**Molecular Mechanisms for Inhibition of Regulators of G-protein Signaling
by Small Molecules**

by

Andrew J. Storaska

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Pharmacology)
in The University of Michigan
2014

Doctoral Committee:

        Professor Richard R. Neubig, Co-Chair, Michigan State University
        Professor John J.G. Tesmer, Co-Chair
        Assistant Professor Tomasz Cierpicki
        Professor Margaret E. Gnegy

For my parents, Jessie, and whole family

Acknowledgements

I would very much like to thank my mentor Dr. Rick Neubig for his constant help and attention to the needs of my project, and his genuine interest in my scientific development. I also appreciate his continual efforts to promote and facilitate the collaborations that have allowed me to expand my knowledge into new areas. I would like to thank my committee, Dr. John Tesmer, Dr. Margaret Gnegy, and Dr. Tomasz Cierpicki, for their support and guidance throughout my project. I thank and appreciate everyone I have worked with in the Neubig Lab over the past four years. Specifically, I thank Dr. Levi Blazer for his early support and guidance, to Sue Wade for all her help and being the "knower" of everything pertaining to the Neubig Lab, I thank Dr. Benita Sjogren for being a good friend and labmate, and Andrew Haak for keeping things interesting in the lab. I also thank Raelene Van Noord, Jason Kehrl, and Dr. Sergio Parra for all their help. I owe a big thank you to Jian Mei who contributed to my projects over a couple of years.

I really appreciate the help provided by all of my collaborators and everyone in the Department of Pharmacology. I am very grateful to the Cierpicki lab for their help with NMR studies over the course of my projects. In addition the

Table of Contents

List of Figures

List of Tables

List of Appendices

List of Abbreviations

1-8ANS: 1-anilinonapthalene-8-sulfonic acid

AF532: AlexaFluor-532

AlCl$_3$: aluminum chloride

ANOVA: Analysis of variance

BSA: Bovine serum albumin

CCG: Center for Chemical Genomics, University of Michigan

CNS: Central nervous system

DMEM: Dulbecco's Modified Eagle Medium

DMSO: Dimethyl sulphoxide

Dox: doxycycline

DRC: Dose-response curve

CPMG: Carr Purcell Meiboom Gill

CRC: Concentration-response curve

CSP: Chemical Shift Perturbation

DTT: dithiothreitol

FCPIA: flow-cytometry protein interaction assay

GAP: GTPase accelerating protein

GDP: Guanosine-5' diphosphate

GEF: Guanine nucleotide exchange factor

GPCR: G protein coupled receptor

GTP: Guanosine-5' triphosphate

HEK: Human embryonic kidney cells

HEPES: 4-(2-hydroxyethyl)-1piperazineethanesulfonic acid

His6: 6x histidine tag

HSQC: Heteronuclear Single Quantum Coherence

IC$_{50}$: Inhibitory concentration, 50%

IP$_6$: Inositol hexakisphosphate, phytic acid

MD: Molecular Dynamics

MgCl$_2$: magnesium chloride

ml: milliliter

MLSMR: Molecular Libraries Small Molecule Repository

NA: Not available

NaCl: sodium chloride

NaF: Sodium fluoride

NIH: National Institutes of Health

NMR: nuclear magnetic resonance spectroscopy

PDB: Protein Databank

PIP$_3$: PI(3,4,5)P$_3$-phosphotidyl inositol

PPI: Protein-protein interaction

RGS: Regulator of G Protein Signaling

SAR: Structure-activity relationship

SD: Standard deviation

SEM: standard error of the mean

SS: Steady-state

TAMD: Temperature Accelerated Molecular Dynamics

TCEP: tris(2-carboxyethyl) phosphine

Tm: melting temperature

TR-FRET: time-resolved fluorescence resonance energy transfer

µl: microliter

Abstract

**Molecular Mechanisms for Inhibition of Regulators of G-protein Signaling
by Small Molecules**

by

Andrew J. Storaska

Co-Chair: Richard R. Neubig
Co-Chair: John J.G. Tesmer

Regulator of G-protein signaling (RGS) proteins potently suppress G-protein coupled receptor (GPCR) signal transduction by accelerating GTP hydrolysis on activated heterotrimeric Gα proteins. RGS proteins have become attractive targets for the purpose of manipulating GPCR-mediated cellular responses. The RGS family comprises thirty-seven proteins with differential expression patterns throughout the body. RGS4 is enriched in the CNS and has been proposed as a therapeutic target for treatment of neurological disorders including epilepsy and Parkinson's disease. Therefore, our lab has focused on the identification of small molecule inhibitors of RGS4. To date, all small molecule inhibitors of RGS proteins function through covalent modification of cysteine residues, yet substantial specificity has been observed for RGS4 over other closely related homologs. The work in this thesis

details the molecular mechanism of inhibition by a potent RGS4 inhibitor (CCG-50014; IC_{50} = 30 nM), and reveals the importance of RGS4 dynamics in the exposure of key cysteine residues that upon binding the inhibitor prevent the protein from reaching native conformations. Elucidating this mechanism has allowed us to propose a novel cryptic site (C-site) that is formed around the buried cysteine residues, and may be more druggable than previously proposed sites on RGS4. In addition, new chemical scaffolds have been identified using a cell-based high-throughput screen that also inhibit RGS4 through a cysteine-dependent mechanism, but are significantly reversible in contrast to the first cell-active RGS4 inhibitors. Furthermore, I employed IP_6 as a derivative of an endogenous negative regulator, PIP_3, to map the binding site on RGS4 and the corresponding effects on protein stability and function. This study shows the direct interactions of a non-covalent small molecule with an allosteric site on the RGS4 structure, and provides insight into the mechanism of endogenous regulation of RGS4. In conclusion, the studies described within this thesis provide new pharmacological tools for the study of RGS function in a cellular context, and describe in detail the molecular interactions of both endogenous and pharmacological modulators of RGS4. These results provide the theoretical framework to pursue drug discovery strategies that are expected to significantly advance the field of RGS drug discovery.

Chapter I

Introduction

The identification and design of small molecule modulators of protein-protein interactions (PPIs) represents a significant challenge for drug discovery projects based on the physical characteristics generally associated with non-enzyme and non-receptor targets. The potential to expanding the number of therapeutic targets and treat a wider range of disease has resulted in substantial interest in pursuing modulators of multi-protein complexes as therapy. The work described in this thesis elucidates a novel model of allosteric inhibition by a PPI inhibitor, and identifies new inhibitors for one therapeutically intriguing PPI involving RGS proteins. As a result of our expanding knowledge of PPI networks, the number of potential therapeutic targets continues to grow. These studies are expected to contribute a significant advancement toward the understanding needed to meet the new physical and biological challenges associated with modulating PPIs.

Therapeutic Relevance of Protein-protein Interactions

Therapeutic development of small molecules has traditionally focused on a narrow subset of molecular targets, with class I GPCRs, voltage and ligand-gated ion channels, and nuclear receptors representing a substantial portion of

1

drug-targeted biomolecules [1]. In addition to these historical targets, a growing number of protein kinases are now successfully being targeted [2]. Current drug data from the Therapeutic Target Database lists 364 successful therapeutic targets of approved drugs in its most recent update [3]. The number of proteins targeted by approved drugs is at odds with the estimated 3,000 human proteins associated with disease through inherited mutations [4], and the total number of disease causing proteins suggested to range between 3,000-10,000 human proteins [5]. It is clear that pursuing traditional targets falls short of ameliorating a large portion of aberrant protein function. This should not be a surprise as cellular responses initiated from a receptor are usually propagated downstream through a complex network of PPIs.

The networks of PPIs that mediate cellular signal transduction can be quite large, which provides an extensive range of possible therapeutic targets. TGFβ-mediated signal transduction pathways are one prominent example of this, where huge networks are linked to a number of important cellular functions including cellular growth, differentiation, and morphogenesis [6]. Disruptions of these pathways are associated with inflammatory diseases, fibrosis, and cancer [6]. TGFβ signaling responses are mediated in part through SMAD protein interaction networks involved in localization of SMAD proteins to various subcellular compartments, including membrane, receptor, and eventually nuclear localization for transcriptional activity [7]. A large-scale yeast two-hybrid PPI mapping study of the SMAD protein signal transduction network resulted in the identification of 755 PPIs involving 591 different proteins [8]. Focusing on PPIs as

2

therapeutic targets creates the daunting task of cataloguing interactions through a signaling pathway and determining which are most relevant as therapeutic targets. The scopes of the networks are often large making it difficult to identify the optimal molecular target within the context of an aberrant signaling pathway, but the application of advanced proteomics and the use of comprehensive PPI databases have made achieving this more feasible [4, 8, 9]. The advantages gained by a PPI targeting strategy include the ability to increase the number of disease treatments by expanding the range of therapeutic targets, and improving therapeutic specificity by targeting proteins that may be more heterogeneously expressed relative to receptors.

Regulators of G-protein Signaling as Therapeutic Targets

Signaling through G-protein coupled receptors (GPCRs) controls many vital physiological functions. As a result, therapeutics that target GPCRs constitute a major portion of the prescription drug market [10]. Despite this success, achieving selectivity among closely related GPCR subtypes has limited therapeutic development in many areas, particularly for Alzheimer's disease [11, 12]. One strategy to attain greater selectivity is to modulate receptor activation by engaging less conserved allosteric surfaces on the receptors, either independently [13], or with the use of bitopic molecules that simultaneously bind the orthosteric and allosteric sites [14]. Another approach is to target the downstream PPIs that regulate GPCR signal transduction [15]. Targeting downstream regulatory PPIs provides an ability to more subtly modulate

3
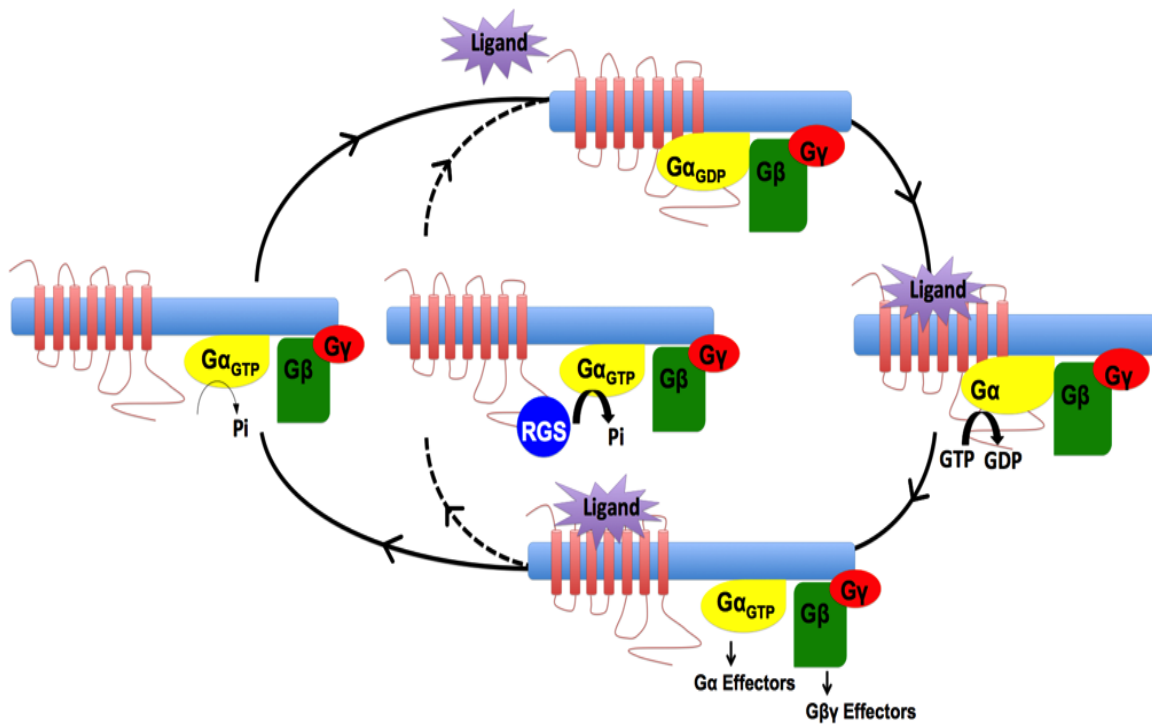
Figure 1.1 GPCR signaling pathway following receptor activation. Ligand binding to the receptors catalyzes nucleotide exchange on the heterotrimeric Gα subunit, which exposes effector interaction surfaces on Gα and Gβγ allowing them to mediate signal transduction. RGS protein GAP activity on the Gα subunits enhances nucleotide hydrolysis over 1000-fold [17-20], thereby decreasing the duration of the signaling cycle (dashed line).

endogenous signaling, similar to the actions of positive or negative allosteric modulators of GPCRs, although one unique advantage is that therapeutic action can also be compartmentalized in distinct tissues where both the receptor and target protein are co-expressed [16].

Regulators of G-protein Signaling (RGS) suppress GPCR signal transduction by selectively interacting with GTP-bound (activated) heterotrimeric Gα subunits to enhance their intrinsic rate of nucleotide hydrolysis [17-20], **Figure 1.1**. The expression profiles among the more than 20 human RGS proteins are relatively diverse, with some homologs being highly expressed in specific tissues such as RGS4 in the CNS [21, 22]. Within these tissues, RGS

proteins can co-localize with specific receptors [23-26], which may in part drive the receptor-specific regulation by certain RGS homologs [24, 27]. Pathway-specific regulation also results from the preference of certain RGS homologs for $G\alpha i/o$ or $G\alpha q$ family members [28-31]. As a result of the expression patterns and pathway-specific effects, modulating GPCR signaling up or down in a particular tissue could be achieved by inhibiting or activating a specific RGS homologs. Therefore, RGS proteins have been proposed as intriguing drug targets [32-34].

RGS4 is highly expressed in cortex, thalamus, and other brain regions [21], and potentially affects numerous centrally-acting GPCR signaling pathways. Within the dorsolateral striatum, RGS4 serves as a bridge between D_2-dopamine and A_2-adenosine receptors and endocannabinoid mobilization driving striatal plasticity associated with normal motor behavior. As a result, RGS4 knockout mice are less sensitive than WT animals to motor behavior deficits occurring from 6-OHDA depletion of dopamine [35]. This suggests that RGS4 may be a new target for treating Parkinson's disease. Additionally, formation of an RGS4-A_1-adenosine receptor complex via the neurabin scaffolding protein negatively regulates the neuroprotective effects of adenosine signaling in a kainate-induced seizure model. Genetic knockout of neurabin or small molecule antagonism of RGS4 reduces seizure severity in this model [36]. In either case, inhibition of RGS4 provides a beneficial enhancement of a particular GPCR signaling pathway in the context of these models. Such studies support the use of RGS inhibitors in therapy. As a result there is a critical need for continued development of selective small molecule RGS modulators.

Challenges Associated with Modulating Protein-protein Interactions

The formation of protein complexes is a ubiquitous mechanism driving cellular signaling events. For this reason small molecule modulators of PPIs are needed both as a therapeutic agents and as probes to identify the function of signal transduction proteins. The significant challenges associated with developing successful small molecules targeting PPIs have limited progress in both areas. Part of this challenge is a result of the overall structural topology generally found in many PPI interfaces. These interfaces typically span large surface areas greater than 1000 $Å^2$ [37]. As a result, small molecules with sufficient affinity to block PPIs competitively at the interface generally require a large molecular weight. Using a cross section of well-characterized PPI inhibitors, Wells and McClendon found a relatively uniform ligand efficiency of 0.24 kcal • mol^{-1} • non-hydrogen $atom^{-1}$ among the group of inhibitors analyzed, which suggests that PPI inhibitors that reach clinical potency (~10 nM) would have a molecular weight larger than 600 Da [38]. Increases in hydrophobicity and molecular weight occuring as a result of potency optimization can impair solubility and permeability [39], leading to downstream attrition as compounds move towards clinical candidacy [40].

One potential reason PPI inhibitors tend to require larger mass to attain clinical potency may be the relatively shallow pockets formed along a PPI

interface. An analysis by Fuller and colleagues found the average pocket volumes for protein-ligand interactions of currently marketed drugs (260 Å^3) is
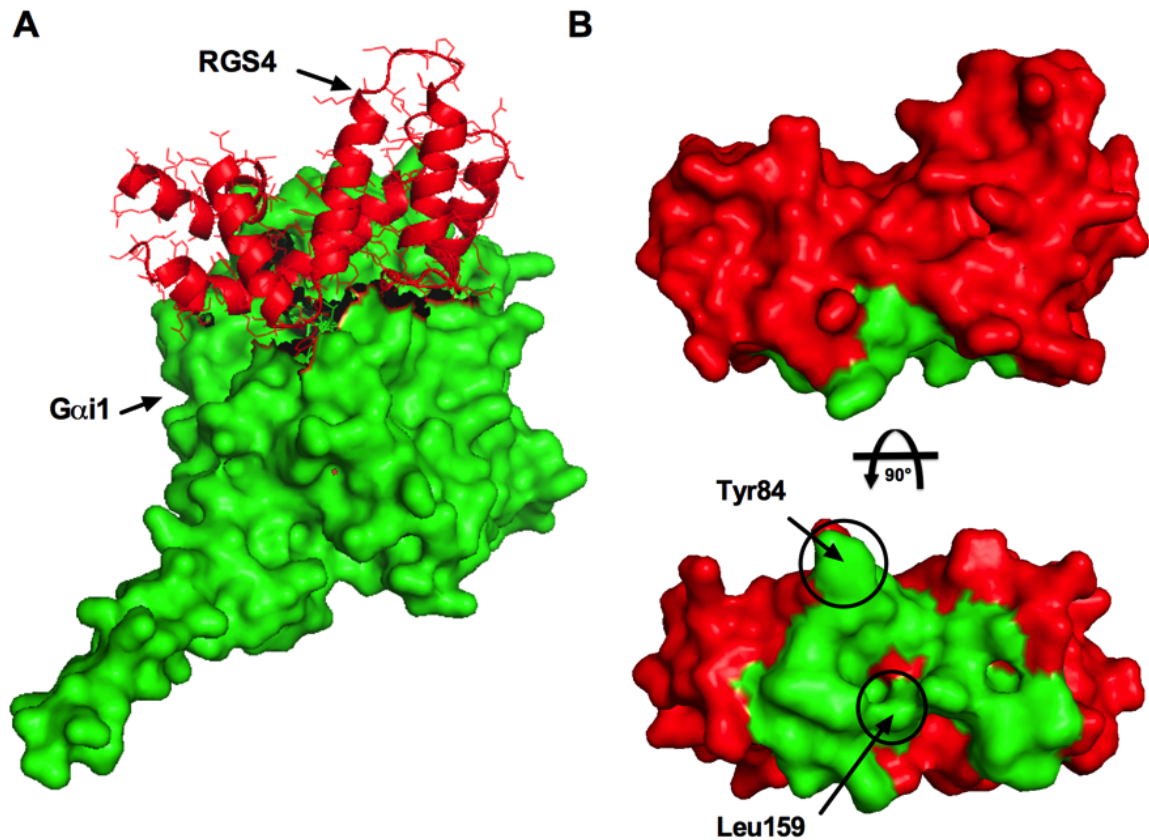
Figure 1.2 The RGS4-Gαi1 binding interface. (PDB code: 1agr). A) The association of RGS4 (red) with Gαi1 (green) forms a majority of the contacts across three switch regions of the G-protein Ras-like domain [40]. B) The G-protein binding interface of RGS4 (highlighted in green) consists primarily of hydrogen-bonding and electrostatic contacts with Gα, with the exception of two hydrophobic residues (Tyr84 and Leu159) in the outer perimeter of the buried interface [40].

approximately five-fold larger than what has been measured across the interfaces of PPIs (54 Å^3) [41]. In addition, protein-drug interactions primarily occur through one or two pockets producing an average occupied volume of 524 Å^3, which is in contrast to a typical PPI interface that has 261 Å^3 of occupied pocket volume across multiple surface pockets [41]. Indeed, a systematic analysis of drug pocket characteristics revealed an overall pocket volume of

7

greater than 593 $\mathring{A}^3$ is required to cover 95% of known drug targets [42]. The physical challenges associated with developing compounds that can bind to PPI interfaces with high potency has likely limited the number of PPI-direct therapeutics [43].

The structural features of RGS proteins are consistent with the challenging characteristics found in most non-traditional drug targets. The interface between the RGS4-Gα$_{i1}$ complex (**figure 1.2A**) is relatively flat, burying 1100 $\mathring{A}^2$ of solvent accessible surface area [44]. In contrast to many PPI interfaces, the interaction is formed primarily through hydrogen-bonding and electrostatic contacts [44]. The majority of the buried hydrophobic surface area results from two resides on the outter perimeter of the G-protein binding interface of RGS4, Tyr84 and Leu159 [44] (**Figure 1.2B**). Sandwiched between Tyr84 and Leu159 is a shallow cavity in the middle of the G-protein binding surface of RGS4 termed the "A-site" [45]. This cavity forms one of two potentially "druggable" sites on RGS4 that have been described previously [33]. The A-site pocket is the smaller of the two, comprising a total volume of 129 $\mathring{A}^3$. While the second pocket termed the "B-site" is 151 $\mathring{A}^3$, and sits within a cleft located on the opposite face of the RGS protein with respect to the G-protein binding interface (**Figure 1.3**). Similar to other PPI targets, these shallow surface pockets found on RGS4 are a major challenge to current drug discovery strategies.

Strategies for Modulating PPIs

Interface Hotspots and Allosteric Pockets

Despite the challenges associated with targeting the relative featureless surface area found in protein-protein interfaces, there are several strategies that can increase success using small molecules. It has been found that protein-protein interfaces contain specific sites that contribute a major portion of the binding energy for the complex. These sites known as "hot spots" are points along the interface that bury large hydrophobic side chains, such as arginine, tyrosine, and tryptophan [46, 47]. As a result, blocking PPIs by targeting these specific sites can make the development of PPI inhibitors more feasible since the molecule needs to cover less overall surface area in order to surmount the interaction. The most widely used method for experimentally determining hot spots across a PPI interface is alanine-scanning mutagenesis [48-50], although computational methods for identifying these specific sites have become more useful [51, 52]. Computational analysis of the RGS4-Gα interaction indicates several sites directly adjacent to the A-site pocket contribute significant binding energy to the formation of the complex [16], which supports targeting the A-site with small molecules as one potential strategy for modulating the interaction.

An approach that may be more promising than directly targeting PPI interfaces may be through the development of small molecule modulators that act at sites that are distant from to the protein-protein interface. Targeting allosteric sites that are functionally coupled [53-55] to the active site provides a
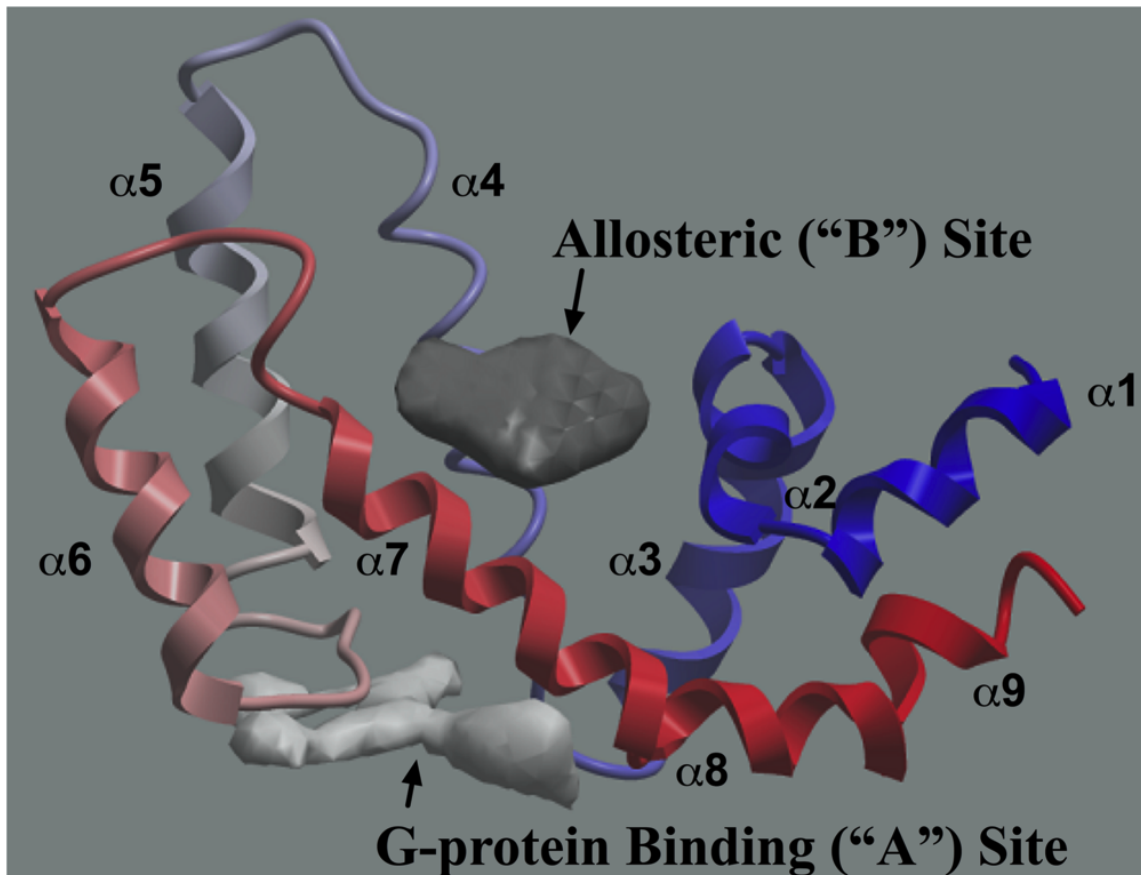
Figure 1.3 Analysis of surface pockets on RGS4, (residues 51-178, PDB code: 1agr). Two potential drug-binding pockets have been previously reported on RGS4 using ICM Pro (Molsoft, San Diego, CA). One pocket in the G-protein binding interface (A-site) comprises a volume of 129 $Å^3$, while a second pocket located on the opposite face of the RGS domain (B-site) is larger with a 151 $Å^3$ volume [33], and may represent a location for allosteric regulation of RGS activity.

mechanism to modulate function without the need to compete with the endogenous binding partner [56, 57]. Additionally, allosteric sites on a protein are typically less conserved since they are not directly involved in protein or ligand interactions. This feature has been particularly useful in the GPCR drug discovery field where selectivity among different receptor subtypes at the orthosteric site has been challenging. The development of allosteric modulators has improved specificity between closely related subtypes [13, 58, 59]. In addition, targeting these sites also provides the ability to modulate activity either

10
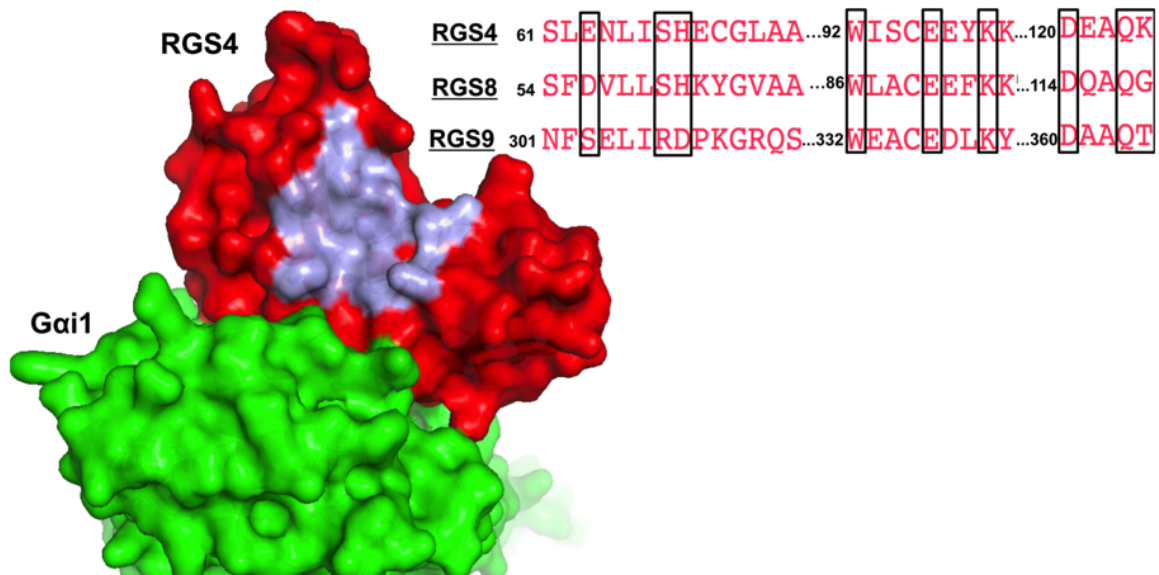
Figure 1.4 Sequence analysis of the residues comprising the RGS B-site pocket across three homologs. The residues on RGS4 (PDB code: 1agr) that form the B-site pocket are highlighted in blue. All three sequences were taken from the human peptide sequence.

up or down, which has led to the development of both positive and negative allosteric modulators of GPCRs [60, 61]. This has also been accomplished recently using disulfide tethering on AGC kinases [62], providing proof-of-concept that differential modulation can be applied to non-receptor targets as well.

An allosteric surface of RGS4 contains a potentially druggable pocket (B-site) with larger volume and more suitable overall geometry than the A-site [33] (**Figure 1.3**). As a result, targeting this site may be advantageous for identifying small molecule interactions. The RGS4 B-site has been identified as a functionally active site on the RGS4 structure. It interacts with phospholipid micelles containing phosphatidylinositol (3,4,5)-triphosphate (PIP$_3$) that inhibit the RGS4 GAP activity [63-65]. In contrast, micelles containing either phosphatidylinositol (4,5)-bisphosphate (PIP$_2$) or L-α-phosphatidyl-choline (PC)

and L-α- phosphatidylethanolamine (PE) alone were unable to inhibit GAP activity [63, 65, 66], suggesting a specific interaction with PIP_3. This interaction is dependent on the presence of positively charged residues (Lys 99 & 100, and to a lesser extent Lys 112 &113), as shown by mutagenesis studies of these potentially important lysine residues [63, 67]. The allosteric surface also appears to form an overlapping binding region with Ca^{2+}/calmodulin, which has been shown to reverse PIP_3-mediated antagonism of GAP activity [63, 64, 66, 67].

The interplay between PIP_3 inhibition of GAP activity and its reversal by Ca^{2+}/calmodulin is suggested to play a role in calcium signaling regulation in cells [67, 68]. Although evidence of the RGS4-PIP_3/Ca^{2+}/calmodulin regulatory mechanism has been demonstrated in cellular systems [64, 69], there is currently no *in vivo* data to support the overall model. Nevertheless, the interactions at this site support the strategy of targeting this surface with small molecules. Since the molecular details of the allosteric mechanism of inhibition by PIP_3 are unknown, it is unclear whether inhibition occurs through a structural change communicated from the allosteric site to the G-protein binding interface, or if the mechanism operates by sequestering the RGS protein in manner that abrogates its interaction with G-proteins. As a result, small molecules that interact in the same manner as PIP_3 could either inhibit the RGS protein through allostery, or alternatively antagonize PIP_3 binding, producing an activation of RGS activity. Identification of small molecules acting at the PIP_3/Ca^{2+}/calmodulin binding site could not only provide starting points for lead molecules, but could also provide improved tools to dissect the relevance of these interactions *in vivo*.

12

An analysis of RGS family peptide sequences surrounding the proposed PIP$_3$/ Ca^{2+}/calmodulin binding site shows strong conservation along the α-helix 4 in the vicinity of the Lys 99 & 100, with more divergence occurring on the N-terminal portion of α-helix 5 and lysine residues 112 & 113 [63]. The strong conservation of residues surrounding the 99/100 positions is indicative of a functional region on the RGS domain. This further supports the advantage of targeting this site to modulate RGS function. Interestingly, an analysis of discontinuous residues that form the B-site pocket shows that side chains comprising the pocket are relatively conserved among three RGS homologs (**Figure 1.4).** A review of additional RGS family members: RSG7, RGS10, and RGS19, similarly shows strong conservation of these pocket residues. As a result, the B-site pocket may represent an interesting site for the development of RGS modulators for a variety of RGS family members. One challenge suggested by the sequence analysis of this site will be acquiring selectivity.

Fragment-based Drug Discovery Methods for Targeting PPIs

Fragment-based drug discovery (FBDD) approaches have emerged as an important strategy for developing small molecule modulators of PPIs. This approach uses fragment-sized compounds (< 300 Da) to identify chemical starting points that interact with small or shallow structural pockets on a protein. The basis of this approach is a focus on efficiency at several levels, rather than a sheer mass of compounds screened on a target in high-throughput. This is achieved in part by the fact that a significantly larger fraction of the relevant

chemical space can be probed using relatively small libraries between 1,000-3,000 compounds. It is estimated that larger (> 450 Da) and more complex drug-like compounds used in HTS can approach 10^{60} possible variations in their chemical space, whereas low molecular weight fragment molecules may only be as diverse as 10^7 different variations [70]. Even though fragment libraries may be one or two orders of magnitude smaller than a commercial or institutional HTS library, the portion of available chemical space applied to a target is fractionally much greater. The biggest advantage FBDD provides for targeting PPIs is the enhanced ligand efficiency that results from screening molecules that form more contacts per unit of atomic mass [71-73]. This rationale has led to the development of a new set of physiochemical parameters for small molecules in fragment screening, described as the "rule of three", including molecules that have < 300 Da, ≤ three hydrogen bond donors/acceptors, ClogP ≤ 3 [74].

Despite the higher ligand efficiency that is a characteristic of fragment starting points, the smaller size compounds still produce relatively weak interactions, typically in micromolar to millimolar affinities in the initial screening process. As a result, direct binding experiments that provide high sensitivity are typically used to screen fragment libraries. Protein-based NMR experiments (e.g. Heteronuclear Single Quantum Coherence, HSQC) are one of the most informative methods, since binding site information is provided as a part of the hit identification process [75, 76]. This formed the basis of "SAR by NMR", where lead compounds are built in a modular fashion by linking fragment molecules that bind in close proximity to one another [76]. More recently, a number of

biophysical techniques have been optimized for FBDD, including surface plasmon resonance (SPR), Thermofluor, and X-ray crystallographic screening [77-79]. Utilization of these techniques has led to therapeutic development of compounds targeting biomolecules previously deemed intractable to drug discovery. One of the earliest examples of success against an intractable target is the development of a Merck inhibitor (MK-8931) against protease β-secretase 1 (BACE1), an enzyme involved in conversion of amyloid precursor protein (APP) into beta-amyloid (Aβ) [80]. The inhibitor was identified in an NMR–based screen of 10,000 fragment molecules, which identified the primary screen hit with a K_d of 550 μM [81]. Lead optimization through crystallographic trials resulted in a compound currently working through PhaseII/III clinical trials. Subsequently, there have been a number of FBDD-derived compounds entering clinical trials [82-84].

As a result of the successes with difficult targets, fragment-based approaches have been suggested as a mechanism to evaluate the druggability of a biomolecule [85]. This assessment is based on the observation that hit rates resulting from fragment screening occur in the 1% range, in contrast to typical HTS screens that return ~0.01% of the library as validated hits. An analysis by Edfeldt and colleagues analyzed a cross section of 36 drug discovery projects that employed fragment screens, allowing them to produce a "ligandability" score based on number of hits, diversity, and affinities [86]. In targets with a low score (hit rate < 0.1%, affinities > 1mM, low diversity), 100% (12/12) projects failed to identify hits using a traditional HTS screen [86]. This would suggest that any

15

molecular target that fails in a fragment screening campaign might be considered undruggable by conventional methods. However, examples exist where results from fragment screening alone were not sufficient to preclude certain targets. A study by Dias et al, found that fragment substituents of a known PPI inhibitor were not detected in an NMR-based fragment screen [87]. Using several biophysical techniques (including three separate ligand-guided NMR experiments) and successive sensitivity optimizations, the authors were able to confidently detect the weakly binding chemical substituents and correlate the biophysical observations with functional PPI displacement [87]. As a result a multifaceted approach may be required to successfully approach difficult targets like RGS proteins, which creates the challenge of utilizing several different techniques to measure binding with a reasonably sized fragment library.

Current RGS Inhibitors

Despite the significant challenge associated with modulating RGS protein activity, several different classes of inhibitors have been reported to date (**Table 1.1**). The first RGS inhibitors were rationally designed based on the peptide sequence of the Gαi switch I region, exemplified by YJ34 [88, 89]. This region of Gαi forms three critical contacts with RGS4, including a Gly183 residue on Gαi [44], that is known to abrogate binding if mutated to serine [90]. Based on the RGS-Gα crystal structural, it is expected the switch 1-mimetics physically occlude the PPI by binding around the RGS4 A-site. This is supported by the loss of activity with a modified peptide that contains a serine residue at the glycine position in YJ34 [89]. A second group identified an RGS4 peptide inhibitor using

a yeast two-hybrid screen, Peptide17 [91]. Although the physical mode of inhibition is unclear, the authors were able to show loss of activity towards RGS4 with a scrambled version of the peptide [91]. The disadvantage of the peptide inhibitors is a lack of potency, solubility, and cellular permeability, which precluded further development of these compounds.

The evolution of RGS inhibitors improved significantly with the identification of novel small molecule scaffolds. The first small molecule RGS inhibitor (CCG-4986) was identified using a flow cytometry protein interaction assay (FCPIA), with a potency of 4 μM [96]. Although potency was significantly improved over the peptide inhibitors, the inhibition of CCG-4986 cannot be reversed in the absence of a strong reducing agent (e.g. dithiothreitol, DTT), and has no activity against an RGS mutant lacking cysteine residues (RGS4 (–Cys)) [92]. Subsequently, a medium-scale (ca. 43K compounds) high-throughput screen using a time-resolved FRET assay identified the first example of an intrinsically reversible (reversible without reducing agent) RGS4 inhibitor with an 8 μM potency (CCG-63802) [94]. The reactivity results from a Michael addition reaction through the vinyl cyanide group, which is less stable [97].

The first sub-micromolar RGS inhibitor shifted the focus to a new family of compounds centered around a thiadiazolidine-3,5-dione ring, represented by CCG-50014, which has an IC_{50} equal to 30 nM to inhibit RGS4-Gαo equilibrium binding [95]. The CCG-50014 compound provided a significant enhancement in specificity (300-fold) for RGS4 over RGS8, and most importantly this compound
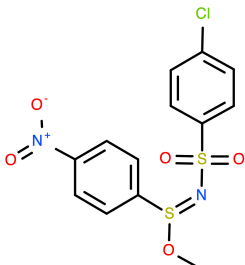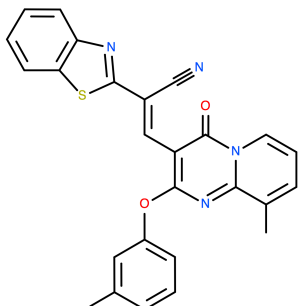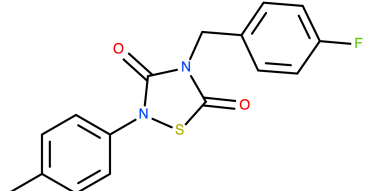
Compound ID	RGS Inhibitor Class	Potency (μM)	Structure	Ref.
YJ34	Cyclic Peptide Inhibitors (S-S)	30	Ac-VKCTGICE-NH$_2$	[88, 89]
Peptide17	Peptide Inhibitors	55	VRHVAVEVGGVVVVVG	[91]
CCG-4986	Covalent, irreversible	4		[92, 93]
CCG-63802	Covalent, reversible	8		[94]
CCG-50014	Covalent, Irreversible, Cell-active	0.03		[95]

Table 1.1 Current classes of RGS inhibitors

showed the first cellular activity towards RGS function [95]. Like all previously reported small molecule inhibitors, CCG-50014 forms a covalent adduct with cysteine residues on RGS proteins through a reactive sulfur moiety on the thiadiazolidine-3,5-dione ring [98]. CCG-50014 irreversibly inhibits WT RGS4 in the absence of reducing agent [95]. Despite the thiol-reactive mechanism, the compound is able to evade the reducing environment of the cell and inhibit RGS activity. Furthermore, derivatives with improved solubility are now being used in animal models (data in preparation), allowing an opportunity to acutely probe RGS function *in vivo*.

