ALLOSTERIC DETERMINANTS OF GUANINE NUCLEOTIDE BINDING PROTEINS AND METHODS TO CRYSTALLIZE THE CYTOSOLIC DOMAINS OF ADENYLYL CYCLASE

APPROVED BY SUPERVISORY COMMITTEE

Alfred G. Gilman M.D., Ph.D.

Paul C. Sternweis Ph. D.

Rama Ranganathan M.D., Ph.D.

Stephen R. Sprang Ph.D.

ACKNOWLEDGMENTS

Many people contributed to making my graduate training an educational and enjoyable experience. It became easy to focus on my research and education while many fantastic people took care of everything else required to keep enrolled. Wendy Deaner and Carla Childers in the pharmacology department skillfully took care of any administrative tasks that were required. As well, Wendy and Carla are experts at efficiently maneuvering through the "system" to insure that the lab had the necessary tools to perform our experiments. Robin Downing in the MSTP office was extremely helpful throughout my training by answering thousands of questions. Surely, a forest worth of paperwork was used to keep me enrolled.

The members of the Gilman lab were a fantastic group of people that established an environment that hardly seemed like work. The technical staff (Helen Aronovich, Linda Hannigan, Jeff Laidlaw, Julie Collins, Eduardo Arteaga, Pam Sternweis, Kevin Vale and Federico Franco) were the constant features that kept the lab operating efficiently. I specifically would like to thank Michelle Jennings for her extremely positive attitude and willingness to eagerly tackle any task asked of her. Without her dedication, I surely would still be finishing experiments. I am sincerely grateful for all her hard work. The scientists in the laboratory provided an intellectual, stimulating environment. Roger Sunahara took me under his wing when I joined the lab and guided me through the first several years of my graduate studies. He always had time to answer my many questions and listen to my ideas. The group of post-docs when I left the lab (Andrejs Krumins, Scott Gibson, and Greg Tall) were especially helpful and supportive. Scott Gibson was a great person to share a lab space and an even better friend. I especially thank Scott and Tara Beers Gibson for critically reading this tome of a dissertation and providing excellent suggestions.

Several faculty members were exceptionally helpful during my training. First I thank my dissertation committee including Paul Sternweis, Stephen Sprang, David Garbers, and Rama Ranganathan. They were very helpful in keeping me focused on my project. My dissertation committee along with Joel Goodman made the qualifying exam an enjoyable experience. I would like to thank Jean Wilson, Rod Ulane, Robert Munford, and Mike Brown for their leadership of the Medical Scientist Training Program that brought me to Dallas.

My collaboration with Rama Ranganathan and his laboratory was one of the highlights of my training. The members of his lab were very tolerant in allowing me to use their equipment, participate in the energetics small group discussions, and at times monopolizing Rama's time. I thank Mark Wall for all the time spent structrure gazing and discussing results. I especially thank Steve Lockless for the exciting collaboration. Rama provided an infectious enthusiasm for the experiments to test our hypotheses. I hope that I can maintain this enthusiasm for my career.

Most of all I thank my mentor Al Gilman. I appreciate his willingness to gamble on taking a graduate student with little experience and his patience and guidance during my education. He has encouraged me to be methodical and meticulous in my approach to scientific problems. Throughout my training, he has been constructively critical and unrelentingly supportive of my work. He has provided me the ideal model of both a scientist and a mentor. The education I received while in the Gilman laboratory was of indescribable value. It was a tremendous privilege to have trained in his laboratory. Thanks Al.

My family and friends both in Oklahoma and here in Dallas have been very supportive throughout this training. My parents, Mike and Jo Anne Hatley, have always encouraged me to work hard and to pursue a career that I enjoy. My wife Ann, whom I met and married while in the lab, provided the encouragement to get me through this portion of my training.

ALLOSTERIC DETERMINANTS OF GUANINE NUCLEOTIDE BINDING PROTEINS AND METHODS TO CRYSTALLIZE THE CYTOSOLIC DOMAINS OF ADENYLYL CYCLASE

by

MARK EDWARD HATLEY

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

June, 2004

ALLOSTERIC DETERMINANTS OF GUANINE NUCLEOTIDE BINDING PROTEINS AND METHODS TO CRYSTALLIZE THE CYTOSOLIC DOMAINS OF ADENYLYL CYCLASE

Publication No. _____

Mark Edward Hatley

The University of Texas Southwestern Medical Center at Dallas, 2004

Supervising Professor: Alfred G. Gilman, M.D., Ph.D.

The cytosolic domains of mammalian adenylyl cyclases, termed C_1 and C_2, are responsible for catalytic activity and most regulatory properties. Crystal structures of the soluble catalytic core of adenylyl cyclase bound to activators $G_{s\alpha}$ and forskolin were previously determined. However, structural information regarding low activity (non-$G_{s\alpha}$ or forskolin bound) states of the enzyme is lacking. Genetic and biochemical methods were utilized to overcome the low affinity of the cytosolic domains in the absence of activators. A genetic screen in *Saccharomyces cerevisiae* identified mutations that activate mammalian adenylyl cyclase in the absence of $G_{s\alpha}$. The increased affinity of the K1014N-C_2 mutant protein for the C_1 domain in the absence of $G_{s\alpha}$ was exploited to isolate a complex containing C_1 and C_2 in the absence of $G_{s\alpha}$**.** Unfortunately, this complex crystallized but

failed to diffract due to heterogeneity. Intein-mediated protein ligation and expression of a C_1-C_2 fusion protein in adenylyl cyclase deficient *Escherichia coli* were explored to circumvent the low affinity of the domains. However, the yields of products were insufficient for crystallization.

Members of the G protein superfamily contain nucleotide-dependent switches that dictate the specificity of their interactions with binding partners. Using a new sequence-based method termed statistical coupling analysis (SCA), I identified the allosteric core of these proteins – the network of amino acid residues that couples the domains responsible for nucleotide binding and protein-protein interactions. One-third of the 38 residues identified by SCA were mutated in the G protein $G_{s\alpha}$, and the interactions of GTPγS- and GDP-bound mutant proteins were tested with both adenylyl cyclase (preferential binding to GTP-$G_{s\alpha}$) and the G protein βγ subunit complex (preferential binding to GDP-$G_{s\alpha}$). A two-state allosteric model predicts that mutation of residues that control the equilibrium between GDP- and GTP-bound conformations of the protein will cause the ratio of affinities of these species for adenylyl cyclase and βγ to vary in a reciprocal fashion. Observed results were consistent with this prediction. The network of residues identified by the SCA appears to comprise a core allosteric mechanism conferring nucleotide-dependent switching; the specific features of different G protein family members are built upon this core.

TABLE OF CONTENTS

ACKNOWLEDGMENTS ii
ABSTRACT vi
TABLE OF CONTENTS viii
PUBLICATIONS PRESENTED IN THIS DISSERTATION x
LIST OF FIGURES xi
LIST OF TABLES xiv
LIST OF ABBREVIATIONS xv

CHAPTER 1 INTRODUCTION 1

G protein regulatory cycle 2
G protein α subunit 3
G protein βγ subunits 6
Mammalian adenylyl cyclases 8
Strutcture of Gsα and the cytoplasmic domains of adenylyl cyclase 12
Statement of purpose 14

CHAPTER 2 ISOLATION, CHARACTERIZATION, AND CRYSTALLIZATION OF CONSTITUTIVELY ACTIVE MUTANTS OF MAMMALIAN ADENYLYL CYCLASE 24
Introduction 24
Experimental procedures 25
Results 29
Discussion 34

CHAPTER 3 INTEIN-MEDIATED PROTEIN LIGATION OF ADENYLYL CYCLASE CYTOSOLIC DOMAINS 50
Introduction 50
Experimental procedures 52
Results 57
Discussion 59

CHAPTER 4 COMPLEMENTATION SCREEN FOR LINKED ADENYLYL CYCLASE CYTOSOLIC DOMAINS 70
Introduction 70
Experimental procedures 72
Results 75
Discussion 78

CHAPTER 5 MAPPING THE ALLOSTERIC DETERMINANTS OF GUANINE NUCLEOTIDE BINDING PROTEINS 84
Introduction 84
Experimental procedures 89
Results 96
Discussion 106

CHAPTER 6 CONCLUSIONS AND FUTURE DIRECTIONS 146

Crystallization of the low activity state if adenylyl cyclase catalytic core 146
Allosteric determinants of G proteins 148

REFERENCES 154

VITA 164

PRIOR PUBLICATIONS

Hatley M.E., Benton B.K., Xu J., Manfredi J.P., Gilman A.G., Sunahara R.K. *Isolation and characterization of constitutively active mutants of mammalian adenylyl cyclase.* Journal Biological Chemistry, 2000; 275(49): 38626-32.

Hatley M.E., Gilman A.G., Sunahara R.K. *Expression, purification, and assay of cytosolic (catalytic) domains of membrane-bound mammalian adenylyl cyclases.* Methods in Enzymology. 2002; 345:127-40.

Hatley, M.E., Lockless, S.W., Gibson, S.K., Gilman, A.G., Ranganathan, R. *Allosteric determinants of guanine nucleotide binding proteins*. Proceedings of the National Academy of Science. 2003; 100: 14445-50.

LIST OF FIGURES

FIGURE 1-1 Guanine nucleotide cycle of heterotrimeric G proteins 16

FIGURE 1-2 Structure of $G_{i\alpha1}$:GTPγS:Mg^{2+} complex 17

FIGURE 1-3 Conformational changes associated transition state of GTP hydrolysis 18

FIGURE 1-4 Structure of G_i heterotrimer and βγ subunits 19

FIGURE 1-5 Conformations of GTPγS- and GDP-bound $G_{i\alpha1}$ 20

FIGURE 1-6 Predicted topology of mammalian adenylyl cyclase 21

FIGURE 1-7 Structure of adenylyl cyclase catalytic core complexed to $G_{s\alpha}$-GTPγS 22

FIGURE 1-8 Model of substrate ATP in active site of adenylyl cyclase 23

FIGURE 2-1 The apparent affinity of IIC_2-K1014N for VC_1 and $G_{s\alpha}$-GTPγS 43

FIGURE 2-2 Substrate kinetics of IIC_2-K1014N 44

FIGURE 2-3 Gel filtration of VC_1, IIC_2-K1014N and $G_{s\alpha}$ in the absence of forskolin 45

FIGURE 2-4 Complexes of VC_1 and IIC_2 46

FIGURE 2-5 Purification of H6-VC1(580):IIC-K1014N:Forskolin:2'd3'-AMP•PPi complex ... 47

FIGURE 2-6 Crystals of VC1:IIC2-K1014N:Forskolin:2'd3'-AMP:PPi 48

FIGURE 2-7 Catalytic core of adenylyl cyclase 49

FIGURE 3-1 Overview of intein-mediated protein splicing 63

FIGURE 3-2 Mechanism of Protein Splicing 64

FIGURE 3-3 Schematic of the purification and ligation of VC_1 and IIC_2 65

FIGURE 3-4 Purification of VC_1 from the intein fusion protein 66

FIGURE 3-5 Characterization of VC_1 intein fusion protein 67

FIGURE 3-6 Activity of Cys-IIC_2 after cleavage with factor Xa 68

FIGURE 3-7 Intein-mediated protein ligation of VC_1 and IIC_2 69

FIGURE 4-1 Phenotypic complementation of adenylyl cyclase-deficient TP2000 *Escherichia coli*... ... 81

FIGURE 4-2 Phenotypic complementation of adenylyl cyclase deficient TP2000 and TP610A *Escherichia coli* constitutively expressing Lac repressor 82

FIGURE 4-3 Phenotypic complementation of adenylyl cyclase-deficient *Escherichia coli* with linked adenylyl cyclase domains 83

FIGURE 5-1 Schematic for extracting coupling information from multiple sequence alignment120

FIGURE 5-2 Mock clustering experiment 121

FIGURE 5-3 Mean $\Delta\Delta G^{stat}$ for least conserved positions following random elimination of sequences from the MSA 122

FIGURE 5-4 Statistical coupling analysis of the GTP binding protein family 123

FIGURE 5-5 Mapping statistically coupled positions on the G alpha subunit 124

FIGURE 5-6 Illustrating non-coupled controls on the G alpha subunit 125

FIGURE 5-7 $\beta\gamma$ affinity assay 126

FIGURE 5-8 Characterization of Q227A and R201A (class 1) 127

FIGURE 5-9 Determination of nucleotide bound to Q227A-$G_{s\alpha}$ 128

FIGURE 5-10 Characterization of G225A and R228A (class 2) 129

FIGURE 5-11 Characterization of W234F and I276A (class 2) 130

FIGURE 5-12 Characterization of R42A and A48H (class 2) 131

FIGURE 5-13 Characterization of C237A and V224A (class 3) 132

FIGURE 5-14 Characterization of C365A (class 3) 133

FIGURE 5-15 Characterization of E50A and F222A (class 4) 134

FIGURE 5-16 Characterization of S205A and I207A (non-coupled controls) 135

FIGURE 5-17 Characterization of D229A and L272A (non-coupled controls) 136

FIGURE 5-18 Characterization of E230A and L46A (non-coupled controls) 137

FIGURE 5-19 Ratios of affinities of mutant and wild type $G_{s\alpha}$ proteins for binding adenylyl cyclase and $\beta\gamma$ 138

FIGURE 5-20 Correlation of nucleotide sensitivities for adenylyl cyclase and βγ 139

FIGURE 5-21 Two state equilibrium model for allosteric regulation of G protein 140

FIGURE 5-22 Illustration of three important features of the allosteric model 141

FIGURE 5-23 Structural map of $G_{s\alpha}$ mutant phenotypes 142

FIGURE 5-24 Interaction of Tryptophan-234 with ileucine-276 and leucine-272 143

FIGURE 5-25 Characterization of L46I (non-coupled control) 144

FIGURE 5-26 Mapping the statistically coupled amino acide network on active and inactive states of distant G protein family members 145

LIST OF TABLES

TABLE 2-1 Growth of yeast harboring plasmids encoding mutant adenylyl cyclases that confer independence of $G_{s\alpha}$-GTPγS and cyclic AMP……..……39

TABLE 2-2 Effect of single amino acid substitutions in adenylyl cyclase 40

TABLE 2-3 Affinity and activity of purified adenylyl cyclase mutants 41

TABLE 2-4 Substrate Kinetics with IIC_2 or IIC_2-K1014N 42

TABLE 3-1 Summary of Purificiation of H6-VC_1-MESNA from *E. coli*. 61

TABLE 3-2 Summary of Purification of H6-Xa-IIC_2-S871C from *E. coli*. 62

TABLE 5-1 Activation of adenylyl cyclase by GTPγS- or GDP-bound $G_{s\alpha}$ 117

TABLE 5-2 Binding of βγ by GDP- and GTPγS-bound $G_{s\alpha}$ 118

TABLE 5-3 Summary of $G_{s\alpha}$ Mutant Analysis 119

LIST OF ABBREVIATIONS

Å	angstrom
AMP	adenosine 5'-monophosphate
ATP	adenosine 5'-triphosphate
ATPαS-Rp	adenosine 5'-(α-thio)-triphosphate
bp	base pair
βLddATP	β-L-2',3'-dideoxyadenosine 5'-triphosphate
cAMP	3',5'-cyclic adenosine monophosphate
CBD	chitin binding domain of *Bacillus circulans*
DTT	dithiothreitol
$\Delta\Delta G^{stat}$	statistical coupling energy
EC_{50}	half maximal effective concentration
ECFP	enhanced cyan fluorescent protein
EDTA	ethylenediaminetetraacetic acid
EYFP	enhanced yellow fluorescent protein
FRET	fluorescence resonance energy transfer
Fsk	forskolin
g	gram
g	force of gravity
G	ground state of G protein
G*	active state of G protein
G protein	guanine nucleotide binding protein
GDP	guanosine 5'-diphosphate
GTP	guanosine 5'-triphosphate

GTPase	enzymatic activity that catalyzes hydrolysis of the γ-phospate of GTP
GTPγS	guanosine 5'-O-3-(thiotriphosphate)
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IC_{50}	half maximal inhibitory concentration
IPL	intein-mediated protein ligation
IPTG	isopropyl-β-D-thiogalactopyranoside
kDa	kilodalton
K_D	equilibrium dissociation constant
K_m	substrate concentration giving half-maximal velocity of reaction
L	liter
L_{GDP}	equilibrium constant for GDP-bound G protein from G to G* state
L_{GTP}	equilibrium constant for GTP-bound G protein from G to G* state
M	molar
MESNA	2-mercaptoethanesulfonic acid
mg	milligram
min	minute
mL	milliliter
mM	millimolar
MSA	multiple sequence alignment
μg	microgram
μL	microliter
μM	micromolar
μmol	micromole
Ni^{2+}-NTA	nitrilotriacetic acid-nickel ion-agarose

nM	nanomolar
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PP_i	pyrophosphate
SCA	statistical coupling analysis
SDS	sodium dodecylsulfate
TCEP	tris(2-carboxyethyl)phosphine
V_{max}	maximal velocity of an enzymatic reaction
WT	wild-type

Amino acids are abbreviated using both single and triple letter codes

Chapter 1 – Introduction

Cells need to sense their environment and respond to external stimuli. Most of these extracellular signals elicit intracellular responses through interactions at the cell surface that allow the cell to integrate these signals and respond appropriately. The human body is composed of trillions of cells composing a variety of tissues. A specific cell must be able to communicate with distant cells in order for the body to perform orchestrated tasks. Hormones released into the bloodstream carry messages that elicit effects on distant target tissues. Several decades of scientific exploration have elucidated the multiple participants in relaying these messages. Early studies of signaling systems centered on glucagon.

Glucagon is released from the endocrine pancreas during fasting and stimulates hepatic glycogenolysis. Glucagon signals the liver to activate glycogen phosphorylase, which catalyzes the cleavage of glucose-1-phosphate from its polymeric storage form, glycogen. Rall and Sutherland observed that glycogen phosphorylase was activated by glucagon in both liver slices and broken cell homogenates (Rall et al., 1957). Importantly, the membrane fraction of the homogenate was required for glucagon to activate glycogen phosphorylase located in the soluble fraction (Rall and Sutherland, 1958). Rall and Sutherland discovered that hepatocyte membranes produced a soluble, heat-stable factor in response to glucagon that activate glycogen phosphorylase in unstimulated extracts, and they eventuallyidentified this heat-stable factor as 3',5'-cyclic adenosine monophosphate (cAMP) (Sutherland and Rall, 1957). Of great interest, stimulation of cyclic AMP synthesis by the enzyme adenylyl cyclase was found to occur in response to several hormones, including epinephrine (Sutherland et al., 1962).

A decade later Rodbell and Birnbaumer made the surprising observation that guanine nucleotides were required for cAMP production in hepatocytes in response to glucagon (Rodbell et al., 1971a). Furthermore, guanine nucleotides reduced the affinity of the hormone for its receptor

(Rodbell et al., 1971b). Maguire and Gilman later showed that guanine nucleotides reduced the affinity of β-adrenergic agonists but not antagonists to their receptor (Maguire et al., 1976). These observations illustrated the functional correlation between guanine nucleotide dependence and hormone-stimulated cAMP production. Using turkey erythrocyte membranes, Cassel and Selinger observed an isoproterenol-dependent activation of GTP hydrolysis that correlated with the kinetic properties of isoproterenol-stimulated cAMP production (Cassel and Selinger, 1976). Cholera toxin, previously known to irreversibly activate adenylyl cyclase, was found to inhibit this GTPase activity (Cassel and Selinger, 1977). From these data, they concluded that the hydrolysis of GTP was the responsible for termination of the signal. Cassel and Selinger also showed that isoproterenol stimulated release of GDP from the membrane and deduced that this release was the rate limiting step in hormonal activation of adenylyl cyclase (Cassel and Selinger, 1978).

The components of the isoproterenol signaling system were isolated and purified using a combination of genetic and biochemical approaches. Three distinct components of this system were isolated: the β-adrenergic receptor (Shorr et al., 1981), adenylyl cyclase itself (Pfeuffer and Metzger, 1982), and the adenylyl cyclase-stimulating guanine nucleotide binding protein G_s (Sternweis et al., 1981). This system represented a conserved paradigm for intercellular communication which is initiated by interaction of an extracellular ligand or hormone with a heptahelical receptor. The hormone-bound receptor activates the G protein, which in turn activates adenylyl cyclase resulting in the conversion of ATP to cAMP.

G protein regulatory cycle

G proteins are heterotrimers, with subunits designated as α, β, and γ. In the basal state, the G protein α subunit is bound to GDP and associated with the G protein β and γ subunits (Figure 1-1). Binding of α to βγ occludes the sites used by α (and βγ) for interactions with at least some effectors

(e.g. adenylyl cyclase). Binding of agonist ligands to a heptahelical G-protein coupled receptor drives the interaction of the receptor with the G protein heterotrimer and elicits a conformational change in the G protein α subunit that stimulates the release of GDP. The nucleotide-free α subunit is then able to bind GTP, which is present at a much higher intracellular concentration than is GDP. GTP binding alters the conformation of the Gα subunit, decreasing its affinity for βγ and increasing affinity for effectors. The free G_α and βγ subunits are liberated to interact with effector molecules. The α subunit contains an intrinsic GTPase activity that hydrolyzes the bound GTP to GDP and phosphate, thereby returning the α subunit to the basal conformation that binds the βγ subunits.

G protein α subunit

Sixteen genes encode G protein alpha subunits in mouse and man. The protein products of these genes are grouped into four families based on sequence homology and function: G_s, G_i, G_q and G_{12}. The G_s family consists of $G_{s\alpha}$ (alternative splicing results in long and short variants) and $G_{olf\alpha}$. The G_i class consists of $G_{i\alpha 1,2 \text{ and } 3}$ (products of separate genes), $G_{o\alpha}$ (A and B forms from alternative splicing), $G_{t\alpha 1 \text{ and } 2}$ (from two genes), $G_{z\alpha}$, and $G_{g\alpha}$ (gustducin). Four genes encode the proteins of the G_q family: $G_{q\alpha}$, $G_{11\alpha}$, $G_{14\alpha}$, and $G_{16\alpha}$. The G_{12} class has only two members, $G_{12\alpha}$ and $G_{13\alpha}$. The protein products of the G_α genes are similar in size, ranging in molecular mass form 39, 000 to 45,000 Daltons. With the exception of the proteins found predominately in sensory organs ($G_{t\alpha}$, $G_{g\alpha}$, and $G_{olf\alpha}$), most of the α subunits have a wide tissue distribution.

G protein α subunits interact with a wide variety of effector molecules. $G_{s\alpha}$ and $G_{olf\alpha}$ activate adenylyl cyclases to generate the second messenger cAMP. Five members of the $G_{i\alpha}$ family ($G_{i\alpha 1\text{-}3}$, $G_{o\alpha}$, and $G_{z\alpha}$) inhibit certain adenylyl cyclases and oppose the action of $G_{s\alpha}$(Taussig et al., 1993a; Taussig et al., 1994). Transducin ($G_{t\alpha}$) activates cyclic GMP phosphodiesterase in retinal rods and

cones (Lerea et al., 1986). Members of the $G_{q\alpha}$ family stimulate phospholipase Cβ (PLC-β) (Strathmann et al., 1989; Smrcka et al., 1991; Taylor et al., 1991). $G_{12\alpha}$ and $G_{13\alpha}$ are involved in cytoskeletal rearrangements through activation of rho exchange factors, including p115RhoGEF (Hart et al., 1998). Recently, both $G_{s\alpha}$ and $G_{i\alpha}$ have been shown to activate the tyrosine kinase c-Src, thus linking the cAMP signaling pathway to the mitogen-activated protein kinase (MAPK) signaling system (Ma et al., 2000).

Several molecules covalently modify G_α subunits. All α subunits with the exception of $G_{t\alpha}$ are palmitoylated at one or more cysteine residues near the amino terminus (Linder et al., 1993). Members of the G_i class are myristoylated at an amino terminal glycine residue (Buss et al., 1987; Mumby et al., 1990b; Jones et al., 1990), and this increases their affinity for both βγ and adenylyl cyclase (Linder et al., 1991; Taussig et al., 1993a). $G_{s\alpha}$ purified from liver has a higher affinity for adenylyl cyclase than does $G_{s\alpha}$ expressed and purified from either Sf9 cells or *Escherichia coli*, suggesting that the native protein contains an unidentified covalent modification (Graziano et al., 1989a; Kleuss and Gilman, 1997). The enterotoxin from *Vibrio cholerae* ADP-ribosylates $G_{s\alpha}$, $G_{olf\alpha}$, and $G_{t\alpha}$ using NAD as the substrate (Abood et al., 1982; Bokoch et al., 1983). This modification inhibits the GTPase activity of G_α proteins and results in their constitutive activation. The toxin from *Bordetella pertussis* catalyzes the ADP-ribosylation of a C-terminal cysteine residue of all members of the $G_{i\alpha}$ family except $G_{z\alpha}$ (Bokoch et al., 1983). This modification prevents interaction of the protein with its cognate receptors (Casey et al., 1990).

High resolution crystal structures of the $G_{i\alpha1}$ and $G_{t\alpha}$ subunits have been solved in the GTPγS-, GDP:AlF_4^--, and GDP-bound states. These structures provide snapshots of the α subunit as it travels through its regulatory cycle. The structure of G_α bound to the non-hydrolyzable GTP analog GTPγS represents the activated G_α protein and provides a model of the enzyme (GTPase)-substrate

complex. The transition state of GTP hydrolysis is illustrated in the G_α-GDP:AlF_4^- structure. The GDP-bound G_α represents in the basal unstimulated state and the enzyme-product complex. The structures of all of the different G protein family members are likely very similar; and, only the $G_{i\alpha}$ structrures will be discussed here (Coleman et al., 1994; Lambright et al., 1994; Sunahara et al., 1997b).

The G_α subunit is composed of two domains, the ras-like and helical domains (Figure 1-2). The structure of the ras-like domain is very similar to that of p21ras (Pai et al., 1990). Similar to ras, the ras-like domain of G_α subunits contains three switch regions whose conformation is dependent on the identity of the bound nucleotide (GTP or GDP). Effectors, βγ, and RGS proteins interact with these switch regions of the G_α subunit (Tesmer et al., 2002; Wall et al., 1995; Lambright et al., 1996; Tesmer et al., 1997a). The guanine nucleotide binds within the cleft that separates the ras-like and helical domains.

G_α proteins contain five highly conserved loops (G-1 through G-5) that are responsible for binding nucleotide. The roles of these sequences were gleaned from the crystal structures. The G-1 region or P-loop contains the conserved sequence GXXXXGK(S/T) and contacts the α and β phosphates of the guanine nucleotide. The G-3, loop which contains the conserved DXXG sequence, is located at the N-terminus of switch II and binds Mg^{2+} and the γ-phosphate of GTP. The G-1 and G-3 loops contain the Walker A and B motifs that are global nucleotide binding sequences. The G-2 loop (switch I) contains a threonine necessary for Mg^{2+} coordination. The guanine ring is recognized by interactions with the G-4 box (conserved sequence NKXD) and the G-5 box, (T/G)(C/S)A.

The transition state structure elucidated the functions of mutations in $G_{s\alpha}$ that cause pituitary and thyroid adenomas (Landis et al., 1989) and McCune-Albright syndrome (Weinstein et al., 1991). Residues that are mutated in these conditions are the arginine in switch I (Arg-201 in $G_{s\alpha}$) and a

glutamine in switch II (Gln-227 in $G_{s\alpha}$) that are conserved in all subunits. These mutations inhibit GTP hydrolysis, resulting in constitutive signaling by $G_{s\alpha}$ (Graziano and Gilman, 1989b; Freissmuth and Gilman, 1989). The $G_{i\alpha1}$:GDP:AlF_4^- structure revealed that these residues stabilize the transition state of GTP hydrolysis (Coleman et al., 1994). Arg-178 of $G_{i\alpha}$ stabilizes the charge of the transition state through interactions with the γ-phosphate of GTP while Gln-204 orients the catalytic water molecule (Figure 1-3). The binding of RGS proteins, which stimulates the GTPase activity of G_α proteins, also stabilizes the transistion state. The structure of RGS4 complexed with $G_{i\alpha1}$:GDP:AlF_4^-:Mg^{2+} revealed that RGS4 does not contribute catalytic residues to the active site but contacts and stabilizes each of the switch regions (Tesmer et al., 1997a).

The structure of $G_{i\alpha}$:GDP differs from those of the GTPγS- and GDP:AlF_4^--bound protein in the switch regions. Switch I moves away from the nucleotide as a result of loss of Mg^{2+} coordination by Thr-181. This structure lacks density in switch II and switch III, suggesting that these regions are disordered. In contrast, switch II and III make extensive contacts in the $G_{i\alpha}$-GTPγS structure and are ordered. The loss of these contacts and the contact between the amide nitrogen of Gly-203 of $G_{i\alpha}$ and the γ-phosphate of GTP explains the flexibility of switch II and, accordingly, switch III. Although the loss of these interactions helps explain the conformational change between the GTP- and GDP-bound states, it is likely that other amino acid residues in G_α also contribute.

G protein βγ subunits

The G protein β and γ subunits are obligate heterodimers that dissociate only on denaturation. The βγ subunits stabilize GDP-bound α by reducing the rate of GDP dissociation. The βγ subunits occlude the binding sites for adenylyl cyclase on α and are required for receptor-catalyzed exchange of GDP for GTP on α. Genes encoding five different β and twelve different γ subunits have been

identified in mammals (Clapham and Neer, 1997). The sequences of β_{1-4} are very similar (80% identical) but that of β_5 is more distinct (50% identity to the others). The primary sequence of G_β is composed of seven repeating sequences of roughly 40 amino acids that include Trp and Asp residues near the end of each repeat; these sequences have thus been called the WD40 repeats. G_γ proteins are more diverse, sharing only sharing 30 to 80% sequence identity. The carboxyl terminus of G_γ contains a CAAX motif that is the site of prenylation.. The γ_1 subunit of the retina is farnesylated, while the remainder of the G_γ subunits are geranylgeranylated (Mumby et al., 1990a; Yamane et al., 1990). G_γ proteins with a CAAX motif cysteine mutated to serine are not prenylated and fail to associate with membranes (Muntz et al., 1992; Simonds et al., 1991)

The βγ subunits are also capable of initiating downstream signaling events by interactions with effectors. βγ interacts with K^+ channels in cardiac atrial myocytes (Logothetis et al., 1987). The βγ subunits were shown to be the primary signaling component in the yeast mating pheromone receptor pathway (Whiteway et al., 1989). βγ subunits can differentially regulate the $G_{s\alpha}$-stimulated activity of adenylyl cyclase (see below). βγ subunits activate PLC-β (Smrcka and Sternweis, 1993), phosphoinositol 3-kinase (Stephens et al., 1994), Bruton's tyrosine kinase (Lowry and Huang, 2002), and G protein- coupled receptor kinases (Pitcher et al., 1995).

Multiple G_β and G_γ subunits could form a total of 60 distinct combinations of complexes offering the possibility of signaling diversity. However, signaling specificity of these combinations is not understood. $G_{\beta1}$ purified from Sf9 cells interacts with G_γ isoforms 1, 2, 3, 5, and 7. $G_{\beta2}$ purified from Sf9 cells interacts with all G_γ isoforms except γ_1. All βγ combinations tested were capable of supporting pertussis toxin-catalyzed ADP-ribosylation of $G_{i\alpha}$, activating adenylyl cyclase type-II, inhibiting adenylyl cyclase type-I, and activating PLC-β (Iñiguez-Lluhi et al., 1992; Ueda et al., 1994).

The structure of the βγ dimer was elucidated both as a component of the heterotrimeric complex of G_i and G_t and by itself (Figure 1-4) (Wall et al., 1995; Lambright et al., 1996; Sondek et al., 1996). The majority of the β subunit forms a β-propeller fold. The seven WD40-repeats form seven blades of the propeller each blade consisting of four antiparallel β strands. The N-terminal α helix of the γ subunit forms a coiled-coil with the N-terminus of β, while the remainder of the γ subunit extends along the bottom surface of the β subunit. The core of the contact of βγ with $G_{i\alpha1}$ is centered on the switch II helix of α (Ile-184, Phe-199, Cys-214, Phe-215, and Lys-210), Figure 1-4B. The structure of the $G_{i\alpha1}$ subunit in the heterotrimer revealed large conformational changes between the GTPγS-bound and the free GDP-bound states (Figure 1-5). The switch II helix rotates 120^o to expose hydrophobic residues interacting with the G_β subunit and removing contacts with the γ-phosphate of GTP. In contrast to the nucleotide-dependent reorganization of the G_α subunit, the βγ dimer is rigid in both the heterotrimic complex and alone as the βγ dimer (Sondek et al., 1996; Lambright et al., 1996). This indicates that the regulation of signaling by βγ is dependent on association with the α subunit and not on intermolecular conformational changes.

Mammalian adenylyl cyclases

Mammalian adenylyl cyclases are membrane-bound enzymes that catalyze the synthesis of the intracellular second messenger cyclic AMP from ATP. Most adenylyl cyclases are activated by forskolin, a vasodilatory diterpine from the root of the plant *Coleus forskolii* (Seamon et al., 1981). Pfeuffer and Metzger developed a forskolin-affinity matrix that allowed the purification of the enzyme (Pfeuffer and Metzger, 1982; Pfeuffer et al., 1985; Smigel, 1986). Krupinski and coworkers purified adenylyl cyclase from bovine brain using this affinity matrix and obtained a partial amino acid sequence. This permitted the cloning of a full-length cDNA encoding type-I adenylyl cyclase

(Krupinski et al., 1989). Subsequently, a total of nine isoforms of the enzyme have now been detected, and they display diverse regulatory properties and patterns of cellular distribution (Taussig and Gilman, 1995; Sunahara et al., 1996).

Cellular rates of cyclic AMP synthesis are controlled by a variety of extracellular ligands that interact with heptahelical receptors in the plasma membrane. Relevant receptors can either stimulate cyclic AMP synthesis, usually via the intermediacy of a G protein (G_s) that activates adenylyl cyclase, or inhibit cyclic AMP synthesis, often by interaction with inhibitory G proteins, G_i. All isoforms of adenylyl cyclase are activated by $G_{s\alpha}$ while only types I, V, and VI are inhibited by the $G_{i\alpha}$ family. Inhibition of adenylyl cyclase requires covalent modification of $G_{i\alpha}$ by myristate (Taussig et al., 1993a; Taussig et al., 1994). Mammalian adenylyl cyclases, except type IX, are activated by the diterpene forskolin and all isoforms are inhibited by certain adenosine analogs and adenine nucleotides called P-site inhibitors (Londos and Wolff, 1977; Dessauer et al., 1999). G protein $\beta\gamma$ subunits inhibit type-I adenylyl cyclase and stimulate type II and IV adenylyl cyclases (Katada et al., 1987; Tang and Gilman, 1991; Taussig et al., 1993b; Federman et al., 1992). The inhibitory effects of $\beta\gamma$ on type I adenylyl cyclase are best observed with the Ca^{2+}•calmodulin-activated enzyme, while the stimulatory effects of $\beta\gamma$ on types II and IV adenylyl cyclase require activation with $G_{s\alpha}$. Certain adenylyl cyclases are also regulated by Ca^{2+}, Ca^{2+}•calmodulin, and phosphorylation (Smit and Iyengar, 1998). The diversity in regulation of the various isoforms of adenylyl cyclase presumably permits integration of signaling pathways appropriate for the cell in which each isoform is localized.

Mammalian adenylyl cyclases are integral membrane proteins (~120 kDa) that appear to contain two sets of six membrane-spanning helices (M_1 and M_2) that are separated by a large (~40 kDa) cytoplasmic loop (C_1) followed by a similarly-sized carboxy-terminal cytosolic domain (C_2) (Krupinski et al., 1989). The predicted topology is reminiscent of pore-forming molecules such as the P-glycoproteins and cystic fibrosis transmembrane conductance regulator, although no transport or

channel activity has been detected (Figure1-6). The cytosolic domains, termed C_1 and C_2, have been extensively studied; they are responsible for catalytic activity and most of the regulatory properties of the enzymes (Tang and Gilman, 1995). The first 200-250 amino acids of each cytosolic domain, designated C_{1a} and C_{2a}, are the most highly conserved regions among adenylyl cyclases. Strikingly, the C_{1a} and C_{2a} domains are approximately 50% similar and 25% identical to each other within a single isoform of adenylyl cyclase, and they are 20-25% similar to the catalytic domains of membrane-bound and cytosolic guanylyl cyclases.

The biochemical and structural characterization of adenylyl cyclase was for some time limited by the low abundance of the native protein (0.01-0.001% of membrane protein), difficulty of expression in heterologous systems, and lability in detergents required during purification. In attempts to circumvent these difficulties, adenylyl cyclase was severed into the N-terminal half, M_1C_1 domains, and the C-terminal half, M_2C_2 domains, and expressed as distinct proteins in Sf9 cells. Neither of these proteins displayed enzymatic activity alone; however, co-expression of the M_1C_1 and the M_2C_2 proteins permitted observation of $G_{s\alpha}$- and forskolin-stimulated adenylyl cyclase activity(Tang et al., 1991). The deletion of the C_{1b} and C_{2b} regions from the M_1C_1 and M_2C_2 proteins had little effect on enzyme activity (Tang et al., 1995). Interestingly, point mutations in either the C_{1a} or C_{2a} domain could increase the K_M for substrate and dramatically reduce the enzymatic activity without affecting affinity for $G_{s\alpha}$ (Tang et al., 1995). These results suggested that the interaction of the C_{1a} and C_{2a} domains is required for catalysis.

It was not possible to purify the M_1C_1 and M_2C_2 domains separately and mix them to reconstitute adenylyl cyclase activity. Therefore, the DNA encoding the C_1 domain from type I adenylyl cyclase was linked to the DNA encoding the C_2 domain from type II adenylyl cyclase. The resulting C_1-C_2 fusion protein was expressed in *E. coli* and was catalytically active. Surprisingly, the adenylyl cyclase activity of the C_1-C_2 fusion protein was activated by both forskolin and $G_{s\alpha}$ and

inhibited by P-site inhibitors (Tang and Gilman, 1995; Dessauer and Gilman, 1996). Low-level accumulation of the protein in the cytosol limited its usefulness. However, the C_{1a} and C_{2a} domains of adenylyl cyclase can easily be expressed separately as recombinant proteins and purified. When mixed together, these domains display most of the characteristics of membrane-bound adenylyl cyclase with respect to regulation by $G_{s\alpha}$, $G_{i\alpha}$, forskolin, and P-site inhibitors (Whisnant et al., 1996; Yan et al., 1996; Sunahara et al., 1997a).

The C_1 domain from type-V adenylyl cyclase (VC_1) and the C_2 domain from type-II adenylyl cyclase (IIC_2) were purified in sufficient yields from *E. coli* to allow detailed biochemical characterization. In the absence of activators, the apparent affinity of VC_1 for IIC_2 was weak (>10 μM), and the catalytic activity was low (1 $\mu mol \bullet min^{-1} \bullet mg^{-1}$). The apparent affinity of VC_1 and IIC_2 was increased to approximately 1 μM in the presence of either $G_{s\alpha}$-GTPγS or forskolin, and the catalytic activity was increased more than 60-fold. In the presence of both $G_{s\alpha}$-GTPγS and forskolin, the apparent affinity of the VC_1 and IIC_2 domains increased to 150 nM and catalytic activity was 160 $\mu mol \bullet min^{-1} \bullet mg^{-1}$ (Sunahara et al., 1997a). Forskolin and $G_{s\alpha}$ appear to increase the interaction between the C_1 and C_2 domains, resulting in increased enzymatic activity.

The similarity in the primary sequence of the C_1 and C_2 domains and the requirement of both domains for enzymatic activity suggests the possibility of more than one binding site for substrate and activators, $G_{s\alpha}$ and forskolin. In fact, a crystal structure of the homodimer formed by the IIC_2 protein revealed two forskolin binding sites without a substrate binding site (Zhang et al., 1997). Sedimentation equilibrium, gel filtration, and equilibrium dialysis studies revealed the presence of one molecule each of $G_{s\alpha}$-GTPγS and forskolin bound to the VC_1:IIC_2 heterodimer(Sunahara et al., 1997a; Dessauer et al., 1997). Equilibrium dialysis of the VC_1 and IIC_2 domains also revealed one binding site for a nonhydrolyzable ATP analog, $Ap(CH_2)pp$, that was not observed with the individual domains (Dessauer et al., 1997). The yields from purification of the C_1 and C_2 domains

and the high affinity of the activated complex permitted the crystallization and structural determination of the VC_1:IIC_2 catalytic core of adenylyl cyclase associated with the activators $G_{s\alpha}$-GTPγS and forskolin (Tesmer et al., 2002).

Structure of $G_{s\alpha}$ and the cytosolic domains of adenylyl cyclase

The crystal structures of the soluble catalytic core of adenylyl cyclase bound to $G_{s\alpha}$ and forskolin (Tesmer et al., 1997b) and of this complex bound with the competitive substrate analogs β-L-2',3'-dideoxyadenosine 5'-triphosphate (βLddATP) and adenosine 5'-(α-thio)-triphosphate (ATPαS-Rp) (Tesmer et al., 1999) have provided detailed insights into mechanisms of catalysis of cyclic AMP synthesis and regulation of the enzyme's activity. The C_1 and C_2 domains are arranged as a pseudo-two-fold symmetrical dimer (Figure 1-7). This symmetry is not surprising because of the similarity of the primary sequence of the C_{1a} and C_2 domains. A central cleft that contains the forskolin and substrate binding sites separates C_1 and C_2. Residues both from the C_1 and C_2 contribute to the active site and the forskolin binding site at the interface. The pseudosymmetry of the VC_1:IIC_2 heterodimer is illustrated by the contacts at the interface. The β2 strand of IIC_2 contacts the β4-β5 loop of VC_1, while the β2 strand of VC_1 contacts the β4-β5 loop of IIC_2. Forskolin interacts with residues in both VC_1 and IIC_2. The increased activity of the enzyme with forskolin results partly from increasing the association of C_1 and C_2 and partly from altering the structure of the active site through contacts with the β2-β3 loop of VC_1.

The VC_1 and IIC_2 domains, and presumably guanylyl cyclases, contain a βαββαβ fold resembling the "palm" domains of DNA polymerase, although there is no significant sequence similarity (Artymiuk et al., 1997). Adenylyl cyclases, guanylyl cyclases, and DNA polymerases catalyze the attack of the 3' hydroxl of ribose on the α-phosphate of a nucleoside triphosphate. DNA

polymerases contain two invariant aspartate residues that coordinate two metal ions (Woody et al., 1996; Kaushik et al., 1996). Only the C_1 domains of adenylyl cyclases contain the analogous aspartate residues. The loops of this domain collapse upon nucleotide binding in both adenylyl cyclase and DNA polymerases resulting in the formation of a competent active site.

The structure of the protein associated with βLddATP revealed the presence of two metal ions (designated A and B) (Tesmer et al., 1999). The structures of C_1:C_2:$G_{s\alpha}$:forskolin soaked with the competitive substrates βLddATP and ATPαS-Rp allowed Tesmer and coworkers to model the binding of the substrate ATP in the active site and propose the mechanism of catalysis (Figure 1-8). Metal A is coordinated by each of the invariant aspartate residues in the C_1 domain (Asp-396 and Asp-440), the unesterified oxygen of the α phosphate of the nucleotide, and the 3' hydroxyl of ribose. Metal A is presumed to act as a Lewis acid and increase the nucleophilicity of the 3' hydroxyl of ATP. Metal B is coordinated by both invariant aspartates and the unesterified oxygens of the α, β, and γ phosphates of the nucleotide.

Adenylyl cyclase interactions with $G_{s\alpha}$ are primarily with the IIC_2 domain. The switch II helix of $G_{s\alpha}$ inserts into a groove between the α2 helix and the α3-β4 loop of IIC_2. The α3-β5 loop of $G_{s\alpha}$ interacts with both the IIC_2 and the VC_1 domains. The $G_{s\alpha}$ and adenylyl cyclase contacts elucidated from the structure correlated well with residues identified by scanning mutatgenesis. The $G_{s\alpha}$ binding region of the IIC_2 domain was previously identified by mutagenesis of the recombinant protein (Yan et al., 1997a). In addition, the switch II helix(Berlot and Bourne, 1992) and the α3-β5 loop (Itoh and Gilman, 1991) of $G_{s\alpha}$ were identified as adenylyl cyclase-interacting regions by scanning mutagenesis. The α4-β6 loop of $G_{s\alpha}$ was also implicated by mutagenesis to interact with adenylyl cyclase (Berlot and Bourne, 1992); however, the lack of interaction in the present structures does not eliminate the validity of this interaction since all regions of adenylyl cyclase are not present

in the structure. To gain insight into the mechanism of activation of adenylyl cyclase by $G_{s\alpha}$, Tesmer and coworkers used the IIC_2 homodimer as an approximate model of the VC_1:IIC_2 inactive state. The insertion of the switch II helix from $G_{s\alpha}$ into IIC_2 displaces the α1-α2 loop 3Å from the α3, helix translating to a 7^o rotation around the central cleft. This rotation results in the β2-β3 loop of the VC_1 domain containing Asp-396 and Asp-440 to move towards the β4 strand of the IIC_2 domain, forming a competent active site.

Inhibition of type-V and type-VI adenylyl cyclase by $G_{i\alpha}$ is noncompetitive, suggesting that $G_{i\alpha}$ binds to a site different from that occupied by $G_{s\alpha}$. The pseudosymetric nature of the adenylyl cyclase catalytic core suggests that $G_{i\alpha}$ could bind to the site in C_1 analogous to the $G_{s\alpha}$ binding site in C_2. Prospective sites of interaction were identified by scanning mutagenesis of the VC_1 domain (Dessauer et al., 1998). Mutations in the α1-α2 loop and the α3 helix decreased the ability of $G_{i\alpha}$ to inhibit adenylyl cyclase. Recently, $G_{i\alpha}$ was show to reduce the interactions of the C_1 and C_2 domains and formation of the catalytic site. Mutations that increase the affinity of the C_1 and C_2 domains decreased the ability of $G_{i\alpha}$ to inhibit adenylyl cyclase activity (Dessauer et al., 2002).

Statement of purpose

The use of the IIC_2 homodimer as a model of the low activity state of adenylyl cyclase is not ideal. The IIC_2 homodimer contains two forskolin molecules and lacks enzymatic activity. A high resolution structure of the VC_1:IIC_2 heterodimer in the absence of activators will give a better approximation of the conformational changes elicited through $G_{s\alpha}$ and forskolin binding. Two major obstacles prevent the purification of useful amounts of a low activity complex of the VC_1 and IIC_2 domains. First, the domains have approximately the same mass (~28 kDa), and form homodimers. This precludes the use of gel filtration to isolate a stoichiometric complex needed for crystallization.

Further complicating the isolation of the VC_1:IIC_2 complex, VC_1 and IIC_2 have low affinity in the absence of activators (>10 μM). Chapters 2, 3, and 4 will explore methods to increase the affinity of the VC_1 and IIC_2 domains to attain a high resolution crystal structure that better approximates the true low activity state of adenylyl cyclase.

Not unlike the elusive changes at the active site of adenylyl cyclase structural rearrangements in the alpha subunit remain to be determined. The G_α subunit passes through many conformations between the GTPγS-bound and the GDP-bound state. The amino acids residues responsible for the transition between these states are elusive. Using a statistical method, several of the amino acids residues responsible for this transition have been characterized in Chapter 5.

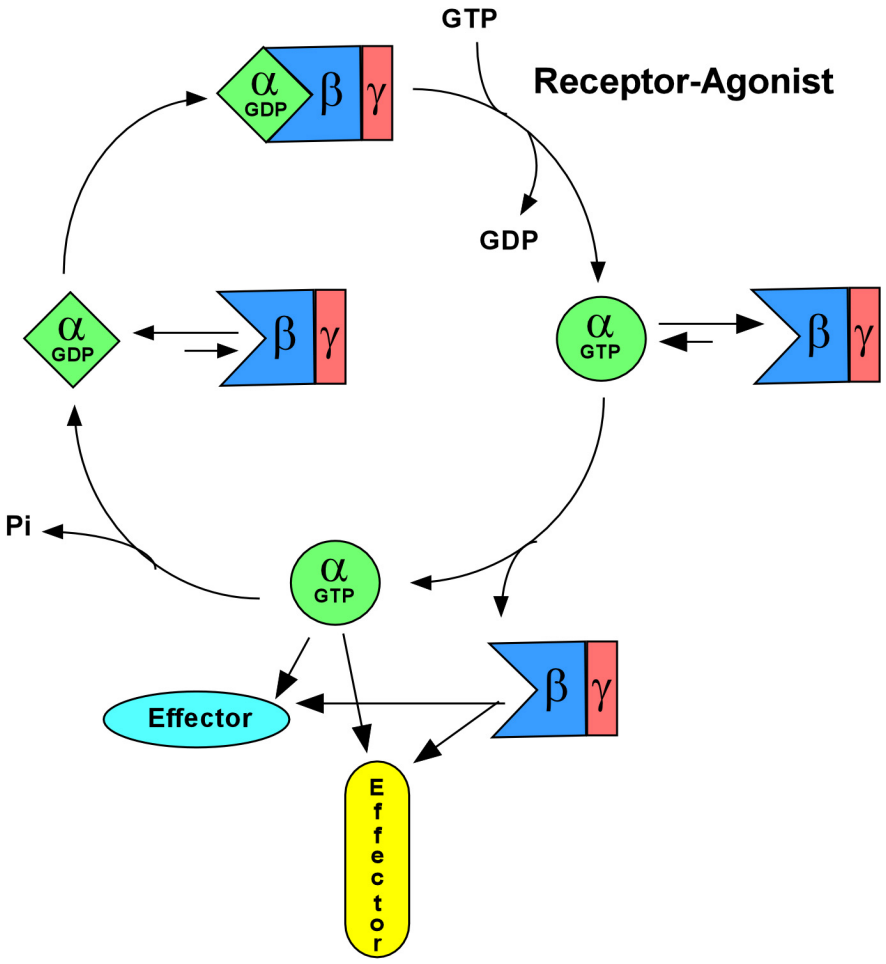


Figure 1-1. Guanine nucleotide cycle of heterotrimeric G proteins. In the basal state, the G protein is a heterotrimeric complex of α, β, and γ subunits with the α bound to GDP. Binding of agonist to the heptahelical G protein coupled receptor elicits a conformational change in α stimulating the release of GDP. The nucleotide free α then binds GTP, which is present at much higher intracellular concentration than GDP. GTP binding alters the conformation of α decreasing the affinity for βγ. Both α and βγ are liberated to interact with effector molecules. The α possesses an intrinsic GTPase activity the hydrolyzes the bound GTP to GDP and phosphate. The GDP-bound α returns to the basal conformation and binds βγ, thus completing the cycle.

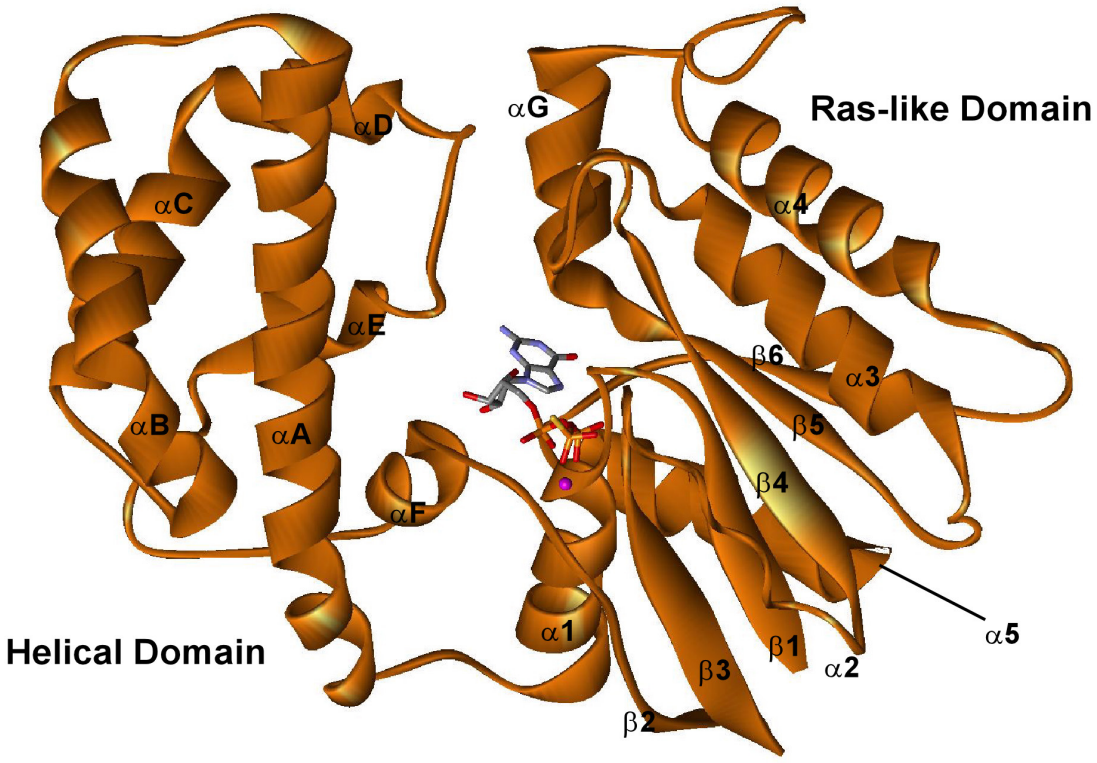


Figure 1-2. Structure of $G_{i\alpha1}$:GTPγS:Mg^{2+} complex(Coleman et al., 1994). The $G_{i\alpha1}$:GTPγS:Mg^{2+} structure is represented as a ribbon diagram (PDB entry 1GIA). The alpha subunit is composed of a helical domain and a ras-like domain. The GTPγS binds in the cleft separating these domains. GTPγS is represented as a stick model and the magnesium atom shown as a purple sphere. The secondary structural elements are labeled.

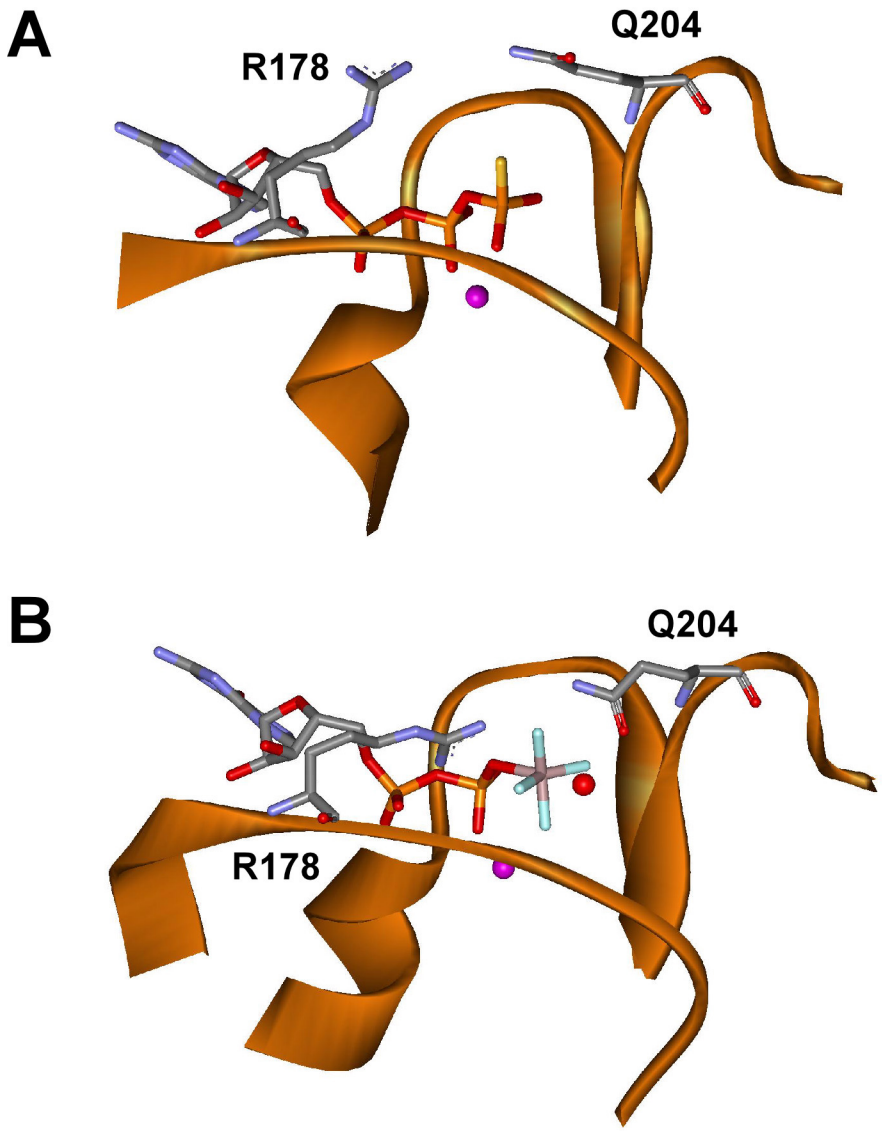


Figure 1-3. Conformational changes associated transition state of GTP hydrolysis. Active site in the $G_{i\alpha1}$-GTPγS structure (**A**, PDB entry 1GIA) and the $G_{i\alpha1}$-GDP:AlF_4^- structure (**B**, PDB entry 1GFI). The structure of $G_{i\alpha1}$ is illustrated as an orange ribbon diagram, the side chains of Arg-178 and Gln-204 and nucleotide are represented as stick models. The magnesium ion and catalytic water molecule are depicted as a magenta and red spheres respectively. The structure of the $G_{i\alpha1}$-GDP:AlF_4^- mimics the transition state for GTP hydrolysis. The Arg-178 and Gln-204 residues rotate toward the AlF4- and the nucleophilic water. The Arg-178 interacts with the γ-phosphate (AlF_4^-) to stabilize the charge of the transition state. The Gln-204 orients the catalytic water.

