, M.; Malia, T. J.; Hopkins, B. T.; Wagner, G. Identification of individual protein-ligand NOEs in the limit of intermediate exchange. *Journal of biomolecular NMR* **2006**, *36*, 1–11.

(12) Kleckner, I. R.; Foster, M. P. An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica et biophysica acta* **2011**, *1814*, 942–968.

(13) Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Progress in nuclear magnetic resonance spectroscopy* **2013**, *73*, 1–16.

(14) Courilleau, D.; Bisserier, M.; Jullian, J.-C.; Lucas, A.; Bouyssou, P.; Fischmeister, R.; Blondeau, J.-P.; Lezoualc'h, F. Identification of a tetrahydroquinoline analog as a pharmacological

inhibitor of the cAMP-binding protein Epac. *The Journal of biological chemistry* **2012**, *287*, 44192–44202.
(15) Courilleau, D.; Bouyssou, P.; Fischmeister, R.; Lezoualc'h, F.; Blondeau, J.-P. The (R)-enantiomer of CE3F4 is a preferential inhibitor of human exchange protein directly activated by cyclic AMP isoform 1 (Epac1). *Biochemical and Biophysical Research Communications* **2013**, *440*, 443–448.
(16) Jahnke, W.; Rüdisser, S.; Zurini, M. Spin label enhanced NMR screening. *Journal of the American Chemical Society* **2001**, *123*, 3149–3150.
(17) Boulton, S.; Akimoto, M.; Akbarizadeh, S.; Melacini, G. Free energy landscape remodeling of the cardiac pacemaker channel explains the molecular basis of familial sinus bradycardia. *The Journal of biological chemistry* **2017**, *292*, 6414–6428.
(18) Boulton, S.; Akimoto, M.; Selvaratnam, R.; Bashiri, A.; Melacini, G. A tool set to map allosteric networks through the NMR chemical shift covariance analysis. *Scientific reports* **2014**, *4*, 7306.
(19) Akimoto, M.; Zhang, Z.; Boulton, S.; Selvaratnam, R.; VanSchouwen, B.; Gloyd, M.; Accili, E. A.; Lange, O. F.; Melacini, G. A mechanism for the auto-inhibition of hyperpolarization-activated cyclic nucleotide-gated (HCN) channel opening and its relief by cAMP. *The Journal of biological chemistry* **2014**, *289*, 22205–22220.
(20) Selvaratnam, R.; VanSchouwen, B.; Fogolari, F.; Mazhab-Jafari, M. T.; Das, R.; Melacini, G. The projection analysis of NMR chemical shifts reveals extended EPAC autoinhibition determinants. *Biophysical journal* **2012**, *102*, 630–639.
(21) Akimoto, M.; Selvaratnam, R.; McNicholl, E. T.; Verma, G.; Taylor, S. S.; Melacini, G. Signaling through dynamic linkers as revealed by PKA. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110*, 14231–14236.
(22) Boulton, S.; Selvaratnam, R.; Ahmed, R.; Melacini, G. Implementation of the NMR CHEmical Shift Covariance Analysis (CHESCA): A Chemical Biologist's Approach to Allostery. *Methods in molecular biology (Clifton, N.J.)* **2018**, *1688*, 391–405.
(23) Selvaratnam, R.; Chowdhury, S.; VanSchouwen, B.; Melacini, G. Mapping allostery through the covariance analysis of NMR chemical shifts. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108*, 6133–6138.
(24) Xu, X.; Marni, F.; Wu, S.; Su, Z.; Musayev, F.; Shrestha, S.; Xie, C.; Gao, W.; Liu, Q.; Zhou, L. Local and global interpretations of a disease-causing mutation near the ligand entry path in hyperpolarization-activated cAMP-gated channel. *Structure (London, England : 1993)* **2012**, *20*, 2116–2123.
(25) Milanesi, R.; Baruscotti, M.; Gnecchi-Ruscone, T.; DiFrancesco, D. Familial sinus bradycardia associated with a mutation in the cardiac pacemaker channel. *The New England journal of medicine* **2006**, *354*, 151–157.
(26) Vauquelin, G. Effects of target binding kinetics on in vivo drug efficacy: Koff , kon and rebinding. *British journal of pharmacology* **2016**, *173*, 2319–2334.

(27) Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug-target residence time and its implications for lead optimization. *Nature reviews. Drug discovery* **2006**, *5*, 730–739.
(28) Parnell, E.; Palmer, T. M.; Yarwood, S. J. The future of EPAC-targeted therapies: Agonism versus antagonism. *Trends in pharmacological sciences* **2015**, *36*, 203–214.
(29) Schmidt, M.; Dekker, F. J.; Maarsingh, H. Exchange protein directly activated by cAMP (epac): A multidomain cAMP mediator in the regulation of diverse biological functions. *Pharmacological reviews* **2013**, *65*, 670–709.
(30) Almahariq, M.; Tsalkova, T.; Mei, F. C.; Chen, H.; Zhou, J.; Sastry, S. K.; Schwede, F.; Cheng, X. A novel EPAC-specific inhibitor suppresses pancreatic cancer cell migration and invasion. *Molecular pharmacology* **2013**, *83*, 122–128.
(31) Tsalkova, T.; Mei, F. C.; Li, S.; Chepurny, O. G.; Leech, C. A.; Liu, T.; Holz, G. G.; Woods, V. L.; Cheng, X. Isoform-specific antagonists of exchange proteins directly activated by cAMP. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 18613–18618.
(32) Brown, L. M.; Rogers, K. E.; Aroonsakool, N.; McCammon, J. A.; Insel, P. A. Allosteric inhibition of Epac: Computational modeling and experimental validation to identify allosteric sites and inhibitors. *The Journal of biological chemistry* **2014**, *289*, 29148–29157.
(33) Brown, L. M.; Rogers, K. E.; McCammon, J. A.; Insel, P. A. Identification and validation of modulators of exchange protein activated by cAMP (Epac) activity: Structure-function implications for Epac activation and inhibition. *The Journal of biological chemistry* **2014**, *289*, 8217–8230.