REGULATION OF GUANINE NUCELOTIDE EXCHANGE IN INHIBITORY G PROTEIN ALPHA SUBUNIT BY ACTIVATOR OF G PROTEIN SIGNALING 3 AND NOVEL REGULATORY PEPTIDES

APPROVED BY SUPERVISORY COMMITTEE

Mentor: Stephen R. Sprang, Ph. D.

Committee Chairperson: Kevin H. Gardner, Ph. D.

Committee Member: Elliott M. Ross, Ph. D.

Committee Member: Johann Deisenhofer, Ph. D.

DEDICATION

To my parents Tulsi Prasad and Sunanda Adhikari

REGULATION OF GUANINE NUCELOTIDE EXCHANGE IN INHIBITORY G
PROTEIN ALPHA SUBUNIT BY ACTIVATOR OF G PROTEIN SIGNALING 3 AND
NOVEL REGULATORY PEPTIDES

by

ANIRBAN ADHIKARI

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

November, 2005

ACKNOWLEDGEMENTS

running my MD simulations. Every member of the Sprang lab, both current and former, has helped in many ways: TC, Xinlin, Dario and Aurora. I thank them as well.

My research benefited a lot from very effective collaborations. I would like to thank Prof. Stephen Lanier for motivating us to work on AGS3 and giving us numerous constructs. I would like to thank Bill in Dr. Richard Robert's lab for encouraging us to investigate R6A and collaborating on the project. Bill has always been very enthusiastic to answer my queries about R6A.

I would like thank my parents and my brother for their affection and encouragement. They always had big hopes for me and encouraged me to do what I want. Lastly, I would like thank my wife, Mekhala, for her love and emotional support. I thank her for helping me deal with the ups and downs of life and being a great companion in graduate school.

REGULATION OF GUANINE NUCELOTIDE EXCHANGE IN INHIBITORY G-PROTEIN ALPHA SUBUNIT BY ACTIVATOR OF G-PROTEIN SIGNALING 3 AND NOVEL REGULATORY PEPTIDES

Publication No. _____

Anirban Adhikari, Ph. D.

The University of Texas Southwestern Medical Center at Dallas, 2005

Supervising Professor: Stephen R. Sprang, Ph. D.

The release of GDP from the nucleotide binding pocket of G protein α subunit (Gα) is accelerated by guanine nucleotide exchange factors (GEFs) and inhibited by guanine nucleotide dissociation inhibitors (GDIs). The $\beta\gamma$ subunit of heterotrimeric G protein and GoLoco motif are GDIs. Activator of G protein signaling 3 (AGS3), a GDI containing four GoLoco motifs in its C-terminal domain (AGS3-C), binds specifically to GDP-bound Gα_{i1} and inhibits the binding of GTP to Gα_{i1}. The stoichiometry and dissociation constant for the binding of AGS3-C to Gα_{i1}, determined using isothermal titration calorimetry (ITC),

indicates the presence of two apparent high affinity ($K_d \sim 20$ nM) and two apparent low affinity ($K_d \sim 300$ nM) binding sites for $G\alpha_{i1}$. Upon deletion of the C-terminal GoLoco motif from AGS3-C, the remaining sites are approximately equivalent with respect to their affinity ($K_d \sim 400$ nM) for $G\alpha_{i1}$. Peptides corresponding to each of the four GoLoco motifs of AGS3 (referred to as GPR1-4, going from the N to C terminus) bind to $G\alpha_{i1}$ with K_d values in the range of 1–8 μM. GPR1, GPR2, and GPR4 were active as GDIs, but GPR3 was inactive. However, addition of N and C terminal flanking residues to the GPR3 GoLoco core increases its affinity for $G\alpha_{i1}$ and conferred GDI activity to the level of AGS3 itself.

R6A is a $G\alpha_{i1}$·GDP selective peptide that has GDI activity towards $G\alpha_{i1}$ despite no sequence similarity with the GoLoco motif. The 1.8 Å crystal structure of $G\alpha_{i1}$·GDP:R6A complex shows that switch II region of $G\alpha_{i1}$, which is disordered in the GDP bound state, adopts an ordered structure through its direct interactions with R6A. Binding of R6A orients residues 202 to 204 of switch II (GG-loop) towards the nucleotide binding pocket and within 3 Å of switch I, which suggests that R6A mediates its activity by occluding the potential exit route of GDP. The structure of $G\alpha_{i1}$·GDP:R6A, when compared to the structures of GoLoco and $G\beta\gamma$ bound $G\alpha_{i1}$·GDP, reveals similarity in the conformation of GG-loop among the GDI bound structures of $G\alpha_{i1}$. A mutation in R6A, designed to position the GG-loop away from the nucleotide binding pocket in $G\alpha_{i1}$·GDP:R6A complex, abolishes the GDI activity and instead confers GEF like activity to R6A. Thus, the conformation of the GG-loop is likely to be a crucial structural determinant of the rate of GDP release from $G\alpha_{i1}$.

TABLE OF CONTENTS

x

PRIOR PUBLICATIONS

1. Adhikari A and Sprang, SR. (2003). Thermodynamic characterization of the binding of activator of G protein signaling 3 (AGS3) and peptides derived from AGS3 with G alpha i1. *J Biol Chem*. **278**, 51825-32.

2. Ja WW, Adhikari A, Austin RJ, Sprang SR, and Roberts RW. (2005). A peptide core motif for binding to heterotrimeric G protein alpha subunits. *J Biol Chem*. **280**, 32057-60.

3. Adhikari A and Sprang, SR. (2005). Structural determinants of guanine nucleotide exchange in G alpha i1. (manuscript under preparation)

LIST OF FIGURES

xviii

LIST OF TABLES

List of Abbreviations

7-AW   7-azatryptophan

A$_{280}$ Absorbance at 280 nm wavelength

AGS3   Activator of G protein Signaling 3

AGS3-C   Residues 465-650 of AGS3

AGS3-Δ34   Residues 465-548 of AGS3

AGS3-Δ4   Residues 465-597 of AGS3

βγ   β and γ subunits of heterotrimeric G protein

BODIPY   4,4-difluoro-4-bora-3a,4a-diaza-s-indacene

cAMP   Cyclic adenosine monophosphate

CD   Circular dichroism

ΔG   Free energy

ΔH   Enthalpy

ΔN-Gα$_{i1}$ Residues 31-354 of Gα$_{i1}$

ΔS   Entropy

DTT   1,4-Dithiothreitol

EDTA   Ethylenediamine tetra-acetic acid

EPPS   N-(2-Hydroxyethyl)piperazine-N′-(3-propanesulfonic acid)

FPLC   Fast protein liquid chromatography

Gα$_{i1}$ Inhibitory G-protein alpha subunit 1

GAP	GTPase activating protein
GDI	Guanine nucleotide dissociation inhibitor
GDP	Guanosine 5'-diphosphate
GEF	Guanine nucleotide exchange factor
GG-loop	A segment of switch II encompassing residues Gly202-Gln204
GMPPNP	Guanosine 5'-[β,γ-imido] triphosphate
GPCR	G protein coupled receptor
GPR	G protein regulatory
GPRex	Extended GPR peptides
GST	Glutathione S-transferase
GST-X	GST-fusion form of protein/peptide X
GTP	Guanosine 5'-triphosphate
GTPγS	Guanosine 5'-O-(3-thiotriphosphate)
HAP	Ceramic hydroxyapatite
HEPES	4-(2-hydroxyethyl)-1-piperizine-ethanesulfonic acid
HPLC	High performance liquid chromatography
IC$_{50}$	Half-maximal inhibitory concentration
IPTG	Isopropyl-β-D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
K$_a$	Association constant
K$_d$	Dissociation constant
kDa	Kilo Dalton

λem	Emission wavelength
λex	Excitation wavelength
LB	Luria broth
MANT	2'(3')-O-N-methylanthraniloyl
Ni-NTA	Ni^{2+}-nitro-triacetic acid
OD600	Optical density at 600 nm wavelength
PAGE	Poly-acrylamide gel electrophoresis
PEG	Poly-ethylene glycol
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulphonylfluoride
PTT	Protease inhibitor cocktail containing PMSF, TPCK, and TLCK(1000X PTT contains 23 mg/ml PMSF, 21 mg/ml TLCK, and 21 mg/ml TPCK)
R6A	9 residue R6A peptide core region: DQLYWWEYL
R6A4	16 residues peptide: SQTKRLDDQLYWWEYL
R6M1	11 residue modified R6A peptide: DQLYWWEYLKK
R6M2	13 residue modified R6A peptide: KRLDDQLYWWEYL
RGS	Regulator of G protein signaling
rpm	Revolution per minute
SDS	sodium dodecyl sulfate
TEV	Tobacco etch virus
TLCK	Nα-p-tosyl-L-lysyl chloromethyl ketone
TPCK	N-tosyl-L-phenylalanine chloromethyl ketone

Tris   Tris(hydroxymethyl) aminomethane

CRYSTALLOGRAPHIC TERMS AND ABBREVIATIONS

B factor   Temperature factor

F$_c$ Calculated structure factor amplitude

F$_o$ Observed structure factor amplitude

I   Intensity of reflections

λ Wavelength

rmsd   Root mean square deviation (Å)

σ standard deviation

V$_m$ Matthews coefficient (Å^3/Dalton)

The crystallographic **structure factor** is calculated from a model using individual atomic form factors f for N atoms:

$$F(hkl) = \sum^{N} f.\exp[2\pi i(hx + ky + lz)]$$

Electron density is calculated from phased structure factors:

$$\rho(xyz) = \frac{1}{V}\sum_{hkl} F(hkl).\exp[-2\pi i(hx + ky + lz)]$$

Rsym is the R factor for comparing the intensities of symmetry related reflections:

$$R_{sym} = \frac{\sum_{hkl}\sum_{i}\left|I_i(hkl) - \overline{I(hkl)}\right|}{\sum_{hkl}\sum_{i}I_i(hkl)}$$

where $I_i(hkl)$ and $<I(hkl)>$ are the i_{th} and mean measurement of the intensity of

reflection hkl.

Patterson function is the Fourier summation using intensities as coefficients:

$$P(uvw) = \frac{1}{V} \sum_{hkl} |F(hkl)|^2 \cos[2\pi(hu + kv + lw)]$$

R factor is the crystallographic residual comparing a model and X-ray data:

$$R = \frac{\sum_{hkl} \left\| F_{obs} \right| - \left| F_{calc} \right\|}{\sum_{hkl} F_{obs}}$$

where Fobs and Fcalc are the observed and calculated structure factors,

respectively.

R$_{free}$ is the R factor value obtained for a test set of reflections consisting of a

randomly selected percentage of the data.

<u>Chapter One</u>

Introduction

The viability of a living cell depends on its ability to adapt to environmental changes outside its cellular boundaries. To achieve this task, cells have evolved capabilities to sense changes in the extracellular environment and modify intracellular processes to adapt to the environmental conditions. Cells perceive extracellular environment as signals. Signals include chemical entities like hormones and nutrients or physical forces like light, sound, pressure, and heat. Cells respond to these extracellular signals through a large number of membrane associated proteins or receptors. These receptors respond to signals and transmit information to the inside of the cell through a vast array of complex signal transduction mechanisms. Cells have also evolved intricate intracellular signaling networks that can communicate with receptors and regulate intracellular pathways. This complex interplay of extracellular signal and intracellular signaling processes regulate virtually every aspect of cellular behavior like metabolism, growth, differentiation, cell division and apoptosis. In addition of responding to environmental signals, multi-cellular organisms have the need to regulate the behavior of individual cells in a manner that is beneficial to the survival and reproducibility of the organism as a whole. The social behavior of cells in multi-cellular organisms is regulated by elaborate networks of intra and extracellular communications. Failures in the cellular communication networks due to genetic changes or pathogenic

assaults usually lead to diseases. Understanding of the molecular basis of these signaling networks has been the prime focus of biomedical research in the last several decades.

1.1   Heterotrimeric G proteins

1.1.1   The G protein cycle

G protein signaling forms an important part of the signaling network in eukaryotic cells. Heterotrimeric G proteins consist of α,β and γ subunits (Figure 1.1). The α subunit is catalytically active and posses GTPase activity, while the β and γ subunits act as an obligate heterodimer with no intrinsic catalytic activity. The α subunits are guanine nucleotide dependent molecular switches; alternating between a GDP-bound inactive state and GTP-bound active state. In the canonical G protein cycle, agonist bound, seven-transmembrane, G protein coupled receptors (GPCR) catalyze the exchange of GDP with GTP at the nucleotide binding pocket of α subunit. Thus activated, Gα subunits dissociate from the hetrodimeric complex of G$\beta\gamma$ (Gilman, 1987; Sprang, 1997a) (Figure 1.1). GTP-bound active α subunits modulate the activity of effector molecules, including adenylyl cyclases, phospholipase Cβ and nucleotide exchange factors for the small GTPase Rho (Hamm, 1998; Sprang, 1997a) . Signaling is terminated by the intrinsic GTPase activity of Gα, which hydrolyzes the bound GTP to GDP. Upon hydrolysis of GTP, the GDP-bound inactive α subunit re-associates with the G$\beta\gamma$ to form the inactive heterotrimer (Gilman, 1987) (Figure 1.1).
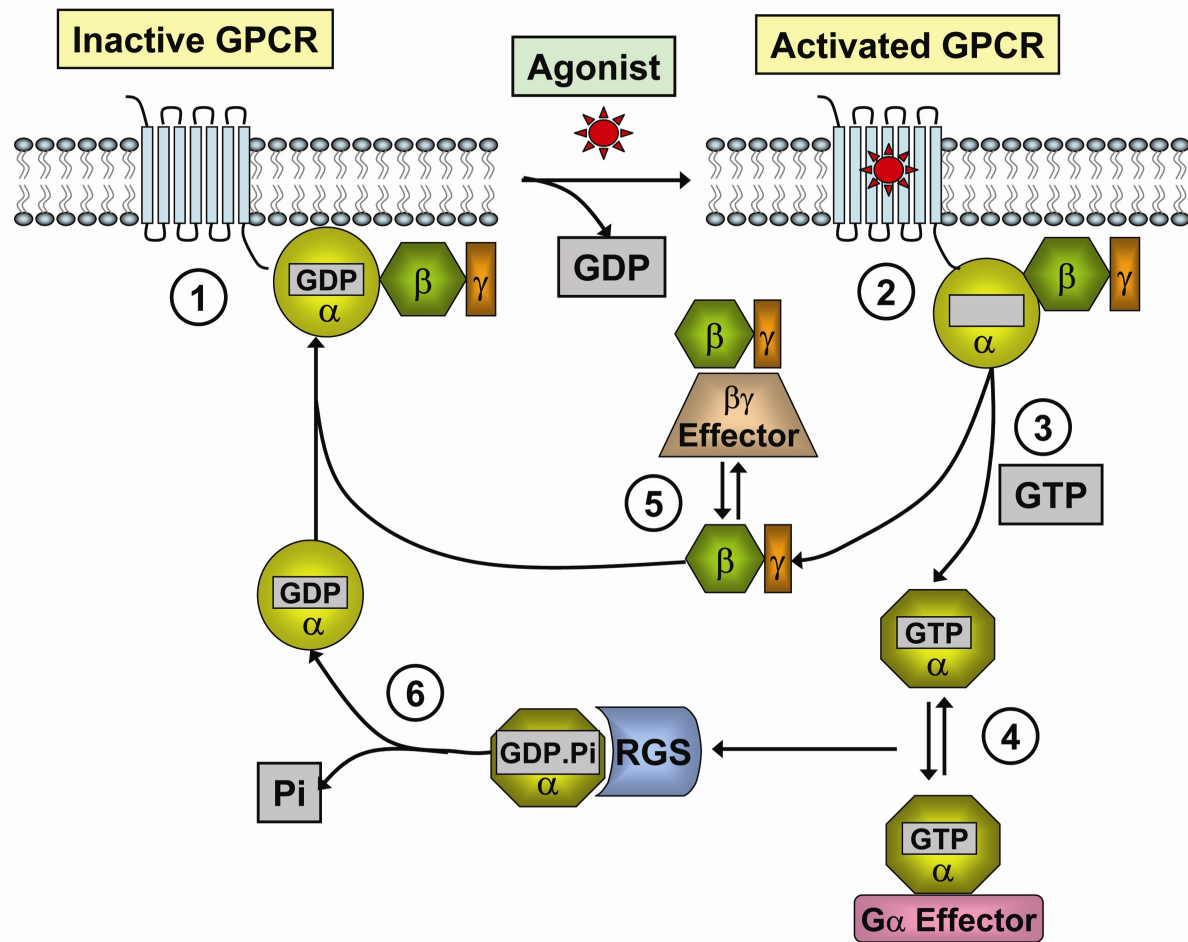
Figure 1.1. The heterotrimeric G-protein cycle. Schematic illustration of heterotrimeric G protein cycle along with its regulators is shown: 1. Inactive heterotrimeric G protein complex consisting of Gα·GDP and Gβγ subunits. 2. Binding of agonist activates GPCR; activated GPCR interacts with the inactive heterotrimer and promotes the release of GDP from the α subunit. 3. Interaction with GPCR leads to the formation of activated Gα·GTP and free βγ subunit. 4 and 5. Gα·GTP and Gβγ interact with their effector molecules. 6. GTP is hydrolyzed into GDP and phosphate (Pi) in the α subunit by itself or with the help of RGS proteins. Pi is released and Gα·GDP re-associates with Gβγ to form the inactive heterotrimer.

1.1.2   Regulators of G protein signaling: GAPs, GEFs, and GDIs

The reaction cycle of heterotrimeric G proteins is regulated by several classes of

molecules; these molecules achieve regulation by modulating two distinct processes: the rate

of GTP hydrolysis and the rate of nucleotide exchange. GTPase Activating Proteins (GAPs) enhance the rate of GTP hydrolysis in Gα subunits by stabilizing a transition state intermediate of the hydrolysis reaction (Ross and Wilkie, 2000; Sprang, 1997a; Sprang, 1997b). RGS proteins, which are GAPs, are among the most familiar modulators of G-protein signaling (Ross and Wilkie, 2000). The intrinsic rate of GTP hydrolysis by Gα subunits is slow; GAPs deactivate Gα by accelerating the rate of GTP hydrolysis (Berman and Gilman, 1998). Apart from hydrolysis, regulation of G proteins is also achieved by modulating the rate of nucleotide exchange. Two classes of regulators influence the process of guanine nucleotide exchange. Guanine nucleotide Exchange Factors (GEFs) accelerate the rate of GDP to GTP exchange in G proteins (Sprang, 1997a). Until the finding that RIC-8 is a cytosolic GEF of Gα subunits, membrane bound GPCRs were the only known GEFs for Gα (Tall et al., 2003). In the absence of receptor activation, the intrinsic rate of nucleotide exchange is slow in Gα subunits (~0.03 min^{-1} at 25 $^{\circ}$C), and almost completely stopped when Gα is in complex with Gβγ (Higashijima et al., 1987; Sprang, 1997a). The other class of modulator of guanine nucleotide exchange, known as Guanine nucleotide Dissociation Inhibitors (GDIs), bind specifically to the GDP bound state of α subunits (or small G proteins) and inhibit the dissociation of GDP from the α subunits by stabilizing the GDP bound state. The βγ heterodimer is a GDI (Sprang, 1997a). In recent years, a new class of GDI has been discovered, typified by Activator of G-protein Signaling 3 (AGS3). Like Gβγ, AGS3 binds specifically to the GDP bound state of Gα and acts as a GDI (De Vries et al.,

2000a; Lanier, 2004). The presence of GEF, GDI, and GAPs allow precise and fast regulation of G-protein signaling.

1.1.3 G protein α and $\beta\gamma$ subunits

Heterotrimeric G proteins form a large family of regulatory proteins. As of this date, more than 16 different genes have been identified that encode 20 different Gα subunits in mammalian organisms (Cabrera-Vera et al., 2003; Sprang, 1997a). Similar diversity is also observed for the G$\beta\gamma$ subunits; 5 genes encode Gβ and 12 genes encode Gγ subunits (Cabrera-Vera et al., 2003).

Based on sequence similarity of the α subunits, G proteins are divided into 4 different classes: 1. Gα_s (Gα_s, Gα_{olf}); 2. G$\alpha_{i/o}$ (Gα_{i1}, Gα_{i2}, Gα_{i3}, Gα_t, Gα_z, Gα_g, Gα_o); 3. Gα_q (Gα_q, Gα_{11}, Gα_{14-16}) ; 4. G$\alpha_{12/13}$ (Cabrera-Vera et al., 2003). All α subunits contain two distinct domains; the Ras-like GTPase domain and the helical domain. The nucleotide binding pocket is formed in a cleft between these domains. The α subunits typically undergo lipid modifications by either palmitoylation or N-myristoylation. These modifications result in membrane localization of α subunits (Sprang, 1997a). The activated α subunits interact with several effectors; stimulatory G proteins like Gα_s activate adenylyl cyclases and increase the production of cAMP, while inhibitory G-proteins like Gα_i inhibit cAMP production. Activated Gα_q up-regulates PLCβs; Gα_{13} can activate p115RhoGEF (a GEF for small G-protein Rho) (Cabrera-Vera et al., 2003; Hamm, 1998; Sprang, 1997a).

The β and γ subunits of heterotrimeric G proteins behave as obligate heterodimers. The β subunit is a WD-40 repeat containing protein; seven WD-40 repeats form a β-propeller domain. The γ subunit is a small protein and stays bound to the β subunit through a coiled-coil interaction (Wall et al., 1995). Much like the α subunit, the βγ subunit also undergoes lipid modification and localizes in membrane (Sprang, 1997a). Upon dissociation from the inactive heterotrimer, Gβγ can regulate effector molecules like PLCβ, G protein inwardly rectifying channel (GIRK), and phosphoinositide-3 kinase (Cabrera-Vera et al., 2003).

1.1.4 Nucleotide-dependent structural transitions in Gα

The α subunit undergoes substantial structural rearrangements based on the nature of the bound nucleotide. These structural changes define the G protein molecular switch and dictate the interaction of the α subunits with effectors, GAPs, and GDIs. The structure of Gα subunits has been studied extensively. The crystal structures of activated and inactive form of $G\alpha_{i1}$ (Coleman et al., 1994; Mixon et al., 1995) (Figure 1.2 A and B) and $G\alpha_t$ (Lambright et al., 1994; Noel et al., 1993) have been determined. The structure of the activated form of $G\alpha_{i1}$ has been studied using two different GTP analogs: GTPγS (Coleman et al., 1994) and GMPPNP (Coleman and Sprang, 1999). The crystal structures of $G\alpha_{i1}$ was also determined in the GDP·AlF$_4$ bound state in complex with a RGS protein (Tesmer et al., 1997a) and the structure of GDP·AlF$_4$ bound transducin has also been determined (Sondek et al., 1994).
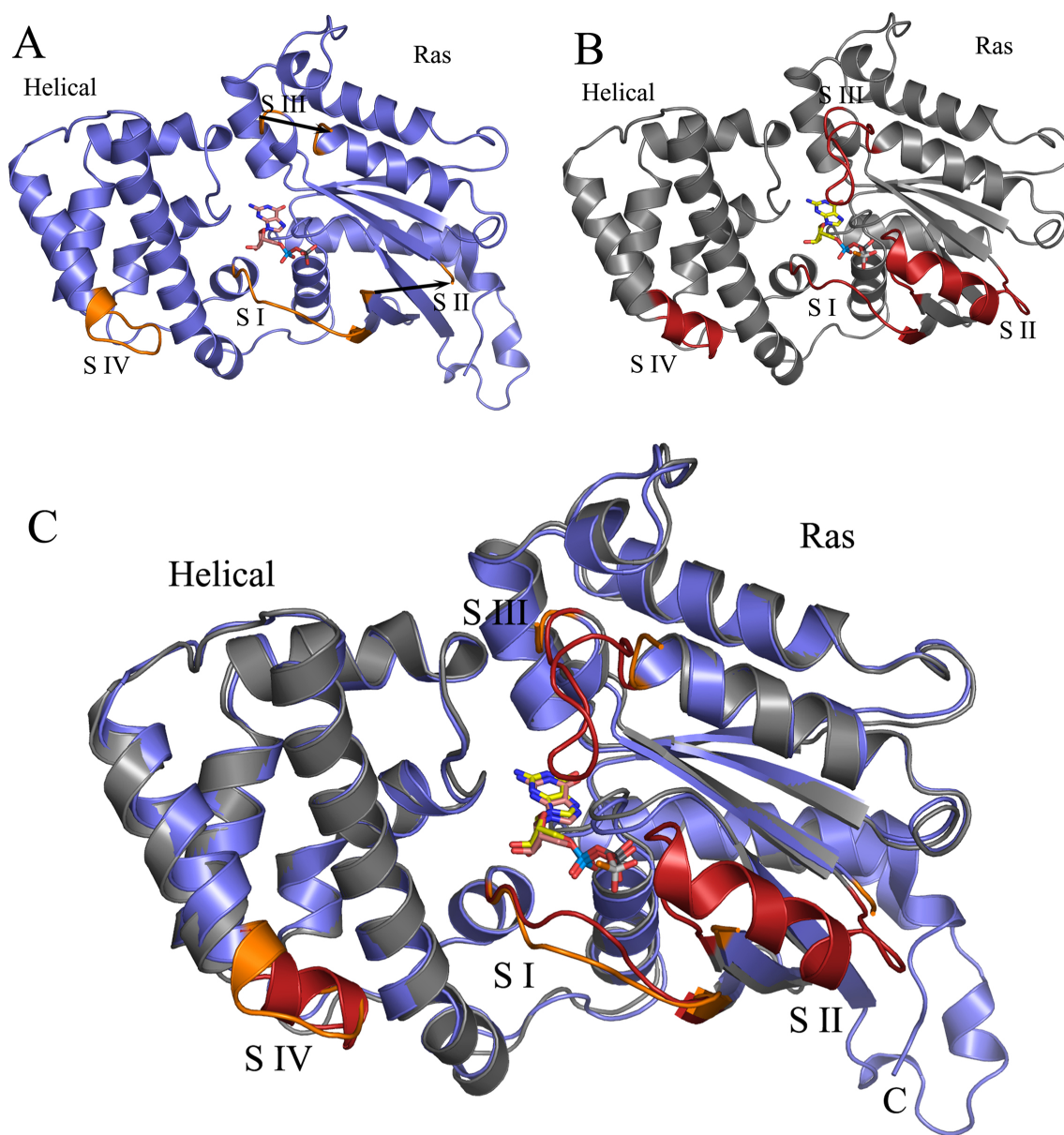
Figure 1.2. Comparison of Gα$_{i1}$ structure in the GDP and GTPγS bound conformations.
A. Crystal structure of Gα$_{i1}$·GDP (PDB Code: 1GDD). Switch regions are shown in orange.
B. Crystal structure of Gα$_{i1}$·GTPγS·Mg^{2+} (PDB Code: 1GIA). Switch regions are shown in red. C. Superposition of A and B. Switch regions are marked as S I (switch I), S II (switch II), S III (switch III), and S IV (switch IV)

Nucleotide dependent structural changes in Gα subunits are localized to certain regions that are commonly referred to as 'switch' regions. There are three prominent switch regions in Gα_{i1} and Gα_t referred to as switch I, switch II and switch III (Lambright et al., 1994; Mixon et al., 1995); and a lesser known switch IV in Gα_{i1} (Mixon et al., 1995). The switch regions in Gα_{i1} are: switch I (residues 177-187), switch II (residues 199-219), switch III (residues 231-242), and switch IV (residues 111-119). In the GDP-bound structure of Gα_{i1} switch II and switch III are completely disordered, but these segments adopt ordered structure in the GTPγS bound state (Mixon et al., 1995) (Figure 1.2). Switch II adopts a helical conformation in the GTPγS bound structure and this structure is stabilized by interaction of the γ-phosphate of GTP and switch II (Sprang, 1997a). In addition to switch regions, structural changes are also observed at the N and C-terminus of Gα_{i1}. The N-terminal 32 residues and the C-terminal 10 residues are disordered in the GTPγS bound structure of Gα_{i1}, but form a compact micro-domain in the GDP bound structure (Mixon et al., 1995; Sprang, 1997a). This region also forms extensive tertiary contacts in the Gα_{i1}.GDP structure, supporting the observation that G proteins exist as oligomeric complexes in the membrane (Jahangeer and Rodbell, 1993). Changes in the switch regions are also observed in the structure of Gα_t, but in the GDP bound structure switch II adopted an ordered conformation. While this is in contrast to the structure of Gα_{i1}·GDP, conformation of the switch regions were different in the GDP and GTPγS bound state (Lambright et al., 1994). The localization of nucleotide dependent structural changes in the switch regions led to the notion that these regions are crucial determinants of effector/regulator binding. This notion

has since been verified through several structural studies, that include the structures of

$G\alpha_{i1}$·GDP·AlF_4 bound to RGS domain of RGS4 (Tesmer et al., 1997a), $G\alpha_{i1}$·GDP in

complex with $G\beta_1\gamma_2$ (Wall et al., 1995), and $G\alpha_s$ in complex with the soluble domain of

adenylyl cyclase (Tesmer et al., 1997b).

1.1.5 Regulation of nucleotide exchange in G protein α subunits: GPCRs and βγ

Heterotrimeric G-proteins have two important regulatory aspects: 1. regulation of the

hydrolysis of GTP that determines the lifespan of activated α subunits; 2. regulation of GDP

to GTP exchange rate that determines the rate at which α subunits are activated. The

mechanism of GTPase activity of G-proteins and its activation by RGS proteins have been

studied in detail and extensively reviewed (Kleuss et al., 1994; Ross and Wilkie, 2000;

Siderovski and Willard, 2005). In this section, the regulation of nucleotide exchange rate will

be focused upon in detail.

Nucleotide exchange in Gα subunits can be thought as a two-step process: release of

GDP from the nucleotide binding pocket leading to the formation of a nucleotide-free

transition state, and binding of GTP to the nucleotide-free Gα. Of these, the release of GDP

from the nucleotide binding pocket is considered to be the rate limiting step of the exchange

reaction (Ferguson et al., 1986; Gilman, 1987; Sprang, 1997a). GEFs and GDIs by definition

regulate the rate of GDP release.

The βγ subunit maintains the G protein heterotrimer in an inactive GDP bound state.

The molecular basis of Gα and Gβγ interaction has been studied extensively and the crystal

structures of $G\alpha_{i1}$ and $G\alpha_t$ in complex with Gβγ has been determined (Lambright et al., 1996;

Wall et al., 1995). The crystal structure of the Gβγ dimer has been determined in isolation as well (Sondek et al., 1996). The βγ subunit is a GDI; binding of Gβγ stabilizes Gα.GDP, resulting in very low rate of GDP dissociation from the heterotrimer (Gilman, 1987; Higashijima et al., 1987; Sprang, 1997a). In the structure of the heterotrimer, Gβγ makes extensive contact with the switch II region and the N-terminal region of $Gα_{i1}$ (Figure 1.3). In the Gβγ bound state, switch II of $Gα_{i1}$ is well ordered and forms a helical structure, but differs from the conformation adopted by switch II in the GTPγS-bound state (Lambright et al., 1996; Wall et al., 1995) (Figure 1.3). Binding of Gβγ promotes the formation of hydrogen bonding interactions between residues R178 and E43 of $Gα_{i1}$, resulting in a 'seatbelt' like conformation which blocks the nucleotide binding pocket and potentially prevents the release of GDP (Wall et al., 1995).

The structure of the heterotrimer led to speculations about possible mechanisms of receptor mediated activation of the heterotrimer. It was thought that βγ, by virtue of its membrane localization, would make the heterotrimer accessible to membrane bound receptors (Wall et al., 1995). Although the interaction site of activated receptor on the heterotrimer was not known, it was suggested that receptors might interact with the N-terminal region of $Gα_{i1}$ and displace Gβγ from the N-terminal binding site. While this might explain activated GPCR induced dissociation of Gβγ from Gα, the enhancement of GDP release rate by GPCRs was thought to be mediated by its interaction with C-terminal region of Gα (Cai et al., 2001; Onrust et al., 1997; Wall et al., 1995). This was also based on the
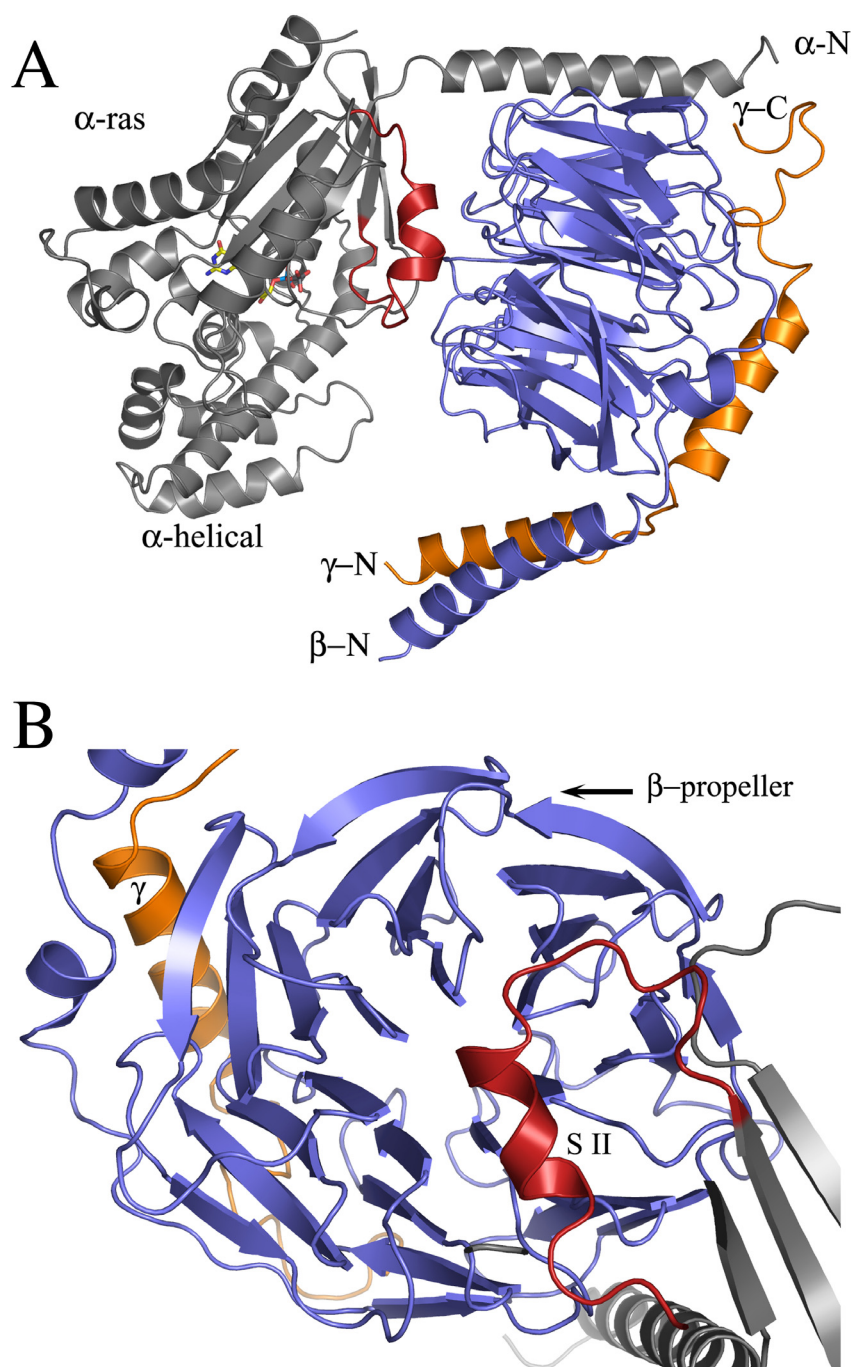
Figure 1.3. Structure of heterotrimeric G protein. *A*. The structure of the complex of $G\alpha_{i1}$ (grey), β_1 (blue), and γ_2 (orange) (PDB code: 1GP2). Switch II region of $G\alpha_{i1}$ (red) forms a helical structure in the complex. *B*. Interaction of switch II (red) with the β-propeller domain (blue). The view is through $G\alpha_{i1}$ towards the β subunit.

observation that, receptor mimetic peptides like mastoparan mediated its GEF activity through interaction with the C-terminal region of Gα (Higashijima and Ross, 1991; Higashijima et al., 1988). The possibility that cytoplasmic loops of activated receptors might directly interact with the switch II region by inserting itself between the Gα and Gβγ subunits was also considered possible (Wall et al., 1995). However, later studies indicated that given the orientation of the heterotrimer in the membrane cytoplasmic loops are not long enough to make direct contacts with switch II (Iiri et al., 1998).

The molecular basis of activated receptor-G protein interaction remains poorly understood. Although the crystal structure of rhodopsin, a light activated GPCR, is available (Palczewski et al., 2000), the crystal structure of GPCR in complex with the G protein heterotrimer remains unknown. Thus, the precise molecular mechanism of GPCR-mediated acceleration of nucleotide exchange remains a mystery. In the absence of structural information, models have been proposed to explain the mechanism of action of GPCRs. These models are based upon a large body of mutational data in Gα, the observation that Gβγ is essential for GPCR activity, and crosslinking studies with receptors (Bourne, 1997; Hamm, 2001; Herrmann et al., 2004; Herrmann R, 2004). There are currently two popular models about the mechanism of GPCR action: the 'lever arm' model (Iiri et al., 1998; Rondard P, 2001) and the 'gear shift' model (Cherfils J, 2003). Both of these models are inspired by the fact that cytoplasmic loops of GPCRs are probably not long enough to contact the nucleotide binding pocket of Gα; the 'action at a distance' (Iiri et al., 1998) problem can be only explained by an allosteric mechanism that involves Gβγ.
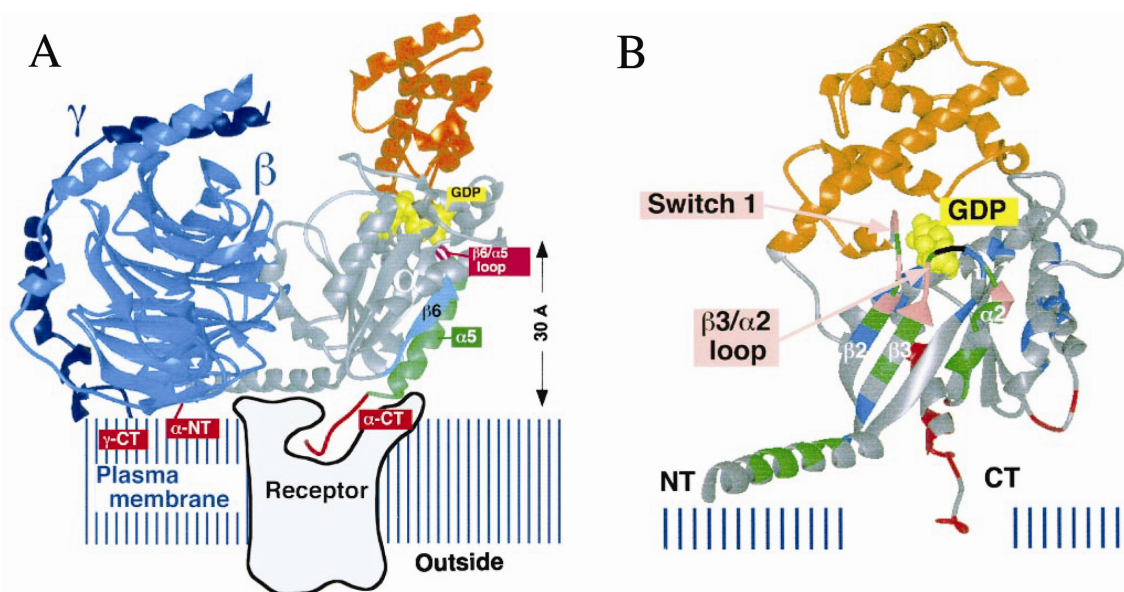
Figure 1.4. The 'lever arm' model of GPCR activity. *A*. The orientation of the heterotrimer is shown with respect to the membrane. The C-terminal tail (red) interacts with the receptor. Gβγ (light and dark blue) is attached to the membrane by lipids. The functionally important α5 helix (green), β6-α5 loop (red), and β6 strand (cyan) is shown. *B*. Residues on the α subunit that contact Gβγ are shown in green and the β3-α2 loop is shown in black. Receptors are proposed to use Gβγ as a lever to pull β3-α2 loop away from the GDP binding pocket. The figure is adopted from (Iiri et al., 1998)

In the 'lever arm' model two regions of Gα are considered important for interaction with GPCRs: The C-terminal region of Gα including the α5 helix and β6-α5 loop, and the β3-α2 loop connecting switch II helix to the β3 strand (Figure 1.4). In the heterotrimer, Gβγ binds to the N-terminal region of Gα on one contact surface, and remains attached to the switch II region on the other contact surface. It was suggested that the activated receptor tilts Gβγ with respect to Gα. This causes the switch II region to be pulled along with Gβγ, which in effect pulls the β3-α2 loop (N-terminal part of switch II) away from the nucleotide binding pocket (Figure 1.4). This creates a potential exit route for GDP from the nucleotide binding

cleft (Iiri et al., 1998; Rondard P, 2001). The receptor was also thought to destabilize the nucleotide binding pocket by pulling the α5 helix (Iiri et al., 1998).

In the 'gear shift' model it is also assumed that interaction of Gβ and Gα plays an important role in mediating GPCR activity, but unlike the 'lever arm' model a role of the γ subunit is also proposed. In the 'gear shift' model, activated receptors are thought to promote more compact interaction of βγ and α subunits. This is achieved by displacement of switch II towards the Ras-like domain core by Gβ and simultaneous interaction of γ subunit with the α-helical domain of the α subunit, such that GDP is pushed away from the nucleotide binding pocket through the opposite side of Gα and Gβγ interface (Figure 1.5). The simultaneous interaction of βγ with both the helical and Ras-like domain is proposed to stabilize the nucleotide-free state (Cherfils J, 2003). The notion that the γ subunit interacts with the helical domain of Gα was inspired by the sequence similarity of γ subunit and C-terminal region of RGS14 GoLoco motif (discussed later), which also interacts with the helical domain (Cherfils and Chabre, 2003; Kimple et al., 2002a).
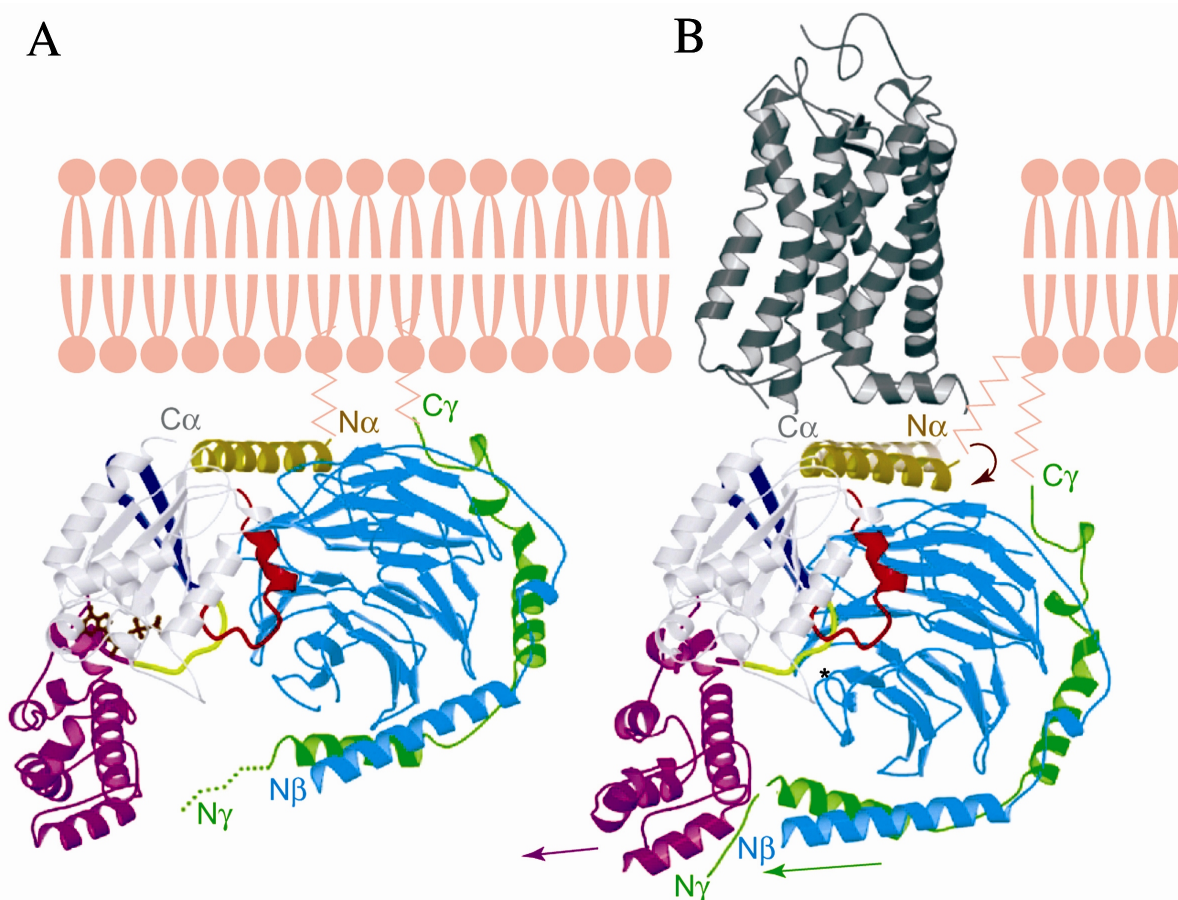
Figure 1.5. The 'gear shift' model of GPCR activity. *A*. The orientation of the heterotrimer in relation to the cell membrane is shown. The Ras-like domain, helical domain, Gβ, and Gγ are shown in light grey, magenta, blue, and green respectively. Switch II is shown in red and C-terminal helix of Gα is shown in yellow. *B*. Interaction of receptor (grey) with N-terminal helix of Gα (yellow) forces Gβ (blue) to pack on the switch II region (red). This leads to interaction of Gγ (green) with the helical domain of Gα (magenta). This conformation stabilizes the nucleotide-free state. The figure is adopted from (Cherfils and Chabre, 2003)

While neither of these models have direct support from structural studies, the role of

Gβγ in promoting receptor enhanced nucleotide exchange is clear. Indeed, a recent NMR

study shows that the α subunit adopts a 'pre-activated' conformation when bound to Gβγ,

which allows favorable interaction with activated receptor and GDP exchange. It was also

suggested that receptor activation might lead to the transient formation of the Gα·GTP:βγ complex (Abdulaev et al., 2005). While this suggestion is contradictory to the fact that Gα effectors cannot bind to activated Gα if Gβγ is bound; the concept that Gβγ dissociates from activated α subunits has been a subject of controversy (Levitzki A, 2002). In any case, it is clear that βγ in addition to being a GDI, plays a dual role of assisting GPCRs in enhancing the nucleotide exchange rate in α subunits. Interestingly, the 'lever arm' model has received some experimental support from the crystal structure of Gα$_{i1}$ in complex with a peptide having GEF activity (detailed discussion in chapter 6) (Johnston CA, 2005).

1.2   Activator of G protein Signaling 3 and the GoLoco motif

1.2.1   Discovery of AGS3 and GoLoco motif

AGS3 is a 650-residue protein from *Rattus norvegicus* that was first identified in a yeast based expression screen as a receptor-independent activator of Gβγ-dependent signaling. In a pioneering study, the pheromone-dependent G protein signaling pathway in yeast was exploited to identify mammalian non-receptor regulators of G protein signaling (Cismowski et al., 1999). The yeast pheromone-dependent G protein signaling pathway shares similarity with mammalian G protein signaling pathways (Bardwell L, 1994). Upon activation by pheromones, the pheromone-associated GPCR in yeast promotes the dissociation of Gβγ subunit from the activated Gα·GTP subunit. Once dissociated, yeast Gβγ activates mating-specific transcription factors through a complex cascade of kinase mediated

signaling. Yeast strains were engineered in such a way that their growth was dependent on the activation of pheromone-dependent G protein signaling pathway (Cismowski et al., 1999). The native GPCR in these strains were knocked out and the cells were transformed with a mammalian cDNA library. Thus, cDNAs encoding mammalian receptor-independent activators would be able to activate $G\beta\gamma$ signaling in the absence of GPCR activation. Several molecules were identified through these screens that were able to activate G-protein signaling and were called as Activators of G-protein Signaling or AGS proteins (Cismowski et al., 2001; Cismowski et al., 1999; Takesono et al., 1999). The most well studied candidate among the AGS family of proteins was activator of G protein signaling 3 (AGS3). Sequence analysis indicated that AGS3 contains tetratricopeptide motifs in its N-terminal domain and GoLoco/GPR motifs (G Protein Regulatory) in its C-terminal domain. It was also found that AGS3 preferred the GDP bound form of $G\alpha_{i3}$ over the GTPγS bound form (Takesono et al., 1999).

GoLoco/GPR motifs found in the C-terminal domain of AGS3 were found earlier in different proteins including RGS12, RGS14, Rap1GAP and LGN (Siderovski et al., 1999). The GoLoco motif was first identified in the *Drosophila* protein 'loco', the homolog of mammalian RGS12. The ~20 residue GoLoco motif was also found in several other proteins that interact with the $G\alpha_{i/o}$ subunits. Since the sequence was identified by comparison with the drosophila protein 'loco', it was named as '$G\alpha_{i/o}$-loco' or 'GoLoco' motif (Siderovski et al., 1999). GoLoco motifs were also referred to as G-protein regulatory motif or GPR motif (Takesono et al., 1999). However, over time the GoLoco nomenclature has been accepted as

standard (Willard et al., 2004). Before the characterization of their function, GoLoco motifs were thought to be GEFs (Siderovski et al., 1999).

1.2.2   Domain architecture of GoLoco motif containing proteins

The analysis of AGS3 sequence revealed the presence of seven tetratricopeptide (TPR) repeats in its amino-terminal half and the presence of four GoLoco motifs in its C-terminal half (Bernard et al., 2001b; Takesono et al., 1999). The third TPR motif shows weak similarity to TPR consensus sequence and was not considered a TPR motif in certain studies (Willard et al., 2004) (Figure 1.6). TPR motifs have been shown to act as protein interaction domains in multi-protein complexes (Blatch GL, 1999). The function of the GoLoco motif was characterized later through a series of biochemical studies (discussed in next section). Sequence analysis and database searches also identified several homologs of AGS3 in different organisms, which had a similar arrangement of TPR and GoLoco rich domains. These homologs include the human protein LGN, *Drosophila* protein Partner of Inscuteable (PINS) and a related protein from *Caenorhabditis elegans* (Bernard et al., 2001a).

GoLoco motifs occur either singly, or as tandem repeats in proteins that interact with G_i and G_o-class α subunits (Willard et al., 2004). Both RGS12 and RGS14 contain one GoLoco motif as do Purkinje cell protein-2 (Pcp2) and Rap1GAP (Kimple et al., 2001; Natochin et al., 2001; Natochin et al., 2000; Willard et al., 2004). GoLoco motif containing proteins with a mix of TPR and GoLoco motifs have different numbers of GoLoco motifs in
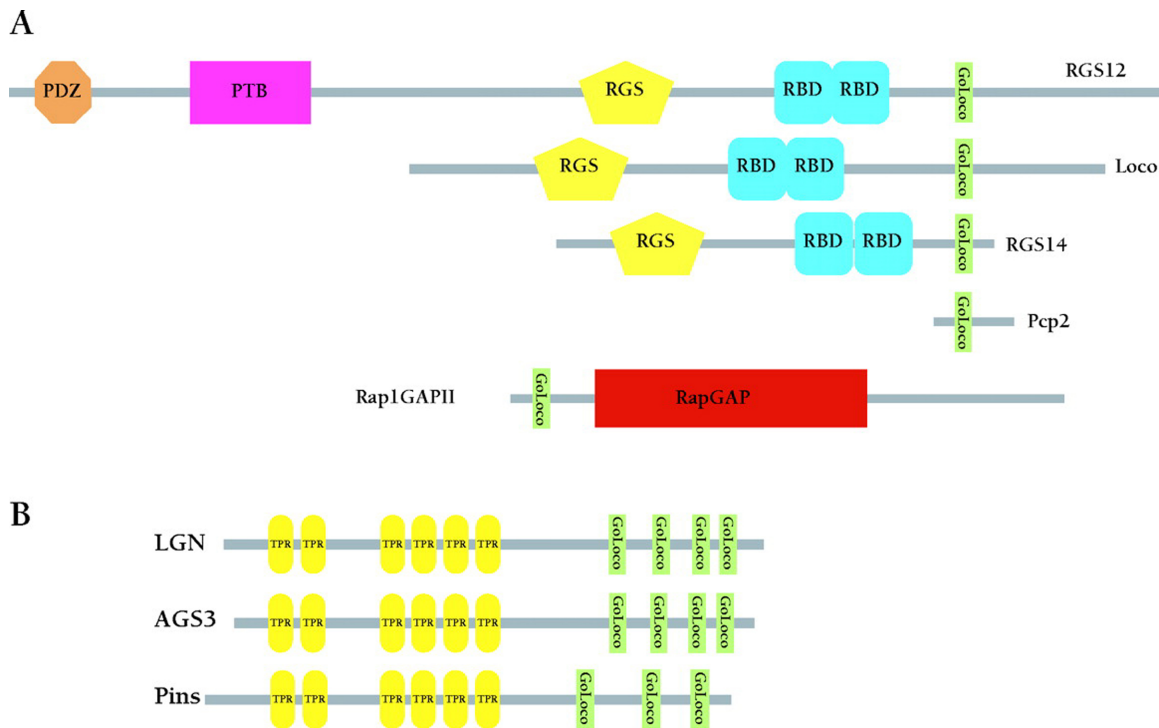
Figure 1.6. Domain architecture of GoLoco motif containing proteins. *A*. Domain organizations of proteins containing one GoLoco motif (green) like RGS12, Loco, RGS14, Pcp2, and Rap1GAPII. *B*. Domain organization of proteins containing TPR repeats (yellow) in their N-terminus and GoLoco (green) repeats in the C-terminus. The figure is adopted from (Kimple et al., 2002b)

their C-terminal domains. Mammalian proteins like AGS3 and LGN have four GoLoco

motifs in the C-terminal domain (Blumer et al., 2002; Takesono et al., 1999). The drosophila

homolog of AGS3 known as PINS (partner of inscuteable) has three GoLoco motifs (Bernard

et al., 2001a) (Figure 1.6). Surprisingly, the *C.elegans* functional homolog of AGS3/PINS

family of proteins known as GPR-1/2 has only one GoLoco motif in the C-terminal domain

(Srinivasan et al., 2003). The number of proteins known to contain GoLoco motif has grown
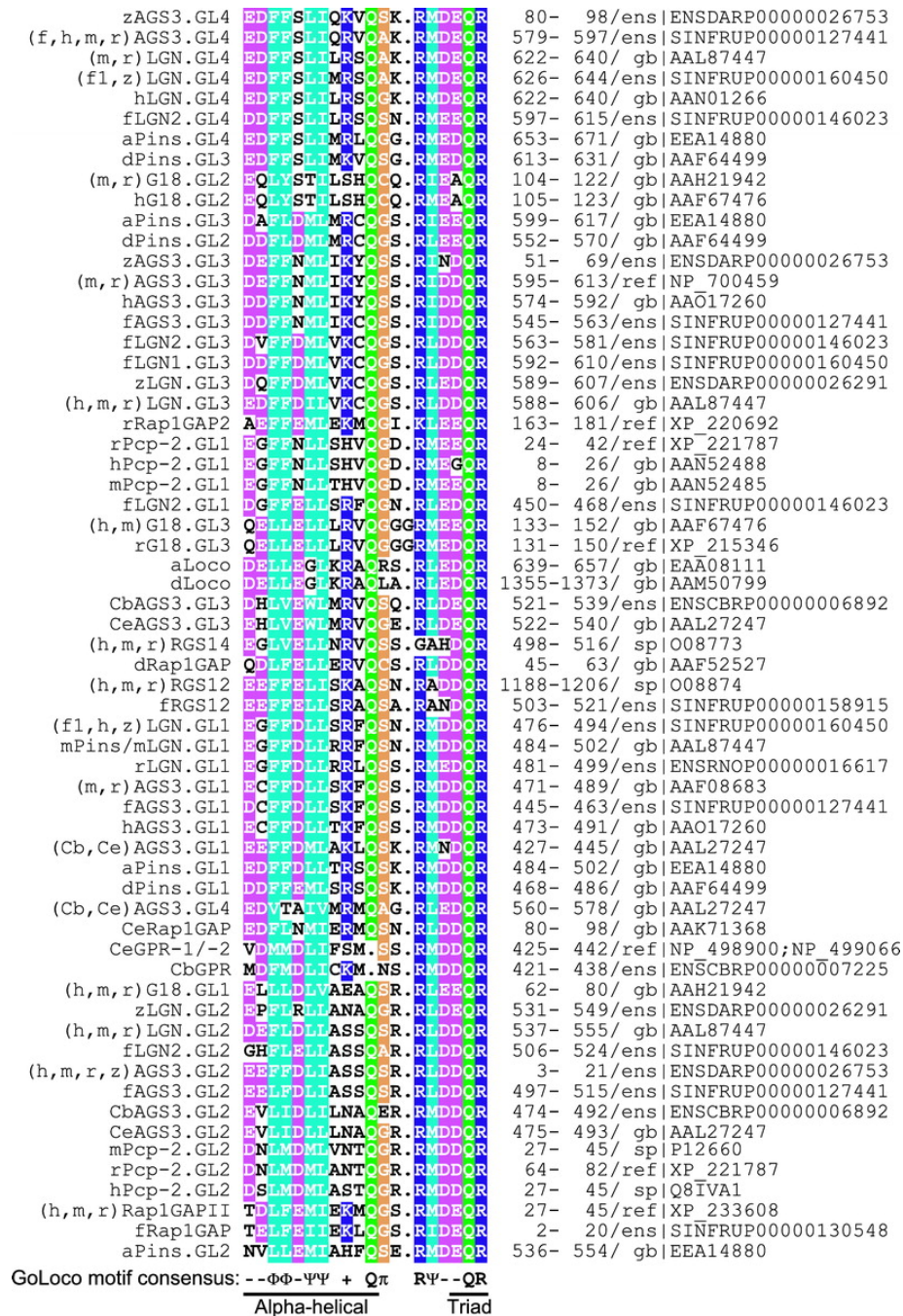
Figure 1.7. Multiple sequence alignment of GoLoco motifs. Individual GoLoco motifs are numbered as GL# going from N to C terminus. The conserved 'DQR triad' and helical region are shown based on the structure of RGS14-GoLoco in complex with $G\alpha_{i1}$ (Kimple et al., 2002a). ϕ and ψ refer to hydrophobic and large aliphatic; - and + indicate acidic and basic. The figure is adopted from (Willard et al., 2004)

rapidly over the last several years. There are 140 proteins in the non-redundant database containing a total of 304 GoLoco motifs. (http://smart.embl-heidelberg.de/smart/do_annotation.pl?DOMAIN=GoLoco&BLAST=DUMMY). A comprehensive multiple sequence alignment of different GoLoco motifs have been reported recently, which suggest the presence of a conserved α-helical region at the N-terminus of the motif and a highly conserved "Asp-Gln-Arg" triad (Willard et al., 2004) (Figure 1.7). The significance of these conserved regions will be discussed in the subsequent sections.

1.2.3   GoLoco motif containing proteins: GDIs of heterotrimeric G proteins

Initially, GoLoco motifs were thought to function as GEFs (Siderovski et al., 1999) but later they were functionally characterized as GDIs. The biochemical function of GoLoco motifs was discovered by studies of a peptide having the consensus sequence of GoLoco motifs in AGS3, and also by studies of the C-terminal GoLoco-containing domain of AGS3. The peptide having a consensus sequence of GoLoco motifs in AGS3, referred to as the 'GPR consensus' peptide, stabilized the GDP bound conformation of $G\alpha_i$ and inhibited the binding of a GTP analog (GTPγS) to $G\alpha_i$ with an IC_{50} of 200 nM (Peterson et al., 2000). Thus, GoLoco motifs are GDIs. It was also found that mutation of the conserved arginine residue at the C-terminus of GPR-consensus peptide significantly reduced its GDI activity (Peterson et al., 2000). Soon after this discovery, several studies demonstrated that AGS3, which has four GoLoco motifs, was also a GDI. In the first study, AGS3 was found to interact specifically with the GDP bound conformation of $G\alpha_{i1}$ and $G\alpha_o$, and inhibited nucleotide exchange in $G\alpha_{i3}$ (De Vries et al., 2000a). In another study it was shown that

AGS3 inhibited the binding of GTPγS to $G\alpha_{i1}$ by stabilizing GDP bound state, but had no effect on the rate of GTPγS binding to $G\alpha_s$ (Natochin et al., 2000). AGS3 competes with Gβγ for binding to $G\alpha_t$ (Bernard et al., 2001a; Natochin et al., 2000). AGS3 also inhibited rhodopsin-mediated activation of $G\alpha_t$ (Natochin et al., 2000) and antagonized the coupling of 5-HT1A receptors and $G\alpha_i$ (Ma et al., 2003). To summarize, AGS3 was found to bind specifically to the GDP-bound forms of all three $G\alpha_i$ isoforms, to $G\alpha_t$, and weakly to $G\alpha_o$ and $G\alpha_q$ (Bernard et al., 2001a; De Vries et al., 2000a; De Vries et al., 2000b; Natochin et al., 2001; Takesono et al., 1999) but exhibited GDI activity only towards $G\alpha_i$ and $G\alpha_t$ (Bernard et al., 2001a; De Vries et al., 2000a; De Vries et al., 2000b; Natochin et al., 2000).

GoLoco-containing proteins Pcp2 and Rap1GAP also display GDI activity towards $G\alpha_o$ (Natochin et al., 2001). None of the GPR/GoLoco-containing proteins heretofore characterized demonstrate GDI activity towards $G\alpha_s$ (Natochin et al., 2001; Natochin et al., 2000). The GoLoco motifs of AGS3-C are necessary and sufficient for the GDI activity of AGS3 (Bernard et al., 2001a; De Vries et al., 2000a; De Vries et al., 2000b). Indeed, a fully active splice variant of AGS3 expressed in cardiac muscle consists only of the GoLoco-containing C-terminal domain of AGS3 (Pizzinat et al., 2001). G18, a protein containing three GoLoco motifs and lacking the TPR repeats, is also a potent GDI (Kimple et al., 2004). Several other GoLoco motif containing proteins were also shown to have GDI activities. GoLoco motif containing proteins like LGN, Pcp2 and Rap1GAP2 exhibits GDI activity towards to $G\alpha_{i1}$ and $G\alpha_o$ (Natochin et al., 2001). Peptides that correspond to individual GPR

motifs of RGS12 and RGS14 were also shown to have potent GDI activity towards $G\alpha_{i1}$ but

not $G\alpha_o$ (Kimple et al., 2001; Natochin et al., 2001; Peterson et al., 2000).

1.2.4   Molecular basis of the GDI activity and specificity of GoLoco motifs

The molecular basis of GoLoco mediated GDI activity was discovered through a

series of biochemical and structural studies. Truncation and alanine scanning mutagenesis

experiments revealed a conserved functional core within the GoLoco motif, corresponding to

the sequence, FFxLLxxxxxxxMxxQR, that is required for GDI activity (Peterson et al.,

2002). Mutation of the conserved arginine residue at the C-terminus of the GoLoco motif

resulted in 70 fold increase in IC_{50} (Peterson et al., 2002). Experiments with chimeric

constructs of $G\alpha_{i1}$ and $G\alpha_s$ identified the switch regions and helical domain of $G\alpha$ as

GoLoco binding sites (Natochin et al., 2002).

The crystal structure of GoLoco motif of RGS14 in complex with $G\alpha_{i1}$ provided the

most detailed insight into the structural basis of GDI activity and specificity of GoLoco

motifs (Kimple et al., 2002a). The crystal structure contained a peptide corresponding to the

RGS14-GoLoco motif and its following 15 C-terminal residues (496-530 of RGS14).

Binding of the GoLoco peptide reorganizes and stabilizes the switch segments of $G\alpha_{i1}$;

switch II region adopts a well ordered helical structure in the complex (Figure 1.8). Switch II

and switch III also undergo conformational changes with respect to the structure of

$G\alpha_{i1}$:$\beta\gamma$ (Kimple et al., 2002a). An arginine residue from the GoLoco motif, located in the

conserved 'Asp-Gln-Arg (DQR) triad', is inserted into the active site of $G\alpha_{i1}$ and interacts

with the β phosphate of GDP. The insertion of the 'arginine finger' from GoLoco resulted in
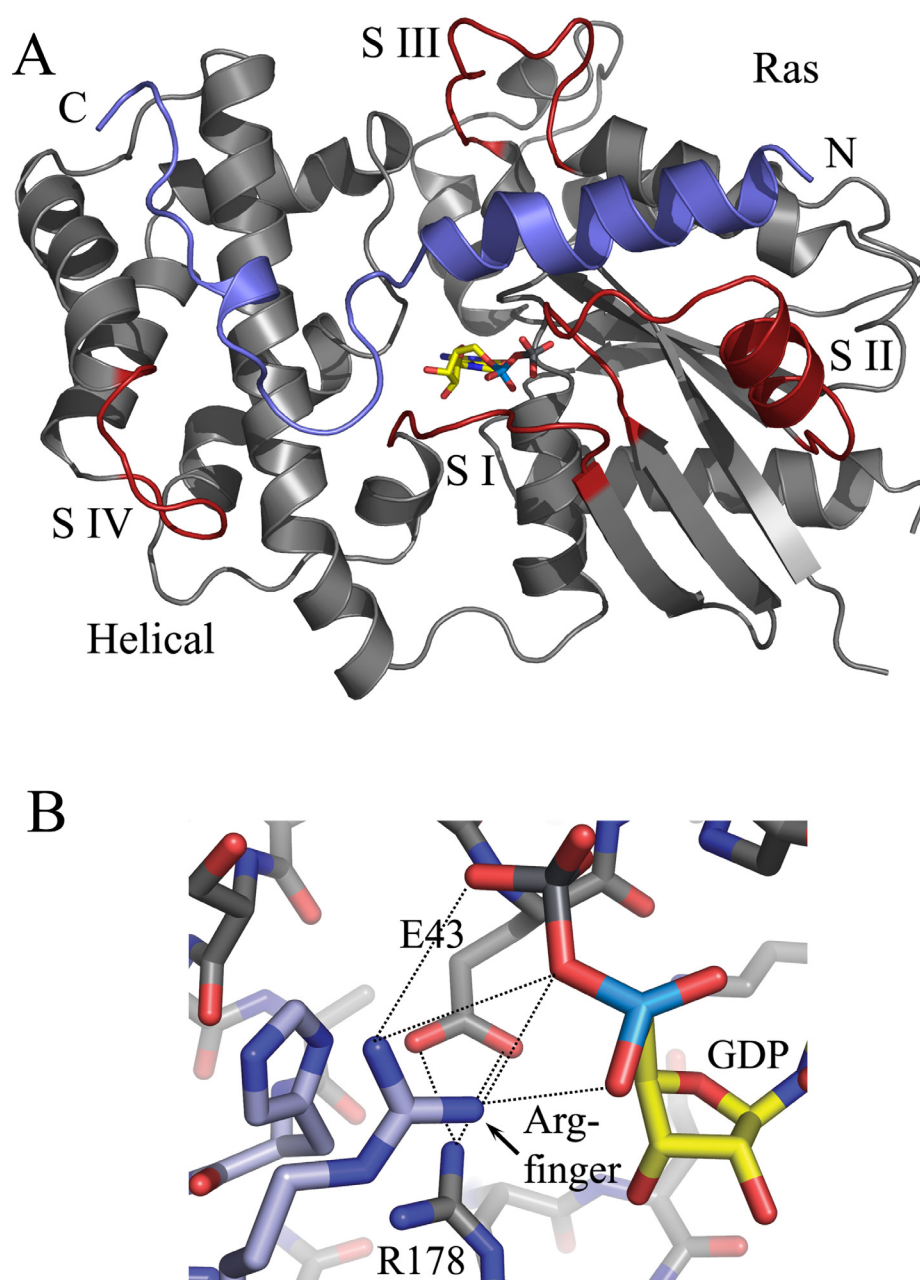
Figure 1.8. Structure of RGS14-GoLoco·Gα$_{i1}$ complex. *A*. The crystal structure of RGS14-GoLoco in complex (PDB code: 1KJY) with Gα$_{i1}$ is shown. Gα$_{i1}$ is shown in dark grey, switch regions (switch I-IV: S I-IV) are shown in red, and the RGS14-GoLoco peptide is shown in blue. *B*. Hydrogen bonding interaction between the conserved arginine (light blue) and GDP; and Gα$_{i1}$ residues R178 and E43 is shown.

the displacement of R178 of $G\alpha_{i1}$, which formed hydrogen bonds with E43 in a 'seatbelt'

like conformation that is highly reminiscent of the $G\beta\gamma$ bound structure of $G\alpha_{i1}$ (Kimple et

al., 2002a; Wall et al., 1995) (Figure 1.8). Residues C-terminal to GoLoco motif crossed over

to the helical domain and formed contacts with the helical domain. Interaction with the C-

terminal domain altered the conformation of the switch IV region (Kimple et al., 2002a;

Mixon et al., 1995). The structure of the complex suggested that GDI activity of GoLoco

motifs are driven by two factors: 1. The guanidinium-βPO_4^- interaction of the conserved

arginine in GoLoco and the E43-R178 ion-pair 'seatbelt'; and 2. The tethering of the helical

domain to the Ras-domain by C-terminal residues of GoLoco motif (Kimple et al., 2002a;

Kimple et al., 2002b).