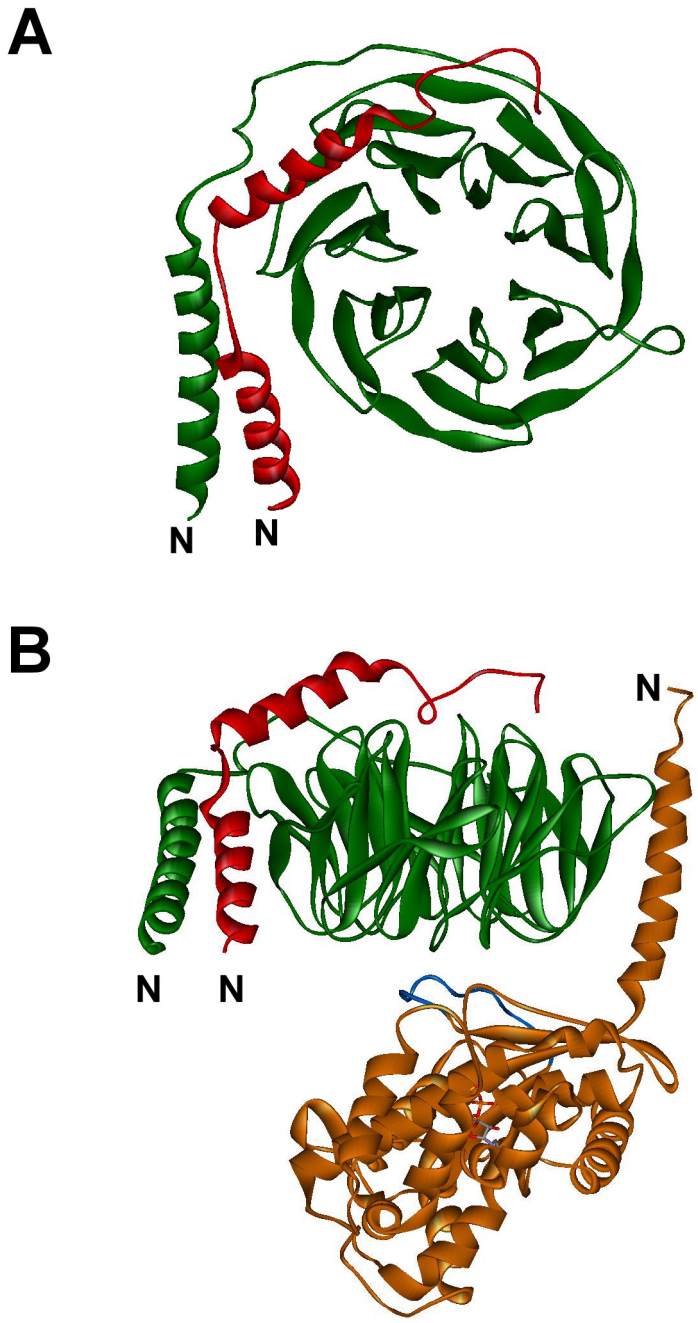


Figure 1-4. Structure of G_i heterotrimer and βγ subunits. A. The $\beta_1\gamma_2$ dimer is illustrated as ribbon diagrams (β_1 as green and γ_2 as red). β forms a propeller fold of seven blades comprised of four antiparallel strands (PDB entry 1GP2 with α removed). **B.** The $G_{i\alpha1}\beta_1\gamma_2$ heterotrimer (PDB entry 1GP2) depicted as ribbon diagrams (α orange, β green, and γ red). The $\beta_1\gamma_2$ contact region of αi1, the switch II region, is colored blue. The amino termini of the proteins are marked with a N.

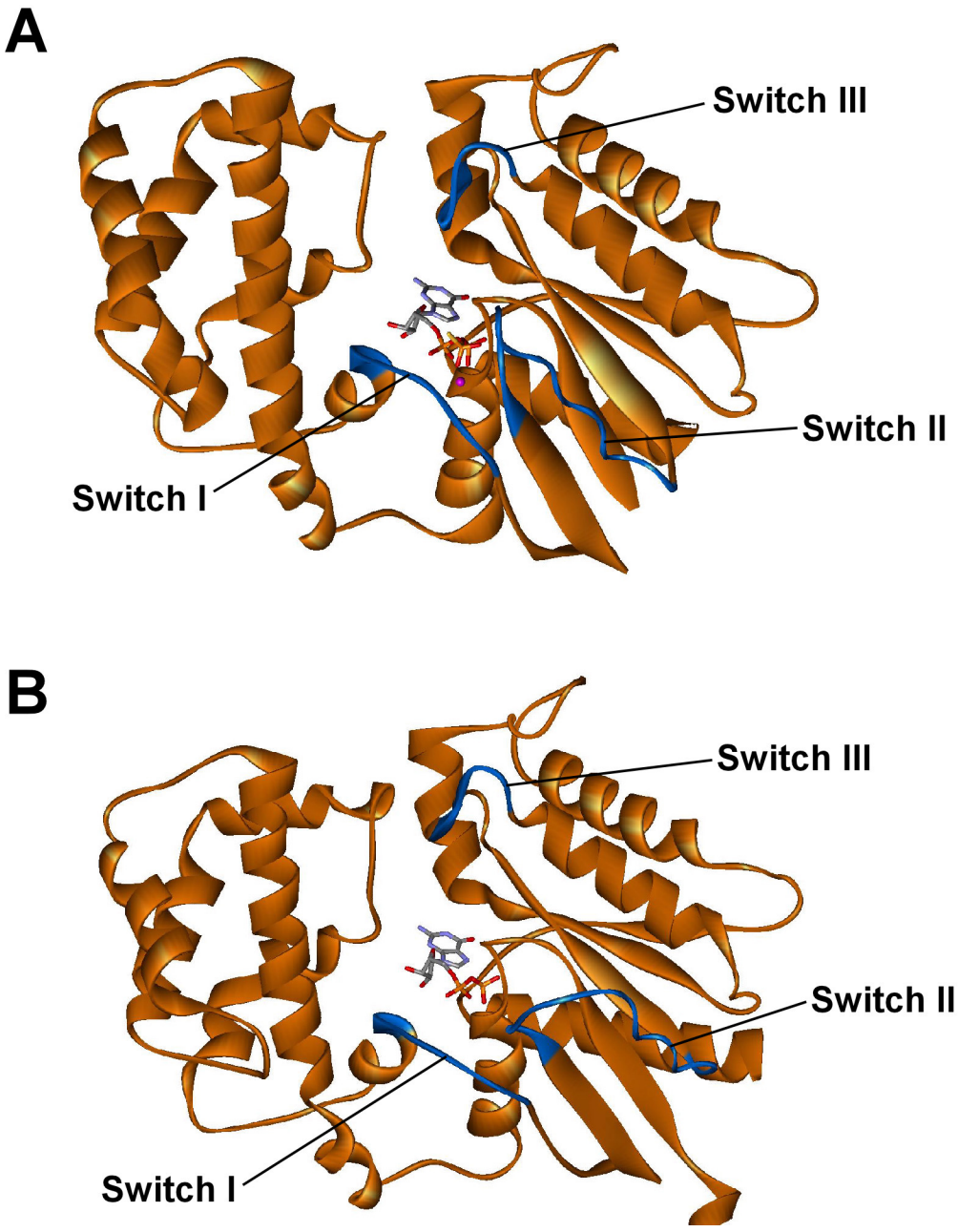


Figure 1-5. Conformations of GTPγS- and GDP-bound $G_{i\alpha 1}$. **A.** $G_{i\alpha 1}$ bound to GTPγS:Mg^{2+} (PDB entry 1GIA) (Coleman et al., 1994). **B.** $G_{i\alpha 1}$ bound to GDP (PDB entry 1GP2)(Wall et al., 1995). The switch regions are colored blue. The GDP-$G_{i\alpha 1}$ is from the heterotrimer structure with the βγ removed. The switch regions are disordered in the GDP-$G_{i\alpha 1}$ (Coleman et al., 1994)suggesting structural flexibility. Binding βγ stabilizes the switch regions allowing observation of the conformational changes in the GDP- and GTP-bound states.

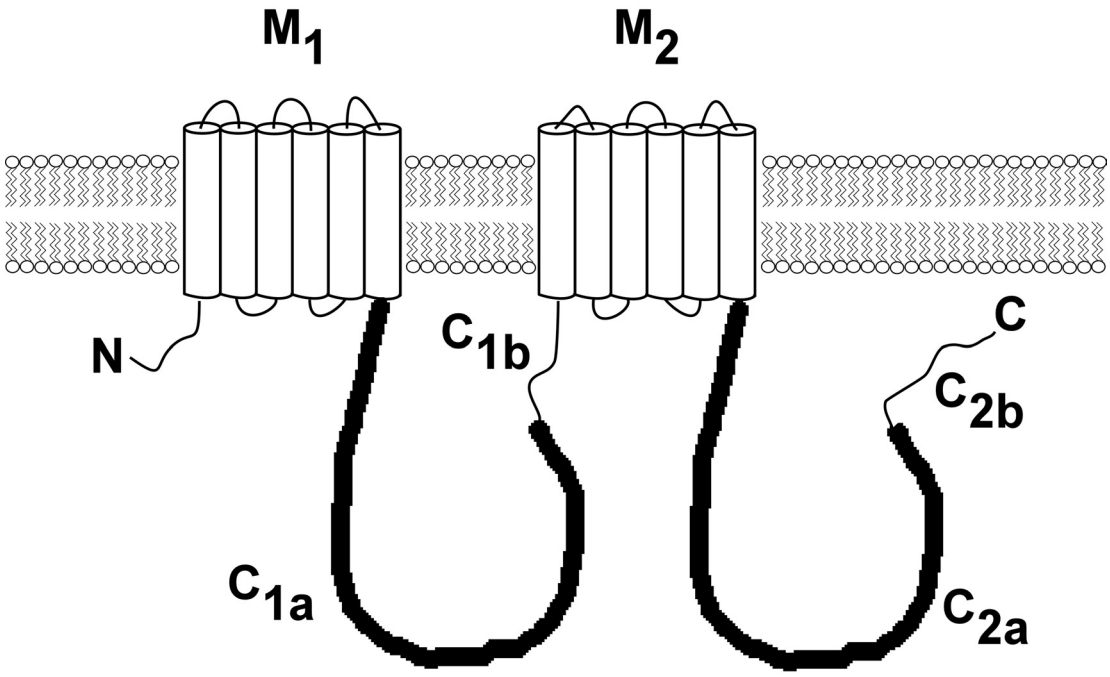


Figure 1-6. Predicted topology of mammalian adenylyl cyclase. The membrane-spanning domains (M_1 and M_2) are separated by cytosolic loop (C_1) and followed by another cytosolic loop (C_2). The C_1 and C_2 cytosolic loops are divided into a and b regions. The C_{1a} and C_{2a} regions are the most conserved regions among adenylyl cyclases, and strikingly 50% similar and 25% identical within a single isoform of adenylyl cyclase.

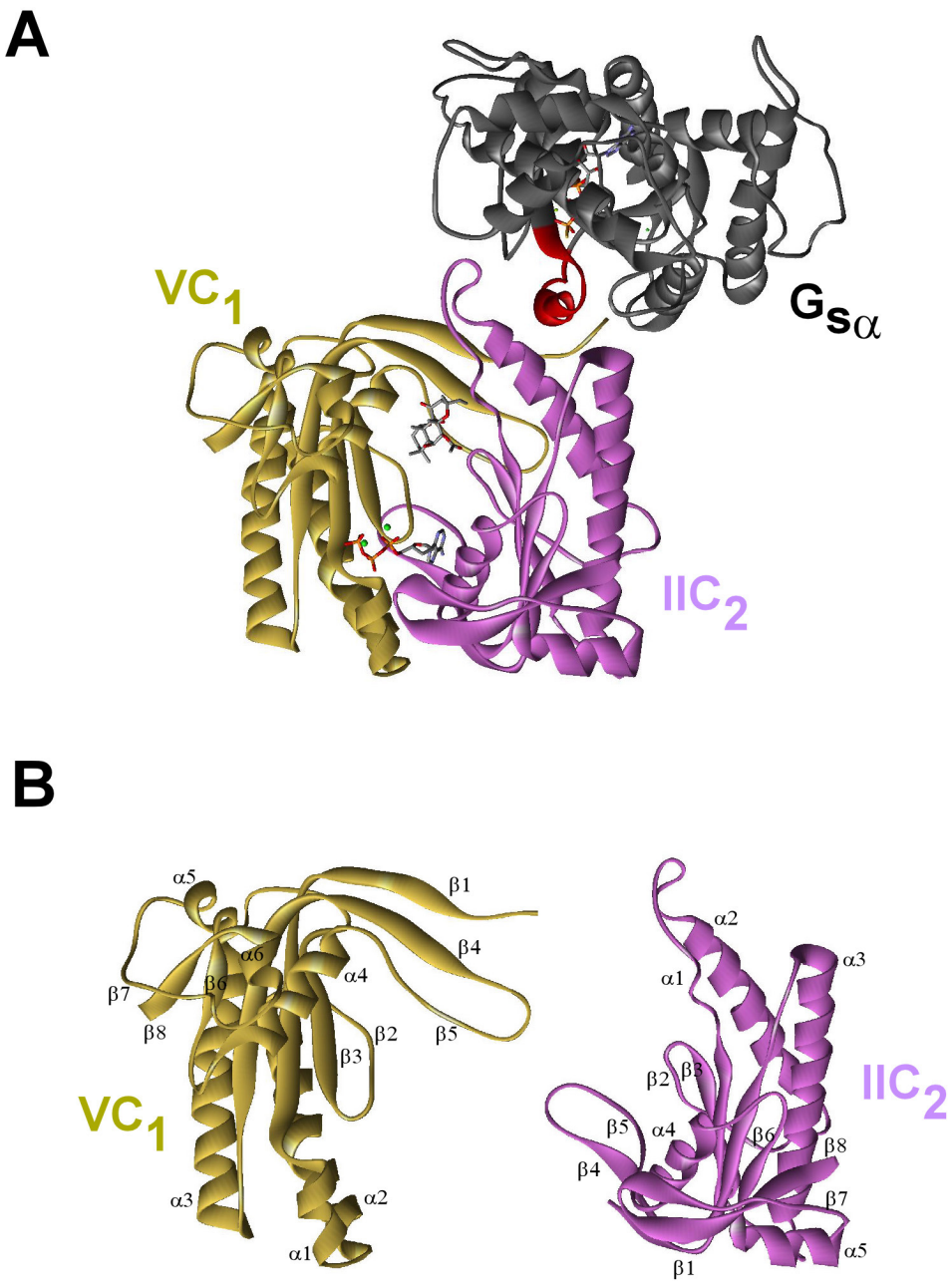


Figure 1-7. Structure of adenylyl cyclase catalytic core complexed to $G_{s\alpha}$-GTPγS. A. The VC1:IIC2:Gsα-GTPγS:forskolin:βLddATP structure (Tesmer et al., 1999) depicted as ribbon structures VC_1 (tan), IIC_2 (purple), and $G_{s\alpha}$-GTPγS (gray). The switch II region of $G_{s\alpha}$ is highlighted red. The substrate analog βLddATP and forskolin are depicted as stick figures. Magnesium ions are shown as green spheres. **B.** The secondary structural elements labeled on the VC1 and IIC2 adenylyl cyclase domains.

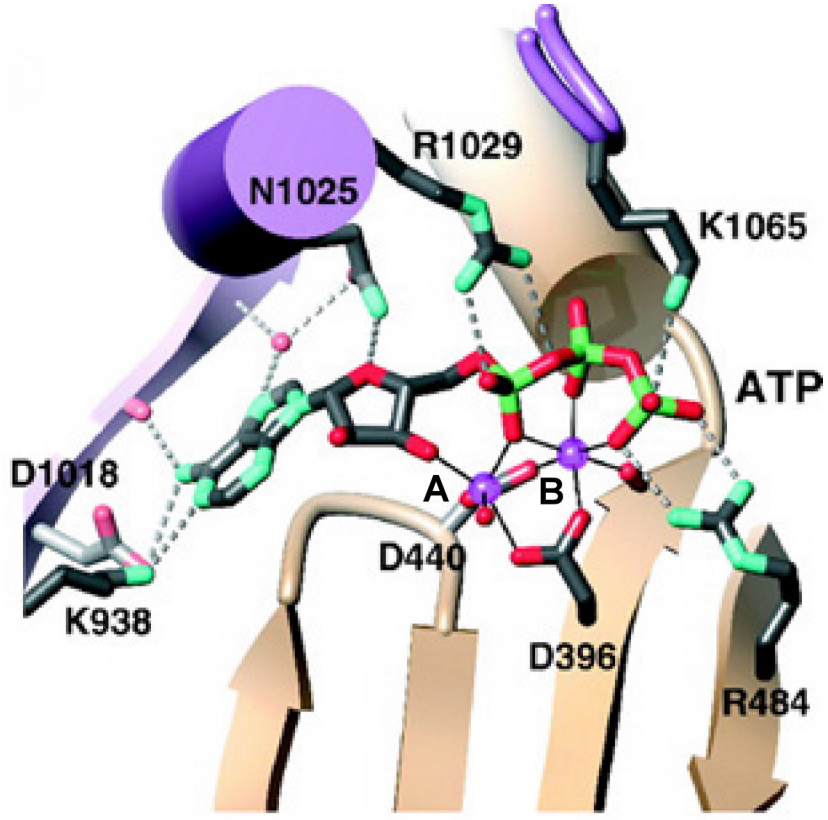


Figure 1-8. Model of substrate ATP in active site of adenylyl cyclase. The structures of VC_1:IIC_2:$G_{s\alpha}$-GTPγS complexed with ATP analogs βLddATP and ATPαS-Rp allowed Tesmer and coworkers to model ATP in the active site of adenylyl cyclase (Figure taken from (Tesmer et al., 1999)). Residues from both VC_1 (tan) and IIC_2 (purple) contribute to the ATP binding site. ATP is depicted as a stick model. The coordination of the metal ions are shown as thin black lines. Magenta spheres represent the Mg^{2+} ions and the red spheres represent water. The metal binding sites are designated A and B. Metal A is presumed to increase the nucleophilicity of the 3' hydroxyl of ATP.

Chapter 2 - Isolation, Characterization, and Crystallization of Constitutively Active Mutants of Mammalian Adenylyl Cyclase.

Introduction

The crystal structures of the soluble catalytic core of adenylyl cyclase activated with bound $G_{s\alpha}$ and forskolin (Tesmer et al., 1997b) and of this complex bound with competitive substrates β-L-2',3'-ddATP and ATPαS (Tesmer et al., 1999) have provided insights into the mechanism of catalysis of cyclic AMP synthesis and regulation of the enzyme. Less information is available about the biochemical and structural properties of low-activity states of adenylyl cyclase. Co-crystallization of the C_1 and C_2 domains of adenylyl cyclase in the absence of $G_{s\alpha}$ and/or forskolin would permit structural comparison of high- and low-activity states of the enzyme. However, the low affinity (>10 μM) of the C_1 and C_2 domains makes difficult the purification and crystallization of a low-activity complex. When expressed separately as recombinant peptides, the C_1 and C_2 domains are homodimers of approximately the same size, precluding gel filtration as a method to purify the heterodimeric complex. Methods to increase the affinity of the C_1 and C_2 domains must be developed to isolate a low activity complex for structural determination.

A combination of genetic and biochemical techniques was used to approach this goal. Full-length mammalian adenylyl cyclase was introduced into a mutant strain of *Saccharomyces cerevisiae* that does not express the endogenous enzyme. Co-expression of the mammalian protein together with $G_{s\alpha}$ relieves dependence on cyclic AMP for growth at non-permissive temperatures. John Manfredi at Cadus Pharmaceutical Corporation developed a genetic screen to isolate constitutively active mutants of adenylyl cyclase within this strain. The isolation of several partially active mutants from Manfredi's screen, the biochemical consequences of these mutations in the context of the soluble, recombinant adenylyl cyclase system, as well as methods used to crystallize a VC_1:IIC_2 heterodimer in the absence of $G_{s\alpha}$ are described in this chapter.

Experimental Procedures

Genetic Screen – John Manfredi, with Benjamin Benton and Jun Xu, developed and performed the following screen at Cadus Pharmaceuticals. To achieve high-level constitutive expression of type II adenylyl cyclase (ACII) in yeast, a 0.94 kb fragment containing the *S. cerevisiae PGK1* promoter and a 3.3 kb fragment encoding rat ACII (PCR-amplified from a template plasmid provided by Randall Reed, John Hopkins University) were subcloned into pRS425 (Christianson et al., 1992). The resulting expression plasmid, Cp1512 (genotype *LEU2 PGK1p-ACII 2μ-ori REP3 bla*), was used as the template for error-prone PCR amplification of two separate fragments: nucleotides 5514 through 6384, which encode the C_{1a} domain, and nucleotides 7033 through 8264, which encode the C_{2a} domain (primer sequence available upon request). The mutagenized fragments encoding the C_{1a} domain were then used to transform yeast strain TC41F2-1 (genotype *MATa cyr1::ura3 cam1 cam2 cam3 leu2-3 leu2-112 his3-532 his4 ura3*; provided by Warren Heideman, University of Wisconsin) along with a "gapped" vector prepared as follows. Cp1512 was digested with ApaI (which cuts at nucleotide 5682 of ACII) and DraIII (which cuts at nucleotide 6254), and the 10.1 kb fragment, which lacks the C_{1a} coding sequence, was isolated. Similarly, TC41F2-1 was cotransformed with mutagenized fragments encompassing the C_{2a} domain and a "gapped" vector prepared as follows. Cp1512 was digested with PvuII (which cuts at nucleotide 7111) and BamHI (which cuts at nucleotide 8168), and the 9.7 kb fragment, which lacks the C_{2a} coding region, was isolated. For both transformations, cells were plated on SC-Leu (Rose et al., 1990) without supplemental cyclic AMP and incubated at room temperature. Colonies from these plates were replica-plated to SC-Leu plates and incubated at 34 ^{o}C. Colonies that grew at 34 ^{o}C were expected to harbor stably replicating episomes that result from *in-situ* recombination between the gapped vector

and a PCR-amplified fragment and that, as a result of this recombination, express an adenylyl cyclase that has constitutive (*i.e.*, $G_{s\alpha}$-independent) activity.

In all cases, the dependence on plasmids for growth under selective conditions (on SC-Leu at 34 ^{o}C) was further tested by allowing spontaneous loss of plasmid under non-selective conditions (on YEPD with supplemental cyclic AMP at 30 ^{o}C), followed by re-testing under selective conditions. Plasmids were rescued from those colonies whose growth under selective conditions was dependent on plasmid; the plasmids were then reintroduced into naïve host strain TC41F2-1, and the transformants were tested for growth on SC-Leu at 34 ^{o}C and at 37 ^{o}C. All plasmids reported here confer plasmid-dependent growth at 34 ^{o}C when introduced into TC41F2-1.

Mutagenesis of adenylyl cyclase domains -- The mutants that appeared to confer the highest level of constitutive activity in yeast were analyzed using the soluble mammalian adenylyl cyclase system. The vectors pQE60-H6-VC_1(364-591)Flag (Sunahara et al., 1997a), pQE60-IIC_2-H6 (Whisnant et al., 1996) and pQE60-ArgC-IIC_2 (Tesmer et al., 1999) served as templates for site-directed mutagenesis (QuikChange™, Stratagene). The mutations in the C_{1a} domain were made in the conserved residues in the C_{1a} domain from canine type V adenylyl cyclase corresponding to the mutations in the rat type IIC_1 domain identified in the yeast screen. The mutations in the C_2 domain identified in the yeast screen were made in the C_2 domain of rat type II adenylyl cyclase using the vector pQE60-IIC_2-H6. The vector pQE60-ArgC-IIC_2 was used to generate a non-tagged IIC_2-K1014N. Sequences of synthetic mutagenic sense and antisense primers are available upon request. All plasmids were prepared using QIAGEN (Valencia, CA) products and sequenced by the McDermott Center DNA sequencing core by Jay Hunter.

Expression and Purification of Proteins -- Wild type and mutant H6-VC_1(591)Flag and IIC_2-H6 were expressed in *E. coli* and purified as previously described (Sunahara et al., 1997a; Whisnant et al., 1996; Hatley et al., 2002). Non-tagged ArgC-IIC_2 wild type and mutant K1014N were

expressed and purified as described (Tesmer et al., 1997b; Tesmer et al., 2002). Bovine $G_{s\alpha}$ (short form) was purified and activated by GTPγS as described (Lee et al., 1994). H6-VC_1(580) was purified as described by Tesmer *et. al* with a few modifications (Tesmer et al., 2002). First, the expression was performed at room temperature. Next, the eluate from the Talon™ column (CLONETECH, Palo Alto, CA) was diluted 5-fold in a buffer containing 20 mM NaHEPES pH 7.0, 2 mM $MgCl_2$, 1 mM EDTA, and 2 mM DTT and applied to a HiTrap Q (1 mL) column (Amersham Pharmacia Biotech, Piscataway, NJ). This HiTrap Q column retained the majority of contaminating proteins, while H6-VC_1(580) flowed through. This flow through was applied to a Mono S HR5/5 column (Amersham Pharmacia Biotech) as previously described (Tesmer et al., 2002).

Adenylyl Cyclase Assays -- Adenylyl cyclase activity was measured as described by Smigel (Smigel, 1986). All assays were performed in a volume of 100 μl for 10 min at 30 ^oC. The final concentration of $MgCl_2$ was 10 mM. The concentration of ATP was 1 mM unless otherwise stated. An ATP regenerating system was not included in the assays. Kinetic constants were determined by varying MgATP from 20 μM to 5 mM with a fixed excess of Mg^{2+}. Unless otherwise stated, assays contained a limiting concentration of VC_1 and an excess of IIC_2 for both wild type and mutant assays. All specific activities reported are with respect to the concentration of the limiting domain. Each experiment was repeated two or three times.

Gel Filtration -- $G_{s\alpha}$ and adenylyl cyclase domains were separated by FPLC with tandemly arranged Superdex 75 and 200 (HR10/30) columns (Amersham Pharmacia Biotech). Proteins were eluted in a buffer containing 20 mM NaHEPES pH 8.0, 2 mM $MgCl_2$, 1 mM EDTA, 2 mM dithiothreitol, and 50 mM NaCl. Flow rates were 0.2 ml/min, and 300-μl fractions were collected. An aliquot of each fraction (15 μl) was analyzed by gel electrophoresis on a 15% SDS-polyacrylamide gel and stained with Coomassie blue.

Talon Column Binding Studies -- H6-VC_1(591)Flag (15 μM) and ArgC-IIC_2 wild type or K1014N (75 μM) were incubated on ice for 15 min in 150 μl of buffer A (20 mM NaHEPES pH 8.0, 2 mM $MgCl_2$, 1 mM EDTA, and 50 mM NaCl) in the presence or absence of 100 μM forskolin, 100 μM 2'd3'-AMP, and 100 μM PP_i. The mixture was applied to a 25μl Talon™ column equilibrated with buffer A. The column was washed twice with 100 μl of buffer B (50 mM Tris-HCl pH 8.0, 50 mM NaCl, 2 mM $MgCl_2$, 100 μM forskolin, 100 μM 2'd3'-AMP, and 100 μM PP_i). The protein was then eluted with 100 μl of buffer B containing 100 mM imidazole. An aliquot (15 μl) of the eluate was resolved by gel electrophoresis on a 15% SDS-polyacrylamide gel and stained with Coomassie blue.

Isolation of the VC_1:IIC_2-K1014N:Forskolin:2'd3'AMP:PP_i Complex for Crystallography -- A method for large-scale production of the VC_1:IIC_2-K1014N complex was adapted from the small-scale procedure detailed above. H6-VC_1(580) was used immediately after purification and not frozen before use. H6-VC_1(580) (100 μM) and ArgC-IIC_2-K1014N (500 μM) were incubated on ice for 30 min in the presence of 500 μM forskolin, 500 μM 2'd3'-AMP, and 500 μM PP_i in 1-2 ml of buffer C (20 mM NaHEPES pH 7.0, 2 mM $MgCl_2$, 1 mM 2-mercaptoethanol, and 50 mM NaCl). The mixture was applied to a 500 μl Talon™ column previously equilibrated with buffer D (20 mM NaHEPES pH 7.0, 2 mM $MgCl_2$, 1 mM 2-mercaptoethanol, 50 mM NaCl, 100 μM forskolin, 100 μM 2'd3'-AMP, and 100 μM PP_i). The column was washed with three column volumes of buffer D. The complex was eluted in 500 μl fractions of buffer D containing 100 mM imidazole. The peak fractions were pooled, concentrated, and exchanged into buffer E (20 mM NaHEPES pH 7.0, 2 mM $MgCl_2$, 1 mM EDTA, 10 mM DTT, 50 mM NaCl, 100 μM forskolin, 100 μM 2'd3'-AMP, and 100 μM PP_i) using an Ultrafree-4 spin concentrator (10 kDa cutoff; Millipore, Bedford, MA).

Crystallization of the VC_1:IIC_2-K1014N:Forskolin:2'd3'AMP:PP_i Complex – Crystals of the complex were obtained by vapor diffusion at 4 °C using either hanging or sitting drops. The hanging drop experiments were performed by depositing 2 μl of concentrated complex on a clean 22 mm siliconized glass coverslip (Hampton Research, Laguna Niguel, CA) and adding an equal volume of well solution. This coverslip was inverted over a well of a VDX plate (Hampton Research) containing 1 ml of well solution. For the sitting drop experiments, the concentrated complex was added to the central reservoir of a Cryschem plate (Hampton Research), and an equal volume of well solution was added. Both hanging and sitting drop plates were sealed with a bead of silicone grease around the lip of the well. Initial screens were performed using the following commercial crystal screens: Crystal Screen Lite (Hampton Research) and Wizard I and Wizard II (Emerald BioStructures, Bainbridge Island, WA). The Hampton Crystal Screen Lite was also performed by diluting 500 μl of the provided well solution with 500 μl of filtered Milli-Q water. The leads from these commercial screens were used to design grid screens. Seeding experiments were performed using a pulled glass capillary or cat's whisker to extract crystals and drag crystal fragments over the surface of the drop to be seeded. Additive screens were conducted using Hampton Research's Additive Screens 1, 2, and 3 following manufacturer's recommendations. The additive screen was done using the sitting drop method and the well condition was as follows: 1.5 M ammonium sulfate, 100 mM cacodylate pH 6.4, and 200 mM sodium chloride.

Results

Genetic Screen – Manfredi's group isolated 24 plasmids that conferred cyclic AMP-independent growth at 34 °C when introduced into strain TC41F2-1; this phenotype is presumed to reflect the expression of $G_{s\alpha}$-independent adenylyl cyclase activity. The mutagenized regions of these plasmids were sequenced, and their mutations are listed in Table 1. Cp4465, which encodes ACII

I259V/Y402C, was used to generate the two plasmids ACII I259V and ACII Y402C; similarly, Cp4522 was used to generate two plasmids encoding ACII M253V and ACII C305R. In this way, it was found that ACII C305R and ACII Y402C are constitutively active mutants. In several cases, mutations that were observed in genetically-selected plasmids were engineered by site-directed mutagenesis into wild type ACII to test their effects independently of coincident mutations. Table 2 lists the mutations and the growth characteristics of all constitutively active ACII single mutants.

It is possible that plasmid-dependent variations in copy number could account for the phenotypes conferred by the ACII mutants. That possibility was addressed by subcloning six of the genetically-selected missense mutants into centromeric vectors whose copy number is stably maintained at 1-3 per cell (Clarke and Carbon, 1980). In each case, the phenotype conferred by the centromeric plasmid was the same as that imparted by its high-copy equivalent (data not shown). Thus, it is suggested that all 14 mutants listed in Table 2 have elevated specific activities.

Adenylyl Cyclase Assays -- The mutations identified in the yeast screen were generated in the soluble mammalian adenylyl cyclase system to facilitate biochemical characterization. The mutations were made in the C_{1a} domain of canine type V adenylyl cyclase at the residues that correspond to those mutated in the C_{1a} domain of rat type II adenylyl cyclase. The mutations F298Y, C305R, N315S, K334R, and V377I from rat $ACIIC_{1a}$ were made to the following residues in canine $ACVC_{1a}$: F400Y, C407R, N417S, K436R, and V479I. A similar screen in yeast using rat type IV adenylyl cyclase (data not shown) generated mutations Y265H, V388I, G968S, and K998N; these mutations were made in VC_1 and IIC_2 respectively as Y383H, V506I, G970S, and K1014N. No conversion was necessary for the remainder of the mutations since they were already in rat $ACIIC_2$. All mutant proteins, with the exception of C407R, K436R, and V506I, were expressed and purified to degrees comparable to levels described previously for the wild type proteins. The mutants were assayed for activity at limiting concentrations of the VC_1 domain and increasing concentrations of the IIC_2

domain in the absence of activators or in the presence of forskolin. The apparent affinity between the two soluble domains of adenylyl cyclase was expressed as the EC_{50}, and the apparent V_{max} was extrapolated from the asymptote of the curve. The apparent EC_{50}'s and V_{max}'s for the mutant proteins are summarized in Table 3.

Surprisingly, most mutant proteins displayed only modest differences in the apparent affinity or V_{max} when compared with the corresponding wild type domain. Mutants C407R and K436R yielded very little protein with no measurable adenylyl cyclase activity. The V506I mutation in VC_1 caused an increase in basal enzymatic activity and a 3-fold decrease in EC_{50} in the presence of forskolin. However, the low level of expression of this protein precluded detailed characterization. The mutations I1010M, K1014N, and P1015Q, all located in the β4-β5 loop of IIC_2, caused an increase in V_{max} in both the basal and the forskolin-stimulated conditions.

The K1014N mutation appeared to cause the largest degree of constitutive activation; IIC_2-K1014N was thus purified to homogeneity by anion-exchange chromatography and characterized biochemically. The apparent affinity of IIC_2-K1014N for wild type VC_1 was determined in various activation states. The activities shown in Fig. 1 were determined with a limiting concentration of VC_1 and either variable concentrations (Figure 2-1A, B, and C) or a saturating concentration (Figure 2-1D) of IIC_2. The apparent affinity for VC_1 and the maximal catalytic activity were both three fold greater for the K1014N mutant compared to wild type IIC_2 in the absence of activators (Figure 2-1A). Similarly, the apparent affinity of IIC_2-K1014N for VC_1 was three-fold higher in the presence of forskolin (Figure 2-1B) and ten-fold higher in the presence of $G_{s\alpha}$-GTPγS (Figure 2-1C) compared to its wild type counterpart. No significant changes were observed in V_{max} under these conditions. There was no difference in either the EC_{50} or the V_{max} in the presence of both forskolin and $G_{s\alpha}$-GTPγS, demonstrating that the mutation does not create a hyperactive enzyme. Similar results were observed using limiting concentrations of IIC_2 and varying concentrations of VC_1 (data not shown).

The apparent affinity of $G_{s\alpha}$-GTPγS for adenylyl cyclase is shown in Figure 2-1D. The EC_{50} for $G_{s\alpha}$-GTPγS was 0.05 μM for K1014N compared to 0.4 μM for the wild type IIC_2. These values were 6 and 25 nM respectively in the presence of forskolin.

Determination of Kinetic Constants - The kinetic constants for substrate were also determined under various conditions (Figure 2-2 and Table 4). Reconstituted adenylyl cyclase containing IIC_2-K1014N exhibited a K_m for ATP that was 6-fold less than that observed with the wild type protein under basal conditions. As noted above, V_{max} is increased under this condition. Assays performed with activated $G_{s\alpha}$ revealed a 6-fold decrease in the K_m for ATP when the mutation was present. No changes were obvious in the presence of forskolin, with or without $G_{s\alpha}$-GTPγS.

Gel Filtration of the VC_1/IIC_2-K1014N/$G_{s\alpha}$ Complex -- Purified H6-VC_1 and wild type or IIC_2-K1014N were combined with $G_{s\alpha}$-GTPγS and gel-filtered using tandem Sephadex 75 and 200 columns. Fractions were analyzed by SDS-PAGE (Figure 2-3). In the absence of forskolin, there was no evidence for formation of a complex between VC_1, wild type IIC_2, and $G_{s\alpha}$ (Figure 2-3B). The largest apparent species (78 kDa) is likely a heterodimer consisting of IIC_2 and $G_{s\alpha}$; similar results have been reported previously (Sunahara et al., 1997a). In contrast, protein in the mixture of VC_1, IIC_2-K1014N, and $G_{s\alpha}$ eluted as two major peaks with the largest species representing a 100-kDa complex. Analysis by SDS-PAGE indicated a complex of VC_1, IIC_2-K1014N, and $G_{s\alpha}$ with an apparent stoichiometry of 1:1:1 (Figure 2-3A). Similar results were observed with wild type IIC_2 only when forskolin was present (Sunahara et al., 1997a).

Isolation of a Complex of VC_1/IIC_2/Forskolin/2'd3'AMP:PP_i -- Purified H6-VC_1 and non-tagged wild type or IIC_2-K1014N were combined and applied to a metal chelate chromatographic column (Talon™,CLONTECH). Samples were eluted with imidazole and analyzed by SDS-PAGE (Figure 2-4). No complex was detected by SDS-PAGE when H6-VC_1 and IIC_2 or IIC_2-K1014N were

incubated with forskolin. Some IIC_2-K1014N was retained on the column in the presence of the P-site inhibitor 2'd3'AMP:PP_i; wild type IIC_2 was not. When both forskolin and 2'd3'AMP:PP_i were present, a complex of H6-VC_1 and either IIC_2 or K1014N was isolated. These complexes had apparent stoichiometries of 1:1 (determined by scanning densitometry). Similar results were obtained when Mn^{2+} replaced Mg^{2+} (data not shown).

Crystallization of VC1/IIC_2/Forskolin/2'd3'AMP:PP_i Complex – The conditions determined in the small-scale experiment were used to isolate a complex of H6-VC_1(580) and IIC_2-K1014N in the presence of forskolin and 2'd3'-AMP:PPi on a large-scale to screen for crystallization conditions. The VC_1 construct terminating at lysine 580 replaced the VC_1(591) construct used above since it has been shown to crystallize more readily in the $G_{s\alpha}$:C_1:C_2 co-crystal (Roger Sunahara, personal communitcation). The complex H6-VC_1(580) and IIC_2-K1014N eluted from the Talon column had an apparent stoichiometry of 1:1 (determined by scanning densitometry) (Figure 2-5). Production of this stoichiometric complex (1.5 to 3 mg) was very reproducible, and the product was used for crystallization trials.

Commercial crystal screens were used to determine well conditions for crystallization of the VC_1:IIC_2-K1014N complex. Screens at either 16 ^oC or 20 ^oC yielded precipitates. When the screens were performed at 4 ^oC, crystals formed under several conditions. More detailed screens were performed based on the leads provided by the commercial screens. However, little improvement in the geometry of the crystals was achieved. A shower of rods formed in a background of precipitate after 4 days at 4 ^oC with the following well conditions: 2.5 M NaCl, 200 mM Li_2SO_4, and 100 mM sodium acetate pH 4.7 (Figure 2-6A). These crystals were not obtained again under similar conditions or when using the well solution from the commercial screen. Several of these rods were harvested and transferred to the well solution containing 15% glycerol. These crystals were extracted with a 100 μm cryoloop (Hampton) and flash-frozen in liquid nitrogen-cooled propane. They

diffracted to 8Å at the Advanced Photon Source synchrotron at Argonne National Laboratory (Andreas Loew, personal communication). Urchin-like crystals formed under several conditions using ammonium sulfate as a precipitate after 24 hours. The Hampton additive screens were performed, although no improvement in the quality of the crystals was achieved. Crystals from Figure 2-6B were used for seeding, resulting in formation of some small plates protruding from the urchin-like crystals and haystacks of thin rods (Figure 2-6C and D). Screens were performed using the sitting drop method, decreasing the ratio of well solution added to the complex, and using an oil-overlay technique to slow the rate of equilibration between the well and the complex. Approximately 2000 conditions were explored in grid screens using the methods detailed above. None of these methods improved the quality of the crystals.

Discussion

The crystal structure of the cytosolic portions of adenylyl cyclase demonstrates that the C_1 and C_2 domains are arranged as a pseudo-two-fold symmetrical dimer (see inset in Figure 2-7) (Tesmer et al., 1997b). The contributions of several residues within each domain to substrate and Mg^{2+} binding, as well as catalysis, have been investigated in previous studies. Adenylyl cyclases, and presumably guanylyl cyclases, contain palm domains. These domains were defined previously in DNA polymerases, enzymes that catalyze very similar reactions (Artymiuk et al., 1997). Crystal structures of the cytosolic domains of adenylyl cyclase have revealed significant conformational changes upon substrate binding (Tesmer et al., 1997b). The α1 and α2 helices and the α3 and β4 helix/strand of C_1 and the β7-β8 loop of C_2 collapse around the nucleotide and align the nucleotide and two metal ions for catalysis. Located within the active site is the β2-β3 loop of C_1, containing aspartate residues 396 and 440 that coordinate two Mg^{2+} ions. These divalent cations participate in deprotonation of the 3' hydroxyl of the ribose moiety (a critical step in the synthesis of cyclic AMP)

and stabilize the pentavalent transition state. The conserved aspartate residues are also found among DNA polymerases and guanylyl cyclases (Tesmer et al., 1997b; Tesmer et al., 1999; Liu et al., 1997; Doublie et al., 1998; Huang et al., 1998; Li et al., 1998).

The β4-β5 loop of C_2 supports the β2-β3 loop of C_1 (Figure 2-7). Perturbations in either the contact regions or the loop-fold could have dramatic effects on C_1:C_2 structure and hence activity. Several residues in both the β2-β3 loop and the α2 helix of C_1 and the β4-β5 loop of C_2 have been investigated by site-directed mutagenesis and have various effects on adenylyl cyclase activity. Structural evidence strongly suggests that D424 in the α2 helix and R434 in the β2 sheet of C_1 engage in extensive hydrogen bonding with the β4-β5 loop of C_2 (Figure 2-7). D424 forms a salt bridge with R434 and a hydrogen bond with the backbone nitrogens of A1012 and Q1013. The side chain of R434 forms a hydrogen bond with the side chain of Q1016; the backbone carbonyl of R434 forms a hydrogen bond with the side chain of Q1013. Mutations of either D424 or R434 have previously been shown to have detrimental effects on adenylyl cyclase activity. Mutations of these residues have broad effects on cyclase activity without affecting the affinity of the enzyme for $G_{s\alpha}$, as follows: R434A increases the IC_{50} for P-site inhibitors (Tang et al., 1995); R434S increases the K_m for MgATP, the K_i for ATPγS, and the EC_{50} with Mg^{2+} (Zimmermann et al., 1998); D424A and D424N decrease forskolin- and $G_{s\alpha}$-stimulated enzymatic activity (Tang et al., 1995; Zimmermann et al., 1998).

Mutations of residues in the β4-β5 loop of C_2 have also been shown to affect adenylyl cyclase activity. The mutations Y1017A and D1018A (Y999A and D1000A in type I adenylyl cyclase) obliterate activity without eliminating $G_{s\alpha}$ binding (Tang et al., 1995). D1018 coordinates substrate binding through the purine ring and is responsible for dictating nucleotide specificity (Tesmer et al., 1997b; Sunahara et al., 1998). Alterations in neighboring side chains that perturb the conformation of the amino acid chain backbone would likely perturb activity. Residues I1010, K1014 and P1015,

investigated in this study, are all located in the β4-β5 loop of C_2, and therefore, are intimately involved in the arrangement of the β2-β3 loop of C_1. The mutations I1010M, K1014N, and P1015Q are all within bonding distance of the β2-β3 loop of C_1. The P1015Q mutation, which displays slightly elevated affinity between C_1 and C_2, was previously identified as a second site suppressor of a catalytically inactive mutant (N1025S), but unfortunately was not characterized alone (Yan et al., 1997b). Prediction of the positions of the side chains of the β4-β5 loop of C_2 is difficult because of its flexibility; however, some explanations of the activating mutations can be extracted from the crystal structure. We expect K1014 upon mutation to asparagine to pack between adjacent glutamines in the β4-β5 loop of C_2 and form a stabilizing hydrogen bond with R434 in the β2-β3 loop of C_1. P1015Q could rearrange the β4-β5 loop, causing a more active conformation. I1010M fills space in the hydrophobic pocket with a larger hydrophobic residue. The introduction of new side chain interactions contributed by mutations I1010M or K1014N or the removal of main-chain constraints with mutation P1015Q may alter the C_1 β2-β3 loop and enhance activity. Taken together, these mutations suggest that proper formation of a competent active site is inhibited by decreased and promoted by increased interactions between the β4-β5 loop of C_2 and the β2-β3 loop of C_1.

The β4-β5 loop of the C_1 domain has a congruous interaction with the β2-β3 loop of C_2 because of the pseudosymmetrical structure of adenylyl cyclase. Of the clones obtained from the genetic screen, only V506I displayed significant enhancement of activity when tested in the soluble adenylyl cyclase system. V506I adds a methyl group that may form a primary contact with forskolin and increase the hydrophobicity of the forskolin-binding pocket. Another possible explanation is that substitution of the valine by isoleucine may enhance C_1:C_2 interactions by altering Van der Waals contacts with neighboring residues in the β4-β5 loop of C_1. The structural effect of this minor change at the base of the loop may be amplified along the length of the loop and thus alter interactions with the β2-β3 loop of C_2.

It is difficult to determine why other mutations displayed strong phenotypes in yeast but failed to produce substantial changes in the soluble adenylyl cyclase assays. As demonstrated in this study, mutations that increase favorable interactions between the β2-β3 loop of C_1 and the β4-β5 loop of C_2 likely account for the increased activity that was observed. Alternatively, increased activity may be related to the fact that the C_1 and C_2 proteins were expressed individually and there form homodimers. Mutations in the interface region may alter homodimerization. Mutations that impair homodimerization may favor heterodimerization and hence increase adenylyl cyclase activity. Perhaps more likely, the lack of observable increased basal activity in the *in-vitro* assays of many of the mutants could be explained by the yeast screen's inherent sensitivity to small changes in cyclic AMP concentrations. The lack of change in activity *in vitro* could also be explained by the ablation of the membrane domains and the putative regulatory C_{1b} domain in the soluble constructs. For example, the F400Y mutation in native type V adenylyl cyclase has been shown to increase both basal activity and sensitivity to the activators $G_{s\alpha}$ and forskolin and to abrogate inhibition by $G_{i\alpha}$ (Zimmermann et al., 1999). However, when assayed in the soluble system, this mutation caused no increase in basal activity or sensitivity to forskolin compared to the wild type enzyme. The contributions of the C_{1b} domain to activity are not known, nor is there any structural information on this domain.

As mentioned previously, all complexes of adenylyl cyclase whose structures have been determined to date contain both $G_{s\alpha}$ and forskolin. Endogenous forskolin-like substances have yet to be discovered, begging the questions of the physiological significance of the C_1:C_2:forskolin:$G_{s\alpha}$-GTPγS structure and the degree to which it resembles the structure of C_1:C_2:$G_{s\alpha}$-GTPγS. Or rather, what is the mechanism of activation of adenylyl cyclase by forskolin? This is a particularly interesting question, since some forms of adenylyl cyclase (types II, IV, V, VI and VII) are activated synergistically by forskolin and $G_{s\alpha}$-GTPγS while others are activated only additively (types I, III and

VIII) (Sunahara et al., 1996). The K1014N mutation facilitated the isolation of a C_1:C_2:$G_{s\alpha}$-GTPγS complex in the absence of forskolin. Determination of the structure of this complex would further our understanding of activation of adenylyl cyclase by the diterpene.

The crystal structure of a homodimer of the C_2 domain bound with two molecules of forskolin has also been determined (Zhang et al., 1997). The structure also contains a two-fold symmetrical arrangement of the domains and has been used as a model for the basal, nonactivated form of adenylyl cyclase. However, the presence of two forskolin molecules in a complex without an active site inherently eliminates this structure as a precise model of nonactivated adenylyl cyclase.

The structure of the constitutively active mutants described herein, particular that of VC_1 associated with IIC_2-K1014N, may overcome the deficient C_1:C_2 affinity and may have structural features that more closely approximate the low activity basal state. Despite the ability to consistently purify a stoichiometric complex VC_1 and IIC_2-K1014N, crystals suitable for diffraction were not obtained. The crystals that were obtained have a striking similarity to the crystals of the IIC_2 homodimer (Zhang et al., 1997). This could suggest that the VC_1:IIC_2 complex dissociated and the crystals obtained were in fact the IIC_2 homodimer. Therefore, the major obstacle to structural determination of a low activity state of the adenylyl cyclase catalytic core remains the low affinity of the VC_1 and IIC_2 heterodimer. The covalent linkage of the VC_1 and IIC_2 domains would increase the local concentration of the two domains and favor heterodimerization. Chapter 3 will explore the use of a new technology to post-translationally, covalently link proteins to link the C1 and C2 domains of adenylyl cylase. This linkage will increase the affinity of the C1 and C2 domains while decreasing the homodimerization

TABLE 2-1.

Growth of yeast harboring plasmids encoding mutant adenylyl cyclases that confer independence of $G_{s\alpha}$-GTPγS and cyclic AMP [a]

Plasmid	Amino Acid Substitutions	Growth at 34°C [a]	Growth at 37°C [a]	Plasmid	Amino Acid Substitutions	Growth at 34°C [a]	Growth at 37°C [a]
Cp4461	Y496C	2	0.5	Cp4478	K896E	1	0
Cp4462	L391F V457A	2	0.25	Cp4521	V377I	2	1.5
Cp4463	V311A K326E Y341H	2	1	Cp4522	M253V C305R	2	1.5
Cp4464	K267E C305R	2	0.5	Cp4525	K334R	2	1.5
Cp4465	I259V Y402C	2	0	Cp4526	E309K	2	1
Cp4466	N315S	2	1	Cp4527	V429A	2	1
Cp4467	F298Y	2	0.25	Cp4528	L210P N224N E329D	0.5	0
Cp4468	T494P	2	1	Cp4530	Y826H E879V	2	0.5
Cp4474	P1015Q	2	0.5	Cp4531	I744T L850P V1026G	2	0.5
Cp4475	L801P M804V R821G L877P Y882H	2	0	Cp4533	I1010M	2	1.5
Cp4476	L877P	2	0	Cp4534	K896E	2	1.5
Cp4477	I740V S788G L877P Y882H	2	1.5	Cp4535	C743R L769P T779A I1010M	2	1

[a] These data were collected by John Manfredi and coworkers at Cadus Pharmaceuticals.

[b] Growth was determined by visual inspection of streaked colonies. The rate of growth was given a relative score so that a colony with a growth of 2 grew to a given size 4 times faster than a colony with a growth of 0.5. TC41F2-1 transformed with CP1512, which encodes wild-type ACII, is 0.25 at 34°C and 0 at 37°C.

TABLE 2-2.

Effect of Single Amino Acid Substitutions in Adenylyl Cyclase [a]

Plasmid	Amino Acid Substitutions	Corresponding residue in VC_1	Growth at 34°C [b]	Growth at 37°C [b]
Cp4461	Y496C	N592	2	0.5
Cp4466	N315S	N417	2	1
Cp4467	F298Y	F400	2	0.25
Cp4468	T494P	G590	2	1
Cp4474	P1015Q	--[c]	2	0.5
Cp4476	L877P	--[c]	2	0
Cp4521	V377I	V479	2	1.5
Cp4525	K334R	K436	2	1.5
Cp4526	E309K	E411	2	1
Cp4527	V429A	A531	2	1
Cp4533	I1010M	--[c]	2	1.5
Cp4534	K896E	--[c]	2	1.5
Cp4563	Y402C	F504	2	1.5
Cp4702	C305R	C407	2	1

[a] These data were collected by John Manfredi and coworkers at Cadus Pharmaceuticals.

[b] Growth was determined by visual inspection of streaked colonies. The rate of growth was given a relative score so that a colony with a growth of 2 grew to a given size 4 times faster than a colony with a growth of 0.5. TC41F2-1 transformed with CP1512, which encodes wild-type ACII, is 0.25 at 34°C and 0 at 37°C.

[c] These mutants are in $ACIIC_2$.

TABLE 2-3.

Affinity and Activity of Purified Adenylyl Cyclase Mutants[a]

Mutant	Basal EC_{50}		Basal V_{max}		Forskolin EC_{50}		Forskolin V_{max}	
	μM	Ratio WT / mutant	nmol/min/mg	% of WT	μM	Ratio WT / mutant	μmol/min/mg	% of WT
Y383H [b]	5 (4.8)	0.96	460 (550)	84	1.7 (1.3)	0.76	85 (130)	65
F400Y [b]	18 (8)	0.44	380 (840)	45	3 (1.4)	0.47	89 (130)	68
C407R [b]	--	--	--	--	--	--	--	--
N417S [b]	7 (8)	1.10	260 (410)	63	1.5 (1.4)	0.93	56 (100)	56
K436R [b]	--	--	--	--	--	--	--	--
V479I [b]	12 (5.7)	0.48	310 (570)	54	1.9 (1.4)	0.74	37 (88)	42
V506I [b]	8 (7)	0.88	980 (660)	150	0.55 (1.7)	3.1	81(130)	62
L877P [c]	3 (2)	0.67	250 (170)	150	1.1 (1.5)	1.4	48 (44)	110
K896E [c]	3 (2.4)	0.80	150 (150)	100	1.3 (1.2)	0.92	34 (55)	62
G970S [b]	3 (4)	1.30	680 (650)	105	0.89 (0.83)	0.93	120 (93)	130
I1010M [c]	4.5 (2.7)	0.60	410 (190)	220	2.2 (1.3)	0.59	110 (65)	170
K1014N [c]	1.6 (2.4)	1.50	1200 (170)	700	0.33 (1.4)	4.20	78 (60)	130
P1015Q [c]	4.2 (2.2)	0.52	510 (160)	320	1.4 (1.2)	0.86	100 (56)	180

[a] The apparent affinities and maximal activities were determined for each mutant protein by assaying a limiting concentration of VC_1 (30 nM for basal; 2 nM with forskolin) with increasing concentrations of IIC_2 in the absence and presence of 100 μM forskolin. The values in parenthesis are the values for the wild type protein in each assay.

