Spring 2016

Structural and dynamic determinants of inhibitor specificity among regulators of G protein signaling

Colin Anthony Higgins
University of Iowa

STRUCTURAL AND DYNAMIC DETERMINANTS OF INHIBITOR SPECIFICITY AMONG REGULATORS OF G PROTEIN SIGNALING

by

Colin Anthony Higgins

A thesis submitted in partial fulfillment
of the requirements for the Doctor of Philosophy
degree in Pharmacy (Medicinal & Natural Products Chemistry) in the
Graduate College of
The University of Iowa

May 2016

Thesis Supervisor: Associate Professor David L. Roman

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

_____

PH.D. THESIS

_____

This is to certify that the Ph.D. thesis of

Colin Anthony Higgins

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Pharmacy (Medicinal & Natural Products Chemistry) at the May 2016 graduation.

Thesis Committee: _____
David L. Roman, Thesis Supervisor

_____
Ernesto J. Fuentes

_____
Jonathan A. Doorn

_____
Robert J. Kerns

_____
M. Ashley Spies

ACKNOWLEDGEMENTS

I want to thank my mentors, teachers, colleagues, friends, and parents. I give special thanks to my graduate advisor, David Roman for sending me on this wild goose chase and giving me the freedom to explore the weeds but keeping an eye on the path—also for being a great cook and host. I would have been utterly lost without the guidance and badgering I received from my honorary co-advisor, Ernesto Fuentes and from Lokesh Gakhar, C. Andrew Fowler, and Liping Yu.

I would also be remiss not to thank my partners in crime in the Roman lab, past and present, Drs. Duncan Mackie and Carlos "Aaron" Monroy, and Michael Hayes and Christopher Bodle. Thanks too, to our hard-working former lab technician, Michelle Kroc. I would also like to thank Tyrell Towle for making sure I didn't blow anything up when I did synthesis that one time.

Of course, I want to thank my girlfriend Samantha Miller, who tolerated invariably inaccurate estimates of "science time" and has been a steadfast partner and kindred competitive, driven spirit. Last but certainly not least, I would like to thank my parents, Blair and Patricia and siblings, Blair, Paul, Nathan, and Erin, for helping to make me the insufferable twit I am today.

ABSTRACT

Regulator of G Protein Signaling 4 (RGS4) mediates motor defects in Parkinson's disease. Small molecule RGS4 inhibitors (e.g. CCG-50014) modify buried cysteine residues, but the structural and dynamic mechanisms underpinning specificity of inhibitors for RGS4 within the RGS family are poorly understood. We used NMR and other biophysical methods to examine ligand-induced structural changes and the dynamics of unliganded RGS4 and RGS8 that allow ligand binding. NMR and fluorescence spectroscopy data reveal details of the hidden, excited conformational state of RGS4 that exposes Cys148, one of the buried cysteines bound by inhibitors. We further show that specificity of RGS4 inhibitors is driven by differential accessibility of the target cysteine compared to its equivalent in RGS8. Cys148 is buried beneath the lid at the center the α4-α7 helix bundle, and this bundle is destabilized in RGS4 compared to RGS8. Notably, helix 6 is highly destabilized in RGS4 compared to RGS8 and is likely the key mediator of access to Cys148. Our findings provide key insight into the mechanism of allosteric RGS4 inhibition and show that dynamics drive inhibitory specificity among RGS proteins.

PUBLIC ABSTRACT

Drugs are chemicals we take for a specific, desired biological effect. The discovery and development of new drugs is a long and difficult process that requires a thorough understanding of the processes that keep our bodies and minds working. Once we understand the processes that break down in a disease and the proteins and cells involved in those processes, scientists discover and/or develop chemicals specifically suited to restore their function.

I have been working on a project to understand how certain chemicals interact with a protein called regulator of G protein signaling 4 (RGS4). RGS4 is overactive in the brains of patients with Parkinson's disease, and we want to develop chemicals into drugs to block RGS4 to help treat these patients. My work on this project has two components: measuring the specific way the chemicals interact with the protein, and identifying factors that contribute to the sensitivity of RGS4 to these chemicals compared to a related protein, RGS8.

I showed that the chemicals destabilized the structure of RGS4 and that RGS4 is more sensitive to these effects because its structure is more malleable than the structure of RGS8. My work will help the next generation of research aimed at creating those drugs and ensuring that they are safe and have minimal side effects. This project is still in its infancy, and it will take years of hard work and good fortune for those drugs to reach patients.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Introduction

1.1. The functional role of RGS Proteins

1.1.1. G protein-coupled receptor signaling

G protein signaling is a highly conserved, principal eukaryotic signaling paradigm essential for myriad fundamental processes including sensation of sight, smell, taste, and neurotransmission (1). G protein signaling is initiated by the formation of a ternary complex of an extracellular stimulus (photon, ion, small molecule, or protein), a transmembrane G protein-coupled receptor (GPCR), and a cytoplasmic heterotrimeric G protein (Figure 1.1.1.1). Interaction with an extracellular stimulus shifts the receptor's conformational ensemble to alter the topology of the intracellular face of the receptor(2). This, in turn, stabilizes an open conformation of the heterotrimeric G protein's α subunit to facilitate release of guanosine diphosphate (GDP)(3), which is the rate-limiting step in the animal G protein cycle in the absence of a guanine nucleotide exchange factor (GEF) like GPCRs(4).

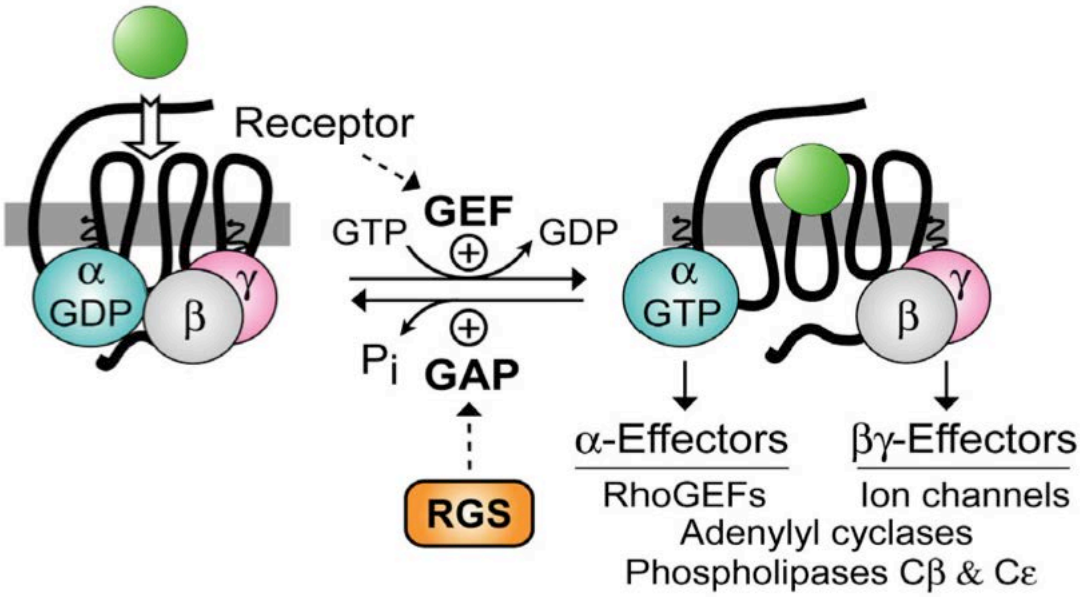

Figure 1.1.1.1 G protein activation cycle diagram

G protein activation occurs upon formation of ternary complex among extracellular ligand (green circle), transmembrane GPCR (black line), and heterotrimeric G protein (shown as individual subunits α (turquoise circle), β (grey circle), and γ (pink circle). G protein deactivation occurs upon hydrolysis of GTP by Gα, accelerated by RGS proteins (orange rounded rectangle). Adapted from (5).

1.1.2. Heterotrimeric G proteins

Heterotrimeric G proteins are made up of α, β, and γ subunits ($\beta\gamma$ form an obligate heterodimeric complex). Upon release of GDP by Gα, guanosine triphosphate (GTP), whose cellular concentration is several-fold higher than GDP, enters the guanine nucleotide binding site. GTP binding alters the conformation of three switch regions in Gα and eliminates the G$\beta\gamma$ binding

surface, leading to the disassociation of the heterotrimer and the ternary complex(6). Following dissociation, Gα and G$\beta\gamma$ diffuse independently along the membrane, binding to and activating numerous effectors including adenylyl cyclase, phospholipase C, ion channels, and various kinases (Figure 1.1.1.1)(7). This process continues until the γ-phosphate of the GTP bound to Gα is hydrolyzed by Gα's intrinsic GTPase activity, converting GTP to GDP, and prompting a rearrangement of Gα switch regions and association with G$\beta\gamma$ into an inactive heterotrimer(8). Intrinsic GTP hydrolysis (k_{cat}) is on the order of 4 min^{-1}(9).

There are four Gα families, G_s, $G_{12/13}$, G_i, and G_q, that signal through distinct downstream effector pathways. Each family also has specific regulators, both GEFs, *vide supra*, and GTPase accelerating proteins (GAPs), though there is some promiscuity among select subsets of Gα families, effectors, and regulators (5). Some effectors of Gα are themselves GAPs, but an important, specialized class of GAP with particular relevance to this work is the eponymous regulator of G protein signaling (RGS) superfamily. RGS proteins are GAPs for G_i and G_q families that, unlike GAPs for small GTPases, facilitate GTP hydrolysis through an indirect mechanism by which residues from RGS proteins coordinate and stabilize flexible switch regions on Gα.

1.2. Conservation within RGS superfamily

1.2.1. RGS protein discovery and classification

The first RGS protein was identified in 1982 in a forward genetic screen in yeast as pheromone-supersensitivity 2 (*sst2*)(10, 11), though it was not recognized as such for a more than decade. The term RGS domain was coined in 1996 by Koelle and coworkers(12), and within a year the first structure of an RGS protein (RGS4 in complex with $G_{i\alpha 1}$) was solved by Tesmer and coworkers (Figure 1.2.1.1)(13). For a historical perspective on the early days of RGS proteins, interested readers may consult an account by Henrik Dohlman(14).

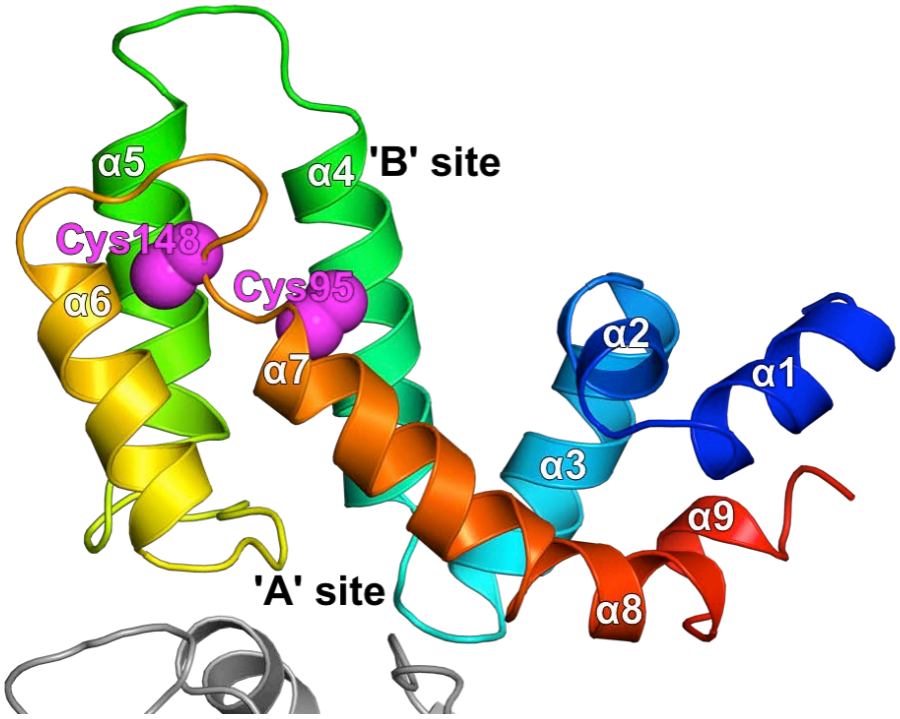

Figure 1.2.1.1 RGS domain cartoon representation

RGS4:Gαi complex (PDB 1AGR, chains E and A). Backbone shown in cartoon representation, Gα (mostly cropped from view) colored grey, RGS4 colored by chainbows in PyMol (N-terminal to C-terminal spectrum from blue to red, respectively). Side chains heavy atoms for Cys95 and Cys148 are shown as spheres, colored either yellow (sulfur) or magenta (carbon).

The constituents of the human RGS superfamily are 37 proteins containing at least one RGS domain. Based on sequence conservation, these are divided into eight subfamilies (Figure 1.2.1.2).

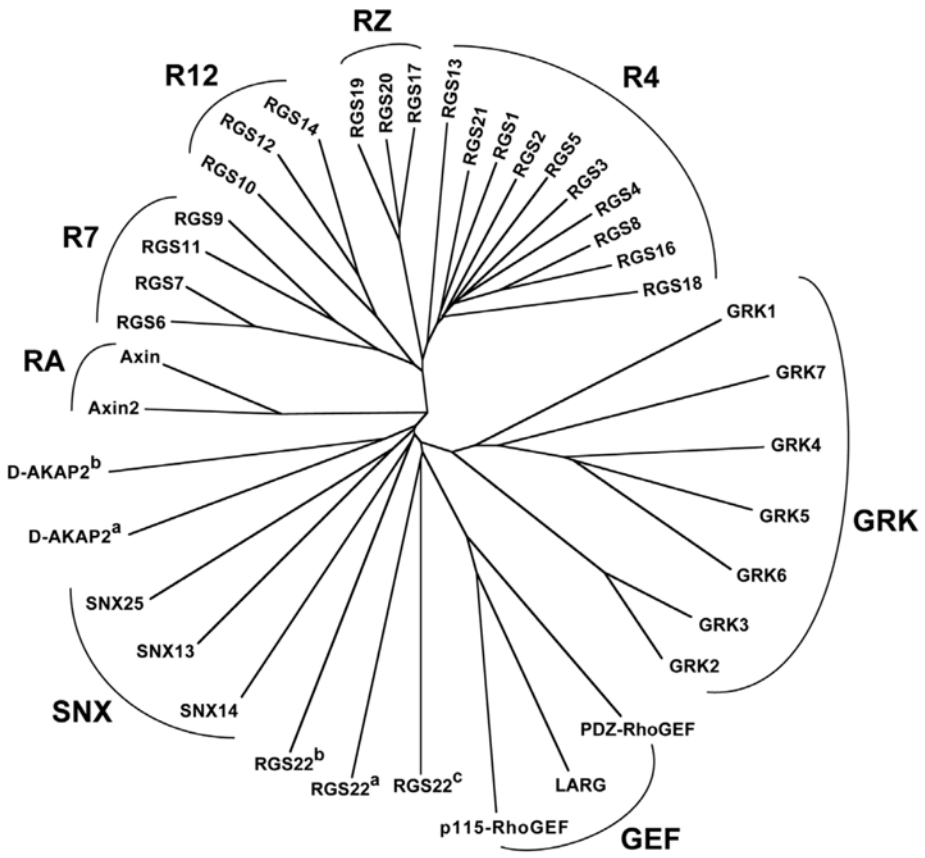

Figure 1.2.1.2 RGS subfamily designations

Unrooted dendrogram showing conservation within each protein's RGS domain. Adapted from (5)

Most subfamilies (all except the R4 and RZ subfamilies) have other domains in addition to the RGS domain, and the RGS domains of some RGS subfamilies lack GAP activity, such as the G protein receptor kinase (GRK) subfamily, leading some to refer to their RGS domain as an RGS homology domain (5). The R4 subfamily (RGSs 1, 2, 3, 4, 5, 8, 13, 16, 18, and 21), of which RGS4 is a member, is made up of RGS proteins that possess no auxiliary domains (except RGS3, which has a post synaptic density protein, Drosophila disc large

tumor suppressor, and zonula occludens-1 protein (PDZ) domain) and whose cellular functions are accordingly restricted to their interaction with Gα proteins.

1.2.2. R4 subfamily sequence conservation

Within the RGS domain (defined here as the 128 residues (51-178) in the canonical RGS4:$G_{i\alpha1}$ crystal structure,), the primary sequence of all 10 proteins in the R4 subfamily is identical at 21% of the positions and similar at 52.3%, though pairwise similarity to the founding member, RGS4, range from 62% similarity (RGS13) to 76% (RGS3 and RGS8)(15) (see Figure 1.2.2.1 for sequence logo showing conservation). Outside of the RGS domain, primary sequences diverge further—including the length of flexible stretches, or arms, extending from the C-terminus of α1 and N-terminus of α9. In Figure 1.2.2.1.

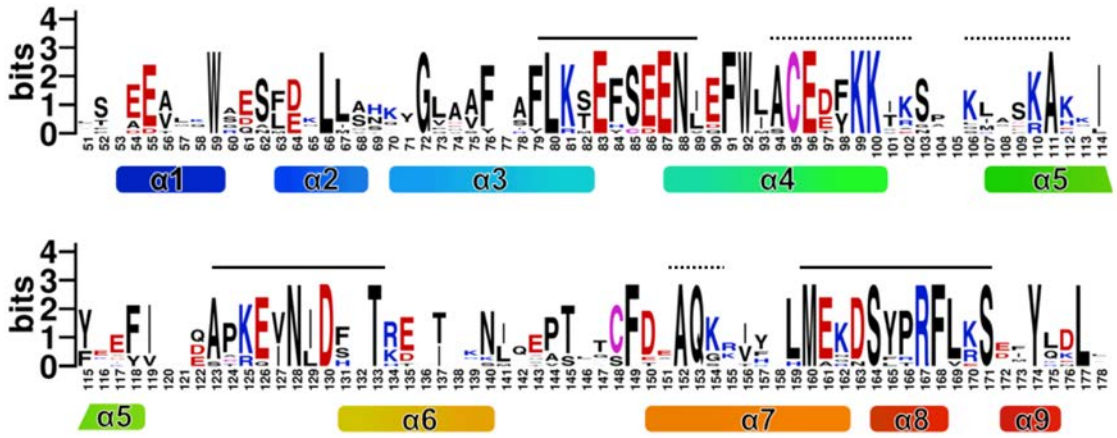

Figure 1.2.2.1 R4 family sequence logo

Sequence logo generated using WebLogo(16) showing conservation of positions from the multiple sequence alignment of human R4 subfamily RGS domains. Conservation is indicated by the overall height (encoded as bits) at each position, and relative heights of stacked 1-letter amino acid symbols indicate relative frequency. Positions with small heights are variable across the sequences in the R4 family, while large heights are conserved. Amino acids are colored according to biophysical properties of side chains: acidic (red), basic (blue), cysteine (magenta), other (black). Secondary structural elements are indicated below each position with colors corresponding to coloring in Figure 1.2.1.1. 'A' and 'B' sites (see 1.4.2 for explanation of these designations) indicated by solid and dotted lines above each position, respectively.

RGS8 and RGS3 are equally similar to RGS4 within their RGS domains, but since RGS3 is an atypical R4 protein given its N-terminal PDZ domain, we have chosen to compare RGS4 with RGS8 to highlight the unique aspects of RGS4.

While both RGS4 and RGS8 have N- and C- terminal extensions, the RGS domain is the only part of RGS proteins required for GAP activity in vitro and folds independently. In all structures reported to date, both the N- and C-terminal extensions are disordered(17). The N-terminus contains an amphipathic helix that is required for membrane localization in cells(18), contributes to GPCR selectivity, and enhances GAP activity in cells by 10000 fold(19), but the physiochemical properties that make this helix essential for biological function render heterologously expressed protein susceptible to isolation in inclusion bodies. As such, nearly all studies on RGS proteins using protein purified from bacteria utilize N-terminal truncation constructs. The number of residues truncated varies, but the convention in the field is to refer to these truncated constructs with a capital Δ followed by the N-terminal residue in the truncation form, e.g. Δ51 RGS4 starts with the 51st residue in the canonical, full length sequence, Val51.

While the C-terminus is similarly disordered in high-resolution structures, this region has historically not been truncated in many studies. After initial studies using Δ51 RGS4, we found that truncating this extension significantly improved purified protein yield from bacteria, reduced spectral crowding in NMR, and improved the quality of SAXS shape reconstruction (see Methods). Both truncation variants were functional in GTPase acceleration assays.

1.2.3. RGS structural conservation

In the Gα-bound RGS4 crystal structure, PDB 1AGR, the RGS domain consists of two subdomains and a total of nine α helices, though α7-α9 are a nearly continuous helix broken only by two bends at conserved serine residues. The larger of the two subdomains is a right-handed, antiparallel four-helix (α4-α7) connected with short loops between α4-5 and α6-7 and a longer loop between α4-5 forming a cleft with the bases of α4 and α7 that make up a substantial portion of the Gα binding interface(Figure 1.2.3.1)(13). The other subdomain, forms a smaller helix bundle made up of α1-3 together with α8-9, which are in turn connected to the disordered N- and C-terminal extensions, respectively.

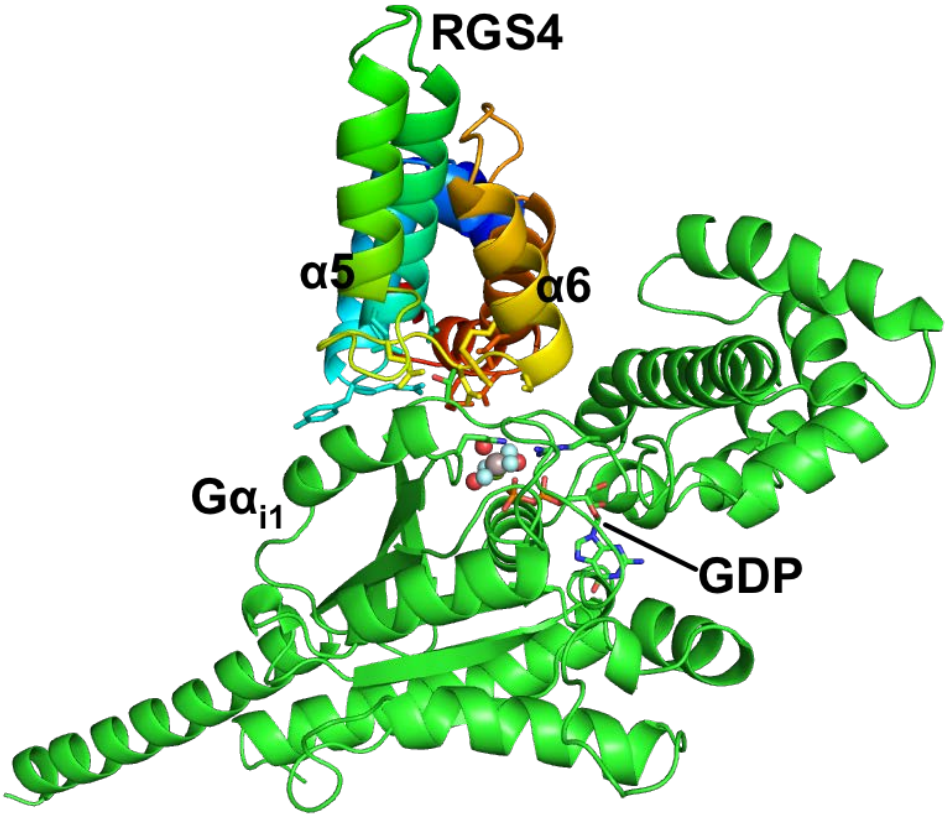

Figure 1.2.3.1 RGS4 Gα-bound structure

RGS4:Gα_{i1} complex (PDB 1AGR, chains E and A) shown with RGS4 α5 and α6 facing toward viewer. Backbone shown in cartoon representation, Gα colored green, RGS4 colored as chainbows (same as Figure 1.2.1.1) Side chain heavy atoms of RGS4 residues involved in Gα binding interface shown as sticks. GDP shown as sticks, AlF_4 (hydrolysis transition state mimic of GTP γ phosphate) and structured, coordinating waters shown as spheres.

As reported by Tesmer and colleagues, the buried surface area of RGS4:Giα1 crystal contacts total 5090 Å^2, of which only 1100 Å^2 consists of a productive interface between RGS4 and Gα switch regions involved in catalysis. The other 3990 Å^2 of buried surface area is ascribed to artifacts of crystal packing given the distance of the contacts from the active site of Gα, but subsequent

findings about dynamics in the major subdomain suggests that these contacts may have contributed substantially to the ordering of RGS4 in the complex. Potentially important contacts include, e.g. chain α5-α6 helix in chain E RGS4 and chain A N-terminal helix Gα and between the chain D Gα αA-αB loop and the RGS4 α4-α5 loop (*vide infra*).

Moy and colleagues solved the RGS4 NMR structure (Figure 1.2.3.2) as part of a drug discovery effort at Wyeth Research in the late 1990s, and they published their structure in 2000(20). The solution structure of RGS4, PDB 1EZT (and ensemble 1EZY), presents an overall similar topology to the Gα-bound form, but has a backbone root mean square deviation (RMSD) of 1.94 Å (21). The major features noted by the authors differing from the bound structure were an extension of α7 (though the numbering used by the original authors was α6 due to interpreting α2-α3 as a single bent secondary structural element), a corresponding rearrangement of the packing of the terminal bundle, and an enlarged binding pocket for Gα Thr182. They proposed a two-stage binding and locking mechanism in which the binding pocket constricts upon the establishment of a hydrogen bond network in RGS4 focused around Gα Thr182.

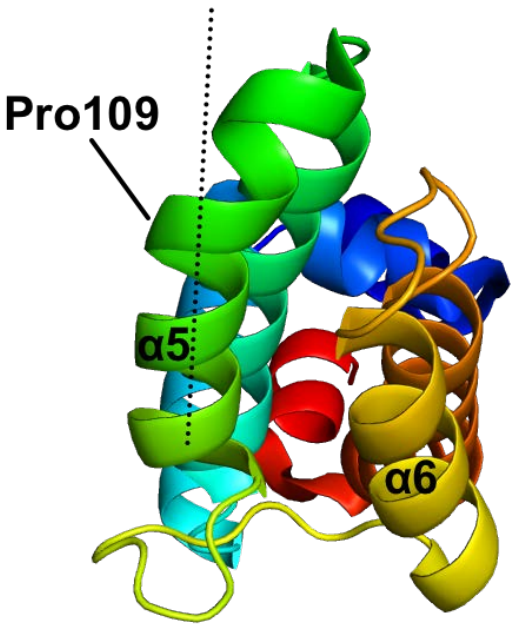

Figure 1.2.3.2 RGS4 apo solution structure

Cartoon representation of solution RGS4 structure (PDB 1EZT) shown with α5 and α6 facing toward viewer. Cartoon colored by chainbows (see Figure 1.2.1.1). Dotted line represents the center of the α5 helix from the RGS4 Gα-bound structure (PDB 1AGR, Figure 1.2.3.1) to demonstrate the bend in the solution structure depicted.

Concerned mainly with the Gα interface of RGS4, the authors neglected to address the other major divergent structural element between the free and bound structures, namely the pronounced bend inward in the top of α5 toward α6, producing a localized 4 Å backbone RMSD compared to the crystal structure, in which the α5 helix is straight. This region in the Gα-bound crystal structure makes extensive crystallographic contacts to regions in Gα (*vide supra*). The bent conformation in the free NMR structure is likely favored given the presence of a

divergent proline (Pro109) on the outer face of the second turn (from the N-terminus) of α5 that abrogates hydrogen bonding in the area.

This topology is conserved across the RGS superfamily with relatively little difference in the apo and $G_{i\alpha}$-bound structures(17). The largest deviations are in the structures of R12 subfamily proteins RGS10 and RGS14, whose α5-α6 loop is lengthened (with a commensurate foreshortening of a canonical α6 helix). Soundararajan and coworkers(17) also report flexibility, indicated by a lack of ordered secondary structure (helix) in the α6-α7 region of RGS10 in their NMR solution structure (Figure 1.2.3.3).

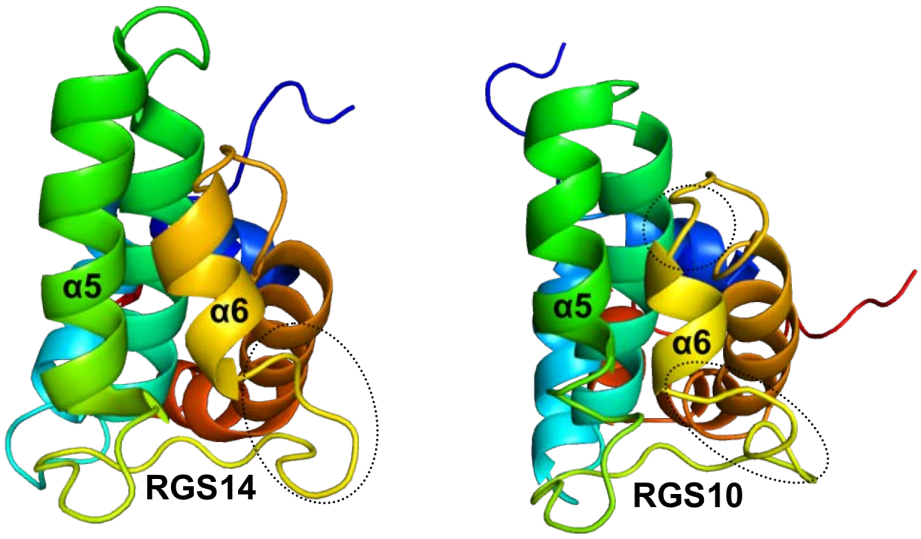

Figure 1.2.3.3 Flexibility in α6 in R12 subfamily

Cartoon representation colored by chainbows. In NMR solution structures of RGS14 (PDB 2JNU) and RGS10 (PDB 2DLR), parts of the α6 region (dotted ellipses) adopt more flexible coil secondary structure instead of the more rigid α helix observed in other RGS structures.