Residues C-terminal to the GoLoco motif interact with non-conserved residues in the

helical domain of $G\alpha_{i1}$, and are therefore proposed to be determinants of specificity (Kimple

et al., 2002a). Replacing these residues with the C-terminal residues of the GoLoco motif of

Pcp2, which is specific towards $G\alpha_o$, resulted in altered specificity towards $G\alpha_o$. Although

the C-terminal residues of the GoLoco motif of RGS14 were found to be important in

determining its specificity, residues that are C-terminal GoLoco motifs in AGS3 shows no

sequence conservation, despite their specificity towards $G\alpha_i$ (Kimple et al., 2002a).

However, a recent study of GoLoco motifs of LGN showed that all the GoLoco motifs of

LGN are specific towards $G\alpha_i$, in spite of significant sequence variation in the C-terminal

residues of its GoLoco motifs (McCudden et al., 2005). The RGS14 GoLoco motif was even

found to be capable of discriminating among different $G\alpha_i$ isoforms (Mittal and Linder,

2004). However, in the absence of the structure of other GoLoco motifs in complex with

their binding partners, the molecular basis of the specificity of GoLoco motifs remains poorly understood.

1.2.5   Receptor-independent G protein signaling: A novel paradigm

The identification of AGS3 through a yeast-based screen represents the first concerted effort to identify receptor-independent regulators of G protein signaling; it suggested that Gβγ mediated signaling can be activated in the absence of receptor signaling (Cismowski et al., 1999; Takesono et al., 1999). In separate experiments, it was found that G protein signaling is involved in asymmetric cell division in *Drosophila* and *C.elegans* (Bellaiche Y, 2001; Knoblich, 2001; Schaefer et al., 2001). The observation that the GoLoco motif containing protein PINS regulates asymmetric cell division in Drosophila neuroblasts, led to the idea that perhaps displacement of Gβγ from Gα by PINS is important for asymmetric cell division (Knust, 2001). The identification of RIC-8, a soluble cytosolic GEF for Gα subunits (Tall et al., 2003), further strengthened the notion that G protein signaling is possible in the absence of GPCR activation. A model for receptor-independent regulation of G protein signaling is shown in Figure 1.9. In this model, binding of GoLoco containing proteins displace Gβγ from Gα·GDP, leading to the formation of a Gα·GDP:GoLoco protein complex and free Gβγ. Gβγ, thus released, can interact with its effectors. The complex of Gα·GDP:GoLoco can be subsequently broken by a GEF such as RIC-8, leading to the formation of activated Gα·GTP, which would allow it to interact with its effectors (Srinivasan et al., 2003). It should be noted that much of this model remains speculative; there is no evidence yet showing that RIC-8 can act upon Gα·GDP:GoLoco complexes.
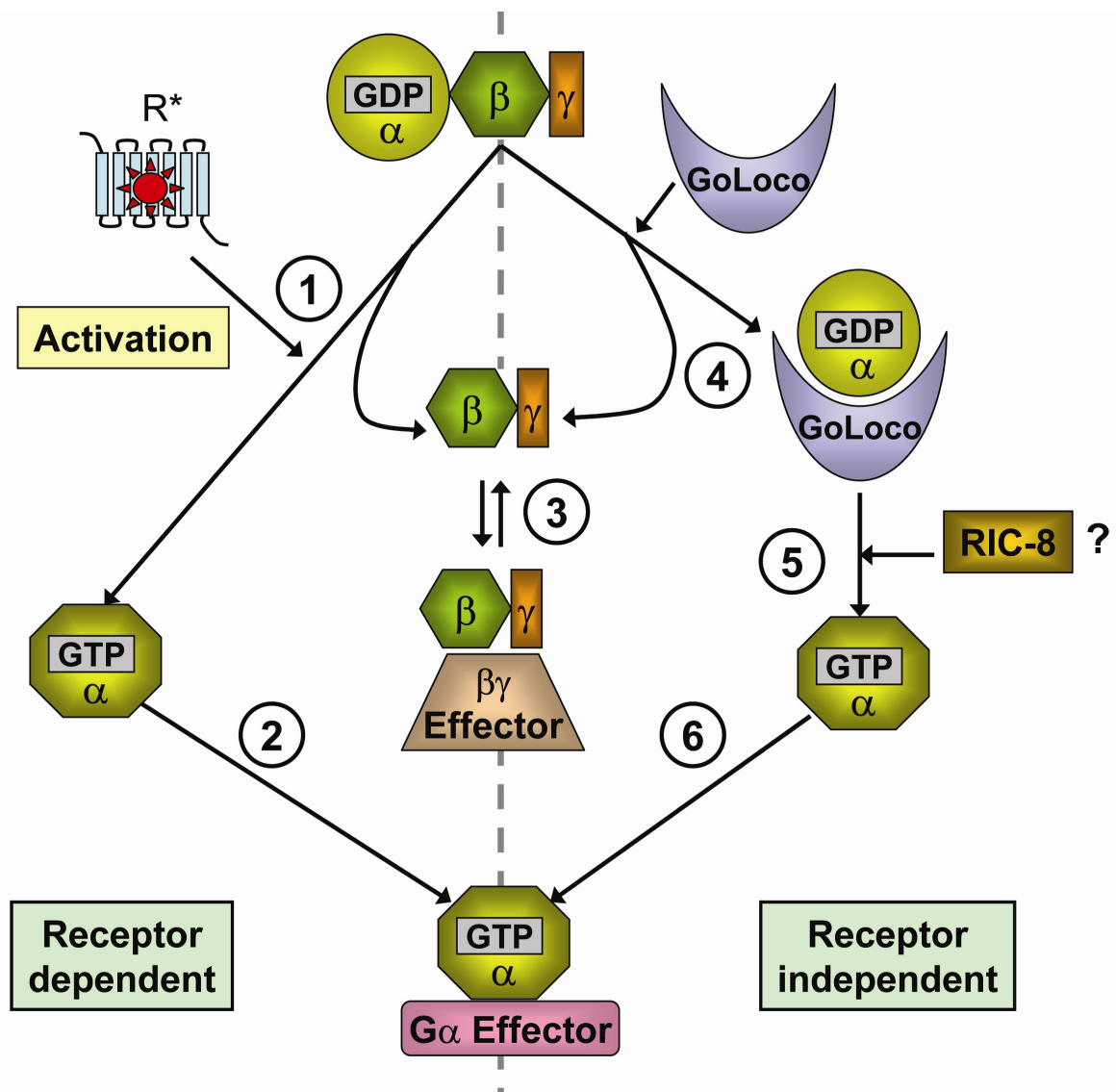
Figure 1.9. Schematic illustration of receptor-independent G protein signaling.
Receptor-dependent (canonical) and receptor-independent G-protein signaling pathways are
shown. In the receptor-independent pathway GoLoco motif containing proteins displace βγ
from Gα·GDP in the absence of receptor activation (4), allowing βγ to interact with its
effectors (3). The inactive Gα·GDP:GoLoco complex can be activated by a GEF like Ric-8
(5) leading to the formation of Gα·GTP, which can interact with its effectors (6). The
analogous components of GoLoco and Ric-8 in the receptor-dependent pathway are βγ (1)
and GPCR. The role of RIC-8 in this scheme is largely speculative (Srinivasan et al., 2003).

The biological implications of receptor-independent signaling remains poorly understood as well. Nevertheless, the role of G protein signaling in asymmetric cell division is profound and will be a topic of discussion in the next section.

1.3   G protein signaling in cell division

1.3.1   GoLoco motif proteins in asymmetric cell division

In recent years G protein signaling has been identified as an important regulator of asymmetric cell division (Hampoelz and Knoblich, 2004). Asymmetric cell division plays a crucial role in generating cellular diversity during embryonic development of metazoan organisms. Under normal circumstances, mitosis results in two identical daughter cells; but at certain stages of development, the daughter cells become unequal (Knoblich, 2001) (Figure 1.10). Most of the understanding of asymmetric cell division is gained from genetic and cell biological studies of the early development of *C.elegans* embryo and Drosophila neuroblast cells (Knoblich, 2001). Asymmetric cell divisions involve several discrete steps: establishment of an axis of polarity; orientation of mitotic spindle along the axis; asymmetric localization of cell fate determinants; and asymmetric cytokinesis resulting in two unequal daughter cells (Betschinger and Knoblich, 2004).

In *C.elegans*, asymmetric distribution of PAR (partition-defective) proteins help to establish anterior/posterior (A/P) polarity (Pellettieri and Seydoux). At the one-cell stage, the asymmetric distribution of PAR proteins leads to asymmetric force generation in the mitotic spindle (Nance, 2005; Srinivasan et al., 2003). This leads to the generation of a larger
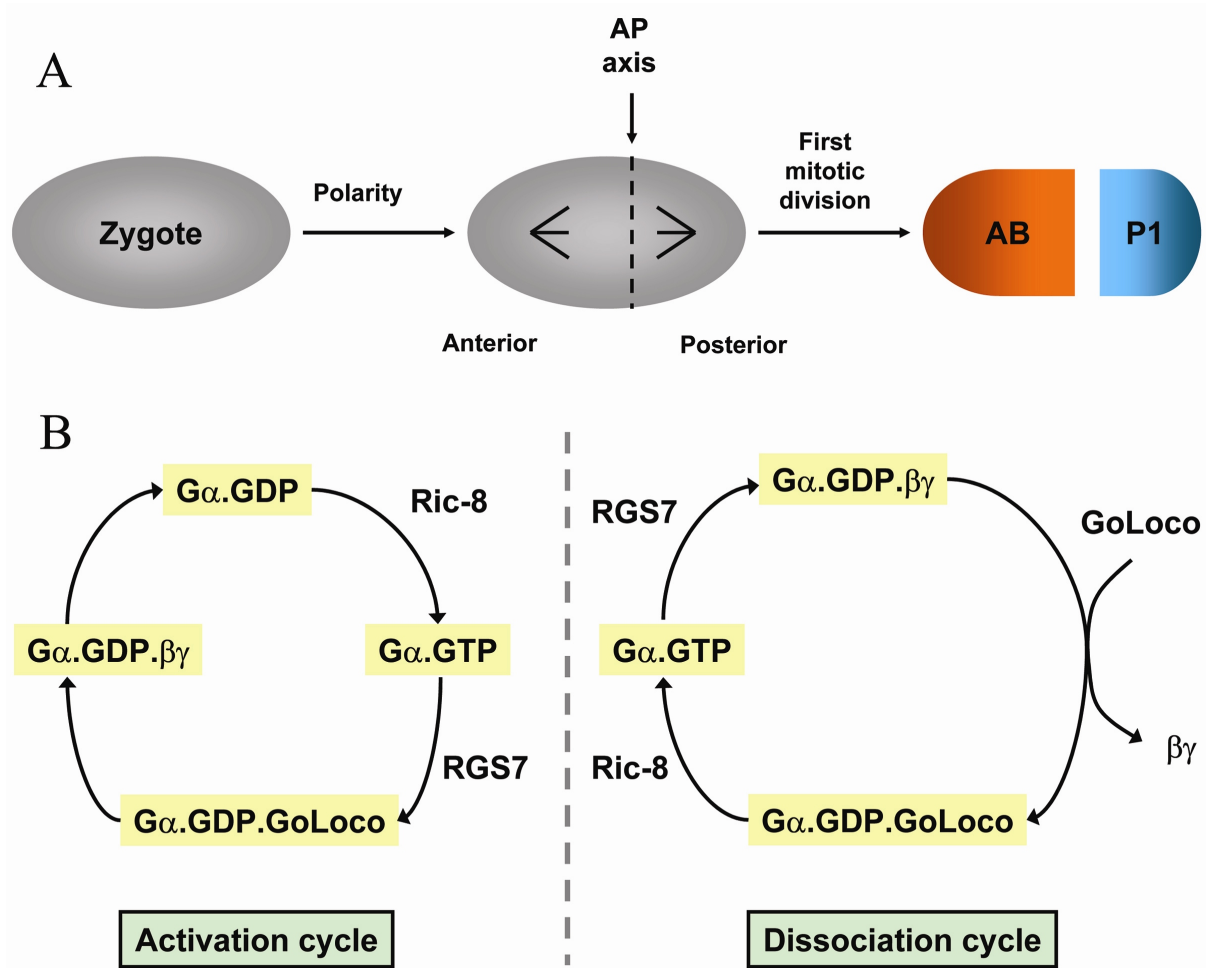
Figure 1.10. G protein signaling in asymmetric cell division. *A*. Asymmetric cell division at the one-cell stage of *C.elegans* embryo results in the formation of large AB and a smaller P1 cell. *B*. Possible regulatory schemes of receptor-independent G-protein signaling during asymmetric cell division in *C.elegans*.(Hampoelz and Knoblich, 2004)

anterior AB and a smaller posterior P1 daughter cell (Betschinger and Knoblich, 2004; Knoblich, 2001) (Figure 1.10). G protein signaling plays an important role in generating spindle forces in *C. elegans*. Knocking out *C. elegans* $G\alpha_{i/o}$ subunit encoding genes *goa-1*

and *gpa-16* by RNAi disrupts asymmetric cell division and results in embryonic lethality. RNAi knock-outs of C. elegans *gpr-1/2*, which is homologous to mammalian AGS3/LGN, also resulted in the same phenotype (Gotta et al., 2003; Srinivasan et al., 2003). Indeed, a direct measurement of spindle force showed that C. elegans embryos lacking gpr-1/2 generate symmetric spindle force (Colombo et al., 2003). While these experiments showed importance of G protein signaling at the one-cell stage of *C. elegans* development, potential GPCRs regulating this process have not been found. GPCR-mediated extracellular signaling is also unlikely at this stage due the presence of an impermeable membrane around the embryo, suggesting that G protein signaling is driven by receptor-independent regulators (Betschinger and Knoblich, 2004; Srinivasan et al., 2003). Several recent studies highlight the role of receptor-independent G protein signaling in *C. elegans* asymmetric cell division, suggesting new roles for RIC-8 and RGS7; however the detailed biochemical nature of this process remains poorly understood (Afshar et al., 2005; Couwenbergs et al., 2004; Hess et al., 2004) (Figure 1.10).

Heterotrimeric G protein signaling also plays a critical role in asymmetric cell division in *Drosophila* neuroblast cells (Betschinger and Knoblich, 2004; Knust, 2001). In *Drosophila*, $G\alpha_i$ localizes apically in neuroblasts and interacts with the *Drosophila* homolog of AGS3 known as PINS (Schaefer et al., 2001). PINS was thought to promote $G\beta\gamma$-mediated signaling in an receptor-independent manner (Schaefer et al., 2001). The C-terminal region of PINS is sufficient for its apical membrane localization in *Drosophila*, suggesting that Drosophila $G\alpha_i$ plays a role in the localization of PINS during asymmetric cell division (Yu

et al., 2002). Much like *C. elegans*, RIC-8 also plays an important role in asymmetric cell division in *Drosophila* (David et al., 2005; Hampoelz et al., 2005; Wang et al., 2005).

Proteins that are implicated in receptor-independent G protein signaling, like GoLoco motif containing proteins and RIC-8, clearly play an important role in regulating asymmetric cell division, although their exact roles remain poorly understood. A putative G protein regulatory cycle in asymmetric cell division in *C. elegans* has been recently proposed (Hampoelz and Knoblich, 2004) and is shown in Figure 1.10.

1.3.2   Role of AGS3/LGN in mammalian cell division

GoLoco motif-containing proteins like AGS3 and LGN play important roles in mammalian cell division, although they are not as well understood as their *Drosophila* or *C. elegans* counterparts. LGN is localized to a constriction between dividing cells during cytokinesis, known as the 'midbody', in PC12 cells (Blumer et al., 2002). LGN is localized in cell cortex during mitosis and this localization is driven by the GoLoco motif containing C-terminal domain (Kaushik et al., 2003). Loss or over-expression of LGN led to mitotic arrest (Kaushik et al., 2003). The most mechanisticically detailed study revealed the interaction between LGN and NuMA (nuclear mitotic apparatus protein) (Du and Macara, 2004). NuMA is a large coiled-coil protein that localizes in spindle poles during mitosis and interacts with LGN through its C-terminal domain. In yeast two-hybrid screens, the GoLoco-rich LGN C-terminal domain interacted with the TPR-rich LGN N-terminal domain. Indeed, FRET (fluorescence resonance energy transfer) based study with N-terminal YFP (yellow fluorescent protein) and C-terminal CFP (cyan fluorescent protein) tagged LGN showed that

LGN stayed in an auto-inhibited conformation within the cell, where the N-terminal domain and the C-terminal interact with each other (Du and Macara, 2004). The auto-inhibited state was disrupted by either $G\alpha_{i1}$ or the C-terminal domain of NuMA. Binding of $G\alpha_{i1}$ to the GoLoco-rich C-terminal domain of LGN allowed the N-terminal domain of LGN to bind NuMA, or vice versa (Du and Macara, 2004). This observation eventually led to the idea that the auto-inhibited conformation of LGN is released by membrane-bound $G\alpha_{i1}$, which then allows the N-terminal TPR domain to recruit NuMA through its interaction with the NuMA C-terminal domain. Upon cortical localization, NuMA regulates mitotic spindle orientation by yet unknown mechanisms (Du and Macara, 2004).

1.4 Novel peptide regulators of G protein signaling

1.4.1 Structure-function studies of G proteins with combinatorial peptide libraries

Peptides obtained from screening combinatorial peptides libraries have been widely used to study structure-function relationship in proteins. 'Phage display' (Scott and Smith, 1990) is the most commonly used tool for generation of novel peptide ligands of proteins (Sidhu et al., 2003; Szardenings, 2003). Phage display has been used widely to identify peptide ligands for proteins in G protein signaling pathways (Ja and Roberts, 2005). Many of these peptides served as excellent tools for understanding G protein signaling (Taylor and Neubig, 1994). Phage display generated peptides, based on the C-terminal residues for $G\alpha_t$, were used for mapping the $G\alpha_t$ interacting residues in rhodopsin (Martin et al., 1996). Peptides that accelerate the binding of GTPγS to $G\alpha_{i1}$ have been identified using phage

display (Hessling et al., 2003). Phage display was used identify a binding 'hot spot' on Gβγ

binding surface (Scott et al., 2001) and structural studies revealed the mechanistic basis of

the interaction of Gβγ with multiple targets (Davis et al., 2005). Recently, phage display was

used to identify a peptide GEF for $G\alpha_{i1}$; the crystal structure of this novel GEF in complex

with $G\alpha_{i1}$ provided insights into the mechanism of GEF activity (Johnston CA, 2005)

(Detailed discussion of the GEF peptide in chapter 6.3.6).

1.4.2 Identification of peptide ligands using "*in vitro* mRNA display"

While phage display has been the most popular method for identification of novel

peptide ligands, *in vitro* techniques like "mRNA display" is gaining popularity (Takahashi et

al., 2003). mRNA display is a very powerful technique for generating large and diverse

libraries; it can generate libraries of $>10^{12}$ complexity, while phage display is limited to 10^9-

10^{10} complexity (Takahashi et al., 2003). In the mRNA display technology, a large double

stranded DNA (dsDNA) library is first created with randomized sequences. The dsDNA

library is transcribed *in vitro* into an mRNA library, and the resulting mRNAs are ligated

with a puromycin group at the 3'-end. The puromycin-coupled mRNA library is *in vitro*

translated to express the encoded peptides; during the translation reaction the encoded

peptides get covalently attached to the encoding mRNA through the puromycin ring. The

peptide-tagged mRNA library is then converted to a cDNA library, which results in a library

of cDNA fused with its encoded peptides. The peptide-DNA fusion is used for selection

experiments, and the positive hits are amplified using PCR to identify the sequence of the

ligand (Roberts and Szostak, 1997; Takahashi et al., 2003).

1.4.3   Identification of R6A: A novel GDI

The '*in vitro* mRNA display' method has been recently used to identify novel

regulators of G protein signaling (Ja and Roberts, 2004). The design of these peptides was

inspired by GoLoco motif sequences. To identify new ligands for G proteins, Ja and

colleagues subjected the C terminal half of 'GPR-consensus' peptide (Peterson et al., 2000)

sequence to randomization and used biotinylated $G\alpha_{i1}$ as a target. Based on the C-terminal

half of the GPR-consensus peptide, new peptides were obtained with mutations at the

position of the conserved arginine (Willard et al., 2004) residue at the C-terminus of GoLoco

motif (Ja and Roberts, 2004). The dominant peptide was a 16 residue peptide (R6A4) with

similarity at N-terminal region to the GPR-consensus peptide, but the C-terminal region with

a completely different sequence (Figure 1.11). The K_d of R6A4 for $G\alpha_{i1}$ binding was only 60

nM, which is comparable with the K_d of the GPR-consensus peptide (82 nM). The 16-residue

R6A4 peptide can be minimized to a 9 residue core region with the sequence

DQLYWWEYL (Figure 1.11), which still binds $G\alpha_{i1}$ with a K_d of 200 nM. The first three N-

terminal residues of the core, which is DQL, replace the conserved 'DQR triad' (Willard et

al., 2004) observed in GoLoco motif. When the leucine residue in R6A4, which corresponds

to the conserved arginine residue (Figure 1.11), was replaced with arginine, no binding was

observed (Ja and Roberts, 2004). R6A4 not only lacked the conserved arginine residue, it

also lacked the N-terminal helical region of GPR-consensus peptide, which has been shown

to form a helix in the RGS14-GoLoco motif when bound to $G\alpha_{i1}$ (Kimple et al., 2002a).

R6A4 only binds to $G\alpha\cdot$GDP and does not bind $G\alpha\cdot$GTP. Despite of sequence

differences, R6A4 was as active as GPR-consensus peptide in inhibiting the binding of

BODIPY-GTPγS (a fluorescent GTP analog) to Gα$_{i1}$. Much like R6A4, the 9 residue

minimized peptide was also a GDI (Ja and Roberts, 2004). The binding of BOIDPY-GTPγS

to Gα$_{i1}$ followed a biphasic pattern in the presence of R6A4, with a fast 'burst' phase a

slower phase. While this is in contrast with the GPR-consensus peptide, the reason behind

this biphasic behavior was not known (Ja and Roberts, 2004). The uniqueness of the R6A4

peptide and complete lack of sequence identity of the 9 residue core with GoLoco motifs

suggests that R6A4 might have a different mechanism of action.
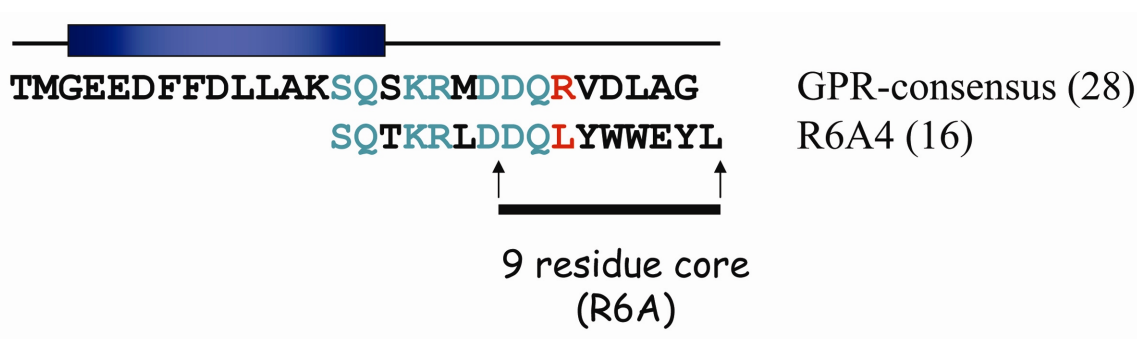
Figure 1.11. Sequence of R6A4 peptide. Sequence alignment of R6A4 and GPR-consensus peptide is shown. The conserved residues are shown in blue; the mutation of the conserved arginine residue of GoLoco to a leucine residue is highlighted in red. The length of the sequences are indicated in parenthesis. The helical region of GPR-consensus peptide is indicated as a blue rectangle. The 9 residue core region is shown.

1.5 Description of dissertation research

The dissertation research described here broadly attempts to investigate the regulation

of nucleotide exchange in Gα$_{i1}$ by GDIs; two distinct aspects are addressed: the

thermodynamic nature of the interaction between the four-GoLoco motif protein AGS3 and

Gα$_{i1}$, and the molecular basis of R6A4-mediated GDI activity.

GoLoco motif-containing proteins are indispensable components in regulating asymmetric cell division. The presence of tandem GoLoco repeats in proteins like AGS3, LGN, and PINS suggests that such domain organization can be functionally important. Inasmuch as a single GoLoco motif can possess GDI activity, the biochemical advantage conferred by proteins such as AGS3, which possess multiple copies of GPR motifs in tandem repeats, is worthy of investigation. The presumption that each of the GPR motifs in AGS3 possesses GDI activity has received experimental support, suggesting that AGS3 and its homologs might be able to bind several GDP-bound $G\alpha$ subunits and serve as scaffolds (Bernard et al., 2001a). Although it is possible that each of the four GPR motifs in AGS3 can bind one molecule of $G\alpha_{i1}$, the actual stoichiometry of the complex had not been determined. It was observed that an AGS3 construct containing only two GoLoco repeats is less potent than one containing all four (Natochin et al., 2000). Therefore, GoLoco motifs might function cooperatively, for example by adopting a more stable tertiary structure in the context of AGS3, than would individual GoLoco motifs in isolation. It is also possible that the GoLoco repeats within AGS3 differ in their $G\alpha$ specificity, affinity or GDI activity. These questions were addressed using isothermal titration calorimetry (ITC) to determine the stoichiometry and thermodynamic parameters associated with binding of $G\alpha_{i1} \cdot GDP$ to the AGS3-C terminal domain and constructs of AGS3 from which one or more GoLoco motifs have been deleted. In a complementary series of experiments, the affinity and GDI activity of peptides derived from each of the GoLoco motifs of AGS3 was measured in order to establish a correlation between affinity and biological activity. These experiments provide direct insight into the thermodynamic basis for the GDI activity of AGS3, and unexpected

insights into the relationship between activity and binding affinity. The results of this part of the research will be the subject of discussion in chapters three and four. The methodologies developed or adopted for the purification of $G\alpha_{i1}$, AGS3, and AGS3 deletion mutants are described in chapter two.

The novel G-protein regulatory peptide R6A4, which is a GDI, has no sequence similarity with naturally occurring GDIs like the GoLoco motif. R6A4 is not only novel in its sequence, but is also shorter in length than GoLoco motifs. Therefore it is possible that R6A4 inhibits nucleotide exchange by a novel mechanism. Although R6A4 has no biological significance, it is possible that R6A4 can be used as a tool to enhance our understanding of the molecular basis of GDI activity. Peptides identified from screening of randomized libraries have been used successfully in many instances to gain insight into G-protein signaling (chapter 1.4.1). To unravel the molecular basis of R6A4 activity, the crystal structure of the 9-residue R6A core region in complex with $G\alpha_{i1}$ was determined. The crystal structure allowed the identification of structural elements in $G\alpha_{i1}$ that participate in nucleotide exchange. Interestingly, mutations in a potentially important region of the peptide converted the GDI peptide to a GEF. The procedure followed to obtain the crystal structure, the description of the structure, and mechanistic insights gained into the regulation of nucleotide exchange are subjects of discussion in chapters five and six.

1.6 References

Abdulaev, N. G., Ngo, T., Zhang, C., Dinh, A., Brabazon, D. M., Ridge, K. D., and Marino, J. P. (2005). Heterotrimeric G-protein alpha -subunit adopts a 'pre-activated' conformation when associated with beta gamma -subunits. J Biol Chem.

Afshar, K., Willard, F. S., Colombo, K., Siderovski, D. P., and Gonczy, P. (2005). Cortical localization of the G{alpha} protein GPA-16 requires RIC-8 function during C. elegans asymmetric cell division. Development *132*, 4449-4459.

Bardwell L, C. J., Inouye CJ, Thorner J. (1994). Signal propagation and regulation in the mating pheromone response pathway of the yeast Saccharomyces cerevisiae. Dev Biol *166*, 363-379.

Bellaiche Y, R. A., Woods DF, Hough CD, Parmentier ML, O'Kane CJ, Bryant PJ, Schweisguth F. (2001). The Partner of Inscuteable/Discs-large complex is required to establish planar polarity during asymmetric cell division in Drosophila. Cell *106*, 355-366.

Berman, D. M., and Gilman, A. G. (1998). Mammalian RGS proteins: Barbarians at the gate. JBiolChem *273*, 1269-1272.

Bernard, M. L., Peterson, Y. K., Chung, P., Jourdan, J., and Lanier, S. M. (2001a). Selective Interaction of AGS3 with G-proteins and the Influence of AGS3 on the Activation State of G-proteins. J Biol Chem *276*, 1585-1593.

Bernard, M. L., Peterson, Y. K., Chung, P., Jourdan, J., and Lanier, S. M. (2001b). Selective interaction of AGS3 with G-proteins and the influence of AGS3 on the activation state of G-proteins. J Biol Chem *276*, 1585-1593.

Betschinger, J., and Knoblich, J. A. (2004). Dare to be different: asymmetric cell division in Drosophila, C. elegans and vertebrates. Curr Biol *14*, R674-685.

Blatch GL, L. M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. Bioessays *21*, 932-939.

Blumer, J. B., Chandler, L. J., and Lanier, S. M. (2002). Expression analysis and subcellular distribution of the two G-protein regulators AGS3 and LGN indicate distinct functionality. Localization of LGN to the midbody during cytokinesis. J Biol Chem *277*, 15897-15903.

Bourne, H. R. (1997). How receptors talk to trimeric G proteins. Curr Opin Cell Biol *9*, 134-142.

Cabrera-Vera, T. M., Vanhauwe, J., Thomas, T. O., Medkova, M., Preininger, A., Mazzoni, M. R., and Hamm, H. E. (2003). Insights into G protein structure, function, and regulation. Endocr Rev *24*, 765-781.

Cai, K., Itoh, Y., and Khorana, H. G. (2001). Mapping of contact sites in complex formation between transducin and light-activated rhodopsin by covalent crosslinking: use of a photoactivatable reagent. Proc Natl Acad Sci U S A *98*, 4877-4882.

Cherfils, J., and Chabre, M. (2003). Activation of G-protein Galpha subunits by receptors through Galpha-Gbeta and Galpha-Ggamma interactions. Trends Biochem Sci *28*, 13-17.

Cherfils J, C. M. (2003). Activation of G-protein Galpha subunits by receptors through Galpha-Gbeta and Galpha-Ggamma interactions. Trends Biochem Sci *28*, 13-17.

Cismowski, M. J., Takesono, A., Bernard, M. L., Duzic, E., and Lanier, S. M. (2001). Receptor-independent activators of heterotrimeric G-proteins. Life Sci *68*, 2301-2308.

Cismowski, M. J., Takesono, A., Ma, C., Lizano, J. S., Xie, X., Fuernkranz, H., Lanier, S. M., and Duzic, E. (1999). Genetic screens in yeast to identify mammalian nonreceptor modulators of G-protein signaling [see comments]. Nat Biotechnol *17*, 878-883.

Coleman, D. E., Berghuis, A. M., Lee, E., Linder, M. E., Gilman, A. G., and Sprang, S. R. (1994). Structures of active conformations of $G_{i\alpha 1}$ and the mechanism of GTP hydrolysis. Sci *265*, 1405-1412.

Coleman, D. E., and Sprang, S. R. (1999). Structure of $G_{i\alpha 1} \cdot$GppNHp, Autoinhibition in a G_{α} Protein-Substrate Complex. J Biol Chem *274*, 16669-16672.

Colombo, K., Grill, S. W., Kimple, R. J., Willard, F. S., Siderovski, D. P., and Gonczy, P. (2003). Translation of polarity cues into asymmetric spindle positioning in Caenorhabditis elegans embryos. Sci *300*, 1957-1961.

Couwenbergs, C., Spilker, A. C., and Gotta, M. (2004). Control of embryonic spindle positioning and Galpha activity by C. elegans RIC-8. Curr Biol *14*, 1871-1876.

David, N. B., Martin, C. A., Segalen, M., Rosenfeld, F., Schweisguth, F., and Bellaiche, Y. (2005). Drosophila Ric-8 regulates Galphai cortical localization to promote Galphai-dependent planar orientation of the mitotic spindle during asymmetric cell division. Nat Cell Biol.

Davis, T. L., Bonacci, T. M., Sprang, S. R., and Smrcka, A. V. (2005). Structural and molecular characterization of a preferred protein interaction surface on G protein betagamma subunits. Biochemistry *44*, 10593-10604.

De Vries, L., Fischer, T., Tronchere, H., Brothers, G. M., Strockbine, B., Siderovski, D. P., and Farquhar, M. G. (2000a). Activator of G protein signaling 3 is a guanine dissociation inhibitor for Galpha i subunits. Proc Natl Acad Sci U S A *97*, 14364-14369.

De Vries, L., Fischer, T., Tronchere, H., Brothers, G. M., Strockbine, B., Siderovski, D. P., and Farquhar, M. G. (2000b). Activator of G protein signaling 3 is a guanine dissociation inhibitor for galpha i subunits [In Process Citation]. Proc Natl Acad Sci U S A *97*, 14364-14369.

Du, Q., and Macara, I. G. (2004). Mammalian Pins is a conformational switch that links NuMA to heterotrimeric G proteins. Cell *119*, 503-516.

Ferguson, K. M., Higashijima, T., Smigel, M. D., and Gilman, A. G. (1986). The influence of bound GDP on the kinetics of guanine nucleotide binding to G proteins. J Biol Chem *261*, 7393-7399.

Gilman, A. G. (1987). G Proteins:transducers of receptor-generated signals. Ann Rev Biochem *56*, 615-649.

Gotta, M., Dong, Y., Peterson, Y. K., Lanier, S. M., and Ahringer, J. (2003). Asymmetrically distributed C. elegans homologs of AGS3/PINS control spindle position in the early embryo. Curr Biol *13*, 1029-1037.

Hamm, H. E. (1998). The many faces of G protein signaling. J Biol Chem *273*, 669-672.

Hamm, H. E. (2001). How activated receptors couple to G proteins. Proc Natl Acad Sci U S A *98*, 4819-4821.

Hampoelz, B., Hoeller, O., Bowman, S. K., Dunican, D., and Knoblich, J. A. (2005). Drosophila Ric-8 is essential for plasma-membrane localization of heterotrimeric G proteins. Nat Cell Biol.

Hampoelz, B., and Knoblich, J. A. (2004). Heterotrimeric G proteins: new tricks for an old dog. Cell *119*, 453-456.

Herrmann, R., Heck, M., Henklein, P., Henklein, P., Kleuss, C., Hofmann, K. P., and Ernst, O. P. (2004). Sequence of interactions in receptor-G protein coupling. J Biol Chem *279*, 24283-24290.

Herrmann R, H. M., Henklein P, Henklein P, Kleuss C, Hofmann KP, Ernst OPJ Biol Chem. (2004). Sequence of interactions in receptor-G protein coupling. J Biol Chem *279*, 24283-24290.

Hess, H. A., Roper, J. C., Grill, S. W., and Koelle, M. R. (2004). RGS-7 completes a receptor-independent heterotrimeric G protein cycle to asymmetrically regulate mitotic spindle positioning in C. elegans. Cell *119*, 209-218.

Hessling, J., Lohse, M. J., and Klotz, K. N. (2003). Peptide G protein agonists from a phage display library. Biochem Pharmacol *65*, 961-967.

Higashijima, T., Ferguson, K. M., Sternweis, P. C., Smigel, M. D., and Gilman, A. G. (1987). Effects of Mg2+ and the beta gamma-subunit complex on the interactions of guanine nucleotides with G proteins. J Biol Chem *262*, 762-766.

Higashijima, T., and Ross, E. M. (1991). Mapping of the mastoparan-binding site on G proteins. Cross-linking of $[1^{25}I-Tyr^3,Cys^{11}]$mastoparan to G_o. JBiolChem *266*, 12655-12661.

Higashijima, T., Uzu, S., Nakajima, T., and Ross, E. M. (1988). Mastoparan, a peptide toxin from wasp venom, mimics receptors by activating GTP-binding regulatory proteins (G proteins). JBiolChem *263*, 6491-6494.

Iiri, T., Farfel, Z., and Bourne, H. R. (1998). G-protein diseases furnish a model for the turn-on switch. Nature *394*, 35-38.

Ja, W. W., and Roberts, R. W. (2004). In vitro selection of state-specific peptide modulators of G protein signaling using mRNA display. Biochemistry *43*, 9265-9275.

Ja, W. W., and Roberts, R. W. (2005). G-protein-directed ligand discovery with peptide combinatorial libraries. Trends Biochem Sci *30*, 318-324.

Jahangeer, S., and Rodbell, M. (1993). The disaggregation theory of signal transduction revisited: Further evidence that G proteins are multimeric and disaggregate to monomers when activated. Proceedings of the National Academy of Sciences USA *90*, 8782-8786.

Johnston CA, W. F., Jezyk MR, Fredericks Z, Bodor ET, Jones MB, Blaesius R, Watts VJ, Harden TK, Sondek J, Ramer JK, Siderovski DP. (2005). Structure of Galpha(i1) bound to a GDP-selective peptide provides insight into guanine nucleotide exchange. Structure (Camb) *13*, 1069-1080.

Kaushik, R., Yu, F., Chia, W., Yang, X., and Bahri, S. (2003). Subcellular localization of LGN during mitosis: evidence for its cortical localization in mitotic cell culture systems and its requirement for normal cell cycle progression. Mol Biol Cell *14*, 3144-3155.

Kimple, R. J., De Vries, L., Tronchere, H., Behe, C. I., Morris, R. A., Gist Farquhar, M., and Siderovski, D. P. (2001). RGS12 and RGS14 GoLoco motifs are G alpha(i) interaction sites with guanine nucleotide dissociation inhibitor Activity. J Biol Chem *276*, 29275-29281.

Kimple, R. J., Kimple, M. E., Betts, L., Sondek, J., and Siderovski, D. P. (2002a). Structural determinants for GoLoco-induced inhibition of nucleotide release by Galpha subunits. Nature *416*, 878-881.

Kimple, R. J., Willard, F. S., Hains, M. D., Jones, M. B., Nweke, G. K., and Siderovski, D. P. (2004). Guanine nucleotide dissociation inhibitor activity of the triple GoLoco motif protein G18: alanine-to-aspartate mutation restores function to an inactive second GoLoco motif. Biochem J *378*, 801-808.

Kimple, R. J., Willard, F. S., and Siderovski, D. P. (2002b). The GoLoco motif: heralding a new tango between G protein signaling and cell division. Mol Interv *2*, 88-100.

Kleuss, C., Raw, A., Lee, E., Sprang, S., and Gilman, A. (1994). Mechanism of GTP hydrolysis by G-protein alpha subunits. Proc Natl Acad Sci U S A *91*, 9828-9831.

Knoblich, J. A. (2001). Asymmetric cell division during animal development. Nat Rev Mol Cell Biol *2*, 11-20.

Knust, E. (2001). G protein signaling and asymmetric cell division. Cell *107*, 125-128.

Lambright, D. G., Noel, J. P., Hamm, H. E., and Sigler, P. B. (1994). Structural determinants for activation of the alpha-subunit of a heterotrimeric G protein. Nature *369*, 621-628.

Lambright, D. G., Sondek, J., Bohm, A., Skiba, N. P., Hamm, H., and Sigler, P. B. (1996). The 2.0Å crystal structure of a heterotrimeric G protein. Nature *379*, 311-319.

Lanier, S. M. (2004). AGS proteins, GPR motifs and the signals processed by heterotrimeric G proteins. Biol Cell *96*, 369-372.

Levitzki A, K. S. (2002). G-protein subunit dissociation is not an integral part of G-protein action. Chembiochem *3*, 815-818.

Ma, H., Peterson, Y. K., Bernard, M. L., Lanier, S. M., and Graber, S. G. (2003). Influence of cytosolic AGS3 on receptor--G protein coupling. Biochemistry *42*, 8085-8093.

Martin, E. L., Rens-Domiano, S., Schatz, P. J., and Hamm, H. E. (1996). Potent peptide analogues of a G protein receptor-binding region obtained with a combinatorial library. J Biol Chem *271*, 361-366.

McCudden, C. R., Willard, F. S., Kimple, R. J., Johnston, C. A., Hains, M. D., Jones, M. B., and Siderovski, D. P. (2005). G alpha selectivity and inhibitor function of the multiple GoLoco motif protein GPSM2/LGN. Biochim Biophys Acta *1745*, 254-264.

Mittal, V., and Linder, M. E. (2004). The RGS14 GoLoco domain discriminates among Galphai isoforms. J Biol Chem *279*, 46772-46778.

Mixon, M. B., Lee, E., Coleman, D. E., Berghuis, A. M., Gilman, A. G., and Sprang, S. R. (1995). Tertiary and quaternary structural changes in $G_{i\alpha1}$ induced by GTP hydrolysis. Sci *270*, 954-960.

Nance, J. (2005). PAR proteins and the establishment of cell polarity during C. elegans development. Bioessays *27*, 126-135.

Natochin, M., Gasimov, K. G., and Artemyev, N. O. (2001). Inhibition of GDP/GTP exchange on G alpha subunits by proteins containing G-protein regulatory motifs. Biochemistry *40*, 5322-5328.

Natochin, M., Gasimov, K. G., and Artemyev, N. O. (2002). A GPR-protein interaction surface of Gi(alpha): implications for the mechanism of GDP-release inhibition. Biochemistry *41*, 258-265.

Natochin, M., Lester, B., Peterson, Y. K., Bernard, M. L., Lanier, S. M., and Artemyev, N. O. (2000). AGS3 inhibits GDP dissociation from galpha subunits of the Gi family and rhodopsin-dependent activation of transducin [In Process Citation]. J Biol Chem *275*, 40981-40985.

Noel, J. P., Hamm, H. E., and Sigler, P. B. (1993). The 2.2 A crystal structure of transducin-alpha complexed with GTPγS. Nature *366*, 654-663.

Onrust, R., Herzmark, P., Chi, P., Garcia, P., Lichtarge, O., Kingsley, C., and Bourne, H. (1997). Receptor and betagamma binding sites in the alpha subunit of the retinal G protein transducin. Sci *275*, 381-384.

Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E*., et al.* (2000). Crystal structure of rhodopsin: A G protein-coupled receptor. Sci *289*, 739-745.

Pellettieri, J., and Seydoux, G. Anterior-posterior polarity in C. elegans and Drosophila--PARallels and differences. Science *298*, 1946-1950.

Peterson, Y. K., Bernard, M. L., Ma, H., Hazard, S., 3rd, Graber, S. G., and Lanier, S. M. (2000). Stabilization of the GDP-bound conformation of Gialpha by a peptide derived from the G-protein regulatory motif of AGS3. J Biol Chem *275*, 33193-33196.

Peterson, Y. K., Hazard, S., 3rd, Graber, S. G., and Lanier, S. M. (2002). Identification of structural features in the G-protein regulatory motif required for regulation of heterotrimeric G-proteins. J Biol Chem *277*, 6767-6770.

Pizzinat, N., Takesono, A., and Lanier, S. M. (2001). Identification of a truncated form of the G-protein regulator AGS3 in heart that lacks the tetratricopeptide repeat domains. J Biol Chem *276*, 16601-16610.

Roberts, R. W., and Szostak, J. W. (1997). RNA-peptide fusions for the in vitro selection of peptides and proteins. Proc Natl Acad Sci U S A *94*, 12297-12302.

Rondard P, I. T., Srinivasan S, Meng E, Fujita T, Bourne HR. (2001). Mutant G protein alpha subunit activated by Gbeta gamma: a model for receptor activation? Proc Natl Acad Sci U S A *98*, 6150-6155.

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

Schaefer, M., Petronczki, M., Dorner, D., Forte, M., and Knoblich, J. A. (2001). Heterotrimeric G proteins direct two modes of asymmetric cell division in the Drosophila nervous system. Cell *107*, 183-194.

Scott, J. K., Huang, S. F., Gangadhar, B. P., Samoriski, G. M., Clapp, P., Gross, R. A., Taussig, R., and Smrcka, A. V. (2001). Evidence that a protein-protein interaction 'hot spot' on heterotrimeric G protein betagamma subunits is used for recognition of a subclass of effectors. Embo J *20*, 767-776.

Scott, J. K., and Smith, G. P. (1990). Searching for peptide ligands with an epitope library. Sci *249*, 386-390.

Siderovski, D. P., Diverse-Pierluissi, M., and De Vries, L. (1999). The GoLoco motif: a Galphai/o binding motif and potential guanine- nucleotide exchange factor. Trends Biochem Sci *24*, 340-341.

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.

Sidhu, S. S., Fairbrother, W. J., and Deshayes, K. (2003). Exploring protein-protein interactions with phage display. Chembiochem *4*, 14-25.

Sondek, J., Bohm, A., Lambright, D. G., Hamm, H. E., and Sigler, P. B. (1996). Crystal structure of a G-protein βγ dimer at 2.1 Å resolution. Nature *379*, 369-374.

Sondek, J., Lambright, D. G., Noel, J. P., Hamm, H. E., and Sigler, P. B. (1994). GTPase mechanism of G proteins from the 1.7 Å crystal structure of transducin α-GDP-AlF$_4^-$. Nature *372*, 276-279.

Sprang, S. R. (1997a). G protein mechanisms: Insights from structural analysis. Annu Rev Biochem *66*, 639-678.

Sprang, S. R. (1997b). G proteins, effectors and GAPs: structure and mechanism. Current Opinion in Structural Biology *7*, 849-856.

Srinivasan, D. G., Fisk, R. M., Xu, H., and van den Heuvel, S. (2003). A complex of LIN-5 and GPR proteins regulates G protein signaling and spindle function in C elegans. Genes Dev *17*, 1225-1239.

Szardenings, M. (2003). Phage display of random peptide libraries: applications, limits, and potential. J Recept Signal Transduct Res *23*, 307-349.

Takahashi, T. T., Austin, R. J., and Roberts, R. W. (2003). mRNA display: ligand discovery, interaction analysis and beyond. Trends Biochem Sci *28*, 159-165.

Takesono, A., Cismowski, M. J., Ribas, C., Bernard, M., Chung, P., Hazard, S., 3rd, Duzic, E., and Lanier, S. M. (1999). Receptor-independent activators of heterotrimeric G-protein signaling pathways. J Biol Chem *274*, 33202-33205.

Tall, G. G., Krumins, A. M., and Gilman, A. G. (2003). Mammalian Ric-8A (synembryn) is a heterotrimeric Galpha protein guanine nucleotide exchange factor. J Biol Chem *278*, 8356-8362.

Taylor, J. M., and Neubig, R. R. (1994). Peptides as probes for G protein signal transduction. Cell Signal *6*, 841-849.

Tesmer, J. J. G., Berman, D. M., Gilman, A. G., and Sprang, S. R. (1997a). Structure of RGS4 bound to AIF$_4^-$-activated G$_{i\alpha1}$:Stabilization of the transition state for GTP hydrolysis. Cell *89*, 251-261.

Tesmer, J. J. G., Sunahara, R. K., Gilman, A. G., and Sprang, S. R. (1997b). Crystal structure of the catalytic domains of adenylyl cyclase in a complex with G$_{s\alpha}$•GTPγS. Sci *278*, 1907-1916.

Wall, M. A., Coleman, D. E., Lee, E., Iñiguez-Lluhi, J. A., Posner, B. A., Gilman, A. G., and Sprang, S. R. (1995). The structure of the G protein heterotrimer G$_{i\alpha1}\beta_1\gamma_2$. Cell *80*, 1047-1058.

Wang, H., Ng, K. H., Qian, H., Siderovski, D. P., Chia, W., and Yu, F. (2005). Ric-8 controls Drosophila neural progenitor asymmetric division by regulating heterotrimeric G proteins. Nat Cell Biol.

Willard, F. S., Kimple, R. J., and Siderovski, D. P. (2004). Return of the GDI: the GoLoco motif in cell division. Annu Rev Biochem *73*, 925-951.

Yu, F., Ong, C. T., Chia, W., and Yang, X. (2002). Membrane targeting and asymmetric localization of Drosophila partner of inscuteable are discrete steps controlled by distinct regions of the protein. Mol Cell Biol *22*, 4230-4240.

Chapter Two

Expression and purification of AGS3 and Gα_{i1}

2.1   Introduction

Crystallization as well as some biochemical studies of proteins requires large scale expression and purification of the proteins involved. To study the interaction between AGS3 and Gα_{i1}, these proteins were expressed in *E.coli* as recombinant proteins. The purification of wild-type Gα_{i1} was optimized earlier (Lee et al., 1994); however the large scale expression of AGS3 needed to be optimized for the purposes of this study. The purification of Gα_{i1} was improved using a GST-tagged construct. Various truncation mutants of AGS3 and Gα_{i1} were also purified. In this chapter, I will describe the creation of several AGS3 constructs, the purification of AGS3, and the purification of Gα_{i1} using the original untagged construct and the modified Gα_{i1} constructs.

2.2   Materials and Methods

2.2.1   Expression and purification of hexa-histidine tagged AGS3

The N-terminally hexa-histidine tagged C-terminal domain of AGS3 (residues 465-650), hereafter referred to as AGS3-C, cloned in the expression vector pQE30 (Figure 2.1) was a gift from Prof. Stephen Lanier (Department of Pharmacology, Louisiana State

University Medical Center, New Orleans, LA). This domain contains four GoLoco motifs:

GoLoco I, residues 470-489; GoLoco II, residues 524-542; GoLoco III, residues 572-590;

and GoLoco IV, residues 606-624.

> **pQE30-6His-AGS3(465-650)**

MRGS**HHHHHH**GSTMAPSSD**EECFFDLLSKFQSSRMDDQR**CPLEEGQAGAAEATAAPTLE
ERAAQPSVTASPQT**EEFFDLIASSQSRRLDDQR**ASVGSLPGLRITLNNVGHLRGDGDPQ
EPG**DEFFNMLIKYQSSRIDDQR**CPPPDVLPRGPTMPD**EDFFSLIQRVQAKRMDEQR**VDL
AGSPDQEASGLPDPRQQCPPGAS

Figure 2.1. Sequence of 6His-AGS3-C construct. The sequence of the N-terminal 6 His-tagged AGS3-C terminal domain (residues 465-650) cloned in the expression vector pQE30 is shown. The 6 His-tag is shown in red and the sequence of AGS3-C is shown in blue. GoLoco I (470-489), GoLoco II (524-542), GoLoco III (572-590), and GoLoco IV (606-624) are underlined in bold font.

E.coli JM109 (DE3) cells harboring the AGS3-C expression vector were grown in LB

media containing 100 µg/ml ampicilin as a selection marker. Cells were grown up to

OD600~0.8 and induced with 1 mM IPTG for 5 h at 37°C. The induced cells were harvested

by centrifugation, flash frozen in liquid nitrogen and kept at -80°C for long term storage.

Screening of small scale cultures suggested that His-AGS3-C is expressed in *E.coli,* but stays

insoluble. Upon lysis of cells under non-denaturing conditions, His-AGS3-C was found in

the insoluble fraction. To purify AGS3-C in soluble form, a urea based denaturing protocol

for purification was adopted from QIAexpressionist handbook (http://www1.qiagen.com-

literature-handbooks-PDF-Protein-Expression-QXP_QIAexpressionist-

1024473_QXPHB_0603.pdf) (Qiagen, 2003), which was followed by refolding. The urea

solubilized AGS3 was refolded by slowly reducing the amount of urea.

Cell lysis was performed by re-suspending the thawed cell stock in lysis buffer (5 ml

buffer/gm of cell) containing 8M urea, 0.1 M NaH_2PO_4, and 10 mM Tris pH 8.0. Upon

incubation for 30 min in the lysis buffer with continuous stirring, the cell lysate was

centrifuged at 10,000 g for 30 min at room temperature. The supernatant was loaded on Ni-

NTA resin pre-equilibrated with wash buffer containing 20 mM Tris pH 7.4, 6 M urea, 500

mM NaCl, and 20 % glycerol. Refolding of the Ni-NTA resin-bound protein was attempted

by running a 0-100% linear gradient (6-1 M urea) of renaturation buffer containing 20 mM

Tris pH 7.4, 1M urea, 500 mM NaCl, and 20% glycerol for 2 h at the flow-rate of 0.5

ml/min. The bound, refolded protein was eluted by a linear gradient of elution buffer

containing 250 mM imidazole in the renaturation buffer (Figure 2.2). Elute protein fractions

containing AGS3-C were dialyzed overnight against 20 mM Tris pH 7.5, 2 mM DTT, 5%

glycerol, 1 mM EDTA, and 1X PTT at 4°C. Dialyzed proteins were purified further using

ion-exchange chromatography. Dialyzed protein sample was loaded on a HiTrapQ™

(Amersham Pharmacia Biosciences) column (column volume: 5 ml) pre-equilibrated with

buffer A containing 20 mM Tris pH 7.5, 2 mM DTT, and 1X PTT. Bound proteins were

eluted using a 0-100% linear gradient of buffer B containing 20 mM Tris pH 7.5, 2 mM

DTT, 1X PTT, and 1M NaCl. HiTrapQ™ elution fractions containing AGS3-C were pooled

together and dialyzed against storage buffer containing 20 mM Tris pH 8.0, 2 mM DTT, 1X

PTT, and 1 mM EDTA overnight at 4°C (Figure 2.2). The dialyzed protein sample was

concentrated using Centriprep™ YM-10 (10 kDa molecular weight cutoff) to a final

concentration of 6-10 mg/ml, flash frozen in liquid nitrogen, and stored at -80°C.
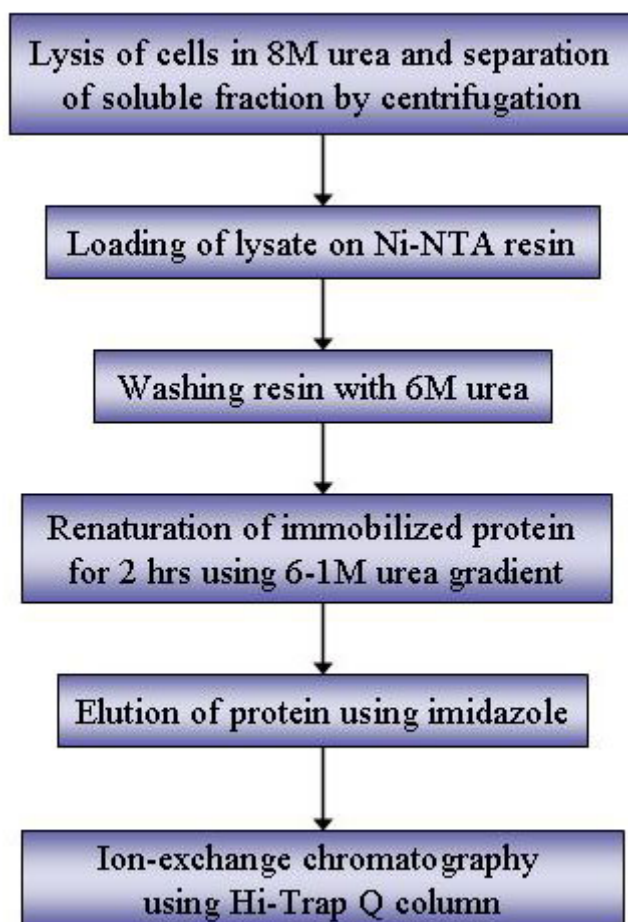
Figure 2.2. Flowchart of 6His-AGS3-C purification.

2.2.2   "Gateway cloning technology" based cloning: A generic strategy

"Gateway cloning technology" (Invitrogen) provides an easy and fast way of creating

constructs with different affinity tags. The reliance of Gateway™ cloning on highly specific

site-directed recombination makes it more efficient than conventional ligation-based cloning
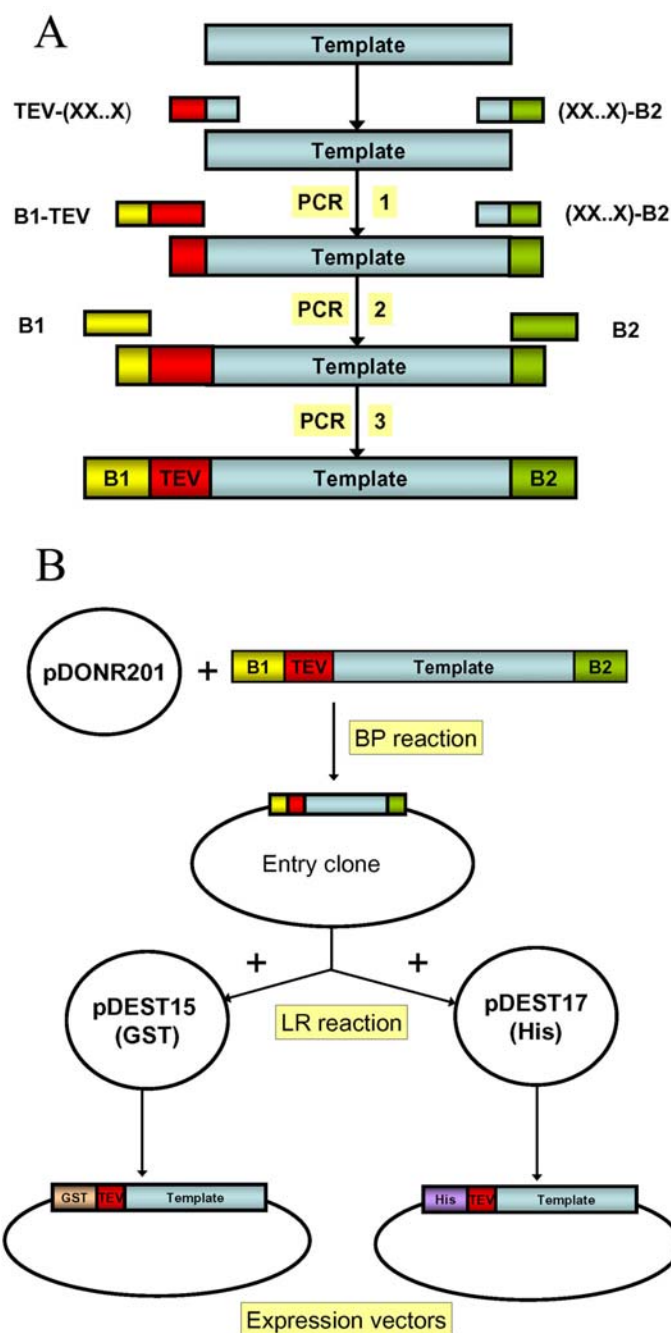
Figure 2.3. Schematic of gateway cloning technology. *A*. Schematic diagram showing the addition of attB1 recombination site and TEV protease cleavage site at the 5' end and attB2 recombination site at the 3'end of DNA template. *B*. Recombination of attB1/B2 flanked DNA template with pDONR201 results in the entry clone which is used for recombination with either His (pDEST17) and GST (pDEST15) based expression vectors to obtain the expression constructs.

strategies. Unlike ligation-based cloning, it does not require specific restriction enzyme sites in the template DNA and the vector (Hartley et al., 2000). Gateway™ reduces the problem of cloning to simply PCR based addition of Gateway™ recombination sites attB1 and attB2 at the 5' and 3' end of a target DNA sequence, respectively (Invitrogen., 2003). To standardize a platform for cloning of different constructs, a PCR based generic strategy was developed with the help of Mr. Ed Belouski (a former technician in the laboratory). The target DNA sequences were amplified in such a manner, that the sequence encoding the desired protein was preceded by a TEV protease recognition sequence for facilitating the removal of the affinity tag during purification. Gateway recombination site attB1 and attB2 was added to the 5' and 3' end of the <TEV recognition site-target DNA> module by performing PCR with B1 and B2 as forward and reverse primers respectively (Figure 2.3). The <attB1-TEV recognition site-target sequence-attB2> module was mixed with the donor vector pDONR201 and entry clones were created using the 'BP reaction' with BP clonase enzyme (Figure 2.3). The products of the BP reactions are referred to as 'entry clones', which were used to transform DH5α cells. Entry clones were purified in larger quantities from successfully transformed DH5α cells. The entry clones were sequenced with forward primer attL1 and reverse primer attL2 to verify the accuracy of the sequences. To make expression constructs, the entry clones were mixed with expression vectors like pDEST 15 (N-terminal GST tag) and pDEST 17 (N-terminal His tag) and expression constructs were obtained by performing the LR reaction with the LR clonase enzyme (Figure 2.3). The protocols for BP and LR reaction were adopted from the Gateway™ manual (Invitrogen., 2003). The products of LR reactions were used to transform DH5α cells. Expression plasmids were purified from DH5α

cells, the accuracy of the DNA sequences were verified, and correct expression vectors were subsequently used for transformation of expression cells like *E.coli* BL21 (DE3). The recombination based cloning strategy is highly effective, resulting in ~100% efficiency in obtaining the correct construct.

The sequences of the generic primers used for Gateway™ cloning are shown below:

<u>Forward primers:</u>

1. 5'-TEV-(18-25 bases from the target sequence):

     5'- TAT TTT CAG GGC-(XX….XX)-3'

2. 5'-B1-TEV:

     5'- AAA AAG CAG GCT CCG AGA ATC TTT ATT TTC AGG GC-3'

3. B1:

     5'- ACA AGT TTG TAC AAA AAA GCA GGC TCC-3'

4. attL1:

     5'- TCG CGT TAA CGC TAG CAT GGA TCT C-3'

<u>Reverse primers:</u>

1. 3'-B2-(18-25 bases from the target sequence):

     5'- AGA AAG CTG GGT-(XX….XX)-3'

2. B2:

     5'- ACC ACT TTG TAC AAG AAA GCT GGG T-3'

3. attL2:

     5'- GTA ACA TCA GAG ATT TTG AGA CAC-3'

2.2.3 Cloning of GST-tagged AGS3 and deletion mutagenesis

The C-terminal domain of AGS3 (465-650), referred to as AGS3-C, was cloned as an N-terminal GST-tagged protein. GST tag is known to enhance solubility of recombinant proteins and also serves as a more specific affinity tag compared to 6- histidine tags. The Gateway™ procedure, as described in the previous section, was used to create the GST-tagged AGS3-C construct. The DNA encoding AGS3-C (Figure 2.4) was amplified by PCR using the 5'-TEV-AGS3 (5'- TAT TTT CAG GGC GCC CCG TCC TCT GAC GAG G – 3') and 3'-B2-AGS3 (5'- AGA AAG CTG GGT TTA GCT GGC ACC TGG GGG – 3') primers. The amplified sequence was used as a template for the subsequent PCR reaction with 5'-B1-TEV and 3'-B2 primers. The product of the second PCR reaction was amplified again using the B1 and B2 primers (primer sequences are described in chapter 2.2.2). The products of the third PCR reaction was purified from agarose gel and used for the BP reaction. Entry clones harboring TEV-AGS3 were used for the LR reaction with N-terminal GST tagged pDEST 15 and N-terminal His-tagged pDEST17 vector to obtain expression constructs. The AGS3-C domain was cloned into GST-tagged pDEST-15 expression vector using the Gateway™ cloning system. A TEV protease cleavage site was inserted between coding regions for GST and AGS3 (Figure 2.4). Cleavage by TEV protease introduces an extra glycine residue at the N terminus.

A

gagaatctttattttcagggcgccccgtcctctgacgaggagtgtttcttcgatctgctg

  E **N** **L** **Y** **F** **Q** **G** A P S S D **E** **E** **C** **F** **F** **D** **L** **L**

      TEV   465

agtaagttccagagcagtcgtatggatgaccaacgctgtcccctggaggaaggccaggct

  S **K** **F** **Q** **S** **S** **R** **M** **D** **D** **Q** **R** C P L E E G Q A

        GoLoco I

ggggctgctgaggctacagctgccccaaccctggaggagagagcggctcagccctctgta

  G A A E A T A A P T L E E R A A Q P S V

acagcttcaccacagactgaagagttctttgacctcattgccagctcccagagccgccgg

  T A S P Q T **E** **E** **F** **F** **D** **L** **I** **A** **S** **S** **Q** **S** **R** **R**

                           GoLoco II

ctggatgaccagagggctagcgtaggcagcctgcctgggcttcgcatcaccctcaacaac

  L **D** **D** **Q** **R** A S V G S L **P** G L R I T L N N

              P548 Stop

gtggggcacctccgaggcgacggggacccccaggagccagggggatgagttttttcaacatg

  V G H L R G D G D P Q E P G **D** **E** **F** **F** **N** **M**

cttatcaaataccagtcctccaggattgatgaccagcgctgtccacccctgatgtgctg

  L **I** **K** **Y** **Q** **S** **S** **R** **I** **D** **D** **Q** **R** C P P P D V L

        GoLoco III

ccccgaggccccaccatgcctgatgaggatttcttcagccttatccagagggtgcaggct

P R G P T M P D **E** **D** **F** **F** **S** **L** **I** **Q** **R** **V** **Q** **A**

P597 Stop   GoLoco IV

aagcggatggatgagcagcgtgtggacctggctgggagtccagaccaagaggccagtggg

  K **R** **M** **D** **E** **Q** **R** V D L A G S P D Q E A S G

ctgcctgatccccggcagcaatgtccgccaggtgccagctaa

  L P D P R Q Q C P P G A S –

                       650

(Figure continued)
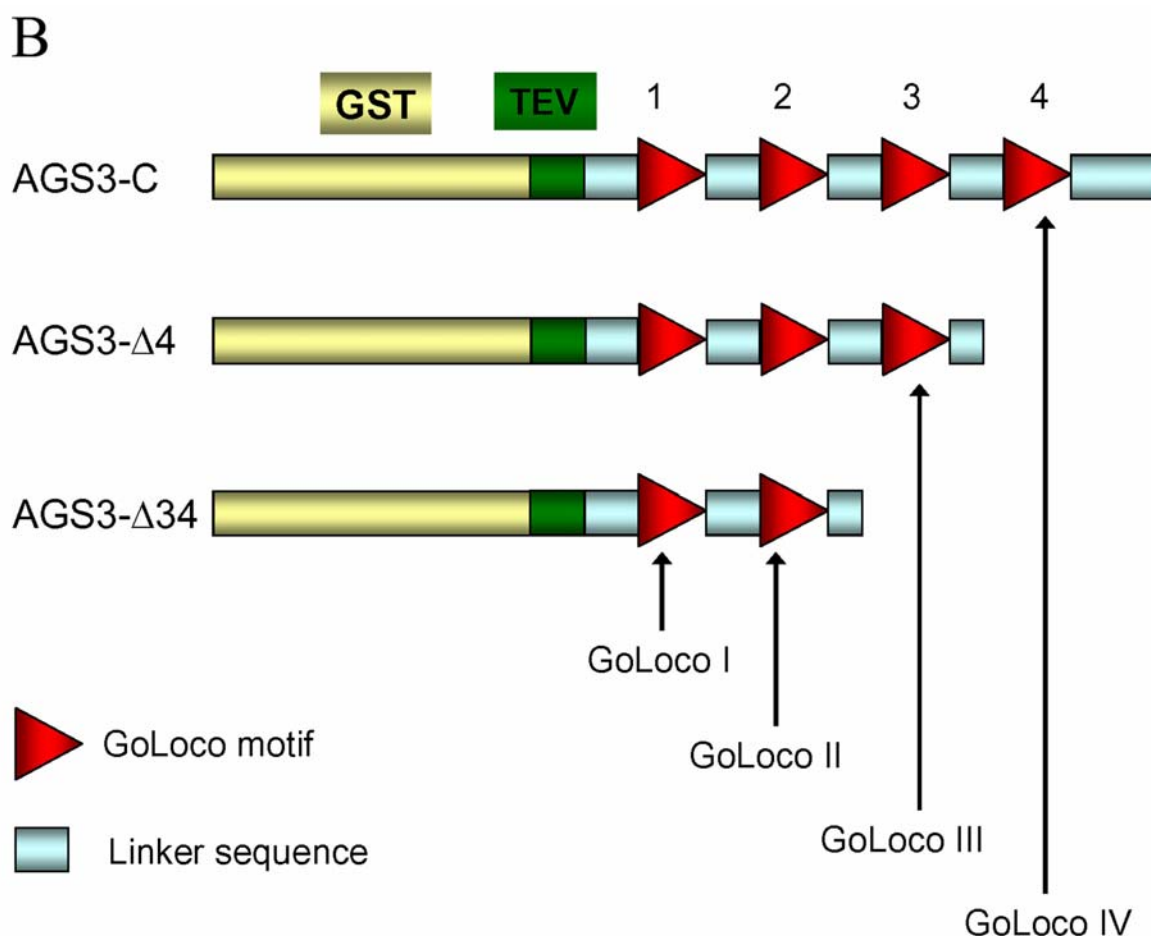
Figure 2.4. Design of GST-fusion AGS3-C and its deletion mutant constructs. *A*. DNA and protein sequence of AGS3-C. TEV site is shown in green and the GoLoco motifs are shown in blue. The stop codons corresponding of AGS3-C (465-650), AGS3-Δ4 (465-597) and AGS3-Δ34 (465-548) are shown in red. *B*. Schematic diagram showing the domain architecture of GST fusion AGS3-C, AGS3-Δ4 and AGS3-Δ34 constructs.