The structural characteristics of RGS proteins have thus far only shown sensitivity towards thiol-reactive small molecules. As a result, the structural pockets available on RGS family members do not appear to provide an easy mechanism to modulate activity with non-covalent small molecules, consistent with the challenges associated with PPI-directed small molecules in general. Considering CCG-50014 is > 300-fold selective for RGS4 over RGS8 and 3000-fold selective for RGS4 over a cysteine protease (papain), excellent progress has been made in obtaining ample potency and selectivity with the thiol-directed RGS inhibitors. The inhibitors are able to effectively block RGS function without modulating G-protein activity. The selectivity of the inhibitors among different RGS homologs is likely a result of the number of cysteine residues present and their structural locations. The degree to which some of the cysteine residues are buried in certain RGS homologs also suggests that dynamical motions may play a role in differential sensitivity to current inhibitors. As a result, elucidating the

conformational changes that play a role in the mechanism of inhibition may be a significant step forward in the design of new small molecule inhibitors with greater specificity.

Time-scales of Protein Motions Relevant for Drug Discovery

Protein motions occur in two general modes involving either high frequency and low amplitude motions or large collective motions. The former involves local structural flexibilities that include bond vibrations, side chain rotamerizations, and loop movements that range in time-scale from $10^{-15} – 10^{-9}$ seconds [99] (**Figure 1.5**). These structural fluctuations result in conformationally similar substates that are present in a single energy well within the free energy landscape of the protein. These motions are easily accessible to experimental methods such as solution-state NMR since the time-scales are much shorter than the overall molecular tumbling rate of the protein in solution [100]. Since the interconversion between these sub-states describes the entropy of the system, changes to these local fluctuations upon ligand binding will affect the overall free energy of the interaction [101-103]. This is an important aspect of a single isoenergetic state of a protein, but does not describe the range of conformational populations within a system.

The local structural fluctuations that occur on very fast time-scales coordinate into collective motions of secondary and tertiary structural elements involved in protein function and ligand binding [99, 104-106]. The height of the energy barriers separating these distinct conformational states result in the
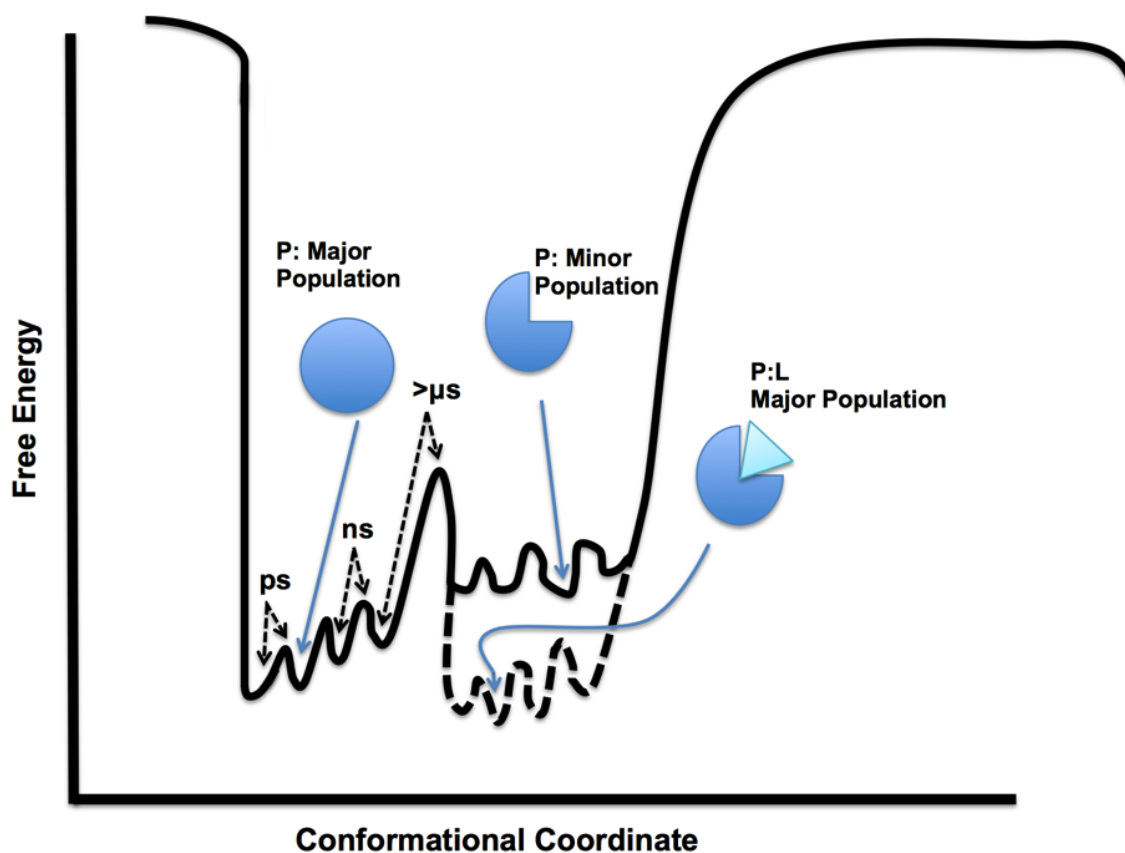
Figure 1.5 Representative free energy landscape as a function of conformational coordinate. Local fluctuations (side chain rotamerization, methyl rotation, and loop movements) occur on the picosecond to nanosecond time-scales. Structural rearrangements of a protein (P) that can reveal cryptic pockets occur on time-scales of microseconds and longer. These may only be present in transient high-energy states (minor population). Ligand (L) selection of these transient states via binding a cryptic pocket can shift the population (dotted line).

interconversion rates of microseconds and slower [99]. Therefore, proteins exist in dynamic equilibrium between kinetically isolated conformations that are populated based on the Boltzmann distribution from the relative free energy level of each state [107, 108]. The presence of different protein conformers during ligand binding has been shown experimentally by monitoring the dynamics of lactate dehydrogenase with NADH, which reveals several kinetically separated bound-conformations occurring over tens of microseconds during the interaction

[109]. This particular experiment suggests that pocket formation may continue to progress even after the first bound-state conformation is formed, further supporting the notion that static ground state structures do not accurately represent all of the possible ligand recognition states. As a result, the longer time-scale motions are most interesting for the discovery of distinct druggable conformations within a single protein (**Figure 1.5**).

Methods for Elucidating The Conformational Landscape of A Protein

Characterizing the conformational landscape of a target can be challenging, although numerous experimental techniques can provide insight into motions leading to alternate conformations. Identifying slowly exchanging and highly populated protein conformers has been particularly effective using hydrogen-deuterium exchange experiments [110, 111], or single molecule FRET techniques [112, 113]. Although, in cases where the difference in free energy between two states is more than approximately 1 kcal•mol^{-1} the excited state may represent only a small fraction of the overall population. Identifying these transient excited-state conformations has been achieved in many cases using solution-state NMR [114-116]. These techniques are useful to validate the presence of low population states of a protein, but for the purpose of structure-based design the limited structural information hampers utility. The structural details can be generated with molecular dynamics (MD) simulations either alone or in combination with experimental data. In one successful example, combining small angle X-ray scattering (SAXS) and course-grained MD simulations revealed

the distribution of several highly populated conformers in dynamic equilibrium during subunit assembly of the Hck tyrosine kinase [107]. Hence, combining experimental and computational techniques can provide atomistic details of protein motions that are validated with physical evidence.

Traditional MD simulations can provide structural details of protein motions at the level of local structural fluctuations, including small loop movements that can reach tens of nanoseconds, but this falls short of reaching biologically relevant time-scales of microseconds and beyond [117]. Instead, accelerated molecular dynamics can resolve the atomistic details of conformational motions several orders of magnitude longer than traditional MD [118]. Enhanced sampling of conformations that would ordinarily require microseconds to milliseconds to overcome large free energy barriers can been achieved by augmenting the energy landscape in order to speed transitions between states [119-122]. Therefore, combining accelerated molecular dynamics with experimental techniques can resolve structural details of biomolecules that experimental techniques alone cannot capture. In doing so a more thorough understanding of both biological and pharmacological mechanism can be understood.

Thesis Overview

The work presented here contributes to our understanding of the molecular mechanisms for inhibition of RGS proteins with small molecules. These studies have focused on RGS4 as a prototypical RGS protein.

Chapter II elucidates the allosteric mechanism of inhibition by a highly potent RGS inhibitor (CCG-50014). Here we employ an RGS4 mutant that contains a single native cysteine in the buried allosteric site at position 95. The cysteine 95 residue is the closest cysteine to the modeled docking site of CCG-50014 [95], and this mutant is less prone to aggregation in the presence of CCG-50014 compared to WT RGS4. Like WT RGS4, the Cys95 mutant is completely inhibited by CCG-50014 and shows a significant reduction in thermal stability in the presence of the compound. We aimed to test the hypothesis that CCG-50014 interacting with Cys95 alters the packing arrangement of alpha helices 4-7 causing disruption of the G-protein binding interface. Consistent with the hypothesis, HSQC-NMR experiments monitoring structural changes associated with CCG-50014 binding to Cys95 show perturbations clustering in alpha helices 4-7 and downstream in the G-protein binding site. Using temperature-accelerated molecular dynamics (TAMD) we show how dynamic motions of the alpha 5-6 helical pair on RGS4 span an open like conformation allowing binding of CCG-50014 to the buried site. Binding of CCG-50014 appears to stabilize this open confirmation resulting in disruption of the G-protein binding interface. This model represents a novel mechanism for allosteric PPI inhibitors, and reveals a cryptic site (C-site) with potentially better properties for the design of small molecule inhibitors.

Chapter III describes the identification of several new cell-active RGS4 inhibitors using a cell-based calcium signaling assay. Most of the RGS inhibitors identified to date lack cellular activity, which may be a result of the challenge for

cysteine-directed small molecules to target the RGS protein in the reducing environment of the cell. In order to identify cell-active chemical modulators of RGS4 we utilized a cell-based calcium signaling assay using an HEK cell line stably expressing the M3-muscarinic receptor with conditional RGS4 expression. This cell–line was screened against MLSMR NIH collection of >300,000 small molecules. Primary screen hits were identified based on their ability to enhance the Gq-mediated calcium response in cells expressing RGS4. Ultimately, fifty-eight compounds were selected and subsequently characterized to test the hypothesis that the enhanced calcium response in the cells was mediated specifically through inhibition of RGS4. Using two biochemical assays of RGS4 activity, thirteen compounds were identified that consistently inhibit RGS4 with IC_{50} values less than 100 μM in both assays. Subsequent biophysical and biochemical characterization of the compounds identified four molecules that could inhibit RGS4 reversibly. The reversible compounds showed relatively superior selectivity towards RGS4, over a panel of other RGS homologs tested. These cell-active RGS4 inhibitors identified using this system provide critical new tools to dissect the role of RSG4 in biology and disease.

In Chapter IV, biophysical and biochemical experiments are used to study the interaction of inositol hexakisphosphate (IP_6) with RGS4. The IP_6 compound is a close analog to the soluble head group of phosphatidylinositol (3,4,5)-trisphosphate (PIP_3), which can inhibit RGS4 GAP activity. Previous biochemical studies suggest that PIP_3 binds to an allosteric site on RGS4 that would overlap the B-site pocket (**Figure 1.3**). I tested the hypothesis that IP_6 would bind to the

B-site on RGS4 in order to understand how B-site interactions might regulate RGS4 activity. Using HSQC-NMR experiments to map the interaction of IP_6 with RGS4, structural perturbations were identified on the C-terminal portion of α-helix 4 and parts of the loop connecting α-helix 6 and 7. Additionally, CSPs were observed within a site centered on Lys 110 of α-helix 5. The NMR binding experiments are supported by changes to RGS4 thermal stability in the presence of IP_6. Biochemical experiments show that IP_6 is unable to antagonize the RGS4 GAP activity under steady-state conditions. The results of the GAP data suggest that PIP_3 may sequester RGS4 away from receptor-G-protein complexes. These data show that small molecules targeting the B-site on RGS4 would provide a mechanism to upregulate RGS4 activity, which is potentially beneficial based on several preclinical animal models. Importantly, these data provide direct evidence of non-covalent small molecule interactions on an allosteric site of RGS4, and reveals a pocket that could be targeted in future structure-based drug discovery efforts.

References

1.	Overington, J.P., B. Al-Lazikani, and A.L. Hopkins, *How many drug targets are there?* Nat Rev Drug Discov., 2006. **5**(12): p. 993-6.
2.	Cohen, P., *Protein kinases--the major drug targets of the twenty-first century?* Nat Rev Drug Discov, 2002. **1**(4): p. 309-15.
3.	Zhu, F., et al., *Therapeutic target database update 2012: a resource for facilitating target-oriented drug discovery.* Nucleic Acids Res, 2012. **40**(Database issue): p. D1128-36.
4.	Vidal, M., M.E. Cusick, and A.L. Barabasi, *Interactome networks and human disease.* Cell, 2011. **144**(6): p. 986-98.
5.	Hopkins, A.L. and C.R. Groom, *The druggable genome.* Nat Rev Drug Discov, 2002. **1**(9): p. 727-30.
6.	Massague, J., *TGF-beta signal transduction.* Annu Rev Biochem, 1998. **67**: p. 753-91.
7.	Miyazono, K., P. ten Dijke, and C.H. Heldin, *TGF-beta signaling by Smad proteins.* Adv Immunol, 2000. **75**: p. 115-57.
8.	Colland, F., et al., *Functional proteomics mapping of a human signaling pathway.* Genome Res, 2004. **14**(7): p. 1324-32.
9.	Griffith, M., et al., *DGIdb: mining the druggable genome.* Nat Methods, 2013.
10.	Williams, C. and S.J. Hill, *GPCR signaling: understanding the pathway to successful drug discovery.* Methods Mol Biol, 2009. **552**: p. 39-50.
11.	Messer, W.S., Jr., *The utility of muscarinic agonists in the treatment of Alzheimer's disease.* J Mol Neurosci, 2002. **19**(1-2): p. 187-93.
12.	Thathiah, A. and B. De Strooper, *The role of G protein-coupled receptors in the pathology of Alzheimer's disease.* Nat Rev Neurosci, 2011. **12**(2): p. 73-87.
13.	Lane, J.R., A. Abdul-Ridha, and M. Canals, *Regulation of G protein-coupled receptors by allosteric ligands.* ACS Chem Neurosci, 2013. **4**(4): p. 527-34.
14.	Lane, J.R., P.M. Sexton, and A. Christopoulos, *Bridging the gap: bitopic ligands of G-protein-coupled receptors.* Trends Pharmacol Sci, 2013. **34**(1): p. 59-66.
15.	Magalhaes, A.C., H. Dunn, and S.S. Ferguson, *Regulation of GPCR activity, trafficking and localization by GPCR-interacting proteins.* Br J Pharmacol, 2012. **165**(6): p. 1717-36.
16.	Blazer, L.L. and R.R. Neubig, *Small molecule protein-protein interaction inhibitors as CNS therapeutic agents: current progress and future hurdles.* Neuropsychopharmacology, 2009. **34**(1): p. 126-41.
17.	Berman, D.M., T.M. Wilkie, and A.G. Gilman, *GAIP and RGS4 are GTPase-activating proteins for the Gi subfamily of G protein alpha subunits.* Cell, 1996. **86**(3): p. 445-52.

18.   Hepler, J.R., et al., *RGS4 and GAIP are GTPase-activating proteins for Gq alpha and block activation of phospholipase C beta by gamma-thio-GTP-Gq alpha.* Proc Natl Acad Sci U S A, 1997. **94**(2): p. 428-32.

19.   Ross, E.M. and T.M. Wilkie, *GTPase-activating proteins for heterotrimeric G proteins: regulators of G protein signaling (RGS) and RGS-like proteins.* Annu Rev Biochem, 2000. **69**: p. 795-827.

20.   Doupnik, C.A., et al., *RGS proteins reconstitute the rapid gating kinetics of gbetagamma-activated inwardly rectifying K+ channels.* Proc Natl Acad Sci U S A, 1997. **94**(19): p. 10461-6.

21.   Larminie, C., et al., *Selective expression of regulators of G-protein signaling (RGS) in the human central nervous system.* Brain Res Mol Brain Res, 2004. **122**(1): p. 24-34.

22.   Gold, S.J., et al., *Regulators of G-protein signaling (RGS) proteins: region-specific expression of nine subtypes in rat brain.* J Neurosci, 1997. **17**(20): p. 8024-37.

23.   Bernstein, L.S., et al., *RGS2 binds directly and selectively to the M1 muscarinic acetylcholine receptor third intracellular loop to modulate Gq/11alpha signaling.* J Biol Chem, 2004. **279**(20): p. 21248-56.

24.   Wang, Q., L.Y. Liu-Chen, and J.R. Traynor, *Differential modulation of mu- and delta-opioid receptor agonists by endogenous RGS4 protein in SH-SY5Y cells.* J Biol Chem, 2009. **284**(27): p. 18357-67.

25.   Kovoor, A., et al., *D2 dopamine receptors colocalize regulator of G-protein signaling 9-2 (RGS9-2) via the RGS9 DEP domain, and RGS9 knock-out mice develop dyskinesias associated with dopamine pathways.* J Neurosci, 2005. **25**(8): p. 2157-65.

26.   Zeng, W., et al., *The N-terminal domain of RGS4 confers receptor-selective inhibition of G protein signaling.* J Biol Chem, 1998. **273**(52): p. 34687-90.

27.   Celver, J., M. Sharma, and A. Kovoor, *RGS9-2 mediates specific inhibition of agonist-induced internalization of D2-dopamine receptors.* J Neurochem, 2010. **114**(3): p. 739-49.

28.   Talbot, J.N., et al., *Differential modulation of mu-opioid receptor signaling to adenylyl cyclase by regulators of G protein signaling proteins 4 or 8 and 7 in permeabilised C6 cells is Galpha subtype dependent.* J Neurochem, 2010. **112**(4): p. 1026-34.

29.   Heximer, S.P., et al., *RGS2/G0S8 is a selective inhibitor of Gqalpha function.* Proc Natl Acad Sci U S A, 1997. **94**(26): p. 14389-93.

30.   Hunt, T.W., et al., *RGS10 is a selective activator of G alpha i GTPase activity.* Nature, 1996. **383**(6596): p. 175-7.

31.   Hooks, S.B., et al., *RGS6, RGS7, RGS9, and RGS11 stimulate GTPase activity of Gi family G-proteins with differential selectivity and maximal activity.* J Biol Chem, 2003. **278**(12): p. 10087-93.

32.   Neubig, R.R. and D.P. Siderovski, *Regulators of G-protein signalling as new central nervous system drug targets.* Nat Rev Drug Discov, 2002. **1**(3): p. 187-97.

33.   Sjogren, B., L.L. Blazer, and R.R. Neubig, *Regulators of G protein signaling proteins as targets for drug discovery.* Prog Mol Biol Transl Sci, 2010. **91**: p. 81-119.

34.   Kimple, A.J., et al., *Regulators of G-protein signaling and their Galpha substrates: promises and challenges in their use as drug discovery targets.* Pharmacol Rev, 2011. **63**(3): p. 728-49.

35.   Lerner, T.N. and A.C. Kreitzer, *RGS4 is required for dopaminergic control of striatal LTD and susceptibility to parkinsonian motor deficits.* Neuron, 2012. **73**(2): p. 347-59.

36.   Chen, Y., et al., *Neurabin scaffolding of adenosine receptor and RGS4 regulates anti-seizure effect of endogenous adenosine.* J Neurosci, 2012. **32**(8): p. 2683-95.

37.   Janin, J., R.P. Bahadur, and P. Chakrabarti, *Protein-protein interaction and quaternary structure.* Q Rev Biophys, 2008. **41**(2): p. 133-80.

38.   Wells, J.A. and C.L. McClendon, *Reaching for high-hanging fruit in drug discovery at protein-protein interfaces.* Nature., 2007. **450**(7172): p. 1001-9.

39.   Lipinski, C.A., *Drug-like properties and the causes of poor solubility and poor permeability.* J Pharmacol Toxicol Methods, 2000. **44**(1): p. 235-49.

40.   Rees, D.C., et al., *Fragment-based lead discovery.* Nat Rev Drug Discov, 2004. **3**(8): p. 660-72.

41.   Fuller, J.C., N.J. Burgoyne, and R.M. Jackson, *Predicting druggable binding sites at the protein-protein interface.* Drug Discov Today., 2009. **14**(3-4): p. 155-61. doi: 10.1016/j.drudis.2008.10.009. Epub 2008 Dec 16.

42.   Perola, E., L. Herman, and J. Weiss, *Development of a rule-based method for the assessment of protein druggability.* J Chem Inf Model, 2012. **52**(4): p. 1027-38.

43.   Morelli, X., R. Bourgeas, and P. Roche, *Chemical and structural lessons from recent successes in protein-protein interaction inhibition (2P2I).* Curr Opin Chem Biol, 2011. **15**(4): p. 475-81.

44.   Tesmer, J.J., et al., *Structure of RGS4 bound to AlF4--activated G(i alpha1): stabilization of the transition state for GTP hydrolysis.* Cell, 1997. **89**(2): p. 251-61.

45.   Zhong, H. and R.R. Neubig, *Regulator of G protein signaling proteins: novel multifunctional drug targets.* J Pharmacol Exp Ther, 2001. **297**(3): p. 837-45.

46.   Bogan, A.A. and K.S. Thorn, *Anatomy of hot spots in protein interfaces.* J Mol Biol, 1998. **280**(1): p. 1-9.

47.   London, N., D. Movshovitz-Attias, and O. Schueler-Furman, *The structural basis of peptide-protein binding strategies.* Structure, 2010. **18**(2): p. 188-99.

48.   Clackson, T. and J.A. Wells, *A hot spot of binding energy in a hormone-receptor interface.* Science., 1995. **267**(5196): p. 383-6.

49.   DeLano, W.L., *Unraveling hot spots in binding interfaces: progress and challenges.* Curr Opin Struct Biol., 2002. **12**(1): p. 14-20.

50.   Thorn, K.S. and A.A. Bogan, *ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions.* Bioinformatics., 2001. **17**(3): p. 284-5.

51.   Kozakov, D., et al., *Structural conservation of druggable hot spots in protein-protein interfaces.* Proc Natl Acad Sci U S A, 2011. **108**(33): p. 13528-33.

52.   Landon, M.R., et al., *Identification of hot spots within druggable binding regions by computational solvent mapping of proteins.* J Med Chem., 2007. **50**(6): p. 1231-40. Epub 2007 Feb 17.

53.   Bowman, G.R. and P.L. Geissler, *Equilibrium fluctuations of a single folded protein reveal a multitude of potential cryptic allosteric sites.* Proc Natl Acad Sci U S A, 2012. **109**(29): p. 11681-6.

54.   Halabi, N., et al., *Protein sectors: evolutionary units of three-dimensional structure.* Cell, 2009. **138**(4): p. 774-86.

55.   Reynolds, K.A., R.N. McLaughlin, and R. Ranganathan, *Hot spots for allosteric regulation on protein surfaces.* Cell, 2011. **147**(7): p. 1564-75.

56.   DeDecker, B.S., *Allosteric drugs: thinking outside the active-site box.* Chem Biol, 2000. **7**(5): p. R103-7.

57.   Nussinov, R. and C.J. Tsai, *Allostery in disease and in drug discovery.* Cell, 2013. **153**(2): p. 293-305.

58.   Conn, P.J., A. Christopoulos, and C.W. Lindsley, *Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders.* Nat Rev Drug Discov, 2009. **8**(1): p. 41-54.

59.   Gao, Z.G. and K.A. Jacobson, *Allosteric modulation and functional selectivity of G protein-coupled receptors.* Drug Discov Today Technol, 2013. **10**(2): p. e237-43.

60.   Shirey, J.K., et al., *A selective allosteric potentiator of the M1 muscarinic acetylcholine receptor increases activity of medial prefrontal cortical neurons and restores impairments in reversal learning.* J Neurosci., 2009. **29**(45): p. 14271-86. doi: 10.1523/JNEUROSCI.3930-09.2009.

61.   Wenthur, C.J., et al., *Discovery of (R)-(2-fluoro-4-((-4-methoxyphenyl)ethynyl)phenyl) (3-hydroxypiperidin-1-yl)methanone (ML337), an mGlu3 selective and CNS penetrant negative allosteric modulator (NAM).* J Med Chem., 2013. **56**(12): p. 5208-12. doi: 10.1021/jm400439t. Epub 2013 Jun 13.

62.   Sadowsky, J.D., et al., *Turning a protein kinase on or off from a single allosteric site via disulfide trapping.* Proc Natl Acad Sci U S A, 2011. **108**(15): p. 6056-61.

63.   Ishii, M., et al., *Phosphatidylinositol 3,4,5-trisphosphate and Ca2+/calmodulin competitively bind to the regulators of G-protein-signalling (RGS) domain of RGS4 and reciprocally regulate its action.* Biochem J, 2005. **385**(Pt 1): p. 65-73.

64.   Ishii, M., A. Inanobe, and Y. Kurachi, *PIP3 inhibition of RGS protein and its reversal by Ca2+/calmodulin mediate voltage-dependent control of the G protein cycle in a cardiac K+ channel.* Proc Natl Acad Sci U S A, 2002. **99**(7): p. 4325-30.

65. Tu, Y. and T.M. Wilkie, *Allosteric regulation of GAP activity by phospholipids in regulators of G-protein signaling.* Methods Enzymol, 2004. **389**: p. 89-105.

66. Ishii, M. and Y. Kurachi, *Assays of RGS protein modulation by phosphatidylinositides and calmodulin.* Methods Enzymol, 2004. **389**: p. 105-18.

67. Popov, S.G., et al., *Ca2+/Calmodulin reverses phosphatidylinositol 3,4, 5-trisphosphate-dependent inhibition of regulators of G protein-signaling GTPase-activating protein activity.* J Biol Chem, 2000. **275**(25): p. 18962-8.

68. Hollinger, S. and J.R. Hepler, *Cellular regulation of RGS proteins: modulators and integrators of G protein signaling.* Pharmacol Rev, 2002. **54**(3): p. 527-59.

69. Ishii, M., M. Ikushima, and Y. Kurachi, *In vivo interaction between RGS4 and calmodulin visualized with FRET techniques: possible involvement of lipid raft.* Biochem Biophys Res Commun, 2005. **338**(2): p. 839-46.

70. Fink, T. and J.L. Reymond, *Virtual exploration of the chemical universe up to 11 atoms of C, N, O, F: assembly of 26.4 million structures (110.9 million stereoisomers) and analysis for new ring systems, stereochemistry, physicochemical properties, compound classes, and drug discovery.* J Chem Inf Model., 2007. **47**(2): p. 342-53. Epub 2007 Jan 30.

71. Murray, C.W., et al., *Fragment-based drug discovery applied to Hsp90. Discovery of two lead series with high ligand efficiency.* J Med Chem, 2010. **53**(16): p. 5942-55.

72. Bembenek, S.D., B.A. Tounge, and C.H. Reynolds, *Ligand efficiency and fragment-based drug discovery.* Drug Discov Today, 2009. **14**(5-6): p. 278-83.

73. Hopkins, A.L., C.R. Groom, and A. Alex, *Ligand efficiency: a useful metric for lead selection.* Drug Discov Today, 2004. **9**(10): p. 430-1.

74. Congreve, M., et al., *A 'rule of three' for fragment-based lead discovery?* Drug Discov Today, 2003. **8**(19): p. 876-7.

75. Hajduk, P.J., et al., *High-throughput nuclear magnetic resonance-based screening.* J Med Chem., 1999. **42**(13): p. 2315-7.

76. Shuker, S.B., et al., *Discovering high-affinity ligands for proteins: SAR by NMR.* Science, 1996. **274**(5292): p. 1531-4.

77. Regnstrom, K., et al., *Label free fragment screening using surface plasmon resonance as a tool for fragment finding - analyzing parkin, a difficult CNS target.* PLoS One., 2013. **8**(7): p. e66879. doi: 10.1371/journal.pone.0066879. Print 2013.

78. Larsson, E.A., et al., *Fragment-based ligand design of novel potent inhibitors of tankyrases.* J Med Chem., 2013. **56**(11): p. 4497-508. doi: 10.1021/jm400211f. Epub 2013 Jun 4.

79. Zhao, L., et al., *Fragment-based drug discovery of 2-thiazolidinones as inhibitors of the histone reader BRD4 bromodomain.* J Med Chem., 2013. **56**(10): p. 3833-51. doi: 10.1021/jm301793a. Epub 2013 May 14.

80. Vassar, R., et al., *Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE.* Science, 1999. **286**(5440): p. 735-41.

81. Wang, Y.S., et al., *Application of fragment-based NMR screening, X-ray crystallography, structure-based design, and focused chemical library design to identify novel microM leads for the development of nM BACE-1 (beta-site APP cleaving enzyme 1) inhibitors.* J Med Chem., 2010. **53**(3): p. 942-50. doi: 10.1021/jm901472u.

82. Woodhead, A.J., et al., *Discovery of (2,4-dihydroxy-5-isopropylphenyl)-[5-(4-methylpiperazin-1-ylmethyl)-1,3-dihydrois oindol-2-yl]methanone (AT13387), a novel inhibitor of the molecular chaperone Hsp90 by fragment based drug design.* J Med Chem, 2010. **53**(16): p. 5956-69.

83. Howard, S., et al., *Fragment-based discovery of the pyrazol-4-yl urea (AT9283), a multitargeted kinase inhibitor with potent aurora kinase activity.* J Med Chem., 2009. **52**(2): p. 379-88. doi: 10.1021/jm800984v.

84. May, P.C., et al., *Robust central reduction of amyloid-beta in humans with an orally available, non-peptidic beta-secretase inhibitor.* J Neurosci., 2011. **31**(46): p. 16507-16. doi: 10.1523/JNEUROSCI.3647-11.2011.

85. Hajduk, P.J., J.R. Huth, and C. Tse, *Predicting protein druggability.* Drug Discov Today., 2005. **10**(23-24): p. 1675-82.

86. Edfeldt, F.N., R.H. Folmer, and A.L. Breeze, *Fragment screening to predict druggability (ligandability) and lead discovery success.* Drug Discov Today, 2011. **16**(7-8): p. 284-7.

87. Dias, D.M., et al., *Is NMR Fragment Screening Fine-Tuned to Assess Druggability of Protein–Protein Interactions?* ACS Medicinal Chemistry Letters, 2013.

88. Roof, R.A., et al., *Mechanism of action and structural requirements of constrained peptide inhibitors of RGS proteins.* Chem Biol Drug Des, 2006. **67**(4): p. 266-74.

89. Roof, R.A., et al., *Novel peptide ligands of RGS4 from a focused one-bead, one-compound library.* Chem Biol Drug Des, 2008. **72**(2): p. 111-9.

90. Lan, K.L., et al., *A point mutation in Galphao and Galphai1 blocks interaction with regulator of G protein signaling proteins.* J Biol Chem., 1998. **273**(21): p. 12794-7.

91. Wang, Y., et al., *Identification of peptides that inhibit regulator of G protein signaling 4 function.* Pharmacology, 2008. **82**(2): p. 97-104.

92. Roman, D., et al., *Allosteric Inhibition of the RGS-G{alpha} Protein-Protein Interaction by CCG-4986.* Mol Pharmacol, 2010.

93. Roman, D.L., S. Ota, and R.R. Neubig, *Polyplexed flow cytometry protein interaction assay: a novel high-throughput screening paradigm for RGS protein inhibitors.* J Biomol Screen, 2009. **14**(6): p. 610-9.

94. Blazer, L.L., et al., *Reversible, allosteric, small-molecule inhibitors of RGS proteins.* Mol Pharmacol, 2010.

95. Blazer, L.L., et al., *A nanomolar-potency small molecule inhibitor of Regulator of G protein Signaling (RGS) proteins.* Biochemistry, 2011.

96. Roman, D.L., et al., *Identification of small-molecule inhibitors of RGS4 using a high-throughput flow cytometry protein interaction assay.* Mol Pharmacol, 2007. **71**(1): p. 169-75.

97. Lee, C.U. and T.N. Grossmann, *Reversible covalent inhibition of a protein target.* Angew Chem Int Ed Engl, 2012. **51**(35): p. 8699-700.

98. Turner, E.M., et al., *Small Molecule Inhibitors of Regulator of G Protein Signalling (RGS) Proteins.* ACS Med Chem Lett, 2012. **3**(2): p. 146-150.

99. Henzler-Wildman, K.A., et al., *A hierarchy of timescales in protein dynamics is linked to enzyme catalysis.* Nature., 2007. **450**(7171): p. 913-6. Epub 2007 Nov 18.

100. Akke, M., *Conformational dynamics and thermodynamics of protein-ligand binding studied by NMR relaxation.* Biochem Soc Trans, 2012. **40**(2): p. 419-23.

101. Diehl, C., et al., *Conformational entropy changes upon lactose binding to the carbohydrate recognition domain of galectin-3.* J Biomol NMR., 2009. **45**(1-2): p. 157-69. doi: 10.1007/s10858-009-9356-5. Epub 2009 Jul 30.

102. Teague, S.J., *Implications of protein flexibility for drug discovery.* Nat Rev Drug Discov., 2003. **2**(7): p. 527-41.

103. Edwards, A.A., et al., *Altered enthalpy-entropy compensation in picomolar transition state analogues of human purine nucleoside phosphorylase.* Biochemistry., 2009. **48**(23): p. 5226-38. doi: 10.1021/bi9005896.

104. Henzler-Wildman, K. and D. Kern, *Dynamic personalities of proteins.* Nature, 2007. **450**(7172): p. 964-72.

105. Frauenfelder, H., S.G. Sligar, and P.G. Wolynes, *The energy landscapes and motions of proteins.* Science, 1991. **254**(5038): p. 1598-603.

106. Boehr, D.D., R. Nussinov, and P.E. Wright, *The role of dynamic conformational ensembles in biomolecular recognition.* Nat Chem Biol, 2009. **5**(11): p. 789-96.

107. Yang, S., et al., *Multidomain assembled states of Hck tyrosine kinase in solution.* Proc Natl Acad Sci U S A, 2010. **107**(36): p. 15757-62.

108. Volkman, B.F., et al., *Two-state allosteric behavior in a single-domain signaling protein.* Science, 2001. **291**(5512): p. 2429-33.

109. Deng, H., N. Zhadin, and R. Callender, *Dynamics of protein ligand binding on multiple time scales: NADH binding to lactate dehydrogenase.* Biochemistry, 2001. **40**(13): p. 3767-73.

110. Mysling, S., et al., *Characterizing the dynamics of alpha-synuclein oligomers using hydrogen/deuterium exchange monitored by mass spectrometry.* Biochemistry, 2013: p. 5.

111. Ahmed, A.H., et al., *Dynamics of cleft closure of the GluA2 ligand-binding domain in the presence of full and partial agonists revealed by hydrogen-deuterium exchange.* J Biol Chem., 2013. **288**(38): p. 27658-66. doi: 10.1074/jbc.M113.495564. Epub 2013 Aug 12.

112. Lamboy, J.A., et al., *Single-molecule FRET reveals the native-state dynamics of the IkappaBalpha ankyrin repeat domain.* J Mol Biol., 2013. **425**(14): p. 2578-90. doi: 10.1016/j.jmb.2013.04.015. Epub 2013 Apr 22.

113.	Fiegland, L.R., et al., *Single-molecule studies of the lysine riboswitch reveal effector-dependent conformational dynamics of the aptamer domain.* Biochemistry., 2012. **51**(45): p. 9223-33. doi: 10.1021/bi3007753. Epub 2012 Oct 30.

114.	Clore, G.M., *Visualizing lowly-populated regions of the free energy landscape of macromolecular complexes by paramagnetic relaxation enhancement.* Mol Biosyst, 2008. **4**(11): p. 1058-69.

115.	Palmer, A.G., 3rd, M.J. Grey, and C. Wang, *Solution NMR spin relaxation methods for characterizing chemical exchange in high-molecular-weight systems.* Methods Enzymol, 2005. **394**: p. 430-65.

116.	Long, D., et al., *A Comparative CEST NMR Study of Slow Conformational Dynamics of Small GTPases Complexed with GTP and GTP Analogues.* Angew Chem Int Ed Engl, 2013.

117.	Zwier, M.C. and L.T. Chong, *Reaching biological timescales with all-atom molecular dynamics simulations.* Curr Opin Pharmacol, 2010. **10**(6): p. 745-52.

118.	Markwick, P.R. and J.A. McCammon, *Studying functional dynamics in bio-molecules using accelerated molecular dynamics.* Phys Chem Chem Phys., 2011. **13**(45): p. 20053-65. doi: 10.1039/c1cp22100k. Epub 2011 Oct 21.

119.	Lange, O.F., L.V. Schafer, and H. Grubmuller, *Flooding in GROMACS: accelerated barrier crossings in molecular dynamics.* J Comput Chem, 2006. **27**(14): p. 1693-702.

120.	Kim, S.Y., D. Perez, and A.F. Voter, *Local hyperdynamics.* J Chem Phys., 2013. **139**(14): p. 144110. doi: 10.1063/1.4824389.

121.	Abrams, C.F. and E. Vanden-Eijnden, *Large-scale conformational sampling of proteins using temperature-accelerated molecular dynamics.* Proc Natl Acad Sci U S A, 2010. **107**(11): p. 4961-6.

122.	Hamelberg, D., J. Mongan, and J.A. McCammon, *Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules.* J Chem Phys, 2004. **120**(24): p. 11919-29.

Chapter II

**Conformational Dynamics of RGS4 Reveals A Mechanism
of Allosteric Inhibition**

Introduction

Protein−protein interactions (PPIs) similar to that between RGS and Gα represent critical therapeutic targets for their direct role in abnormal signaling mechanisms [1-3], although targeting PPIs is a significant challenge [4, 5]. The many challenges associated with inhibiting PPIs are outlined in Chapter I. Despite these challenges, several compound families of RGS inhibitors have been identified and characterized (**Table 1.1**). Relative to the first reported peptide inhibitors, the advantages of small molecule inhibitors include greater potency and, most recently, the ability to inhibit RGS4 inside a cell [6]. As a result, small molecule modulators of RGS proteins have been a primary focus towards developing new biological probes of RGS function, as well as being chemical starting points for RGS-directed therapeutics.