[b] protein purified by metal affinity column followed by anion exchange chromatography

[c] protein purified by metal affinity column

TABLE 2-4.

Substrate Kinetics with IIC_2 or IIC_2-K1014N[a]

Condition	Protein	$K_{m, ATP}$ (μM)	V_{max} (μmol/min/mg)
Basal	K1014N	210 ± 28	0.78 ± 0.23
	IIC2	1300 ± 310	0.20 ± 0.07
Forskolin	K1014N	380 ± 40	83 ± 8
	IIC2	840 ± 85	66 ±2
$G_{s\alpha}$	K1014N	35 ± 5	26 ± 1
	IIC2	220 ± 45	28 ± 4
$G_{s\alpha}$+ Forskolin	K1014N	160 ± 27	92 ± 18
	IIC2	210 ± 26	106 ± 24

[a] Experimental values are reported as the mean and range for two experiments or the ssandard error of the mean for three experiments.

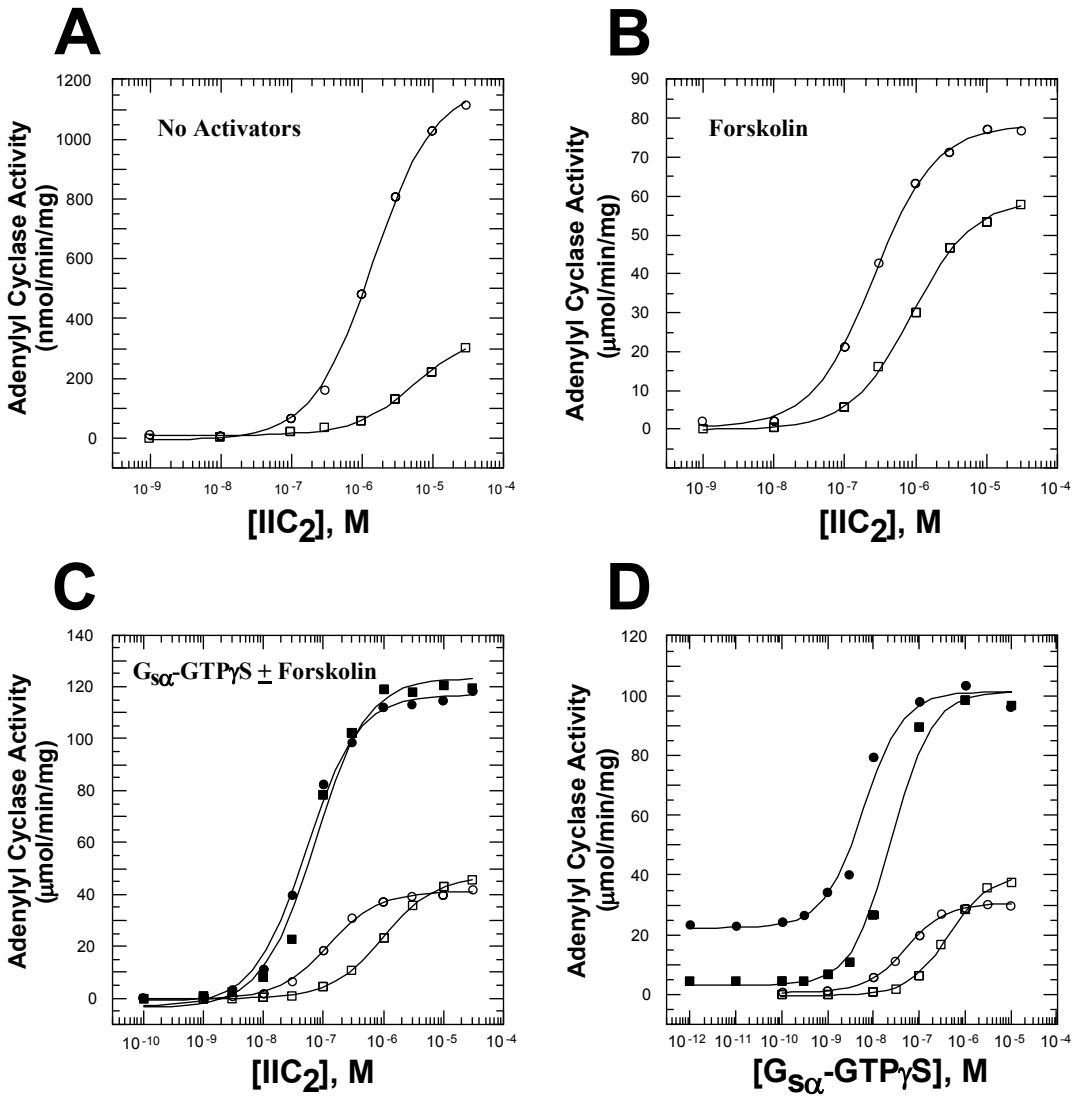


Figure 2-1. The apparent affinity of IIC_2-K1014N for VC_1 and $G_{s\alpha}$-GTPγS. VC_1 was mixed with wild type IIC_2 or IIC_2-K1014N and assayed for adenylyl cyclase activity in the presence of the indicated activators. **A.** Adenylyl cyclase activity was assayed with VC_1 (30 nM) and increasing concentrations of IIC_2 (□) or IIC_2-K1014N (○) in the absence of activators. **B.** Adenylyl cyclase was assayed with VC_1 (2 nM) and increasing concentrations of IIC_2 (□) or IIC_2-K1014N (○) in the presence of 100 µM forskolin. **C.** Adenylyl cyclase was assayed with VC_1 (2 nM) and increasing concentrations of IIC_2 (□) or IIC_2-K1014N (○) in the presence of 400 nM $G_{s\alpha}$-GTPγS with (filled symbols) and without (open symbols) 100 µM forskolin. **D.** Adenylyl cyclase was assayed with 2 nM VC_1 and 2 µM IIC_2 (□) or IIC_2-K1014N (○) with increasing concentrations of $G_{s\alpha}$-GTPγS in the presence (filled symbols) and absence (open symbols) of 100 µM forskolin.

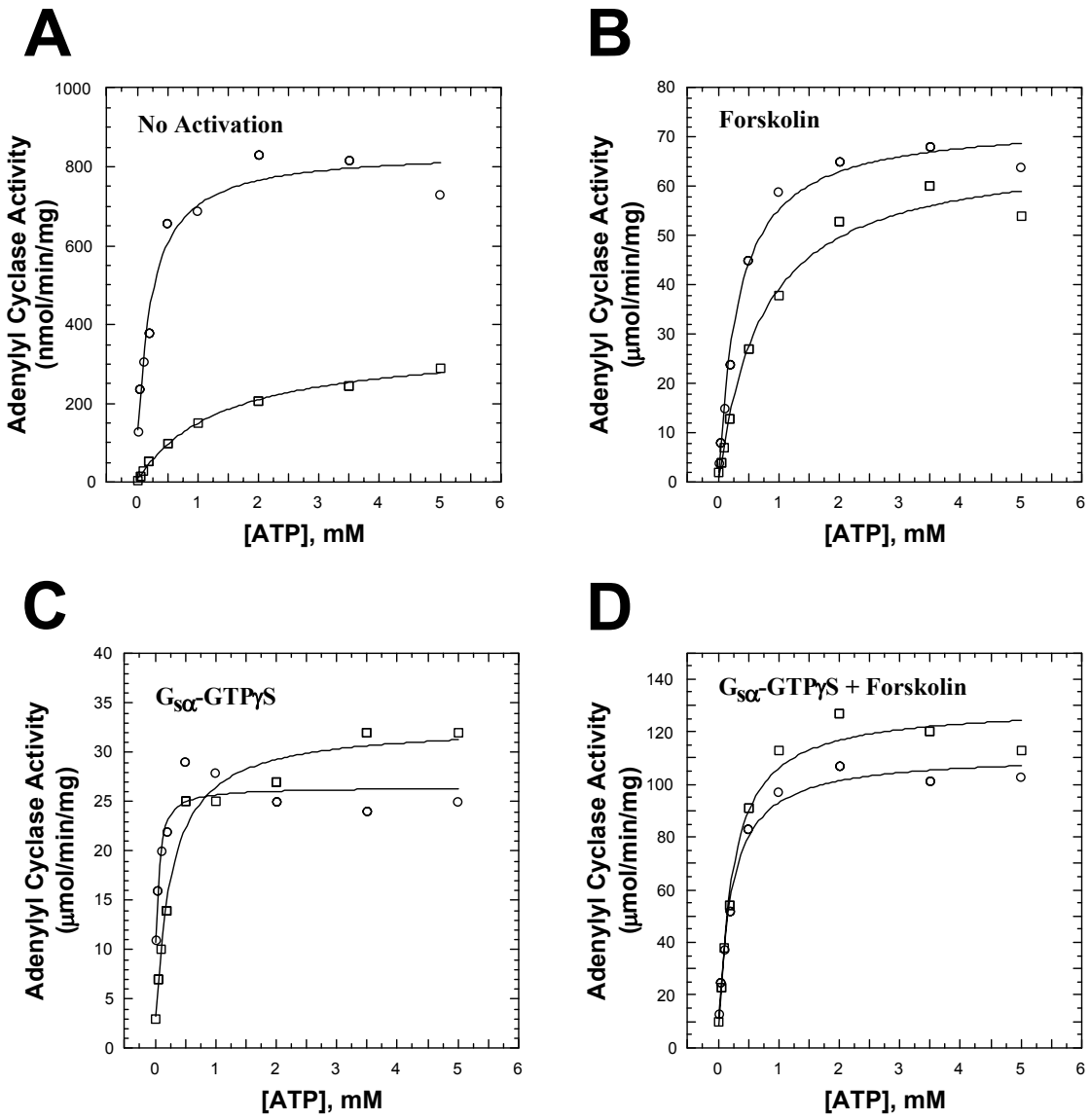


Figure 2-2. Substrate kinetics of IIC_2-K1014N. The kinetic constants for ATP were determined from the Michaelis-Menten plots generated by measuring adenylyl cyclase activity with a limiting concentration of VC_1 and an excess of IIC_2 while varying the ATP concentration from 20 μM to 5 mM. **A.** 30 nM VC_1 and 10 μM IIC_2 (□) or IIC_2-K1014N (○) were assayed in the absence of activators. **B-D.** 2 nM VC_1 and 3 μM IIC_2 (□) or IIC_2-K1014N (○) were assayed in the presence of 100 μM forskolin (**B**), 400 nM $G_{s\alpha}$-GTPγS (**C**), or 100 μM forskolin and 400 nM $G_{s\alpha}$-GTPγS (**D**).

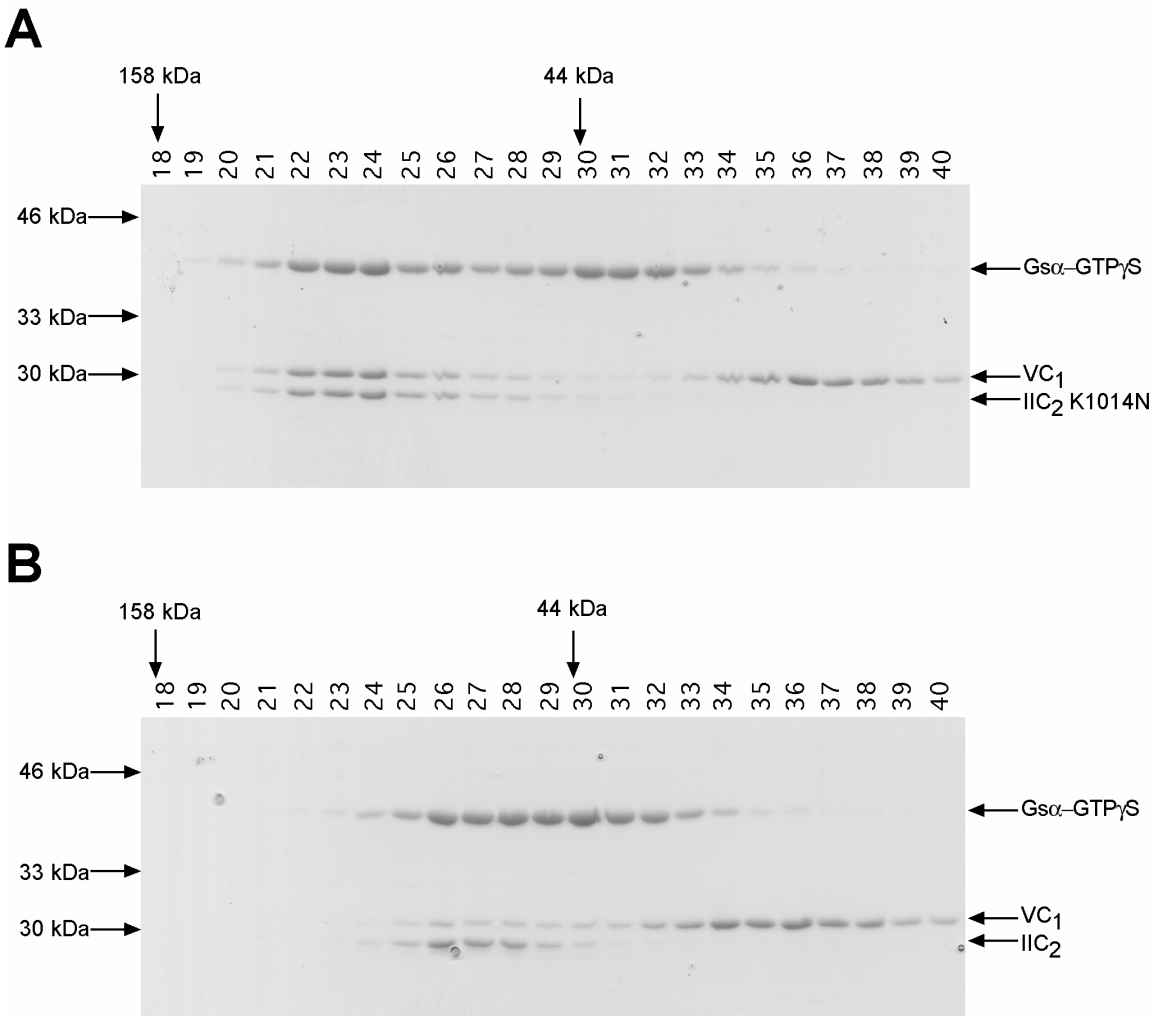


Figure 2-3. Gel filtration of VC_1, IIC_2-K1014N and $G_{s\alpha}$ in the absence of forskolin.
A mixture of VC_1 (100 μM), IIC_2-K1014N (50 μM) (**A**) or wild type IIC_2 (50 μM) (**B**) and $G_{s\alpha}$-GTPγS (50 μM) was applied to a Superdex 75 (HR10/30) gel filtration column in tandem with a Superdex 200 (HR10/30). Fractions 18-40 (15 μl of 300μl fractions) were resolved by SDS-PAGE on a 15% polyacrylamide gel and stained with Coomassie blue. The positions of elution of two gel filtration standards are indicated.

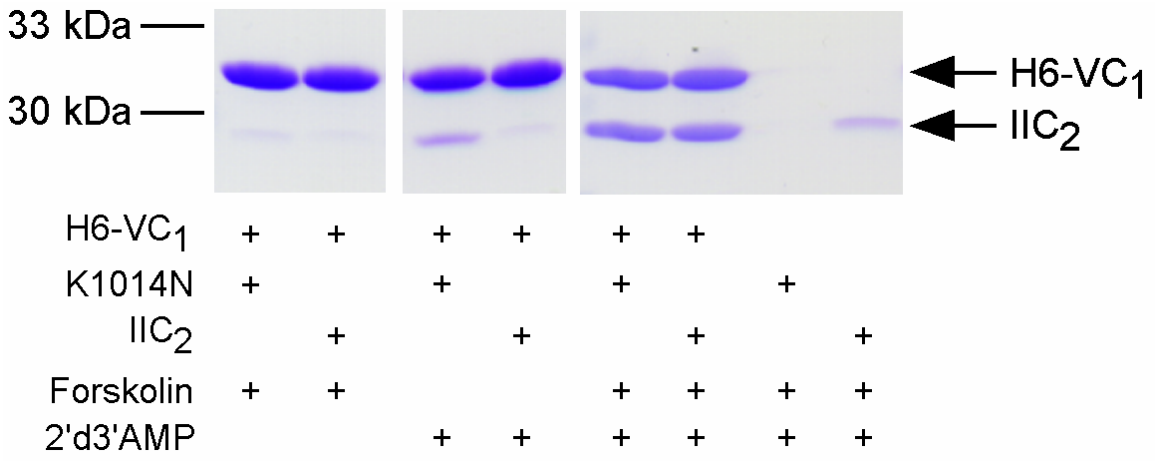


Figure 2-4. Complexes of VC_1 and IIC_2. IIC_2 or IIC_2-K1014N (75 μM) were incubated for 15 min on ice in the presence or absence of H6-VC_1 (15 μM) with 100 μM forskolin and/or 100 μM 2'd3'AMP•PP_i as indicated. The mixtures were applied to 25 μl Talon™ columns, which were washed and then eluted with imidazole. Aliquots of the eluates were resolved by SDS-PAGE on a 15% polyacrylamide gel and visualized by staining with Coomassie blue.

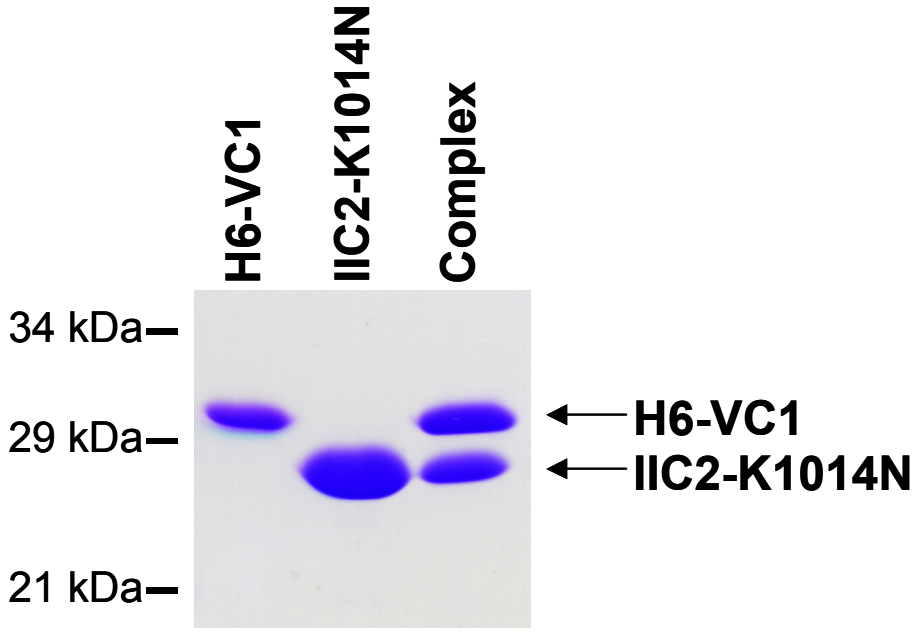


Figure 2-5. Purification of H6-VC1(580):IIC-K1014N:Forskolin:2'd3'-AMP•PPi complex. H6-VC1 (100 μM) and IIC2-K1014N (500 μM) were incubated on ice for 30 min in the presence of 500 μM forskolin and 500 μM 2'd3'-AMP•PPi. The VC1:IIC2-K1014N complex was purified on a 500 μl Talon™ column. 2 μg of H6-VC1(580), 3 μg of non-tagged IIC2-K1014N, and 4 μg of purified VC1:IIC2 complex were resolved by SDS-PAGE on a 15% polyacrylamide gel and stained with Coomassie Blue.

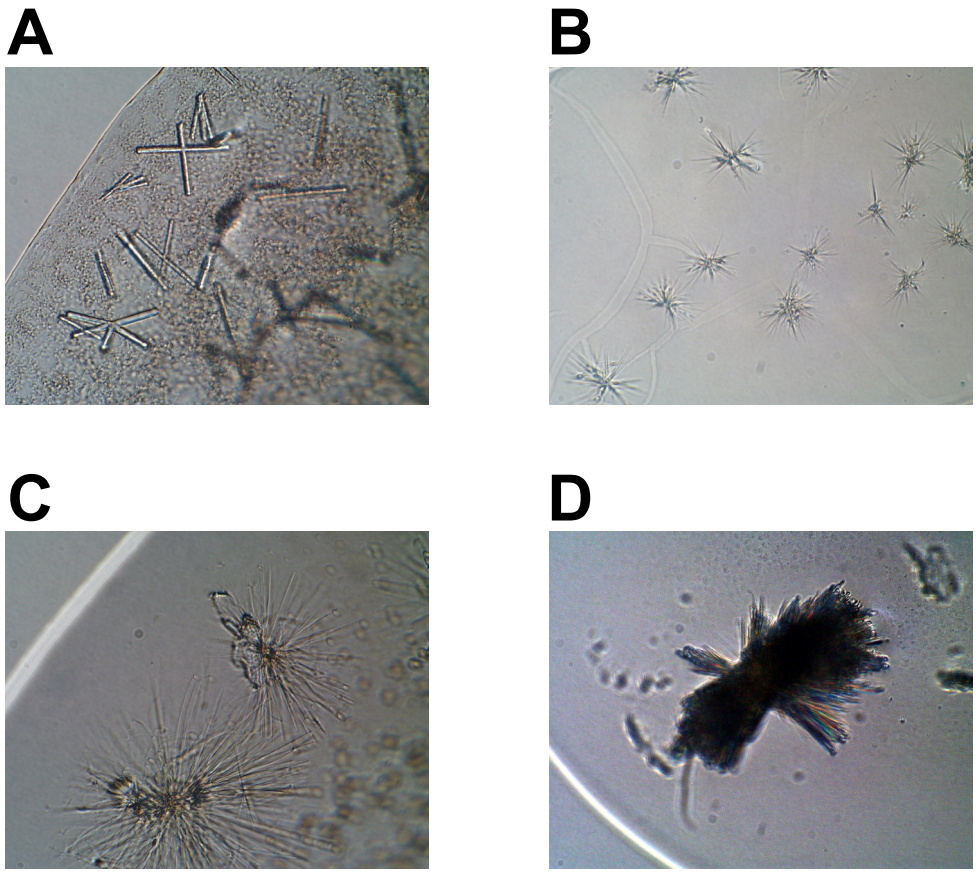


Figure 2-6. Crystals of VC1:IIC2-K1014N:Forskolin:2'd3'-AMP:PPi. Crystals in all panels were obtained by the hanging drop vapor diffusion method at 4 ^{o}C and viewed at 40X. The well solutions are as follows: **A**, 2.5 M NaCl, 200 mM Li_2SO_4, and 100 mM sodium acetate pH 4.7; **B**, 1.7 M $(NH_4)_2SO_4$, 100 mM cacodylate pH 6.4, 200 mM NaCl, 2.5% glycerol; **C**, 1.2 M $(NH_4)_2SO_4$, 100 mM cacodylate pH 6.4, 200 mM NaCl, 2.5% glycerol seeded with crystals from B. **D**, 800 mM $(NH_4)_2SO_4$, 100 mM cacodylate pH 6.4, 200 mM NaCl, 5% glycerol seeded with crystals from B.

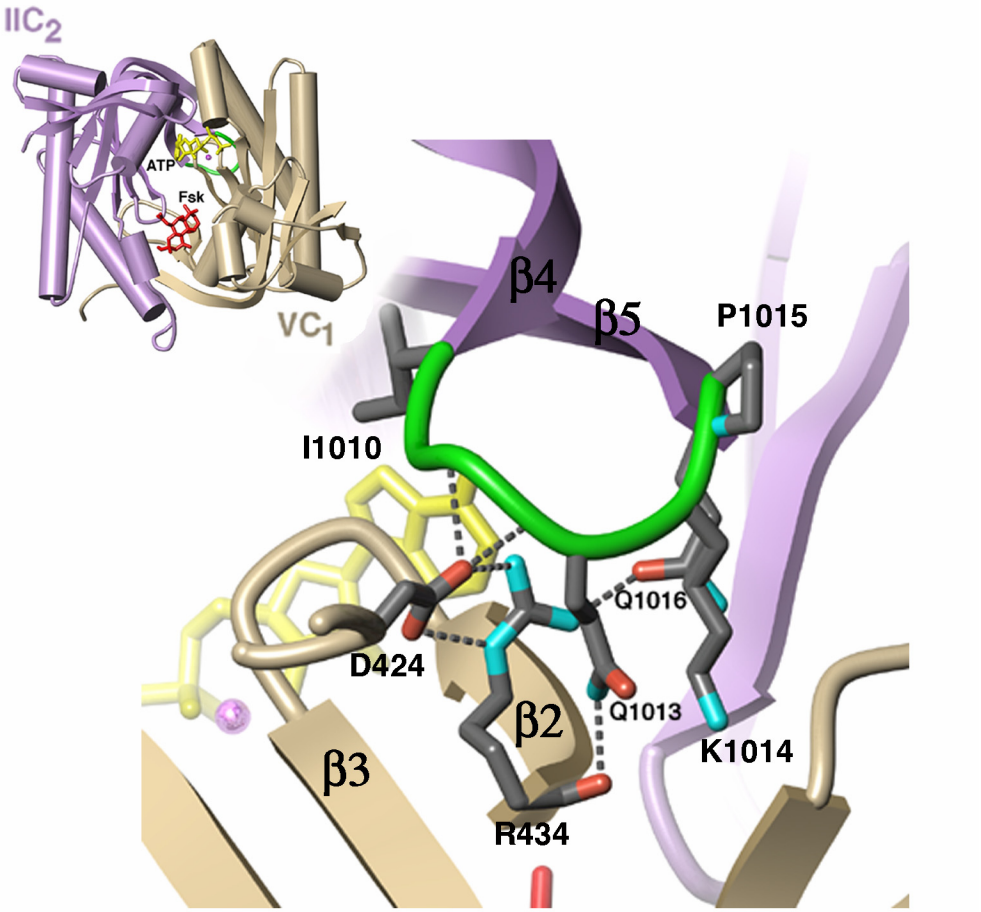


Figure 2-7. Catalytic core of adenylyl cyclase. *Inset,* the heterodimeric complex formed by VC_1 (khaki) and IIC_2 (mauve) viewed along its pseudo-twofold axis toward the hypothesized cytoplasmic face. Forskolin (Fsk) and ATP bind between VC_1 and IIC_2 and are shown as stick models. The β4-β5 loop of IIC_2 containing I1010M, K1014N, and P1015Q is highlighted in green. *Interactions of the β4-β5 loop of IIC_2.* The side chains of the β4-β5 loop (green) of IIC_2 (mauve) and their interactions with VC_1 (khaki) are shown. Dashed gray lines show side chain-side chain and side chain-main chain hydrogen bonds. Carbon atoms are gray, nitrogens blue, and oxygens red.

Chapter 3 – Intein-Mediated Protein Ligation of Adenylyl Cyclase Cytosolic Domains.

Introduction

Hypothetically, the low affinity of the C_1 and C_2 domains of adenylyl cyclase for each other is considered the major obstacle to the solution of a structure of the low activity state of the enzyme. The increased affinity upon addition of forskolin and the P-site inhibitor 2'd3'-AMP was insufficient to permit crystallization of the VC_1:IIC_2-K1014N heterodimer in the absence of $G_{s\alpha}$. The formation of a covalently linked VC_1-IIC_2 polypeptide could drive formation of a properly folded polypeptide with basal catalytic activity. However, previous attempts to express a fusion of the C_1 and C_2 domains were frustrated by a proteolysis and low expression levels (Tang and Gilman, 1995; Dessauer and Gilman, 1996).

A recent advance in protein engineering could provide a means of linking the two domains after they are expressed as individual peptides. A post-translational process termed protein splicing has recently been described. Protein splicing involves the excision of an internal peptide or intein and the ligation of the two flanking external peptides or exteins (Figure 3-1) (Perler and Adam, 2000). The mechanism of protein splicing has been elucidated by site-directed mutagenesis. Protein splicing occurs through a series of four nucleophilic substitution reactions by the mechanism outlined in Figure 3-2 (Perler, 1998; Shao and Kent, 1997; Noren et al., 2000). The side chain hydroxyl or thiol of a conserved serine or cysteine residue at the N-terminus of the intein attacks the carbonyl of the preceding amino acid. This results in the formation of a thioester bond at the N-terminal splice site. Next, this carbonyl is attacked by the hydroxyl or thiol of the serine, threonine, or cysteine residue at the C-terminus of the intein. This results in N-terminal cleavage and formation of a branched intermediate. Then, an invariant asparagine at the C-terminus of the intein undergoes cyclization to form a succinimide. This reaction cleaves the peptide bond and excises the intein. Finally, the linkage between the exteins undergoes a spontaneous acyl rearrangement that forms a peptide bond.

New England Biolabs Inc. has used inteins to design a protein purification kit called IMPACT-CN (Intein Mediated Purification with an Affinity Chitin-binding Tag). The *Sce* VMA intein from *Saccharomyes cerevisiae* and the *Mxe* GyrA intein from *Mycobacterium xenopi* have been modified for use in this affinity chromatography system (Chong et al., 1997; Evans et al., 1998; Southworth et al., 1999). The protein of interest is fused to the amino terminus of the intein. The C-terminal invariant asparagine of the intein has been mutated to alanine, blocking succinimde formation without affecting the acyl rearrangement at the N-terminus of the intein. Therefore, the process is stalled at the thioester intermediate. This thioester bond is more reactive than the native peptide bond and is readily cleaved with sulfur or nitrogen nucleophiles. The chitin binding domain (CBD) from *Bacillus circulans* is fused to the carboxyl terminus of the intein to facilitate separation by binding to chitin resin. The target protein is cleaved from the intein-CBD fusion protein following overnight incubation with dithiothreitol (DTT). DTT forms a thioester linkage with the C-terminus of the target protein. This thioester was shown to be reactive to nucleophilic attack by free cysteine (Chong et al., 1997).

Several groups quickly adapted this technique to ligate synthetic peptides containing N-terminal cysteines to the thioester-modified protein (Evans et al., 1998; Muir et al., 1998; Severinov and Muir, 1998). This process has been termed intein-mediated protein ligation or IPL (Evans and Xu, 1999b). IPL has been used to purify cytotoxic proteins that cannot be expressed in bacteria. Truncations of the cytotoxic proteins that lack enzymatic activity have been expressed and purified using the IMPACT system. The thioester-modified proteins are reacted with synthetic peptides with an N-terminal cysteine. The ligation of the peptide restores the amino acids missing from the truncations. Bovine pancreatic ribonuclase A and a restriction endonuclease from *Haemophilus parainfluenzae (Hpa*I) have been purified and enzymatic activity restored using this technique (Evans et al., 1998). Proteolytic cleavage of bacterially expressed proteins with a protease site adjacent to a

cysteine provides a recombinant protein with an amino-terminal cysteine that is capable of ligating thioester modified proteins (Xu et al., 1999). IPL may provide a method to chemically link the VC_1 and IIC_2 domains of adenylyl cyclase at the protein level, thus circumventing the problems of expressing the fusion protein in *E. coli.* A scheme for the linkage of VC1 and IIC2 is shown in Figure 3-3 and the results of that effort are described in this chapter.

Experimental Procedures

Construction of VC_1 intein fusion vector – The polymerase chain reaction was used to remove the FLAG epitope tag (Sigma-Aldrich, St. Louis, MO) from pQE60-H6-VC_1(591)FLAG. The sequence of the sense primer was 5'-GTCTTCACCTCCGAGAAATCATA-3' corresponding to the *Xho*I restricition endonuclease site in pQE60-H6. The sequence of the antisense primer was 5' GCGGCGAAGCTTGTGCCCAATGGAGTTGG-3' corresponding to the last 17 base pairs of VC_1(591) followed immediately by a *Hin*dIII restriction endonuclease site. The polymerase chain reaction product was digested with *Nco*I and *Hin*dIII and ligated to pQE60-H6 digested with the same enzymes to yield pQE60-H6-VC_1(591). A 3' SapI restriction endonuclease site was added to this vector using two complementary oligonucleotides designed with 5' *Hin*dIII and 3' *Blp*I cohesive ends. The sequence of the sense oligo was 5'-AGCTTGGTTGCGGAAGAGCAATTAGC-3'. The sequence of the antisense oligo was 5'-TCAGCTAATTGCTCTTCCGCAACCA-3'. The two oligonucleotides were phosphorylated with T4 polynucleotide kinase, annealed, and ligated into pQE60-H6-VC_1(591) digested with *Hin*dIII and *Blp*I. The resulting vector was digested with *Nco*I and partially digested with *Sap*I and ligated to the vector pTXB3 (New England Biolabs) digested with the same enzymes to give pTXB3-VC_1(591). An amino-terminal hexahistidine tag was added using complementary oligonuleotides with *Nco*I cohesive ends. The sequence of the sense oligonucleotide was 5'-ATGCATCACCATCACCATCACGG-3'. The sequence of the antisense

oligonucleotide was 5'-CATGCCGTGATGGTGATGGTGATG-3'. These oligonucleotides were phosphorylated, annealed, and ligated to pTXB3-VC_1(591) digested with *Nco*I. The resulting bacterial expression vector, pTXB3-H6-VC_1(591), encodes the protein sequence MHHHHHHA followed by amino acids 364 to 591 of canine type V adenylyl cyclase linked with amino acids KLG in frame with the *Mxe* GyrA intein and chitin-binding domain from the pTXB3 vector. Linkers were added between His591 of VC_1 and the N-terminal cysteine of *Mxe* GyrA intein using complementary oligonucleotides. The L1 linker encoding amino acids KLGSGTSSGTSSTGSAG was added using oligonucleotides witrh *Hin*dIII and *Spe*I cohesive ends. The L1 sense oligonucleotide was 5'-AGCTTGGTTCTGGTACCAGGTCTGGTACTTCTTCTACTGGTTCTGCTGGTTGCATCA CGGGAGATGCA-3'. The L1 antisense oligonucleotide was 5'-CTAGTGCATCTCCCGTGATGCAACCAGCAGAACCAGTAGAAGAAGTACCAGACCTGGTA CCAGAACCA-3'. These oligonucleotides were phosphorylated, annealed, and ligated to pTXB3-VC1(591) digested with *Hin*dIII and *Spe*I. The L3 linker encoding amino acids KLAAAGGMPPAAAGGMKL was added with oligonucleotides with *Hin*dIII cohesive ends. The L3 sense oligonucleotide was 5'-AGCTTGCTGCTGCTGGTGGTAGTATGCCGCCGGCTGCTGCTGGTGGTATGA-3'. The L3 antisense oligonucletide was 5'-AGCTTCATACCACCAGCAGCAGCCGGCGGCATACCACCAGCAGCAGCA-3'. These oligonucleotides were phosphorylated, annealed, and ligated to pTXB3-VC_1(591) digested with *Hin*dIII. The subcloning strategy outlined above was also used to create a vector encoding a fusion of the VC_1 domain with the *Sce* VMA intein in the pTYB3 vector from NEB.

Construction of Xa-cys-IIC_2 vector – The vector pQE60-H6-IIC_2(847-1090) was a gift from Roger Sunahara. This vector was constructed by digesting pQE60-IIC_2-H6 (Whisnant et al., 1996) with *Bgl*II and partially digesting with *Nco*I. The resulting fragment was ligated to pQE60-H6 (Lee et

al., 1994) digested with the same enzymes. A factor Xa cleavage site and a S871C point mutation were added to pQE60-H6-IIC_2(847-1090) by Quikchange™ mutagenesis (Stratagene, La Jolla, CA). The sequence of the sense primer was 5'-GCACACGTGGCTGAACACATCGAAGGTAGGTGCCTGAAAAATGAGGAGC-3'. The sequence of the antisense primer was 5'-GCTCCTCATTTTTCAGGCACCTACCTTCGATGTGTTCAGCCACGTGTGC-3'. The resulting vector encodes amino-terminal hexahistidine tagged IIC_2(847-1090) with a S871C point mutation and the residues directly amino terminal of this mutation altered to encode a factor Xa recognition site (Ile-Glu-Gly-Arg).

Expression and Purification of VC1-intein-CBD – The bacterial strain BL21(DE3) was transformed with the pTXB3-H6-VC_1(591) expression vector. These colonies were used to inoculate 100 ml of Luria broth containing 50 μg/ml of ampicillin. This starter culture was shaken overnight at 37 ^oC. The starter culture was pelleted at 1000xg and resuspended in T7 medium. This was used to inoculate 12 L of T7 medium containing 50 μg/ml of ampicillin. These cultures were shaken at 30 ^oC until the OD_{600} reached 1.1. Protein expression was induced by the addition of 30 μM isopropyl-β-D-thiogalactopyranoside (IPTG), and cultures were incubated at room temperature for 5 hours. Cells were collected by centrifugation and stored at –80 ^oC.

All purification steps were performed at 4 ^oC. The cells were suspended in 800 ml of lysis buffer (50 mM Tris-HCl pH8, 120 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP, Pierce, Rockford, IL), and a mixture of protease inhibitors. The mixture of protease inhibitors used throughout the purification was 34 μg/ml each of L-tosylamido-2-phenylethyl chloromethyl ketone, 1-chloro-3-tosylamidio-7-amino-2-heptanone, and phenylmethylsulfonyl fluoride and 3 μg/ml each of lima bean trypsin inhibitor and leupeptin. Cells were lysed by addition of 0.40 mg/ml lysozyme and incubated for 40 min on ice. The lysate was treated with 8 μg/ml of DNase I and 1 mM $MgCl_2$ for an

additional 20 min. The cellular debris and membranes were removed by centrifugation at 100,000xg for 35 min at 4 ^{o}C. The supernatant was collected. The pellets were snap frozen in liquid nitrogen and lysed again by thawing in 400 ml of lysis buffer and treating with DNase I as described above. This lysate was clarified by ultracentrifugation and pooled with the first supernatant and applied to a 5-ml column of nitrilotriacetic acid-nickel ion-agarose (Ni^{2+}-NTA, Qiagen) equilibrated with lysis buffer.

The Ni^{2+}-NTA column was washed with 20 column volumes of buffer A (50 mM Tris-HCl pH 8, 500 mM NaCl, 1 mM TCEP, and protease inhibitors) followed by 10 column volumes of buffer B (50 mM Tris-HCl pH 8, 50 mM NaCl, 1 mM TCEP, 5 mM imidazole, and protease inhibitors). The column was eluted with 2.5-ml volumes of buffer B containing 150 mM imidazole. The peak fractions were pooled and diluted 10-fold in buffer C (50 mM Tris-HCl pH 8, 100 mM NaCl, 1 mM TCEP, and protease inhibitors). The VC_1-intein fusion was applied to a 20-ml chitin column (New England Biolabs) equilibrated with 10 volumes of buffer C. The chitin column was washed with 20 column volumes of buffer D (50 mM Tris-HCl pH 8, 250 mM NaCl, 1 mM TCEP, and protease inhibitors). The column was quickly equilibrated and incubated for 48 hours at 4 ^{o}C in buffer E (20 mM NaHEPES pH 8, 2 mM $MgCl_2$, 1 mM EDTA, 250 mM NaCl, 50 mM 2-mercaptoethanesulfonic acid (MESNA, Sigma), and protease inhibitors). The protein was eluted with 10-ml fractions in buffer F (20 mM NaHEPES pH 8, 2 mM $MgCl_2$, 1 mM EDTA, 250 mM NaCl, and 10 mM MESNA). This protein was concentrated in a Millipore Ultrafree-15 10 kDa cut-off spin concentrator and used in the ligation reaction.

Purification and cleavage of Xa-Cys-IIC_2 –The H6-Xa-IIC_2-S871C/K1014N was purified as described for the IIC_2-H6 domain (Hatley et al., 2002). H6-Xa-IIC_2-S871C/K1014N (5 mg) was cleaved with 25 μg of factor Xa protease (New England Biolabs, Beverly, MA) in buffer G (20 mM NaHEPES pH 8, 2 mM $CaCl_2$, and 100 mM NaCl) for 16 hours at 4 ^{o}C on ice. The reaction was

stopped by addition of an inhibitor, 1,5-dansyl-glu-gly-arg-chloromethyl ketone (2 μM)(Calbiochem, San Diego, CA). To remove the uncut and cleaved amino termi, the reaction mixture was applied to a 1-ml Talon™ column equilibrated with 10 ml of buffer H (20 mM NaHEPES pH 8, 2 mM $MgCl_2$, 50 mM NaCl), and the flow through was collected. The column was washed with 10 ml of buffer H. This wash was pooled with the flow through, diluted to 50 ml with buffer I (20 mM NaHEPES pH 8, 2 mM $MgCl_2$, 1 mM EDTA, and 2 mM DTT), and applied to a Source 15Q HR5/5 column (Amersham Pharmacia Biotech). The cleaved IIC_2 was eluted in a gradient of 0 to 300 mM NaCl over 25 column volumes. The peak fractions were pooled and concentrated using a Millipore Ultrafree-4 spin concentrator (10 kDa cut-off).

Adenylyl Cyclase Assays -- Adenylyl cyclase activity was measured as described by Smigel (Smigel, 1986). All assays were performed in a volume of 100 μl for 10 min at 30 ^oC. The final concentration of $MgCl_2$ was 10 mM. Assays contained 1 mM ATP, a limiting concentration of VC_1, and an excess of IIC_2 unless otherwise stated. The assays did not contain an ATP regenerating system. All specific activities reported are with respect to the concentration of the limiting domain.

Ligation of VC_1-MESNA with Cys-IIC_2 – The MESNA-modified VC_1 (70 μM final) from the chitin column was added immediately to Cys-IIC_2 at either 150 μM or 300 μM in the absence or presence of 100 μM forskolin and 100 μM 2'd3'-AMP:PPi. The mixture was incubated at 4 ^oC for 40 hours. Ligated proteins and the unligated VC_1 and IIC_2 (4 μg each) were resolved by SDS-PAGE on a 15% polyacrylamide gel and stained with Coomassie blue. The unligated VC_1 and IIC_2 (1 μg and 200 ng, respectively) and ligation products (1 μg each) were resolved by SDS-PAGE on a 15% polyacrylamide gel, transferred to nitrocellulose, and blotted with N-196 and X-400 antibodies. The N-196 antibody was raised against amino acids 375-388 of the VC_1 domain (Christiana Kleusse, unpublished results). The X-400 antibody was targeted to the C-terminus of the IIC_2 domain (John Krupinski, unpublished results).

Results

Purification of C-terminal thioester VC_1 – The fractions from the Ni^{2+}-NTA column are shown in Figure 3-4A. The Ni^{2+}-NTA pool contained 42% of the activity of VC_1 and resulted in a 28-fold purification of the protein (Table 3-1). The fractions from the chitin column are shown in Figure 3-4B. Both the *Mxe* GyrA intein-chitin binding domain fusion cleaved from the VC_1 domain and the VC_1 domain have a molecular weight of 27 kDa and co-migrate on SDS-PAGE polyacrylamide gels. The chitin resin was stripped with SDS. This showed that the cleavage with MESNA was only 50% efficient, as illustrated by the amounts of full-length fusion protein and the cleaved *Mxe* GyrA intein-chitin binding domain remaining on the chitin column. A major loss of activity and protein was experienced during chitin-column chromatography (Table 3-1). There was no difference in the apparent affinity of the VC_1-MESNA from the chitin pool and purified H6-VC_1(591)FLAG for the IIC2 domain; however, the apparent V_{max} of the chitin pool was 3-fold less than that of H6-VC_1(591)FLAG (Figure 3-5A).

Different thiol reagents were tested for deleterious effects on adenylyl cyclase activity. DTT (50 mM) had no effect on adenylyl cyclase activity or the apparent affinity of the VC_1 and IIC_2 domains (Figure 3-5B). In contrast, increasing concentrations of MESNA decreased the affinity of purified VC_1 for IIC_2 (Figure 3-5C). Purified H6-VC_1(591)FLAG domain (50 μM) was incubated with 50 mM MESNA, 50 mM DTT, or no treatment, and aliquots were removed at various times and assayed for adenylyl cyclase activity. Even after 50 hours at 4 ^{o}C, neither MESNA nor DTT treated VC_1 exhibited a difference from untreatment samples when reconstituted with the IIC_2 domain and assayed for adenylyl cyclase activity (Figure 3-5D). MESNA, DTT, and thiophenol were equally efficacious in cleaving the VC_1 from the intein fusion protein. Standard reducing agents, 2-mercaptoethanol and DTT, cleave the intein fusion; therefore, TCEP was used as a reducing agent in

the purification buffers. Incubation of VC_1 in 1 mM TCEP for 3 hours at 4 ^oC had no effect on adenylyl cyclase activity.

Purification of IIC_2 with N-terminal cysteine - The H6-Xa-IIC_2-S871C was readily purified to homogeneity (Table 3-2). Cleavage with factor Xa protease and subsequent passage through Talon™ and Mono Q columns yielded IIC_2 protein with an amino-terminal cysteine for the ligation reaction, Cys-IIC_2 (Figure 3-7B). The enzymatic activity of both the H6-Xa-IIC_2-S871C and the Xa-digested Cys-IIC_2 were indistinguishable from wild type IIC_2-H6 (Figure 3-6).

Ligation of VC_1 and IIC_2 - The method used to ligate the VC_1 and IIC_2 domains using Intein-mediated protein ligation is illustrated in Figure 3-6A. All ligation reactions contained 70 μM concentration of freshly isolated MESNA-modified VC_1. Ligation 1 and ligation 2 contained 150 μM and 300 μM respectively of Cys-IIC_2. Ligation 3 contained 150 μM Cys-IIC_2 and 100 μM each of forskolin and 2'd3'-AMP:PPi. These three ligation conditions resulted in the same amount of ligated VC_1 and IIC_2 after a 40-hour incubation at 4 ^oC (Figure 3-7B). The Coomassie blue-stained gel showed that the ligation reactions had a band slightly above the 48-kDa molecular weight marker that was not present in either the VC_1-MESNA chitin pool or the factor Xa cleaved IIC_2 that was consistent with the molecular weight expected for the ligated product, 51 kDa. This band was immunoreactive with both anti-VC_1 and anti-IIC_2 antibodies (Figure 3-7C and D), indicating ligation of the peptides. The products from the ligation reaction were applied to a Talon™ column. This removed any unligated Cys-IIC_2. Next, this Talon™ eluate was applied to a forskolin-agarose column (Sigma-Aldrich, St. Louis, MO). However, neither the ligated VC_1-IIC_2 nor the free VC_1 domain were retained by the column.

Linkers were added to the VC_1-intein fusion between the C-terminus of the VC_1 and the N-terminus of the intein. The first linker, L3 (AAAGGMPPAAAGGM), was used by Wei-Jen Tang in the construction of a fused IC_1-IIC_2 peptide (Tang and Gilman, 1995). The other linker was

KLGSGTSSGTSSTGSAG. This linker was used by Stephen Michnick in his dihydrofolate reductase complementation assay (Pelletier et al., 1999). The addition of these linkers increased neither the yield of the VC_1 from the chitin column nor the efficiency of the ligation. The VC_1 domain was also fused to the *Sce* VMA intein using the pTYB3 expression plasmid from NEB. The VC_1-*Sce* VMA-CBD fusion was not expressed as a soluble protein in bacteria.

Discussion

The major limitations of this approach are the poor yields observed during both the cleavage of the VC_1 domain from the chitin column and the subsequent ligation reaction. MESNA, DTT, and thiophenol all cleaved the VC_1-intein fusion equally well. Neither MESNA nor DTT treatment alone denatured (Figure 3-5C). It has been reported that MESNA-modified peptides are better substrates for the ligation reaction (Evans et al., 1998; Evans et al., 1999a); therefore, MESNA was used as the cleaving reagent. MESNA reacts covalently with the C-terminus of the VC_1 domain and provides an excellent leaving group when Cys-IIC_2 performs the nucleophilic substitution addition. The VC_1-intein fusion expressed well in *E. coli,* but the cleavage of VC_1 from the chitin column was inefficient (50%), and the eluted product had reduced enzymatic activity compared to the traditionally purified H6-VC_1(591)FLAG. MESNA did not have a deleterious effect on the activity of cytosolic domains of adenylyl cyclase. Direct exposure of purified H6-VC_1(591)FLAG to either MESNA or DTT for the duration of time required for cleavage from the chitin column had no effect on the enzymatic activity when mixed with IIC_2.

The inefficiency of the cleavage reaction could result from limited access of the thiol reagent to the substrate. As well, the ligation reaction could be inhibited in a similar fashion. The MESNA-modified C-terminus of VC_1 and the N-terminal cysteine of IIC_2 could be folded such that the N-terminal cysteine of IIC_2 is never exposed to the MESNA-modified C-terminus of VC_1. In the

ligation reaction, the VC_1 and IIC_2 domains are present at concentrations (70 and 150 μM) well above the apparent affinity of the proteins for each other in the absence of activators (5 μM). This should drive formation of the VC_1:IIC_2 heterodimer. The formation of the heterodimer could constrain the termini of the domains, leading to inefficient ligation. To test this hypothesis, linkers were added between VC_1 and the *Mxe* GyrA intein to allow more flexibility in the termini. The L3 linker previously used in a IC_1-IIC_2 fusion yielded no greater cleavage or ligation. Another linker was chosen based on its ability to be expressed successfully in *E. coli* without proteolysis. This linker also failed to increase the amount of cleavage or ligation. Inteins exhibit sequence specificity N-terminal to splice junctions that promote or inhibit splicing. Therefore, a VC_1 fusion with the *Sce* VMA intein was constructed. This VC_1-intein fusion protein was expressed at a low level and was largely insoluble; this limited its usefulness. Other vectors for expression of proteins fused to other inteins are not commercially available.

The yield of the ligated VC_1 and IIC_2 was insufficient to be useful for biochemical characterization or structural determination of the adenylyl cyclase catalytic core in the absence of $G_{s\alpha}$. It was not possible to purify the ligated VC_1-IIC_2 from the unligated domains due to the low yield. The N-terminal sequence of the *Mxe* GyrA intein was thought to be the major determinant of both the efficiency of cleavage of the VC_1-intein fusion protein and the ligation of the VC_1 and IIC_2 domains. It was not deemed prudent to screen linkers by expressing protein and testing cleavage ligation efficiency. In order to continue this line of investigation, a method to rapidly screen linkers that would facilitate the functional interaction of the VC_1 and IIC_2 domains might be developed. Alternative methods for the production of a linked VC_1-IIC_2 peptide must also be explored.

Table 3-1.

Summary of Purificiation of H6-VC_1-MESNA from E. coli.

Pool	mg/ml	Volume (ml)	Yield (mg)	Specific Activity (μmol/min/mg)	Total Activity μmol/min	Recovery %	Purification -fold
Clarified lysate	4.4	1200	5700	0.4	2500	100	----
Ni-NTA	5.5	17.5	96	11	1056	42	28
Chitin	3.9	0.45	1.8	1	1.8	0.1	2.5

Pools from the lysate, Ni^{2+}-NTA column, and chitin column were assayed for $G_{s\alpha}$-GTPγS- and forskolin-stimulated adenylyl cyclase activity. The VC_1 concentration in the assay for each pool was 2 nM. Fractions were assayed in a total volume of 100 μl in the presence of 2 μM IIC_2, 100 μM forskolin, 400 nM $G_{s\alpha}$-GTPγS, and 1 mM [^{32}P]ATP for 10 min at 30 ^{o}C.

Table 3-2.

Summary of Purification of H6-Xa-IIC_2-S871C from E. coli.

Pool	mg/ml	Volume (ml)	Yield (mg)	Specific Activity (μmol/min/mg)	Total Activity μmol/min	Recovery %	Purification -fold
Clarified lysate	5	250	1260	8.9	11200	100	----
Ni^{2+}-NTA	16	5	80	74	5900	53	8
Source 15Q	7.4	7.8	58	97	5600	50	11

Pools from the lysate, Ni^{2+}-NTA column, and Source 15Q column were assayed for $G_{s\alpha}$-GTPγS- and forskolin-stimulated adenylyl cyclase activity. The IIC_2 concentration in the assay for each pool was 2 nM. Fractions were assayed in a total volume of 100 μl in the presence of 2 μM VC_1, 100 μM forskolin, 400 nM $G_{s\alpha}$-GTPγS, and 1 mM [^{32}P]ATP for 10 min at 30 ^{o}C.

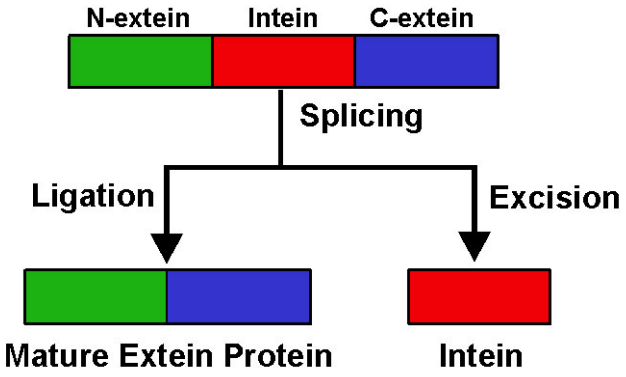


Figure 3-1. Overview of intein-mediated protein splicing. An internal protein sequence or intein is excised from a protein. The flanking external protein sequences or exteins are ligated through a series of four nucleophilic substitution reactions. The resulting ligated exteins are linked with a native peptide bond.

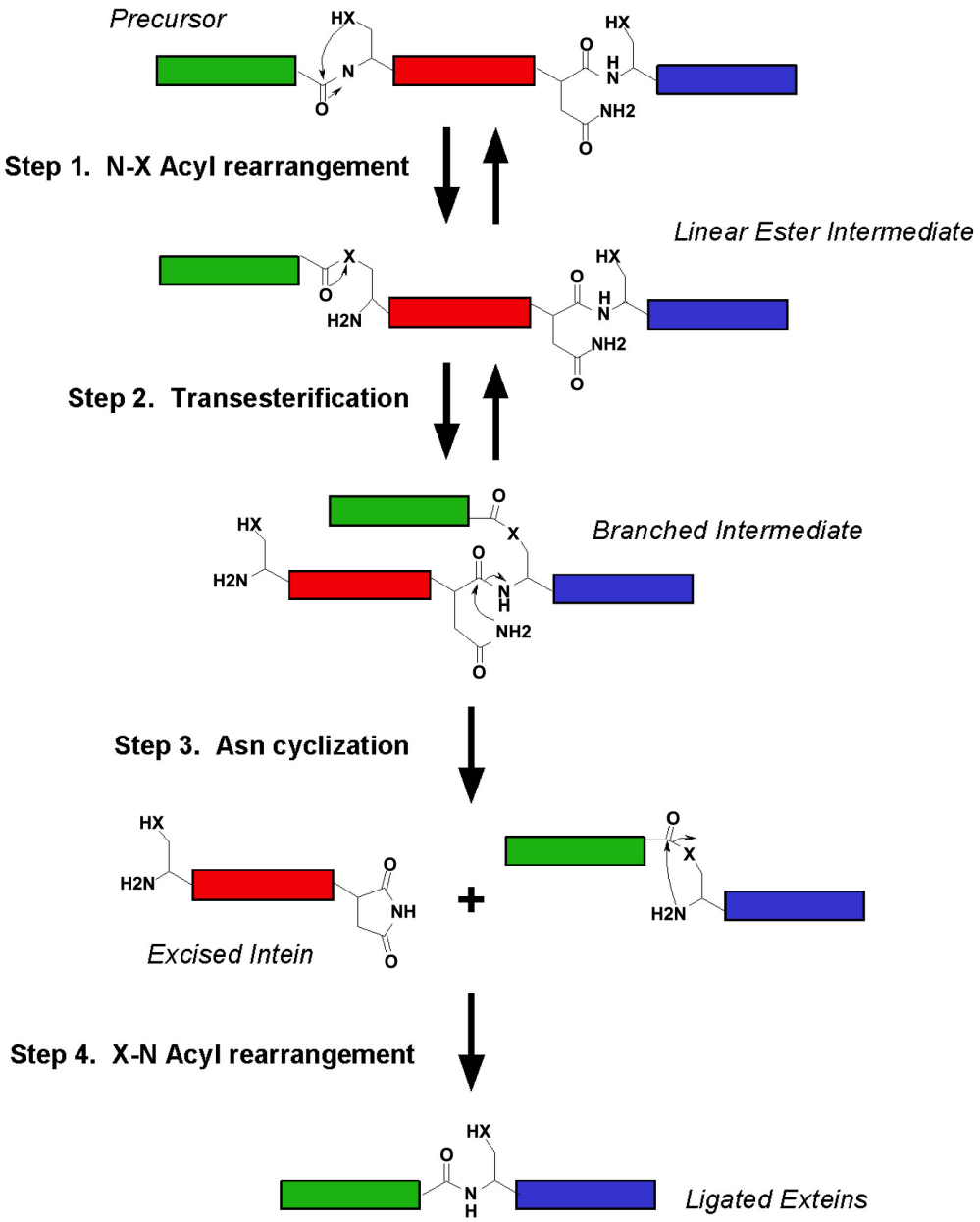


Figure 3-2. Mechanism of Protein Splicing. The two exteins are ligated with a native peptde bond and the intervening intein is excised through a series of four nucleophilc subsitution reactions. The X represents either the sulfur from a cysteine side chain or an oxygen from a serine of threonine side chain. (This figure was adapted from (Perler, 1998)).