Two different deletion mutants were also created: AGS3-Δ34 encompasses the first two (N-terminal) GoLoco motifs of AGS3-C (residues 465-548); and AGS3-Δ4 includes the first three GoLoco motifs (residues 465-597) (Figure 2.4). Both mutants were created by inserting stop codons at the desired C terminus of the open reading frame of AGS3-C using the QuikChange™ site-directed mutagenesis kit (Stratagene).

The following primers were used for site-directed mutagenesis:

AGS3-Δ34:

Forward:

5' GAG GGC TAG CGT AGG CAG CCT GTA GGG GCT TCG CAT CAC CCT CAA C-

3'

Reverse:

5'- GTT GAG GGT GAT GCG AAG CCC CTA CAG GCT GCC TAC GCT AGC CCT C-

3'

AGS3-Δ4:

Forward:

5'- GCT GTC CAC CCC CTG ATG TGC TGT GAC GAG GCC CCA CCA TGC CTG

ATG-3'

Reverse:

5'- CAT CAG GCA TGG TGG GGC CTC GTC ACA GCA CAT CAG GGG GTG GAC

AGC -3'

The QuickChange^{TM} PCR cycle was performed with the following parameters:

Step 1 (1 cycle)   95 ^{o}C: 10 min

Step 2 (20 cycles)   95 ^{o}C: 1 min

                      55 ^{o}C: 2 min

                      68 ^{o}C: 6 min

Step 3 (1 cycle)   68 ^{o}C: 10 min

Step 4 (1 cycle)   6 ^{o}C: indefinite period

2.2.4   Expression and purification of GST-AGS3-C and deletion mutants

AGS3-C and the deletion mutants of AGS3-C were expressed in transformed BL21 (DE3) strains of *E. coli* cells as GST-fusion proteins. BL21 (DE3) cells were transformed by pDEST15 vectors harboring AGS3-C, AGS3-Δ4, and AGS3-Δ34. Cells were grown in 1-3 liters of LB medium at 37 °C to OD600~0.9 and induced with 200 µM IPTG for 5 h at 25 or 30°C for expression of recombinant proteins. Induced cells were harvested by centrifugation, flash frozen in liquid nitrogen, and stored at -80 °C. Frozen cells were thawed and re-suspended in lysis buffer (50 mM HEPES, pH 7.5, 100 mM NaCl, 5 mM DTT, and 1x PTT). Cell lysis was initiated by addition of 3 mg/ml hen egg white lysozyme followed by continuous stirring for 30 min at 4 °C. Lysozyme treated cell lysate was sonicated for 5 min on ice (5-s pulse and 5-s idle cycle). The cell lysate was centrifuged at 35,000 rpm for 40 min at 4 °C in a Beckman Ti45 rotor. Clear supernatant was filtered using a 0.45-µm syringe filter and loaded on glutathione-Sepharose 4B resin (Amersham Pharmacia Biosciences). The resin was washed with lysis buffer, and the GST-tagged proteins were eluted with elution buffer (20 mM Tris, pH 8.0, 100 mM NaCl, 5 mM DTT, 25 mM reduced glutathione, and 1x PTT) (Figure 2.5). The purified GST-fusion proteins were cleaved using 10 µg of recombinant TEV protease/mg of fusion protein at 4 °C overnight and dialyzed against low salt buffer (20 mM Tris, pH 8.0, 2 mM DTT, 1 mM EDTA, 1x PTT, 5% glycerol). The protein was loaded on a HiTrapQ™ column and eluted with a 75-ml linear gradient of 100-750 mM NaCl. Fractions containing AGS3-C or its deletion mutants were pooled, concentrated, and loaded on tandem Superdex™ 200 and 75 gel filtration columns (Amersham Pharmacia Biosciences)
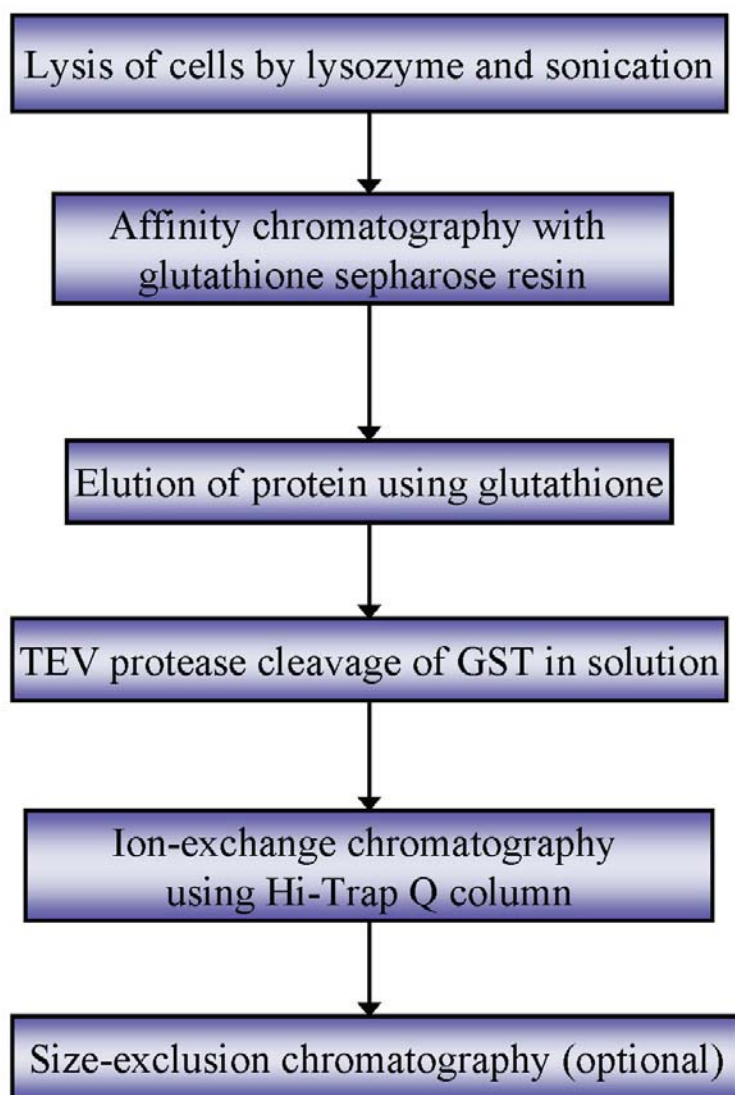
Figure 2.5. Flowchart of GST-AGS3 purification

with low salt buffer supplemented with 100 mM NaCl. The peak fractions were pooled, concentrated, and stored at 4 °C (Figure 2.5). Protein concentration was estimated by the 'Bradford' assay using the Bio-Rad protein assay kit.

2.2.5   Expression and purification of $G\alpha_{i1}$

$G\alpha_{i1}$ was purified as an untagged protein using methods described earlier (Lee et al.,

1994). *E.coli* JM109 cells harboring $G\alpha_{i1}$ encoded in a pQE60 vector (Figure 2.6) were

grown in T7 media containing 100 μg/ml ampicilin. Cells were induced at OD600~0.4 using

30 μM IPTG and 1 μg/ml of chloramphenicol at 30 ^{o}C for 12-16 hrs. Cells were harvested by

centrifugation, flash frozen in liquid nitrogen and stored at -80 ^{o}C. The purification of $G\alpha_{i1}$

from cells involved three different chromatographic steps, following the lysis of cells. Frozen

cell pellets were thawed and re-suspended in lysis buffer containing 20 mM HEPES pH 8.0,

1 mM EDTA, 3 mM DTT, and protease inhibitor cocktail (PTT or Sigma protease inhibitor

cocktail). Cell lysis was initiated by addition of 3 mg of lysozyme/gm of cells, followed by

incubation at 4 ^{o}C for 30 min with continuous stirring. DNase was added in conjunction with

$MgCl_2$ to ensure removal of DNA. To ensure the complete lysis of cells, lysozyme/DNase

treated cell lysate was also sonicated for 5 min (with alternating 5s pulse and 5s idle cycle).

The cell lysate was centrifuged either at 18,000 rpm for 45 min in a Srovall SS34 rotor or at

40,000 rpm for 30 min in a Beckman Ti45 rotor. The supernatant containing the soluble

fraction of the cell lysate was loaded on a Q-sepharose ion-exchange column (column

volume: 25 ml) pre-equilibrated with Q-buffer A containing 20 Tris pH 8.0 and 3 mM DTT

(Figure 2.7). Upon loading, the column was washed with 3-5 column volumes of Q-buffer A

and the bound proteins were eluted using a 10-20 column volume 0-100% linear gradient of

Q-buffer B containing 20 mM Tris pH 8.0, 3 mM DTT, and 500 mM NaCl. The Q-sepharose

column fractions containing $G\alpha_{i1}$ were dialyzed against HAP-buffer A (10 mM Tris pH 7.2,

>gi|6980962| *Rattus norvegicus* Gα_{i1}

```
atgggctgcacactgagcgctgaggacaaggcggccgtggagcgcagcaagatgatcgac
 M  G  C  T  L  S  A  E  D  K  A  A  V  E  R  S  K  M  I  D

cgcaacctccgggaggacggagagaaggcagcgcgcgaggtcaagctgctgctgctgggt
 R  N  L  R  E  D  G  E  K  A  A  R  E  V  K  L  L  L  L  G

gctggtgaatccgggaagagcacaattgtgaagcagatgaaaattatccacgaggctggc
 A  G  E  S  G  K  S  T  I  V  K  Q  M  K  I  I  H  E  A  G

tactcagaggaagagtgtaagcagtacaaagcagtggtctacagcaacaccatccagtcc
 Y  S  E  E  E  C  K  Q  Y  K  A  V  V  Y  S  N  T  I  Q  S

atcattgccatcattagagccatgggggagattgaaaatcgactttggagacgctgctcgt
 I  I  A  I  I  R  A  M  G  R  L  K  I  D  F  G  D  A  A  R

gcggatgatgctcgccaactcttcgtgcttgctggggctgcagaggaaggctttatgacc
 A  D  D  A  R  Q  L  F  V  L  A  G  A  A  E  E  G  F  M  T

gcggagctcgccggcgtcataaagagactgtggaaggacagcggtgtgcaagcctgcttc
 A  E  L  A  G  V  I  K  R  L  W  K  D  S  G  V  Q  A  C  F

aacagatcccgggagtaccagctgaacgattcggcggcgtactacctgaatgacttggac
 N  R  S  R  E  Y  Q  L  N  D  S  A  A  Y  Y  L  N  D  L  D

agaatagcacaaccaaattacatcccaacccagcaggatgttctcagaactagagtgaaa
 R  I  A  Q  P  N  Y  I  P  T  Q  Q  D  V  L  R  T  R  V  K

acgacgggaattgtggaaacccactttactttcaaagatcttcattttaaaatgtttgac
 T  T  G  I  V  E  T  H  F  T  F  K  D  L  H  F  K  M  F  D

gtgggaggccagagatcagagcggaagaagtggattcactgctttgaaggcgtgactgcc
 V  G  G  Q  R  S  E  R  K  K  W  I  H  C  F  E  G  V  T  A

atcatcttctgtgtggccctgagtgactatgacctggttcttgctgaggatgaagaaatg
 I  I  F  C  V  A  L  S  D  Y  D  L  V  L  A  E  D  E  E  M

aaccggatgcatgaaagcatgaagctgttcgatagcatatgtaacaacaagtggtttacg
 N  R  M  H  E  S  M  K  L  F  D  S  I  C  N  N  K  W  F  T

gacacatccatcatcctttcctgaacaagaaggacctcttcgaagagaagatcaaaaag
 D  T  S  I  I  L  F  L  N  K  K  D  L  F  E  E  K  I  K  K

agtcccctcacgatatgctatccagaatatgcaggctcaaacacatatgaagaggcggct
 S  P  L  T  I  C  Y  P  E  Y  A  G  S  N  T  Y  E  E  A  A

gcgtatatccagtgtcagtttgaagacctcaataaaaggaaggacacaaaggaaatttac
 A  Y  I  Q  C  Q  F  E  D  L  N  K  R  K  D  T  K  E  I  Y

acccacttcacttgcgccacggatacgaagaatgtgcagtttgtgttcgatgctgtaacg
 T  H  F  T  C  A  T  D  T  K  N  V  Q  F  V  F  D  A  V  T

gacgtcatcataaagaataacctaaaagactgtggtctcttctaa
 D  V  I  I  K  N  N  L  K  D  C  G  L  F  -
```

Figure 2.6. Sequence of Gα_{i1}. The protein (blue) and DNA (black) sequence of Gα_{i1} of *Rattus norvegicus* is shown along with the GI number.

3 mM DTT, and 100 mM NaCl) to prepare the sample for the subsequent hydroxyapatite column. The dialyzed sample was loaded on a hydroxyapatite column (column volume: 10 ml) pre-equilibrated with HAP-buffer A and the column was washed with two column volumes of HAP-buffer A. The bound proteins were eluted by 0-100% linear gradient with HAP-buffer B containing 10 mM Tris pH 7.2, 3 mM DTT, and 200 mM phosphate buffer pH 7.2. HAP column fractions containing $G\alpha_{i1}$ were prepared for the subsequent hydrophobic phenyl-superose chromatography by addition of ammonium sulfate to a final concentration of 1.2 M. The phenyl superose column was equilibrated with Phe-buffer A containing 50 mM Tris pH 8.0, 2 mM DTT, 1 mM EDTA, and 1.2 M ammonium sulfate; the protein sample was loaded on the column, and the column was washed with 1-2 column volumes of Phe-buffer A. The bound protein was eluted by a linear gradient of 0-100% Phe-buffer B containing 50 mM Tris pH 8.0, 2 mM DTT, and 1 mM EDTA (Figure 2.7). The elute fractions containing $G\alpha_{i1}$ were pooled together and dialyzed against the dialysis buffer containing 50 mM Tris pH 8.0, 1 mM EDTA, 2 mM DTT, and 10 μM GDP for the removal of ammonium sulfate from the protein solution. Dialyzed protein was concentrated using Centriprep™ YM 10 concentrator and the protein was buffer exchanged with the final storage buffer containing 50 mM EPPS pH 8.0, 2 mM DTT, 1 mM EDTA, and 10 μM GDP. The concentrated protein was stored at 4 ^{o}C for short-term storage or flash-frozen in liquid nitrogen and stored at -80 ^{o}C for long-term storage.
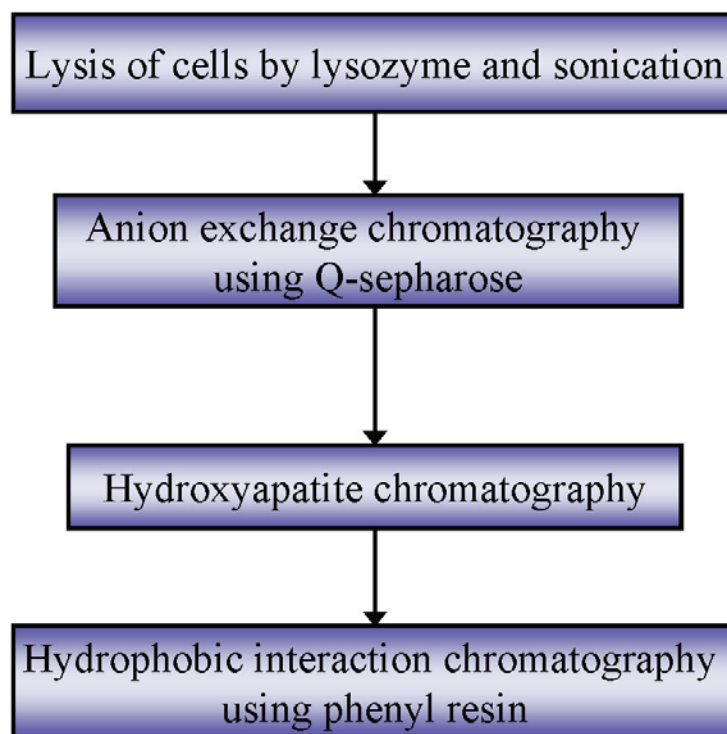
Figure 2.7. Flowchart of Gα$_{i1}$ purification

2.2.6 Modification of Gα$_{i1}$ expression system

To avoid the pitfalls associated with purification of Gα$_{i1}$, the expression and

purification method was modified by creating an N-terminal GST-tagged construct of Gα$_{i1}$.

An N-terminal GST-tagged pDEST15 vector encoding Gα$_{i1}$ residues G2-F354 was created

using Gateway™ cloning techniques as described earlier (Figure 2.3). The DNA encoding

Gα$_{i1}$ was amplified by PCR using the 5'-TEV-Gi (5'- TAT TTT CAG GAC TGC ACA CTG

AGC GCT GAG G– 3') and 3'-B2-Gi (5'- AGA AAG CTG GGT GCT TTT AGA AGA

GAC CAC AGT C – 3') primers. The amplified sequence was used as a template for the

subsequent PCR reaction with 5'-B1-TEV and 3'-B2 primers. The product of the second

PCR reaction was amplified again using the B1 and B2 primers. The subsequent steps were

similar to the methodology described earlier (chapter 2.2.2 and 2.2.3) whereby TEV protease

cleavage site was inserted between coding regions for GST and $G\alpha_{i1}$. The cleavage site of

TEV protease includes the first glycine residue of the $G\alpha_{i1}$ construct (G2), thus no extra

residue is introduced upon cleavage by TEV protease (Figure 2.8).
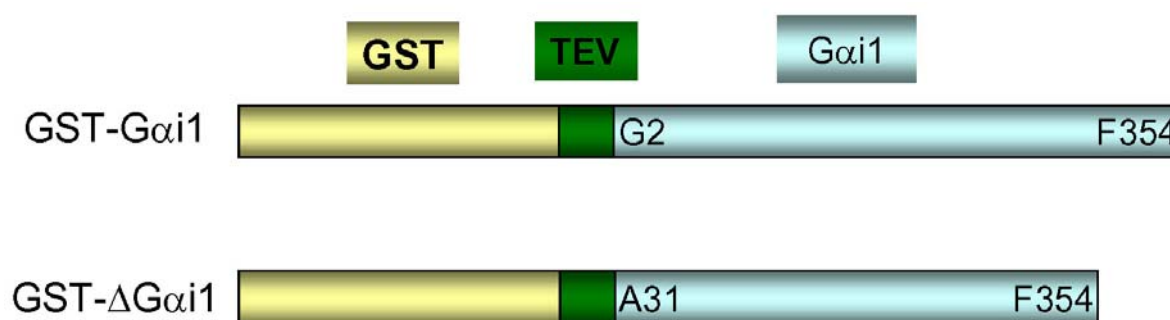
Figure 2.8. Schematic diagram of GST-fusion $G\alpha_{i1}$ and ΔN-$G\alpha_{i1}$.

An N-terminal truncation mutant encompassing $G\alpha_{i1}$ residues A31-F354 (ΔN-$G\alpha_{i1}$),

which has been shown earlier to be useful for crystallization (Kimple et al., 2002), was also

created using Gateway™. The DNA sequence encoding $G\alpha_{i1}$ was amplified by PCR using

the forward primer 5-TEV-ΔGi (5'- TAT TTT CAG GGC GCG CGC GAG GTC AAG CTG

G-3') and the reverse primer 3'-B2-ΔGi (5'- AGA AAG CTG GGT GCT TTT AGA AGA

GAC CAC AGT C-3'). The subsequent steps were carried out as described earlier. ΔN-$G\alpha_{i1}$

was cloned as an N-terminal GST-fusion protein using the pDEST15 expression vector.

2.2.7 Expression and purification of GST-Gα$_{i1}$/ΔN−Gα$_{i1}$

N-terminal GST-tagged pDEST 15 vectors encoding Gα$_{i1}$ (G2-F354) and ΔN- Gα$_{i1}$ (A31-F354) were used to transform *E.coli* BL21 (DE3) cells. Cells were grown in 3-6 liters of LB medium at 30 °C to OD600~0.4-0.5 and induced with 30 μM IPTG at 25/30 °C overnight (12-16 h) for expression of recombinant proteins (in certain preps 10 mM Tris pH 8.8 was added to the LB media at the suggestion of Dr. Celestine Thomas). Induced cells were harvested by centrifugation, flash frozen in liquid nitrogen, and stored at -80 °C. Frozen cells were thawed, re-suspended in lysis buffer (20 mM HEPES, pH 8.0, 2 mM DTT, 10% glycerol, and Sigma protease inhibitor cocktail), and lysozyme was added at 3 mg/gm of cell. After incubation with lysozyme for 30 min at 4 °C with continuous stirring, DNase (5mg/gm of cell) and MgCl$_2$ was added and the sample was incubated for another 30 min. The cell lysate was sonicated for 2-5 min on ice (5-s pulse and 5-s idle cycle) followed by centrifugation at 18,000 rpm for 45 min at 4 °C in a Sorvall SS34 rotor. Clear supernatant was loaded on glutathione-Sepharose 4B resin (Amersham Pharmacia Biosciences) (Figure 2.9). The resin was washed with three different wash buffers: wash buffer1 (20 mM HEPES pH 8.0, 500 mM NaCl, and 2 mM DTT), wash buffer 2 (20 mM HEPES pH 8.0, 300 mM NaCl, 2 mM DTT, 11 mM MgCl2, and 2 mM ATP) and wash buffer 3 (50 mM Tris pH 8.0 and 2 mM DTT). The resin bound to the GST-fusion protein was re-suspended in wash buffer 3 and incubated with TEV protease overnight at 4°C for cleavage of the fusion proteins on GST-resin (Figure 2.9). Upon cleavage of the GST-fusion tag on the resin, protein was isolated from the resin by passing the TEV incubated resin slurry over a gravity based column and collecting the flow-through. In certain preps, the GST-fusion protein was
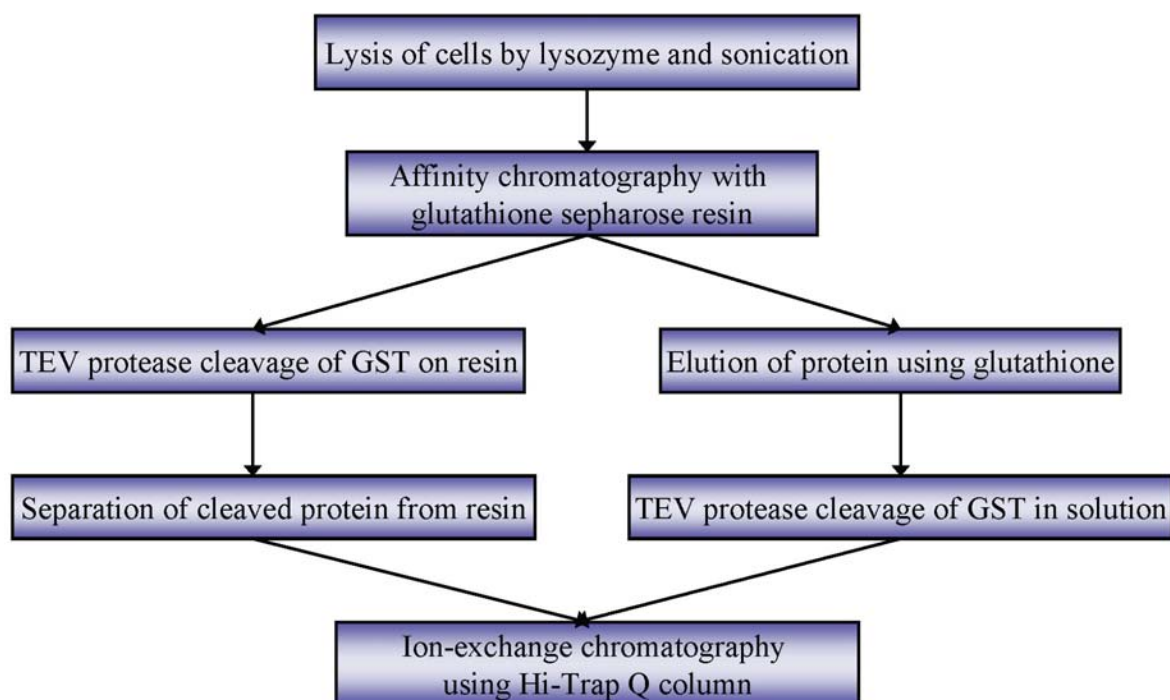
Figure 2.9. Flowchart of GST-Gα_{i1} /ΔN-Gα_{i1} purification.

eluted with a buffer containing 50 mM Tris pH 8.0, 2 mM DTT, and 20 mM reduced

glutathione and the fusion proteins were cleaved by overnight incubation with TEV protease

at 4 $^{\circ}$C (Figure 2.9). The protein sample, either cleaved on the resin or in solution, was

loaded on a HiTrapQ™ column and the column was washed with buffer A (50 mM EPPS pH

8.0 and 2 mM DTT). Bound proteins were eluted with a 100 ml linear gradient of 0-100%

buffer B (50 mM EPPS pH 8.0, 2 mM DTT, and 1 M NaCl). Fractions containing Gα_{i1} were

pooled together, concentrated using a Centriprep™ YM-10 or Amicon concentrator, buffer

exchanged with the storage buffer (50 mM EPPS pH 8.0, 5 mM DTT, 1 mM EDTA, and 10

µM GDP), and stored at 4 $^{\circ}$C. The GDP concentration (10 µM, 100 µM, 5 mM etc.) of the

storage buffer was changed in different preps to facilitate crystallization trials.

2.3 Results and Discussion

2.3.1 Purification of denatured AGS3 and refolding

6-His tagged AGS3-C (residues 465-650) cloned in pQE30 vector was used to

express AGS3-C in *E.coli* JM109 cells. Initial characterization of the construct with small

scale screening studies suggested that it can be used to express AGS3-C in JM109 cells, but

no expression was observed in BL21 (DE3) cells. The protein expressed in JM109 cells was

insoluble and upon lysis of the cells under native conditions, the protein localized in the

insoluble fraction. To purify AGS3-C in soluble form, a urea based denaturing protocol was

used, followed by refolding as described in chapter 2.2.1.

The chromatograms and gels corresponding to the purification of refolded AGS3-C

are shown in the subsequent pages. In the affinity chromatography step, His-tagged AGS3-C

binds to the Ni-NTA resin along with some contaminating proteins. The bound AGS3-C

elute over a wide range of imidazole concentrations (Figure 2.10). The yield of protein from

the Ni-NTA column was ~35-50 mg/l of culture. The protein eluted from the ion-exchange

column in two different peaks around ~200-250 mM NaCl (Figure 2.11). There was also a

large peak containing AGS3-C that eluted at very low salt concentrations, which might

represent a fraction of AGS3-C which did not bind due to saturation of the column or

adopted a conformation that differed from the bound AGS3-C in their ionic properties. It is
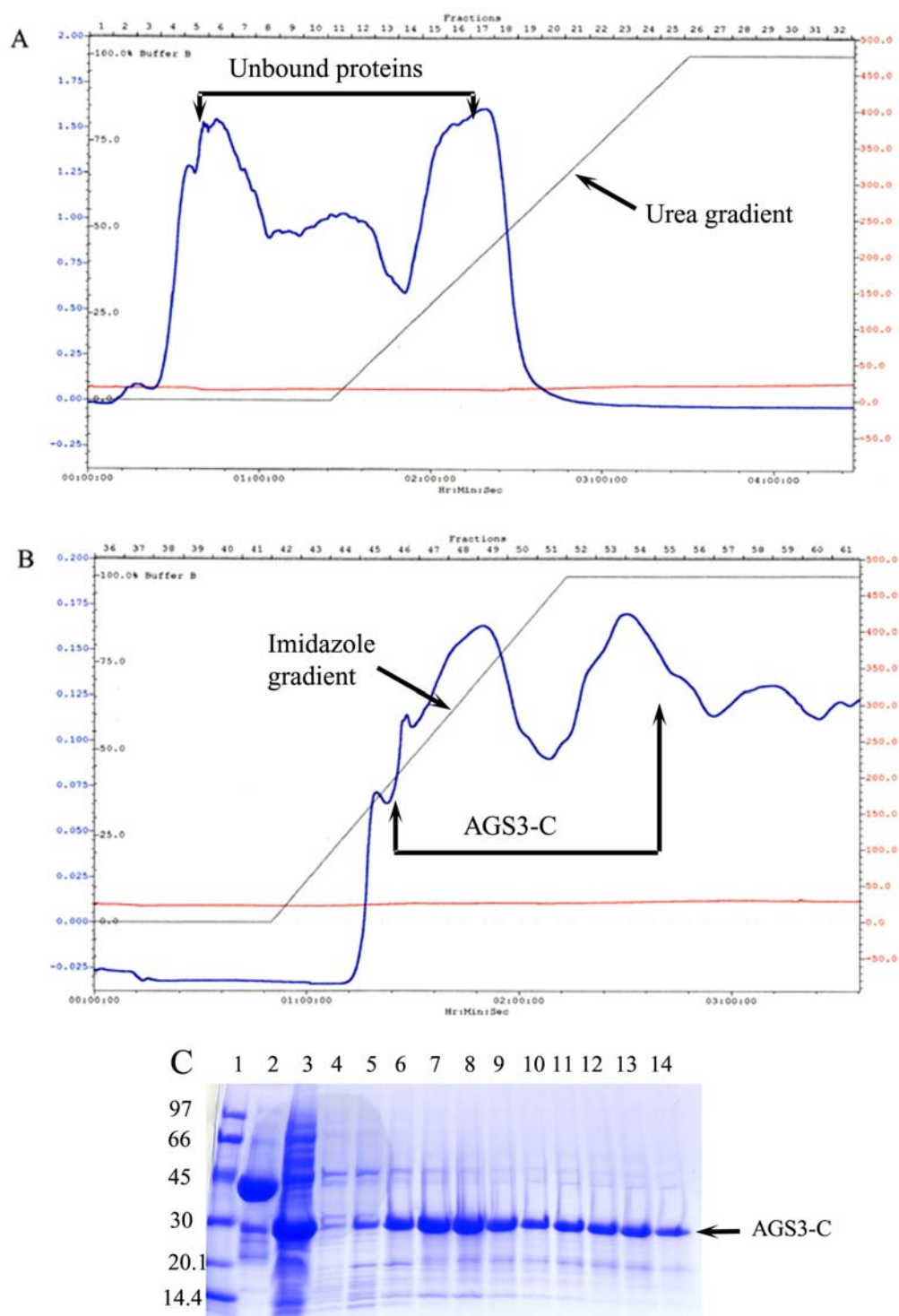
Figure 2.10. Affinity purification of denatured 6His-AGS3-C and refolding. *A*. Chromatogram showing the renaturation of proteins on Ni-NTA column. Unbound and poorly bound proteins are eluted during this time. *B*. Chromatogram showing the elution of

refolded protein with imidazole gradient. AGS3-C is eluted in multiple peaks over a broad range of imidazole concentration. *C*. Coomassie stained SDS-PAGE of nickel column elution fractions: lane 1, Molecular weight marker; lane 2, $G\alpha_{i1}$:AGS3 complex (control); lane 3, cell lysate; lanes 4-14, fraction number 45-55. The X-axis and Y-axis in A and B represent A_{280} and time, respectively.

not clear why the protein separated in multiple peaks within a short range of salt

concentration, although it might indicate the presence of partly folded molecules or

aggregated AGS3-C. Certain contaminating proteins are not removed from AGS3-C during

ion-exchange purification (Figure 2.11) but AGS3-C is more than 90% pure at this point and

the yield is ~30 mg/l of culture.

Although the purification of the refolded AGS3-C posed the problems mentioned

above, the refolded AGS3-C was capable of forming a stable complex with $G\alpha_{i1}$. It also

inhibited the binding of $GTP\gamma S^{35}$ to $G\alpha_{i1} \cdot GDP$, suggesting that it was active as GDI (data not

shown). Upon the development of the expression protocol for soluble GST-fusion AGS3-C

(discussed in chapter 2.3.2), the refolded AGS3-C was no longer used.
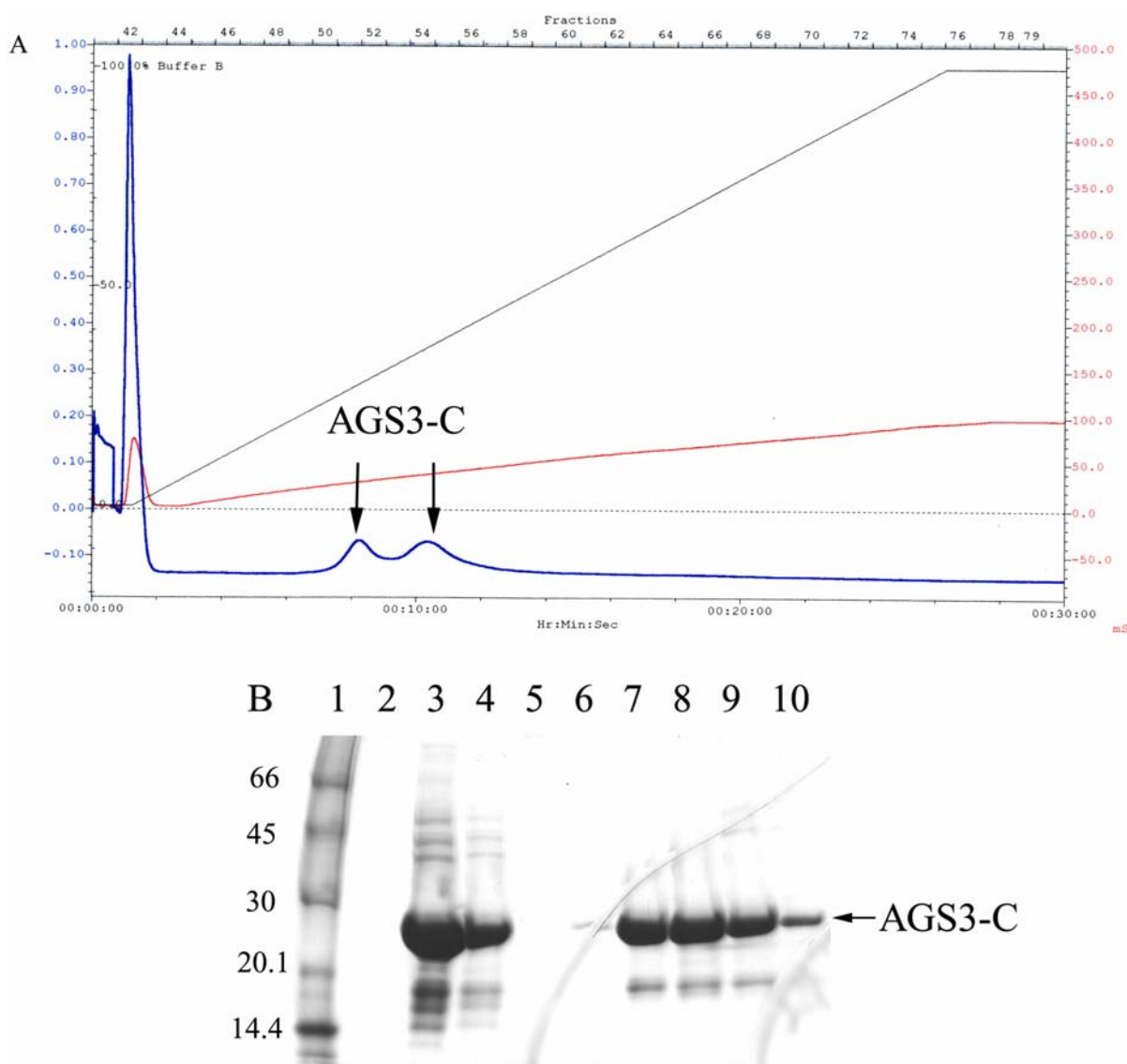
Figure 2.11. Ion-exchange purification of 6His-AGS3-C. *A*: Chromatogram showing the elution of 6His-AGS3-C from HiTrapQ column. X and Y axis represent A_{280} and time, respectively. The blue curve represents the change of A_{280} with time. *B*: Coomassie stained SDS-PAGE of HiTrapQ column fractions: lane 1, marker; lanes 2-4, fraction 41-43; lane 5, fraction 46; lane 6; fraction 49; lanes 7-10, fractions 51, 52, 54 and 56.

2.3.2   Purification of soluble AGS3-C and AGS3-C deletion mutants

Although the refolded 6-His tagged AGS3-C was functionally active, the process of refolding raises the possibility of mis-folding or partial folding. In addition to folding related problems, the incomplete removal of urea might pose a challenge for subsequent crystallographic studies. To avoid the potential problems associated with the refolding of the protein, purification of AGS3-C in soluble from was undertaken using an N-terminal GST-fusion construct of AGS3-C. The GST construct also allowed the removal of the GST-tag through a TEV protease cleavage site, unlike the 6His-tagged AGS3-C construct, which did not allow the removal of the affinity tag. GST-AGS3-C is expressed as a soluble protein in very high quantity in *E.coli* cells (Figure 2.12). The yield of the GST fusion protein was approximately 15-20 mg/l of culture. The fusion protein can be cleaved completely by TEV protease after overnight incubation (Figure 2.12), and GST can be separated from AGS3 by ion-exchange chromatography. AGS3-C elutes from HiTrapQ™ ion-exchange column at approximately 250-300 mM NaCl (Figure 2.13). The elution peak of GST and AGS3-C from the HiTrapQ™ column tends to partly overlap, and a trace amount of GST is observed in the AGS3-C peak (Figure 2.13). This trace amount of GST can be removed by passing the mixture again through GST-affinity resin and collecting the flow-through containing pure AGS3-C. The final yield of purified AGS3-C was ~5-10 mg/l of culture.

The deletion mutants of AGS3-C, AGS3-Δ4 and AGS3-Δ34, were also purified using a similar strategy (Figure 2.14). The deletion mutants were less stable than AGS3-C and more prone to degradation (Figure 2.14); however, this problem can be solved by addition of excess protease inhibitors and keeping the temperature of induction at or below 25 $^{\circ}$C. To
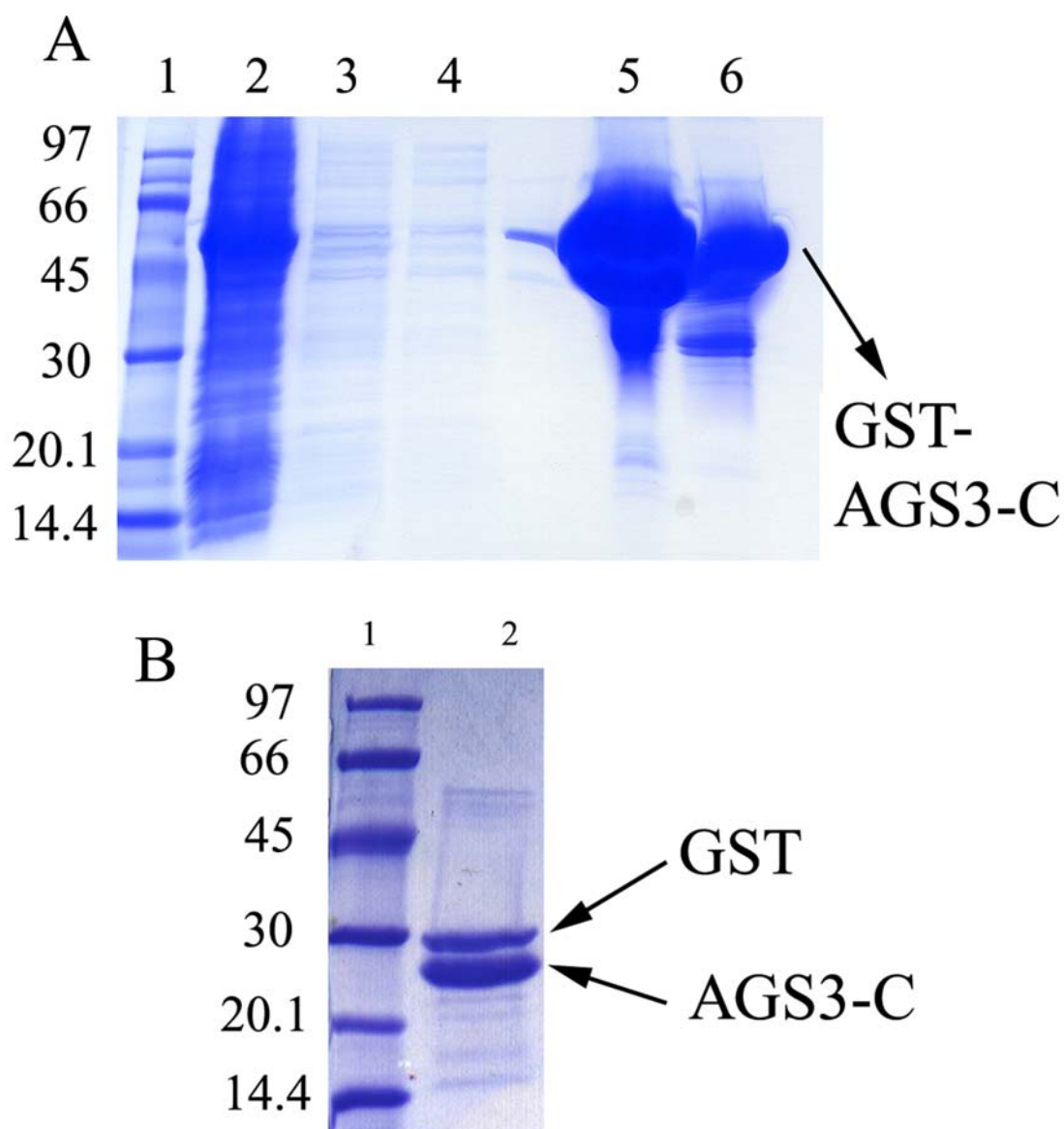
Figure 2.12. Affinity purification and TEV protease cleavage of GST-AGS3-C. *A*. Coomassie stained SDS-PAGE of the fractions collected during affinity purification of GST-AGS3-C fusion protein: lane 1, molecular weight marker; lane 2, flow-through fraction; lane 3, first wash fraction; lane 4, second wash fraction; lane 5-6, 20 and 10 µl of GST-AGS3-C elution fraction. *B*. Coomassie stained SDS-PAGE showing the TEV protease cleavage products of GST-AGS3-C: lane 1, molecular weight marker; lane 2, upper band corresponds to GST and lower band corresponds to AGS3-C.
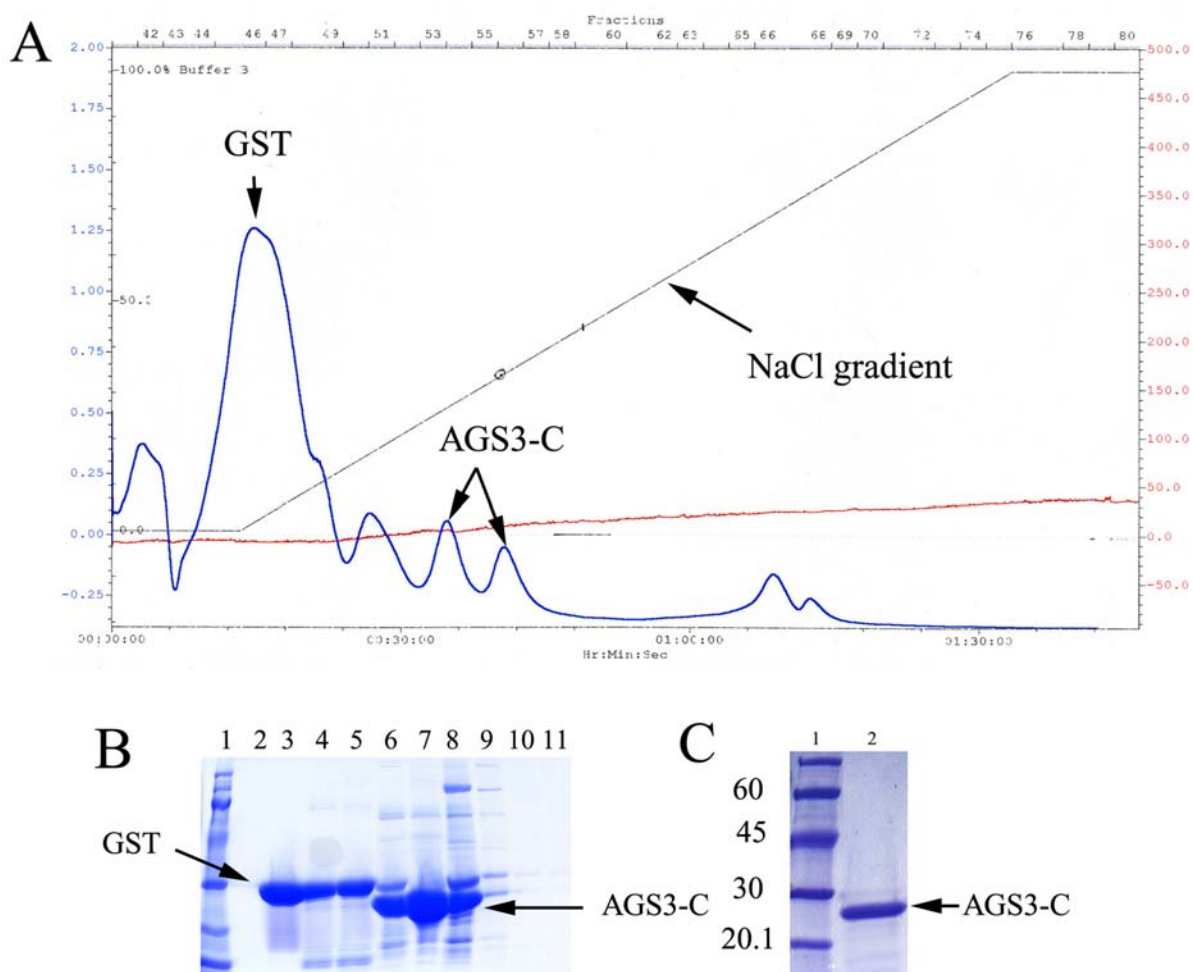
Figure 2.13. Ion-exchange purification of AGS3-C and removal of GST.
A. Chromatogram showing the purification of AGS3-C on HiTrapQ column after TEV protease cleavage of GST-AGS3-C. X and Y axis are A_{280} and time, respectively. *B*. Coomassie stained SDS-PAGE of elute fractions in A: lane 1, molecular weight marker; 2, Flow through fraction; lane 3, fraction 47; lane 4, fraction 51; lane 5, fraction 53; lane 6-7, fraction 53 and 54; lane 8, fraction 56; lane 9, fraction 58; lane 10, fraction 66 and lane 11, fraction 68. *C*. Pure AGS3-C (lane 2) after removal of free GST by affinity chromatography.
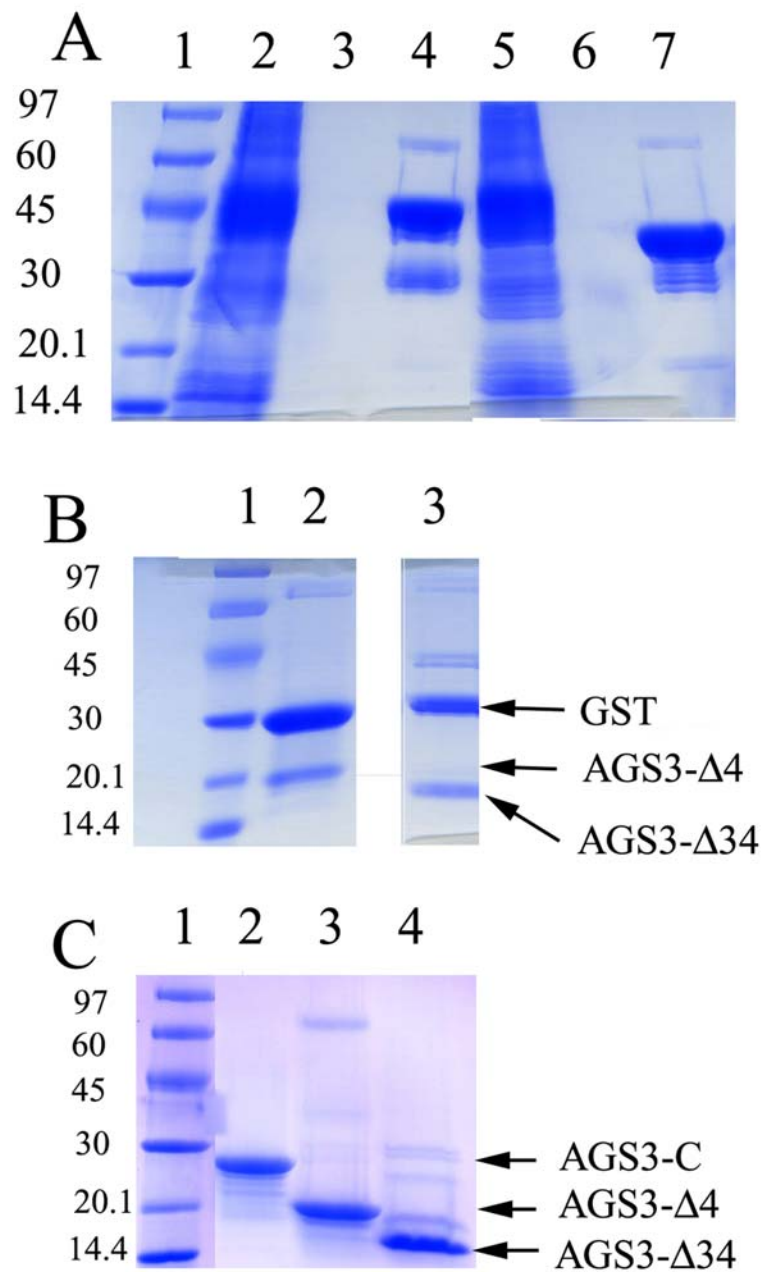
Figure 2.14. Purification of AGS3-C deletion mutants. *A*. Coomassie stained SDS-PAGE of affinity purification of AGS3-Δ4 and AGS3-Δ34: lane 1, marker; lane 2-4, flow through, wash and elution fractions of GST-AGS3-Δ4; lane 5-7, flow through, wash and elution fractions of GST-AGS3-Δ34. *B*. Coomassie stained SDS-PAGE of TEV protease cleaved GST-AGS3-Δ4 (lane 2) and GST-AGS3-Δ34 (lane 3). *C*. SDS-PAGE showing AGS3-C (lane 2), AGS3-Δ4 (lane 3), and AGS3-Δ34 (lane 4) after final purification.

remove the contaminating proteins and potential degradation products, in some preps, AGS3-C and the deletion mutants were purified further using size-exclusion chromatography (data not shown). The purity of the deletion mutants shown in Figure 2.14 is from a representative prep; considerable variations were observed among different preps.

In most of the AGS3-C or deletion mutant preps, complete removal of GST by ion-exchange chromatography was often problematic due to very close elution profiles of these proteins. However, the remaining GST can be almost completely removed by running the ion-exchange eluant over a GST affinity column. The problem associated with complete removal of GST can be solved by cleaving the GST-tag while the fusion protein is bound to the resin. Although, this strategy was not used while purifying AGS3-C and the deletion mutants, the strategy has been very effective for other GST fusion proteins (chapter 2.3.4).

2.3.3   Purification of Gα_{i1}

Wild-type Gα_{i1} without any affinity tag was purified using three chromatography steps as described in the methods section. The protocol for the purification of untagged Gα_{i1} (Lee et al., 1994) has not been modified to any substantial extent from the previously optimized method. The first step of the purification based on Q-sepharose column is highly reproducible among different preps, but the subsequent hydroxyapatite chromatography was variable among preps. The binding of Gα_{i1} to the hydroxyapatite column is highly pH sensitive. The protein fails to bind the column effectively at pH>7.2. The lowering of the pH of the buffer (pH<7.2) improves the binding of Gα_{i1} as well as that of contaminating proteins, and pH can be lowered up to 6.8 to improve binding. A substantial amount of protein is lost during the last stage of hydrophobic interaction based purification. This happens presumably due to precipitation of poorly folded Gα_{i1} while mixing with ammonium sulfate. In certain instances, Gα_{i1} aggregated and remained bound to the hydrophobic phenyl-matrix even at 100% Phe-buffer B lacking ammonium sulfate. The final yield of purified untagged Gα_{i1} after three steps of purification ranged from 10 to 20 mg/l of culture.

The following chromatograms and gels (Figures 2.15, 2.16, and 2.17) describe the purification of untagged Gα_{i1}.
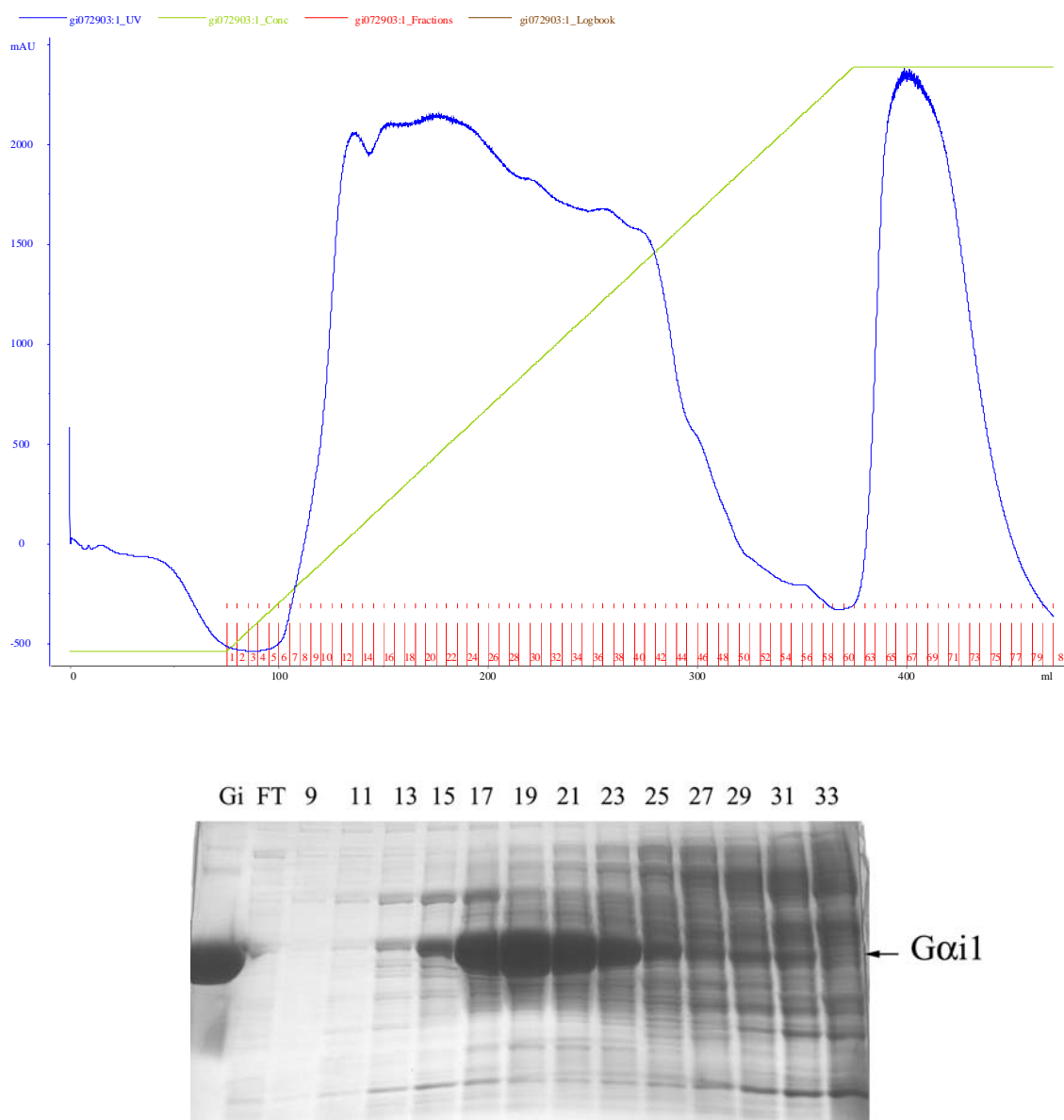
Figure 2.15. Purfication of untagged-Gα_{i1} using Q-sepharose column. *Upper panel.* Chromatogram showing the elution of proteins from Q-sepharose column. *Lower panel.* Coomassie stained SDS-PAGE of eluant fractions with fraction numbers displayed at the top of each lanes. Purified Gα_{i1} was used as a control (Gi) and flow through fraction is shown as FT.
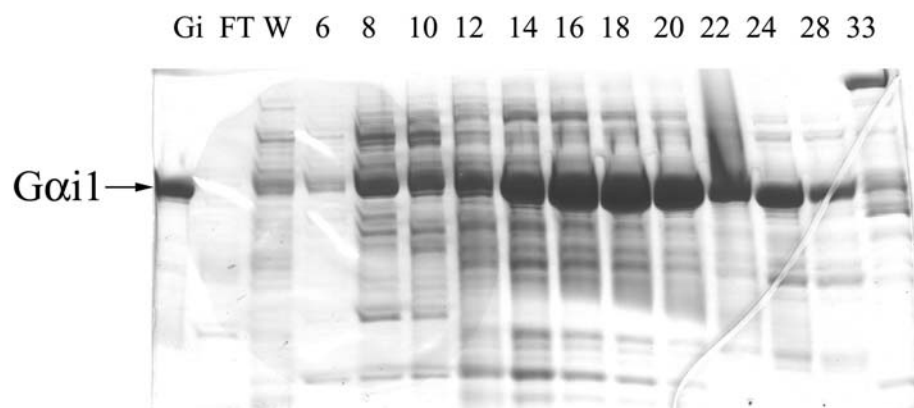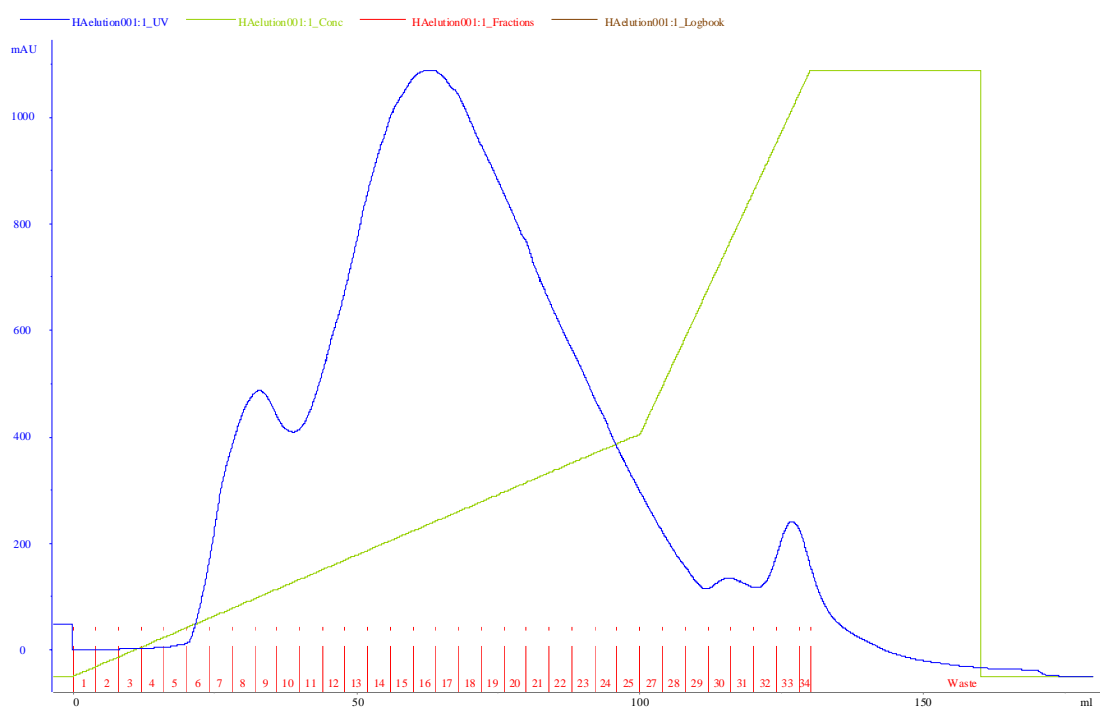
Figure 2.16. Purfication of untagged-Gα_{i1} using hydroxyapatite column. *Upper panel.* Chromatogram showing the elution of proteins from hydroxyapatite column. *Lower panel.* Coomassie stained SDS-PAGE of elute fractions with fraction numbers displayed at the top of each lanes. Purified Gα_{i1} was used as a control (Gi). Flow through and wash fractions are shown as FT and W respectively.
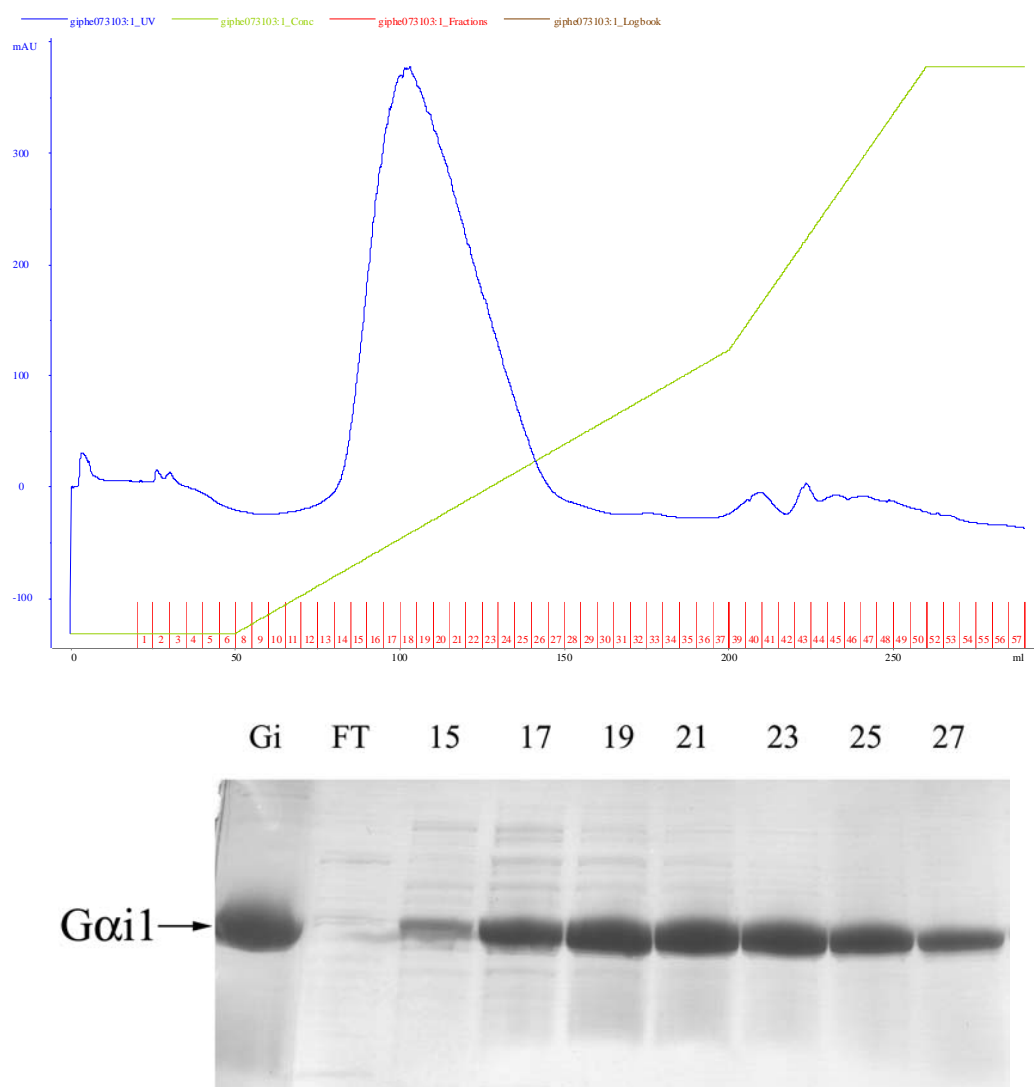
Figure 2.17. Purfication of untagged-Gα_{i1} using phenyl superose column. *Upper panel.* Chromatogram showing the elution of proteins from Source 15-phenyl column. *Lower panel.* Coomassie stained SDS-PAGE of elute fractions with fraction numbers displayed at the top of each lanes. Purified Gα_{i1} was used as a control (Gi). Flow through is shown as FT.

