Activation of G Protein-Coupled Receptors and Heterotrimeric G Proteins

by

Brian Thomas DeVree

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in the University of Michigan
2013

Doctoral Committee:

Associate Professor Roger K. Sunahara, Chair
Professor Charles L. Brooks, III
Associate Professor Jason E. Gestwicki
Professor Richard R. Neubig

DEDICATION

To my parents, who taught me to love learning.

ACKNOWLEDGMENTS

I am deeply indebted to Roger Sunahara for his support over many years. He hired me as green scientist recently returned from a challenging Peace Corps assignment and it was only through his mentorship, passion for science, and patient encouragement that I am now leaving his lab as a Ph.D. I look forward to many more years of working with him as a collaborator, colleague, and friend.

I am also sincerely gracious for the aid of my committee; Drs. Charles L. Brooks, Jason Gestwicki, and Rick Neubig. Their thoughtful suggestions throughout the process both challenged me and enriched my understanding in many ways.

I am lucky to have had the best labmates that one could ever ask for: Remy Brim, Diane Calinski, Elín Edwald, Faiza Haji-Abdi, Adam Kuszak, Jacob Mahoney, Søren Midtgaard, Diwa Narasimhan, Joey Nichols, Prahnesh Venkataraman, Nico Villanueva, Gisselle Vélez-Ruiz, and Matt Whorton. You guys were always available, whether it be for an in-depth discussion of technical details or grandiose ideas, encouragement after a failed experiment, or just hanging out and goofing around. In particular, I want to acknowledge Sanna-Paula Pehkonen for her tireless and excellent work in making several rHDL technology improvements along with me. Her enthusiasm kept me inspired and her intelligent questions kept me on my toes.

I wouldn't have been able to get very far without the support of my friends. Many thanks to the 2008 cohort and to all of the ChemBio program for always being up for a good time. Laura Matney, Chris Scannell, Matt Thayer, Chris Tom, and Ningkun Wang, many thanks for all the nights of drinks and discussion, where the real thinking usually got done.

To my family, you have my most heartfelt appreciation for your immeasurable support. My grandparents, sister, brother-in-law, aunts, uncles, and cousins; you all have made me the person I am today and were unwavering in your encouragement and your help in becoming a scientist.

To my parents, there are no words that are adequate to express my gratitude. You sacrificed much so that I could be given a good education. You taught me how to be creative in thought, curious about the world, rigorous in method, and ambitious with my goals. Most importantly, I always knew that I was dearly loved. Thank you.

And last, but certainly not least, I thank Laura Cesa. More than anybody else, you bore the brunt of the actual effects of writing. Throughout all, you supported me, cheered me, helped me, and loved me. I am truly blessed to have found you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

ABBREVIATIONS

Note: text in brackets (e.g. [number]) denote a variable part of an abbreviation's name. The meaning of the bracketed text is included in the explanation of the abbreviation.

[^{3}H]DHAP – (-)-[^{3}H]-dihydroalprenolol, a radiolabeled antagonist for the β_2AR.

AC – Adenylyl cyclase, an enzyme that produces the second messenger cAMP.

ALP – Alprenolol, an antagonist for the β_2AR.

apoAI – Apolipoprotein A-I, the major protein found in high density lipoporteins.

β_2AR – β_2-adrenergic receptor, a hormone-binding GPCR.

BSA – Bovine serum albumin, a protein that is commonly added to *in vitro* systems to decrease non-specific binding of reagents.

$C_{[12]}E_{[10]}$ – Polyoxyethylene-10-dodecyl ether, an inexpensive class of detergents. The number following the "C" denotes the length of the hydrocarbon chain, and the number following the "E" denotes the number of polyoxyethelene repeats in the head group.

cAMP – Cyclic adenosine monophosphate, a second messenger.

CDR – Complementarity determining region, the variable portions of a camilid antibody.

CHAPS – 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate, a zwitterionic detergent commonly used in protein biochemistry.

DAG – Diacylglycerol, a lipophilic second messenger.

DDM – *n*-Dodecyl β-D-maltoside, a detergent commonly used for membrane protein work.

DMSO – Dimethyl sulfoxide, a solvent.

ECL[number] – Extracellular Loop [number], a designation for the extracellular-facing loop regions in-between the transmembrane helices of a GPCR.

EPI – (-)-epinephrine, an endogenous agonist for the β_2AR.

EPR – Electron paramagnetic resonance, a method that measures chemical shifts in the local environment of unpaired electrons.

G protein – Guanine nucleotide binding protein, a signaling protein that binds GDP or GTP

GAP – GTPase-accelerating protein, a protein that enhances the rate of GTP hydrolysis of G proteins.

GDP – Guanosine diphosphate, a common nucleotide in the cell.

GPCR – G protein coupled receptor, a major class of cell-surface transmembrane receptors.

GRK – G protein-coupled receptor kinase, a class of kinases that phosphorylate activated GPCRs.

GTP – Guanosine triphosphate, a common nucleotide in the cell.

GTPase – Guanosine triphosphate hydrolase, a class of enzymes that convert GTP to GDP and P_i via hydrolysis.

GTPγS – Guanosine 5'-O-[γ-thio]triphosphate), a GTP mimic that is slowly or not-at-all hydrolyzed by many GTPase enzymes.

H8 – Helix 8, one of the conserved α-helices found in GPCR.

HEPES – (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), an aqueous buffering reagent.

Hi5 or High5 or High Five – BTI-TN-5B1-4 insect ovary cells isolated from *Trichopulsia ni* and commonly used for recombinant protein overexpression.

HPLC – High performance liquid chromatography, a liquid-phase separation method.

ICL[number] – Intercellular Loop [number], a designation for the cytosol-facing loop regions in-between the transmembrane helices of a GPCR.

IP3 – Inositol 1,4,5-trisphosphate, a soluble second messenger.

ISO – (-)-isoproterenol, a synthetic agonist for the β_2AR.

LCP – Lipidic cubic phase, a highly curved hydrated lipid phase.

MAG – Monoacylglycerol, a class of lipid molecules commonly used to make lipidic cubic phase.

MES – 2-(N-morpholino)ethanesulfonic acid, an aqueous buffering reagent.

MNG – Maltose neopentyl glycol detergent, a class of detergent used for membrane protein work.

NB[number] – Nanobody [number], identifier for a camilid antibody fragment clone.

Ni-NTA resin – Nickel nitrilotriacetic acid resin, a chromatography resin that binds His-tagged proteins.

PEG [number] – Polyethylene glycol [approximate molecular weight], a soluble polymer commonly used in crystallography as a precipitant.

P_i – Inorganic phosphate, the free phosphate ion.

PLCβ – Phospholipase C isoform β, an enzyme that produces the second messengers IP3 and DAG.

POPC – 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine, a common zwitterionic phospholipid.

POPG – 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1'-rac-glycerol), a common negatively charged phospholipid.

PPase – Protein phosphatase, a general term for an enzyme that removes a phosphate group from a protein or peptide substrate

PP_i – Inorganic pyrophsophate, the free pyrophosphate ion.

RGS – Regulator of G protein signaling protein, a class of GAPs for heterotrimeric G proteins.

rHDL – Recombinant high density lipoprotein, a platform for studying integral membrane proteins in solution.

SDS-PAGE – Sodium dodecyl sulfate polyacrylamide gel electrophoresis, a standard technique for separation of proteins by molecular weight.

Sf9 – *Spodoptera frugiperda* insect ovary cells, commonly used for recombinant protein overexpression.

TCEP – Tris(2-carboxyethyl)phosphine, a non-sulfur reducing agent.

TM[number] – Transmembrane domain [number], a designation for the transmembrane helices of a GPCR.

ABSTRACT

G protein-coupled receptors (GPCRs) are an important class of cell-surface transmembrane receptors that pass an activation signal to the interior of the cell through heterotrimeric G proteins. In this work, we study the human β_2-adrenergic receptor (β_2AR) and stimulatory G protein (Gs) as examples in order to understand the molecular basis of this signal transfer event. We solved a 3.2 Å crystal structure of β_2AR and Gs in a nucleotide-free, intermediate signaling complex, revealing the interaction between the proteins at atomic resolution. The structure was consistent with previous biochemical knowledge, but also revealed several previously unknown features of the activation process. We used deuterium/hydrogen exchange and electron microscopy in order to find regions in the complex that change conformation during the activation process. These regions are highly conserved within the GPCR and G protein families, and his work shows the central role that they play in the process of GPCR signal transduction. The binding of drugs to the receptor in the fully activated state, as seen in the β_2AR-Gs complex, was also characterized by radioligand and antibody fragment binding. A full kinetic model was developed for drug binding to the activated receptor which demonstrated how the ligand is held very tightly in the receptor binding pocket. This tight ligand binding can be relieved by the addition of GDP, demonstrating a direct allosteric link between the G protein nucleotide binding site and the receptor ligand binding site. Overall, this work demonstrates how the GPCR signal transduction machinery operates in high-resolution structural, kinetic, and pharmacological detail. It advances our understanding of how GPCRs and G proteins pass a signal across the cellular membrane.

CHAPTER 1
INTRODUCTION AND LITERATURE REVIEW

1.1 – Introduction and importance.

G protein coupled receptors (GPCRs) are a large and diverse group of cell-surface receptors which can bind a variety of ligands, including hormones, metabolites, ions, odorants, therapeutic drugs and even photons. These receptors share a common topology of seven transmembrane helices and an almost universal ability to activate one or more subtypes of heterotrimeric G proteins. GPCR-G protein signaling cascades are known to play extremely important roles in basic physiological functions such as sensatory perception, neuronal communication, hormone response, and development.[1] These receptors have long been recognized to be highly druggable, to the extent that of drugs that act at known targets, the single largest group is the GPCRs, with about 27% of the total (in 2005).[2]

The work contained in this dissertation concerns the molecular structure, function, and dynamics of how the prototypical GPCR and G protein pair, the β_2-adrengeric receptor (β_2AR) and stimulatory G protein (Gs), interact during the activation process. This complex of the receptor and G protein is the crucial intermediate state where the information contained in the chemical structure of the receptor-bound agonist is passed across the cellular membrane to the G protein, the first step of a cellular signaling event. This process is studied with modern molecular and structural methods as well as from a pharmacological perspective, leading to a much fuller understanding of the G protein activation process.

1.2 – Evolution of receptor theory.

The discovery of GPCRs is deeply intertwined in the discovery of receptors in general and the development of the pharmacological sciences. Although many, many scientists contributed research pointing towards the existence of the proteins that we now call receptors, the idea that compounds act at specific sites of action on cells can be traced to the work of both Paul Ehrlich and John N. Langley in the late 1870's. Langley did so in the context of studying the action of pilocarpine and atropine, now known as a muscarinic acetylcholine receptor agonist and antagonist, respectively, on the salivary secretion of the cat sub-maxillary gland. Langley's observations on the competing activities of the two drugs led him to suspect that they competed via the laws of mass action for sites of action on the gland.[3] Unfortunately, acceptance of Langley's idea was hampered by the difficulty of resolving questions related to the complexity of organ and tissue systems, such as whether the drugs were acting upon the nerve cells going to the gland or to these proposed sites of action on the gland itself. For this reason, Paul Ehrlich is usually credited as the first person to seriously study the binding of biologically active molecules to specific sites of interaction,[4, 5] as his system of study ended up being much more amenable to isolation from the organism it came from. Around the turn of the century, he showed that the toxicity of diphtheria toxin could be neutralized with antibodies produced against the toxin, known as “anti-toxins” initially, even when the reaction was performed *in vitro.* Ehrlich followed up this observation with extensive research on the nature of the neutralization reaction. The work led him to propose the theory that the toxin molecules interacted with a specific antitoxin molecule, in the same manner that chemicals can bind and react with one another. In contrast, most other theories about how biologically active compounds worked at the time, including toxins, assumed that the effect of the compound was generalized across a tissue, and that the tissue-specific responses were due to how the compound was differentially distributed in the animal. Ehrlich’s work was clearly at odds with this distributive explanation, and in his extremely productive career he built up incontrovertible evidence that many different kinds of biologically active compounds bind to specific molecules in order to elicit their function,[6] opening up the possibility

of studying these responsive molecules as well as the compounds themselves. It was not Ehrlich, but Langley that first started calling these responsive molecules a “receptive substance” (which Ehrlich quickly shortened to just “receptor”).[4] Ehrlich was honored with the Nobel Prize in Physiology or Medicine in 1908 for these very important discoveries.

John Langley continued research to advance receptor theory while studying both nicotinic and muscarinic acetylcholine systems during the same time Ehrlich was working with his toxins and anti-toxins. Langley carefully studied how the binding of one drug to a receptor could block the action of another one, establishing evidence that the competing effects of such drugs on a receptor was dependent on both the affinity of each drug for the receptor and the relative concentrations of the two drugs.[7] His qualitative descriptions of drug saturation and competition were upgraded to a quantitative mathematical description by A.J. Clark and J.H. Gaddum in the mid 1920s. Clark applied the laws of mass action to the nicotinic receptor system,[8, 9] showing that the occupancy of the receptor obeyed the same laws that Irving Langmuir determined for the binding of gas to solid surfaces.[10] (Clark was apparently aware of A.V. Hill's work on the same nicotinic receptor system in 1909,[11] but did not either mention or notice that Hill had also derived the Langmuir equation to describe his response curve. His colleagues apparently didn't notice it either, so the credit commonly goes to Clark.) Gaddum extended Clark’s full mathematical treatment to the competition between drugs that bound the same site.[12] Over the next several decades, the concept of efficacy was built when differences between the occupancy of receptors predicted from Clark and Gaddum’s basic description of drug binding and the biological response of the system were rigorously compared.[4] This effort included the work of such prominent pharmacologists as E.J. Ariëns,[13] R.P. Stephenson,[14] and R.F. Furchgott.[15] In this way, over the span of about half a century the idea of a drug receptor was developed from the basic concept of Erhlich’s specific anti-toxin molecules to a physical entity whose behavior could be quantitatively described and modeled.[16] Although the researches did not know it at the time, GPCRs were some of the most commonly used receptors for performing these studies, with the α- and β-adrenergic and muscarinic acetylcholine receptors

being the most prominent examples, along with the non-GPCR nicotinic acetylcholine receptors.

The next major advances in receptor theory were inspired from work in enzymology with hemoglobin. A.V. Hill had introduced an equation in 1910 (now known at the Hill equation) that described the binding of oxygen to hemoglobin, which showed a marked degree of cooperatively in binding.[17] His observation and equation were given their modern interpretation much later by F. Jacob, J. Monod, J. Wyman, and J.P. Changeux in 1963 and 1965 when they described hemoglobin as a cooperative, allosteric protein. In their model, the individual subunits can exist in two different conformational states, one of which binds oxygen at a higher affinitiy.[18, 19] This model, which is known today as the Monod-Wyman-Changeux (MWC) model of allostery, made the important advance of allowing the subunits to adopt both conformational states even in the absence of ligand. Several years before the MWC model was presented, J. Del Castillo and B. Katz had also described the activity of the nicotinic acetylcholine receptor ion channels using a two-state model, where acetylcholine binding allowed the channel to transition from an open to a closed state, but they did not allow the unligaded channel to transition to an open state.[20] That development was made by A. Karlin in 1967, who applied the full MWC model to the nicotinic receptor.[21] This was the general state of the field when the G proteins were found, a discovery which also revealed the general nature of the GPCR signal transduction system.

1.3 – Discovery of G proteins and GPCRs.

The discovery of GPCRs as a distinct type of receptor was made possible by the earlier discovery of cyclic AMP (cAMP). In 1957, Earl Sutherland and Theodore Rall identified a soluble molecule that acted to stimulate glycogenolysis in dog liver homogenates.[22] This molecule was produced when the hormones epinephrine and glucagon interacted with some unknown receptor the membrane portion of the homogenate, and was identified a year later as cAMP.[23] The field started to study both the unknown receptor and the cAMP response, and about a decade later Martin Rodbell demonstrated that the cAMP generating enzyme and the receptors that

bound the hormones were most likely separate entities.[24] Soon after, he also demonstrated that GTP was somehow involved in the transmission of the hormone signal from the receptor to the enzyme that stimulated cAMP production.[25]

In the early to mid 1970's, Alfred Gilman started the search for Rodbell's GTP-dependent transmission protein. His search was greatly aided by two new scientific developments. At this point, good antagonists for the β-adrenergic receptors had been developed, most prominently through the work of James Black.[26] Gerald Aurbach created a radiolabeled version of one of these drugs and showed it could be used to directly measure the presence of the receptors.[27] Also, a pair of S49 lymphoma cell lines that died from cAMP elevation were isolated by Henry Bourne, Phil Coffino, and Gordon Tomkins; one cell line had an intact cAMP response to the β-adrenergic receptor agonist isoproterenol, and one was a mutant where the response was broken.[28] In a collaborative effort between the labs of Gilman and Kenneth Melman, Paul Insel used these new tools to show that the mutant lymphoma cells still had perfectly functional β-adrenergic receptors.[29] Soon after, Gilman's postdoctoral fellow Elliot Ross demonstrated that the mutant cells also retained functional adenylyl cyclase enzyme as well, but instead were missing a novel GTP binding protein that could be resupplied using non-mutant, detergent solubilized membrane extracts.[30] At this time, Ross also described a model of receptor action that was to first to recognize that agonist binding to the receptor was negatively cooperative with binding of guanine nucleotides. This model became known as the "ternary complex model" of receptor function,[31] and is described in more detail below in section 1.6. In 1980, shortly after the finding that the GTP binding proteins existed, Gilman's postdoctoral students John Northup and Paul Sternweis purified the protein,[32] which today is known as the G protein Gs. Gilman had proven the existence of Rodbell's GTP-dependent transmission protein, eventually earning both him and Rodbell a Nobel Prize in 1995.

The discovery of the G proteins and their relationship with the GPCRs happened at the beginning of the modern era of molecular biology. Over the 1980s, the ability to clone genes, sequence them, and produce recombinant proteins expanded from the

few specialized labs that developed the technology to biochemistry labs in general. By the end of the decade, most of the genes for the G proteins and the adenylyl cyclases were found through the work of Gilman, Mel Simon, Randy Reed, and other colleagues. In 1986, the β_2-adrenergic receptor was also cloned by Brian Kobilka from Robert Lefkowitz's lab and Richard Dixon from Merk,[33] using peptide sequences from a receptor purification method worked out several years earlier by Marc Caron.[34] The receptor's sequence showed that both the photoreceptor rhodopsin and the β_2-adrenergic receptor, the prototypical GPCR, shared the same 7 transmembrane domain topology. This was a major surprise, as it was generally assumed at the time that the hormone receptors and the photoreceptors were not related to one another.[5] Following this revelation, other 7 transmembrane receptors were quickly cloned, and the work that followed in the 1990s established our modern understanding of the GPCR family and the workings of the classical G-protein mediated signaling pathways.

1.4 – Classes of GPCRs and G proteins.

GPCRs have been classified by several different methods, but the most well recognized system divides the family into 6 classes: Class A (Rhodopsin-like), Class B (Secretin-like), Class C (Metabotropic glutamate-like), Class D (Fungal pheromone), Class E (cAMP), Class F (Frizzled/smoothened).[35] Only classes A, B, C, and F are found in vertebrates, with approximately 800 different receptors in humans.[36] It is clear from sequence homology that most of these classes are very old, from before nematode and chordate lineages split.[37] Classes A, B, and C are all known to signal primarily through heterotrimeric G proteins. Class F instead is associated with signaling through the Wnt pathway, although evidence for some level of interaction with G proteins has been observed.[38, 39] Class A is the largest class by far, including about 270 non-olfactory and 400 olfactory receptors in humans. This class generally binds small to medium size ligands that fully or partially interact with residues deep within the seven transmembrane domain core of the protein.[40] Classes B and C are smaller, with 48 and 22 members in humans, respectively,[36] and have large N-terminal domains that are essential for ligand binding.[41, 42]

Heterotrimeric G proteins are much less numerous than GPCRs. Typically, the G proteins are classified by the identity of the Gα subunit, of which the 16 different human genes are divided into 4 classes.[43] The Gs class contains Gαs and Gαolf; the Gi/o class contains Gαo, Gαi1, Gαi2, Gαi3, Gαt1, Gαt2, Gαz, and Gαgust; the Gq/11 class contains Gαq, Gα11, Gα14, Gα15/16; and the G12/13 class contains, not surprisingly, Gα12 and Gα13. There is also a fifth class of heterotrimeric G proteins, known as Gv (named for the Roman numeral "V"), which is not observed in humans but found in other animals, including other vertebrates. Gv is just as ancient as the other four classes, traceable back to the earliest metazoans.[44] Classically, the Gs class is characterized by its ability to stimulate cAMP production by adenylyl cyclases (AC) and the Gi class by the opposite, to inhibit AC function. The Gq class activates phospholipase Cβ (PLCβ) enzymes and the G12/13 class is liked to activation of RhoGEF proteins. However, the list of known Gα effectors has grown to include many other signaling pathways, and classification is also complicated by the fact that the Gβγ subunits are able to activate effectors in their own right once they are released from the heterotrimeric complex with Gα.[43] The analysis of the myriad signaling pathways associated with heterotrimeric G proteins is out of the scope of this dissertation, but instead we focus on the very early event common to all heterotrimeric G protein pathways, the activation of the G protein by a GPCR.

There are 5 Gβ and 12 Gγ genes in humans, which are generally expressed widely in overlapping patterns (the exception is Gγ1, which is restricted to the brain and eye).[45] Many, but not all combinations of Gβγ dimers can form, and there have been some indications that specific combinations are preferred for some receptors and/or Gα subunits.[46] However, there are also many cases where the different combinations of Gβ and Gγ subunits do not seem make much of a difference for either receptors or the effectors they interact with.[47] Likely because of this uncertainty in the role of Gβγ subtype specificity, no classification system exists for them as it does for the Gα subunits.

1.5 – Canonical GPCR signaling.

In the canonical GPCR-mediated G protein activation scheme, a GPCR binds an agonist (or in the case of rhodopsin, absorbs a photon to create a covalently bound agonist) on its extracellular side, which enhances the propensity of the receptor to enter an activated conformation. This change is recognized from the intercellular side of the receptor by a heterotrimeric G protein, which binds to the activated receptor. Normally, a molecule of guanine diphosphate (GDP) is tightly bound to the Gα subunit of the G protein, but when the G protein binds to the activated receptor it causes the GDP to dissociate. A guanine triphosphate (GTP) molecule from the cytosol can then bind to the vacant site on the Gα subunit, causing the G protein to dissociate from the receptor and the alpha subunit to dissociate from the Gβγ subunit. The GTP-bound Gα and Gβγ subunits both go on to activate various other signaling proteins in the cell, causing a cell-specific cascade of reactions that leads to an appropriate behavior (or an inappropriate one, in the case of disease) in response to the GPCR agonist.[48] This process may or may not involve full dissociation of the Gα and Gβγ subunits.[49]

The Gα subunit contains an intrinsic GTPase activity, enabling it to convert the bound GTP molecule to a GDP molecule. However, this reaction is typically quite slow for heterotrimeric G proteins, and in a cellular context the rate of the GTPase reaction is usually enhanced by the interaction of the Gα subunit with a member of the regulator of G protein signaling (RGS) proteins.[50] The GDP-bound Gα subunit is then able to bind the Gβγ subunit again, leading to the termination of both Gα and Gβγ subunit-based signaling cascades.[51] This cycle is shown below in figure 1-1.

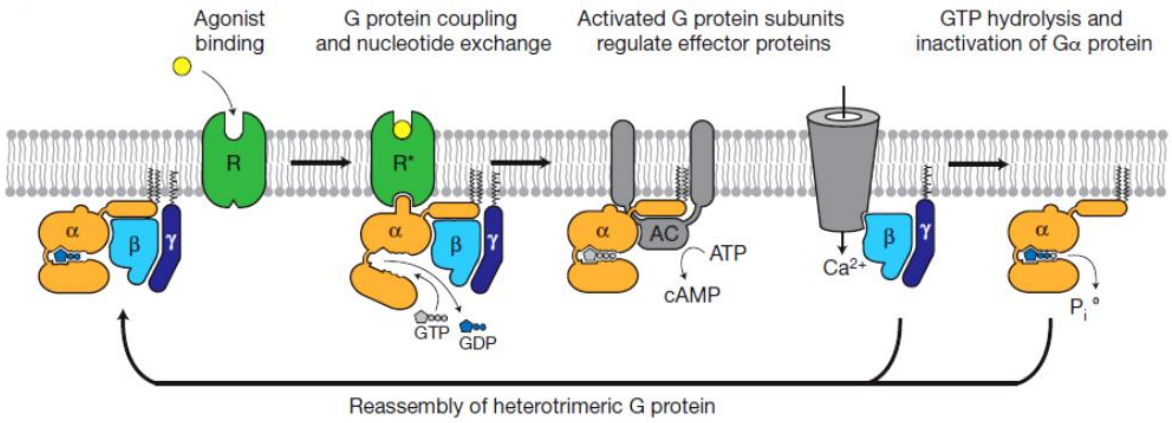


Figure 1-1. The canonical GPCR and G protein signaling cycle. Please see text for details. Figure reprinted from Rasmussen et al.[52]

In addition to the signal termination by RGS-mediated or intrinsic GTP hydrolysis, GPCR signaling is often accompanied by a cellular process that pulls the activated receptors off of the plasma membrane and into endosomes, known as desensitization. In the most well established pathway, which is present for the majority of GPCR systems studied to date, an agonist-bound, activated receptor is recognized by a G protein-coupled receptor kinase (GRK). The kinase phosphorylates several residues on the receptor, typically on intercellular loop 3 or the C-terminus. This phosphorylation promotes the removal of the receptor from the plasma membrane into clathrin-dependent vesicles.[53] The arrestin class of proteins is usually involved in this process, although recent research has also ascribed it a large number of additional roles in GPCR-mediated signaling.[54]

1.6 – GPCR activation and binding theory.

The advent of good radiolabeled antagonists for GPCRs allowed for a very important observation about the mechanism of G protein activation. Competition binding assays with unlabeled agonists demonstrated that a second, higher affinity binding site for agonists to their receptor appeared when their cognate G proteins were present in the samples. If the G protein was dissociated by the addition of the non-hydrolyzable GTP analogue GTPγS, the high-affinity binding was lost, shown in figure 1-2A, below. This was evidence that not only is the G protein activated by the agonist-bound receptor, but it can also induce the receptor itself to adopt an

activated conformation, indicating a true allosteric link between the ligand and G protein binding sites. This observation was the basis for the "ternary complex model" (TCM) between receptor, G protein, and agonist, as formulated by Eliot Ross[31] and Andre De Lean.[55] The TCM and its descendents, shown in figure 1-2B and C, below, are the most common mathematical models for GPCR function in current use.

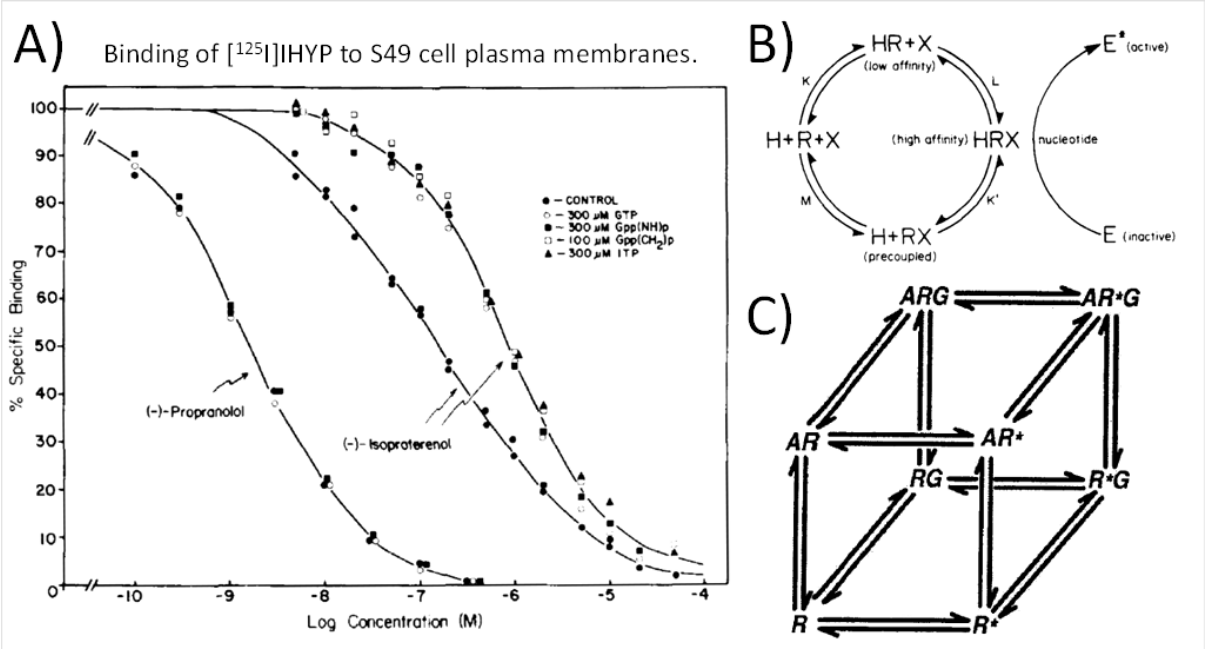


Figure 1-2. Development of the ternary complex model of GPCR function. A) An early example of high-affinity agonist binding. The agonist isoproterenol binds β-adrenoreceptors with two different affinities, leading to a binding curve with a shallower-than-normal slope. Reprinted from Ross et al.[31] This research was originally published in the Journal of Biological Chemistry, © the American Society for Biochemistry and Molecular Biology. B) An early formulation of the ternary complex model reprinted from De Lean et al,[55] where the binding of the G protein (denoted as "X" in the diagram) enhances the affinity of the receptor ("R") for the agonist ("H"). In the description of the model, the effector is activated by interaction with X after its nucleotide-catalyzed release from the high affinity complex, although that is not immediately apparent in the provided diagram. This research was originally published in the Journal of Biological Chemistry, © the American Society for Biochemistry and Molecular Biology. C) A modern schematic of the cubic ternary complex model (cTCM) which incorporates a two-state model of receptor activation into the basic conceptual framework of the TCM. Figure redrawn based on Christopoulos and Kenakin.[56]

The TCM as originally described did not explicitly apply a two-state model to the receptor, although agonist and G protein binding were proposed to cause conformational changes to the receptor, which was represented implicitly in the mechanism. It was not until 1993 that Philippe Samama and colleagues applied a

full 2-state model to GPCRs, called the extended ternary complex model (eTCM),[57] even though similar models were applied 26 years earlier to the nicotinic receptors.[21] The application of the model was made in the context of research from Tommaso Costa and Albert Herz that showed high levels of basal activity from overexpression of receptors, and these systems facilitated the discovery of inverse agonists as a way that drugs could interact with GPCRs .[58] A few years later, the eTCM model was upgraded again to the cubic ternary complex model (cTCM) by Terry Kenakin, drawn above in figure 1-2C, which allowed the G protein to interact with a receptor in the inactive conformation as well.[59-61] The cTCM represents a thermodynamically complete two-state model that interacts with one ligand and one G protein. As discussed at length in section 5.8, it is unclear how prevalent the states with interacting inactive receptor and G protein are in reality. The cTCM also forms the basis for more models that describe more complex situations, such as binding of allosteric GPCR ligands or ligand biased signalling.[56]

1.7 – GPCR and G protein structure.

All GPCRs share a core fold that consists of 7 transmembrane α-helices (numbered TM1 to TM7), connected in a serpentine fashion with the N-terminus of the protein on the extracellular side of the membrane and the C-terminus on the intercellular side. A short, amphiphatic helix known as helix 8 (H8) is also found directly after TM7, where it lies parallel with the membrane. On most receptors, palmitoylation site(s) and/or a stretch of cationic residues immediately follow H8 and help keep the helix on the membrane. The residues of the transmembrane helices are primarily hydrophobic in nature, although a number of polar or charged residues are typically found in the center of the transmembrane bundle.[62, 63] These residues form a network of hydrogen bonds with each other and with several water molecules that help hold the hydrophobic transmembrane helices together in the hydrophobic membrane layer, but is also flexible enough to allow the helices to move as required during the conformational changes involved in GPCR function.[63, 64] The only class of GPCRs that we have high-resolution structural data in the transmembrane region is for class A receptors, but it is likely that class B and C have similar overall structures, since they interact with and activate the exact same G proteins as the

class A receptors. Shown below in figure 1-3 are some of the key features of the GPCR fold.

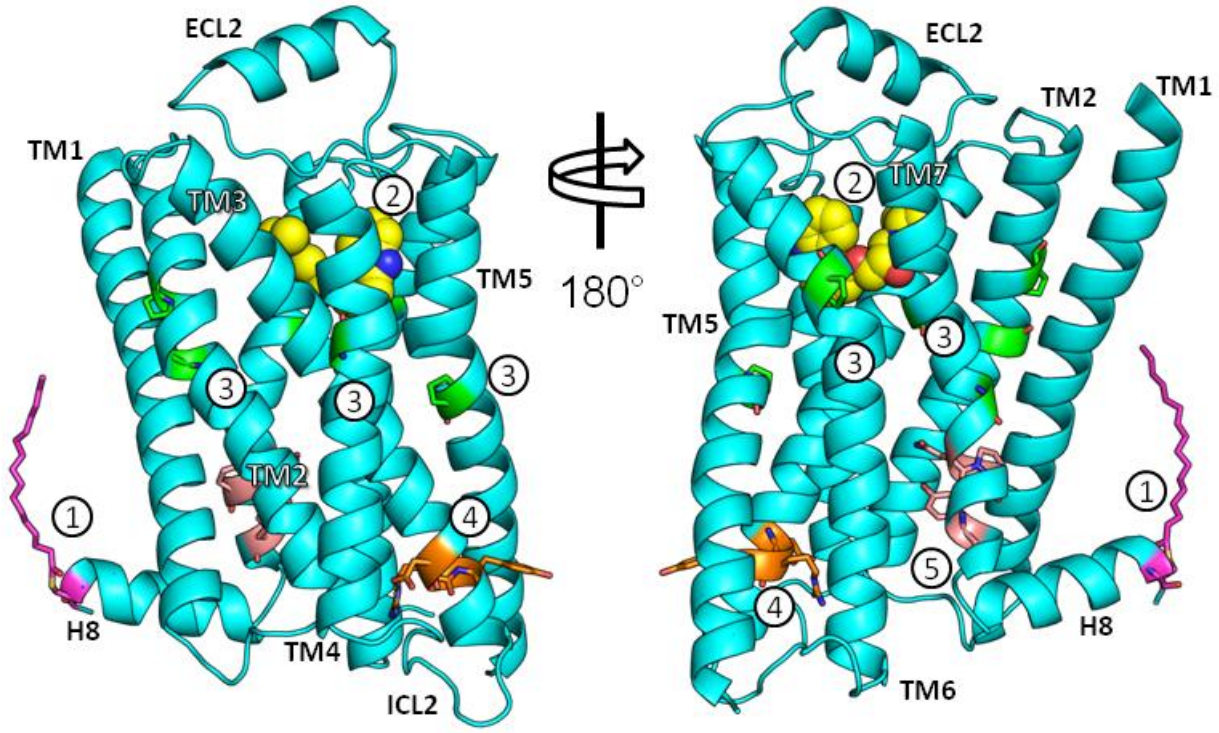


Figure 1-3. Structural overview of GPCRs. The inactive β_2-adrenergic receptor (PDB 2RH1) is used as an example. Numbered features are as follows: 1) Palmitoylation at the end of helix 8, colored in magenta. 2) Ligand binding site, ligand colored in yellow. 3) Transmembrane helix kicks induced by proline or glycine residues, colored in green. 4) (E/D)RY motif, colored in orange. 5) NPxxY motif, colored in salmon. Please see text for additional details.

TM3 is the longest helix in the transmembrane region, and it is tipped by ~35° from the normal vector of the membrane. It makes well-conserved contacts with TM2-6, forming the structural core that the other helices pack and move against. In the extracellular loop 1 region (ECL1) between TM2 and TM3, a highly conserved disulfide bond is formed, which also serves to stabilize the GPCR fold. The ECL2 region is highly variable in sequence and structure for GCPRs in general, but it tends to be conserved for each receptor family. Similarly, the ligand binding site is conserved within the receptor families but not highly in general, and is usually formed between TM3, TM5, TM6 and TM7.[40]

Towards the center of the receptor, most of the TM helixes have a single kink induced by proline and/or glycine residues. These kinks serve as fulcrums that amplify small movements in the extracellular side of the receptor to larger ones on the intracellular side, discussed more in chapter 3. On the intercellular side of the receptor, there are two prominent structural and sequence features. One feature, the (E)DRY motif, located on the intracellular side of TM3, is part of the so-called "ionic lock" interaction. This interaction was seen in the first GPCR structure solved, bovine rhodopsin, which showed a salt bridge between the TM3 DRY sequence and a glutamic acid on TM6 that helped to prevent the receptor from reaching its activated conformations.[65] However, it is not obvious how important similar interactions are for other GPCRs, as such salt bridges are not observed in most other receptor structures and the effects of mutations in the motif are complex.[66] The other feature has a consensus sequence motif of NPxxY, and is found at the end of TM7, right before the bend that allows Helix 8 to lie along the membrane. This sequence is important for stabilizing the active receptor conformations, discussed further in chapter 3.

Heterotrimeric G proteins consist of 3 different polypeptide chains, the Gα, Gβ, and Gγ subunits. The Gβ and Gγ subunits bind very tightly together, a process that is typically aided by specific chaperone molecules.[67, 68] The Gβ and Gγ polypeptide chains do not dissociate from one another, so the dimer is often treated as a single subunit, the Gβγ subunit. The C-terminus of Gγ is either farnesylated or geranylgeranylated, and serves to anchor the subunit in the membrane.[47] Shown below in figure 1-4A, the Gβ subunit has a 7-bladed WD-repeat fold, also called a β-propeller fold, and an N-terminal alpha helix. The Gγ is mostly alpha helical in structure, and it packs along the Gβ α-helix and the 5th, 6th, and 7th WD repeat. The Gα subunit binds to the Gβγ subunit through contacts between the 1st, 2nd, and 3rd WD repeat of Gβγ and the switch II and N-terminal region of Gα,[69] shown below in figure 1-4B. The interaction with the switch II region is always broken when the Gα is bound to GTP,[51] but for the N-terminal region there is evidence that suggests some G proteins can maintain the interaction even after activation with GTP, and other G

proteins tend to lose all interactions and physically dissociate into separate Gα and Gβγ subunits.[70, 71]

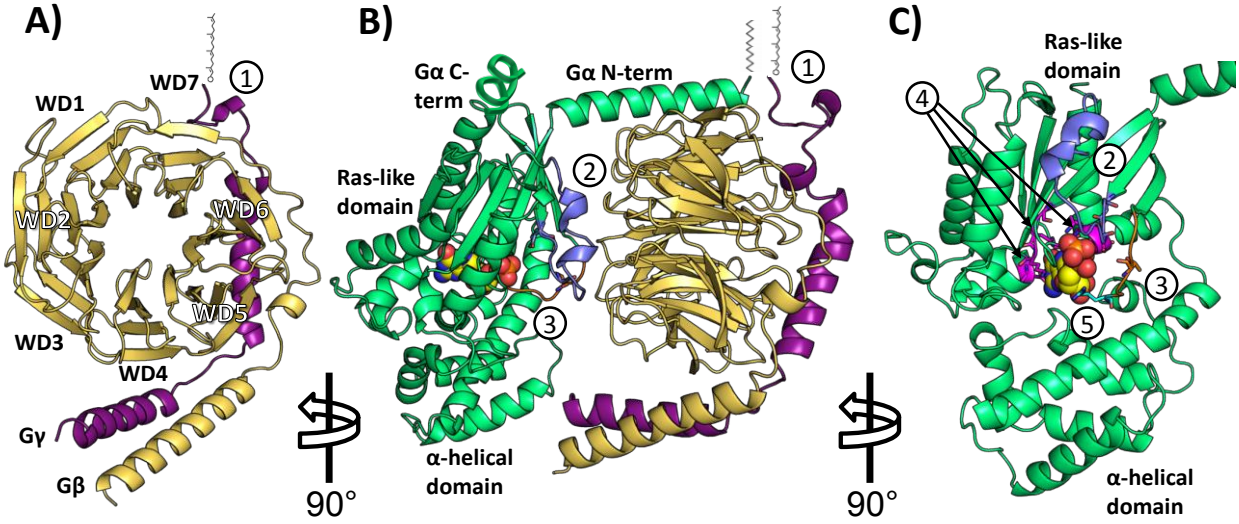


Figure 1-4. Structural overview of G proteins. The GDP-bound Gt/Gi chimera heterotrimer structure (PDB 1GOT) is used as an example. Numbered features are as follows: 1) Farnesylation or geranylgeranylation of the Gγ C-terminus and palmitoylation and/or myristoylation of the Gα N-terminus. 2) The Gα switch II domain (shown in periwinkle blue) includes G-box sequence 2 and binds to Gβ when the protein is GDP bound. 3) The Gα switch I domain (shown in orange) includes G-box sequence 3. 4) G-box sequences 1, 4, and 5 (shown in magenta) form the rest of the guanine nucleotide binding site. 5) The catalytic arginine residue (shown in cyan) is at the c-terminal end of the α-helical domain and is followed by switch 1. Please see text for additional details.

The heterotrimeric Gα subunits are part of a class of GTPase enzymes that include the EF-Tu and EF-G families of elongation factors and the $p21^{ras}$ (Ras) small GTPase homologs. All of these proteins contain a core fold of a twisted, 6-strand β-sheet domain surrounded by α-helices and a highly conserved 5-loop structure that forms the guanine nucleotide binding site, shown above in figure 1-4C. The consensus sequences of the loops are called the G-box domains, and are given the designations G-1 through G-5. Two regions that include the G-box domain G-2 and G-3 change conformation due to the presence of the γ-phosphate of GTP and are known as switch I and switch II, respectively. The functions of these G-box domains are discussed in more depth in sections 3. and 5.4 where their participation in G protein activation is addressed.

Heterotrimeric Gα subunits have an additional domain composed of a 6 α-helix bundle which is grafted into a loop of the core GTPase fold. This forms a protein with two distinct sections; the Ras-like GTPase domain and the Gα specific α-helical domain.[51] The α-helical domain has a key residue, arginine 201 on Gαs, that is necessary for hydrolysis of GTP to GDP. The homologous arginine for the Ras family is not found on the Ras proteins themselves, but instead is supplied by a separate protein required for hydrolysis, known as Ras-GAP.

1.8 – Motivation for research.

At the beginning of my research, there was high resolution structural information for only two GPCRs, the human β_2-adrenergic receptor and bovine rhodopsin. These receptors were in an inactive conformation, but it was known from biochemical and biophysical work that they must change conformation quite significantly during activation. In particular, the transmembrane helix 6 of the receptor was known to move outward, which would open up a binding site on the receptor that almost certainly seemed to be the site of interaction with the G protein α-subunit c-terminus. However, several major questions about the activated receptor and its interaction with G protein remained, which made the nucleotide-exchange step the least well-understood part of the G protein signaling cycle. The orientation between the interacting receptor and the G protein was unknown, and it was unclear how the activated receptor actually caused the changes to the G protein that led to nucleotide exchange. The leading models suggested that the receptor either pulled the Gβγ subunit away from the Gα subunit, the “lever-arm” model, or tilted it in towards the Gα subunit, the “gear-shift” model. Both models suggested that the accompanying distortion of the Gα switch I and II regions was also responsible for forcing out the nucleotide. There were also several models of activation that required a dimer of receptors or had a single receptor bind to both Gα and Gγ subunits simultaneously. In addition, it was unclear if the simple act of binding to a receptor was enough to cause nucleotide exchange on a G protein, or if the exchange step induced by receptor activation was transmitted separately from the G protein binding step.

I started the project with the end goal of being able to obtain a crystal structure of the β_2AR and Gs interacting together. Brian Kobilka and his postdoc, Søren G. F. Rasmussen, recently crystallized the β_2AR and knew how to make it in quantity. Roger Sunahara's lab had also started making significant quantities of the Gs heterotrimer, so I was confident that if we could figure out how to make a β_2AR-Gs complex sample, we could actually produce enough sample to complete a crystallography project. We hoped that seeing the interaction between the two proteins in high structural resolution would either solve or greatly aid in answering the questions mentioned in the previous paragraph about the nucleotide exchange step of the G protein signaling cycle.

CHAPTER 2
G PROTEIN PRODUCTION AND RECEPTOR-G PROTEIN COMPLEX OPTIMIZATION

2.1 – Introduction.

This section contains work that was done in order to determine how receptor and G protein could be made to interact to form a molecular complex, figure out how to produce the large quantities of G protein needed, and how to stabilize the receptor-G protein complex for crystallography. In most cases, the data in this section are the first indications of findings that were more conclusively proved in the work presented in the following chapter. This work represents the majority of my time invested throughout the project. The first experiments that devised the overall nucleotide-depletion strategy for making receptor-G protein complex and proving it could be purified (given in sections 2.3 to 2.9) were performed by me using my own heterotrimer and receptor produced and purified by Søren G. F. Rasmussen from Dr. Brian Kobilka's lab. However, once we started crystallography trials, Søren performed the final sample preparation so it could go directly into Brian Kobilka's crystallization trays, and I focused on making enough G protein to support the crystallography effort as well as worked together with Søren to assess the sample quality and find ways it could be improved. Later on in the project, many other collaborators joined to contribute techniques and reagents that proved to be critical to success, but this was only after Søren and I had worked for a couple years to optimize the basic receptor-G protein coupling protocol and were making high-quality samples that still refused to crystallize into useful crystals.

2.2 – Overall Gs heterotrimer purification strategy.

The overall strategy used for Gs heterotrimer expression and purification was pioneered by Gilman lab 20-30 years ago.[72, 73] The three G protein subunits are co-expressed in insect cells by infection with recombinant *Autographa californica* multiple nuclear polyhedrosis viruses, with a single N-terminal hexahistidine tag on the Gβ subunit. Membranes from the infected cells are isolated and then the intact heterotrimer is solubilized with detergent. This is followed by chromatographic purification using Ni-NTA resin, strong anion exchange, and gel filtration. The bulk of the purification is done on the Ni-NTA column and the following anion exchange step is effective at separation of the heterotrimer from excess Gβγ subunits. The gel filtration step serves mainly to reduce the concentration of detergent and exchange the sample into the final buffer.

In the original procedure that our protocol was based on,[73] the protein was initially solubilized in sodium cholate, exchanged into $C_{12}E_{10}$ (closely related to another historically used detergent Lubrol, or $C_{12}E_9$), and finally into CHAPS. For our samples, however, we knew that the final detergent in our protocol had to be dodecylmaltoside (DDM) because of the receptor's detergent requirements. Therefore, we re-worked the sequence of detergents so that the protein was exchanged from sodium cholate into anzergent 3-12 and then into DDM. This allowed us to add together purified receptors and G proteins with no fear of detergent mixing problems. Because directly exchanging the solubilized membranes into DDM was prohibitively expensive, anzergent 3-12 was used as an intermediate detergent. The critical micelle concentration of anzergent 3-12 (0.094 %) is in-between between that of sodium cholate (0.41 %) and DDM (0.0087 %),[74] which facilitates quick and complete detergent exchange.

2.3 – Optimization of transfection and expression conditions.

The incubation time for protein expression was determined by partially purifying G protein from membranes harvested from Hi5 insect cells between 27 and 76 hours post transfection, shown below in figure 2-1. The relative amounts of Gα and Gβ subunits were quantified by Coomassie blue staining of SDS-PAGE gels[75] of

samples obtained by following the first part of the Gs purification method given in method 6.1 up through elution from the Ni-NTA column. The partial purification of samples was necessary because gels and western blots of crude membrane fractions showed very large amounts of G protein subunits being produced, but only a minor fraction could be solubilized with detergent, indicating that most of the protein was in an insoluble, aggregated form. From inspection of the graph, harvesting the sample during the window of time between 40 and 48 hours post-infection was determined to be optimal for routine protein production.

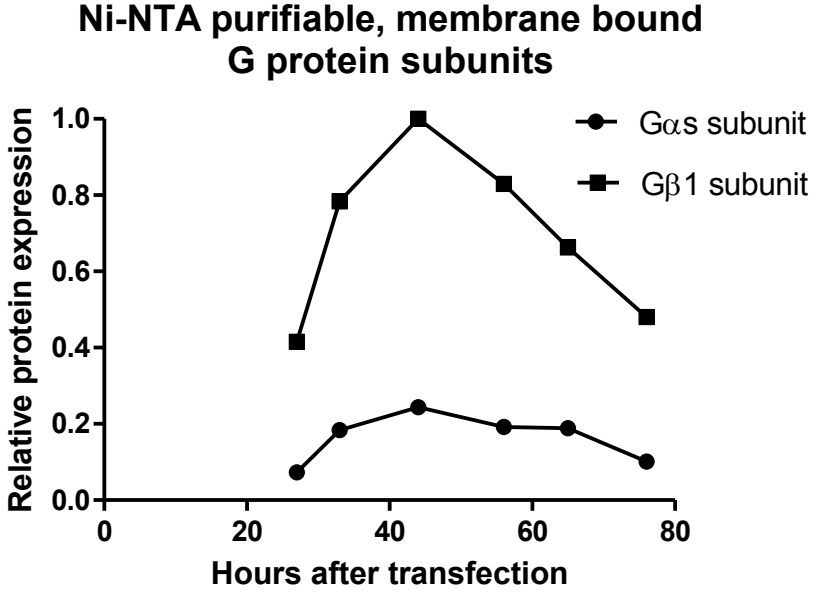


Figure 2-1. Expression of G protein subunits in Hi5 insect cell membranes. Samples were purified by Ni-NTA chromatography, run on a SDS-PAGE gel, and stained with Coomassie blue. The window between 40 and 48 hours was chosen for harvesting membranes for optimal protein production.

2.4 – Optimization of G protein solubilization conditions.

The temperature and detergent used to solubilize the Gs heterotrimer was also optimized. For this assay, a single preparation of membranes with 5 mg/ml membrane protein was incubated for 1 hour with 1% concentration of various detergents in Gs purification wash buffer at the specified temperature. The samples were centrifuged at high speed and analyzed by SDS-PAGE. The ratio of insoluble (pelleted) verses soluble (supernatant) G protein subunits were determined by

densitometry of the Coomassie-stained gel bands, shown below in figure 2-2. The detergents used were chosen because of their historical use in G protein or GPCR purifications or because of their low cost. It can be seen that in all cases, sodium cholate was the best performing detergent, sometimes by a wide margin. Although extremely inexpensive, the polyoxyethylene-based detergents and Triton X-100 both did not perform well in general. Anzergent 3-12, a zwitterionic detergent, and DDM, the required detergent for purification of the receptor, also were not able to outperform cholate for the initial membrane solubilization steps. It can also be seen from the graph that the best temperature for protein extraction was 0°C, so in the end the chosen conditions for Gs solubilization were the same as those determined by Kozasa & Gilman[73] for G_{12}, G_q, G_z, and G_{i1}.

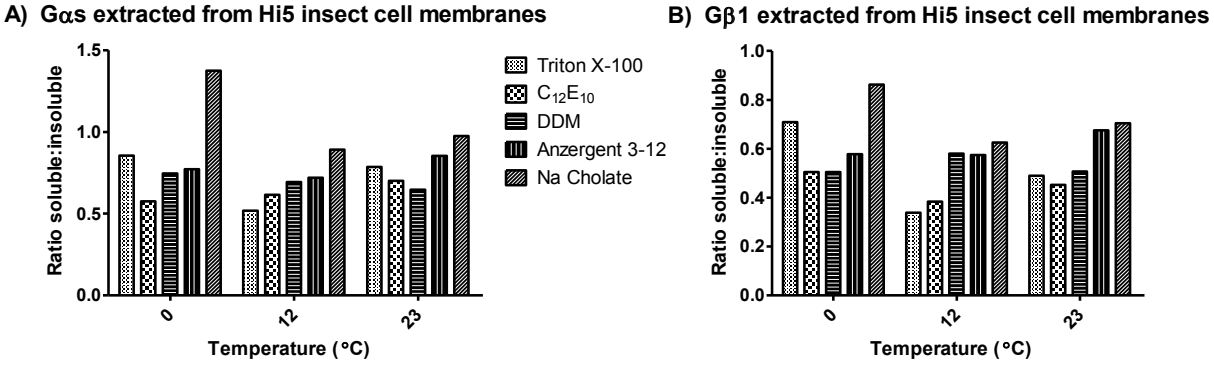


Figure 2-2. Solubilization of Gs subunits from Hi5 insect cell membranes. A) Ratio of soluble to insoluble Gαs subunit. B) Ratio of soluble to insoluble Gβ1 subunit. For both proteins, the ratio was determined with densitometry of Coomassie-stained SDS-PAGE gels. The chosen solubilization condition was incubation of membranes with sodium cholate at 0°C for 1 hour.