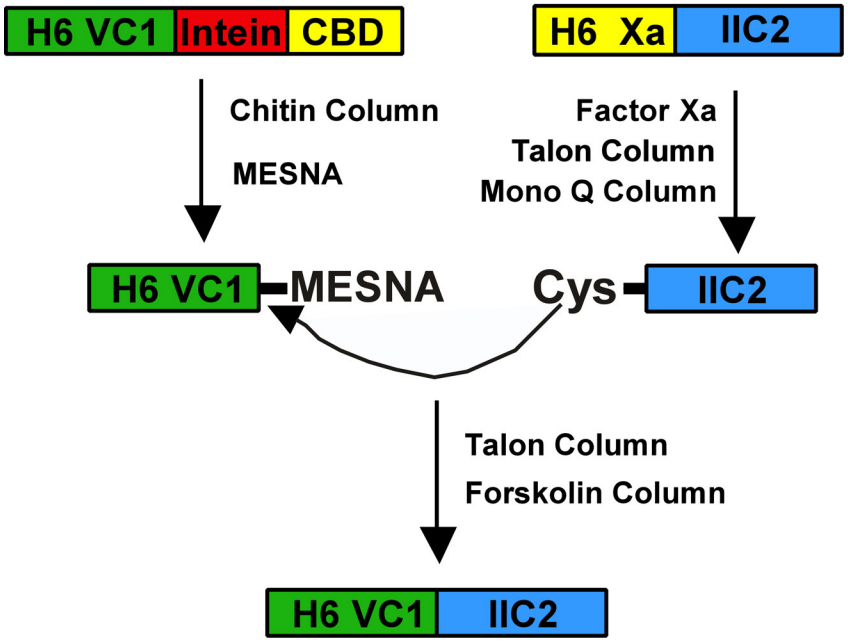


Figure 3-3. Schematic of the purification and ligation of VC_1 and IIC_2. The VC_1-intein-chitin binding domain fusion protein was purified by binding to chitin resin. The VC_1 protein was liberated by overnight incubation with 2-mercaptoethanesulfonic acid, MESNA. MESNA reacts covalently with the C-terminus of the VC_1 protein providing an excellent leaving group for the forthcoming reaction. An N-terminal cysteine residue was generated on the IIC_2 protein by proteolysis with factor Xa. The Cys-IIC_2 was purified by column chromatography. The Cys-IIC_2 protein was ligated to the VC_1 protein by the nucleophilic substitution of the N-terminal cysteine of the IIC_2 protein on the MESNA-modified C-terminus of the VC_1 protein resulting in a VC_1-IIC_2 fusion protein linked by a peptide bond.

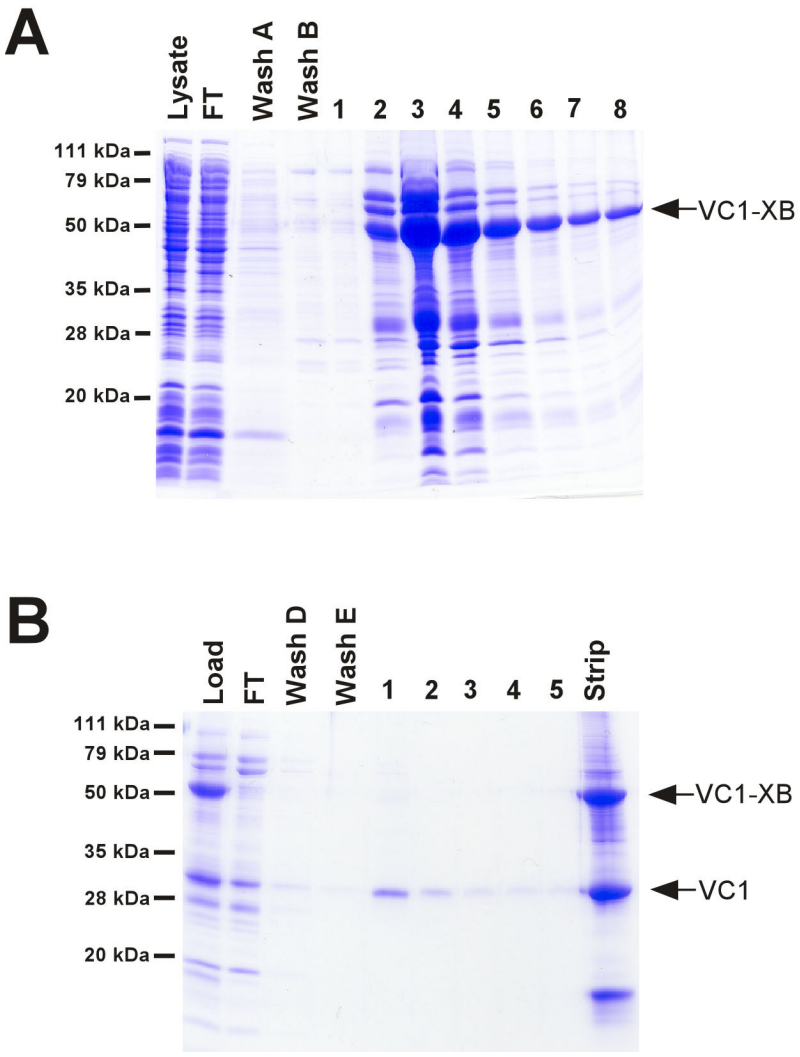


Figure 3-4. Purification of VC_1 from the intein fusion protein. A. The VC_1-*Mxe* GyrA-CBD VC_1-intein fusion protein (VC_1-XB) was expressed in *E. coli.* Aliquots of the lysate (5 µg) and fractions from the Ni^{2+}-NTA column (5 µl) were resolved by SDS-PAGE on a 15% polyacrylamide gel and stained with Coomassie blue. **B.** The eluate from the Ni^{2+}-NTA column was applied to a chitin column (NEB). Aliquots of the fractions and SDS-stripped chitin resin (15 µl) were resolved by SDS-PAGE on a 15% polyacrylamide gel and stained with Coomassie blue. The chitin resin (300 µl) was stripped with 100 µl of 1% SDS and 15 µl of this solution was resolved by SDS-PAGE. Molecular weight markers are indicated.

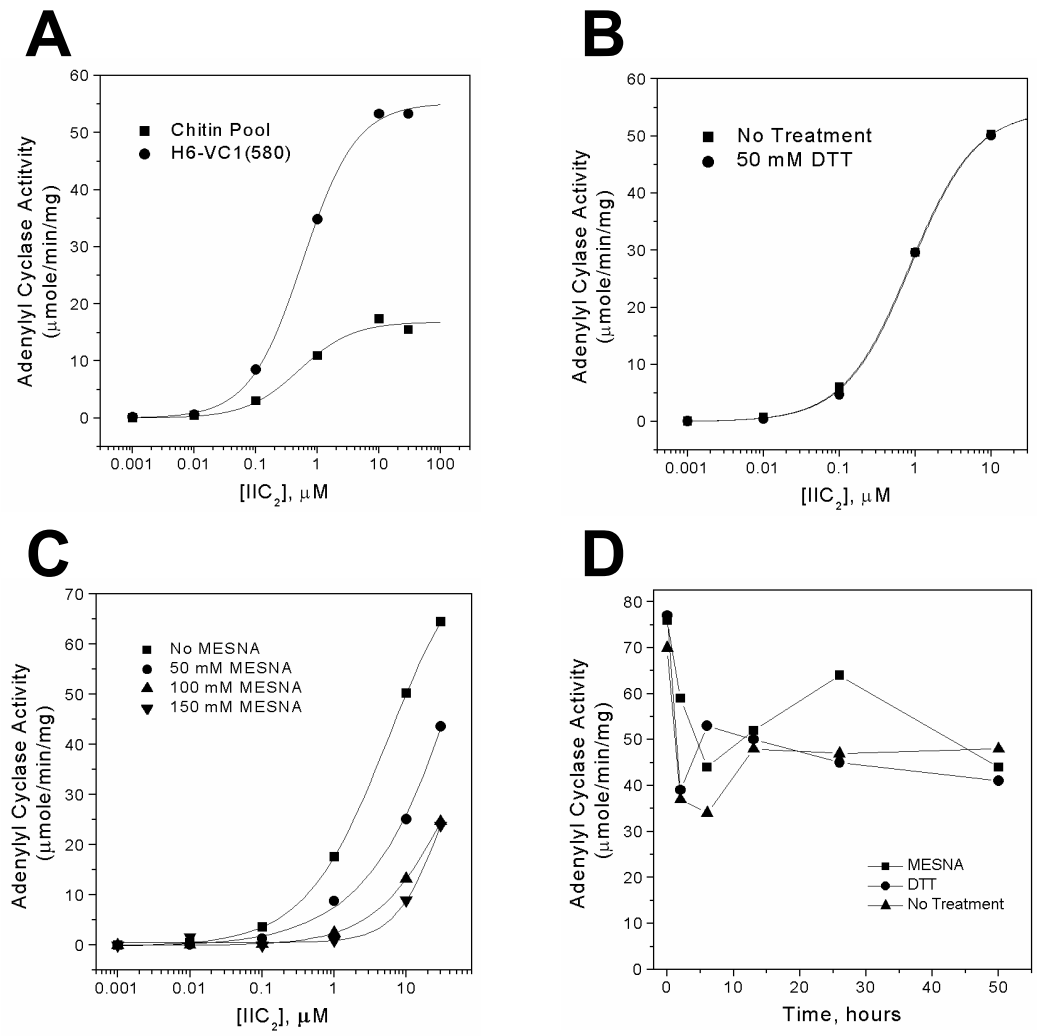


Figure 3-5. Characterization of VC_1 intein fusion protein. A. Adenylyl cyclase activity of MESNA-modified VC_1 from the chitin column compared to H6-VC_1(591)FLAG. The MESNA-cleaved VC_1 pooled from the chitin column and purified H6-VC_1(591)FLAG were limited at 2 nM, and adenylyl cyclase activity was assayed with increasing concentrations of IIC_2 in the presence of 100 μM forskolin. **B.** The effect of DTT on the activity of VC_1 and IIC_2. Adenylyl cyclase activity was assayed in the presence or absence of 50 mM DTT. **C.** Effect of increasing MESNA concentrations on adenylyl cyclase activity of VC_1 and IIC_2. **D.** Time course of MESNA and DTT treatment of VC_1 and IIC_2. Purified H6-VC_1(591)FLAG (50 μM) was incubated at 4 ^{o}C in the absence or presence of either 50 mM DTT or MESNA. Aliquots were taken at times indicated. Adenylyl cyclase activity measured at a limiting concentration of the treated VC_1 domain (2 nM) and an excess of the IIC_2 domain (2 μM) in the presence of 100 μM forskolin and 400 nM $G_{s\alpha}$-GTPγS.

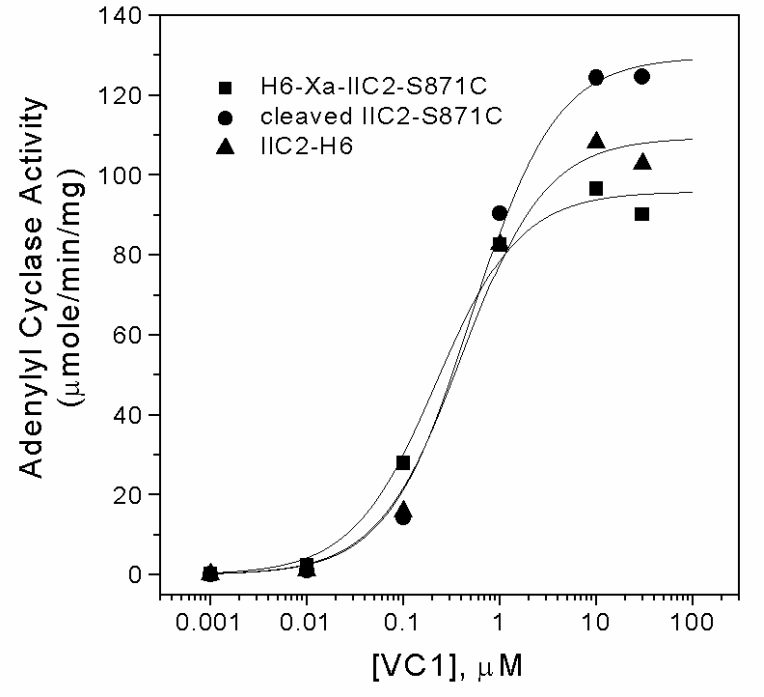


Figure 3-6. Activity of Cys-IIC_2 after cleavage with factor Xa. The reconstituted adenylyl cyclase activity of purified Cys-IIC_2 after proteolysis with factor Xa was compared with the unproteolyzed H6-Xa-IIC_2-S871C and wild type IIC_2. Adenylyl cyclase activity was assayed with limiting IIC_2 (2 nM) and increasing concentrations of H6-VC_1(591)FLAG in the presence of 100 μM forskolin.

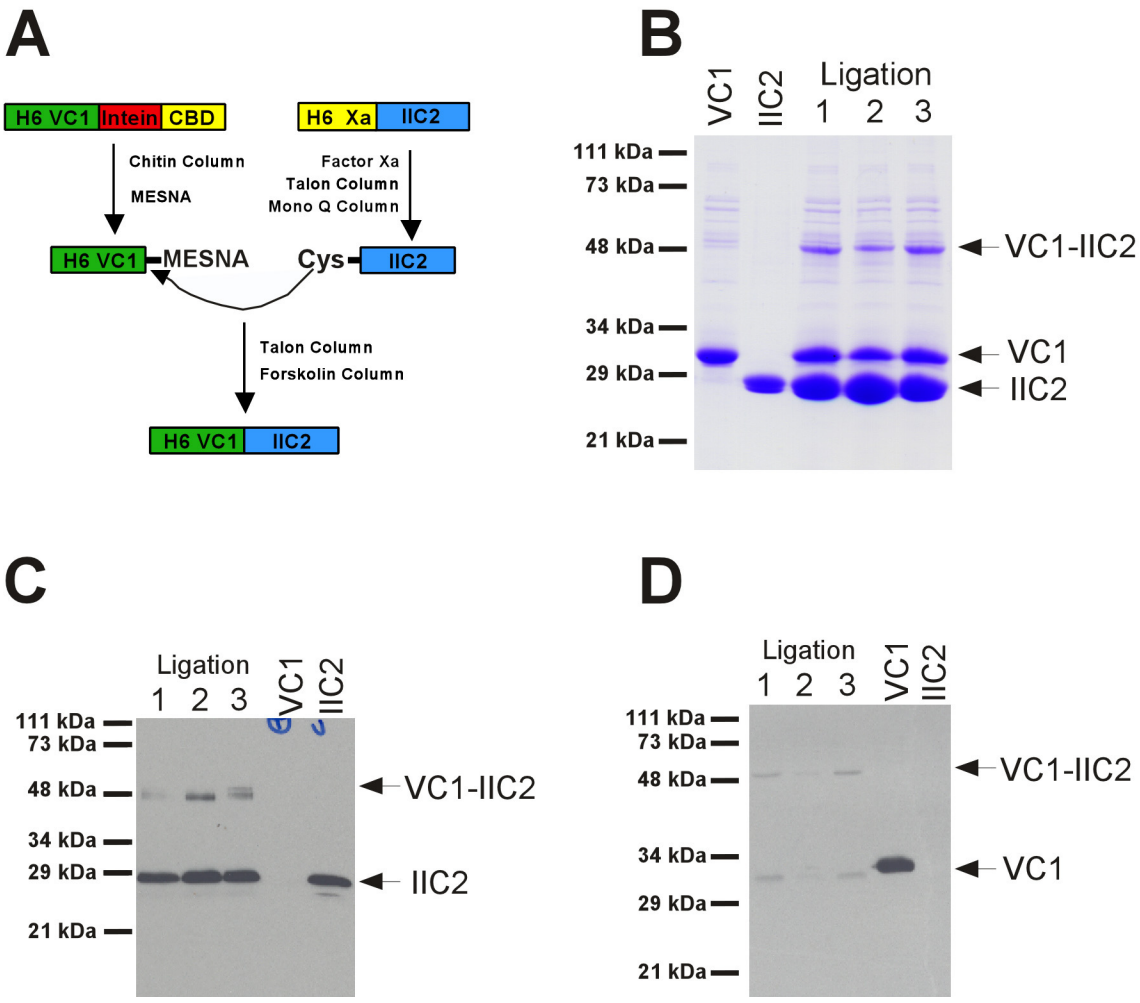


Figure 3-7. Intein-mediated protein ligation of VC_1 and IIC_2. A. Schematic of the purification and ligation of VC_1 and IIC_2. **B.** MESNA modified VC_1 from the chitin column (3µg), factor Xa cleaved IIC_2 (3µg), and the VC_1:IIC_2 ligation reaction (4µg) were resolved by SDS-PAGE on 15% polyacrylamide gel and stained with Coomassie blue. **C,D.** Components of the VC_1:IIC_2 ligation reaction were resolved by SDS-PAGE on a 15% polyacrylamide gel and blotted with X400 anti-IIC_2 antibody (**C**) and N-196 anti-VC_1 antibody (**D**).

Chapter 4 – Complementation Screen for Linked Adenylyl Cyclase Cytosolic Domains

Introduction

The IC_{1a} and IIC_{2a} domains have been expressed as a chimeric fusion protein in *E. coli.* This linked IC_1-IIC_2 protein displays characteristics of membrane-bound adenylyl cyclase with respect to regulation by $G_{s\alpha}$, forskolin, and so-called P-site inhibitors such as 2'd3'-AMP (Tang and Gilman, 1995; Dessauer and Gilman, 1996; Dessauer and Gilman, 1996). However, it was not possible to purify large amounts of the protein limiting its usefulness for structural determination. The C_{1a} and C_2 domains were then expressed separately as recombinant proteins. When mixed together, the domains associate to reconstitute an enzyme with the characteristics of the membrane-bound adenylyl cyclase (Whisnant et al., 1996; Yan et al., 1996; Sunahara et al., 1997a; Yan et al., 1996; Sunahara et al., 1997a). The low apparent affinity of the two purified proteins for each other (>5 μM), is the major obstacle for purification of a low activity complex (Sunahara et al., 1997a). While identification of the constitutively active K1014N mutant and purification of the complex of VC_1 and IIC_2-K1014N containing forskolin and 2'd3'-AMP:PPi permitted crystallization of a protein complex, these crystal were not of a quality sufficient to permit determination of protein structure (see Chapter 2).

As described in Chapter 3, I attempted to exploit the greater solubility and expression level of the individual VC_1 and IIC_2 domains to link post-translationally the VC_1 and IIC_2 domains using intein-mediated protein splicing. Unfortunately, these attempts were plagued with technical difficulties. Incomplete cleavage of the VC_1-*Mxe* GyrA fusion protein and low ligation efficiency of the VC_1 and IIC_2 domains using intein-mediated protein ligation were hypothesized to result from the lack of a linker between the two domains. The addition of a linker might allow better access of both the thiol reagent MESNA for cleavage of VC_1 from the intein fusion and Cys-IIC_2 for the ligation reaction. The laborious nature of IPL precludes screening linkers using this technique.

The amount of IC_1-IIC_2 chimeric fusion protein purified from *E. coli* was insufficient for structural determination. However, large quantities of the IIC_2 domain can be purified from *E. coli* (~20 mg/l culture) when expressed as an individual protein (Hatley et al., 2002). In contrast, the IC_1 domain is much less soluble and less active than the VC_1 domain when expressed as an individual protein. A fusion protein of the VC_1 and IIC_2 domains could circumvent the problems associated with the IC_1-IIC_2 fusion protein. A major obstacle in the production of a functional VC_1-IIC_2 fusion protein is the selection of an appropriate linker peptide. A short peptide linker separating the VC_1 and IIC_2 domains could facilitate the functional interaction of the proteins. A phenotypic complementation assay using adenylyl cyclase deficient *Escherichia coli* has been developed as a genetic screening method for selecting linkers that allow productive interaction of the adenylyl cyclase domains.

In *E. coli*, cAMP serves as a coactivator of transcription of many inducible genes. The cAMP binds to a transcriptional factor called cAMP receptor protein or CRP. CRP is a homodimer with each subunit containing two domains: one for DNA binding and one for interaction with cAMP. Binding of cAMP elicits a conformational change in CRP that increases its affinity for DNA (Botsford and Harman, 1992). The interaction of cAMP with CRP controls the transcription of many catabolic enzymes, including the enzymes required for the fermentation of maltose. Therefore, strains deficient in adenylyl cyclase activity form white colonies on MacConkey agar plates and fail to grow on minimal medium containing maltose as the sole carbon source. Complementation of such deficient strains with adenylyl cyclase suppresses this phenotype and permits growth on minimal medium and growth of red colonies on MacConkey agar plates (Roy and Danchin, 1981; Roy and Danchin, 1982; Roy and Danchin, 1982; Roy and Danchin, 1982). Adenylyl cyclase deficient *E. coli* have been used to identify the genes encoding adenylyl cyclase in several bacterial species (Roy and Danchin, 1982; Beuve et al., 1990; Beuve et al., 1990) and to develop a bacterial two-hybrid screen

(Karimova et al., 1998; Karimova et al., 2000). This Chapter will explore the use of these bacteria to develop a complementation screen for linkers that allow functional interaction of the VC_1 and the IIC_2 domains.

Experimental Methods

Materials and general methods – Molecular biological procedures were performed as outlined by Sambrook and Russell (Sambrook and Russell, 2001). Quikchange™ mutagenesis (Stratagene, La Jolla, CA) was performed using *Pfu* polymerase and the manufacturer's recommended conditions. All plasmids were prepared using QIAGEN products and sequenced by the McDermott Center DNA sequencing core by Jay Hunter.

Bacterial strains and growth conditions – The strains used were TP2000 (*F* $^-$, *xyl, cya*Δ*, argH1, lac*Δ*X74* (Roy and Danchin, 1981) and TP610A (*F* $^-$*, thi-1, leuB6, pro, lacY1, tonA21, supE44, hsdR, hsdM, recBC, lop-11, lig*$^+$*, cya-610*) (Beuve et al., 1990). The MacConkey agar plates were prepared with Bacto MacConkey agar base (Difco, Detroit, MI) and contained 0.4% maltose, 50 μg/ml each of carbenicillin and kanamycin, and either 0, 10, or 100 μM of IPTG. The synthetic medium M63 agar plates contained 0.4% maltose, 1 mM $MgSO_4$, 10 μg/ml of thiamine, 50 μg/ml each of carbenicillin and kanamycin, and either 0, 10 or 100 μM of IPTG. M63 agar plates for TP2000 cells contained 100 μg/ml L-arginine. M63 agar plates for TP610A cells contained 100 μg/ml of each threonine, leucine, and proline.

Construction of pACYC-T5-IIC$_2$-K1014N – The pACYC-T5(RS1) vector was a gift from Roger Sunahara. This vector contains the T5 promoter, lac operators, transcriptional terminators, and multiple cloning sites from pQE60 and the *kan* gene for kanamycin resistance and the p15A origin of replication from pACYC177. The p15A origin of replication is compatible with plasmids containing the Col E1 origin of replication. The pQE60-IIC_2-K1014N plasmid was digested with *Eco*RI and

*Blp*I and ligated to pACYC-T5 digested with the same enzymes. The resulting plasmid was designated pACYC-T5-IIC_2-K1014N.

Construction of linked VC_1-IIC_2 – A *Kpn*I site was added immediately 5' to the bases encoding Arg870 in pQE60-H6-IIC_2(847-1090) by Quikchange™ mutagenesis. The sequence of the sense primer was 5'-CACTTCCTGGGTACCAGCCTGAAAAATGAG-3'. The sequence of the antisense primer was 5'-CTCATTTTTCAGGCTGGTACCCAGGAAGTG-3'. The base pairs of IIC_2 encoding Arg870 and Ser871 were changed by Quikchange™ mutagenesis to encode a *Bgl*II restriction site without changing the encoded amino acid residues. The sequence of the sense primer was 5'-CACTTCCTGGGTACCAGATCTCTGAAAAATGAGGAG-3'. The sequence of the antisense primer was 5'-CTCCTCATTTTTCAGAGATCTGGTACCCAGGAAGTG-3'. The resulting vector was designated pQE60-IIC_2-KpnI/BglII. A *Kpn*I restriction site encoding a glycine and threonine residue was added 3' to the bases encoding Lys580 of VC_1 in pQE60-H6-VC_1(364-580) by Quikchange™ mutagenesis. The sequence of the sense primer was 5'-GGCCATGATCGCCAAGGGTACCAAGCTTAATTAGCTGAGC-3'. The sequence of the antisense primer was 5'-GCTCAGCTAATTAAGCTTGGTACCCTTGGCGATCATGGCC-3'. The resulting plasmid was designated pQE60-H6-VC_1(364-580)-KpnI. The plasmid pQE60-IIC_2-KpnI/BglII was digested with *Kpn*I and *Nhe*I and ligated to pQE60-H6-VC_1(364-580)-KpnI which had been digested with the same enzymes. The resulting plasmid, pQE60-H6-VC_1-IIC_2, encodes the protein sequence MHHHHHHA followed in order by amino acids 364 to 580 of canine type V adenylyl cyclase, a glycine threonine linker (from addition of *Kpn*I site), and amino acids 870-1090 of rat type II adenylyl cyclase. Complementary oligonucleotides designed with a 5' *Kpn*I and 3' *Bgl*II cohesive end were to be ligated to pQE60-H6-VC_1-IIC_2 digested with *Kpn*I and *Bgl*II to create different peptide linkers between VC_1 and IIC_2. However, the *Bgl*II restriction endonuclease requires bases flanking the recognition site for efficient digestion. Therefore, the bases encoding a glycine

residue were inserted between the *Kpn*I and *Bgl*II restriction sites by Quikchange™ mutagenesis. The sequence of the sense primer was 5'-TCGCCAAGGGTACCGGCAGATCTCTGAAAAATGAG-3'. The sequence of the antisense primer was 5'-CTCATTTTTCAGAGATCTGCCGGTACCCTTGGCAG-3'. The resulting plasmid was pQE60-H6-VC_1-Gly-IIC_2 encodes the protein sequence MHHHHHHA followed in order by amino acids 364 to 580 of canine type V adenylyl cyclase, a glycine threonine glycine linker, and amino acids 870-1090 of rat type II adenylyl cyclase..

The pQE80L plasmid from QIAGEN was modified to contain a *Nco*I site at the 5' cloning site by Quikchange™ mutagenesis. The sequence of the sense oligo was 5'-CACCATCACCATCACGCCATGGGATCCGCATGCG-3'. The sequence of the antisense oligo was 5'-CGCATGCGGATCCCATGGCGTGATGGTGATGGTG-3'. Another *Nco*I site in the pQE80L at base pairs 861 was eliminated by Quikchange™. The sequence of the sense primer was 5'- CCCCCGTTTTCAGTATGGGCAAATATT-3'. The sequence of the antisense primer was 5'-AATATTTGCCCATACTGAAAACGGGGG-3'. The resulting plasmid, pQE80-Nco, allowed convenient shuttling from pQE60 as *Nco*I-*Hin*dIII fragments. Genes subcloned into the *Nco*I site will encode a protein with the sequence MRGSHHHHHHA followed by the protein of interest. The plasmid pQE60-H6-VC_1(364-580) was digested with *Nco*I and *Hin*dIII and ligated to pQE80-Nco digested with the same enzymes. The resulting plasmid was designated pQE80-H6-VC1(364-580).

Complementation Screen – The compatible plasmids encoding the VC_1 and IIC_2 domains respectively were co-transformed in chemically competent TP2000 or TP610A *E. coli,* plated on LB agar plates containing 50 μg/ml each of carbenicillin and kanamycin, and incubated at 37 ^{o}C overnight. Single colonies were picked and suspended in 10 μl of M63 salts. The suspension (1 μl) was spotted onto MacConkey and M63 minimal plates. Forskolin (5 μl of a 10 mM solution) was spotted onto the plates prior to the addition of the bacteria. Samples that were spotted onto forskolin

were suspended in M63 salts containing 0.33 μg/ml polymixin B nonapeptide. The plates were incubated at 30 ^{o}C. Photographs were obtained using a Kodak DC3400 digital camera.

Results

Different combinations of the cytosolic domains of adenylyl cyclase and of $G_{s\alpha}$ were transformed into the TP2000 strain of *E. coli* (Figure 4-1A). The bacteria were spotted onto both MacConkey and M63 minimal agar plates in the absence and presence of forskolin following the template shown in Figure 4-1B. Bacteria spotted onto forskolin were treated with polymyxin B nonapeptide to permeabilize the outer membrane and thereby increase the uptake of forskolin (Vaara, 1992). Bacteria containing the wild type VC_1 and IIC_2 domains failed to ferment maltose in the absence of forskolin, illustrated by the inability to form red colonies on MacConkey agar (Fig 4-1C) or to grow on minimal medium (Fig 4-1D). In contrast, bacteria containing the wild type VC_1 domain and the IIC_2-K1014N mutant (see Chapter 2) with increased basal affinity and activity fermented maltose shown by the growth of red colonies on MacConkey agar and growth on minimal medium in the absence of forskolin. The increased basal activity of the IIC_2-K1014N mutant protein in combination with the VC_1 protein produced sufficient cAMP to complement the metabolic deficiency in the TP2000 strain of *E. coli.* The co-expression of $G_{s\alpha}$ in bacteria expressing VC_1 and IIC_2 proteins allowed the fermentation of maltose shown by the red growth on MacConkey agar and growth on minimal medium. The bacteria containing IC_1, $G_{s\alpha}$, and either wild type IIC_2 or the mutant IIC_2-K1014N formed red colonies on MacConkey agar; however, only the bacteria containing the IIC_2-K1014N mutant protein grew weakly on minimal medium. The growth patterns in the presence and absence of forskolin were the same with one exception. In the presence of forskolin, bacteria expressing VC_1 and IIC_2 formed red colonies on MacConkey agar but failed to grow on minimal medium. These results suggest that the increased cAMP resulting from enhanced adenylyl cyclase

activity either from the presence of the IIC_2-K1014N mutant protein or $G_{s\alpha}$ was sufficient to complement the catabolic defect in the TP2000 strain of *E. coli.*

When repeated, this assay failed to yield consistent results. Variable results occurred when plating multiple clones containing the same adenylyl cyclase domains. The above experiment was performed in the absence of Lac repressor. This allowed unregulated, uninducible expression from the T5 promoter of the pQE60 and pACYC-T5. The phage T5 promoter is recognized by the endogenous *E. coli* RNA polymerase. Transcription from this promoter is regulated by two *lac* operators that bind Lac repressor and block transcription. Without Lac repressor, transcription is constitutive and unregulated by IPTG. The VC_1 domains were subcloned into a pQE80 vector from QIAGEN that constitutively expresses Lac repressor *in cis* to VC_1. These new vectors were assayed for complementation in the TP2000 strain and another strain as well TP610A (Beuve et al., 1990)in the absence or presence of IPTG (Figure 4-2). The expression of Lac repressor led to more consistent results. The TP2000 and TP610A bacteria exhibited different phenotypes when co-expressed with various combinations of the C_1 and C_2 domains. In the absence of IPTG, the catabolic defect in the TP2000 strain was complemented by the addition of VC_1 and IIC_2-K1014N but not by the VC_1 and IIC_2. However, in the presence of 10 μM IPTG, the catabolic defect of the TP2000 strain was complemented by the combinations of VC_1 and IIC_2-K1014N and the combinations of VC_1 and IIC_2 with or without the addition of $G_{s\alpha}$. At 100 μM IPTG, the metabolic defect in the TP610A strain was complemented by the combination of VC_1 and IIC_2-K1014N or the combination of VC_1, IIC_2 and $G_{s\alpha}$; however, the combination of VC_1 and IIC_2 failed to complement. All combinations of proteins that complemented the metabolic defects in the TP2000 and TP610A *E. coli* illustrated by red growth on MacConkey agar resulted in growth on minimal medium.

This phenotypic difference in adenylyl cyclase-deficient bacteria was exploited to select a functional, linked VC_1-IIC_2 fusion protein. The VC_1 and IIC_2 domains were subcloned to yield a

VC_1-IIC_2 fusion protein separated only by the glycine and threonine residues from the *Kpn*I restriction site. The parent VC_1-IIC_2 vector without an intervening linker was transformed into the TP2000 and TP610A adenylyl cyclase-deficient *E. coli.* Unexpectedly, the cAMP produced from this simple fusion complemented the catabolic defect of these bacteria (Figure 4-2C). In the TP2000 *E. coli* strain, the basal, uninduced expression of the VC_1-IIC_2 fusion protein produced sufficient cAMP to complement the adenylyl cyclase-deficient bacterium. However, TP2000 bacteria expressing the VC_1-IIC_2 fusion protein formed white colonies at 10 μM IPTG and failed to grow at 100 μM IPTG. The linked VC_1-IIC_2 protein conferred a different phenotype when expressed in the TP610A strain. The bacterium grew as white colonies with no IPTG, red colonies with 10 μM IPTG, and white less dense colonies with 100 μM IPTG.

This complementation assay compared the VC_1-IIC_2 fusion protein to the previously described IC_1-IIC_2 fusion protein (Tang and Gilman, 1995). Both TP2000 and TP610A adenylyl cyclase deficient *E. coli* were transformed with vectors encoding either the VC_1-IIC_2, IC_1-IIC_2, or VC_1-Gly-IIC_2 fusion proteins. The VC_1-Gly-IIC_2 fusion vector was made to facilitate the addition of random linkers. In the absence of IPTG, the activity of both VC_1-IIC_2 and the VC_1-Gly-IIC_2 fusion proteins complemented the metabolic defect in the TP2000 strain. Weaker complementation or pink growth was evident at 10 μM IPTG, and no growth was seen at 100 μM IPTG. The IC_1-IIC_2 fusion protein failed to complement the adenylyl cyclase deficiency at all concentrations of IPTG. The TP610A bacteria were weakly complemented by the activity of both VC_1-IIC_2 and VC_1-Gly-IIC_2 fusion proteins in the absence of IPTG and strongly complemented with the addition of 10 μM IPTG. However, only sparse white bacterial growth was seen at 100 μM. The activity of the IC_2-IIC_2 protein failed to complement the metabolic deficiency and formed white colonies at all concentrations of IPTG in the TP610A strain. The red growth on MacConkey agar correlated with growth on

minimal medium. These data suggests that the VC1-IIC2 fusion protein has greater adenylyl cyclase activity than the IC1-IIC2 fusion protein.

Discussion

The phenotypic complementation of the adenylyl cyclase-deficient *E. coli* with the activity of the VC_1-IIC_2 fusion protein was unexpected. The construction of the pQE60-H6-VC_1-IIC_2 chimeric fusion vector was to serve as a parent vector to screen for linkers that allow functional interaction of the VC_1 and IIC_2 domain when expressed as a linked peptide. This parent vector encoded the VC_1 domain (residues 364 to 580) followed by the IIC_2 domain (residues 870 to 1090) separated by only glycine and threonine residues.

The results from the complementation assay suggest that the survival of the bacteria was related to the amount of VC_1-IIC_2 produced. With increasing concentrations of IPTG, more of the VC_1-IIC_2 fusion protein is expressed resulting in more adenylyl cyclase activity and production of cAMP. It has been shown that *E. coli* are sensitive to excess cAMP, and overexpressing adenylyl cyclase in an *E. coli* strain containing a *cya* deletion with an intact *crp* gene has been shown to be lethal (Roy et al., 1983). The CRP transcription factor induces the expression of genes in response to cAMP. Therefore, the lack of growth of bacteria expressing the VC1-IIC2 fusion protein at 100 μM IPTG results from cAMP toxicity (Danchin et al., 1984; Roy et al., 1983; Roy et al., 1983).

As shown in the bacterial screen, the basal expression of the linked VC_1-IIC_2 from the pQE60 or pQE80 expression vectors resulted in selection of poorly expressing bacteria due the activity of the fusion protein and the resulting cAMP toxicity. An expression system with tighter control of basal expression was explored. The VC_1-IIC_2 fusion was subcloned into a T7 promoter expression vector. The constitutive expression of T7 lysozyme from the pLysS vector reduced the basal expression of the protein. This system was used to express VC_1-IIC_2 in *E. coli* strain BL21(DE3), and an initial attempt to purify the fusion protein made. The expression level of the VC_1-IIC_2 fusion was poor.

Adenylyl cyclase activity was enriched following purification on Ni^{2+}-NTA, although many contaminating proteins remained. Alternative methods would have to be considered to circumvent the cAMP toxicity to allow sufficient expression and purification of the VC_1-IIC_2 fusion protein.

Expression of the linked proteins in the adenylyl cyclase deficient strains should be explored. In the absence of IPTG, the enzymatic activity of the VC_1-IIC_2 fusion was insufficient to complement the TP610A strain; however, the adenylyl cyclase deficiency was complemented by the addition of 10 μM IPTG. The lack of phenotypic complementation in the absence of IPTG should eliminate the selection of poorly expressing clones and allow induction of protein expression with IPTG. Another *E. coli* strain, TP2339, contains deletions in both the *cya* and the *crp* genes (Roy et al., 1983). The *crp* gene encodes the cAMP receptor protein (CRP) which is a cAMP responsive transcription factor. The TP2339 strain should exhibit less toxicity to the cAMP produced from the VC_1-IIC_2 fusion due to the loss of CRP. The complementation assay could be performed with the TP2339 strain transformed with the linked VC_1-IIC_2. The expression of the VC_1-IIC_2 fusion in both of these strains could be explored.

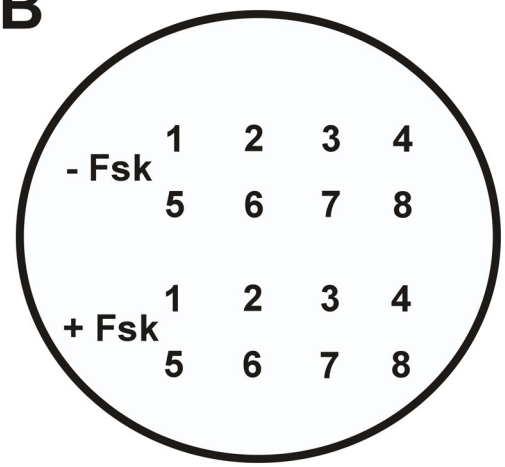
Reducing the enzymatic activity of the VC_1-IIC_2 fusion would be another approach to reduce the cAMP toxicity resulting from the expression of the VC_1-IIC_2 fusion protein. The generation of a mutation in adenylyl cyclase that decreases enzymatic activity without altering ATP affinity could be useful. Mutation of asparagine 1025 in the IIC_2 domain reduces K_{cat} without affecting K_M for ATP or its apparent affinity for the C_1 domain, forskolin, or $G_{s\alpha}$ (Yan et al., 1997b). Asparagine 1025 is hypothesized to form a hydrogen bond with the O4' oxygen of the ribose of ATP based on the structure of $G_{s\alpha}$:VC_1:IIC_2:forskolin soaked with the substrate analog adenosine 5'-(α-thio)-triphosphate (R_p) and manganese (Tesmer et al., 1999). The mutation of this residue would not allow the proper orientation of ATP in the active site. Mutations of asparagine 1025 were constructed and could be assayed for complementation of the adenylyl cyclase deficient *E. coli*.

Expression of the linked VC_1-IIC_2 in the adenylyl cyclase deficient *E. coli* strains, use of the asparagine 1025 mutation, or a combination could lead to the purification of a linked VC_1-IIC_2 chimera. Biochemical characterization will determine if the linked VC_1-IIC_2 is regulated by forskolin, $G_{s\alpha}$, and P-site inhibitors similarly to wild type adenylyl cyclase. This linked VC_1-IIC_2 could lead to the structure of a functional C_1-C_2 heterodimer in the absence of $G_{s\alpha}$ and/or forskolin. These structures would elucidate the conformational changes that activate adenylyl cyclase in response to $G_{s\alpha}$ and/or forskolin.

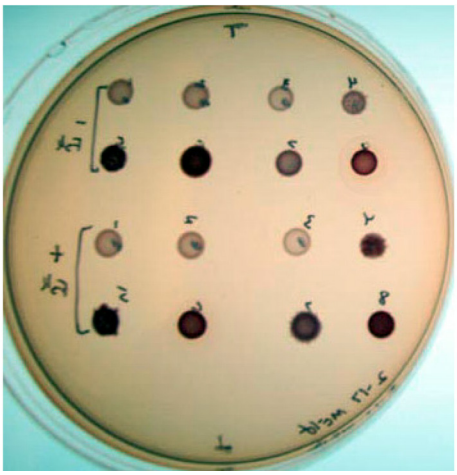
A

Position	pQE60 plasmid	pACYC-T5 plasmid
1	vector	vector
2	VC1	vector
3	vector	IIC2
4	VC1	IIC2
5	VC1	IIC2-K1014N
6	VC1	IIC2 + Gsα
7	IIC2	IC1 + Gsα
8	IIC2-K1014N	IC1 + Gsα

B



C



D

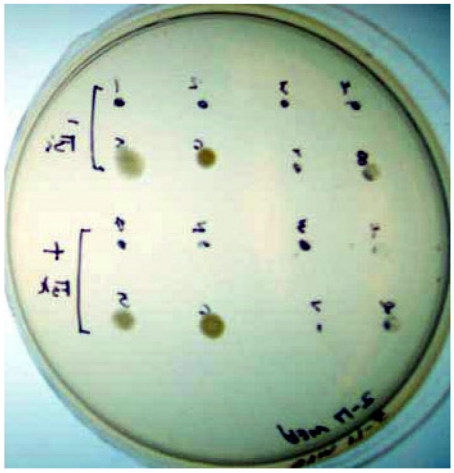
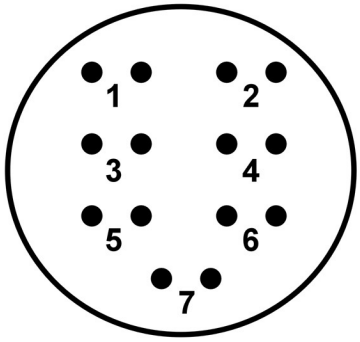


Figure 4-1. Phenotypic complementation of adenylyl cyclase-deficient TP2000 *Escherichia coli*. **A.** TP2000 bacteria (F-, *xyl, cya*Δ*, argH1, lac*Δ*X74)* were cotransformed with the compatible plasmids listed. The pQE60 plasmid and pACYC-T5 contained either the C_1 or C_2 domain of adenylyl cyclase. Clones 6, 7, and 8 dicistronically express $G_s\alpha$. **B.** Single colonies from the transformations in A were resuspended in buffer and spotted (1 µl) onto MacConkey agar (**C**) and M63 minimal medium (**D**) in the presence and absence of forskolin according to the template. Forskolin was spotted (5µl, 10 mM solution) onto the plates prior to the addition of bacteria. The plates contained 0.4% maltose , 50 µM carbenicillin, 50 µM kanamycin, and 100 µM IPTG. Minimal plates also contained 40 µg/ml arginine. Plates were incubated at 30 ^{o}C for 24 and 48 hours for MacConkey and M63 plates respectively.

A

Position	pQE80 plasmid	pACYC-T5 plasmid
1	vector	vector
2	VC1	vector
3	vector	IIC2
4	VC1	IIC2
5	VC1	IIC2-K1014N
6	VC1	IIC2 + Gsα
7	VC1-IIC2	vector

B



C

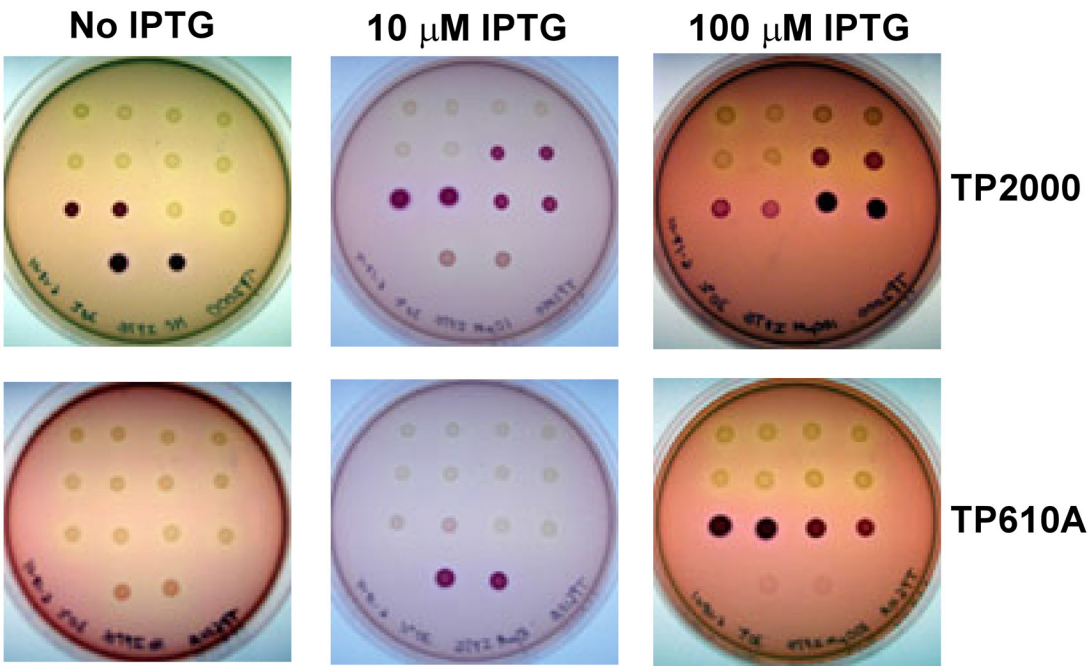
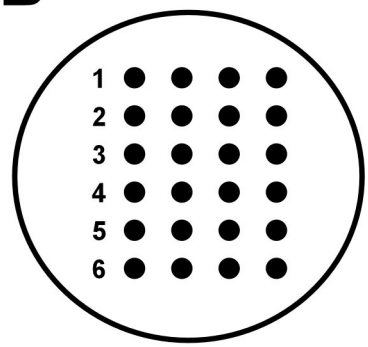


Figure 4-2. Phenotypic complementation of adenylyl cyclase deficient TP2000 and TP610A *Escherichia coli* constitutively expressing Lac repressor. **A.** Both TP2000 and TP610A (F^-, *thi-1, leuB6, pro, lacY1, tonA21, supE44, hsdR, hsdM, recBC, lop-11, lig*$^+$*, cya-610)* were transformed with the compatible plasmids listed in the table and plated on LB plates containing 50 μg/ml of both carbenicillin and kanamycin. Bacteria harboring the pQE80 vector constitutively express Lac repressor. **C.** Single colonies from the transformations in A were resuspended in buffer and spotted (1 μl) onto MacConkey plates following the template (**B**). The plates contained 0.4% maltose, 50 μM carbenicillin, 50 μM kanamycin, and either 0, 10, or 100 μM IPTG. Plates were incubated at 30 ^oC for 24 hours.

A

Row #	Linked Construct	E. coli strain
1	VC1-IIC2	TP2000
2	VC1-Gly-IIC2	TP2000
3	IC1-IIC2-L3	TP2000
4	VC1-IIC2	TP610A
5	VC1-Gly-IIC2	TP610A
6	IC1-IIC2-L3	TP610A

B



C

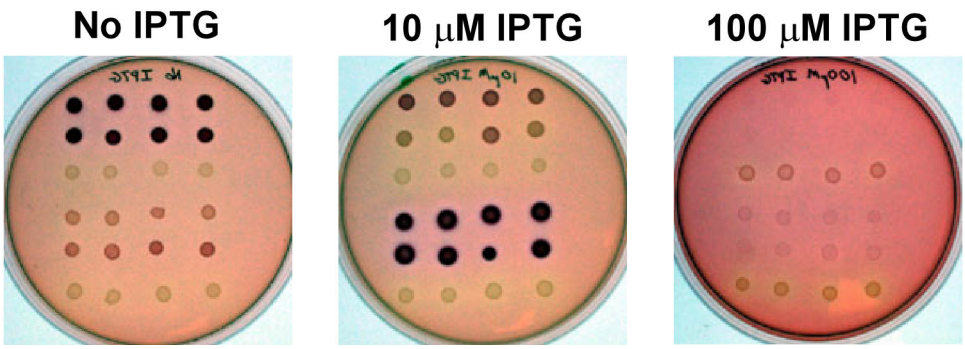


Figure 4-3. Phenotypic complementation of adenylyl cyclase-deficient *Escherichia coli* with linked adenylyl cyclase domains. **A.** Both TP2000 and TP610A (*F⁻, thi-1, leuB6, pro, lacY1, tonA21, supE44, hsdR, hsdM, recBC, lop-11, lig⁺, cya-610)* were co-transformed with pQE60 plasmids encoding the linked cyclase domains listed in the table and pREP4 and plated on LB plates containing 50 μg/ml of both carbenicillin and kanamycin. The pREP4 vector constitutively expresses Lac repressor. **C.** Single colonies from the transformations in A were resuspended in buffer and spotted (1 μl) onto MacConkey plates following the template (**B**). The plates contained 0.4% maltose, 50 μM carbenicillin, 50 μM kanamycin, and either 0, 10, or 100 μM IPTG. Plates were incubated at 30 °C for 24 hours.

Chapter 5 – Mapping the Allosteric Network of Gsα

Introduction

The heterotrimeric G protein alpha subunit (G_α) acts as a binary switch that relays signals dependent on the identity of the bound nucleotide. In the basal state, G_α is bound to GDP and associated with a complex of the G protein β and γ subunits. Binding of α to βγ occludes the sites on α used for interaction with at least some effectors (e.g. adenylyl cyclase). Binding of agonists ligands to heptahelical G protein coupled receptors elicits a conformational change in the G protein alpha subunit that stimulates the release of GDP. The nucleotide free alpha subunit is then able to bind GTP, which is at a much higher intracellular concentration than is GDP. GTP binding alters the conformation of the α subunit, decreasing the affinity for βγ and increasing the affinity for effectors. While much data has been accumulated regarding the kinetics and sites of interaction of effectors, little is know about the molecular network in G_α responsible for relaying the identity of the bound nucleotide to the sites of interaction of G_α with βγ and effectors.

It is easy to hypothesize that this relay is conducted through a network of amino acid residues connecting the nucleotide-binding site and the effector-binding surface. Site-directed mutagenesis could identify these residues; however, mutations at single sites will not reveal the how these amino acid residues interact with other positions. Thermodynamic mutant cycle analysis has been used to study the cooperative interactions of amino acid residues in proteins (Carter et al., 1984a; Hidalgo and MacKinnon, 1995; Schreiber and Fersht, 1995a; Carter et al., 1984b; Schreiber and Fersht, 1995b). This method compares the effect of a mutation in the wild type background with that in the background of a second mutation. Non-additivity (logarithmic) of the effects of two concurrent mutations indicates that the mutated residues are energetically coupled. Thermodynamic mutant cycle analysis was used to identify positions involved in catalysis in the active site of tyrosyl-tRNA

synthetase from *Bacillus stearothermophilus* (Carter et al., 1984a). Energetically coupled residues spanning protein interfaces have also been identified using this method. Schreiber and Fersht used this method to functionally map the interface between barnase and barstar, an extracellular RNase from *Bacillus amylolquefaciens* and its intracellular inhibitor. Residues located in close proximity across the barnase-barstar interface were energetically coupled, while residues separated by more than 7Å were energetically independent (Schreiber and Fersht, 1995a). The functional interaction sites of the Shaker K^+ channel and scorpion toxin were also mapped using this technique (Hidalgo and MacKinnon, 1995; Ranganathan et al., 1996). To perform a complete mutant cycle analysis of the $G_{s\alpha}$ would require over 100,000 mutations; therefore, these studies have been restricted to small proteins, domains or protein interfaces.

Lockless and Ranganathan have described a more focused method for mapping the energetic coupling of proteins (Lockless and Ranganathan, 1999). The method assumes that nature has performed a large-scale mutagenesis experiment during evolution, and by using a diverse structure-based multiple sequence alignment (MSA) of a protein family, the covariance of residues can be calculated. This covariance represents the probability that a change of amino acid sequence at one position in the alignment correlates with a change at another site. The magnitude of covariance is a measure of statistical coupling of amino acid residues ($\Delta\Delta G^{stat}$). An example is illustrated in Figure 5-1. In this example, site *i* is energetically coupled to site *j*. Site *l* has a distribution of amino acids equivalent to the mean of all proteins. Sites *k, j* and *i* have amino acids frequencies that deviate from this mean. A perturbation to glutamine is made at site *i*. The perturbation to glutamine generates a sub-alignment by extracting only those sequences that contain glutamine at position *i* (Figure 5-1C). This sub-alignment must contain a sufficient number of sequences to maintain the diversity representative of the whole protein family. The sub-alignment is considered of sufficient size and diversity if the amino acid distribution of unconserved sites in the sub-alignment shows no change in

amino acid distribution compared the parent alignment. If the sub-alignment meets these criteria, the following properties are valid. The lack of change in distribution at position *l* illustrates that the sub-alignment maintains diversity. Positions are statistically coupled to perturbation *i*-Q if the distribution of amino acid residues has changed from the whole alignment. Figure 5-1 illustrates that the distribution of amino acids at position *j* displays covariance with site *i*; however, the distribution of amino acids at site *k* remains unchanged. Therefore, site *i* is statistically coupled to site *j* but not to site *k* or *l*. The magnitude of covariance of site *j* to the perturbation *i*-Q is measured as a statistical coupling energy ($\Delta\Delta G^{stat}$).

This evolutionary covariance method was used to predict the energetically coupled positions responsible for peptide binding in the PDZ domain (Lockless and Ranganathan, 1999). A single perturbation of an amino acid residue in the active site of the PDZ domain was made, and the statistical coupling to all amino acid positions of the PDZ domain was calculated. Amino acid positions with an increased statistical coupling ($\Delta\Delta G^{stat}$) to the perturbation were studied through mutagenesis and peptide binding. The statistical analysis of energetic coupling correlated well the mutant cycle analysis of the domain. Strikingly, this analysis identified sites distant to the site of perturbation to be energetically coupled. These distant sites were connected through a continuous network of coupled residues. These data suggest that energy distribution in a protein fold is a conserved property that can be elucidated through this covariance analysis.

Suel and coworkers expanded on this statistical covariance method (Suel et al., 2003). In contrast to measuring the statistical coupling to one perturbation as was done for the PDZ domain analysis, several perturbations were made and the statistical coupling for each amino acid position in the alignment was calculated for each perturbation. Each sub-alignment generated by the perturbations maintained the size and diversity of the parent alignment measured by the lack of change in amino acid distribution at the least conserved sites. These data were assembled into a

coupling matrix of $\Delta\Delta G^{stat}$ values where the columns of the matrix represent each individual perturbation experiment and the rows of the matrix represent the amino acid positions in the MSA. An example of a mock matrix is illustrated in Figure 5-2A. The statistical coupling ($\Delta\Delta G^{stat}$) of each amino acid position from N-terminus (top) to C-terminus (bottom) for the perturbation in each column is illustrated as a gradient from blue (low coupling) to red (high coupling). The position in each column that represents the site of perturbation is colored red. For example, the third column (D5) represents the statistical coupling of all positions in the MSA (rows 1-15) calculated by generating a sub-alignment that contained aspartate at position 5. Position 5 is the site of perturbation for D5 and is colored red. Perturbation D5 exhibits coupling to positions 1 and 10 (yellow) and increased coupling to position 7 (orange).