2.3.4   Improvement of $G\alpha_{i1}$ expression and purification

Although untagged wild-type $G\alpha_{i1}$ can be purified to very high-degree in large

amount using the method described in the last section, the purification process involves three

different steps, is time consuming, and often involved poor binding of $G\alpha_{i1}$ to the hydroxyl-

apatite and phenyl-superose column. To avoid these problems, an N-terminal GST-fusion

construct of $G\alpha_{i1}$ was created with a TEV protease cleavage site between GST and $G\alpha_{i1}$. The

cleavage of the fusion protein with TEV protease does not add any residues to the $G\alpha_{i1}$

sequence because the first residue (G2) of $G\alpha_{i1}$ is also the most C-terminal residue of the

TEV protease cleavage site. The GST-fusion form of $G\alpha_{i1}$ is expressed at very high

quantities in *E. coli* BL21 (DE3) cells. The yield of fusion protein is ~30-50 mg/l of culture

(Figure 2.18). The fusion protein is cleaved easily by overnight incubation with TEV

protease, either in solution or while bound to the resin (Figure 2.18). Cleavage of GST tag on

the resin avoids subsequent problems of removing GST from the protein mixture, and was

adopted with the suggestion of Dr. Celestine Thomas (Postodoctoral fellow in the lab). The

purity of $G\alpha_{i1}$ after one step of affinity purification and TEV cleavage was very high with

minor contamination from the TEV protease (Figure 2.18). The TEV and other minor

contaminants are removed in the following ion-exchange purification using HiTrapQ™

column. The proteins were eluted at 150-250 mM NaCl (Figure 2.19). When the GST fusion

was cleaved in solution after elution from the glutathione-resin, GST was removed from

protein mix by using a HiTrapQ™ ion-exchange column. The elution profile of $G\alpha_{i1}$ partly

overlapped with the elution profile of GST. The GST remaining in the HiTrapQ™ eluant was

removed by passing the mixture over glutathione-resin to remove free GST. In comparison
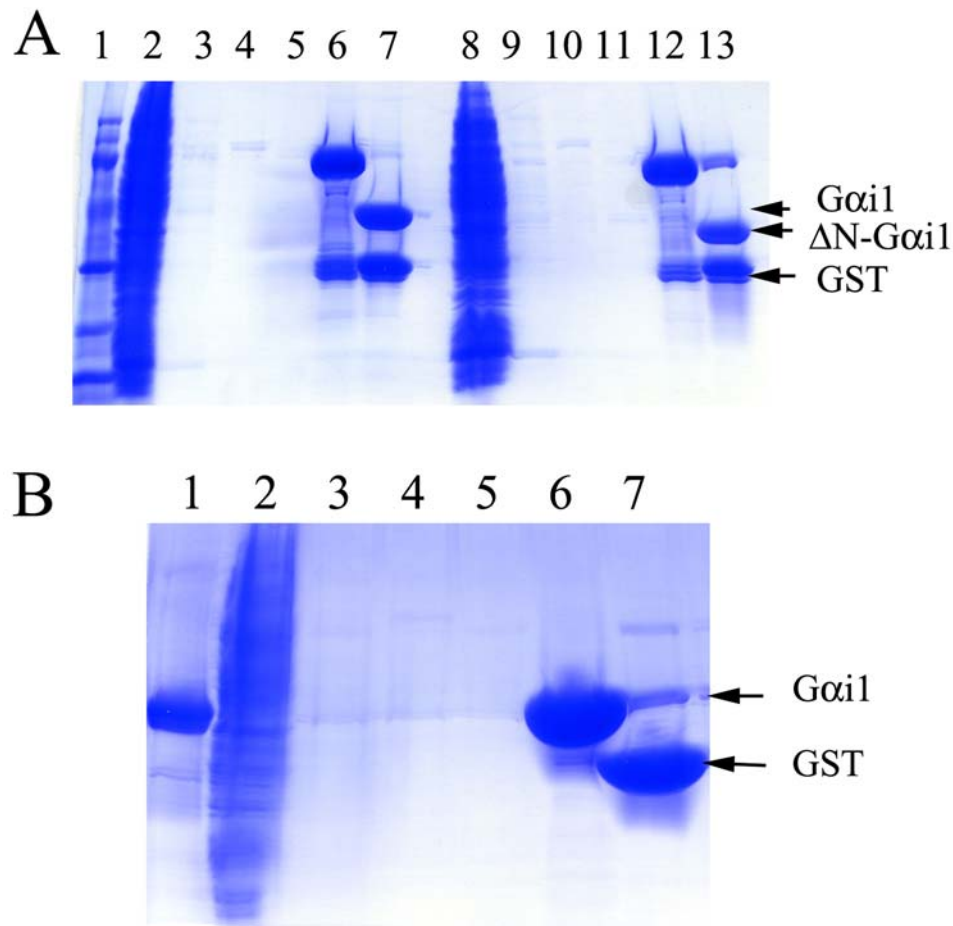
Figure 2.18. Purification of GST-Gα$_{i1}$ and TEV protease cleavage. *A*. Coomassie stained SDS-PAGE of affinity purification fractions of GST-Gα$_{i1}$ /ΔN-Gα$_{i1}$ and TEV protease cleavage: lane 2-7, flow through, first wash, second wash, third wash, elute and TEV cleavage of GST- Gα$_{i1}$ respectively; lane 8-13, flow through, first wash, second wash, third wash, elute and TEV cleavage of GST-ΔN- Gα$_{i1}$ respectively. *B*. Coomassie stained SDS-PAGE of affinity purification fractions of GST-Gα$_{i1}$ when the fusion protein was cleaved on the resin: lane 1, Gα$_{i1}$ control; lane 2-5, flow through, first wash, second wash and third wash respectively; lane 6, Gα$_{i1}$ after cleavage of the fusion protein on the resin; lane 7, elution fraction containing GST after complete cleavage of fusion protein.
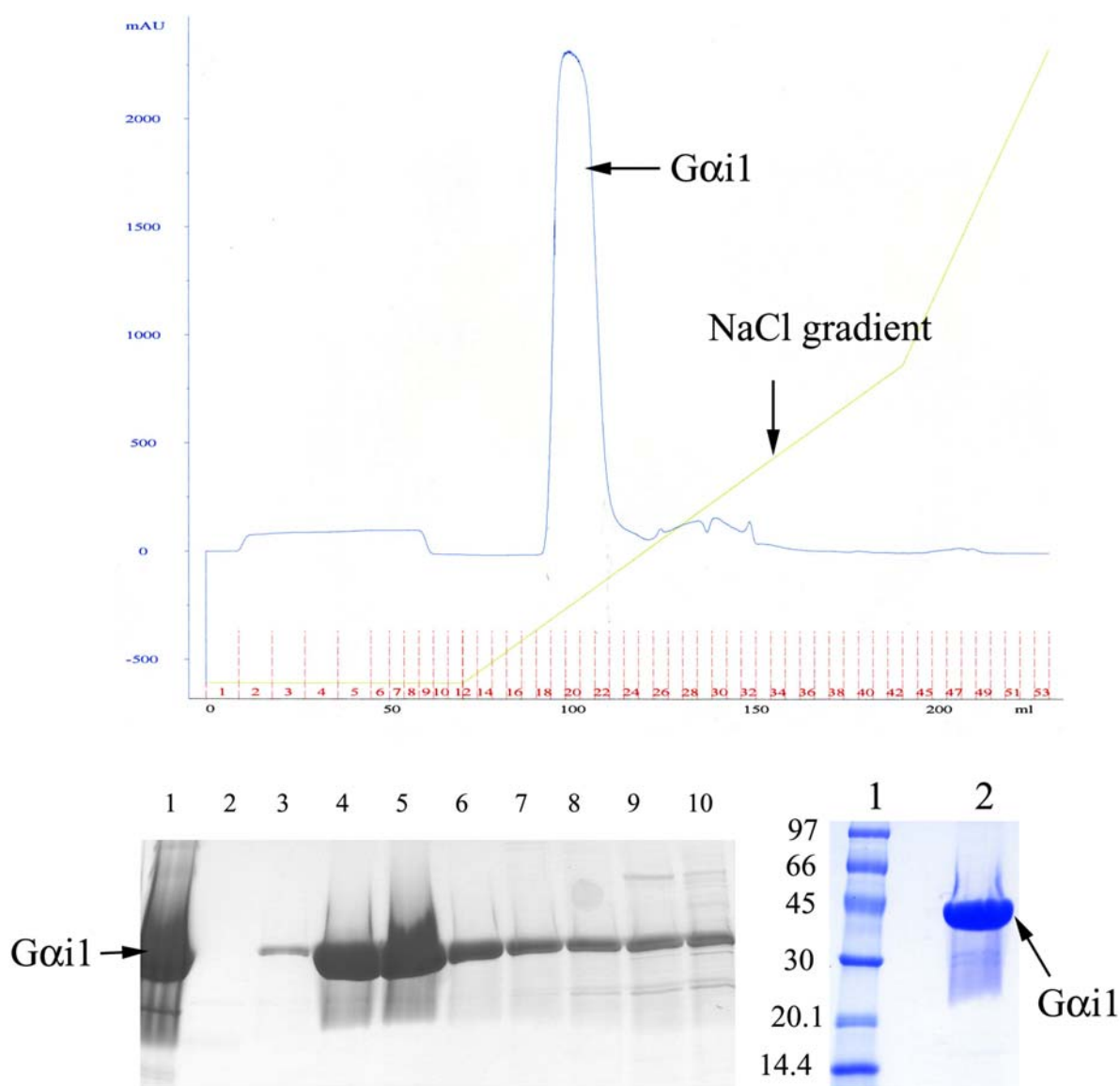
Figure 2.19. Ion-exchange purification of Gα$_{i1}$. *A*. Chromatogram showing the elution of Gα$_{i1}$ from HiTrapQ column. *B*. Coomassie stained SDS-PAGE of HiTrapQ elute fractions: lane 1, Gα$_{i1}$ control; lane 2, flow-through; lanes 3-10, fractions 18, 20, 21, 22, 24, 26, 28, and 30. *C*. SDS-PAGE showing the level of purity of Gα$_{i1}$ after the final purification step.

with the purification of the untagged version of $G\alpha_{i1}$, the new protocol reduced the time required for the purification substantially. In addition, the final yield of purified $G\alpha_{i1}$ was much higher (20-35 mg/l of culture). The $G\alpha_{i1}$ purified through this modified protocol was identical to untagged $G\alpha_{i1}$ in terms of GTPγS binding properties (data not shown), and behaved more reproducibly than the untagged $G\alpha_{i1}$ in terms of crystallization in the GDP or GTPγS bound form (data not shown) (Coleman et al., 1994).

Apart from the full-length $G\alpha_{i1}$, an N-terminal truncation mutant of $G\alpha_{i1}$ (ΔN-$G\alpha_{i1}$) consisting of residues A31-F354 was also expressed using a similar GST-fusion system (Figure 2.18). This construct favors crystallization of $G\alpha_{i1}$ in complex with the GoLoco motif of RGS14 (Kimple et al., 2002). The purification of ΔN-$G\alpha_{i1}$ was essentially identical to the purification of $G\alpha_{i1}$.

2.4 References

Coleman, D. E., Lee, E., Mixon, M. B., Linder, M. E., Berghuis, A., Gilman, A. G., and Sprang, S. R. (1994). Crystallization and preliminary crystallographic studies of $G_{i\alpha 1}$ and mutants of $G_{i\alpha 1}$ in the GTP and GDP-bound states. JMolBiol *238*, 630-634.

Hartley, J. L., Temple, G. F., and Brasch, M. A. (2000). DNA cloning using in vitro site-specific recombination. Genome Res *10*, 1788-1795.

Invitrogen. (2003). Gateway® Technology: A universal technology to clone DNA sequences for functional analysis and expression in multiple systems).

Kimple, R. J., Kimple, M. E., Betts, L., Sondek, J., and Siderovski, D. P. (2002). Structural determinants for GoLoco-induced inhibition of nucleotide release by Galpha subunits. Nature *416*, 878-881.

Lee, E., Linder, M. E., and Gilman, A. G. (1994). Expression of G-protein alpha subunits in Escherichia coli. Methods Enzymol *237*, 146-164.

Qiagen (2003). The QIAexpressionist™, In A handbook for high-level expression and purification of 6xHis-tagged proteins (Qiagen Inc.).

<u>**Chapter Three**</u>

Thermodynamic characterization of the binding of AGS3 with $G\alpha_{i1}$ and

attempts to crystallize $G\alpha_{i1}$:AGS3 complex.

3.1   Introduction

The C-terminal domain of AGS3 interacts with GDP-bound $G\alpha_{i/o}$ subunits through

the four GoLoco/GPR motifs in its C-terminal domain (Bernard et al., 2001; De Vries et al.,

2000). The binding of AGS3 leads to the stabilization of GDP bound $G\alpha$ subunits resulting

in decrease of the rate of GDP dissociation from the α subunits (Natochin et al., 2000). It was

also known that a peptide having the consensus sequence of GoLoco motifs in AGS3 can

interact specifically with the GDP-bound α subunit and act as a GDI (Peterson et al., 2000).

Although, the mechanism of action of isolated GoLoco motifs have been studied, the

biochemical properties of proteins containing tandem GoLoco repeats remained poorly

understood. The experiments described in this chapter attempts to address several questions

regarding the mechanism of AGS3 action:

- What is the arrangement of GoLoco motifs in the tertiary structure of AGS3?

- What are the biochemical consequences of having multiple copies of GoLoco motif in
  one protein?

- What is the affinity and stoichiometry of AGS3:$G\alpha_{i1}$ binding?

- Is there cooperativity in the binding of $G\alpha_{i1}$ to AGS3?

To address these questions, I used isothermal titration calorimetry (ITC) and size-exclusion/gel-filtration chromatography to determine the stoichiometry and thermodynamic parameters associated with binding of $G\alpha_{i1}$.GDP to AGS3-C and constructs of AGS3-C from which one or more GoLoco repeats had been deleted.

3.2   Materials and Methods

3.2.1   Purification of $G\alpha_{i1}$:AGS3-C/AGS3 deletion mutant complexes

Since all of the four GoLoco motifs are individually capable of binding to $G\alpha_{i1}$ (Bernard et al., 2001), AGS3-C might be able to bind four molecules of $G\alpha_{i1}$ under saturating amounts of $G\alpha_{i1}$. The complex of $G\alpha_{i1}$·GDP with AGS3-C was formed by incubating the proteins together in molar ratios ($G\alpha_{i1}$:AGS3-C) exceeding 4:1 to ensure saturation of AGS3-C by $G\alpha_{i1}$. Molar ratios exceeding 3:1 and 2:1 were used for complex formation of $G\alpha_{i1}$ with AGS3-Δ4 and AGS3-Δ34, respectively. In a typical experiment, 100 μl of 500 μM AGS3-C was mixed with 500 μl of 350 μM $G\alpha_{i1}$ in buffer containing 50mM Tris, pH 8.0, 2mM DTT, 1mM EDTA, and 1 mM GDP and incubated on ice for 2 h. The resulting complex was separated at 4 °C on tandemly connected Superdex 200 and 75 gel filtration columns (Amersham Pharmacia Biosciences) at a flow rate of 0.4 ml/min with 20 mM Tris, pH 8.0, 100 mM NaCl, 1 mM EDTA, 2 mM DTT, and 1X PTT as the running buffer. The composition of the gel-filtration peaks was analyzed by SDS-PAGE. The apparent molecular

weights of the complexes were estimated by comparing the elution volume of the complexes with the elution volumes of gel filtration standards (Bio-Rad) of known molecular weight.

3.2.2 Introduction to ITC experimental procedure, data analysis, and obtaining binding models

Isothermal titration calorimetry (ITC) is a thermodynamic technique to directly measure binding enthalpy. In a typical ITC experiment, a small amount of ligand is injected repeatedly at certain time intervals into a sample cell containing a binding partner. Depending on the nature of the interaction between the ligand and the sample, heat is either evolved (exothermic) or absorbed (endothermic) (Pierce et al., 1999). Because the sample cell in the calorimeter (Figure 3.1) is maintained at isothermal conditions (at a constant temperature), heat is either supplied or taken away from the sample cells (Jelesarov and Bosshard, 1999; Pierce et al., 1999). The temperature difference between the sample and the reference cell is measured by sensitive thermocouple circuits, which regulates the heaters attached to the cells through a feedback controlled mechanism (Pierce et al., 1999). The calorimeter accurately measures the amount heat absorbed or released at every injection. As the injection proceeds and the sample becomes saturated by the ligand, the amount of heat required to maintain isothermal condition decreases. If the ligand and the sample are kept in identical solutions, no heat is evolved due to mixing (Pierce et al., 1999). The thermal titration data thus obtained, can be used to obtain thermodynamic parameters of the binding process like enthalpy (ΔH), association constant (K_a), and stoichiometry (N). The free energy
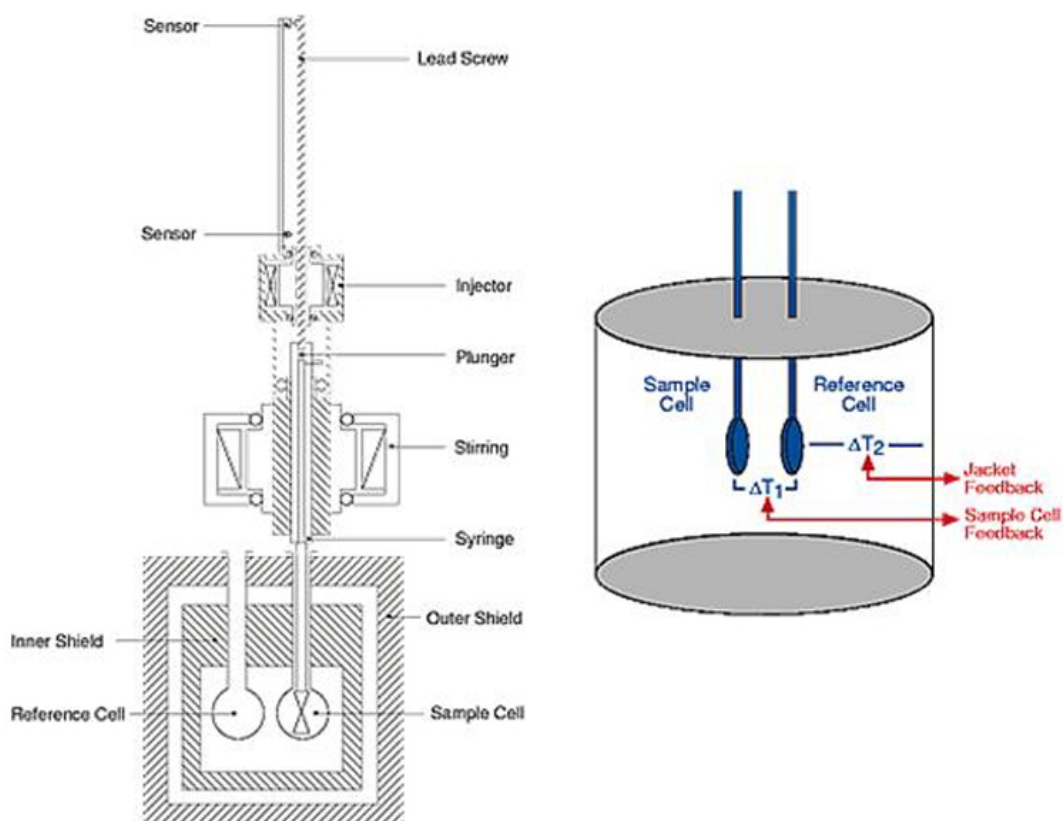
Figure 3.1. Schematic diagram of a microcalorimeter. *Left panel.* The important components of the VP-ITC equipment (Micorcal Inc.) including the sample cell, reference cell, syringe, plunger etc. are shown. The computer controlled plunger injects a small amount of sample in the sample cell through the syringe with continuous stirring (adopted from http://www.microcalorimetry.com/images/illustrations/figure3.gif). *Right panel.* A computer controlled device measures the temperature difference between the sample and reference cell ($\Delta T1$) and also the difference between the reference cell and the jacket ($\Delta T2$). $\Delta T1$ is maintained constant through the equipments feedback system (adopted from http://www.microcalorimetry.com/images/illustrations/figure2.gif).

of binding (ΔG) and entropy (ΔS) can be calculated from standard thermodynamic relationships with the following equations:

$\Delta G = \Delta H - T \cdot \Delta S$

$\Delta G = -R \cdot T \cdot \ln K_a$

ITC instruments are capable of determining binding constants over a broad range of binding association constants (10^2-10^8 M^{-1}), making ITC one of the most powerful tools available to study molecular interactions (Fisher and Singh, 1995; Pierce et al., 1999). ITC is a direct thermodynamic measurement in contrast to other biophysical techniques for measuring binding constants that rely mostly upon indirect measurement (such as change in fluorescence) or hydrodynamic parameters (such as sedimentation equilibrium or size-exclusion chromatography). ITC has been used widely to measure molecular interactions between protein and ligands, such as proteins, small peptides, DNA, small molecule ligands, metal ions, and antibodies (Jelesarov and Bosshard, 1999; Pierce et al., 1999).

Three different kinds of binding models (Microcal, 1998; Wiseman T, 1989) are employed by the Microcal-Origin ITC data analysis software that is available with the VP-ITC equipment: 'single set of identical sites', 'two sets of independent sites', and 'sequential binding sites'. The theoretical aspects of these models are reviewed extensively in several articles (Fisher and Singh, 1995; Indyk and Fisher, 1998; Wiseman et al., 1989). The discussion in this section will be primarily limited to the case of single binding sites to provide an overview of the derivation of binding parameters from titration data. The mathematical derivation of the specific binding models implemented in the Origin™ software provided by Microcal, is available in the user manual (Microcal, 1998) and also

online at Microcal website

(http://www.microcalorimetry.com/files/tech_docs2/itc_tutorial_origin_5.pdf)

1. 'Single set of identical sites': In this model a single set of identical binding sites is assumed to be present. The stoichiometry or the number of binding sites can be any number but the thermodynamic parameters describing all of the binding sites are identical. The simplest case is the binding of one ligand molecule with one protein molecule:

$$P + L \leftrightarrow PL \quad (1)$$
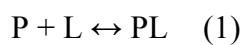
The reaction in (1) is a specific case of the more generalized case, where 'n' ligand molecules bind to a protein molecule:

$$P + nL \leftrightarrow PL_n \quad (2)$$
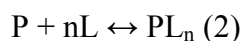
The concentration of bound ligand in (1) can be expressed as:

$$[PL] = [P]_T \frac{K[L]}{1 + K[L]} \quad (3)$$

where K is the binding constants and $[P]_T = [P] + [PL] =$ total protein concentration and $[L]_T = [L] + [PL] =$ total ligand concentration.

The fractional saturation of the protein, Θ can be defined as:

$$\Theta = \frac{[PL]}{[P]_T} \quad (4)$$

Combining equations (3) and (4):

$$K = \frac{\Theta}{(1 - \Theta)L} \quad (5)$$

The total ligand concentration can be expressed as:

$$[L]_T = [L] + \Theta[P]_T \quad (6)$$

From (5) and (6), fractional saturation Θ can be expressed as a quadratic equation in the following manner:

$$\Theta^2 - \Theta\left[1 + \frac{[L]_T}{[P]_T} + \frac{1}{K[P]_T}\right] + \frac{[L]_T}{[P]_T} = 0 \quad (7)$$

The total heat evolved due to binding can be expressed as a function of enthalpy (ΔH) and fractional saturation (Θ) as:

$$Q = \Theta[P]_T \Delta H V_0 \quad (8)$$

where, V_0 is total volume of the cell.

Because, Θ is a function of the unknown parameter K, total heat Q in equation (8) becomes a function with two unknown parameters K and ΔH:

$$Q = \frac{[P]_T \Delta H V_0}{2}\left[1 + \frac{[L]_T}{[P]_T} + \frac{1}{K[P]_T} - \sqrt{\left(1 + \frac{[L]_T}{[P]_T} + \frac{1}{K[P]_T}\right)^2 - 4\frac{[L]_T}{[P]_T}}\right] \quad (9)$$

In a general case, where 'n' identical binding sites are present, equation (9) can be modified as:

$$Q = n\frac{[P]_T \Delta H V_0}{2}\left[1 + \frac{[L]_T}{[P]_T n} + \frac{1}{K[P]_T n} - \sqrt{\left(1 + \frac{[L]_T}{[P]_T n} + \frac{1}{K[P]_T n}\right)^2 - 4\frac{[L]_T}{[P]_T n}}\right] \quad (10)$$

Equation (10) is the expression of total heat at the end of ith injection and can be referred to as Q(i). However, ITC instruments measured the amount of heat at every injection, and the important parameter is Q(i)-Q(i-1). It is important to note that the volume of the ith injection

(ΔV_i) results in a change in the total volume at the end of every injection. In that case, the expression for the heat in the ith injection can be expressed as:

$$\Delta Q(i) = Q(i) + \frac{\Delta V_i}{V_0}\left[\frac{Q(i) + Q(i-1)}{2}\right] - Q(i-1) \qquad (11)$$

The curve fitting process involves: A) estimation of the initial values of K and ΔH for the first injection using equation (10); B) calculation of the ΔQ(i) value of for every injection using the initial estimation of the parameters n, K, and ΔH and comparison of the estimated value with the experimental data; C) improvement of the initial value of n, K, and ΔH using non-linear least-square minimization methods; D) iteration of steps A through C to determine parameter values that provide best fit with the data obtained from all the injections. The deviation of the fitted isotherm from the measured binding data is used for estimating errors in the parameters.

2. 'Two sets of independent sites': In this model two sets of identical binding sites are assumed. Each of the sets can have any number of binding sites but the thermodynamic parameters of each of the binding sites within a set are identical (Figure 3.2). The two sets of sites are completely independent of each other, such that binding to any one site is independent of binding to any other sites.

'Two sets of independent sites' model is basically an extension of the 'single set of sites' model. The equations related to the 'two sets of independent sites' model are shown in Figure 3.3. The 'two sets of independent sites' model may not be adequate for systems where more than two independent binding sites are present. However, for a more complex model, each set of binding sites require optimization of 3 different parameters (n, K, and ΔH) and the

number of parameters become too high in the relation to the number of injections that are

typically carried out in ITC experiments.

<div style="border:1px solid black; padding:10px;">

<center>Two Sets of Independent Sites</center>

K_1, K_2 = binding constant of two independent sets of sites
n_1 and n_2 = number of sites in the first and second sets
M_t is concentration of macromolecule
Xt and X is concentration of total and free ligand
Θ_1 and Θ_2 is fraction of sites occupied by ligand in first and second sets
V_o is volume of cell
ΔH_1 and ΔH_2 is enthalpy

$$K_1 = \frac{\Theta_1}{(1-\Theta_1)[X]} \qquad K_2 = \frac{\Theta_2}{(1-\Theta_2)[X]} \qquad (1)$$

$$X_t = [X] + M_t(n_1\Theta_1 + n_2\Theta_2) \qquad (2)$$

From, (1) and (2)

$$X_t = [X] + \frac{n_1 M_t [X] K_1}{1+[X]K_1} + \frac{n_2 M_t [X] K_2}{1+[X]K_2} \qquad (3)$$

Equation (3) can be converted to a cubic equation of the form:

$$[X]^3 + p[X]^2 + q[X] + r = 0 \qquad (4)$$

where

$$p = \frac{1}{K_1} + \frac{1}{K_2} + (n_1 + n_2)M_t - X_t$$

$$q = \left(\frac{n_1}{K_2} + \frac{n_2}{K_1}\right)M_t - \left(\frac{1}{K_1} + \frac{1}{K_2}\right)X_t + \frac{1}{K_1 K_2}$$

$$r = \frac{-X_t}{K_1 K_2}$$

Total heat content can be expressed as:

$$Q = M_t V_o (n_1 \Theta_1 \Delta H_1 + n_2 \Theta_2 \Delta H_2) \qquad (5)$$

</div>

Figure 3.2. Equations describing 'two sets of independent sites' model. (adopted from ITC tutorial http://www.microcalorimetry.com/files/tech_docs2/itc_tutorial_origin_5.pdf)

Sequential Binding Sites

$K_1, K_2 \ldots\ldots K_n$= binding constants relative to progress of saturation of n sites
M and M_t is free and total concentration of macromolecule
Xt and X is concentration of total and free ligand
V_o is volume of cell
$\Delta H_1, \Delta H_2 \ldots \Delta H_n$ is enthalpy associated with 1st, 2nd and n^{th} level of saturation

$$K_1 = \frac{[MX]}{[M][X]}$$

$$K_2 = \frac{[MX_2]}{[MX][X]}$$

$$K_3 = \frac{[MX_3]}{[MX_2][X]} \quad \text{and so on} \qquad (1)$$

Fraction of total macromolecule having i bound ligands can be expressed as:

$$F_o = \frac{1}{P}$$

$$F_1 = \frac{K_1[X]}{P}$$

$$F_2 = \frac{K_1 K_2 [X]^2}{P}$$

$$F_n = \frac{K_1 K_2 \ldots K_n [X]^n}{P} \qquad (2)$$

where

$$P = 1 + K_1[X] + K_1 K_2 [X]^2 + \ldots + K_1 K_2 \ldots K_n [X]^n$$
and
$$X_t = [X] + M_t \sum_{i=1}^{n} i F_i \qquad (3)$$

The total heat content can be expressed as:

$$Q = M_t V_o \left(F_1 \Delta H_1 + F_2 [\Delta H_1 + \Delta H_2] + \ldots + F_n [\Delta H_1 + \Delta H_2 + \Delta H_3 + \ldots + \Delta H_n] \right) (4)$$

Figure 3.3. Equations describing 'sequential binding sites' model. (adopted from ITC tutorial http://www.microcalorimetry.com/files/tech_docs2/itc_tutorial_origin_5.pdf)

3. 'Sequential binding sites': In the 'sequential binding sites' model binding affinity of a particular site is dependent upon the fraction of occupied sites (Figure 3.3). The affinities of binding reflect the binding affinity at a particular saturation level, but not that of a binding site per se. This model can be used to model cooperativity in binding. The binding constants become progressively larger, when positive cooperativity is present, and vice versa (Fisher and Singh, 1995).

The ITC binding isotherm is fitted to any of the three models using non-linear least square curve fitting techniques. The fitting procedure is performed in such a manner, that the sum of the square of deviations of the fitted curve from the experimental data points in minimized. A general equation used for non-linear least square fit can be described as:

$$y = f(x; p_1, p_2, p_3, ...)$$

Where, y is dependent upon an independent variable x and several fitting parameters p_i. The deviation of the fitted data from the experimental measurements is represented by x^2 in the following way:

$$\chi^2 = \frac{1}{n^{eff} - p} \sum [y_i - f(x_i; p_1, p_2, ...)]^2$$

Where, n_{eff} and p are total number of data points and parameters respectively, y_i are experimental data points, and f is the fitting function. The accuracy of the curve-fitting procedure is monitored by the lowering of x^2. The value of x^2 is dependent upon the quality of the data, appropriateness of the binding model, and the degrees of freedom of the data (the difference of the number of data points and the number of fitted parameters).

3.2.3   ITC studies of Gα_{i1}-AGS3 binding

Isothermal titration calorimetry studies were carried out to characterize the binding of

Gα_{i1} with AGS3-C, AGS3-Δ4, and AGS3-Δ34. ITC experiments were performed at 20 °C

(293 K) using a MicroCal VP-ITC (MicroCal, Northhampton, MA) calorimeter. Protein

samples were dialyzed overnight against titration buffer (20 mM Tris, pH 8.0, 2 mM DTT,

and 10 μM GDP) with 2-3 changes in buffer to ensure complete equilibration. Samples were

centrifuged at 13,000 rpm in a bench-top microcentrifuge (Forma Scientific) for 5 min before

loading in the sample cell or syringe. Contents of the sample cell were stirred continuously at

300 rpm during the experiment. A typical titration of AGS3-C with Gα_{i1} involved 35-50

injections at 3-min intervals of 5-6 μl of Gα_{i1} (125-200 μM) into a sample cell containing 1.4

ml of AGS3-C (5-10 μM). Similar ITC injection parameters were also used for the titration

of AGS3-Δ4 and AGS-Δ34 with Gα_{i1}. The heats of dilution of the titrants were subtracted

from the titration data for base-line correction. The base-line corrected data were analyzed

with MicroCal Origin™ 5.0 software to determine the enthalpy (ΔH), association constant

(K$_a$), and stoichiometry of binding (N). Thermal titration data were fit to one or more of the

three association models available in the software: "single set of identical sites," "two sets of

independent sites," and "sequential binding sites". The models were compared by visual

inspection of the fitted curves and by comparing the χ^2 values obtained after the computation.

The model resulting in the lowest value of χ^2 was considered the best model to describe the

molecular mechanism of binding. Free energy change (ΔG) and entropy change (ΔS) were

calculated from ΔH and K$_a$ using standard free energy relationships. Several titrations were

performed (two to four) for each sample set to evaluate reproducibility.

3.2.4 Monitoring GDI activity by fluorescence spectroscopy

The rate of binding of GTP to Gα subunits can be measured in various ways. Typically, a non-hydrolyzable analog of GTP is used to measure the rate of GTP binding to prevent the hydrolysis of bound GTP by Gα subunit. The binding of the GTP analog is measured either by labeling the GTP analog with a radioactive probe or by using fluorescent GTP analogs that undergo change in fluorescence upon binding to Gα (McEwen et al., 2001). GTP binding can also be monitored by the change in intrinsic tryptophan fluorescence of Gα. The intrinsic tryptophan fluorescence of $Gα_{i1}$ increases upon binding of GTP and has also been used to assay the binding GTP binding (Higashijima and Ferguson, 1991; Higashijima et al., 1987). Several fluorescent analogs of GTP are available commercially. These GTP-analogs are based on chemical coupling of either MANT (*N*-methyl-3'-*O*-anthraniloyl) or BODIPY fluorophores to non-hydrolyzable GTP-analogs like GTPγS (guanosine 5'-*O*-(3-thiotriphosphate)) or GMPPNP (guanosine-5'-[(β,γ)-imido]triphosphate) (Gille A, 2003; McEwen et al., 2001; Remmers et al., 1994; Rojas et al., 2003).

Nucleotide exchange in Gα·GDP involves at least two distinct steps. In the first step, GDP is released spontaneously from the Gα·GDP complex, presumably resulting in a transient nucleotide-free state (Ferguson et al., 1986). In the second step, nucleotide-free Gα associated with GTP (Ferguson et al., 1986). The rate of GDP release is the rate limiting step in the nucleotide exchange process. Thus, the rate of nucleotide exchange would be a reasonable approximation of the rate of GDP release (Ferguson et al., 1986). Guanine nucleotide dissociation inhibitors (GDI) inhibit the rate of GDP release by stabilizing the

GDP bound conformation of Gα, and consequently reduce the rate of nucleotide exchange. GDI activity has been measured by monitoring the rate of release of GDP from the Gα·GDP complex (Natochin et al., 2000) and also by monitoring the rate of radioactive or fluorescent GTP-analog binding (Kimple et al., 2002b; Peterson et al., 2000).

The rate of guanine nucleotide exchange in Gα_{i1} was assayed by monitoring the rate at which the fluorescent GTP analog BODIPY-GTPγS (Molecular Probes) replaces bound GDP in the catalytic site of Gα_{i1} (McEwen et al., 2001). Binding of the fluorescent nucleotide analog is accompanied by an increase in fluorescence of the BODIPY moiety (Figure 3.4). Gα_{i1} (final concentration, 200 nM) in assay buffer (50 mM Tris, pH 8.0, 1 mM EDTA, and 10 mM MgCl$_2$) was incubated in the presence or absence of AGS3-C, AGS3-Δ34, and AGS3- Δ4 for 1-2 h at 25 °C. Samples were transferred into a 3-ml quartz cuvette with a path length of 1 cm containing 1 μM BOPDIPY-GTPγS in assay buffer, and nucleotide exchange was initiated by rapid mixing. Binding of fluorescent nucleotide to Gα_{i1} was monitored by the intensity of fluorescence emission at 510 nm using a PerkinElmer Life Sciences LS50B spectrophotometer. The excitation wavelength was set to 485 nm with the slit widths for excitation and emission maintained at 2.5 nm. All assays were performed at 30 °C. Typically fluorescence intensity was recorded at 30-s intervals with 10-s averaging time over a period of 1 h after mixing of the samples. Two to three data sets were measured for each experiment and the base-line fluorescence (intensity at time $t = 0$) was subtracted from the data sets. The data were averaged and smoothed using five-point adjacent averaging. The data were fit to a first order exponential association model, $Y = Y_s(1 - e^{-kt})$, where k is the rate constant (s^{-1}) and

Y and Y_s represent concentrations of BODIPY-GTPγS-bound Gα$_{i1}$ at time t and at maximum

saturation, respectively. Initial rates were estimated by linear approximation to the change in

fluorescence intensity during the first 10 min after initiation of the exchange reaction.
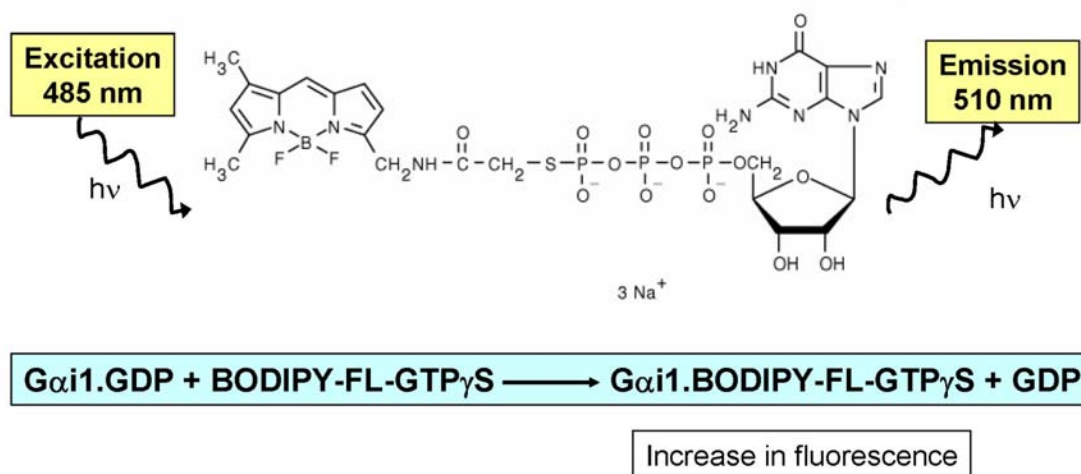
Figure 3.4. BODIPY-GTPγS based nucleotide exchange assay. *Upper.* The structure of
BODIPY and fluorescence excitation and emission parameters are shown. *Lower.* Binding of
BOIDIPY-FL-GTPγS to Gα$_{i1}$.GDP leads to the formation of Gα$_{i1}$.BODIPY-FL-GTPγS,
which has higher fluorescence intensity compared to free BOIDIPY-FL-GTPγS.

3.2.5   Crystallization screening of Gα$_{i1}$:AGS3 complex

The purified complex of Gα$_{i1}$:AGS3-C was used for screening crystallization

conditions. Crystallization trials were set up with 10-20 mg/ml of complex at 20 $^{\circ}$C and 4 $^{\circ}$C.

Several crystallization screens were performed, including standard sparse-matrix screens

(Jancarik and Kim, 1991) obtained from Hampton Research and Emerald Biosciences. In

addition to standard screens, several grid screens were designed based on previously known

conditions or using recipes developed previously in the laboratory. Randomized screens were

Screen name	Manufacturer	Method	Number of conditions
Crystal Screen™	Hampton Research Inc.	Sparse matrix screen	50
Crystal Screen2™	Hampton Research Inc.	Sparse matrix screen	48
Crystal Screen Lite ™	Hampton Research Inc.	Sparse matrix screen	50
PEG/Ion Screen™	Hampton Research Inc.	PEG/ion grid screen	48
Grid Screen Ammonium Sulfate™	Hampton Research Inc.	Grid screen	24
Grid Screen MPD™	Hampton Research Inc.	Grid screen	24
Grid Screen PEG/LiCl™	Hampton Research Inc.	Grid screen	24
Grid Screen PEG 6000™	Hampton Research Inc.	Grid screen	24
Wizard™ I	Emerald Biostructures Inc.	Sparse matrix screen	48
Wizard™ II	Emerald Biostructures Inc.	Sparse matrix screen	49
JBScreen Classic	Jena Biosciences	Based on Biological Macromolecules Crystallization Database	240
Highthroughout screen	Haumptman Woodward Institue	Microbatch screen under oil	1536

Table 3.1. Table of standard commercially available crystallization screens. The name of the screen, name of the commercial supplier, nature of the screen, and the number of unique conditions in each of the screens are shown.

also designed using a program called Crystool (Segelke, 2001). An extensive micro-batch

crystallization screening under oil was also performed in the high-throughput crystallization

facility at Hauptman Woodward Institute, Buffalo, NY.

(http://www.hwi.buffalo.edu/ProductsServices/highthroughput/HighThroug.htm)

The discussion of the exhaustive list of crystallization screens is beyond the scope of this

dissertation, examples of the important and unique screens are listed in Table 3.1 and 3.2.

Screen name	Strategy
Crystool-PEG based	Crystool program
Crystool-salt based	Crystool program
Ammonium Sulfite grid	Based on $G\alpha_{i1}$·GDP Crystallization[1]
Ammonium Sulfate grid	Grid screen[2]
PEG 1.5K, 3K, 6K, 8K	Grid screen[2]
Cesium Sulfate grid	Based on $G\alpha_{i1}$:GoLoco Crystallization[3]

Table 3.2. Examples of personally designed crystallization screens. Randomized salt and
PEG based crystallization screens were designed using a program Crystool, available from
http://porter.llnl.gov/crystool4.1/. Other grid screens and the rationale of performing those
screens are indicated. [1] (Coleman et al., 1994); [2] (McPherson, 1999); [3] (Kimple et al., 2002a)

In addition to $G\alpha_{i1}$:AGS3-C complex, the complexes of $G\alpha_{i1}$ with AGS3-Δ4 and

AGS3-Δ34 were also used for crystallization screening. Similar crystallization screenings

were also performed using complexes in which ΔN-$G\alpha_{i1}$ was used for complex formation

instead of the full-length $G\alpha_{i1}$.

3.2.6   Circular dichroism spectroscopic studies of AGS3 secondary structure

The secondary structure of AGS3-C was analyzed using CD spectroscopy. CD

spectra was also collected for the 'GPR consensus' peptide, which has the consensus

sequence of GoLoco motifs in AGS3 (Peterson et al., 2000). CD spectra were measured with a Jasco Model J715 spectropolarimeter (Jasco Inc., Easton, MD) using a 0.05-cm path length cylindrical cell. AGS3-C was dialyzed overnight in 10 mM Tris, pH 8.0, and lyophilized GPR consensus peptide was dissolved directly in the buffer. Spectra were recorded of 8 and 16 μM AGS3-C and of 100 μM GPR consensus peptide. CD spectra were measured at 25 °C with a 1-nm spectral bandwidth, scan speed of 50 nm/min, and a response time of 1 s. Data were collected at 0.1-nm intervals, and 15 accumulations were averaged to obtain each spectrum. The spectra of AGS3-C at the two different concentrations were averaged to obtain the final spectrum. CD data were smoothed by the Savitzky-Golay method using the program provided by Jasco and ε_L-ε_R was calculated (in units of M^{-1} cm^{-1}/residue) at 1-nm intervals. The CD spectra over the range of 250-190 nm were analyzed for fractional content of secondary structures using CDPRO software CONTINLL (Provencher and Glöckner, 1981), SELCON3 (Sreerama et al., 1999), and CDSSTR (Johnson, 1999) with a reference set containing 43 proteins.

3.3   Results and Discussion

3.3.1   Determination of Gα_{i1}:AGS3 binding stoichiometry using gel-filtration chromatography

AGS3-C (residues 465-650) contains four tandemly repeated GPR motifs, each of which has been shown in immunoprecipitation assays to be capable of binding to Gα_{i1} (Bernard et al., 2001; Natochin et al., 2001), suggesting that a single AGS3-C domain can

bind up to four molecules of the α subunit. Gel filtration chromatography was used to estimate the stoichiometry of the interaction between AGS3-C and $G\alpha_{i1}$. Upon incubation together in the presence of GDP, $G\alpha_{i1}$ and AGS3-C formed a stable complex. The complex could be purified by gel filtration chromatography (Figure 3.5) and eluted with an apparent molecular mass (~207 kDa) consistent with four $G\alpha_{i1}$ subunits bound to one molecule of AGS3-C (molecular mass ~180 kDa). Hence it appears that the each of the GPR motifs binds to $G\alpha_{i1}$.

To further test this hypothesis, C-terminal deletion mutants of AGS3-C were created that contain the first two (AGS3-Δ34) or the first three (AGS3-Δ4) GPR motifs. Both deletion mutants formed complexes with $G\alpha_{i1}$ in the presence of GDP and could be separated from their constituents by gel filtration chromatography. An estimate of the molecular masses of the complexes by gel filtration using standards of known molecular mass was consistent with a stoichiometry of 1:3 for AGS3-Δ4:$G\alpha_{i1}$ (~151-kDa complex) and 1:2 for AGS3-Δ34:$G\alpha_{i1}$ (~110 kDa). The estimated molecular masses for both of these complexes were ~20 kDa higher than the masses expected for the predicted complexes (Figure 3.6).

It was shown by gel-filtration chromatography that the stoichiometry of binding of AGS3-C and the deletion mutants to $G\alpha_{i1}$ was directly related to the number of GoLoco motifs in the construct. It is important to note however, that determination of molecular weight from gel-filtration elution volume is imprecise. The elution volume of a protein or protein complex in a gel-filtration column is determined by the hydrodynamic radius of the molecule rather than the absolute molecular weight. Determination of accurate molecular weight using gel-filtration standards is dependent on the globular nature of the molecule.
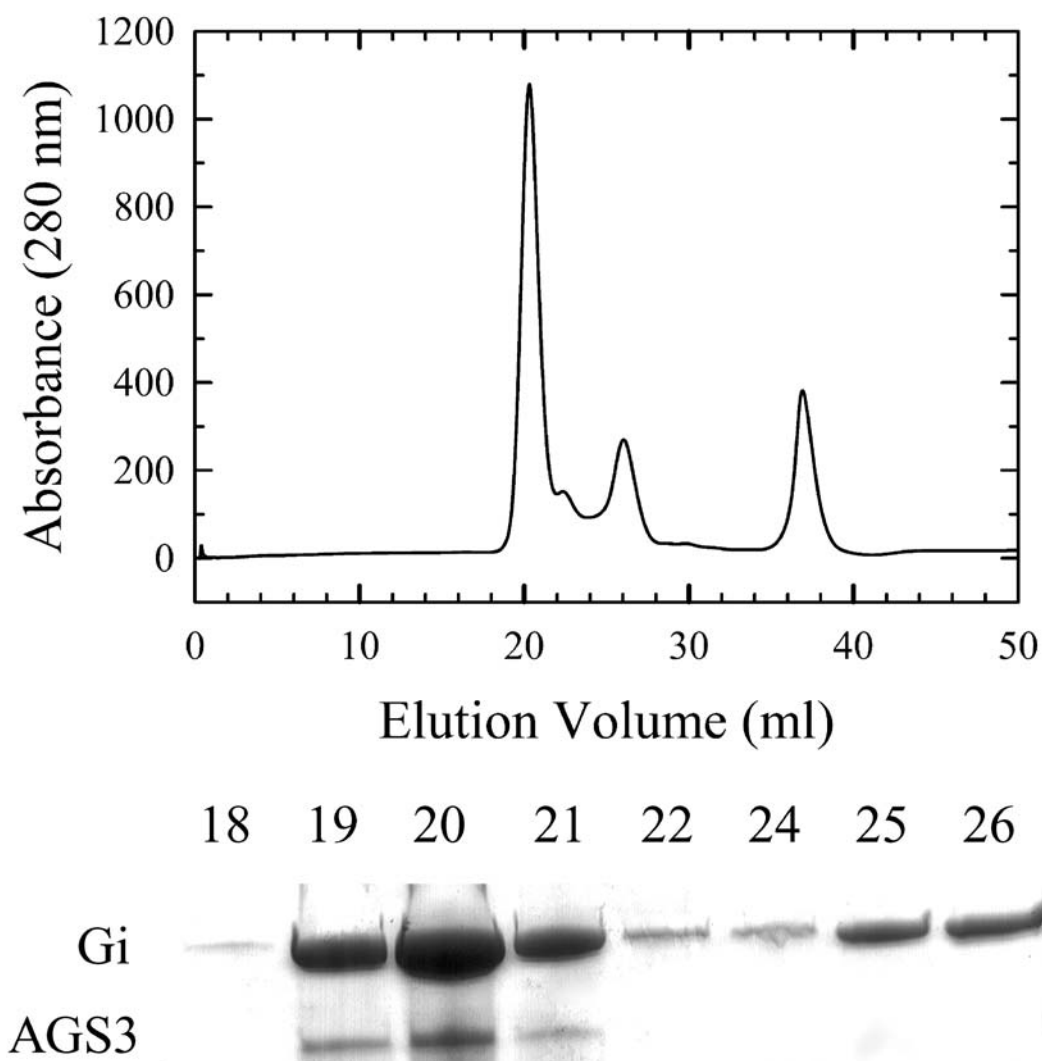
Figure 3.5. Purification of Gα_{i1}:AGS3 complex using gel-filtration chromatography.
upper panel, Gα_{i1} and AGS3 were mixed at a molar ratio>4:1 (see chapter 3.2.1), separated on a tandemly connected Superdex 200/Superdex 75 gel filtration column and 1 ml fractions were collected. The profile of absorbance at 280 nm shows three distinct peaks, from left to right: the AGS3:Gα_{i1}·GDP complex, unbound Gα_{i1}·GDP and free GDP. The *lower panel* shows the SDS-PAGE separation of different fractions after gel filtration. The elution volumes of the fractions are indicated at the top of each lane. The 40 kDa upper band corresponds to Gα_{i1} and the 20 kDa lower band correspond to AGS3-C. The molecular masses of the proteins in each band were verified by mass spectroscopy. Fractions 19-21 contain the complex and fractions 24-26 contain free Gα_{i1}.

Molecules that are elongated in shape will migrate with higher apparent molecular weight in a gel-filtration column, compared to more globular (spherical) molecules of the same molecular weight. In spite of these limitations, the apparent molecular weight of the AGS3-C:$G\alpha_{i1}$, AGS3-Δ4:$G\alpha_{i1}$, and AGS3-Δ34:$G\alpha_{i1}$ complexes strongly suggested the presence of four, three and two binding sites for $G\alpha_{i1}$ in AGS3-C, AGS3-Δ4, and AGS3-Δ34 respectively.

	AGS3-C	AGS3-Δ4	AGS3-Δ34
Estimated (kDa)	207	151	110
Predicted (kDa)	180	134	90
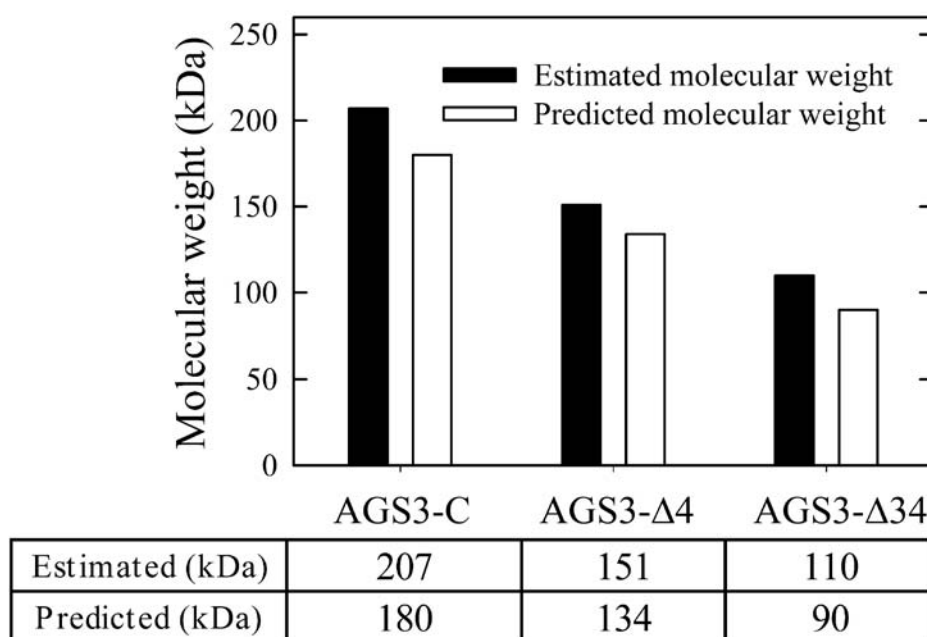
Figure 3.6. Estimation of the stoichiometry of binding of AGS3-C, AGS3-CΔ4 and AGS3-CΔ34 with $G\alpha_{i1}$ using gel-filtration chromatography. The molecular weights of the AGS3-C:$G\alpha_{i1}$, AGS3-Δ4:$G\alpha_{i1}$, and AGS3-Δ34:$G\alpha_{i1}$ complexes were estimated by comparing their elution volume with the elution volume of standards of known molecular weight. Estimated molecular weight refers to the molecular weight estimated from gel filtration results. Predicted molecular weight refers to the molecular weight of the complexes assuming AGS3-C binds to 4, AGS3-Δ4 binds to 3 and AGS3-Δ34 binds to 2 molecules of $G\alpha_{i1}$ respectively. The molecular weights of $G\alpha_{i1}$, AGS3-C, AGS3-Δ4, and AGS3-Δ34 are 40.2 kDa, 20.4 kDa, 14.6 kDa, and 9.1 kDa respectively.

3.3.2 Characterization of Gα_{i1}:AGS3 binding using ITC: "Two sets of independent

sites"

Since the migration rate of solutes through gel filtration media is dependent upon the

hydrodynamic radius rather than mass, we used isothermal titration calorimetry (ITC) to

precisely determine the stoichiometry and affinity of the interaction between AGS3-C and

Gα_{i1}. The dependence of heat evolved upon titration of Gα_{i1}·GDP into a solution containing

AGS3-C (Figure 3.7) yielded an overall stoichiometry of 4.3 (Gα_{i1}:AGS3-C), consistent with

a 4:1 stoichiometry of binding. Because the amino acid sequences of the four GPR repeats

differ from each other, the most general model for binding of Gα_{i1} to AGS3-C would require

four association constants. If some of the affinity constants are of similar magnitude, their

individual values cannot be accurately determined from the titration data. The determination

of four independent sites using ITC would require optimization of 12 independent

parameters, which would result in inaccuracies due to the limitation of non-linear least square

curve fitting algorithms (Fisher and Singh, 1995). On the other hand, binding of Gα_{i1} to

AGS3-C was not well approximated by a model in which four identical independent binding

sites are assumed (Table 3.3). A sequential four-site binding model, in which the dissociation

constant is a function of the number of binding sites occupied, provided a better fit to the data

(Table 3.3), but the values of enthalpies and dissociation constants were aberrant and showed

wide fluctuations/errors during the curve-fitting procedure, which is indicative of problems

with convergence of the least-square fitting (Table 3.3). Thus, the fit obtained with the

sequential binding model, which requires the determination of 8 independent parameters for

four binding sites, could be incorrect and misleading. However, a simpler model, in which
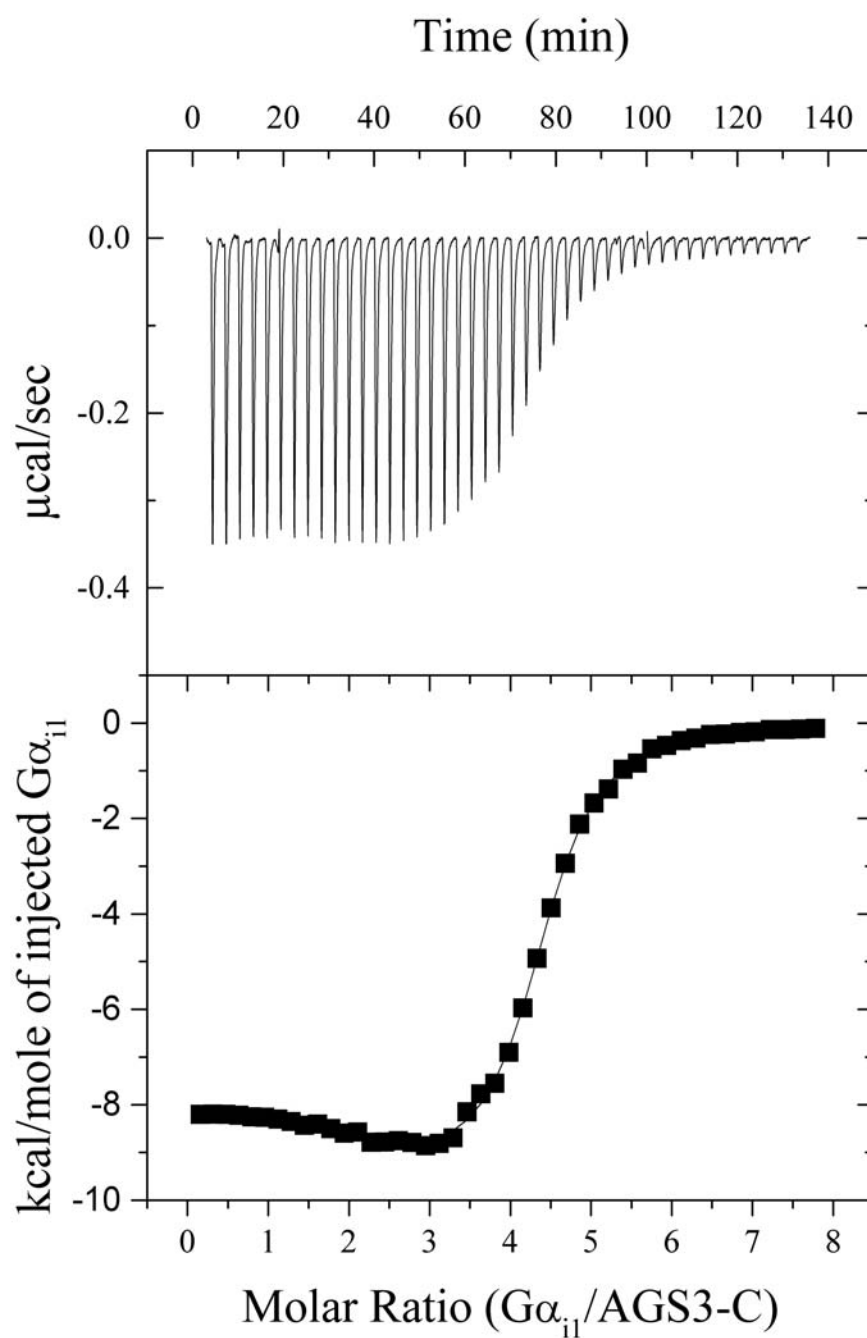
Figure 3.7. ITC profile for the binding of Gα$_{i1}$ to AGS3-C. Aliquots (6 μl) of 0.2 mM Gα$_{i1}$ were injected into an ITC cell containing 1.4 ml of 5.4 μM AGS3-C. The heats associated with the injections appear as peaks in the upper panel. Nonlinear least square fit using "two sets of independent sites" model resulted in the fit shown in the lower panel.

Single set of identical sites	Two sets of independent sites	Sequential binding sites
χ^2=38255	χ^2=4661	χ^2=9487
N= 3.9 ± 0.02 K_d= 128 ± 24 ΔH= -6.57 ± 0.07	N= 2.21 ± 0.12 K_d= 326 ± 25 ΔH= -9.69 ± 0.16	N= 1 K_d = 421 ±157 ΔH= -8.43 ± 4.82
	N= 2.13 ±0.11 K_d = 19 ±8 nM ΔH= -8.13 ±0.16	N= 1 K_d= 68 ± 37 ΔH= 10.95 ±11.83
		N= 1 K_d= 2 ± 1 ΔH= -22.31 ±8.39
		N= 1 K_d= 631 ±70 ΔH= -6.09 ±0.15

Table 3.3. Comparison of the different binding models of Gα_{i1} and AGS3-C interaction.
The thermodynamic parameters for Gα_{i1} and AGS3-C binding obtained using 'single set of identical sites', 'two sets of independent sites', and 'sequential binding sites' are compared. The units of enthalpy (ΔH) and dissociation constants (K_d) are kcal mol^{-1} and nM.

two sets of independent binding sites are assumed, provided the best fit to the titration data

(Figure 3.8 and Table 3.3). Accordingly the data were consistent with two strong (K_d ~20

nM) and two weak (K_d ~300 nM) Gα_{i1} binding sites in AGS3-C (Table 3.4). The two types

of sites differed only by ~1.5 kcal/mol in binding free energy and enthalpy. Binding to the

higher affinity sites was accompanied by a larger positive change in entropy but a smaller

negative enthalpy than is binding to the low affinity sites (Table 3.4).

3.3.3 Structural perturbation of AGS3 resulting in one set of identical binding sites

Binding of the two AGS3-C deletion mutants to $G\alpha_{i1}$ was also investigated using

ITC. Analysis of the thermal titration data indicated that AGS3-Δ4 and AGS3-Δ34 bind to

three and two molecules of $G\alpha_{i1}\cdot$GDP (Table 3.4), respectively, which again is consistent

with the gel filtration data. However, unlike the binding of AGS3-C to $G\alpha_{i1}$, the titration data

for both of the deletion mutants was most consistent with a model in which a single set of

identical binding sites is present (Figure 3.8). Although the deletion mutants share a similar

	N	K_d (nM)	ΔH (kcal mol^{-1})	ΔG (kcal mol^{-1})	ΔS (cal mol^{-1}K^{-1})
AGS3-C					
1. Strong site	2.13±0.11	19 ± 8	-8.13 ± 0.06	-10.30	7.41
2. Weak site:	2.21±0.12	326 ± 25	-9.69 ± 0.16	-8.67	-3.48
AGS3-Δ4	3.07 ± 0.01	384 ± 10	-6.96 ± 0.02	-8.57	5.49
AGS3-Δ34	2.03 ± 0.01	137 ± 8	-9.98 ± 0.04	-9.17	-2.76

Table 3.4. Thermodynamic parameters of the binding of AGS3-C and its deletion mutants to $G\alpha_{i1}$.
Thermodynamic parameters for the binding of AGS3-C and its deletion mutants with $G\alpha_{i1}$ was determined using ITC at 20°C in 20 mM Tris, pH 8.0, 2 mM DTT and 10 μM GDP. A binding model that assumes 'two sets of independent sites' best described the titration data for AGS3-C. Two different types of sites are referred to as 'strong site' and 'weak site' in the table. . Titration data for AGS3-Δ4 and AGS3-Δ34 were fit to a 'one set of identical sites' model. K_d, ΔH, ΔG, ΔS and N represents the dissociation constant, enthalpy, free energy, entropy and stoichiometry respectively and the units corresponding to these parameters are shown in the table.

binding mechanism, AGS3-Δ34 had almost 3-fold greater affinity for Gα$_{i1}$ than did AGS3-Δ4. The apparent change in the Gα$_{i1}$ binding mechanism of AGS3-C upon deletion of one or two of its C-terminal GPR domains could reflect an alteration of the tertiary structure or structural environment of the remaining GPR repeats. Alternatively the four repeats may differ substantially in their affinity for Gα$_{i1}$·GDP. Overall, the ITC data are consistent with the presence of four Gα$_{i1}$ binding sites in one molecule of AGS3, such that one binding site is lost with each deletion of a GPR motif.
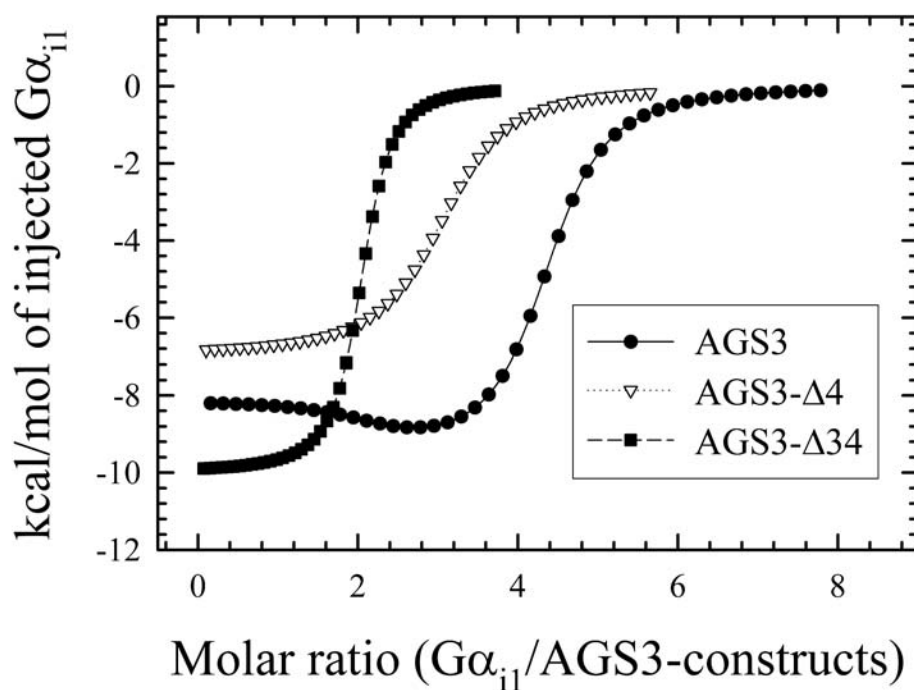
Figure 3.8. ITC analysis of the binding of AGS3-C and its deletion mutants with Gα$_{i1}$. The fitted titration profiles of AGS3-C (*closed circles*), AGS3-Δ4 (*open triangles*) and AGS3-Δ34 (*closed squares*) with Gα$_{i1}$ are superimposed. The difference in the shape of the curves for 'two sets of sites' (AGS3-C) and 'single set of sites' (AGS3-Δ4 and AGS3-Δ34) are clearly illustrated. With increasing number of Gα$_{i1}$ binding sites, the mid-point of the curves shift towards a higher molar ratio.

3.3.4 Significance of 'two sets of independent sites'

Using isothermal titration calorimetry and gel filtration analysis, I have demonstrated that AGS3-C can bind up to four molecules of $G\alpha_{i1}$, equal to the number of GPR motifs present in its amino acid sequence. Successive deletion of GPR motifs from the C-terminus of AGS3 does not abrogate the $G\alpha_{i1}$- binding activity of the remaining motifs. Hence GPR repeats can function independently within the context of the AGS3-C scaffold. Indeed the naturally occurring splice-variant of AGS3 known as AGS3-SHORT contains three complete GPR motifs and is known to be a functionally active molecule (Pizzinat et al., 2001). Nevertheless, the ITC titration data for the AGS3-C and its truncation mutants strongly suggest that the structural context of the GPR repeats within AGS3-C affects their affinity for $G\alpha_{i1}$.

Although all conform to the GPR-consensus sequence, the four GPR repeats in AGS3-C otherwise differ in sequence from each other, and therefore AGS3-C must have four non-equivalent binding sites; thermodynamic linkage among them is also possible. However, the experimental errors inherent in the ITC measurements do not allow thermodynamic parameters to be extracted for each binding site in the most general four-site model. On the other hand, the binding data clearly do not fit to the most simple model in which four identical, independent sites is assumed. Although the titration data can be fit to a sequential four-site binding model (8 free parameters), the uncertainties associated with derived constants are smaller if a two-site model (6 free parameters) is assumed. From this analysis, two high affinity ($K_d \sim 20$ nM) and two low affinity ($K_d \sim 300$ nM) binding sites for $G\alpha_{i1}$ can be derived. The deletion of the fourth GPR motif from the sequence results in a

conformation in which all the motifs behave identically but with a decreased binding affinity ($K_d \sim 400$ nM) for $G\alpha_{i1}$. Further deletion of the third motif does not change the binding mechanism, but increases $G\alpha_{i1}$ affinity and binding enthalpy for the remaining two sites ($K_d \sim 150$ nM).

Several protein families have evolved multiple GoLoco motifs, even though a single GoLoco motif is necessary and sufficient for biochemical activity. Molecules such as AGS3, could act as scaffolding molecules (Blumer et al., 2002; Pizzinat et al., 2001) to bring several $G\alpha$ subunits together and thereby enhance the efficiency of signaling. The 'strong sites' in AGS3-C differ from the 'weak sites' by 1.5 kcal/mol in binding free energy for $G\alpha_{i1}$, (~15 fold difference in K_d) suggesting that a somewhat linear response in GDI activity is possible over a 10-500 nM range in effective $G\alpha_i$ concentration. The broader dynamic range of AGS3-C might play a critical role in regulation of signaling *in vivo*.

3.3.5   GDI activity of AGS3 and AGS3 deletion mutants

The AGS3-C construct used in these experiments inhibit the GDP exchange activity of $G\alpha_{i1}$ as expected from previous studies (Figure 3.9). The apparent association rate for the binding of a fluorescent non-hydrolyzable GTP analog (BODIPY-GTPγS) to $G\alpha_{i1}$ was measured in the presence of AGS3-C and the two C-terminal truncation mutants. In keeping with earlier reports (Natochin et al., 2000), the GDP exchange rate of $G\alpha_{i1}$ was reduced 50% at 100 nM AGS3-C (Figure 3.9). Deletion of one or two GPR repeats from AGS3-C did not have a severe effect on nucleotide exchange inhibitory activity. Both AGS3-Δ4 and AGS3-Δ34 potently inhibited the binding of BODIPY-GTPγS to $G\alpha_{i1}$·GDP (Figure 3.10).