2.5 – Gs dephosphorylation.

Because of data shown in section 2.10, we determined that we needed to insure that both our receptor and G protein were not phosphorylated in order to interact efficiently. While we routinely added protein phosphatases to the reaction mixture during incubations for forming receptor-G protein complex, we also found it convenient and beneficial to dephosphorylate the G protein sample during purification. Evidence for this is found in figure 2-3 below, where several anion exchange chromatrographs for Gs purifications are shown.

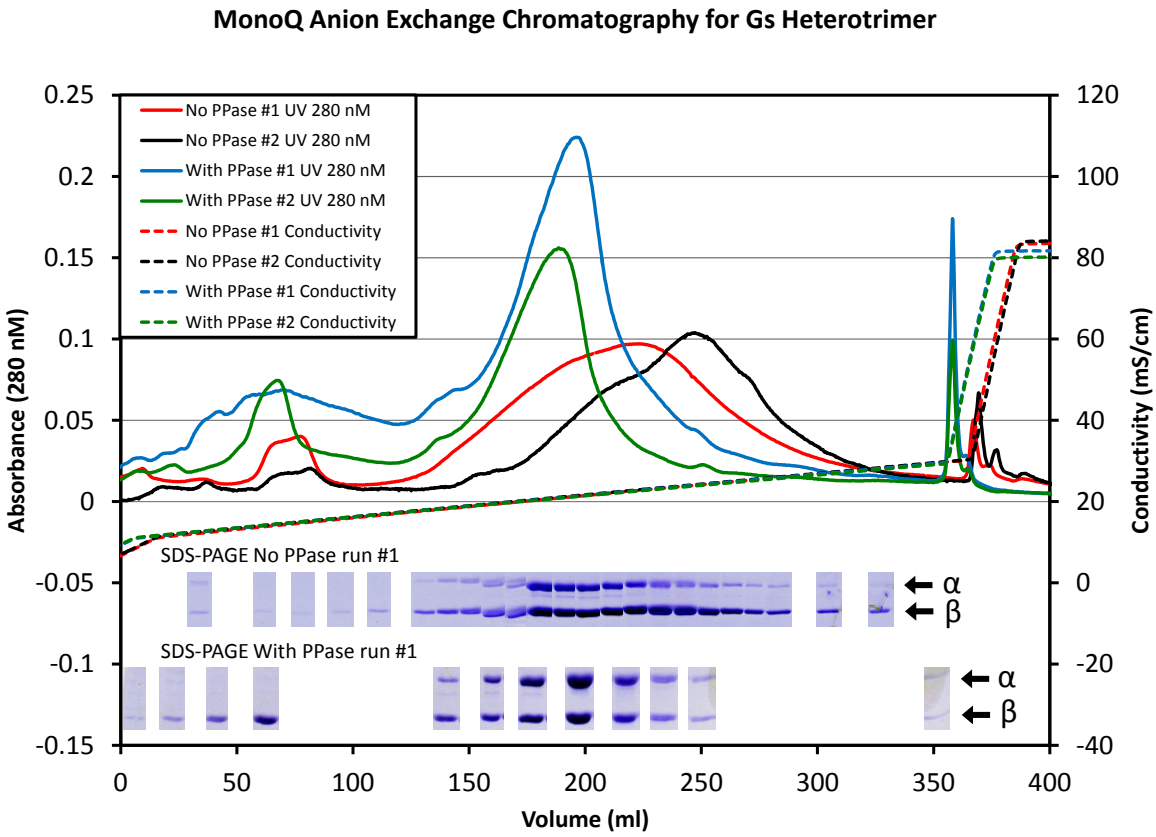


Figure 2-3. MonoQ anion exchange chromatography of Gs heterotrimer. Shown are 280 nm absorbance and conductivity traces from two purification runs without any protein phoshatases and from two runs that had 5 units per liter original culture volume of lambda protein phosphatase (λPPase) added before the column. For the first run of both the λPPase and non- λPPase treated samples, fractions were analyzed by SDS-PAGE gels and stained with Coomassie blue, shown below the traces in their corresponding position.

Although the chromatographs from individual purification runs can vary quite significantly from one another, the general elution patterns shown in the traces above were typical. Without any additional protein phosphatases added to the samples, the Gs eluted in a broad peak from the column around 23 millisiemens per centimeter (mS/cm) with sodium chloride as the eluent salt. When similar samples were treated for 1 hour with 5 units per liter of cell culture volume of lambda phage protein phosphatase (λPPase, made and assayed in-house by methods given in sections 6.3 and 6.4) the Gs eluted reliably at 20.5 mS/cm. The λPPase treated protein came off the column in a tighter peak, significantly improving the chromatography and reducing eluate volume. In addition, λPPase treatment reduced the amount of free βγ subunits that co-elute with the heterotrimer, which can be seen

by comparison of the intensity of the SDS-PAGE bands for the Gα and Gβ subunits. Although we never systematically determined why the λPPase treatment made such an improvement in the anion exchange chromatography, we assume it is because the removal of the phosphate groups from the proteins in the sample reduces the overall negative charge of the proteins as well as eliminates the charge heterogeneity that arises from non-uniform phosphorylation. Both these effects would lead to the observed elution profiles which show the protein tends to interact with the resin a little weaker than without λPPase treatment and that it elutes in a sharper and more consistent peak.

2.6 – Receptor and G protein interact in rHDL particles.

In order to test if the purified G protein could interact with a receptor and allow agonist-induced nucleotide loss of the G protein, we used a conformationally-sensitive fluorescent receptor that had been developed in the Kobilka lab. Previously, the Kobilka lab together with our own lab showed that versions of the β_2AR can be made which have 5 of the most reactive cysteines mutated into non-reactive residues, allowing facile labeling on any other accessible cysteine residue with mono-bromobimane, a very small, environmentally sensitive fluorescent dye.[76, 77] It was found that the large conformational changes that the receptor experiences upon activation could be detected using a bimane-labeled β_2AR with a unique reactive cysteine at residue 265, which is located towards the cytoplasmic end of transmembrane helix 6 (TM6). Fortunately, the bimane label also did not interfere with G protein binding when attached to that residue, allowing us to confirm receptor and G protein interaction, shown below in figure 2-4.

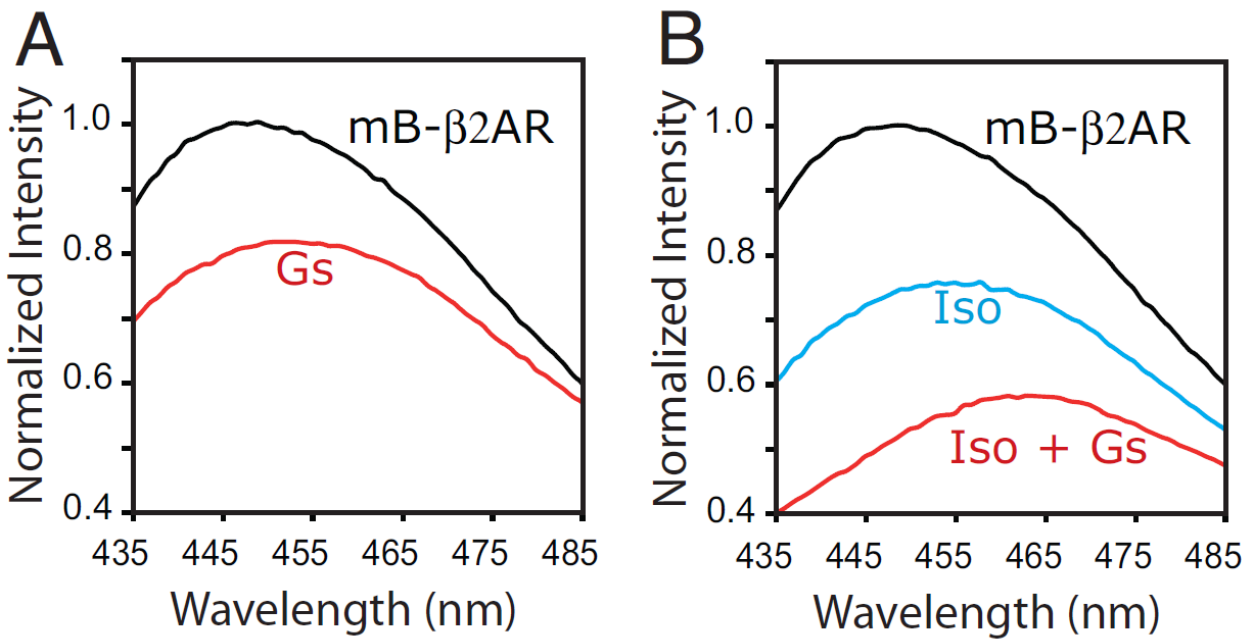


Figure 2-4. Conformational change of the receptor in the presence of agonist and G protein in rHDL discs. In figure A, bimane-labeled β_2AR in rHDL particles is mixed with excess Gs heterotrimer without any drugs. In figure B, the labeled receptor is incubated with saturating amounts (100 μM) of the agonist isoproterenol (Iso) or with Iso and Gs together. Full experimental details given in method 6.5. Figure reprinted from Yao et al, 2009.[77]

In the bimane-receptor assay shown above (described in methods section 6.5), the conformational changes of the fluorophore-labled TM6 domain of the receptor forces the flourophore into a more polar environment, detected as a generalized decrease in fluorescence intensity as well as a redshift in the λ_{max} of the fluorescence spectrum. Both Gs (figure A) and the agonist isoproterenol (figure B) cause the environment of the flourophore to change to a similar extent when added to the receptor alone. However, when both the agonist and the G protein are added together, the environment of the flourophore changes to an even greater extent than either the drug or G protein can do so by itself (figure B). We interpret this phenomenon as evidence for cooperativity of receptor activation between the agonist and G protein. Such behavior is expected, as agonists by definition cause the GPCR to interact with and activate G protein. This experiment also confirmed that both the receptor prepared by the Kobilka lab and the Gs heterotrimer prepared by our lab were completely functional.

2.7 – Apyrase helps stabilize the receptor and G protein interaction.

We also hypothesized that since GDP release is the consequence of G protein binding to the receptor, known from the cononical G protein cycle, we might be able to stabilize the interaction of the two proteins by removing GDP from the sample as it is relased upon receptor-G protein complex formation. Removal of the GDP would drive the equlibrium towards complex formation by preventing GDP re-association to already formed complexes and causing them to dissociate in the reverse reaction. We tested this idea using the same bimane-labeled receptor assay as in section 2.6 above. However, this time we used samples of the receptor with Gs heterotrimer in rHDL discs and then added the enzyme apyrase, a non-specific nucleotide hydrolase, to convert the free GDP to GMP and P_i, shown below in figure 2-5.

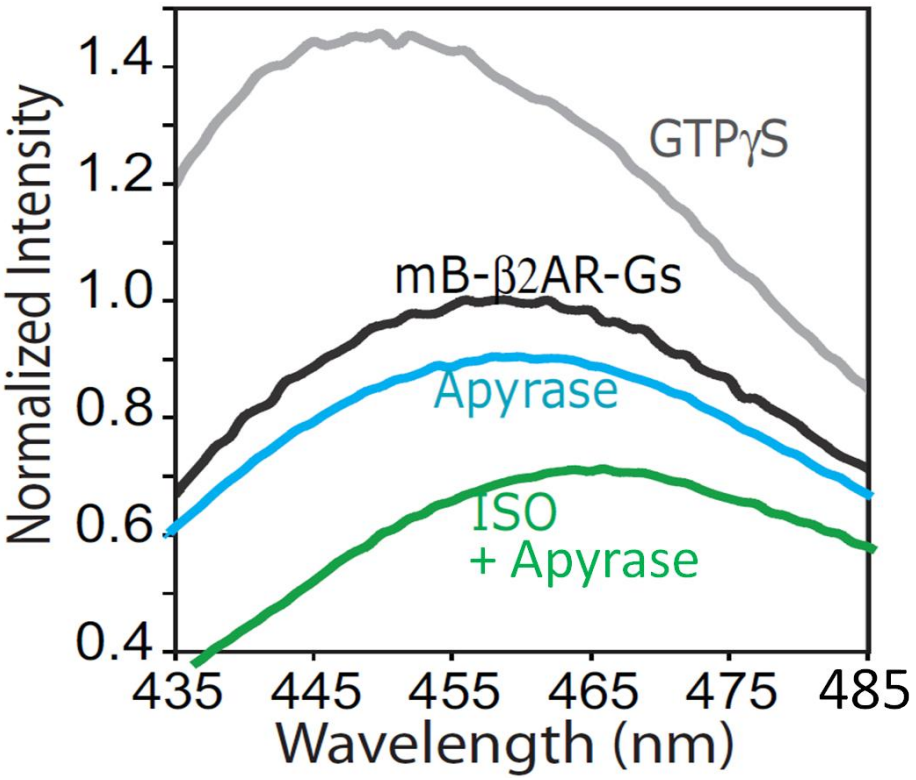


Figure 2-5. Conformational change of the receptor and G protein in rHDL discs when treated with apyrase to remove GDP. Shown in black is the fluorescence emission spectrum of samples with only receptor and G protein. When the sample was incubated with 10 μM GTPγS for 40 minutes in order to uncouple all the G protein from the receptor, the spectrum changed to the one shown in grey. When it was incubated with apyrase or apyrase plus 100 μM of the agonist isoproterenol (ISO), the spectra changed to the ones shown in blue and green, respectively. Figure reprinted from Yao et al, 2009.[77]

When apyrase was added to the sample with receptor and G protein, it caused a small but noticeable decrease in fluorescence intensity and a slight redshift. This effect was also present when apyrase was added to samples with the agonist isoproterenol, causing the intensity of the fluorescence signal to decrease to 48% of the non-activated receptor signal as opposed to 55% for samples that were only incubated with the agonist and G protein but no apyrase (figure 2-4). More evidence in support of using apyrase to help drive the receptor-G protein coupling reaction to completion was found using radioligand binding of an antagonist to the receptor, but the interpretation of the data is dependent on the overall model of receptor and G protein allostery and so it is discussed in chapter 4, which deals at length with such experiments.

2.8 – Receptor and G protein can interact in detergent.

Initially, we were not certain if it would be possible for the Gs heterotrimer and the β_2-adrenergic receptor to interact productively while solubilized in detergent instead of being in an rHDL particle or other bilayer membrane environment. In order to measure receptor-G protein interaction in detergent, the bimane-receptor assay used in section 2.6 above was adapted to use with highly concentrated, detergent solubilized proteins in a 384-well plate format. Details of the assay are given in method 6.6, and a timecourse of representative data are shown below in figure 2-6.

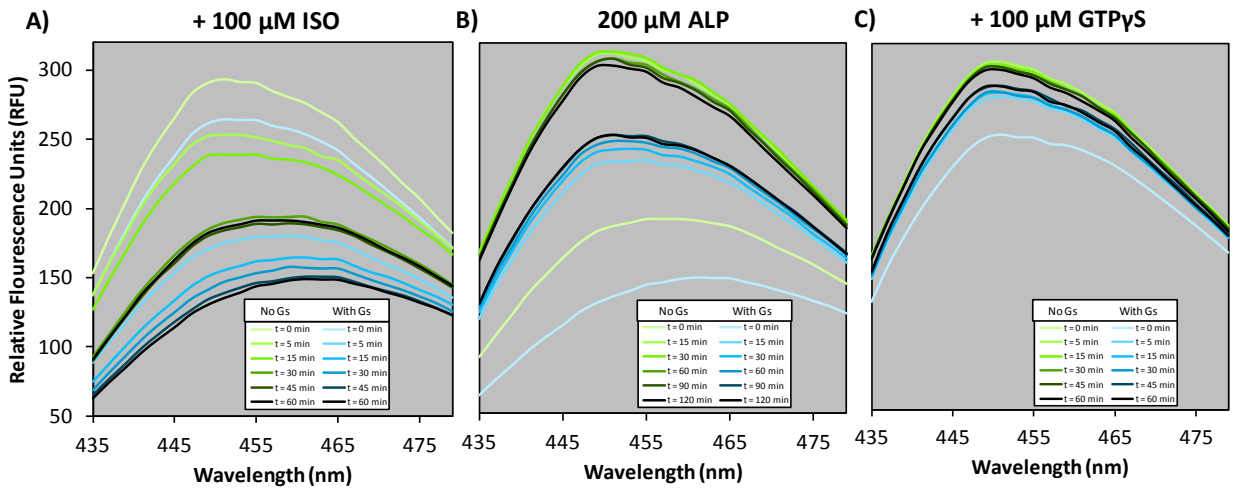


Figure 2-6. Conformational change of the receptor in the presence of agonist and G protein in DDM detergent. A) 100 µM isoproterenol (ISO, an agonist) was added to samples of 1 µM bimane-labeled β_2AR either with or without 5 µM Gs. The fluorescence spectra were taken at the specified timepoints after the addition of drug. B) 200 µM alprenolol (ALP, an antagonist, used at a concentration that should compete 400× more effectively than the ISO concentration used for binding the receptor) was added to the same samples. C) GTPγS was added to the same samples to force G protein to uncouple from the receptor.

We found that we could get the two proteins to interact in DDM detergent when they were at micromolar concentrations. In part A of figure 2-6, one can see that upon addition of the agonist isoproterenol to the samples, we saw a decrease in fluorescence intensity and redshift of λ_{max} that is also seen in rHDL. In the samples where Gs is included along with ISO, the effect is more extreme than when it is not included, indicative of interaction between receptor and G protein just as in figure 2-4. The isoproterenol-induced changes in fluorescence can be fully reversed in the sample without G protein by addition of alprenolol, a competing antagonist, shown in part B. Only a fraction of the signal can be reversed with antagonist when G protein was included in the initial incubation with agonist, and even extended incubation up to 2 hours does not overcome the effect. However, addition of GTPγS, which is known to bind to activated G protein and cause it to dissociate from the receptor and into separate subunits, rapidly reversed the G protein effect. All together, the response of the receptor and G protein containing sample indicated that the two proteins were able to form a stable complex in detergent. The decrease in decrease in fluorescence intensity and redshift of λ_{max} suggests that the receptor was in an active conformation and the quick response of the sample to guanine nucleotide

suggests that G protein was in a conformation that can rapidly bind guanine nucleotides.

2.9 – Receptor and G protein complex is stable and purifiable in detergent.

After confirming our basic strategy of forming the β_2AR and Gs complex in detergent using the bimane-labeled receptor assay shown above in section 2.8, we needed to determine if the complex was stable enough to purify away from any remaining contaminating proteins and coupling reaction byproducts. To do so, we made a small sample of complex and purified it with anion exchange chromatography. The preparation of the sample was almost exactly as described in method described in section 6.9, except that the sample was incubated for 30 minutes at 30°. Receptor alone and Gs heterotrimer alone were also run on the anion exchange column under identical conditions to confirm the ability of the column to separate the complex from the excess proteins that did not form a complex. Chromatographs of the three samples are shown in figure 2-7.

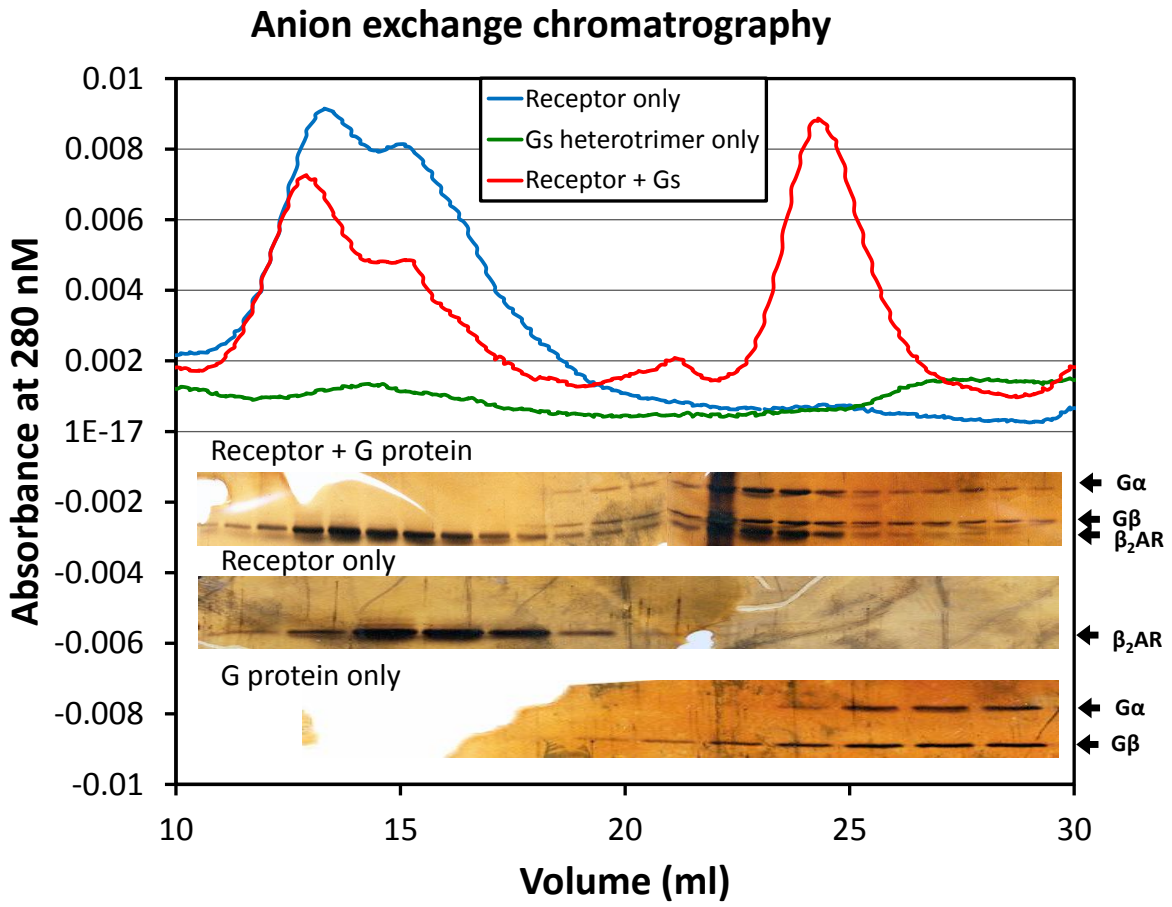


Figure 2-7. Anion exchange chromatography of β_2AR, Gs heterotrimer, and β_2AR-Gs complex. Shown are absorbance traces at 280 nM for each of the indicated samples along with silver stained SDS-PAGE analysis of representative fractions along each chromatogram, lined up to correspond to the approximate position of the fractions in the chromatogram. In these gels, the topmost band corresponds to Gαs, the middle band to Gβ1, and the lowest and slightly fuzzy band to the β_2AR. The extremely dark band on the receptor + Gs gel is due to contaminating protein and is not a feature in the chromatogram.

Formation of a stable receptor and G protein complex was confirmed by the co-migration of receptor with both the Gα and Gβ subunits of the G protein on the anion exchange column. The complexed receptor and G protein eluted at a fraction that neither the receptor nor the Gα subunit was found at when the proteins were run separately. The height of the peak and intensity of the bands also told us that the desired sample of the receptor and G protein complex was forming at efficiencies in the neighborhood of 40%, a perfectly workable starting point for further optimization. The experiment also confirmed that the complex could survive purification by anion

exchange and can be separated from leftover unreacted receptor and G proteins. At this stage, we started preparing milligram quantities of this complex and attempted crystallization trials with the sample.

2.10 – Protein phosphorylation inhibits complex formation.

Relatively soon after we started to prepare large amounts of β_2AR-Gs complex, we tested the sample for phosphorylated residues. The idea we had was that because the receptor is known to be phosporylated as part of normal receptor physiology leading to the downregulation of signaling,[78] and that there were some other reports of Gs proteins being phosphorylated,[79] perhaps we could increase the efficiency and stability of complex formation by removing any phosphate modifications that were added by the insect cell kinases during protein overexpression. We used the commercially available phosphoprotein stain Pro-Q Diamond (Molecular Probes, Invitrogen) on samples analyzed with SDS-PAGE, shown below in figure 2-8.

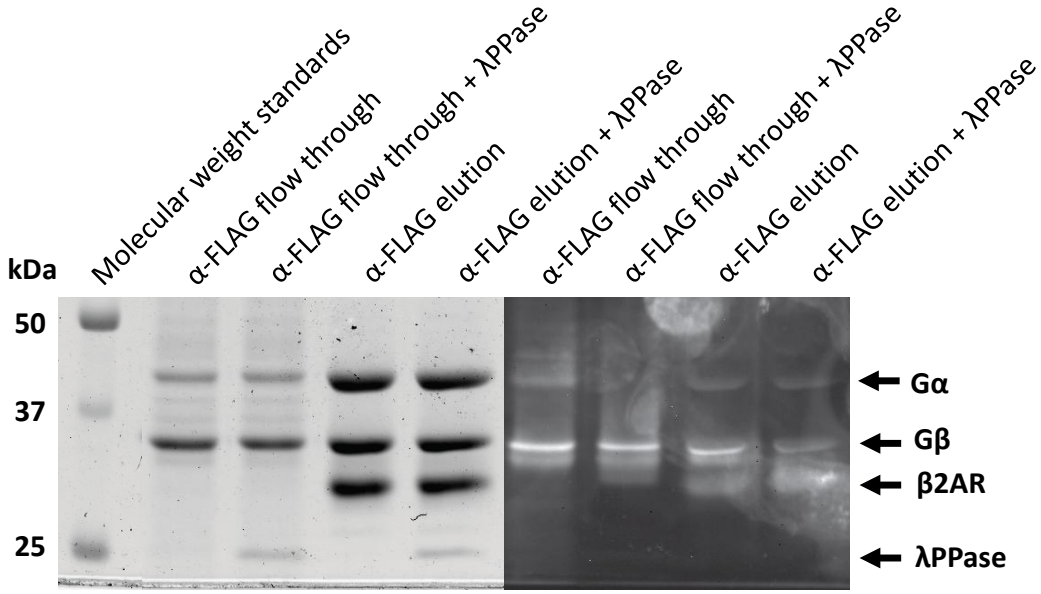


Figure 2-8. Pro-Q Diamond phosphoprotein staining of β_2AR-Gs complex samples. On the left, Coomassie blue stain was used to test for total protein on an SDS-PAGE gel. On the right, Pro-Q Diamond phosphoprotein imagining was used to detect phosphorylation of the same gel. Please see text for the description of the experiment. Data was obtained by Søren Rasmussen with samples prepared by Søren Rasmussen and Brian DeVree.

In the experiment, a sample containing both β_2AR-Gs complex and uncomplexed reactants was purified using a M1-FLAG antibody affinity column for the N-terminal FLAG tag on the receptor. The receptor was previously dephosphorylated during

purification, but the G protein still had any phosphate modifications that were made to it by the insect cells that were used to overexpress it. Both the eluted sample and the flow-through during the loading of the column were analyzed by SDS-PAGE, with some of each sample pre-treated with a large excess of λ-phage protein phosphatase (λPPase), an enzyme that removes all phosphates from tyrosine, threonine, serine, and histidine residues. The gels were stained with ProQ dye according to the manufacturer's directions and then with Coomassie blue stain. The dye will often stain each protein at a certain background level, so it is important to confirm the presence of a phosphate group by seeing that the dye-stained band decreases intensity upon extensive λPPase treatment.

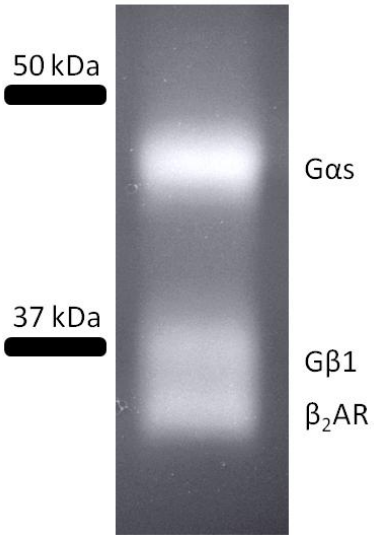
In figure 2-8, one can see that the some of the Gαs protein was not successfully complexed with the receptor and instead remained in the flow-through. This non-reactive Gα subunit stained well with the ProQ dye, but the band completely disappeared after λPPase treatment, indicating that the subunit was probably phosphorylated. In contrast, the Gα subunit that co-purified with the receptor as a complex stained relatively poorly with the phosphoprotein dye (compare dye intensity with coomassie blue stain intensity), and the intensity of the band did not change after λPPase treatment. These data suggested to us that the phosphorylated G protein probably does not efficiently interact with the receptor. For this reason and because of the chromatography improvements outlined in section 2.5, we routinely treated the G protein with λPPase during purification. As a precaution against the possibility that the dephosphorylation of the receptor or G protein was not fully completed during the purification process, we also added the phosphatase into the complexing reaction mixture to ensure the removal of any remaining phosphates on either the receptor or G protein.

2.11 – Receptor and G protein complex has 1:1 stoichiometry.

Another question we needed answered about our β_2AR-Gs complex samples was the overall stoichiometry of the different subunits. In some of the lab's previous work it was shown that a single receptor was all that was required to activate the G protein,[80, 81] but at the time there were also many papers suggesting that normally a

dimer of receptors was responsible for activation in physiological contexts.[82, 83] We could not use Coomassie blue staining to test the ratio of β_2AR to Gs because the receptor stained poorly, so instead we used two other techniques to estimate the stoichiometry of the complex; SDS-PAGE analysis of fluorescamine labeled proteins and total amino acid analysis.

Fluorescamine is a non-fluorescent spiro compound that will quickly react with primary amines, such as those found on lysine residues or non-capped peptide N-termini, to form a fluorophore that is covalently attached through the amine. It absorbs photons from the UVA region and emits blue light, allowing detection of proteins in gel using standard UV transilluminators. We labeled and purified an SDS-denatured sample of β_2AR-Gs complex with fluorescamine and ran the sample on an SDS-PAGE gel. The ratio of the intensities of the bands is compared with the theoretical predictions based on the number of lysine residues in each protein and a stoichiometry of receptor to G protein of either 1:1 or 2:1. In figure 2-9 below, it can be seen that the predicted signal ratio is much closer to the experimentally derived ratio if only a single receptor is assumed to be included in the β_2AR-Gs complex.



Theoretical ratios of fluorescamine signal for labeled subunits:

Monomeric receptor

Gαs	:	Gβ1	:	β_2AR
1.00	:	0.44	:	0.60

Dimeric receptor

Gαs	:	Gβ1	:	β_2AR
1.00	:	0.44	:	1.20

Experimental ratio of fluorescamine signal for labeled subunits:

Gαs	:	Gβ1	:	β_2AR
1.00	:	0.51	:	0.45

Figure 2-9. Fluorescamine labeling of β_2AR-Gs complex. A sample of purified β_2AR-Gs complex was labeled using the method given in section 6.10 and analyzed by SDS-PAGE. Quantification of the signal shows much better agreement with a 1:1 ratio of receptor to G protein that the alternate model of a 2:1 ratio. Data was obtained by Søren Rasmussen according to the method derived by both Brian DeVree and Søren Rasmussen.

In addition to fluorescamine labeling, we also tested the ratio of receptor to G protein in our samples by amino acid analysis. The sample was hydrolyzed in hot HCl and analyzed by HPLC by the University of Michigan Protein Structure Facility. Tryptophan and cysteine cannot be measured by this technique, and asparagine and glutamine are converted to aspartic acid and glutamic acid, respectively, so 16 values incorporating the abundance of 18 of the amino acids are reported from the analysis. We used an ordinary least squares optimization routine and the known amino acid sequences to find the most likely ratio of the different proteins in the sample. The optimized Gαs : Gβ1-Gγ2 : β_2AR ratio from the measurement of three replicates was 0.75 : 1.45 : 1.00. In figure 2-10 below, the least squares error is shown as a surface function of the ratio of Gαs and Gβ1-Gγ2 to the β_2AR, which is held at a relative concentration equal to 1.00.

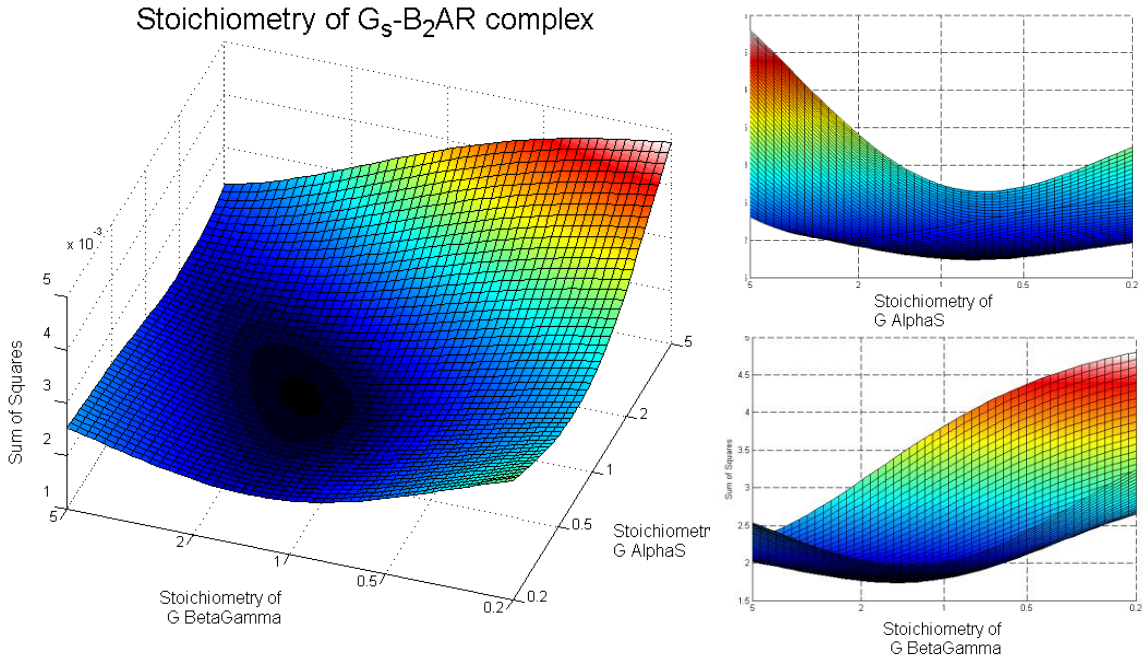


Figure 2-10. Ordinary least squares optimization of amino acid analysis on β_2AR-Gs complex. The receptor is assumed to be at a stoichiometry = 1.00, and the error of the predicted amino acid composition versus the experimental measurement is plotted as a function of the stoichiometery of the Gαs and Gβ1-Gγ2 subunits.

It can be seen that the amino acid analysis shows an excess of Gβ1-Gγ2 subunits in our preparations, but is not highly suggestive of having multiple receptors per complex. Taken together with the fluorescamine data, which indicated a slight deficit

of receptor, we determined that we probably did not have 2 receptors per complex molecule. However, both assays showed higher than expected levels of Gβγ subunits, so we added an additional M1 FLAG antibody affinity column to our evolving purification protocol for separating the FLAG-tagged receptor-G protein complex from unreacted G protein subunits. It was considered highly unlikely that the excess of Gβγ subunits was due to any true physiological reason, since it had long been known that the alpha and beta subunits associate in a 1:1 ratio using both biochemical and crystallographic methods.[69, 84]

2.12 – Maltose neopentyl detergents enhance complex stability.

After we had confirmation that we could trap the β_2AR and Gs proteins together in a purifiable complex with reasonable stoichiometry, we started working to enhance the stability of our molecular complex while simultaneously attempting crystallography using the best protocol that we had developed at that point. One of the problems with our samples was that the β_2AR-Gs complex in DDM detergent shows significant dissociation after 48 h at 4 °C, shown in figure 2-11A below. We screened and characterized existing detergents as well several new families of amphiphiles made by Samuel Gellman and identified the new MNG class[85] of amphiphiles as detergents that were substantially better than DDM at stabilizing the complex, shown below.

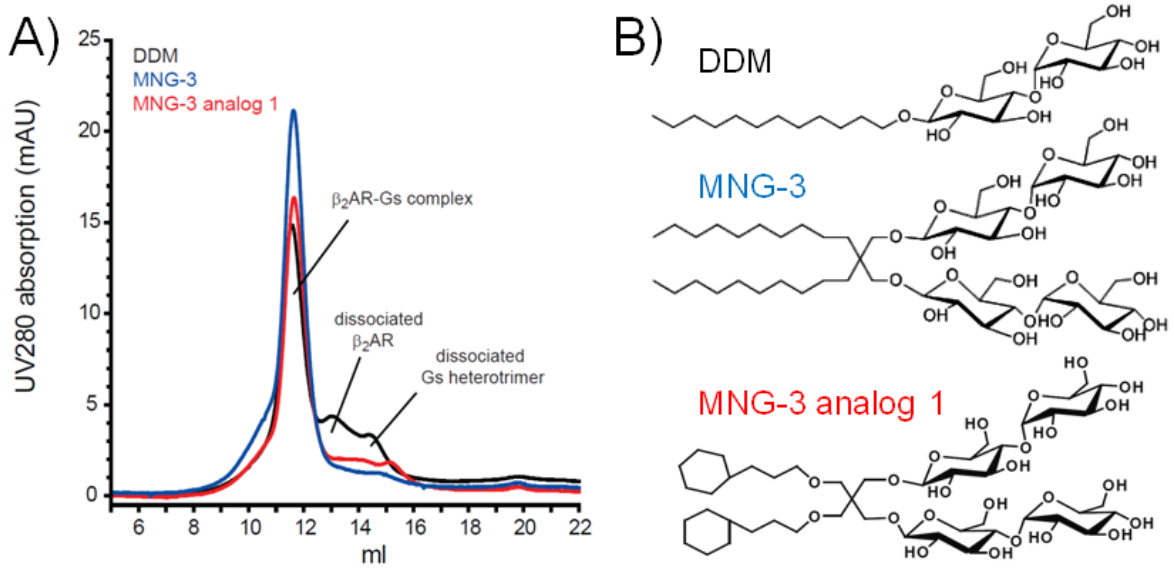


Figure 2-11. MNG amphiphiles enhance receptor-G protein stability. A) Gel filtration of β_2AR-Gs complex samples in various detergents after incubation at 4° for 48 hours. Notice there is significant amounts of dissociated receptor and G protein in the typically used DDM detergent, but less in the two MNG family detergents tested. Figure reprinted from Rasmussen et al.[86] B) Structures of the three amphiphiles used in part A. Data was collected by Søren Rasmussen using samples prepared by Søren Rasmussen and Brian DeVree. MNG detergents were synthesized by Pil Seok Chae.

In addition to keeping the β_2AR-Gs complex together better than DDM during purification, the MNG detergents also had the beneficial property that they tended not to dissociate very well from the receptor when the solution was diluted below the detergent's critical micelle concentration. However, free micelles containing only detergent and no receptors rapidly dissociated upon dilution, allowing us to prepare samples for electron microscopy that contained β_2AR-Gs complexes with no free micelles to complicate particle identification.

2.13 – Nanobody 35 enhances complex stability.

We also worked with Jan Steyaert to develop camilid antibody fragments (nanobodies) to stabilize the complex. Work with our laboratories[87, 88] and many other ones[89-92] has shown that these nanobodies can bind proteins and protein complexes in ways that enhance the stability of particular conformations of the target protein. They often also promote crystallization, in most cases providing numerous crystal contacts. We immunized two llamas with crosslinked β_2AR-Gs complex,

generated phage display libraries of nanobodies from cDNA libraries of their peripheral lymphocytes, and screened the library according to the method given in section 6.11. Several nanobodies were found that bound to the complex, including nanobody 35 (NB35) and 37 (NB37). NB35 prevented complex dissociation, shown below, and was eventually found to be particularly good at promoting high-quality crystal formation.

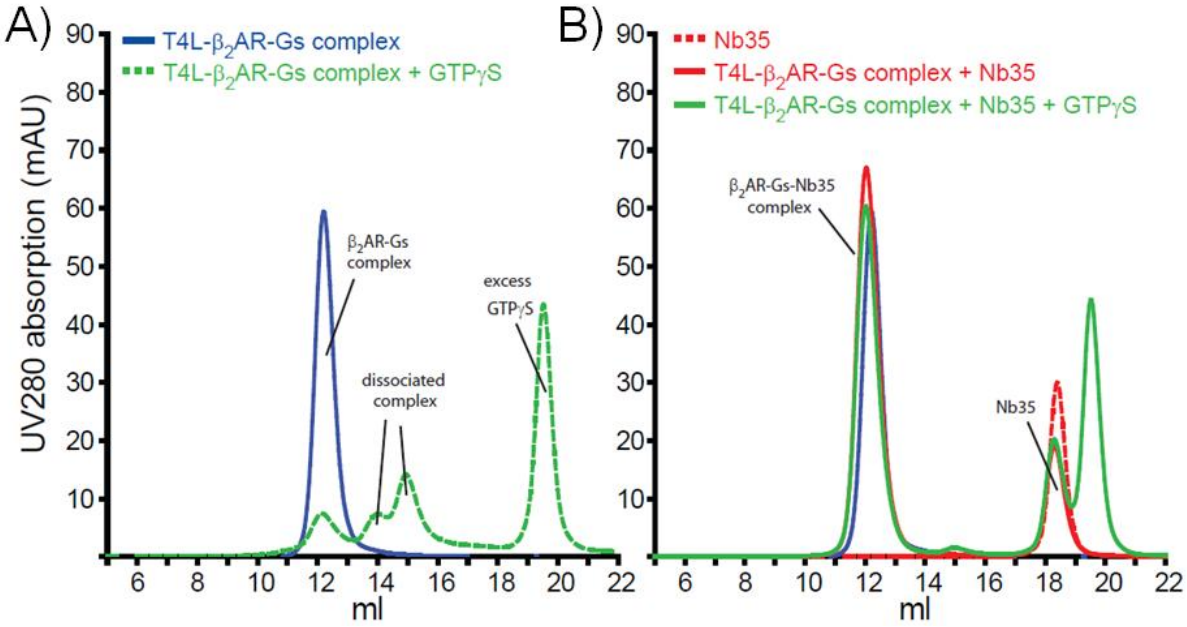


Figure 2-12. Nanobody 35 prevents dissociation of receptor-G protein complex. A) Gel filtration of a non-nanobody bound β_2AR-Gs complex. The complex can be very effectively dissociated into free receptor and G protein subunits by adding 100 μM of the non-hydrolysable GTP analogue GTPγS. B) Gel filtration of a β_2AR-Gs complex bound to NB35. The nanobody-bound complex is very slightly larger than the complex alone, and cannot be dissociated into components by 100 μM GTPγS. Data was collected by Søren Rasmussen using samples prepared by Søren Rasmussen and Brian DeVree. Figure reprinted from Rasmussen et al.[86]

In part A) of figure 2-12 above, a sample of β_2AR-Gs complex can be dissociated with GTPγS, but in part B) the same β_2AR-Gs complex is resistant to dissociation when NB35 is added at a 1:1.2 complex:NB35 stoichiometry. The nanobody clearly helped the complex stay together, so we added it into our protocol for complex preparation for crystallography.

2.14 – High affinity agonists, lipidic cubic phase, and T4L fusion proteins promote crystallogenesis.

Some other refinements we made to our protocol for making β_2AR-Gs complex samples for crystallography was to use a very-high affinity agonist to stabilize the activated form of the receptor, using a specialized cubic-phase forming lipid, and to fuse a T4 lysozyme protein on the N-terminal (extracellular) side of the receptor. The agonist BI-167107, shown below in figure 2-13A, was identified in Dr. Kobilka's lab using a receptor melting temperature screen of several dozen proprietary and non-proprietary high-affinity agonists. It was one of a series of ligands provided by Boehringer Ingelheim and was previously used to crystallize the receptor bound to a G protein mimic nanobody, NB80.[87]

In the past, we had also had good results incorporating the receptor into lipidic cubic phase (LCP) based matricies to promote crystallogenesis,[87, 93, 94] so for our crystallography of the β_2AR-Gs complex samples we also focused our efforts on using LCP matricies. However, the typical lipid used for forming LCP, 9.9 MAG, shown in figure 2-13B, formed a cubic phase with aqueous channels of 40 Å diameter.[95] We thought such a small channel was likely to restrict diffusion of the soluble parts of the complex, which totaled about 120 kDa of protein. (For comparison, the Stokes diameter of carbonic anhydrase, a 29 kDa protein, is also 40 Å.) Instead, we used the lipid 7.7 MAG, also shown below in figure 2-13B, to make a cubic phase which had larger channels of about 62 Å.[95] These channels were evidently large enough not to inhibit diffusion and crystal nucleation of the β_2AR-Gs complex.

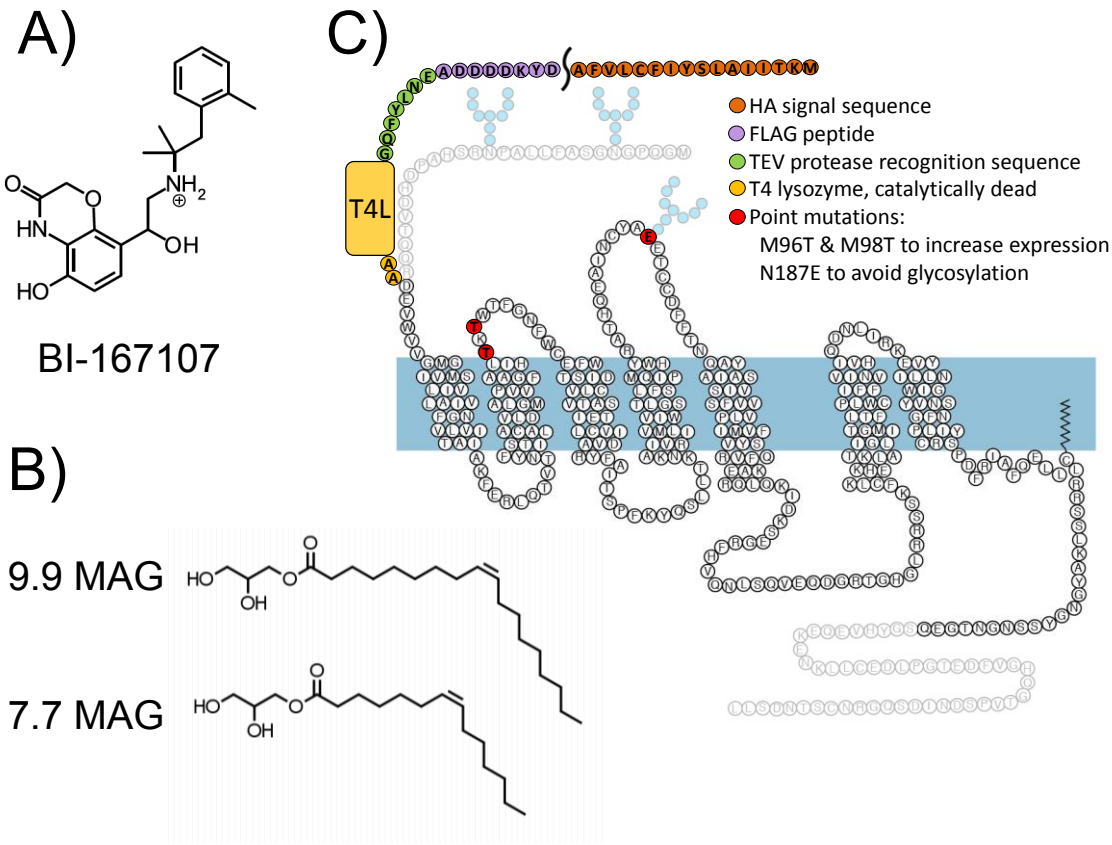


Figure 2-13. Ligand, lipid matrix, and receptor construct used for crystallography. A) The high-affinity agonist BI-167107 used to stabilize the receptor-G protein complex. It has a K_D of 84 pM for the receptor. B) Lipids used for making lipidic cubic phase. The 9.9 MAG (monoacylglycerol) lipid was initially used, but was replaced by the 7.7 MAG lipid in order to get larger aqueous pores in the cubic phase. C) Schematic of the receptor construct used for overexpression. Faded residues represent parts of the wild type receptor that were not included in the crystallization construct. Figure modified from Kashai et al.[96] The G proteins were completely wild type, with the exception of 6 histidine residues added to the N-terminus of the Gβ subunit.

Dr. Kobilka also established the strategy of replacing the intercellular loop three of the receptor with a T4 lysozyme fusion protein to promote receptor crystallogenesis.[87, 93, 94] However, for the β_2AR-Gs complex we fused the lysozyme to the N-terminal extracellular portion of the receptor and we left the third loop unmodified in order to allow unhindered interaction of the receptor with the G protein. A diagram of the entire β_2AR used for crystallography is given above in figure 2-13.

2.15 – Section summary.

At this point in the project, we were able to routinely make and purify milligram quantities of β_2AR-Gs complex. We regularly attempted crystallography screens with the best protocols we had available, but the process of getting well diffracting crystals took several years. During that time, we would also send small amounts of the β_2AR-Gs complex samples to Dr. Virgil Woods and Dr. Georgios Skiniotis for analysis by deuterium exchange and electron microscopy, respectively, in order to determine what could be improved in our sample preparations. The data they collected with our samples were also useful on their own right aside from helping with the crystallography effort, and they are described in detail in section 3 along with the crystal structure.

CHAPTER 3
STRUCTURE AND DYNAMICS OF A RECEPTOR-G PROTEIN COMPLEX

3.1 – Introduction.

This section describes data that was collected as part of a large collaboration between 8 laboratories and published in a series of high-impact publications in 2011.[86, 88, 97] The completion of the project was a longstanding goal in the GPCR field and was cited as a major factor that led to the awarding of the 2012 Nobel Prize in Chemistry to Dr. Brian Kobilka along with Dr. Robert Lefkowitz.

My own part in the work was to determine the initial conditions and strategy for coupling receptor and G protein, described in the previous chapter, and later on to produce G protein in large quantities for crystallography sample preparation. I also helped with crystallographic data collection and figure preparation. In lieu of detailing each collaborator's contribution in the text of the chapter, I have instead compiled a list of principal investigators and students/postdocs that were involved in the project. Also listed are the major contributions from each person.

Principal Investigators:

Martin Caffrey - Lipid synthesis and production for lipidic cubic phase crystallography.

Samuel H. Gellman - Detergent synthesis and production.

Brian Kobilka - Project design, major funding, receptor-G protein biochemistry, receptor production, crystallography and crystallographic data collection, manuscript preparation.

Georgios Skiniotis - Project design, funding, electron microscopy, manuscript preparation.

Jan Steyaert - Nanobody screening and production.

Roger Sunahara - Project design, major funding, receptor-G protein biochemistry, G protein production, crystallographic data collection and analysis, manuscript preparation.

William Weis - Crystallographic data analysis.

Virgil L. Woods Jr. - Deuterium exchange by mass spectrometry.

Students/Postdocs:

Diane Calinski - Sunahara lab; G protein production.

Pil Seok Chae - Gellman lab; Detergent design and synthesis.

Ka Young Chung - Kobilka lab; Deuterium exchange by mass spectrometry, receptor-G protein biochemistry .

Brian T. DeVree - Sunahara lab; G protein production, receptor-G protein complexation and purification, receptor-G protein biochemistry, crystallographic data collection and analysis, figure preparation.

Somnath Dutta - Skiniotis lab; electron microscopy data collection.

Tong Sun Kobilka - Kobilka lab; receptor production.

Andrew C. Kruse - Kobilka lab; crystallography, crystallographic data collection and analysis.

Sheng Li - Woods lab; Deuterium exchange by mass spectrometry.

Tong Liu - Woods lab; Deuterium exchange by mass spectrometry.

Joseph A. Lyons - Caffrey lab; lipid production for lipidic cubic phase crystallography.

Jesper M. Mathiesen - Kobilka lab; receptor biochemistry.

Austin N. Oleskie - Skiniotis lab; electron microscopy data collection.

Els Pardon - Steyaert lab; nanobody screening and production.