All of the small molecule inhibitors thus far identified require the presence of a reactive cysteine residue on the RGS protein for functional activity [6-8]. For many RGS homologs, cysteine residues are located in functionally sensitive areas, as indicated by the ability of thiol-reactive compounds to inhibit G-protein binding [7]. Part of the successful inhibition achieved by the cysteine-reactive

compounds may result from the presence of cysteine residues located in allosteric sites that are distal to the G-protein binding interface. Targeting allosteric sites would avoid competition with the G-protein and may provide greater potency. Furthermore, the RGS4 allosteric site (also known as the "B-site") [9], located away from the Gα binding site (termed as the "A-site"), is a promising site since physiological regulators of RGS4 directly interact with the B-site [10, 11].

The structural details by which cysteine-directed inhibitors modulate RGS function at sites distal to the G-protein binding interface are unknown. This has made it difficult to pursue structure-based approaches directed towards the RGS B-site. A recently identified RGS inhibitor (CCG-50014) is a potent (IC50 = 30 nM) and selective inhibitor of RGS4 that was discovered in a high-throughput biochemical screen [12]. As with all previously reported RGS inhibitors, it acts by forming a covalent adduct to cysteine residues in RGS proteins [6-8, 13]. A previous docked model of CCG-50014 binding to RGS8 (a homolog of RGS4) [6] has been proposed, which suggested that the inhibitor molecule docks near the Cys107 residue of RGS8 (Cys95 in RGS4; see **Figure 2.1**). In this model, the reactive group of CCG-50014 is located ~8−13 Å away from two cysteine residues in RGS8, where formation of a covalent adduct is unlikely. Therefore, it is hypothesized that a conformational change in the helix bundle must occur to allow binding of the inhibitor molecule to the buried side-chain of Cys95/Cys107 in RGS4/RGS8. In this work, the existence of a previously uncharacterized open-like conformation of RGS4 is exposed using atomistic simulations and NMR

experiments. In characterizing the molecular mechanism of CCG-50014, it is revealed that the compound acts as a chemical probe of this novel conformational state, which may provide an invaluable model for structure-based design of novel inhibitors.

In this chapter, I performed all of the protein expression and purification, and all of the biochemical and biophysical experiments. Based on my experimental data we were able to formulate a hypothesis on the mechanism of inhibition by CCG-50014. This hypothesis was tested using computational modeling of the interaction, which was performed by Harish Vashitsh (post-doctoral fellow in Professor Charles Brooks' Lab, Dept. of Biophysics, UM). This modeling includes MD, TAMD, principle component analysis, normal mode analysis, and theoretical chemical shift perturbation calculations. Parts of this chapter have been compiled into a publication in ACS Chemical Biology [14].

Materials and Methods

Molecular Dynamics (MD) Simulations

All MD trajectories were generated using NAMDv2.8 [15] and the CHARMM force-field with CMAP correction [16, 17]. VMDv1.9 [17] was used for system creation and protein rendering. Initial coordinates from two different structures of RGS4 with Protein Data Bank (PDB) codes 1AGR (crystal structure) and 1EZT (NMR structure) were used. All systems were solvated using explicit (TIP3P) water and all hydrogen atoms were included. Charge neutrality was

maintained by adding counterions. Similar protocols were followed for all equilibration simulations: 500-1000 cycles of energy-minimization via conjugate-gradient optimization and a constant temperature (310K) via the Langevin thermostat with damping coefficient of 5 ps^{-1}. The equilibration phases were carried out initially in the NPT ensemble to adjust the box volume, and thereafter continued in the NVT ensemble using a time-step of 2-fs with rigid bonds in all simulations. Periodic boundary conditions were used throughout. Non-bonded interactions were cut-off beyond 10 Å with smooth switching taking effect at 8.0 Å. Long-range electrostatic interactions were handled using the particle mesh Ewald (PME) method. The equilibration runs for wild-type RGS4 were used to sample initial conditions for all temperature accelerated molecular dynamics (TAMD) trajectories. Details of TAMD simulations are described in the following. The inhibitor molecule (CCG-50014; see **Figure 2.1**) was parameterized using MATCH [18].

Temperature Accelerated Molecular Dynamics

Temperature-accelerated molecular dynamics (TAMD) is an enhanced sampling approach to explore the physical free-energy landscape in a pre-defined set of collective variables (CVs) [19-21]. In this work, we have used a conformational sampling algorithm for proteins based upon TAMD [22], which has been successfully applied to several systems recently [23-26]. The coupled system of equations describing TAMD are as follows:

$$m_i \ddot{x}_i = -\frac{\delta V(x)}{\delta x_i} - \kappa \sum_{j=1}^{m} [\theta_j^*(x) - \theta_j] \frac{\delta \theta_j^*(x)}{\delta x_i} - \gamma m_i \dot{x}_i + \eta_i\,(t;\beta)$$

$$\bar{\gamma}\bar{m}_j \dot{\theta}_j = \kappa [\theta_j^*(x) - \theta_j] + \xi_j(t;\bar{\beta})$$

where $\theta^*(x) = (\theta^*_1(x), \theta^*_2(x), \ldots\ldots, \theta^*_m(x))$ are collective variables that are functions of atom Cartesian coordinates, m_i and $\bar{m}_j$ are the masses of x_i and θ_j, V(x) is the interatomic MD potential, κ is the "coupling spring-constant", γ is the Langevin friction coefficient, η is the white noise satisfying fluctuation-dissipation theorem at physical temperature $\bar{\bar{\beta}}^{-1}$, $\bar{\gamma}$ and ξ respectively are fictitious friction and thermal noise at artificial temperature $\bar{\bar{\beta}}^{-1}$.

The aforementioned set of equations describe the motion of *x(t)* and θ*(t)* over the extended potential:

$$U_\kappa(x,\theta) = V(x) + \frac{\kappa}{2} \sum_{j=1}^{m} [\theta_j^*(x) - \theta_j]^2$$

As shown earlier by Maragliano and Vanden-Eijnden [19], by choosing κ such that $\theta^*(x(t)) \approx \theta(t)$ and fictitious friction coefficient $\bar{\gamma}$ such that θ moves slower than *x*, one can generate a trajectory at artificial temperature $\bar{\beta}^{-1}$ subject to the free energy computed at the physical temperature β^{-1}. To ensure these conditions, we have chosen a TAMD friction $\bar{\gamma}$ of 500 ps^{-1} and a spring constant κ of 100 kcal/mol·Å^2. In this work, we choose the Cartesian coordinates of centers-of-mass of spatially contiguous groups of residues as CVs. The entire RGS4 structure (residues 51-178) was divided into 6 subdomains (18 CVs). Residue

memberships for mutually exclusive and collectively exhaustive subdomains of RGS4 are listed in **Appendix II** and the location of each subdomain is shown. A total of four independent TAMD trajectories were carried out (**Table 2.1**).

Theoretical Chemical Shift Perturbation (CSP) Analysis

We used software packages SHIFTX2 [27] and SPARTA+ [28] for theoretical prediction of chemical shifts for ^{15}N and ^{1}H atoms based upon our MD simulations. Specifically, starting with two different initial structures (PDB codes 1AGR and 1EZT), we carried out at least three independent simulations of mutant RGS4 with and without inhibitor CCG-50014 (runs #2, 7, 11 and 13 in **Table 2.1**). The predicted values of chemical shifts were further time-averaged over each trajectory and ensemble-averaged over independent trajectories. We used following equation to calculate chemical shift perturbation (CSP) for each residue.

$$CSP = \sqrt{\left(\frac{\Delta\delta^{15}N}{5}\right)^2 + \left(\Delta\delta\ ^{1}H\right)^2}$$

The predicted CSPs for each residue (in ppm) are also calculated using this equation, while key residues are summarized in below.

Protein Expression and Purification

RGS4 was expressed and purified for NMR studies as previously described [29]. Briefly, N-terminally truncated ($\Delta51$) RGS4 was expressed with an N-terminal maltose binding protein (MBP) fusion containing a 10x histidine tag

and tobacco etch virus protease (TEV) site using a pMalC2H10 vector. Two RGS4 mutants were prepared with either all native cysteines mutated to alanine (cysless RGS4), or all native cysteines except cysteine 95 mutated to alanine (Cys95 RGS4), using methods employed previously [8]. Purified ^{15}N-labeled Δ51 RGS4 (cysless and Cys95) was prepared by removing the MBP tag via addition of TEV protease and purified by ion exchange chromatography. The samples were buffer exchanged into 50 mM sodium phosphate buffer with 50 mM NaCl at pH 6.0 by dialyzing overnight at 4°C. The concentration was determined using a NanoDrop spectrophotometer.

Intrinsic Fluorescence Experiment

Fluorescence experiments were performed on a Photon Technology International AlphaScan Spectrofluorometer (Photon Technology International, Lawrenceville, NJ) with 2.5-nm slits using wavelengths of 285 and 340 nm for excitation and emission, respectively. (Δ51) RGS4 Cys95 or Cysless mutants at 1 µM in 500 µl of buffer containing 20 mM HEPES, 100 mM NaCl, pH 7.4 was added to a 5-mm cylindrical quartz cell with micro stir bar. The baseline reading was allowed to stabilize for 1 min followed by injections of CCG-50014 or equivalent volume of DMSO up to a final concentration of 30 µM. The fluorescence was recorded for 30 seconds at each concentration of CCG-50014.

Flow Cytometry Protein Interaction Assay (FCPIA)

Biotinylated (Δ51) RGS4 Cys95 or Cysless (N-terminally tagged with maltose binding protein) was immobilized on avidin-coated microspheres from Luminex, and incubated with compound or DMSO for 15 minutes at room temperature on a 96-well plate. AlexaFluor 532-tagged Gαo (30 nM final) activated with GDP, magnesium, and AlF^-_4 is then added to each well and incubated at room temperature for 30 minutes. Compound activity is determined using a Luminex 200 flow cytometer to detect changes microsphere-associated fluorescence. CCG-50014 was tested in a concentration-response format ranging from 100 μM to 30 nM. Concentration-response curves for each compound, performed in duplicate, were fit to IC_{50} curves (constrained to 100% inhibition) using Graphpad software v5.04.

Thermal Stability Assay

Δ51 RGS4 Cys95 (10 μM) was incubated in the presence of CCG-50014 (100 μM) or DMSO for 15 minutes black 384-well PCR microtiter plate (ThermoFisher catalog no. TF-0384/K). 1-anilinonaphthalene-8-sulfonic acid (ANS) was added to each well and all wells were overlaid with 5μl silicon oil. Fluorescence was measured at 25°C in up/down mode with increasing temperature increments of 1°C from 30°-90°C using ThermoFluor Instrument (Johnson & Johnson, Langhorne, PA). Melting curves were fit to a Boltzmann equation using Graphpad software v5.04 to determine the melting temperature (T_m).

NMR Spectroscopy

NMR samples were prepared with 0.1 mM ^{15}N-labeled RGS4 in the same buffer with the addition of 7% (v/v) D_2O and either DMSO (2% v/v) or 0.1 mM CCG-50014. The heteronuclear single-quantum coherence (HSQC) NMR data were collected at 25°C with a 600 MHz Bruker Avance III magnet with cryogenically cooled sample probe. Data were processed with Bruker TopSpin 1.3 software and analyzed with computer aided resonance assignment (CARA) software. The ^{15}N chemical shifts were referenced from the previous solution structure data of RGS4 [30]. ^{1}H-^{15}N HSQC spectra of cysless RGS4 and Cys95 RGS4 are shown as normalized intensity ratios of the HSQC peaks after and before addition of inhibitor CCG-50014. We note that chemical shift changes for each residue in the HSQC spectra have not been identified since the irreversible nature of the compound (CCG-50014) prevents titrating peak shift trajectories as a function of molar ratio of compound to protein. Instead, peak perturbations are analyzed via a loss in intensity resulting from the peaks becoming split between two (or possibly more) chemical states. We also note that the HSQC spectra of cysless RGS4 and Cys95 RGS4 show similar signal dispersion compared to wild-type RGS4 indicating that the overall fold is preserved. Furthermore, peak coordinates in the HSQC spectra are similar to wild-type RGS4 for more than 80% of the residues in both mutant proteins. Peak intensities in each spectrum were normalized to residue 201 in order to compare the intensity change before and after addition of CCG-50014. Peak intensities were compared either as a

ratio, or as % peak attenuation $(100-((I_{50014}/I_0)*100))$. A number of peaks were removed from the peak intensity analysis due to ambiguity resulting from overlap in the HSQC spectrum: 58, 73, 78, 79, 95, 96, 97, 107, 114, 117, 119, 132, 133, 135, 136, 152, 156, 161, 163, 183, 186, 188, 189, 192, 203, 204. Nearly all of these peaks are found to be overlapped in the WT RGS4 spectrum, published previously.

Principal Component Analysis Of The NMR Ensemble (PDB code 1EZY)

We carried out principal component analysis (PCA) of the NMR ensemble (PDB code 1EZY) to better understand correspondence of principal modes with elastic network model based normal modes. Such analysis has been quite useful in other proteins as well [31, 32]. These data are presented in **Appendix III**. By matching a single principal component (PC) with a single low-frequency mode via computation of overlaps, we find that PC 1 is significantly correlated to low-frequency Mode 2 in the normal mode analysis (**Figure 2.5**). This suggests that principal motions present in RGS4 can be explained by a single low-frequency mode of RGS4. This also provides a structure-based explanation of underlying dynamics.

Results

RGS4 Motions Allowing Inhibitor Access To Cysteine 95

A single-cysteine mutant of RGS4 was created to specifically understand the effect of CCG-50014 binding to Cys95 (**Figure 2.1**). CCG-50014 is able to perturb the intrinsic fluorescent properties of RGS4 (**Figure 2.2 A**), and

completely block RGS4-Gα binding (**Figure 2.2 B**) through covalent modification of the Cys95 residue. Both effects are lost when an RGS4 mutant lacking cysteine residues is employed. Additionally, CCG-50014 produces a significant decrease in thermal stability of RGS4 Cys95 compared to DMSO treated protein (**Figure 2.2 C**). The decrease in melting temperature is comparable to what has been reported for WT RGS4 in the presence of CCG-50014 [6]. Such a significant decrease in melting temperature suggests a significant change to the overall tertiary structure in the presence of the compound.

Using two independent starting structures of RGS4 (PDB codes 1AGR and 1EZT) [33, 34], we first conducted molecular dynamics (MD) simulations (**Table 2.1**) of apo wild-type RGS4, and a mutant form of RGS4, where all cysteine residues except Cys95 were mutated to Ala (**Figure 2.1**). Conventional MD simulations reveal no significant change in the overall structure of RGS4 over the 20 nanosecond simulation (**Figure 2.3 A**). This is indicated by the relatively low root- mean-squared deviation (RMSD; Cα) comparing the starting and ending structures (red and green traces in **Figure 2.3 C**). To probe the solvent-accessibility of the residue Cys95, we measured the buried surface area (BSA)
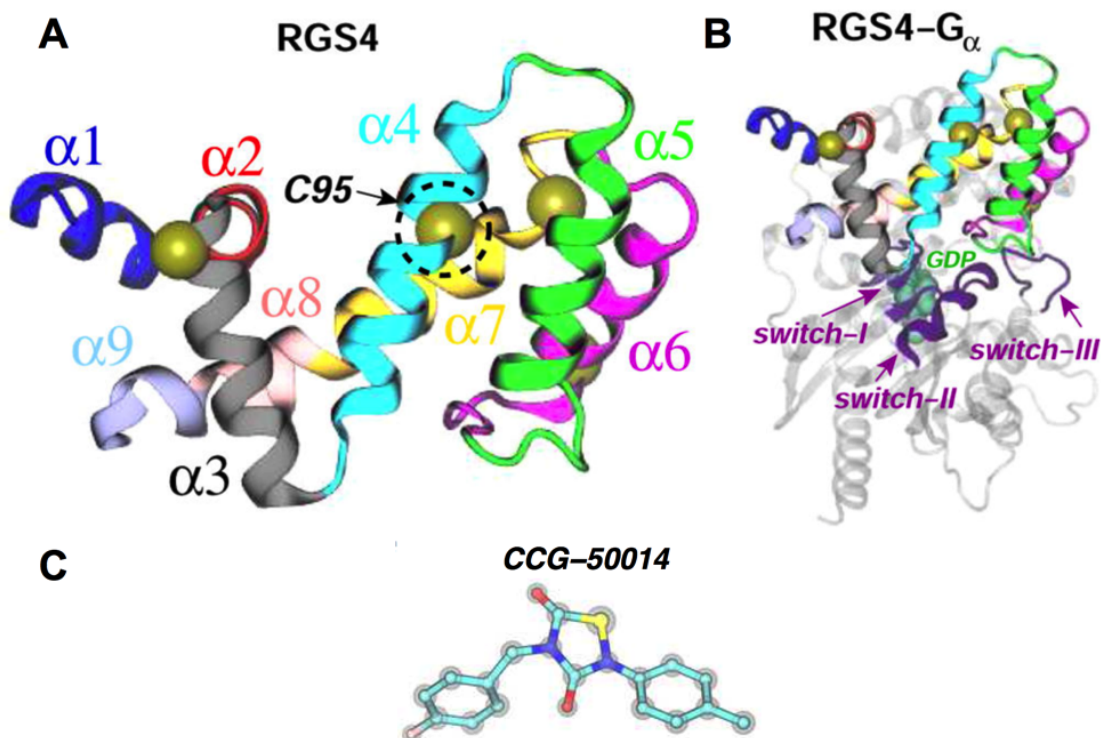
Figure 2.1 Cartoon representations of the box domain of RGS4 and the RGS4-Giα1 complex (**A & B**). Each of the nine helices of RGS4 is distinctly colored and labeled. The Cα atoms of four cysteine residues in RGS4 are shown as yellow spheres, while the buried residue Cys95 (C95) is labeled and highlighted in a dotted circle. In panel B, three switch regions of Gα$_{i1}$ are shown as magenta cartoons and labeled, and the GDP molecule is rendered as green space-filling spheres. **C**) Unconjugated CCG-50014 molecule.

between the α_5-α_6 helix pair and the rest of RGS4 (red and green traces in

Figure 2.3 D) along with the solvent-accessible-surface-area (SASA) of residue

Cys95 (**Figure 2.4**). Throughout each 20-ns MD simulation, we observe no

change in the aforementioned BSA, and the SASA of residue Cys95 remains

negligible(~2−3 Å^2). Therefore, the underlying dynamics of RGS4 that facilitate

access to Cys95 are not yet clear. However, this is not surprising because long

time-scale conformational flexibility in proteins is seldom revealed using short

unbiased MD simulations due to the underlying free-energy barriers. Therefore,

Run No. System	Initial Coordinates	Run Type	Run Length (ns)	Ligand CCG-50014	Construct Type	System Size (atoms)
1 RGS4	1AGR	MD	20	X	WT	28076
2 RGS4	1AGR	MD	20	X	M	28073
3 RGS4	1AGR	TAMD	35	X	WT	28076
4 RGS4	1AGR	TAMD	35	X	WT	28076
5 RGS4	1AGR	TAMD	25	X	WT	28076
6 RGS4	TAMD (Run#4)	MD	40	X	M	31585
7 RGS4	TAMD (Run#4)	MD	40	YES	M	31529
8 RGS4-G_α	1AGR	MD	20	X	WT	67368
9 RGS4-G_α	1AGR[d]	MD	45	YES	WT	69259
10 RGS4	1EZT	MD	20	X	WT	29269
11 RGS4	1EZT	MD	20	X	M	29269
12 RGS4	1EZT	TAMD	50	X	WT	29269
13 RGS4	TAMD	MD	40	YES	M	39436

Table 2.1 Details of all MD and TAMD simulations.

[a]all cysteines but Cys95 were mutated to Ala in all mutant simulations
[b]$\bar{\beta}^{\wedge}(-1)$ = 2 kcal/mol; all other TAMD runs were carried out $\bar{\beta}^{\wedge}(-1)$= 3 kcal/mol
[c]three independent simulations were carried out in each case except for run#13 where four independent simulations were carried out
[d]initial coordinates of RGS4 were taken from TAMD (run#4), and of Gα were taken from the PDB coordinate file 1AGR; restraints on all
Cα-atoms were enforced for the first 20 ns to maintain RGS4-Gα interaction which was followed by a 25-ns long unrestrained MD equilibration

we used an enhanced sampling algorithm for proteins [22] based upon temperature-accelerated molecular dynamics (TAMD) [19, 20] to carry out conformational sampling of RGS4 (see supporting methods for details). In this algorithm, enhanced conformational sampling of proteins is achieved by exploring the dynamics of relatively rigid subdomains (spatially contiguous groups of residues) by accelerating, via high fictitious temperature, the Cartesian coordinates of the centers-of-mass (COM) of subdomains as collective variables
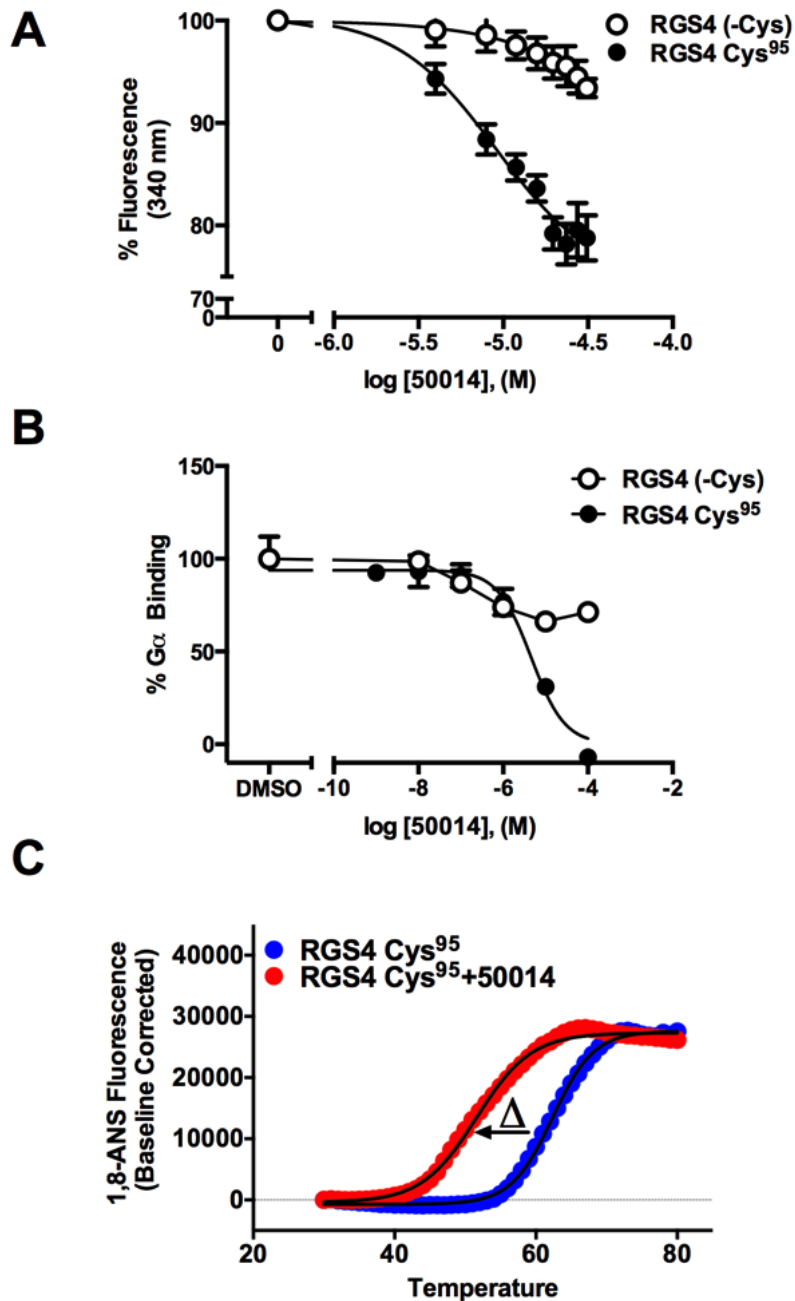
Figure 2.2 The biophysical and functional effects of CCG-50014 targeting Cys95 on RGS4. **A**) CCG-50014 produces strong intrinsic fluorescence changes (>20% reduction) indicating significant structural perturbations resulting specifically from modifying Cys95 **B**) The RGS4-Gα interaction is inhibited by CCG-50014 with 4 µM potency at the Cys95 site. **C**) CCG-50014 decreases the melting temperature of RGS4 Cys95 by 11°C (Tm=62°C for apo vs 51°C in the presence of 50014). In the presence of 50014, the reduction in melting temperature is very close to the 10°C decrease observed for WT RGS4, compared to no change in Gαo melting.

(CVs) [23-26, 35]. We conducted four independent TAMD simulations (**Table 2.1**) of RGS4 at fictitious thermal energies of $\bar{\beta}^{-1}$= 2 and 3 kcal/mol, where $\bar{\beta}$= 1/kB$\bar{T}$, kB is Boltzmann's constant, and $\bar{T}$ is the fictitious temperature. The RMSD and BSA data (magenta traces in **Figure 2.3 C and D**) from a typical TAMD run at $\bar{\beta}^{-1}$ = 3 kcal/mol reveal the existence of relatively large-scale conformational flexibility in the $\alpha_5-\alpha_6$ helix pair, where these helices span open-like conformational states away from the rest of RGS4 (**Figure 2.3 B**). However, similar data from a TAMD run at a lower $\bar{\beta}^{-1}$ = 2 kcal/mol (blue traces in **Figure 2.3 C and D**) suggest that the conformational change in the $\alpha_5-\alpha_6$ helix pair cannot be observed on similar time-scale with thermal energies lower or comparable to 2 kcal/mol, which also explains why this conformational change was not observed in conventional MD simulations (vide supra). We note that this conformational change is reproducibly observed in multiple independent TAMD simulations of RGS4.

Collective Motions Drive Exposure Of Cysteine 95 To CCG-50014

It is well established that large-amplitude collective motions in proteins are governed by their intrinsically accessible low- frequency normal modes [31, 36-39], and conformational sampling via TAMD in Cartesian CVs is also governed by these modes [24]. Therefore, we carried out a Cα-based elastic network normal-mode analysis (NMA) of RGS4 with both initial structures (PDB codes 1AGR and 1EZT), and computed the individual and cumulative projections or
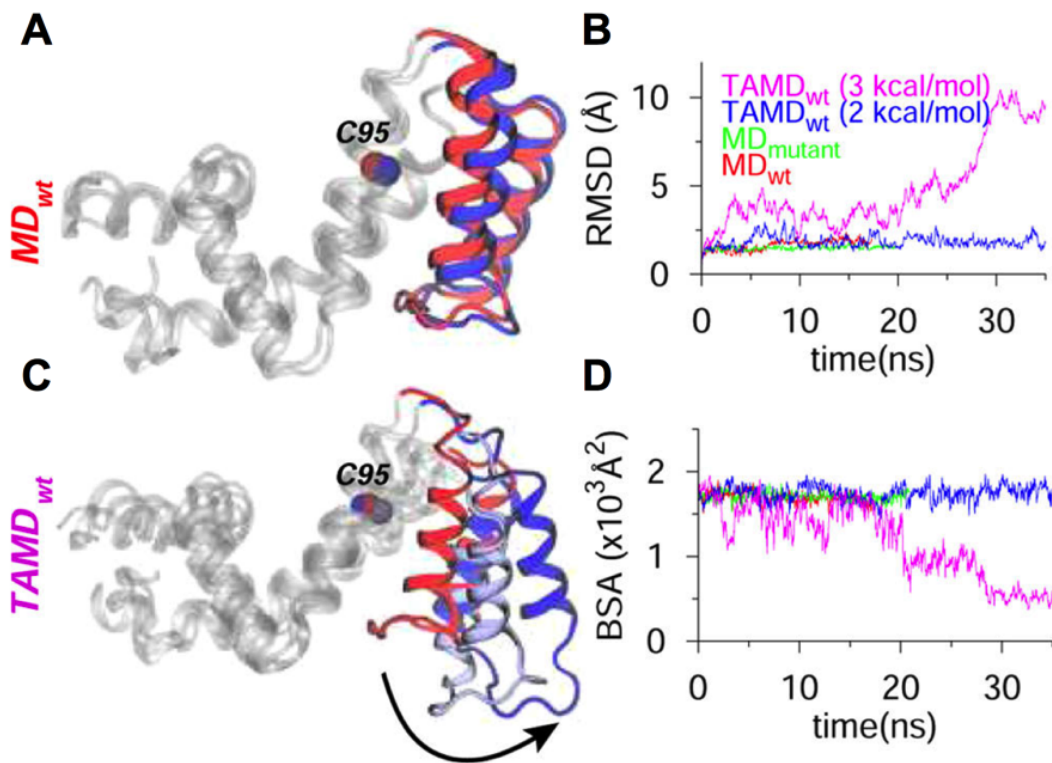
Figure 2.3 MD and TAMD simulation data for RGS4 runs with initial coordinates from PDB code 1AGR. (**A** and **B**) Overlay of cartoon representations of apo-RGS4 (red, beginning; blue, end of simulations). All helices of RGS4, except the α5−α6 pair, are shown in white cartoons. **C**) The Cα-RMSD traces with reference to starting conformations. **D**) Buried surface area (BSA) between the α5−α6 helix pair and the rest of RGS4.

overlaps (θ_m) for 20 low- frequency (nonzero) normal modes of apo RGS4 onto the functional displacement vector ($\vec{d}$) between initial and final conformations generated by a typical TAMD simulation of each structure (**Figure 2.5**). These data suggest that low-frequency modes 1, 3, and 5 consistently contribute to this conformational change (see overlaps, θ_m, for these modes in Figure 3), while either mode 7 (**Figure 2.5 B**) or modes 8, 9, and 10 (**Figure 2.5 A**) can contribute significantly to this conformational change. Also, the first 10 low-

50
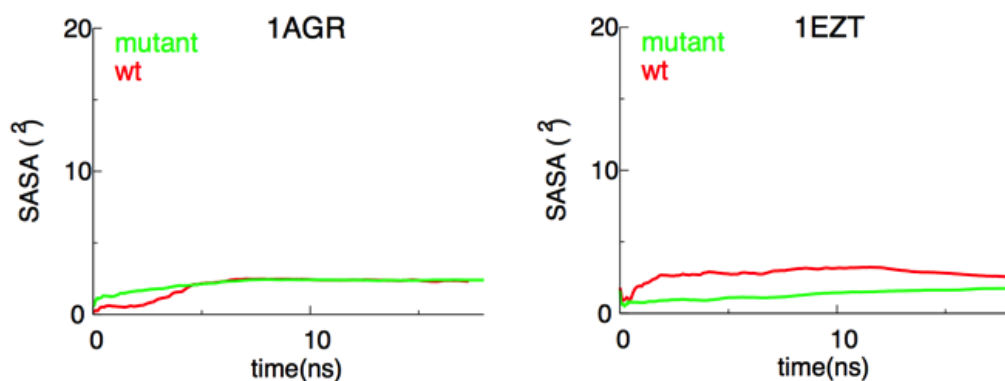
Figure 2.4 Solvent Accessible surface area (SASA) for residue Cys95 from MD equilibration trajectories of wild-type and mutant-RGS4 (runs#1, 2, 10 and 11 in Table S1). Data for MD simulations starting with two different initial coordinates (PDB codes 1AGR and 1EZT) are shown. The SASA traces for mutant-RGS4 runs are averages over three independent MD runs.

frequency modes can describe the majority of the conformational change in the $\alpha_5-\alpha_6$ helix pair, as indicated by the cumulative overlap traces (red dotted lines in **Figure 2.5**). These results reinforce the point that the conformational flexibility observed in the TAMD simulations is intrinsic to the structure of RGS4 due to the presence of such low-frequency (high-amplitude) modes. This also means that our NMA analysis supports the definition of subdomain COM as our collective variables in TAMD since both calculations show significant overlap in conformational changes.

A key implication of the conformational change in the $\alpha_5-\alpha_6$ helix bundle is that the side-chain of residue Cys95 is now accessible, and the inhibitor molecule could be successfully accommodated in the TAMD-generated conformations of RGS4. To understand specifically the effect of CCG-50014 binding to Cys95, we
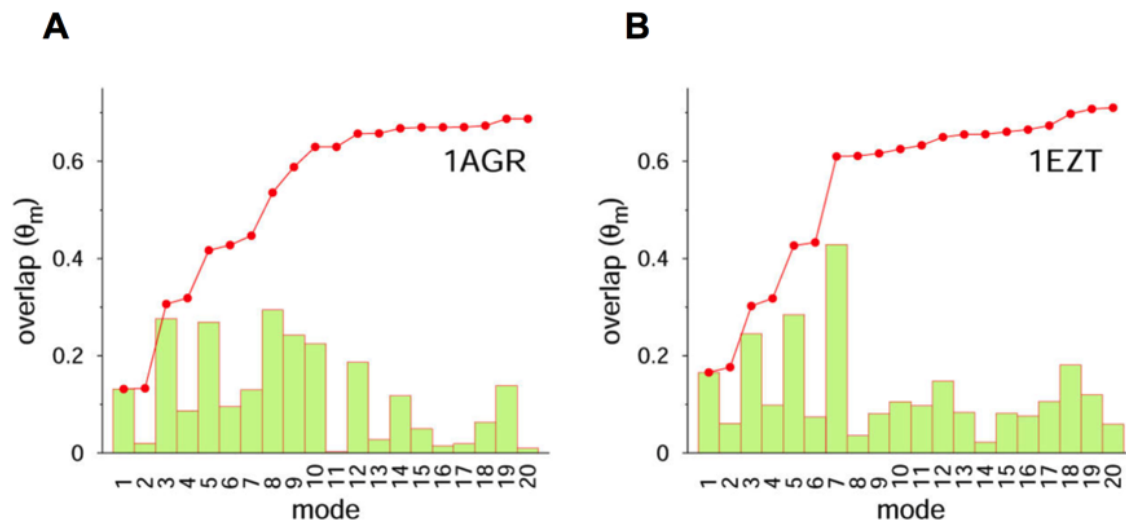
Figure 2.5 Individual (bars) and cumulative (red dotted line) projections of 20 low-frequency normal modes of RGS4 onto TAMD generated conformations; projection of the nth mode θ_{mn} is the normalized scalar product between the eigen vector $(\vec{I}_n)$ of the corresponding mode and functional displacement $\vec{d}$, that is, $\theta = \vec{I}_n \cdot \vec{d}/(|\vec{I}_n| \times |\vec{d}|)$, while the cumulative overlap is $(\theta_{m1}^2 + \theta_{m2}^2 + \cdots + \theta_{mn}^2)^{1/2}$. The functional displacement vector $\vec{d}$ is defined as the vector difference between starting and ending conformations from a typical TAMD simulation.

carried out seven independent 40-ns long MD equilibration trajectories (**Table 2.1**) of the single-cysteine mutant of RGS4 with a covalently linked inhibitor. In 6 out of 7 simulations, we observe that intercalation of CCG-50014 into the helix bundle restricts closure of the $\alpha_5-\alpha_6$ helix pair (**Figure 2.6 A**), as indicated by a nearly constant BSA between these two helices and the rest of RGS4 (magenta traces in **Figure 2.6 D**). This is primarily because the covalently linked (to Cys95) inhibitor molecule experiences only small conformational rearrangements due to minor rotations of aromatic groups (**Appendix IV A**). However, in one simulation, we observe that CCG-50014 rotates (around the disulfide bond) by ~90° with respect to its initially docked conformation (**Appendix IV B**), which also allows

52
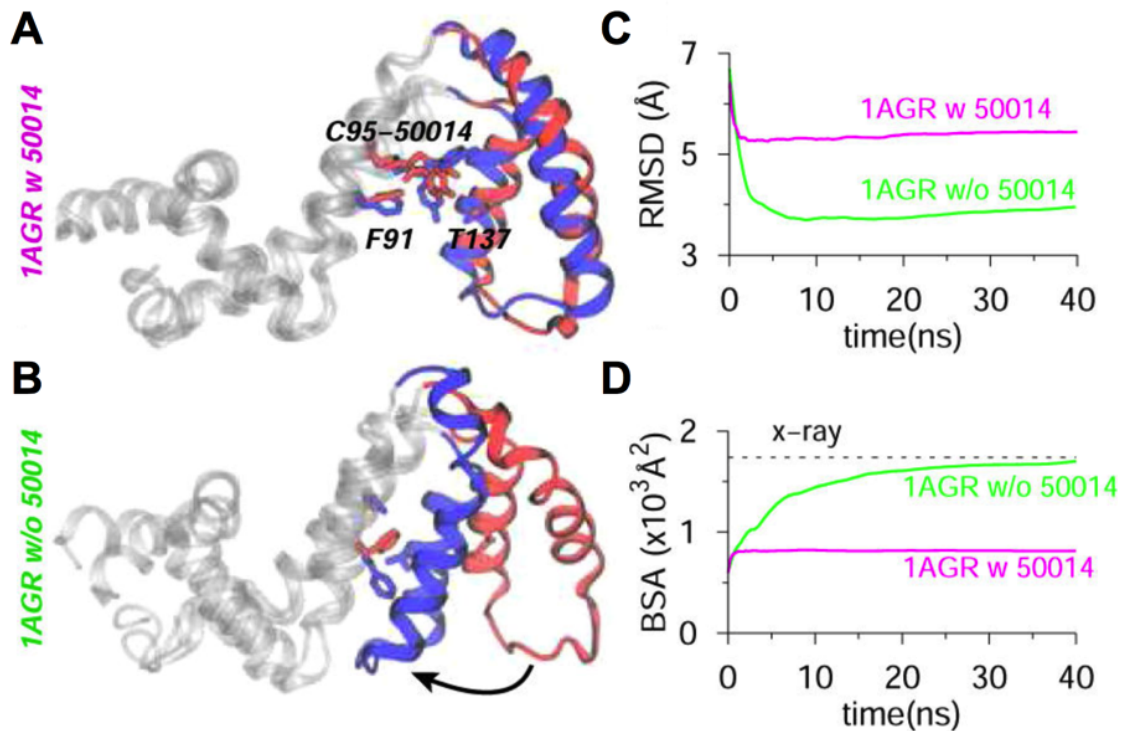
Figure 2.6 Data from MD equilibrations of TAMD-generated conformations of RGS4 (PDB code 1AGR) with **A**) and without **B**) CCG-50014 (red, beginning; blue, end of simulations). **C**) RMSD with reference to the crystallographic conformation of RGS4 (PDB code 1AGR). **D**) Same BSA as in Figure 2.3 are shown.

partial closing of the $\alpha_5-\alpha_6$ helix pair (see snapshots, and the BSA trace in **Appendix V**). To further understand the role of the inhibitor in keeping the $\alpha_5-\alpha_6$ helix pair in a relatively open conformation, we also carried out a 40-ns long MD equilibration of the TAMD-generated open-conformation of RGS4 in the absence of CCG-50014. The data from this simulation showed a spontaneous recovery of the $\alpha_5-\alpha_6$ helix pair to a near crystallographic conformation (snapshot in **Figure 2.6 B**, and green traces for RMSD and BSA in **Figure 2.6 C & D**) unlike the inhibitor-bound simulation (magenta traces for RMSD and BSA in **Figure 2.6 C & D**). This suggests that apo-RGS4 likely exists predominantly in a closed-

53

conformation, but the open-conformation that exposes Cys95 is favored when RGS4 is bound to CCG-50014.