Most R4 proteins are promiscuous GAPs for both $G_{i\alpha}$ and $G_{q\alpha}$, with the notable exception of RGS2, which is $G_{q\alpha}$-specific. Recently, a structure of this interaction showed that RGS2 adopts a tilted binding pose that, while maintaining conserved contacts between the α5-α6 loop and switch regions of $G_{q\alpha}$, adds additional contacts to the helical domain of $G_{q\alpha}$. This complex also reveals that α6 and the C-terminal portion of the α5-α6 loop adopt a novel conformation of among bound RGS structures that is consistent with its apo structure(22). This, coupled with evidence that α6 is disordered in a chimeric RGS2 engineered to bind to Giα, points to an important role for α6 in conferring biological specificity among RGS proteins(23).

1.3. Parkinson disease and the case for RGS4 inhibitors

1.3.1. RGS4 in Parkinson's disease

Many human diseases involve dysfunctional GPCR signaling pathways, and ligands for GPCRs have been a mainstay of clinical pharmacotherapy for decades(24). Despite the success of these compounds in the clinic, limits in receptor and tissue specificity present a significant obstacle to growth in this class through off-target effects. RGS proteins provide a key control mechanism for these GPCR cascades, and due to their differential receptor and tissue distribution pattern, offer a compelling alternative to modulate GPCR pathways

in many disorders(25, 26). Efforts to target RGS proteins for drug discovery started within years of their discovery(20), but recent findings have reinforced a clear need for therapeutically useful RGS inhibitors. Notable among these, is RGS4's involvement in Parkinson's disease (PD).

In the etiology of PD, degeneration of dopaminergic neurons in the substantia nigra pars compacta results in decreased levels of dopamine and increased release of acetylcholine in the striatum(27). While this rise in acetylcholine has been attributed to loss of activity of inhibitory D2 dopamine receptors, which populate cholinergic interneurons, the basal level of interneuron spiking following dopamine depletion does not change, challenging that hypothesis(28). Interestingly, Lerner and Kreitzer recently published their findings that RGS4 is required for dopaminergic control of both striatial long-term depression as well as susceptibility to Parkinsonian motor deficits(29).

This work provides several key pieces of data that validate RGS4 as a clinical target for treatment of PD. First, they demonstrated that both D2 dopamine and A2A adenosine receptors regulate striatial long-term plasticity in an RGS4-dependent manner. Second, they showed that RGS4 knockout mice are resistant to dopamine depletion, a hallmark of PD, and third, RGS4 mice have fewer behavioral deficits in a mouse model (6-OHDA lesion) of PD. The evidence for a role of RGS4 and the utility of an RGS4 inhibitor is demonstrated in that the

direct neuronal injection of the RGS inhibitor CCG-63802 at very high concentration (100 μM) showed disruption of the long-term depression. Unfortunately, this compound does not work in cells (it doesn't cross the plasma membrane, and gets quickly reduced to an inactive form when directly injected)(29). In fact, the authors end that paper by stating, "Hopefully, RGS4 inhibitors with suitable characteristics for clinical use are on the horizon and can be tested as Parkinson's disease therapeutics or for other conditions in which RGS4 is involved."

1.4. RGS4 inhibitors, discovery and specificity

1.4.1. Obstacles to discovery

RGS proteins have several characteristics that render them difficult to inhibit with traditional drug-like molecules though. RGS proteins' eponymous function occurs via a highly conserved protein-protein interaction (PPI) interface (see 1.1.2), and unlike receptors and enzymes, the RGS domain lacks a clear active site or pocket selected for binding small molecules amenable to rational drug design. Clues to the existence of a few potential sites exist, however, based on semi-selective modulation of RGS activity by the endogenous molecules.

1.4.2. Endogenous small molecule interactions

Phosphatidylinositol 3,4,5-triphosphate (PIP_3) inhibits RGS4 upon binding to a region spanning the tops of α4-α5 through an allosteric mechanism(30). Calmodulin competes for binding with PIP_3 at this site in a Ca^{2+}-dependent manner, relieving PIP_3 inhibition of RGS4 activity. This process is proposed to constitute a negative feedback mechanism for $G_{q\alpha}$-mediated Ca^{2+} signaling by accelerating the signal deactivation. The PIP_3-RGS4 interaction is quite specific, as related phospholipids showed no RGS4 binding. Structural evidence for this interaction exists in the form of titratable localized chemical shift perturbations in RGS4 interacting with a soluble mimetic of the PIP_3 head group, inositol hexakiphosphate (IP_6)(31). Though IP_6 was shown to have specific interactions with RGS4, this study failed to demonstrate any functional inhibition of RGS4 GAP activity due to IP_6 binding.

Tightly controlled palmitoylation is important for proper subcellular localization of other constituents of the GPCR signaling cascade including Gα(32), and adjacent to the PIP_3 binding site in α4, palmitoylation of a conserved cysteine (Cys95) has pleiotropic effects on RGS activity. Cys95 palmitoylation enhances membrane localization and GAP activity of RGS10, and RGS16 in mammalian cells, and though RGS4 membrane localization is also enhanced, reports conflict as to the effects of palmitoylation on RGS4 GAP activity(33, 34).

Because localization is an essential feature of GAP activity in cells, a preponderance of evidence points to a positive role palmitoylation on RGS4 function. The case for modulating Cys95 plamitoylation with small molecules as a selective mechanism for RGS4 inhibition is dubious given the high level of homology in the α4 helix around Cys95, however.

Interaction with these endogenous molecules, and a crystal structure of RGS superfamily member Axin bound to a peptide on the opposite face of the RGS domain from the Gα interface led some researchers in the field to denote this back side of the RGS domain the 'B-site.' The 'A-site,' of course, being the interaction interface with Gα(15). These two sites were thought to be the most promising sites in the RGS domain as targets for small molecule inhibitors (see Figure 1.2.1.1 and Figure 1.2.2.1 for a mapping of these sites onto RGS4 tertiary and primary structure, respectively).

1.4.3. Small molecule inhibitors

High-throughput screens (HTS) out of the Neubig lab and the Center for Chemical Genomics (CCG) targeting RGS4 produced a collection of small molecule inhibitors(35-39) that act via covalent modification of cysteine residues and inhibit RGS4 through an allosteric mechanism involving Cys95 and Cys148(40-42). There have been other reports of RGS4 inhibitors(20), though no structures for these molecules were released, and their mechanism of action is

unknown. Nonetheless, the known RGS4 inhibitors offer an unequaled opportunity to learn more about the structural mechanisms by which small molecules can inhibit GAP activity by interacting with the RGS domain.

The first disclosed RGS4 small molecule inhibitor, CCG-4986 (Figure 1.4.3.1), inhibits RGS4/Gα protein-protein interaction (PPI) and RGS4 GAP activity with an IC_{50} of 4 μM but neither RGS8/Gα PPI nor GAP activity(35). Our lab, in collaboration with the Neubig lab, demonstrated that the majority of inhibition was attributable to modification of a single cysteine in the allosteric pocket, Cys148. That study also used mass spectrometry to confirm that CCG-4986 binds RGS8's Cys148 equivalent residue, Cys142, despite not inhibiting Gα binding or GAP function(41).

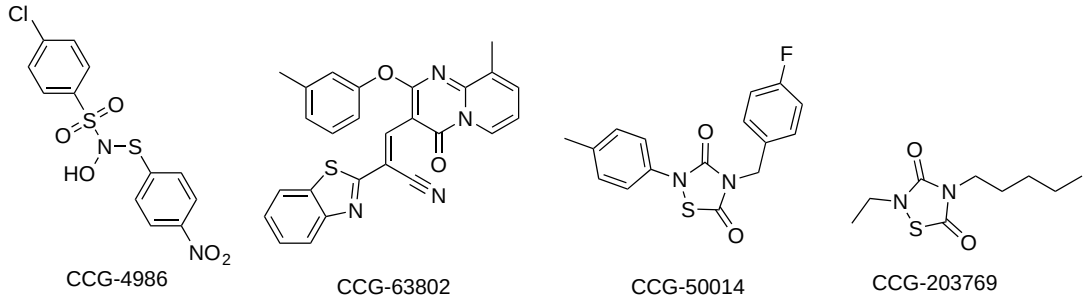

Figure 1.4.3.1 Structures of selected RGS4 inhibitors

CCG-4986 structure depicts the hydrolysis product of published structure that forms spontaneously in aqueous solutions. All others are as previously reported, citations in text.

The second disclosed RGS4 small molecule inhibitor was CCG-63802 (Figure 1.4.3.1), which inhibits RGS4 but also inhibits related R4 proteins and

even more distantly related RZ family proteins with similar potency. CCG-63802 is cysteine dependent and requires binding to the allosteric site on RGS4 for full activity, causing destabilization of target proteins(37). The third RGS4 inhibitor reported was CCG-50014 (Figure 1.4.3.1), a much more potent inhibitor of RGS4 (30 nM) that has modest potency for RGS8, (11 μM) and its binding destabilizes and can even cause certain point mutants of RGS8 to precipitate. Virtual docking of CCG-50014 on RGS8 placed it on the 4-helix bundle containing the allosteric site within 12 Å of Cys142(38). A medicinal chemistry effort generated analogs of CCG-50014 with improved potency and specificity for RGS4 versus RGS8, though all retain cysteine dependence(43). Recently, a lead compound from this series, CCG-203769 (Figure 1.4.3.1), showed in vivo efficacy in a mouse model of PD-like bradykinesia(44).

1.4.4. Specificity and dynamics

That these three inhibitors share a mechanism is not coincidence; all known RGS4 inhibitors follow a pattern of covalent cysteine modification, but the source of differential selectivity of inhibitors for RGS4 versus RGS8 (and other RGS proteins) is an open question. Structural studies support the idea that RGS4 and RGS8, while nearly identical when bound to Gα, exist in different conformations when free from their cognate Gα subunits (see 1.2.3).

In all published structures of RGS proteins, the Cys residues in the B-site are buried entirely from solvent(13, 17, 21), but allosteric-site cysteines must be at least transiently solvent exposed in order to covalently bind inhibitors. Therefore, there is obvious conflict between these structural models and the wealth of evidence from these chemical probes.

A reasonable hypothesis to resolve this discrepancy is the existence of an excited state conformation of the RGS domain that is hidden from biophysical techniques, such as crystallography, that selectively observe low energy conformers. The Neubig lab recently proposed a model, based on accelerated molecular dynamics simulations, by which inhibitors bind to a hidden, excited state conformation of RGS4 wherein Cys95 is exposed to solvent through a rigid body translation of the bottom of the α5-α6 helix pair and connecting loop away from the α4-α7 helix pair in a swinging hinge motion(42). The authors go on to show that covalent attachment of CCG-50014 in the excited state to Cys95 prevents closure of this open hinge in traditional molecular dynamics (MD) simulations. Vashisth and coworkers observed a more uniform distribution of chemical shift perturbation (CSP) in in vitro Cys95-CCG-50014 adducted RGS4 than predicted by these simulations(42). An in vitro analysis of inhibited states would enhance our ability to ascertain the validity of this in silico model.

1.5. Statement of hypothesis

Our long-term goal is to develop RGS4 inhibitors for clinical use in PD. The current best RGS4 inhibitors depend on covalent modification for their efficacy. We lack structural models of the pre-encounter complex (i.e. excite state) and of the inhibited state(s), which are vital to future efforts to improve specificity and/or design non-covalent inhibitors. We will use these inhibitors as chemical probes of RGS4 and RGS8 to establish structural and dynamic determinants of specificity.

Specifically, we will study the effects of inhibitor binding to Cys148 and interrogate the excited state hypothesis by testing the hinge model in the context of motions exposing Cys148 using extensive biophysical characterization of RGS4 and RGS8. We have chosen to focus our efforts at characterizing inhibited states to those stemming from modification of Cys148 as opposed to Cys95 because the conservation of Cys95 and residues surrounding it in all R4, RZ and R12 proteins makes the prospect of obtaining specificity through interactions at that site dubious. That being said, our comparison of the dynamics in RGS4 and RGS8 has relevance for targeting small molecule inhibitors at all sites in the RGS domain.

Our hypothesis has two elements: 1) that inhibitors bind to Cys148 in an excited state conformation of RGS4, disrupting the α4-α7 helical bundle to

increase sampling of an ensemble of non-Gα-binding RGS4 conformations and 2) that inhibitor specificity is driven by a greater ability of RGS4 to adopt the excited state conformation through dynamic rearrangement of the helix bundle.

Chapter 2. Assignment of RGS4 and RGS8 backbone NMR chemical shifts

2.1. Background on NMR

2.1.1. Theory

Peak intensity at a given chemical shift in an NMR spectrum is proportional to the population of nuclei in a particular chemical environment. Chemical exchange between the native state and one or more non-native states (e.g. ligand bound) causes attenuation of peaks at the native chemical shift as the population is split. Therefore, the ^{1}H-^{15}N heteronuclear single quantum coherence spectroscopy (HSQC) spectrum of a ^{15}N-enriched protein is diagnostic of the protein's folded state since it reports the chemical environment of backbone amides.

Each peak in an HSQC can be assigned to its backbone amide using sequential resonance assignment techniques. These techniques make use of intra-residue and inter-residue through-bond *J* couplings among ^{1}H, ^{15}N, $^{13}C_O$ (carbonyl), $^{13}C_\alpha$ (α carbon), and $^{13}C_\beta$ (except for glycine) atoms in uniformly ^{15}N and ^{13}C-labeled protein.

2.2. WT RGS4 Assignment

2.2.1. Assignment completeness

The ^{1}H-^{15}N HSQC spectra (Figure 2.2.1.1) of the RGS4 RGS domain is expected to have 128 assignable peaks out of 131 residues (the three Pro residues do not produce peaks due to their lack of an amide proton). Of these, 125 amides were assigned based on sequential assignment of HNCACB and CBCAcoNH spectra (Table 2.2.1.1). Backbone carbonyl carbon chemical shifts were assigned via HNCO.

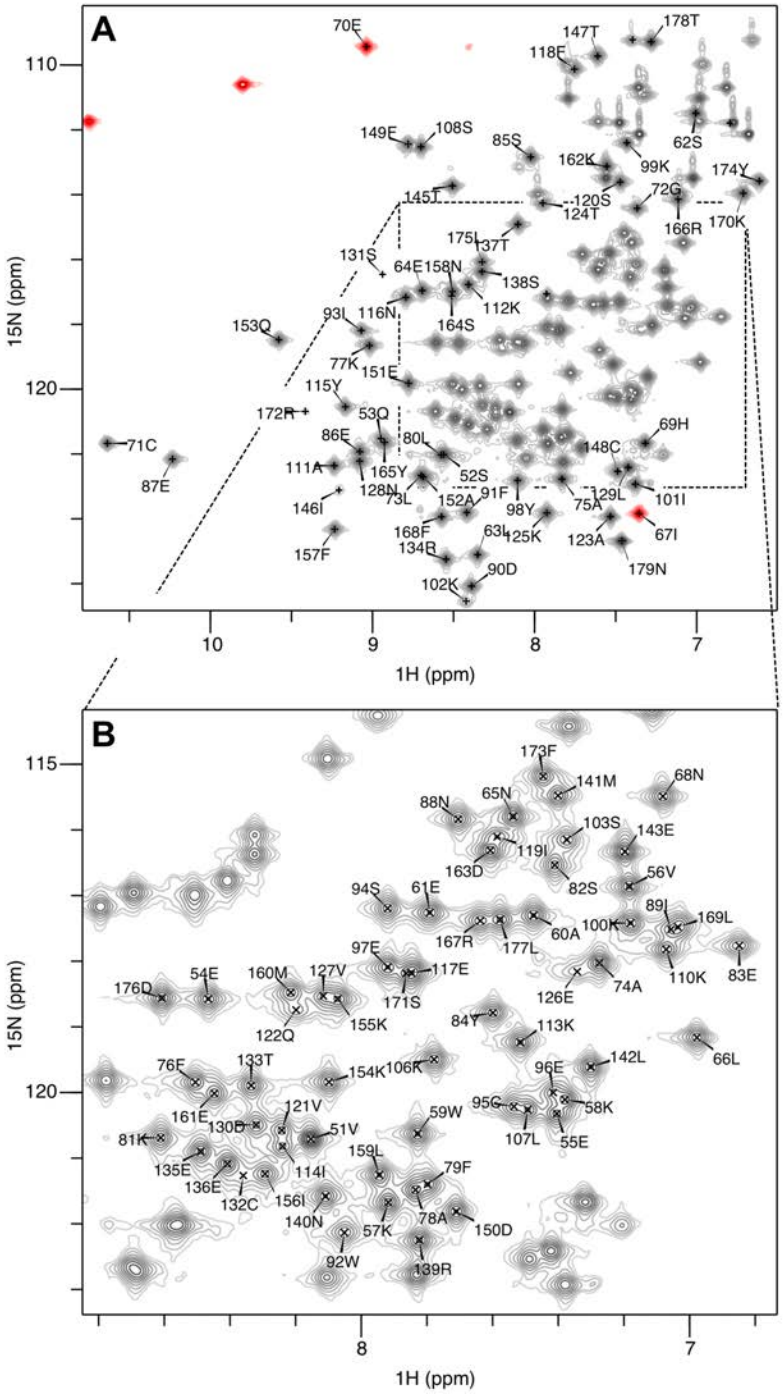

Figure 2.2.1.1 RGS4 WT ^{1}H-^{15}N HSQC

A) Peaks labeled with backbone residue assignments. Red peaks aliased in ^{15}N dimension.

B) Zoom of region indicated by dotted box in A. Peaks lacking assignments in both A&B are side chain resonances. Sample: 1 mM RGS4, pH 7, 25 C, see methods for details.

Category	Total	Assigned
All Residues	131	129 (98.5%)
(Backbone) N+H	258	250 (96.9%)
Backbone	393	379 (96.9%)

Table 2.2.1.1 RGS4 assignment quality report

Quality report generated by CCPN Analysis(45) for RGS4 WT RGS domain combining assigned resonances from HSQC, HNCACB, CBCAcoNH, and HNCO spectra.

Residues present in the construct but lacking detectable peaks include the two N-terminal residues, Ser49 and Met50, both of which are artifacts of the expression vector, and Ser105, which sits atop the α4-α5 loop.

Details for the assignment completeness for each atom type can be found in Table 2.2.1.1.

Atom Type	Total	Assigned	Minimum Shift (ppm)	Mean Shift (ppm)	Maximum Shift (ppm)
H(N)	128	125 (97.7%)	6.62	8.04	10.64
N(H)	131	125 (95.4%)	109.3	118.9	126.5
C(O)	131	125 (95.4%)	173.3	177.3	181.0
Cα	131	129 (98.5%)	46.52	58.95	67.85
Cβ	130	124 (95.4%)	17.66	37.56	72.38

Table 2.2.1.2 RGS4 assignment quality report details

Quality report generated by CCPN Analysis(45) for RGS4 WT RGS domain combining assigned resonances from HSQC, HNCACB, CBCAcoNH, and HNCO spectra.

2.2.2. Comparison to published chemical shifts

Backbone assignments for RGS4 WT RGS domain match published chemical shifts for the majority of peaks (21). Notable exceptions were noted near the termini, where the constructs used diverge. The published assignments

employed a construct with 26 amino acids from the C-terminal extension, whereas our assignments were performed on a construct truncated at the boundary of the RGS domain. Other slight differences likely reflect the pH and temperature difference between the samples. We performed assignments at pH 7 at 25°C, whereas Moy and coworkers used a sample at pH 6 at 30°C(21). A complete list of chemical shift assignments can be found in the appendix.

2.3. WT RGS8 assignment

2.3.1. Assignment completeness

Since RGS8 has no publicly available chemical shift assignments, we carried out *de novo* backbone assignments. The HSQC spectra (Figure 2.2.1.1) of the RGS8 RGS domain is expected to have 135 assignable peaks (non-Pro) out of 138 residues, and of these, 129 amides were assigned with sequential resonance linkage using HNCO, HNcaCO, HNCACB, and CBCAcoNH experiments. We obtained complete or partial assignments for 97.8% of residues (Table 2.3.1.1).

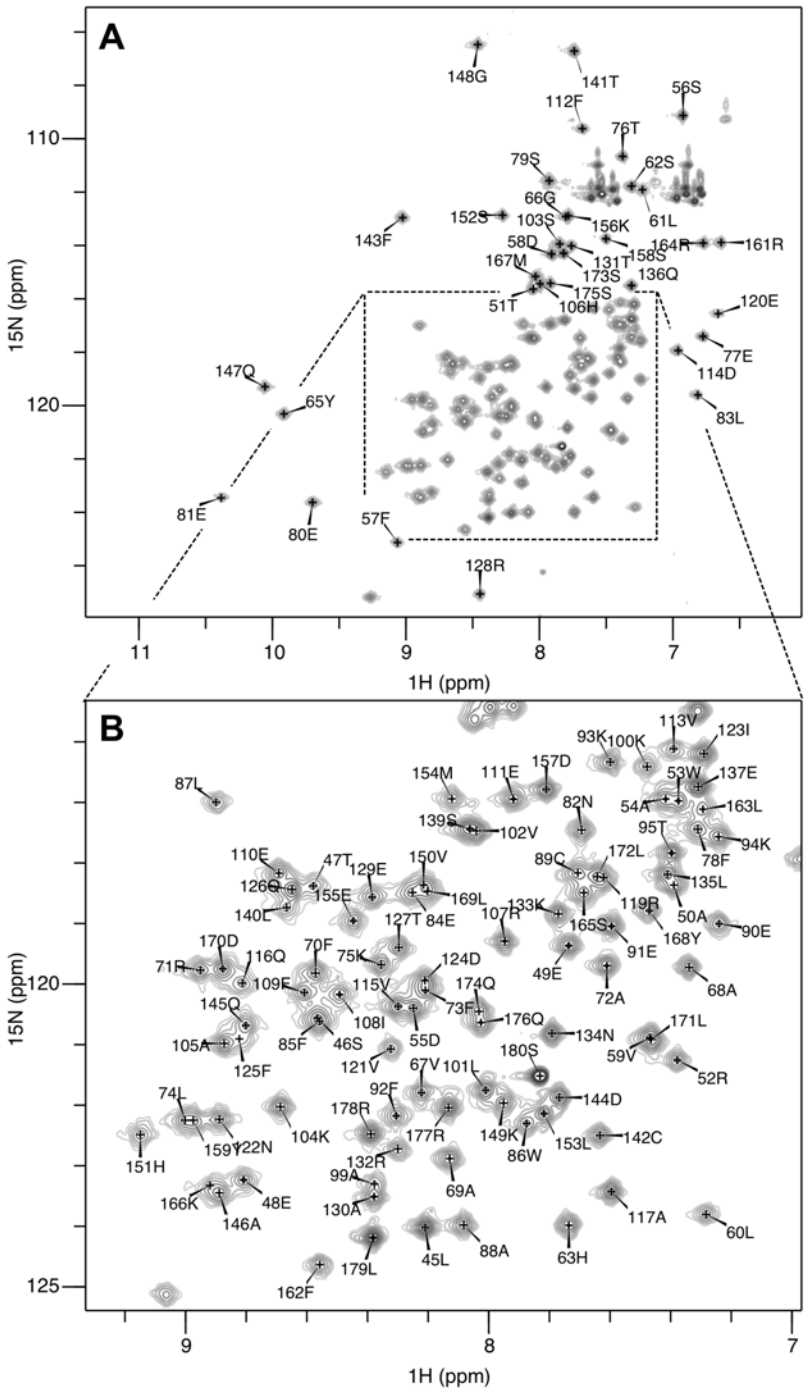

Figure 2.3.1.1 RGS8 WT ^{1}H-^{15}N HSQC

A) Peaks labeled with backbone residue assignments. B) Zoom of region indicated by dotted box in A. Peaks lacking assignments in both are side chain resonances. Sample: 1 mM RGS8, pH 7, 25 C, see methods for details.

Category	Total	Assigned
All Residues	138	135 (97.8%)
(Backbone) N+H	272	258 (94.9%)
Backbone	414	399 (96.4%)

Table 2.3.1.1 RGS8 assignment quality report

Quality report generated by CCPN Analysis(45) for RGS8 WT RGS domain combining assigned resonances from HSQC, HNCACB, CBCAcoNH, HNcaCO, and HNCO spectra.

Residues present in the construct but lacking detectable peaks include two N-terminal residues, Ser43 and Met44, both of which are artifacts of the expression vector; Lys64, at the kink between α2 and α3; and Arg96, Ser97, and Thr98, which sit atop the α4-α5 loop, and the three prolines, Pro118, Pro138, and Pro160. Details of the completeness for each atom type are listed in Table 2.3.1.2.

Atom Type	Total	Assigned	Minimum Shift (ppm)	Mean Shift (ppm)	Maximum Shift (ppm)
H(N)	135	129 (95.6%)	6.64	8.05	10.38
N(H)	138	129 (93.5%)	106.5	118.8	127.1
C(O)	138	135 (97.8%)	172.2	177.2	181.3
Cα	138	135 (97.8%)	46.15	58.90	67.51
Cβ	136	131 (96.3%)	17.34	37.80	69.29

Table 2.3.1.2 RGS8 assignment quality report details

Quality report generated by CCPN Analysis(45) for RGS8 WT RGS domain combining assigned resonances from HSQC, HNCACB, CBCAcoNH, HNcaCO, and HNCO spectra.

2.4. Feature extraction from assigned spectra

2.4.1. Line shape and exchange broadening

Chemical exchange between multiple conformations where the nuclei are in differing chemical environments can broaden peak line shapes. In RGS4, several peaks showed significant line broadening (Figure 2.4.1.1A). These include 126Glu, 131Ser, and 132Cys in the long α5-α6 loop, 102Lys and 146Ile in the loops between α4-α5 and α6-7, respectively; and 172Arg at the bend separating α7 and α8 (see 2.5.1 for explanation of updated helix numbering). RGS8 showed a tighter distribution of peak heights (Figure 2.4.1.1B), though there were a few marginally broadened peaks. These include 81Glu at the base of α4, 120Glu in the long α5-α6 loop, and 125Phe at the base of α6.

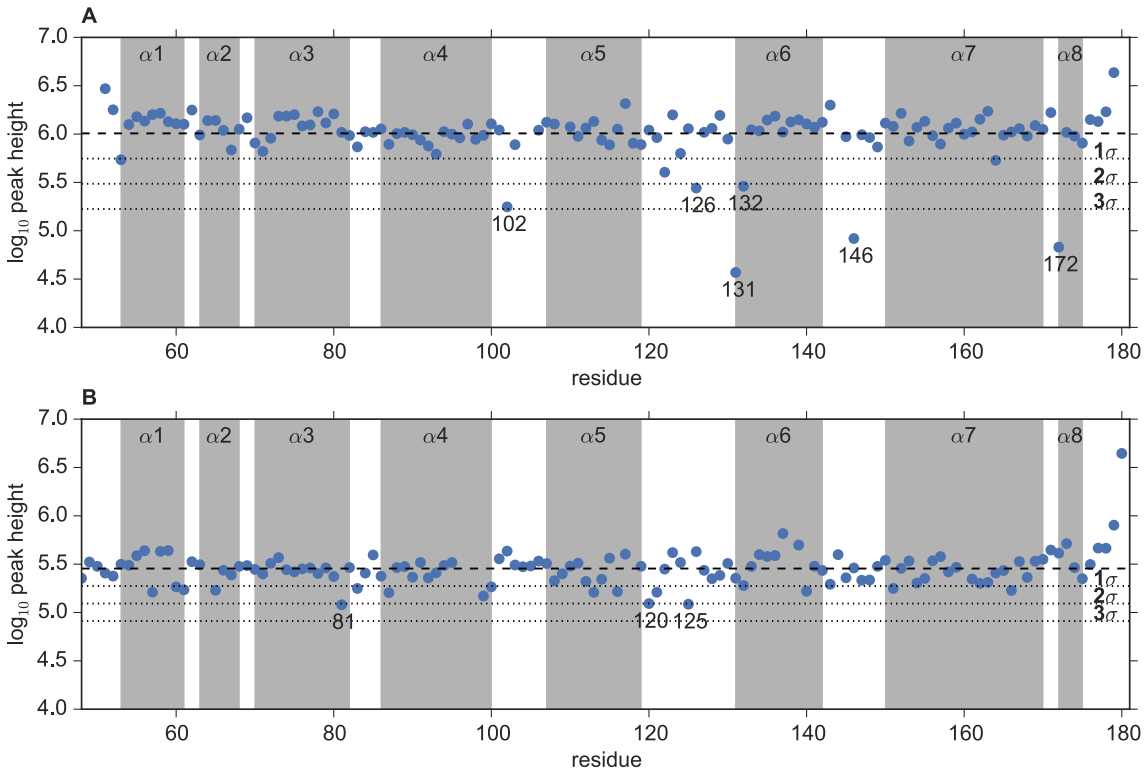

Figure 2.4.1.1 Peak height in WT HSQCs

Height (arbitrary units) for all assigned crosspeaks from representative RGS4 (A) and RGS8 (B) ^{1}H-^{15}N HSQCs. For each spectrum, the mean peak height, μ, is indicated with a dashed line, and the mean minus the standard deviation, σ, and 2σ, and 3σ are shown as dotted lines. Peaks with intensity less than (μ-2σ) are labeled with their assigned residue number. Grey vertical spans indicate helical segments described below (2.5.1).

2.5. Estimation of 2° structure based on chemical shifts

2.5.1. Reconciling differences in 2 structure between RGS4 structures

The free and Gα-bound RGS4 structures have differences in their secondary structural assignments, most notably in the separation of α7-α8 (1AGR) into two distinct helices or as a single extended, bent α7 helix (1EZT).

Modern statistical techniques that make use of the large body of high quality, publicly available datasets with NMR structures and chemical shift assignments make it possible to predict the secondary structural elements with high precision and accuracy from backbone chemical shifts. These algorithms, such as δ2D and CSI 2.0, use probabilistic models and machine learning to give secondary structure prediction accuracy comparable to algorithms that use high resolution 3D coordinates like DSSP(46, 47).