Surprisingly, AGS3-C and its truncation mutants have nearly equal potency as GDIs when tested at a concentration near their average K_d. At 100 nM concentration, AGS3-Δ4 and AGS3-Δ34 were almost equally potent inhibitors of nucleotide exchange of Gα_{i1} (Figure 3.11) as AGS3-C, indicating that the GPR repeats can function independently and that deletion of individual repeats from AGS3-C does not result in loss of its GDI activity.
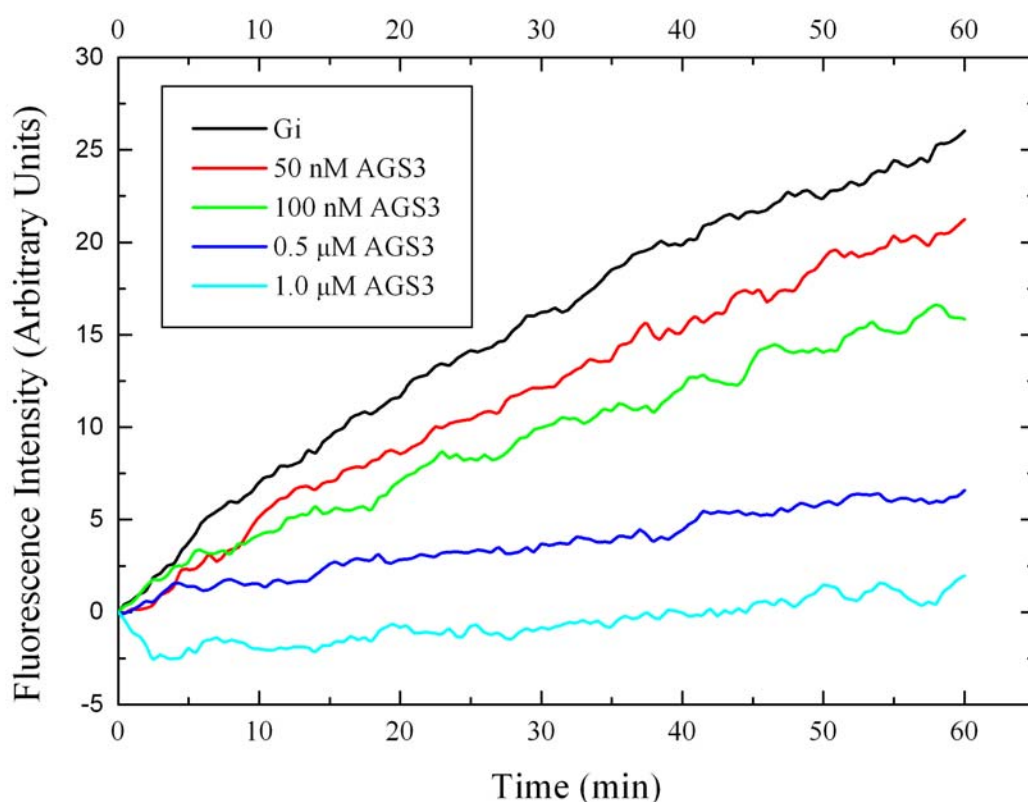
Figure 3.9. Inhibition of BOIDPY-GTPγS binding to Gα$_{i1}$ by AGS3-C. Time course of BODIPY-GTPγS (1 μM) binding to Gα$_{i1}$ (200 nM) in the absence (*black*) or presence (*red*, 50 nM; *green*, 100 nM; *blue*, 500 nM; *cyan*, 1 μM) of AGS3-C. Reactions were initiated by addition of pre-incubated AGS3:Gα$_{i1}$ complex to a cuvette containing BODIPY-GTPγS and fluorescence was monitored at 510nm for 60 min.

A

B
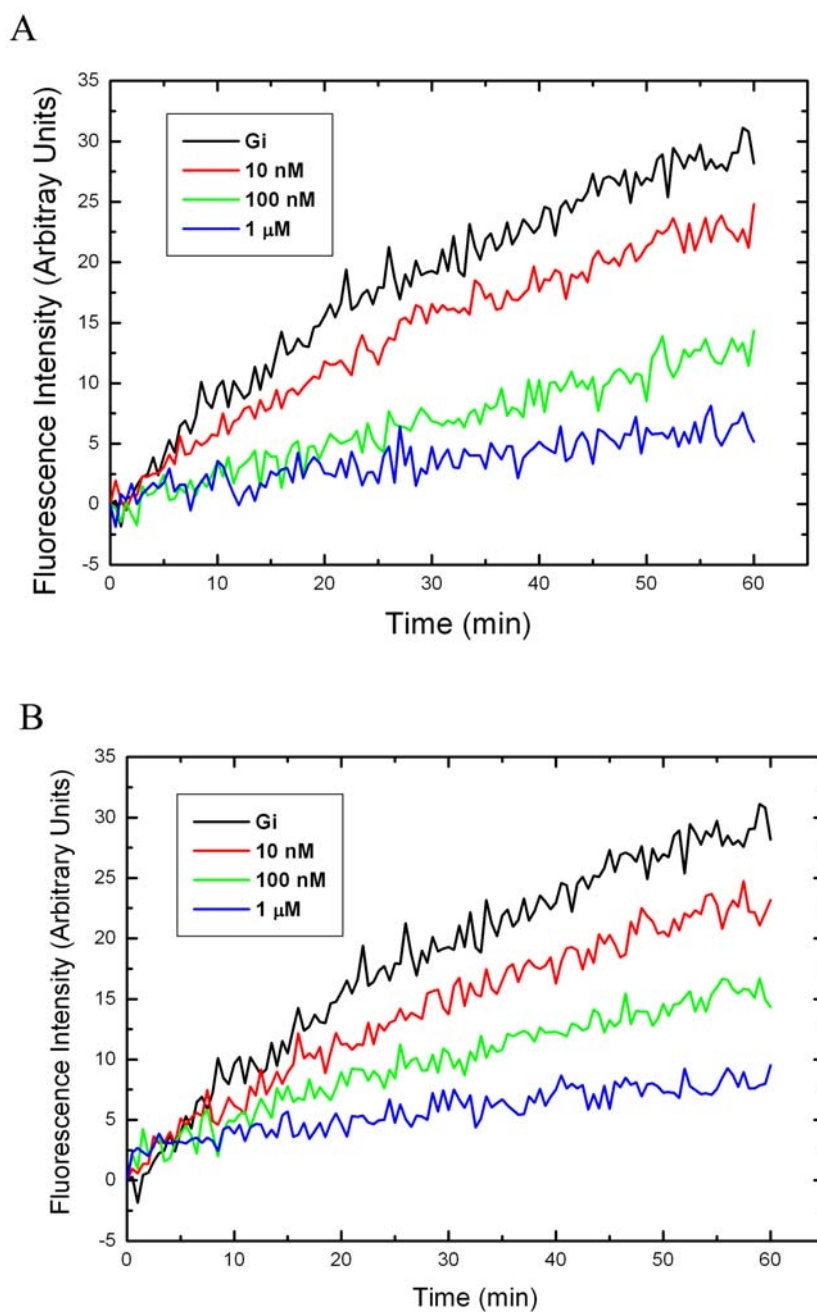
Figure 3.10. Effect of AGS3-C deletion mutants on the kinetics of BOIDPY-GTPγS binding to Gα$_{i1}$. *A,* Time course of BODIPY-GTPγS (1 μM) binding to Gα$_{i1}$ (200 nM) in the absence (*black*) or presence (*red,* 10 nM; *green,* 100 nM; *blue,* 1 μM) of AGS3-Δ4. *B,* Time course of BODIPY-GTPγS (1 μM) binding to Gα$_{i1}$ (200 nM) in the absence (*black*) or presence (*red,* 10 nM; *green,* 100 nM; *blue,* 1 μM) of AGS3-Δ34.
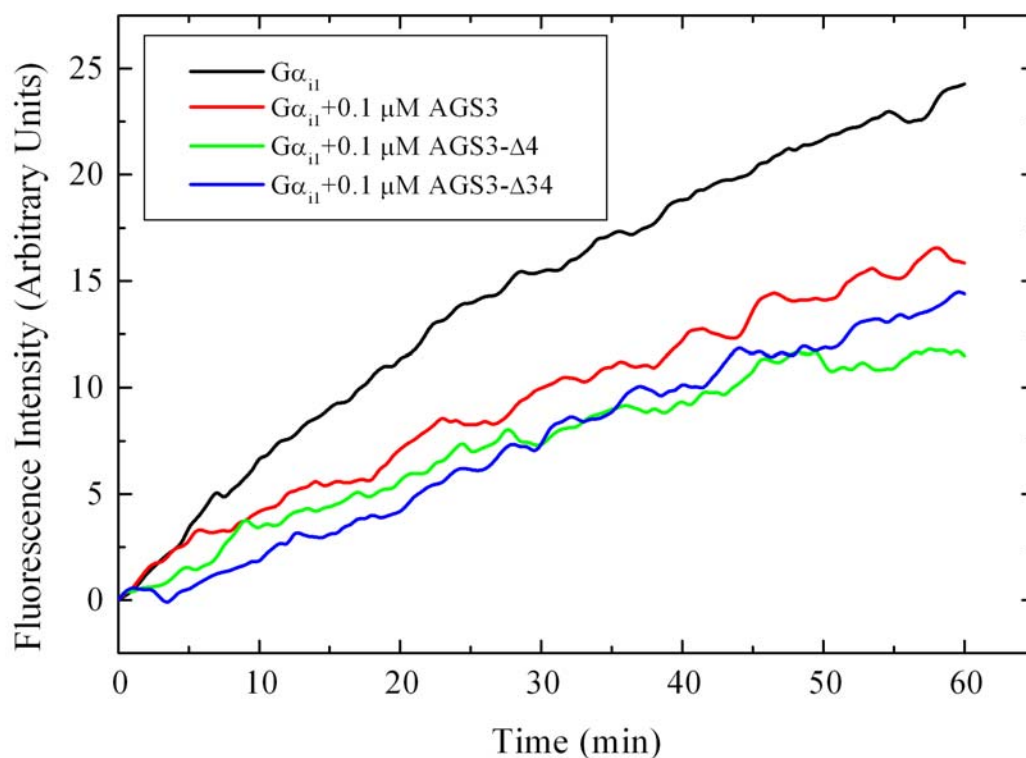
Figure 3.11. Comparison of the GDI activity of AGS3-C and the deletion mutants. Comparison of the time course of inhibition of BODIPY-GTPγS binding to Gα_{i1} (200 nM) by 100 nM AGS3-C (*red*), 100 nM AGS3-Δ4 (*green*) and AGS3-Δ34 (*blue*). The *black* curve represents BODIPY-GTPγS binding in the presence of Gα_{i1} only.

3.3.6   Predominantly helical structure of AGS3-C

In the absence of crystallographic structural information about AGS3-C, CD spectrometry was used to elucidate the secondary structure elements in AGS3-C. In addition the secondary structure of AGS3-C was also compared with the structure of individual GoLoco motifs. CD spectra were obtained for AGS3-C and a peptide having the consensus sequence of GoLoco motifs in AGS3 in the wavelength range of 190-250 nm. Analysis of the CD spectra indicated presence of significant fraction of disordered loop like region in both

AGS3-C and the peptide. Analysis of the CD spectra (Figure 3.12) also indicated higher α-helical content in AGS3-C (20.4% or ~38 residues of 186 residues) than in the GPR consensus peptide (8% or ~2 residues of 28 residues), suggesting that AGS3-C could have native helical structures that are in the proper conformation to interact with $G\alpha_{i1}$. The analysis also suggested that AGS3-C and the GPR peptide have approximately equal fractions of turn (~25%) and disordered structure (~33%), but the latter has a higher content of β-strand structure (33.5%) than AGS3-C (22.2%). The implications of the difference in secondary structure of AGS3-C and the GPR peptide is discussed in chapter 4.
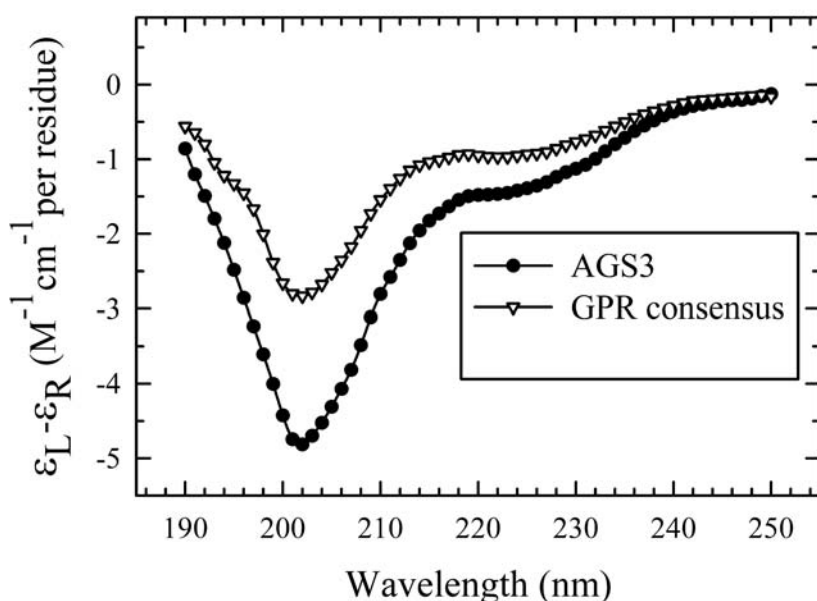
Figure 3.12: Analysis of secondary structure using CD spectroscopy. CD spectra of AGS3-C (*closed circles*) and GPR-consensus peptide (*open triangles*) in the wavelength range of 250-190 nm are shown. Secondary structure analysis (see chapter 3.2.6) of the spectra and averaging the result of the programs CDSSTR and CONTINLL suggested: 20.4 ± 0.2 % α-helix, 22.2 ± 0.3 % β-strand, 25.1 ± 0.1 % turn and 32.2 ± 0.6 % disordered structure for AGS3-C; 8 ± 0.8 % α-helix, 33.5 ± 1.0 % β-strand, 24.1 ± 0.1 % turn and 34.2 ± 0.5 % disordered structure for GPR-consensus peptide.

3.3.7   Difficulties in crystallization of Gα$_{i1}$:AGS3 complex

Although the complex of Gα$_{i1}$ and AGS3-C was isolated in a very high degree of

purity and homogeneity (Figure 3.5), crystals of the complex were not obtained upon

extensive screening of crystallization conditions. Only a few conditions appeared promising,

but careful pursuit of these conditions did not lead to crystals. Crystallization of Gα$_{i1}$ in

complex with AGS3-Δ4 and AGS3-Δ34 was also attempted, but the complex with the

deletion mutants did not lead to crystals either. The number of crystallization screens

performed covered a very wide range of crystallization conditions (Table 3.1 and 3.2) and

further screening is unlikely to lead to crystallization. To investigate whether homologs of

AGS3 might be more suitable crystallization targets, the C-terminal domain of human protein

LGN (contains four GoLoco motifs in C-terminal domain) and another human protein known

as G18.1b (contains three GoLoco motifs) were expressed and purified in *E.coli*. Both of

these homologs were poorly expressed and were degraded heavily within the cell. These

homologs appeared inferior crystallization targets to begin with and were not pursued further.

The most important requirement of crystallization is high degree of purity and

conformational homogeneity (McPherson, 1999). Although high degree of purity was

ensured for the Gα$_{i1}$:AGS3-C complex, the potential for conformational heterogeneity

remained. Computational prediction of AGS3-C secondary structure using the program Jpred

(Cuff et al., 1998), available at http://www.compbio.dundee.ac.uk/~www-jpred/, suggested

the presence of significant stretches of disordered and loop like elements in addition to

helical structures in the GoLoco regions. CD analysis of AGS3-C also indicated the presence

of ~60% disordered and loop like region (Figure 3.12). The linker regions connecting the
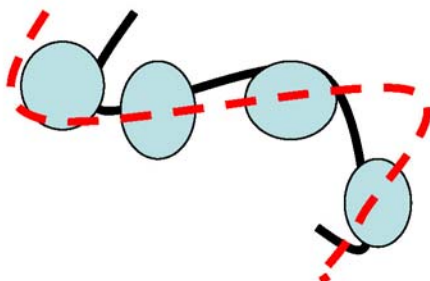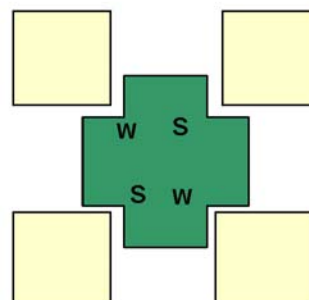
Figure 3.13. Models of Gα_{i1}/AGS3-C complex structure. *Right panel*. Schematic illustration of a conformationally flexible Gα_{i1}:AGS3-C complex structure is shown, that is behaving in a 'beads on a string' like fashion. *Left panel*. Schematic illustration of a compact Gα_{i1}:AGS3-C complex structure highlighting the presence of two strong (s) and two weak (w) sites.

GoLoco motifs might be flexible in conformation, allowing inter-domain movements. Gel-filtration chromatography of AGS3-C suggested an elongated and flexible structure, which eluted with an apparent molecular weight almost twice that of the real molecular mass. Interestingly, the hexa-histidine tagged form of AGS3-C is poorly soluble and can only be purified through a refolding protocol (chapter 2.2.1), indicating that AGS3-C in itself is probably a poorly folded molecule. AGS3-C by itself behaved poorly under crystallization screens, with significant precipitation over a broad range of precipitant concentration. Gα_{i1}:AGS3-C complex was significantly different from AGS3-C, in terms of behavior in crystallization screens as well as in gel-filtration purification. While this was promising behavior, structural flexibility of AGS3-C even in complex with Gα_{i1} is possible, which might interfere with crystallization. These observations leads to the possibility that individual

GoLoco motifs on AGS3-C can behave in a 'beads on a string' fashion instead of forming a well ordered, globular domain (Figure 3.13). This model is also supported by the independent nature of the individual binding sites as observed through ITC studies. Whether, AGS3-C forms a well ordered structure upon binding to $G\alpha_{i1}$ or whether it remains as a flexible 'beads on a string'-structure is not understood yet. Further structural studies of $G\alpha_{i1}$:AGS3-C complex might provide insight into the structural nature of the interaction between AGS3-C and $G\alpha_{i1}$. Some of the potentially promising approaches to crystallize AGS3 or its homologs will be discussed in Chapter 7 (Conclusion and future directions).

3.4   References

Bernard, M. L., Peterson, Y. K., Chung, P., Jourdan, J., and Lanier, S. M. (2001). Selective Interaction of AGS3 with G-proteins and the Influence of AGS3 on the Activation State of G-proteins. J Biol Chem *276*, 1585-1593.

Blumer, J. B., Chandler, L. J., and Lanier, S. M. (2002). Expression analysis and subcellular distribution of the two G-protein regulators AGS3 and LGN indicate distinct functionality. Localization of LGN to the midbody during cytokinesis. J Biol Chem *277*, 15897-15903.

Coleman, D. E., Lee, E., Mixon, M. B., Linder, M. E., Berghuis, A., Gilman, A. G., and Sprang, S. R. (1994). Crystallization and preliminary crystallographic studies of $G_{i\alpha 1}$ and mutants of $G_{i\alpha 1}$ in the GTP and GDP-bound states. JMolBiol *238*, 630-634.

Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M., and Barton, G. J. (1998). JPred: a consensus secondary structure prediction server. Bioinformatics *14*, 892-893.

De Vries, L., Fischer, T., Tronchere, H., Brothers, G. M., Strockbine, B., Siderovski, D. P., and Farquhar, M. G. (2000). Activator of G protein signaling 3 is a guanine dissociation inhibitor for galpha i subunits. Proc Natl Acad Sci U S A *97*, 14364-14369.

Ferguson, K. M., Higashijima, T., Smigel, M. D., and Gilman, A. G. (1986). The influence of bound GDP on the kinetics of guanine nucleotide binding to G proteins. J Biol Chem *261*, 7393-7399.

Fisher, H. F., and Singh, N. (1995). Calorimetric methods for interpreting protein-ligand interactions. Methods Enzymol *259*, 194-221.

Gille A, S. R. (2003). Low-affinity interactions of BODIPY-FL-GTPgammaS and BODIPY-FL-GppNHp with G(i)- and G(s)-proteins. Naunyn Schmiedebergs Arch Pharmacol *68*, 210-215.

Higashijima, T., and Ferguson, K. M. (1991). Tryptophan fluorescence of G proteins: analysis of guanine nucleotide binding and hydrolysis. Methods in Enzymology *195*, 321-328.

Higashijima, T., Ferguson, K. M., Sternweis, P. C., Ross, E. M., Smigel, M. D., and Gilman, A. G. (1987). The effect of activating ligands on the intrinsic fluorescence of guanine nucleotide-binding regulatory proteins. J Biol Chem *262*, 752-756.

Indyk, L., and Fisher, H. F. (1998). Theoretical aspects of isothermal titration calorimetry. Methods Enzymol *295*, 350-364.

Jancarik, J., and Kim, S.-H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. Journal of Applied Crystallography *24*, 409-411.

Jelesarov, I., and Bosshard, H. R. (1999). Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. J Mol Recognit *12*, 3-18.

Johnson, W. C. (1999). Analyzing protein circular dichroism spectra for accurate secondary structures. Proteins *35*, 307-312.

Kimple, R. J., Kimple, M. E., Betts, L., Sondek, J., and Siderovski, D. P. (2002a). Structural determinants for GoLoco-induced inhibition of nucleotide release by Galpha subunits. Nature *416*, 878-881.

Kimple, R. J., Willard, F. S., and Siderovski, D. P. (2002b). The GoLoco motif: heralding a new tango between G protein signaling and cell division. Mol Interv *2*, 88-100.

McEwen, D. P., Gee, K. R., Kang, H. C., and Neubig, R. R. (2001). Fluorescent BODIPY-GTP analogs: real-time measurement of nucleotide binding to G proteins. Anal Biochem *291*, 109-117.

McPherson, A. (1999). Crystallization of biological macromolecules (New York: Cold Spring Harbor Laboratory Press).

Microcal (1998). Appendix: Equations used for fitting ITC binding data. ITC Data Analysis in Origin *5.0*, 73-78.

Natochin, M., Gasimov, K. G., and Artemyev, N. O. (2001). Inhibition of GDP/GTP exchange on G alpha subunits by proteins containing G-protein regulatory motifs. Biochemistry *40*, 5322-5328.

Natochin, M., Lester, B., Peterson, Y. K., Bernard, M. L., Lanier, S. M., and Artemyev, N. O. (2000). AGS3 inhibits GDP dissociation from galpha subunits of the Gi family and rhodopsin-dependent activation of transducin [In Process Citation]. J Biol Chem *275*, 40981-40985.

Peterson, Y. K., Bernard, M. L., Ma, H., Hazard, S., 3rd, Graber, S. G., and Lanier, S. M. (2000). Stabilization of the GDP-bound conformation of Gialpha by a peptide derived from the G-protein regulatory motif of AGS3. J Biol Chem *275*, 33193-33196.

Pierce, M. M., Raman, C. S., and Nall, B. T. (1999). Isothermal titration calorimetry of protein-protein interactions. Methods *19*, 213-221.

Pizzinat, N., Takesono, A., and Lanier, S. M. (2001). Identification of a truncated form of the G-protein regulator AGS3 in heart that lacks the tetratricopeptide repeat domains. J Biol Chem *276*, 16601-16610.

Provencher, S. W., and Glöckner, J. (1981). Estimation of globular protein secondary structure from circular dichroism. Biochemistry *20*, 33-37.

Remmers, A. E., Posner, R., and Neubig, R. R. (1994). Fluorescent guanine nucleotide analogs and G protein activation. JBiolChem *269*, 13771-13778.

Rojas, R. J., Kimple, R. J., Rossman, K. L., Siderovski, D. P., and Sondek, J. (2003). Established and emerging fluorescence-based assays for G-protein function: Ras-superfamily GTPases. Comb Chem High Throughput Screen *6*, 409-418.

Segelke, B. W. (2001). Efficiency Analysis of Screening Protocols Used in Protein Crystallization. Journal of Crystal Growth *232*, 553-562.

Sreerama, N., Venyaminov, S. Y., and Woody, R. W. (1999). Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy. Protein Sci *8*, 370-380.

Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Anal Biochem *179*, 131-137.

Wiseman T, W. S., Brandts JF, Lin LN (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Anal Biochem *179*, 131-137.

Interaction of AGS3-derived peptides with Gα_{i1}

4.1 Introduction

The amino acid sequences of the four GPR motifs in AGS3 are similar and residues deemed to be functionally critical are absolutely conserved (Peterson et al., 2002). However, residues that are not conserved might lead to variation in their interaction with Gα_{i1}. Earlier studies also suggested that each individual GoLoco motifs of AGS3, expressed as GST fusion proteins are capable of binding to Gα_{i1} (Bernard et al., 2001). A peptide having the consensus sequence of GoLoco motifs in AGS3 is a potent GDI (Peterson et al., 2000). In spite of this preliminary information, the affinities and activities of individual GoLoco motifs of AGS3 were not studied in detail. Characterization of the interaction of individual GoLoco motifs of AGS3 and Gα_{i1} was pursued to address the following questions:

- What is the affinity and activity of individual GoLoco motifs of AGS3?

- Does the variation of sequence among GoLoco motifs affect their binding affinity towards Gα_{i1} and their GDI activity?

- Is the presence of two sets of independent binding sites in AGS3-C a result of sequence variation in GoLoco motifs?

- What is the contribution of each GoLoco motif in AGS3-C towards the binding of AGS3-C to Gα_{i1}?

- Does the structural context of GoLoco motifs in AGS3 affect its activity?

- What are the roles of residues flanking GoLoco motifs?

To address these questions, the interaction between peptides corresponding to each of the four GoLoco motifs of AGS3 and $G\alpha_{i1}$ was studied using ITC and their activities were studied by fluorescence spectroscopy. In this chapter, I will describe the thermodynamic nature of the interaction of GoLoco motif peptides with $G\alpha_{i1}$, the role of the structural context of GoLoco motifs in their activity, and also the role of sequences outside of GoLoco motifs in their activity.

4.2   Materials and methods

4.2.1   Synthesis of GPR-peptides derived from AGS3

Peptides that correspond in amino acid sequence to the GoLoco motifs 1-4 (Figure 4.1 A) of rat AGS3 were synthesized in the Protein Chemistry Technology Center at The University of Texas Southwestern Medical Center at Dallas, using solid-phase fmoc synthetic chemistry. These peptides are referred to as: GPR1, ^{469}EECFFDLLSKFQSSRMDDQRCPL^{492}; GPR2, ^{524}EEFFDLIASSQSRRLDDQRASV^{545}; GPR3, KK^{571}GDEFFNMLIKYQSSRIDDQRCPP^{593}; and, GPR4, ^{606}EDFFSLIQRVQAKRMDEQRVDL^{627}. GPR3, containing only the residues of the third GoLoco motif, was poorly soluble. The solubility of the GPR3 peptide was enhanced by the addition of two lysine residues at the N-terminus. In addition, the GPR consensus peptide (Peterson et al., 2000), TMGEEDFFDLLAKSQSKRMDDQRVDLAG, was also synthesized.

The C terminus of each peptide was blocked by amidation. The purity of the synthetic

peptides was verified using electrospray mass spectrometry and analytical high performance

liquid chromatography. Peptides were desalted by elution through Sep-pack C18 cartridges

(Waters), lyophilized, and stored at -20 °C.


```
A
GPR1          ---------EECFFDLLSKFQ SSRMDDQRCPL---- --------  23
GPR2          ---------EE-FFDLIASSQ SRRLDDQRASV---- --------  22
GPR3          ------KKGDE-FFNMLIKYQ SSRIDDQRCPP---- --------  25
GPR4          ---------ED-FFSLIQRVQ AKRMDEQRVDL---- --------  22
GPR-consensus -----TMGEED-FFDLLAKSQ SKRMDDQRVDLAG-- --------  28
RGS14-GoLoco  -------DIEG-LVELLNRVQ SSGAHDQRGLLRKED LVLPEFLQ  36
B
GPR1ex        --PRAPSSDEECFFDLLSKFQ SSRMDDQRCPLEEGQ AGAAE---  39
GPR2ex        -SVTASPQTEE-FFDLIASSQ SRRLDDQRASVGSLP GLRITL--  40
GPR3ex        GDGDPQEPGDE-FFNMLIKYQ SSRIDDQRCPPPDVL PRGPT---  40
```

Figure 4.1. Multiple sequence alignment of GPR peptides. *A*, Sequences of the four synthetic GPR peptides of AGS3-C are shown along with the GoLoco peptide derived from RGS14 (Kimple et al., 2002) and the GPR consensus peptide (Peterson et al., 2000). The highly conserved residues that are critical for GDI activity (Peterson et al., 2002) are colored *red*, absolutely conserved but functionally non-critical residues are colored *green*, other highly conserved but functionally non-critical positions are colored *blue*, and non-native residues added to enhance solubility are colored *orange*. The length of each peptide is listed at the *right* of its sequence. *B*, Multiple sequence alignment of the extended GPR peptides GPR1ex, GPR2ex, and GPR3ex. Flanking residues are colored *gray*. Extended GPR peptides were expressed as GST fusion proteins.

4.2.2 Purification of GST-fusion extended GPR peptides

pGEX-4T-1 expression vectors (Amersham Biosciences) encoding peptides

corresponding to AGS3 GoLoco motifs extended with native N- and C-terminal flanking

residues (Figure 4.1 B) were a gift from Dr. Stephen M. Lanier (Department of

Pharmacology, Louisiana State University Health Sciences Center, New Orleans, LA). These

peptides were expressed as N-terminal fusion proteins with GST (Bernard et al., 2001). The

fusion peptides corresponding to the extended forms of the first three GoLoco motifs are

referred to as GPR1ex (Pro[463]-Glu[501]), GPR2ex (Ser[516]-Leu[555]), and GPR3ex (Gly[563]-Thr[602]),

respectively. GST fusion proteins with extended GPR peptides were expressed in BL21

(DE3) strains. Cells (1 liter) were grown to an OD600 of 0.5-0.6 and induced with 100 μM

IPTG at 30°C for 3 h. Cells were harvested by centrifugation, flash frozen in liquid nitrogen,

and stored at -80 °C. Frozen cells were thawed and resuspended in lysis buffer containing 1x

PBS (1 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 2 mM KCl, 140 mM NaCl, pH 7.4), 5 mM EDTA,

5 mM DTT, and complete protease inhibitor mixture (Roche Applied Science). Cell lysis was

performed by brief 1-min sonication on ice followed by incubation with 1% Triton X-100

with continuous stirring at 4 °C. The cell lysate was centrifuged at 35,000 rpm for 40 min at 4

°C in a Beckman Ti45 rotor. Cleared lysate was loaded on glutathione-Sepharose 4B resin

pre-equilibrated with lysis buffer. The resin was washed with lysis buffer, and the fusion

proteins were eluted using buffer containing 50 mM Tris, pH 8.0, 100 mM NaCl, 5 mM DTT,

5 mM EDTA, and 25 mM reduced glutathione. The fusion proteins were dialyzed against a

buffer containing 20 mM Tris, pH 8.0, and 1 mM EDTA, concentrated, and stored at 4 °C.

Protein concentration was estimated by measuring A_{280} and using theoretically predicted

extinction coefficients.

4.2.3 ITC studies of $G\alpha_{i1}$:GPR peptide binding

Isothermal titration calorimetry studies and data analysis was performed as described earlier (Chapter 3.2.2). ITC experiments were performed at 20 °C (293 K) using a MicroCal VP-ITC (MicroCal, Northhampton, MA) calorimeter. Protein samples were dialyzed against titration buffer (20 mM Tris, pH 8.0, 2 mM DTT, and 10 µM GDP). Because peptides could not be dialyzed due to their lower molecular weight, lyophilized peptides were dissolved directly in titration buffer. Samples were centrifuged at 13,000 rpm in a bench-top microcentrifuge (Forma Scientific) for 5 min before loading in the sample cell or syringe. Contents of the sample cell were stirred continuously at 300 rpm during the experiment. A typical titration of $G\alpha_{i1}$ with GPR peptides involved 25-30 injections at 3-min intervals of 8-10 µl of peptide solution (1 mM) into a sample cell containing 1.4 ml of $G\alpha_{i1}$ (45-50 µM). The result of the titration experiments were analyzed using Mircocal Origin software.

4.2.4 ITC studies of the binding of $G\alpha_{i1}$ to GST-GPRex fusion proteins

$G\alpha_{i1}$ (175-225 µM) was injected (35 injections of 8 µl) into the sample cell containing GPRex fusion peptide (20-25 µM), and NaCl (100 mM) was added to the titration buffer to avoid nonspecific interactions with GST.

4.2.5 Monitoring GDI activity by fluorescence spectroscopy

GDI activity of GPR peptides and extended GPR peptides were studied by measuring the rate of the binding of fluorescent GTP analog BODIPY-GTPγS to $G\alpha_{i1}$. The details of BODIPY-GTPγS based GDI assay was described in chapter 3.2.4. $G\alpha_{i1}$ (final concentration,

200 nM) in assay buffer (50 mM Tris, pH 8.0, 1 mM EDTA, and 10 mM $MgCl^2$) was

incubated in the presence or absence of GPR peptides and GST-GPRex for 1–2 h at 25 °C.

Samples were transferred into a 3-ml quartz cuvette with a path length of 1 cm containing 1

μM BODIPY-GTPγS in assay buffer, and nucleotide exchange was initiated by rapid mixing.

Binding of fluorescent nucleotide to $G\alpha_{i1}$ was monitored by the intensity of fluorescence

emission at 510 nm using a PerkinElmer Life Sciences LS50B spectrophotometer (data

shown in Figures 4.3, 4.4, and 4.5 A) or a custom-built fluorescence spectrophotometer

(Photon Technology International, Figure 4.8) The excitation wavelength was set to 485 nm

with the slit widths for excitation and emission maintained at 2.5 nm. All assays were

performed at 30 °C. Typically fluorescence intensity was recorded at 30 s intervals with a 10

s averaging time over a period of 1 h after mixing of the samples.

4.3   Results and discussion

4.3.1   Stoichiometry and affinity of $G\alpha_{i1}$:GPR peptide binding

The apparent differences in the binding models for AGS3-C and that of its truncation

mutants could reflect differences in the affinities of individual GoLoco motifs for $G\alpha_{i1}$. These

might arise from sequence variation at positions other than those that are conserved and

critical for function (Peterson et al., 2002). To investigate this possibility, the energetic

contribution of each GoLoco motif to the binding of AGS3-C to $G\alpha_{i1}$ was determined.

Peptides corresponding to each of the four GPR motifs in AGS3 were synthesized, and the

binding of these peptides to $G\alpha_{i1}$ was analyzed by ITC (Figure 4.2). The interaction of the
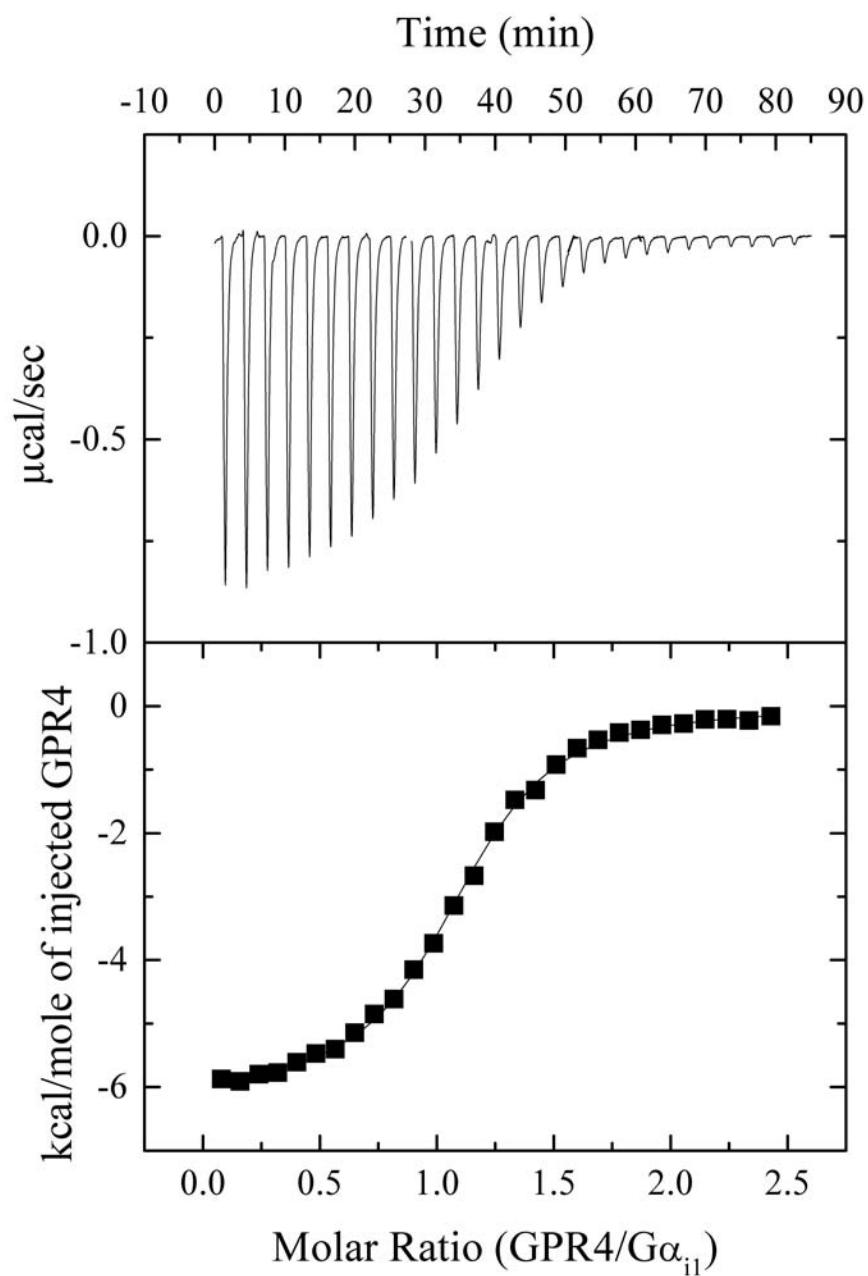
Figure 4.2. Typical ITC profile showing the titration of Gα$_{i1}$ with GPR4. Aliquots (10 μl) of 0.5 mM GPR4 were injected into an ITC cell containing 1.4 ml of 45 μM of Gα$_{i1}$. Nonlinear least square fit using the "one set of sites" model yielded a dissociation constant (K$_d$) of 1.9 μM, enthalpy (ΔH) of -6.12 kcal/mol and stoichiometry (N) of 1.09.

previously described GPR consensus peptide (Peterson et al., 2000) with $G\alpha_{i1}$ was also

analyzed. The titration data for the GPR peptides are best described by a single binding site

model. As expected, each of the peptides bound to $G\alpha_{i1}$ with a stoichiometry of ~1:1. The

dissociation constant (K_d) for the peptides ranged from 1 to 8 μM, and the enthalpy of

binding ranged from -3.5 to -9.8 kcal/mol (Table 4.1). The differences in the enthalpy

changes were balanced by entropic terms such that binding free energies of individual

peptides differed by no more than 1.2 kcal/mol. In the case of GPR2, the binding was

enthalpically least favorable among all the GPR peptides, but this was compensated by a

larger positive change in entropy (Table 4.1). The affinity of the consensus peptide for $G\alpha_{i1}$

was somewhat higher than that of the four GPR peptides. This maybe due to the presence of

additional residues at the N and C termini of the consensus peptide compared to the other

GPR peptides. The average free energy of binding for the four GPR peptides was -7.3

kcal/mol, which is significantly less in absolute value than the average free energy of binding

per site for AGS3-C (-10.3 kcal/mol) and for AGS3-Δ4 and AGS3-Δ34 (-8.6 and -9.2

kcal/mol, respectively) (Chapter 3.3.2). Since sequence variation among the AGS3-C

GoLoco motifs appears to have little effect on their affinity for $G\alpha_{i1}$, it is possible that

residues outside of the GoLoco motif are involved in binding or that the three-dimensional

structure of AGS3-C contributes to the stability of the complex. Indeed, CD studies indicate

the presence of a higher proportion of α-helical structure in AGS3 in comparison with the

GPR consensus peptide (chapter 3.3.6). The existence of "pre-formed" secondary structure

consistent with the requirements of the $G\alpha_{i1}$ binding site may explain the 2.0 kcal/mol (per

binding site) increase in binding energy for GPR repeats in AGS3-C relative to that for

individual GPR peptides. GPR peptides potentially have to undergo a disorder to order

transition before binding to $G\alpha_{i1}$. This will incur an entropic cost that might lower the overall

binding affinity. Alternately, residues outside of the GoLoco motif might also contribute to

the binding of AGS3-C to $G\alpha_{i1}$ (discussion in chapter 4.3.3). In essence, the results of the

ITC studies demonstrate that the structural context of GoLoco motifs affects its binding

characteristics.

Peptide	N	K_d (μM)	ΔH (kcal mol^{-1})	ΔG (kcal mol^{-1})	ΔS (cal mol^{-1} K^{-1})
GPR1	1.23 ± 0.01	8.18 ± 0.57	-9.87 ± 0.17	-6.80	-10.49
GPR2	1.05 ± 0.02	3.29 ± 0.19	-3.57 ± 0.09	-7.33	12.82
GPR3	1.02 ± 0.01	2.98 ± 0.13	-5.12 ± 0.04	-7.38	7.72
GPR4	1.09 ± 0.01	1.90 ± 0.08	-6.11 ± 0.03	-7.65	5.21
GPR-consensus	1.13 ± 0.01	1.02 ± 0.14	-7.32 ± 0.10	-8.01	2.39

Table 4.1. Thermodynamic parameters of GPR peptide $G\alpha_{i1}$ binding. Thermodynamic
parameters describing the binding of GPR peptides (GPR1, GPR2, GPR3, GPR4 and GPR-
consensus) and $G\alpha_{i1}$ were determined at 20°C using ITC in 20 mM Tris, pH 8.0, 2 mM DTT
and 10 μM GDP. The titration data were fitted to a single site binding model.

4.3.2 GDI activity of GPR peptides derived from AGS3

The GPR consensus peptide has been shown earlier to be a potent GDI (Peterson et al., 2000), but the activities of individual GoLoco motifs were not measured. To investigate the possibility that GPR peptides differ in GDI activity, the nucleotide exchange rate of $G\alpha_{i1}$ was determined in the presence of each peptide. For some of the peptides, the rate of nucleotide exchange was measured at different concentrations of the peptides to establish the concentration dependence of activity (Figure 4.3). To obtain a comparative measure of the activity of different GPR peptides, peptides were used in the assay at a concentration of 10 μM, which exceeds the K_d of GPR1 (which has the lowest affinity of the four for $G\alpha_{i1}$). Under these conditions, all the GPR peptides with the exception of GPR3 demonstrated significant nucleotide exchange inhibitory activity (Figure 4.3 and 4.4). Relative GDI activity of the peptides followed the rank order: GPR consensus > GPR1 > GPR2 > GPR4 >> GPR3 (Figure 4.4). In contrast, the rank order of binding affinity as determined by ITC was: GPR consensus > GPR4 > GPR3 ⁓GPR2 > GPR1, indicating that the binding affinity and activity were not well correlated.

Given the sequence similarity of the GPR peptides, the observation that GPR3 failed to inhibit nucleotide exchange at 10 μM was intriguing. At higher concentrations GPR3 exhibited some GDI activity (Figure 4.3), but this was not observed in a reproducible manner; perhaps aggregation of the $G\alpha_{i1}$-GPR3 complex at higher peptide concentrations resulted in a loss of the active pool of $G\alpha_{i1}$. GPR3 also failed to inhibit the increase in intrinsic tryptophan fluorescence of $G\alpha_{i1}$ in the presence of excess GTPγS, again indicating that it cannot inhibit nucleotide exchange (data not shown) (Higashijima et al., 1987).
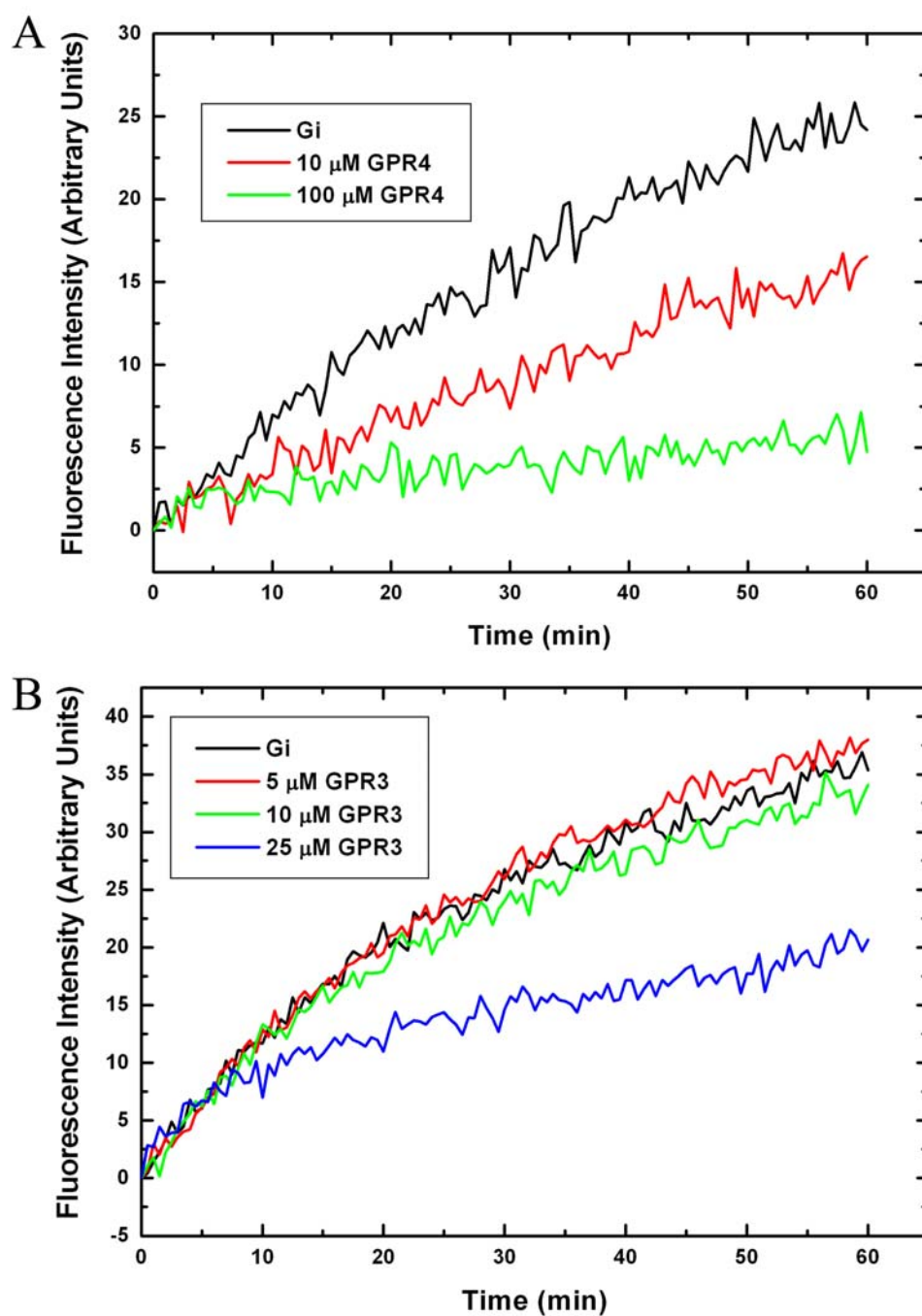
Figure 4.3. Inhibition of BODIPY- GTPγS binding to Gα$_{i1}$ by GPR peptides. *A*. Time course experiment showing inhibition of nucleotide exchange by GPR4 at 10 and 100 μM concentrations.*B*. GPR3 does not inhibit nucleotide exchange in the 5-10 μM concentration range but inhibits at 25 μM.
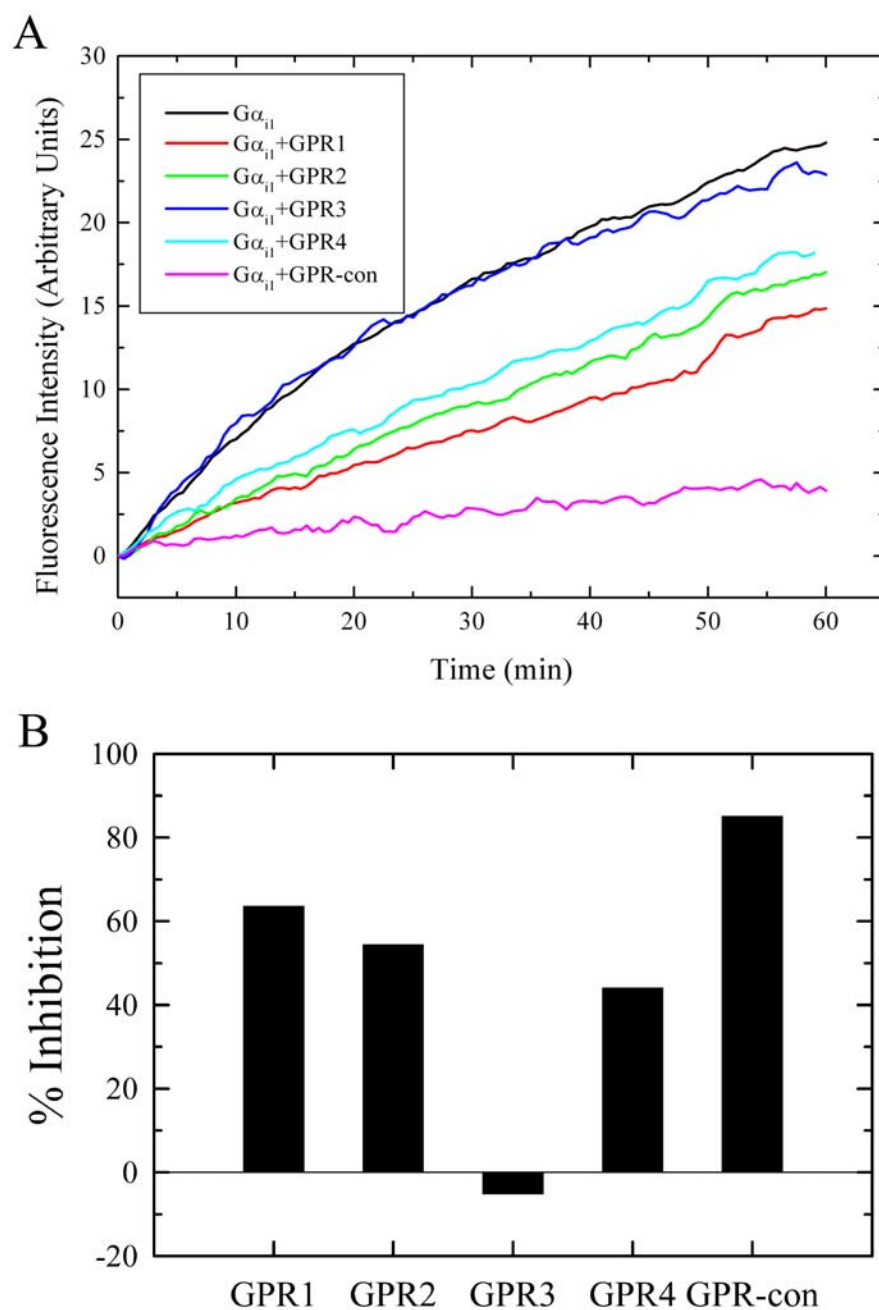
Figure 4.4. Comparison of the inhibition of BODIPY-GTPγS binding to Gα$_{i1}$ by GPR peptides. *A,* Comparison of the time course of BODIPY-FL-GTPγS binding to Gα$_{i1}$ in the presence of various GPR peptides. Gα$_{i1}$ (200 nM) was incubated with 10 μM of GPR1 (*red*), GPR2 (*green*), GPR3 (*blue*), GPR4 (*cyan*) and GPR-consensus (*magenta*). The pre-incubated solutions were mixed with 1 μM BODIPY-GTPγS and fluorescence intensity was monitored at 510 nm. The *black* curve represents the binding of BODIPY-GTPγS with Gα$_{i1}$ in the

absence of any peptide. *B,* Histogram showing GDI the activity of different peptides at 10 µM concentration. Fractional exchange at any time point was calculated as a ratio of fluorescence intensity of the sample with peptide and without peptide. % Inhibition = (1-fractional exchange)*100. The peptides can be arranged in descending order of activity as GPR-consensus>GPR1>GPR2>GPR4>>GPR3.

Because GPR3 bound to $G\alpha_{i1}$ in an unproductive manner, it was investigated whether GPR3 shares the same binding sites as the other GPR peptides. If GPR3 shares the same binding site as other GPR peptides, it is conceivable that it will compete with other GPR peptides for binding to $G\alpha_{i1}$. In such a situation, it is expected that GPR3 would be able to release the inhibition of nucleotide exchange by other GPR peptides. GPR1 strongly inhibited the binding of BODIPY-GTPγS to $G\alpha_{i1}$. But when the same assay was performed in the presence of 25 µM GPR3 in addition to GPR1, the inhibition of BOIDPY-GTPγS binding was significantly lower (Figure 4.5). The competition assay unambiguously demonstrated that GPR3 does not inhibit nucleotide exchange, even though it binds to the same site as other GPR peptides.

The lack of activity of GPR3 suggests that, it adopts a binding mode that is apparently unproductive for GDI activity. The conserved core of residues common to all GPR motifs is retained in GPR3 (Figure 4.1). The only significant differences between the sequence of GPR3 and those of the other peptides are: *1*, presence of a hydrophobic isoleucine residue (I 579) at a position that is normally occupied by either a solvent exposed polar residue or alanine (Kimple et al., 2002); *2,* substitution of a isoleucine for a methionine at position (I 586) which was shown to be important for GDI activity(Peterson et al., 2002), and *3*, the occurrence of a pair of proline residues immediately following the catalytic arginine that

binds the β phosphate of GDP in the $G\alpha_{i1}$•GDP complex (Kimple et al., 2002). It is also

possible that GPR3 possesses GDI activity towards a Gα subunit other than $G\alpha_i$ (although

AGS3-C has no GDI activity towards $G\alpha_o$), or is active only in the context of the AGS3-C

domain.
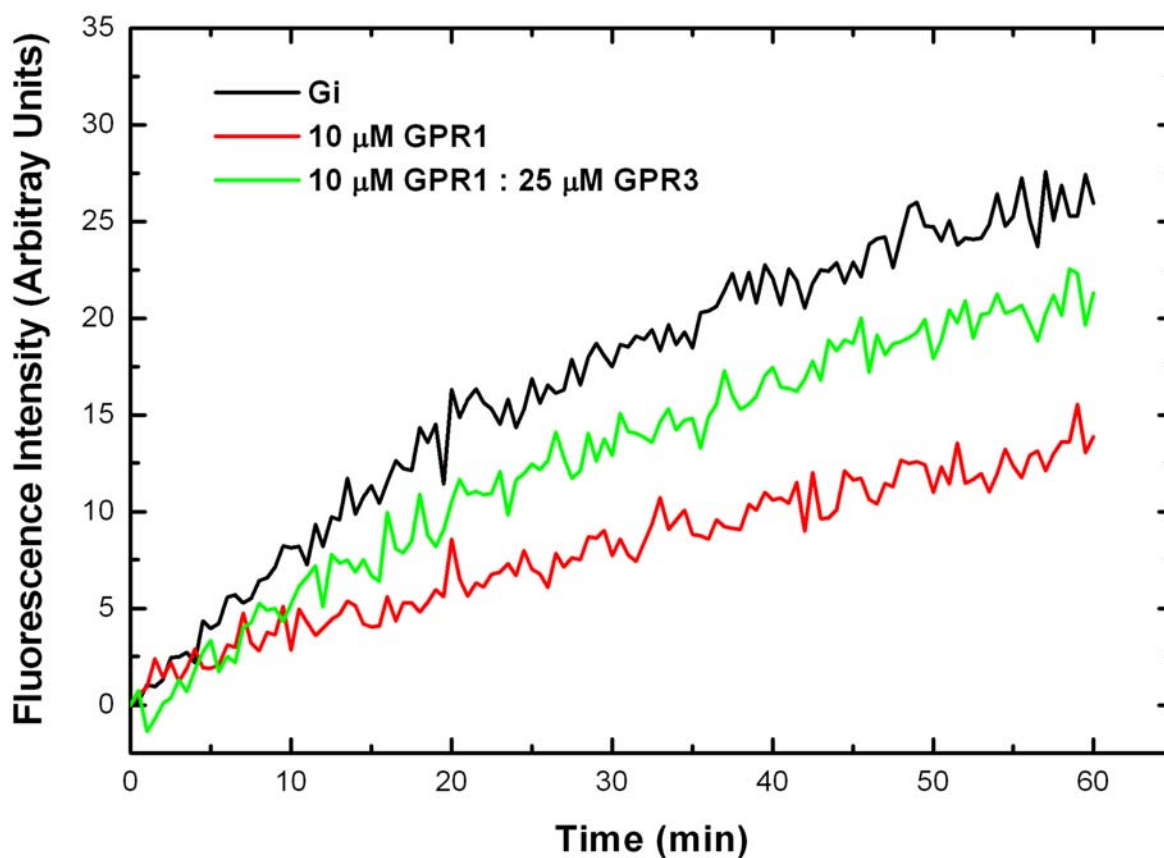
Figure 4.5. GPR3 competitively releases the inhibition of nucleotide exchange by GPR1.
The binding of BODIPY-GTPγS to 200 nM $G\alpha_{i1}$ (black) is strongly inhibited in the presence
of 10 μM GPR1 (red). Addition of a mixture of 10 μM GPR1 and 25 μM GPR3 (green) has
lower inhibition of BODIPY-GTPγS binding compared to that of 10 mM GPR1.

4.3.3   Effect of C-terminal residues of GoLoco motifs on binding of $G\alpha_{i1}$

In light of the observation that residues extending C-terminal to the GoLoco motif of

RGS14 contribute to $G\alpha$ specificity and GDI activity (Kimple et al., 2002), the effect of

residues outside of the GoLoco motifs was investigated. The extended peptides (Figure 4.1)

were expressed as N-terminally GST-tagged fusion proteins. The proteins were purified to

near homogeneity by glutathione affinity chromatography and used without further

purification (Figure 4.6). Because GST-GPR4ex was rapidly degraded despite the presence of

protease inhibitors, it was not tested for $G\alpha_{i1}$ binding or GDI activity. The affinity of the

extended GPR peptides towards $G\alpha_{i1}$ was measured using ITC. Initially, injecting highly

concentrated GST-GPRex to a sample cell containing $G\alpha_{i1}$ resulted in poor quality ITC data.

This was possibly caused by non-specific interaction with GST or GST dimerization.

Addition of 100 mM NaCl to the titration buffer and use of lower concentration of GST-

GPRex peptides in the sample cell instead in the syringe allowed the collection of high

quality ITC data (Figure 4.7).
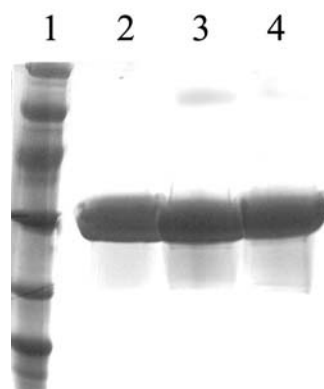
Figure 4.6. Purification of GST-fusion extended GPR peptides. SDS-PAGE showing
GST-extended GPR peptides after elution from the column. 1, molecular weight marker; 2,
GST-GPR1ex; 3, GST-GPR2ex; 4, GST-GPR3ex.

Peptide	N	K_d (μM)	ΔH (kcal mol^{-1})	ΔG (kcal mol^{-1})	ΔS (cal mol^{-1} K^{-1})
GPR1ex	0.74 ± 0.00	0.25 ± 0.01	-8.62 ± 0.02	-8.82	0.66
GPR2ex	0.88 ± 0.00	0.07 ± 0.00	-4.49 ± 0.02	-9.58	17.37
GPR3ex	0.98 ± 0.00	0.27 ± 0.01	-10.34 ± 0.02	-8.78	-5.32

Table 4.2. Thermodynamic parameters of extended GPR peptide Gα_{i1} binding.
Thermodynamic parameters describing the binding of extended GPR peptides (GPR1ex,
GPR2ex, GPR3ex) with Gα_{i1} were determined at 20°C using ITC in 20 mM Tris, pH 8.0, 2
mM DTT and 10 μM GDP. The titration data were fitted to a single site binding model.
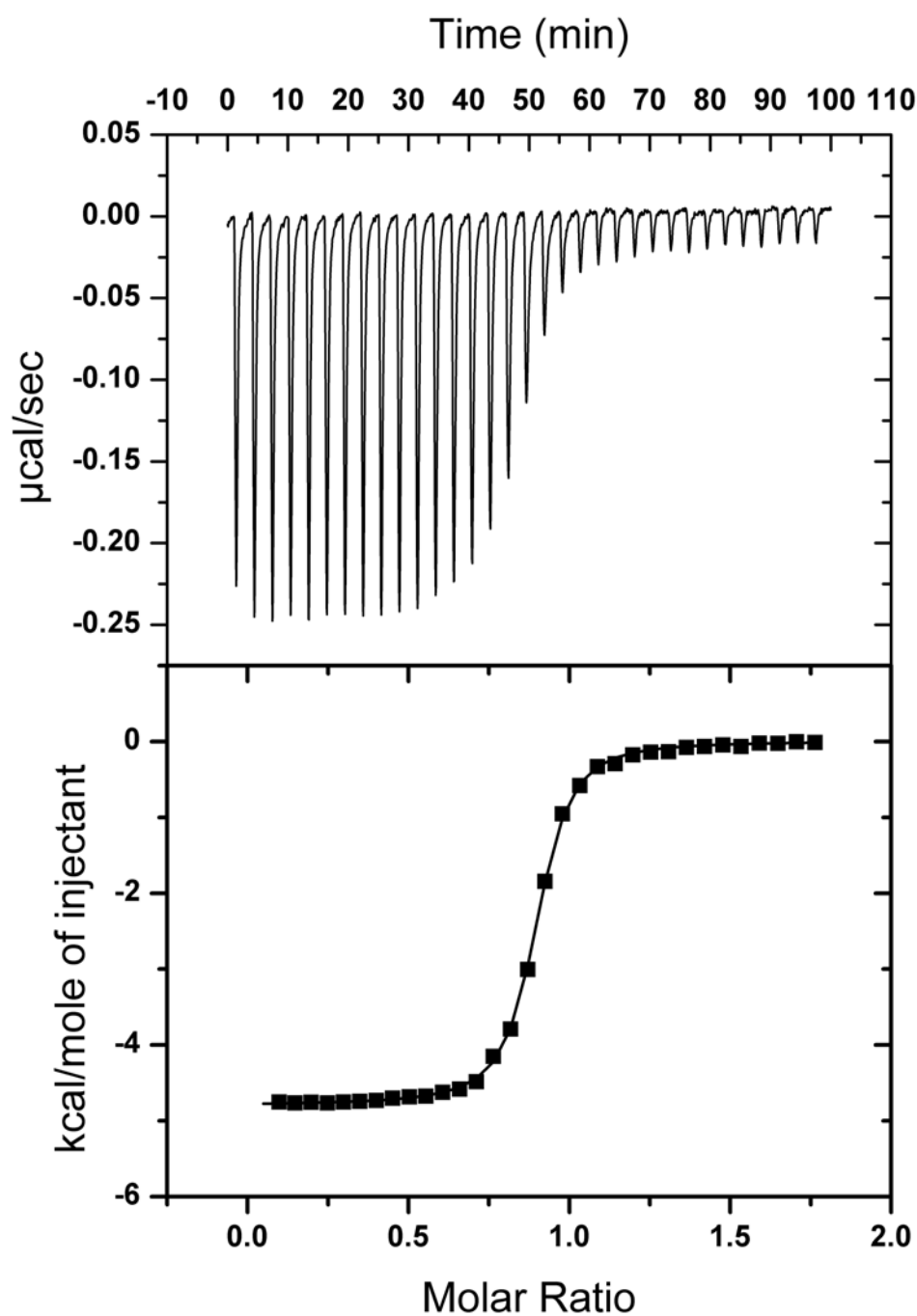
Figure 4.7. Typical titration profile of a GST-fusion extended GPR peptide. Titration of GST-GPR2ex with $G\alpha_{i1}$ is shown. Aliquots (8 µl) of 215 µM $G\alpha_{i1}$ were injected into an ITC cell containing 1.4 ml of 25 µM of $G\alpha_{i1}$. Nonlinear least squares fit using the "single set of identical sites" model yielded a dissociation constant (K_d) of 0.07 µM, enthalpy (ΔH) of -4.49 kcal/mol and stoichiometry (N) of 0.88.

The extended GPR peptides bound with 10-50-fold higher affinity to $G\alpha_{i1}$·GDP than did the smaller peptides that contain only few (approximately three) flanking residues in addition to the core GoLoco consensus motif (Table 4.1 and 4.2). Binding constants for these extended peptides were within the range measured for AGS3 itself, with GST-GPR2ex having the strongest affinity of 70 nM (Kd). Interestingly, much like GPR2, which had the lowest enthalpy of binding among the GPR peptides, GPR2ex also had a low enthalpy of binding, and binding was to a great extent driven by entropy. The greater affinity of GPR1ex relative to GPR1 was also marked by a large increase in the entropy of binding. In contrast, the binding enthalpy of GPR3ex was substantially greater than that of GPR3. The rank order of extended GPR peptides in terms of binding affinity was: GPR2ex>GPR1ex~GPR3ex The stronger binding affinity of the extended GPR peptides suggest that residues outside of the GoLoco motifs can make significant interactions with $G\alpha_{i1}$ and contribute towards binding affinity. GST by itself did not show any interaction with $G\alpha_{i1}$ in control experiments (data not shown), suggesting that the presence of the GST tag has no effect on the binding free energy of the extended GPR peptides.

The affinity of any one of the extended GPR peptides for $G\alpha_{i1}$ was comparable to that of the high (GPR2ex) or low (GPR1ex and GPR3ex) affinity sites in AGS3-C. Even the GPR consensus peptide, which is longer by a few residues at both termini than that of the other four GPR peptides, had a greater free energy of binding to $G\alpha_{i1}$. The dissociation constants derived for the binding of AGS3-C to $G\alpha_{i1}$ may therefore simply represent the distribution of affinities of the individual extended GPR motifs. Nevertheless the possibility that the tertiary organization of the $G\alpha_{i1}$ binding motifs in AGS3-C influences their affinity for $G\alpha_{i1}$ cannot

be dismissed. While it is clear that residues outside of the GoLoco motif contribute binding energy toward $G\alpha_{i1}$, it is not apparent whether the N- or C-terminal residues are equally important. The structure of the complex between $G\alpha_{i1}$ and the GoLoco motif of RGS14 reveals a substantial interface between the helical domain of $G\alpha_{i1}$ and the residues that extend from the C terminus of the GoLoco motif (Kimple et al., 2002). In contrast, the N-terminal boundary of the GoLoco motif almost extends beyond the surface of the Ras-like domain of $G\alpha_{i1}$. Hence it is probable that it is the C-terminal flanking residues of GPRex peptides that contribute the additional binding energy toward $G\alpha_{i1}$. It is remarkable that these residues are not conserved among the four GPR repeats in AGS3-C (Figure 4.1 B). The structures of the complexes between extended peptides and $G\alpha_{i1}$ should reveal the structural basis for their contribution to binding.