Søren G. F. Rasmussen - Kobilka lab; receptor production, receptor-G protein complexation and purification, receptor-G protein biochemistry, crystallography, crystallographic data collection, manuscript and figure preparation.

Syed T. A. Shah - Caffrey lab; lipid production for lipidic cubic phase crystallography.

Min Su - Skiniotis lab; electron microscopy data collection.

Foon Sun Thian - Kobilka lab; receptor production.

Gisselle A. Vélez-Ruiz - Sunahara lab; receptor-G protein biochemistry.

Gerwin H. Westfield - Skiniotis lab; electron microscopy data collection and data analysis, low-resolution structure reconstruction.

Yaozhong Zou - Kobilka lab; receptor biochemistry, crystallographic data collection and analysis.

3.2 – Crystallographic data collection, model solving, and statistics.

The highly-stabilized T4L-β_2AR-Gs-Nb35 complex samples crystallized in the primitive monoclinic space group $P2_1$ from a PEG 400 and potassium nitrate based crystallization buffer at pH 6.5, shown below in figure 3-1 (detailed methods given in section 6.13). Diffraction patterns from the crystal were collected at Argonne National Laboratory on the Advanced Photon Source beamline 23 ID-B. The beamline's microfocus capabilities (5 μm diameter beam) and ability to search for

diffraction spots by screening through the volume of the opaque lipidic cubic phase matrix that the crystals were embedded in were critical to successful data collection.

Figure 3-1. T4L-β_2AR-Gs-Nb35 complex crystals used for data collection. Scale bar represents 100 µm. Figure reprinted from Rasmussen et al.[86]

Reflections from 20 different crystals were merged to create the final data set, and the electron density of the crystal was determined by molecular replacement with previously determined crystal structures of the individual proteins, followed with an iterative process of automatic atomic position refinement and manual model adjustment and residue building. The crystals were anisotropic, with reflections along one axis going out to 2.9 Å and the other two going out to only 3.2 Å. Although the structure is reported as a 3.2 Å structure, the extra 2.9 Å reflections were included in the solution and aided in defining the electron density more accurately than would normally be expected from a 3.2 Å data set. The crystallographic statistics are shown below in table 3-1, and full details about solving the structure are given in method 6.14.

Data collection	
Number of crystals	20
Space group	P 2_1
Cell dimensions	
a, b, c (Å)	119.3, 64.6, 131.2
a, b, g (°)	90.0, 91.7, 90.0
Resolution (Å)	41 – 3.2 (3.26 – 3.20)
R_{merge} (%)	15.6 (55.3)
<I>/<σI>	10.8 (1.8)
Completeness (%)	91.2 (53.9)
Redundancy	6.5 (5.0)
Refinement	
Resolution (Å)	41 – 3.2
No. reflections	31075 (1557 in test set)
R_{work}/R_{free} (%)	22.6 / 27.8
No. atoms	10,275
No. protein residues	1,318
Anisotropic *B* tensor	B_{11}= -6.4 / B_{22}= 3.8 / B_{33}= 2.6 / B_{13}= 1.9
Unmodelled sequences	
β_2 adrenergic receptor	29, 176-178, 240-264, 342-365
Gs α	1-8, 60-88, 203-204, 256-262
Gs γ	1-4, 63-68
T4 lysozyme	161
Average *B*-factors (Å^2)	
β_2 adrenergic receptor	131.6
Gs α, ras domain	81.4
Gs α, helical domain	121.9
Gs β	63.0
Gs γ	83.6
Nanobody 35	59.5
T4 lysozyme	112.1
R.m.s. deviation from ideality	
Bond length (Å)	0.007
Bond angles (°)	0.71
Ramachandran statistics	
Favored regions (%)	95.4
Allowed regions (%)	4.6
Outliers (%)	0

Table 3-1. Crystallographic statistics for T4L-β2AR-Gs-Nb35 complex crystals. Highest shell statistics are in parentheses. All purification tags and the listed regions were omitted from the model due to poorly resolved electron density. Table reprinted from Rasmussen et al.[86]

The R_{work} and R_{free} factors for the refinement of the crystal structure are typical or slightly below average for a structure of 3.2 Å,[98] indicating with a well optimized model and consistent with the fact that some reflections of higher resolution than 3.2 Å were included in the final data set. It also help that the crystal contains many regions of well ordered and well packed domains, such as the lysozyme and nanobody that were added to promote crystallography and the very stable β-propeller fold of the Gβ subunit. The quite low r.m.s. deviation from ideality statistics are a function of our conservative strategy while model building. We generally preferred to keep bond lengths and angles close the average values since there were no truly high-resolution structures available of the biologically interesting molecules in the crystal to accurately guide modeling of non-standard residue orientations.

3.3 – Crystallographic packing.

The β_2AR-Gs complex packed into a crystal with alternating flat layers of lipid and aqueous domains, shown below in figure 3-2.

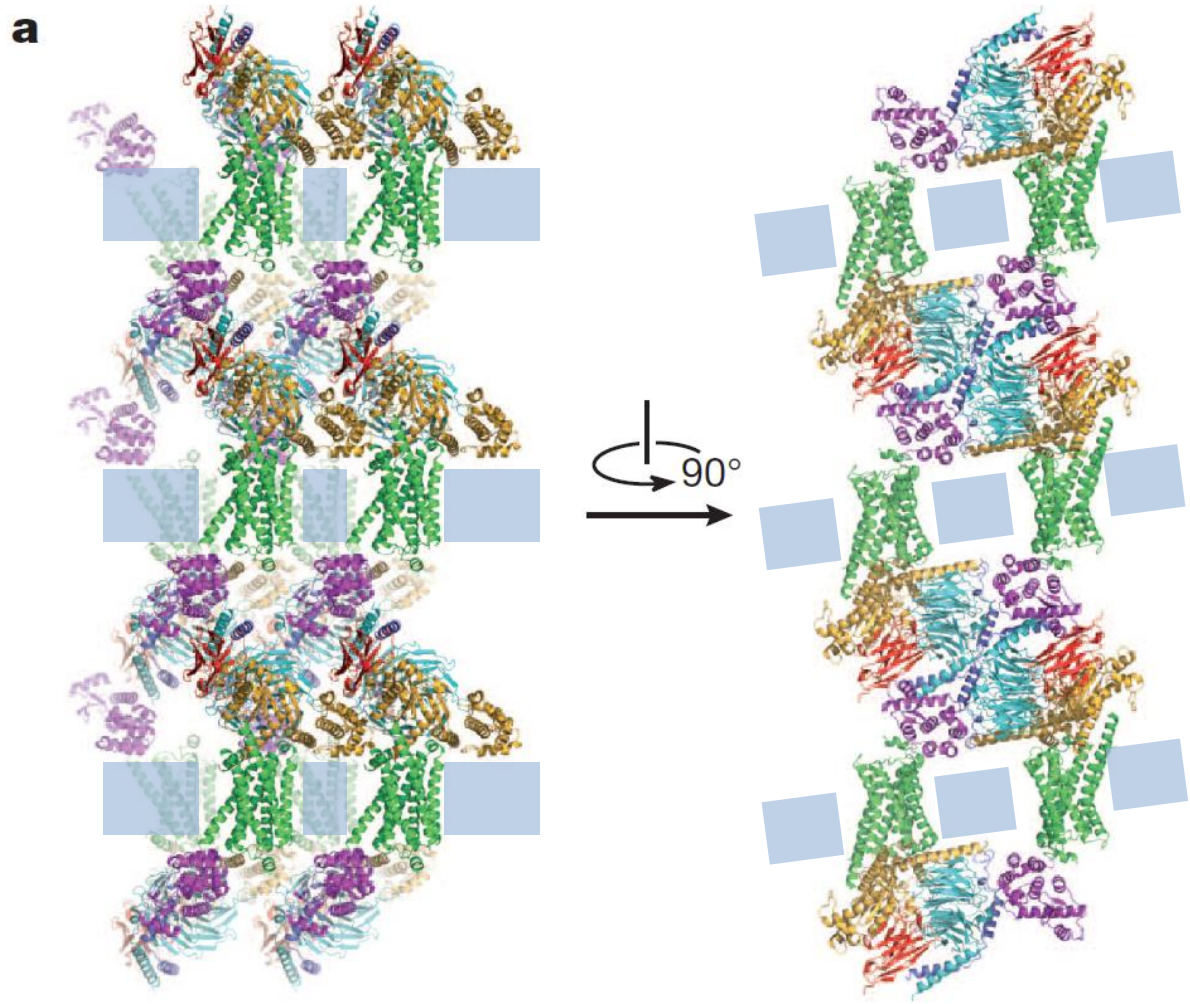


Figure 3-2. Crystal packing lattice of the β_2AR-Gs complex crystals. The receptor, shown in green, packs into the lipid layers, indicated as grey squares. All other proteins pack into the soluble layer of the crystal. Figure reprinted from Rasmussen et al.[86]

In the aqueous domain the various proteins are packed very tightly, utilizing many crystal contacts that formed with residues from the nanobody and lysozyme proteins, as expected. Only the transmembrane part of the receptor packed into the lipid domain of the crystal, with the receptors spaced widely apart from each other. Absolutely no contacts are made between adjacent receptors in the crystal, which is consistent with previous findings from our lab that a single receptor is sufficient to activate a G protein.[77, 80, 81] Also, the lack of crystal contacts between the receptor and the rest of the lattice results in relatively high *B*-factors for the receptor compared to the rest of the structure. A color map of the *B*-factors is shown below in figure 3-3.

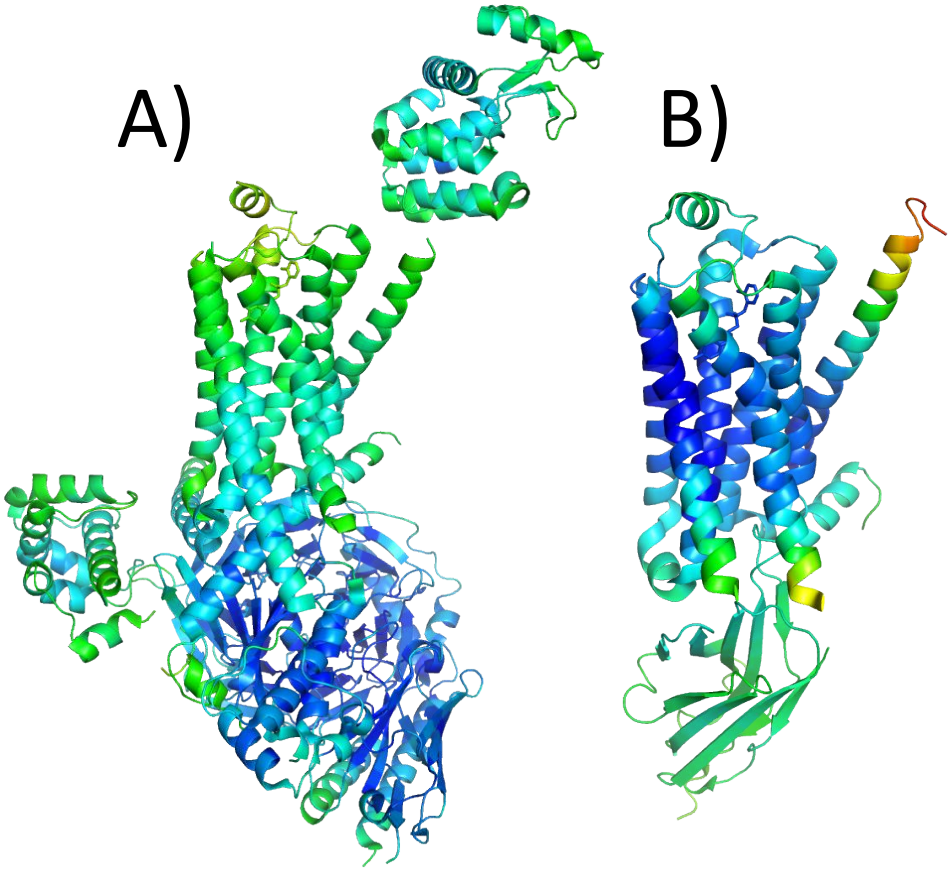


Figure 3-3. *B*-factor heatmaps of active-state β_2AR crystal structures. Blue residues represent the lowest *B*-factors and red residues are the highest *B*-factors. A) *B*-factors of the β_2AR-Gs complex structure (PDB 3SN6). B) *B*-factors of the β_2AR-NB80 complex structure (PDB 3P0G).

It can also be seen that the extracellular portion of the receptor has particularly high B-factors, which makes detailed conclusions about the side chain orientation in receptor's ligand binding site difficult. However, comparison of this structure with the structure of the nanobody 80 bound receptor, with was crystallized with the exact same ligand, allows us to determine that the overall shape of the ligand binding site is very similar between the two structures and almost all the differences between them are in the G-protein binding intercellular half of the receptor. One can also see that the α-helical domain of the Gα subunit also has somewhat elevated *B*-factors relative to the rest of the G protein, which is possibly a reflection of its overall increased mobility after nucleotide loss, which is discussed more in the context of electron microscopy in section 3.9.

3.4 – Overview of structure.

The overall structure is consistent with previously determined biochemical information about the GPCR-G protein complex. Techniques like electron paramagnetic resonance[99] and fluorescence quenching[100] have shown that that activation of the receptor causes large movements of the cytoplasmic side of transmembrane helix 6 (TM6) away from the core of the receptor. The complex structure also shows a 14 Å outward movement of TM6. This movement creates a large binding pocket on the intercellular face of the receptor, in which the C-terminus of the Gα subunit binds. Work with domain swapping of different Gα subunits,[101, 102] NMR and peptide binding,[103] mutational analysis,[104] and more recently crystallography[105] has established that the C-terminus of the Gα subunit and the cleft opened by TM6 movement are the primary, though not the only, regions of interaction between the receptor and G protein. Both the Gα and Gβγ subunits are present in the structure, which we expected because both subunits were found in an approximately 1:1:1 ratio of Gα : Gβγ : receptor during our purification (section 2.11) and because activation of the Gα subunit has been shown to be quite in absence of the βγ subunit.[106] An overview of the structure is shown below in figure 3-4.

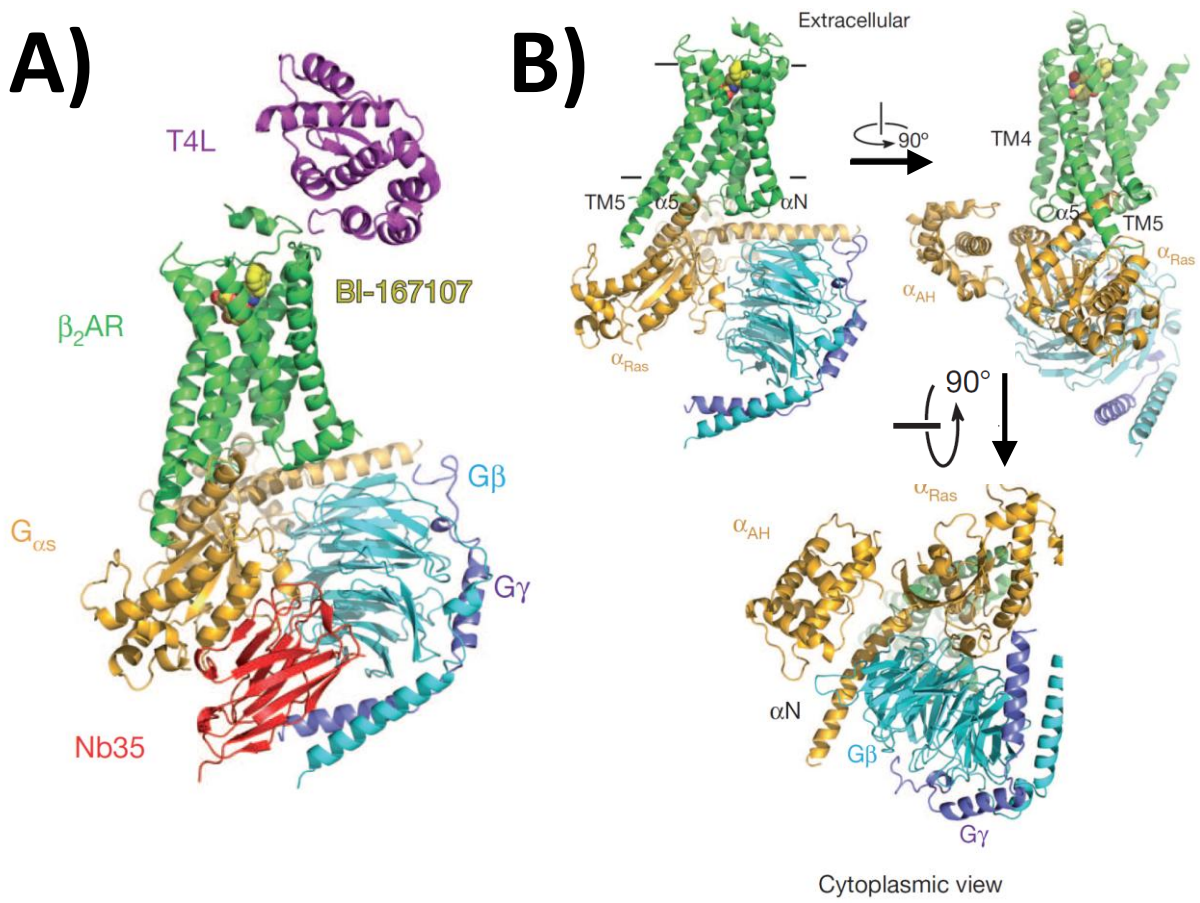


Figure 3-4. Overview of the β_2AR-Gs crystal structure. Figure modified from Rasmussen et al.[86] Please see text for details.

A genuine surprise in the structure was the extremely large movement of the Gsα α-helical domain relative to the Ras-like domian. The center of the domain moves by about 45 Å, with the individual residues moving by 12 to 86 Å. The domain is rotated by 127° around the hinge region where it is attached to the Ras-like domain. The movement removes the sugar binding part of the GTP/GDP binding site on the Gα subunit, disrupting it in a manner consistent with the nucleotide-free nature of the β_2AR-Gs complex.[77] Previously, most proposed models of G protein activation assumed little or no movement of the α-helical domain, instead suggesting that nucleotide entered and exited the Gα subunit near the switch I and II domains after Gβγ-aided movement of the β3-α2 loop away from the nucleotide binding site.[107] Recently, a couple studies suggested that much larger movements of the α-helical

domain are associated with GPCR-stimulated nucleotide exchange[108, 109] and that the nucleotide-free receptor and G protein complex is quite conformationally flexible.[110] Our data clearly support the latter findings.

The complex structure also shows how nanobody 35 (NB35) binding enhances the stability of the complex and prevents dissociation. It binds at the interface of both the Gα and Gβ subunits, with the complementarity determining region (CDR) 1 interacting primarily with Gβ and a long CDR3 loop interacting with both Gα and Gβ. These interactions help hold the two subunits together, keeping the complex intact for the several days that are needed for crystallization. The nanobody also prevents GTP binding (shown in section 2.13) by forcing the loop 2 region of the Gα subunit to remain bound to the β subunit, keeping it in the inactive conformation that cannot bind the third phosphate group of GTP.

3.5 – Structural analysis of the receptor in the receptor-G protein complex.

One of the most important questions that this structure helps answer is how the binding of the β_2AR's natural agonists, epinephrine and norepinephrine, in the extracellular part of the receptor leads to the opening of the G protein binding site in the intercellular part of the receptor. No single dataset can tell us everything, but the structure of the β_2AR-Gs complex and the structure of the NB80 bound β_2AR[97] show the receptor's fully activated G protein-stimulating conformation. That information, combined with a structure from the highly similar β_1AR co-crystallized with isoproterenol,[111] a synthetic epinephrine derivative, gives us a very detailed look into the structural changes that happen in response to epinephrine binding. Comparison of these activated or partially activated conformations with the inactive structure of the receptor bound to the inverse agonist carazolol[52, 94] or antagonist alprenolol[112] shows a significant tightening of the binding pocket is associated with agonist binding. As shown in figure 3-5A below, binding of both alprenolol and the agonist BI-167107 involves interaction of the drugs' hydroxyethylamine backbone to residue D133 on transmembrane domain 3 (TM3) and N312 on TM7. At the other end of the binding pocket, TM5 moves and rotates inward, allowing S203 and S207, which are both on the TM5 helix, to interact with the hydroxyl and amine of the BI-167107

benzoxazinone ring. For native agonists, the catechol hydroxyls would make a similar reaction, shortening the S207-D113 distance from 12.0 Å to 11.4 Å. In contrast, it can be seen that for alprenolol, there are neither any interactions to stabilize the TM5 movement, nor are there any steric hindrances to prevent it, true to the drug's nature as a neutral antagonist.

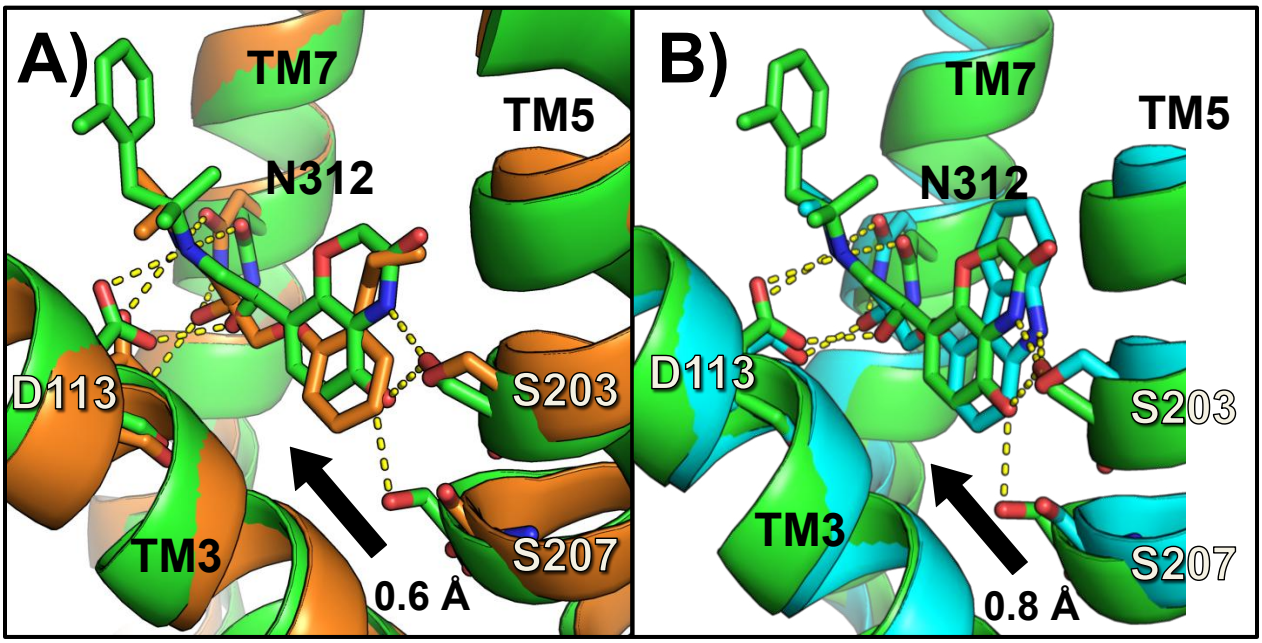


Figure 3-5. Hydrogen bonding of β_2AR ligands. A) BI-167107 and the NB80-bound β_2AR are drawn in green, and the antagonist alprenolol bound to an inactive β_2AR structure is drawn in orange. B) BI-167107 and the NB80-bound β_2AR are drawn in green, and the inverse agoinist carazolol bound to an inactive β_2AR structure is drawn in cyan.

The tightening of the receptor ligand binding pocket upon activation is slightly more evident when the BI-167107 bound active receptor is compared to the inactive carazolol bound receptor. The S207-D113 distance shortens by ≈0.8 Å, shown above in figure 3-5B. One of the major differences between the antagonist alprenolol and the inverse agonist carazolol is that while alprenolol has no hydrogen bonding interactions with S203, carazolol does hydrogen bond with the residue, but in the inactive position. This prevents the inward movement of the residue during receptor activation and stabilizes the ligand binding site in its inactive conformation, as expected.

The binding site for the adrenergic receptors (and many similar hormone binding GPCRs) is buried partway within the transmembrane region of the protein, so upon activation the tightening of the binding pocket restricts the ability of ligands to freely exchange between the pocket and the bulk solvent. This effect is explored in detail using traditional pharmacological methods in chapter 4, but the structures of the active and inactive receptor easily show why this phenomenon occurs. In figure 3-6 below, the surface representation of the active and inactive receptors are shown. In the active structure (figure part A), the atoms that lie above the binding pocket move close enough together to prevent the agonist from leaving the binding site. They would likely also prevent ligands from entering the binding site if the receptor became activated through basal activity. In particular, the structure shows that the TM7 movement during activation breaks a salt bridge that between residues K305 (on TM7) and D192 (on ECL2), and instead K305 forms a hydrogen bond with the backbone carbonyl of F193 (on ECL2) right above the binding site. This aids the movement of F193 and Y308 (on TM7) towards each other, further capping off the catechol binding-part of the ligand binding site from bulk solvent. The breaking of the K305-D192 salt bridge during activation is also corroborated with solution state NMR studies which demonstrated that an ionic bond involving K305 is broken in an agonist-dependent manner.[113] In comparison, the inactive conformation of the ligand binding pocket (figure part B) has a much larger opening for the ligand to diffuse in and out of. Unlike the F193-K305 hydrogen bond, the D192-K305 salt bridge in the inactive receptor lies off to the right side of the view shown and is not expected to restrict diffusion significantly.

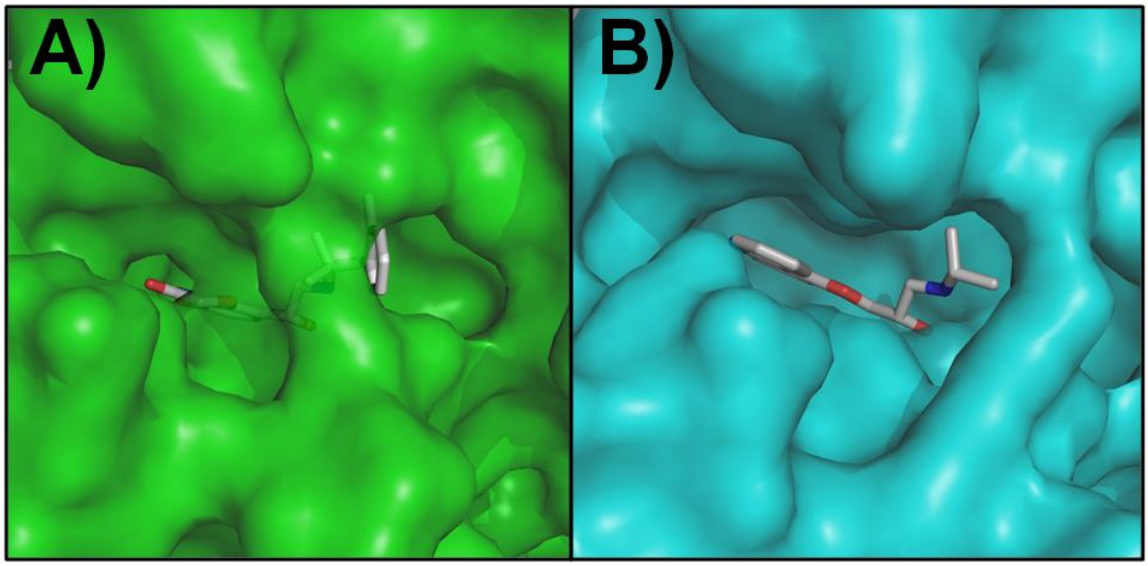


Figure 3-6. Tightening of the β_2AR ligand binding pocket upon activation. A) Surface detail of activated receptor binding pocket. Shown is agonist BI-167107 in NB80-bound β_2AR (PDB 3P0G). B) Surface detail of inactive receptor binding pocket. Shown is inverse agonist carazolol in β_2AR (PDB 2RH1).

The tightening of the ligand binding pocket also promotes movements of the transmembrane helices that ultimately causes the G protein binding site to open on the intercellular face of the receptor, diagrammed below in figure 3-7. The rotation of the extracellular half of TM5 caused by agonist binding also pulls TM4 in slightly closer toward the ligand, allowing for the formation of a new hydrogen bonding contact between Y199 and the backbone carbonyl of T164. At the proline kink of TM5, the helix rotates around the mostly stable TM3 like the pivot of a cantilever, causing the intercellular half of the helix to move away from the receptor core by about 4 Å. This movement is followed and amplified by TM6, which packs along TM5 and moves outward by about 12-14 Å at its intercellular end. The shifting positions of TM3, TM4, and TM5 allow the ICL2 loop to adopt a helical conformation, which is stabilized by a new hydrogen bond that forms between the aspartate of the DRY motif on TM3 and Y141 on the ICL2 helix. TM7 then moves in towards the core of the receptor by ≈1-2 Å along its length, a movement with is aided by agonist binding to N312. The tyrosine residue from the conserved NPxxY motif on TM7 also bulges into the upper part of the gap vacated by TM6 movement, helping to stabilize the open G protein binding site.

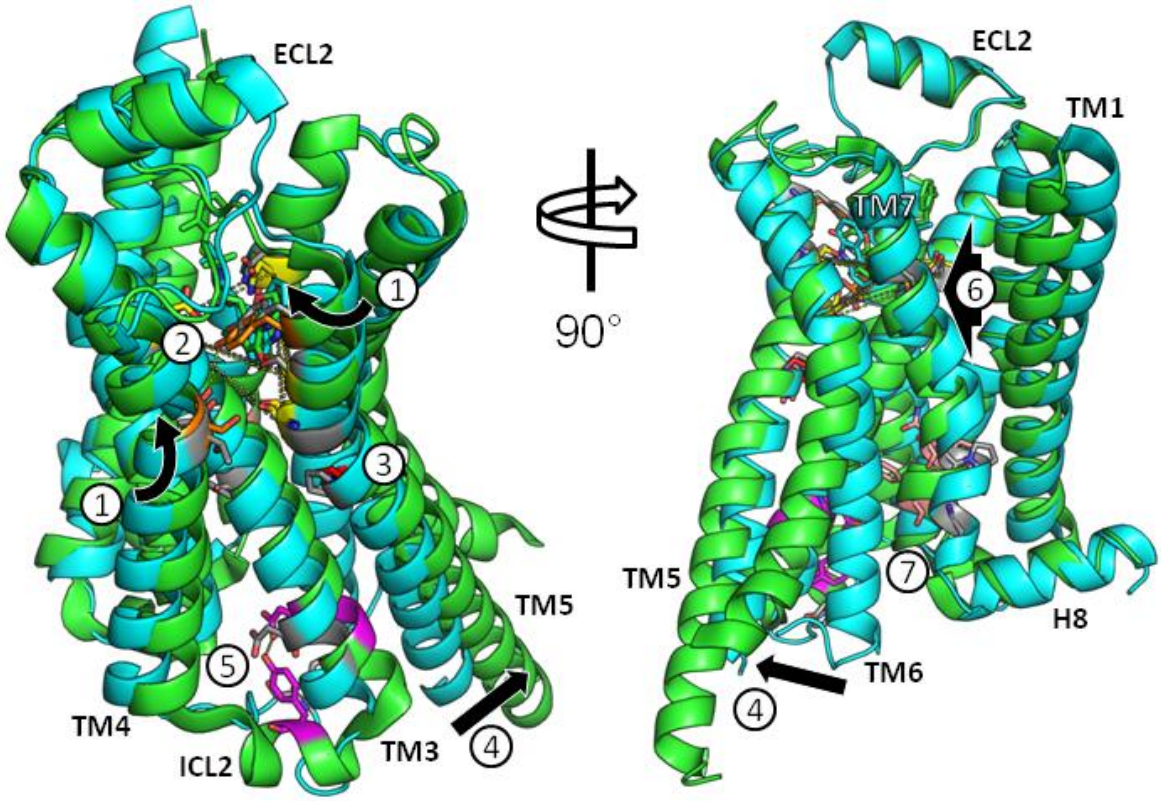


Figure 3-7. Diagram of structural changes during receptor activation. The inactive β_2AR is drawn in cyan (PDB 2RH1) and the active β_2AR is drawn in green (PDB 3SN6). Receptors are aligned using residues 50-120, the most stable part of the receptor that includes TM2 and parts of TM1 and TM3. Numbered features are as follows: 1) Rotation of TM4 and TM5 caused by agonist binding. 2) Formation of hydrogen bond between Y199 and T164. 3) TM5 pivots against TM3 at the helix's proline kink. 4) Intercellular end of TM5 and TM6 moves outward to form G protein binding cleft. 5) ICL2 adopts a helical conformation stabilized by a hydrogen bond between D130 and Y141. 6) TM7 moves in towards the core of the receptor. 7) Y326 of the NPxxY motif moves 7.0 Å into part of the gap created by TM6 movement. Please see text for additional details.

The movement of TM5&6 is also helped by rearrangement of the hydrophobic residues near the proline kink of TM5. In the inactive receptor, I121 from TM3 packs side-by-side with F282, separating it from P211 on TM5. In the activated receptor, I121 moves up towards the ligand binding site and F282 packs underneath it, moving much closer to P211 and allowing TM6 to rotate along with TM5. These changes are shown below in figure 3-8 below.

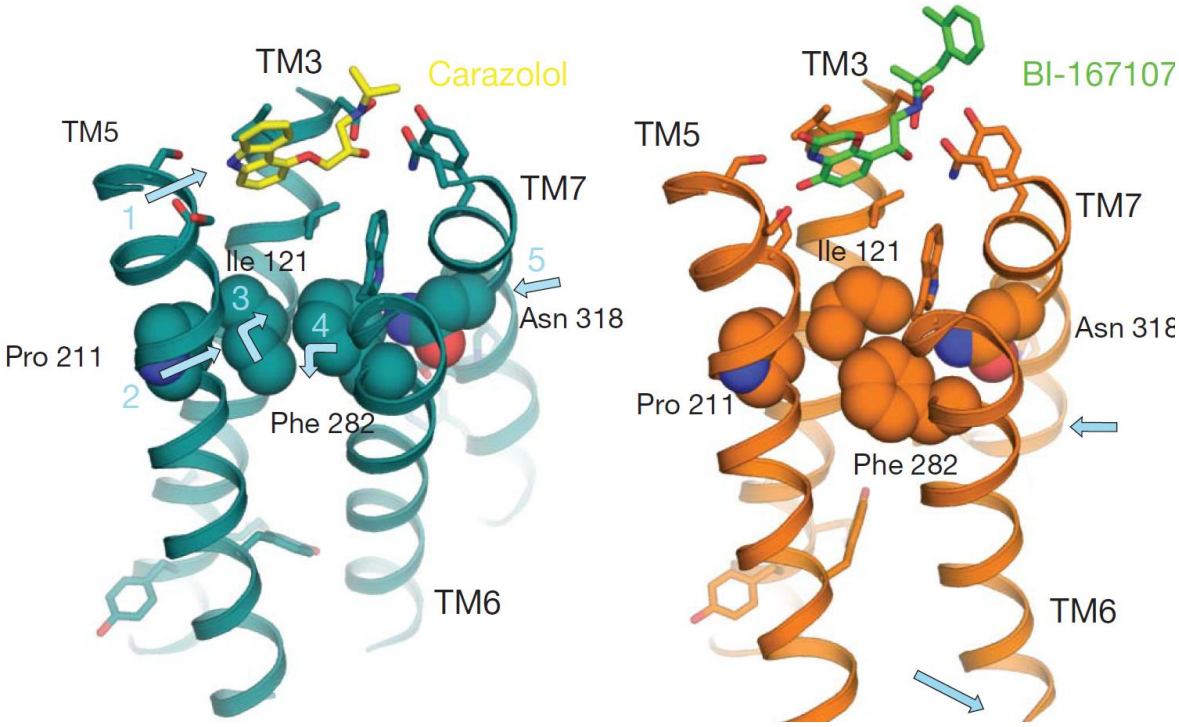


Figure 3-8. Hydrophobic core repacking during receptor activation. Inactive β_2AR is drawn in blue (PDB 2RH1) on the left, and active β_2AR is drawn in orange (PDB 3P0G). Figure taken from Rasmussen et al.[87]

Overall, these TM domain rearrangements lead to an opening of a large cleft between TM5&6 and the core of the receptor, in which the C-terminus of the Gα subunit can bind. Both the NB80 and G protein bound structures show how activation involves many sections of the receptor that are highly conserved in the GPCR family, such as the DRY and NPxxY motifs, demonstrating the reason that these sequences are under evolutionary pressure to remain unchanged.

3.6 – Structural analysis of the receptor-G protein interface.

The interface where the β_2AR and Gαs proteins interact has several structural features that show previously unknown aspects of how the receptor communicates an activation signal to the G protein. As expected, the G protein's C-terminal helix binds to the pocket formed from TM5&6 movement, but the structure also shows a 2-turn extension of the TM5 helix that would not have been able to form in the inactive structures due to the presence of a T4 lysozyme domain fused between TM5 and TM6 that replaced residues in the receptor's intercellular loop 3 (IL3). The TM5 helix extends all the way over the Gα subunit C-terminal helix and terminates

very close to an L-shaped pocket formed by the α4 helix and part of the loop between the α4 and αG helices, shown below in figure 3-9. Although the shape of the pocket is closely contoured to fit the residues of the TM5 helix with many residues within 4 Å of one another, there are no obvious hydrogen bonding or ionic interactions that would enhance the binding of TM5 to the Gα subunit in this region. Instead, it could be that the most important structural aspect in this region is that there are not any steric clashes between the receptor and G protein surfaces. This is discussed more in the context of G protein subtype-specificity in section 5.7.

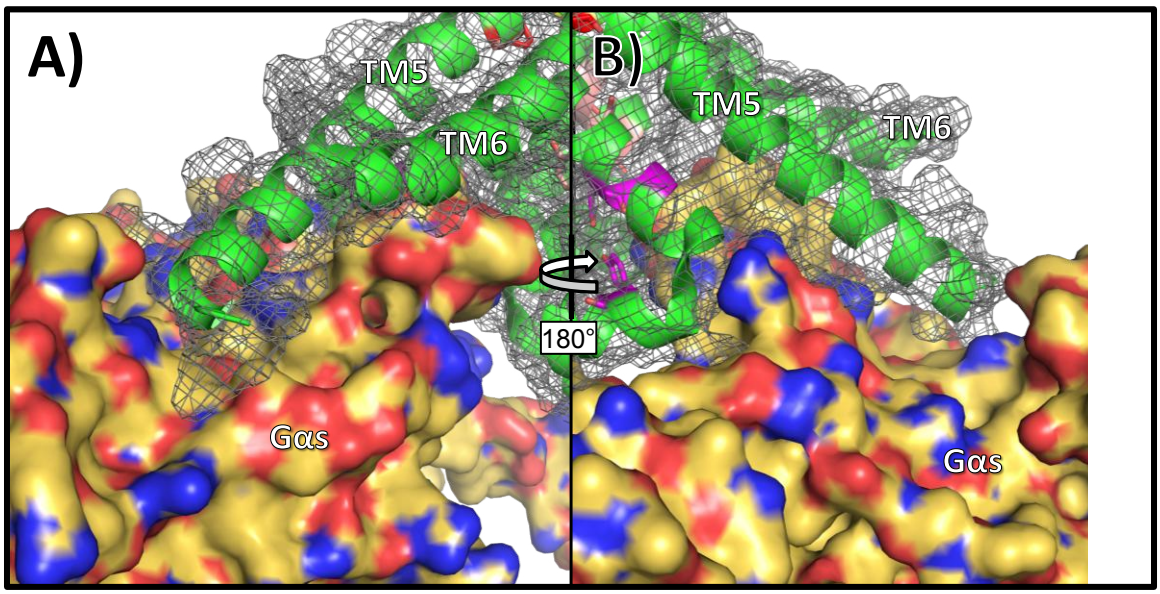


Figure 3-9. Interface between TM5 of the receptor and the α4-αG region of the G protein. The receptor is drawn in green with a grey mesh indicating the Van der Waals surface of the residues. The Gαs subunit is drawn as an orange Van der Waals surface.

The Gα subunit C-terminal helix interacts extensively with the activated receptor, contributing the majority of the 2,576 $Å^2$ of total buried surface area of the β_2AR-Gαs interface. There are two different regions and modes of interaction between the C-terminus and the receptor, drawn below in figure 3-10. In the region shown in figure section A, the middle part of the C-terminal helix forms an extensive ionic and hydrogen-bonding network with residues along the receptor TM5 helix and the backbone amines of the TM3 helix. These interactions involve 4 side chains from the Gα subunit and 4 side chains from the receptor, forming a patch of positive and

negative charges that would be expected to bind relatively strongly together, but would also have to be fully solvated when the two proteins are not interacting.

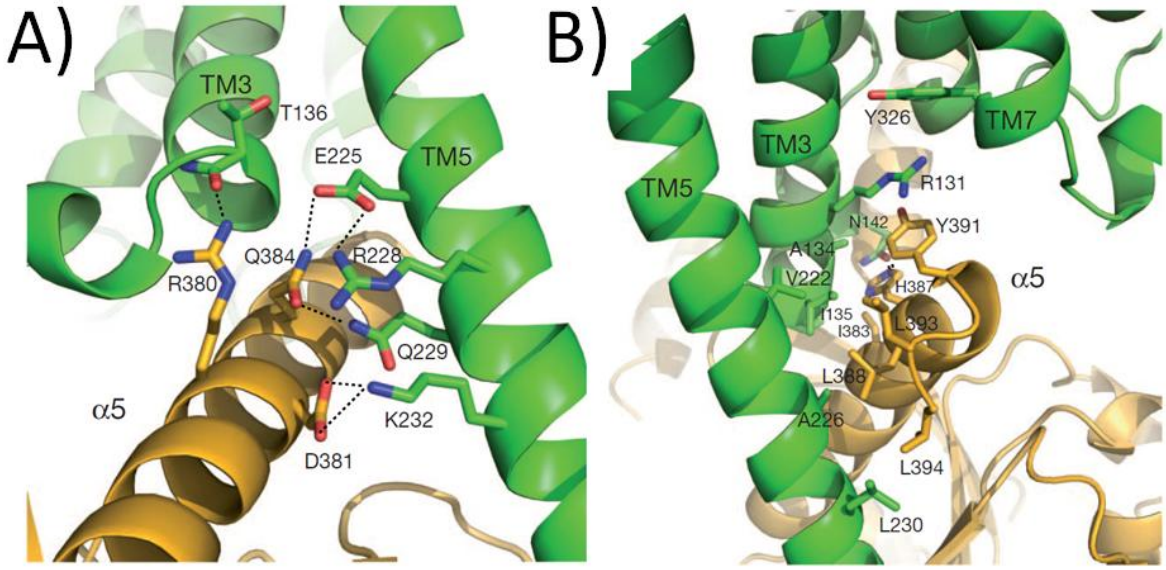


Figure 3-10. Gαs C-terminal helix interactions with the activated β_2AR. A) Network of charged and polar residues that interact via ionic and hydrogen bonds in the middle region of the C-terminal helix. B) Hydrophobic interactions that dominate the interaction between the extreme C-terminal end of the helix and the core of the receptor. Figure reprinted from Rasmussen et al.[86]

At the end of the Gα C-terminus, the interaction between the receptor and the G protein is much more hydrophobic in character. The C-terminal helix extends until the last three residues, which wrap back around the helix as it nears the TM7-H8 turn of the receptor to form a cap at the end of the helix. The binding pocket that surrounds this region of the C-terminus is formed by TM6, TM5, and TM3 and is lined with mostly alanine and leucine/isoleucine residues, which pack along a patch of three leucine residues on the C-terminal helix and cap. The interaction is shown above in figure 3-10B.

Another very important interaction between the receptor and G protein shown below in figure 3-11 involves F139, which is located at the beginning of the ICL2 helix and sits in a hydrophobic pocket formed by Gαs H41 at the beginning of the β1-strand, V217 at the start of the β3-strand and F376, C379, R380 and I383 in the C-terminal helix. This interaction is likely to explain the finding that the β_2AR mutant F139A has severely impaired coupling to Gs.[114] Also of interest, the ICL2 helix is stabilized by

an interaction between D130 of the conserved DRY sequence and Tyr 141 in the middle of the ICL2 helix.

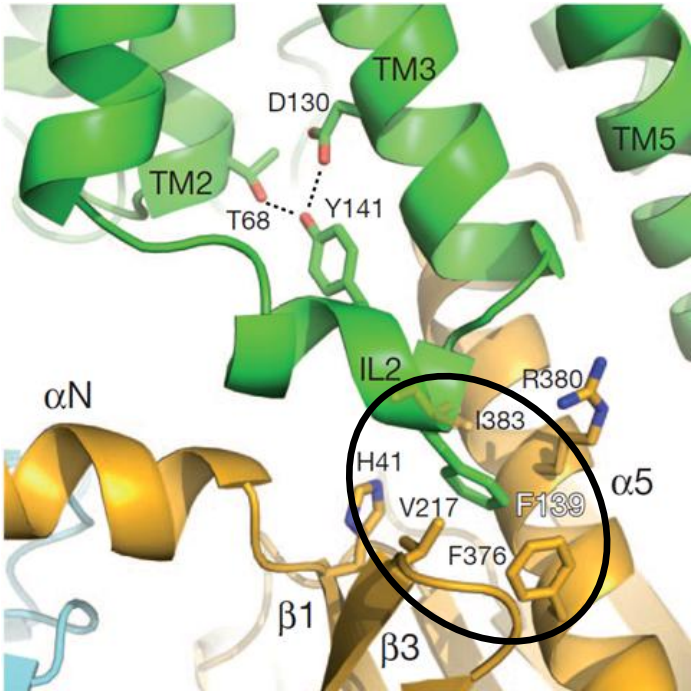


Figure 3-11. Interactions of ICL2 that stabilize Gαs binding. Tyrosine 141 binds to the aspartate of residue of the receptor DRY motif to stabilize formation of the ICL2 helical conformation, and phenylalanine 139 binds a hydrophobic pocket on Gαs formed between the C-terminal helix and the β1-β3 strands. Figure reprinted from Rasmussen et al.[86]

Overall, these interactions lead to stabilization of the activated conformational state of the receptor and a nucleotide-free Gα subunit. The interaction surface between the two proteins includes a large amount of buried surface area and numerous hydrogen and ionic bonds. All these interactions are not accessible on the inactive receptor or the nucleotide-bound G protein, so they serve to stabilize the conformational states of the proteins that pass the activation signal from the agonist binding site of the receptor to the nucleotide binding site of the G protein.

3.7 – Structural analysis of the G protein in the receptor-G protein complex.

The extensive interactions between the β_2AR and Gαs Ras-like domain cause several structural re-arrangements in the Gα subunit that ultimately lead to nucleotide loss. The most direct effect of receptor binding is a displacement of the C-terminal helix by 6 Å towards the receptor as it projects into the transmembrane core

of the β_2AR, accompanied by a rotation of the helix by about 30°, shown below in figure 3-12A.

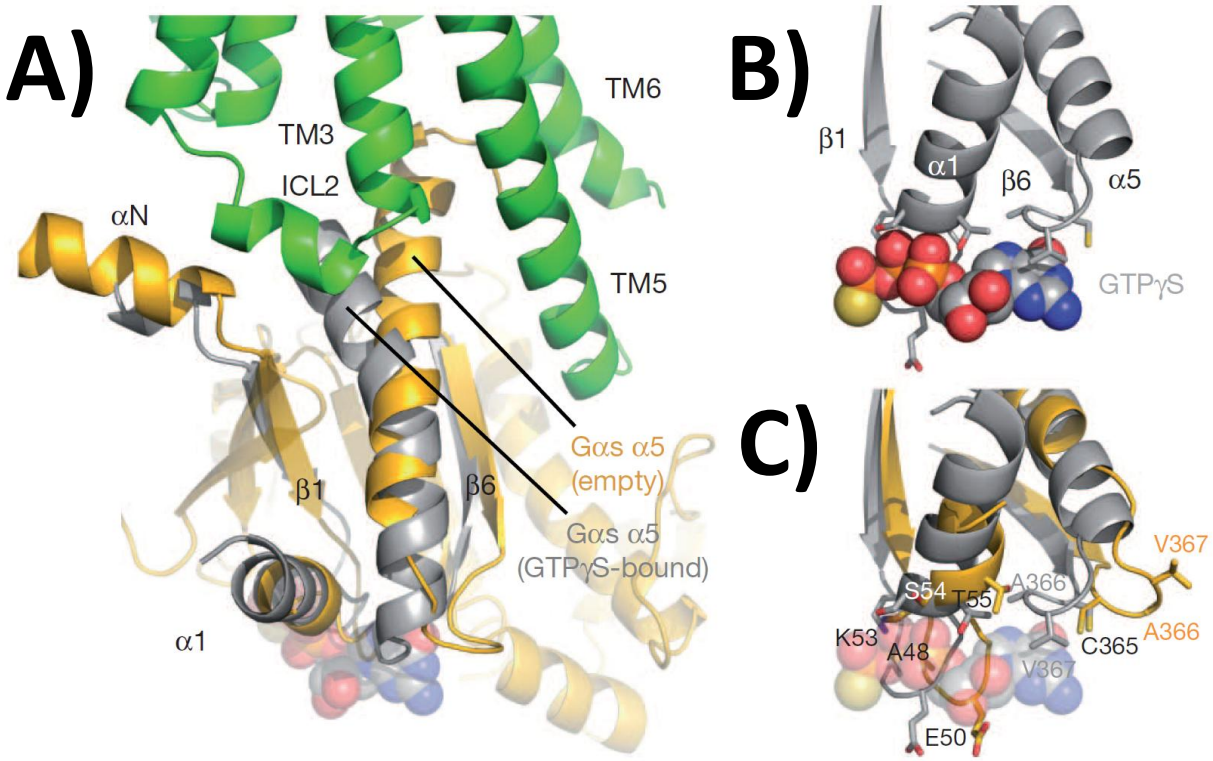


Figure 3-12. Structural re-arrangements of the Gαs C-terminal helix in the β_2AR-Gs complex. A) The Gαs subunit from the β_2AR–Gs complex is drawn in orange, and it is aligned with the GTPγS-bound Gαs drawn in grey,[115] showing overall movement of the helix into the transmembrane core of the receptor. The helix moves by about 6 Å towards the receptor and rotates by about 30°. B) The loop between the C-terminal helix and the β6 strand contains the G-box sequence 5 and forms part of the guanine base-binding portion of the nucleotide binding site. C) Movement of the helix disrupts the structure of loop, which destabilizes nucleotide binding. Figure reprinted from Rasmussen et al.[86]

The movement of the C-terminal helix disrupts the position of the loop that links the beginning of the helix to the β6 strand. When GTP or GDP is bound to the Gα subunit, this loop interacts with the guanine ring of the nucleotide. Residues A366 and V367 provide a hydrophobic surface that interacts with the π-cloud of the ring, and the backbone amide of A366 forms a hydrogen bonding interaction with the guanine nitrogen base carbonyl, shown above in figure 3-12B. Upon movement of the C-terminal helix, these two resides are forced to adopt new positions as shown in figure 3-12C. This is necessary to accommodate the extended position of the

helix, with the α-carbons of the alanine and valine residues moving 7.7 and 6.8 Å away from their nucleotide-binding positions, respectively.

Another change in the Gα conformation is caused by the binding of F139 on the receptor ICL2 to the hydrophobic pocket on the Gα subunit formed by residues from the C-terminal helix and the β1 and β3 strands, shown in detail above in figure 3-11. The binding of this residue pulls the β1, β2, and β3 strands up towards the receptor by 2-3 Å, a motion that is helped along by the rotation of the c-terminal helix, shown below in figure 3-13A.