We compared the agreement of secondary structure predictions derived from our backbone chemical shifts to the segment assignments from the headers of the 1EZT and 1AGR PDB files. In general, our data support the designations in 1AGR for the boundaries of α1-α6, but they do not support the split of α7 and α8 into distinct structural elements (Figure 2.5.1.1A). Neither, however, do our data validate the lack of loop/coil regions separating α1 from α2, α2 from α3, and α7 from α8 in 1EZT (Figure 2.5.1.1B). Without a full structural assignment, we cannot confidently assert that the differences in helix formation between the solution and Gα-bound RGS4 structures noted by Moy and coworkers(21) is artifactual, and we recognize that determinations of secondary structure are subjective to some extent. We have chosen to delineate secondary structure segments with a hybrid of the two structures. This segmentation consists of 8 helical segments where α1-α6 reflect α1-α6 in 1AGR, the split α7-α8 1AGR

helices are combined into a single helix, numbered α7, and α8 is the α9 segment from 1AGR. These mappings are corroborated by predictions from δ2D and that match more closely the segments for RGS8 (see 2.5.2), which allows for a more straightforward comparison of elements in the two proteins (see Figure 2.4.1.1 for an example using the shared secondary structural elements).

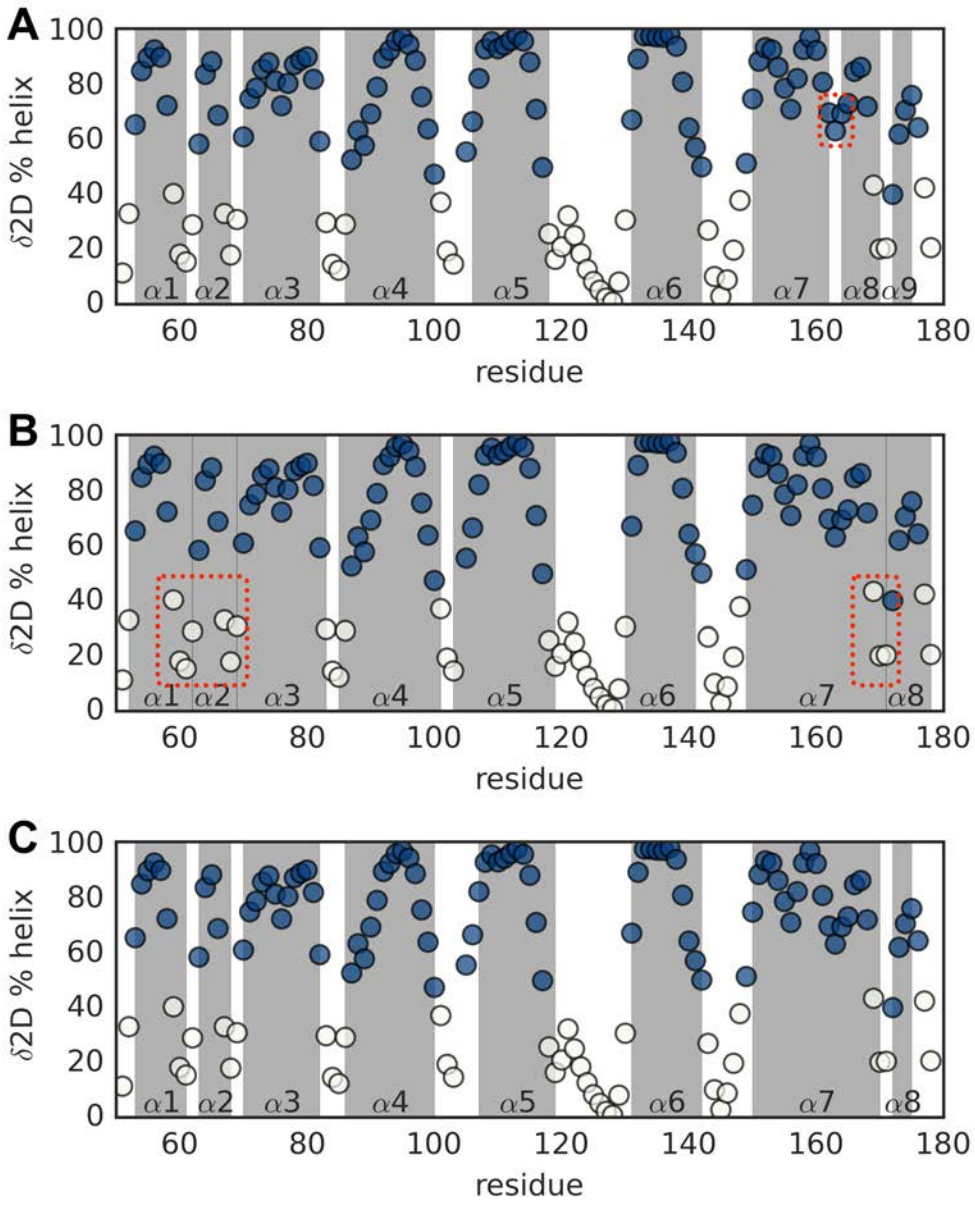

Figure 2.5.1.1 δ2D-predicted RGS4 secondary structure

A-C) Classification (helical: blue, coil: white) and population probability of α helix (ordinate) for each residue (abscissa) based on RGS4 WT chemical shifts at pH 7, 25 C, predicted by δ2D algorithm(47). Grey spans indicate helical segments defined by A) PDB 1AGR, B) PDB 1EZT, and C) hybrid consistent with δ2D predictions for both RGS4 and RGS8. Red dotted outlines indicate regions in A and B inconsistent with δ2D predictions and with classification in other RGS structures (e.g. RGS8, PDB 2IHD)

2.5.2. RGS8 WT secondary structure from chemical shifts

The secondary structure predicted from RGS8 RGS domain backbone chemical shifts correlated reasonably well to the segments defined in the apo RGS8 structure (PDB 2IHD) (Figure 2.5.2.1A). The notable exception is in the α6-α7 loop, δ2D predicts a two-residue pseudo β strand while helices α6 and α7 propagate an extra one-quarter turn into the loop. Here again, the chemical shifts of backbone atoms at the kink separating α7 from α8 do not fall below the threshold constituting a break contiguous secondary structure. Accordingly, we designate this a single helix α7, yielding helical definitions identical to those used for RGS4 (2.5.1) apart from a difference of indexing by 6 residues (e.g. Cys148 in RGS4 is the positional equivalent of Cys142 in RGS8). CSI2.0 gives equivalent results (not shown).

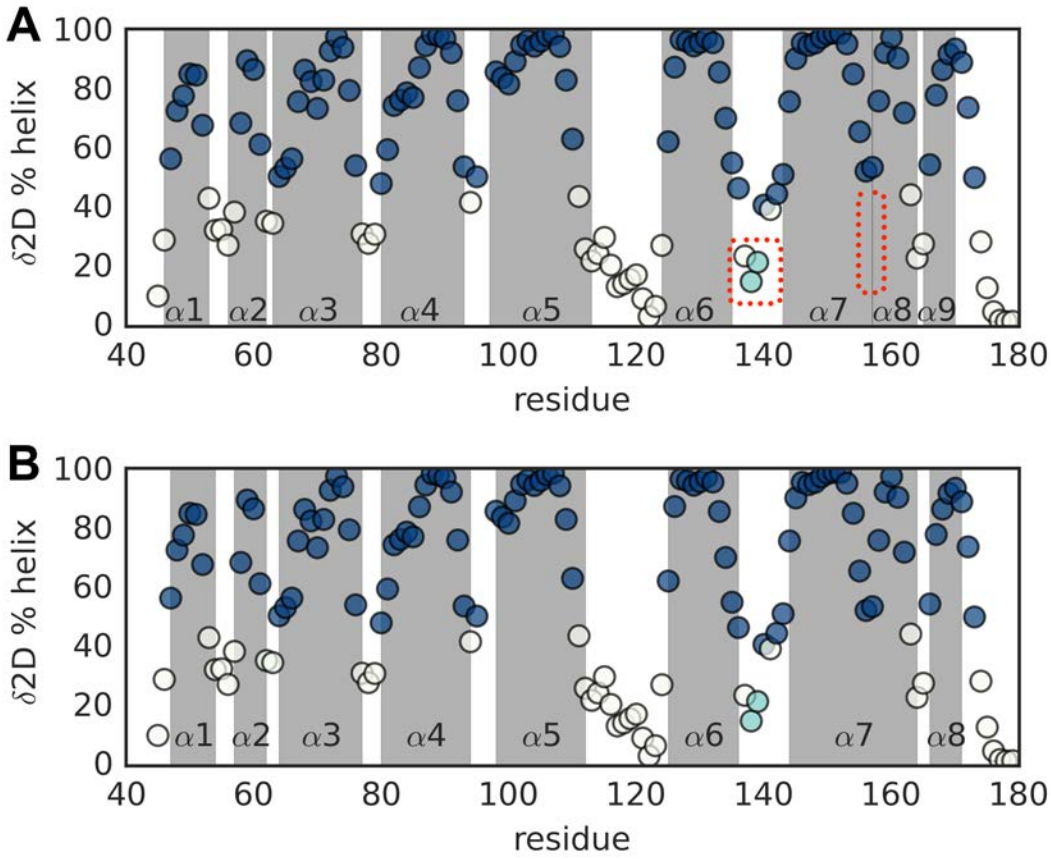

Figure 2.5.2.1 δ2D-predicted RGS8 secondary structure

A-B) Classification (helical: blue, extended-β: cyan, coil: white) and population probability of α helix (ordinate) for each residue (abscissa) based on RGS8 WT chemical shifts at pH 7, 25 C, predicted by δ2D algorithm(47). Grey spans indicate helical segments defined by A) PDB 2IHD and B) a hybrid consistent with δ2D predictions for both RGS4 and RGS8. Red dotted outlines indicate regions in (A) inconsistent with δ2D predictions.

2.6. Transferring assignments to other conditions

2.6.1. RGS4 mutants

We employed multiple variants of RGS4 and RGS8. In some cases with large separation from nearby peaks in the ^{1}H-^{15}N spectra, we heuristically propagated assignments from the unambiguously assigned WT spectra to mutants. In light of the fact that we did not perform the requisite experiments to unambiguously assign these peaks for the mutants, we erred on the side of caution when propagating assignments. The primary factors influencing our judgment were distance to nearest neighbor peaks in chemical shift space, chemical shift perturbation magnitude between WT and mutant spectra, and real-space Euclidian distance from the site of nearest mutation. When propagated assignments are used in future chapters, they will be noted.

2.6.2. Different pH samples

For hydrogen exchange experiments, a series of ^{1}H-^{15}N HSQCs was acquired over time for samples with different pHs. Similar heuristic assignment propagation was employed to transfer assignments to these spectra. In addition to chemical shift perturbation, propagation of these peak assignments included trends in chemical shift perturbation versus pH and the exchange rates of ambiguous peaks compared to real space Euclidian nearest neighbors.

Nevertheless, a small minority of peaks remained ambiguous even with these additional features and will be noted in Chapter 7.

Chapter 3. Effects of modification of RGS4 Cys148

3.1. Background

3.1.1. NMR as a structural indicator

We studied ligand binding via solution NMR spectroscopy to pursue our goal of elucidating the structural mechanism of RGS4 inhibition following binding to Cys148. As described in 2.1.1, chemical exchange between the native state and one or more non-native states (e.g. ligand bound) causes attenuation of peaks at the native chemical shift as the population is split. In the simplistic limiting case of two discrete states A and B, the rate of their interconversion, k_{ex}, is defined by Equation 3.1.1.1, their populations are defined by Equation 3.1.1.2, and the difference in their resonance frequencies, $\Delta\nu$, is defined by Equation 3.1.1.3:

$$A \underset{k_a}{\overset{k_b}{\rightleftharpoons}} B \quad \Rightarrow \quad k_{ex} = \frac{k_a}{k_a + k_b}$$

Equation 3.1.1.1

$$P_a = \frac{k_b}{k_a + k_b} \quad P_b = \frac{k_a}{k_a + k_b} \quad : \quad P_a + P_b = 1$$

Equation 3.1.1.2

$$\Delta\nu = \mid \nu_a - \nu_b \mid$$

Equation 3.1.1.3

Both k_{ex} and $\Delta\nu$ can be expressed in units of s^{-1}. The ratio of these quantities determines the behavior of peaks subject to chemical exchange, and though the ratio is a continuous variable, these behaviors are broadly classified into three exchange regimes, defined in Equation 3.1.1.4.

$$\begin{aligned} k_{ex} &\ll \mid \Delta\nu \mid \quad \Rightarrow \quad slow \\ k_{ex} &\approx \mid \Delta\nu \mid \quad \Rightarrow \quad intermediate \\ k_{ex} &\gg \mid \Delta\nu \mid \quad \Rightarrow \quad fast \end{aligned}$$

Equation 3.1.1.4

Nuclei in slow exchange show discrete peaks at ν_a and ν_b whose areas are proportional to the relative populations. Nuclei in fast exchange show a single peak at a the population-weighted average of ν_a and ν_b. Intermediate exchange is characterized by a single peak at ν_{obs} between ν_a and ν_b whose line width is broadened by interconversion between states during the acquisition window (~150 ms), and this exchange broadening effect sometimes renders these peaks undetectable. For systems characterized by more than two states, the relative populations of each state are further decreased, and deconvolution of populations becomes problematic. For more details, interested readers are referred to an excellent review by Kleckner and Foster(48).

3.1.2. RGS cysteine mutants

In order to isolate structural perturbations induced by ligands to just modification of Cys148, we used an RGS4 Cys-null variant (RGS4c), an RGS4 variant with just Cys148 in a null background (RGS4 c148C), a Cys-null RGS8 (RGS8c), and a variant of RGS8 with just Cys142 in a null background (RGS8 c142C). These mutations do not impair RGS4 GAP activity and have served as useful tools to delineate the effects of site-specific cysteine modification (35-37, 39, 49). Both the RGS4c and RGS8c variants exhibited folded HSQC spectra, and neither of the inhibitors tested had any effect (data not shown). This is in agreement with published findings showing that these inhibitors have no activity against RGS4 or RGS8 lacking all cysteines(41, 42).

3.2. Effects of CCG-4986 and CCG-50014 on RGS4 ^{1}H-^{15}N HSQC spectra

3.2.1. Effects of modification of Cys148 by CCG-4986

Stoichiometric treatment of the RGS4 c148C RGS domain with RGS4-specific inhibitor CCG-4986 produced a rapid (<30 minutes) and striking change in its HSQC spectrum (Figure 3.2.1.1 A). Most of backbone and side chain amide resonances (~90%) were attenuated to beneath the limit of detection in the CCG-4986-adducted spectrum (Figure 3.2.1.1 B).

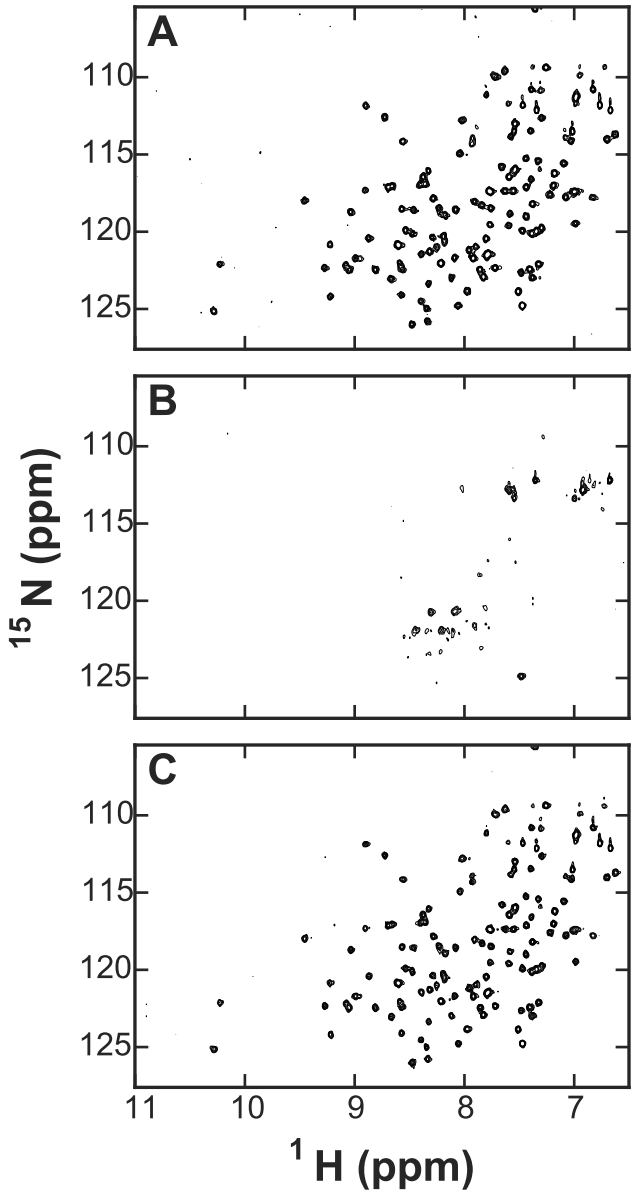

Figure 3.2.1.1 RGS4 c148C HSQC with CCG-4986

A) Folded, native HSQC of 150 μM RGS4 c148C at pH 7, 25 C (B) Partially unfolded HSQC of sample from (A) 60 minutes after addition of 450 μM (final) CCG-4986 (C) Folded, native HSQC acquired of sample from (B) 60 minutes after addition of 5 mM (final) DTT.

Large chemical shift perturbations from the native state and potential for ensuing overlap rendered the assignments of the remaining peaks (~10%) in the

CCG-4986-adducted spectrum ambiguous. Since as a covalent inhibitor, CCG-4986's k_{off} is effectively zero, excursions from the native state are not reversible, so the classic slow, intermediate, and fast regimes of chemical exchange are not applicable *per se*. However, covalent ligand binding is approximated by the slow exchange limit where P_{native} is zero. In the case of CCG-4986-bound RGS4, attenuation of native peaks is not matched by a commensurate intensification of sharply defined peaks as would be expected if there was a single inhibited state. Instead a relatively barren spectrum results, with the vast majority of peaks broadened to the noise background.

Disappearance of peaks in the HSQC spectrum is a characteristic of molten globules(50); for example, the molten globule form of human serum retinol binding protein, has only one visible peak(51). A reasonable hypothesis to explain this observation is thus that CCG-4986 binding to Cys148 causes RGS4 to partially unfold to a molten globule. Molten globules are partially unfolded proteins that retain generally similar secondary structure but lose the defined, rigid tertiary contacts of the native state and instead have considerable flexibility on the μs-ms time scale. We can test this hypothesis by measuring these characteristics (secondary and tertiary structure) in the native and inhibited states.

In contrast, unfolded random coil states of proteins under conditions such as at high concentrations of chaotropic agents like urea and guanidinium chloride, HSQC peaks tend to be sharp and show low dispersion. Full unfolding, then, would not produce the spectra observed upon reaction with CCG-4986. The conspicuous absence of even weak, broad peaks for most residues indicates that Cys148-modified RGS4 exists in an ensemble of two or more interconverting non-native states in intermediate exchange.

The departure from the native state was readily reversible upon addition of a reducing agent (DTT), reproducing the native spectrum *in situ* (Figure 3.2.1.1 C). This finding is consistent with previous reports that DTT-washout fully restores RGS $G\alpha_o$ binding activity by breaking the disulfide between adduct and thiol(41). Critically, for full reversibility of this phenomenon, the protein concentration was held below 200 μM because concentrations above this resulted in aggregation and precipitation of a portion of the protein-compound complex that was insoluble even with 10-fold molar excess of reducing agent.

3.2.2. Effects of modification of Cys148 by CCG-50014

Treatment of RGS4 c148C with the inhibitor CCG-50014 also denatured RGS4 and led to a ^{1}H-^{15}N HSQC spectrum (Figure 3.2.2.1 B) indistinguishable from that induced by CCG-4986 (Figure 3.2.1.1 B). This, too, was also reversible

upon addition of DTT, allowing RGS4 to return to its native state *in situ* (Figure 3.2.2.1 C).

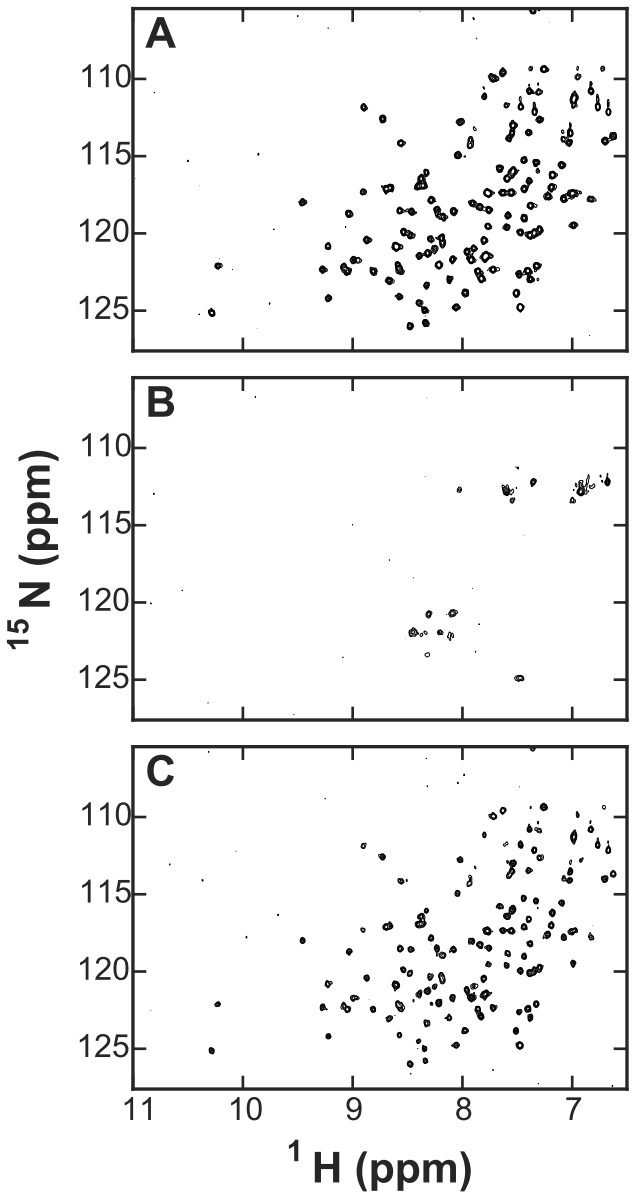

Figure 3.2.2.1 RGS4 c148C HSQC with CCG-50014

A-C) Same sequence as Figure 3.2.1.1 with 450 μM CCG-50014 instead of CCG-4986.

3.3. Identification of RGS4 C148Y as an adduct-mimetic mutant

3.3.1. Rationale for C148Y mutation

To rule out any potential effects of inhibitors binding to other sites on RGS4 in addition to modifying Cys148, we sought to create a mutant that mimics the effects of inhibitor binding to that site alone. The cysteine adduct of CCG-4986 is a disulfide-linked 4-nitrothiophenyl moiety(40) and the natural amino acid that most closely resembles it is tyrosine (Figure 3.3.1.1). We hypothesized that a C148Y mutant would recapitulate a CCG-4986-bound c148C RGS4.

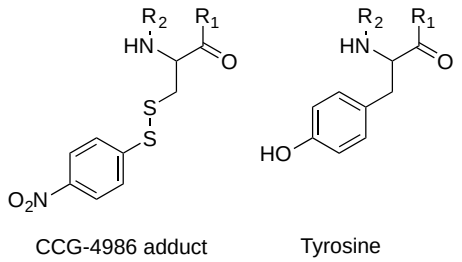

Figure 3.3.1.1 Structures of Cys adduct and Tyrosine

R_1 and R_2 represent the rest of the protein. The structure of the CCG-4986 adduct was confirmed by mass spectrometry(40)

3.3.2. Characterization of C148Y mutation

As we hypothesized, RGS4 c148Y was inhibited to a significant extent; it was 100-fold less potent in a GTPase acceleration assay (Figure 3.3.2.1). The c148Y variant also had an ^{1}H-^{15}N HSQC spectrum similar to those of CCG-4986-

and CCG-50014-treated c148C RGS4—though with more peaks detectable at an identical protein concentration, number of scans, and contour level.

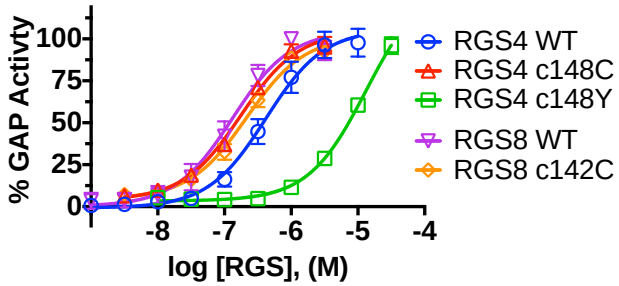

Figure 3.3.2.1 GAP activity of RGS4 c148Y

In vitro GAP activity of purified RGS domains for indicated variants. GAP activity was measured by enhanced phosphate liberation from GTP to GDP hydrolysis as described previously using colorimetric molybdate malachite green assay(49).

Probing the RGS4 c148Y HSQC with greater sensitivity revealed a subset of peaks whose distinct pattern of chemical shifts overlapped between RGS4 WT and RGS4 c148Y with only minor chemical shift perturbations: 54Glu, 57Lys, 59Trp, 76Phe, 85Ser, 178Thr, 179Asn (Figure 3.3.2.2). These all reside in the minor subdomain (α1-α3 and α8), among the furthest residues from Cys148. Several other peaks that occupied more crowded spectral regions, and so for whom a putative assignment was not possible without prior information, had plausible neighboring peaks in the minor subdomain as well and likely represent resonances from those residues in the inhibited ensemble. A subset of peaks in c148Y RGS4 also overlap with the residual peaks from CCG-4986 and CCG-50014-modified spectra.

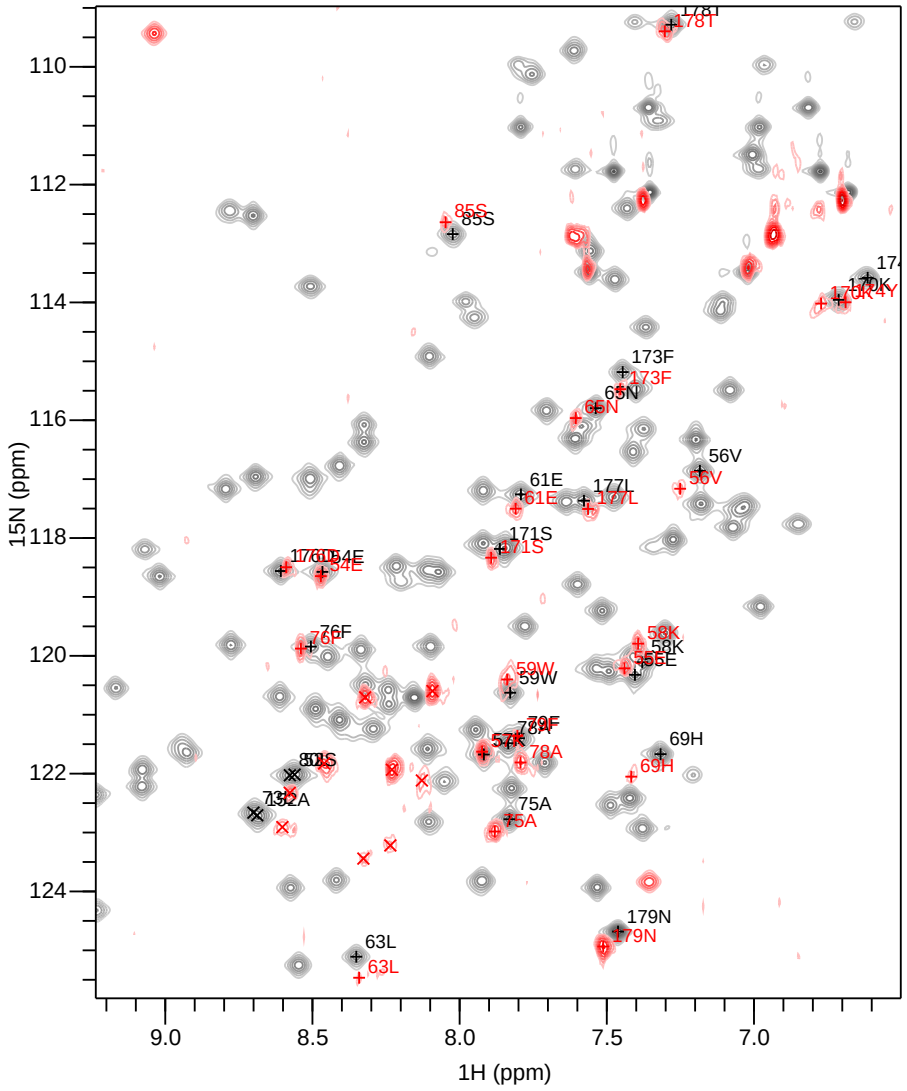

Figure 3.3.2.2 Overlapping peaks in RGS4 c148Y

Overlaid region of ^{1}H-^{15}N HSQCs from WT (grey) and c148Y (red) RGS4 RGS domains. Putative assignments, based on resonance overlap, for subset of peaks in c148Y spectrum labeled with red '+', residue number and 1-letter amino acid symbol. Unambiguous assignments of corresponding peaks from WT RGS4 spectrum indicated with black '+.' c148Y peaks with ambiguous assignments labeled with red 'x,' and peaks in WT RGS4 spectrum that could plausibly represent those ambiguous peaks labeled with black 'x.'

3.4. Discussion

3.4.1. Partially unfolded states of RGS4

The partially unfolded ensemble adopted by RGS4 upon modification or mutation of Cys148 undergoes exchange on an intermediate time scale. The inhibitor-bound ensembles rapidly refold to the native state upon removal of the adduct, suggesting either that RGS4 can refold spontaneously or that the partially unfolded states maintain significant structural elements similar to the native state. The tendency of inhibitor-bound RGS4 to aggregate above 200 μM is also consistent with the behavior of partially unfolded proteins whose hydrophobic interiors are transiently exposed to solvent and aggregate in a concentration-dependent fashion between 50-300 μM(52). Based on these characteristics, we hypothesize that the inhibited state adopted by RGS4 upon CCG-4986 binding to Cys148 is a molten globule (i.e. is not fully unfolded), and we will test specific aspects of this hypothesis in Chapter 4, namely that it retains significant secondary structure and that it is a relatively compact particle.