4.3.4   Role of C-terminal residues of GoLoco motifs on GDI activity.

The ability of the extended GPR peptides to inhibit nucleotide exchange was also tested by monitoring the rate of the binding of BODIPY-GTPγS to $G\alpha_{i1}$. All of three extended GPR peptides exhibited GDI activity. Surprisingly, in contrast to GPR3, GST-GPR3ex was a potent GDI toward $G\alpha_{i1}$ (Figure 4.8) with an apparent IC_{50} in the micromolar range. The addition of GPR-flanking residues also increased the potency of GPR2 ~10-fold but had less effect upon the activity of GPR1. The rank order of the extended peptides with respect to GDI activity (GPR3ex >=GPR2ex > GPR1ex) was almost the reverse of that for the core GPR peptides. GST itself had no affinity or GDI activity toward $G\alpha_{i1}$ (data not shown).
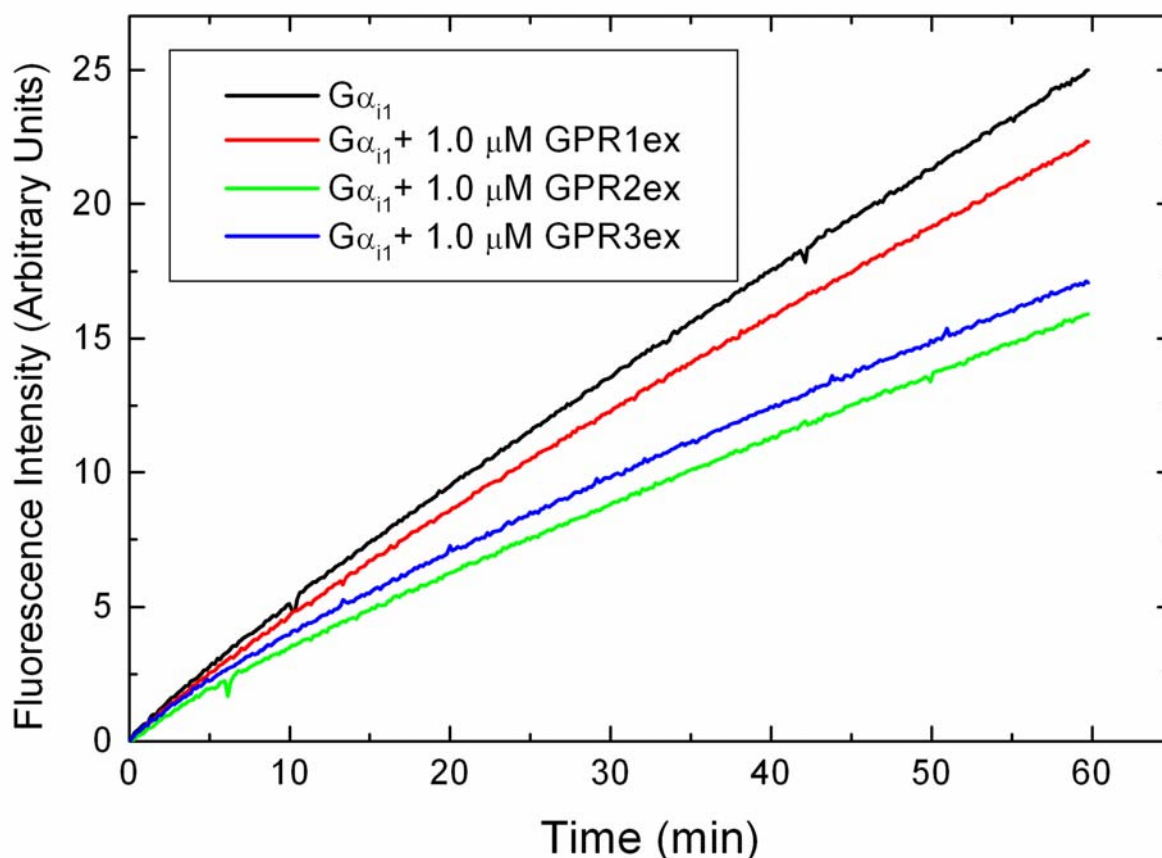
Figure 4.8: Inhibition of BODIPY-GTPγS binding to Gα$_{i1}$ by GST-fusion extended GPR peptides. Comparison of the time course of BODIPY-FL-GTPγS binding to Gα$_{i1}$ in the presence of various extended GPR peptides. Gα$_{i1}$ (200 nM) was incubated with 1 μM of GPR1ex (*red*), GPR2ex (*green*) and GPR3ex (*blue*). The pre-incubated solutions were mixed with 1 μM BODIPY-FL-GTPγS and fluorescence intensity was monitored at 510 nm. The *black* curve represents the binding of BODIPY-FL-GTPγS with Gα$_{i1}$ in the absence of any peptide.

Residues that flank the GPR/GoLoco motif not only confer specificity, as demonstrated for RGS14 (Kimple et al., 2002), but may be critical for GDI activity as the data indicated for GPR3 of AGS3-C. The factors that might render GPR3 ineffective as a GDI was discussed earlier (Chapter 4.3.2). Whether these or other factors account for the

inactivity of GPR3, residues that flank this motif are compensatory to the extent that GPR3ex

exhibits potent GDI activity.

4.4   References

Bernard, M. L., Peterson, Y. K., Chung, P., Jourdan, J., and Lanier, S. M. (2001). Selective Interaction of AGS3 with G-proteins and the Influence of AGS3 on the Activation State of G-proteins. J Biol Chem *276*, 1585-1593.

Higashijima, T., Ferguson, K. M., Sternweis, P. C., Ross, E. M., Smigel, M. D., and Gilman, A. G. (1987). The effect of activating ligands on the intrinsic fluorescence of guanine nucleotide-binding regulatory proteins. J Biol Chem *262*, 752-756.

Kimple, R. J., Kimple, M. E., Betts, L., Sondek, J., and Siderovski, D. P. (2002). Structural determinants for GoLoco-induced inhibition of nucleotide release by Galpha subunits. Nature *416*, 878-881.

Peterson, Y. K., Bernard, M. L., Ma, H., Hazard, S., 3rd, Graber, S. G., and Lanier, S. M. (2000). Stabilization of the GDP-bound conformation of Gialpha by a peptide derived from the G-protein regulatory motif of AGS3. J Biol Chem *275*, 33193-33196.

Peterson, Y. K., Hazard, S., 3rd, Graber, S. G., and Lanier, S. M. (2002). Identification of structural features in the G-protein regulatory motif required for regulation of heterotrimeric G-proteins. J Biol Chem *277*, 6767-6770.

Chapter Five

Crystallization of $G\alpha_{i1}$:R6A4 complex, X-ray diffraction data collection and optimization of crystals

5.1   Introduction

The structure of $G\alpha_{i1}\cdot GDP$:R6A4 complex was determined in order to elucidate the mechanism by which R6A4 inhibits nucleotide exchange. The peptide R6A4 identified by William Ja in Dr. Richard Robert's group in Caltech is a GDI (Ja WW, 2004). Although R6A4 shares functional similarity with GoLoco motifs, R6A4 is shorter in sequence, lacks the conserved and functionally important arginine residue of the GoLoco motif (Peterson et al., 2002; Willard et al., 2004), and also lacks the N-terminal helical region present in GoLoco motifs (Kimple et al., 2002). Because of the differences of R6A4 peptide with GoLoco motif, the structure of $G\alpha_{i1}$ in complex with R6A4 might allow the identification of structural elements in $G\alpha_{i1}$ that are important for interaction with GDIs.

$G\alpha_{i1}\cdot GDP$:R6A4 crystals diffract only to 3.2 Å. Although this low resolution data was good enough for unambiguous determination of space group and obtaining a molecular replacement solution for the complex; the poor quality of the resulting electron density map precluded the determination of an unambiguous structure. Several strategies were adopted to improve the diffraction quality of $G\alpha_{i1}\cdot GDP$:R6A4 crystals, which included both modification of the crystallization conditions and also mutagenesis of the peptide to promote

alternate crystal forms. In this chapter, I will discuss the crystallization, diffraction data collection, and analysis of the diffraction data of $G\alpha_{i1} \cdot GDP:R6A4$ crystals. In addition, I will also discuss the strategies used to obtain better quality diffraction data.

5.2   Materials and methods

5.2.1   Synthesis of a soluble R6A peptide

R6A4, a 16-residue peptide with the sequence of SQTKRLDDQLYWWEYL, was synthesized in the Protein Chemistry Technology Core facility at the UT Southwestern Medical Center by solid-phase Fmoc chemistry. The C-terminus of the peptide was amidated. Crude peptides were purified using reverse-phase HPLC and the mass of the purified peptide was verified by electro-spray mass spectrometry.

5.2.2   Modification of R6A peptide

Modification of the crystallization target is a standard approach in protein crystallography to obtain crystals or modifying existing crystals. Although R6A4 was identified as a 16 residue peptide, it is known that the C-terminal 9 residues of R6A4 comprising the sequence DQLYWWEYL is sufficient for binding to $G\alpha_{i1}$ (Ja WW, 2004). This 9-residue C-terminal segment forms a core motif for interacting with $G\alpha$ subunits. However, the 9-residue core consists of several large hydrophobic residues, which makes it insoluble in water. The shortest soluble and biologically active peptide is a 13 residue peptide, which lacks the three most N-terminal residues of R6A4. The 13-residue version of

R6A4 (hereby referred to as R6M2) and the 9-residue C-terminal core with two lysine

residues at the C-terminus (hereby referred to as R6M1) to enhance solubility, were chosen

as the best candidate peptides for crystallization trials. Both, R6M1 and R6M2 (Figure 5.1)

were synthesized in the Protein Chemistry Technology Core facility at the UT Southwestern

Medical Center. The peptides were purified using HPLC and their purity was verified using

electro-spray mass spectrometry.

```
R6A4 (16):    SQTKRLDDQLYWWEYL
R6M1 (11):           DQLYWWEYLKK
R6M2 (13):        KRLDDQLYWWEYL
```

Figure 5.1. Modification of R6A4. Sequence alignment of R6A4 with two modified
versions of R6A4 referred to as R6M1 and R6M2. The conserved 9 residue core region is
shown in red and the additional lysine residues added to enhance solubility is shown in blue.

5.2.3   Formation and crystallization of the $G\alpha_{i1}$/ΔN-$G\alpha_{i1}$:R6A4 complex

$G\alpha_{i1}$ and ΔN-$G\alpha_{i1}$ was purified to a high-level of homogeneity as described in chapter

2. The complex of $G\alpha_{i1}$ and ΔN-$G\alpha_{i1}$ with R6A4 was formed by addition of a 1.5 fold molar

excess of R6A4 to $G\alpha_{i1}$, followed by the incubation of the mixture for 1-2 hr in ice. The

protein was stored in a buffer containing 50 mM EPPS pH 8.0, 1 mM EDTA, 3 mM DTT,

and 10 μM GDP. Crystallization screening was performed using 10-15 mg/ml of

$G\alpha_{i1}$·GDP:R6A4 complex. Crystals were obtained by hanging-drop vapor diffusion by

mixing 1-2 μl of complex solution with an equal volume of reservoir solution containing 1.8-

2.2 M ammonium sulfite and 0.1 M sodium acetate (pH-5.5-6.5) at 20°C. Large, hexagonal-

rod like crystals were obtained in 1-2 days. The crystal dimensions varied in the range of 0.3-1.2 mm X 0.1-0.4 mm X 0.1-0.4 mm.

Crystallization screen was also performed with $G\alpha_{i1}$:R6M1/R6M2 complexes. Preparations of the complex and crystallization conditions were similar to what was followed for $G\alpha_{i1}$:R6A4 complex.

The use of additives in the initial crystallization conditions has often proven to be successful in obtaining better crystals or crystals in alternate forms. Useful additives are usually identified through randomized screening approaches or based on prior knowledge of ligands and protein stability. Standard additive screens are commercially available from Hampton Research; three additive screens available from Hampton Research were used for screening purposes. In addition, a salt based customized additive screen available from Nextal Biotechnology was also used.

5.2.4 Stabilization and freezing of crystals

The crystals were stable in the mother liquor over a long period of time (3-30 days). Crystals were transferred from the crystallization drop to mother liquor containing 2.0 M ammonium sulfite and 0.1 sodium acetate pH 6.0. Cryoprotectant solution contained 20-25% glycerol in addition to the components of the mother liquor. Crystals were directly transferred from mother liquor to the cryoprotectant solution using cryo-loops and frozen in liquid nitrogen. The crystals were either allowed a brief exposure to the cryoprotectant or soaked for 15 min-1hr. For the purposes of screening, crystals were often frozen directly

under the liquid nitrogen vapor stream at the X-ray beam or stored in liquid nitrogen dewars for long term storage.

5.2.5 Screening of crystals and data collection using synchrotron X-ray source

Native crystals of $G\alpha_{i1}\cdot GDP$:R6A4 complex were screened either using the in-house x-ray facility, which is equipped with an RU-H3R (Rigaku) rotating anode X-ray generator (wavelength 1.54 Å) and a R-Axis IV detector, or at the Advanced Light Source (ALS) synchrotron beamline 8.2.1 (Lawrence Berkeley National Laboratory, Berkeley, CA). The beamline at ALS was equipped with ADSC Q210 2x2 CCD detector. Several crystals were screened both at the in-house facility and at the ALS. Data collected at ALS was used for crystallographic analysis. The wavelength was 1.00 Å (12.398 keV) and the crystal-to-detector distance was 250 mm. Ninety frames of diffraction data were measured with 1° oscillation and exposure time of 2 s/frame.

5.2.6 Indexing, integration and scaling of data

The first data frame was used for auto-indexing using the DENZO component of the HKL2000 suite of programs (Otwinowski, 1993; Otwinowski and Minor, 1997). The program DENZO allows the determination of crystal unit cell parameters from a diffraction image that can then be used to identify, which of the 14 possible Bravais lattice correctly describes the crystal. The initial crystal and detector parameters are refined in DENZO (Otwinowski and Minor, 1997). Upon the completion of data collection, all the frames were integrated and scaled using the HKL2000 suite of programs. The integration component of

DENZO predicts diffraction spots for every frame based on geometric parameters, uses 'profile fitting' to identify the collected spots, and integrates the diffraction maxima of the spots according to their Miller indices (h, k, and l) (Otwinowski and Minor, 1997). Integration is followed by scaling using the SCALEPAK component of the HKL2000 program. Scaling is used for finding relative scale factor between measurements, refining the whole data set using crystal parameters, and merging of reflections based on space-group symmetry. During the data collection process many symmetry-related reflections are recorded independently. In principle these symmetry-related reflections should be equal in intensity, but differences arise due to radiation damage to the crystals, x-ray absorption, and error in integration. Scaling applies correction factors for the errors and scales symmetry-related reflections to their average value.

The statistical parameters that reflect the quality of the data are also computed during scaling. The overall differences between symmetry-related reflections are quantified by the factor R_{sym}, defined as:

$$R_{sym} = \frac{\sum_{hkl} \sum_i \left| I_i(hkl) - \overline{I(hkl)} \right|}{\sum_{hkl} \sum_i I_i(hkl)}$$

Scaling also estimates the signal to noise ratio (I/σ) and completeness of the data set. It can be also used to merge data sets collected from different crystals by putting them in an uniform scale. Because, scaling is based on crystal symmetry, it can be also used for the identification of the correct space-group in many situations.

The diffraction data for $G\alpha_{i1}$:R6A4 crystal was indexed as a hexagonal lattice and can be scaled in the hexagonal point-group P622. To identify the possible presence of screw-axis, the data was scaled in five-different space-groups within the P622 point group: $P6_122$, $P6_222$, $P6_322$, $P6_422$, and $P6_522$. The correct space-group was determined at a later stage by using molecular replacement programs and observing systematic absences.

Scaled intensities obtained from the diffraction data was used for calculating structure factor amplitudes. The temperature factor or Wilson B-factor was estimated using the Wilson plot. The intensity of a particular reflection depends on the scattering angle (θ) as well as a constant (B) related to the temperature dependent disorder in the system according to the following equation:

$$\overline{I(S)} = K \exp\left[-2B \frac{\sin^2 \theta}{\lambda^2} \right] \sum_i (f_i^o)^2$$

Where, K is scale factor, B is temperature/Wilson factor, I(S) is intensity, and f_i^0 is the atomic scattering factor at rest.

5.2.7 The "phase problem" of crystallography and determination of phase using molecular replacement

The basic goal of macromolecular crystallography is to determine the structure of the macromolecule from the diffraction patterns of its crystals. Mathematically, three-dimensional structure and the diffraction pattern are related by the 'Fourier' transformation. The crystallographic structure factor is a contribution of individual atomic structure factors and the position of those atoms in space.

The structure factor corresponding to a particular reflection is calculated as:

$$F(h,k,l) = \sum_{j=1}^{N} f_j . e^{2\pi i (h.x_j + k.y_j + l.z_j)} \qquad (1)$$

Where, N is the total number of atoms; x_j, y_j, and z_j are the coordinates of the jth atom; f_j is the atomic scattering factor for the jth atom; h, k, and l are 'Miller indices'; and the structure factor vector F(h,k,l) is the structure factor corresponding to the reflection with Miller indices of h, k, and l.

Electron density can be calculated from the structure factors as:

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F(h,k,l) e^{-2\pi i (h.x + k.y + l.z)} \qquad (2)$$

Where, V is the volume of the unit cell and $\rho(x,y,z)$ is electron density at coordinates x,y,z.

The diffraction data collection equipment allows only the measurement of the intensity of the spot which is related to the structure factor as:

$$I(h,k,l) = |F(h,k,l)|^2$$

Thus, the amplitude of the structure factor can be computed from the measured intensities but the phase information associated with the structure factor vectors are lost. This is commonly referred to as the 'phase problem' in crystallography.

Solving the 'phase problem' is one of the most important aspects of protein structure determination. There are three major ways of determining structures: 1. Molecular replacement; 2. Multiple isomorphous replacement (MIR); 3. Multi-wavelength anomalous dispersion (MAD). Besides these, data obtained from single isomorphous replacement (SIR)

and single-wavelength anomalous dispersion (SAD) techniques can provide phasing information; these techniques are particularly valuable when combined with other techniques or when high resolution diffraction data can be collected. Apart from these, there are 'direct methods' of phase determination, but these methods are limited to molecules of relatively smaller size and requires very high resolution diffraction data. Direct methods are impractical for most protein molecules.

Molecular replacement is the most straightforward of structure determination when the structure of a homologous molecule is known. The technique of molecular replacement is extensively reviewed in textbooks (Rossman and Arnold, 2001) and several online resources (http://www-structmed.cimr.cam.ac.uk/Course/MolRep/molrep.html#concepts and http://www.ocms.ox.ac.uk/mirrored/xplor/manual/htmlman/node327.html). Unlike other phasing methods, such as MIR or MAD, molecular replacement does not require heavy-atom derivatives or collection of multiple data sets. Molecular replacement can be performed when the structure of a homologous molecule or part of the molecule is known. For molecular replacement to succeed, the difference in atomic position between the search model and the crystal structure has to be within 1 Å (Brunger, 1997), and the sequence identity should be above 50% (Brunger, 1997).

The goal of the molecular replacement procedure is to position the search model in such a manner that its position overlaps with the unknown structure in the crystals. The first step in this process involves orienting the search model in the unknown crystal by use of rotation functions. This is followed by finding the location of the search-model in the unit cell using translation functions (Figure 5.2). This reduces the problem of orientating the
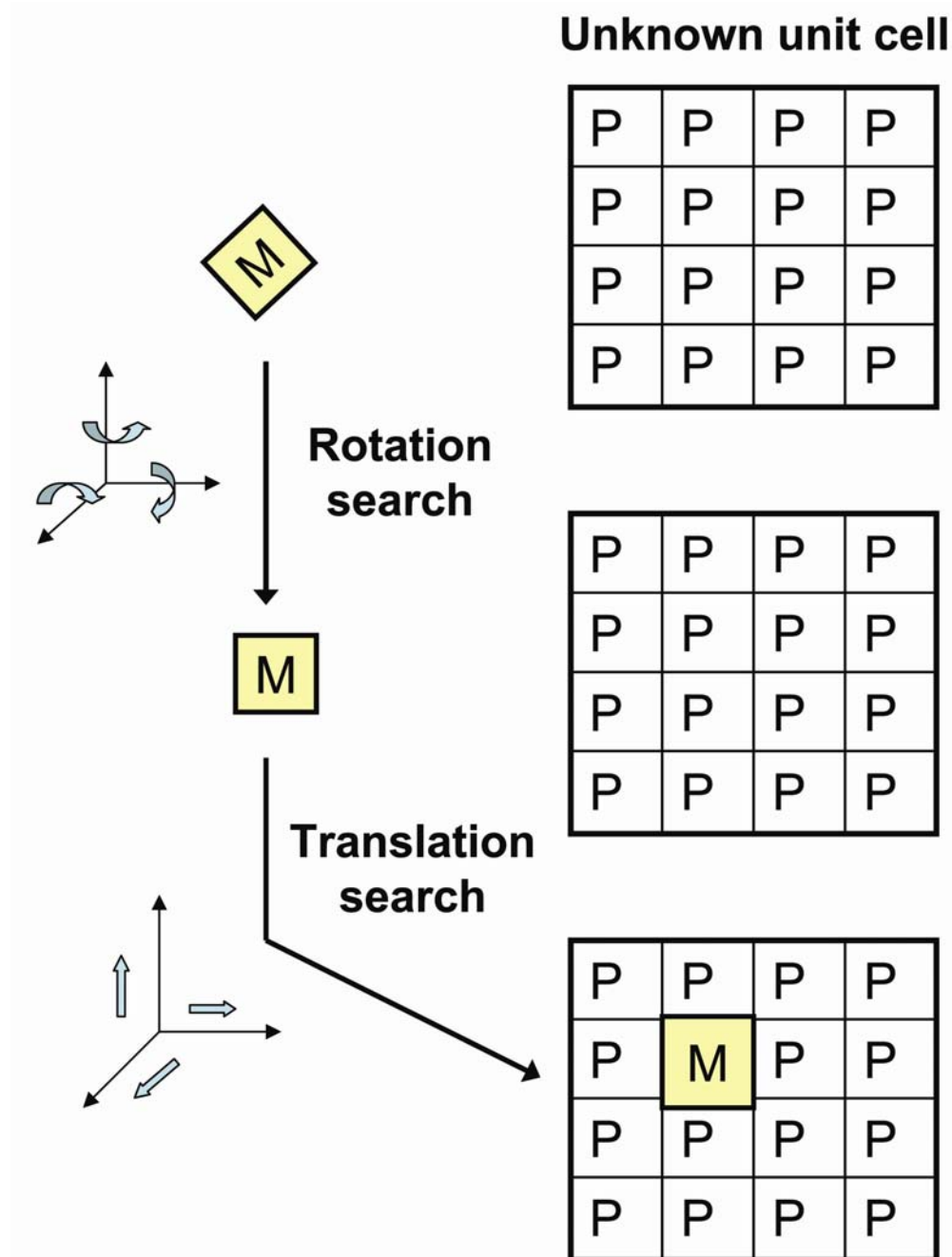
Figure 5.2. Schematic illustration of molecular replacement search. Molecular replacement search is performed for protein P in the unknown unit cell using M as search model. The proper orientation of M is determined through rotation search and location of P in the unit cell is determined through translation search.

search-model on the unknown structure to a six dimensional problem: three rotational and three translational degrees of freedom (Rossmann, 1972; Rossmann and Blow, 1962). Molecular replacement rotation and translation searches are performed using Patterson maps, which are Fourier maps calculated with square of the structure factor amplitude, such that each peaks in the map correspond to inter-atomic vectors. Rotation search algorithms are aimed at maximizing the overlap between the patterson map of the unknown structure and that of the search model (Brunger, 1997). The search model is effectively oriented along all three possible rotational axes until a best overlap with the unknown structure is found (Figure 5.2). Once the orientation of the search model is defined, the search model is translated along three possible axes to position it in the proper place in the unit cell (Figure 5.2). The quality of rotation and translation search results is monitored by various correlation coefficients.

Molecular replacement procedures are implemented in several commonly used crystallographic computing packages. Because of differences in their algorithm, certain programs perform better in particular situations. Typically, alternate programs are tried when a solution is not obtained using one of the programs. The CNS suite (Brunger et al., 1998) of programs use a 'direct rotation' function (DeLano and Brunger, 1995) and a 'fast translation' function (Navaza and Vernoslova, 1995). The CCP4 suite of programs (Bailey, 1994) include a molecular replacement program called AMoRe (automated molecular replacement) (Navaza, 1994; Navaza, 2001). Unlike the rotation search in CNS, AMoRe uses a 'fast rotation function', which uses spherical harmonics to approximate the patterson function (Navaza, 1993). The use of the 'fast rotation function' makes AMoRe considerably faster than the CNS rotation search. Another popular program for molecular replacement includes

the recently developed program PHASER. PHASER uses 'maximum likelihood, based fast

rotation and fast translation functions (McCoy et al., 2005; Storoni et al., 2004). The concept

of 'likelihood' is based on the observation, that given a model, certain data are more probable

than the other data. Likelihood is the probability that the data would have been measured,

given a model. The aim of maximum likelihood estimation is to find the parameter values,

that makes the observed data most likely (McCoy, 2004). In the case of molecular

replacement, data is typically the structure factor amplitudes, the model is the search model,

and the parameters are related to orientations. PHASER is a relatively new program and a

detailed review of PHASER performance is not available yet, but anecdotal evidence

suggests that PHASER is highly successful in obtaining molecular replacement solution in

many difficult cases. The success of PHASER in obtaining a solution is monitored by two

different parameters: the Z-score and the log-likelihood gain (LLG). Z-score is defined as the

number of standard deviations over the mean, and LLG refers to the difference between the

likelihood of the model and the likelihood with a model in which atoms are randomly

distributed (Wilson distribution). Detailed discussion of the program PHASER is available at

the website of the PHASER software. (http://www-

structmed.cimr.cam.ac.uk/phaser/documentation/phaser-1.3.html.)

5.2.8 Molecular replacement using AMoRE/PHASER and determination of space group

The diffraction data obtained from $G\alpha_{i1}$·GDP:R6A4 crystals were scaled in the point-

group P622. The structure of $G\alpha_{i1}$·GDP (PDB code: 1GDD) was used as a search model.

Initial molecular replacement attempts in CNS with the space-group P622 failed to provide any solution. To determine the correct space group, a faster program AMoRe was used; molecular replacement searches were performed in different space groups: $P6_122$, $P6_222$, $P6_322$, $P6_422$, and $P6_522$. The correct space group was identified by comparing the translation search solution in different space groups. Molecular replacement search was also performed using PHASER in all possible hexagonal space groups to check if the best PHASER solution is in agreement with the AMoRe solution.

5.2.9   Choosing an appropriate search model for molecular replacement

Although molecular replacement is an extremely powerful tool to obtain phase information using a related structure, results of molecular replacement often suffer from the problem of 'model bias'. The structure of $G\alpha_{i1}$ has been solved in different states including the structure of $G\alpha_{i1} \cdot GDP$ (PDB code: 1GDD), $G\alpha_{i1} \cdot GTP\gamma S$ (PDB code: 1GIA), $G\alpha_{i1}$: RGS14 GoLoco (PDB code: 1KJY) and $G\alpha_{i1}$:$\beta\gamma$ (PDB code: 1GP2). To avoid problems associated with 'model bias', all of these structures were used as search models for molecular replacement and their outcome was compared to obtain the best search-model. Only the $G\alpha_{i1}$ portion of these structures was included in the search model; bound nucleotides and other proteins in the complex were removed. All molecular replacement calculations were performed using the program PHASER.

5.2.10 Model refinement, map calculation and model building

The correct molecular replacement solution was used to compute trial phases. The

initial model was refined by one cycle of rigid-body refinement using the program CNS. This

was followed by one cycle of simulated annealing and energy minimization using the CNS

package (Brünger, 1992; Brünger et al., 1998). The refined model was used to calculate

electron density maps using CNS. Sigma-A weighted difference fourier maps (Read, 1986)

were calculated and the resulting maps were inspected using the program O (Jones et al.,

1991).

The agreement of the structure factor amplitudes in the model and observed data was

monitored using R and R_{free} factors. The crystallographic R factor is defined as:

$$R = \frac{\sum_{hkl} \left\| F_{obs} \right| - \left| F_{calc} \right\|}{\sum_{hkl} F_{obs}}$$

Where, F_{obs} and F_{calc} are observed and calculated structure factor amplitudes.

R_{free} was calculated with the above equation with a randomly selected test set of

reflections consisting of 10% of the data.

5.2.11 Screening for new crystallization conditions

Since the crystals were obtained from a salt-based condition, various other salts were

also used to evaluate their viability as precipitants. The 'Salt Rx' screen, available from

Hampton Research, which is a collection of 96 different salt based conditions, was used for

screening. Ammonium sulfate was also used for extensive screening of crystallization

conditions. Apart from more focused salt-based screening, randomized crystal screens were

also performed to obtain alternate crystal forms. Commercially available crystals screens like the Hampton Crystal Screen 1, Wizard screen I and II (Emerald Biosciences Inc.), and sodium malonate grid screen (Hampton Research Inc.) were used for screening. In addition to these, high-throughput micro-batch crystallization screening under oil was also performed at 1536 different conditions using the high-throughput crystallization facility at the Hauptman Woodward Institute (Buffalo, NY)

5.3 Results and discussion

5.3.1 Crystallization of the Gα_{i1}·GDP:R6A4 complex and data collection

Crystals of Gα_{i1}·GDP in complex with R6A4 were obtained by using ammonium sulfite as precipitant. The crystallization condition was similar to that used to obtain crystals of Gα_{i1}·GDP itself (Coleman et al., 1994). However, the crystal morphology was different; Gα_{i1}·GDP:R6A4 crystals were hexagonal in cross-section and were rod-like in shape with tapered ends (Figure 5.3). The crystals had hexagonal lattice in contrast to the tetragonal lattice observed for Gα_{i1}·GDP crystals (Coleman et al., 1994). The crystals grew very large within 1-2 days; the largest crystals were more than 1 mm in the longest dimension. Most crystals were 0.4-1.0 mm in the longest dimension and 0.1-0.4 mm in the other dimensions.

Although the crystals of Gα_{i1}·GDP:R6A4 were good in appearance, the quality of diffraction obtained from these crystals was poor. For most crystals, diffraction was observed to a maximum resolution of 4-7 Å, when screened in the in-house X-ray facility. To check if cryoprotection conditions were affecting resolution, crystals were frozen with different

cryoprotectants (glycerol, ethylene glycol, and PEG 400) and soaked for various lengths of time in the cryoprotectant. The quality of diffraction was not improved by changing the cryoprotectants, but when crystals were soaked in the cryoprotectant for longer period of time, the resolution improved. Soaking the crystals for 15 min to 1h improved diffraction allowing the collection of 3.5-4.0 Å data. To obtain higher resolution data, several crystals were screened at the ALS 8.2.1 synchrotron beamline. The diffraction quality was improved at the ALS and diffraction to 3.5 Å resolution was observed for most crystals (Figure 5.4). The best crystal allowed the collection of data set that was complete up to 3.2 Å resolution with good scaling statistics (Table 5.1).
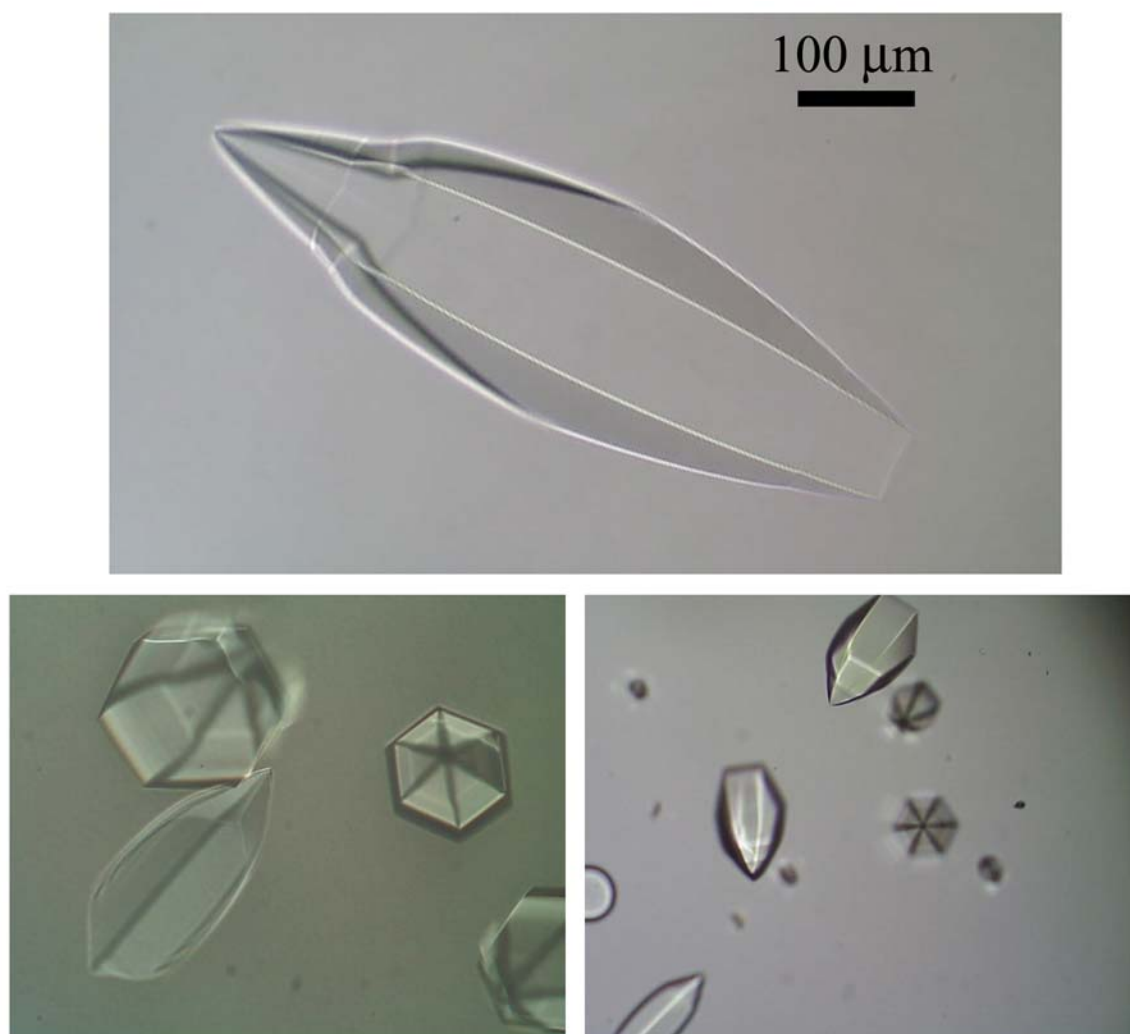
Figure 5.3. Crystals of Gα$_{i1}$·GDP:R6A4 complex. Upper panel. A large single crystal of Gα$_{i1}$·GDP:R6A4 complex grown under 2.0 M ammonium sulfite and 0.1 M sodium acetate, pH 6.0. Lower panel. Several crystals in crystallization drops are shown and the hexagonal cross section of the crystals appears as dark lines.
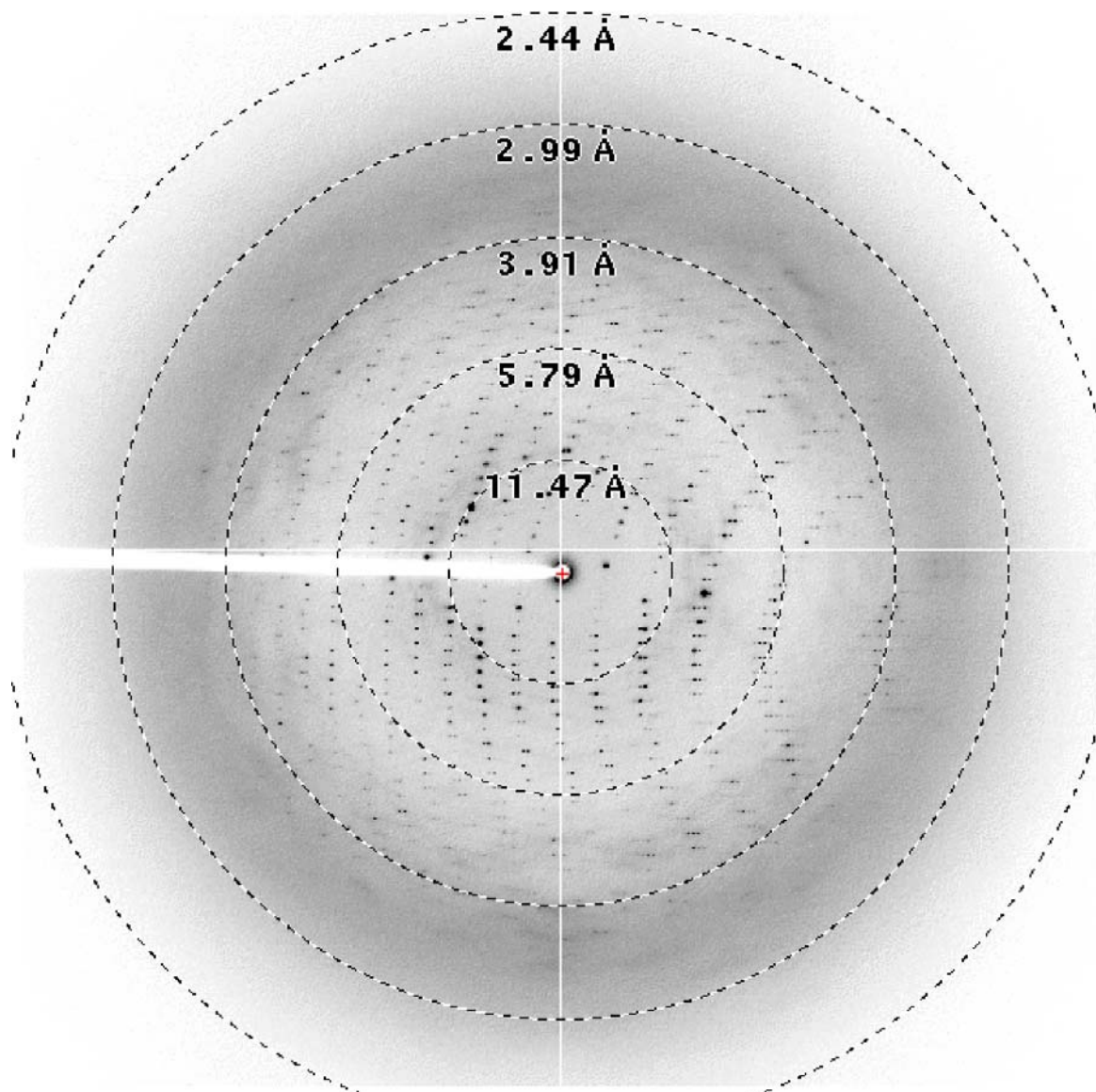
Figure 5.4. Typical diffraction image collected from Gα$_{i1}$·GDP:R6A4 crystals. A typical diffraction image collected at ALS 8.2.1 with 2s exposure time. Resolution circles are shown (dashed) and their values are shown in Å.

	$G\alpha_{i1}$:R6A4
Wavelength (λ)	1.00 Å
X-ray	ALS 8.2.1
a, b, c (Å)	103.611, 103.611, 243.253
α, β, γ (°)	90.0, 90.0, 120.0
Resolution (Å)	50.0-3.2 (3.31-3.2) *
R_{sym}[1] (%)	6.1 (57.3)
I / σI	39.2 (4.0)
Completeness (%)	99.8 (100)
Redundancy	10.4 (4.0)
Reflections:	
Measured	139,743
Unique	13,494
Wilson B-factor	95.34

Table 5.1. Crystallographic data collection statistics.

* Numbers in parentheses correspond to the highest resolution shell, 3.31-3.2 Å
[1] Defined in the crystallographic terms and abbreviations section

5.3.2   Molecular replacement and identification of the correct space group

In the absence of high resolution diffraction data, the 3.2 Å resolution data collected at ALS was used for structure determination. Initial phases were obtained by molecular replacement using the structure of $G\alpha_{i1}$·GDP as a search model. When molecular replacement was performed in CNS using the space group P622, the rotation search was able to identify an unambiguous solution; the correlation coefficient for the best solution (0.064) was almost twice that of the second best (0.038) solution. However, the result of the translation search was ambiguous with no distinct solution; the standard deviation of the scoring function for several solutions was ~10 % of the mean value. The failure of the translation search indicated incorrect identification of the space group. Systematic absences observed after scaling could not identify the correct space group unambiguously. To identify the correct space group, all space groups belonging to the point group P622 such as $P6_122$, $P6_222$, $P6_322$, $P6_422$, and $P6_522$ was chosen as space groups and molecular replacement search was performed using the program AMoRe. The result of molecular replacement with AMoRe indicated $P6_522$ to be the correct space group. The top rotation search result in AMoRe allowed the identification of a distinct translation search result (Table 5.2), when $P6_522$ was used as space group. The result of AMoRe was also verified using the program PHASER. PHASER search was performed with all possible hexagonal space groups, and correct solution was obtained in the $P6_522$ space group. The Z-score for this solution was 26.9 (typically >8 represents a definite solution) and log-likelihood gain (LLG) was 710, indicating a distinct solution. Expected systematic absences were also observed in the $P6_522$ space group (Table 5.3)

	α	β	γ	CC_F	CC_P
1	23.5	30.92	173.24	13.9	18.8
2	8.92	9.89	37.87	12.7	13.6
3	11.22	59.54	119.33	12.5	10.8
4	39.67	77.3	76.19	12.1	13.8
5	56.94	18.46	287.33	11.9	10.4

	α	β	γ	Tx	Ty	Tz	CC_F
1	23.5	30.92	173.24	0.25	0.589	0.0174	25.3
2	23.5	30.92	173.24	0.2493	0.5889	0.0525	19.9
3	23.5	30.92	173.24	0.2482	0.5856	0.4819	19.5
4	23.5	30.92	173.24	0.2481	0.5855	0.1674	19.4

Table 5.2. Result of molecular replacement search using AMoRe in the space group P6$_5$22. The rotation and translation search solutions with highest correlation coefficients are highlighted in grey. *Upper panel*, Top five rotation search solutions are shown. Rotation angles are α, β, and γ; CC_F is the linear correlation coefficient between observed and calculated amplitudes; CC_P is the Patterson correlation coefficient. *Lower panel*, Top four translation search solutions obtained using the first rotation solution are shown. Tx, Ty, and Tz are translations along X, Y, and Z axis, respectively. The CC_F of the first translation search solution is distinctively higher than the other three.

h	k	l	Intensity	Sigma	I/Sigma
0	0	20	0.3	4.6	0.1
0	0	21	-4.5	5.3	-0.9
0	0	22	4.4	5.9	0.7
0	0	23	-5.1	10.8	-0.5
0	0	25	7.8	8	1
0	0	26	-0.6	6.1	-0.1
0	0	27	-3.2	7.7	-0.4
0	0	28	-13.5	7.9	-1.7
0	0	31	22.5	14.1	1.6
0	0	32	-23.2	9.3	-2.5
0	0	33	-11.2	8.1	-1.4
0	0	34	-0.8	8.5	-0.1
0	0	35	-14.5	8.8	-1.7
0	0	37	-16.3	10.3	-1.6
0	0	38	-1.2	8.3	-0.1
0	0	39	17.3	9.1	1.9
0	0	40	8.6	11.4	0.8
0	0	41	-15.8	9.9	-1.6
0	0	43	-33.4	11.1	-3
0	0	44	-38.2	10.5	-3.6
0	0	45	-39	15.1	-2.6
0	0	46	-31.5	13.7	-2.3
0	0	47	-1.4	25	-0.1
0	0	50	-86.4	17.6	-4.9
0	0	51	-33.3	13.5	-2.5
0	0	52	-5.3	16.3	-0.3
0	0	53	-32.8	13.6	-2.4
0	0	55	-28.9	16.9	-1.7
0	0	56	-32.1	16	-2
0	0	57	10.6	17.6	0.6
0	0	58	-39.8	16.6	-2.4
0	0	59	-29.9	19	-1.6
0	0	61	-69.7	17	-4.1
0	0	62	-24.4	20.5	-1.2
0	0	63	-24.2	21.1	-1.1
0	0	64	40	24.1	1.7
0	0	65	-50.9	17.8	-2.9
0	0	67	-3	21.8	-0.1
0	0	68	-24.6	22.6	-1.1
0	0	69	-27.3	25.9	-1.1
0	0	70	23.1	22.1	1
0	0	71	-12.3	28.5	-0.4
0	0	73	-17.7	24.6	-0.7
0	0	74	-37.9	27.5	-1.4
0	0	75	18.3	27	0.7
0	0	76	124.8	33.7	3.7

Table 5.3. Intensity of systematic absences. Intensity and intensity/sigma for systematic absences in the P$6_5$22 space group. The condition for systematic absences in this space group is (0,0,l) where l=6n+1, 2, 3, 4, 5.

5.3.3 High solvent content in $G\alpha_{i1} \cdot GDP:R6A4$ crystals

The solvent content of protein crystals can be estimated using the Matthew's

coefficient (Matthews, 1968). Matthew's coefficient (V_m), which is the ratio of the volume of

the unit cell and the molecular weight of total number of protein molecules in the unit cell, is

defined as:

$$V_m = \frac{V_{cell}}{M_r \times Z}$$

Where, V_{cell} is the volume of the unit cell; M_r is the molecular weight of the protein; and Z is

the number of protein monomers in the unit cell. V_m is inversely related to the solvent

content of the unit cell. For most proteins V_m varies between 2-3.5 Å^3/Da. Molecular

replacement solution of $G\alpha_{i1}.GDP:R6A4$ crystals suggested the presence of only one

molecule in the asymmetric unit, which corresponds to a V_m of 4.45 or 72.36% solvent

content (Table 5.4), which is unusually high for protein crystals. Although the V_m for

$G\alpha_{i1}.GDP:R6A4$ crystals were within normal range when two molecules were present in the

asymmetric unit, molecular replacement clearly indicated the presence of only one molecule.

Higher solvent content often means poor crystal packing and higher thermal disorder, which

results in poor resolution. Indeed, the thermal disorder of $G\alpha_{i1}.GDP:R6A4$ crystals were very

high. The Wilson B-factor, which is an estimate of the thermal disorder in crystals, was 95.34

for $G\alpha_{i1}.GDP:R6A4$ crystals, indicating significant disorder. This resulted in the steep drop-

off of diffraction intensity with increasing resolution. Efforts to reduce the solvent content by

dehydration were not successful in improving the diffraction quality of the crystals.

N	Probability	Vm ($\mathring{A}^3$/Da)	Solvent content (%)
1	0.04	4.45	72.36
2	0.96	2.23	44.72
3	0.01	1.48	17.08

Table 5.4. Solvent content of Gα_{i1}·GDP:R6A4 crystals. The Matthew's coefficient (V$_m$) and solvent content of Gα_{i1}·GDP:R6A4 crystals are shown with different number (N) of molecules in the asymmetric unit. The probability of the occurrence of a certain V$_m$ is calculated based on the V$_m$ of known structures in the Protein Data Bank.

5.3.4 Identification of the R6A4 binding region

Initial phases calculated using molecular replacement was used to obtain a model for the Gα_{i1}.GDP:R6A4 complex. The initial model was refined through one cycle of rigid-body refinement, simulated annealing, and coordinate minimization. The resulting model was used for calculation of electron density maps. The overall quality of the electron density map allowed unambiguous identification of most regions of Gα_{i1} and the bound GDP. The N-terminal region of Gα_{i1} spanning from residue G2-A31 was disordered in the structure and no electron density was observed. Visual inspection of the electron density maps suggested ordered structure in the switch II region and additional electron density for the peptide around the switch II region. But the map quality in the peptide binding region was poor and did not allow model building. To obtain a better model for the conformation of the switch II region, the structure of Gα_{i1} in different states were used as molecular replacement search models. Of these, the structures of Gα_{i1}·GDP:GoLoco complex appeared to be the best model in terms of the value of R and R$_{free}$ parameters after refinement (Table 5.5). Although this allowed identification of the approximate orientation of the switch II region, the putative electron density of the peptide appeared discontinuous with poor features for side chains. The

	Z-Score	LLG	R-factor (%)	R-free (%)
1GDD (Gα_{i1}.GDP)	26.97	710	37.3	42.3
1KJY (Gα_{i1}.GDP.RGS14 GoLoco)	25.26	552	33.1	38.0
1GIA (Gα_{i1}.GTPγS)	26.14	594	35.7	40.8
1GP2 (Gα_{i1}.GDP.$\beta\gamma$)	22.88	228	37.7	43.0

Table 5.5. Molecular replacement using different search models. The PDB code and the state of different search models used for molecular replacement is shown, along with the Z-score and log-likelihood gain (LLG) estimated by PHASER. The value of R and R_{free} after one cycle of refinement is shown; and the lowest values are highlighted in the table in grey.
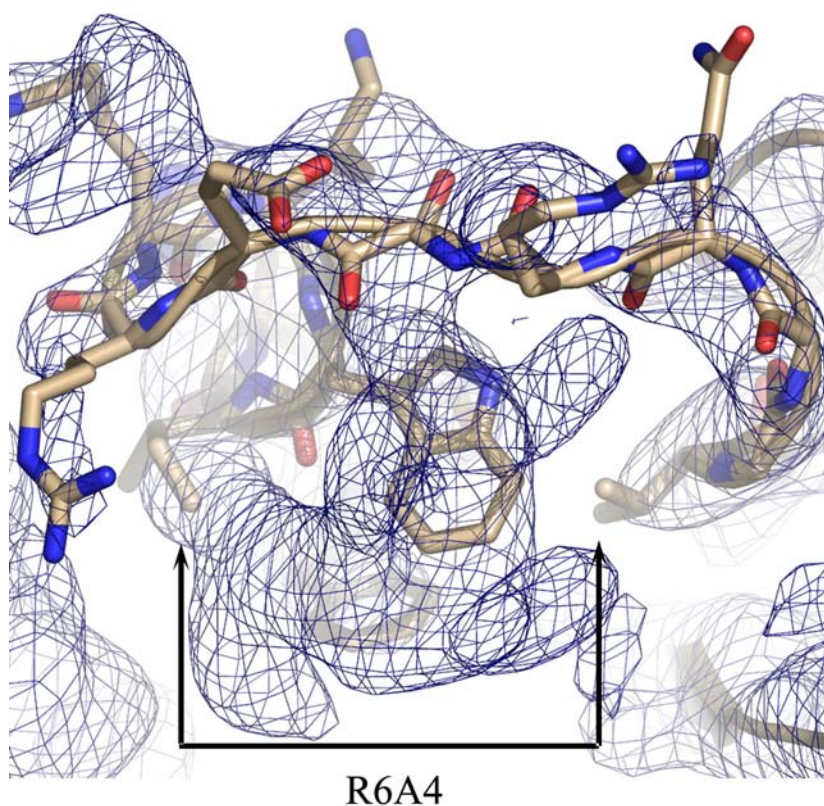
R6A4

Figure 5.5. Electron density of R6A4 peptide. SigmaA-weighted $2F_o$-F_c electron density map contoured at 1.2 σ.calculated from the molecular replacement solution obtained with the structure of GoLoco bound Gα_{i1}·GDP as search model. Electron density map is displayed around the switch II region and the putative R6A4 electron density is indicated.

R6A4 peptide main-chain was partly built through the electron density but an unambiguous model for the peptide could not be built (Figure 5.5).

Irrespective of the poor quality of the electron density maps, analysis of the crystals confirmed the presence of bound R6A4 in the crystals, allowed the mapping of the overall binding region of R6A4, and indicated that switch II adopts an ordered structure. Detailed structural characterization of R6A4 and $G\alpha_{i1}$·GDP interactions would require higher resolution diffraction data.

5.3.5   Result of crystallization screens

To promote the crystallization of $G\alpha_{i1}$.GDP:R6A4 complex in alternate crystal forms, additive screens were performed. Although crystals were obtained with several additives, none of these crystals resulted in better diffraction quality. All of the crystals retained the original unit cell dimensions and diffraction characteristics. In the absence of good additives, crystallization screens were performed to obtain new crystallization conditions. The salt-based screen Salt Rx provided several new crystallization conditions; crystals were obtained with sodium formate, sodium citrate, sodium/potassium tartarate, and phosphate based conditions (Figure 5.6). The tartarate-based crystals were later found to be salt crystals. Even though several crystals were obtained with new precipitants, all of these crystals were hexagonal, similar to the crystals obtained with ammonium sulfite, and did not exhibit better diffraction qualities.
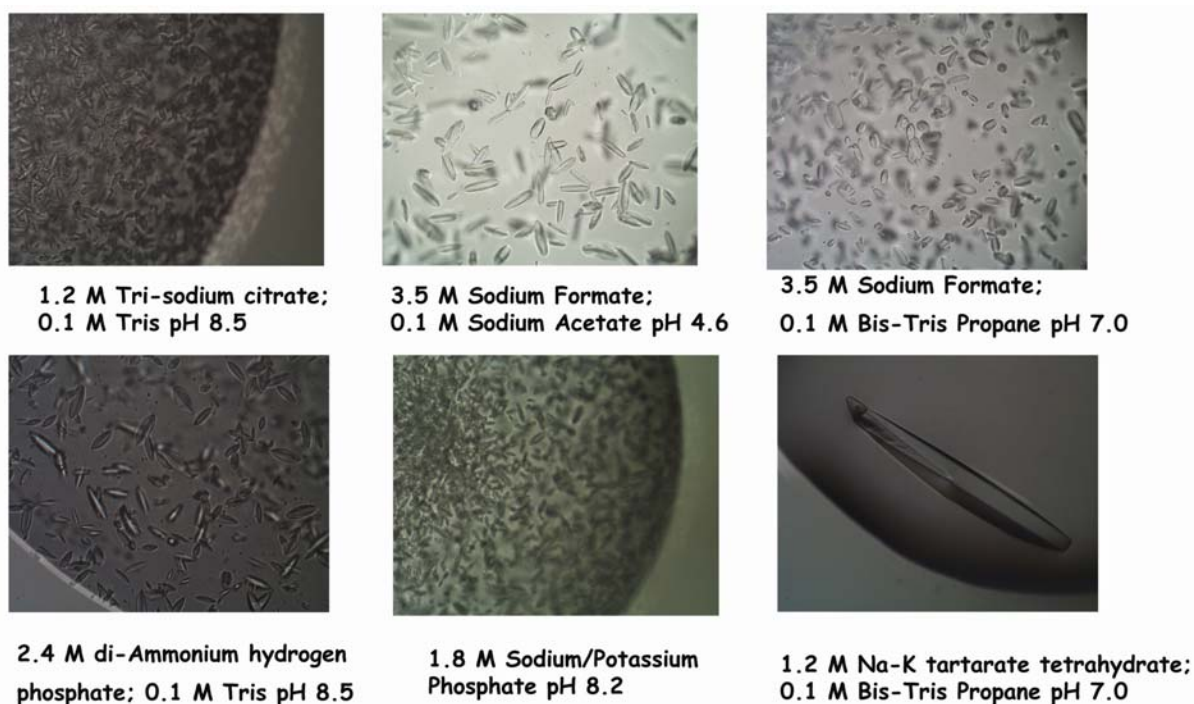
Figure 5.6. New crystallization conditions. Crystallization of $G\alpha_{i1}$·GDP:R6A4 complex in various precipitant and pH conditions available from the Salt Rx screen.

Extensive crystallization screening with commercially available screens did not result in new crystal forms. This suggested that the $G\alpha_{i1}$.GDP:R6A4 complex inherently preferred the hexagonal crystal form and altering this would require modification of either $G\alpha_{i1}$ or the R6A4 peptide. Alternately, the peptide can be soaked into $G\alpha_{i1}$.GDP crystals. However, $G\alpha_{i1}$.GDP crystals fractured after 15 min of soaking in solution containing the peptide. The soaking conditions were not optimized as other approaches turned out to be more promising (discussed in chapter 6).

5.3.6 Crystallization of Gα_{i1} in complex with modified R6A4 peptides

Two different variants of the R6A4 peptides were used for additional crystallization screening. These modified peptides were also active as GDIs and inhibited the binding of BODIPY-GTPγS to Gα_{i1} (Figure 5.7). The R6M1 peptide, containing the 9 residue core region and two lysine residues, failed to crystallize in complex with Gα_{i1}. R6M2, which lacked the first three N-terminal residues of R6A4, crystallized in complex with Gα_{i1} under conditions that were identical to the crystallization condition of the Gα_{i1}·GDP:R6A4 complex. This indicated that residues N-terminal to the core region of R6A4 were important in making crystal contacts in the hexagonal crystal form. The best data collected using the Gα_{i1}·GDP:R6M2 crystals went only up to 3.6 Å in resolution.
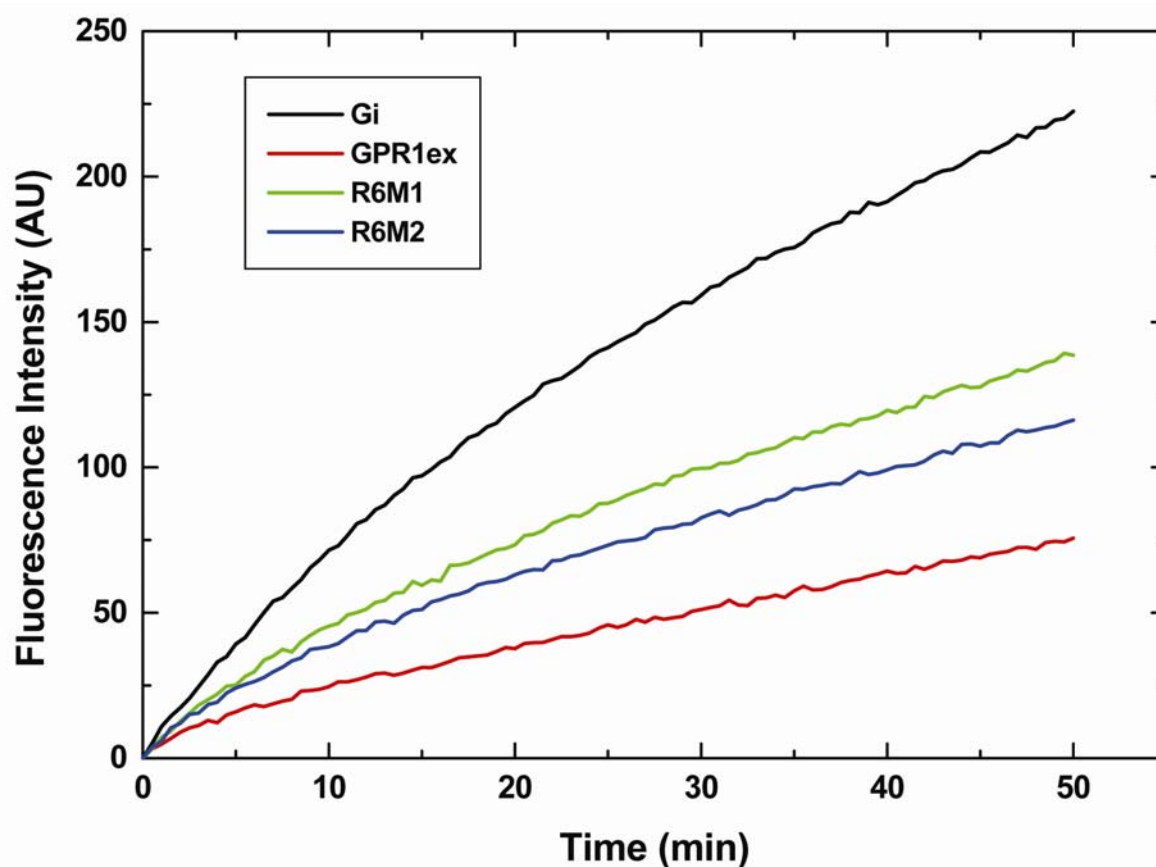
Figure 5.7. Inhibition of BODIPY-GTPγS binding to Gα$_{i1}$ by modified R6A4 peptides.
Typical time course experiment showing the binding of 1 μM **BODIPY-GTPγS** to 200 nM
Gα$_{i1}$ (black) and binding to 200 nM Gα$_{i1}$ in the presence of 10 mM R6M1 (green), R6M2
(blue), and GPR1ex (red). GPR1ex is a peptide corresponding to the most N-terminal
GoLoco motif of AGS3. Assays were performed as described in chapter 4.2.5.

5.4   References

Bailey, S. (1994). The CCP4 suite- programs for protein crystallography. Acta Cryst *D50*, 760-763.

Brunger, A. T. (1997). Patterson Correlation Searches and Refinement. Methods in Enzymology *276*, 558-580.

Brünger, A. T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472-475.

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S.*, et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr *54*, 905-921.

Brünger, A. T., Adams, P. D., Clore, G. M., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J.*, et al.* (1998). Crystallography and NMR system (CNS): A new software suite for macromolecular structure determination. Acta Cryst *D54*, 905-921.

Coleman, D. E., Lee, E., Mixon, M. B., Linder, M. E., Berghuis, A., Gilman, A. G., and Sprang, S. R. (1994). Crystallization and preliminary crystallographic studies of $G_{i\alpha1}$ and mutants of $G_{i\alpha1}$ in the GTP and GDP-bound states. JMolBiol *238*, 630-634.

DeLano, W. L., and Brunger, A. T. (1995). The direct rotation function: Patterson correlation search applied to molecular replacement. Acta Crystallogr D Biol Crystallogr *51*, 740-748.

Ja WW, R. R. (2004). In vitro selection of state-specific peptide modulators of G protein signaling using mRNA display. Biochemistry *43*, 9265-9275.

Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991). Improved methods for the building of protein models in electron density maps and the location of errors in these models. Acta Crystallogr *A47*, 110-119.

Kimple, R. J., Kimple, M. E., Betts, L., Sondek, J., and Siderovski, D. P. (2002). Structural determinants for GoLoco-induced inhibition of nucleotide release by Galpha subunits. Nature *416*, 878-881.

Matthews, B. W. (1968). Solvent content of protein crystals. JMolBiol *33*, 491.

McCoy, A. J. (2004). Liking likelihood. Acta Crystallogr D Biol Crystallogr *60*, 2169-2183.

McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005). Likelihood-enhanced fast translation functions. Acta Crystallogr D Biol Crystallogr *61*, 458-464.

Navaza, J. (1993). On the computation of the fast rotation function. Acta Crystallogr D Biol Crystallogr *49*, 588-591.

Navaza, J. (1994). *AMoRe*: an automated package for molecular replacement. Acta Cryst *A50*, 157-163.

Navaza, J. (2001). Implementation of molecular replacement in AMoRe. Acta Crystallogr D Biol Crystallogr *57*, 1367-1372.

Navaza, J., and Vernoslova, E. (1995). On the fast translation function for molecular replacement. Acta Cryst *A51*, 445-449

Otwinowski, Z. (1993). Oscillation data reduction program, In Data collection and processing, N. I. L. Sawyer, and S. W. Bailey, eds. (Daresbury, U.K.: Science and Engineering Council Daresbury Laboratory), pp. 56-62.

Otwinowski, Z., and Minor, W. (1997). Processing of x-ray diffraction data collected in oscillation mode. Methods in Enzymology *276*, 307-326.

Peterson, Y. K., Hazard, S., 3rd, Graber, S. G., and Lanier, S. M. (2002). Identification of structural features in the G-protein regulatory motif required for regulation of heterotrimeric G-proteins. J Biol Chem *277*, 6767-6770.

Read, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. Acta Crystallographica *A42*, 140-149.

Rossman, M. G., and Arnold, E. (2001). International Tables for Crystallography: Crystallography of Biological Macromolecules, Vol F).

Rossmann, M. G. (1972). The Molecular Replacement Method: A collection of papers on the use of non-crystallographic symmetry. (New York: Gordon and Breach).

Rossmann, M. G., and Blow, D. M. (1962). The detection of sub-units within the crystallographic asymmetric unit. Acta Crystallographica *15*, 24-30.

Storoni, L. C., McCoy, A. J., and Read, R. J. (2004). Likelihood-enhanced fast rotation functions. Acta Crystallogr D Biol Crystallogr *60*, 432-438.

Willard, F. S., Kimple, R. J., and Siderovski, D. P. (2004). Return of the GDI: the GoLoco motif in cell division. Annu Rev Biochem *73*, 925-951.

Chapter Six

Crystal structure of Gα$_{i1}$ in complex with R6A: Identification of the structural determinants of guanine nucleotide exchange in Gα$_{i1}$

6.1   Introduction

The strategies described in Chapter 5 to obtain the high resolution crystal structure of Gα$_{i1}$ in complex with R6A peptides were not successful. The primary causes of the poor diffraction quality can be attributed the high solvent content of the hexagonal crystal form and possibly non-uniform packing interactions that cause long-range disorder in the crystals. Although the crystals of Gα$_{i1}$:R6A4 and Gα$_{i1}$:R6M2 were obtained with different precipitants, all of the crystals had similar morphology, belonged to the same space-group, and exhibited poor diffraction quality. The packing interactions in the hexagonal crystal form were contributed to a large extent by a single tryptophan residue (Trp258) in the Ras-like domain of Gα$_{i1}$ (Figure 6.1). If this residue is mutated to a smaller residue; it could potentially disrupt the crystal packing interactions and promote crystallization in alternate forms. To test this idea, Trp258 of Gα$_{i1}$ was mutated to an alanine residue and crystallization screening was performed with the mutant protein:R6A peptide complexes. This strategy was successful in promoting an alternate crystal form, from which high resolution data was collected and the structure of the (W258A) Gα$_{i1}$·GDP:R6A complex was determined. The structure was determined with a peptide (R6M1, described in chapter 5.2.2) that possessed

two lysine residues at the C-terminus of the DQLYWWEYL-core sequence of R6A.

However, the lysine residues were not observed in the structure. Thus, the peptide will be

simply referred to as R6A throughout this chapter.

In this chapter, I will describe the determination of the structure of the complex,

analysis of the structure, and use the structure to propose models of inhibition of nucleotide

exchange by the peptide.
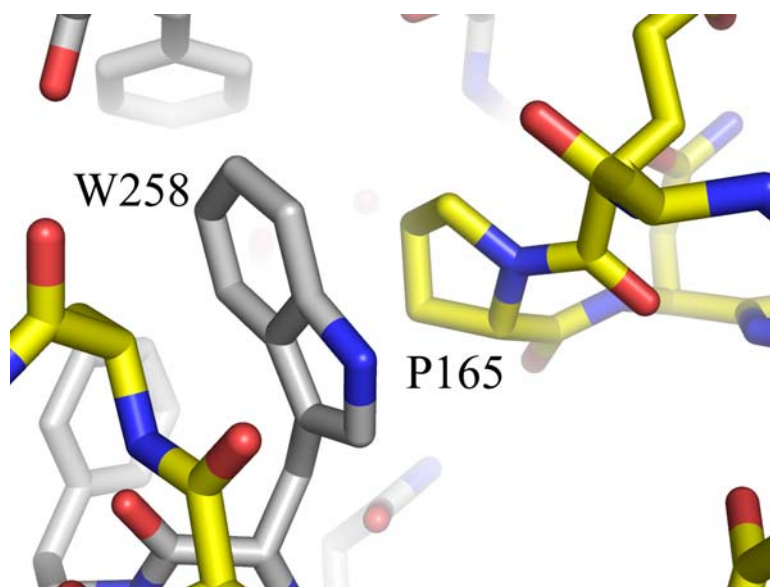
Figure 6.1. Intra-molecular contacts in Gα_{i1}:R6A4 crystals. Stick representation of the crystal contact surface between two Gα_{i1} subunits in the hexagonal (P6522) crystals of Gα_{i1}:R6A4 is shown. The W258 side-chain of one Gα_{i1} (shown in gray) stacks between the P165 ring and peptide-plane (formed between T120 and A121) of a neighboring Gα_{i1} (shown in yellow).

6.2 Materials and methods

6.2.1 Mutagenesis of Gα$_{i1}$-W258 to alanine

An N-terminal GST-tagged expression vector pDEST15, encoding Gα$_{i1}$ residues G2-F354 (with a TEV protease cleavage site between GST and Gα$_{i1}$), was used a template for mutagenesis. Site-directed mutagenesis was performed using the Quick-change™ site-directed mutagenesis kit (Stratagene). The tryptophan codon (TGG) was changed to an alanine codon (GCG) using the following primers:

Forward: 5'-CAT ATG TAA CAA CAA GGC GTT TAC GGA CAC ATC C-3'

Resverse: 5'-GGA TGT GTC CGT AAA CGC CTT GTT GTT ACA TAT G-3'

Site-directed mutagenesis was performed as described in chapter 2.2.3 and the sequence of mutant constructs were verified by DNA sequencing.