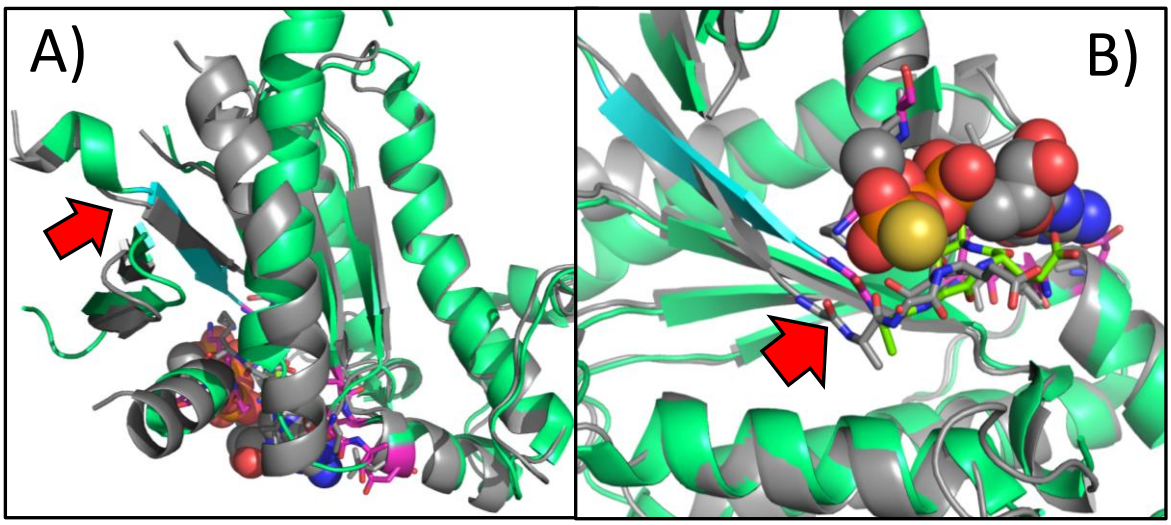


Figure 3-13. Conformational changes involving the β1 strand of the Gαs subunit. Gαs bound to GTPγS is drawn in grey (PDB 1AZT), showing the position of the β1 strand and P-loop of the GTP or GDP bound state. Drawn in color is the nucleotide-free Gαs subunit from the β_2AR-Gs structure (PDB 3SN6). The β1 strand is colored in cyan, the G-box consensus sequences are colored in magenta, and the P-loop residues are colored in bright green. A) The β1 strand and nearby β2 and β3 strands are all moved in towards the core of the fold upon receptor binding. B) The warping of the β1 strand pushes the position of the P-loop about 2.5 Å towards the nucleotide, disrupting binding of the GTP or GDP β-phosphate.

Although the movement of the β1 strand is not large, it is enough to warp the positioning of the β1-α1 loop, shown in figure 3-13B. This loop, also known as the "P-loop," contains the G-box sequence 1 that binds the β-phosphate of the GTP or GDP molecule. The position of the loop is shifted towards the guanine-base end of the binding pocket by about 2.5Å. However, this change in position is probably not as important as the overall flexibility that is imparted on the region by the interaction

with receptor. As discussed below in section 3.8, we also observed a dramatic increase in deuterium exchange of the β1 strand peptide, indicating increased disorder and solvent accessibility. Our interpretation of these data is that the movement of the β1-β3 strands associated with F139 binding from the receptor ICL2 domain is incompatible with the proper positioning of the P-loop relative to the rest of the nucleotide binding site, leading to a situation where the GTP/GDP β-phosphate binding and ICL2 binding to the Gα subunit are mutually exclusive.

The final difference between the nucleotide-free Gαs structure in the β_2AR-Gs complex and the GTPγS bound Gαs is the most dramatic change seen in the whole structure. The α-helical domain of the Gαs is normally positioned next to the Ras-like domain at the site of nucleotide binding, forming a loop that packs around the ribose group of the nucleotide and positioning the catalytic residue R201 next to the β- and γ-phosphates, shown below in figure 3-14A. However, in the complex structure the center of the α-helical domain moves away by 45 Å, rotating by 127° around the hinge region where it is attached to the Ras-like domain, shown below in figure 3-14B.

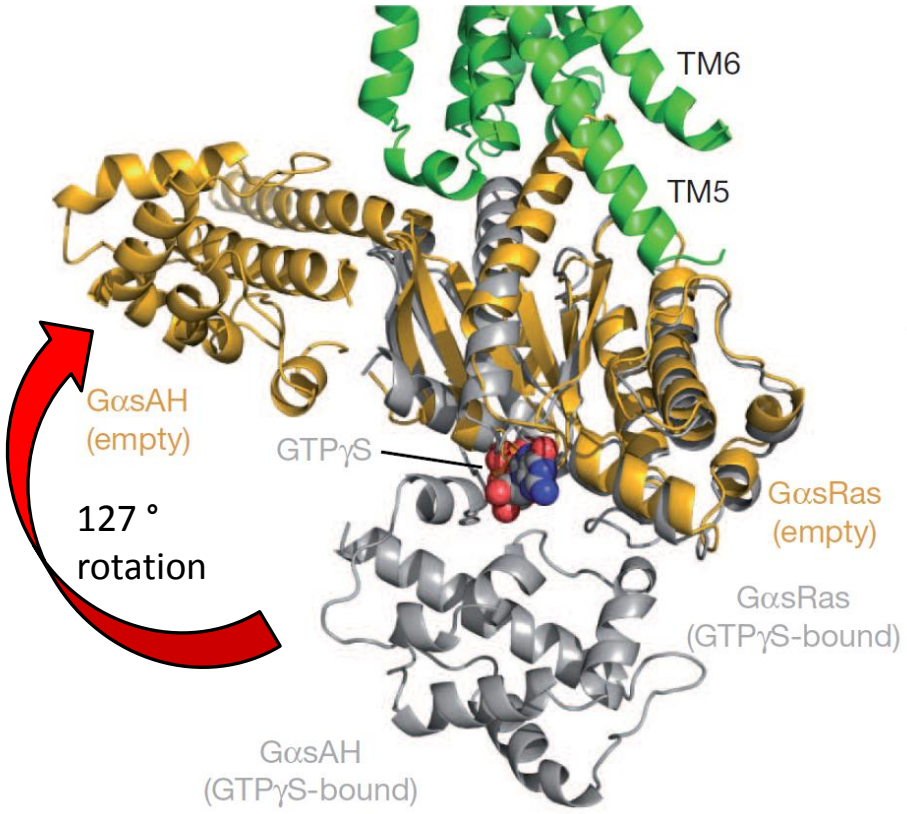


Figure 3-14. Movement of the Gαs α-helical domain in the β_2AR-Gs complex. Gαs bound to GTPγS is drawn in grey (PDB 1AZT), and drawn in yellow is the nucleotide-free Gαs subunit from the β_2AR-Gs structure (PDB 3SN6). When bound to GTP or GDP, the domain packs against the nucleotide's ribose group, forming a portion of the binding site. The position of the α-helical domain in the nucleotide-free Gαs subunit of the β_2AR-Gs complex structure is much different, far removed from the nucleotide binding site. Figure reprinted from Rasmussen et al.[86]

The movement of the Gαs α-helical domain disrupts the nucleotide binding site and contributes to the loss of nucleotide caused by binding to the receptor. Based on electron microscopy work described below in section 3.9, We propose that the observed position of the α-helical domain in the β_2AR-Gs complex crystal is only one of many positions relative to the Ras-like domain that it can take, and that in solution the domain is expected to sample the alternate positions rapidly. This view is supported by recent EPR data taken using transducin and rhodopsin, where the distance distribution between the tranducin α-helical domain and the Ras-like domain broadens out and increases on average when the transducin interacts with activated rhodopsin.[108]

3.8 – Dynamics of the receptor-G protein complex determined by deuterium exchange.

In addition to crystallizing the β_2AR-Gs complex, we also performed studies on the samples using deuterium exchange by mass spectrometry (DXMS). The sample was incubated in deuterated buffer, quenched in acid, digested with an immobilized pepsin column, and finally analyzed by LC-MS/MS to determine the identity of each peptide as well as the number of incorporated deuterium atoms (full methods given in section 6.15). The technique can count the total number of deuteriums on each peptide, but cannot localize them to a specific residue. Very few peptides from the receptor were recoverable in the assay without greatly sacrificing the yield of G protein peptides, so we chose to optimize the assay to obtain good sequence coverage for G protein subunits. For the Gα subunit, 80% of the sequence was measured by at least one peptide, and for the Gβγ subunit, the coverage was 75%. The raw data was converted to % deuteration values based on the theoretical maximum number of exchangeable hydrogen atoms on each peptide. Data for three different time points of deuterium incubation are shown below as a ribbon diagram for the Gαs subunit in figure 3-15 and for the Gβγ subunits in figure 3-16. The DXMS measurements were repeated for samples of heterotrimer alone, the β_2AR-Gs complex, and the complex treated under two conditions that are known to disrupt the receptor-G protein coupling. The addition of GDP plus AlF_3 mimics the binding of GTP and fully dissociates both the G protein from the receptor and the Gα subunit from the Gβγ subunit.[116] The addition of GDP by itself allows the coupling reaction between the receptor and the G protein to run in reverse, yielding free receptor and heterotrimeric G protein.

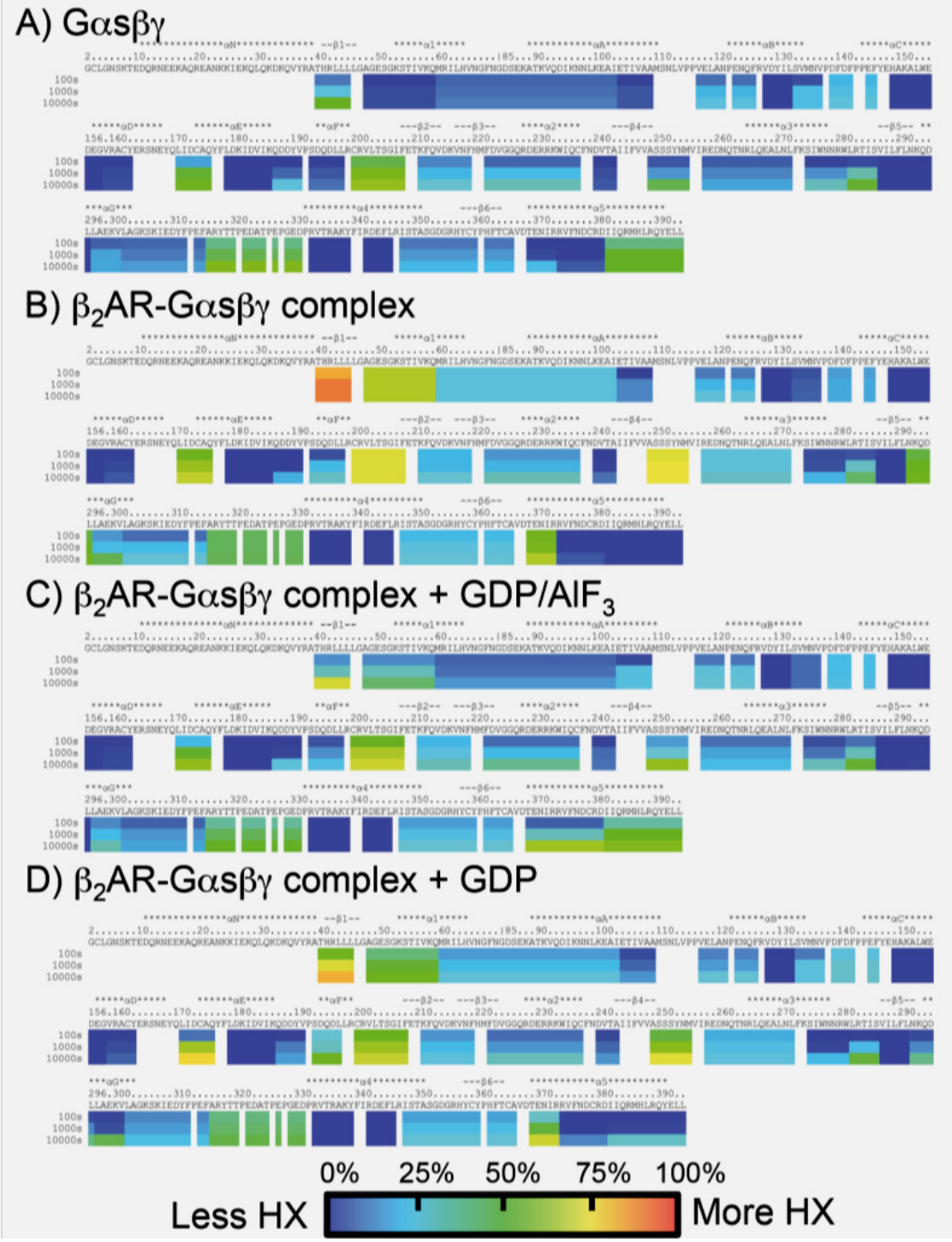


Figure 3-15. Ribbon diagram of the hydrogen-deuterium exchange levels of Gαs. A) Exchange levels in heterotrimer alone. B) Exchange levels in the β_2AR-Gs complex. C) Exchange levels in the β_2AR-Gs complex treated with GDP/AlF_3. D) Exchange levels in the β_2AR-Gs complex treated with GDP. Indicated is the amino-acid sequence and secondary structure. Exchange levels are color-coded according to the indicated heat map. Residues not colored represent fragments where no mass information was obtained. Figure reprinted from Chung et al.[97]

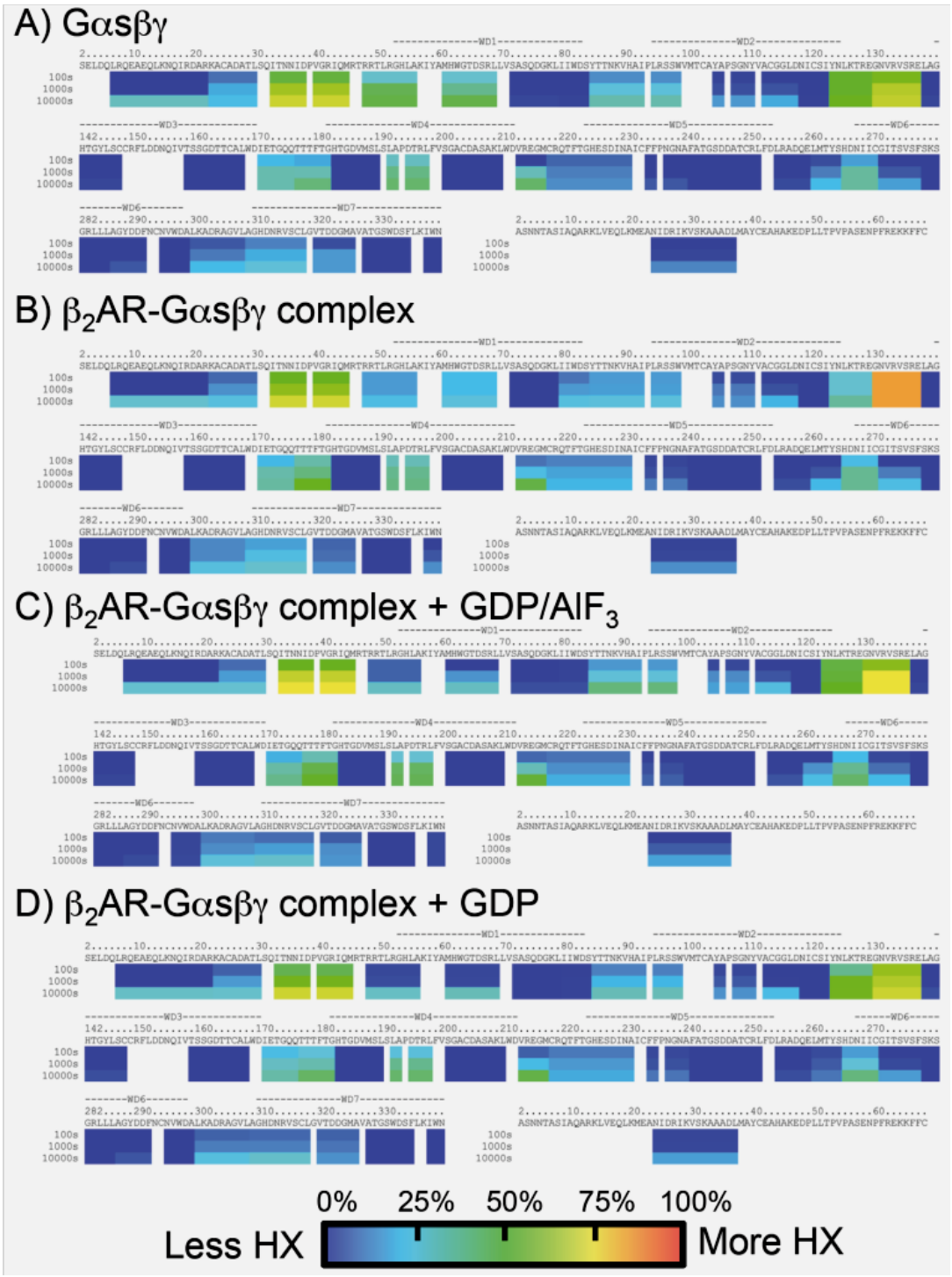


Figure 3-16. Ribbon diagram of the hydrogen-deuterium exchange levels of Gβγ. A) Exchange levels in heterotrimer alone. B) Exchange levels in the β_2AR-Gs complex. C) Exchange levels in the β_2AR-Gs complex treated with GDP/AlF_3. D) Exchange levels in the β_2AR-Gs complex treated with GDP. Indicated is the amino-acid sequence and secondary structure. Exchange levels are color-coded according to the indicated heat map. Residues not colored represent fragments where no mass information was obtained. Figure reprinted from Chung et al.[97]

An overall visual survey of the data shows that there are only a few domains on either of the proteins that exchange over 50% of their available hydrogen atoms in the time frame of the study. These regions include domains that are known to be relatively flexible on the Gα subunit, such as the hinge region between the Ras and α-helical domain and the C-terminus. Interestingly, a large amount of exchange was found on the loop that contains the αG helix of the protein, one of the more variable regions for the different α subunits. On the Gβ subunit, a stretch of residues that form an extended random coil near the C-terminal region of the Gγ subunit are also rather flexible. Most curiously, two loop regions in the Gα α-helical domain and the Gβ subunit, shown below in figure 3-17, also showed completely unexpected high basal levels of exchange and a small increase in exchange when they were bound in the nucleotide-free receptor-G protein complex. The regions are in very close proximity to each other in the complex crystals, although they do not pack against one another. However, the DXMS data was obtained without NB35 bound, so the relative position of the Gα and Gβ domains may be more flexible in these samples compared to the nanobody-bound sample used for crystallography. It is tempting to speculate that the two domains may transiently interact or collide, and this might lead to the elevated deuterium exchange. However, we know of no studies where an interaction like such has been proposed, and so at this point we show these data only as an interesting finding and not as any basis for a particular interpretation of the G protein activation mechanism.

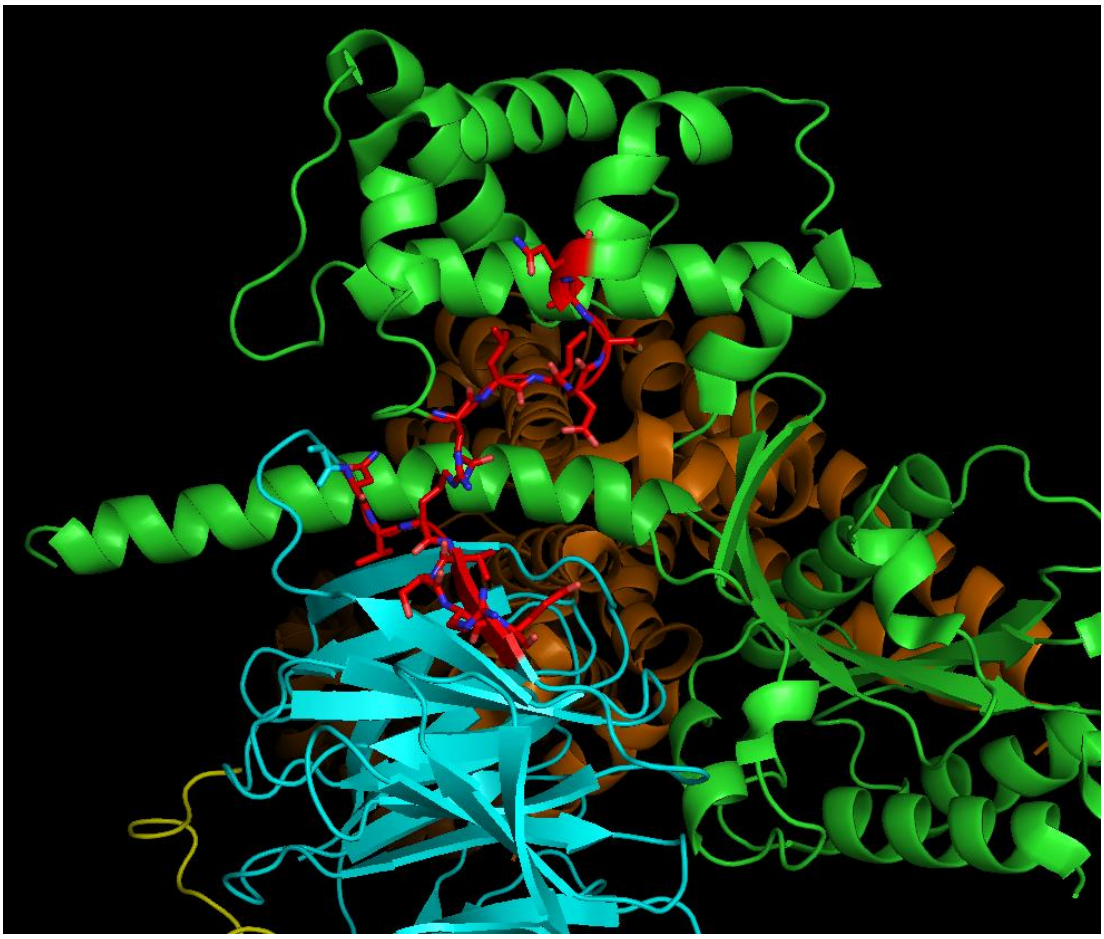


Figure 3-17. Regions on the Gα α-helical domain and the Gβ subunit that show unexpectedly high deuterium exchange values. The regions are drawn in red, with the Gα subunit in green and the Gβ subunit in cyan.

There are also several other regions on the Gα subunit that do not exchange hydrogens with solvent much in a nucleotide-bound conformation, but do so when the protein is nucleotide-free in the β_2AR-Gs complex. They are best seen in figure 3-18 below, where the data from the 100 second timepoint are used to make a heatmap showing the difference in deuterium exchange levels between the Gα subunit bound in the complex and one of the three different nucleotide-bound states.

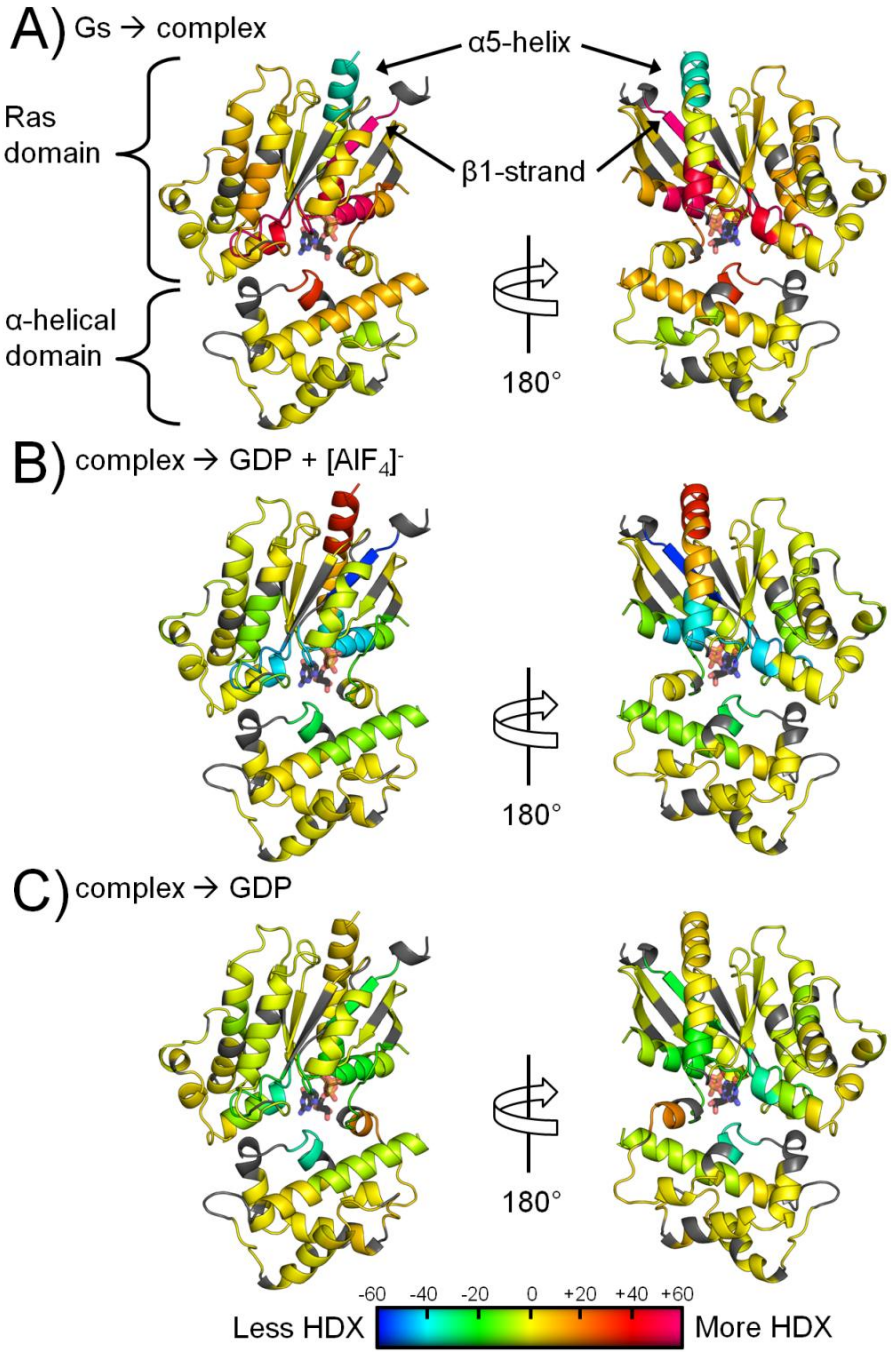


Figure 3-18. Pairwise comparisons of deutrium exchanged levels for Gαs under different conditions. A) Changes in DXMS measurements for Gαs in the β_2AR-Gs complex compared to the Gs heterotrimer. B) Changes in DXMS measurements for Gαs in the β_2AR-Gs complex caused by dissociation with GDP/AlF_3. C) Changes in DXMS measurements for Gαs in the β_2AR-Gs complex caused by dissociation with GDP alone. The changes in hydrogen-deuterium exchange (HDX) given as changes in the percentage of the theoretical maximum number of deuterons incorporated per peptide were mapped on to the crystal structure of Gαs based on the GTPγS bound form (PDB 1AZT) as indicated by included heatmap. Regions where no mass spectrometry data were obtained are indicated in grey. Figure reprinted from Chung et al.[97]

As expected, it can easily be seen in the figure part A that the residues next to the GDP/GTP binding site all increase their deuterium exchange levels when going from GDP bound heterotrimer to the nucleotide-free β_2AR-Gs complex. The opposite is true in figure parts B and C, where the β_2AR-Gs complex is treated with GDP/AlF_3 or GDP and the residues near the binding site all decrease their levels of deuterium exchange. These changes in the deuterium exchange levels of the nucleotide-binding residues are consistent with a large increase in solvent accessibility and flexibility of the residues when the nucleotide is absent from the binding site.

The DXMS data also show that the C-terminal region of the α-subunit is less flexible and solvent exposed in the β_2AR-Gs complex, consistent with the position of the helix in the crystal structure, where is it bound in the transmembrane core of the receptor. When the complex is dissociated with GDP/AlF_3, the opposite is true and the C-terminal residues return to the high basal rate of exchange. However, the addition of GDP alone is not as efficacious at enhancing the exchange rates on the helix, even though it is just as good at stabilizing the residues in the nucleotide binding pocket as GDP/AlF_3 treatment. One possible explanation for this discrepancy is that there might be a transient, GDP-bound conformation of the G protein that can still interact with the receptor through its C-terminus. Several other labs have proposed such a pre-bound conformation,[117, 118] and if it exists, it is likely to enhance response times for GPCR signaling by removing the need for G protein to diffuse to the receptor before it can interact. Nevertheless, it is likely that the GDP-bound β_2AR-Gαs complex is not very stable, because increased exchange at the C-terminal helix could be observed following longer exchange durations or incubations at higher (but still physiological) temperatures.

The most unexpected finding with the DXMS data was that the β1 strand of the Gα subunit, which links the N-terminal helix and the P-loop for the nucleotide-binding site, was found to undergo extremely rapid deuterium exchange in the β_2AR-Gαs complex samples. Shown below in figure 3-19, the position of the β1 strand differs slightly from the GDP bound conformation, but by enough to disrupt the positioning of the attached P-loop.

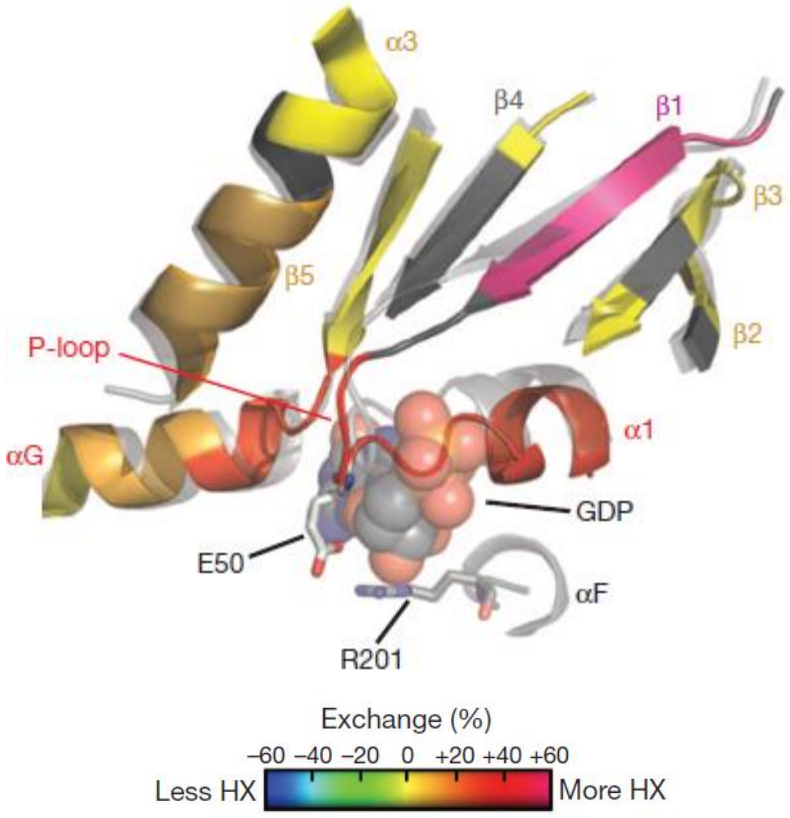


Figure 3-19. Comparison of the structure and dynamics of Gαs in β_2AR-Gαs complex with Gαi bound to GDP. The change in deuterium exchange levels from the Gs heterotrimer to the β_2AR-Gαs complex is indicated according to the indicated heatmap. The structure of Gαi bound to GDP is shown in grey for comparison. Figure reprinted from Chung et al.[97]

Given the extreme change in deuterium exchange levels, it seems likely that in solution the β1 strand is undergoing much more movement than is seen in the crystal structure. The β1 strand residues are very highly conserved in the Gα protein subtypes. For Gαs and Gαolf, the sequence is RLLLL, but in most receptors the arginine is substituted for a lysine residue. Evidently, this domain has evolved to be an important conformational link between the areas were the Gα subunit interacts with receptor and the phosphate binding loop of the nucleotide binding site. In this case, the DXMS data was crucial to determining this aspect of GPCR-mediated nucleotide exchange on the G protein, as the changes in the crystal structure were subtle and one would usually assume that a β-sheet would not normally be associated with such high flexibility.

3.9 – Analysis of the receptor-G protein complex by electron microscopy.

The β_2AR-Gs complex samples for crystallography were also studied by electron microscopy. For these studies, the sample was bound to a formvar-coated grid and negative stained with uranyl formate (full details given in section 6.16). The samples were all solubilized in MNG-3 detergent, which stays bound to the protein-detergent complex for very long periods of time after the detergent is diluted underneath its critical micelle concentration (cmc). Because of this, we can see significant density that corresponds to the MNG-3 micelle surrounding the receptor, making it look larger than its molecular weight would suggest. Unlike the receptor-bound micelles, the free micelles of MNG-3 rapidly dissociate when diluted underneath the cmc, which helped greatly to reduce the number of non-target particles in the EM images.

Imaging of the complex shows a low-density electron map that agrees in general with the crystal structure of the complex. Shown below in figure 3-20, the 2D reprojections of the crystal structure appear very similar to the experimental EM data, with the exception that the EM data includes a large micelle of MNG-3 bond to the receptor. From the analysis of the class averages, we identified two major and distinct conformational subpopulations of the complex. One conformation, which accounted for about a third of the analyzed particles, includes clear density for the Gαs α-helical domain in an orientation similar to its position in GDP or GTP bound Gα subunits. For the other two-thirds of the particles, the density corresponding to the α-helical domain was highly delocalized and it was averaged out in the maps.

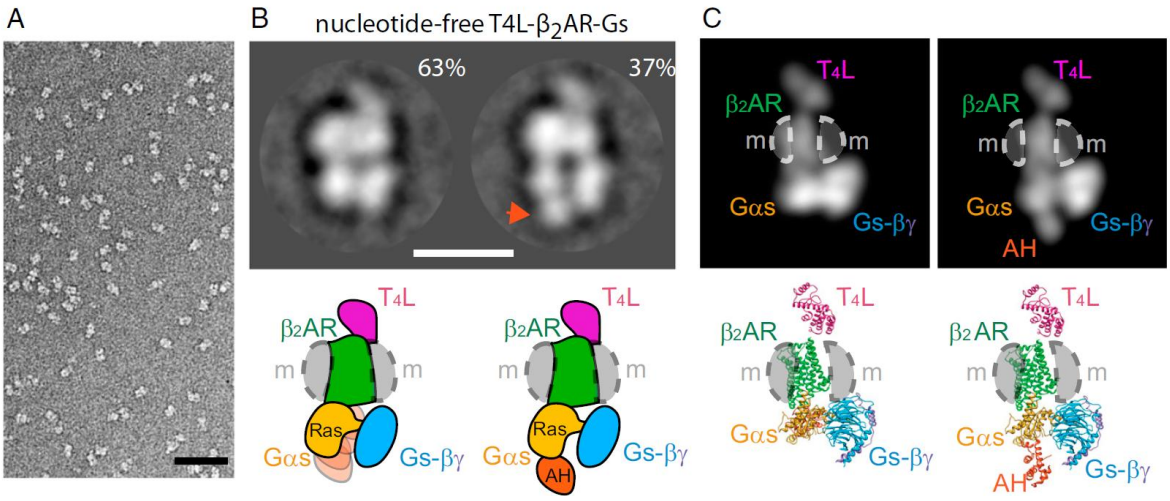


Figure 3-20. Two-dimensional projection analysis of the T4L-β_2AR-Gs complex in the nucleotide-free state. A) Raw EM image of detergent-solubilized T4L-β_2AR-Gs complex embedded in negative stain (scale bar, 50 nm). B) Representative EM class averages of the nucleotide-free complex with the projection profile of the α-helical (AH) domain not visible (left), or visible on the Ras-like domain (right, AH indicated by arrow). Cartoon models that represent the conformations reflected by the EM averages are shown below the averaged images (scale bar, 10 nm). C) Reprojections (upper section) of the β_2AR-Gs crystal structure (lower section) in the same overall orientation as figure part B reveal the identity of each EM density component. The crystal structure on the right shows the AH domain in the same position relative to the Ras-like domain as determined in the crystal structure of Gαs bound to GTPγS alone. Figure reprinted from Westfield et al.[88]

The data were also used to construct a 3D map of the electron density of the two classes of particles. The maps are shown below in figure 3-21, with the ribbon representation of the constituent proteins superimposed in the density maps. Again, the maps are in good agreement with the crystal structure of the β_2AR-Gs complex, with the exception of the α-helical domain of the Gαs protein.

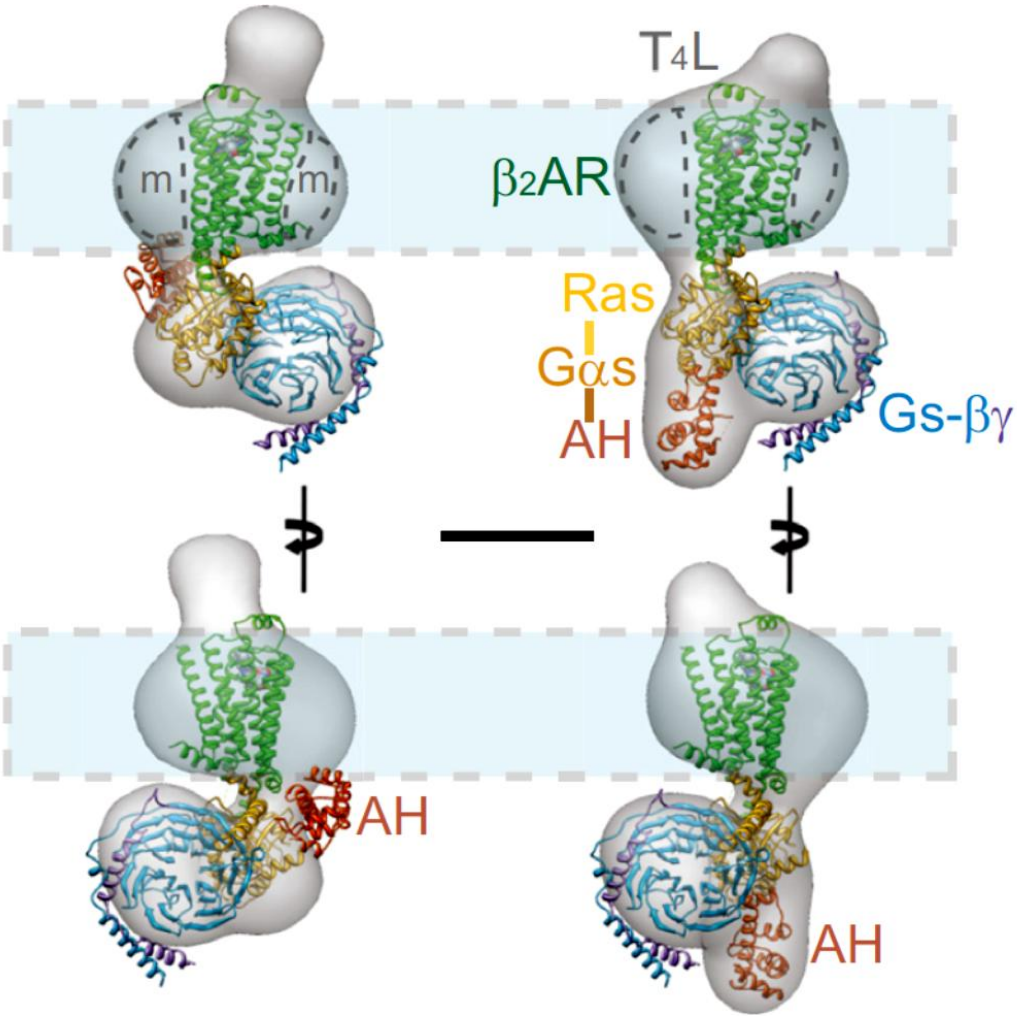


Figure 3-21. Three-dimensional reconstructions of the T4L-β_2AR-Gs complex in the nucleotide-free state. In the reconstruction to the left, the α-helical (AH) domain (orange ribbon) is shown in the same position as found in the docked crystal structure. Absence of sufficient density to accommodate this domain indicates that its position is highly variable in this particle population. In the reconstruction to the right, the AH domain is modeled within the available EM density right below the Ras-like domain of Gαs, as also suggested by the 2D averages. Figure reprinted from Westfield et al.[88]

In order to further determine what conformations the α-helical domain was taking in the two-thirds of the β_2AR-Gs complex particles where it could not be localized, we incubated the samples with a nanobody that bound to the α-helical domain and increased its molecular weight. This nanobody, designated as nanobody 37 (NB37), was found to bind to the complex in the same screen that found NB35, but did not end up aiding the quality of the crystals significantly. However, it did prove to be useful to visualize the position of the α-helical domain, shown in figure 3-22 below.

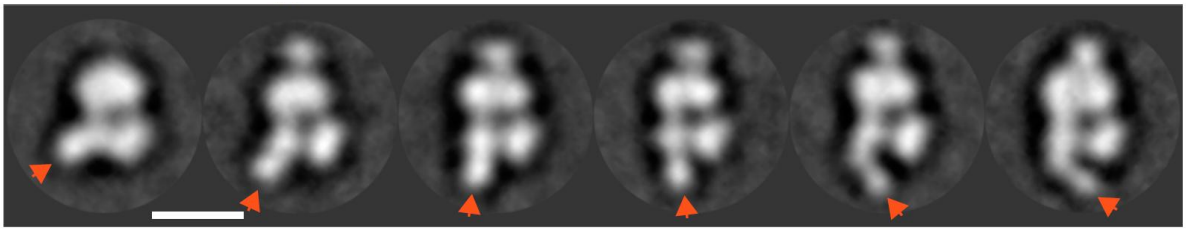


Figure 3-22. 2D class averages of the T4L-β_2AR-Gs complex bound to NB37. NB37 binds to the α-helical domain, indicated by an arrow, and allows identification and classification of several different positions of the domain. Figure reprinted from Westfield et al.[88]

The relative location of the domain is variable, and can extend all the way from a position close to where it was found in the crystal structure to the position it adopts when a nucleotide is bound. We used these class averages to reconstruct a 3D map of the electron density for each of the variable α-helical domain positions. The maps are shown below in figure 3-23.

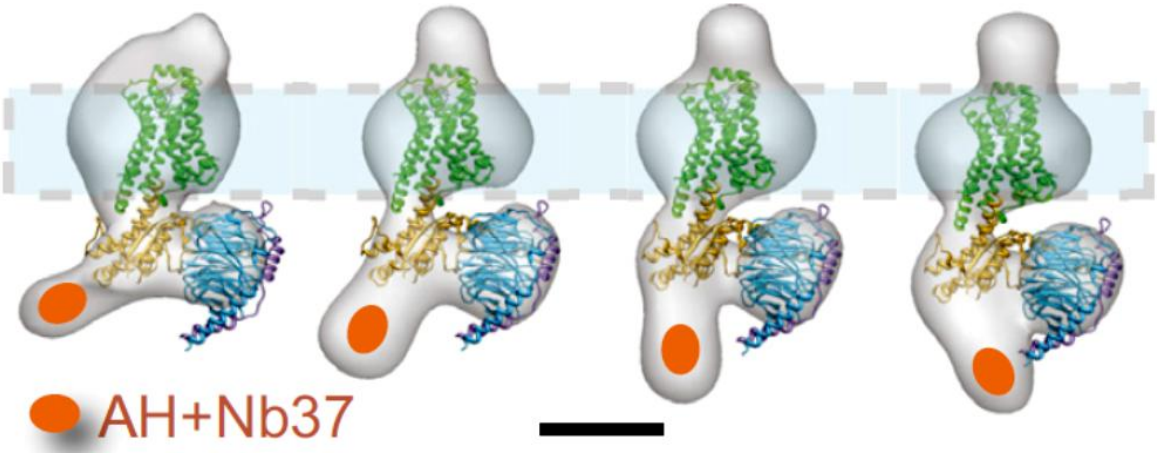


Figure 3-23. Three-dimensional reconstructions of the T4L-β_2AR-Gs complex with NB37 bound to the α-helical (AH) domain. The Nb37-enhanced density of the AH domain (marked with an oval) shows variable positioning around the Ras-like domain of Gαs (scale bar, 5 nm). Figure reprinted from Westfield et al.[88]

With the aid of NB37 binding, the α-helical domain can be localized in the many different position shown. It is likely that the domain is actually fully flexible and can be found in all the positions along the trajectory of movement, but these particular structures are averages of similar conformations that were binned together because they were indistinguishable within the resolution limits of the technique.

We also wanted to determine what factors stabilized the α-helical domain in its "normal" position packed against the Ras-like domain. To assess this, we subjected

a sample of the complex to several different treatments and measured how many of the imaged particles contained a stabilized α-helical domain, shown below in figure 3-24.

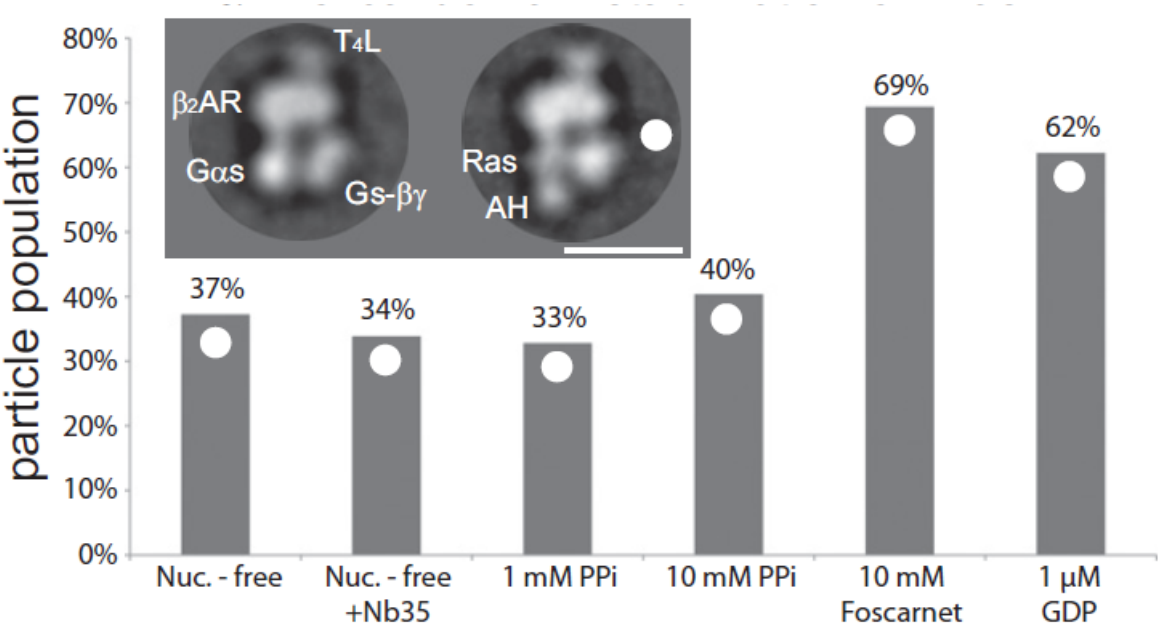


Figure 3-24. Nucleotide-dependent positioning of the Gαs AH domain. The distribution of particles with a distinct projection profile of the α-helical (AH) domain stabilized on the Ras-like domain across different conditions is quantified (inset right, marked with a white dot. A class average of a particle with a nonvisible AH domain is shown for comparison, inset left). The presence of foscarnet and GDP significantly increases the number of particles with stabilized AH domain. Figure reprinted from Westfield et al.[88]

Treatment with either NB35 or pyrophosphate did not seem to alter the distribution of α-helical domain positions significantly, but the addition of Foscarnet or GDP showed a clear increase in the number of particles with stabilized α-helical domains. GDP treatment caused partial dissociation of the complex, but enough particles were still intact to obtain the data shown. It is not surprising that GDP causes the α-helical domain to close, as the domain is normally positioned in that orientation when nucleotide is bound to the G protein. However, the most interesting finding is that Foscarnet also stabilized the domain. Foscarnet, or phosphonoformic acid, is a pyrophosphate mimic that is assumed to bind where the β-phosphate of GDP or GTP binds the P-loop of the Gαs subunit. We propose that the changes in the domain position distribution due to the drug is because it is stabilizing the phosphate binding part of the nucleotide-binding site without also requiring the guanine base-

binding residues of the site to adopt their GDP or GTP bound position. This allows partial ordering of the nucleotide binding site without also requiring the Gαs C-terminal helix to adopt its GDP/GTP bound position, thus allowing the extensive interaction between the receptor and the C-terminus to remain intact. However, it is unclear why pyrophosphate did not also show a similar effect the in experiment.

We also imaged the complex after treatment with GTPγS, shown below in figure 3-25. As expected, the complex rapidly dissociates. However, if the sample preparation was done quickly, the particles could be visualized in the various stages of disassembly.

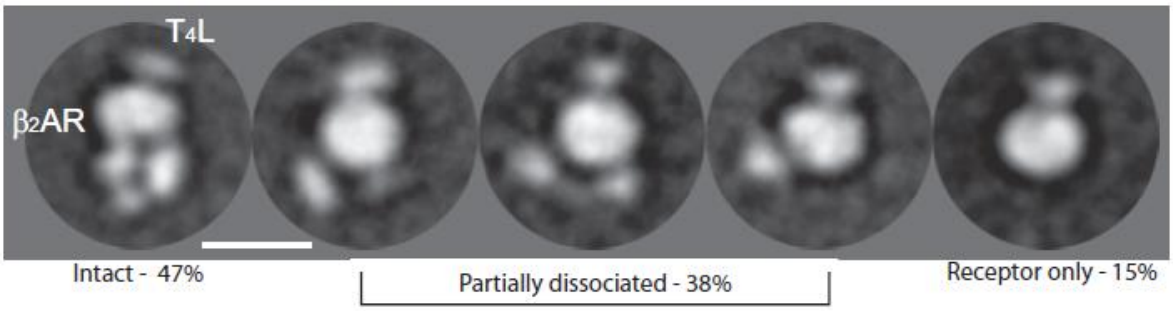


Figure 3-25. Dissociation of the β_2AR-Gs complex with GTPγS. Representative class averages of the T4L-β_2AR-Gs complex after rapid mixing with GTPγS (1 μM) and immediate stain embedding reveal both intact as well as partially or fully dissociated complexes (scale bars, 10 nm). Figure reprinted from Westfield et al.[88]

Although this technique is not appropriate for detailed study of the dissociation pathway due to the difficulty of accurately identifying which particles belong to the various stages of dissociation, it does give us a general idea about the process. In particular, it can be seen that the Gαs and the Gβγ subunits dissociate from one another before the entire complex dissociates from the micelle. Unfortunately, it is not clear if the Gβγ subunit dissociates before the Gα-receptor interaction is broken, or if the entire G protein dissociates from the receptor before subsequently breaking apart as well. Such questions are likely best addressed by other techniques, such as kinetic fluorescence energy transfer methods or EPR-based studies with labeled G protein subunits.

3.10 – Section summary.

The data presented in the sections above give us a detailed view of how the activation signal carried by a GPCR agonist is transferred across the plasma membrane and causes the first step of the G protein activation reaction. As one of the main players of a large collaborative effort, my research has given the field much new information about the structure and dynamics of GPCR and G protein activation. We have shown the molecular interface between receptor and G protein, detailed the structural changes that happen during receptor-G protein interaction in atomic detail, and determined regions of the proteins that undergo changes in flexibility and position while the complex is in solution, free from artifacts of crystallization. The structural changes that we observed involved some domains of the proteins that have long been known to be involved in the interaction, but we also showed several previously unknown aspects of the activation process. Overall, this work is a major addition to the understanding of how GPCRs and G proteins function and should greatly aid future work on these important classes of cell signaling proteins.

CHAPTER 4
PHARMACOLOGY OF A RECEPTOR-G PROTEIN COMPLEX

4.1 – Introduction.

This section explores the nature of the activation signal that is passed during the interaction between receptor and G protein. We propose that the binding of the two proteins to one another in the manner described in chapters 2 and 3 is, in itself, this activation signal. This would imply that there is a negatively cooperative interaction between receptor ligand binding site and G protein nucleotide binding site in the receptor-G protein complex. If this is true, then not only would the G protein be forced to lose its bound GDP during interaction with the activated receptor, but the binding of a nucleotide-free G protein must also induce the receptor itself to adopt an activated conformation. As discussed in the main introduction, this relationship is classically seen as the appearance of a second, higher affinity binding site for agonists to receptors when their cognate G proteins are present in the samples. This capacity of G proteins to allosterically modulate agonist binding serves as the basis for the “ternary complex model” (TCM) between receptor, G protein, and agonist.[119]

Shown in the following experiments, we find that formation of the G protein-receptor complex stabilizes a conformation of the receptor that restricts both access to and dissociation from the ligand binding site of the receptor for an antagonist, which we identify as the activated receptor conformation. We also show that GDP binding to the complex reverses this unique pharmacological signature of the activated receptor. Moreover, we show that a G-protein mimic, NB80, is able to force the receptor into this activated conformation and can use it to make significant progress towards finding a complete kinetic solution of how the activated receptor interacts

with ligands. These findings support the basic idea behind the TCM by linking the new structural information about receptor and G protein activation to both well known and underappreciated aspects of the receptor's pharmacology.