Both functionally and biophysically, c148Y recapitulates inhibitor-bound RGS4, though it appears that the mutant is a moderate form of the inhibitor-adducted RGS4 given the larger number of peaks in the HSQC that are not fully attenuated. This constitutively inactive mutant and others like it could help delineate the minimal steric and electrostatic requirements for inhibition at this

site and to assess the conformational ensemble of inhibited states. Moreover, it might also be a useful tool to study this mode of inhibition in reducing environments inimical to stable disulfide adducts, e.g. in cells.

Because none of the inhibited states of either RGS4 were sufficiently populated to yield peaks intense enough from which to extract even basic NMR observables like line shape or chemical shift for the vast majority of the protein, we turned to other approaches to obtain characterize the inhibited ensemble.

Chapter 4. Assessing 2° and 3° structure of inhibited RGS4

4.1. Background

4.1.1. Assessing properties of proteins through low resolution techniques

While NMR and X ray crystallography offer high resolution structural information at the residue and atomic level, these techniques are ill-equipped to describe certain dynamic systems that are not "well-behaved". Crystallography is more limited owing to its requirement for highly ordered proteins, but even NMR is limited in its ability to directly report on proteins in molten globule states without a manipulating the experimental system(53). Conversely, there are low-resolution techniques that can yield substantive information about the solution-averaged secondary and tertiary structure of proteins that do not conform to the sample requirements of NMR or X ray crystallography.

4.1.2. Tertiary structure from hydrodynamics

The size and shape of particles affects their diffusion through liquid media, and the gross size and shape of proteins is determined by their tertiary structure. The relationship between the length of a polypeptide and its excluded volume follow power law relationships(54). The radius of hydration, R_h, of a protein can be estimated by $4.75N^{0.29}$ Å for folded globular state and $2.21N^{0.58}$ Å in

an unfolded random coil, where *N* is the number of amino acids(55). For a protein the size of RGS4 (131 residues), this equates to a difference between 20 Å and 37 Å. For soluble proteins that fold into compact, globular native states, the degree of foldedness can be measured by changes in R_h. Two well-established methods used to measure this property are size exclusion chromatography (SEC) and dynamic light scattering (DLS), both of which are noninvasive (require no tags or markers) and nondestructive.

SEC separates macromolecules by frictional coefficient(56), which for similarly shaped molecules is proportional to molecular weight, based on their tendency to be excluded from pores in the chromatography medium. The preferred medium for separating proteins is typically dextran-coated agarose beads(57). In addition to its wide use as a polishing step in protein purification, SEC can be used analytically to assess a protein's hydrodynamic radius (also called Stokes radius, R_S), which is the radius of a hard sphere that would diffuse at the same rate as the protein.

In this capacity, SEC can be used to monitor native, unfolded, and molten globule states(58) in response to changes in sample conditions or detect the presence of dimers and higher order oligomers. DLS is a orthogonal detection method that also measures hydrodynamic radius by tracking microsecond

fluctuations in scattered light caused by the translational diffusion of proteins through a laser light source(59).

4.1.3. Secondary structure of proteins from circular dichroism

Circular dichroism (CD) is a classic method for determining protein secondary structure that relies upon characteristic absorbance patterns of polarized far UV light (<240 nm) by peptides with specific torsional angles. The CD spectra of a large collection of proteins with known secondary structure can be used as reference spectra to deconvolute an unknown CD spectrum into secondary structure components of the protein as a whole. The information content (i.e. number of components) of a CD spectrum depends on the wavelength range over which a protein's CD spectrum is obtained. The typical lower bound of the spectral range obtainable with widely available laboratory equipment using Xe arc light sources is 185 nm, and this range is sufficient to describe α helix, β sheet, and unordered regions(60). Estimates that are more reliable are possible using data at lower wavelengths (as low as 165 nm) from synchrotron light sources.

Various deconvolution algorithms have been developed, and while their performance varies depending on the reference set and wavelength available, most give reliable correlations to standards for α helical content greater than 0.9(61). Accuracy in predicting other components (β strand, turns, and unordered

regions) lags the accuracy of α helix because of the strong and characteristic signal produced by α helix. Fortunately for this study, RGS proteins are predominantly α helical, so standard methods will work well.

4.1.4. Thermal denaturation tracked by circular dichroism

In addition to extracting secondary structure components from CD spectra under static conditions, the analysis of CD spectra obtained as a function of temperature or denaturant can yield important information about thermodynamics and reversibility of protein unfolding and about folding intermediates throughout the unfolding process.

4.2. Hydrodynamic properties of RGS4

4.2.1. Hydrodynamic radius and shape

Partial unfolding of RGS4 either to a molten globule or random coil state would increase its R_S, but oligomerization would also increase the observed R_S. We compared size SEC elution profiles of c148Y and inhibitor-treated c148C RGS4 to untreated c148C and WT RGS4. A representative set of chromatograms shows that CCG-4986 and CCG-50014 increase the Stokes radius of c148C RGS4 significantly. We have converted the retention time into R_S based on a standard curve of proteins with known molecular weights and retention times.

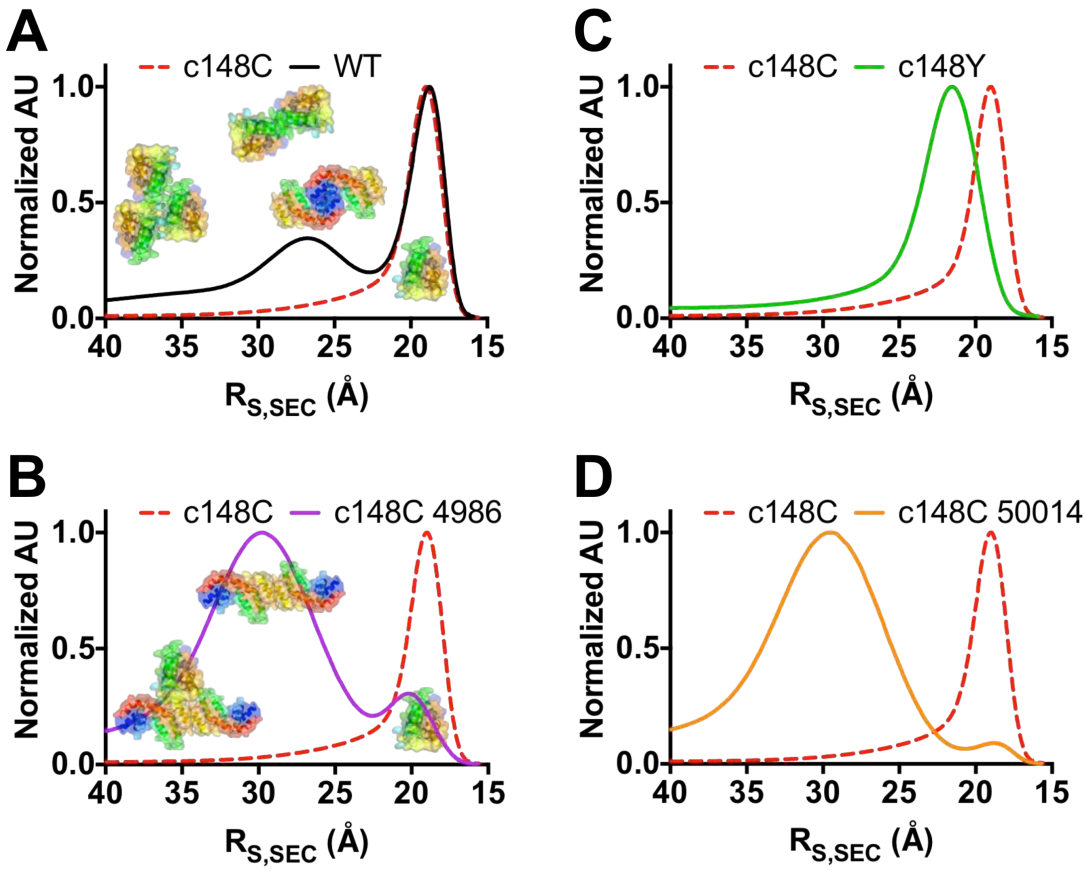

Figure 4.2.1.1 Representative SEC elution profiles

A-D) SEC retention time converted to Stokes radius by calibration with Bio-Rad molecular weight standards as described by le Maire and coworkers(62). A) RGS4 WT spontaneously forms dimer under running conditions (no reducing agent) required for stable inhibitor adducts. RGS4 c148C elutes as a monomer. B) CCG-4986 adducted RGS4 c148C shows bimodal distribution of sizes, though 90%+ of protein elutes at 30 Å, possibly a dimer or mostly unfolded monomer. C) RGS4 c148Y elutes as a monomer with slightly extended tertiary structure. D) CCG-50014-adducted RGS4 c148C elutes with a highly similar profile to CCG-4986-adducted RGS4 c148C (B), with 90%+ of protein eluting at a radius inconsistent with a folded, globular monomer.

The R_S of the predominant peak of WT (and the only peak of c148C) RGS4 match the predicted 20 Å for a globular protein with their molecular weight. Careful observers may note a minor peak at ~27 Å in the WT curve. This artifact of the running conditions is eliminated by inclusion of DTT in the buffer (data not shown). Since this peak appears only for WT RGS4 in the absence of reducing agent, we attribute it to a disulfide-linked dimer (presumably between Cys71 and/or Cys132, both of which are surface-exposed) that forms spontaneously in the absence of reducing agent. The observed R_S for this species is in agreement with the predicted R_S for a dimer (~26 Å). The presence of this additional background peak—and the fact that the inhibitors could potentially modify any combination of Cys 95, Cys132, and Cys148—prevented analysis of inhibitor-treated WT RGS4 because the adducts are not stable in the presence of DTT or other reducing agents.

The R_S of the predominant peak in both CCG-4986 and CCG-50014-treated samples is 30 Å, slightly larger than the WT dimer, though smaller than would be expected for a random coil (37 Å). A minor species is also observed at the size of monomeric RGS4, and represents a pool of either monomeric, modified RGS4 or residual unmodified protein. Intermediate between these extremes, c148Y showed an increased hydrodynamic radius of 22 Å. This suggests that c148Y does not either unfold to the extent that inhibitor-bound c148C RGS4 does or that

it does not dimerize. We collected these fractions for analysis by CD to determine whether their properties were consistent with either a random coil or dimer of folded proteins (see below).

We analyzed elution fractions from SEC by DLS to monitor the sample to ensure that they remained consistent during CD measurements. In general, there was very good agreement between R_S determined by SEC and DLS (Table 4.2.1.1).

RGS4 variant	$R_{S,DLS}$ (Å)	$R_{S,SEC}$ (Å)
WT	19.4 +/- 0.4	19.5 +/- 0.2
c148C	19.4 +/- 0.6	19.8 +/- 0.1
c148Y	21.1 +/- 0.2	21.4 +/- 0.2

Table 4.2.1.1 Hydrodynamic properties of RGS4

Correspondence of Stokes radius measured by SEC and DLS for non-aggregating RGS4 variants. $R_{s,DLS}$ measurements acquired on samples from highest concentration elution fraction from SEC UV curve (see Figure 4.2.1.1). Measurements represent mean and standard deviation from independent experiments and samples.

Samples of CCG-4986 and CCG-50014-treated c148C RGS4 spontaneously formed small populations (<1% of mass) of strongly scattering (>30% of total scattered intensity) aggregates that precluded DLS measurements.

4.3. RGS4 secondary structure from circular dichroism

4.3.1. Measurement of changes in CD spectra

We hypothesized that RGS4 inhibited through modification of Cys148 adopts a molten globule state, that is, a partially folded state with a high degree of residual secondary structure with a dynamic hydrophobic core and unstable tertiary contacts. We used CD to determine how inhibitor-induced perturbations in RGS4 effect secondary structure. The classic CD signature of α helix is the presence of strong negative absorption bands at 208 and 222 nm with a positive band at 192 nm. As expected based on the known structures of RGS4 and secondary structure predictions from NMR chemical shifts measured for WT RGS4, WT (not shown) and c148C RGS4 have a high degree of α helical character based on the far UV (185-250 nm) CD signal (Figure 4.3.1.1).

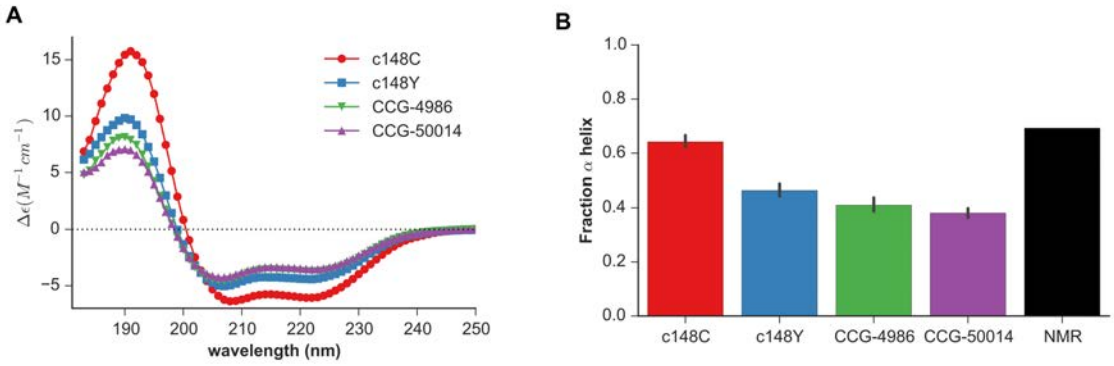

Figure 4.3.1.1 CD spectra and deconvoluted helicity of RGS4

A) Representative far-UV CD spectra acquired on elution fractions from SEC (Figure 4.2.1.1) showing decreased helical character of RGS4 inhibited by small-molecule covalent modification or mutation. B) Fractional helical content predicted by deconvolution of spectra from (A) and comparison to calculated helical content from NMR chemical shifts (Figure 2.5.1.1).

We deconvoluted these spectra to get a quantitative measure of the secondary structure changes. Since the secondary structure of RGS4 is predominantly α-helical (as well as loops and coil connecting the helices), we have summarized changes in secondary structure caused by mutation and inhibitor binding as reduction in the proportion of α-helix. RGS4 is the predicted to be 62% helical, and experimentally observed. CCG-50014 and CCG-4986 adducted c148C RGS4 helicity was (38% and 41%, respectively), and c148Y RGS4 helicity was 46%, all significantly lower than untreated c148C RGS4 ($p<0.0001$, ANOVA with Sidak multiple comparisons correction).

4.4. Thermal denaturation of RGS4 measured by circular dichroism

4.4.1. Residual secondary structure in inhibited RGS4

Further support for the hypothesis that RGS4 inhibited through modification of Cys148 adopts the partially folded ensemble of a molten globule comes from thermal denaturation experiments tracked by CD. We measured the far-UV CD spectrum of samples of RGS4 c148C, c148Y, and c148C bound to both CCG-4986 and CCG-50014—all purified by SEC prior to the experiments. Representative spectra for these unfolding curves are shown in (Figure 4.4.1.1), and the spectrum of the sample after cooling back to the initial temperature is shown as an overlay in black. Each sample undergoes a melting transition whereby helical character (negative intensity at 208 nm and 222 nm) is lost, showing that the inhibited states retain significant secondary structure and can be further unfolded. Refolding for RGS4 c148C and c148Y did not exceed 50% of initial secondary structure, and was below 25% for inhibitor-bound RGS4.

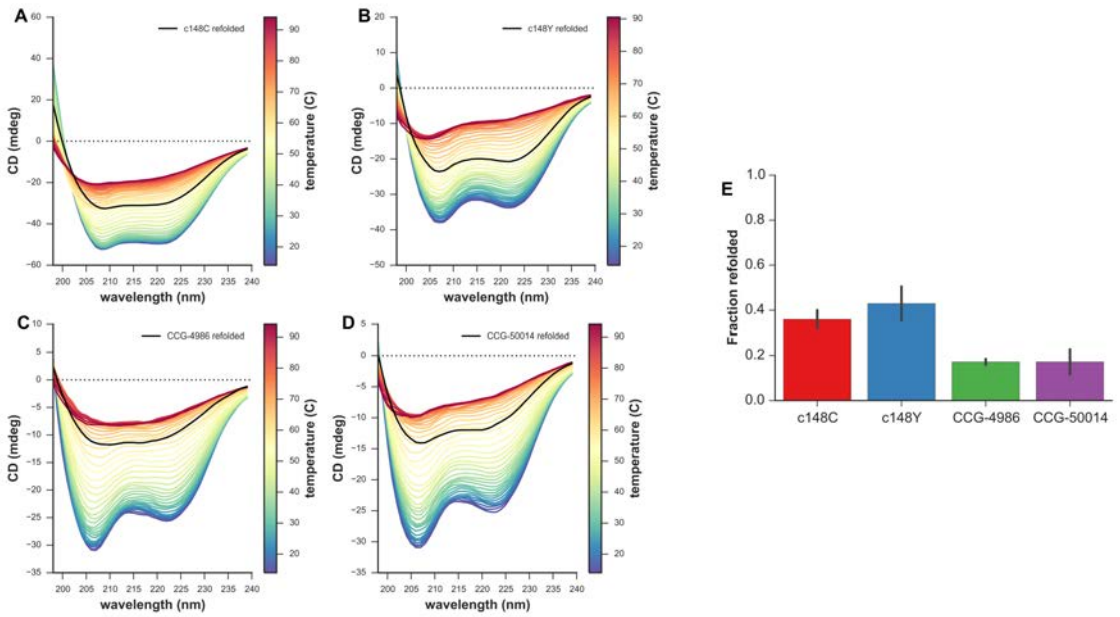

Figure 4.4.1.1 Representative melting CD spectra

A-D) Far-UV CD spectra (colored blue->yellow->red from low to high temperature) acquired as a function of temperature showing inhibited RGS4 contains significant residual secondary structure that undergoes unfolding transition in response to thermal denaturation. Following melting experiment, samples were cooled back to 15 C to assess ability to refold native secondary structure (black lines). E) Average refolding from independent melting/refolding experiments.

In addition to a leftward shift in the melting temperatures indicative of destabilization for the inhibited states, they also show a significant decrease in the sigmoid character in their thermal denaturation curves (Figure 4.4.1.2), suggestive of a partial loss of cooperativity in the unfolding process characteristic of molten globules(63, 64).

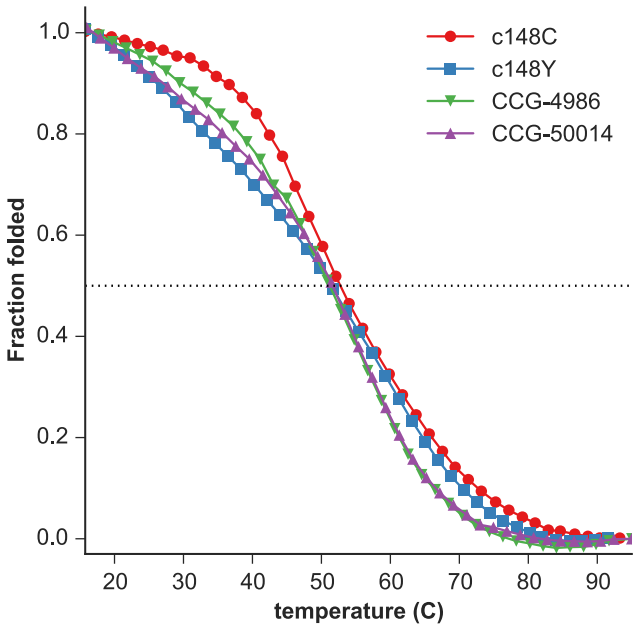

Figure 4.4.1.2 Average unfolding curves

Average of fraction folded for each sample as a function of temperature. Fraction folded calculated by linear combination of fully folded (15 C) and fully unfolded (95 C) CD spectra at each temperature.

4.5. Discussion

4.5.1. Residual structure in the inhibited ensemble

The high degree of residual secondary structure indicates that the large particles detected in inhibitor-bound SEC are dimers of molten globules, not random coil unfolded state. Association of transient exposed hydrophobic patches near Cys148 would explain both the dimerization as well as the formation of higher order oligomeric species (aggregates) detected both by DLS and during NMR experiments.

The hydrodynamic data also suggests that the packing of helices around the inhibitor (or mutated side chain in the case of c148Y) in individual monomers is disrupted such that the monomer explores a broad ensemble of extended conformational states that produce a significantly larger average hydrodynamic radius than that of the native structure. This larger average size is still, for c148Y at least, consistent with a monomeric molten globule and not a random coil unfolded state. Evidence from refolding of thermally denatured RGS4 shows that it cannot spontaneously refold, so the recovery of the native HSQC spectrum following addition of reducing agent to inhibited RGS4 indicates that even though the inhibitor-bound RGS4 appears to be more highly destabilized than c148Y, it is not unfolded to such an extent as to be beyond rescue.

Chapter 5. Characterization of RGS8 following exposure to inhibitors

5.1. Background

5.1.1. RGS4 inhibitor specificity

The determinants of RGS4 inhibitor specificity are unclear. The cysteines to which the molecules bind (Cys95, Cys148) are conserved in many RGS proteins, but the sensitivity of these proteins to inhibition varies over four orders of magnitude. RGS8 is one of the RGS proteins that is least sensitive to cysteine modifiers but retains a high degree of homology in the RGS domain. We sought to study the effects RGS4 inhibitors have on RGS8 to gather evidence about the source of this sensitivity.

5.2. Characterization of inhibitor effects on RGS8 ^{1}H-^{15}N HSQC spectra

5.2.1. CCG-4986

In contrast to the rapid molten globule transition in RGS4 c148C, the HSQC spectrum of RGS8 c142C showed no denaturation when treated with CCG-4986 under conditions identical to those used for RGS4, though there was a small, uniform reduction in native-state peak intensity (Figure 5.2.1.1)

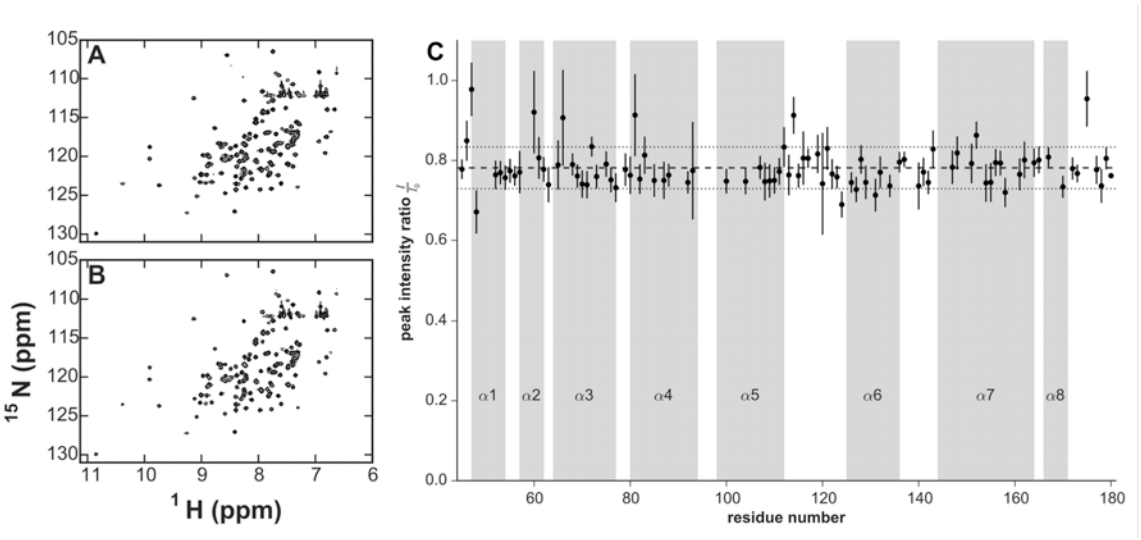

Figure 5.2.1.1 CCG-4986 effect on RGS8 c142C

A) ^{1}H-^{15}N HSQC of 150 μM RGS8 c142C showing folded, native spectra. B) Sample from (A) 60 minutes after exposure to 450 μM CCG-4986. C) Quantitation of peak attenuation (intensity of B/A) for peaks whose assignments were transferred from RGS8 WT spectra unambiguously (no overlapping peaks and no/minor chemical shift difference between mutant and WT).

That the structure of RGS8 is not perturbed to the same extent as RGS4 is not necessarily surprising since in vitro G*α* binding assays have shown that RGS8 is not sensitive to inhibition from CCG-4986(35).

5.2.2. CCG-50014

Since CCG-50014 can inhibit RGS8 in vitro, though with 300-fold lower potency than against RGS4(38), we sought to characterize differences between CCG-50014-bound and CCG-4986-bound RGS8. HSQCs immediately after adding CCG-50014 showed no change (data not shown), but incubating RGS8 with inhibitors for 6 hours before acquiring an HSQC showed that CCG-50014

reduced HSQC peak intensities more than CCG-4986 (60% and 10% attenuation, respectively), though both did so uniformly across the protein sequence (Figure 5.2.2.1). We did not detect any nascent peaks from an inhibited state for either inhibitor. Longer incubations (overnight) eventually led to visual precipitation of protein from solution.

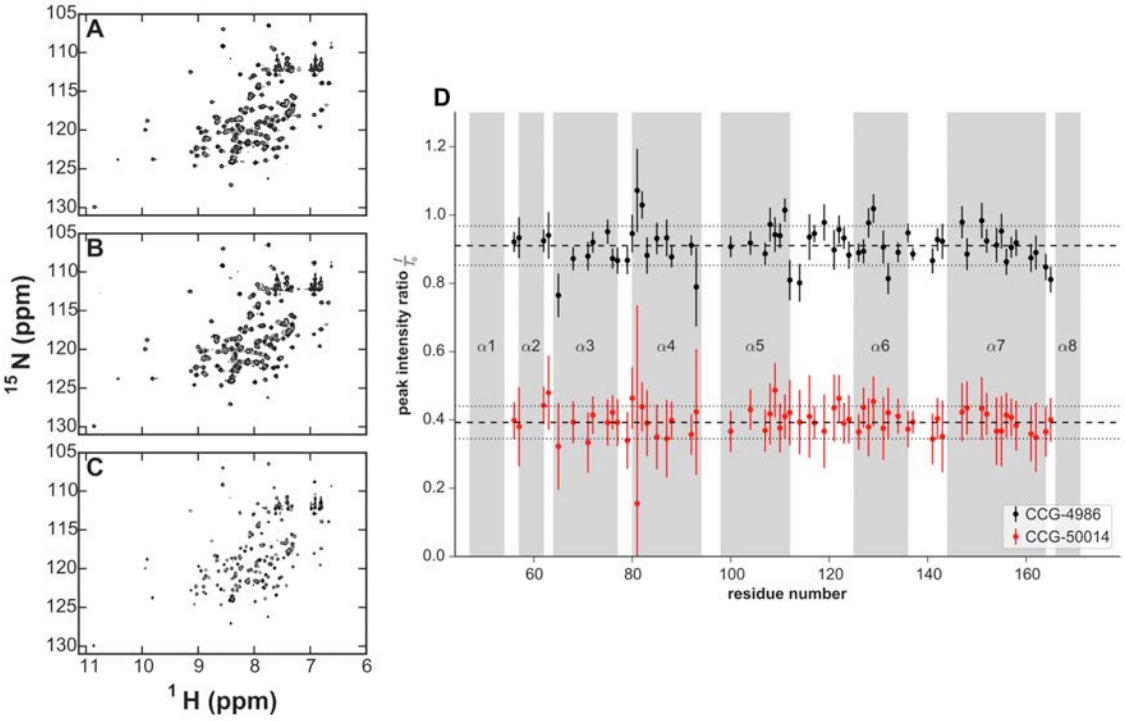

Figure 5.2.2.1 CCG-50014 effect on RGS8 c142C

A) HSQC of 150 μM RGS8 c142C. B) HSQC of 150 μM RGS8 c142C + 450 μM CCG-4986 after 6 hour incubation. C) HSQC of 150 μM RGS8 c142C + 450 μM CCG-50014 after 6 hour incubation. D) Attenuation of peaks in B and C compared to A

5.3. Attempt to make adduct mimetic of RGS8

5.3.1. Failure to fold

Since the RGS4 C148Y mutant shed valuable insight onto inhibited states of RGS4, we hypothesized that making a c142Y mutation would force RGS8 into an inhibited structure or ensemble as well. The c142Y RGS8 mutant was robustly expressed in *E. coli*, but unexpectedly, the protein was exclusively found in insoluble inclusion bodies. All attempts to increase the proportion of folded c142Y by modifying expression conditions (e.g. reducing induction temperature) or refold c142Y from inclusion bodies were unsuccessful (data not shown). We attempted both on-column and stepwise dialysis refolding from urea and guanidinum chloride.

5.4. Discussion

5.4.1. Unfolding response to inhibitor

The precipitation of CCG-50014-bound RGS8 after long incubations and the failure of RGS8 c142Y to fold indicates that the structure of RGS8 is no better able to accommodate modification of Cys142 than RGS4 is able for Cys148. In fact, we were unable to detect any soluble, non-aggregated, CCG-50014-modified RGS8. SEC of the incubated mixture showed a time-dependent decrease in UV absorbance but no change in retention. At the same time, after repeated

injections, aggregated protein caught on the pre-column filter led to increased back-pressure. Blazer and coworkers noted the same phenomenon when performing SEC on this mutant(38).

One of the major differences we noted during our studies was that RGS4 showed rapid changes in response to inhibitor addition. Within the time it took to acquire even a 1D HSQC trace, the effects of either CCG-50014 or CCG-4986 on the ^{1}H dispersion was noticeable. Conversely, RGS8 showed little effect without extended incubation. It is possible that the two proteins undergo a similar transition to partially unfolded ensembles, but that the rate of reaction with the covalent modifiers differs. We hypothesized that differences in cysteine accessibility could account for differential sensitivity to covalent inhibitors, so we sought to quantify differences in reactivity and accessibility.