To further dissect the effect of inhibitor binding on the structure of RGS4, we computed the root-mean-squared- fluctuation (RMSF) per residue from three independent MD simulations of mutant-RGS4 with and without CCG-50014 (**Figure 2.7**). Based upon these data, we show a change in RMSF per residue on binding of the inhibitor (**Figure 2.7**). On inhibitor binding, simulations from both RGS4 structures (PDB codes 1AGR and 1EZT) reveal a consistent and significant increase in the fluctuations of residues located in the α_1-helix (51−60), and residues (~110 to 135) in the helices α_5, α_6, and the loop region connecting these two helices (see **Figure 2.1** for designation of helices). Additionally, simulations of the NMR structure (PDB code 1EZT) suggest increased flexibility in the C-terminus of the α_3-helix, and helices α_7 and α_8 (bottom panel in **Figure 2.7**) on binding of CCG-50014. These results are consistent with the observed inhibition of binding between the inhibitor-bound RGS4 and the activated Gα-subunit because the residues in the helices α_3, α_7, α_8 of RGS4 directly contact switch-I and switch-II regions on Gα, while residues in the loop region connecting helices α_5 and α_6 directly contact the switch- III region on Gα (see **Figure 2.1 B** for RGS4-Gα complex). The increase in the fluctuations of residues located away from the inhibitor binding site suggests an allosteric effect of CCG- 50014 binding, which may contribute to an unstable RGS4-G$_\alpha$ interface.
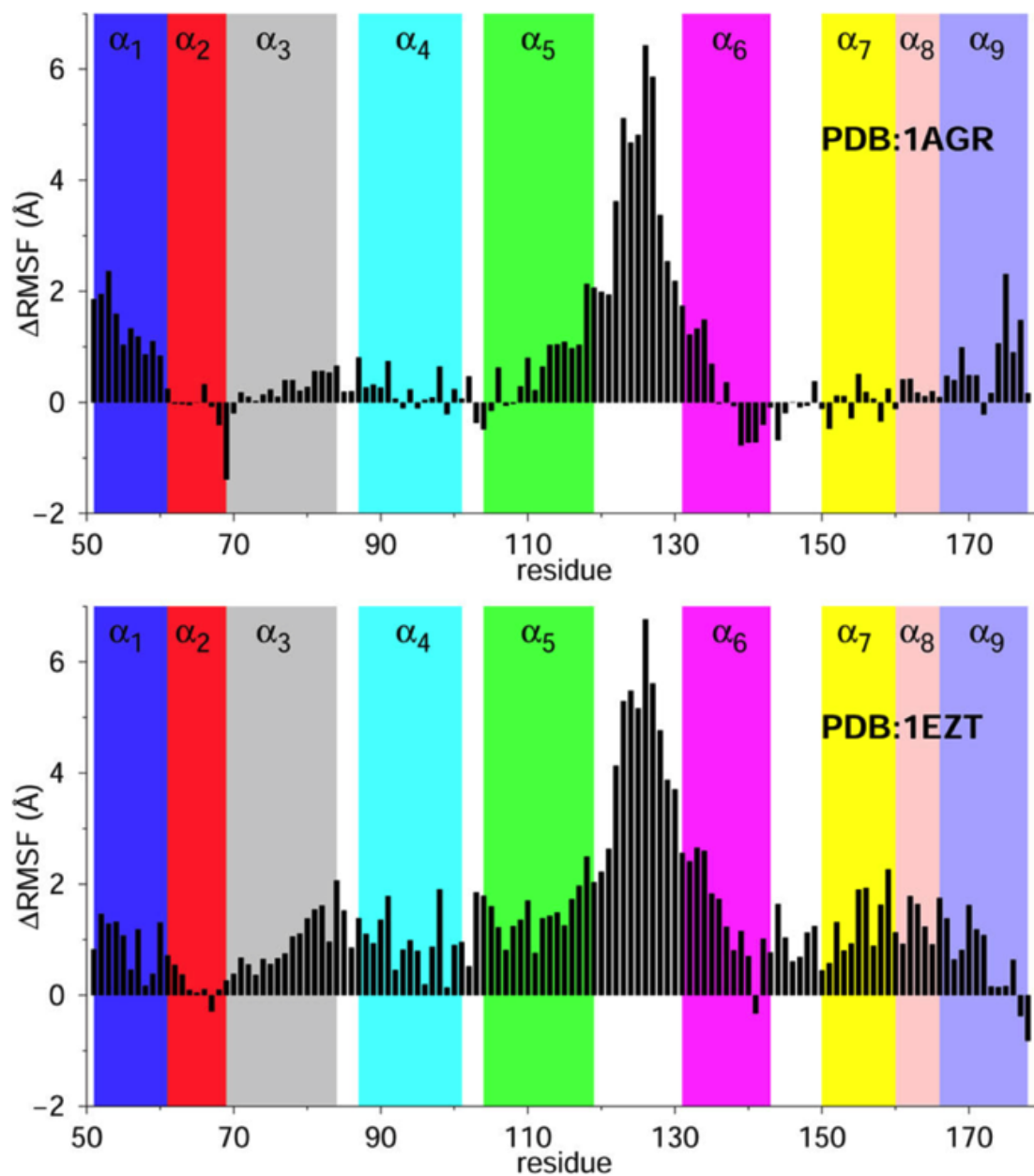
Figure 2.7 Change in RMSF per residue (ΔRMSF) on binding of CCG-50014 to RGS4. Data represent difference in RMSF from at least three independent MD simulations, each for apo and inhibitor-bound structures, of RGS4.

Structural Perturbations Induced By CCG-50014

Using NMR methods for RGS4 [29, 33] structural analysis of the effects of

CCG-50014 on RGS4 were carried out. Specifically, we recorded ^{1}H$-^{15}$N HSQC

spectra (**Figure 2.8**) for two different constructs of RGS4 (cysteine-less RGS4, cysless-RGS4; and single-cysteine RGS4, Cys95-RGS4) in the presence of CCG-50014 at a 1:1 molar ratio. The absence of peak shifts, new peaks, or significant change in signal intensity in the HSQC spectrum of cysless-RGS4 in the presence of CCG-50014 (top panel in **Figure 2.8**) suggest that the compound is not interacting with the protein. This is consistent with the mechanism of (cysteine) thiol modification necessary for activity of CCG-50014 [6, 13]. However, CCG-50014 modification of Cys95-RGS4 produces significant perturbations in the HSQC spectrum in the form of attenuation of existing peaks and the appearance of new peaks, as observed in the spectral overlay of CCG-50014-treated protein with the DMSO-treated reference sample (bottom panel in **Figure 2.8**). The appearance of new peaks with lower dispersion in the presence of CCG-50014 demonstrates that the perturbed regions of the structure undergo greater conformational averaging than in apo-RGS4. Consistent with extensive conformational changes in the presence of CCG-50014, the normalized intensity ratios of the HSQC peaks before and after the CCG-50014 modification (**Figure 2.9**) indicate that signal intensity decreased for most of the peaks between residues 61 and 182 (**Figure 2.9 B**) with the exception of those in the unstructured portion of the C-terminus of RGS4 starting with Gly184.
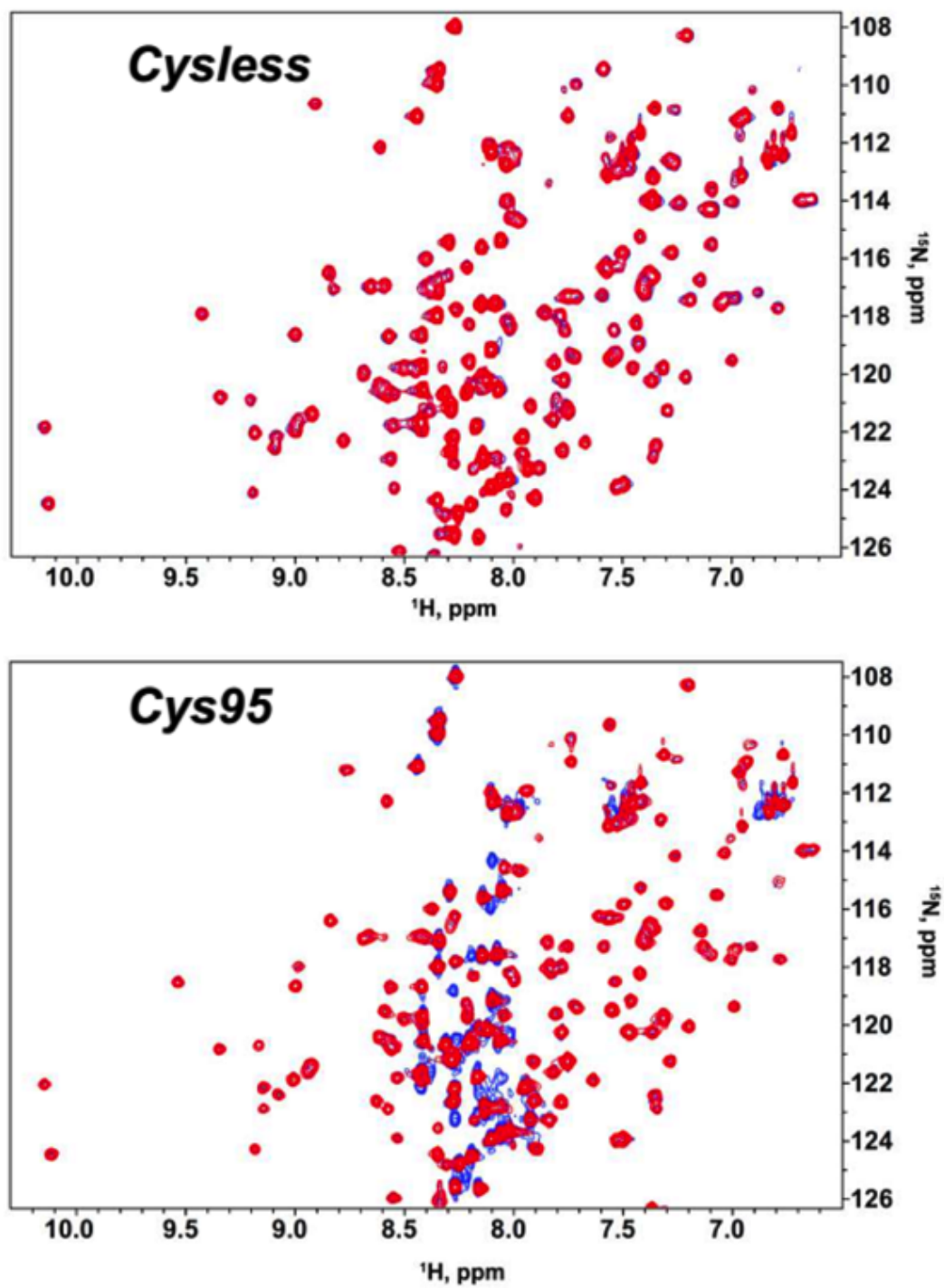
Figure 2.8 ^{1}H-^{15}N HSQC spectra of Cysteine-less (Cysless) and single Cysteine (Cys95) RGS4 are shown in the top and bottom panels. Spectra were recorded before (red) and after (blue) CCG-50014 exposure.

In the presence of the covalent inhibitor, where $k^{off} \ll$ the chemical shift time scale, structural perturbations or inhibitor interactions result in partial or full attenuation of peaks as they are split between two (or possibly more) chemical states [40, 41]. Therefore, we computed % attenuation in peak intensity per residue on CCG-50014 binding (**Figure 2.9 C**), and mapped residues experiencing above-average structural perturbations on the RGS4 structure (**Figure 2.10 A**). These data suggest significant perturbations in many residues (91, 101, 110, 142, and 153) in the vicinity of Cys95, which is likely a result of direct interaction with the aromatic rings of the inhibitor, as seen in MD simulations (**Appendix IV & V**). It is worth noting that the most perturbed peak in the HSQC spectrum comes from Phe91 (**Figure 2.9 C**), a residue located exactly below Cys95, and in direct contact with one of the phenyl rings of CCG-50014 in our simulations (**Appendix IV**). Consistent with RMSF results (**Figure 2.7**), additional significant perturbations are also seen in residues (61, 63, 64, 66, 77, 167, and 178) in the N- and C-terminal helices of RGS4, as well as residues (124 and 126) in the loop region connecting helices α_5 and α_6. Based upon our MD trajectories, we also carried out a theoretical chemical shift perturbation (CSP) analysis using software packages SHIFTX2 [27] and SPARTA+ [28] (**Appendix VI**). Overall, we find that ~66% of perturbed residues seen in NMR experiments are also consistently observed in MD simulations (**Figure 2.10 A and B**). The correlations between strong perturbations observed experimentally and predicted from the simulations are plotted together in **Figure 2.10 C**. Many of the strongly perturbed residues in theoretical CSP analysis that do not directly

58
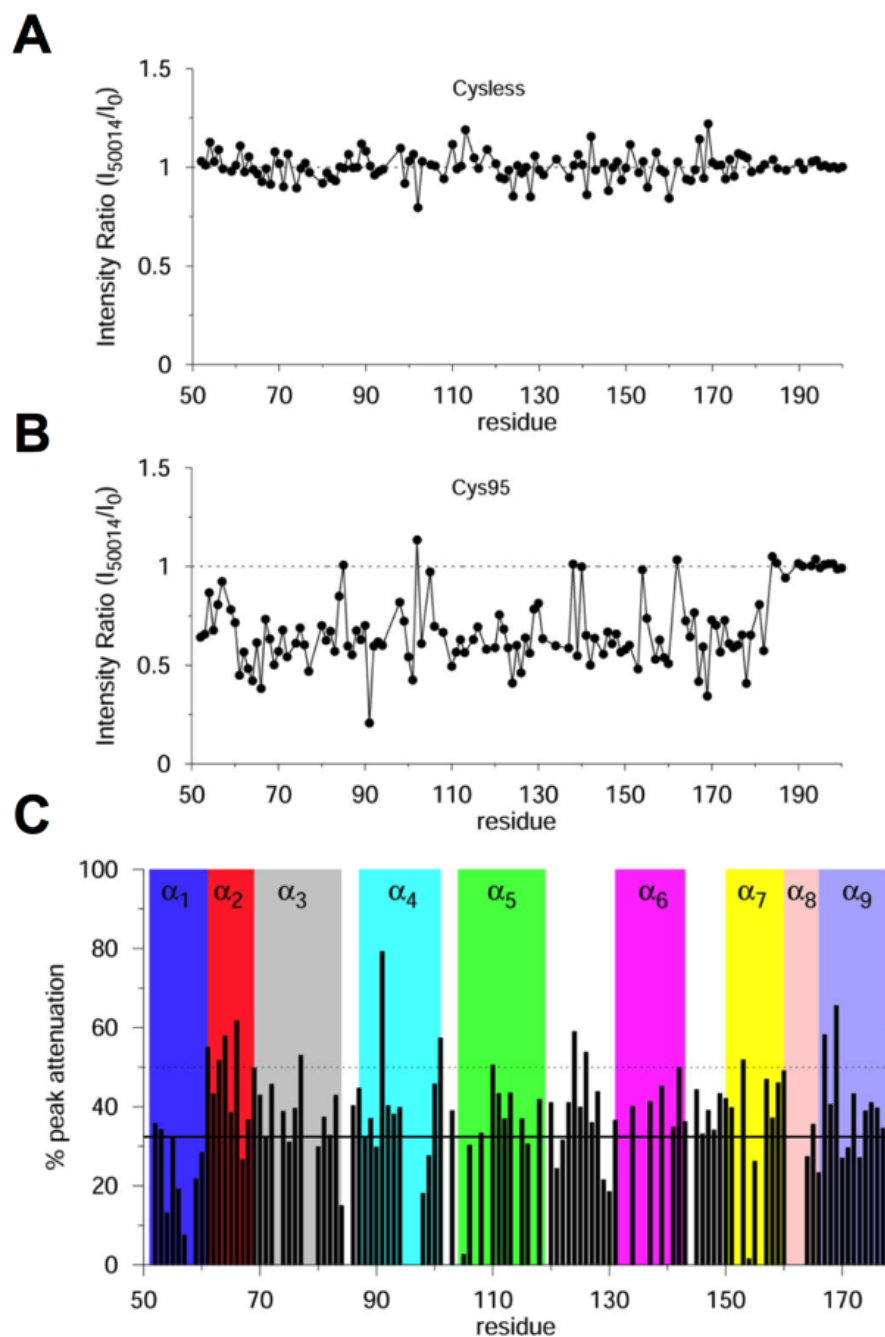
Figure 2.9 Effect of CCG-50014 on the $^{1}H-^{15}N$ HSQC spectra of RGS4. (**A and B**) Normalized intensity ratios of the HSQC peaks after and before CCG-50014 modification as a function of residue number for the cysteine-less (cysless) and single cysteine (Cys95) RGS4. **C**) Percentage attenuation in peak intensity for each residue after treatment with CCG-50014. Solid and dotted horizontal lines indicate values at mean and 1 standard deviation (SD).
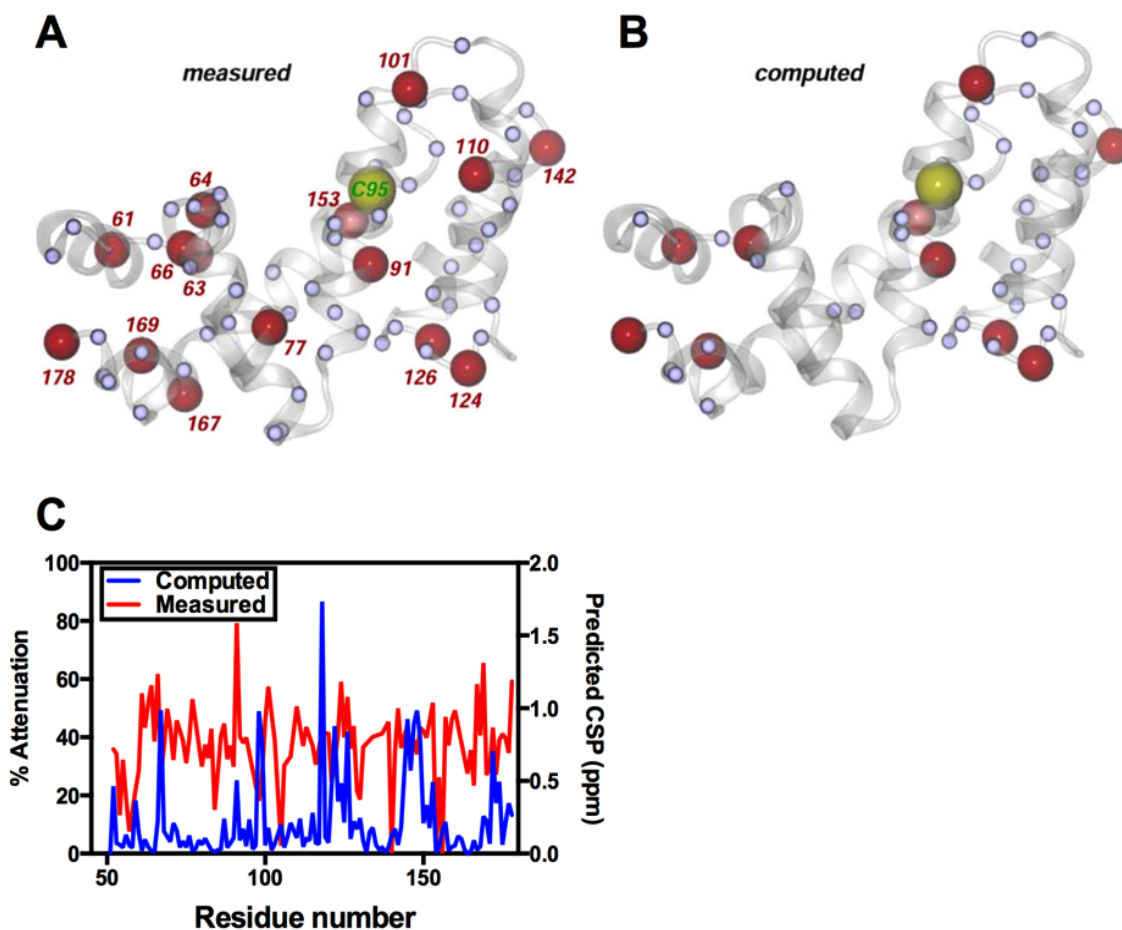
Figure 2.10 Comparison of perturbed residues on the RGS4 structure (PDB code 1AGR). **A**) Measured (from NMR experiments) and **B**) computed (from MD simulations). Red labeled spheres indicates residues with % peak attenuation larger than 1 SD, while small blue spheres are residues with % peak attenuation between mean and 1 SD in NMR experiments. Residue Cys95, which is covalently modified by CCG-50014, is shown as a labeled yellow sphere. The protein backbone is rendered as a transparent white cartoon. **C**) Plot comparing the predicted chemical shift perturbations measure in ppm with the attenuated HSQC cross peaks in the presence of CCG-50014.

correlate with strongly perturbed residues in the HSQC experiment show increased flexibility in the RMSF data as a result of the interaction with CCG-50014 (**Figure 2.7**). Some disagreement between the experimental observations and predictions from simulations may be due to a variety of factors such as difficulty in unambiguously resolving overlapping peak positions in the NMR

60

spectra, and the absence of certain rotameric conformational states of residues that occur on long time-scale [42]. Consistent with a loss of G-protein binding and GAP activity [6], residues in the G-protein binding interface, such as Thr124 and Glu126, are perturbed both in experiments and simulations in the presence of CCG-50014.

CCG-50014 Prevents Key Contacts Between RGS4 And Gα

Finally, we carried out MD simulations of the RGS4-Gα complex with and without inhibitor to better understand the role of CCG-50014 in inhibiting this protein−protein interaction. For each simulation, we assess the stability of interaction between RGS4 and Gα by measuring (a) BSA between the entire RGS4 and Gα structures in the complex and (b) total nonbonded interaction energy and its components between key residues (see **Figure 2.11 A**) in the RGS4- Gα interface. These metrics, computed from a 20-ns MD equilibration of the apo crystal structure of RGS4-Gα complex (PDB code 1AGR), indicate that the interface between RGS4 and Gα remains preserved (**Figure 2.11 B**, red trace for BSA in **Figure 2.11 D**, and the interaction energy traces in **Figure 2.11 E**). The interaction energy traces (**Figure 2.11 E**) suggest that electrostatic interactions significantly contribute to the total non-bonded energy. Because the conformation of RGS4 used in the inhibitor-bound complex simulation was generated via TAMD, we enforced restraints (on all Cα atoms) for the first 20 ns of MD simulation to maintain the integrity of RGS4-Gα interface, and thereafter continued a 25-ns long free (unrestrained) MD equilibration. We observe that although RGS4-Gα interface remains largely intact during the 20-ns long

restrained simulation (green trace in the gray background in (**Figure 2.11 D**), it
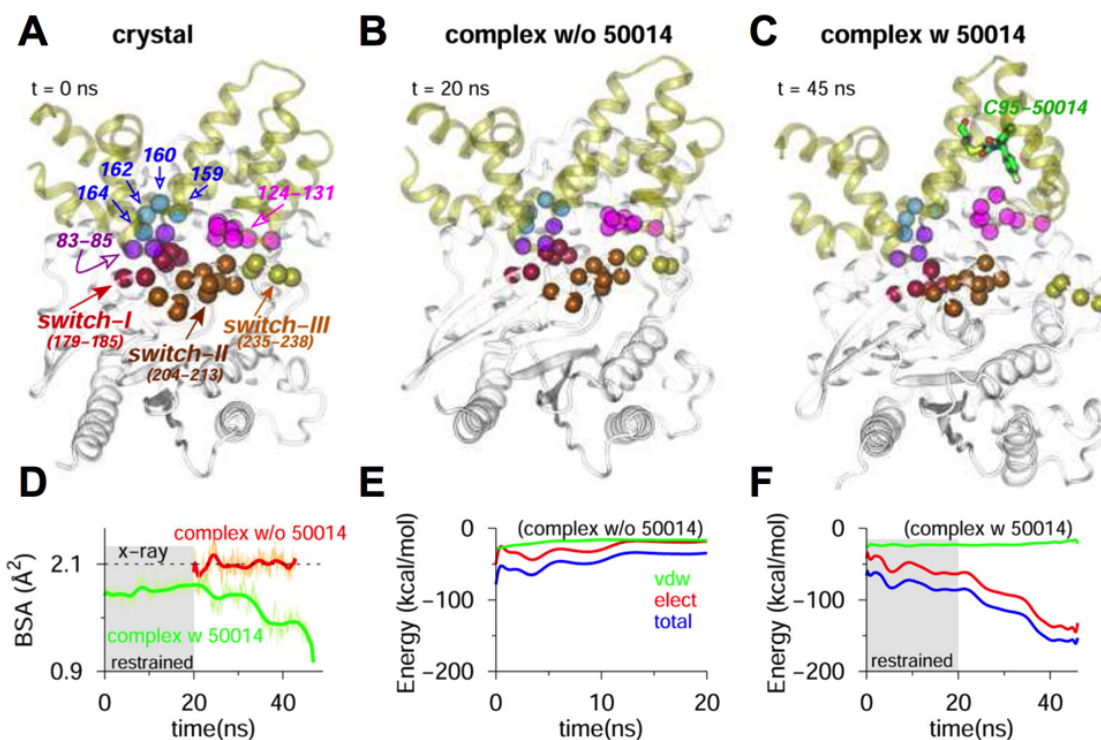
Figure 2.11 Snapshots of RGS4-G$_\alpha$ complex in **A**) crystallographic conformation (PDB code 1AGR); and conformations at the end of MD equilibrations without (**B**) and with (**C**) inhibitor. The Cα-atoms of key interacting residues in the RGS4-Gα interface are labeled (panel A), and rendered as space-filling spheres in each snapshot. **D**) Traces of BSA between RGS4 and Gα in the complex (green, with inhibitor; and red, without inhibitor). **E−F**) Traces for nonbonded interaction energy and its components computed between residues of RGS4 and Gα (see panel A) from MD simulations without (panel E) and with (panel F) inhibitor. (Gray background in panels D and F) The first 20 ns of restrained MD simulation with inhibitor. The red trace in panel D is deliberately shifted by 20 ns to highlight the fact that simulation was not restrained.

gradually becomes unstable on removing the restraints (green trace in the white background in **Figure 2.11 D and C**) due likely to a loss of dominant electrostatic interactions in this interface (**Figure 2.11 F**). These data further suggest that binding of CCG-50014 to Cys95 on RGS4 results in allosteric perturbations to residues in the RGS4-Gα interface thereby inhibiting this protein−protein interaction.

Conclusions

In summary, this study has determined the effect of CCG-50014 (a covalent inhibitor) binding to RGS4 as well as on the interaction of RGS4 with the Gα-subunit. Enhanced conformational sampling of RGS4 using TAMD consistently reveals the existence of open-like conformations of the $\alpha_5-\alpha_6$ helix pair, which facilitates binding of CCG-50014 to the buried side-chain of Cys95. These methods consistently identify a large number of key perturbed residues from NMR experiments and MD simulations, especially in the vicinity of the inhibitor binding site and in the RGS4-Gα interface. The steric effects of CCG-50014 interacting with specific residues in $\alpha_4-\alpha_7$, particularly with Phe91, contribute to the stabilization of the RGS4 open-conformation by restricting the closure of the $\alpha_5-\alpha_6$ helix pair. This allosteric effect translates into structural perturbations downstream in the Gα binding interface of RGS4. Finally, results from a simulation of the RGS4-Gα complex in the presence of CCG-50014 suggest a relatively unstable interface between RGS4 and Gα, unlike the simulation of the apo complex where a stable interface is observed.

References

1.	Ryan, D.P. and J.M. Matthews, *Protein–protein interactions in human disease.* Current Opinion in Structural Biology, 2005. 15(4): p. 441-446.
2.	Chen, Y., et al., *Neurabin scaffolding of adenosine receptor and RGS4 regulates anti-seizure effect of endogenous adenosine.* J Neurosci, 2012. 32(8): p. 2683-95.
3.	Lerner, T.N. and A.C. Kreitzer, *RGS4 is required for dopaminergic control of striatal LTD and susceptibility to parkinsonian motor deficits.* Neuron, 2012. 73(2): p. 347-59.
4.	Blazer, L.L. and R.R. Neubig, *Small molecule protein-protein interaction inhibitors as CNS therapeutic agents: current progress and future hurdles.* Neuropsychopharmacology, 2009. 34(1): p. 126-41.
5.	Arkin, M.R. and J.A. Wells, *Small-molecule inhibitors of protein-protein interactions: progressing towards the dream.* Nat Rev Drug Discov, 2004. 3(4): p. 301-17.
6.	Blazer, L.L., et al., *A nanomolar-potency small molecule inhibitor of Regulator of G protein Signaling (RGS) proteins.* Biochemistry, 2011.
7.	Blazer, L.L., et al., *Reversible, allosteric, small-molecule inhibitors of RGS proteins.* Mol Pharmacol, 2010.
8.	Roman, D., et al., *Allosteric Inhibition of the RGS-G{alpha} Protein-Protein Interaction by CCG-4986.* Mol Pharmacol, 2010.
9.	Zhong, H. and R.R. Neubig, *Regulator of G protein signaling proteins: novel multifunctional drug targets.* J Pharmacol Exp Ther, 2001. 297(3): p. 837-45.
10.	Ishii, M., et al., *Phosphatidylinositol 3,4,5-trisphosphate and Ca2+/calmodulin competitively bind to the regulators of G-protein-signalling (RGS) domain of RGS4 and reciprocally regulate its action.* Biochem J, 2005. 385(Pt 1): p. 65-73.
11.	Popov, S.G., et al., *Ca2+/Calmodulin reverses phosphatidylinositol 3,4, 5-trisphosphate-dependent inhibition of regulators of G protein-signaling GTPase-activating protein activity.* J Biol Chem, 2000. 275(25): p. 18962-8.
12.	Roman, D.L., S. Ota, and R.R. Neubig, *Polyplexed flow cytometry protein interaction assay: a novel high-throughput screening paradigm for RGS protein inhibitors.* J Biomol Screen, 2009. 14(6): p. 610-9.
13.	Turner, E.M., et al., *Small Molecule Inhibitors of Regulator of G Protein Signalling (RGS) Proteins.* ACS Med Chem Lett, 2012. 3(2): p. 146-150.
14.	Vashisth, H., et al., *Conformational Dynamics of a Regulator of G-Protein Signaling Protein Reveals a Mechanism of Allosteric Inhibition by a Small Molecule.* ACS Chemical Biology, 2013.
15.	Phillips, J.C., et al., *Scalable molecular dynamics with NAMD.* J Comput Chem, 2005. 26(16): p. 1781-802.
16.	MacKerell, A.D., et al., *All-atom empirical potential for molecular modeling and dynamics studies of proteins.* Journal of Physical Chemistry B, 1998. 102(18): p. 3586-3616.

17. Mackerell, A.D., M. Feig, and C.L. Brooks, *Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations.* Journal of Computational Chemistry, 2004. 25(11): p. 1400-1415.

18. Yesselman, J.D., et al., *MATCH: An Atom-Typing Toolset for Molecular Mechanics Force Fields.* Journal of Computational Chemistry, 2012. 33(2): p. 189-202.

19. Maragliano, L. and E. Vanden-Eijnden, *A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations.* Chemical Physics Letters, 2006. 426(1-3): p. 168-175.

20. Maragliano, L. and E. Vanden-Eijnden, *Single-sweep methods for free energy calculations.* J Chem Phys, 2008. 128(18): p. 184110.

21. Vanden-Eijnden, E., *Some Recent Techniques for Free Energy Calculations.* Journal of Computational Chemistry, 2009. 30(11): p. 1737-1747.

22. Abrams, C.F. and E. Vanden-Eijnden, *Large-scale conformational sampling of proteins using temperature-accelerated molecular dynamics.* Proc Natl Acad Sci U S A, 2010. 107(11): p. 4961-6.

23. Vashisth, H. and C.F. Abrams, *All-atom structural models of insulin binding to the insulin receptor in the presence of a tandem hormone-binding element.* Proteins-Structure Function and Bioinformatics, 2013. 81(6): p. 1017-1030.

24. Vashisth, H. and C.L. Brooks, *Conformational Sampling of Maltose-Transporter Components in Cartesian Collective Variables Is Governed by the Low-Frequency Normal Modes.* Journal of Physical Chemistry Letters, 2012. 3(22): p. 3379-3384.

25. Vashisth, H., L. Maragliano, and C.F. Abrams, *"DFG-Flip" in the Insulin Receptor Kinase Is Facilitated by a Helical Intermediate State of the Activation Loop.* Biophysical Journal, 2012. 102(8): p. 1979-1987.

26. Vashisth, H., G. Skiniotis, and C.L. Brooks, *Using Enhanced Sampling and Structural Restraints to Refine Atomic Structures into Low-Resolution Electron Microscopy Maps.* Structure, 2012. 20(9): p. 1453-1462.

27. Han, B., et al., *SHIFTX2: significantly improved protein chemical shift prediction.* Journal of Biomolecular Nmr, 2011. 50(1): p. 43-57.

28. Shen, Y. and A. Bax, *SPARTA plus : a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network.* Journal of Biomolecular Nmr, 2010. 48(1): p. 13-22.

29. Storaska, A.J. and R.R. Neubig, *Chapter Eight - NMR Methods for Detection of Small Molecule Binding to RGS4*, in *Methods in Enzymology*, P.M. Conn, Editor. 2013, Academic Press. p. 133-152.

30. Moy, F.J., et al., *1H, 15N, 13C, and 13CO assignments and secondary structure determination of RGS4.* J Biomol NMR, 1999. 15(4): p. 339-40.

31. Yang, L., et al., *Close correspondence between the motions from principal component analysis of multiple HIV-1 protease structures and elastic network modes.* Structure, 2008. 16(2): p. 321-330.

32. Yang, L.W., et al., *Principal component analysis of native ensembles of biomolecular structures (PCA_NEST): insights into functional dynamics (vol 25, pg 606, 2009).* Bioinformatics, 2009. 25(16): p. 2147-2147.

33. Moy, F.J., et al., *NMR structure of free RGS4 reveals an induced conformational change upon binding Galpha.* Biochemistry, 2000. 39(24): p. 7063-73.

34. Tesmer, J.J., et al., *Structure of RGS4 bound to AlF4--activated G(i alpha1): stabilization of the transition state for GTP hydrolysis.* Cell, 1997. 89(2): p. 251-61.

35. Vashisth, H., G. Skiniotis, and C.L. Brooks, *Enhanced Sampling and Overfitting Analyses in Structural Refinement of Nucleic Acids into Electron Microscopy Maps.* Journal of Physical Chemistry B, 2013. 117(14): p. 3738-3746.

36. Hayward, S. and N. Go, *Collective Variable Description of Native Protein Dynamics.* Annual Review of Physical Chemistry, 1995. 46: p. 223-250.

37. Berendsen, H.J.C. and S. Hayward, *Collective protein dynamics in relation to function.* Current Opinion in Structural Biology, 2000. 10(2): p. 165-169.

38. Tama, F. and Y.H. Sanejouand, *Conformational change of proteins arising from normal mode calculations.* Protein Engineering, 2001. 14(1): p. 1-6.

39. Yang, L., G. Song, and R.L. Jernigan, *How well can we understand large-scale protein motions using normal modes of elastic network models?* Biophysical Journal, 2007. 93(3): p. 920-929.

40. Matsuo, H., et al., *Identification by NMR spectroscopy of residues at contact surfaces in large, slowly exchanging macromolecular complexes.* Journal of the American Chemical Society, 1999. 121(42): p. 9903-9904.

41. Girotto, S., et al., *Dopamine-derived Quinones Affect the Structure of the Redox Sensor DJ-1 through Modifications at Cys-106 and Cys-53.* Journal of Biological Chemistry, 2012. 287(22): p. 18738-18749.

42. Robustelli, P., K.A. Stafford, and A.G. Palmer, *Interpreting Protein Structural Dynamics from NMR Chemical Shifts.* Journal of the American Chemical Society, 2012. 134(14): p. 6365-6374.

Chapter III

**Identification and Biochemical Evaluation of Reversible RGS
Inhibitors with Cellular Activity**

Introduction

In order for small molecules to have utility as biological probes or chemical starting points from which to develop therapeutics, the compounds must have cellular activity. Although moderate potency has been reached with early small molecule RGS inhibitors, the lack of cellular activity observed from these compounds limited their utility for further development [1-4]. A major component of the challenge for cysteine-reactive compounds to be effective in a cellular environment is the ability to function in the presence of endogenous reducing agents (e.g. glutathione) that can reach cytoplasmic concentrations of several millimolar [5]. As a result, the cellular reducing environment can quickly inactivate many cysteine-modifying compounds. Despite this challenge, the irreversible inhibitor CCG-50014 was identified as the first example of a small molecule capable of inhibiting RGS function inside a cell [6].

In an effort to identify new RGS inhibitors with cellular activity we employed a novel cell-based calcium assay with regulated RGS4 expression. This system mitigates a major challenge to screening in cellular systems, which is the multiple potential sites of action of the compound in the pathway. By screening compounds in an inducible RGS4 cell line (Doxycycline treated cells), followed by

a counter-screen of the hits in the absence of RGS4 (untreated cells) we could enrich for those that are actual RGS4 inhibitors. Using this approach to identify new RGS4 inhibitors, we screened >300,000 compounds from NIH small molecule repository (MLSMR). This chapter describes the identification process and biochemical characterization of several new RGS4 inhibitors with cellular activity. Like all previously reported RGS4 inhibitors, these compounds are dependent on covalent modification of cysteine residues for activity.

The findings in this chapter show that identified RGS inhibitors operate through a mechanism similar to what has been reported for previous RGS inhibitors [1, 3, 6]. Importantly, several of the covalent RGS inhibitors identified are reversible and have selectivity for RGS4 over other RGS homologs tested. Covalent drugs can provide advantages over nonreactive compounds including greater potency and prolonged therapeutic effect [7], though risk of toxicity associated with reactive compounds [8] has deterred widespread use. Reversible covalent molecules may provide the benefits associated with the covalent interaction, but with reduced physiological and toxicological problems resulting from a permanent association [9-11]. As a result, the identification of compounds with cellular activity that are also significantly reversible provides a step forward for the development of RGS inhibitors. Such compounds should provide important new tools to dissect the role of RGS4 in biology and as a therapeutic target.

The work described in the chapter is the result of a multi-institutional effort involving University of Michigan, Johns Hopkins Ion Channel Screening Center

(JHICC), and the Vanderbilt Specialized Chemistry Center. The high-throughput screen implementation and analysis was performed at JHICC and chemical triage of the hits was assessed at Vanderbilt. I performed the biochemical validation and analysis of RGS inhibitors in this chapter with the following exceptions: 1) FCPIA validation of the primary screening hits was performed by Jian Mie and Levi Blazer. 2) Jian Mei performed the thermal stability measurements of RGS4 under my supervision. The contents of this chapter have been compiled into a publication in Cellular Signaling [12].