Most of the work shown in this chapter was produced by me, again with the exception that the receptor used throughout was made by Søren G. F. Rasmussen. Any other exceptions are noted in the text, including major contributions from Gisselle Vélez-Ruiz.

4.2 – Nucleotide sensitivity of antagonist binding the β2AR-Gs complex.

During the process of figuring out what we needed to do to get a sample of β_2AR-Gs complex, described in chapter 2 above, we used the neutral antagonist [^{3}H]-dihydroalprenolol ([^{3}H]DHAP) to label our receptors. Initially, we expected the antagonist to bind fully to all samples no matter if they were complexed with the G protein or not, as it should not matter if the receptor is stabilized in the active conformation since antagonists should bind both active and inactive receptor states equally well. It became apparent, however, that this assumption was incorrect and that we were getting reduced levels of binding to the G protein-complexed receptor in our binding assays. To determine what was happing, a series of experiments were performed by Gisselle Vélez-Ruiz where a monomeric β_2AR preparation in rHDL discs experiences G protein coupling and removal of free GDP through the addition of apyrase (full methods given is section 6.18), followed by re-addition of high concentrations of GDP. The capacity of [^{3}H]DHAP to bind to the receptor was assessed at each step and is illustrated in figure 4-1, shown below.

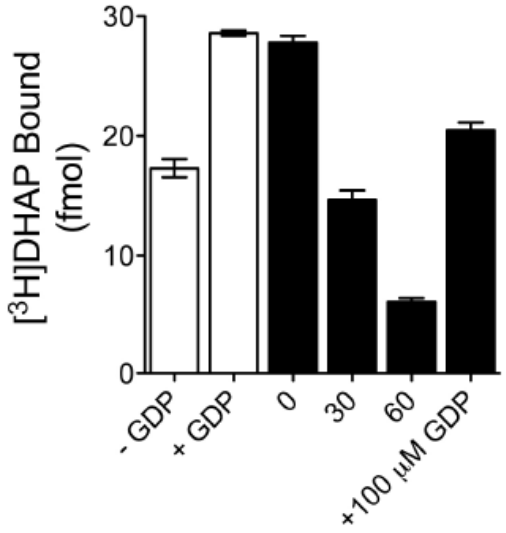


Figure 4-1. Binding of 2 nM [^{3}H]DHAP to Gs-β_2AR-rHDL samples. The samples were incubated with 100 nM GDP and then treated with apyrase for the indicated times. Following apyrase treatment, 100 µM GDP was added back to the samples. Data obtained by Gisselle Vélez-Ruiz with proteins purified by Søren G. F. Rasmussen and Brian DeVree.

Counter to our initial expectations, removal of free GDP by apyrase decreases [^{3}H]DHAP binding in a time-dependent manner. The release of nucleotide appears to be rate-limiting since apyrase is capable of completely hydrolyzing 100 µM GDP within 10 min under identical assay conditions, as determined by anion-exchange chromatography of the nucleotides (data not shown). Since no exogenous GDP is initially added, the GDP concentration before apyrase addition is equal to the concentration of G protein added to the mixture (~1-3 nM). If GDP is resupplied at high concentration (100 µM), the [^{3}H]DHAP binding is partially restored, albeit not completely since apyrase will likely continue to degrade GDP during the entirety of the [^{3}H]DHAP binding assay. To explore this effect further, Gisselle measured saturation isotherms (method given in section 6.22) on apyrase-treated Gs-β_2AR complex samples that had been purified away from the remaining apyrase enzyme, shown below in figure 4-2A.

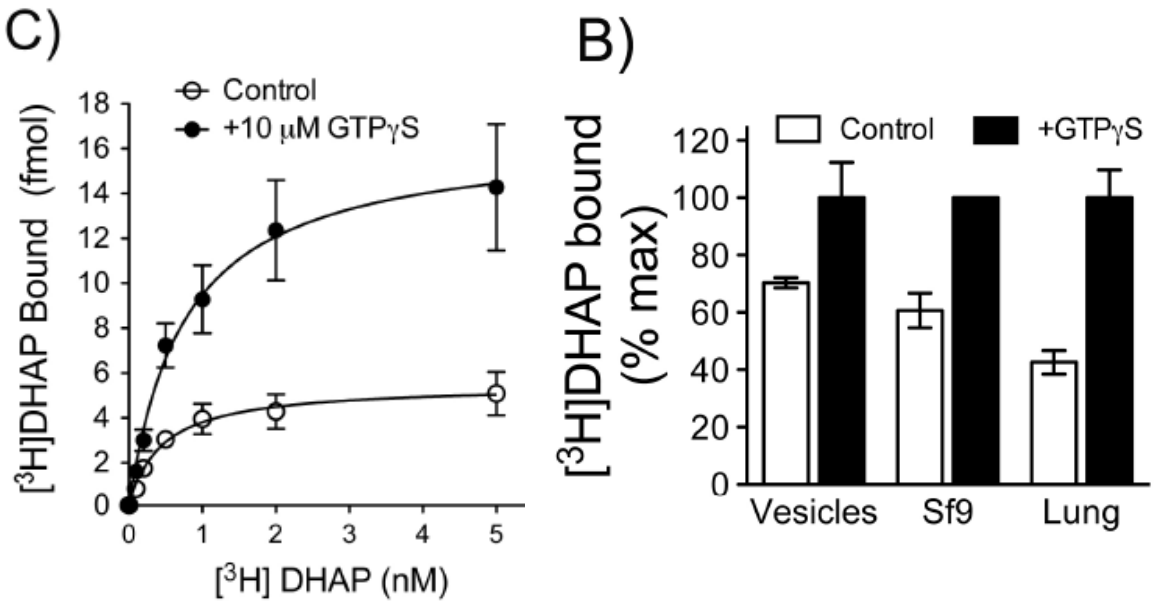


Figure 4-2. The effect of apyrase treatment on [^{3}H]DHAP binding to Gs and β_2AR containing samples. A) Saturation binding analysis of [^{3}H]DHAP to apyrase-treated, Gs-β_2AR complexes in rHDL in the presence or absence of GTPγS. Data obtained by Gisselle Vélez-Ruiz with proteins purifited by Søren G. F. Rasmussen and Brian DeVree. B) Binding of 2 nM [^{3}H]DHAP to β_2AR and Gs reconstituted in vesicles, in membranes prepared from Sf9 cells expressing β_2AR and Gs, and in native membranes prepared from rat lung. Data obtained by Gisselle Vélez-Ruiz.

The saturation curves suggest that [^{3}H]DHAP binds with comparable affinities in the absence or presence of 10 μM GTPγS: 0.41 nM and 0.66 nM, respectively. However, uncoupling Gs from the β_2AR with GTPγS results in a dramatic increase in the maximal number of [^{3}H]DHAP binding sites (Bmax ≈ 16.6 fmol), compared to control (Bmax ≈ 5.5 fmol). A similar phenomenon can be seen in vesicle reconstitutions with model bilayers (prepared according to method 6.19), membranes prepared from β_2AR and Gs-expressing Sf9 cells (prepared according to method 6.20), and even in native murine lung membranes (prepared according to method Sf9 cells growing in mid-log phase were infected at a multiplicity of infection of 0.5-1 with recombinant baculoviruses containing expression constructs for FLAG-His_{10}-mECFP-β_2AR (CBAR) and Gs subunits. 48–60 hours later, the cells were pelleted by centrifugation for 10 min at 500 × g. They were resuspended in 1/10 the original culture volume of TBS buffer (25 mM Tris-Cl pH 7.4, 137 mM NaCl, and 3 mM KCl) + 1× PTT (35 μg ml^{-1} phenylmethanesulphonyl fluoride, 32 μg ml^{-1} tosyl

phenylalanyl chloromethyl ketone, 32 μg ml^{-1} tosyl lysyl chloromethyl ketone) and 1× LS (3.2 μg ml^{-1} leupeptin and 3.2 μg ml^{-1} soybean trypsin inhibitor). They were lysed by sonication and centrifuged at 500 × g for 10 min to pellet unlysed cells and nuclei. The supernatant was centrifuged at 35 min at 100,000 × g to pellet the membranes. The membranes were resuspended in 1/20 the original culture volume with low salt buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1× LS, 1× PTT), flash frozen with liquid nitrogen, and stored at −80°C until use.

6.21), shown above in figure 4-2B. The figure shows strong GTPγS-dependent increases in the number of [^{3}H]DHAP binding sites, which provides support that the effect is not an artifact of the rHDL reconstitution but instead an intrinsic property of the β_2AR and Gs proteins under nucleotide-depleted conditions.

It is also well established that high-affinity agonist binding may be disrupted by guanine nucleotides binding to Gα subunits, and the triphosphate form is known to be more effective. Shown below in figure 4-3, we also found that GTPγS uncouples the nucleotide-free form of Gs from the receptor with lower apparent effective concentration than GDP (EC_{50} = 5.0 and 39 nM, respectively), as indicated by the capacity of the guanine nucleotides to increase [^{3}H]DHAP binding. These values are slightly lower than previously reported ones for a somewhat similar experiment, where GTP and GDP were found to cause a β_2AR-Gαs (the short isoform) fusion protein to uncouple with an EC_{50} of 50 and 80 nM, respectively.[120] Although it is hard to directly compare the two experiments due to the different protein constructs used and the use of GTP instead of GTPγS, it seems clear that guanine nucleotides can start affecting nucleotide-free Gαs subunits in the mid to low nanomolar range of concentrations, and that GTP or GTP mimics are more potent than GDP.

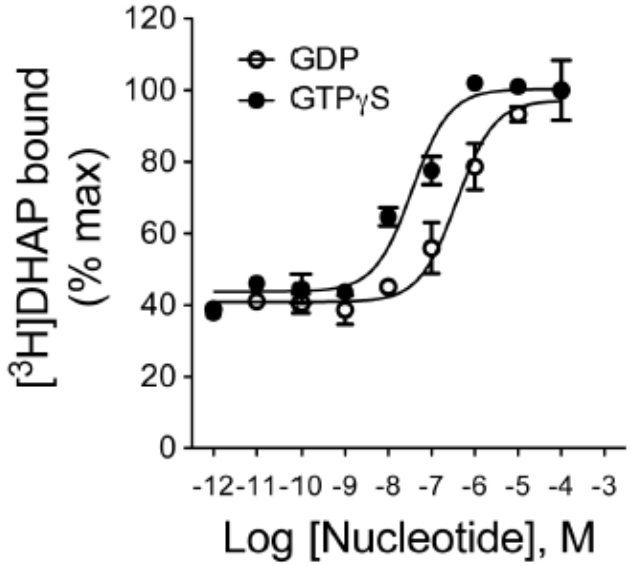


Figure 4-3. Concentration-dependent effect of nucleotides on 2 nM [^{3}H]DHAP binding to nucleotide-free β_2AR-Gs in rHDL particles. Samples were apyrase-treated and purified by gel filtration. Data obtained by Gisselle Vélez-Ruiz with proteins purified by Søren G. F. Rasmussen and Brian DeVree

4.3 – Reduced antagonist binding the β_2AR-Gs complex is a kinetic phenomenon.

In principle, there are two ways that Gs might be able to affect the binding of ligands at the β_2AR orthosteric binding site. One way is that the binding of the G protein stabilizes a conformation of the receptor that has a lower affinity for the ligand, in which case the ligand should act as an inverse agonist with regards to G protein activation. However, a large body of research shows that DHAP and its parent compound, alprenolol, are neutral antagonists[121] or possibly weak partial agonists in some tissue systems,[122] so such an explanation is unlikely to be the cause of our observed reduction in the number of [^{3}H]DHAP binding sites. Instead, we investigated the second possibility, which is that the binding of the G protein does not cause any change in affinity for the ligand, but instead slows the kinetics of drug binding to timescales longer than the length of a typical binding assay itself.

In order to test for kinetic phenomena as the basis of our observations of reduced [^{3}H]DHAP binding, we used the nanobody NB80 which we had previously employed for crystallography purposes. NB80 behaves like a G protein mimic for the receptor and supports high affinity agonist binding.[87] However, it is small, soluble, and has no

nucleotide binding activity, which makes it better for studying receptor activation than the much more complex G protein heterotrimer. Here, we use NB80 as a G protein surrogate to test whether forcing the receptor into the NB80-bound conformational state, with is very similar to the agonist bound, activated, and G protein bound state, also changes the kinetics of [^{3}H]DHAP binding. Shown below, figure 4-4 and 4-5 are the association and disassociation, respectively, of [^{3}H]DHAP from the β_2AR in the presence of increasing amounts of NB80 (full methods given in section 6.24).

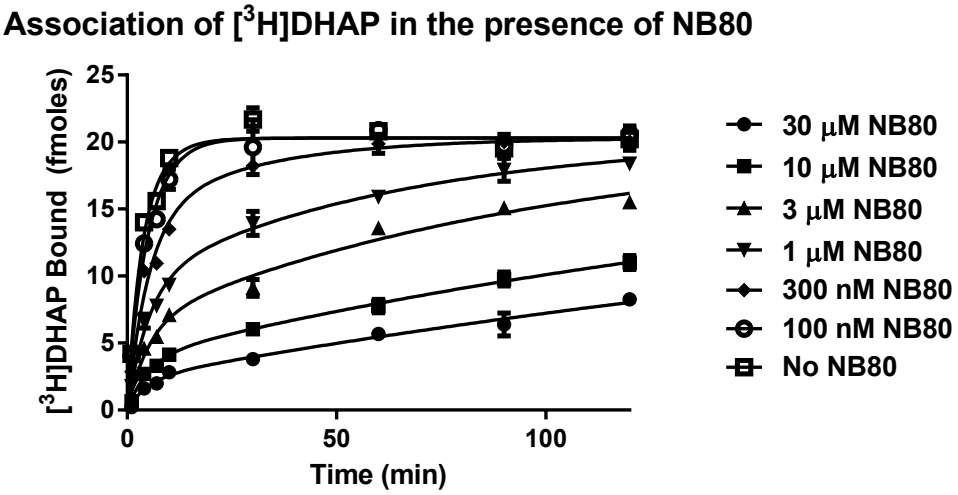


Figure 4-4. Association of [3H]DHAP to β2AR-rHDL the presence of varying NB80 concentrations. 5 nM [^{3}H]DHAP was used and the data was fit to a biphasic association curve, given below in table 4-1. The data points shown are the averages of 3 replicates with the error bars indicating the SEM.

	30 μM NB80	10 μM NB80	3 μM NB80	1 μM NB80	300 nM NB80	100 nM NB80	No NB80
PercentFast	11.12	18.74	31.25	46.86	70.78	93.74	56.21
k_{fast}	0.1924	0.1924	0.1924	0.1924	0.1924	0.1924	0.1924
k_{slow}	0.003179	0.004766	0.01012	0.01564	0.03447	1.006	0.3699

Table 4-1. Rate constants and percent fast phase fitted values for figure 4-4. Rates constants are given in fM^{-1} min^{-1}.

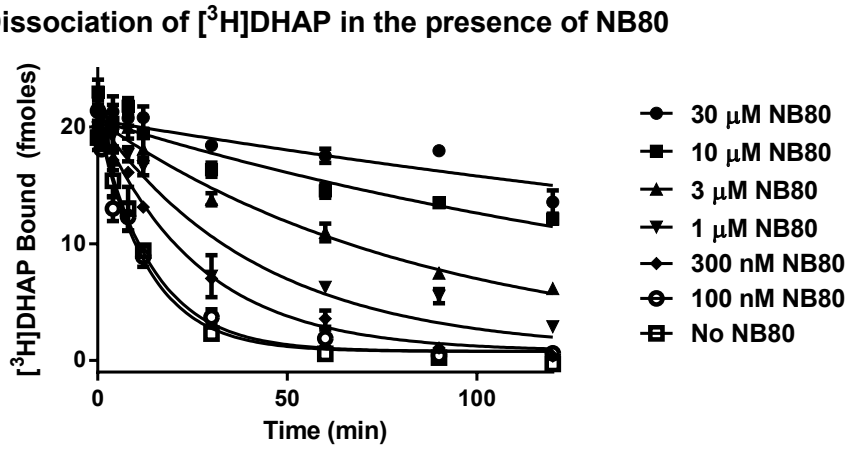


Figure 4-5. Dissociation of [3H]DHAP to β2AR-rHDL the presence of varying NB80 concentrations. 5 nM [^{3}H]DHAP was used and the data was fit to a monophasic dissociation curve, given below in table 4-2. The data points shown are the averages of 3 replicates with the error bars indicating the SEM.

	30 μM NB80	10 μM NB80	3 μM NB80	1 μM NB80	300 nM NB80	100 nM NB80	No NB80
k	0.002768	0.005145	0.01157	0.02325	0.03756	0.07674	0.07048

Table 4-2. Rate constants and percent fast phase fitted values for figure 4-5. Rate constants are given in min^{-1}.

In the figure, the data are fit to a two phase association model and a single phase dissociation model. The fast phase of the association curves was constrained to be the same in all curves and is the rate of [^{3}H]DHAP association to receptors that are unbound to NB80. At high concentrations of NB80, the kinetics of drug binding and release slows dramatically, suggesting that the activated conformation of the receptor does not allow drugs to enter or exit the orthosteric binding site. The effect is large. The sample with the highest nanobody concentration tested, 30 μM, would take about 14.5 hours to reach 95% of its equilibrium binding value, as compared to the sample without nanobody, which does so in only 10.5 minutes, or less than 1/80th of the time.

Although the crystal structures of NB80 and Gs bound β_2AR both show very similar conformations, we wanted to be sure that the kinetic effect on ligand binding we obtained using NB80 was truly representative of the behavior of the Gs-β_2AR complex. To do so, we prepared a sample of pure, nucleotide-free, and ligand-free

Gs-β_2AR complex in rHDL and tested the association kinetics of [^{3}H]DHAP binding after the addition of guanine nucleotides, shown in figure 4-6 below (full methods given in sections 6.18 and 6.25).

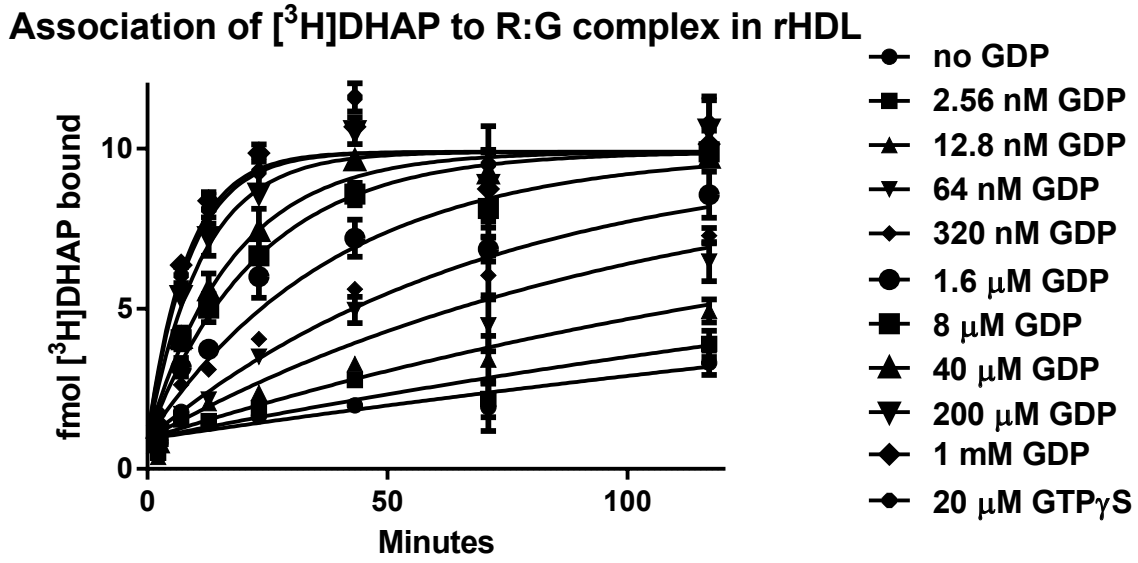


Figure 4-6. Association of 5 nM [^{3}H]DHAP to β_2AR-Gs complex in the presence of varying guanine nucleotide concentrations. Fitted rate constants are given below in table 4-3. The data points shown are the averages of 3 replicates with the error bars indicating the SEM.

	no GDP	2.56 nM GDP	12.8 nM GDP	64 nM GDP	320 nM GDP	1.6 μM GDP	8 μM GDP	40 μM GDP	200 μM GDP	1 mM GDP	20 μM GTPγS
k	0.002475	0.003366	0.005398	0.009404	0.01413	0.02551	0.04368	0.05603	0.08790	0.1208	0.1136
Half-time	280.1	205.9	128.4	73.70	49.07	27.17	15.87	12.37	7.886	5.738	6.102

Table 4-3. Rate constants and percent fast phase fitted values for figure 4-6. Rate constants are given in min^{-1}.

It is clear that for [^{3}H]DHAP association, the same general trend is evident for both G protein and NB80 binding. When either protein is bound to the receptor, it forces it into the activated conformation, which greatly restricts diffusion of the into the orthosteric binding site. However, when GDP or GTP is present, the G protein is able to dissociate from the receptor, leading to quick ligand exchange as the receptor has access to its full ensemble of conformational states.

4.4 – Cooperativity of nanobody 80 and agonist binding.

We also measured the extent of cooperativity between nanobody 80 and agonist binding using competition binding with [^{3}H]DHAP as a radiolabel. An assay was performed where the full agonist (-)-epinephrine (EPI) added at increasing

concentrations to samples of 15 fmol β_2AR in rHDL in the presence of 2 nM [^{3}H]DHAP and varying concentrations of NB80 (full methods given in section 6.23). This experiment was incubated for an extended amount of time (6 hours) in order to minimize the effect of the slow drug binding kinetics due to NB80 binding, described in section 4.3 above. The data are graphed below in figure 4-7. As expected, the presence of the NB80 dramatically increases the affinity of epinephrine by over 2 orders of magnitude.

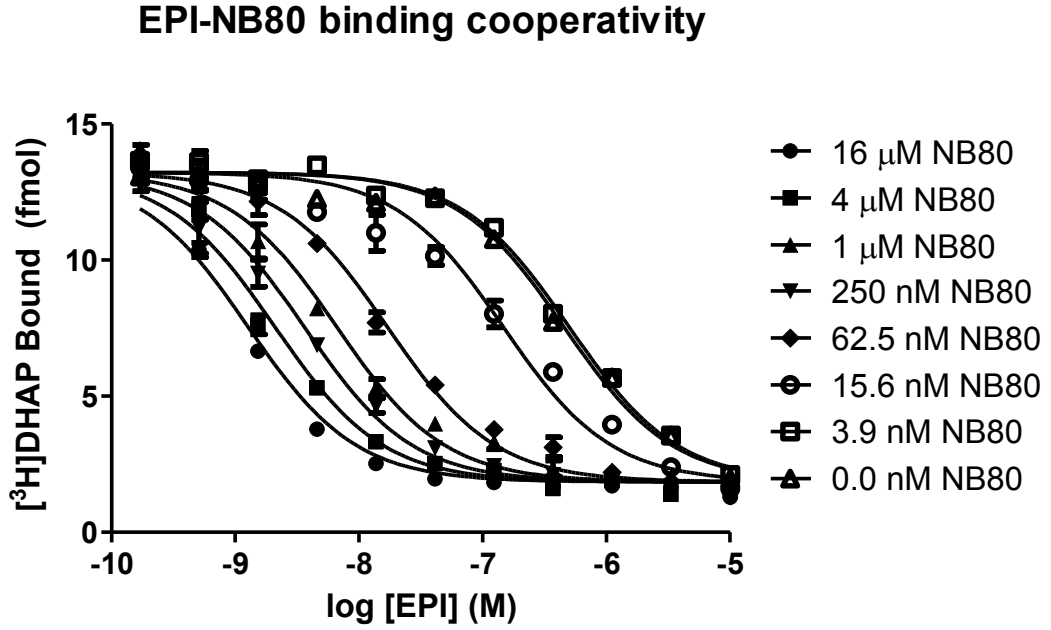


Figure 4-7. Competition binding of 2 nM [^{3}H]DHAP and epinephrine (EPI) in the presence of NB80. The fitted Ki estimate values are given below in table 4-4.The data points shown are the averages of 3 replicates with the error bars indicating the SEM.

	16 μM NB80	4 μM NB80	1 μM NB80	250 nM NB80	62.5 nM NB80	15.6 nM NB80	3.9 nM NB80	No NB80
Ki	2.554e-010	4.203e-010	1.325e-009	7.246e-010	3.406e-009	2.945e-008	1.003e-007	9.112e-008

Table 4-4. Ki estimate values for figure 4-7. Equilibrium constants are given in molar units.

4.5 – Agonist enhancement of nanobody 80 affinity is primarily due to faster on rates.

Despite our best efforts, we were unable to directly measure the binding of tritiated agonists (either epinephrine or 4-methoxyfenotrol) to the receptor in the presence of NB80. However, we were able to use a relatively new technology, bio-layer

interferometry, to measure the binding of NB80 in the presence of an agonist. We used a Blitz instrument (FortéBio) to determine how NB80 binds to receptors in rHDL discs. While the operation of the Blitz is not the focus of this section, a summary of its operation is appropriate given the novelty of the instrument and technique. Bio-layer interferometry relies on using a specialized glass rod which is coated with a partially reflective reference surface and an immobilized layer of proteins. When white light is sent down the glass rod, the beam hits either the reference or protein layer, and is reflected back to the detector. Using inferometery, slight differences between the two reflected light beams are compared, and in this way changes in the optical density of the protein layer can be detected. In our setup, biotin-labeled rHDL discs are bound to a streptavadin layer immobilized on the glass rods, and the binding of NB80 to receptor incorporated into the discs is followed as a function of time.

Previous work done by Jacob Mahoney from the lab has established that the association rate of NB80 to receptor increases dramatically when the agonist isoproterenol is bound to the receptor.[123] For example, when no ligand is bound, the G protein mimicking nanobody binds with a rate constant of about 3.1×10^4. In contrast, when the receptor is bound to saturating amounts of isoproterenol, NB80 binds much quicker, with a rate constant of 1.1×10^6. Overall, the affinity of NB80 for the receptor increases from 760 nM for binding to the unliganded receptor to 2.9 nM for binding an isoproterenol-saturated receptor. In other experiments, he also noticed when the receptor is bound to the inverse agonist ICI-118,551, negligible binding is detected.

We wanted to confirm that the natural hormone agonist epinephrine (EPI) would also increase the rate of NB80 binding. Although we were unable to run full sets of experiments to optimize the assay for epinephrine and obtain a full characterization of NB80 binding, we were able to perform some initial trials that appear to confirm similar behavior of epinephrine, shown below in figure 4-8A.

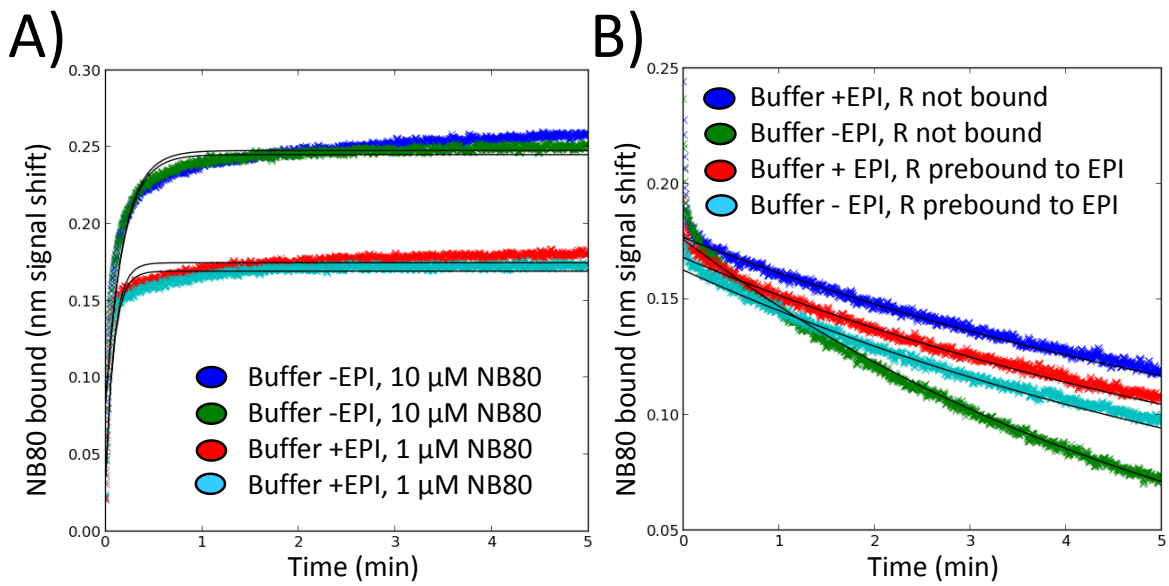


Figure 4-8. Kinetic binding assay of NB80 to β_2AR in rHDL discs. A) Association of 1 or 10 µM NB80 to receptor in the presence or absence of a saturating concentration of epinephrine. Differences in the maximal signal are due to non-specific binding of the nanobody to the glass rod, which finishes in the first 2 seconds of the traces. B) Dissociation of NB80 from receptor in the in the presence or absence of a saturating concentration of epinephrine.

This data shows that the bound agonist causes 1 µM NB80 to be able to bind receptor at a rate faster than even 10 µM NB80 can bind to the unliganded recpetor. The effect of the ligand bound to the receptor on the dissociation rate of NB80, shown above in figure 4-8B, is much smaller compared to changes in the association.

4.6 – Global modeling of receptor binding with agonist, antagonist, and nanobody.

The data presented in the previous sections can tell us several qualitative aspects about the activated conformation of the receptor and how it interacts with G proteins and NB80, listed below:

- Antagonist associates with the receptor very slowly when it is bound to either NB80 or nucleotide-free Gs.
- Antagonist dissociates very slowly from the receptor when it is bound to NB80.

- The slow antagonist binding to the receptor due to interaction with nucleotide-free Gs can be reversed with by re-addition of GDP or GTP.
- NB80 binding cooperatively enhances the affinity of agonist (or possibly decreases the antagonist affinity).
- Agonist binding cooperatively enhances the affinity of NB80.
- Most of the NB80 affinity enhancement is due to an increased rate of association.
- The dissociation rates of NB80 from agonist bound receptor and non-ligand bound receptor are similar.

We wanted to use these clues to build a full kinetic model of how the receptor interacts with antagonists, agonists, and NB80. At this point, we know the model cannot be fully optimized because we are not able to get direct binding data of an agonist to the receptor and because we need a fuller set of curves to define NB80 binding in the presence of both agonist and antagonist. However, we do have enough data to prototype the overall strategy of the data analysis. In addition to giving us some partial indications of the validity of the model, starting on the modeling and analysis was also a way for me to learn the programming and numerical computation skills that are necessary for completion of such a project.

This model takes advantage of our ability to use NB80 binding to manipulate the conformational state of the receptor in a way that simply is not possible with a heterotrimeric G protein. The model we used is identical in form to the extended ternary complex model (eTCM) of GPCR action, except that NB80 binds to the receptor instead of G protein. We chose to use a model based on the eTCM instead of the more thermodynamically complete cubic TCM model because we could find no indication that NB80 could actually bind to the inactive receptor. In this model, designated as model #1, we did not constrain any parameters to be identical to one another as we wanted to avoid making assumptions about the nature of the activated receptor conformation. The model is shown below, in figure 4-9.

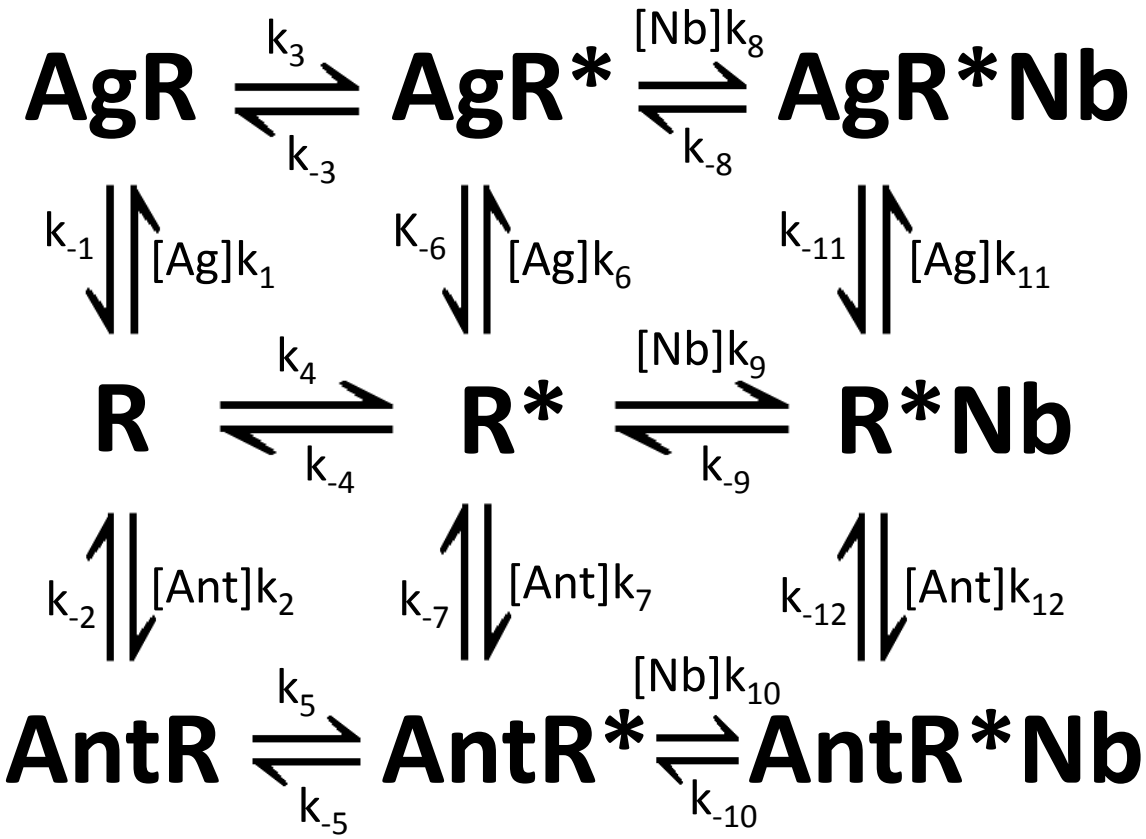


Figure 4-9. Model #1 used for global fitting of data. The bound agonist is represented by the symbol “Ag”, the bound antagonist is represented by the symbol “Ant”, the inactive and active conformations of the receptor are represented by the symbols “R” and “R*”, respectively, and the bound NB80 is represented by the symbol “Nb”.

The set of rate equations that the model describes can be represented as the matrix equation *d/dt* (**S**) = **MS** where **M** is a 9×9 matrix of the rate constants and **S** is a vertical vector of the concentration of the states. The matrices are shown below in equations 4-1 and 4-2.

$$\begin{bmatrix}
-k_{-1}-k_3 & [Ag]k_1 & 0 & k_{-3} & 0 & 0 & 0 & 0 & 0 \\
k_{-1} & -[Ag]k_1-k_4 \\ -[Ant]k_2 & k_{-2} & 0 & k_{-4} & 0 & 0 & 0 & 0 \\
0 & [Ant]k_2 & -k_{-2}-k_5 & 0 & 0 & k_{-5} & 0 & 0 & 0 \\
k_3 & 0 & 0 & -k_{-3}-k_{-6} \\ -[NB80]k_8 & [Ag]k_6 & 0 & k_{-8} & 0 & 0 \\
0 & k_4 & 0 & k_{-6} & -k_{-4}-[Ag]k_6 \\ -[Ant]k_7-[NB80]k_9 & k_{-7} & 0 & k_{-9} & 0 \\
0 & 0 & k_5 & 0 & [Ant]k_7 & -k_{-5}-k_{-7} \\ -[NB80]k_{10} & 0 & 0 & k_{-10} \\
0 & 0 & 0 & [NB80]k_8 & 0 & 0 & -k_{-8}-k_{-11} & [Ag]k_{11} & 0 \\
0 & 0 & 0 & 0 & [NB80]k_9 & 0 & k_{-11} & -k_{-9}-[Ag]k_{11} \\ -[Ant]k_{12} & k_{-12} \\
0 & 0 & 0 & 0 & 0 & [NB80]k_{10} & 0 & [Ant]k_{12} & -k_{-10}-k_{-12}
\end{bmatrix}$$

Equation 4-1. **M**, the model matrix. The matrix describes the mass transfer relationship between the different states of the model.

$$\begin{bmatrix}
[\mathrm{AgR}] \\
[\mathrm{R}] \\
[\mathrm{AntR}] \\
[\mathrm{AgR^*}] \\
[\mathrm{R^*}] \\
[\mathrm{AntR^*}] \\
[\mathrm{AgR^*NB80}] \\
[\mathrm{R^*NB80}] \\
[\mathrm{AntR^*NB80}]
\end{bmatrix}$$

Equation 4-2. **S**, the state matrix. The matrix describes the concentrations of the different states of the model.

Because all our experiments were performed under pseudo-first-order conditions, we could solve any equilibrium binding problem by finding the eigenvectors of the model matrix. Full details about the equilibrium solving method are given in section 6.29. Kinetic binding problems were solved by first calculating the initial distribution of states by assuming full equilibration with the given conditions before the reaction was started, followed by numerical integration of the rate equations given by *d/dt* (**S**) = **MS**. Full details about the kinetic solving method are given in section 6.30. Initial estimates for rate constants and other experimental variables were determined by analysis of the individual experiments and knowledge of several equilibrium constants. The initial estimates used for optimizing this model are given below, in table 4-5.

k_1	1.0×10^{8}	k_{-1}	1.0×10^{1}
k_2	1.4×10^{8}	k_{-2}	7.0×10^{-2}
k_3	3.0×10^{6}	k_{-3}	3.0×10^{5}
k_4	3.0×10^{6}	k_{-4}	3.0×10^{7}
k_5	3.0×10^{6}	k_{-5}	3.0×10^{7}
k_6	1.0×10^{6}	k_{-6}	5.0×10^{-4}
k_7	1.0×10^{6}	k_{-7}	5.0×10^{-4}
k_8	1.0×10^{5}	k_{-8}	2.0×10^{-2}
k_9	1.0×10^{5}	k_{-9}	2.0×10^{-2}
k_{10}	1.0×10^{5}	k_{-10}	2.0×10^{-2}
k_{11}	1.0×10^{6}	k_{-11}	5.0×10^{-4}
k_{12}	1.0×10^{6}	k_{-12}	5.0×10^{-4}

Table 4-5. Initial estimates for the optimization of model #1.

A custom Python program was designed to perform the data fitting and global optimization. The program weighted each curve to a similar extent, so that there was information from as many experimental conditions and measurements as possible incorporated in the final goodness-of-fit value. Full details on the program are given in section 6.28. The data shown in figures 4-4 through 4-8 were fitted to the model and used to optimize the values for the 24 rate constants, determined with repeated rounds of simulated annealing. The best solution found so far is given below in table

4-6. Also shown are the calculated microequlibrium constants for each binding or conformational change step in the model.

k_1	8.33×10^{6}	k_{-1}	2.73×10^{1}	K_1	3.28×10^{-6}
k_2	7.87×10^{7}	k_{-2}	4.97×10^{-2}	K_2	6.31×10^{-10}
k_3	6.08×10^{4}	k_{-3}	8.61×10^{2}	K_3	1.42×10^{-2}
k_4	1.53×10^{5}	k_{-4}	1.91×10^{5}	K_4	1.25
k_5	9.61×10^{4}	k_{-5}	3.95×10^{5}	K_5	4.11
k_6	3.08×10^{5}	k_{-6}	4.99×10^{-6}	K_6	1.62×10^{-11}
k_7	2.30×10^{3}	k_{-7}	2.19×10^{-6}	K_7	9.52×10^{-10}
k_8	1.14×10^{7}	k_{-8}	1.35×10^{-1}	K_8	1.18×10^{-8}
k_9	1.20×10^{6}	k_{-9}	1.76×10^{-1}	K_9	1.47×10^{-7}
k_{10}	5.53×10^{6}	k_{-10}	4.30×10^{-1}	K_{10}	7.78×10^{-8}
k_{11}	3.83×10^{5}	k_{-11}	7.89×10^{-6}	K_{11}	2.06×10^{-11}
k_{12}	3.41×10^{3}	k_{-12}	2.91×10^{-5}	K_{12}	8.51×10^{-9}

Table 4-6. Rate constants and the associated microequlibrium constants for the NB80 binding model #1 obtained by global fitting. Please refer to figure 4-9 for a diagram of the model.

A graph of simulated vs. real data using the best values we obtained for model #1 is shown below in figure 4-10.

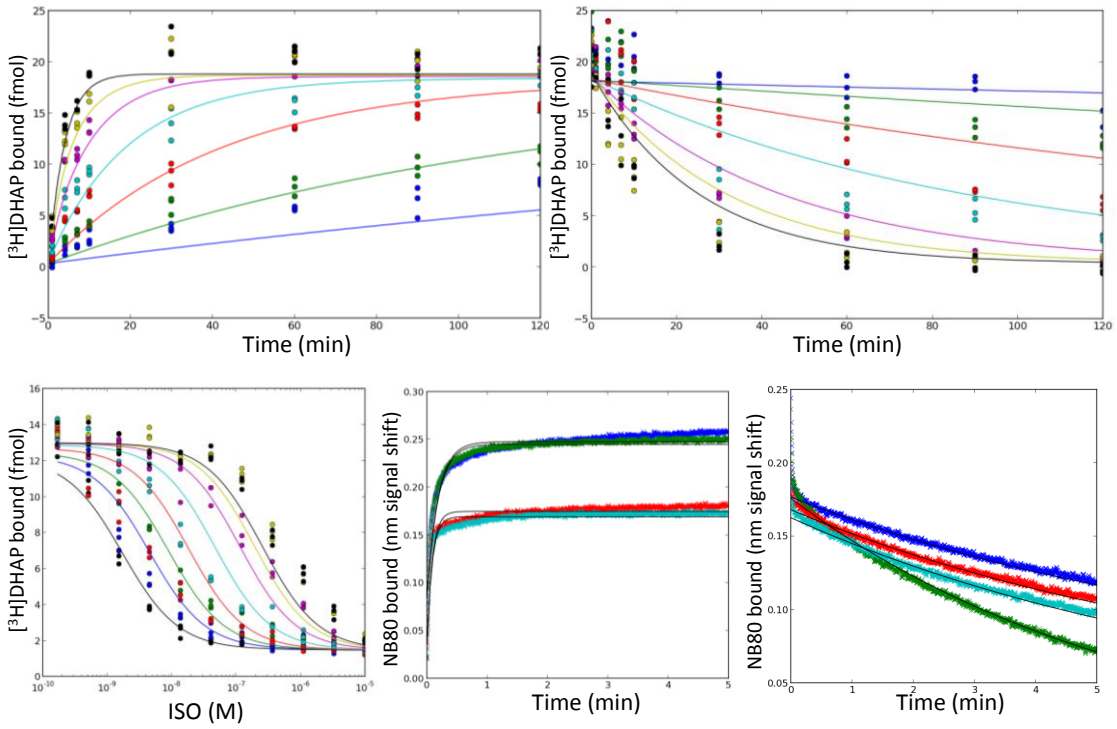


Figure 4-10. Graph of the best global fit obtained. Please see figures 4-4 through 4-8 for full descriptions of the experiments.

Several important aspects about the solution are apparent from the fitted rate constants and the data. In general, the NB80 modification of the eTCM model appears to be sufficient to explain the general pattern of the data. Although the fit is not perfect, the overall pattern of curves matches the data very well. Also, all of the rate constants that describe binding to an activated receptor are slower than the corresponding constants for the inactive receptor. This difference is more pronounced for the antagonist than the agonist, with may be a reflection of the larger molecular weight of the antagonist compared to the agonist used.

There also seems to be an indication of a partially active state for non-drug-bound receptor. Both agonist and antagonist bound receptors transitioned to the active state at slower rates than the unbound receptor. The agonist bound receptor showed a very slow transition back to the inactive state as well as binding the nanobody at very fast rates. Once formed, the antagonist-bound active receptor showed fast transition to either the inactive or nanobody-bound states, indicating that it is likely not a preferred conformation. However, the receptor adopts an approximately equal

distribution of active and inactive conformations, but binds nanobody about 10 fold slower than the agonist bound receptors. This indicates that the unliganded "active" conformation and the agonist-bound "active" conformation may not have exactly the same degree of active character.

After seeing how model #1 tended towards having partially activated states for unliganded and antagonist bound receptor, we wanted to see how well a model that required all the “active” receptor states to be conformationally equivalent. This could be accomplished by modifying model #1 so that the nanobody was forced to bind the antagonist-bound, agonist-bound, and unliganded receptor at exactly the same rate. To do so, all that was required was to replace the rate constants k_9 and k_{10} with k_8 and k_{-9} and k_{-10} with k_{-8}, giving us a new model, designated as model #2, drawn below in figure 4-11.

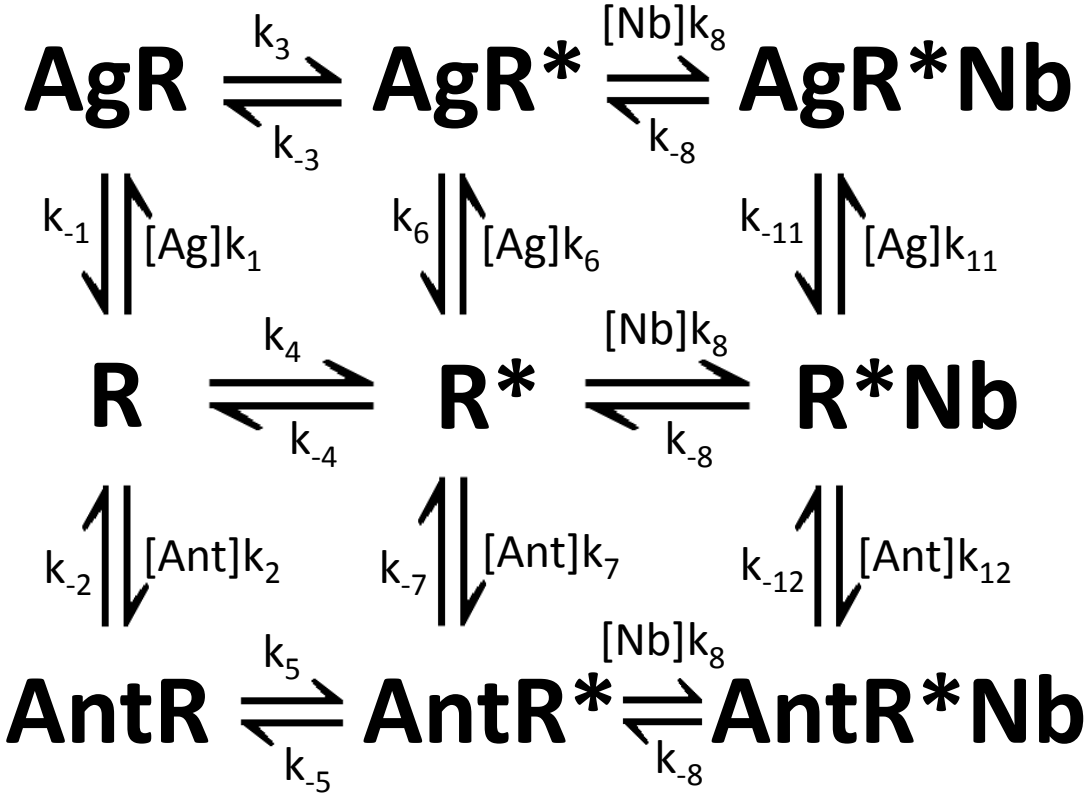


Figure 4-11. Model #2 used for global fitting of data. The bound agonist is represented by the symbol “Ag”, the bound antagonist is represented by the symbol “Ant”, the inactive and active conformations of the receptor are represented by the symbols “R” and “R*”, respectively, and the bound NB80 is represented by the symbol “Nb”. This model forces the R* state to be conformationally equivalent with respect to nanobody binding.

Model #2 was optimized in the same manner as model #1, using the best estimates for the rate constants of model #1 as the initial guesses, with the exception that the initial guesses for k8 and k-8 were set at 6.0 $x10^6$ and 0.2, respectively. The best estimate of the model #2 parameters are given below in table 4-7.

k_1	2.91×10^7	k_{-1}	1.47×10^1	K_1	5.04×10^{-7}
k_2	2.86×10^8	k_{-2}	8.23×10^{-2}	K_2	2.88×10^{-10}
k_3	1.36×10^5	k_{-3}	3.22×10^3	K_3	2.37×10^{-2}
k_4	6.58×10^4	k_{-4}	2.01×10^5	K_4	3.05
k_5	2.77×10^4	k_{-5}	2.62×10^5	K_5	9.44
k_6	9.92×10^5	k_{-6}	2.86×10^{-6}	K_6	2.89×10^{-12}
k_7	1.12×10^4	k_{-7}	3.11×10^{-6}	K_7	2.78×10^{-10}
k_8	5.03×10^6	k_{-8}	1.40×10^{-1}	K_8	2.79×10^{-8}
k_{11}	2.87×10^5	k_{-11}	4.06×10^{-6}	K_{11}	1.42×10^{-11}
k_{12}	4.25×10^3	k_{-12}	5.56×10^{-5}	K_{12}	1.31×10^{-8}

Table 4-7. Rate constants and the associated microequlibrium constants for the NB80 binding model #2 obtained by global fitting. Please refer to figure 4-11 for a diagram of the model.

A graph of simulated vs. real data using the best values we obtained for model #2 is shown below in figure 4-12.

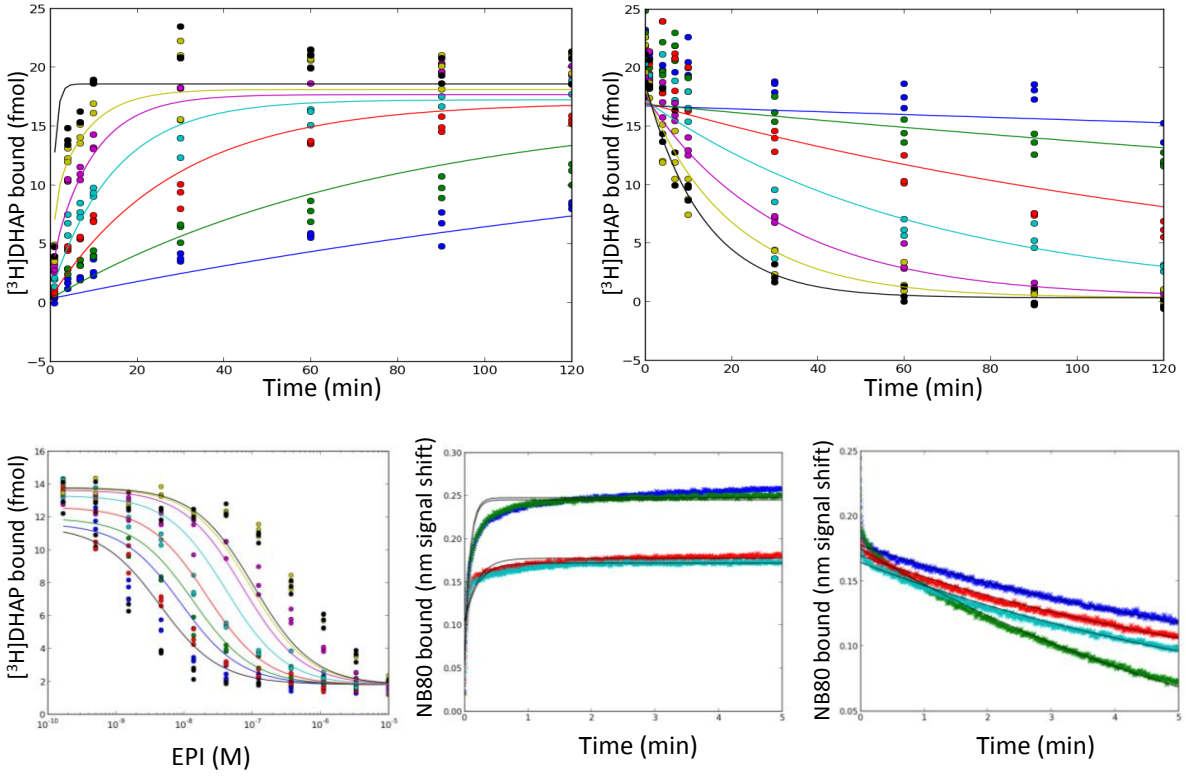


Figure 4-12. Graph of the best global fit obtained. Please see figures 4-4 through 4-8 for full descriptions of the experiments.