Chapter 6. Evaluation of cysteine accessibility in RGS4 and RGS8

6.1. Background

6.1.1. Reactivity of buried groups

How does Cys148 come to interact with either CCG-4986 or CCG-50014 in the first place? According to the native state as reported in every crystal structure of an RGS protein, Cys148 in RGS4 and Cys142 in RGS8 are buried in the center of the α4-α7 helix bundle. The problem is analogous to the problem of native state hydrogen exchange (HX). The prevailing model for HX is described by the Linderstrøm-Lang equation(65). According to this model, a protected group (e.g. a backbone amide NH) can exchange with solvent only when it is exposed to the solvent through a conformational change. The same relationship also holds for reactive cysteine side chains, in what is termed sulfhydryl exchange (SX)(66).

In either case, the nascent exposed group reacts according to its intrinsic reactivity (k_{int}). Conformational fluctuations are stochastic processes that are described by an opening (k_{op}) and closing rate (k_{cl}). The model, using a thiol as the reactive group, is shown in Equation 6.1.1.1.

$$Protected - SH \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} Accessible - SH \overset{k_{int}}{\rightarrow} S - Labeled$$

Equation 6.1.1.1

The observed exchange rate (k_{obs}) is either EX1 limited or EX2 limited as described in Equation 6.1.1.2, in which case k_{obs} provides either thermodynamic information, reflecting the fraction of molecules in the open state (EX2) or kinetic information, reflecting the rate of structural opening (EX1).

$$k_{int} \ll k_{cl} \quad \Rightarrow \quad EX2 \qquad k_{int} \gg k_{cl} \quad \Rightarrow \quad EX1$$

Equation 6.1.1.2

EX2 conditions usually hold in a native ensemble, where by definition the equilibrium favors the closed conformation. However, this assumption must be tested experimentally in order to draw conclusions about whether k_{obs} is reporting on either thermodynamics or kinetics of opening.

6.1.2. Size of opening fluctuation

Protected groups in proteins that undergo reaction with solvent (or solutes) can be exposed to bulk solvent by transient local or global unfolding events. Because these experiments were carried out under native conditions (neutral pH, room temperature), and there is no evidence for global unfolding for either RGS4 or RGS8, it is unlikely that the conformational fluctuations detected by SX are global.

6.1.3. Intrinsic reactivity

A difference in the rate of cysteine modification at this site between RGS4 and RGS8 could be the result of two factors: protection or intrinsic reactivity. Thiol reactivity (k_{int}) with a thiol probe is dependent on the probe concentration, [P], and the biomolecular rate constant for the reaction of an unprotected cysteine (*k*). Values for *k* can be obtained from the reactivity of unfolded proteins(67).

6.1.4. Choice of thiol probe

Unlike HX, where external reactant is bulk water, SX requires that the experimenter choose an appropriate probe for the question to be addressed. We have chosen as our thiol probe the environmentally sensitive fluorophore N-(7-dimethylamino-4-methylcoumarin-3-yl))maleimide (DACM). DACM is only nominally fluorescent in solution and undergoes a large increase in quantum yield upon reaction with thiols(68). Maleimides are more selective for thiol nucleophiles over other protein nucleophiles like Lys, His and Tyr side chains than are iodoacetamides(69). DACM is also closer in size to CCG-4986 and CCG-50014 than large dyes like Texas Red. The structure of DACM compared to CCG-4986 and CCG-50014 is shown in Figure 6.1.4.1.

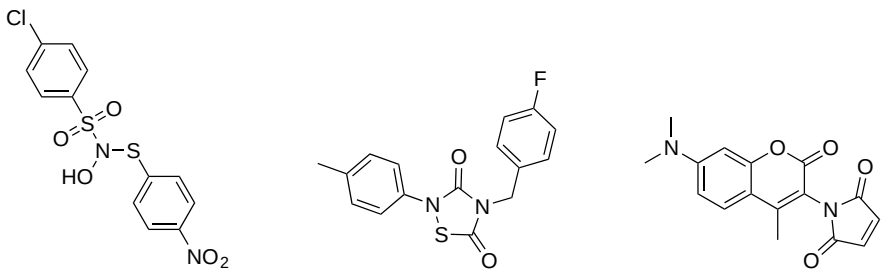

Figure 6.1.4.1 Structure of DACM compared to RGS4 inhibitors

6.2. Intrinsic SX in unfolded state

6.2.1. Equivalent reactivity in unfolded state

In order to make accurate comparisons about reactivity in the native ensemble, it is important to account for differences in the intrinsic reactivity of RGS4 Cys148 and RGS8 Cys142. We measured the intrinsic reactivity of Cys142 and Cys148 by denaturing RGS4 and RGS8 with 6M urea. Because the thiolate anion is the primary reactive form of a thiol, the reaction rate should show linear pH-dependence up to its pKa (8-9 for unprotected cysteine)(67). Representative progress curves for RGS4 Cys148 (A) and RGS8 Cys142 (B) and a linear fit of the natural log of rate constants to pH for both cysteines (C) are shown in Figure 6.2.1.1. These important controls show that any differences measured in the native state reflect differences in protection, not intrinsic reactivity.

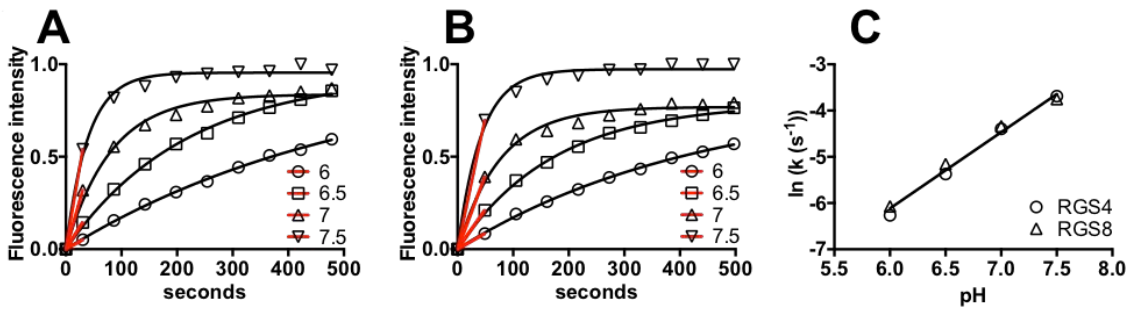

Figure 6.2.1.1 Intrinsic reactivity of unfolded RGS4 Cys148 and RGS8 Cys142

A-B) Progress curves of DACM reaction with cysteines in proteins unfolded with 6 M urea at pH 6, 6.5, 7, and 7.5. Red lines show linear fits of initial rates for (A) RGS4 c148C and (B) RGS8 c142C.

C) Relationship of unfolded rates and pH. Both RGS4 and RGS8 are fit by a single linear model (AICc)

6.3. Native SX

6.3.1. Differences in protection

Representative progress curves for the reactions of single-Cys mutants of RGS4 and RGS8 with DACM show the stark difference in native reactivity of RGS4 compared to RGS8 (Figure 6.3.1.1). Under native conditions, RGS4 Cys148 reacted with DACM robustly in a pH-dependent fashion, and its linear correspondence with intrinsic reactivity confirms EX2 behavior from pH 6 to 7.5. RGS8 Cys142 reactivity with DACM was below the limit of detection below pH 7, and significantly lower than RGS4 at both pH 7 and 7.5.

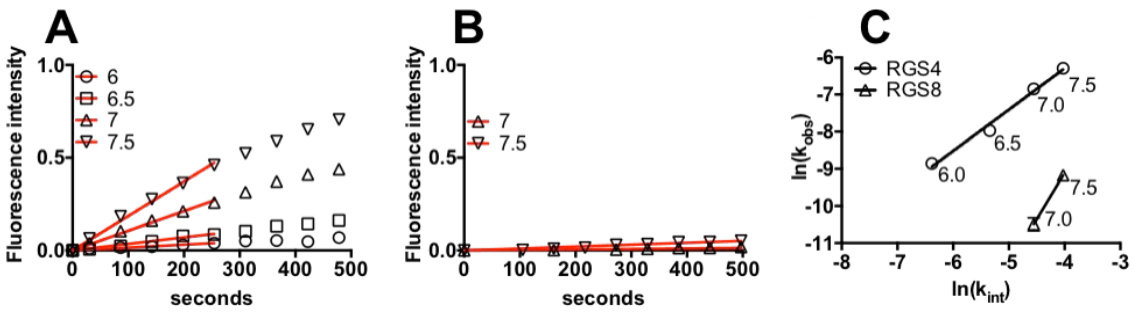

Figure 6.3.1.1 DACM progress curves

A-B) Progress curves of DACM reaction with cysteines in native proteins at pH 6, 6.5, 7, and 7.5. Red lines show linear fits of initial rates for (A) RGS4 c148C and (B) RGS8 c142C. C) Relationship of folded rates (k_{obs}) to unfolded rates.

These data indicate that the stability of the native states protecting (ΔG_{SX}) of RGS4 Cys148 and RGS8 Cys142 from exposure are 1.52±0.03 kcal mol^{-1} and 3.37±0.20 kcal mol^{-1}, respectively, a difference of 1.85 kcal mol^{-1} (Figure 6.3.1.2). This difference is highly significant, indicating that as a proportion of the total population at steady state, 1 in 14 RGS4 molecules are in an excited state conformation compared to only 1 in 327 for RGS8.

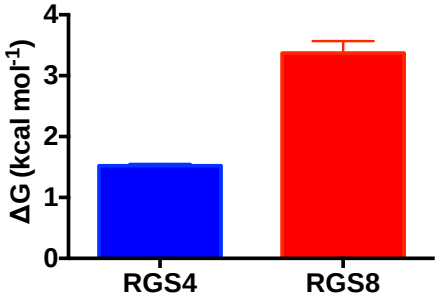

Figure 6.3.1.2 Native state SX free energy

ΔG calculated using equation analogous to Equation 7.1.3.2 for SX instead of HX.

6.4. Discussion

6.4.1. Differences in the native ensemble

Quantitative SX demonstrates that there is a fundamental difference in the protection provided by the α4-α7 bundle to the thiol buried at its center in the native ensemble of RGS4 compared to RGS8. The conformational change that exposes Cys148 must involve the concerted motion of side chains and backbone atoms from several residues. A model of the conformational change exposing nearby Cys95 proposed by Vashisth and coworkers(42) involved the concerted motion of helices α5 and α6. Such a motion would expose a host of other exchangeable groups along the interface between the helix pairs α4-α7 and α5-α6. Native HX would either corroborate this model or suggest other regions that may be participating in concerted motion to expose Cys148. Based on the large difference observed in the free energy from SX, the critical region should show an equally large difference in free energy from HX between RGS4 and RGS8.

Chapter 7. Measurement of slow backbone dynamics by hydrogen exchange

7.1. Background

7.1.1. Time-scales of protein dynamics

Proteins are inherently flexible, and they continuously sample conformations from a hierarchical set of spatiotemporal coordinates. A global motion, e.g. the folding of an entire domain, can take anywhere from milliseconds to hours(48), and in either state a local fluctuation, e.g. rotation of a bond in an individual side chain, occurs on the picosecond time scale. According to the energy landscape theory, the stochastic rate of change between two conformations is related to the height of the energy barrier separating them, and the conformations are populated according to the Boltzmann distribution of the relative free energies of the conformations(70).

In the native state ensemble of RGS4, there is at least one conformation that exposes Cys148 by the collective rearrangements of multiple residues in the α4-α7 helix bundle. This conformation is sufficiently populated to allow cysteine modification to occur, indicating that the relative energies of this excited state and the native, protected state are similar. The analogous excited state in RGS8 is much less populated, leading to a large degree of selectivity in cysteine

modification. We seek to determine the regions of RGS4 that account for these differences and to quantify the relative population of these excited states.

7.1.2. Methods of measuring backbone hydrogen exchange

Real time HX (RT-HX) is a classic, powerful technique for quantifying site-specific differences in backbone dynamics on the timescale of protein folding and solvent exchange(71). This technique monitors the exchange of highly protected backbone amide hydrogens for deuterons when a protein's H_2O-based solvent is replaced with D_2O-based solvent. For quickly exchanging amides, the limit of detection, in terms of the exchange rate, for RT-HX is a function of the minimum time needed to acquire a 2D HSQC spectrum and the signal-to-noise ratio. At the slow limit, there is no theoretical limit, though practical considerations define an upper bound correlating to the maximum total observed time and the signal-to-noise ratio.

A RT-HX experiment is begun by either dissolving lyophilized protein in D_2O or by diluting a concentrated H_2O sample with a large volume of D_2O. We found the former approach to work better for RGS proteins due to limitations in the maximum concentration achievable before aggregation occurred in the latter. Once dissolved, HX commences and a series of 2D HSQC spectra are acquired. From these spectra, exponential decay rates, k_{ex}, for each crosspeak are fit by

Equation 7.1.2.1, where t is time, I_0 is the initial peak intensity, and I(t) is the peak intensity at time t.

$$I(t) = I_0 e^{-k_{ex}t}$$

Equation 7.1.2.1

There are also techniques that measure considerably less protected amides using an alternative detection scheme. Instead of measuring signal intensity decrease as a function of hydrogen/deuteron exchange, zz-exchange spectroscopy can be used to quantify HX. A robust protocol for performing this technique is the Phase-Modulated CLEAN chemical EXchange clean (CLEANEX-PM)(72). In CLEANEX-PM, solvent exchange is quantified from the transfer of magnetized protons from bulk water to amides. The exchange rate, k_{ex}, is fit along with the relaxation rate constant during the spinlock, $R_{1A,app}$, as a function of the mixing time, τ_m, with Equation 7.1.2.2, where $R_{1b,app}$ is the relaxation rate constant for water, estimated to be $0.6 s^{-1}$, though the value for $R_{1b,app}$ has been shown to have little to no effect on k_{ex}(73). Fitted rates were then corrected for water saturation as in(72).

$$\frac{I}{I_0} = \frac{k_{ex}}{R_{1A,app} + k_{ex} - R_{1B,app}} \times e^{-(R_{1B,app})\tau_m} - e^{-(R_{1A,app}+k_{ex})\tau_m}$$

Equation 7.1.2.2

7.1.3. Extracting thermodynamic information from exchange rates

For structurally protected amides, HX is modeled as a steady state process analogous to Equation 6.1.1.1. Under EX2 conditions (see 6.1.1 for description), the equilibrium constant for open and closed forms is given by Equation 7.1.3.1.

$$k_{ex}^{EX2} = \frac{k_{op}k_{int}}{k_{cl}} = K_{op}k_{int}, \quad K_{op} = \frac{k_{ex}^{EX2}}{k_{int}}$$

Equation 7.1.3.1

Therefore, in order to extract thermodynamic information from native HX, the observed exchange rates must be divided by the intrinsic (or unprotected) rates, which are dependent on the sequence identity of the amino acid triplet around the exchanging amide and upon the temperature and pH of the solvent(74). Intrinsic rates have been determined empirically and are readily calculated using the Server Program for Hydrogen Exchange Rate Estimation (SPHERE)(75). The free energy of protection from exchange is calculated using Equation 7.1.3.2.

$$\Delta G_{HX} = -RTln(\frac{k_{ex}^{EX2}}{k_{int}})$$

Equation 7.1.3.2

7.2. First order data processing for HX

7.2.1. RT-HX peak fitting and rate extraction

In NMR spectra, the signal to noise ratio is proportional to sample concentration and the square root of the number of scans. All other things being equal, as the number of scans increases, the total time for each 2D spectrum increases, and the frequency with which a process, e.g. HX, can be sampled decreases. Therefore, there is a tradeoff between the minimum rate of exchange that can be experimentally determined (the laboratory time frame, or laboratory frame) and the precision of the measured intensity of any crosspeak in the series. We achieved a dead time (that is, the time between initiating exchange and starting the first 2D spectrum acquisition) of two minutes and an acquisition time of 4 minutes. We assigned the time for each spectrum as the midpoint of acquisition, so the first point in each RT-HX series was 4 minutes. Following the initial spectrum, spectra were acquired at 8, 12, 16, 20, 30, 40, 50, 60, 80, 100, 120, 150, 180, 210, and 240 minutes. For samples where more than 30% of peaks remained at 240 minutes, we acquired additional spectra at 270, 300, 330, and 360 minutes.

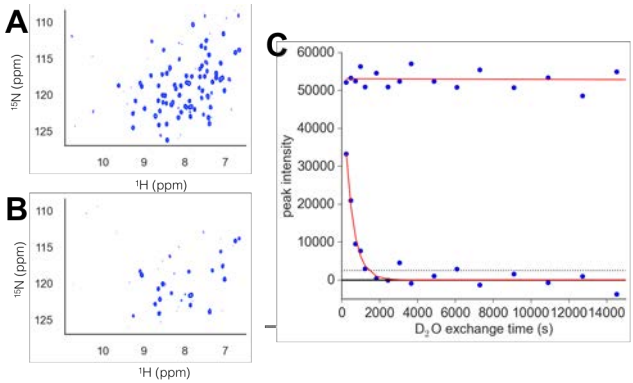

Figure 7.2.1.1 Peak intensity rate extraction

A) Representative HSQC 4 minutes after D_2O exchange during RT-HX time course.

B) Representative HSQC 240 minutes after D_2O exchange.

C) Extracted peak intensities (blue circles) and best-fit least squares regression (red solid lines) to Equation 7.1.2.1. for two assigned peaks over the time course of RT-HX. These examples are near the limits of detection (see below) of measurable rates. The dotted black line is the standard deviation of intensity of fully decayed peaks (also equal to the standard deviation of spectral regions without peaks).

7.2.2. CLEANEX-PM rate extraction

We fit exchange rates to buildup curves of peak intensity to Equation 7.1.2.2. A representative example of is shown in Figure 7.2.2.1.

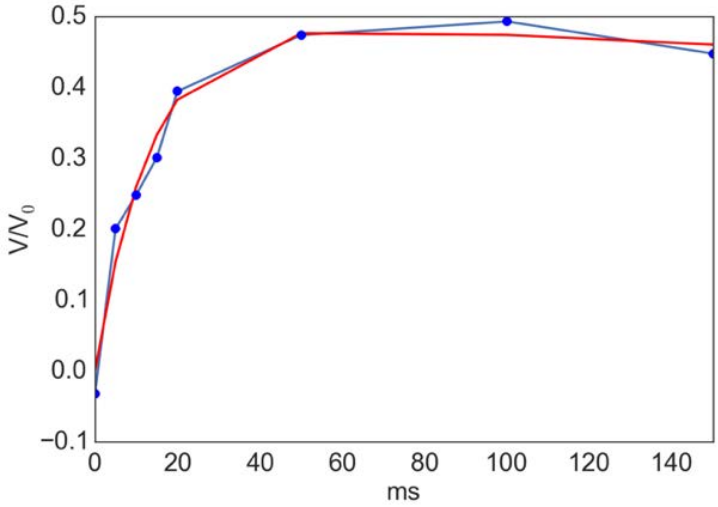

Figure 7.2.2.1 CLEANEX-PM rate extraction

Representative build-up curve from CLEANEX-PM experiment. Relative peak intensities compared to reference HSQC (blue circles) at indicated mixing times were fit to Equation 7.1.2.2. The best fit to this equation is shown as a solid red line.

7.2.3. Limits of detection

The level of noise in our spectra, defined as the standard deviation of a region containing no peaks, was 2.5×10^3 (arbitrary units). The distribution of peak heights measured at the first HSQC had a mean, μ, of 5.1×10^4 and a standard deviation, σ, of 2.48×10^4. With these sampling parameters, as the exchange rate approaches ~5×10^{-3} (s^{-1}), the peak intensity decays so quickly that by the time the first spectrum is complete, the peak has decayed by 80% of its initial amplitude. The intensity of even the most intense peaks (1.2×10^6) is below 2σ at the second time point. On the other hand, for peaks that approach the slow limit of detection, ~5×10^{-6} (s^{-1}), the decay in intensity at the last time point

approaches or is smaller than I_0-2σ. We have shown that the coefficient of variation (CV=σ/μ) in the fitted exchange rate from our data approaches unity as measured rates approach these limits (Figure 7.2.3.1). CV is a concise way to compare the uncertainty (σ) of fit to the fitted value (μ) over a wide range of μ, since CV is scaled and dimensionless. A CV of 0.1 or 0.01 means that the fitted value has an uncertainty of 10% or 1%, respectively.

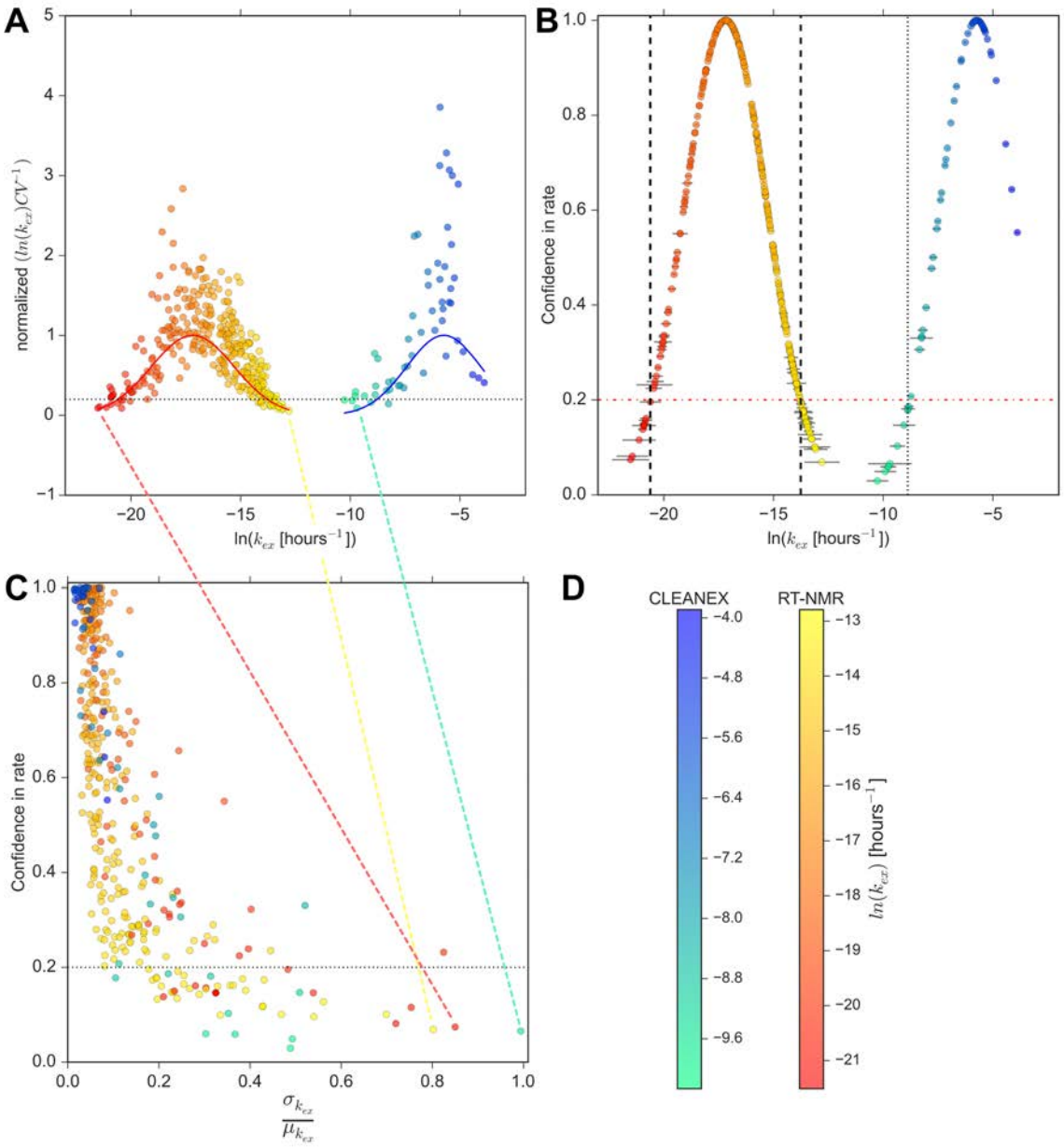

Figure 7.2.3.1 Uncertainty in rate and limits of detection for HX

Each circle shows the quality of the fit of an individual data curve (one residue at one pH)

A) Relationship of normalized CV^{-1} (μ, the natural log of the k_{ex} fit from data, divided by σ, the natural log of its standard deviation from the least squares fit—all divided by the max CV^{-1}), to k_{ex}. Larger values of this quantity are indicative of better fits. (continued next page)

(continued from previous page) CV^{-1} trends toward 0 (where μ and σ are equal, and thus the fit is a poor one) at the upper and lower extremes of the RT-NMR range (natural log of k_{ex} < -20 and > -14). This shows the limits of detection for this experiment given the acquisition parameters. Solid lines are scaled (max of 1) Gaussian functions fit to the data (see (B)). B) Confidence in rate (scaled Gaussian from (A)) as a function of k_{ex}. Horizontal error bars are σ of least squares fits. Vertical, black dashed lines are estimates of limits of detection, set where the confidence functions intersect confidence = 0.2. C) Relationship of confidence score to σ/μ (or CV^{-1}) showing asymptotic trend toward large fit uncertainty below confidence = 0.2. D) Colorbars indicating fitted k_{ex} of circles in A-C.

The distribution of the coefficient of variation within these bounds approximates a Gaussian function, so we fit these values to assign a confidence rating (from 0 to 1) for each measured rate to use as a weighting factor when incorporating rates from multiple experimental conditions (*vide infra*). A fast limit of detection was found for rates measured by CLEANEX-PM of 0.9 s^{-1} according to the same principle.

To visually represent residues for which we could not measure accurate decay rates because they were outside the limit of detection, we estimated their upper or lower bound. For peaks that decayed more quickly but still had intensity greater than 3σ in the initial spectrum, the rate and fitting error was estimated as the average of the highest 5^{th} percentile of measured rates below the fast limit of detection. The true exchange rates are higher, but we could not determine them under our sampling conditions. We did the same for peaks that

did not appreciably decay, assigning the lowest 5th percentile above the slow limit of detection to their decay rates as an upper bound. Their presence on the raw rate plot (below) merely signifies that peaks with unambiguous assignments were detected for those residues. The confidence score for these projected values is near zero though, so these rates are heavily down-weighted in thermodynamic calculations where values measured under other conditions for that residue were within the limits of detection. In cases where no measured k_{ex} was within the limits of detection, the calculated thermodynamic parameters should be considered an upper (or lower) bound on the true value.

7.3. Second order data processing

7.3.1. Positional effects on exchange rate

A plot of the natural log of k_{ex} at each measured position in RGS4 and RGS8 is shown in Figure 7.3.1.1. To facilitate direct comparison because the numbering of equivalent positions in the RGS domain of RGS4 and RGS8 is different, the residue numbers in the figure (and subsequent figures) have been re-indexed to the first conserved position in the RGS domain. To back-calculate the RGS-specific numbering, add 50 and 44 to RGS4 and RGS8, respectively (e.g. RGS4 Cys148 and RGS8 Cys142 are both shown on the X-axis as residue 98). The dashed lines indicate the conservative estimates of limits of detection (*vide supra*).

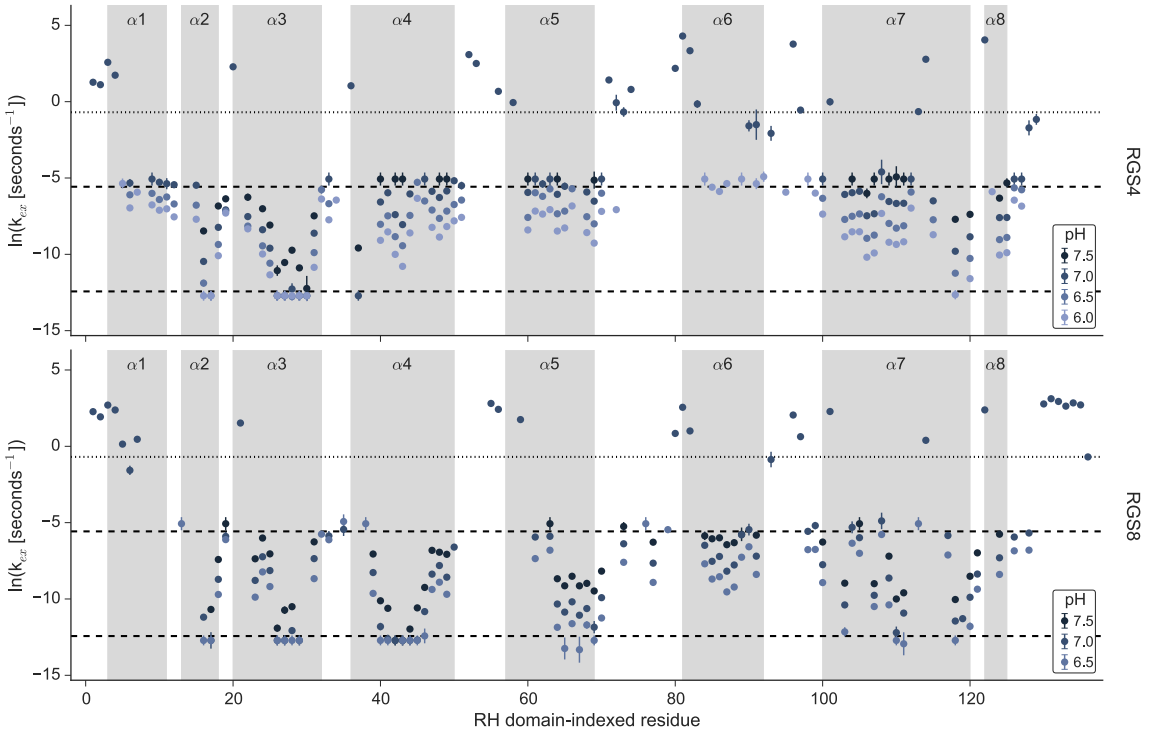

Figure 7.3.1.1 HX rates by residue

All measured backbone HX rates for RGS4 (top) and RGS8 (bottom). Each point has an associated error bar (standard deviation from least squares fitting), though some do not extend beyond the marker. Color encodes the pH of the measurement (uncorrected). Horizontal lines show estimates of limits of detection for CLEANEX-derived rates (dotted) and RT-NMR-derived rates (dashed). Grey spans indicate secondary structure elements described in Chapter 2.5)

At neutral pH (7), RGS4 and RGS8 both contained regions that were fully protected from exchange within our sampling window for RT-HX. These regions spanned the α2-α4 and α7 bundle that joins the two subdomains. This area constitutes the bulk of the hydrophobic core of the RGS domain, and its lack of exchange indicates that these proteins do not undergo significant transient global

unfolding under these conditions, i.e. the conditions are appropriate for native HX.