The GST-fusion (W258A) Gα$_{i1}$ mutant protein was purified in the same way as wild-type GST-fusion Gα$_{i1}$ (as described in chapter 2.2.7).

6.2.2 Crystallization of (W258A) Gα$_{i1}$ in complex with R6A

(W258A) Gα$_{i1}$:R6A complex was formed by incubating (W258A) Gα$_{i1}$ and R6A together in a molar ratio of 1:1.5 for 1 h in ice. The concentration of the complex was adjusted to 15-18 mg/ml and was used for setting up crystallization experiments without further purification. Initial screening was performed at 20°C with an ammonium sulfite gird

screen based on the crystallization condition of wild-type $G\alpha_{i1} \cdot GDP$ (Coleman et al., 1994). The concentration of ammonium sulfite was varied from 1.7-2.2 M (in 0.1 M increments) and 0.1 M sodium acetate pH 5.0-6.5 (in 0.5 pH unit increments) was used as buffer. Large, single crystals grew in the range of 1.9-2.1 M ammonium sulfite and 0.1 M sodium acetate pH 5.5 and 6.0 in 2-3 days.

6.2.3   Collection of diffraction data

To collect diffraction data, (W258A) $G\alpha_{i1}$:R6A crystals were transferred into a cryoprotectant solution containing 2.1 M ammonium sulfite, 0.1 M sodium acetate pH 6.0, and 25% glycerol. Crystals were soaked for 1-5 min in the cryoprotectant solution and frozen in liquid nitrogen for long term storage. Alternately, crystals were briefly exposed to the cryoprotectant solution and frozen under liquid nitrogen cryo-stream (XTREAM™) in the rotating anode x-ray generator.

Data was collected both at the in-house rotating anode x-ray generator using an R-Axis IV detector and at the ALS 8.2.1 synchrotron beam-line (as described in chapter 5.2.4). 90-100 frames of data were measured with an exposure time of 2 s at the synchrotron or 15 min in the in-house x-ray facility with 1° oscillation angle. The detector distance was maintained at 100 mm with the R-Axis IV detector and 150 mm at the synchrotron.

6.2.4   Processing of diffraction data

Diffraction data collected in-house or at the synchrotron was indexed, integrated, and scaled using the HKL2000 suite of programs (as described in chapter 5.2.5) (Otwinowski, 1993; Otwinowski and Minor, 1997).

6.2.5   Molecular replacement

Initial phases were determined by molecular replacement using the program PHASER (as described in Chapter 5.2.7) (McCoy AJ, 2005). The structure of $G\alpha_{i1}\cdot GDP$ (PDB code: 1GDD) was used as a search model. The model included the coordinate of residues 9-201, 218-233, and 240-354 of $G\alpha_{i1}$. The coordinates of GDP were excluded from the search model.

6.2.6   Refinement of initial model and model building

The molecular replacement solution was used to obtain an initial model. The initial model was refined by one cycle of rigid-body refinement using the program CNS . This was followed by one cycle of simulated annealing and energy minimization using the CNS package (Brünger, 1992; Brünger et al., 1998). The refined model was used to calculate electron density maps using CNS. Sigma-A weighted difference fourier maps (Read, 1986) were calculated and the resulting maps were inspected using the program O (Jones et al., 1991) (Figure 6.2). The part of the electron density map corresponding to the peptide was identified by visual inspection and a model of the peptide was built using the model building features of the program O. Initially, all 11 residues of the peptide were modeled as a poly-

alanine chain. After building of the main-chain, the residues were identified by inspection of the side chain electron densities. Upon identification of the correct side-chains, alanine residues in the poly-alanine model were replaced by the actual peptide sequence. A similar strategy was adopted to rebuild the missing regions of $G\alpha_{i1} \cdot GDP$ structure (particularly the switch II region). The structure of $G\alpha_{i1} \cdot GDP$ was also partly rebuilt to model the conformational changes that occurred due to the binding of the peptide.

The initial model of the peptide conformation was obtained from the 2 Å data collected in the in-house x-ray facility. This model was used later as a molecular replacement search-model for the higher resolution data collected at ALS.
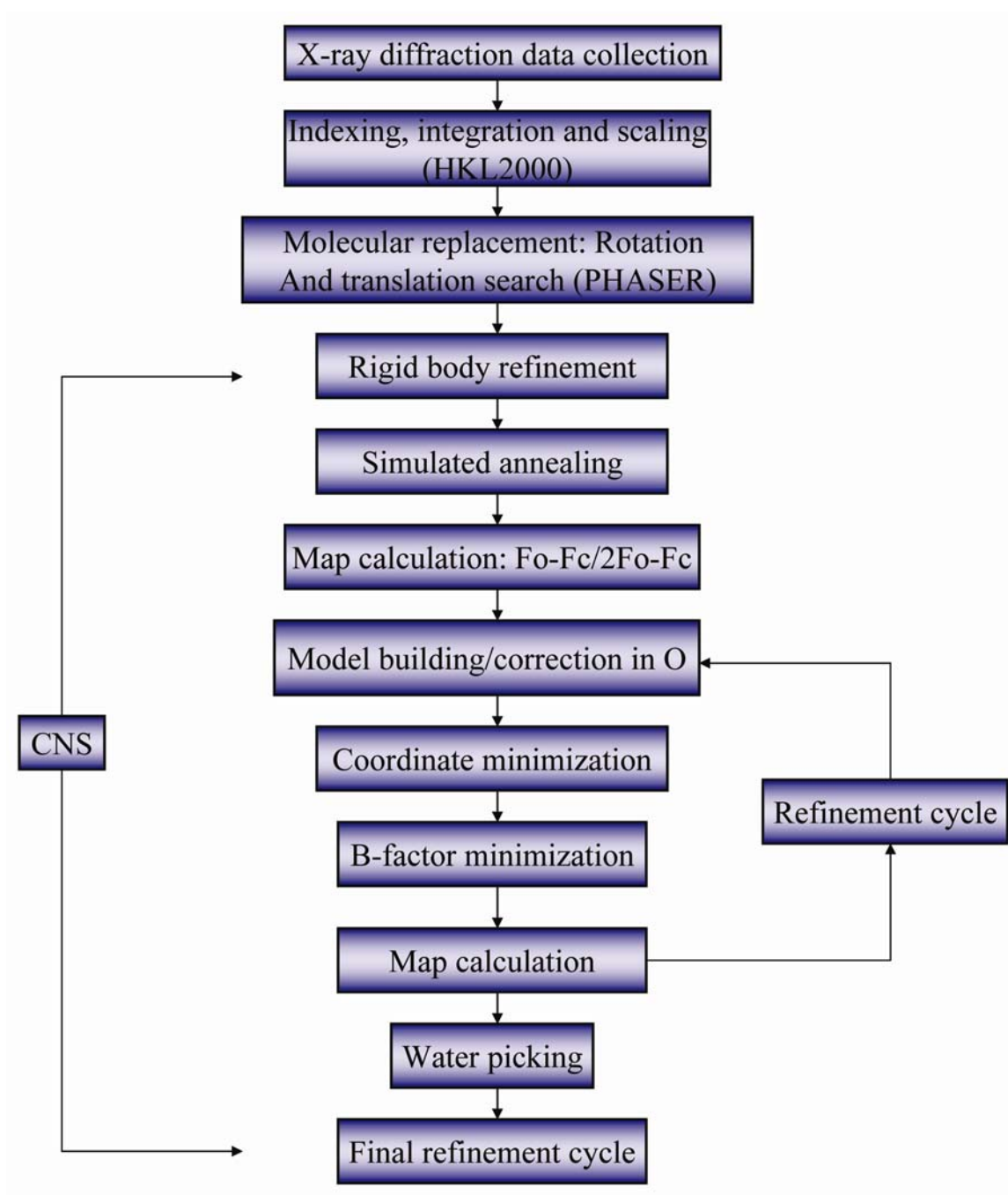
Figure 6.2. Structure determination by molecular replacement. Flowchart showing the methodology of obtaining the crystal structure of (W258A) $G\alpha_{i1}$:R6Acomplex using molecular replacement.

6.2.7   Refinement of model and structure validation

The initial model obtained after building the peptide and $G\alpha_{i1}$ was refined through an iterative cycle of model building, energy minimization, and individual B-factor minimization. All refinements were performed using the program CNS, and after every cycle of refinement the agreement of the model with the electron density was inspected using the program O. The progress of the refinement was monitored by the R and R_{free} values (Brünger, 1992) and also by visual inspection of the map quality. Ordered water molecules were modeled in the electron density map using automated water identification program in CNS and visual inspection of the map. Putative water molecules within hydrogen-bonding distance of at least one protein atom or other water oxygen atom and with refined *B*-factors <100 $Å^2$ were included in the model. Several ordered sulfite ions (precipitant) were also identified in the electron density map and modeled using O. At the final stages of refinement the stereo-chemical weight term (wa) was optimized to obtain lowest R_{free} value while maintaining proper stereo-chemical restraints. The restraints associated with B-factor refinement (r_{weight}) were also optimized at this step.

The quality of the final crystallographic model was assessed using the program PROCHECK (Laskowski et al., 1993), which is part of the CCP4 suite (Bailey, 1994) of programs. PROCHECK was used to validate proper stereochemistry of the proteins by analyzing the distribution of residues in the Ramachandran plot (Ramachandran et al., 1963; Ramachandran and Sassiekharan, 1968), main-chain bond lengths/angles, planarity of aromatic rings, distance of non-bonded interactions etc.

6.2.8. Analysis of the structure

The structure of (W258A) $G\alpha_{i1}$:R6A complex was analyzed by comparing the structure to other structures of $G\alpha_{i1}$ in various states. Structure superpositions were performed using DeepView Swiss-PDB viewer (Guex et al., 1999) (http://au.expasy.org/spdbv/) or PyMol (http://pymol.sourceforge.net/). Secondary structure assignment was verified using the program DSSP (Kabsch and Sander, 1983) available at http://bioweb.pasteur.fr/seqanal/interfaces/dssp-simple.html and also using the program STRIDE (Frishman and Argos, 1995) available at http://bioweb.pasteur.fr/seqanal/interfaces/stride.html. Solvent accessible surface area was computed in CNS using the Lee and Richards method (Lee and Richards, 1971) with a probe radius of 1.4 Å.

6.2.9 Mutagenesis of peptide

Two mutant versions of R6A peptide were synthesized: R6A-G; DQLGWWEYLKK and R6A-T; DQLTWWEYLKK. The peptides were synthesized by solid-phase fmoc synthesis in the Protein Chemistry Technology Core facility (UT Southwestern Medical Center). The peptides were purified by desalting and their molecular masses were verified using electro-spray mass spectrometry.

6.2.10 Nucleotide exchange assay

Nucleotide exchange assays were performed using the procedure described in chapter 3.2.4. Instead of BODIPY-GTPγS, BODIPY-GMPPNP was used for the assays; both of these

have identical absorption and emission characteristics (λ_{ex}: 485 nm, λ_{em}: 510 nm). Nucleotide

exchange assays were also performed using MANT-GMPPNP. The λ_{ex} for MANT was set at

355 nm and the λ_{em} was set at 448 nm. In an alternate approach, 7-azatryptophan (7-AW)

labeled (W258Y) $G\alpha_{i1}$ was used to monitor the change in intrinsic tryptophan fluorescence

during nucleotide exchange. The protein was a gift from Professor Elliott Ross at the

Department of Pharmacology, UT Southwestern Medical Center. The λ_{ex} and λ_{em} for 7-AW

$G\alpha_{i1}$ were set at 285 nm and 390 nm respectively. Nucleotide exchange reaction was initiated

by the addition of 10 μM GTPγS to 1 μM 7-AW $G\alpha_{i1}$ with rapid mixing by stirring. All

assays were performed using a 0.5 ml quartz cuvette; excitation and emission slit-widths

were maintained at 2.5 nm; data collection was performed using a Perkin Elmer LS55

Luminescence Spectrometer. Data analysis was performed using the software Origin 5.0.

6.3   Results and discussion

6.3.1   Crystallization of (W258A) $G\alpha_{i1}$:R6A complex and structure determination

The structure of $G\alpha_{i1}$ in complex with the core region of the R6A4 peptide was

determined using the crystals of the (W258A) $G\alpha_{i1}$:R6A complex. The crystals of wild type

$G\alpha_{i1}$ in complex with R6A4 peptide were of poor quality. These crystals were hexagonal

with the space group P6$_5$22, had large unit cell dimensions, and very high solvent content.

Use of the mutant $G\alpha_{i1}$ allowed crystallization in a different crystal form. Crystals of the

mutant protein in complex with R6A was obtained under crystallization conditions similar to

that of wild-type $G\alpha_{i1}$:R6A4 complex, but the crystals were morphologically different and

belonged to the space group I 4 (Figure 6.3). High resolution diffraction data (Figure 6.4)

were collected using these crystals.

The crystals of the complex were essentially identical to those of wild-type $G\alpha_{i1}\cdot$GDP

with respect to unit cell dimensions (Table 6.1). However, their overall morphology was

different (Coleman et al., 1994). Wild-type $G\alpha_{i1}\cdot$GDP crystals were of tetragonal-rod shape

and grew as long as 1 mm in the longest dimension; but the crystals of (W258A) $G\alpha_{i1}$:R6A

complex had a bi-pyramidal structure with a tetragonal cross-section (Figure 6.3) and grew

only up to 0.3-0.4 mm in the longest dimension. The crystals of the complex also produced

higher resolution diffraction compared to the crystals of wild-type $G\alpha_{i1}$.GDP. The quality of

the data in terms of signal to noise ratio, redundancy, and R_{sym} were excellent (Table 6.1).
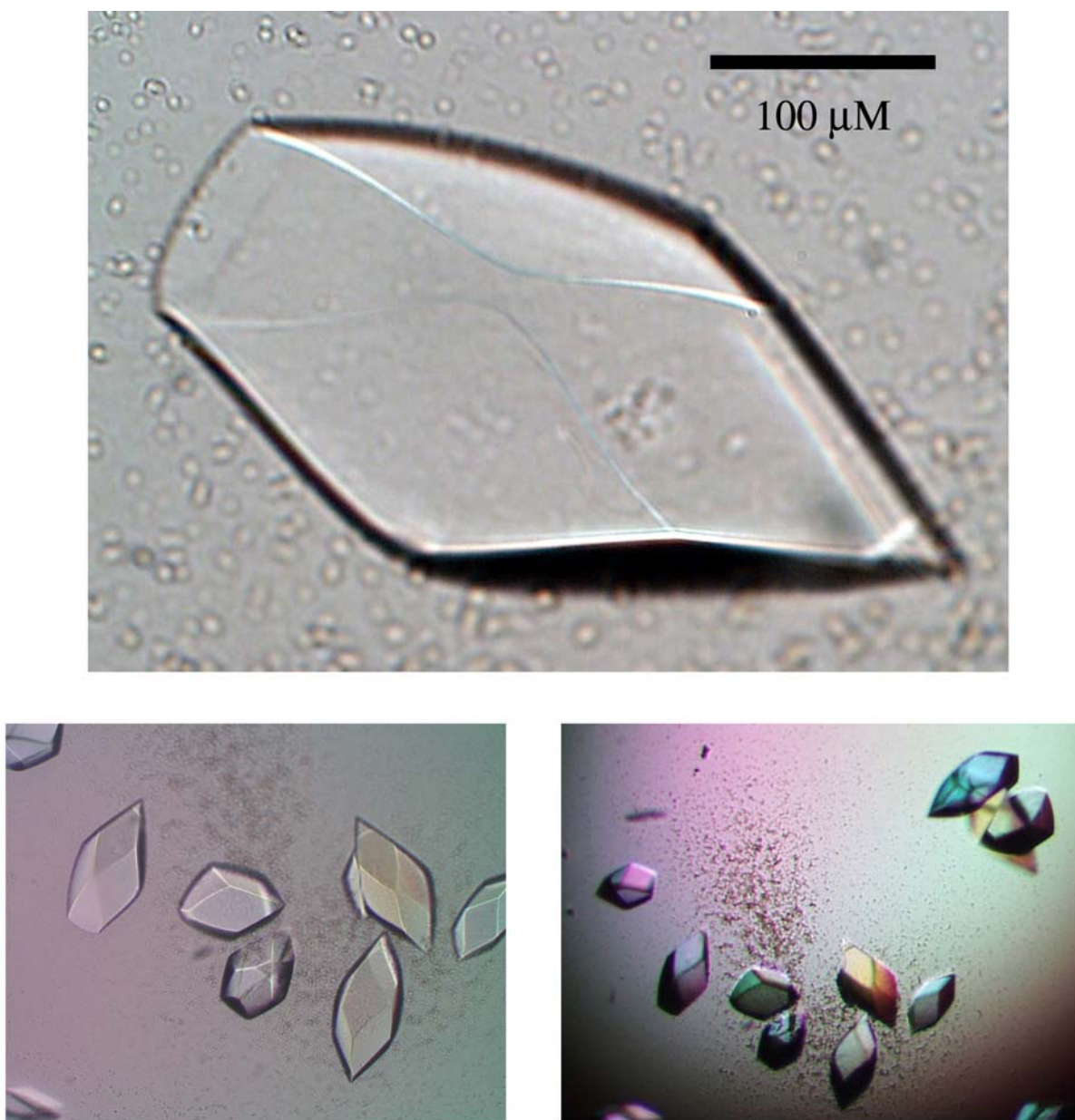
Figure 6.3. Crystals of (W258A) Gα_{i1}:R6A complex. *Upper panel*. A large single crystal of (W258A) Gα_{i1}:R6A complex grown under 1.9 M ammonium sulfite and 0.1 M sodium acetate, pH 6.0. *Lower panel*. Several crystals in a drop (left) and the birefringence of the crystals (right) are shown.
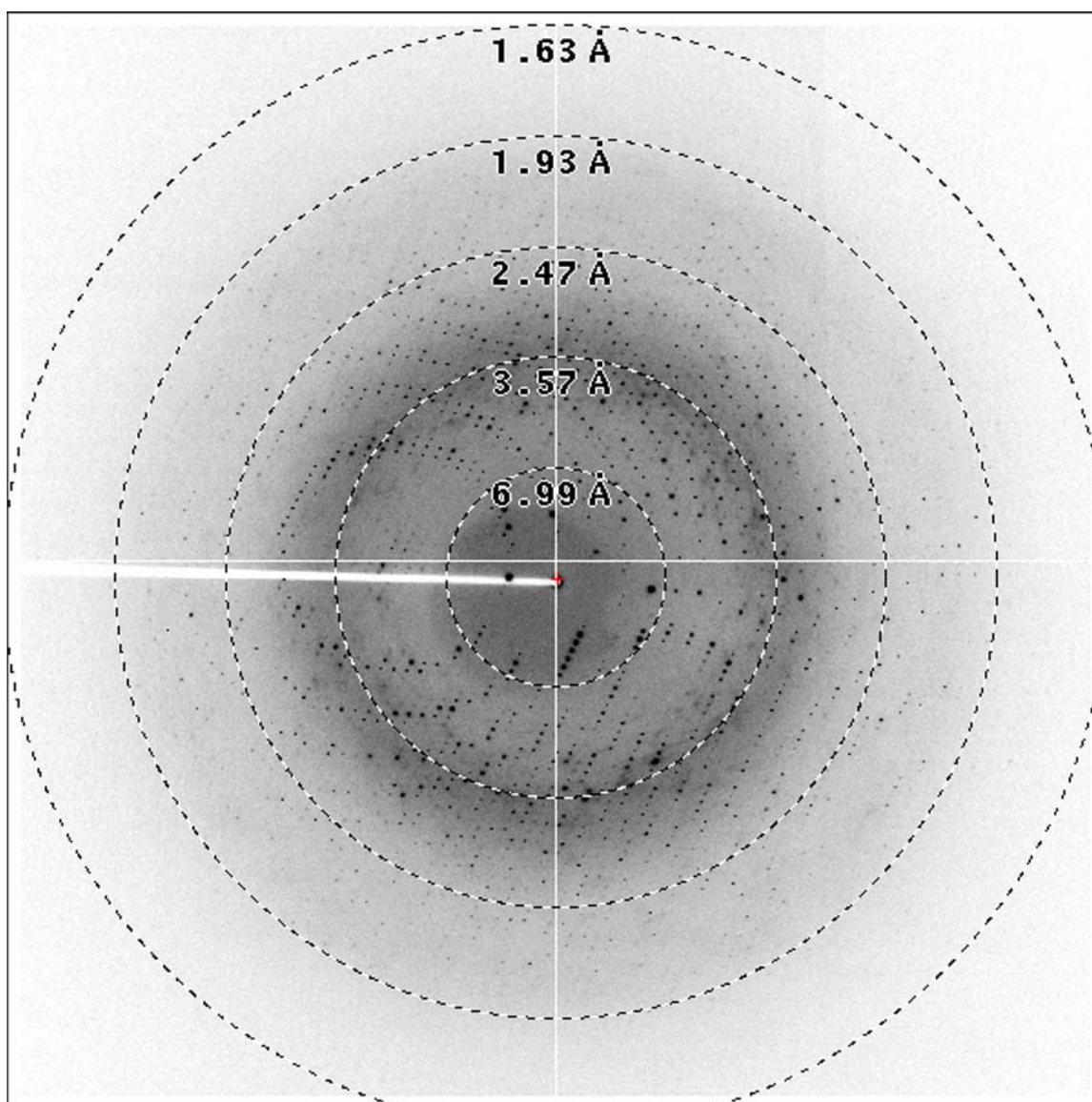
Figure 6.4: A typical high resolution diffraction image. A typical diffraction image collected at ALS 8.2.1 with 2s exposure time. Resolution circles are shown (dashed) and their values are shown in Å.

	(W258A) Gα$_{i1}$:R6A
Data collection	
Wavelength (λ)	1.00 Å
X-ray source	ALS 8.2.1
a, b, c (Å)	120.215, 120.215, 70.019
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	50.0-1.8 (1.86-1.8) *
Linear R-factor[1]	0.036 (0.319)
$I / \sigma I$	37.3 (4.1)
Completeness (%)	99.9 (100)
Redundancy	4.1 (4.0)
Wilson B-factor	25.21
Refinement	
Resolution (Å)	50.0-1.8
No. reflections (working/test)	41589 / 4649
R$_{work}$ / R$_{free}$ (%)[2]	20.85 / 23.39
No. atoms	3206
No. protein atoms	2834
Ligand/ion	1 GDP / 2 sulfite ions
Water	336
B-factor	
Protein	25.99
Ligand (peptide)	31.63
GDP	16.37
Water	37.46
R.m.s deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.321
Ramachandran plot (% in region)	
Most favored	95.3
Allowed	4.7
Generously allowed	0
Disallowed	0

Table 6.1. Crystallographic data collection and refinement statistics.
* Numbers in parentheses correspond to the highest resolution shell, 1.86-1.80 Å
[1,2] These terms are defined in the crystallographic terms and abbreviations section
[2] R_{free} was calculated by selecting 10% of the data.

The three dimensional structure of this complex was determined at 1.8Å resolution using phases derived from molecular replacement (the procedure is described in chapter 5.2.6). The R_{work} and R_{free} (%) values of the final model are 20.85 and 23.39 respectively. The stereochemical properties of the structure including bond lengths, bond angles and distribution of torsion angles (Ramachandran plot) are excellent. The overall value of B-factor for the structure was low, which is expected given the high resolution. The B-factor values for the bound peptide (31 Å^2) were somewhat higher than the average B-factor for the protein (26 Å^2). The final values of the different statistical parameters at the end of refinement are indicated in Table 6.1.

6.3.2 Structure of (W258A) $G\alpha_{i1}$:R6A complex

The high quality of the electron density maps (Figure 6.5) obtained after cycles of model building and refinement allowed unambiguous determination of the structure of the (W258A) $G\alpha_{i1}$:R6A complex. The final model of (W258A) $G\alpha_{i1}$:R6A complex consists of residues Leu5-Phe354 of $G\alpha_{i1}$, 9 residues of R6A (Asp1-Leu9), GDP, two sulfite ions and 336 water molecules. There is one $G\alpha_{i1}$:R6A complex in the asymmetric unit. The structure of $G\alpha_{i1}$ in the complex is highly similar to the structure of wild-type $G\alpha_{i1}$·GDP (Mixon et al., 1995) with C_α rmsd of only ~0.5 Å. Much like the $G\alpha_{i1}$·GDP structure, the Switch III region (residues 234-240) had weak and disconnected electron density; these residues were removed from the final model. The first few residues at the N-terminal of the $G\alpha_{i1}$ (Leu5-Asp9) also had weak electron density and only the main chain can be built accurately. Two lysine
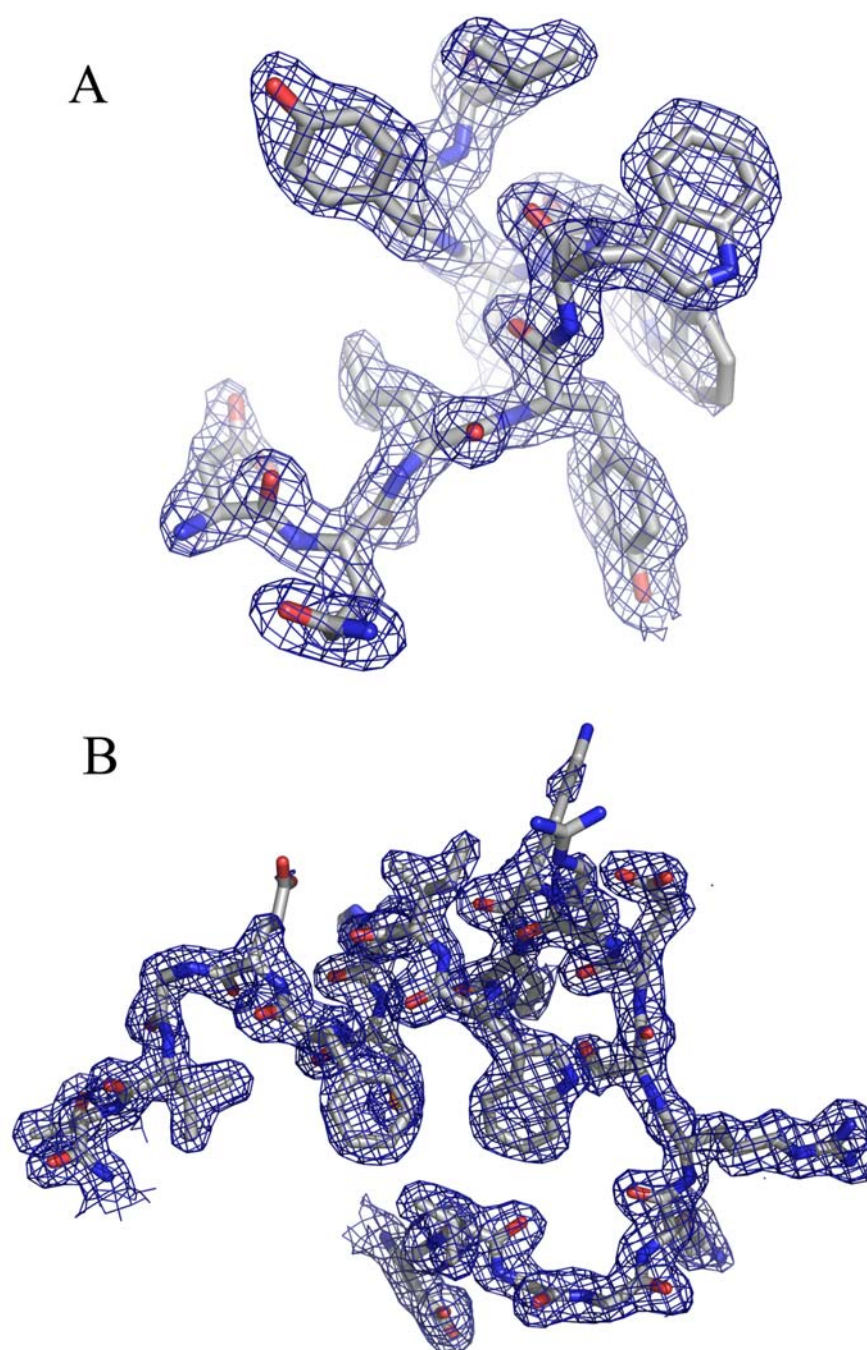
Figure 6.5. Representative electron density map. SigmaA-weighted 2Fo-Fc electron density map contoured at 1.2 σ. *A*. Electron density map is displayed around the model of R6A peptide. *B*. Electron density map is displayed around the switch II region of Gα$_{i1}$.

residues were added at the C-terminus of R6A peptide to enhance the solubility of 9-residue core region. These lysine residues were completely disordered in the structure and did not appear in the electron density map.

R6A interacts with switch II region on one side of the binding pocket, and β1 and α3 of the ras-like domain on the other side (Figure 6.6). The switch II region, which was found to be in a disordered state in the $G\alpha_{i1}$·GDP structure, undergoes significant conformational change upon binding to R6A and adopts an ordered structure. Switch II also adopts an ordered structure when $G\alpha_{i1}$ is bound to other GDIs like the GoLoco motif (Kimple et al., 2002) or Gβγ (Wall et al., 1995), and is also ordered in the $G\alpha_{i1}$·GTPγS structure (Mixon et al., 1995). Upon binding to R6A, residues 209-216 of switch II forms a helical structure, residues 200-205 forms a stable loop region and residues 206-207 forms a β-strand. Both the helical and loop like region of switch II form extensive interactions with the peptide. Binding of R6A pushes switch II away from the Ras-domain core and closer towards the switch I region (Figure 6.6). In essence, R6A acts like a wedge between the switch II region and the Ras-domain core by inserting itself directly between these regions of $G\alpha_{i1}$. When compared to the structure of $G\alpha_{i1}$·GDP, binding of R6A does not induce any conformational change on either the β1 strand or the α3 helix of the ras-domain.
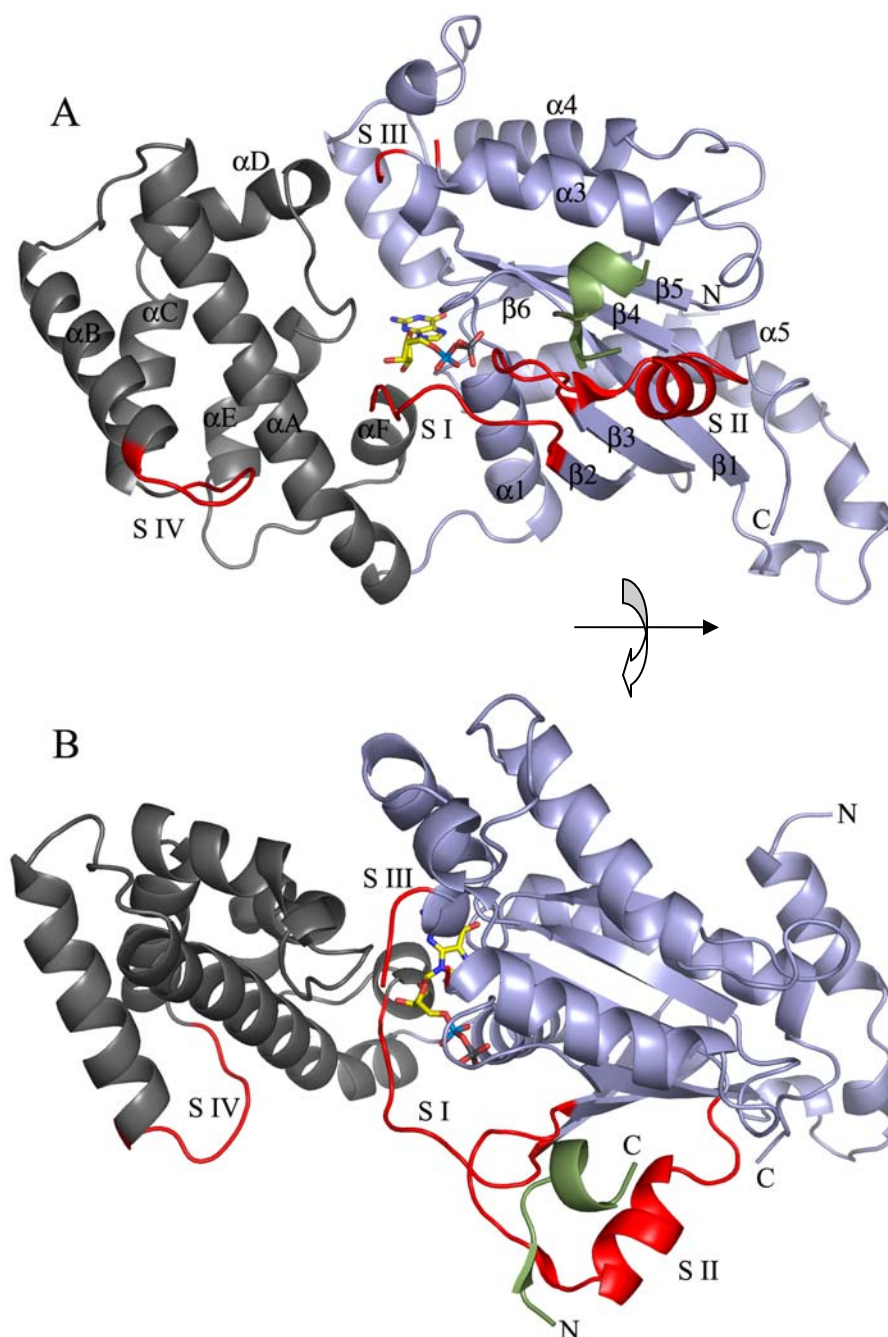
Figure 6.6. Crystal structure of (W258A) Gα$_{i1}$:R6A complex at 1.8 Å resolution. *A*. Ribbon diagram of the structure (W258A) Gα$_{i1}$:R6A is shown. The Ras-like domain is shown in blue, helical domain is shown in gray, switch regions (I, II, III, and IV) are colored in red, and the peptide is shown in green. Secondary structure elements are labeled based on the nomenclature followed in describing Gα$_{i1}$·GDP structure ((Mixon et al., 1995). *B*. View of the complex rotated by 90° around the horizontal axis shown in the image.

R6A adopts a stable structure with two distinct secondary structure elements. All the nine residues corresponding to the functional core of R6A are well ordered. The N-terminal region of R6A (Gln2-Leu3) forms a short anti-parallel β-sheet-like structure with switch II (Ser206-Glu207) (Figure 6.6). This short strand constitutes only two residues. Although the length of the β-strand regions is short, hydrogen bonding interactions clearly identify this region as a β-strand. Secondary structure identification programs like DSSP (Kabsch and Sander, 1983) and STRIDE (Frishman and Argos, 1995) unambiguously assigns a β-strand structure to this region. The C-terminal region of the peptide forms a short α-helical structure consisting of one complete helical turn. The helical region encompasses only four residues (Trp5-Tyr8) of the R6A peptide.

The W258A mutation has no effect on the overall structure of $G\alpha_{i1}$, although it affects the conformation of the loop region that connects the α3 helix and the β5 strand. The mutation destabilizes the preceding N-terminal residue Lys257; however this solvent exposed residue is more than 8 Å away from the peptide binding region and switch II, and is unlikely to affect the conformation of other parts of the structure. The mutation has no effect on the nucleotide exchange rate of $G\alpha_{i1}$ or its ability to interact with R6A (data not shown).

Upon the determination of the 1.8 Å crystal structure of (W258A) $G\alpha_{i1}$:R6A complex, the 3.2 Å data collected from $G\alpha_{i1}$:R6A4 crystals was revisited (discussed in chapter 5). The electron density map of $G\alpha_{i1}$:R6A4 at 3.2 Å was of poor quality, particularly in the peptide binding region (chapter 5.3.4). Molecular replacement was used to recalculate initial phases for the $G\alpha_{i1}$:R6A4 complex using the structure of (W258A) $G\alpha_{i1}$:R6A
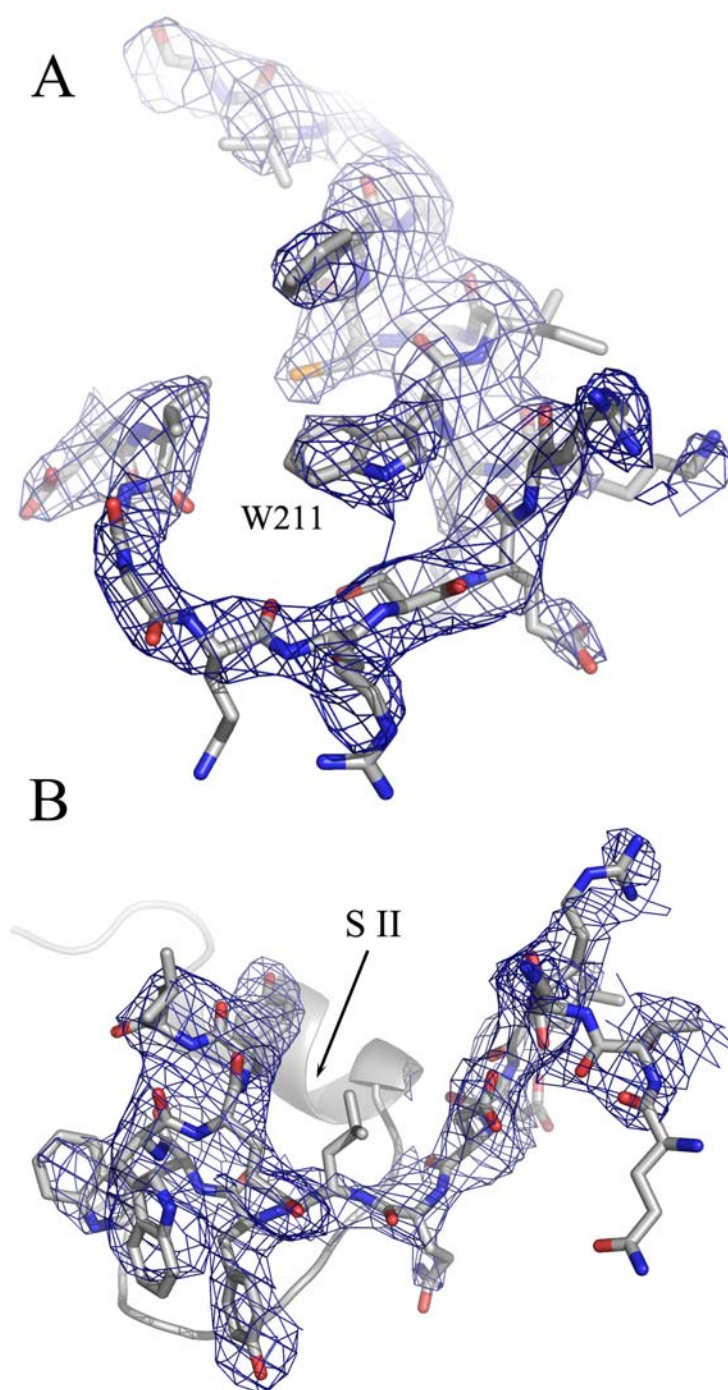
Figure 6.7. Electron density map of the peptide binding region in Gα_{i1}:R6A4 structure.
SigmaA-weighted 2Fo-Fc electron density map contoured at 1.0 σ. *A*. Electron density map is displayed around the switch II region of Gα_{i1}. *B*. Electron density map is displayed around the model of R6A4 peptide and switch II is shown in ribbon representation.

as a search model. The new phases resulted in a more interpretable electron density map

(Figure 6.7). Although, the structure of $G\alpha_{i1}$:R6A4 complex has not been refined,

preliminary analysis shows well ordered main-chain residues for R6A4 and the switch II

region (Figure 6.7). Thus, reducing 'model bias' allowed the interpretation of the electron

density map even at low resolution. The conformation of the peptide and switch II in the

$G\alpha_{i1}$:R6A4 structure is highly similar to the structure of (W258A) $G\alpha_{i1}$:R6A, indicating that

conformation in the peptide binding region is not influenced by different crystal packing

interactions in these two different crystal forms.

6.3.3 Molecular basis of $G\alpha_{i1}$:R6A interaction

R6A peptide forms extensive interactions with the binding pocket in $G\alpha_{i1}$. The

binding of the peptide to $G\alpha_{i1}$ results in the burial of 1271 $Å^2$ solvent accessible surface area

in both molecules (635 $Å^2$ in both R6A and $G\alpha_{i1}$). Analysis of the surface electrostatics of

$G\alpha_{i1}$ and R6A reveals that the interaction of the peptide with $G\alpha_{i1}$ is mediated by both

hydrophobic contacts and polar/hydrogen-bonding interactions. R6A has two distinct binding

surfaces: a hydrophobic surface facing the hydrophobic binding pocket formed by $\beta1$, $\alpha3$,

and switch II in $G\alpha_{i1}$; and a polar surface that stays solvent exposed (Figure 6.8). The

hydrogen bonding interactions are mostly limited to the periphery of the hydrophobic binding

surface.
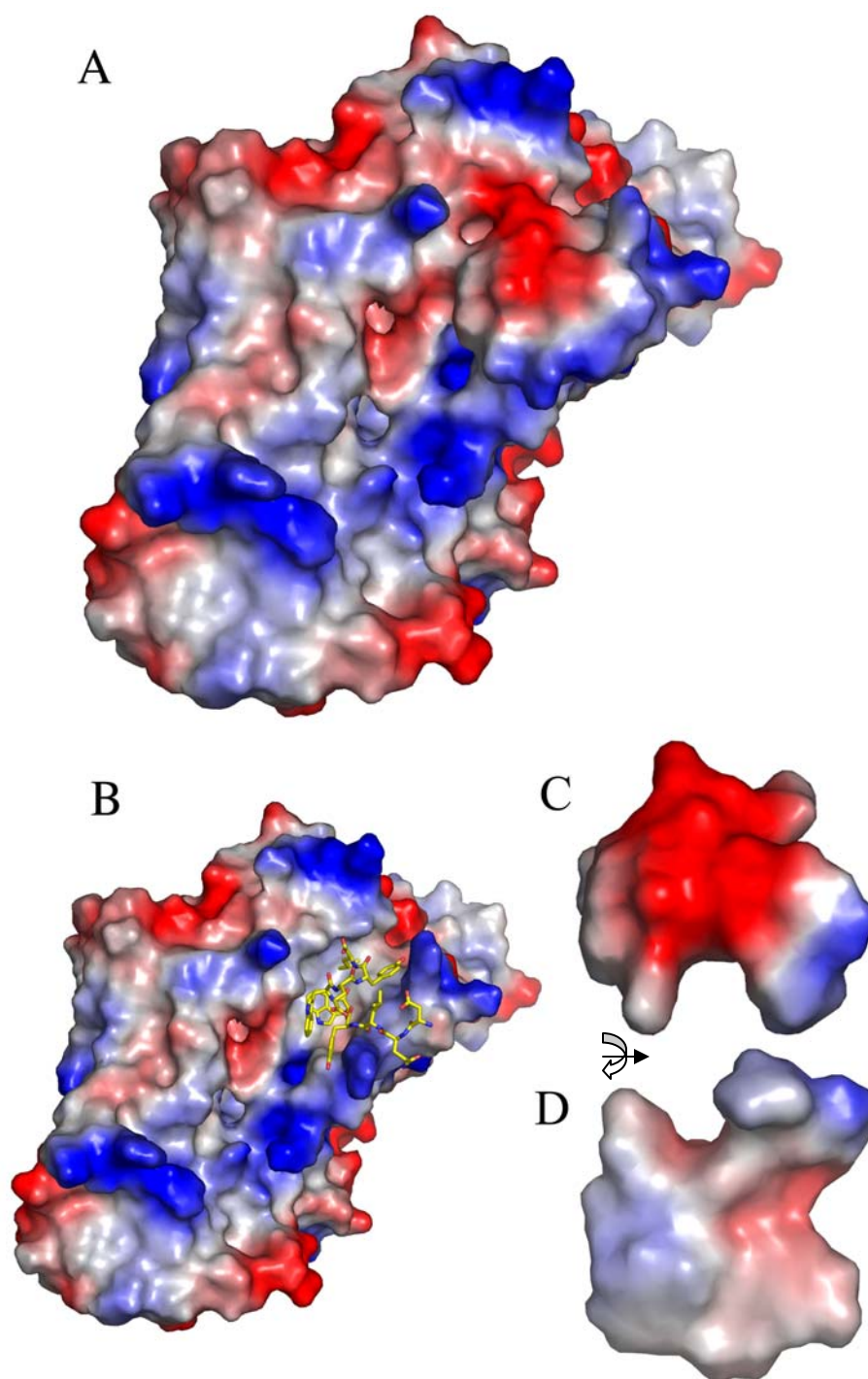
Figure 6.8. Surface electrostatics representation of Gα_{i1}:R6A complex. Electrostatic potential of Gα_{i1}:R6A complex computed using PyMol is shown as a molecular surface in an arbitrary scale of -15 (red) to +15 (blue). *A*. Overall electrostatic potential of the complex. *B*. R6A is shown as stick model and Gα_{i1} is shown as a surface. The peptide binds to a

hydrophobic groove appearing white in color. *C*. Electrostatic potential surface of the solvent-exposed hydrophilic side of R6A is involved in forming hydrogen-bonding interactions with $G\alpha_{i1}$. *D*. Electrostatics of hydrophobic side of R6A involved in packing on the hydrophobic groove in $G\alpha_{i1}$. The view is rotated by 180° along the horizontal axis with respect to C.
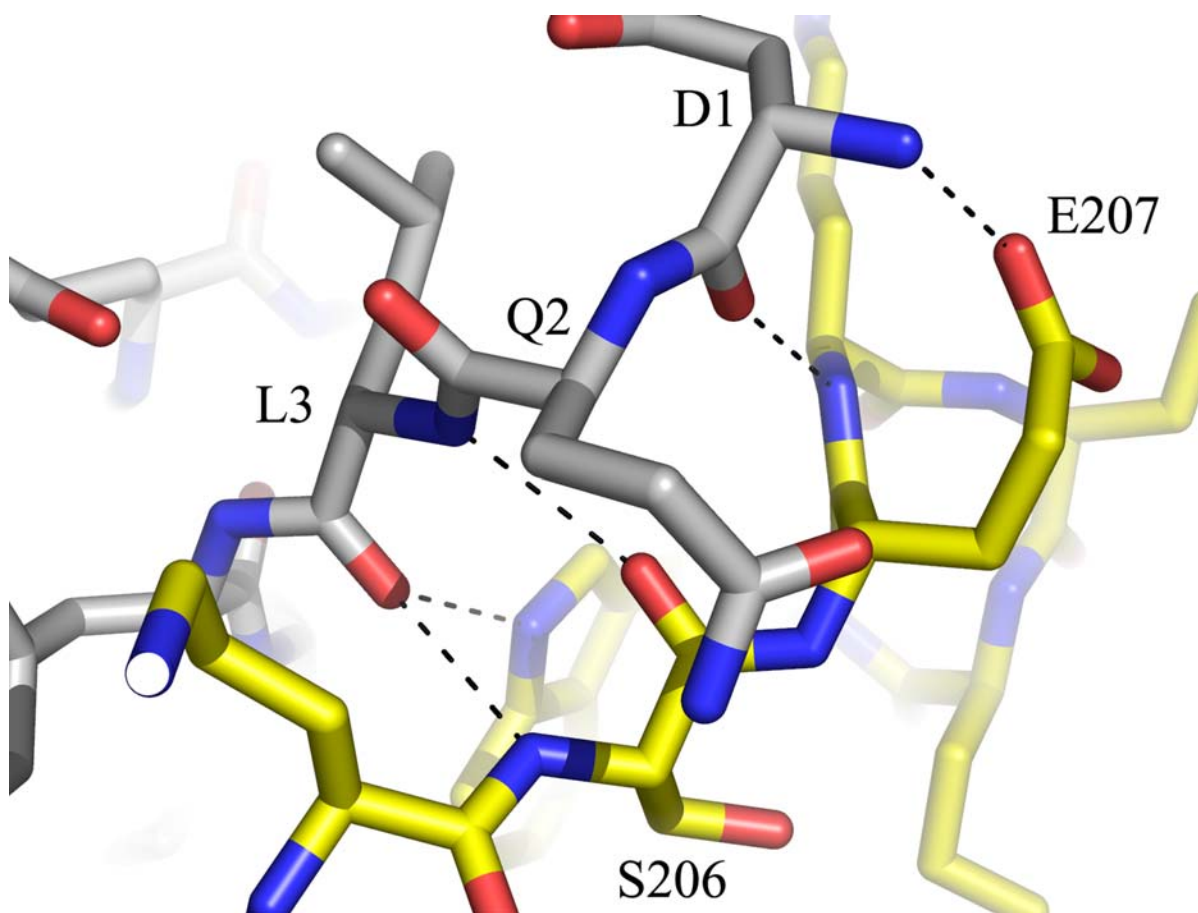
Figure 6.9. Hydrogen-bonding interaction between $G\alpha_{i1}$ and R6A. Anti-parallel β-sheet like hydrogen bonding interaction between the N-terminal region of R6A (grey) and switch II of $G\alpha_{i1}$ (yellow).

R6A makes several hydrogen bonding interactions with $G\alpha_{i1}$ (Table 6.2). Most of the hydrogen bonding interactions involves inter-main chain hydrogen bonds that allow the formation of the β-sheet structure (Figure 6.9). Apart from these main-chain interactions, side-chains of both $G\alpha_{i1}$ and R6A participate in inter-molecular hydrogen bonding. The carboxyl group of Glu207 forms a hydrogen bond with the N-terminus of the peptide. The indole ring of Trp211 forms a hydrogen bond with the carbonyl group of Leu3 in R6A. The indole nitrogen of Trp5 in R6A can potentially form hydrogen bonds with the main-chain carbonyl groups of Gly40 and Gly202; however these interactions are longer in range than typical hydrogen bond lengths and probably represent van der Waals interactions. The intermolecular hydrogen-bonds observed between R6A and $G\alpha_{i1}$ are summarized in Table 6.2.

Peptide residue	Atom	Interacting $G\alpha_{i1}$ residue	Atom	Distance	Type of contact
Asp1	N	Glu207	OE2	2.81	Hydrogen bond
	O	Arg208	N	2.85	Hydrogen bond
Leu3	N	Ser206	O	2.95	Hydrogen bond
	O	Ser206	N	2.97	Hydrogen bond
	O	Trp211	NE1	2.83	Hydrogen bond
Trp5	NE1	Gly40	O	3.99	Polar
	NE2	Gly202	O	3.86	Polar
Leu9	O	Ser252	OG	3.38	Hydrogen bond

Table 6.2. Intermolecular polar interactions between $G\alpha_{i1}$ and R6A.

R6A is a very hydrophobic peptide; six out of the nine residues of R6A core region are hydrophobic in nature. Not surprisingly, the interaction of R6A with $G\alpha_{i1}$ involves
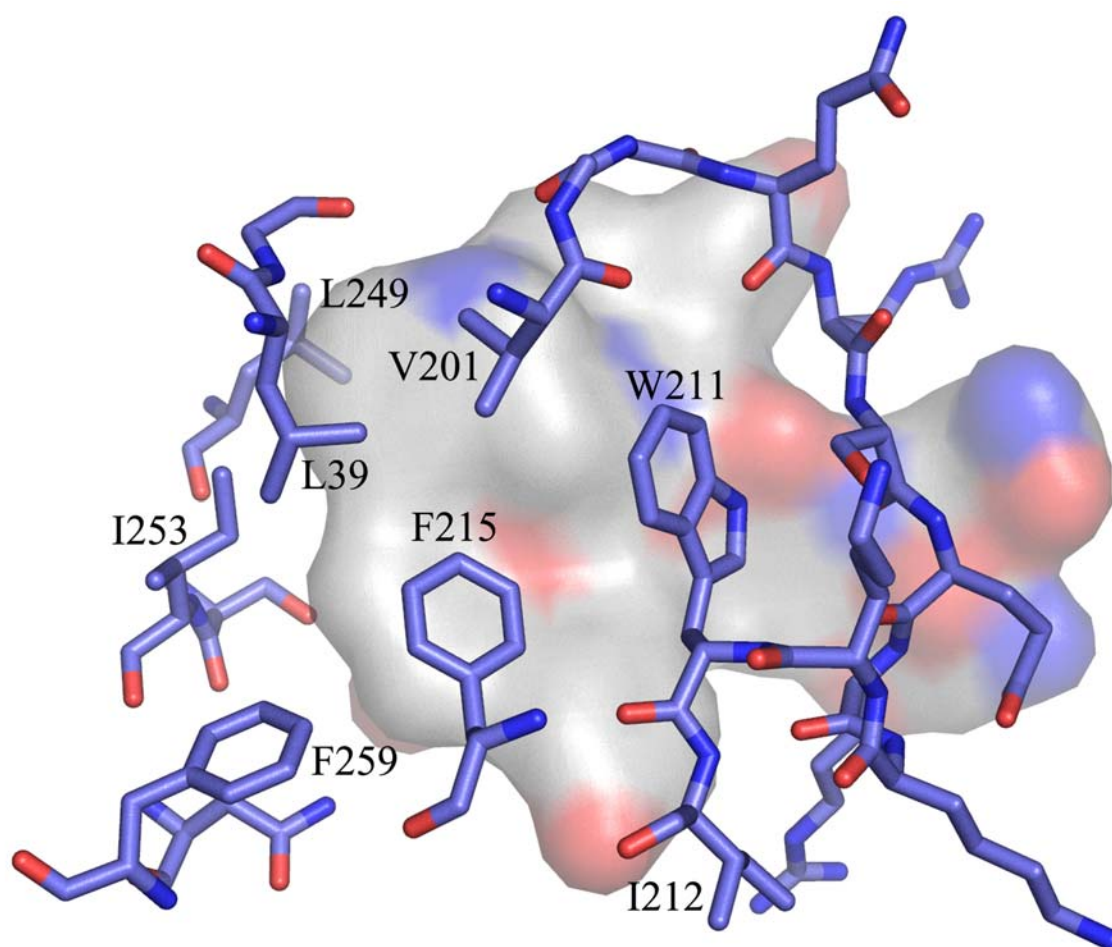
Figure 6.10. Hydrophobic interactions in R6A binding pocket. Gα_{i1} residues making hydrophobic contacts with R6A is shown in blue and R6A is shown as surface. White color represents hydrophobic surface. Blue and red color represents polar surface.

burial of the hydrophobic residues of R6A in a hydrophobic binding pocket in Gα_{i1}. Burial of

these hydrophobic residues potentially contribute a significant fraction of the binding free

energy. The hydrophobic binding pocket in Gα_{i1} is created by the switch II region, β1 strand

and the α3 helix. One of the most prominent hydrophobic contacts is driven by the conserved

tryptophan residue in position 5 of the peptide. This tryptophan is buried in a hydrophobic

environment created by residue Trp211 in switch II and residues Leu39 and I253 of the Ras-domain core. Phe215 of switch II forms hydrophobic interactions with the Tyr8 residue of R6A, and Leu249 of α3 makes hydrophobic contacts with Trp5 and Trp6 of R6A (Figure 6.10, Table 6.3).

Gα_{i1} residue	Peptide Residue
Gly40	Trp5
Val201	Trp5
Gly202	Trp5
Gly203	Tyr4
Gln204	Tyr4
Arg205	Tyr4
	Leu3
Trp211	Trp5
	Tyr8
	Leu3
Ile212	Tyr8
Phe215	Leu9
	Tyr8
Leu 249	Trp6
Ile253	Trp5
	Leu9

Table 6.3. Intermolecular hydrophobic interactions between Gα_{i1} and R6A. Residues involved in forming intra-molecular hydrophobic contacts (within a distance of 4 Å) are shown.

The structure of Gα_{i1}:R6A explains the selectivity of R6A peptides towards the GDP-bound form of Gα_{i1} (Ja WW, 2004). If the structure of Gα_{i1}·GTPγS and Gα_{i1}:R6A are superimposed, the conformation adopted by switch II in the GTPγS bound state overlaps with R6A; indicating that R6A would not be able to bind Gα_{i1}.GTPγS due to steric clashes with
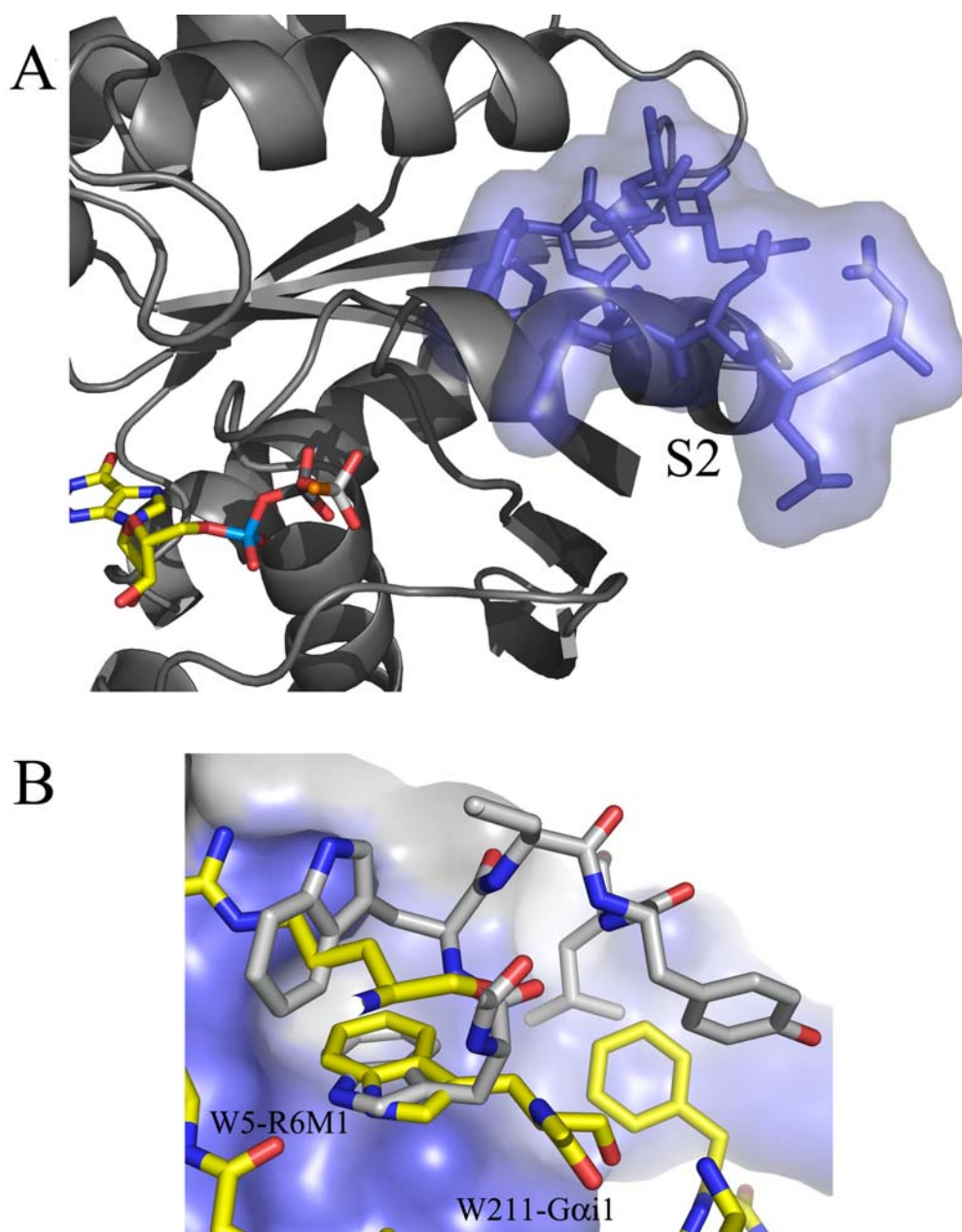
Figure 6.11. Molecular basis of GDP selectivity of R6A peptide. Superposition of
Gα_{i1}:R6A and Gα_{i1}·GTPγS structure. *A*, R6A, shown as transparent surface (blue) overlaps
with switch II (S2) in the Gα_{i1}·GTPγS structure (dark gray). *B*, Trp5 in R6A (gray) binds to
the same residues in Gα_{i1} (shown as blue surface) that are involved in burying W211
(yellow) of switch II in the GTPγS bound structure.

the switch II region (Figure 6.11). Interestingly, the Trp5 side chain of R6A is buried in the

same hydrophobic surface that mediates the burial of Trp211 of switch II in the structure of

GTPγS-bound $G\alpha_{i1}$ (Figure 6.11). The set of hydrophobic residues in $G\alpha_{i1}$ (Leu39, Ile253,

and Leu249) that mediate the mediate the binding of R6A, are also utilized by a variety of

other regulatory proteins for binding to $G\alpha_{i1}$ (discussion in chapter 6.3.8).

6.3.4. GDI activity of R6A peptides

The 16-residue R6A4 peptide has been shown to be a potent GDI and the peptide can

be minimized to a 9 residue core (R6A) while retaining $G\alpha_{i1}$ binding affinity and GDI

activity (Ja WW, 2004). In the R6A peptide, two lysine residues were added to the C-

terminal to enhance the solubility. Initial characterization showed that the lysine-modified

R6A was able to inhibit binding of BOIDPY-GTPγS to $G\alpha_{i1}$ (chapter 5.3.6). To obtain a

biochemical model for GDI activity, the activity of R6A and R6A4 were studied in further

details. The characterization of the biochemical activity of the R6A peptides is still at a

preliminary stage and only the initial results will be discussed.

The spontaneous binding of BODIPY-GMPPNP to $G\alpha_{i1}$ was inhibited by both R6A4

and R6A (Figure 6.12). However, R6A has low GDI activity. Although inhibition of

nucleotide exchange was observed with only 10 μM R6A4, no significant inhibition was

observed when R6A concentration was less than ~ 25 μM (Figure 6.12). The lower activity

of R6A was expected; the 9-residue R6A peptide had 3 fold weaker affinity towards $G\alpha_{i1}$ (K_d

~ 200 nM) in comparison with the 16 residue R6A4 peptide (K_d ~ 60 nM), based on surface

Figure 6.12. Inhibition of BOIDPY-GMPPNP binding to $G\alpha_{i1}$ by R6A peptides. The binding of 1 μM BODIPY-FL-GMPPNP to 200 nM $G\alpha_{i1}$ (black) at 25 $^{\circ}$C in the presence of 10 μM (red), 50 μM (green), and 100 μM (blue) R6M1; 10 μM (cyan) and 100 μM (magenta) R6A4.

plasmon resonance studies (Ja WW, 2004). The activity of R6A showed significant variation

from earlier assays performed with BODIPY-GTPγS (chapter 5.3.6). In the BODIPY-GTPγS

binding assay, R6A strongly inhibited nucleotide exchange at 10 μM. The IC_{50} of R6A

appears much lower than the previously described K_d of R6A:$G\alpha_{i1}$ binding (Ja WW, 2004).

However, ITC studies indicate that the affinity of R6A and $G\alpha_{i1}$ binding is ~ 2 μM (data not

shown), which is 10 fold higher than what was observed by surface plasmon resonance. The

kinetics of nucleotide exchange was different from the expected single rate behavior, which is typically observed in nucleotide binding assays. The binding seems to have two components: a fast binding rate dominating the initial period of the exchange reaction and a slow binding rate where the binding essentially plateaus. The data can be fitted to a biphasic exponential association model.

To investigate whether the presence of BODIPY moiety influences the biphasic kinetic behavior, exchange of GDP for MANT-GMPPNP was also measured. The binding of MANT-GMPPNP to $G\alpha_{i1}$ was inhibited in the presence of R6A4 and R6A (Figure 6.13). Both of these peptides appeared to have higher GDI activity in the MANT based assay in comparison with the BODIPY based assay, but their relative activities were similar (R6A4>R6M1). 5-10 μM R6A4 strongly inhibited the binding of MANT-GMPPNP; significant inhibition was also observed with 25 μM R6A. The change in the fluorescence of MANT-GMPPNP upon binding to $G\alpha_{i1}$ was smaller than the change in BODIPY fluorescence, which caused the MANT-based data to be noisy. The MANT-GMPPNP binding kinetics also appeared to exhibit biphasic exchange kinetics, much like its BOIDPY counterpart; but the biphasic nature was less prominent in comparison with BOIDPY.