Model #2 gave a slightly worse fit to the data, but not terribly so. One way to measure this is to look at the final goodness-of-fit parameter for the given solution, the value of which is smaller for better fits. For model 1 this parameter was 180, and for model 2 this parameter was 268 (for comparison, the absolute minimum for the goodness-of-fit parameter is 30). One of the places where model 2 seems to be unable to accommodate the data well is where the on-rate of NB80 is measured using the Blitz instrument. This is not terribly surprising, as this is exactly where we constrained model 2 relative to model 1. In general, it appears that model 2 forces NB80 to bind to the unliganded receptor too quickly and to the agonist bound receptor too slowly. Again, this is suggestive that there is probably is a real difference between the active, agonist-bound conformation of the receptor and the "active" conformation that the receptor is able to reach without any ligand bound. The other experiment that model #2 seems to have trouble predicting is the equilibrium competition binding assay where NB80 causes an increase in agonist affinity for the receptor. In particular, it appears that forcing nanobody to bind

identically to all the different activated receptor conformations limits the amount of cooperativity that can be simulated between the nanobody and agonist. However, the overall difference between the models does not seem to be terribly large. Better quality and more data are needed before we would be confident in choosing between them or determining that neither is sufficient to describe the system.

It is worthwhile to look at how well defined each parameter is in the two models presented. A full estimation of the error in each measurement is beyond the scope of this study, but one can get a general idea of what the relative errors of the different parameters are by perfuming a simple sensitivity analysis. For this analysis, one simply increases or decreases the value of each optimized parameter one at a time and records the overall goodness-of-fit parameter. In general, the more well-defined a value is, the greater the increase of the goodness-of-fit parameter will be when the value is changed from its optimized value. Given in figure 4-13 below is the sensitivity analysis for increasing or decreasing each fitted parameter in the two models by an order of magnitude.

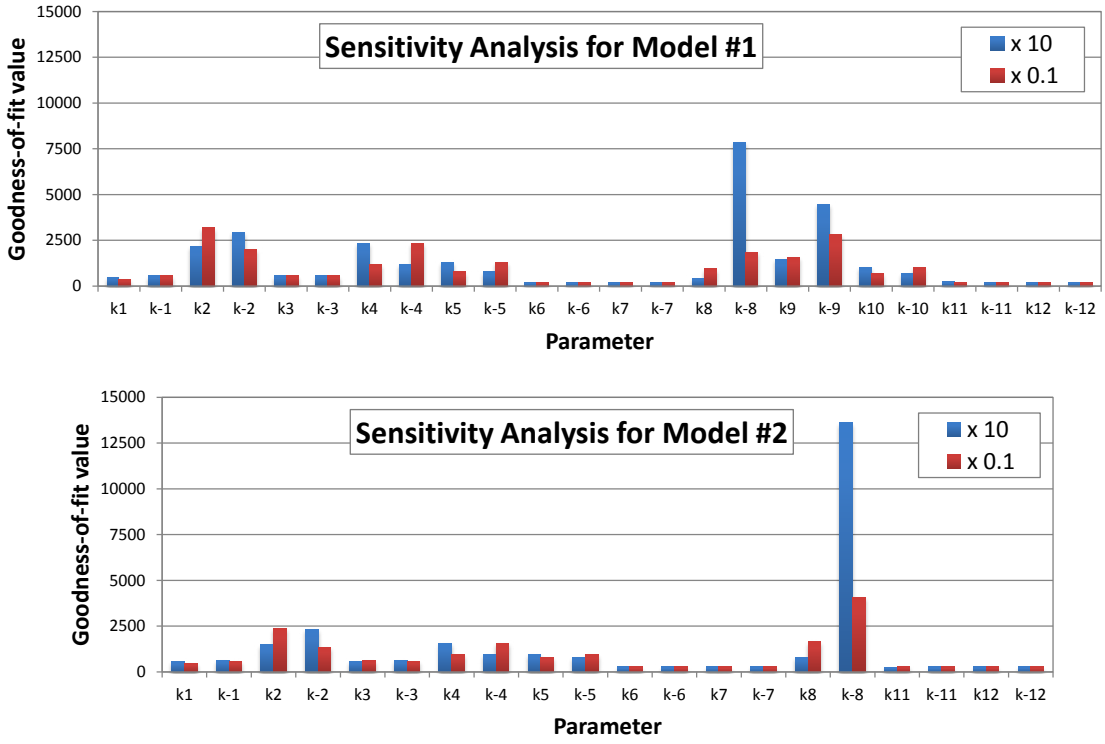


Figure 4-13. Sensitivity analysis of global model parameters for model #1 and #2. The goodness-of-fit value was recorded when the value of each parameter was changed an order of magnitude higher (blue) and lower (red) of the optimized value.

It can be seen in the sensitivity analysis that not all values for the parameters are likely to be well fitted. In particular, the rate constants for drug association and dissociation from the activated conformations of the receptor are not well determined by these models. This is related to the fact that these rate constants are expected to be quite slow in general, as access to the ligand binding site should be obscured when the receptor is in an activated conformation, as discussed earlier in this chapter and in chapter 3. Most likely, the best one will ever be able to do even with more extensive data is to get an upper limit for these values. It can also be seen that the most well estimated parameters are the rates of nanobody binding and dissociation, and in particular that of dissociation.

4.7 – Section summary.

In this chapter, we complemented our crystallographic evidence with pharmacological assays that also showed strong evidence that an integral part of

receptor activation is the tightening of the ligand binding pocket. This tightening is stabilized by agonists and promoted by binding of nucleotide-free G proteins or a G protein mimic nanobody. The tightening of the ligand binding pocket that is promoted by nucleotide-free G protein can be reversed by the addition of GDP as well as GTP, which is evidence that the G protein cannot bind both the receptor and GDP at the same time.

In an effort towards determining the full kinetic model of how a receptor and G protein interact, we collected data and started modeling the interaction of the receptor and a much simpler G-protein mimicking nanobody. There is still more work to be done before even the simplified nanobody system can be confidently solved in full kinetic detail, but so far the results are consistent with our overall hypothesis that the binding of G protein to the activated receptor is coupled the loss of GDP from the G protein.

CHAPTER 5
EXTENDED DISCUSSION AND CONCLUSIONS

5.1 – Overview of discussion.

This section extends the analysis of the data presented in chapters 2-4 above. There are three general themes that are discussed. The first theme, found in sections 5.2 and 5.3, concerns the practical advances in scientific techniques that this work demonstrates, as well as some of the direct questions that they raise. The second theme, sections 5.4 and 5.5, is a thorough analysis of the major findings in chapters 3 and 4 from a more holistic viewpoint. The third theme, found in sections 5.6, 5.7, and 5.8, is speculative in nature and regards some new avenues of research that this work suggests might be productive.

5.2 – Advances in G protein purification.

One of the major technical achievements in this body of work is the re-working of our Gs heterotrimer purification scheme. Originally, we started with a protocol based off of the work of Tohru Kozasa,[73] but we required that the protein to be in the detergent DDM instead of Tohru's Lubrol detergent. This was a surprisingly difficult change to make. In our first iterations of making the detergent switch, we simply substituted DDM for $C_{12}E_{10}$ (the very close relative of Lubrol we had been using), but the switch ended up causing the final samples to be significantly dirtier. Eventually, we figured out two details that were probably causing the problem and how to compensate for them. Some examples of the purity issue are shown below, in figure 5-1.

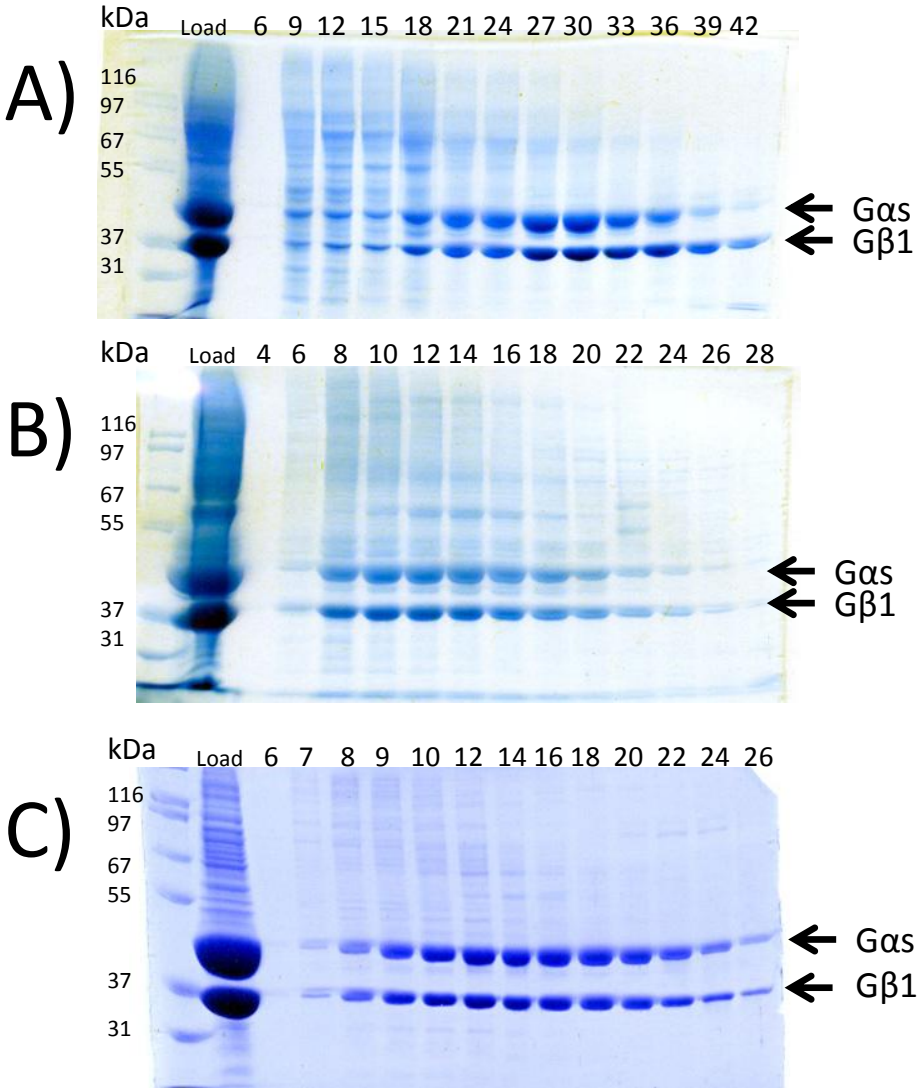


Figure 5-1. Purity of Gs heterotrimer samples. Shown are the SDS-PAGE analyses of the final gel filtration step for several different Gs purification runs. A) Purification with cholate-CHAPS-Lubrol detergent sequence as in Kozasa et al.[73] Notice how the remaining impurities are mostly at higher molecular weights than the G protein and can easily be excluded from the final sample. B) Purification with cholate-DDM detergent sequence as initially attempted. Notice how the G protein sample is no longer easily separable from the remaining impurities. C) Purification with optimized cholate-Anzergent-DDM detergent sequence as given in section 6.1. The increase in purity is mainly due to overall reduction of the level of impurities in the sample before gel filtration.

One reason why we were getting reduced purity was due to the nature of the detergents themselves. Both $C_{12}E_{10}$ and DDM are nonionic detergents with low critical micelle concentration (cmc) values (0.013% and 0.0087%,[74] respectively) and rather large micelle sizes (the micelle size for $C_{12}E_{10}$ is not known but would be slightly smaller than the 83 kDa for $C_{12}E_9$, 71 kDa for DDM). However, the

alternating hydrophobic-hydrophilic nature of the $C_{12}E_{10}$ polyoxyethylene head group allows the $C_{12}E_{10}$ micelles to fuse and split easily, since the headgroups neither attract nor repel their neighbors strongly. In contrast, detergents with extremely hydrophilic sugar-headgroups like DDM create a large hydration shell and resist being too close to neighboring headgroups, so their micelles tend to avoid fusion and splitting, favoring instead association and dissociation of single molecules from the micelle.[124] A similar phenomenon can also occur with ionic detergents where the headgroups repel one another based on like charges.[125] Consequently, when our Gs heterotrimer was solubilized in DDM, the only way for any contaminating membrane proteins embedded in the same micelle as a G protein to move to a new micelle would be for either the contaminant or the G protein to dissociate into aqueous solution and back into a different micelle, an inherently slow process. In contrast, when the Gs was solubilized in $C_{12}E_{10}$, the micelle could split and fuse with empty micelles, separating contaminating proteins without ever forcing them to dissociate into aqueous solution.

Another reason why we saw increased contamination was due to the abrupt exchange into DDM that the sample underwent during purification. In the purification scheme for Gs heterotrimer, the protein is solublized from membranes with sodium cholate as the detergent. This detergent is cheap and highly effective at solublizing phospholipid membranes. However, it is neither compatible with anion exchange columns nor does it support ligand binding of β_2AR or most other GPCRs. In order to exchange the cholate for another detergent that is more compatible with the columns and receptors, the solublized, clarified membranes are diluted underneath the cmc of cholate using buffers with the new detergent. When we initially diluted the membranes directly in DDM, we did so by simple mixing. Eventually, we figured out that this was a mistake. The quick dilution caused the cholate micelles to rapidly disperse, but the previously solublized membrane proteins seemed not to be able to transfer to the DDM micelles very effectively, with the samples becoming somewhat cloudy upon dilution (this observation was not immediately identified because solutions with DDM inherently tend to be somewhat cloudy to begin with). If the dilution solution was slowly added to the solublized membranes over about 20-30

minutes on ice, the turbidity of the resulting solution was markedly reduced. Chromatrography of the G protein was also improved after using the slow dilution method.

While it was effective to dilute the soluble membrane samples directly into DDM, the large volumes of buffer needed for the dilution step made the direct use of DDM prohibitively expensive for routine purifications. Instead, we used an intermediate detergent, anzergent 3-12, for the dilution and then exchanged our sample into DDM on-column during the Ni-NTA purification step. Anzergent 3-12 was an ideal choice because it is relatively inexpensive, and has a cmc that is intermediate between cholate and DDM, which promotes good detergent exchange of samples from cholate but does not inhibit later exchange of the sample into DDM. We also made sure to wash the Ni-NTA column free of diluted, solublized membrane after the loading step. We used buffer that contained the exact same concentrations of Anzergent, cholate (now under its cmc but likely still interacting somewhat with the micellar phase), and ionic strength (which is known to effect both micelle size and cmc for many detergents). Only after we knew that as much contaminant protein as possible was washed through the column did we then exchange the sample into DDM, which appeared to prevent most of the co-localization of contaminant and target protein into the same DDM micelle.

After switching to buffers with rationally designed detergent sequences, the variability in the purity of the final protein sample was markedly reduced. Unfortunately, the yield of the purifications was still quite variable. From the analysis of SDS-PAGE gels, it appeared that all the protein that was extractable from the membranes with cholate was also being successfully purified, with minimal losses after extraction. However, large amounts of protein were left in the insoluble fraction after cholate solubilization of the membranes, and extended extraction at a variety of temperatures and in several detergents could not enhance the solublization (data not shown). This led us to conclude that this insoluble protein was denatured. In this case, enhancing the yield of G protein beyond the average of 0.75 mg/liter culture volume is primarily dependent on finding a way to improve the ability of insect cells

to successfully fold the Gα and Gβ subunits of the G protein. Luckily, the path forward to higher expression levels has probably already been shown.

Folding of the Gβ subunit, and of β-propeller protein in general, is difficult and generally requires the cellular chaperone machinery. Specific chaperone proteins for both Gβ and a Gγ subunits have been identified as phosducin-like protein (PhLP),[67] and dopamine receptor interacting protein 78 (DRiP78),[68] respectively. It is quite likely that coexpession of these proteins during overexpression of the Gβγ subunit would increase the yield dramatically. A similar finding has also been made with Gα subunits. There are two chaperone proteins that help fold Gα; Ric8B helps Gαs to fold, and Ric8A helps most of the other Gα subunits. Specifically, it has been found that the Ric8 proteins bind to the nucleotide-free Gα proteins and are released upon GTP or GDP binding.[126]

Currently, we use three high-quality baculoviruses to express G protein in Hi5 insect cells. There is one virus that carries each gene for the Gα, Gβ, or Gγ subunit that we wish to express, and the appropriate combination is used to co-infect the cells all at the same time. However, advances in commercially available baculovirus transfer vectors now allow one to insert two genes for overexpression in each virus. It is highly likely that if each of our subunit viruses were remade so that their specific chaperone protein was co-expressed, we could increase yields of the heterotrimer. Of course, there are many other possible ways to improve G protein yields, but the proposed change would not require any other modifications of the overall procedure which we already know produces high-quality protein.

5.3 – Technological advances for membrane-protein crystallography.

The ability to crystallize the β_2AR and other GPCRs has so far required the use of varied stabilization techniques with the exception of only two receptors; bovine rhodopsin and squid rhodopsin. While it is unfortunate that as of yet nobody can get structural information with wild type receptors that are not rhodopsins, the information gained by analysis of even imperfect structures has proven invaluable to our understanding of GPCR function. Hopefully, within the near future less invasive strategies to obtain membrane protein structural information will be discovered. For

example, crystallographic data collection with the recently developed free-electron laser[127] might be able to compensate for the general small size and high sensitivity to radiation damage characteristic to GPCR crystals. Alternatively, recent work with patterned nanoarrays has yielded some that hold small molecules in crystal-like arrays that are capable of producing Bragg reflections for X-ray analysis, but do not require actual crystallization of the compound.[128] Similar arrays may soon be able to hold proteins as well; a leading candidate for making such an array are folded DNA scaffolds because of the rapid progress in the field and the compatible length scales of the arrays with typical protein dimensions.[129] However, it is our laboratory's opinion that the pressing need for structural information of GPCRs, and membrane proteins in general, necessitates the application of the proven but imperfect technologies described below to an ever widening circle of biologically important proteins.

One of most important techniques for getting GPCRs to form crystals is genetic fusion of the receptor with easily crystallizable protein domains. The first example of this strategy was in 2007, when Brian Kobilka and Ray Stevens replaced the flexible ICL3 region of the β_2AR with a catalytically dead but well-folded mutant of T4 lysozyme (T4L). A very similar strategy uses a receptor fusion construct with a thermostabilized apocytochrome (b_{562}RIL, or BRIL) domain in the same position. So far, these two IL3 fusion protein replacement strategies have been used to successfully crystallize 15 other GPCRs.[94, 130-143] In addition, this work and a more recent structure by Yaozhong Zou et al[144] also show that simply fusing the T4L to the N-terminus of the GPCR can be sufficient to promote crystallogenesis, avoiding the structural perturbations caused by ICL3 replacement.

Another strategy that promotes crystallization of difficult targets and rare conformations is antibody stabilization, in particular using nanobodies derived from camelids. The use of traditional antibody F_{ab} fragments for promoting crystallization is well known,[145] and they have even been used for successful GPCR crystallization as well.[52] However, the small size of nanobody fragments appears to make them particularly adept at stabilizing various 3-dimensional conformations of the target

protein, as opposed to the linear peptide epitopes that are often recognized in the binding cleft of traditional antibodies.[146] In addition, the single-polypeptide, single-domain nature of these fragments allows for efficient selection of clones by screening of phage display libraries as well as fast expression and purification of the selected clones.

The use of bicelles and lipidic cubic phase (LCP) based matrices for membrane protein crystallography has also been helpful for GPCRs. Traditional membrane protein crystallization techniques are pretty much identical to those used for soluble proteins, but performed in the presence of detergent. As discussed above in section 5.2, the maltose-based headgroups on the detergents that have been the most successful for GPCR purification generally make micelles that are resistant to fusion and splitting. This includes the new MNG class of detergents used in this study. It is expected that this property would typically be detrimental to the process of crystallization, as the receptor molecules in solution must fuse their associated micelle with the growing crystal's detergent layer in order to pack in any orientations that have contacts between the receptor transmembrane regions. In contrast, the bilayer-like nature of both bicelles and LCP easily allow for receptors to contact the growing crystal in a parallel or antiparallel orientation to the receptors already incorporated in the crystal. Additionally, this bilayer-like character is more similar to the receptor's native environment of a phospholipid membrane, which should protect the receptor against denaturation during the crystallization process.

LCP in particular has proven to be very useful for crystallization of transmembrane proteins.[147] Likely, this is related to the 3-dimenionsal nature of the matrix, which enhances diffusion and stacking of these proteins, which are typically restricted to 2-dimentional movement in their native lamellar bilayers. However, the narrow aqueous channels in the typical monoolein-based matrix appear to be inhibitory when the transmembrane protein of interest also has large soluble domains, as was the case for our work with the large Gs heterotrimer in complex with the receptor. For this situation, the use of different lipids that form a cubic phase with larger aqueous channels was critical to successfully using the method for crystallography.

In general, this also seems to be the case for other transmembrane proteins with large soluble domains as well.[147] Even with this improvement, though, the LCP method suffers from a serious drawback. The crystals grow surrounded by the extremely viscous cubic phase, and no amount of careful washing can remove the crystal from the surrounding matrix without destroying the crystal. Instead, one must simply scoop the lipid matrix and embedded crystals into a crystallography loop and freeze it as is. The lipid matrix is both optically dense and highly birefringent, making it impossible to see the embedded crystals in typical X-ray crystallography beamline setups. This is a formidable challenge, as is it pretty much impossible to obtain high-quality data when the crystals cannot be accurately positioned in the X-ray beam.

For this work, that challenge was overcome by the use of the GM/CA-CAT beamline at Argonne National Lab. The work of the beamline staff on the controlling software package, JBluce, allowed us to efficiently screen through the entire volume of the frozen LCP blob with a low-power X-ray beam, looking for places where the invisible crystals produced Bragg peaks in the diffraction pattern. Because the target protein was so difficult to crystallize in general, we needed to locate the few good crystals in a very large pool of poor quality crystals. The GM/CA-CAT beamline is also one of the few microfocus beamlines available for protein crystallography. At the time we collected data for the β_2AR-Gs complex, beam diameters down 5 μm were available for use, which allowed us to focus on individual regions of our small crystals. This allowed us to take diffraction patterns at very high beam intensity, which promotes high-resolution diffraction, on one region of the crystal until it was too damaged to continue, then move to a different region of the same crystal and continue. Using judicious selection of crystal rotation as we consumed the different regions of each crystal, we were able to obtain a complete dataset with the highest possible resolution for the crystals, even though we were very limited on the amount of good quality crystals. The same challenges will most likely exist for many other GPCR crystallization projects, so the technological advancements made by GM/CA-CAT and similar facilities to lower the barrier of what quality of crystals can practically yield useful data will allow increasingly challenging target proteins to finally be structurally characterized.

5.4 – Structural basis for GPCR-mediated G protein activation.

One of the most important parts of this work is its contribution to our understanding of how the structures of a GPCR and a G protein change during the process of activation. Some aspects of the findings, like the movement of TM6 during activation and the interaction of the Gα C-terminal helix, have been well characterized in the past and are confirmed again here. However, several new findings can also been seen from the structure, such as the exact orientation of the G protein to the receptor and the wide separation between Gα subunit domains. Shown below in figure 5-2 is a summary of the structural changes that allow the agonist binding to the β_2AR to cause the nucleotide loss on the Gαs subunit of the Gs heterotrimer.

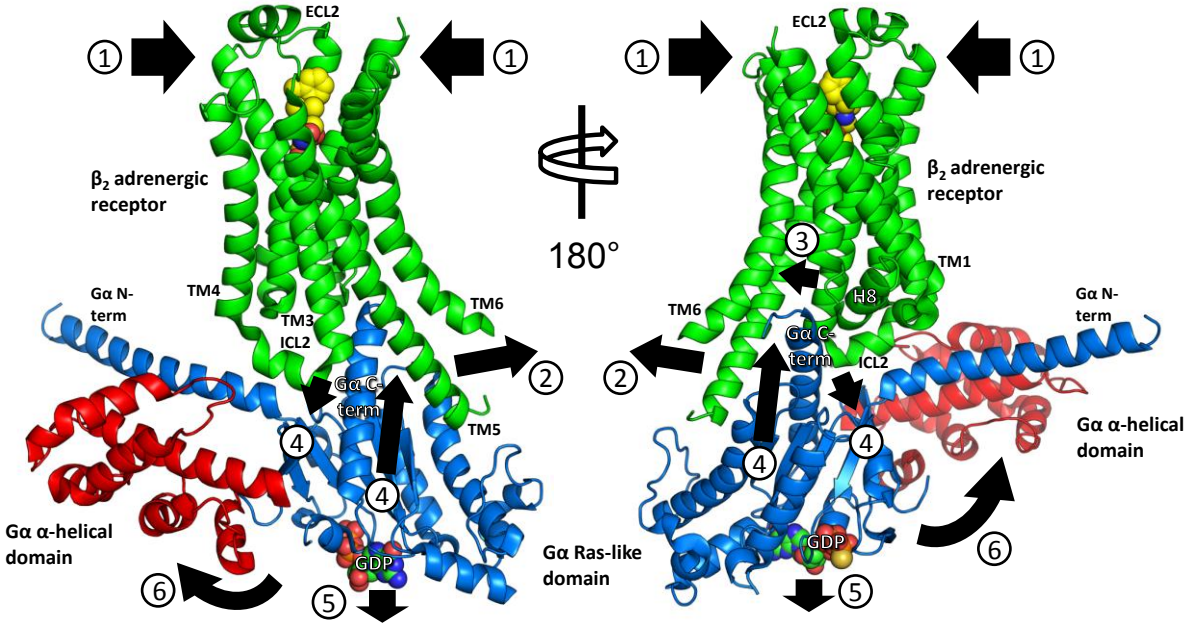


Figure 5-2. Overview of structural changes involved in activation of β_2AR and Gs to the β_2AR-Gs nucleotide-free intermediate complex. Numbered features are as follows: 1) Agonist causes ligand binding site to contract by 1-2 Å, mainly by pulling TM5 and TM7 towards the stable TM3. 2) The ligand binding site contraction on the extracellular side of the receptor causes TM5 and TM6 to be pushed outward on the intercellular side, similar to the working of a lever. 3) TM7 moves in and the NPxxY motif bulges in order to stabilize the outward movement of TM6. 4) The cleft opened by TM6 movement binds the Gα C-terminal helix and the ICL2 helix binds next to the Gα N-terminal helix. 5) The binding mode of the Gα to the receptor causes major disruption of the guanine nucleotide binding residues, leading to the loss of GDP. 6) Loss of nucleotide also disrupts the interface between the Ras-like domain and the α-helical domain of Gα, causing the α-helical domain to disengage and adopt a flexible position relative to the Ras-like domain.

There are many more details about the structural changes during β_2AR-Gs complex formation in sections 3.5 through 3.9, but here we summarize some more general structural aspects of the interaction between β_2AR and Gs. One issue that this structure does not fully answer but instead leads to more questions is that of the role of Gβγ subunits in G protein activation. In all our work, Gβγ was required for the interaction between the β_2AR and Gs and it did not dissociate after Gα nucleotide loss, but in the complex structure Gβγ doesn't actually contact the receptor. As far as we can tell, its primary role seems to be binding the Gα N-terminal helix and helping to present it on the membrane in the correct orientation for receptor interaction. This is consistent with several Gαi structures that show the N-terminus folding up into a microdomain that also binds the end of the C-terminal helix. This microdomain covers up most of the surfaces on the C-terminal helix that interact with the receptor as well as the binding site for ICL2 binding to Gα, but in the heterotrimer the microdomain does not form and instead the N-terminus binds to Gβγ, shown below in figure 5-3.

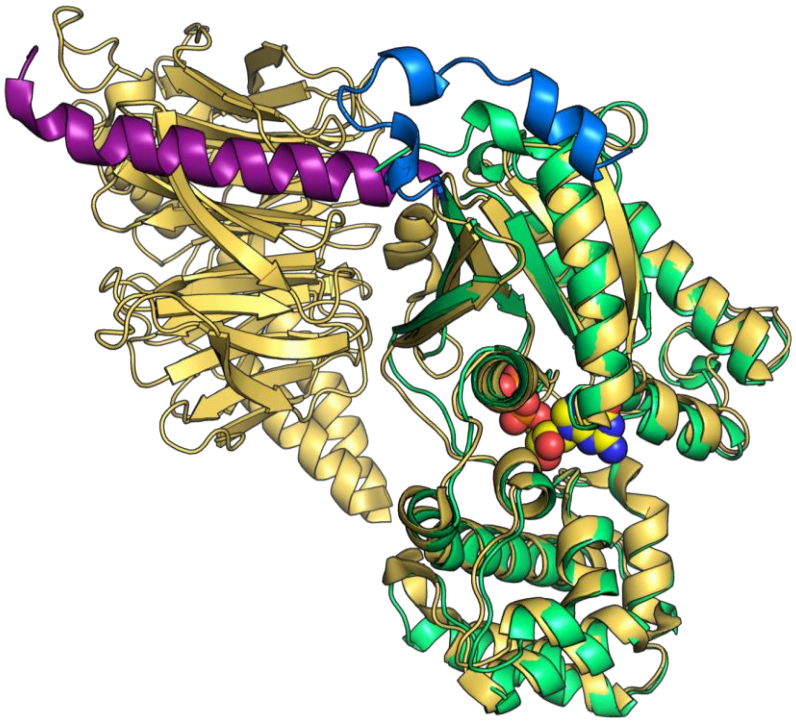


Figure 5-3. N-terminal microdomain formation on Gαi. The structure of Gαi1 bound to GDP and Mg^{++} (PDB 1BOF)[148] is drawn in green with a blue N-terminal domain. The structure of the Gαi1 heterotrimer (PDB 1GP2)[69] is drawn in tan with a purple N-terminal domain.

Even if a Gα subunit does not form an N-terminal microdomain similar to the one observed for Gαi1, it could likely still benefit from being displayed in the correct conformation for receptor interaction next to the membrane by a Gβγ subunit. It also seems likely that Gβγ-binding of the N-terminus would aid in putting strain on the β1 strand of the Ras-like domain, since it severely limits the ability of the N-terminus to adopt alternate positions that would accommodate the binding of the receptor ICL2 domain. However, most of these proposed roles for Gβγ are still quite speculative and more work is needed to determine why it is generally so difficult to get Gα subunits to interact with receptors without its presence.

The receptor and G protein complex structure also shows why several regions of both proteins are so highly conserved. Most of these conserved residues are ones that participate directly in the movements of the proteins. In order to help visualize this relationship, shown below in figure 5-4 is a map of the relative entropy of each position in the 7TM receptor or the Gα family, as determined by the Pfam hidden markov model for each family, onto the β_2AR-Gs complex structure. Relative entropy is basically a measure of sequence conservation, except that it takes into account how likely one is to see a particular type of residue in the first place given its overall frequency in the sequence.

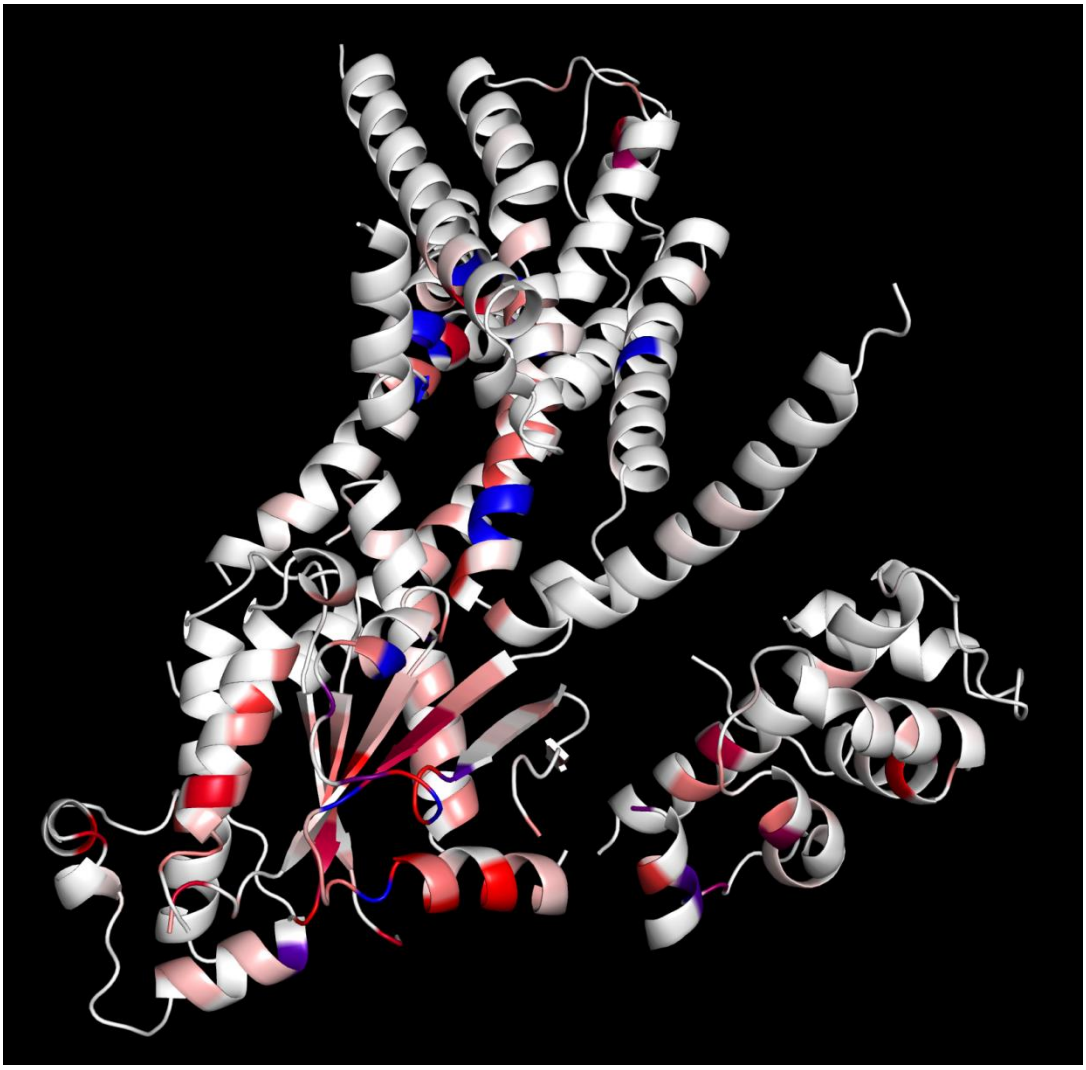


Figure 5-4. Relative entropy of the 7TM receptor or the Gα family. Places with low relative entropy (low sequence conservation) are white, changing to red and then blue as the residues is more conserved.

The figure above shows the region with the most conserved residues includes the G-box domains, which make up the nucleotide-binding site and nearby residues in the Gα subunit. On the receptor, the NPxxY and DRY motifs also have very low relative entropy, as expected. It can be seen that other residues in the receptor are also fairly conserved, mainly those that are involved in packing interfaces with other TM helices. Similar analyses have determined that the residues that pack in the central part of the receptor tend to make contacts with one another that do not require the residues to be completely conserved, but instead have a more generalized pattern,[40, 149] which is consistent with the observed intermediate relative

entropies that suggest partial conservation. Another observation that can be made with this type of analysis is that unlike the residues involved in the conformational changes that happen during activation, the actual protein-protein interaction surface of the G protein and the receptor is not terribly well conserved within the protein families. This suggests that the identity of the interacting residues is likely less important than the overall topology of the surfaces.

5.5 – Allosteric communication between agonists and guanine nucleotides.

Consistent with the tighter ligand binding pocket on the receptor shown by crystallography, we saw slower association and dissociation of ligands into the receptor when it was forced into the active conformation by binding the G protein mimic NB80. The same behavior was seen when heterotrimeric G protein was used instead of NB80, but the G protein had to be nucleotide-free in order to interact with the receptor in a manner that closes the ligand binding site. Both GTP and GDP can bind to the nucleotide-free G protein and allow it to be released from the receptor, but GTP (and GTPγS) will cause the G protein subunits to split from each other, which prevents reengagement of the G protein with the receptor. However, GDP binds reversibly, allowing G protein and receptor to periodically disengage, but then come back together as the G protein-receptor nucleotide-free complex, which releases the GDP back into solution. These observations lead us to propose that the model needed to describe the relationship between the GPCR and G protein is *at least* as complicated as a version of the eTCM model modified to include the binding of GDP as an integral part of the interaction. This model is shown below, in figure 5-5.

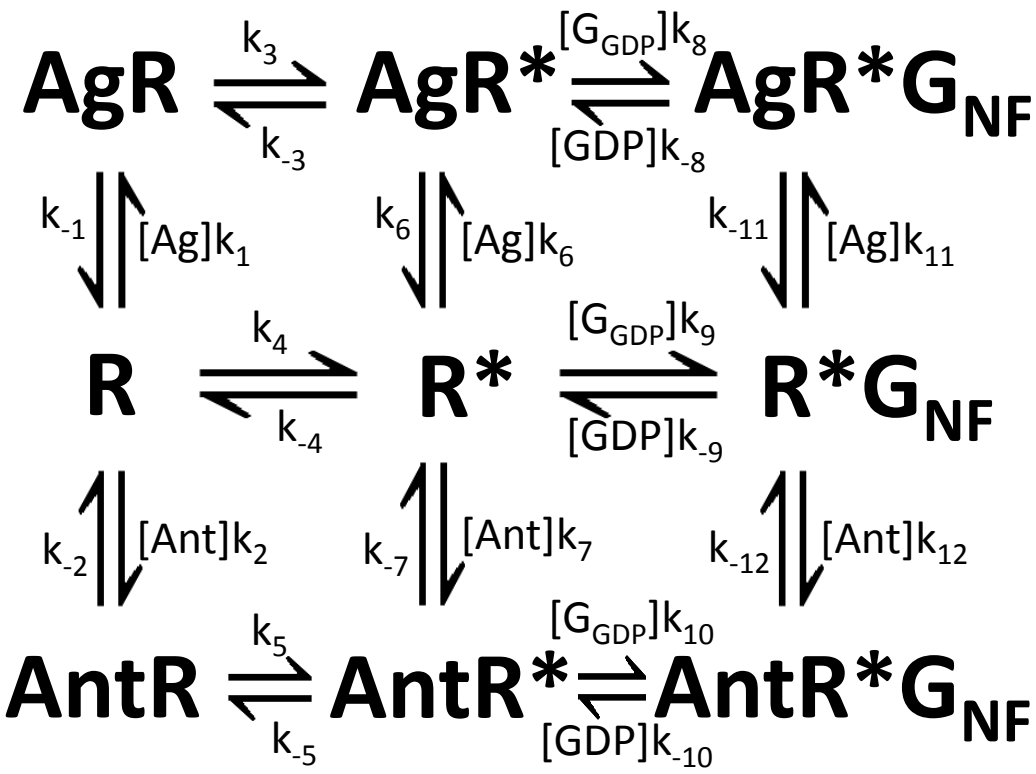


Figure 5-5. Minimum possible model of GPCR and G protein interaction. The G protein is assumed to be nucleotide-free, and the loss of GDP and the binding of receptor are modeled as a concerted reaction.

This model of GPCR function can include reduced diffusion of ligands into activated receptors and can also account for all features of our observed competition binding curves between radiolabled antagonist and unlabeled agonist of samples that contain Gs-β_2AR nucleotide-free complexes in figure 4-6. When there is no nucleotide present in the assay, only one low affinity site exists for the agonist which corresponds to any receptor in the sample that is not interacting with G protein. Receptors that are complexed to nucleotide-free G protein and assayed without any nucleotide present are stuck in an activated, closed conformation and thus cannot bind the drug in the time frames that are typically used for such assays. Under conditions where GDP is added in concentrations approaching its apparent affinity for Gs, radiolabeled antagonist can bind to the receptor when the G protein binds GDP, allowing the receptor to access its open, inactive conformation before being forced back into the closed, active conformation when GDP is lost from the G protein. In this way, a moderate level of GDP uncovers the receptors that had been

invisible to the assay because they were locked in the active conformation in a nucleotide-free receptor-G protein complex. The receptors are then able to bind competing radiolabeled antagonist and unlabeled agonist concentrations, and the positive cooperativity between agonist and G protein binding causes the appearance of a high-affinity binding site in the competition assay. However, very large amounts of GDP would increase the off-rate of the G protein and eventually reduce the affinity of the high-affinity state so that it would become indistinguishable from the low-affinity state.

The sensitivity of the high-affinity agonist binding site to GDP in compeitition assays needs to be characterized for other receptor systems and in more detail than reported here. A major prediction of the model presented above would be that in a true equilibrium assay, the K_d value of the high-affinity state should vary with the GDP concentration, becoming lower affinity with higher amounts of GDP present. If this cannot be found, it is likely that a model akin to the cubic ternary complex model would be needed to explain the results. This type of model would allow the G protein to remain bound to the receptor and GDP at the same time. Most likely, this receptor and GDP-bound G protein would be in a conformation that is much different than our nucleotide-free receptor-G protein complex structure. It is also probable that new biophysical assays will have to be developed in order to test these hypotheses, as the ligand binding assays used extensively in this work offer only an indirect measurement about what is happening on the G-protein binding side of the GPCR.

5.6 – Flexible α-helical domains and the mechanism RGS protein activity.

The surprising finding that the Gα α-helical domain moves so dramatically during the activation process suggests the interesting possibility that the domain may also be flexible in the GDP or GTP bound states as well. Although at first glance it seems that the Gα α-helical domain must contribute significant surface area to the nucleotide binding site, the only contacts made between the domain and the GDP/GTP are in a small area on the ribose portion of the nucleotide. Most other contacts which define the binding site, including 4 of the 5 highly conserved G-box sequences, are completely located on the Ras-like domain. Since Ras is able to bind

GDP at nanomolar concentrations and it is also known that the α-helical domain can be removed from Gαs and the Ras-like domain expressed separately as a functional domain,[150] it seems clear that the α-helical domain and the nucleotide ribose group need not be in constant interaction in order to maintain the integrity of the protein.

This hypothesis that the α-helical domain can move away from the Ras-like domain while in a "normal" GDP or GTP bound state could have physiological implications. It is known that in the typical cell, most G proteins do not become deactivated by hydrolysis of GTP to GDP by their own intrinsic GTPase activity, but instead by hydrolysis stimulated by the binding of an RGS protein. Based on structural analysis of GDP, $[AlF_4]^-$, and RGS bound Gα structures, the consensus explanation of how RGS proteins stimulate GTPase activity has been that the RGS binding protein subtly pushes the catalytic residues into optimal positions for enhancing the reaction.[151] However, it is possible that an additional job that RGS binding does to the Gα subunit is simply to force the α-helical domain to remain in close proximity to the Ras-like domain, allowing the catalytic arginine to stay near the γ-phosphate of the GTP. Structures of several RGS-Gα complexes show residues of RGS and the Gα α-helical domain in close proximity (≤ 4 Å) to each other, and the nature of these interactions varies with the specific pair of proteins used.[152] An addition, a recent structure of RGS2 and Gαq shows extensive contact between the α-helical domain and the RGS protein, and mutations in the interface are associated with large changes in GTPase activating activity (both in positive and negative directions).[153]

In many ways, a mechanism that works this way is more consistent with the hydrolysis method of the Ras family of small GTPases. When GTP binds to the Ras protein, it cannot be hydrolyzed quickly without the binding of another protein that carries the catalytic arginine residue, known as RasGAP. RasGAP accelerates the GTPase activity partly by simply putting the missing catalytic arginine near the rest of the enzyme's active site, but also by stabilizing the switch II region of the Ras protein into a conformation that favors the transition state for hydrolysis.[154] For heterotrimeric Gα subunits, these two jobs are separated so that the α-helical domain supplies the catalytic arginine and the RGS protein provides the switch

domain stabilization. However, if the α-helical domain is normally not positioned correctly for catalysis and is instead moving around, having RGS bind the α-helical domain to keep it in place would achieve the same end as the RasGAP protein.

5.7 – Subtype specificity of G protein binding to GPCRs.

One of the outstanding unresolved questions about GPCR-mediated activation of G proteins is how the ~800 human GPCRs are able to signal through the appropriate combination of the 27 isoforms of the 16 G proteins. For some pairs of receptor and G protein, the specificity seems to be accounted for by highly restricted expression patterns, as in the case for the olfactory G protein, Golf. Golf is typically expressed exclusively in olfactory sensory neurons and is very similar to Gs. When the Gαolf subunit is artificially expressed in S49 cyc- cells which naturally lack the Gαs subunit, it can substitute for Gαs in order to stimulate cAMP production.[155] Thus, it is likely that if Golf was expressed widely in tissues, it would interact with many receptors that it normally would not interact with solely due to its very specific expression patter. However, for most cell types, there are at least several different of subtypes of G proteins and hundreds of different GPCRs that are expressed in the same cell. The cell maintains the correct signaling network of receptors and G proteins by using both the intrinsic affinity of specific GPCR-G protein pairs for each other and by subcellular localization processes.[156]

Our work does not offer any insight to how localization contributes to the specificity of G protein signaling *in vivo*, but it does give us a better idea about where to look for structural determinants that determine the intrinsic ability of any particular receptor to activate a G protein subtype. Many of these regions have been previously identified as important for coupling using various techniques on different receptor families,[107] but our structure now allows us to link specific receptor and G protein regions as places of contact. In order to aid in the interpretation of the role for each of the contact regions, we generated a structure-based alignment of the Gαs (PDB 1AZT), Gαi1 (PDB 1AS0), and Gαq (PDB 4EKD) proteins with 3D-Coffee,[157] shown below in figure 5-6. Also labeled on the sequence are 10 main regions of interest on the G protein.

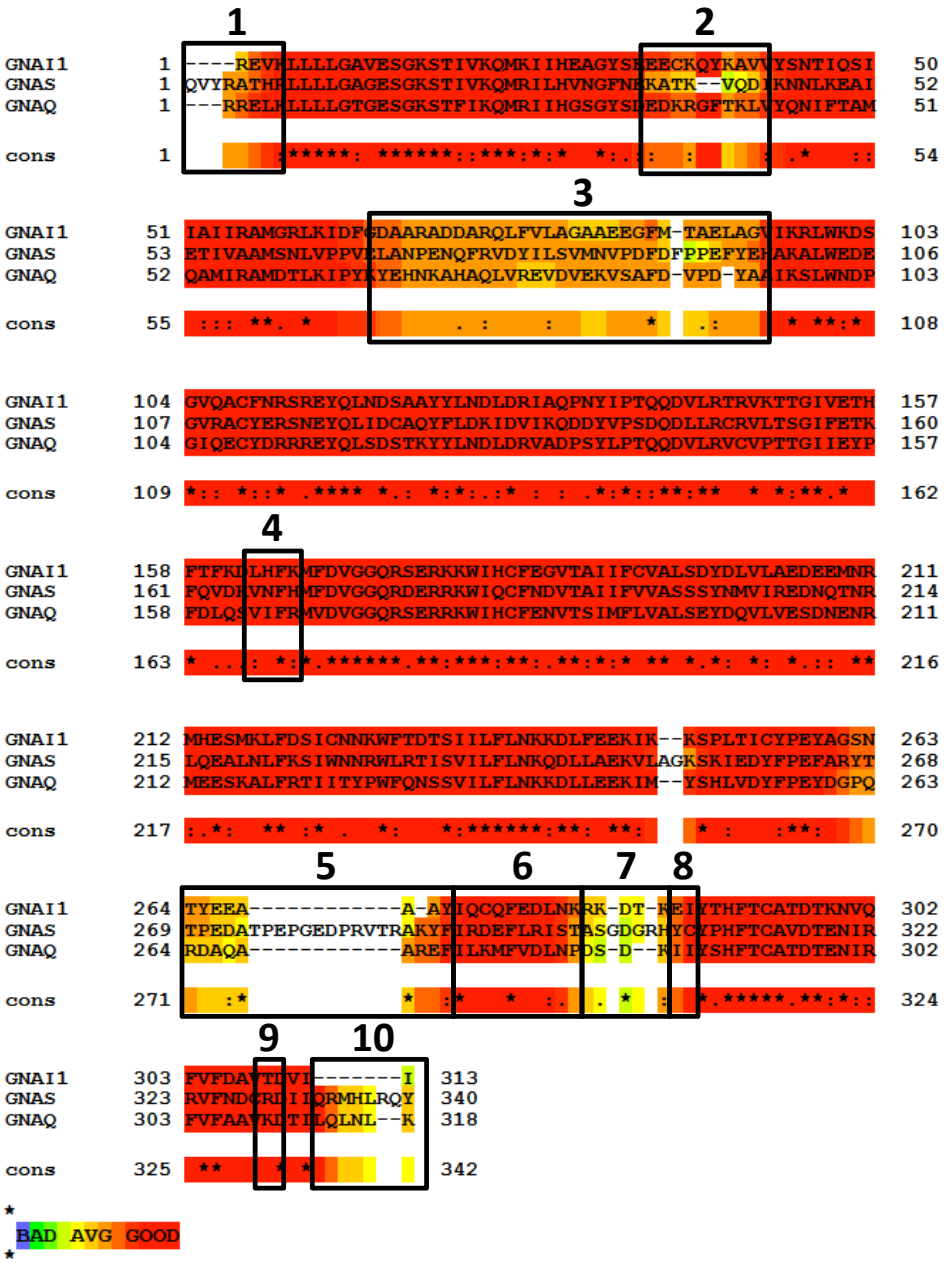


Figure 5-6. Structure based alignment of Gαs, Gαi1, and Gαq. The coloring of the residues indicates the degree of structural homology between the different proteins and the "cons" column in the alignment indicates the degree of sequence similarity.

The areas of interest are generally those on the G protein that are close the receptor in the β_2AR-Gs crystal structure. They are listed below in table 5-1, along with the general region of the receptor that that region would interact with.

Number	Region of G protein	Region of Receptor
1	Turn between N-terminal helix and β1 strand	Intercellular loop 2, F139 on β_2AR
2	1st hinge region between Ras and α-helical domain	None, but should effect region 3 due to the general mobility of the α-helical domain
3	Loop between αB and αC helices	Intercellular loop 2, also could affect region 4 of G protein
4	Turn between β2 and β3 strand	Intercellular loop 2, F139 on β_2AR
5	αG helix and associated region	Extensions of TM5 helix and proximal region of intercellular loop 3
6	α4 helix	TM5 helix
7	α4 helix and α4-β6 turn	Possibly TM6 and proximal region of intercellular loop 3
8	β6 strand	TM5 helix
9	C-terminal helix	TM5 helix, Intercellular loop 2, F139 on β_2AR
10	C-terminal helix	TM5, TM6, and TM3 helices

Table 5-1. Regions of interest on G proteins for the determination of specificity of interaction with GPCRs.

In the β_2AR-Gs crystal structure, regions 1, 4, and 9 are all involved in forming the binding pocket for F139 from the ICL2 domain on the receptor (described in section 3.6). The location of the regions can be seen in figure 5-7, below. A bulky hydrophobic residue (I, L, V, M, or F) is typically found on most GPCRs in the same position. Substitution of one bulky hydrophobic residue for another has much less of an effect on receptor-G protein coupling than substituting a small, polar, or charged residue.[114] This suggests that these regions may tend to be more important for the conserved mechanism of G protein activation than for selection between different G protein subtypes, which is also supported by the regions' involvement in the disruption of the β1-strand and P-loop interaction with nucleotide discussed in chapter 3.