There were also regions where all of the peaks had exchanged before the initial spectrum at pH 7, notably in α6 in RGS4, in loop regions, and in the unstructured C-terminus of RGS8. In order to maximize the coverage of residues for which we could draw conclusions, we carried out RT-HX at pH 7.5 and 6.5 for both RGS4 and RGS8. We also obtained a dataset at pH 6 (RGS4 only) where the fast rates (e.g. RGS4 α6 residues) were slowed to the laboratory frame.

7.3.2. Evidence for EX2 regime

We performed these measurements under conditions favorable for the native state and have strong evidence from orthogonal techniques (e.g. CD) to that effect. Nonetheless, we sought to validate the assumption that the observed k_{ex} is EX2-limited. As expressed in Equation 7.1.3.1, a characteristic of the HX in the EX2 limit is a linear relationship between k_{int} and k_{ex}. Since k_{int} varies linearly with pH and we measured k_{ex} at multiple pHs, validation of EX2 conditions can be obtained by inspection of a plot of k_{ex} vs k_{int} for each residue—where a slope of 1 indicates the residue is in EX2(76). We found this relationship to hold for nearly all residues with at least 2 measured values for k_{ex} within the limits of detection (median slopes 1.03 and 1.14 for RGS4 and RGS8, respectively) (Figure 7.3.2.1).

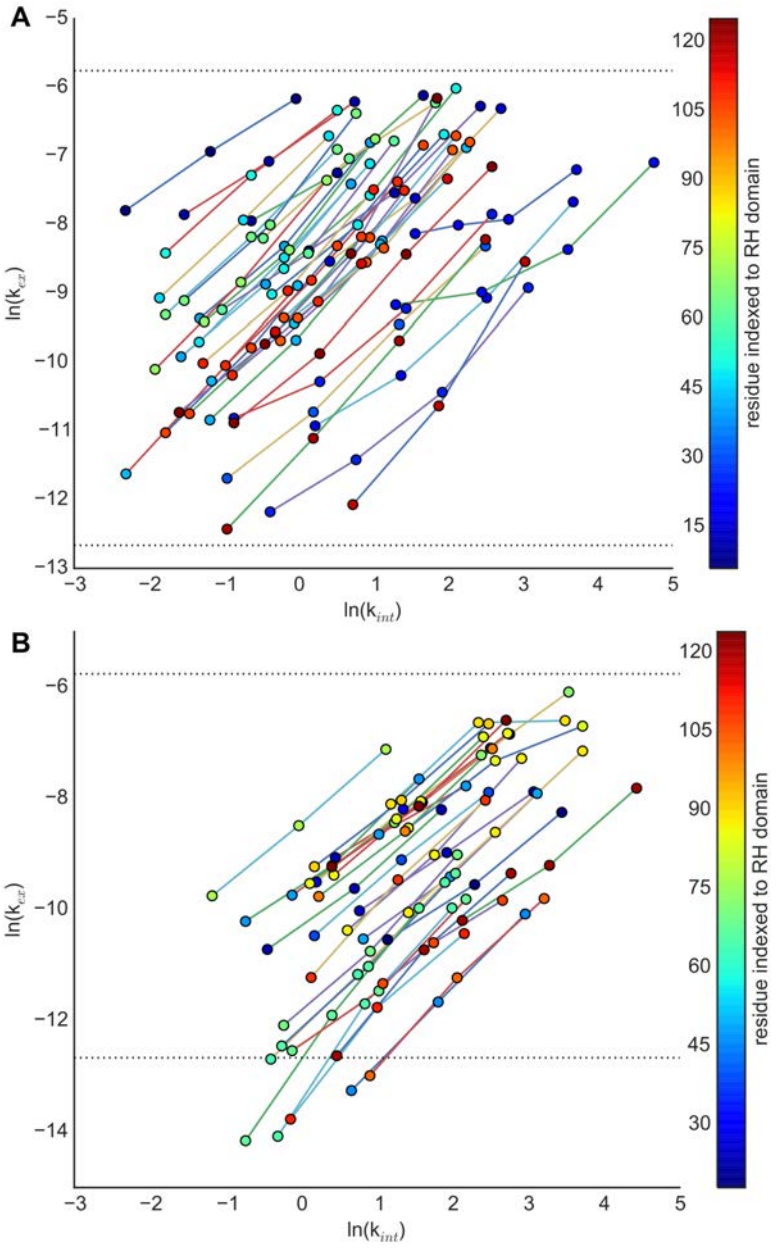

Figure 7.3.2.1 EX2 validation for RGS4 and RGS8

A-B) Log-log plot of k_{ex} [s^{-1}] vs. k_{int} [s^{-1}] for (A) RGS4 and (B) RGS8 from RT-HX. Each circle is the rate for a given condition (one residue, one pH) and is colored by the position in the RGS domain analogous to chainbow coloring. To guide the eye, solid lines connect data points of a given residue with measurements at multiple pHs (not necessarily the same color as circles). Horizontal dotted lines indicate estimated limits of detection.

7.3.3. Confidence-weighted average for equilibrium constant

We calculated the apparent equilibrium constant, K_{op}, according to Equation 7.1.3.1, for each independent measurement of k_{ex}. Values for K_{op} calculated from conditions in which k_{ex} could be measured with higher confidence (i.e. farther from the limits of detection) are in turn more robust estimators of K_{op}. Where we had multiple measurements of K_{op}, we calculated the weighted average of lnK_{op} using Equation 7.3.3.1, where c is the confidence associated with each measurement, μ is the weighted arithmetic mean, and σ is the weighted propagated uncertainty, and subscripts index the repeated measures. A plot of the correlation of the simple arithmetic mean and the weighted arithmetic mean is shown in Figure 7.3.3.1. In general, weighting serves to decrease uncertainty of residues with multiple measurements because the propagated uncertainty is also weighted toward those with smaller fitting uncertainties.

$$w_i = \frac{c_i}{\sum_i^n c_i}, \quad \mu = \sum_i^n w_i (lnK_{op})_i, \quad \sigma = \sum_i^n \sqrt{w_i^2 \sigma_i^2}$$

Equation 7.3.3.1

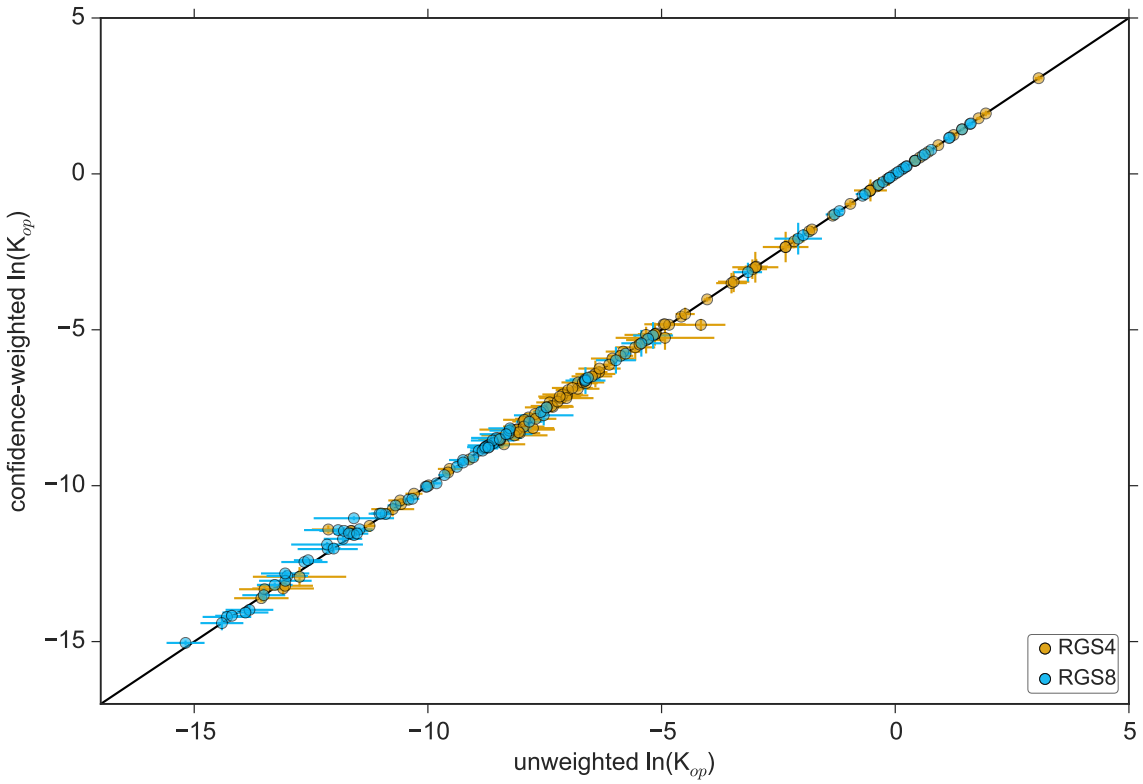

Figure 7.3.3.1 Weighted vs. simple mean of $\ln K_{op}$

Log-Log plot of weighted mean and simple arithmetic mean (c_i is 1 for all measurements). Each circle represents a unique condition (one residue, one pH), and the vertical and horizontal error bars represent σ calculated from Equation 7.3.3.1. The diagonal black line has a slope of 1 and indicates perfect correspondence.

7.3.4. Benefits and assumptions of weighting

Weighting serves two purposes: 1) Weighting allows us to directly compare the stability of residues that exchange on different laboratory frames (e.g. RGS4 α6 exchanges too fast to observe above pH 6 while α3 is slower than the slow limit below pH 7.5). 2) Weighting is a systematic, unbiased way to give priority to measurements in the optimal laboratory frame (i.e. $-13 < ln k_{ex}\ [s^{-1}] < -6$)

for residues that have measurements inside and outside these optimal limits of detection. Justification for this mathematical treatment follows from the same assumptions governing EX2-limited kinetics, which we have already validated for both RGS4 and RGS8, namely that k_{ex} has a 1:1 correspondence with k_{int}, and thus lnK_{op} (and thus ΔG_{HX}) is constant across the measured conditions. This method avoids underestimating ΔG_{HX} for residues highly stable residues, see

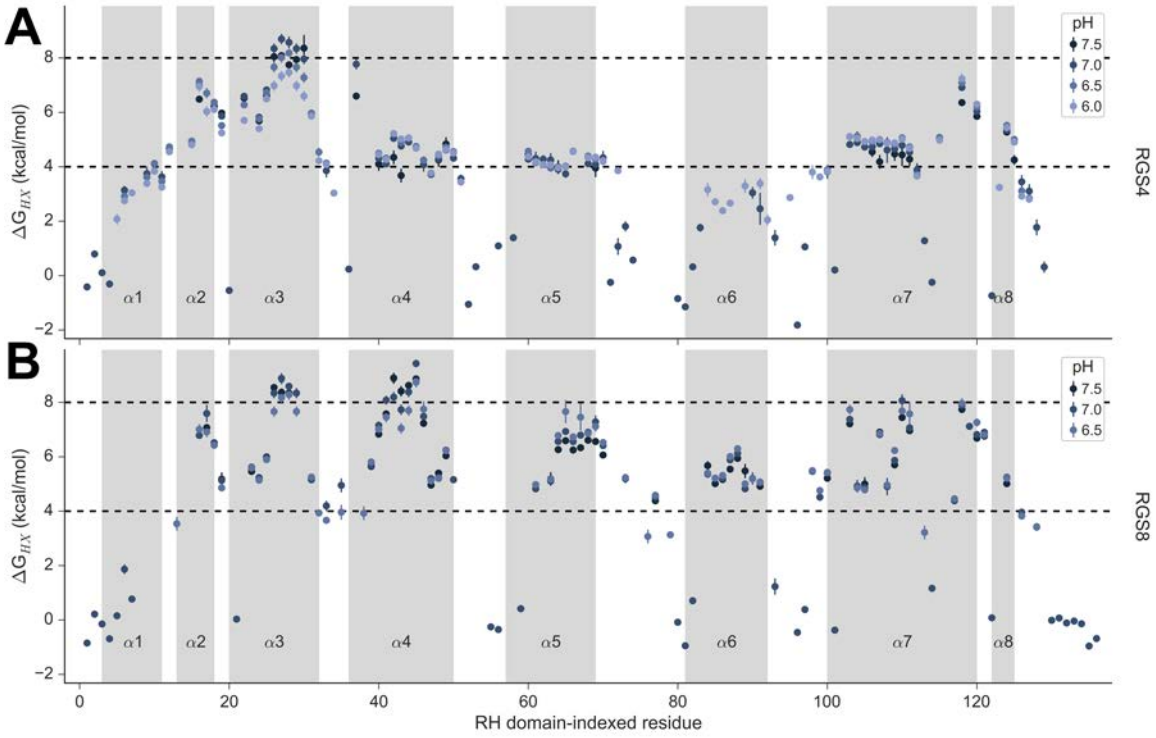

Figure 7.3.4.1 All un-weighted ΔG_{HX} measurements

A-B) Un-weighted, independent calculated free energies for each condition (residue and pH combination) with available data for RGS4 (A) and RGS8 (B). Dashed lines drawn at 4 and 8 kcal/mol to draw the eye. Note systematic underestimation of ΔG_{HX} in (A) at pH 6 in α3 compared to ideal agreement in α5 across all conditions measured, all of which had a k_{ex} within the limits of detection for the laboratory frame (see Figure 7.3.1.1).

7.4. Thermodynamic determinants of RGS buried cysteine accessibility

7.4.1. Overall stability of the α4-α7 helix bundle

As expected—based on hydrogen bonding networks that stabilize secondary structural elements—the backbone amides in the helices are much more protected against exchange than are the loop regions (Figure 7.3.4.1). Our data show that on the whole, the entire α4-α7 helix bundle in RGS4 is less stable than the bundle in RGS8. The median ΔG_{HX} of α4, α5, and α6 in RGS8 is greater than the respective median in RGS4 by 2.4, 1.7, and 2.5 kcal/mol (Table 7.4.3.1).

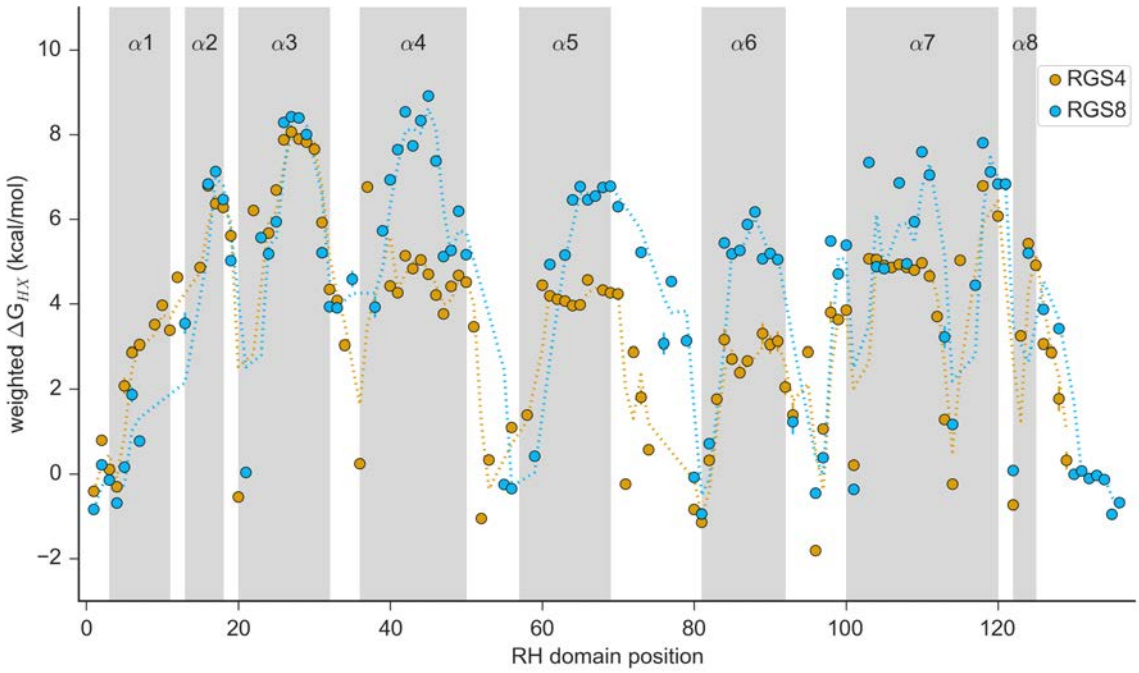

Figure 7.4.1.1 Weighted ΔG_{HX}

Circles represent weighted free energies. Error bars represent propagated uncertainty (most do not extend beyond the symbol). Dotted lines represent rolling mean of data to guide the eye.

This strong relative preference for the closed conformation in RGS8 is in close agreement with our findings from the relative stability measured by SX ($\Delta\Delta G_{SX}$ 1.9 kcal/mol), and the spatial resolution of the HX data (Figure 7.4.1.1) paint a much richer picture of the collective motions in these two proteins that explain why Cys148 in RGS4 is so much more accessible to thiol modification.

7.4.2. Contiguous stability

The most protected amide, with the largest ΔG_{HX}, in a protein is indicative of the protein's stability toward global unfolding(65). The difference in global unfolding is modest; $\Delta G_{HX,max}$ is 8.9 kcal/mol for RGS8 and 8.0 kcal/mol for RGS4 (Table 7.4.3.1). The difference in the number of highly stable residues, though, is marked. RGS8 has seven residues with $\Delta G_{HX} > 8$ kcal/mol to RGS4's one, and the trend holds for 7 kcal/mol (16 vs. 5), 6 kcal/mol (31 vs. 13), and 5 kcal/mol (53 vs. 22).

These stable residues (take for example those with $\Delta G_{HX} > 6$ kcal/mol) are also differentially distributed between the two proteins. In RGS4, the stable core is isolated to α2-α3, two residues the nearby C-terminal portion of α7, and a single residue at the base of α4 adjacent to α3 (Figure 7.4.2.1). ΔG_{HX} is mapped to the structures by colored spheres at the positions of the amide nitrogen atoms from blue (>7 kcal/mol) to white (3.5 kcal/mol) to yellow (0 kcal/mol).

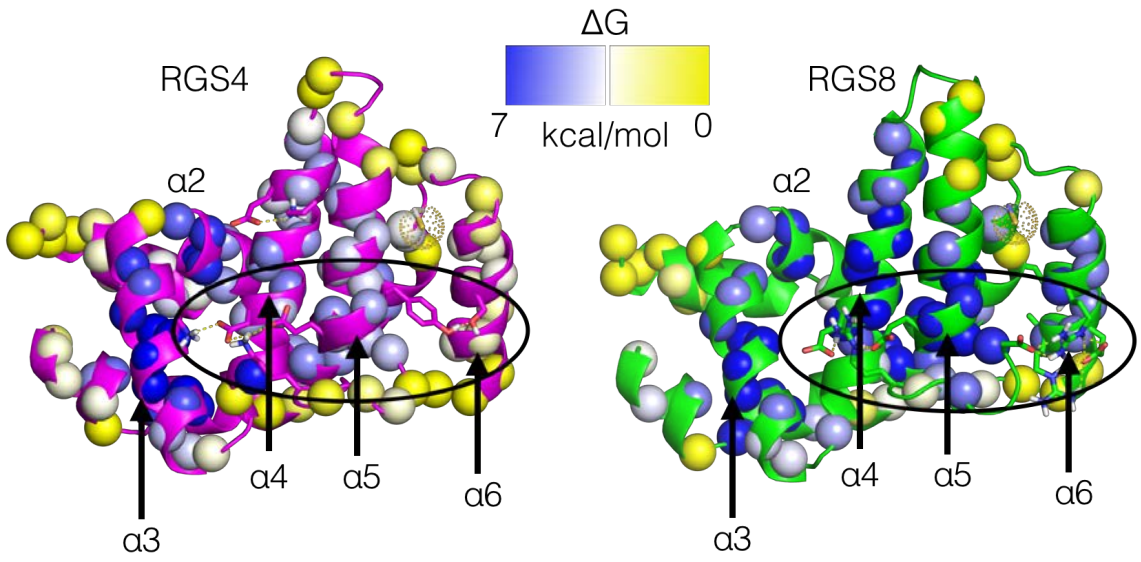

Figure 7.4.2.1 ΔG_{HX} mapped onto RGS4 and RGS8 structures

Weighted ΔG_{HX} encoded as colored nitrogen atom spheres. Spheres for residues whose HSQC crosspeaks were overlapped are omitted for clarity. Stable core from α3 does not extend into RGS4 α4-α7 bundle as it does for RGS8 (black ellipses). In this and the following figures detailing each of the interfaces, RGS4 (PDB 1AGR, chain E) and RGS8 (PDB 2IHD) will be shown side by side with ΔG_{HX} mapped onto from blue (most stable) to white (average) to yellow (least stable) on the nitrogen atoms (shown as spheres for residues where data were acquired). Hydrogen bonds (H-bonds) are shown as yellow dashes, and the sulfur atom of Cys148/Cys142 is rendered as a sphere of yellow dots.

In stark contrast, highly protected residues in RGS8 are distributed in a contiguous network spanning α2, α3, α4, α5, α6, and α7. This extended network suggests that only cooperative, global unfolding exposes these highly protected NHs to exchange. Conversely, it appears that RGS4 has several regions that undergo sub-global cooperative unfolding.

7.4.3. The weak link: α6

By a wide margin, α6 is the least stable helix in the RGS4 α4-α7 bundle, and in fact, of any helix in the protein in terms of median (2.7 kcal/mol) or maximum (3.3 kcal/mol) ΔG_{HX} (Table 7.4.3.1). At equilibrium, this equates to a partially unfolded form populated to ~1% of the native state in which the α6 backbone amides are available for exchange.

helix	RGS4 median	RGS8 median	RGS4 maximum	RGS8 maximum	RGS4 values	RGS8 values
1	3.0	0.2	4.0	1.9	8	5
2	6.3	6.6	6.8	7.1	4	4
3	6.7	5.6	8.1	8.4	11	11
4	4.5	6.9	6.8	8.9	13	13
5	4.1	5.8	4.6	6.8	10	10
6	2.7	5.2	3.3	6.2	11	10
7	4.9	5.7	6.8	7.8	17	16
8	4.1	3.9	5.4	5.2	4	3

Table 7.4.3.1 ΔG_{HX} (kcal/mol) by helix

The dynamic modes of α helices have been exhaustively characterized(77), and common characteristics of highly dynamic α helices (those prone to helix-coil transitions that facilitate HX) are short length, depletion of hydrophobic residues, lower amphipathic moments, and higher solvent accessible surface area. Differences in these parameters might explain why α6 is so much less stable in both RGS4 and RGS8, and why it is less stable in RGS4 than in RGS8. Because the length and solvent accessibility of α6 is equivalent in RGS4 and RGS8, we focused on the hydrophobic and amphipathic properties of the residues in the

α4-α7 bundle that form the interfaces between the helices (removing from calculations the section of α7 that interacts with the α1-3 and α8 bundle).

Helix properties were calculated with an in-house Python implementation of the algorithms used by Gautier and coworkers(78). The mean hydrophobicity for each helix was calculated according to Equation 7.4.3.1 using calculated amino acid hydrophobicities(79), and the amphipathic moment was caluclated using Equation 7.4.3.2, where δ is the angle separating side chains along the backbone (assumed to 100° for α helices).

$$\overline{H} = \frac{\sum_{i=1}^{n} H_i}{n}$$

Equation 7.4.3.1

$$\overline{\mu H} = \frac{\sqrt{[\sum_{i=1}^{n} H_i \times sin(i\delta)]^2 + [\sum_{i=1}^{n} H_i \times cos(i\delta)]^2}}{n}$$

Equation 7.4.3.2

	Mean hydrophobicity		Amphipathic moment	
helix	RGS4	RGS8	RGS4	RGS8
4	0.28	0.35	0.29	0.30
5	0.24	0.32	0.62	0.84
6	0.08	-0.03	0.09	0.19
7	0.27	0.25	0.52	0.34

Table 7.4.3.2 Helical hydrophobicity in α4-α7 bundle

Mean hydrophobicity of residues in each helical segment and amphipathic moment resulting from the orientation of hydrophobic residues and hydrophilic residues toward opposite faces of the helix. The range of possible mean hydrophobicity is -1.01 to 2.25, and the range of possible amphipathic moment is 0 to 3.26. Larger amphipathic moments indicate stronger partitioning.

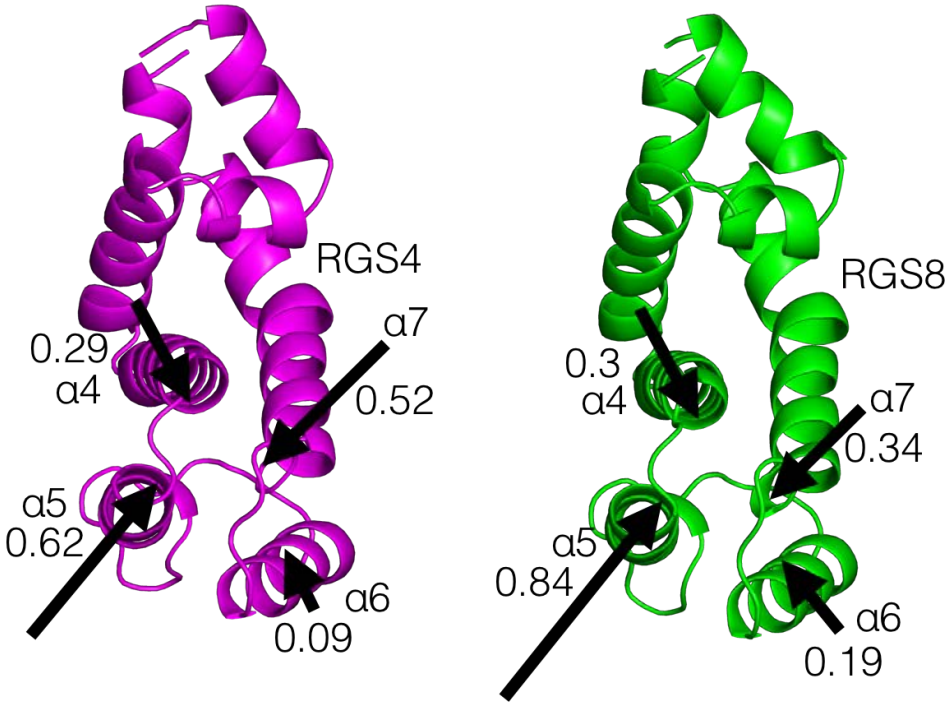

Figure 7.4.3.1 α4-α7 bundle amphipathic moments

Approximate directions and magnitudes of amphipathic moment vectors for helical segments in α4-α7 bundles in RGS4 (PDB 1AGR, chain E, left) and RGS8 (PDB 2IHD, right). Magnitude of vectors indicated by length of arrow tail (not including the arrow head). In both RGS4 and RGS8, α6 has the smallest amphipathic moment in the bundle.

This analysis shows that α6 has the lowest amphipathic moment and mean hydrophobicity in the bundle for both RGS4 and RGS8 and that α6 in RGS4 has a smaller amphipathic moment than α6 in RGS8. The amphipathic moment of a helix is the driving force that orients its most hydrophobic face toward the center of the bundle. Since RGS4 α6 has almost no amphipathic moment, it lacks a strong force anchoring it to the hydrophobic core of the α4-α7 bundle. This is reflected in the lack of stabilization of inner-facing residues compared to residues whose backbone amides are solvent exposed. We explore this further in 7.4.5 and 7.4.6.

RGS8 α6 is also significantly less stable than its neighbors and has a lower amphipathic moment, but it displays evidence of protection from tighter helix packing. The two residues whose backbone amides are oriented toward the core of the α4-α7 bundle and that lie closest to the midpoint in the helix are Thr131 and Arg132. These amides are stabilized by ~1 kcal/mol compared to their neighbors.

7.4.4. The α4-α5 interface

In RGS8, α4 shows what appears to be a classic example of helix fraying—where the residues at the ends of the helix are less stable than residues toward the center. Examination of α4 and the juxtaposed residues in α5 shows a pattern that appears to demonstrate key stabilizing contacts locking these helices

together. The interfacial hydrophobic contacts are completely conserved in α4 and α5 and a salt bridge triad mediated by Arg119 between Glu84 and Glu111 in RGS8 is substituted by a salt bridge triad mediated by Lys125 between Asp90 and Glu117 (Figure 7.4.4.1). The stability of these H-bond networks may be unequal, however, due to the geometry of these interactions. The distances between donor and acceptor in the RGS8 triad are 2.0 Å and 1.8 Å for Glu84 and Glu111 carboxylates with the guanidinum group of Arg119, respectively. In RGS4, due to Asp90's shorter side chain and the substitution of a primary amine for the conjugated guanidinum, the distances are 2.4 Å and 2.5 Å, respectively.