Two-dimensional clustering analysis was used by Suel and coworkers to extract patterns from the coupling matrix (Figure 5-2B) (Suel et al., 2003). First, the perturbations (columns) are sorted by the pattern of coupling to positions (rows). Then, the positions (rows) are sorted by the pattern of coupling to the perturbations (columns). The example coupling matrix in Figure 5-2A was clustered using this method and the resulting clustered matrix is shown in Figure 5-2B. Most residues show low statistical coupling (blue); however, two distinct groups distinguished themselves from the rest of the matrix. The first cluster contains perturbations A2, E6, and I14 that are coupled to positions 2, 6, and 14. The second cluster contained perturbations D5 and G10 that are coupled to positions 1, 5, 7, and 10. If the network of coupled residues responsible for energy distribution is a conserved property of the protein family, perturbations to multiple positions in the network should identify each other and contain the same coupling pattern. In the example in Figure 5-2B, the perturbation I14 is coupled to position 6, and perturbation E6 is coupled to position 14. This illustrates self-consistency in the coupling pattern.

Using the method outlined in the mock experiment in Figure 5-2, Suel and coworkers analyzed three distinct and diverse protein families (serine proteases, hemoglobins, and G protein coupled receptors)(Suel et al., 2003). Only a fraction of the amino acid residues in the MSA of these families had significant coupling to the perturbations. The identified residues formed discrete connected networks through each protein linking distant functional regions that correlated with mechanistic analysis of each protein family. For example, the analysis of the G protein coupled receptors (GPCR) identified a discrete network of residues connecting the ligand-binding pocket with the proposed cytoplasmic G protein interaction sites through a series of van der Waals interactions. In response to ligand, the amino acid residues in this network undergo dynamic re-organization as determined by electron paramagnetic resonance (EPR) (Altenbach et al., 2001a; Altenbach et al., 2001b). These studies illustrate that the networks of coupling identified through covariance analyses are evolutionarily conserved in protein families and form physical interactions responsible for long-range communication between functional sites within proteins.

The statistical analysis of covariance of amino acid residues described above will be used to identify the network of residues in G_α responsible for relaying the identity of bound nucleotide to the effector interaction surfaces. The GTP binding protein family consists of the heterotrimeric G protein alpha subunit, the ras family of small GTPases, and the translation elongation factors. The members of this family bind and hydrolyze guanine nucleotides; as well, these proteins bind other proteins depending on the identity of the nucleotide bound. A multiple sequence alignment of the GTP binding protein family was assembled. The statistical analysis of the covariance of amino acid residues in this alignment should identify residues that are important for the common features of the family. However, proteins of this family bind effector molecules through a wide variety of interactions; therefore, I hypothesize that the residues important in effector binding will not be identified by the coupling analysis. It is tempting to speculate that the statistically coupled residues in

the alignment will identify those required for relaying information about the presence or absence of the γ-phosphate of the guanine nucleotide to the effector binding surfaces. These residues would be responsible for the guanine nucleotide-dependent allostery of the G proteins. This chapter describes the role of many positions in $G_{s\alpha}$ identified in the statistical analysis using mutagenesis and assays for interactions with the βγ subunits and adenylyl cyclase.

Experimental Methods

G protein multiple sequence alignment – The multiple sequence alignment (MSA) of the G protein family was assembled by Steve Lockless (Lockless, 2002). G protein sequences were collected using PSI-BLAST (e-score < 0.001) (Altschul et al., 1997) from a non-redundant database using proteins with known structures as queries. The alignment contained 363 ras family members, 124 heterotrimeric G protein alpha subunits, and 230 translation elongation factors. Alignments were generated using ClustalW (Thompson et al., 1994), followed by manual adjustments using guidelines described Methods in Enzymology Volume 266 (Doolittle et al., 1996). Several stretches of residues of $G_{s\alpha}$ were not included in the alignment due to lack of structural conservation. Residues 1-37, 65-196 (helical domain), 253-266 (switch III), 302-336, 349-358, and 392-393 of bovine $G_{s\alpha}$ were not represented in the alignment since these structural elements have no counterparts in the ras-like GTPases or the elongation factors. The G protein MSA contained 163 positions.

Statistical Coupling Analysis – The calculation of statistical coupling was performed as described previously (Lockless and Ranganathan, 1999; Lockless, 2002). The statistical coupling energy ($\Delta\Delta G^{stat}$) between two amino acid residues is a measure of the magnitude of change in amino acid frequency at one position in the MSA to a perturbation at another site. The MSA must be diverse and large in order to glean functional information from the statistical analysis. The diversity of the alignment is illustrated by many positions having amino acid distributions close to the mean

frequencies in all proteins. The MSA should be large so that the random elimination of sequences has little effect on the amino acid distributions. If these criteria are satisfied, the MSA has reached a state of statistical equilibrium in sequence space which allows the use of the Boltzmann equation to describe the statistical energy of observing an amino acid distribution at a particular site (Lockless, 2002). Thus, the statistical coupling energy is represented in kT^*. A perturbation to the MSA must generate a sub-alignment that is also large and diverse such that the equilibrium of the MSA is not altered. If a sub-alignment is in statistical equilibrium, the coupling of least conserved positions in the parent alignment should remain unconserved in the sub-alignment. The minimum number of sequences that generate a meaningful sub-alignment was determined as described (Lockless, 2002). The coupling for the five least conserved residues in the alignment was calculated. The least conserved residues from the alignment correspond to $G_{s\alpha}$ residues D215, L297, D343, E370, and I382. Next, random sequences were eliminated from the alignment and the coupling re-calculated. This process was repeated by serially, randomly eliminating more sequences from the alignment. A plot of coupling, $\Delta\Delta G^{stat}$, versus number of sequences in the sub-alignment following random elimination illustrated that as the number of sequences decreased the coupling of least conserved positions increased (Figure 5-3). This plot was used to determine the 300-sequence sub-alignment size cutoff for maintaining the statistical equilibrium.

Perturbations were only made to amino acid residues that were present at a position in the alignment in at least 300 (~41%) of the 717 sequences. Thirty-five perturbations satisfied this criterion. The statistical coupling to all positions for each of these perturbations was calculated as described (Lockless and Ranganathan, 1999; Lockless, 2002). The coupling matrix of coupling values for all perturbations was assembled. The columns in the matrix are the individual 35 perturbation experiments, and the rows in the matrix are the 163 $G_{s\alpha}$ positions from amino-terminus to

carboxyl-terminus. The magnitude of statistical coupling is illustrated with a gradient from low (blue) to high (red). The sites of perturbation are colored red.

Matrix Clustering – The coupling matrix was clustered using the statistics package in MATLAB 6.1 (MathWorks Inc.) similarly to the clustering methods used for DNA microarray analysis (Getz et al., 2000). Hierarchial clustering was performed by using city-block distance metric and complete linkage. The 35 perturbations were clustered by grouping perturbations that had similar coupling patterns in the positions of the alignment. Likewise, the 163 positions were clustered by grouping the positions that displayed similar coupling patterns to the perturbations. An initial round of clustering columns was used to eliminate 10 perturbations experiments at highly conserved sites that consequently contained very low coupling values. The positions and perturbations with the highest coupling values were selected and extracted for two rounds of iterative clustering (Suel et al., 2003; Lockless, 2002).

Purification of $G_{s\alpha}$ mutations – Several positions identified in the statistical analysis were mutated to alanine in bovine $G_{s\alpha}$. Amino acid 48 of $G_{s\alpha}$ is alanine; therefore, this position was mutated to the next most conserved residue at this position in the MSA, histidine. Trp-234 was mutated to Phe because this mutation allowed for proper folding of the protein when characterized previously in $G_{t\alpha}$ and $G_{o\alpha}$ (Faurobert et al., 1993; Natochin et al., 1998; Lan et al., 1998). Mutations in $G_{s\alpha}$ were made by Quikchange™ (Stratagene, La Jolla, CA) using pQE60- $G_{s\alpha(s)}$-H6 (Lee et al., 1994) as a template. All plasmids were sequenced in the McDermott Center DNA sequencing core by Jay Hunter. Both wild type and mutant $G_{s\alpha}$ proteins were expressed and purified as previously described (Lee et al., 1994) with the following alterations. The pool from the Mono Q HR10/10 column was applied to 1-ml of ceramic hydroxyapatite (40 μm)(Bio-Rad) packed in a HR5/5 column (Amersham Pharmacia Biotech). The $G_{s\alpha}$ was eluted from the column in 50 mM Tris, pH 8, and 2 mM DTT with a linear gradient of potassium phosphate from 0 to 300 mM over 30 column volumes.

The $G_{s\alpha}$ peak was pooled, concentrated, and buffer exchanged into 20 mM NaHEPES pH 8, 1 mM EDTA, 2 mM DTT, and 50 mM NaCl using a 30 kDa cutoff Ultrafree-15 spin concentrator from Millipore. The $G_{s\alpha}$ was aliquoted, flash frozen in liquid nitrogen, and stored at –80 ^{o}C.

The extinction coefficient of $G_{s\alpha}$ bound to GDP was determined to be 51660 $M^{-1}cm^{-1}$ as described (Edelhoch, 1967; Pace et al., 1995). The protein concentration was determined by UV spectrophotometry and by Bradford assay using Bio-Rad protein assay reagent and BSA as a standard. The value from the Bradford assay overestimated the protein concentration compared to the value from UV spectrophotometry, which was deemed the true value. The concentration of all $G_{s\alpha}$ mutants was determined by Bradford assay using BSA as the standard and corrected using a 0.62 correction factor to account for this overestimation.

GTPγS was bound to wild type and mutant $G_{s\alpha}$ as follows. The $G_{s\alpha}$ (typically 4 mg) was diluted to 125 μM in 50 mM NaHEPES pH 8, 10 mM $MgSO_4$, 1 mM EDTA, and 10 mM DTT and was incubated with GTPγS (800 μM) at 30 ^{o}C for 2.5 hours. The free GTPγS was removed by gel filtration through Sephadex G50 fine resin (Amersham Pharmacia Biotech). The pool from the G50 column was concentrated to 100 μl in a Millipore Ultrafree 0.5 ml spin concentrator. The sample was diluted to 500 μl with 50 mM NaHEPES pH 8, 10 mM $MgSO_4$, 1 mM EDTA, 10 mM DTT, and 100 mM NaCl and concentrated to 100 μl. This process was repeated five times. The GTPγS-bound $G_{s\alpha}$ was aliquoted, flash frozen in liquid nitrogen, and stored at –80 ^{o}C.

GTPγS binding and tryptophan fluorescence – $G_{s\alpha}$ (400 nM) was diluted in 50 mM NaHEPES pH 8, 1 mM EDTA, 10 mM DTT, 0.1% Lubrol ($C_{12}E_{10}$), 100 mM NaCl, and 10 mM $MgSO_4$. The reaction was performed at 20 ^{o}C and initiated by the addition of 20 μM [^{35}S]GTPγS (specific activity of 2000 cpm/pmol). Duplicate 25-μl samples were removed at various time points and binding of radioactive nucleotide was stopped with the addition of excess unlabelled GTP and

$MgSO_4$. Protein was separated from free nucleotide by adsorption to BA-85 nitrocellulose filters (Schleicher and Schuell). The amount of GTPγS bound was determined by scintillation counting. The data were fit using a nonlinear least-squares method to the following equation: $B = B_{max}(1 - e^{-kt})$ where B is the concentration of bound GTPγS, B_{max} is the apparent maximum bound, *k* is the rate constant and t is time.

Intrinsic tryptophan fluorescence of $G_{s\alpha}$ was measured as follows. $G_{s\alpha}$ (400 nM) was diluted in 100 µl 50 mM NaHEPES pH 8, 1 mM EDTA, 10 mM DTT, 0.1% Lubrol ($C_{12}E_{10}$), 100 mM NaCl, and 10 mM $MgSO_4$. The fluorescence was monitored at 350 nm on a Photon Technology International fluorometer following excitation at 295 nm with a 1 nm excitation and emission slit width. The reaction was performed at 20 ^oC and was started with the addition of 100 µM GTPγS. The data were fit using a non-linear least squares method to the following equation: $B = B_{max}(1 - e^{-kt})$ where B is the concentration of bound GTPγS, B_{max} is the apparent maximum bound, *k* is the rate constant and t is time.

Determination of nucleotide bound to Q227A mutant protein – The identity of the nucleotide bound to the Q227A-$G_{s\alpha}$ purified from *E. coli* was determined by anion exchange chromatography. The Q227A-$G_{s\alpha}$ mutant protein purified from *E. coli* and stored in 20 mM NaHEPES pH 8, 1 mM EDTA, 2 mM DTT, and 50 mM NaCl was diluted to 30 µM protein with 100 mM potassium phosphate pH 7.0. The protein was heated at 100 ^oC for 30 minutes. The guanine nucleotide was separated from denatured protein with three phenol:chloroform extractions followed by three chloroform extractions. 200 µl of the resulting solution was applied to a Synchropak AX300 100 x 4.6 mm anion exchange column (Micra Scientific), eluted at 1 ml/minute with a gradient of potassium phosphate pH 7.0 from 100 mM to 400 mM. The absorbance at 253 nm was recorded. GDP and GTP, 200 µl of 15 µM, were applied and eluted with the same potassium phosphate pH 7.0 gradient.

Purification of H6-$G_{i\alpha 1}$-ECFP and $\beta_1\gamma_2$-EYFP – The pQE60-H6-$G_{i\alpha 1}$-ECFP plasmid was constructed by Scott Gibson (unpublished results). Nucleotides encoding enhanced cyan fluorescent protein, ECFP, were inserted into the sequence of rat $G_{i\alpha 1}$ between Ala-121and Glu-122 of the helical domain. The H6-$G_{i\alpha 1}$-ECFP was expressed and purified as previously described for H6- $G_{i\alpha 1}$ (Lee et al., 1994). Enhanced yellow fluorescent protein, EYFP, was added onto the amino-terminus of rat β_1. The β_1-EYFP was coexpressed in Sf9 insect cells with rat γ_2 and $G_{i\alpha 1}$-H6. The $\beta_1\gamma_2$-EYFP complex was purified as described by Kozasa (Kozasa and Gilman, 1995). Protein concentrations were determined by UV absorbance using extinction coefficients of 26,000 M^{-1} at 430 nm for ECFP and 84,000 M^{-1} at 514 nm for EYFP.

Adenylyl cyclase assay – The cytosolic domains of adenylyl cyclase were used to measure the apparent affinities of GDP- and GTPγS-bound $G_{s\alpha}$ for adenylyl cyclase. The H6-VC_1(591)FLAG and IIC_2-H6 proteins were purified as previously described (Hatley et al., 2002). Adenylyl cyclase activity was measured as the conversion of [α-^{32}P]ATP to [^{32}P]cAMP as reported by Smigel (Smigel, 1986). All assays were performed in a volume of 50 µl for 10 minutes at 30 ^{o}C in the absence of an ATP regenerating system. The final concentrations of free $MgCl_2$ and ATP were 5 mM and 1 mM, respectively. IIC_2 protein was held at a limiting concentration (1 nM) and VC_1 was present in excess (10 µM). $G_{s\alpha}$ was diluted in 20 mM NaHEPES pH 8, 2 mM $MgCl_2$, 1 mM EDTA, and 2 mM DTT. Either GDP- or GTPγS-bound $G_{s\alpha}$ was added at increasing concentrations. Assays including GDP-bound $G_{s\alpha}$ also contained a 50-µM excess of GDP. The adenylyl cyclase domain proteins and $G_{s\alpha}$ were mixed and incubated on ice for 30 min. The reaction was started with the addition of [α-^{32}P]ATP. All specific activities reported are with respect to the concentration of the limiting IIC_2. The data were fit to a one site binding model to the equation $y = Vmax*[G_s]/([G_s] + EC_{50})$ using a non-linear least squares method in Microcal Origin (Northhampton, MA). The activity of each

mutant was measured two or three times and that of the wild type protein was measured twenty six times. The mean and range of the EC_{50} from the fit was calculated for both the GDP- and GTPγS-bound $G_{s\alpha}$ proteins. Ratios of the EC_{50} values for GDP- and GTPγS-bound proteins were calculated by dividing the EC_{50} for $G_{s\alpha}$-GDP by that for $G_{s\alpha}$-GTPγS with propogation of errors. Each ratio was tested for significant deviation from the wild type $G_{s\alpha}$ value with a Student's t test and the significance was reported as a *P* value.

The affinity of $G_{s\alpha}$ for $\beta\gamma$ -- A fluorescence assay to measure the apparent affinity of $G_{s\alpha}$ for $\beta_1\gamma_2$ was developed. Binding of $G_{i\alpha1}$-ECFP and $\beta_1\gamma_2$-EYFP chimeric fusion proteins has been monitored by the fluorescence resonance energy transfer (FRET) from ECFP to EYFP (Scott Gibson, unpublished results). The disappearance of FRET was used to monitor competition between $G_{s\alpha}$ and $G_{i\alpha1}$-ECFP for binding $\beta_1\gamma_2$-EYFP. Due to the high affinity of myristoylated $G_{i\alpha1}$ for βγ (<1 nM), unmyristoylated H6-$G_{i\alpha1}$-ECFP was purified from *E. coli* and used in this assay. The lack of the myristate reduced the affinity of $G_{i\alpha}$ for $\beta_1\gamma_2$ to allow the binding reaction to reach equilibrium within 16 hours at 18 ^{o}C and allow the affinity of GTPγS-bound $G_{s\alpha}$ for $\beta_1\gamma_2$-YFP to be determined. GDP- and GTPγS-bound $G_{s\alpha}$, H6-$G_{i\alpha1}$-ECFP, and $\beta_1\gamma_2$-EYFP were diluted in 50 mM NaHEPES pH 8, 1 mM EDTA, 5 mM DTT, 100 mM NaCl, 2 mM $MgSO_4$, 0.1% Lubrol ($C_{12}E_{10}$), and 200 μM GDP. Increasing concentrations of $G_{s\alpha}$-GDP or $G_{s\alpha}$-GTPγS were added to H6-$G_{i\alpha1}$-ECFP and $\beta_1\gamma_2$-EYFP (200 nM each) and incubated at 18 ^{o}C for 16 hours. The fluorescence spectra were obtained using a Photon Technology International fluorometer. The sample was excited at 410 nm and the emission was recorded from 461 to 535 nm at a 3 nm slit width at 20 ^{o}C. The ratio of the fluorescence intensity of EYFP at 522 nm to ECFP at 474 nm was plotted versus the $G_{s\alpha}$ concentration. The data were fit to the following equation using non-linear least squares method in Microcal Origin: $y = (B_{max}*(IC_{50}/[G_{s\alpha}]))/(1+(IC_{50}/[G_{s\alpha}]) + y0$ where y0 is the offset of the lower asymptote and B_{max} is the

upper asymptote. The apparent affinity of each mutant $G_{s\alpha}$ protein for $\beta\gamma$ was measured twice and the value for the wild type protein was determined ten times. The mean and range of the IC_{50} from the fit was calculated for both the GDP- and GTPγS-bound proteins. Ratios of the IC_{50} values for GDP- and GTPγS-bound proteins were calculated as the magnitude difference between the IC_{50} of GTPγS-bound protein and that of the GDP-bound protein with propogation of errors. Each ratio was tested for significant deviation from the wild type $G_{s\alpha}$ value with a Student's t test and the significance was reported as a *P* value.

Results

Assembly of the coupling matrix – The statistical analysis of covariance of amino acid residues in a MSA described previously (Lockless and Ranganathan, 1999; Lockless, 2002; Suel et al., 2003) was used to identify coupled residues responsible for relaying the identity of bound nucleotide to the effector binding surface in $G_{s\alpha}$. The structure-based sequence alignment of the GTP binding protein family includes 717 sequences: 363 ras-like small GTPases, 124 heterotrimeric G protein alpha subunits, and 230 translation elongation factors. In order to extract functional information from this matrix, the MSA must be diverse and large. This insures that the alignment is in statistical equilibrium in sequence space, which is required for the calculation of coupling ($\Delta\Delta G^{stat}$). If the MSA is in statistical equilibrium, the random elimination of sequences from the MSA should have little effect on the amino acid distribution of the positions. This was tested by calculating the coupling of the five least conserved positions in the MSA (D215, L297, D343, E370, and I382 of $G_{s\alpha}$). If the MSA is diverse, the coupling of these least conserved positions should be low as is shown in Figure 5-3 for the G protein MSA. The coupling of these least conserved residues was calculated after randomly eliminating an increasing number of sequences (Figure 5-3) to determine the minimum number of sequences in the MSA that is representative of the whole G protein family.

As more sequences are randomly eliminated from the MSA, the remaining sequences are less representative of the whole MSA. If the MSA is large and in statistical equilibrium, the random elimination of sequences should have little effect on the coupling of the least conserved positions until the number of sequences remaining is no longer representative of the whole family. The coupling of the least conserved sites in the G protein family shows little change until only 200 sequences remain at which point the coupling of these positions increases. Therefore, the G protein alignment was sufficiently large and diverse to perform the statistical analysis.

The coupling between two positions was calculated by making perturbations to the MSA as described (Lockless and Ranganathan, 1999; Lockless, 2002). A perturbation is a sub-alignment made by extracting sequences containing a specific amino acid residue at a position. The coupling of positions to the perturbation was a measure of the change in amino acid distribution between the sub-alignment and the whole MSA. The number of sequences in the sub-alignment must be in statistical equilibrium to calculate the coupling. From inspection of Figure 5-3, the minimum number of sequences for generating a sub-alignment was determined to be 300 sequences. According to Figure 5-3, when 300 or more sequences were present, the alignment was in equilibrium as represented by the low coupling values of the least conserved positions. Therefore, perturbations were only made at positions in the MSA that contained a specific amino acid residue in 300 or more of the sequences. The value of 300 sequences represents 41% of the alignment. Thirty-five perturbations satisfied these criteria. The coupling data were presented as a coupling matrix of perturbations (columns) versus positions (rows) of the G protein alignment (Figure 5-4A). In this coupling matrix, each column represents a separate perturbation experiment, and each row represents coupling of each position in the primary sequence of $G_{s\alpha}$ from amino-terminus (top) to carboxyl-terminus (bottom) to that perturbation. The magnitude of coupling is illustrated in a linear scale from blue ($0kT^*$) to red ($5.7kT^*$).

Clustering of the coupling matrix - The coupling matrix was subjected to two-dimensional cluster analysis to analyze patterns in the coupling (Figure 5-4B) (Suel et al., 2003; Lockless, 2002). This analysis sorts the 35 perturbations that contain similar coupling patterns to positions. Next, the 163 positions are sorted by the patterns of coupling to multiple perturbations. Most positions exhibited no coupling to the perturbations in the first round of clustering illustrating the diversity of the MSA and the evolutionary independence of these positions. Only 21% of the positions in the sequence exhibited statistical coupling. The positions with large coupling values were extracted for iterative clustering. Iterative clustering was performed similarly to the DNA microarray analysis (Getz et al., 2000). Two rounds of iterative clustering of the matrix in Figure 5-4B yielded two overlapping groups (Figure 5-4C). The first group contained many positions that were coupled to perturbations 224T, 225A, and 228E. The second group was coupled to perturbations in the first group and coupled to a distinct group of perturbations at the left of the matrix. Both of these clusters were self consistent. For example, in the first group perturbation 224T is highly coupled to position 228 and perturbation 228E is highly coupled to position 224. In the second group perturbation 42K is strongly coupled to position 376 and perturbation 376F is strongly coupled to position 42. This self consistent property was expected if the clustering pattern identified the amino acids comprising a co-evolving network.

The positions identified in the iterative cluster have been mapped on the structures of $G_{s\alpha}$-GTPγS (Sunahara et al., 1997b) and $G_{i\alpha}$-GDP in a co-crystal with βγ (Wall et al., 1995) as shown in Figure 5-5. In the $G_{s\alpha}$-GTPγS structure (Figure 5-5A&B), the coupled residues are within van der Waals contact of each other, forming a network from the back of the nucleotide pocket through the core and terminating at switch II. The statistically coupled residues in switch II (W234 and Q227) are in contact with coupled residues in the α3 helix (I276) and the β1-α1 loop (A48 and E50, respectively). When bound to GDP, a fracture is formed between the coupled residues in switch II

and the core of the alpha subunit (Figure 5-5C&D). It is tempting to speculate that these statistically coupled amino acid residues identified in cluster analysis comprise the network of residues responsible for relaying the identity of bound nucleotide to the effector binding sites and thus regulating the allostery. The functions of these coupled positions were characterized by mutagenesis and biochemical assays to determine binding to adenylyl cyclase and βγ.

Six non-coupled positions were chosen for characterization as controls for the method (L46, S205, I207, D229, E230, and L272 in $G_{s\alpha}$). These non-coupled positions exhibited low statistical coupling in the clustered matrix (Figure 5-4B) and as a result were excluded from the iterative matrix (Figure 5-4C). These non-coupled control residues are located adjacent to statistically coupled positions in the structure (Figure 5-6). Serine-205 in switch I contacts glutamate-230 in switch II in a nucleotide dependent manner. The residues are in contact in the GTP-bound state (Figure 5-6A), and this contact is broken in the GDP-bound state (Figure 5-6B). This state-dependent contact of serine-205 and glutamate-230 is similar to the state-dependent contact of the coupled positions tryptophan-234 and isoleucine-276. The mutations of these non-coupled positions and subsequent characterization will test the robustness of the statistical method.

Mutational analysis – To gain insight into regulation of $G_{s\alpha}$, the positions identified by the statistical method were mutated in $G_{s\alpha}$ and assayed for allosteric interactions. The predicted coupled positions and the non-coupled positions were mutated to alanine (with the exception of Ala-48 and Trp-234 which were mutated to His and Phe, respectively) in $G_{s\alpha(short)}$, expressed in *E. coli,* and purified to homogeneity. Allosteric regulation of the $G_{s\alpha}$ protein was assessed by measuring interactions of $G_{s\alpha}$-GTPγS and $G_{s\alpha}$-GDP with both adenylyl cyclase and the G protein βγ subunit complex. Adenylyl cyclase activity was measured with increasing concentrations of GTPγS-bound and GDP-bound $G_{s\alpha}$ using the cytosolic domains of adenylyl cyclase (Hatley et al., 2002). There is a 5.7-fold difference between the EC_{50} for activation of adenylyl cyclase for GTPγS-bound compared to

GDP-bound $G_{s\alpha}$ (Table 5-1). This difference in EC_{50} values for the GTPγ-S bound and GDP-bound $G_{s\alpha}$ was used for comparison of the wild type protein with the mutants.

$G_{s\alpha}$ also binds to the βγ subunits in a nucleotide-dependent fashion. $G_{s\alpha}$-GDP has a much greater affinity for βγ than does $G_{s\alpha}$-GTPγS (Gilman, 1987). Previous determinations of the affinity of G_{α} and βγ subunits measured the indirect functional effect of βγ inhibition of steady state GTP hydrolysis (Ueda et al., 1994). These assays would not measure the apparent affinity difference of GDP- and GTPγS-bound $G_{s\alpha}$ for βγ. A sensitive fluorescence-based competition assay was developed to directly measure the interactions of $G_{s\alpha}$ and the βγ subunits. Chimeras of $G_{i\alpha1}$-ECFP and $\beta_1\gamma_2$-EYFP were constructed and purified to homogeneity (Scott Gibson, unpublished results). The binding of the $G_{i\alpha1}$-ECFP and $\beta_1\gamma_2$-EYFP chimeras was monitored by the fluorescence resonance energy transfer, FRET, from ECFP to EYFP (Scott Gibson, unpublished results). Disappearance of FRET was used to monitor competition between $G_{s\alpha}$ and $G_{i\alpha1}$-ECFP for binding to $\beta_1\gamma_2$-EYFP (Figure 5-7). Myristoylated $G_{i\alpha}$ has a very high affinity for the βγ subunits, less than 1 nM (Sarvazyan et al., 1998). The affinity of the $G_{i\alpha1}$-ECFP and βγ-EYFP was altered by purifying non-myristoylated $G_{i\alpha1}$-ECFP from *E. coli*. The reduced affinity allowed the assay to reach equilibrium within 16 hours and allowed the determination of apparent affinity for the GTPγS-$G_{s\alpha}$. The apparent affinities determined in this assay are not reflective of the actual affinity of $G_{s\alpha}$ for βγ in the cell, but provided a means to compare the difference in apparent affinity of GDP- and GTPγS-bound wild type and mutant $G_{s\alpha}$ proteins for βγ. $G_{s\alpha}$-GDP had a 19-fold greater apparent affinity for βγ than did $G_{s\alpha}$-GTPγS (Figure 5-7C and Table 5-2). The $G_{s\alpha}$ mutants were assayed for deviations from the wild type fold effect.

The EC_{50} and IC_{50} values for the adenylyl cyclase activation and βγ-binding assays for the $G_{s\alpha}$ mutants are summarized in Tables 5-1 and 5-2 respectively. The phenotypes of the statistically coupled mutants can be grouped into five classes. The first class of mutations consists of Q227A and

R201A. Mutations at both of these positions have been previously identified to create a constitutively active $G_{s\alpha}$ proteins by decreasing the GTPase activity (Graziano and Gilman, 1989b; Masters et al., 1989; Freissmuth and Gilman, 1989). Strikingly, opposite phenotypes were observed in the adenylyl cyclase activation assay and the βγ binding assay. The GDP- and GTPγS-bound mutant proteins exhibited a decrease in the difference of the apparent affinity for activation of adenylyl cyclase and an increase in the difference of apparent affinity for βγ binding.

The Q227A mutant protein displayed a complete loss of allosteric activation of adenylyl cyclase (Figure 5-8A) since the GDP-bound form of Q227A activated adenylyl cyclase as well as the GTPγS-bound protein. Since the Q227A mutant protein has decreased GTPase activity, the loss of allosteric activation of adenylyl cyclase could be explained if the protein purified from *E. coli* was bound to GTP. However, the Q227A mutant protein was denatured, and the guanine nucleotide separated by anion exchange chromatography was GDP (Figure 5-9). Curiously, the Q227A mutant protein demonstrated an increase in tryptophan fluorescence upon binding GTPγS similar to that of the wild type protein (Table 5-3). Therefore, the Q227A mutant protein undergoes a conformational change associated with the binding of GTP, but shows no discrimination in the activation of adenylyl cyclase. The GTPγS-bound Q227A mutant protein bound βγ less well than the wild type protein resulting in a 50-fold difference in the apparent affinities for binding of βγ compared to the GDP-bound proteins (Figure 5-8B). The diffence in binding of the GDP- and GTPγS-bound Q227A mutant protein reinforced that the mutant protein purified from *E. coli* was bound to GDP in spite of the decreased GTPase activity. The GDP-bound R201A had an increased apparent affinity for adenylyl cyclase resulting in a 2.8-fold difference between GTPγS- and GDP-bound proteins (Figure 5-8C). Both GTPγS- and GDP-bound R201A had a decreased apparent affinity for βγ. However, the difference between the apparent affinities of the GTPγS- and GDP-bound proteins was increased to

45-fold (Figure 5-8D). The Q227A and R201A mutant proteins exhibit diametric phenotypes. These mutant proteins displayed decreased nucleotide sensitivity in binding adenylyl cyclase and increased nucleotide sensitivity in binding βγ.

The second class of mutants (R42A, A48H, G225A, R228A, W234F, and I276A) had the opposite effects as the first class on the activation of adenylyl cyclase and binding βγ thus maintaining the reciprocal regulation. The GDP- and GTPγS-bound mutant proteins exhibited an increase in the difference of the apparent affinities for activation of adenylyl cyclase and a decrease in the difference of apparent affinities for βγ binding. The GTPγS- and GDP-bound mutants G225A, R228A, W234F, and I276A exhibited an increased EC_{50} for the activation of adenylyl cyclase; however, the difference between the apparent affinities of the GTPγS- and GDP-bound mutant proteins was greater than that of the wild type protein (Figures 5-10A&C and Figure 5-11A&C). Both GDP-bound R42A and A48H mutant proteins displayed a decreased apparent affinity for adenylyl cyclase while the GTPγS-bound proteins showed no difference in apparent affinity compared to the wild type protein. This change resulted in an increase in the nucleotide-dependent difference between the apparent affinities of the mutant proteins compared to the wild type protein (Figure 5-12A&C). All the mutant proteins in this class displayed a decrease in the difference between the apparent affinity of GDP- and GTPγS-bound $G_{s\alpha}$ for βγ compared to the wild type protein; however, the positions of the curves varied. The GTPγS-bound A48H and G225A mutant proteins bound βγ with the same apparent affinity as the GDP-bound proteins (Figure 5-10B and Figure 5-12D). The R42A, W234F, and I276A mutant proteins had a decreased apparent affinity for the GTPγS-bound mutant and a more dramatic decreased affinity for the GDP-bound mutants. This resulted in a reduced fold difference compared to the wild type protein (Figure 5-11B&D and Figure 5-12B). The R228A mutant protein exhibited a decreased apparent affinity for βγ when bound to

GDP and an increased apparent affinity for βγ when bound to GTPγS (Figure 5-10D). This class of mutant proteins maintained the phenotypic reciprocity toward binding of adenylyl cyclase and βγ.

The third class of coupled mutants (V224A, C237A, and C365A) had an effect on the activation of adenylyl cyclase without affecting the binding to βγ. The difference between the apparent affinity of adenylyl cyclase for the GDP- and GTPγS-bound V224A and C237A mutant proteins was increased compared to the wild type protein. The difference in apparent affinities for βγ of the GDP- and GTPγS-bound mutant proteins was the same as the wild type $G_{s\alpha}$. The GDP-bound C237A protein had a decreased apparent affinity for adenylyl cyclase resulting in a 12-fold difference from the GTPγS-bound protein (Figure 5-13A). Both the GDP- and GTPγS-bound C237A protein displayed a decrease in the apparent affinity for the βγ subunits without a statistically significant change in fold difference between the nucleotide states (Figure 5-13B). The GTPγS-bound V224A mutant protein had an increased apparent affinity while the GDP-bound V224A mutant protein had a decreased apparent affinity for adenylyl cyclase resulting in a 11-fold affinity difference (Figure 5-13C). The binding of the βγ subunits was indistinguishable from wild type in the V224A mutant protein (Figure 5-13D). The GDP-bound C365A mutant protein had an increased apparent affinity for adenylyl cyclase resulting in a decreased difference between the apparent affinities of GDP- and GTPγS-bound C365A mutant protein compared to the wild type $G_{s\alpha}$ (Figure 5-14A). Both the GDP- and GTPγS-bound C365A mutant protein displayed a decrease in the apparent affinity for the βγ subunits without a statistically significant change in fold difference between the nucleotide states (Figure 5-14B). Although the V224A, C237A, and C365A mutant proteins exhibited effects on adenylyl cyclase activation, they lacked the reciprocal effects on binding βγ that was observed in mutant proteins in class 1 and 2.

The fourth class of coupled mutants, E50A and F222A, were not statistically different from wild type with respect to either adenylyl cyclase or βγ affinity (Figure 5-15). The GDP-bound E50A mutant protein appears to have an increased apparent affinity compared to the wild type protein. Although the errors of the points on the two curves do not overlap, the difference in the apparent affinity of the GTPγS- and GDP-bound E50A mutant protein was not significantly different from the wild type protein with the current number of observations. The final class contains a mutant protein that expressed poorly and could not be purified in a sufficient quantity to determine apparent affinity for adenylyl cyclase or βγ. The F376A mutant protein was not soluble when expressed in *E. coli*.

Statistically non-coupled positions were chosen as controls based on their low statistical coupling and proximity to statistically coupled positions. These non-coupled residues are illustrated in Figure 5-6. The S205A mutant protein was indistinguishable from the wild type protein in either of the biochemical assays (Figure 5-16A&B). The I207A, E230A, and L272A mutant proteins all had an effect on the activation of adenylyl cyclase without affecting the binding to βγ. The difference between the apparent affinity of adenylyl cyclase for the GDP- and GTPγS-bound mutant proteins was increased compared to the wild type protein (Figures 5-16C, 5-17C, and 5-18A). Although there were no alterations in the fold difference between the GDP- and GTPγS-bound I207A and E230A mutant proteins compared to the wild type protein, the apparent affinity for βγ was altered for both I207A and E230A mutant proteins (Figures 5-16D and Figure 5-18B). Both the GDP- and GTPγS-bound D229A mutant proteins had an increased affinity for adenylyl cyclase; however, the fold diffence was not significantly different from the wild type protein (Figure 5-17A). The GDP-bound D229A mutant protein increased the apparent affinity for βγ while the GTPγS-bound protein decreased the apparent affinity resulting in a 36-fold difference in the binding of the βγ subunits compared to the wild type protein (Figure 5-17B). Similar to the coupled mutant proteins of class 2, the L46A GDP- and GTPγS-bound mutant protein exhibited an increase in the nucleotide sensitivity

for activation of adenylyl cyclase and a decrease in the nucleotide sensitivity for βγ binding (Figure 5-18 C&D). The GDP-bound L46A mutant protein displayed a decreased apparent affinity for adenylyl cyclase while the GTPγS-bound proteins showed no difference in apparent affinity compared to the wild type protein resulting in an increase in the nucleotide-dependent difference compared to the wild type protein. The L46A mutant protein had a decreased apparent affinity for the GTPγS-bound mutant protein and a more dramatic decreased affinity for the GDP-bound mutant protein resulting in a reduced fold difference compared to the wild type protein. The V247A and N292A mutant proteins were insoluble when expressed in *E. coli* (Table 5-3). Although these non-coupled positions pack against coupled positions, only one of these mutant proteins exhibited the reciprocal effects on the apparent affinities of adenylyl cyclase and βγ.

The overall results of adenylyl cyclase and βγ binding experiments are displayed in Tables 5-1 and 5-2, respectively. The fold-AC for adenylyl cyclase is the fold difference between the apparent affinity of GDP- and GTPγS-bound proteins for adenylyl cyclase and was calculated by dividing the EC_{50} of GDP-bound protein by the EC_{50} of the GTPγS-bound protein. Similarly, the fold-βγ for βγ is the fold difference between the apparent affinity of the GTPγS- and GDP-bound proteins for βγ and was calculated by dividing the IC_{50} of GTPγS-bound protein by the IC_{50} of the GDP-bound protein. The statistical significance of these ratios for the $G_{s\alpha}$ mutant proteins was compared to the ratio of the wild type protein using a Student's t test and the resulting *P* values are listed in Tables 5-1 and 5-2. The deviation of the mutant proteins compared to wild type protein was illustrated in Figure 5-19A&B as the ratio of the fold-AC/βγ of the mutant to the fold-AC/βγ of the wild type protein in the adenylyl cyclase and βγ binding assays respectively. The wild type value is represented as a line at y = 1. Deviations from this line were deemed significant. In this representation, a value >1 indicates increased nucleotide sensitivity relative to the wild-type protein or a greater separation of affinities

between the GDP- and GTPγS-bound states. Similarly, a value <1 indicates decreased nucleotide sensitivity relative to wild type or a less separation of affinities between the GDP- and GTPγS-bound states. The mutants are arranged in order of increasing nucleotide sensitivity for adenylyl cyclase (black bars), which illustrates the reciprocal effects observed in the class 1 and 2 coupled mutants. As the ratio for adenylyl cyclase increases, the ratio for βγ decreases. This reciprocity was only observed in one of the non-coupled mutant proteins, L46A (Figure 5-19B). Although there is an increasing effect on adenylyl cyclase for the remaining non-coupled controls, there is no reciprocal decrease in βγ for the non-coupled mutant proteins.

The normalized ratios for the mutant proteins for βγ were plotted against the normalized ratios for adenylyl cyclase (Figure 5-20). The data for the coupled mutant proteins were fit to an inverse relationship (y = 1/x) demonstrating that the mutations predicted from the statistical coupling analysis are consistent (r = 0.85) with a reciprocal relationship of the nucleotide sensitivities for the two binding partners, adenylyl cyclase and βγ. The data for the control or non-coupled proteins deviated from this relationship (r = 0.42).

Discussion

The G protein alignment used in this study contained ras-like small G proteins, heterotrimeric alpha subunits, and translation elongation factors. These classes of G proteins all interact with a different complement of proteins including effectors, guanine nucleotide dissociation and exchange factors, and GTPase activating proteins. However, all members of the alignment share the properties of interacting with proteins in a nucleotide-dependent manner, nucleotide binding, and nucleotide hydrolysis. The statistical analysis performed here is proposed to identify covarying amino acid residues in a MSA that are important to the functional properties of the protein family. Therefore, the statistical coupling calculated from this alignment was likely to identify the amino acid residues

important for the allosteric properties of G proteins and/or GTPase activity. Strong evidence from Geyer *et al.* indicates that allosteric coupling in G proteins is described by a two-state equilibrium between an inactive state (G, normally stabilized by GDP), and an activated state (G*, normally stabilized by GTP)(Figure 5-21)(Geyer et al., 1996). Mechanistically, the model implies that residues participating in the two-state allosteric mechanism should comprise a cooperative unit making a concerted conformational change between the inactive and active states of the protein. In terms of evolution, it is tempting to hypothesize that this cooperativity forces mutual co-evolution of the constituent sites. Such a network should emerge as a cluster of residues displaying a similar and inter-connected pattern of significant $\Delta\Delta G^{stat}$ values.

To test these predictions, site-directed mutagenesis of $G_{s\alpha}$ at statistically coupled residues was performed. The allosteric model allows specific predictions of the phenotypes expected from mutations within the protein's allosteric core. There are two major features of this model. First, the activated G* state binds adenylyl cyclase and the basal G state binds βγ; these interactions occur at an overlapping surface of $G_{s\alpha}$ and are thus mutually exclusive. Second, the equilibrium between G and G* is given by the allosteric constants L_{GDP} and L_{GTP}, and the essence of nucleotide-dependent switching is in the ratio of these two equilibrium constants. For example, L_{GDP} must be less than 1 (GDP binding stabilizes the G state) and L_{GTP} must be greater than 1 (GTP stabilizes the G* state). It is tempting to hypothesize that mutation of residues in the allosteric core of $G_{s\alpha}$ will perturb the allosteric constants, causing a change of the equilibrium between G and G*. Thus, if a mutation locks the protein in the G* state, it should bind adenylyl cyclase equally well when associated with either GTP or GDP. Similarly, proteins locked in the G state should interact with βγ without regard to the identity of the bound nucleotide. Since the same allosteric mechanism is expected to underlie both L_{GDP} and L_{GTP}, it is predicted that mutations that perturb this mechanism will affect these two equilibrium constants similarly.

At the limit, where the effect of a mutation on L_{GDP} and is equal, the model makes the strong prediction that the mutation will cause the ratio of affinities of $G_{s\alpha}$-GTP and $G_{s\alpha}$-GDP for adenylyl cyclase on the one hand and for $\beta\gamma_{\gamma}$ on the other to vary in a strictly reciprocal fashion. Rama Ranganathan and his laboratory proposed a model for this allosteric system. The apparent equilibrium binding constant (dissociation constant) for the interaction of adenylyl cyclase or βγ with the GTP- or GDP-bound states of $G_{s\alpha}$ will depend on the allosteric constants in the following manner:

$$K_{app,GTP}^{AC} = K_{AC}\left[1+\frac{1}{L_{GTP}}\right]$$

$$K_{app,GDP}^{AC} = K_{AC}\left[1+\frac{1}{L_{GDP}}\right]$$

$$K_{app,GTP}^{\beta\gamma} = K_{\beta\gamma}\left[1+L_{GTP}\right]$$

$$K_{app,GDP}^{\beta\gamma} = K_{\beta\gamma}\left[1+L_{GDP}\right]$$

where K_{AC} is the intrinsic binding constant of adenylyl cyclase to the G* state, and $K_{\beta\gamma}$ is the intrinsic binding constant of βγ to the G state. To understand the effects of mutations that alter the allosteric binding constants, the apparent binding constants for adenylyl cyclase and βγ were plotted for all values of L (Figure 5-22). Mutations in $G_{s\alpha}$ that cause equal left or right shifts of the allosteric constants L_{GDP} and L_{GTP} on the binding affinities of adenylyl cyclase or βγ would have the following effects. Shifts to the right will cause the ratio of affinities of adenylyl cyclase for GTPγS and GDP states (the nucleotide sensitivity) to decrease relative to wild type $G_{s\alpha}$ but will cause the same ratio for βγ to increase relative to wild type $G_{s\alpha}$. Shifts of L_{GDP} and L_{GTP} to the left will cause exactly the opposite effect.

In contrast, mutations that perturb the absolute affinities for adenylyl cyclase or βγ need not display reciprocity toward nucleotide sensitivities. Note that this model makes no claim about the

effect of such mutations on the absolute affinities of $G_{s\alpha}$ for adenylyl cyclase or βγ. Indeed, mutations that do not influence the allosteric core but rather directly alter the binding energy of $G_{s\alpha}$ for adenylyl cyclase or βγ should cause isolated effects on the binding of $G_{s\alpha}$ to one partner with little effect on the interaction with the other. In addition, mutations that affect the allosteric core could also have an independent, direct effect on the affinity of $G_{s\alpha}$ for either adenylyl cyclase or βγ.

Several observations can be made by inspection of the phenotypes of the mutations revealed by the adenylyl cyclase and βγ binding assays. In general, the $G_{s\alpha}$ with mutations at coupled positions that decreased the allosteric difference for adenylyl cyclase increased the allosteric difference of the βγ subunits; conversely, the proteins with mutations at coupled positions that increased the allosteric difference of adenylyl cyclase decreased the allosteric difference of the βγ subunits. This diametric property was only observed in one of the proteins with mutations in non-coupled positions. The two classes of coupled mutations that exhibited this diametric property were mapped on the structure of G protein in the GTPγS- and GDP-bound states, Figure 5-23. The selection of these coupled positions from the surrounding positions and the phenotypes of mutations of these positions could not be predicted from structural analysis alone.

The phenotypes of the mutant proteins can be explained using the allosteric model (Figure 5-21). The mutant proteins of class one, Q227A and R201A, increase both L_{GDP} and L_{GTP} in favor of G* resulting in a decreased allosteric difference for adenylyl cyclase; there will be decreased nucleotide selectivity in the G* state. On the other hand, the shifting of the equilibrium to the G* state results in a concomitant increase in the difference in apparent affinity for GDP- and GTPγS-bound proteins binding to the βγ subunits due to destabilization of the ground state, G. In this case, GTP binding will have a larger impact on G. This model is supported by the structural analysis of Q227A and R201A. Arg-201 forms a contact with Glu-50 in the GDP-bound state that prevents GDP from dissociating. Both Arg-201 and Gln-227 are important in the stabilization of the transition state

of GTP hydrolysis and mutations at either position dramatically reduce the stability of the transition state (Coleman et al., 1994). Proteins with mutations at either Arg-201 or Gln-227 have been shown to dramatically reduce the GTPase activity of $G_{s\alpha}$ (Graziano and Gilman, 1989b; Masters et al., 1989; Freissmuth and Gilman, 1989). Loss of these interactions would favor the G* state and discrimination between nucleotides bound would be lost.

Mutation proteins that decrease the equilibrium constants, L_{GDP} and L_{GTP}, favor of the ground state, G. The shift in equilibrium to the ground state should result in an increased difference between the apparent affinities of GDP- and GTPγS- bound proteins for adenylyl cyclase and decreased difference between the apparent affinity for GDP- and GTPγS-bound proteins for binding the βγ subunits. Since the ground state is favored, there should be less discrimination between nucleotides for binding the βγ subunits. In contrast, there would be a greater difference between nucleotides in the G* state. The phenotypes of the mutation proteins of class two (R42A, A48H, G225A, R228A, W234F, and I276A) displayed these properties. These residues make several structural contacts in the GTP-bound state. Ile-276 and Trp-234 contact each other. Trp-234 also contacts the non-coupled positions Leu-46 and Leu-272. Gly-225 forms alternate state contacts; Gly-225 interacts with Trp-234 when bound to βγ in the GDP state and forms a main chain H-bond with GTP in the GTP-bound state. R228 stabilizes switch II through ionic interactions with Glu-259 of switch III. Although Ala-48 does not directly contact the guanine nucleotide, Ala-48 is located in the G-1 region (GXXXXGK(S/T)) that contacts the α and β phosphates of the guanine nucleotide. These interactions stabilize the G* state; therefore, loss of these interactions could alter L_{GDP} and L_{GTP} in favoring the ground state.

Generally, mutations in non-coupled positions deviated from the diametric effects on the allosteric binding of adenylyl cyclase and βγ subunits. Proteins with mutations at these non-coupled positions altered the nucleotide sensitivity for the interaction with one binding partner, but five of six

control mutations failed to display significant reciprocal changes in the other partner. Leu-46 was the exception. This result is particularly significant since the control positions are all direct packing neighbors of the statistically coupled positions. Comparison of the GDP- and GTP-bound structures illustrates that some of these control residues display nucleotide-dependent conformational changes similar to the statistically coupled residues. In fact, Ser-205 and Glu-230 contact each other in the GTP-bound state, and this contact is broken in the GDP-bound state (Figure 5-6). This illustrates that state dependent contact and motion is not sufficient for inclusion in the allosteric core. It is difficult to predict from structural analysis why some only select positions have effects. For example, the non-coupled Leu-272 contacts coupled positions Trp-234 and Ile-276 in the GTP-bound state (Figure 5-24). From the structure, it is not clear why mutations at the coupled positions Trp-234 and Ile-276 have this reciprocal property while a mutation at the non-coupled Leu-272 does not. This illustrates the discrete manner in which energy is propagated through proteins.

The non-coupled controls should only exert their effects in the state that they contact the coupled residue. This is illustrated by the interactions of Leu-272. Leu-272 contacts both Ile-276 and Trp-234 in the G* state (Figure 5-24). The replacement of Leu-272 with alanine could result in the loss of contacts with both coupled positions. The loss of volume created with the alanine substitution could result in compensatory repositioning of the coupled Trp-234 and Ile-276 in the G* state. This could result in the destabilization of the G* state and thus increased nucleotide selectivity for adenylyl cyclase. The statistical coupling suggests that Trp-234 and Ile-276 are affected similarly by the perturbations; however, Leu-272 showed no significant coupling. Leu-272 and Trp-234 are structurally coupled but not necessarily energetically coupled. This implies that a thermodynamic mutant cycle of these residues would reveal energetic non-additivity for Trp-234 and Ile-276 while Leu-272 would have additive affects on both position Trp-234 and Ile-276.

Mutations at statistically non-coupled residues are hypothesized to exert their effects through alterations of the neighboring coupled residues. Site-directed mutagenesis is an imprecise tool to assess protein function. Alanine substitutions at a given site are commonly considered a loss-of-function mutation due to the loss of the side chain; however, this is not always the case. A mutation of a given residue can exhibit a gain-of-function phenotype by its effect on the surrounding residues. For example, the replacement of hydrophobic residues in the core of proteins results in the creation of a cavity, altering the structure of the adjacent residues. Alanine mutations in the core of T4 lysozyme resulted in a free energy change that was directly proportional to the size of the cavity created by the mutation (Eriksson et al., 1992). This work illustrated that the protein compensates for a loss of volume created by the alanine mutation by repositioning the adjacent residues to fill the void left by the mutation. If the void is not filled, the protein is less stable. A mutation at any given site will have radiative effects on its immediate neighbors. Through double mutant cycle analysis of the barnase-barstar interface, Schreiber and Fersht showed that the free energy of a mutation radiates 4 to 5 angstroms (Schreiber and Fersht, 1995a). The phenotype of a mutation at a residue that packs against a coupled position could result from compensation for the mutation made. For example, the predicted non-coupled position Leu-46 contacts the coupled Trp-234 in the GTPγS-bound state. Leu-46 is located in the β_1 sheet in the core of $G_{s\alpha}$ and contributes to the hydrophobicity of the pocket that Trp-234 binds in the GTP-bound state (Figure 5-25C). The effect of mutating this residue to alanine could result not from the loss of interaction with Leu-46 but from the adjustment of neighboring side chains to account for the volume lost by replacement of a leucine with an alanine. Therfore, the reciprocal effect of L46A is thought to result from the formation of a cavity in the core of $G_{s\alpha}$ adjacent to many statistically coupled residues. The replacement of Leu-46 with isoleucine displays no change in the nucleotide selectivity for binding adenylyl cyclase or βγ (Figure 5-25A&B). Suggesting that the reciprocal effect of mutating Leu-46 to alanine results from the loss of volume

due to the substitution and the resulting restructuring of the adjacent coupled residues. The mutation of other core non-coupled residues Val-247 and Asn-292 to alanine resulted in insoluble protein when expressed in *E. coli*. The instability of Asn-292 was not unexpected; it is located in the G-4 loop and forms a hydrogen bond with the guanine ring of the bound nucleotide.

Several coupled positions contact the β subunit (Trp-234, Phe-222, and Cys-237). The central contact for βγ is Trp-234. The W234F mutant protein reduced the nucleotide selectivity in the apparent affinity for βγ when GDP-or GTPγS-bound. The coupled C237A mutant protein had a decreased apparent affinity for βγ when bound to either GDP or GTPγS although the fold difference was not statistically significant from that of the wild type protein. Proteins with mutations of Phe-222 and Ile-207 had no effect on the difference in apparent affinity of GDP- and GTPγS-bound proteins binding βγ. Strikingly, the non-coupled I207A mutation altered the apparent affinity of both GDP- and GTPγS-bound $G_{s\alpha}$ without perturbing the fold difference. This shift of the affinity curves was speculated to result from the loss of a non-coupled contact that acted solely through the loss of a non-coupled interaction.

This study has also provided further insight to phenotypes of previously studied mutant proteins. Mutation of glutamine 227 to leucine has been studied extensively and shown to GTPase deficient (Graziano and Gilman, 1989b; Masters et al., 1989). This mutant protein appeared to activate adenylyl cyclase when bound to GDPβS; however, it was determined this resulted from the ATP regenerating system converting contaminating GDP to GTP. Trypsin digestion distinguishes between the active and inactive conformation of $G_{s\alpha}$. The Q227L mutant protein was protected from trypsin digestion in the presence of GTPγS and GTP but not with GDP. The Q227A mutant protein in this study exhibited increased tryptophan fluorescence upon addition of GTPγS indicating the change in conformation of switch II (Trp-234). Therefore, the constitutive activity of Q227A-GDP,

the increased tryptophan fluorescence and lack of trypsin protection presents a paradox. The GDP-bound Q227A-$G_{s\alpha}$ activated adenylyl cyclase as well as the GTP-bound protein suggesting the mutation stabilizes the active conformation. However, the increased tryptophan fluorescence upon binding GTPγS suggests that switch II undergoes a conformation change associated with the transition from the inactive to active state; therefore the GDP-bound Q227A-$G_{s\alpha}$ was not locked in the active conformation. The constitutive activity of the GDP-bound Q227A-$G_{s\alpha}$ would not display an effect in a cell since the mutant protein binds the βγ subunits normally thus preventing the interaction with adenylyl cyclase.

$G_{s\alpha}$ containing a mutation of glycine-225 to threonine has been reported to function as a dominant negative (Osawa and Johnson, 1991). When transfected into COS cells, the mutant protein fails to activate adenylyl cyclase upon isoproterenol treatment; however, no further biochemical characterization of this mutant protein has been published. Mutation of the analogous position in Ras, A59T, creates a constitutively active protein illustrated by strong activity in cell transformation assays due to a reduced GTPase activity (Gibbs et al., 1984). The G225A-$G_{s\alpha}$ mutant protein exhibits increased tryptophan fluorescence upon binding GTPγS similar to wild type $G_{s\alpha}$. Therefore, G225A undergoes the conformational change associated with formation of the active state. However, the G225A mutant protein has a complete loss of discrimination between GDP- and GTPγS-bound binding of βγ. The GTPγS-bound G225A-$G_{s\alpha}$ binds βγ as well as the GDP-bound protein. This loss of a nucleotide-dependent binding for the βγ subunits explains the dominant negative phenotype; the G225A mutant never releases βγ and is thus prevented from interacting with adenylyl cyclase.