Materials and methods

Materials

Chemicals were purchased from Fisher Scientific (Hampton, NH) or Sigma-Aldrich (St. Louis, MO). All materials are at least reagent grade. Avidin-coated Luminex beads were purchased from Luminex (Austin, TX). Ni-NTA resin was purchased from Qiagen (Valencia, CA). Amylose resin was obtained from New England Biolabs (Ipswich, MA). Antisera were from Santa Cruz Biotechnology (Santa Cruz, CA).

M3-R4 Cell-Line Development And Characterization

The Invitrogen Flp-In T-Rex HEK 293 cells stably expressing the Tet repressor (pcDNA6/TR) and lacZ-Zeocin fusion gene (pFRT/lac-Zeo), containing the Flp Recombination Target (FRT) site, were used as host cells. HA-tagged RGS4 (C2S) was ligated into a pCDNA5/FRT/TO vector. Flp-In cells were plated

in 6-well plates at 400,000 cells/well and co-transfected with 0.4 μg of RGS4-pCDNA5/FRT/TO and 3.6 μg of pOG44 (expressing Flp recombinase) using 10 μL Lipofectamine 2000 reagent. Stable integration of the RGS4-containing vector occurs between the FRT sites orienting the SV40 promoter and initiation codons in frame with the Hygromycin resistance gene, while inactivating the lacZ-Zeocin fusion gene, making the stably transfected cells Hygromycin resistant and Zeocin sensitive. Two days after transfection, 200 μg/mL Hygromycin was added to the wells to select for stably transfected cells. Cell pools were tested for Zeocin sensitivity and Doxycycline induced RGS4 expression was verified via Western blot. RGS4 expressing cells were then transfected with human M$_3$-muscarinic receptor cloned into pCDNA3.1(+) using 4 μg of plasmid and 10 μL of Lipofectamine 2000, followed by selection of neomycin resistant clones using G418. Single M3-R4 cells were then flow sorted into two 96-well plates. Isolated single clones were expanded and tested for carbachol response and an RGS effect using the Fluo4 NW calcium signaling assay according to the manufacturer's instructions.

Western Blot For RGS4 Expression

Cell transfections and preparation of lysates were performed as described previously [13]. In brief, HEK 293T cells maintained in DMEM plus 10% FBS were grown to confluence in 6-well plates and transiently transfected with 2.5 μg of HA-RGS4 or empty vector (mock) with 4 μL of Lipofectamine 2000 per microgram of DNA, followed by incubation for 48 hours. The HEK293-FlpIn-

TREx/M$_3$R/RGS4 cells were similarly plated in the presence or absence of 1 μg/mL Doxycycline for 45 hours to induce RGS4 expression. Cell lysates were prepared by removing DMEM/FBS medium and rinsing wells with PBS at room temperature followed by the addition of 350 μL of lysis buffer plus protease inhibitors at 4ºC. Cells were scraped and transferred to a microcentrifuge tube and allowed to tumble at 4ºC for 1 hour. Lysates were pelleted at 13,000 rpm for 15 minutes and the supernatant protein was quantified using Bradford reagent. Cell lysates were then run on a 12% SDS-PAGE gel and transferred to Immobilon-P transfer membrane (Millipore, Billerica, MA) and probed with either rabbit anti-HA at 1:400 and rabbit anti-Actin at 1:500, followed by probing both with anti-rabbit HRP secondary at 1:8,000.

High-Throughput Cellular Screen

High-throughput screening was performed at Johns Hopkins Ion Channel Center, Johns Hopkins University, School of Medicine. The Molecular Libraries Small Molecule Repository (MLSMR) collection of > 300,000 compounds was used to screen for inhibitors of RGS4 using the HEK293-FlpIn-TREx/M$_3$R/RGS4 cell line described above. Cell plating was achieved using high density cell freezes that were diluted to 200,000 cells/mL in DMEM (high glucose with glutamine) 10% FBS, 1% Pen/Strep, 15 μg/mL Blasticidin, 400 μg/mL G418, and 200 μg/mL Hygromycin. RGS4 expression was induced with the addition of 10 ng/mL Doxycycline. The day before the assay, 50 μl/well of the diluted cells was plated in black, clear bottom, poly-D-lysine coated 384-well plates and incubated

overnight at 37°C with 5% CO_2. On the day of the experiment the cell medium was removed and 20 μl/well of Fluo4-AM dye solution was added, followed by a 30 minute incubation at 37°C. The dye solution was removed followed by addition of 20 μL of assay buffer (Hank's Balanced Salt Solution in HEPES, pH 7.4) with a second 30 minute incubation at room temperature. The cells were then treated with 4 μL of 6x compound in assay buffer (10 μM final) for 20 minutes at room temperature. Plates were imaged using a Hamamatsu FDSS 6000 kinetic plate reader and a baseline was recorded for 10s at 1Hz. This was followed by injection of 4μl of 7x carbachol (7 nM) and calcium-mediated fluorescence was recorded for 100 seconds. The fluorescence ratio (maximum minus the minimum intensity divided by the baseline over the 100 seconds) was integrated for each well, followed by B-score normalization [14]. Hit selection was based on two criteria, that a compound's initial-fluorescence B-score is within five standard deviations of the mean initial fluorescence ratio B-score of the library, and the B-score of the compound following carbachol injection is greater than three standard deviations above the mean. Compounds that lacked activity in the absence of Doxycycline were confirmed in concentration-response format using concentrations 1 nM – 30 μM in a 1:3 serial dilution in duplicate.

Protein Expression And Purification

N-terminally truncated Δ51 RGS4 (residues 52-205) was expressed as a single open reading frame with maltose binding protein (MBP) fusion protein N-terminally fused via a linker region containing a 10x histidine sequence and a

tobacco etch virus (TEV) cleavage site. Cysless Δ51 RGS4 RGS4 (cys(-) RGS4), described previously [3], was prepared via mutation of the seven cysteines within the RGS-box to alanine: Cys 71, 95, 132, 148, 183, 197, 204. RGS19 C-terminally truncated (ΔC11) and full length RGS16 were both expressed as MBP fusions. All MBP fusion proteins were prepared as previously described using the pMALC2H10 vector [3]. RGS8 (residues 61-198) was expressed with an N-terminal 6x his-tag in a pQE80 vector, as described previously [15]. $G\alpha_{i1}$ protein was expressed with an internal 6x his-tag and $G\alpha o$ using an N-terminal 6x his-tag, both using a pQE80 vector, as described previously [16].

Chemical Labeling Of Purified RGS And $G\alpha_o$ Proteins

RGS proteins were biotinylated on free amines and $G\alpha_o$ was labeled with AlexaFluor-532 on free thiols exactly as described before [17].

Flow Cytometry Protein Interaction Assay (FCPIA)

FCPIA was performed as previously described with the use of AlexaFluor 532-tagged Gαo and biotinylated RGS proteins [17]. To assay the compounds in concentration-response format, RGS protein was immobilized on streptavidin-coated Luminex beads in assay buffer containing 50 mM HEPES buffer pH 7.4 with 100 mM NaCl, 0.1% lubrol and 1% BSA, and treated with compound or DMSO for 15 minutes at room temperature in 96-well plates (Genemate T-3082-1). Gαo-tagged AlexFluor 532 (30 nM final) was mixed with $GDP \cdot AlF_4^-$ and 1 mM MgCl, then added to the RGS beads and allowed to incubate at room

73

temperature for 30 minutes before measuring binding. Compounds were tested in concentration-response from 100 nM-100 µM. To determine reversibility, RGS-coated beads were incubated with 100 µM compound or DMSO for 30 minutes at room temperature. The beads were then washed three times with assay buffer. The beads solution was then split into two groups in which one group was retreated with 100 µM compound (to determine maximum inhibition) and the other DMSO, then each added to a separate Gαo mixture as before. Binding between RGS and G-protein was observed using a Luminex 200 flow cytometer. GraphPad Prism software v5.01 was used for the non-linear regression analysis of inhibition curves.

Receptor-independent, Steady-state GTPase Accelerating Protein Assay

In order to measure changes to RGS4 stimulation of G-protein steady-state GTP hydrolysis, independent of a receptor, an R178M/A326S mutant of $G\alpha_{i1}$ was utilized [18]. GTP hydrolysis was monitored colorimetrically using a malachite green dye that increases absorbance at 630 nm in complex with free inorganic phosphate. Malachite green was prepared using the method described previously [19]. Compounds were serially diluted in 50 mM HEPES at pH 7.4, including 100 mM NaCl, 0.01% Lubrol, 5 mM MgCl, and 10 µg/mL BSA using half-log steps between 100 nM-100 µM in 384-well plates (Corning 3680 clear flat-bottom plates, Corning, NY). To each well, 200 nM Δ51 RGS4-MBP and 6 µM $G\alpha_{i1}$ (R178M/A326S) was added before initiating the reaction with 300 µM GTP in 8 µL final volume. The reaction was incubated at room temperature for

120 minutes followed by quench with 10 µL of HCl/Malachite green dye, then immediately followed by 2uL of a 32% w/v stock solution of sodium citrate as a colorimetric stabilizer. Each plate was incubated at room temperature for 20 minutes before reading Malachite green absorbance at 630 nm using a Victor II plate reader (Perkin Elmer). GraphPad Prism software v5.01 was used for the non-linear regression analysis of inhibition curves.

Thermal Stability Measurements

Thermal stability changes of $G\alpha_{i1}$ or $\Delta51$ RGS4 (cleaved from MBP) in the presence of compound were determined using a Thermofluor instrument (Johnson & Johnson, Langhorne, PA). Each protein (10 µM final) was mixed with compound (100 µM final) or DMSO, 1-anilinonaphthalene-8-sulfonic acid (1,8 ANS) at a final concentration of 200 µM, and overlaid with 5 µL of silicon oil before incubating for 15 minutes at room temperature in a black 384-well PCR plate (Fisher Scientific, Cat# TF-0384K). Thermal stability was measured between 30-90 °C increasing the temperature by 1 °C with intervening fluorescence measurements at 25 °C for each point. The melting temperature (Tm) was determined by applying a sigmoidal fitting procedure to determine the midpoint transition from the folded to unfolded state, using GraphPad Prism v5.01.

Results

High-Throughput Cellular Screen For RGS4 Inhibitors

An M3-R4 cell-line was developed to efficiently identify compounds capable of inhibiting RGS4 in a cellular environment. These cells stably express human M_3-muscarinic receptor with Doxycycline (Dox)-regulated RGS4 expression. The full length RGS4 expressed in these cells contains a cysteine to serine point mutation at position two in the peptide sequence (C2S), which abrogates proteosomal degradation to enhance cellular levels of RGS4 [13]. Stimulation of the cells with carbachol produces a $G\alpha_q$-mediated calcium transient that is monitored using the calcium-sensitive Fluo4 fluorescent dye. Overnight Dox treatment induces RGS4 expression leading to a marked suppression of calcium release through GAP activity on $G\alpha_q$ (**Figure 3.1 A & B**). In the screen cells were plated in 384-well plates and stimulated with an approximate EC_{20} concentration of carbachol (1 nM) in order to obtain an optimal RGS4-effect. This produced an observed Z-factor of 0.5-0.6 across three separate plates (approximately 80 –Dox and 80 +Dox wells on each plate).

In the primary screen (PubChem AID: 463165), 305,721 compounds were tested for their ability to increase carbachol-stimulated calcium transient in RGS4-expressing cells. Compounds that produce high background fluorescence during the 10 second baseline recording were removed from further analysis. Hit identification followed with selection of compounds producing a B-score [14] greater than three standard deviations above the mean for the library. This

76

	Compounds Tested	Active/ Confirmed	Percent of Tested
Primary Screen	305,721	1,365	0.45
Counter Screen: (-)Dox, (+)Carb.	1,365	1,078	79
Counter Screen: (-)Dox, (-)Carb.	1,078	467	43
CRC [a] Confirmation	467	58	12
Biochemical Confirmation	58	13	22

Table 3.1 Screening and Biochemical Confirmation Summary

[a]Concentration-response curves.

resulted in 1,365 compounds capable of enhancing carbachol-stimulated fluorescence above the threshold B-score value (**Table 3.1**). Results from representative hits are illustrated in **Figure 3.1 C**.

To exclude compounds that increase calcium flux through a mechanism unrelated to RGS4 (e.g. positive allosteric modulator of the receptor or through interactions with downstream targets in the calcium signaling pathway), the primary hits were tested in a counter-screen in the absence of Dox induction. Compounds that enhanced the carbachol-stimulated fluorescence more than five standard deviations above the mean for the negative controls were discarded at this stage (**Table 3.1**). Furthermore, in order to remove muscarinic receptor agonists, the cells were left untreated with Dox and compound alone was injected before recording data. Compounds were discarded based on the same activity criteria as before. Finally, the remaining hits were confirmed in a concentration-response format to evaluate the EC_{50} concentrations. As a result of the primary screen validation steps, a total of 58 compounds were selected on the basis of their ability to enhance carbachol-stimulated calcium signaling
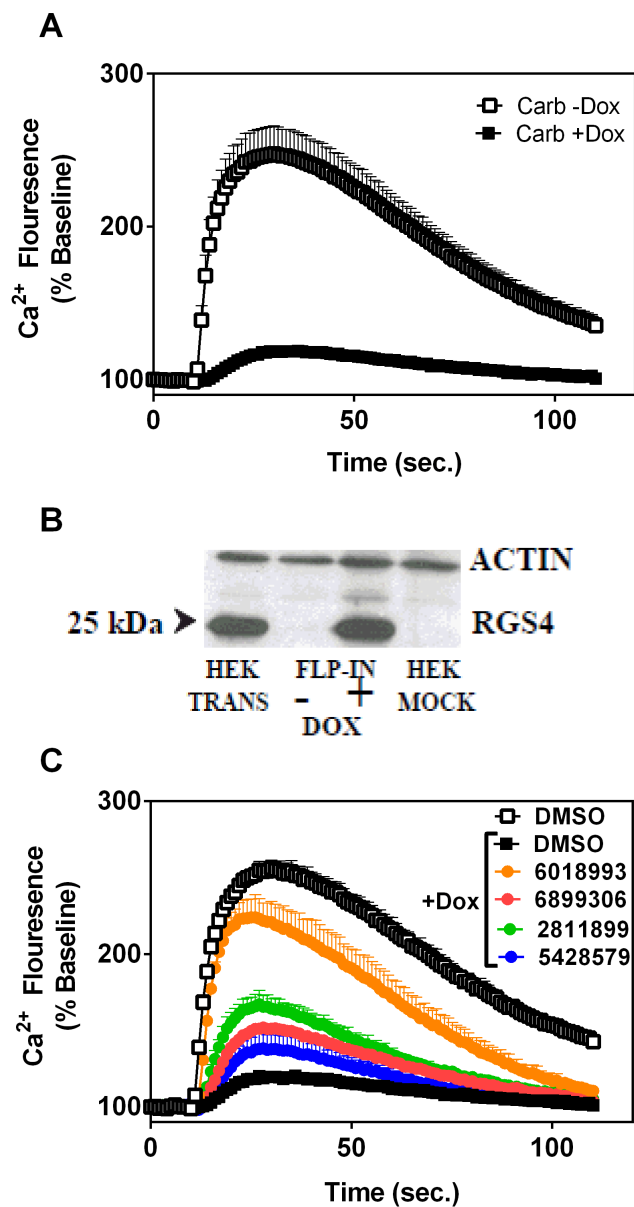
Figure 3.1 Inhibitors targeting RGS4 enhance carbachol-stimulated calcium signaling in cells treated with Doxycycline. A) HEK-FlpIn cells stably transfected with M3-muscarinic receptor produce a calcium transient in response to carbachol. Overnight treatment with Doxycycline (Dox) markedly suppresses the Gαq-mediated calcium response through induction of RGS4 protein. B) HA-tagged RGS4 is minimally detectable in HEK-FlpIn cells prior to treatment with Dox. Probing with an anti-HA antibody after a 45 hour treatment with Dox shows a marked increase in RGS4 protein, producing levels similar to transiently transfected cells (HEK Trans). C) Representative primary screening hits that increase the carbachol-stimulated calcium transient towards the –Dox control (open squares) by antagonizing RGS4-mediated suppression of Gαq signaling. The primary screening hits were tested in duplicate wells at 10 µM (final). Carb, carbachol; Dox, Doxycycline; HEK Trans, transiently transfected cells; HEK Mock, mock transfection with empty pCDNA vector.

78

specifically in Dox treated cells and in a concentration-dependent manner (**Table 3.1**).

Inhibition Of The RGS-Gα Interaction

Two assays were employed to determine whether the primary screen hits directly block the RGS-Gα interaction. The first, a multiplexed Flow Cytometry Protein Interaction Assay (FCPIA) monitors the equilibrium association between RGS-coupled Luminex beads and GDP-AlF$_4^-$-activated Gα$_o$ [17]. Each of the 58 compounds were reordered and dissolved in DMSO prior to testing in multiplexed FCPIA, which initially tested activity against MBP-tagged RGS4, RGS4 lacking cysteines (-Cys), and RGS8. The multiplexing is achieved by coupling each protein to individually identifiable bead sets and then adding the bead sets to the same well with fluorescently-labeled Gα$_o$. Thirteen compounds inhibited RGS4-Gα$_o$ binding with IC$_{50}$ values less than 100 μM, though most were less than 15 μM (**Figure 3.2 A**). Inhibition of Gα$_o$ binding to RGS4 (-Cys) or RGS8 occurred with only two of the 58 compounds tested (CID: 11957531 and 10069059, **Table 3.2**), which are structurally similar trihydroxy apomorphine derivatives (**Appendix VII**). The results of the multiplexed FCPIA suggest that, with the exception of the apomorphine compounds, all of the active inhibitors covalently modify cysteine residues on RGS4. This has been the common mechanism for all small molecule RGS4 inhibitors to date [1, 3, 6].
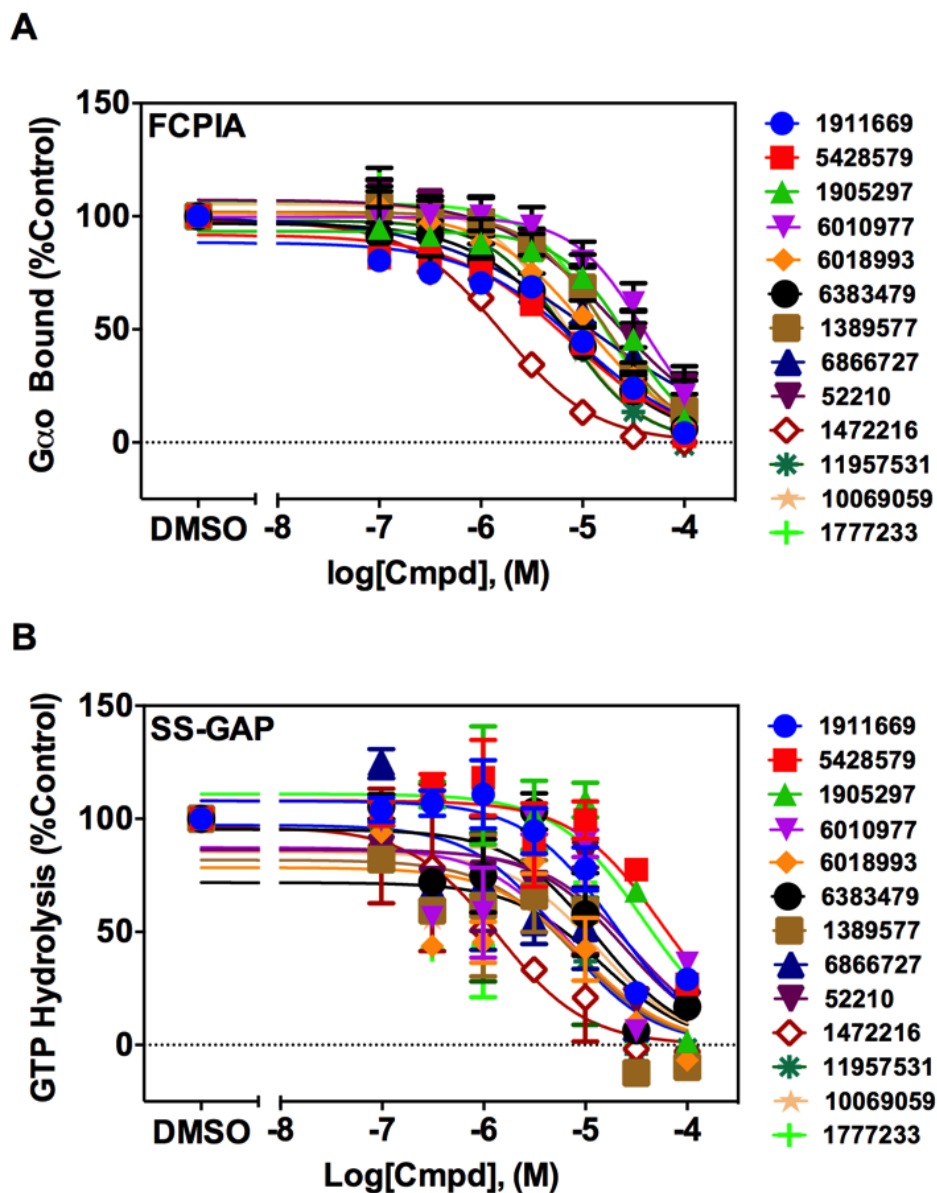
Figure 3.2 Activity summary of the compounds tested on RGS4 in FCPIA and SS-GAP. All 58 primary screen hits were assayed for an ability to block RGS4-Gαo equilibrium binding in FCPIA and inhibit RGS4-stimulation of Gαi1 nucleotide hydrolysis in SS-GAP. Compounds with an IC50 value less than 100 μM were selected for further analysis, resulting in 27 compounds meeting this criterion in SS-GAP and 13 in FCPIA. All 13 compounds active in (**A**) FCPIA were also active in (**B**) SS-GAP. The IC50 values range from approximately 2-55 μM in both assays. CID 1472216 is the most potent RGS4 inhibitor tested, IC50= 1.6±0.4 μM in SS-GAP and 1.7±0.5 μM in FCPIA. The data are the mean ± S.E.M. of three independent experiments in duplicate wells. SS-GAP, steady-state GTPase acceleration protein assay; FCPIA, flow cytometry protein interaction assay; pIC50, log IC50.

CID	RGS4		RGS4 (-Cys)		RGS7		RGS8		RGS16		RGS19	
1911669	9.9	± 3.4	NA[b]		27.1	± 7.2	NA		86.1	± 3.1	17.1	± 0.5
5428579	8.2	± 4.4	NA		10.8	± 4.1	NA		20.1	± 6.0	5.5	± 0.7
1905297	27.9	± 13.9	NA		NA		NA		NA		NA	
6010977	39.5	± 13.6	NA		NA		NA		NA		NA	
6018993	11.3	± 4.5	NA		NA		NA		NA		23.5	± 5.8
6383479	8.0	± 5.3	NA		NA		NA		NA		NA	
1389577	18.3	± 8.4	NA		29.6	± 2.1	NA		NA		19.4	± 2.1
6866727	3.4	± 2.2	NA		NA		NA		NA		NA	
52210	21.8	± 10.6	NA		NA		NA		NA		NA	
1472216	1.7	± 0.5	NA		12.7	± 4.2	NA		NA		6.7	± 1.6
11957531	6.9	± 1.1	27.6	± 8.4	7.1	± 1.1	22.4	± 4.2	7.3	± 1.1	8.1	± 1.4
10069059	7.5	± 4.0	81.4	± 22.5	23.1	± 1.4	24.8	± 13.2	16.5	± 5.1	19.6	± 4.6
1777233	15.3	± 0.6	NA		NA		NA		76.0	± 13.1	29.9	± 2.9

Table 3.2 RGS Specificity (FCPIA)[a]
[a]Flow cytometry protein interaction assay IC50 in micromolar; [b]NA, not active: compounds with IC50 greater than 100 µM, or with inhibition less than 50% at 100 µM. Reversible inhibitors are bolded.

In order to test the compounds in a system that more closely resembles the function of RGS4 inside the cell, all 58 compounds were tested in an assay that monitors RGS4 stimulation of steady-state GTP hydrolysis by Gα (SS-GAP). The reaction was performed in the absence of a receptor using a $Gα_{i1}$ mutant (R178M, A326S) capable of spontaneous nucleotide exchange [18]. Using this simplified system it is possible to rapidly evaluate large compound sets in 384-well concentration-response format using malachite green dye to quantify GTP hydrolysis [19]. In contrast to FCPIA, 27 of the 58 compounds produced IC_{50}

	Total Number of Targets				
CID	Active	Inactive	Inconclusive	Unspecified	Bioactivity[a]
1911669	69	526	31	1	13.1%
5428579	30	560	27	2	5.4%
1905297	30	542	28	4	5.5%
6010977	46	537	34	1	8.6%
6018993	87	429	28	3	20.3%
6383479	18	426	14	1	4.2%
1389577	35	542	30	1	6.5%
6866727	68	500	33	2	13.6%
52210	15	502	17	1	3.0%
1472216	185	547	26	9	33.8%
11957531	47	357	10	8	13.2%
10069059	2	11	NR[b]	NR	18.2%
1777233	25	396	22	1	6.3%

Table 3.3 PubChem Bioactivity Analysis for RGS4 Inhibitors

[a]Bioactivity is the percent active of both active and inactive biological test results in PubChem.
[b]None reported. Reversible inhibitors bolded

values less than 100 µM. All 13 compounds active in FCPIA were also active in SS-GAP (**Figure 3.2 B and Appendix VIII**).

Although there are more compounds active in SS-GAP, the majority are weak inhibitors with IC_{50} values greater than 30 µM. The more transient association of RGS4 with Gα during nucleotide turnover in SS-GAP may be more sensitive to inhibitors compared to the equilibrium association of RGS4 with a permanently activated $Gα_o$ protein in FCPIA. A second possibility is that the greater amount of bovine serum albumin required for FCPIA (1% w/v in FCPIA vs. 0.01% w/v in SS-GAP) may bind some of the compounds and reduce their effective concentrations. The 13 compounds that are active in both assays were relatively consistent in potency between the two assays. One compound in

particular stood out with high potency in both assays (CID: 1472216). This higher potency could result from being significantly more reactive than most of the compounds tested, which is consistent with the level of promiscuity reported for this compound in PubChem (active in 187 bioassays, **Table 3.3**). Further characterization focused on the thirteen compounds active in both assays.

Reversibility Of RGS Inhibitors

Since eleven out of thirteen compounds show a loss of activity against RGS4 (-Cys), it is likely the compounds work through a covalent mechanism. Reversibility of the compounds was tested using FCPIA since the compounds can be easily removed from the bead-coupled protein by pelleting and re-suspending the beads three times in fresh buffer. Less than half of the compounds showed significant reversibility, although four compounds produced greater than 85% inhibition at 100 µM, yet were more than 50% reversible within the ten minute washing period (CID: 5428579, 1905297, 1389577, & 1777233 in **Figure 3.3**). Interestingly, the two apomorphine compounds that showed activity on RGS4 (-Cys) were not reversible in this assay. The products formed through autoxidation of hydroxy apomorphine derivatives are highly reactive [20], thus inhibition of both RGS and Gα by these two compounds would explain this observation, but that has not been tested.
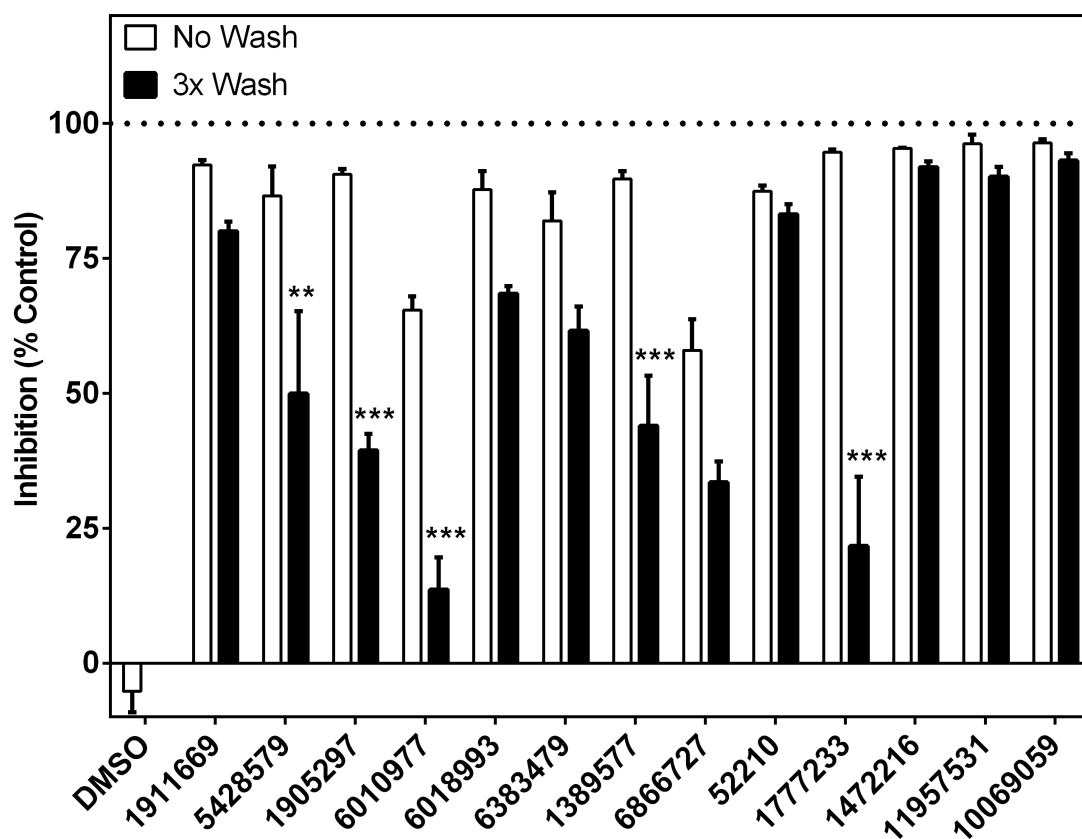
Figure 3.3 Determination of RGS4 inhibitor reversibility using FCPIA. RGS4-labeled Luminex beads were treated with 100 µM compound or DMSO for 30 minutes at room temperature before washing the beads three times. The beads were then split in two sets with DMSO or 100 µM compound added back to the no wash set (open bar). Both treated and washed RGS4-coupled beads were tested for inhibition of Gαo binding using FCPIA. Statistics: One-Way ANOVA with Bonferroni's multiple comparisons *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. Data are the result of three independent experiments ± S.E.M.

Selectivity Of Reversible RGS Inhibitors

Although the four reversible compounds do not appear to functionally modulate the G-protein since there is no effect on RGS8 or RGS4 (-Cys) binding in FCPIA, their binding to both RGS4 and Gα was assessed using a thermal stability shift assay. This assay provides a measure of direct protein-ligand

84

interactions by monitoring changes in thermal stability in the presence of compound. The melting temperature of each protein was determined with and without 100 μM compound using DMSO as control. Consistent with a mechanism of covalent modification [21], all four compounds significantly destabilize RGS4 by as much as 13 °C. In contrast, none of the compounds tested significantly alter Gα thermal melting temperature (**Figure 3.4**).

Selectivity was also determined using multiplexed FCPIA with a panel of RGS proteins (RGS7, RGS16, & RGS19). The four reversible RGS inhibitors showed variable activity, with 1905297 being relatively selective for RGS4 (**Figure 3.5**). RGS19 appears to be highly sensitive to cysteine modification as it is consistently inhibited by most of the compounds tested (**Table 3.2**). Even though RGS4 and RGS19 show nearly comparable sensitivity to most of the inhibitors, other closely related RGS4 homologs (i.e. RGS8 and RGS16) remain relatively insensitive at these concentrations. It is also notable that several compounds (including 5428579 and 1389577) showed activity towards RGS7, which lacks cysteine residues. It is unclear from these data whether the compounds acting on RGS7 bind to a defined pocket, or are reactive towards other groups besides thiols.

Inhibition Of RGS4 Activity In M3-R4 Cells

The challenge facing thiol-reactive compounds to permeate the cell membrane and overcome the reductive cytoplasmic environment has limited the

number of cell-active RGS inhibitors. As a result, this system monitoring RGS4

activity within intact cells was employed to identify novel RGS inhibitors. The
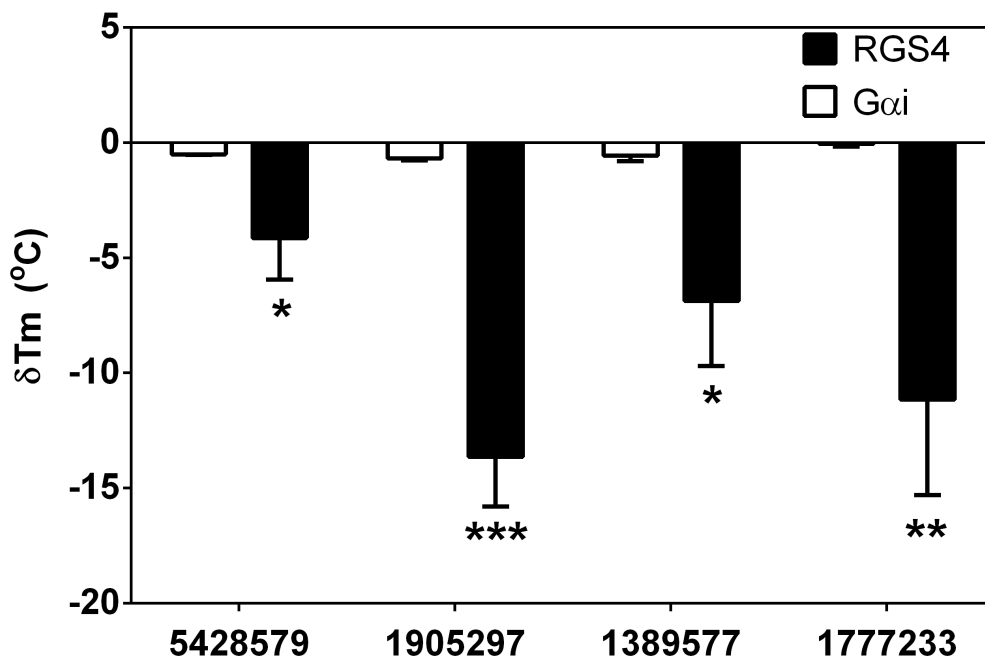
Figure 3.4 The reversible RGS4 inhibitors alter the thermal melting temperature of RGS4, but not Gα. RGS4 and Gαi1 were treated with 100μM compound or DMSO for 30 minutes prior to measuring thermal stability changes of each protein across a temperature gradient of 30-90˚C. Melting was monitored as an increase in 1,8 ANS fluorescence using a Thermofluor instrument. Data are plotted as temperature change relative to the DMSO control for each protein. None of the compounds significantly changed the thermal stability of Gαi1 compared to DMSO control. Statistics: One-Way ANOVA with Bonferroni's multiple comparisons *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. Data are the result of three independent experiments ± S.E.M.

thirteen RGS inhibitors identified in this screen antagonize RGS4 cellular activity

at multiple concentrations, and this is illustrated specifically for the four reversible

RGS inhibitors (**Figure 3.6**). The cellular EC_{50} concentrations (**Table 3.4**) are

very close to the values observed using FCPIA and SS-GAP, which suggests the

mechanism of disrupting the RGS-Gα interaction is the same inside the cells.
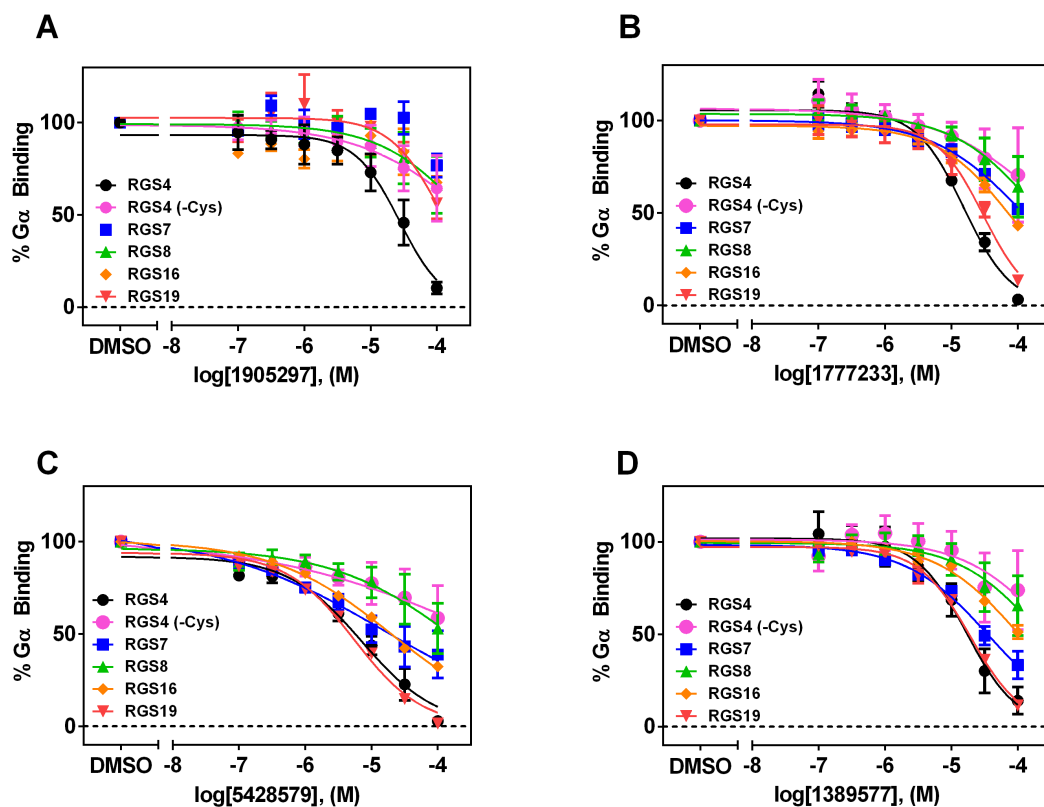
Figure 3.5 Compounds 1905297 and 1777233 show selectivity for RGS4 over four other RGS homologs tested in FCPIA. (**A**) Compound 1905297 is at least 3.6-fold selective for RGS4 over all other RGS proteins tested. (**B**) Compound 1777233 is slightly more potent towards RGS4, although there is only a 2-fold potency difference from RGS19 and 5-fold from RGS16. (**C**) & (**D**) Both 5428579 and 1389577 have approximately equal potency towards RGS4 and RGS19. In addition, activity was observed against RGS7 and RGS16 for 5428579 and RGS7 for 1389577. Compounds were tested from 100 nM to 100 µM and data are the mean of three independent experiments.
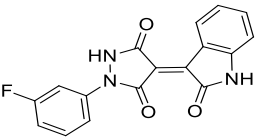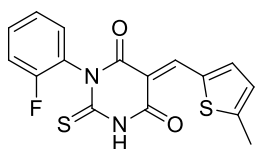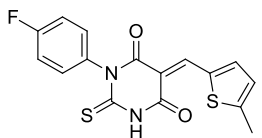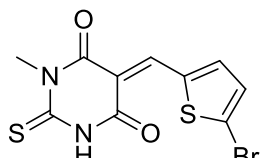
CID	FCPIA IC$_{50}$ (μM)	SS-GAP[a] IC$_{50}$ (μM)	Cellular Activity IC$_{50}$ (μM)	Specificity (FCPIA)[b]
5428579	8.2	55.8	35.5	R19 ≈ R4 > R7 > R16 > R8
1905297	27.9	35.6	23.8	R4 > (R7, R8, R16, R19)
1389577	18.3	10.9	15.0	R4 ≈ R19 > R7 > (R8, R16)
1777233	15.3	9.4	20.0	R4 > R19 > R16 > (R7, R8)

Table 3.4 Activity Summary of Reversible Compounds

[a]Steady-state GTPase Activating Protein Assay; [b]Flow cytometry protein interaction assay.
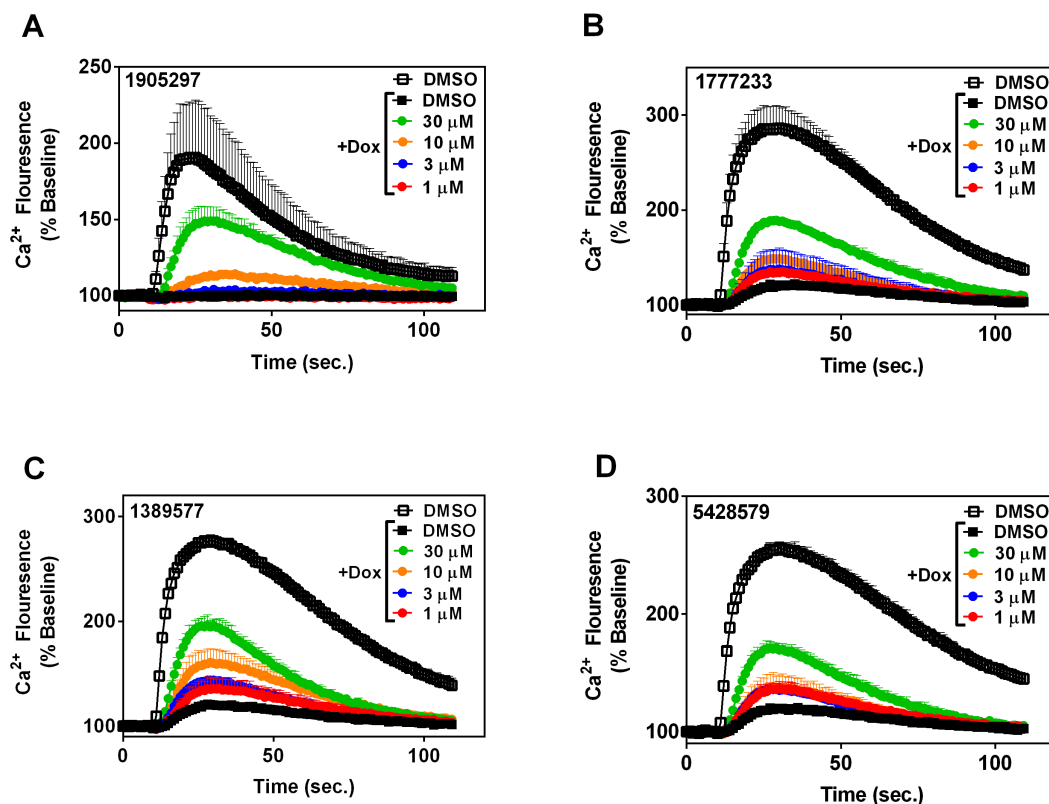
Figure 3.6 Cellular activity of the four reversible RGS4 inhibitors. (**A**) 1905297 increased the calcium response over Dox-treated control cells at 30 and 10 μM, although effects at lower concentrations may be mitigated as a result of lower overall carbachol response. (**B**)–(**D**) The carbachol response is increased over Dox-treated control cells at 1 μM compound and above. Data shown are representative responses performed in duplicate wells.