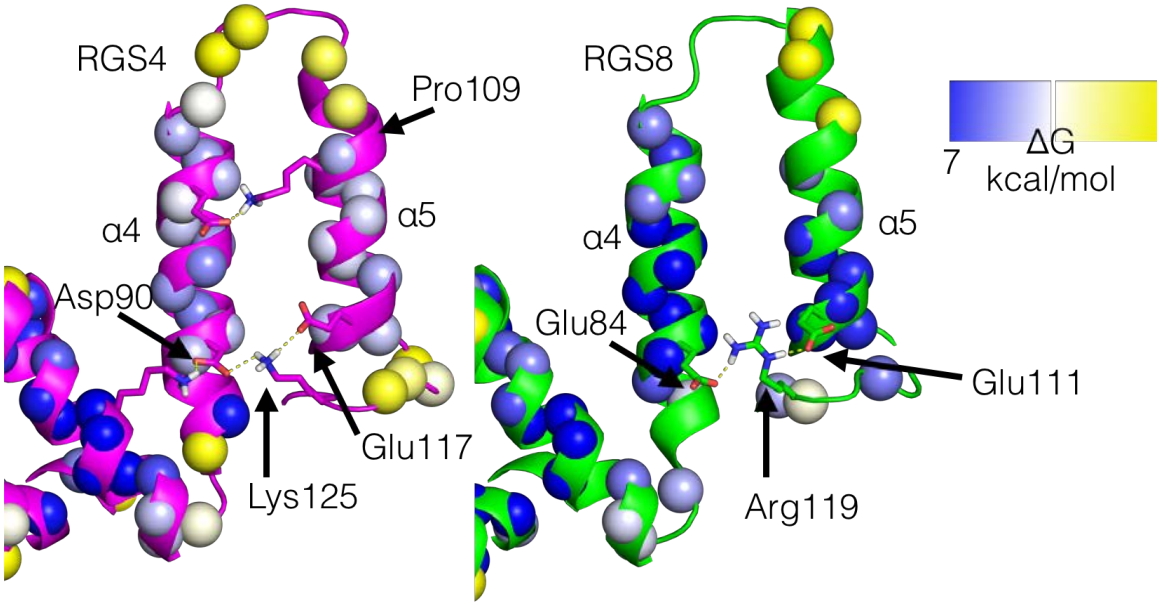

Figure 7.4.4.1 α4-α5 interface comparison

RGS4's α4-α5 interface has a weaker salt bridge than RGS8, has significantly lower stability in the lower portion of the interface, and the stability of RGS4 α5 is also less correlated to stability in α4 than in RGS8. See Figure 7.4.2.1 legend for description of representations and coloring.

Another difference in this region is the substitution of Ser103 for Pro109 in RGS8 and RGS4, respectively. The effect of Pro109 in RGS4 is less straightforward than a weaker salt bridge. While prolines have a reputation for destabilizing α helices, there are many situations in which prolines in the first turn of an α helix or loop region is stabilizing against helix-coil transitions because of the marginal entropic difference for proline due to its conformationally restricted side chain(80) and because prolines stabilize secondary structure through rigid dihedral angles instead of H-bonds(81).

7.4.5. The α5-α6 interface

The interface between α5 and α6 is less conserved between RGS4 and RGS8 than the α4-α5 interface. As discussed in 7.4.3, α6 has a very low amphipathic moment, and its helix packing contacts reflect that. The only major hydrophobic contacts come from Phe109 and Leu135 in RGS8 and Tyr115 and Met141 in RGS4, from α5 and α6 respectively (not shown). Like the α4-α5 interface, the α5-α6 interface in RGS8 is stabilized by a salt bridge triad mediated by Arg132 between Asp114 and Glu129 (Figure 7.4.5.1). This network is further extended by a H-bond between the Asp114 side chain carboxylate and the backbone amide of Val115 and between the carboxylate of Glu129 and the side chain of Arg128. This extensive H-bond network is absent in RGS4, where Arg132 and Asp114 are substituted by Ser138 and Ser120.

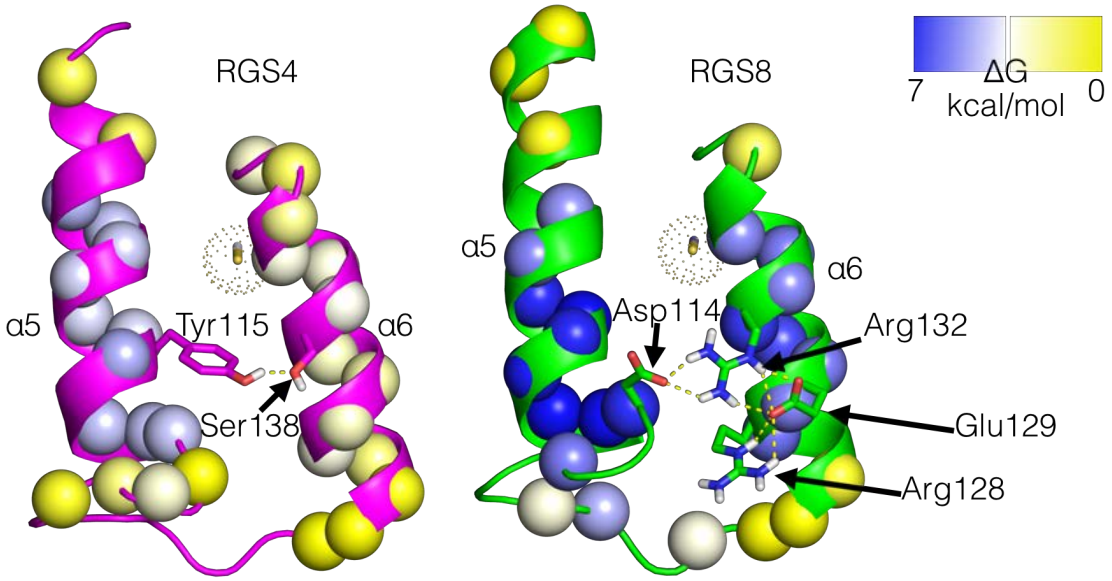

Figure 7.4.5.1 α5-α6 interface comparison (RGS4 left)

RGS8 α5 and α6 are linked by a stabilizing salt bridge, and the free energy of unfolding is significantly higher in the region spanning the salt bridge. RGS4 lacks a salt bridge at this interface and has significantly less stability in these positions. The stability of RGS4 residues in α6 are less correlated to the α5 residues across the interface at the same vertical position than are those in RGS8 (e.g. Arg132 with Asp114). See Figure 7.4.2.1 legend for description of representations and coloring.

The difference in these interfacial contacts stands out as a key difference in the α4-α7 bundle. Though Ser138 makes a H-bond with Tyr115, the interaction is much less favorable than the salt bridge triad in RGS8. Recall that the ΔG_{HX} of Arg 132 (6.2 kcal/mol) is stabilized by 1 kcal/mol compared to the median for α6 (5.2 kcal/mol). There is evidence that the contribution of a second ion pair in a surface-exposed salt bridge can stabilize the initial ion pair by up to 1

kcal/mol(82). In contrast, RGS4 α6 does not show significant correlations in ΔG_{HX} or interactions indicative of inter-helix cooperativity with α5.

7.4.6. The α6-α7 interface

This interface is entirely bereft of hydrophobic packing contacts from α6 in both RGS4 and RGS8. The interface is also comparatively conserved, but the paired mutations in interfacial residues may contribute to the differential stability. In RGS8, the side chains for Ala130 and Gln145 are oriented ~40° from a plane connecting the center of each helix. Due to the tilt of α7 compared to α6, the orientation of the side chain for Gln145 brings it in close proximity to the space that would be occupied by the side chain for Ala130 (if it had one), even though Ala130 is in the second turn from the α6-α7 loop and Gln145 is in the first turn. Even still, the side chain of Gln145 makes a H-bond to the side chain of Thr127, in the third turn of α6.

In RGS4 though, Ala130 is replace by Glu136, and Gln145 is replaced by Glu151. The aforementioned orientation of the side chains for these two positions dictated by backbone geometry would bring together the two glutamate side chains—a highly unfavorable clash due to charge-charge repulsion. In 1AGR, Gln151 bends to form an ion pair with the side chain of Lys154, but in 1EZT (the solution NMR structure of RGS4), these glutamate carboxylates are within 2.4 Å. In 1AGR, there is also a salt bridge between Lys155 and Asp130 with an

additional H-bond between Lys155 and Thr133, but the low ΔGHX in this lower portion of α6 indicates that these interactions may not stabilize to a significant extent.

Both sides of the top of this interface, encompassing the two C-terminal turns of α6 and the first turn of α7, share common stability to HX despite the differing magnitudes of this stability in RGS4 and RGS8. The relative difference average of ΔG_{HX} between RGS4 and RGS8, $\Delta\Delta G_{HX}$, of the residues enclosed in the ellipses in Figure 7.4.6.1 is 1.8 kcal/mol. Since the difference in free energy for accessibility to Cys148/142 by DACM, $\Delta\Delta G_{SX}$, was 1.85 kcal/mol, this suggests that rearrangement of the top of the α6-α7 helix pair is strongly correlated to cysteine accessibility. These findings suggest an alternative entry vector for small molecules into the α4-α7 bundle than the bottom-entry vector proposed by Vashisth and coworkers(42).

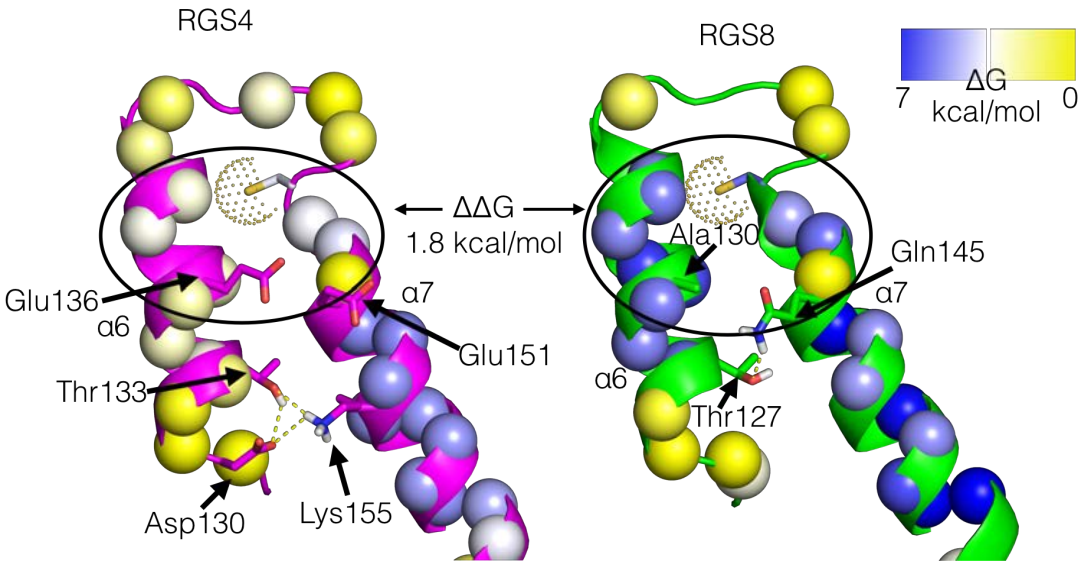

Figure 7.4.6.1 α6-α7 interface comparison

The top of α6 and α7 have similar stabilities in both RGS4 and RGS8 suggestive of inter-helical cooperativity in unfolding. See Figure 7.4.2.1 legend for description of representations and coloring.

7.5. Discussion

7.5.1. The path of least resistance: access to Cys148

We sought to characterize the relative population of excited states that afford access to small molecules that covalently modify Cys148 (RGS4) or Cys142 (RGS8) and to determine which regions within the α4-α7 bundle are displaced in this excited state. We identified several inter-helical interactions that appear to govern cooperativity of these partial unfolding processes. With these data, we are able to significantly revise the model for how dynamics determine the

specificity of inhibitors for RGS proteins, and we can rationalize a number of disparate pieces of experimental evidence.

In the model of the excited state proposed by Vashisth and coworkers(42), α5 and α6 (along with the α5-α6 loop) pivot on a hinge made up of the α4-α5 and α6-α7 loops. Small molecule inhibitors, e.g. CCG-50014, then transit up the groove along these axes to reach Cys95 (or Cys148) in a lower entry vector. There are several inconsistencies between the rigid helix pair motion model and our measurements of actual residue-level solvent exposure in the native ensemble. Firstly, the observed stabilities of α4 and α5 are highly correlated, whereas α5 and α6 are not. If α5 and α6 moved cooperatively in a large motion that exposed their inner backbone amides to exchange, their stabilities would be correlated. It is difficult to reconcile our data to this entry vector for small molecules, and indeed we propose an alternative model of the excited state and entry vector.

On the other hand, multiple residues at the top of α7 show similar exchange kinetics to α6 (see Figure 7.3.1.1, upper panel, pH 6 and region enclosed by ellipse in Figure 7.4.6.1). It is possible that displacement of this region opens a groove for access to Cys148 either between the α4-α5:α6-α7 (see Figure 7.5.1.1 for cartoon representation) interface or by the separation of the tops of α6 and α7, opening a channel under the α6-α7 loop.

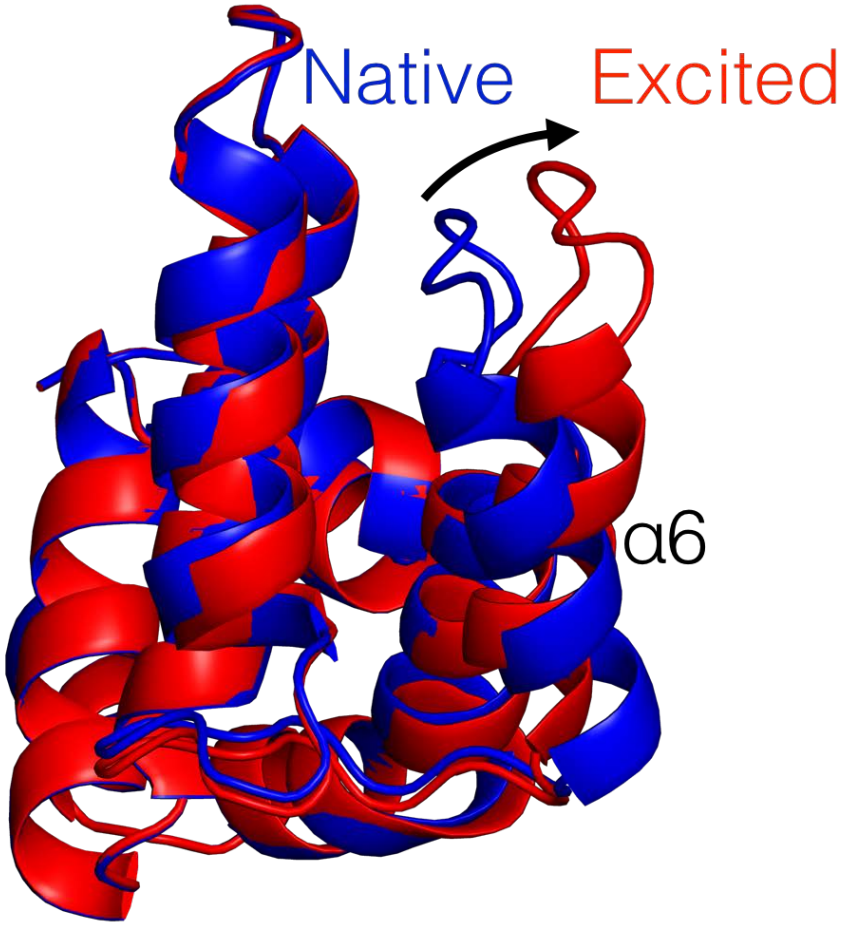

Figure 7.5.1.1 Diagram of top-groove excited state

Cartoon representation of an excited state model in which the α6-α7 loop and adjacent α6 and α7 segments move outward away from the α4-α5 helix pair. This motion would create a groove between the helix pairs and expose Cys148 and Cys95 to solvent and is consistent with our HX and SX data for RGS4 and RGS8.

Due to the instability of the α6 bundle and its position occluding Cys148, its displacement is almost surely involved motions exposing Cys148 because conformational states are populated according to Boltzmann distributions, so conformations with these lower energy changes are more highly populated in the ensemble. This is possible in any number of ways, from the aforementioned

openings to a rotation of the entire α6 helix away from either α5 or α7. It is more difficult to rationalize an entry vector from these latter motions for molecules interacting with Cys95, since it is physically separated from the center of the α4-α7 bundle by the side chain of Phe149 in RGS4 (Phe143 in RGS8). In this case, an opening of the groove between the α4-α5 and α6-α7 pairs by the concerted movement of the top of α6-α7 away from α4-α5 seems more likely. Unfortunately, our current data do not allow us to invalidate any of these hypotheses without concrete measurements of cooperativity in unfolding.

It is also possible that both the top- and bottom-entry models of conformational opening both occur, but that these motions occur on different time scales. The greater protection from exposure of backbone residues at the interaction between α4-α5 compared to that between α5-α6 means that the former would occur on a slower time scale but may still have some relevance to larger-scale openings (e.g. leading to exposure of Cys95).

Our model also meshes well with what is known about PIP_3 binding. PIP_3 binds near the top of the α4-α5 helix pair and intervening loop and inhibits RGS GAP activity. Native HX demonstrates that this region is relatively more flexible than the lower part of α4 and α5 in both RGS4 and RGS8. If PIP_3 binding induced an outward bend in the top of α4-α5, it would expose the hydrophobic inner leaflet of α4-α5 and widen the groove between α4-α5 and α6-α7. Since PIP_3 is

tethered to the membrane, this could have the effect of encouraging palmitoylation of Cys95. Supporting this idea, previous work in our lab also showed that 4-hydroxy-2-nonenal (4-HNE), a lipid peroxidation product produced at the membrane during oxidative stress, modifies Cys148 in cells(83).

Chapter 8. Conclusions and relevance for drug discovery

8.1. Effects of inhibitors on RGS Structure

RGS4 undergoes a transition to a molten globule upon modification of Cys148. While it retains a significant portion of native-like helical character, tertiary contacts in this conformational ensemble are highly unstable—especially in the α4-α7 helix bundle. This instability exposes hydrophobic patches that lead to dimerization and, to a lesser extent, oligomerization. Mutation of Cys148 to Tyr recapitulates this molten globule with a lower propensity to aggregate and without the tendency to dimerize. When RGS8 Cys142 is modified, we did not detect any soluble inhibited species—molten globule or otherwise. RGS8 is profoundly destabilized, however, upon modification or mutation of Cys142 to Tyr and inevitably aggregates.

In light of what we learned from our study of the thermodynamics of the native state ensemble, these findings are quite easy to rationalize. RGS8 has a rigid α4-α7 helix bundle that is stabilized by large amphipathic moments and distributed salt bridge triads, so the energy barrier that must be surmounted in order to expose Cys142 to modification involves the collective unfolding of a large portion of the bundle. Once the cysteine is modified, RGS8 becomes trapped in a state with very few native-like contacts and a configuration that

likely exposes a number of highly hydrophobic residues, e.g. in α5 with its large amphipathic moment and high solvent exposure. Unable to refold to any semblance of its highly stable native state due to steric occlusion of the central hydrophobic packing vestibule by the Cys adduct, RGS8 is driven to aggregate to minimize the solvent exposure of these hydrophobic patches.

On the other hand, RGS4 has a much more flexible α4-α7 bundle with a more shallow energy landscape in the native ensemble. Smaller collective motions expose Cys148, and so the nascent adducted species retains more native-like tertiary contacts. Modification of Cys148 is still destabilizing, but the helices are more able to accommodate the loss of native tertiary contacts. Pro104 and Pro109 may facilitate this stabilization because they retain more native-like secondary structure and can help resist the helix-coil transition. It is also possible that α6's flexibility coupled with a lower amphipathic moment simply adjusts to accommodate a Cys adduct. In any case, as a result of its lessened driving force toward rigid bundle packing, this ensemble has a relatively stabile molten globule.

It is possible to test this hypothesis and refine our understanding of RGS4's molten globule using HX pulse-labeling techniques. In these experiments, RGS4 would first be thiol-modified, then exchanged into D_2O for various times at various pHs, after which the solvent would be re-exchanged for H_2O, add

reducing agent to remove the adduct and reform the native fold, and finally measure the degree of HX from the molten globule state(65).

8.2. Flexibility, dynamics, and inhibitor specificity

Molten globules and transient sub-global unfolding are all very interesting from a biochemical perspective, but do these findings have relevance for RGS drug discovery? We believe the answer is a resounding yes. There are numerous examples where flexibility drives ligand specificity among otherwise highly homologous proteins including some for clinically important drugs(84). If our goal is to develop specific inhibitors of RGS4 activity in vivo, we need to leverage any differentiating factors we can find because of the structural conservation of the Gα binding interface and the RGS domain in general. In the extended family of RGS proteins, RGS4 and RGS8 are siblings, and we have shown that their dynamics are profoundly different. Exploration of other more distantly related RGS proteins is likely to reveal hitherto unimaginable diversity in cryptic binding pockets.

On some level, it is surprising that it has taken nearly 20 years since the structure of RGS4 became known for the importance of dynamics for regulation and specificity to be recognized. Take, for example, palmitoylation of Cys95, which is completely buried in the hydrophobic interior of the RGS domain. It is indeed difficult to conceive of a way that Cys95 palmitoylation could occur

without a major unfolding event. Palmitoylation of Cys95, which occurs in at least RGS4, RGS10, and RGS16(85), may be an indicator of α6 flexibility. We have shown that α6 in RGS4 is highly flexible in the apo state, and α6 in RGS10 is almost completely disordered two different solution NMR structures (2I59 and 2DLR) and lacks electron density in the $G\alpha_{i3}$-bound crystal structure (2IHB), and the B factors in α6 in the apo RGS16 crystal structure (3C7L) are the highest in the RGS domain—especially near the α6-α7 loop.

There is strong evidence to suggest that α6 may play an important role in specificity between RGS proteins outside of the context of the canonical families. For example, RGS2, which is the only $G\alpha_q$-specific R4 protein, has an alternate conformation for α6, and α6 is disordered in the distantly related p115-RhoGEF crystal structure bound to $G\alpha_{i/13}$ (1SHZ). Indeed, John Tesmer, the crystallographer who released the first crystal structure of any RGS protein, noted in a review on the structure and function of RGS homology domains, "[α6] has the greatest tendency to be disordered or variable in structure among the characterized RH domains"(86). Disorder in α6 may be the key factor that has prevented successful crystallization of RGS4 in the apo state despite exhaustive attempts by many skilled crystallographers.

The holy grail of these inquiries is a high resolution model of the excited states of these proteins. This would enable virtual screening and rational drug

design for molecules that enrich the population of these unique excited state conformations through conformational selection. These excited state conformations need not necessarily even allosterically inhibit Gα binding to be successful RGS inhibitors. If they block palmitoylation or enhance PIP_3 binding, the net result—an inhibition of GAP activity—achieves the same end.

Chapter 9. Future Directions

9.1. Details of RGS4 dynamics

9.1.1. Cooperativity of specific interactions

Our work has laid the foundation for empirical dissection of the cooperativity and concerted motions that lead to excited states in RGS4. As such, specific mechanisms we have proposed are hypothetical and based on correlations from the native state ensemble. To test these hypotheses will require more rigorous experimental evidence. Two promising lines of inquiry include mutations to test the importance of the residues we have implicated and quantitative measures of cooperativity using chemical denaturants.

In the case of RGS4 and RGS8, there are 34 divergent positions in the bundle domain, and approximately two thirds of those are similar (e.g. Ile vs. Leu). We have identified a smaller subset of residues that, based on their interactions and proximity to regions with divergent thermodynamic stability, have a high probability of driving differential dynamics. Systematic mutation of these positions in RGS4 to their cognate RGS8 residue would delineate the specific contribution each plays. An alternative approach that may be useful would be to swap entire helices to create chimeric RGS proteins (e.g. RGS8 with

RGS4's α6). The downside to this approach is that the paired interactions we have identified between helices may limit the effect size.

To validate the constituent residues of subglobal unfolding units, chemical denaturation (e.g. with guanidinium chloride or urea) is a powerful technique when combined with RT-HX that has been used to great effect historically(87, 88). This relatively simple addition to our current experimental protocol allows the determination of residues' dependence of the free energy of global unfolding on denaturant concentrations (the so called 'm-value'). M-values correlate with the degree of surface area exposure in an unfolding event, and residues that participate in concerted unfolding events have the same (or similar) m-values.

9.1.2. Faster time scale dynamics

In theory, these collective, slow backbone motions we have characterized by HX and SX should have contributions from faster, smaller motions on the μs-ms and even ps-ns time scales. We began initial studies to characterize these using Carr Purcell Meiboom Gill (CPMG) relaxation dispersion and nuclear spin relaxation but found marginal exchange broadening (R_{ex} 1.5-3 Hz on average across the α4-α7 bundle for RGS4 and <1 Hz for RGS8) and relatively little change in order parameters (not shown) at 25 C. It may be possible to alter the sample conditions in order to enhance the sensitivity of this technique (e.g. changing temperature or perdeuteration of non-exchangeable protons).

Given our evidence that the excited state(s) of RGS4 are within ~2 kcal/mol of the native state, we expect that the RGS4 excited state should be detectable by CPMG relaxation dispersion because the technique has been successfully used to characterize the excited states of proteins with equivalent energy differentials(89). This technique would be a very powerful tool to study the RGS4 excited state and can even be used to calculate chemical shifts of the excited state, from which it is possible to generate excited state structures(90). CPMG relaxation dispersion has the additional benefit that it can be used to measure the motions of side chains, which would greatly enhance our understanding of the RGS4 excited state(91).

9.2. Dynamic diversity in the RGS domain

From our studies, we have concrete evidence of diversity in the dynamics of RGS proteins, but we have only explored this diversity in any detail for RGS4—a relatively flexible protein and RGS8—a comparatively rigid one. Data from structural genomics efforts indicate that R12 family proteins, at least, have more in common with RGS4 than with RGS8 (see Section 1.2.3). What of the other 20+ RGS proteins though? How common is a highly flexible bundle domain? If excited states are prevalent among RGS proteins, it will likely be important to understand their structural details in order to avoid off-target effects of ligands designed to bind to RGS4's excited state. Since RGS domains

are very amenable to study by NMR (there are entries for 7 RGS proteins in the Biological Magnetic Resonance Databank and structures in the PDB of 2 others), it is conceivable that a representative survey of their dynamic natures would be tractable without needing to acquire full resonance assignments—which is time consuming and expensive compared to HX.

Chapter 10. Methods

10.1. Construct design and molecular biology

Rat Δ50 RGS4 (residues 51-205) was PCR-amplified (primers: 5′: GAGCCATGGTGAGCCAAGAAGAAGT and 3′: CGGGCTTTGTTAGCAGCCGG) from pMALC$_2$H$_{10}$T-RGS4 into pET28a(+) using NcoI and XhoI restriction sites to obtain a C-terminal 6X His-tagged bacterial expression construct closely matching that used by Powers and coworkers in solving the NMR structure of RGS4(21). The RGS4 domain (residues 51-179) was subsequently sub-cloned into pET28 derivative pNIC28-Bsa4 (a gift from Opher Gileadi; Addgene plasmid # 26103) using ligation independent cloning(92), to afford a tobacco etch virus (TEV)-cleavable N-terminal 6X His-tagged construct. Briefly, PCR was used to amplify the desired region of rat RGS4 with flanking sequences (primers: 5′: TACTTCCAATCCATGGTGAGCCAAGAAGAAGTCAAG and 3′: TATCCACCTTTACTGTCAATTGGTCAGGTCAAGATAGA). The pNIC28-Bsa4 plasmid was digested with BsaI to generate a blunt, linearized vector. Cohesive ends on both vector and insert were generated by the action of T4 DNA polymerase supplemented with either dCTP (insert) or dGTP (vector). Insert and vector with cohesive ends were mixed and immediately used to transform DH5α

E. coli. Human RGS8 (residues 45-180) in pNIC-Bsa4 was a kind gift from Stefan Knapp (University of Frankfurt).

Oligonucleotide-directed mutagenesis (Quik Change II, Agilent) was used to create the cysteine-deletion mutant of RGS4 designated RGS4c (C71A, C95A, C132A, C148A, C183A, C196A, C204A), c148C (all deleted except C148A), and c148Y (same as c148C but also bearing C148Y mutation) according to following the manufacturer's protocols and primer design software. Cysteine deletion mutants (C89A, C142A, and combination RGS8c) and c142Y were also generated for RGS8 using the same method.

10.2. Protein expression and Purification

10.2.1. Protein Expression

DNA was used to transform BL21-CodonPlus (DE3)-RIPL *E. coli* (Agilent). All cultures included 50 μg mL^{-1} kanamycin and 50 μg mL^{-1} chloramphenicol. Initially, colony optimization in Lysogeny broth (LB) was used to generate glycerol stocks of highly-expressing bacterial populations of each protein (93). Subsequent large-scale growths used glycerol stocks to inoculate starter cultures (5-10 mL LB) the night before protein production. Starter cultures were expanded to intermediate cultures of 100-200 mL terrific broth (TB) to re-establish log-phase growth then expanded to 1-6 L of TB until absorbance at 600 nm (OD600)

reached 0.2. These large-scale cultures were shaken at 275 rpm and grown to OD600 of 2 at 37°C before centrifugation at 3650 *g* to pellet bacteria. Pellets were resuspended in M9 minimal media (either ^{15}N-labeled or natural abundance) of equal volume to the pre-pellet culture and returned to shaking incubator for 1 hour at either 30°C (RGS4 mutants except c148Y) or 18°C (c148Y and all RGS8 mutants). Following this hour, during which the OD600 increased 0.2-1 units, protein expression was induced by addition of β-D-1-thiogalactopyranoside solution (freshly prepared in water by dissolving powder (RPI Corp)) to a concentration of 500 μM (final) with continued agitation for 16 hours at aforementioned temperatures. Following induction, cultures were centrifuged at 3600 *g* to pellet cells, resuspended to a 2:1 buffer:pellet volume ratio in 25 mM HEPES pH 7.5 500mM NaCl 10mM imidazole supplemented with protease inhibitors (100 μg mL^{-1} phenylmethylsulfonyl fluoride, 1 μM Leupeptin, 1 μg mL^{-1} Pepstatin, and 1 μM E-64), and frozen in liquid nitrogen for storage at -80°C until purification (up to two weeks).

10.2.2. Protein Purificaiton

All purifications were carried out at 4°, and all buffers and samples were filtered to remove particles larger than 0.45 μm using polyethersulfone membrane filters. During purification of RGS4 WT, purification buffers were supplemented with dithiothreitol (DTT) to a final concentration of 1 mM to

prevent the formation of disulfide bonds. Frozen, resuspended pellets were thawed and lysed using two cycles of freeze-thaws using liquid nitrogen and a 37°C water bath, adding 1 mg lysozyme (Research Products International) and 20 μg DNaseI (Roche Applied Science) per mL pellet after the first and second thaws, respectively. Following these additions but before refreezing, samples were mixed for 10 minutes using a HulaMixer (Life Technologies) at 50 rpm. The lysate was clarified by centrifugation at 100,000 *g* for 1 hour in a Sorvall T-647.5 rotor (Thermo Fisher Scientific) and filtration of the supernatant (0.45 μm).