Other of the statistically coupled positions predicted by the method described in this study have previously been shown to have important roles in the alpha subunit. $G_{s\alpha}$ with a mutation of Arg-201, the site of ADP-ribosylation by cholera toxin, to alanine was shown to decrease GTPase activity

and to decrease affinity for the βγ subunits (Freissmuth and Gilman, 1989). Arg-201 stabilizes the transition state of GTP hydrolysis illustrated by the coordination of AlF_4^- and the β-phosphate in the $G_{i\alpha1}$-GDP:AlF_4^- structure (Coleman et al., 1994). The structure of $G_{i\alpha1}\beta_1\gamma_2$ heterotrimer illustrated that the interaction of the Arg-201 and another coupled residue Glu-50 forms a "seatbelt" that prevents the dissociation of GDP (R187 and E43 in $G_{i\alpha1}$) (Wall et al., 1995). Cys-365 buttresses the site for guanine ring recognition; mutation of this position in $G_{o\alpha}$ to alanine (C325) decreased the protein's affinity for GDP 10-fold without altering the affinity for GTPγS (Thomas et al., 1993). Arg-228 of switch II contacts residues in switch III in the GTP-bound state; the mutation of this residue in $G_{t\alpha}$ (R201) to alanine resulted in a protein with decreased affinity for the phosphodiesterase γ-subunit and decreased trypsin protection in the presence of GDP:AlF_4^- (Natochin et al., 1998). Strikingly, this statistical method identified many positions in $G_{s\alpha}$ that have not been previously characterized by mutational analysis or predicted to be required for the nucleotide-dependent binding differences for adenylyl cyclase and βγ.

Figure 5-20 provides a summary of the mutant analysis by plotting the nucleotide sensitivity for adenylyl cyclase versus that of βγ. The data for the mutants at coupled positions were fit to an inverse relationship (y = 1/x). These data demonstrate that the full set of coupled positions predicted by the statistical coupling analysis were consistent with a reciprocal relationship of the nucleotide sensitivities for adenylyl cyclase and βγ (r = 0.85). These results support the hypothesis that mutations of these sites cause varying degrees of perturbation to the allosteric mechanism controlling the equilibrium between G and G*. Mutations at control positions deviated from the reciprocal relationship (r = 0.42). Although some of the mutations at the control non-coupled positions caused significant change in the nucleotide sensitivity in the interaction with one binding partner, five out of six of the control non-coupled positions failed to exhibit the reciprocal change for the other partner.

These results are particularly significant since all the control residues form contacts with the coupled residues. This mutant analysis illustrates that the residues identified by the statistical coupling analysis specifically participate in the mechanism of G protein allostery.

The statistical coupling analysis used in this study relies on the long-term evolutionary record of the G protein family. Thus, the network of co-evolving residues may be a canonical structural feature of all members of the G protein family. Therefore, the structural features of this network, the nucleotide-dependent physical connectivity, should be conserved in distantly related famlily members. The statistically coupled residues have been mapped on the structure inactive (GDP-bound) and active (GTPγS or Gpp(NH)p-bound) members of four distant members of the G protein family (Figure 5-26). Similar to the heterotrimeric G protein alpha subunits, the network connectivity is intact in the active-state structures, but this network is broken in the inactive-state structures. The similarity of the network of the statistically coupled residues occurs despite substantial divergence of other structural features and very different functions of the members of the G protein family. The network of residues identified in this study is proposed to comprise the core allosteric mechanism conferring nucleotide-dependent switching of the G protein family, and the specific features of different G protein family members were built on this core.

Table 5-1.

Activation of adenylyl cyclase by GTPγS- or GDP-bound $G_{s\alpha}$

Mutant	EC_{50} $G_{s\alpha}$-GTPγS [a]	EC_{50} $G_{s\alpha}$-GDP [a]	Fold-AC[b]	*P* value [c]
WT	0.23 ± 0.03	1.3 ± 0.39	5.7 ± 1.85	
R42A	0.3 ± 0.007	4.3 ± 0.42	14.3 ± 1.44	0.0001
A48H	0.19 ± 0.042	2.6 ± 1.66	13.7 ± 9.25	0.0002
E50A	0.2 ± 0.007	0.68 ± 0.078	3.4 ± 0.41	0.096
R201A	0.19 ± 0.031	0.53 ± 0.057	2.8 ± 0.55	0.039
F222A	0.4 ± 0.078	2.2 ± 0.35	5.5 ± 1.38	0.88
V224A	0.16 ± 0.035	1.7 ± 0.51	10.6 ± 3.94	0.002
G225A	0.4 ± 0.106	4.4 ± 0.35	11 ± 3	0.0008
Q227A	0.26 ± 0.015	0.38 ± 0.029	1.5 ± 0.14	0.0006
R228A	0.3 ± 0.014	6.9 ± 0.57	23 ± 2.18	0.0001
W234F	1.2 ± 0.07	21.5 ± 4.95	17.9 ± 4.26	0.0001
C237A	0.29 ± 0.042	3.5 ± 0.71	12.1 ± 3.01	0.0001
I276A	0.65 ± 0.127	20 ± 9.9	30.8 ± 16.37	0.0001
C365A	0.2 ± 0.007	0.56 ± 0.021	2.8 ± 0.14	0.039
L46A	0.28 ± 0	5.3 ± 0.42	18.9 ± 1.5	0.0001
S205A	0.21 ± 0.025	1.7 ± 0.53	8.1 ± 2.7	0.095
I207A	0.27 ± 0.028	4.4 ± 0.42	16.3 ± 2.3	0.0001
D229A	0.14 ± 0.049	0.71 ± 0.085	5.1 ± 1.88	0.66
E230A	0.21 ± 0.026	2.2 ± 0.92	10.5 ± 4.57	0.003
L272A	0.76 ± 0.092	10.1 ± 0.57	13.3 ± 1.77	0.0001

a. The EC_{50} values for individual experiments were calculated by fitting the data to a one site binding model as described in the Experimental Methods. The mean values for all experiments of the sample are reported above with the associated error. The wild type is represented by the mean and standard deviation of 26 experiments. The mutants are represented as the mean and range of 2 to 3 experiments.

b. The fold-AC was calculated by dividing the EC_{50} of $G_{s\alpha}$-GDP by the EC_{50} of $G_{s\alpha}$-GTPγS and propagating the associated error.

c. The fold-AC of the mutants was compared to the fold potency of the wild type using a two-tailed Student's t test. The significance is reported as the *P* value.

Table 5-2.

Binding of βγ by GDP- and GTPγS-bound $G_{s\alpha}$

Mutant	IC_{50} $G_{s\alpha}$-GDP[a]	IC_{50} $G_{s\alpha}$-GTPγS[a]	Fold-βγ[b]	*P* value[c]
WT	0.7 ± 0.066	13.3 ± 3.29	19 ± 5.03	
R42A	3.4 ± 0.39	33 ± 10.1	9.6 ± 3.13	0.03
A48H	1.3 ± 0.12	1.7 ± 0.11	1.3 ± 0.14	0.0007
E50A	0.41 ± 0.071	8.8 ± 0.07	21.5 ± 3.72	0.53
R201A	2.8 ± 0.07	126 ± 15.6	45 ± 5.69	0.0001
F222A	0.87 ± 0.17	20.2 ± 2.47	23.2 ± 5.35	0.31
V224A	1.2 ± 0.07	17.5 ± 0.71	14.6 ± 1.05	0.26
G225A	0.82 ± 0.064	0.9 ± 0.06	1.1 ± 0.11	0.0007
Q227A	0.98 ± 0.021	49.5 ± 6.36	50.5 ± 6.58	0.0001
R228A	1.6 ± 0.07	8.8 ± 0.14	5.5 ± 0.26	0.004
W234F	5.4 ± 0.64	18 ± 5.7	3.3 ± 1.12	0.002
C237A	2.9 ± 0.42	33 ± 5.7	11.4 ± 2.55	0.07
I276A	4.2 ± 0.35	29.5 ± 0.71	7 ± 0.61	0.009
C365A	1.2 ± 0.21	29 ± 11.8	24.8 ± 10.9	0.23
L46A	4.2 ± 0.41	36 ± 9.2	8.5 ± 2.33	0.018
S205A	0.78 ± 0.163	12.5 ± 0.71	16 ± 3.47	0.45
I207A	2.8 ± 0.14	47 ± 7.1	16.8 ± 2.66	0.57
D229A	0.52 ± 0.042	19 ± 4.24	36.5 ± 8.67	0.002
E230A	0.36 ± 0.085	4 ± 1.06	11.1 ± 3.94	0.066
L272A	1.4 ± 0.07	24.5 ± 0.71	17.5 ± 1.01	0.69

a. The IC_{50} values for individual experiments were calculated by fitting the data to a one site binding model as described in the Experimental Methods. The mean values for all experiments of the sample are reported above with the associated error. The wild type is represented by the mean and standard deviation of 10 experiments. The mutants are represented as the mean and range of 2 to 3 experiments.

b. The fold-βγ was calculated by dividing the IC_{50} of $G_{s\alpha}$-GTPγS by the IC_{50} of $G_{s\alpha}$-GDP and propagating the associated error.

c. The fold-βγ of the mutants was compared to the fold-βγ of the wild type using a two-tailed Student's t test. The significance is reported as the *P* value.

TABLE 5-3.

Summary of $G_{s\alpha}$ Mutant Analysis

Mutant	Yield (mg) [a]	Coupled [b]	Mutant Class	GTPγS Binding (min^{-1}) [c]	Tryptophan Fluorescence (min^{-1}) [d]	Fold-AC [e]	Fold-βγ [f]
WT	190	----		0.69	0.10	5.7 ± 1.85	19 ± 5.03
R42A	3	C	2	ND [h]	ND [h]	14.3 ± 1.44	9.6 ± 3.13
A48H	2	C	2	0.30	ND [h]	13.7 ± 9.25	1.3 ± 0.14
E50A	13	C	4	0.31	0.97	3.4 ± 0.41	21.5 ± 3.72
R201A	18	C	1	0.12	ND [h]	2.8 ± 0.55	45 ± 5.69
F222A	25	C	4	ND [h]	0.19	5.5 ± 1.38	23.2 ± 5.35
V224A	29	C	3	0.071	0.12	10.6 ± 3.94	14.6 ± 1.05
G225A	14	C	2	ND [h]	0.11	11 ± 3	1.1 ± 0.11
Q227A	15	C	1	0.097	0.14	1.5 ± 0.14	50.5 ± 6.58
R228A	42	C	2	ND [h]	ND [h]	23 ± 2.18	5.5 ± 0.26
W234F	155	C	2	0.10	ND [h]	17.9 ± 4.26	3.3 ± 1.12
C237A	160	C	3	0.067	0.11	12.1 ± 3.01	11.4 ± 2.55
I276A	9	C	2	0.033	0	30.8 ± 16.37	7 ± 0.61
C365A	3	C	3	ND [h]	ND [h]	2.8 ± 0.14	24.8 ± 10.9
F376A	----[g]	C	5	----	----	----	----
L46A	10	NC		0.047	0.027	18.9 ± 1.5	8.5 ± 2.33
S205A	26	NC		0.061	0.11	8.1 ± 2.7	16 ± 3.47
I207A	26	NC		ND [h]	ND [h]	16.3 ± 2.3	16.8 ± 2.66
D229A	13	NC		0.075	0.17	5.1 ± 1.88	36.5 ± 8.67
E230A	18	NC		ND [h]	ND [h]	10.5 ± 4.57	11.1 ± 3.94
V247A	-----[g]	NC		----	----	----	----
L272A	21	NC		ND [h]	ND [h]	13.3 ± 1.77	17.5 ± 1.01
N292A	-----[g]	NC		----	----	----	----

a. The yield of protein purified from 12 L of *E.coli.*
b. C= statistically coupled mutant, NC = non-coupled control
c. The rate of GDP release calculated by GTPγS binding. The rate was determined as described in the Experimental Methods.
d. Rate of increased intrinsic tryptophan fluorescence upon GTPγS binding. The rate was determined as described in the Experimental Methods.
e. The fold-AC for adenylyl cyclase activation was calculated by dividing the EC_{50} of $G_{s\alpha}$-GDP by the EC_{50} of $G_{s\alpha}$-GTPγS and propagating the associated error (see Table 5-1).
f. The fold-βγ for βγ binding was calculated by dividing the IC_{50} of $G_{s\alpha}$-GTPγS by the IC_{50} of $G_{s\alpha}$-GDP and propagating the associated error (see Table 5-2).
g. Not soluble when expressed in *E. coli.*
h. Not Determined (ND)

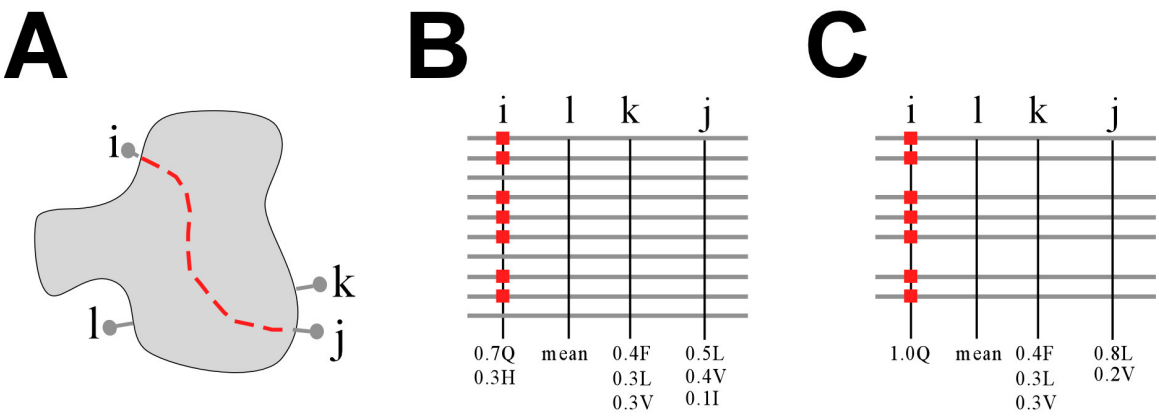


Figure 5-1. Schematic for extracting coupling information from multiple sequence alignment. **A.** A hypothetical protein A is depicted containing amino acids i, j, k, and l. The position i is coupled to position j but not coupled to positions k or l. Position k contains a frequency of amino acids that deviate from the mean frequency found in all proteins. Position l contains amino acids at the same frequency as found in all proteins. **B. MSA of protein family.** A hypothetical multiple sequence alignment of the protein family containing protein A. The horizontal lines represent the primary sequence of members of the protein family from amino- to carboxy-terminus. Above the sequences, the position of the amino acids labeled in **A** are shown. The vertical lines show the position of the labeled amino acids in all sequences and below these lines the frequency of amino acids at these positions is shown. The red boxes at position i indicate the sequences that contain a glutamine. The mean frequency indicated for position l represents the mean frequency of amino acids in all proteins in the database. **C. Subalignment of B.** A perturbation is made at site i by extracting the sequences that contain a glutamine at position i. The change in frequency of amino acids at position j upon perturbation of position i illustrates coupling. The frequency at positions k and l does not change upon perturbation of position i. (This figure was adapted from Suel, G. et al(Suel et al., 2003)).

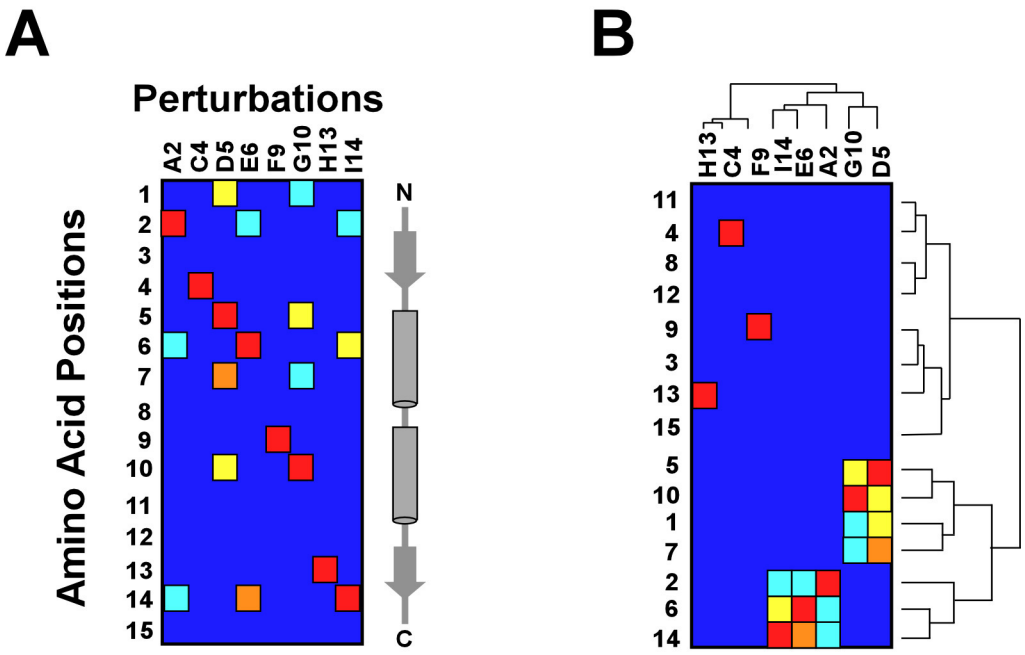


Figure 5-2. Mock clustering experiment. A. Coupling matrix. The statistical coupling ($\Delta\Delta G^{stat}$) of each amino acid position from N-terminus (top) to C-terminus (bottom) for the perturbation in each column is illustrated as a gradient from blue (low coupling) to red (high coupling). The position in each column that represents the site of perturbation is colored red. For example, the third column (D5) represents the statistical coupling of all positions in the MSA (rows 1-15) calculated by generating a sub-alignment that contained aspartate at position 5. Position 5 is the site of perturbation for D5 and is colored red. Perturbation D5 exhibits coupling to positions 1 and 10 (yellow) and increased coupling to position 7 (orange) **B. Matrix Clustering.** Two-dimensional clustering analysis was used to extract patterns from the coupling matrix (**A**). First, the perturbations (columns) are sorted by the pattern of coupling to positions (rows). Then, the positions (rows) are sorted by the pattern of coupling to the perturbations (columns). Most residues show low statistical coupling (blue); however, two distinct groups distinguished themselves from the rest of the matrix. The first cluster contains perturbations A2, E6, and I14 that are coupled to positions 2, 6, and 14. The second cluster contained perturbations D5 and G10 that are coupled to positions 1, 5, 7, and 10. If the network of coupled residues responsible for energy distribution is a conserved property of the protein family, perturbations to multiple positions in the network should identify each other and contain the same coupling pattern. In the example in Figure 5-2B, the perturbation I14 is coupled to position 6, and perturbation E6 is coupled to position 14. This illustrates self-consistency in the coupling pattern.

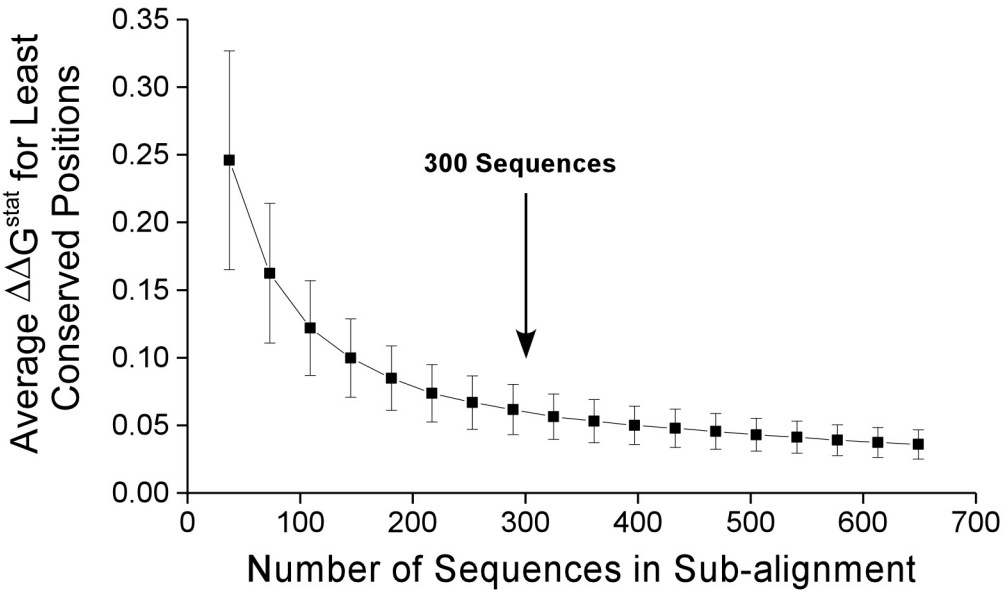


Figure 5-3. Mean $\Delta\Delta G^{stat}$ for least conserved positions following random elimination of sequences from the MSA. The coupling of the five least conserved positions in the G protein alignment was calculated as sequences were randomly eliminated from the alignment. The asymptotic nature of the curve indicates the alignment is in statistical equilibrium. The coupling of the least conserved positions increases as the number of sequences in the sub-alignment no longer represents the whole alignment. Perturbations made to the alignment must contain over 300 sequences to maintain the statistical equilibrium.

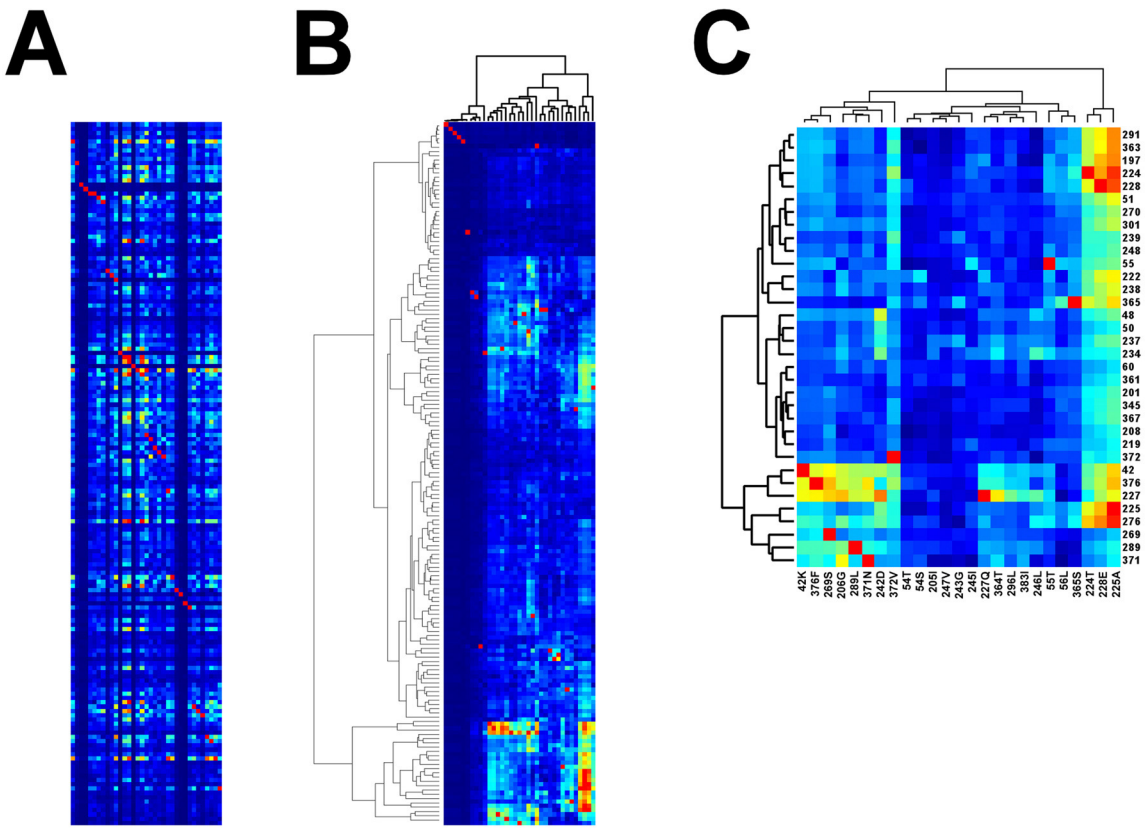


Figure 5-4. Statistical coupling analysis of the GTP binding protein family. *Coupling matrix* (**A**). The statistical coupling data for perturbations of the G protein alignment are represented as a matrix. The rows represent the primary sequence of $G_{s\alpha}$ from amino (top) to carboxyl-terminus (bottom). The columns are the individual perturbation experiments. The $\Delta\Delta G^{stat}$ value for each position is displayed coliormetrically in a linear scale from blue (0.0kT*) to red (5.7kT*). The sites of perturbations are illustrated in red. *Clustered Matrix* (**B**)**.** The two-dimensional clustering of the matrix in **A** reveals relationships between the patterns of perturbations and positions. *Iterative Matrix* (**B**)**.** The positions with the highest coupling values from the matrix in **B** were selected, extracted, and re-clustering**.**

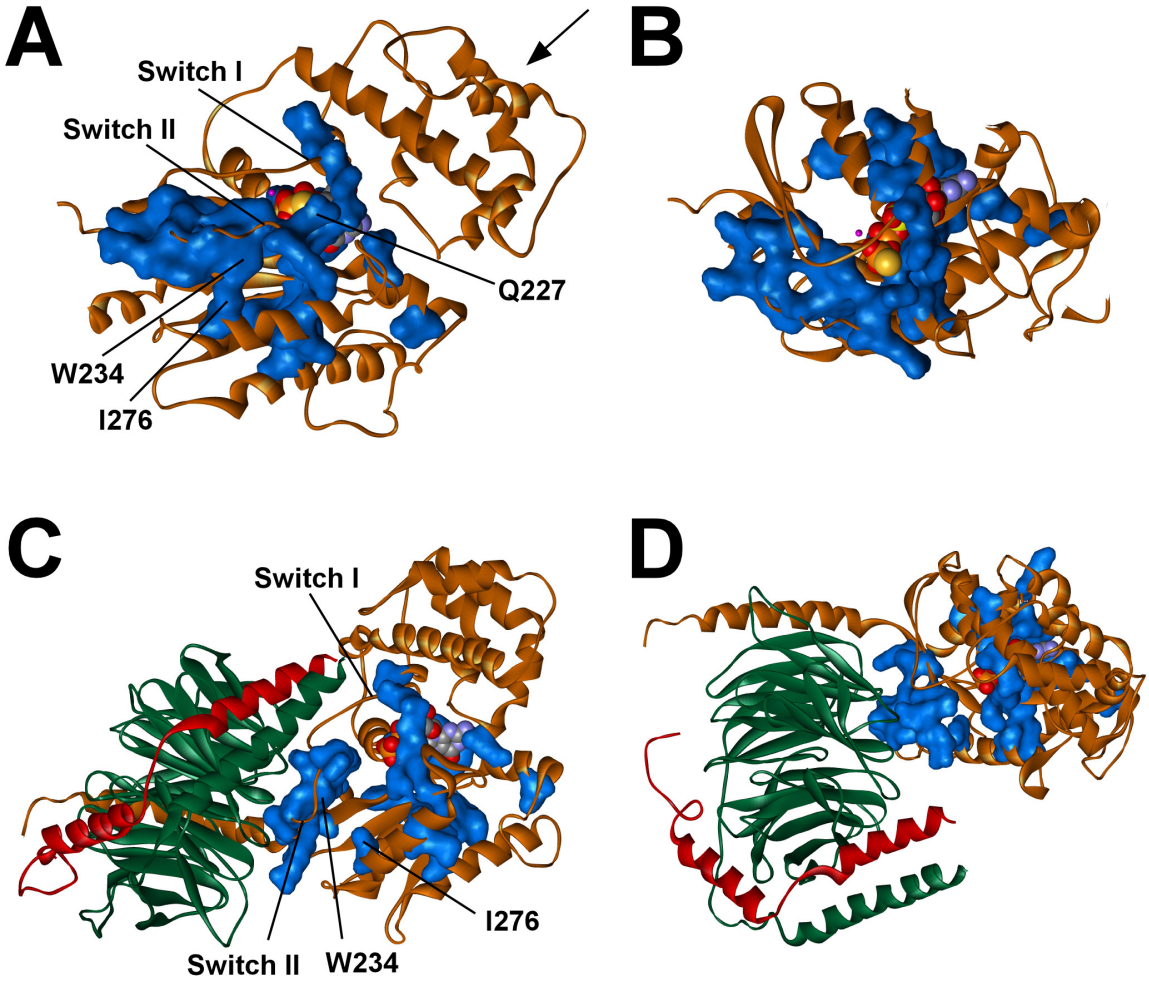


Figure 5-5. Mapping statistically coupled positions on the G alpha subunit. *Active conformation* (**A** and **B**). The statistically coupled residues are represented as blue van der Waals surfaces on a tan ribbon structure of $G_{s\alpha}$ bound to GTPγS (Protein Databank entry 1AZT). The helical domain has been removed in B and the protein viewed from the position of the arrow in A. *Inactive conformation* (**C** and **D**). The homologous residues were mapped on the GDP-bound $G_{i\alpha}$ in complex with the $\beta_1\gamma_2$ subunits, green and red respectively (Protein Databank entry 1GP2). The heterotrimer is rotated out of the page 90° in panel D.

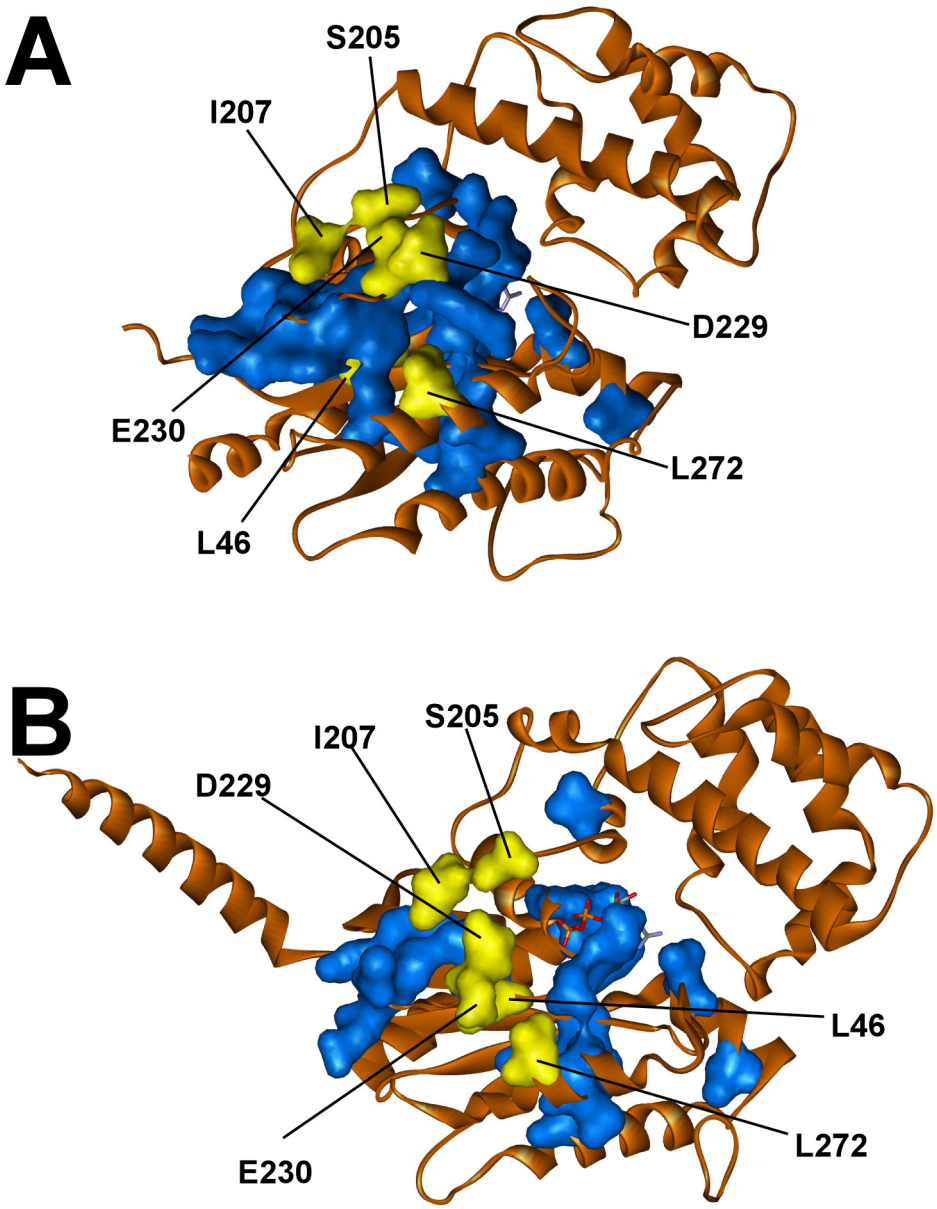


Figure 5-6. Illustrating non-coupled controls on the G alpha subunit. *Active conformation* (**A**). The non-coupled controls are represented as yellow van der Waals surfaces and statistically coupled residues are represented as blue van der Waals surfaces on a tan ribbon structure of $G_{s\alpha}$ bound to GTPγS (Protein Databank entry 1AZT). *Inactive conformation* (**B**). Homologous residues mapped on GDP-bound $G_{i\alpha}$ in complex with the $\beta_1\gamma_2$ subunits removed to facilitate viewing (Protein Databank entry 1GP2). Labeling is according to the $G_{s\alpha}$ sequence.

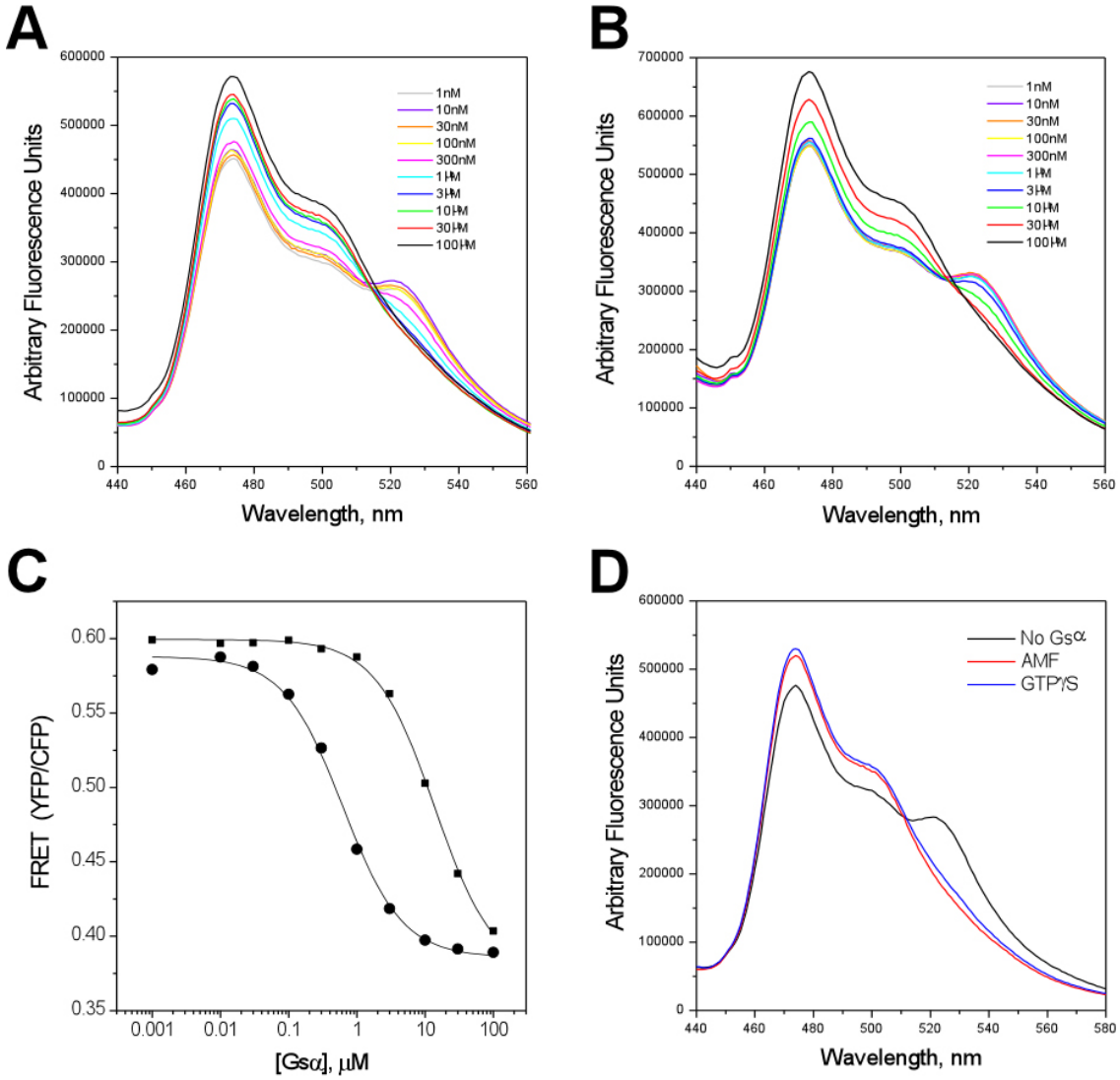


Figure 5-7. βγ affinity assay. Binding of $G_{i\alpha1}$-ECFP and $\beta_1\gamma_2$-EYFP chimeric fusion proteins has been monitored by the fluorescence resonance energy transfer (FRET) from ECFP to EYFP (Scott Gibson, unpublished results). The disappearance of FRET was used to monitor competition between $G_{s\alpha}$ and $G_{i\alpha1}$-ECFP for binding $\beta_1\gamma_2$-EYFP. GDP- and GTPγS-bound $G_{s\alpha}$, H6-$G_{i\alpha1}$-ECFP, and $\beta_1\gamma_2$-EYFP were diluted in 50 mM NaHEPES pH 8, 1 mM EDTA, 5 mM DTT, 100 mM NaCl, 2 mM $MgSO_4$, 0.1% Lubrol ($C_{12}E_{10}$), and 200 μM GDP. Increasing concentrations of $G_{s\alpha}$-GDP (**A**) or $G_{s\alpha}$-GTPγS (**B**) were added to H6-$G_{i\alpha1}$-ECFP and $\beta_1\gamma_2$-EYFP (200 nM each) and incubated at 18 °C for 16 hours. The fluorescence spectra were obtained using a Photon Technology International fluorometer. The sample was excited at 410 nm and the emission was recorded from 461 to 535 nm at a 3 nm slit width at 20 °C. The ratio of the fluorescence intensity of YFP at 522 nm to CFP at 474 nm was plotted versus the $G_{s\alpha}$ concentration (**C**). The data were fit to the following equation using non-linear least squares method in Microcal Origin: $y = (B_{max}*([IC_{50}/[G_{s\alpha}]))/(1+([IC_{50}/[G_{s\alpha}]) + y0$ where y0 is the offset of the lower asymptote and B_{max} is the upper asymptote. **D.** Fluorescence spectra of 200nM H6-$G_{i\alpha1}$-ECFP and $\beta_1\gamma_2$-EYFP alone, with AMF (10mM $MgCl_2$, 10mM NaF, and 50μM $AlCl_3$) or with 400μM GTPγS.

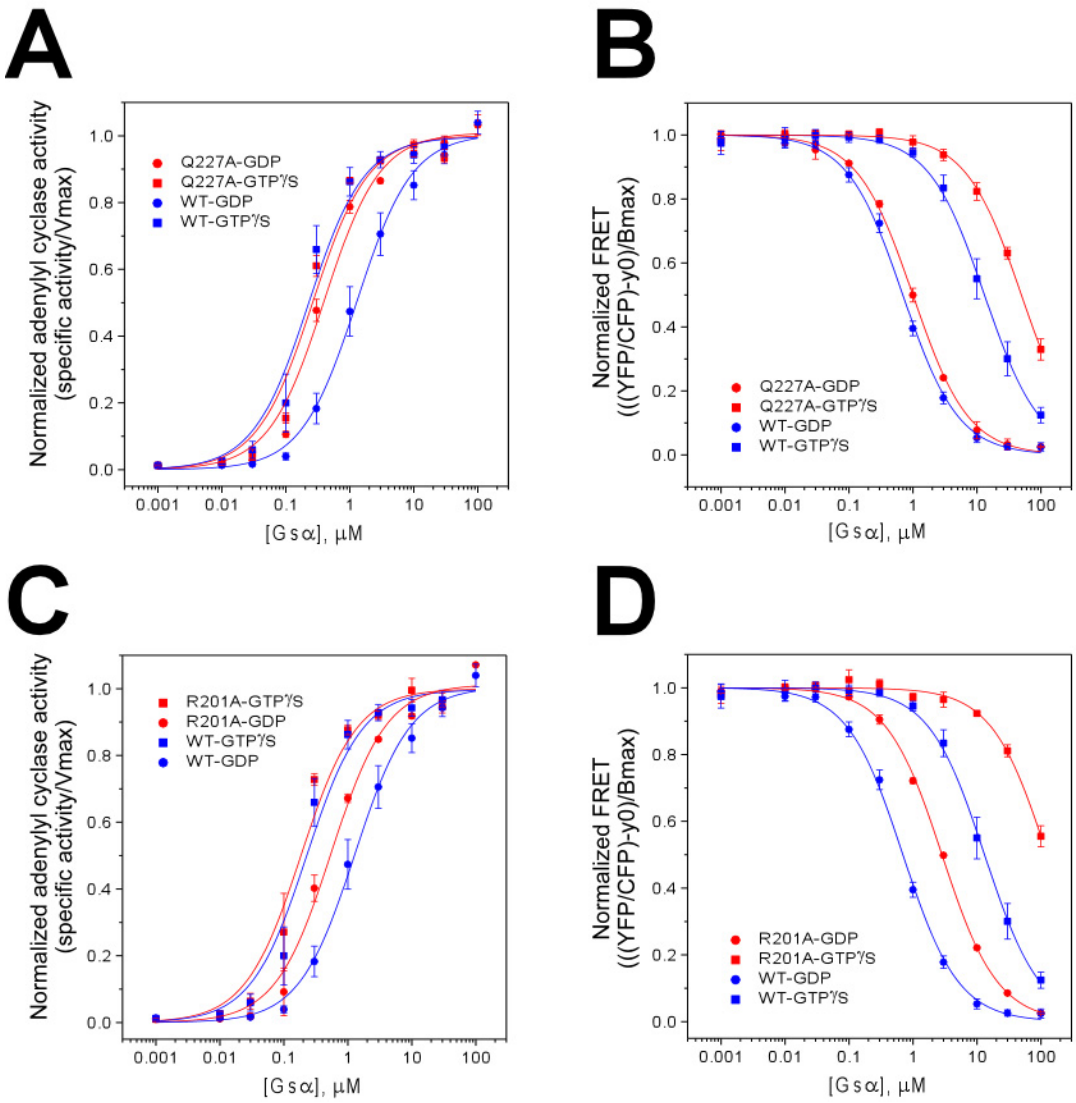


Figure 5-8. Characterization of Q227A and R201A (class 1). *Adenylyl cyclase.* Activation of adenylyl cyclase was assayed with increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 1 nM IIC_2 and 10 μM VC_1. The 50 μl reaction was initiated with the addition of 1 mM [^{32}P]ATP and proceeded for 10 min. The data were fit using a non-linear least squares fit to the equation $y=(V_{max}*x)/(EC_{50}+x)$. The Q227A (**A**) and R201A (**C**) mutant proteins are the mean and range of two experiments performed in duplicate. The wild type curve is the mean and standard deviation of 26 independent experiments performed in duplicate. The specific activities were normalized by dividing by the Vmax. *Binding of βγ.* Increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 200nM of H6-Giα-CFP and β1γ2-YFP. The samples were incubated at 18 ^{o}C for 16 hours. The fluorescence was recorded at 20 ^{o}C from 440 to 560 nm following excitation at 410 nm. The FRET is represented as the ratio of the fluorescence intensity of YFP at 522nm to that of CFP a 474nm. The Q227A (**B**) and R201A (**D**) mutant proteins are the mean and range of two experiments. The wild type curve is the mean and standard deviation of 10 independent experiments.

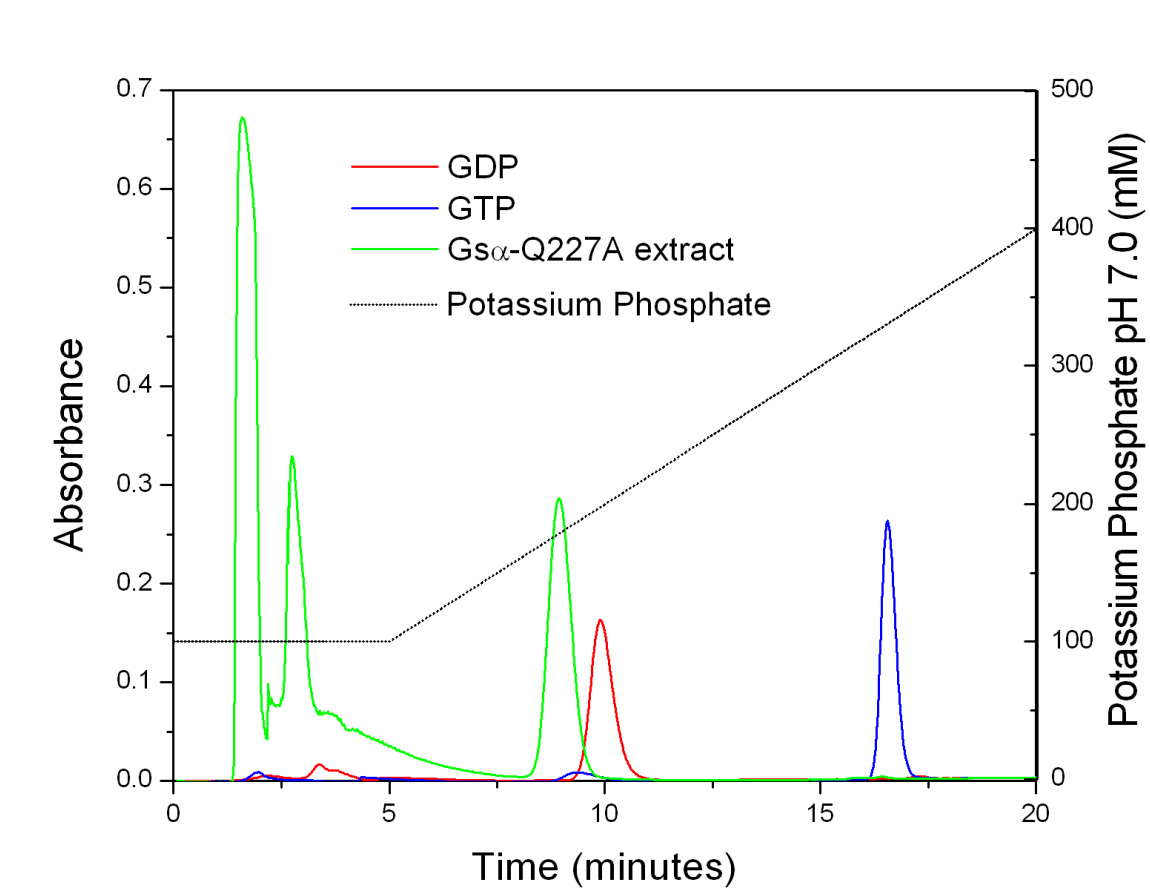


Figure 5-9. Determination of nucleotide bound to Q227A-$G_{s\alpha}$. The identity of the nucleotide bound to the Q227A-$G_{s\alpha}$ purified from *E. coli* was determined by high-pressure liquid chromatography. The Q227A-$G_{s\alpha}$ mutant protein purified from *E. coli* and stored in 20 mM NaHEPES pH 8, 1 mM EDTA, 2 mM DTT, and 50 mM NaCl was diluted to 30µM protein with 100mM potassium phosphate pH 7.0. The protein was heated at 100°C for 30 minutes. The guanine nucleotide was separated from denatured protein with three phenol:chloroform extractions followed by three chloroform extractions. 200 µl of the resulting solution was applied to a Synchropak AX300 100 x 4.6 mm anion exchange column (Micra Scientific), eluted at 1 ml/minute with a gradient of potassium phosphate pH 7.0 from 100 mM to 400 mM. The absorbance at 253 nm was recorded. GDP and GTP, 200 µl of 15 µM in 100 mM potassium phosphate pH 7.0, were serparately applied and eluted with the same potassium phosphate pH 7.0 gradient.

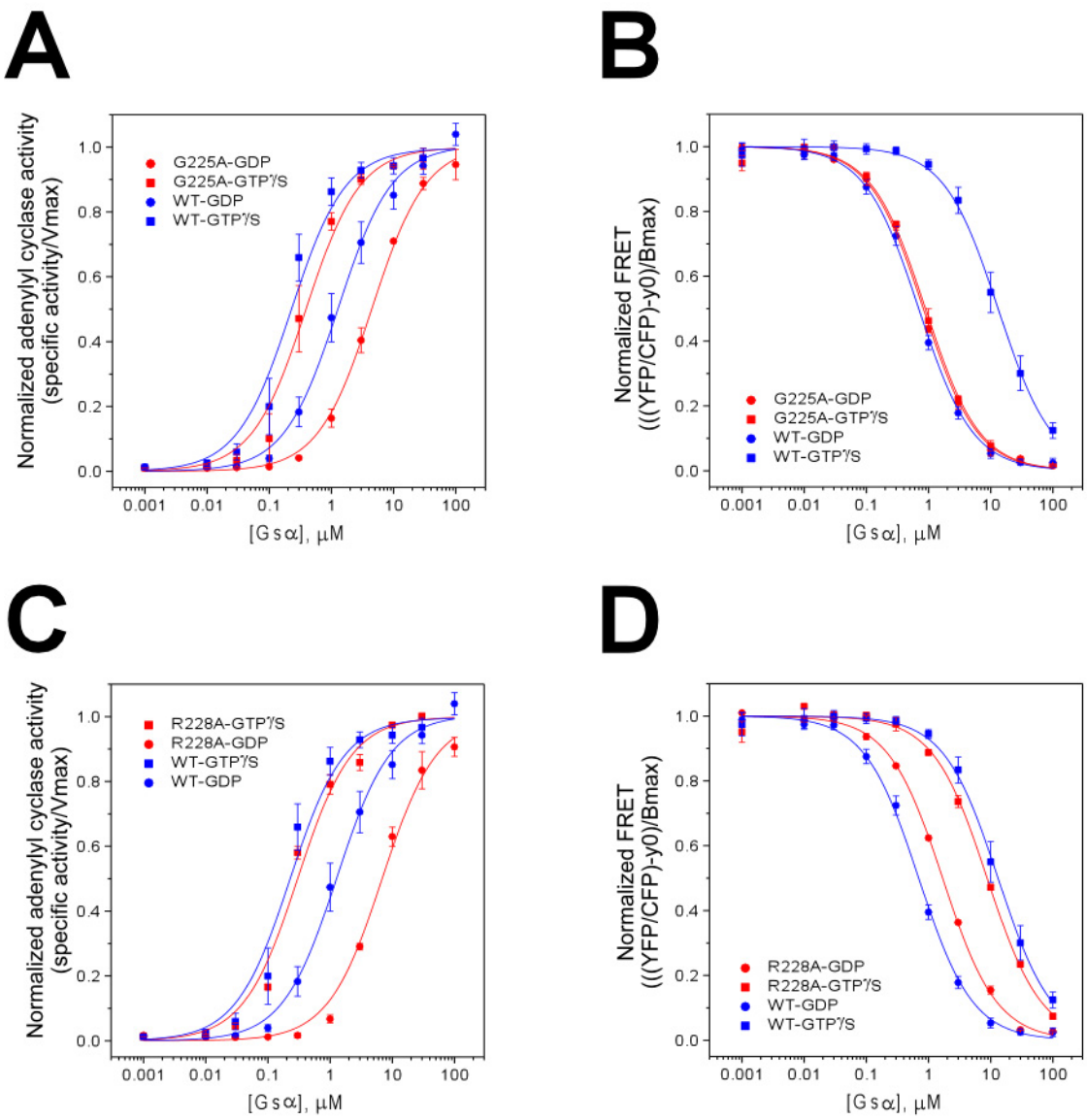


Figure 5-10. Characterization of G225A and R228A (class 2). *Adenylyl cyclase.* Activation of adenylyl cyclase was assayed with increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 1 nM IIC_2 and 10 μM VC_1. The 50 μl reaction was initiated with the addition of 1 mM [^{32}P]ATP and proceeded for 10 min. The data were fit using a non-linear least squares fit to the equation $y=(V_{max}*x)/(EC_{50} + x)$. The G225A (**A**) and R228A (**C**) mutant proteins are the mean and range of two experiments performed in duplicate. The wild type curve is the mean and standard deviation of 26 independent experiments performed in duplicate. The specific activities were normalized by dividing by the Vmax. *Binding of βγ.* Increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 200nM of H6-Giα-CFP and β1γ2-YFP. The samples were incubated at 18 °C for 16 hours. The fluorescence was recorded at 20 °C from 440 to 560 nm following excitation at 410 nm. The FRET is represented as the ratio of the fluorescence intensity of YFP at 522nm to that of CFP a 474nm. The G225A (**B**) and R228A (**D**) mutant proteins are the mean and range of two experiments. The wild type curve is the mean and standard deviation of 10 independent experiments.

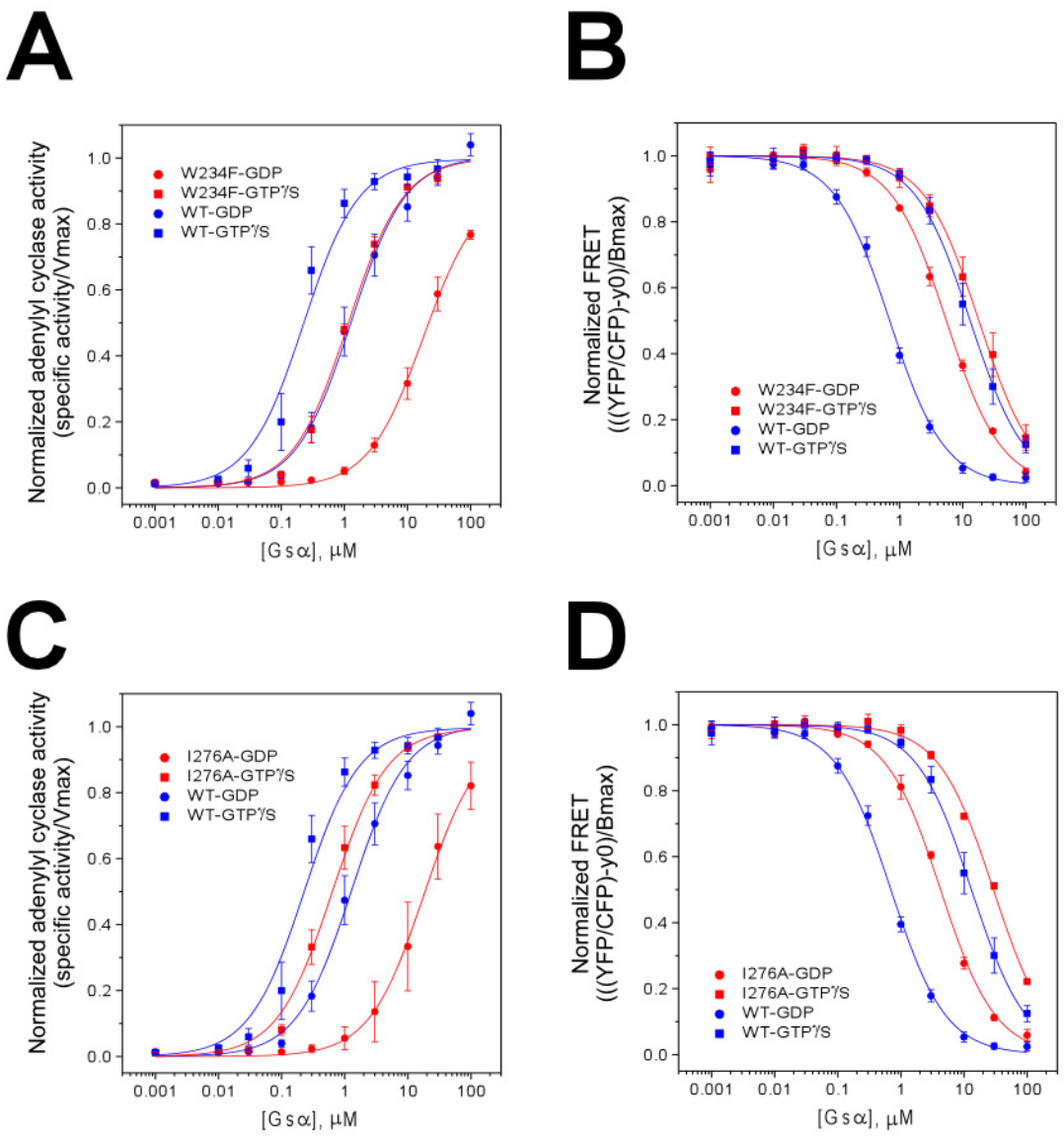


Figure 5-11. Characterization of W234F and I276A (class 2). *Adenylyl cyclase.* Activation of adenylyl cyclase was assayed with increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 1 nM IIC_2 and 10 μM VC_1. The 50 μl reaction was initiated with the addition of 1 mM [^{32}P]ATP and proceeded for 10 min. The data were fit using a non-linear least squares fit to the equation $y=(V_{max}*x)/(EC_{50}+x)$. The W234F (**A**) and I276A (**C**) mutant proteins are the mean and range of two experiments performed in duplicate. The wild type curve is the mean and standard deviation of 26 independent experiments performed in duplicate. The specific activities were normalized by dividing by the Vmax. *Binding of βγ.* Increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 200nM of H6-Giα-CFP and β1γ2-YFP. The samples were incubated at 18 °C for 16 hours. The fluorescence was recorded at 20 °C from 440 to 560 nm following excitation at 410 nm. The FRET is represented as the ratio of the fluorescence intensity of YFP at 522nm to that of CFP a 474nm. The W234F (**B**) and I276A (**D**) mutant proteins are the mean and range of two experiments. The wild type curve is the mean and standard deviation of 10 independent experiments.