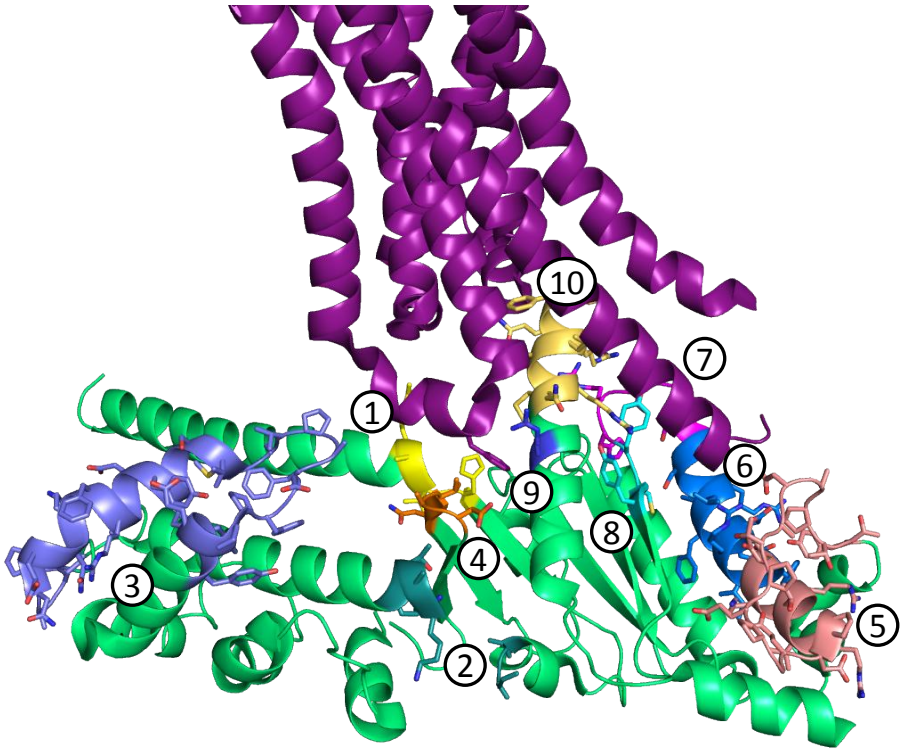


Figure 5-7. Location of domains involved in G protein signaling specificity. The domains are numbered as in table 5-1 and mapped onto the β_2AR-Gs crystal structure in varying colors, given as follows: 1 – yellow, 2 – teal, 3 – periwinkle, 4 – orange, 5 – salmon, 6 – blue, 7 – magenta, 8 – cyan, 9 – purple, 10 – tan.

Regions 2 and 3 might be involved in interactions between the α-helical domain and the IL2 region of the receptor. Region 2 is one of the hinge regions between the α-helical domain and the Ras-like domain. This is the region on Gs where the “short” and “long” isoforms differ, where the “long” isoform has 15 extra amino acids inserted at this position. The long isoform promotes the constitutive activation of the G protein, presumably by promoting the flexibility of the α-helical domain.[120] The region 3 domain contains parts of the αB and αC helices of the α-helical domain as well as the loop connecting them. The large movement of the α-helical domain in the β_2AR-Gs crystal structure brings this region up to close proximity to the ICL 2 domain of the receptor and the β2-β3 loop region of the Ras-like domain. This region on the α-helical domain is one of the more structurally divergent parts of the Gα subunits, shown below in figure 5-8.

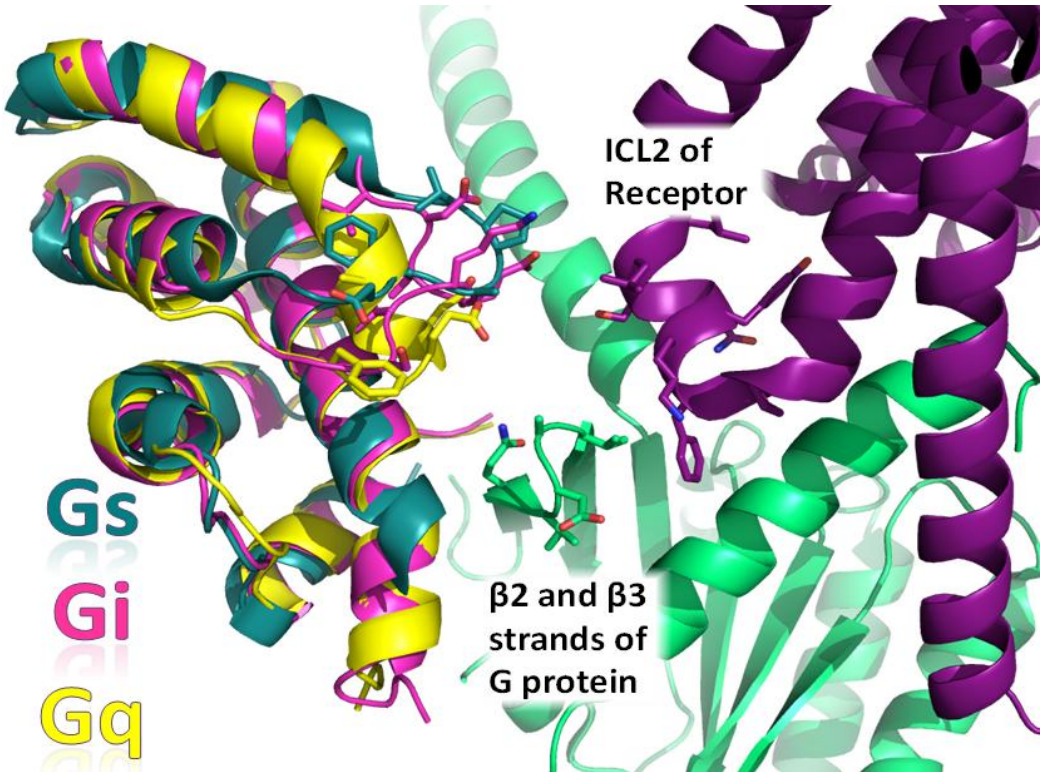


Figure 5-8. Alignment of Gα α-helical domains in a possible ICL2-interacting conformation. Gαs (PDB 3SN6) is shown in green (Ras domain) and blue (α-helical domain), Gαi1 (PDB 1AS0) α-helical domain is shown in pink, and Gαq (PDB 4EKD) α-helical domain is shown in yellow. β_2AR (PDB 3SN6) is shown in purple.

In the figure above, one can see that this variable domain on the α-helical domain is positioned quite close to the receptor, about 6-8 Å. Given the general flexibility of the α-helical domain when the G protein is nucleotide-free, one would imagine that contact could easily be made in this region when the receptor-G protein complex is in solution and not trapped in a crystal.

Regions 6, 8, and 9 all contribute to large contact surfaces between the G protein and the TM5 helix of the receptor. These regions have previously been linked with G protein specificity, especially region 6.[107] While the different G protein subtypes show very high structural conservation in this region, the sequence conservation is strongly maintained only on the side of the helix that is buried in the core protein fold. As noted in chapter 3, the surface between the two proteins in the β_2AR-Gs crystal structure does not contain any obviously strong interactions that would likely

dominate the nature of the interaction, but that certainly does not mean that other receptor-G protein pairs would not have such an interaction.

Region 5 is near the end of the TM5 helix in the β_2AR-Gs crystal structure. For the Gαs protein, there are about a dozen extra residues in this region, known as the αG helix region. These residues form a large loop structure that sits on top of the α4 helix which is not present in most other Gα subunits. Alignment of the Gαi and Gαq proteins in the same position suggests that these G protein subtypes could tolerate at least one or two more helix turns on TM5 before it would cause a steric hindrance with their smaller αG helix region. This region would also be in a prime position to interact with any residues in the TM5 proximal region of the receptor ICL3 domain.

Region 7 is structurally divergent between the G proteins studied, and is positioned right below the TM6 region of the receptor. While there are no contacts between this region and the receptor in the β_2AR-Gs crystal structure, it would be in an obvious place to look for interactions with the TM6 proximal region of the ICL3 loop on other receptor-G protein pairs. Also, if the proposed receptor-G protein “pre-coupled” complex with the receptor in a more inactive conformation does exist, it seems very likely that this region would be a likely site of interaction between the G-protein and the receptor TM6.

Region 10 is the C-terminus that forms extensive interactions between the receptor TM5, TM6, and TM3 helices and the G protein. The region has long been known to drive much, though certainly not all, of the specificity in G protein interaction and has been studied extensively. The interaction is formed by helices packed against each other in perpendicular fashion, so without structural data or high-quality homology models, it would be hard to predict exactly which residues contact each other.

The β_2AR-Gs complex structure has allowed us to determine these aforementioned contact regions and other structurally linked domains that should be involved in the process of GPCR signaling specificity. It is out of the scope of this analysis to attempt to derive rules governing why a particular receptor signals through one G

protein or another, as it is clearly a complex process that involves at least several non-adjacent domains on both proteins.

5.8 – Existence of a GDP-liganded pre-coupled state between receptor and G protein.

There have been several reports that suggest the receptor and G protein and interact even when the G protein is in a GDP-bound conformation. This GDP-bound G protein would almost certainly interact with the receptor in a different manner than what we observe in our nucleotide-free complex structure. From a physiological standpoint, this type of interaction would make a lot of sense. If the G protein was always literally right next to the receptor, it would be able to convey the activation signal more quickly because it would avoid the delay caused by diffusing to the receptor before it could interact. A relatively stable interaction between receptor and GDP-bound G proteins could also contribute to subtype-specific G protein signaling, especially for receptors with a large receptor reserve. If a particular subtype of G protein interacted with the receptor in the GDP-bound state, it would sterically hinder the interaction of the receptor with other G proteins until it was forced to leave the site following activation and GTP binding. If certain ligands promoted this interaction more so than others, it could also explain many cases of ligand-biased signaling through GPCRs without requiring that the evolutionarily conserved receptor and nucleotide-free G protein complex transmit any of the biased-ligand signal. A similar mechanism could also exist for preventing receptor binding with other types of proteins, such as arrestins and GRKs, that in general should not interact with the receptor until after an agonist binds to it. For all these reasons and more, it is sensible to entertain the idea that the GDP-bound G protein and receptor might be able to interact.

Although our data does not go far in proving or disproving such a hypothesis, we are still able to shed some light on several aspects of the issue. From the data given in sections 4-3 and 4-6, we know that if an interaction between the receptor and the GDP bound state of the G protein exists, it does not restrict access to the receptor's ligand binding site like the interaction of the nucleotide-free G protein or NB80 does.

This is suggestive that the receptor is likely to in a conformation that is most similar to its inactive state. The same conclusion was also arrived at with a photobleaching recovery study of muscarinic receptors and Gq.[117]

We also saw an indication that there might be a weak interaction between the GDP-bound heterotrimer and the receptor in our DSMS data, shown in section 3.8. When our samples of nucleotide-free receptor-G protein complex were treated with GDP, the 100 second timepoint sample indicated that the GDP was able to fully stabilize the nucleotide-binding site, but it also still showed partial stabilization of the Gα C-terminus. In contrast, treatment of the same samples with GDP and AlF_3 caused the C-terminus to be completely flexible and solvent exposed.

5.9 – Conclusion.

In the work presented in this dissertation, significant progress has been made in elucidating the mechanism for the first, key step of G protein activation by a GPCR. I have detailed the process and rationale behind the production of the G protein heterotrimer, as well as the strategy we used for making the receptor-G protein complex signaling intermediate. With a large, multi-lab collaboration, we were also able to obtain both structural and dynamic information about this complex. This has allowed us to form a plausible qualative model about the molecular nature of G protein activation that incorporates both previously known and new aspects of the process.

I have used this model to study the pharmacology of the receptor while incorporated into the receptor-G protein complex and a similar receptor-nanobody complex. Both equlibirum and kinetic data were used to create quantative models of drug and binding and receptor activation. These models are consistant with the qualitative model inferred from structural analysis, but they represent only the beginning of the process of getting a full mathematical understanding of the activation process.

There are two areas of future work that should naturally follow this effort. More structures of GPCR-G protein complexes need to be obtained, so that one can compare and contrast them to get a better idea what aspects of the activation we

have seen in this structure are general to all GPCRs and which ones are specific to our particular proteins. The methods developed here should guide and aid this effort considerably. Additionally, new biophysical assays that probe the binding and conformational changes of the G protein during activation need to be developed in order to complete the full mathematical model of the activation process. When such a model is obtained, it can be used to unambiguously predict the G protein's contribution to such complex processes like biased signaling and downregulation of GPCRs that are currently not well understood.

CHAPTER 6
MATERIALS AND METHODS

6.1 – Gs heterotrimer expression and purification.

Bovine Gαs short, His6-rat Gβ1 and bovine Gγ2 were expressed in HighFive insect cells (Invitrogen) grown in Insect Xpress serum-free media (Lonza). Cultures were grown to a density of 1.5 million cells per ml and then infected with three separate *Autographa californica* nuclear polyhedrosis viruses each containing the gene for one of the G protein subunits at a 1:1 multiplicity of infection (the viruses were a gift from A. Gilman). After 40–48 h of incubation the infected cells were harvested by centrifugation and resuspended in 75 ml lysis buffer (50 mM HEPES, pH 8.0, 65 mM NaCl, 1.1 mM $MgCl_2$, 1 mM EDTA, 1× PTT (35 μg ml^{-1} phenylmethanesulphonyl fluoride, 32 μg ml^{-1} tosyl phenylalanyl chloromethyl ketone, 32 μg ml^{-1} tosyl lysyl chloromethyl ketone), 1× LS (3.2 μg ml^{-1} leupeptin and 3.2 μg ml^{-1} soybean trypsin inhibitor), 5 mM β-mercaptoethanol (β-ME), and 10 μM GDP) per liter of culture volume. The suspension was pressurized with 600 p.s.i. N_2 for 40 min in a nitrogen cavitation bomb (Parr Instrument Company). After depressurization, the lysate was centrifuged to remove nuclei and unlysed cells, and then ultracentrifuged at 180,000g for 40 min. The pelleted membranes were resuspended in 30 ml wash buffer (50 mM HEPES, pH 8.0, 50 mM NaCl, 100 μM MgCl2, 1× PTT, 1× LS, 5 mM β-ME, 10 μM GDP) per liter culture volume using a Dounce homogenizer and centrifuged again at 180,000g for 40 min. The washed pellet was resuspended in a minimal volume of wash buffer and flash-frozen with liquid nitrogen.

The frozen membranes were thawed and diluted to a total protein concentration of 5 mg ml^{-1} with fresh wash buffer. Sodium cholate detergent was added to the

suspension at a final concentration of 1.0%. The sample was stirred on ice for 40 min, and then centrifuged at 180,000g for 40 min to remove insoluble debris. The supernatant was diluted fivefold with Ni-NTA load buffer (20 mM HEPES, pH 8.0, 363 mM NaCl, 1.25 mM $MgCl_2$, 6.25 mM imidazole, 0.2% Anzergent 3-12, 1× PTT, 1× LS, 5 mM β-ME, 10 μM GDP), taking care to add the buffer slowly to avoid dropping the cholate concentration below its critical micelle concentration too quickly. Ni-NTA resin (3 ml; Qiagen) pre-equilibrated in Ni-NTA wash buffer 1 (20 mM HEPES, pH 8.0, 300 mM NaCl, 2 mM $MgCl_2$, 5 mM imidazole, 0.2% cholate, 0.15% Anzergent 3-12, 1× PTT, 1× LS, 5 mM β-ME, 10 μM GDP) per liter culture volume was added and the sample was stirred on ice for 20 min. The resin was collected into a gravity column and washed with 4× column volumes of Ni-NTA wash buffer 1, Ni-NTA wash buffer 2 (20 mM HEPES, pH 8.0, 50 mM NaCl, 1 mM $MgCl_2$, 10 mM imidazole, 0.15% Anzergent 3-12, 0.1% DDM, 1× PTT, 1× LS, 5 mM β-ME, 10 μM GDP), and Ni-NTA wash buffer 3 (20 mM HEPES, pH 8.0, 50 mM NaCl, 1 mM $MgCl_2$, 5 mM imidazole, 0.1% DDM, 1× PTT, 1× LS, 5 mM β-ME, 10 μM GDP). The protein was eluted with Ni-NTA elution buffer (20 mM HEPES, pH 8.0, 40 mM NaCl, 1 mM $MgCl_2$, 200 mM imidazole, 0.1% DDM, 1× PTT, 1× LS, 5 mM β-ME, 10 μM GDP). Protein-containing fractions were pooled and $MnCl_2$ was added to a final concentration of 100 μM. Purified lambda protein phosphatase (5 units or about 50 μg; prepared according to method 6.3) was added per liter of culture volume and the eluate was incubated on ice with stirring for 30 min. The eluate was passed through a 0.22-μm filter and loaded directly onto a MonoQ HR 16/10 column (GE Healthcare) equilibrated in MonoQ buffer A (20 mM HEPES, pH 8.0, 50 mM NaCl, 100 μM $MgCl_2$, 0.1% DDM, 5 mM β-ME, 1× PTT). The column was washed with 150 ml buffer A at 5 ml min^{-1} and bound proteins were eluted over 350 ml with a linear gradient up to 28% MonoQ buffer B (same as buffer A except with 1 M NaCl). Fractions were collected in tubes spotted with enough GDP to make a final concentration of 10 μM. The Gs-containing fractions were concentrated to 2 ml using a stirred ultrafiltration cell (Amicon) with a 10-kDa nominal molecular weight cut-off (NWCO) regenerated cellulose membrane (Millipore). The concentrated sample was run on a Superdex 200 prep grade XK 16/70 column (GE Healthcare) equilibrated in

S200 buffer (20 mM HEPES, pH 8.0, 100 mM NaCl, 1.1 mM $MgCl_2$, 1 mM EDTA, 0.012% DDM, 100 µM TCEP, 2 µM GDP). The fractions containing pure Gs were pooled, glycerol was added to 10% final concentration, and then the protein was concentrated to at least 10 mg ml^{-1} using a 30 kDa MWCO centrifugal ultrafiltration device (Millipore). The concentrated sample was then aliquoted, flash frozen, and stored at −80 °C.

6.2 – Expression and purification of β_2-adrenergic receptor constructs.

Note: This protocol was developed by Kobilka lab, and is based primarily off of the method given in Zao et al.[144] *as well as optimizations by Søren G. F. Rasmussen*.

Recombinant baculovirus for expressing the β_2AR receptor for bimane labeling was made using the Bac-to-Bac system with the transfer vector pFastbac1 (Invitrogen). Baculovirus for expressing the T4L-β_2AR receptor constructs for crystallography was made using the Best-Bac system with the transfer pvl1393 (Expression Systems). All β_2AR constructs start with an overexpression and purification tag which begins with the HA signal sequence (MKTIIALSYIFCLVFA) followed the FLAG tag antibody binding epitope (DYKDDDDA). During overxpression, the signal sequence, which was originally derived from human influenza hemagglutinin protein, directs the protein to the plasma membrane with high efficiency and is cleaved off during the process. This leaves the FLAG-tag at the extreme N-terminus of the receptor for only the proteins that were successfully trafficked to the plasma membrane. Subsequent purification with M1 anti-FLAG antibody affinity columns selects against any receptor that did not reach the plasma membrane because it can only bind the epitope when it is displayed on the N-terminus.[158]

The β_2AR protein was expressed by infecting Sf9 cells at a density of 4 million ml^{-1} in ESF 921 Insect Cell Culture Medium (Expression Systems) with a second passage baculovirus stock using 1 ml of virus stock per 50 ml of cell culture. 1 µM alprenolol was included to enhance the receptor stability and yield. The infected cells were harvested after 48 hours of incubation at 27°C and frozen.

Cell pellets were thawed and lysed by vigorous stirring in lysis buffer (10 mM Tris-Cl pH 7.5, 2 mM EDTA, 10 ml of buffer per gram of cell pellet) supplemented with protease inhibitor leupeptin (2.5 µg/ml final concentration, Sigma) and benzamindine (160 µg/ml final concentration, Sigma) for 15 minutes. The β_2AR was extracted from the cell membrane by dounce homogenization in solubilization buffer (100 mM NaCl, 20 mM Tris-Cl, pH 7.5, 1% DDM) supplemented with leupeptin and benzamindine. 10 ml of solubilization buffer was used for each gram of cell pellet. The DDM-solubilized receptor was then purified by M1 anti-FLAG antibody affinity chromatography (Sigma). Extensive washing using HLS buffer (100 mM NaCl, 20 mM HEPES pH 7.5, 0.1%DDM) was performed to remove alprenolol. The protein was then eluted with HLS buffer containing a saturating concentration of cholesterol hemisuccinate (CHS) and supplemented with 5 mM EDTA and 200 µg ml^{-1} FLAG peptide. Fractions containing the protein were pooled, and 100 µM TCEP was added, which is a concentration that is sufficient to reduce any inter-molecular disulfide bonds without breaking the intra-molecular disulfide bridges in the receptor's extracellular loops.

The receptor was further purified by affinity chromatography using alprenolol-sepharose resin (ALP-resin). NaCl was added at a final of 350 mM, and the sample was loaded onto ALP resin (≈1 ml resin per 10 nMol of soluble receptor) at room temperature in a gravity column. The column was washed with ample HHS buffer (350 mM NaCl, 20 mM HEPES pH 7.5, 0.1%DDM) + CHS for 1 hour at room temperature, followed by HHS buffer + CHS + 2 mM $CaCl_2$ 1 hour at room temperature. The ALP column was eluted directly onto M1 anti-FLAG resin using approximately 3× the column volume of HHS buffer + CHS + 300 µM Alprenolol + 2 mM $CaCl_2$. The elution buffer was cycled through the columns for 3 hours at room temperature, and then the column was washed with HLS buffer + CHS + 2 mM $CaCl_2$ (+ ligand of choice, if desired) 1 hour at room temperature. The sample was eluted with HLS buffer + CHS (+ ligand of choice) + 5 mM EDTA + 0.2 mg/ml FLAG peptide at room temperature and collected fractions were stored on ice. The sample was dialyzed against HLS buffer (+ ligand of choice) overnight at 4 °C to remove EDTA and FLAG peptide. Optionally, PNGaseF (New England Biosciences) can be

added to the sample to remove glycosylation and/or the FLAG tag can be removed by treatment with tobacco etch virus (TEV) protease (Invitrogen) during this step.

Finally, the purified protein is concentrated using a 100 kDa molecular weight cut-off Vivaspin concentrator (Vivascience), exchanged into SEC buffer (10 mM HEPES pH 7.5, 100 mM NaCl, 0.1% DDM, + desired ligand) using size-exclusion chromatography (SEC) on a Superdex 200 HR 30/100 column (GE healthcare) equilibrated in SEC buffer. The final sample is concentrated to >5 mg/ml using a 100 kDa molecular weight cut-off Vivaspin concentrator, aliquoted, flash frozen with liquid nitrogen, and stored at -80 °C until use.

6.3 – Expression and purification of lambda protein phosphatase.

Note: This protocol was derived from work by Zuho et al.,[159] but was not optimized much. In particular, we avoided eluting the protein with 50% glycerol as suggested because we thought we could not get such a viscous buffer to work properly on our FPLC columns. However, the protein interacts extremely well with the phenyl sepharose column even with 10% glycerol buffers, possibly indicating that it is binding the phenyl groups in a similar manner as it would a phosphotyrosine residue, using the resin as a ligand-affinity column. It is suggested for any future purifications, one might investigate if a phosphotyrosine mimic like benzoic acid or phenylacetic acid might be more effective at protein elution.

Lambda phage protein phosphatase (λPPase, ORF221) was cloned from a commercial lambda DNA preparation (New England Biolabs) and heterologously expressed in *Escherichia coli* strain BL21 using the pET15b vector modified to remove the hexahistidine tag. Cell cultures in LB medium supplemented with 50 µg/ml ampicillin were grown to an optical density at 600 nm of 0.7 at 37°. Expression was induced with 400 µM IPTG and the media was supplemented with 2 mM $MnCl_2$. The cultures were incubated overnight at 23°. The following morning, the cells were harvested by centrifugation at 4500g for 15 minutes and resuspended in 30 ml lysis buffer (50 mM Tris pH 8.0, 30 mM NaCl, 2 mM EDTA, 3 mM $MnCl_2$, 1 mM $MgCl_2$, 4 µg/ml DNAse I (Roche Diagnostics), 2 mM DTT, 0.4 mM phenylmethanesulfonyl fluoride, 10% glycerol) per liter of culture volume. Cells were lysed by passage twice

through a French press at 10,000 psi. Insoluble derbies were removed by ultracentrifugation at 180,000 g for 40 minutes. NaCl was added to the supernatant to a final concentration of 500 mM and then loaded onto a Phenyl Sepharose XK 26/40 column (GE Healthcare) pre-equlibrated in buffer A (20 mM Tris pH 8.0, 500 mM NaCl, 1 mM EDTA, 2 mM $MnCl_2$, 1 mM DTT, 0.4 mM phenylmethanesulfonyl fluoride, 10% glycerol) The sample was then washed with 100 ml buffer A and bound proteins were eluted over 1200 ml with a linear gradient up to 100% Q buffer B (same as buffer A except with 30 mM NaCl). Fractions that contained λPPase were concentrated to 10 ml using using a stirred ultrafiltration cell with a 10 kDa MWCO regenerated cellulose membrane (Millipore). The concentrated protein was run in 2 ml injections on a Superdex 200 prep grade XK 16/70 column (GE Healthcare) equilibrated in S200 buffer (36 mM HEPES pH 8.0, 90 mM NaCl, 0.9 mM EDTA, 1.8 mM $MnCl_2$, 9 mM β-ME, 10% glycerol). Fractions containing the λPPase were collected and concentrated to ≈10 mg/ml, glycerol was added to a final concentration of 50%, and the sample was then aliquoted, flash frozen, and stored at -80°. Throughout the purification, the enzyme activity was assayed using the protocol given below in section 6.4.

6.4 – Lambda protein phosphatase enzyme assay.

The activity of the λ protein phosphatase was determined by measuring the hydrolysis of the substrate *para*-nitrophenylphosphate (pNPP) into inorganic phosphate and *para*-nitrophenol (pNP) in a manner similar to Zhuo et al.[159] The appearance of the intensely yellow pNP molecule was quantified by absorbance measurements at 410 nM in a 96 well plate reader. For the assay, the enzyme sample dilutions were added to the bottom of the assay plate in 2 µl of assay buffer (50 mM Tris pH 8.0, 50 mM NaCl, 2 mM $MnCl_2$, 2 mM DTT). 200 µl of assay buffer plus 20 mM pNPP was rapidly added to the samples and the plate was measured every 15 seconds for 10 minutes at room temperature. The rate of the reaction was calculated using the initial linear portion of the curve. Assuming an extinction coefficient at 410 nm for pNPP of 16.2 cm^{-1} mM^{-1} and that the definition of a unit to be µmol of pNPP hydrolyzed per minute, typical preps had a specific activities between 70 and 250 units per mg.

6.5 – Bimane-β_2AR activation assay in rHDL particles.

A β_2AR construct with 4 reactive cysteines mutated to non-reactive residues (C77V, C327S, C378A, and C406A) was expressed, purified, and labeled with mono-bromobimane on C365 according to protocols in Yao et al 2006.[76] The labeled receptor (mbb-β_2AR) was incorporated into rHDL discs according to method 6.7. Purified, concentrated Gs heterotrimer in DDM prepared method according to method 6.1 was added directly to the discs at a ratio of 10 Gs proteins to 1 receptor so that the detergent was below its critical micelle concentration and did not disrupt the discs. The samples were diluted to 100 nM mbb-β_2AR and excited with light at 370 nm with a bandpass of 4 nm. The fluorescence emission spectrum was recorded from 435 to 485 nM with a bandpass of 4 nm.

6.6 – Bimane-β_2AR activation assay in detergent

A β_2AR construct with 4 reactive cysteines mutated to non-reactive residues (C77V, C327S, C378A, and C406A) was expressed, purified, and labeled with mono-bromobimane on C365 according to protocols in Yao et al 2006.[76] Gs heterotrimer was expressed and purified according to method 6.1. Both the labeled receptor and Gs heterotrimer were solubilized in DDM detergent and the assay was run in a final detergent concentration of 0.05-0.2% plus 0.01% cholesteryl hemisuccinate (CHS) and HN buffer (20 mM HEPES pH 7.5, 100 mM NaCl). The proteins were mixed along with the indicated reagents in a BD Falcon 384-well small-volume black microplate to a total volume of 25 µl. The fluorescence spectra were read in a SpectraMax M5 plate reader (Molecular Devices) using an excitation wavelength of 370 nM and emission range from 435 to 470 nm.

6.7 – Incorporation of β_2AR into rHDL particles

To incorporate the β_2AR into rHDL particles, we followed almost exactly the method given in Whorton et al 2007[80] and Kuszak et al 2009.[160] The synthetic lipids 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC, Avanti Polar Lipids) and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) (POPG, Avanti Polar Lipids) were mixed in a 3:2 POPC:POPG ratio, dried completely, and solubilized in $HNEC_{50}$ buffer (20 mM HEPES pH 8.0, 100 mM NaCl, 1 mM EDTA, 50 mM sodium

cholate). The lipid mixture, HNE buffer (20 mM HEPES pH 8.0, 100 mM NaCl, 1 mM EDTA), ApoAI protein, and receptor were mixed together to yield final concentrations of 24 mM sodium cholate, 8 mM lipid, 100 μM ApoAI, and 5 μM receptor. The mix was incubated on ice for 1 hour. The cholate was removed by addition of 50 mg per 100 μl of mix of Bio-Beads SM-2 (Bio-Rad Laboratories) washed extensively in HNE buffer followed by incubation overnight with shaking at 4°. The Bio-Beads were removed after the incubation and the rHDL samples containing 50 μM HDL discs and 5 μM receptor were stored on ice.

6.8 – Expression and purification of Apolipoprotein A-I.

Note: This protocol is adapted from Kuszak et al. 2009.[160] Improvements to the protocol were developed with the aid of Sanna-Paula Pehkonen.

A pET15 bacterial expression plasmid was created that expressed a fusion protein with an N-terminal hexahistidine tag, TEV protease cleavage site, and residues 44-267 of the human Apolipoprotein A-I (ApoAI) gene. The plasmid was transformed into BL21(DE3) *Escherichia coli* cells and several liters of cells were grown from a single colony in Luria broth (LB) media at 37° with shaking at 250 rpm. When the cultures reached an optical density of 0.6 at 600 nm, protein expression was induced with 1 mM IPTG for 4 hours. After induction, the cells were harvested by centrifugation and resuspended in buffer A (100 mM NaH_2PO_4, 10 mM Tris, 6M GuHCl, pH 8.0). The cells were lysed by the addition of 1% final concentration Triton X-100 and treatment with a polytron homogenizer. Insoluble derbies were removed by ultracentrifugation at 180,000 g for 40 minutes, and the lysate was loaded onto 1.25 ml of Ni-NTA resin (Qiagen) per gram of per gram wet weight of pellet, pre-equlibrated in buffer a plus 1% Triton X-100. The column was washed with 10 column volumes of buffer B (100 mM NaH_2PO_4, 10 mM Tris, 6M GuHCl, pH 7.0) plus 0.2% Triton X-100 followed by 10 column volumes of buffer C (50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0) plus 0.2% Triton X-100. The protein was eluted from the column with 10 ½ column volumes of buffer D (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0) plus 0.2% Triton X-100. Elution fractions that contained significant amounts of protein were pooled together and the sample was

heated for 70° for an hour. After the heat treatment, the precipitated contaminant proteins were removed by centrifugation and the sample was concentrated to 2-4 ml using a positive pressure ultrafiltration cell (Amicon) with a 10 kDa MWCO regenerated cellulose membrane. The concentrated sample was again heated at 70° for an hour and the additional precipitated contaminant proteins were removed by centrifugation. Sodium cholate was added to a final concentration of 20 mM to the concentrated, clarified sample and it was loaded onto a Superdex 200 XK 16/70 gel filtration column (GE Healthcare) equilibrated in $HNEC_{20}$ buffer (20 mM HEPES pH 8.0, 100 mM NaCl, 1 mM EDTA, 20 mM sodium cholate). The ApoAI eluted from the column as a dimer and was well separated from the Triton X-100 peak. The fractions containing ApoAI were pooled and dialyzed 1× against $HNEC_{20}$ buffer and then twice against $HNEC_5$ buffer (same as $HNEC_{20}$, but with 5 mM sodium cholate). After dialysis, the sample was concentrated to 10-15 mg/ml, snap frozen in liquid nitrogen, and stored at -80°.

6.9 – Initial receptor and G protein complex coupling and purification method.

For the initial preparation of milligram quantities of β_2AR and Gs complex, we used the following protocol. It was slowly changed and refined over several years in order to arrive at the final complexing reaction given in method 6.12. β_2AR (prepared according to method 6.2), Gs heterotrimer (prepard according to method 6.1), and buffer were added together to make a solution with final concentrations of 50 µM β_2AR, 25 µM Gs, 20 mM HEPES pH 7.5, 100 mM NaCl, 1 mM EDTA, 2 mM MgCl2, 0.1% DDM, and 0.01% CHS. 100 µM agonist and 25 mU/ml apyrase (purchased from NEB) were added and the sample was incubated overnight at 18°.

After incubation the receptor-G protein complex was diluted 10× with IEC Buffer A (20 mM HEPES pH 7.2, 20 mM NaCl, 1 mM EDTA, 1.1 mM MgCl2, 0.05% DDM) and loaded onto a MonoQ HR 5/5 column. The sample was eluted with a linear gradient from 0% to 25% IEC Buffer B (same as IEC Buffer A except with 1000 mM NaCl) at a flow rate of 0.6 ml/min. Fractions with the purified receptor-G protein complex were collected and concentrated to at least $10\,mg\,ml^{-1}$ using a 30 kDa MWCO centrifugal ultrafiltration device (Millipore).

6.10 – Fluorescamine labeling for SDS-PAGE analysis.

For labeling protein with fluorescamine, the sample was put into HEPES or Tris buffer at pH 8.0 and incubated with 1% SDS for 5 minutes at room temperature. An equal volume of freshly prepared 0.5 mg/ml fluorescamine dissolved in acetone was added to the protein sample and it was immediately mixed by vortexing. The sample was allowed to react for another 5 minutes before proceeding with SDS-PAGE analysis. To detect the fluorescent bands in the gel, a UV transilluminator was used for excitation and no cutoff filter was used for the detection of emitted blue light. Bands were quantified using the ImageJ software package provided by the NIH.

6.11 – Nanobody generation against the receptor and G protein complex.

Two llamas were immunized (*Lama glama*) with the bis(sulphosuccinimidyl)glutarate (BS2G, Pierce) cross-linked β_2AR-Gs-BI-167107 ternary complex. Peripheral blood lymphocytes were isolated from the immunized animals to extract total RNA, prepare cDNA, and construct a nanobody phage display library according to published methods.[90] Nb35 and Nb37 were enriched by two rounds of biopanning on the β_2AR-Gs-BI-167107 ternary complex embedded in biotinylated high-density lipoprotein particles.[80] Nb35 and Nb37 were selected for further characterization because they bind the β_2AR-Gs-BI-167107 ternary complex but not the free receptor in an ELISA assay.

6.12 – Receptor and G protein complex coupling and purification for crystallography.

Approximately 100 μM Gs heterotrimer and 130 μM BI-167107-bound T4L–β_2AR was mixed together in incubation buffer (10 mM HEPES, pH 7.5, 100 mM NaCl, 0.1% DDM, 1 mM EDTA, 3 mM $MgCl_2$, 10 μM BI-167107) and incubated for 3 hourr at room temperature. After the first 90 minutes of the incubation, apyrase (25 mU ml^{-1}, NEB) was added to hydrolyse residual GDP released from Gαs upon binding to the receptor. The sample was the exchanged into the detergent MNG-3 by adding the β_2AR-Gs mixture to 4× the sample volume of MNG buffer (20 mM HEPES, pH 7.5, 100 mM NaCl, 10 μM BI-167107, 1% MNG-3) and further incubating for 1 hour at room temperature. To complete the detergent exchange, the β_2AR-Gs

complex was immobilized on M1 FLAG resin and washed in MNG-FLAG buffer (20 mM HEPES, pH 7.5, 100 mM NaCl, 10 μM BI-167107, and 3 mM $CaCl_2$, 0.2% MNG-3). To prevent cysteine bridge-mediated aggregation of β2AR-Gs complexes, 100 μM TCEP was added to the eluted protein before concentrating it with a 50 kDa MWCO Millipore concentrator. The final size exclusion chromatography procedure to separate excess free receptor from the β2AR-Gs complex was performed on a Superdex 200 10/300 GL column (GE Healthcare) equilibrated with buffer containing 0.02% MNG-3, 10 mM HEPES, pH 7.5, 100 mM NaCl, 10 μM BI-167107 and 100 μM TCEP. Peak fractions were pooled and concentrated to approximately 90 mg ml−1 with a 100 kDa MWCO Viva-spin concentrator.

6.13 – NB35 stabilization and lipidic cubic phase based crystallography of receptor-G protein complex.

BI-167107 bound T4L-β2AR-Gs complex and Nb35 were mixed in 1:1.2 molar ratio. The mixture was incubated for 1 hr at room temperature before mixing with 7.7 MAG (1-(7Z-tetradecenoyl)-rac-glycerol) containing 10% cholesterol in a 1:1 protein solution to lipid ratio (w/w) using the twin-syringe mixing method reported previously.[161] The concentration of T4L-β2AR-Gs-Nb35 complex in 7.7 MAG was approximately 25 mg ml^{-1}. The protein-lipid mixture was delivered through an LCP dispensing robot (Gryphon, Art Robbins Instruments) in 40 nl drops to either 24-well or 96-well glass sandwich plates and overlaid en-bloc with 0.8 μl precipitant solution. Multiple crystallization leads were initially identified using in-house screens partly based on reagents from the StockOptions Salt kit (Hampton Research). Crystals for data collection were grown in 18 to 22% PEG 400, 100 mM MES, pH 6.5, 350 to 450 mM potassium nitrate, 10 mM foscarnet, 1 mM TCEP, and 10 μM BI-167107. Crystals reached full size within 3–4 days at 20 °C and were picked from a sponge-like mesophase and flash-frozen in liquid nitrogen without additional cryoprotectant.

6.14 – Crystallographic data collection and model solving.

The structure was solved by molecular replacement using Phaser.[162, 163] The order of the molecular replacement search was found to be critical in solving the structure. In order, the search models used were: the β and γ subunits from a Gi heterotrimer

(PDB ID: 1GP2), the Gs α Ras-like domain (PDB ID: 1AZT), the active-state β_2AR (PDB ID: 3P0G), a β_2AR binding nanobody (PDB ID: 3P0G), T4 lysozyme (PDB ID: 2RH1), and the Gs α-helical domain (PDB ID: 1AZT). Regions that were expected to differ in the complex structure compared to the reference models were removed from the models and manually re-built during refinement. Such regions included the complementarity determining regions of the nanobody, several residues in the hinge region between the Gα Ras-like and α-helical domains as well as the N-terminal helix, the C-terminal region of the Gγ subunit, and residues on the receptor intercellular loop 3 (IL3) that were in contact with the lysozyme fusion on that structure. Following the determination of the initial structure by molecular replacement, rigid body refinement and simulated annealing were performed in Phenix[164] and BUSTER,[165] followed by restrained refinement and manual rebuilding in Coot.[166] After iterative refinement and manual adjustments, the structure was refined in CNS[167] using the deformable elastic network (DEN) method.[168] Although the resolution of this structure exceeds that for which DEN is typically most useful, the presence of several poorly resolved regions indicated that the incorporation of additional information to guide refinement could provide better results. The DEN reference models used were those used for molecular replacement, with the exception of NB35, which was well ordered and for which no higher resolution structure is available.

6.15 – Analysis of the receptor-G protein complex with deuterium exchange by mass spectrometry analysis.

1.5 ml of β_2AR-Gs complex (10 mg ml^{-1}) or 1.5 ml of Gs heterotrimer (7 mg ml^{-1}) was mixed with 4.5 ml of D_2O buffer (20mM HEPES, pH 7.5, 100mM NaCl, 10 mM BI-167107, 100 mM TCEP, 0.0015% MNG-3 in D_2O) and incubated for 100, 1,000 and 10,000 seconds on ice. At the indicated times, the sample was quenched by 15 ml of ice-cold quench solution (0.1M NaH_2PO_4, 20mM TCEP, 16.6% glycerol, pH 2.4), immediately frozen on dry ice, and stored at -80 °C. Non-deuterated control was prepared in H_2O buffer (20mM HEPES, pH 7.5, 100mM NaCl, 10 mM BI-167107, 100mM TCEP, 0.0015% MNG-3 in H2O), mixed with quench solution, and snap-frozen on dry ice. Samples were thawed and immediately passed through an

immobilized porcine pepsin column (16 ml bed volume) at a flow rate of 20 ml min^{-1} of 0.05% trifluoroacetic acid. Peptide fragments were collected on a C18 trap column for desalting and separated by a Magic C18AQ column (Michrom BioResources Inc.) using a linear gradient of acetonitrile from 6.4% to 38.4% over 30 min. Mass spectrometric analysis was performed using LCQ Classic mass spectrometer (Thermo Finnigan), with capillary temperature of 200 °C. Deuterium quantification data were collected in MS1 profile mode, and peptide identification data were collected in data dependent MS/MS mode. Recovered peptide identification and analysis were carried out using DXMS Explorer (Sierra Analytics Inc.), specialized software for processing DXMS data.[169]

6.16 – Analysis of the receptor-G protein complex by negative-stain electron microscopy.

Samples of T4L-β_2AR-Gs complex prepared for crystallography (method given in section 6.12) were also used for electron microcroscopy. For experiments where the complex was labeled with nanobody, the T4L-β_2AR-Gs complex was incubated for 15 min at room temperature with approximately equimolar concentrations of Nb35 or Nb37 and subsequently prepared by negative staining. For samples that were treated with nucleotide (GDP or GTPγS) or nucleotide mimicking fragments (PPi or foscarnet), these compounds were rapidly mixed with the complex and the sample was immediately fixed by negative stain.

For staining and visualization, several microliters of a sample solution was adsorbed to a glow-discharged carbon-coated copper grid, washed with two drops of deionized water, and stained with two drops of freshly prepared 0.75% uranyl formate. Dried grids were imaged at room temperature with a Tecnai T12 electron microscope operated at 120 kV using low-dose procedures. Images were recorded at a magnification of 71,138× and a defocus value of ~1.5 μm on a Gatan US4000 CCD camera. All images were binned (2 × 2 pixels) to obtain a pixel size of 4.16 Å on the specimen level. Tilt-pair particles from 60° and 0° images were selected using WEB.[170] Particles for only 2D classification of 0° projections were excised using Boxer (part of the EMAN 1.9 software suite).[171]

6.17 – Image processing and low-resolution 3D reconstruction for electron microscopy.

The 2D reference-free alignment and classification of particle projections were performed using SPIDER.[170] For all conditions, the 0° particle projections were iteratively classified into multiple classes for 10 cycles. For AH conformation assignments,.we used the first classification to select only the particles from averages clearly displaying the profiles of Ras-like, Gβγ, β2AR, and T4L domain densities in the same position, thereby restricting the range of particle projection orientations. These projections were pulled together and subjected to a second iterative classification (referred to as the "secondary" classification). For counting the numbers of particles with and without stabilized AH domain on the Ras-like domain, three different operators examined each secondary classification and assigned each class average according to the projection profile of the specific region. The assignment from the different operators was in good agreement, and the particle numbers belonging to individual classes were added to calculate percentages for each conformation. Assignments for each full individual dataset were done in addition to the secondary classification, and the results were in agreement.

To test any bias, the particles from nucleotide-free, 1 mM PPi, 10 mM PPi, and foscarnet conditions were combined into a single dataset of 15,753 particles and were classified into 200 classes. The individual class averages were assigned as before according to the visibility of the AH domain, and the percentage of projections from each condition was determined according to the number of projections contributing to the assigned class averages. The results of this "blind" test showed very good agreement with our assignments from individual classifications. For 3D reconstructions, in a first step we used the random conical tilt technique[172] to determine initial 3D maps by back-projection of tilted particle images belonging to individual classes. After a first round of angular refinement, corresponding particles from the images of the untilted specimen were added, and the images were subjected to another cycle of refinement. We thus generated reliable initial models for complexes with variability in the positioning of the AH domain of Gαs. After contrast transfer function (CTF) correction according to local defocus values

obtained by CTFTILT,[173] the full dataset from each condition was subjected to multiple reference-supervised alignment [174, 175] with the "multirefine" routine in EMAN (1.9) by using our initial models as reference maps. This approach allowed us to separate particles from the entire dataset (of each condition) according to the positioning of the AH domain of Gαs. For final maps, we used the separated datasets, as provided by the multiple reference-supervised alignment, and used FREALIGN[176] for further refinement of the orientation parameters and reconstruction. The resolution for each map was determined at FSC = 0.5.

6.18 – Preparation of receptor-G protein complex in rHDL particles.

Purified Gs heterotrimer was added to the pre-formed β_2AR-rHDL particles (method 6.7), incubated for 2 hours at 4°, and BioBeads were added to remove the excess detergent. Nucleotide-free Gs-β_2AR complex was prepared by incubating β_2AR-Gs-rHDL particles with apyrase in the presence of 1 mM $MgCl_2$ for 30 minutes at room temperature, or alternately, 2 hours at 4°. If needed, the sample was passed through a Superdex 200 HR 30/100 gel filtration column to remove free nucleotide and apyrase.

6.19 – Reconstitution of β_2AR and Gs in vesicles.

Lipid stock mixtures of DOPC (3 mg/ml) and CHS (0.3 mg/ml) were prepared in 20 mM HEPES pH 7.5, 100 mM NaCl including 1% β-OG, as previously described.[177, 178] Breifly, the lipid stocks were removed from storage, vortexed, and sonicated for 30 min in an ice bath. The reconstitution mixture was prepared in 20 mM HEPES pH 7.5, 100 mM NaCl + 0.1% DDM containing a 10-fold dilution of the lipid stock and β_2AR +/- Gs. The final DOPC and CHS concentrations in this mixture were 0.3 and 0.03 mg/ml, respectively. The reconstitution mixture was inverted several times and incubated for 2 h on ice. Detergent removal and formation of vesicles were attained by gel filtration chromatography on a Sephadex G-50 Fine column.[177, 178]

6.20 – Preparation of β_2AR and Gs containing membranes from Sf9 cells.

Sf9 cells growing in mid-log phase were infected at a multiplicity of infection of 0.5-1 with recombinant baculoviruses containing expression constructs for FLAG-His_{10}-mECFP-β_2AR (CBAR) and Gs subunits. 48–60 hours later, the cells were pelleted

by centrifugation for 10 min at 500 × g. They were resuspended in 1/10 the original culture volume of TBS buffer (25 mM Tris-Cl pH 7.4, 137 mM NaCl, and 3 mM KCl) + 1× PTT (35 μg ml^{-1} phenylmethanesulphonyl fluoride, 32 μg ml^{-1} tosyl phenylalanyl chloromethyl ketone, 32 μg ml^{-1} tosyl lysyl chloromethyl ketone) and 1× LS (3.2 μg ml^{-1} leupeptin and 3.2 μg ml^{-1} soybean trypsin inhibitor). They were lysed by sonication and centrifuged at 500 × g for 10 min to pellet unlysed cells and nuclei. The supernatant was centrifuged at 35 min at 100,000 × g to pellet the membranes. The membranes were resuspended in 1/20 the original culture volume with low salt buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1× LS, 1× PTT), flash frozen with liquid nitrogen, and stored at −80°C until use.

6.21 – Preparation of β_2AR and Gs containing membranes from rat lung.

Lungs from freshly sacrificed rats (*Rattus norvegicus*) was homogenized in TBS buffer (25 mM Tris-Cl pH 7.4, 137 mM NaCl, and 3 mM KCl) + 1× PTT (35 μg ml^{-1} phenylmethanesulphonyl fluoride, 32 μg ml^{-1} tosyl phenylalanyl chloromethyl ketone, 32 μg ml^{-1} tosyl lysyl chloromethyl ketone) and 1× LS (3.2 μg ml^{-1} leupeptin and 3.2 μg ml^{-1} soybean trypsin inhibitor) using a Tissue-Tearor homogenizer (Biospec). The sample was centrifuged at 500 × g for 10 min to pellet large debris and the supernatant was centrifuged again at 35 min at 100,000 × g to pellet the membranes. The membranes were resuspended in 1/20 the original culture volume with low salt buffer (50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1× LS, 1× PTT), flash frozen with liquid nitrogen, and stored at −80°C until use.

6.22 – Saturation binding of radioligand to β_2AR in rHDL particles.

Saturation binding assays were performed with [^{3}H]-dihydroalprenolol ([^{3}H]DHAP). Assays were carried-out in a volume of 100 μl in a 96-well format. Samples were incubated for 90 minutes (standard time) or longer (as indicated) at room temperature with or without GTPγS and the specified concentrations of [^{3}H]DHAP in TBS buffer (25 mM Tris-Cl pH 7.4, 137 mM NaCl, and 3 mM KCl). Reactions were terminated and free [^{3}H]DHAP was removed by rapid filtration through Whatman GF/B filters pre-treated with 0.3% polyethylamine in TBS and washed with 3× 200 μl ice-cold TBS buffer. Non-specific binding was determined in the presence of 10 μM

(+/-)-propranolol. Data were fit to a one-site binding model by using Prism 5.0 (GraphPad) to determine K_d and B_{max}. For single point binding assays, a concentration of 2 nM [^{3}H]DHAP was used and the points were measured in triplicate.

6.23 – Competition binding of radioligand with β_2AR in rHDL particles.

Competition experiments were performed with 2 nM [^{3}H]DHAP as the radiolabel tracer. Assays were run at a concentration of 200 µl in ME-TK8 buffer (2 mM $MgCl_2$, 1 mM EDTA, 20 mM Tris-Cl pH 8.0, and 135 mM KCl) + 0.02 % sodium ascorbate + 100 µM TCEP + 0.05% BSA. Increasing concentrations of the competing agonist, (-)-epinephrine (EPI) were added along with the specified amount of NB80. Samples were incubated for 6 hours at room temperature and filtered through Whatman GF/B filters pre-treated with 0.3% polyethylamine in TBS and washed with 3× 200 µl ice-cold TBS buffer. Data were fit to a one or two-site competition binding model with Prism, using the Cheng-Prusoff correction for estimation of K_i values (assuming the K_d of [^{3}H]DHAP = 0.5 nM).

6.24 – Kinetic binding of radioligand to β_2AR in rHDL with NB80.

For association binding experiments with [^{3}H]DHAP and NB80, β_2AR-rHDL (20 fmol per timepoint, 20 pM final concentration) was pre-incubated with varying concentrations of NB80 for 30 minutes at room temperature in TBS buffer (25 mM Tris-Cl pH 7.4, 137 mM NaCl, and 3 mM KCl) + 0.05% BSA. 5 nM [^{3}H]DHAP was then added and timepoints were taken by filtering the samples on Whatman GF/B filters pre-treated with 0.3% polyethylamine and washed with 3× 200 µl ice-cold TBS buffer. Dissociation experiments were done in a similar manner, except that first the samples were incubated with 5 nM [^{3}H]DHAP for 30 minutes, then incubated with varying NB80 concentrations for 30 minutes. After the incubations, a zero timepoint was taken and the dissociation experiment was started by adding 50 µM unlabled alprenolol.

6.25 – Kinetic binding of radioligand to β_2AR-Gs complex in rHDL with guanine nucleotides.

For association binding experiments with [^{3}H]DHAP and β_2AR-Gs nucleotide-free complexes, a gel-filtered sample of apyrase-treated β_2AR-Gs-rHDL particles was incubated with 5 nM [^{3}H]DHAP in TBS buffer (25 mM Tris-Cl pH 7.4, 137 mM NaCl, and 3 mM KCl) + 0.05% BSA for 30 minutes to bind any receptor that was not complexed with Gs. After the incubation, a zero timepoint was taken and the association experiment was started immediately afterward by adding varying amounts of either GDP or GTPγS. Timepoints were taken by filtering the samples on Whatman GF/B filters pre-treated with 0.3% polyethylamine and washed with 3× 200 µl ice-cold TBS buffer.

6.26 – Preperation of biotinyated apolipoprotein A-I.

Purified apolipoprotein A-I in $HNEC_5$ buffer (20 mM HEPES pH 8.0, 100 mM NaCl, 1 mM EDTA, 5 mM sodium cholate) was biotinylated via amine coupling with N-hydroxysuccinimide-biotin (NHS-biotin). A 10× molar excess of dry NHS-biotin was weighed and dissolved in a volume of dry DMSO equal to that of the ApoAI sample. The freshly prepared NHS-biotin in DMSO and the apoAI were mixed, allowed to react for an hour at room temperature, and then quenched with ethanolamine. The sample was gel filtered on a Superdex 200 HR10/30 column in $HNEC_{20}$ buffer (20 mM HEPES pH 8.0, 100 mM NaCl, 1 mM EDTA, 20 mM sodium cholate) to remove free biotin and DMSO. Fractions containing apoAI protein were collected, dialized 2x against $HNEC_5$ buffer, and concentrated to a protein concentration equal to or greater than 10 mg/ml.