Discussion

Although therapeutics targeting GPCRs have been remarkably successful, there remains a critical need for improved therapeutic options to selectively modulate GPCR signaling. A promising new approach towards this goal is to selectively modulate mechanisms that regulate these receptors. Among these, RGS proteins are recognized as important targets for their role in negatively

regulating the amplitude and duration of GPCR output. The varying tissue expression profiles and receptor coupling efficiencies suggest RGS modulation may provide therapeutic specificity towards individual receptors, or receptor populations. This effect has been shown using genetic models of RGS "insensitivity" with knock-in mice that carry a $G\alpha_{i2}$ (G184S) mutant allele in place of the WT gene. This mutation renders $G\alpha_{i2}$ incapable of binding to any RGS protein [22]. Animals carrying this knock-in mutation show potentiated responses that are specific to 5-HT_{1A}-mediated anti-depressant behavior [23], with no significant increase in the hypothermic response that is also known to be mediated through these receptors [24].

Targeting PPIs, like the RGS-Gα interaction, is a significant challenge due to the structural topology generally associated with signal transduction proteins. In particular, the RGS-Gα interface covers a surface area of 1100 Å^2 [25] without well-defined pockets into which small molecules can be targeted to block the interaction. The absence of any non-covalent small molecules capable of acting on RGS proteins is one potential consequence of the relatively featureless structural surface. In addition, some RGS homologs show relatively high sensitivity to thiol-reactive compounds, in particular RGS4 and RGS19. As a result, thiol-reactive compounds have predominated as RGS inhibitors. The challenge of blocking RGS function within the reducing cellular environment has likely limited the number of RGS inhibitors thus far identified.

A cell-based high-throughput screen was implemented to directly identify cell-active RGS inhibitors. This assay takes advantage of the strong inhibitory

effects of RGS4 on $G\alpha_q$ signaling, which substantially suppresses the calcium fluorescence readout show in **Figure 3.1**. A critical component of the assay is the Dox-inducible RGS4 expression, which provides an efficient means of eliminating compounds targeting other parts of the calcium signaling pathway. Additional steps were taken to rule out muscarinic receptor agonists, positive allosteric modulators, or partial agonists in order to ensure calcium signaling effects were mediated through antagonism of RGS4. As a result of these triage steps, approximately half of the 58 compounds evaluated here were able to directly block the RGS-Gα interaction in a biochemical assay (SS-GAP). For reasons outlined above, there were fewer active compounds in FCPIA, but there was direct overlap between the two assays.

It remains possible the compounds found inactive in FCPIA and SS-GAP could still inhibit RGS4 effects in the cell by modulating localization to receptor or G-protein pools at the membrane. The N-terminal 33 residues of RGS4 contain an amphipathic helical domain that plays an important role in trafficking [26] and localization to both the membrane and specific receptors [27, 28]. A compound that binds to this critical region could inhibit RGS4 localization to receptor complexes at the membrane, thereby preventing or reducing association of RGS4 with $G\alpha_q$. In FCPIA and SS-GAP the proteins are freely diffusible, so compounds that only affect localization would not have activity. Furthermore, the $\Delta51$ RGS4 protein used in these assays lacks the N-terminal localization domain. In this report we have characterized the compounds that specifically block the

RGS-Gα interaction, which provide the advantage of not being limited to a specific receptor or receptor-scaffolding complex.

Like all previously reported small molecule RGS inhibitors, most compounds in this study lack activity towards RGS4 (-Cys). This suggests that the mechanism of inhibition is, at least in part, due to a covalent interaction with cysteine sidechains on the RGS proteins. The thermal destabilization of RGS4 in the presence of the four reversible inhibitors (**Figure 3.4**) is also consistent with a covalent mechanism of action, and has been observed with all previous RGS4 inhibitors [1, 3, 6]. Despite the reactive mechanism, the compounds are specific for RGS4 in the thermofluor assay as they do not significantly change the Gα melting temperature. A reaction with Gα cannot be completely ruled out as an interaction with a solvent exposed and functionally isolated cysteine sidechain is still possible without significantly altering melting temperature. The lack of an effect on the folding→unfolding transition energy of Gα strongly suggests there is no functional influence. Furthermore, the specificity for RGS is also corroborated by the selectivity for certain RGS homologs. If the compounds were able to modulate Gα function, it is unlikely specificity would be observed (**Table 3.2** and **Table 3.4**).

The biggest hurdle associated with the use of covalent inhibitors as therapeutics or even biological probes is the concern over the specificity and toxicity. As a therapeutic, irreversible compounds or metabolites sometimes form the basis of immune-related adverse drug reactions [8]. In a research setting toxicity can also obscure interpretation of biological data. These problems can

often be mitigated if the covalent compounds are reversible over a biologically relevant time scale [9, 11]. Reversibility of the RGS inhibitors was determined using FCPIA, which identified four compounds with both strong inhibition and significant reversibility over a 10 minute washing period. Additionally, these four compounds were some of the least promiscuous [29] based on biological activity reported in the PubChem project (*pubchem.ncbi.nlm.nih.gov),* which is compiled in **Table 3.3**.

Conclusions

This report details the identification and mechanistic evaluation of several new reversible, cell active RGS inhibitors. Thirteen compounds were identified that consistently inhibit the RGS-Gα interaction, although most lost activity towards an RGS4 mutant lacking cysteine residues. This is a common characteristic among all RGS4 small molecule inhibitors, and most cysteine-reactive compounds in general. Despite the apparent reactivity, many of the inhibitors show specificity for RGS4 over other RGS homologs tested. We further determined that four compounds are significantly reversible, which is an important characteristic both as a therapeutic or pharmacological tool. The significance of this study is that now a greater number of pharmacological tools are available to disseminate the roles of RGS4 in biology and disease. Recent publications have already begun using RGS inhibitors previously described, but the lack of activity in intact cells has no doubt limited the use of most RGS inhibitors currently available. Finally, as a part of the identification process we

have also described a novel cellular signaling assay that provides an efficient mechanism to rapidly evaluate large numbers of compounds for activity against RGS4 (or possibly other RGS proteins) within the context of the calcium signaling cascade.

References

1. Blazer, L.L., et al., *Reversible, allosteric, small-molecule inhibitors of RGS proteins.* Mol Pharmacol, 2010.
2. Kimple, A.J., et al., *The RGS protein inhibitor CCG-4986 is a covalent modifier of the RGS4 Galpha-interaction face.* Biochim Biophys Acta, 2007. **1774**(9): p. 1213-20.
3. Roman, D., et al., *Allosteric Inhibition of the RGS-G{alpha} Protein-Protein Interaction by CCG-4986.* Mol Pharmacol, 2010.
4. Roman, D.L., et al., *Identification of small-molecule inhibitors of RGS4 using a high-throughput flow cytometry protein interaction assay.* Mol Pharmacol, 2007. **71**(1): p. 169-75.
5. Maher, P., *The effects of stress and aging on glutathione metabolism.* Ageing Res Rev., 2005. **4**(2): p. 288-314.
6. Blazer, L.L., et al., *A nanomolar-potency small molecule inhibitor of Regulator of G protein Signaling (RGS) proteins.* Biochemistry, 2011.
7. Singh, J., et al., *The resurgence of covalent drugs.* Nat Rev Drug Discov, 2011. **10**(4): p. 307-17.
8. Park, B.K., et al., *Managing the challenge of chemically reactive metabolites in drug development.* Nat Rev Drug Discov, 2011. **10**(4): p. 292-306.
9. Lee, C.U. and T.N. Grossmann, *Reversible covalent inhibition of a protein target.* Angew Chem Int Ed Engl, 2012. **51**(35): p. 8699-700.
10. Smith, A.J., et al., *Beyond picomolar affinities: quantitative aspects of noncovalent and covalent binding of drugs to proteins.* J Med Chem, 2009. **52**(2): p. 225-33.
11. Tummino, P.J. and R.A. Copeland, *Residence time of receptor-ligand complexes and its effect on biological function.* Biochemistry, 2008. **47**(20): p. 5481-92.
12. Storaska, A.J., et al., *Reversible inhibitors of regulators of G-protein signaling identified in a high-throughput cell-based calcium signaling assay.* Cell Signal., 2013. **25**(12): p. 2848-55. doi: 10.1016/j.cellsig.2013.09.007. Epub 2013 Sep 14.
13. Bodenstein, J., R.K. Sunahara, and R.R. Neubig, *N-terminal residues control proteasomal degradation of RGS2, RGS4, and RGS5 in human embryonic kidney 293 cells.* Mol Pharmacol, 2007. **71**(4): p. 1040-50.
14. Brideau, C., et al., *Improved statistical methods for hit selection in high-throughput screening.* J Biomol Screen, 2003. **8**(6): p. 634-47.
15. Soundararajan, M., et al., *Structural diversity in the RGS domain and its interaction with heterotrimeric G protein alpha-subunits.* Proc Natl Acad Sci U S A, 2008. **105**(17): p. 6457-62.
16. Lee, E., M.E. Linder, and A.G. Gilman, *Expression of G-protein alpha subunits in Escherichia coli.* Methods Enzymol, 1994. **237**: p. 146-64.
17. Blazer, L.L., et al., *Use of flow cytometric methods to quantify protein-protein interactions.* Curr Protoc Cytom, 2010. **Chapter 13**: p. Unit 13 11 1-15.
18. Zielinski, T., et al., *Two Galpha(i1) rate-modifying mutations act in concert to allow receptor-independent, steady-state measurements of RGS protein activity.* J Biomol Screen, 2009. **14**(10): p. 1195-206.

19.   Chang, L., et al., *High-throughput screen for small molecules that modulate the ATPase activity of the molecular chaperone DnaK.* Anal Biochem, 2008. **372**(2): p. 167-76.

20.   El-Bacha, R.S., P. Netter, and A. Minn, *Mechanisms of apomorphine cytoxicity towards rat glioma C6 cells: protection by bovine serum albumin and formation of apomorphine-protein conjugates.* Neurosci Lett, 1999. **263**(1): p. 25-8.

21.   Cimmperman, P., et al., *A quantitative model of thermal stabilization and destabilization of proteins by ligands.* Biophys J, 2008. **95**(7): p. 3222-31.

22.   Fu, Y., et al., *RGS-insensitive G-protein mutations to study the role of endogenous RGS proteins.* Methods Enzymol, 2004. **389**: p. 229-43.

23.   Talbot, J.N., et al., *RGS inhibition at G(alpha)i2 selectively potentiates 5-HT1A-mediated antidepressant effects.* Proc Natl Acad Sci U S A, 2010. **107**(24): p. 11086-91.

24.   Richardson-Jones, J.W., et al., *5-HT1A autoreceptor levels determine vulnerability to stress and response to antidepressants.* Neuron, 2010. **65**(1): p. 40-52.

25.   Tesmer, J.J., et al., *Structure of RGS4 bound to AlF4--activated G(i alpha1): stabilization of the transition state for GTP hydrolysis.* Cell, 1997. **89**(2): p. 251-61.

26.   Bastin, G., et al., *Amino-terminal cysteine residues differentially influence RGS4 protein plasma membrane targeting, intracellular trafficking, and function.* J Biol Chem, 2012. **287**(34): p. 28966-74.

27.   Srinivasa, S.P., et al., *Plasma membrane localization is required for RGS4 function in Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 1998. **95**(10): p. 5584-9.

28.   Zeng, W., et al., *The N-terminal domain of RGS4 confers receptor-selective inhibition of G protein signaling.* J Biol Chem, 1998. **273**(52): p. 34687-90.

29.   McGovern, S.L., et al., *A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening.* J Med Chem, 2002. **45**(8): p. 1712-22.

Chapter IV

**Mapping The Interaction of Inositol Hexakisphosphate
on RGS4**

Introduction

The previous chapters contribute a greater understanding of the molecular

mechanism of cysteine-directed inhibitors of RGS proteins. In this chapter I

explore a non-covalent interaction on RGS4 based on a physiologic allosteric

mechanism. Targeting allosteric sites on RGS4 may provide several advantages

towards the goal of identifying non-covalent small molecule RGS modulators. In

particular, the largest surface pocket exists on a surface that is proposed to be

involved in allosteric regulation of RGS4, termed the B-site (see **Figure 1.3**) [1].

The size of the pocket is important for accommodating a compound with a

sufficient number of specific contacts to provide high affinity. As discussed in

Chapter I, the B-site is the putative binding site for physiological regulators of

RGS4, including Ca^{2+}/Calmodulin and PIP_3 [2, 3]. Inhibition of RGS4 GAP activity

occurs in the presence of PIP_3 *in vitro* [3-5]. In light of this, mapping the

molecular details of acidic phospholipid interactions with RGS4 will elucidate the

exact allosteric sites of interaction. This would provide novel insight into the most

appropriate physical areas for structure-based design strategies for new

compounds. Additionally, uncovering these molecular details may also

provide a better understanding of how non-covalent compounds inhibit GAP activity.

In order to understand the interactions of RGS4 with acidic phospholipids, I employed a highly soluble inositol phosphate derivative (inositol hexakisphosphate, IP_6) (see **Figure 4.1**). This compound contains similar poly-phosphate charge density as PIP_3 across one side of the myo-inositol ring, and has been used in biophysical studies of acidic phospholipid interactions for other proteins [6, 7]. IP_6 itself is present in animal cells at concentrations up to 0.1 millimolar [8], and has been shown to be a regulator of cell cycle progression and neurotransmitter release *in vitro* [8, 9].

The work in this chapter presents direct evidence of inositol phosphate binding to RGS4 using solution NMR. The binding interaction of IP_6 on RGS4 is mapped through the chemical shift perturbations produced in the NMR-HSQC titration. The structural effects of IP_6 on RGS4 are also assessed using Thermofluor experiments. This study identifies residues involved in the interaction. Characterization at this resolution using PIP_3 has not been possible due to the difficulty of using micelles in biophysical measurements. As a result, this study provides novel insight into the sites on RGS4 interacting with endogenous regulators, which defines sites for future structure-based studies to identify novel small molecule modulators.

I performed all of the work described in this chapter, and this chapter is currently in preparation to be submitted for publication.

Materials and Methods

All reagents are obtained through Fisher Scientific (Pittsburgh, PA) or Sigma (St. Louis, MO), unless otherwise specified. $\gamma[32P]$GTP (10 mCi/mL) was obtained from Perkin Elmer Life and Analytical Sciences, (Boston, MA) and was isotopically diluted before use. Amylose resin was purchased from New England Biolabs (Ipswich, MA). Phytic acid (IP_6) sodium salt hydrate was purchased from Sigma-Aldrich. $[^{15}N]$-$(NH_4)_2SO_4$ was purchased from Cambridge Isotope Laboratories (Andover, MA).

Protein Expression And Purification

$\Delta51$ RGS4 and ^{15}N-$\Delta51$ RGS4 were expressed in E. coli as a fusion construct with an N-terminal maltose-binding protein (MBP) containing a Tabacco Etch Virus (TEV) protease cleavage site and a 10x poly-Histidine tag, using a pMalC2H10 vector. The details of expression and purification were followed exactly as described previously [10, 11]. The resulting RGS4 protein is nearly identical in sequence to the construct previously used for NMR spectral assignment and solution structure determination [12, 13].

NMR Spectroscopy

^{1}H-^{15}N-HSQC experiments were used to monitor IP_6-induced chemical shift perturbations (CSP) of ^{15}N-RGS4 ($\Delta51$) peaks in 50 mM MES (2-(N-morpholino)ethanesulfonic acid) buffer at pH 6.0 and 50 mM NaCl. Experiments were conducted at 25°C using a Bruker Avance III 600 MHz spectrometer with a

cryogenically cooled 3 mm sample probe. IP$_6$ was prepared as a 10X stock in the NMR buffer and adjusted to pH 6.0. NMR samples were prepared with 0.1 mM ^{15}N-labeled RGS4 in the above buffer plus 7% (v/v) D$_2$O and increasing concentrations of IP$_6$. All RGS4-IP$_6$ HSQC spectra were compared to a reference spectrum of a sample containing RGS4 alone. Chemical shift perturbations were calculated using the following equation:

$$CSP = \sqrt{\left(\frac{\Delta\delta\,^{15}N}{5}\right)^2 + \left(\Delta\delta\,^1H\right)^2}$$

Thermal Stability Shift Assay

RGS4 (Δ51) at a concentration of 10 μM final was mixed in a buffer containing 50 mM HEPES, pH 7.4 and 100 mM NaCl. RGS4 was incubated in a black 384-well PCR microtiter plate (ThermoFisher catalog no. TF-0384/K) and in the presence of increasing concentration of IP$_6$ (0.1 mM to 10 mM) for 20 minutes at room temperature, or buffer as control. 1-anilinonaphthalene-8-sulfonic acid (ANS) was added to each well at a final concentration of 100 μM, and all wells were overlaid with 5 μl silicon oil before measuring changes in thermostability. Fluorescence image collection was taken at 25°C for each point between denaturation steps, which ramped up 1° between 30°-90°C using a ThermoFluor Instrument (Johnson & Johnson, Langhorne, PA). Melting temperature was calculated using the Thermofluor+ software and plotted using Graphpad Prism 4.0.

[32]P Steady-State GTPase Activating Protein Assay (SS-GAP)

Hydrolysis of $[\gamma\text{-}^{32}P]$-GTP (Perkin Elmer) was measured under receptor-independent steady-state conditions by employing a mutant Gαi1 (R178M, A326S) [1]. The mutations produce spontaneous release of GDP, making GTP hydrolysis the rate-limiting step [1]. The steady-state GAP assay was performed in 96-well PCR plates in reaction buffer comprising 50 mM Hepes, 100 mM NaCl, 0.01% Lubrol, 5 mM $MgCl_2$ at pH 7.4. IP_6 was added to RGS4 at concentrations ranging from 0.5 mM to 10 mM (final). The reaction was initiated for each row of the plate at one minute intervals with a 4x solution of reaction buffer containing 32 μM GTP and 4 nM $[\gamma\text{-}32P]$-GTP to act as a tracer, resulting in a final volume of 40 μl. The reaction was stopped after 100 minutes by quenching 30 μl of the reaction (one minute /row) into 300 μl of ice-cold 20 mM sodium phosphate, pH 2.0 containing a 5% w/v slurry of activated charcoal in a 96-deepwell 1000 μl plate (Eppendorff). The plate was then centrifuged for 15 minutes at 1100 x g at a temperature of 4 °C, so 50 μl of the charcoal-free supernatant (containing only free inorganic phosphate) could be transferred to a Perkin Elmer 96-well OptiPlate. To each well, 200 μl of Microscint 40 fluid was added and allowed to thoroughly mix overnight at room temperature. The radiolabel activity (counts per minute) from $[^{32}P]$-inorganic phosphate was recorded using a Packard TopCount NXT plate reader. RGS4 GAP activity was determined by as the number of picomoles of GTP hydrolyzed in 100 minutes. The data were analyzed using Graphpad Prism v5.01.

In Silico Docking Of IP$_6$ Onto The RGS4 Structure

The PDB file of RGS4 (code: 1agr) was prepared for docking using the Chimera.33 DockPrep tool (UCSF) [14]. This tool was used to add hydrogen atoms and remove water molecules from the structure. Partial charges on the structure were assigned using AMBER99 force field. The three-dimensional structure of IP$_6$ was extracted as a .mol file from PDB 2xam. Automated flexible ligand docking was performed using a web-based docking server: SwissDock (http://www.swissdock.ch/), which is based on EADock DSS.34 [15]. The coordinate files of RGS4 and IP$_6$ were docked under the "Accurate" setting, which provides the most exhaustive sampling of binding modes. The docking was unbiased in the sense that IP$_6$ was allowed to explore the entire surface of RGS4 in the simulation. The output clusters were ranked according to binding free energy.

Results

RGS4 functions as a regulator of calcium signaling through GAP-mediated inhibition of Gα_q activity [16]. *In vitro* experiments show that an increased fraction of PIP$_3$ in PC:PIP$_3$ micelles is able to antagonize GAP activity towards Gα_i [3], and is expected to also inhibit RGS4 activity towards Gα_q. This action of PIP$_3$ is proposed to limit the activity of RGS4 in a low-calcium basal state. Activation of intracellular calcium release is suggested to reverse PIP$_3$ inhibition of RGS4 by activating calmodulin, leading to negative feedback inhibition of calcium signaling [17]. Interestingly, lipid co-sedimentation assays show that RGS4 binding is
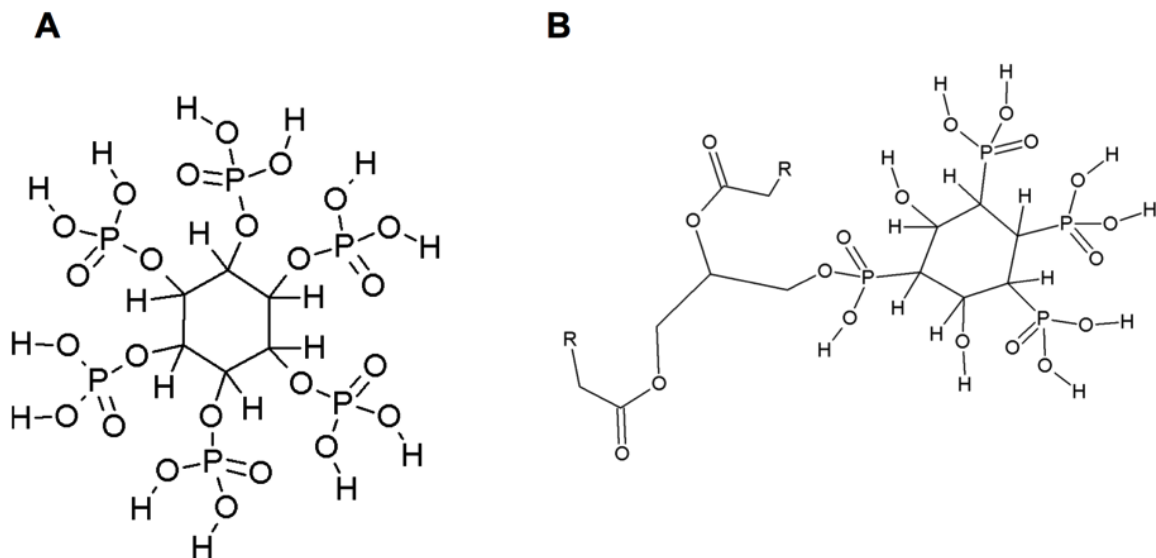
102

Figure 4.1 Structural comparison of IP6 and PIP3. **A**) IP_6 is a close structural analog to the hydrophilic groups found on acid phospholipids (e.g. PIP_3). **B**) The hydrophilic portion of PIP_3 contains phosphate groups situated at positions 3, 4, and 5 on the six membered ring, which results in one side of the molecule containing similar anionic charge density as IP6.

specific to PIP_3 and not other phospholipid derivatives(e.g. $PI(3,4)P_2$ and $PI(4,5)P_2$) [2]. This suggests that the charge density resulting from phosphates at positions 3, 4, and 5, as in PIP_3, is critical for the interaction. As a result, IP_6 should be able to interact with RGS4 in a manner similar to that of PIP_3, given the similarity in structure and charge density (**Figure 4.1**).

IP_6 Binding Increases RGS4 Thermal Stability at High Concentrations

Since small molecules can influence the thermal stability of a protein through direct binding interactions [18, 19], a Thermofluor assay was used to determine whether IP_6 binds to RGS4. Previous studies show that PIP_3 is capable of interacting with a truncated version of RGS4 containing the RGS homology domain plus the C-terminus (residues 52-205) [2, 3]. This construct
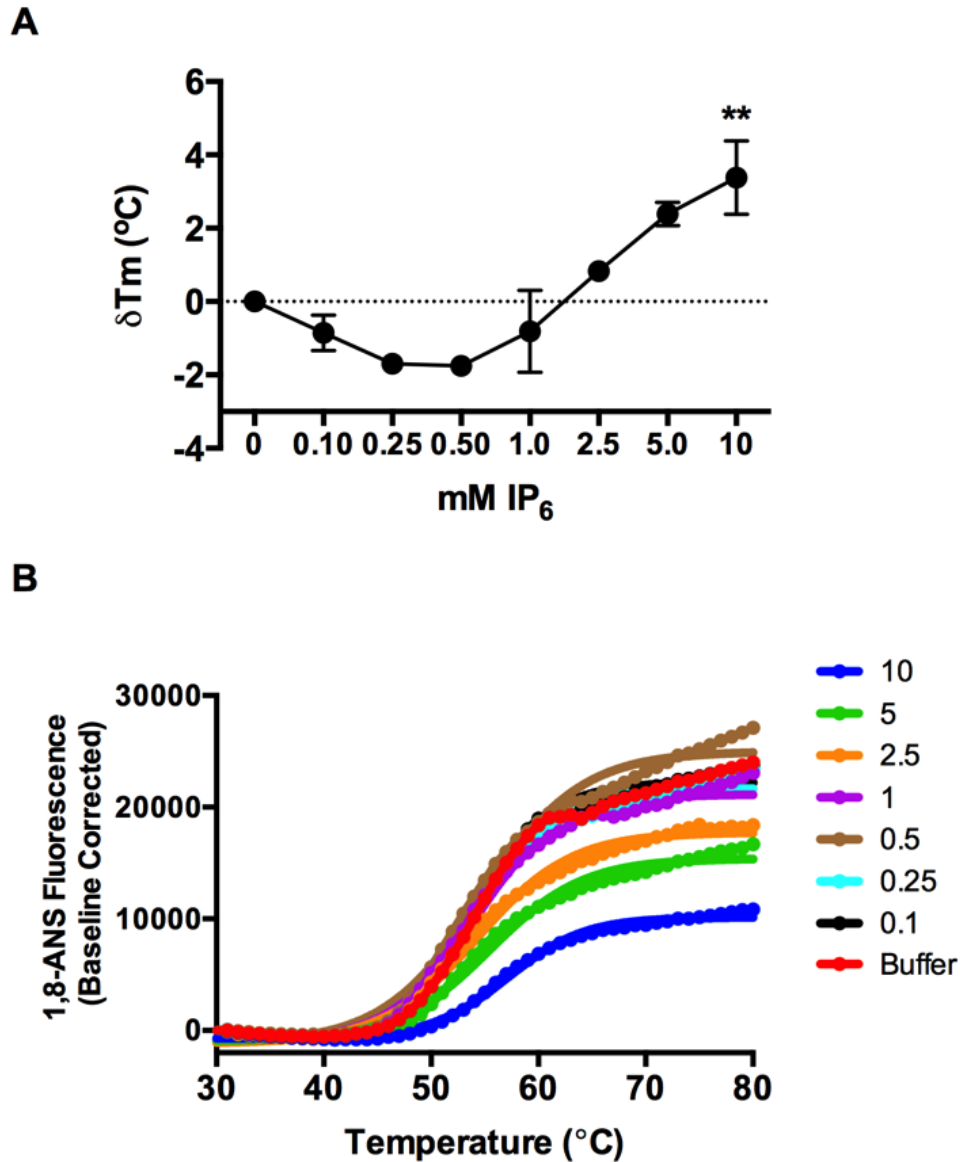
Figure 4.2. Thermal stability changes of RGS4 as a function of IP6 concentration. **A**) Changes to RGS4 thermal stability in response to IP_6. Low IP_6 concentrations down to 0.5 mM show a downward trend in RGS4 thermal stability by as much as -1.7 °C, while RGS4 is stabilized at concentrations of 2.5 mM and above, to a maximum of 3.4 °C at 10 mM**. B**) Representative thermal melting curves from a single experiment showing changes in the melting temperature (Tm) of RGS4 in control buffer (Red) compared to IP_6-treated RGS4. Statistics: One-Way ANOVA with Bonferroni's multiple comparisons **, $P < 0.01$. Data are the result of three independent experiments ± S.E.M.

was used for all of the biophysical studies due to better expression and solubility [11]. At the lowest concentrations of IP_6 tested (1 mM and below) there is a trend towards a decrease in thermal stability of RGS4 by as much as -1.7 °C, although these points are not statistically significant compared to the buffer control in the ANOVA analysis. In contrast, higher concentrations of IP_6 increase the thermal stability of RGS4 by as much as 3.4 °C at 10 mM ($p < 0.01$; **Figure 4.2**). The possibility of IP_6 affecting the fluorescence of the 1,8 ANS dye was ruled out by concurrently running the experiment with ANS and IP_6 alone. The significant increase in thermal stability observed at 10 mM shows that IP_6 is interacting with RGS4.

Mapping IP_6-induced Chemical Shift Perturbations

^{1}H-^{15}N-HSQC NMR experiments were used to assess changes to the chemical environment of backbone amide cross peaks of RGS4 that result from interactions with IP_6. Every residue in RGS4, with the exception of proline, provides a cross-peak in the HSQC spectrum. The specific binding sites of ligands on a protein can be determined from the changes in the chemical shift of the HSQC cross peaks when spectra in the presence and absence of the ligand binding partner are compared in an overlay [24]. The chemical shift perturbations (CSPs) of RGS4 in the presence (red) and absence (black) of IP_6 are shown for saturating concentrations of IP_6 (**Figure 4.3)**. In this experiment, side chain interactions with IP_6 are not directly detectable, so only resonances from
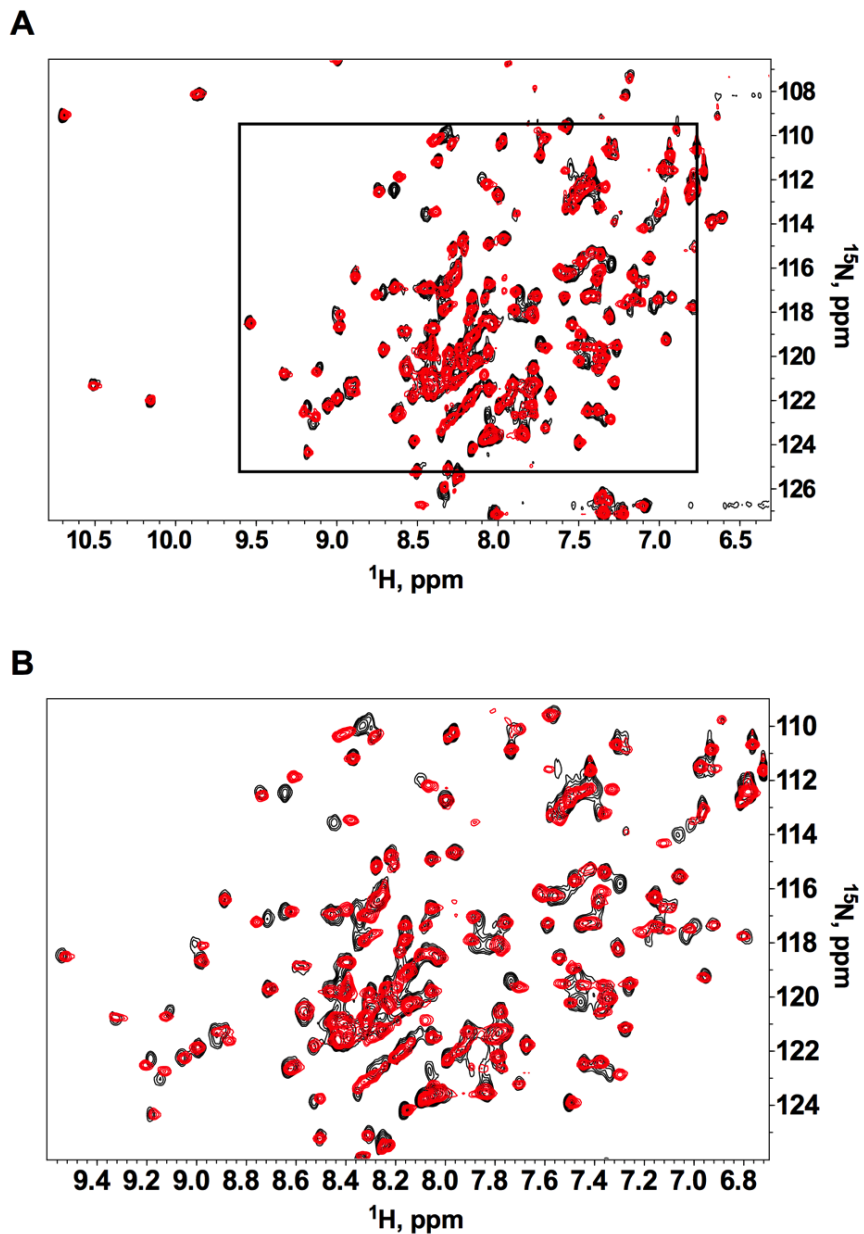
Figure 4.3. ^{1}H-^{15}N HSQC overlay of ^{15}N-labeled RGS4 (0.1 mM) in the presence and absence of IP$_6$. **A**) Spectra of ^{15}N-RGS4 (residues 52-205) were collected at pH 6.0 at 25°C, in the presence (Red) and absence (Black) of IP$_6$ at a concentration of 10 mM. The solution containing IP$_6$ was buffer matched to the reference sample. **B**) Zoomed view of the boxed region in panel A.

backbone amides are being detected and analyzed in this plot. The presence of

peak shifts across the spectrum validates that IP$_6$ is binding to RGS4.

In order to determine the residues involved in the interaction, RGS4 was titrated with IP_6 using concentrations from 0.25 mM to 10 mM. Changes in chemical shift for each cross peak were followed as a function of IP_6 concentration, which allowed accurate determination of the final position of each cross peak at saturation. A plot of CSPs as a function of RGS4 residue number shows that residues that undergo large CSPs upon IP_6 binding cluster to four areas (**Figure 4.4 A**). These large CSPs, denoted as greater than one standard deviation above the average CSP of all residues (solid line, **Figure 4.4 A**), occur along the C-terminal portion of α-helix 4 and the N-terminal portion of α-helix 5. Two other clusters of perturbations are observed along the loop connecting α-helix 6 & 7, and one C-terminal residue in the unstructured portion of RGS4.