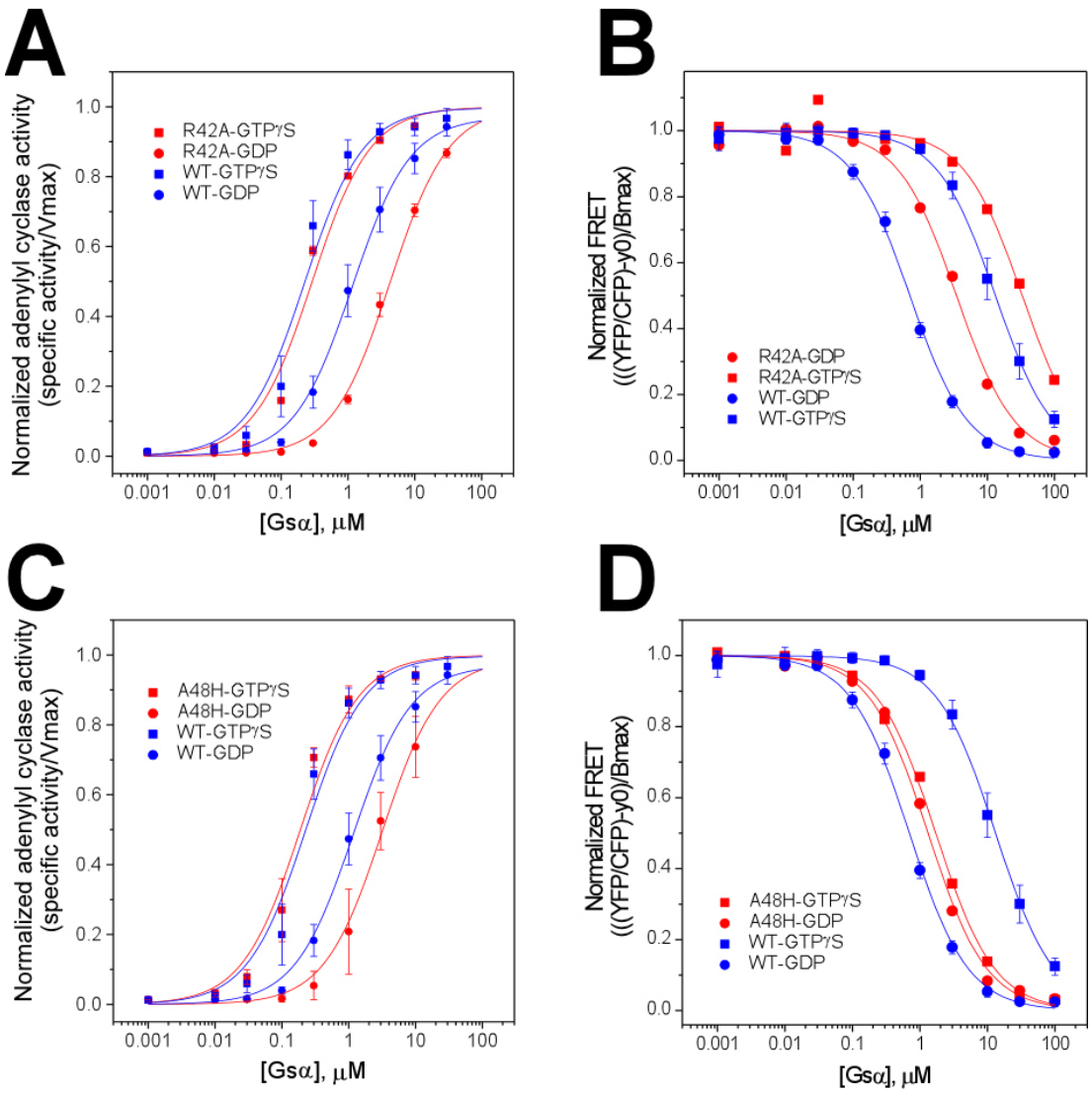


Figure 5-12. Characterization of R42A and A48H (class 2). *Adenylyl cyclase.* Activation of adenylyl cyclase was assayed with increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 1 nM IIC_2 and 10 μM VC_1. The 50 μl reaction was initiated with the addition of 1 mM [^{32}P]ATP and proceeded for 10 min. The data were fit using a non-linear least squares fit to the equation $y=(V_{max}*x)/(EC_{50} + x)$. The R42A (**A**) and A48H (**C**) mutant proteins are the mean and range of two experiments performed in duplicate. The wild type curve is the mean and standard deviation of 26 independent experiments performed in duplicate. The specific activities were normalized by dividing by the Vmax. *Binding of βγ.* Increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 200nM of H6-Giα-CFP and β1γ2-YFP. The samples were incubated at 18 ^{o}C for 16 hours. The fluorescence was recorded at 20 ^{o}C from 440 to 560 nm following excitation at 410 nm. The FRET is represented as the ratio of the fluorescence intensity of YFP at 522nm to that of CFP a 474nm. The R42A (**B**) and A48H (**D**) mutant proteins are the result of one experiment due to lack of protein. The wild type curve is the mean and standard deviation of 10 independent experiments.

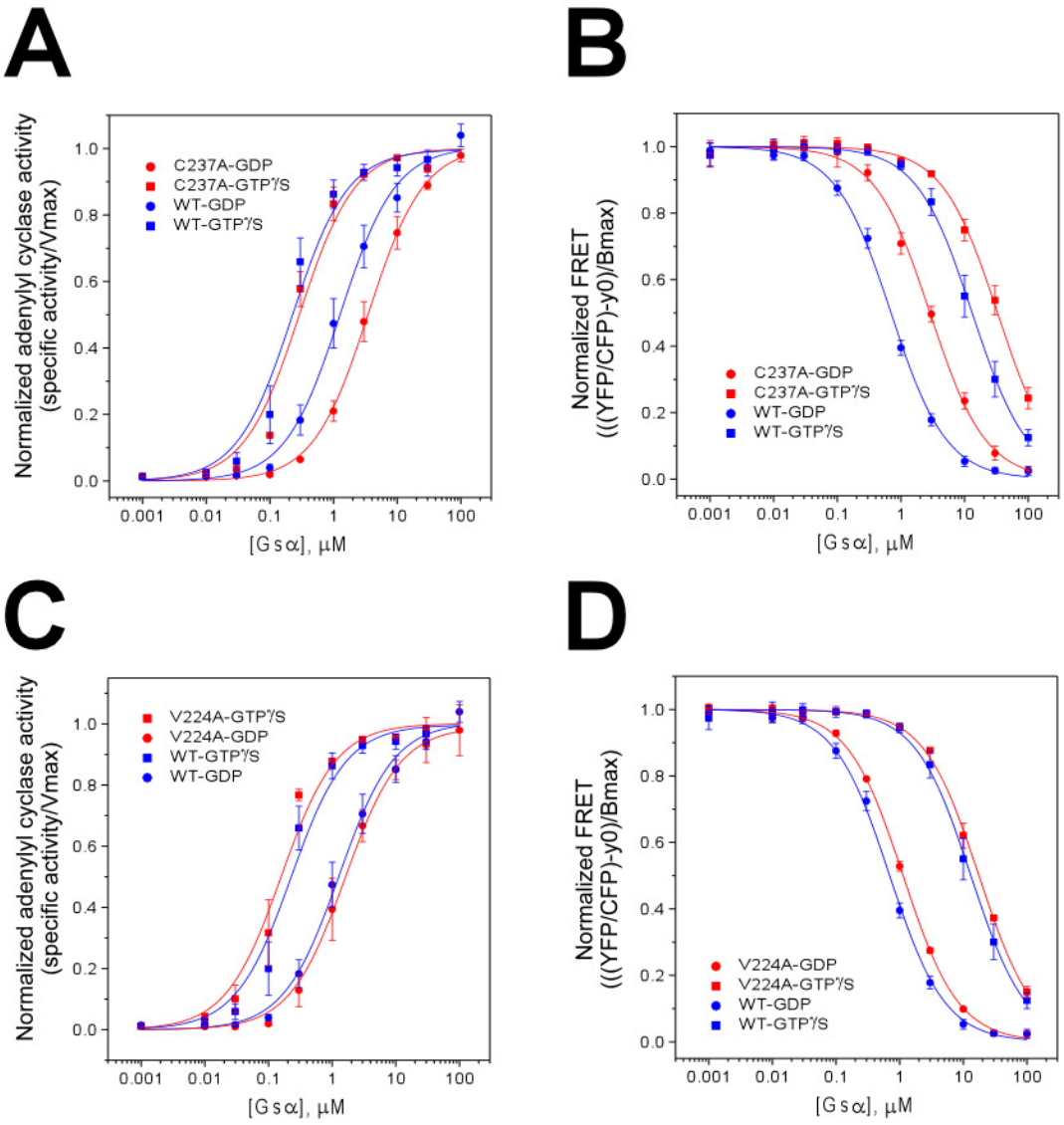


Figure 5-13. Characterization of C237A and V224A (class 3). *Adenylyl cyclase.* Activation of adenylyl cyclase was assayed with increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 1 nM IIC_2 and 10 μM VC_1. The 50 μl reaction was initiated with the addition of 1 mM [^{32}P]ATP and proceeded for 10 min. The data were fit using a non-linear least squares fit to the equation $y=(V_{max}*x)/(EC_{50}+x)$. The C237A (**A**) and V224A (**C**) mutant proteins are the mean and range of two experiments performed in duplicate. The wild type curve is the mean and standard deviation of 26 independent experiments performed in duplicate. The specific activities were normalized by dividing by the Vmax. *Binding of βγ.* Increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 200nM of H6-Giα-CFP and β1γ2-YFP. The samples were incubated at 18 °C for 16 hours. The fluorescence was recorded at 20 °C from 440 to 560 nm following excitation at 410 nm. The FRET is represented as the ratio of the fluorescence intensity of YFP at 522nm to that of CFP a 474nm. The C237A (**B**) and V224A (**D**) mutant proteins are the mean and range of two experiments. The wild type curve is the mean and standard deviation of 10 independent experiments.

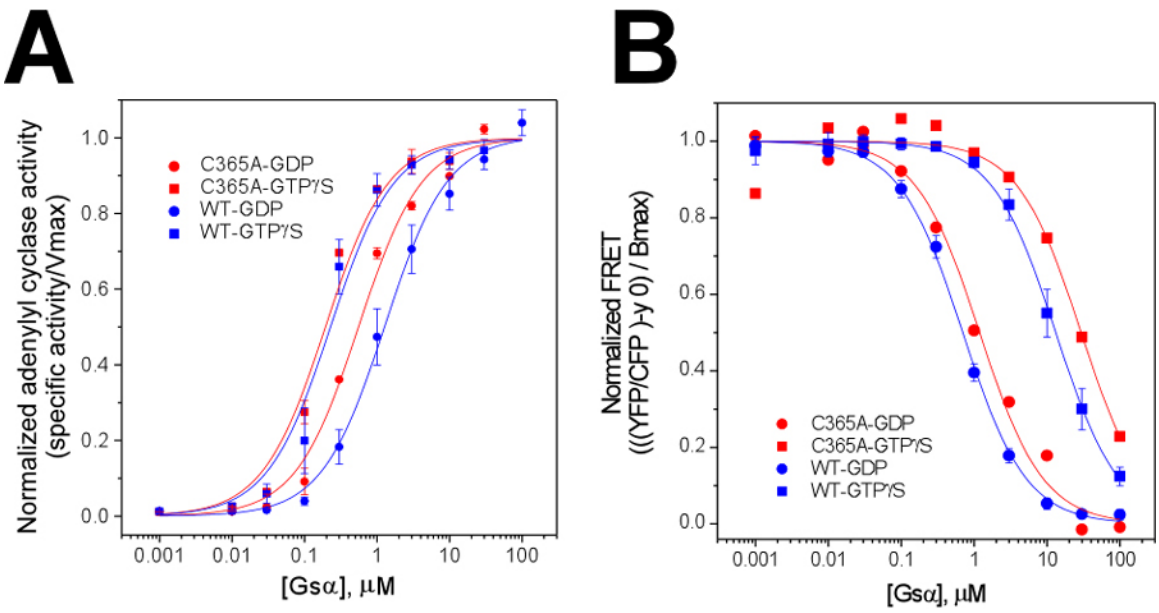


Figure 5-14. Characterization of C365A (class 3). *Adenylyl cyclase.* Activation of adenylyl cyclase was assayed with increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 1 nM IIC_2 and 10 μM VC_1. The 50 μl reaction was initiated with the addition of 1 mM [^{32}P]ATP and proceeded for 10 min. The data were fit using a non-linear least squares fit to the equation $y=(V_{max}*x)/(EC_{50} + x)$. The C365A (**A**) mutant protein is the mean and range of two experiments performed in duplicate. The wild type curve is the mean and standard deviation of 26 independent experiments performed in duplicate. The specific activities were normalized by dividing by the Vmax. *Binding of βγ.* Increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 200nM of H6-Giα-CFP and β1γ2-YFP. The samples were incubated at 18 ^{o}C for 16 hours. The fluorescence was recorded at 20 ^{o}C from 440 to 560 nm following excitation at 410 nm. The FRET is represented as the ratio of the fluorescence intensity of YFP at 522nm to that of CFP a 474nm. The C365A (**B**) mutant protein was the result of one experiment due to lack of protein. The wild type curve is the mean and standard deviation of 10 independent experiments.

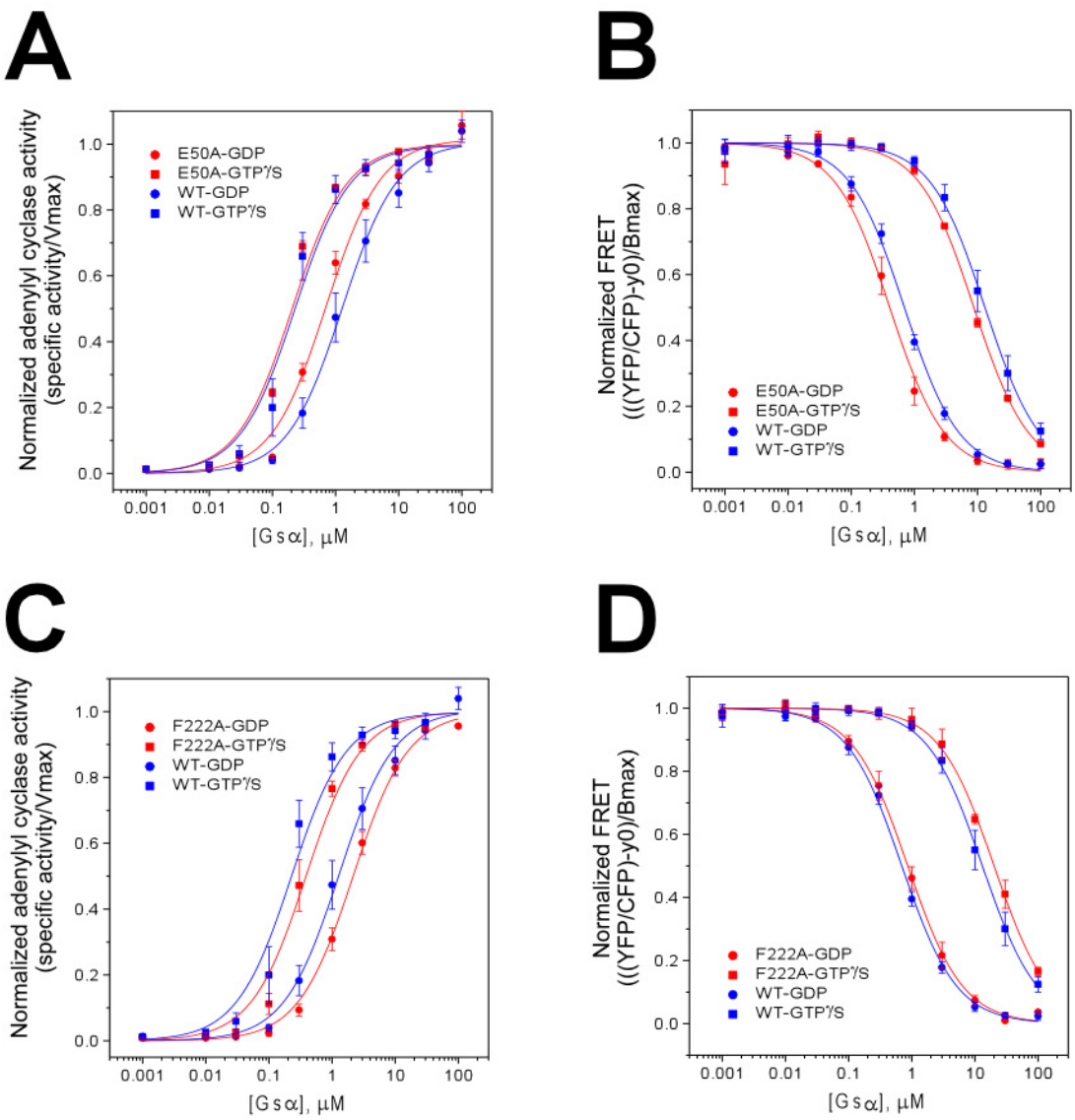


Figure 5-15. Characterization of E50A and F222A (class 4). *Adenylyl cyclase.* Activation of adenylyl cyclase was assayed with increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 1 nM IIC_2 and 10 μM VC_1. The 50 μl reaction was initiated with the addition of 1 mM [^{32}P]ATP and proceeded for 10 min. The data were fit using a non-linear least squares fit to the equation $y=(V_{max}*x)/(EC_{50}+x)$. The E50A (**A**) and F222A (**C**) mutant proteins are the mean and range of two experiments performed in duplicate. The wild type curve is the mean and standard deviation of 26 independent experiments performed in duplicate. The specific activities were normalized by dividing by the Vmax. *Binding of βγ.* Increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 200nM of H6-Giα-CFP and β1γ2-YFP. The samples were incubated at 18 ^{o}C for 16 hours. The fluorescence was recorded at 20 ^{o}C from 440 to 560 nm following excitation at 410 nm. The FRET is represented as the ratio of the fluorescence intensity of YFP at 522nm to that of CFP a 474nm. The E50A (**B**) and F222A (**D**) mutant proteins are the mean and range of two experiments. The wild type curve is the mean and standard deviation of 10 independent experiments.

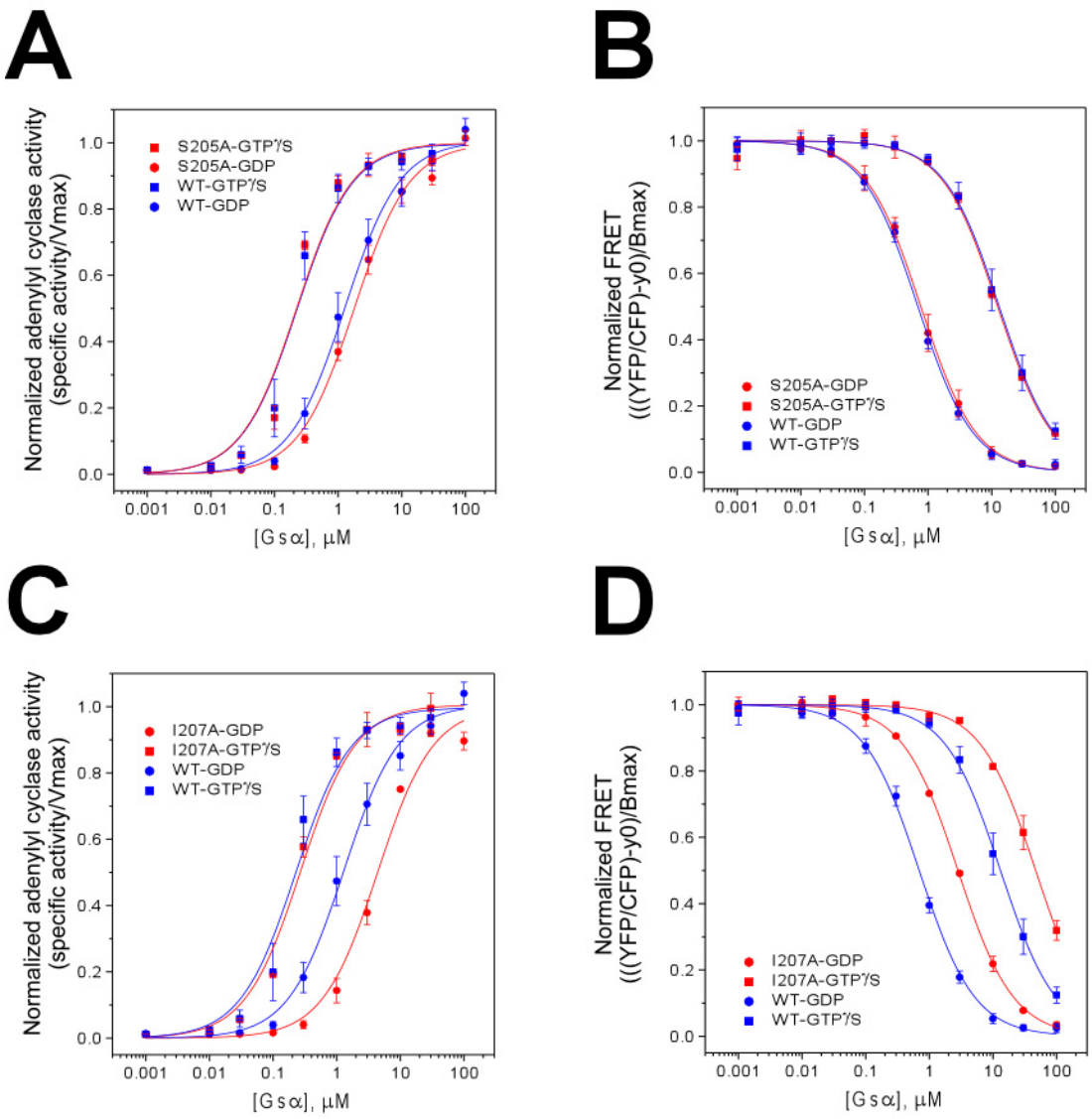


Figure 5-16. Characterization of S205A and I207A (non-coupled controls). *Adenylyl cyclase.* Activation of adenylyl cyclase was assayed with increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 1 nM IIC_2 and 10 μM VC_1. The 50 μl reaction was initiated with the addition of 1 mM [^{32}P]ATP and proceeded for 10 min. The data were fit using a non-linear least squares fit to the equation $y=(V_{max}*x)/(EC_{50}+x)$. The S205A (**A**) and I207A (**C**) mutant proteins are the mean and range of two experiments performed in duplicate. The wild type curve is the mean and standard deviation of 26 independent experiments performed in duplicate. The specific activities were normalized by dividing by the Vmax. *Binding of βγ.* Increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 200nM of H6-Giα-CFP and β1γ2-YFP. The samples were incubated at 18 °C for 16 hours. The fluorescence was recorded at 20 °C from 440 to 560 nm following excitation at 410 nm. The FRET is represented as the ratio of the fluorescence intensity of YFP at 522nm to that of CFP a 474nm. The S205A (**B**) and I207A (**D**) mutant proteins are the mean and range of two experiments. The wild type curve is the mean and standard deviation of 10 independent experiments.

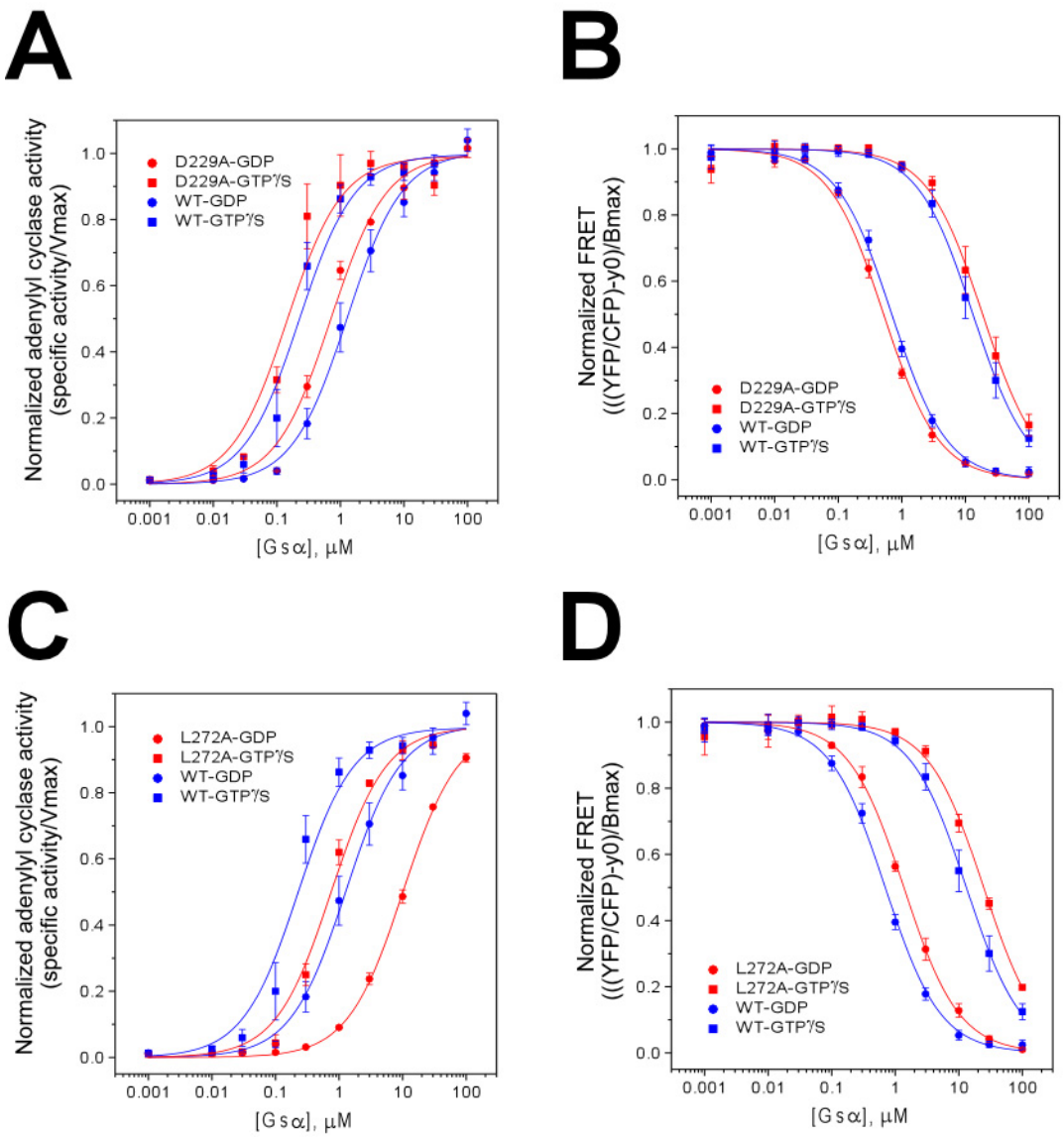


Figure 5-17. Characterization of D229A and L272A (non-coupled controls). *Adenylyl cyclase.* Activation of adenylyl cyclase was assayed with increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 1 nM IIC_2 and 10 μM VC_1. The 50 μl reaction was initiated with the addition of 1 mM [^{32}P]ATP and proceeded for 10 min. The data were fit using a non-linear least squares fit to the equation $y=(V_{max}*x)/(EC_{50} + x)$. The D229A (**A**) and L272A (**C**) mutant proteins are the mean and range of two experiments performed in duplicate. The wild type curve is the mean and standard deviation of 26 independent experiments performed in duplicate. The specific activities were normalized by dividing by the Vmax. *Binding of βγ.* Increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 200nM of H6-Giα-CFP and β1γ2-YFP. The samples were incubated at 18 °C for 16 hours. The fluorescence was recorded at 20 °C from 440 to 560 nm following excitation at 410 nm. The FRET is represented as the ratio of the fluorescence intensity of YFP at 522nm to that of CFP a 474nm. The D229A (**B**) and L272A (**D**) mutant proteins are the mean and range of two experiments. The wild type curve is the mean and standard deviation of 10 independent experiments.

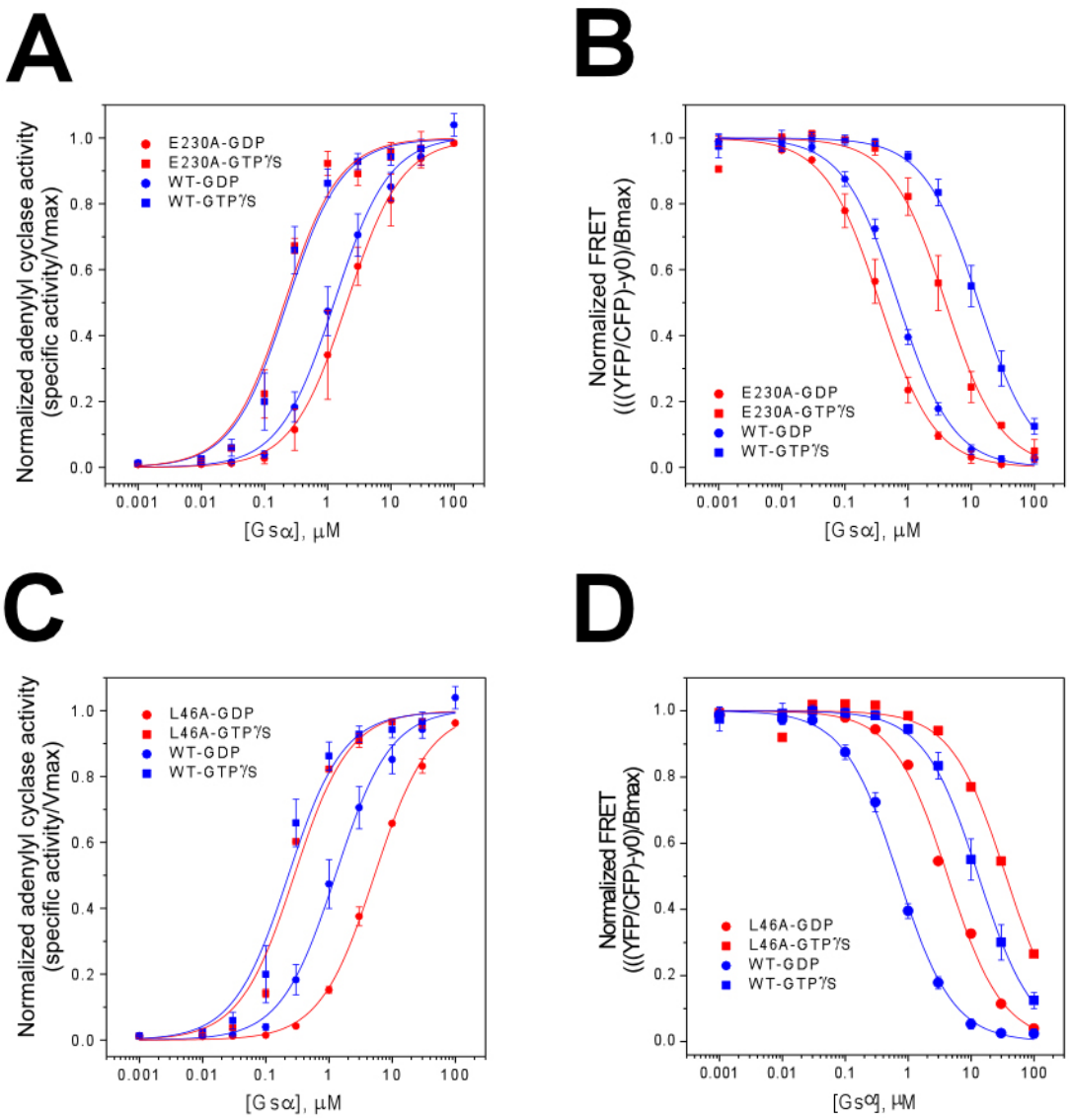


Figure 5-18. Characterization of E230A and L46A (non-coupled controls). *Adenylyl cyclase.* Activation of adenylyl cyclase was assayed with increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 1 nM IIC_2 and 10 μM VC_1. The 50 μl reaction was initiated with the addition of 1 mM [^{32}P]ATP and proceeded for 10 min. The data were fit using a non-linear least squares fit to the equation $y=(V_{max}*x)/(EC_{50}+x)$. The E230A (**A**) and L46A (**C**) mutant proteins are the mean and range of two experiments performed in duplicate. The wild type curve is the mean and standard deviation of 26 independent experiments performed in duplicate. The specific activities were normalized by dividing by the Vmax. *Binding of βγ.* Increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 200nM of H6-Giα-CFP and β1γ2-YFP. The samples were incubated at 18 °C for 16 hours. The fluorescence was recorded at 20 °C from 440 to 560 nm following excitation at 410 nm. The FRET is represented as the ratio of the fluorescence intensity of YFP at 522nm to that of CFP a 474nm. The E230A (**B**) mutant protein is the mean and range of two experiments and the L46A (**D**) mutant protein is the result of one experiment due to lack of protein. The wild type curve is the mean and standard deviation of 10 independent experiments.

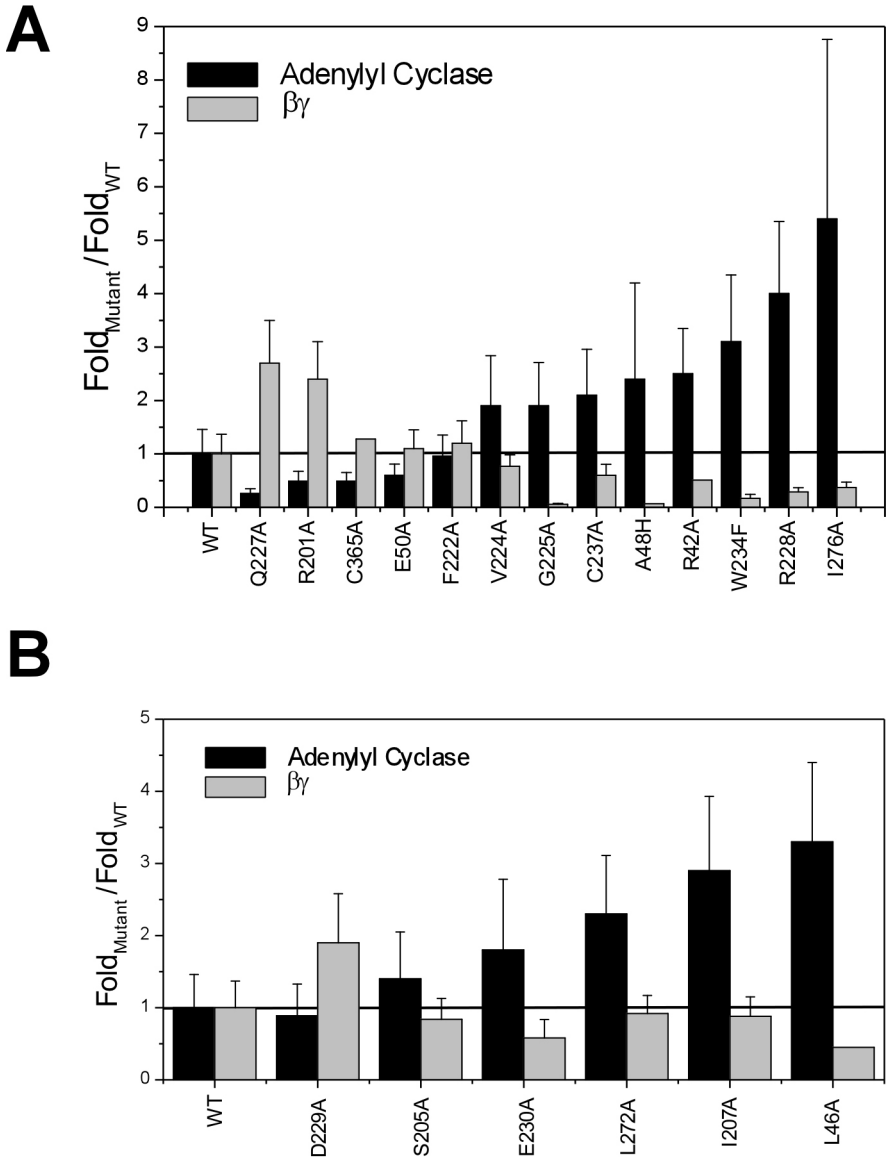


Figure 5-19. Ratios of affinities of mutant and wild type $G_{s\alpha}$ proteins for binding adenylyl cyclase and βγ. The fold-AC and fold-βγ of the mutant $G_{s\alpha}$ proteins from Tables 5-1 and 5-2 were normalized by divided by that of the wild type protein. The error bars are the propagated error from this calculation. This normalized the fold-AC and fold-βγ of the wild type protein to one; the line at y = 1 represents the level of wild type. The normalized fold-AC for adenylyl cyclase is represented as black bars and the normalized fold-βγ for βγ is represented as gray bars. The results of mutations at coupled position are shown in **A** and the results of mutations at non-coupled positions are shown in **B.** The mutations in both **A** and **B** are arranged in order of increasing sensitivity for adenylyl cyclase. As the normalized fold-AC for adenylyl cyclase increased, the fold-βγ for βγ decreased with proteins with mutations in coupled positions (**A**). This diametric regulation was only observed in one of the proteins with mutations in non-coupled positions, L46A (**B**).

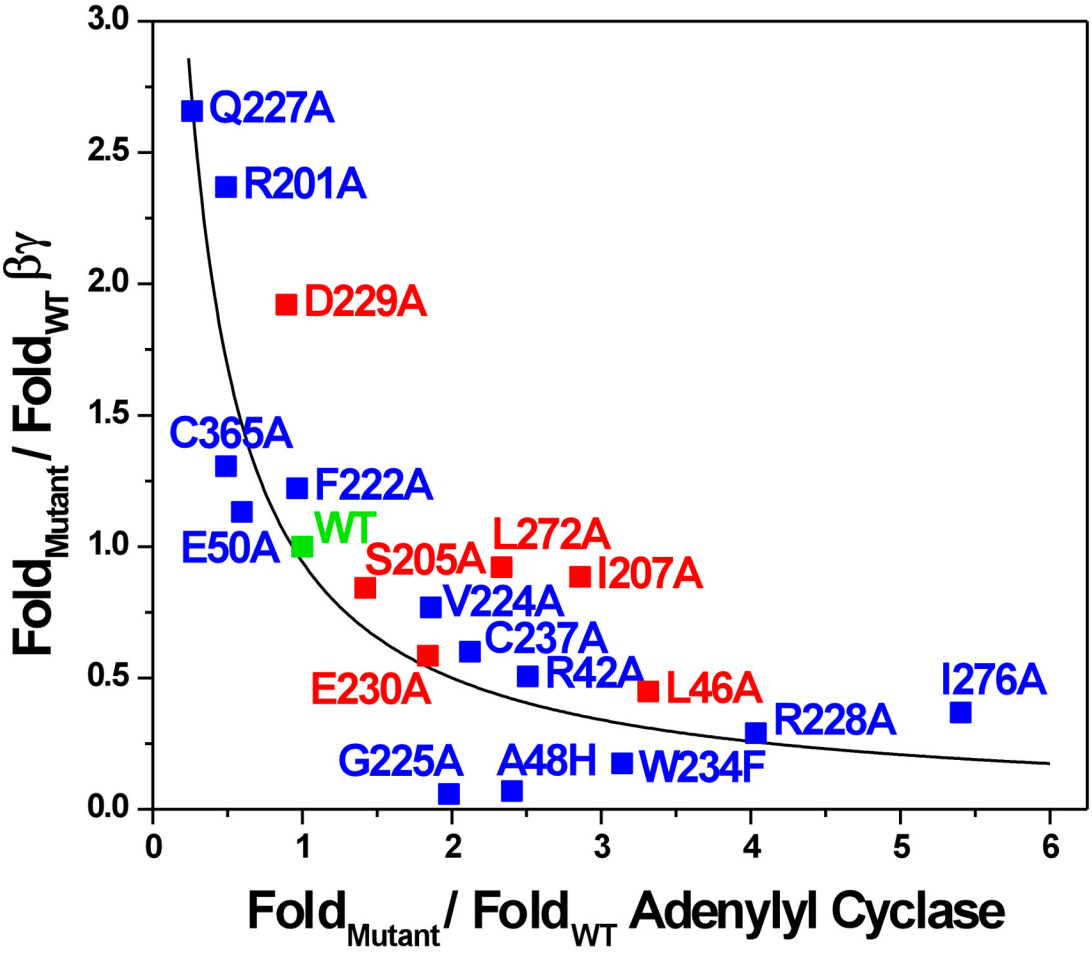


Figure 5-20. Correlation of nucleotide sensitivities for adenylyl cyclase and βγ. The values for both coupled and non-coupled mutant proteins from the bar graphs in Figure 5-19 were used to generate a plot of the normalized fold-βγ versus the normalized fold-AC. The values of statistically coupled positions are blue while the non-coupled controls are red. The wild type is shown in green. The mutant proteins of class 1(R201A and Q227A) and the class 2 (R42A, A48H, G225A, R228A, W234F, and I276A) that display the diametric regulation of adenylyl cyclase and βγ apparent affinities group by class separate from the other mutant proteins. The mutations of positions predicted from the statistical coupling analysis fit an inverse relationship (y = 1/x) with a correlation coefficient of 0.85 while the mutations of the control non-coupled positions deviate from this inverse relationship with a correlation coefficient of 0.42.

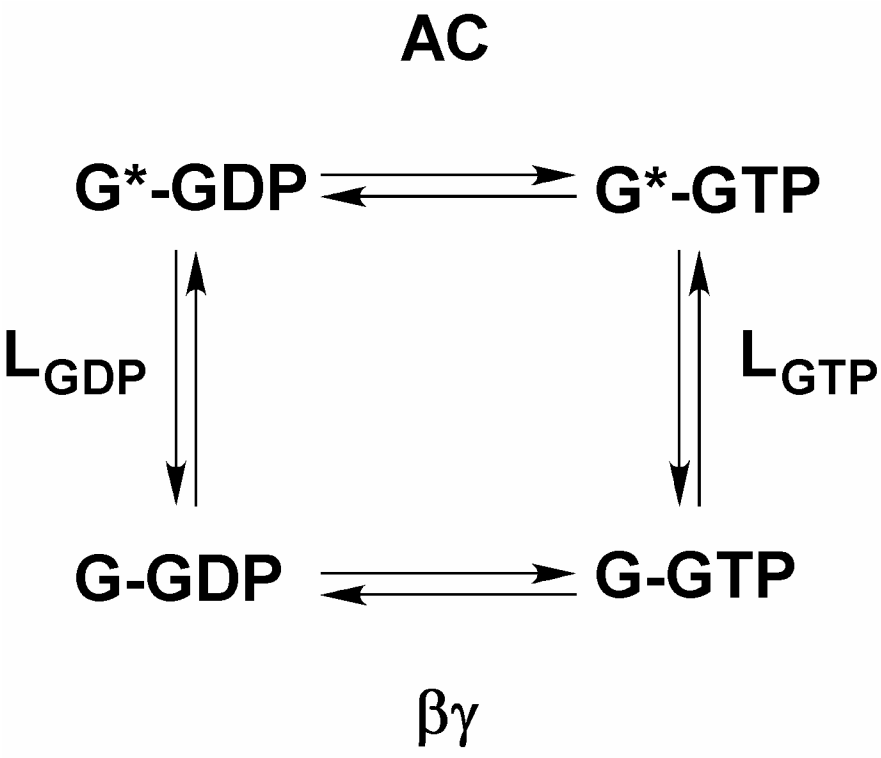


Figure 5-21. Two state equilibrium model for allosteric regulation of G protein. The equilibrium between the ground state, G, and the active state, G*, is affected by the equilibrium constants L_{GTP} and L_{GDP}. In the ground state, both G-GDP and G-GTP bind the βγ subunits. Both G*-GDP and G*-GTP bind and activate adenylyl cyclase (AC). The mutations identified in this study are suggested to alter the equilibrium constants L_{GTP} and L_{GDP}, thus shifting the equilibrium between the ground state, G, and the active state, G*. Mutations that alter L_{GTP} and L_{GDP} in favor of the G* state would decrease the nucleotide dependent difference in the activation of adenylyl cyclase by stabilizing G*-GDP. The shift in equilibrium to G* would result in a concomitant destabilization of G. This destabilization of G results in an increase in nucleotide dependent binding of the βγ subunits. Mutations that alter L_{GTP} and L_{GDP} in favor of the G state would have the reciprocal effect.

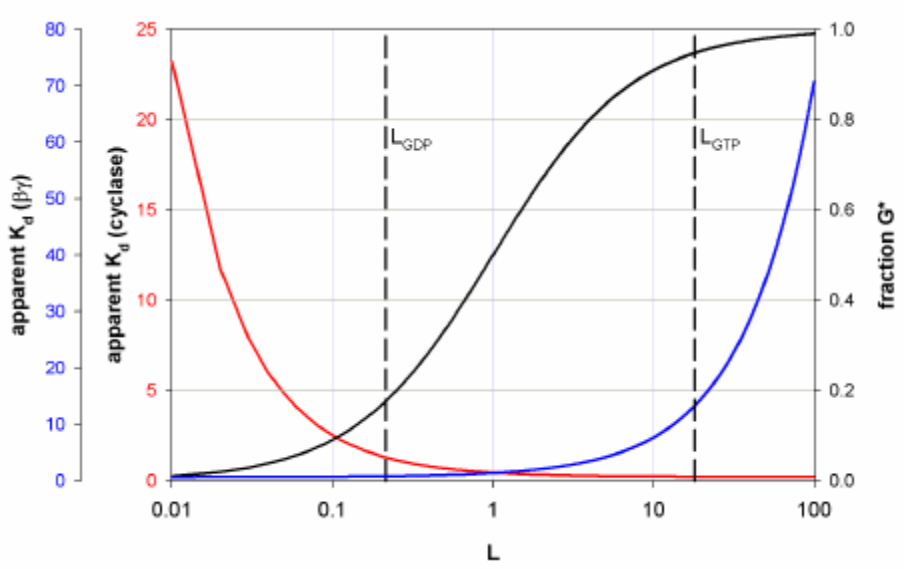


Fig. 5-22. Illustration of three important features of the allosteric model: (*i*) $L<1$ results in high-affinity binding of βγ (blue curve) and a steeply decreasing affinity for AC (red curve); (*ii*) $L>1$ results in high affinity binding of AC and a steeply decreasing affinity for βγ ; and (*iii*) L = 1 results in complete uncoupling of nucleotide exchange and effector interactions. Quite reasonably, the values for L_{GDP} and L_{GTP} for WT $G_{s\alpha}$ are 0.21 and 18, respectively. These values were estimated by assuming that the equilibrium dissociation constant of adenylyl cyclase for GTPγ S-bound WT $G_{s\alpha}$ is a good estimate of the intrinsic binding affinity for the G* state, and that the equilibrium dissociation constant of βγ for GDP-bound WT $G_{s\alpha}$ is a good estimate of its intrinsic binding affinity for the G state. (Equilibrium equations solved and Figure assembled by Rama Ranganathan).

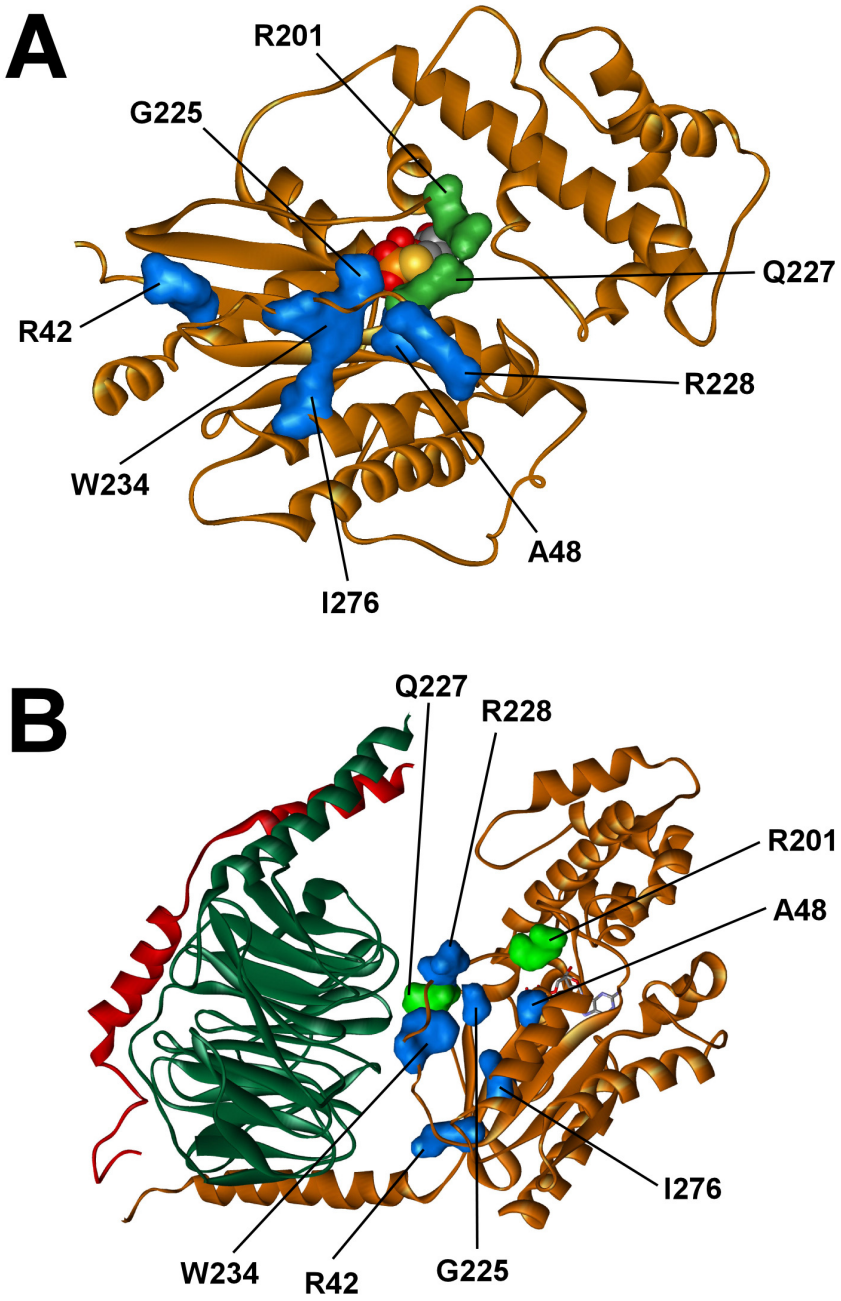


Figure 5-23. Structural map of $G_{s\alpha}$ mutant phenotypes. *Active conformation* (**A**). Mutations that decrease the allosteric activation of adenylyl cyclase and decrease the allosteric binding of the βγ subunits (class 1) are illustrated as green van der Waals surfaces. Mutations that increase the allosteric activation of adenylyl cyclase and decrease the allosteric binding of the βγ subunits (class 2) are illustrated as blue van der Walls surfaces. These groups are shown on a tan ribbon structure of $G_{s\alpha}$ bound to GTPγS (Protein Databank entry 1AZT). *Inactive conformation* **(B).** The homolgous residues were mapped on GDP-bound $G_{i\alpha}$ in complex with the $\beta_1\gamma_2$ subunits, green and red respectively (Protein Databank entry 1GP2). Residues were labeled according to the $G_{s\alpha}$ residues.

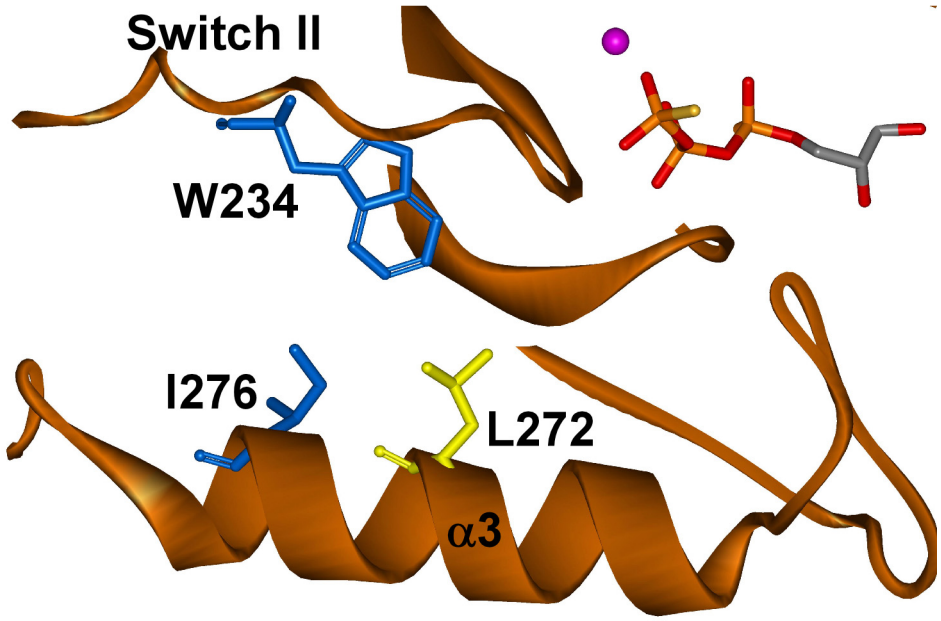


Figure 5-24. Interaction of Tryptophan-234 with ileucine-276 and leucine-272. Trp-234 of switch II contacts both the coupled Ile-276 (blue) and the non-coupled Leu-272 (yellow) of the α3 helix in the $G_{s\alpha}$-GTPγS structure (Protein Databank entry 1AZT).