RGS protein was purified from clarified lysate using two to three stages of column chromatography consisting of 6xHis affinity capture, cation exchange, and/or SEC until judged 99%+ pure by SDS-PAGE depending on the construct. Affinity purifications for all constructs were performed using 25 mL of Ni Sepharose 6 FF resin self-packed in an XK 16/20 column with an ÄKTApurifier (G.E. Healthcare). FPLC affinity purifications used segmented gradients of increasing imidazole concentration optimized for target protein purity. Fractions containing 90+% pure target protein were pooled and concentrated in an Amicon stirred-cell concentrator to 20-50 mL.

Δ50 RGS4 constructs were diluted 15-fold in 50 mM HEPES, pH 7 to reduce ionic strength and loaded onto a 20 mL SP FF-filled XK 16/20 column equilibrated with the same buffer. Pure RGS4 was eluted using buffer with 1 M

NaCl over segmented gradients. A peak containing 99%+ pure RGS4, as judged by SDS-PAGE, eluted at 100 mM NaCl.

RGS4 and RGS8 RGS domain constructs were purified by SEC following affinity capture. The pooled affinity fractions were diluted in 25 mM HEPES, pH 7.5, 500 mM NaCl to reduce the imidazole concentration below 25 mM and incubated for 8-12 hours with a 20:1 molar ratio of RGS:TEV protease tagged with a 6xHis, purified as previously described(94). Cleaved RGS was purified by passing through the Ni column and collected in the flow-through (uncleaved RGS, TEV, and nonspecific Ni-binding proteins are retained on the column). To remove trace contaminants, cleaved RGS was further purified using a preparatory-grade Superdex 75 column equilibrated with 25 mM HEPES pH 7.5 500 mM NaCl. Following purifications, RGS4 (10 [for Δ50 constructs] to 100 mg protein [for RGS domain constructs] per L of bacterial culture), was then pooled, concentrated to up to 15 mg mL^{-1} (measured by absorbance at 280 nm using molar extinction coefficients calculated by Protein Calculator v3.4: protcalc.sourceforge.net), and dialyzed into the desired buffer for 12 hours at 4°C for use immediately or frozen as aliquots in liquid nitrogen for storage at -80°C.

10.3. NMR spectroscopy

10.3.1. Instrumentation

All experiments were performed on either a Bruker Avance II 800 MHz spectrometer equipped with a 5 mm TCI cryoprobe, Bruker Avance II 500 MHz equipped with a 5 mm triple resonance probe, or a Varian Unity Inova 600 MHz spectrometer equipped with a 5 mm triple resonance probe at the University of Iowa Carver College of Medicine Nuclear Magnetic Resonance Facility.

10.3.2. HSQC samples

^{15}N-labeled samples were prepared from purified protein by dilution to 200 μM in 50 mM HEPES, pH 7.5, 100 mM NaCl, 10% D_2O, and compound or vehicle (DMSO, final concentration ≤2%) in a final volume of 500 μL. Compounds from DMSO stocks were added to a final concentration of 1.1 molar equivalents (unless otherwise noted) and incubated at RT for 1 hr after brief mixing. All spectra were acquired at 25°C.

Backbone resonance assignments for the WT RGS4 RGS domain and WT RGS8 RGS domain were obtained using standard triple-resonance assignment experiments (CBCA(CO)NH, HNCACB, HNCO, and HN(CA)CO). The samples for backbone assignment sample was ^{15}N, ^{13}C labeled 1 mM WT RGS in 20 mM $NaHPO_4$, pH 7, 250 mM NaCl, 10 mM DTT, and 2 mM NaN_3. Spectra were

processed using NMR Pipe(95) and analyzed using CCPN Analysis(45). HSQC spectra for figures were visualized at matched contour levels using nmrglue 0.5(96). Peak intensities were extracted using nlinLS, part of the NMR Pipe package.

Predicted secondary structure based on chemical shifts was calculated using δ2D(47) using backbone H, N, Co, Cα, and Cβ chemical shifts referenced to water.

10.4. Chemical reagents

CCG-4986 was purchased from Chembridge. CCG-63802 was purchased from Tocris. DACM was purchased from Anaspec. CCG-50014 was synthesized as reported previously(38).

10.5. GTPase acceleration assay

Activity assays were performed as previously published(49) with the indicated final concentration of RGS protein and 5 μM Gαi. All liquid additions to assay plates were followed by centrifugation at 1000 *g* for 60 seconds and protein stock solutions were prepared at 4X final assay concentration. Briefly, 10 μL of assay buffer (50 mM HEPES pH 7.5, 100 mM NaCl, 5 mM EDTA, 10 mM MgCl2) was added to a clear 96-well plate. 10 μL of RGS protein in assay buffer was added, followed by 10 μL of human Gαi (R178M, A326S), and finally 10 μL

of 1.2 mM GTP. The plate was incubated at 25 C for 50 minutes. Then, 10 μL of developing/stop reagent was added followed by incubation for 50 min at 25 C.

Absorbance was measured at 642 nm. GTPase activity of Gα was calculated by subtracting A_{642} for GTP-only (30 μL assay buffer, 10 μL GTP) wells from A_{642} for Gα+GTP wells (20 μL assay buffer). GAP activity was calculated by dividing GTP-subtracted A_{642} for RGS-containing wells by A_{642} from Gα+GTP wells. GAP activity reported as % of maximal RGS8 GAP activity.

10.6. Dynamic light scattering

Protein samples used for analyses were obtained directly from size exclusion fractions following inhibitor treatment and filtered with 0.02 μm Anotop syringe filters (G.E. Healthcare). Data were acquired at 15°C on a DynaPro NanoStar (Wyatt) equipped with a 665 nm laser. Data analysis was performed with Dynamics V7 software (Wyatt) using spherical R_g models and standard data filters.

10.7. Circular dichroism

SEC fractions were diluted and dispensed into a 1 mm path length cuvette (110-QS, Hellma) at 10 μM. Data were acquired on a J-815 (Jasco) spectrophotometer. Each spectrum was the average of three accumulations using a scanning speed of 50 nm min^{-1}, 1 nm data pitch, 1 nm bandwidth, and 4 second

data integration time at 25°C. Deconvolution of secondary structure composition using CONTINLL and CDSSTR algorithms through DichroWeb(97) gave self-consistent results using reference set 6 (185-240 nm) (98) and using CDSSTR with the SMP180 reference set (190-240 nm)(99). Fractional helical content reported is the average of these three measures. The DSSP algorithm was used through the 2Struc web server (100) to estimate the secondary structure composition of 1AGR-RH, the RGS4 crystal structure (PDB ID: 1AGR) appended with two N-terminal residues (SM-) and one C-terminal residue (-N) in the RH construct but not found in the crystal structure as random coil in MacPyMol v. 1.7.4.3 (Schrödinger).

Thermal denaturation by CD was performed by collecting single scan spectra from 190 nm to 240 nm collected at 2 degree increments from 10 to 96 °C with a ramp rate of 0.67 °C/min. Refolding was monitored after cooling back to 10 °C over 10 minutes and incubation at 10 °C for 5 minutes. Samples were 16 μM diluted from stock solution measured by A_{280} using molar absorbance calculated by VectorNTI (Thermo Fisher Scientific). Cuvettes used for thermal denaturation were 4 mm path length stirred quartz cells (119-QS, Hellma).

REFERENCES

1. de Mendoza, A., Sebe-Pedros, A., and Ruiz-Trillo, I. (2014) The Evolution of the GPCR Signaling System in Eukaryotes: Modularity, Conservation, and the Transition to Metazoan Multicellularity. *Genome Biology and Evolution*. **6**, 606–619
2. Manglik, A., Kim, T. H., Masureel, M., Altenbach, C., Yang, Z., Hilger, D., Lerch, M. T., Kobilka, T. S., Thian, F. S., Hubbell, W. L., Prosser, R. S., and Kobilka, B. K. (2015) Structural Insights into the Dynamic Process of β2-Adrenergic Receptor Signaling. *Cell*. **161**, 1101–1111
3. Rasmussen, S. G., Choi, H.-J., Rosenbaum, D. M., Kobilka, T. S., Thian, F. S., Edwards, P. C., Burghammer, M., Ratnala, V. R., Sanishvili, R., and Fischetti, R. F. (2007) Crystal structure of the human beta-2 adrenergic G-protein-coupled receptor. *Nature*. **450**, 383–387
4. Bradford, W., Buckholz, A., Morton, J., Price, C., Jones, A. M., and Urano, D. (2013) Eukaryotic G Protein Signaling Evolved to Require G Protein-Coupled Receptors for Activation. *Sci Signal*. **6**, ra37–ra37
5. Siderovski, D. P., and Willard, F. S. (2005) The GAPs, GEFs, and GDIs of heterotrimeric G-protein alpha subunits. *Int. J. Biol. Sci.* **1**, 51–66
6. Oldham, W. M., and Hamm, H. E. (2008) Heterotrimeric G protein activation by G-protein-coupled receptors. *Nat Rev Mol Cell Biol*. **9**, 60–71
7. Ford, C. E., Skiba, N. P., Bae, H., Daaka, Y., Reuveny, E., Shekter, L. R., Rosal, R., Weng, G., Yang, C.-S., Iyengar, R., Miller, R. J., Jan, L. Y., Lefkowitz, R. J., and Hamm, H. E. (1998) Molecular Basis for Interactions of G Protein βγ Subunits with Effectors. *Science*. **280**, 1271–1274
8. Coleman, D. E., Berghuis, A. M., Lee, E., Linder, M. E., Gilman, A. G., and Sprang, S. R. (1994) Structures of active conformations of Gi alpha 1 and the mechanism of GTP hydrolysis. *Science*. **265**, 1405–1412
9. Gilman, A. G. (1987) G proteins: transducers of receptor-generated signals. *Annu. Rev. Biochem.* **56**, 615–649
10. Chan, R. K., and Otte, C. A. (1982) Isolation and genetic analysis of Saccharomyces cerevisiae mutants supersensitive to G1 arrest by a factor and alpha factor pheromones. *Mol. Cell. Biol.* **2**, 11–20
11. Chan, R. K., and Otte, C. A. (1982) Physiological characterization of Saccharomyces cerevisiae mutants supersensitive to G1 arrest by a factor and alpha factor pheromones. *Mol. Cell. Biol.* **2**, 21–29
12. Koelle, M. R., and Horvitz, H. R. (1996) EGL-10 regulates G protein signaling in the C. elegans nervous system and shares a conserved domain with many mammalian proteins. *Cell*. **84**, 115–125
13. Tesmer, J. J. J., Berman, D. M. D., Gilman, A. G. A., and Sprang, S. R. S.

(1997) Structure of RGS4 bound to AlF4--activated G(i alpha1): stabilization of the transition state for GTP hydrolysis. *Cell*. **89**, 251–261
14. Dohlman, H. G. (2009) RGS proteins the early days. *Prog Mol Biol Transl Sci*. **86**, 1–14
15. Neubig, R. R., and Siderovski, D. P. (2002) Regulators of G-protein signalling as new central nervous system drug targets. *Nature Reviews Drug Discovery*. **1**, 187–197
16. Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190
17. Soundararajan, M., Willard, F. S., Kimple, A. J., Turnbull, A. P., Ball, L. J., Schoch, G. A., Gileadi, C., Fedorov, O. Y., Dowler, E. F., Higman, V. A., Hutsell, S. Q., Sundström, M., Doyle, D. A., and Siderovski, D. P. (2008) Structural diversity in the RGS domain and its interaction with heterotrimeric G protein alpha-subunits. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 6457–6462
18. Bernstein, L. S., Grillo, A. A., Loranger, S. S., and Linder, M. E. (2000) RGS4 binds to membranes through an amphipathic alpha -helix. *J. Biol. Chem.* **275**, 18520–18526
19. Zeng, W., Xu, X., Popov, S., Mukhopadhyay, S., Chidiac, P., Swistok, J., Danho, W., Yagaloff, K. A., Fisher, S. L., Ross, E. M., Muallem, S., and Wilkie, T. M. (1998) The N-terminal domain of RGS4 confers receptor-selective inhibition of G protein signaling. *J. Biol. Chem.* **273**, 34687–34690
20. Moy, F. J., Haraki, K., Mobilio, D., Walker, G., Powers, R., Tabei, K., Tong, H., and Siegel, M. M. (2001) MS/NMR: A Structure-Based Approach for Discovering Protein Ligands and for Drug Design by Coupling Size Exclusion Chromatography, Mass Spectrometry, and Nuclear Magnetic Resonance Spectroscopy. *Anal. Chem.* **73**, 571–581
21. Moy, F. J., Chanda, P. K., Cockett, M. I., Edris, W., Jones, P. G., Mason, K., Semus, S., and Powers, R. (2000) NMR structure of free RGS4 reveals an induced conformational change upon binding Galpha. *Biochemistry*. **39**, 7063–7073
22. Nance, M. R., Kreutz, B., Tesmer, V. M., Sterne-Marr, R., Kozasa, T., and Tesmer, J. J. G. (2013) Structural and functional analysis of the regulator of G protein signaling 2-gαq complex. *Structure*. **21**, 438–448
23. Kimple, A. J., Soundararajan, M., Hutsell, S. Q., Roos, A. K., Urban, D. J., Setola, V., Temple, B. R. S., Roth, B. L., Knapp, S., Willard, F. S., and Siderovski, D. P. (2009) Structural Determinants of G-protein α Subunit Selectivity by Regulator of G-protein Signaling 2 (RGS2). *Journal of Biological Chemistry*. **284**, 19402–19411
24. Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006) How many

drug targets are there? *Nature Reviews Drug Discovery*. **5**, 993–996
25. Zhang, P., and Mende, U. (2011) Regulators of G-protein signaling in the heart and their potential as therapeutic targets. *Circ. Res.* **109**, 320–333
26. Roman, D. L., and Traynor, J. R. (2011) Regulators of G Protein Signaling (RGS) Proteins as Drug Targets: Modulating G-Protein-Coupled Receptor (GPCR) Signal Transduction
27. Aosaki, T., Miura, M., Suzuki, T., Nishimura, K., and Masuda, M. (2010) Acetylcholine-dopamine balance hypothesis in the striatum: An update. *Geriatrics & Gerontology International*. **10**, S148–S157
28. Ding, J., Guzman, J. N., Tkatch, T., Chen, S., Goldberg, J. A., Ebert, P. J., Levitt, P., Wilson, C. J., Hamm, H. E., and Surmeier, D. J. (2006) RGS4-dependent attenuation of M4 autoreceptor function in striatal cholinergic interneurons following dopamine depletion. *Nat. Neurosci.* **9**, 832–842
29. Lerner, T. N., and Kreitzer, A. C. (2012) RGS4 is required for dopaminergic control of striatal LTD and susceptibility to Parkinsonian motor deficits. *Neuron*. **73**, 347–359
30. Popov, S. G., Krishna, U. M., Falck, J. R., and Wilkie, T. M. (2000) Ca2+/Calmodulin reverses phosphatidylinositol 3,4, 5-trisphosphate-dependent inhibition of regulators of G protein-signaling GTPase-activating protein activity. *J. Biol. Chem.* **275**, 18962–18968
31. Storaska, A. J. (2014) *Molecular Mechanisms for Inhibition of Regulators of G-protein Signaling by Small Molecules*. Ph.D. thesis, University of Michigan, Ann Arbor
32. Chisari, M., Saini, D. K., Kalyanaraman, V., and Gautam, N. (2007) Shuttling of G protein subunits between the plasma membrane and intracellular membranes. *J. Biol. Chem.* **282**, 24092–24098
33. Osterhout, J. L., Waheed, A. A., Hiol, A., Ward, R. J., Davey, P. C., Nini, L., Wang, J., Milligan, G., Jones, T. L. Z., and Druey, K. M. (2003) Palmitoylation Regulates Regulator of G-protein Signaling (RGS) 16 Function: II. palmitoylation of a cysteine residue in the RGS box is critical for RGS16 GTPase accelerating activity and regulation of Gi-coupled signaling. *Journal of Biological Chemistry*. **278**, 19309–19316
34. Tu, Y., Popov, S., Slaughter, C., and Ross, E. M. (1999) Palmitoylation of a Conserved Cysteine in the Regulator of G Protein Signaling (RGS) Domain Modulates the GTPase-activating Activity of RGS4 and RGS10. *Journal of Biological Chemistry*. **274**, 38260–38267
35. Roman, D. L., Talbot, J. N., Roof, R. A., Sunahara, R. K., Traynor, J. R., and Neubig, R. R. (2007) Identification of small-molecule inhibitors of RGS4 using a high-throughput flow cytometry protein interaction assay. *Mol Pharmacol*. **71**, 169–175

36. Roman, D. L., Ota, S., and Neubig, R. R. (2009) Polyplexed Flow Cytometry Protein Interaction Assay: A Novel High-Throughput Screening Paradigm for RGS Protein Inhibitors. *Journal of Biomolecular Screening*. **14**, 610–619
37. Blazer, L. L., Roman, D. L., Chung, A., Larsen, M. J., Greedy, B. M., Husbands, S. M., and Neubig, R. R. (2010) Reversible, allosteric small-molecule inhibitors of regulator of G protein signaling proteins. *Mol Pharmacol*. **78**, 524–533
38. Blazer, L. L., Zhang, H., Casey, E. M., Husbands, S. M., and Neubig, R. R. (2011) A nanomolar-potency small molecule inhibitor of regulator of G-protein signaling proteins. *Biochemistry*. **50**, 3181–3192
39. Storaska, A. J., Mei, J. P., Wu, M., Li, M., Wade, S. M., Blazer, L. L., Sjögren, B., Hopkins, C. R., Lindsley, C. W., Lin, Z., Babcock, J. J., McManus, O. B., and Neubig, R. R. (2013) Reversible inhibitors of regulators of G-protein signaling identified in a high-throughput cell-based calcium signaling assay. *Cell. Signal.* **25**, 2848–2855
40. Kimple, A. J., Willard, F. S., Giguère, P. M., Johnston, C. A., Mocanu, V., and Siderovski, D. P. (2007) The RGS protein inhibitor CCG-4986 is a covalent modifier of the RGS4 Galpha-interaction face. *Biochim. Biophys. Acta*. **1774**, 1213–1220
41. Roman, D. L., Blazer, L. L., Monroy, C. A., and Neubig, R. R. (2010) Allosteric inhibition of the regulator of G protein signaling-Galpha protein-protein interaction by CCG-4986. *Mol Pharmacol*. **78**, 360–365
42. Vashisth, H., Storaska, A. J., Neubig, R. R., and Brooks, C. L. (2013) Conformational dynamics of a regulator of G-protein signaling protein reveals a mechanism of allosteric inhibition by a small molecule. *ACS Chem. Biol.* 10.1021/cb400568g
43. Turner, E. M., Blazer, L. L., Neubig, R. R., and Husbands, S. M. (2012) Small Molecule Inhibitors of Regulators of G Protein Signaling (RGS) Proteins. *ACS Med. Chem. Lett.* **3**, 146–150
44. Blazer, L. L., Storaska, A. J., Jutkiewicz, E. M., Turner, E. M., Calcagno, M., Wade, S. M., Wang, Q., Huang, X.-P., Traynor, J. R., Husbands, S. M., Morari, M., and Neubig, R. R. (2015) Selectivity and Anti-Parkinson's Potential of Thiadiazolidinone RGS4 Inhibitors. *ACS Chem. Neurosci.* **6**, 911–919
45. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J., and Laue, E. D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*. **59**, 687–696
46. Hafsa, N., and Wishart, D. (2014) CSI 2.0: a significantly improved

version of the Chemical Shift Index. *J. Biomol. NMR*. **60**, 131–146
47. Camilloni, C., De Simone, A., Vranken, W. F., and Vendruscolo, M. (2012) Determination of Secondary Structure Populations in Disordered States of Proteins Using Nuclear Magnetic Resonance Chemical Shifts. *Biochemistry*. **51**, 2224–2231
48. Kleckner, I. R., and Foster, M. P. (2011) An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. **1814**, 942–968
49. Monroy, C. A., Mackie, D. I., and Roman, D. L. (2013) A High Throughput Screen for RGS Proteins Using Steady State Monitoring of Free Phosphate Formation. *PLoS One*. **8**, e62247
50. Baum, J., Dobson, C. M., Evans, P. A., and Hanley, C. (1989) Characterization of a partly folded protein by NMR methods: studies on the molten globule state of guinea pig alpha-lactalbumin. *Biochemistry*. **28**, 7–13
51. Greene, L. H., Wijesinha-Bettoni, R., and Redfield, C. (2006) Characterization of the Molten Globule of Human Serum Retinol-Binding Protein Using NMR Spectroscopy. *Biochemistry*. **45**, 9475–9484
52. Dyson, H. J., and Wright, P. E. (2001) Nuclear magnetic resonance methods for elucidation of structure and dynamics in disordered states. *Meth. Enzymol.* **339**, 258–270
53. Redfield, C. (2004) Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins. *Methods*. **34**, 121–132
54. Kohn, J. E., Millett, I. S., Jacob, J., Zagrovic, B., Dillon, T. M., Cingel, N., Dothager, R. S., Seifert, S., Thiyagarajan, P., Sosnick, T. R., Hasan, M. Z., Pande, V. S., Ruczinski, I., Doniach, S., and Plaxco, K. W. (2004) Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 12491–12496
55. Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999) Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques. *Biochemistry*. **38**, 16424–16431
56. Erickson, H. P. (2009) Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy. *Biological Procedures Online*. **11**, 32
57. Kågedal, L., Engström, B., Elelgren, H., Lieber, A.-K., Lundström, H., Sköld, A., and Schenning, M. (1991) Chemical, physical and chromatographic properties of Superdex 75 prep grade and Superdex 200 prep grade gel filtration media. *J Chromatogr A*. **537**, 17–32
58. Uversky, V. N. (1993) Use of fast protein size-exclusion liquid

chromatography to study the unfolding of proteins which denature through the molten globule. *Biochemistry*. **32**, 13288–13298
59. Murphy, R. M. (1997) Static and dynamic light scattering of biological macromolecules: what can we learn? *Cell Metab*. **8**, 25–30
60. Kelly, S. M., Jess, T. J., and Price, N. C. (2005) How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. **1751**, 119–139
61. Sreerama, N., and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–260
62. le Maire, M., Arnou, B., Olesen, C., Georgin, D., Ebel, C., and Møller, J. V. (2008) Gel chromatography and analytical ultracentrifugation to determine the extent of detergent binding and aggregation, and Stokes radius of membrane proteins using sarcoplasmic reticulum Ca2+|[ndash]|ATPase as an example. *Nat Protoc*. **3**, 1782–1795
63. Demarest, S. J. (2004) Packing, specificity, and mutability at the binding interface between the p160 coactivator and CREB-binding protein. *Protein Science*. **13**, 203–210
64. Kjaergaard, M., Teilum, K., and Poulsen, F. M. (2010) Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP. *PNAS*. **107**, 12535–12540
65. Englander, S. W., Mayne, L., Bai, Y., and Sosnick, T. R. (1997) *Hydrogen exchange: the modern legacy of Linderstrøm-Lang.*, **6**, 1101–1109
66. Feng, Z., Butler, M. C., Alam, S. L., and Loh, S. N. (2001) On the nature of conformational openings: native and unfolded-state hydrogen and thiol-disulfide exchange studies of ferric aquomyoglobin 1. *Journal of Molecular Biology*. **314**, 153–166
67. Isom, D. G., Vardy, E., Oas, T. G., and Hellinga, H. W. (2010) Picomole-scale characterization of protein stability and function by quantitative cysteine reactivity. *PNAS*. **107**, 4908–4913
68. Liu, D. F., Wang, D., and Stracher, A. (1990) The accessibility of the thiol groups on G- and F-actin of rabbit muscle. *Biochem. J.* **266**, 453–459
69. Johnson, I., and Spence, M. (2010) The molecular probes handbook. *Life Technologies Corporation*
70. Boehr, D. D., Nussinov, R., and Wright, P. E. (2009) The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796
71. Krishna, M. M. G., Hoang, L., Lin, Y., and Englander, S. W. (2004) Hydrogen exchange methods to study protein folding. *Methods*. **34**, 51–64

72. Hwang, T.-L., van Zijl, P. M., and Mori, S. (1998) Accurate Quantitation of Water-amide Proton Exchange Rates Using the Phase-Modulated CLEAN Chemical EXchange (CLEANEX-PM) Approach with a Fast-HSQC (FHSQC) Detection Scheme. *J. Biomol. NMR*. **11**, 221–226–226
73. Bertini, I., Ghosh, K., Rosato, A., and Vasos, P. R. (2003) A high-resolution NMR study of long-lived water molecules in both oxidation states of a minimal cytochrome c. *Biochemistry*. **42**, 3457–3463
74. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins*. **17**, 75–86
75. Zhang, Y.-Z. (1995) *Protein and peptide structure and interactions studied by hydrogen exchange and NMR*. Ph.D. thesis, University of Pennsylvania, Ann Arbor
76. Xu, J., Lee, Y., Beamer, L. J., and Van Doren, S. R. (2015) Phosphorylation in the catalytic cleft stabilizes and attracts domains of a phosphohexomutase. *Biophys. J.* **108**, 325–337
77. Sandhu, K. S., and Dash, D. (2007) Dynamic α-helices: Conformations that do not conform. *Proteins: Structure, Function, and Bioinformatics*. **68**, 109–122
78. Gautier, R., Douguet, D., Antonny, B., and Drin, G. (2008) HELIQUEST: a web server to screen sequences with specific alpha-helical properties. *Bioinformatics*. **24**, 2101–2102
79. Fauchere, J. L., and Pliska, V. (1983) Hydrophobic parameters-pi of amino-acid side-chains from the partitioning of n-acetyl-amino-acid amides. *European Journal of Medicinal Chemistry*. **18**, 369–375
80. Reiersen, H., and Rees, A. R. (2001) The hunchback and its neighbours: proline as an environmental modulator. *Trends Biochem. Sci.* **26**, 679–684
81. Wedemeyer, W. J., Welker, E., and Scheraga, H. A. (2002) Proline cis-trans isomerization and protein folding. *Biochemistry*. **41**, 14637–14644
82. Lebbink, J. H. G., Consalvi, V., Chiaraluce, R., Berndt, K. D., and Ladenstein, R. (2002) Structural and thermodynamic studies on a salt-bridge triad in the NADP-binding domain of glutamate dehydrogenase from Thermotoga maritima: cooperativity and electrostatic contribution to stability. *Biochemistry*. **41**, 15524–15535
83. Monroy, C. A., Doorn, J. A., and Roman, D. L. (2013) Modification and Functional Inhibition of Regulator of G-Protein Signaling 4 (RGS4) by 4-Hydroxy-2-nonenal. *Chem. Res. Toxicol.* **26**, 1832–1839
84. Teague, S. J. (2003) Implications of protein flexibility for drug discovery. *Nature Reviews Drug Discovery*. **2**, 527–541
85. Hiol, A., Davey, P. C., Osterhout, J. L., Waheed, A. A., Fischer, E. R., Chen, C. K., Milligan, G., Druey, K. M., and Jones, T. L. Z. (2003)

Palmitoylation Regulates Regulators of G-protein Signaling (RGS) 16 Function: I. mutation of amino-terminal cysteine residues on RGS16 prevents its targeting to lipid rafts and palmitoylation of an internal cysteine residue. *Journal of Biological Chemistry*. **278**, 19301–19308

86. Tesmer, J. J. G. (2009) Structure and function of regulator of G protein signaling homology domains. *Prog Mol Biol Transl Sci*. **86**, 75–113
87. Yan, S., Kennedy, S. D., and Koide, S. (2002) Thermodynamic and Kinetic Exploration of the Energy Landscape of Borrelia burgdorferi OspA by Native-state Hydrogen Exchange. *Journal of Molecular Biology*. **323**, 363–375
88. Fuentes, E. J., and Wand, A. J. (1998) Local dynamics and stability of apocytochrome b562 examined by hydrogen exchange. *Biochemistry*. **37**, 3687–3698
89. Vallurupalli, P., Hansen, D. F., and Kay, L. E. (2008) Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy. *PNAS*. **105**, 11766–11771
90. Baldwin, A. J., and Kay, L. E. (2009) NMR spectroscopy brings invisible protein states into focus. *Nat. Methods*. **5**, 808–814
91. Mulder, F. A., Mittermaier, A., Hon, B., Dahlquist, F. W., and Kay, L. E. (2001) Studying excited states of proteins by NMR spectroscopy. *Nat. Struct. Mol. Biol.* **8**, 932–935
92. Savitsky, P., Bray, J., Cooper, C. D. O., Marsden, B. D., Mahajan, P., Burgess-Brown, N. A., and Gileadi, O. (2010) High-throughput production of human proteins for crystallization: The SGC experience. *Journal of Structural Biology*. **172**, 3–13
93. Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., and Li, Q. (2009) Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. *Protein Sci.* **18**, 936–948
94. Higgins, C., Vermeer, L. M., Doorn, J. A., and Roman, D. L. (2012) Expression and purification of recombinant human tyrosine hydroxylase as a fusion protein in Escherichia coli. *Protein Expr Purif*. **84**, 219–223
95. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*. **6**, 277–293
96. Helmus, J. J., and Jaroniec, C. P. (2013) Nmrglue: an open source Python package for the analysis of multidimensional NMR data. *J. Biomol. NMR*. **55**, 355–367
97. Whitmore, L., and Wallace, B. A. (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers*. **89**, 392–400

98. Sreerama, N., Venyaminov, S. Y., and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis. *Anal. Biochem.* **287**, 243–251
99. Abdul-Gader, A., Miles, A. J., and Wallace, B. A. (2011) A reference dataset for the analyses of membrane protein secondary structures and transmembrane residues using circular dichroism spectroscopy. *Bioinformatics*. **27**, 1630–1636
100. Klose, D. P., Wallace, B. A., and Janes, R. W. (2010) 2Struc: the secondary structure server. *Bioinformatics*. **26**, 2624–2625