The factors that cause the biphasic behavior is not understood yet, but it is likely that conditions under which the assay is performed influence such behavior. The affinity of binding of R6A peptide to $G\alpha_{i1}\cdot$GDP increases in the presence of Mg^{++} (Ja et al., 2005). Because $MgCl_2$ is present in the assay buffer at 10 mM concentration, it might influence the behavior of $G\alpha_{i1}$:R6A complex. At high concentration, Mg^{2+} alters the conformation of
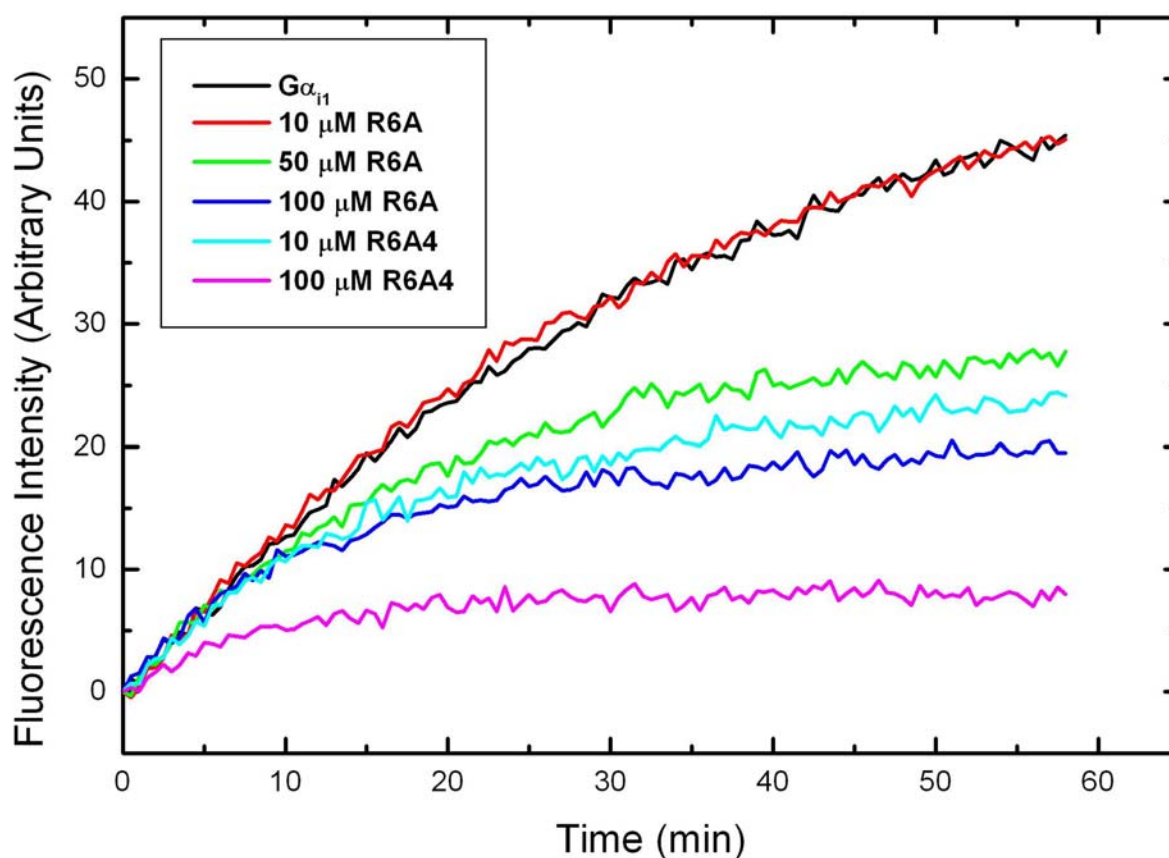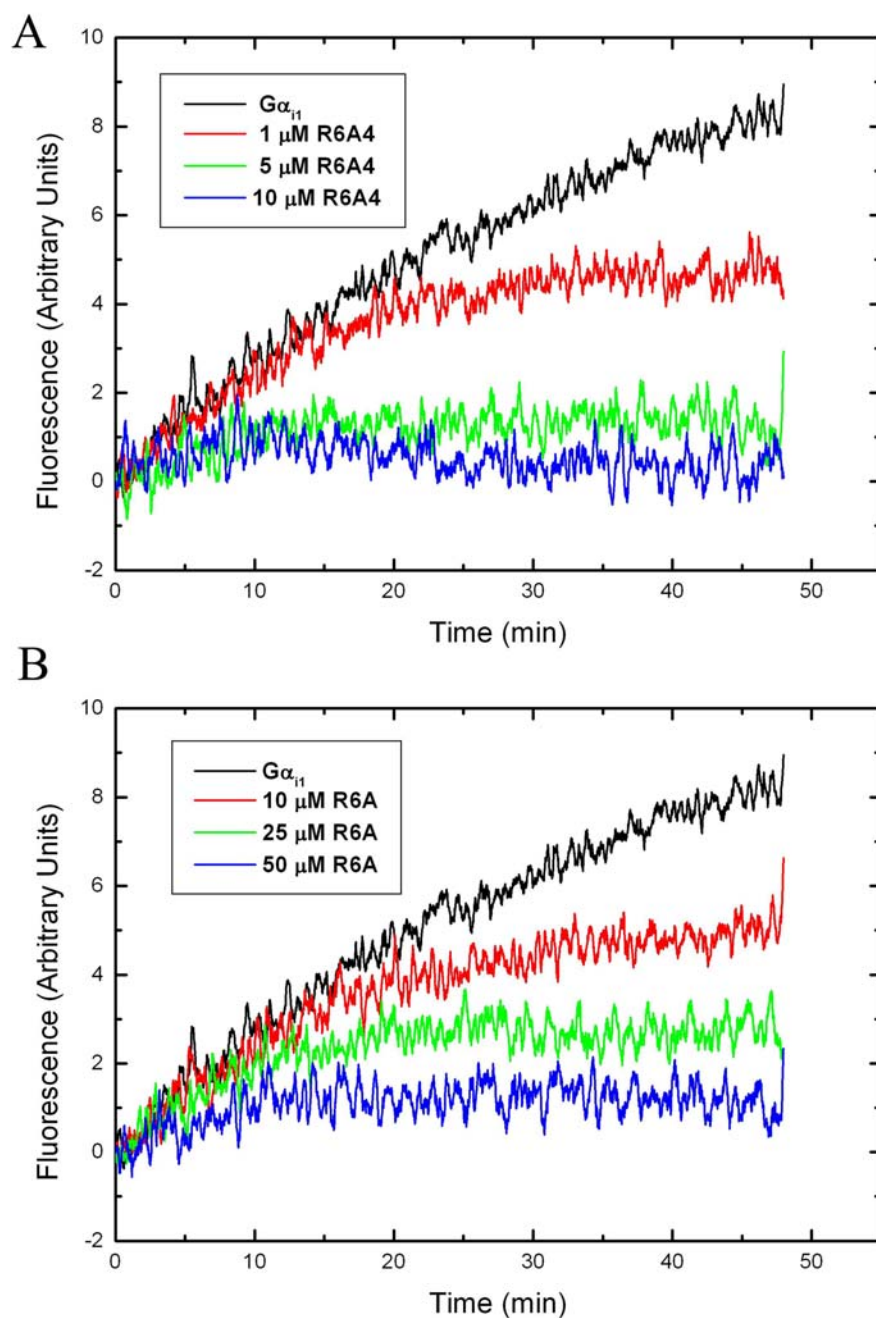
Figure 6.13. Inhibition of MANT-GMPPNP binding to Gα_{i1} by R6A peptides. The binding of 1 μM MANT-GMPPNP to 200 nM Gα_{i1} (black) at 25 ^{o}C in the presence of: A, 1 μM (red), 5 μM (green), and 10 μM (blue) R6A4; B, 10 μM (red), 25 μM (green), and 50 μM (blue) R6M1.

switch I in the GDP-bound state, and promotes switch I to adopt a conformation resembling the GTPγS-bound state (Coleman and Sprang, 1998). High concentration of Mg^{2+} also promotes the release of GDP from the Gα·GDP:βγ complex (Higashijima et al., 1987b). Comparison of the structure of $Gα_{i1}$·GDP·Mg^{2+} (PDB code: 1BOF) and $Gα_{i1}$·GDP:R6A reveals, that interaction of Mg^{2+} with switch I might disrupt the hydrogen bonding interaction observed between switch I and switch II in the $Gα_{i1}$·GDP:R6A structure (discussed in chapter 6.3.5). Thus, presence of Mg^{2+} in the assay can alter the kinetics of GTP binding to $Gα_{i1}$·GDP:R6A complex by promoting two distinct populations of $Gα_{i1}$·GDP:R6A complex. While this might account for the aberrant binding kinetics, the effect of increasing Mg^{++} concentration on binding kinetics has not been tested yet. The binding of fluorescent GTP analogs to $Gα_{i1}$·GDP:R6A complex plateaus at a lower level than what is observed for $Gα_{i1}$·GDP. Failure of $Gα_{i1}$ to attain stoichiometric level of GTPγS binding is not unusual; a fraction of $Gα_{i1}$ often becomes irreversibly inactivated under assay conditions (Zelent et al., 2001). The nucleotide-free intermediate in nucleotide exchange is susceptible to deactivation. The kinetics of such irreversible deactivation during nucleotide exchange has been well characterized in $Gα_q$ (Chidiac P, 1999). The binding of R6A peptides are specific towards the GDP-bound state and the affinity of R6A towards a nucleotide-free state of $Gα_{i1}$ is not known. Whether the binding of R6A irreversibly deactivates $Gα_{i1}$ remain to be tested.

The intrinsic tryptophan fluorescence of $Gα_{i1}$ increases upon the binding of GTP, and nucleotide exchange can be measured by monitoring the intrinsic tryptophan fluorescence (Higashijima et al., 1987a). The fluorescence change is primarily contributed by
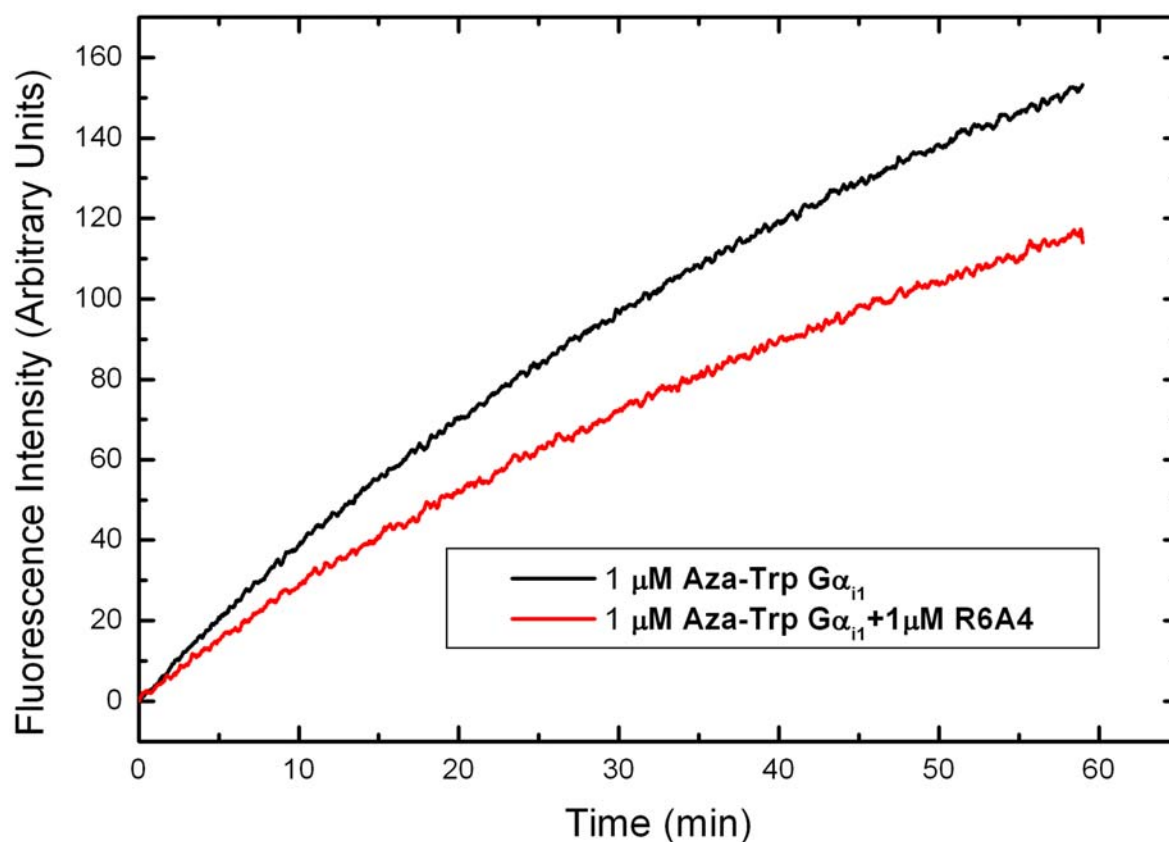
Figure 6.14. Inhibition of GTPγS binding to Gα$_{i1}$ by R6A4. The binding of 10 μM GTPγS to 1 μM 7-AW Gα$_{i1}$ (black) is inhibited in the presence of 1 μM R6A4 (red).

Trp211 in the switch II region. The presence of two tryptophan residues in R6A, with strong

intrinsic fluorescence, precludes the performance of a tryptophan-fluorescence based

nucleotide exchange assay. This problem can be solved by using 7-azatryptophan (7-AW)

labeled Gα$_{i1}$, where the fluorescence emission wavelength (390 nm) is red-shifted from the

tryptophan emission wavelength (340 nm). Results from preliminary studies with 7-AW

labeled Gα$_{i1}$ (a gift from Prof. Elliott Ross, UT Southwestern Medical Center) was

encouraging. The increase in the fluorescence of 7-AW Gα$_{i1}$ upon binding of GTPγS was

inhibited in the presence of R6A4 (Figure 6.14). The tryptophan fluorescence of R6A4 is not

completely diminished at 390 nm. To maximize the fluorescence contribution from 7-AW, instead of using large excess of R6A4, both R6A4 and 7-AW $G\alpha_{i1}$ was used at 1 μM concentration. Even at equimolar conditions, R6A4 inhibited nucleotide exchange (Figure 6.14). Interestingly, the binding followed single exponential association kinetics, and the biphasic behavior observed with fluorescent GTP analogs was absent. In light of this observation, it seems possible that the two-rate behavior of the fluorescent GTP analog binding was influenced by the attached fluorophores. R6A has not been tested yet using the 7-AW $G\alpha_{i1}$ based assay. The activity of R6A peptides needs further investigation to establish a kinetic model of R6A mediated inhibition of nucleotide exchange (discussion in chapter 7).

6.3.5   Structural basis of the GDI activity of R6A

Nucleotide dependent structural changes in $G\alpha_{i1}$ are mostly localized to the switch regions (Mixon et al., 1995; Sprang, 1997). To gain insight into the molecular basis of the GDI activity of R6A, the conformation of the switch regions in $G\alpha_{i1}$:R6A complex was compared to the conformation of switch regions in the structures of $G\alpha_{i1}$·GDP (PDB code: 1GDD), $G\alpha_{i1}$·GTPγS (PDB code: 1GIA), $G\alpha_{i1}$:RGS14 GoLoco (PDB code: 1KJY), and $G\alpha_{i1}$:G$\beta_1\gamma_2$ (PDB code: 1GP2). Apart from the switch regions, the interaction of the bound nucleotide with $G\alpha_{i1}$ was also analyzed. The overall structure of $G\alpha_{i1}$ in all of these forms was very similar and the differences were limited to the switch regions. The rmsd of C_α atoms between the structure of $G\alpha_{i1}$:R6A and the structures of $G\alpha_{i1}$.GDP, $G\alpha_{i1}$.GTPγS, $G\alpha_{i1}$.GoLoco, and $G\alpha_{i1}$.G$\beta_1\gamma_2$ were 0.56, 0.85, 1.23, and 1.04 Å respectively (Table 6.4).

The stabilization of the GDP bound state by the GoLoco motif of RGS14 and Gβγ were attributed primarily to stabilization of switch II and stabilization of GDP in the nucleotide binding pocket (Kimple et al., 2002; Wall et al., 1995). In the GoLoco-bound structure of Gα$_{i1}$, switch II is stabilized through its interaction with the N-terminal helical segment of the GoLoco motif. In addition, a conserved arginine residue at the C-terminus of GoLoco motif forms hydrogen bonds with the β-phosphate of GDP (Kimple et al., 2002). Residues that are C-terminal to the GoLoco motif of RGS14 interact with the helical domain of Gα$_{i1}$ and tether the helical domain and Ras-domain to each other, resulting in further stabilization of the nucleotide binding pocket (Kimple et al., 2002; Kimple RJ, 2002; Willard et al., 2004). Unlike the GoLoco motif, Gβγ neither forms a direct contact with the bound GDP, nor does it interact with the helical domain of Gα$_{i1}$. However, Gβγ stabilizes the switch II region by binding directly to switch II (Wall et al., 1995). Although Gβγ does not interact directly with GDP, it promotes hydrogen bond formation between Arg178 in the helical domain and Glu43 in the Ras-domain. This hydrogen bond, originally referred to as a 'seatbelt', was thought to stabilize GDP in the nucleotide binding pocket (Wall et al., 1995). The 'seatbelt' is also observed in the structure of GoLoco-bound Gα$_{i1}$·GDP (Kimple et al., 2002).

The only common theme in the interaction of R6A, GoLoco, and Gβγ with Gα$_{i1}$·GDP is stabilization of switch II through direct interactions (Figure 6.15). Interactions that stabilize the nucleotide binding pocket in GoLoco and Gβγ-bound structures are not observed in the structure of Gα$_{i1}$·GDP:R6A complex. In the R6A-bound structure of Gα$_{i1}$·GDP,
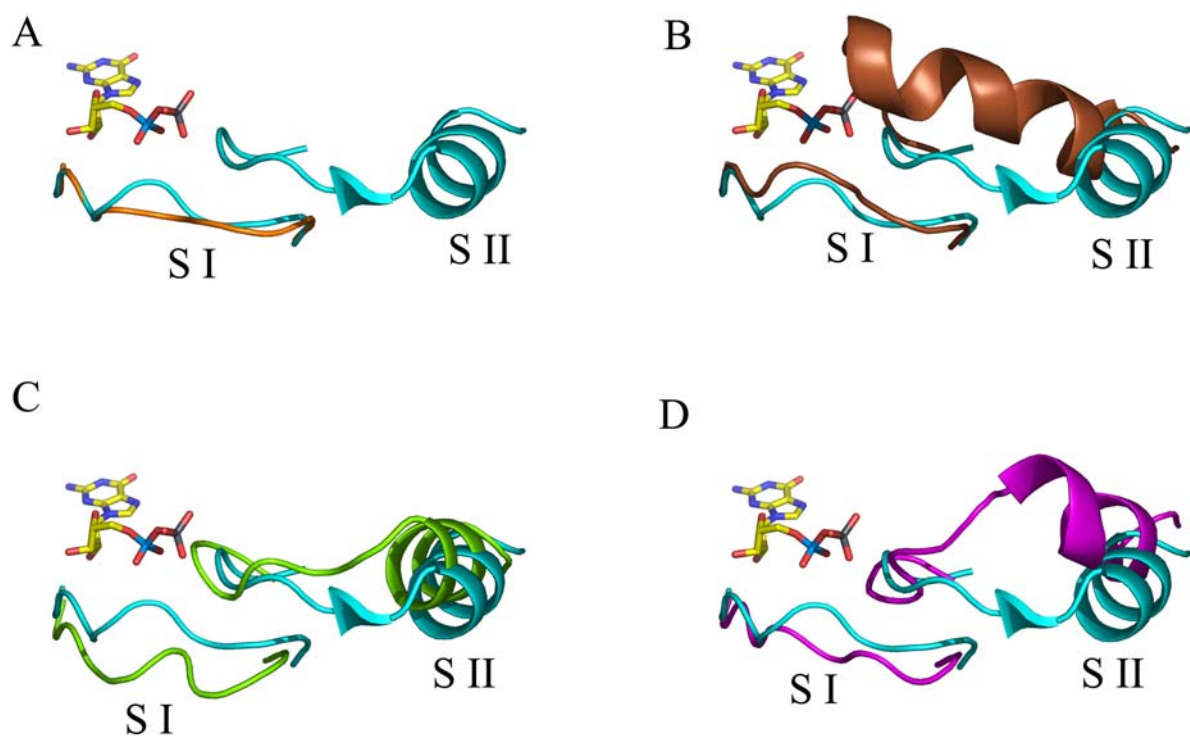
Figure 6.15. Comparison of switch region conformations. The structure of $G\alpha_{i1}$:R6A complex is superimposed with the structures of $G\alpha_{i1}$·GDP (A), $G\alpha_{i1}$·GTPγS (B), $G\alpha_{i1}$:GoLoco (C), and $G\alpha_{i1}$:G$\beta\gamma$ (D). The conformations of switch I (S I) and switch II (S II) of $G\alpha_{i1}$:R6A (cyan), $G\alpha_{i1}$·GDP (orange), $G\alpha_{i1}$·GTPγS (brown), $G\alpha_{i1}$.GoLoco (green), and $G\alpha_{i1}$.$\beta\gamma$ (magenta) are displayed in ribbon representation. $G\alpha_{i1}$:R6A switch regions are shown in the all (A, B, C, D) the panels for comparison. Switch II adopts helical structure in all except $G\alpha_{i1}$·GDP.

instead of adopting a 'seatbelt conformation', Arg 178 forms direct hydrogen bonds with

GDP. While this interaction might stabilize GDP, the same salt bridge is observed in the

structure of $G\alpha_{i1}$·GDP itself, indicating that additional stability is unlikely to be gained from this interaction. The comparison of R6A, $G\beta\gamma$, and GoLoco-bound structures of $G\alpha_{i1}$ indicates that stabilization of switch II by GDIs is necessary and sufficient to inhibit GDP release. The conformation of switch regions show significant variation among the GDI-bound structures, but the binding of R6A primarily affects switch II and parts of switch I (residues 180-184); the rest of $G\alpha_{i1}$·GDP:R6A structure is almost identical to the structure of $G\alpha_{i1}$·GDP (C_α rmsd ~ 0.56 Å) (Figure 6.15).

Inasmuch as the stabilization of switch II is important for GDI activity, the conformation of switch II among the GDI bound structures differ significantly (Figure 6.16). The differences are predominantly in the helical region of switch II (residues 207-218); the N-terminal region (residues 202-204) adopts relatively similar conformation (rmsd calculations are shown in Table 6.4) (Figure 6.16). The conformation of the N-terminal region of switch II was thought to play important role in occluding the GDP exit route from $G\alpha$ (Iiri et al., 1998). This region, referred to earlier as 'β3/α2 loop' or 'lip of switch 2' (Iiri et al., 1998; Johnston CA, 2005), contains two glycine residues (Gly202 and Gly203). For convenience, this region will be referred to as the "GG-loop" in this discussion. GG-loop adopts radically different conformation depending on whether switch II is stabilized by a GDI in the GDP-bound state or is stabilized by a non-hydrolyzable GTP analog. Even though the mean rmsd of GG-loop in GDI-bound structures is ~2 Å (Table 6.4), their overall orientation is similar (Figure 6.16) In the GDI-bound structures, the GG-loop is close to switch I and the GDP binding pocket, suggesting that this conformation can occlude the potential GDP release route (Figure 6.17). Binding of GTPγS induces almost a 90° rotation of
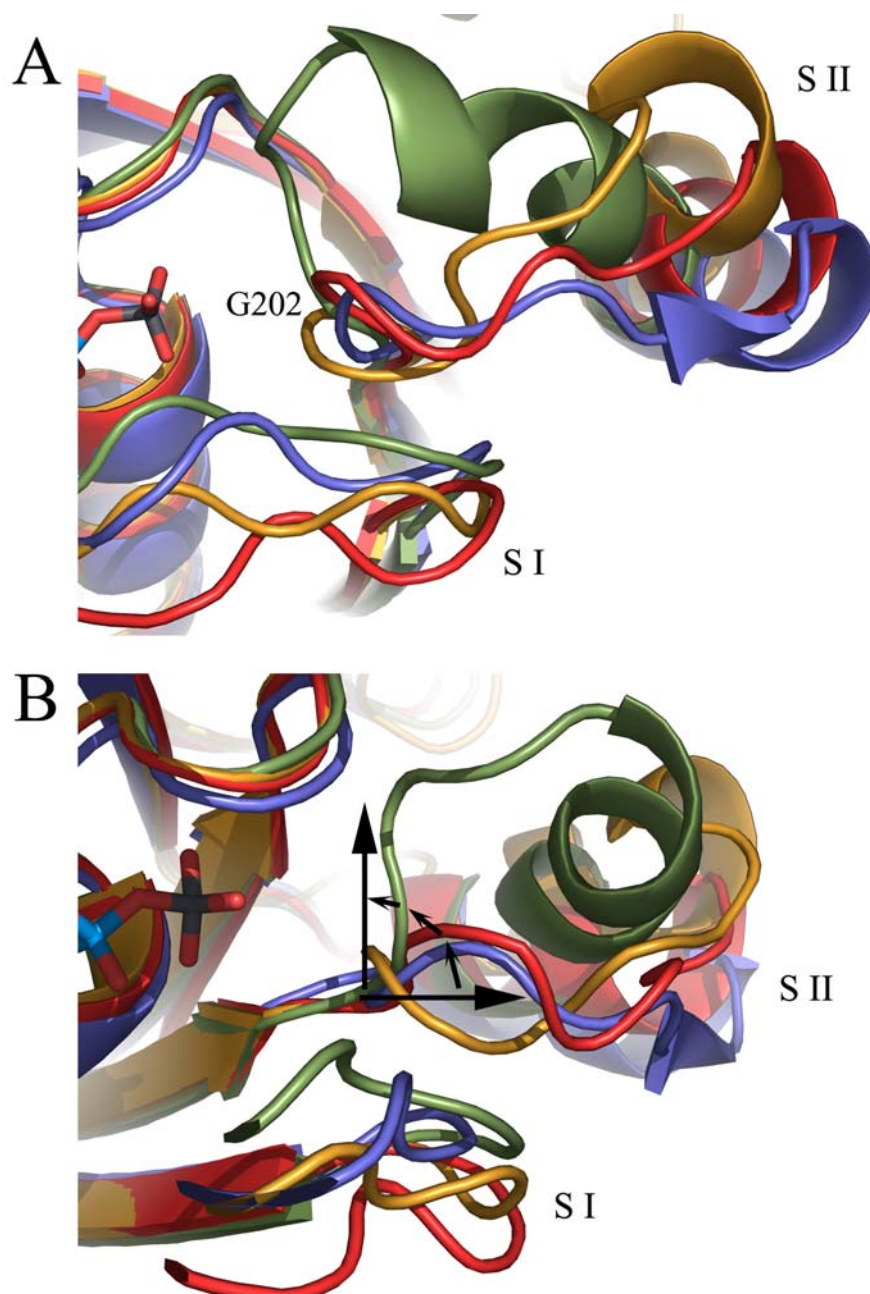
Figure 6.16. Conformation of the GG-loop in GDI bound structures. The structures of
$G\alpha_{i1}$:R6M1 (blue), $G\alpha_{i1}$:GoLoco (red), $G\alpha_{i1}$:G$\beta\gamma$ (orange), and $G\alpha_{i1}$·GTPγS (green) are
superimposed and shown in ribbon representation; switch I (S I) and switch II (S II) regions
are displayed. A. Gly202 is at the base of GG-loop. GG-loop in the $G\alpha_{i1}$·GTPγS structure is
drastically displaced from the rest. B. The displacement of GG-loop in the $G\alpha_{i1}$·GTPγS
structure in relation to the GDI bound structures are shown by arrows. The putative rotational
movement of switch II during nucleotide exchange is shown by dashed arrows.

RMSD (Å)	GoLoco	Gβγ	GTPγS	KB-752
Overall C_α[1]	1.24	1.04	0.85	0.69
Switch II (V201-V218)[2]	2.14	4.77	6.59	1.94
GG-loop (G202-Q204)[3]	2.00	2.07	5.47	2.82
Switch II helix (E207-V218)[4]	2.16	5.41	6.46	1.47
GG-loop similarity index[5]	0.9	0.4	0.8	1.9

Table 6.4. Quantitative analysis of conformational changes in switch II. The root mean square deviation (rmsd) of four different $G\alpha_{i1}$ structures from the $G\alpha_{i1}$·GDP:R6A structure is shown in Å. The structures of $G\alpha_{i1}$·GDP (PDB code: 1GDD), $G\alpha_{i1}$·GTPγS (PDB code: 1GIA), $G\alpha_{i1}$·GDP:RGS14 GoLoco (PDB code: 1KJY), $G\alpha_{i1}$·GDP:Gβ$_1$γ$_2$ (PDB code: 1GP2), and $G\alpha_{i1}$·GDP:KB-752 (PDB code: 1Y3A) were superimposed on the structure of $G\alpha_{i1}$·GDP:R6A.

[1] Structure superposition was performed using all the C_α atoms of $G\alpha_{i1}$ that are common in the structures.

[2] The rmsd of switch II region (residues 201-218) was calculated after the structure superposition.

[3,4] Switch II was divided into two parts: the N-terminal GG-loop region (residues 202-204) and the C-terminal switch II helix (residues 207-218).

[5] GG-loop similarity index is the ratio of the rmsd of GG-loop and switch II helix. The GG-loop similarity index is used to identify the relative variation of switch II with respect to the overall variation in switch II. Index <1 indicates that GG-loop conformation is relatively unchanged and overall switch II rmsd has higher contribution from the switch II helix region and vice versa. The GoLoco and Gβγ bound structures had similarity index <1, indicating that the GG-loop conformation in those structures are more similar to GG-loop conformation in R6A-bound structure

the GG-loop towards the nucleotide (Figure 6.16). Although the GG-loop adopts two distinct orientations in the GDI-bound and GTPγS-bound states, it is disordered in the structure of $G\alpha_{i1} \cdot GDP$ (Mixon et al., 1995). It is possible that GG-loop does not effectively occlude the GDP release route when it is disordered, which results in the higher GDP release rate that is observed when $G\alpha_{i1} \cdot GDP$ is not interacting with a GDI. In contrast to the $G\alpha_{i1} \cdot GDP$ structure, switch II is ordered in the structure of $G\alpha_t \cdot GDP$ (Lambright et al., 1994). The intrinsic rate of GDP release is very slow in $G\alpha_t$, suggesting that disorder of switch II in $G\alpha_{i1} \cdot GDP$ is more favorable for GDP release.

Among all the GDI-bound structures compared, GG-loop is closest to switch I in the R6A bound structure of $G\alpha_{i1} \cdot GDP$ (minimum distance ~2.9 Å) (Figure 6.17). This allows the formation of additional hydrogen bonding interactions between switch II and switch I (carboxyl group of Thr182 forms hydrogen bond with the amide nitrogen of Gly203 and Gln204). The additional hydrogen bonds alter the conformation of switch I in the $G\alpha_{i1} \cdot GDP$:R6A structure; switch I is closer to the nucleotide binding pocket than its position in the $G\alpha_{i1} \cdot GDP$ structure (Figure 6.17). Thus, binding of R6A to $G\alpha_{i1} \cdot GDP$ effectively narrows the potential GDP exit route by bringing GG-loop and switch I closer towards the nucleotide binding pocket.

The crystal structure of $G\alpha_{i1}$ in complex with R6A allows the identification of the crucial structural determinants of R6A activity. In addition, it highlights the role of GG-loop in occluding the nucleotide binding pocket, which is also observed in naturally occurring GDIs.

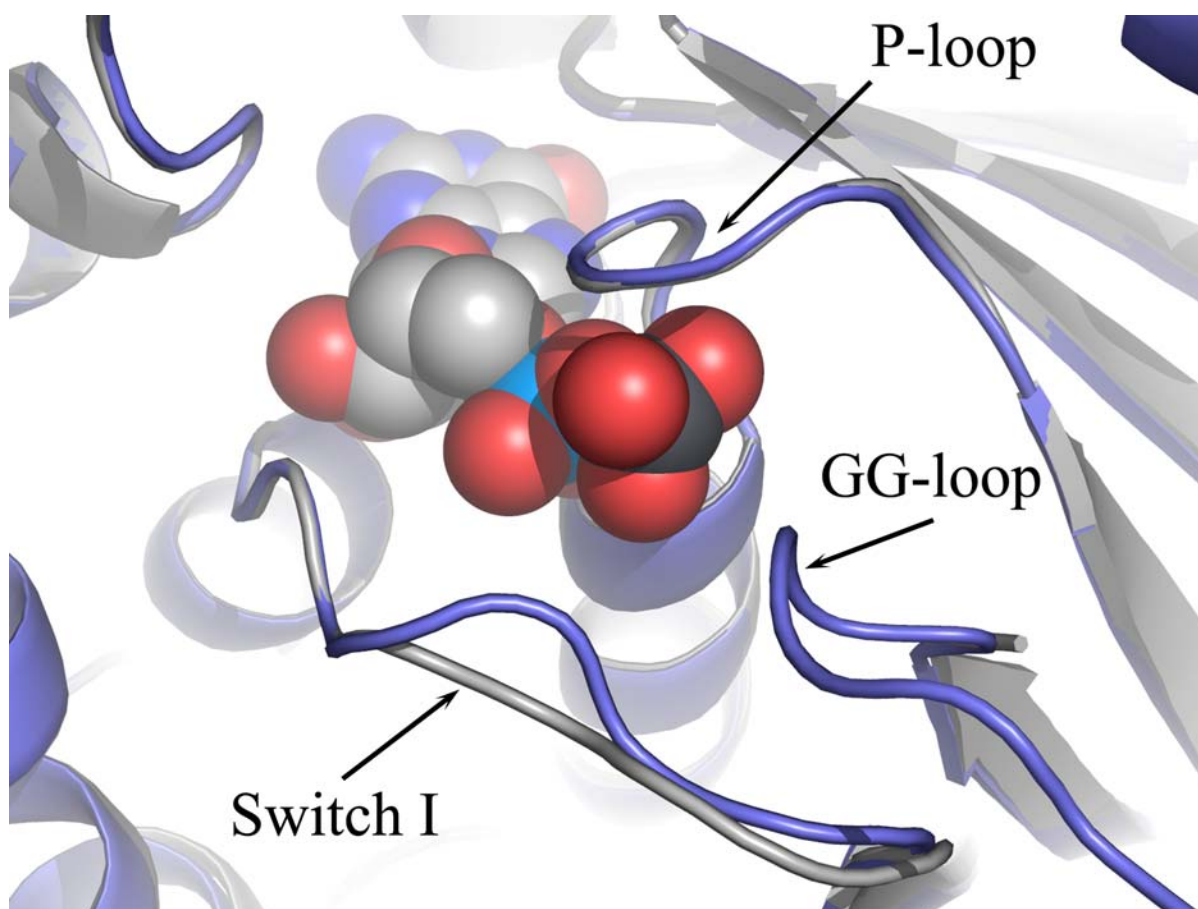**Figure 6.17. Nucleotide binding pocket in the structure of Gα$_{i1}$·GDP:R6A complex.**
Superposition of the structures of Gα$_{i1}$·GDP (grey) and Gα$_{i1}$·GDP:R6A (blue) complex is shown. The bound GDP is shown as spheres. GG-loop and switch I move closer to the nucleotide binding pocket in the Gα$_{i1}$·GDP:R6A complex and narrows the potential GDP exit route.

6.3.6   Comparison of GDI and GEF peptide bound structures of Gα$_{i1}$

The structure of Gα$_{i1}$ in complex with a peptide (KB-752) with low GEF activity has been determined recently (Johnston et al., 2005). While GEF activity is directly in contrast with GDI activity, the similarities in the sequence of KB-752 and R6A are surprising. Sequence alignment of KB-752 and R6A shows variation in only one conserved residue. The tyrosine residue in position 4 of the R6A is replaced by a threonine in the KB-752 peptide. Even more, the structure of Gα$_{i1}$:R6A and Gα$_{i1}$:KB-752 are remarkably similar with C$_\alpha$ rmsd of only 0.69 Å over the entire structure (Table 6.4). The structural differences between these two complexes are localized to the conformation of switch II and the position of the peptide in relation to Gα$_{i1}$ (Figure 6.18). Even though the residues involved in the interaction of Gα$_{i1}$ with R6A and KB-752 are almost identical, R6A and KB752 peptides are ~2-2.5 Å apart when the structures are superimposed on the basis of C$_\alpha$ residues. The conformation of the GG- loop is also significantly different among the structures (Table 6.4). The GG-loop in the KB-752 bound structure is displaced ~2-3.2 Å away from the R6A bound structure (Figure 6.19). The GG-loop in the KB-752 bound structure is pulled away from the nucleotide binding pocket. This suggests that in the GEF peptide bound structure, GG-loop is poised in such a manner, that the potential exit route of GDP will be wider compared to the GDP exit route in Gα$_{i1}$:R6A complex (Figure 6.17)

The conformational difference between KB752 and R6A bound structure might be a result of sequence difference at position four of the peptides. The tyrosine residue at position four in R6A point outwards from the peptide and the side chain is located at ~3 Å distance from the GG-loop. If the Gα$_{i1}$:R6A structure is superimposed on KB-752 bound structure,
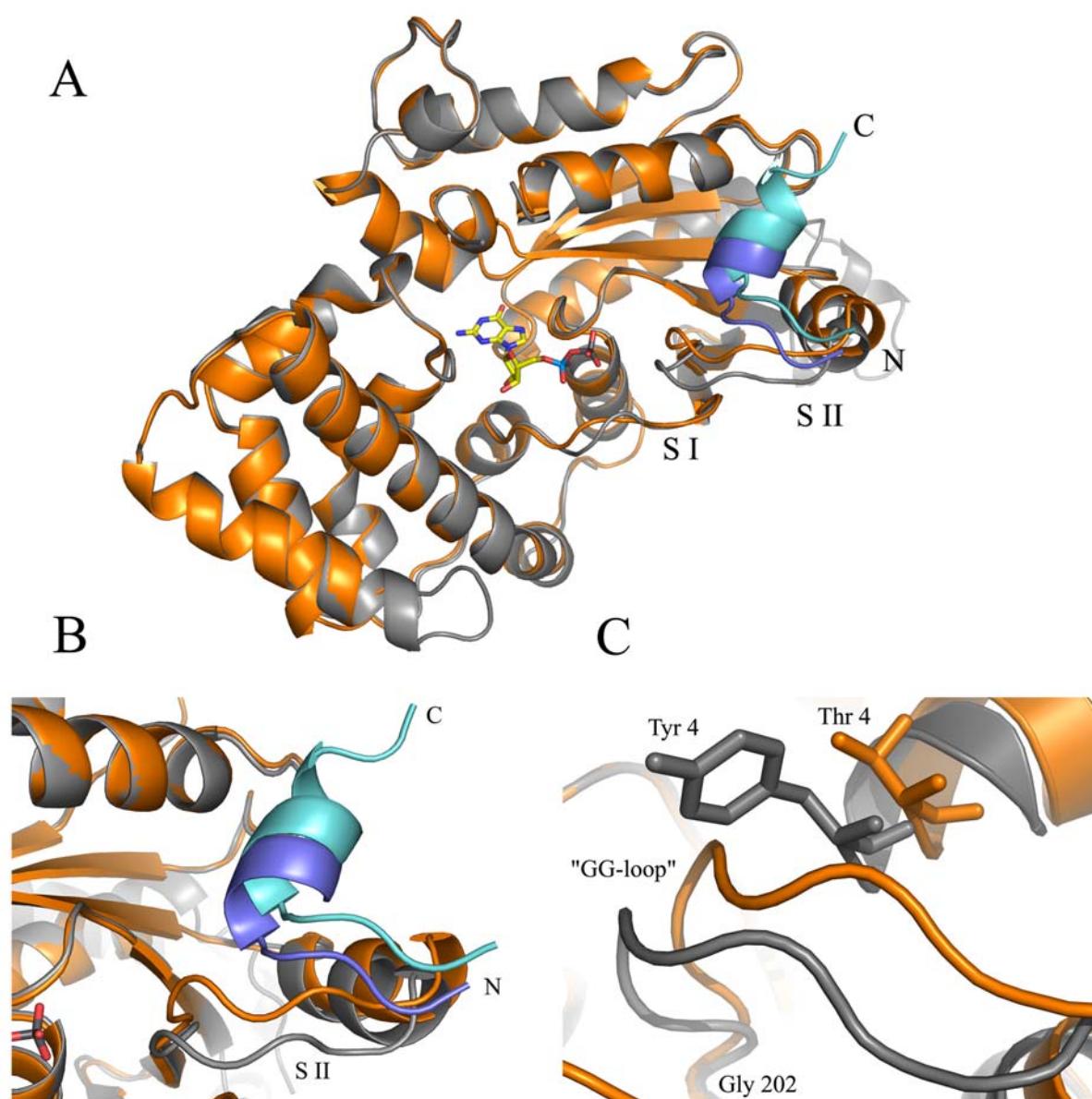
Figure 6.18. Comparison of Gα$_{i1}$:R6A and Gα$_{i1}$:KB752 structures. A. Superposition of the structures of Gα$_{i1}$:KB752 (orange) and Gα$_{i1}$:R6M1 (dark grey) (rmsd~0.69 Å). KB-752 is shown in cyan and R6A is shown in blue. B. Detailed view of structural differences in the GG-loop region. C. Difference in the GG-loop conformation between Gα$_{i1}$:KB752 (orange) and Gα$_{i1}$:R6M1(dark grey). Tyr4 in R6A pushes the GG-loop towards switch I, while Thr4 in KB-752 allows more flexibility. Tyr4 side-chain would have steric clashes with the 'GG-loop' conformation in the KB-752 bound structure.

there will be steric clashes between the Tyr4 side-chain and the GG-loop of KB-752 bound

structure (Figure 6.18). In the KB-752 structure, the threonine residue in the same position

points in the opposite direction and forms an intra-molecular hydrogen bond (Johnston et al.,

2005). This removes the conformational restraint on the GG-loop region, allowing it to move

closer towards the GTPγS bound state.

It was suggested that the KB-752 peptide accelerates the rate of nucleotide exchange

by disrupting the 'seatbelt' that was observed in the Gβγ–bound structure of $G\alpha_{i1}$ (Wall et al.,

1995). It was also suggested that KB-752 peptide opens up the nucleotide binding pocket by

pulling the β2/α3 loop (GG-loop) away from the nucleotide binding pocket (Johnston et al.,

2005). In the structure of $G\alpha_{i1}$:R6A complex the opposite phenomena is observed, where the

GG-loop is pushed towards the nucleotide binding pocket. The mean distance of the Cα

atoms in GG-loop and the β-phosphate of the bound GDP is 7.2 Å in the R6A-bound

structure, but increases to 9.3 Å in the KB-752-bound structure. The structure of $G\alpha_{i1}$:R6A

complex, in conjunction with the structure of $G\alpha_{i1}$:KB752 complex provides crystallographic

snapshots of the potential conformational changes that occur during nucleotide exchange

(Figure 6.19). It also suggests that conformation of GG-loop is perhaps a crucial determinant

of the rate of GDP release from $G\alpha_{i1}$.
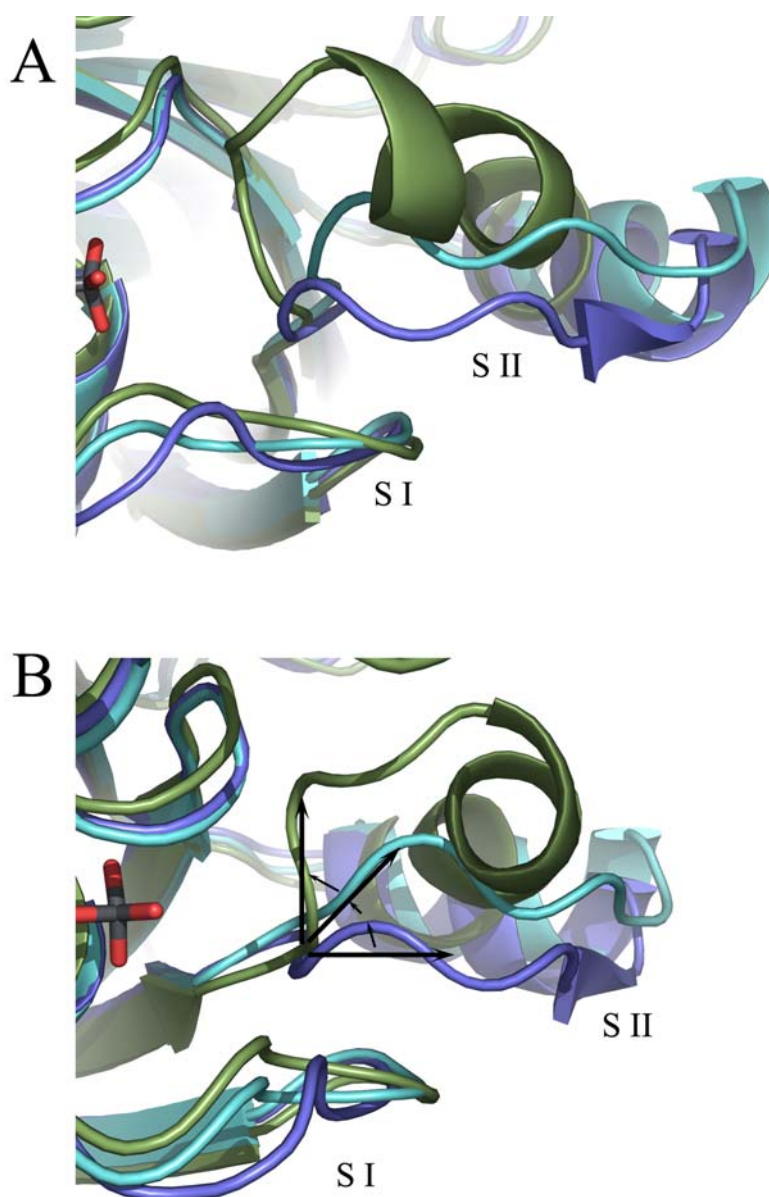
Figure 6.19. Crystallographic snapshots of conformational changes during nucleotide exchange. The structures of $G\alpha_{i1}$:R6A (blue), $G\alpha_{i1}$:KB-752 (cyan), and $G\alpha_{i1}\cdot GTP\gamma S$ (green) are superimposed and shown in ribbon representation; switch I (S I) and switch II (S II) regions are displayed. *A*. GG-loop in KB-752 bound structure is positioned away from the bound GDP. *B*. The displacement of GG-loop in the $G\alpha_{i1}\cdot GTP\gamma S$ structure in relation to the R6A and KB-752 bound structures are shown by arrows. The putative movement of GG-loop during nucleotide exchange is shown by dashed arrows.

6.3.7   Mutational analysis of the peptide: A fine balance between GDI and GEF

In light of the observation that the orientation of the GG-loop affects the rate of GDP release, altering the conformation of this loop might provide a way of probing the structural processes involved in nucleotide exchange. While crystal contacts might influence the orientation of GG-loop in the KB-752 bound structure of Gα_{i1}, it appeared that the difference of the peptides at position four played an important role. Since, KB-752 has a threonine residue at position four, mutating the Tyr4 residue of R6A to a smaller residue might convert it into a GEF. The Tyr4 residue of R6A was mutated either to a threonine (R6A-T) or a glycine (R6A-G), with the anticipation that these mutations would abolish the GDI activity of R6A peptides (Figure 6.20).

```
R6M1:      DQLYWWEYLKK
R6A-G:     DQLGWWEYLKK
R6A-T:     DQLTWWEYLKK
KB752:     SRVTWYDFLMEDTKSR
```

Figure 6.20. Sequence alignment of mutant R6A peptides.

The effect of these mutant peptides on the binding of BODIPY-GMPPNP binding to Gα_{i1} was monitored by fluorescence spectroscopy. Not only did these mutations abolish the GDI activity of R6A, but the mutant peptides also accelerated the rate of binding of BODIPY-GMPPNP to Gα_{i1} (Figure 6.21). It is striking that the activity of the R6A peptide was dramatically altered by the mutation of a single residue. The mutation of Tyr4 to Thr or Gly probably allows GG-loop to move away from switch I and the nucleotide binding pocket

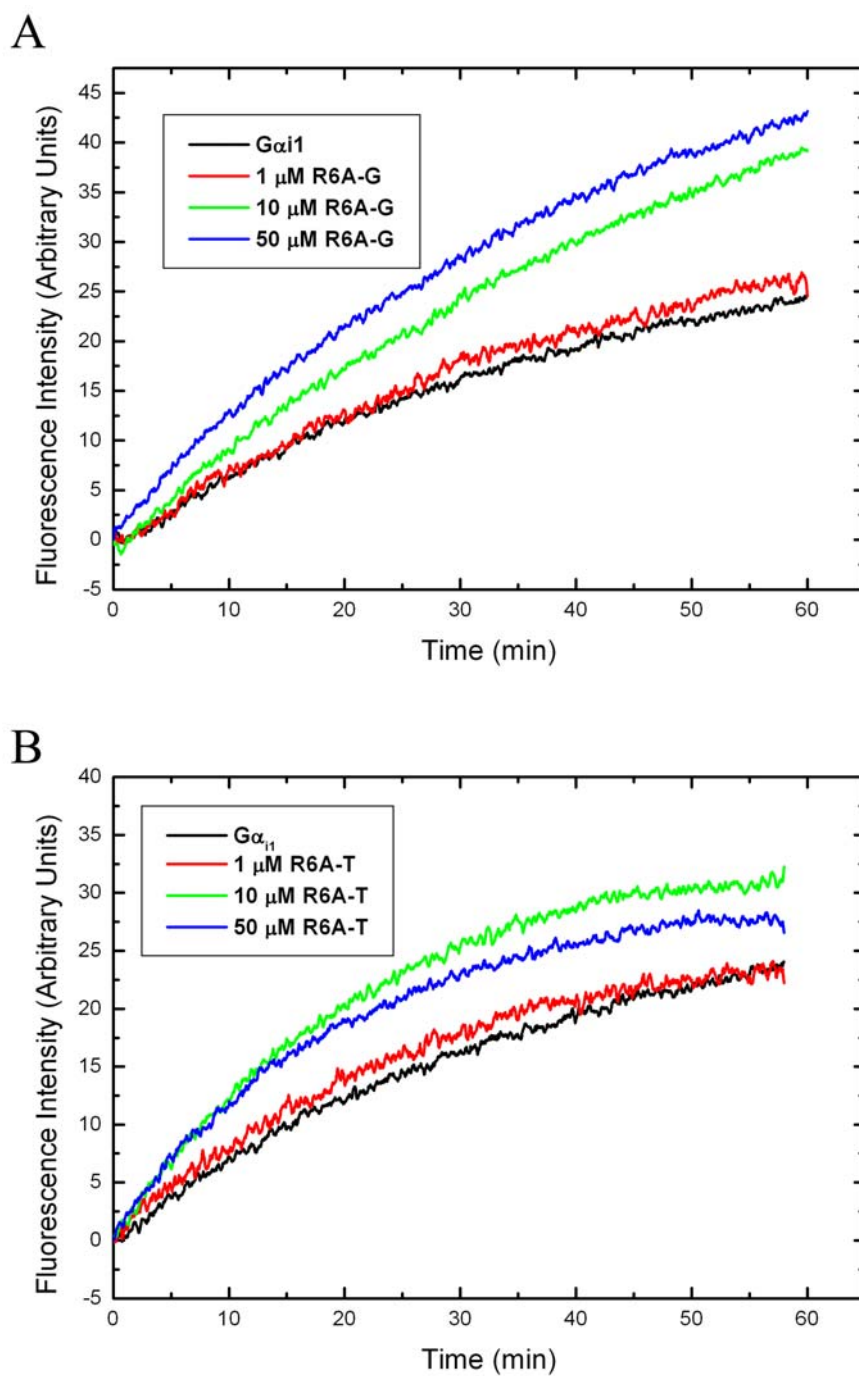**Figure 6.21. GEF activity of mutant R6A peptides.** The binding of 1 μM BODIPY-GMPPNP to 200 nM Gα$_{i1}$ (black) at 25 ^{o}C in the presence of: A, 1 μM (red), 10 μM (green), and 50 μM (blue) R6A-G; B, 1 μM (red), 10 μM (green), and 50 μM (blue) R6A.

, which allows the mutant peptides to behave like KB-752. It is interesting to note that R6A-G appears to be a stronger GEF than R6A-T; suggesting that a Gly residue at position 4 might allow even more conformational flexibility of the GG loop. The initial rates of BODIPY-GMPPNP binding to $G\alpha_{i1}$ at 10 and 50 µM R6A-T was similar. This might indicate that 10 µM is saturating under the assay conditions. Alternately, it could mean that R6A-T associates with the nucleotide-free intermediate during nucleotide exchange and release of R6A-T from the complex becomes rate limiting. The biochemical data shown here is still at a preliminary stage; more experiments are required to obtain a better biochemical description of the regulation of nucleotide exchange by R6A peptides.

The preliminary biochemical analysis of the mutant peptides suggests that nucleotide exchange rate can be modulated by simply reorienting the GG-loop. This observation lends support to the 'lever arm' model (discussed in chapter 1.1.5) of nucleotide exchange (Iiri et al., 1998). The structures of KB-752 and R6A-bound $G\alpha_{i1}$ suggests that the conformation of GG-loop is an important determinant of the rate of GDP release from the nucleotide binding pocket. Another part of $G\alpha_{i1}$ has been implicated in nucleotide exchange as well; the $\alpha5$ helix and $\beta6$-$\alpha5$ loop play important role in regulating GDP release (Iiri et al., 1998; Iiri et al., 1994). Several mutations in the $\alpha5$ helix and $\beta6$-$\alpha5$ loop are known to accelerate nucleotide exchange (Marin et al., 2001; Posner et al., 1998). These mutants are though to act allosterically to lower the affinity of $G\alpha$ for GDP and possibly mimic the activated GPCR bound conformation of $G\alpha$ (Marin et al., 2001). It was also shown that GPCRs interact with the C-terminal region of $G\alpha$ and possibly destabilize the GDP binding pocket by pulling the $\alpha5$ helix (Hamm, 2001; Iiri et al., 1998). In addition, GPCRs can potentially utilize $G\beta\gamma$ as a

lever to move the $\beta3$-$\alpha2$ /GG-loop away from the nucleotide binding pocket (Iiri et al.,

1998). It seems likely that GPCRs use both mechanisms synergistically to mount a concerted

assault on the GDP binding pocket.

6.4   References

Bailey, S. (1994). The CCP4 suite- programs for protein crystallography. Acta Cryst *D50*, 760-763.

Brünger, A. T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472-475.

Brünger, A. T., Adams, P. D., Clore, G. M., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J.*, et al.* (1998). Crystallography and NMR system (CNS): A new software suite for macromolecular structure determination. Acta Cryst *D54*, 905-921.

Chidiac P, M. V., Ross EM. (1999). Kinetic control of guanine nucleotide binding to soluble Galpha(q). Biochem Pharmacol *58*, 39-48.

Coleman, D. E., Lee, E., Mixon, M. B., Linder, M. E., Berghuis, A., Gilman, A. G., and Sprang, S. R. (1994). Crystallization and preliminary crystallographic studies of $G_{i\alpha1}$ and mutants of $G_{i\alpha1}$ in the GTP and GDP-bound states. JMolBiol *238*, 630-634.

Coleman, D. E., and Sprang, S. R. (1998). Crystal structures of the G-protein Gia1 complexed with GDP and Mg^{2+}: A crystallographic titration experiment. Biochemistry *37*.

Frishman, D., and Argos, P. (1995). Knowledge-based protein secondary structure assignment. Proteins *23*, 566-579.

Guex, N., Diemand, A., and Peitsch, M. C. (1999). Protein modelling for all. Trends Biochem Sci *24*, 364-367.

Hamm, H. E. (2001). How activated receptors couple to G proteins. Proc Natl Acad Sci U S A *98*, 4819-4821.

Higashijima, T., Ferguson, K. M., Sternweis, P. C., Ross, E. M., Smigel, M. D., and Gilman, A. G. (1987a). The effect of activating ligands on the intrinsic fluorescence of guanine nucleotide-binding regulatory proteins. J Biol Chem *262*, 752-756.

Higashijima, T., Ferguson, K. M., Sternweis, P. C., Smigel, M. D., and Gilman, A. G. (1987b). Effects of Mg2+ and the beta gamma-subunit complex on the interactions of guanine nucleotides with G proteins. J Biol Chem *262*, 762-766.

Iiri, T., Farfel, Z., and Bourne, H. R. (1998). G-protein diseases furnish a model for the turn-on switch. Nature *394*, 35-38.

Iiri, T., Herzmark, P., Nakamoto, J. M., van Dop, C., and Bourne, H. R. (1994). Rapid GDP release from Gs alpha in patients with gain and loss of endocrine function. Nature *371*, 164-168.

Ja, W. W., Adhikari, A., Austin, R. J., Sprang, S. R., and Roberts, R. W. (2005). A peptide core motif for binding to heterotrimeric G protein alpha subunits. J Biol Chem *280*, 32057-32060.

Ja WW, R. R. (2004). In vitro selection of state-specific peptide modulators of G protein signaling using mRNA display. Biochemistry *43*, 9265-9275.

Johnston CA, W. F., Jezyk MR, Fredericks Z, Bodor ET, Jones MB, Blaesius R, Watts VJ, Harden TK, Sondek J, Ramer JK, Siderovski DP. (2005). Structure of Galpha(i1) bound to a GDP-selective peptide provides insight into guanine nucleotide exchange. Structure (Camb) *13*, 1069-1080.

Johnston, C. A., Willard, F. S., Jezyk, M. R., Fredericks, Z., Bodor, E. T., Jones, M. B., Blaesius, R., Watts, V. J., Harden, T. K., Sondek, J*., et al.* (2005). Structure of Galpha(i1) bound to a GDP-selective peptide provides insight into guanine nucleotide exchange. Structure (Camb) *13*, 1069-1080.

Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991). Improved methods for the building of protein models in electron density maps and the location of errors in these models. Acta Crystallogr *A47*, 110-119.

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers *22*, 2577-2637.

Kimple, R. J., Kimple, M. E., Betts, L., Sondek, J., and Siderovski, D. P. (2002). Structural determinants for GoLoco-induced inhibition of nucleotide release by Galpha subunits. Nature *416*, 878-881.

Kimple RJ, W. F., Siderovski DP (2002). The GoLoco motif: heralding a new tango between G protein signaling and cell division. Mol Interv *2*, 88-100.

Lambright, D. G., Noel, J. P., Hamm, H. E., and Sigler, P. B. (1994). Structural determinants for activation of the alpha-subunit of a heterotrimeric G protein. Nature *369*, 621-628.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography *26*, 283-291.

Lee, B., and Richards, F. M. (1971). The interpretation of protein structure: estimation of static accessibility. JMolBiol *55*, 379-400.

Marin, E. P., Krishna, A. G., and Sakmar, T. P. (2001). Rapid activation of transducin by mutations distant from the nucleotide- binding site: evidence for a mechanistic model of receptor-catalyzed nucleotide exchange by G proteins. J Biol Chem *276*, 27400-27405.

McCoy AJ, G.-K. R., Storoni LC, Read RJ. (2005). Likelihood-enhanced fast translation functions. Acta Crystallogr D Biol Crystallogr, 458-464.

Mixon, M. B., Lee, E., Coleman, D. E., Berghuis, A. M., Gilman, A. G., and Sprang, S. R. (1995). Tertiary and quaternary structural changes in $G_{i\alpha1}$ induced by GTP hydrolysis. Science *270*, 954-960.

Otwinowski, Z. (1993). Oscillation data reduction program, In Data collection and processing, N. I. L. Sawyer, and S. W. Bailey, eds. (Daresbury, U.K.: Science and Engineering Council Daresbury Laboratory), pp. 56-62.

Otwinowski, Z., and Minor, W. (1997). Processing of x-ray diffraction data collected in oscillation mode. Methods in Enzymology *276*, 307-326.

Posner, B. A., Mixon, M. B., Wall, M. A., Sprang, S. R., and Gilman, A. G. (1998). The A326S mutant of $G_{i\alpha1}$ as an approximation of the receptor-bound state. J Biol Chem *273*, 21752-21758.

Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. JMolBiol *7*, 95-99.

Ramachandran, G. N., and Sassiekharan, V. (1968). Conformation of polypeptides and proteins. Advances in Protein Chemistry *28*, 283-437.

Read, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. Acta Crystallographica *A42*, 140-149.

Sprang, S. R. (1997). G protein mechanisms: Insights from structural analysis. Annu Rev Biochem *66*, 639-678.

Wall, M. A., Coleman, D. E., Lee, E., Iñiguez-Lluhi, J. A., Posner, B. A., Gilman, A. G., and Sprang, S. R. (1995). The structure of the G protein heterotrimer $G_{i\alpha1}\beta_1\gamma_2$. Cell *80*, 1047-1058.

Willard, F. S., Kimple, R. J., and Siderovski, D. P. (2004). Return of the GDI: the GoLoco motif in cell division. Annu Rev Biochem *73*, 925-951.

Zelent, B., Veklich, Y., Murray, J., Parkes, J., Gibson, S., and Liebman, P. (2001). Rapid irreversible G protein alpha subunit misfolding due to intramolecular kinetic bottleneck that precedes Mg2+ "lock" after GTP/GDP exchange. Biochemistry *40*, 9647-9656.

Chapter Seven

Conclusion and future directions

Nucleotide exchange is one of the most important aspects of heterotrimeric G protein signaling. Even after more than two decades of extensive studies, the precise molecular mechanism of nucleotide exchange remains elusive. To gain better understanding of the regulation of nucleotide exchange, I focused on investigating GDIs, which inhibit nucleotide exchange by stabilizing the GDP-bound conformation of $G\alpha$. To that end, my dissertation research focused on two molecules: AGS3 and R6A. To characterize the binding of AGS3 and $G\alpha_{i1}$, the thermodynamic properties of the interaction between AGS3 and $G\alpha_{i1}$ was studied using ITC. In another approach, R6A was utilized as a molecular probe to identify elements in $G\alpha_{i1}$ that are important in regulation of nucleotide exchange. The conclusions from both of these projects have been discussed in the previous chapters. In this chapter, I will focus more on discussing my ideas about future projects that are logical continuation of the ideas developed in this dissertation.

Future studies on AGS3

The biochemical studies of AGS3 and $G\alpha_{i1}$ interaction provide a detailed picture about the role played by the four GoLoco motifs of AGS3 in influencing its behavior. ITC studies not only allowed the determination of the binding stoichiometry and affinity of AGS3

towards $G\alpha_{i1}$, it also revealed the biochemical consequence of having four GoLoco motifs in one protein. AGS3-C possesses two high-affinity and two low-affinity sites for binding to $G\alpha_{i1}$. Analysis of the binding affinities of isolated GoLoco motifs suggested that this behavior was not a consequence of sequence differences among the GoLoco motifs of AGS3. Because, deletion of the most C-terminus GoLoco motif eliminated the biphasic binding behavior of AGS3, it is likely that tertiary organization of GoLoco motifs in AGS3-C influenced the biphasic behavior. The peptides corresponding to the GoLoco motifs of AGS3-C were weaker compared to AGS3-C in terms of their affinity for $G\alpha_{i1}$ and GDI activity. But, the residues flanking the GoLoco motifs of AGS3, increased the affinity and potency of these motifs to the level of AGS3 itself. While, this revealed the role of the flanking sequences, it did not explain the presence of two distinct and independent sets of binding sites in the four-GoLoco version of AGS3. Several questions remain unanswered about the interaction of AGS3-C and $G\alpha_{i1}$:

1. What is the tertiary organization of GoLoco motifs in AGS3-C?

2. What determines the $G\alpha$ selectivity of the GoLoco motifs in AGS3?

3. What is the structural basis of two sets of independent binding sites?

The C-terminal residues of GoLoco motifs are important in determining the binding specificity of GoLoco motifs, but these residues are poorly conserved among the GoLoco motifs in AGS3 despite their specificity towards $G\alpha_i$. The molecular basis of GoLoco specificity can be addressed by determining the crystal structures of $G\alpha_{i1}$ in complex with the extended GPR peptides (described in chapter 4). Several attempts to crystallize with $G\alpha_{i1}$ in complex with GPR1ex were not successful. However, other extended GPR peptides were not

evaluated as crystallization targets. Thus, $G\alpha_{i1}$:GPR2ex/GPR3ex complexes are potential

crystallization targets. The structure of $G\alpha_{i1}$:GPR3ex might also reveal the molecular basis of

the requirement of flanking sequences for its activity.

AGS3-C in itself or in complex with $G\alpha_{i1}$ are poor crystallization targets. However, in

the full-length AGS3, the C-terminal GoLoco domain is likely to be stabilized by its

interaction with the N-terminal TPR domain (discussion in chapter 1). Thus, the full-length

AGS3 or LGN can be good crystallization targets; it is also more biologically relevant than

AGS3-C. Preliminary expression studies suggested that full-length AGS3 is not expressed in

E.coli, but the TPR-rich and GoLoco-rich domains in isolation can be expressed in *E.coli*.

However, the expression of the TPR-rich domain was poor and needs optimization. The TPR

and GoLoco-rich domains, once purified, can be used to make a complex that can potentially

resemble the structure of full-length AGS3. AGS3-C and its deletion constructs are relatively

small in size and can be purified in high amounts. Thus, they can be good candidates for

NMR-based structural studies. Although, CD studies and structure predictions indicated large

segments of disordered loop like regions in AGS3-C, NMR studies can be used to validate

those predictions.

While, the effect of C-terminal truncations on the binding behavior of AGS3-C was

studied, N-terminal truncation constructs were not made. It would be worthwhile to

investigate whether N-terminal truncations affect the two-binding site behavior of AGS3-C

in the same way as C-terminal truncations do. This approach might help to elucidate the

physical basis of the two sets of binding sites observed in AGS3-C. Alternately, if a construct

containing four identical GoLoco sequences exhibits two sets of binding sites, it would suggest that the biphasic behavior is solely contributed by the tertiary structure.

Future studies with R6A peptides

The structure of $G\alpha_{i1}$:R6A complex along with the mutagenesis data suggests that the conformation of GG-loop is a critical determinant of the rate of GDP release from $G\alpha_{i1}$. This provides support to the 'lever arm' model of nucleotide exchange proposed by Bourne and colleagues. While this is encouraging, the biochemical activity of R6A needs further investigation.

The biphasic behavior of nucleotide exchange kinetics in the presence of R6A needs to be addressed. The biphasic behavior is indicative of two distinct processes; a fast GTP binding process and a slow process where $G\alpha_{i1}$ poorly participates in binding to GTP. It raises the important question about which of these two processes actually represent the effect of R6A peptides. The crystal structure of R6A represents a GDP-bound form, but does not eliminate the possibility of other modes of interaction between $G\alpha_{i1}$ and R6A. Because, the activity of R6A peptides showed considerable variation depending on the nature of the GTP-analog used in the assay, it raises the possibility that conditions in the assay influence the behavior of R6A. To eliminate these influences, assays can be performed to measure the rate of GDP release itself, instead of measuring the rate of GTP binding. This can be performed by using $G\alpha_{i1}$ that is pre-loaded with radio-labeled GDP and the rate of GDP release can be monitored by performing competition assays with unlabeled GDP. Alternately, GTPγS

binding assays could be performed for all of the R6A peptides using 7-AW labeled $G\alpha_{i1}$, which eliminates the possible influence of BODIPY and MANT fluorophores.

Although the differences in the GG-loop conformation between $G\alpha_{i1}$:KB752 and $G\alpha_{i1}$:R6A complexes appear significant (rmsd ~ 3 Å), the role of crystal contacts in influencing these differences cannot be eliminated. The N-terminus of the KB-752 peptide is involved in making inter-subunit contacts in the $G\alpha_{i1}$:KB-752 crystals. Besides, KB-752 possesses six additional residues at its C-terminus that are not present in R6A. Thus, the structure of $G\alpha_{i1}$:R6A complex should be compared to the structures of $G\alpha_{i1}$:R6A-G/T complexes to evaluate the conformational changes in the GG-loop in an identical crystallographic context. Determination of the crystal structure of $G\alpha_{i1}$ in complex with R6A-G or R6A-T should be a priority.

Switch II adopts a stable conformation in both R6A and KB-752 bound structures, suggesting that stability of switch II is required for GDI and GEF activity. But the relative stabilities of switch II in the GEF and GDI bound state cannot be inferred from the crystal structures. It is possible that switch II would be less stable in GEF peptide bound state, to allow a more energetically favorable transition to the GTP-bound state. The 9-residue R6A core has weaker affinity for $G\alpha_{i1}$ compared to the 16-residue R6A4 peptide. But the affinities of the R6A-T and R6A-G mutants for $G\alpha_{i1}$ have not been measured yet. If the mutants have lower affinity compared to R6A, it might indicate that stability of switch II is lower in the GEF-bound state in comparison with the GDI-bound state.

Nucleotide exchange is associated with dynamic changes in the structure of $G\alpha_{i1}$. The crystal structures of $G\alpha_{i1}$ in complex with R6A or KB-752 provide only static snapshots of a

particular state. It is likely that binding of these peptides will influence the dynamics of $G\alpha_{i1}$, and switch II region in particular. To investigate, the effect of R6A and KB-752 binding on the dynamics of $G\alpha_{i1}$, molecular dynamics simulation studies were initiated with the help of Arne Strand in the lab. Preliminary simulations indicated higher mean rmsd of KB-752 bound structure in comparison to the R6A bound structure. Interestingly, the simulation suggests that in the $G\alpha_{i1}$:KB-752 complex parts of switch II and the peptide itself is more labile compared to the same regions in the $G\alpha_{i1}$:R6A complex. However, the differences in the mean rmsd of $G\alpha_{i1}$:R6A and $G\alpha_{i1}$:KB-752 during molecular dynamics simulation was very small (0.2 Å). The interpretation of the molecular dynamic simulations is at a preliminary stage and more sophisticated analysis of the data is needed.

Although, the structure of $G\alpha_{i1}$:R6A complex sheds light on the role of GG-loop, binding of R6A did not influence the structure of the C-terminal region of $G\alpha_{i1}$, which is also important for nucleotide exchange. The C-terminal $\alpha 5$ helix and $\beta 6$-$\alpha 5$ loop are utilized by GPCRs. Several mutations in these regions destabilize the GDP bound state and strongly accelerate the basal rate of GDP release. These mutations are thought to mimic the receptor bound conformation of $G\alpha$. The activity of R6A peptides can be studied in these mutant backgrounds. If the GEF peptides accelerate the rate of GDP release further in these mutant backgrounds, it would validate the notion that GPCRs utilize both $\beta\gamma$-lever and the C-terminus synergistically. The mechanism of GPCR activity is probably the most poorly understood process in heterotrimeric G-protein signaling. Structural probes like R6A would perhaps, enhance our understanding of GPCR mediated nucleotide exchange in the near future.

VITAE

Anirban Adhikari was born in the city of Tartous, Syria, on December 23, 1976, the son of Tulsi Prasad Adhikari and Sunanda Adhikari. Anirban moved to Tamluk, West Bengal (India) in 1983 and graduated from Tamluk Hamilton High School in 1995. He received the degree of Bachelor of Technology (Honours) from Indian Institute of Technology, Kharagpur (India), in 1999. While an undergraduate, he had the opportunity to work with Dr. Hemanta Majumder at the Indian Institute of Chemical Biology, Calcutta and also worked as summer intern in the computational chemistry division of AstraZeneca R&D, Bangalore, under the supervision of Dr. Indira Ghosh. Anirban entered Division of Cellular and Molecular Biology at the University of Texas Southwestern Medical Center at Dallas in 1999 and was admitted to candidacy in the Molecular Biophysics graduate program in 2001. On December 8, 2001, he married Mekhala Maiti who is currently a graduate student in the Genetics and Development program at University of Texas Southwestern Medical Center. His parents live in Tamluk, India, and he has a brother who lives in Bristol, U.K.

Permanent Address:
C/o Tulsi Prasad Adhikari
Tamluk-721636
West Bengal
India