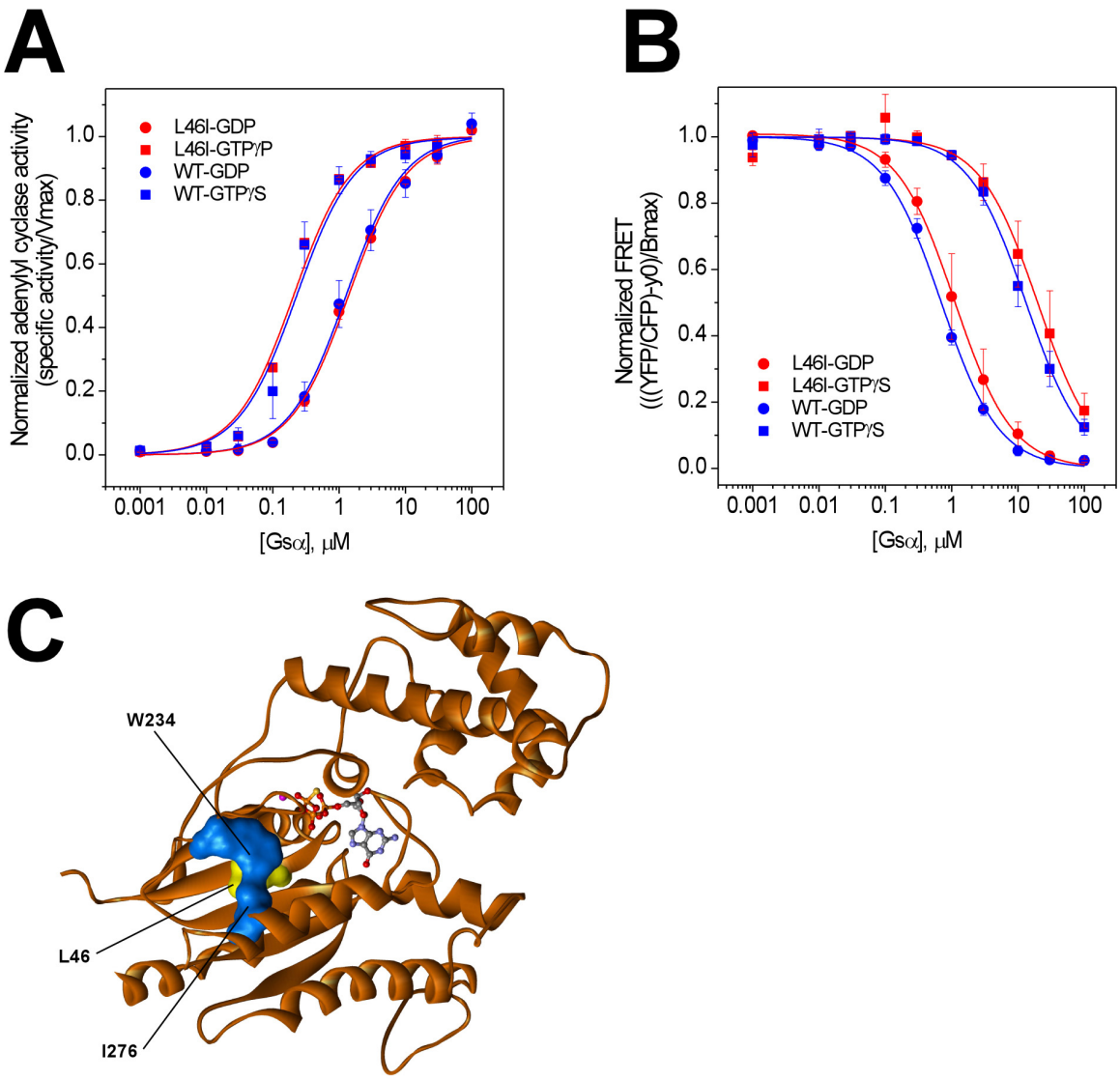


Figure 5-25. Characterization of L46I (non-coupled control). *Adenylyl cyclase* (**A**). Activation of adenylyl cyclase was assayed with increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 1 nM IIC_2 and 10 μM VC_1. The 50 μl reaction was initiated with the addition of 1 mM [^{32}P]ATP and proceeded for 10 min. The data were fit using a non-linear least squares fit to the equation $y=(V_{max}*x)/(EC_{50}+x)$. The mutant protein is the mean and range of two experiments performed in duplicate. The wild type curve is the mean and standard deviation of 26 independent experiments performed in duplicate. The specific activities were normalized by dividing by the Vmax. *Binding of βγ* (**B**). Increasing concentrations of either GTPγS-bound or GDP-bound $G_{s\alpha}$ were added to 200nM of H6-Giα-CFP and β1γ2-YFP. The samples were incubated at 18 °C for 16 hours. The fluorescence was recorded at 20 °C from 440 to 560 nm following excitation at 410 nm. The FRET is represented as the ratio of the fluorescence intensity of YFP at 522nm to that of CFP a 474nm. The mutant protein is the mean and range of two experiments. The wild type curve is the mean and standard deviation of 10 independent experiments. *Position of L46* (**C**). The statistically coupled residues are represented as blue and the non-coupled as yellow van der Waals surfaces on a tan ribbon structure of $G_{s\alpha}$ bound to GTPγS (Protein Databank entry 1AZT).

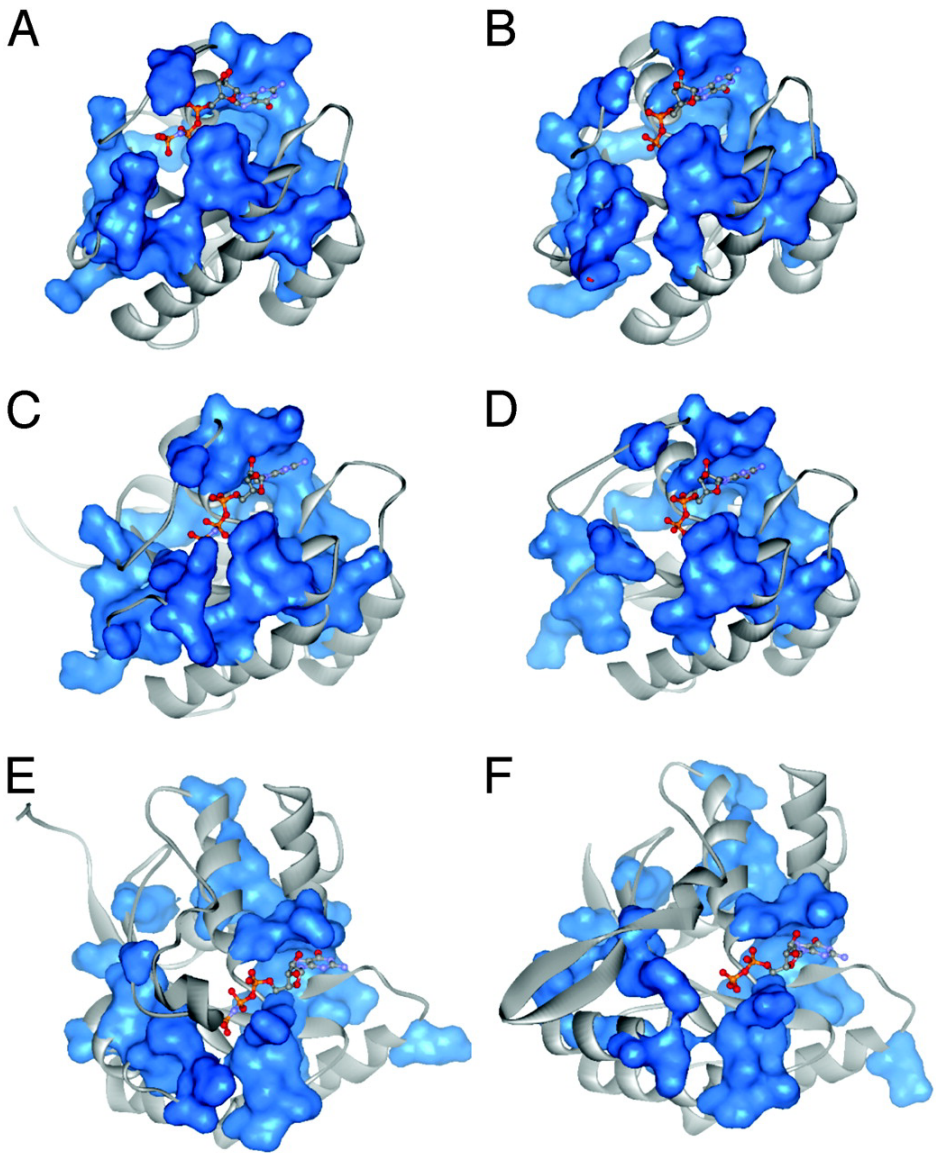


Figure 5-26. Mapping the statistically coupled amino acide network on active and inactive states of distant G protein family members. GTPγS- [or Gpp(NH)p-] bound structures (**A, C,** and **E**) and GDP-bound structures (**B, D,** and **F**) of H-Ras (**A** and **B**), Ypt7p (**C** and **D**), and EF-Tu (**E** and **F**) are illustrated. PBD ID codes: 5P21 (**A**), 4Q21 (**B**), 1KY2 (**C**), 1KY3 (**D**), 1EFT (**E**), and 1TUI (**F**). This figure was assembled by Rama Ranganathan.

Chapter 6 – Conclusions and Future Directions

Crystallization of the low activity state of adenylyl cyclase catalytic core

The attempts to determine the structure of a low activity state of adenylyl cyclase have been hampered by the low affinity of the C_1 and C_2 domains in the absence of activators (>10 μM). Biochemical methods to purify a heterodimeric complex of the VC_1 and IIC_2 domains is complicated by the similarity in molecular mass and the formation of homodimers by each of the domains; thus, eliminating gel filtration as a purification technique. A genetic screen in *Saccharomyces cerevisiae* identified mutations in the C_1 and C_2 domains of mammalian adenylyl cyclase that activate the enzyme in the absence of $G_{s\alpha}$. Several of these mutations were located at the interface between the VC_1 and IIC_2 domain in the $G_{s\alpha}$ co-crystal. It was hoped that the constitutive activity of at least some of these mutant proteins resulted from an increased affinity of the C_1 and C_2 domains. The IIC_2-K1014N mutant protein displayed an increased affinity for the VC_1 protein and increased enzymatic activity in the absence of activators. The identification of the K1014N mutation in the IIC_2 domain along with a successful purification scheme reported in Chapter 2 allowed the crystallization of the VC_1:IIC_2:Fsk:2'd3'-AMP:PP_i complex. Unfortunately, these crystals failed to diffract resulting from inadequate size and morphology of the crystals and heterogeneity resulting from the lack of affinity.

Alternative methods for increasing the affinity of the cytosolic domains were explored. Intein-mediated protein ligation was used to post-translationally covalently link the VC_1 and IIC_2 domains as described in Chapter 3. This method was plagued with technical difficulties. The VC_1 domain was inefficiently cleaved from the intein fusion and had reduced activity compared to the wild type protein. The ligation of the VC_1 and IIC_2 domains was inefficient and the low yield was not sufficient for further purification or crystal screens. The linker length between the two domains was

thought to be insufficient to allow successful interaction of the modified termini. Linkers would need to be added to increase the efficiency of the ligation reaction.

A rapid method for screen for functional interaction of the C_1 and C_2 domains was developed. Danchin and colloegues have previously developed strains of *E. coli* lacking adenylyl cyclase activity (Roy and Danchin, 1981; Roy and Danchin, 1982). Complementation assays using these strains had been used to isolate the genes encoding several bacterial adenylyl cyclases. Wei-Jen Tang had also used these strains to assay enzymatic activity of the linked IC_1-IIC_2 fusion protein (Tang and Gilman, 1995). Serendipitously, it was discovered that the enzymatic activity of a simple fusion of the VC_1 and IIC_2 domains was sufficient to complement the deficiency in these bacteria in the absence of $G_{s\alpha}$ or forskolin (see Chapter 4). The VC_1-IIC_2 fusion protein had much greater enzymatic activity than the previously studied IC_1-IIC_2 fusion protein. When expressed independently, the VC_1 domain is much more soluble than the IC_1 domain. The use of the VC_1 domain in the C_1-C_2 fusion could increase the solubility. The high level of cAMP produced from increased activity of the VC_1-IIC_2 fusion protein was toxic to the bacteria even in the absence of IPTG. This precludes expression in conventional bacteria.

Several approaches could circumvent the cAMP toxicity. First, the mutation of Asn-1025 of the IIC_2 domain would reduce the k_{cat} without affecting the K_M for ATP. Although the substrate would not be in the proper orientation, structural determination of a VC_1-IIC_2 protein could reveal the conformational changes in the C_1 and C_2 domains upon $G_{s\alpha}$ binding. The TP2339 strain of *E. coli* lacks a functional adenylyl cyclase and CRP; therefore, this strain would not be able to respond to the high level of cAMP produced by the VC_1-IIC_2 fusion protein. The purification of a linked construct would solve the affinity problem and could allow the structural determination of the low activity state of adenylyl cyclase. The linked VC_1-IIC_2 protein could provide insight to the mechanism of $G_{i\alpha}$ inhibition. $G_{i\alpha}$ is presumed to act by decreasing the affinity of the C_1 and C_2 domains. The linked

construct could elucidate the conformational changes in the VC_1:IIC_2 complex upon $G_{i\alpha}$ binding. As well, it could be useful in determination of the postulated quaternary complex of $G_{i\alpha}$:C_1:C_2:$G_{s\alpha}$.

Allosteric determinants of G Proteins

Members of the G protein superfamily contain nucleotide-dependent switches that dictate the specificity of their interactions with binding partners. Chapter 5 described a sequence-based method termed statistical coupling analysis (SCA) used to identify the allosteric core of these proteins – the network of amino acid residues that couples the domains responsible for nucleotide binding and protein-protein interactions. One-third of the 38 residues identified by SCA were mutated in the G protein $G_{s\alpha}$, and the interactions of GTPγS- and GDP-bound mutant proteins were tested with both adenylyl cyclase (preferential binding to GTP-$G_{s\alpha}$) and the G protein βγ subunit complex (preferential binding to GDP-$G_{s\alpha}$). A two-state allosteric model predicts that mutation of residues that control the equilibrium between GDP- and GTP-bound conformations of the protein will cause the ratio of affinities of these species for adenylyl cyclase and $G_{\beta\gamma}$ to vary in a reciprocal fashion. The observed results presented in Chapter 5 were consistent with this prediction. The network of residues identified by the SCA appears to comprise a core allosteric mechanism conferring nucleotide-dependent switching; the specific features of different G protein family members are built upon this core. Several questions from the studies presented in Chapter 5 remain. First, how does the substitution of alanine at coupled and non-coupled positions affect the interpretation of the analysis? Second, what is the significance of the grouping and magnitude of the statistical coupling matrix? Finally, does this allosteric core regulate the other members of the G protein family?

Accounting for phenotypes of alanine substitutions – Mutation analysis is an imprecise tool for studying the function of amino acid residues in proteins. The effect of a substitution at a position could result either from the loss or gain of function due to that substitution. Several of the mutations

in non-coupled residues resulted in proteins with effects on binding adenylyl cyclase or βγ. However, only mutation of one of the non-coupled residues, L46A, resulted in the reciprocal effects on the nucleotide sensitivity for the two binding partners. This was particularly striking since all of the non-coupled positions contacted at least one of the coupled positions, thus indicating the discrete manner that enegy progates through proteins.

One could argue that the reciprocity for binding adenylyl cyclase and βγ exhibited by the L46A mutant protein is due to a gain-of-function mutation as a result of replacing a large hydrophobic residue in the core of $G_{s\alpha}$ with an alanine. Leu-46 contacts the coupled residue Trp-234 and contributes to the hydrophobic pocket where Trp-234 binds in the GTPγS-bound state. Eriksson et al. demonstrated that proteins compensate for the loss of volume created by an alanine mutation in the core of proteins repositioning the neighboring residues to fill the void created by the mutation. If the void remains unfilled, the protein is less stable (Eriksson et al., 1992). Therefore, the phenotype resulting from the replacement of Leu-46 with alanine could result from the repositioning of the adjacent coupled residues, Trp-234 and Ile-276. Consistent with this idea, the relatively isovolemic replacement of Leu-46 with isoleucine displayed no difference from the wild type in binding adenylyl cyclase and βγ.

Schreiber and Fersht illustrated that the free energy of a mutation radiates 4 to 5 angstroms and decreases with distance; therefore, a mutation of a non-coupled residue in the core of $G_{s\alpha}$ separated from a coupled residue by another non-coupled residue should result in less of an effect on the diametric regulation of binding adenylyl cyclase and βγ (Schreiber and Fersht, 1995a). Both Ile-244 and Ile-245 are located in the core of $G_{s\alpha}$ and in close proximity of Leu-46; however, both residues are separated from coupled residues. Since these residues were not identified in the SCA as coupled, the mutation of Ile-244 and Ile-245 to either alanine or leucine should have no effect on the diametric regulation of $G_{s\alpha}$. The replacement of these hydrophobic residues in the core of $G_{s\alpha}$ with

alanine could cause repositioning of local residues, but should have less effect on the diametric regulation since the residues are not directly packing against coupled residues.

From inspection of the crystal structure, it is difficult to explain the selection of some residues as coupled and others non-coupled. For example, Trp-234 in switch II contacts both Ile-276 and Leu-272 in the α3 helix in the GTP-bound structure (Figure 5-24). Both Ile-276 and Leu-272 contribute to the hydrophobic environment of Trp-234. However, only Trp-234 and Ile-276 were predicted as coupled from the SCA. The coupled W234F and I276A mutant proteins exhibited the diametric binding to adenlylyl cyclase and βγ while the non-coupled L272A mutant protein did not. Therefore, Trp-234 and Leu-272 are structurally coupled but not necessarily energetically coupled. If this were true, thermodynamic mutant cycle analysis of these residues should reveal energetic non-additivity for Trp-234 and Ile-276 while Leu-272 would have additive effects on both Trp-234 and Ile-276.

The L272A mutant protein did increase the nucleotide sensitivity for adenylyl cyclase although it did not exhibit the reciprocal decrease in nucleotide sensitivity for βγ. This effect on binding adenylyl cyclase was thought to result from the repositioning of the coupled Trp-234 and Ile-276 residues in the G* state in response to a loss of volume from the alanine substitution. Therefore, the creation of a cavity and the possible compensation by the adjacent coupled positions could account for the phenotype. The isovolemic replacement of Leu-272 with isoleucine should decrease the effect on the nucleotide sensitivity of adenylyl cyclase similar to the phenotype of the replacement of Leu-46 with isoleucine. It is tempting to speculate that the coupled residues would be more sensitive to mutations than the non-coupled residues. To address this, the coupled position Ile-276 could be mutated to leucine. This conservative replacement of the side chain should strengthen this argument if the I276L mutant protein maintains the reciprocal allosteric regulation of adenylyl cyclase and βγ.

Significance of clustering in matrix analysis – It is impossible to relate the phenotypes of the mutant proteins from the SCA to the pattern of the clustered matrix of statistical coupling. The grouping of positions in the clustered matrix does not correlate with the phenotypes of the $G_{s\alpha}$ mutant proteins. For example, positions Gln-227 and Arg-42 cluster together (Figure 5-4); however, mutations of the postions to alanine yielded opposite phenotypes (Table 5-3). The positions Arg-42, Gln-227, and Phe-376 appear significantly clustered from the matrix (Fifure 5-4). Perturbations in each of these positions yield high coupling values in the other two positions, and these positions cluster together as a feature of the coupling values for all perturbations. These residues do not form direct contacts in the structure. Arginine-42 and Gln-227 appear to be connected through van der Waals contacts with other statistically coupled residues. Unfortunately, only the Q227A and R42A mutant proteins were appreciably purified when expressed in *E. coli.* The F376A and F376Y mutant proteins were insoluble when expressed in *E. coli.* Phenylalanine-376 does not appear to form any contacts with other statistically coupled residues, and the inability to express mutations of Phe-376 could indicate that the residue is important for protein folding. A more conservative replacement of Phe-376 with leucine could be made. The mutant cycle analysis of R42A and Q227A could reveal the coupling of residues that are distantly separated in the structure; however, the interpretation of a double mutant protein could be difficult since the single mutant proteins have opposite phenotypes.

Completion of a few thermodynamic mutant cycles with a series of double mutants that are statistically coupled but not directly structurally coupled will provide insight into the significance of the clustering and mechanism of allostery. For example, positions Gly-225 and Ile-276 cluster together and have the same level of coupling to all the perturbations. These two residues are not in contact in either the GDP- or GTP-bound states; however, both residues contact Trp-234 in the GTP-bound state. The analysis of this mutant cycle will address whether the statistical coupling pattern is predictive of thermodynamic coupling between residues.

Perturbation 242 results in a large coupling value for position Gln-227. However, position Thr-242 has low to moderate coupling across all perturbations and is not part of the functional cluster (Figure 5-4). The Gln-227 is located at the amino-terminus of the α2 helix and Thr-242 is located at the carboxy-terminus of α2 in the α2/ β4 loop. A double mutant of these residues will determine the significance of high statistical coupling of a position to the position of the perturbation made. The completion of these experiments should determine how clustering of statistically coupled residues correlates with allostery of the G protein alpha subunit and give insight to the mechanism of allostery.

Allosteric core in other G protein family members - The statistical coupling analysis used a multiple sequence alignment containing diverse members of the G protein family including the heterotrimeric G protein alpha subunits, ras-like small G proteins, and the translation elongation factors. Although the analysis described in Chapter 5 tested the coupling in heterotrimeric G proteins, the nature of the method suggests that the analogous network of residues would compose the allosteric determinants of the entire family. Analysis of the allosteric mechanism in the ras family of low-molecular weight GTPases using H-Ras and its state dependent binding partner Raf-1 is being performed in the Ranganathan laboratory. Unlike the heterotrimeric G proteins, H-Ras does not have a βγ-like GDP-dissociation inhibitor to assess interactions in the GDP-bound state. However, the nucleotide dependence of the H-Ras/Raf-1 interaction should allow determination of perturbation of the allosteric core without measuring the reciprocity of binding as in the heterotrimer G protein analysis.

The residues identified as coupled from the SCA should represent the features of the entire family. In the G protein analysis, I propose that the residues comprise the allosteric core that controls the nucleotide-dependent switching between G and G* states. The residues responsible for interaction with binding partners would not be identified with such analysis since those residues would be specific for only a small subset of the family. Since these allosteric core residues are a

conserved property of the entire G protein superfamily, it is tempting to speculate that the allosteric core from one member of the family could be substituted for the core of another member of the family. The sequential replacement of the identified allosteric core of $G_{s\alpha}$ with the analogous residues from H-Ras could result in the restoration of allostery as measured by the nucleotide dependent interactions with adenylyl cyclase and $\beta\gamma$ as well as GTP hydrolysis and GTP binding.

The Ranganathan laboratory has mapped the enegetic connectivity of the GPCR family (Suel et al., 2003). The analysis of the statistically coupled residues in Drosophilia rhodopsin is underway in the Ranganathan laboratory. Further investigation of the SCA could reveal the flow of information elicited by ligand binding to the G protein coupled receptor and through G protein alpha subunit. The next step in the flow of information in hormone signaling is adenylyl cyclase. A multiple sequence alignment containing 325 adenylyl cyclase and guanylyl cyclase cytosolic domains has been assembled. The statistical coupling analysis of the cyclase family could reveal the network of residues responsible for relaying information from the binding surface to the active site of the enzyme. The SCA of adenylyl cyclase could enhance our understanding of $G_{s\alpha}$ activation as well as $G_{i\alpha}$ inhibition of adenylyl cyclase. The completion of these analyses will illuminate the flow of information from the extracellular surface of the cell to the site of production of the second messenger cyclic AMP.

REFERENCES

Abood, M.E., Hurley, J.B., Pappone, M.-C., Bourne, H.R., and Stryer, L. (1982). Functional homology between signal-coupling proteins. J.Biol.Chem. *257*, 10540-10543.

Altenbach, C., Cai, K., Klein-Seetharaman, J., Khorana, H.G., and Hubbell, W.L. (2001a). Structure and function in rhodopsin: mapping light-dependent changes in distance between residue 65 in helix TM1 and residues in the sequence 306-319 at the cytoplasmic end of helix TM7 and in helix H8. Biochemistry *40*, 15483-15492.

Altenbach, C., Klein-Seetharaman, J., Cai, K., Khorana, H.G., and Hubbell, W.L. (2001b). Structure and function in rhodopsin: mapping light-dependent changes in distance between residue 316 in helix 8 and residues in the sequence 60-75, covering the cytoplasmic end of helices TM1 and TM2 and their connection loop CL1. Biochemistry *40*, 15493-15500.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic.Acids.Res. *25*, 3389-3402.

Artymiuk, P.J., Poirrette, A.R., Rice, D.W., and Willett, P. (1997). A polymerase I palm in adenylyl cyclase? [letter; comment]. Nature *388*, 33-34.

Berlot, C.H. and Bourne, H.R. (1992). Identification of effector-activating residues of $G_{s\alpha}$. Cell *68*, 911-922.

Beuve, A., Boesten, B., Crasnier, M., Danchin, A., and O'Gara, F. (1990). Rhizobium meliloti adenylyl cyclase is related to eukaryotic adenylate and guanylate cyclases. J.Bacteriol. *172*, 2614-2621.

Bokoch, G.M., Katada, T., Northup, J.K., Hewlett, E.L., and Gilman, A.G. (1983). Identification of the predominant substrate for ADP-ribosylation by islet activating protein. J.Biol.Chem. *258*, 2072-2075.

Botsford, J.L. and Harman, J.G. (1992). Cyclic AMP in prokaryotes. Microbiol.Rev. *56*, 100-122.

Buss, J.E., Mumby, S.M., Casey, P.J., Gilman, A.G., and Sefton, B.M. (1987). Myristoylated α subunits of guanine nucleotide-binding regulatory proteins. Proc.Natl.Acad.Sci.U.S.A. *84*, 7493-7497.

Carter, P.J., Winter, G., Wilkinson, A.J., and Fersht, A.R. (1984a). The use of double mutants to detect structural changes in the active site of the tyrosyl-tRNA synthetase (Bacillus stearothermophilus). Cell *38*, 835-840.

Casey, P.J., Fong, H.K.W., Simon, M.I., and Gilman, A.G. (1990). G_z, a guanine nucleotide-binding protein with unique biochemical properties. J.Biol.Chem. *265*, 2383-2390.

Cassel, D. and Selinger, Z. (1976). Catecholamine-stimulated GTPase activity in turkey erythrocyte membranes. Biochem.Biophys.Acta *452*, 538-551.

Cassel, D. and Selinger, Z. (1977). Mechanism of adenylate cyclase activation by cholera toxin: An inhibition of GTP hydrolysis at the regulatory site. Proc.Natl.Acad.Sci.U.S.A. *74*, 3307-3311.

Cassel, D. and Selinger, Z. (1978). Mechanism of adenylate cyclase activation through the β-adrenergic receptor: catecholamine-induced displacement of bound GDP by GTP. Proc.Natl.Acad.Sci.U.S.A. *75*, 4155-4159.

Chong, S., Mersha, F.B., Comb, D.G., Scott, M.E., Landry, D., Vence, L.M., Perler, F.B., Benner, J., Kucera, R.B., Hirvonen, C.A., Pelletier, J.J., Paulus, H., and Xu, M.Q. (1997). Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. Gene *192*, 271-281.

Christianson, T.W., Sikorski, R.S., Dante, M., Shero, J.H., and Hieter (1992). Multifunctional yeast high-copy-number shuttle vectors. Gene *110*, 119-122.

Clapham, D.E. and Neer, E.J. (1997). G protein beta/gamma subunits. Annu.Rev.Pharm.Toxicol. *37*, 167-203.

Clarke, L. and Carbon, J. (1980). Isolation of a yeast centromere and construction of functional small circular chromosomes. Nature *287*, 504-509.

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. Science *265*, 1405-1412.

Danchin, A., Guiso, N., Roy, A., and Ullmann, A. (1984). Identification of the Escherichia coli cya gene product as authentic adenylate cyclase. J.Mol.Biol. *175*, 403-408.

Dessauer, C.W., Chen-Goodspeed, M., and Chen, J. (2002). Mechanism of Galpha i-mediated Inhibition of Type V Adenylyl Cyclase. J.Biol.Chem *277*, 28823-28829.

Dessauer, C.W. and Gilman, A.G. (1996). Purification and characterization of a soluble form of mammalian adenylyl cyclase. J.Biol.Chem. *271*, 16967-16974.

Dessauer, C.W., Scully, T.T., and Gilman, A.G. (1997). Interactions of forskolin and ATP with the cytosolic domains of mammalian adenylyl cyclase. J.Biol.Chem. *272*, 22272-22277.

Dessauer, C.W., Tesmer, J.J., Sprang, S.R., and Gilman, A.G. (1999). The interactions of adenylate cyclases with P-site inhibitors. Trends.Pharmacol.Sci *20*, 205-210.

Dessauer, C.W., Tesmer, J.J.G., Sprang, S.R., and Gilman, A.G. (1998). Identification of a $G_{i\alpha}$ binding site on type V adenylyl cyclase. J.Biol.Chem. *273*, 25831-25839.

Doolittle, R. F., Abelson, J. N., and Simon, M. I. Computer methods for macromolecular sequence analysis. Methods in Enzymology (266). 1996. Academic Press. Methods in Enzymology. (GENERIC) Ref Type: Serial (Book,Monograph)

Doublie, S., Tabor, S., Long, A.M., Richardson, C.C., and Ellenberger (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution [see comments]. Nature *391*, 251-258.

Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. Biochemistry *6*, 1948-1954.

Eriksson, A.E., Baase, W.A., Zhang, X.J., Heinz, D.W., Blaber, M., Baldwin, E.P., and Matthews, B.W. (1992). Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. Science *255*, 178-183.

Evans, T.C.J., Benner, J., and Xu, M.Q. (1998). Semisynthesis of cytotoxic proteins using a modified protein splicing element. Protein Sci *7*, 2256-2264.

Evans, T.C.J., Benner, J., and Xu, M.Q. (1999a). The in vitro ligation of bacterially expressed proteins using an intein from Methanobacterium thermoautotrophicum. J.Biol.Chem. *274*, 3923-3926.

Evans, T.C.J. and Xu, M.Q. (1999b). Intein-mediated protein ligation: harnessing nature's escape artists. Biopolymers *51*, 333-342.

Faurobert, E., Ottobruc, A., Chardin, P., and Chabre, M. (1993). Tryptophan W207 in transducin T α is the fluorescence sensor of the G protein activation switch and is involved in the effector binding. EMBO J. *12*, 4191-4198.

Federman, A.D., Conklin, B.R., Schrader, K.A., Reed, R.R., and Bourne, H.R. (1992). Hormonal stimulation of adenylyl cyclase through G i-protein $\beta\gamma$ subunits. Nature *356*, 159-161.

Freissmuth, M. and Gilman, A.G. (1989). Mutations of $G_{s\alpha}$ designed to alter the reactivity of the protein with bacterial toxins: substitutions at Arg^{187} result in loss of GTPase activity. J.Biol.Chem. *264*, 21907-21914.

Getz, G., Levine, E., and Domany, E. (2000). Coupled two-way clustering analysis of gene microarray data. Proc.Natl.Acad.Sci U.S.A. *97*, 12079-12084.

Geyer, M., Schweins, T., Herrmann, C., Prisner, T., Wittinghofer, A., and Kalbitzer, H.R. (1996). Conformational transitions in p21ras and in its complexes with the effector protein Raf-RBD and the GTPase activating protein GAP. Biochemistry *35*, 10308-10320.

Gibbs, J.B., Sigal, I.S., Poe, M., and Scolnick, E.M. (1984). Intrinsic GTPase activity distinguishes normal and oncogenic *ras* p21 molecules. Proc.Natl.Acad.Sci.U.S.A. *81*, 5704-5708.

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

Graziano, M.P., Freissmuth, M., and Gilman, A.G. (1989a). Expression of $G_{s\alpha}$ in *Escherichia coli*: purification and properties of two forms of the protein. J.Biol.Chem. *264*, 409-418.

Graziano, M.P. and Gilman, A.G. (1989b). Synthesis in *Escherichia coli* of GTPase-deficient mutants of $G_{s\alpha}$. J.Biol.Chem. *264*, 15475-15482.

Hart, M.J., Jiang, X., Kozasa, T., Roscoe, W., Singer, W.D., Gilman, A.G., Sternweis, P.C., and Bollag, G. (1998). Direct stimulation of the guanine nucleotide exchange activity of p115 RhoGEF by $G\alpha_{13}$. Science *280*, 2112-2114.

Hatley, M.E., Gilman, A.G., and Sunahara, R.K. (2002). Expression, purification, and assay of cytosolic (catalytic) domains of membrane-bound mammalian adenylyl cyclases. Methods Enzymol. *345*, 127-140.

Hidalgo, P. and MacKinnon, R. (1995). Revealing the architecture of a K+ channel pore through mutant cycles with a peptide inhibitor. Science *268*, 307-310.

Huang, H., Chopra, R., Verdine, G.L., and Harrison, S.C. (1998). Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance [see comments]. Science *282*, 1669-1675.

Iñiguez-Lluhi, J.A., Simon, M.I., Robishaw, J.D., and Gilman, A.G. (1992). G protein βγ subunits synthesized in Sf9 cells: Functional characterization and the significance of prenylation of γ. J.Biol.Chem. *267*, 23409-23417.

Itoh, H. and Gilman, A.G. (1991). Expression and analysis of $G_{s\alpha}$ mutants with decreased ability to activate adenylyl cyclase. J.Biol.Chem. *266*, 16226-16231.

Jones, T.L.Z., Simonds, W.F., Merendino, J.J., Jr., Brann, M.R., and Spiegel, A.M. (1990). Myristoylation of an inhibitory GTP-binding protein α subunit is essential for its membrane attachment. Proc.Natl.Acad.Sci.U.S.A. *87*, 568-572.

Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. (1998). A bacterial two-hybrid system based on a reconstituted signal transduction pathway. Proc.Natl.Acad.Sci U.S.A. *95*, 5752-5756.

Karimova, G., Ullmann, A., and Ladant, D. (2000). A bacterial two-hybrid system that exploits a cAMP signaling cascade in Escherichia coli. Methods Enzymol *328*, 59-73.

Katada, T., Kusakabe, K., Oinuma, M., and Ui, M. (1987). A novel mechanism for the inhibition of adenylate cyclase via inhibitory GTP-binding proteins. J.Biol.Chem. *262*, 11897-11900.

Kaushik, N., Rege, N., Yadav, P.N., Sarafianos, S.G., Modak, M.J., and Pandey, V.N. (1996). Biochemical analysis of catalytically crucial aspartate mutants of human immunodeficiency virus type 1 reverse transcriptase. Biochemistry *35*, 11536-11546.

Kleuss, C. and Gilman, A.G. (1997). $G_{s\,\alpha}$ contains an unidentified covalent modification that increases its affinity for adenylyl cyclase. Proc.Natl.Acad.Sci.U.S.A. *94*, 6116-6120.

Kozasa, T. and Gilman, A.G. (1995). Purification of recombinant G proteins from Sf9 cells by hexa-histidine tagging of associated subunits. Characterization of α_{12} and inhibition of adenylyl cyclase by α_z. J.Biol.Chem. *270*, 1734-1741.

Krupinski, J., Coussen, F., Bakalyar, H.A., Tang, W.-J., Feinstein, P.G., Orth, K., Slaughter, C., Reed, R.R., and Gilman, A.G. (1989). Adenylyl cyclase amino acid sequence: possible channel- or transporter-like structure. Science *244*, 1558-1564.

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

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

Lan, K.L., Remmers, A.E., and Neubig, R.R. (1998). Roles of G(o)alpha tryptophans in GTP hydrolysis, GDP release, and fluorescence signals. Biochemistry *37*, 837-843.

Landis, C.A., Masters, S.B., Spada, A., Pace, A.M., Bourne, H.R., and Vallar, L. (1989). GTPase inhibiting mutations activate the α chain of G_s and stimulate adenylyl cyclase in human pituitary tumors. Nature *340*, 692-696.

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

Lerea, C.L., Somers, D.E., Hurley, J.B., Klock, I.B., and Bunt-Milam, A.H. (1986). Identification of specific transducin α subunits in retinal rod and cone photoreceptors. Science *234*, 77-80.

Li, Y., Korolev, S., and Waksman, G. (1998). Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of Thermus aquaticus DNA polymerase I: structural basis for nucleotide incorporation. EMBO J. *17*, 7514-7525.

Linder, M.E., Middleton, P., Hepler, J.R., Taussig, R., Gilman, A.G., and Mumby, S.M. (1993). Lipid modifications of G proteins: α subunits are palmitoylated. Proc.Natl.Acad.Sci.U.S.A. *90*, 3675-3679.

Linder, M.E., Pang, I.-H., Duronio, R.J., Gordon, J.I., Sternweis, P.C., and Gilman, A.G. (1991). Lipid modifications of G protein subunits. Myristoylation of $G_{o\alpha}$ increases its affinity for $\beta\gamma$. J.Biol.Chem. *266*, 4654-4659.

Liu, Y., Ruoho, A.E., Rao, V.D., and Hurley, J.H. (1997). Catalytic mechanism of the adenylyl and guanylyl cyclases: modeling and mutational analysis. Proc.Natl.Acad.Sci.U.S.A. *94*, 13414-13419.

Lockless, S. W. Networks of evloutionarily coupled residues: relating protein sequence, structure and function. 2002. University of Texas Southwestern Medical School at Dallas. (GENERIC)
Ref Type: Thesis/Dissertation

Lockless, S.W. and Ranganathan, R. (1999). Evolutionarily conserved pathways of energetic connectivity in protein families. Science *286*, 295-299.

Logothetis, D.E., Kurachi, Y., Galper, J., Neer, E.J., and Clapham, D.E. (1987). The βγ subunits of GTP-binding proteins activate the muscarinic K^{+} channel in heart. Nature *325*, 321-326.

Londos, C. and Wolff, J. (1977). Two distinct adenosine-sensitive sites on adenylate cyclase. Proc.Natl.Acad.Sci.U.S.A. *74*, 5482-5486.

Lowry, W.E. and Huang, X.Y. (2002). G Protein beta gamma subunits act on the catalytic domain to stimulate Bruton's agammaglobulinemia tyrosine kinase. J.Biol.Chem. *277*, 1488-1492.

Ma, Y.C., Huang, J., Ali, S., Lowry, W., and Huang, X.Y. (2000). Src tyrosine kinase is a novel direct effector of G proteins. Cell *102*, 635-646.

Maguire, M.E., Van Arsdale, P.M., and Gilman, A.G. (1976). An agonist-specific effect of guanine nucleotides on binding to the β-adrenergic receptor. Mol.Pharmacol. *12*, 335-339.

Masters, S.B., Miller, R.T., Chi, M.-H., Chang, F.-H., Beiderman, B., Lopez, N.G., and Bourne, H.R. (1989). Mutations in the GTP-binding site of $G_{s\alpha}$ alter stimulation of adenylyl cyclase. J.Biol.Chem. *264*, 15467-15474.

Muir, T.W., Sondhi, D., and Cole, P.A. (1998). Expressed protein ligation: a general method for protein engineering. Proc.Natl.Acad.Sci U.S.A. *95*, 6705-6710.

Mumby, S.M., Casey, P.J., Gilman, A.G., Gutowski, S., and Sternweis, P.C. (1990a). G protein γ subunits contain a 20-carbon isoprenoid. Proc.Natl.Acad.Sci.U.S.A. *87*, 5873-5877.

Mumby, S.M., Heuckeroth, R.O., Gordon, J.I., and Gilman, A.G. (1990b). G protein α subunit expression, myristoylation, and membrane association in COS cells. Proc.Natl.Acad.Sci.U.S.A. *87*, 728-732.

Muntz, K.H., Sternweis, P.C., Gilman, A.G., and Mumby, S.M. (1992). Influence of γ subunit prenylation on association of guanine nucleotide-binding regulatory proteins with membranes. Mol.Biol.Cell *3*, 49-61.

Natochin, M., Granovsky, A.E., and Artemyev, N.O. (1998). Identification of effector residues on photoreceptor G protein, transducin. J.Biol.Chem. *273*, 21808-21815.

Noren, C.J., Wang, J., and Perler, F.B. (2000). Dissecting the Chemistry of Protein Splicing and Its Applications. Angew.Chem.Int.Ed.Engl. *39*, 450-466.

Osawa, S. and Johnson, G.L. (1991). A dominant negative gas mutant is rescued by secondary mutation of the α chain amino terminus. J.Biol.Chem. *266*, 4673-4676.

Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. Protein Sci *4*, 2411-2423.

Pai, E.F., Krengel, U., Petsko, G.A., Goody, R.S., Kabsch, W., and Wittinghofer, A. (1990). Refined crystal structure of the triphosphate conformation of H-*ras* p21 at 1.35 Å resolution: Implications for the mechanism of GTP hydrolysis. EMBO J. *9*, 2351-2359.

Pelletier, J.N., Arndt, K.M., Pluckthun, A., and Michnick, S.W. (1999). An in vivo library-versus-library selection of optimized protein-protein interactions. Nat.Biotechnol. *17*, 683-690.

Perler, F.B. (1998). Protein splicing of inteins and hedgehog autoproteolysis: structure, function, and evolution. Cell *92*, 1-4.

Perler, F.B. and Adam, E. (2000). Protein splicing and its applications. Curr.Opin.Biotechnol. *11*, 377-383.

Pfeuffer, E., Mollner, S., and Pfeuffer, T. (1985). Adenylate cyclase from bovine brain cortex: purification and characterization of the catalytic unit. EMBO J. *4*, 3675-3679.

Pfeuffer, T. and Metzger, H. (1982). 7-*O*-Hemisuccinyl-deacetyl forskolin-Sepharose: a novel affinity support for purification of adenylate cyclase. FEBS Lett. *146*, 369-375.

Pitcher, J.A., Touhara, K., Payne, E.S., and Lefkowitz, R.J. (1995). Pleckstrin homology domain-mediated membrane association and activation of the β-adrenergic receptor kinase requires coordinate interaction with G(beta-gamma) subunits and lipid. J.Biol.Chem. *270*, 11707-11710.

Rall, T.W. and Sutherland, E.W. (1958). Formation of a cyclic adenine ribonucleotide by tissue particles. J.Biol.Chem. *232*, 1065-1091.

Rall, T.W., Sutherland, E.W., and Berthet, J. (1957). The relationship of epinephrine and glucagon to liver phosphorylase. IV. Effect of epinephrine and glucagon on the reactivation of phosphorylase in liver homogenates. J.Biol.Chem. *224*, 463-475.

Ranganathan, R., Lewis, J.H., and MacKinnon, R. (1996). Spatial localization of the K+ channel selectivity filter by mutant cycle-based structure analysis. Neuron *16*, 131-139.

Rodbell, M., Birnbaumer, L., Pohl, S.L., and Krans, H.M.J. (1971a). The glucagon-sensitive adenyl cyclase system in plasma membranes of rat liver. V. An obligatory role of guanyl nucleotides in glucagon action. J.Biol.Chem. *246*, 1877-1882.

Rodbell, M., Krans, H.M.J., Pohl, S.L., and Birnbaumer, L. (1971b). The glucagon-sensitive adenyl cyclase system in plasma membranes of rat liver. IV. Effects of guanyl nucleotides on binding of ^{125}I-glucagon. J.Biol.Chem. *246*, 1872-1876.

Rose, M.D., Winston, F., and Hieter, P. (1990). Methods in Yeast Genetics (Cold Spring Harbor Laboratory Press).

Roy, A. and Danchin, A. (1981). Restriction map of the cya region of the Escherichia coli K12 chromosome. Biochimie *63*, 719-722.

Roy, A. and Danchin, A. (1982). The cya locus of Escherichia coli K12: organization and gene products. Mol.Gen.Genet. *188*, 465-471.

Roy, A., Haziza, C., and Danchin, A. (1983). Regulation of adenylate cyclase synthesis in Escherichia coli: nucleotide sequence of the control region. EMBO J. *2*, 791-797.

Sambrook, J. and Russell, D.W. (2001). Molecular Cloning (cold springs harbor: cold springs harbor laboratory press).

Sarvazyan, N.A., Remmers, A.E., and Neubig, R.R. (1998). Determinants of gi1alpha and beta gamma binding. Measuring high affinity interactions in a lipid environment using flow cytometry. J.Biol.Chem. *273*, 7934-7940.

Schreiber, G. and Fersht, A.R. (1995a). Energetics of protein-protein interactions: analysis of the barnase-barstar interface by single mutations and double mutant cycles. J Mol.Biol *248*, 478-486.

Schreiber, G. and Fersht, A.R. (1995b). Energetics of protein-protein interactions: analysis of the barnase-barstar interface by single mutations and double mutant cycles. J Mol.Biol *248*, 478-486.

Seamon, K.B., Padgett, W., and Daly, J.W. (1981). Forskolin: unique diterpene activator of adenylate cyclase in membranes and in intact cells. Proc.Natl.Acad.Sci.U.S.A. *78*, 3363-3367.

Severinov, K. and Muir, T.W. (1998). Expressed protein ligation, a novel method for studying protein-protein interactions in transcription. J.Biol.Chem. *273*, 16205-16209.

Shao, Y. and Kent, S.B. (1997). Protein splicing: occurrence, mechanisms and related phenomena. Chem.Biol. *4*, 187-194.

Shorr, R.G.L., Lefkowitz, R.J., and Caron, M.G. (1981). Purification of the β-adrenergic receptor. Identification of the hormone binding subunit. J.Biol.Chem. *256*, 5820-5826.

Simonds, W.F., Butrynski, J.E., Gautam, N., Unson, C.G., and Spiegel, A.M. (1991). G-protein βγ dimers. Membrane targeting requires subunit coexpression and intact γ C-A-A-*X* domain. J.Biol.Chem. *266*, 5363-5366.

Smigel, M.D. (1986). Purification of the catalyst of adenylate cyclase. J.Biol.Chem. *261*, 1976-1982.

Smit, M.J. and Iyengar, R. (1998). Mammalian adenylyl cyclases. Adv.Second Messenger Phosphoprotein Res. *32*, 1-21.

Smrcka, A.V., Hepler, J.R., Brown, K.O., and Sternweis, P.C. (1991). Regulation of polyphosphoinositide-specific phospholipase C activity by purified G_q. Science *251*, 804-807.

Smrcka, A.V. and Sternweis, P.C. (1993). Regulation of purified subtypes of phosphatidylinositol-specific phospholipase C β by G protein α and βγ subunits. J.Biol.Chem. *268*, 9667-9674.

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

Southworth, M.W., Amaya, K., Evans, T.C., Xu, M.Q., and Perler, F.B. (1999). Purification of proteins fused to either the amino or carboxy terminus of the Mycobacterium xenopi gyrase A intein. BioTechniques *27*, 110-120.

Stephens, L., Smrcka, A.V., Cooke, F.T., Jackson, T.R., Sternweis, P.C., and Hawkins, P.T. (1994). A novel phosphoinositide 3 kinase activity in myeloid-derived cells is activated by G protein βγ subunits. Cell *77*, 83-93.

Sternweis, P.C., Northup, J.K., Smigel, M.D., and Gilman, A.G. (1981). The regulatory component of adenylate cyclase: purification and properties. J.Biol.Chem. *256*, 11517-11526.

Strathmann, M., Wilkie, T.M., and Simon, M.I. (1989). Diversity of the G-protein family: sequences from five additional α subunits in the mouse. Proc.Natl.Acad.Sci.U.S.A. *86*, 7407-7409.

Suel, G.M., Lockless, S.W., Wall, M.A., and Ranganathan, R. (2003). Evolutionarily conserved networks of residues mediate allosteric communication in proteins. Nat.Struct.Biol. *10*, 59-69.

Sunahara, R.K., Beuve, A., Tesmer, J.J., Sprang, SR, Garbers, D.L., and Gilman, A.G. (1998). Exchange of substrate and inhibitor specificities between adenylyl and guanylyl cyclases. J.Biol.Chem. *273*, 16332-16338.

Sunahara, R.K., Dessauer, C.W., and Gilman, A.G. (1996). Complexity and diversity of mammalian adenylyl cyclases. Annu.Rev.Pharm.Toxicol. *36*, 461-480.

Sunahara, R.K., Dessauer, C.W., Whisnant, R.E., Kleuss, C., and Gilman, A.G. (1997a). Interaction of $G_{s\alpha}$ with the cytosolic domains of mammalian adenylyl cyclase. J.Biol.Chem. *272*, 22265-22271.

Sunahara, R.K., Tesmer, J.J., Gilman, A.G., and Sprang, S.R. (1997b). Crystal structure of the adenylyl cyclase activator $G_{s\alpha}$. Science *278*, 1943-1947.

Sutherland, E.W. and Rall, T.W. (1957). The properties of an adenine ribonucleotide produced with cellular particles, ATP, Mg++, and epinephrine or glucagon. J.Am.Chem.Soc. *79*, 3608

Sutherland, E.W., Rall, T.W., and Menon, T. (1962). Adenyl cyclase. I. Distribution, preparations, and properties. J.Biol.Chem. *237*, 1220-1227.

Tang, W.-J. and Gilman, A.G. (1991). Type-specific regulation of adenylyl cyclase by G protein βγ subunits. Science *254*, 1500-1503.

Tang, W.-J. and Gilman, A.G. (1995). Construction of a soluble adenylyl cyclase activated by $G_{s\alpha}$ and forskolin. Science *268*, 1769-1772.

Tang, W.-J., Krupinski, J., and Gilman, A.G. (1991). Expression and characterization of calmodulin-activated (Type-I) adenylyl cyclase. J.Biol.Chem. *266*, 8595-8603.

Tang, W.-J., Stanzel, M., and Gilman, A.G. (1995). Truncation and alanine-scanning mutants of type I adenylyl cyclase. Biochemistry *34*, 14563-14572.

Taussig, R. and Gilman, A.G. (1995). Mammalian membrane-bound adenylyl cyclases. J.Biol.Chem. *270*, 1-4.

Taussig, R., Iñiguez-Lluhi, J., and Gilman, A.G. (1993a). Inhibition of adenylyl cyclase by $G_{i\alpha}$. Science *261*, 218-221.

Taussig, R., Quarmby, L.M., and Gilman, A.G. (1993b). Regulation of purified type-I and type-II adenylyl cyclases by G protein βγ subunits. J.Biol.Chem. *268*, 9-12.

Taussig, R., Tang, W.-J., Hepler, J.R., and Gilman, A.G. (1994). Distinct patterns of bidirectional regulation of mammalian adenylyl cyclases. J.Biol.Chem. *269*, 6093-6100.

Taylor, S.J., Chae, H.Z., Rhee, S.G., and Exton, J.H. (1991). Activation of the β1 isozyme of phospholipase C by α subunits of the G_q class of G proteins. Nature *350*, 516-518.

Tesmer, J.J., Sunahara, R.K., Fancy, D.A., Gilman, A.G., and Sprang, S.R. (2002). Crystallization of complex between soluble domains of adenylyl cyclase and activated Gs alpha. Methods Enzymol. *345*, 198-206.

Tesmer, J.J., Sunahara, R.K., Johnson, R.A., Gosselin, G., Gilman, A.G., and Sprang, S.R. (1999). Two-metal-Ion catalysis in adenylyl cyclase. Science *285*, 756-760.

Tesmer, J.J.G., Berman, D.M., Gilman, A.G., and Sprang, S.R. (1997a). Structure of RGS4 bound to AlF_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 $G_{s\alpha}$.GTPgammaS. Science *278*, 1907-1916.

Thomas, T.C., Schmidt, C.J., and Neer, E.J. (1993). G protein ao subunit: mutation of conserved cysteines identifies a subunit contact surface and alteres GDP affinity. Proc.Natl.Acad.Sci.U.S.A. *90*, 10295-10299.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic.Acids.Res. *22*, 4673-4680.

Ueda, N., Iñiguez-Lluhi, J.A., Lee, E., Smrcka, A.V., Robishaw, J.D., and Gilman, A.G. (1994). G protein βγ subunits: Simplified purification and properties of novel isoforms. J.Biol.Chem. *269*, 4388-4395.

Vaara, M. (1992). Agents that increase the permeability of the outer membrane. Microbiol.Rev. *56*, 395-411.

Wall, M.A., Coleman, D.E., Lee, E., Iniguez-Lluhi, J.A., Posner, B.A., Gilman, A.G., and Sprang, S.R. (1995). The structure of G protein heterotrimer $G_{i\alpha1}\beta_1Gamma_2$. Cell *83*, 1047-1058.

Weinstein, L.S., Shenker, A., Gejman, P.V., Merino, M.J., Friedman, E., and Spiegel, A.M. (1991). Activating mutations of the stimulatory G protein in the McCune-Albright syndrome. N.Engl.J.Med. *325*, 1688-1695.

Whisnant, R.E., Gilman, A.G., and Dessauer, C.W. (1996). Interaction of the two cytosolic domains of mammalian adenylyl cyclase. Proc.Natl.Acad.Sci.U.S.A. *93*, 6621-6625.

Whiteway, M., Hougan, L., Dignard, D., Thomas, D.Y., Bell, L., Saari, G.C., Grant, F.J., O'Hara, P., and MacKay, V.L. (1989). The *STE4* and *STE18* genes of yeast encode potential β and τ subunits of the mating factor receptor-coupled G protein. Cell *56*, 467-477.

Woody, A.Y., Eaton, S.S., Osumi-Davis, P.A., and Woody, R.W. (1996). Asp537 and Asp812 in bacteriophage T7 RNA polymerase as metal ion-binding sites studied by EPR, flow-dialysis, and transcription. Biochemistry *35*, 144-152.

Xu, R., Ayers, B., Cowburn, D., and Muir, T.W. (1999). Chemical ligation of folded recombinant proteins: segmental isotopic labeling of domains for NMR studies. Proc.Natl.Acad.Sci U.S.A. *96*, 388-393.

Yamane, H.K., Farnsworth, C.C., Xie, H., Howald, W., Fung, B.K.K., Clarke, S., Gelb, M.H., and Glomset, J.A. (1990). Brain G protein τ subunits contain an all-*trans*-geranylgeranyl-cysteine methyl ester at their carboxyl termini. Proc.Natl.Acad.Sci.U.S.A. *87*, 5868-5872.

Yan, S.Z., Hahn, D., Huang, Z.H., and Tang, W.-J. (1996). Two cytoplasmic domains of mammalian adenylyl cyclase form a $G_{s\alpha}$- and forskolin-activated enzyme in vitro. J.Biol.Chem. *271*, 10941-10945.

Yan, S.Z., Huang, Z.H., Rao, V.D., Hurley, J.H., and Tang, W.J. (1997a). Three discrete regions of mammalian adenylyl cyclase form a site for $G_{s\alpha}$ activation. J.Biol.Chem. *272*, 18849-18854.

Yan, S.Z., Huang, Z.H., Shaw, R.S., and Tang, W.-J. (1997b). The conserved asparagine and arginine are essential for catalysis of mammalian adenylyl cyclase. J.Biol.Chem. *272*, 12342-12349.

Zhang, G., Liu, Y., Ruoho, A.E., and Hurley, J.H. (1997). Structure of the adenylyl cyclase catalytic core. Nature *386*, 247-253.

Zimmermann, G., Zhou, D., and Taussig, R. (1998). Mutations uncover a role for two magnesium ions in the catalytic mechanism of adenylyl cyclase. J.Biol.Chem. *273*, 19650-19655.

Zimmermann, G., Zhou, D., and Taussig, R. (1999). Activating mutation of adenylyl cyclase reverses its inhibition by G proteins. Mol.Pharmacol. *56*, 895-901.

VITAE

Mark Edward Hatley was born on August 27, 1971 in Edmond, Oklahoma, the son of Michael Edward Hatley and Jo Anne Hatley. As a child he lived in Edmond, Oklahoma. He graduated from Edmond Memorial High School in 1990. He obtained a Bachelor of Science degree in chemistry from the University of Central Oklahoma in 1995. While an undergraduate, he had the opportunity to work with Fakhrildeen Albahadily, Ph.D. at the University of Central Oklahoma and Richard A. J. O'Hair, Ph.D. at Kansas State University. Mark entered the Medical Scientist Training Program at the University of Texas Southwestern Medical School at Dallas in 1995. After completing two years of medical school, he began work towards his Ph.D. degree in the laboratory of Alfred G. Gilman, M.D., Ph.D. On June 12, 1999, he married Ann Elizabeth Melvin. A daughter, Mary Elizabeth Hatley, was born June 12, 2002.

Current Address: 4318 Merrell Road
Dallas, Texas 75229