The CSPs are mapped onto the backbone structure of RGS4 (red > average + 1σ, solid line; average + 1σ > orange > average, dotted line) (**Figure 4.4 B**). One cluster of highly perturbed residues occurs along parts of α-helices 4 & 5, with the most perturbed residues centered around Lys110. This creates a hot spot of CSPs ranging from Ser103 to Lys113. Perturbations also appear on the portion of α-helix 4 that faces helix 5, including residues Ser94 and Glu97. A second cluster of CSPs occurs on the C-terminus of α-helix 4, particularly residues Tyr98 and Lys 99/100. Part of this cluster is formed through adjacent residues in the flexible loop connecting α-helix 6 & 7, involving residues Thr145 and Ile146. It is unclear why perturbations are observed in the C-terminal unstructured portion of RGS4, but only one residue (Gly190) shows a strong perturbation suggesting this is not a binding site for IP_6.
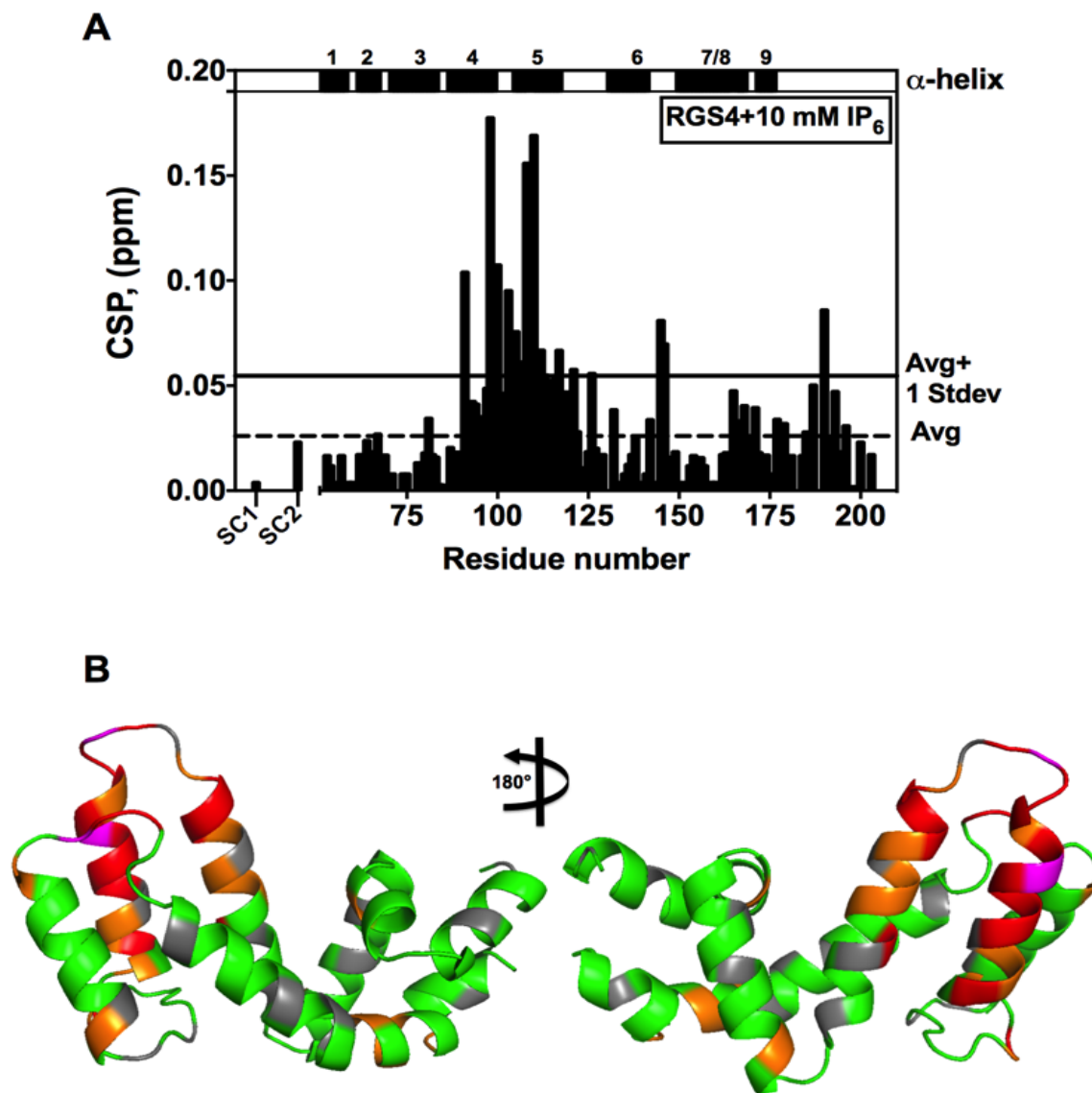
Figure 4.4 Chemical shift perturbations of ^{15}N-RGS4 HSQC spectra in the presence of IP$_6$. **A**) Chemical shift perturbations (CSPs) are plotted as a function of RGS4 residue number (SC1 and SC2 refer to indole N-H resonances from the two tryptophan side chains). CSP calculation was determined based upon the equation provided in materials and methods. CSPs strongly cluster to two sites on the RGS4 backbone, including a cluster at the C-terminal portion of α-helix 4 and N-term of α-helix 7, and a cluster on α-helix 5. **B**) CSPs are highlighted on the RGS4 backbone (red > average + 1σ, solid line; average + 1σ > orange > average, dotted line; grey are overlapping residues, magenta indicates proline residues). Structures are from PDB: 1agr.

Quantification Of IP$_6$ Binding Affinity

Binding of IP$_6$ to RGS4 appears to be in fast exchange relative to the NMR time-scale based on the linear trajectory of CSPs as a function of IP$_6$

concentration. All of the CSPs observed in the presence of IP$_6$ are fully saturated at 10 mM. The broad range of CSPs across α-helices 4, 5, and parts of α-helix 7 suggest the possibility of two binding sites for IP$_6$ on RGS4. The N-terminal portion of α-helices 5 and parts of α-helix 4 that face α-helix 5 (e.g. residues Ser94 and Glu97), appear to form one binding site since large CSPs are clustered broadly from Ser103 to Lys113 (*vide supra*). The second binding site might involve an adjacent structural region based on a set of CSPs that are clustered around Lys99/100, as well as a loop region connecting α-helices 6 and 7.

To determine whether IP$_6$ binds to two distinct regions of RGS4, saturation curves were fit using residues from the two putative binding sites. Disassociation constants were determined by fitting IP$_6$ concentration-dependent CSPs from residues at α-helix 5 as one independent site (**Figure 4.5 A**) and α-helix 4 and 7 for the second putative site (**Figure 4.5 B**). The result from fitting the binding isotherms shows similar affinities of IP$_6$ for the two sites, K_d = 0.35 ± 0.27 mM for α-helix 5 CSPs and 0.59 ± 0.34 mM for α-helix 4 and 7. The overlapping affinities that are observed from fitting saturation curves for each cluster of CSPs are consistent with a single binding site that produces this broad range of perturbations. Using a global fitting procedure for all ten residues, based upon a 1:1 binding interaction, results in an affinity of 0.56 ± 0.06 mM. Thus, the simplest
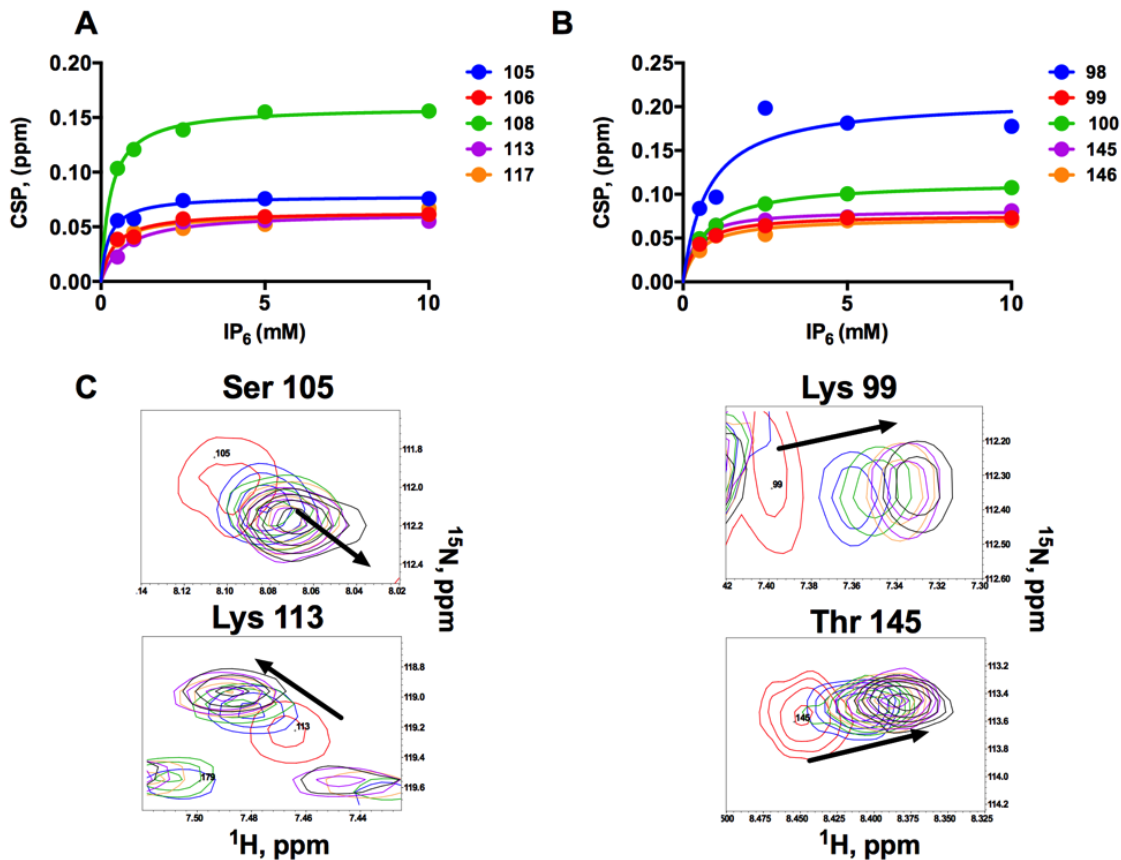
Figure 4.5 Saturation binding curves of residues from each cluster of CSPs. **A**) CSPs of residues clustered in the α-helix 5 IP6 binding site are plotted as a function of IP6 concentration showing a Kd = 0.35 ± 0.27 mM, compared to **B**) α-helix 4/7 cluster Kd = 0.59 ± 0.34 mM. C) Representative CSPs for residues in each cluster, which show a saturatable chemical shift change from buffer control (red) to 10 mM IP6 (black). Binding isotherms were each fit to a one site specific binding curve with global fitting.

model is one in which binding of IP_6 produces allosteric structural changes through one binding site. It seems most likely that binding occurs along α-helix 5 where the largest cluster of CSPs is observed (Ser103 to Lys113), although this cannot be definitively concluded based upon the NMR analysis alone.
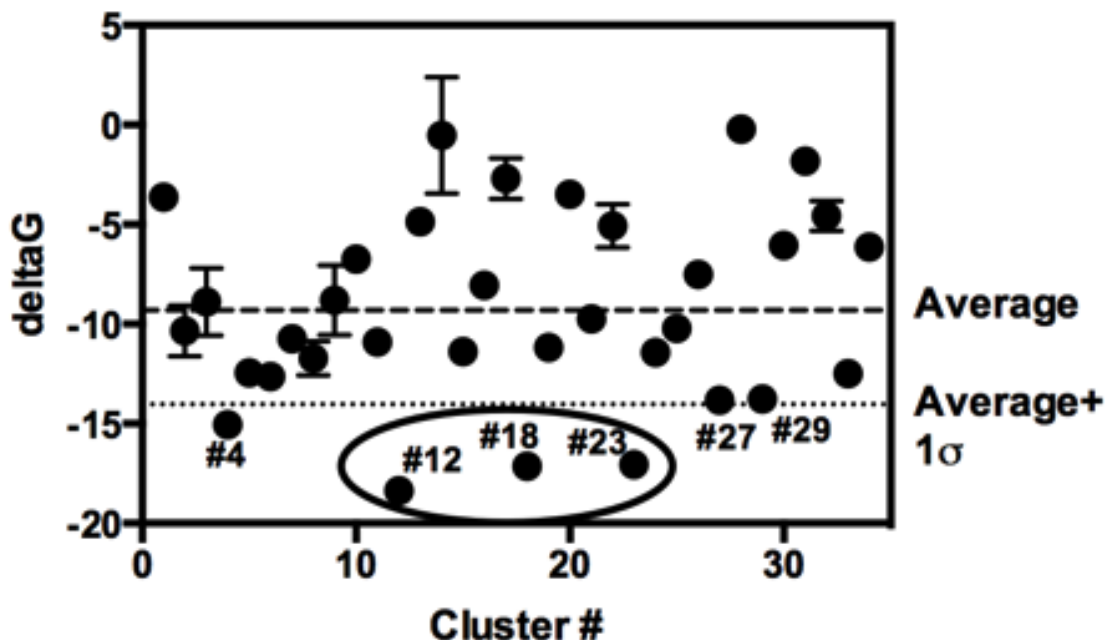
Figure 4.6 Docking statistics for the various IP_6 poses as a function of cluster number. The lowest energy docking poses of IP_6 are comprised of three clusters that dock consistently to the α-helix 5 binding site (circled). Clusters 4, 27, and 29 correspond to poses that dock to an area of RGS4 centered around Lys99/100, although these poses are not predicted to be as energetically favorable as 12, 18, and 23.

Blind Docking Of IP_6 On RGS4

In silico docking was performed using the SwissDock automated webserver [25] in order to predict the most energetically favorable binding site of IP_6 on RGS4. Binding modes were generated using a blind docking procedure that allowed IP_6 to explore the entire surface of RGS4. The binding energies of the poses are estimated using the CHARMM algorithm, and then each pose is evaluated using a generalized Born implicit solvent model (FACTS) [26] before clustering. The docking run resulted in 34 total clusters (**Figure 4.6**), where a cluster is defined by a group of IP_6 molecules that bind to a single site on RGS4 and differ only by rotational changes around the bonds. Clusters are limited to

111
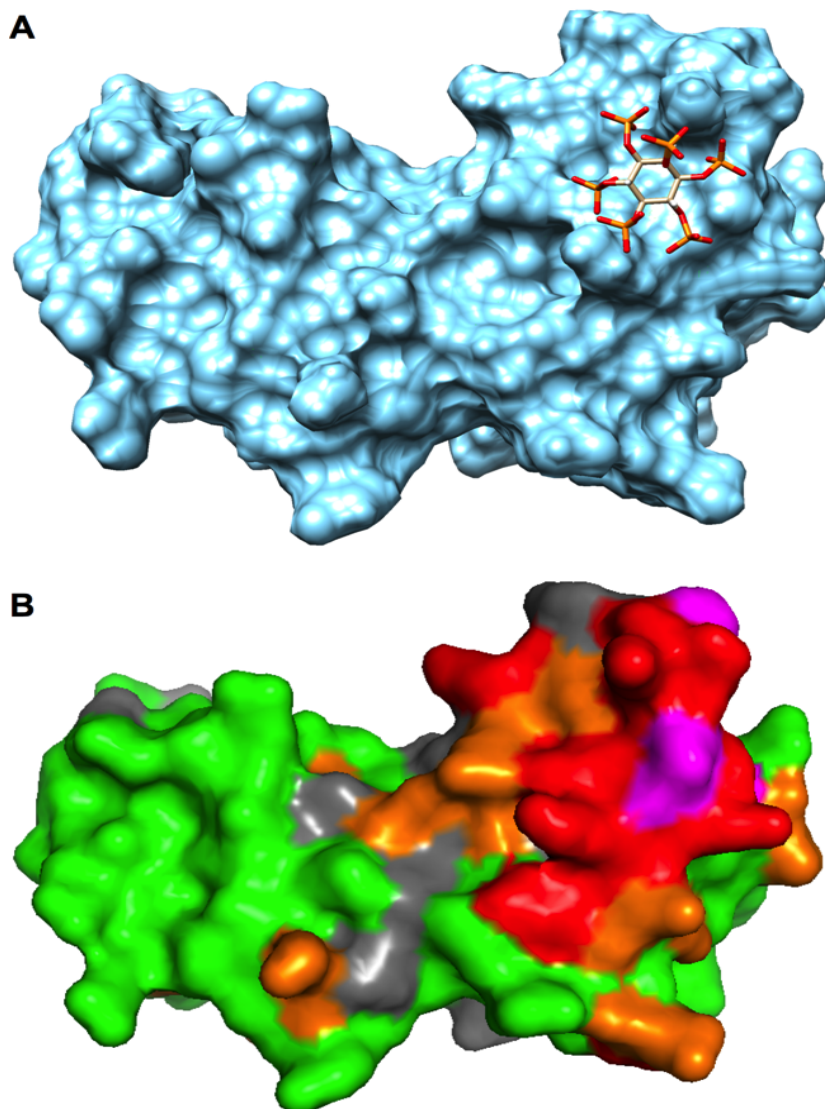
Figure 4.7 Docking model of IP$_6$ across α-helix 5 on RGS4. **A**) Representative pose from cluster 12. This site overlaps a previously reported docking site for the CCG-50014 RGS4 inhibitor, resulting in the first experimental evidence of the allosteric surface pocket. Blind docking of the IP$_6$ molecule on RGS4 was performed using the Swiss dock web server. **B**) CSPs are highlighted on the RGS4 backbone (red > average + 1σ; average + 1σ > orange > average; grey are overlapping residues, magenta indicates proline residues). The most robust CSPs (red) directly overlap with the modeled IP$_6$ binding site.

series of eight rotamers by the server, resulting in several clusters docking to a single site.

The three clusters (clusters: 12, 18, and 23) with the lowest predicted binding free energy (ΔG) all docked to the α-helix 5 site (**Figure 4.7 A**) that

directly overlaps with the largest cluster of chemical shift perturbations that are mapped on the RGS4 structure (**Figure 4.7 B**). The representative pose illustrated in **Figure 4.7 A** corresponds to cluster 12 in **Figure 4.6**. As observed in the docking model, and consistent with the NMR data, IP$_6$ docks in a small crevice formed between Lys110 and Lys 112/113 on α-helix 5. This model would also explain the perturbations observed on α-helix 4 that directly face helix 5, as the IP$_6$ molecule extends over the structural density formed by both helices. Structural perturbations that result from an interaction at this site could translate indirect structural changes on the C-terminal portion of α-helix 4, surrounding residues Lys99/100. Structural changes occurring around the Lys99/100 side chains might also drive changes to the chemical environment of residues Thr145 and Ile146, where perturbations are observed within the loop connecting α-helix 6 and 7. Interestingly, this pocket formed by Lys110 and Lys 112/113 in the protein overlaps with a site on RGS8 that has been proposed as the initial docking site for CCG-50014 [22], which is adjacent to Cys95.

The IP$_6$ molecule also docked to a site involving α-helix 4 and 7 (**Figure 4.8 A**), which corresponds to clusters 4, 27, and 29. The docked model similarly shows overlap with CSPs observed in the HSQC experiments, (**Figure 4.8 B**) with primary sites of contact on Lys99/100, Thr145 and Ile146 across α-helices 4 and 7. However, these poses represented by cluster four (**Figure 4.8 A**) are not as favorable as those docking to the α-helix 5 site, and only cluster four has a predicted binding free energy that is below the dotted line representing one standard deviation beyond the mean docking energies (**Figure 4.6)**. Therefore,
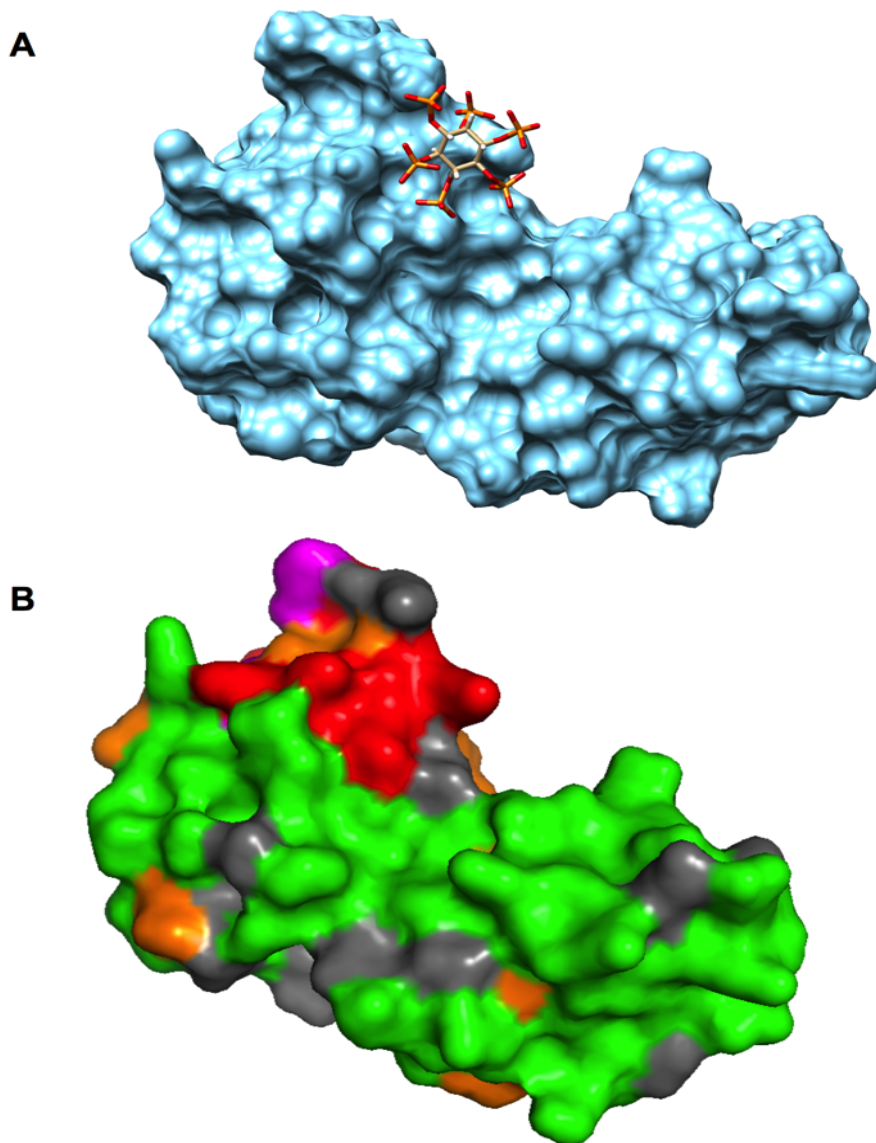
Figure 4.8 Docking model of IP$_6$ at α-helices 4 and 7 on RGS4. **A**) Representative pose from cluster four of the same blind docking run as in figure 4.7. This site overlaps with the putative PIP$_3$/calmodulin binding sites from past reports. **B**) Consistently, the docking model binding site overlaps with CSPs observed in the NMR titration of RGS4 with IP$_6$.

based upon the predicted docking energies and the extent of CSPs, it is not possible to establish the existence of more than one binding site from these data.
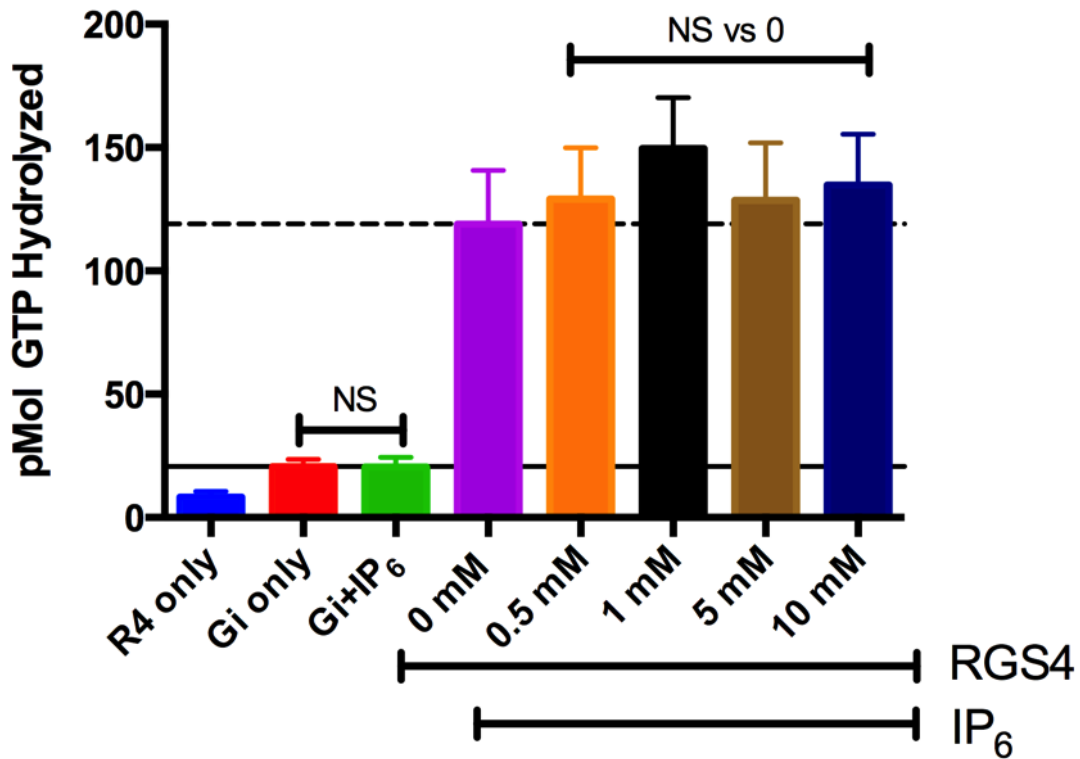
Figure 4.9 Steady-state GAP activity of RGS4 in the presence of IP_6. Although CSPs are observed in the vicinity of the G-protein binding interface on RGS4 (Val121 and Gln122), no effect on RGS4 GAP activity is observed in the presence of IP_6.

GAP Activity Of RGS4 Is Unaffected By IP_6

In order to determine whether IP_6 binding to RGS4 affects GAP activity towards $G_{\alpha i1}$, a steady-state GAP assay was employed to measure changes in RGS-accelerated GTP hydrolysis. A mutant form of $G\alpha_{i1}$ (R178M, A326S) that is capable of spontaneous (receptor-independent) guanine nucleotide exchange was incubated with a saturating concentration of GTP for 100 minutes in order to measure RGS4 stimulation with and without IP_6. The basal GTP hydrolysis of the $G_{\alpha i1}$ mutant is unaffected by 10 mM IP_6 (**Figure 4.9**). The IP_6 effect on RGS4 GAP activity showed a trend upward at lower concentrations, although this is not

observed above 1 mM, and none of the effects were significant compared to the buffer controls (**Figure 4.9**).

Discussion

In order to elucidate the binding site of lipids that are proposed to act as endogenous regulators of RGS4 [2, 3], I employed IP_6 as a soluble derivative of PIP_3 to map the interactions sites and the corresponding effects on protein stability and function. Using solution NMR, the direct interactions of a non-covalent small molecule with an allosteric site are observed on the RGS4 structure at single residue resolution. Based on the combination of NMR studies and docking models the binding site of IP_6 on RGS4 occurs on α-helix 5 within a small pocket formed by three lysine residues (Lys110 &112/113). The largest perturbations in this cluster are Ser108 and Lys110, which are directly in the middle of the pocket (**Figure 4.6 B**). Based on the docking model, IP_6 also appears to directly overlap Pro109, although this residue in not visible in the amide spectrum.

Interestingly, the pocket characterized in this study along α-helix 5 forms the site where CCG-50014 is predicted to dock to RGS8 [22]. This groove may be an important first contact site for CCG-50014 before intercalating into the helix bundle to react with Cys95, as discussed in Chapter II. Despite the overlapping binding site between IP_6 and CCG-50014, there was no inhibition of RGS4 GAP activity observed in the presence of IP_6. Based on the NMR data, the IP_6 interaction produces more localized structural perturbations around α-helix 4 and

116

5, which comprise the binding site. In contrast, binding of CCG-50014 shows extensive structural perturbations throughout all four α-helices in the bundle regions (α_{4-7}). The steric effects of CCG-50014 when bound to Cys95 within the four-helix bundle of RGS4 stabilize an open-like state that results in inhibition of GAP activity. Such a drastic structural change is not indicated by the IP_6 binding data in the HSQC experiments. However, the IP_6 binding site may represent an important new allosteric surface for structure-based ligand identification. Identifying non-covalent ligand interactions on RGS4 would provide an important first step toward inhibitor/activator development, which could then be followed by potency and efficacy optimization in further studies.

The data in Chapter II provide a model where a small molecule binding at the IP_6 pocket could inhibit RGS4 GAP activity through an insertion of a large hydrophobic functional group that would stabilize a more open conformation of RGS4 with respect to the four-helix bundle. Alternatively, if PIP_3 and IP_6 bind to overlapping areas of RGS4 then IP_6 would be expected to reverse the inhibition of PIP_3 micelles based on the results shown in **Figure 4.9**. As a result, a small molecule interaction at the IP_6 binding site could act as an "activator" of RGS4 by blocking the actions of endogenous phospholipids. IP_6 could be used as a proof of concept for this effect in GAP assays using either membrane preparations or pure proteins with PC/PIP_3 micelles. Decreases in RGS4 mRNA levels have been found in the post-mortem brains of patients with Alzheimer's disease [27, 28]. This correlation suggests that RGS4 upregulation could be beneficial in the treatment of Alzheimer's disease. Additionally, in genetic mouse models RGS4

appears to be a positive modulator of the anti-depressant and anti-allodynic actions of monoamine-targeted drugs, suggesting that upregulation could be beneficial for depression or neuropathic pain [29]. However, the opposite appears to be the case for SNC80 and ketamine in this same model, complicating the therapeutic modality of RGS4 in pain management [29].

The data described in this chapter also provides important insight towards the mechanism of RGS4 inhibition by endogenous phospholipids. The inhibition of RGS4 activity by PIP_3 could potentially operate through two mechanisms, in which RGS4 is inhibited through an allosteric structural change or by sequestering RGS4 away from the G-protein-receptor complex. The lack of activity of IP_6 towards RGS4 GAP function suggests the latter mechanism. The localization of RGS4 to receptor complexes has been shown to be a critical component of RGS4 activity. As a result, small molecules that compete with PIP_3 may provide a beneficial upregulation of RGS4 activity, which could be useful based on evidence from preclinical animal models outlined above.

References

1. Sjogren, B., L.L. Blazer, and R.R. Neubig, *Regulators of G protein signaling proteins as targets for drug discovery.* Prog Mol Biol Transl Sci, 2010. **91**: p. 81-119.
2. Ishii, M., et al., *Phosphatidylinositol 3,4,5-trisphosphate and Ca2+/calmodulin competitively bind to the regulators of G-protein-signalling (RGS) domain of RGS4 and reciprocally regulate its action.* Biochem J, 2005. **385**(Pt 1): p. 65-73.
3. Popov, S.G., et al., *Ca2+/Calmodulin reverses phosphatidylinositol 3,4, 5-trisphosphate-dependent inhibition of regulators of G protein-signaling GTPase-activating protein activity.* J Biol Chem, 2000. **275**(25): p. 18962-8.
4. Ishii, M. and Y. Kurachi, *Assays of RGS protein modulation by phosphatidylinositides and calmodulin.* Methods Enzymol, 2004. **389**: p. 105-18.
5. Tu, Y. and T.M. Wilkie, *Allosteric regulation of GAP activity by phospholipids in regulators of G-protein signaling.* Methods Enzymol, 2004. **389**: p. 89-105.
6. Campbell, S., et al., *Modulation of HIV-like particle assembly in vitro by inositol phosphates.* Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10875-9.
7. Chen, H.C., et al., *Membrane docking geometry of GRP1 PH domain bound to a target lipid bilayer: an EPR site-directed spin-labeling and relaxation study.* PLoS One, 2012. **7**(3): p. e33640.
8. Sasakawa, N., M. Sharif, and M.R. Hanley, *Metabolism and biological activities of inositol pentakisphosphate and inositol hexakisphosphate.* Biochem Pharmacol, 1995. **50**(2): p. 137-46.
9. Morrison, R.S., et al., *Inositolhexakisphosphate (InsP6): an antagonist of fibroblast growth factor receptor binding and activity.* In Vitro Cell Dev Biol Anim, 1994. **30A**(11): p. 783-9.
10. Blazer, L.L., et al., *Reversible, allosteric, small-molecule inhibitors of RGS proteins.* Mol Pharmacol, 2010.
11. Storaska, A.J. and R.R. Neubig, *Chapter Eight - NMR Methods for Detection of Small Molecule Binding to RGS4*, in *Methods in Enzymology*, P.M. Conn, Editor. 2013, Academic Press. p. 133-152.
12. Moy, F.J., et al., *NMR structure of free RGS4 reveals an induced conformational change upon binding Galpha.* Biochemistry, 2000. **39**(24): p. 7063-73.
13. Moy, F.J., et al., *1H, 15N, 13C, and 13CO assignments and secondary structure determination of RGS4.* J Biomol NMR, 1999. **15**(4): p. 339-40.
14. Lang, P.T., et al., *DOCK 6: Combining techniques to model RNA-small molecule complexes.* Rna-a Publication of the Rna Society, 2009. **15**(6): p. 1219-1230.

15. Grosdidier, A., V. Zoete, and O. Michielin, *Fast Docking Using the CHARMM Force Field with EADock DSS.* Journal of Computational Chemistry, 2011. **32**(10): p. 2149-2159.

16. Hepler, J.R., et al., *RGS4 and GAIP are GTPase-activating proteins for Gq alpha and block activation of phospholipase C beta by gamma-thio-GTP-Gq alpha.* Proc Natl Acad Sci U S A, 1997. **94**(2): p. 428-32.

17. Hollinger, S. and J.R. Hepler, *Cellular regulation of RGS proteins: modulators and integrators of G protein signaling.* Pharmacol Rev, 2002. **54**(3): p. 527-59.

18. Vedadi, M., et al., *Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination.* Proc Natl Acad Sci U S A., 2006. **103**(43): p. 15835-40. Epub 2006 Oct 11.

19. Ericsson, U.B., et al., *Thermofluor-based high-throughput stability optimization of proteins for structural studies.* Anal Biochem., 2006. **357**(2): p. 289-98. Epub 2006 Aug 10.

20. Williamson, M.P., *Using chemical shift perturbation to characterise ligand binding.* Prog Nucl Magn Reson Spectrosc., 2013. **73:1-16.**(doi): p. 10.1016/j.pnmrs.2013.02.001. Epub 2013 Mar 21.

21. Grosdidier, A., V. Zoete, and O. Michielin, *SwissDock, a protein-small molecule docking web service based on EADock DSS.* Nucleic Acids Research, 2011. **39**: p. W270-W277.

22. Zoete, V., et al., *Use of the FACTS solvation model for protein-ligand docking calculations. Application to EADock.* Journal of Molecular Recognition, 2010. **23**(5): p. 457-461.

23. Blazer, L.L., et al., *A nanomolar-potency small molecule inhibitor of Regulator of G protein Signaling (RGS) proteins.* Biochemistry, 2011.

24. Saetre, P., E. Jazin, and L. Emilsson, *Age-Related Changes in Gene Expression are Accelerated in Alzheimer's Disease.* Synapse, 2011. **65**(9): p. 971-974.

25. Emilsson, L., P. Saetre, and E. Jazin, *Low mRNA levels of RGS4 splice variants in Alzheimer's disease: Association between a rare haplotype and decreased mRNA expression.* Synapse, 2006. **59**(3): p. 173-176.

26. Stratinaki, M., et al., *Regulator of G protein signaling 4 [corrected] is a crucial modulator of antidepressant drug action in depression and neuropathic pain models.* Proc Natl Acad Sci U S A., 2013. **110**(20): p. 8254-9. doi: 10.1073/pnas.1214696110. Epub 2013 Apr 29

Chapter V

Conclusions

Modulating protein complexes with small molecules is a major challenge that has limited the number of potential drug targets being pursued in industry and academia. RGS proteins are representative of the many challenges associated with targeting PPIs outlined in Chapter I. As a result, very few RGS inhibitors have been described (see table 1.1), and previous to this work only one chemical scaffold had been identified with cellular activity. The primary goal of the RGS studies in our lab is to identify RGS modulators for use as pharmacological tools to dissect the role of RGS proteins in biology and disease, and also as chemical starting points to develop RGS-directed therapeutics. The main objective of my research has been to pursue innovative drug discovery strategies that will advance the development of novel RGS modulators. I have focused on RGS4 as a prototypical RGS protein that has been well studied in many systems. RGS4 is highly enriched in the CNS and appears to be an intriguing target for neurological disorders based on recent evidence from animal models [1, 2] that have been highlighted in previous chapters.

Through the work in these chapters we now understand the molecular mechanism of covalent thiol-directed RGS inhibitors, which has revealed a new druggable site on RGS4 and provided the theoretical framework to begin

studying RGS dynamics as a mechanism of inhibitor selectivity. The identification and characterization of the first cell active and reversible small molecule inhibitors of RGS4 have contributed new pharmacological tools that can contribute to the study of RGS function in whole cells. Finally, I have provided the first direct evidence of non-covalent small molecule interactions at two allosteric sites on the RGS4 structure.

Summary of Results

In chapter II, the allosteric mechanism of inhibition by a highly potent RGS inhibitor (CCG-50014) was elucidated. The mechanism of inhibition occurs through a conformational change that exposes cysteine 95 to the CCG-50014 molecule. Binding of the compound to RGS4 stabilizes an open structure that translates into downstream structural perturbations in the G-protein binding interface. This study reveals the importance of RGS4 dynamics in the exposure of key cysteine residues to the inhibitor, which results in this stabilization of non-native conformations. These studies reveal the potential for a novel cryptic site (C-site) that is formed around the buried cysteine residues, and may be more druggable than previously identified sites (**Appendix I**).

The third chapter focuses on identifying new cell-active RGS4 inhibitors using a cell-based calcium signaling assay. Several new cell-active RGS inhibitors were identified that rely on the presence of cysteine residues in RGS4 for activity. The mechanism of the identified inhibitors further illustrates the highly sensitive nature of RGS4 towards cysteine-directed compounds, even in the context of the

reducing environment of the cell. In this study, RGS19 was also particularly sensitive to the inhibitors compared to other RGS homologs. It is possible that RGS19 is also a very dynamic protein, and this could be a mechanism that differentiates RGS family members in their sensitivity to cysteine-directed compounds. Furthermore, these compounds are also substantially reversible, in contrast to CCG-50014, which is irreversible. These compounds provide important new tools to probe RGS biology in cells, and potentially *in vivo* in future studies.

In chapter IV I used solution NMR to show direct evidence of a non-covalent small molecule interaction on an allosteric site of RGS4. A soluble analog of PIP_3 was employed to map the interaction interfaces of RGS4 with polyanionic lipids. This study identified the site on RGS4 to which IP_6 binds. The lack of a functional effect on RGS4 GAP activity by IP_6 suggests that the mechanism of PIP_3 inhibition of RGS4 depends on sequestration, rather than producing a structural change affecting function. Alternatively, it could be that structural changes leading to inhibition of RGS4 GAP activity need to occur within the context of the membrane. Interestingly, the IP_6 binding site directly overlaps the proposed docking site of CCG-50014. This may be a potentially new surface on RGS4 that could be utilized for structure-based screening for novel RGS modulators.

Future Research Directions

Probing Conformational Dynamics As a Way to Target Cryptic Pockets

Proteins exist as a collection of ensembles that can have various representative (low energy) states across a free energy pathway. The conformational forms and accessibility of these states play a large role in molecular recognition for protein and ligand binding [3]. The structural dynamics of RGS4 appear to play a significant role in exposing cysteine residues buried in an allosteric site to cysteine-reactive RGS inhibitor molecules. This represents an important model for allosterically acting PPI modulators, where small molecules targeting allosteric pockets present in transient excited-state conformations can modulate protein function by transmitting structural changes to an active site or interaction interface. Understanding the various conformational populations of RGS4 has revealed a discrete pocket with potentially more advantageous drug-binding properties than those found in the X-ray or NMR structures. These excited-state "cryptic" pockets can be used for structure-based drug design strategies [4]. The existence of cryptic pockets should be identifiable using biophysical screens, but access to enough resources to cover the chemical space needed to identify such pockets can limit success. Therefore, probing structural dynamics as a part of a structure-based strategy provides a mechanism to identify the most druggable pockets within the conformational space of a protein. Virtual compound databases that are orders of magnitude larger that what would be possible using physical methods alone can be applied towards these unique conformations sites on the protein.