6.27 – Kinetic binding of NB80 to β_2AR in rHDL.

Samples of β_2AR were prepared in biotinylated rHDL discs following the method given in section 6.7, except that the apoA-I used was biotinylated according to the method given in section 6.26. The biotinylated discs were loaded onto biosensor tips (FortéBio) by incubation of the tip with the stock solution of discs for 5-10 minutes at room temperature. The tips were placed on the Blitz machine (FortéBio) and washed in excess buffer that included any drugs that bound to the receptor. The NB80

association reaction was started by incubating the appropriate concentration of nanobody with the tip while observing the interferometry signal. The dissociation reaction was started by incubation of the NB80 bound tip into an excess of non-nanobody containing buffer while observing the interferometry signal.

6.28 – Global optimization of NB80 binding model.

An object-oriented program was created in Python to run the optimization. The program is organized around the requirements of fitting a globally shared theoretical model to the data from each individual experiment. The experimental data is supplied to the program in a multi-dimensional array that is attached to lists of experimental conditions and independent variable values as metadata along each array axis.[179] The program is also given an object that contains the theoretical model that is used to simulate the different states of the system and other objects that provide methods to convert the collection of predicted states for each type of experiment into simulated data. The simulated and real data are subtracted to calculate residuals for the optimization routine to minimize.

The general structure of the optimization software divides the variables needed for data fitting into two categories. For each experiment, there is a set of variables that is specific to that particular experiment, such as the level of non-specific binding, the maximal amount of ligand binding, etc. These variables are called “level 2” variables and are optimized for each experiment separately. Level 2 variables are not interesting from a theoretical standpoint, but they must still be estimated in order to successfully fit the data. There is also a single set of “level 1” variables, which are shared between all experiments and describe the behavior of the underlying physical model used to simulate the data. In this case, the level 1 variables are the 24 rate constants that describe the model shown in figure 4-9. To perform optimization of the level 2 variables, of which there were typically only a few per experiment, we used a simple least-squares optimization routine based on the Newton-Raphson method supplied by the software package Scipy.[180]

The global optimization of the level 1 variables was more complicated. One of the challenges of doing a global optimization with both the radioligand binding and Blitz data sets was that each type of data gave very different values for the sum of squares (ssq) on the calculated residuals, and each experiment also contained variable numbers of data points. In order to correct for this and make each experiment contribute an appropriate fraction of the total ssq value, we used two separate normalizations for each experiment. The first normalization divided the calculated ssq for each experiment by the minimum possible ssq for that dataset. For radioligand binding data, this was calculated by setting the fitted data as the mean of replicates for each data point. For Blitz data, we calculated the minimal ssq value using smoothed curve generated with a LOWESS routine provided by the Biopython software package.[181] This normalized the raw ssq values to a number that represented the goodness of fit as a fold-over-minimum for each experiment which could, in theory, be compared with each other. However, some experiments contained more information about the system than others, and we wanted these to contribute proportionally more to the final goodness-of-fit value in the optimization. We did this by multiplying the fold-over-minimum value obtained with the first normalization by the number of "curves" that each experiment contained. For example, a competition assay run in the presence of 6 different NB80 concentrations would typically be graphed as 6 different curves, one for each NB80 condition, so its fold-over-minimum value was multiplied by 6. In contrast, a typical Blitz curve only gave information about a single experimental condition, so its fold-over-minimum value was multiplied by 1. In this way, we were able to combine multiple experiments with diverse types of data into a single goodness-of-fit parameter for use in an optimization routine.

Initial estimates for the level 1 rate constant variables were determined by analysis of the individual experiments and knowledge of several equilibrium constants. The estimates were then used as a starting point for optimization of the level 1 values using the simulated annealing optimization routine provided by Scipy. The range that the values were allowed to change over for each optimization run was small, about 1/3 of a base 10 logarithm unit in each direction. However, the routine was run many

times, with periodic re-optimizations of the level 2 parameters in-between the level 1 optimization runs.

6.29 – Simulation of equilibrium binding problems.

The model we used to for fitting the data was composed solely of elementary steps and all data was taken under pseudo-first order conditions, so this allowed us to solve any equilibrium binding problems as an eigenvector problem. What we are seeking are the concentrations of each state in the vector **S** that gives us a net rate of change equal to zero for all states, consistent with the definition of chemical equilibrium. Mathematically, we are searching for the eigenvector corresponding to the eigenvalue = $\lambda = 0$ for this equation where **M** = the model matrix and **S** = a column vector of the model states:

$$\mathbf{MS} = \lambda\mathbf{S}$$

Equation 6-1. Eigenvalue and eigenvector equation that must be solved to determine the equilibrium state of the model.

The solutions to this equation are calculated numerically for each experimental condition using the linalg package given in Scipy, which uses the Fortran-based LAPACK library. Theoretically, the equilibrium solution to the problem should be the eigenvector associated with the eigenvalue $\lambda = 0$, but in practice we determine the desired eigenvector differently. Because of the numerical nature of the calculation, it can sometimes be difficult to determine the difference between a small eigenvalue and one that is truly equal to zero. In some cases, it is also possible that the matrix has multiple eigenvectors for the same eigenvalue, a situation that is also made worse by numerical instability in the returned eigenvectors. To overcome these issues, we instead had the program simply sort through the entire set of returned eigenvectors, looking for the one that implied exclusively positive or zero values for all concentrations of the vector **S**.

6.30 – Simulation of kinetic binding problems.

The simulation of kinetic data was performed by numerical integration of the rate equations using the odeint function supplied in Scipy, which is based on the fortran

library ODEPACK. The initial concentrations of each state were calculated by solving the system for equilibrium with the experimental conditions given. The conditions were then changed to their values at the beginning of the experiment and the predicted concentrations of the model states were determined by odeint to a relative tolerance of 1×10^{-6}.

BIBLIOGRAPHY

1. Rosenbaum, D.M., S.G. Rasmussen, and B.K. Kobilka, *The structure and function of G-protein-coupled receptors.* Nature, 2009. **459**(7245): p. 356-63.

2. Overington, J.P., B. Al-Lazikani, and A.L. Hopkins, *How many drug targets are there?* Nat Rev Drug Discov, 2006. **5**(12): p. 993-6.

3. Maehle, A.H., *"Receptive substances": John Newport Langley (1852-1925) and his path to a receptor theory of drug action.* Medical history, 2004. **48**(2): p. 153-74.

4. Limbird, L.E., *Cell surface receptors : a short course on theory & methods*. 2005, Springer: New York. p. 1-28.

5. Lefkowitz, R.J., *Historical review: a brief history and personal retrospective of seven-transmembrane receptors.* Trends in pharmacological sciences, 2004. **25**(8): p. 413-22.

6. Ehrlich, P., *The Nobel Lectures in Immunology - Lecture for the Nobel-Prize for Physiology or Medicine, 1908 - Partial Cell Functions.* Scandinavian Journal of Immunology, 1990. **31**(1): p. 1-13.

7. Langley, J.N., *On the contraction of muscle, chiefly in relation to the presence of receptive substances. Part iv The effect of curari and of some other substances on the nicotine response of the sartorius and gastrocnemius muscles of the frog.* Journal of Physiology-London, 1909. **39**(4): p. 235-295.

8. Clark, A.J., *The reaction between acetyl choline and muscle cells: Part II.* J Physiol, 1927. **64**(2): p. 123-43.

9. Clark, A.J., *The reaction between acetyl choline and muscle cells.* J Physiol, 1926. **61**(4): p. 530-46.

10. Langmuir, I., *THE CONSTITUTION AND FUNDAMENTAL PROPERTIES OF SOLIDS AND LIQUIDS. PART I. SOLIDS.* Journal of the American Chemical Society, 1916. **38**(11): p. 2221-2295.

11. Hill, A.V., *The mode of action of nicotine and curari, determined by the form of the contraction curve and the method of temperature coefficients.* The Journal of physiology, 1909. **39**(5): p. 361-73.

12. Gaddum, J.H., *The quantitative effects of antagonistic drugs.* J Physiol, 1937. **89**: p. 7P-9P.

13. Ariens, E.J., *Affinity and intrinsic activity in the theory of competitive inhibition. I. Problems and theory.* Arch Int Pharmacodyn Ther, 1954. **99**(1): p. 32-49.

14. Stephenson, R.P., *A modification of receptor theory.* Br J Pharmacol Chemother, 1956. **11**(4): p. 379-93.

15. Furchgott, R.F., *The pharmacology of vascular smooth muscle.* Pharmacol Rev, 1955. **7**(2): p. 183-265.

16. Ruffolo, R.R., Jr., *Review important concepts of receptor theory.* J Auton Pharmacol, 1982. **2**(4): p. 277-95.

17. Hill, A.V., *The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves.* J Physiol, 1909. **40 (Suppl)**: p. iv–vii.

18. Monod, J., J.P. Changeux, and F. Jacob, *Allosteric proteins and cellular control systems.* Journal of molecular biology, 1963. **6**: p. 306-29.

19. Monod, J., J. Wyman, and J.P. Changeux, *On the Nature of Allosteric Transitions: A Plausible Model.* Journal of molecular biology, 1965. **12**: p. 88-118.

20. Del Castillo, J. and B. Katz, *Interaction at end-plate receptors between different choline derivatives.* Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society, 1957. **146**(924): p. 369-81.

21. Karlin, A., *On the application of "a plausible model" of allosteric proteins to the receptor for acetylcholine.* Journal of theoretical biology, 1967. **16**(2): p. 306-20.

22. Berthet, J., T.W. Rall, and E.W. Sutherland, *The relationship of epinephrine and glucagon to liver phosphorylase. IV. Effect of epinephrine and glucagon on the reactivation of phosphorylase in liver homogenates.* The Journal of biological chemistry, 1957. **224**(1): p. 463-75.

23. Sutherland, E.W. and T.W. Rall, *Fractionation and characterization of a cyclic adenine ribonucleotide formed by tissue particles.* The Journal of biological chemistry, 1958. **232**(2): p. 1077-91.

24. Birnbaumer, L. and M. Rodbell, *Adenyl cyclase in fat cells. II. Hormone receptors.* The Journal of biological chemistry, 1969. **244**(13): p. 3477-82.

25. Rodbell, M., L. Birnbaumer, S.L. Pohl, and H.M. Krans, *The glucagon-sensitive adenyl cyclase system in plasma membranes of rat liver. V. An obligatory role of guanylnucleotides in glucagon action.* The Journal of biological chemistry, 1971. **246**(6): p. 1877-82.

26. Black, J., *Nobel lecture in physiology or medicine--1988. Drugs from emasculated hormones: the principle of syntopic antagonism.* In vitro cellular & developmental biology : journal of the Tissue Culture Association, 1989. **25**(4): p. 311-20.

27. Brown, E.M. and G.D. Aurbach, *Beta-Adrenergic receptor interactions. Characterization of iodohydroxybenzylpindolol as a specific ligand.* The Journal of biological chemistry, 1976. **251**(5): p. 1232-8.

28. Bourne, H.R., P. Coffino, and G.M. Tomkins, *Selection of a variant lymphoma cell deficient in adenylate cyclase.* Science, 1975. **187**(4178): p. 750-2.

29. Insel, P.A., M.E. Maguire, A.G. Gilman, H.R. Bourne, P. Coffino, and K.L. Melmon, *Beta adrenergic receptors and adenylate cyclase: products of separate genes?* Molecular pharmacology, 1976. **12**(6): p. 1062-9.

30. Ross, E.M. and A.G. Gilman, *Reconstitution of catecholamine-sensitive adenylate cyclase activity: interactions of solubilized components with receptor-replete membranes.* Proceedings of the National Academy of Sciences of the United States of America, 1977. **74**(9): p. 3715-9.

31. Ross, E.M., M.E. Maguire, T.W. Sturgill, R.L. Biltonen, and A.G. Gilman, *Relationship between the beta-adrenergic receptor and adenylate cyclase.* The Journal of biological chemistry, 1977. **252**(16): p. 5761-75.

32. Northup, J.K., P.C. Sternweis, M.D. Smigel, L.S. Schleifer, E.M. Ross, and A.G. Gilman, *Purification of the regulatory component of adenylate cyclase.* Proceedings of the National Academy of Sciences of the United States of America, 1980. **77**(11): p. 6516-20.

33. Dixon, R.A., B.K. Kobilka, D.J. Strader, J.L. Benovic, H.G. Dohlman, T. Frielle, M.A. Bolanowski, C.D. Bennett, E. Rands, R.E. Diehl, R.A. Mumford, E.E. Slater, I.S. Sigal, M.G. Caron, R.J. Lefkowitz, and C.D. Strader, *Cloning of the gene and cDNA for mammalian beta-adrenergic receptor and homology with rhodopsin.* Nature, 1986. **321**(6065): p. 75-9.

34. Caron, M.G., Y. Srinivasan, J. Pitha, K. Kociolek, and R.J. Lefkowitz, *Affinity chromatography of the beta-adrenergic receptor.* The Journal of biological chemistry, 1979. **254**(8): p. 2923-7.

35. Kolakowski, L.F., Jr., *GCRDb: a G-protein-coupled receptor database.* Receptors Channels, 1994. **2**(1): p. 1-7.

36. Bjarnadottir, T.K., D.E. Gloriam, S.H. Hellstrand, H. Kristiansson, R. Fredriksson, and H.B. Schioth, *Comprehensive repertoire and phylogenetic analysis of the G protein-coupled receptors in human and mouse.* Genomics, 2006. **88**(3): p. 263-73.

37. Schioth, H.B. and R. Fredriksson, *The GRAFS classification system of G-protein coupled receptors in comparative perspective.* Gen Comp Endocrinol, 2005. **142**(1-2): p. 94-101.

38. Riobo, N.A. and D.R. Manning, *Pathways of signal transduction employed by vertebrate Hedgehogs.* Biochem J, 2007. **403**(3): p. 369-79.

39. Kikuchi, A., H. Yamamoto, and S. Kishida, *Multiplicity of the interactions of Wnt proteins and their receptors.* Cell Signal, 2007. **19**(4): p. 659-71.

40. Venkatakrishnan, A.J., X. Deupi, G. Lebon, C.G. Tate, G.F. Schertler, and M.M. Babu, *Molecular signatures of G-protein-coupled receptors.* Nature, 2013. **494**(7436): p. 185-94.

41. Harmar, A.J., *Family-B G-protein-coupled receptors.* Genome Biol, 2001. **2**(12): p. REVIEWS3013.

42. Chun, L., W.H. Zhang, and J.F. Liu, *Structure and ligand recognition of class C GPCRs.* Acta Pharmacol Sin, 2012. **33**(3): p. 312-23.

43. Milligan, G. and E. Kostenis, *Heterotrimeric G-proteins: a short history.* Br J Pharmacol, 2006. **147 Suppl 1**: p. S46-55.

44. Oka, Y., L.R. Saraiva, Y.Y. Kwan, and S.I. Korsching, *The fifth class of Galpha proteins.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(5): p. 1484-9.

45. Schwindinger, W.F. and J.D. Robishaw, *Heterotrimeric G-protein betagamma-dimers in growth and differentiation.* Oncogene, 2001. **20**(13): p. 1653-60.

46. Dupre, D.J., M. Robitaille, R.V. Rebois, and T.E. Hebert, *The role of Gbetagamma subunits in the organization, assembly, and function of GPCR signaling complexes.* Annu Rev Pharmacol Toxicol, 2009. **49**: p. 31-56.

47. Clapham, D.E. and E.J. Neer, *G protein beta gamma subunits.* Annu Rev Pharmacol Toxicol, 1997. **37**: p. 167-203.

48. Lambert, N.A., *Dissociation of heterotrimeric g proteins in cells.* Sci Signal, 2008. **1**(25): p. re5.

49. Gales, C., J.J. Van Durm, S. Schaak, S. Pontier, Y. Percherancier, M. Audet, H. Paris, and M. Bouvier, *Probing the activation-promoted structural rearrangements in preassembled receptor-G protein complexes.* Nature structural & molecular biology, 2006. **13**(9): p. 778-86.

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

51. Sprang, S.R., *G protein mechanisms: insights from structural analysis.* Annu Rev Biochem, 1997. **66**: p. 639-78.

52. Rasmussen, S.G., H.J. Choi, D.M. Rosenbaum, T.S. Kobilka, F.S. Thian, P.C. Edwards, M. Burghammer, V.R. Ratnala, R. Sanishvili, R.F. Fischetti, G.F. Schertler, W.I. Weis, and B.K. Kobilka, *Crystal structure of the human beta2 adrenergic G-protein-coupled receptor.* Nature, 2007. **450**(7168): p. 383-7.

53. Kohout, T.A. and R.J. Lefkowitz, *Regulation of G protein-coupled receptor kinases and arrestins during receptor desensitization.* Molecular pharmacology, 2003. **63**(1): p. 9-18.

54. DeWire, S.M., S. Ahn, R.J. Lefkowitz, and S.K. Shenoy, *Beta-arrestins and cell signaling.* Annu Rev Physiol, 2007. **69**: p. 483-510.

55. Delean, A., J.M. Stadel, and R.J. Lefkowitz, *A Ternary Complex Model Explains the Agonist-Specific Binding-Properties of the Adenylate Cyclase-Coupled Beta-Adrenergic-Receptor.* Journal of Biological Chemistry, 1980. **255**(15): p. 7108-7117.

56. Christopoulos, A. and T. Kenakin, *G protein-coupled receptor allosterism and complexing.* Pharmacol Rev, 2002. **54**(2): p. 323-74.

57. Samama, P., S. Cotecchia, T. Costa, and R.J. Lefkowitz, *A mutation-induced activated state of the beta 2-adrenergic receptor. Extending the ternary complex model.* The Journal of biological chemistry, 1993. **268**(7): p. 4625-36.

58. Costa, T. and A. Herz, *Antagonists with negative intrinsic activity at delta opioid receptors coupled to GTP-binding proteins.* Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(19): p. 7321-5.

59. Weiss, J.M., P.H. Morgan, M.W. Lutz, and T.P. Kenakin, *The Cubic Ternary Complex Receptor-Occupancy Model III. Resurrecting Efficacy.* Journal of theoretical biology, 1996. **181**(4): p. 381-397.

60. Weiss, J.M., P.H. Morgan, M.W. Lutz, and T.P. Kenakin, *The Cubic Ternary Complex Receptor–Occupancy Model I. Model Description.* Journal of theoretical biology, 1996. **178**(2): p. 151-167.

61. Weiss, J.M., P.H. Morgan, M.W. Lutz, and T.P. Kenakin, *The Cubic Ternary Complex Receptor–Occupancy Model II. Understanding Apparent Affinity.* Journal of theoretical biology, 1996. **178**(2): p. 169-182.

62. Angel, T.E., M.R. Chance, and K. Palczewski, *Conserved waters mediate structural and functional activation of family A (rhodopsin-like) G protein-coupled receptors.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(21): p. 8555-60.

63. Wootten, D., J. Simms, L.J. Miller, A. Christopoulos, and P.M. Sexton, *Polar transmembrane interactions drive formation of ligand-specific and signal pathway-biased family B G protein-coupled receptor conformations.* Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(13): p. 5211-6.

64. Liu, W., E. Chun, A.A. Thompson, P. Chubukov, F. Xu, V. Katritch, G.W. Han, C.B. Roth, L.H. Heitman, I.J. AP, V. Cherezov, and R.C. Stevens, *Structural basis for allosteric regulation of GPCRs by sodium ions.* Science, 2012. **337**(6091): p. 232-6.

65. Palczewski, K., T. Kumasaka, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, I. Le Trong, D.C. Teller, T. Okada, R.E. Stenkamp, M. Yamamoto, and M. Miyano, *Crystal structure of rhodopsin: A G protein-coupled receptor.* Science, 2000. **289**(5480): p. 739-45.

66. Rovati, G.E., V. Capra, and R.R. Neubig, *The highly conserved DRY motif of class A G protein-coupled receptors: beyond the ground state.* Molecular pharmacology, 2007. **71**(4): p. 959-64.

67. Lukov, G.L., T. Hu, J.N. McLaughlin, H.E. Hamm, and B.M. Willardson, *Phosducin-like protein acts as a molecular chaperone for G protein betagamma dimer assembly.* The EMBO journal, 2005. **24**(11): p. 1965-75.

68. Dupre, D.J., M. Robitaille, M. Richer, N. Ethier, A.M. Mamarbachi, and T.E. Hebert, *Dopamine receptor-interacting protein 78 acts as a molecular chaperone for Ggamma subunits before assembly with Gbeta.* The Journal of biological chemistry, 2007. **282**(18): p. 13703-15.

69. Wall, M.A., D.E. Coleman, E. Lee, J.A. Iniguez-Lluhi, B.A. Posner, A.G. Gilman, and S.R. Sprang, *The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2.* Cell, 1995. **83**(6): p. 1047-58.

70. Digby, G.J., R.M. Lober, P.R. Sethi, and N.A. Lambert, *Some G protein heterotrimers physically dissociate in living cells.* Proceedings of the National

Academy of Sciences of the United States of America, 2006. **103**(47): p. 17789-94.

71. Levitzki, A. and S. Klein, *G-protein subunit dissociation is not an integral part of G-protein action.* Chembiochem, 2002. **3**(9): p. 815-8.

72. Bokoch, G.M., T. Katada, J.K. Northup, M. Ui, and A.G. Gilman, *Purification and properties of the inhibitory guanine nucleotide-binding regulatory component of adenylate cyclase.* The Journal of biological chemistry, 1984. **259**(6): p. 3560-7.

73. Kozasa, T. and A.G. Gilman, *Purification of recombinant G proteins from Sf9 cells by hexahistidine tagging of associated subunits. Characterization of alpha 12 and inhibition of adenylyl cyclase by alpha z.* The Journal of biological chemistry, 1995. **270**(4): p. 1734-41.

74. Affymetrix, I., *2011-12 Anatrace Products Catalog*. 2011.

75. Sambrook, J. and D.W. Russell, *Molecular cloning : a laboratory manual*. 3rd ed. 2001, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

76. Yao, X., C. Parnot, X. Deupi, V.R. Ratnala, G. Swaminath, D. Farrens, and B. Kobilka, *Coupling ligand structure to specific conformational switches in the beta2-adrenoceptor.* Nat Chem Biol, 2006. **2**(8): p. 417-22.

77. Yao, X.J., G. Velez Ruiz, M.R. Whorton, S.G. Rasmussen, B.T. DeVree, X. Deupi, R.K. Sunahara, and B. Kobilka, *The effect of ligand efficacy on the formation and stability of a GPCR-G protein complex.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(23): p. 9501-6.

78. Gurevich, E.V., J.J. Tesmer, A. Mushegian, and V.V. Gurevich, *G protein-coupled receptor kinases: more than just kinases and not only for GPCRs.* Pharmacol Ther, 2012. **133**(1): p. 40-69.

79. Hausdorff, W.P., J.A. Pitcher, D.K. Luttrell, M.E. Linder, H. Kurose, S.J. Parsons, M.G. Caron, and R.J. Lefkowitz, *Tyrosine phosphorylation of G protein alpha subunits by pp60c-src.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(13): p. 5720-4.

80. Whorton, M.R., M.P. Bokoch, S.G. Rasmussen, B. Huang, R.N. Zare, B. Kobilka, and R.K. Sunahara, *A monomeric G protein-coupled receptor isolated in a high-density lipoprotein particle efficiently activates its G protein.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(18): p. 7682-7.

81. Whorton, M.R., B. Jastrzebska, P.S. Park, D. Fotiadis, A. Engel, K. Palczewski, and R.K. Sunahara, *Efficient coupling of transducin to monomeric*

rhodopsin in a phospholipid bilayer. The Journal of biological chemistry, 2008. **283**(7): p. 4387-94.

82. Fotiadis, D., B. Jastrzebska, A. Philippsen, D.J. Muller, K. Palczewski, and A. Engel, *Structure of the rhodopsin dimer: a working model for G-protein-coupled receptors.* Current opinion in structural biology, 2006. **16**(2): p. 252-9.

83. Milligan, G., *G protein-coupled receptor dimerisation: molecular basis and relevance to function.* Biochim Biophys Acta, 2007. **1768**(4): p. 825-35.

84. Northup, J.K., P.C. Sternweis, and A.G. Gilman, *The subunits of the stimulatory regulatory component of adenylate cyclase. Resolution, activity, and properties of the 35,000-dalton (beta) subunit.* The Journal of biological chemistry, 1983. **258**(18): p. 11361-8.

85. Chae, P.S., S.G. Rasmussen, R.R. Rana, K. Gotfryd, R. Chandra, M.A. Goren, A.C. Kruse, S. Nurva, C.J. Loland, Y. Pierre, D. Drew, J.L. Popot, D. Picot, B.G. Fox, L. Guan, U. Gether, B. Byrne, B. Kobilka, and S.H. Gellman, *Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins.* Nat Methods, 2010. **7**(12): p. 1003-8.

86. Rasmussen, S.G., B.T. DeVree, Y. Zou, A.C. Kruse, K.Y. Chung, T.S. Kobilka, F.S. Thian, P.S. Chae, E. Pardon, D. Calinski, J.M. Mathiesen, S.T. Shah, J.A. Lyons, M. Caffrey, S.H. Gellman, J. Steyaert, G. Skiniotis, W.I. Weis, R.K. Sunahara, and B.K. Kobilka, *Crystal structure of the beta2 adrenergic receptor-Gs protein complex.* Nature, 2011. **477**(7366): p. 549-55.

87. Rasmussen, S.G., H.J. Choi, J.J. Fung, E. Pardon, P. Casarosa, P.S. Chae, B.T. Devree, D.M. Rosenbaum, F.S. Thian, T.S. Kobilka, A. Schnapp, I. Konetzki, R.K. Sunahara, S.H. Gellman, A. Pautsch, J. Steyaert, W.I. Weis, and B.K. Kobilka, *Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor.* Nature, 2011. **469**(7329): p. 175-80.

88. Westfield, G.H., S.G. Rasmussen, M. Su, S. Dutta, B.T. DeVree, K.Y. Chung, D. Calinski, G. Velez-Ruiz, A.N. Oleskie, E. Pardon, P.S. Chae, T. Liu, S. Li, V.L. Woods, Jr., J. Steyaert, B.K. Kobilka, R.K. Sunahara, and G. Skiniotis, *Structural flexibility of the G alpha s alpha-helical domain in the beta2-adrenoceptor Gs complex.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(38): p. 16086-91.

89. Baranova, E., R. Fronzes, A. Garcia-Pino, N. Van Gerven, D. Papapostolou, G. Pehau-Arnaudet, E. Pardon, J. Steyaert, S. Howorka, and H. Remaut, *SbsB structure and lattice reconstruction unveil Ca2+ triggered S-layer assembly.* Nature, 2012. **487**(7405): p. 119-22.

90. Domanska, K., S. Vanderhaegen, V. Srinivasan, E. Pardon, F. Dupeux, J.A. Marquez, S. Giorgetti, M. Stoppini, L. Wyns, V. Bellotti, and J. Steyaert,

Atomic structure of a nanobody-trapped domain-swapped dimer of an amyloidogenic beta2-microglobulin variant. Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(4): p. 1314-9.

91. De Genst, E.J., T. Guilliams, J. Wellens, E.M. O'Day, C.A. Waudby, S. Meehan, M. Dumoulin, S.T. Hsu, N. Cremades, K.H. Verschueren, E. Pardon, L. Wyns, J. Steyaert, J. Christodoulou, and C.M. Dobson, *Structure and properties of a complex of alpha-synuclein and a single-domain camelid antibody.* Journal of molecular biology, 2010. **402**(2): p. 326-43.

92. Vercruysse, T., E. Pardon, E. Vanstreels, J. Steyaert, and D. Daelemans, *An intrabody based on a llama single-domain antibody targeting the N-terminal alpha-helical multimerization domain of HIV-1 rev prevents viral production.* The Journal of biological chemistry, 2010. **285**(28): p. 21768-80.

93. Rosenbaum, D.M., C. Zhang, J.A. Lyons, R. Holl, D. Aragao, D.H. Arlow, S.G. Rasmussen, H.J. Choi, B.T. Devree, R.K. Sunahara, P.S. Chae, S.H. Gellman, R.O. Dror, D.E. Shaw, W.I. Weis, M. Caffrey, P. Gmeiner, and B.K. Kobilka, *Structure and function of an irreversible agonist-beta(2) adrenoceptor complex.* Nature, 2011. **469**(7329): p. 236-40.

94. Cherezov, V., D.M. Rosenbaum, M.A. Hanson, S.G. Rasmussen, F.S. Thian, T.S. Kobilka, H.J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka, and R.C. Stevens, *High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor.* Science, 2007. **318**(5854): p. 1258-65.

95. Misquitta, L.V., Y. Misquitta, V. Cherezov, O. Slattery, J.M. Mohan, D. Hart, M. Zhalnina, W.A. Cramer, and M. Caffrey, *Membrane protein crystallization in lipidic mesophases with tailored bilayers.* Structure, 2004. **12**(12): p. 2113-24.

96. Kahsai, A.W., K. Xiao, S. Rajagopal, S. Ahn, A.K. Shukla, J. Sun, T.G. Oas, and R.J. Lefkowitz, *Multiple ligand-specific conformations of the beta2-adrenergic receptor.* Nat Chem Biol, 2011. **7**(10): p. 692-700.

97. Chung, K.Y., S.G. Rasmussen, T. Liu, S. Li, B.T. DeVree, P.S. Chae, D. Calinski, B.K. Kobilka, V.L. Woods, Jr., and R.K. Sunahara, *Conformational changes in the G protein Gs induced by the beta2 adrenergic receptor.* Nature, 2011. **477**(7366): p. 611-5.

98. Kleywegt, G.J. and A.T. Brunger, *Checking your imagination: applications of the free R value.* Structure, 1996. **4**(8): p. 897-904.

99. Farrens, D.L., C. Altenbach, K. Yang, W.L. Hubbell, and H.G. Khorana, *Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin.* Science, 1996. **274**(5288): p. 768-70.

100. Ghanouni, P., J.J. Steenhuis, D.L. Farrens, and B.K. Kobilka, *Agonist-induced conformational changes in the G-protein-coupling domain of the beta 2 adrenergic receptor.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(11): p. 5997-6002.

101. Liu, J., B.R. Conklin, N. Blin, J. Yun, and J. Wess, *Identification of a receptor/G-protein contact site critical for signaling specificity and G-protein activation.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(25): p. 11642-6.

102. Blahos, J., 2nd, S. Mary, J. Perroy, C. de Colle, I. Brabet, J. Bockaert, and J.P. Pin, *Extreme C terminus of G protein alpha-subunits contains a site that discriminates between Gi-coupled metabotropic glutamate receptors.* The Journal of biological chemistry, 1998. **273**(40): p. 25765-9.

103. Kisselev, O.G., J. Kao, J.W. Ponder, Y.C. Fann, N. Gautam, and G.R. Marshall, *Light-activated rhodopsin induces structural binding motif in G protein alpha subunit.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(8): p. 4270-5.

104. Wess, J., *G-protein-coupled receptors: molecular mechanisms involved in receptor activation and selectivity of G-protein recognition.* FASEB J, 1997. **11**(5): p. 346-54.

105. Scheerer, P., J.H. Park, P.W. Hildebrand, Y.J. Kim, N. Krauss, H.W. Choe, K.P. Hofmann, and O.P. Ernst, *Crystal structure of opsin in its G-protein-interacting conformation.* Nature, 2008. **455**(7212): p. 497-502.

106. Phillips, W.J., S.C. Wong, and R.A. Cerione, *Rhodopsin/transducin interactions. II. Influence of the transducin-beta gamma subunit complex on the coupling of the transducin-alpha subunit to rhodopsin.* The Journal of biological chemistry, 1992. **267**(24): p. 17040-6.

107. Cabrera-Vera, T.M., J. Vanhauwe, T.O. Thomas, M. Medkova, A. Preininger, M.R. Mazzoni, and H.E. Hamm, *Insights into G protein structure, function, and regulation.* Endocr Rev, 2003. **24**(6): p. 765-81.

108. Van Eps, N., A.M. Preininger, N. Alexander, A.I. Kaya, S. Meier, J. Meiler, H.E. Hamm, and W.L. Hubbell, *Interaction of a G protein with an activated receptor opens the interdomain interface in the alpha subunit.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(23): p. 9420-4.

109. Jones, J.C., J.W. Duffy, M. Machius, B.R. Temple, H.G. Dohlman, and A.M. Jones, *The crystal structure of a self-activating G protein alpha subunit reveals its distinct mechanism of signal initiation.* Sci Signal, 2011. **4**(159): p. ra8.

110. Abdulaev, N.G., T. Ngo, E. Ramon, D.M. Brabazon, J.P. Marino, and K.D. Ridge, *The receptor-bound "empty pocket" state of the heterotrimeric G-protein alpha-subunit is conformationally dynamic.* Biochemistry, 2006. **45**(43): p. 12986-97.

111. Warne, T., R. Moukhametzianov, J.G. Baker, R. Nehme, P.C. Edwards, A.G. Leslie, G.F. Schertler, and C.G. Tate, *The structural basis for agonist and partial agonist action on a beta(1)-adrenergic receptor.* Nature, 2011. **469**(7329): p. 241-4.

112. Wacker, D., G. Fenalti, M.A. Brown, V. Katritch, R. Abagyan, V. Cherezov, and R.C. Stevens, *Conserved binding mode of human beta2 adrenergic receptor inverse agonists and antagonist revealed by X-ray crystallography.* J Am Chem Soc, 2010. **132**(33): p. 11443-5.

113. Bokoch, M.P., Y. Zou, S.G. Rasmussen, C.W. Liu, R. Nygaard, D.M. Rosenbaum, J.J. Fung, H.J. Choi, F.S. Thian, T.S. Kobilka, J.D. Puglisi, W.I. Weis, L. Pardo, R.S. Prosser, L. Mueller, and B.K. Kobilka, *Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor.* Nature, 2010. **463**(7277): p. 108-12.

114. Moro, O., J. Lameh, P. Hogger, and W. Sadee, *Hydrophobic amino acid in the i2 loop plays a key role in receptor-G protein coupling.* The Journal of biological chemistry, 1993. **268**(30): p. 22273-6.

115. Sunahara, R.K., J.J. Tesmer, A.G. Gilman, and S.R. Sprang, *Crystal structure of the adenylyl cyclase activator Gsalpha.* Science, 1997. **278**(5345): p. 1943-7.

116. Sondek, J., D.G. Lambright, J.P. Noel, H.E. Hamm, and P.B. Sigler, *GTPase mechanism of Gproteins from the 1.7-A crystal structure of transducin alpha-GDP-AIF-4.* Nature, 1994. **372**(6503): p. 276-9.

117. Qin, K., C. Dong, G. Wu, and N.A. Lambert, *Inactive-state preassembly of G(q)-coupled receptors and G(q) heterotrimers.* Nat Chem Biol, 2011. **7**(10): p. 740-7.

118. Scheerer, P., M. Heck, A. Goede, J.H. Park, H.W. Choe, O.P. Ernst, K.P. Hofmann, and P.W. Hildebrand, *Structural and kinetic modeling of an activating helix switch in the rhodopsin-transducin interface.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(26): p. 10660-5.

119. De Lean, A., J.M. Stadel, and R.J. Lefkowitz, *A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled beta-adrenergic receptor.* The Journal of biological chemistry, 1980. **255**(15): p. 7108-17.

120. Seifert, R., K. Wenzel-Seifert, T.W. Lee, U. Gether, E. Sanders-Bush, and B.K. Kobilka, *Different effects of Gsalpha splice variants on beta2-adrenoreceptor-mediated signaling. The Beta2-adrenoreceptor coupled to the long splice variant of Gsalpha has properties of a constitutively active receptor.* The Journal of biological chemistry, 1998. **273**(18): p. 5109-16.

121. Randall, M.H., L.J. Altman, and R.J. Lefkowitz, *Structure and biological activity of (-)-[3H]dihydroalprenolol, a radioligand for studies of beta-adrenergic receptors.* Journal of medicinal chemistry, 1977. **20**(8): p. 1090-4.

122. Abrahamsson, T., *The beta 1- and beta 2-adrenoceptor stimulatory effects of alprenolol, oxprenolol and pindolol: a study in the isolated right atrium and uterus of the rat.* British journal of pharmacology, 1986. **87**(4): p. 657-64.

123. Irannejad, R., J.C. Tomshine, J.R. Tomshine, M. Chevalier, J.P. Mahoney, J. Steyaert, S.G. Rasmussen, R.K. Sunahara, H. El-Samad, B. Huang, and M. von Zastrow, *Conformational biosensors reveal GPCR signalling from endosomes.* Nature, 2013. **495**(7442): p. 534-8.

124. Aoudia, M. and R. Zana, *Aggregation Behavior of Sugar Surfactants in Aqueous Solutions: Effects of Temperature and the Addition of Nonionic Polymers.* J Colloid Interface Sci, 1998. **206**(1): p. 158-167.

125. le Maire, M., P. Champeil, and J.V. Moller, *Interaction of membrane proteins and lipids with solubilizing detergents.* Biochim Biophys Acta, 2000. **1508**(1-2): p. 86-111.

126. Chan, P., C.J. Thomas, S.R. Sprang, and G.G. Tall, *Molecular chaperoning function of Ric-8 is to fold nascent heterotrimeric G protein alpha subunits.* Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(10): p. 3794-9.

127. Chapman, H.N., P. Fromme, A. Barty, T.A. White, R.A. Kirian, A. Aquila, M.S. Hunter, J. Schulz, D.P. DePonte, U. Weierstall, R.B. Doak, F.R. Maia, A.V. Martin, I. Schlichting, L. Lomb, N. Coppola, R.L. Shoeman, S.W. Epp, R. Hartmann, D. Rolles, A. Rudenko, L. Foucar, N. Kimmel, G. Weidenspointner, P. Holl, M. Liang, M. Barthelmess, C. Caleman, S. Boutet, M.J. Bogan, J. Krzywinski, C. Bostedt, S. Bajt, L. Gumprecht, B. Rudek, B. Erk, C. Schmidt, A. Homke, C. Reich, D. Pietschner, L. Struder, G. Hauser, H. Gorke, J. Ullrich, S. Herrmann, G. Schaller, F. Schopper, H. Soltau, K.U. Kuhnel, M. Messerschmidt, J.D. Bozek, S.P. Hau-Riege, M. Frank, C.Y. Hampton, R.G. Sierra, D. Starodub, G.J. Williams, J. Hajdu, N. Timneanu, M.M. Seibert, J. Andreasson, A. Rocker, O. Jonsson, M. Svenda, S. Stern, K. Nass, R. Andritschke, C.D. Schroter, F. Krasniqi, M. Bott, K.E. Schmidt, X. Wang, I. Grotjohann, J.M. Holton, T.R. Barends, R. Neutze, S. Marchesini, R. Fromme, S. Schorb, D. Rupp, M. Adolph, T. Gorkhover, I. Andersson, H. Hirsemann,

G. Potdevin, H. Graafsma, B. Nilsson, and J.C. Spence, *Femtosecond X-ray protein nanocrystallography.* Nature, 2011. **470**(7332): p. 73-7.

128. Inokuma, Y., S. Yoshioka, J. Ariyoshi, T. Arai, Y. Hitora, K. Takada, S. Matsunaga, K. Rissanen, and M. Fujita, *X-ray analysis on the nanogram to microgram scale using porous complexes.* Nature, 2013. **495**(7442): p. 461-6.

129. Han, D., S. Pal, J. Nangreave, Z. Deng, Y. Liu, and H. Yan, *DNA origami with complex curvatures in three-dimensional space.* Science, 2011. **332**(6027): p. 342-6.

130. Jaakola, V.P., M.T. Griffith, M.A. Hanson, V. Cherezov, E.Y. Chien, J.R. Lane, A.P. Ijzerman, and R.C. Stevens, *The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist.* Science, 2008. **322**(5905): p. 1211-7.

131. Wu, B., E.Y. Chien, C.D. Mol, G. Fenalti, W. Liu, V. Katritch, R. Abagyan, A. Brooun, P. Wells, F.C. Bi, D.J. Hamel, P. Kuhn, T.M. Handel, V. Cherezov, and R.C. Stevens, *Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists.* Science, 2010. **330**(6007): p. 1066-71.

132. Chien, E.Y., W. Liu, Q. Zhao, V. Katritch, G.W. Han, M.A. Hanson, L. Shi, A.H. Newman, J.A. Javitch, V. Cherezov, and R.C. Stevens, *Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist.* Science, 2010. **330**(6007): p. 1091-5.

133. Shimamura, T., M. Shiroishi, S. Weyand, H. Tsujimoto, G. Winter, V. Katritch, R. Abagyan, V. Cherezov, W. Liu, G.W. Han, T. Kobayashi, R.C. Stevens, and S. Iwata, *Structure of the human histamine H1 receptor complex with doxepin.* Nature, 2011. **475**(7354): p. 65-70.

134. Hanson, M.A., C.B. Roth, E. Jo, M.T. Griffith, F.L. Scott, G. Reinhart, H. Desale, B. Clemons, S.M. Cahalan, S.C. Schuerer, M.G. Sanna, G.W. Han, P. Kuhn, H. Rosen, and R.C. Stevens, *Crystal structure of a lipid G protein-coupled receptor.* Science, 2012. **335**(6070): p. 851-5.

135. Wu, H., D. Wacker, M. Mileni, V. Katritch, G.W. Han, E. Vardy, W. Liu, A.A. Thompson, X.P. Huang, F.I. Carroll, S.W. Mascarella, R.B. Westkaemper, P.D. Mosier, B.L. Roth, V. Cherezov, and R.C. Stevens, *Structure of the human kappa-opioid receptor in complex with JDTic.* Nature, 2012. **485**(7398): p. 327-32.

136. Thompson, A.A., W. Liu, E. Chun, V. Katritch, H. Wu, E. Vardy, X.P. Huang, C. Trapella, R. Guerrini, G. Calo, B.L. Roth, V. Cherezov, and R.C. Stevens, *Structure of the nociceptin/orphanin FQ receptor in complex with a peptide mimetic.* Nature, 2012. **485**(7398): p. 395-9.

137. Granier, S., A. Manglik, A.C. Kruse, T.S. Kobilka, F.S. Thian, W.I. Weis, and B.K. Kobilka, *Structure of the delta-opioid receptor bound to naltrindole.* Nature, 2012. **485**(7398): p. 400-4.

138. Manglik, A., A.C. Kruse, T.S. Kobilka, F.S. Thian, J.M. Mathiesen, R.K. Sunahara, L. Pardo, W.I. Weis, B.K. Kobilka, and S. Granier, *Crystal structure of the micro-opioid receptor bound to a morphinan antagonist.* Nature, 2012. **485**(7398): p. 321-6.

139. Kruse, A.C., J. Hu, A.C. Pan, D.H. Arlow, D.M. Rosenbaum, E. Rosemond, H.F. Green, T. Liu, P.S. Chae, R.O. Dror, D.E. Shaw, W.I. Weis, J. Wess, and B.K. Kobilka, *Structure and dynamics of the M3 muscarinic acetylcholine receptor.* Nature, 2012. **482**(7386): p. 552-6.

140. Haga, K., A.C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W.I. Weis, T. Okada, B.K. Kobilka, T. Haga, and T. Kobayashi, *Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist.* Nature, 2012. **482**(7386): p. 547-51.

141. Wacker, D., C. Wang, V. Katritch, G.W. Han, X.P. Huang, E. Vardy, J.D. McCorvy, Y. Jiang, M. Chu, F.Y. Siu, W. Liu, H.E. Xu, V. Cherezov, B.L. Roth, and R.C. Stevens, *Structural features for functional selectivity at serotonin receptors.* Science, 2013. **340**(6132): p. 615-9.

142. Wang, C., Y. Jiang, J. Ma, H. Wu, D. Wacker, V. Katritch, G.W. Han, W. Liu, X.P. Huang, E. Vardy, J.D. McCorvy, X. Gao, X.E. Zhou, K. Melcher, C. Zhang, F. Bai, H. Yang, L. Yang, H. Jiang, B.L. Roth, V. Cherezov, R.C. Stevens, and H.E. Xu, *Structural basis for molecular recognition at serotonin receptors.* Science, 2013. **340**(6132): p. 610-4.

143. Wang, C., H. Wu, V. Katritch, G.W. Han, X.P. Huang, W. Liu, F.Y. Siu, B.L. Roth, V. Cherezov, and R.C. Stevens, *Structure of the human smoothened receptor bound to an antitumour agent.* Nature, 2013.

144. Zou, Y., W.I. Weis, and B.K. Kobilka, *N-terminal T4 lysozyme fusion facilitates crystallization of a G protein coupled receptor.* PloS one, 2012. **7**(10): p. e46039.

145. Griffin, L. and A. Lawson, *Antibody fragments as tools in crystallography.* Clinical and experimental immunology, 2011. **165**(3): p. 285-91.

146. Steyaert, J. and B.K. Kobilka, *Nanobody stabilization of G protein-coupled receptor conformational states.* Current opinion in structural biology, 2011. **21**(4): p. 567-72.

147. Caffrey, M., D. Li, and A. Dukkipati, *Membrane protein structure determination using crystallography and lipidic mesophases: recent advances and successes.* Biochemistry, 2012. **51**(32): p. 6266-88.

148. Coleman, D.E. and S.R. Sprang, *Crystal structures of the G protein Gi alpha 1 complexed with GDP and Mg2+: a crystallographic titration experiment.* Biochemistry, 1998. **37**(41): p. 14376-85.

149. Chelikani, P., V. Hornak, M. Eilers, P.J. Reeves, S.O. Smith, U.L. RajBhandary, and H.G. Khorana, *Role of group-conserved residues in the helical core of beta2-adrenergic receptor.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(17): p. 7027-32.

150. Markby, D.W., R. Onrust, and H.R. Bourne, *Separate GTP binding and GTPase activating domains of a G alpha subunit.* Science, 1993. **262**(5141): p. 1895-901.

151. De Vries, L., B. Zheng, T. Fischer, E. Elenko, and M.G. Farquhar, *The regulator of G protein signaling family.* Annu Rev Pharmacol Toxicol, 2000. **40**: p. 235-71.

152. Soundararajan, M., F.S. Willard, A.J. Kimple, A.P. Turnbull, L.J. Ball, G.A. Schoch, C. Gileadi, O.Y. Fedorov, E.F. Dowler, V.A. Higman, S.Q. Hutsell, M. Sundstrom, D.A. Doyle, and D.P. Siderovski, *Structural diversity in the RGS domain and its interaction with heterotrimeric G protein alpha-subunits.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(17): p. 6457-62.

153. Nance, M.R., B. Kreutz, V.M. Tesmer, R. Sterne-Marr, T. Kozasa, and J.J. Tesmer, *Structural and functional analysis of the regulator of g protein signaling 2-galphaq complex.* Structure, 2013. **21**(3): p. 438-48.

154. Scheffzek, K., M.R. Ahmadian, W. Kabsch, L. Wiesmuller, A. Lautwein, F. Schmitz, and A. Wittinghofer, *The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants.* Science, 1997. **277**(5324): p. 333-8.

155. Jones, D.T. and R.R. Reed, *Golf: an olfactory neuron specific-G protein involved in odorant signal transduction.* Science, 1989. **244**(4906): p. 790-5.

156. Albert, P.R. and L. Robillard, *G protein specificity: traffic direction required.* Cellular signalling, 2002. **14**(5): p. 407-18.

157. O'Sullivan, O., K. Suhre, C. Abergel, D.G. Higgins, and C. Notredame, *3DCoffee: combining protein sequences and structures within multiple sequence alignments.* Journal of molecular biology, 2004. **340**(2): p. 385-95.

158. Kobilka, B.K., *Amino and carboxyl terminal modifications to facilitate the production and purification of a G protein-coupled receptor.* Analytical biochemistry, 1995. **231**(1): p. 269-71.

159. Zhuo, S., J.C. Clemens, D.J. Hakes, D. Barford, and J.E. Dixon, *Expression, purification, crystallization, and biochemical characterization of a recombinant protein phosphatase.* The Journal of biological chemistry, 1993. **268**(24): p. 17754-61.

160. Kuszak, A.J., S. Pitchiaya, J.P. Anand, H.I. Mosberg, N.G. Walter, and R.K. Sunahara, *Purification and functional reconstitution of monomeric mu-opioid receptors: allosteric modulation of agonist binding by Gi2.* The Journal of biological chemistry, 2009. **284**(39): p. 26732-41.

161. Caffrey, M. and V. Cherezov, *Crystallizing membrane proteins using lipidic mesophases.* Nat Protoc, 2009. **4**(5): p. 706-31.

162. McCoy, A.J., R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, and R.J. Read, *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.

163. McCoy, A.J., *Solving structures of protein complexes by molecular replacement with Phaser.* Acta Crystallogr D Biol Crystallogr, 2007. **63**(Pt 1): p. 32-41.

164. Adams, P.D., P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.W. Hung, G.J. Kapral, R.W. Grosse-Kunstleve, A.J. McCoy, N.W. Moriarty, R. Oeffner, R.J. Read, D.C. Richardson, J.S. Richardson, T.C. Terwilliger, and P.H. Zwart, *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.

165. Blanc, E., P. Roversi, C. Vonrhein, C. Flensburg, S.M. Lea, and G. Bricogne, *Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2210-21.

166. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.

167. Brunger, A.T., *Version 1.2 of the Crystallography and NMR system.* Nat Protoc, 2007. **2**(11): p. 2728-33.

168. Schroder, G.F., M. Levitt, and A.T. Brunger, *Super-resolution biomolecular crystallography with low-resolution data.* Nature, 2010. **464**(7292): p. 1218-22.

169. Englander, J.J., C. Del Mar, W. Li, S.W. Englander, J.S. Kim, D.D. Stranz, Y. Hamuro, and V.L. Woods, Jr., *Protein structure change studied by hydrogen-deuterium exchange, functional labeling, and mass spectrometry.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(12): p. 7057-62.

170. Frank, J., M. Radermacher, P. Penczek, J. Zhu, Y. Li, M. Ladjadj, and A. Leith, *SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields.* Journal of structural biology, 1996. **116**(1): p. 190-9.

171. Ludtke, S.J., P.R. Baldwin, and W. Chiu, *EMAN: semiautomated software for high-resolution single-particle reconstructions.* Journal of structural biology, 1999. **128**(1): p. 82-97.

172. Radermacher, M., T. Wagenknecht, A. Verschoor, and J. Frank, *Three-dimensional structure of the large ribosomal subunit from Escherichia coli.* The EMBO journal, 1987. **6**(4): p. 1107-14.

173. Mindell, J.A. and N. Grigorieff, *Accurate determination of local defocus and specimen tilt in electron microscopy.* Journal of structural biology, 2003. **142**(3): p. 334-47.

174. Brink, J., S.J. Ludtke, Y. Kong, S.J. Wakil, J. Ma, and W. Chiu, *Experimental verification of conformational variation of human fatty acid synthase as predicted by normal mode analysis.* Structure, 2004. **12**(2): p. 185-91.

175. Menetret, J.F., R.S. Hegde, S.U. Heinrich, P. Chandramouli, S.J. Ludtke, T.A. Rapoport, and C.W. Akey, *Architecture of the ribosome-channel complex derived from native membranes.* Journal of molecular biology, 2005. **348**(2): p. 445-57.

176. Grigorieff, N., *FREALIGN: high-resolution refinement of single particle structures.* Journal of structural biology, 2007. **157**(1): p. 117-25.

177. Swaminath, G., X. Deupi, T.W. Lee, W. Zhu, F.S. Thian, T.S. Kobilka, and B. Kobilka, *Probing the beta2 adrenoceptor binding site with catechol reveals differences in binding and activation by agonists and partial agonists.* The Journal of biological chemistry, 2005. **280**(23): p. 22165-71.

178. Fung, J.J., X. Deupi, L. Pardo, X.J. Yao, G.A. Velez-Ruiz, B.T. Devree, R.K. Sunahara, and B.K. Kobilka, *Ligand-regulated oligomerization of beta(2)-adrenoceptors in a model lipid bilayer.* The EMBO journal, 2009. **28**(21): p. 3315-28.

179. Campagnola, L. *MetaArray*. 2008; Available from: www.scipy.org/Cookbook/MetaArray.

180. Various. *Scipy*. 2013; Available from: www.scipy.org.

181. Cock, P.J., T. Antao, J.T. Chang, B.A. Chapman, C.J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski, and M.J. de Hoon, *Biopython: freely available Python tools for computational molecular biology and bioinformatics.* Bioinformatics, 2009. **25**(11): p. 1422-3.