The TAMD simulations performed in Chapter II provide a plethora of poses across the conformational trajectories defined for the RGS4 crystal and NMR starting structures. These can be used to provide unique conformations for small molecule docking studies. The simplest mechanism for selecting poses would be to run docking studies on poses at a specific time interval, for example every 50-100 picoseconds across the total simulation length, but a major drawback is that certain poses may not closely represent real conformations in solution. Alternatively, poses can be selected by clustering similar conformations within the MD trajectory in order to generate a series of predicted metastable states. Clustering is based on several methods, including geometric clustering that expresses each conformation as a set of points from which a distance metric (e.g. Cα RMSD) can be applied to minimize distances within a conformational subset relative to others [5]. Although there is an implication for kinetic similarity based on geometric matching, actual kinetic clustering is achieved by calculating transition probabilities between different states [6]. Markov state models [7] of the transition rates across the free energy space can be used to partition the MD dataset into kinetically meaningful clusters, such that the probability of exchange between two states within a cluster is significantly higher than exchange with a conformational state in an adjacent cluster [8-10]. A range of clustering methodology can applied to MD datasets [10-12], but the principle strategy is to generate a set of unique poses that represent relevant conformations (energy minimums across the sampling space) with more advantageous small molecule

binding properties than what is apparent from the NMR or x-ray crystal structures alone.

Experimental Molecular Dynamics of RGS Proteins

Based on the results from TAMD on RGS4, it is apparent that RGS4 may be very dynamic. An important question that arises from the results shown in Chapter II is whether differences in the dynamic motions of RGS proteins play a role in the differential sensitivity to thiol-reactive compounds. This is a particularly intriguing question in view of the large difference in sensitivity to CCG-50014 between RGS4 and RGS8. The thermal stability of RGS4 and RGS8 is approximately 55 °C for both proteins. However, other biophysical indications such as the ease with which RGS8 undergoes crystal lattice formation, whereas RGS4 does not, suggest RGS8 may be much more rigid.

Molecular dynamics simulations would be an excellent starting point to begin examining this question, but ultimately experimental dynamics are needed to validate the physical differences between RGS4 and other RGS family members. This question can be answered from several experiments, one of which would be through hydrogen-deuterium exchange followed by mass spectrometry [13, 14]. It would be expected that amide protons that are buried in the RGS helix bundle comprising helices 4-7 would show greater deuterium exchange on RGS4 compared RGS8, or possibly other RGS homolgs. These results could be compared to positive and negative controls, either cross linking the helix in a closed conformation, or incubating the protein with CCG-50014 to

126

produce the opposite effect. Thus, hydrogen-deuterium exchange experiments may provide a relatively straightforward approach to validating differences between RGS4 and other RGS family members. Complications could arise from the resolution obtained from this method, which is typically on the order of several residues. Thus only large differences in the dynamics of the RGS structure may be observed.

A higher resolution method for exploring the molecular dynamics is to employ solution NMR studies that can quantify protein motions out to the microsecond time-scale. Carr Purcell Meiboom Gill (CPMG)-NMR relaxation dispersion experiments can quantify relaxation rate constants at a single residue resolution out to these time-scales [15, 16]. Preliminary relaxation data has already been obtained for RGS4 in collarboration with Dr. Erik Zuiderweg at University of Michigan. The T_2 relaxation data were collected by running the HSQC-based experiments on RGS4 with and without the CPMG spin-echo pulse sequence. This experiment employed only three delay points in order to increase acquisition time and reduce data analysis time. As a result, we can get a qualitative picture of residues on RGS4 that are slowly exchanging between two distinct chemical states. It is expected that peaks in the NMR spectrum that show broadening should correspond to areas of the RGS4 structure involved in relatively large conformational shifts. This experiment was employed to determine whether areas of the RGS4 structure involved in the motions predicted in the TAMD simulations would show broadening in the CPGM experiment. Interestingly, broadened peaks clustered around the C-terminus of helix 4 and

residues within the loop connecting helix 4 and 5 on RGS4. A second small cluster was observed from residues on the N and C-terminus that fold in close proximity to one another (**Appendix IX**). This data is consistent with the alpha helix 5 and 6 motions predicted in the TAMD simulations in Chapter II that resulted in an open conformation. It is unclear how the N- and C-terminal residues relate to the dynamic motions observed in the TAMD trajectories, although large fluctuations are seen in these residues in several of the simulations outlined in Chapter II. Future studies will seek to repeat this experiment on RGS4 with increased pulse-delays in order to fit relaxation rates of each backbone amide in the RGS4 structure. Importantly, these preliminary results support using CPMG-relaxation dispersion NMR experiments to quantifiably compare dynamical differences in RGS homologs in future studies.

Fragment Screening for RGS4 Ligands

Two successive NMR-based fragment-screening studies were performed on RGS4 by our lab in order to identify new chemical scaffolds on which to build RGS modulators using SAR by NMR. Approximately 3600 fragments in total failed to identify any ligands for the protein after successive validation steps were applied to preliminary hits. Previous high-throughput biochemical screens against RGS4 have only identified covalent inhibitors [17-19]. On the basis of ligandability assessment via fragment screening results [20], this would suggest that RGS proteins are conventionally undruggable. However, based on literature examples indicating the difficulty in identifying suitable ligands from a single

approach [21], it appears that a multifaceted strategy that employs different detection methods may be required. This creates the challenge of utilizing several different techniques to measure binding with a reasonably sized fragment library.

Concluding Perspective on RGS-directed Therapeutic Strategies

G-protein coupled receptors (GPCRs) are medically significant since ~30% of prescription drugs target these receptors [22]. Signal transduction through GPCRs is tightly regulated through a number of intracellular PPIs that affect both localization and downstream signaling [23-25]. RGS proteins are less broadly expressed than the GPCRs they regulate [26-28], and there is specificity for both the G-alpha subtypes and receptors to which they couple [29, 30]. This suggests small molecules targeting RGS proteins could provide a novel way to selectively augment signaling through a specific GPCR [31-33].

Therapeutic strategies for modulating activity or cellular levels of certain RGS homologs are becoming clearer with genetic and pharmacological approaches delineating their individual roles. Humans with low RGS2 levels [34] and RGS2 knockout mice are hypertensive as a result of enhanced $G\alpha_q$-mediated vasoconstriction signaling in peripheral vascular smooth muscle cells [35]. Increasing RGS2 levels would be particularly useful as an anti-hypertensive therapy, and recent data showing elevation of RGS2 protein in response to cardiotonic steroids bolsters the use of small molecules for this approach [36]. Additionally, reduced RGS4 levels linked to schizophrenia [37] suggest that

upregulation is a potentially beneficial therapeutic strategy. In contrast, RGS4 mRNA is upregulated in animal models of neuropathic pain [38], and may contribute to opioid insensitivity observed in these conditions. As a result, the preclinical data currently available suggest a potential benefit from both RGS activators and inhibitors depending on the RGS homolog and the context of disease.

In order to advance RGS proteins as a viable therapeutic target, there is a critical need to pursue innovative strategies to develop non-covalent RGS-specific modulators. Although there are effective covalent drugs currently on the market (for example, clopidogrel, Plavix; Sanofi-Aventis, Bristol-Myers Squibb), concerns over off-target reactivity make the transition to commercial drug-candidate stages difficult. Covalent, reversible compounds can provide some additional advantage over irreversible compounds, as they may be less cytotoxic as a pharmacological tool, or less prone to idiopathic drug-induced toxicity as a therapeutic [39-41]. Compounds with these characteristics will provide a step forward for RGS modulation. Both the identification and validation of new druggable pockets on RGS4 through the course of these studies provide important information for future work. In conclusion, these studies provide the critical information needed to pursue new and innovative approaches to selectively and non-covalently target RGS proteins.

References

1. Chen, Y., et al., *Neurabin scaffolding of adenosine receptor and RGS4 regulates anti-seizure effect of endogenous adenosine.* J Neurosci, 2012. **32**(8): p. 2683-95.
2. Lerner, T.N. and A.C. Kreitzer, *RGS4 is required for dopaminergic control of striatal LTD and susceptibility to parkinsonian motor deficits.* Neuron, 2012. **73**(2): p. 347-59.
3. Boehr, D.D., R. Nussinov, and P.E. Wright, *The role of dynamic conformational ensembles in biomolecular recognition.* Nat Chem Biol, 2009. **5**(11): p. 789-96.
4. Anderson, A.C., *The process of structure-based drug design.* Chem Biol, 2003. **10**(9): p. 787-97.
5. Chema, D. and A. Goldblum, *The "nearest single neighbor" method-finding families of conformations within a sample.* J Chem Inf Comput Sci., 2003. **43**(1): p. 208-17.
6. Bowman, G.R., X. Huang, and V.S. Pande, *Using generalized ensemble simulations and Markov state models to identify conformational states.* Methods., 2009. **49**(2): p. 197-201. doi: 10.1016/j.ymeth.2009.04.013. Epub 2009 May 4.
7. Pande, V.S., K. Beauchamp, and G.R. Bowman, *Everything you wanted to know about Markov State Models but were afraid to ask.* Methods., 2010. **52**(1): p. 99-105. doi: 10.1016/j.ymeth.2010.06.002. Epub 2010 Jun 4.
8. Bowman, G.R., *Improved coarse-graining of Markov state models via explicit consideration of statistical uncertainty.* J Chem Phys., 2012. **137**(13): p. 134111. doi: 10.1063/1.4755751.
9. Noe, F., et al., *Hierarchical analysis of conformational dynamics in biomolecules: transition networks of metastable states.* J Chem Phys., 2007. **126**(15): p. 155102.
10. Keller, B., X. Daura, and W.F. van Gunsteren, *Comparing geometric and kinetic cluster algorithms for molecular simulation data.* J Chem Phys., 2010. **132**(7): p. 074110. doi: 10.1063/1.3301140.
11. Wolf, A. and K.N. Kirschner, *Principal component and clustering analysis on molecular dynamics data of the ribosomal L11.23S subdomain.* J Mol Model, 2013. **19**(2): p. 539-49.
12. Bowman, G.R., L. Meng, and X. Huang, *Quantitative comparison of alternative methods for coarse-graining biological networks.* J Chem Phys, 2013. **139**(12): p. 121905.
13. Ahmed, A.H., et al., *Dynamics of cleft closure of the GluA2 ligand-binding domain in the presence of full and partial agonists revealed by hydrogen-deuterium exchange.* J Biol Chem., 2013. **288**(38): p. 27658-66. doi: 10.1074/jbc.M113.495564. Epub 2013 Aug 12.
14. Mysling, S., et al., *Characterizing the dynamics of alpha-synuclein oligomers using hydrogen/deuterium exchange monitored by mass spectrometry.* Biochemistry, 2013: p. 5.
15. Kovrigin, E.L., et al., *Faithful estimation of dynamics parameters from CPMG relaxation dispersion measurements.* J Magn Reson, 2006. **180**(1): p. 93-104.

16. Vallurupalli, P., D.F. Hansen, and L.E. Kay, *Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy.* Proc Natl Acad Sci U S A, 2008. **105**(33): p. 11766-71.
17. Blazer, L.L., et al., *Reversible, allosteric, small-molecule inhibitors of RGS proteins.* Mol Pharmacol, 2010.
18. Blazer, L.L., et al., *A nanomolar-potency small molecule inhibitor of Regulator of G protein Signaling (RGS) proteins.* Biochemistry, 2011.
19. Roman, D.L., et al., *Identification of small-molecule inhibitors of RGS4 using a high-throughput flow cytometry protein interaction assay.* Mol Pharmacol, 2007. **71**(1): p. 169-75.
20. Edfeldt, F.N., R.H. Folmer, and A.L. Breeze, *Fragment screening to predict druggability (ligandability) and lead discovery success.* Drug Discov Today, 2011. **16**(7-8): p. 284-7.
21. Dias, D.M., et al., *Is NMR Fragment Screening Fine-Tuned to Assess Druggability of Protein–Protein Interactions?* ACS Medicinal Chemistry Letters, 2013.
22. Williams, C. and S.J. Hill, *GPCR signaling: understanding the pathway to successful drug discovery.* Methods Mol Biol, 2009. **552**: p. 39-50.
23. McCoy, K. and J. Hepler, *Regulators of G protein signaling proteins as central components of G protein-coupled receptor signaling complexes.* Prog Mol Biol Transl Sci, 2009. **86**: p. 49-74. Epub 2009 Oct 7.
24. Pierce, K.L., R.T. Premont, and R.J. Lefkowitz, *Seven-transmembrane receptors.* Nat Rev Mol Cell Biol, 2002. **3**(9): p. 639-50.
25. Tesmer, J.J., *Structure and function of regulator of G protein signaling homology domains.* Prog Mol Biol Transl Sci, 2009. **86**: p. 75-113.
26. Doupnik, C.A., T. Xu, and J.M. Shinaman, *Profile of RGS expression in single rat atrial myocytes.* Biochim Biophys Acta, 2001. **1522**(2): p. 97-107.
27. Gold, S.J., et al., *Regulators of G-protein signaling (RGS) proteins: region-specific expression of nine subtypes in rat brain.* J Neurosci, 1997. **17**(20): p. 8024-37.
28. Larminie, C., et al., *Selective expression of regulators of G-protein signaling (RGS) in the human central nervous system.* Brain Res Mol Brain Res, 2004. **122**(1): p. 24-34.
29. Anger, T., et al., *RGS protein specificity towards Gq- and Gi/o-mediated ERK 1/2 and Akt activation, in vitro.* J Biochem Mol Biol, 2007. **40**(6): p. 899-910.
30. Wang, Q., L.Y. Liu-Chen, and J.R. Traynor, *Differential modulation of mu- and delta-opioid receptor agonists by endogenous RGS4 protein in SH-SY5Y cells.* J Biol Chem, 2009. **284**(27): p. 18357-67.
31. Blazer, L.L. and R.R. Neubig, *Small molecule protein-protein interaction inhibitors as CNS therapeutic agents: current progress and future hurdles.* Neuropsychopharmacology, 2009. **34**(1): p. 126-41.
32. Neubig, R.R. and D.P. Siderovski, *Regulators of G-protein signalling as new central nervous system drug targets.* Nat Rev Drug Discov, 2002. **1**(3): p. 187-97.
33. Sjogren, B., L.L. Blazer, and R.R. Neubig, *Regulators of G protein signaling proteins as targets for drug discovery.* Prog Mol Biol Transl Sci, 2010. **91**: p. 81-119.

34. Yang, J., et al., *Genetic variations of regulator of G-protein signaling 2 in hypertensive patients and in the general population.* J Hypertens, 2005. **23**(8): p. 1497-505.

35. Heximer, S.P., et al., *Hypertension and prolonged vasoconstrictor signaling in RGS2-deficient mice.* J Clin Invest, 2003. **111**(4): p. 445-52.

36. Sjogren, B., et al., *Cardiotonic steroids stabilize regulator of G protein signaling 2 protein levels.* Mol Pharmacol, 2012. **82**(3): p. 500-9.

37. Mirnics, K., et al., *Disease-specific changes in regulator of G-protein signaling 4 (RGS4) expression in schizophrenia.* Mol Psychiatry, 2001. **6**(3): p. 293-301.

38. Garnier, M., et al., *Up-regulation of regulator of G protein signaling 4 expression in a model of neuropathic pain and insensitivity to morphine.* J Pharmacol Exp Ther, 2003. **304**(3): p. 1299-306.

39. Lee, C.U. and T.N. Grossmann, *Reversible covalent inhibition of a protein target.* Angew Chem Int Ed Engl, 2012. **51**(35): p. 8699-700.

40. Smith, A.J., et al., *Beyond picomolar affinities: quantitative aspects of noncovalent and covalent binding of drugs to proteins.* J Med Chem, 2009. **52**(2): p. 225-33.

41. Tummino, P.J. and R.A. Copeland, *Residence time of receptor-ligand complexes and its effect on biological function.* Biochemistry, 2008. **47**(20): p. 5481-92.
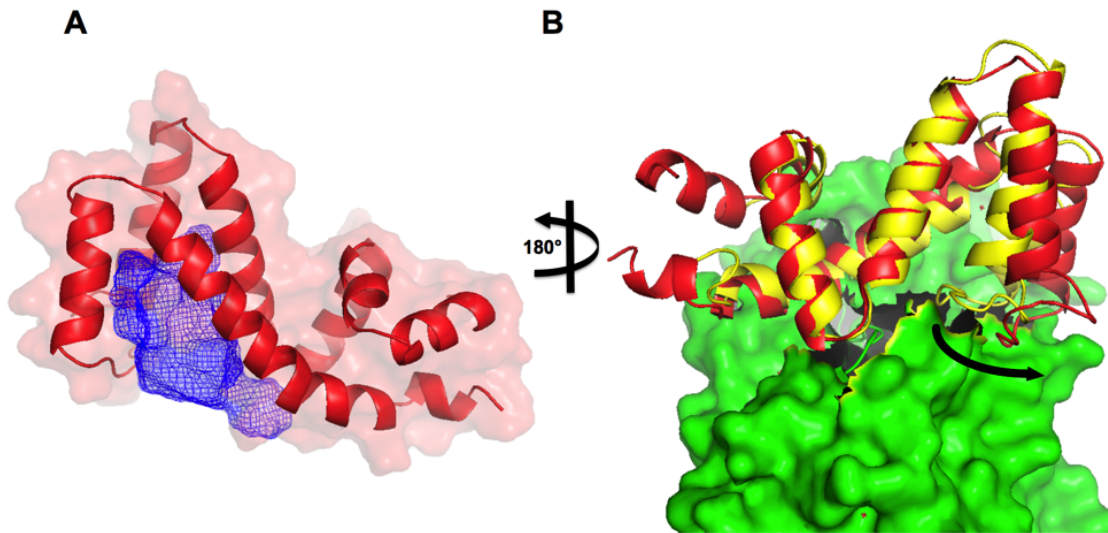
Figure AI Open conformation of RGS4 selected from a single TAMD simulation frame reveals a newly predicted C-site, which results from of the outward displacement of α-helices 5 & 6. **A**) The selected MD frame was used to calculate pocket size using PocketAnnotate and PocketDepth finder, resulting in a pocket volume of 1131 Å3, which is more than 5-fold larger than previously identified pockets on RGS4. **B**) Alignment of the TAMD-generated open conformation of RGS4 (Red) with the Gαi-RGS4 complex (Green & Yellow, respectively; Pdb:1agr). A small molecule that could stabilize this conformational would likely inhibit the PPI via a 3.5 Å displacement of α helix 5 and 6.5 Å change in α helix 6, both rotating causing a disruption in the interface.
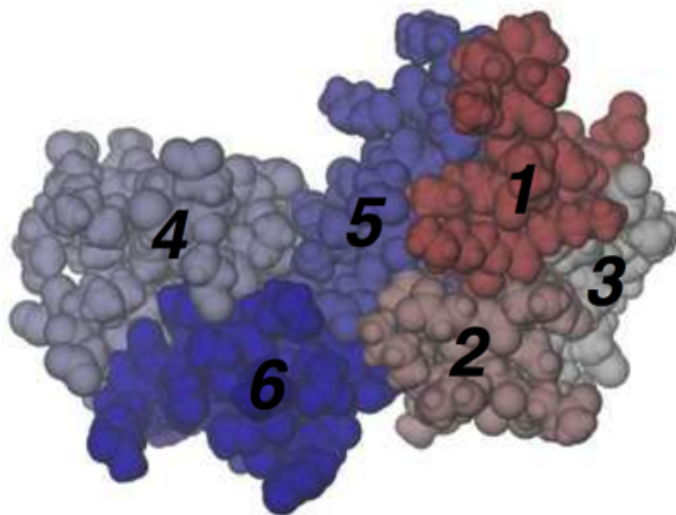
Subdomain	Mass (kDa)	Residues
1	1.56	103-114, 143-145
2	1.63	115-128
3	1.63	129-142
4	3.32	51-74, 174-178
5	3.43	88-102, 146-158
6	3.49	75-87, 159-173

Figure AII Subdomain partitions of RGS4 are shown for TAMD simulations. Each subdomain is colored and labeled with a total of 6 subdomains (18 CVs).

Appendix III

	Mode 1	Mode 2	Mode 3
PC 1	0.01	**0.41**	0.23
PC 2	0.12	0.20	0.03
PC 3	0.10	0.17	0.15

Figure AIII Overlaps between the first three principal components (PCs) and the first three low-frequency normal modes.
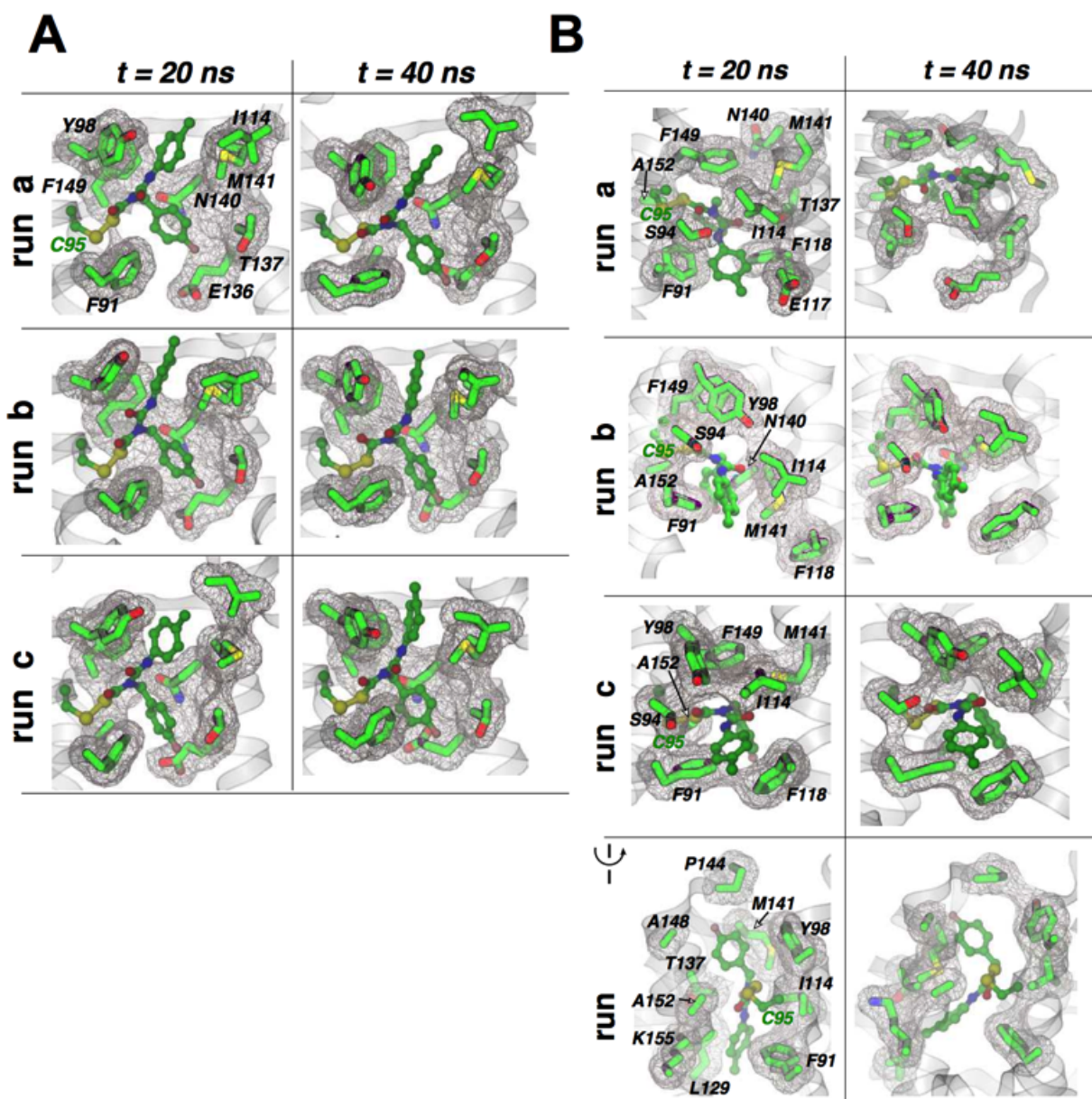
Figure AIV Residues interacting with the inhibitor molecule (CCG-50014). **A**) At two different time-points during MD equilibration, highlighted are the positions of the side-chains of residues directly in contact with CCG-50014. Covalently linked (to Cys95) CCG-50014 is shown in small spheres, and the labeled side-chains of RGS4 residues are rendered as transparent green sticks. The surface corresponding to each RGS4 residue is also rendered as a gray mesh. The snapshots are based upon three independent 40-ns long MD simulations. **B**) Same data are shown for four independent 40-ns long MD simulations. The bottom-most panels are ~90° rotated in comparison to all top panels for a better view of CCG-50014 and surrounding residues.
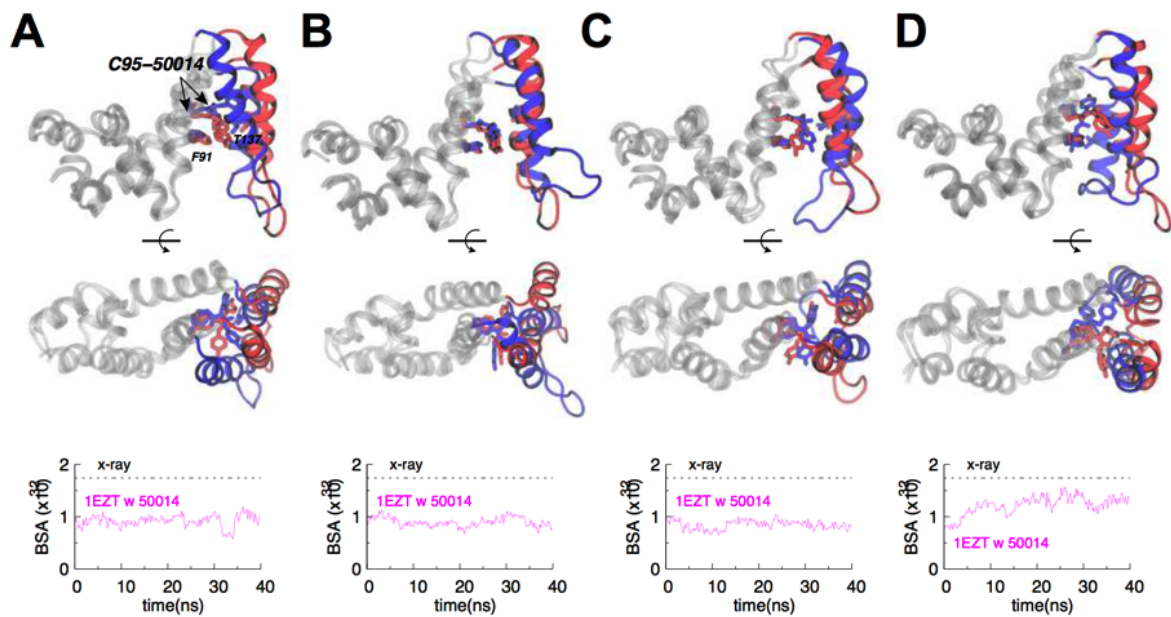
Figure AV Snap shots from the beginning (red) and end (blue) of four independent 40-nslong CCG-50014 bound MD simulations
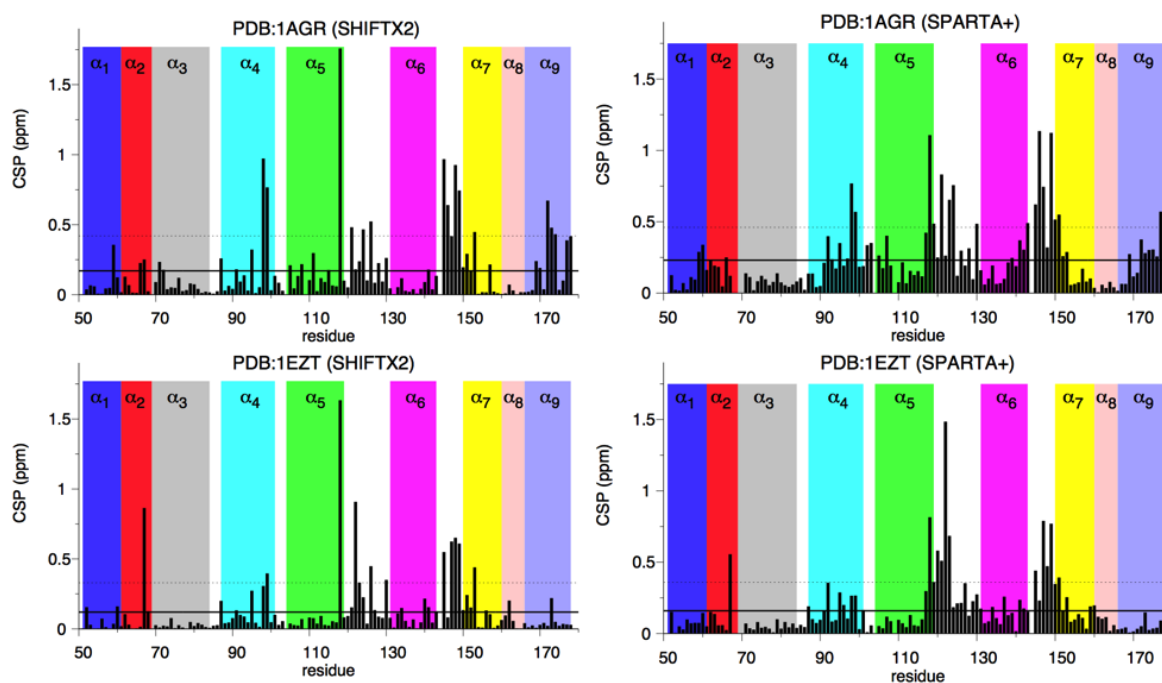
Figure AVI Chemical shift perturbation (CSP; ppm) per residue computed using software package SHIFTX2 and SPARTA+. Computed CSPs are based upon MD simulations of mutant (Cys95) RGS4 with and without CCG-50014. Top and bottom panels, respectively, correspond to CSP predictions based upon MD trajectories starting with PDB structures 1AGR and 1EZT. Solid and dotted horizontal lines indicate values at mean and 1 standard deviation (SD).
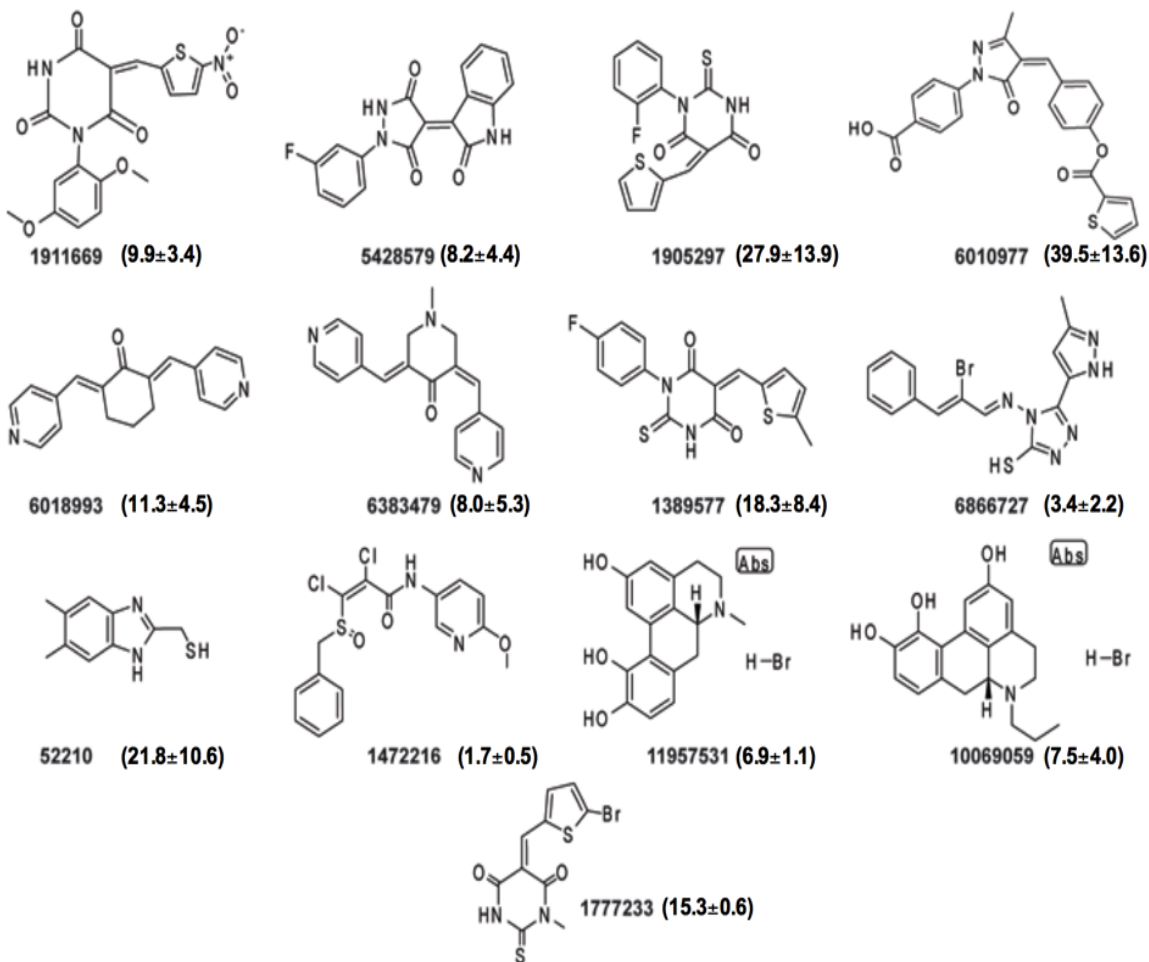
Figure AVII Chemical structures of the thirteen inhibitors directly targeting the RGS-G-protein interaction. The IC_{50}'s in FCPIA for each compound are indicated next to the compound ID (CID) in micromolar.
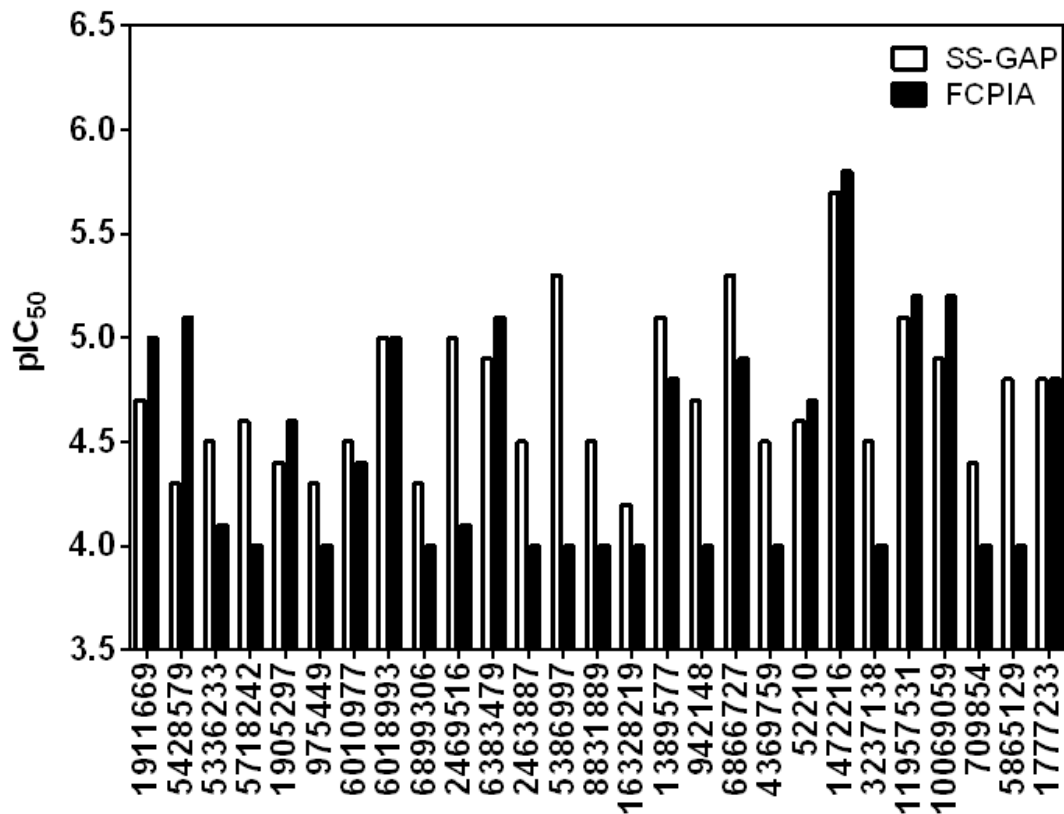
Figure AVIII The log IC$_{50}$ values of each compound are compared between two biochemical assays: SS-GAP (open bars) and FCPIA (close bars).
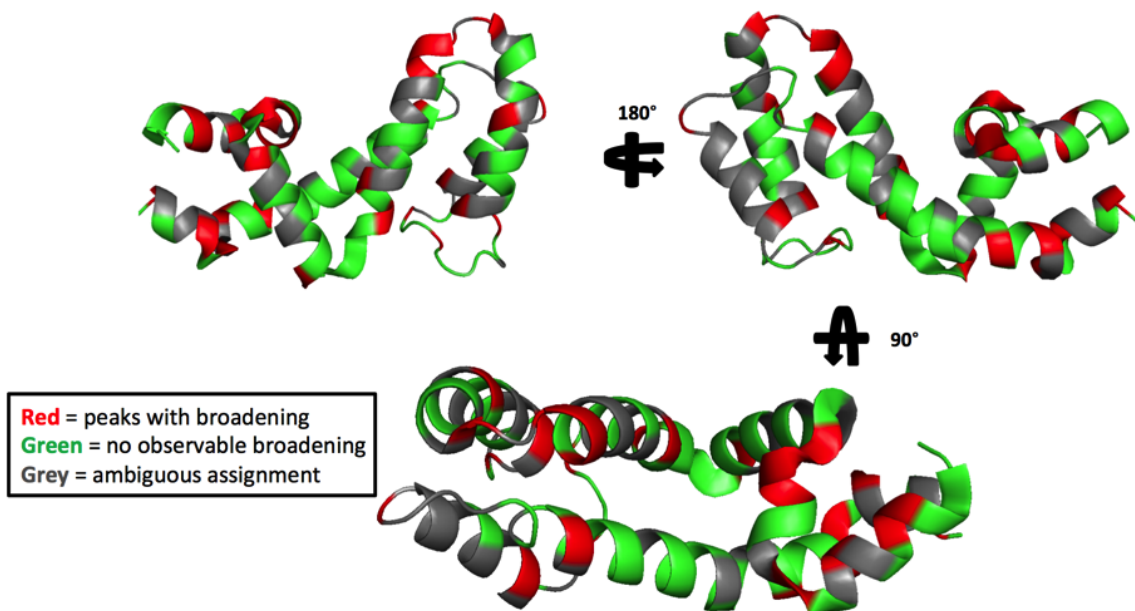
Figure AIX Spin-spin ^{1}H-^{15}N relaxation (T$_2$) CPMG spin echo experiment was used with three pulse delays to qualitatively determine areas of the RGS4 backbone undergoing microsecond or greater dynamics. Areas of the NMR spectrum with broadening correspond to sites on the RGS4 backbone highlighted in Red. Areas in green show no broadening and thus are not likely undergoing long-timescale exchange. The areas highlighted in grey were overlapped in the NMR spectrum and could not be reliably identified.